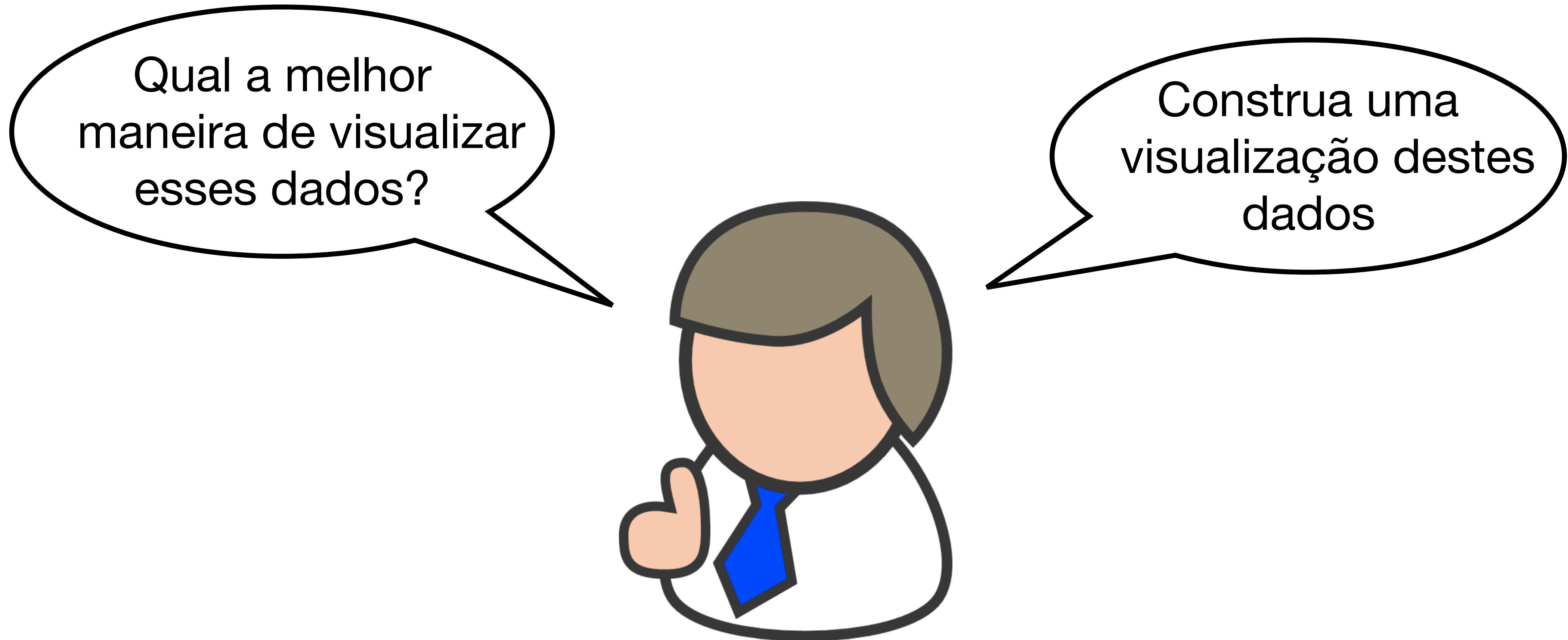


Analise Visual de Dados

Princípios de Visualização de Dados

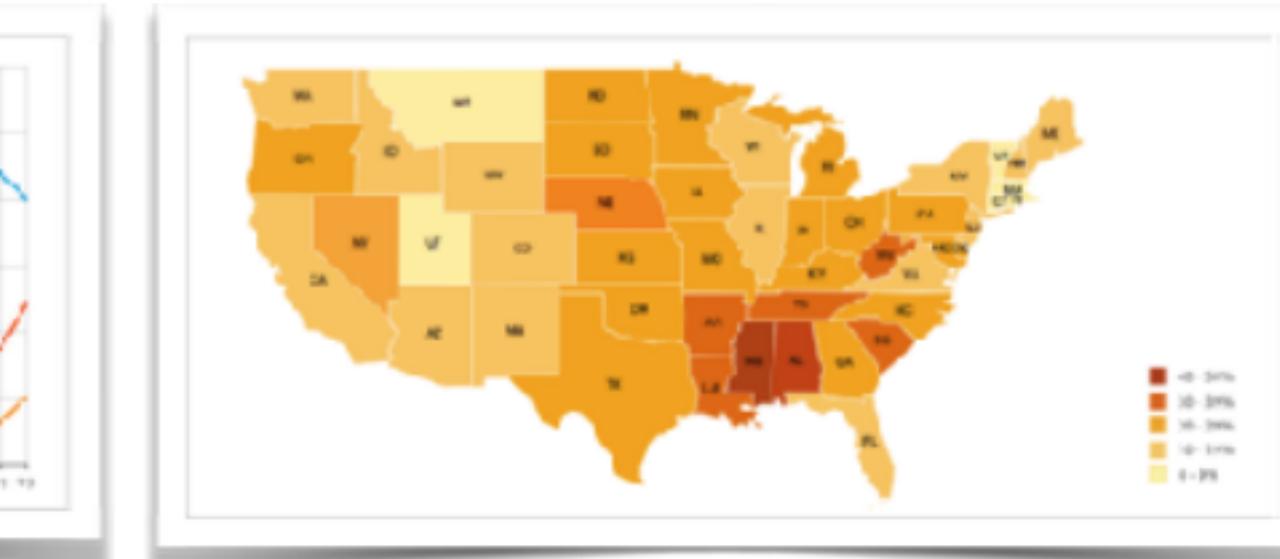
