



PONTÍFICIA UNIVERSIDADE CATÓLICA DE SÃO PAULO
Faculdade de Economia, Administração, Contabilidade e Atuariais.

ATLAS BRASIL 2013

DIMENSÕES DEMOGRAFIA E HABITAÇÃO

Disciplina: Métodos Quantitativos

Professor: Dr. Arnaldo Jose de Hoyos

Cleonice Nazaré do Nascimento

1º Semestre 2014

1. INTRODUÇÃO

O presente trabalho tem por objetivo efetuar uma análise discriminante e verificar principais componentes para as dimensões Demografia e Habitação, baseado no Índice de Desenvolvimento Humano Municipal Brasileiro do banco de dados original do Atlas Brasileiro 2013.

Para tal, iniciamos com o entendimento dos dados, incluindo a definição dos indivíduos e das variáveis, suas classificações em variáveis categóricas ou quantitativas, os significados e unidades de medida, além da apresentação da tabela de dados, gráficos e mapas.

Apresentamos na sequência Comparação de Médias, Intervalos de Confiança, Variâncias, Análise Discriminante, Regressão Logística, Árvores de Classificação e Análise de Correspondência de dados quantitativos sobre duas dimensões, sendo Demografia e Habitação dos Municípios do Brasil, respectivamente. O processo lógico é a análise discriminante nas funções linear e quadrática considerando as regiões brasileiras agrupadas no trabalho anterior (nº 08). Em seguida um agrupamento de dados apresentado x Brasis. Neste trabalho abordaremos as variáveis referentes ESPVIDA(1), MORT1(i) e MORT5(i) na dimensão Demografia e T_BANAGUA, T_DENS e T_AGUA na dimensão Habitação O software estatístico utilizado é o MINITAB16 e SPSS.

2. ENTENDENDO OS DADOS

2.1 Os Indivíduos

Os indivíduos desta análise são os municípios brasileiros, nas dimensões Demografia e Habitação, com origem no Banco de Dados Atlas Brasil 2013, com dados referentes ao ano de 2010. Trata-se de um total de 5565 municípios brasileiros e os dados analisados de cada município são as variáveis que descrevemos a seguir. (<http://www.atlasbrasil.org.br/2013/>)

2.2 As Variáveis

São 13 as variáveis desta pesquisa, sendo 08 na Dimensão Demografia e 05 na Dimensão Habitação, sendo a variável categórica o nome do Município. Todos os dados desta pesquisa são referentes ao ano de 2010 e a tabela 1, logo abaixo, apresenta a descrição de cada uma das variáveis, tipo e respectiva unidade de medida.

Tabela 1. As Variáveis

DIMENSÃO	VARIAVEL	SIGNIFICADO	TIPO	UNIDADE DE MEDIDA
Ambas	Município	Nome do Município	Categórica	N/A
Demografia	ESPVIDA	Esperança de Vida ao Nascer	Quantitativa	Anos
Demografia	MORT1_pt	Mortalidade Infantil	Quantitativa	Percentual
Demografia	MORT5-pt	Mortalidade até os 5 anos de idade	Quantitativa	Percentual
Demografia	T_ENV_pt	Taxa de Envelhecimento	Quantitativa	Percentual
Demografia	RAZDEP_pt	Razão de Dependência	Quantitativa	Índice
Demografia	FECTOT	Taxa de Fecundidade Total	Quantitativa	Percentual
Demografia	SOBRE40	Probabilidade de sobrevivência até 40 anos	Quantitativa	Percentual
Demografia	SOBRE60	Probabilidade de sobrevivência até 60 anos	Quantitativa	Percentual
Habitação	T_BANAGUA	% da população em domicílios com banheiro e água encanada	Quantitativa	Percentual
Habitação	T_LUZ	% da população em domicílios com energia elétrica	Quantitativa	Percentual
Habitação	T_DENS_pt	% da população em domicílios com densidade > 2	Quantitativa	Percentual
Habitação	T_AGUA	% da população em domicílios com água encanada	Quantitativa	Percentual
Habitação	T_LIXO	% da população em domicílios com coleta de lixo	Quantitativa	Percentual

2.3 A Tabela de Dados

Results for: **Demografia** (1).MTW

Descriptive Statistics: ESPVIDA; FECTOT; MORT1; MORT5; RAZDEP; SOBRE40; SOBRE60, T_ENV_pt

Variable	Mean	Minimum	Q1	Median	Q3	Maximum
ESPVIDA	73,089	65,300	71,150	73,470	75,160	78,640
FECTOT	2,1901	1,2100	1,8500	2,1300	2,4600	4,8900

MORT1	19,247	8,490	13,800	16,900	23,800	46,800
MORT5	21,533	9,980	16,040	19,420	25,825	50,940
RAZDEP	51,486	29,170	44,880	49,590	56,870	118,040
SOBRE40	93,781	88,830	92,950	93,890	94,680	97,090
SOBRE60	82,755	71,980	81,040	82,920	84,595	90,810

Results for: **Habitação (1).MTW**

Descriptive Statistics: T_AGUA; T_BANAGUA; T_DENS; T_LIXO; T_LUZ

Variable	Mean	Minimum	Q1	Median	Q3	Maximum
T_AGUA	85,598	0,150	79,635	90,280	96,260	100,000
T_BANAGUA	80,872	3,260	67,770	91,250	98,000	100,000
T_DENS	25,127	0,650	15,410	23,070	32,580	88,640
T_LIXO	94,047	0,000	93,720	98,030	99,490	100,000
T_LUZ	97,190	27,410	97,645	99,390	99,870	100,000

3. ANÁLISE DAS VARIÁVEIS

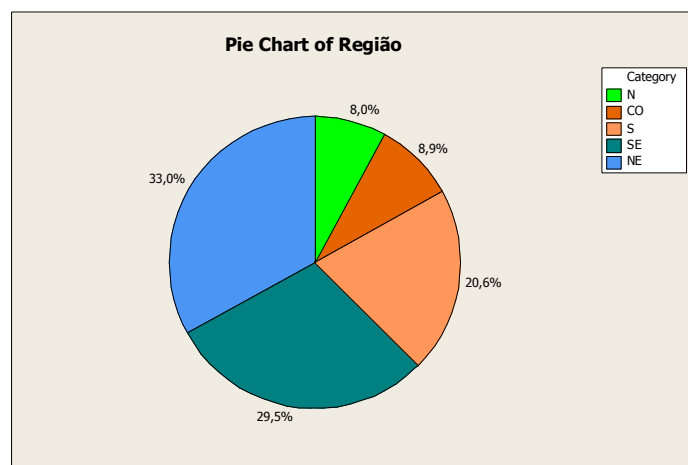
3.1 Variáveis Categóricas

Este tipo de variável indica que o foco de concentração deve ser a análise de gráficos do tipo *pie chart* e/ou barras.

3.1.1 Variável: “Município”

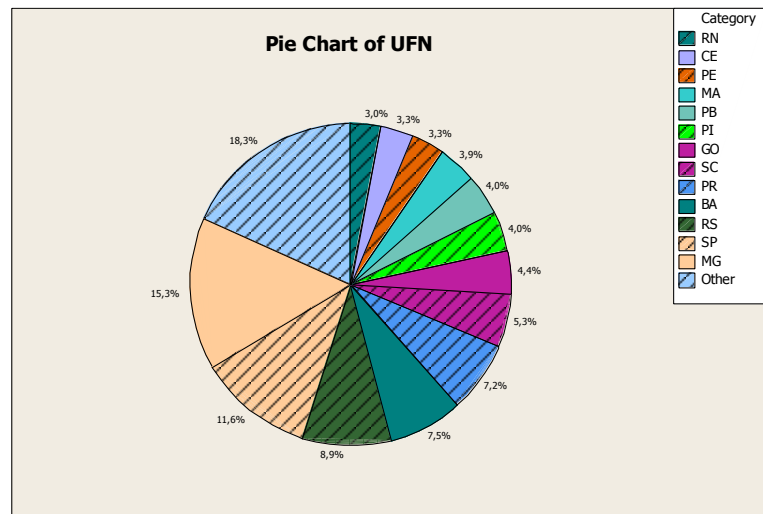
Nossa amostra totaliza 5565 municípios. Veremos nos mapas abaixo a distribuição dos municípios no território brasileiro.

O Gráfico 1 demonstra a distribuição dos municípios nas Regiões Brasileiras.



Podemos observar no Gráfico 1 que as maiores concentrações de municípios brasileiros estão nas regiões do Nordeste com 33,0% e Sudeste com 29,5% somando juntas mais de 50% dos municípios pesquisados (62,50%).

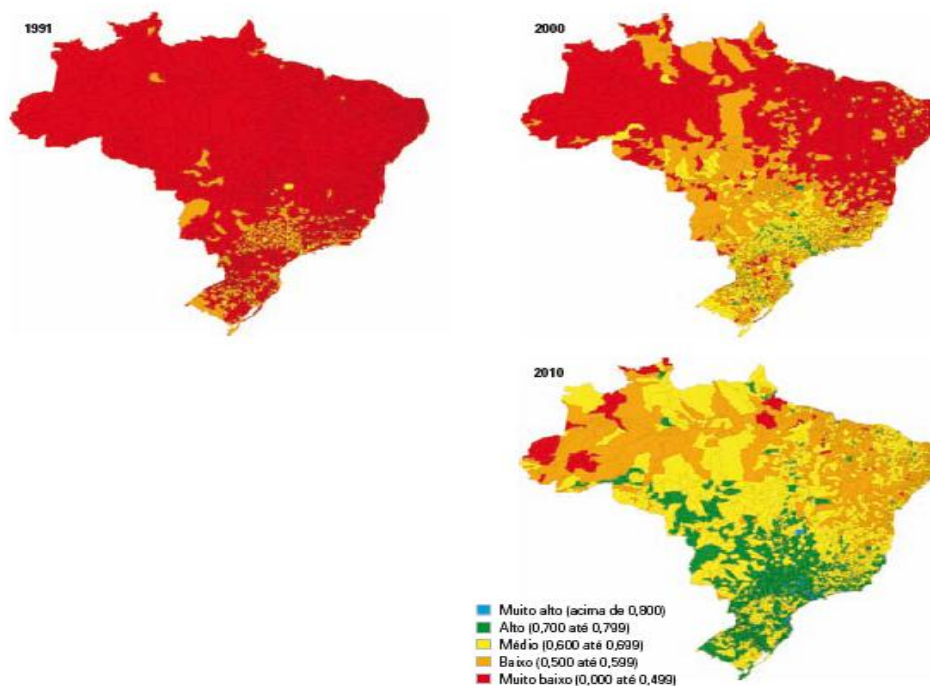
O Gráfico 2 demonstra a distribuição dos municípios pelas Unidades Federais Brasileiras.



Podemos observar no Gráfico 2 que as Unidades Federais mais representativas são Minas Gerais (15,3%), São Paulo (11,6%) e Rio Grande do Sul (8,9%). As menos expressivas são as Unidades Federais de Roraima, Amapá e Acre tendendo a zero pontos percentuais.

O Mapa 1 (conforme Atlas 2013), apresenta o IDHM (Índice de Desenvolvimento Humano Municipal) dos municípios brasileiros em 1991, 2000 e 2010. Observa-se que existe a concentração dos municípios que possuem IDHM elevados na região centro-sul do Brasil. Por outro lado, as regiões Norte concentram os municípios que apresentam Muito Baixo o índice e a maioria dos municípios que possuem Baixo Desenvolvimento Humano.

MAPA 1: IDHM DO BRASIL - 1991, 2000 E 2010



A tabela 2, ilustrada pelo gráfico 1, apresenta a evolução, entre 1991 e 2010, da classificação dos municípios brasileiros para as faixas de desenvolvimento humano propostas pelo Atlas Brasil 2013. A variação está entre Muito Baixo Desenvolvimento Humano (IDHM inferior a 0,500) a Muito Alto Desenvolvimento Humano (IDHM igual ou superior a 0,800). Em 1991, mais de 85% dos municípios encontravam-se na faixa de Muito Baixo Desenvolvimento Humano. Já nos anos 2000, pouco mais que 70% deles encontravam-se nas faixas de Baixo e Muito Baixo Desenvolvimento Humano. Na última análise referente a 2010, apenas um quarto (25%) dos municípios brasileiros encontrava-se nessas faixas e mais de 70% deles já figuravam nas faixas de Médio e Alto Desenvolvimento Humano. Segundo o Atlas Brasileiro 2013 isso ilustra os avanços do desenvolvimento humano no país nas últimas duas décadas.

**TABELA 2 | GRÁFICO 1: DISTRIBUIÇÃO DOS MUNICÍPIOS
FAIXAS DE DESENVOLVIMENTO HUMANO MUNICIPAL - 1991, 2000 E 2010**

Desenvolvimento Humano	1991		2000		2010	
	Nº de municípios	%	Nº de municípios	%	Nº de municípios	%
Muito alto	0	0,0	1	0,0	44	0,8
Alto	0	0,0	133	2,4	1.889	33,9
Médio	43	0,8	1.451	26,1	2.233	40,1
Baixo	745	13,4	1.852	29,7	1.367	24,6
Muito baixo	4.777	85,8	2.328	41,8	32	0,6



3.2 VARIÁVEIS QUANTITATIVAS

Veremos a seguir a utilização da Análise Discriminante separadamente para as Dimensões Demografia e Habitação, nos Métodos Linear e Quadrático. Partimos do princípio da parcimônia, onde se busca a simplicidade na análise, sendo assim caso a diferença seja irrelevante entre os métodos adotaremos o Método Linear como base.

