

## RESOLUÇÃO DO EXERCÍCIO: "TURISMO"

Este documento deve ser distribuído aos participantes pelos Responsáveis do Seminário



## Programme PRESTA

PROGRAMME DE RECHERCHE ET D'ENSEIGNEMENT EN STATISTIQUE APPLIQUÉE

### Resolução de um exercício de aplicação da Análise Fatorial de Correspondências Simples

**Objetivo do exercício:** verificar a possibilidade de construir uma topologia dos "lugares de procedência" dos turistas, segundo as preferências por certos períodos do ano para realizar a visita de Foz de Iguaçu pelo lado argentino.

#### III. Etapas de resolução do exercício

##### III.1. Verificação formal dos dados

1. Quantas linhas figuram no arquivo TURISMO.DON?

**Resposta :** 9 linhas.

2. Quantas colunas figuram no arquivo TURISMO.DON?

**Resposta :** 13 colunas.

3. Qual é o significado da primeira coluna do arquivo TURISMO.DON?

**Resposta :**

A primeira coluna desse arquivo contém as etiquetas rótulos das linhas da Tabela Contingência, isto é os "lugares de procedência" dos turistas.

4. Indique o conteúdo e explique o significado da última coluna do arquivo TURISMO.DON.

**Resposta :**

Conteúdo: 17, 56, 17, 27, 13, 14, 5, 36, 10

Trata-se da distribuição por "lugar de procedência" da amostra de turistas que visitaram Foz de Iguaçu durante o mês de Dezembro.

5. Indique o conteúdo e explique o significado da primeira linha do arquivo TURISMO.DON.

**Resposta :**

Conteúdo: CapFeder, 29, 29, 37, 41, 23, 29, 68, 30, 33, 28, 17, 17

Trata-se da distribuição mensal de visitantes de Foz de Iguaçu que provêm da Capital Federal (Argentina). Os oito primeiros dígitos são a etiqueta correspondente a essa linha. Note-se que, nesse arquivo, a etiqueta das linhas não está entre aspas e que o separador dos valores em coluna é um caracter branco.

6. Indique o conteúdo e explique o significado da primeira linha do arquivo TURISMO<DIC.

**Resposta :** Conteúdo: 00001 Janeiro

Ao ser a primeira linha desse arquivo, seu conteúdo refere-se à primeira “variável” da tabela de dados. Os quatro primeiros dígitos indicam que a primeira variável é declarada como uma variável contínua (valor 1), dado que é uma distribuição de frequências brutas. Trata-se da distribuição por “lugar de procedência” do total de turistas da amostra que declararam ter visitado o lugar durante o mês de “Janeiro”.

7. Indique o conteúdo e explique o significado da última linha do arquivo TURISMO.DIC.

**Resposta :**

Conteúdo : DICI Diciembre

Ao ser a linha N°24 desse arquivo, seu conteúdo refere-se à variável N°12 e última da tabela de dados. Os quatro primeiros dígitos constituem a “etiqueta curta”: associada a essa variável (modalidade da variável em coluna da Tabela de Contingência). A etiqueta completa associada à mesma variável é “Diciembre”.

### III.2 Primeira leitura dos dados

1. Quais são os três meses do ano em que houve maior afluência de turistas em Foz de Iguaçu, pelo lado argentino?

**Resposta :**

No mês de Julho: 15,4%; no mês de Janeiro: 12,1% e no mês de Fevereiro: 10,6%. A afluência de turistas às cachoeiras durante esses meses representam 38% do total de turistas que visitaram Foz de Iguaçu durante todo o ano.

2. Em 1993, Semana Santa caiu no mês de Abril.

Por outro lado, consideramos que os turistas de uma procedência dada manifestam uma “preferencia” particular por visitar Foz de Iguaçu em um período dado do ano, quando a proporção observada de turistas dessa procedência, nesse período do ano, é superior à proporção de turistas dessa procedência durante todo o ano.

Qual é o “lugar de procedência” dos turistas que “preferem” visitar Foz de Iguazu durante as férias de Semana Santa (a Páscoa)?

**Resposta :**

Em primeiro lugar, são os turistas argentinos quem preferem visitar Foz de Iguazu durante as férias da Semana Santa. Com efeito, a distribuição de turistas pelo lugar de procedência no mês de Abril, mostra que 26,2% provêm de Buenos Aires; 21,4% do resto da Argentina, 17,6% das províncias da Mesopotâmia argentina e 14,1 % da Capital Federal. Isto é que 79% do total de turistas que visitaram Foz de Iguazu durante as férias da Semana Santa são turistas argentinos. Mas deve-se ter em conta que estes turistas constituem 75% do fluxo anual de turistas de Foz de Iguazu.

