

**UNIVERSIDADE FEDERAL DE SANTA MARIA
CENTRO DE TECNOLOGIA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA CIVIL**

**APLICAÇÃO DE TÉCNICAS ESTATÍSTICAS MULTIVARIADAS
EM DADOS DE CERÂMICA VERMELHA PRODUZIDA
NA REGIÃO CENTRAL DO RIO GRANDE DO SUL**

DISSERTAÇÃO DE MESTRADO

Danielle de Souza Saad

Santa Maria, RS, Brasil

2009

**APLICAÇÃO DE TÉCNICAS ESTATÍSTICAS
MULTIVARIADAS EM DADOS DE CERÂMICA
VERMELHA PRODUZIDA NA REGIÃO CENTRAL
DO RIO GRANDE DO SUL**

por

Danielle de Souza Saad

Dissertação apresentada ao Curso de Mestrado do Programa de Pós-Graduação em Engenharia Civil da Universidade Federal de Santa Maria (UFSM, RS), como requisito parcial para obtenção do grau de **Mestre em Engenharia Civil.**

Orientador: Prof. Dr. José Mario Doleys Soares

Santa Maria, RS, Brasil

2009

**Universidade Federal de Santa Maria
Centro de Tecnologia
Programa de Pós-Graduação em Engenharia Civil**

A Comissão Examinadora, abaixo assinada,
aprova a Dissertação de Mestrado

**APLICAÇÃO DE TÉCNICAS ESTATÍSTICAS
MULTIVARIADAS EM DADOS DE CERÂMICA
VERMELHA PRODUZIDA NA REGIÃO CENTRAL
DO RIO GRANDE DO SUL**

elaborada por
Danielle de Souza Saad

como requisito parcial para obtenção do grau de
Mestre em Engenharia Civil

COMISSÃO EXAMINADORA:

Prof. Dr. José Mario Doleys Soares - UFSM
(Presidente/Orientador)

Prof. Dr. Adriano Mendonça Souza - UFSM

Prof. Dr. Gerson Sisniegas Alva - UFSM

Santa Maria, 10 de setembro de 2009.

RESUMO

Dissertação de Mestrado
Programa de Pós-Graduação em Engenharia Civil
Universidade Federal de Santa Maria

APLICAÇÃO DE TÉCNICAS ESTATÍSTICAS MULTIVARIADAS EM DADOS DE CERÂMICA VERMELHA PRODUZIDA NA REGIÃO CENTRAL DO RIO GRANDE DO SUL

AUTORA: DANIELLE DE SOUZA SAAD

ORIENTADOR: PROF. DR. JOSÉ MARIO DOLEYS SOARES

Data e Local da Defesa: Santa Maria, 10 de setembro de 2009.

Este trabalho visou empregar técnicas estatísticas multivariadas através do *Software STATISTICA 7.0 for Windows*, na análise de dados de cerâmica vermelha produzidos na região Central do estado do Rio Grande do Sul. As variáveis utilizadas foram: produção mensal total, número de indústrias cerâmicas, tijolos maciços, blocos de vedação e blocos estruturais. As técnicas utilizadas foram Análise de Agrupamento, Análise Fatorial, e Análise de Componentes Principais. O objetivo da técnica de Análise de Agrupamento é determinar o grau de similaridade entre as variáveis. A Análise Fatorial visa reduzir o número de variáveis analisadas em concordância com a Análise de Agrupamento. O grau de contribuição das variáveis na formação dos fatores é identificado pela técnica de Análise de Componentes Principais. No trabalho concluiu-se que as técnicas podem ser aplicadas em dados de produtos cerâmicos, pois os resultados obtidos confirmaram resultados e conclusões obtidas em trabalhos anteriores. As técnicas empregadas demonstraram ser pertinentes aos objetivos propostos.

Palavras-chave: cerâmica vermelha; análise de agrupamentos; análise fatorial; análise de componentes principais

ABSTRACT

Dissertação de Mestrado
Programa de Pós-Graduação em Engenharia Civil
Universidade Federal de Santa Maria

APLICAÇÃO DE TÉCNICAS ESTATÍSTICAS MULTIVARIADAS EM DADOS DE CERÂMICA VERMELHA PRODUZIDA NA REGIÃO CENTRAL DO RIO GRANDE DO SUL

AUTHOR: DANIELLE DE SOUZA SAAD

ADVISOR: PROF. DR. JOSÉ MARIO DOLEYS SOARES

Date and Place of Defense: Santa Maria, september 10th, 2009.

This work aimed the application of multivariate statistical techniques using Software STATISTICA 7.0 for Windows, in the analysis of red ceramics data produced in the Central region of the state of the Rio Grande do Sul. The used variable had been: total monthly production, massive number of ceramic industries, bricks, structural blocks. The used techniques had been Cluster Analysis, Factor Analysis, and Principal Components Analysis. The objective of the technique of Cluster Analysis is to determine the degree of similarity between the variables. The Factor Analysis aims to reduce the number variable analyzed in agreement with the Cluster Analysis. The degree of contribution of the variable in the formation of the factors is identified by the technique of Principal Components Analysis. The work concluded that the techniques can be applied in data of ceramic products, because of the results had confirmed previous works. The employed techniques had demonstrated to be pertinent to the considered objectives.

Keywords: red ceramics; cluster analysis; factor analysis; principal components analysis

AGRADECIMENTOS

À Universidade Federal de Santa Maria, instituição que viabilizou minha formação neste Curso de Pós-Graduação.

Ao profº José Mario Doleys Soares por possibilitar a realização deste trabalho, pela dedicação na minha orientação, horas dispensadas e confiança na minha capacidade durante a execução deste trabalho.

Ao profº Adriano Mendonça Souza pelo ensino e fundamental colaboração, apoio e amizade dispensada.

Ao Mestre Rodrigo Rosa Tomazetti que me possibilitou a utilização de sua dissertação desenvolvida neste curso.

Ao secretário do PPGEC, Sr. Eliomar pela atenção e amizade.

De um modo geral, a todas as pessoas que de alguma forma colaboraram na realização deste trabalho.

De modo especial aos meus pais Elias e Terezinha (in memoriam), eternamente responsáveis por todos os acontecimentos da minha vida.

A minha irmã Denise, que sempre me incentivou e me apoiou na minha formação, sem medir esforços. Obrigada pela paciência e compreensão.

Ao Vini, amor, companheiro, amigo e irmão em todos os momentos.

Acima de tudo agradeço a Deus, pela força e coragem e por todas as oportunidades recebidas.

Muito obrigada.

LISTA DE ILUSTRAÇÕES

FIGURA 1 – LOCALIZAÇÃO DA PRODUÇÃO BRASILEIRA DE ALGUNS SEGMENTOS CERÂMICOS	9
FIGURA 2 – CLASSIFICAÇÃO DE ARGILAS PARA USO CERÂMICO COM BASE NA COR APÓS QUEIMA A 950°C, 1250°C E 1450°C, EM ATMOSFERA OXIDANTE	13
FIGURA 3 – PASSOS NA ANÁLISE DE AGRUPAMENTOS	31
FIGURA 4 - MAPA DE LOCALIZAÇÃO DA REGIÃO CENTRAL DO ESTADO DO RS.....	41
FIGURA 5 – DADOS DA DISTRIBUIÇÃO DAS EMPRESAS E PRODUÇÃO DE CERÂMICA VERMELHA NOS MUNICÍPIOS DA REGIÃO CENTRAL DO ESTADO DO RS	52
FIGURA 6 - DENDOGRAMA VERTICAL DOS DADOS DA PRODUÇÃO DE CERÂMICA VERMELHA	53
FIGURA 7 - LIGAÇÃO ENTRE AS VARIÁVEIS DA PRODUÇÃO DE CERÂMICA VERMELHA	53
FIGURA 8 – LIGAÇÃO DAS VARIÁVEIS DA PRODUÇÃO DE CERÂMICA VERMELHA	54
FIGURA 9 – MATRIZ DE DISTÂNCIAS PARA PRODUÇÃO DE CERÂMICA VERMELHA.....	54
FIGURA 10 – ESTATÍSTICA DESCRITIVA (MÉDIA E DESVIO PADRÃO) DOS DADOS DE CERÂMICA VERMELHA.....	55
FIGURA 11 – DENDOGRAMA VERTICAL DOS DADOS, EXCLUINDO-SE A VARIÁVEL PRODUÇÃO MENSAL TOTAL	55
FIGURA 12 – SELEÇÃO DAS VARIÁVEIS DA PRODUÇÃO DE CERÂMICA VERMELHA	56
FIGURA 13 – AUTOVALOR, % DA VARIÂNCIA TOTAL, AUTOVALOR ACUMULADO E % ACUMULADA PARA DADOS DE CERÂMICA VERMELHA.....	56
FIGURA 14 – COMPOSIÇÃO DOS FATORES PARA AS VARIÁVEIS DE CERÂMICA VERMELHA ANALISADAS	57
FIGURA 15 – PLANOS FATORIAIS PARA PRODUÇÃO DE CERÂMICA VERMELHA	58
FIGURA 16 – PLANOS FATORIAIS CONTENDO PERPENDICULARES EM RELAÇÃO AO FATOR 1, PARA PRODUÇÃO DE CERÂMICA VERMELHA.....	59
FIGURA 17 - PLANOS FATORIAIS CONTENDO PERPENDICULARES EM RELAÇÃO AO FATOR 2, PARA PRODUÇÃO DE CERÂMICA VERMELHA.....	59
FIGURA 18 – PLANOS FATORIAIS, COM ROTAÇÃO <i>VARIMAX NORMALIZED</i> PARA DADOS DE CERÂMICA VERMELHA	60
FIGURA 19 – AUTOVALORES PARA DADOS DE CERÂMICA VERMELHA.....	61
FIGURA 20 - COMUNALIDADES PARA PRODUÇÃO DE CERÂMICA VERMELHA.....	61
FIGURA 21 - GRÁFICO TRI-DIMENSIONAL DOS PLANOS FATORIAIS PARA DADOS DE CERÂMICA VERMELHA.....	62
FIGURA 22 - RESULTADOS DOS AUTOVETORES PARA PRODUÇÃO DE CERÂMICA VERMELHA.....	62
FIGURA 23 – RESULTADOS DAS COMPONENTES PRINCIPAIS PARA PRODUÇÃO DE CERÂMICA VERMELHA.....	63
FIGURA 24 - CORRELAÇÕES PARA PRODUÇÃO DE CERÂMICA VERMELHA	64
FIGURA 25 – COMPONENTES PRINCIPAIS PARA PRODUÇÃO DE CERÂMICA VERMELHA	65
FIGURA 26 - CÍRCULO DE CORRELAÇÃO UNITÁRIO PARA PRODUÇÃO DE CERÂMICA VERMELHA	66
FIGURA 27 - DISTRIBUIÇÃO DA NUVEM DE PONTOS PARA PRODUÇÃO DE CERÂMICA VERMELHA.....	67

FIGURA 28 – VARIÁVEIS ORIGINAIS E COMPONENTES PRINCIPAIS (<i>FACTOR SCORES</i>) PARA PRODUÇÃO DE CERÂMICA VERMELHA	68
FIGURA 29 – MATRIZ DE CORRELAÇÃO ENTRE VARIÁVEIS ORIGINAIS E COMPONENTES PRINCIPAIS PARA PRODUÇÃO DE CERÂMICA VERMELHA	68
FIGURA 30 - DADOS DAS DIMENSÕES MÉDIAS E CARACTERÍSTICAS GEOMÉTRICAS DOS TIJOLOS MACIÇOS	69
FIGURA 31 - DENDOGRAMA VERTICAL DOS DADOS DE TIJOLOS MACIÇOS	70
FIGURA 32 - LIGAÇÃO ENTRE AS VARIÁVEIS ANALISADAS PARA TIJOLOS MACIÇOS	70
FIGURA 33 - LIGAÇÃO DAS VARIÁVEIS ANALISADAS PARA TIJOLOS MACIÇOS	71
FIGURA 34 - MATRIZ DE DISTÂNCIAS PARA VARIÁVEIS DE TIJOLOS MACIÇOS.....	71
FIGURA 35 – ESTATÍSTICA DESCRITIVA (MÉDIA E DESVIO PADRÃO) DOS DADOS PARA TIJOLOS MACIÇOS	72
FIGURA 36 - AUTOVALOR, % DA VARIÂNCIA TOTAL, AUTOVALOR ACUMULADO E % ACUMULADA, PARA TIJOLOS MACIÇOS	72
FIGURA 37 - COMPOSIÇÃO DOS FATORES PARA AS VARIÁVEIS DE TIJOLOS MACIÇOS	73
FIGURA 38 - PLANOS FATORIAIS PARA TIJOLOS MACIÇOS	74
FIGURA 39 - PLANOS FATORIAIS CONTENDO PERPENDICULARES EM RELAÇÃO AO FATOR 1, PARA TIJOLOS MACIÇOS	75
FIGURA 40 - PLANOS FATORIAIS CONTENDO PERPENDICULARES EM RELAÇÃO AO FATOR 2, PARA TIJOLOS MACIÇOS	75
FIGURA 41 - PLANOS FATORIAIS, COM ROTAÇÃO <i>VARIMAX NORMALIZED</i> , PARA OS DADOS DE TIJOLOS MACIÇOS	76
FIGURA 42 - MATRIZ FATORIAL ROTACIONADA PARA TIJOLOS MACIÇOS.....	76
FIGURA 43 - AUTOVALORES PARA TIJOLOS MACIÇOS	77
FIGURA 44 - COMUNALIDADES PARA TIJOLOS MACIÇOS.....	77
FIGURA 45 – GRÁFICO TRI-DIMENSIONAL DOS PLANOS FATORIAIS PARA TIJOLOS MACIÇOS	78
FIGURA 46 - RESULTADOS DOS AUTOVETORES PARA TIJOLOS MACIÇOS.....	78
FIGURA 47 – RESULTADOS DAS COMPONENTES PRINCIPAIS PARA TIJOLOS MACIÇOS.....	79
FIGURA 48 - CORRELAÇÕES PARA TIJOLOS MACIÇOS.....	80
FIGURA 49 – COMPONENTES PRINCIPAIS PARA TIJOLOS MACIÇOS	81
FIGURA 50 - CÍRCULO DE CORRELAÇÃO UNITÁRIO PARA TIJOLOS MACIÇOS	82
FIGURA 51 - DISTRIBUIÇÃO DA NUVEM DE PONTOS POR MUNICÍPIOS, PARA TIJOLOS MACIÇOS..	83
FIGURA 52 - DISTRIBUIÇÃO DA NUVEM DE PONTOS POR OLARIAS, PARA TIJOLOS MACIÇOS.....	83
FIGURA 53 - MATRIZ DE CORRELAÇÃO ENTRE VARIÁVEIS ORIGINAIS E COMPONENTES PRINCIPAIS, PARA TIJOLOS MACIÇOS.....	84
FIGURA 54 - DADOS DAS DIMENSÕES MÉDIAS PARA BLOCOS DE VEDAÇÃO E BLOCO PORTANTE	85
FIGURA 55 - DENDOGRAMA VERTICAL DOS DADOS DE BLOCOS.....	86
FIGURA 56 - LIGAÇÃO ENTRE AS VARIÁVEIS ANALISADAS PARA BLOCOS.....	87
FIGURA 57 - LIGAÇÃO DAS VARIÁVEIS ANALISADAS PARA BLOCOS	87
FIGURA 58 - MATRIZ DE DISTÂNCIAS PARA VARIÁVEIS DOS BLOCOS.....	88
FIGURA 59 - ESTATÍSTICA DESCRITIVA (MÉDIA E DESVIO PADRÃO) DOS DADOS PARA BLOCOS .	88
FIGURA 60 - AUTOVALOR, % DA VARIÂNCIA TOTAL, AUTOVALOR ACUMULADO E % ACUMULADA, PARA BLOCOS.....	89
FIGURA 61 - COMPOSIÇÃO DOS FATORES PARA AS VARIÁVEIS DOS BLOCOS	89
FIGURA 62 - PLANOS FATORIAIS PARA OS BLOCOS	90
FIGURA 63 - PLANOS FATORIAIS, COM ROTAÇÃO <i>VARIMAX NORMALIZED</i> , PARA OS DADOS DE BLOCOS	91
FIGURA 64 - MATRIZ FATORIAL ROTACIONADA PARA OS BLOCOS	91
FIGURA 65 - AUTOVALORES PARA OS BLOCOS	92
FIGURA 66 - COMUNALIDADES PARA OS BLOCOS	92

FIGURA 67 - GRÁFICO TRI-DIMENSIONAL DOS PLANOS FATORIAIS PARA OS BLOCOS.....	93
FIGURA 68 - RESULTADOS DOS AUTOVETORES PARA OS BLOCOS.....	93
FIGURA 69 – RESULTADOS DAS COMPONENTES PRINCIPAIS PARA OS BLOCOS.....	95
FIGURA 70 - CORRELAÇÕES PARA OS BLOCOS	95
FIGURA 71 - COMPONENTES PRINCIPAIS PARA OS BLOCOS.....	96
FIGURA 72 - CÍRCULO DE CORRELAÇÃO UNITÁRIO PARA OS BLOCOS	97
FIGURA 73 - DISTRIBUIÇÃO DA NUVEM DE PONTOS POR MUNICÍPIOS, PARA OS BLOCOS.....	98
FIGURA 74 - DISTRIBUIÇÃO DA NUVEM DE PONTOS POR CASOS, PARA OS BLOCOS.....	98
FIGURA 75 - MATRIZ DE CORRELAÇÃO ENTRE VARIÁVEIS ORIGINAIS E COMPONENTES PRINCIPAIS, PARA OS BLOCOS.....	99
FIGURA 76 - DADOS DE ABSORÇÃO DE ÁGUA, IRA E RESISTÊNCIA À COMPRESSÃO PARA TIJOLOS MACIÇOS.....	100
FIGURA 77 - DADOS PADRONIZADOS PARA TIJOLOS MACIÇOS.....	101
FIGURA 78 - DENDOGRAMA VERTICAL DOS DADOS DE ABSORÇÃO DE ÁGUA, IRA E RESISTÊNCIA À COMPRESSÃO DE TIJOLOS MACIÇOS	102
FIGURA 79 - LIGAÇÃO ENTRE AS VARIÁVEIS DE ABSORÇÃO DE ÁGUA, IRA E RESISTÊNCIA À COMPRESSÃO PARA TIJOLOS MACIÇOS	102
FIGURA 80 - LIGAÇÃO DAS VARIÁVEIS DE ABSORÇÃO DE ÁGUA, IRA E RESISTÊNCIA À COMPRESSÃO PARA TIJOLOS MACIÇOS	103
FIGURA 81 – MATRIZ DE DISTÂNCIAS PARA ABSORÇÃO DE ÁGUA, IRA E RESISTÊNCIA À COMPRESSÃO PARA TIJOLOS MACIÇOS	103
FIGURA 82 - CÍRCULO DE CORRELAÇÃO UNITÁRIO PARA ABSORÇÃO DE ÁGUA, IRA E RESISTÊNCIA À COMPRESSÃO PARA TIJOLOS MACIÇOS	104
FIGURA 83 - DISTRIBUIÇÃO DA NUVEM DE PONTOS POR MUNICÍPIOS, PARA TIJOLOS MACIÇOS	105
FIGURA 84 - DISTRIBUIÇÃO DA NUVEM DE PONTOS POR CASOS, PARA TIJOLOS MACIÇOS.....	105
FIGURA 85 - DADOS DE ABSORÇÃO DE ÁGUA, IRA E RESISTÊNCIA À COMPRESSÃO PARA BLOCOS	106
FIGURA 86 - DADOS PADRONIZADOS PARA BLOCOS.....	107
FIGURA 87 - DENDOGRAMA VERTICAL DOS DADOS DE ABSORÇÃO DE ÁGUA, IRA E RESISTÊNCIA À COMPRESSÃO DE BLOCOS.....	108
FIGURA 88 - LIGAÇÃO ENTRE AS VARIÁVEIS DE ABSORÇÃO DE ÁGUA, IRA E RESISTÊNCIA À COMPRESSÃO PARA BLOCOS.....	108
FIGURA 89 - LIGAÇÃO DAS VARIÁVEIS DE ABSORÇÃO DE ÁGUA, IRA E RESISTÊNCIA À COMPRESSÃO PARA BLOCOS.....	109
FIGURA 90 - MATRIZ DE DISTÂNCIAS PARA ABSORÇÃO DE ÁGUA, IRA E RESISTÊNCIA À COMPRESSÃO DE BLOCOS.....	109
FIGURA 91 - CÍRCULO DE CORRELAÇÃO UNITÁRIO PARA ABSORÇÃO DE ÁGUA, IRA E RESISTÊNCIA À COMPRESSÃO DE BLOCOS.....	110
FIGURA 92 - DISTRIBUIÇÃO DA NUVEM DE PONTOS POR MUNICÍPIOS, PARA BLOCOS.....	111
FIGURA 93 - DISTRIBUIÇÃO DA NUVEM DE PONTOS POR CASOS, PARA BLOCOS	111
FIGURA 94A - MÓDULOS DISPONÍVEIS NO <i>STATISTICA 7.0</i>	124
FIGURA 94B - MÓDULOS DISPONÍVEIS NO <i>STATISTICA 7.0</i>	124
FIGURA 95 - ENTRADA DOS DADOS.....	125
FIGURA 96 - SELEÇÃO DA ANÁLISE ESTATÍSTICA	125
FIGURA 97 - ACESSO AO <i>SOFTWARE STATISTICA</i>	126
FIGURA 98 – CAIXA DE SELEÇÃO PARA TRANSPORTAR O BANCO DE DADOS DO EXCEL PARA O <i>STATISTICA</i>	126
FIGURA 99 – CAIXA DE SELEÇÃO PARA IMPORTAR TODAS AS PLANILHAS OU AS PLANILHAS SELECIONADAS PARA O <i>STATISTICA</i>	127

FIGURA 100 – CAIXA DE SELEÇÃO PARA IMPORTAR OS DADOS DO EXCEL PARA O <i>STATISTICA</i> , POR LINHAS E POR COLUNAS.....	127
FIGURA 101 – CAIXA DE SELEÇÃO DAS TÉCNICAS MULTIVARIADAS.....	128
FIGURA 102 – MÉTODOS DE <i>CLUSTER</i>	128
FIGURA 103 - CAIXA DE SELEÇÃO PARA ANÁLISE DE AGRUPAMENTO, SELECIONANDO-SE <i>QUICK</i>	129
FIGURA 104 - CAIXA DE SELEÇÃO PARA ANÁLISE DE AGRUPAMENTO, SELECIONANDO-SE <i>ADVANCED</i> , PARA A OPÇÃO <i>INPUT FILE</i>	129
FIGURA 105 - CAIXA DE SELEÇÃO PARA ANÁLISE DE AGRUPAMENTO, SELECIONANDO-SE <i>ADVANCED</i> , PARA A OPÇÃO <i>CLUSTER</i>	130
FIGURA 106 - CAIXA DE SELEÇÃO PARA ANÁLISE DE AGRUPAMENTO, SELECIONANDO-SE <i>ADVANCED</i> , PARA A OPÇÃO <i>AMALGAMATION (LINKAGE) RULE</i>	130
FIGURA 107 - CAIXA DE SELEÇÃO PARA ANÁLISE DE AGRUPAMENTO, SELECIONANDO-SE <i>ADVANCED</i> , PARA A OPÇÃO <i>DISTANCE MEASURE</i>	131
FIGURA 108 - CAIXA DE SELEÇÃO DAS VARIÁVEIS.....	131
FIGURA 109 -- CAIXA DE SELEÇÃO PARA ELABORAÇÃO DO DENDOGRAMA, SELECIONANDO-SE <i>QUICK</i>	132
FIGURA 110 - CAIXA DE SELEÇÃO PARA ELABORAÇÃO DO DENDOGRAMA, SELECIONANDO-SE <i>ADVANCED</i>	132
FIGURA 111 - DADOS EMPREGADOS NA ELABORAÇÃO DE DENDOGRAMAS.....	133
FIGURA 112 - DENDOGRAMA VERTICAL.....	133
FIGURA 113 - DENDOGRAMA HORIZONTAL.....	134
FIGURA 114 - MATRIZ DE DISTÂNCIAS EUCLIDIANAS.....	134
FIGURA 115 – GRÁFICO DE LIGAÇÃO ATRAVÉS DOS PASSOS.....	135
FIGURA 116 - CAIXA DE SELEÇÃO PARA REALIZAÇÃO DA ANÁLISE FATORIAL.....	136
FIGURA 117 - CAIXA DE SELEÇÃO PARA ANÁLISE FATORIAL, OPÇÕES PARA <i>INPUT FILE</i>	137
FIGURA 118 - CAIXA DE SELEÇÃO DAS VARIÁVEIS.....	137
FIGURA 119 – TELA PARA REALIZAÇÃO DE UMA ANÁLISE FATORIAL, OPÇÃO <i>QUICK</i>	138
FIGURA 120 - TELA PARA REALIZAÇÃO DE UMA ANÁLISE FATORIAL, OPÇÃO <i>ADVANCED</i>	139
FIGURA 121 - TELA PARA REALIZAÇÃO DE UMA ANÁLISE FATORIAL, OPÇÃO <i>DESCRIPTIVES</i>	139
FIGURA 122 - <i>EXPLAINED VARIANCE / EINGENVALUES</i>	140
FIGURA 123 - <i>SCORES / FACTOR SCORE COEFFICIENTS</i>	141
FIGURA 124 - SELEÇÃO DA ANÁLISE DE COMPONENTES PRINCIPAIS.....	142
FIGURA 125 - TELA DE ANÁLISE DE COMPONENTES PRINCIPAIS, PARA OPÇÃO <i>QUICK</i>	142
FIGURA 126 - TELA DE SELEÇÃO DAS VARIÁVEIS.....	143
FIGURA 127 - TELA PARA DETERMINAÇÃO DAS COMPONENTES PRINCIPAIS.....	143
FIGURA 128 - CAIXA DE PADRONIZAÇÃO DAS VARIÁVEIS.....	144
FIGURA 129 - TELA DE ANÁLISE DE COMPONENTES PRINCIPAIS, PARA OPÇÃO <i>ADVANCED</i>	145
FIGURA 130 - ESTATÍSTICAS BÁSICAS E TABELAS.....	145
FIGURA 131 - MATRIZES DE CORRELAÇÃO.....	146
FIGURA 132 - OPÇÕES PARA MATRIZ DE CORRELAÇÃO.....	146
FIGURA 133 - TELA DE SELEÇÃO DAS VARIÁVEIS E DAS COMPONENTES.....	147
FIGURA 134 - RESULTADOS PARA AF: <i>LOADINGS / VARIMAX NORMALIZED</i>	148
FIGURA 135 - CÍRCULO DE CORRELAÇÃO UNITÁRIO.....	149

LISTA DE FLUXOGRAMAS

FLUXOGRAMA 1 – PROCESSO DE FABRICAÇÃO DE CERÂMICA VERMELHA	8
--	---

LISTA DE TABELAS

TABELA 1 – PRODUÇÃO DE CERÂMICA VERMELHA NO ANO DE 2001	11
TABELA 2 – PRODUÇÃO RELATIVA DE CERÂMICA VERMELHA NO ANO DE 2001	12
TABELA 3 - PROPRIEDADES DOS PRINCIPAIS ELEMENTOS CONSTITUINTES DAS ARGILAS	13
TABELA 4 - CLASSIFICAÇÃO DO SOLO SEGUNDO A GRANULOMETRIA.....	16
TABELA 5 – VALORES DE RESISTÊNCIA À FLEXÃO E DE ABSORÇÃO DE ÁGUA ESPECIFICADOS PARA ARGILAS UTILIZADAS NA FABRICAÇÃO DE ALGUNS PRODUTOS DE CERÂMICA VERMELHA.	18
TABELA 6 – DIMENSÕES DE FABRICAÇÃO DE BLOCOS CERÂMICOS SEGUNDO A NBR 15270 (ABNT, 2005)	18
TABELA 7 – CATEGORIZAÇÃO DE TIJOLOS MACIÇOS COMUNS, EM FUNÇÃO DA RESISTÊNCIA À COMPRESSÃO, SEGUNDO A NBR 7170 (ABNT, 1983).....	19
TABELA 8 – RESISTÊNCIA À COMPRESSÃO MÍNIMA DE BLOCOS, SEGUNDO A NBR 15270 (ABNT, 2005)	19
TABELA 9 – CARACTERÍSTICAS EXIGIDAS PARA BLOCOS CERÂMICOS, SEGUNDO A NBR 15270 (ABNT, 2005)	19
TABELA 10 – CARGAS DE RUPTURA A FLEXÃO PARA TELHAS CERÂMICAS, SEGUNDO A NBR 15310 (ABNT, 2005).....	20
TABELA 11 – EXIGÊNCIAS DE FABRICAÇÃO PARA TELHAS CERÂMICAS, SEGUNDO A NBR 15310 (ABNT, 2005)	20
TABELA 12 – ESTATÍSTICAS ASSOCIADAS À ANÁLISE FATORIAL	35
TABELA 13 - DISTRIBUIÇÃO DAS EMPRESAS E PRODUÇÃO DE CERÂMICA VERMELHA NOS MUNICÍPIOS.	45
TABELA 14 – DIMENSÕES MÉDIAS PARA BLOCOS DE VEDAÇÃO E BLOCO PORTANTE	46
TABELA 15 - ABSORÇÃO DE ÁGUA E IRA DOS TIJOLOS MACIÇOS	47
TABELA 16 - ABSORÇÃO DE ÁGUA E IRA DOS BLOCOS DE VEDAÇÃO E BLOCO PORTANTE	48
TABELA 17 - RESISTÊNCIA MECÂNICA DE TIJOLOS MACIÇOS	49
TABELA 18 - RESISTÊNCIA MECÂNICA DE BLOCOS DE VEDAÇÃO E BLOCO PORTANTE	50

LISTA DE APÊNDICES

APÊNDICE A – INTRODUÇÃO AO SOFTWARE STATISTICA 7.0 E TUTORIAL E UTILIZAÇÃO.....123

SUMÁRIO

1 INTRODUÇÃO	1
1.1 OBJETIVOS	2
1.2 JUSTIFICATIVA	2
1.3 DELIMITAÇÕES DO TRABALHO	2
1.4 ESTRUTURA DO TRABALHO	3
2 REVISÃO SOBRE CERÂMICA VERMELHA	4
2.1 INTRODUÇÃO	4
2.2 ORIGEM DA CERÂMICA	9
2.3 ARGILAS	12
2.4 COMPOSIÇÃO DA ARGILA	15
3 REVISÃO SOBRE ANÁLISE ESTATÍSTICA MULTIVARIADA	22
3.1 TÉCNICAS MULTIVARIADAS	22
4 METODOLOGIA	40
4.1 LOCALIZAÇÃO GEOGRÁFICA DA REGIÃO EM ESTUDO	40
4.2 DADOS UTILIZADOS NA AVALIAÇÃO DOS PRODUTOS CERÂMICOS	41
4.3 ENSAIOS REALIZADOS NO CONTROLE TECNOLÓGICO DOS PRODUTOS	42
4.4 ANÁLISES ESTATÍSTICAS	50
5 RESULTADOS E DISCUSSÃO	52
5.1 ANÁLISE ESTATÍSTICA DOS DADOS TOTAIS DA DISTRIBUIÇÃO DAS EMPRESAS E PRODUÇÃO DE CERÂMICA VERMELHA	52
5.1.1 ANÁLISE DE AGRUPAMENTO DAS VARIÁVEIS DA PRODUÇÃO DE CERÂMICA VERMELHA ...	52
5.1.2 ANÁLISE FATORIAL DAS VARIÁVEIS DA PRODUÇÃO DE CERÂMICA VERMELHA	56
5.1.3 ANÁLISE DAS COMPONENTES PRINCIPAIS DAS VARIÁVEIS DA PRODUÇÃO DE CERÂMICA VERMELHA	64
5.2 ANÁLISE ESTATÍSTICA DOS DADOS REFERENTES ÀS DIMENSÕES MÉDIAS E CARACTERÍSTICAS GEOMÉTRICAS DOS TIJOLOS MACIÇOS	69
5.2.1 ANÁLISE DE AGRUPAMENTO PARA AS DIMENSÕES MÉDIAS E CARACTERÍSTICAS GEOMÉTRICAS DOS TIJOLOS MACIÇOS	70
5.2.2 ANÁLISE FATORIAL PARA AS DIMENSÕES MÉDIAS E CARACTERÍSTICAS GEOMÉTRICAS DOS TIJOLOS MACIÇOS	72
5.2.3 ANÁLISE DE COMPONENTES PRINCIPAIS PARA AS DIMENSÕES MÉDIAS E CARACTERÍSTICAS GEOMÉTRICAS DOS TIJOLOS MACIÇOS	80
5.3 ANÁLISE ESTATÍSTICA DOS DADOS REFERENTES ÀS DIMENSÕES MÉDIAS PARA BLOCOS DE VEDAÇÃO E BLOCO PORTANTE	85
5.3.1 ANÁLISE DE AGRUPAMENTO DAS DIMENSÕES MÉDIAS E CARACTERÍSTICAS GEOMÉTRICAS PARA BLOCOS DE VEDAÇÃO E BLOCO PORTANTE	86
5.3.2 ANÁLISE FATORIAL DAS DIMENSÕES MÉDIAS E CARACTERÍSTICAS GEOMÉTRICAS PARA BLOCOS DE VEDAÇÃO E BLOCO PORTANTE	88
5.3.3 ANÁLISE DE COMPONENTES PRINCIPAIS DAS DIMENSÕES MÉDIAS E CARACTERÍSTICAS GEOMÉTRICAS PARA BLOCOS DE VEDAÇÃO E BLOCO PORTANTE	95
5.4 ANÁLISE ESTATÍSTICA DOS DADOS REFERENTES À ABSORÇÃO DE ÁGUA, TAXA DE ABSORÇÃO INICIAL (IRA) E RESISTÊNCIA À COMPRESSÃO PARA TIJOLOS MACIÇOS, BLOCOS DE VEDAÇÃO E BLOCO PORTANTE	100
5.4.1 ANÁLISE DE AGRUPAMENTO PARA OS DADOS DE ABSORÇÃO DE ÁGUA, IRA E RESISTÊNCIA À COMPRESSÃO PARA TIJOLOS MACIÇOS	101
5.4.2 ANÁLISE DE AGRUPAMENTO PARA OS DADOS DE ABSORÇÃO DE ÁGUA, IRA E RESISTÊNCIA À COMPRESSÃO PARA BLOCOS DE VEDAÇÃO E BLOCO PORTANTE	106
6. CONCLUSÕES	113

7. BIBLIOGRAFIA	119
APÊNDICE A - INTRODUÇÃO AO SOFTWARE STATISTICA 7.0 E TUTORIAL DE UTILIZAÇÃO.....	123

1 INTRODUÇÃO

Devido às mudanças no cenário econômico mundial e do Brasil, no início da década de 90, passou-se a perceber a importância das chamadas cerâmicas tradicionais para a economia do país como um todo, como a indústria de refratários, e os segmentos de revestimento cerâmico e cerâmica vermelha que exercem enorme influência na cadeia produtiva do macro-complexo da construção civil. Além de serem segmentos-chaves para setores importantes da economia nacional (refratários – siderurgia), são responsáveis pela geração de centenas de milhares de empregos diretos e demandam uma forte interação com a área de pesquisa face às suas necessidades de desenvolvimento tecnológico.

As ferramentas estatísticas, embora possam ser aplicadas em qualquer área do conhecimento científico, são dificilmente empregadas nas empresas do setor cerâmico. Elas podem contribuir na caracterização dos materiais utilizados e posteriormente nos produtos obtidos.

Atualmente devido à grande amplitude e complexidade dos fenômenos que são objeto de estudo e, também do volume cada vez maior de informações que são coletadas a fim de serem analisadas, as técnicas estatísticas assumiram um papel fundamental pois na maior parte das pesquisas realizadas é necessário que seja feito um tratamento estatístico seja na fase de planejamento, na organização de dados, na análise ou na divulgação dos resultados.

Como as diversas áreas de Engenharia Civil empregam a Estatística na maioria dos trabalhos realizados, faz-se necessário o conhecimento não só da teoria, mas também da forma como devem ser realizadas as análises estatísticas, uma vez que nem sempre é fácil escolher a melhor técnica a ser utilizada.

Dentre as diversas técnicas estatísticas destaca-se a análise multivariada ou multidimensional que compreende um conjunto de técnicas onde duas ou mais variáveis poderão ser analisadas simultaneamente.

Assim, esse trabalho visa apresentar uma fundamentação teórica e as aplicações de algumas técnicas estatísticas multivariadas na Engenharia Civil, bem como algumas noções de interpretação dos resultados obtidos, de modo a trazer uma contribuição que venha a facilitar a utilização das mesmas.

1.1 Objetivos

a) Objetivo geral

Desenvolver um estudo teórico-prático sobre os métodos estatísticos multivariados mais freqüentemente utilizados e suas principais aplicações na análise de resultados obtidos para a Engenharia Civil.

b) Objetivos Específicos

- a) Desenvolver um capítulo teórico sobre cerâmica e teoria da Estatística, através de uma revisão bibliográfica.
- b) Destacar a importância e as vantagens das ferramentas estatísticas, mostrando sua aplicabilidade na solução de problemas de Engenharia Civil.
- c) Apresentar resultados das aplicações das principais ferramentas estatísticas multivariadas no desenvolvimento de resultados experimentais provenientes de uma pesquisa realizada por Tomazzeti (2003), empregando o programa computacional *Statistica 7.0*.

1.2 Justificativa

Justifica-se este trabalho pela importância e aplicabilidade das ferramentas estatísticas nas diversas áreas do conhecimento científico, especificamente na área de materiais para Engenharia Civil, pela obtenção de resultados provenientes da aplicação de técnicas estatísticas multivariadas à análise de resultados obtidos experimentalmente.

Para a solução dos exemplos e problemas propostos posteriormente, será empregado o Software *Statistica 7.0* por ser atualmente um dos mais completos Softwares de Estatística, devido a sua ampla aplicabilidade e relativa facilidade de execução, sendo totalmente compatível com ambiente *Windows*.

1.3 Delimitações do trabalho

O presente trabalho tem como enfoque principal a aplicação de técnicas estatísticas multivariadas aos resultados das propriedades físico-químicas e mecânicas obtidos a partir da caracterização dos produtos de cerâmica vermelha. Também serão abordados aspectos relativos ao processo produtivo e formas de aplicação dos resultados obtidos.

A restrição que o trabalho apresenta é que no estudo multivariado sempre se tem uma perda de informação, mas quando as variáveis estão correlacionadas é a maneira correta de se fazer o estudo, mesmo havendo essa imprecisão.

1.4 Estrutura do trabalho

Este trabalho será composto por seis capítulos sendo duas partes distintas: a primeira destinada a uma breve revisão da literatura que servirá de fundamentação teórica para a discussão do tema proposto; a segunda destinada à parte prática, onde serão relacionados os métodos estatísticos multivariados, suas formas de execução, análise dos resultados e as conclusões.

Capítulo 1 será apresentada uma introdução, constando a delimitação do tema, os objetivos geral e específico, a justificativa, o referencial teórico e uma síntese relacionando as partes constituintes do trabalho.

Capítulo 2 será descrito um breve histórico da indústria de cerâmica vermelha, conceitos e aspectos gerais das matérias-primas empregadas, a constituição das massas cerâmicas e o processo de fabricação de tijolos. Também serão abordados conceitos e as principais técnicas utilizadas na caracterização de argilas.

Capítulo 3 serão apresentados os principais conceitos de estatística, o método estatístico e as ferramentas estatísticas que serão empregadas para o desenvolvimento do trabalho.

Capítulo 4, metodologia, será descrita a metodologia utilizada para a elaboração deste trabalho, além das técnicas e procedimentos empregados.

Capítulo 5, resultados e discussão, serão apresentados os resultados obtidos durante o desenvolvimento da pesquisa.

Capítulo 6 serão expostas as conclusões a partir dos resultados obtidos e ainda sugestões de temas para a realização de trabalhos futuros.

No apêndice será apresentada uma introdução ao *software Statistica 7.0* e tutorial de utilização do mesmo.

2 REVISÃO SOBRE CERÂMICA VERMELHA

2.1 Introdução

Segundo a Associação Brasileira de Cerâmica – ABC (2002), a cerâmica tem uma participação no PIB (Produto Interno Bruto) do Brasil estimada em um por cento (1%), correspondendo a cerca de seis bilhões de dólares (US\$ 6.000.000.000,00). A grande disponibilidade de recursos naturais, as fontes alternativas de energia e a disponibilidade de tecnologias fizeram com que as indústrias brasileiras evoluíssem rapidamente e diversos tipos de produtos dos segmentos cerâmicos atingiram níveis de qualidade mundial com uma quantidade exportada importante.

Segundo Kazmierczak (2007), no Brasil o setor industrial fabricante de cerâmica vermelha possui uma importância relevante devido à grande quantidade de produtos empregados na construção civil. A Associação Nacional da Indústria Cerâmica (ANICER, 2007), afirma que existem cerca de 5.500 empresas do setor cerâmico no Brasil, consumindo cerca de 10.300.000 toneladas de argila por mês.

Ainda segundo a ABC (2002), o setor industrial da cerâmica é bastante diversificado podendo ser dividido nos segmentos: cerâmica vermelha, materiais de revestimento, materiais refratários, louça sanitária, isoladores elétricos de porcelana, louça de mesa, cerâmica artística (decorativa e utilitária), filtros cerâmicos de água para uso doméstico, cerâmica técnica e isolantes térmicos. O Brasil conta com todos estes segmentos e existem fabricantes de matérias-primas sintéticas para cerâmica.

Dentre as regiões brasileiras, a Sul e a Sudeste são aquelas onde existe maior densidade demográfica, maior atividade industrial e agropecuária, melhor infra-estrutura e melhor distribuição de renda, o que acarreta numa concentração de indústrias de todos os segmentos cerâmicos associado às facilidades de matérias-primas, energia, Centros de Pesquisa, Universidades e Escolas Técnicas. Deve-se salientar que as demais regiões do país têm apresentado certo grau de desenvolvimento, principalmente a Nordeste, onde muitas fábricas de diversos setores industriais têm-se instalado e onde o setor de turismo tem crescido acentuadamente. Assim, a demanda de materiais cerâmicos tem aumentado, principalmente nos setores ligados à construção civil (ABC, 2002).

O setor cerâmico brasileiro apresenta uma grande deficiência de dados estatísticos e indicadores de desempenho, imprescindíveis para acompanhar o seu desenvolvimento e melhorar a competitividade.

A ABC (2002) faz referência a uma definição e a classificação da cerâmica. A cerâmica abrange todos os materiais inorgânicos, não metálicos, obtidos normalmente após tratamento térmico em temperaturas elevadas. O setor cerâmico é amplo e heterogêneo, o que induz a dividi-lo em segmentos de acordo com diversos fatores como matérias-primas, propriedades e áreas de utilização. Assim, tem-se a seguinte classificação:

- Cerâmica vermelha compreende os materiais com coloração avermelhada utilizados na construção civil (tijolos, blocos, telhas, elementos vazados, lajes, tubos cerâmicos e argilas expandidas) e também utensílios de uso doméstico e de adorno.
- Materiais de revestimento (placas cerâmicas) são os materiais em forma de placas empregados na construção civil para revestimento de paredes, pisos, bancadas e piscinas de ambientes internos e externos. Recebem as denominações de azulejo, pastilha, porcelanato, grês, lajota, piso, etc.
- Cerâmica branca é um grupo bastante diversificado, compreendendo os materiais formados por um corpo branco e geralmente recobertos por uma camada vítrea transparente e incolor. Foram agrupados pela cor branca da massa, necessária devido a razões estéticas e/ou técnicas. Com o aparecimento dos vidrados opacificados, muitos dos produtos enquadrados neste grupo passaram a ser fabricados, sem perda das características para uma determinada aplicação, com matérias-primas contendo grau de impurezas, responsáveis pela coloração.

Assim, é adequado subdividir este grupo em: louça sanitária, louça de mesa, isoladores elétricos para alta e baixa tensão, cerâmica artística (decorativa e utilitária) e cerâmica técnica para fins diversos, tais como: químico, elétrico, térmico e mecânico.

- Material refratário compreende uma variedade de produtos, que têm como objetivo tolerar temperaturas elevadas nas condições específicas de processo e de operação dos equipamentos industriais, que geralmente envolvem esforços mecânicos, ataques químicos, variações bruscas de temperatura entre outras solicitações. Para suportar tais solicitações e em função da natureza das mesmas, desenvolveram-se vários tipos de produtos a partir de diferentes matérias-primas ou uma combinação destas. Assim, de acordo com a matéria-prima ou componente químico principal, os produtos refratários podem ser classificados em sílica, sílico-aluminoso, aluminoso, mulita, magnesianocromítico, cromítico-magnésiano, carbeto de silício, grafita, carbono, zircônia, zirconita, espinélio e outros.
- Isolantes térmicos podem ser classificados em:

a) refratários isolantes que se enquadram no segmento de refratários;

b) isolantes térmicos não refratários, que compreendem produtos tais como vermiculita expandida, sílica diatomácea, diatomito, silicato de cálcio, lã de vidro e lã de rocha, os quais são obtidos por processos distintos ao do item anterior e que podem ser empregados, dependendo do tipo de produto, até uma temperatura de 1100 °C; e

c) fibras ou lãs cerâmicas com características físicas semelhantes às citadas no item anterior, todavia possuem composições tais como sílica, sílica-alumina, alumina e zircônia, que segundo o tipo, podem atingir temperaturas de utilização de 2000° C ou maiores.

- Frita (ou vidrado fritado). Algumas indústrias especializadas fabricam um vidro moído a partir da fusão da mistura de diferentes matérias-primas. Sua aplicação é feita na superfície do corpo cerâmico que, após a queima, adquire um aspecto vítreo. Este tipo de acabamento tem como objetivos o aperfeiçoamento da estética, tornar a peça impermeável, o aumento da resistência mecânica e a melhoria ou desenvolvimento de outras características.

- Corantes constituem-se de óxidos puros ou pigmentos inorgânicos sintéticos obtidos através da mistura de óxidos ou de seus compostos. Os pigmentos são produzidos por empresas especializadas, inclusive aquelas que produzem fritas, cuja obtenção abrange a mistura das matérias-primas, calcinação e moagem. Os corantes são adicionados aos esmaltes (vidrados) ou aos corpos cerâmicos com o objetivo de proporcionar diversas tonalidades de colorações e alguns efeitos especiais.

- Abrasivos. Uma fração da indústria de abrasivos compõe um segmento cerâmico devido o emprego de matérias-primas e processos análogos aos da cerâmica. Dentre os produtos podem ser citados o óxido de alumínio eletrofundido e o carbetto de silício.

- Vidro, cimento e cal são importantes segmentos cerâmicos e, ocasionalmente, são analisados à parte da cerâmica devido às particularidades apresentadas.

- Cerâmica de alta tecnologia/cerâmica avançada. Com o desenvolvimento de novas tecnologias e o aperfeiçoamento das tecnologias existentes, são necessários materiais com qualidade excepcionalmente elevada. Esses materiais passaram a ser desenvolvidos a partir de matérias-primas sintéticas de elevada pureza e através de processos rigorosamente controlados. Tais produtos são fabricados pelo segmento cerâmico de alta tecnologia ou cerâmica avançada. Segundo as suas funções, são classificados em: eletroeletrônicos, magnéticos, ópticos, químicos, térmicos, mecânicos, biológicos e nucleares. Cabe salientar que os produtos decorrentes deste segmento são de uso intenso e tendem a uma maior ampliação. Podem ser citados como exemplos: naves espaciais, satélites, usinas nucleares, materiais para implantes em seres humanos, aparelhos de som e de vídeo, suporte de

catalisadores para automóveis, sensores (umidade, gases e outros), ferramentas de corte, brinquedos, acendedor de fogão, etc.

Ainda conforme a ABC (2002), as matérias-primas utilizadas na produção de materiais cerâmicos podem ser classificadas em naturais ou sintéticas.

As matérias-primas naturais são aquelas utilizadas na forma como foram extraídas da natureza ou que sofreram algum tratamento físico para eliminação de impurezas indesejáveis, sem alteração da composição química e mineralógica dos componentes principais.

As principais matérias-primas naturais são: agalmatolito, andalusita, argila, bauxita, calcita, cianita, cromita, dolomita, fedspato, filitos cerâmicos, grafita, magnesita, materiais fundentes, pirofilita, quartzo, silimanita, talco, wollastonita e zirconita.

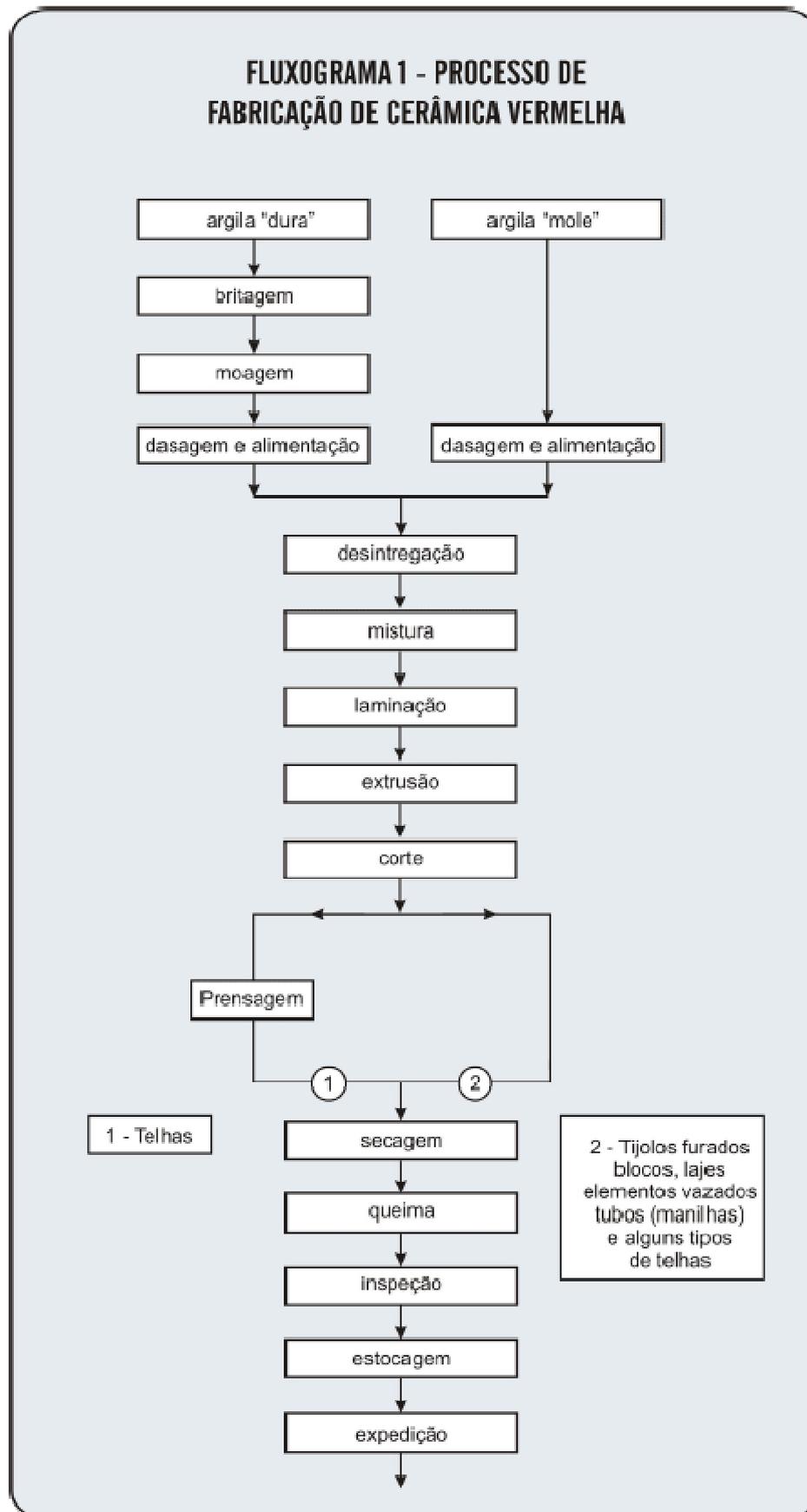
Dentre essas se destaca a argila utilizada para a fabricação dos produtos de cerâmica vermelha. Posteriormente, a argila será detalhada através de conceitos, características, propriedades e aplicações da mesma.

As matérias-primas sintéticas são aquelas que, individualmente ou em mistura, sofreram um tratamento térmico, podendo ser calcinação, sinterização, fusão e fusão/redução e as produzidas por processos químicos.

As principais matérias-primas sintéticas são: alumina, alumina calcinada para cerâmica, alumina eletrofundida marrom, alumina eletrofundida branca, alumina tabular, carbetto de silício, cimento aluminoso, espinélio, magnésia, mulita sintética, mulita- zircônia, óxido de zinco e sílica ativa.

O mercado de matérias-primas naturais é de aproximadamente US\$ 750 milhões e de matérias-primas sintéticas de US\$ 70 milhões (BUSTAMANTE & BRESSIANE, 2000).

No Fluxograma (1) a seguir está descrito o processo de fabricação de cerâmica vermelha.



Fonte: ABC (2002)

No Brasil convencionou-se definir o setor cerâmico em segmentos que se diferenciam segundo os produtos obtidos e mais, especificamente, pelos mercados onde estão inseridos.

Na Figura 1 estão indicadas as localizações geográficas da produção de alguns segmentos cerâmicos no Brasil.



Figura 1 – Localização da produção brasileira de alguns segmentos cerâmicos

Bustamante & Bressiani (2000) afirmaram que o segmento de cerâmica estrutural produz tijolos furados, tijolos maciços, telhas ou lajes, blocos de vedação e estruturais, telhas, manilhas e pisos rústicos. Constitui uma atividade de base ao possibilitar a construção civil, em geral, desde a mais simples até a mais sofisticada. Distribui-se por todo país em micro e pequenas empresas, quase sempre de organização simples e familiar.

2.2 Origem da cerâmica

O termo cerâmica é dado à pedra artificial obtida por meio da moldagem, secagem e cozedura de argila ou mistura contendo argila.

Segundo Pauletti (2001) cerâmica deriva do grego “*kerameikos*”, que significa “feito de terra”. A cerâmica vermelha é assim denominada porque possui uma coloração avermelhada no produto final, em função do tipo de matéria-prima empregada. Na cerâmica vermelha incluem-se produtos como o tijolo e suas variações, constituindo-se por um grupo de produtos rústicos onde o acabamento dificilmente ocorre. Também é denominada de cerâmica estrutural, devido ao emprego na estrutura de edificações.

O Instituto Brasileiro de Geografia e Estatística – IBGE classifica a indústria de cerâmica vermelha ou estrutural como sendo uma indústria de transformação no ramo de atividades denominada transformação de materiais não metálicos.

O processo de fabricação de produtos provenientes da cerâmica vermelha é amplamente conhecido, sendo dominado por diversas civilizações antigas como a China, Babilônia e Grécia Helênica há cerca de sete mil anos (SECTME, 1990).

Os produtos de cerâmica vermelha tiveram seu desenvolvimento no seio de todos os povos, pois a falta de pedra para as construções era constante. Porém, foram os romanos, que implementaram novas técnicas e aprimoraram a fabricação da cerâmica como atividade industrial, através do crescimento das grandes construções dessa época.