3.2.1 ANÁLISE DISCRIMINANTE LINEAR POR REGIÃO (DIMENSÃO DEMOGRAFIA)

A análise discriminante é uma técnica da estatística multivariada utilizada para discriminar e classificar objetos, e estuda a separação de objetos de uma população em duas ou mais classes. Neste caso queremos discriminar os valores das variáveis ESPVIDA(1), MORT1(i) e MORT5(i) do Brasil, e utilizaremos inicialmente a variável categórica Região. Para geração de análise discriminante utilizaremos o comando do Minitab:

STAT >> MULTIVARIATE >> DISCRIMINANT ANALYSIS

Discriminant Analysis: Região versus MORT1(i); MORT5(i); ESPVIDA(1)

Linear Method for Response: Região

Predictors: MORT1(i); MORT5(i); ESPVIDA(1)

Group	CO	N	NE	S	SE
Count	461	431	1813	1183	1676

Summary of classification

Put into Group	True Group				
	CO	N	NE	S	SE
CO	217	28	104	136	528
N	32	257	538	70	94
NE	38	83	930	89	116
S	133	29	113	628	223
SE	41	34	128	260	715
Total N	461	431	1813	1183	1676
N correct	217	257	930	628	715
Proportion	0,471	0,596	0,513	0,531	0,427

N = 5564

N Correct = 2747

Proportion Correct = 0,494

Squared Distance Between Groups

	CO	N	NE	S	SE
CO	0,00000	4,27549	4,22993	0,94304	0,56339
N	4,27549	0,00000	0,73269	2,06538	2,48452
NE	4,22993	0,73269	0,00000	3,14829	2,66362
S	0,94304	2,06538	3,14829	0,00000	0,31082
SE	0,56339	2,48452	2,66362	0,31082	0,00000

AGRUPAMENTO= SUL+SUDESTE

Linear Discriminant Function for Groups

	CO	N	NE	S	SE
Constant	-18,36	-19,21	-14,28	-21,29	-18,18
MORT1(i)	86,21	-42,48	-35,94	30,85	34,39
MORT5(i)	18,94	155,97	134,29	81,74	66,74
ESPVIDA(1)	-69,57	-77,82	-67,75	-72,95	-64,04

Com base nas informações apresentadas acima pode ser notado que a região que acertou mais é Norte (0,596), seguida pela região Sul (0,531). A região com maior erro é a Sudeste (0,427). As informações ainda exibem o cruzamento de dados entre as regiões, por exemplo, a região Nordeste possui 1813 municípios e apenas 930 correspondem a região. Essa matriz é denominada *confusion matrix* (matriz de confusão). Podemos analisar a possibilidade de agrupar regiões com maior similaridade.

3.2.2. ANÁLISE DISCRIMINANTE LINEAR POR “4 BRASIS” (Demografia)

Esta segunda análise está interessada em verificar os possíveis agrupamentos dos dados utilizando a variável “4 Brasis”, calculada a partir do exercício anterior, e demonstra os agrupamentos do Brasil segundo sua proximidade de dados de demografia. Para esta análise

foram agrupadas as regiões de Sul, Sudeste como SSE, e as regiões de Norte e Nordeste e Centro_Oeste separadamente.

Discriminant Analysis: 4 BRASIS versus ESPVIDA(1); MORT1(i); MORT5(i)

Linear Method for Response: 4 BRASIS

Predictors: ESPVIDA(1); MORT1(i); MORT5(i)

Group	CO	N	NE	SSE
Count	461	431	1813	2859

Summary of classification

Put into Group	True Group			
	CO	N	NE	SSE
CO	272	37	124	828
N	32	258	539	203
NE	38	83	930	205
SSE	119	53	220	1623
Total N	461	431	1813	2859
N correct	272	258	930	1623
Proportion	0,590	0,599	0,513	0,568

N = 5564

N Correct = 3083

Proportion Correct = 0,554

Squared Distance Between Groups

	CO	N	NE	SSE
CO	0,00000	4,25710	4,23036	0,64201
N	4,25710	0,00000	0,70830	2,22938
NE	4,23036	0,70830	0,00000	2,78369
SSE	0,64201	2,22938	2,78369	0,00000

Linear Discriminant Function for Groups

	CO	N	NE	SSE
Constant	-17,85	-18,55	-13,79	-18,86
ESPVIDA(1)	-66,45	-74,27	-64,69	-64,44
MORT1(i)	87,47	-41,08	-34,73	34,24
MORT5(i)	13,68	150,00	129,14	67,40

Existem duas possibilidades de realizar a análise discriminante que são a linear e a quadrática. Dependendo da variável deve-se dar mais peso e mais atenção a um método em detrimento do outro. Neste caso a linear não apresenta dados satisfatórios. Podemos observar que alguns estados e municípios da região SSE tem características da região CO, visto pelo número 828 municípios foram encontrados na intersecção entre os dois agrupamentos (regiões).

3.2.3. ANÁLISE DISCRIMINANTE QUADRÁTICA POR “4 BRASIS” (Demografia)

Esta segunda análise está interessada em verificar os agrupamentos dos dados utilizando a variável “4 Brasís”, utilizando o Método Quadrático para Análise Discriminante. Foram mantidos os agrupamentos utilizados no item anterior, no Método Linear.

Discriminant Analysis: 4 BRASIS versus ESPVIDA(1); MORT1(i); MORT5(i)

Quadratic Method for Response: 4 BRASIS

Predictors: ESPVIDA(1); MORT1(i); MORT5(i)

Group	CO	N	NE	SSE
Count	461	431	1813	2859

Summary of classification

Put into Group	True Group			SSE
	CO	N	NE	
CO	159	16	52	229
N	37	296	785	242
NE	32	45	670	148
SSE	233	74	306	2240
Total N	461	431	1813	2859
N correct	159	296	670	2240
Proportion	0,345	0,687	0,370	0,783

N = 5564

N Correct = 3365

Proportion Correct = 0,605

From Generalized Squared Distance to Group

Group	CO	N	NE	SSE
CO	-17,46	-14,58	-14,76	-17,82
N	-15,37	-18,99	-17,69	-15,57
NE	-14,56	-18,12	-18,16	-14,89
SSE	-17,08	-16,97	-16,06	-18,76

No modelo quadrático a proporção subiu de forma inexpressiva saindo de 0,554 na linear atingindo o índice de 0,605 na quadrática. Como comentamos anteriormente decidimos escolher o método linear por sua simplicidade.

3.2.4. ANÁLISE DISCRIMINANTE LINEAR PARA DADOS AGRUPADOS (Dimensão Demografia)

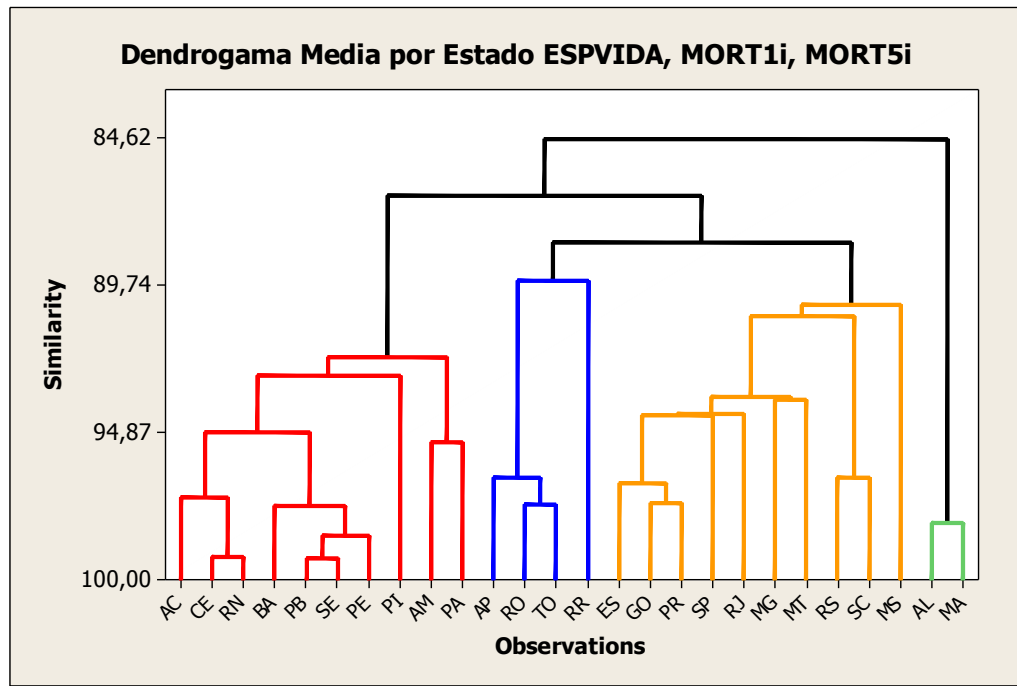


Grafico3. Dendrograma da Média ESPVIDA x MORT1x MORT5 por grupos do Brasil (classificação supervisionada)

Neste exemplo abaixo vamos através do dendrograma pesquisar o grau de similaridade das médias das variáveis ESPVIDA(1), MORT1(i) e MORT5(i) nos agrupamentos. Com base na análise discriminante poderemos verificar a proporção correta dos agrupamentos.

Discriminant Analysis: Agrupamentos versus M_ESPVIDA; M_MORT1i; M_MORT5i

Linear Method for Response: Agrupamentos

Predictors: M_ESPVIDA; M_MORT1i; M_MORT5i

Group	G1	G2	G3	G4
Count	10	2	4	10

Summary of classification

Put into Group	True Group			
	G1	G2	G3	G4
G1	10	0	0	0
G2	0	2	0	0
G3	0	0	4	0
G4	0	0	0	10
Total N	10	2	4	10
N correct	10	2	4	10
Proportion	1,000	1,000	1,000	1,000

N = 26

N Correct = 26

Proportion Correct = 1,000

Squared Distance Between Groups

	G1	G2	G3	G4
G1	0,000	14,247	17,541	131,577
G2	14,247	0,000	60,340	205,125
G3	17,541	60,340	0,000	62,305
G4	131,577	205,125	62,305	0,000

Linear Discriminant Function for Groups

	G1	G2	G3	G4
Constant	-97,8	-54,4	-164,4	-302,0
M_ESPVIDA	11,5	41,9	45,2	138,1
M_MORT1i	1347,1	1038,1	1709,1	2733,4
M_MORT5i	-973,2	-782,8	-1247,1	-2164,4

No mapa abaixo apresentamos visualmente a divisão dos Grupos G1, G2, G3 e G4 divididos por cores dos Estados, conforme o agrupamento por similaridade. Nesta representação observamos a concentração nos estados do meio ao sul do país com significativa concentração. Quanto aos estados de Maranhão e Alagoas aparecem de forma isolada e seria interessante se aprofundar para melhores informações e características que justifiquem tais números.



Mapa(2). Brasil por Agrupamentos – Média Variáveis ESPVIDA(1), MORT1(i) e MORT5(i)

Legenda: G1= Azul, G2= Laranja, G3= Verde, G4= Rosa

3.2.5 ANÁLISE DISCRIMINANTE LINEAR POR REGIÃO (DIMENSÃO HABITAÇÃO)

A análise discriminante é uma técnica da estatística multivariada utilizada para discriminar e classificar objetos, e estuda a separação de objetos de uma população em duas ou mais classes. Neste caso queremos discriminar os valores das variáveis T_BANAGUA, T_DENS e T_AGUA do Brasil, e utilizaremos inicialmente a variável categórica Região. Para geração de análise discriminante utilizaremos o comando do Minitab:

STAT >> MULTIVARIATE >> DISCRIMINANT ANALYSIS

Results for: Habitação_OFICIAL.MTW

Discriminant Analysis: Região versus T_AGUA(1); T_BANAGUA(1); T_DENS(1)

Linear Method for Response: Região

Predictors: T_AGUA(1); T_BANAGUA(1); T_DENS(1)

Group	CO	N	NE	S	SE
Count	461	431	1813	1183	1676

Summary of classification

Put into Group	True Group				
	CO	N	NE	S	SE
CO	44	40	113	76	91
N	54	233	509	65	126
NE	39	46	696	73	144
S	189	51	226	700	561
SE	135	61	269	269	754
Total N	461	431	1813	1183	1676
N correct	44	233	696	700	754
Proportion	0,095	0,541	0,384	0,592	0,450

N = 5564

N Correct = 2427

Proportion Correct = 0,436

Squared Distance Between Groups

	CO	N	NE	S	SE
CO	0,00000	2,22565	1,06063	0,21879	0,08375
N	2,22565	0,00000	0,88923	3,70667	2,59706
NE	1,06063	0,88923	0,00000	1,70670	1,40630
S	0,21879	3,70667	1,70670	0,00000	0,27337
SE	0,08375	2,59706	1,40630	0,27337	0,00000

Com base nas informações apresentadas acima pode ser notado que a região que acertou mais é Sul (0,592), seguida pela região Norte (0,541). A região com maior erro é o Centro Oeste (0,095). As informações ainda exibem o cruzamento de dados entre as regiões, por exemplo, a região Nordeste possui 1813 municípios e apenas 696 correspondem a região. Essa matriz é denominada *confusion matrix* (matriz de confusão). Podemos analisar a possibilidade de agrupar regiões com maior similaridade.

3.2.6. ANÁLISE DISCRIMINANTE LINEAR POR “4 BRASIS” (Habitação)

Esta segunda análise está interessada em verificar os possíveis agrupamentos dos dados utilizando a variável “4 Brasis”, calculada a partir do exercício anterior, e demonstra os agrupamentos do Brasil segundo sua proximidade de dados de habitação. Para esta análise foram agrupadas as regiões de Sul, Sudeste como SSE, e as regiões de Norte e Nordeste e Centro_Oeste separadamente.

Discriminant Analysis: Agrupamentos versus T_AGUA(1); T_BANAGUA(1); ...

