

O TRATAMENTO GRÁFICO DE UM CONJUNTO DE DADOS - ESTUDO DA TÉCNICA MATRIZ ORDENÁVEL QUANTITATIVA

MÁRCIA MARIA DUARTE DOS SANTOS*

MIGUEL CÉSAR SANCHEZ**

Resumo

Enfoca-se a modalidade gráfica de tratamento de dados, a partir de princípios e regras da Semiologia Gráfica, sistematizada por Bertin. Descreve-se a técnica Matriz Ordenável Quantitativa, fundamentada na construção de matrizes permutáveis. E exemplifica-se sua utilização a partir do tratamento de um conjunto de dados.

Palavras-chave: Matriz Ordenável Quantitativa; Tratamento Gráfico; Semiologia Gráfica

Le Traitement Graphique d'un Ensemble de Données - Étude de la Technique Matrice Ordennable Quantitative

Resumé

On aborde la modalité de traitement graphique de données, en considérant les règles et les principes de la Sémiologie Graphique, systématisée par Bertin. On décrit la technique Matrice Ordennable Quantitative, établie sur la construction de matrices permutables. Et on demontre leur utilisation à partir du traitement d'un ensemble de données.

Mots-clé: Matrice Ordennable Quantitative; Traitement Graphique; Sémiologie Graphique.

* Departamento de Geografia Instituto de Geociências Universidade Federal de Minas Gerais

** Departamento de Planejamento Regional Instituto de Geociências e Ciências Exatas UNESP-Campus de Rio Claro

INTRODUÇÃO

A Semiologia Gráfica, sistematizada por BERTIN (1973, 2ªed.; 1977) tem sido amplamente utilizada e divulgada, em razão da eficácia do conjunto de princípios e regras que abrange aspectos semânticos, sintáticos e pragmáticos, a propósito da construção de mensagens gráficas temáticas.

Além comunicação de uma informação através de mensagens gráficas, a Semiologia aborda outros aspectos de interesse, relacionados ao levantamento, à armazenagem e à análise de dados, utilizando, sempre, recursos gráficos.

Esse campo abrangido pela Semiologia Gráfica é pouco divulgado, mas alguns dos seus procedimentos e técnicas têm se tornado mais conhecidos. É o caso, por exemplo, das técnicas de análise de um conjunto de dados, conhecidas como Matrizes Permutáveis.

Tem sido observado, entretanto, que a utilização de técnicas gráficas suscita reações adversas, que se sobrepõem às considerações das vantagens ou da pertinência do seu emprego. Uma dessas reações relaciona-se ao tempo e às habilidades pessoais demandados para se elaborar o material gráfico de tratamento - diagramas, mapas e redes. E outra diz respeito aos aspectos qualitativos do tratamento, considerado, por isso, sem objetividade.

Atualmente, os aspectos relacionados ao tempo gasto e às dificuldades de se elaborar material gráfico já estão, em grande parte, superados. Com a generalização do uso dos microcomputadores surgiram, e já são bastante conhecidos e utilizados, aplicativos computacionais que geram produtos gráficos e podem auxiliar o preparo do material especificado pelas técnicas da Semiologia.

Recentemente, iniciou-se a comercialização de programas que tornam possível a completa automação dos procedimentos que correspondem a algumas das técnicas de tratamento gráfico da informação.(1)

No momento, então, que já se pode contar com recursos que otimizam o trabalho do pesquisador interessado em aplicar técnicas gráficas, considera-se de interesse explorar os fundamentos desse tipo de tratamento, para patentear seu alcance.

Neste trabalho, serão enfocados os princípios do tratamento gráfico de um conjunto de dados e será descrita uma técnica de tratamento, fundamentada na construção de Matrizes Permutáveis, a Matriz Ordenável Quantitativa. Essa descrição compreenderá regras e procedimentos de processamento de um conjunto de dados, a partir de um exemplo de aplicação.

Este trabalho pode proporcionar, então, uma contribuição para a divulgação dos procedimentos gráficos de análise.

O TRATAMENTO GRÁFICO DE UM CONJUNTO DE DADOS

A Semiologia Gráfica estabelece um conjunto de princípios e define alguns procedimentos, que devem ser considerados para o entendimento e aplicação de um tratamento gráfico.

Esses princípios e procedimentos, que formam um verdadeiro sistema teórico, foram apresentados por BERTIN, num livro, cuja primeira edição data de 1967, no qual o autor aborda, também, aspectos da representação gráfica temática. Mais tarde, eles foram tratados, exaustiva e exclusivamente, pelo mesmo autor, em outro livro editado em 1977.

Esse sistema tem sido abordado, também, por BONIN (1975), que tem desenvolvido pesquisas originais sobre o assunto e contribuído, significativamente, para sua divulgação e aceitação nos meios científicos, não só na França, como em outros países. No Brasil, destaca-se o trabalho de LE SANN, através de cursos sobre o assunto e de orientações sobre a aplicação das técnicas gráficas de tratamento.(2)

O princípio mais básico desse sistema expressa que tratar graficamente um conjunto de dados compreende a simplificação pertinente desse conjunto, visando ao patenteamento de uma informação útil (BERTIN, 1977, p.12).

No sistema, o processo de simplificação é entendido como um procedimento que permite a redução do número de elementos de um conjunto de dados, quer os referidos como sujeitos- seres que se pretende analisar, ou seja, itens, localidades, fenômenos, etc, quer os que correspondem às características dos sujeitos- as variáveis ou atributos reunidos para caracterizá-los. E um processo de simplificação pertinente é compreendido como o que não suprime dados, reagrupando-os em novas categorias ou subconjuntos, diferentes das que possibilitaram sua apresentação no documento que os armazenou originalmente, tais como, uma matriz, um quadro ou uma tabela.

Considerando a Matriz 1, onde se apresenta os dados estudados neste trabalho, observa-se que estes estão dispostos segundo a classe a que pertencem. Cada coluna compreende um agrupamento, reunindo um subconjunto dos dados, que são da mesma espécie ou natureza, o que no sistema se denomina *componente*.(3)

Assim, todos os dados relacionados na primeira coluna referem-se aos cursos da UFMG- os seres que foram caracterizados nas colunas subseqüentes, através das percentagens de aprovados que obtiveram a metade dos pontos ou mais, em cada uma das provas realizadas na primeira etapa do vestibular de 1994 da UFMG. Nessas colunas dispõem-se, em cada uma delas, apenas os dados referentes aos resultados obtidos na prova, cuja denominação aparece no cabeçalho.