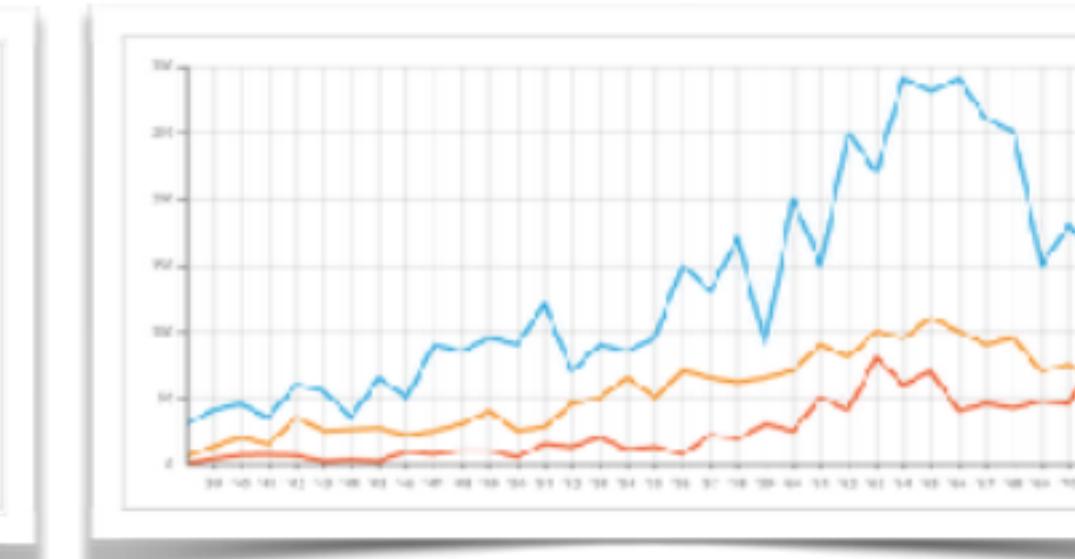
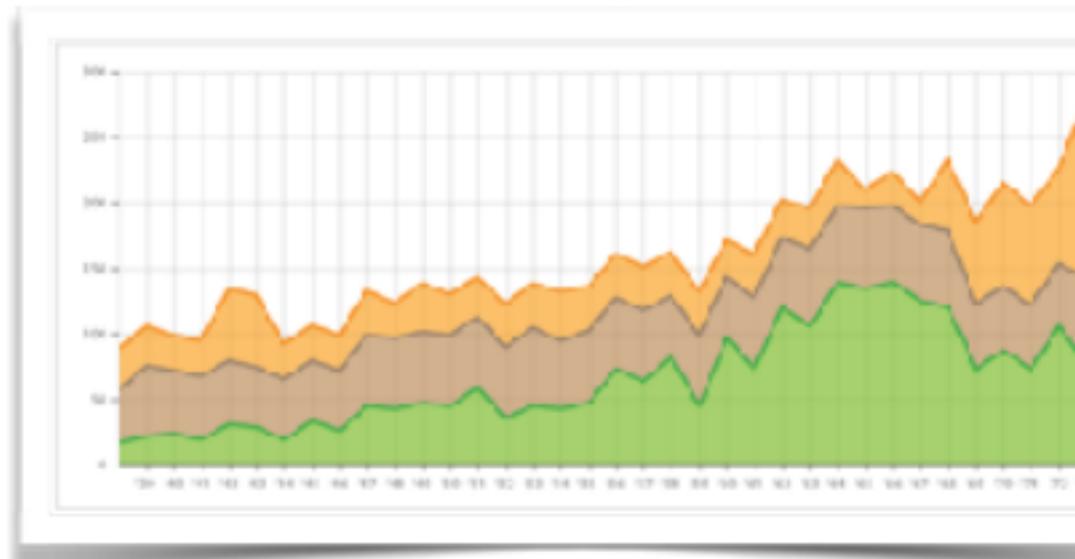
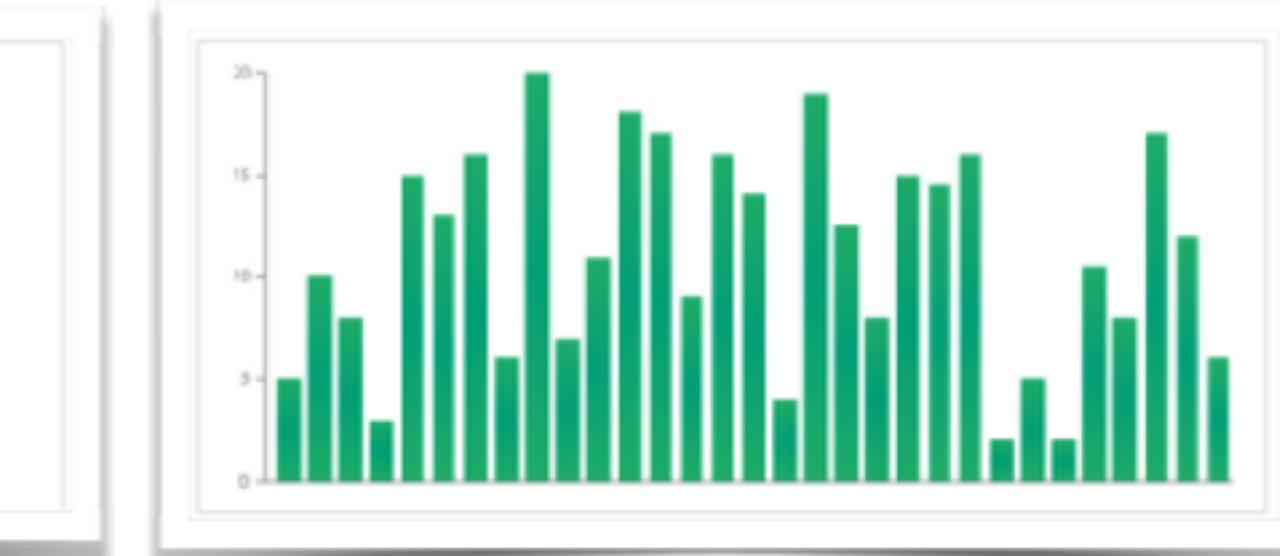
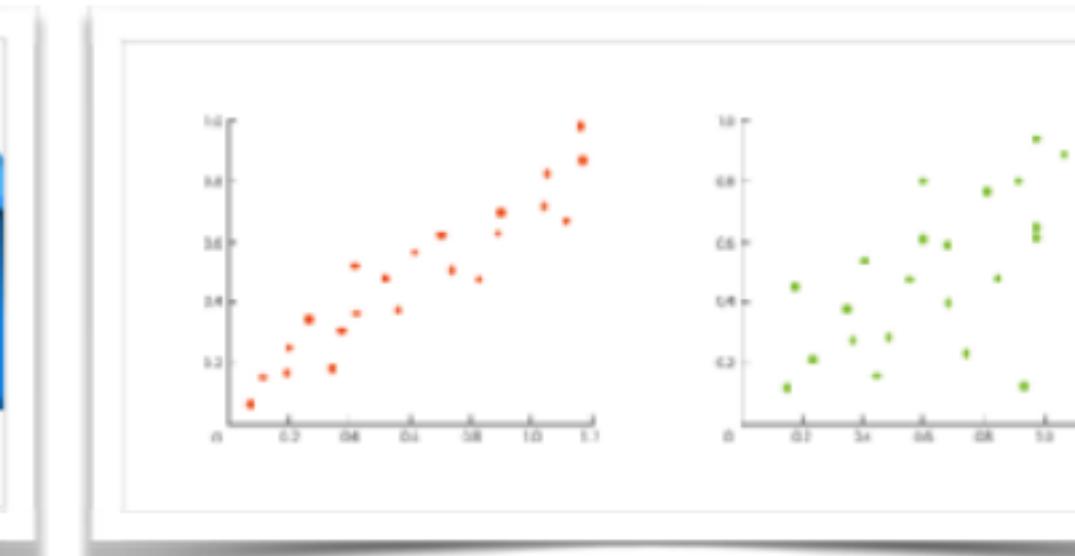
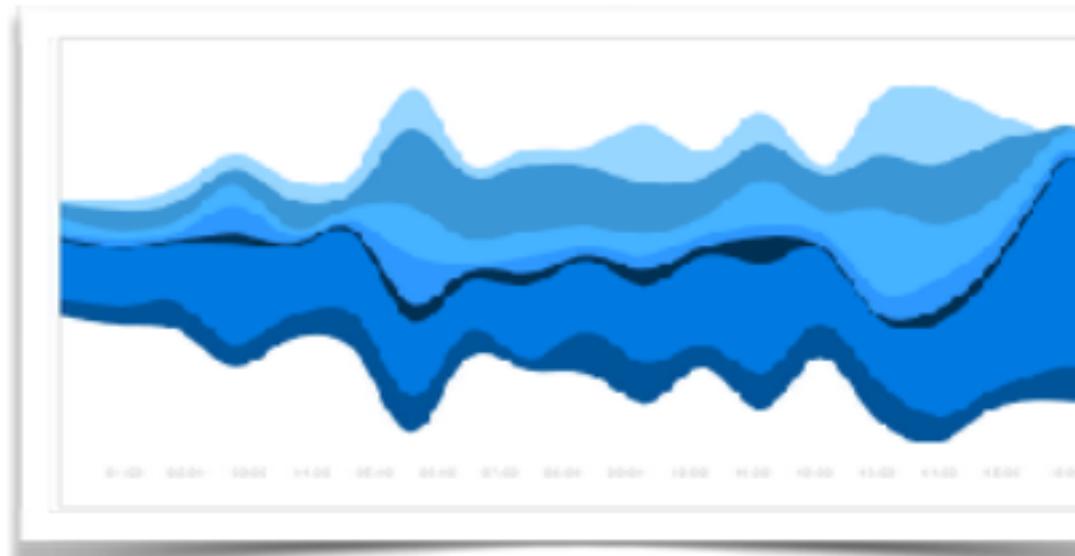
