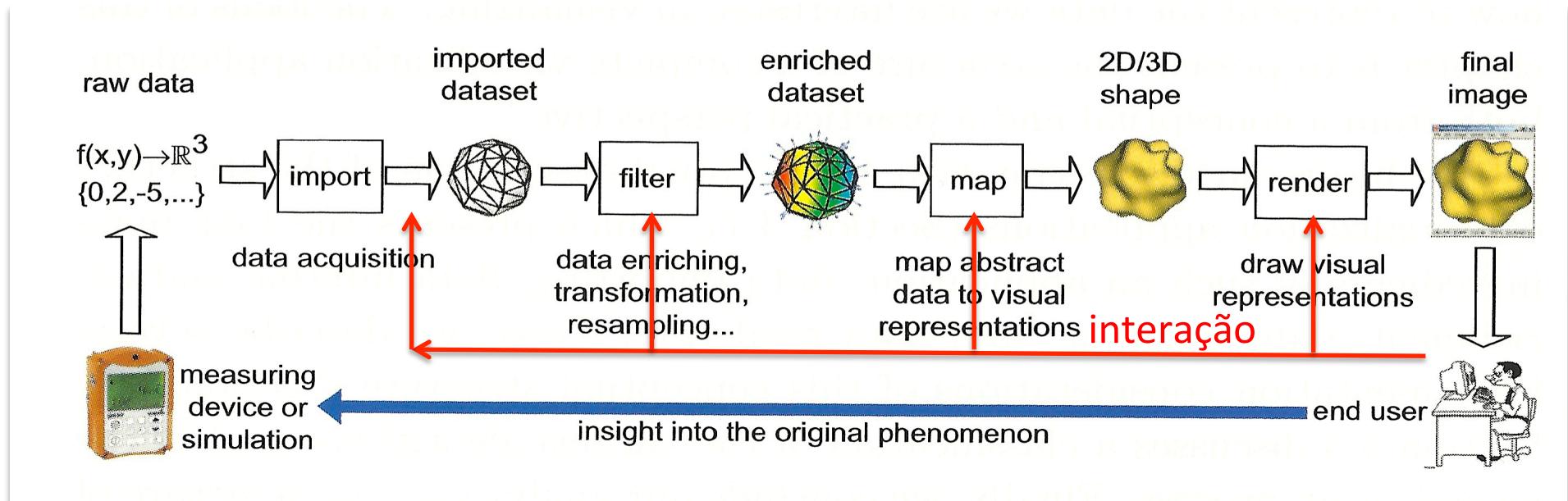


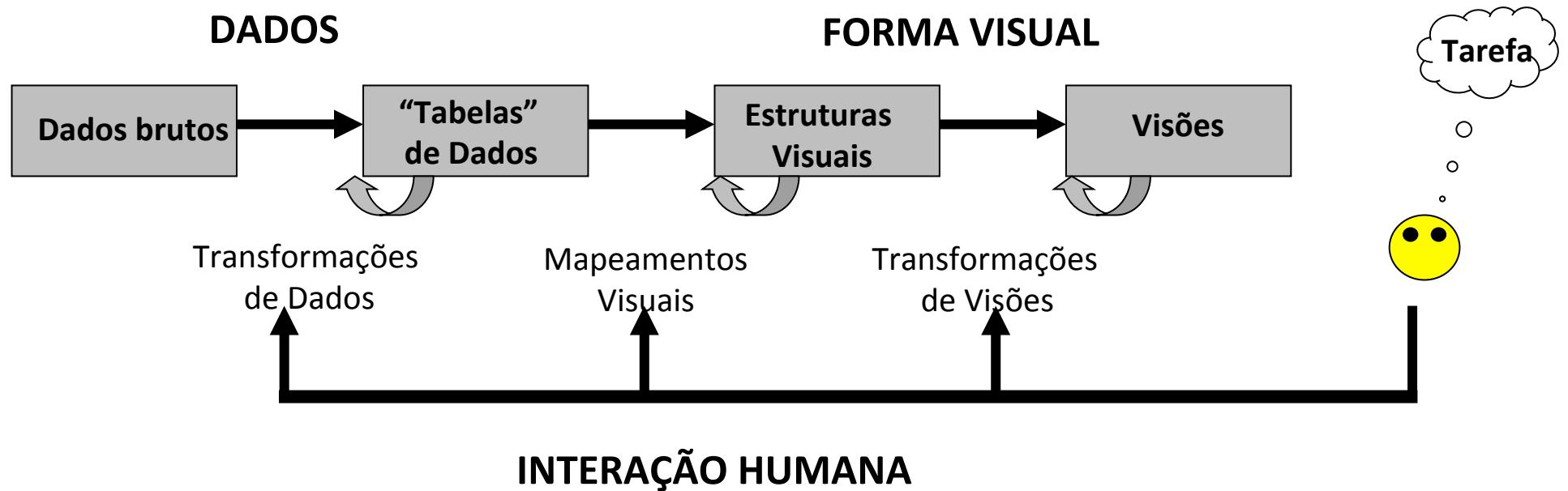
# Caracterização de dados

---

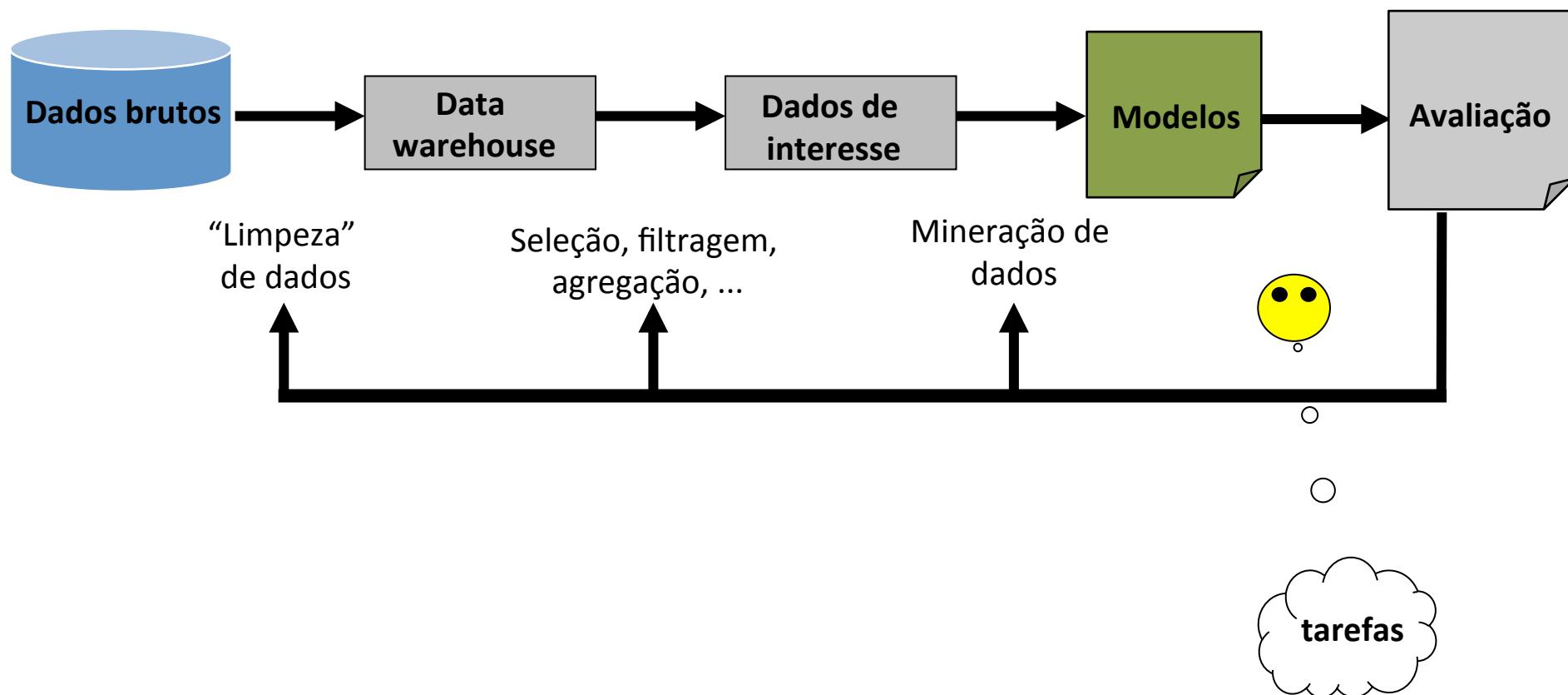
# Caracterização da área



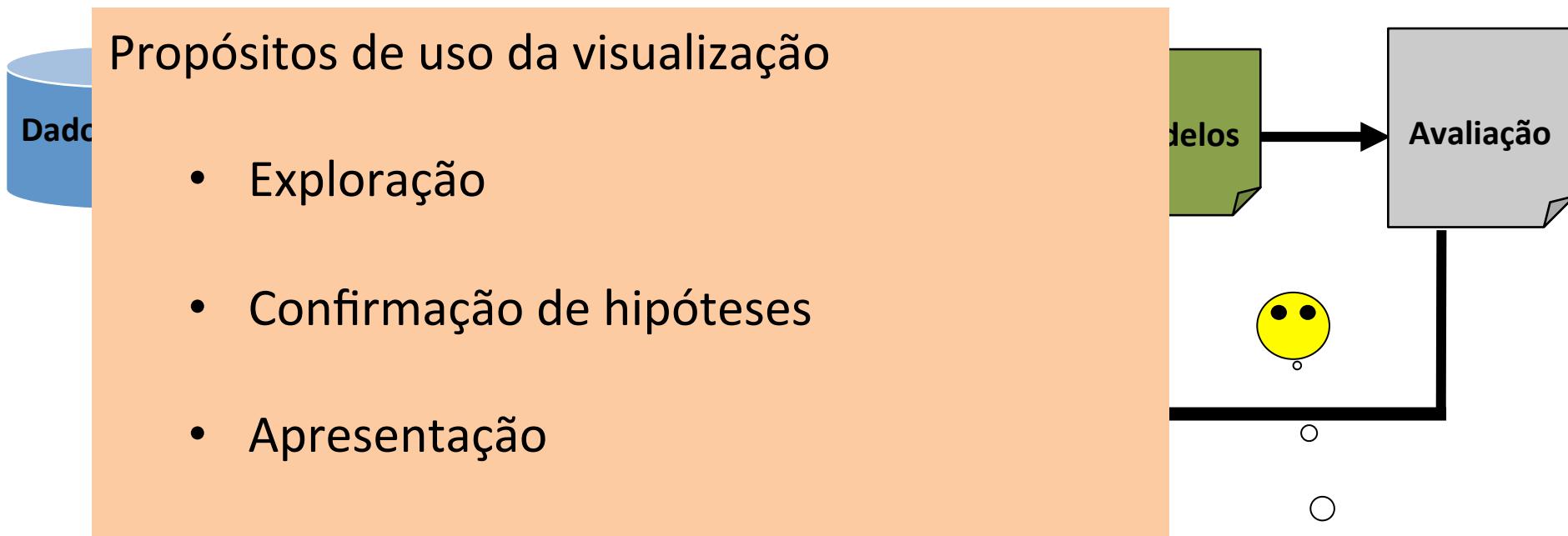
# Modelo de referência de visualização



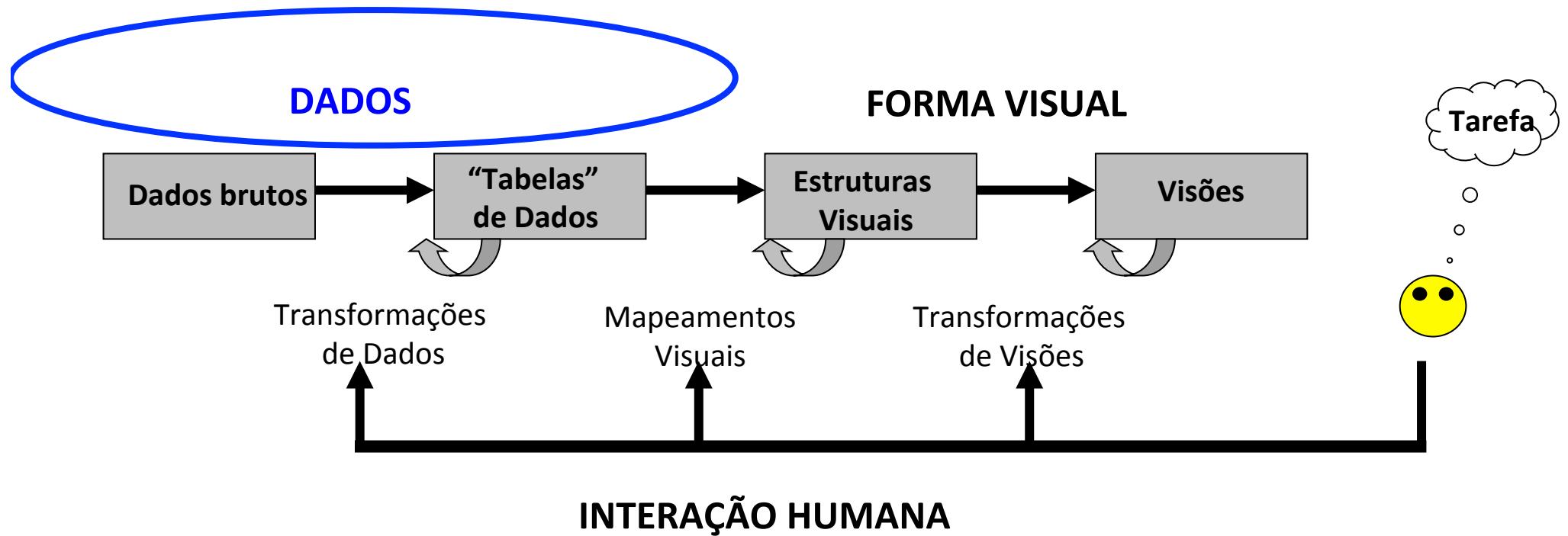
# Pipeline de descoberta de conhecimento



# Pipeline de descoberta de conhecimento



# Modelo de referência de visualização

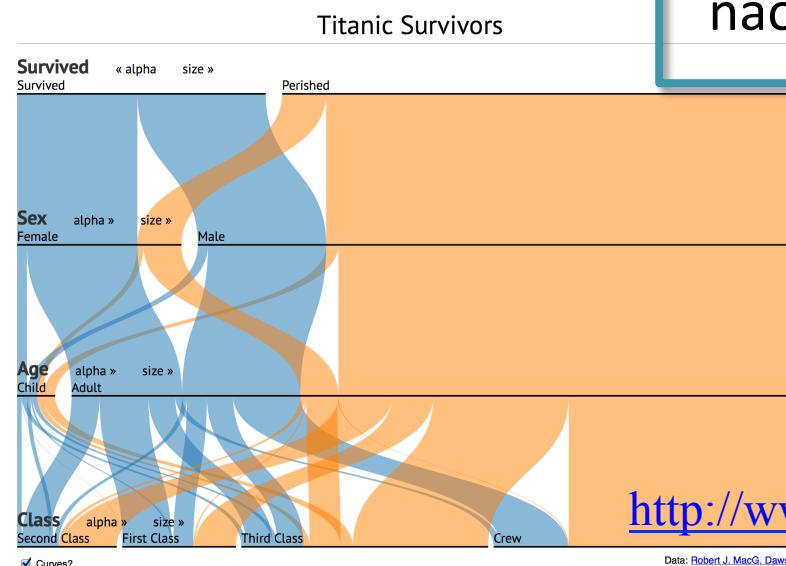


# Princípios

- Caracterização de dados
- Percepção
- Interação e tarefas de usuário



SCI - Univ. of Utah



<http://www.jasondavies.com/parallel-sets/>

Data: Robert J. MacG. Dawson.

**Dados “científicos”:**  
informações associadas a posições/regiões de um espaço (geometria implícita ou explícita)

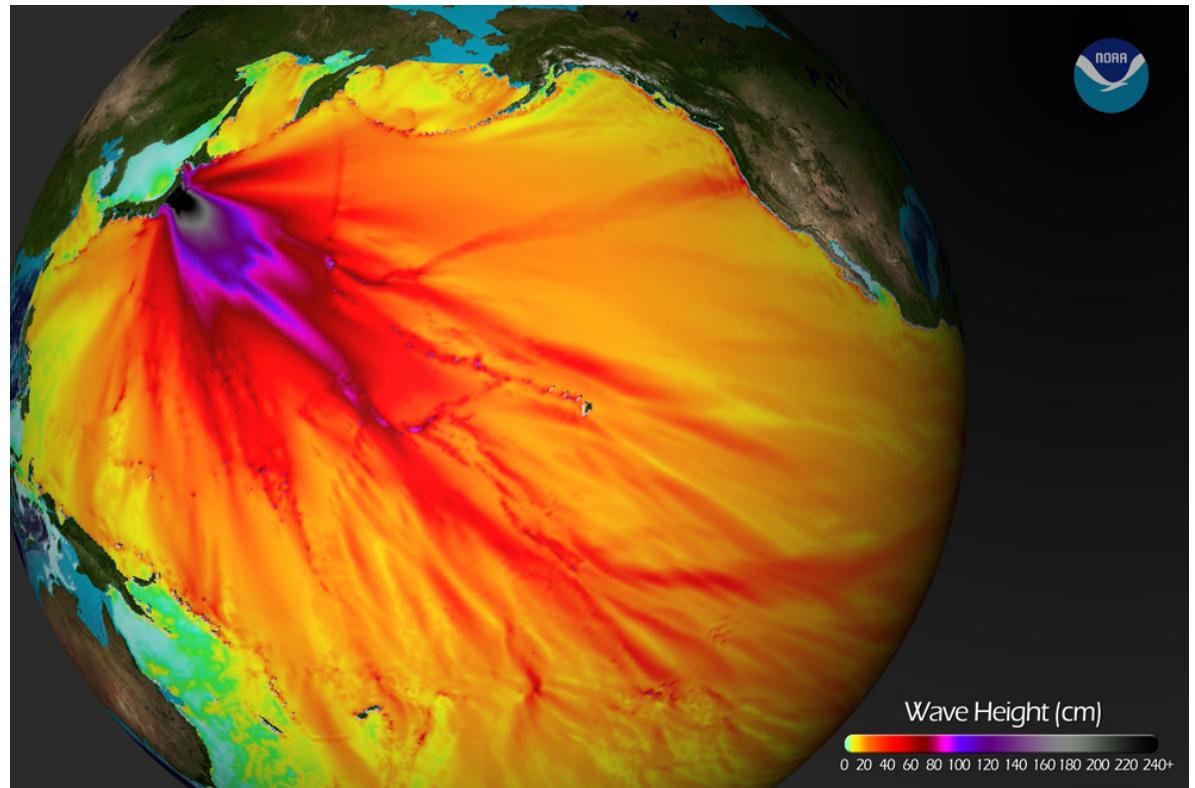
**Dados “abstratos”:**  
informações associadas a entidades de algum domínio não necessariamente espacial

# Caracterização de dados

- Várias classificações ...

