

ANÁLISE DE HOMOGENEIDADE DOS PAÍSES DA AMÉRICA LATINA QUANTO A INDICADORES SOCIAIS

Giovani Glaucio de Oliveira Costa
Universidade Federal Rural do Rio de Janeiro
Instituto Multidisciplinar
Rua Governador Roberto Silveira.S/N
giovaniglaucio@ufrj.br

Resumo: Este artigo apresenta e explora as potencialidades de utilização de um método original e inovador de Análise de Dados- Análise de Homogeneidade (procedimento HOMALS no SPSS) também conhecido por Análise de Correspondência Múltipla, particularmente apropriada à abordagem simultânea de múltiplos indicadores e ao tratamento de variáveis qualitativas ou tratadas como tal. A Análise de Homogeneidade é particularmente interessante para reproduzir num espaço de menores dimensões a multidimensionalidade que sustenta o espaço original. Não sendo este o único objetivo que leva à seleção de uma HOMALS é, sem dúvida, um dos mais habituais e cujo enfoque foi privilegiado neste estudo. A partir das configurações definidas pelas categorias das variáveis qualitativas ou tratadas como tal espera-se conhecer de forma aproximada a estrutura que sustenta o espaço em análise. As configurações nos planos ao exibirem diferentes combinações das propriedades envolvidas, refletem a presença de grupos de indivíduos relativamente homogêneos. Seja, então, qual for à natureza do campo de estudo e das unidades de análise, o que determina a opção pela Análise de Homogeneidade é o propósito de realizar uma abordagem relacional sobre as múltiplas variáveis que caracterizam os objetos, podendo vir a serem definidos diferentes grupos. Este artigo tem o objetivo de apresentar os fundamentos, os conceitos, o método, as fases de análise e a utilização da Análise de Homogeneidade com SPSS, exemplificando a aplicação e interpretação dos resultados através de um estudo de caso: Análise de Homogeneidade dos Países da América do Sul quanto a Indicadores Sociais.

Palavras-chaves: análise de dados qualitativos, análise de homogeneidade, homogeneidade de países da América do Sul, indicadores sociais.

Abstract: this article introduces and explores the potential of using an original method and innovative data analysis-analysis of homogeneity (HOMALS procedure in SPSS) also known by multiple correspondence analysis, particularly suitable for simultaneous multiple indicators and approach to treatment of qualitative variables or treated as such. The analysis of homogeneity is particularly interesting to play in a smaller space the multidimensionality that sustains the original space. Is not this the only goal that leads to the selection of a HOMALS is undoubtedly one of the most usual and whose focus was privileged in this study. From the settings defined by categories of qualitative variables or treated as such is expected to know roughly the structure that sustains the space under consideration. The settings in the plans to display different combinations of the properties involved, reflect the presence of relatively homogeneous groups of

individuals. Is, then, what is the nature of the field of study and units of analysis, which determines the option for Analysis of homogeneity is the purpose to perform a relational approach about the multiple variables that characterize the objects, and may come to be defined by different groups. This article aims to present the fundamentals, concepts, the method, the stages of analysis and use of homogeneity Analysis with SPSS, exemplifying the application and interpretation of the results through a case study: analysis of the homogeneity of the countries of South America and the Social Indicators.

Key Words: qualitative data analysis, homogeneity analysis, homogeneity of countries of South America, social indicators.

1-Introdução

Na elaboração de um plano de pesquisa que planeja uma coleta extensa de informações, a seleção dos métodos de análise de dados é uma tarefa (entre outras) de importância decisiva, quaisquer que seja o campo em estudo. É essencial garantir que os instrumentos que venham a ser selecionados permitam operacionalizar as orientações sistematizadas no modelo de análise, e sejam consistentes com a natureza da informação escolhida(Michel Greenacre,1996).

Muitas das vezes quando o pesquisador elabora esse plano de pesquisa quantitativa, utilizando o método estatístico, ele se depara com questionários onde se não todas, a maioria das variáveis são qualitativas, isto é, as questões que envolvem o questionário têm como respostas, muitas das vezes estimuladas, a forma categorial. As técnicas tradicionais de Análise Multivariada pressupõem que as variáveis sejam quantitativas e que sigam à Curva Normal. O analista, então, tem utilizado como alternativa o recurso analítico trivial do exame de tabelas de frequência, tabelas de contingência com testes de independência do qui-quadrado. A Análise de Homogeneidade ou Análise de Correspondência Múltipla, técnica recente e inovadora, pouco difundida ainda principalmente no Brasil, permite que os pesquisadores ao utilizar questionários que envolvam perguntas com respostas nominais ou ordinais, saiam do campo meramente descritivo e desenvolvam técnicas multivariadas para o tratamento avançado dos múltiplos indicadores envolvidos na pesquisa. A Análise de Homogeneidade trata múltiplas variáveis categoriais e realiza análises estruturais(John Geer,1993).

A Análise de Homogeneidade é um método de análise quantitativa para variáveis qualitativas e tem a finalidade de atingir a solução ótima na quantificação de dados qualitativos(Helena Carvalho,2004). A Análise de Homogeneidade submete os dados qualitativos de *input* a um processo de quantificação(Albert Gifi,1996). As quantificações são estimadas para as categorias e objetos(Helena Carvalho,2004). Cada categoria tem a sua quantificação e cada objeto também. As quantificações ótimas representam o comportamento geral de cada categoria e de cada objeto no espaço original(Helena Carvalho,2004).

2- A Análise de Dados via HOMALS

A análise de dados através da HOMALS tem duas fases(Helena Carvalho,2004):

- **1ª Fase- Identificação dos Eixos Estruturantes do Espaço de Análise;**
- **2ª Fase- Identificação de Configurações ou de Grupos Homogêneos no Espaço de Análise.**

1ª Fase- Identificação dos Eixos Estruturantes do Espaço de Análise:

A- Seleção das Dimensões mais Representativas

- 1- Identificação das dimensões com valores próprios mais elevados, com a ajuda de gráfico, feito em Excel, destas estatísticas;
- 2- Identificação das variáveis com medidas de discriminação mais elevadas, privilegiando as dimensões da análise feita em 1, podendo ser feito também com o auxílio do gráfico de variáveis, que também é chamado de gráficos das medidas de discriminação.

B- Interpretação/Nomeação das Dimensões

- 3- Leitura do significado das dimensões por via das variáveis selecionadas em 2. O objetivo desta fase é nomear as dimensões;
- 4- Leitura do significado das dimensões por via também das categorias das variáveis selecionadas usando as suas quantificações. O objetivo desta fase é analisar as associações e as oposições entre as categorias selecionadas.