As primeiras máquinas moldadoras para fabricação de tijolos surgiram por volta de 1850 e eram movimentadas por força animal. Depois, foram sendo substituídas por máquinas a vapor, o que possibilitou um aumento significativo da produção. Esse fato pode ser considerado como sendo o primeiro grande avanço para a indústria da cerâmica vermelha, pois tornou possível a fabricação de peças especiais e dos tijolos ocós ou furados (STEIL, 2000).

Foram os imigrantes europeus que trouxeram a tecnologia da produção de tijolos e telhas para ao Brasil e, através de organizações familiares, deu-se o surgimento de pequenas olarias em diversas regiões do país.

Com o emprego de estruturas metálicas nas construções, os tijolos passaram a ser utilizados principalmente com a função de elementos de alvenaria de vedação, perdendo a função estrutural. Já as telhas, ainda possuem uma boa penetração no mercado, devido, principalmente, à função estética, embora existam materiais com custos inferiores e resultados funcionais similares (SECTME, 1990).

Atualmente, os materiais cerâmicos são muito empregados na construção civil ainda que com uma grande quantidade de com produtos concorrentes. Isso ocorre devido à estabilidade de resistência mecânica e durabilidade desses materiais, evitando a deterioração por agentes externos e ocasionando vantagens de uso e qualidade nas edificações. Porém, as ameaças que rondam os produtos de cerâmica vermelha, devido ao aprimoramento e desenvolvimento de novos materiais, não podem ser desprezadas (OLIVEIRA, 1993).

Devido a diversos fatores mercadológicos e culturais, as indústrias de cerâmica vermelha no Brasil evoluíram muito pouco em relação ao passado. Diferentemente do mercado europeu, a maioria das empresas brasileiras continua adotando um processo produtivo artesanal em pequena escala e obtendo baixa rentabilidade nos negócios. Isso

acarreta em poucos investimentos na automatização do processo, e conseqüentemente, leva a dificuldades na adequação e evolução tecnológica necessárias ao atendimento do mercado.

Os principais produtos gerados pelo setor de cerâmica vermelha no Brasil são tijolos, blocos, telhas, elementos vazados, lajes, lajotas, ladrilhos vermelhos, tubos e agregados leves.

Segundo o relatório elaborado pelo LACER (Laboratório de Materiais Cerâmicos) da Escola de Engenharia da UFRGS (2004), um levantamento das empresas do setor ceramista (excluído parte de pisos e azulejos) do estado do Rio Grande do Sul demonstrou que o setor de cerâmica vermelha tem uma grande importância não somente no setor da construção civil, mas também na economia do estado do RS, onde esse setor representa entre 0,4 e 0,6% do total arrecadado em ICMS (Imposto sobre Circulação de Mercadorias e Serviços) entre os anos de 2000 a 2003. Cabe salientar que um dos principais tributos que incidem no segmento de cerâmica vermelha é o ICMS.

Na Tabela 1 encontram-se os valores da produção de cerâmica vermelha para o ano de 2001.

Tabela 1 – Produção de cerâmica vermelha no ano de 2001

	Nº de Empresas	Blocos-mês	Telhas-mês	ton Matéria- prima-mês	Empregos diretos
BR	6.860	2.102.000.000	387.000.000	6.855.000	214.000
RS	1.250	240.000.000	10.000.000	531.000	25.000

Fonte: Anuário Brasileiro de Cerâmica (2002).

Pode-se constatar através da tabela a seguir (Tabela 2), confrontando-se a população do Brasil e do RS em relação à produção de cerâmica vermelha, que o RS tem uma produção de blocos-mês próxima a duas vezes maior do que a média brasileira, já a quantidade produzida de telhas por mês em relação aos dados populacionais é visivelmente inferior, o que fornece indícios de que o RS é um potencial consumidor de telhas, isto amparado na localização geográfica do estado e em dados da ABC relativos à produção de cerâmica vermelha.

Tabela 2 - Produção relativa de cerâmica vermelha no ano de 2001.

	População	Blocos-mês/pessoa	Telhas-mês/pessoa
BR	169.799.170	12,38	2,28
RS	10.187.798	23,56	0,98

Fonte: Adaptado da Tabela 4 anterior

A arrecadação do ICMS tanto de blocos quanto de telhas, no RS, indica uma redução na produção desses produtos nos anos de 2000 a 2003.

Segundo o Sindicato das Indústrias de Olaria e de Cerâmica para Construção no Estado do Rio Grande do Sul (SINDICER/RS), o Governo do Estado publicou a lei 353/2005 que isenta o ICMS para tijolos, contemplando o setor cerâmico no Rio Grande do Sul. A Lei determina a isenção de ICMS de tijolos de cerâmica classificados no código 6904.10.00 da Nomenclatura Brasileira de Mercadorias Nomenclatura Brasileira de Mercadorias (NBM)/ Sistema Harmonizado de Designação e Classificação de Mercadorias (SH) - Nomenclatura Comum do Mercosul (NCM). Este benefício foi concedido para conter a entrada de tijolos produzidos em Santa Catarina. As olarias gaúchas não recolhem mais este imposto, nas operações internas, desde 01 de janeiro de 2006. Isso representa um grande avanço para o desenvolvimento e competitividade do setor cerâmico gaúcho.

2.3 Argilas

A elevada gama de valores que definem as propriedades das argilas que faz com que existam diversas variedades de argilas e, conseqüentemente, diversas aplicações destes materiais naturais.

A Figura 2 apresenta de forma sintética um diagrama ternário que apresenta uma classificação para a maioria das argilas utilizadas em cerâmica, baseada na coloração decorrente da queima em atmosfera oxidante. São consideradas as argilas que são aplicadas em três indústrias: cerâmica branca, cerâmica vermelha ou estrutural e materiais refratários.

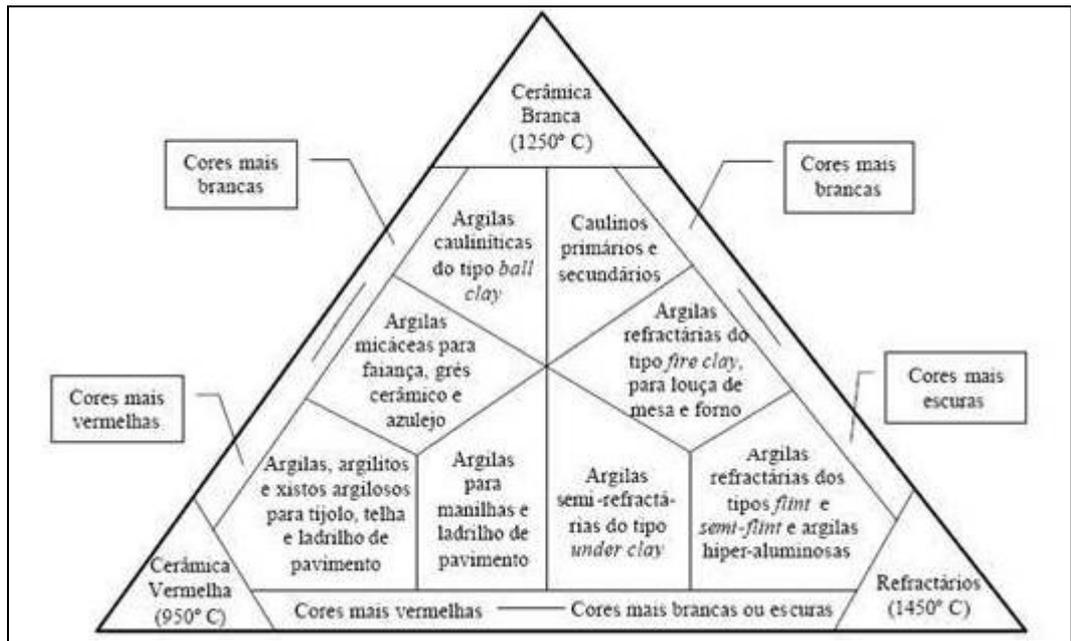


Figura 2 – Classificação de argilas para uso cerâmico com base na cor após queima a 950°C, 1250°C e 1450°C, em atmosfera oxidante

Na Tabela 3 estão sumarizadas as propriedades dos principais elementos constituintes das argilas.

Tabela 3 - Propriedades dos principais elementos constituintes das argilas

Elementos	Principais propriedades
Alumina	Propicia estabilidade dimensional em temperaturas elevadas
Carbonato e sulfato de cálcio e magnésio	Resultam em expansão volumétrica; agem como fundentes.
Matéria orgânica	Resulta em retração, fissuras durante os processos de sazonalidade e queima e diferenças de coloração em um mesmo componente cerâmico.
Sílica livre	Diminui a retração durante os processos de sazonalidade e queima; reduz a plasticidade da argila.
Silicatos e fosfatos	São fundentes, alguns aumentam a resistência da cerâmica.
Sais solúveis	Propiciam o aparecimento de eflorescências nos componentes cerâmicos.

Fonte: Kazmierczak (2007)

Com isso, pode-se dizer que a argila é constituída por uma mistura natural polimérica relativamente complexa e com textura própria, uma vez que cada uma das espécies minerais

apresenta cristais com dimensões, hábitos e formas próprias, características essas que dependem do processo de formação. Até mesmo cristais de uma mesma espécie mineral apresentam maior ou menor variabilidade dimensional, hábito, forma e organização estrutural.

A argila apresenta uma textura terrosa e granulação fina. É formada principalmente por argilominerais, mas pode conter outros minerais como quartzo, mica, pirita, hematita, entre outros, matéria orgânica e outras impurezas.

Os minerais característicos das argilas são os argilominerais, que quimicamente são silicatos de alumínio ou magnésio hidratados, podendo conter outros elementos como ferro, potássio, lítio e outros.

Os principais grupos de argilominerais são caulinita, illita e esmectitas ou montmorilonita. Devido a esses, as argilas em contato com a água desenvolvem propriedades como plasticidade, resistência mecânica a úmido, retração linear de secagem, compactação, tixotropia e viscosidade de suspensões aquosas, as quais explicam a sua ampla variedade de aplicações tecnológicas.

A diferença básica entre estes argilominerais está no tipo de estrutura e nas substituições que podem ocorrer, dentro da estrutura, do alumínio por magnésio ou ferro, e do silício por alumínio ou ferro, e a conseqüente neutralização das cargas residuais geradas pelas diferenças de cargas elétricas dos íons por alguns cátions.

Para a caulinita praticamente não acontecem substituições, já na illita ocorrem substituições sendo o potássio o cátion neutralizante; na montmorilonita também ocorrem substituições e os cátions neutralizantes podem ser sódio, cálcio, potássio entre outros. Isto acarreta em diferentes características de interesse nas várias aplicações tecnológicas.

As argilas constituídas basicamente pela caulinita são as mais refratárias, pois são constituídas basicamente de sílica (SiO_2) e alumina (Al_2O_3), ao passo que os demais, devido à presença de outros elementos, têm a refratariedade sensivelmente reduzida.

A presença de outros minerais, diversas vezes considerados como impurezas, pode afetar fundamentalmente as características de uma argila em uma determinada aplicação. Por isso em muitas aplicações os minerais indesejáveis são eliminados por processos físicos. Esse processo é denominado de beneficiamento.

Segundo as possibilidades de emprego tecnológico influenciadas pela origem e composição mineralógica do material, em diversas situações as argilas recebem designações tais como caulins, bentonitas, argilas refratárias, *flint-clays* e *ball clays*.

As argilas oferecem uma ampla gama de aplicações, tanto na área de cerâmica como em outras áreas tecnológicas.

As argilas mais adequadas à fabricação dos produtos de cerâmica vermelha apresentam em sua composição os argilominerais illita, de camadas mistas illita-montmorilonita e clorita-montmorilonita, além de caulinita, pequenos teores de montmorilonita e compostos de ferro. As argilas apropriadas para materiais refratários são essencialmente cauliníticas, devendo apresentar baixos teores de compostos alcalinos, alcalinos-terrosos e de ferro, podendo conter ainda em alguns tipos a gibbsita ($\text{Al}_2\text{O}_3 \cdot 3\text{H}_2\text{O}$). Nos materiais de revestimento são utilizadas argilas semelhantes àquelas empregadas na produção de cerâmica vermelha ou para cerâmica branca e materiais refratários.

As argilas de queima vermelha ou argilas comuns são especialmente utilizadas na cerâmica vermelha e de revestimento. Essas argilas são caracterizadas como matérias-primas de baixo valor unitário, o que não viabiliza o seu transporte a grandes distâncias, condicionando a instalação de unidades industriais cerâmicas nas proximidades das jazidas.

Segundo Dias, Santos & Goto (2006), com relação à argila utilizada para produção de cerâmica vermelha pode-se afirmar que:

- existem 7.000 minas em atividade,
- são gerados 400 mil empregos diretos e 1,2 milhões de empregos indiretos,
- o faturamento anual é em torno de R\$ 6 bilhões,
- são gerados R\$ 600 milhões em impostos,
- constitui a base da cadeia produtiva da construção civil,
- existe um grande número de empresas de pequeno porte,
- mão-de-obra é intensiva, o que constitui um grande gerador de empregos,
- fixa o trabalhador rural fora dos grandes centros.

2.4 Composição da argila

Segundo Araujo, Rodrigues & Freitas (2000) a indústria cerâmica necessita quantidades suficientes de solo apropriado, água e combustível. O solo de natureza argilosa apresenta características de plasticidade, ou seja, ao ser misturado à água adquire a forma desejada, a qual se mantém após secagem e cozimento. Os constituintes do solo podem ser classificados de acordo com o tamanho de suas partículas, conforme a Tabela (4) a seguir.

Tabela 4 - Classificação do solo segundo a granulometria

Fração	Dimensões (mm)	
Areia	Grossa	2,0 – 0,6
	Média	0,6 – 0,2
	Fina	0,2 – 0,06
Silte	Grosso	0,06 – 0,02
	Médio	0,02 – 0,006
	Fino	0,006 – 0,002
Argila	---	menor que 0,002

Fonte: Araujo, Rodrigues & Freitas (2000)

O solo para fabricação da cerâmica deve conter uma fração de argila, juntamente com silte e areia, de modo a configurar as características desejáveis de plasticidade, bem como de não trincamento e retração, de vitrificação, entre outras. Geologicamente, as argilas são solos residuais ou sedimentares que se formam em consequência da ação do intemperismo físico e/ou químico sobre rochas cristalinas e sedimentares. Devido à grande quantidade de rochas que podem originar as argilas, os processos de formação e seu grau de pureza, os materiais argilosos disponíveis podem ser dotados de diferentes características, tais como:

- cerâmica branca (caulim residual e sedimentar);
- cerâmica refratária (caulim sedimentar e argila refratária);
- cerâmica vermelha (argila de baixa plasticidade, contendo fundentes);
- cerâmica de louça (argila plástica, com fundentes e vitrificantes).

Segundo Kipper (2004), historicamente diversos materiais já foram utilizados na construção civil. A maior parte deles não sobreviveu porque suas qualidades intrínsecas se mostraram insuficientes para atender os requisitos exigidos para o espaço onde o homem passa a maior parte do tempo. As exigências mais tradicionais, como resistência mecânica elevada, resistência ao fogo, durabilidade, isolamento térmico e acústico, impermeabilidade, flexibilidade, etc. são facilmente obtidos pela maioria dos materiais empregados atualmente. Entretanto, somente o tijolo e a telha fabricados a partir de argilas naturais apresentam um desempenho que atende aos requisitos necessários simultaneamente, inclusive respondendo ao problema da umidade em climas como o existente no Sul do País. Talvez por isso, os materiais cerâmicos, especialmente o tijolo e a telha, nunca foram abandonados e ainda se mantêm como o melhor material de construção, mesmo em regiões desenvolvidas como a Europa.

Ainda segundo o mesmo autor, a maioria dos materiais apresenta bom desempenho térmico. Entretanto, somente essa propriedade é insuficiente quando não existe um clima qualitativamente agradável no habitat. As paredes de tijolos funcionam quase como um sistema automático de regulação de umidade, acumulando vapor úmido do ambiente, e liberando-o posteriormente, de maneira que em condições higrométricas desfavoráveis, o ambiente interno da construção com tijolos se mantém agradável. A umidade pode mover-se num material poroso tanto em forma de vapor d'água, em baixas concentrações, como em forma de água líquida, quando em concentrações mais elevadas. A zona limítrofe entre o transporte em forma de vapor e em forma líquida pode ser considerada como um nível crítico de umidade. Abaixo deste nível, a umidade é transportada como vapor d'água por difusão e a quantidade transportada varia segundo a resistência à difusão do vapor d'água do material considerado. Além dos materiais cerâmicos, quase todos os outros materiais de construção são capazes de difusão, sendo o sistema capilar de poros decisivo para o transporte de água na forma líquida. Na cerâmica, este sistema faz com que, por efeito da capilaridade, a umidade seja succionada para fora do material. O mais importante e decisivo na opção do melhor material de construção é que esta é uma propriedade específica da cerâmica em geral, e do tijolo e da telha em particular.

Devido às suas características de transporte de umidade, a argila queimada representa um surpreendente material de construção apresentando baixa higroscopia, isto é, absorção de água, e muito baixo teor de umidade. Pode-se afirmar que, em comparação com outros materiais de construção, a absorção de umidade e posterior liberação são, aproximadamente, dez vezes mais rápidos nos produtos cerâmicos.

Os poros da argila queimada são produzidos pelo fato da água ainda presente nas peças secas ser expelida a altas temperaturas provocando cavidades ou bolsas anteriormente ocupadas pela água combinada. Esta importante característica da cerâmica possui a vantagem de absorver a umidade do ar, mas também possui a capacidade de desenvolver, novamente, esta umidade rapidamente. Por isso, uma parede ou cobertura cerâmica fica umedecida com a chuva ou outras influências, mas seca tão rapidamente quanto se formou a umidade. Este constitui um fator decisivo na escolha do material para a construção civil, especialmente onde a umidade representa uma presença quase constante. E precisamente neste detalhe os materiais de argila cozida apresentam desempenho singular, diferentemente de outros materiais de construção.

Na Tabela 5 estão referidos os valores de resistência e de absorção de água recomendados para argilas utilizadas na fabricação de componentes de cerâmica vermelha.

Tabela 5 – Valores de resistência à flexão e de absorção de água especificados para argilas utilizadas na fabricação de alguns produtos de cerâmica vermelha.

Ensaio	Blocos	Lajotas	Telhas	Tijolos
Tensão de ruptura à flexão em corpos-de-prova secos a 110°C	≥ 2,5 MPa	-	≥ 3,0 MPa	≥ 1,5 MPa
Tensão de ruptura à flexão após a queima	≥ 5,5 MPa	-	≥ 6,5 MPa	≥ 2,0 MPa
Absorção de água após a queima	≤ 25%	≤ 1%	≤ 20%	-

Fonte: Kazmierczak (2007)

As principais dimensões de fabricação de blocos cerâmicos especificadas pela NBR 15270 (ABNT, 2005) estão especificadas na Tabela 6.

Tabela 6 – Dimensões de fabricação de blocos cerâmicos segundo a NBR 15270 (ABNT, 2005)

Largura (cm)	Altura (cm)	Comprimento do bloco principal (cm)	Comprimento do ½ bloco (cm)
9	9	19 ou 24	9 ou 11,5
	14	19 ou 24 ou 29	9 ou 11,5 ou 14
	19	19 ou 24 ou 29 ou 39	9 ou 11,5 ou 14 ou 19
11,5	11,5	24	11,5
	14	24	11,5
	19	19 ou 24 ou 29	9 ou 11,5 ou 14 ou 19
14	19	19 ou 24 ou 29 ou 39	9 ou 11,5 ou 14 ou 19
19	19	19 ou 24 ou 29 ou 39	9 ou 11,5 ou 14 ou 19
24	24	24 ou 29 ou 39	11,5 ou 14 ou 19

Fonte: Kazmierczak (2007)

Kazmierczak (2007) afirma que os tijolos maciços comuns e os blocos cerâmicos devem apresentar resistência à compressão superior às especificadas nas Tabelas 7 e 8 seguintes.

Tabela 7 – Categorização de tijolos maciços comuns, em função da resistência à compressão, segundo a NBR 7170 (ABNT, 1983)

Categoria	Resistência à compressão (MPa)
A	≥ 1,5
B	≥ 2,5
C	≥ 4,0

Fonte: Kazmierczak (2007)

Tabela 8 – Resistência à compressão mínima de blocos, segundo a NBR 15270 (ABNT, 2005)

Tipo de bloco	Resistência à compressão (MPa)*
Blocos de vedação utilizados com furos na horizontal	$\geq 1,5$
Blocos de vedação utilizados com furos na vertical	$\geq 3,0$
Blocos estruturais	$\geq 3,0$

Fonte: Kazmierczak (2007)

onde: * representa a referida área bruta (largura x comprimento do bloco, em cm^2).

Quanto às características exigidas para blocos cerâmicos, segundo a NBR 15270 (ABNT, 2005) as mesmas estão relacionadas na tabela 9 a seguir.

Tabela 9 – Características exigidas para blocos cerâmicos, segundo a NBR 15270 (ABNT, 2005)

Características do bloco	Especificação da norma
Medida das faces	tolerância dimensional de ± 5 mm
Espessura dos septos e paredes externas dos blocos de vedação	septos: espessura ≥ 6 mm paredes externas: espessura ≥ 7 mm
Espessura dos septos e paredes externas dos blocos estruturais de paredes vazadas	septos: espessura ≥ 7 mm paredes externas: espessura ≥ 8 mm
Espessura dos septos e paredes externas dos blocos estruturais de paredes maciças	septos: espessura ≥ 8 mm paredes externas: espessura ≥ 20 mm
Espessura das paredes externas dos blocos estruturais perfurados	espessura ≥ 8 mm
Desvio em relação ao esquadro	desvio ≤ 3 mm
Planeza das faces	flecha ≤ 3 mm
Índice de absorção de água	entre 8% e 22%

Fonte: Kazmierczak (2007)

Kazmierczak (2007) destaca que na NBR 15270-3 (ABNT, 2005) fica especificado um procedimento para a determinação da absorção inicial de água de blocos cerâmicos, denominado de índice de absorção de água inicial. Este ensaio consiste na imersão parcial de um tijolo ou bloco em um recipiente com água, durante um minuto, e na determinação da quantidade de água absorvida durante a imersão, expresso em $\text{g}/193,55 \text{ cm}^2$. Já a NBR 15270-3 (ABNT, 2005) prevê a determinação do módulo de deformação longitudinal e do coeficiente de Poisson do componente cerâmico.

Ainda segundo Kazmierczak (2007), a NBR 15310 (ABNT, 2005) classifica as telhas em função de suas características geométricas e do tipo de fixação, podendo ser dos tipos: telhas planas de encaixe, telhas compostas de encaixe, telhas simples de sobreposição e telhas planas de sobreposição. Nas Tabelas 10 e 11 seguintes estão relacionadas às principais exigências de fabricação para telhas cerâmicas, segundo a NBR 15310 (ABNT, 2005).

Tabela 10 – Cargas de ruptura a flexão para telhas cerâmicas, segundo a NBR 15310 (ABNT, 2005)

Tipo de telha	Exemplo	Carga de ruptura a flexão (N), em ensaio de três pontos
Plana de encaixe	telha francesa	1000
Composta de encaixe	telha romana	1300
Simplex de sobreposição	telha capa e canal colonial telha plana telha paulista telha Piauí	1000
Plana de sobreposição	telha alemã	1000

Fonte: Kazmierczak (2007)

Tabela 11 – Exigências de fabricação para telhas cerâmicas, segundo a NBR 15310 (ABNT, 2005)

Características da telha	Especificação da norma
Retilineidade	telhas planas: $\leq 1\%$ do comprimento efetivo e $\leq 1\%$ da largura efetiva telhas simples de sobreposição e telhas compostas de encaixe: $\leq 1\%$ da largura efetiva
Planaridade	≤ 5 mm
Tolerância dimensional	$\pm 2,0\%$ em relação as dimensões especificadas na norma
Altura do pino	em telhas prensadas: ≥ 7 mm em telhas extrudadas: ≥ 3 mm
Massa da telha seca	no máximo 6% superior ao valor especificado no projeto do modelo da telha
Absorção de água	em clima temperado ou tropical: $\leq 20\%$ em climas frios e temperados: $\leq 12\%$ em clima muito frio ou úmido, durante longos períodos, sujeitos a ciclos frequentes de gelo e degelo: $\leq 7\%$
Impermeabilidade	não apresentar pequenos vazamentos ou formação de gotas em sua face inferior, ao longo do ensaio
Características visuais	são admitidos pequenos defeitos que não prejudiquem seu desempenho
Sonoridade	a telha deve apresentar som semelhante ao metálico, quando suspensa por uma extremidade e percutida

Fonte: Kazmierczak (2007)

A grande variedade de argilas encontradas permite que se obtenham produtos cerâmicos com as mais diversas características tecnológicas, abrangendo:

- a) produtos rústicos, como tijolos e telhas, até produtos de fino acabamento, como porcelana;
- b) produtos permeáveis, como velas de filtros, até produtos impermeáveis, como as louças sanitárias e de grés cerâmico;
- c) produtos frágeis ao fogo até elementos refratários e resistentes a altas temperaturas;
- d) produtos usados como isoladores elétricos até os supercondutores, uma das maiores inovações tecnológicas deste final de século.

Desse modo, o material utilizado no princípio da civilização por nossos ancestrais, encontra atualmente inúmeras aplicações nas indústrias de construção civil, automobilística, eletroeletrônica, espacial e biomédica, entre outras.

3 REVISÃO SOBRE ANÁLISE ESTATÍSTICA MULTIVARIADA

Segundo Christofolletti *et al.* (2005), as empresas do setor cerâmico dificilmente usam as ferramentas estatísticas na caracterização do material e dos produtos obtidos. Essas ferramentas podem ser a análise estatística univariada ou multivariada.

Para Kendall (1969), a análise univariada tem por objetivo a análise de uma única variável. Entretanto, se duas ou mais variáveis forem obtidas, essas poderão ser analisadas pela estatística multivariada.

A análise estatística multivariada ou multidimensional se preocupa com as relações entre as variáveis e apresenta duas características: os valores das diferentes variáveis devem ser obtidos sobre os mesmos indivíduos e estas devem ser independentes e consideradas simultaneamente (KENDALL, 1969). Entre os métodos mais utilizados em Geociências destacam-se a análise de agrupamentos, a análise fatorial e a análise discriminante.

A realização de uma análise múltipla sobre uma amostra fornece um melhor entendimento na razão direta do número de variáveis utilizadas e permite considerar, respectivamente, a variabilidade existente nas diversas propriedades avaliadas.

Dentre as técnicas estatísticas de análise multivariada estão a análise de componentes principais e a análise de agrupamento hierárquico, que são técnicas fundamentalmente diferentes e que podem ser aplicadas independentemente podendo ser complementares. Ambas fornecem uma visão global das amostras dentro do conjunto de dados, segundo as variáveis usadas (CAZAR, 2003).

Cabe salientar que a caracterização da matéria-prima é fundamental para a indústria cerâmica visando o controle de qualidade do produto final. A partir do conhecimento das propriedades e do comportamento da matéria-prima, é possível otimizar a mistura do material buscando uma maior uniformidade nos produtos. A partir disso, Monteiro, Bernardes & Masson (2000) aplicaram uma análise de agrupamentos para material cerâmico.

3.1 Técnicas multivariadas

Os métodos estatísticos para analisar as variáveis estão dispostos em dois grupos:

- estatística univariada que analisa as variáveis de maneira isolada, e
- estatística multivariada que analisa as variáveis de forma conjunta.

A denominação Análise Multivariada corresponde a um grande número de métodos e técnicas que utilizam, simultaneamente, todas as variáveis na interpretação teórica do conjunto de dados obtidos (NETO, 2004).

Os métodos multivariados são escolhidos de acordo com os objetivos da pesquisa, pois se sabe que a análise multivariada é uma análise exploratória de dados. Portanto, na estatística multivariada cada método tem sua fundamentação teórica e aplicabilidade. Quando o interesse é verificar como as amostras se relacionam, ou seja, o quanto estas são semelhantes, destacam-se dois métodos: a análise de agrupamento hierárquico e a análise fatorial com análise de componentes principais.

Quando um estudo estatístico univariado ou multivariado é realizado, ocorre uma perda de informações.

A análise multivariada refere-se a todos os métodos analíticos que analisam simultaneamente múltiplas medidas em cada indivíduo ou objeto sob investigação. Qualquer análise simultânea de mais de duas variáveis pode ser considerada análise multivariada.

Muitas técnicas multivariadas são extensões das análises univariadas e bivariadas. Entretanto, para ser considerado verdadeiramente multivariado, todas as variáveis devem ser caracterizadas e inter-relacionadas de tal modo que seus diferentes efeitos não possam ser interpretados separadamente. O propósito da análise multivariada é medir, explicar e prever o grau de relação entre variáveis. O caráter multivariado surge nas variáveis múltiplas (múltiplas combinações de variáveis), e não somente no número de variáveis ou observações.

A análise multivariada conduz aos seguintes objetivos:

- a) Redução de dados ou simplificação estrutural - o fenômeno estudado deve ser representado da maneira mais simples possível, sem sacrificar valiosas informações.
- b) Ordenação e agrupamento - agrupamento de objetos, tratamentos, ou variáveis similares baseados em dados amostrais ou experimentais.
- c) Investigação da dependência entre variáveis - o estudo das relações estruturais entre variáveis muitas vezes é de interesse do pesquisador.
- d) Predição - relações entre variáveis devem ser determinadas para o propósito de predição de uma ou mais variáveis com base na observação de outras variáveis.
- e) Construção e teste de hipóteses.

A classificação das técnicas multivariadas depende:

- da dependência ou independências das variáveis;
- do número de variáveis dependentes;
- do tipo de escala (métrica, não-métrica).

Técnica de dependência é aquela na qual uma variável ou um conjunto de variáveis é identificado como variável dependente a ser predita ou explicada por outras variáveis independentes. Podem ser classificadas pelo número de variáveis dependentes e pelo tipo de escala das variáveis. Podem-se citar como exemplos: regressão linear múltipla, MANOVA, análise discriminante, análise de correlação canônica e análise conjunta.

Técnica de interdependência é aquela na qual nenhuma das variáveis é definida como dependente ou independente, mas o procedimento envolve a análise simultânea de todas as variáveis no conjunto. Dependem da estrutura procurada: estrutura de variáveis, de casos ou de objetos (da matriz de dados). Podem-se citar como exemplos: análise fatorial, análise de *cluster*, escalonamento multidimensional e análise de correspondências.

Como as técnicas multivariadas que serão empregadas para o desenvolvimento deste trabalho são a análise de agrupamento, a análise de componentes principais e a análise fatorial, as mesmas serão expostas detalhadamente a fim de que sejam mais bem compreendidas.

- Análise de agrupamentos (AA)

Análise de agrupamentos (*cluster analysis*) é usada para descrever diversas técnicas numéricas cujo principal propósito é classificar os valores de uma matriz de dados em estudo em grupos discretos. A técnica multivariada da análise de agrupamentos pode ser empregada para explorar as similaridades entre os indivíduos (modo Q) ou entre as variáveis (modo R). Nesse método buscam-se agrupamentos homogêneos de itens representados por pontos num espaço n-dimensional em um número conveniente de grupos.

Os métodos para a análise de agrupamentos podem ser classificados em quatro tipos, segundo Davis (1986):

- métodos de partição: procuram classificar regiões no espaço, definido em função de variáveis, que sejam densamente ocupados em termos de observações daqueles com ocupações mais esparsa;
- métodos com origem arbitrária: procuram classificar as observações segundo “k” conjuntos previamente definidos;
- métodos por similaridade mútua: procuram agrupar observações que tenham uma similaridade comum com outras observações;
- métodos por agrupamentos hierárquicos: são as técnicas mais empregadas em Geologia. A partir da matriz inicial de dados obtém-se uma matriz simétrica de similaridades e inicia-se a detecção de pares de casos com a mais alta similaridade ou mais baixa distância.

A análise de *cluster* busca agrupar elementos de dados baseando-se na similaridade entre eles. Aqui, ocorre uma ordenação ou um agrupamento de objetos ou variáveis. Os grupos são determinados de forma a obter-se um elevado grau de homogeneidade dentro dos grupos e um alto nível de heterogeneidade entre eles.

A técnica de agrupamento hierárquico interliga as amostras por suas associações, produzindo um dendograma onde as amostras semelhantes, segundo as variáveis escolhidas, são agrupadas entre si. A suposição básica de sua interpretação é: quanto menor a distância entre os pontos, maior a semelhança entre as amostras.

Os dendogramas são especialmente úteis na visualização de semelhanças entre amostras ou objetos.

Em alguns estudos, se torna necessário conhecer algumas características de determinado grupo de um conjunto de elementos amostrais, principalmente quando é resultante de uma ou mais variáveis. Quando se obtém mensurações de diferentes naturezas, pode-se observar se existem similaridades no conjunto de dados.

A finalidade da análise de agrupamentos é reunir os objetos verificados nos grupos em que exista homogeneidade dentro do grupo e heterogeneidade entre os grupos. Os objetos de um grupo são relativamente semelhantes e diferentes de objetos de outros grupos. Quando utilizada dessa forma, a AA é inversa a Análise de Fatores, pelo fato de reduzir o número de objetos e não o número de variáveis, concentrando-se em um número muito menor de grupos.

Para a aplicação da AA, as estatísticas e os conceitos seguintes serão utilizados.

- Dendograma ou fenograma: também denominado gráfico em árvore. Representa uma síntese gráfica do trabalho desenvolvido, acarretando uma perda de informações. Mesmo com essa perda, o gráfico é muito útil para a classificação, comparação e discussão de agrupamentos.

A representação de um dendograma pode ser horizontal ou vertical.

No dendograma horizontal, o eixo y representa os grupos unidos por ordem decrescente de semelhança, e a posição da reta, no eixo x, indica as distâncias entre os grupos que foram formados. O dendograma é lido de cima para baixo.

O dendograma vertical é lido da direita para a esquerda. A interpretação dos resultados é análoga.

Para a realização da AA, os dados são reunidos numa tabela de m colunas (variáveis) x n linhas (objetos). Antes de escolher a medida de distância para a análise dos dados, é necessário verificar se os mesmos se encontram com a mesma unidade de medida. Quando isso não ocorrer, deverá ser feita a padronização dos mesmos.

A partir disso, é escolhida uma medida de distância apropriada que determinará semelhanças ou diferenças entre os objetos agrupados.

Existem diversos métodos de aglomeração e deve-se escolher o mais adequado ao estudo.

Sabe-se que um método é melhor do que outro quando o dendograma fornece uma imagem menos distorcida da realidade. O cálculo do coeficiente de correlação cofenético avalia o grau de deformação causado pela construção do dendograma (VALENTIM, 2000). Isto é, o menor grau de distorção será refletido pelo maior coeficiente cofenético, fornecido pela matriz fenética F, onde os seus valores são obtidos a partir da matriz de distâncias inicial, e pela matriz cofenética C, onde os seus valores são obtidos a partir da matriz de distâncias final.

A escolha das variáveis empregadas no processo de aglomeração constitui a parte mais importante da análise de agrupamentos. Segundo Bussab *et al* (1990) critérios distintos levam a grupos homogêneos distintos, e o tipo de homogeneidade depende dos objetivos a serem alcançados.

Inicialmente deve-se verificar se os dados necessitam de tratamento. Quando as variáveis tiverem unidades de medidas distintas, deve-se proceder a padronização das mesmas pois as unidades das variáveis podem alterar o grau de similaridade dos conglomerados. A padronização faz com que as variáveis tenham a mesma contribuição no cálculo do coeficiente de similaridade. Isso é feito segundo a equação (1):

$$Z_{ij} = \frac{X_{ij} - \bar{X}_j}{S_j} \quad (1)$$

Na análise de agrupamentos é fundamental a escolha do coeficiente que quantifique quanto dois objetos são semelhantes. O coeficiente pode ser de similaridade ou de dissimilaridade.

Coeficiente de similaridade: quanto maior o valor, mais semelhantes serão os objetos.

Coeficiente de dissimilaridade: quanto maior o valor, menos semelhantes serão os objetos.

Medidas de similaridade:

A similaridade entre objetos é uma medida de correspondência ou semelhança entre os objetos a serem agrupados, podendo ser avaliada de diversas maneiras:

- medidas correlacionais, como por exemplo correlação de Pearson;
- medidas de distância, como por exemplo distância euclidiana;

- medidas de associação, como por exemplo índice de Jaccard.

Cada uma dessas formas representa uma perspectiva particular da similaridade, de acordo com os seus objetivos e do tipo de dados.

Tanto as medidas correlacionais quanto as medidas de distância requerem dados métricos, ao passo que as medidas de associação requerem dados não-métricos.

As medidas correlacionais representam similaridades pela correspondência de padrões ao longo dos atributos.

As medidas de distância representam as similaridades como sendo a proximidade entre as observações ao longo dos atributos. Na realidade, essas medidas são de dissimilaridade na qual os valores maiores denotam uma menor similaridade. A distância é convertida em similaridade usando-se uma relação inversa ($1 - \text{distância}$).

Segundo Souto (2005) as medidas de distância se concentram na importância dos valores e representam casos similares que estão próximos, mas podem ter padrões muito diferentes ao longo dos atributos. Já os agrupamentos baseados em medidas correlacionais podem ter padrões similares.

Em geral, a distância entre pares de objetos é usada para avaliar a semelhança ou diferença entre os objetos. Menores distâncias indicam maiores semelhanças e vice-versa. A distância euclidiana fornece tal medida de semelhança.

Malhotra (2001) afirma que quando diferentes medidas de distância são empregadas, os resultados de aglomeração podem ser diferentes. Por isso, devem-se utilizar medidas diferentes e comparar os resultados obtidos.

Alguns coeficientes utilizados para determinar o conceito de distância entre os objetos são: distância euclidiana, distância euclidiana média, distância de Mahalanobis e coeficiente de Pearson. O coeficiente a ser utilizado no desenvolvimento deste trabalho será distância euclidiana

a) Distância euclidiana

É a medida de distância mais utilizada na análise de agrupamentos.

Dado n indivíduos, onde cada um deles possui valores para p variáveis, a distância euclidiana entre eles é determinada pelo teorema de Pitágoras, num espaço multidimensional.

Segundo Manly (1986) quando a distância euclidiana é determinada a partir das variáveis originais, pode ser influenciada pela escala de medida, pelo número de variáveis e pela correlação existente entre as mesmas. Assim, é feita a padronização das variáveis, para que possuam variância igual a um (1).

Dados dois indivíduos i e i' , a distância entre eles é dada pela equação (2):

$$d_{ii'} = \left[\sum_{j=1}^p (X_{ij} - X_{i'j})^2 \right]^{1/2} \quad (2)$$

Após a escolha de uma medida de distância, a próxima etapa é a escolha de um processo de agrupamento, uma vez que existem vários métodos.

Os processos de agrupamentos podem ser hierárquicos ou não-hierárquicos.

Malhotra (2001) afirma que o processo de aglomeração hierárquica é caracterizado pelo estabelecimento de uma hierarquia, ou uma estrutura em forma de árvore. Os métodos hierárquicos dividem-se em aglomerativos e divisivos. O agrupamento aglomerativo inicia em um grupo separado. Os grupos são formados, e os objetos são reunidos em grupos cada vez maiores. O método é seqüencial e continua até que todos os objetos estejam reunidos em um único grupo, segundo o critério de similaridade entre si.

O agrupamento divisivo inicia em um grande grupo, onde os objetos são subdivididos em dois subgrupos, nos quais deve haver a máxima semelhança entre os elementos de um mesmo subgrupo e a maior dissimilaridade entre os elementos dos subgrupos diferentes. Posteriormente, os subgrupos são subdivididos em outros subgrupos dissimilares e, assim, sucessivamente até que existam tantos subgrupos quanto o número de elementos (MALHOTRA, 2001).

Os métodos aglomerativos são formados por métodos de encadeamento, métodos de erros de somas de quadrados, métodos de variância e métodos centróides.

Os métodos de encadeamento podem ser de encadeamento único, médio e completo.

O método do encadeamento_único (single linkage) tem como base a distância mínima entre os objetos, ou a regra do vizinho mais próximo, sendo a distância entre dois grupos calculada a partir dos dois pontos mais próximos.

O método do encadeamento completo (complete linkage) é semelhante ao anterior, porém tem como base a distância máxima entre os objetos, ou a regra do vizinho mais afastado, sendo a distância entre dois grupos calculada a partir dos dois pontos mais afastados.

O método do encadamento_médio é similar aos anteriores. Nesse método, a distância entre dois grupos é definida como a média da distância entre todos os pares de objetos. Aqui, cada membro de um par de objetos provém de cada um dos grupos.

Aqui serão apresentados os métodos de agrupamento hierárquicos aglomerativos (*Linkage Methods*). Inicialmente, objetos mais similares são agrupados formando um único

grupo. Esse processo é repetido até que todos os subgrupos estejam formados, originando um grupo único contendo todos os objetos.

Serão empregados os métodos de ligação simples (distância mínima ou do vizinho mais próximo) e de ligação completa (distância máxima ou do vizinho mais distante).

Ferreira (1996) apresentou um algoritmo geral para os agrupamentos hierárquicos aglomerativos com n objetos (itens ou variáveis). As etapas são:

- Iniciar o agrupamento com n grupos, cada um contendo um único elemento, e com uma matriz simétrica $n \times n$ de dissimilaridades (distâncias) $D = \{d_{hi}\}$.
- Procurar na matriz D o par de grupos mais similar e fazer a distância entre os grupos mais similares U e V igual à d_{uv} .
- Unir os grupos U e V e nomeá-los por (UV) . Novamente, calcular e arranjar as distâncias na matriz D , eliminando as linhas e colunas correspondentes a U e V , e adicionando uma linha e coluna com as distâncias, entre o grupo (UV) e os demais grupos.
- Repetir a segunda e a terceira etapa $(n - 1)$ vezes até que todos os objetos estejam num grupo único.

O desenvolvimento da análise de agrupamentos pode ser realizado pelo método de ligação simples ou método de ligação completa.

Independentemente do método empregado para resumir os dados, devem ser feitas medidas do grau de ajuste entre a matriz original dos coeficientes de distância e a matriz resultante do agrupamento (ROHLF, 1970, apud REGAZZI, 2001).

Segundo alguns autores, o grau de agrupamento pode ser considerado bom quando estiver acima de 7,0 e o método será considerado inadequado para valores menores do que 7,0.

Cabe ressaltar que quanto maior for o grau de ajuste, menor é a distorção resultante da aplicação do método.

Para Valentim (2000), um método é considerado melhor do que outro quando o dendograma produzido apresentar uma imagem menos distorcida da realidade. O autor afirma que sempre existirá uma distorção, pois o valor de r nunca será igual à unidade. Assim, quanto maior for o valor de r , menor será a distorção resultante.

O grau de deformação pode ser avaliado pelo coeficiente de correlação cofenético que mede o grau de ajuste entre a matriz de dissimilaridade (matriz fenética F) e a matriz resultante do método de agrupamento (matriz cofenética C). Tal coeficiente é o coeficiente de correlação de Pearson (r).

Conclui-se que quando o coeficiente de correlação de Pearson (r) for menor do que 0,7, o método empregado é inadequado para resumir a informação do conjunto de dados. Com isso, outros métodos devem ser utilizados para a análise dos dados.

Valentim (2000) define três regras básicas que devem ser utilizadas na interpretação de um dendograma:.

- na frente de cada amostra anotar, no dendograma, as suas características, o que poderá revelar as igualdades entre as amostras de um mesmo grupo e as diferenças com relação às amostras de outro grupo.
- iniciar a leitura do dendograma a partir dos menores valores de similaridade. Com isso, inicialmente os grupos maiores deverão ser interpretados, a partir da formulação de uma hipótese razoável, para posteriormente explicar os grupos menores.
- se possível desenvolver, simultaneamente, um análise de ordenação que evidenciará os fatores responsáveis pelos agrupamentos.

A análise de agrupamentos classifica os objetos de modo que exista grande semelhança no agrupamento (grupo ou *cluster*) em relação a algum critério de seleção estabelecido anteriormente. Os grupos formados apresentam uma elevada homogeneidade dentro dos grupos e elevada heterogeneidade entre os grupos.

O objetivo principal da análise de agrupamentos é definir a estrutura dos dados dispondo as observações mais semelhantes em grupos. Aqui, três questões básicas deverão ser abordadas:

- as formas de medir a similaridade: correlação, distância, medida de associação
- a formação dos grupos (*clusters*): procedimento para reunir as observações mais similares
- o número de grupos formados: menor número de grupos e maior homogeneidade.

Uma maneira para representar um agrupamento hierárquico é através de um dendograma. O eixo horizontal representa o coeficiente de aglomeração, nesse caso a distância usada para unir grupos.

A Figura 3 apresenta os passos realizados durante uma análise de agrupamentos.

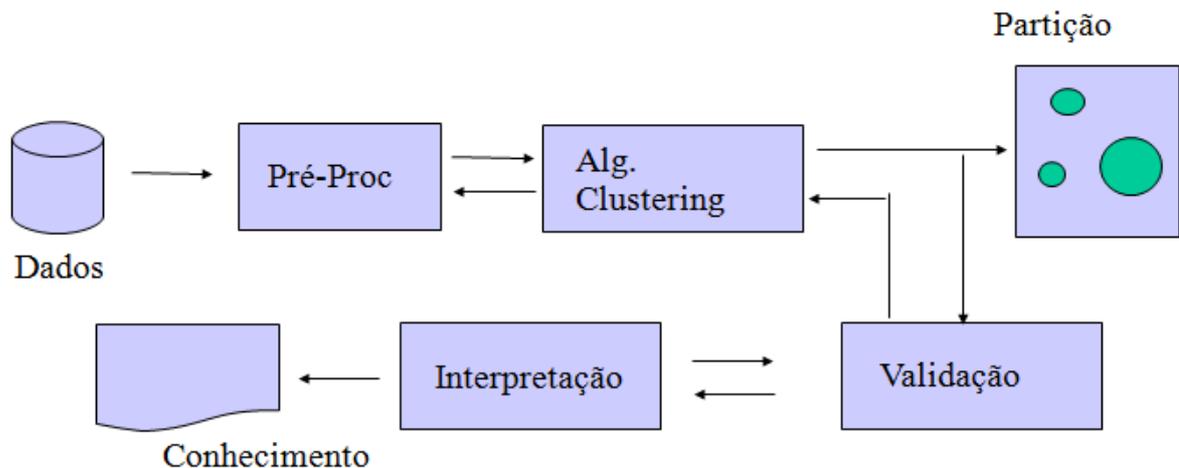
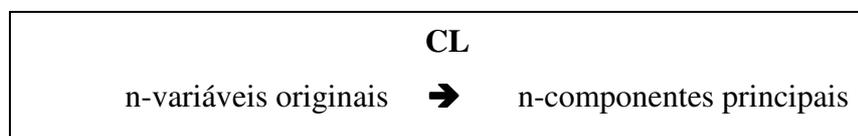


Figura 3 – Passos na análise de agrupamentos (adaptado de Souto, 2006)

- Análise de componentes principais (ACP)

Ocorre uma redução de dados ou uma simplificação estrutural. Consiste em reescrever as coordenadas das amostras em outro sistema de eixos mais conveniente para a análise dos dados.

Através de combinações lineares, as n variáveis originais são transformadas em n componentes principais.



As características da ACP são:

- Requisitos da amostra: nº de amostras deveria ser 10 a 25 vezes maior que o nº de variáveis (por isso utiliza-se o método das componentes principais).
- Normalização das Variáveis - $Z = (X - \bar{X}) / S'$.
- Adequabilidade do método fatorial – $KMO \geq 0,6$ (pelo menos, razoável).
- Cálculo da matriz de correlações.
- Cálculo dos autovalores e autovetores - proporção da variância explicada.
- Seleção do nº de componentes: critério de Kaiser (autovalores >1) e critério do *Scree Plot*
- Interpretação dos factor loadings (com rotação varimax – maximiza a associação das variáveis com um único fator).

Segundo Pontes (2005), a ACP compreende uma das mais antigas técnicas de análise multivariada que tem por finalidade a análise de estruturas de covariâncias e correlações em

matrizes simétricas positivas. Ou seja, essa ferramenta realiza a transformação de variáveis correlacionadas em outro grupo de variáveis não correlacionadas e ainda tem a capacidade de reduzir a dimensionalidade dos dados. Nesse sentido, que a ACP pode ser usada na análise final ou intermediária dos dados.

A análise de agrupamento hierárquico (análise de *cluster*) e de componentes principais são técnicas de estatística multivariada complementares que têm grande aceitação na análise de dados.

É uma técnica que possibilita a análise com um grande número de dados disponíveis e, também, identificar as medidas responsáveis pelas variações relevantes entre os resultados, com as menores perdas de informações possíveis.

A ACP transforma um conjunto de variáveis originais em outro, os componentes principais (CP) com dimensões equivalentes, buscando eliminar algumas das variáveis originais que contenham pouca informação. Cabe ressaltar que a redução de variáveis iniciais somente é possível quando essas não forem independentes e possuírem coeficiente de correlação não-nulos.

Recentemente, uma das principais aplicações da ACP ocorre quando as variáveis originam-se a partir de processos, onde diversas características devem ser observadas simultaneamente. A técnica vem sendo desenvolvida por Morrison (1976), Jackson (1980, 1981), Johnson & Wichern (1992, 1998), entre outros.

Uma análise de componentes principais tem por objetivo a substituição de um conjunto de variáveis correlacionadas por um outro conjunto de novas variáveis não correlacionadas, onde essas são combinações lineares das variáveis originais, dispostas em ordem decrescente de variâncias. (VERDINELLI, 1980).

$$\text{VAR CP}_1 > \text{VAR CP}_2 > \dots > \text{VAR CP}_p$$

onde:

VAR CP₁ representa a variância da componente principal 1

VAR CP₂ representa a variância da componente principal 2

...

VAR CP_p representa a variância da componente principal p

Na determinação das CP, calcula-se a matriz de variância-covariância (Σ) ou a matriz de correlação (R), encontram-se os autovalores (Λ) e os autovetores (x) e, finalmente,

escrevem-se as combinações lineares, ou seja, as componentes principais, onde cada uma delas é uma combinação linear de todas as variáveis originais (REGAZZI, 2001).

A primeira etapa da ACP é a transformação da matriz de dados $n \times p$ em uma matriz quadrada, onde n é o número de indivíduos e p representa um conjunto de variáveis. A interdependência entre as variáveis é representada pela matriz de variância-covariância (Σ) ou a matriz de correlação R .

Souza (2000) afirma que, em casos populacionais, para realizar a ACP é utilizada a matriz Σ , associada ao vetor aleatório $\vec{X} = [X_1, X_2, \dots, X_p]$. No entanto, em casos amostrais, utiliza-se a matriz de variância-covariância amostral (S), e o vetor média é dado por $\vec{\bar{X}} = [\bar{X}_1, \bar{X}_2, \dots, \bar{X}_p]$.

Cabe salientar que a maioria dos problemas práticos só são resolvidos quando mais de uma técnica multivariada é aplicada. Dentre as técnicas mais empregadas estão: análise de agrupamentos (AA), análise de componentes principais (ACP), análise discriminante (AD), análise de correspondência (AC), entre outras.

Segundo Reis (1997), a aplicação da ACP e AF deverá incluir:

- as variáveis envolvidas na análise;
- as percentagens da variância explicadas pelas componentes principais;
- o número de componentes retidas e a proporção de variância total explicada pelas mesmas;
- uma tabela onde conste a contribuição de cada variável para cada componente (*factor loadings*), antes e após a aplicação de um método de rotação de fatores, e
- a interpretação e cada componente principal retida.

No emprego da ACP algumas etapas deverão ser seguidas para a obtenção do resultado final.

Primeiramente, calcula-se a matriz S (matriz de variância-covariância amostral) ou R (matriz de correlação) e verifica-se se existe uma correlação entre as variáveis. Se as variáveis não estiverem correlacionadas, aplica-se o teste do KMO ou outro teste para verificar se as correlações são significativas ou não, para proceder a análise dos dados através desta técnica.

Também cabe ao pesquisador investigar a escala em que foram medidas as variáveis. Quando estiverem em escalas diferentes, as variáveis devem ser padronizadas para que se evitem erros nos resultados.

A próxima etapa serve para decidir o número total de componentes que explicarão da melhor forma o conjunto de variáveis originais. As duas maneiras para selecionar os componentes são:

- através dos autovalores, pelo critério de Kaiser (1960) apud Mardia (1979) que inclui somente as componentes com valores próprios superiores a um (1). Geralmente as componentes utilizadas conseguem sintetizar uma variância acumulada aproximadamente igual a setenta por cento (70%);

- através do método gráfico, o critério sugerido por Cattell (1966) e exemplificado por Pla (1986) considera as componentes anteriores ao ponto de inflexão da curva.

Uma vez selecionado o número de componentes, a próxima etapa consiste em encontrar os autovetores que formarão as combinações lineares e, posteriormente, as novas variáveis.

A última etapa constitui a normalização e a ortogonalização dos autovetores, o que garantirá uma solução única das componentes principais e uma independência entre elas.

A matriz de variância-covariância é expressa pelas ligações realizadas entre as p variáveis, apanhadas duas a duas e sendo resumidas por suas covariâncias.

A matriz S é chamada de matriz de variância-covariância das p variáveis, onde os termos da diagonal principal são as variâncias (S_i^2) e os demais termos são as covariâncias (S_{ij}).

A matriz de correlação é empregada quando é necessária uma padronização dos dados, para que sejam evitados problemas tais como a influência da magnitude das variáveis (JACKSON, 1981 apud SOUZA, 2000).

Segundo Johnson & Wichern (1992), a solução empregando a matriz de correlação poderá ser utilizada quando as variáveis possuírem escalas muito distintas, uma vez que essa matriz é equivalente a matriz das variáveis padronizadas.

- Análise Fatorial (AF)

Segundo Pereira, Júnior & Ramos (2006) a análise fatorial é um processo estatístico que se destina à redução e a sintetização dos dados. Foi desenvolvida por Charles Spearman em 1904.

Malhotra (2001) afirma que a análise fatorial deve ser utilizada nas seguintes situações:

- identificação das dimensões latentes ou fatores que expliquem as correlações entre um conjunto de variáveis;
- identificação de um novo conjunto, menor do que o inicial, de variáveis não correlacionadas em substituição ao conjunto de variáveis correlacionadas na análise multivariada seguinte (regressão ou análise discriminante);

- identificação, em um conjunto maior, de um conjunto menor de variáveis que se destacam para uso em uma análise multivariada posterior.

Cada variável pode ser expressa como uma combinação linear de fatores subjacentes. Entende-se por comunalidade como sendo o compartilhamento de uma variável com todas as outras variáveis analisadas.

Segundo Malhotra (2001), a análise fatorial possui duas abordagens básicas na escolha do método de aplicação: a análise de componentes principais (ACP) e a análise fatorial (AF).

A ACP leva em consideração a variância total dos dados, onde sua maior preocupação é a identificação do número mínimo de fatores que levam a máxima variância nos dados para o emprego na análise que se pretende realizar. Já a AF leva em consideração a variância comum nos dados. Sua principal função é a identificação das dimensões subjacente, onde a variância comum é um fator de interesse

Segundo Kassai & Kassai (2000), o desenvolvimento de técnicas estatísticas a partir de cálculos de regressão linear admite a solução de problemas que contenham variáveis quantitativas e qualitativas.

A Tabela (12) a seguir demonstra as principais estatísticas de análise fatorial, determinando o número de fatores, contribuindo para a escolha de um modelo mais adequado.

