

**FCT/Unesp – Presidente Prudente**  
**Departamento de Matemática e Computação**

# Técnicas de Visualização para Dados Multivariados

## Parte 1

Prof. Danilo Medeiros Eler  
danilo.eler@unesp.br

# Sumário

## ■ Parte 1

- ❑ Técnicas Orientadas a Pixel
- ❑ Técnicas Iconográficas

## ■ Parte 2

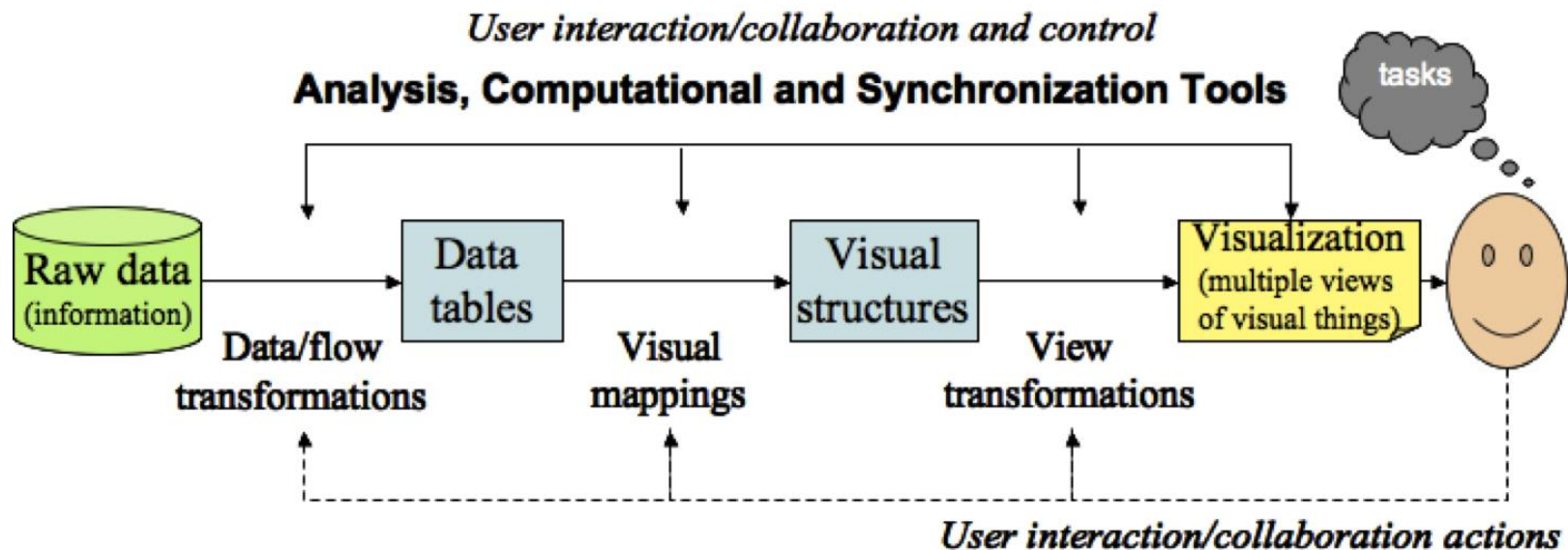
- ❑ Técnicas Baseadas em Grafos
- ❑ Técnicas Hierárquicas

## ■ Parte 3

- ❑ Técnicas de Projeção Geométrica

# Processo de Visualização

- Primeiramente, vamos retomar o pipeline de visualização utilizado pela maioria dos sistemas



# Taxonomia

- Foram propostas seis classes para categorizar as técnicas de visualização
  - Displays 2D/3D padrão
    - Gráficos de barra, gráficos de linhas
  - Displays Geometricamente Transformados
    - Técnicas de Projeção Geométrica
  - Displays Iconográficos
    - Técnicas Iconográficas

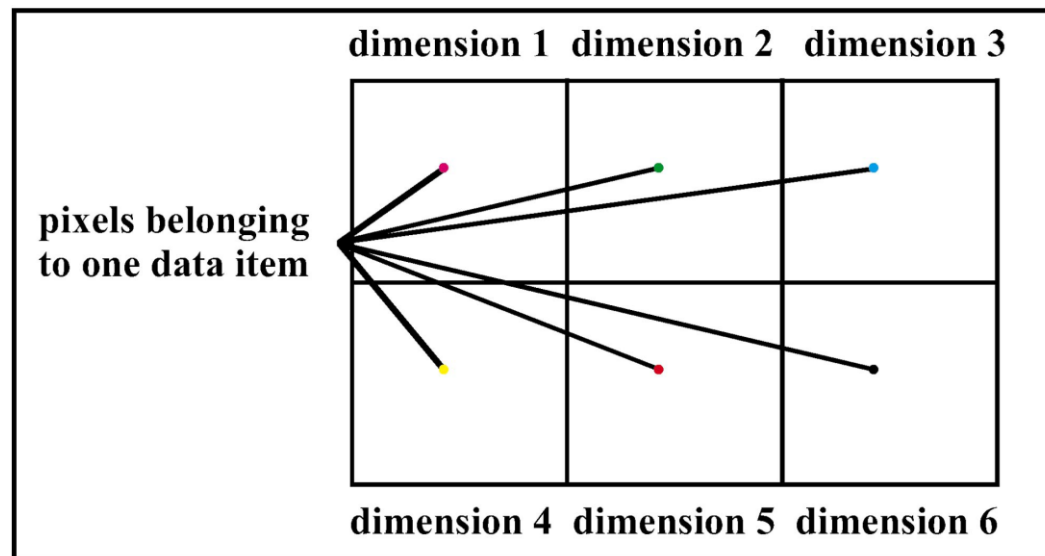
# Taxonomia

- Foram propostas seis classes para categorizar as técnicas de visualização
  - Display Denso de Pixels
    - Técnicas orientadas a pixel
  - Displays Empilhados
    - Técnicas Hierárquicas
  - Técnicas Baseadas em Grafos
    - Não estava na taxonomia

# Técnicas Orientadas a Pixel

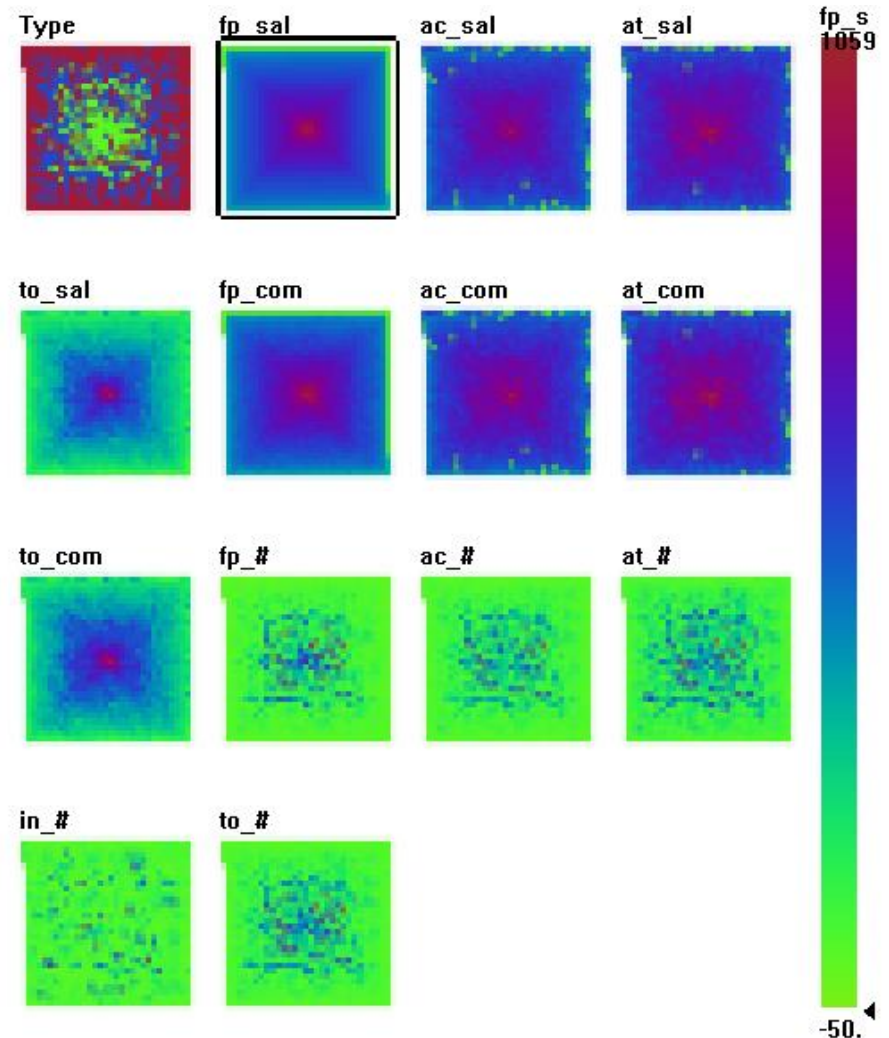
# Técnicas Orientadas a Pixel

- A ideia básica das técnicas orientadas a pixel é mapear cada valor de dado a um pixel colorido e apresentar os valores de dados pertencente a uma dimensão (atributo) em uma sub-janela separada



# Técnicas Orientadas a Pixel

- Uma das vantagens dessa técnica é poder ordenar por uma dimensão para que algum padrão seja revelado
  - Ordenado pelo salário do Professor Titular





# Técnicas Orientadas a Pixel

## ■ Ordenado pela quantidade de Professores Titulares

Type: (I, IIA, or IIB)