As associações são analisadas a partir da presença de categorias que detenham quantificações com mesmo sinal e as oposições pela presença de categorias que possuem sinais contrários.

2ª Fase- Identificação de Configurações ou de Grupos Homogêneos no Espaço de Análise

C-Interpretação dos Planos das Categorias

- 5-Leitura de diferentes configurações de categorias por via das projeções de todas as categorias identificadas em 4. O Objetivo desta fase é identificar grupos homogêneos.

D-Interpretação dos Planos de Objetos

- 6-Observação da intensidade da presença dos grupos já identificados;
- 7-Identificação dos objetos quando tal for possível;
- 8-Observações de *outliers*.

3-Estudo de Caso: Análise de Homogeneidade dos Países da América do Sul Quanto à Indicadores Sociais

Com base no IBGE, foram coletados os dados abaixo que se referem a indicadores sociais dos países da América do Sul. O objetivo da aplicação da Análise de Homogeneidade seria a busca da estruturação destes países no espaço de desenvolvimento social. Os países da América do Sul podem possivelmente ser agrupados em termos de suas similaridades no padrão de indicadores sociais.

Quais blocos de países da América do Sul apresentam perfis semelhantes de indicadores sociais?

A verificação de blocos de países da América do Sul com indicadores sociais homogêneos é o objetivo deste trabalho.

A Tabela 1 apresenta os dados coletados. Estes dados já disponíveis (dados secundários) foram levantados pelo Instituto Brasileiro de Geografia e Estatística e sua consulta pode ser realizada pelo site www.ibge.gov.br/paisesat.

Tabela 1- Indicadores Sociais dos Países da América do Sul

Países	IDH	EV	POPS	Kcal	DAAP	DARS	TAM15	TBM
Argentina	0,866	75,0	menor que 5 %	3004	96,0	91,0	97,0	89,7
Bolívia	0,729	65,1	22	2175	86,0	43,0	87,0	86,0
Brasil	0,813	72,0	6	3094	91,0	77,0	88,0	87,5
Chile	0,878	78,4	menor que 5 %	2976	95,0	94,0	96,0	82,9
Colômbia	0,807	72,5	10	2666	93,0	78,0	94,0	75,1
Equador	0,806	74,8	15	2299	95,0	84,0	91,0	SEM DADO
Guiana	0,729	65,8	6	2827	93,0	81,0	97,0	85,0
Paraguai	0,761	71,5	11	2593	77,0	70,0	92,0	69,1
Peru	0,806	71,0	15	2450	84,0	72,0	88,0	85,8
Suriname	0,769	69,8	7	2712	92,0	82,0	88,0	77,1
Uruguai	0,865	76,1	menor que 5 %	2923	100,0	100,0	98,0	88,9
Venezuela	0,844	73,4	12	2453	SEM DADO	SEM DADO	93,0	75,5
Unidade	-	Anos	%	Kcal/dia	%	%	%	%

Fonte: IBGE: <http://www.ibge.gov.br/paisesat/>

No Quadro 1 são descritos os indicadores estruturantes das possíveis configurações dos planos a serem analisados

Quadro 1- Descrição das Variáveis da Base de Dados

Variáveis	Descrição
IDH	Índice de desenvolvimento humano - 2007
EV	Esperança de vida ao nascer - 2006
POPS	População subnutrida - 2005
Kcal	Calorias consumidas - 2005
DAAP	Domicílios com acesso a água potável - 2006
DARS	Domicílios com acesso a rede sanitária - 2006
TAM15	Taxa de alfabetização das pessoas de 15 anos ou mais de idade - 2005
TBM	Taxa bruta de matrículas para todos os níveis de ensino - 2005

Como os as variáveis são quantitativas, elas precisam ser tratadas como variáveis qualitativas e em seguida, codificadas, para ser realizada a Análise de Homogeneidade. A Tabela 2 indica como ponto de corte para categorização das variáveis a mediana.

Tabela 2- Mediana das Variáveis Consideradas

Variáveis	Mediana
IDH	0,807
EV	72,3
POPS	11,0
Kcal	2689
DAAP	93,0
DARS	81,0
TAM15	92,5
TBM	85,0

A Tabela 3 informa os indicadores abordados na análise tratados como variáveis nominais, isto é, categorizadas.

Tabela 3- Matriz dos Dados Categorizados

Países	IDH	EV	POPS	Kcal	DAAP	DARS	TAM15	TBM
Argentina	>0,807IDH	>72,3Anos	<5%POPS	>2689Kcal	>93,0%DAAP	>81%DARS	>92,5%TAM15	>85,0%TBM
Bolívia	<=0,807IDH	<=72,3Anos	>11%POPS	<=2689Kcal	<=93,0%DAAP	<=81%DARS	<=92,5%TAM15	>85,0%TBM
Brasil	>0,807IDH	<=72,3Anos	<=11%POPS	>2689Kcal	<=93,0%DAAP	<=81%DARS	<=92,5%TAM15	>85,0%TBM
Chile	>0,807IDH	>72,3Anos	<5%POPS	>2689Kcal	>93,0%DAAP	>81%DARS	>92,5%TAM15	<=85,0%TBM
Colômbia	<=0,807IDH	>72,3Anos	<=11%POPS	<=2689Kcal	<=93,0%DAAP	<=81%DARS	>92,5%TAM15	<=85,0%TBM
Equador	<=0,807IDH	>72,3Anos	>11%POPS	<=2689Kcal	>93,0%DAAP	>81%DARS	<=92,5%TAM15	S DTBM
Guiana	<=0,807IDH	<=72,3Anos	<=11%POPS	>2689Kcal	<=93,0%DAAP	<=81%DARS	>92,5%TAM15	<=85,0%TBM
Paraguai	<=0,807IDH	<=72,3Anos	<=11%POPS	<=2689Kcal	<=93,0%DAAP	<=81%DARS	<=92,5%TAM15	<=85,0%TBM
Peru	<=0,807IDH	<=72,3Anos	>11%POPS	<=2689Kcal	<=93,0%DAAP	<=81%DARS	<=92,5%TAM15	>85,0%TBM
Suriname	<=0,807IDH	<=72,3Anos	<=11%POPS	>2689Kcal	<=93,0%DAAP	>81%DARS	<=92,5%TAM15	<=85,0%TBM
Uruguai	>0,807IDH	>72,3Anos	<5%POPS	>2689Kcal	>93,0%DAAP	>81%DARS	>92,5%TAM15	>85,0%TBM
Venezuela	>0,807IDH	>72,3Anos	>11%POPS	<=2689Kcal	SDDAAP	S DDARS	>92,5%TAM15	<=85,0%TBM

Primeiramente foi realizada a HOMLAS com todas as variáveis disponíveis para a análise. Através do exame da distribuição de frequência, das medidas de discriminação e do gráfico de variáveis foi possível detectar que a variável TBM, taxa bruta de matrícula para todos os níveis de ensino, tem pouco poder de discriminação pois é uma variável que tem um nível de “consenso” junto aos países da base de dados (comportamento típico). Esta variável se direciona ao perfil médio dos países da base de dados e são pouco diferenciados dos mesmos, não colaborando para a formação de diferentes configurações ou blocos de nações com perfis de indicadores sociais homogêneos.