Linear Method for Response: 4 BRASIS

Predictors: T_AGUA(1); T_BANAGUA(1); T_DENS(1)

Group	CO	N	NE	SSE
Count	461	431	1813	2859

Summary of classification

Put into Group	True Group			
	CO	N	NE	SSE
CO	126	66	230	463
N	57	236	536	230
NE	41	49	718	232
SSE	237	80	329	1934
Total N	461	431	1813	2859
N correct	126	236	718	1934
Proportion	0,273	0,548	0,396	0,676

N = 5564

N Correct = 3014

Proportion Correct = 0,542

Squared Distance Between Groups

	CO	N	NE	SSE
CO	0,00000	2,19740	1,06000	0,07310
N	2,19740	0,00000	0,86963	2,95637
NE	1,06000	0,86963	0,00000	1,46262
SSE	0,07310	2,95637	1,46262	0,00000

Linear Discriminant Function for Groups

	CO	N	NE	SSE
Constant	-38,295	-40,682	-34,113	-38,140
T_AGUA(1)	34,328	36,771	31,840	32,656
T_BANAGUA(1)	35,256	31,806	31,967	36,898
T_DENS(1)	62,536	70,978	64,008	62,102

Existem duas possibilidades de realizar a análise discriminante que são a linear e a quadrática. Dependendo da variável deve-se dar mais peso e mais atenção a um método em detrimento do outro. Neste caso a linear não apresenta dados satisfatórios. Podemos observar que alguns estados e municípios da região SSE tem características da região CO, visto pelo número 463 municípios foram encontrados na intersecção entre os dois agrupamentos (regiões).

3.2.7. ANÁLISE DISCRIMINANTE QUADRÁTICA POR “4 BRASIS” (Habitação)

Esta segunda análise está interessada em verificar os agrupamentos dos dados utilizando a variável “4 Brasis”, utilizando o Método Quadrático para Análise Discriminante. Foram mantidos os agrupamentos utilizados no item anterior, no Método Linear.

Discriminant Analysis: Agrupamentos versus T_AGUA(1); T_BANAGUA(1); ...

Quadratic Method for Response: 4 Brasis

Predictors: T_AGUA(1); T_BANAGUA(1); T_DENS(1)

Group	CO	N	NE	SSE
Count	461	431	1813	2859

Summary of classification

Put into Group	True Group			
	CO	N	NE	SSE
CO	92	78	388	414
N	45	221	346	151
NE	39	42	677	240
SSE	285	90	402	2054
Total N	461	431	1813	2859
N correct	92	221	677	2054
Proportion	0,200	0,513	0,373	0,718

N = 5564

N Correct = 3044

Proportion Correct = 0,547

From Generalized Squared Distance to Group

Group	CO	N	NE	SSE
CO	-13,35	-10,40	-11,10	-13,44
N	-11,02	-11,32	-11,19	-9,48
NE	-12,01	-10,58	-11,97	-11,40
SSE	-13,25	-10,02	-10,75	-13,55

No modelo quadrático a proporção subiu de forma inexpressiva saindo de 0,542 na linear atingindo o índice de 0,547 na quadrática. Decidimos escolher o método linear por sua simplicidade.

3.2.8. ANÁLISE DISCRIMINANTE LINEAR PARA DADOS AGRUPADOS (Habitação)

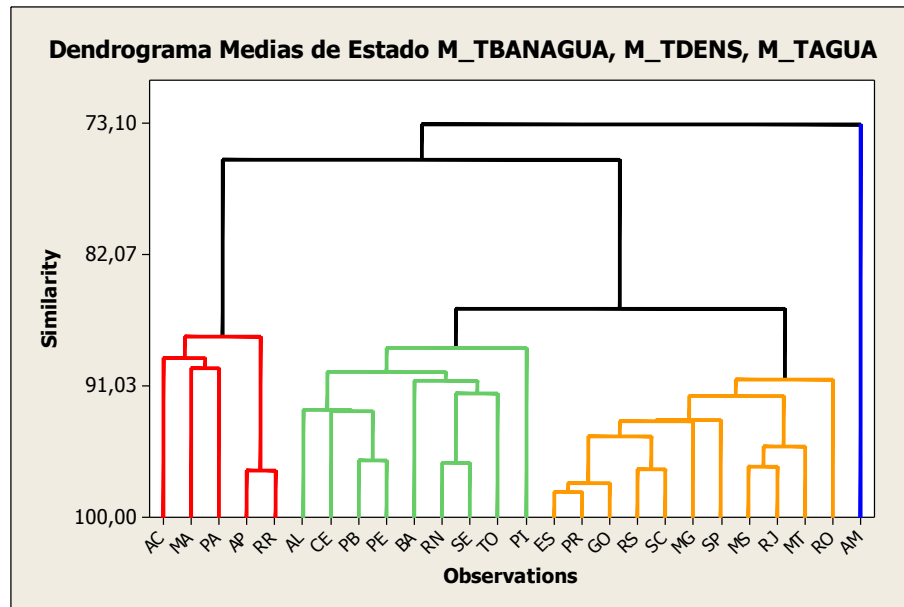


Grafico4. Dendrograma da Média T_BANAGUA x T_DENS x T_AGUA por grupos do Brasil (classificação supervisionada)

Neste exemplo abaixo vamos através do dendrograma pesquisar o grau de similaridade das médias das variáveis T_BANAGUA, T_DENS e T_AGUA nos agrupamentos. Com base na análise discriminante poderemos verificar a proporção correta dos agrupamentos.

Discriminant Analysis: Grupos versus M_TABANAGUA; M_TDENS; M_TAGUA (Amazonas incluído no G1, por não ser possível realizar análise apenas com 01 variável por agrupamento)

Linear Method for Response: Grupos

Predictors: M_TABANAGUA; M_TDENS; M_TAGUA

Group	G1	G2	G3
Count	6	9	11

Summary of classification

Put into Group	True Group		
	G1	G2	G3
G1	6	0	0
G2	0	9	0
G3	0	0	11
Total N	6	9	11
N correct	6	9	11
Proportion	1,000	1,000	1,000

N = 26

N Correct = 26

Proportion Correct = 1,000

Squared Distance Between Groups

	G1	G2	G3
G1	0,000	40,141	108,193
G2	40,141	0,000	21,348
G3	108,193	21,348	0,000

Linear Discriminant Function for Groups

	G1	G2	G3
Constant	-159,16	-122,85	-165,70
M_TABANAGUA	14,13	97,16	160,99
M_TDENS	220,38	148,55	120,80
M_TAGUA	226,58	168,25	168,88

No mapa abaixo apresentamos visualmente a divisão dos Grupos G1, G2 e G3 divididos por cores dos Estados, conforme o agrupamento por similaridade. Nesta representação observamos uma divisão bem estruturada não apresentando uma concentração significativa em determinado estado.



Mapa(3). Brasil por Agrupamentos – Média Variáveis T_BANAGUA, T_DENS e T_AGUA

Legenda: G1= Azul, G2= Laranja, G3= Verde

3.3 REGRESSÃO LOGÍSTICA

A regressão logística é um método estatístico que produz a partir de um conjunto de variáveis e observações a predição de valores. A seguir apresentamos a Regressão Logística para as Dimensões Demografia e Habitação, mantendo as variáveis que estão sendo utilizadas regularmente nesta apresentação.

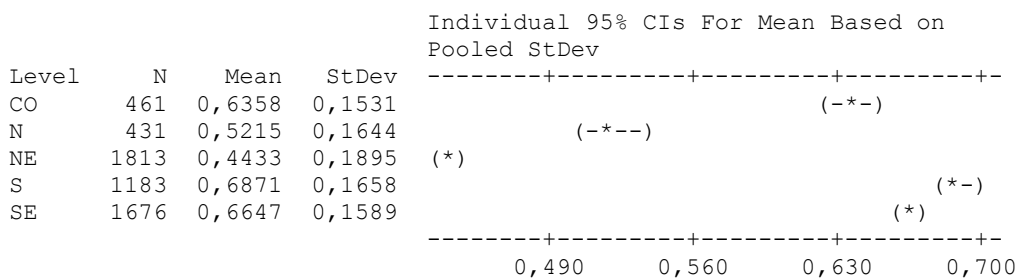
3.3.1 DIMENSÃO DEMOGRAFIA

3.3.1.1 ANOVA das Variáveis

One-way ANOVA: **ESPVIDA(1)** versus Região

Source	DF	SS	MS	F	P
Região	4	62,2950	15,5738	533,46	0,000
Error	5559	162,2892	0,0292		
Total	5563	224,5843			

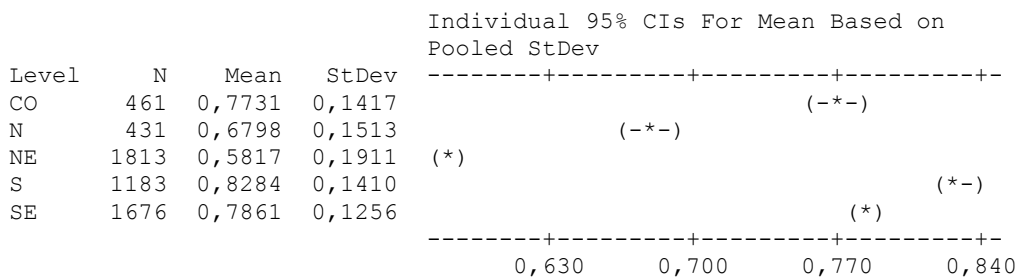
S = 0,1709 R-Sq = 27,74% R-Sq(adj) = 27,69%



One-way ANOVA: **MORT1(i)** versus Região

Source	DF	SS	MS	F	P
Região	4	57,8850	14,4712	595,09	0,000
Error	5559	135,1823	0,0243		
Total	5563	193,0673			

S = 0,1559 R-Sq = 29,98% R-Sq(adj) = 29,93%



Pooled StDev = 0,1559

One-way ANOVA: **MORT5(i)** versus Região

Source	DF	SS	MS	F	P
Região	4	48,3230	12,0807	518,80	0,000
Error	5559	129,4453	0,0233		
Total	5563	177,7683			

S = 0,1526 R-Sq = 27,18% R-Sq(adj) = 27,13%

Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev
CO	461	0,7550	0,1382
N	431	0,6917	0,1467
NE	1813	0,5920	0,1875
S	1183	0,8216	0,1373
SE	1676	0,7775	0,1231

0,630 0,700 0,770 0,840

Pooled StDev = 0,1526

Variável x Região	CO	N	NE	S	SE
ESPVIDA(1)	3	2	1	5	4
MORT1(i)	3	2	1	5	4
MORT5(i)	3	2	1	5	4
Total	9	6	3	15	12
Nova Ordem	3	2	1	5	4

Tabela(3) Sequencia de Regiões por Variável (Ordem Crescente) – ESPVIDA(1), MORT1(i) e MORT5(i)

Podemos observar na tabela (3) que nesta dimensão, as três variáveis apresentaram sequencia idêntica nos níveis, por região.

3.3.1.2 Regressão Logística (nova sequencia crescente, ordinal)

Ordinal Logistic Regression: C50 versus ESPVIDA(1); MORT1(i); MORT5(i)

Link Function: Logit

Response Information

Variable	Value	Count
C50	1	431
	2	1813
	3	461
	4	1676
	5	1183
Total		5564

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI	
						Lower	Upper
Const (1)	-0,0978639	0,141301	-0,69	0,489			
Const (2)	2,53302	0,148815	17,02	0,000			
Const (3)	3,01572	0,150426	20,05	0,000			
Const (4)	4,64684	0,155015	29,98	0,000			
ESPVIDA (1)	-0,227234	0,484388	-0,47	0,639	0,80		0,31
MORT1 (i)	-34,4719	1,68277	-20,49	0,000	0,00		0,00
MORT5 (i)	30,5612	1,66963	18,30	0,000	1,87312E+13		7,10191E+11

Predictor	Upper
Const (1)	
Const (2)	
Const (3)	
Const (4)	
ESPVIDA (1)	2,06
MORT1 (i)	0,00
MORT5 (i)	4,94031E+14

Log-Likelihood = -7247,892

Test that all slopes are zero: G = 1756,823, DF = 3, P-Value = 0,000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	22286,8	17065	0,000
Deviance	12997,0	17065	1,000

Measures of Association:

(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures
Concordant	8180499	70,9	Somers' D 0,43
Discordant	3245737	28,1	Goodman-Kruskal Gamma 0,43
Ties	105954	0,9	Kendall's Tau-a 0,32
Total	11532190	100,0	

Comparativamente aos resultados observados na Análise discriminante Linear podemos verificar um crescimento significativo no percentual de concordância, que passa de 49,40% na Análise Discriminante Linear para 70,90% na Regressão Logística. A regressão logística se apresentou de uma forma mais confiável no que se refere ao comportamento destas variáveis.

3.3.1.3 Regressão Logística por Agrupamento – “4 BRASIS”

Ordinal Logistic Regression: C51 versus ESPVIDA(1); MORT1(i); MORT5(i)

Link Function: Logit

Response Information

Variable	Value	Count
C51	1	431
	2	1813
	3	461
	4	2859
Total		5564

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI	
						Lower	Upper
Const (1)	-1,07770	0,151298	-7,12	0,000			
Const (2)	1,54314	0,154073	10,02	0,000			
Const (3)	2,02509	0,155205	13,05	0,000			
ESPVIDA (1)	-3,99952	0,563914	-7,09	0,000	0,02		0,01
MORT1 (i)	-36,9104	1,95594	-18,87	0,000	0,00		0,00
MORT5 (i)	37,4295	1,94537	19,24	0,000	1,80062E+16	3,97670E+14	

Predictor	Upper
Const (1)	
Const (2)	
Const (3)	
ESPVIDA (1)	0,06
MORT1 (i)	0,00
MORT5 (i)	8,15309E+17

Log-Likelihood = -5299,827

Test that all slopes are zero: G = 1774,976, DF = 3, P-Value = 0,000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	17605,6	12798	0,000
Deviance	9441,2	12798	1,000

Measures of Association:

(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures
Concordant	6971690	73,0	Somers' D 0,47
Discordant	2518497	26,4	Goodman-Kruskal Gamma 0,47
Ties	59295	0,6	Kendall's Tau-a 0,29
Total	9549482	100,0	

Comparativamente aos resultados observados na Análise discriminante Linear podemos verificar um crescimento significativo no percentual de concordância, que passa de 55,40% na Análise Discriminante Linear para 73,00% na Regressão Logística. A regressão logística se apresentou de uma forma mais confiável no que se refere ao comportamento destas variáveis.