Keller e Keller (1994)

- Campos escalares
- Dados nominais
- Campos direcionais
- Forma
- Posição
- Região



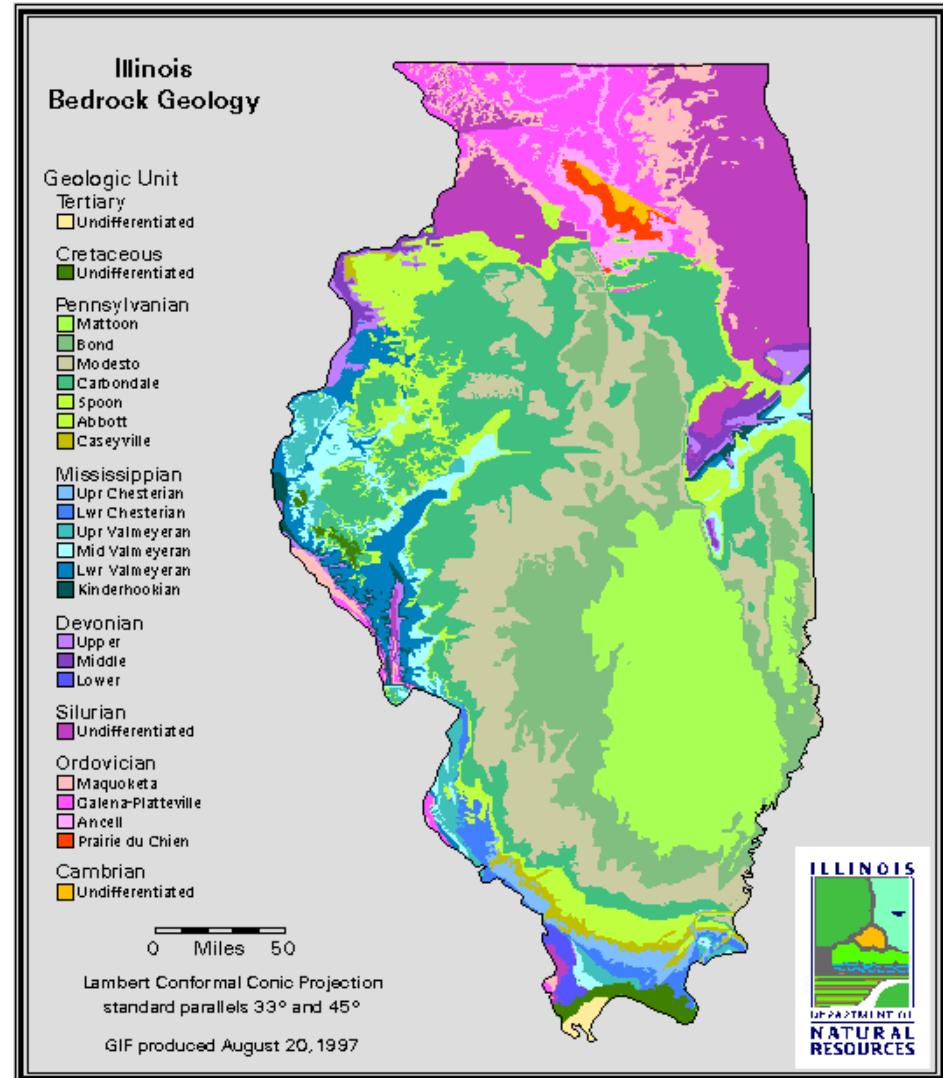
**Campo escalar:** mapa de altitude de ondas  
(tsunami de 11.03.2012)

# Caracterização de dados

- Várias classificações ...

Keller e Keller (1994)

- Campos escalares
- Dados nominais
- Campos direcionais
- Forma
- Posição
- Região



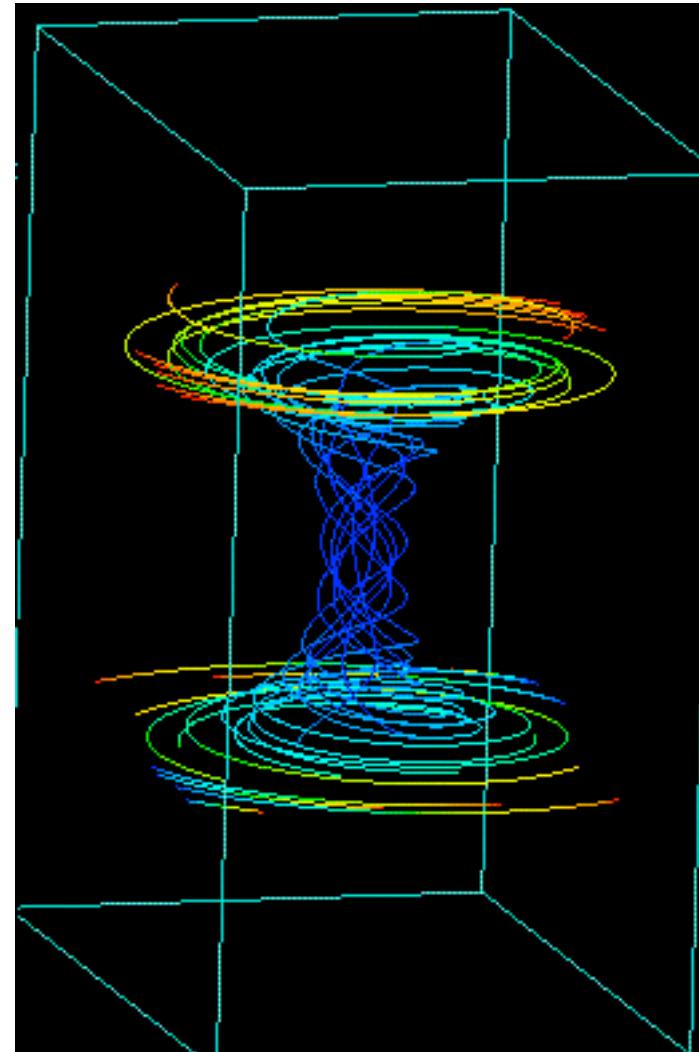
Dados nominais: tipos de rochas

# Caracterização de dados

- Várias classificações ...

Keller e Keller (1994)

- Campos escalares
- Dados nominais
- Campos direcionais
- Forma
- Posição
- Região

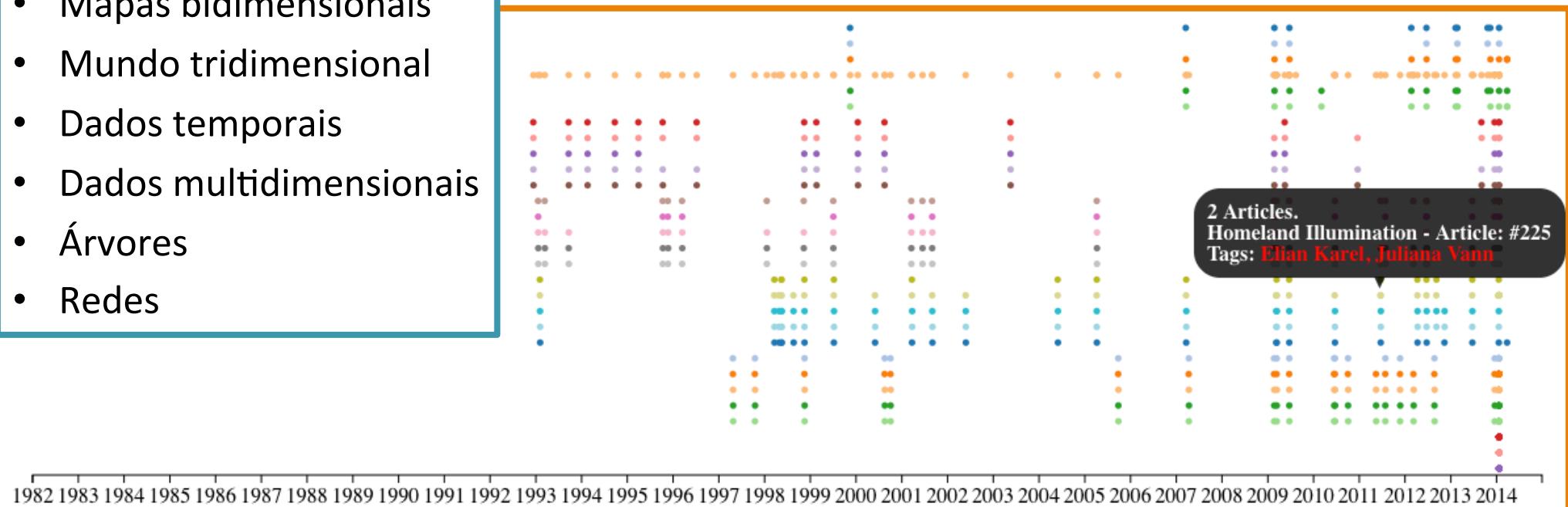


**Campo vetorial de velocidade (representada por cor)**  
<http://web.cs.wpi.edu/~matt/courses/cs563/talks/flowvis/flowvis.html>

# Caracterização de dados

Shneiderman (1996)

- Dados unidimensionais
- Mapas bidimensionais
- Mundo tridimensional
- Dados temporais
- Dados multidimensionais
- Árvores
- Redes



Dados unidimensionais e temporais

Exemplo: VAST 2014 mini-challenge 1

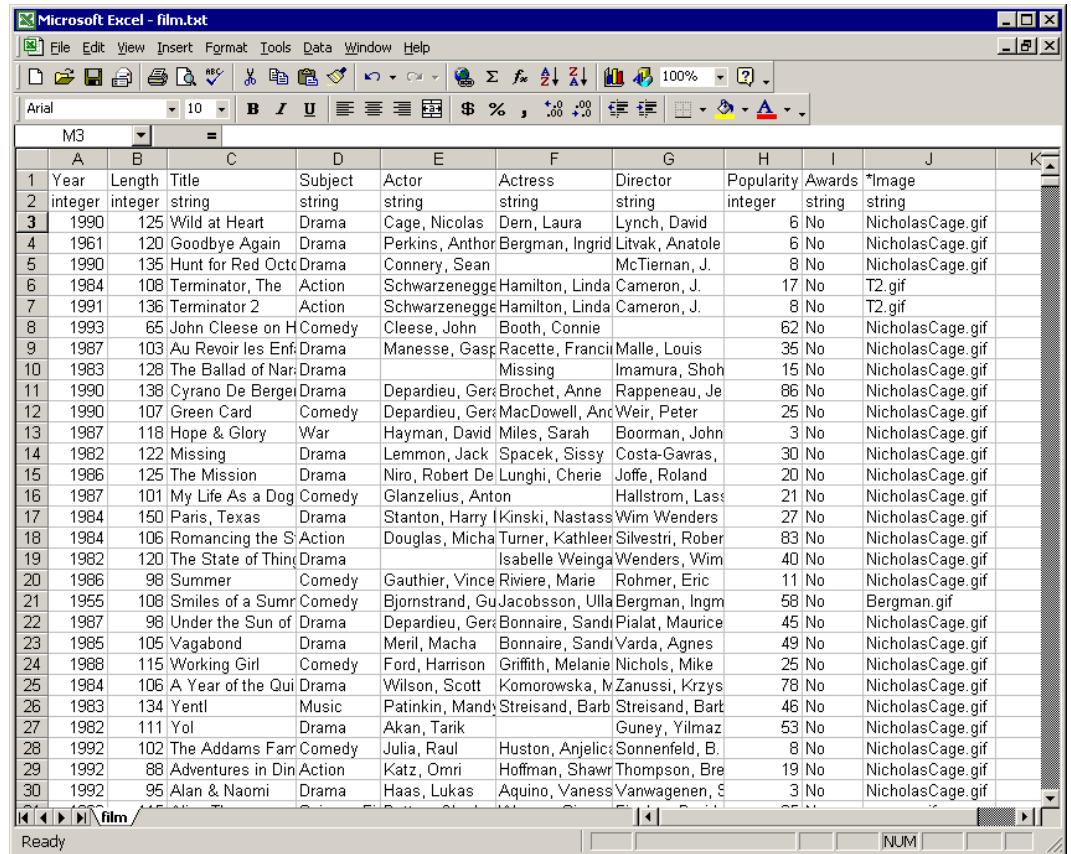
Shneiderman, Ben *The Eyes Have it: A Task by Data Type Taxonomy for Information Visualization*. 1996 IEEE Symposium on Visual Language, pp336-343

Guedes, L. et al. *Interactive Timeline Visualization of Documents*. SIBGRAPI 2014, Works in Progress (poster).

# Caracterização de dados

Shneiderman (1996)

- Dados unidimensionais
- Mapas bidimensionais
- Mundo tridimensional
- Dados temporais
- Dados multidimensionais
- Árvores
- Redes



The screenshot shows a Microsoft Excel spreadsheet titled "Microsoft Excel - film.txt". The data is presented in a tabular format with columns labeled A through K. The first few rows define the schema:

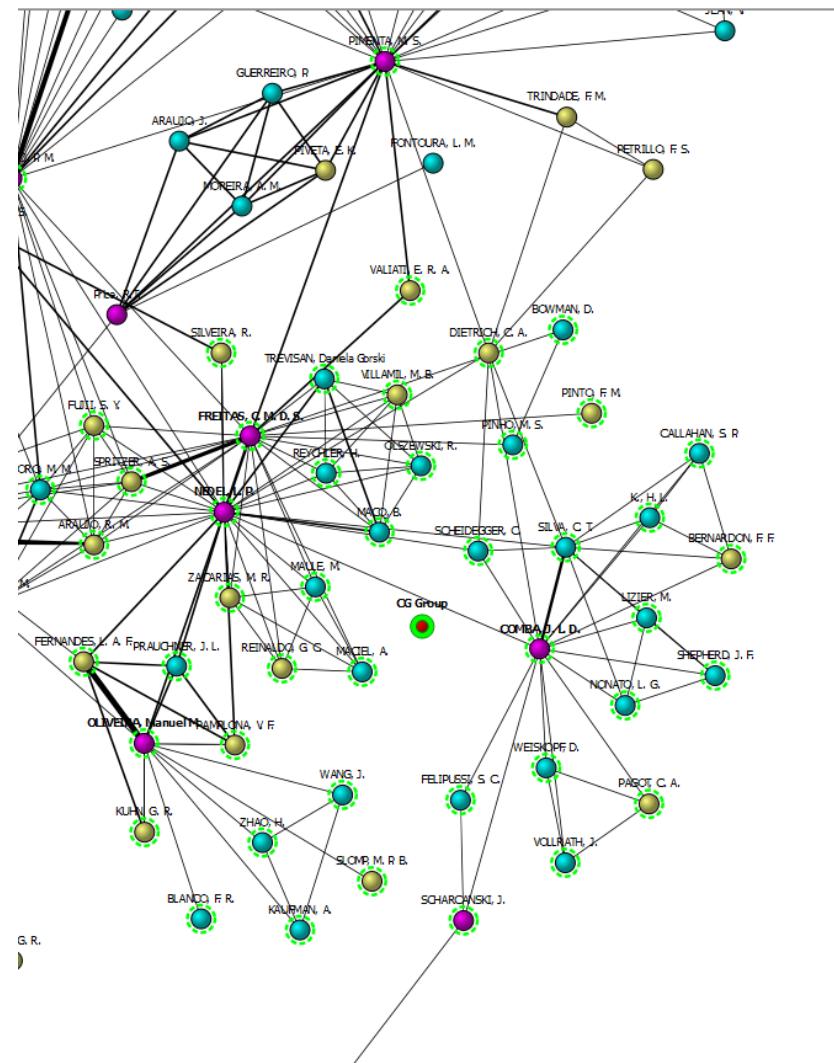
A	B	C	D	E	F	G	H	I	J	K
1	Year	Length	Title	Subject	Actor	Actress	Director	Popularity	Awards	*Image
2	integer	integer	string	string	string	string	string	integer	string	string
3	1990	125	Wild at Heart	Drama	Cage, Nicolas	Dern, Laura	Lynch, David	6	No	NicholasCage.gif
4	1961	120	Goodbye Again	Drama	Perkins, Anthony	Bergman, Ingrid	Litvak, Anatole	6	No	NicholasCage.gif
5	1990	135	Hunt for Red Octo	Drama	Connery, Sean		McTiernan, J.	8	No	NicholasCage.gif
6	1984	108	Terminator, The	Action	Schwarzenegge	Hamilton, Linda	Cameron, J.	17	No	T2.gif
7	1991	136	Terminator 2	Action	Schwarzenegge	Hamilton, Linda	Cameron, J.	8	No	T2.gif
8	1993	65	John Cleese on H	Comedy	Cleese, John	Booth, Connie		62	No	NicholasCage.gif
9	1987	103	Au Revoir les Enf	Drama	Manesse, Gaspard	Racette, Francine	Malle, Louis	35	No	NicholasCage.gif
10	1963	128	The Ballad of Nar	Drama			Imamura, Shoh	15	No	NicholasCage.gif
11	1990	138	Cyrano De Berger	Drama	Depardieu, Gerard	Brochet, Anne	Rappeneau, Je	86	No	NicholasCage.gif
12	1990	107	Green Card	Comedy	Depardieu, Gerard	MacDowell, Andie	Weir, Peter	25	No	NicholasCage.gif
13	1987	118	Hope & Glory	War	Hayman, David	Miles, Sarah	Boorman, John	3	No	NicholasCage.gif
14	1982	122	Missing	Drama	Lemmon, Jack	Spacek, Sissy	Costa-Gavras,	30	No	NicholasCage.gif
15	1986	125	The Mission	Drama	Niro, Robert De	Lunghi, Cherie	Joffe, Roland	20	No	NicholasCage.gif
16	1987	101	My Life As a Dog	Comedy	Glanzelius, Anton		Hallstrom, Lasse	21	No	NicholasCage.gif
17	1984	150	Paris, Texas	Drama	Stanton, Harry	Ikinski, Nastassja	Wim Wenders	27	No	NicholasCage.gif
18	1984	106	Romancing the S	Action	Douglas, Michael	Turner, Kathleen	Silvestri, Robert	83	No	NicholasCage.gif
19	1982	120	The State of Thing	Drama			Isabelle Weingard	40	No	NicholasCage.gif
20	1986	98	Summer	Comedy	Gauthier, Vince	Riviere, Marie	Rohmer, Eric	11	No	NicholasCage.gif
21	1955	108	Smiles of a Summ	Comedy	Bjornstrand, Gunnar	Jacobsson, Ulla	Bergman, Ingmar	58	No	Bergman.gif
22	1987	98	Under the Sun of	Drama	Depardieu, Gerard	Bonnaire, Sandi	Pialat, Maurice	45	No	NicholasCage.gif
23	1985	105	Vagabond	Drama	Meril, Macha	Bonnaire, Sandi	Varda, Agnes	49	No	NicholasCage.gif
24	1988	115	Working Girl	Comedy	Ford, Harrison	Griffith, Melanie	Nichols, Mike	25	No	NicholasCage.gif
25	1984	106	A Year of the Qui	Drama	Wilson, Scott	Komorowska, M	Zanussi, Krzysztof	78	No	NicholasCage.gif
26	1983	134	Yentl	Music	Patinkin, Mandy	Streisand, Barbra	Streisand, Barbra	46	No	NicholasCage.gif
27	1982	111	Yol	Drama	Akan, Tarik		Guney, Yilmaz	53	No	NicholasCage.gif
28	1992	102	The Addams Fam	Comedy	Julia, Raul	Huston, Anjelica	Sonnenfeld, Barry	8	No	NicholasCage.gif
29	1992	88	Adventures in Din	Action	Katz, Omri	Hoffman, Shaw	Thompson, Bruce	19	No	NicholasCage.gif
30	1992	95	Alan & Naomi	Drama	Haas, Lukas	Aquino, Vanessa	Vanwagenen, Sam	3	No	NicholasCage.gif