fp\_sal: Average salary - full professors

ac\_sal: Average salary - associate professors

at\_sal: Average salary - assistant professors

to\_sal: Average salary - all ranks

fp\_com: Average compensation - full professors

ac\_com: Average compensation - associate professors

at\_com: Average compensation - assistant professors

to\_com: Average compensation - all ranks

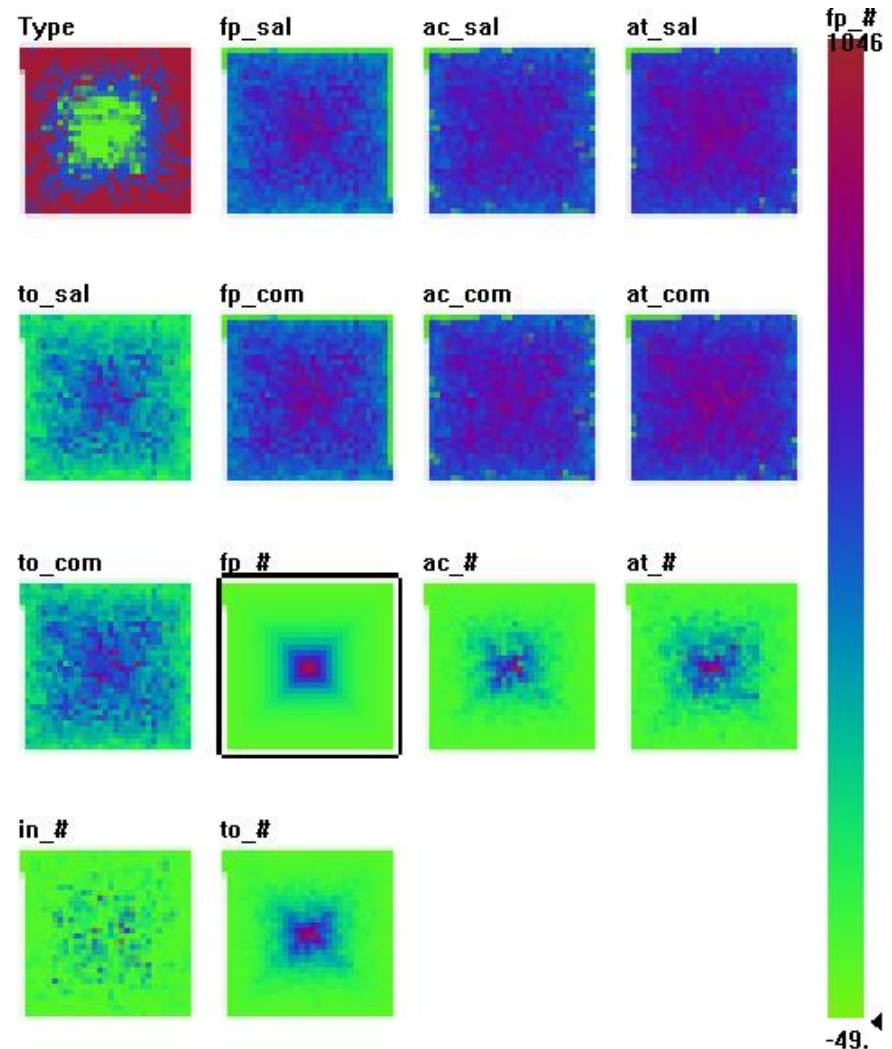
fp\_#: Number of full professors

ac\_#: Number of associate professors

at\_#: Number of assistant professors

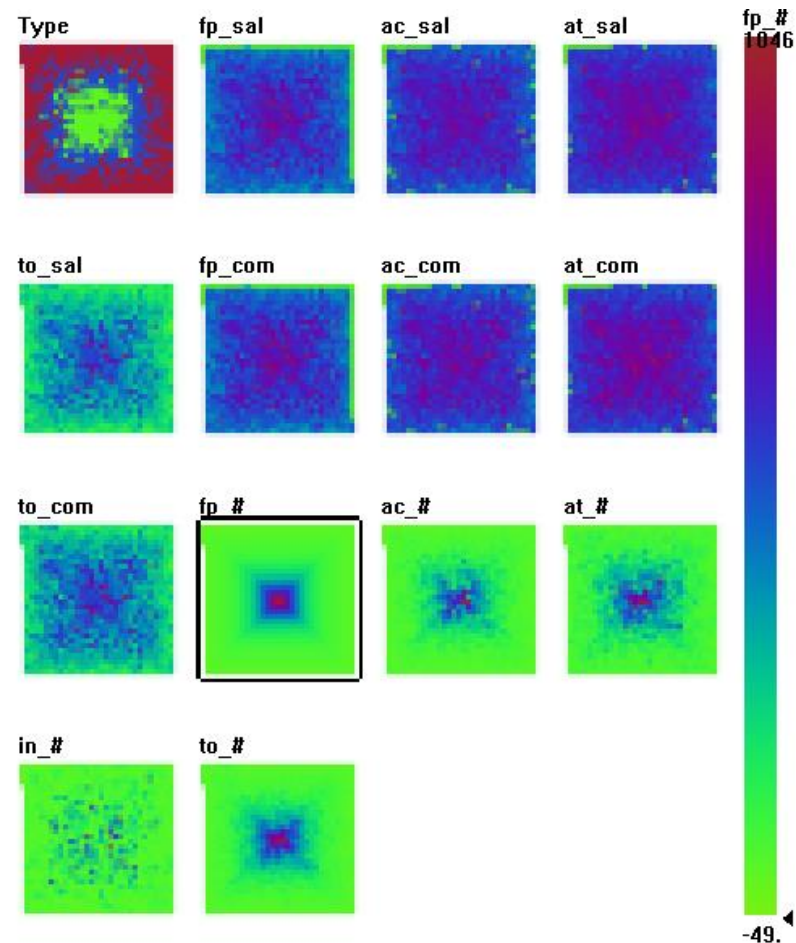
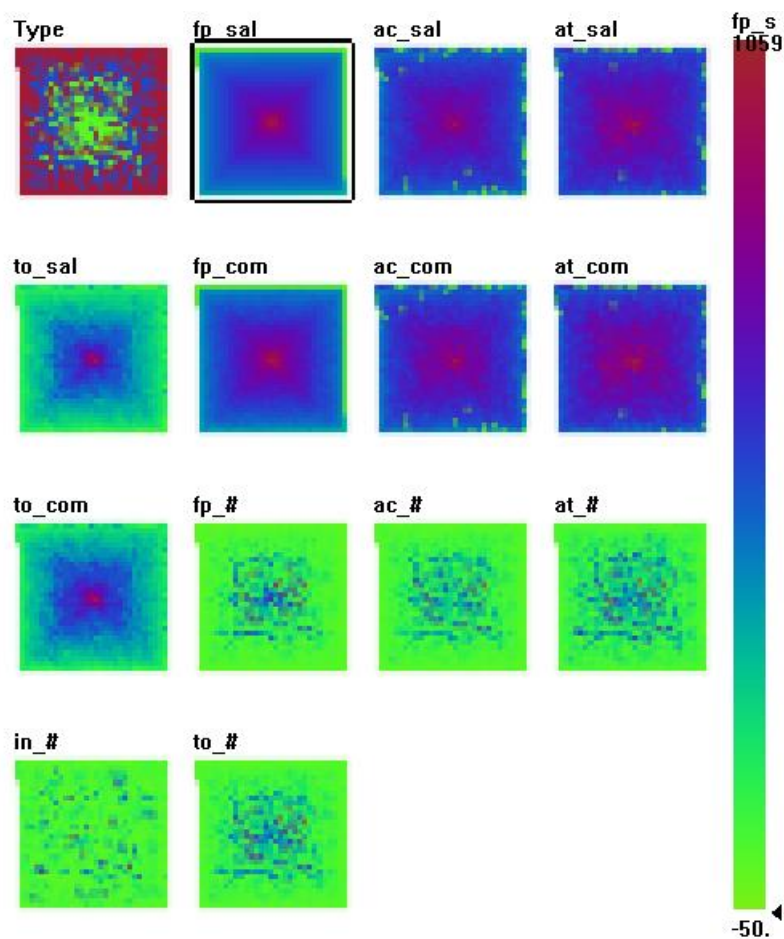
in\_#: Number of instructors

to\_#: Number of faculty - all ranks



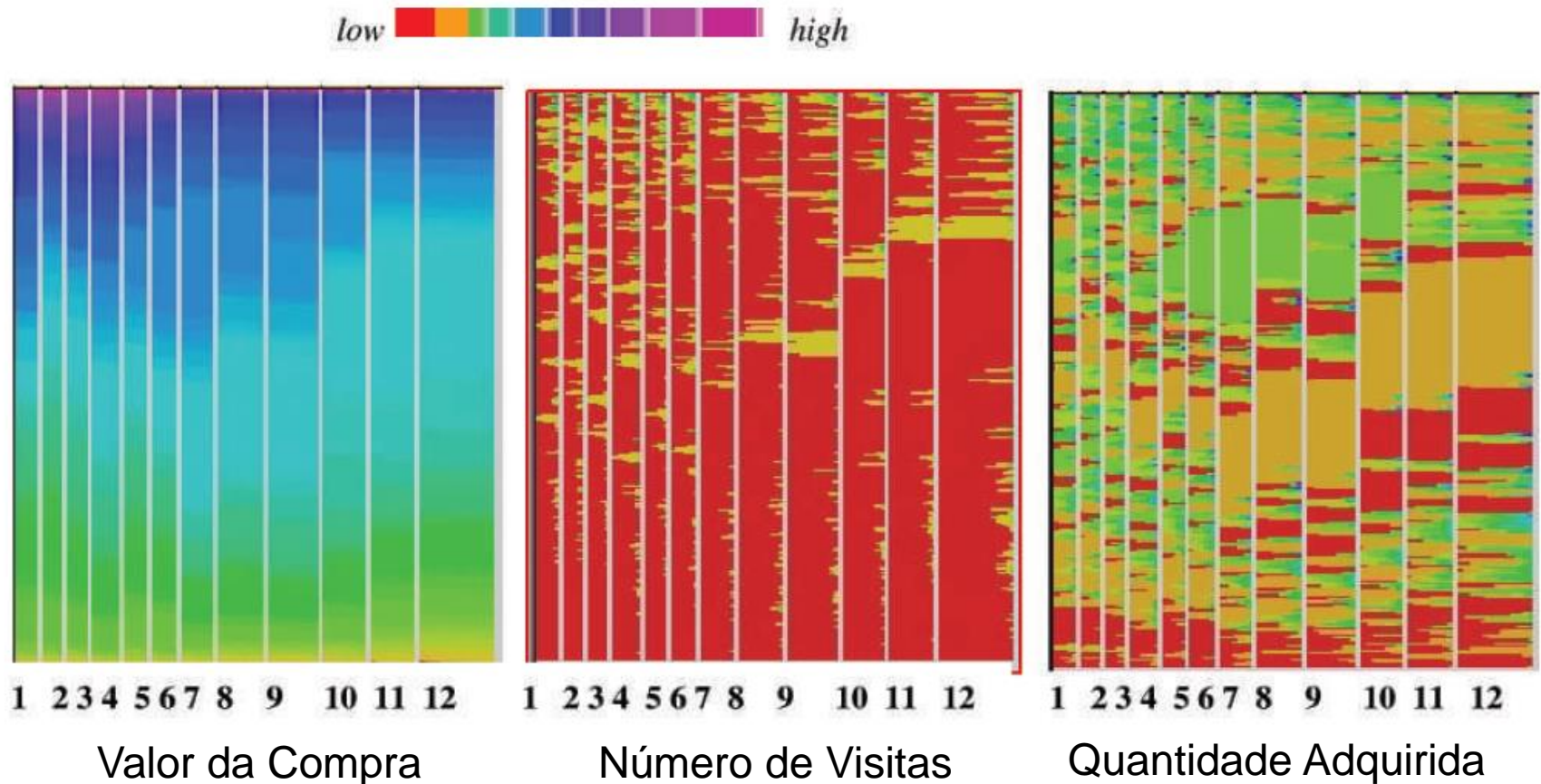
# Comparação entre as ordenações

- salário x quantidade prof. titulares



# Técnicas Orientadas a Pixel

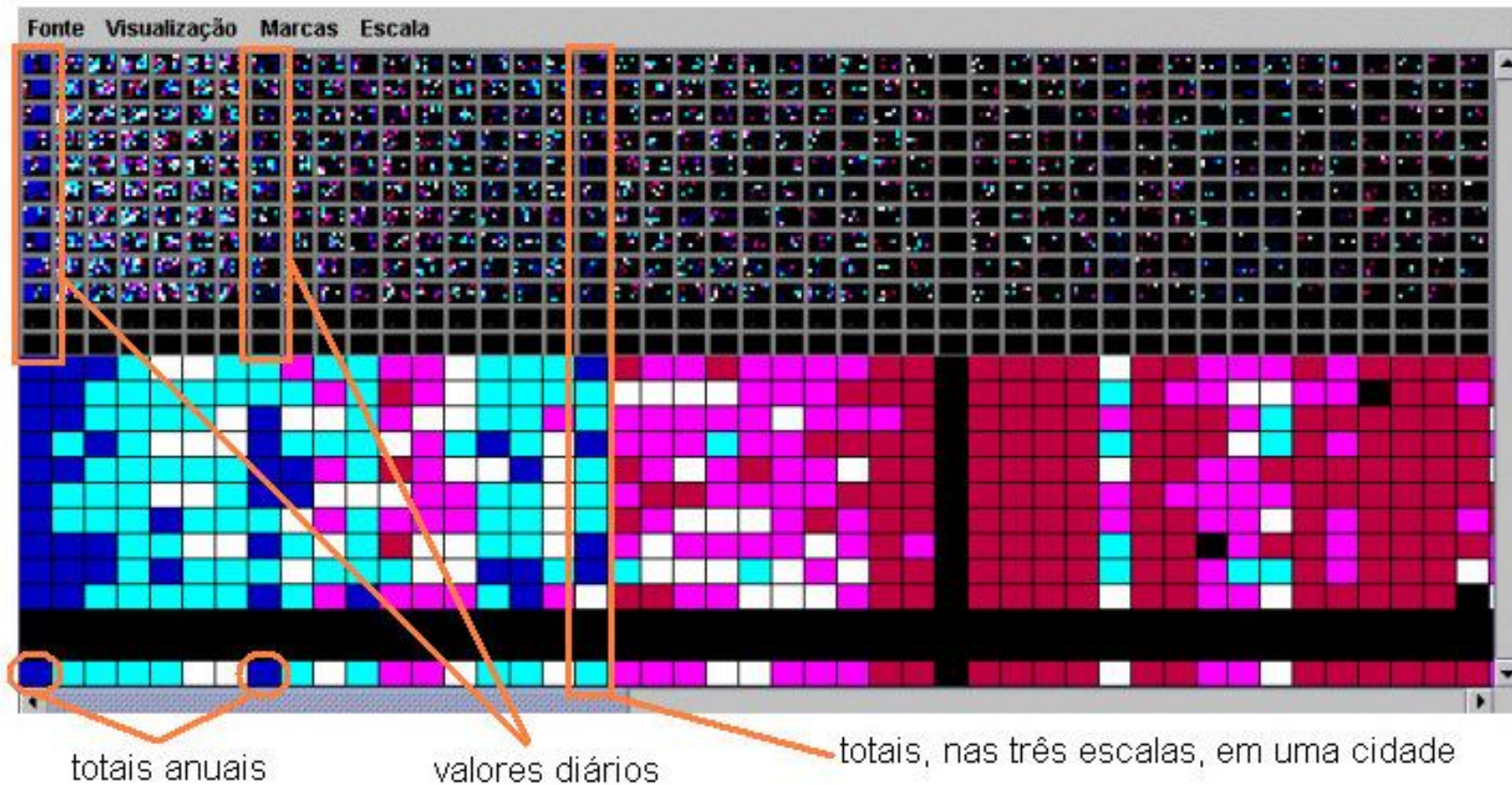
- Atividades de 150 mil clientes em um site de compras
  - X: meses; Y: valor da compra





# Técnicas Orientadas a Pixel

- Valor das transações (escalas diária, mensal e anual) em várias cidades (colunas) em um dado ano

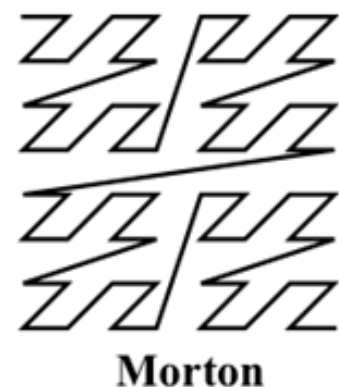
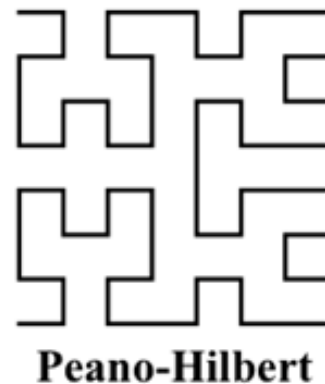
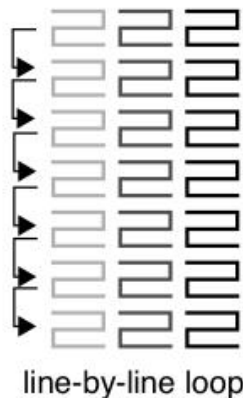
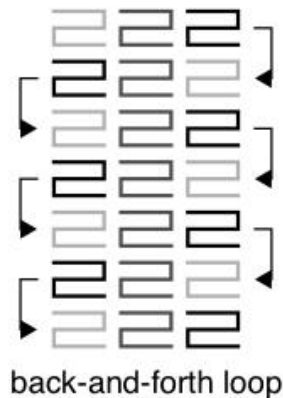
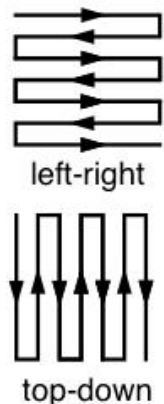