MATRIZ 1 - MATRIZ DE DADOS (Percentagem de aprovados no vestibular de 1994 da UFMG que obtiveram metade ou mais pontos na primeira etapa, por curso e prova)

Nº	CURSOS	PROVAS							
		Port.	Mat.	Geog	Hist.	L.Est	Fis.	Quím.	Biol.
1	Administração D.	100	92	80	78	98	78	52	74
2	Administração N.	96	96	74	66	92	78	58	64
3	Arquitetura	100	95	85	67,5	97,5	92,5	53,7	76,2
4	Belas Artes (L/B)	88,3	23,3	56,6	35	70	11,6	8,8	31,6
5	Biblioteconomia	91,4	20	40	32,8	45,7	4,2	4,2	14,2
6	C. da Computação	100	100	92,8	87,1	100	100	68,5	94,2
7	C.Biológ. D.(L/B)	100	76,2	88,7	53,7	93,7	41,2	50	95
8	C. Biológ. N.(L)	95	52,5	75	55	62,5	27,5	35	77,5
9	C. Contábeis	98,7	65	58,7	57,5	86,2	50	18,7	30
10	C. Econômicas	98,7	73,7	72,5	53,7	93,7	66,2	32,5	61,2
11	C. Sociais (L/B)	98,4	27,6	61,5	70,7	76,9	21,5	16,9	35,3
12	Comunicação Social	100	91,6	100	91,6	100	70	55	95
13	Direito	100	86,6	91,3	85,6	98	77,3	49,3	81,3
14	Ed. Física (L/B)	99	23	31	22	74	28,3	19	40
15	Enfermagem (L/B)	97,5	42,5	61,2	26,2	86,2	8,7	16,2	70
16	Eng. Civil	97,5	89,5	64	47	92,5	85,5	50,5	57,5
17	Eng. Elétrica	95	100	81,2	66,2	93,7	100	70	77,5
18	Eng. Mecânica	98,7	98,7	78,7	52,5	95	98,7	66,2	81,2
19	Eng. Metalúrgica	96	82	88	68	84	76	52	64
20	Eng. de Minas	95	75	75	32,5	62,5	45	25	42,5
21	Eng. Química	98	92	64	60	92	84	70	78
22	Estatística	93,3	83,3	63,3	46,6	96,6	23,3	23,3	53,3
23	Farmácia	100	51,6	52,5	37,5	90	60,8	56,6	81,6
24	Filosofia (L/B)	95	25	77,5	62,5	72,5	15	12,5	40
25	Física D. (B)	100	86,6	90	33,5	86,6	86,6	56,6	53,3
26	Física N. (L)	96,6	76,6	66,6	43,3	60	80	16,6	43,3
27	Fisioterapia	100	70	75	42,5	92,5	62,5	35	85
28	Geografia D.(L/B)	77,5	25	70	62,5	45	5	10	30
29	Geografia N.(L)	90	16,6	90	63,3	86,6	16,6	6,6	26,6
30	Geologia	63,3	60	76,6	46,6	83,3	33,3	26,6	33,3
31	História D. (L/B)	100	50	82,5	82,5	80	25	15	42,5
32	História N. (L/B)	95	30	97,5	82,5	77,5	25	12,5	40
33	Letras (L)	96,2	20,8	55,8	40,8	90,4	4,5	8,7	30,8
34	Matemática D.(L)	96,6	66,6	26,6	16,6	66,6	46,6	16,6	36,6
35	Matemática D.(B)	95	80	65	35	60	60	20	30
36	Matemática N.(L)	96,6	86,6	56,6	40	73,3	90	26,6	33,3
37	Medicina	100	99,3	99,6	90	100	95,3	96,5	100
38	Med. Veterinária	99,1	76,6	90	65,8	94,1	58,3	54,1	94,1
39	Música (L/B)	71,8	21,8	31,2	37,5	46,8	12,5	6,2	12,5
40	Odontologia	100	100	95	74,1	100	84,1	74,1	99,1
41	Pedagogia D.(L)	95	20	58,3	38,3	70	6,6	5	31,6
42	Pedagogia N.(L/B)	95	15	58,3	30	55	3,3	5	33,3
43	Psicologia (L/B)	100	61,6	83,3	63,3	94,1	24,1	26,6	77,5
44	Química D.(L/B)	95	62,5	87,5	35	55	32,5	52,5	52,5
45	Química N.(L)	96,6	60	80	56,6	60	46,6	66,6	56,6
46	Terap. Ocupacional	97,5	55	62,5	37,5	77,5	17,5	17,5	67,5
	UFMG	94,3	63,1	72	53,7	80,6	49,1	35,2	57,1

Fonte: COPEVE/UFMG

A apresentação desses dados, segundo as categorizações gerais de cursos e provas, decorre de uma escolha do pesquisador, que registrou, como ponto de partida para seu estudo, uma entre as várias relações que podem ser estabelecidas entre os dados. Nesse caso, foi escolhida a relação mais elementar existente entre os dados, que independe de um tratamento para ser patenteada.

A reunião e apresentação dos dados sob novas categorizações pertinentes, expressando outras relações que os dados estabelecem entre si, as de nível médio e superior, serão alcançadas, apenas, após um tratamento.

As relações de dados correspondentes a esses níveis são consideradas como as mais proveitosas ou úteis, uma vez que vão permitir o patenteamento de distribuições, agrupamentos, correlações e, conseqüentemente, a obtenção de classificações, tipologias, a consecução de generalizações, etc.

Mas as relações que um tratamento permite evidenciar de nível médio e superior, assim como as relações elementares, podem ser especificadas *a priori*, sob a forma de questões, pois correspondem a informações que se deseja obter. (BERTIN, 1973, p.141; BERTIN, 1977, p.11-13).

Essas informações correspondem, fundamentalmente, aos resultados visados pelo pesquisador. Elas referem-se às questões que são colocadas, objetivando-se o norteamento de uma análise, tendo em vista descrever, explicar ou compreender aspectos de um problema pesquisado.

Níveis e Tipo de Informação

Na Semiologia Gráfica, a informação corresponde a uma relação, referente à associação de dados, que pode ser definida pelo nível de abrangência dessas associações, a saber, o elementar, o médio e o superior.

Esses níveis dizem respeito, respectivamente, as correspondências que podem ser estabelecidas entre: um elemento do conjunto de dados, referido como sujeito, e um elemento de uma determinada característica; um número parcial de elementos do conjunto de sujeitos e um número parcial, também, de elementos de algumas características estudadas; e, por fim, de todos os elementos do conjunto de sujeitos e de todo o conjunto das características.

Assim, qualquer conjunto de dados encerra tantas informações elementares, quanto o número de elementos que compreende, e contém tantas informações de nível médio, quanto os subconjuntos de dados que possam ser discernidos, em função de objetivos de um tratamento. Mas, um conjunto só apresenta uma informação de nível superior e "*a partir du tableau des données, on ne peut accéder à un niveau plus élevé da la connaissance*" (BERTIN, 1977, p.13).

No sistema, uma informação pode ser definida, também, pela maneira como são formuladas as questões que orientam associações dos dados que se quer patentear. Há um tipo de informação que corresponde à apresentação de uma questão, em que os elementos referidos como sujeitos, ou um elemento desse subconjunto de dados, correspondem a uma variável dependente. E há a que introduz uma questão, considerando como variável dependente, ou variáveis dependentes, um ou vários elementos das características.

Os níveis e tipos de questões podem ser exemplificados, graficamente, como no esquema mostrado na Figura 1. E podem ser exemplificados, também, verbalmente, considerando o conjunto de dados que se apresenta na Matriz 1 e os objetivos da pesquisa que propiciaram seu levantamento.