Por outro lado, essa apreciação global da preferência dos turistas argentinos pelo período das férias da Semana Santa, deve ser corrigida se considerarmos os diferentes “lugares de procedência” para o conjunto de turistas argentinos.

Lugar de procedência	% de visitantes na Semana Santa	% de visitantes durante o ano
em Argentina		
CapFeder	14.1	12.1 (baixa preferência)
BueAires	26.2	34.1 (sem preferência)
Mesopota	17.6	9.4 (alta preferência)
RestPaís	21.4	14.9 (preferência marcada)

Também pode-se considerar que os turistas latino-americanos (com exceção dos brasileiros e paraguaios) “preferem” visitar Foz de Iguazu durante as férias de Semana Santa. Com efeito, no mês de Abril se observou 5,5% de turistas provenientes dos países latino-americanos, enquanto que estes constituem só 2,8% do fluxo anual de turistas em Foz de Iguazu.

3. Quais são os meses “preferidos” pelos turistas europeus para visitar Foz de Iguazu pelo lado argentino?

**Resposta :**

Os turistas europeus “preferem” visitar Foz de Iguazu em todo o mês de Novembro (14%, enquanto que os turistas do mês de Novembro representam só 17% do total de turistas do ano observados na amostra), em Agosto (13,6%, enquanto que os turistas do mês de Agosto representam só 8% do total de turistas do ano observados na amostra) e em Março (12,2%, enquanto que os turistas do mês de Março representam só 8% do total de turistas do ano observados na amostra).

A distribuição de turistas dessa procedência, no ano, mostra uma “preferência” dos mesmos por visitar Foz de Iguazu durante os períodos de Agosto a Dezembro e Janeiro a Fevereiro. Por outro lado, dos 71% dos turistas europeus que visitam Foz de Iguazu pelo lado argentino o fazem nesses períodos, enquanto que — sobre o total de turistas do ano, observado na amostra — só 54,4% deles visitam Foz de Iguazu durante os meses de Agosto a Dezembro e Janeiro a Fevereiro

4. Quais são os meses “preferidos” pelos turistas norte-americanos para visitar Foz de Iguaçu pelo lado argentino?

**Resposta :**

Os turistas norte-americanos “preferem” visitar Foz de Iguaçu em Dezembro (8%, quando os turistas do mês de Dezembro representam 6,2% do total de turistas do ano, observado na amostra) Janeiro (16,7%, quando os turistas do mês de Janeiro representam 12,1% do total de turistas do ano, observado na amostra) e em Fevereiro (13,8%, quando os turistas do mês de Fevereiro representam 10,6% do total de turistas do ano, observado na amostra).

**III.3 Aplicação do método de A.F.C. à tabela de dados e interpretação dos resultados da análise**

1. O arquivo TURISMO.DON tem “formato livre”, com os rótulos das linhas sem aspas?

**Resposta :**

Sim, trata-se de um arquivo em formato ASCII, com o caracter branco como separador de valores em coluna e com os rótulos de linhas em primeira coluna, sem aspas.

2. Quantos indivíduos ativos devem ser declarados na análise dessa tabela de dados?

**Resposta :**

9 linhas ou ‘indivíduos’ ativos. São as 9 linhas da Tabela de Contingência estudada.

3. Consideram-se “indivíduos” suplementares ou ilustrativos?

**Resposta :**

Não. Neste estudo não são considerados distribuições em linha suplementar. Os dados não contêm perfis linha suplementar.

4. Quantas ‘variáveis’ ativas e suplementares serão declaradas na análise dessa tabela de dados?

**Resposta :**

12 “variáveis” ativas correspondentes às 12 modalidades da variável nominal localizada em colunas da tabela de dados.

Não são declaradas “variáveis” suplementares.

5. Quantos eixos fatoriais poderão ser definidos, como máximo, no estudo dessa tabela de dados?

**Resposta :**

A descomposição da inércia em um sistema de eixos ortogonais não poderá produzir mais de 9 eixos, o que corresponde à menor dimensão da tabela de dados.

6. Qual é a sequência de procedimentos que executará o programa SPAD.N a partir desse arquivo de parâmetros?

**Resposta :** ARDIC - ARDON - SELEC - CORBI -DEFAC - GRAPH

7. Como se chama (por defeito) o arquivo de resultados e em que diretório foi criado pelo programa SPAD.N?

**Resposta :**

O arquivo de resultados chama-se TURISMO.LST e figura no diretório corrente: "TURISMO".