Dados multidimensionais

# Caracterização de dados

Shneiderman (1996)

- Dados unidimensionais
- Mapas bidimensionais
- Mundo tridimensional
- Dados temporais
- Dados multidimensionais
- Árvores
- Redes



Shneiderman, Ben *The Eyes Have it: A Task by Data Type*  
*Taxonomy for Information Visualization*. 1996 IEEE Symposium  
on Visual Language, pp336-343

**Rede** de co-autorias do grupo de  
CG (Spritzer e Freitas 2012)

# Caracterização de dados

Keim (2002)

- Dados unidimensionais
- Dados bidimensionais
- Dados multidimensionais
- Texto e hipertexto
- Hierarquias e grafos
- Algoritmos e software



Dados unidimensionais, Software

# Caracterização de dados

---

Ward et al. (2010)

- Não é uma taxonomia, mas uma caracterização sob diversos aspectos
- Dados ordinais (numéricos)
  - Binários, discretos, contínuos
- Dados nominais
  - Categóricos, categórico-ordenado, arbitrário
- Dados estruturados
  - Escalares, vetoriais, tensoriais
- Geometria e grids

# Caracterização de dados

---

- Classes de dados
- Tipos de dados
- Natureza do domínio
- Dimensão do domínio

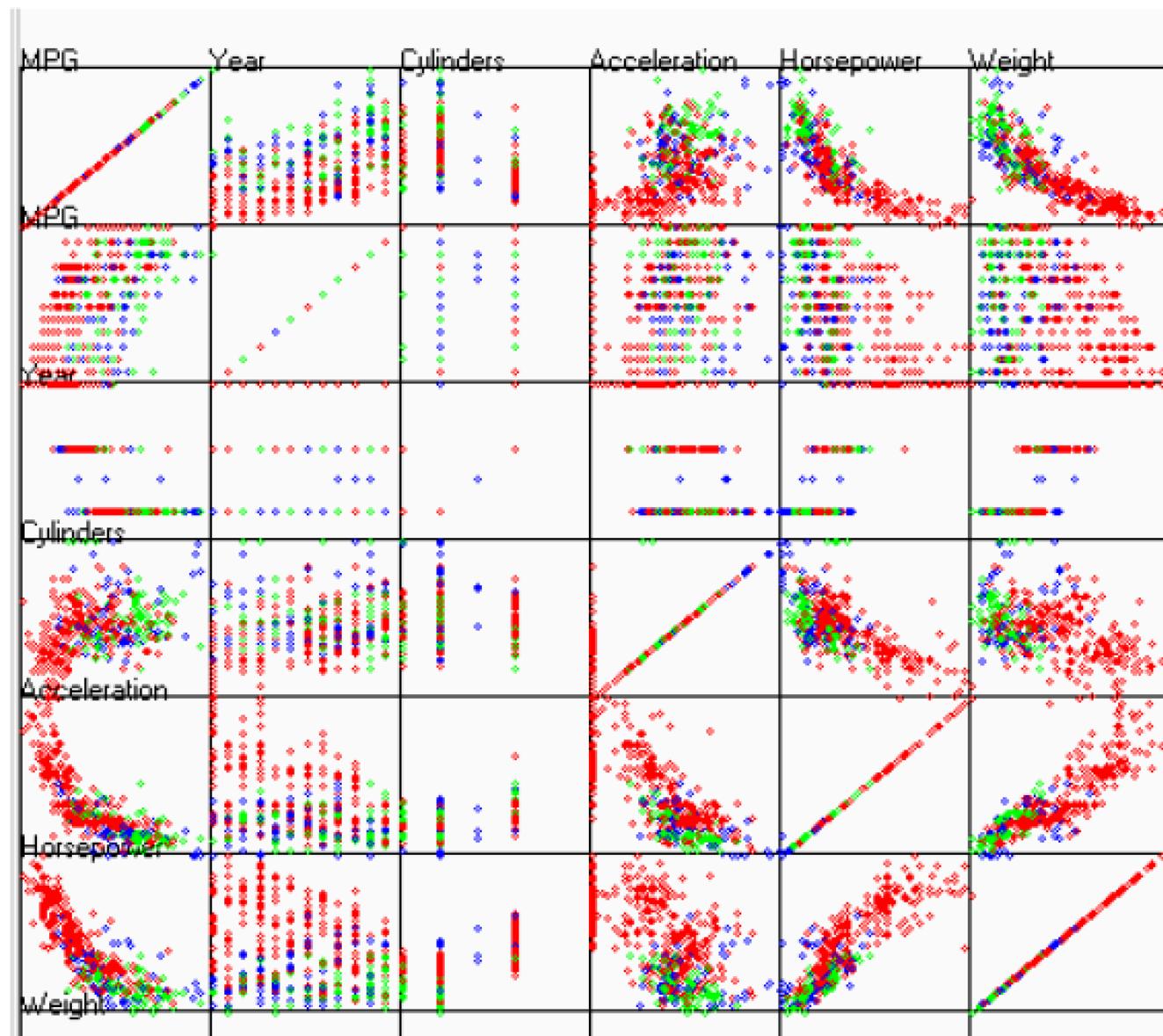
# Classes de dados

---

- Dado = característica de uma entidade
  - exemplo: entidade = carro
  - informações sobre carros
    1. mpg: continuous
    2. cylinders: multi-valued discrete
    3. horsepower: continuous
    4. weight: continuous
    5. acceleration: continuous
    6. model year: multi-valued discrete
    7. origin: multi-valued discrete
    8. car name: string (unique for each instance)
  - ....

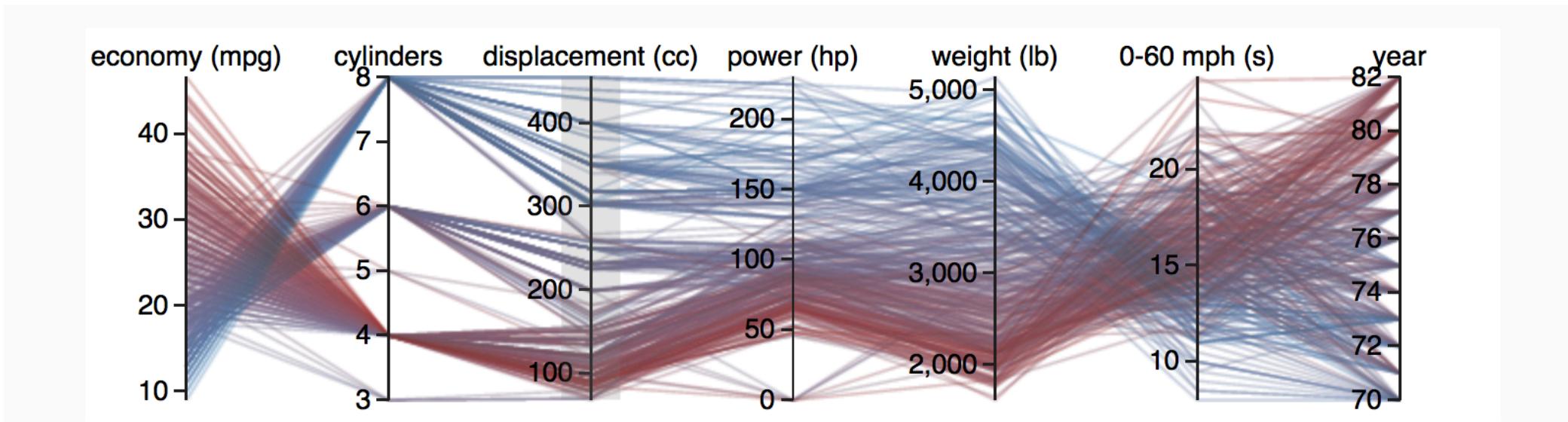
<http://archive.ics.uci.edu/ml/datasets/Auto+MPG>

# Matriz de scatterplots



# Coordenadas paralelas

<https://syntagmatic.github.io/parallel-coordinates/>



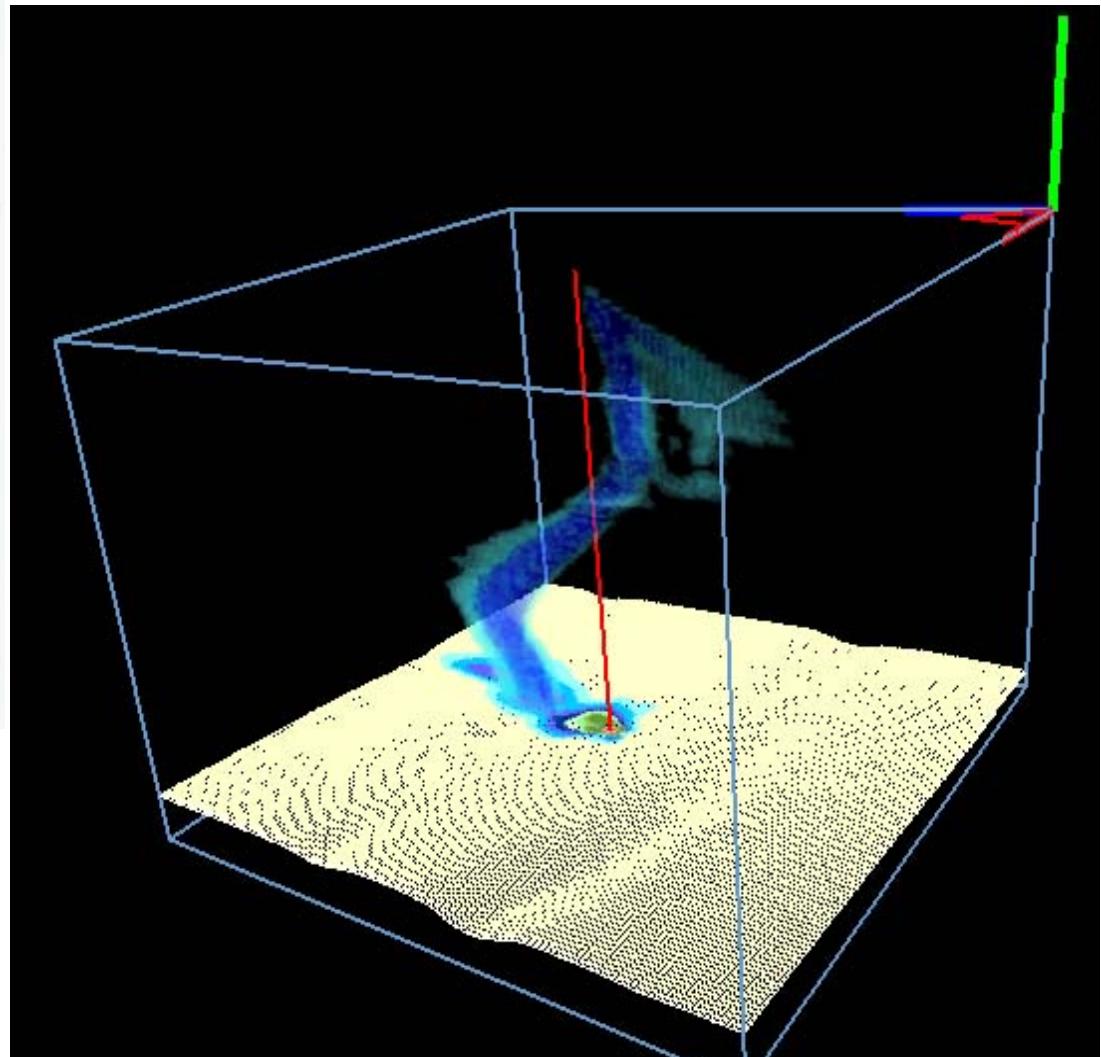
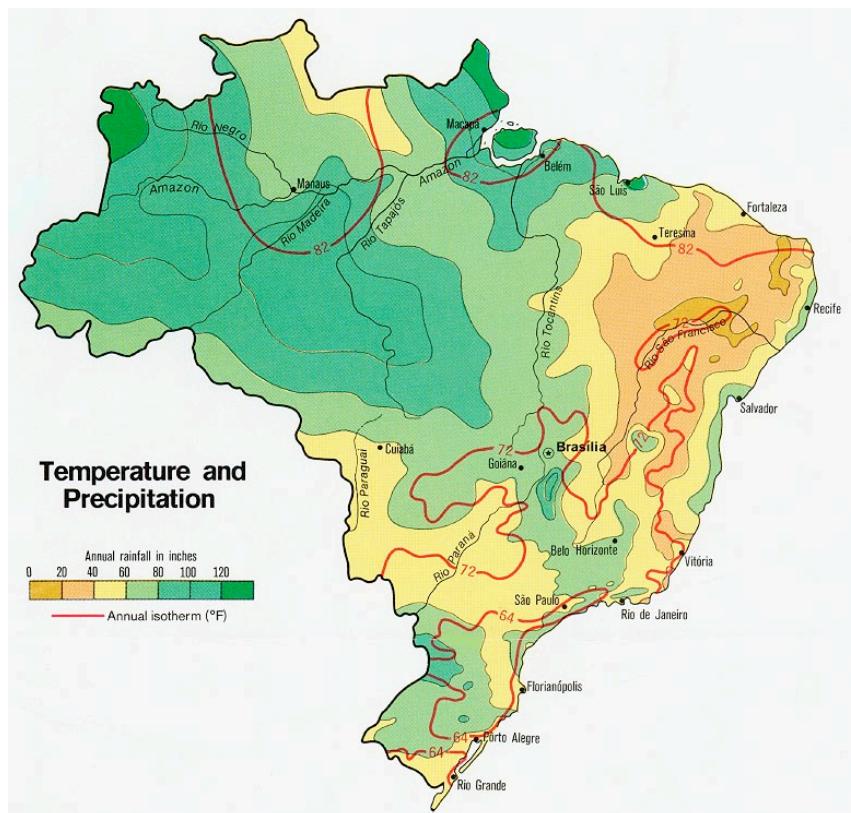
displacement = *total air displaced by the pistons in all of an engine's cylinders and is expressed in liters, cubic centimeters or cubic inches*

# Classes de dados

---

- Dado = grandeza escalar
  - entidade = região do espaço
  - exemplo: medidas físicas nas posições do espaço
    - Temperatura
    - Pressão
  - Cada dado corresponde ao valor escalar de uma função naquele ponto
    - Campo escalar

# Grandeza escalar



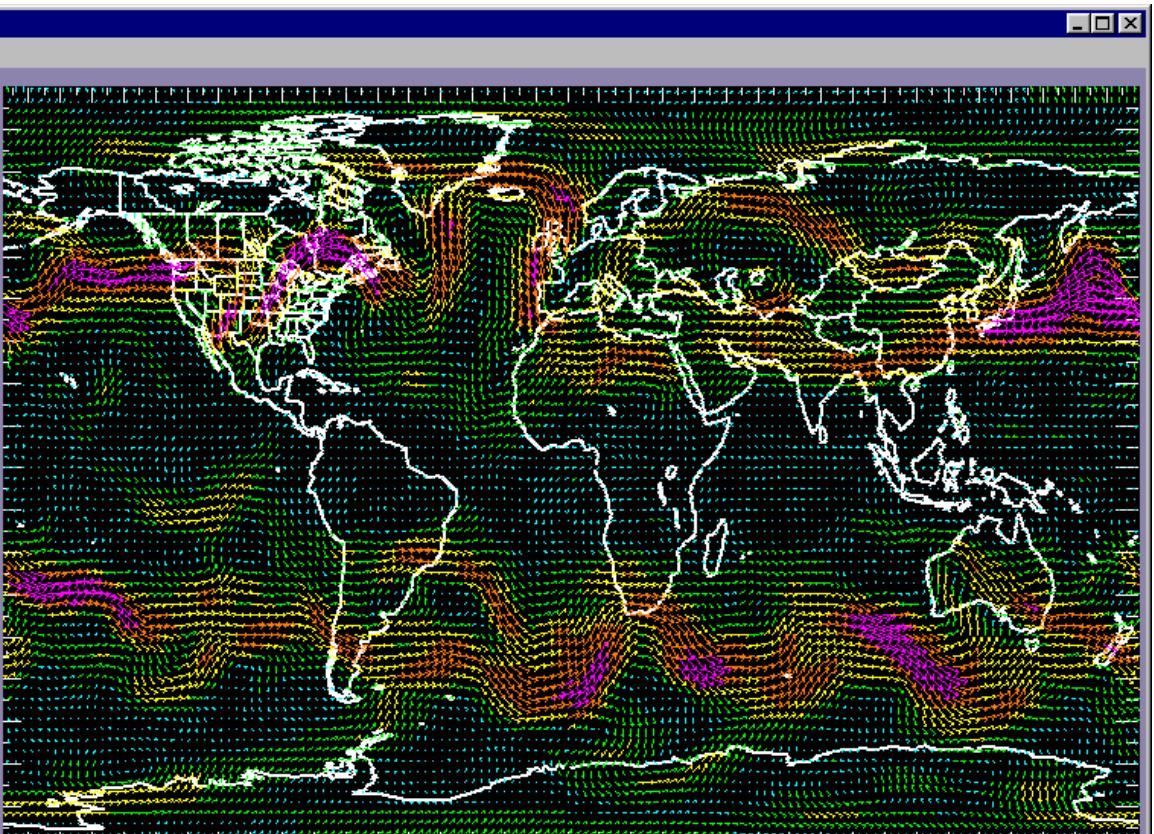
Densidade de material de perfuração de poço de petróleo em alto mar descartado num certo intervalo de tempo (Binotto et al., 2003)

