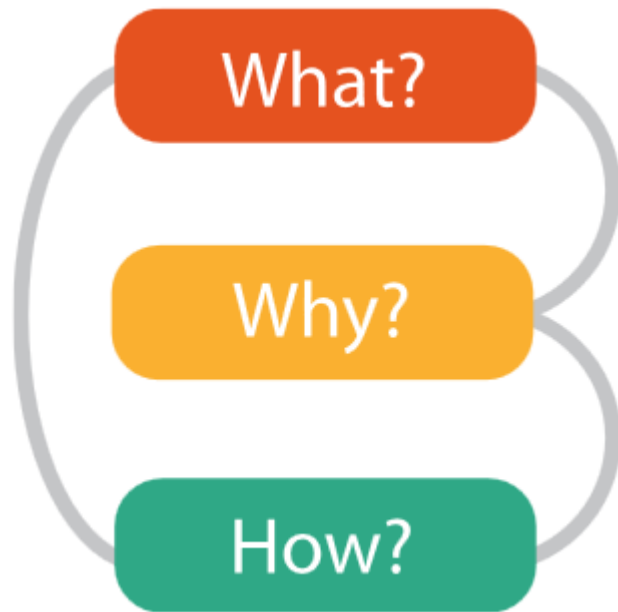


Visualização de Dados – Parte 02

Prof. Wellington Franco

Análise de Visualização: *Framework What-Why-How*

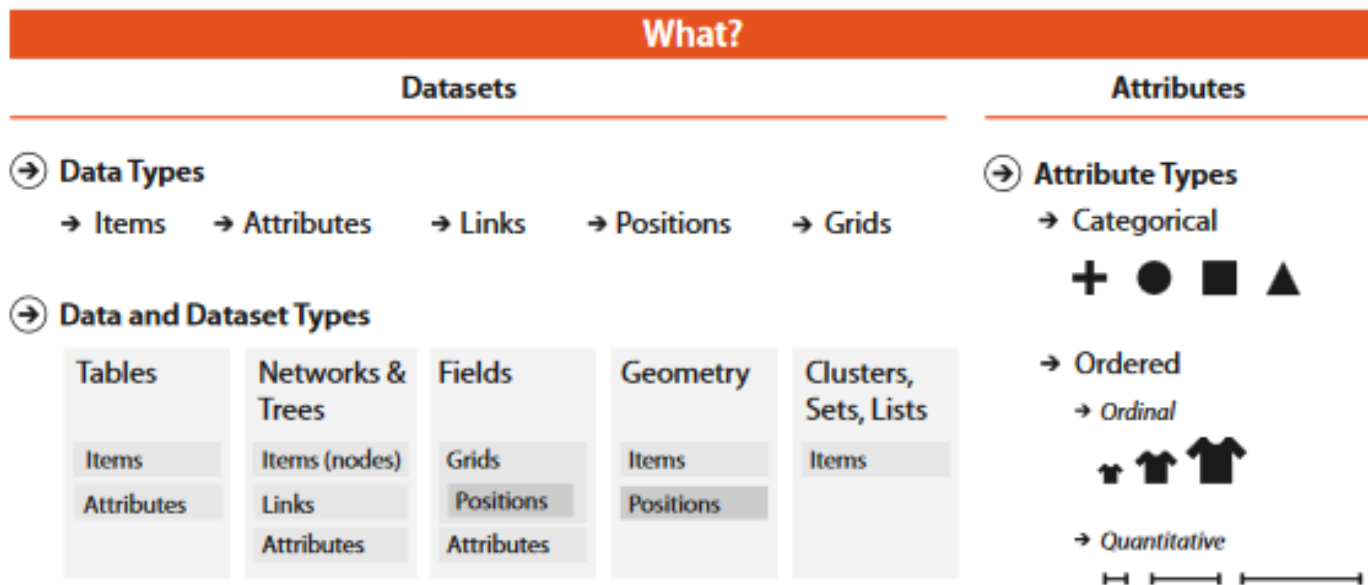
- **O que** é mostrado?
 - Abstração de **dados**.
- **Por que** o usuário está olhando?
 - Abstração de **tarefas**.
- **Como** é mostrado?
 - Codificação visual e interação.



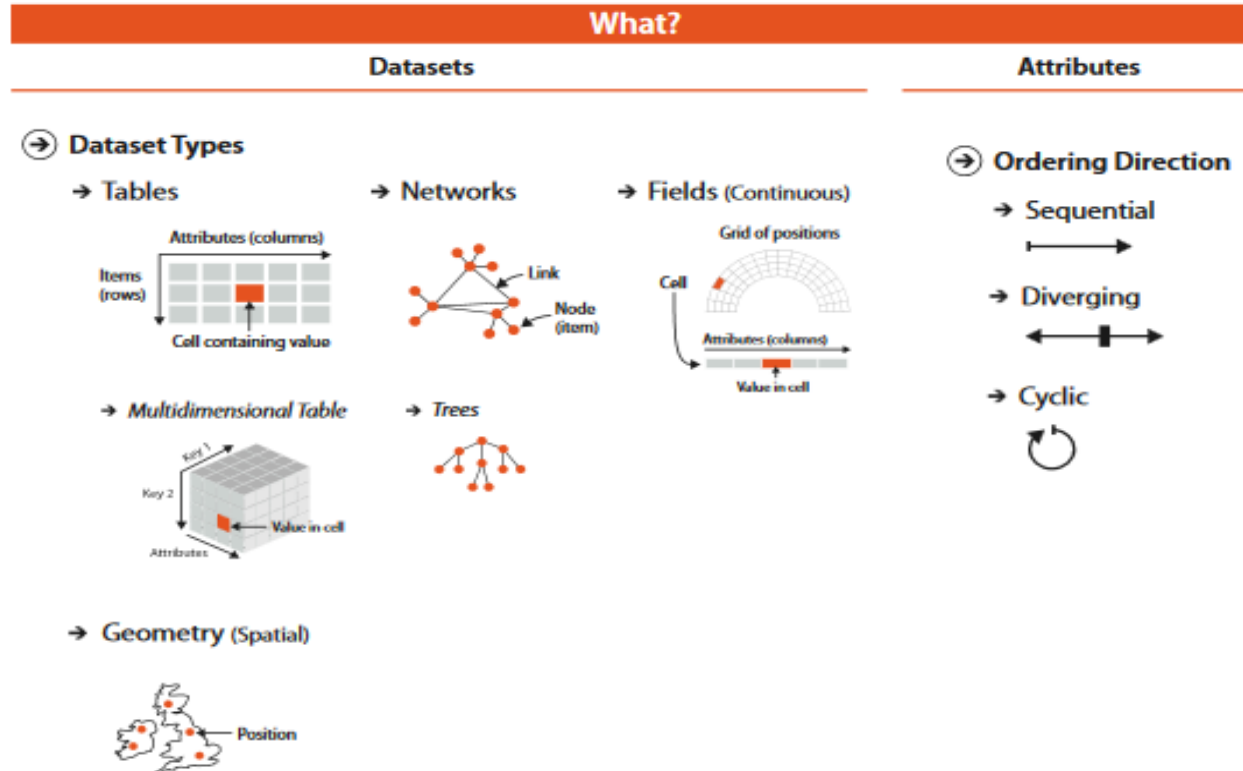
Abstração de Dados

- Tipos abstratos de dados que podem ser visualizados:
 - Conjunto de Dados
 - Atributos
- Responde à pergunta do tipo “**O QUE é mostrado?**”

Abstração de Dados: Visão Geral



Abstração de Dados: Visão Geral



Abstração de Dados: Visão Geral



→ Dataset Availability

→ Static



→ Dynamic



Tipos e Semântica de Dados

- Um dos aspectos do *design de vis* é o tipo de dados disponível;
 - Que informação podemos extrair dos dados e que significados precisamos saber explicitamente?
 - Exemplo: o que os dados abaixo significam?

Basil, 7, S, Pear

Tipos e Semântica de Dados

Basil, 7, S, Pear

- É necessário saber a semântica e os tipos dos dados;
- **Semântica:** significado no mundo real;
 - Exemplo:
 - A palavra representa o primeiro nome de uma pessoa, o nome fantasia de uma empresa, um objeto, uma fruta, ...?
 - O número representa um dia? Um mês? Uma idade?

Tipos e Semântica de Dados

Basil, 7, S, Pear

- O **tipo** do dado é a sua interpretação estrutural ou matemática;
 - No nível dos dados, que tipo de coisa é?
 - *Um ítem, um relacionamento ou um atributo?*
 - No nível dos conjuntos de dados, como esses tipos de dados são combinados em uma estrutura maior?
 - *É uma tabela, uma árvore ou um campo de valores amostrados?*
 - No nível de atributos, que tipos de operações matemáticas fazem sentido para eles?
 - *Quantidades x Categorias?*

Tipos e Semântica de Dados

Basil, 7, S, Pear

ID	Name	Age	Shirt Size	Favorite Fruit
1	Amy	8	S	Apple
2	Basil	7	S	Pear
3	Clara	9	M	Durian
4	Desmond	13	L	Elderberry
5	Ernest	12	L	Peach
6	Fanny	10	S	Lychee
7	George	9	M	Orange
8	Hector	8	L	Loquat
9	Ida	10	M	Pear
10	Amy	12	M	Orange

Tipos de Dados

- 5 tipos básicos:
 - Item;
 - Atributo;
 - *Link* (relacionamento)
 - Posição
 - *Grid* (grade)

Tipos de Dados

- 5 tipos básicos:
 - **Item:** entidade individual discreta, ex: uma linha em uma tabela, um nó de uma árvore, representando uma pessoa, uma ocorrência de um crime;
 - Em *data science* é mais conhecido por instância.
 - **Atributo:** alguma propriedade específica que pode ser medida, observada ou registrada, ex: salário, preço, temperatura;
 - Também chamado de variável, dimensão.
 - **Link:** relacionamento entre itens, normalmente em uma rede;

Tipos de Dados

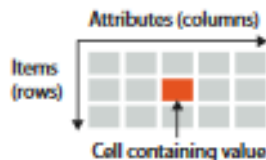
- 5 tipos básicos:
 - **Posição:** um dado espacial, fornecendo uma localização num espaço bidimensional (2D) ou tridimensional (3D), ex: pares latitude-longitude representando coordenadas na superfície terrestre;
 - **Grid:** especifica uma estratégia para *amostrar* dados contínuos em termos de relacionamento geométrico e tipológico entre suas células.

Tipos de Conjunto de Dados

- Um conjunto de dados (*dataset*) é qualquer coleção de informações que pode ser alvo de análise.

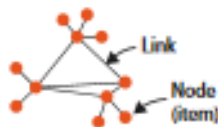
→ Dataset Types

→ Tables

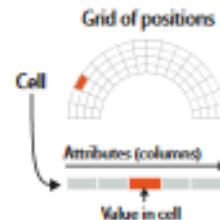


Usado em
Machine Learning

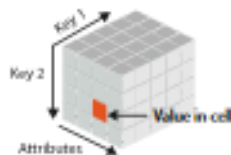
→ Networks



→ Fields (Continuous)



→ Multidimensional Table



Usado em
Deep Learning (Tensor)

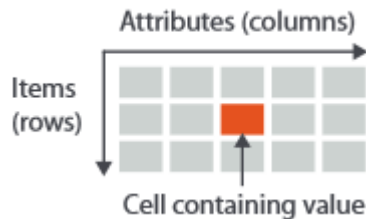
→ Trees



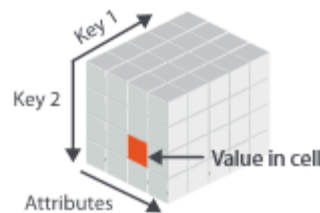
Usado em
Ontologia (RDF)

Tabelas

- Tipos de conjunto de dados mais comuns;
- São formadas por linhas (itens) e colunas (atributos);
- Cada célula da tabela é especificada pela combinação de uma linha e uma coluna - um item e um atributo, e contém um valor para esse par;
- Uma tabela multidimensional possui uma estrutura mais complexa de indexação de uma célula, com múltiplas chaves.