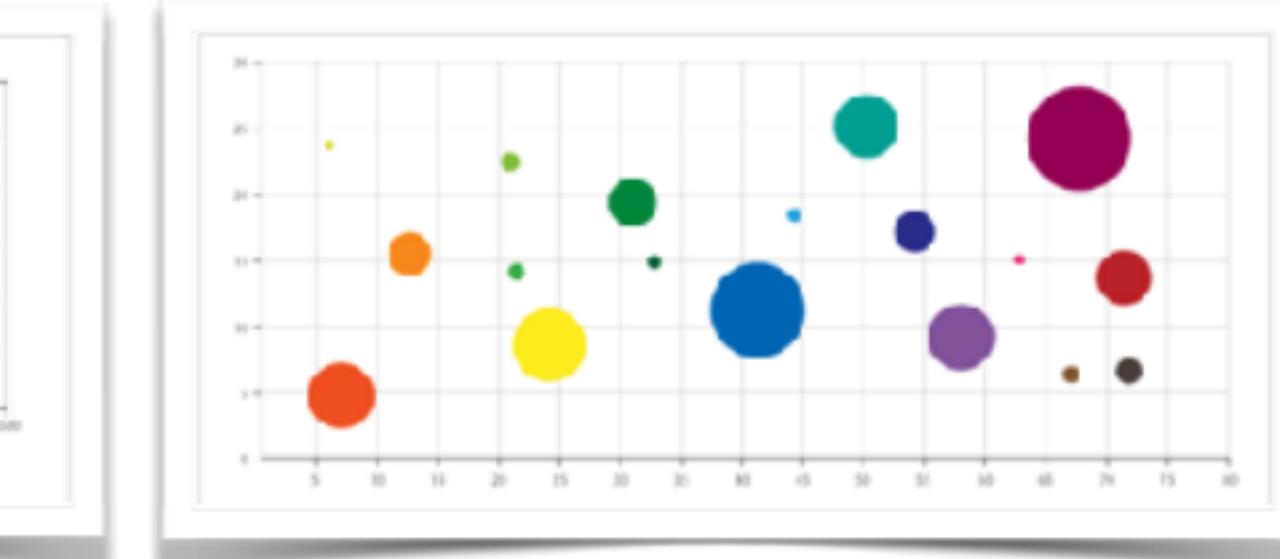
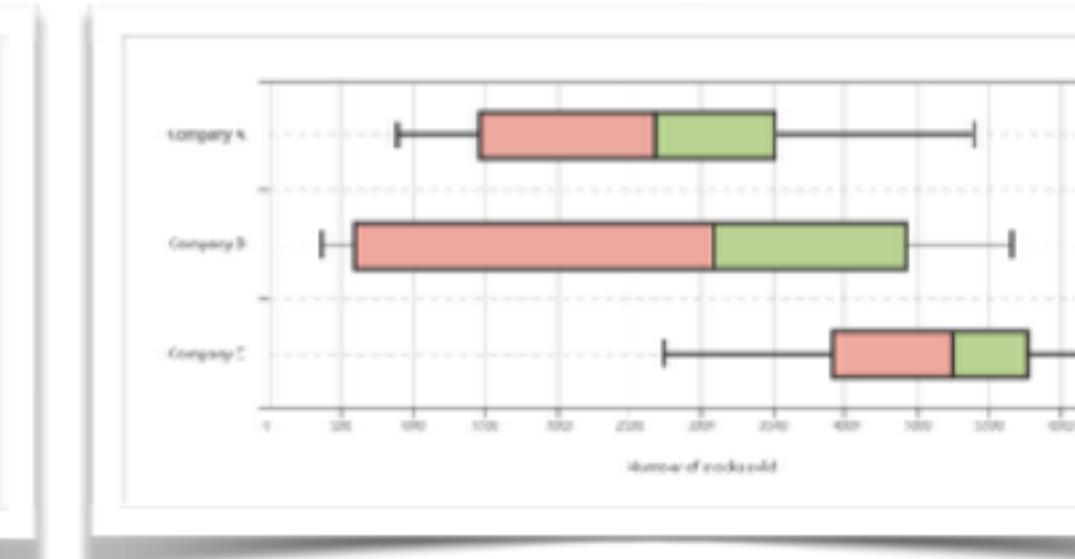
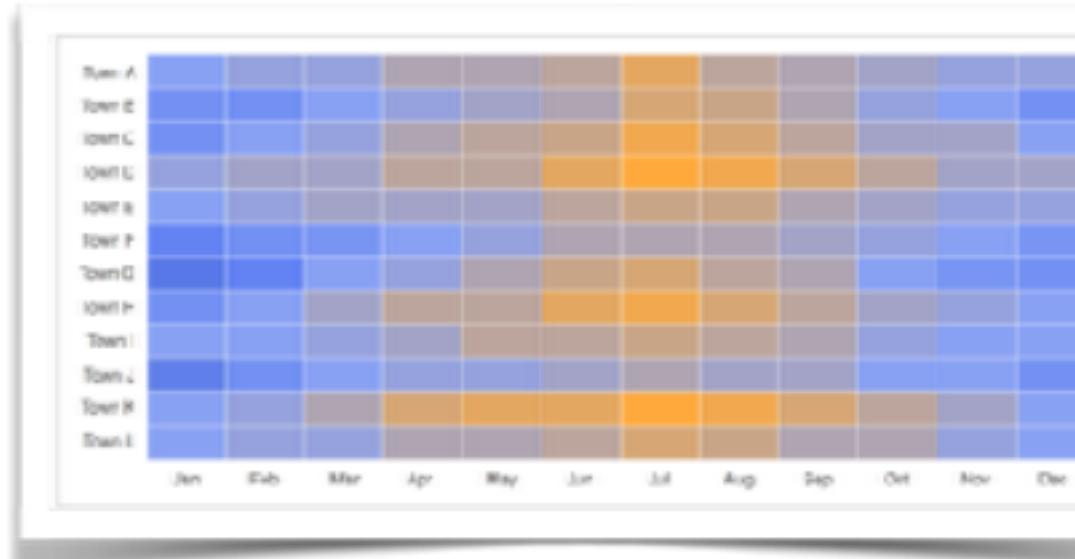
Cenário Recorrente