Os dados apresentados na Matriz 1 foram reunidos, entre outros, para o desenvolvimento de uma pesquisa que visou aos seguintes objetivos gerais (SANTOS, 1994):

- estabelecer as características do vestibular de 1994 da UFMG, para os cursos de Geografia Diurno e Noturno, considerando o número de vagas ofertadas, de candidatos inscritos, de candidatos por vaga e de classificados por sexo, comparando-as com as dos demais cursos;
- definir o desempenho dos candidatos e dos aprovados no vestibular de 1994 da UFMG, para os cursos de Geografia, Diurno e Noturno, comparando-o com o dos candidatos e aprovados dos outros cursos, analisando-se o número de desistentes, o número de classificados, o de eliminados, de não-classificados, por curso e para as duas etapas do vestibular, e as percentagens de aprovados que obtiveram metade ou mais pontos por prova, segundo o curso e etapa.

Mas, no que toca particularmente aos dados mostrados na Matriz 1, sobre a percentagem de aprovados com metade ou mais pontos, por prova da primeira etapa e segundo os cursos da UFMG, observa-se que foram levantados para a consecução dos seguintes objetivos específicos, expressos através das questões que seguem:

- Como se classificam os cursos da UFMG, segundo o desempenho dos seus aprovados, nas provas realizadas na primeira etapa do vestibular de 1994?
- Nessa classificação, qual é a posição dos cursos de Geografia, em relação aos outros cursos da UFMG?
- Os aprovados nos cursos da UFMG apresentam desempenho diferenciado em função do tipo de prova, ou de um conjunto de provas, entre as realizadas na primeira etapa do vestibular?

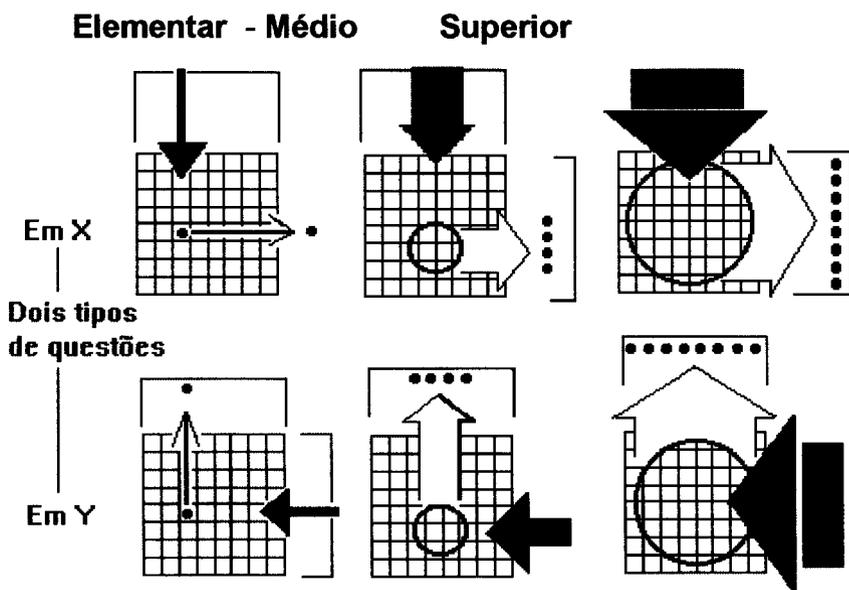
-Em quais provas, entre as realizadas na primeira etapa, os aprovados nos cursos de Geografia da UFMG apresentam um desempenho melhor?

Esses objetivos específicos correspondem, então, às informações que o pesquisador deseja obter. Nota-se que nenhuma das questões propostas pode ser respondida a partir de uma visualização rápida da matriz, indicando que elas devem corresponder ao nível médio e/ou superior de informação.

FIGURA 1

NÍVEIS E TIPOS DE INFORMAÇÃO

Níveis de Informação



Fonte: Bertin, J. *La Graphique et le Traitement Graphique de L' Information*. Paris: Flammarion, 1977. p. 13.

De fato, a primeira e a terceira questão correspondem a uma informação de nível superior, e a segunda e a terceira, a uma de nível médio. Essas questões exemplificam, também, os tipos de informação, apresentados anteriormente. Assim, a primeira e segunda referem-se ao tipo de questão formulada a partir das características, e a terceira e a quarta, a partir dos sujeitos.

Considerando os dados da Matriz 1 podem ser propostas, também, questões de nível elementar, tais como:

-Qual foi o desempenho dos aprovados no vestibular do curso de Geografia Diurno, na prova de Geografia, da primeira etapa? E o do curso Noturno?

Entretanto, o pesquisador não incluiu nenhuma questão de nível elementar entre os objetivos específicos que nortearam a análise dos dados apresentados na matriz. Mas, caso tivesse incluído, essa questão poderia ser visualizada, também, numa representação de tratamento gráfico.

Num tratamento gráfico, propicia-se a organização de um conjunto de dados numa representação que permite manipulações e a obtenção do nível superior da informação. Esse nível é o privilegiado, em razão da constatação de que *“lorsque la construction permet ce niveau de lecture tous les niveaux inférieurs sont lisibles alors que l’inverse n’est pas vrai”* (BERTIN, 1977, p.13).

Tipos de Representações e Tratamentos Gráficos

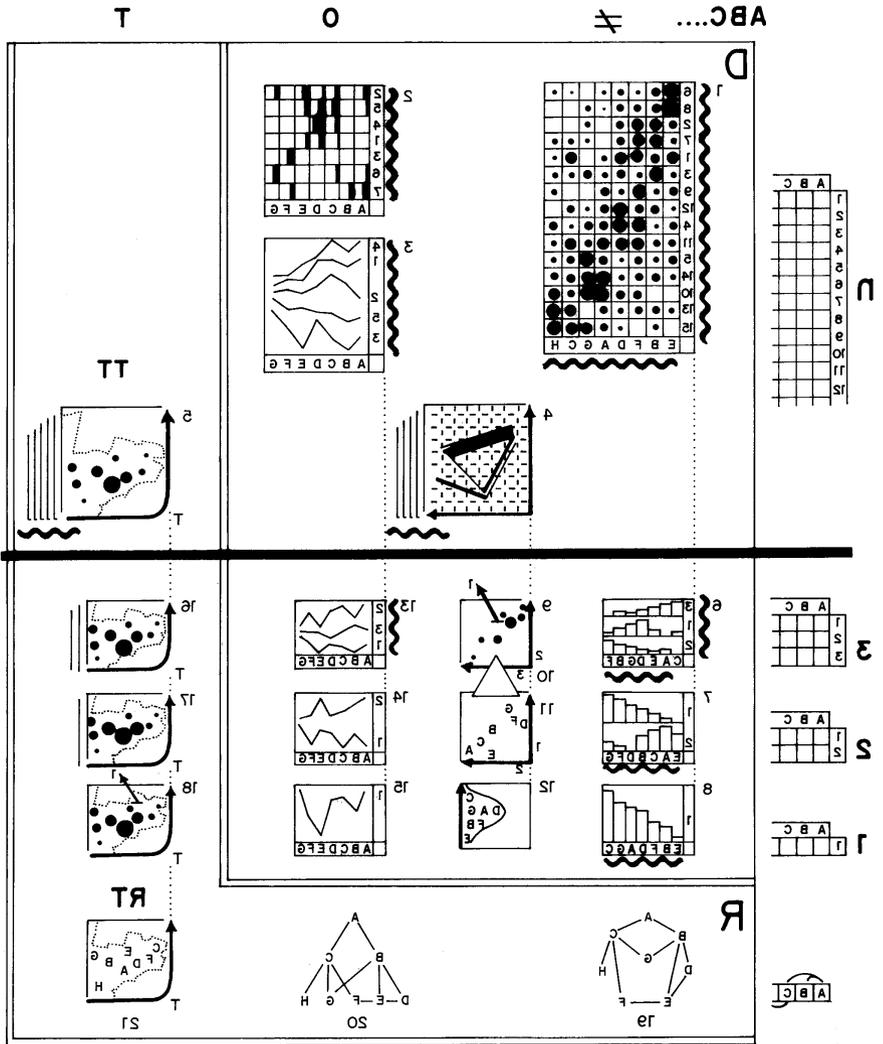
As representações que tornam possível tratamentos gráficos correspondem às mensagens que são construídas através de um sistema semiológico gráfico, monossêmico, tendo como suporte o plano. Elas abrangem os diagramas, as redes e os mapas.