→ *Multidimensional Table*



Tabelas

A	B	C	S	T	U
Order ID	Order Date	Order Priority	Product Container	Product Base Margin	Ship Date
3	10/14/06	5-Low	Large Box	0.8	10/21/06
6	2/21/08	4-Not Specified	Small Pack	0.55	2/22/08
32	7/16/07	2-High	Small Pack	0.79	7/17/07
32	7/16/07	2-High	Jumbo Box		7/17/07
32	7/16/07	2-High	Medium Box		7/18/07
32	7/16/07	2-High	Medium Box	0.65	7/18/07
35	10/23/07	4-Not Specified	Wrap Bag	0.52	10/24/07
35	10/23/07	4-Not Specified	Small Box	0.58	10/25/07
36	11/3/07	1-Urgent	Small Box	0.55	11/3/07
65	3/18/07	1-Urgent	Small Pack	0.49	3/19/07
66	1/20/05	5-Low	Wrap Bag	0.56	1/20/05
69		5 4-Not Specified	Small Pack	0.44	6/6/05
69		5 4-Not Specified	Wrap Bag	0.6	6/6/05
70	12/18/06	5-Low	Small Box	0.59	12/23/06
70	12/18/06	5-Low	Wrap Bag	0.82	12/23/06
96	4/17/05	2-High	Small Box	0.55	4/19/05
97	1/29/06	3-Medium	Small Box	0.38	1/30/06
129	11/19/08	5-Low	Small Box	0.37	11/28/08
130	5/8/08	2-High	Small Box	0.37	5/9/08
130	5/8/08	2-High	Medium Box	0.38	5/10/08
130	5/8/08	2-High	Small Box	0.6	5/11/08
132	6/11/06	3-Medium	Medium Box	0.6	6/12/06
132	6/11/06	3-Medium	Jumbo Box	0.69	6/14/06
134	5/1/08	4-Not Specified	Large Box	0.82	5/3/08
135	10/21/07	4-Not Specified	Small Pack	0.64	10/23/07
166	9/12/07	2-High	Small Box	0.55	9/14/07
193	8/8/06	1-Urgent	Medium Box	0.57	8/10/06
194	4/5/08	3-Medium	Wrap Bag	0.42	4/7/08

attribute

cell

item

instância

Redes (ou Grafos) e Árvores

- Tipo de conjunto de dado apropriado para quando existe algum tipo de relacionamento entre dois ou mais itens;
- Um item em um **grafo** é chamado nó;
- Um link é um relacionamento entre dois itens;
 - Exemplo: Redes sociais, Redes de Computadores;
- Nós e links podem ter atributos associados a eles.
- **Árvores:** redes hierárquicas acíclicas;
 - Cada nó-filho possui um único nó-pai apontando para ele.
 - Exemplo: Estrutura hierárquica de uma empresa, árvore genealógica

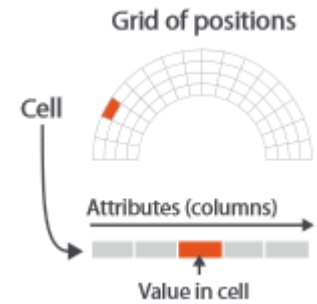


→ Trees

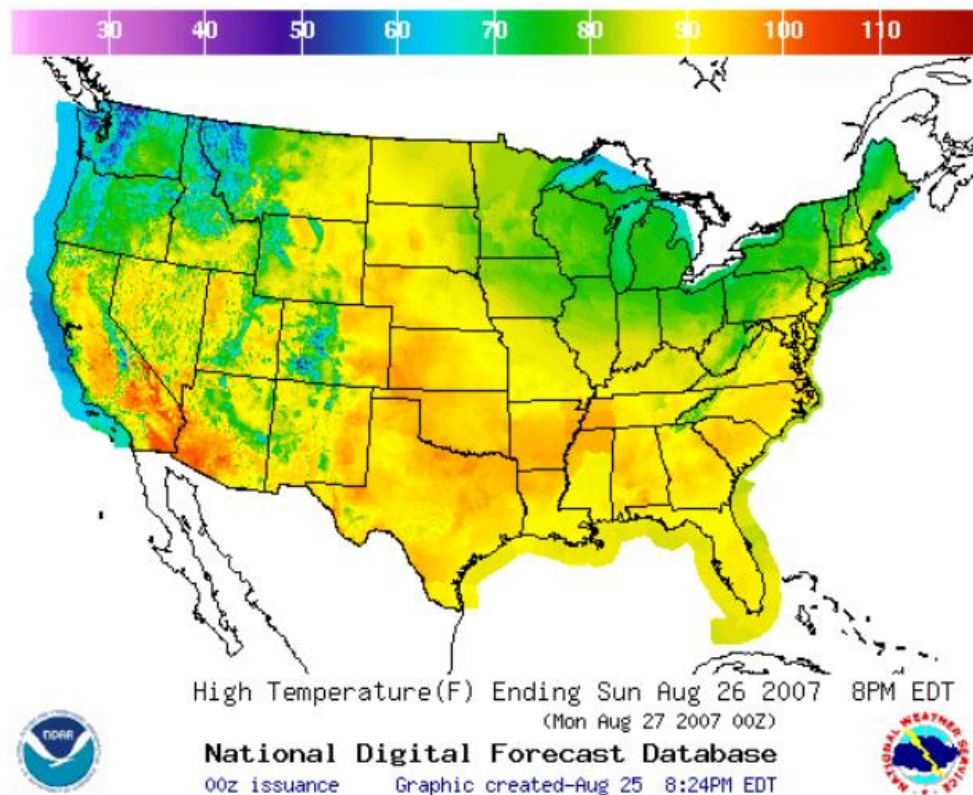


Grid (Campos)

- Também possui valores de atributos associados a células;
- Cada célula em um campo contém medidas ou cálculos de um **domínio contínuo** (infinitos valores);
- Exemplos de medidas: temperatura, pressão, velocidade, força e densidade;
- Exemplo de campo: tomografia computadorizada com $64 \times 64 \times 64 = 262.144$ células representando um volume cúbico;

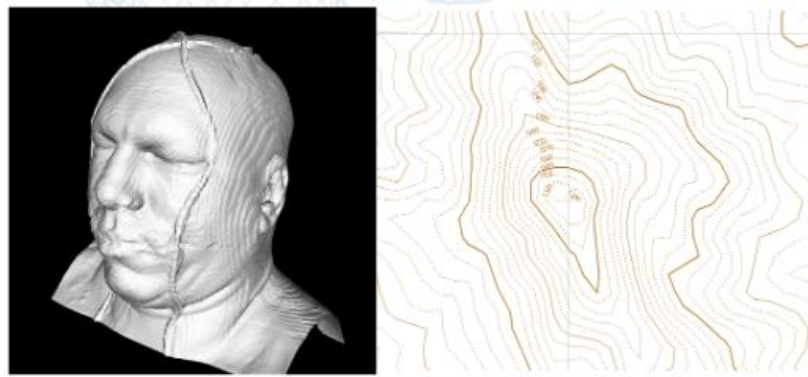


Campos



Geometria

- Especifica informação sobre a forma dos itens com posições espaciais explícitas;
- Itens são pontos, linhas ou curvas, superfícies 2D ou volumes 3D; São tipos de conjuntos de dados essencialmente espaciais e não necessariamente possuem atributos.



Disponibilidade dos Dados

- **Dados estáticos:** O conjunto de dados inteiro está disponível de uma vez (offline);
- **Dados dinâmicos ou *stream* de dados:** os dados são obtidos sob demanda ao longo de uma sessão (*online*)
 - Itens são adicionados ou removidos ao longo da visualização;
 - Qualquer um dos tipos de conjuntos de dados pode ser estático ou dinâmico.

→ Static



→ Dynamic



Tipos de Atributos

➔ Attribute Types

➔ Categorical

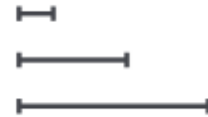


➔ Ordered

➔ Ordinal



➔ Quantitative



➔ Ordering Direction

➔ Sequential



➔ Diverging



➔ Cyclic



Tipos de Atributos

- **Categóricos (ou nominais):**

- Não possuem ordem inerente;
- Exemplos: nomes de frutas, títulos de filmes, nomes de cidades, etc.



Tipos de Atributos

- **Ordenados:**

- Todos os dados ordenados possuem uma ordem implícita;
- Dados ordinais ou quantitativos;
- Dados ordinais: operações matemáticas não são possíveis.
 - Exemplos: tamanhos de camisa (pequeno, médio e grande)



Tipos de Atributos

- **Ordenados:**

- Dados quantitativos: dão suporte a comparação aritméticas;
 - Exemplos: altura, idade, peso, temperatura, tempo de execução de um algoritmo, etc.
- Números inteiros e reais



Tipos de Atributos

- **Ordenados Direcionais:**

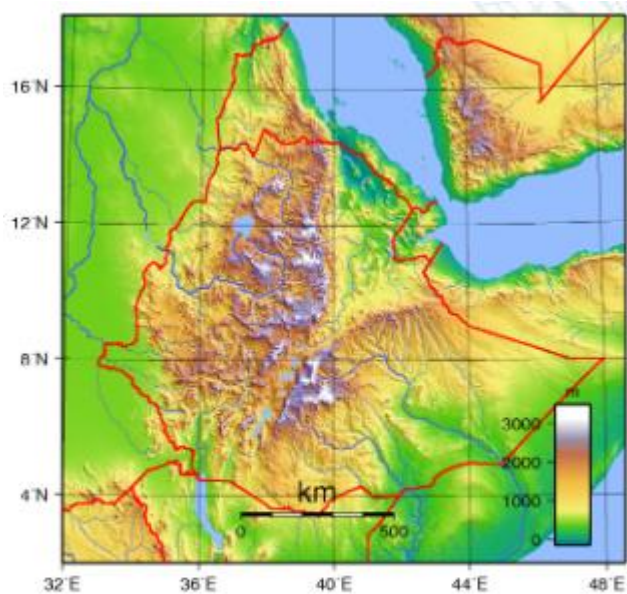
- Sequenciais:

- Existe um intervalo homogêneo de um valor mínimo a um valor máximo.

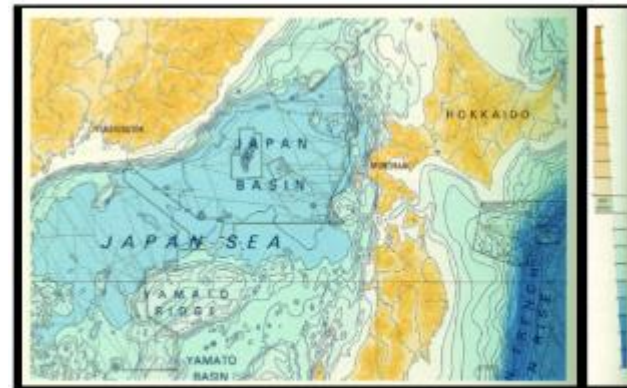
- Divergentes:

- Pode ser separado em duas sequências apontando em direções opostas partindo de um ponto inicial comum