As categorias “SEM DADO(SDDAAP e SDDARS)” e “<5%POPS” constituem não respostas e de frequência residuais respectivamente e, optou, então, por não incluí-las na análise de dados, atribuindo-lhes códigos -1 para melhorar o Fit.

Ao realizar, então, em segundo momento, a análise sem a variável TBM e sem as categorias “SEM DADO(SDDAAP e SDDARS)” e “<5%POPS” o *Fit* aumentou e se concentrou em 0,827735.

Por fim, a Tabela 4 indica a matriz de *input* para realização da HOMALS no SPSS.

Tabela 4- Matriz de Input no SPSS

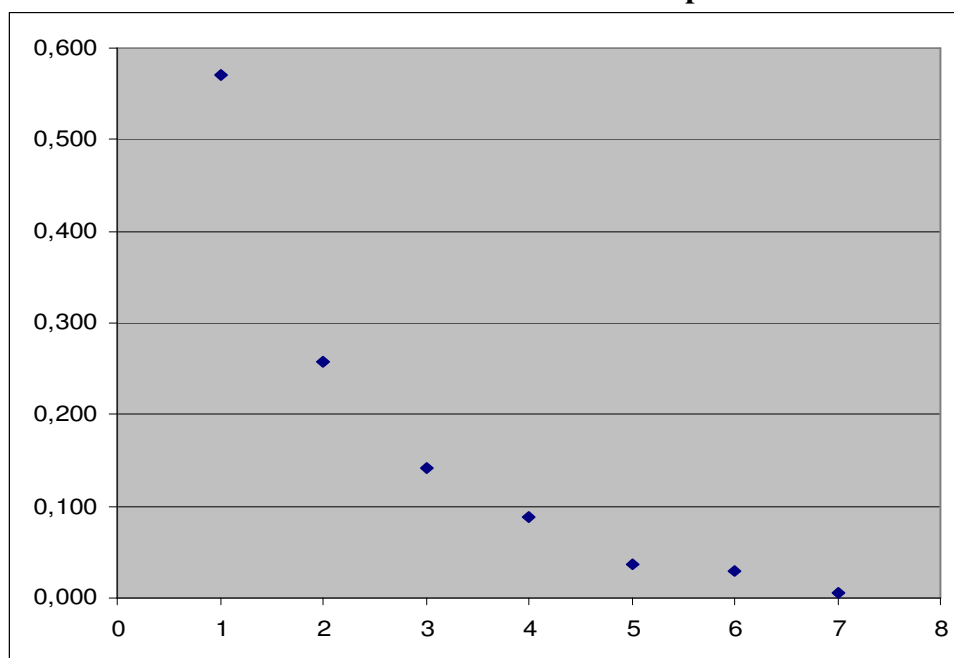
Código	Países	IDH	EV	POPS	Kcal	DAAP	DARS	TAM15	TBM
1	Argentina	2	2	3	2	2	2	2	2
2	Bolívia	1	1	2	1	1	1	1	2
3	Brasil	2	1	1	2	1	1	1	2
4	Chile	2	2	3	2	2	2	2	1
5	Colômbia	1	2	1	1	1	1	2	1
6	Equador	1	2	2	1	2	2	1	-1
7	Guiana	1	1	1	2	1	1	2	1
8	Paraguai	1	1	1	1	1	1	1	1
9	Peru	1	1	2	1	1	1	1	2
10	Suriname	1	1	1	2	1	2	1	1
11	Uruguai	2	2	3	2	2	2	2	2
12	Venezuela	2	2	2	1	-1	-1	2	1

A Tabela 5 indica os valores próprios das dimensões para a análise da HOMALS. Foi solicitado ao SPSS o número máximo de soluções(14-7=7). O Gráfico 23 apresenta a representação dos valores próprios via imagem.

Tabela 5- Valores Próprios das Dimensões

Dimensões	Valores Próprios
1	0,570
2	0,258
3	0,142
4	0,088
5	0,037
6	0,030
7	0,006

Gráfico 1- Gráfico dos Valores Próprios



Pela análise da Tabela 5 e do Gráfico 1, conclui-se pela retenção para o problema das duas primeiras dimensões, dada a sua nítida superioridade sobre as demais.

As duas primeiras dimensões acumulam um FIT de 0,827735, isto é, uma variância explicada percentual de 73,16%. As duas primeiras dimensões retêm mais de 70% das informações originais da base de dados de partida.

Com esta seção finalizada, foi feita a análise das soluções ou dimensões mais importantes para o problema. As dimensões, então, associadas aos dois primeiros maiores valores próprios serão os dois eixos estruturantes para a análise das configurações do espaço indicadores sociais dos países da América do Sul.

Com esta seção finalizada, foi feita a análise das soluções ou dimensões mais importantes para o problema.

Vai-se partir para análise das variáveis mais importantes em cada uma das dimensões retidas na análise.