# Classes de dados

---

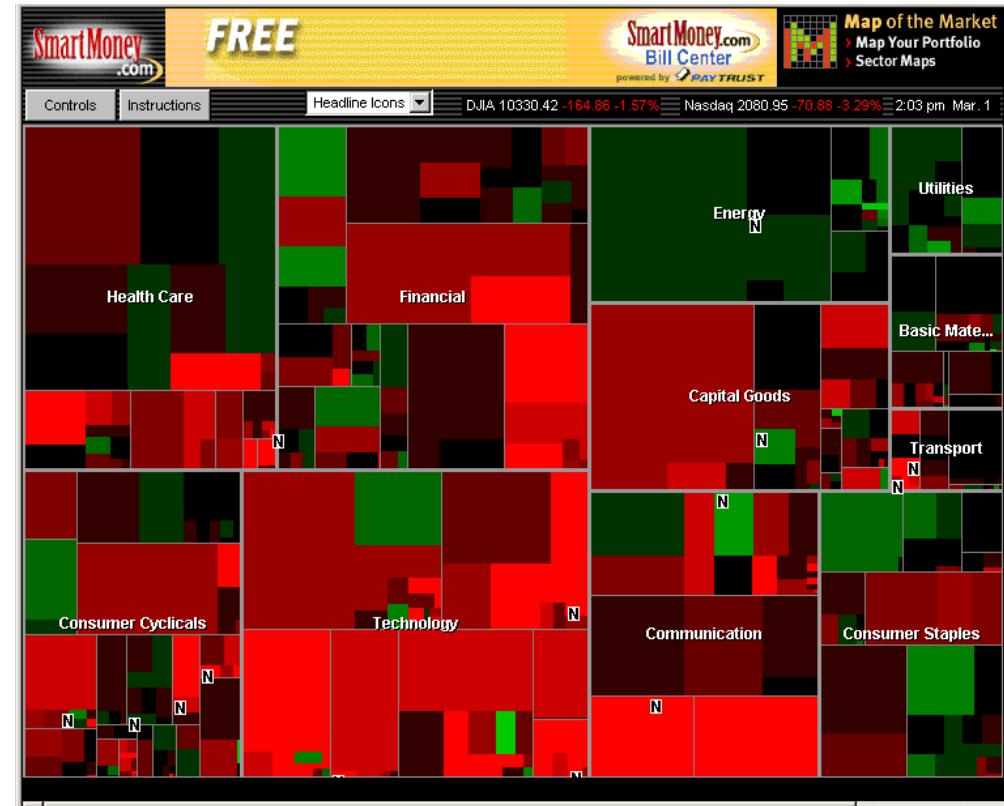
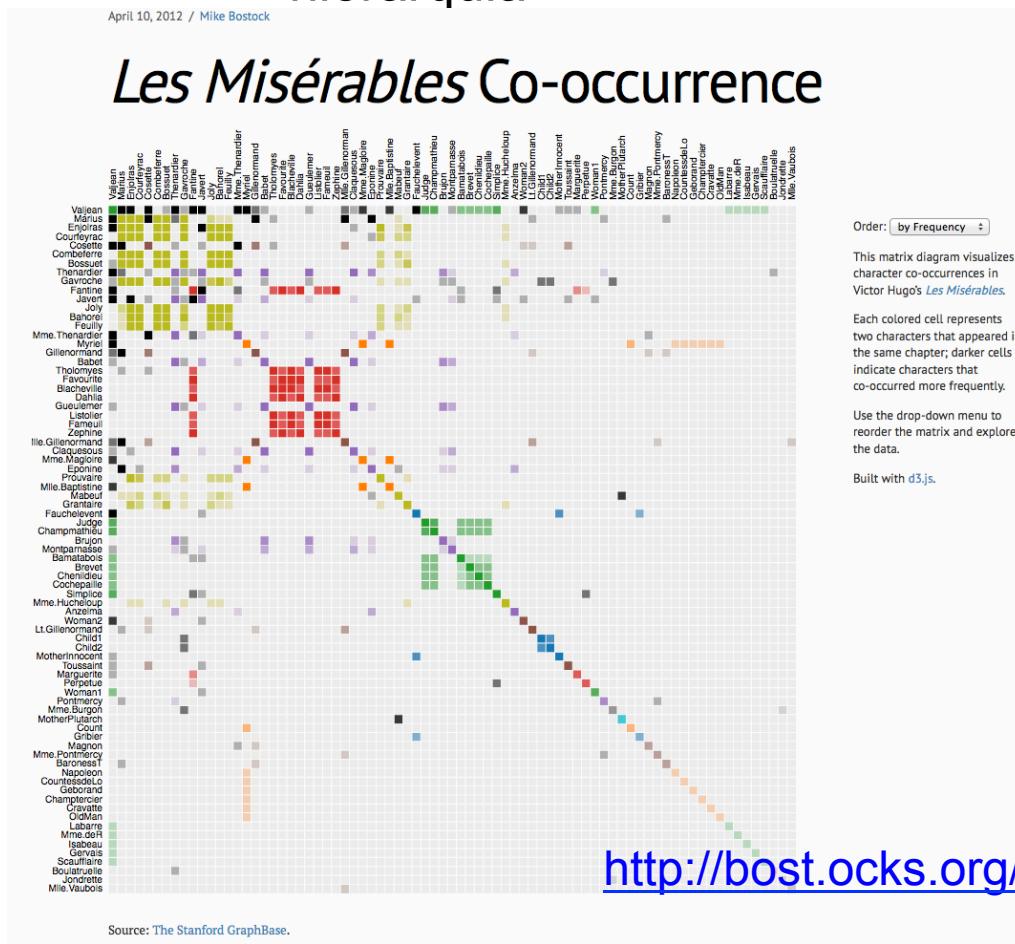
- Dado = grandeza vetorial
  - entidade = região do espaço
  - exemplo: medidas físicas nas posições do espaço
    - Fluxo de vento (magnitude, direção, sentido)
  - Cada dado corresponde ao valor vetorial de uma função naquele ponto
    - Campo vetorial

# Grandeza vetorial



# Classes de dados

- Dado = relacionamento
  - associação/ligação
  - hierarquia



<http://www.smartmoney.com/marketmap/>  
<https://finviz.com/map.ashx>  
<http://www.bewitched.com/marketmap.html>

# Classes de dados

---

- Característica
  - Dados multivariados, multidimensionais
- Campo escalar
  - Dados correspondem a uma função escalar definida num domínio espacial
- Campo vetorial, campo tensorial
  - ...
- Relacionamento
  - O conjunto de dados descreve relacionamentos entre entidades

# Caracterização dos dados

---

- Classes de dados
- Tipos de dados
- Natureza do domínio
- Dimensão do domínio

# Tipos de dados

- Quantitativos
- Categóricos
- Ordenados
- Relações

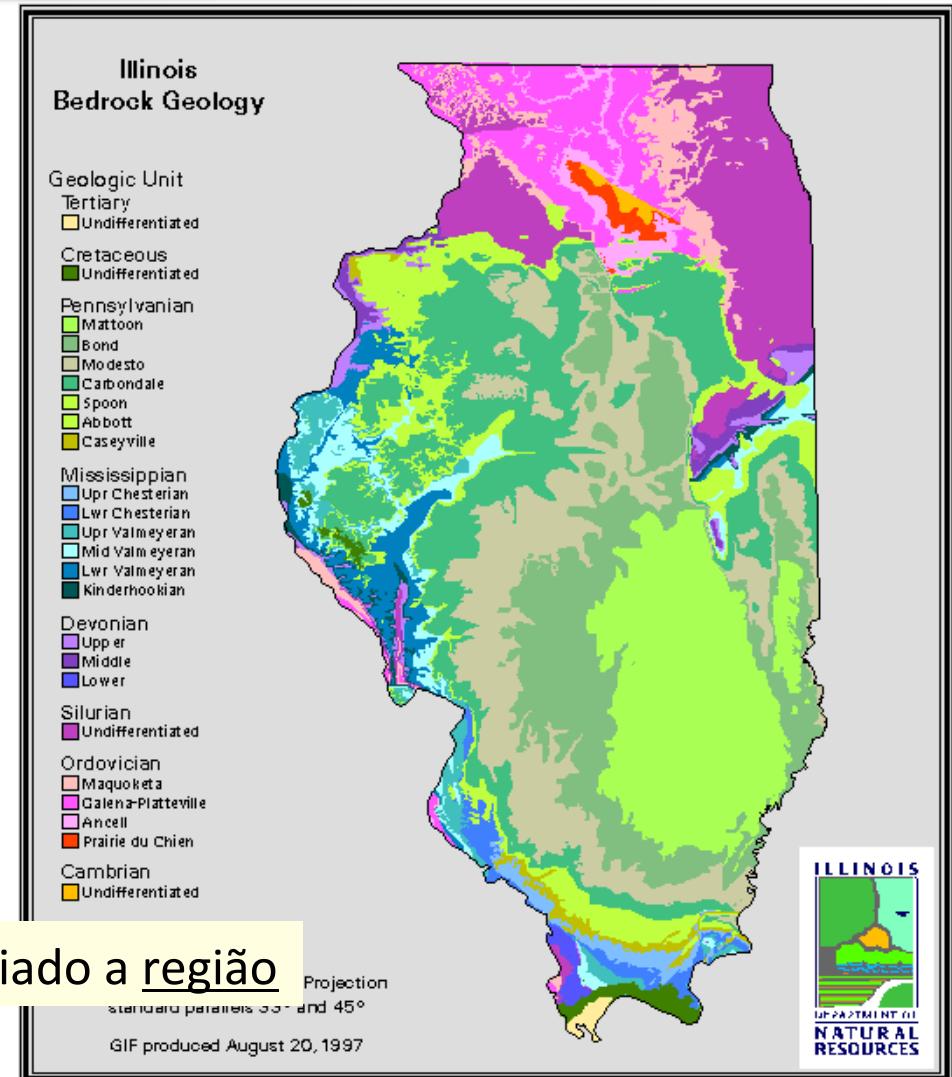
The screenshot shows a Microsoft Excel window titled "Microsoft Excel - film.txt". The window includes a menu bar with File, Edit, View, Insert, Format, Tools, Data, Window, and Help. Below the menu is a toolbar with various icons for file operations, cell selection, and data manipulation. The main area displays a table with data in rows and columns labeled A through K. The first few rows define the schema:

	A	B	C	D	E	F	G	H	I	J	K
1	Year	Length	Title	Subject	Actor	Actress	Director	Popularity	Awards	Image	
2	integer	integer	string	string	string	string	string	integer	string	string	
3	1990	125	Wild at Heart	Drama	Cage, Nicolas	Dem, Laura	Lynch, David	6	No	NicholasCage.gif	

The data continues with many more rows, each representing a movie entry with fields such as Year, Length, Title, Subject, Actor, Actress, Director, Popularity, Awards, and Image path.

# Tipos de dados x classes de dados

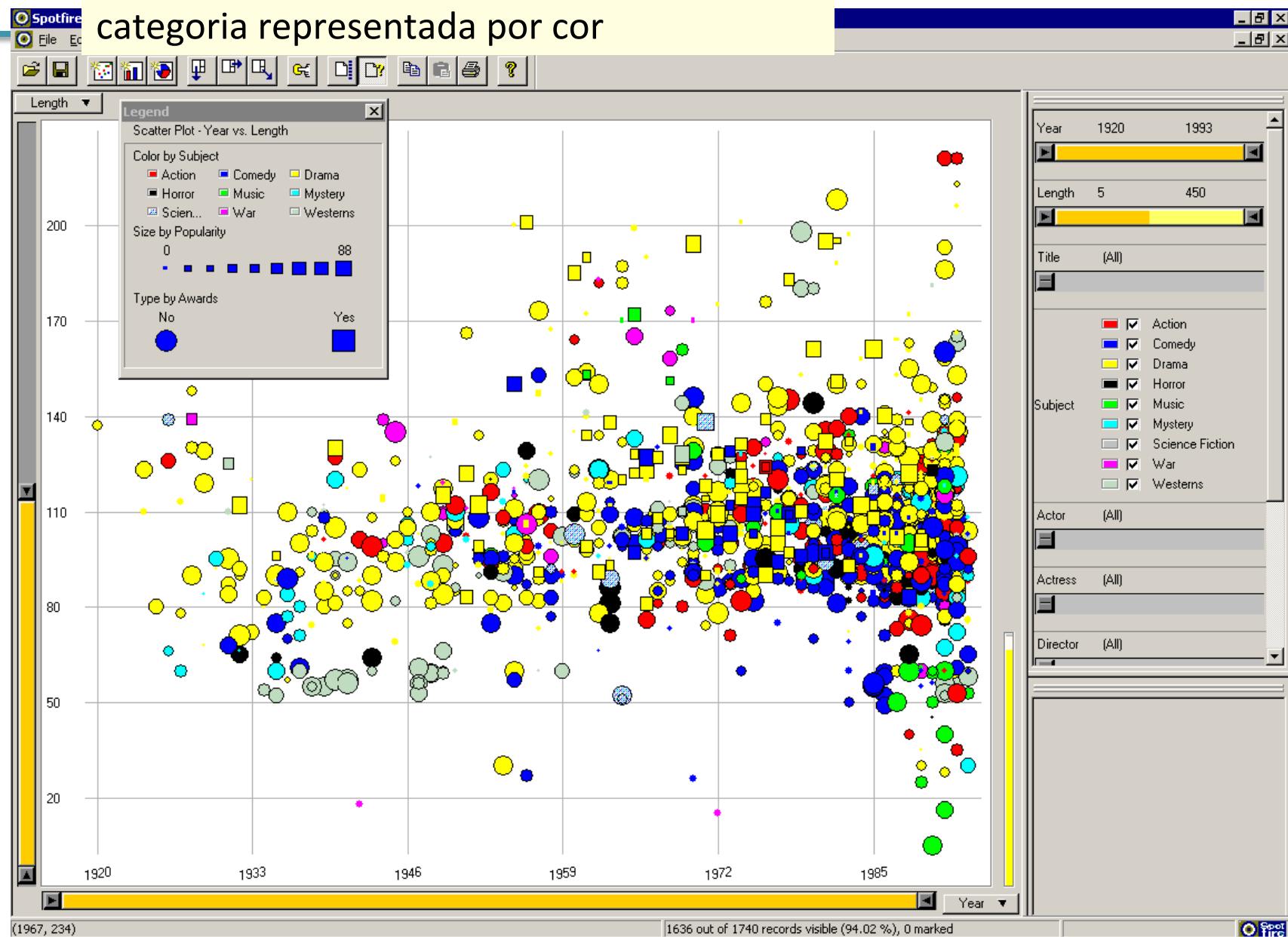
- Características
  - Numéricos, quantitativos ou ordenados
  - Categóricos (tipicamente membros de enumeração)



Tipo de rocha associado a região

Spotfire

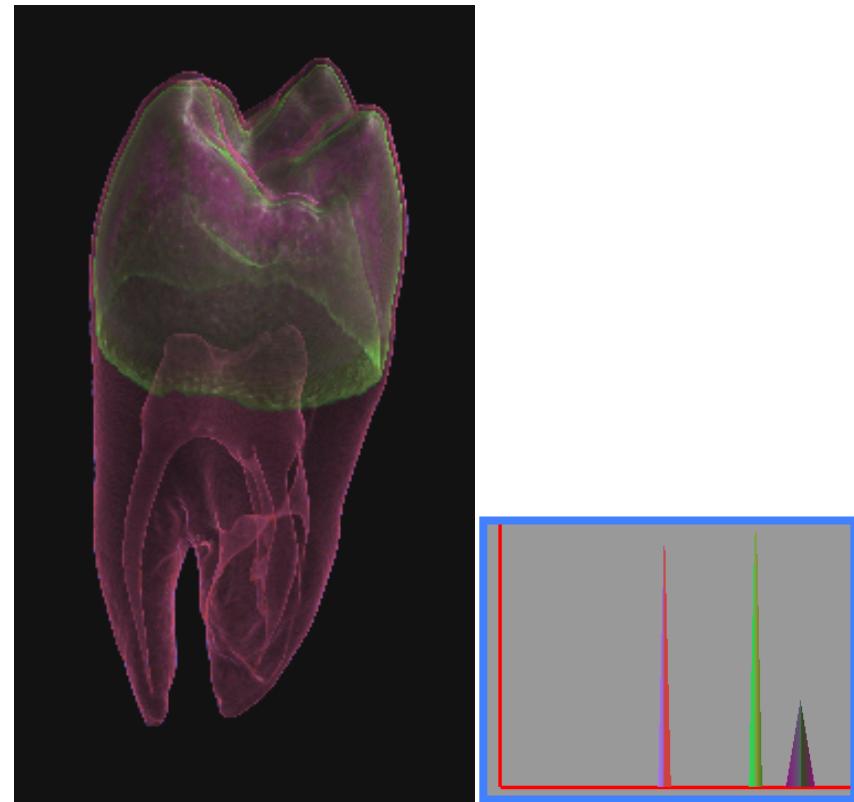
## Tipo de filme (representado por símbolo): categoria representada por cor



# Tipos de dados (escalares)

---

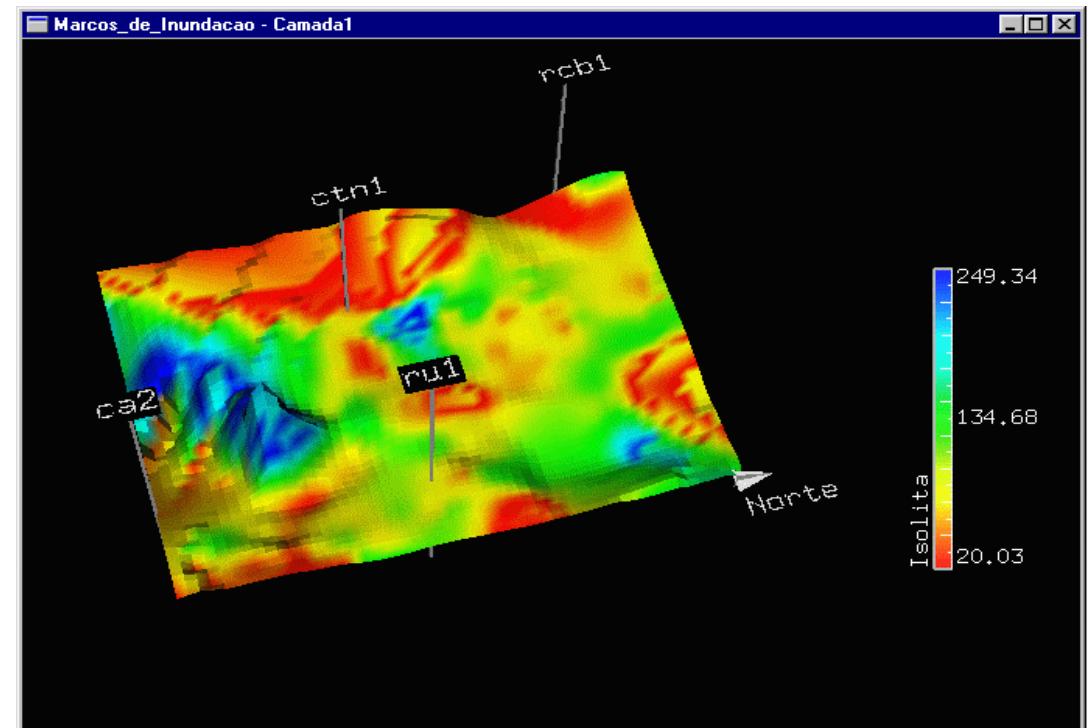
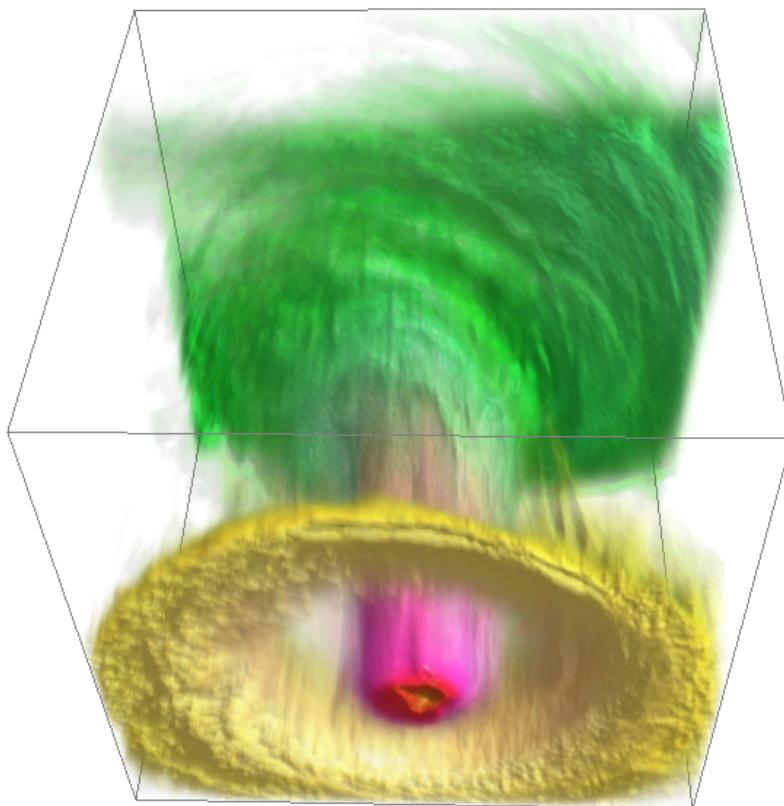
- Características
  - Numéricos, quantitativos ou ordenados
  - Categóricos (tipicamente membros de enumeração)
- Grandezas escalares, vetoriais, tensoriais
  - Numéricos, quantitativos



