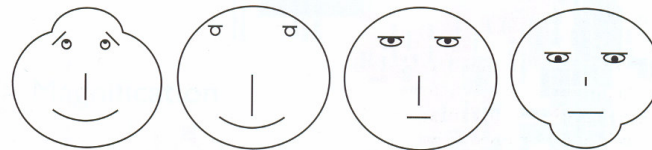




Universidade de Aveiro  
Departamento de Electrónica,  
Telecomunicações e Informática

# Características dos Dados



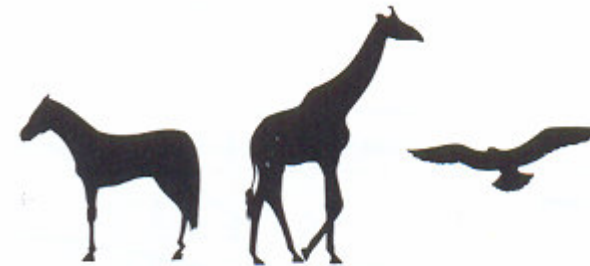
- Os dados a visualizar podem ter diversas formas e têm-se desenvolvido muitas técnicas e sistemas para os visualizar
- É vantajoso estabelecer uma **classificação dos dados** por várias razões:
  - para poder prever que técnicas de Visualização serão adequadas
  - para facilitar a comunicação sobre os dados
  - para permitir progresso na visualização de dados multidimensionais
  - para permitir uma abordagem sistemática à Visualização
  - ....

# Uma classificação geral dos dados

- **Natureza dos dados:**
  - Contínuos
  - Discretos
- **Nível de representação:**
  - Qualitativos (ou categóricos)
  - Quantitativos (ou numéricos)
- **Escala de medida:**
  - Nominal,
  - Ordinal
  - De intervalo
  - De razão

4.1 27 102 3.14  
-0.1 16

Numerical data



Categorical data

Monday Wednesday  
Tuesday Thursday

Ordinal data

*(Spence, 2001)*

- exemplo de escalas:
  - **nominal** --> marcas de automóveis, sexo, tipos de animais...
  - **ordinal** --> dias da semana, preferências (1º, 2º...)
  - **de intervalo** --> data, QI, temperaturas em °C
  - **de razão** --> temperaturas em °K, peso, altura
- a escala de razão representa o **nível mais elevado da representação**, tem um zero não arbitrário (ao contrário da escala de intervalo)
- Esta classificação é geral; serve também para escolha de técnicas estatísticas a aplicar aos dados

# Modelo, estrutura e formato dos dados a Visualizar

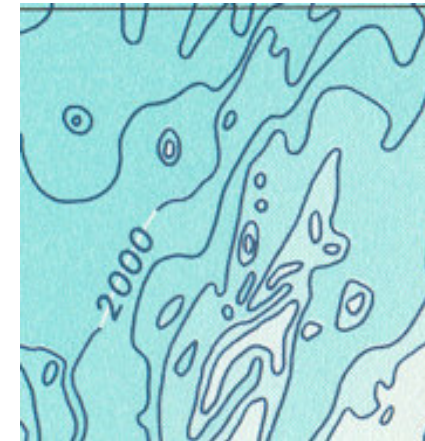
- Considere-se um conjunto de dados com três colunas:

*latitude*      *longitude*      *d*

- Qual é a forma mais apropriada de visualizar estes dados?

- Se *d* for profundidade, é provável que a técnica de visualização escolhida envolva interpolação (ex: contornos)

- Se os dados representarem localização e nº de mortes em acidentes de viação, usar contornos não faz sentido



**Conhecer a estrutura dos dados não é suficiente!  
É necessário conhecer o fenómeno**

!

- Cada conjunto de dados tem um campo subjacente (*underlying field*)
- São diferentes nos dois casos do exemplo anterior
- A tentativa de obter um *modelo empírico* faz parte do processo de Visualização
- Deve ser guiada por um cientista da área de aplicação
- Os dados passam por várias formas:
  - *underlying field*
  - *visualization input data*
  - *empirical model*
  - *abstract visualization object*

# Funções de muitas variáveis

- Caracterize-se o campo subjacente (*underlying field*) dos dados a visualizar
- Podemos usar uma função  $F(X)$  em que