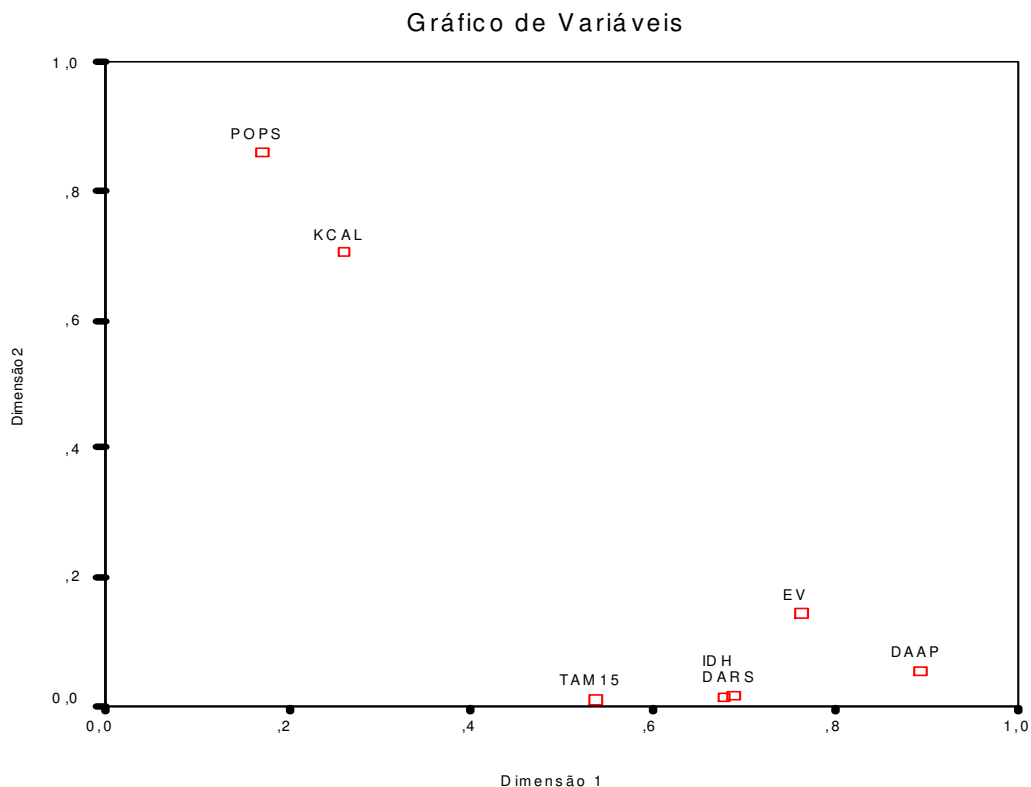
Tabela 6- Medidas de Discriminações

Variáveis	Medidas de Discriminação	
	Dimensão 1	Dimensão 2
IDH	0,677	0,013
EV	0,761	0,145
POPS	0,171	0,860
Kcal	0,263	0,704
DAAP	0,893	0,053
DARS	0,689	0,018
TAM15	0,536	0,011

Na Tabela 6, foram assinaladas em negrito as medidas de discriminação mais importantes em cada dimensão. As variáveis mais importantes numa dada dimensão são aquelas com medidas de discriminação maior ou igual ao valor próprio da dimensão.

O Gráfico 2 confirma, então, as informações da Tabela 6: as variáveis mais importantes para a dimensão 1 estão próximas desta no gráfico e as mais importantes para dimensão 2 estão próximas da mesma no gráfico.

Gráfico 2- Gráfico de Variáveis



Com esta seção concluída, realizou-se a análise das variáveis mais importantes para cada dimensão e por conseguinte para a HOMALS em questão.

O Quadro 1 informa as variáveis mais importantes para a dimensão 1. Em função de suas descrições, temos que a dimensão 1 discrimina os países em função de sua qualidade de vida, isto é, do índice de desenvolvimento humano, da esperança de vida ao nascer, de domicílios com acesso a água potável, de domicílios com acesso a rede sanitária e da taxa de alfabetização das pessoas de 15 anos ou mais de idade.

Quadro 1: Variáveis Importantes para a Dimensão 1

Variáveis	Descrição
IDH	Índice de desenvolvimento humano - 2007
EV	Esperança de vida ao nascer - 2006
DAAP	Domicílios com acesso a água potável - 2006
DARS	Domicílios com acesso a rede sanitária - 2006
TAM15	Taxa de alfabetização das pessoas de 15 anos ou mais de idade - 2005
NOMEAÇÃO	<i>Qualidade de Vida</i>

Já o Quadro 2, informa as variáveis mais importantes para a dimensão 2. Em função de suas descrições, temos que a dimensão 2 discrimina os países em função de sua qualidade nutricional, isto é, do percentual da população subnutrida e do total de calorias consumidas.

Quadro 2 - Variáveis Importantes para a Dimensão 2

Variáveis	Descrição
POPS	População subnutrida - 2005
Kcal	Calorias consumidas - 2005
NOMEAÇÃO	<i>Qualidade Nutricional</i>

Portanto, os dois eixos estruturantes do espaço de análise são “Qualidade de Vida” e “Qualidade Nutricional”. São estes dois eixos temáticos que sustentam o espaço de indicadores sociais dos países da América do Sul. “Qualidade de Vida” e “Qualidade Nutricional” são os novos indicadores sociais dos países da América do Sul. Isto quer dizer que cerca de 70% do desenvolvimento social da América do Sul é refletida pela qualidade de vida e nutricional dos habitantes de seus países. Com estas duas variáveis no plano, é possível se estruturar o espaço de análise.

O objetivo desta fase é analisar as associações e as oposições entre as categorias selecionadas, bem como analisar as categorias mais importantes a HOMALS (Helena Carvalho, 2004). As categorias mais importantes são aquelas com quantificações afastadas de zero, em módulo (Helena Carvalho, 2004).

As associações são analisadas a partir da presença de categorias que detenham quantificações com mesmo sinal e as oposições pela presença de categorias que possuem sinais contrários (Helena Carvalho, 2004).

Este passo da análise permite que o analista verifique como as dimensões “Qualidade de Vida” e “Qualidade Nutricional” discriminam os países em foco, ou seja, como aquelas separam as nações no plano em blocos ou perfis distintos. Os países com patamares de qualidade de vida e nutricional semelhantes tenderão a

ficar juntos no plano formando blocos de países com indicadores sociais homogêneos.

Para descrever a dimensão 1 pode-se proceder à análise da quantificação das categorias (Helena Carvalho, 2004). Nesta análise, primeiro, sinalizam-se todas as variáveis importantes para esta dimensão (Helena Carvalho, 2004). Somente vão-se considerar as quantificações das categorias destes indicadores. Segundo, consideram-se na coluna da dimensão 1 em cada variável assinalada anteriormente as categorias afastadas de zero, separando-as em dois blocos: as de **sinal negativo**, em negrito, vão para o da dimensão $1 < 0$, e as de **sinal positivo**, para a dimensão $1 > 0$. Os resultados desta operação então registrados no Quadro 3.

Quadro 3 - Descrição da Dimensão 1

Dimensão $1 < 0$
>0,807IDH >72,3Anos >93,0%DAAP >81,0%DARS >92,5%TAM15
Dimensão $1 > 0$
<=0,807IDH <=72,3Anos <=93,0%DAAP <=81,0%DARS <=92,5%TAM15

Pela análise da dimensão 1, pode-se constatar que os países com indicadores sociais em patamar superior tendem a estar próximos no plano (associados) e os países com indicadores em patamares inferiores também seguem o mesmo comportamento. Os países com indicadores sociais em patamares contrários opõem-se no plano, por isso que estão em níveis de dimensões diferentes (Dimensão < 0 e Dimensão > 0). Estas características se refletirão em como as categorias e os objetos se posicionarão nos quadrantes dos planos em análise contribuirão para a formação de configurações.