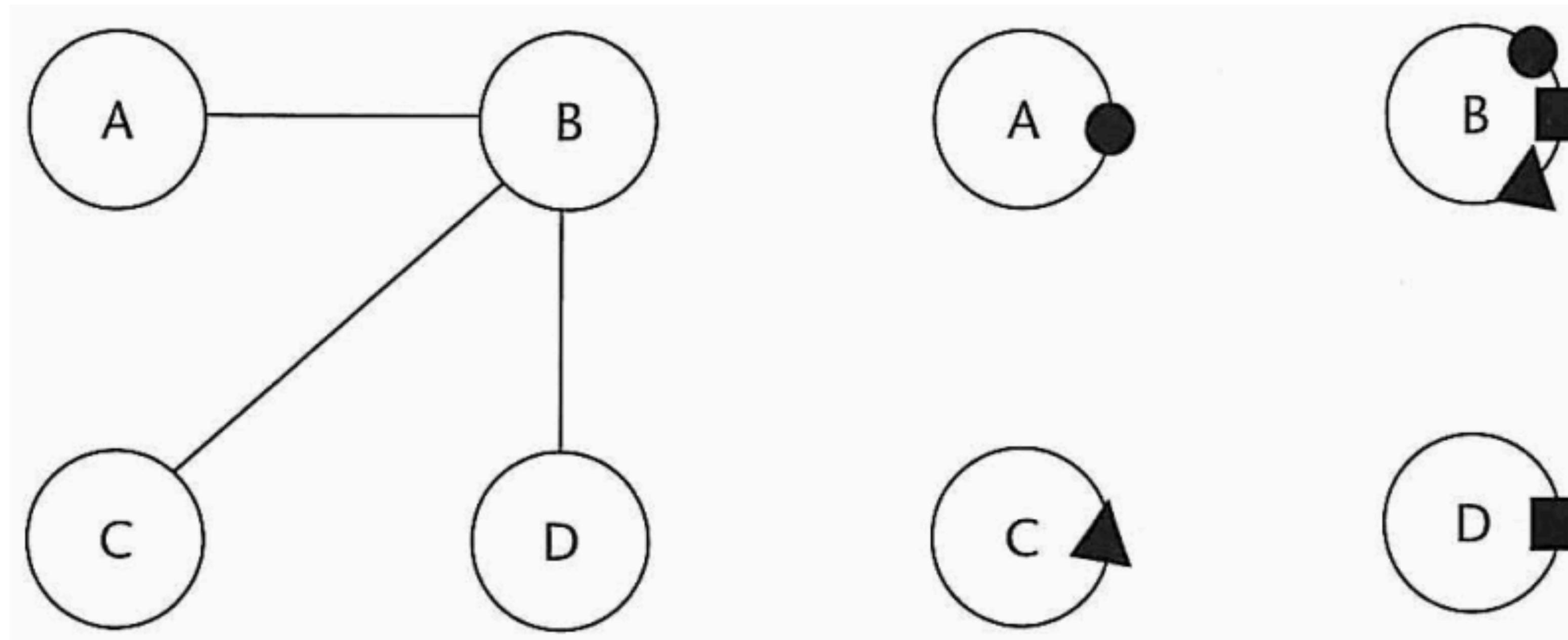
O espaço de possibilidades de design é enorme!