Tabela 12 – Estatísticas associadas à análise fatorial

Determinação	Considerações
A priori	Em virtude de conhecimento prévio, o pesquisador sabe quantos fatores pode esperar, o que permite especificar o número de fatores a serem extraídos.
Autovalores	São retidos apenas os fatores com autovalores superiores a 1,0.
Porcentagem da variância	Recomenda-se que os fatores respondam, no mínimo, por 60% da variância.
Confiabilidade meio a meio	A amostra é dividida ao meio, fazendo-se uma análise fatorial sobre cada metade. São retidos apenas os fatores com elevada correspondência de cargas fatoriais ao longo de duas subamostras.
Testes de significância	É possível determinar a significância estatística dos autovalores separados, retendo apenas os fatores que são estatisticamente significativos.
Kaisen – Meyer – Olkin (KMO)	Valores altos (entre 0,5 e 1,0) indicam que a análise fatorial é adequada.
Comunalidades	Porção da variância explicada pelos

fatores comuns.

Fonte: Corrar (2002)

A análise fatorial tem como objetivos principais:

- analisar as inter-relações entre um grande número de variáveis em termos de poucas, mas não observáveis, variáveis chamadas fatores.
- verificar se é possível descrever um conjunto de p variáveis em um conjunto menor de índices ou fatores que explicam tanto sobre o fenômeno, que o conjunto original.

Supõe-se que as variáveis possam ser agrupadas segundo suas correlações, isto é, que todas as variáveis dentro de um grupo sejam altamente correlacionadas entre si, mas tenham correlações muito baixas com as variáveis de outros grupos. A idéia através da análise fatorial é que cada grupo de variáveis representa um *constructo* básico, que é o responsável pelas correlações observadas nas respostas.

Não existe uma solução única para a AF de um conjunto de dados, mas sim dois princípios básicos que devem ser considerados:

- Princípio de Parcimônia: tem-se que explicar as correlações existentes entre as variáveis observadas utilizando o menor número de fatores possível.
- Interpretabilidade: deseja-se que os fatores tenham um significado no contexto estudado, conservando em si uma coerência lógica.

Os tipos de aplicação e de entendimento da AF são:

- análise fatorial exploratória: é a mais antiga. Buscam-se dimensões subjacentes, para que se conheça o que é mais importante ou mais significativo em um conjunto de variáveis.
- análise fatorial confirmatória: esboça-se uma estrutura dos fatores e, a seguir, busca-se a confirmação desta, analisando as variáveis observadas (usado em modelagem de comportamento).

Inicialmente, o pesquisador possui várias medidas realizadas, não sendo possível identificar quais variáveis poderão ser reunidas num fator. Isso é feito pela AF, que permite identificar um número reduzido de novas variáveis, sem perda significativa de informações.

A AF se refere a um conjunto de técnicas relacionadas que tornam mais clara a interpretação dos dados observados. Os inter-relacionamentos entre as variáveis são analisados, de maneira que essas variáveis possam ser descritas por um grupo de categorias básicas chamado fatores.

Na extração de fatores, um dos métodos mais empregados é o da análise de componentes principais. Os fatores gerados são empregados de maneira representativa do processo.

A AF é aplicada na identificação de fatores para um determinado conjunto de medidas realizadas. Também é utilizada como ferramenta na tentativa de redução de um grande conjunto de variáveis em um conjunto mais significativo, que é representado pelos fatores. Determina-se que variáveis pertencem aos fatores, e o quanto cada fator é explicado por cada uma das variáveis.

Ressalta-se que quando a AF é empregada para amostras pequenas, os coeficientes de correlação tendem a ser menos confiáveis. Assim, para cada variável observada deve-se ter um mínimo de cinco (5) casos.

Quando a AF é aplicada, inicialmente verificam-se as relações entre as variáveis. Para cada par de variáveis, pode-se empregar o coeficiente de correlação linear como uma medida de associação.

A AF é dita exploratória quando for utilizada com o intuito de redução da dimensão dos dados. Também pode ser confirmatória quando for utilizada para testar a hipótese inicial de redução dos dados a uma determinada dimensão e de qual a distribuição de variáveis, segundo essa dimensão (REIS, 1997).

Antes da aplicação da AF, consideram-se algumas premissas com relação à natureza dos dados.

Quanto à adequação dos dados, uma medida importante foi sugerida por Kaiser-Meyer-Olkin Measure of Adequacy (KMO). O valor do KMO é determinado pela expressão (3):

$$KMO = \frac{\sum_i \sum_j r_{ij}^2}{\sum_i \sum_j r_{ij}^2 + \sum_i \sum_j a_{ij}^2} \quad (3)$$

Se os intervalos obtidos estiverem no intervalo de 0,5 a 0,9, as variáveis poderão ser utilizadas na AF.

Segundo Pereira (2001), o KMO é uma medida de adequação dos dados, na qual todas as medidas são utilizadas simultaneamente. O teste do KMO estabelece que:

- valores próximos a 0,90: a adequação dos dados é considerada ótima;
- valores próximos a 0,80: a adequação dos dados é considerada boa;

- valores próximos a 0,70: a adequação dos dados é considerada razoável;
- valores próximos a 0,60: a adequação dos dados é considerada medíocre;
- valores próximos a 0,50: a adequação dos dados é considerada imprópria.

O KMO é uma medida de adequação na qual o ajuste dos dados é verificado, empregando simultaneamente todas as variáveis.

O teste de Bartlett (BTS – *Bartlett Test of Sphericity*) também pode ser empregado para AF, no qual é testada a hipótese de que a matriz de correlação seja uma matriz identidade.

O teste de Bartlett pressupõe que a hipótese nula seja rejeitada.

O princípio da AF pressupõe que cada variável pode ser decomposta em: uma parte comum, onde a sua variação é partilhada com outras variáveis, e uma parte única, onde a sua variação é específica.

Assim, uma diferença entre a ACP e a AF a partir da variância analisada parte do princípio de que enquanto a primeira considera a variação total presente no conjunto de variáveis originais, na segunda somente é retida a variação comum a todas as variáveis (REIS, 1997).

Os fatores podem estar relacionados com uma variável individual (fatores únicos) ou com duas ou mais variáveis originais (fatores comuns). As cargas relacionam uma associação específica entre os fatores e as variáveis originais. Assim, encontram-se as cargas e a solução para os fatores, aproximando a relação entre as variáveis originais e os fatores encontrados, sendo as cargas derivadas dos autovalores, associados às variáveis individuais.

Para uma melhor visualização das variáveis que melhor representam os fatores, realiza-se uma rotação dos eixos para auxiliar na interpretação dos fatores. Diversas rotações podem ser feitas na matriz fatorial: varimax, quartimax e equimax, que são rotações ortogonais e as rotações oblíquas que são não-ortogonais.

Segundo Pereira (2001) uma rotação na matriz fatorial não afeta as comunalidades das variáveis (inércia) nem as percentagens das variações explicadas pelos fatores.

Dois conceitos são importantes na interpretação da AF: ortogonalidade e carga fatorial.

A ortogonalidade está relacionada com a independência, onde deve haver uma dissociação entre as variáveis. Para isso, geralmente se realiza uma ACP para a extração de fatores, onde cada componente é independente.

A carga fatorial é um coeficiente, positivo ou negativo, normalmente menor do que um, que expressa o quanto um teste está carregado em um fator. Ou seja, quanto maior a carga em cima de um fator, mais a variável se identifica com o fator.

Assim, a AF é um método empregado na determinação do número de fatores que existem em um conjunto de dados. Também determina que testes ou variáveis pertencem aos respectivos fatores.

De acordo com Pereira (2001) serão apresentados alguns conceitos resultantes da AF, para uma interpretação correta dos dados. São eles:

- *Eigenvalue* (valor próprio): correlativo aos autovalores e a variância total, que pode ser explicada pelo fator. Isto é, avalia a contribuição do fator ao modelo estabelecido através da análise fatorial. Quando a explicação da variância pelo fator for alta, significa que existe uma alta explicação desse fator ao modelo. Quando a explicação for baixa, significa que existe uma baixa explicação do fator ao modelo.
- *Factor loading* (carregamento de fator): corresponde à proporção de variação da variável que é explicada pelo fator, ou também, o quanto cada variável colabora na formação de cada componente.
- *Factor Score* (contagem do fator): os autovetores determinam as direções dos eixos da máxima variabilidade. Representam a medida assumida pelos objetos examinados na função derivada da análise.
- *Communality* (comunalidade): é a medida de quanto da variância de uma variável é explicada pelos fatores derivados pela análise fatorial. A contribuição da variável ao modelo construído pela AF é avaliada, isto é, determina o quanto cada variável tem participação na formação da outra. Para as comunalidades, os maiores valores são os mais importantes para análise.
- *Factor matrix* (Matriz de fator): representa a matriz de correlação entre as variáveis originais e os fatores encontrados

Em uma matriz de correlação a multicolinearidade e a singularidade podem estar presentes. A multicolinearidade ocorre quando as variáveis são altamente correlacionadas (acima de 0,90). Nela, existe uma inter-relação entre as variáveis independentes. A singularidade está presente quando uma variável está perfeitamente relacionada às outras.

4 METODOLOGIA

Neste capítulo será apresentada a localização geográfica da região em estudo bem como a metodologia utilizada para a definição do universo da pesquisa e avaliação estatística dos resultados de ensaios de produtos cerâmicos realizados por Tomazzeti (2003).

Durante o desenvolvimento teórico deste trabalho serão apresentados os conteúdos referentes a algumas técnicas de análise multivariada, resultantes de uma revisão bibliográfica.

Após o desenvolvimento teórico, mostrar-se-ão os empregos práticos para Engenharia Civil, resolvidos através do *software STATISTICA* para *Windows-Versão 7.0 (StatSoft, Inc.)*, com o objetivo de proporcionar a aplicação dos módulos adequados ao desenvolvimento da análise estatística de cada um dos exemplos.

4.1 Localização geográfica da região em estudo

Na Figura (4) a seguir é apresentada a localização geográfica da região central do estado do RS, onde estão localizadas as indústrias cerâmicas consultadas por Tomazzeti (2003).

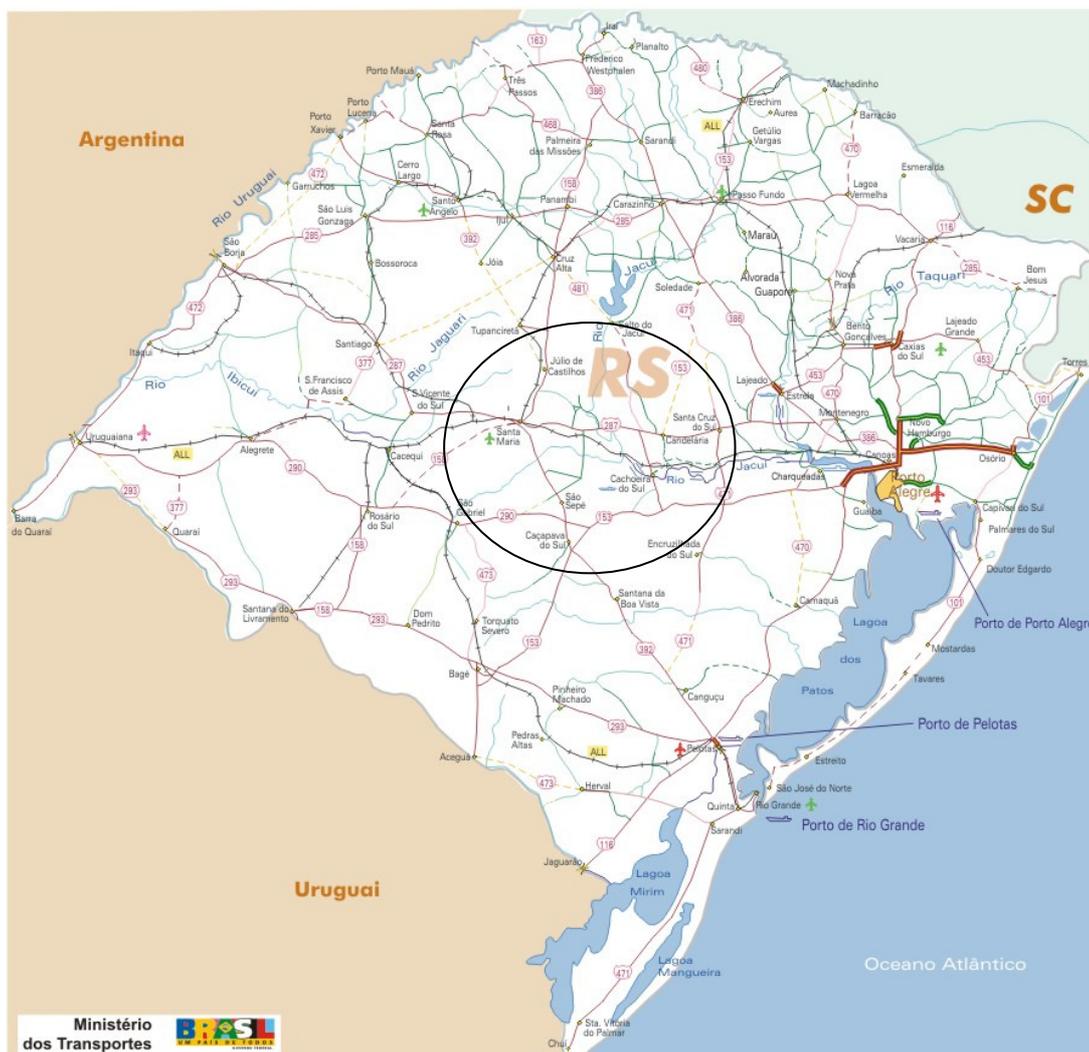


Figura 4 - Mapa de localização da região central do estado do RS

4.2 Dados utilizados na avaliação dos produtos cerâmicos

Os dados empregados neste trabalho são provenientes da análise de ensaios de produtos cerâmicos fabricados na Região Central do Estado do Rio Grande do Sul, resultantes de um trabalho desenvolvido por Tomazzeti (2003). São decorrentes dos ensaios de sessenta e seis (66) amostras provenientes de trinta e duas (32) indústrias cerâmicas localizadas nos vinte e cinco (25) municípios da região central do RS.

Para a definição do universo da pesquisa, Tomazzeti (2003) adotou como critério um número de indústrias proporcional a quantidade existente nos municípios considerados, adotando uma (1) olaria por município, obtendo trinta e duas (32) indústrias cerâmicas no total. A partir disso, foram coletadas as amostras dos produtos nas olarias, utilizando aleatoriedade na escolha das mesmas, totalizando vinte e quatro (24) unidades para cada tipo. Foram coletadas um total de sessenta e seis (66) amostras de produtos assim constituídas:

vinte e oito (28) unidades de tijolos cerâmicos, trinta e duas (32) unidades de blocos de vedação e uma (1) unidade de bloco estrutural.

a) Delimitação do Universo

Segundo Tomazetti (2003) o levantamento das indústrias de cerâmica vermelha existentes na região foi realizado junto aos cadastros das Prefeituras Municipais, Sindicatos das Indústrias de Olaria e de Cerâmica para Construção no Estado do Rio Grande do Sul – SIOCERGS e informações obtidas junto às próprias olarias.

Com as informações obtidas, através da aplicação de questionários as sessenta e cinco (65) empresas identificadas, foi elaborado um panorama sócio-econômico da indústria de cerâmica vermelha da região.

As oitenta e quatro (84) indústrias cerâmicas identificadas estão presentes em vinte e seis (26) dos trinta e cinco (35) municípios que compõem a região, representando setenta e quatro por cento (74%).

Serão analisados resultados dos ensaios dimensionais e de caracterização, ensaios físicos e mecânicos para tijolos maciços, blocos de vedação e blocos portante. A partir disso, serão aplicadas técnicas estatísticas multivariadas aos dados disponíveis, a fim de que sejam realizadas as análises sugeridas e a conseqüente apreciação dos resultados obtidos para verificação dos objetivos propostos no presente estudo.

4.3 Ensaios realizados no controle tecnológico dos produtos

Segundo Tomazetti (2003), ao final do processo produtivo, deve haver uma verificação dos produtos quanto à resistência mecânica, absorção de água, dimensões e cor para que seja realizado um controle técnico dos produtos. Esse controle serve para examinar se os produtos acabados seguem ou não as normas técnicas exigidas.

A seguir, estão descritas as metodologias empregadas nos ensaios dimensionais e de caracterização, ensaios físicos e ensaios mecânicos, todos esses realizados para a avaliação dos produtos analisados.

a) Ensaios dimensionais e de caracterização.

a.1) Avaliação dimensional. Segue as recomendações das normas NBR 7170/83 para tijolos maciços e NBR 7171/92 para blocos. Obtiveram-se a largura (L), altura (H) e comprimento (C) de tijolos, blocos e tabelas. Para as telhas foi empregada a norma NBR 8038/83 para telhas francesas e NBR 13582/96 para telhas romanas.

Para os blocos de vedação e tabelas foram determinadas as espessuras das paredes internas e externas.

Foram determinados os desvios de esquadro e planeza das faces para tijolos, blocos e tabelas.

a.2) Inspeção visual e sSom. A inspeção visual verifica a existência de fissuras, trincas, quebras e a cor predominante dos produtos. Com relação ao som, os produtos foram classificados em metálico forte, metálico e cavo, segundo o grau de percussão emitido pela peça cerâmica analisada.

b) Ensaio Físicos

b.1) Absorção de Água. Utilizada para quantificar a porosidade do produto cerâmico, a qual está relacionada à resistência e a aderência. O ensaio foi realizado para todas as amostras: tijolos, blocos, tabelas e telhas segundo a norma NBR 8947/85.

A absorção de água é determinada segundo a equação (4):

$$\text{Abs}(\%) = \frac{M_u - M_s}{M_s} \times 100 \quad (4)$$

onde:

Abs (%) representa a absorção, medida em porcentagem

Mu representa a massa do tijolo ou bloco saturado, medida em kilograma (kg)

Ms representa a massa do tijolo seco, medida em kilograma (kg)

b.2) Taxa de absorção inicial (IRA). Representa uma medida normalizada da velocidade inicial de absorção de água. Significa a avidez de água que o tijolo ou bloco possui e sua tendência de retirá-la da argamassa de assentamento. O IRA indica se existe a necessidade de molhar o bloco ou o tijolo antes do assentamento.

O ensaio é realizado segundo a norma americana ASTM C67-78.

O cálculo do IRA é realizado pela equação (5):

$$\text{IRA} = \frac{193,55(m_i - m_s)}{A} \quad (5)$$

onde:

IRA representa a taxa de absorção inicial (g/193,55 cm²/min)

mi representa a massa do tijolo imerso sob lâmina de água durante 1 minuto, medida em gramas (g)

ms representa a massa do tijolo seco em estufa durante 24 horas , medida em gramas (g)

A representa a área da face do tijolo que está em contato com a água, em centímetros quadrados (cm²)

b.3) Impermeabilidade. O ensaio de verificação da impermeabilidade de telhas cerâmicas é realizado segundo a norma NBR 8948/85.

c) Ensaio Mecânicos

c.1) Resistência à compressão. É determinada pela norma NBR 6460/83 para tijolo maciço e NBR 6461/83 para bloco cerâmico. Foram realizados ensaios de resistência à compressão para tijolos maciços, blocos de vedação e tabelas.

c.2) Resistência à flexão. Foi determinada para telhas e tabelas segundo a norma NBR 6462/83.

A Tabela (13) a seguir apresenta as informações sobre a distribuição das empresas e a produção de cerâmica vermelha nos municípios.

Tabela 13 - Distribuição das empresas e produção de cerâmica vermelha nos municípios.

Municípios	Nº de Indústrias Cerâmicas	Produção Mensal (milheiros)					Produção Mensal Total
		Tijolos Maciços	Blocos de Vedação	Bloco Estrutural	Telhas	Tabelas	
Agudo	1	80	150	-	-	-	230
Cacequi	1	40	-	-	-	-	40
Cachoeira do Sul	4	134	577	6	113	23	853
Capão do Cipó	1	20	-	-	-	-	20
Cerro Branco	-	-	-	-	-	-	0
Dilermando de Aguiar	-	-	-	-	-	-	0
Dona Francisca	1	5	-	-	-	-	5
Faxinal do Soturno	2	50	130	-	-	130	310
Formigueiro	1	40	60	-	-	-	100
Itaara	1	50	-	-	-	-	50
Ivorá	-	-	-	-	-	-	0
Jaguari	2	40	50	-	-	-	90
Jari*	14	213	5	-	-	-	218
Júlio de Castilhos	5	147	-	-	-	-	147
Mata	1	30	45	-	-	-	75
Nova Esperança do Sul	-	-	-	-	-	-	0
Nova Palma	1	55	-	-	-	-	55
Novo Cabrais	1	70	-	-	-	-	70
Paraíso do Sul	1	-	90	-	-	-	90
Pinhal Grande	-	-	-	-	-	-	0
Quevedos*	1	-	-	-	-	-	0
Restinga Seca	4	52	350	-	-	-	402
Santa Maria	19	850	2.724	108	1.054	370	5.106
Santiago*	5	144	371	-	-	-	515
São Francisco do Sul	2	25	-	-	-	-	25
São João do Polesine	4	227	-	-	-	-	227
São Martinho da Serra	1	7	30	-	-	-	37
São Pedro do Sul	3	64	-	-	-	-	64
São Sepé	4	250	40	-	-	10	300
São Vicente do Sul	1	-	85	-	-	-	85
Silveira Martins	-	-	-	-	-	-	0
Toropi	-	-	-	-	-	-	0
Tupanciretã	-	-	-	-	-	-	0
Unistalda	-	-	-	-	-	-	0
Vila Nova do Sul	3	95	-	-	-	-	95
Total	84	2.688	4.707	114	1.167	533	9.209

Fonte: TOMAZETTI (2003)

As Tabelas (14) a (18), extraídas de Tomazetti (2003), contendo os dados a serem utilizados no desenvolvimento do presente trabalho são apresentadas a seguir.

Tabela 14 – Dimensões médias para blocos de vedação e bloco portante

Município	Olaria	Bloco	L (mm)	H (mm)	C (mm)	Ih (mm)	Iv (mm)	Eh (mm)	Ev (mm)
Agudo	1.1	4F	115,00	104,38	221,25	10,35	10,65	13,53	13,05
Agudo	1.1	6F	94,17	142,13	188,92	10,35	9,25	11,30	11,28
Agudo	1.1	6F	104,46	153,33	201,54	8,98	8,95	9,75	9,23
Cachoeira	3.1	6F	102,50	174,42	216,21	12,40	8,90	15,08	11,63
Cachoeira	3.1	6F	112,75	173,88	230,63	11,13	10,55	13,35	12,10
Cachoeira	3.1	10F	124,75	257,92	251,42	10,00	8,03	10,38	11,30
Cachoeira	3.1	BE	141,46	193,17	289,13	-	-	20,65	20,65
Cachoeira	3.3	2F	118,04	61,29	235,63	-	13,35	10,58	14,28
Cachoeira	3.3	4F	120,58	121,83	232,21	14,28	13,08	12,80	13,68
Cachoeira	3.3	6F	120,04	148,33	206,04	11,28	8,85	11,63	11,93
Cachoeira	3.3	6F	114,58	173,08	232,67	13,33	10,95	13,25	13,13
Faxinal	7.2	6F	91,67	135,38	177,33	10,50	9,00	10,78	10,23
Faxinal	7.2	6F	104,13	148,04	206,17	9,75	9,25	10,15	9,90
Formigueiro	8.1	4F	103,33	103,75	209,58	9,55	8,95	10,70	11,43
Formigueiro	8.1	6F	105,00	151,04	206,88	9,38	9,05	11,55	10,18
Jaguari	11.1	6F	97,88	135,67	192,92	10,93	12,70	10,05	11,35
Mata	14.1	6F	87,08	135,46	199,96	10,50	9,60	11,10	10,40
Mata	14.1	6F	86,79	135,75	202,54	9,53	8,25	8,95	8,55
Paraíso	18.1	4F	113,08	113,00	234,79	8,73	8,85	11,00	11,28
Paraíso	18.1	6F	112,29	168,63	234,96	10,58	9,65	11,73	10,35
Paraíso	21.1	6F	94,29	130,46	190,46	9,58	8,25	10,73	10,95
Rest. Seca	21.3	2F	110,17	61,33	232,75	-	10,90	12,30	12,55
Rest. Seca	21.3	4F	109,13	107,92	221,04	9,85	9,18	13,23	12,78
Rest. Seca	21.3	6F	92,71	128,17	189,17	10,33	11,79	11,30	11,38
Rest. Seca	21.3	6F	102,25	148,33	213,04	10,05	9,30	11,45	12,13
Santiago	22.1	6F	97,96	146,58	207,50	9,95	8,78	11,95	12,00
Santiago	22.2	4F	99,71	103,50	208,13	12,88	11,28	14,03	12,78
Santiago	22.2	6F	101,21	149,17	210,83	9,93	10,30	12,13	12,48
Santiago	22.3	4F	119,46	121,88	224,46	12,40	12,05	14,60	14,73
Santiago	22.3	6F	101,33	144,38	224,42	10,95	12,15	12,88	14,03
Santiago	22.3	6F	119,29	175,75	222,00	13,58	11,60	13,98	14,65
São Martinh	25.1	6F	90,96	140,21	182,71	11,65	10,50	12,80	11,70
São Sepé	27.2	6F	98,79	133,21	197,08	10,30	11,38	12,88	14,88

FONTE: TOMAZETTI (2003)

onde: L = largura, H = Altura, C = comprimento, Ih = Espessura das paredes internas horizontais, Iv = espessura das paredes internas verticais, Eh = Espessura das paredes externas horizontais, Ev = espessura das paredes externas verticais, 2F, 4F, 6F 10F = número de furos dos blocos de vedação, BE = bloco portante de alvenaria.

Tabela 15 - Absorção de água e IRA dos tijolos maciços

Município	Olaria	Absorção de água (%)	Absorção coef. de variação (%)	IRA	IRA coef. de variação(%)
Agudo	1.1	11,38	11,73	37,89	18,26
Cachoeira	3.1	13,30	8,96	37,25	46,73
Cachoeira	3.3	12,36	2,00	19,04	27,60
Capão Cipó	4.1	19,66	3,47	35,00	16,87
DonaFrancisca	6.1	14,12	1,91	58,78	20,08
Formigueiro	8.1	10,31	3,32	40,81	23,74
Itaara	9.1	24,24	5,96	50,36	13,87
Jaguari	11.2	11,04	7,64	53,15	47,92
Jarí	12.1	25,40	2,55	35,31	15,35
Jari	12.2	23,61	7,93	53,54	22,46
J. de Castilhos	13.1	17,54	2,26	31,46	10,47
J. de Castilhos	13.8	16,19	2,04	39,86	19,52
Mata	14.1	13,11	4,49	93,66	34,34
Nova Palma	16.1	17,42	3,94	39,52	41,42
Novo Cabrais	17.1	11,88	2,79	49,74	32,16
Restinga Seca	21.3	17,13	3,22	24,41	37,27
Santiago	22.1 L	19,98	1,31	47,47	16,32
Santiago	22.1 R	20,52	1,71	32,56	19,77
Santiago	22.2	20,83	5,33	96,61	13,67
São Francisco	23.2	14,01	6,70	54,87	18,05
São Francisco	23.3	10,33	5,79	43,44	19,29
São João Polês	24.2	14,48	6,14	62,89	36,11
São João Polês	24.4	12,00	6,53	67,99	38,53
São Martinho	25.1	21,68	4,05	32,24	37,11
São P. do Sul	26.3	15,96	5,35	60,17	29,88
São Sepé	27.2	15,65	5,67	41,56	44,75
São Sepé	27.4	14,50	16,47	25,22	36,76
V. Nova Sul	33.2	18,04	7,56	37,67	28,35

FONTE: TOMAZETTI (2003)

onde: IRA = taxa de absorção inicial de água

Tabela 16 - Absorção de água e IRA dos blocos de vedação e bloco portante

Município	Olaria	Bloco	Absorção de água (%)	Absorção coef. de variação (%)	IRA	IRA coef. de variação(%)
Agudo	1.1	4F	11,95	5,56	26,92	21,39
Agudo	1.1	6F	12,04	2,72	32,30	32,97
Agudo	1.1	6F	11,09	2,40	24,22	10,29
Cachoeira	3.1	6F	14,79	11,60	36,55	34,39
Cachoeira	3.1	6F	15,44	2,60	33,27	35,25
Cachoeira	3.1	10F	13,28	7,62	28,27	43,38
Cachoeira	3.1	BE	14,74	4,40	19,57	23,37
Cachoeira	3.3	2F	12,26	2,05	18,40	22,08
Cachoeira	3.3	4F	13,66	4,21	18,36	24,44
Cachoeira	3.3	6F	13,33	5,52	20,87	25,81
Cachoeira	3.3	6F	13,92	3,31	22,26	31,05
Fax. do Sot.	7.2	6F	11,16	7,14	18,81	38,61
Fax. do Sot.	7.2	6F	11,52	6,11	24,60	55,09
Formigueiro	8.1	4F	9,73	3,01	24,32	11,88
Formigueiro	8.1	6F	10,24	3,08	31,69	22,02
Jaguari	11.1	6F	11,60	5,22	17,12	25,63
Mata	14.1	6F	10,01	19,92	23,54	33,98
Mata	14.1	6F	13,64	4,85	17,56	35,48
Paraíso	18.1	4F	12,16	5,10	34,27	18,76
Paraíso	18.1	6F	13,13	4,01	27,03	10,73
Rest. Seca	21.1	6F	17,26	2,14	14,63	13,30
Rest. Seca	21.3	2F	16,77	3,34	18,79	34,63
Rest. Seca	21.3	4F	16,59	3,17	15,33	26,04
Rest. Seca	21.3	6F	16,75	4,96	19,07	16,24
Rest. Seca	21.3	6F	16,31	3,44	16,72	22,59
Santiago	22.1	6F	20,71	2,51	27,33	10,28
Santiago	22.2	4F	19,00	4,48	61,24	19,86
Santiago	22.2	6F	19,61	4,93	45,84	29,30
Santiago	22.3	4F	19,41	1,35	12,42	14,74
Santiago	22.3	6F	19,32	7,59	13,49	26,16
Santiago	22.3	6F	20,72	1,92	58,45	11,72
S. Martinho	25.1	6F	23,08	3,39	25,06	20,34
São Sepé	27.2	6F	15,45	7,51	31,34	31,99

FONTE: TOMAZETTI (2003)

onde: IRA = taxa de absorção inicial de água

Tabela 17 - Resistência mecânica de tijolos maciços

Município	Olaria	Resist. a compressão (MPa)	Coefficiente de variação (%)
Agudo	1.1	9,83	13,55
Cach. do Sul	3.1	8,28	22,37
Cach. Do Sul	3.3	11,22	7,01
Capão do Cipó	4.1	16,57	16,40
Dona Francisca	6.1	10,58	10,32
Formigueiro	8.1	15,18	8,68
Itaara	9.1	8,84	41,28
Jaguari	11.2	10,47	26,99
Jari	12.1	11,20	10,79
Jarí	12.2	15,82	39,07
J. de Castilhos	13.1	3,23	13,67
J. de Castilhos	13.8	6,04	21,72
Mata	14.1	8,66	27,88
Nova Palma	16.1	16,43	16,60
Novo Cabrais	17.1	8,94	18,98
Restinga Seca	21.3	10,77	22,64
Santiago	22.1 L	18,68	6,73
Santiago	22.1 R	13,31	8,55
Santiago	22.2	12,40	13,96
São Francisco	23.2	8,76	37,36
São Francisco	23.3	9,33	26,13
S. J. do Polêsine	24.2	11,40	19,18
S. J. do Polêsine	24.4	9,95	24,45
São Martinho	25.1	8,03	18,81
São Pedro do Sul	26.3	10,74	11,95
São Sepé	27.2	4,97	33,87
São Sepé	27.4	10,97	24,32
Vila Nova do Sul	33.2	9,98	14,42

FONTE: TOMAZETTI (2003)

Tabela 18 - Resistência mecânica de blocos de vedação e bloco portante

Município	Olaria	Bloco	Resist. a compressão (MPa)	Coefficiente de variação (%)
Agudo	1.1	4F	1,60	19,67
Agudo	1.1	6F	1,68	25,21
Agudo	1.1	6F	0,50	14,43
Cachoeira	3.1	6F	1,23	25,69
Cachoeira	3.1	6F	1,08	16,14
Cachoeira	3.1	10F	1,28	27,17
Cachoeira	3.1	BE	9,93	20,44
Cachoeira	3.3	2F	3,00	18,63
Cachoeira	3.3	4F	1,52	7,70
Cachoeira	3.3	6F	1,03	29,55
Cachoeira	3.3	6F	0,90	26,63
Fax. do Soturno	7.2	6F	0,75	25,91
Fax. do Soturno	7.2	6F	0,83	25,17
Formigueiro	8.1	4F	1,67	10,14
Formigueiro	8.1	6F	0,58	27,29
Jaguari	11.1	6F	0,65	21,76
Mata	14.1	6F	2,21	28,51
Mata	14.1	6F	1,96	34,75
Paraíso	18.1	4F	1,75	18,22
Paraíso	18.1	6F	0,82	33,67
Rest. Seca	21.1	6F	0,71	31,60
Rest. Seca	21.3	2F	2,22	8,92
Rest. Seca	21.3	4F	1,56	21,05
Res. Seca	21.3	6F	0,78	17,20
Rest. Seca	21.3	6F	0,74	27,04
Santiago	22.1	6F	0,87	23,18
Santiago	22.2	4F	1,92	17,08
Santiago	22.2	6F	1,84	25,04
Santiago	22.3	4F	1,93	23,27
Santiago	22.3	6F	1,65	51,48
Santiago	22.3	6F	1,16	24,11
São Martinho	25.1	6F	0,93	37,08
São Sepé	27.2	6F	1,76	30,80

FONTE: TOMAZETTI (2003)

4.4 Análises estatísticas

Para a análise estatística do conjunto de dados cerâmicos selecionado serão empregadas algumas técnicas de análise multivariada, adequadas para cada caso.

Para os dados de produção mensal total de cerâmica vermelha, número de indústrias cerâmicas, tijolos maciços, blocos de vedação, bloco estrutural, telhas e tabelas serão aplicadas as técnicas de análise de agrupamentos, análise fatorial e análise de componentes principais.

Para os dados de dimensões médias e características geométricas dos tijolos maciços e de dimensões médias de blocos de vedação e bloco portante, serão aplicadas as técnicas de análise de agrupamentos, análise fatorial e análise de componentes principais.

Para os dados referentes à absorção de água, IRA e resistência à compressão para tijolos maciços, blocos de vedação e bloco portante serão aplicadas as técnicas de análise de agrupamentos, círculo de correlação unitário e planos fatoriais.

A seguir, para o desenvolvimento da análise dos dados relativos à cerâmica vermelha será utilizado o *software Statistica 7.0* na determinação das estatísticas a serem determinadas.

5 RESULTADOS E DISCUSSÃO

Neste capítulo serão apresentados os resultados decorrentes da aplicação das técnicas estatísticas multivariadas aos dados referidos anteriormente e a discussão dos mesmos, para que posteriormente sejam obtidas as conclusões a partir dos objetivos inicialmente propostos.

5.1 Análise estatística dos dados totais da distribuição das empresas e produção de cerâmica vermelha

Para os dados contidos na Figura (5) seguinte, serão aplicadas as técnicas de análise multivariada: análise de agrupamento, análise fatorial e análise de componentes principais.

	1	2	3	4	5	6	7
	Nº de Ind. Cerâm.	Tijolos maciços	Blocos de ved.	Bloco estrut.	Telhas	Tavelas	Prod. mensal tot.
Agudo	1	80	150	0	0	0	230
Cacequi	1	40	0	0	0	0	40
Cachoeira do Sul	4	134	577	6	113	23	853
Capão do Cipó	1	20	0	0	0	0	20
Dona Francisca	1	5	0	0	0	0	5
Faxinal do Soturno	2	50	130	0	0	130	310
Formigueiro	1	40	60	0	0	0	100
Itaara	1	50	0	0	0	0	50
Jaguari	2	40	50	0	0	0	90
Jari	13	213	5	0	0	0	218
Júlio de Castilhos	5	147	0	0	0	0	147
Mata	1	30	45	0	0	0	75
Nova Palma	1	55	0	0	0	0	55
Novo Cabrais	1	70	0	0	0	0	70
Paraíso do Sul	1	0	90	0	0	0	90
Restinga Seca	4	52	350	0	0	0	402
Santa Maria	19	850	2724	108	1054	370	5106
Santiago	5	144	371	0	0	0	515
São Francisco de Assis	2	25	0	0	0	0	25
São João do Polésine	4	227	0	0	0	0	227
São Martinho da Serra	1	7	30	0	0	0	37
São Pedro do Sul	3	64	0	0	0	0	64
São Sepé	4	250	40	0	0	10	300
São Vicente do Sul	1	0	85	0	0	0	85
Vila Nova do Sul	3	95	0	0	0	0	95

Figura 5 – Dados da distribuição das empresas e produção de cerâmica vermelha nos municípios da região central do estado do RS

5.1.1 Análise de agrupamento das variáveis da produção de cerâmica vermelha

Para a realização da análise de agrupamento  Cluster Analysis, seleciona-se o método *Joining (tree clustering)*, as variáveis de interesse e as opções de análise *Advanced*.

Selecionando-se todas as variáveis e a opção  Vertical icicle plot obtém-se o dendograma da Figura (6) a seguir.

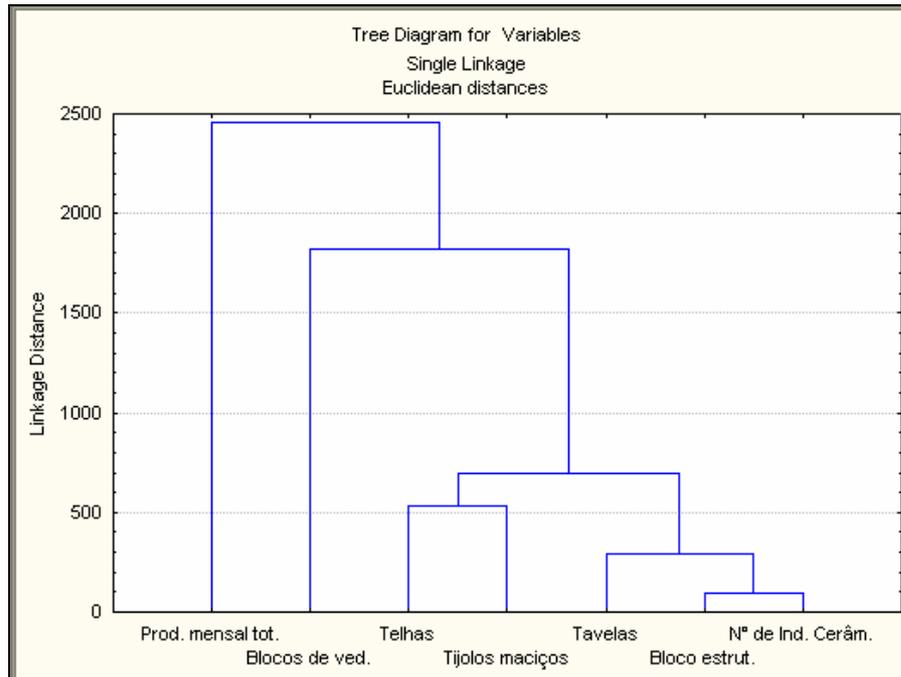


Figura 6 - Dendograma vertical dos dados da produção de cerâmica vermelha

No dendograma elaborado verifica-se a formação de 4 grupos distintos: um grupo composto pelas variáveis número de indústrias cerâmicas, bloco estrutural e tabelas; um grupo composto pelas variáveis tijolos maciços e telhas, um grupo composto pela variável bloco de vedação e outro grupo composto pela variável produção mensal total, da direita para a esquerda.

Selecionado-se a opção  **Amalgamation schedule** obtém-se a Figura (7):

Amalgamation Schedule (Spreadsheet1 in Dados dissertação.stw)							
Single Linkage							
Euclidean distances							
linkage distance	Obj. No. 1	Obj. No. 2	Obj. No. 3	Obj. No. 4	Obj. No. 5	Obj. No. 6	Obj. No. 7
90,74139	N° de Ind. Cerâm.	Bloco estrut.					
293,1433	N° de Ind. Cerâm.	Bloco estrut.	Tabelas				
534,5549	Tijolos maciços	Telhas					
696,7159	N° de Ind. Cerâm.	Bloco estrut.	Tabelas	Tijolos maciços	Telhas		
1824,750	N° de Ind. Cerâm.	Bloco estrut.	Tabelas	Tijolos maciços	Telhas	Blocos de ved.	
2455,360	N° de Ind. Cerâm.	Bloco estrut.	Tabelas	Tijolos maciços	Telhas	Blocos de ved.	Prod. mensal tot.

Figura 7 - Ligação entre as variáveis da produção de cerâmica vermelha

Selecionado-se a opção  **Graph of amalgamation schedule** obtém-se a Figura (8):

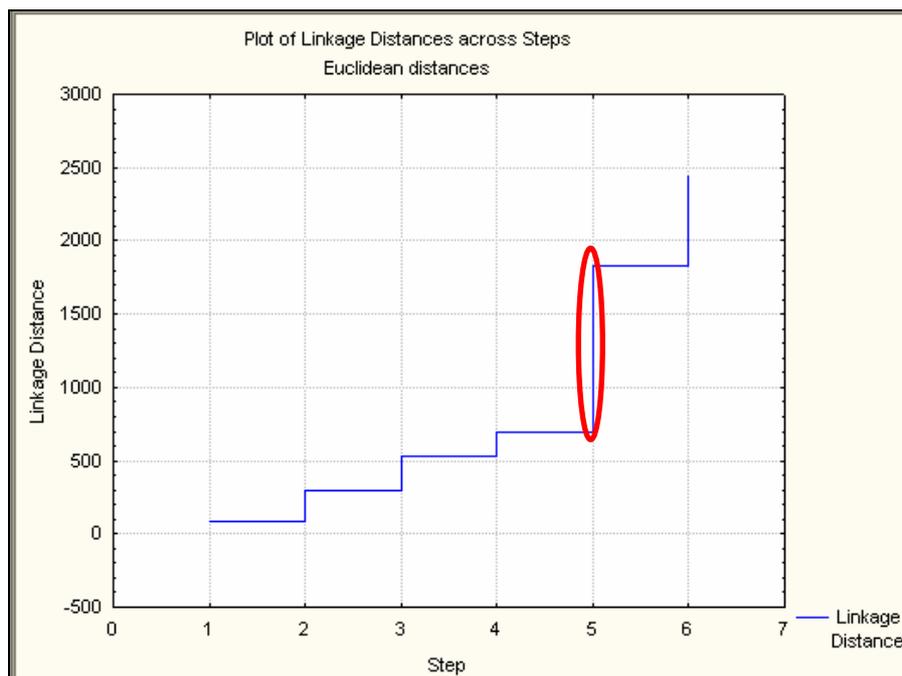


Figura 8 – Ligação das variáveis da produção de cerâmica vermelha

Analisando-se a figura 8, verifica-se que o maior salto ocorre entre as distâncias 696,7159 e 1824,750. Nesta parte do gráfico, poderia ser feito um corte entre essas distâncias onde observa-se a formação dos grupos distintos, conforme o histograma.

Selecionado-se a opção  Distance matrix obtem-se a matriz de distâncias entre as variáveis, conforme a Figura (9).

Variable	Euclidean distances (Spreadsheet1 in Dados dissertação.stw)						
	Nº de Ind. Cerâm.	Tijolos maciços	Blocos de ved.	Bloco estrut.	Telhas	Tabelas	Prod. mensal tot.
Nº de Ind. Cerâm.	0	968	2822	91	1041	375	5237
Tijolos maciços	968	0	2016	900	535	697	4360
Blocos de ved.	2822	2016	0	2738	1825	2481	2455
Bloco estrut.	91	900	2738	0	952	293	5153
Telhas	1041	535	1825	952	0	702	4221
Tabelas	375	697	2481	293	702	0	4889
Prod. mensal tot.	5237	4360	2455	5153	4221	4889	0

Figura 9 – Matriz de distâncias para produção de cerâmica vermelha

Selecionado-se a opção  Descriptive statistics obtem-se os valores de média e desvio padrão das variáveis, para um maior conhecimento e compreensão dos resultados obtidos, conforme a Figura (10).

Variable	Mean	Std.Dev.
Nº de Ind. Cerâm.	3,2800	4,169
Tijolos maciços	107,5200	170,188
Blocos de ved.	188,2800	547,433
Bloco estrut.	4,5600	21,583
Telhas	46,6800	211,070
Tabelas	21,3200	77,208
Prod. mensal tot.	368,3600	1005,396

Figura 10 – Estatística descritiva (média e desvio padrão) dos dados de cerâmica vermelha

Os resultados demonstram que a maior parte das variáveis apresenta um desvio padrão elevado, o que significa uma grande variação entre os resultados. Como a produção mensal total apresenta a maior variabilidade, optou-se por retirar essa variável algumas vezes durante a execução das análises.

Analisando-se os dados, excluindo-se a variável produção mensal total, obtem-se um novo dendograma conforme a Figura (11).

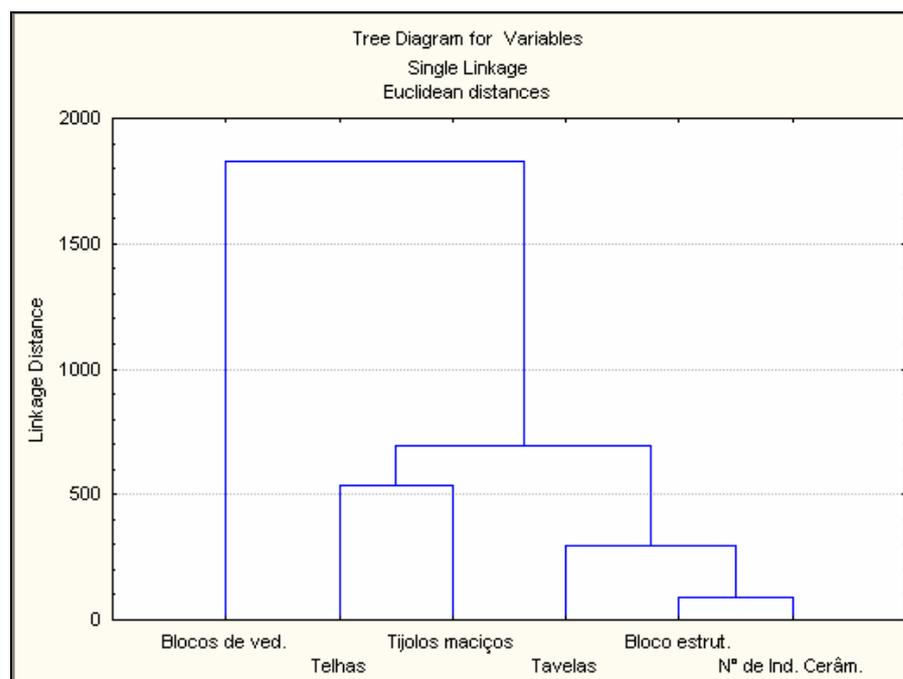


Figura 11 – Dendograma vertical dos dados, excluindo-se a variável produção mensal total

Analisando-se o dendograma anterior, verifica-se a formação de três grupos descritos, da direita para a esquerda:

- um grupo formado pelas variáveis número de indústrias cerâmicas, bloco estrutural e tabelas;
- um grupo formado pelas variáveis tijolos maciços e telhas; e
- um grupo formado pela variável bloco de vedação.

5.1.2 Análise fatorial das variáveis da produção de cerâmica vermelha

Para a realização da análise fatorial  Factor Analysis, selecionando-se as variáveis de interesse e a opção *Raw data*, surge uma tela com as opções *Quick*, *Advanced* e *Descriptives*, conforme a Figura (12).

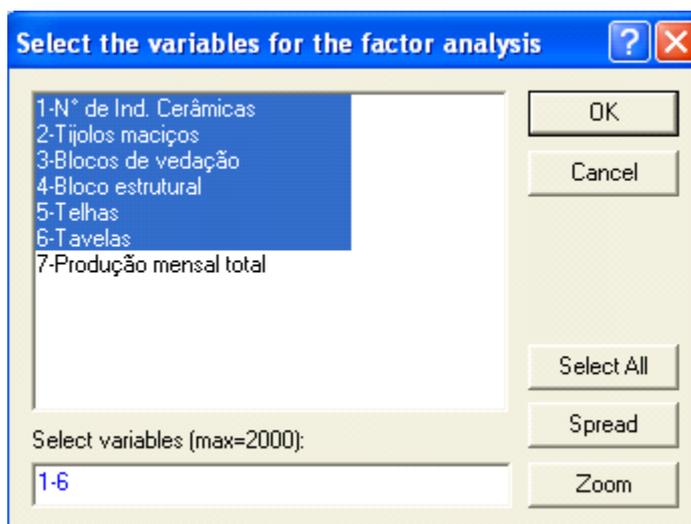


Figura 12 – Seleção das variáveis da produção de cerâmica vermelha

Seleciona-se *Quick*. Clicando-se em *OK* obtém-se a tela onde observa-se que existem várias opções para a realização da análise fatorial, que são: *Quick*, *Explained variance*, *Loadings*, *Scores* e *Descriptives*. Na opção *Quick*, existem os itens *Eigenvalues*, *Factor rotation*, *Summary: Factor loadings* e *Plot of factor loadings, 2D*.

Selecionando-se *Eigenvalues* (autovalores) obtém-se os seguintes resultados conforme a Figura (13):

Eigenvalues (Spreadsheet1 in Dados dissertação.stw)				
Extraction: Principal components				
Value	Eigenvalue	% Total variance	Cumulative Eigenvalue	Cumulative %
1	5,449461	90,82435	5,449461	90,82435
2	0,368834	6,14723	5,818295	96,97158
3	0,087356	1,45593	5,905651	98,42752
4	0,065888	1,09813	5,971539	99,52565
5	0,027520	0,45866	5,999059	99,98431

Figura 13 – Autovalor, % da variância total, autovalor acumulado e % acumulada para dados de cerâmica vermelha

Considerando-se as 6 variáveis e analisando-se os resultados obtidos verifica-se que somente o 1º autovalor é maior do que 1 e explica 90,82% da variância. Os demais autovalores são menores do que 1. Portanto, os dados serão resumidos pela 1ª componente principal.

Clicando-se em  ou em  obtém-se a tela a seguir (Figura 14).

Factor Loadings (Unrotated) (Spreadsheet1 in Dados dissertação.stw)					
Extraction: Principal components					
(Marked loadings are > ,700000)					
Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Nº de Ind. Cerâmicas	-0,869597	0,476525	0,063636	0,111141	-0,017959
Tijolos maciços	-0,955582	0,200470	-0,057797	-0,205021	0,036065
Blocos de vedação	-0,973455	-0,137945	-0,096176	0,100980	0,117908
Bloco estrutural	-0,984084	-0,138335	-0,065352	0,009068	-0,087599
Telhas	-0,984801	-0,139276	-0,076197	0,019678	-0,063704
Tabelas	-0,945651	-0,209776	0,246250	-0,028907	0,016197
Expl.Var	5,449461	0,368834	0,087356	0,065888	0,027520
Prp.Totl	0,908244	0,061472	0,014559	0,010981	0,004587

Figura 14 – Composição dos fatores para as variáveis de cerâmica vermelha analisadas

Na figura anterior verifica-se o resultado dos *Factor loadings* sendo que os valores destacados possuem uma significância maior do que 0,7, o que poderá ser alterado de acordo com as necessidades do pesquisador. Observa-se que todos os valores para o fator 1 são significativos, pois todos são maiores do que 0,7. No caso dos demais fatores os valores não são significativos.

Selecionando-se a opção , obtém-se a Figura (15) a seguir.

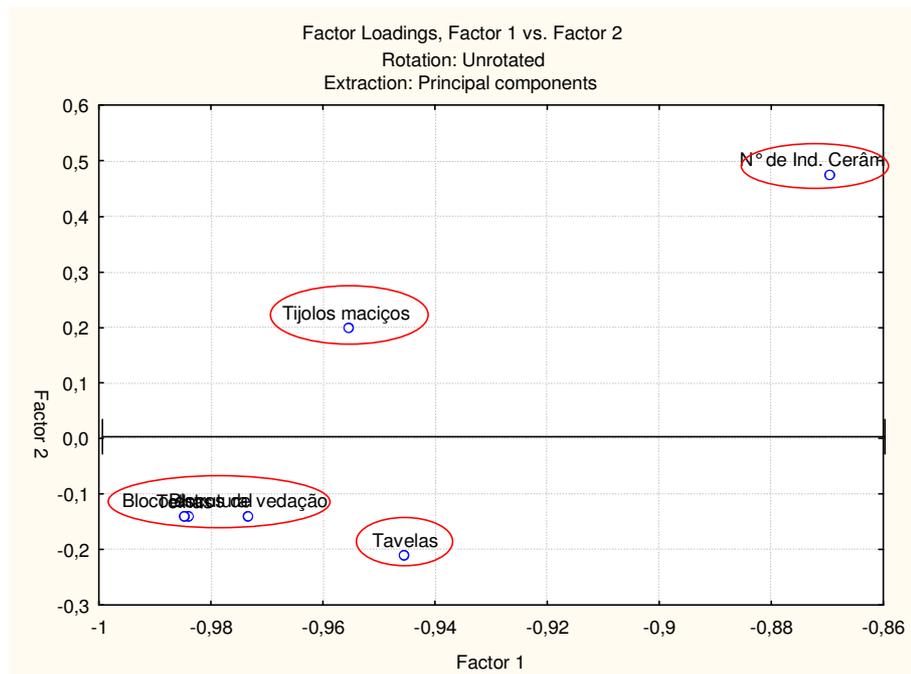


Figura 15 – Planos fatoriais para produção de cerâmica vermelha

Analisando a figura anterior, pode-se observar que as variáveis se agrupam por similaridades de explicação, ou seja, estão agrupadas por fatores.

As variáveis que explicam ou representam melhor o fator 1 são aquelas que estiverem localizadas mais distantemente da origem, em relação ao eixo x. As variáveis que estiverem localizadas próximas à origem do plano fatorial possuem baixa representatividade. Neste caso, num prolongamento do eixo x até a origem dos eixos, estas variáveis são Bloco de vedação e Bloco estrutural. Do mesmo modo, a variável que melhor representa o fator 2 é o número de indústrias cerâmicas.

Para encontrar as distâncias de cada uma das variáveis analisadas, traçam-se segmentos de reta perpendiculares ao eixo considerado e determina-se aquela ou aquelas variáveis que estiverem localizadas mais distantemente da origem, à partir da perpendicular traçada. Essas terão maior influência sobre o fator analisado. Isso pode ser observado nas Figuras (16) e (17) seguintes.