Para gerar as informações do Quadro 4, segue-se procedimento análogo ao realizado para a dimensão 1.

Quadro 4- Descrição da Dimensão 2

Dimensão $2 < 0$
<=11%POPS >2689Kcal
Dimensão $2 > 0$
>11%POPS <=2689Kcal

Pelo Quadro 4, pode observar que estão na Dimensão $2 < 0$, portanto associados, os países com indicadores positivos de qualidade de nutrição e na Dimensão $2 > 0$, os países com aqueles indicadores negativos. Os países com indicadores positivos se opõem aos com indicadores sociais negativos.

Com esta tarefa concluída, efetuou-se, além de uma análise interna das dimensões mais importantes para o problema, a verificação das categorias mais relevantes para a Análise da Homogeneidade em questão. Estas categorias são as categorias das variáveis mais importantes em cada dimensão que possuem valores afastados de zero, em módulo e são as categorias estruturantes das configurações do plano multidimensional.

Até agora se investigou as dimensões mais importantes, as variáveis mais importantes e as categorias mais importantes para a topologia do espaço de análise. Realizou-se também para um conhecimento maior de como as dimensões retidas atuam na diferenciação dos países uma análise interna das mesmas através da quantificação das categorias. Os próximos passos da análise consistirão da estruturação propriamente dita do espaço social dos países da América do Sul, no sentido que se procurará averiguar grupos de países da América do Sul em termos de suas similaridades no padrão de indicadores sociais. A idéia é investigar a existência de grupos distintos de países em função de sua economia. Observa-se no Gráfico 3 quatro perfis ou configurações bem distintas e claras.

O *Perfil 1 ou Configuração 1* está concentrado nos 2º e 3º quadrantes e é formado pelos países com índice de desenvolvimento humano maior de 0,807, com esperança de vida maior de 72,3 anos, com percentual de domicílios com acesso à água potável maior de 93,0% e com percentual de domicílios com acesso à rede sanitária maior 81,0%. Estes países são os que apresentam altos indicadores sociais, com excelentes qualidades de vida e nutrição para os padrões da América do Sul.

O *Perfil 2 ou Configuração 2* é está concentrada nos 1º e 2º quadrantes. É formado pelos países com esperança de vida maior de 72,3 anos, com percentual de população subnutrida maior de 11% e com total de calorias consumidas menor ou igual a 2689 Kcal/dia. São países que precisam melhorar os seus indicadores sociais.

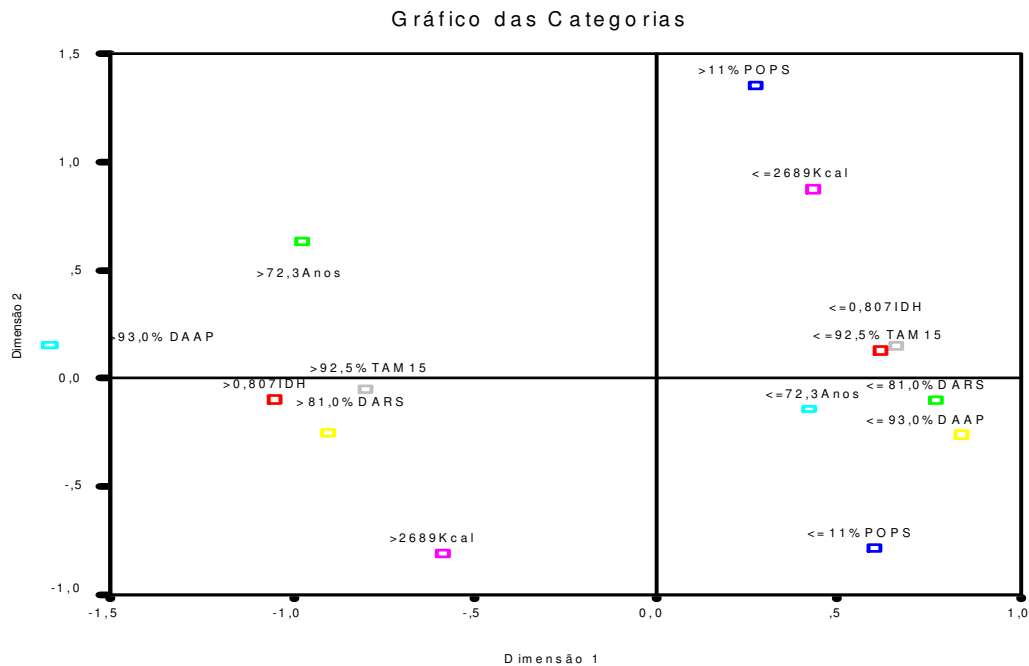
O *Perfil 3 ou Configuração 3* está concentrada nos 1º, 3º e 4º quadrantes. É formado pelos países com índice de desenvolvimento humano menor ou igual 0,807, com esperança de vida menor ou igual a 73,2 anos, com percentual de população subnutrida menor ou igual a 11%, com percentual de domicílios com acesso à água potável menor de 93,0%, com percentual de domicílios com acesso à rede sanitária menor de 81,0%, com taxa de alfabetização das pessoas adultas menor de 93,5% e com total de calorias consumidas maior do que 2689 Kcal/dia. São países que têm indicadores positivos, mas tem que melhorar muito os seus padrões de indicadores sociais.

O *Perfil 4 ou Configuração 4* está concentrado 1º e 4º quadrantes. O Perfil 4 ou Configuração 4 está concentrado 1º e 4º quadrantes. É formado pelos países com índice de desenvolvimento humano menor ou igual 0,807, com esperança de vida menor ou igual a 73,2 anos, com percentual de domicílios com acesso à água potável menor de 93,0%, com percentual de domicílios com acesso à rede sanitária menor de 81,0%, com taxa de alfabetização das pessoas adultas menor de 93,5% e com total de calorias consumidas menor do que 2689 Kcal/dia. Estes

países têm perfil muito semelhante aos países do perfil 3, diferindo em indicadores de qualidade nutricional.