$X$  pode representar muitas variáveis independentes  $(x_1, x_2, \dots, x_n)$

$F$  pode representar muitas funções  $(f_1, f_2, \dots, f_n)$

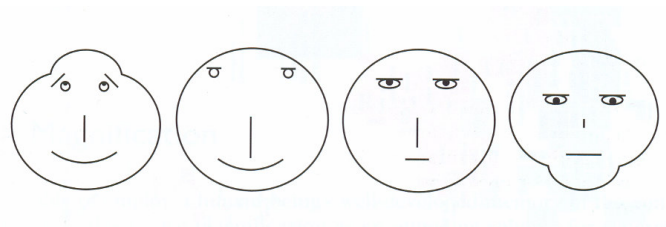
- Assim, para caracterizar um conjunto de dados é necessário identificar e caracterizar as variáveis dependentes e independentes

## Exemplos:

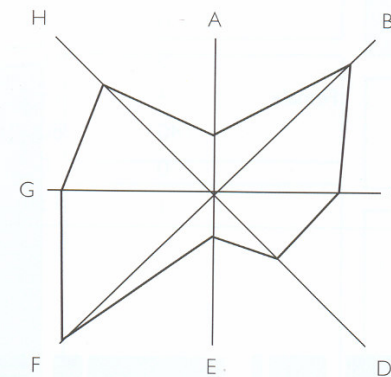
- muito simples:
  - $X$  - ponto 2D
  - $F$  - temperatura
- campo de fluxo 3D variável no tempo, numa reação química:
  - $X$  - ponto 3D; tempo
  - $F$  - velocidade do fluxo; temperatura; concentração de cada reagente
- campo magnético:
  - $X$  - ponto 3D
  - $F$  - vector campo magnético; identificador do material
- profundidade do mar:
  - $X$  - ponto 2D
  - $F$  - profundidade
- acidentes de viação - nº de mortes e localização de acidentes:
  - $X$  - identificador
  - $F$  - latitude; longitude; mortes

- Note-se que existe uma diferença entre os casos
  - Muitas variáveis dependentes
  - Muitas variáveis independentes
- Por exemplo se tivermos ao todo seis variáveis, é diferente termos:
  - 4 independentes e 2 dependentes:  
temperatura e pressão (dependentes)  
em 3D + tempo (independentes)
  - 2 independentes e 4 dependentes:  
temperatura, pressão, humidade e velocidade do vento  
em 2D

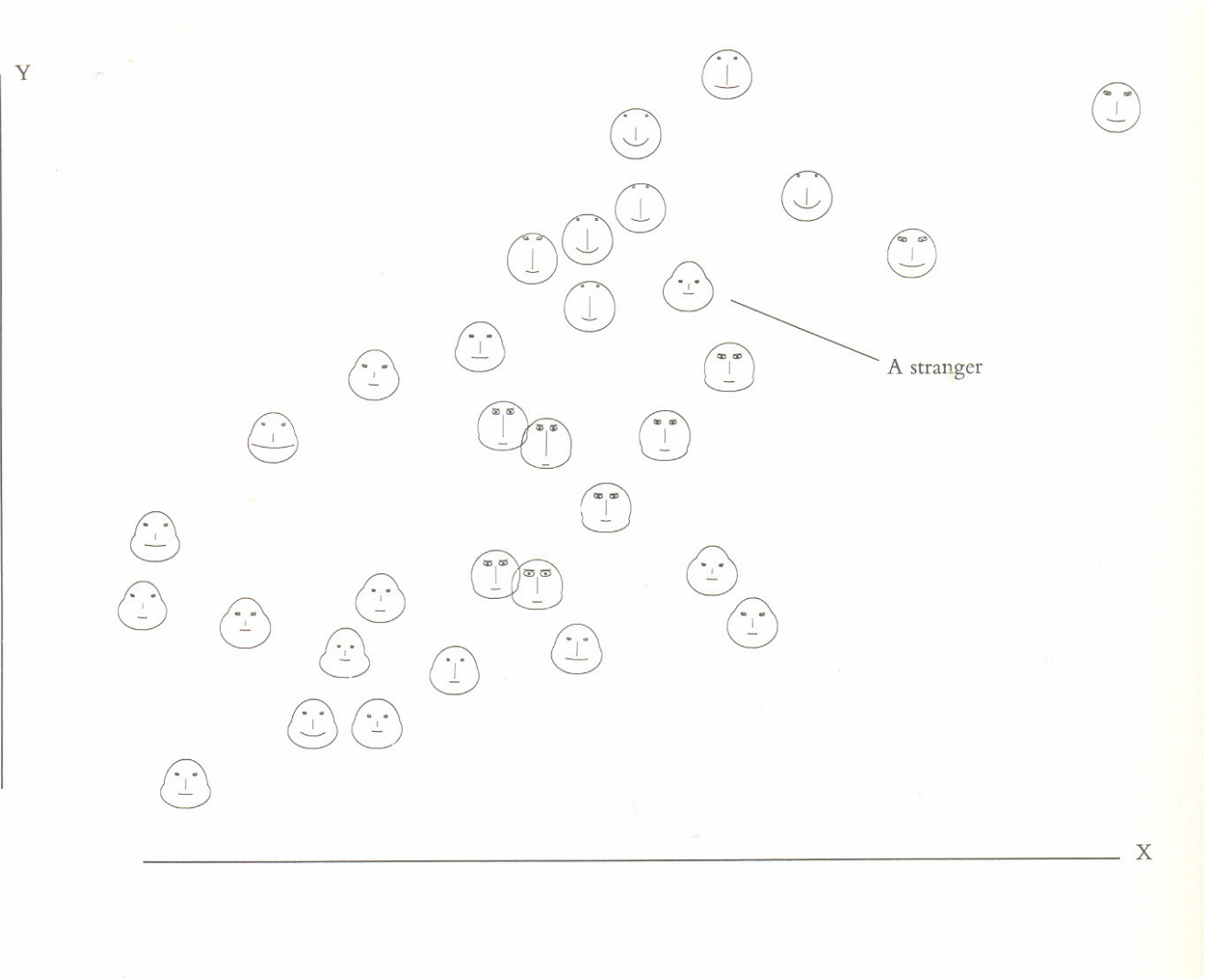
- As caras de Chernoff e os polígonos estrelados são exemplos de **glifos** (ícones) que permitem a visualização de dados multidimensionais (muitas variáveis dependentes)