Tipos de Atributos



Atributos sequenciais:
somente altitude



Atributos divergentes:
altitude e batimetria

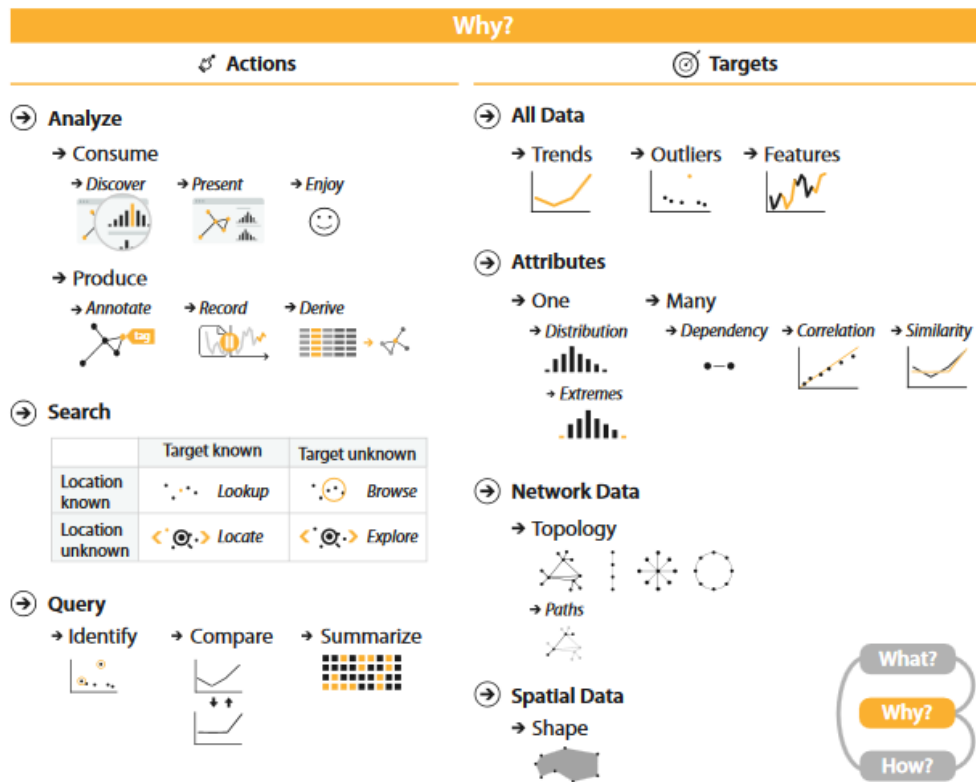
Tipos de Atributos

- **Ordenados Direcionais:**

- Cíclicos:

- Valores voltam ao ponto de partida ao invés de crescerem indefinidamente;
 - Várias medidas de tempo são cíclicas: horas do dia, dias da semana e meses do ano

Abstração de Tarefas: Visão Geral



Ações (Verbos)

Três níveis de ações:

→ Analyze

→ Search

→ Query

Ações (Verbos)

Em Análise, temos:

- Consumir:
 - Descobrir
 - Apresentar
 - Apreciar
- Produzir:
 - Anotar
 - Registrar
 - Derivar

→ Analyze

→ Consume

→ Discover



→ Present



→ Enjoy



→ Produce

→ Annotate



→ Record



→ Derive



→ Search

→ Query

Descobrir

```
In [46]: file.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 24 entries, 0 to 23
Data columns (total 5 columns):
 #   Column
---  ---
 0   YEARMONTH
 1   INCIDENTCLASSIFICATION
 2   INCIDENTBOROUGH
 3   INCIDENTCOUNT
 4   AVERAGERESPONSETIME
dtypes: int64(1), object(4)
memory usage: 1.1+ KB
```

```
In [53]: file['INCIDENTCOUNT'].agg(np.max)
```

```
Out[53]: 40850
```

```
In [54]: file['INCIDENTCOUNT'].agg(np.min)
```

```
Out[54]: 10
```

```
In [55]: file['INCIDENTCOUNT'].agg(np.mean)
```

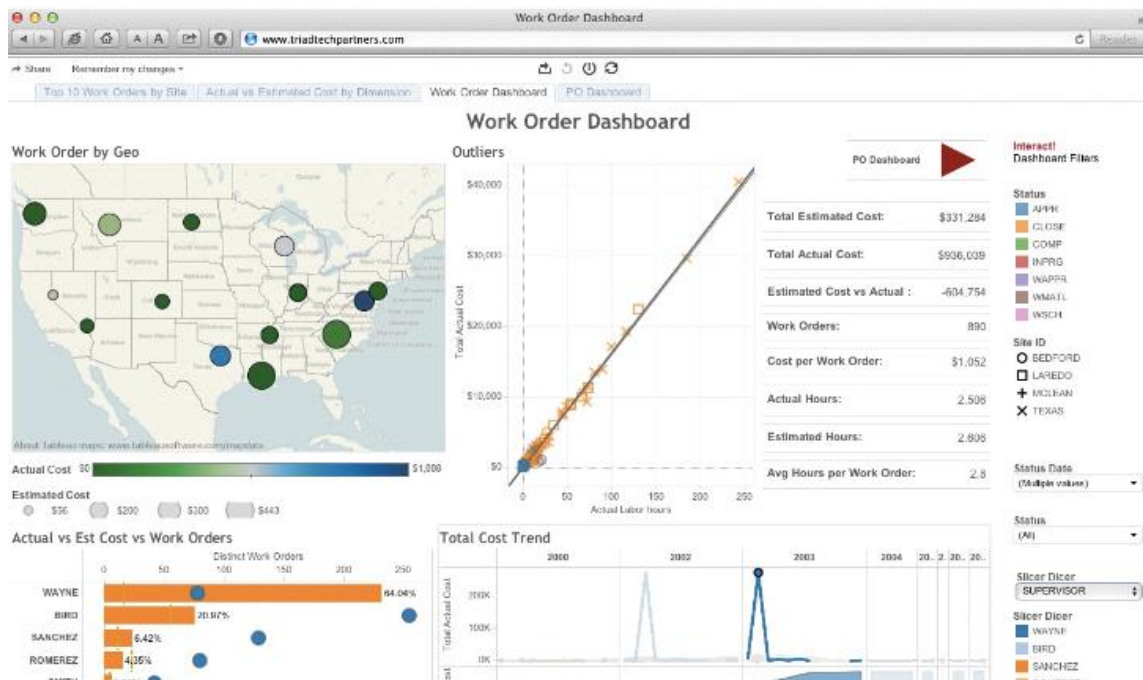
```
Out[55]: 5209.916666666667
```

```
In [56]: file['INCIDENTCOUNT'].agg(np.std)
```

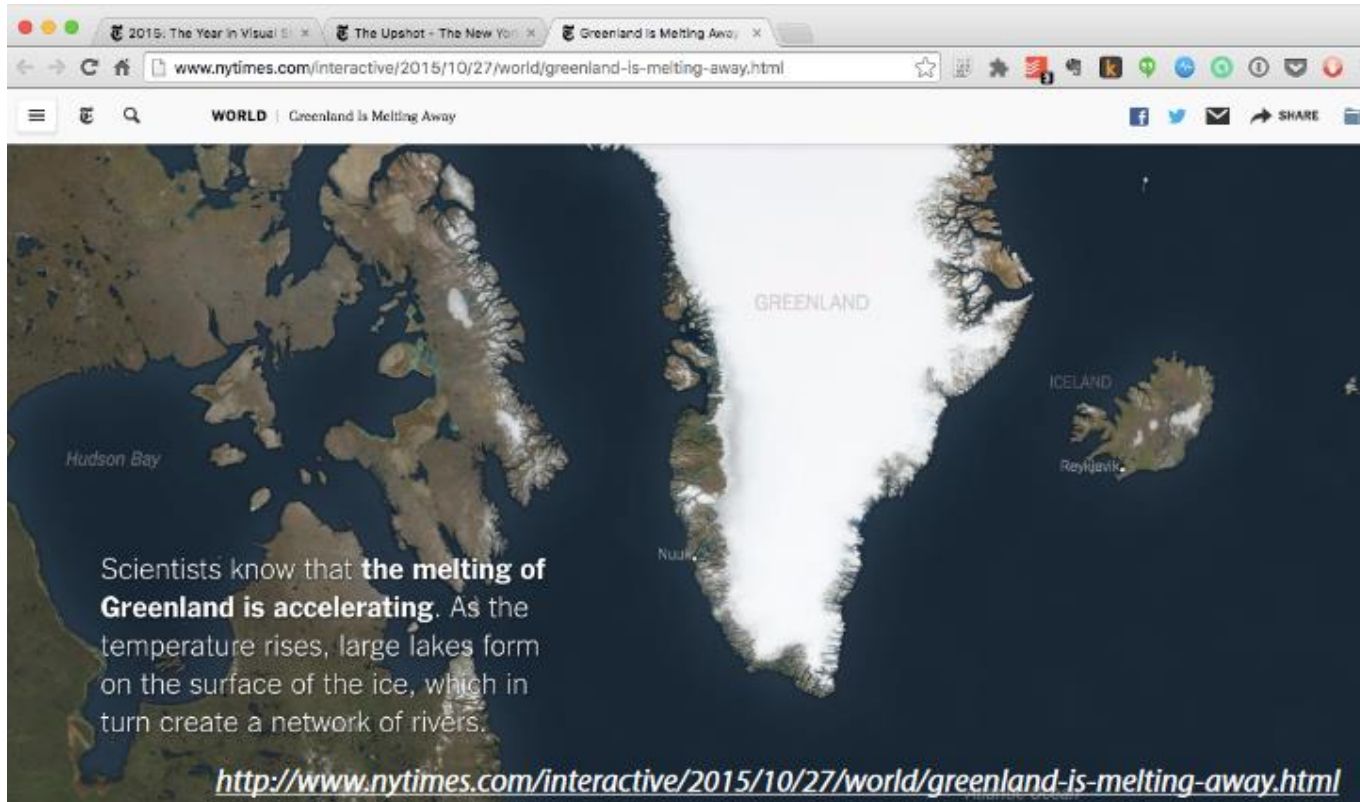
```
Out[56]: 8945.296111057072
```

```
In [57]: file['INCIDENTCOUNT'].agg(np.sum)
```

```
Out[57]: 125038
```



Apresentar



Apreciar

wind map

<http://hint.fm/wind/>

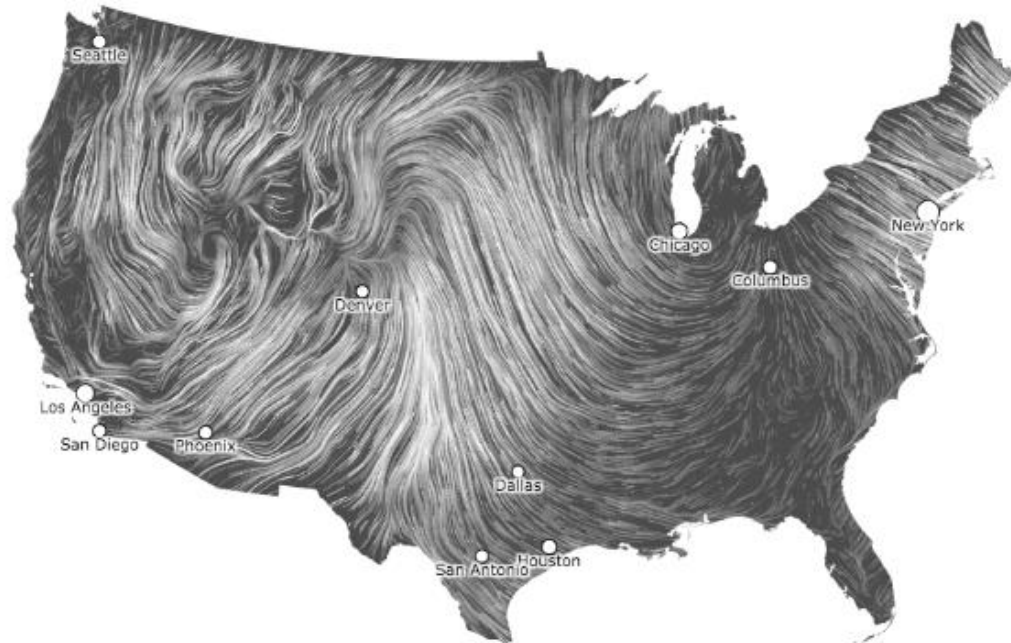
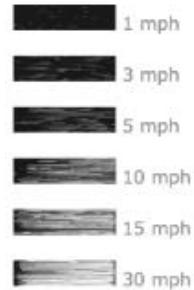
March 29, 2016

3:35 pm EST

(time of forecast download)

top speed: **33.8 mph**

average: **12.0 mph**



Ações (Verbos)

Em Análise, temos:

- Consumir:
 - Descobrir
 - Apresentar
 - Apreciar
- Produzir:
 - Anotar
 - Registrar
 - Derivar

➔ Analyze

➔ Consume

➔ Discover



➔ Present



➔ Enjoy



➔ Produce

➔ Annotate



➔ Record



➔ Derive



➔ Search

➔ Query

Anotar

A meta de anotação refere-se à adição de anotações gráficas ou textuais associadas a um ou mais elementos de visualização preexistentes, normalmente como uma ação manual do usuário. Quando uma anotação é associada a itens de dados, a anotação pode ser considerada um novo atributo para eles.