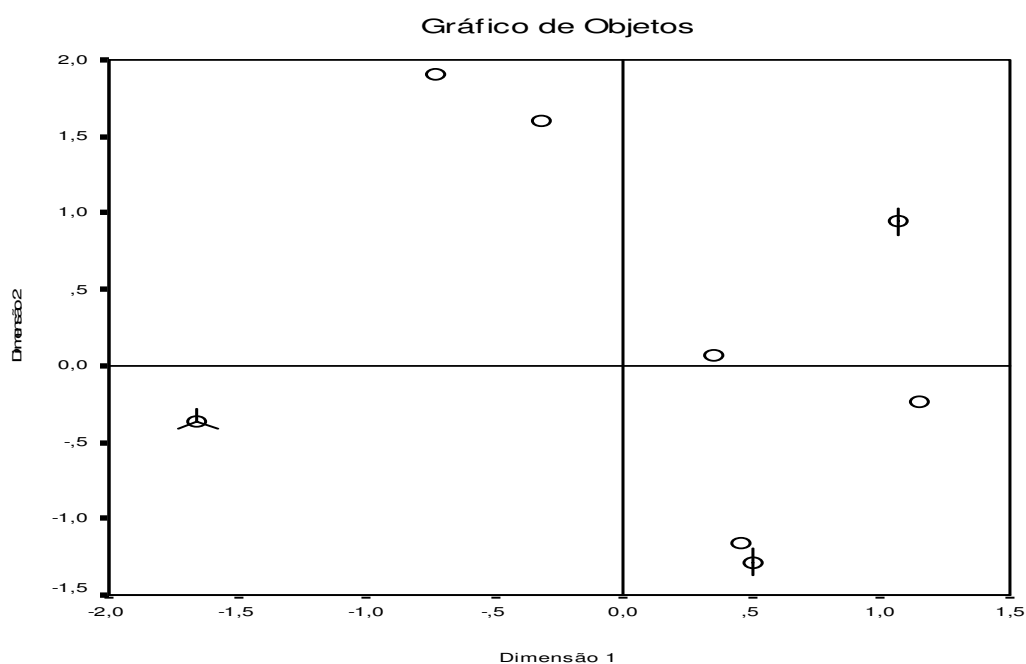
Os perfis estão definidos. A estruturação do espaço de análise foi, então, realizada. Restar indicar que países da América do Sul se encaixam em cada perfil detectado. Para isto, os gráficos dos objetos desempenham um papel fundamental.

Gráfico 3 – Gráfico das Categorias



Observa-se pelo Gráfico 4 a presença nítida de quatro grupos de países no gráfico de objetos. Tem um grupo com três objetos no 3º quadrante, um grupo com dois objetos no 2º quadrante, um grupo com quatro objetos nos 1º e 4º quadrantes e no final do 4º quadrante tem um grupo de objetos com três objetos.

Gráfico 4- Gráfico dos Objetos



O gráfico de objeto simplesmente indica grupos de objetos, no caso países e o número de componentes em cada grupo ou configuração. Entretanto, não informa que países são os formadores de cada grupo ou configuração. Com o Gráfico 5 fica claro, então, quem são os formadores de cada grupo evidenciado no Gráfico 4.

O primeiro perfil é formado pelo Uruguai, Argentina e Chile. O segundo perfil pela Equador e Venezuela. O terceiro perfil é formado pelo Suriname, Guiana e Brasil e o quarto perfil é constituído pela Bolívia, Peru, Colômbia e Paraguai.

Cada grupo de países forma um bloco de nações com similaridades de padrão de indicadores sociais. Com isso, foi concluída a Análise de Homogeneidade dos países da América do Sul quanto aos padrões de indicadores sociais.

A Análise de Homogeneidade informou as características marcantes de cada perfil ou configuração coexistindo no mesmo espaço através do gráfico das categorias. O gráfico de objetos identificou quatro grupo de países similares em indicadores sociais, mas não pode fazer a associação com os perfis do gráfico de categorias. Qual perfil que cabe a cada grupo de países? Na articulação entre a HOMALS e a Análise de Clusters esta pergunta pode ser respondida. O procedimento consiste em associar os perfil do plano das categorias com os clusters de países da Análise de Clusters. Espera-se, entretanto, que os grupos com países similares detectados nos gráficos de objetos possam ser os mesmos gerados pela Análise de clusters.

Gráfico 5- Gráfico de Objetos por Países

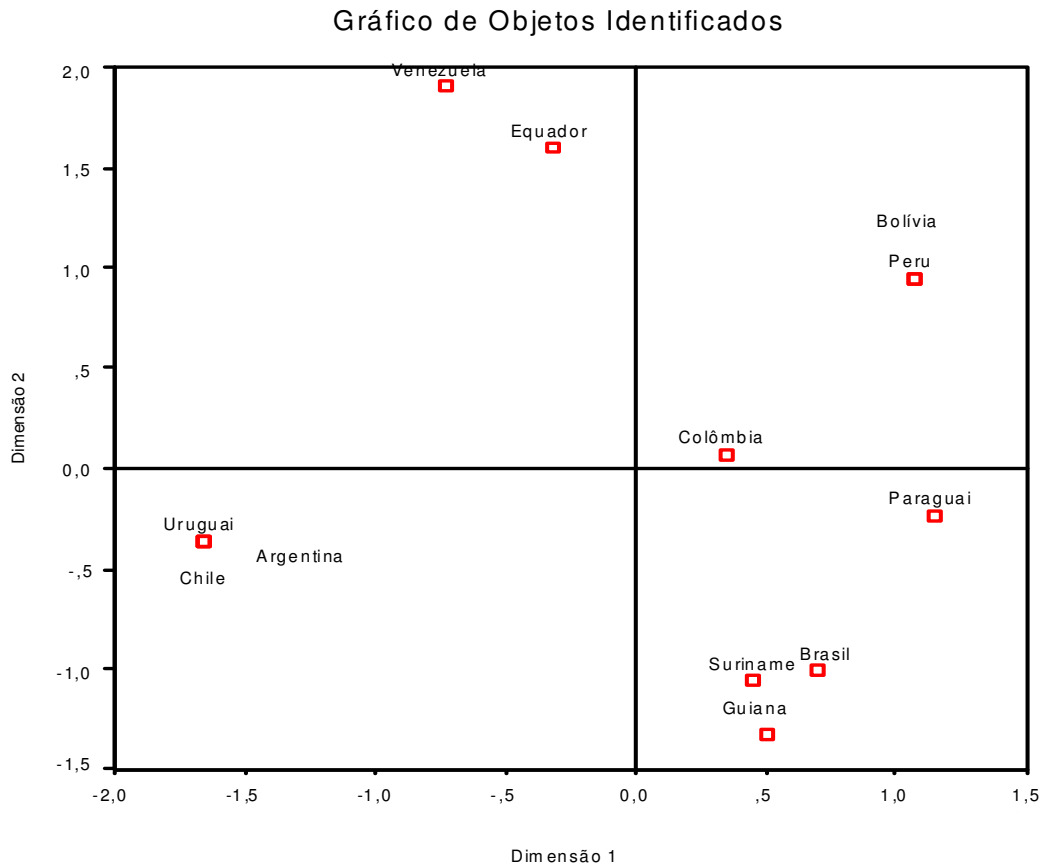
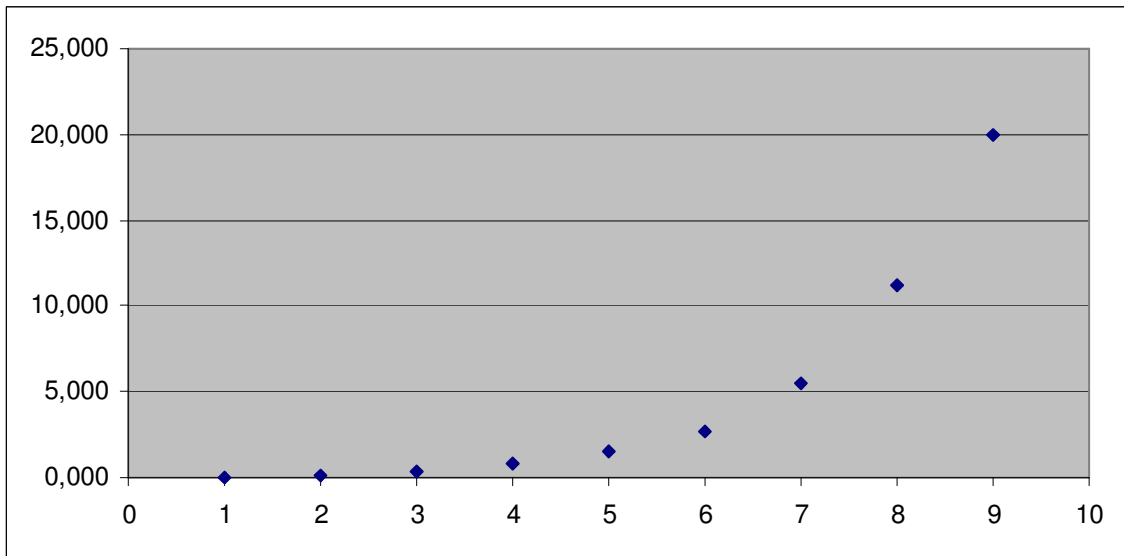