### III.4 Interpretação dos resultados da análise fatorial

1. Qual é o valor do índice de associação  $\chi_{\text{obs}}^2$  associado a tabela de dados? Justifique a resposta.

**Resposta:**

Trata-se do índice de associação:

$$\varphi_{\text{obs}}^2 = \sum_{\alpha} \lambda_{\alpha} = 0.1098$$

Seu valor mínimo é 0 (associação inexistente, independência de caracteres observados) e seu valor máximo pode ser definido assim :

$$\text{Max}(\varphi_{\text{Obs}}^2) = \min[(9-1); (11-1)] = 8 .$$

Lembre-se também que :  $\chi_{\text{Obs}}^2 = \varphi_{\text{Obs}}^2 \times n_{..} = 0.1098 \times 3150 = 345.87 .$

Pode-se definir um índice de associação relativo a esse valor máximo:

$$\varphi_{\text{rel}}^2 = \frac{\varphi_{\text{Obs}}^2}{8} \times 100 = \frac{0.1098}{8} = 1.4\% .$$

Observa-se que a associação entre o "lugar de procedência" dos turistas e os "meses" de visita Foz de Iguaçu, é uma associação bastante fraca.

2. Considerando a tabela dos auto-valores que resultam da decomposição dessa inércia total, quantos eixos fatoriais devem ser interpretados nesta análise? Justifique sua resposta.

**Resposta :**

Aplicando a regra "do cotovelo" e a regra de seleção dos eixos que apresentam auto-valores associados superiores ao auto-valor médio, devem-se considerar nesta análise os três primeiros eixos fatoriais.

O auto-valor associado ao quarto eixo é já bastante baixo e inferior ao auto-valor médio.

$\lambda_{\text{prom}} = 0,1098/8$ , os auto-valores de ordem superior ao 4º, apresentam já um decréscimo regular.

3. Quais são os 3 “lugares de procedência” dos turistas cujos perfis mensais de visita a Foz de Iguazu são os mais diferentes do perfil-marginal correspondente? Justifique sua resposta.

**Resposta :**

Punto-Perfil	DISTO	P.Rel.
OtroPais	0.35	2.16
RestAmer	0.25	14.03
EuroOcci	0.24	14.03

4. Quais são os pontos-perfis “lugares de procedência” de maior coordenada e de maior contribuição à inércia, sobre o primeiro eixo fatorial?

**Resposta :**

Punto-Perfil	Coord.	Contribución
RestPais	0.23	15.5%
Mesopota	0.23	9.9%
BueAires	0.08	4.1
BrasPara	-0.13	1.7
OtroPais	-0.37	5.8%
EuroOcci	-0.47	62.0%

5. Tendo em conta as coordenadas, as contribuições e as qualidades de representação de todos os pontos perfis “lugares de procedência” sobre o primeiro eixo fatorial, que interpretação pode dar desse eixo no espaço dos lugares de procedência dos turistas?

**Resposta :**

O primeiro eixo fatorial opõe os perfis dos turistas ‘argentinos’( com exceção daqueles que provêm da Capital Federal) aos perfis dos turistas “estrangeiros” (com exceção dos turistas norte-americanos e latino-americanos que não provêm do Brasil ou do Paraguai).

6. Quais são os 3 “meses” nos quais a distribuição de turistas segundo o “lugar de procedência” dos mesmos, são os mais diferentes do perfil-marginal correspondente? Justifique sua resposta.

**Resposta :**

Punto-Perfil	DISTO	P.Rel.
JUNIO	0.29	3.97
NOVIEMBRE	0.22	7.05
ABRIL	0.21	9.21

7. Quais são os 3 'meses' de maior afluência de turistas?

Resposta :

Punto-Perfil	DISTO	P.Rel.
JULIO	0.09	15.40
ENERO	0.08	12.06
FEBRERO	0.07	10.57

8. Quais são os pontos-perfis 'meses' de maior coordenada e de maior contribuição à inércia, sobre o primeiro eixo fatorial?

Resposta :

Punto-Perfil	Coord.	Contribución
ABRIL	0.27	13.2
JULIO	0.25	19.2
FEBRERO	0.18	6.7
ENERO	0.14	4.5
MAYO	-0.13	1.6
DICIEMBRE	-0.20	4.8
AGOSTO	-0.26	10.8
MARZO	-0.31	15.2
NOVIEMBRE	-0.40	21.8

9. Tendo em conta as coordenadas, as contribuições e as qualidades de representação de todos os pontos perfis "meses" sobre o primeiro eixo fatorial, que interpretação pode se dar desse primeiro eixo no espaço dos "meses"?

Resposta :

No espaço dos 'meses', o primeiro eixo fatorial opõe os perfis-meses que correspondem às férias de verás e interrupção sa Semana Santa aueles que más correspondem as férias escolares e profissionais, quer no hemisfério norte quer no hemisfério sul. 'lugares de procedência' dos turistas.

10. Interprete a segundo eixo fatorial no espaço dos puntos-perfis "lugares de procedência" dos turistas. Justifique sua respuesta.