# Técnicas Orientadas a Pixel

- Nas sub-janelas, os valores dos dados são arranjados de acordo com uma ordenação do conjunto ou alguma ordem definida pelo usuário
- Correlações, dependências e outros relacionamentos podem ser percebidos pelo desenho formado entre as múltiplas janelas

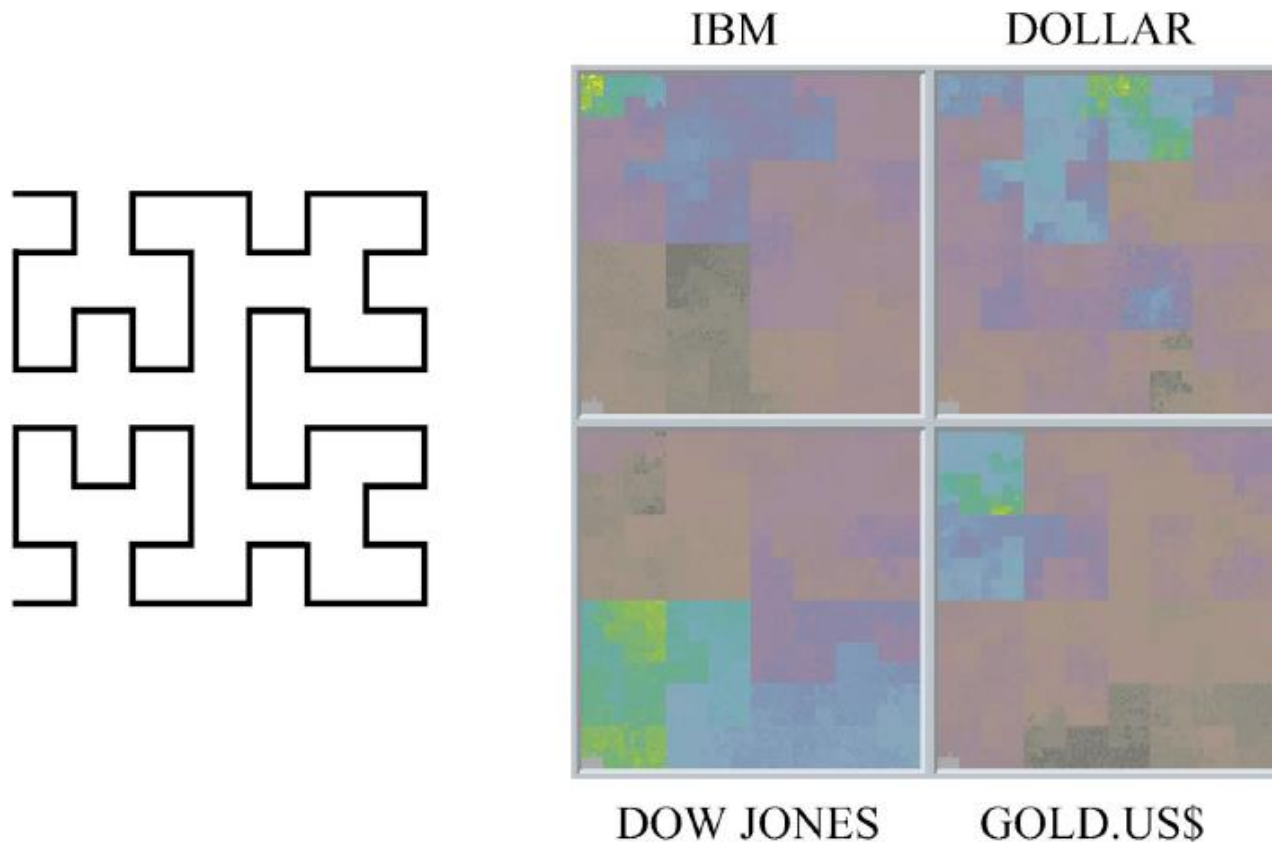
# Técnicas Orientadas a Pixel

- Uma decisão de que deve ser tomada, além do mapeamento de cores, é ordem com que os registros mapeados nas sub-janelas



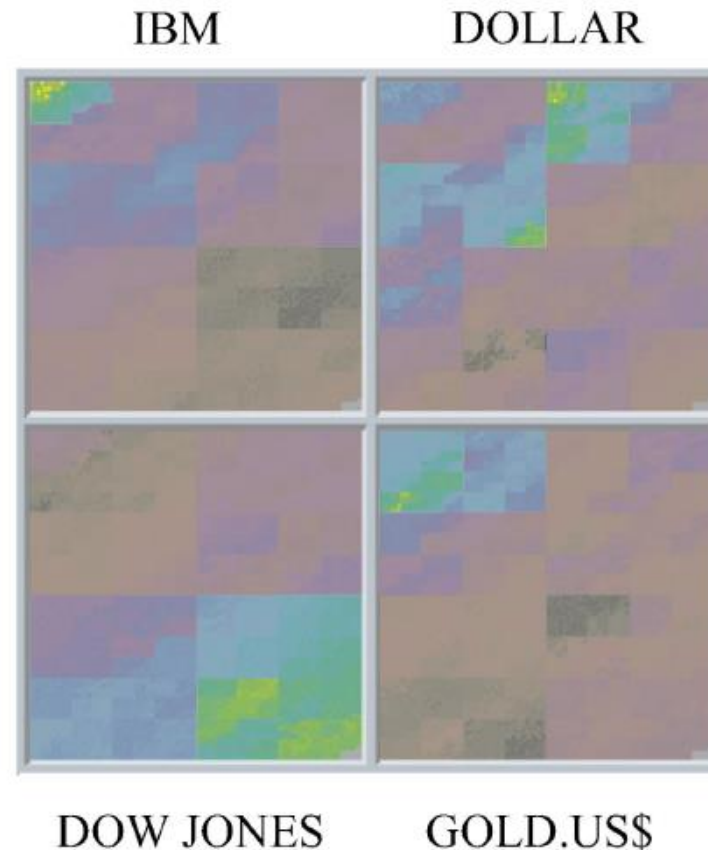
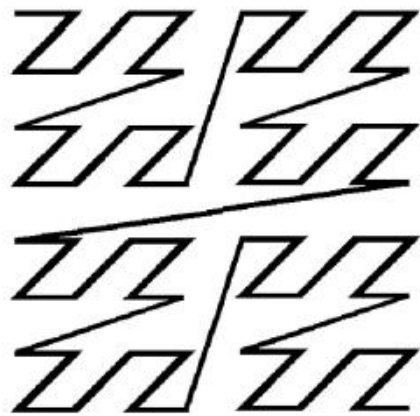
# Técnicas Orientadas a Pixel

- Dados financeiros (16200 registros) de Setembro de 1987 a Fevereiro de 1995



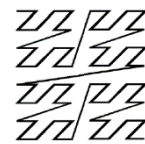
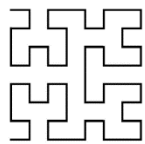
# Técnicas Orientadas a Pixel

- Dados financeiros (16200 registros) de Setembro de 1987 a Fevereiro de 1995



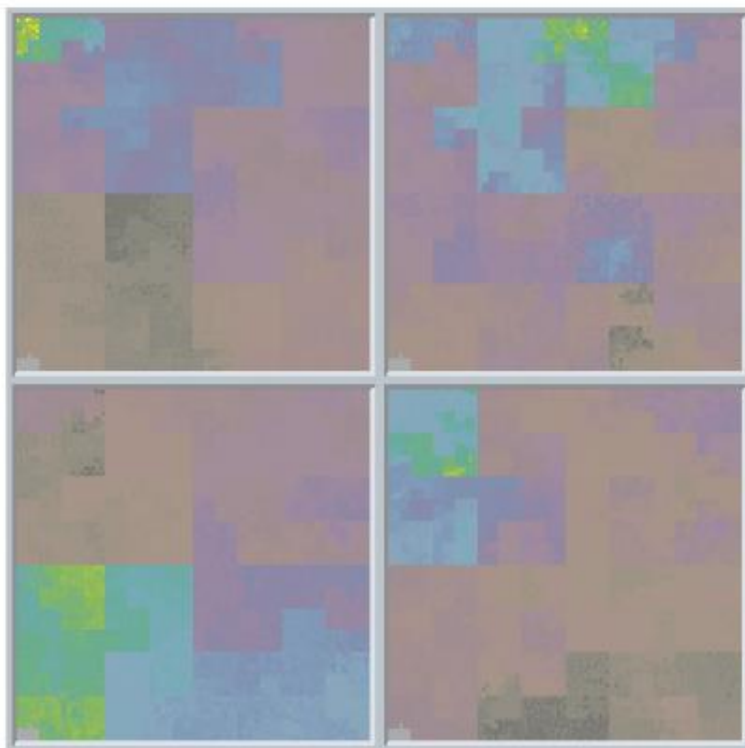


# Técnicas Orientadas a Pixel



IBM

DOLLAR

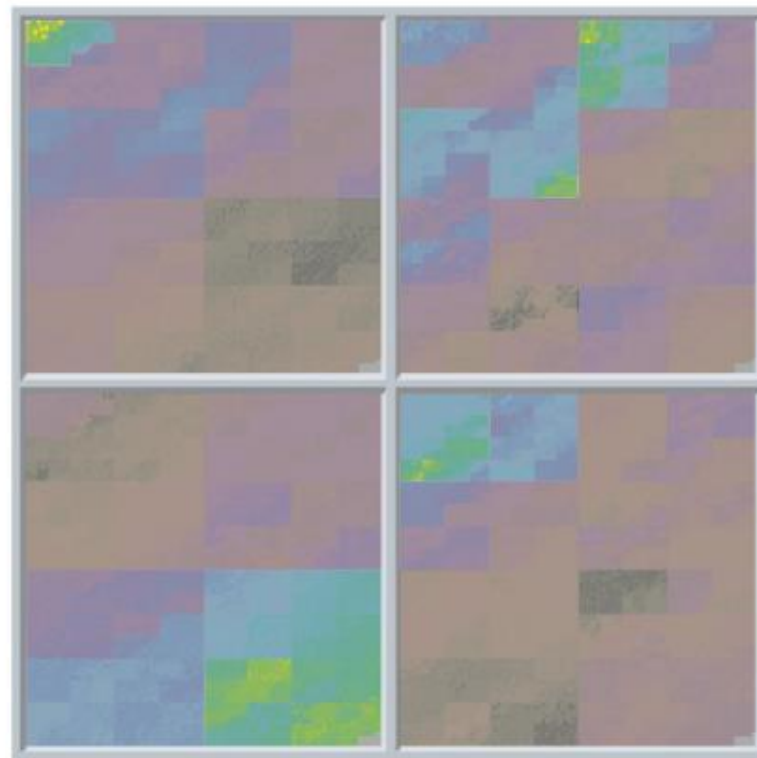


DOW JONES

GOLD.US\$

IBM

DOLLAR

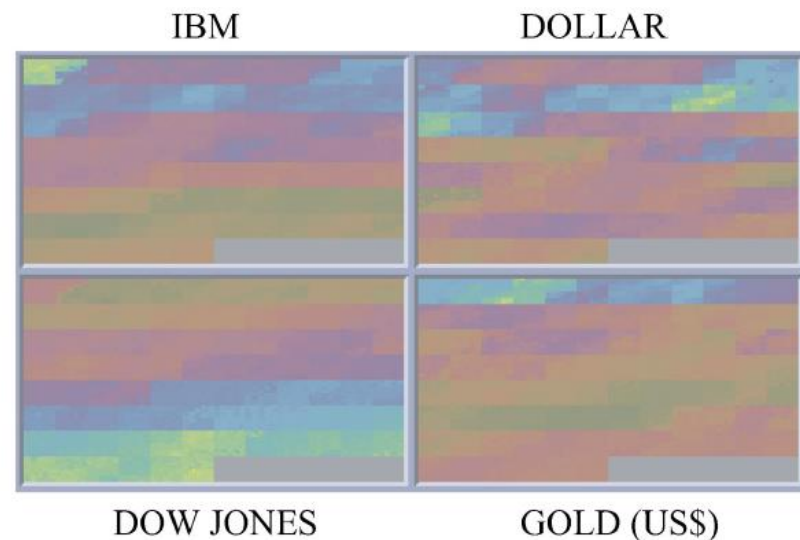
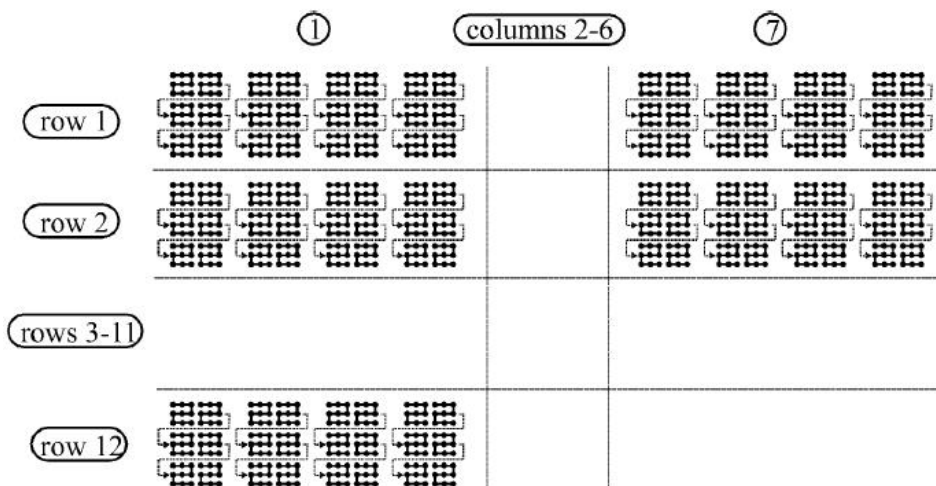