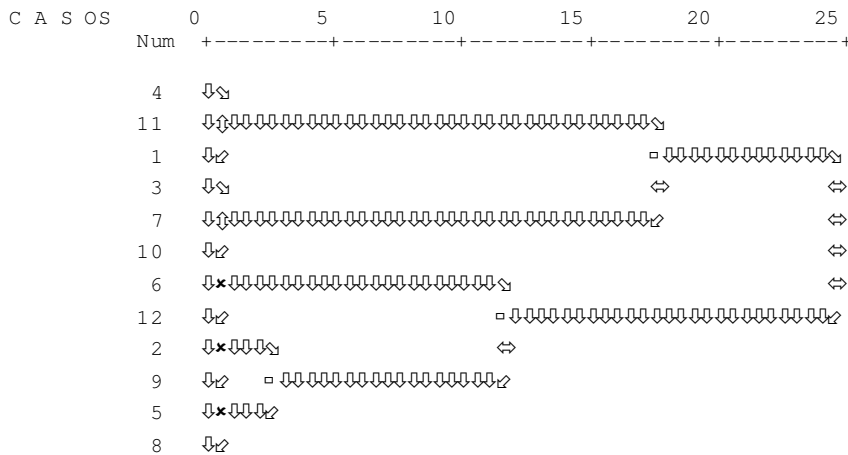
Gráfico 6 – Gráfico dos Coeficientes de Fusão



Concluída esta demonstração retoma-se a solução otimizada com a realização do método de *K-Means*. Classificados os casos, é então necessário fazer a correspondência entre a solução de três clusters e os três perfis configurados pelo plano da HOMALS. Esta operação pode ser realizada através do cruzamento de cada uma das variáveis usadas na definição dos perfis, via HOMALS com os clusters.

Para identificar os quartos clusters resultantes, apresenta-se o Gráfico 7, o Dendograma. O Dendograma é lido da esquerda para a direita. A posição na escala horizontal indica a distância a que os clusters são agrupados. Nos primeiros estágios, muitas linhas têm magnitudes idênticas e por isso é difícil saber a seqüência pela qual os primeiros clusters foram criados. Pelo contrário, nos últimos estágios, a distância a que os clusters são combinados é, tipicamente, grande. É fácil identificar em qual clusters cada objeto pertence.

Gráfico 7 -Dendograma



Assim, o *Clusters 1* é formado pelos países Uruguai, Chile e Argentina; o *Clusters 2* é formado pelos países Venezuela e Equador; O *Clusters 3* pelos países Suriname, Guiana e Brasil e o *Clusters 4* pelos países Paraguai, Colômbia, Peru e Bolívia (ver a codificação dos países na Tabela 1). Observe que a formação dos clusters Hierárquico coincide com os perfis da HOMALS. Falta agora associar as características dos perfis da HOMALS com os clusters da Análise de Clusters. Para isso se irá cruzar as variáveis da HOMAS com os clusters da Análise de Clusters através do SPSS.

O resultado encontra-se no Tabela 7. A partir das percentagens sombreadas nesta tabela, as mais elevadas em termos relativos para cada um dos dois grupos, podem fazer-se uma síntese das principais características dos clusters. As percentagens que não foram sombreadas não discriminam os grupos gerados.

Como se pretende caracterizar cada um deles, no sentido de identificar os perfis configurados pela Análise de Homogeneidade, procedeu-se a uma análise comparativa.

Tem-se então:

Clusters 1 com 25%- identifica o bloco de países com índice de desenvolvimento humano maior de 0,807, com esperança de vida maior de 72,3 anos, com percentual de domicílios com acesso à água potável maior de 93,0% e com percentual de domicílios com acesso à rede sanitária maior 81,0%.

Clusters 2 com 17% - identifica o bloco de países com esperança de vida maior de 72,3 anos, com percentual de população subnutrida maior de 11% e com total de calorias consumidas menor ou igual a 2689 Kcal/dia.

Clusters 3 com 25% - identifica o bloco de países com índice de desenvolvimento humano menor ou igual 0,807, com esperança de vida menor ou

igual a 73,2 anos, com percentual de população subnutrida menor ou igual a 11%, com percentual de domicílios com acesso à água potável menor de 93,0%, com percentual de domicílios com acesso à rede sanitária menor de 81,0%, com taxa de alfabetização das pessoas adultas menor de 93,5% e com total de calorias consumidas maior do que 2689 Kcal/dia.