Por exemplo, o usuário pode anotar todos os pontos em um cluster com um rótulo de texto

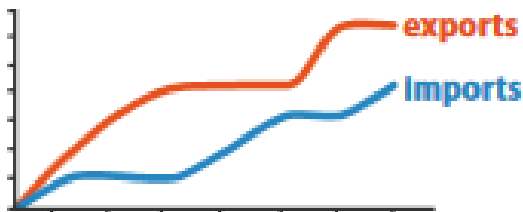
Registrar

- O objetivo do registro salva ou captura elementos de visualização como artefatos persistentes. Esses artefatos incluem capturas de tela, listas de elementos ou locais marcados, configurações de parâmetros, logs de interação ou anotação
- Em *Data Science* usamos o termo **reprodutibilidade**;
- Registrar é diferente da etapa anterior (anotar).



Derivar

- Não mostre apenas o que você recebeu!
 - Decida qual é a coisa certa a ser exibida;
 - Crie novas exibições a partir de transformações dos dados originais.
 - Esses recursos acima podem ajudar a lidar com a complexidade do problema.



Original Data



$$\text{trade balance} = \text{exports} - \text{imports}$$

Derived Data

Alvos (Substantivos)

- A ideia de um alvo é explícita nas pesquisas e nas consultas.
- Para todos os tipos de dados, temos:
 - *Trends* (ou *Patterns*)
 - *Outliers*
 - *Features*

→ ALL DATA

→ Trends



→ Outliers



→ Features



→ ATTRIBUTES

→ One

→ Distribution



↓ Extremes

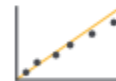


→ Many

→ Dependency



→ Correlation

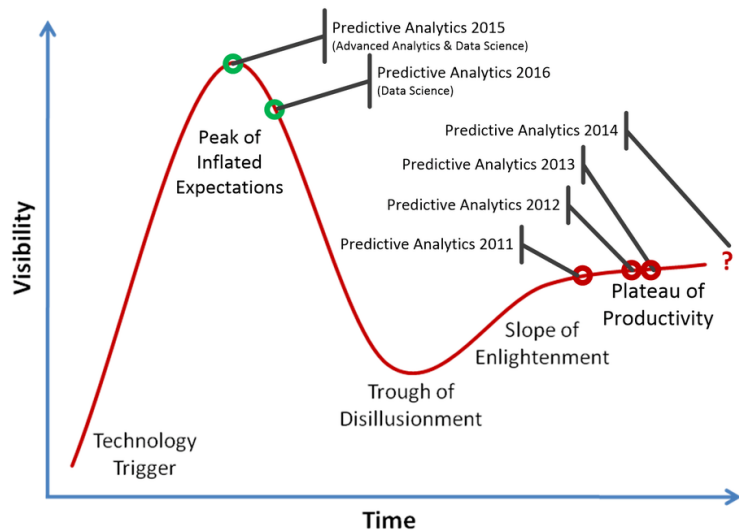


→ Similarity



Alvos

- *Trend (Pattern)*: Um *trend* é uma “tendência” é uma caracterização de alto nível de um padrão dos dados. Exemplos de *trends* incluem aumento, decréscimo, picos, *plateaus*...

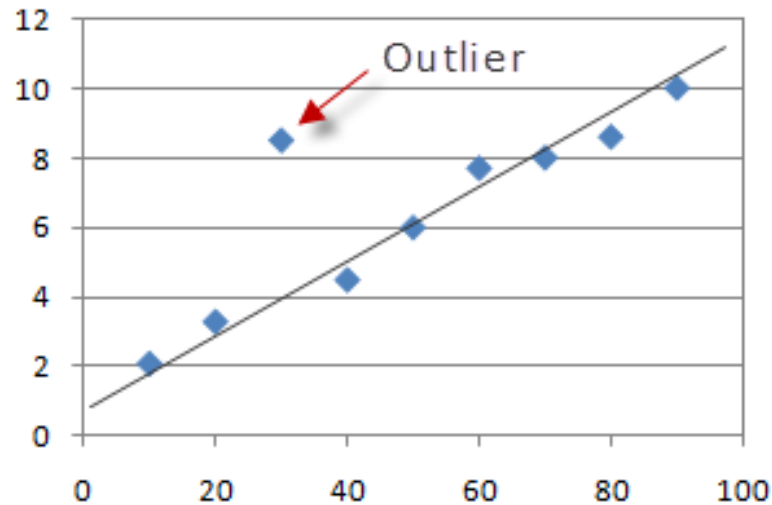


Exemplo:

<https://www.datasciencecentral.com/profiles/blogs/predictive-analytics-takes-a-victory-lap>

Alvos

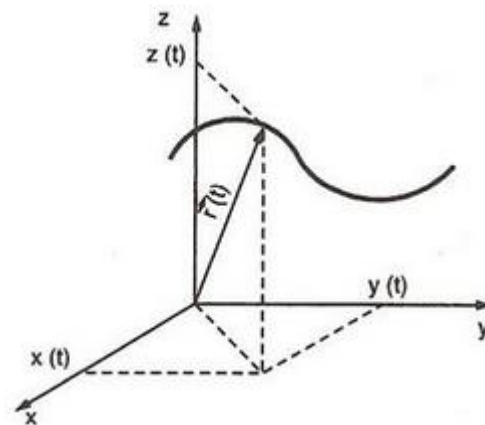
- *Outliers*: Elementos que não segue o padrão dos outros dados.



Alvos

- *Features*: características dos dados (que dão semântica a eles).

O gráfico ao lado mostra um dataset com duas variáveis independentes (features) x , z e uma independente y .



Alvos (Substantivos)

- Em atributos temos a busca de comportamentos e padrões através de técnicas exploradas na aula de Análise Exploratória de Dados.

→ ALL DATA

→ Trends



→ Outliers



→ Features



→ ATTRIBUTES

→ One

→ Distribution



↓ Extremes

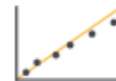


→ Many

→ Dependency



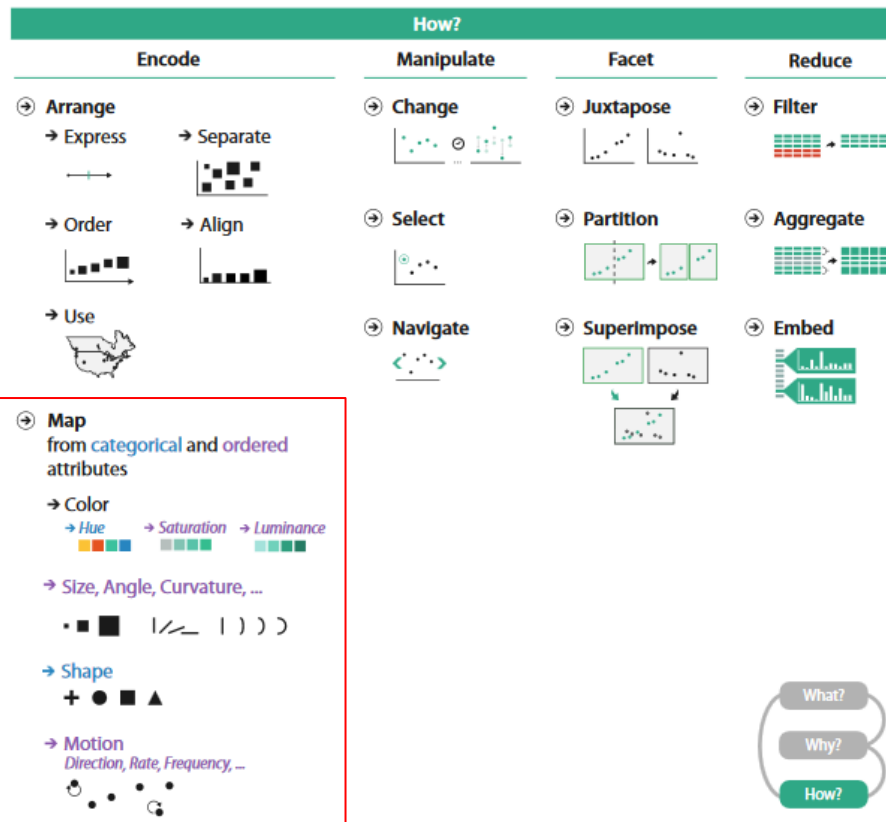
→ Correlation



→ Similarity



Codificação Visual: Visão Geral



Este tópico foi
abordado em
Marcadores e
Canais!



Introdução

Este tópico visa orientar na escolha de design para organizar dados tabulares espacialmente.

→ Express Values



→ Separate, Order, Align Regions

→ Separate



→ Order



→ Align



→ Axis Orientation

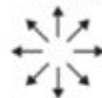
→ Rectilinear



→ Parallel

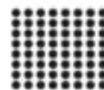


→ Radial



→ Layout Density

→ Dense

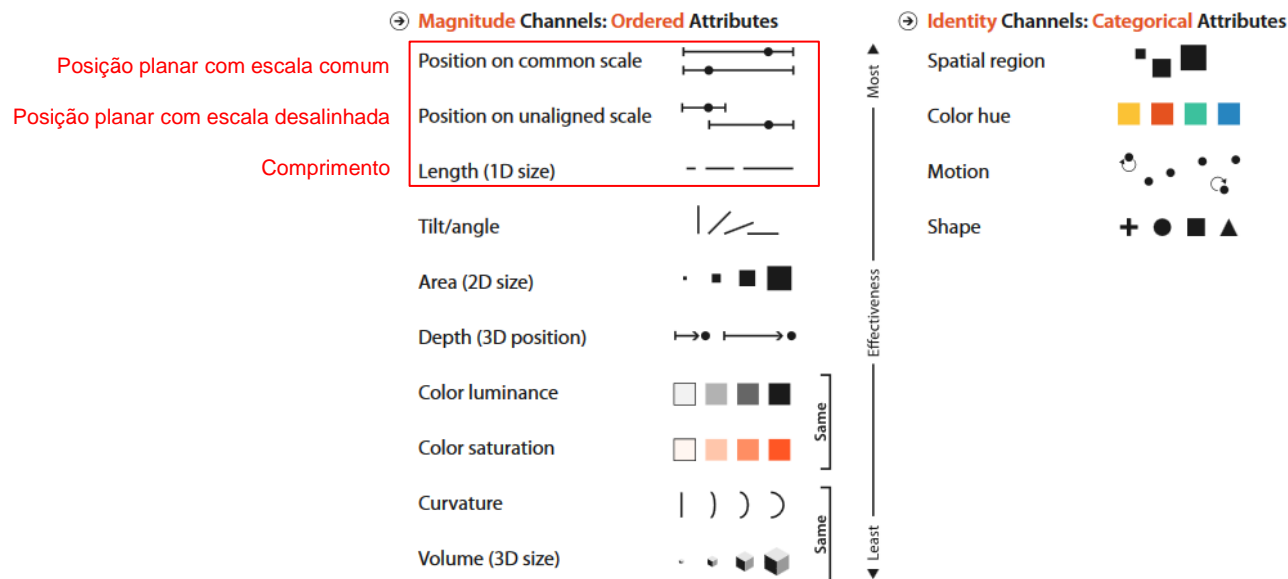


→ Space-Filling



Introdução: Por que organizar espacialmente?

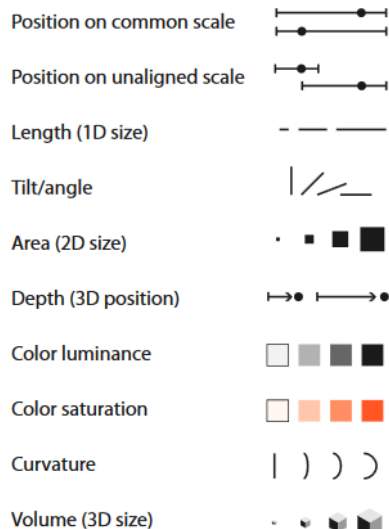
- O uso de espaço domina o modelo mental do usuário para o dataset;
 - Os três canais melhor ranqueados com respeito à efetividade para atributos **quantitativos e ordenados** são todos relacionados à posição espacial, lembra?