Essas representações podem fundamentar uma classificação das técnicas de tratamento gráfico, como a que pode ser vista na Figura 2, num inventário simplificado.

No campo da Figura 2, designado por **D**, apresentam-se os tratamentos propiciados por tipos de diagramas, suas construções de base. No campo definido por **R**, estão reunidos os tratamentos que são possíveis a partir da construção de redes que podem ser, ou não, de natureza topográfica. As redes que correspondem à construção de mapas são designadas por **T**.

Essas construções de base, e somente elas, permitem manipulações de suas imagens, ou seja, dos seus elementos constitutivos, visando à simplificação dos dados e ao patenteamento das relações procuradas. Em função do modo como as imagens dessas construções são manipuladas, podem ser identificados os seguintes tipos de tratamento, denominados:

Fig. 5 - Inventário dos Tipos de Tratamento Gráfico



ABC... ≠: Ordenável O: Ordenado
 D: Diagramas e Classificações
 R: Redes RT: Redes Topológicas TT: Topologias Temáticas

Fonte: Bertini, J. La Graphique et le Traitement Graphique de la Information
 Paris: Flammarion, 1977, p. 29
 deê - Arnaldo Rosolem

- matriciais*, quando as manipulações transformam as imagens de suas construções de base, através de permutações de linhas e/ou de colunas dos diagramas (Figura 2, em 1, 2, 3, 6, 7, 8, 13, 14 e 15);
- sobreposição* bicromática ou tricromática, quando as manipulações modificam as imagens topográficas ao sobrepô-las (Figura 2, em 16 e 17);
- coleções*, no caso em que as manipulações das construções de base não as transformam, limitando-se a propiciar comparações das imagens, para classificá-las, segundo determinados critérios (Figura 2, em 4 e 5);

Há, ainda, tratamentos em que a simplificação dos dados é direta. As relações que um pesquisador procura evidenciar podem ser visualizadas, sem que sejam efetuadas manipulações. Mas, a visualização dessas relações é possível apenas a partir das construções de base propostas pelo sistema e em dois casos particulares:

- na construção de diagramas de correlação, quando o subconjunto de dados referentes às características são de nível de organização quantitativo e correspondem a três atributos, no máximo (Figura 2, em 9, 10, 11 e 12);
- e na construção de redes, topográficas ou não, a partir de um conjunto de dados que compreende dois subconjuntos de dados, um abrangendo os sujeitos e outro uma série de atributos (Figura 2, em 19, 20, 21 e 18).

A escolha dos tipos de tratamento gráfico, que estão associados a tipos de construções gráficas, é definida, considerando-se algumas propriedades dos dados a serem tratados. Essas características dizem respeito, de modo geral:

- ao número de subconjuntos de dados ou componentes que podem ser identificados no conjunto a ser analisado, referentes às categorizações originais escolhidas pelo pesquisador, anteriores a um tratamento;
- ao número de elementos do subconjunto de sujeitos e ao número de atributos destes;
- ao nível de organização desses subconjuntos, sujeitos e características, o que vem a ser a natureza da relação lógica que pode ser estabelecida, quando seus elementos são comparados, ou seja, uma relação de:
- simples diferenciação ou similitude, que define o nível de organização seletivo ou associativo (.,);
- ordem, que expressa o nível de organização ordenado (O);
- métrica, que define o nível de organização quantitativo (Q).(4)

Como pode ser visto na Figura 2, os tipos de tratamentos são definidos, considerando-se, primeiramente, o número de subconjuntos de dados que

correspondem às características e o nível de organização dos componentes sujeitos e características. Mas deve-se atentar para as dimensões da matriz de dados. Um número excessivamente grande de sujeitos e características inviabiliza ou dificulta as manipulações de elementos das representações de base que correspondem aos diagramas. E, desse modo, inviabiliza ou dificulta a utilização de algumas técnicas, uma vez que as manipulações é que tornam possível a realização de um tratamento gráfico.

De acordo com BERTIN (1977, p.31), as dimensões de uma matriz superiores às especificadas na Figura 3 podem dificultar as manipulações necessárias à aplicação de algumas técnicas, apresentadas, também, na figura citada.

FIGURA 3

LIMITES DAS DIMENSÕES DE UMA MATRIZ, SEGUNDO
O TIPO DE TRATAMENTO GRÁFICO

Técnicas	Sujeitos	Características
	Nº de Elementos	Nº de Elementos
Matriz Ordenável	120	120
Fichário-Matriz	1000	30
Fichário-Imagem	1000	Ilimitado
Inventário de Curvas	1000	Ilimitado

Ao se definir um tipo de tratamento, uma outra característica dos dados deve, também, ser considerada, quando o conjunto a ser tratado compreender um componente de natureza topográfica. Um subconjunto de dados dessa natureza apresenta o nível de organização seletivo, que permite tratar o conjunto de dados através de diagramas ou mapas-redes topográficas, ou seja, através de técnicas de tratamento matriciais, de coleções ou sobreposições cromáticas.

A escolha entre as construções de base, diagramas ou mapas deve ser baseada no número de elementos do subconjunto de dados de natureza topográfica e no número de atributos. Segundo BERTIN (1977, p.31), tendo-se um número muito grande de sujeitos, a coleção de cartas se justifica como instrumento de manipulação, mas se ocorrer o contrário, deve-se escolher o tratamento matricial, pois trata-se de

“un instrument de manipulation plus puissant, plus souple et plus précis que la carte. Celle-ci n’intervient qu’après la manipulation. Pour introduire les relations de proximité, on projette sur la carte les groupements découverts par la manipulation.”