3.3.1.3 Regressão Logística para Dados Agrupados (Média-“4 Brasis)

Link Function: Logit

Response Information

Variable	Value	Count
Agrup_n	1	10
	2	2
	3	4
	4	10
Total		26

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI	
						Lower	Upper
Const (1)	1396,95	6776,78	0,21	0,837			
Const (2)	1436,23	6955,83	0,21	0,836			
Const (3)	1506,62	50268,2	0,03	0,976			
M_ESPVIDA	7329,04	36110,0	0,20	0,839	*	0,00	*
M_MORT1i	-134920	655328	-0,21	0,837	0,00	0,00	*
M_MORT5i	123725	600888	0,21	0,837	*	0,00	*

Log-Likelihood = -0,000
 Test that all slopes are zero: G = 63,454, DF = 3, P-Value = 0,000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	0,0001491	72	1,000
Deviance	0,0002982	72	1,000

Measures of Association:
 (Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures
Concordant	228	100,0	Somers' D 1,00
Discordant	0	0,0	Goodman-Kruskal Gamma 1,00
Ties	0	0,0	Kendall's Tau-a 0,70
Total	228	100,0	

Comparativamente aos resultados observados na Análise discriminante Linear podemos verificar o resultado idêntico nas duas situações com percentual de concordância, em ambos os casos, de 100%.

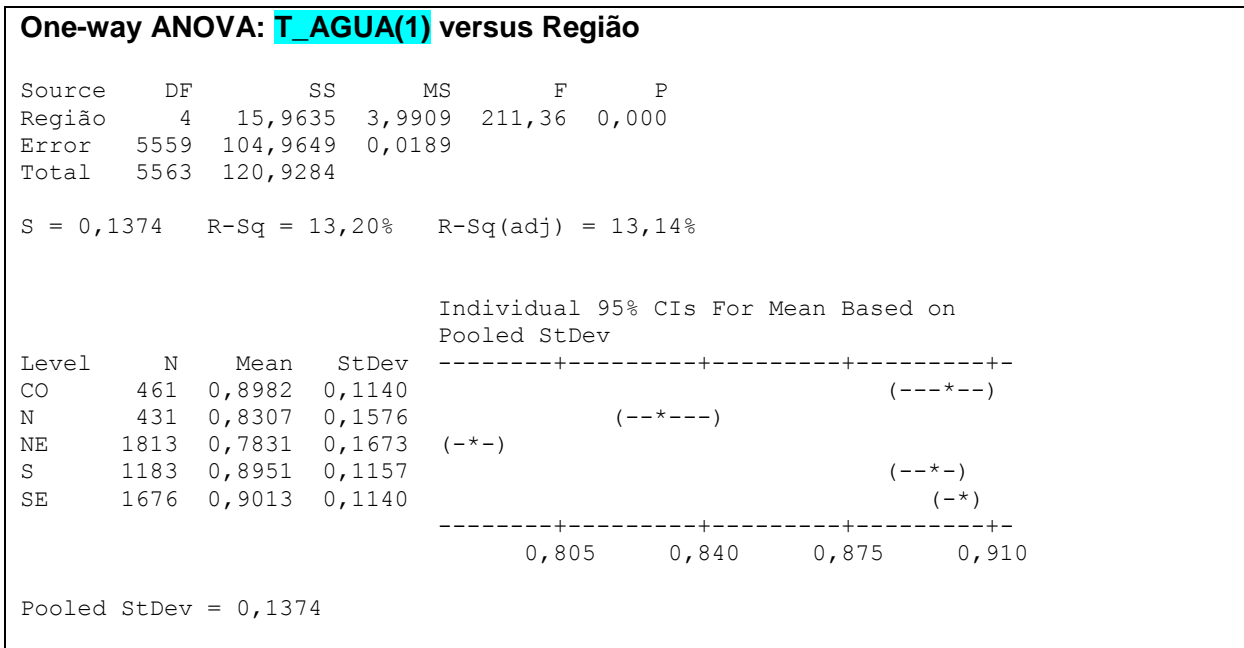
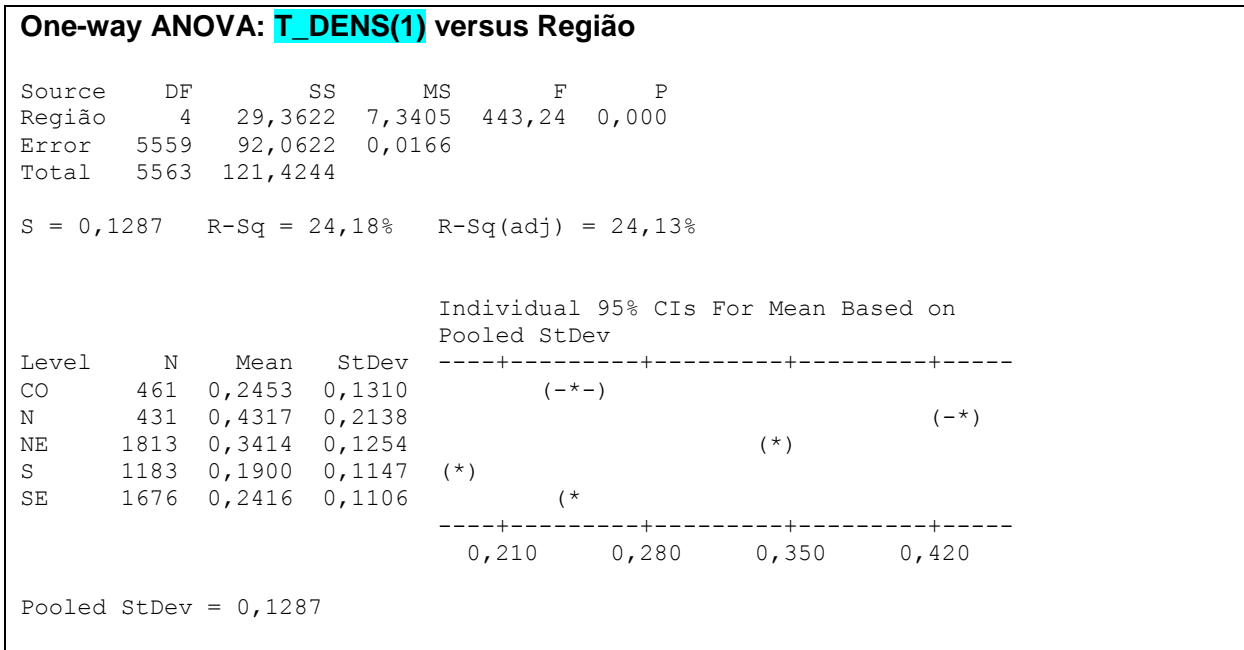
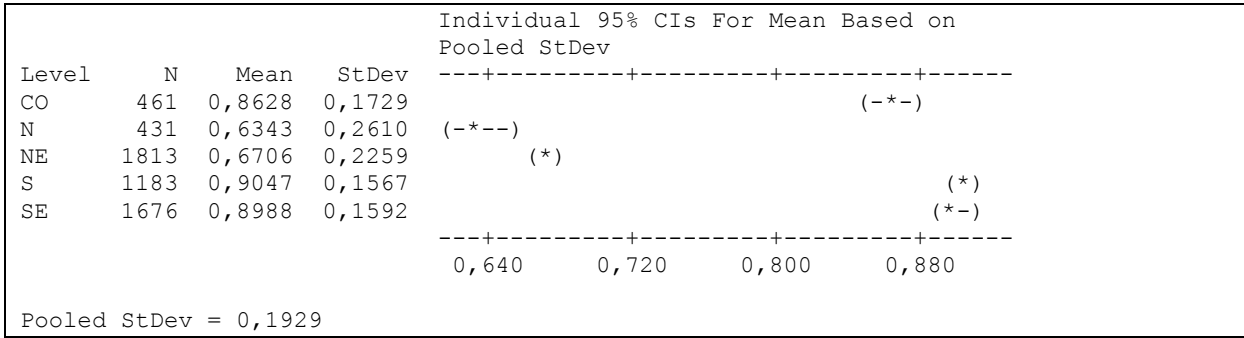
3.3.2 DIMENSÃO HABITAÇÃO

3.3.2.1 ANOVA das Variáveis

One-way ANOVA: T_BANAGUA(1) versus Região

Source	DF	SS	MS	F	P
Região	4	73,3161	18,3290	492,37	0,000
Error	5559	206,9401	0,0372		
Total	5563	280,2562			

S = 0,1929 R-Sq = 26,16% R-Sq(adj) = 26,11%



Variável x Região	CO	N	NE	S	SE
T_BANAGUA	3	1	2	5	4
T_DENS	3	5	4	1	2
T_AGUA	4	2	1	3	5
Total	10	8	7	9	11
Nova Ordem	4	2	1	3	5

Tabela(4) Sequencia de Regiões por Variável (Ordem Crescente) – T_BANAGUA, T_DENS e T_AGUA

Podemos observar na tabela (4) que nesta dimensão as três variáveis apresentaram sequencias diferenciadas nos níveis, por região. Desta forma por meio do calculo de média aritmética encontramos uma nova ordem para a composição da análise das três variáveis.

3.3.2.2 Regressão Logística (nova sequência crescente, ordinal)

Ordinal Logistic Regression: Prioridade versus T_BANAGUA(1); T_AGUA(1); ...

Link Function: Logit

Response Information

Variable	Value	Count
Prioridade	1	1813
	2	431
	3	1183
	4	461
	5	1676
Total		5564

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI	
						Lower	Upper
Const (1)	3,77944	0,240415	15,72	0,000			
Const (2)	4,20173	0,241730	17,38	0,000			
Const (3)	5,24366	0,245043	21,40	0,000			
Const (4)	5,65384	0,246156	22,97	0,000			
T_BANAGUA (1)	-4,43771	0,245216	-18,10	0,000	0,01	0,01	0,02
T_AGUA (1)	-0,811576	0,276926	-2,93	0,003	0,44	0,26	0,76
T_DENS (1)	-1,24265	0,273567	-4,54	0,000	0,29	0,17	0,49

Log-Likelihood = -7521,262

Test that all slopes are zero: G = 1210,084, DF = 3, P-Value = 0,000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	23396,0	22249	0,000
Deviance	15042,5	22249	1,000

Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures
Concordant	8210459	71,2	Somers' D 0,43
Discordant	3257400	28,2	Goodman-Kruskal Gamma 0,43
Ties	64331	0,6	Kendall's Tau-a 0,32
Total	11532190	100,0	

Comparativamente aos resultados observados na Análise Discriminante Linear podemos verificar um crescimento significativo no percentual de concordância, que passa de 43,60% na Análise Discriminante para 71,20% na Regressão Logística. A Regressão Logística se apresentou de uma forma mais confiável no que se refere ao comportamento destas variáveis.

3.3.2.3 Regressão Logística por Agrupamento – “4 BRASIS”

Ordinal Logistic Regression: Agrup_1 versus T_BANAGUA(1); T_AGUA(1); ...

Link Function: Logit

Response Information

Variable	Value	Count
Agrup_1	1	1813
	2	431
	3	2859
	4	461
Total		5564

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI	
						Lower	Upper
Const (1)	2,62865	0,244961	10,73	0,000			
Const (2)	3,05301	0,245971	12,41	0,000			
Const (3)	6,20241	0,254363	24,38	0,000			
T_BANAGUA (1)	-3,18278	0,251135	-12,67	0,000	0,04	0,03	0,07
T_AGUA (1)	-1,35271	0,288570	-4,69	0,000	0,26	0,15	0,46
T_DENS (1)	0,868627	0,293904	2,96	0,003	2,38	1,34	4,24

Log-Likelihood = -5586,663

Test that all slopes are zero: G = 1201,303, DF = 3, P-Value = 0,000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	20141,2	16686	0,000
Deviance	11173,3	16686	1,000

Measures of Association:
(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures
Concordant	6918601	72,5	Somers' D 0,46
Discordant	2570228	26,9	Goodman-Kruskal Gamma 0,46
Ties	60653	0,6	Kendall's Tau-a 0,28

Total	9549482	100,0
-------	---------	-------

Comparativamente aos resultados observados na Análise Discriminante Linear podemos verificar um crescimento significativo no percentual de concordância, que passa de 54,20% na Análise Discriminante para 72,5% na Regressão Logística. A Regressão Logística se apresentou de uma forma mais confiável no que se refere ao comportamento destas variáveis.

3.3.2.4 Regressão Logística para Dados Agrupados (Média – “3 BRASIS” – junção Amazonas”

Ordinal Logistic Regression: Grup2 versus M_TABANAGUA; M_TDENS; M_TAGUA

Link Function: Logit

Response Information

Variable	Value	Count
Grup2	1	6
	2	9
	3	11
Total		26

Logistic Regression Table

Predictor	Coef	SE Coef	Z	P	Odds Ratio	95% CI	
						Lower	Upper
Const (1)	-3,19425	53529,3	-0,00	1,000			
Const (2)	71,9873	56230,0	0,00	0,999			
M_TABANAGUA	-105,304	29523,4	-0,00	0,997	0,00	0,00	*
M_TDENS	229,969	53299,6	0,00	0,997	7,48755E+99	0,00	*
M_TAGUA	-69,0616	31675,6	-0,00	0,998	0,00	0,00	*

Log-Likelihood = -0,000

Test that all slopes are zero: G = 55,616, DF = 3, P-Value = 0,000

Goodness-of-Fit Tests

Method	Chi-Square	DF	P
Pearson	0,0000002	47	1,000
Deviance	0,0000004	47	1,000

Measures of Association:

(Between the Response Variable and Predicted Probabilities)

Pairs	Number	Percent	Summary Measures
Concordant	219	100,0	Somers' D 1,00
Discordant	0	0,0	Goodman-Kruskal Gamma 1,00
Ties	0	0,0	Kendall's Tau-a 0,67
Total	219	100,0	

Comparativamente aos resultados observados na Análise discriminante linear podemos verificar o resultado idêntico nas duas situações com percentual de concordância, em ambos os casos, de 100%.