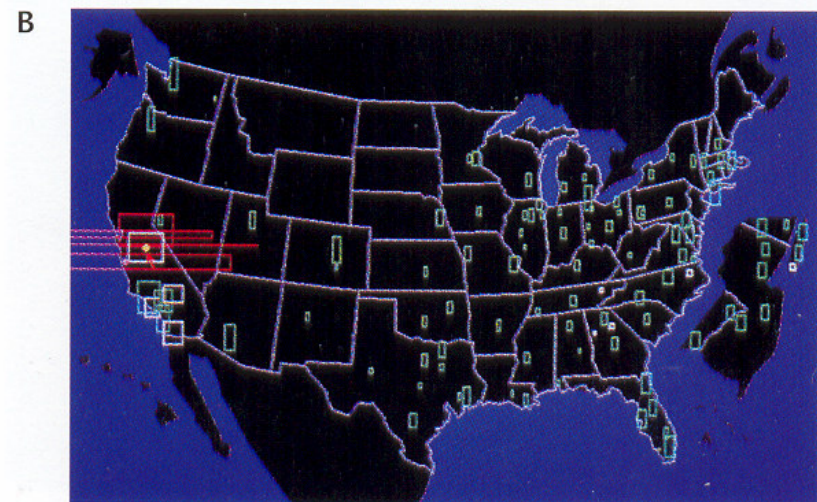
*(Spence, 2001)*



- Esta técnica explora a grande capacidade que nós temos de distinguir faces
- Chernoff conseguiu representar 18 características diferentes através das caras