APLICAÇÃO DAS REGRAS E PRINCÍPIOS DE UM TRATAMENTO GRÁFICO NO ESTUDO DE UM CONJUNTO DE DADOS

Os dados que são estudados neste trabalho, mostrados na Matriz 1, têm propriedades que devem ser conhecidas para se definir o tipo de tratamento que tornará possível a obtenção das informações especificadas nos objetivos, apresentados anteriormente.

Verifica-se que esse conjunto de dados reúne os seguintes *componentes*:

- curros*, composto por 47 elementos da mesma espécie ou natureza, ou seja, os cursos ofertados pela UFMG, para os quais são realizadas provas de vestibular;
- provas*, abrangendo 8 elementos, a saber, as provas que são realizadas na primeira etapa do vestibular da UFMG;
- percentagens de aprovados que obtiveram metade ou mais pontos*, por prova.

Esses *componentes* devem ser classificados, também, segundo o seu *nível de organização*. Assim, os componentes cursos e provas são seletivos (*S*), pois a relação que se pode estabelecer entre seus elementos é de simples diferenciação. Os componentes percentagem de aprovados com metade ou mais pontos, por prova, são quantitativos (*Q*), pois pode-se estabelecer uma relação métrica entre seus elementos.

As características dos dados descritas anteriormente podem ser expressas graficamente, como se apresenta na Figura 4.

Essas características indicam que o tratamento a ser aplicado ao conjunto de dados corresponde ao matricial, de acordo com o inventário mostrado na Figura 2. O tipo de tratamento recomendado, mais especificamente, é o que se denomina Matriz Quantitativa Ordenável.

FIGURA 4

PROPRIEDADES DOS DADOS TRATADOS GRAFICAMENTE

CURSOS \neq	PROVAS $\neq ; 8 \text{ el.}$
47 el	PERCENTAGEM DE APROVADOS Q

A Construção e a Análise de uma Matriz Ordenável Quantitativa

Todo tratamento gráfico da informação inicia-se pela organização do conjunto de dados, numa matriz ou tabela, onde os dados classificados como características e sujeitos são colocados nas dimensões x e y e as mensurações, em qualquer escala, em Z ou no interior das células.

Em função dos objetivos do pesquisador e na medida em que o tipo de mensuração dos dados permita, pode ser necessário realizar algum tipo de tratamento matemático e/ou estatístico dos dados originais, tais como: a efetuação de operações aritméticas simples; a extração de informações numéricas, como contagens simples, cálculos de percentagens ou outras taxas; ou ainda, a indicação de medidas estatísticas, como medidas de tendência central, de variabilidade e outras. Desse modo, o pesquisador pode ser levado a incluir linhas ou colunas, na tabela ou matriz original, ou, até mesmo, reconstruí-la, para expressar as novas mensurações. E essa construção será, daí em diante, considerada a tabela original, a partir da qual será construída a Matriz Ordenável Quantitativa.

No caso do conjunto de dados estudado neste trabalho, o pesquisador apresentou-o na forma comumente indicada, ou seja, numa matriz de dupla entrada,

na Matriz 1. Nota-se que os dados são apresentados em percentagens, indicadas para cada prova, em relação aos totais de aprovados, por curso. Observa-se, também, que o pesquisador julgou necessário apresentar a média dos aprovados da UFMG no vestibular de 1994, que obtiveram a metade ou mais pontos, por prova.

A etapa seguinte à organização dos dados numa tabela ou matriz corresponde à sua transcrição gráfica, ou seja, à construção das representações de base do tratamento escolhido.

Como uma Matriz Ordenável Quantitativa constitui um tipo de tratamento, fundamentado na construção de diagramas, trata-se, então, nessa etapa, de aplicar as regras e princípios gerais de construção de diagramas, sistematizados pela Semiologia Gráfica, que não são, fundamentalmente, muito diferentes das orientações clássicas da cartografia.

Essas regras e princípios observam que, dadas as propriedades dos dados, relativas ao número de componentes e seu nível de organização, os diagramas adequados à sua transcrição são do tipo ortogonal, colunas ou barras, cujo número total corresponde ao de elementos de um dos componentes seletivos. Assim, serão construídos tantos gráficos de colunas, quanto o número de elementos do componente sujeito, com um número de colunas correspondente ao de elementos do componente característica seletivo. Ou serão construídos, tantos diagramas de barras, quantos forem os elementos do componente característica seletivo, apresentando um número de barras igual ao de elementos do componente sujeitos.

O tipo de gráfico a ser construído, barras ou colunas, deve ser escolhido de modo a propiciar comparações significativas, tendo em vista os objetivos do pesquisador. Uma regra prática para essa escolha, quando se trabalha com dados percentuais, enuncia que os perfis que irão representar as quantidades devem ser paralelos ao sentido em que foram calculadas as percentagens ou ser paralelos à série de 100%, quando esta está indicada na matriz. Outra regra observa que, quando se tem interesse em indicar uma média nos perfis que serão construídos para se representar as quantidades, essa média deve ser única para todas as linhas ou colunas do diagrama, o que determina o tipo de representação - barras ou colunas.

É importante considerar, na organização de uma Matriz Ordenável Quantitativa, além da escolha do tipo de diagrama a ser construído, o problema da escala desses diagramas. Como se trata de um conjunto de diagramas, eles podem ser construídos cada um com a sua própria escala ou podem ser organizados a partir de uma escala comum. Dessa escolha dependerá o favorecimento de determinadas percepções.

A percepção dos totais é imediata, quando se escolhe uma escala comum. Nesse caso, comparações tipicamente quantitativas são favorecidas entre os perfis

de um e outro gráfico, ou ainda, entre todos os perfis ou todos os gráficos da matriz de tratamento. A percepção de distribuições e agrupamentos é favorecida, ao se escolher escalas próprias para cada série de dados. Nessas representações, os totais são expressos nos eixos dos gráficos, e, embora não possam ser visualizados, podem ser lidos.

Na Figura 5, exemplifica-se as afirmações anteriores, apresentando em **1** e **2**, respectivamente, diagramas construídos em escala comum e própria, relativos aos resultados das provas de Português e Química, de acordo com os dados mostrados na Matriz 1.

Como o objetivo que se pretende atingir com a construção de uma matriz é o patenteamento da distribuição dos dados, tendo em vista o estabelecimento de agrupamentos, os diagramas que irão compô-la devem ser organizados em escala própria e seus eixos devem se cruzar no valor mínimo da série de dados quantitativos representada, como se apresenta na Figura 5, em 2. Entretanto, quando os dados que se deseja tratar apresentam componentes quantitativos que não mostram valores máximos e mínimos muito diferenciados, a representação desses dados através da escala comum pode propiciar a percepção de distribuições, além, é claro, dos totais.