DOW JONES

GOLD.US\$

# Técnicas Orientadas a Pixel

- Esquema de desenho recursivo para permitir que padrões de nível inferior sejam usados como blocos de construção para padrões de nível superior



# Técnicas Orientadas a Pixel

- Query Dependent

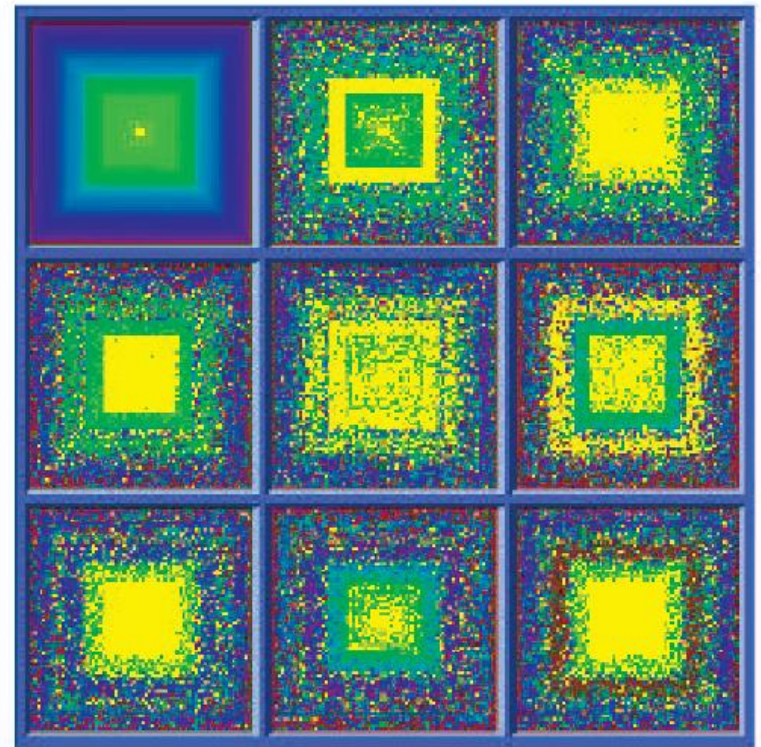
- São apresentados os itens mais relacionados com aquele utilizado como referência (consulta)

- Janela de resultados (superior esquerda)

- Posiciona as instâncias de acordo com sua relação geral para a consulta formulada

- Demais Janelas

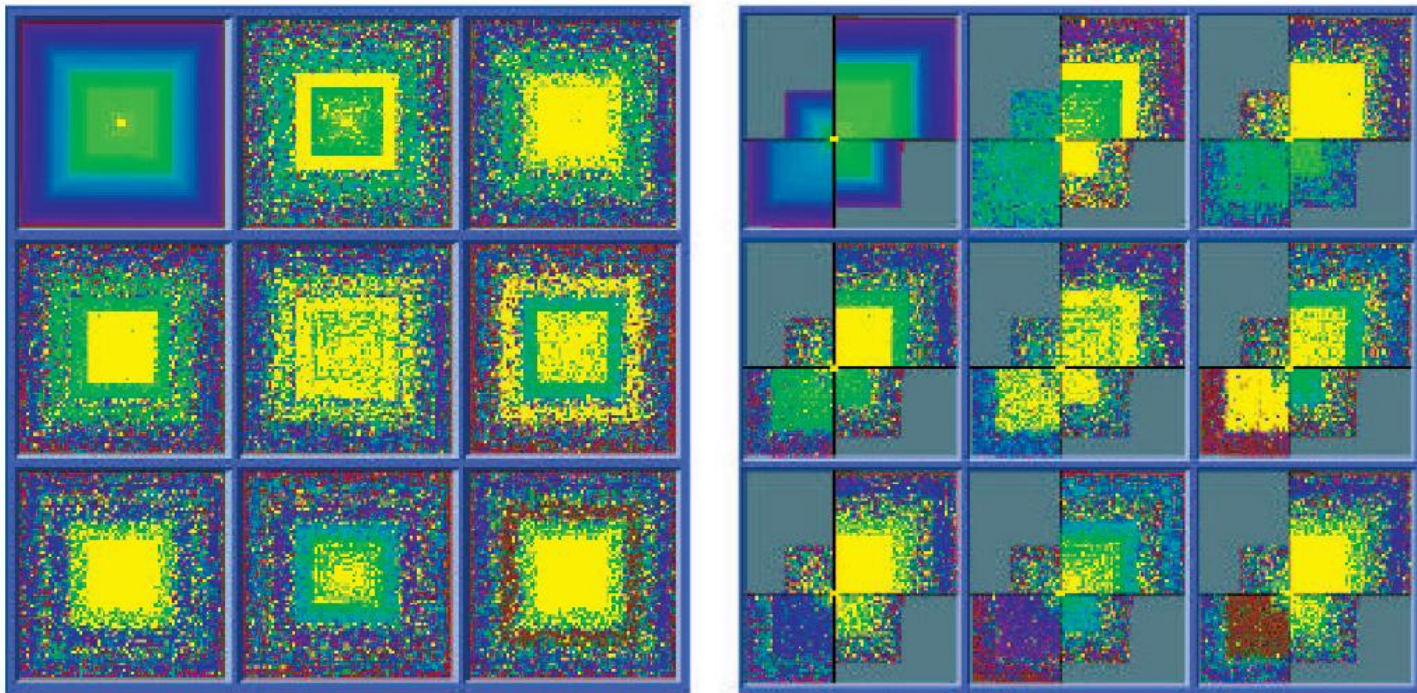
- Comportamento dos atributos





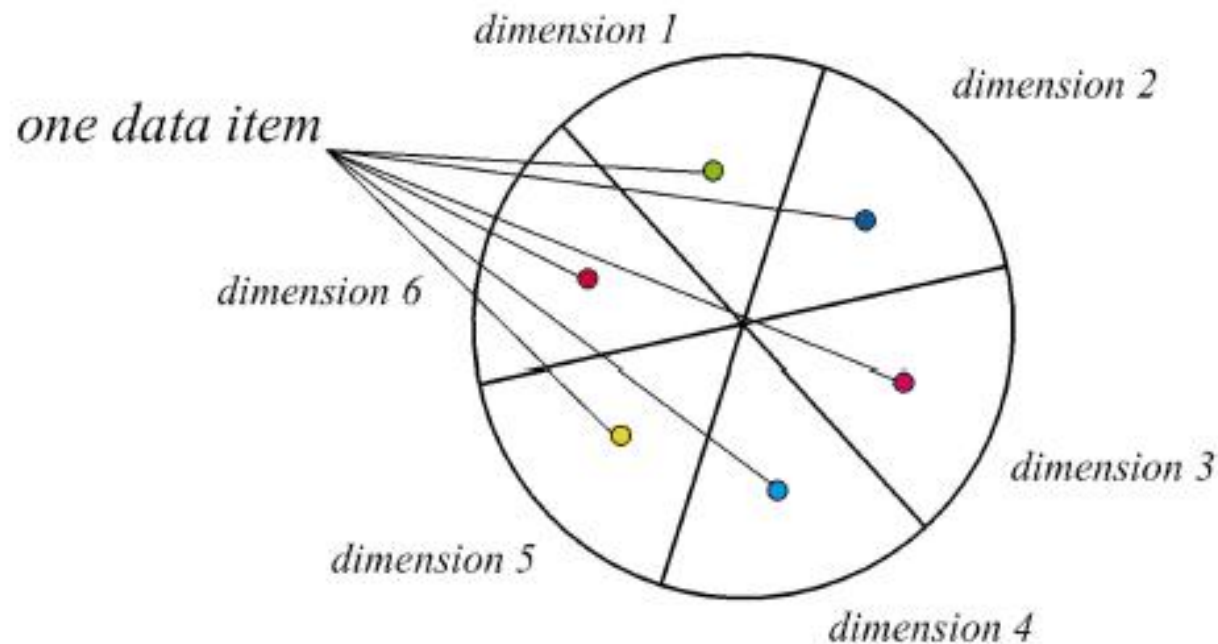
# Técnicas Orientadas a Pixel

- Sete mil instâncias com oito atributos
  - Escala de cor: do amarelo brilhante para verde, azul, vermelho escuro e preto
    - Itens de dados que melhor atendem a uma consulta são coloridos com amarelo brilhante e aqueles mais distantes são coloridos preto



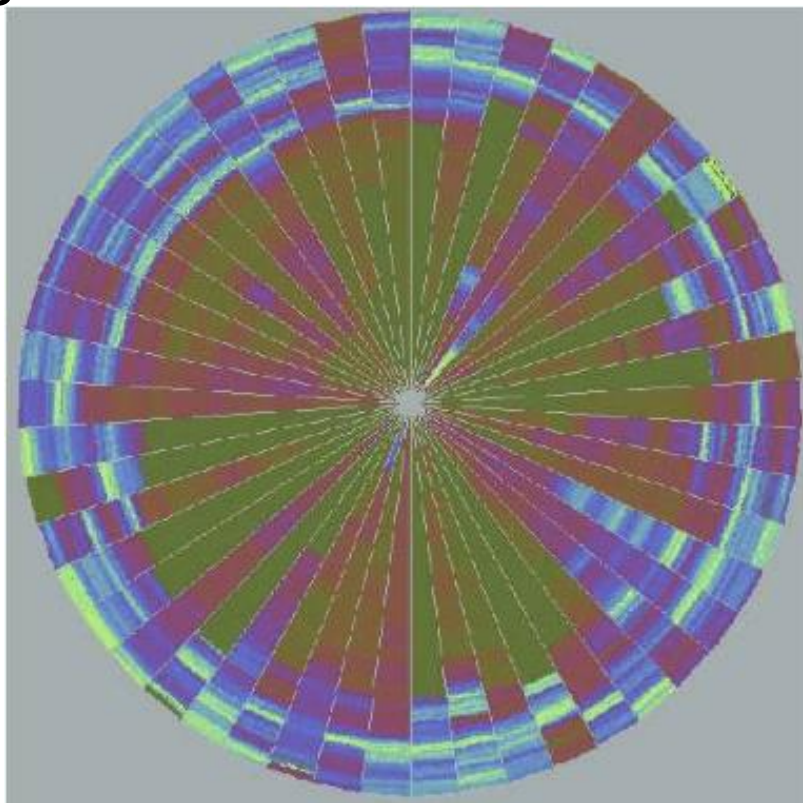
# Técnicas Orientadas a Pixel

- Uma alternativa é desenhar os pixels em segmentos de um círculo



# Técnicas Orientadas a Pixel

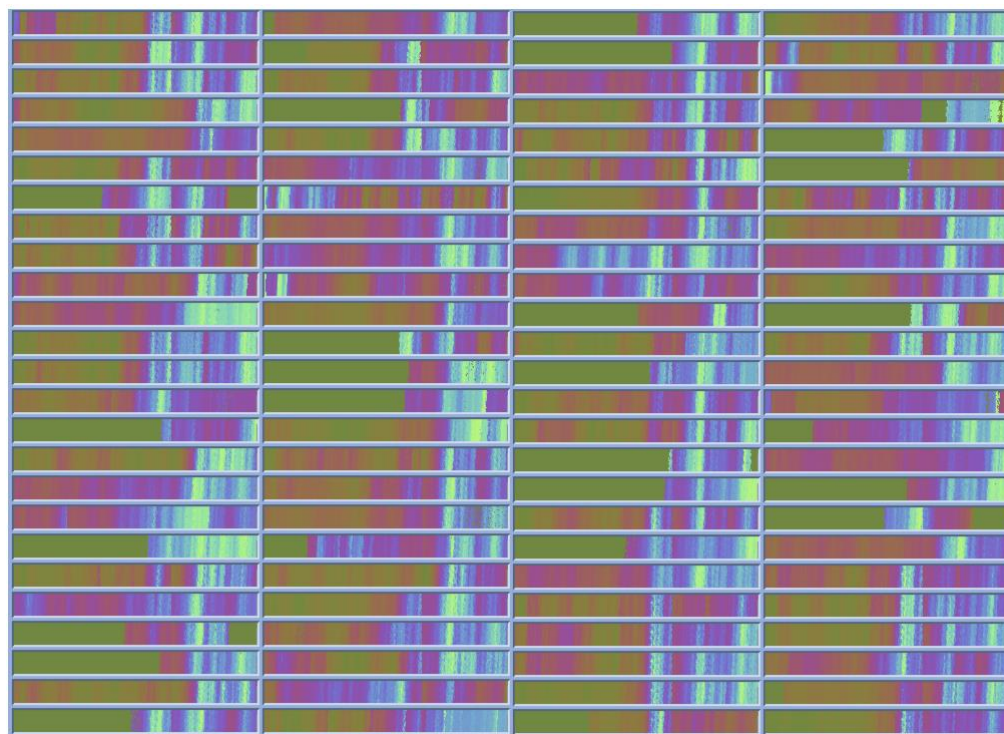
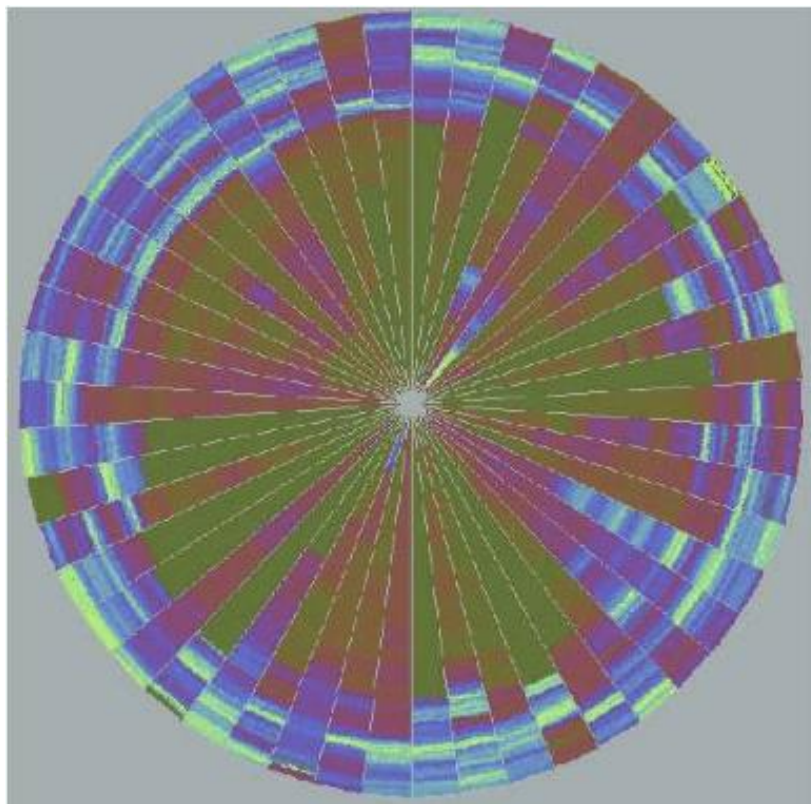
- Visualização apresentando o preço diário de 20 anos de 50 estoques de Frankfurt
  - 265 mil registros