O espaço de possibilidades de design é enorme!



O espaço de possibilidades de design é enorme!



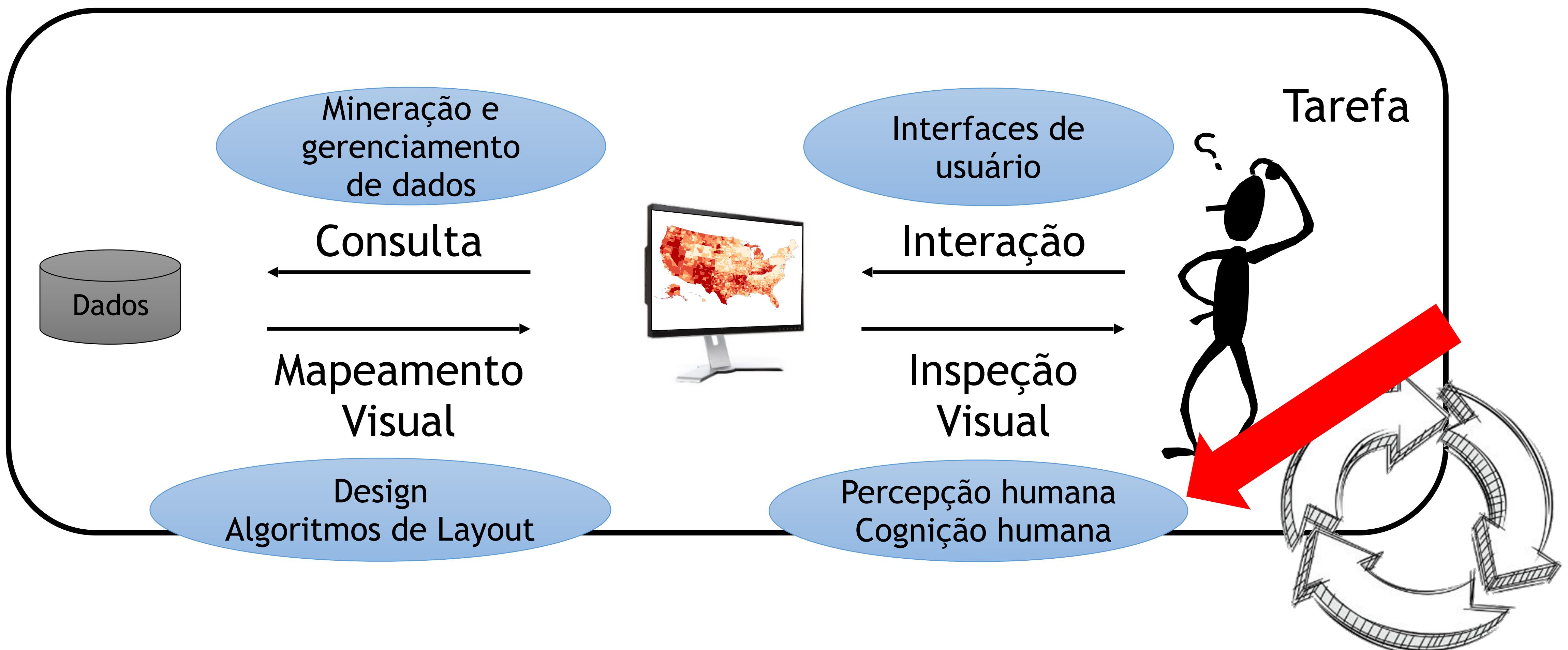
Como navegar esse grande espaço?

“Today's first task is not to **invent** wholly new [graphical] techniques, though these are needed. Rather we need most vitally to **recognize** and **reorganize** the **essentials** of old **techniques**, to make easy their **assembly in new ways**, and to **modify** their **external appearances** to fit the new opportunities.”