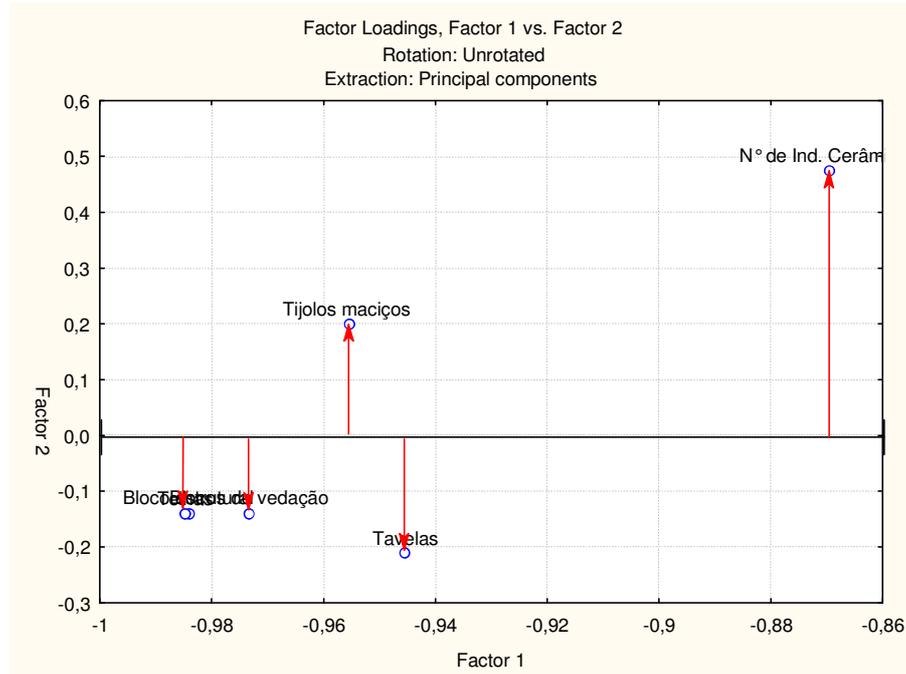


Figura 16 – Planos fatoriais contendo perpendiculares em relação ao fator 1, para produção de cerâmica vermelha

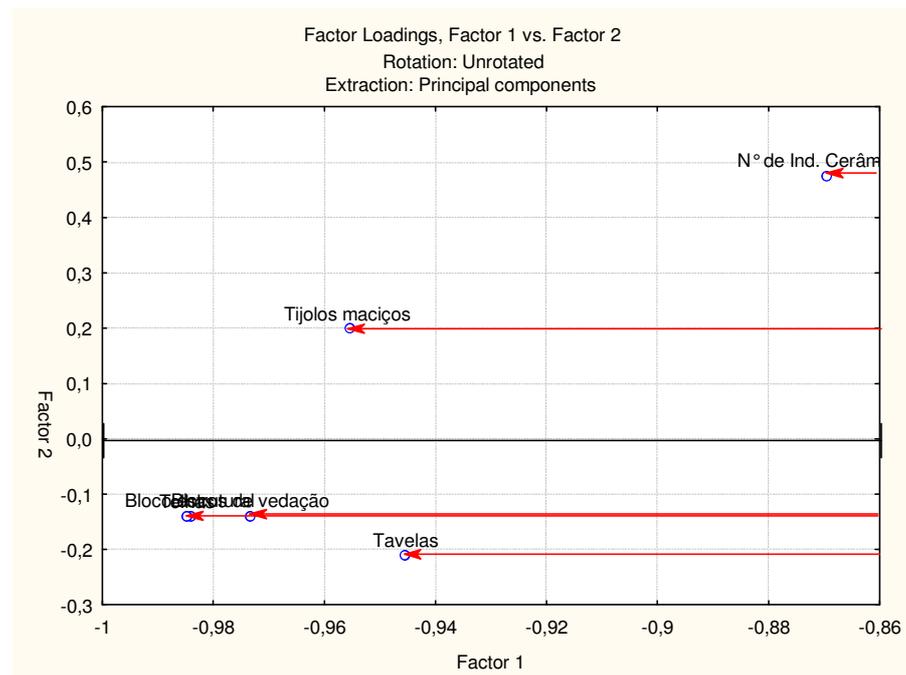


Figura 17 - Planos fatoriais contendo perpendiculares em relação ao fator 2, para produção de cerâmica vermelha

A visualização de ambos os gráfico, Figuras (16) e (17), confirmam o que foi colocado anteriormente. Aqui deve-se considerar que se a variância for nula ou aproximadamente nula, todos os indivíduos estarão próximos ou sobre a origem do plano principal, possuindo baixa representatividade.

Quando a identificação das variáveis que representam um fator tornar-se difícil, seja quando existir mais de um grupo de variáveis significativas no mesmo fator ou em fatores diferentes, realizam-se as rotações pois assim ocorre uma mais perfeita visualização da disposição dos pontos, ou seja, das variáveis de maior representatividade para cada componente. O tipo de rotação mais empregado é a *Varimax normalized*, que mantém os eixos ortogonais. Isto é feito selecionando-se *Factor rotation: Varimax normalized*.

Para uma melhor visualização dos fatores, pode-se utilizar os gráficos de dispersão ou os planos fatoriais. Seleciona-se *Plot of loadings, 2D*, conforme a Figura (18) a seguir.

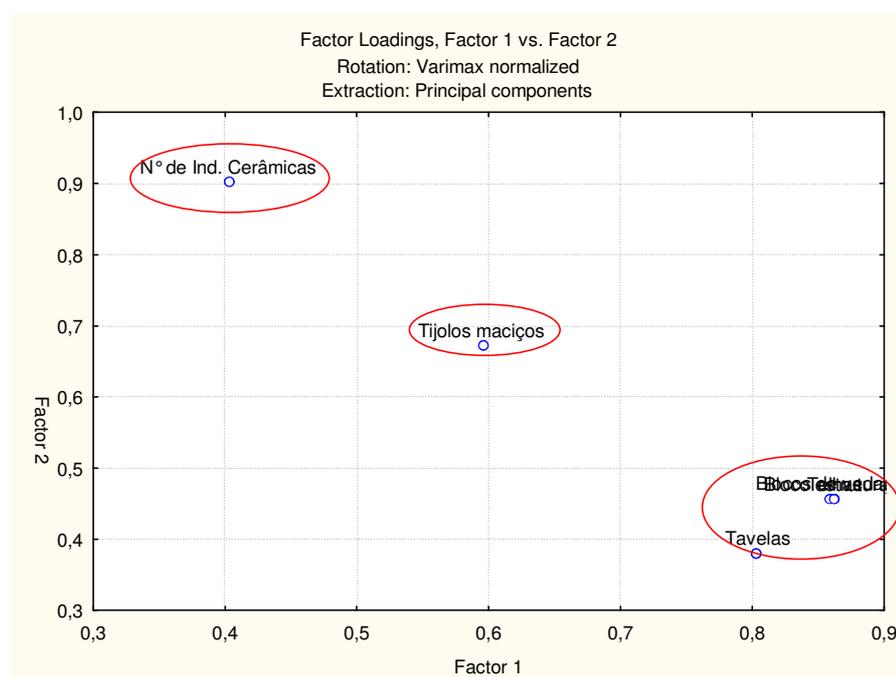


Figura 18 – Planos fatoriais, com rotação *varimax normalized* para dados de cerâmica vermelha

Assim, poderão ser identificadas as variáveis que possuem uma maior representatividade nos planos fatoriais. Analisando-se o gráfico verifica-se que as variáveis reúnem-se em grupos por similaridades de explicação, ou seja, por fatores. As variáveis de maior representatividade são aquelas localizadas mais distantemente da origem do plano fatorial. Aquelas que estiverem localizadas próximas à origem do plano fatorial, possuem baixa representatividade.

Na análise fatorial seleciona-se a opção *Explained variance*. A seguir, selecionando-se *Eigenvalues*  *Eigenvalues* obtém-se os mesmos resultados determinados anteriormente, segundo a Figura (13) anterior.

Selecionado-se *Scree plot*  *Scree plot* obtém-se o resultado conforme a Figura (19) a seguir.

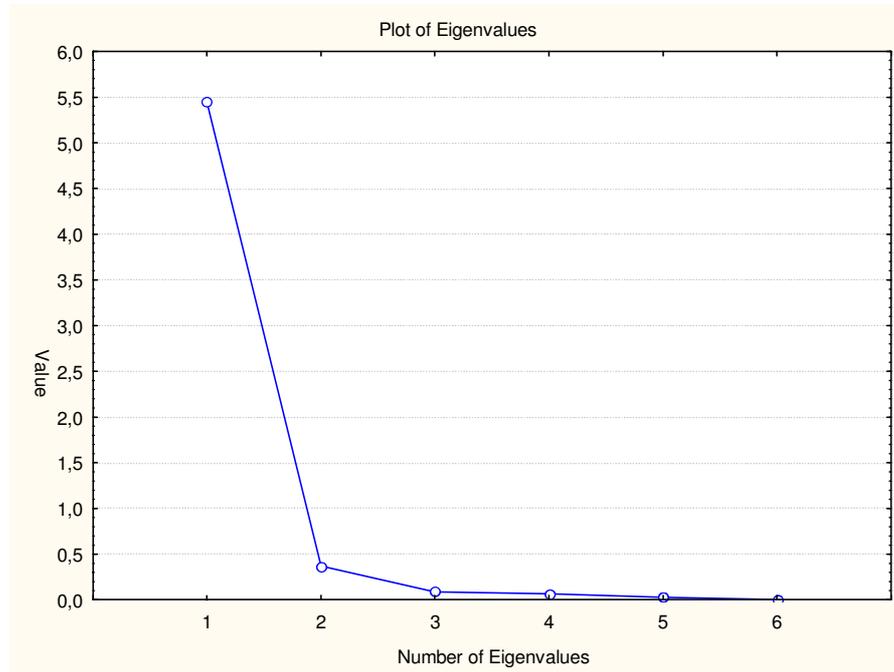


Figura 19 – Autovalores para dados de cerâmica vermelha

Este método gráfico sugerido por Cattell (1966) é utilizado para a seleção dos componentes principais. Examinando o gráfico anterior, observa-se uma queda brusca entre o 1º e o 2º fator. E, analisando os autovalores superiores a 1, observa-se que somente o 1º fator deverá ser considerado.

As comunalidades são determinadas porque são índices que expressam quanto da variância é explicada pela AF, ou seja, quanto mais próximo de 1, melhor será o ajuste da AF.

Selecionando-se *Communalities*  *Communalities*, ou seja comunalidades, obtém-se o resultado conforme a Figura (20) a seguir.

Variable	Communalities (Spreadsheet1 in Dados dissertação.stw)					
	From 1 Factor	From 2 Factors	From 3 Factors	From 4 Factors	From 5 Factors	Multiple R-Square
Nº de Ind. Cerâmicas	0,756199	0,983276	0,987325	0,999677	1,000000	0,818863
Tijolos maciços	0,913137	0,953325	0,956666	0,998699	1,000000	0,917883
Blocos de vedação	0,947614	0,966643	0,975893	0,986090	0,999992	0,968076
Bloco estrutural	0,968421	0,987558	0,991829	0,991911	0,999585	0,997916
Telhas	0,969833	0,989231	0,995037	0,995424	0,999483	0,998306
Tabelas	0,894257	0,938263	0,998902	0,999737	1,000000	0,889684

Figura 20 - Comunalidades para produção de cerâmica vermelha

Na análise fatorial selecionando-se a opção *Loadings* e *Summary:Factor loadings* obtêm-se os mesmos resultados obtidos anteriormente na opção *Quick*. Selecionando *Plot of loadings, 2D* obtêm-se a Figura (15), determinada anteriormente. Selecionando *Plot of loadings, 3D*, obtêm-se uma figura tri-dimensional com a localização das variáveis analisadas e de suas respectivas distâncias, conforme a Figura (21) abaixo.

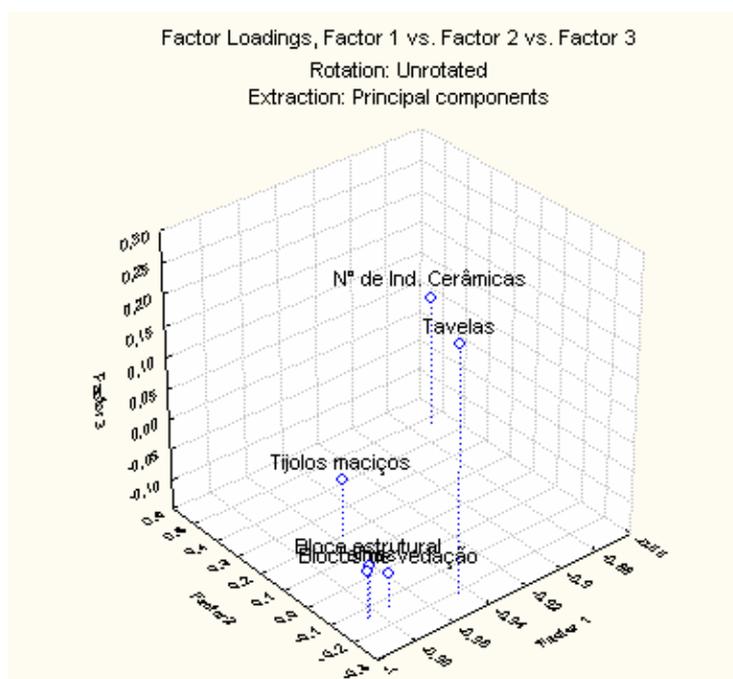


Figura 21 - Gráfico tri-dimensional dos planos fatoriais para dados de cerâmica vermelha

Selecionando-se *Scores* e *Factor score coefficients* surge a Figura (22) a seguir.

Variable	Factor Score Coefficients (Spreadsheet1 in Dados dissertação.stw)					
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
N° de Ind. Cerâmicas	-0,159794	1,292201	0,79541	1,63054	-0,69219	-0,2343
Tijolos maciços	-0,175600	0,542364	-0,76577	-3,04623	1,36623	0,3469
Blocos de vedação	-0,178663	-0,368375	-1,08656	1,69047	3,62274	3,7409
Bloco estrutural	-0,180285	-0,377235	-0,68899	-0,01389	-3,42814	9,9640
Telhas	-0,180900	-0,377338	-0,90557	0,36358	-1,49596	-13,4746
Tavelas	-0,173820	-0,573231	2,81630	-0,52309	0,64497	-0,2913

Figura 22 - Resultados dos autovetores para produção de cerâmica vermelha

Como pode ser observado, existem tantos componentes quantas variáveis existirem. Neste caso como existem seis variáveis, tem-se seis componentes principais. A partir dos autovalores, determinam-se os autovetores que formarão a combinação linear denominada de

componentes principais ou variáveis de referência. A Figura (22) contém os autovetores, derivados dos autovalores, que definem a direção dos eixos para AF e ACP.

Baseando-se nestes autovetores, pode-se escrever as seguintes combinações lineares:

$$Y_1 = -0,1598X_1 - 0,1756X_2 - 0,1787X_3 - 0,1803X_4 - 0,1809X_5 - 0,1738X_6$$

$$Y_2 = 1,2922X_1 + 0,5424X_2 - 0,3684X_3 - 0,3772X_4 - 0,3773X_5 - 0,5732X_6$$

$$Y_3 = 0,7954X_1 - 0,7658X_2 - 1,0866X_3 - 0,6890X_4 - 0,9056X_5 + 2,8163X_6$$

$$Y_4 = 1,6305X_1 - 3,0462X_2 + 1,6905X_3 - 0,0139X_4 + 0,3636X_5 - 0,5231X_6$$

$$Y_5 = -0,6922X_1 + 1,3662X_2 + 3,6227X_3 - 3,4281X_4 - 1,4960X_5 + 0,6450X_6$$

$$Y_6 = -0,2343X_1 + 0,3469X_2 + 3,7409X_3 + 9,9640X_4 - 13,4746X_5 - 0,2913X_6$$

Selecionando-se *Factor scores* surgem os resultados dos componentes principais, segundo a Figura (23) a seguir.

Case	Factor Scores (Spreadsheet1 in Dados dissertação.stw)					
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Agudo	0,25235	-0,45153	-0,674853	-0,45064	0,74276	0,87804
Cacequi	0,34258	-0,47806	-0,197148	-0,19787	-0,57101	-0,22851
Cachoeira do Sul	-0,20640	-0,00968	-0,838946	1,11366	2,89415	-3,56050
Capão do Cipó	0,36322	-0,54180	-0,107157	0,16011	-0,73156	-0,26928
Dona Francisca	0,37869	-0,58960	-0,039664	0,42860	-0,85198	-0,29985
Faxinal do Soturno	-0,04117	-1,18887	4,432633	-0,46504	1,28951	0,13352
Formigueiro	0,32300	-0,51844	-0,316237	-0,01259	-0,17395	0,18150
Itaara	0,33226	-0,44619	-0,242143	-0,37686	-0,49073	-0,20813
Jaguari	0,28793	-0,20172	-0,105575	0,34768	-0,40617	0,05695
Jari	-0,29755	3,78977	1,304275	1,41486	-1,14172	-0,51626
Júlio de Castilhos	0,07884	1,10289	0,084657	-0,54846	-0,37624	-0,23527
Mata	0,33821	-0,54021	-0,241469	0,12008	-0,35349	0,05861
Nova Palma	0,32710	-0,43026	-0,264641	-0,46636	-0,45059	-0,19794
Novo Cabrais	0,31163	-0,38246	-0,332134	-0,73484	-0,33017	-0,16737
Paraíso do Sul	0,35448	-0,66610	-0,195800	0,79602	-0,29652	0,30497
Restinga Seca	0,10097	0,25463	-0,373389	1,84160	1,34337	2,01904
Santa Maria	-4,70996	-0,66744	-0,284416	-0,00445	-0,55094	0,16227
Santiago	-0,03914	0,84368	-0,638214	0,65088	2,05484	2,29384
São Francisco de Assis	0,31972	-0,21588	0,061160	0,46177	-0,85747	-0,31530
São João do Polêsine	0,03463	1,04785	-0,466120	-2,37155	0,43203	-0,01600
São Martinho da Serra	0,36684	-0,60342	-0,108207	0,48544	-0,63739	-0,09077
São Pedro do Sul	0,24115	0,21840	0,076491	0,15486	-0,71044	-0,29202
São Sepé	-0,02467	1,01998	-0,284234	-2,72747	0,96491	0,26649
São Vicente do Sul	0,35611	-0,66273	-0,185876	0,78058	-0,32961	0,27080
Vila Nova do Sul	0,20916	0,31719	-0,062994	-0,40002	-0,46158	-0,22883

Figura 23 – Resultados das componentes principais para produção de cerâmica vermelha

Como existem seis variáveis são encontrados seis componentes mas, segundo a análise fatorial, somente a 1ª componente principal deve ser considerada.

Selecionando-se *Descriptives* surgem as opções de estatísticas descritivas. Na tela contendo *Review correlations* (correlações da revisão), *means* (médias), *standard deviations* (desvio padrão) surgem as opções *Quick*, *Advanced*, *Pairwise e Save*.

As correlações são determinadas a partir do ícone , conforme a Figura (24).

Correlations (Spreadsheet1 in Dados dissertação.stw)						
Casewise deletion of MD						
N=25						
Variable	Nº de Ind. Cerám.	Tijolos maciços	Blocos de ved.	Bloco estrut.	Telhas	Tabelas
Nº de Ind. Cerám.	1,00	0,90	0,78	0,79	0,79	0,73
Tijolos maciços	0,90	1,00	0,89	0,91	0,91	0,85
Blocos de ved.	0,78	0,89	1,00	0,97	0,98	0,92
Bloco estrut.	0,79	0,91	0,97	1,00	1,00	0,94
Telhas	0,79	0,91	0,98	1,00	1,00	0,94
Tabelas	0,73	0,85	0,92	0,94	0,94	1,00

Figura 24 - Correlações para produção de cerâmica vermelha

Analisando-se os resultados na matriz de correlação, observa-se que todos os valores são superiores a 0,7, indicando uma correlação de moderada a forte entre as variáveis. Assim, conclui-se que as variáveis estão interligadas entre si e a análise das mesmas não poderá ser realizada de maneira isolada e sim conjuntamente. Salienta-se que pode ser realizado um teste de significância das correlações, podendo ser o teste *KMO* e o teste de *Bartlett*.

5.1.3 Análise das componentes principais das variáveis da produção de cerâmica vermelha

Para determinar os componentes principais seleciona-se no *software Statistica* o módulo referente a esta análise.

Selecionam-se as variáveis de interesse e clica-se em *OK*. Aqui não serão usadas variáveis suplementares. Selecionam-se as variáveis de 1 a 6.

Surge uma tela com as opções disponíveis para a determinação dos componentes principais. Para encontrar os componentes principais selecionam-se as opções *Cases/Factor scores* e *OK*

A Figura (25) a seguir apresenta os resultados encontrados para os componentes.

Case	Factor scores, based on correlations					
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6
Agudo	0,25416	-0,44858	-0,64191	-0,50060	0,86737	0,75359
Cacequi	0,34433	-0,47385	-0,18473	-0,18919	-0,61463	-0,10682
Cachoeira do Sul	-0,25430	-0,11390	-1,02154	0,98894	2,20661	-4,00895
Capão do Cipó	0,36493	-0,53772	-0,10698	0,17648	-0,76864	-0,12069
Dona Francisca	0,38039	-0,58563	-0,04866	0,45074	-0,88415	-0,13109
Faxinal do Soturno	-0,03887	-1,17844	4,43611	-0,34214	1,31426	-0,21599
Formigueiro	0,32475	-0,51484	-0,30540	-0,02121	-0,14504	0,22625
Itaara	0,33402	-0,44191	-0,22361	-0,37203	-0,53762	-0,09988
Jaguari	0,28973	-0,19808	-0,11054	0,35544	-0,37986	0,13348
Jari	-0,29493	3,79447	1,22970	1,51759	-1,12195	-0,40616
Júlio de Castilhos	0,08095	1,10761	0,09831	-0,52692	-0,41689	-0,18164
Mata	0,33994	-0,53653	-0,23636	0,11963	-0,33944	0,13605
Nova Palma	0,32887	-0,42595	-0,24305	-0,46345	-0,49912	-0,09641
Novo Cabrais	0,31341	-0,37804	-0,30136	-0,73770	-0,38361	-0,08601
Paraíso do Sul	0,35617	-0,66309	-0,21023	0,79412	-0,21826	0,36504
Restinga Seca	0,10291	0,25516	-0,41102	1,78523	1,74740	1,73262
Santa Maria	-4,70583	-0,65717	-0,25669	-0,01145	-0,54274	0,32049
Santiago	-0,03702	0,84457	-0,63617	0,56658	2,46366	1,87573
São Francisco de Assis	0,32150	-0,21182	0,04834	0,48972	-0,88669	-0,15447
São João do Polêsine	0,03681	1,05317	-0,38746	-2,39427	0,35571	-0,08890
São Martinho da Serra	0,36854	-0,59974	-0,11677	0,49816	-0,63395	0,03683
São Pedro do Sul	0,24304	0,22267	0,07147	0,18131	-0,74292	-0,16468
São Sepé	-0,02242	1,02563	-0,19221	-2,75964	0,92211	0,08752
São Vicente do Sul	0,35780	-0,65967	-0,20017	0,78012	-0,25739	0,33729
Vila Nova do Sul	0,21109	0,32167	-0,04904	-0,38548	-0,50421	-0,14319

Figura 25 – Componentes principais para produção de cerâmica vermelha

Analisando-se os resultados, observa-se que os mesmos são iguais aos obtidos anteriormente pela AF, assim como os demais resultados tanto para as variáveis quanto para os casos.

Na interpretação dos componentes principais dois aspectos devem ser explorados, ou seja, a correlação das variáveis originais com as componentes, as quais fornecerão a contribuição mais relevante, e a realização das análises levando em consideração os indivíduos pesquisados.

Neste caso, verifica-se que pode haver uma redução do número de variáveis de seis para um componente, pois ao invés de se trabalhar com as seis variáveis passa-se a trabalhar apenas com o primeiro componente, representado pela combinação linear.

Outro gráfico que pode ser utilizado como complemento para um maior esclarecimento das análises realizadas é o círculo de correlação unitário. Selecionando-se *Variables/Plot var. factor coordinates, 2D* e os fatores a serem plotados, obtém-se o círculo de correlação (Figura 26).

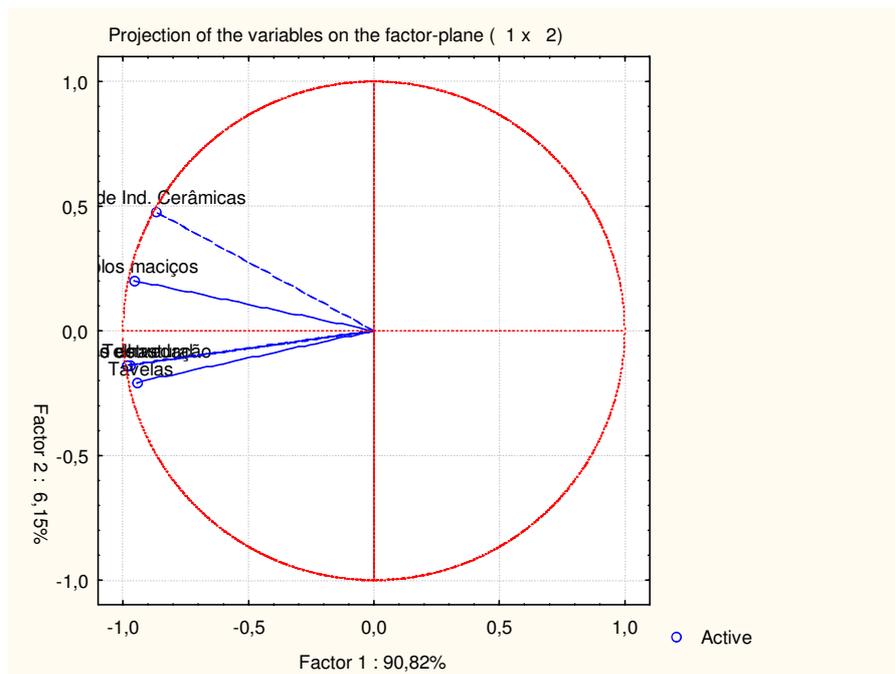


Figura 26 - Círculo de correlação unitário para produção de cerâmica vermelha

Analisando-se a figura anterior observa-se que algumas variáveis estão sobrepostas. Isso demonstra uma igual representatividade no gráfico. As variáveis que estiverem bem próximas ao círculo unitário, possuem uma maior contribuição em relação aquelas que estiverem mais afastadas. Também deve ser observado o quadrante no qual as variáveis se encontram. Na observação da disposição das variáveis, quanto menor o ângulo entre 2 variáveis, maior a correlação entre as mesmas. Quanto maior o ângulo entre 2 variáveis, menor a correlação entre as mesmas.

Fazendo-se uma sobreposição deste sobre o 1º plano fatorial, pode-se identificar visualmente quais as variáveis que estão relacionadas com os casos em estudo. Para construir o gráfico dos planos principais seleciona-se *Cases/Plot case. factor coordinates, 2D* e os fatores a serem plotados, conforme a Figura (27).

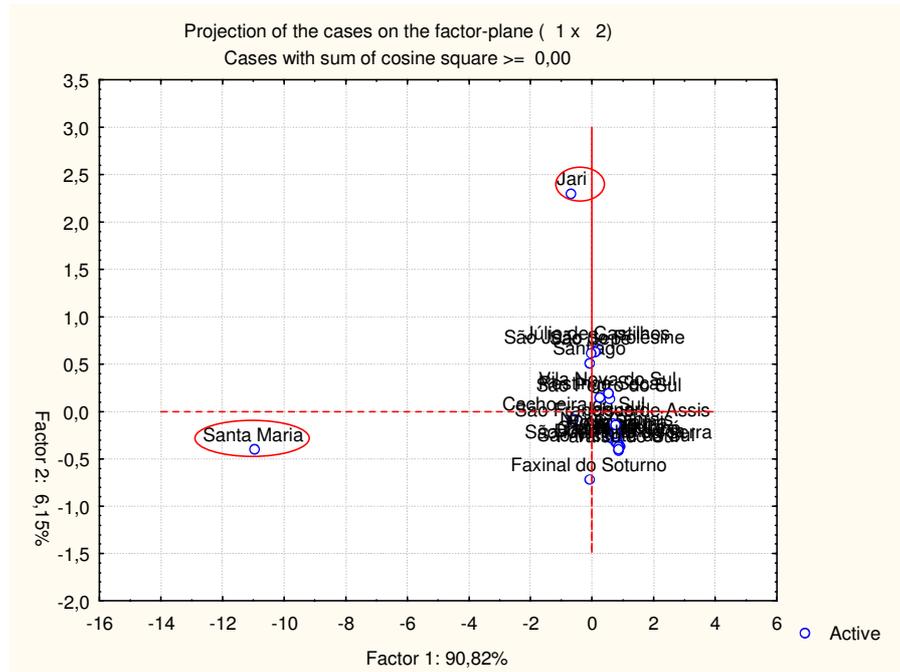


Figura 27 - Distribuição da nuvem de pontos para produção de cerâmica vermelha

Analisando-se o gráfico anterior, conclui-se que o caso que mais contribui na formação da combinação linear do fator 1 é Santa Maria. Da mesma forma pode-se concluir que o caso Jari possui a maior contribuição na combinação linear do fator 2. Os demais casos não apresentam representatividade significativa, em relação a esses dois fatores.

Outra forma de encontrar a contribuição de cada variável em relação aos fatores formados é através da matriz de correlação entre as variáveis originais e os componentes principais. Selecionando-se as variáveis surge uma tela contendo as variáveis originais e os componentes principais (Figura 28).

	Spreadsheet1 in Dados dissertação.stw									
	1	2	3	4	5	6	7	8	9	Fa
	Ind. Cerâm	blo maciç	os de ved	co estrut	Telhas	Tabelas	Factor1	Factor2	Factor3	Fa
Agudo	1	80	150	0	0	0	0,254	-0,449	-0,642	
Cacequi	1	40	0	0	0	0	0,344	-0,474	-0,185	
Cachoeira do Sul	4	134	577	6	113	23	-0,254	-0,114	-1,022	
Capão do Cipó	1	20	0	0	0	0	0,365	-0,538	-0,107	
Dona Francisca	1	5	0	0	0	0	0,380	-0,586	-0,049	
Faxinal do Soturno	2	50	130	0	0	130	-0,039	-1,178	4,436	
Formigueiro	1	40	60	0	0	0	0,325	-0,515	-0,305	
Itaara	1	50	0	0	0	0	0,334	-0,442	-0,224	
Jaguari	2	40	50	0	0	0	0,290	-0,198	-0,111	
Jari	13	213	5	0	0	0	-0,295	3,794	1,230	
Júlio de Castilhos	5	147	0	0	0	0	0,081	1,108	0,098	
Mata	1	30	45	0	0	0	0,340	-0,537	-0,236	
Nova Palma	1	55	0	0	0	0	0,329	-0,426	-0,243	
Novo Cabrais	1	70	0	0	0	0	0,313	-0,378	-0,301	
Paraíso do Sul	1	0	90	0	0	0	0,356	-0,663	-0,210	
Restinga Seca	4	52	350	0	0	0	0,103	0,255	-0,411	
Santa Maria	19	850	2724	108	1054	370	-4,706	-0,657	-0,257	
Santiago	5	144	371	0	0	0	-0,037	0,845	-0,636	
São Francisco de As	2	25	0	0	0	0	0,321	-0,212	0,048	
São João do Polésin	4	227	0	0	0	0	0,037	1,053	-0,387	
São Martinho da Ser	1	7	30	0	0	0	0,369	-0,600	-0,117	

Figura 28 – Variáveis originais e componentes principais (*Factor scores*) para produção de cerâmica vermelha

Para construir a matriz de correlação seleciona-se *Statistics/Basic Statistics/Tables* na barra principal do *Statistica*.

Surge uma caixa contendo as estatísticas descritivas disponíveis. Seleciona-se *Correlation matrices* e *OK*. Surge uma caixa de diálogo contendo as opções para a seleção das variáveis. Seleciona-se a opção *Two lists (rect. matrix)*. Surge uma tela contendo as opções para matriz de correlação.

A seguir clica-se em *Summary: Correlation matrix* para obter-se a matriz de correlação (Figura 29).

Correlations (Spreadsheet5)						
Marked correlations are significant at $p < ,05000$						
N=25 (Casewise deletion of missing data)						
Variable	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6
Nº de Ind. Cerâmicas	-0,87	0,48	0,06	0,11	-0,02	-0,00
Tijolos maciços	-0,96	0,20	-0,06	-0,21	0,04	0,00
Blocos de vedação	-0,97	-0,14	-0,10	0,10	0,12	0,00
Bloco estrutural	-0,98	-0,14	-0,07	0,01	-0,09	0,02
Telhas	-0,98	-0,14	-0,08	0,02	-0,06	-0,02
Tabelas	-0,95	-0,21	0,25	-0,03	0,02	-0,00

Figura 29 – Matriz de correlação entre variáveis originais e componentes principais para produção de cerâmica vermelha

Cabe ressaltar que as correlações significativas aparecem em vermelho e ocorrem quando o valor de p for menor do que 0,05. Os valores encontrados representam o quanto cada variável contribui para cada fator, sendo os maiores valores absolutos os que melhor explicam o fator. Ou seja, para o fator 1: -0,87; -0,96; -0,97; -0,98; -0,98 e -0,95. Para o fator 2: 0,48.

5.2 Análise estatística dos dados referentes às dimensões médias e características geométricas dos tijolos maciços

Para os dados referentes às dimensões médias e características geométricas dos tijolos maciços contidos na Figura (30) novamente serão aplicadas as técnicas de análise multivariada: análise de agrupamento, análise fatorial e análise de componentes principais.

	1 Olaria	2 L	3 H	4 C	5 Desvio Esquadro	6 Planeza Faces
Agudo	1.1	107,92	54,58	231,75	3,00	5,00
Cachoeira de	3.1	117,29	56,83	230,38	2,00	1,00
Cachoeira de	3.3	121,96	62,54	230,42	2,00	2,00
Capão do Ci	4.1	98,08	59,33	219,54	2,00	3,00
Dona Francis	6.1	114,83	59,67	237,46	2,00	2,00
Formigueiro	8.1	104,96	53,75	228,92	2,00	4,00
Itaara	9.1	107,04	52,29	222,38	3,00	4,00
Jaguari	11.2	105,92	54,79	222,92	1,00	2,00
Jari	12.1	107,54	55,5	216,75	7,00	3,00
Jari	12.2	108,92	54,58	221,46	3,00	2,00
Júlio de Cas	13.1	108,42	58,46	245,21	2,00	3,00
Júlio de Cas	13.8	119,79	54,67	226,04	4,00	4,00
Mata	14.1	104,21	50,75	226,17	4,00	4,00
Nova Palma	16.1	107,92	52,5	215,83	2,00	2,00
Novo Cabrais	17.1	112,75	56,46	232,42	2,00	4,00
Restinga Se	21.3	117,25	56,83	247,08	2,00	3,00
Santiago	22.1L	106,83	58,29	222,08	2,00	2,00
Santiago	22.1R	105,67	55,63	218,04	1,00	1,00
Santiago	22.2	110,13	58,00	226,67	3,00	3,00
São Francis	23.2	103,92	51,04	215,58	2,00	4,00
São Francis	23.3	113,83	56,25	232,33	3,00	4,00
São João do	24.2	119,67	60,79	245,42	2,00	4,00
São João do	24.4	106,5	53,75	232,21	3,00	2,00
São Martinh	25.1	130,21	53,54	210,83	2,00	1,00
São Pedro d	26.3	113,58	57,21	228,50	2,00	3,00
São Sepé	27.2	104,96	55,67	223,58	2,00	2,00
São Sepé	27.4	106,46	55,08	229,58	4,00	2,00
Vila Nova do	33.2	103,25	53,54	214,29	2,00	2,00

Figura 30 - Dados das dimensões médias e características geométricas dos tijolos maciços

5.2.1 Análise de agrupamento para as dimensões médias e características geométricas dos tijolos maciços

Para a realização da análise de agrupamento foram realizadas as mesmas etapas descritas anteriormente.

Selecionando-se todas as variáveis, exceto a olaria, e a opção



Vertical icicle plot

obtem-se o dendograma a seguir (Figura 31)

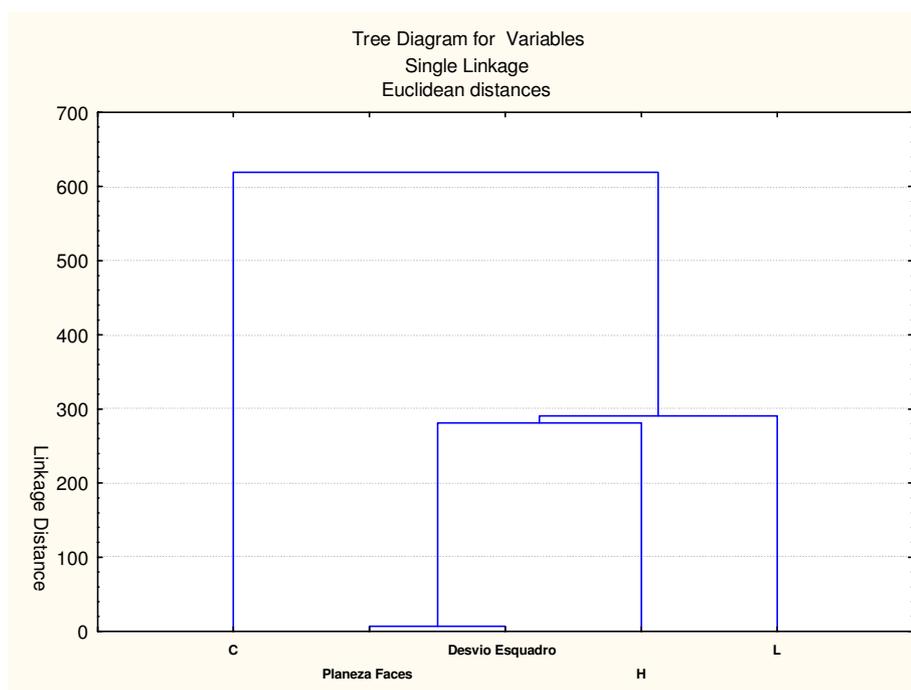


Figura 31 - Dendograma vertical dos dados de tijolos maciços

No dendograma elaborado verifica-se a formação de 2 grupos: um grupo composto pelas variáveis L, H, Desvio esquadro e Planeza faces e outro grupo composto pela variável C, da direita para a esquerda.

Selecionado-se a opção



Amalgamation schedule

obtem-se a Figura (32):

Amalgamation Schedule					
Single Linkage					
Euclidean distances					
linkage distance	Obj. No. 1	Obj. No. 2	Obj. No. 3	Obj. No. 4	Obj. No. 5
7,141428	Desvio Esquadro	Planeza Faces			
280,9984	H	Desvio Esquadro	Planeza Faces		
290,6735	L	H	Desvio Esquadro	Planeza Faces	
619,0370	L	H	Desvio Esquadro	Planeza Faces	C

Figura 32 - Ligação entre as variáveis analisadas para tijolos maciços

Selecionado-se a opção  **Graph of amalgamation schedule** obtém-se a Figura (33):

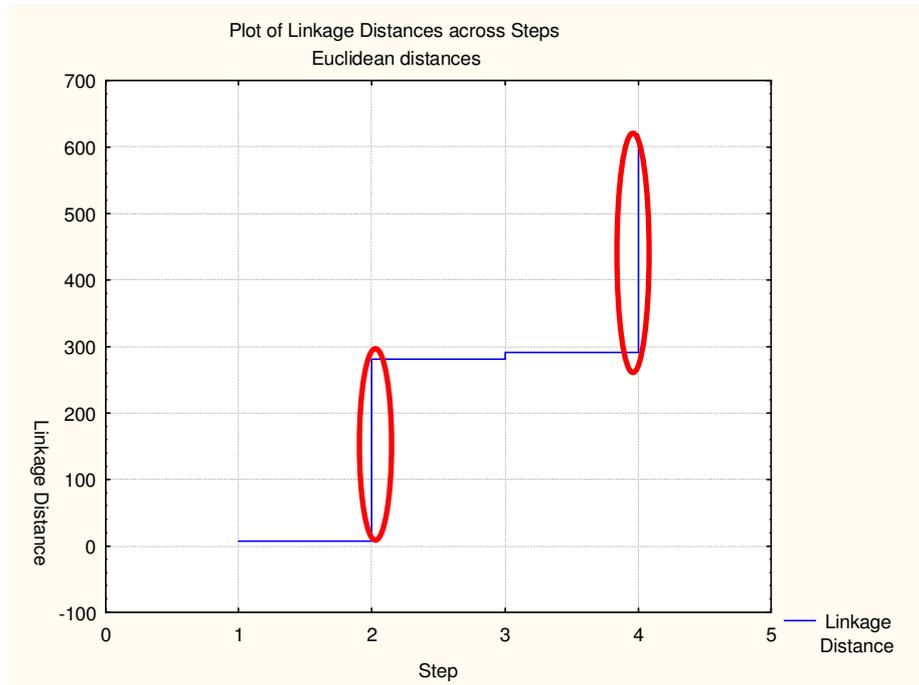


Figura 33 - Ligação das variáveis analisadas para tijolos maciços

Analisando-se o gráfico anterior verifica-se que os maiores saltos ocorrem entre as distâncias 7,141428 e 280,9984 e 290,6735 e 619,0370. Entre as distâncias 300 e 400 pode ser feito um corte no gráfico onde observa-se a formação dos grupos distintos, conforme o histograma.

Selecionado-se a opção  **Distance matrix** obtém-se a matriz de distâncias entre as variáveis (Figura 34).

Variable	Euclidean distances				
	L	H	C	Desvio Esquadro	Planeza Faces
L	0	291	619	572	570
H	291	0	907	282	281
C	619	907	0	1188	1187
Desvio Esquadro	572	282	1188	0	7
Planeza Faces	570	281	1187	7	0

Figura 34 - Matriz de distâncias para variáveis de tijolos maciços

Selecionado-se a opção  obtem-se os valores de média e desvio padrão das variáveis, para um maior conhecimento e compreensão dos resultados obtidos, conforme a Figura (35).

Variable	Means and Standard Deviations	
	Mean	Std.Dev.
L	110,3604	6,959670
H	55,7971	2,832009
C	226,9229	9,333445
Desvio Esquadro	2,5357	1,170063
Planeza Faces	2,7857	1,100745

Figura 35 – Estatística descritiva (média e desvio padrão) dos dados para tijolos maciços

Os resultados anteriores demonstram que a maior parte das variáveis apresenta um desvio padrão baixo, o que significa uma pequena variação entre os resultados.

5.2.2 Análise fatorial para as dimensões médias e características geométricas dos tijolos maciços

Para a realização da análise fatorial , selecionam-se as variáveis de interesse e a opção *Raw data*.

Selecionando-se *Quick* e clicando-se em *OK* obtem-se a tela onde observa-se que existem as opções para a realização da análise fatorial. Selecionando-se *Eigenvalues* (autovalores) obtem-se a Figura (36):

Eigenvalues				
Extraction: Principal components				
Value	Eigenvalue	% Total variance	Cumulative Eigenvalue	Cumulative %
1	1,832635	36,65270	1,832635	36,6527
2	1,378234	27,56469	3,210870	64,2174
3	0,879235	17,58470	4,090105	81,8021
4	0,597564	11,95128	4,687669	93,7534
5	0,312331	6,24663	5,000000	100,0000

Figura 36 - Autovalor, % da variância total, autovalor acumulado e % acumulada, para tijolos maciços

Considerando-se as 5 variáveis e analisando-se os resultados obtidos verifica-se que o 1º e o 2º autovalor são maiores do que 1 e explicam 64,22% da variância. Os demais

autovalores são menores do que 1. Portanto, os dados serão resumidos pela 1ª e 2ª componente principal.

Clicando-se em  ou em  obtém-se a Figura (37) a seguir.

Variable	Factor Loadings				
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
L	-0,636275	-0,059513	0,637931	0,426159	0,055182
H	-0,843137	-0,004661	-0,030596	-0,451541	0,290297
C	-0,731582	0,525077	-0,217121	-0,010172	-0,376612
Desvio Esquadro	0,405689	0,557592	0,614760	-0,374669	-0,078746
Planeza Faces	0,130819	0,887725	-0,215109	0,267550	0,277446
Expl.Var	1,832635	1,378234	0,879235	0,597564	0,312331
Prp.Totl	0,366527	0,275647	0,175847	0,119513	0,062466

Figura 37 - Composição dos fatores para as variáveis de tijolos maciços

Na figura anterior verifica-se o resultado dos *Factor loadings* sendo que os valores destacados possuem uma significância maior do que 0,7, o que poderá ser alterado conforme descrito anteriormente. Observa-se que os valores de H e C para o fator 1 e o valor de Planeza faces para o fator 2 são significativos. No caso dos demais fatores os valores não são significativos.

Selecionando-se a opção , obtém-se a Figura (38) a seguir.

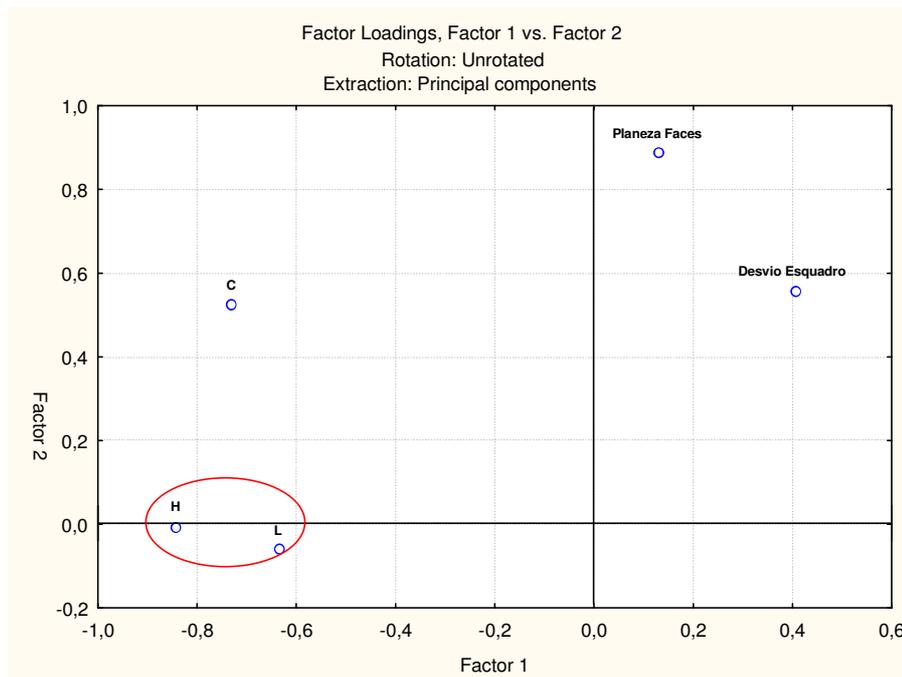


Figura 38 - Planos fatoriais para tijolos maciços

Analisando-se a figura anterior, pode-se observar que as variáveis se agrupam por similaridades de explicação, ou seja, estão agrupadas por fatores. As variáveis que explicam ou representam melhor o fator 1 são aquelas que estiverem localizadas mais distantemente da origem, em relação ao eixo x. As variáveis que estiverem localizadas próximas à origem do plano fatorial possuem baixa representatividade. Neste caso, num prolongamento do eixo x até a origem dos eixos, estas variáveis são H, C e L. Do mesmo modo, a variável que melhor representa o fator 2 é Planeza faces.

Como descrito anteriormente, as distâncias de cada uma das variáveis analisadas são encontradas traçando-se segmentos de reta perpendiculares ao eixo considerado e determina-se aquela ou aquelas variáveis que estiverem localizadas mais distantemente da origem, à partir da perpendicular traçada, sendo que essas terão maior influência sobre o fator analisado. Isso pode ser observado nas Figuras (39) e (40) seguintes.

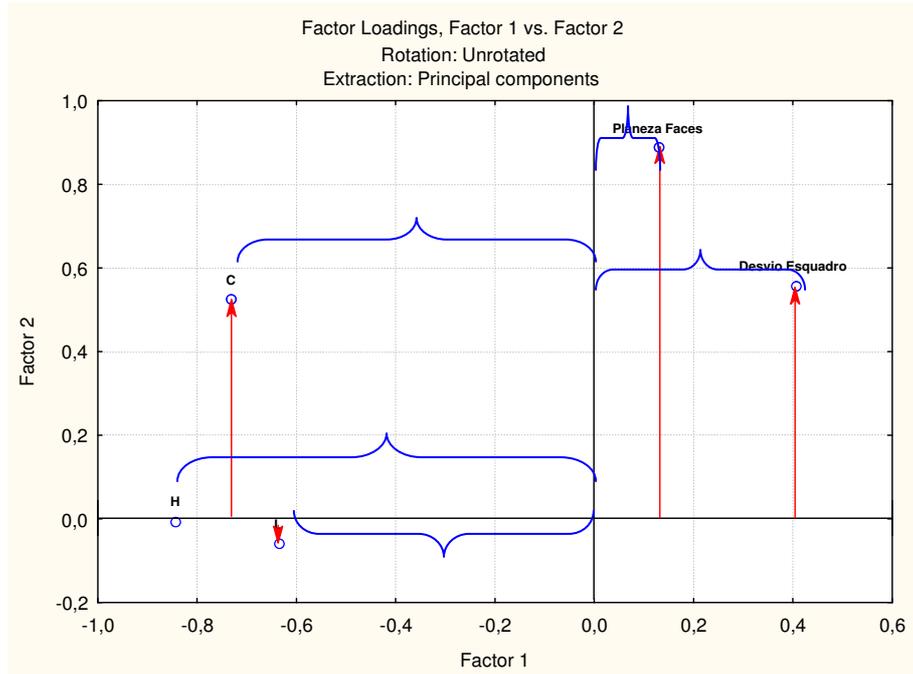


Figura 39 - Planos fatoriais contendo perpendiculares em relação ao fator 1, para tijolos maciços

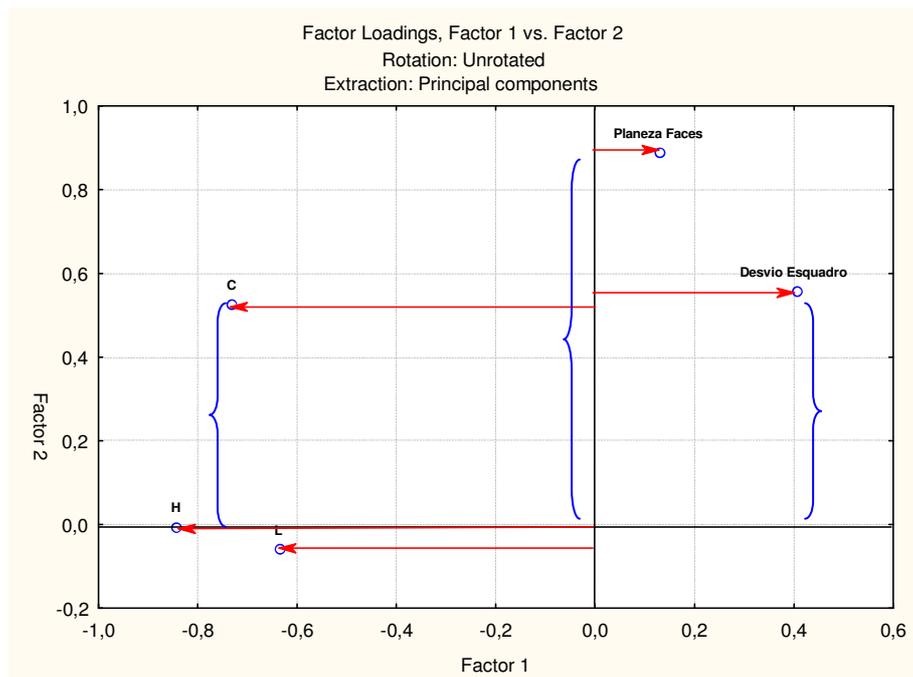


Figura 40 - Planos fatoriais contendo perpendiculares em relação ao fator 2, para tijolos maciços

A visualização de ambos os gráficos anteriores confirma o que já foi descrito.

Novamente quando a identificação das variáveis que representam um fator tornar-se difícil realizam-se as rotações pois assim ocorre uma mais perfeita visualização da disposição dos pontos. Como já citado o tipo de rotação mais empregado é a *Varimax normalized* que mantém os eixos ortogonais. Para uma melhor visualização dos fatores seleciona-se *Plot of loadings, 2D*, conforme a Figura (41) a seguir.

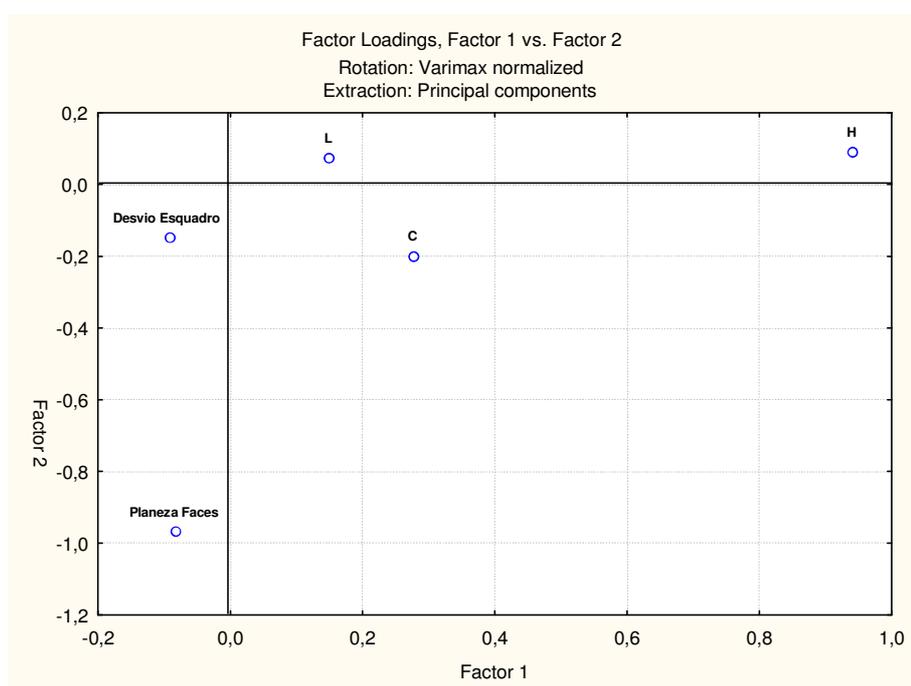


Figura 41 - Planos fatoriais, com rotação *varimax normalized*, para os dados de tijolos maciços

Analisando-se o gráfico anterior, novamente as variáveis de maior representatividade são aquelas localizadas mais distantemente da origem do plano fatorial. Aquelas que estiverem localizadas próximas à origem do plano fatorial, possuem baixa representatividade.

Abaixo são apresentados os resultados dos *Factor Loadings* rotacionados (Figura 42).

Variable	Factor Loadings (Varimax normalized) Extraction: Principal components (Marked loadings are > ,700000)				
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
L	0,148226	0,075116	-0,978756	-0,016633	0,118938
H	0,939884	0,091241	-0,164866	-0,103298	0,265411
C	0,275887	-0,197689	-0,136001	-0,058603	0,928910
Desvio Esquadro	-0,091324	-0,146014	0,016518	0,983574	-0,051467
Planeza Faces	-0,081987	-0,965742	0,079038	0,155424	0,173831
Expl.Var	0,996528	1,007026	1,010161	1,005956	0,980330
Prp.Totl	0,199306	0,201405	0,202032	0,201191	0,196066

Figura 42 - Matriz fatorial rotacionada para tijolos maciços

Selecionando-se *Scree plot*



obtem-se o resultado conforme a

Figura (43) a seguir.