# Técnicas Orientadas a Pixel

- A correlação e as dependências são mais facilmente percebidas, pois podemos ter uma visão geral, quando comparada com a abordagem em barras

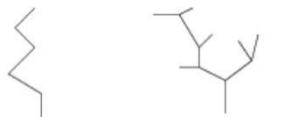


# Técnicas Iconográficas

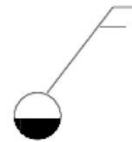


# Técnicas Iconográficas

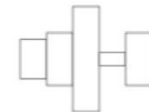
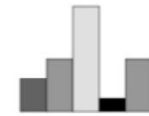
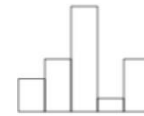
- As Técnicas Iconográficas utilizam ícones ou glifos (*glyphs*), que são controlados por um ou mais atributos dos dados
  - Por exemplo



STICKS AND TREES



ARROWS/WEATHERVANES



PROFILE GLYPHS



AUTOGLYPH/BOX GLYPH



STARS AND METROGLYPHS



FACE GLYPHS

# Técnicas Iconográficas

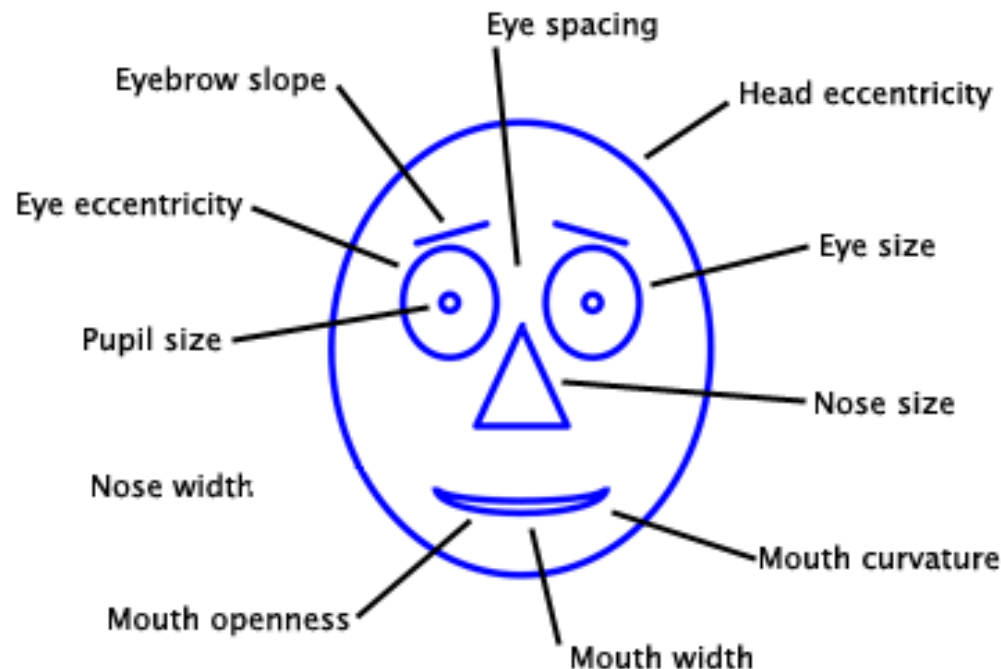
- Algumas das técnicas mais mais conhecidas são
  - Faces de Chernoff
  - Star Plot
    - Técnica Estrela
  - Stick Figures

# Faces de Chernoff

- Apresentada em 1971 por Herman Chernoff, a técnica Faces de Chernoff é baseada no desenho de faces para representar dados multivariados
- O propósito é ajudar o usuário a encontrar similaridade dos dados

# Faces de Chernoff

- Cada atributo é mapeado para um elemento da face, sendo modificado de acordo com o valor do atributo



# Faces de Chernoff

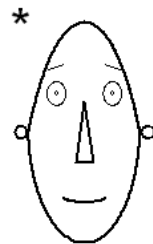
- Um exemplo de uso foi apresentado por Astel et al. (2006) para representar a concentração de determinadas substâncias na água potável em doze distritos poloneses a partir de 6 anos de monitoramento
  - Monitorando a qualidade da água e a distribuição dos componentes

# Faces de Chernoff

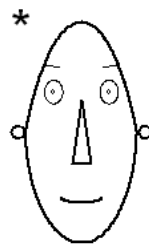
As dimensões foram mapeadas conforme o seguinte:

- Curvatura da boca: Concentração de  $CHCl_3$ ;
- Comprimento do nariz: Concentração de  $CHBrCl_2 + C_2HCl_3$ ;
- Largura do nariz: Concentração de  $CHBr_2Cl$ ;
- Ângulo da sobrancelha: Concentração de  $CH_2Cl_2$ ;
- Comprimento da sobrancelha: Concentração de  $CH_2Cl_4$ .

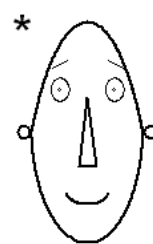
# Faces de Chernoff



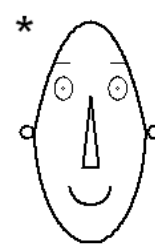
Chelm



Morena



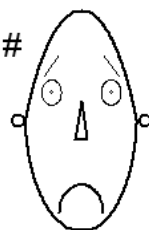
Suchanino



Niedźwiednik



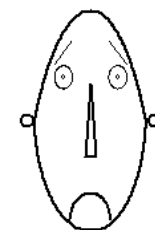
Przymorze



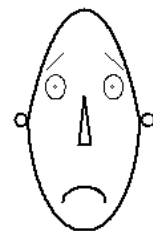
Old-Town



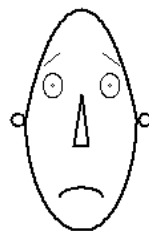
Zaspá



Żabianka



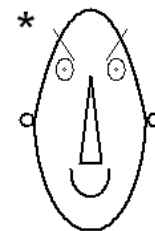
Wrzeszcz-GUT



Wrzeszcz



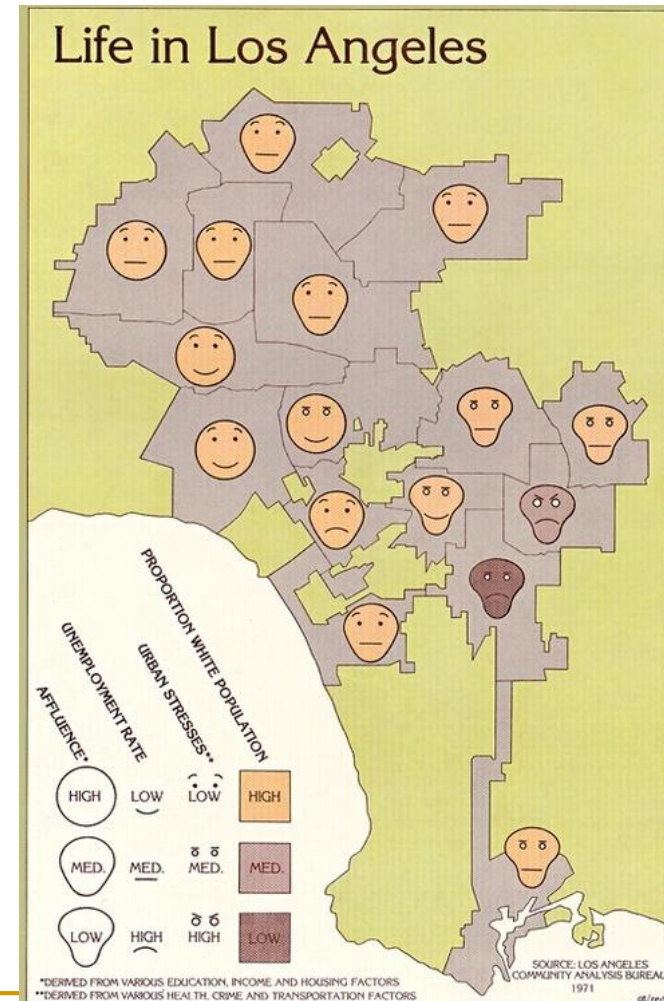
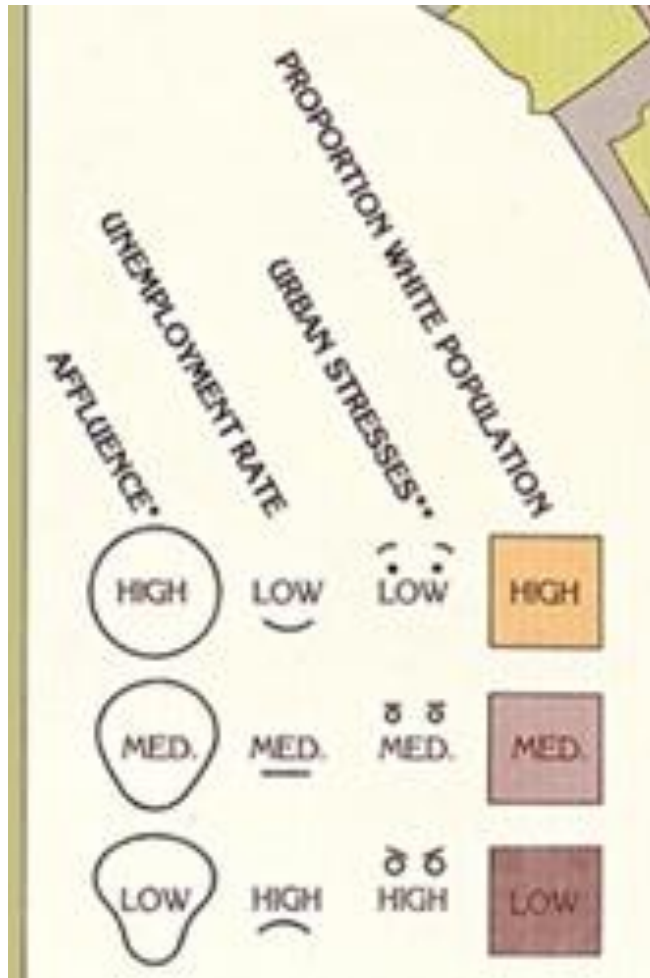
Stogi