Resposta :

O segundo eixo fatorial opõe os perfis dos turistas 'argentinos' que residem perto de Foz de Iguazu e os turistas estrangeiros europeus e latino-americanos (com exceção dos turistas brasileiros e paraguaios), aos perfis dos turistas norte-americanos e argentinos de Buenos Aires.

Punto-Perfil	Coord.	Contribución	Coseno Cuadrado
RestAmer	0.28	10.1	0.33
Mesopota	0.24	24.1	0.36
EuropOcc	0.12	9.4	0.06
RestPais	0.08	4.8	0.10
BueAires	-0.09	12.1	0.32
USACanad	-0.38	35.8	0.65

11. Interprete o segundo eixo fatorial no espaço dos pontos-perfis meses. Justifique sua resposta.

**Resposta :**

No espaço dos 'meses', o segundo eixo fatorial opõe os meses de turismo de interrupções de curta duração (Semana Santa e Festas Religiosas de Novembro) aos meses de turismo de férias estáveis (Fevereiro - Janeiro) ou do bimestre Maio - Junho.

Punto-Perfil	Coord.	Contribución	Coseno Cuadrado
ABRIL	0.35	50.6	0.57
NOVIEMBRE	0.13	5.1	0.07
AGOSTO	0.08	2.2	0.07
OCTUBRE	0.06	1.3	0.20
ENERO	-0.06	1.6	0.04
MAYO	-0.12	3.1	0.16
FEBRERO	-0.13	8.1	0.23
JUNIO	-0.38	26.5	0.50

12. Interprete o terceiro eixo fatorial no espaço dos pontos-perfis "lugares de procedência" dos turistas. Justifique sua resposta.

**Resposta :**

O terceiro eixo fatorial opõe os perfis dos turistas da capital federal e dos países latino-americanos (com exceção do Brasil e do Paraguai), aos perfis dos turistas argentinos da Mesopotamia, dos turistas norte-americanos e dos europeus.

A inércia projetada sobre este eixo é constituída essencialmente pelos pontos-perfis CapFeder e Mesopota (77,2%). Deste modo, este eixo acrescenta uma correção importante à informação dada pelo o primeiro eixo fatorial, o qual agrupa os pontos-perfis de todos os lugares de procedência argentinos no lado positivo do mesmo. Este eixo mostra que os turistas da Capital Federal apresentam um perfil mensal de frequência de Foz de Iguaçu que é diferente do resto dos turistas argentinos e em particular oposto ao perfil mensal de visita de Foz de Iguaçu dos turistas da mesopotamia argentina.

Punto-Perfil	Coord.	Contribución	Coseno Cuadrado
Mesopota	0.16	16.4	0.15
USACanad	0.14	8.1	0.09
RestAmer	-0.22	9.5	0.19
CapFeder	-0.26	60.8	0.80

13. Interprete o terceiro eixo fatorial no espaço dos pontos-perfis "meses". Justifique sua resposta.

**Resposta :**

No espaço dos 'meses', o terceiro eixo fatorial opõe os meses de turismo de férias estáveis no hemisfério sul, aos meses que (na Argentina) apresentam interrupções de curta duração (feriado escolar devido festas cívicas ou religiosas).



Punto-Perfil	Coord.	Contribución	Coseno Cuadrado
ENERO	0.22	42.6	0.65
FEBRERO	0.09	6.6	0.12
DICIEMBRE	0.10	4.3	0.12
SEPTIEMBRE	-0.07	2.5	0.13
ABRIL	-0.09	5.3	0.04
MAYO	-0.13	5.7	0.19
MARZO	-0.12	8.0	0.10
JULIO	-0.10	10.1	0.10
JUNIO	-0.20	11.9	0.14

14. Interprete o primeiro eixo fatorial. Justifique sua resposta.

**Resposta :**

Pela posição dos pontos-perfis linha e coluna, localizados respeitando uma dupla relação baricêntrica, pode-se formular as seguintes hipóteses sobre as preferências dos turistas segundo seus lugares de procedência:

- Os turistas argentinos (residentes perto de Foz de Iguaçu : Mesopotâmia argentina, ou do resto do país salvo de Buenos Aires) preferem visitar Foz de Iguaçu durante os meses de férias escolares, mas sobretudo durante as férias da Semana Santa.

- Os turistas norte-americanos preferem visitar Foz de Iguaçu particularmente durante o mês de Junho.

- Os turistas de Europa Ocidental visitam Foz de Iguaçu preferencialmente nos meses de Novembro, Agosto e Dezembro, igual aos turistas brasileiros e paraguaios.