*Campo escalar*  
(UFRGS – Pinto e Freitas, 2006)

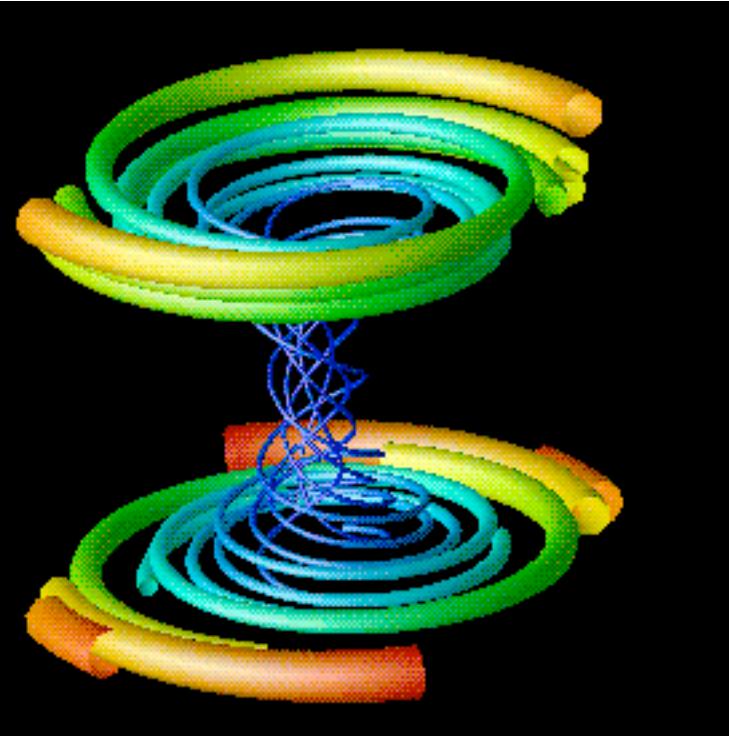
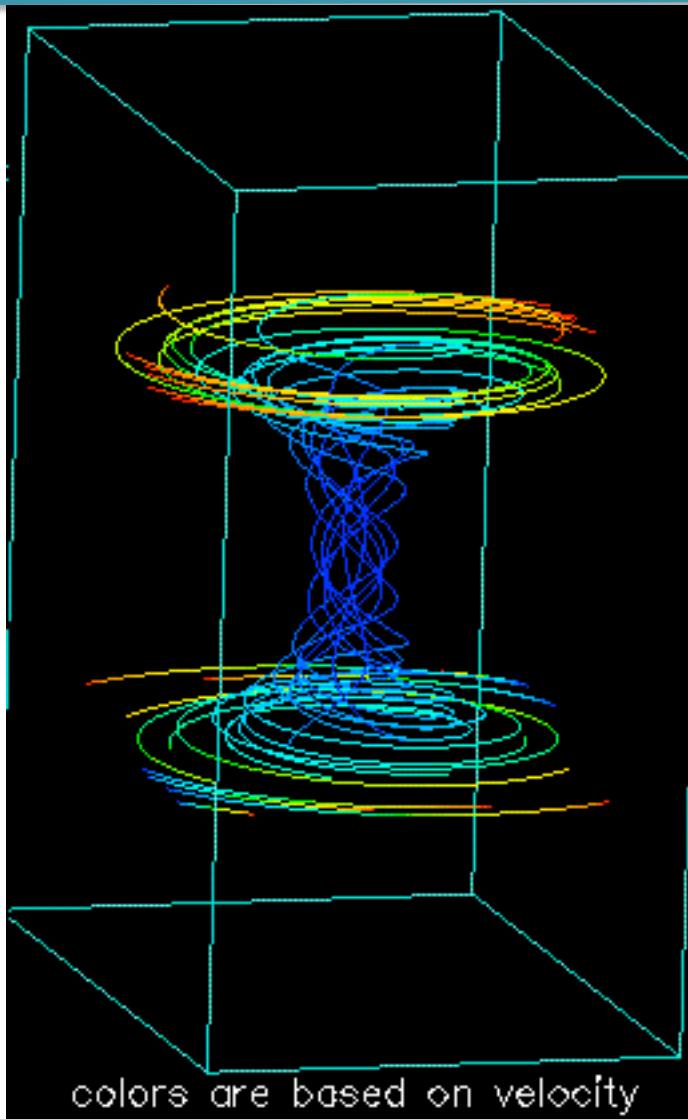
# Tipos de dados (escalares)

Pressão, temperatura e magnitude da velocidade do vento  
(UFRGS – Pinto e Freitas, 2007)



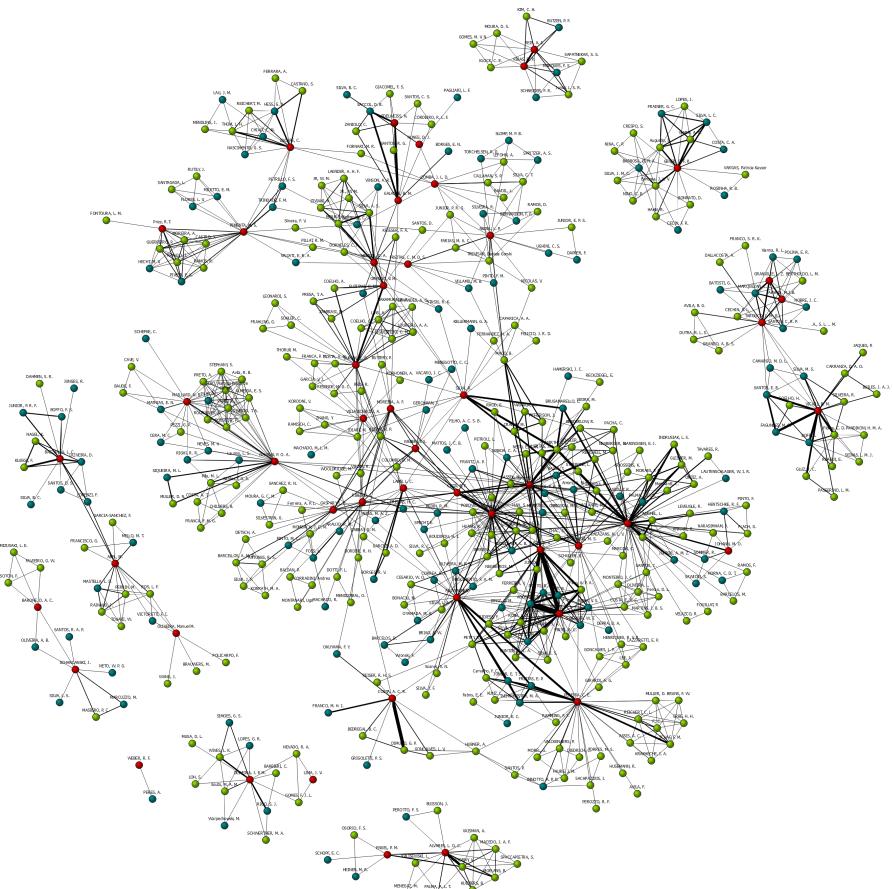
Basso e Freitas 1998

# Tipos de dados (vetoriais)



Imagens geradas com IBM Visualization Data Explorer obtidas de Ward (1999)  
<http://web.cs.wpi.edu/~matt/courses/cs563/talks/flowvis/flowvis.html>

# Tipos de dados (relacionamentos)



- 2007 UFRGS Informatics Institute publications
  - **474 Nodes**
    - Professors, students and external collaborators
  - **1252 Edges**
    - Proceedings, journals and books

*Spritzer and Freitas (2012)*

# Caracterização de dados

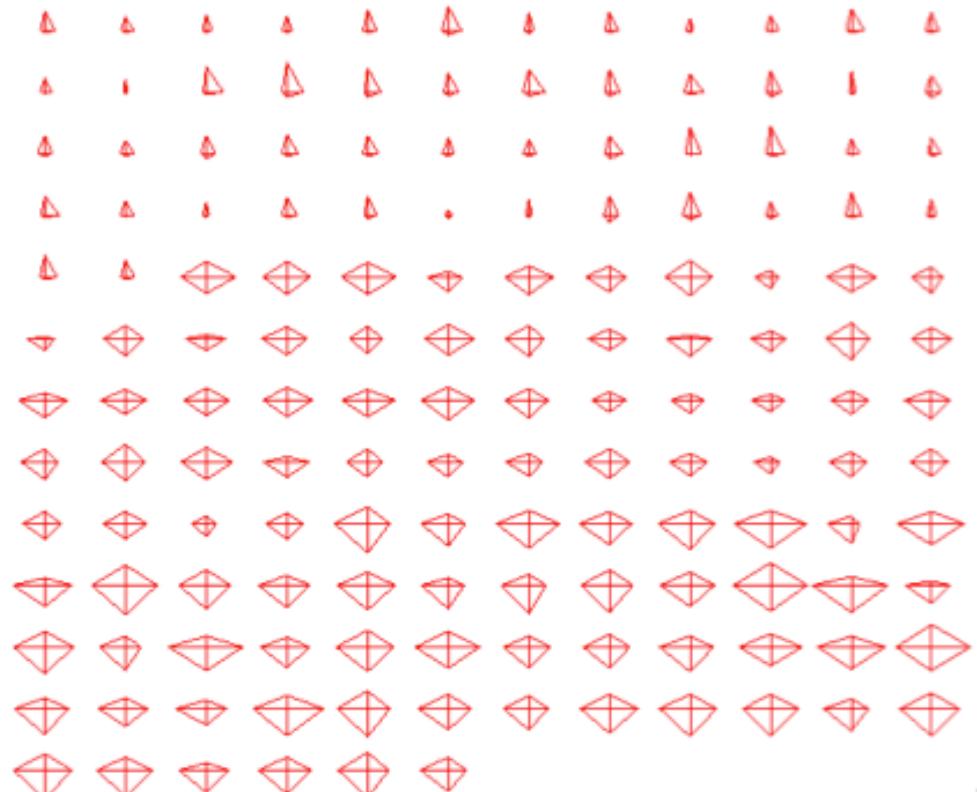
---

- Classes de dados
- Tipos de dados
- **Natureza do domínio**
- Dimensão do domínio

# Natureza do domínio

---

- Discreto
  - Conjunto de entidades
  - Cada entidade corresponde a um “ponto” no espaço de amostragem
- Exemplo: Iris dataset
  1. sepal length in cm
  2. sepal width in cm
  3. petal length in cm
  4. petal width in cm
  5. class: Iris Setosa, Iris Versicolour, Iris Virginica
- Cada “glifo” representa uma flor do conjunto

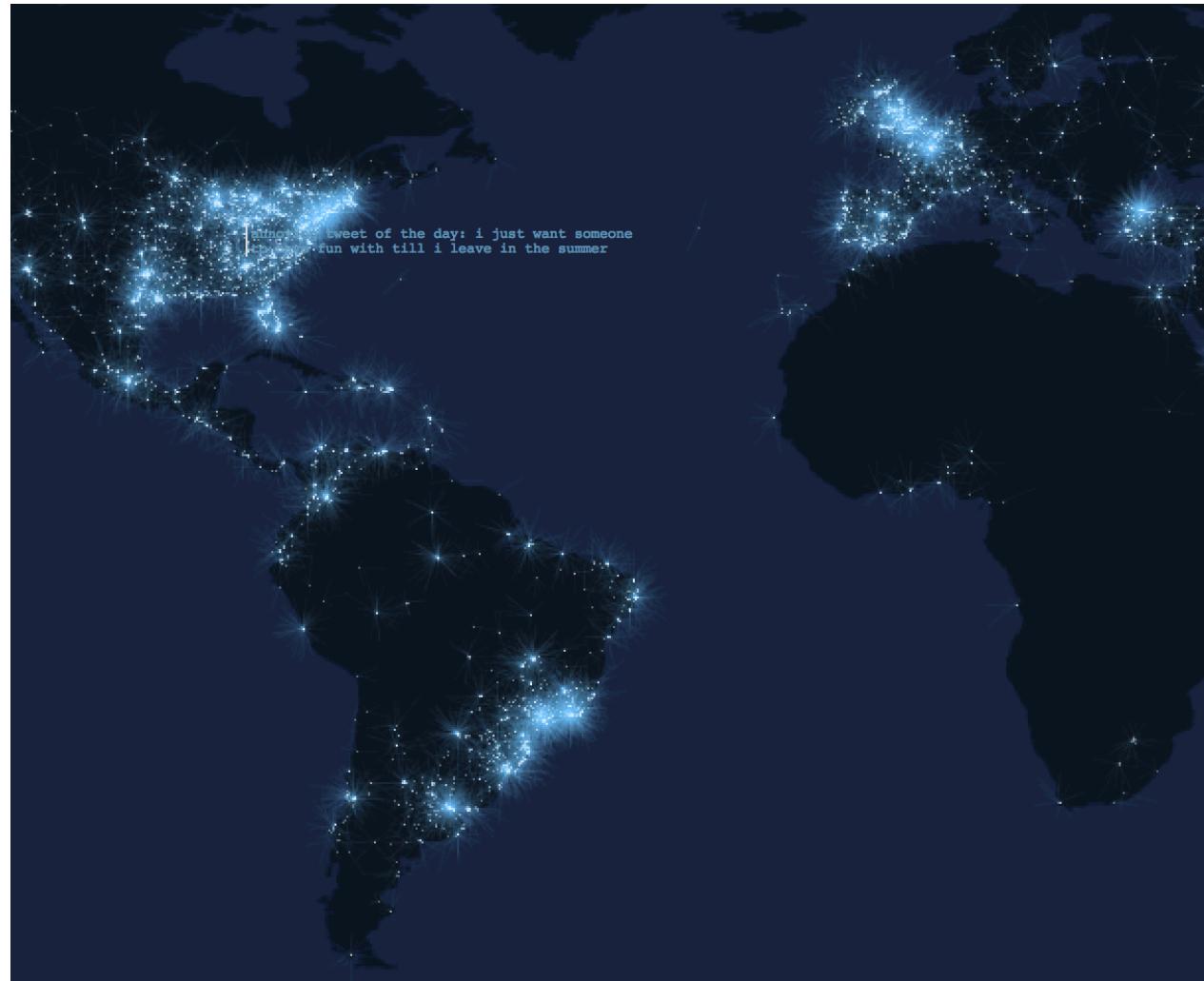


<http://archive.ics.uci.edu/ml/datasets/Iris>

# Natureza do domínio

---

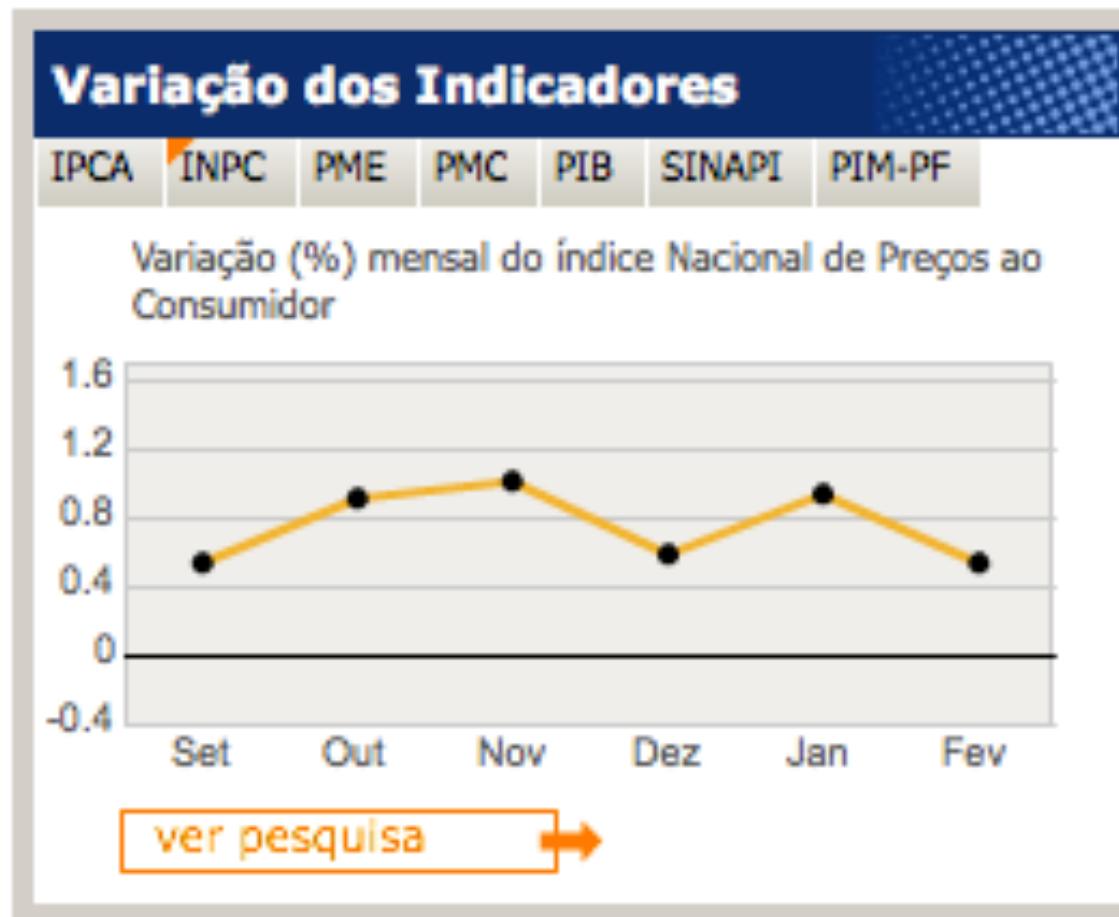
- Contínuo
- Exemplo
  - tempo



<http://tweetping.net/#>

# Natureza do domínio

- Contínuo-discretizado



# Caracterização de dados

---

- Classes de dados
- Tipos de dados
- Natureza do domínio
- **Dimensão do domínio**