Siedlce

# Faces de Chernoff

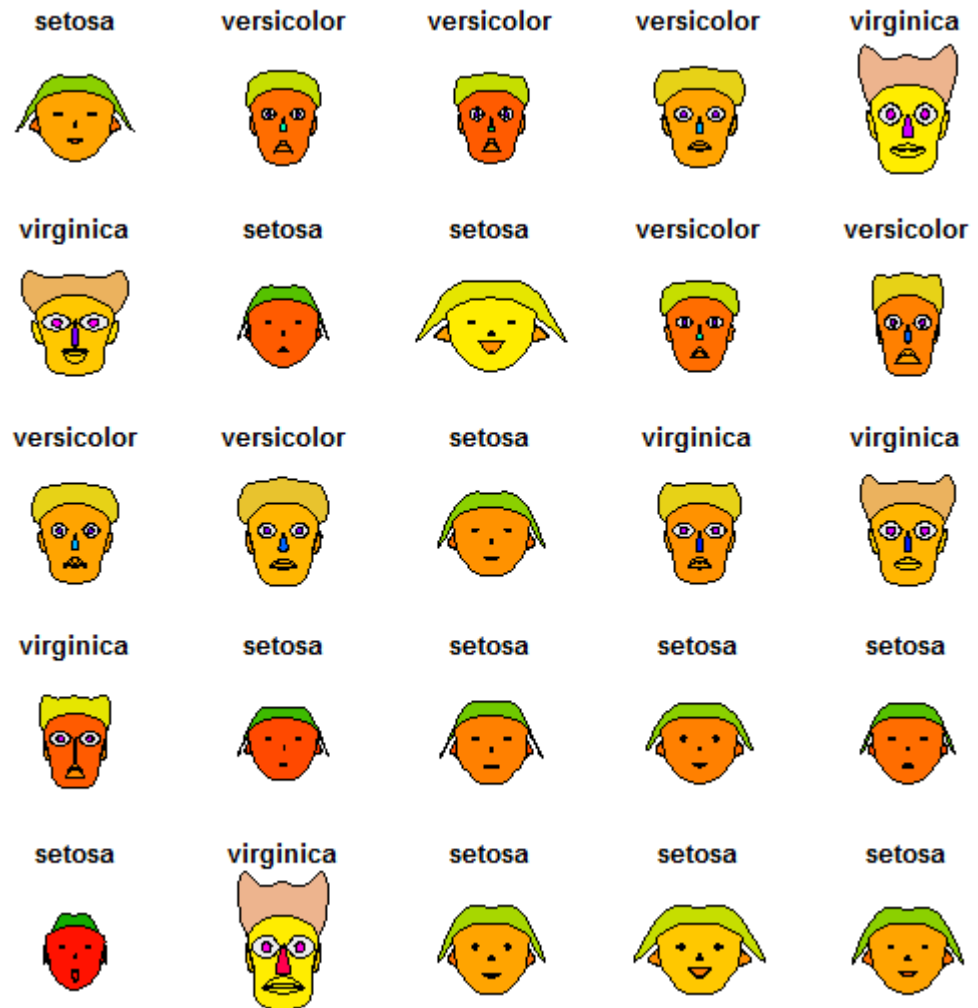
## ■ Análise sócio econômica





# Faces de Chernoff

## ■ Conjunto Iris

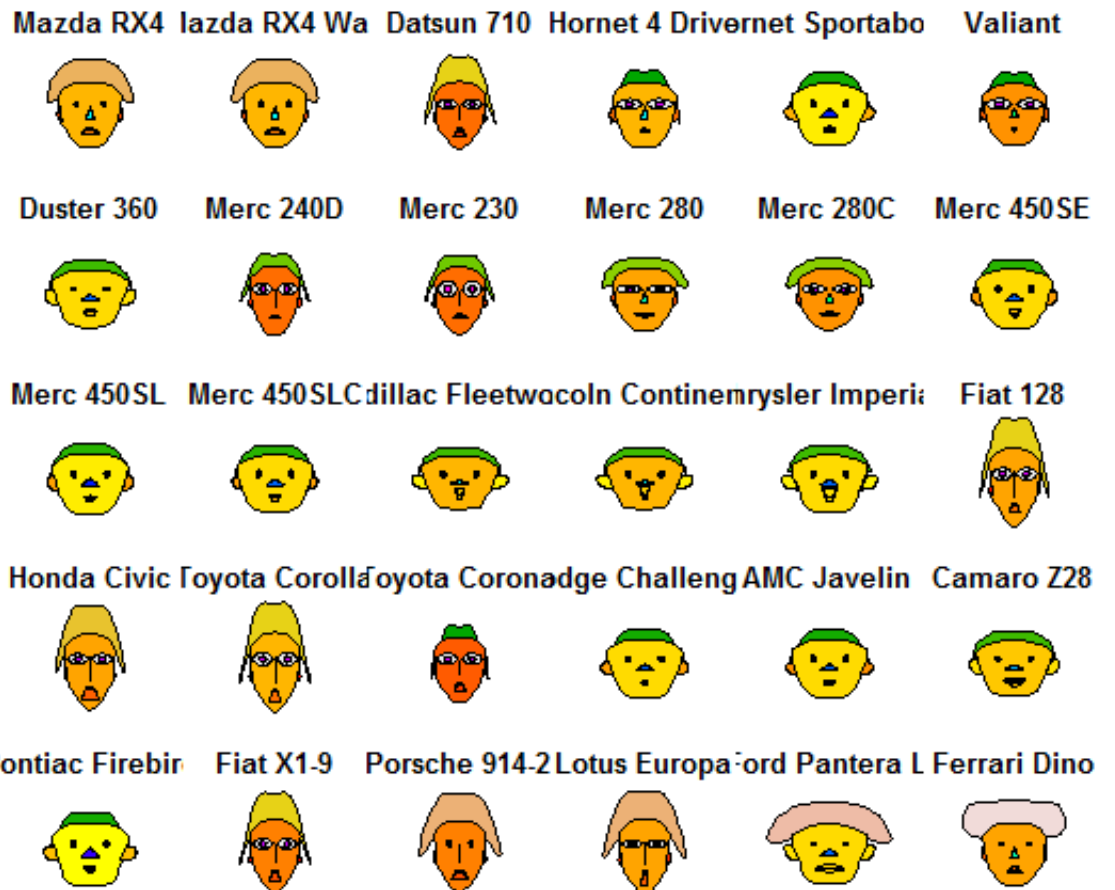


# Faces de Chernoff

## ■ Conjunto de dados de Carros

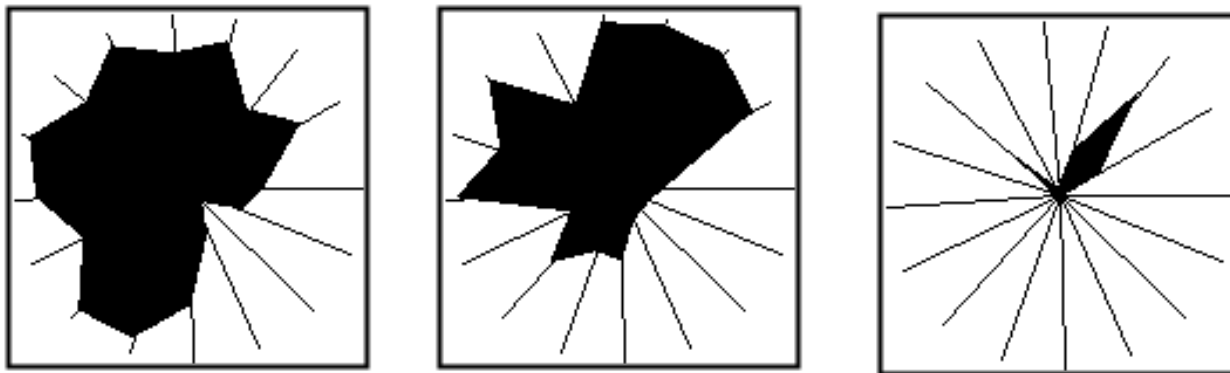
effect of variables:

modified item	Var
"height of face"	"mpg"
"width of face"	"cyl"
"structure of face"	"disp"
"height of mouth"	"hp"
"width of mouth"	"drat"
"smiling"	"wt"
"height of eyes"	"qsec"
"width of eyes"	"vs"
"height of hair"	"am"
"width of hair"	"gear"
"style of hair"	"carb"
"height of nose"	"mpg"
"width of nose"	"cyl"
"width of ear"	"disp"
"height of ear"	"hp"



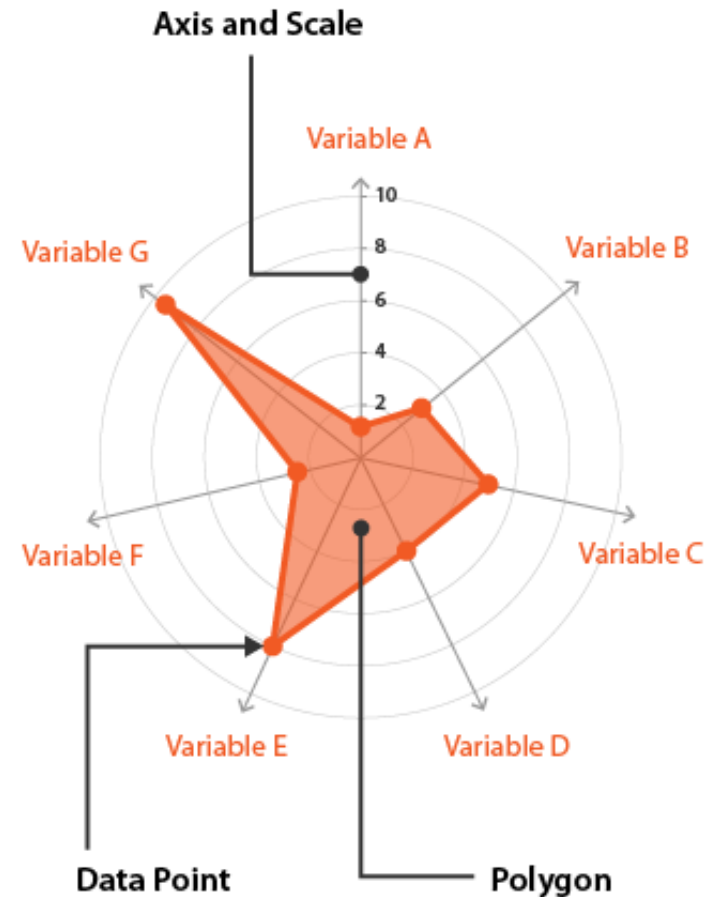
# Star Plot

- Star Plot, também conhecida como Spider Chart ou Radar Chart, permite a comparação de vários atributos quantitativos
- É útil para comparar variáveis com atributos similares e identificar *outliers*



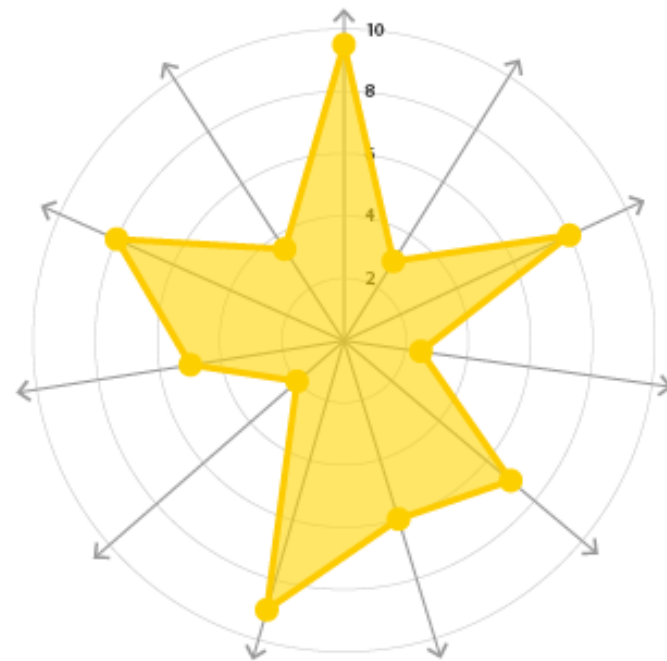
# Star Plot

- Para cada atributo há uma aresta que parte do centro de uma circunferência
- O valor de cada atributo é desenhado ao longo do eixo correspondente
- Em seguida, todos os atributos são conectados, formando um polígono (ícone) que representará a instância