- Os turistas provenientes dos países latino-americanos não limítrofes com Foz de Iguaçu, visitam Foz de Iguaçu preferentemente nos meses de Novembro, Agosto e Dezembro, mas também parecem ter uma certa preferência pelo período das férias da Semana Santa (Abril).

- Quanto aos restantes turistas estrangeiros, a posição do ponto-perfil OUTROS no primeiro plano fatorial sugere que os meses preferidos por eles para visitar Foz de Iguaçu são: Março, Maio e Dezembro.

Os períodos de preferência dos turistas provenientes da Capital Federal não podem ser determinados a partir desse plano fatorial, dado que esse ponto-perfil está mal representado no mesmo. Por outro lado, observa-se diferentes pontos-perfis em posição “intermédia” (tanto no espaço dos “meses”, como no espaço dos “lugares de procedência”) nesse primeiro plano fatorial. De modo que não é possível definir uma tipologia precisa e completa dos “lugares de procedência” dos turistas, segundo os períodos do ano que preferem visitar Foz de Iguaçu, referindo-se unicamente às duas aproximações da tabela de dados (os dois primeiros eixos fatoriais).

### III.5 Construção de uma tipologia dos “lugares de procedência” dos turistas segundo os períodos do ano em que preferem visitar as Cataratas

1. A classificação e a partição dos perfis-linhas da tabela de Dados será feita a partir das coordenadas fatoriais. É razoável conservar os 3 primeiros eixos fatoriais para realizar a classificação hierárquica ascendente dos “lugares de procedência” dos turistas?

**Resposta :**

Poderia-se conservar todos os eixos fatoriais para a classificação dos “lugares de procedência” dos turistas. Sem dúvida, utilizando os escores fatoriais sobre os três primeiros eixos, estamos analisando numa óptica de classificação uma tabela de dados aproximada (correspondente à somadas três tabelas similares à tabela de dados, mas de dimensão 1, que podemos construir com os escores fatoriais dos três primeiros eixos) a qual conserva a estrutura colocada em evidência por esses três eixos, mas elimina os efeitos parasitários que mostram os eixos de ordem superior.

Trabalha-se então com uma tabela aproximada de dados que conserva a informação essencial da tabela de dados original “(a inércia total será então menor que a inércia total original, mas suficientemente grande) e evita-se o “ruído” que comporta inevitavelmente toda a tabela bruta observada.

2. Como se chama o arquivo que contém as coordenadas fatoriais produzidas na etapa anterior? Em que área foi gravado esse arquivo?

**Resposta :**

O arquivo que contém as coordenadas fatoriais se chama TURISMO.GUS e será criado pelo programa SPAD.N na área corrente de trabalho “TURISMO”.

3. A que corresponde a soma dos índices de nível que figura ao pé do histograma correspondente?

**Resposta :**

A soma dos índices de nível constitui a inércia total da tabela aproximada de dados utilizada. Neste caso, essa soma é igual a 0,087, o que corresponde à soma dos três primeiros valores próprios da análise fatorial de correspondências, dado que se conservaram esses três eixos de análise da tabela de dados original para estudar a classificação dos “lugares de procedência” dos turistas.

4. Considerando o dendograma da classificação hierárquica, observe a composição e o índice de nível correspondente ao último nó. Considera você que o último nó da classificação, à qual corresponde o índice de nível mais elevado, traduz em termos de diferenciação de perfis-linha a mesma informação que a análise fatorial mostrava em termos de oposição de perfis-linha sobre o primeiro eixo?

**Resposta :**

Não. Lembre-se que o primeiro eixo fatorial opõe os pontos-perfis “turistas argentinos” a todos os pontos-perfis “turistas estrangeiros”.

Por outro lado, a composição do último nó da classificação mostra uma diferenciação máxima entre o nó primogênito 13 e o nó benjamim 16. O nó 13, por sua vez, resulta da agregação (a muito baixos níveis do índice) do perfil Europa Ocidental com o nó 10, o qual resulta da agregação (a níveis ainda mais baixos do índice) dos elementos Outros Países e Brasil-Paraguai. De modo que o último nó da classificação mostra uma diferenciação máxima entre o grupo de perfis Europa Ocidental, Outros Países e Brasil-Paraguai e o resto dos lugares de procedência dos turistas.

5. A partir do histograma dos índices de nível e do dendograma da classificação hierárquica, quantas classes deveria ter a topologia a construir?

**Resposta :**

Parece razoável encurtar a árvore de classificação em quatro classes, isto é identificando os nós que apresentam os níveis mais baixos do índice, o que corresponde a agregações dos perfis-linha mais similares.

De modo que a partição será realizada a partir de uma configuração com base em quatro classes.

6. Qual é o valor da inércia total da partição em quatro classes e a que corresponde a mesma na etapa de classificação.

**Resposta :**

Inércia total = 0,087, corresponde à soma dos índices de nível que resultam da classificação.