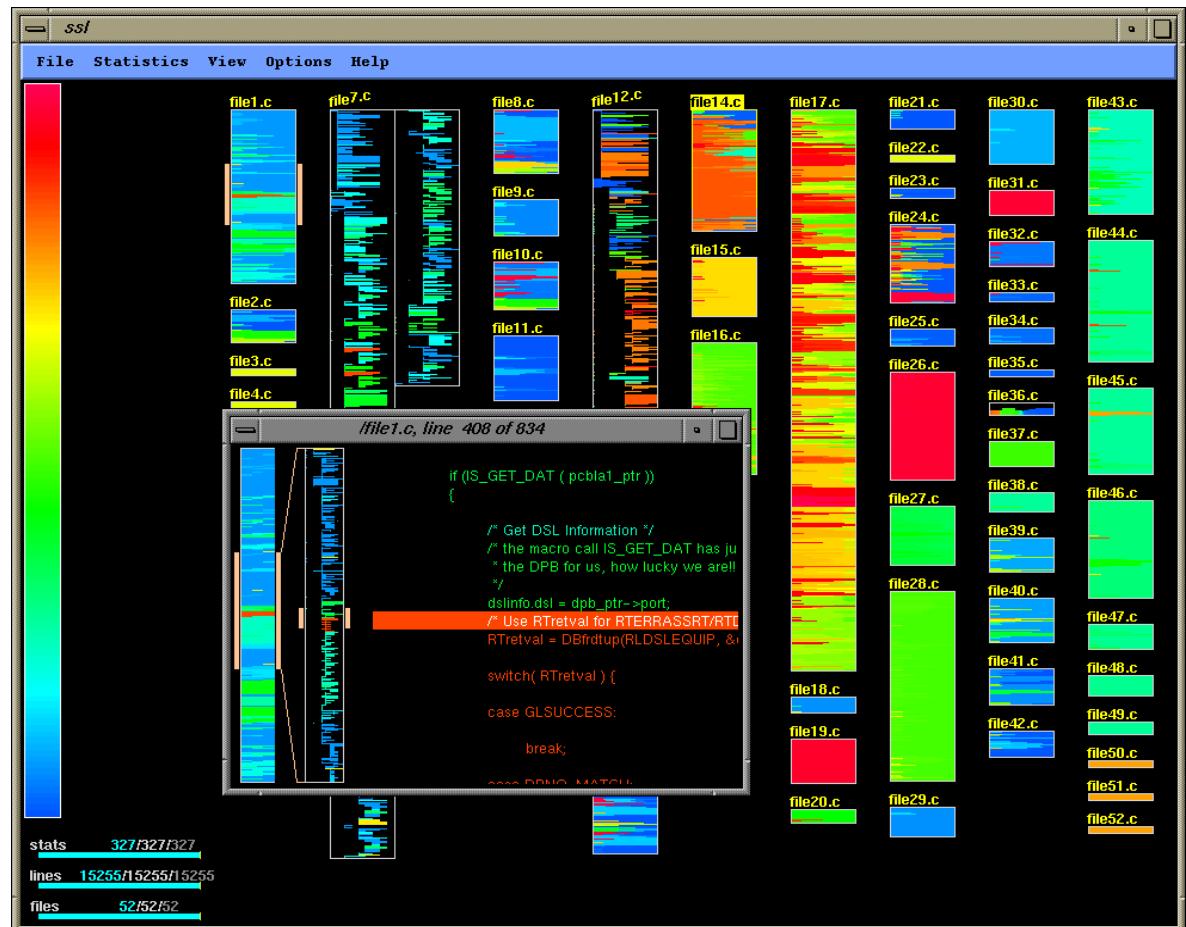
# Dimensão do domínio

---

- 1D
  - Informação temporal (fatos históricos)
- 2D
  - Dados associados a regiões planas (mapas)
- 3D
  - Dados associados a regiões no espaço (volumes)
- Multidimensional
  - Dados tabulares

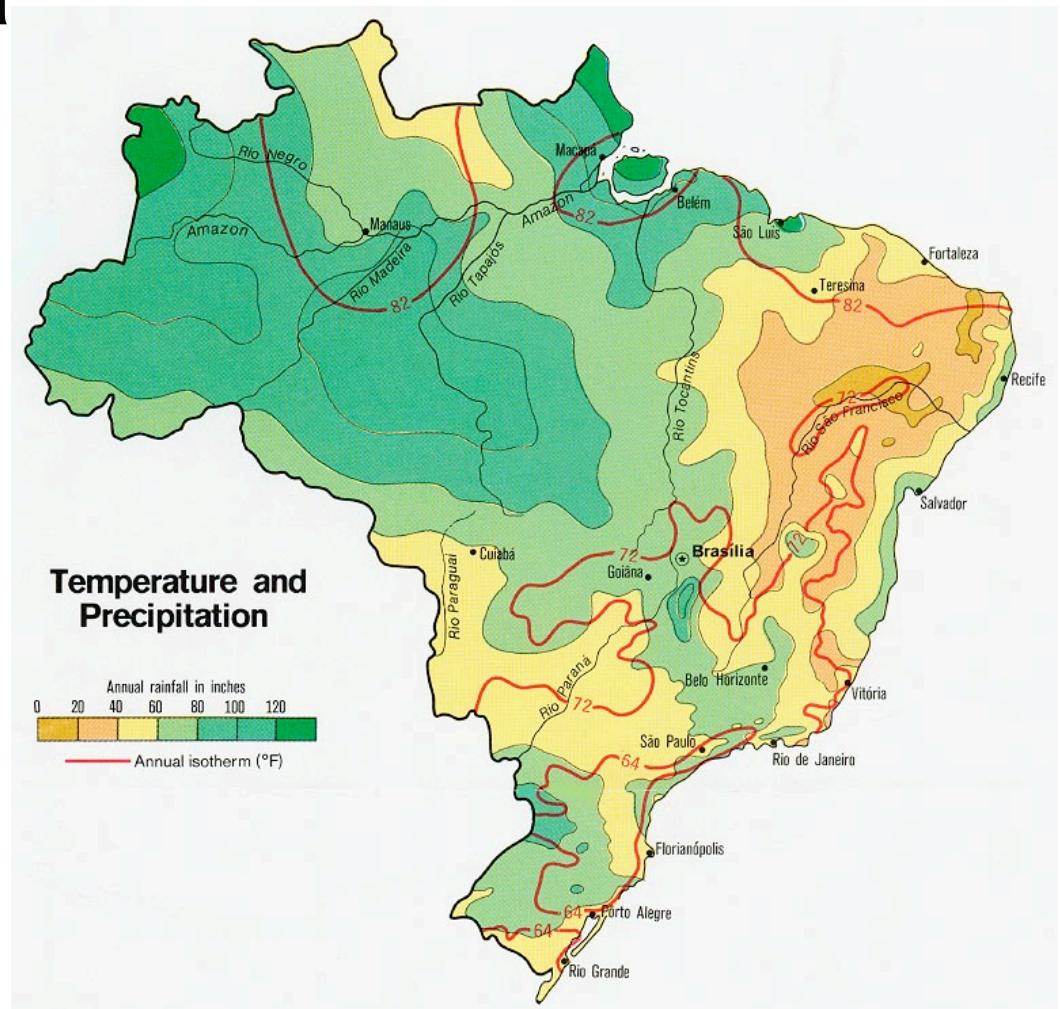
# Dados 1D

- Textos
- Informação temporal
- Exemplo:
  - Seesoft
    - Informação temporal indicada pela cor



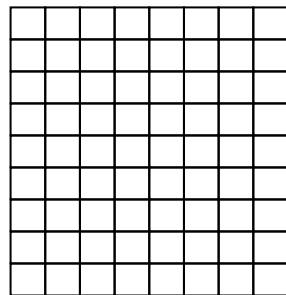
# Dados 2D

- Informação associada a posições, regiões num plano

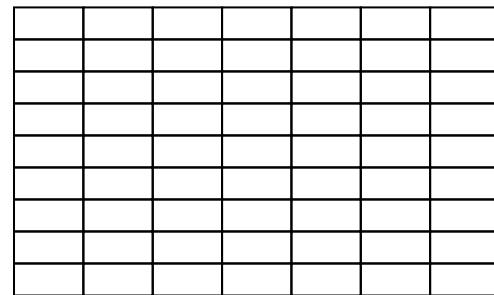


# Representação do espaço (2D ou 3D)

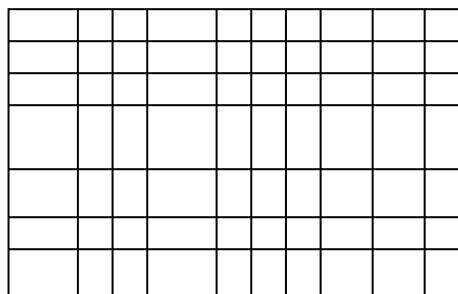
---



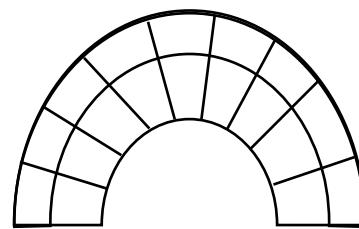
cartesiana



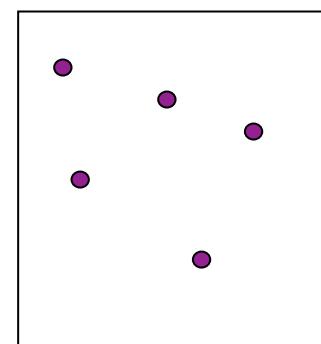
regular



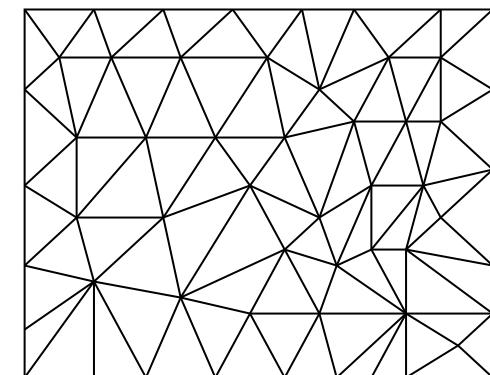
retilínea



curvilínea

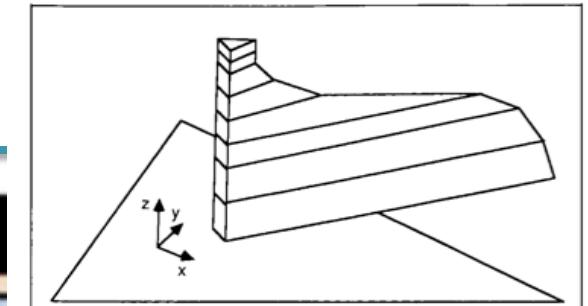
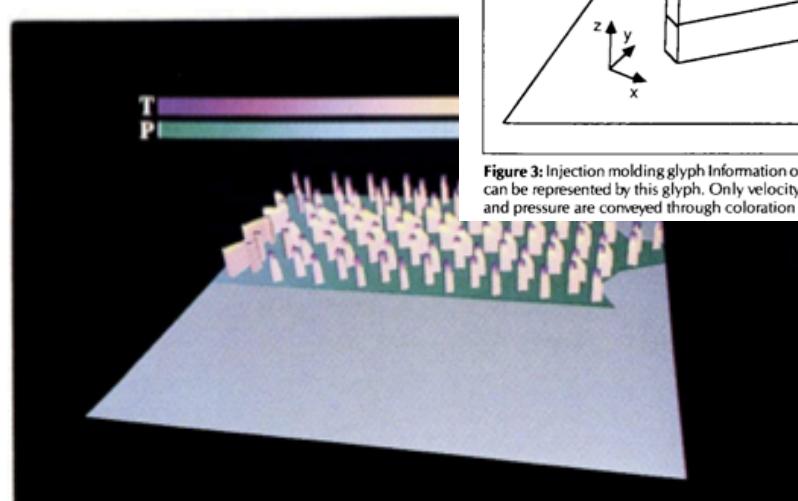
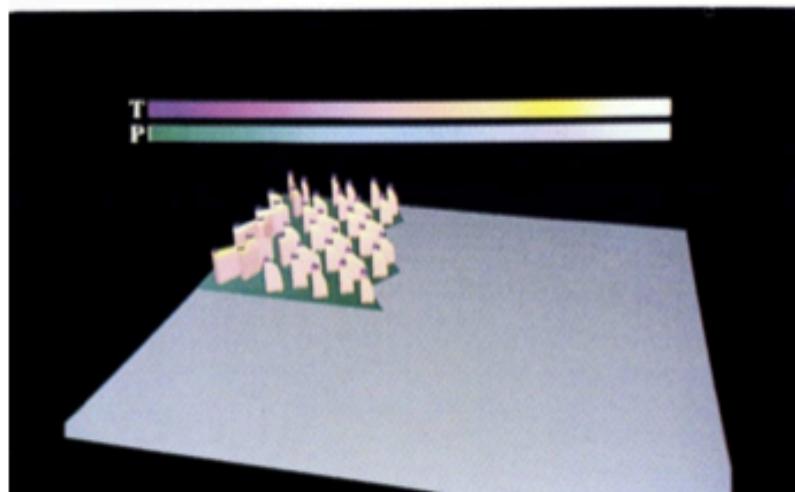


pontos esparsos

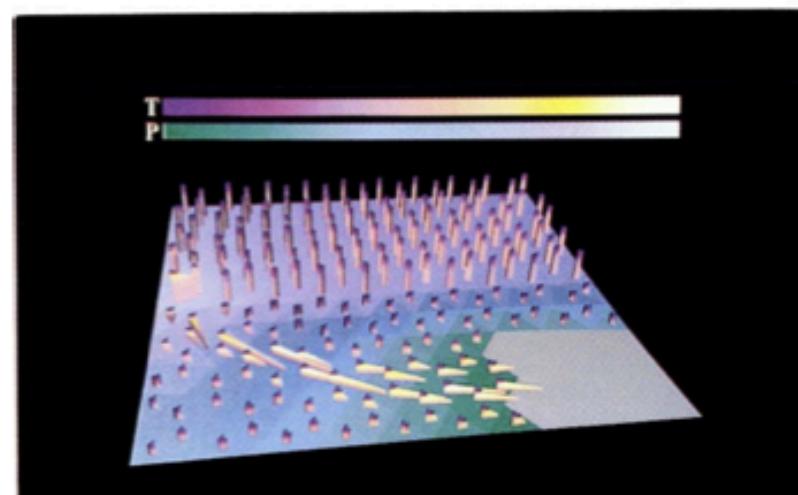
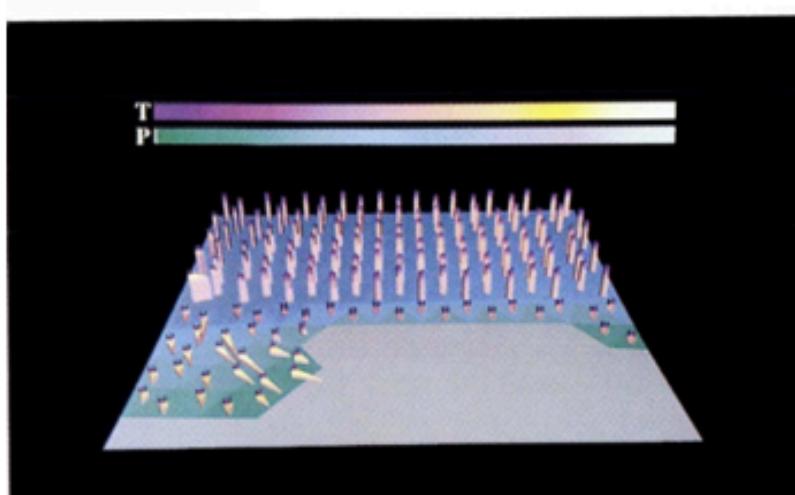


não-estruturada

# Dados 2D/3D

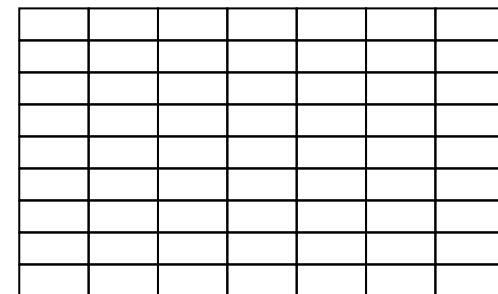
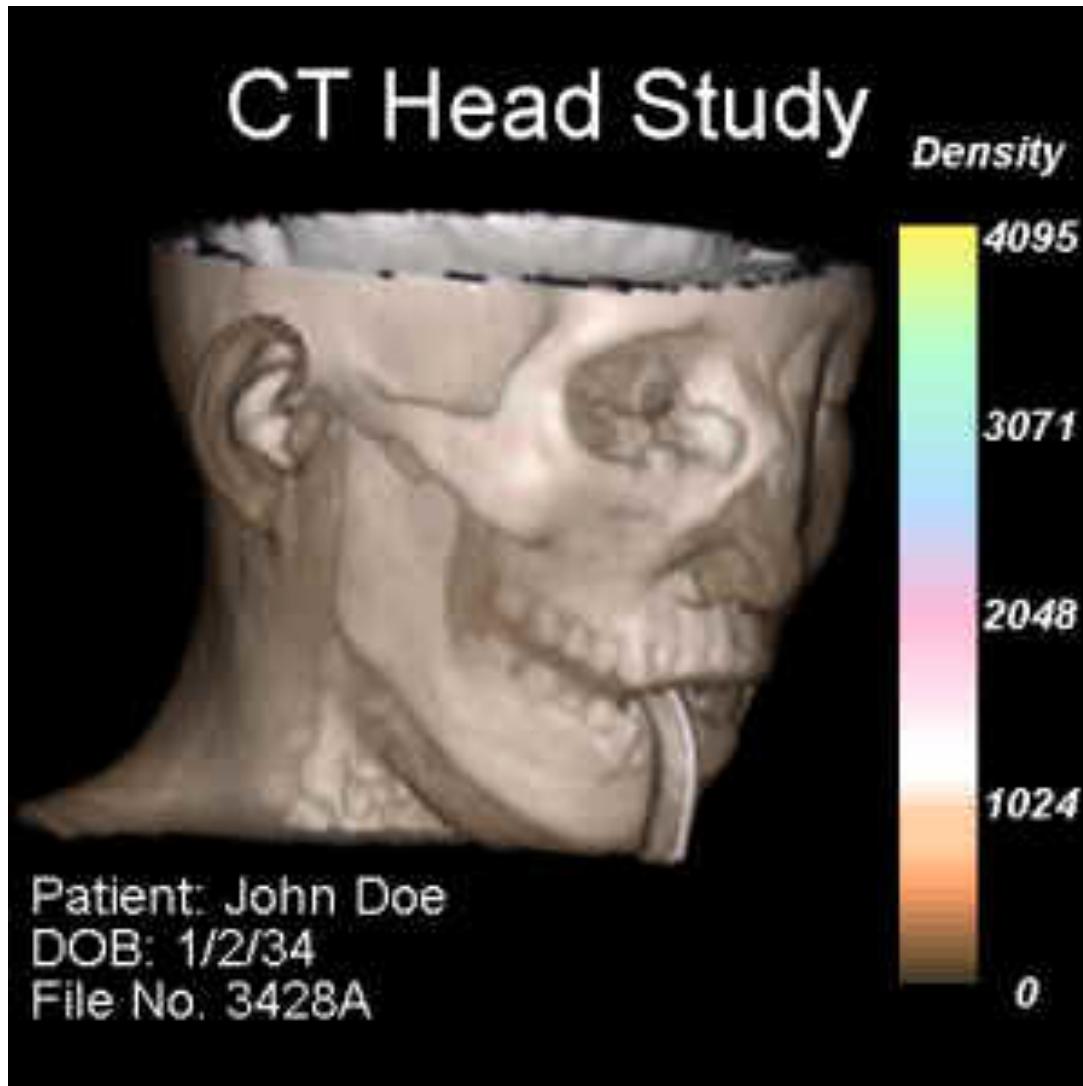


**Figure 3:** Injection molding glyph Information on pressure, temperature and velocity can be represented by this glyph. Only velocity is shown here as both temperature and pressure are conveyed through coloration of the glyph structure.