Introdução: Por que organizar espacialmente?

- O uso de espaço domina o modelo mental do usuário para o dataset;
 - O canal melhor ranqueado para **atributos categóricos** também é relacionado ao espaço!

➔ Magnitude Channels: Ordered Attributes



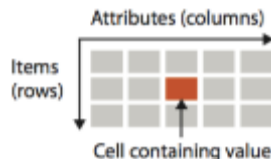
➔ Identity Channels: Categorical Attributes



Agrupamento de itens em uma mesma região

Organizando Dados por CHAVES e VALORES

- A distinção entre CHAVE e VALOR é muito relevante para a codificação visual de dados tabulares;
- Uma **CHAVE** é um atributo independente que pode ser usado como ÍNDICE para identificar itens em uma tabela;
 - Podem ser categóricos ou ordinais.
- Um é um atributo dependente: o valor de uma célula em uma tabela;
 - Podem ser categóricos, ordinais ou quantitativos.
- Os diferentes valores para um atributo categórico ou ordinal são chamados de níveis.

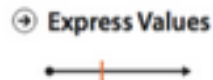


Organizando Dados por CHAVES e VALORES

- As **escolhas de design** vão depender da semântica dos atributos das tabelas: **Quantas chaves e quantos valores existem na tabela?**
 - Idiomas que mostram somente valores, sem chave: Scatterplot;
 - Idiomas que mostram uma chave e um valor: Barras;
 - Idiomas que mostram duas chaves e um valor: Mapa de Calor;
 - Idiomas para muitas chaves e muitos valores: Matriz de Scatterplot.
- Normalmente chaves são usadas para definir uma região no espaço para cada item na qual os atributos-valores são mostrados.

Expressando Valores Quantitativos

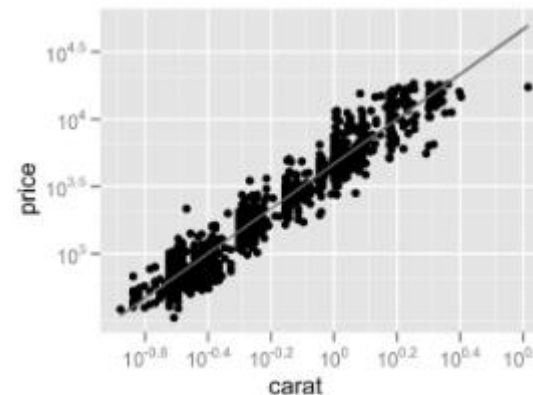
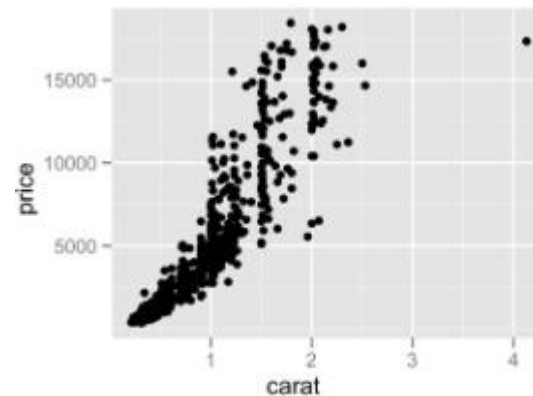
- Mapeamento direto de valores de atributos em posição espacial;
 - Posicionamento de uma marcação ao longo do eixo.



- Outros atributos podem ser mapeados na mesma marcação usando outros canais não espaciais (cor, tamanho ou um *glyph*).
 - Exemplos: Scatterplot, gráfico de bolha.

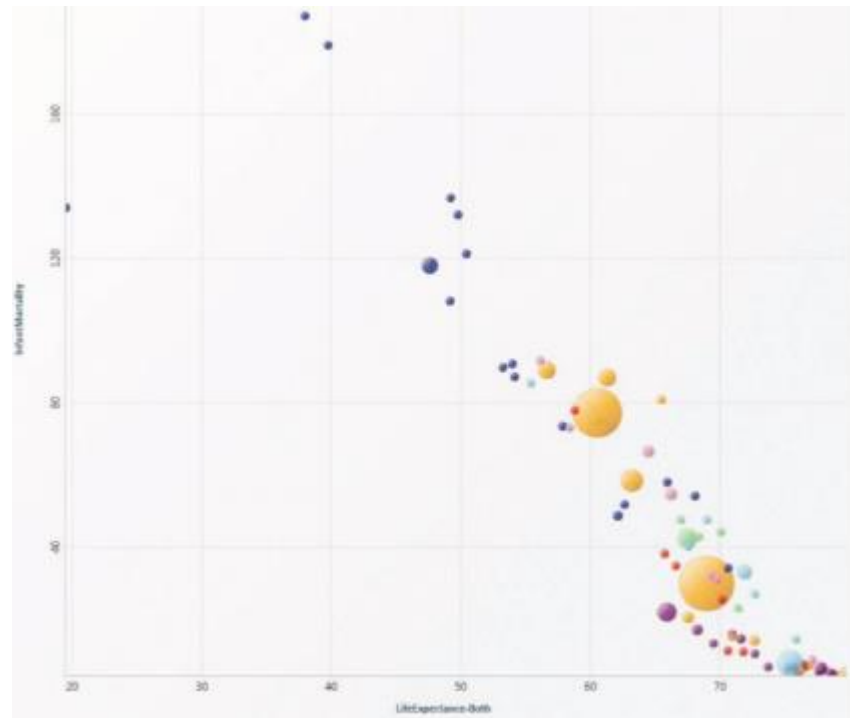
Idioma: *Scatterplots*

- **Sem chaves, apenas valores:**
 - Dados:
 - 2 atributos quantitativos.
 - Marcação: Pontos;
 - Canais:
 - Posição Horizontal e Vertical;
 - Tarefas:
 - Achar correlação, *outlier*, distribuição, cluster, tendência.
 - Escala: Centenas de itens



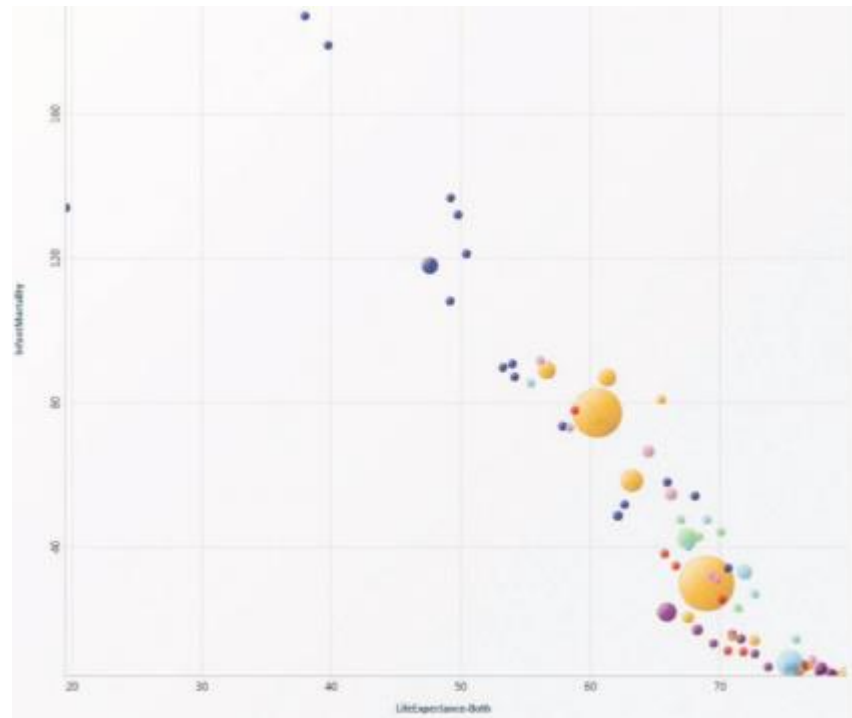
Idioma: Gráfico de Bolha

- **Sem chaves, apenas valores;**
 - Dados:
 - 2 atributos quantitativos.
 - Marcação: Pontos;
 - Canais:
 - Posição Horizontal e Vertical;
 - Tamanho;
 - Cor
 - Tarefas:
 - Achar correlação, *outlier*, distribuição, cluster, tendência.
 - Escala: Centenas de itens



Idioma: Gráfico de Bolha

- Gráfico de Bolha é um *Scatterplot* aumentado com uso do tamanho e cor;
- Na figura ao lado: cada marcação representa um país, com as posições horizontal e vertical codificando os atributos quantitativos de expectativa de vida e mortalidade infantil, a cor representa um atributo categórico (região) e o tamanho codifica a população.



Regiões Categóricas

→ Separate



→ Order



→ Align

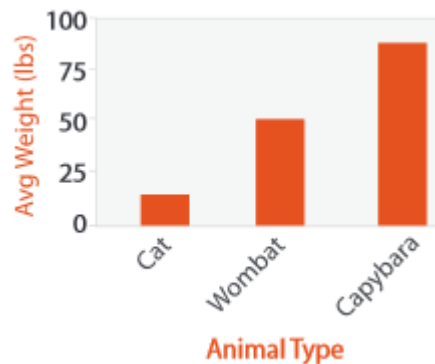
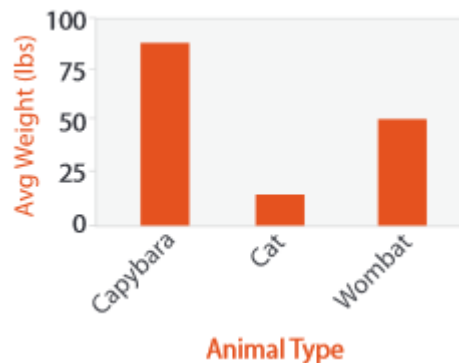


- O uso do espaço para codificar atributos categóricos é mais complexo do que no caso de atributos quantitativos;
 - Posição espacial é um canal visual de magnitude ordenada, mas atributos categóricos não possuem ordem.
 - Usar espaço para **separar** os dados em **regiões**.

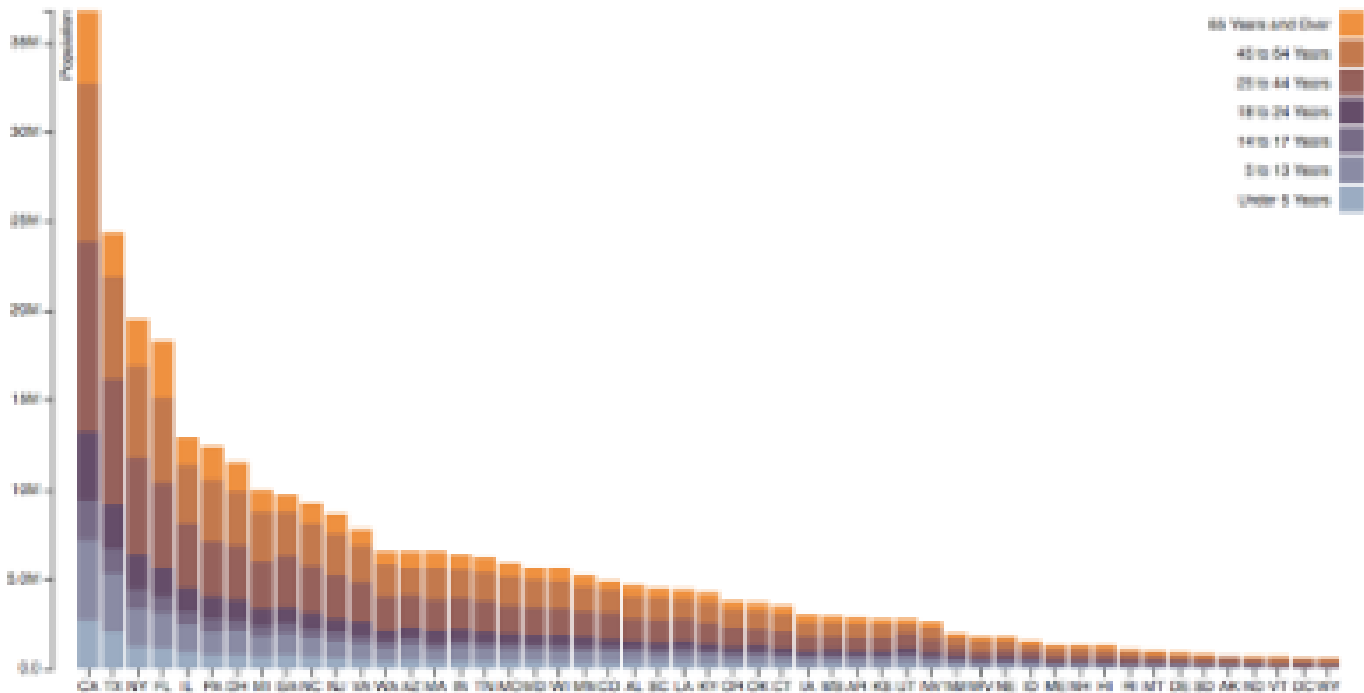
Idioma: Gráfico de Barras

- **Uma chave, um valor:**

- Dados:
 - 1 atributo categórico e 1 atributo quantitativo;
- Marcação: Linhas;
- Canais:
 - Comprimento para expressar valor quantitativo;
 - Regiões espaciais: uma por marcação:
 - Separadas horizontalmente, alinhadas verticalmente;
 - Ordenadas por atributos quantitativo: pelo label (alfabeticamente) ou pelo atributo quantitativo.
- Tarefas:
 - Comparar, encontrar valores.
- Escala: Dezenas a centenas de níveis por atributo-chave.