7. As quatro classes resultantes da partição lhe parecem suficientemente homogêneas e bem diferenciadas entre si? Justifique sua resposta.

**Resposta :** Sim,

- As inércias intra-classe da quatro classes são pequenas (as classes são homogêneas).
- O coeficiente Inércia inter/ Inércia total = 82%, isto é a diferenciação das classes absorve 82% da inércia total (as classes estão bem diferenciadas).

8. Qual é a composição das classes da partição em quatro classes?

Resposta :

- Classe I: Capital Federal, Resto de América Latina (exceto Brasil e Paraguai).
- Classe II: Buenos Aires (Argentina), Brasil-Paraguai, USA-Canadá.
- Classe III: Resto do País (Argentina), Mesopotamia (Argentina).
- Classe IV: Europa Ocidental, Outros Países.

9. Comparando a composição de classes por partição com os resultados da classificação hierárquica ascendente, indique os nós da classificação que foram recompostos pela partição e os nós que se modificaram no final da mesma.

Resposta:

- Classe I: seus componentes correspondem à composição do nó 11 da classificação que agrupa os elementos Capital Federal e Resto de América Latina (exceto Brasil e Paraguai).

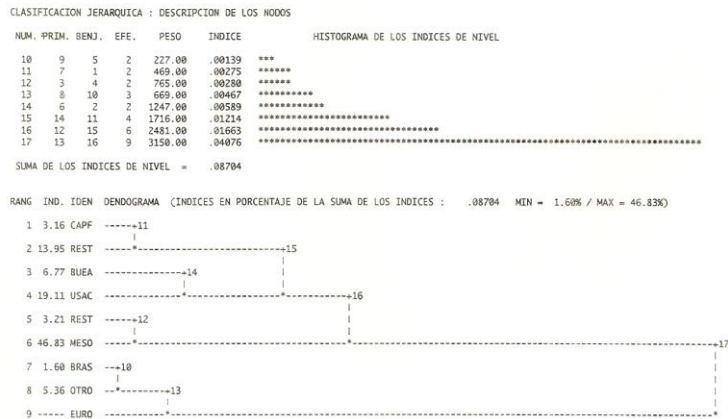
- Classe III: seus componentes correspondem à composição do nó 12 da classificação que agrupa os elementos Resto do País (Argentina), Mesopotamia (Argentina).

Note-se que os nós 11 e 12 da classificação hierárquica ascendente apresentaram índices de nível muito baixos e praticamente iguais (0,003).

- Os componentes do nó 14 da classificação hierárquica ascendente ficam incluídos na classe II da partição, mas agrega-se a eles o elemento Brasil - Paraguai. O índice de nível correspondente a esse nó é duas vezes maior do que os índices de nível dos nós 11 e 12 dessa classificação hierárquica ascendente.

- Os componentes do nó 13 da classificação hierárquica ascendente ficam incluídos na classe IV da partição, com excepção do elemento Brasil-Paraguai. O índice de nível correspondente a esse nó é também quase duas vezes maior que os índices de nível dos nós 11 e 12 dessa classificação hierárquica ascendente.

De modo então que as Classes I e III são as mais homogêneas do que as outras duas Classes.



### III.6 Retorno aos dados e apresentação das conclusões do estudo.

1. Quais são os meses “preferidos” pelos turistas de cada classe para visitar Foz de Iguaçu?

Resposta :

- Classe I : Capital Federal, Resto de América Latina (exceto Brasil e Paraguai). Na primeira linha da tabela de caracterização pelas frequências da tipologia em quatro classes, observa-se que (FRE/CLA = 6,8 %) 6,8% do efetivo de turistas que pertencem a esta Classe, visitam Foz de Iguaçu no mês de JUNHO, enquanto que a proporção de turistas desse mês, no total de turistas da amostra é só do 3,97%. Por outro lado, no total de turistas do mês de JUNHO, aqueles que pertencem à Classe I representam 25,6% (CLA/FRE = 25,6%), enquanto que a proporção de turistas dessa Classe no total de turistas da amostra é só de 14,89%.

De modo que, os turistas procedentes dos lugares que pertencem à Classe I “preferem” visitar Foz de Iguaçu pelo lado argentino durante mês de Junho. Essa frequência caracteriza suficientemente essa Classe de turistas. Por outro lado, um turista em cada quatro do mês de Junho, pertence à Classe I.

- Classe II : Buenos Aires (Argentina), Brasil - Paraguai, USA - Canadá. Fazendo um raciocínio similar, vê-se que os turistas procedentes dos lugares que pertencem a esta Classe II “preferem” visitar Foz de Iguaçu durante o mês de Fevereiro. Essa frequência caracteriza suficientemente essa Classe de turistas. Por outro lado, mais de um turista em cada dois do mês de Fevereiro, pertencem a essa Classe.