Clusters 4 com 33% - identifica o bloco de países com índice de desenvolvimento humano menor ou igual 0,807, com esperança de vida menor ou igual a 73,2 anos, com percentual de domicílios com acesso à água potável menor de 93,0%, com percentual de domicílios com acesso à rede sanitária menor de 81,0%, com taxa de alfabetização das pessoas adultas menor de 93,5% e com total de calorias consumidas menor do que 2689 Kcal/dia.

Estes perfis de clusters coincidem exatamente as configurações evidenciadas pelos planos da HOMALS.

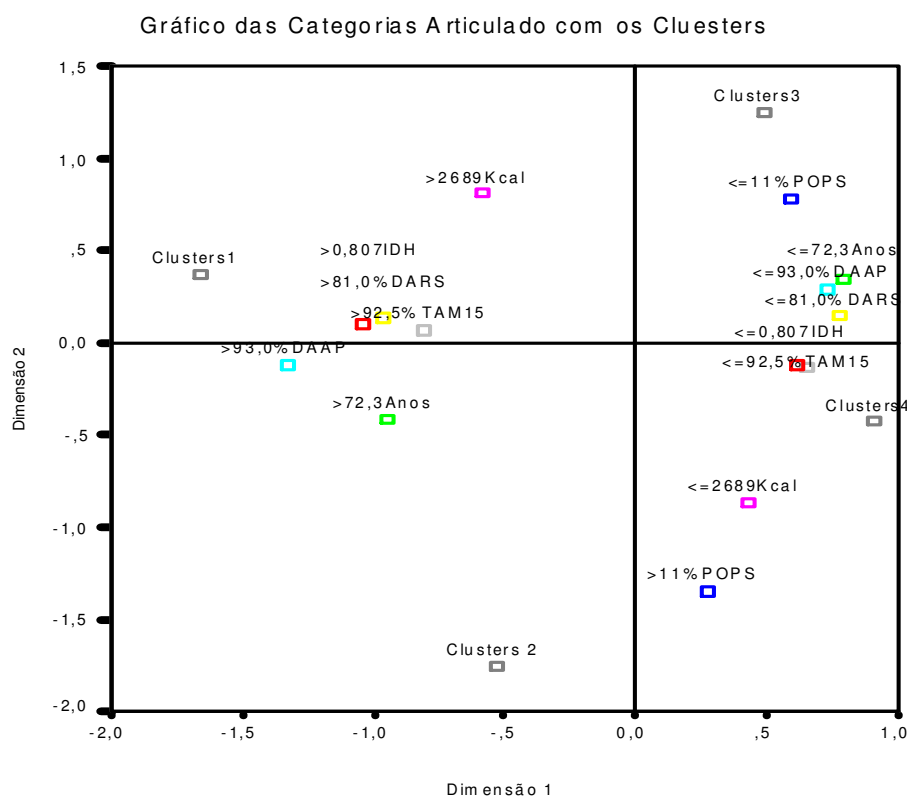
Tabela 7- Caracterização dos Clusters

Variáveis	Clusters1	Clusters 2	Clusters3	Clusters4
<=0,807IDH	0,0	50,0	66,7	100,0
>0,807IDH	100,0	50,0	33,3	0,0
<=72,3Anos	0,0	0,0	100,0	75,0
>72,3Anos	100,0	100,0	0,0	25,0
<=11%POPS	0,0	0,0	100,0	50,0
>11%POPS	0,0	100,0	0,0	50,0
<=2689Kcal	0,0	100,0	0,0	100,0
>2689Kcal	100,0	0,0	100,0	0,0
<=93,0%DAAP	0,0	0,0	100,0	100,0
>93,0%DAAP	100,0	50,0	0,0	0,0
<=81,0%DARS	0,0	0,0	66,7	100,0
>81,0%DARS	100,0	50,0	33,3	0,0
<=92,5%TAM15	0,0	50,0	66,7	75,0
>92,5%TAM15	100,0	50,0	33,3	25,0

Para validar graficamente a consistência desta classificação, pode ainda proceder-se à projeção dos quatro clusters no plano cuja configuração os havia sugerido. É o que mostra o Gráfico 8. O gráfico permite verificar que a posição ocupada por cada um dos clusters está muito próxima das diversas categorias que caracterizam os respectivos perfis da HOMALS. Conforme evidencia a representação gráfica, a tipologia de segmentação afigura-se coerente com os resultados logo obtidos com a realização da Análise de Homogeneidade.

Do ponto de vista técnico, consiste numa representação gráfica na qual estão simultaneamente representadas as variáveis que configuram os perfis, como variáveis ativas, e a variável que identifica a tipologia de segmentação, tendo esta o *status* de variável passiva (Helena Carvalho, 2004). O objetivo é observar a disposição dos quatro clusters. O resultado da operação se encontra no Gráfico 8. Percebe-se que as categorias da variável passiva (clusters 1, clusters 2, clusters 3 e clusters 4) estão exatamente juntas com as configurações da HOMALS.

Gráfico 8 - Disposição dos Clusters no Espaço de Países



4-Conclusão

Na Análise de Homogeneidade, as estatísticas são geradas por médias e seus valores finais vêm de uma iteração ótima.

As localizações das categorias e objetos nos planos são estabelecidas através de medias de distância que neste caso são métricas do qui-quadrado.

Os valores próprios são estatísticas que indicam a qualidade das dimensões, isto é, indicam as dimensões mais importantes para a estruturação do espaço de análise. As medidas de discriminação avaliam as variáveis mais importantes em cada dimensão mais importante. Já as quantificações das categorias indicam as categorias mais importantes nas variáveis mais importantes em cada uma das dimensões mais importantes. Pode-se também realizar uma análise interna das dimensões através das quantificações das

Os planos são estatísticas visuais, gráficas, que permitem a análise de estruturação do espaço de análise.

Nos planos das categorias as configurações presentes são “desenhadas” por categorias do espaço de análise. Este plano permite saber **como** é cada grupo evidenciado.

Nos gráficos dos objetos, as configurações são constituídas por objetos do espaço de partida. Este plano permite saber **quem** é cada grupo evidenciado.

Os dois gráficos devem ser analisados conjuntamente para que se tenha uma análise completa da estrutura do espaço que se estuda.

5-Referências Bibliográficas

- [1]CARVALHO, HELENA. *Análise Multivariada de Dados Qualitativos-Utilização da HOMALS com SPSS*. Edições Silabo. Lisboa, 2004.
- [2]GEER, JOHN. *Multivariate Analysis of Categorical Data: Theory*. Sage Publications. USA,1993.
- [3]GIFI, ALBERT. *Nonlinear Multivariate Analysis*. England. John Wiley & Sons.1996.
- [4]GREENACRE, MICHEL. *Theory and Applications of Correspondence Analysis*. Academic Press. London, 1996.