Idioma: Gráfico de Barras Empilhadas

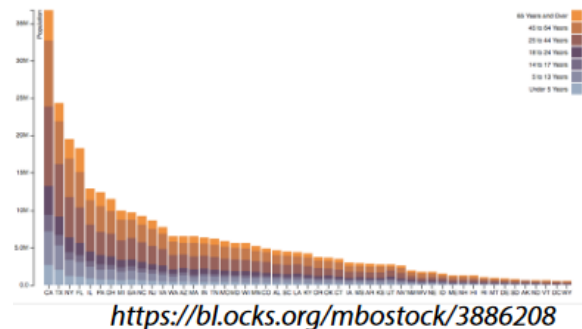


<https://bl.ocks.org/mbostock/3886208>

Idioma: Gráfico de Barras Empilhadas

- **Uma chave ou mais:**

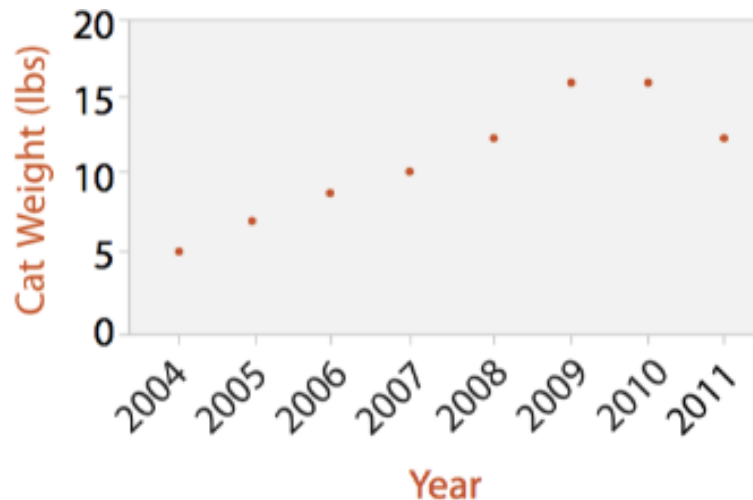
- Dados:
 - 2 atributos categóricos e 1 atributo quantitativo;
- Marcação: Pilha vertical de Marcação de Linhas;
- Canais:
 - Comprimento de Matiz e Cor;
 - Regiões Espaciais: Uma por Glyph;
 - Alinhadas: glyph completo, componentes mais baixo da barra;
 - Não alinhadas: os demais componentes da barra.
- Tarefas:
 - Relacionamentos parte-do-todo.
- Escala: Até uma dúzia de níveis por atributos empilhado.



Idioma: Gráfico de Pontos

- **Uma chave, um valor:**

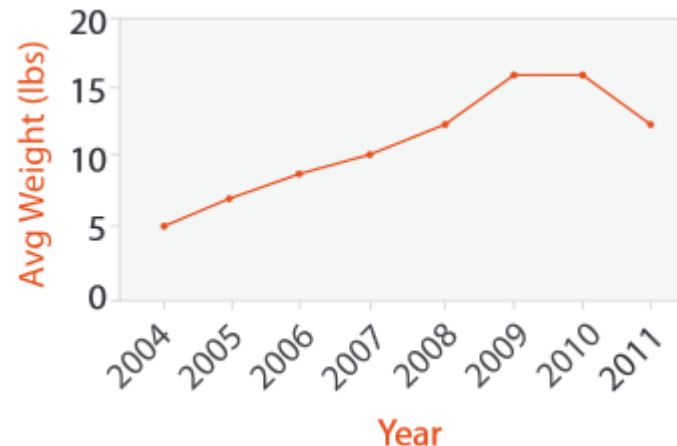
- Dados:
 - 1 atributo chave ordenado e 1 atributo valor quantitativo.
- Marcação: Pontos
- Canais:
 - Comprimentos alinhados para expressar valor quantitativo;
 - Separados e ordenados pelo atributo-chave em regiões horizontais.
- Tarefas:
 - Encontrar e comparar valores.
- Escala: Dezenas a centenas de níveis por atributo-chave.



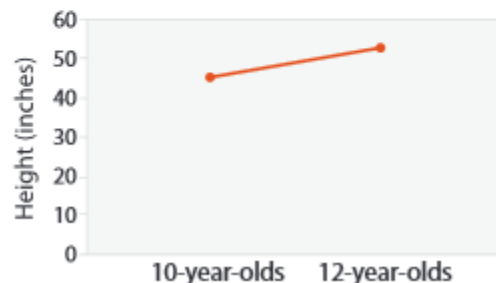
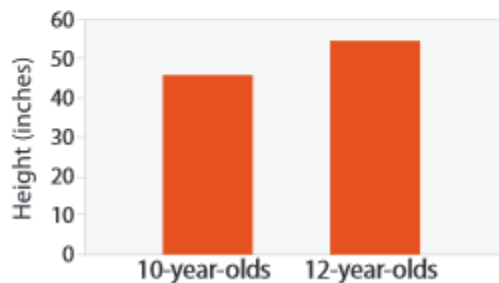
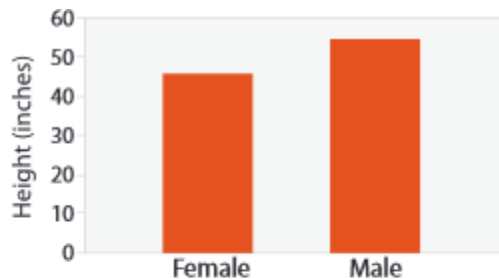
Idioma: Gráfico de Linhas

- **Uma chave, um valor:**

- Dados:
 - 1 atributo chave ordenado e 1 atributo valor quantitativo.
- Marcação: Conexão de marcações em linhas entre pontos.
- Canais:
 - Comprimentos alinhados para expressar valor quantitativo;
 - Separados e ordenados pelo atributo-chave em regiões horizontais.
- Tarefas:
 - Achar tendências.
 - As linhas enfatizam a ordem dos itens e o relacionamento entre itens.
- Escala: Dezenas a centenas de níveis por atributo-chave.



Gráficos de Barras X Gráfico de Linhas

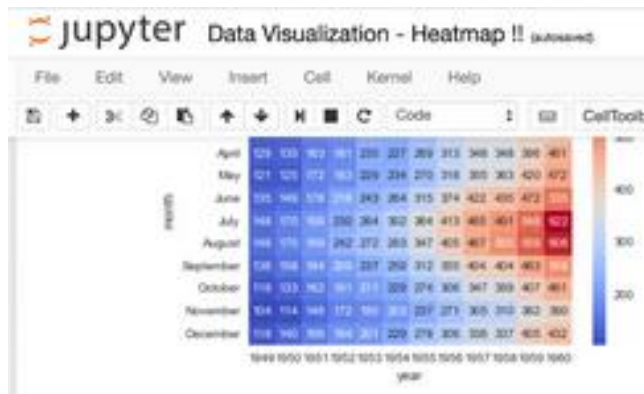


after [Bars and Lines: A Study of Graphic Communication. Zacks and Tversky. Memory and Cognition 27:6 (1999), 1073–1079.]

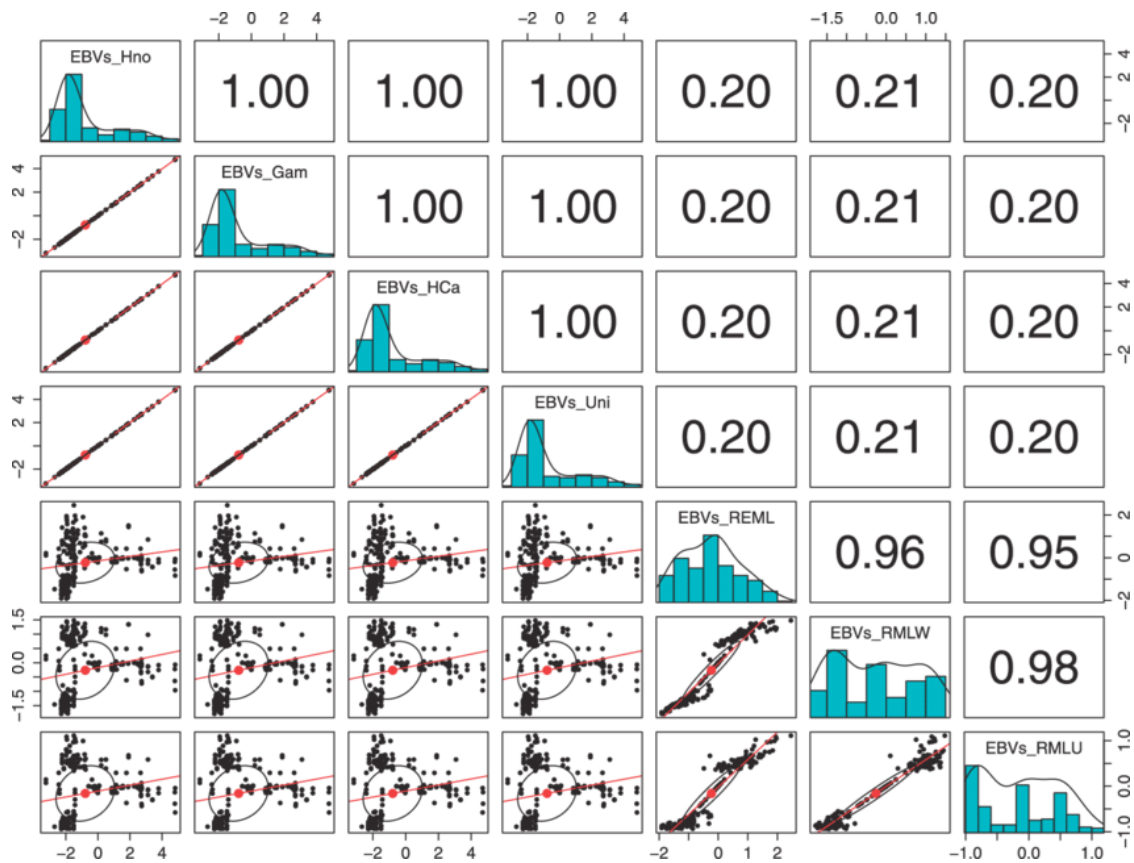
- Depende do tipo do atributo-chave:
 - Categórico: Barras;
 - Ordenados: Linhas
- Não usar gráfico de linhas para atributos-chave categóricos.
 - Violação do princípio da expressividade:
 - Implicação de tendência é tão forte que sobreescreve a semântica.
 - Ex: quanto mais masculina a pessoa for, mais alta ela é.