3.3 Classification Tree

3.3.1 DIMENSÃO DEMOGRAFIA

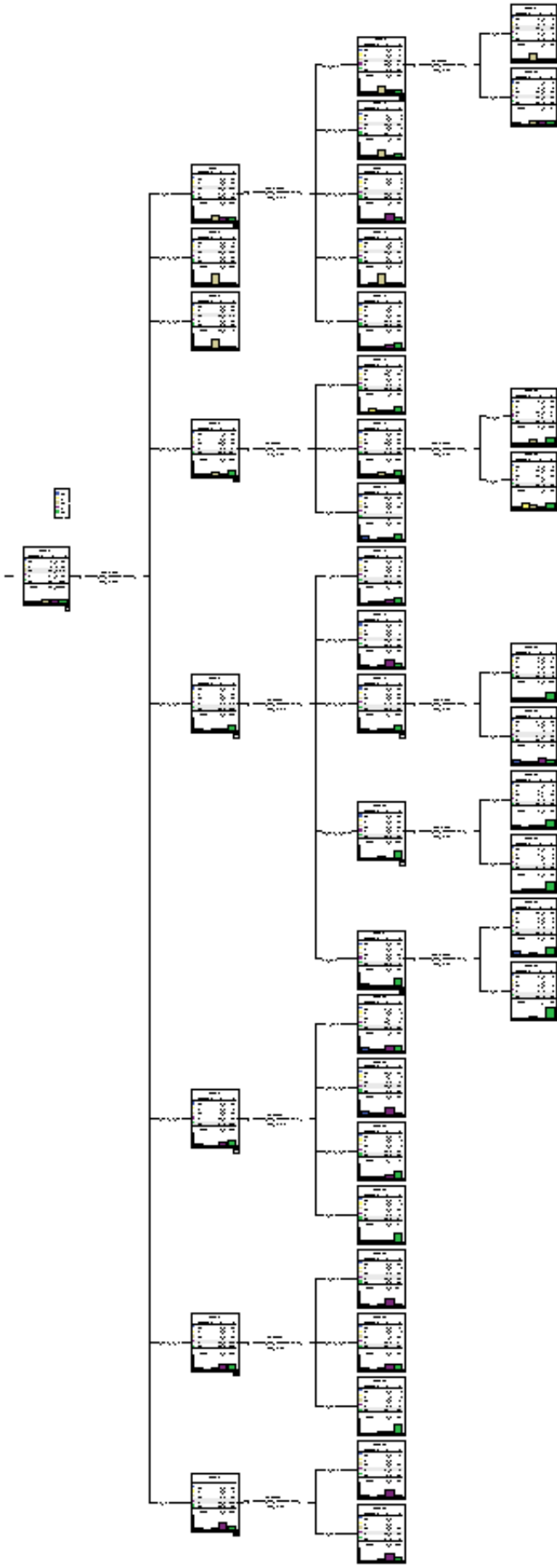
Classification Tree

Warnings

Gain summary Tables are not displayed because profits are undefined.
 Target category gains tables are not displayed because target categories are undefined.

Model Summary

Specifications	Growing Method	CHAID	
	Dependent Variable	Região	
	Independent Variables	ESPVIDAn, MORTum, MORTcinco	
	Validation	None	
	Maximum Tree Depth		3
	Minimum Cases in Parent Node		100
	Minimum Cases in Child Node		50
Results	Independent Variables Included	MORTum, MORTcinco, ESPVIDAn	
	Number of Nodes		41
	Number of Terminal Nodes		29
	Depth		3



Risk	
Estimate	Std. Error
,430	,007

Growing Method: CHAID

Dependent Variable: Região

Observed	Predicted					
	CO	N	NE	S	SE	Percent Correct
CO	0	0	79	162	220	,0%
N	0	0	198	50	183	,0%
NE	0	0	1292	177	344	71,3%
S	0	0	131	742	310	62,7%
SE	0	0	230	307	1139	68,0%
Overall Percentage	,0%	,0%	34,7%	25,8%	39,5%	57,0%

Growing Method: CHAID

Dependent Variable: Região

3.3.2 DIMENSÃO HABITAÇÃO

Classification Tree

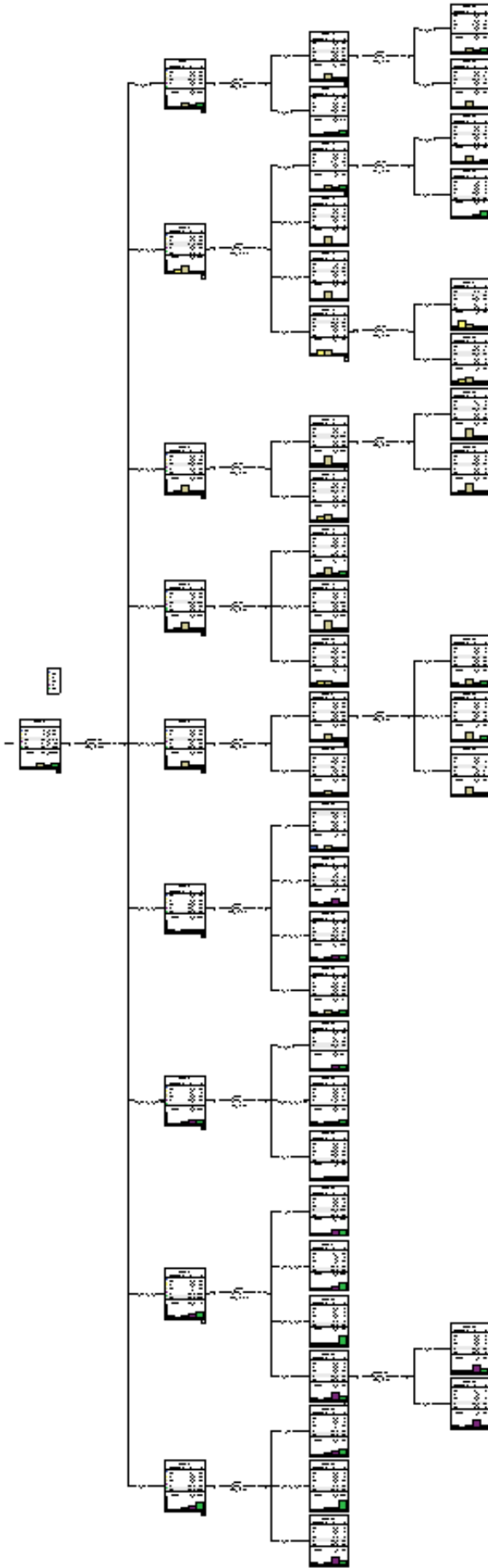
Warnings

Gain summary Tables are not displayed because profits are undefined.

Target category gains tables are not displayed because target categories are undefined.

Model Summary

Specifications	Growing Method	CHAID
	Dependent Variable	Região
	Independent Variables	T_AGUA, T_BANAGUA, T_DENS
	Validation	None
	Maximum Tree Depth	3
Results	Minimum Cases in Parent Node	100
	Minimum Cases in Child Node	50
	Independent Variables Included	T_BANAGUA, T_AGUA, T_DENS
	Number of Nodes	50
	Number of Terminal Nodes	34
	Depth	3



Risk	
Estimate	Std. Error
,491	,007

Growing Method: CHAID

Dependent Variable: Região

Observed	Classification					
	Predicted					
	CO	N	NE	S	SE	Percent Correct
CO	16	3	128	163	151	3,5%
N	4	45	260	47	75	10,4%
NE	15	25	1324	177	272	73,0%
S	11	8	218	598	348	50,5%
SE	12	7	352	458	847	50,5%
Overall Percentage	1,0%	1,6%	41,0%	25,9%	30,4%	50,9%

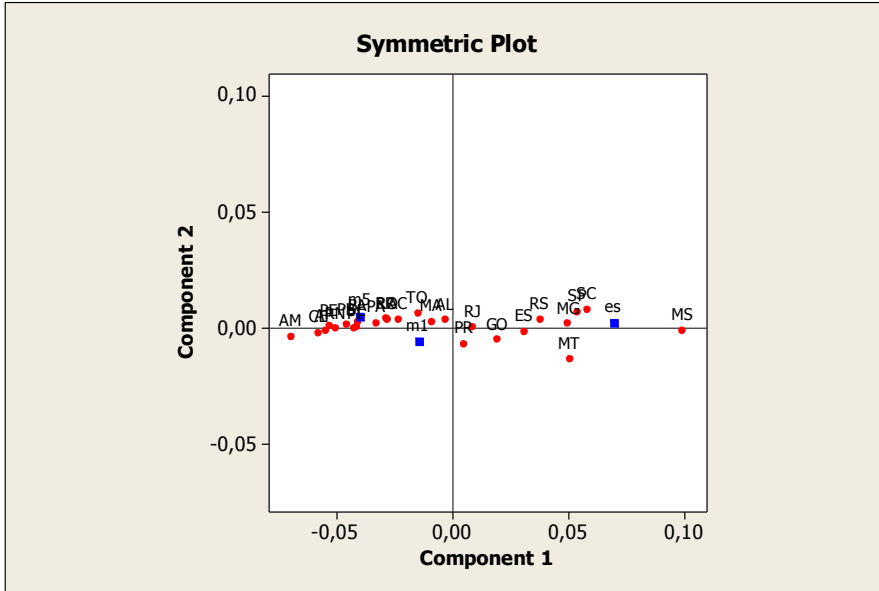
Growing Method: CHAID

Dependent Variable: Região

3.4 Análise de Correspondência

3.4.1 Dimensão **Demografia** por Estado

3.4.1.1 Análise de Correspondência da Média por Estado das Variáveis M_ESPVIDA; M_MORT1i; M_MORT5i



Simple Correspondence Analysis: M_ESPVIDA; M_MORT1i; M_MORT5i

Relative Inertias

	es	m1	m5	Total
AC	0,005	0,000	0,003	0,009
AL	0,000	0,000	0,000	0,000
AM	0,057	0,005	0,020	0,081
AP	0,040	0,003	0,015	0,059
BA	0,016	0,001	0,008	0,025
CE	0,036	0,003	0,013	0,052
ES	0,015	0,000	0,008	0,023
GO	0,006	0,000	0,005	0,010
MA	0,001	0,000	0,000	0,001
MG	0,039	0,003	0,015	0,057
MS	0,139	0,007	0,061	0,206
MT	0,033	0,000	0,027	0,060
PA	0,012	0,000	0,006	0,019
PB	0,020	0,001	0,009	0,030
PE	0,028	0,001	0,012	0,041
PI	0,016	0,001	0,006	0,024
PR	0,000	0,001	0,001	0,002
RJ	0,001	0,000	0,001	0,002
RN	0,027	0,002	0,011	0,040
RO	0,010	0,000	0,005	0,015
RR	0,011	0,000	0,006	0,017
RS	0,026	0,002	0,009	0,037
SC	0,063	0,007	0,018	0,089
SE	0,016	0,001	0,007	0,024
SP	0,051	0,006	0,015	0,072
TO	0,002	0,000	0,002	0,005
Total	0,671	0,044	0,285	1,000

Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0020	0,9889	0,9889	*****
2	0,0000	0,0111	1,0000	
Total	0,0020			