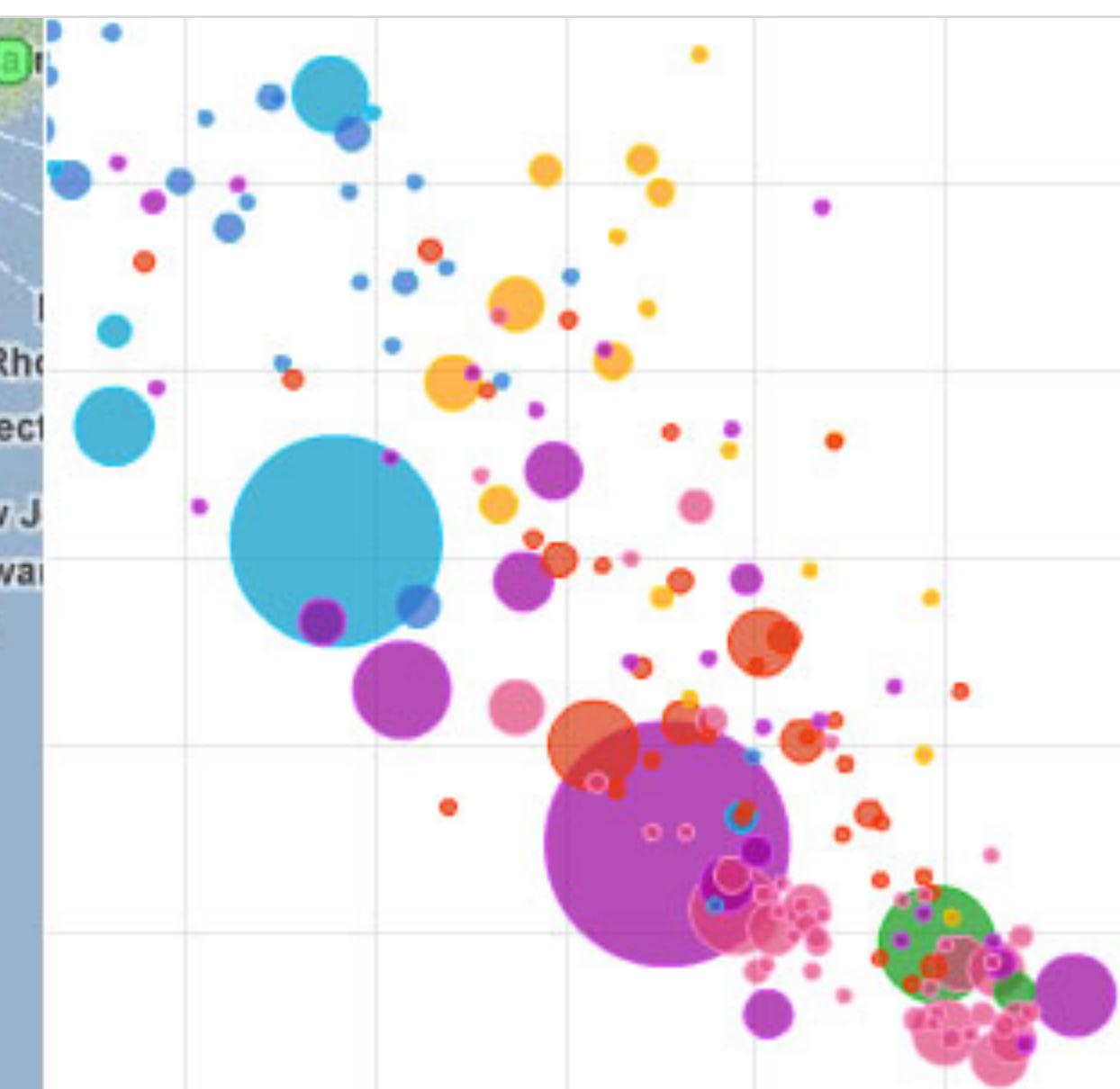
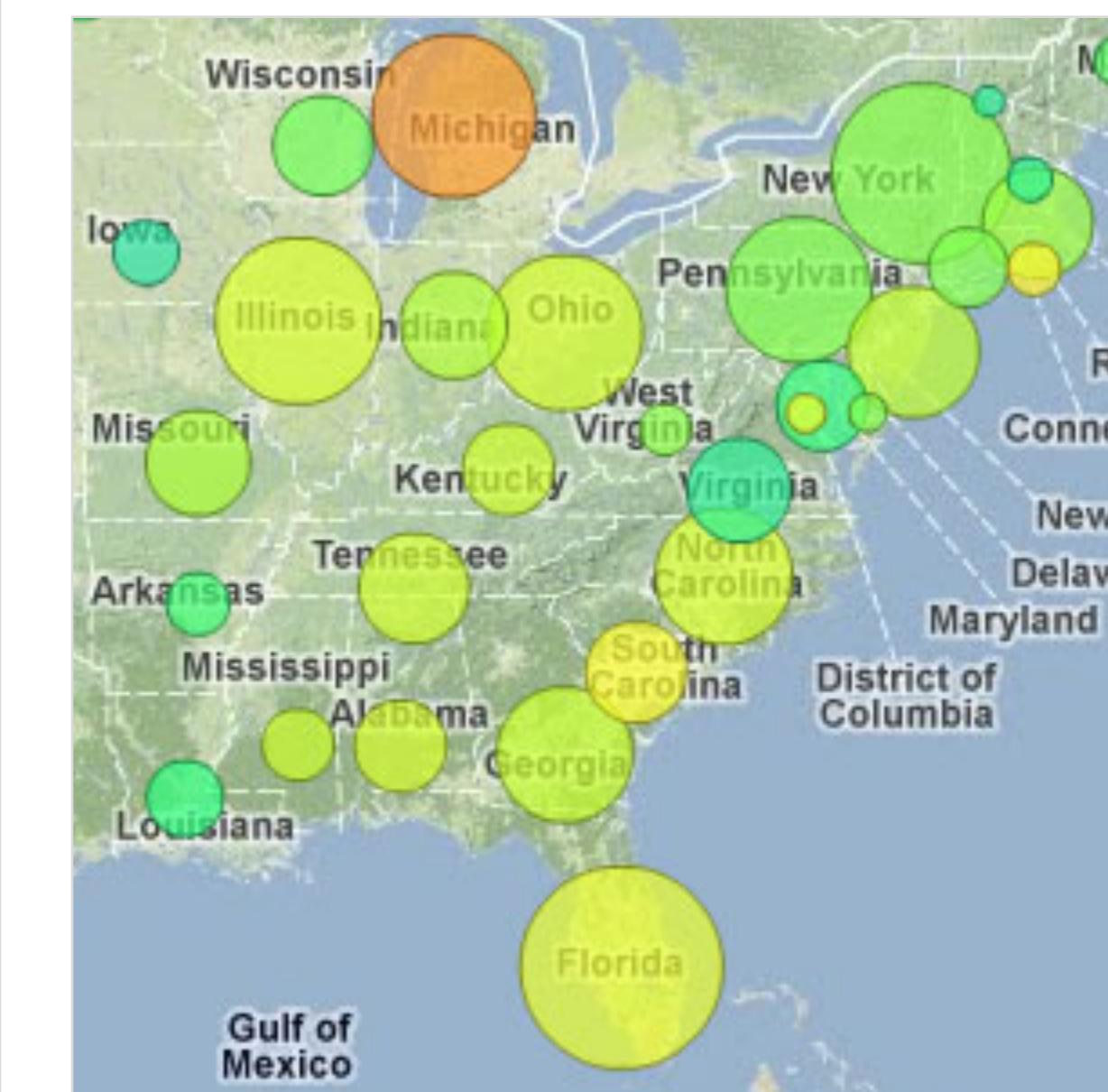