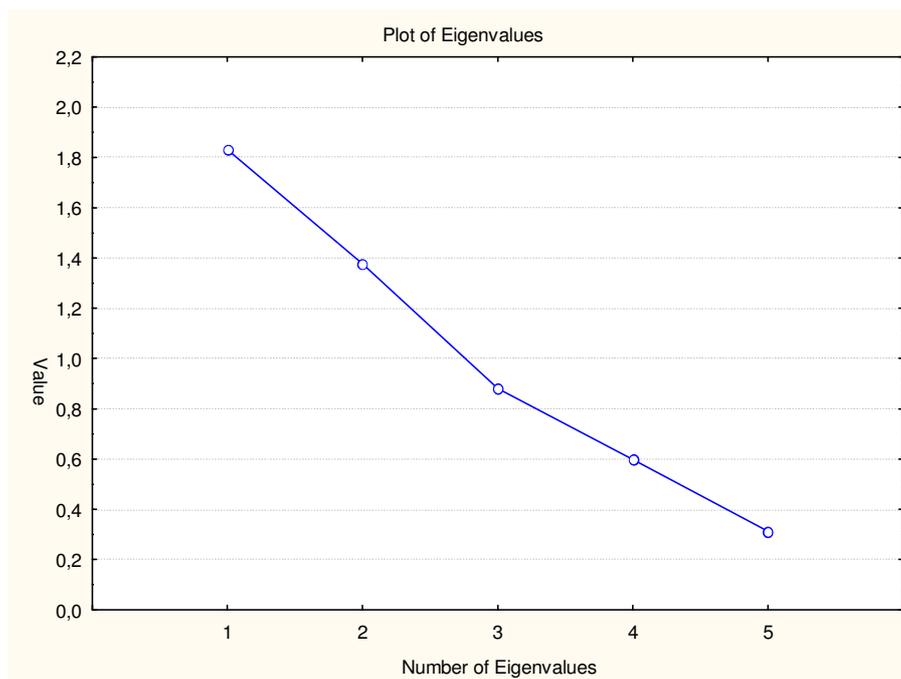


Figura 43 - Autovalores para tijolos maciços

Examinando-se o gráfico anterior, não se observa uma queda brusca entre os fatores e, analisando-se os autovalores superiores a 1, observa-se que o 1º e o 2º fator deverão ser considerados.

Selecionando-se *Communalities*  *Communalities*, comunalidades, obtém-se o resultado a seguir (Figura 44).

Communalities						
Extraction: Principal components						
Rotation: Unrotated						
Variable	From 1 Factor	From 2 Factors	From 3 Factors	From 4 Factors	From 5 Factors	Multiple R-Square
L	0,404846	0,408388	0,815344	0,996955	1,000000	0,158706
H	0,710880	0,710902	0,711838	0,915727	1,000000	0,393114
C	0,535212	0,810918	0,858060	0,858163	1,000000	0,450473
Desvio Esquadro	0,164583	0,475492	0,853422	0,993799	1,000000	0,136626
Planeza Faces	0,017114	0,805169	0,851441	0,923024	1,000000	0,319438

Figura 44 - Comunalidades para tijolos maciços

Cabe lembrar que as comunalidades são índices que expressam quanto da variância é explicada pela AF, ou seja, quanto mais próximo de 1, melhor será o ajuste da AF.

Aqui selecionando-se *Summary:Factor loadings* obtém-se os mesmos resultados obtidos anteriormente na opção *Quick e* selecionando-se *Plot of loadings, 2D* obtém-se a

Figura (41), determinada anteriormente. Na opção *Plot of loadings, 3D*, obtém-se uma figura tri-dimensional, conforme a Figura (45) abaixo.

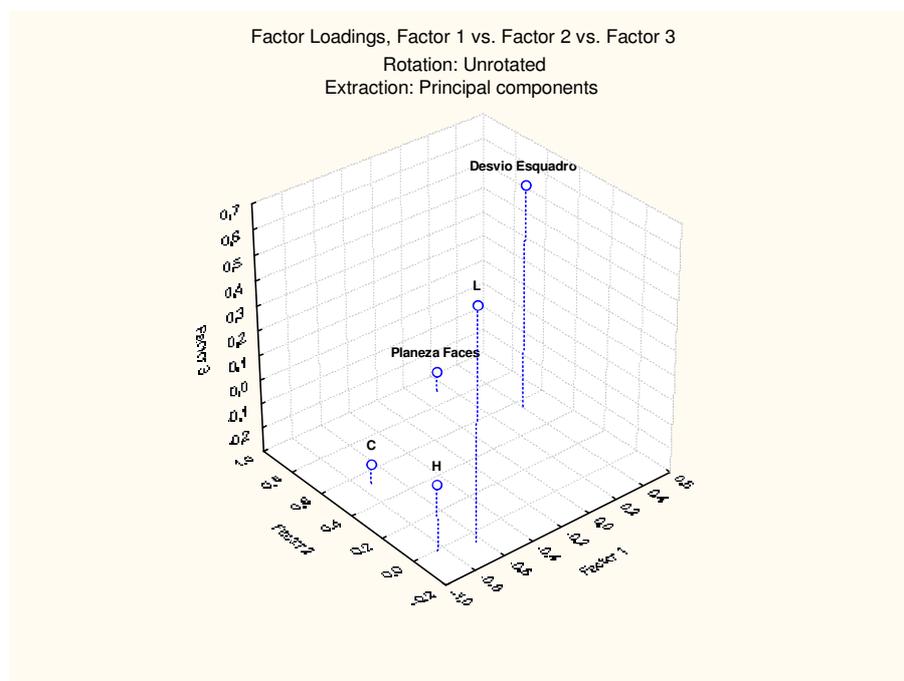


Figura 45 – Gráfico tri-dimensional dos planos fatoriais para tijolos maciços

Selecionando-se *Factor score coefficients* surgem os autovetores segundo a Figura (46) a seguir.

Factor Score Coefficients					
Rotation: Unrotated					
Extraction: Principal components					
Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
L	-0,347191	-0,043181	0,725552	0,713160	0,17668
H	-0,460068	-0,003382	-0,034798	-0,755637	0,92945
C	-0,399197	0,380978	-0,246943	-0,017022	-1,20581
Desvio Esquadro	0,221369	0,404570	0,699199	-0,626993	-0,25212
Planeza Faces	0,071383	0,644103	-0,244654	0,447734	0,88831

Figura 46 - Resultados dos autovetores para tijolos maciços

Como pode ser observado, existem tantos componentes quantas variáveis existirem. Neste caso como existem cinco variáveis, tem-se seis componentes principais. A partir dos autovalores, determinam-se os autovetores que formarão a combinação linear denominada de componentes principais ou variáveis de referência. A tabela anterior (Figura 46) contém os autovetores, derivados dos autovalores, que definem a direção dos eixos para AF e ACP.

Baseando-se nestes autovetores, pode-se escrever as seguintes combinações lineares:

$$Y_1 = -0,3472X_1 - 0,4601X_2 - 0,3992X_3 + 0,2214X_4 + 0,0714X_5$$

$$Y_2 = -0,0432X_1 - 0,0034X_2 + 0,3810X_3 + 0,4046X_4 + 0,6441X_5$$

$$Y_3 = 0,7255X_1 - 0,0348X_2 - 0,2469X_3 + 0,6992X_4 - 0,2446X_5$$

$$Y_4 = 0,7132X_1 - 0,7556X_2 - 0,0170X_3 - 0,6270X_4 + 0,4477X_5$$

$$Y_5 = 0,1767X_1 + 0,9294X_2 - 1,2058X_3 - 0,2521X_4 + 0,8883X_5$$

Selecionando-se *Factor scores* surgem os resultados dos componentes principais, segundo a Figura (47) a seguir.

Case	Factor Scores				
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Agudo	0,34395	1,66980	-0,58083	0,71879	0,60211
Cachoeira do Sul	-0,87900	-1,13332	0,69607	-0,01006	-1,25713
Cachoeira do Sul	-1,97644	-0,58233	0,88944	-0,64839	1,53725
Capão do Cipó	0,26651	-0,28929	-1,49503	-1,81229	2,09015
Dona Francisca	-1,45562	-0,24730	-0,00487	-0,62607	-0,49518
Formigueiro	0,49344	0,64272	-1,17965	0,77121	0,02865
Itaara	1,09577	0,71037	-0,17426	0,84997	0,23173
Jaguari	0,21433	-1,12546	-1,08665	0,32538	-0,22903
Jari	1,48208	1,27154	2,59993	-2,49523	0,35636
Jari	0,53962	-0,51189	0,46245	-0,38024	-0,46413
Júlio de Castilhos	-1,20590	0,69541	-1,08556	-0,56743	-1,24926
Júlio de Castilhos	0,10574	1,12359	1,62643	0,97890	0,64818
Mata	1,51422	1,23024	0,04693	0,42811	-1,05065
Nova Palma	0,97901	-1,07877	-0,06485	0,61841	-0,22933
Novo Cabrais	-0,48512	0,73401	-0,49344	0,83999	0,66364
Restinga Seca	-1,46158	0,71890	-0,19447	0,76889	-1,80165
Santiago	-0,17453	-0,82381	-0,41499	-1,04957	0,83580
Santiago	0,23421	-1,90926	-0,77166	-0,32222	-0,13624
Santiago	-0,23432	0,27434	0,18647	-0,77152	0,82293
São Francisco de Assis	1,55613	0,10789	-0,90183	1,41205	0,83626
São Francisco de Assis	-0,31183	1,06966	0,22168	0,47099	0,41829
São João do Polesine	-2,08977	1,21655	-0,16918	0,37005	0,58090
São João do Polesine	0,33540	-0,05708	-0,06406	-0,42636	-2,18678
São Martinho da Serra	-0,15290	-2,00756	2,60067	2,22735	0,51680
São Pedro do Sul	-0,54555	-0,01719	-0,09015	0,32532	0,63029
São Sepé	0,28023	-0,74785	-0,61744	-0,54486	-0,26534
São Sepé	0,42301	0,18000	0,58259	-1,31639	-1,62700
Vila Nova do Sul	1,10890	-1,11390	-0,52374	-0,13481	0,19239

Figura 47 – Resultados das componentes principais para tijolos maciços

Como existem cinco variáveis são encontrados cinco componentes mas, segundo a análise fatorial, somente a 1ª e a 2ª componente principal devem ser consideradas.

Selecionando-se *Descriptives* surgem as opções de estatísticas descritivas, correlações da revisão (*Review correlations*), médias (*means*), desvio padrão (*standard deviations*) e regressão múltipla (*Multiple regression*).

As correlações são determinadas a partir do ícone , segundo a Figura (48).

		Correlations				
		Casewise deletion of MD				
		N=28				
Variable	L	H	C	Desvio Esquadro	Planeza Faces	
L	1,00	0,34	0,27	-0,06	-0,14	
H	0,34	1,00	0,52	-0,22	-0,15	
C	0,27	0,52	1,00	-0,10	0,31	
Desvio Esquadro	-0,06	-0,22	-0,10	1,00	0,29	
Planeza Faces	-0,14	-0,15	0,31	0,29	1,00	

Figura 48 - Correlações para tijolos maciços

Analisando-se os resultados na matriz de correlação, observa-se que todos os valores são inferiores a 0,7 o que indica uma correlação fraca entre as variáveis. Assim, poder-se-ia afirmar que neste caso a realização de uma análise fatorial seria insatisfatória. Mas, como forma de exemplificação a técnica será aplicada.

5.2.3 Análise de componentes principais para as dimensões médias e características geométricas dos tijolos maciços

Determinam-se os componentes principais selecionando-se no *software Statistica* o módulo referente a esta análise, como descrito anteriormente.

Selecionam-se as variáveis de 2 a 5 e clica-se em *OK*. Aqui novamente não serão usadas variáveis suplementares. Na figura seguinte, onde aparecem as variáveis selecionadas, clica-se novamente em *OK*. Surge uma tela contendo as opções disponíveis para a determinação dos componentes principais. Para encontrar os componentes principais selecionam-se as opções *Cases/Factor scores* e *OK*

A Figura (49) a seguir apresenta os resultados encontrados para os componentes.

Case	Factor scores, based on correlations				
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5
Agudo	0,34395	1,66980	-0,58083	0,71879	0,60211
Cachoeira do Sul	-0,87900	-1,13332	0,69607	-0,01006	-1,25713
Cachoeira do Sul	-1,97644	-0,58233	0,88944	-0,64839	1,53725
Capão do Cipó	0,26651	-0,28929	-1,49503	-1,81229	2,09015
Dona Francisca	-1,45562	-0,24730	-0,00487	-0,62607	-0,49518
Formigueiro	0,49344	0,64272	-1,17965	0,77121	0,02865
Itaara	1,09577	0,71037	-0,17426	0,84997	0,23173
Jaguari	0,21433	-1,12546	-1,08665	0,32538	-0,22903
Jari	1,48208	1,27154	2,59993	-2,49523	0,35636
Jari	0,53962	-0,51189	0,46245	-0,38024	-0,46413
Júlio de Castilhos	-1,20590	0,69541	-1,08556	-0,56743	-1,24926
Júlio de Castilhos	0,10574	1,12359	1,62643	0,97890	0,64818
Mata	1,51422	1,23024	0,04693	0,42811	-1,05065
Nova Palma	0,97901	-1,07877	-0,06485	0,61841	-0,22933
Novo Cabrais	-0,48512	0,73401	-0,49344	0,83999	0,66364
Restinga Seca	-1,46158	0,71890	-0,19447	0,76889	-1,80165
Santiago	-0,17453	-0,82381	-0,41499	-1,04957	0,83580
Santiago	0,23421	-1,90926	-0,77166	-0,32222	-0,13624
Santiago	-0,23432	0,27434	0,18647	-0,77152	0,82293
São Francisco de Assis	1,55613	0,10789	-0,90183	1,41205	0,83626
São Francisco de Assis	-0,31183	1,06966	0,22168	0,47099	0,41829
São João do Polesine	-2,08977	1,21655	-0,16918	0,37005	0,58090
São João do Polesine	0,33540	-0,05708	-0,06406	-0,42636	-2,18678
São Martinho da Serra	-0,15290	-2,00756	2,60067	2,22735	0,51680
São Pedro do Sul	-0,54555	-0,01719	-0,09015	0,32532	0,63029
São Sepé	0,28023	-0,74785	-0,61744	-0,54486	-0,26534
São Sepé	0,42301	0,18000	0,58259	-1,31639	-1,62700
Vila Nova do Sul	1,10890	-1,11390	-0,52374	-0,13481	0,19239

Figura 49 – Componentes principais para tijolos maciços

Analisando-se os resultados, observa-se que estes são iguais aos obtidos anteriormente pela AF, assim como os demais resultados tanto para as variáveis quanto para os casos.

Na interpretação dos componentes principais dois aspectos devem ser explorados, ou seja, a correlação das variáveis originais com as componentes, as quais fornecerão a contribuição mais relevante, e a realização das análises levando em consideração os indivíduos pesquisados.

Neste caso, verifica-se que pode haver uma redução do número de variáveis de cinco para dois componentes, pois ao invés de se trabalhar com as cinco variáveis passa-se a trabalhar com o primeiro e segundo componente, representados pelas combinações lineares.

Outro gráfico que pode ser utilizado como complemento para um maior esclarecimento das análises realizadas é o círculo de correlação unitário.

Selecionando-se *Variables/Plot var. factor coordinates, 2D* e os fatores a serem plotados, obtém-se o círculo de correlação (Figura 50).

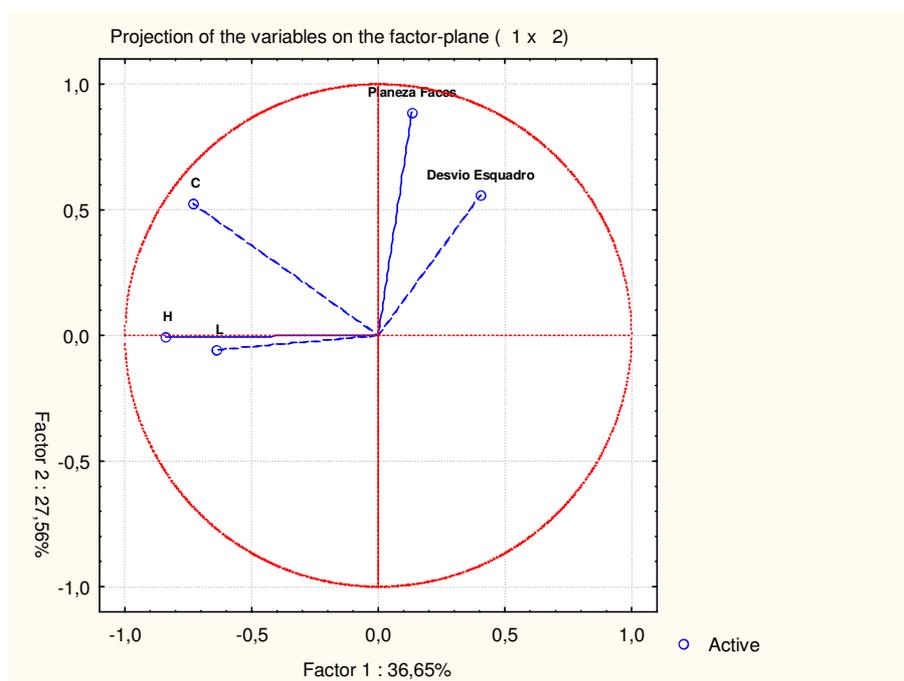


Figura 50 - Círculo de correlação unitário para tijolos maciços

Analisando-se a figura anterior observa-se que as variáveis que estiverem mais próximas ao círculo unitário, possuem uma maior contribuição em relação aquelas que estiverem mais afastadas. As variáveis *Planeza faces* e *Desvio esquadro* encontram-se no 1º quadrante. A variável *C* encontra-se no 2º quadrante, a variável *H* encontra-se sobre o eixo e a variável *L* encontra-se no 3º quadrante. Também deve ser observado o quadrante no qual as variáveis se encontram e o ângulo formado entre 2 variáveis, demonstrando maior ou menor correlação entre as mesmas.

Fazendo-se uma sobreposição deste sobre o 1º plano fatorial, pode-se identificar visualmente quais as variáveis que estão relacionadas com os casos em estudo. Para construir o gráfico dos planos principais (Figuras 51 e 52) seleciona-se *Cases/Plot case. factor coordinates, 2D* e os fatores a serem plotados.

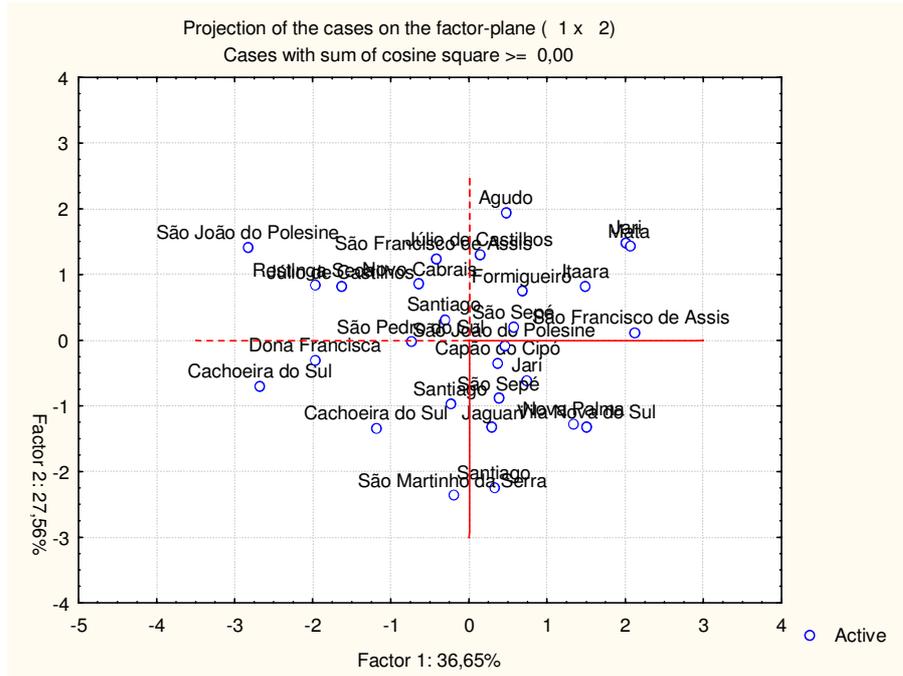


Figura 51 - Distribuição da nuvem de pontos por municípios, para tijolos maciços

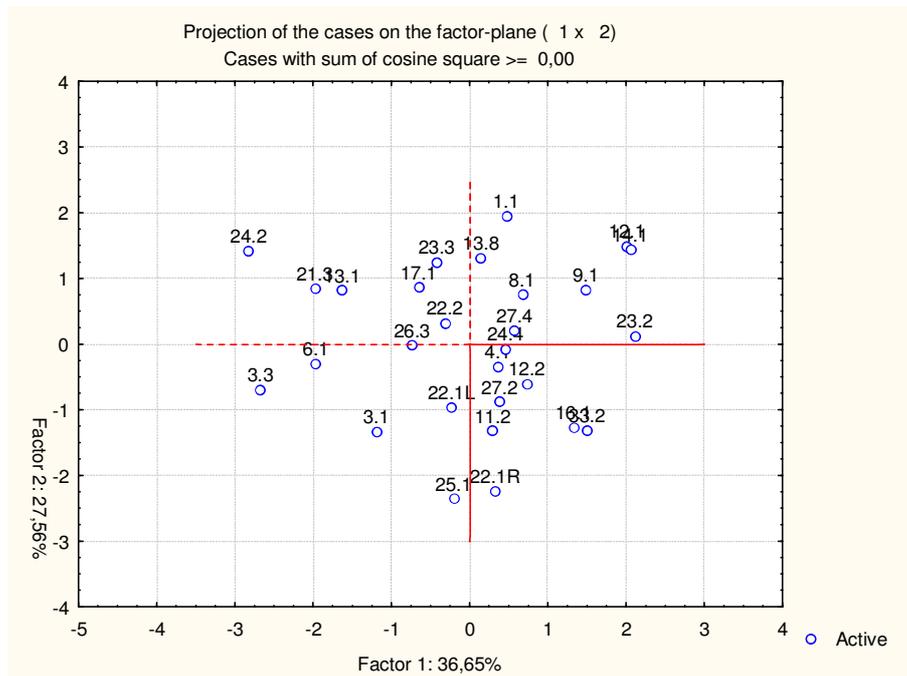


Figura 52 - Distribuição da nuvem de pontos por olarias, para tijolos maciços

Analisando-se o gráfico anterior, conclui-se que diversos casos contribuem na formação da combinação linear dos fatores 1 e 2, tornando-se muito difícil a identificação dos casos mais representativos podendo-se citar: Desvio esquadro: Mata e Jari, Planeza faces: Agudo, C: São João do Polêsine, L: Cachoeira do Sul

Observa-se uma distribuição bastante uniforme das olarias, ocorrendo uma maior concentração em torno da origem. Esses casos não apresentam representatividade significativa, em relação a esses dois fatores. Também observa-se que ocorrem algumas sobreposições das mesmas.

Outra forma de encontrar a contribuição de cada variável em relação aos fatores formados é através da matriz de correlação entre as variáveis originais e os componentes principais. Para determinar essa correlação seleciona-se *Cases/Save case statistics* e a opção *Factor Scores*. Clica-se em *OK*. A seguir, selecionam-se as variáveis e *OK*. Surge uma tela contendo as variáveis originais e os componentes principais.

Para construir a matriz de correlação seleciona-se *Statistics/Basic Statistics/Tables* na barra principal do *Statistica*. Surge uma caixa contendo as estatísticas descritivas disponíveis. Seleciona-se *Correlation matrices* e *OK*. Aparece uma caixa de diálogo contendo as opções para a seleção das variáveis.

Seleciona-se a opção *Two lists (rect. matrix)*. Surge uma caixa contendo as variáveis originais e os componentes principais.

A seguir clica-se em *Summary: Correlation matrix* para obter-se a matriz de correlação (Figura 53).

Correlations (Spreadsheet22)					
Marked correlations are significant at p < ,05000					
N=28 (Casewise deletion of missing data)					
Variable	Factor1	Factor2	Factor3	Factor4	Factor5
L	-0,64	-0,06	0,64	0,43	0,06
H	-0,84	-0,00	-0,03	-0,45	0,29
C	-0,73	0,53	-0,22	-0,01	-0,38
Desvio Esquadro	0,41	0,56	0,61	-0,37	-0,08
Planeza Faces	0,13	0,89	-0,22	0,27	0,28

Figura 53 - Matriz de correlação entre variáveis originais e componentes principais, para tijolos maciços

Ressalta-se que as correlações significativas aparecem em vermelho e ocorrem quando o valor de p for menor do que 0,05. Os valores encontrados representam o quanto cada variável contribui para cada fator, sendo os maiores valores absolutos os que melhor explicam o fator.

Para o fator 1: -0,64; -0,84; -0,73 e 0,41.

Para o fator 2: 0,53; 0,56 e 0,89.

Para o fator 3: 0,64 e 0,61.

Para o fator 4: 0,43; -0,45 e -0,37.

Para o fator 5: -0,38.

Comparando-se os resultados aqui obtidos com os *factor loadings* obtidos anteriormente, observa-se que os valores são iguais. A diferença entre as análises reside no fato de que aqui existem mais correlações significativas.

5.3 Análise estatística dos dados referentes às dimensões médias para blocos de vedação e bloco portante

Para os dados referentes as dimensões médias para blocos de vedação e bloco portante contidos na Figura (54) seguinte, novamente serão aplicadas as técnicas de análise multivariada: análise de agrupamento, análise fatorial e análise de componentes principais.

	1 Olaria	2 Bloco	3 L (mm)	4 H (mm)	5 C (mm)	6 Ih (mm)	7 Iv (mm)	8 Eh (mm)	9 Ev (mm)
Agudo	1.1	4F	115,00	104,38	221,25	10,35	10,65	13,53	13,05
Agudo	1.1	6F	94,17	142,13	188,92	10,35	9,25	11,30	11,28
Agudo	1.1	6F	104,46	153,33	201,54	8,98	8,95	9,75	9,23
Cachoeira	3.1	6F	102,50	174,42	216,21	12,40	8,90	15,08	11,63
Cachoeira	3.1	6F	112,75	173,88	230,63	11,13	10,55	13,35	12,10
Cachoeira	3.1	10F	124,75	257,92	251,42	10,00	8,03	10,38	11,30
Cachoeira	3.1	BE	141,46	193,17	289,13			20,65	20,65
Cachoeira	3.3	2F	118,04	61,29	235,63		13,35	10,58	14,28
Cachoeira	3.3	4F	120,58	121,83	232,21	14,28	13,08	12,80	13,68
Cachoeira	3.3	6F	120,04	148,33	206,04	11,28	8,85	11,63	11,93
Cachoeira	3.3	6F	114,58	173,08	232,67	13,33	10,95	13,25	13,13
Faxinal	7.2	6F	91,67	135,38	177,33	10,50	9,00	10,78	10,23
Faxinal	7.2	6F	104,13	148,04	206,17	9,75	9,25	10,15	9,90
Formigueiro	8.1	4F	103,33	103,75	209,58	9,55	8,95	10,70	11,43
Formigueiro	8.1	6F	105,00	151,04	206,88	9,38	9,05	11,55	10,18
Jaguari	11.1	6F	97,88	135,67	192,92	10,93	12,70	10,05	11,35
Mata	14.1	6F	87,08	135,46	199,96	10,50	9,60	11,10	10,40
Mata	14.1	6F	86,79	135,75	202,54	9,53	8,25	8,95	8,55
Paraíso	18.1	4F	113,08	113,00	234,79	8,73	8,85	11,00	11,28
Paraíso	18.1	6F	112,29	168,63	234,96	10,58	9,65	11,73	10,35
Paraíso	21.1	6F	94,29	130,46	190,46	9,58	8,25	10,73	10,95
Rest. Seca	21.3	2F	110,17	61,33	232,75		10,90	12,30	12,55
Rest. Seca	21.3	4F	109,13	107,92	221,04	9,85	9,18	13,23	12,78
Rest. Seca	21.3	6F	92,71	128,17	189,17	10,33	11,79	11,30	11,38
Rest. Seca	21.3	6F	102,25	148,33	213,04	10,05	9,30	11,45	12,13
Santiago	22.1	6F	97,96	146,58	207,50	9,95	8,78	11,95	12,00
Santiago	22.2	4F	99,71	103,50	208,13	12,88	11,28	14,03	12,78
Santiago	22.2	6F	101,21	149,17	210,83	9,93	10,30	12,13	12,48
Santiago	22.3	4F	119,46	121,88	224,46	12,40	12,05	14,60	14,73
Santiago	22.3	6F	101,33	144,38	224,42	10,95	12,15	12,88	14,03
Santiago	22.3	6F	119,29	175,75	222,00	13,58	11,60	13,98	14,65
São Martin	25.1	6F	90,96	140,21	182,71	11,65	10,50	12,80	11,70
São Sené	27.2	6F	98,79	133,21	197,08	10,30	11,38	12,88	14,88

Figura 54 - Dados das dimensões médias para blocos de vedação e bloco portante

5.3.1 Análise de agrupamento das dimensões médias e características geométricas para blocos de vedação e bloco portante

Para a realização da análise de agrupamento foram realizadas as mesmas etapas descritas anteriormente.

Selecionando-se todas as variáveis, exceto a olaria e bloco, e a opção



obtem-se o dendograma a seguir (Figura 55).

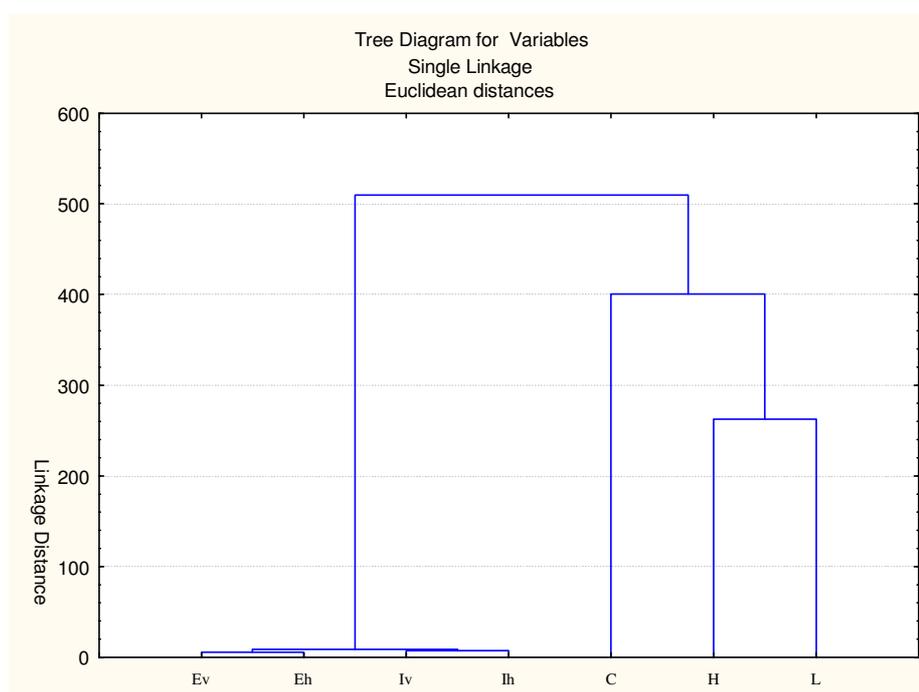


Figura 55 - Dendograma vertical dos dados de blocos

No dendograma elaborado verifica-se a formação de 2 grupos: um grupo composto pelas variáveis L, H e C e outro grupo composto pelas variáveis Ev, Eh, Iv e Ih, da direita para a esquerda.

Selecionado-se a opção  obtem-se a Figura (56):

Amalgamation Schedule							
Single Linkage							
Euclidean distances							
linkage distance	Obj. No. 1	Obj. No. 2	Obj. No. 3	Obj. No. 4	Obj. No. 5	Obj. No. 6	Obj. No. 7
5,674020	Eh	Ev					
7,560192	Ih	Iv					
9,105624	Ih	Iv	Eh	Ev			
262,6664	L	H					
400,6914	L	H	C				
510,1620	L	H	C	Ih	Iv	Eh	Ev

Figura 56 - Ligação entre as variáveis analisadas para blocos

Selecionado-se a opção  Graph of amalgamation schedule obtém-se a Figura (57):

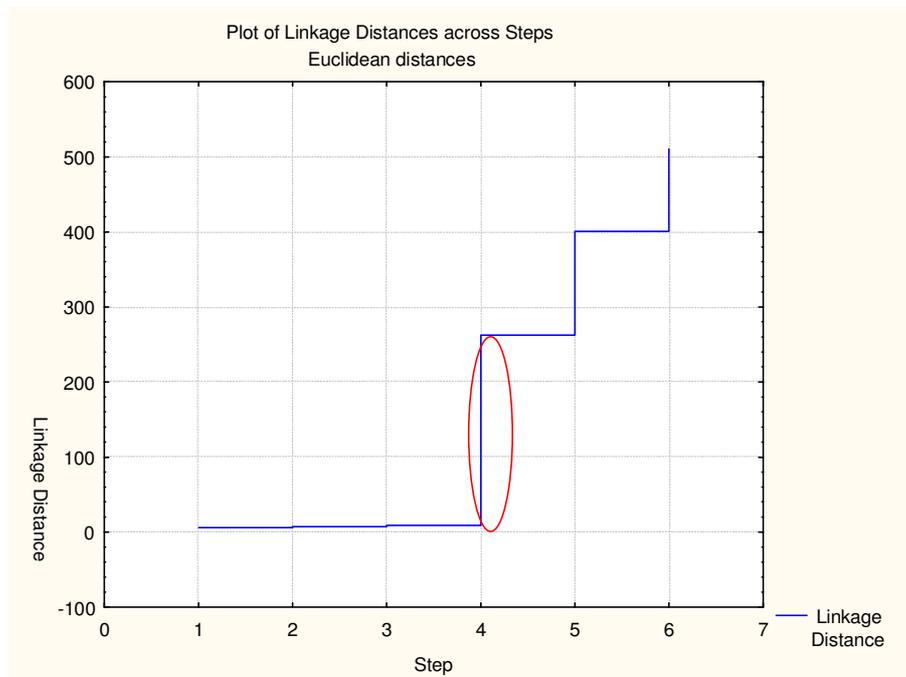


Figura 57 - Ligação das variáveis analisadas para blocos

Analisando-se o gráfico anterior verifica-se que os maiores saltos ocorrem entre as distâncias 9,1056 e 262,6664 e 262,6664 e 400,6914. Entre as distâncias 300 e 400 pode ser feito um corte no gráfico onde observa-se a formação dos grupos distintos, conforme o histograma.

Selecionado-se a opção  Distance matrix obtém-se a matriz de distâncias entre as variáveis (Figura 58).

Variable	Euclidean distances						
	L	H	C	Ih	Iv	Eh	Ev
L	0	263	587	517	521	510	511
H	263	0	401	745	749	739	740
C	587	401	0	1102	1106	1095	1096
Ih	517	745	1102	0	8	9	9
Iv	521	749	1106	8	0	13	12
Eh	510	739	1095	9	13	0	6
Ev	511	740	1096	9	12	6	0

Figura 58 - Matriz de distâncias para variáveis dos blocos

Selecionado-se a opção  **Descriptive statistics** obtêm-se os valores de média e desvio padrão das variáveis, para um maior conhecimento e compreensão dos resultados obtidos, conforme a Figura (59).

Variable	Means and Standard Deviations			
	Mean	Std.Dev.		
L	104,5723	10,63015		
H	143,5193	30,17184		
C	211,2287	17,62392		
Ih	10,7667	1,39979		
Iv	10,0357	1,42471		
Eh	11,9680	1,51833		
Ev	11,8497	1,58217		

Figura 59 - Estatística descritiva (média e desvio padrão) dos dados para blocos

Os resultados anteriores demonstram que as variáveis Ih, Iv, Eh e Ev apresentam um desvio padrão em torno de 1, o que significa uma pequena variação entre os resultados. Já as variáveis L, H e C apresentam um desvio padrão maior que representa uma maior variação entre os resultados.

5.3.2 Análise fatorial das dimensões médias e características geométricas para blocos de vedação e bloco portante

Para a realização da análise fatorial  **Factor Analysis**, selecionam-se as variáveis de interesse e a opção *Raw data*.

Selecionando-se *Quick* e clicando-se em *OK* obtêm-se a tela onde observa-se que existem as opções para a realização da análise fatorial. Selecionando-se *Eigenvalues* (autovalores) obtêm-se a Figura (60):

Value	Eigenvalues Extraction: Principal components			
	Eigenvalue	% Total variance	Cumulative Eigenvalue	Cumulative %
1	3,375093	48,21561	3,375093	48,2156
2	1,814450	25,92072	5,189543	74,1363
3	0,698671	9,98101	5,888214	84,1173
4	0,488358	6,97654	6,376571	91,0939
5	0,323667	4,62382	6,700239	95,7177
6	0,199571	2,85101	6,899809	98,5687
7	0,100191	1,43129	7,000000	100,0000

Figura 60 - Autovalor, % da variância total, autovalor acumulado e % acumulada, para blocos

Considerando-se as 7 variáveis e analisando-se os resultados obtidos verifica-se que o 1° e o 2° autovalor são maiores do que 1 e explicam 74,14% da variância. Os demais autovalores são menores do que 1. Portanto, os dados serão resumidos pela 1ª e 2ª componente principal.

Clicando-se em  ou em  obtém-se a Figura (61) a seguir.

Variable	Factor Loadings Extraction: Principal components (Marked loadings are > ,700000)						
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
L	-0,673209	0,599018	0,267871	0,132156	0,113194	0,272029	0,109243
H	-0,084122	0,766255	-0,595107	-0,022092	-0,223394	0,003257	0,034925
C	-0,581983	0,695512	0,273177	0,089981	0,051623	-0,290188	-0,089231
Ih	-0,807039	-0,196018	-0,409662	0,011713	0,353944	0,048406	-0,121181
Iv	-0,686630	-0,501948	-0,114169	0,473678	-0,114206	-0,107866	0,120431
Eh	-0,833148	-0,198456	0,006419	-0,480171	0,008420	-0,107210	0,155916
Ev	-0,871317	-0,234485	0,131297	-0,085065	-0,346257	0,126014	-0,159925
Expl.Var	3,375093	1,814450	0,698671	0,488358	0,323667	0,199571	0,100191
Prp.Totl	0,482156	0,259207	0,099810	0,069765	0,046238	0,028510	0,014313

Figura 61 - Composição dos fatores para as variáveis dos blocos

Na figura anterior verifica-se o resultado dos *Factor loadings* sendo que os valores destacados possuem uma significância maior do que 0,7, o que poderá ser alterado conforme descrito anteriormente. Observa-se que os valores de Ih, Eh e Ev para o fator 1 e o valor de H para o fator 2 são significativos. No caso dos demais fatores os valores não são significativos.

Selecionando-se , obtém-se a Figura (62) a seguir.

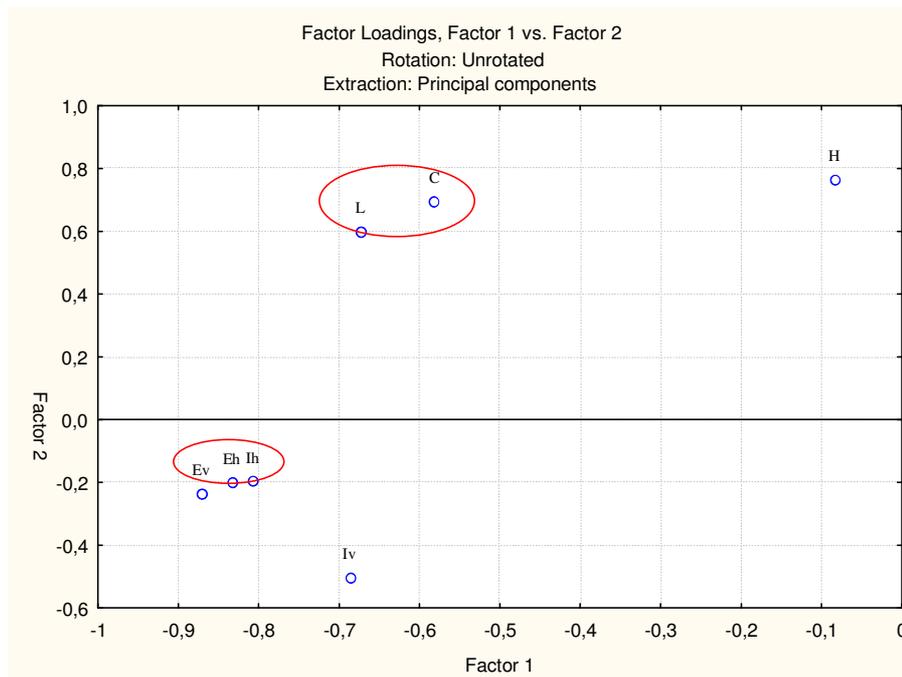


Figura 62 - Planos fatoriais para os blocos

Analisando-se a figura anterior, pode-se observar que as variáveis se agrupam por similaridades de explicação. As variáveis que explicam ou representam melhor o fator 1 são aquelas que estiverem localizadas mais distantemente da origem, em relação ao eixo x. As variáveis que estiverem localizadas próximas à origem do plano fatorial possuem baixa representatividade. Neste caso, num prolongamento do eixo x até a origem dos eixos, estas variáveis são Ih, Eh e Ev. Do mesmo modo, a variável que melhor representa o fator 2 é H.

Como descrito anteriormente, as distâncias de cada uma das variáveis analisadas são encontradas traçando-se segmentos de reta perpendiculares ao eixo considerado e determina-se aquela ou aquelas variáveis que estiverem localizadas mais distantemente da origem, à partir da perpendicular traçada, sendo que essas terão maior influência sobre o fator analisado.

Do mesmo modo, quando a identificação das variáveis que representam um fator tornar-se difícil realizam-se as rotações pois assim ocorre uma mais perfeita visualização da disposição dos pontos. Como já citado o tipo de rotação mais empregado é a *Varimax normalized* que mantém os eixos ortogonais. Para uma melhor visualização dos fatores seleciona-se *Plot of loadings, 2D*, conforme a Figura (63) a seguir.

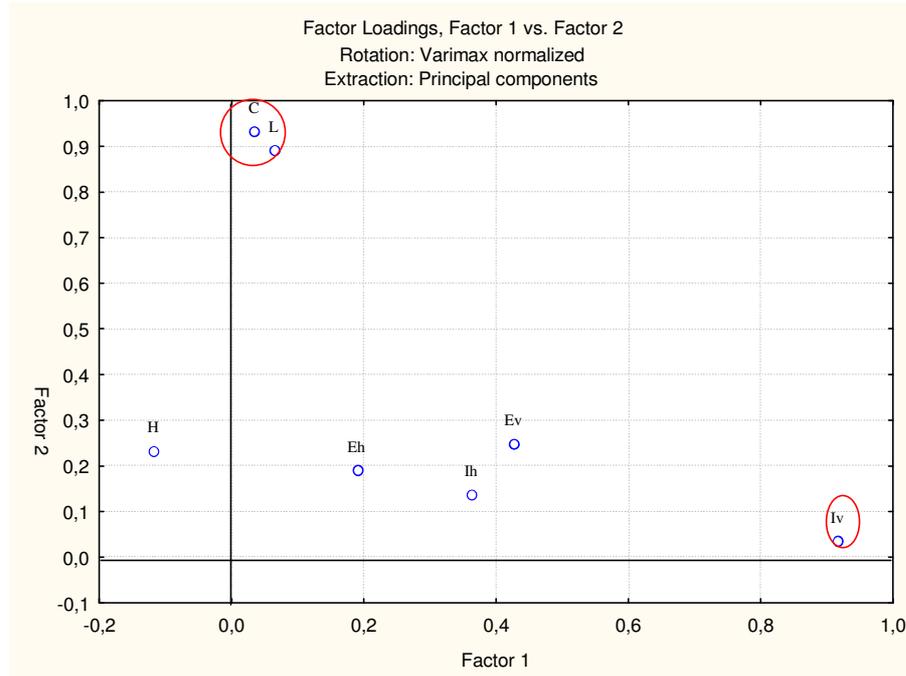


Figura 63 - Planos fatoriais, com rotação *varimax normalized*, para os dados de blocos

Analisando-se o gráfico, novamente as variáveis de maior representatividade são aquelas localizadas mais distantemente da origem do plano fatorial, neste caso Iv e L e C. Aquelas que estiverem localizadas próximas à origem do plano fatorial, possuem baixa representatividade.

Abaixo são apresentados os resultados dos *Factor Loadings* rotacionados (Figura 64).

Variable	Factor Loadings (Varimax normalized)						
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
L	0,064046	0,891719	0,141844	0,137732	0,153993	0,339053	0,151575
H	-0,118313	0,232644	0,962248	-0,053018	0,046068	0,004042	-0,031741
C	0,033088	0,933235	0,204831	0,152289	0,033073	-0,246458	0,031553
Ih	0,362725	0,137308	0,075959	0,364841	0,836368	0,018966	0,104057
Iv	0,916710	0,035009	-0,142358	0,198956	0,274466	0,005834	0,152325
Eh	0,189817	0,190364	-0,056675	0,901798	0,293755	0,005398	0,157980
Ev	0,427013	0,249104	-0,080069	0,536137	0,179389	0,025968	0,654903
Expl.Var	1,209491	1,838583	1,023656	1,318349	0,920245	0,176812	0,512865
Prp.Totl	0,172784	0,262655	0,146237	0,188336	0,131464	0,025259	0,073266

Figura 64 - Matriz fatorial rotacionada para os blocos

Selecionando-se *Scree plot*   obtém-se o resultado conforme a Figura (65) a seguir.

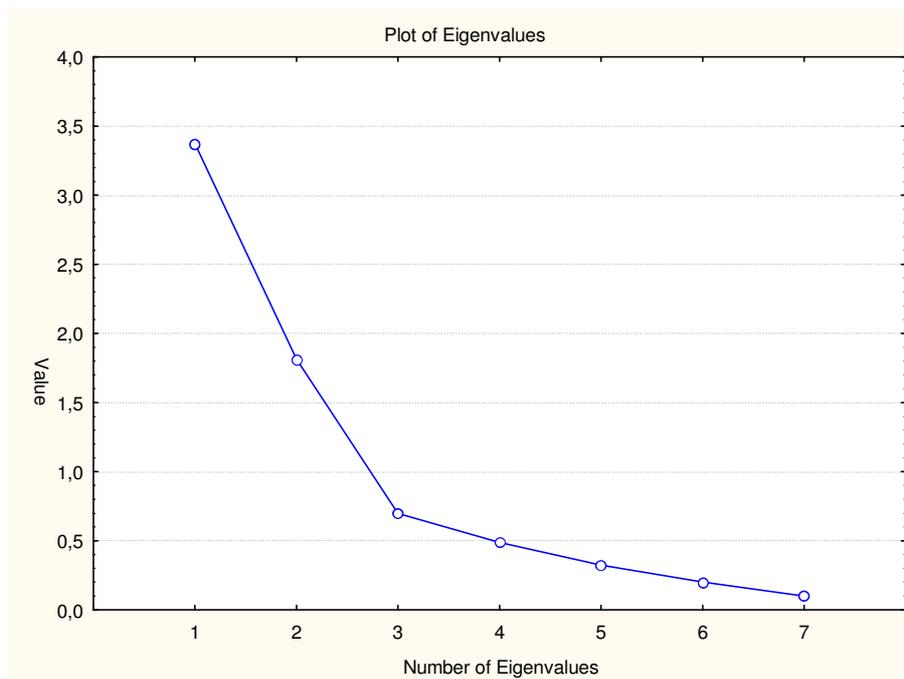


Figura 65 - Autovalores para os blocos

Examinando-se o gráfico anterior, observa-se uma queda brusca entre os fatores 1 e 3 e, analisando-se os autovalores superiores a 1, observa-se que o 1º e o 2º fator deverão ser considerados.

Selecionando-se *Communalities*  *Communalities*, comunalidades, obtem-se o resultado a seguir (Figura 66).

Communalities								
Extraction: Principal components								
Rotation: Varimax normalized								
Variable	From 1 Factor	From 2 Factors	From 3 Factors	From 4 Factors	From 5 Factors	From 6 Factors	From 7 Factors	Multiple R-Square
L	0,004102	0,799265	0,819384	0,838354	0,862068	0,977025	1,000000	0,717366
H	0,013998	0,068121	0,994043	0,996854	0,998976	0,998993	1,000000	0,335405
C	0,001095	0,872022	0,913977	0,937169	0,938263	0,999004	1,000000	0,696639
Ih	0,131569	0,150423	0,156192	0,289302	0,988812	0,989172	1,000000	0,680592
Iv	0,840357	0,841582	0,861848	0,901432	0,976763	0,976797	1,000000	0,660659
Eh	0,036030	0,072269	0,075481	0,888721	0,975013	0,975042	1,000000	0,733387
Ev	0,182340	0,244393	0,250804	0,538247	0,570427	0,571102	1,000000	0,764154

Figura 66 - Comunalidades para os blocos

Cabe lembrar que as comunalidades são índices que expressam quanto da variância é explicada pela AF, ou seja, quanto mais próximo de 1, melhor será o ajuste da AF.

Na opção *Plot of loadings, 3D*, obtem-se uma figura tri-dimensional, conforme a Figura (67) abaixo.

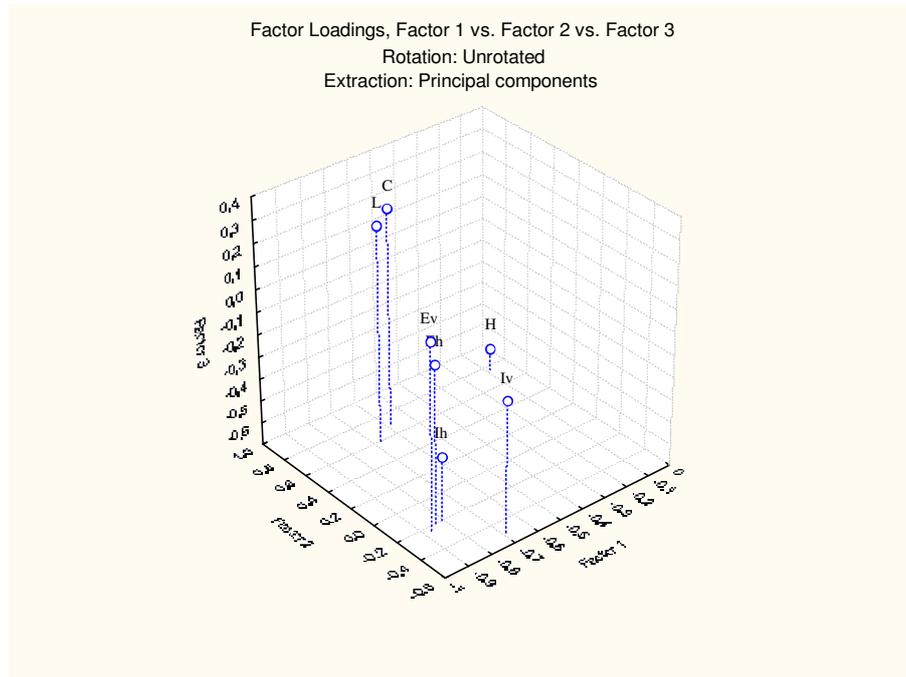


Figura 67 - Gráfico tri-dimensional dos planos fatoriais para os blocos

Selecionando-se *Factor score coefficients* surgem os autovetores segundo a Figura (68) a seguir.

Factor Score Coefficients							
Rotation: Unrotated							
Extraction: Principal components							
Variable	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
L	-0,199464	0,330138	0,383401	0,270613	0,34972	1,36307	1,09035
H	-0,024924	0,422307	-0,851771	-0,045238	-0,69020	0,01632	0,34859
C	-0,172435	0,383318	0,390995	0,184253	0,15950	-1,45406	-0,89061
Ih	-0,239116	-0,108031	-0,586344	0,023984	1,09354	0,24255	-1,20951
Iv	-0,203440	-0,276639	-0,163409	0,969940	-0,35285	-0,54049	1,20202
Eh	-0,246852	-0,109375	0,009188	-0,983236	0,02601	-0,53720	1,55620
Ev	-0,258161	-0,129232	0,187924	-0,174185	-1,06979	0,63142	-1,59621

Figura 68 - Resultados dos autovetores para os blocos

Como pode ser observado, existem tantos componentes quantas variáveis existirem. Neste caso como existem sete variáveis, tem-se sete componentes principais. A partir dos autovalores, determinam-se os autovetores que formarão a combinação linear denominada de componentes principais ou variáveis de referência. A tabela anterior (Figura 68) contém os autovetores, derivados dos autovalores, que definem a direção dos eixos para AF e ACP.

Baseando-se nestes autovetores, pode-se escrever as seguintes combinações lineares:

$$Y_1 = -0,1995X_1 - 0,0249X_2 - 0,1724X_3 - 0,2391X_4 - 0,2034X_5 - 0,2468X_6 - 0,2582X_7$$

$$Y_2 = 0,3301X_1 + 0,4223X_2 + 0,3833X_3 - 0,1080X_4 - 0,2766X_5 - 0,1094X_6 - 0,1292X_7$$

$$Y_3 = 0,3834X_1 - 0,8518X_2 + 0,3910X_3 - 0,5863X_4 - 0,1634X_5 + 0,0092X_6 + 0,1879X_7$$

$$Y_4 = 0,2706X_1 - 0,0452X_2 + 0,1842X_3 + 0,0240X_4 + 0,9699X_5 - 0,9832X_6 - 0,1742X_7$$

$$Y_5 = 0,34977X_1 - 0,6902X_2 + 0,1595X_3 + 1,0935X_4 - 0,3528X_5 + 0,0260X_6 - 1,0700X_7$$

$$Y_6 = 1,3631X_1 + 0,0163X_2 - 1,4541X_3 + 0,2425X_4 - 0,5405X_5 - 0,5372X_6 + 0,6314X_7$$

$$Y_7 = 1,0903X_1 + 0,3486X_2 - 0,8906X_3 - 1,2095X_4 + 1,2020X_5 + 1,5562X_6 - 1,5962X_7$$

Selecionando-se *Factor scores* surgem os resultados dos componentes principais, segundo a Figura (69) a seguir.