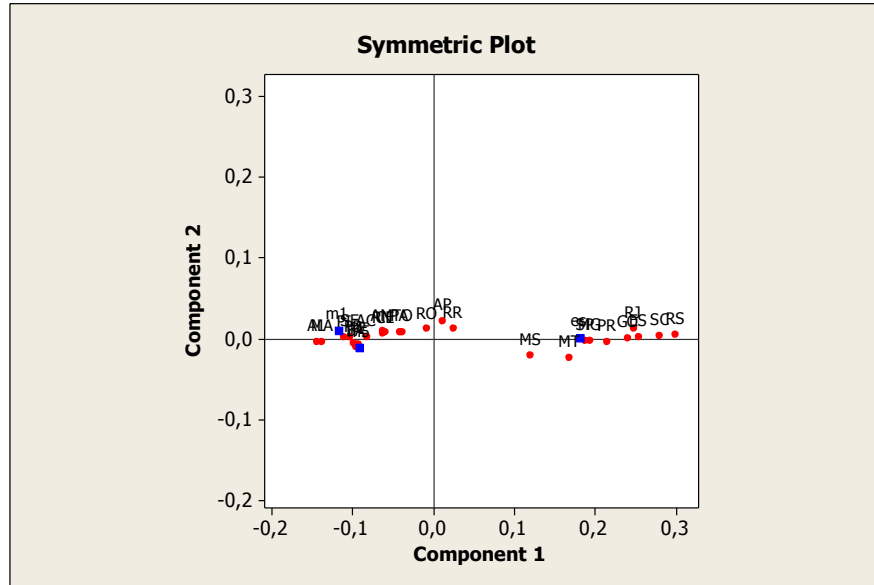
Row Contributions

ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	AC	1,000	0,032	0,009	-0,023	0,983	0,008	0,003	0,017	0,013
2	AL	1,000	0,022	0,000	-0,003	0,513	0,000	0,003	0,487	0,008
3	AM	1,000	0,034	0,081	-0,069	0,996	0,082	-0,004	0,004	0,026
4	AP	1,000	0,040	0,059	-0,055	0,999	0,059	-0,002	0,001	0,006
5	BA	1,000	0,031	0,025	-0,040	0,997	0,025	0,002	0,003	0,007
6	CE	1,000	0,032	0,052	-0,057	0,998	0,053	-0,003	0,002	0,009
7	ES	1,000	0,048	0,023	0,031	0,995	0,023	-0,002	0,005	0,010
8	GO	1,000	0,049	0,010	0,020	0,927	0,010	-0,006	0,073	0,067
9	MA	1,000	0,023	0,001	-0,009	0,951	0,001	0,002	0,049	0,004
10	MG	1,000	0,046	0,057	0,050	0,999	0,058	0,001	0,001	0,004
11	MS	1,000	0,042	0,206	0,099	1,000	0,208	-0,002	0,000	0,005
12	MT	1,000	0,044	0,060	0,051	0,929	0,057	-0,014	0,071	0,386
13	PA	1,000	0,035	0,019	-0,033	0,997	0,019	0,002	0,003	0,005
14	PB	1,000	0,030	0,030	-0,045	1,000	0,030	0,001	0,000	0,001
15	PE	1,000	0,030	0,041	-0,053	1,000	0,042	0,000	0,000	0,000
16	PI	1,000	0,027	0,024	-0,042	1,000	0,024	-0,001	0,000	0,001
17	PR	1,000	0,048	0,002	0,005	0,313	0,001	-0,007	0,687	0,118
18	RJ	1,000	0,047	0,002	0,009	0,999	0,002	-0,000	0,001	0,000
19	RN	1,000	0,032	0,040	-0,050	1,000	0,040	-0,001	0,000	0,001
20	RO	1,000	0,040	0,015	-0,028	0,986	0,015	0,003	0,014	0,019
21	RR	1,000	0,043	0,017	-0,028	0,985	0,017	0,003	0,015	0,023
22	RS	1,000	0,052	0,037	0,038	0,994	0,038	0,003	0,006	0,021
23	SC	1,000	0,052	0,089	0,058	0,985	0,088	0,007	0,015	0,123
24	SE	1,000	0,029	0,024	-0,041	1,000	0,025	-0,000	0,000	0,000
25	SP	1,000	0,049	0,072	0,054	0,986	0,072	0,006	0,014	0,088
26	TO	1,000	0,039	0,005	-0,015	0,869	0,004	0,006	0,131	0,057

Column Contributions

ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	es	1,000	0,281	0,671	0,070	0,999	0,678	0,002	0,001	0,041
2	m1	1,000	0,358	0,044	-0,015	0,847	0,038	-0,006	0,153	0,604
3	m5	1,000	0,361	0,285	-0,040	0,986	0,284	0,005	0,014	0,355

3.4.1.2 Análise de Correspondência do Desvio Padrão por Estado das Variáveis M_ESPVIDA; M_MORT1i; M_MORT5i



Simple Correspondence Analysis: DP_ESPVIDA; DP_MORT1i; DP_MORT5i

Relative Inertias

	es	m1	m5	Total
AC	0,010	0,004	0,002	0,017
AL	0,042	0,014	0,011	0,066
AM	0,006	0,003	0,001	0,009
AP	0,000	0,000	0,001	0,001
BA	0,016	0,004	0,005	0,025
CE	0,004	0,002	0,001	0,007
ES	0,043	0,015	0,010	0,069
GO	0,036	0,013	0,008	0,056
MA	0,033	0,011	0,008	0,052
MG	0,046	0,017	0,010	0,073
MS	0,015	0,009	0,001	0,026
MT	0,021	0,011	0,003	0,035
PA	0,002	0,001	0,000	0,003
PB	0,018	0,005	0,005	0,028
PE	0,017	0,005	0,005	0,027
PI	0,020	0,007	0,004	0,032
PR	0,039	0,015	0,008	0,062
RJ	0,043	0,014	0,011	0,068
RN	0,006	0,003	0,001	0,009
RO	0,000	0,000	0,000	0,000
RR	0,001	0,000	0,000	0,001
RS	0,074	0,025	0,018	0,117
SC	0,087	0,030	0,020	0,137
SE	0,017	0,007	0,004	0,027
SP	0,032	0,012	0,007	0,050
TO	0,003	0,002	0,000	0,004
Total	0,629	0,228	0,143	1,000

Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0190	0,9958	0,9958	*****
2	0,0001	0,0042	1,0000	
Total	0,0190			

Row Contributions										
ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	AC	1,000	0,048	0,017	-0,081	1,000	0,017	0,001	0,000	0,000
2	AL	1,000	0,062	0,066	-0,143	0,999	0,066	-0,005	0,001	0,017
3	AM	1,000	0,045	0,009	-0,062	0,984	0,009	0,008	0,016	0,036
4	AP	1,000	0,030	0,001	0,013	0,274	0,000	0,021	0,726	0,172
5	BA	1,000	0,053	0,025	-0,094	0,988	0,025	-0,010	0,012	0,072
6	CE	1,000	0,037	0,007	-0,058	0,987	0,007	0,007	0,013	0,021
7	ES	1,000	0,020	0,069	0,255	1,000	0,069	0,001	0,000	0,000
8	GO	1,000	0,018	0,056	0,242	1,000	0,057	-0,000	0,000	0,000
9	MA	1,000	0,053	0,052	-0,136	0,999	0,052	-0,005	0,001	0,015
10	MG	1,000	0,037	0,073	0,195	1,000	0,073	-0,003	0,000	0,003
11	MS	1,000	0,032	0,026	0,121	0,969	0,025	-0,022	0,031	0,192
12	MT	1,000	0,023	0,035	0,169	0,980	0,034	-0,024	0,020	0,166
13	PA	1,000	0,035	0,003	-0,041	0,972	0,003	0,007	0,028	0,021
14	PB	1,000	0,054	0,028	-0,098	0,996	0,028	-0,006	0,004	0,024
15	PE	1,000	0,061	0,027	-0,091	0,992	0,027	-0,008	0,008	0,049
16	PI	1,000	0,050	0,032	-0,109	1,000	0,032	0,001	0,000	0,000
17	PR	1,000	0,025	0,062	0,216	0,999	0,062	-0,005	0,001	0,008
18	RJ	1,000	0,021	0,068	0,248	0,998	0,068	0,011	0,002	0,034
19	RN	1,000	0,044	0,009	-0,062	0,991	0,009	0,006	0,009	0,020
20	RO	1,000	0,038	0,000	-0,007	0,276	0,000	0,011	0,724	0,058
21	RR	1,000	0,023	0,001	0,026	0,835	0,001	0,011	0,165	0,037
22	RS	1,000	0,025	0,117	0,301	1,000	0,117	0,004	0,000	0,006
23	SC	1,000	0,033	0,137	0,281	1,000	0,137	0,002	0,000	0,002
24	SE	1,000	0,050	0,027	-0,102	1,000	0,027	0,002	0,000	0,002
25	SP	1,000	0,027	0,050	0,188	1,000	0,051	-0,003	0,000	0,003
26	TO	1,000	0,056	0,004	-0,038	0,961	0,004	0,008	0,039	0,040

Column Contributions										
ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	es	1,000	0,365	0,629	0,181	1,000	0,632	0,001	0,000	0,004
2	m1	1,000	0,313	0,228	-0,117	0,992	0,227	0,011	0,008	0,460
3	m5	1,000	0,322	0,143	-0,091	0,984	0,142	-0,011	0,016	0,536

3.4.2 Dimensão **Demografia** por Região

3.4.2.1 ANOVA por região das Variáveis M_ESPVIDA; M_MORT1i; M_MORT5i

One-way ANOVA: M_ESPVIDA versus REGIAO_N

Source	DF	SS	MS	F	P
REGIAO_N	4	0,54443	0,13611	66,61	0,000
Error	21	0,04291	0,00204		
Total	25	0,58734			

S = 0,04520 R-Sq = 92,69% R-Sq(adj) = 91,30%

Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev
CO	3	0,67313	0,01835
N	7	0,49153	0,05601
NE	9	0,36724	0,03815
S	3	0,73857	0,06200
SE	4	0,69105	0,03689

0,36 0,48 0,60 0,72

Pooled StDev = 0,04520

One-way ANOVA: M_MORT1i versus REGIAO_N

Source	DF	SS	MS	F	P
REGIAO_N	4	0,51157	0,12789	34,75	0,000
Error	21	0,07728	0,00368		
Total	25	0,58885			

S = 0,06066 R-Sq = 86,88% R-Sq(adj) = 84,38%

Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev
CO	3	0,78608	0,06832
N	7	0,66889	0,06756
NE	9	0,50543	0,06887
S	3	0,88203	0,02301
SE	4	0,82244	0,02276

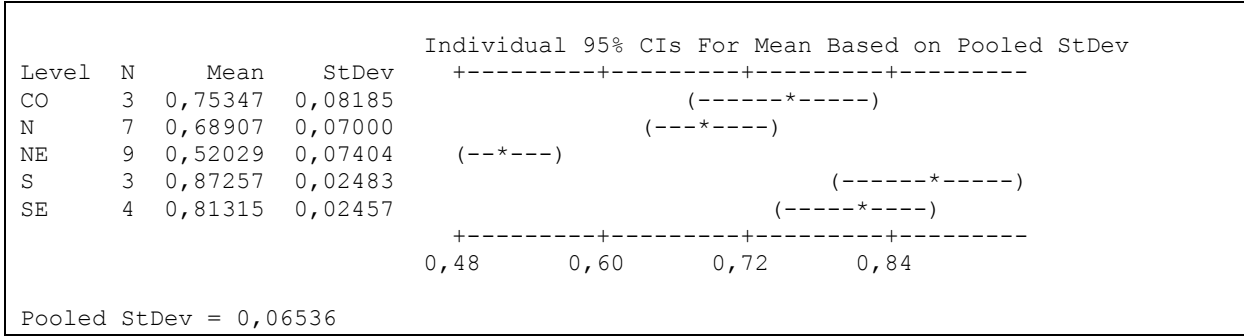
0,60 0,75 0,90 1,05

Pooled StDev = 0,06066

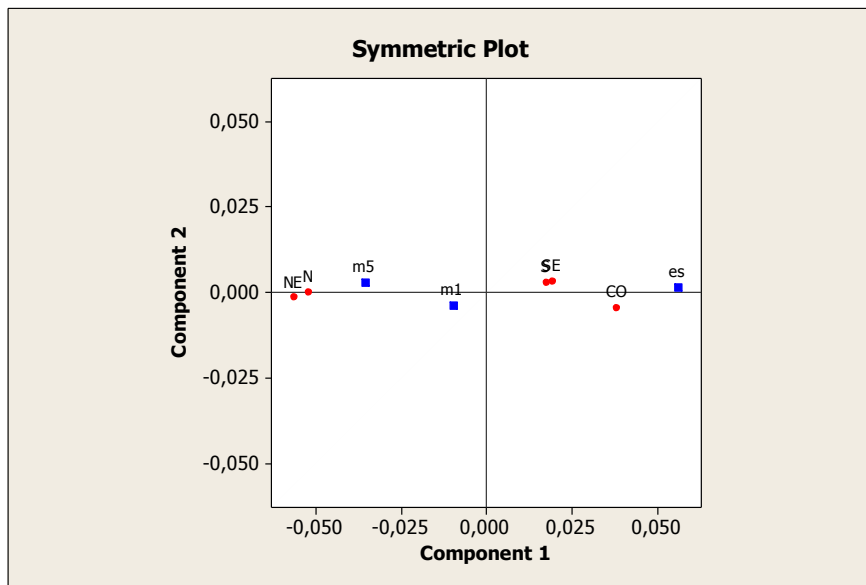
One-way ANOVA: M_MORT5i versus REGIAO_N

Source	DF	SS	MS	F	P
REGIAO_N	4	0,42843	0,10711	25,07	0,000
Error	21	0,08970	0,00427		
Total	25	0,51813			

S = 0,06536 R-Sq = 82,69% R-Sq(adj) = 79,39%



3.4.2.2 Análise de Correspondência da Média por região das Variáveis M_ESPVIDA; M_MORT1i; M_MORT5i



Symmetric Plot

Simple Correspondence Analysis: M_ESPVIDA1_reg; M_MORT1i_reg; M_MORT5i_reg

Relative Inertias

	es	m1	m5	Total
CO	0,137	0,001	0,092	0,230
N	0,226	0,009	0,111	0,346
NE	0,204	0,010	0,093	0,307
S	0,039	0,004	0,013	0,056
SE	0,043	0,005	0,014	0,062
Total	0,648	0,029	0,323	1,000

Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0014	0,9937	0,9937	*****
2	0,0000	0,0063	1,0000	
Total	0,0014			