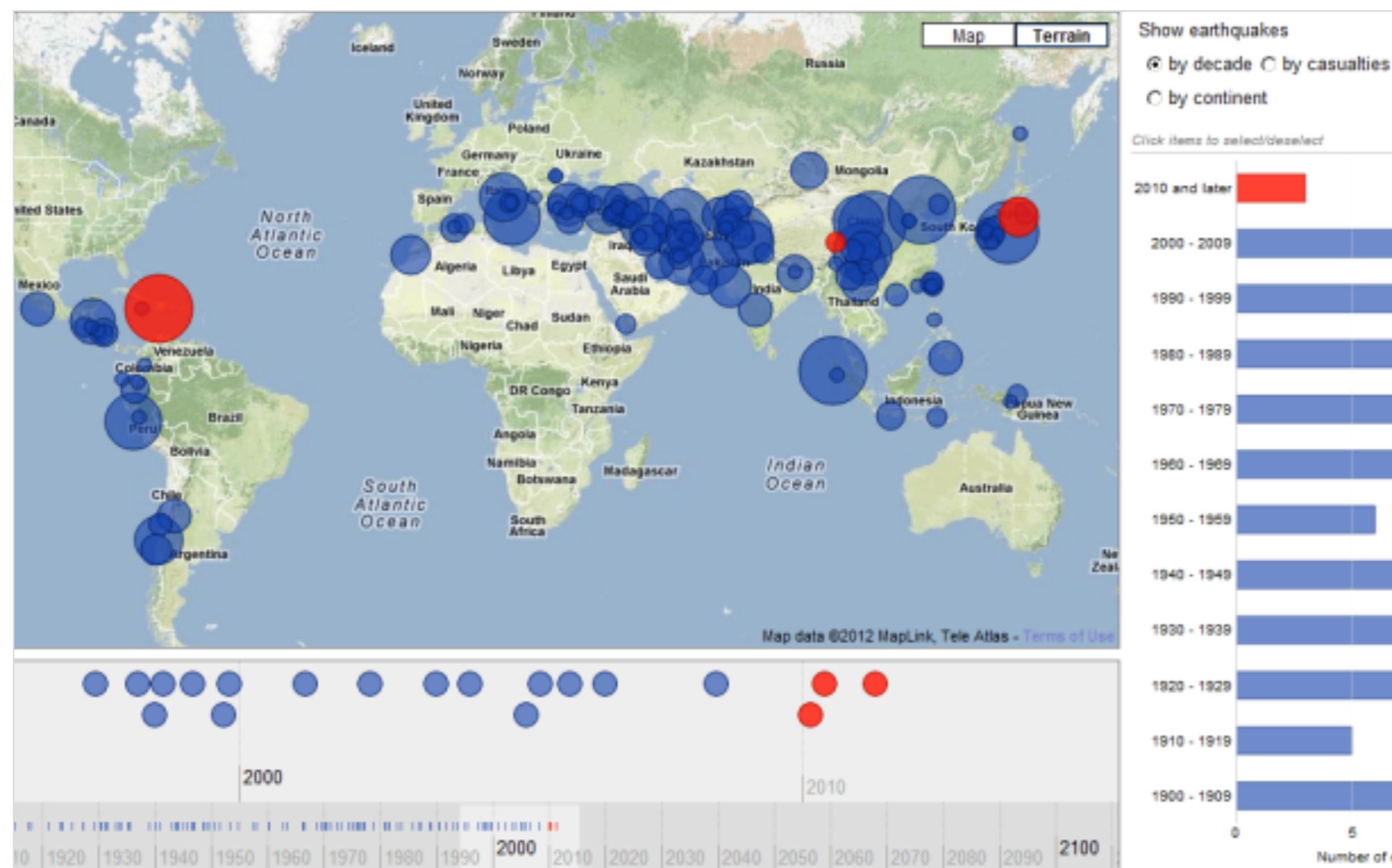
Jonh Tukey