Após a definição do número de diagramas necessários à transcrição dos dados e da escala a ser adotada, eles devem ser construídos, observando-se que:

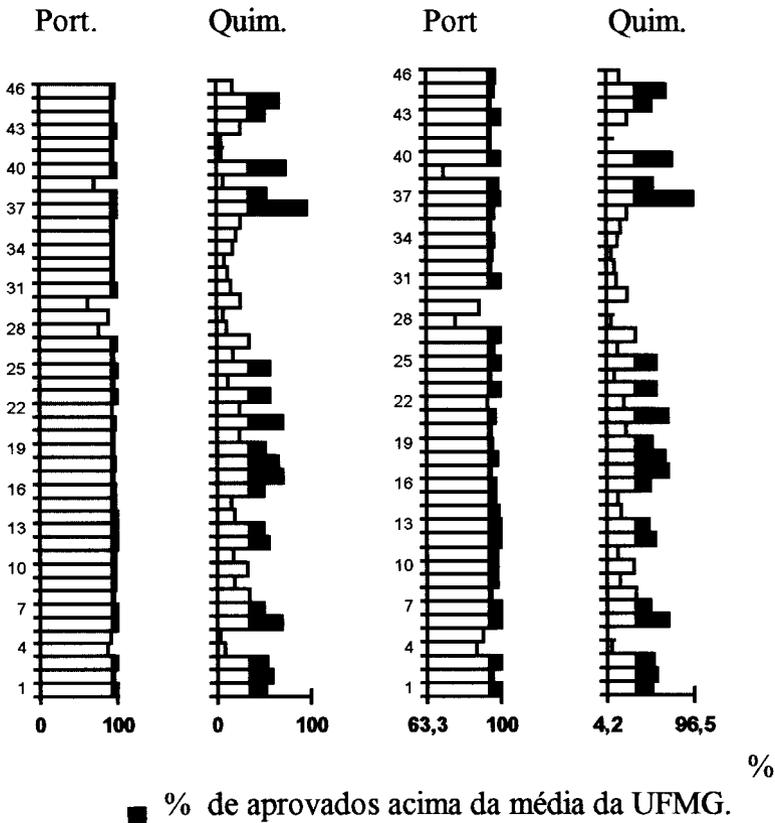
- as colunas ou barras, de uma mesma série, são desenhadas justapostas nos eixos correspondentes, mas sem intervalo, para facilitar as comparações no interior de cada representação;
- as colunas ou barras são preenchidas por um elemento de uma variável visual, ou seja, de um meio gráfico de representação, para destacá-las do fundo da imagem;
- se for o caso, as médias, relativas a cada série de dados, são indicadas, usando-se dois elementos de uma variável visual ordenada, para se diferenciar as porções das colunas ou barras que se situam, respectivamente, acima ou abaixo da média;
- os diagramas são apresentados, ou lado a lado, no caso de se construir gráficos de barras, ou um abaixo do outro, quando se constrói gráficos de colunas, mas sempre separados por um pequeno intervalo. A extensão desse intervalo deve ser suficiente para permitir a visualização de cada um dos diagramas, em separado, sem, no entanto, causar transtornos visuais para sua comparação;
- os diagramas são justapostos, inicialmente, na mesma ordem em que estão apresentados, na matriz de dados, os elementos do componente que representam;

FIGURA 5

ESCALAS POSSÍVEIS DAS REPRESENTAÇÕES DE BASE DA MATRIZ ORDENÁVEL QUANTITATIVA

ESCALA COMUM

ESCALA PRÓPRIA



-os conceitos que identificam os gráficos e suas colunas ou linhas são anotados de modo esquemático, mas de forma bem legível, para evitar confusões na fase de manipulação da matriz de tratamento;

-os conceitos que especificam os sujeitos são anotados apenas para a primeira série de características.

Todas essas regras e princípios foram observados na construção da matriz de tratamento, apresentada na Figura 6, que transcreve os dados apresentados na Matriz 1. Essa matriz está pronta para ser manipulada, iniciando-se uma nova etapa de tratamento.

A etapa de manipulação da matriz consiste na permutação de linhas e de colunas, procurando-se aproximar aquelas que são semelhantes. Essa semelhança é verificada a partir da observação do padrão formado pelos elementos das variáveis visuais utilizadas na construção dos diagramas. Na matriz de tratamento, mostrada na Figura 6, em função dos elementos da variável ordenada usada, patenteiam-se linhas mais claras, linhas mais escuras, linhas que são mais escuras ou claras, considerando-se os quadrantes em que pode ser dividida a matriz, e ainda, evidenciam-se colunas mais claras e outras mais escuras.

A partir dessas observações, a matriz de tratamento será reconstruída, alterando-se a posição das linhas e das colunas. Nota-se que será necessário reconstruir a matriz tantas vezes quantas forem necessárias para se obter agrupamentos de linhas. Nesses agrupamentos as semelhanças entre as linhas devem ser sempre maiores do que as diferenças percebidas.

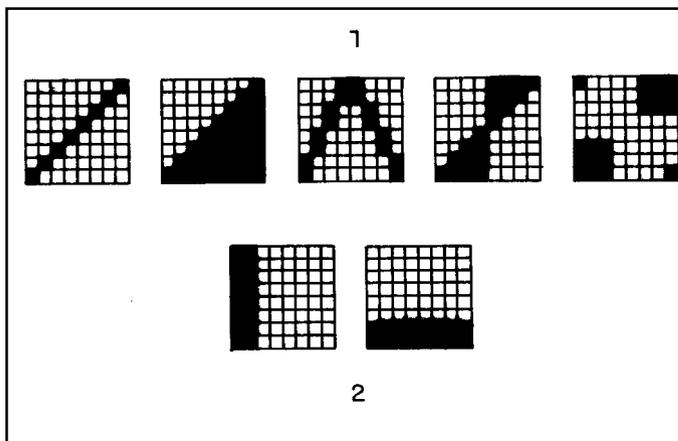
Nessa tarefa de reconstruções sucessivas, em busca de uma imagem ideal da matriz, o pesquisador deve lembrar que toda operação de classificação, a partir de coordenadas ortogonais “tend vers la réalisation de la diagonale du champ ou de formes simples que en sont déduites (ou) à une forme parallèle à l’une des coordonnées” do plano, tal como se mostra, respectivamente, em 1 e 2, na Figura 7, segundo BERTIN (1973, p. 168).

Assim, o pesquisador deve procurar tornar mais nítida, ou melhor configurada, a forma ou formas que estão sendo evidenciadas na imagem da matriz, através das permutações sucessivas. A imagem ideal da matriz de tratamento será, então, aquela que apresentar, com maior nitidez, as formas que as associações dos dados estabelecem.

Essas formas irão, finalmente, indicar se existe correlação entre os elementos expressos nos eixos ortogonais e, em função de sua existência, irão permitir a identificação de agrupamentos. Na Figura 7, por exemplo, os esquemas, apresentados em 1, mostram correlações mais e menos perfeitas, em 2, correlações inexistentes. Nos esquemas apresentados em 1, é possível a identificação de agrupamentos, em 2, ao contrário, a não variação de um dos componentes, representados em x ou y, não permite que agrupamentos sejam identificados. Nesse caso, indicam a não pertinência do problema colocado inicialmente pelo pesquisador (Bertin, 1973, p. 168).