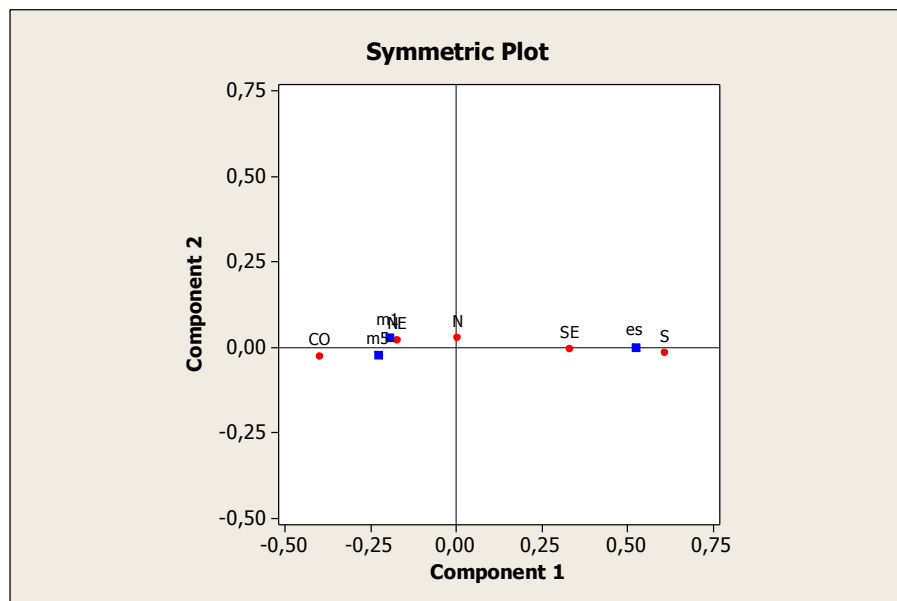
Row Contributions

ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	CO	1,000	0,215	0,230	0,038	0,985	0,228	-0,005	0,015	0,547
2	N	1,000	0,180	0,346	-0,052	1,000	0,348	-0,000	0,000	0,000
3	NE	1,000	0,136	0,307	-0,056	0,999	0,308	-0,002	0,001	0,049
4	S	1,000	0,243	0,056	0,018	0,979	0,055	0,003	0,021	0,189
5	SE	1,000	0,226	0,062	0,019	0,978	0,061	0,003	0,022	0,215

Column Contributions

ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	es	1,000	0,288	0,648	0,056	0,999	0,652	0,001	0,001	0,060
2	m1	1,000	0,357	0,029	-0,010	0,866	0,025	-0,004	0,134	0,618
3	m5	1,000	0,355	0,323	-0,036	0,994	0,323	0,003	0,006	0,322

3.4.2.3 Análise de Correspondência do Desvio Padrão por região das Variáveis M_ESPVIDA; M_MORT1i; M_MORT5i



Simple Correspondence Analysis: DP_ESPVIDA1_reg; DP_MORT1i_reg; DP_MORT5i_reg

Relative Inertias

	es	m1	m5	Total
CO	0,225	0,026	0,069	0,320
N	0,000	0,001	0,001	0,002
NE	0,044	0,011	0,007	0,062
S	0,358	0,067	0,077	0,502

SE	0,081	0,015	0,018	0,114
Total	0,709	0,118	0,173	1,000

Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,1114	0,9956	0,9956	*****
2	0,0005	0,0044	1,0000	
Total	0,1119			

Row Contributions

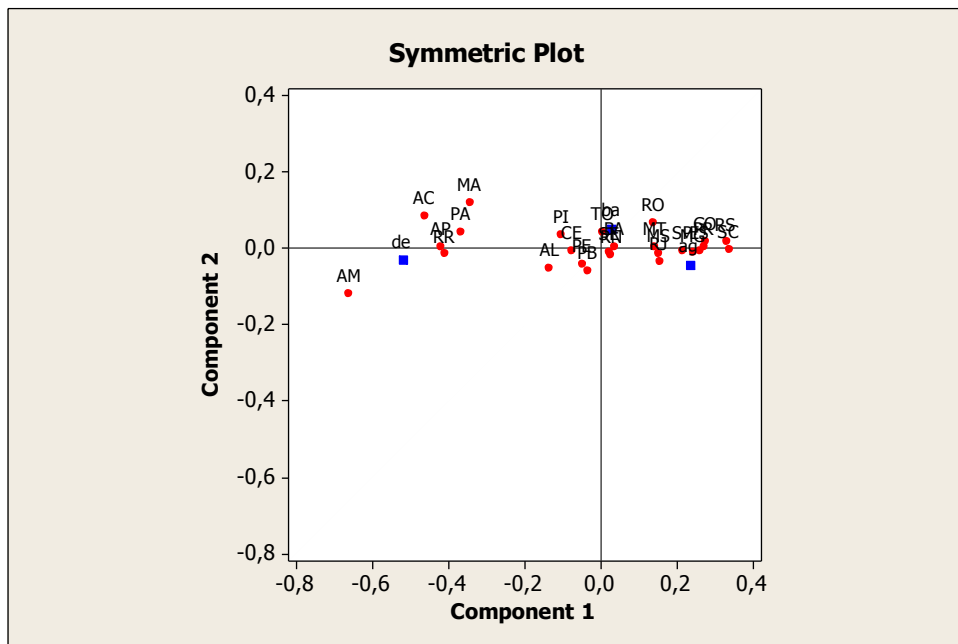
ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	CO	1,000	0,229	0,320	-0,395	0,994	0,320	-0,030	0,006	0,415
2	N	1,000	0,263	0,002	0,007	0,070	0,000	0,024	0,930	0,319
3	NE	1,000	0,246	0,062	-0,168	0,990	0,062	0,017	0,010	0,141
4	S	1,000	0,149	0,502	0,614	0,999	0,504	-0,019	0,001	0,114
5	SE	1,000	0,114	0,114	0,334	1,000	0,115	-0,007	0,000	0,012

Column Contributions

ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	es	1,000	0,287	0,709	0,526	1,000	0,712	-0,001	0,000	0,001
2	m1	1,000	0,340	0,118	-0,195	0,980	0,116	0,028	0,020	0,544
3	m5	1,000	0,373	0,173	-0,226	0,988	0,172	-0,025	0,012	0,455

3.4.3 Dimensão Habitação por Estado

3.4.3.1 Análise de Correspondência da Média por estado das Variáveis M_TBANAGUA; M_TAGUA; M_TDENS



Simple Correspondence Analysis: M_TABANAGUA; M_TDENS; M_TAGUA

Relative Inertias

	ag	de	ba	Total
AC	0,039	0,061	0,000	0,101
AL	0,001	0,008	0,001	0,010
AM	0,035	0,160	0,008	0,204
AP	0,026	0,067	0,000	0,093
BA	0,000	0,001	0,000	0,001
CE	0,001	0,002	0,000	0,003
ES	0,012	0,027	0,000	0,039
GO	0,010	0,032	0,001	0,043
MA	0,029	0,029	0,002	0,059
MG	0,010	0,022	0,000	0,033
MS	0,005	0,009	0,000	0,014
MT	0,003	0,009	0,000	0,012
PA	0,023	0,043	0,000	0,066
PB	0,000	0,001	0,001	0,002
PE	0,000	0,001	0,001	0,002
PI	0,002	0,003	0,000	0,005
PR	0,012	0,030	0,000	0,042
RJ	0,006	0,009	0,000	0,015
RN	0,000	0,000	0,000	0,001
RO	0,001	0,010	0,002	0,013
RR	0,023	0,067	0,001	0,091
RS	0,015	0,044	0,001	0,060
SC	0,018	0,044	0,000	0,062
SE	0,000	0,000	0,000	0,000
SP	0,009	0,019	0,000	0,028
TO	0,000	0,000	0,001	0,001
Total	0,283	0,698	0,019	1,000

Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0720	0,9735	0,9735	*****
2	0,0020	0,0265	1,0000	
Total	0,0740			

Row Contributions

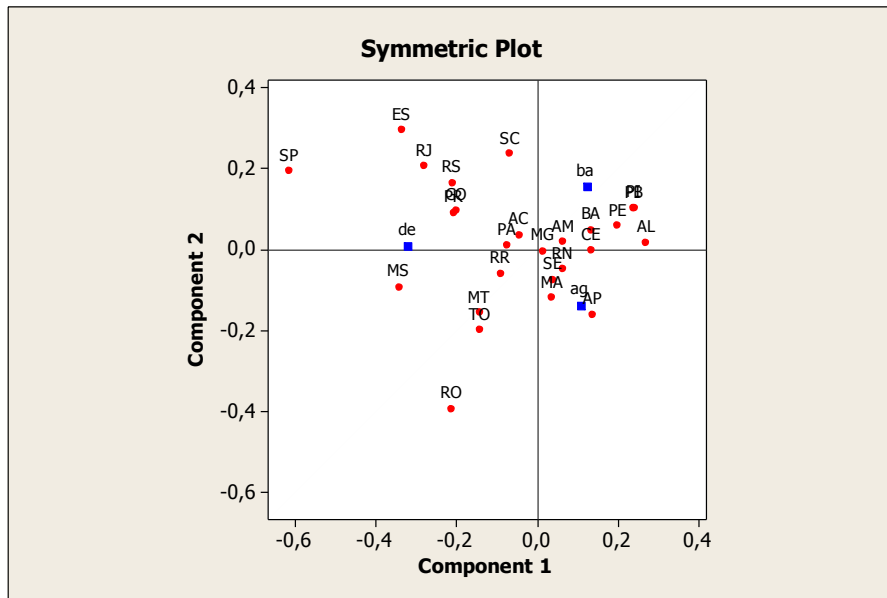
ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	AC	1,000	0,034	0,101	-0,459	0,968	0,100	0,083	0,032	0,120
2	AL	1,000	0,035	0,010	-0,132	0,862	0,009	-0,053	0,138	0,050
3	AM	1,000	0,034	0,204	-0,657	0,968	0,203	-0,120	0,032	0,249
4	AP	1,000	0,040	0,093	-0,417	1,000	0,096	0,001	0,000	0,000
5	BA	1,000	0,035	0,001	0,040	1,000	0,001	0,001	0,000	0,000
6	CE	1,000	0,036	0,003	-0,075	0,989	0,003	-0,008	0,011	0,001
7	ES	1,000	0,042	0,039	0,262	0,999	0,040	-0,010	0,001	0,002
8	GO	1,000	0,041	0,043	0,276	0,996	0,044	0,017	0,004	0,006
9	MA	1,000	0,034	0,059	-0,341	0,890	0,054	0,119	0,110	0,246
10	MG	1,000	0,040	0,033	0,245	0,997	0,034	-0,014	0,003	0,004
11	MS	1,000	0,043	0,014	0,155	0,990	0,014	-0,015	0,010	0,005
12	MT	1,000	0,042	0,012	0,144	1,000	0,012	0,003	0,000	0,000
13	PA	1,000	0,036	0,066	-0,366	0,987	0,067	0,042	0,013	0,033
14	PB	1,000	0,033	0,002	-0,033	0,236	0,001	-0,060	0,764	0,061
15	PE	1,000	0,034	0,002	-0,047	0,550	0,001	-0,043	0,450	0,032
16	PI	1,000	0,032	0,005	-0,103	0,898	0,005	0,035	0,102	0,020
17	PR	1,000	0,042	0,042	0,273	1,000	0,043	0,001	0,000	0,000
18	RJ	1,000	0,044	0,015	0,156	0,947	0,015	-0,037	0,053	0,030
19	RN	1,000	0,038	0,001	0,028	0,711	0,000	-0,018	0,289	0,006
20	RO	1,000	0,040	0,013	0,141	0,830	0,011	0,064	0,170	0,084
21	RR	1,000	0,040	0,091	-0,408	0,998	0,093	-0,017	0,002	0,006
22	RS	1,000	0,040	0,060	0,333	0,998	0,062	0,014	0,002	0,004

23	SC	1,000	0,040	0,062	0,339	1,000	0,063	-0,004	0,000	0,000
24	SE	1,000	0,039	0,000	0,025	0,834	0,000	-0,011	0,166	0,003
25	SP	1,000	0,044	0,028	0,217	0,998	0,029	-0,009	0,002	0,002
26	TO	1,000	0,040	0,001	0,006	0,019	0,000	0,042	0,981	0,036

Column Contributions

ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	ag	1,000	0,373	0,283	0,233	0,968	0,281	-0,043	0,032	0,346
2	de	1,000	0,191	0,698	-0,520	0,996	0,714	-0,031	0,004	0,095
3	ba	1,000	0,436	0,019	0,028	0,237	0,005	0,050	0,763	0,559

3.4.3.2 Análise de Correspondência do Desvio Padrão por estado das Variáveis M_TBANAGUA; M_TAGUA; M_TDENS



Simple Correspondence Analysis: DP_TBANAGUA; DP_TDENS; DP_TAGUA

Relative Inertias

	ag	de	ba	Total
AC	0,001	0,001	0,000	0,003
AL	0,008	0,057	0,013	0,078
AM	0,000	0,003	0,001	0,004
AP	0,020	0,011	0,003	0,035
BA	0,000	0,011	0,006	0,017
CE	0,002	0,010	0,002	0,014
ES	0,034	0,029	0,002	0,066
GO	0,009	0,014	0,000	0,022
MA	0,008	0,001	0,005	0,014
MG	0,000	0,000	0,000	0,000
MS	0,002	0,049	0,023	0,073
MT	0,002	0,008	0,017	0,027
PA	0,001	0,004	0,001	0,006
PB	0,000	0,043	0,026	0,069
PE	0,001	0,028	0,013	0,042
PI	0,000	0,044	0,028	0,073
PR	0,007	0,013	0,000	0,021

RJ	0,025	0,026	0,001	0,052
RN	0,003	0,003	0,000	0,005
RO	0,019	0,014	0,062	0,095
RR	0,000	0,007	0,007	0,014
RS	0,016	0,016	0,001	0,032
SC	0,016	0,002	0,009	0,027
SE	0,003	0,001	0,001	0,005
SP	0,048	0,111	0,003	0,162
TO	0,006	0,009	0,028	0,044
Total	0,232	0,515	0,253	1,000

Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0369	0,6999	0,6999	*****
2	0,0158	0,3001	1,0000	*****
Total	0,0527			