Case	Factor Scores						
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
Agudo	-0,72774	-0,30370	1,95945	-0,30365	0,06658	0,11026	1,37928
Agudo	0,79953	-0,54836	-0,63795	-0,54268	-0,26953	0,74082	-0,36846
Agudo	1,33708	0,64555	0,05239	0,83614	0,29060	0,63182	1,58886
Cachoeira	-0,62233	0,36472	-1,39775	-2,78326	1,02914	-1,13487	0,93449
Cachoeira	-0,76939	0,85299	-0,30482	-0,20064	-0,23902	-1,05691	1,49304
Cachoeira	-0,10113	3,70995	-1,13390	0,47261	-1,34717	0,30375	-0,74245
Cachoeira	-1,95653	-0,42180	0,05673	2,05186	1,98006	0,19977	-1,13001
Cachoeira	-0,11998	0,64325	0,23542	-0,25606	0,98644	3,10446	0,03296
Cachoeira	-1,40777	0,61861	-1,01665	0,12985	0,77942	-0,31510	-1,13668
Faxinal	1,23136	-0,81237	-0,95671	-0,43267	0,57790	1,25860	0,06863
Faxinal	0,95362	0,46074	0,01760	0,76871	0,52372	0,35012	0,58173
Formigueiro	0,70979	-0,20076	1,61797	0,11813	0,43438	0,43741	-1,24451
Formigueiro	0,74633	0,48886	0,19969	-0,28615	0,08529	0,03290	1,97314
Jaguari	0,29620	-1,06688	-0,87093	2,76372	-0,43357	0,14491	0,79310
Mata	0,93056	-0,61504	-0,66915	-0,13051	0,37180	-1,46999	-0,88215
Mata	1,92040	0,07911	-0,30215	0,54892	1,16958	-1,35344	-1,67697
Paraíso	0,40259	0,85317	2,60674	0,35614	0,26243	-0,65748	-0,32855
Paraíso	-0,02740	1,33624	0,03882	0,46039	0,85387	-1,35508	0,98743
Paraíso	1,21269	-0,35286	0,12459	-0,79457	-0,12520	0,93875	-0,99845
Rest. Seca	-0,23031	-0,07334	1,98730	-1,24593	-0,05854	-0,15377	-0,01459
Rest. Seca	0,46041	-1,28348	-0,56203	1,16157	-0,70826	-0,40172	1,36724
Rest. Seca	0,28783	0,24718	0,23536	-0,25593	-0,74616	0,00543	-1,08933
Santiago	0,45523	0,05224	0,09624	-1,08564	-0,75027	-0,13738	-0,97831
Santiago	-0,87112	-1,40779	-0,01932	-0,65055	1,47654	-0,85362	-0,40589
Santiago	0,03831	-0,08391	0,10636	-0,10692	-1,38600	-0,44621	0,21661
Santiago	-1,85547	-0,49479	0,88423	-0,07242	-0,02079	0,54277	0,68834
Santiago	-0,90621	-0,47022	0,09661	0,66602	-1,84602	-1,72651	-0,62874
Santiago	-1,89618	0,24794	-1,15319	-0,05907	-0,20419	1,31568	-0,53642
São Martinho	0,20914	-1,29540	-1,46625	-0,83078	0,06031	0,22847	0,63894
São Sepé	-0,49953	-1,16984	0,17533	-0,29665	-2,81331	0,71614	-0,58227

Figura 69 – Resultados das componentes principais para os blocos

Como existem sete variáveis são encontrados sete componentes mas, segundo a análise fatorial, somente a 1ª e a 2ª componente principal devem ser consideradas.

Selecionando-se *Descriptives* surgem as opções de estatísticas descritivas, correlações da revisão (*Review correlations*), médias (*means*), desvio padrão (*standard deviations*) e regressão múltipla (*Multiple regression*).

As correlações são determinadas a partir do ícone , conforme a Figura (70).

		Correlations						
		Casewise deletion of MD						
		N=30						
Variable		L	H	C	Ih	Iv	Eh	Ev
L		1,00	0,33	0,81	0,36	0,16	0,37	0,45
H		0,33	1,00	0,40	0,08	-0,24	-0,07	-0,11
C		0,81	0,40	1,00	0,24	0,08	0,32	0,33
Ih		0,36	0,08	0,24	1,00	0,64	0,68	0,60
Iv		0,16	-0,24	0,08	0,64	1,00	0,47	0,67
Eh		0,37	-0,07	0,32	0,68	0,47	1,00	0,77
Ev		0,45	-0,11	0,33	0,60	0,67	0,77	1,00

Figura 70 - Correlações para os blocos

Analisando-se os resultados na matriz de correlação, observa-se que alguns valores são superiores a 0,7 o que indica uma correlação de moderada a forte entre as variáveis. Assim, conclui-se que as variáveis estão interligadas entre si e a análise das mesmas não poderá ser realizada de maneira isolada e sim conjuntamente. Salienta-se que pode ser realizado um teste de significância das correlações, podendo ser o teste *KMO* e o teste de *Bartlett*.

5.3.3 Análise de componentes principais das dimensões médias e características geométricas para blocos de vedação e bloco portante

Determinam-se os componentes principais selecionando-se no *software Statistica* o módulo referente a esta análise, como descrito anteriormente.

Selecionam-se as variáveis e clica-se em *OK*. Aqui novamente não serão usadas variáveis suplementares. Na figura seguinte, onde aparecem as variáveis selecionadas, clica-se novamente em *OK*. Surge uma tela contendo as opções disponíveis para a determinação dos componentes principais. Para encontrar os componentes principais selecionam-se as opções *Cases/Factor scores* e *OK*

A Figura (71) a seguir apresenta os resultados encontrados para os componentes.

Case	Factor scores, based on correlations						
	Factor 1	Factor 2	Factor 3	Factor 4	Factor 5	Factor 6	Factor 7
Agudo	-0,72774	-0,30370	1,95945	-0,30365	0,06658	0,11026	1,37928
Agudo	0,79953	-0,54836	-0,63795	-0,54268	-0,26953	0,74082	-0,36846
Agudo	1,33708	0,64555	0,05239	0,83614	0,29060	0,63182	1,58886
Cachoeira	-0,62233	0,36472	-1,39775	-2,78326	1,02914	-1,13487	0,93449
Cachoeira	-0,76939	0,85299	-0,30482	-0,20064	-0,23902	-1,05691	1,49304
Cachoeira	-0,10113	3,70995	-1,13390	0,47261	-1,34717	0,30375	-0,74245
Cachoeira	-1,95653	-0,42180	0,05673	2,05186	1,98006	0,19977	-1,13001
Cachoeira	-0,11998	0,64325	0,23542	-0,25606	0,98644	3,10446	0,03296
Cachoeira	-1,40777	0,61861	-1,01665	0,12985	0,77942	-0,31510	-1,13668
Faxinal	1,23136	-0,81237	-0,95671	-0,43267	0,57790	1,25860	0,06863
Faxinal	0,95362	0,46074	0,01760	0,76871	0,52372	0,35012	0,58173
Formigueiro	0,70979	-0,20076	1,61797	0,11813	0,43438	0,43741	-1,24451
Formigueiro	0,74633	0,48886	0,19969	-0,28615	0,08529	0,03290	1,97314
Jaguari	0,29620	-1,06688	-0,87093	2,76372	-0,43357	0,14491	0,79310
Mata	0,93056	-0,61504	-0,66915	-0,13051	0,37180	-1,46999	-0,88215
Mata	1,92040	0,07911	-0,30215	0,54892	1,16958	-1,35344	-1,67697
Paraíso	0,40259	0,85317	2,60674	0,35614	0,26243	-0,65748	-0,32855
Paraíso	-0,02740	1,33624	0,03882	0,46039	0,85387	-1,35508	0,98743
Paraíso	1,21269	-0,35286	0,12459	-0,79457	-0,12520	0,93875	-0,99845
Rest. Seca	-0,23031	-0,07334	1,98730	-1,24593	-0,05854	-0,15377	-0,01459
Rest. Seca	0,46041	-1,28348	-0,56203	1,16157	-0,70826	-0,40172	1,36724
Rest. Seca	0,28783	0,24718	0,23536	-0,25593	-0,74616	0,00543	-1,08933
Santiago	0,45523	0,05224	0,09624	-1,08564	-0,75027	-0,13738	-0,97831
Santiago	-0,87112	-1,40779	-0,01932	-0,65055	1,47654	-0,85362	-0,40589
Santiago	0,03831	-0,08391	0,10636	-0,10692	-1,38600	-0,44621	0,21661
Santiago	-1,85547	-0,49479	0,88423	-0,07242	-0,02079	0,54277	0,68834
Santiago	-0,90621	-0,47022	0,09661	0,66602	-1,84602	-1,72651	-0,62874
Santiago	-1,89618	0,24794	-1,15319	-0,05907	-0,20419	1,31568	-0,53642
São Martinho	0,20914	-1,29540	-1,46625	-0,83078	0,06031	0,22847	0,63894
São Sené	-0,49953	-1,16984	0,17533	-0,29665	-0,81331	0,71614	-0,58277

Figura 71 - Componentes principais para os blocos

Analisando-se os resultados, observa-se que estes são iguais aos obtidos anteriormente pela AF, assim como os demais resultados tanto para as variáveis quanto para os casos.

Neste caso na interpretação dos componentes principais, verifica-se que pode haver uma redução do número de variáveis de sete para dois componentes, pois ao invés de se trabalhar com as sete variáveis passa-se a trabalhar com o primeiro e segundo componente, representados pelas combinações lineares.

Outro gráfico que pode ser utilizado como complemento para um maior esclarecimento das análises realizadas é o círculo de correlação unitário.

Selecionando-se *Variables/Plot var. factor coordinates, 2D* e os fatores a serem plotados, obtém-se o círculo de correlação (Figura 72).

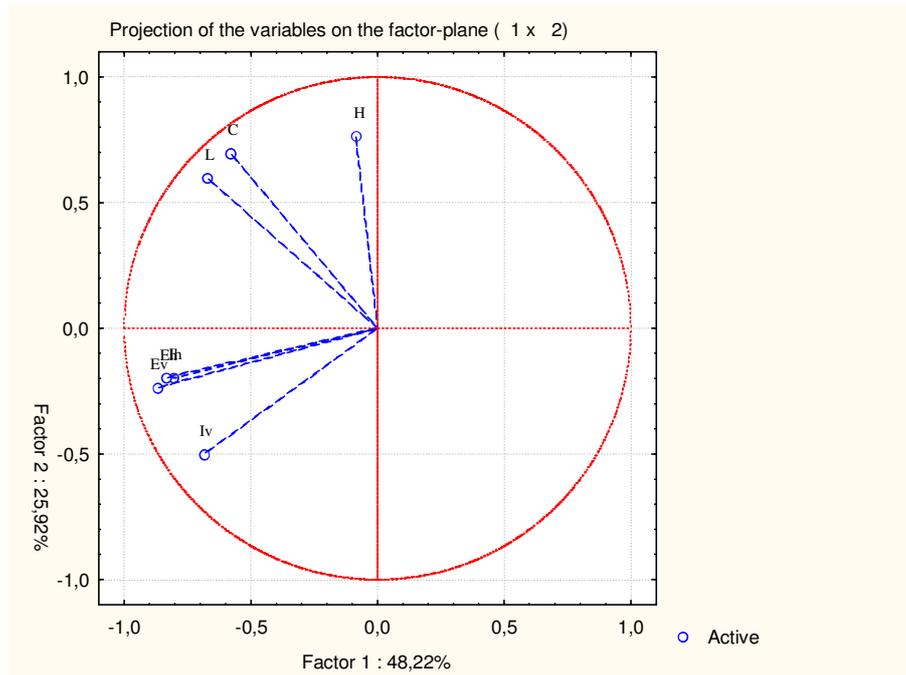


Figura 72 - Círculo de correlação unitário para os blocos

Analisando-se a figura anterior, e conforme descrito anteriormente, observa-se que as variáveis que estiverem mais próximas ao círculo unitário, possuem uma maior contribuição em relação aquelas que estiverem mais afastadas. Ocorre uma concentração das variáveis no 2º e 3º quadrantes. Nota-se também que as variáveis estão localizadas praticamente a mesma distância do círculo e as variáveis Ih, Eh e Ev estão quase sobrepostas. Também deve ser analisado o quadrante no qual as variáveis se encontram e o ângulo formado entre 2 variáveis, indicando uma maior ou menor correlação entre as mesmas.

Fazendo-se uma sobreposição deste sobre o 1º plano fatorial, pode-se identificar visualmente quais as variáveis que estão relacionadas com os casos em estudo. Para construir o gráfico dos planos principais (Figuras 73 e 74), seleciona-se *Cases/Plot case. factor coordinates, 2D* e os fatores a serem plotados.

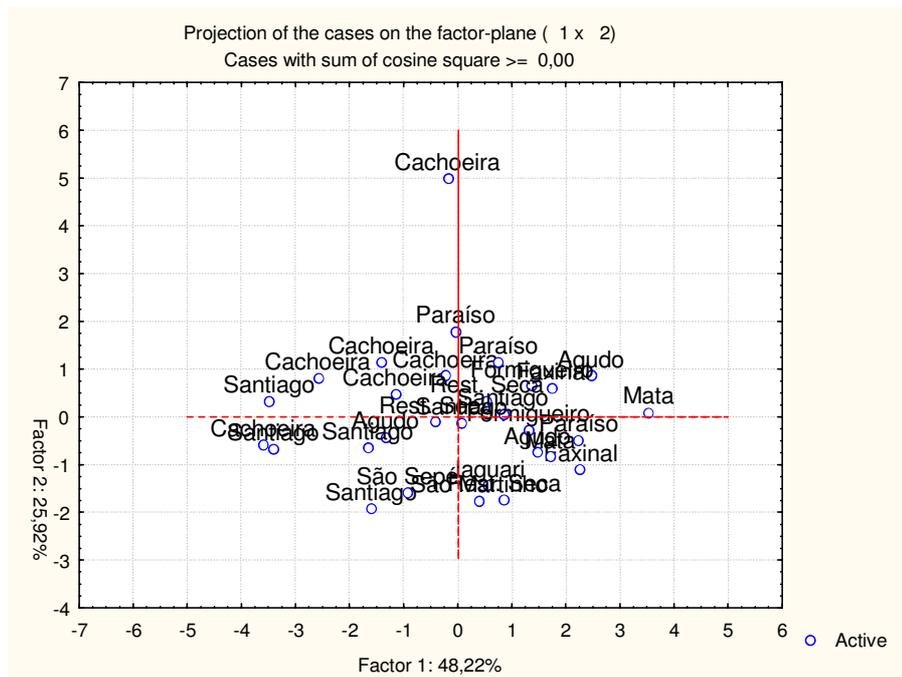


Figura 73 - Distribuição da nuvem de pontos por municípios, para os blocos

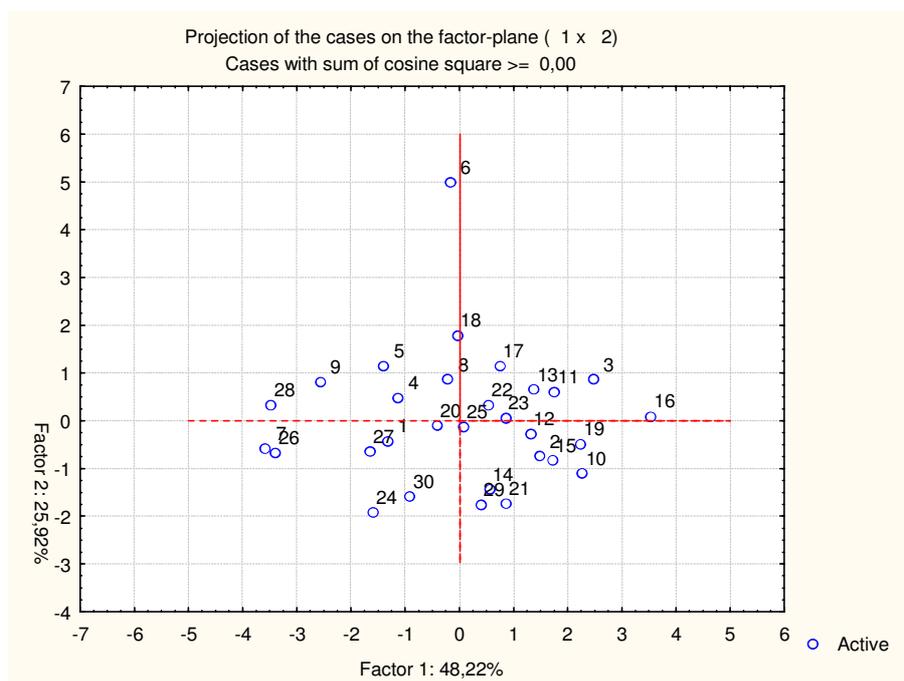


Figura 74 - Distribuição da nuvem de pontos por casos, para os blocos

Analisando-se o gráfico anterior, conclui-se que a maior parte dos casos estão concentrados em torno da origem. Esses, não apresentam representatividade significativa em relação a esses dois fatores. Também observa-se que ocorrem várias sobreposições. O único caso que se encontra mais afastado da nuvem formada é Cachoeira do Sul, podendo ser o mais representativo.

Outra forma de encontrar a contribuição de cada variável em relação aos fatores formados é através da matriz de correlação entre as variáveis originais e os componentes principais. Para determinar essa correlação seleciona-se *Cases/Save case statistics* e a opção *Factor Scores*. Clica-se em *OK*. A seguir, selecionam-se as variáveis e *OK*. Surge uma tela contendo as variáveis originais e os componentes principais.

Para construir a matriz de correlação seleciona-se *Statistics/Basic Statistics/Tables* na barra principal do *Statistica*. Surge uma caixa contendo as estatísticas descritivas disponíveis. Seleciona-se *Correlation matrices* e *OK*. Aparece uma caixa de diálogo contendo as opções para a seleção das variáveis.

Seleciona-se a opção *Two lists (rect. matrix)*. Surge uma caixa contendo as variáveis originais e os componentes principais.

A seguir clica-se em *Summary: Correlation matrix* para obter-se a matriz de correlação (Figura 75).

Correlations							
Marked correlations are significant at $p < ,05000$							
N=30 (Casewise deletion of missing data)							
Variable	Factor1	Factor2	Factor3	Factor4	Factor5	Factor6	Factor7
L	-0,67	0,60	0,27	0,13	0,11	0,27	0,11
H	-0,08	0,77	-0,60	-0,02	-0,22	0,00	0,03
C	-0,58	0,70	0,27	0,09	0,05	-0,29	-0,09
Ih	-0,81	-0,20	-0,41	0,01	0,35	0,05	-0,12
Iv	-0,69	-0,50	-0,11	0,47	-0,11	-0,11	0,12
Eh	-0,83	-0,20	0,01	-0,48	0,01	-0,11	0,16
Ev	-0,87	-0,23	0,13	-0,09	-0,35	0,13	-0,16

Figura 75 - Matriz de correlação entre variáveis originais e componentes principais, para os blocos

Ressalta-se que as correlações significativas aparecem em vermelho e ocorrem quando o valor de p for menor do que 0,05. Os valores encontrados representam o quanto cada variável contribui para cada fator, sendo os maiores valores absolutos os que melhor explicam o fator.

Para o fator 1: -0,67; -0,58; -0,81; -0,69; -0,83 e -0,87.

Para o fator 2: 0,60; 0,77; 0,70 e 0,50.

Para o fator 3: -0,60 e -0,41.

Para o fator 4: 0,47 e -0,48.

Para o fator 5: -0,38.

Comparando-se os resultados aqui obtidos com os *factor loadings* obtidos anteriormente, observa-se que os valores são iguais. A diferença entre as análises reside no fato de que aqui existem mais correlações significativas.

5.4 Análise estatística dos dados referentes à absorção de água, taxa de absorção inicial (IRA) e resistência à compressão para tijolos maciços, blocos de vedação e bloco portante

Para os dados de tijolos maciços, blocos de vedação e bloco portante referentes a absorção de água, taxa de absorção inicial (IRA) e resistência à compressão contidos na tabela seguinte (Figura 76), serão aplicadas as técnicas de análise de agrupamento. Outros recursos gráficos utilizados serão o círculo de correlação unitário e os planos fatoriais.

	1	2	3	4	5	6	7	8
	Olaria	Bloco	Absorção de água	CV Absorção de água	IRA	CV IRA	Resist. a compressão	CV Resist. a compressão
Agudo	1.1	4F	11,95	5,56	26,92	21,39	1,6	19,67
Agudo	1.1	6F	12,04	2,72	32,3	32,97	1,68	25,21
Agudo	1.1	6F	11,09	2,4	24,22	10,29	0,5	14,43
Cachoeira	3.1	6F	14,79	11,6	36,55	34,39	1,23	25,69
Cachoeira	3.1	6F	15,44	2,6	33,27	35,25	1,08	16,14
Cachoeira	3.1	10F	13,28	7,62	28,27	43,38	1,28	27,17
Cachoeira	3.1	BE	14,74	4,4	19,57	23,37	9,93	20,44
Cachoeira	3.3	2F	12,26	2,05	18,4	22,08	3	18,63
Cachoeira	3.3	4F	13,66	4,21	18,36	24,44	1,52	7,7
Cachoeira	3.3	6F	13,33	5,52	20,87	25,81	1,03	29,55
Cachoeira	3.3	6F	13,92	3,31	22,26	31,05	0,9	26,63
Fax. do Sot	7.2	6F	11,16	7,14	18,81	38,61	0,75	25,91
Fax. do Sot	7.2	6F	11,52	6,11	24,6	55,09	0,83	25,17
Formigueiro	8.1	4F	9,73	3,01	24,32	11,88	1,67	10,14
Formigueiro	8.1	6F	10,24	3,08	31,69	22,02	0,58	27,29
Jaguari	11.1	6F	11,6	5,22	17,12	25,63	0,65	21,76
Mata	14.1	6F	10,01	19,92	23,54	33,98	2,21	28,51
Mata	14.1	6F	13,64	4,85	17,56	35,48	1,96	34,75
Paraíso	18.1	4F	12,16	5,1	34,27	18,76	1,75	18,22
Paraíso	18.1	6F	13,13	4,01	27,03	10,73	0,82	33,67
Rest. Seca	21.1	6F	17,26	2,14	14,63	13,3	0,71	31,6
Rest. Seca	21.3	2F	16,77	3,34	18,79	34,63	2,22	8,92
Rest. Seca	21.3	4F	16,59	3,17	15,33	26,04	1,56	21,05
Rest. Seca	21.3	6F	16,75	4,96	19,07	16,24	0,78	17,2
Rest. Seca	21.3	6F	16,31	3,44	16,72	22,59	0,74	27,04
Santiago	22.1	6F	20,71	2,51	27,33	10,28	0,87	23,18
Santiago	22.2	4F	19	4,48	61,24	19,86	1,92	17,08
Santiago	22.2	6F	19,61	4,93	45,84	29,3	1,84	25,04
Santiago	22.3	4F	19,41	1,35	12,42	14,74	1,93	23,27
Santiago	22.3	6F	19,32	7,59	13,49	26,16	1,65	51,48
Santiago	22.3	6F	20,72	1,92	58,45	11,72	1,16	24,11
S. Martinho	25.1	6F	23,08	3,39	25,06	20,34	0,93	37,08
São Sené	27.2	6F	15,45	7,51	31,34	31,99	1,76	30,8

Figura 76 - Dados de absorção de água, IRA e resistência à compressão para tijolos maciços

5.4.1 Análise de agrupamento para os dados de absorção de água, IRA e resistência à compressão para tijolos maciços

Como as variáveis se apresentam em unidades de medidas diferentes, realizou-se a padronização das mesmas. Isso foi feito selecionando-se todas as variáveis e com o botão direito do *mouse*, selecionou-se a opção de padronização.

As variáveis padronizadas estão apresentadas na Figura (77) a seguir.

	1 Absorção de água	2 IRA	3 Resist. a compressão
Agudo	-0,827	0,076	-0,005
Agudo	-0,802	0,545	0,045
Agudo	-1,071	-0,159	-0,690
Cachoeira	-0,022	0,915	-0,235
Cachoeira	0,162	0,629	-0,329
Cachoeira	-0,450	0,193	-0,204
Cachoeira	-0,037	-0,565	5,188
Cachoeira	-0,740	-0,667	0,868
Cachoeira	-0,343	-0,670	-0,054
Cachoeira	-0,436	-0,451	-0,360
Cachoeira	-0,269	-0,330	-0,441
Fax. do Sot	-1,051	-0,631	-0,534
Fax. do Sot	-0,949	-0,126	-0,485
Formigueiro	-1,457	-0,151	0,039
Formigueiro	-1,312	0,491	-0,640
Jaguari	-0,927	-0,778	-0,597
Mata	-1 377	-0 219	0 376
Mata	-0,348	-0,740	0,220
Paraíso	-0,768	0,716	0,089
Paraíso	-0,493	0,085	-0,491
Rest. Seca	0,678	-0,995	-0,559
Rest. Seca	0,539	-0,633	0,382
Rest. Seca	0,488	-0,934	-0,029
Rest. Seca	0,533	-0,608	-0,516
Rest. Seca	0,409	-0,813	-0,541
Santiago	1,656	0,112	-0,460
Santiago	1,171	3,066	0,195
Santiago	1,344	1,724	0,145
Santiago	1,287	-1,188	0,201
Santiago	1,262	-1,094	0,027
Santiago	1,659	2,823	-0,279
S. Martinho	2,328	-0,086	-0,422
São Sepé	0,165	0,461	0,095

Figura 77 - Dados padronizados para tijolos maciços

Selecione-se a opção  Vertical icicle plot obtem-se o dendograma da Figura (78).

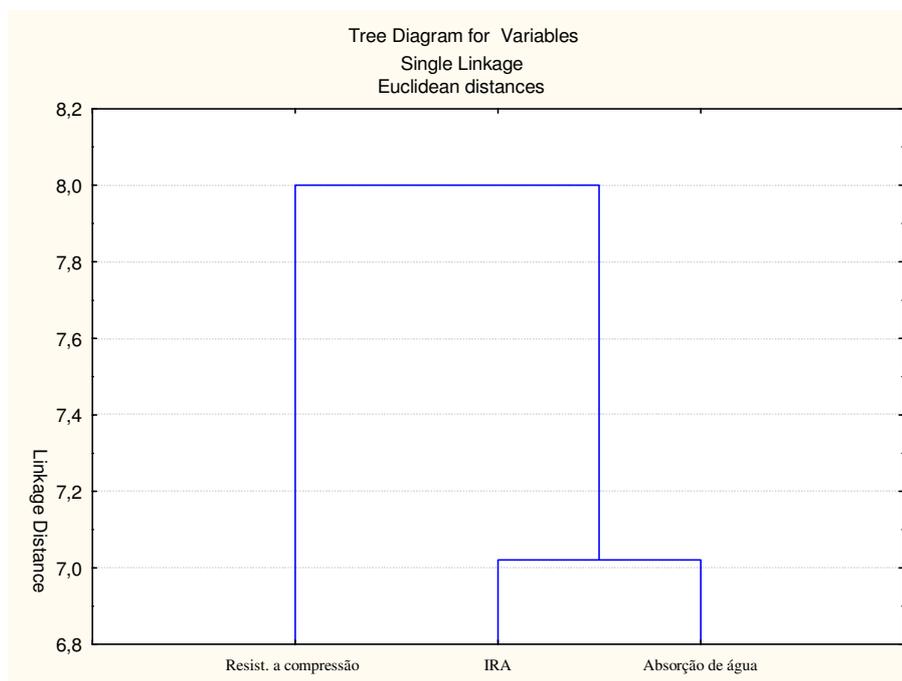


Figura 78 - Dendograma vertical dos dados de absorção de água, IRA e resistência à compressão de tijolos maciços

Em ambos os dendogramas elaborados verifica-se a formação de 2 grupos descritos da seguinte forma:

- para o dendograma horizontal de cima para baixo: um grupo composto pela variável Absorção e IRA e outro grupo composto pela variável Resistência à compressão.

Para o dendograma vertical observa-se a formação dos mesmos grupos, da direita para a esquerda.

Selecione-se a opção  Amalgamation schedule obtem-se a Figura (79):

Amalgamation Schedule Single Linkage Euclidean distances			
linkage distance	Obj. No. 1	Obj. No. 2	Obj. No. 3
7,020831	Absorção de água	IRA	
8,000903	Absorção de água	IRA	Resist. a compressão

Figura 79 - Ligação entre as variáveis de absorção de água, IRA e resistência à compressão para tijolos maciços

Selecione-se a opção  Graph of amalgamation schedule obtem-se a Figura (80).

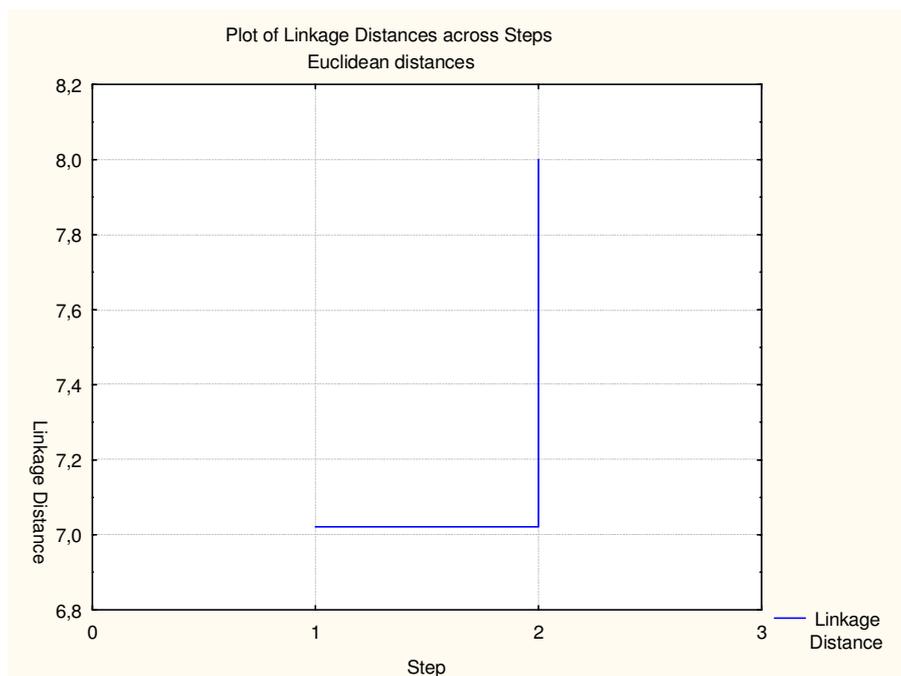


Figura 80 - Ligação das variáveis de absorção de água, IRA e resistência à compressão para tijolos maciços

Analisando-se o gráfico anterior verifica-se que o salto ocorre entre as distâncias 7 e 8. Entre essas distâncias pode ser feito um corte no gráfico onde observa-se a formação dos grupos distintos, conforme o histograma apresenta.

Selecionado-se a opção  Distance matrix obtém-se a matriz de distâncias entre as variáveis (Figura 81).

Variable	Euclidean distances		
	Absorção de água	IRA	Resist. a compressão
Absorção de água	0,00	7,02	8,00
IRA	7,02	0,00	8,28
Resist. a compressão	8,00	8,28	0,00

Figura 81 – Matriz de distâncias para absorção de água, IRA e resistência à compressão para tijolos maciços

O círculo de correlação unitário (Figura 82) e o plano principal (Figuras 83 e 84) são utilizados como complemento para um maior esclarecimento das análises realizadas.

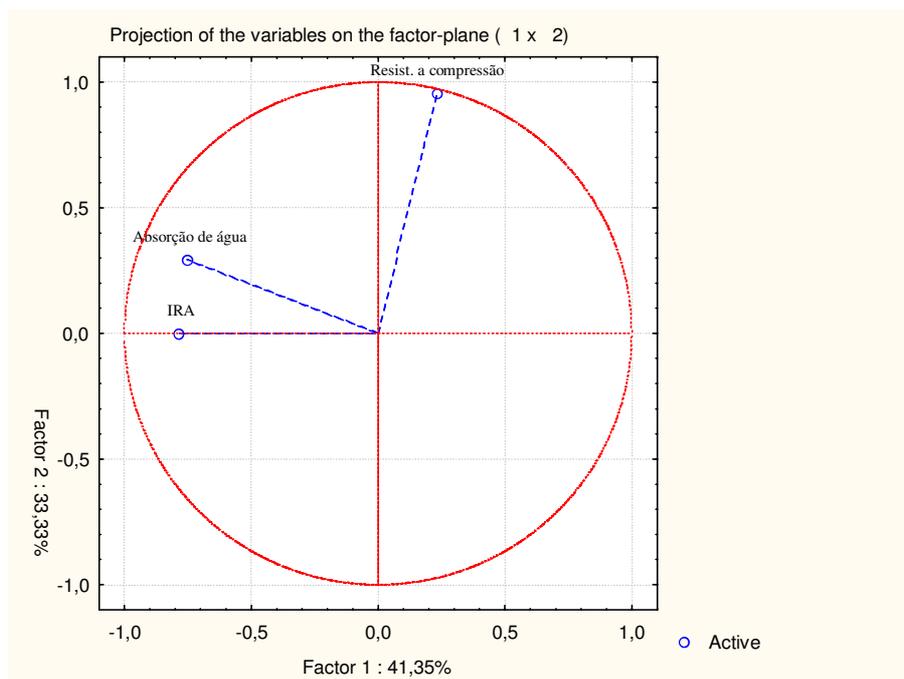


Figura 82 - Círculo de correlação unitário para absorção de água, IRA e resistência à compressão para tijolos maciços

Analisando-se a figura anterior observa-se que a variável resistência à compressão está localizada no 1º quadrante e sobre o círculo de correlação, indicando uma maior contribuição. A variável absorção de água está localizada no 2º quadrante próxima a variável IRA, localizada sobre o eixo. Observa-se que essas duas variáveis encontram-se mais afastadas do eixo significando uma menor contribuição. Nota-se que os agrupamentos são semelhantes aos formados na análise de *cluster*.

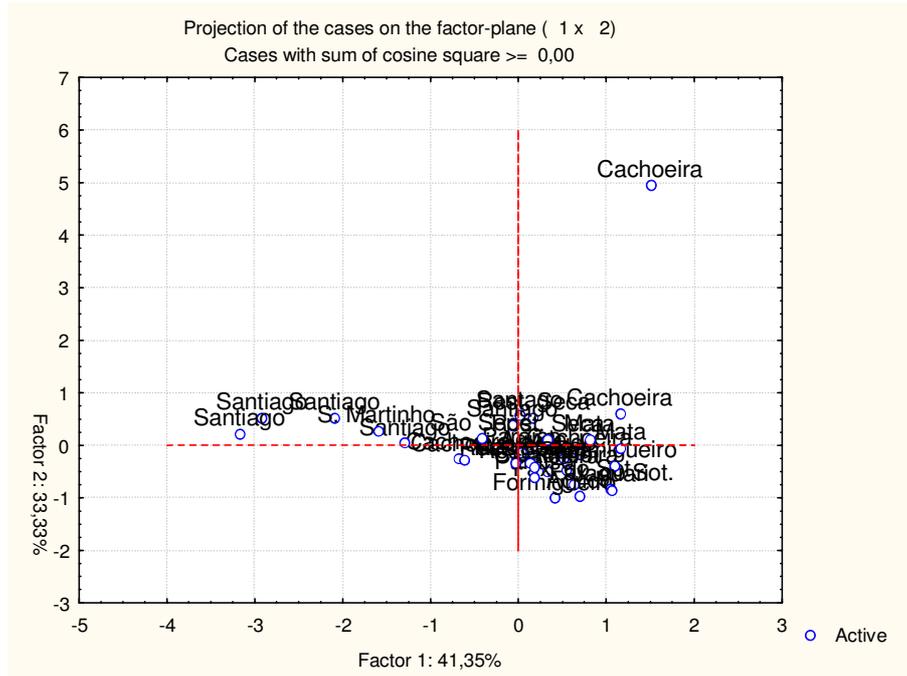


Figura 83 - Distribuição da nuvem de pontos por municípios, para tijolos maciços

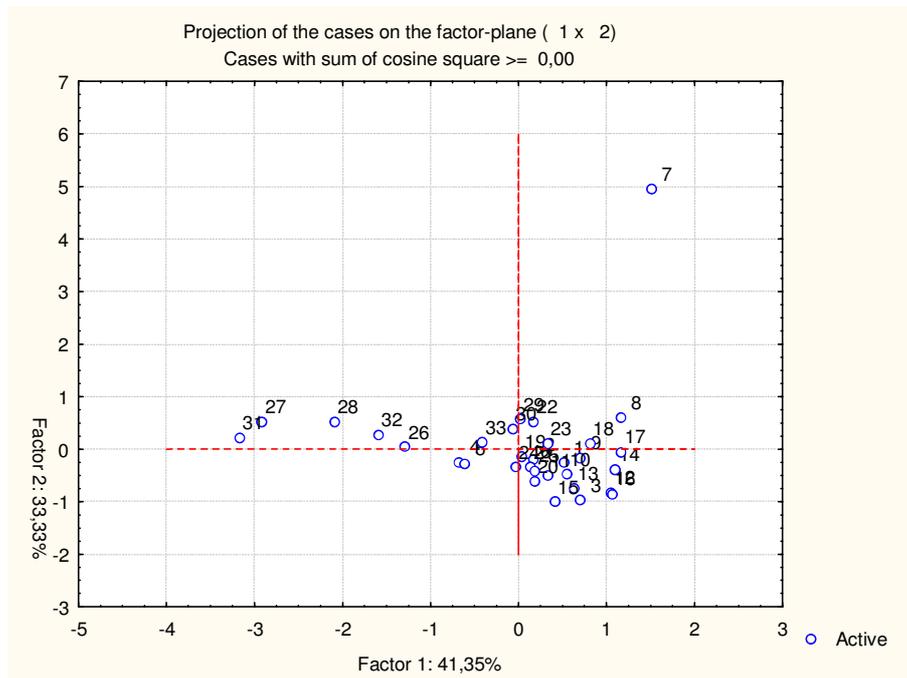


Figura 84 - Distribuição da nuvem de pontos por casos, para tijolos maciços

Analisando-se as figuras anteriores, através de uma sobreposição observa-se que o caso que está representando a variável resistência à compressão é Cachoeira do Sul (caso 7). Da mesma forma, os casos de Santiago e São Martinho (casos 26, 27, 28, 31 e 32) estão representando as variáveis absorção de água e IRA. Os demais casos encontram-se

sobrepostos e próximos a origem, não apresentando representatividade significativa em relação a estes dois fatores.

5.4.2 Análise de agrupamento para os dados de absorção de água, IRA e resistência à compressão para blocos de vedação e bloco portante

Para os dados de blocos de vedação e bloco portante (Figura 85) serão aplicadas as mesmas técnicas descritas no item anterior.

	1 Absorção de água	2 IRA	3 Resist. a compressão
Agudo	11,38	37,89	9,83
Cachoeira	13,30	37,25	8,28
Cachoeira	12,36	19,04	11,22
Capão Cipó	19,66	35,00	16,57
DonaFrancisca	14,12	58,78	10,58
Formigueiro	10,31	40,81	15,18
Itaara	24,24	50,36	8,84
Jaguari	11,04	53,15	10,47
Jarí	25,40	35,31	11,20
Jari	23,61	53,54	15,82
J. de Castilhos	17,54	31,46	3,23
J. de Castilhos	16,19	39,86	6,04
Mata	13,11	93,66	8,66
Nova Palma	17,42	39,52	16,43
Novo Cabrais	11,88	49,74	8,94
Restinga Seca	17,13	24,41	10,77
Santiago	19,98	47,47	18,68
Santiago	20,52	32,56	13,31
Santiago	20,83	96,61	12,40
São Francisco	14,01	54,87	8,76
São Francisco	10,33	43,44	9,33
São João Polês	14,48	62,89	11,40
São João Polês	12,00	67,99	9,95
São Martinho	21,68	32,24	8,03
São P. do Sul	15,96	60,17	10,74
São Sepé	15,65	41,56	4,97
São Sepé	14,50	25,22	10,97
V. Nova Sul	18,04	37,67	9,98

Figura 85 - Dados de absorção de água, IRA e resistência à compressão para blocos

Como aqui as variáveis também se apresentam em unidades de medidas diferentes, realizou-se a padronização das mesmas conforme já descrito (Figura 86).

	1 Absorção de água	2 IRA	3 Resist. a compressão
Agudo	-1,13	-0,47	-0,26
Cachoeira	-0,69	-0,51	-0,70
Cachoeira	-0,91	-1,51	0,14
Capão Cipó	0,77	-0,63	1,67
DonaFrancisca	-0,50	0,67	-0,04
Formigueiro	-1,38	-0,31	1,27
Itaara	1,82	0,21	-0,54
Jaguari	-1,21	0,36	-0,08
Jari	2,09	-0,61	0,13
Jari	1,68	0,39	1,46
J. de Castilhos	0,28	-0,83	-2,15
J. de Castilhos	-0,03	-0,37	-1,34
Mata	-0,74	2,59	-0,59
Nova Palma	0,26	-0,38	1,63
Novo Cabrais	-1,02	0,18	-0,51
Restinga Seca	0,19	-1,21	0,01
Santiago	0,84	0,05	2,27
Santiago	0,97	-0,77	0,74
Santiago	1,04	2,75	0,48
São Francisco	-0,53	0,46	-0,57
São Francisco	-1,37	-0,17	-0,40
São João Polês	-0,42	0,90	0,19
São João Polês	-0,99	1,18	-0,22
São Martinho	1,23	-0,78	-0,77
São P. do Sul	-0,08	0,75	0,00
São Sepé	-0,15	-0,27	-1,65
São Sepé	-0,42	-1,17	0,07
V. Nova Sul	0,40	-0,49	-0,22

Figura 86 - Dados padronizados para blocos

Selecionando-se a opção  Vertical icicle plot obtém-se o dendograma da Figura (87) a seguir.

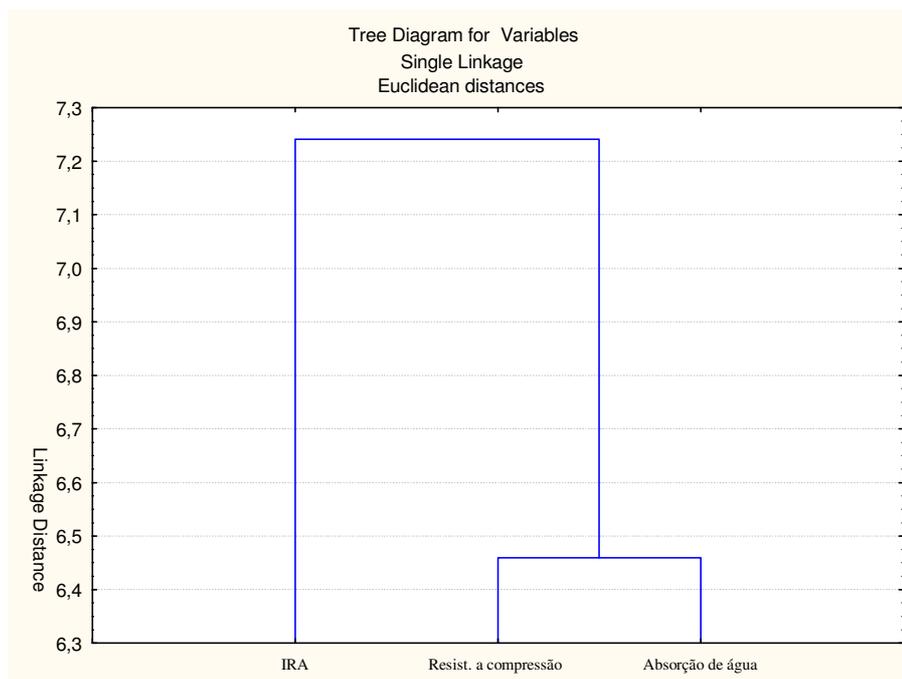


Figura 87 - Dendrograma vertical dos dados de absorção de água, IRA e resistência à compressão de blocos

No dendrograma elaborado verifica-se a formação de 2 grupos descritos da seguinte forma:

- para o dendrograma horizontal de cima para baixo: um grupo composto pela variável Absorção e Resistência à compressão e outro grupo composto pela variável IRA.

Para o dendrograma vertical observa-se a formação dos mesmos grupos, da direita para a esquerda.

Selecionado-se a opção  Amalgamation schedule obtém-se a Figura (88):

Amalgamation Schedule Single Linkage Euclidean distances			
linkage distance	Obj. No. 1	Obj. No. 2	Obj. No. 3
6,459263	Absorção de água	Resist. a compressão	
7,240633	Absorção de água	Resist. a compressão	IRA

Figura 88 - Ligação entre as variáveis de absorção de água, IRA e resistência à compressão para blocos

Selecionado-se a opção  Graph of amalgamation schedule obtém-se a Figura (89):

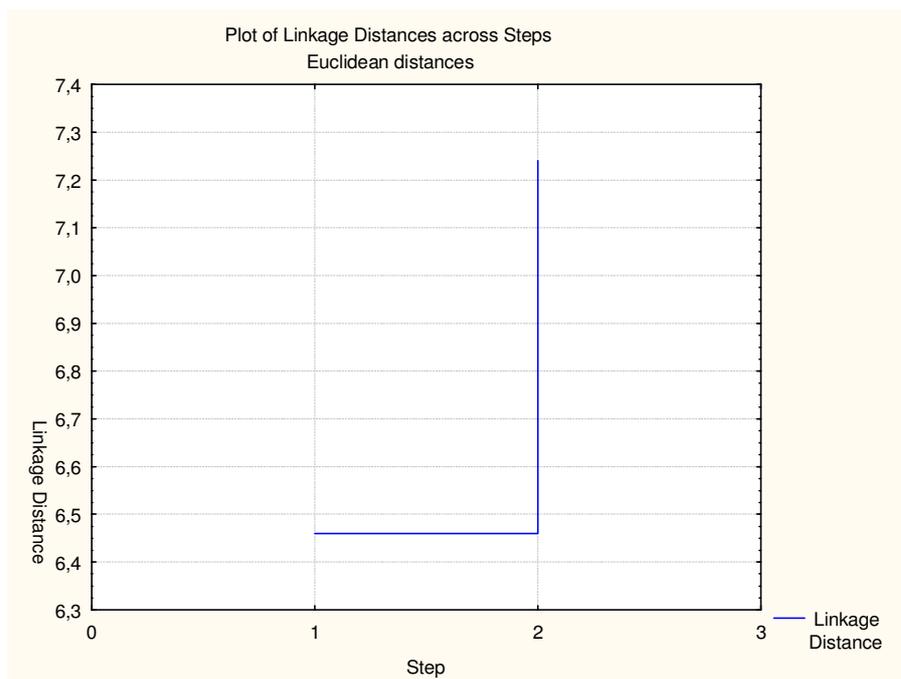


Figura 89 - Ligação das variáveis de absorção de água, IRA e resistência à compressão para blocos

Analisando-se o gráfico anterior verifica-se que o salto ocorre entre as distâncias 6,46 e 7,24. Entre essas distâncias pode ser feito um corte no gráfico onde observa-se a formação dos grupos distintos, conforme o histograma apresenta.

Selecionado-se a opção  Distance matrix obtém-se a matriz de distâncias entre as variáveis (Figura 90).

Variable	Euclidean distances		
	Absorção de água	IRA	Resist. a compressão
Absorção de água	0,00	7,56	6,46
IRA	7,56	0,00	7,24
Resist. a compressão	6,46	7,24	0,00

Figura 90 - Matriz de distâncias para absorção de água, IRA e resistência à compressão de blocos

Comparando-se os resultados obtidos, verifica-se que estes são bastante semelhantes tanto em disposição gráfica quanto em valores. Para os agrupamentos das Figuras 78 e 87, observa-se que:

- para os tijolos maciços, um grupo é formado pelas variáveis absorção de água e IRA e outro grupo pela variável resistência à compressão;
- para os blocos, um grupo é formado pelas variáveis absorção de água e resistência à compressão e outro grupo pela variável IRA.

Aqui também serão utilizados o círculo de correlação unitário (Figura 91) e os planos principais (Figuras 92 e 93) como complemento para um maior esclarecimento das análises realizadas.

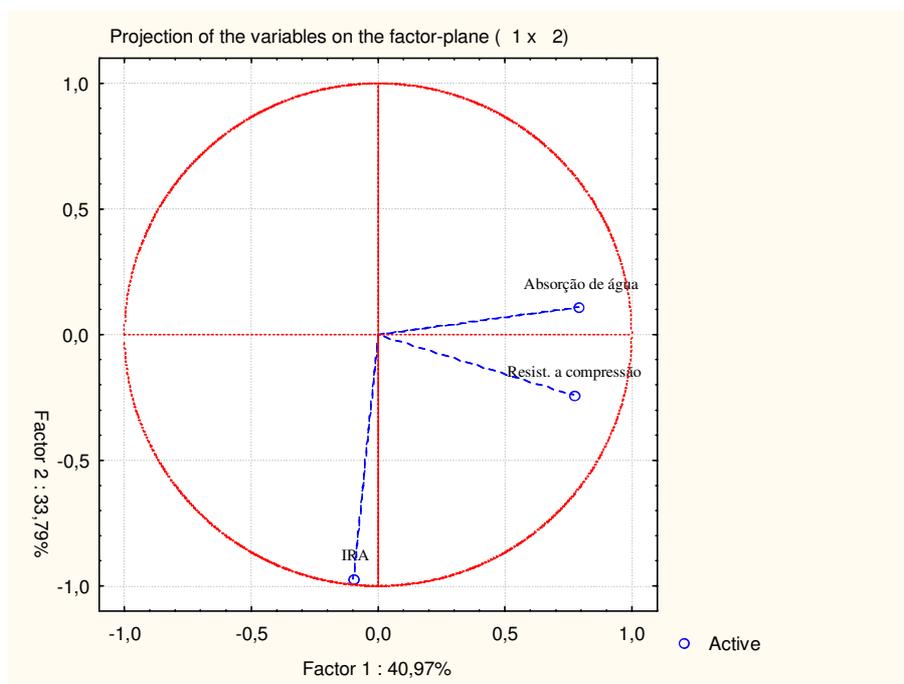


Figura 91 - Círculo de correlação unitário para absorção de água, IRA e resistência à compressão de blocos

Analisando-se a figura anterior observa-se que a variável IRA está localizada no 3º quadrante e sobre o círculo de correlação, indicando uma maior contribuição. A variável absorção de água está localizada no 1º quadrante próxima a variável IRA, localizada no 4º quadrante. Observa-se que essas duas variáveis encontram-se mais afastadas do eixo significando uma menor contribuição. Nota-se que os agrupamentos são semelhantes aos formados na análise de *cluster*.

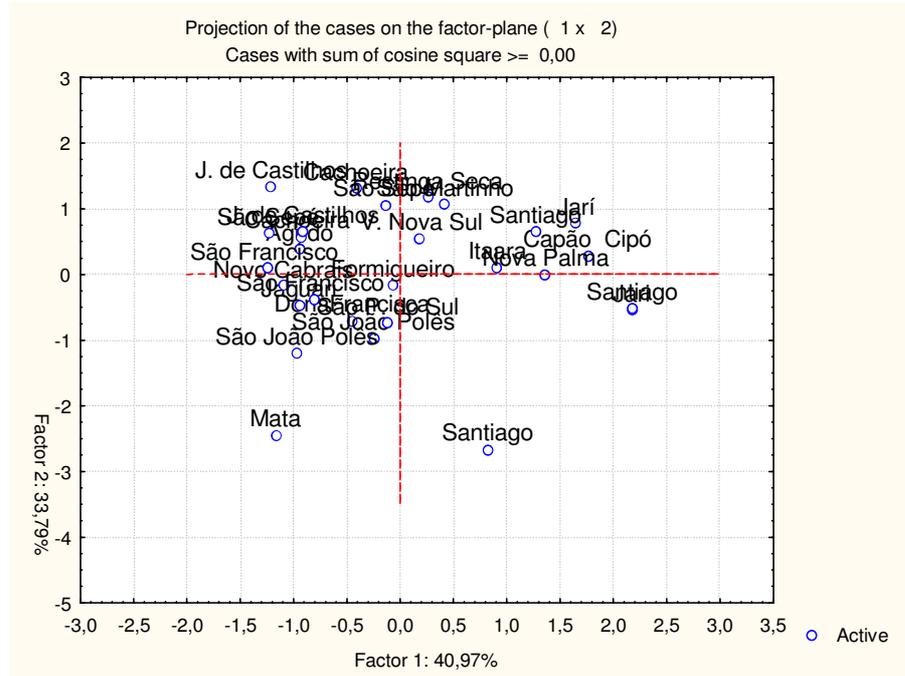


Figura 92 - Distribuição da nuvem de pontos por municípios, para blocos

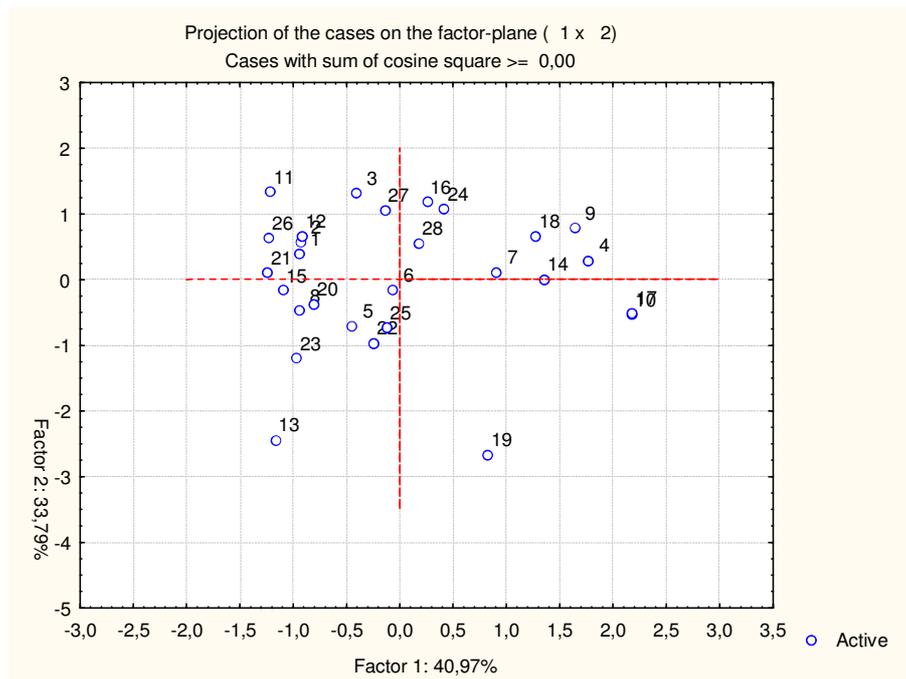


Figura 93 - Distribuição da nuvem de pontos por casos, para blocos

Analisando-se as figuras anteriores referentes aos blocos, através de uma sobreposição observa-se que o caso que está representando a variável IRA é Mata (caso 13). Igualmente, o caso de Santiago (caso 10) está representando a variável resistência à compressão. Os demais casos encontram-se sobrepostos e próximos a origem, não apresentando representatividade significativa em relação a estes dois fatores.

A partir da análise do conjunto de resultados obtidos, serão extraídas as conclusões possíveis bem como as sugestões para a realização de trabalhos futuros.

6. CONCLUSÕES

No presente trabalho procedeu-se a aplicação de técnicas de análise multivariada ao conjunto de dados cerâmicos selecionado.

A partir dos resultados obtidos e das discussões realizadas, obtiveram-se as conclusões descritas a seguir.

A técnica de análise de agrupamentos (*cluster*) com o objetivo de identificar as variáveis pertencentes ao mesmo grupo, ou seja, quais variáveis apresentavam maiores similaridades.

A técnica de análise fatorial foi utilizada na tentativa de reduzir um grande número de variáveis para um número mais significativo, representado pelos fatores, onde se pode verificar que os agrupamentos formados pela análise foram semelhantes aos formados na análise de agrupamentos.

A técnica de análise de componentes principais demonstrou ser eficiente na identificação das variáveis que apresentavam uma maior contribuição na formação dos fatores.

A partir dos resultados obtidos e discussões realizadas, obtiveram-se as conclusões descritas a seguir.

Para os dados de produção mensal total de cerâmica vermelha, número de indústrias cerâmicas, tijolos maciços, blocos de vedação, bloco estrutural, telhas e tabelas concluiu-se que:

- na análise de agrupamentos a partir do dendograma obtido observou-se a formação de três grupos distintos: um 1º grupo formado pelas variáveis número de indústrias cerâmicas, bloco estrutural e tabelas; um 2º grupo formado pelas variáveis tijolos maciços e telhas e um 3º grupo formado pela variável bloco de vedação. Cabe ressaltar que esses grupos foram formados pelas similaridades existentes entre as variáveis. Também se observou que os valores de desvio padrão eram elevados nos casos de tijolos maciços, blocos de vedação e telhas significando uma grande variabilidade entre os resultados. Optou-se pela retirada da variável produção mensal total devido ao valor do desvio padrão ser muito elevado e por ser uma variável que representava a soma de todas as demais.