A matriz apresentada na Figura 8 mostra a imagem ideal dos dados que são estudados neste trabalho, após sucessivas manipulações das linhas e colunas da matriz mostrada na Figura 6.

Fig. 7 - Exemplos de Resultados da Manipulação da Matriz de Tratamento



Fonte: Bertin, J. *Sémiologie Graphique Les Diagrammes, Les Réseaux, Les Cartes*. Paris: La Haye, Mouton; Paris: Gauthier Villars, 1973 p.168. des-A.Rosalem

A partir da imagem da matriz manipulada, percebe-se uma linha diagonal, separando o plano da matriz em dois triângulos retângulos - um mais escuro, outro mais claro. De um outro ponto de vista, podem ser visualizados vários retângulos que se apresentam cada vez mais claros, desde o início da matriz.

Essa linha e as formas que se patenteiam indicam que o problema colocado inicialmente pelo pesquisador, expresso por objetivos específicos, é pertinente. Ou seja, através da análise da imagem manipulada, podem ser identificados grupos de cursos com desempenho variado e provas em que grupos de cursos se destacam, considerando a média verificada para a UFMG.

A análise, sempre buscando os padrões claro e escuro, formados pelas variáveis visuais, aponta a existência de seis grupos de cursos, correspondentes aos retângulos que podem ser visualizados de cima para baixo, na matriz. Esses grupos se individualizam pela tonalidade que, por sua vez, é um indicador de desempenho.

De fato, analisando em detalhe a matriz, observa-se que o retângulo mais escuro, formado pelas primeiras oito linhas, corresponde ao grupo de cursos em que seus aprovados, com metade ou mais pontos nas provas da primeira etapa do vestibular, se situaram sempre acima da média da UFMG. Nota-se, também, que o

retângulo mais claro, configurado pelas cinco últimas linhas da matriz, compreende os cursos em que os aprovados se situaram abaixo da média da UFMG, em todas as provas.

Examinando, ainda, os retângulos que podem ser visualizados de cima para baixo na matriz, as tonalidades no seu interior patenteiam as provas em que os aprovados se sobressaíram.

Assim é que podem ser identificados grupos de cursos em que os aprovados se destacaram em todas as provas, grupos em que o desempenho é melhor nas provas da área de ciências humanas, grupos em que o desempenho mais destacado pode ser relacionado às provas da área de ciências físicas, etc.

Os grupos de cursos com desempenho opostos, citados anteriormente, assim como, os que se caracterizam por apresentar desempenho intermediário, são mostrados na Figura 9. E na Figura 10, apresenta-se as provas de maior destaque, por grupo de curso.

As Figura 9 e 10 poderiam registrar outros resultados da análise, mas, nelas, são mostrados apenas os resultados considerados importantes pelo pesquisador. Elas representam um importante documento de comunicação, onde são registradas as informações úteis, de modo simplificado.

Essas figuras denominadas, também, matrizes de interpretação são, então, tão importantes e úteis ao pesquisador, quanto as matrizes que reúnem os dados de um problema e as matrizes de tratamento, apresentadas anteriormente.

O ALCANCE DE UM TRATAMENTO GRÁFICO

Os princípios e as regras que norteiam a modalidade gráfica de análise de dados, ainda que tenham sido expostos suscintamente, propiciam elementos para que sua propriedade e eficácia sejam afirmadas.

Num contexto científico, a propriedade e a eficácia de qualquer linguagem de tratamento podem ser relacionadas com sua capacidade de consecução dos objetivos que são inerentes aos procedimentos de processamento e análise de dados. Através da modalidade gráfica de tratamento é possível, conforme foi exposto anteriormente, organizar e reduzir pertinentemente dados, tendo em vista o patenteamento de informações úteis ao pesquisador.

Nesse contexto, a propriedade e eficácia de uma linguagem de tratamento podem ser consideradas, também, em relação à sua objetividade. Essa característica, por sua vez, pode ser atribuída à modalidade gráfica de tratamento, na medida em que apresenta regras bem definidas, a propósito de todo o processo analítico e de todos os procedimentos que o compõe.

FIGURA 8
MATRIZ DE TRATAMENTO MANIPULADA

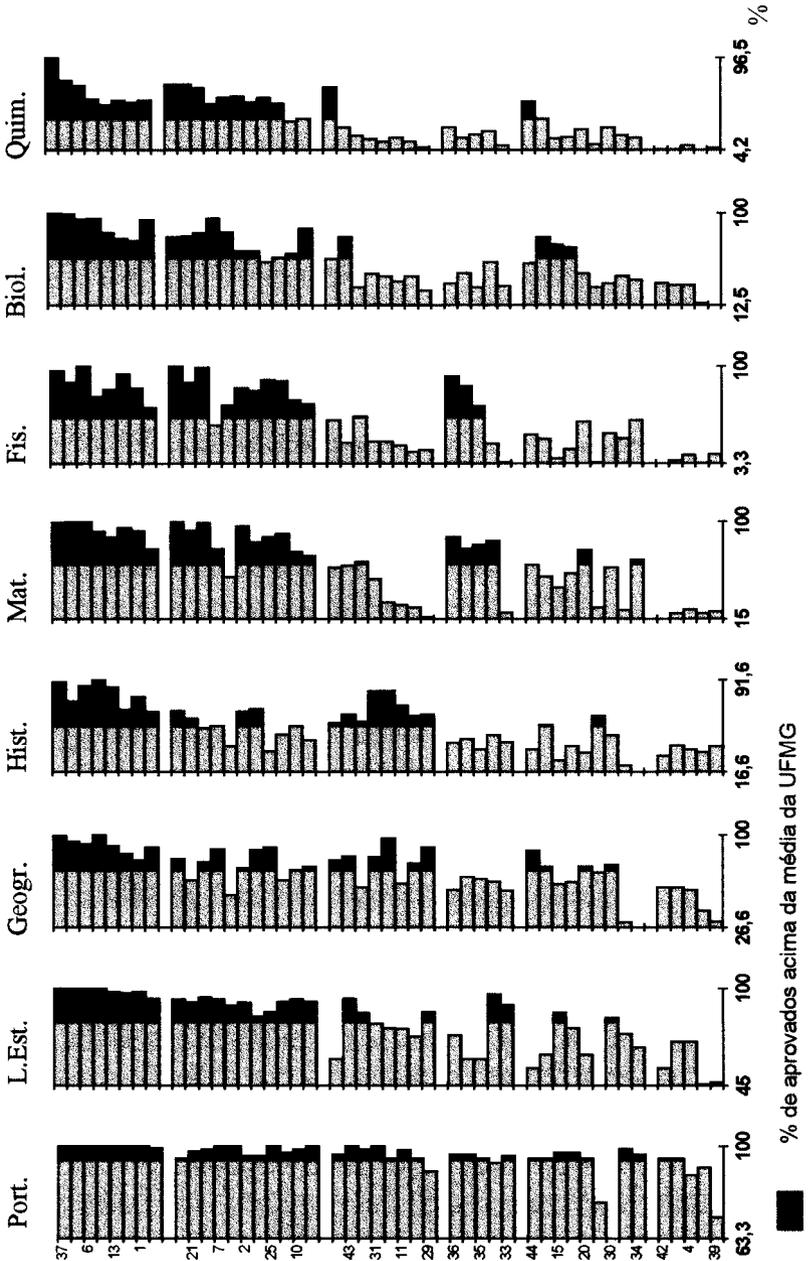


FIGURA 10 - MATRIZ DE COMUNICAÇÃO DE INFORMAÇÕES ÚTEIS (Provas de melhor desempenho por grupo de cursos)