# Star Plot

- Exemplo do mapeamento de duas instâncias de diferentes conjuntos de dados



# Star Plot



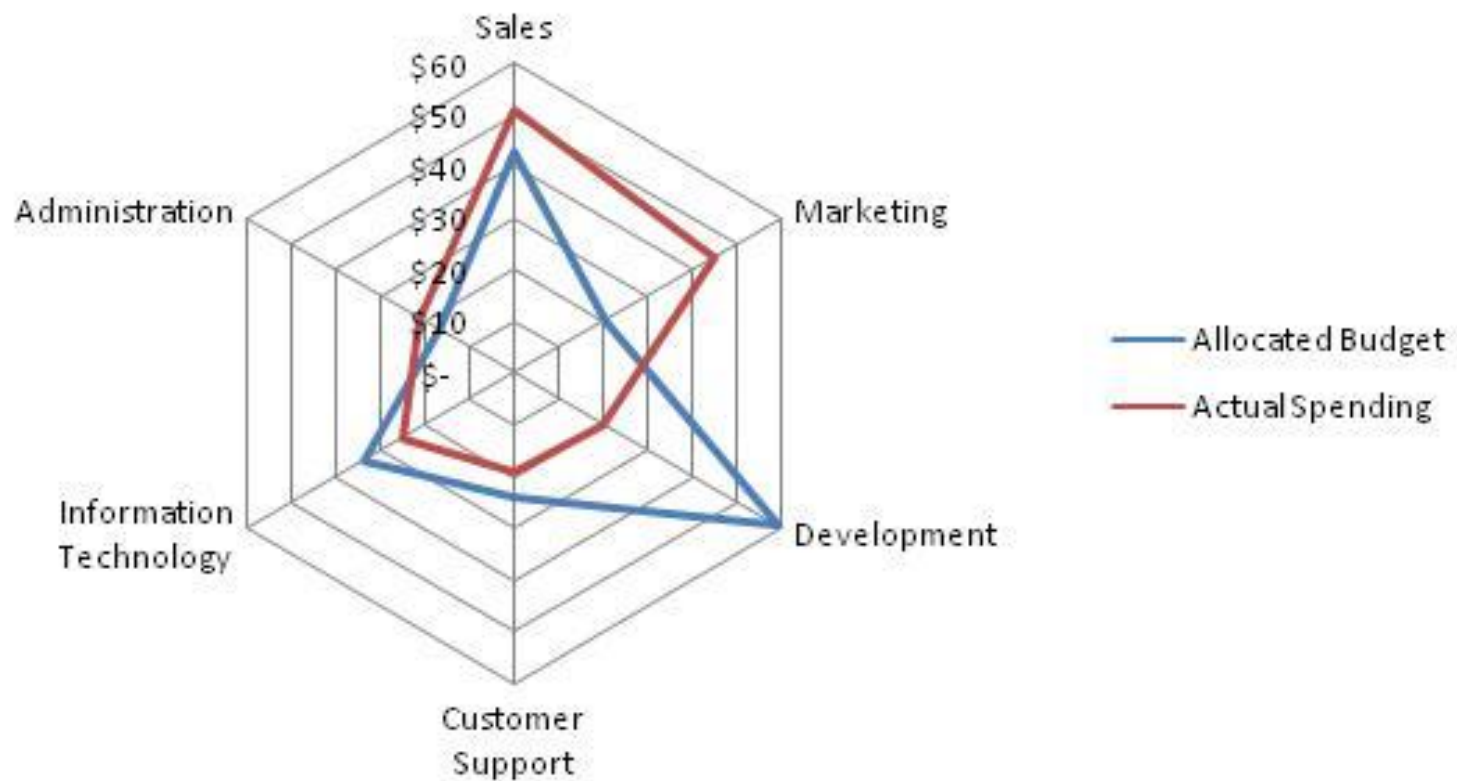


# Star Plot



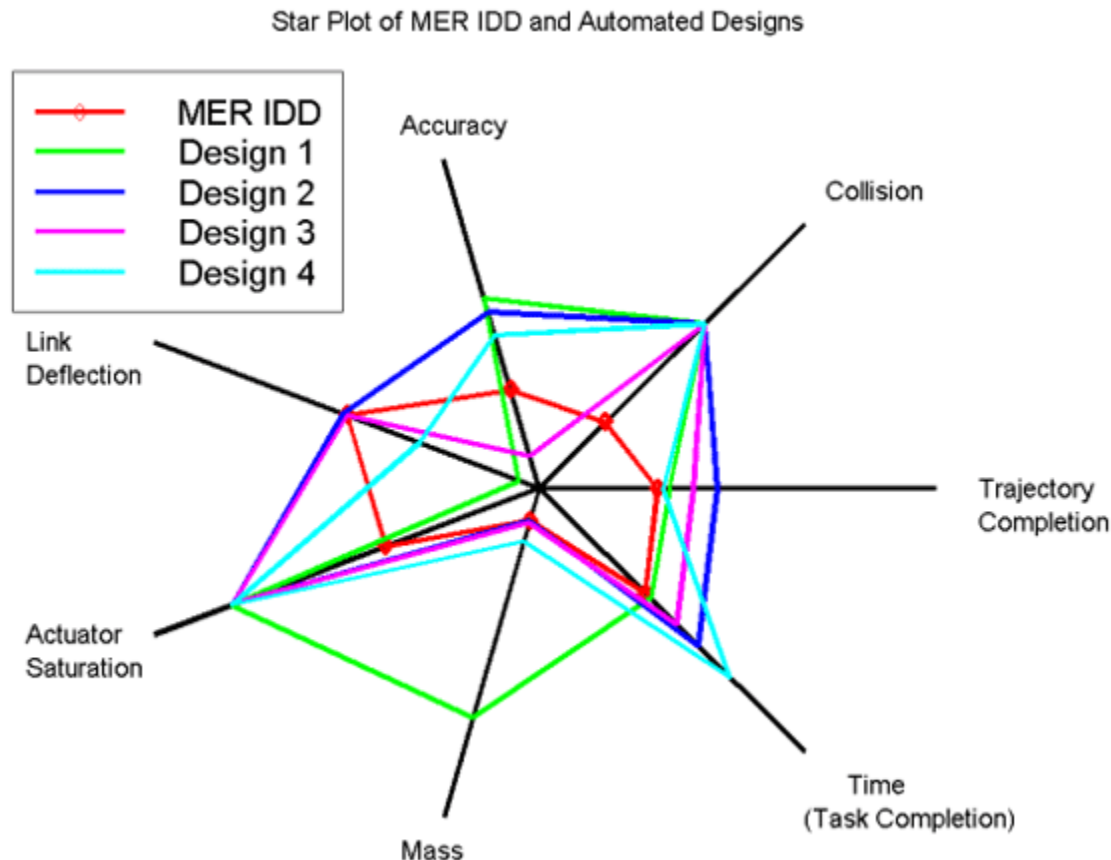
# Star Plot

- A comparação entre instâncias ou alvos pode ser realizada por meio de sobreposição



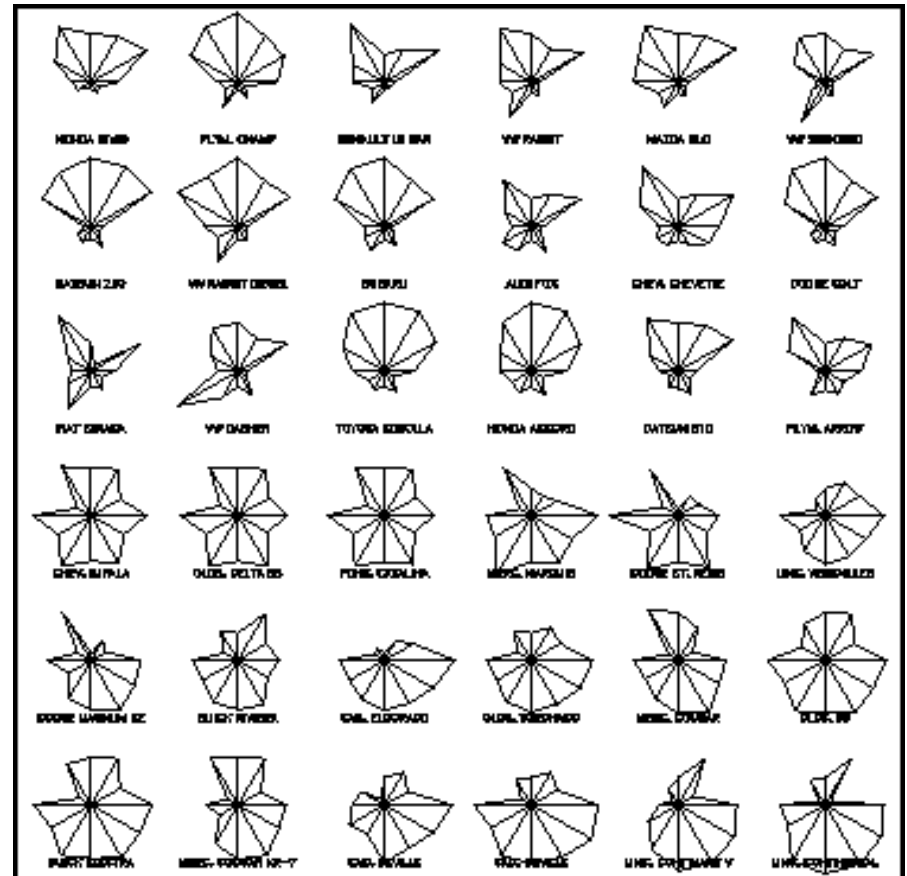
# Star Plot

- A comparação entre instâncias ou alvos pode ser realizada por meio de sobreposição



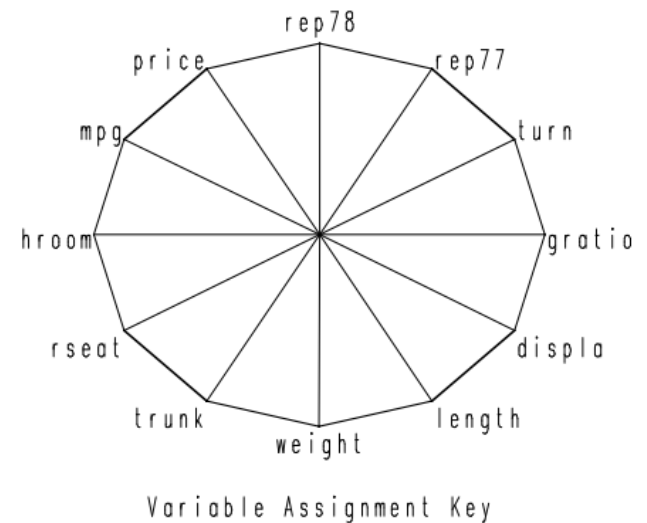
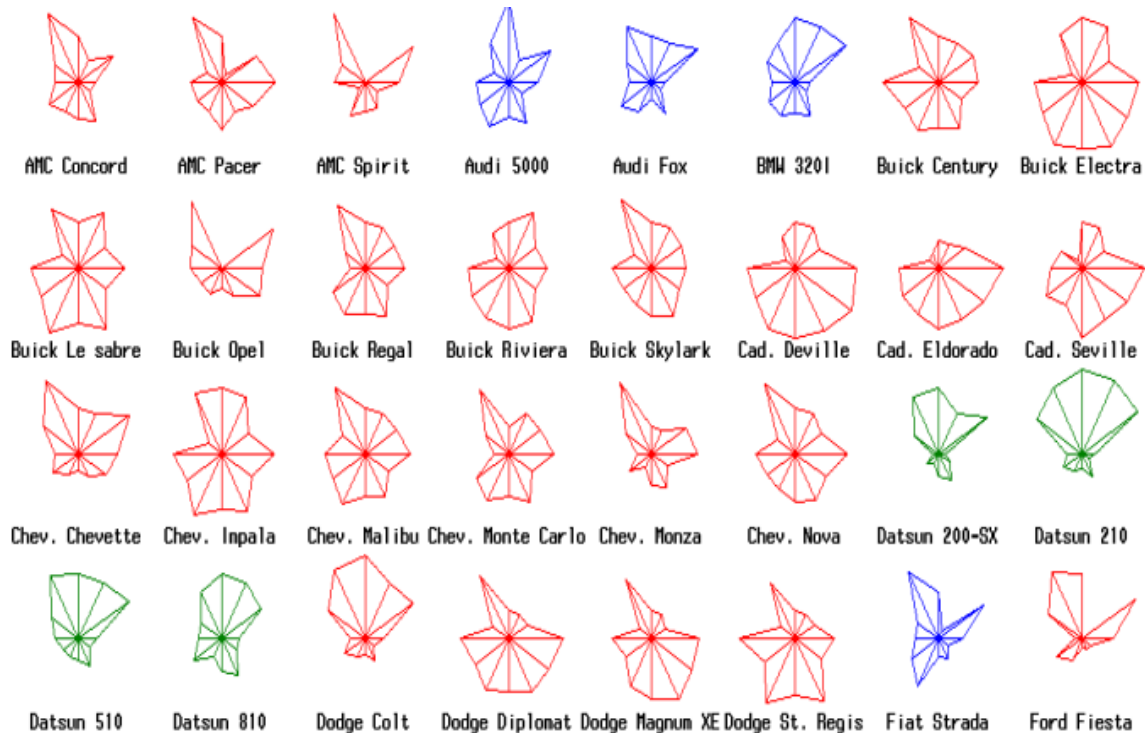
# Star Plot

- Também podemos comparar diferentes instâncias por meio do posicionamento lado a lado
  - Base de dados de carros



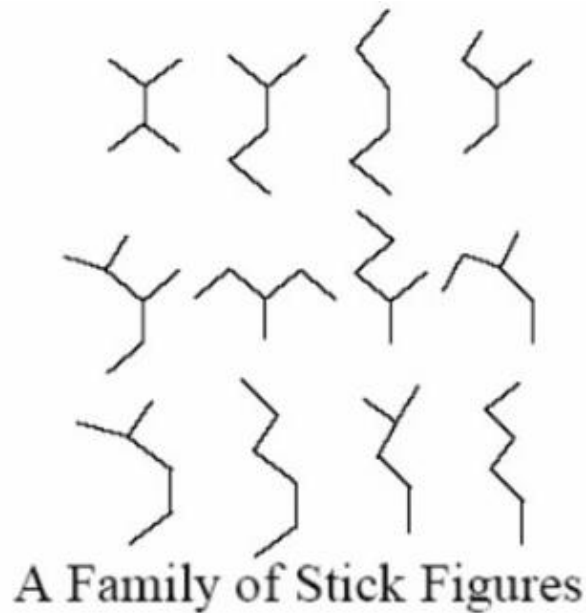
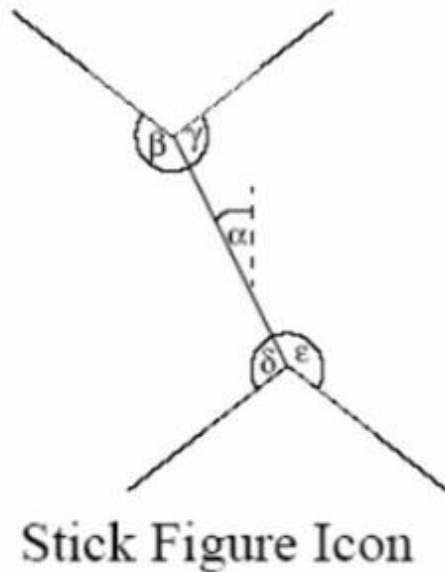
# Star Plot