Idioma: Mapas de calor (*heatmaps*)

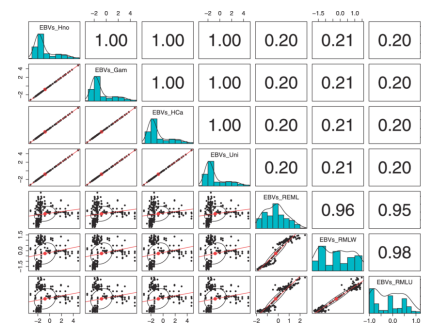
- Duas chaves, um valor:
 - Dados:
 - 2 atributos categóricos;
 - 1 atributo quantitativo.
- Marcação: Área.
 - Separar e alinhar em matriz 2D;
 - Indexados pelos 2 atributos categóricos.
- Canais:
 - Cor pelo atributo quantitativo (ordenado usando mapa de cor divergente)
- Tarefa: encontrar clusters, outliers
- Escala: Milhares de itens, centenas de níveis categóricos, aprox. 10 níveis de atributos quantitativos.



Idioma: Matriz de *Scatterplots* (SPLOM)



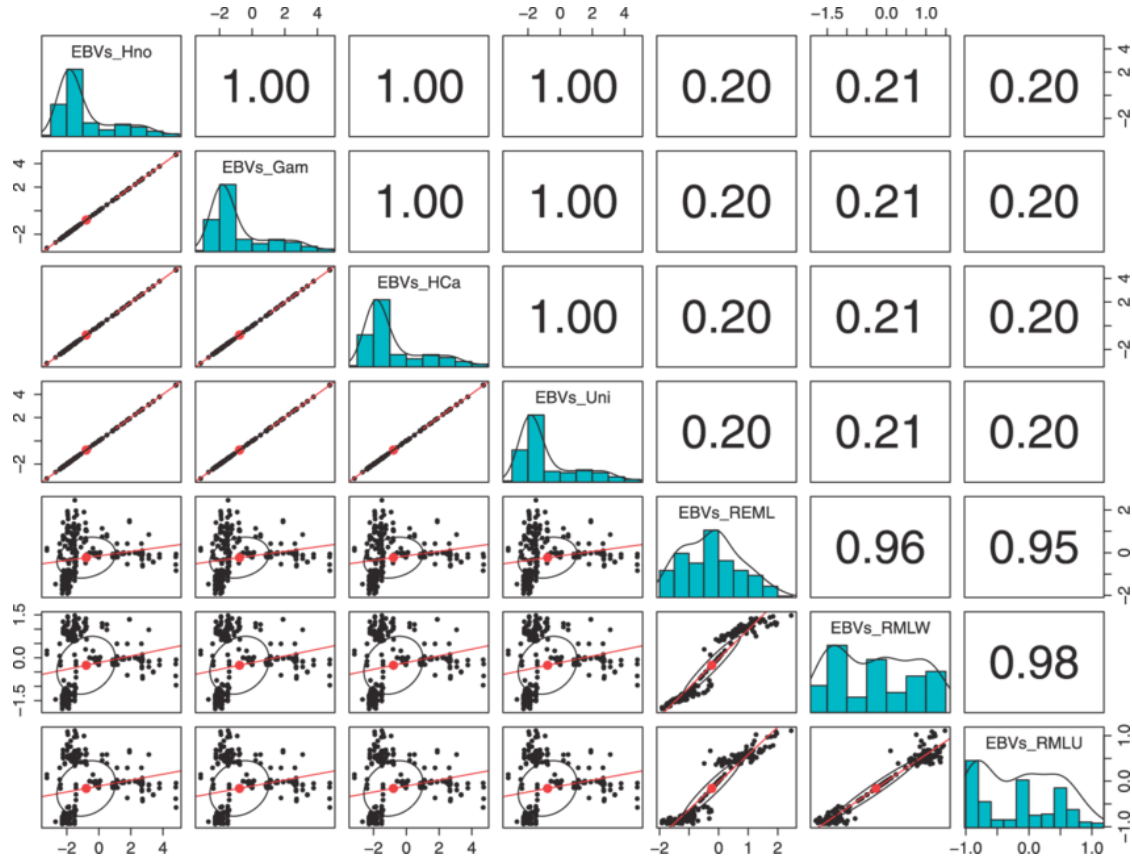
Idioma: Matriz de *Scatterplots* (SPLOM)



- Uma matriz onde cada célula contém um gráfico de *scatterplot* completo;
- Uma SPLOM mostra todas as combinações de pares de atributos;
- A chave é um atributo derivado simples que é o mesmo para linhas e colunas;
- Normalmente, apenas um triângulo da matriz é mostrado (devido à simetria) com os nomes dos atributos na diagonal.
 - Tarefa: encontrar correlações, tendências e outliers;
 - Escala: Aproximadamente uma dúzia de atributos e centenas de itens

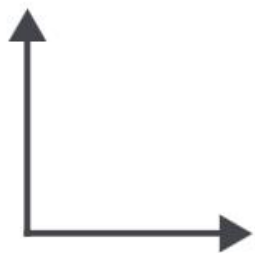
Idioma: Matriz de *Scatterplots* (SPLOM)

- Diagonal principal pode vir apenas com os rótulos, mas também pode trazer a distribuição dos seus dados;
- Como a matriz é simétrica, uma ideia interessante é trazer na diagonal superior os valores das correlações (numericamente falando).



Orientação Espacial de Eixo

→ Rectilinear

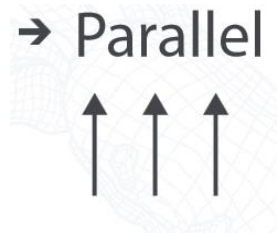


→ Parallel



Idioma: Coordenadas Paralelas

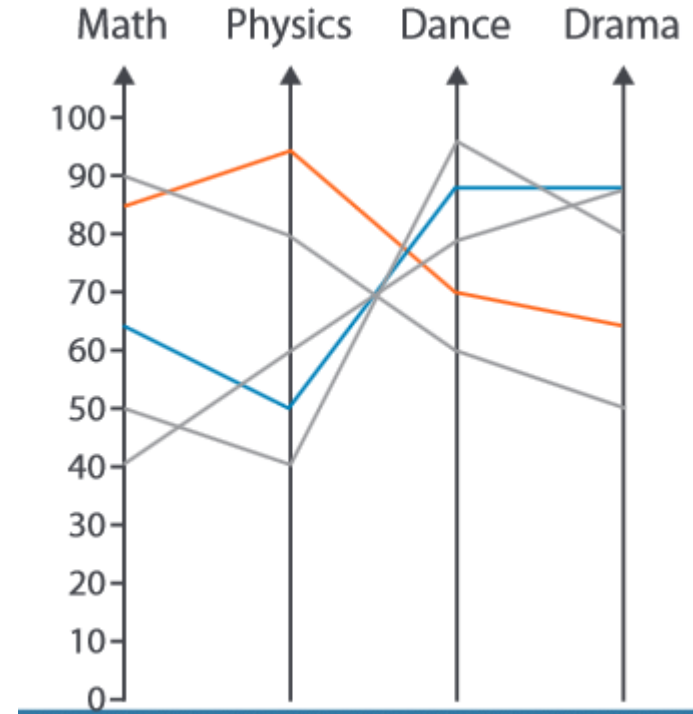
- Visualizar muitos atributos quantitativos de uma vez usando posição espacial.
 - Dados:
 - Vários atributos valor
 - Layout paralelo:
 - Posição espacial horizontal usada para separar eixos;
 - Posição espacial vertical expressa o valor do atributo ao longo de cada eixo alinhado;
 - Marcações em linhas conectam os atributos de um mesmo item.
 - Tarefa:
 - Encontrar tendências, outliers, extremos, correlação.
 - Escala:
 - Dezenas de atributos ao longo do eixo secundário e centenas de itens.



Idioma: Coordenadas Paralelas

Table

Math	Physics	Dance	Drama
85	95	70	65
90	80	60	50
65	50	90	90
50	40	95	80
40	60	80	90



Idioma: Coordenadas Paralelas - Correlação

Em coordenadas paralelas, o comportamento das linhas nos diz sobre a correlação entre os atributos:

- **Correlação positiva:** segmentos de retas paralelas;
- **Correlação negativa:** todos os segmentos se cruzam no meio;
- **Sem correlação:** cruzamentos espalhados.

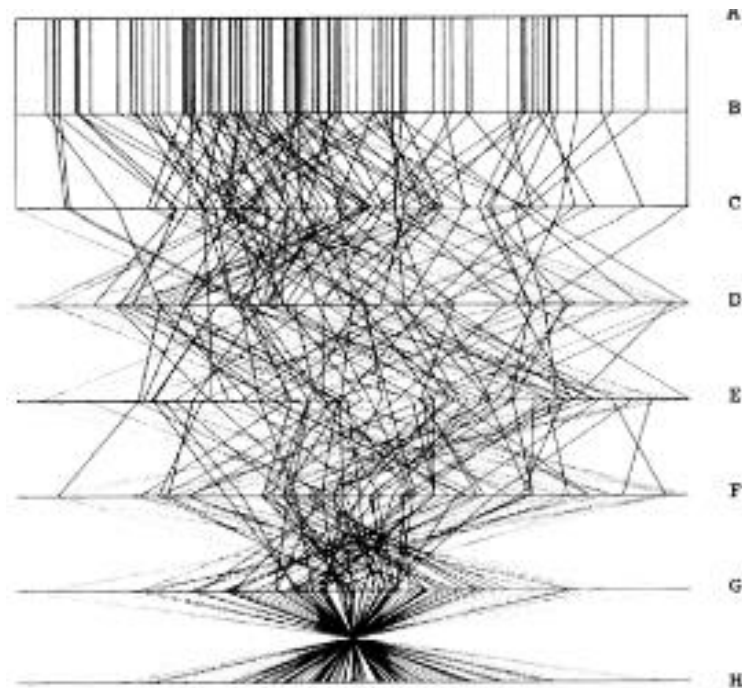
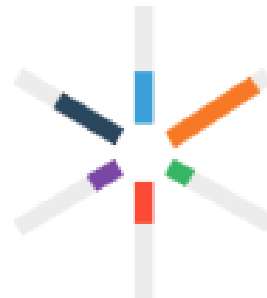


Figure 3. Parallel Coordinate Plot of Six-Dimensional Data Illustrating Correlations of $\rho = 1, .8, .2, 0, -.2, -.8, \text{ and } -1$.

Idioma: Gráfico de Barras Radiais

- Gráfico de Barras Radiais:
 - Eixos radiais se encontram no anel central;
 - Marcação em linhas.
- Acurácia:
 - Comprimento desalinhado com posição radial.
 - Menos acurado do que alinhado com posição retilínea.



Idiomas: Pie Chart, Polar Area Chart

- **Pie Chart:**

- Marcações de área com canal de ângulo;
- Acurácia: ângulo/ área muito menos acurado do que comprimento de linha.

- **Polar Area Chart:**

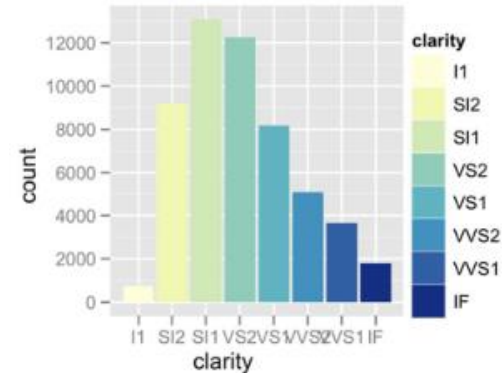
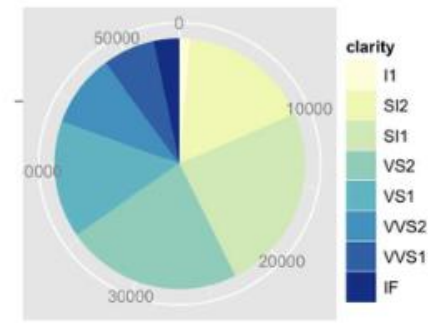
- Marcação área com canal comprimento.
- Analogia mais direta com gráfico de barras

- **Dados:**

- 1 Atributo chave categórico,
- 1 Atributo valor quantitativo.