- Na análise fatorial verificou-se que somente o 1º autovalor era maior do que 1 e explicava 90,82% da variância total. Assim, os dados seriam resumidos pela 1ª componente principal.
- Os resultados dos *factor loadings* para o fator 1 foram significativos. No caso dos demais fatores os valores não foram significativos.
- No caso dos planos fatoriais, as variáveis que melhor explicaram o fator 1 foram blocos de vedação, bloco estrutural e telhas, e a variável que melhor explicou o fator 2 foi o número de indústrias cerâmicas.
- Após a rotação *varimax normalized*, para melhoraria da visualização, as variáveis que melhor explicaram o fator 1 foram blocos de vedação, bloco estrutural, telhas e tabelas, e a variável que melhor explicou o fator 2 foi o número de indústrias cerâmicas, resultando em praticamente as mesmas variáveis determinadas anteriormente.
- Como foram consideradas 6 variáveis, foram determinados 6 componentes principais, mas segundo a análise fatorial, somente a 1ª componente principal deveria ser considerada.
- Analisando-se os resultados da matriz de correlação entre as variáveis, observou-se que todos os valores eram superiores a 0,7 indicando uma correlação de moderada a forte. Logo, concluiu-se que existia uma interligação entre as variáveis e que a análise das mesmas deveria ser realizada conjuntamente e não de forma isolada.
- A sobreposição do círculo de correlação unitário sobre o 1º plano principal foi utilizado como complemento para um melhor esclarecimento das análises realizadas. Constatou-se que o caso que mais contribuiu na formação da combinação linear do fator 1 foi o de Santa Maria e o que mais contribuiu na formação da combinação linear do fator 2 foi o de Jari. Os demais casos não apresentaram representatividade significativa em relação a esses 2 fatores.
- A matriz de correlação entre as variáveis originais e os componentes principais também possibilitou determinar a contribuição de cada variável em relação aos fatores formados, ou seja, o quanto cada variável contribuiu na formação de cada fator, sendo os mais relevantes aqueles com os maiores valores absolutos. Os resultados encontrados foram similares aos determinados anteriormente.

Para as dimensões médias e características geométricas dos tijolos maciços concluiu-se que:

- na análise de agrupamentos a partir do dendograma obtido verificou-se a formação de dois grupos distintos: um 1º grupo formado pelas variáveis L, H, Desvio esquadro e Planeza faces e um 2º grupo formado pela variável C. Cabe ressaltar que esses grupos foram formados pelas

similaridades existentes entre as variáveis. Também se observou que os valores de desvio padrão eram pequenos significando uma pequena variabilidade entre os resultados.

- Na análise fatorial verificou-se que o 1º e o 2º autovalores eram maiores do que 1 e explicavam 64,22% da variância total. Portanto, os dados seriam resumidos pela 1ª e 2ª componente principal.
- Os resultados dos *factor loadings* para o fator 1 demonstram ser significativos para as variáveis H e C e para o 2º fator, a variável Planeza faces foi significativa. Os demais fatores não foram significativos.
- Para os planos fatoriais, as variáveis que melhor representaram o fator 1 foram H, C e L. Do mesmo modo, a variável que melhor explicou o fator 2 foi Planeza faces.
- Após a rotação *varimax normalized*, a variável que melhor representou o fator 1 H e a variável que melhor representou o fator 2 foi Planeza faces. Isso ficou confirmado pelos valores dos *factor loadings* rotacionados.
- Como existiam 5 variáveis, foram determinados 5 componentes principais, mas segundo a análise fatorial, somente a 1ª e a 2ª componente principal deveriam ser consideradas.
- Analisando-se os resultados da matriz de correlação entre as variáveis, observou-se que todos os valores eram inferiores a 0,7, indicando uma correlação fraca entre as variáveis. Assim, a realização de uma análise fatorial seria insatisfatória mas a técnica foi aplicada como forma de exemplificação.
- Na sobreposição do círculo de correlação unitário sobre o 1º plano principal constatou-se que os casos mais representativos para a variável Desvio esquadro foram Mata e Jari; para a variável Planeza faces foi Agudo; para a variável C foi São João do Polêsine; para a variável L foi Cachoeira do Sul ou ainda considerando-se a olaria especificada segundo as Figuras 50 e 51. Também se observou uma maior concentração de olarias em torno da origem, não apresentando representatividade significativa em relação aos 2 fatores, e ainda algumas sobreposições das mesmas.
- A matriz de correlação entre as variáveis originais e os componentes principais foi outra forma de encontrar a contribuição de cada variável em relação aos fatores formados. Comparando-se os resultados obtidos com os *factor loadings* obtidos anteriormente, observou-se que os valores eram iguais, com um maior número de correlações significativas no caso da matriz de correlação.

Quanto às dimensões médias para blocos de vedação e bloco portante, foram observadas as seguintes conclusões:

- A partir do dendograma obtido pela análise de agrupamentos verificou-se a formação de dois grupos distintos: um 1º grupo formado pelas variáveis L, H e C e um 2º grupo formado pelas variáveis Ev, Eh, Iv e Ih. Esses grupos foram formados pelas similaridades existentes entre as variáveis.
- Observou-se que os valores de desvio padrão para as variáveis Eh, Ev, Ih e Iv eram pequenos, em torno de 1, significando uma pequena variabilidade entre os resultados. Já os valores de desvio padrão para as variáveis L, H e C eram bem mais elevados, o que representava uma maior variabilidade entre os resultados.
- Na análise fatorial verificou-se que o 1º e o 2º autovalores eram maiores do que 1 e explicavam 74,14% da variância total. Portanto, os dados seriam resumidos pela 1ª e 2ª componente principal.
- Os resultados dos *factor loadings* para o fator 1 demonstram ser significativos para as variáveis Ih, Eh e Ev e para o 2º fator, a variável H foi significativa. Os demais fatores não foram significativos.
- Para os planos fatoriais, as variáveis que melhor representaram o fator 1 foram Ih, Eh e Ev. Do mesmo modo, a variável que melhor explicou o fator 2 foi H.
- Após a rotação *varimax normalized*, a variável que melhor representou o fator 1 foi Iv e a variável que melhor representou o fator 2 foi L e C. Isso ficou confirmado pelos valores dos *factor loadings* rotacionados.
- Como existiam 5 variáveis, foram determinados 7 componentes principais, mas segundo a análise fatorial, somente a 1ª e a 2ª componente principal deveriam ser consideradas.
- Analisando-se os resultados da matriz de correlação entre as variáveis, observou-se que existiam valores superiores a 0,7, indicando uma correlação de moderada a forte entre as variáveis. Assim, concluiu-se que existia uma interligação entre as variáveis e que a análise das mesmas deveria ser realizada conjuntamente e não de forma isolada.
- Na sobreposição do círculo de correlação unitário sobre o 1º plano principal constatou-se que a maior parte dos casos estavam concentrados em torno da origem, não apresentando representatividade significativa em relação aos 2 fatores, e ainda, observou-se a ocorrência de várias sobreposições. O único caso que se destacou foi o de Cachoeira do Sul como sendo o caso mais representativo para a variável H ou considerando-se a olaria especificada segundo as Figuras 72 e 73.
- A matriz de correlação entre as variáveis originais e os componentes principais foi outra forma de encontrar a contribuição de cada variável em relação aos fatores formados. Comparando-se os resultados obtidos com os *factor loadings* obtidos anteriormente,

observou-se que os valores eram iguais, com um maior número de correlações significativas no caso da matriz de correlação.

Quanto aos dados referentes à absorção de água, IRA e resistência à compressão para tijolos maciços, blocos de vedação e bloco portante foram aplicadas as técnicas de análise de agrupamento, círculo de correlação unitário e os planos fatoriais, após a realização da padronização das mesmas. Foram observadas as seguintes conclusões:

- no caso dos tijolos maciços, observou-se um dendograma composto por 2 grupos: um 1º grupo formado pelas variáveis absorção de água e IRA e um 2º grupo composto pela variável resistência à compressão. Numa sobreposição do círculo de correlação unitário sobre os planos fatoriais, observou-se que o caso representativo para a variável resistência à compressão era Cachoeira do Sul. Da mesma forma, as variáveis absorção de água e IRA estavam representadas pelos casos de Santiago e São Martinho. Os demais casos estavam sobrepostos e próximos à origem, não apresentando representatividade significativa.

- No caso de blocos de vedação e bloco portante, observou-se um dendograma composto por 2 grupos: um 1º grupo formado pelas variáveis absorção de água e resistência à compressão e um 2º grupo composto pela variável IRA. Numa sobreposição do círculo de correlação sobre os planos fatoriais, observou-se que o caso representativo para a variável IRA era Mata. Igualmente, a variável resistência à compressão estava representada pelo caso de Santiago. A variável absorção de água estava representada por diversos casos, entre os quais se destacavam Capão Cipó, Jari e Santiago. Nos demais casos, existiam sobreposições e proximidades à origem, não apresentando representatividade significativa.

Com isso, o conjunto de dados utilizado neste trabalho demonstrou ser adequado para a realização das técnicas propostas.

Assim, analisando-se os resultados obtidos conclui-se que as técnicas multivariadas descritas neste trabalho podem ser aplicadas em dados de produtos cerâmicos, pois os resultados obtidos confirmaram resultados e conclusões obtidas em trabalhos realizados anteriormente. Cabe ressaltar que o mais importante foi demonstrar o comportamento de dados já existentes frente às técnicas estatísticas multivariadas apresentadas.

A partir da metodologia desenvolvida neste trabalho e das conclusões alcançadas, fazem-se sugestões de trabalhos futuros tais como:

- a aplicação de outras técnicas de análise multivariada disponíveis ao conjunto de dados utilizado neste trabalho;

- a aplicação das técnicas, empregadas no presente trabalho, a outros conjuntos de dados disponíveis; e

- a aplicação das técnicas de análise multivariada para análise de outros materiais de Engenharia Civil assim como nas demais áreas da Engenharia Civil.

7 BIBLIOGRAFIA

Anuário Brasileiro de Cerâmica. Disponível em: <http://www.abceram.org.br/asp/abc_0.asp>. Acesso em: 25 fev. 2007.

ARAÚJO, R. C. L.; RODRIGUES, E. H. V.; FREITAS, E. G. A. **Produtos cerâmicos**. Disponível em: <<http://www.ufrj.br/institutos/it/dau/.../Materiais%20cer%20micos.pdf>> Acesso em: 05 ago. 2008.

Associação Brasileira de Cerâmica. Disponível em: <http://www.abceram.org.br/asp/abc_0.asp>. Acesso em: 31 ago. 2008.

BUSSAB, W. O.; MIAZAKI, E. S.; ANDRADE, D. F. Introdução à análise de agrupamentos. In: Simpósio Brasileiro de Probabilidade e Estatística. 9., 1990. São Paulo: Resumos:...São Paulo, 1990.

BUSTAMANTE, G. M. & BRESSIANE, J. C. A Indústria Cerâmica Brasileira. **Revista Cerâmica Industrial**. São Paulo, v.5, n.3, mai/jun. 2000.

CATTEL, R. B. The scree test for the number of factors. In:..., **Multivariate behavior research**. V. 1, p. 245-276. 1966.

CAZAR, R. A. An Exercise on Chemometrics for a Quantitative Analysis Course. **Journal of Chemical Education**, Madison, WI: v. 80, n. 9. 2003. p. 1026 - 1029.

CHRISTOFOLETTI, S. R. *et al.* Utilização da Análise Estatística Multivariada no Tratamento de Dados Aplicados a Matérias-Primas Cerâmicas. **Revista do Instituto Geológico**, 2005. São Paulo, 26 (1/2), p. 19-29.

Confederação Nacional da Indústria. Disponível em: <<http://www.cni.org.br/portal/data/pages/FF80808121B517F40121B54C2EAA4940.htm>> Acesso em: 25 fev. 2007.

CORRAR, J. L. **Metodologia de identificação dos principais indicadores financeiros para acompanhamento e análise de empresas**: uma aplicação em companhias seguradoras. Encontro Anual da ANPAD – Associação Nacional de Programas em Pós-Graduação em Administração, 2002.

DAVIS, J. C. **Statistics and data analysis in Geology**. 2. ed. John Wiley and Sons, Inc. 1986.

- DIAS, E. G.; SANTOS M. C.; GOTO, M. M. **Reflexos da aplicação da resolução Conama nº 369/2006 na mineração de agregados.** Disponível em: <<http://www.ibram.org.br/sites/700/784/00000946.pdf>>. Acesso em: 31 jul. 2008.
- FERREIRA, D. F. **Análise multivariada.** Lavras, 1996. Disponível em: <<http://www.dex.ufla.br/~danielff/dex522.pdf>>. Acesso em: 27 nov. 2007.
- KAZMIERCZAK, C. S. Produtos de Cerâmica Vermelha. In: ISAIA, G. C. (Org.). **MATERIAIS DE CONSTRUÇÃO CIVIL e Princípios de Ciência e Engenharia de Materiais.** São Paulo: IBRACON, 2007. v.1. p. 563-586.
- JACKSON, J. E. Principal components and factor analysis: Part I – principal components. **Journal of Quality Technology.** v.12, n.4, p. 201-213, Oct. 19
- JOHNSON, R. A; WICHERN, D. W. **Applied multivariate statistical analysis.** 4. ed. New Jersey: Prentice-Hall, 1998.
- KASSAI, S.; KASSAI, J. R. **Desvendando o termômetro de insolvência de Kanitz.** Encontro Nacional da ANPAD – Associação Nacional de Programas em Pós-Graduação em Administração, 2003.
- KENDALL, M.G. **Discrimination and classification.** In: Krishnaian (Ed.), Proc. Int. Symp. on Multivariate Analysis, New York: Academic Press, 1969. p. 165-185.
- KIPPER, A. C. **Cerâmica é opção natural.** Disponível em: <<http://www.siocergs.org.br/martin/?page=shop/template&t=noticia¬icia=25>>. Acesso em: 30 set. 2008.
- KRUMBEIN, W. C.; GRAYBILL. F. A. An introduction to Statistical Model in Geology. Data: Elsevier. 1982.
- LANDIM, P. M. B. **Análise Estatística de Dados Geológicos Multivariados.** Universidade Estadual Paulista /UNESP: Campus de Rio Claro. Departamento de Geologia Aplicada, 2000. Disponível em: < <http://www.cfh.ufsc.br/gcn3506/documents/LandimMultivariada2002.pdf>>. Acesso em: 05 jul. 2008.
- MAFRA, A. T. **Proposta de indicadores de desempenho para indústria de cerâmica vermelha do estado de Santa Catarina.** 1999. 126 p. Dissertação (Mestrado em Engenharia de Produção) - Universidade Federal de Santa Catarina, Florianópolis, 1999.
- MALHOTRA, N. K. **Pesquisa de marketing: uma orientação aplicada.** Porto Alegre: Bookman, 2001.
- MANLY, B. F. J. **Multivariate statistical methods: a primer.** London: Chapman and Hall, 3. Ed. 2004.
- MARDIA, K. V. **Multivariate analysis.** London: Academic, 1979.

- MONTEIRO, R. C.; BERNARDES, E. S.; MASSON, M. R.. Análise estatística multivariada para materiais cerâmicos. **Geociências**, 19(1): 2000. p. 133-140.
- MORRISON, D. F. **Multivariate statistical methods**. New Delhi: McGraw-Hill. 1976.
- NETO, M. M. J. Estatística multivariada. **Revista de Filosofia e ensino**. 2004. Disponível em: <http://www.criticanarede.com/cien_estatistica.html>. Acesso em: 09 mai. 2008.
- OLIVEIRA, S. M. **Avaliação dos blocos e tijolos cerâmicos do Estado de Santa Catarina**. 1993. 132 p. Dissertação (Mestrado em Engenharia Civil) - Universidade Federal de Santa Catarina, Florianópolis, 1993.
- PAULETTI, M. C. **Modelo para introdução de nova tecnologia em agrupamentos de micro e pequenas empresas: estudo de caso das indústrias de cerâmica vermelha no Vale do rio Tijucas**. 2001. 168 p. Dissertação (Mestrado em Engenharia de Produção) – Universidade Federal de Santa Catarina. Florianópolis, 2001.
- PEREIRA, J. C. R. **Análise de dados qualitativos: estratégias metodológicas para as ciências da saúde, humanas e sociais**. São Paulo: Edusp, 2001.
- Pereira, A. F. O. A.; Júnior, C. P.; Ramos, E. J. S. Modelo e análise de previsão de desempenho pela metodologia de análise multivariada de dados: um estudo empírico do setor de energia elétrica. **Revista Contemporânea de Contabilidade**, ISSN 1807-1821, v. 3, n. 5, 2006 , p. 59-74.
- PLA, L. E. **Analysis multivariado: Método de componentes principais**. Washington: Secretaria General de La Organización de Los Estados Americanos, D. C. 1986.
- PONTES, A. C. F. **Análise de variância multivariada com a utilização de testes não-paramétricos e componentes principais baseados em matrizes de postos**. 2005. 106 p. Tese (Doutorado em Agronomia) – Escola Superior de Agricultura Luiz de Queiroz, Piracicaba, 2005.
- REGAZZI, A. J. **INF 766 – Análise multivariada**. Viçosa: Universidade Federal de Viçosa, Centro de Ciências Exatas e Tecnológicas. Departamento de Informática, 2001. 166 p. Apostila de disciplina.
- REIS, E. **Estatística multivariada aplicada**. Lisboa, 1997.
- SECTME. **Diagnóstico do setor de cerâmica vermelha em Santa Catarina**. Florianópolis: Secretaria de Estado de Ciência e Tecnologia, das Minas e Energia de Santa Catarina, 1990.
- Serviço Nacional de Aprendizagem Industrial. Disponível em: <<http://www.senai.br/br/home/index.aspx>>. Acesso em: 10 jul. 2008.
- SOUTO, M. C. P. **Análise de Agrupamentos (Clusters)**. Disponível em: <www.dimap.ufrn.br/~marcilio/AM/analise-de-agrupamentos.ppt>. Acesso em: 25 set. 2008.

SOUZA, A. M. **Monitoração e ajuste de realimentação em processos produtivos multivariados**. 2000. 166 p. Tese (Doutorado em Engenharia de Produção) – Universidade Federal de Santa Catarina, Florianópolis, 2000.

SOUZA, A. M.; VICINI, L. **Análise multivariada**: da teoria à prática. Caderno didático. Santa Maria: UFSM, 2005. 215 p.

STEIL, M. C. Fabricação de produtos cerâmicos. In: 44º Congresso Brasileiro de *Cerâmica*, 2000. São Pedro. **Anais...** São Paulo: Associação Brasileira de Cerâmica, 2000. 1 CD.

TOMAZETTI, R. S. **Análise da produção de cerâmica vermelha da região central do estado do Rio Grande do Sul**. 2003. 190 p. Dissertação (Mestrado em Engenharia Civil) – Universidade Federal de Santa Maria, Santa Maria, 2003.

VALENTIM, J. L. **Ecologia numérica**: uma introdução à análise multivariada de dados ecológicos. Rio de Janeiro: Interciência, 2000.

VERDINELLI, M. A. **Análise inercial em ecologia**. 1980. 162 p. Tese (Doutorado em Ciências Biológicas) - USP, São Paulo, 1980.

APÊNDICE A - INTRODUÇÃO AO *SOFTWARE STATISTICA 7.0* E TUTORIAL DE UTILIZAÇÃO

O *software Statistica* é muito utilizado para análises estatísticas devido ao grande número de módulos disponíveis bem como a facilidade na utilização do mesmo por ser compatível com ambiente *Windows*. Ressalta-se que talvez um inconveniente seja que o *software* é apresentado em idioma inglês.

A seguir são apresentados os módulos disponíveis no *Software Statistica 7.0*:

- Estatística Básica
- Distribuições Não-paramétricas
- Anova-Manova
- Regressão Múltipla
- Estimção Não linear
- Séries temporais-Previsões
- Análise de Cluster
- Gerenciamento de dados (*Data Management*)
- Análise de fatores
- Análise canônica
- Escala Multidimensional
- Análise Correspondente
- *SEPATH*
- Análise Discriminante
- Análise Log-linear
- Análise de sobrevivência
- Controle de Qualidade
- Análise de Processos
- Design experimental
- Linguagem de comandos

Na Figura (94a) e (94b) seguinte serão apresentadas às telas do *Statistica 7.0*, onde estão listados os módulos disponíveis citados anteriormente.

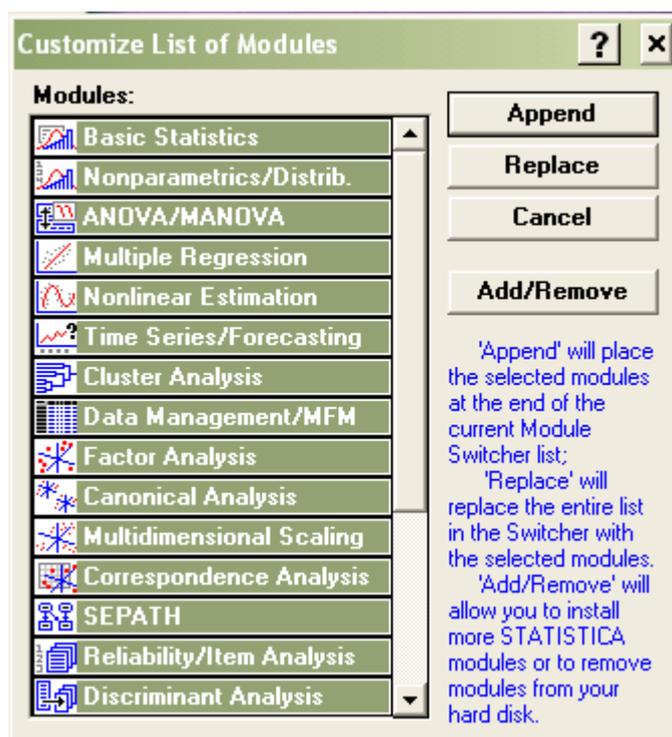


Figura 94a - Módulos disponíveis no *Statistica 7.0*

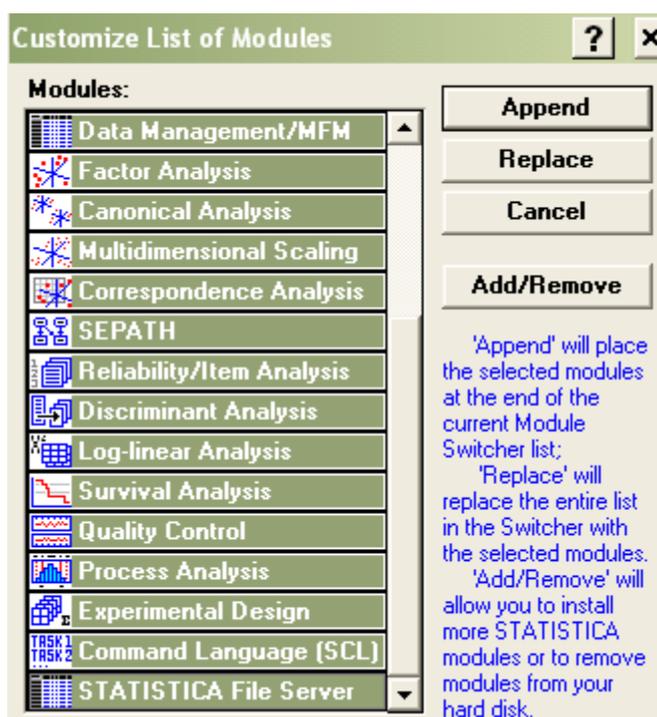


Figura 94b - Módulos disponíveis no *Statistica 7.0*

Assim, dependendo do tipo de dados disponíveis e das análises a serem desenvolvidas, escolhem-se os módulos apropriados para a realização das mesmas.

Como o objetivo deste trabalho é a apresentação de algumas técnicas de análise multivariada, proceder-se-á ao desenvolvimento das mesmas. Para isso, demonstrar-se-á

passo-a-passo as etapas para uma análise de agrupamentos (*cluster*), análise fatorial e análise de componentes principais.

Nas Figuras (95) e (96) seguintes são apresentadas as telas referentes à entrada dos dados e a seleção da análise estatística.

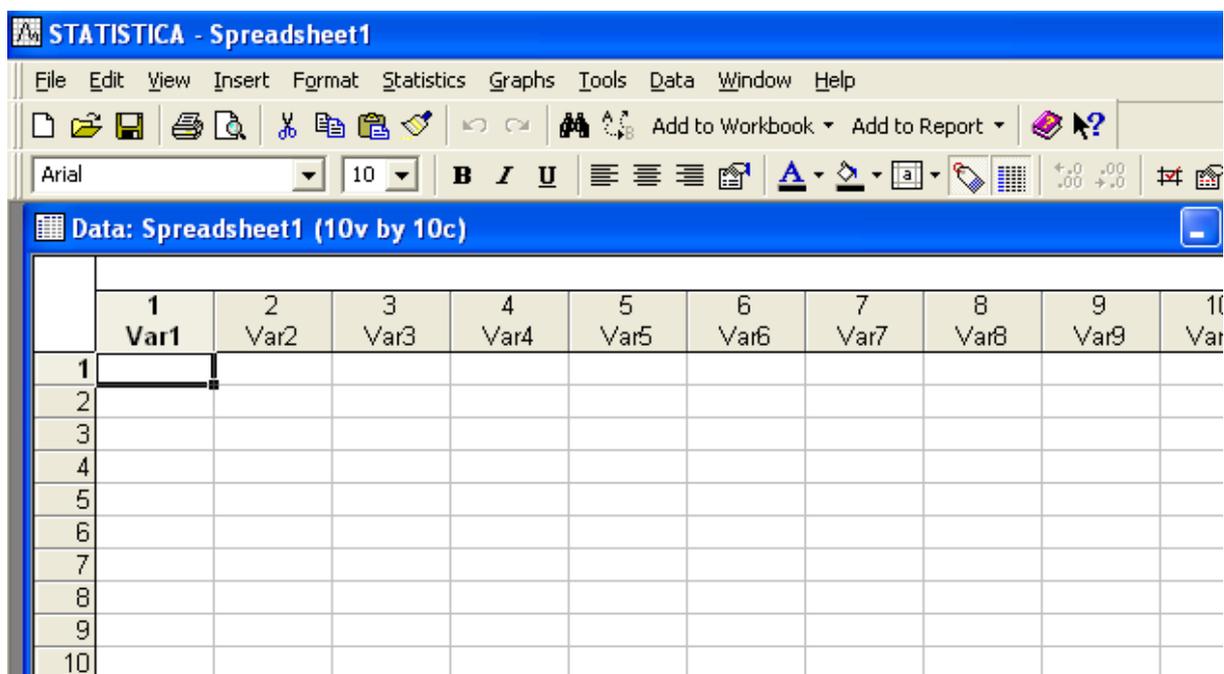


Figura 95 - Entrada dos dados

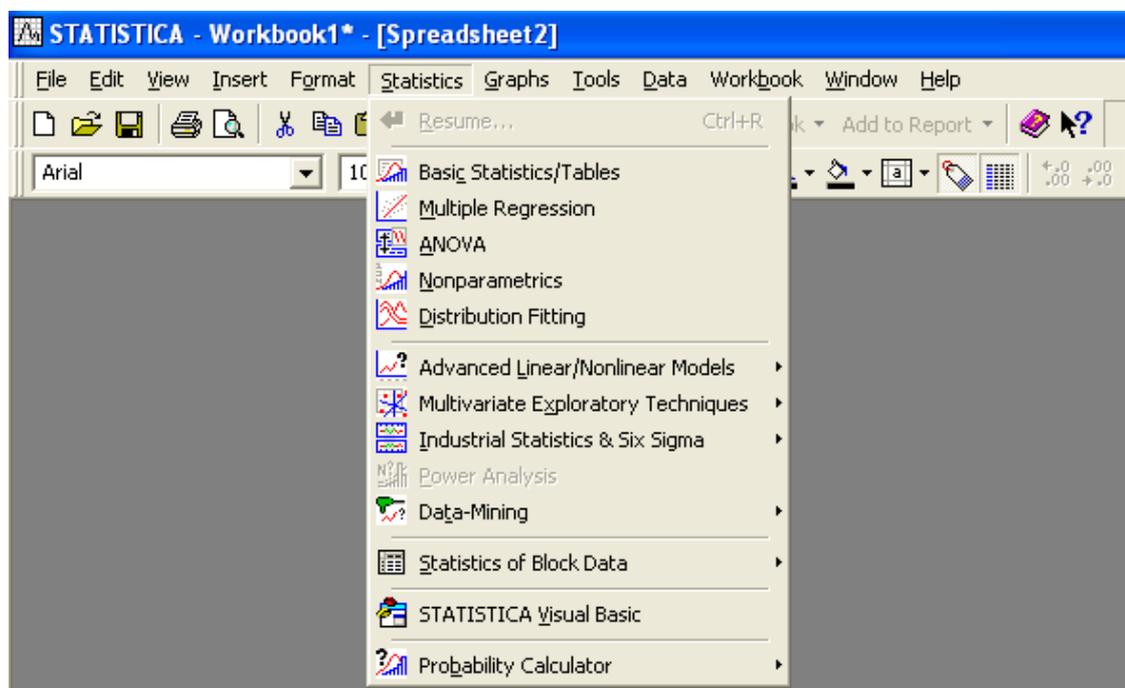


Figura 96 - Seleção da análise estatística

Inicialmente serão detalhados os procedimentos para a realização da análise de agrupamento (AA), empregando-se o método do vizinho mais próximo.

Para acessar o programa utiliza-se uma sequência: clica-se em Iniciar, localizam-se os programas e em seguida o *Statistica* ou seleciona-se o programa na tela inicial do computador, quando existir um ícone disponível, conforme a Figura (97).



Figura 97 - Acesso ao Software *Statistica*

Segundo SOUZA & VICINI (2005) pode-se importar banco de dados do Excel para o *Statistica* sem a necessidade de copiar as variáveis individualmente. Para isso, deve-se clicar na opção Abrir Arquivos do tipo: e selecionar a opção Excel *Files (.xls)* e clicar em Abrir, conforme a Figura (98).

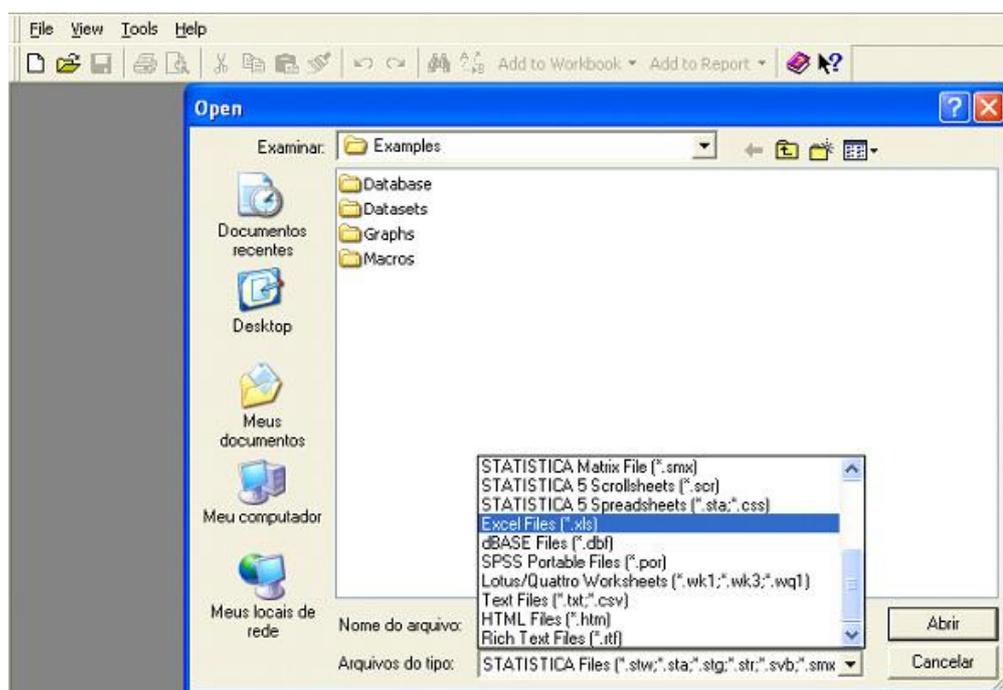


Figura 98 – Caixa de seleção para transportar o banco de dados do Excel para o *Statistica*

Em seguida surge a caixa de seleção das planilhas, conforme a Figura (99).

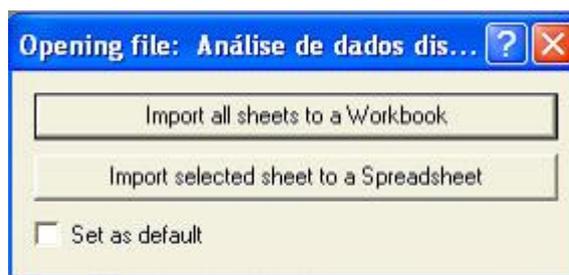


Figura 99 – Caixa de seleção para importar todas as planilhas ou as planilhas selecionadas para o *Statistica*

Na primeira opção todas as planilhas podem ser importadas para a área de trabalho. Na segunda opção, importam-se as planilhas selecionadas, conforme a Figura (100).

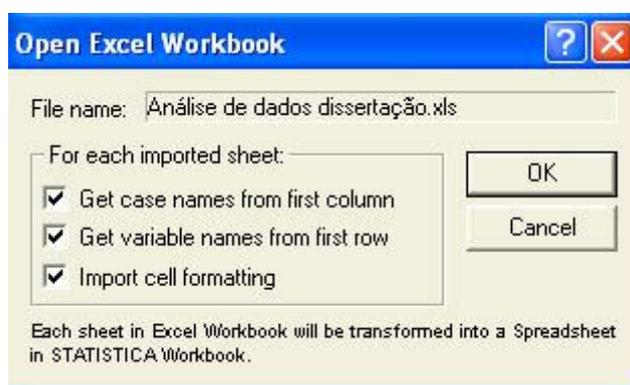


Figura 100 – Caixa de seleção para importar os dados do Excel para o *Statistica*, por linhas e por colunas

Na figura anterior selecionando-se:

- a primeira opção, serão importados os nomes da 1ª coluna que normalmente constituem variáveis qualitativas;
- a segunda opção, serão importados os nomes das variáveis que estão na 1ª linha da planilha;
- a terceira opção, serão importados os formatos dos dados.

Elaborou-se o banco de dados onde as variáveis estão representadas nas colunas e os objetos nas linhas.

Para realizar a análise, seleciona-se na barra superior a opção *Statistics* e a seguir seleciona-se a opção *Multivariate Exploratory Techniques*. Assim, surgem as técnicas multivariadas disponíveis conforme a Figura (101) a seguir.

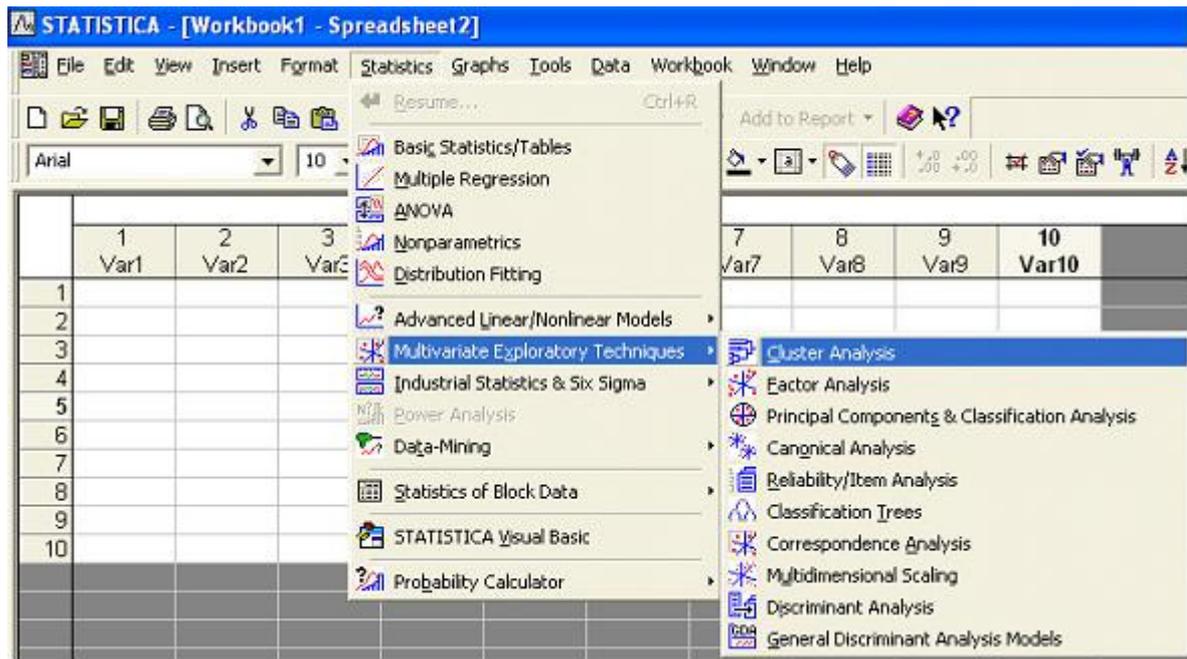


Figura 101 – Caixa de seleção das técnicas multivariadas

Selecionando-se *Cluster Analysis*, surge a tela da Figura (102):

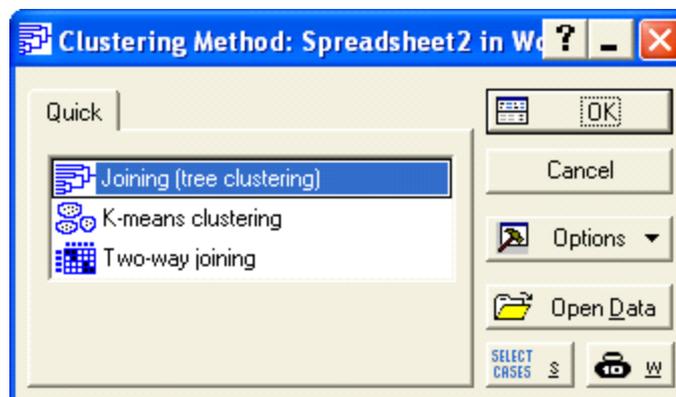


Figura 102 – Métodos de Cluster

Clicando-se na primeira opção *Joining (tree clustering)* e em seguida em Ok, surge a tela da Figura (103) com as opções para a análise de *cluster*. Na opção *Joining (tree clustering)* determina-se o dendograma que apresenta o número de grupos formados segundo as mesmas características. Na opção *K-means clustering* define-se o número de grupos a serem utilizados na análise definidos pelas médias determinadas a partir do banco de dados inicial. Pela opção *Two-way joining* é possível fazer um mapa associativo entre cada variável e a unidade amostral o que permite, através de uma inspeção visual, que variável apresenta maior representatividade no conjunto de dados analisado.

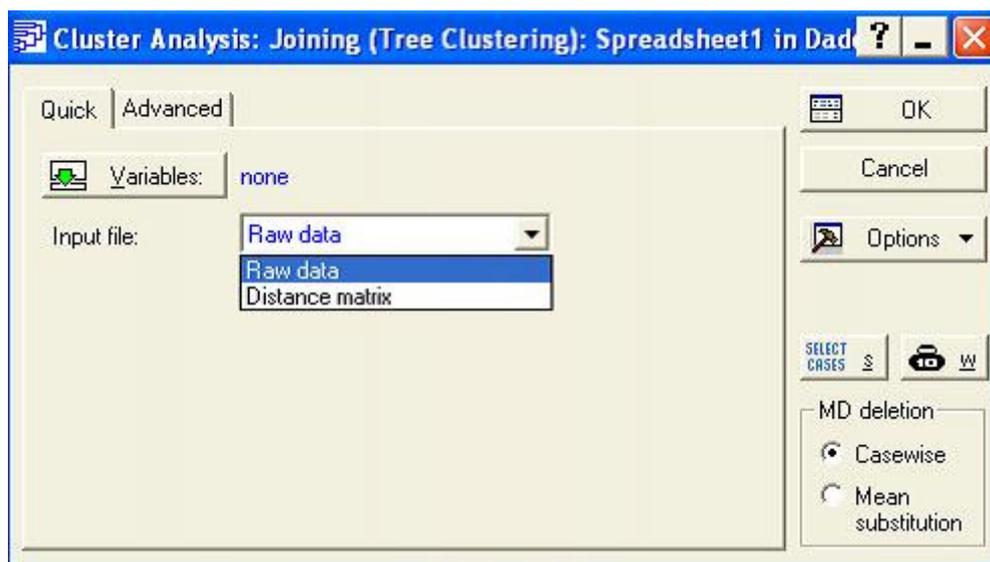


Figura 103 - Caixa de seleção para análise de agrupamento, selecionando-se *Quick*

Na opção *Variables*, selecionam-se as variáveis a serem incluídas na análise.

Na opção *Input file* encontram-se as opções *Raw-data* e *Distance matrix*. A opção *Raw data* utiliza os dados brutos do banco de dados.

Selecionando-se *Advanced* surgem as opções da Figura (104) a seguir.

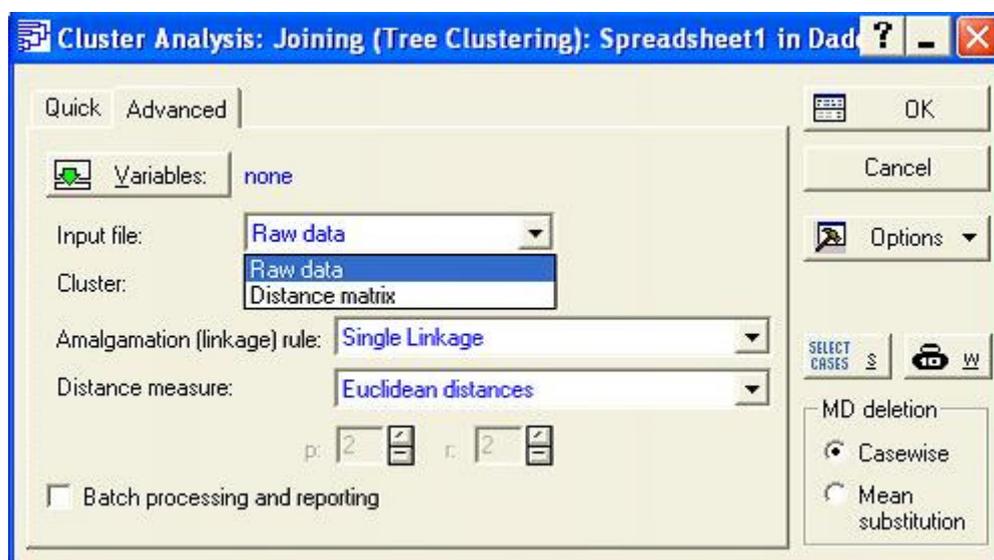


Figura 104 - Caixa de seleção para análise de agrupamento, selecionando-se *Advanced*, para a opção *Input file*

O item *Cluster* apresenta as opções *Variables* e *Cases*. Quando se seleciona *Variables*, o agrupamento é feito por colunas; quando se seleciona *Cases*, o agrupamento é feito por linhas, conforme a Figura (105).

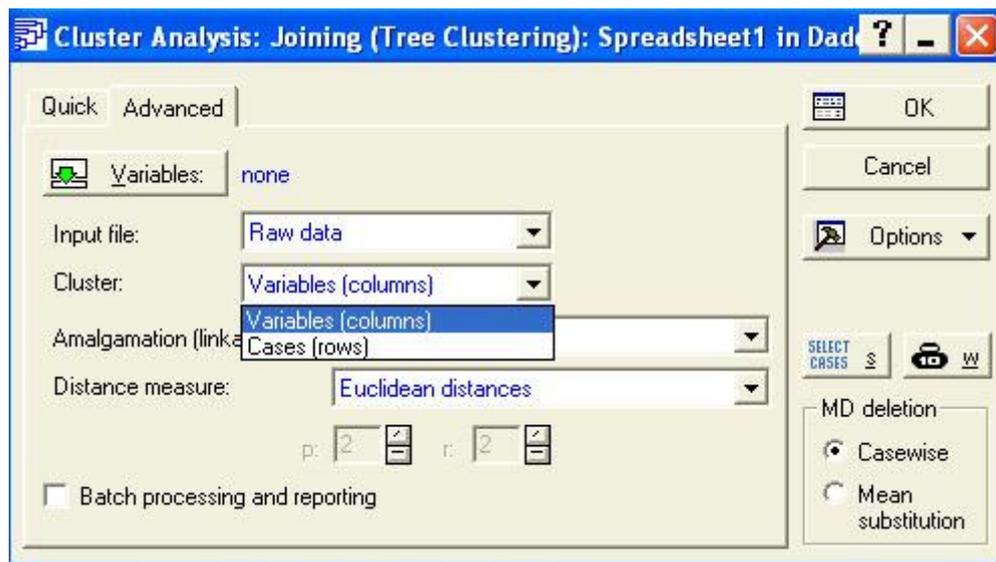


Figura 105 - Caixa de seleção para análise de agrupamento, selecionando-se *Advanced*, para a opção *Cluster*

O item *Amalgamation (linkage) rule*, Figura (106), contém as opções dos métodos de encadeamento:

- *Single Linkage*, que se fundamenta na distância mínima entre os objetos, e
- *Complete Linkage*, que se fundamenta na distância máxima entre os objetos.

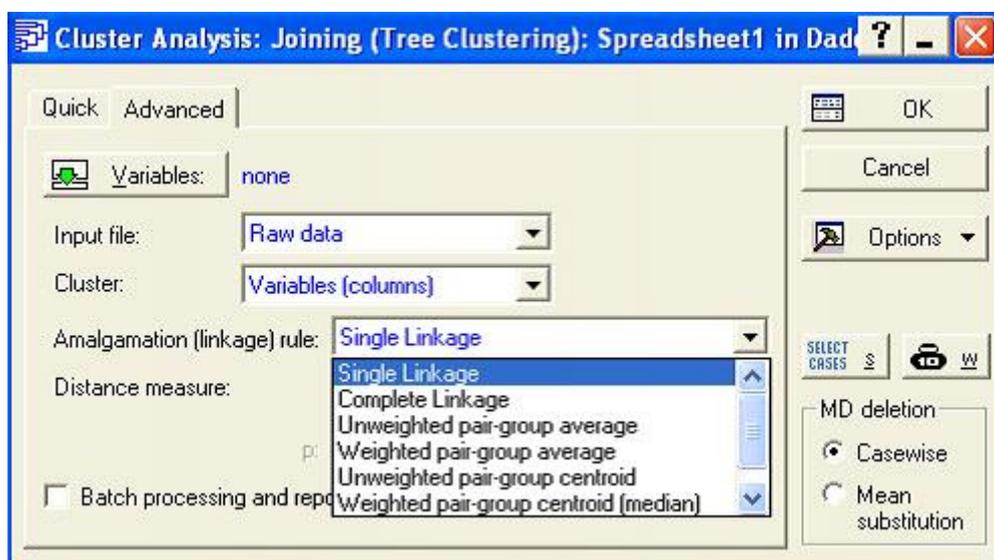


Figura 106 - Caixa de seleção para análise de agrupamento, selecionando-se *Advanced*, para a opção *Amalgamation (linkage) rule*

No item *Distance Measure*, Figura (107), pode-se selecionar o tipo de distância a ser utilizado. Cabe ressaltar que a distância euclidiana (*Euclidean distances*) é a mais empregada.

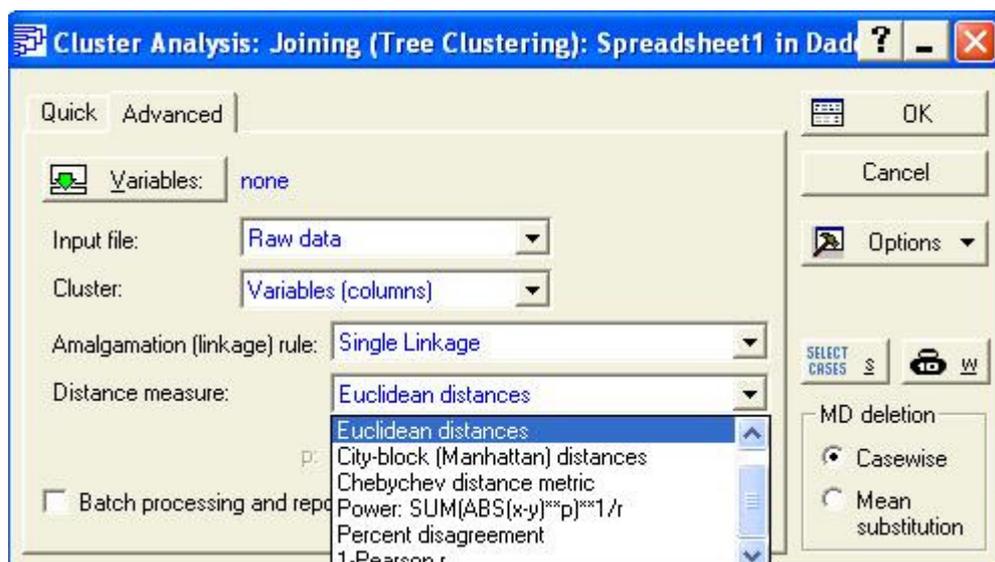


Figura 107 - Caixa de seleção para análise de agrupamento, selecionando-se *Advanced*, para a opção *Distance measure*

Na Figura (108) selecionam-se as variáveis para a análise. Para selecionar todas, clica-se em *Select All* e *OK*. Para selecionar algumas, com a tecla Ctrl pressionada, selecionam-se as variáveis desejadas.

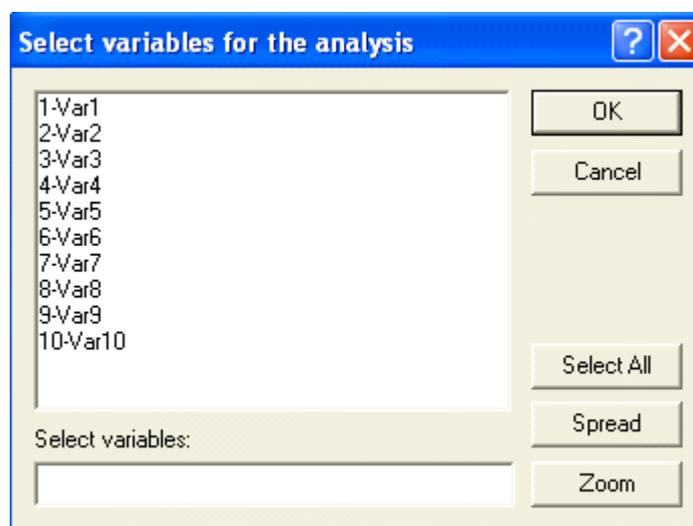


Figura 108 - Caixa de seleção das variáveis

Na Figura (109) são apresentadas as opções do dendograma, matriz de distâncias e estatísticas descritivas. Selecionando-se *Quick* surgem *Horizontal hierarchical tree plot* e *Vertical icicle plot*.

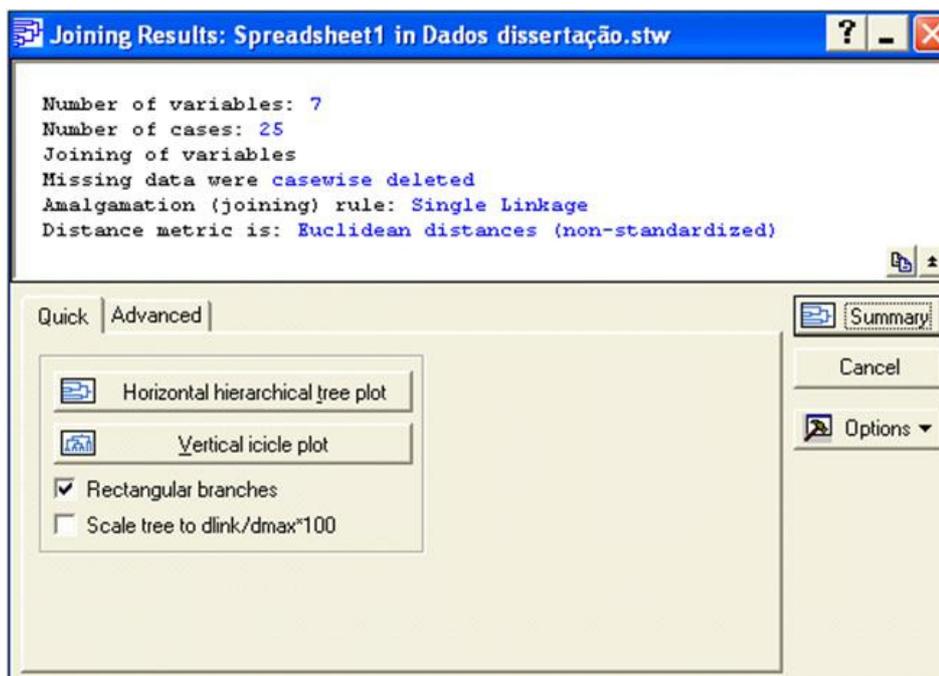


Figura 109 — Caixa de seleção para elaboração do dendograma, selecionando-se *Quick*

Selecionando-se *Advanced* surgem *Horizontal hierarchical tree plot*, *vertical icicle plot*, *Amalgamation Schedule*, *Graph of amalgamation Schedule*, *Distance matrix*, *Descriptive statistics* e *Matrix*, conforme a Figura (110).



Figura 110 - Caixa de seleção para elaboração do dendograma, selecionando-se *Advanced*

Escolhendo-se a opção *Horizontal hierarchical tree plot* obtém-se o dendograma horizontal, escolhendo-se a opção *Vertical icicle plot*, obtém-se o dendograma vertical. A

opção *Distance matrix* determina a matriz de distâncias entre as variáveis. Selecionando-se *Descriptive statistics* é possível determinar uma estatística descritiva dos dados.

O item *Vertical hierarchical tree plot* origina um dendograma empregando o método do vizinho mais próximo para o agrupamento dos dados e a distância euclidiana como medida de dissimilaridade. Os grupos são formados através de uma similaridade decrescente. Ou seja, dos mais similares (menor distância entre eles) até os menos similares (maior distância entre eles). Como exemplo, os dados a seguir (Figura 111) foram empregados na elaboração do dendograma vertical (Figura 112) e horizontal (Figura 113).