## ■ Outro exemplo do conjunto de dados de carros



# Stick Figures

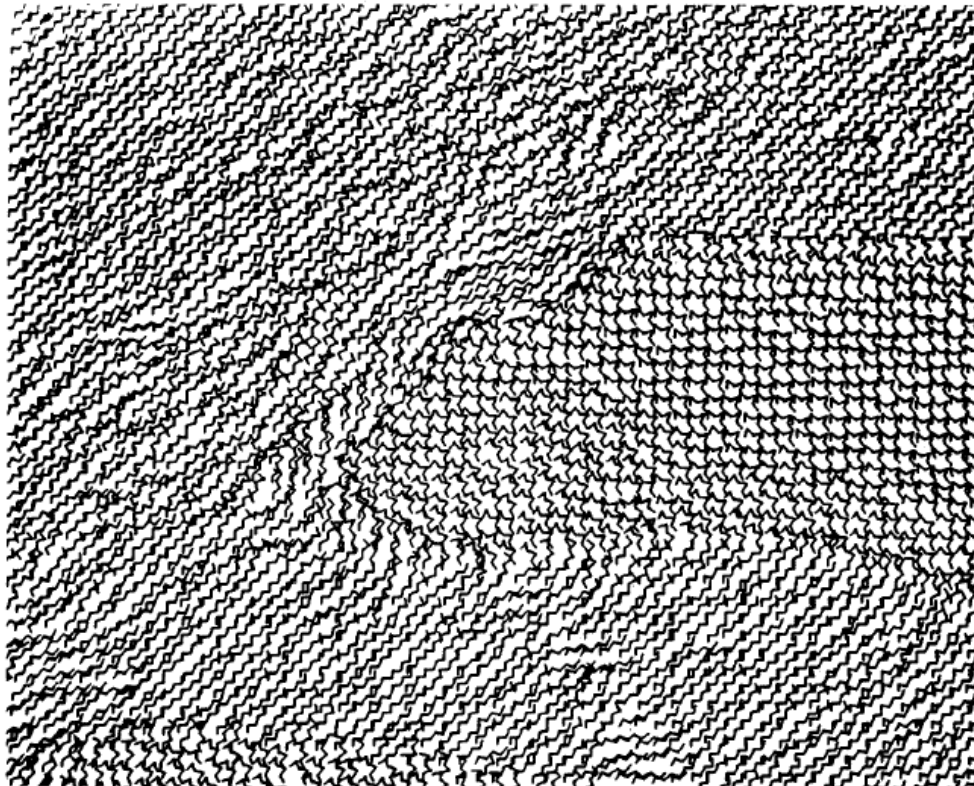
- A técnica *Stick Figures* mapeia os dados para figuras compostas por 5 arestas
  - Cada figura tem um corpo e quatro membros





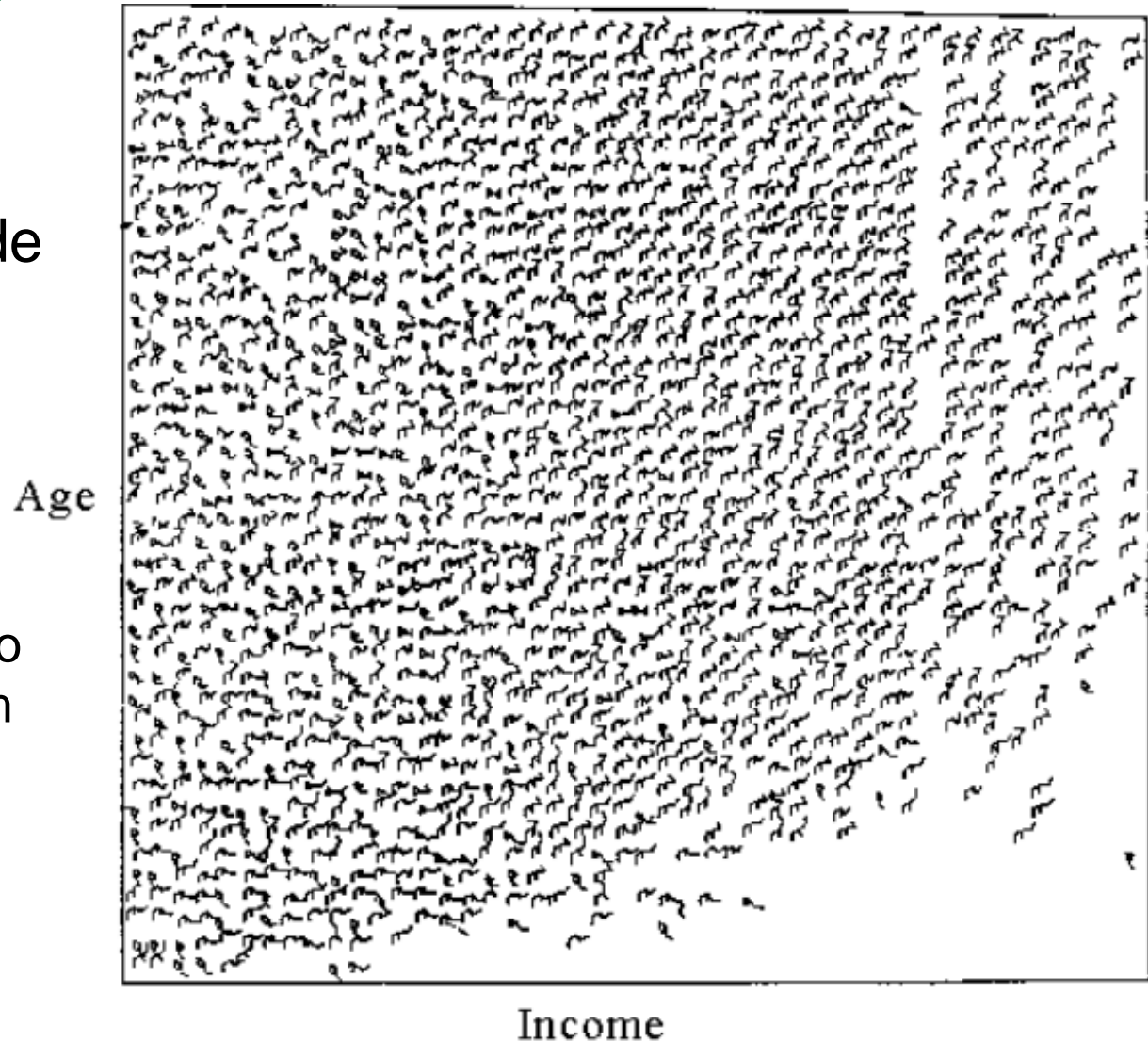
# Stick Figures

- Duas dimensões mapeiam a posição  $(x,y)$  e as restantes são mapeadas para ângulos e comprimentos dos membros
  - Assim, texturas e padrões são interpretados para obter informações sobre os dados



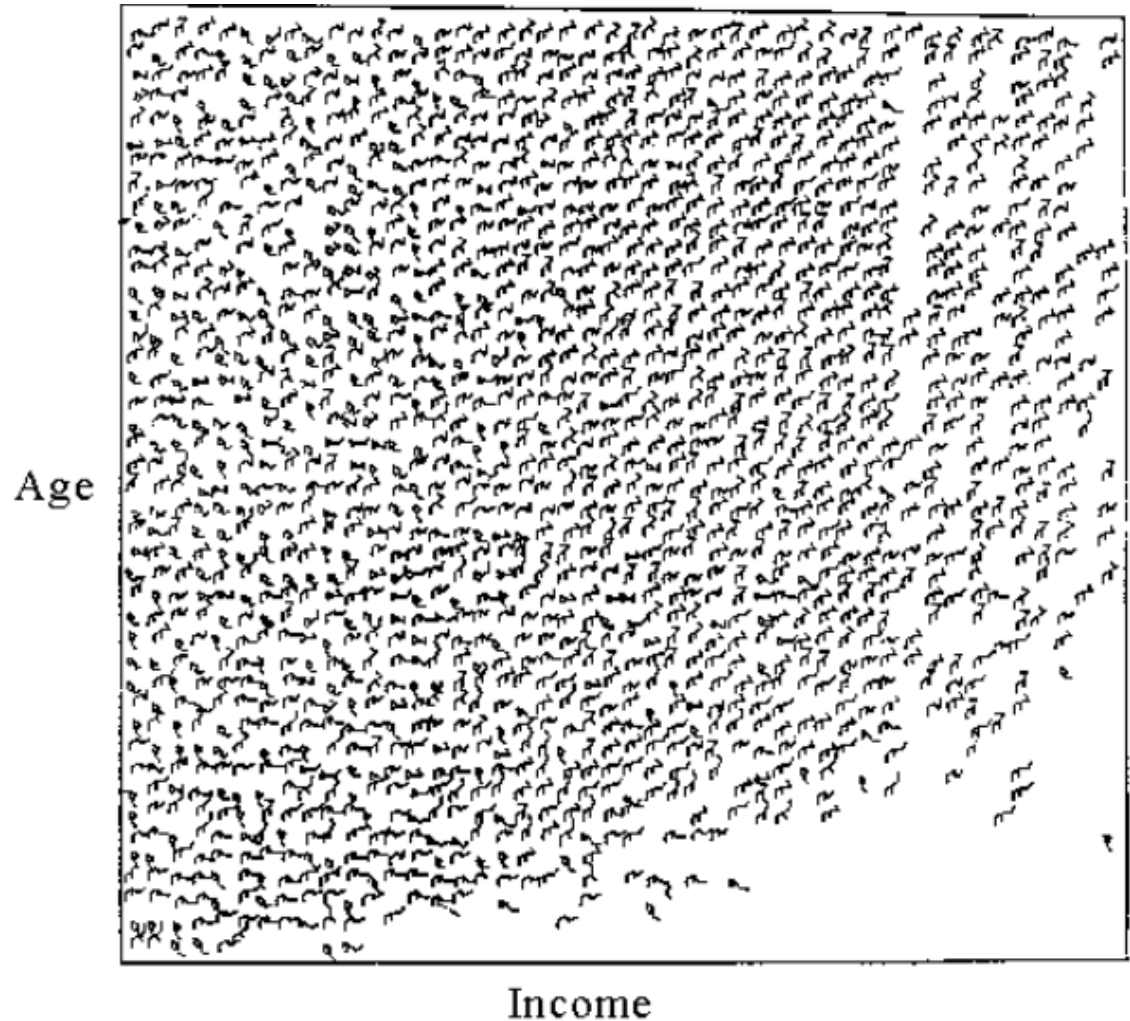
# Stick Figures

- Por exemplo, os dados do censo Norte Americano de 1998
  - A idade e renda foram mapeados para dimensões
  - Ocupação, nível educacional, estado civil e gênero foram mapeados para as características das figuras



# Stick Figures

- O agrupamento de figuras similares na parte superior indica uma dependência entre idade e renda



# Referências

- Ward, M., Grinstein, G. G., Keim, D.
  - Interactive data visualization foundations, techniques, and applications. Natick, Mass., A K Peters, 2010.
    - Capítulos 8 e 9
- Daniel A. Keim
  - Designing Pixel-Oriented Visualization Techniques: Theory and Applications. IEEE TRANSACTIONS ON VISUALIZATION AND COMPUTER GRAPHICS, VOL. 6, NO. 1, 2000
- Daniel A. Keim and Hans-Peter Kriegel
  - VisDB: Database Exploration Using Multidimensional Visualization. IEEE Computer Graphics and Applications, vol. 14, no. 5, pp. 40-49, Sept. 1994



# Referências

- Maria Cristina Ferreira de Oliveira e Haim Levkowitz
  - From visual data exploration to visual data mining: a survey. IEEE Transactions on Visualization and Computer Graphics 9 (3), 378-394, 2003
  
- Tese de Doutorado
  - Milton Hirokazu Shimabukuro, Visualizações Temporais em uma Plataforma de Software Extensível e Adaptável, ICMC/USP, 2004
  
- Ronald M Pickett, Georges G. Grinstein
  - Iconographic Displays for Visualizing Multidimensional Data, Proceedings of IEEE International Conference on Systems, Man, and Cybernetics, Beijing and Shenyang, China, 514-419, 1998

# Referências

- Alfred Inselberg
  - “The Plane with Parallel Coordinates.” *The Visual Computer* 1:2 (1985), 69–91
- Patrick Hoffman, Georges Grinstein, Kenneth Marx, Ivo Grosse, and Eugene Stanley
  - “DNA Visual and Analytic Data Mining.” In *VIS '97: Proceedings of the 8th Conference on Visualization '97*, pp. 437–ff. Los Alamitos, CA: IEEE Computer Society Press, 1997.
- Brian Johnson and Ben Shneiderman
  - “Tree-Maps: A Space-Filling Approach to the Visualization of Hierarchical Information Structures.” In *VIS '91: Proceedings of the 2nd Conference on Visualization '91*, pp. 284–291. Los Alamitos, CA: IEEE Computer Society Press, 1991.



# Referências

- John Stasko and Eugene Zhang.
  - “Focus+Context Display and Navigation Techniques for Enhancing Radial, Space-Filling Hierarchy Visualizations.” In Proceedings of the IEEE Symposium on Information Visualization, pp. 57–65. Los Alamitos, CA: IEEE Computer Society, 2000.
- George G. Robertson, Jock D. Mackinlay, and Stuart K. Card.
  - “Cone Trees: Animated 3D Visualizations of Hierarchical Information.” In Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, pp. 189–194. New York: ACM Press, 1991.

# Referências

- A. Astel, K. Astel, M. Biziuk, J. Namieśnik.
  - ❑ Clasifcation of Drinking Water Samples Using the Chernoff's Faces Visualization Approach, Polish J. of Environ. Stud. Vol. 15, No. 5 (2006), 691-697
- Shneiderman
  - ❑ Bem Shneiderman. The eyes have it: a task by data type taxonomy for information visualization. In Proceedings of the 1996 IEEE Symposium on Visual Languages, pp. 336-343. Washington, DC: IEEE Computer Society, 1996
- Keim
  - ❑ Daniel A. Keim. Information Visualization and Visual Data Mining. IEEE Transactions on Visualization and Computer Graphics, 8:1 (2002), 1-8

# Referências

- J. B. Kruskal and M. Wish
  - Multidimensional Scaling. Quantitative Applications in the Social Sciences Series, Newbury Park: Sage Publications, 1978.
- História da Treemap
  - <http://www.cs.umd.edu/hcil/treemap-history/>