- Classe III : Resto do País (Argentina), Mesopotamia (Argentina). Os turistas procedentes dos lugares desta Classe III “preferem” visitar Foz de Iguaçu durante os meses de ABRIL e de JANEIRO. A frequência de ABRIL é a mais característica dessa Classe de turistas. Por outro lado, 39% dos turistas das férias de Semana Santa, pertencem à Classe III.

- Classe IV : Europa Ocidental, Outros Países. Os turistas procedentes dos lugares que pertencem a esta Classe IV “preferem” visitar Foz de Iguaçu do lado argentino durante os meses de NOVEMBRO, AGOSTO, MARÇO e DEZEMBRO. As frequências desses meses são suficientemente características dessa classe de turistas. Por outro lado, entre 24% e 30% dos turistas desses meses, pertencem a essa Classe.

2. Quais são, ao contrario, os meses menos “preferidos” pelos turistas de cada classe para visitar as Cataratas?

Resposta :

- Classe I : Capital Federal, Resto de América Latina (exceto Brasil e Paraguai). Na segunda linha da tabela de caracterização pelas frequências da tipologia em quatro classe, observa-se que (FRE/CLA = 6,82%) o 6,8% do efetivo de turistas que pertencem a esta Classe, visitam as Cataratas no mês de JANEIRO, enquanto que a proporção de turistas nesse mês, no total de turistas da amostra é de 12,06%. Por outro lado, do total de turistas do mês de Janeiro, aqueles que pertencem à Classe I representam 8,42% (CLA/FRE =

8,42%), enquanto que proporção de turistas dessa Classe no total de turistas da amostra é de 14,89%.

De modo que, os turistas procedentes dos lugares que pertencem à Classe I visitam as Cataratas do lado argentino durante o mês de JANEIRO numa proporção deficitária relativamente ao valor médio marginal. O mês de JANEIRO caracteriza suficientemente essa Classe de turistas como um período do ano “menos preferido” por esses turistas para visitar Foz de Iguaçu. Contudo, 31% dos turistas de Semana Santa, pertencem à Classe II.

- Classe III : Resto do país (Argentina), Mesopotamia (Argentina). Os turistas procedentes dos lugares que pertencem à Classe III visitam as Cataratas do lado argentino durante os meses de NOVENBRO e MARÇO numa proporção deficitária relativamente ao valor médio marginal. Esses meses caracterizam suficientemente essa Classe de turistas como períodos do ano “menos preferidos” por estes turistas para visitar Foz de Iguaçu. Por outro lado, vemos que só 15% dos turistas de NOVENBRO pertencem à Classe III e só 12% dos turistas de MARÇO pertencem a essa Classe

- Classe IV : Europa Ocidental, Outros Países. Os turistas procedentes dos lugares que pertencem à Classe IV visitam as Cataratas do lado argentino durante os meses de ABRIL, FEVEREIRO e JULHIO numa proporção deficitária relativamente ao valor médio marginal. Esses meses caracterizam suficientemente essa Classe de turistas como períodos do ano “menos preferidos” pelos mesmos para visitar Foz de Iguaçu. Por outro lado, vemos que só 10% dos turistas de Semana Santa pertencem à Classe IV, 9% dos turistas de FEVEREIRO pertencem a essa Classe e só 7% dos turistas de Julho pertencem a essa Classe.

3. Resuma suas conclusões com uma Tabela de Contingência de menores dimensões que apresente esse esquema de preferências para os diferentes lugares de procedência dos turistas e comente esses resultados. Essa Tabela de Contingência, assim como as tabelas de perfis correspondentes foram editadas pelo programa SPAD.N na listagem de resultados.

Como formularia sua resposta, destinada aos operadores de turismo da região de Posadas em Argentina, que encomendaram este trabalho...?

**Tabela de Dados: Frequência mensal de turistas segundo as regiões de procedência dos mesmos**

	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SETEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE	CONJUNTO
Clase I	32	38	43	57	28	32	81	34	41	32	29	22	469
Clase II	187	183	116	91	66	61	236	95	99	95	94	83	1406
Clase III	113	83	31	113	31	21	134	55	44	62	34	44	765
Clase IV	48	29	62	29	31	11	34	68	46	41	65	46	510
Total	380	333	252	290	156	125	485	252	230	230	222	195	3150