Ellson, R.; Cox, D. Visualization of injection molding. *Simulation*, 1988, 51:184-188

# Dados 3D

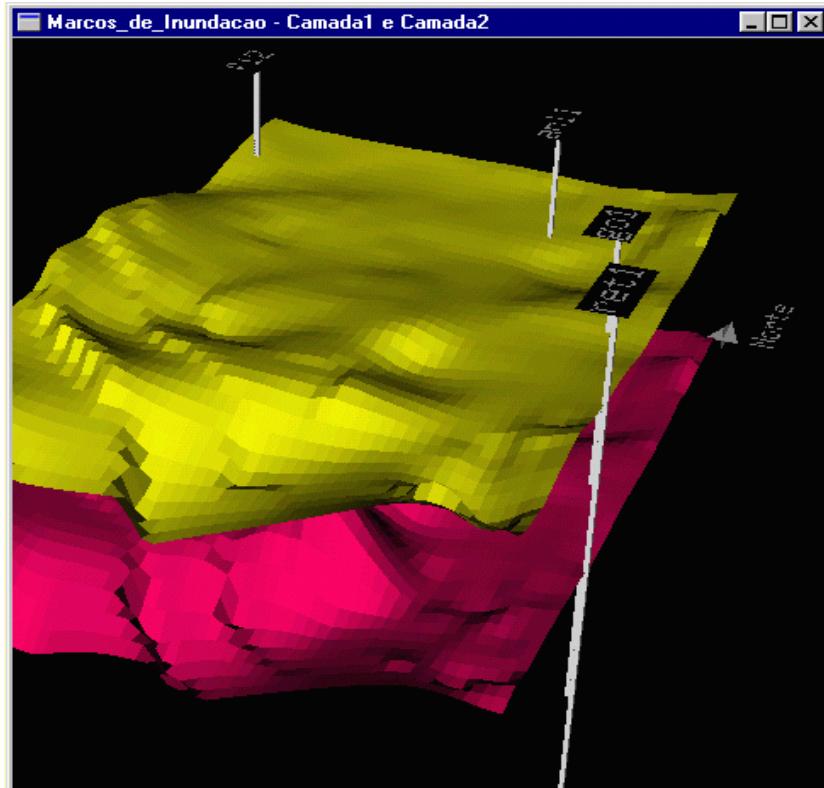


regular

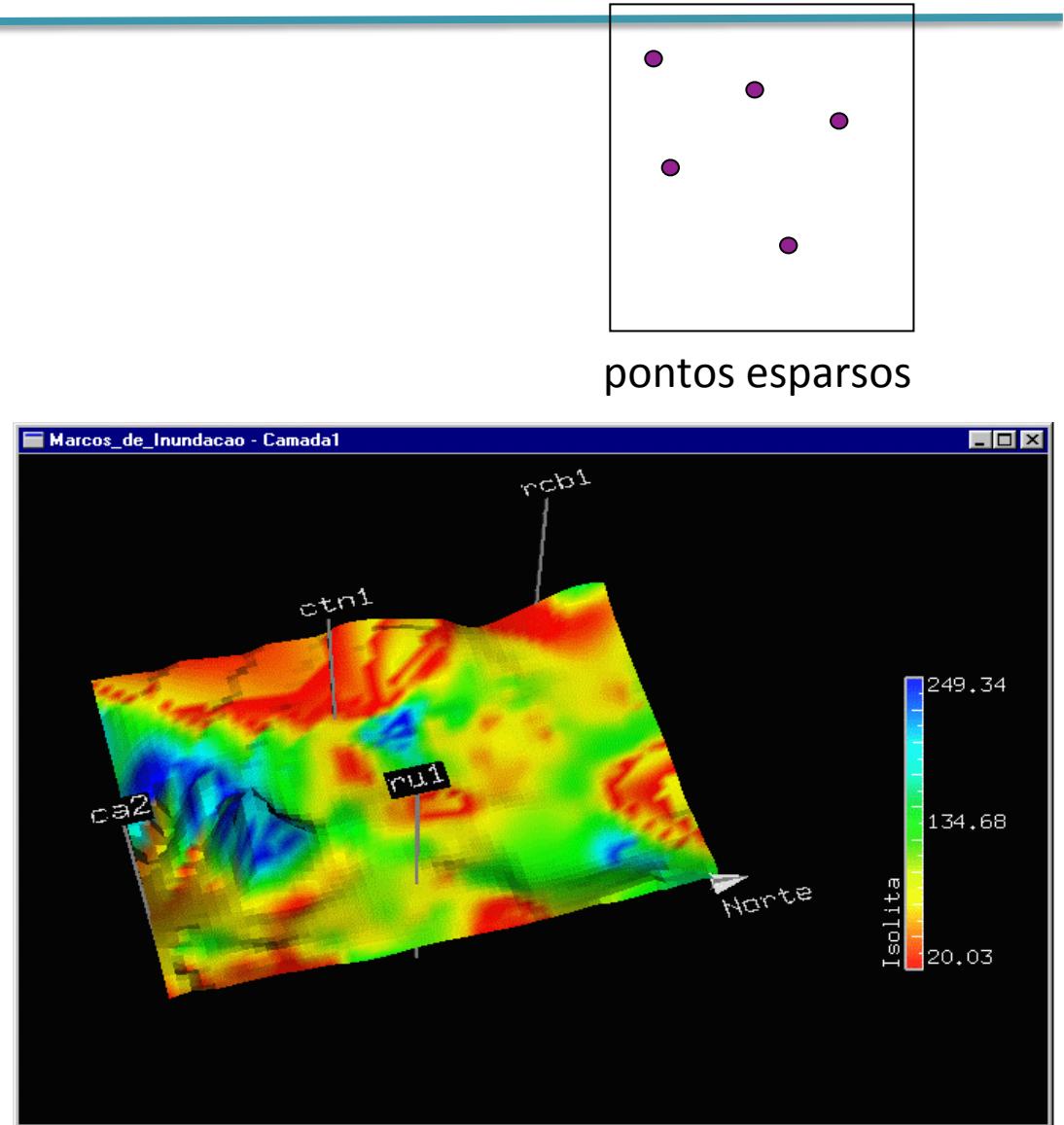
Valores de densidade dos tecidos estão associados a pontos regularmente espaçados no espaço 3D.

Imagen gerada com VolView  
([www.kitware.com](http://www.kitware.com))

# Representação do espaço (2D ou 3D)



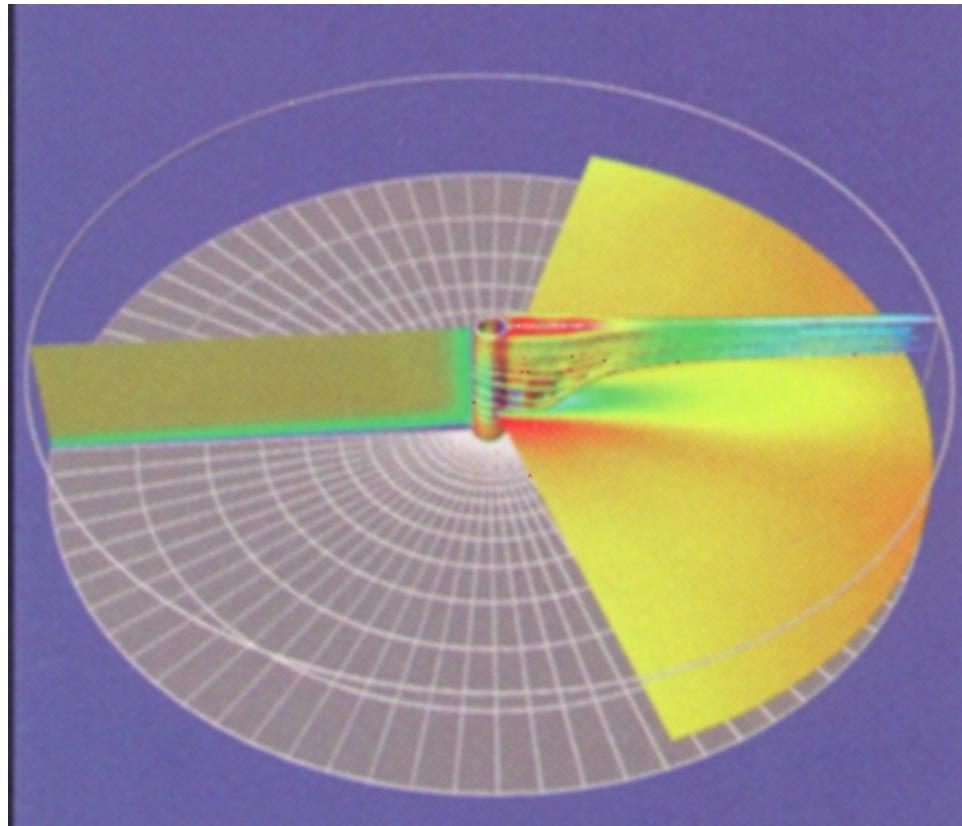
Geovis-UFRGS (1998)



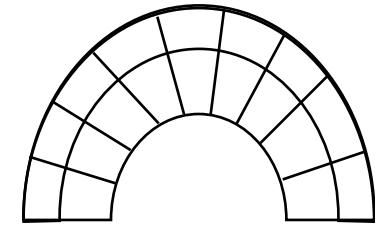
pontos esparsos

# Representação do espaço (2D ou 3D)

---



Simulação de fluxo  
(VTK)



curvilínea

# Dimensão do domínio?

---

- Hierarquias
  - Diretórios
- Redes
  - Sites web, redes de comunicação, comunidades sociais
- Coleção de documentos
  - Bibliotecas digitais
- Coleção de entidades num banco de dados
  - Cadastro de pessoas

# Dados nD

---

- Tabelas
  - atributos associados a entidades isoladas
- Árvores
  - entidades com relacionamentos hierárquicos
- Grafos
  - entidades com múltiplos relacionamentos

# Tabelas: cada atributo é uma dimensão

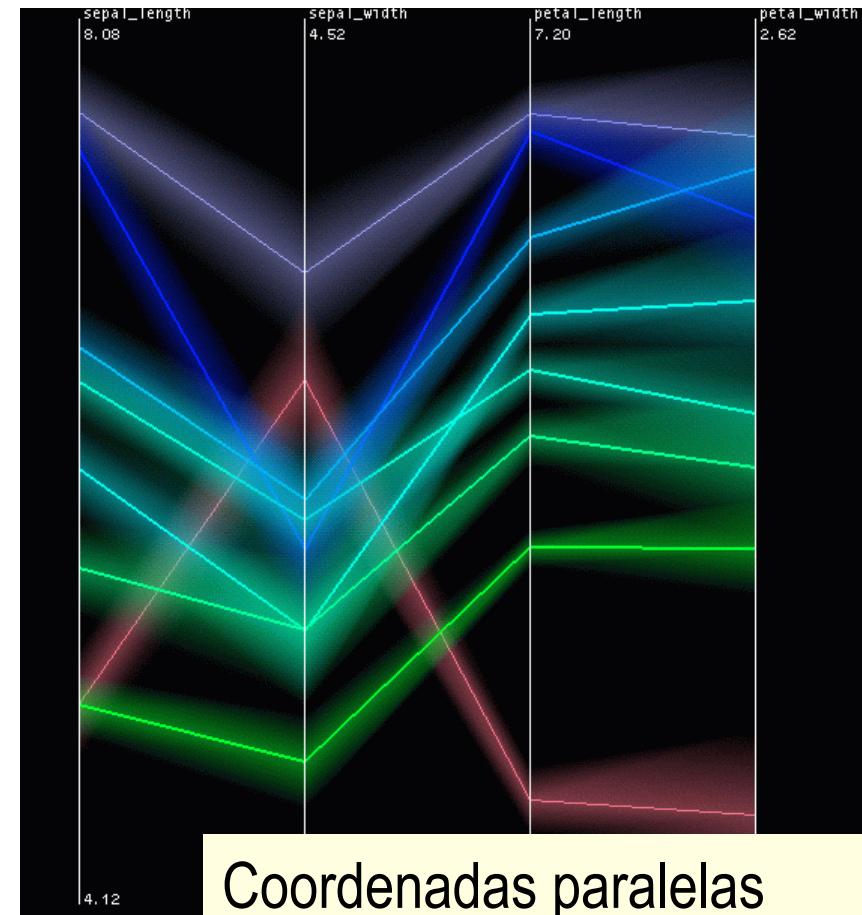
Microsoft Excel - film.txt

File Edit View Insert Format Data Window Help

Arial 10 B I U =

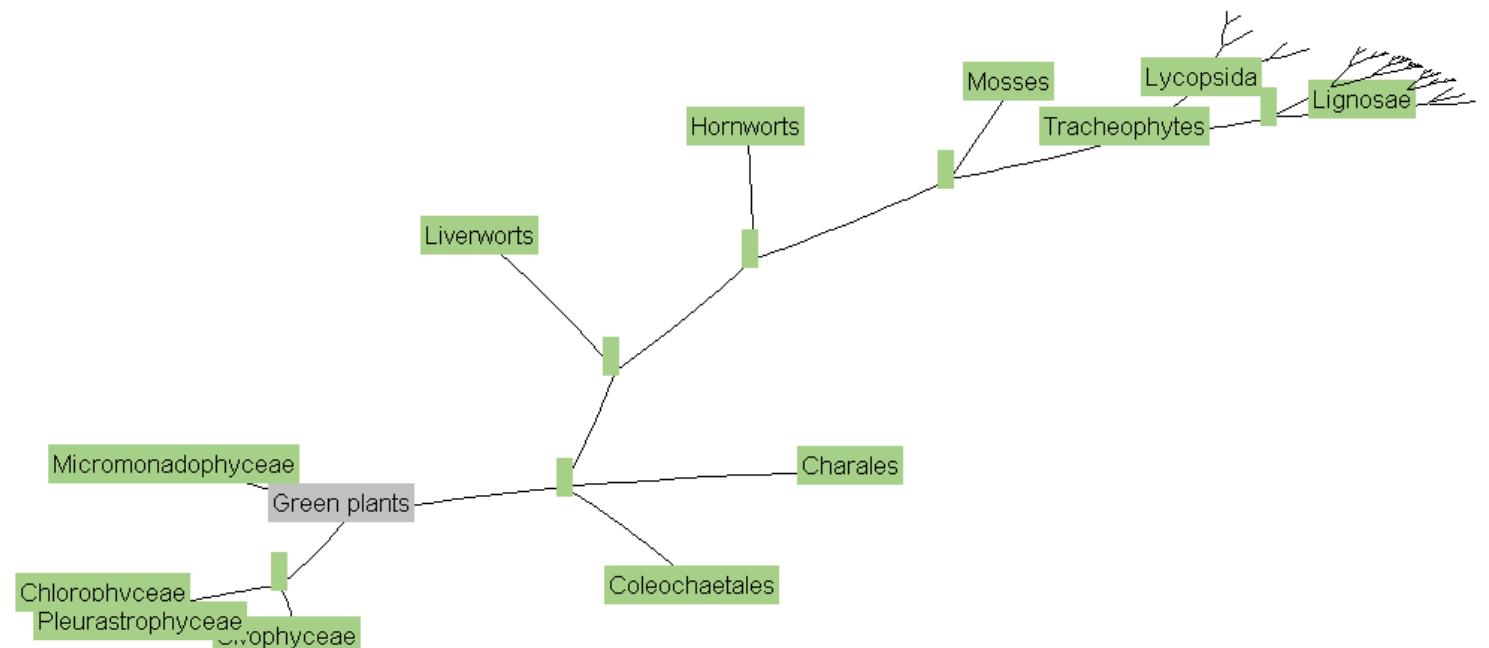
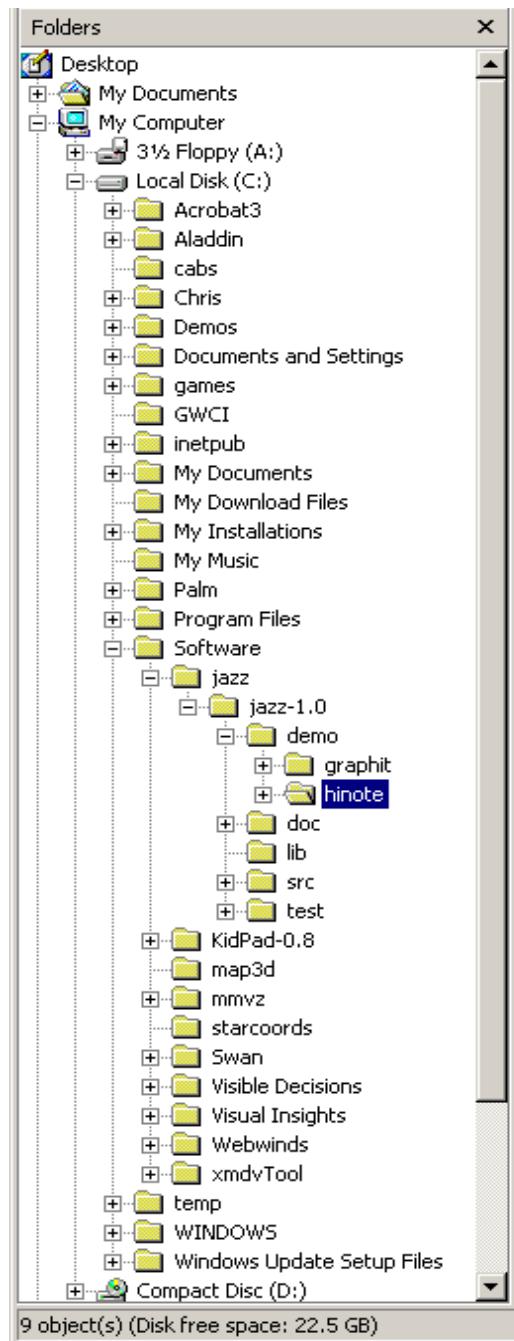
M3