Row Contributions

ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	AC	1,000	0,052	0,003	-0,042	0,621	0,002	0,033	0,379	0,004
2	AL	1,000	0,055	0,078	0,272	0,997	0,111	0,014	0,003	0,001
3	AM	1,000	0,049	0,004	0,066	0,944	0,006	0,016	0,056	0,001
4	AP	1,000	0,040	0,035	0,138	0,415	0,021	-0,164	0,585	0,068
5	BA	1,000	0,046	0,017	0,135	0,901	0,022	0,045	0,099	0,006
6	CE	1,000	0,040	0,014	0,135	0,998	0,019	-0,006	0,002	0,000
7	ES	1,000	0,018	0,066	-0,333	0,566	0,053	0,291	0,434	0,095
8	GO	1,000	0,025	0,022	-0,197	0,814	0,026	0,094	0,186	0,014
9	MA	1,000	0,045	0,014	0,038	0,090	0,002	-0,122	0,910	0,042
10	MG	1,000	0,034	0,000	0,015	0,803	0,000	-0,008	0,197	0,000
11	MS	1,000	0,031	0,073	-0,340	0,925	0,097	-0,097	0,075	0,018
12	MT	1,000	0,032	0,027	-0,141	0,441	0,017	-0,159	0,559	0,051
13	PA	1,000	0,058	0,006	-0,071	0,991	0,008	0,007	0,009	0,000
14	PB	1,000	0,054	0,069	0,242	0,858	0,085	0,098	0,142	0,033
15	PE	1,000	0,051	0,042	0,201	0,929	0,056	0,055	0,071	0,010
16	PI	1,000	0,057	0,073	0,240	0,851	0,088	0,101	0,149	0,036
17	PR	1,000	0,022	0,021	-0,205	0,847	0,025	0,087	0,153	0,010
18	RJ	1,000	0,023	0,052	-0,278	0,652	0,048	0,203	0,348	0,060
19	RN	1,000	0,041	0,005	0,064	0,602	0,005	-0,052	0,398	0,007
20	RO	1,000	0,025	0,095	-0,211	0,220	0,030	-0,397	0,780	0,247
21	RR	1,000	0,064	0,014	-0,088	0,661	0,013	-0,063	0,339	0,016
22	RS	1,000	0,024	0,032	-0,209	0,626	0,029	0,161	0,374	0,040
23	SC	1,000	0,024	0,027	-0,066	0,074	0,003	0,234	0,926	0,082
24	SE	1,000	0,035	0,005	0,041	0,214	0,002	-0,078	0,786	0,013
25	SP	1,000	0,021	0,162	-0,613	0,911	0,211	0,191	0,089	0,048
26	TO	1,000	0,038	0,044	-0,140	0,324	0,020	-0,202	0,676	0,098

Column Contributions

ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	ag	1,000	0,392	0,232	0,108	0,374	0,124	-0,140	0,626	0,484
2	de	1,000	0,264	0,515	-0,321	1,000	0,735	0,007	0,000	0,001
3	ba	1,000	0,344	0,253	0,123	0,389	0,140	0,154	0,611	0,515

3.4.4 Dimensão Habitação por Região

3.4.4.1 ANOVA por região das Variáveis M_TBANAGUA; M_TAGUA; M_TDENS

One-way ANOVA: M_TABANAGUA versus REGIAO_N

Source	DF	SS	MS	F	P
REGIAO_N	4	0,8865	0,2216	15,65	0,000
Error	21	0,2975	0,0142		
Total	25	1,1839			

S = 0,1190 R-Sq = 74,88% R-Sq(adj) = 70,09%

Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev
CO	3	0,9068	0,0325
N	7	0,5094	0,1778
NE	9	0,6152	0,1138
S	3	0,9531	0,0019
SE	4	0,9501	0,0259

0,60 0,80 1,00 1,20

Pooled StDev = 0,1190

One-way ANOVA: M_TDENS versus REGIAO_N

Source	DF	SS	MS	F	P
REGIAO_N	4	0,5071	0,1268	11,02	0,000
Error	21	0,2415	0,0115		
Total	25	0,7486			

S = 0,1072 R-Sq = 67,74% R-Sq(adj) = 61,59%

Individual 95% CIs For Mean Based on Pooled StDev

Level	N	Mean	StDev
CO	3	0,2507	0,0639
N	7	0,5583	0,1816
NE	9	0,3773	0,0582
S	3	0,1496	0,0315
SE	4	0,2368	0,0463

0,16 0,32 0,48 0,64

Pooled StDev = 0,1072

One-way ANOVA: M_TAGUA versus REGIAO_N

Source	DF	SS	MS	F	P
REGIAO_N	4	0,14467	0,03617	6,72	0,001
Error	21	0,11304	0,00538		
Total	25	0,25771			

S = 0,07337 R-Sq = 56,14% R-Sq(adj) = 47,78%

Individual 95% CIs For Mean Based on Pooled StDev

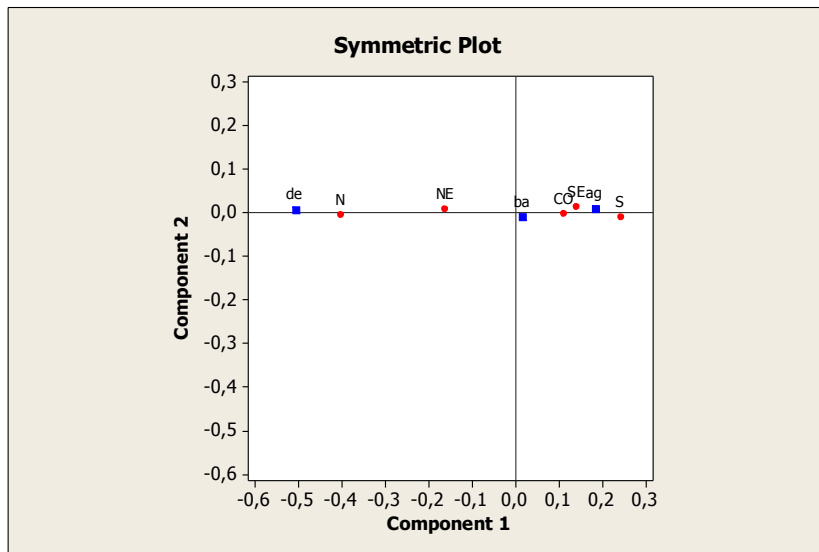
Level	N	Mean	StDev
CO	3	0,9307	0,0052
N	7	0,8065	0,1172

NE	9	0,7517	0,0573	(-----*-----)
S	3	0,9117	0,0212	(-----*-----)
SE	4	0,9239	0,0335	(-----*-----)

-----+-----+-----+-----+-----
 0,720 0,800 0,880 0,960

Pooled StDev = 0,0734

3.4.4.2 Análise de Correspondência da Média por região das Variáveis M_TBANAGUA; M_TAGUA; M_TDENS



Simple Correspondence Analysis: M_TBAN_r; M_TDEN_r; M_TAGU_r

Relative Inertias

	ag	de	ba	Total
CO	0,011	0,039	0,000	0,050
N	0,144	0,416	0,001	0,561
NE	0,018	0,065	0,001	0,084
S	0,050	0,173	0,001	0,224
SE	0,025	0,056	0,000	0,081
Total	0,248	0,749	0,003	1,000

Analysis of Contingency Table

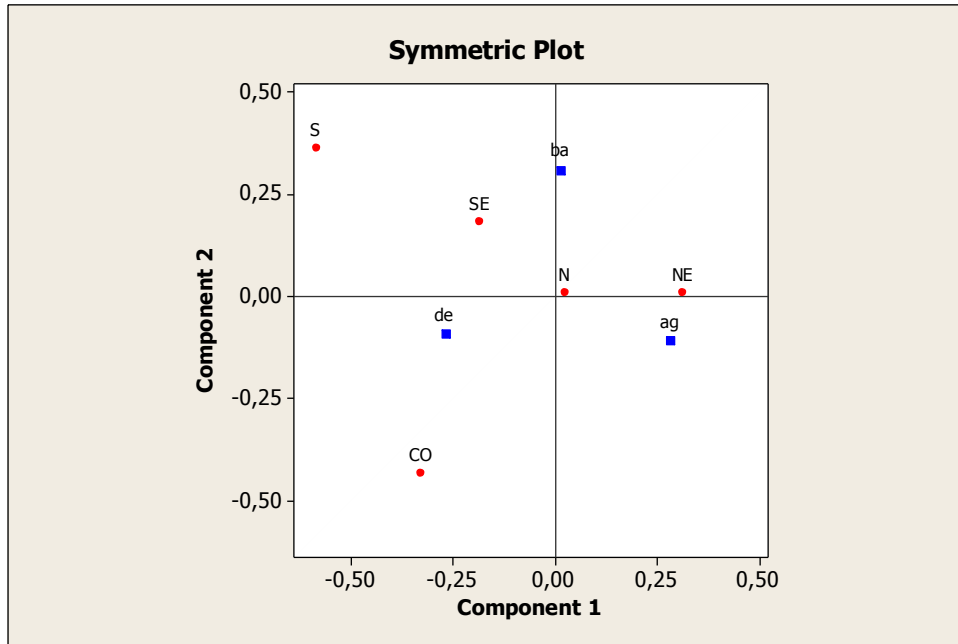
Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0544	0,9985	0,9985	*****
2	0,0001	0,0015	1,0000	
Total	0,0545			

Row Contributions

ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	CO	1,000	0,212	0,050	0,114	0,998	0,050	-0,005	0,002	0,056
2	N	1,000	0,191	0,561	-0,400	1,000	0,562	-0,006	0,000	0,080
3	NE	1,000	0,177	0,084	-0,160	0,998	0,084	0,008	0,002	0,134
4	S	1,000	0,205	0,224	0,244	0,998	0,224	-0,010	0,002	0,270
5	SE	1,000	0,215	0,081	0,143	0,991	0,080	0,013	0,009	0,460

Column Contributions					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	ag	1,000	0,400	0,248	0,184	0,998	0,248	0,009	0,002	0,352
2	de	1,000	0,160	0,749	-0,505	1,000	0,750	0,007	0,000	0,090
3	ba	1,000	0,440	0,003	0,016	0,722	0,002	-0,010	0,278	0,558

3.4.4.3 Análise de Correspondência do Desvio Padrão por região das Variáveis M_TBANAGUA; M_TAGUA; M_TDENS



Simple Correspondence Analysis: DP_TBAN_r; DP_TDEN_r; DP_TAGU_r

Relative Inertias

	ag	de	ba	Total
CO	0,006	0,166	0,179	0,350
N	0,001	0,002	0,000	0,004
NE	0,130	0,134	0,001	0,264
S	0,189	0,054	0,056	0,299
SE	0,048	0,006	0,028	0,082
Total	0,374	0,362	0,264	1,000

Analysis of Contingency Table

Axis	Inertia	Proportion	Cumulative	Histogram
1	0,0577	0,6519	0,6519	*****
2	0,0308	0,3481	1,0000	*****
Total	0,0886			

Row Contributions

Row Contributions					Component 1			Component 2		
ID	Name	Qual	Mass	Inert	Coord	Corr	Contr	Coord	Corr	Contr
1	CO	1,000	0,105	0,350	-0,327	0,361	0,194	-0,435	0,639	0,643
2	N	1,000	0,492	0,004	0,026	0,919	0,006	0,008	0,081	0,001
3	NE	1,000	0,237	0,264	0,314	1,000	0,405	0,007	0,000	0,000
4	S	1,000	0,056	0,299	-0,583	0,722	0,331	0,362	0,278	0,239
5	SE	1,000	0,109	0,082	-0,184	0,508	0,064	0,181	0,492	0,116

Column Contributions										
ID	Name	Qual	Mass	Inert	Component 1			Component 2		
					Coord	Corr	Contr	Coord	Corr	Contr
1	ag	1,000	0,364	0,374	0,282	0,875	0,502	-0,107	0,125	0,134
2	de	1,000	0,394	0,362	-0,270	0,895	0,497	-0,092	0,105	0,109
3	ba	1,000	0,242	0,264	0,015	0,002	0,001	0,310	0,998	0,757

4. CONSIDERAÇÕES FINAIS

Realizar a análise discriminante dos dados busca verificar qual dos métodos entre as funções linear e quadrática, demonstra maior assertividade nos agrupamentos e grupos de similaridade.

Foram realizadas análises nas dimensões Demografia e Habitação, utilizando agrupamentos em 4 Brasis. Para a dimensão Habitação foi necessário agrupar o estado do Amazonas no Grupo 1 pois o software necessita de pelo menos dois registros dentro de cada grupo para gerar a devida análise discriminante.

Na dimensão Demografia as variáveis ESPVIDA1, MORT1(i) e MORT5(i) utilizamos a análise discriminante linear e conseguimos um resultado de 49,40% de proporção correta, valor muito baixo para confiança nos dados em análise. Para análise Discriminante Linear para os “4 Brasis” obtivemos o percentual de 55,40% enquanto no método quadrático 60,50%. Na análise regressão logística apresenta-se um ganho significativo de confiabilidade atingindo o percentual de 70,90% para as variáveis ESPVIDA1, MORT1(i) e MORT5(i) e 73,00% para os “4 Brasis”.

Na análise realizada por agrupamentos de regiões apresentado no dendrograma, e na regressão logística com o cenário 4 “Brasis”, a proporcionalidade ficou em 100%, para os dois modelos.

Na dimensão Habitação as variáveis T_BANAGUA, T_DENS e T_AGUA utilizamos a análise discriminante linear e conseguimos um resultado de 43,60% de proporção correta, valor muito baixo para confiança nos dados em análise. Para análise Discriminante Linear para os “4 Brasis” obtivemos o percentual de 54,20% enquanto no método quadrático 54,70%. Na análise regressão logística apresenta-se um ganho significativo de confiabilidade atingindo o percentual de 71,20% para as variáveis T_BANAGUA, T_DENS e T_AGUA e 72,50% para os “4 Brasis”.

Na análise realizada por agrupamentos apresentado no dendrograma, e na regressão logística com o cenário 3 “Brasis” (junção do AM), a proporcionalidade ficou em 100%, para os dois modelos.

Para a análise realizada pela metodologia *Classification Tree*, por meio do software SPSS, a mesma apresentou uma queda na confiabilidade dos índices também não apresentando índices satisfatórios. Para a Dimensão Demografia apresentou o percentual de 57,00% como *Percent Correct* enquanto na Dimensão Habitação 50,9%.