**Tabela de perfis en linha**

	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SETEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE	CONJUNTO
Clase I	7%	8%	9%	12%	6%	7%	17%	7%	9%	7%	6%	5%	100%
Clase II	13%	13%	8%	6%	5%	4%	17%	7%	7%	7%	7%	6%	100%
Clase III	15%	11%	4%	15%	4%	3%	18%	7%	6%	8%	4%	6%	100%
Clase IV	9%	6%	12%	6%	6%	2%	7%	13%	9%	8%	13%	9%	100%
Total	12%	11%	8%	9%	5%	4%	15%	8%	7%	7%	7%	6%	100%

**Tabela de perfis em coluna**

	ENERO	FEBRERO	MARZO	ABRIL	MAYO	JUNIO	JULIO	AGOSTO	SETEMBRE	OCTUBRE	NOVIEMBRE	DICIEMBRE	CONJUNTO
Clase I	8%	11%	17%	20%	18%	26%	17%	13%	18%	14%	13%	11%	15%
Clase II	49%	55%	46%	31%	42%	49%	49%	38%	43%	41%	42%	43%	45%
Clase III	30%	25%	12%	39%	20%	17%	28%	22%	19%	27%	15%	23%	24%
Clase IV	13%	9%	25%	10%	20%	9%	7%	27%	20%	18%	29%	24%	16%
Total	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%	100%

Formulação da resposta:

O estudo dos resultados da entrevista mostra certas “preferências” por diferentes épocas do ano por parte de turistas provenientes das regiões indicadas nas quatro classes definidas anteriormente.

Contudo, as “preferências” que foram evidenciadas nesta análise correspondem a “atrações” relativamente fracas entre períodos do ano e “lugares de procedência”, isto é baixa percentagens de incremento de turistas de certos lugares em certos períodos do ano (cf. &III.4.Q.1.).

Como não se dispõe das informações necessárias sobre o método de amostragem empregado nesse estudo, não é possível projetar essas “preferências” fracas à escala da população de referência. Mas, considera-se que a maioria das “preferências” colocadas em evidência nesta análise são da mesma ordem de importância que o erro ligado ao método de amostragem empregado.

Por isso, considera-se que este estudo das “preferências” dos turistas segundo as regiões de procedência permite:

- indicar as “preferências” observadas detalhadas nas conclusões e nas tabelas anteriores, mostrando que não existe homogeneidade entre as distribuições mensais, por região de procedência, dos turistas que visitam Foz de Iguaçu.
- indicar que as diferenças observadas nessas distribuições são menores, isto é que não justificam — em termos do mercado — uma adaptação específica da oferta mensal de serviços turísticos segundo a região de procedência dos turistas.



**NOTA :**

Observe que o valor  $\varphi_{\text{obs}}^2$  da Tabela I é igual a 0.08  $\varphi_{\text{obs}}^2 = \frac{\chi_{\text{obs}}^2}{n_{..}}$ , isto é igual à inércia total da etapa de partição e à soma dos índices de nível da etapa de classificação hierárquica ascendente.

Por outro lado, o  $\varphi_{\text{obs}}^2$  dessa Tabela é ligeiramente inferior ao  $\varphi_{\text{obs}}^2$  da tabela de Dados brutos.

Com efeito, sabemos que comparando os perfis de uma Tabela de Contingência com a distância do  $\chi^2$ , criamos um espaço de representação que respeita a propriedade de equivalência distribucional. Esta propriedade justifica que se adicionem dois perfis idênticos, sem que se produza uma perda de informação (diminuição do  $\varphi_{\text{obs}}^2$  da tabela).

Da mesma maneira, a propriedade de equivalência distribucional que caracteriza o espaço de representação proposto, justifica também a redução das dimensões da Tabela original adicionando os perfis mais similares. A redução da inércia total da tabela será tanto mais pequena quanto mais similares sejam entre eles os perfis que tem sido adicionados.

A partição que foi feita sobre as linhas da tabela de dados identificou os perfis mais similares, ainda que já se tenha consentido uma certa perda de informação tomando somente os três primeiros eixos fatoriais como base para essa partição.

A perda de informação resultante é então a seguinte :

Inércia total $\varphi_{\text{obs}}^2 = \sum_{\alpha} \lambda_{\alpha} = 0.1098$	Porcentagem com respeito a la inércia original 100%
Inércia conservada na classificação- partição-do conjunto de linhas $\sum (\text{ndices de nivel}) = \sum_{\alpha} \lambda_{\alpha} = 0.087$	Porcentagem com respeito a la inércia original 79.2%
Inércia da Tabela resumo $\varphi_{\text{obs}}^2 = \frac{\chi_{\text{obs}}^2}{n_{..}} = 0.08$	Porcentagem com respeito a la inércia original 69%

Como se pode observar, a parte mais importante da redução da inércia observada é devido à redução do número de eixos utilizados no processo de classificação - partição. Por outro lado, a agregação de linhas com perfis semelhantes produz uma redução mínima da informação contida na tabela.