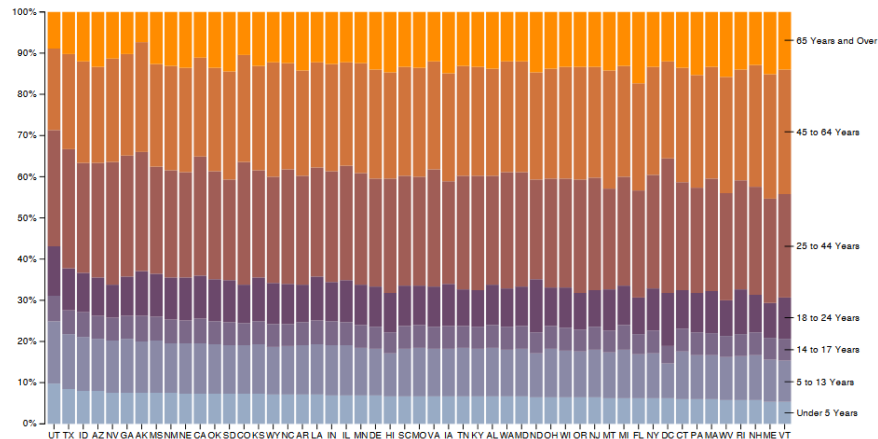
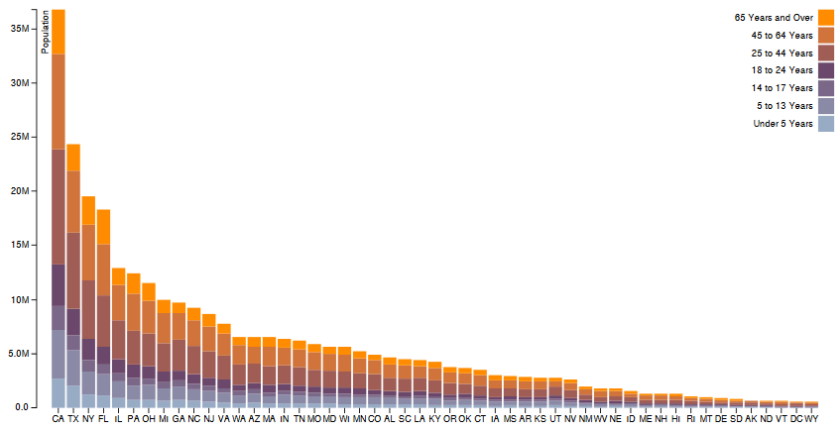
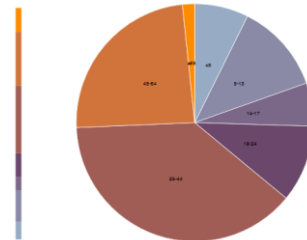
- **Tarefa:**

- Julgamento parte do todo.

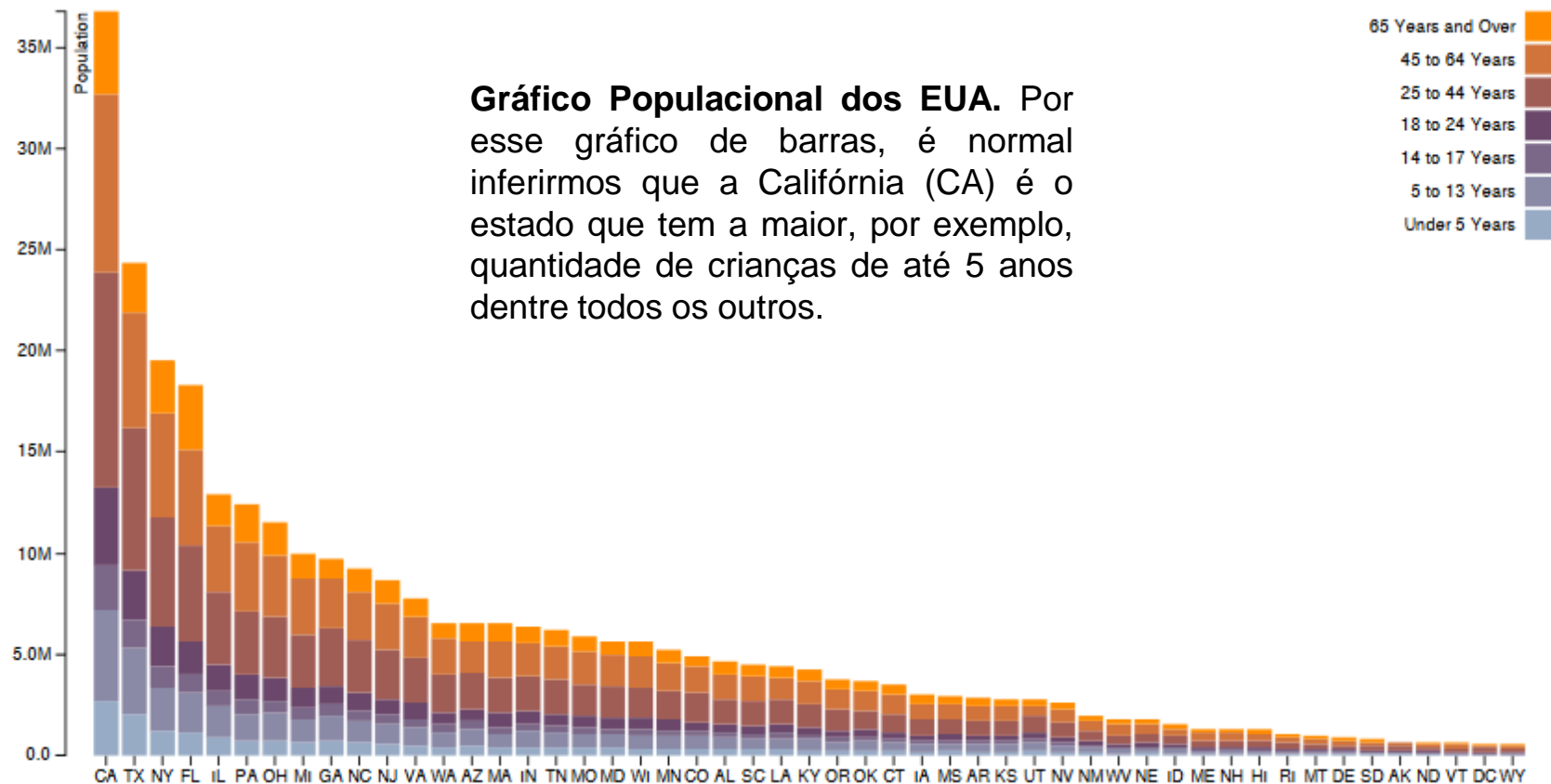


Idiomas: Gráfico de Barras Empilhadas

- Tarefa:
 - Julgamento parte do todo!
- Gráfico de barras Empilhadas normalizadas:
 - Normalizadas para o tamanho máxima da barra;
 - Uma única barra é equivalente a um Pie Chart completo.
 - Alta densidade de informação requer retângulo estreito.

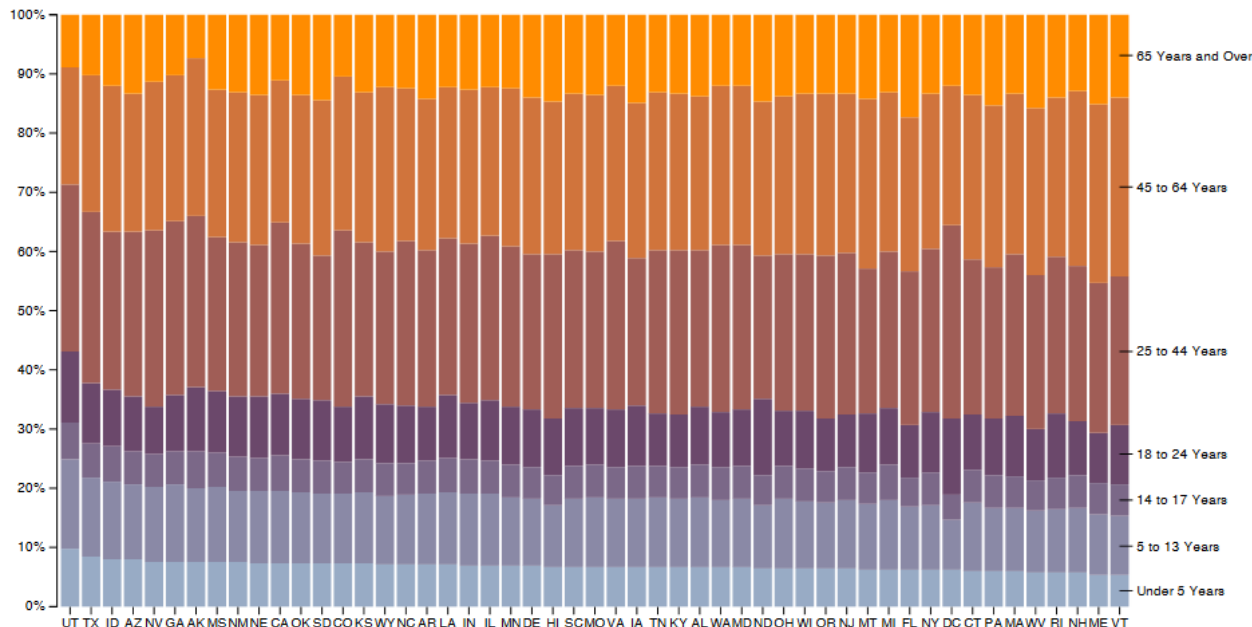


Idiomas: Gráfico de Barras Empilhadas



Idiomas: Gráfico de Barras Empilhadas

Ao normalizarmos cada estado, entretanto, podemos ver que Utah (UT) é o estado que, proporcionalmente falando, possui mais crianças de até 5 anos comparada às outras faixa etária. Califórnia apenas tem a maior quantidade absoluta, mas é relativamente menor comparada à outras.



Limitações de Orientação

- **Retilíneo:** escala com relação ao número de eixos:
 - 2 eixos: Melhor caso
 - 3 eixos: Problemático
 - 4+ eixos: impossível
- **Paralelo:** não é familiar, requer treinamento.
- **Radial:** limites perceptivos
 - Assimetria: menor precisão com ângulos do que com comprimentos.

➔ Axis Orientation

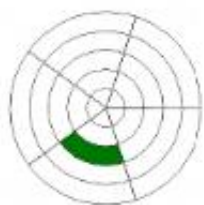
→ Rectilinear



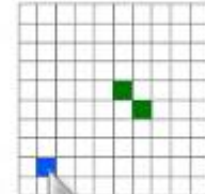
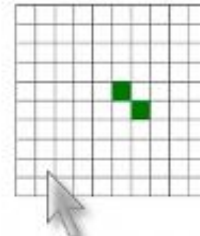
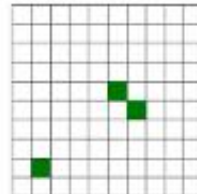
→ Parallel



→ Radial



clicked at wrong cell



clicked at correct cell

Uma região define um conjunto de células e isso fica impreciso! Os pontos em verde na matriz foram convertidos para região da primeira figura.

Visão Geral

- Algumas combinações de Marcadores e Canais já são clássicas no *estado-da-arte*; segundo Evergreen¹, a seguir, vamos mostrar algumas que podem auxiliar no seu processo de visualização;

¹ “Effective Data Visualization: The Right Chart for the Right Data” by Stephanie Evergreen

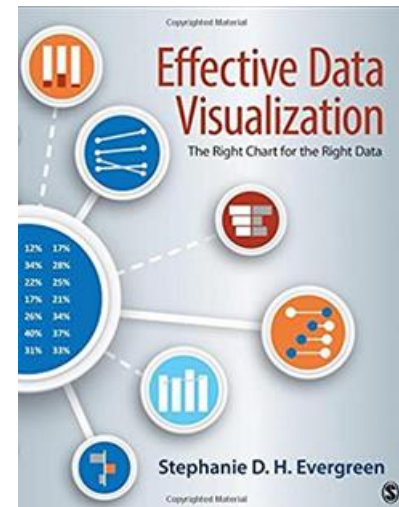


Chart Chooser Cheat Sheet