	A	B	C	D	E	F	G	H	I	J	K
1	Year	Length	Title	Subject	Actor	Actress	Director	Popularity	Awards	*Image	
2	integer	integer	string	string	string	string	string	integer	string	string	
3	1990	125	Wild at Heart	Drama	Cage, Nicolas	Dern, Laura	Lynch, David	6	No	NicholasCage.gif	
4	1961	120	Goodbye Again	Drama	Perkins, Anthon	Bergman, Ingrid	Litvak, Anatole	6	No	NicholasCage.gif	
5	1990	135	Hunt for Red Octo	Drama	Connery, Sean		McTiernan, J.	8	No	NicholasCage.gif	
6	1984	108	Terminator, The	Action	Schwarzenegge	Hamilton, Linda	Cameron, J.	17	No	T2.gif	
7	1991	136	Terminator 2	Action	Schwarzenegge	Hamilton, Linda	Cameron, J.	8	No	T2.gif	
8	1993	65	John Cleese on H	Comedy	Cleese, John	Booth, Connie		62	No	NicholasCage.gif	
9	1987	103	Au Revoir les Enf	Drama	Manesse, Gasp	Racette, Franci	Malle, Louis	35	No	NicholasCage.gif	
10	1983	128	The Ballad of Nar	Drama		Missing	Imamura, Shoh	15	No	NicholasCage.gif	
11	1990	138	Cyrano De Berger	Drama	Depardieu, Ger	Brochet, Anne	Rappeneau, Je	86	No	NicholasCage.gif	
12	1990	107	Green Card	Comedy	Depardieu, Ger	MacDowell, And	Weir, Peter	25	No	NicholasCage.gif	
13	1987	118	Hope & Glory	War	Hayman, David	Miles, Sarah	Boorman, John	3	No	NicholasCage.gif	
14	1982	122	Missing	Drama	Lemmon, Jack	Spacek, Sissy	Costa-Gavras,	30	No	NicholasCage.gif	
15	1986	125	The Mission	Drama	Niro, Robert De	Lunghi, Cherie	Joffe, Roland	20	No	NicholasCage.gif	
16	1987	101	My Life As a Dog	Comedy	Glanzelius, Anton		Hallstrom, Lasse	21	No	NicholasCage.gif	
17	1984	150	Paris, Texas	Drama	Stanton, Harry I	Kinski, Nastass	Wim Wenders	27	No	NicholasCage.gif	
18	1984	106	Romancing the S	Action	Douglas, Micha	Turner, Kathleen	Silvestri, Rober	83	No	NicholasCage.gif	
19	1982	120	The State of Thing	Drama	Isabelle Weinga	Wenders, Wim		40	No	NicholasCage.gif	
20	1986	98	Summer	Comedy	Gauthier, Vince	Riviere, Marie	Rohmer, Eric	11	No	NicholasCage.gif	
21	1955	108	Smiles of a Summ	Comedy	Bjornstrand, Gu	Jacobsson, Ulla	Bergman, Ingmar	58	No	Bergman.gif	
22	1987	98	Under the Sun of	Drama	Depardieu, Ger	Bonnaire, Sandi	Pialat, Maurice	45	No	NicholasCage.gif	
23	1985	105	Vagabond	Drama	Meril, Macha	Bonnaire, Sandi	Varda, Agnes	49	No	NicholasCage.gif	
24	1988	115	Working Girl	Comedy	Ford, Harrison	Griffith, Melanie	Nichols, Mike	25	No	NicholasCage.gif	
25	1984	106	A Year of the Qui	Drama	Wilson, Scott	Komorowska, M	Zanussi, Krzys	78	No	NicholasCage.gif	
26	1983	134	Yentl	Music	Patinkin, Mand	Streisand, Barb	Streisand, Bart	46	No	NicholasCage.gif	
27	1982	111	Yol	Drama	Akan, Tarik		Guney, Yilmaz	53	No	NicholasCage.gif	
28	1992	102	The Addams Fam	Comedy	Julia, Raul	Huston, Anjelic	Sonnenfeld, B.	8	No	NicholasCage.gif	
29	1992	88	Adventures in Din	Action	Katz, Orn	Hoffman, Shaw	Thompson, Bre	19	No	NicholasCage.gif	
30	1992	95	Alan & Naomi	Drama	Haas, Lukas	Aquino, Vanessa	Vanwagenen, S	3	No	NicholasCage.gif	

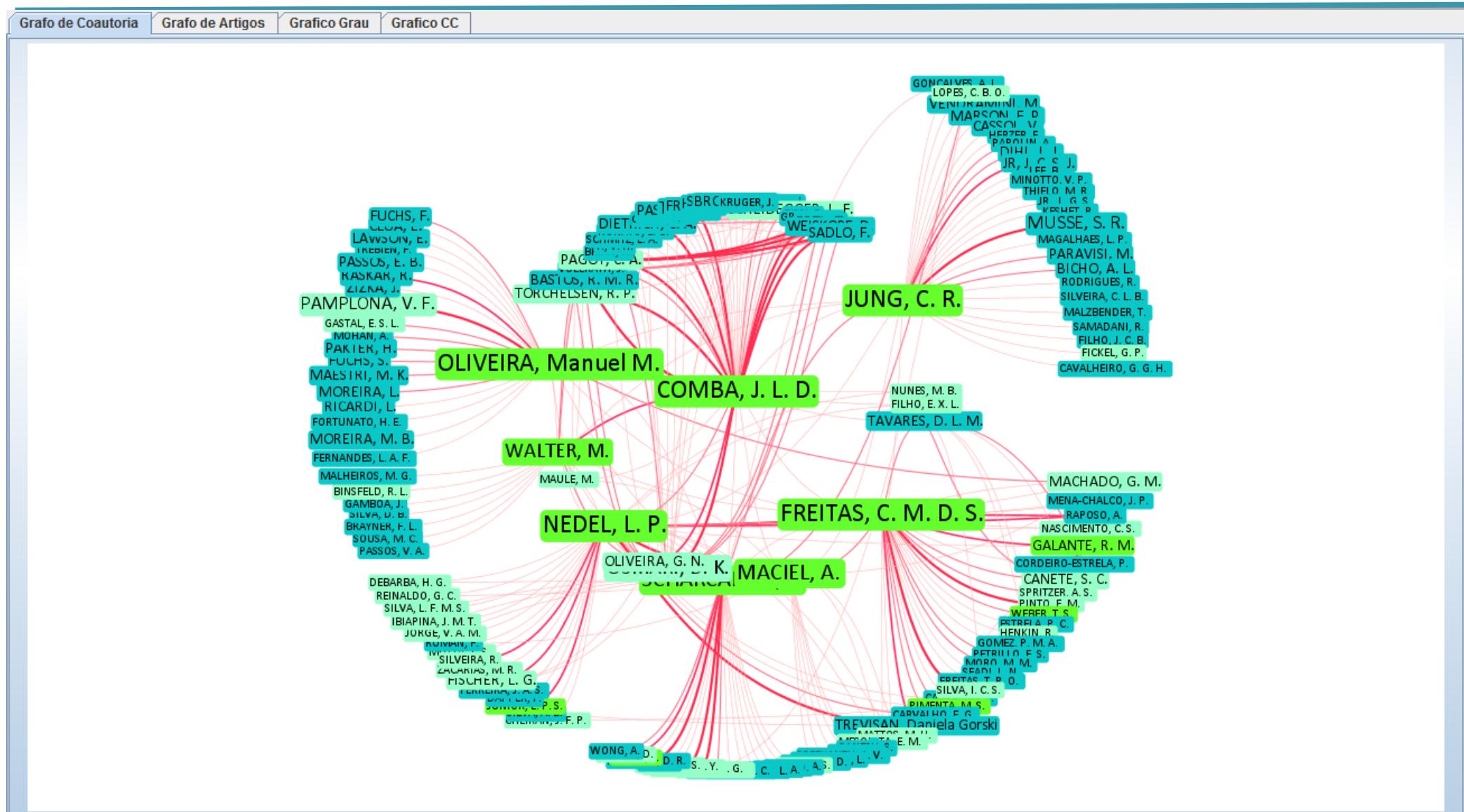


Coordenadas paralelas  
(Inselberg)

# Árvores

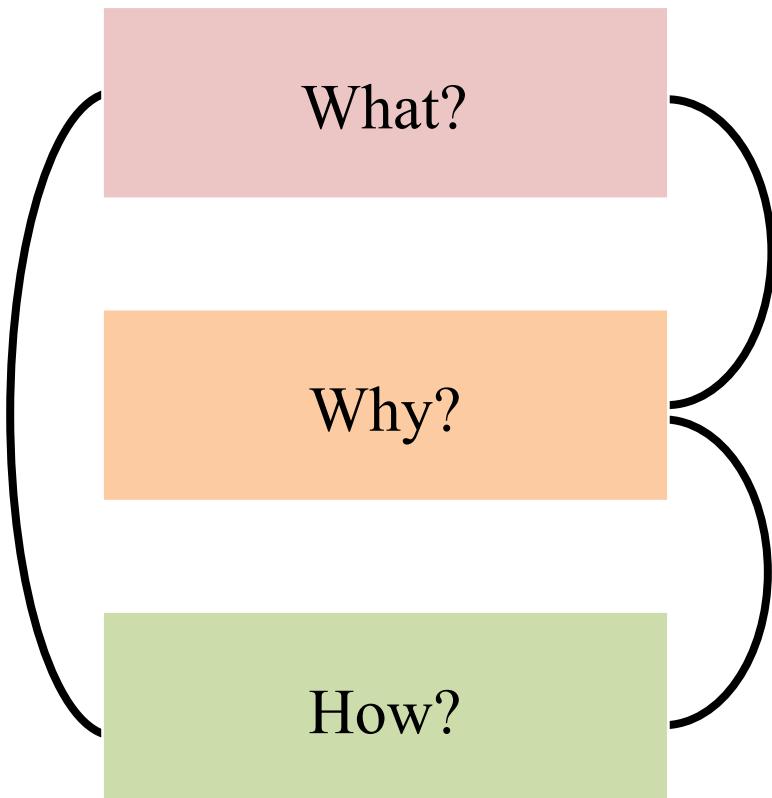


# Grafos



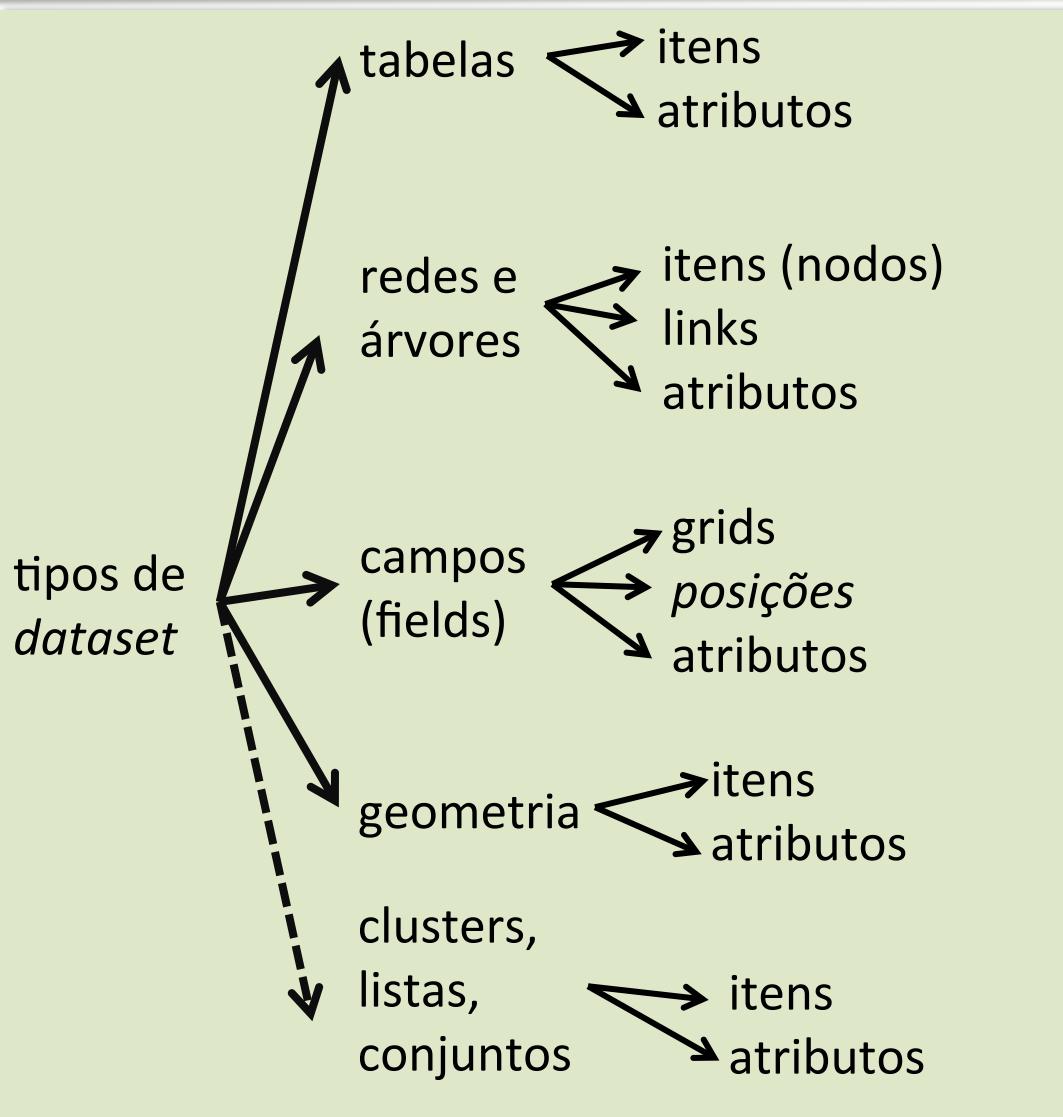
# Abordagem “what-why-how”

---

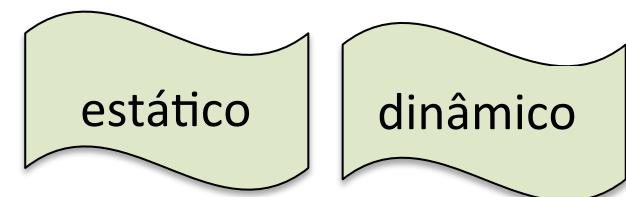


- Abordagem para analisar técnicas de acordo com as 3 questões
- **What**
  - Que dados os usuários vêem
- Why
  - Quais as intenções dos usuários em usar uma técnica de vis
- How
  - Quais são as codificações visuais e as técnicas de interação (*idiomas*)

# Caracterização de dados



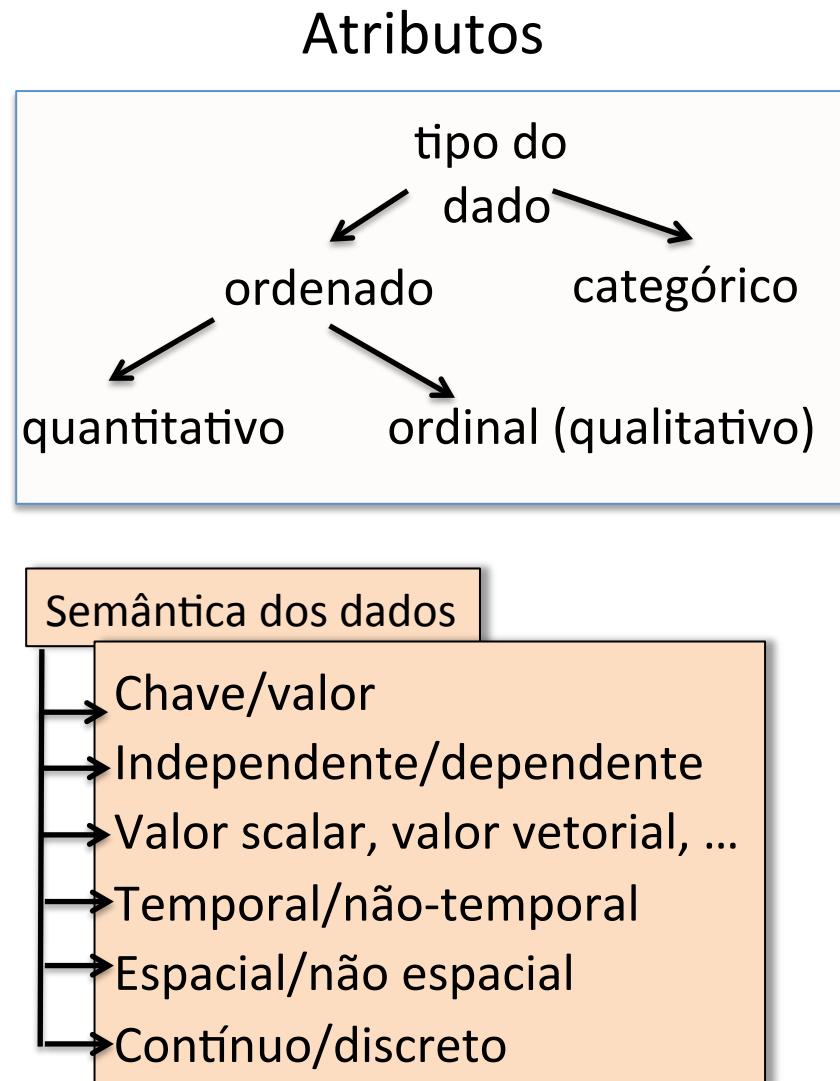
Aplicações podem requerer combinações desses tipos básicos



Munzner, Tamara *Visualization Analysis and Design*.  
Boca Raton, FL:Taylor & Francis, 2014

# Caracterização de dados

- Tipos de dados básicos
  - Item (nodo)
  - Atributo
  - Aresta
  - Posição
  - Grid



# Tarefa 1

---

- Caracterizar um conjunto de dados de sua escolha:
  - Classe do dado
  - Tipo dos dados
  - Natureza e dimensão do domínio