	1 Var1	2 Var2	3 Var3	4 Var4	5 Var5	6 Var6	7 Var7
1	5	30	35	85	110	140	145
2	10	40	45	90	115	150	155
3	15	50	55	95	120	160	165
4	20	60	65	100	125	170	175
5	25	70	75	105	130	180	185

Figura 111 - Dados empregados na elaboração de dendogramas

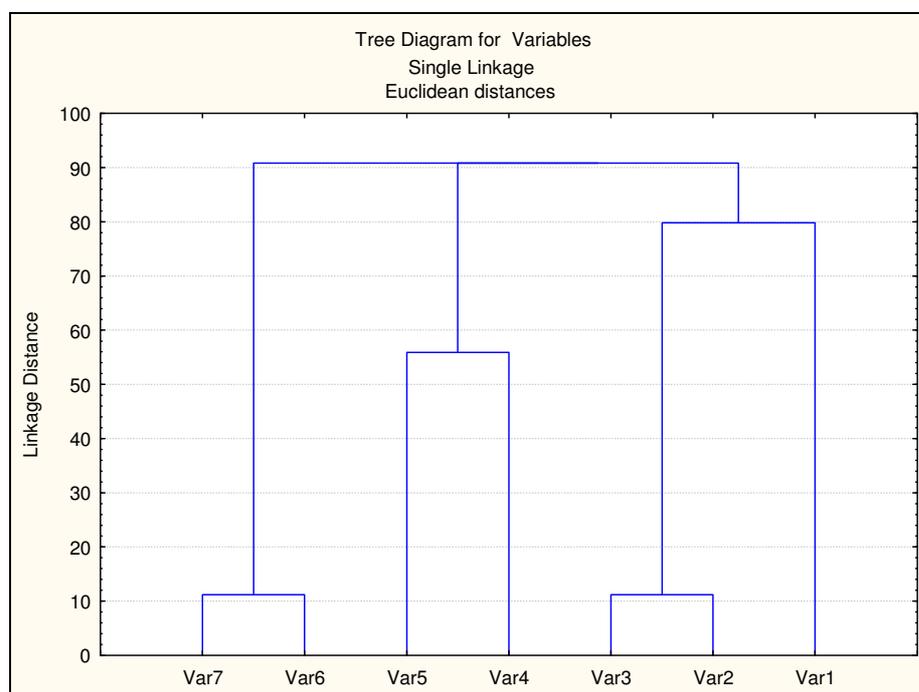


Figura 112 - Dendograma vertical

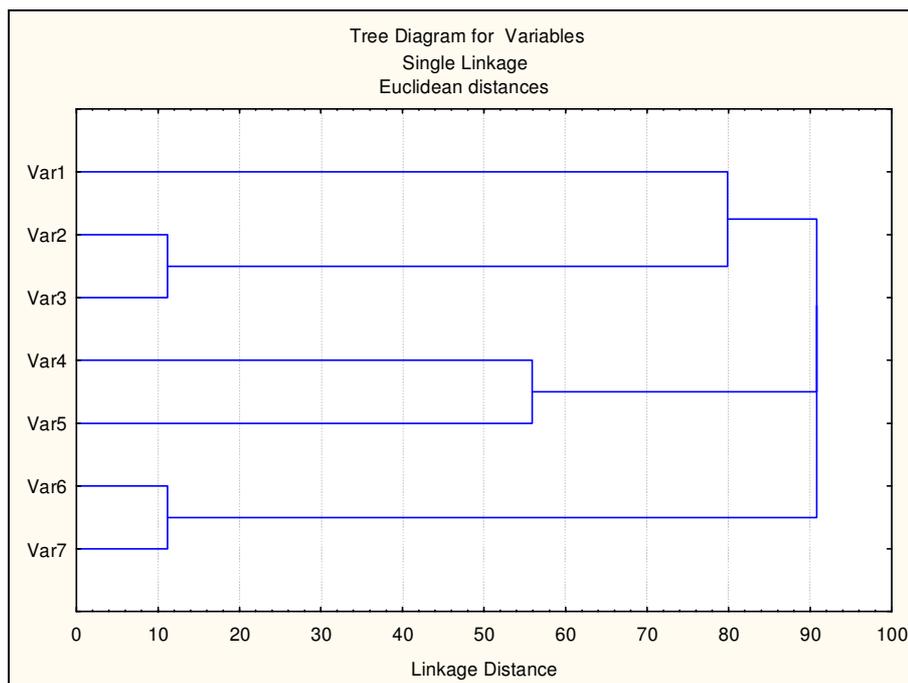


Figura 113 - Dendrograma horizontal

Para os dados utilizados, a matriz de distâncias euclidianas é uma matriz (7x7), conforme a Figura (114) a seguir:

Variable	Euclidean distances						
	Var1	Var2	Var3	Var4	Var5	Var6	Var7
Var1	0	80	91	179	235	325	336
Var2	80	0	11	102	157	246	257
Var3	91	11	0	91	146	235	246
Var4	179	102	91	0	56	146	157
Var5	235	157	146	56	0	91	102
Var6	325	246	235	146	91	0	11
Var7	336	257	246	157	102	11	0

Figura 114 - Matriz de distâncias euclidianas

As distâncias são calculadas da seguinte maneira:

$$d_{\text{var1,var1}} = \sqrt{(5-5)^2 + (10-10)^2 + (15-15)^2 + (20-20)^2 + (25-25)^2} = 0$$

$$d_{\text{var1,var2}} = \sqrt{(30-5)^2 + (40-10)^2 + (50-15)^2 + (60-20)^2 + (70-25)^2} = 79,84 = 80$$

$$d_{\text{var1,var3}} = \sqrt{(35-5)^2 + (45-10)^2 + (55-15)^2 + (65-20)^2 + (75-25)^2} = 90,83 = 91$$

e assim sucessivamente. Pela matriz anterior verifica-se que:

- os elementos da diagonal principal são nulos, ou seja, $d_{\text{var1}, \text{var1}} = 0$; $d_{\text{var2}, \text{var2}} = 0$; $d_{\text{var3}, \text{var3}} = 0$; ...; $d_{\text{var7}, \text{var7}} = 0$
- existe uma igualdade entre as linhas e as colunas; linha 1 = coluna 1, linha 2 = coluna 2, linha 3 = coluna 3, ..., linha 7 = coluna 7, ou seja, $d_{\text{var1}, \text{var2}} = d_{\text{var2}, \text{var1}}$; $d_{\text{var1}, \text{var3}} = d_{\text{var3}, \text{var1}}$; ...; $d_{\text{var1}, \text{var7}} = d_{\text{var7}, \text{var1}}$ e assim sucessivamente.

No dendograma formado, o eixo vertical indica o grau de similaridade e o eixo horizontal indica os indivíduos, segundo a ordem de agrupamento. As linhas verticais têm origem nos indivíduos e uma altura que corresponde ao nível no qual os indivíduos são considerados semelhantes. No dendograma é feito um corte através de uma linha denominada fenon que serve para determinar o número de grupos formados. Esse corte geralmente é feito em relação às maiores distâncias, considerando-se critérios pré-estabelecidos. Pelo dendograma observa-se que os indivíduos pertencentes ao mesmo grupo apresentam valores semelhantes e aqueles que apresentam valores distintos, formam outros grupos. Assim, fica constatado que existe homogeneidade dentro dos grupos e heterogeneidade entre os grupos.

O gráfico a seguir (Figura 115) auxilia nos casos em que o dendograma não mostra claramente onde ocorre um maior salto entre as distâncias.

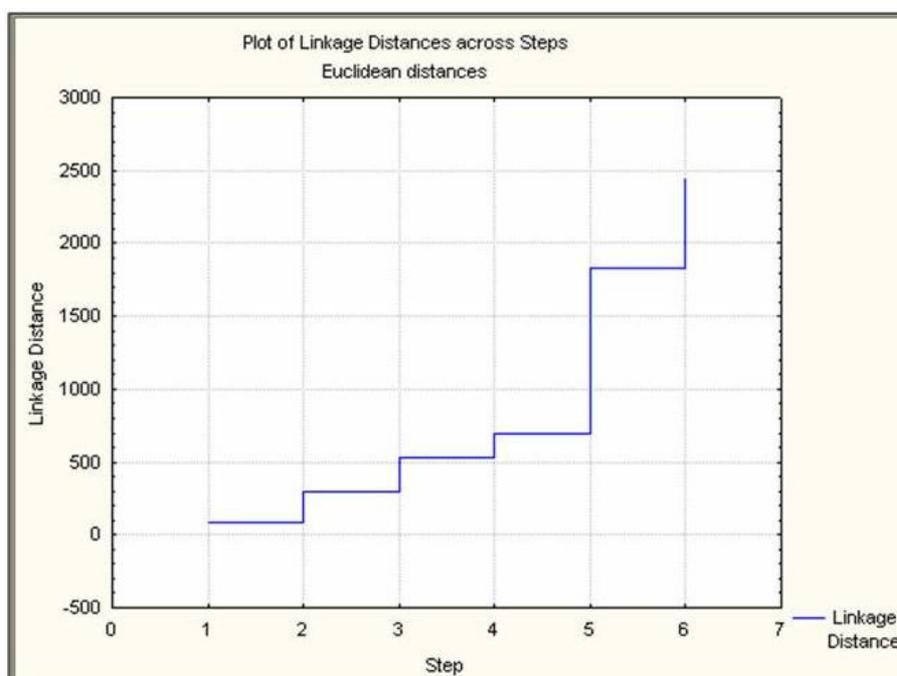


Figura 115 – Gráfico de ligação através dos passos

Para a realização de uma análise fatorial (AF) seleciona-se no módulo principal do *Statistica* a opção: *Multivariate Exploratory Techniques* e clica-se em *Factor Analysis*, conforme a Figura (116).

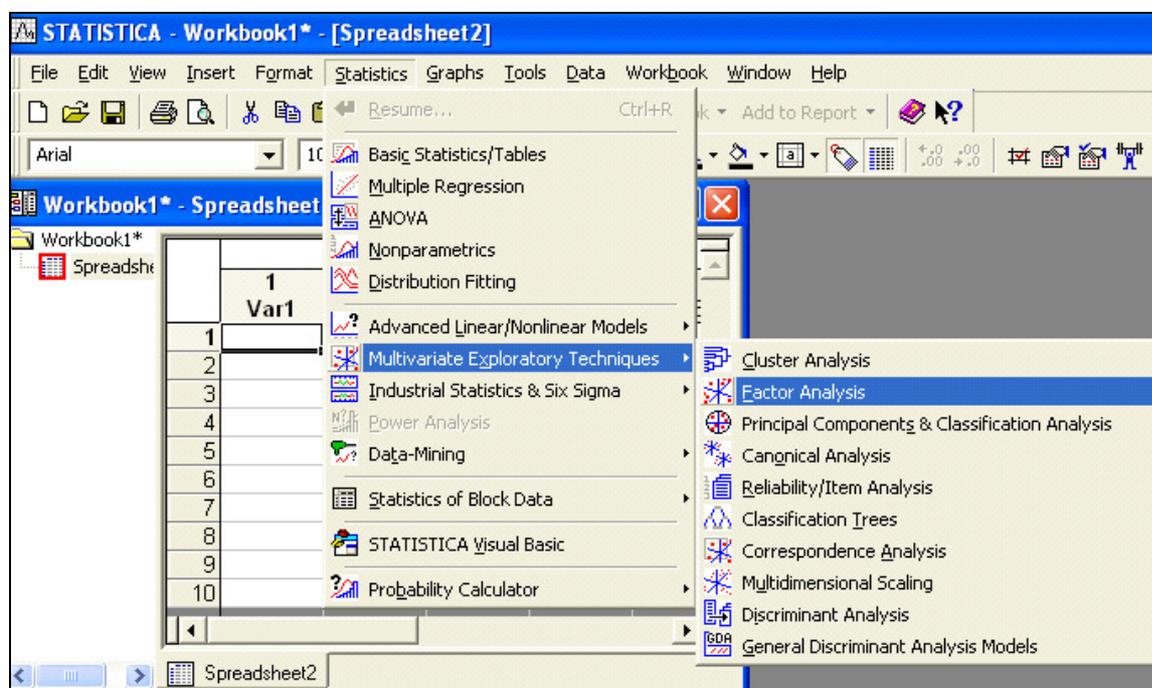


Figura 116 - Caixa de seleção para realização da análise fatorial

Surge a tela para a seleção das variáveis e as opções para *Input file*, conforme a Figura (117) a seguir. Na opção *Input file* surgem as opções *Raw Data* e *Correlation Matrix*. Se os dados forem empregados na forma como foram coletados, seleciona-se *Raw Data* e clica-se em *Ok*.

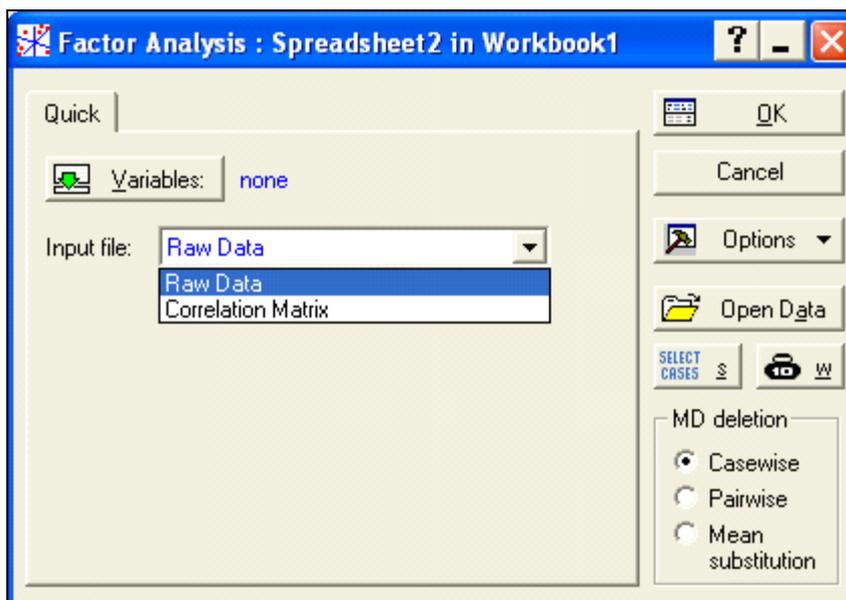


Figura 117 - Caixa de seleção para análise fatorial, opções para *Input file*

Clicando-se em *Variables* aparece uma lista de variáveis disponíveis (Figura 118). Para selecionar todas as variáveis, clica-se em *Select All*. No canto inferior esquerdo da tela aparecem as variáveis selecionadas. Quando existirem algumas variáveis sobre as quais se deseja verificar o comportamento numa etapa posterior a análise inicial, estas são denominadas variáveis suplementares, as quais deverão ser analisadas no círculo unitário. Clica-se em *OK*.

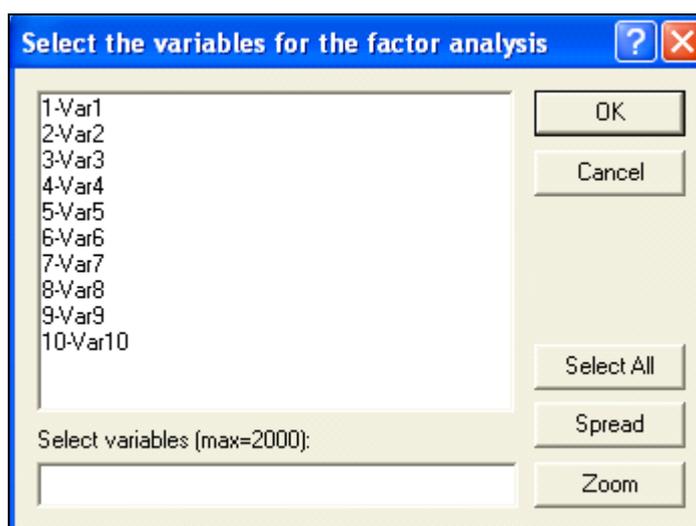


Figura 118 - Caixa de seleção das variáveis

A seguir, surgem os métodos utilizados para a extração dos fatores. Pela opção *Quick* (Figura 119) é possível definir o número de fatores para AF e ACP. No item *Maximum no. of*

factors define-se o número, que deverá ser igual ou inferior ao número de variáveis, pois não é possível ter um número de fatores maior do que o número de variáveis. No item *minimum eingevalue* é aconselhável a indicação de um valor bem pequeno tipo 0,001. Com isso, é obtido o número máximo de autovalores possíveis que resulta em melhores resultados. No caso de especificar um valor igual a 1, são obtidos os autovalores superiores a 1, segundo o critério de Kaiser (1960 *apud* MARDIA, 1979). Definido o valor, clica-se em *Ok*.

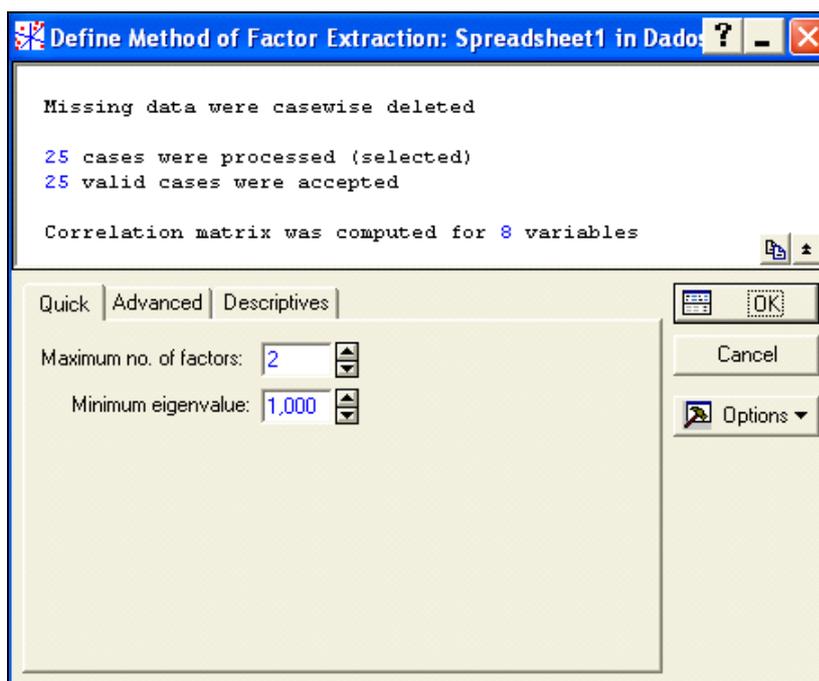


Figura 119 – Tela para realização de uma análise fatorial, opção *Quick*

A opção *Advanced* apresenta os métodos de extração que poderão ser empregados na análise fatorial, conforme a Figura (120).

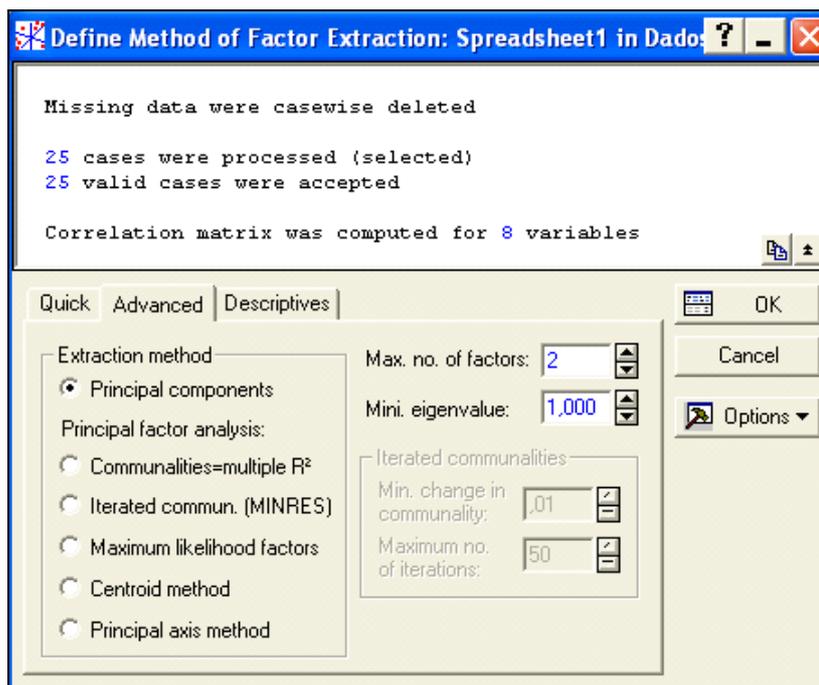


Figura 120 - Tela para realização de uma análise fatorial, opção *Advanced*

A opção *Descriptives* (Figura 121) apresenta as estatísticas descritivas dos dados, conforme a figura a seguir.

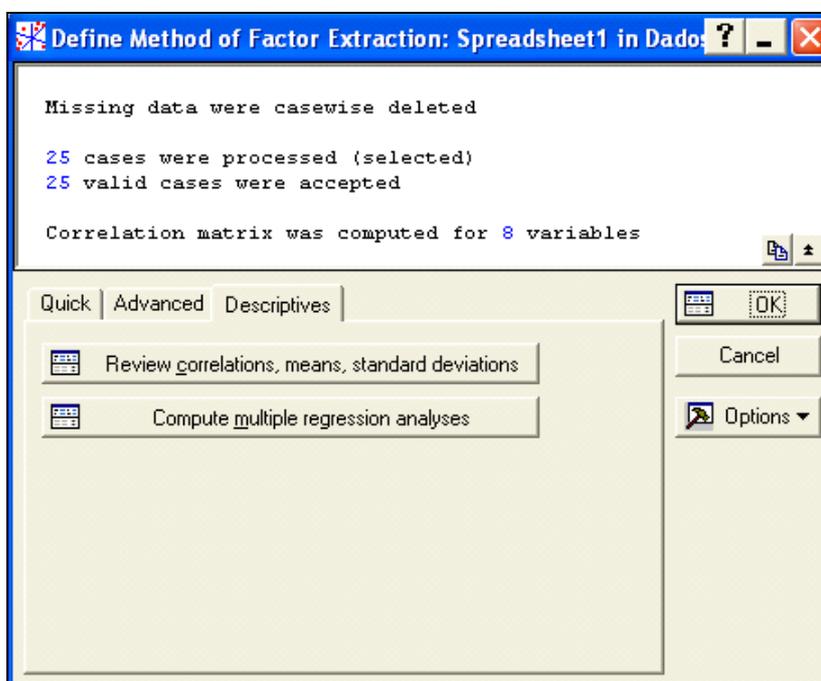


Figura 121 - Tela para realização de uma análise fatorial, opção *Descriptives*

Na Figura (122) selecionando-se *Explained variance* e em seguida *Eigenvalues* são obtidos os autovalores.

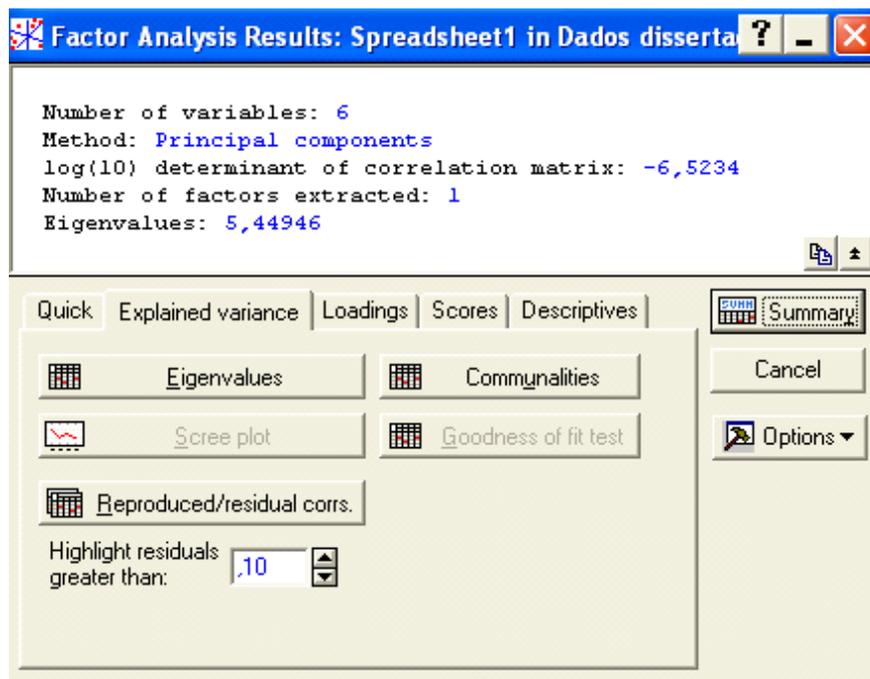


Figura 122 - Explained variance / Eigenvalues

A opção *Scree plot* é uma representação gráfica na qual o eixo das ordenadas contém a porcentagem de variação explicada pela componente e o eixo das abcissas contém os autovalores em ordem decrescente.

Na opção *Communalities* são determinadas as comunalidades que representam a proporção de contribuição de cada uma das variáveis (*factor loadings*). A partir dos resultados pode-se construir uma tabela contendo:

- número de componentes,
- autovalores,
- porcentagem da variância explicada por cada componente,
- autovalores acumulados, e
- porcentagem da variância explicada acumulada.

A partir disso, deve-se fazer a escolha do número de fatores podendo-se considerar diversos critérios. Segundo um critério, para a análise deve-se selecionar as componentes capazes de sintetizar uma variância acumulada em torno de 70%. Segundo outro critério, a seleção deve incluir as componentes com valores superiores a um (1).

Sendo a AF e ACP técnicas exploratórias de dados, é fundamental a realização de uma estatística descritiva das variáveis para uma melhor complementação dos resultados. Isto poderá ser obtido através da opção *Descriptives*.

Selecione-se a opção *Advanced/ correlations* obtém-se a matriz de correlação entre as variáveis.

Na Figura (123) seleciona-se a opção *Scores/ Factor score coefficients* para determinar os autovalores que definem a direção dos eixos para AF e ACP.

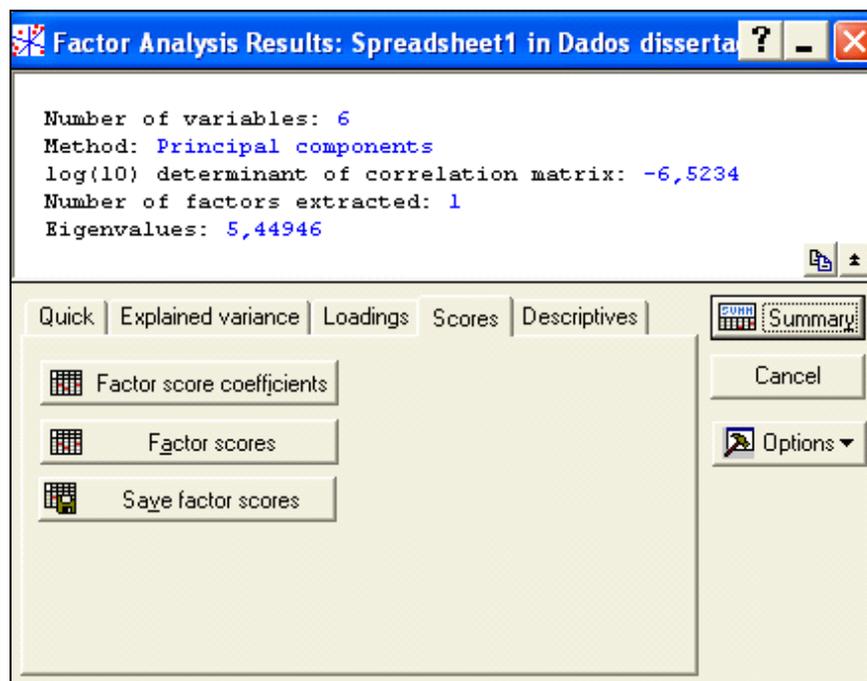


Figura 123 - Scores/ Factor score coefficients

Para a realização de uma análise de componentes principais, seleciona-se na barra superior do *Statistica* as opções *Statistics/ Multivariate Exploratory Techniques/ Principal Components & Classification Analysis*, conforme a Figura (124).

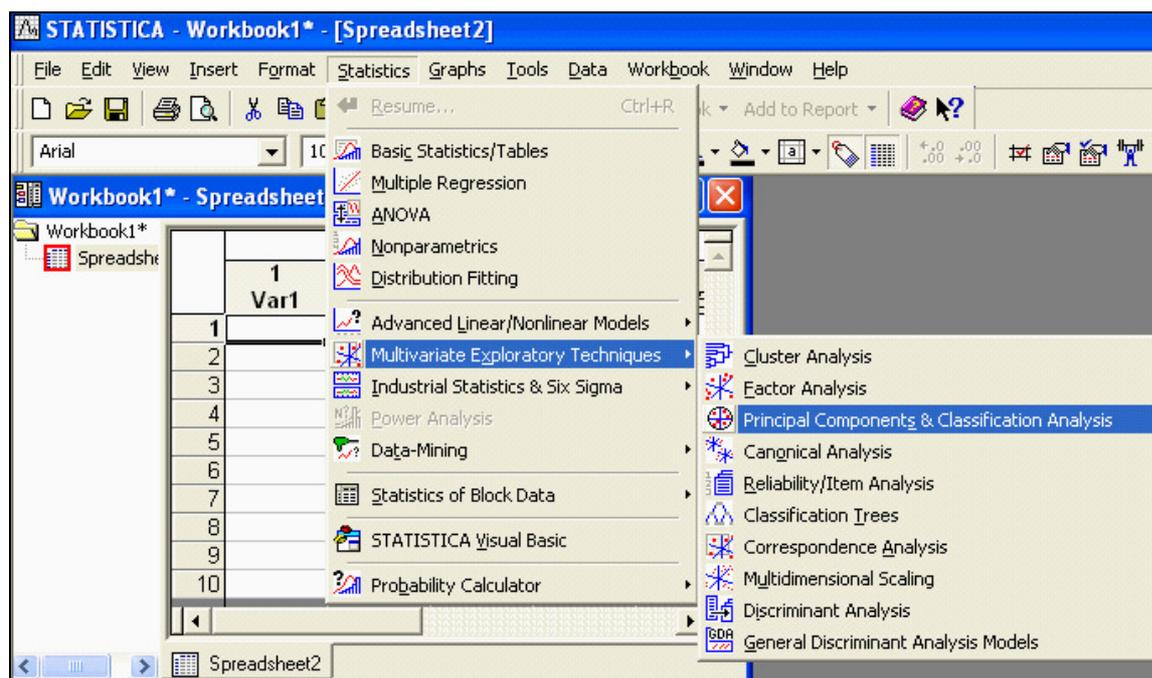


Figura 124 - Seleção da análise de componentes principais

A análise de componentes principais apresenta as opções *Quick* e *Advanced*. Selecionando-se *Quick* surge a Figura (125) a seguir.

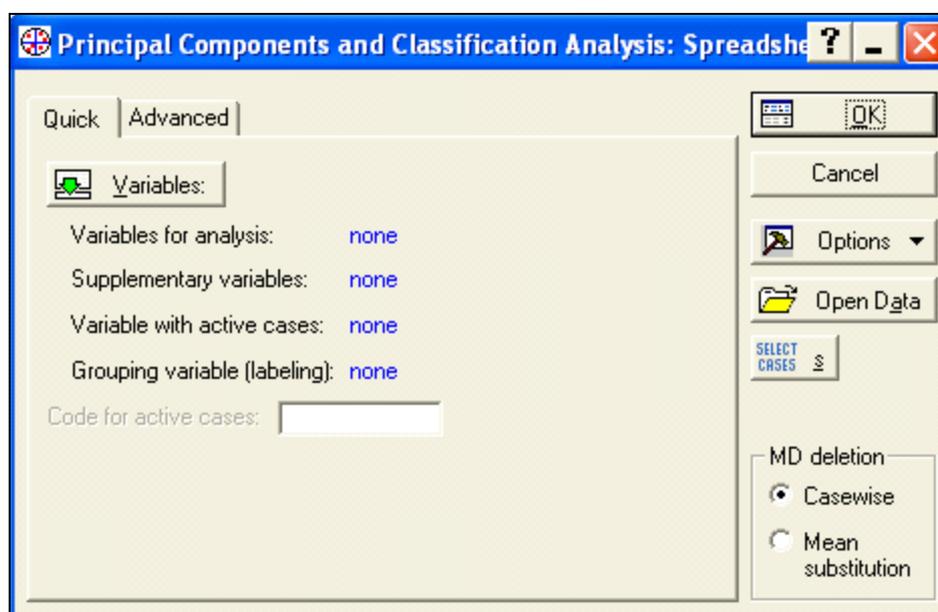
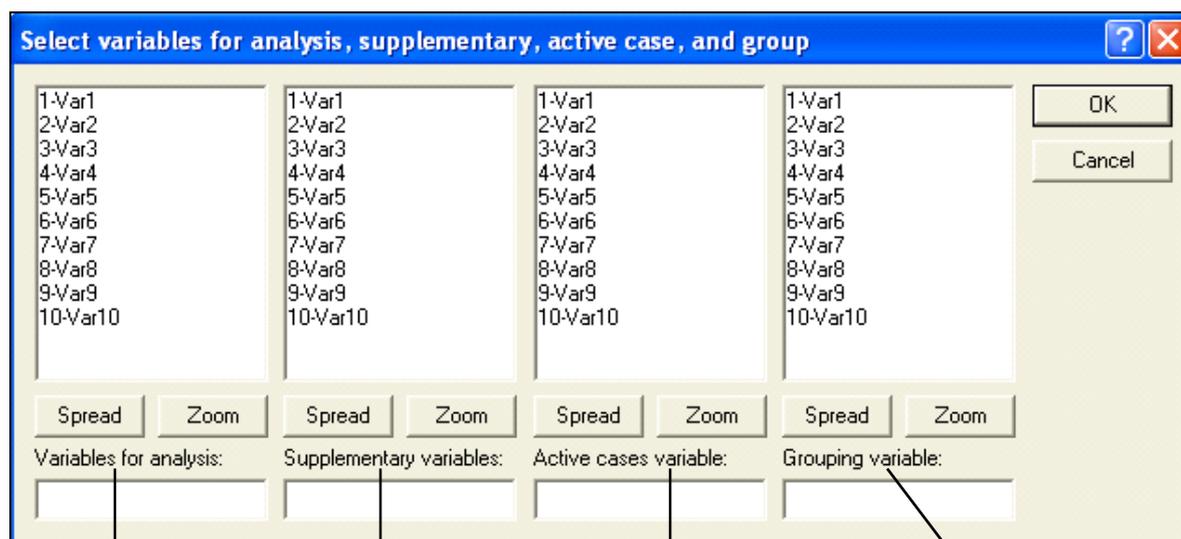


Figura 125 - Tela de análise de componentes principais, para opção *Quick*

No ícone *Variables* é feita a seleção das variáveis. Existem as opções: variáveis para análise, variáveis suplementares, casos ativos e grupo (Figura 126).



Variáveis para análise variáveis suplementares variável com casos ativos agrupando variável

Figura 126 - Tela de seleção das variáveis

Selecionam-se as variáveis e clica-se em *Ok*.

A Figura (127) apresenta a caixa de seleção para determinar as componentes principais. Seleciona-se *Cases* e em seguida *Factor scores*. Clica-se em *Ok*.

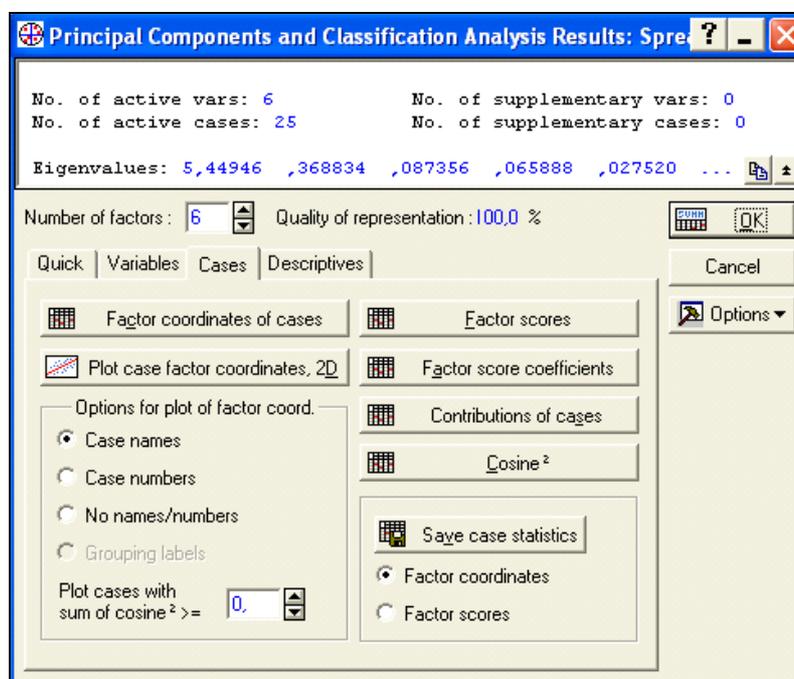


Figura 127 - Tela para determinação das componentes principais

Quando os dados estiverem dispostos em unidades diferentes, é necessário que se faça uma padronização dos mesmos. Isso faz com que a possível influência de uma variável sobre

a outra, na constituição das componentes, seja eliminada. Para isso, selecionam-se as variáveis originais e clica-se com o botão direito do mouse sobre a tela. Aparecerá a caixa de seleção para a padronização das variáveis, conforme a Figura (128) a seguir.

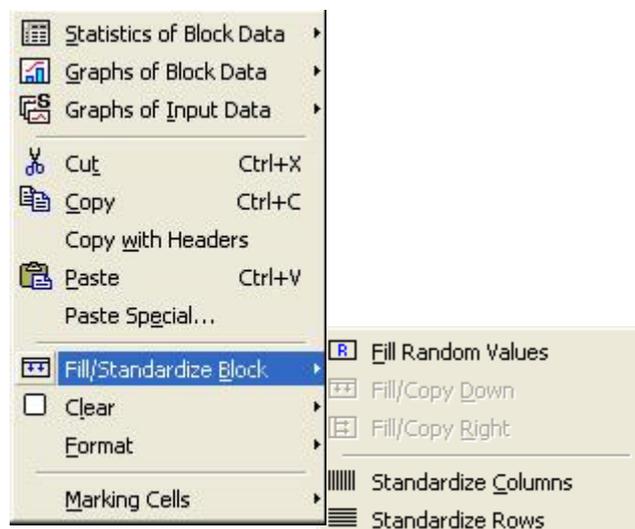


Figura 128 - Caixa de padronização das variáveis

Selecionando-se *Fill/ Standardizide Block* surgem duas opções de padronização:

- *Standardizide Columns*: a padronização é feita por colunas;
- *Standardizide Rows*: a padronização é feita por linhas.

Com a padronização das variáveis, determina-se a contribuição de cada variável para a formação dos fatores nos *Factor Loading*. Para isso, existem duas maneiras:

- 1º) Através da matriz de correlação entre as variáveis originais e as componentes principais.
- 2º) Através dos *Factor loadings*.

Selecionando-se a opção *Advanced* surge a Figura (129) a seguir.

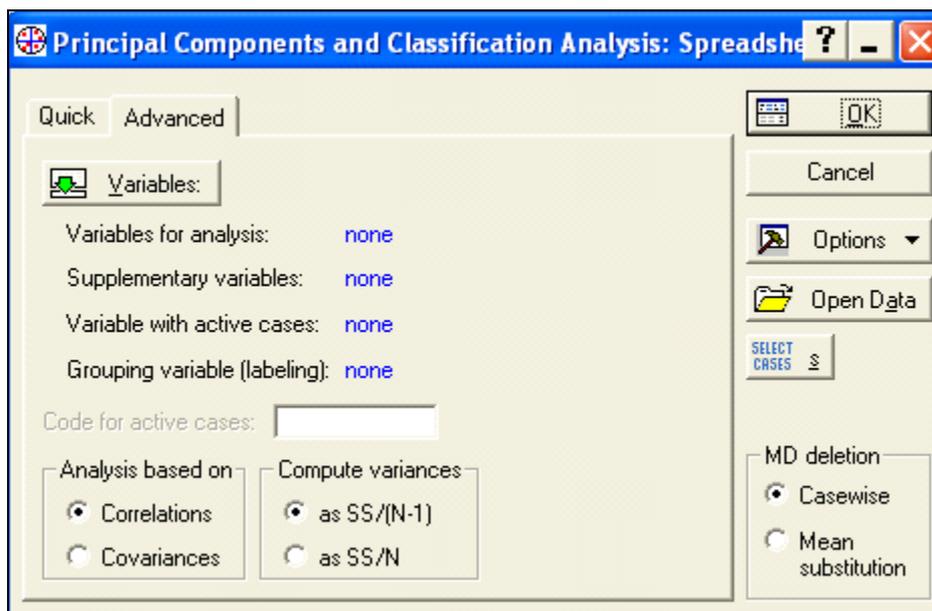


Figura 129 - Tela de análise de componentes principais, para opção *Advanced*

Para a determinação da matriz de correlação citada anteriormente, seleciona-se a opção *Save case statistics* e a seguir a opção *Factor Scores*. Clica-se em *Ok*.

A construção da matriz de correlação é feita selecionando-se *Statistics/ Basic Statistics/ Tables* segundo a Figura (130).

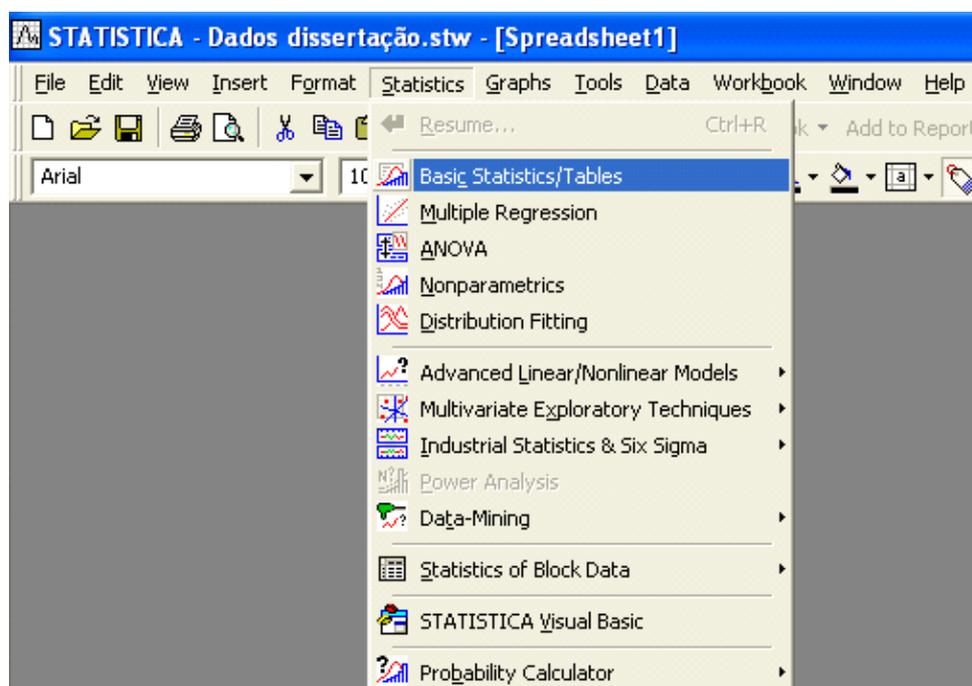


Figura 130 - Estatísticas básicas e tabelas

Surge uma tela com diversas opções (Figura 131). Seleciona-se *Correlation matrices* e *Ok*.

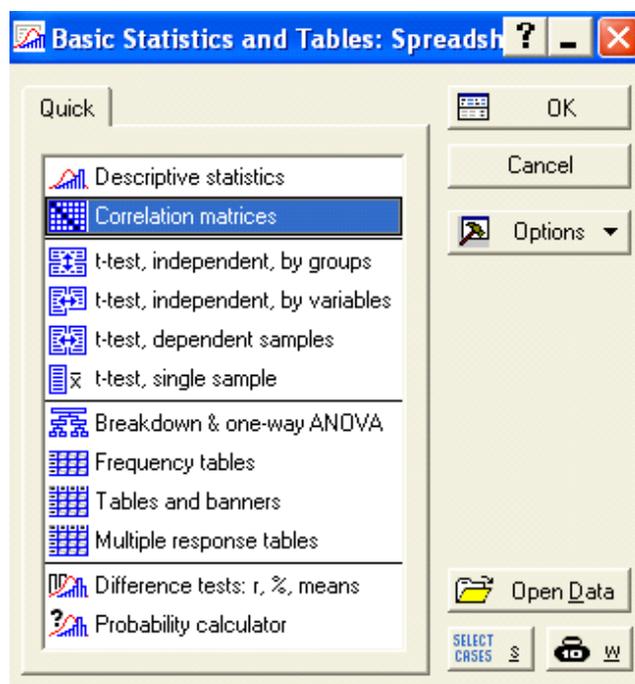


Figura 131 - Matrizes de correlação

Na Figura (132) seleciona-se a opção *Two lists (rect. matrix)*.

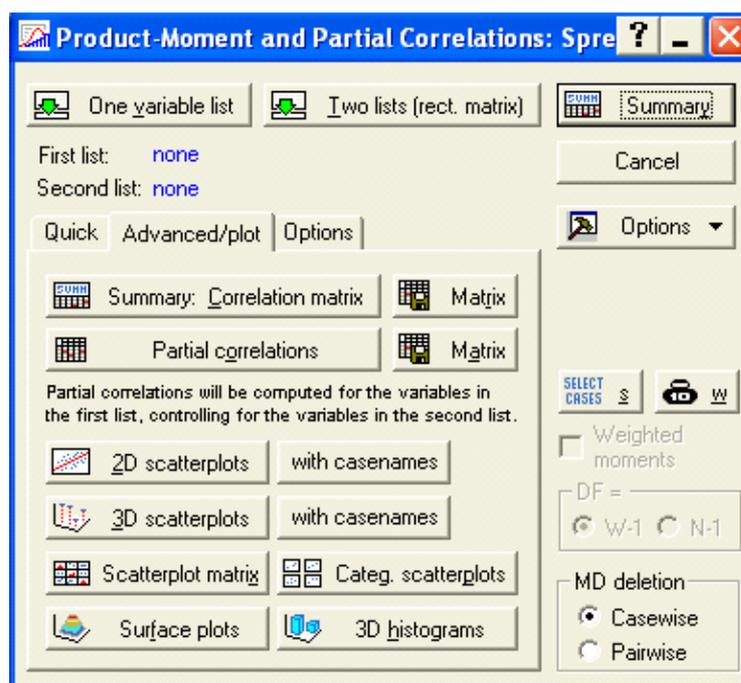


Figura 132 - Opções para matriz de correlação

Surge a Figura (133) onde é feita a seleção das variáveis e das componentes e clica-se em *Ok*.

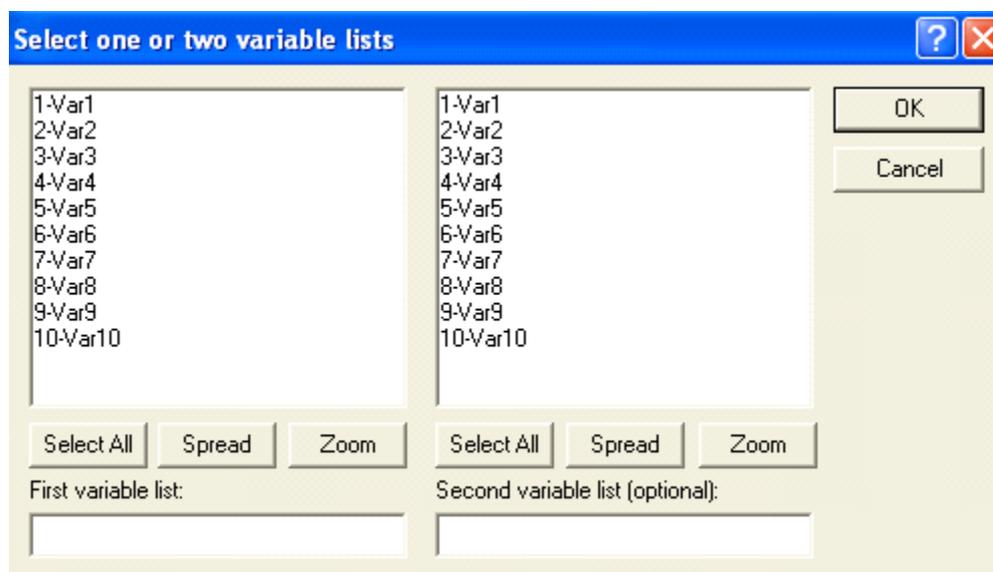


Figura 133 - Tela de seleção das variáveis e das componentes

Na Figura (132) onde estão as variáveis selecionadas (*first list*) e as componentes selecionadas (*second list*) clica-se em *Summary: Correlation matrix* para obter-se a matriz de correlação.

Cabe ressaltar que as correlações significativas aparecem em vermelho e ocorrem quando o valor de p for menor do que 0,05. Os valores encontrados representam o quanto cada variável contribui para cada fator, sendo os maiores valores absolutos os que melhor explicam o fator.

Como citados anteriormente, os *Factor loadings* possibilitam determinar a contribuição das variáveis relativa aos fatores formados.

Retornando a AF, na caixa de seleção das ferramentas de análise seleciona-se *Loadings*. Na opção *Factor rotation* seleciona-se *unrotated* e em seguida *Summary: factor loadings* determinando-se quanto cada variável contribui na formação das componentes. Na ACP as rotações de eixos servem para melhor demonstrar a dispersão dos dados.

Quando a identificação das variáveis que representam um fator tornar-se difícil, seja quando existir mais de um grupo de variáveis significativas no mesmo fator ou em fatores diferentes, realizam-se as rotações pois assim ocorre uma mais perfeita visualização da disposição dos pontos, ou seja, das variáveis de maior representatividade para cada

componente. O tipo de rotação mais empregado é a *Varimax normalized*, que mantém os eixos ortogonais. Isto é feito selecionando-se *Factor rotation: Varimax normalized*.

Para uma melhor visualização dos fatores, pode-se utilizar os gráficos de dispersão ou os planos fatoriais. Seleciona-se *Plot of loadings, 2D*, conforme a Figura (134) a seguir.

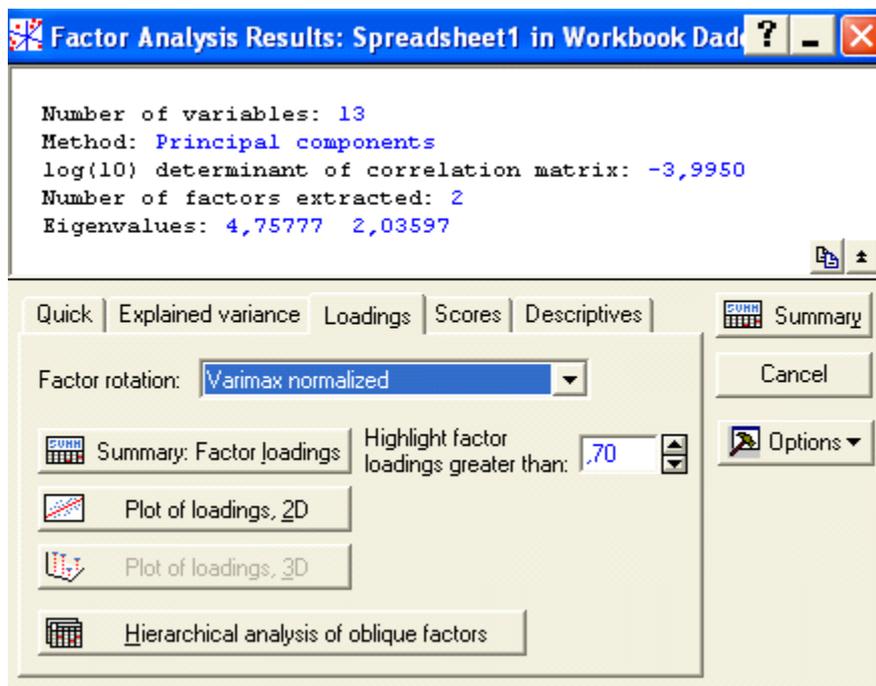


Figura 134 - Resultados para AF: *Loadings / Varimax normalized*

Assim, poderão ser identificadas as variáveis que possuem uma maior representatividade nos planos fatoriais.

Para auxiliar na interpretação dos planos fatoriais usa-se a análise de agrupamentos, que poderá confirmar se as variáveis contidas num mesmo grupo são aquelas que explicam determinado fator. Traçam-se perpendiculares em relação aos eixos x e y. Observando-se as distâncias das variáveis ou do conjunto de variáveis até a origem dos eixos, constata-se que aquelas que estiverem mais distantes possuem a maior influência sobre o fator analisado.

Selecionando-se *Loadings* e a seguir *Plot of loadings 3D* surge uma figura tridimensional, onde pode-se fazer uma visualização das variáveis em relação aos fatores. Encontrando-se os planos principais é possível visualizar uma nuvem de variáveis que representam cada plano e, também, uma nuvem de pontos onde se encontra a localização de cada objeto em relação às variáveis dos planos principais. Para isso, na caixa de seleção de variáveis para ACP, clica-se em *Variables*. Assim, surgem todas as variáveis para análise.

Quando se deseja analisar todas as variáveis, faz-se a seleção das mesmas na opção *Variables for analysis* e clica-se em *Ok*. Se houverem variáveis suplementares, as mesmas devem ser selecionadas na opção *Supplementary variables*.

Na caixa de seleção de variáveis e comandos para ACP selecionado-se *Variables*, escolhe-se a opção *Plot var. factor coordinates; 2D* para a construção dos planos principais contendo a nuvem de variáveis. Surge a seguir, a caixa de seleção dos fatores a serem relacionados. Selecionam-se os fatores e clica-se em *Ok*.

Na Figura (135) aparece o círculo de correlação unitário, contendo a nuvem de variáveis.

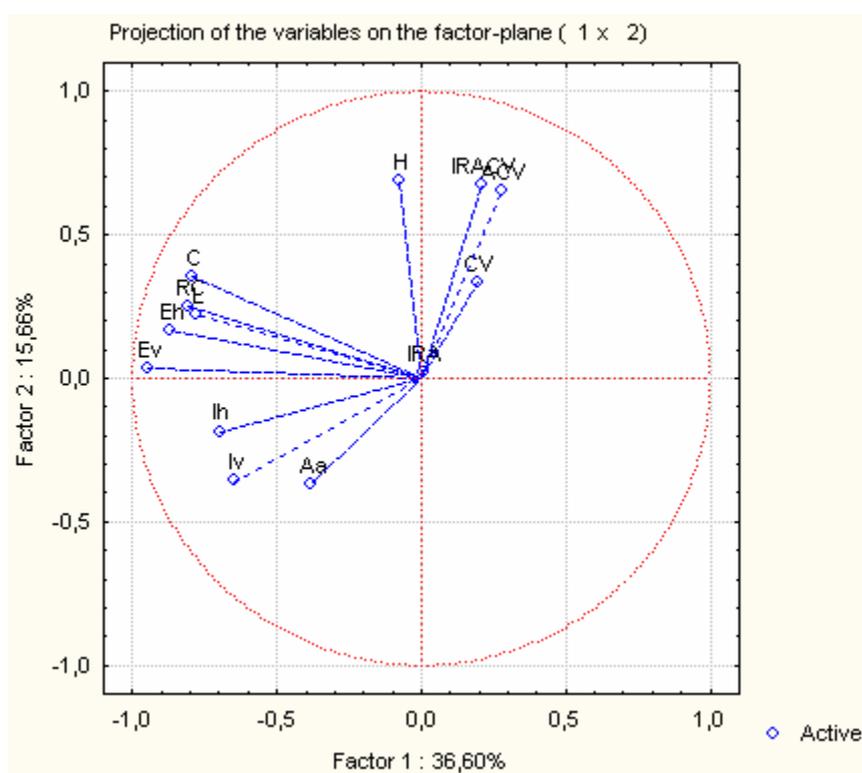


Figura 135 - Círculo de correlação unitário

Fazendo-se uma sobreposição deste círculo sobre os planos fatoriais pode-se identificar visualmente quais as variáveis que estão relacionadas com os casos em estudo.

Segundo Souza & Vicini (2005) cabe ressaltar que na interpretação da ACP é muito importante a definição do que cada eixo representa. Ainda, na ACP deve-se analisar os seguintes princípios:

- uma maior ou menor proximidade entre dois pontos-variáveis, no plano, demonstra uma maior ou menor correlação entre essas variáveis, sobretudo quando estiverem afastadas do centro do plano;

- a proximidade entre dois pontos demonstra uma similaridade entre esses dois objetos, em termos de variáveis.