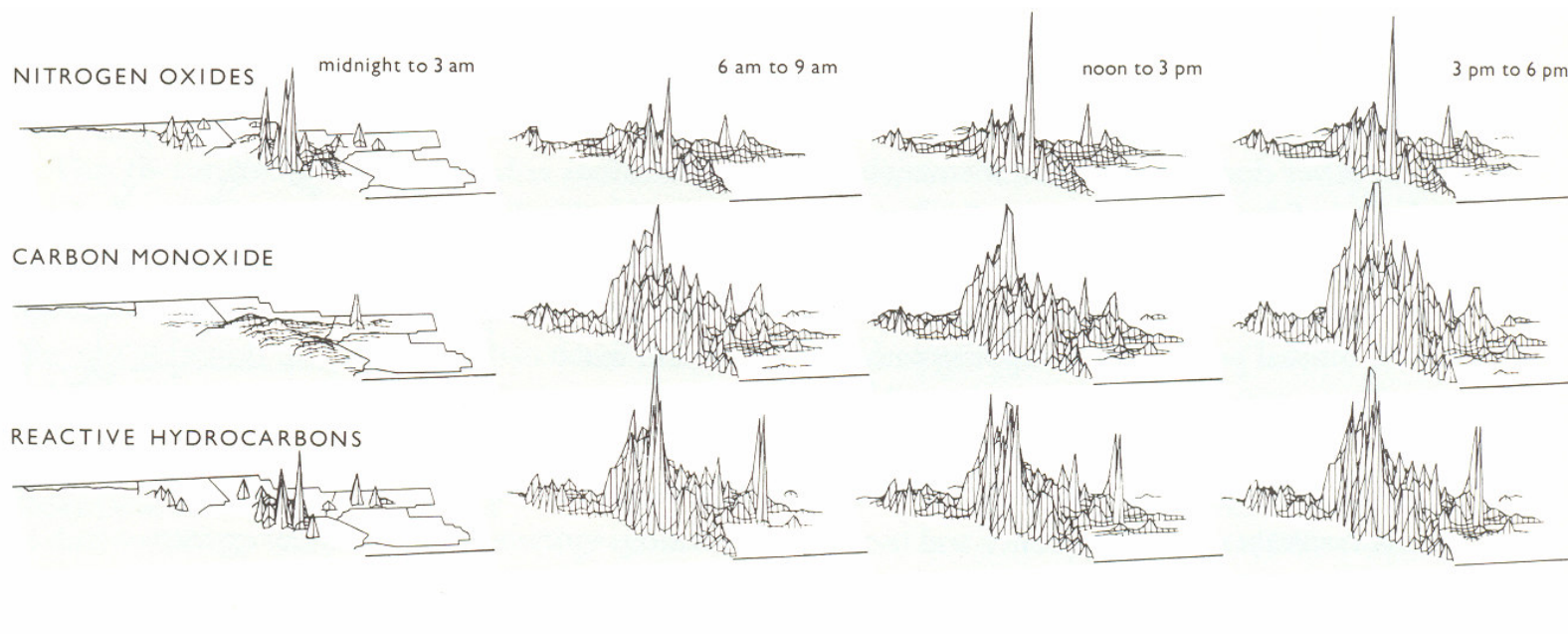


- Quatro variáveis independentes (cidade, lat., long., momento)
- Duas variáveis dependentes (tráfego originado e recebido)
- Tráfego telefónico:
  - A- normal
  - B- depois de um sismo
- Altura do rectângulo – nº de chamadas feitas
- Largura do rectângulo- nº de chamadas recebidas
- Cor – redundante (vermelho-recebeu + chamadas)



*(Keller e Keller, 1993)*

- Três variáveis independentes (latitude, longitude, tempo)
- Três variáveis dependentes (poluentes)



*small multiple*

*(Tuft, 1990)*

- São ainda importantes os seguintes aspectos dos dados:
  - Origem
  - Formato
  - Compressão
  - Transformações
  - ...

# Dados Discretos e interpolação

- Os dados a visualizar são discretos, mesmo quando o fenómeno é contínuo
- Mas pretende-se visualizar nos “intervalos”
- Este problema trata-se através da interpolação,
- Podem usar-se vários tipos de funções:
  - Lineares
  - Quadráticas
  - Cúbicas
  - *Splines*
  - .....

# Estrutura

- Os dados correspondentes a fenómenos com uma estrutura espacial inerente podem ser :
  - Estruturados (i.e., ter uma **estrutura regular**)
  - Não estruturados (i.e., ter uma **estrutura irregular**)
- Os dados regulares têm uma relação inerente entre os vários pontos
- Por exemplo, não é necessário armazenar as coordenadas de todos os pontos (bastam a posição inicial, o o espaçamento e o n<sup>o</sup> de pontos)
- Os dados irregulares permitem representar informação mais densamente onde varia mais

# Dimensão

- Os dados podem ser :
  - 1D (linha)
  - 2D (superfície)
  - 3D (volume)
  - ....
  - N D
- A dimensão é importante na escolha das técnicas de visualização

## Principal Bibliografia

- Brodlie, K., et. Al., *Scientific Visualization, Techniques and Applications*, Springer-Verlag, 1992
- Gallop, J., “Underlying data models and structures for Visualization”, in: Rosenblum, L., R. Earnshaw, J. Encarnação, H. Hagen, A. Kaufman, S. Klimenko, G. Nielson, F. Post, D. Thalmann (eds.), *Scientific Visualization, Advances and Challenges*, IEEE Computer Society Press, Academic Press, 1994
- Spence, R., *Information Visualization*, ACM Press Books, Addison Wesley, 2001
- Tufte, E., *Envisioning Information*, Graphics Press, 1990