Nº	CURSOS	PROVAS							
		P.	M.	G.	H.	L.E.	F.	Q.	B.
37	Medicina	■	■	■	■	■	■	■	■
40	Odontologia	■	■	■	■	■	■	■	■
6	C. da Computação	■	■	■	■	■	■	■	■
12	Comunicação Social	■	■	■	■	■	■	■	■
13	Direito	■	■	■	■	■	■	■	■
3	Arquitetura	■	■	■	■	■	■	■	■
1	Administração D.	■	■	■	■	■	■	■	■
38	Med. Veterinária	■	■	■	■	■	■	■	■
17	Eng. Elétrica	■	■	■	■	■	■	■	■
21	Eng. Química	■	■	■	■	■	■	■	■
18	Eng. Mecânica	■	■	■	■	■	■	■	■
7	C.Biológ. D.(L/B)	■	■	■	■	■	■	■	■
23	Farmácia	■	■	■	■	■	■	■	■
2	Administração N.	■	■	■	■	■	■	■	■
19	Eng. Metalúrgica	■	■	■	■	■	■	■	■
25	Física D. (B)	■	■	■	■	■	■	■	■
16	Eng. Civil	■	■	■	■	■	■	■	■
10	C. Econômicas	■	■	■	■	■	■	■	■
27	Fisioterapia	■	■	■	■	■	■	■	■
45	Química N.(L)	■	■	■	■	■	■	■	■
43	Psicologia (L/B)	■	■	■	■	■	■	■	■
9	C. Contábeis	■	■	■	■	■	■	■	■
31	História D. (L/B)	■	■	■	■	■	■	■	■
32	História N. (L/B)	■	■	■	■	■	■	■	■
11	C. Sociais (L/B)	■	■	■	■	■	■	■	■
24	Filosofia (L/B)	■	■	■	■	■	■	■	■
29	Geografia N.(L)	■	■	■	■	■	■	■	■
36	Matemática N.(L)	■	■	■	■	■	■	■	■
26	Física N. (L)	■	■	■	■	■	■	■	■
35	Matemática D.(B)	■	■	■	■	■	■	■	■
22	Estatística	■	■	■	■	■	■	■	■
33	Letras (L)	■	■	■	■	■	■	■	■
44	Química D.(L/B)								
8	C. Biológ. N.(L)								
15	Enfermagem (L/B)								
46	Terap. Ocupacional								
20	Eng. de Minas								
28	Geografia D.(L/B)								
30	Geologia								
14	Ed. Física (L/B)								
34	Matemática D.(L)								
42	Pedagogia N.(L/B)								
41	Pedagogia D.(L)								
4	Belas Artes (L/B)								
5	Biblioteconomia								
39	Música (L/B)								

 Provas de melhor desempenho

Se a modalidade de tratamento gráfica reúne características importantes do domínio científico, não deixa, entretanto, assim como suas congêneres, de apresentar limites analíticos, colocar obstáculos e proporcionar facilidades aos seus usuários, tendo em vista sua formação, seus problemas de pesquisa e/ou os recursos para solucioná-los.

Nesse sentido, sua escolha não dispensa uma reflexão. Não obstante, as técnicas gráficas podem ter sua aplicação muito ampliada na construção do conhecimento geográfico, dadas as afinidades da Geografia e dos geógrafos com as questões gráficas.

NOTAS

- (1) O software Amado, comercializado pelo Centre International de Statistique et Informatique Appliquée - CISIA, Saint Mandé, France, é um exemplo de produto computacional, idealizado para permitir a automação de técnicas de tratamento gráfico, compreendidas pela Semiologia.
- (2) Cita-se, como exemplo, os trabalhos de dissertação de Carmo (1994) e de Moura (1993), orientados por Le Sann, nos quais foram utilizadas técnicas de tratamento gráfico.
- (3) A Matriz 1 e todas as demais figuras apresentadas neste trabalho foram organizadas, utilizando-se recursos do Microsoft Word 6.0, com exceção das Figuras 2 e 7, que foram desenhadas à mão.
- (4) O conceito nível de organização, elaborado pela Semiologia Gráfica, guarda uma certa relação com o de escala de mensuração, usado pela Estatística. Assim, os dados obtidos na escala de mensuração nominal podem ser classificados como seletivos ou associativos. Os dados obtidos na escala ordinal podem ser definidos como ordenados e os levantados a partir das escalas de mensuração intervalar e razão podem ser classificados como quantitativos.

BIBLIOGRAFIA

BERTIN, J. *La Graphique et le Traitement Graphique de l'Information*. Paris: Flammarion, 1977.

_____ *Sémiologie Graphique. Les Diagrammes, Les Réseaux, Les Cartes*. Paris-La Haye: Mouton; Paris: Gauthier-Villars, 1973.

_____ "Un Outil Cartographique: la Synthèse Trichromatique." *Bulletin du Comité Français de Cartographie*, Paris, vol. 94, nº 4, (1982), 18-21.

BONIN, S. *Initiation à la Graphique*. Paris: Epi, 1975.

_____ "Méthodes d'Approche pour la Réalisation des Cartes de Superposition." *Bulletin du Comité Français de Cartographie*, Paris, vol. 94, nº 4, (1982), 12-17.

CARMO, V. A. "Proposta Gráfica para um Mapa Politemático: O Caso do Mapa Diagnóstico Geo-Ambiental da Usina Hidrelétrica de Igarapava". Dissertação de Mestrado, Departamento de Geografia, Instituto de Geociências, Universidade Federal de Minas Gerais, Belo Horizonte, 1994.

MOURA, A. C. M. "Estudo Metodológico da Aplicação da Cartografia Temática às Análises Urbanas". Dissertação de Mestrado, Departamento de Geografia, Instituto de Geociências, Universidade Federal de Minas Gerais, Belo Horizonte, 1993.

SANCHEZ, M. C. "A Cartografia como Técnica Auxiliar da Geografia." *Boletim de Geografia Teórica*, Rio Claro, vol.3, nº6, (1973), 31-46.

SANTOS, M. M. D. "Desempenho dos Candidatos e dos Aprovados para os Cursos de Geografia, Diurno e Noturno, do Instituto de Geociências da Universidade Federal de Minas Gerais, no Vestibular de 1994." Belo Horizonte: UFMG, IGC, Departamento de Geografia, 1994. (Relatório de Pesquisa)

_____ "A Representação Gráfica da Informação Geográfica." *Geografia*, Rio Claro, vol.12, nº.23,(abril 1987), 1-13.