The Future of Data Analysis, 1962

Processo de Visualização Interativa de Dados



Percepção e Cognição Humana



Psicofísica

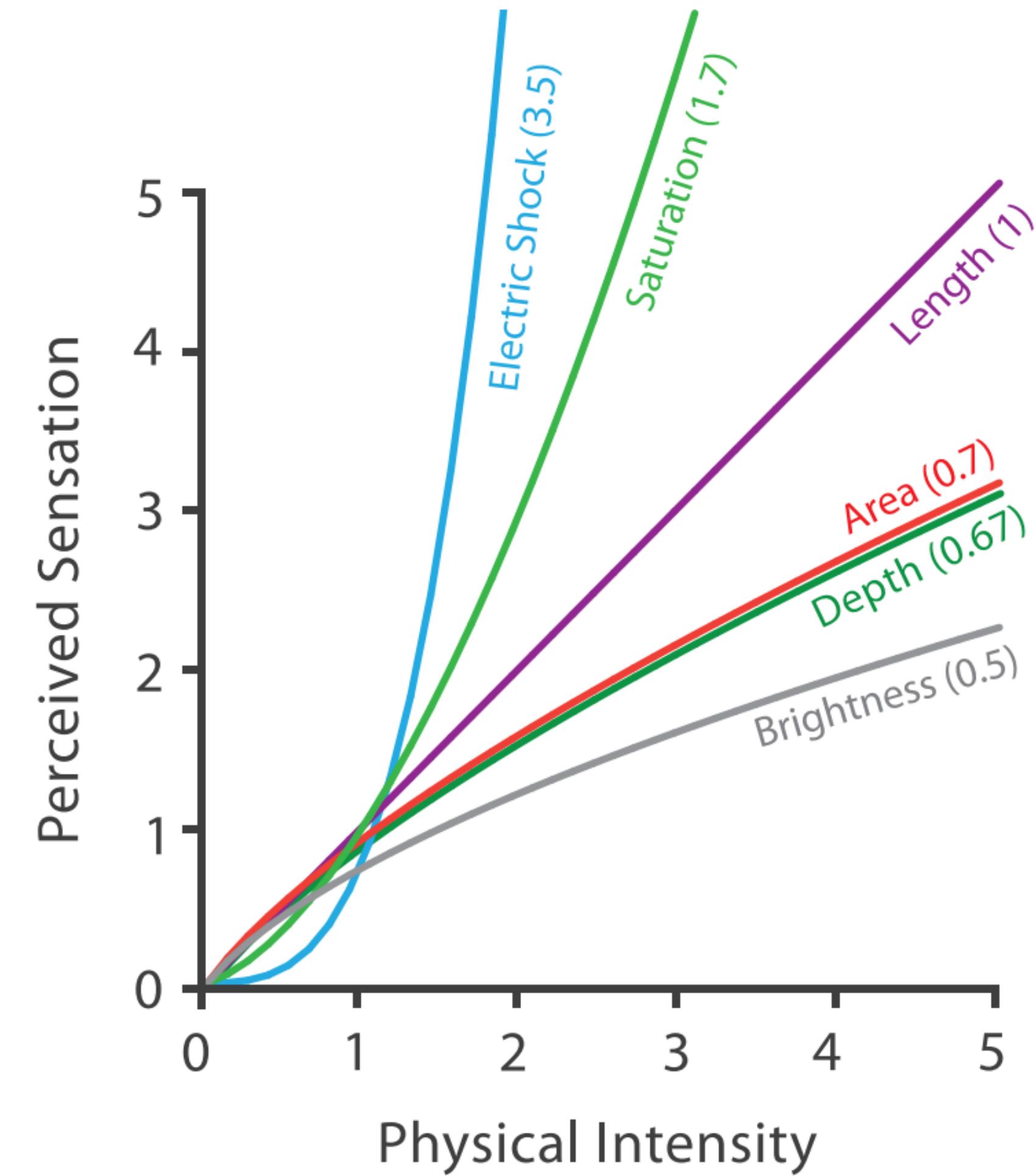


- Estudo científico da relação entre estímulo externos e a sensação (percepção)
- A lei de Weber: "quanto mais intenso um estímulo maior a diferença necessária para que seja percebida uma diferença entre eles"

$$dp = k \frac{dS}{S}$$

Psicofísica

Steven's Psychophysical Power Law: $S = I^n$



Aplicação em Percepção Visual

Graphical Perception: Theory, Experimentation, and Application to the Development of Graphical Methods

WILLIAM S. CLEVELAND and ROBERT McGILL*

The subject of graphical methods for data analysis and for data presentation needs a scientific foundation. In this article we take a few steps in the direction of establishing such a foundation. Our approach is based on *graphical perception*—the visual decoding of information encoded on graphs—and it includes both theory and experimentation to test the theory. The theory deals with a small but important piece of the whole process of graphical perception. The first part is an identification of a set of

largely unscientific. This is why Cox (1978) argued, “There is a major need for a theory of graphical methods” (p. 5), and why Kruskal (1975) stated “in choosing, constructing, and comparing graphical methods we have little to go on but intuition, rule of thumb, and a kind of master-to-apprentice passing along of information. . . . there is neither theory nor systematic body of experiment as a guide” (p. 28–29).

There is, of course, much good common sense about

Aplicação em Percepção Visual

④ Position

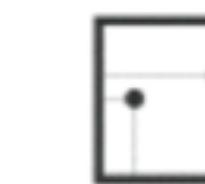
→ Horizontal



→ Vertical



→ Both



④ Color



④ Shape



④ Tilt

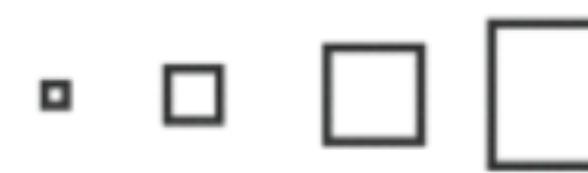


④ Size

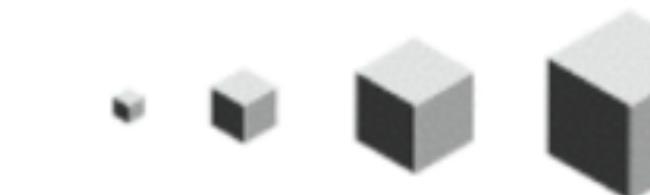
→ Length



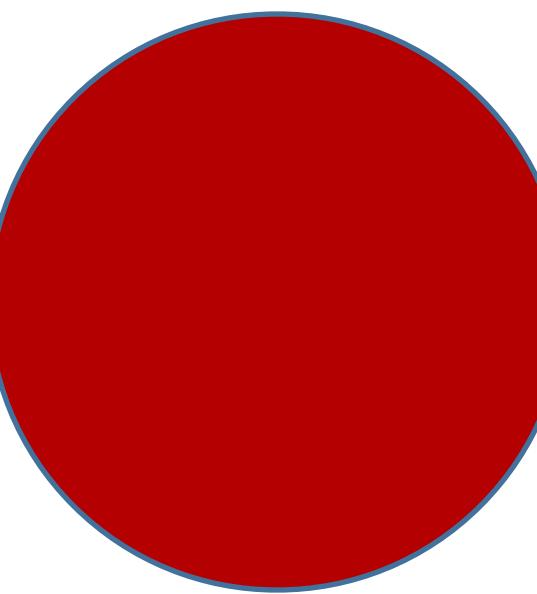
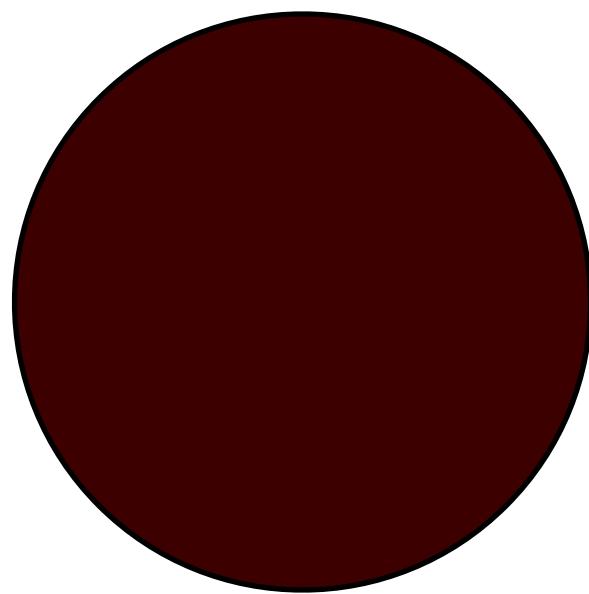
→ Area



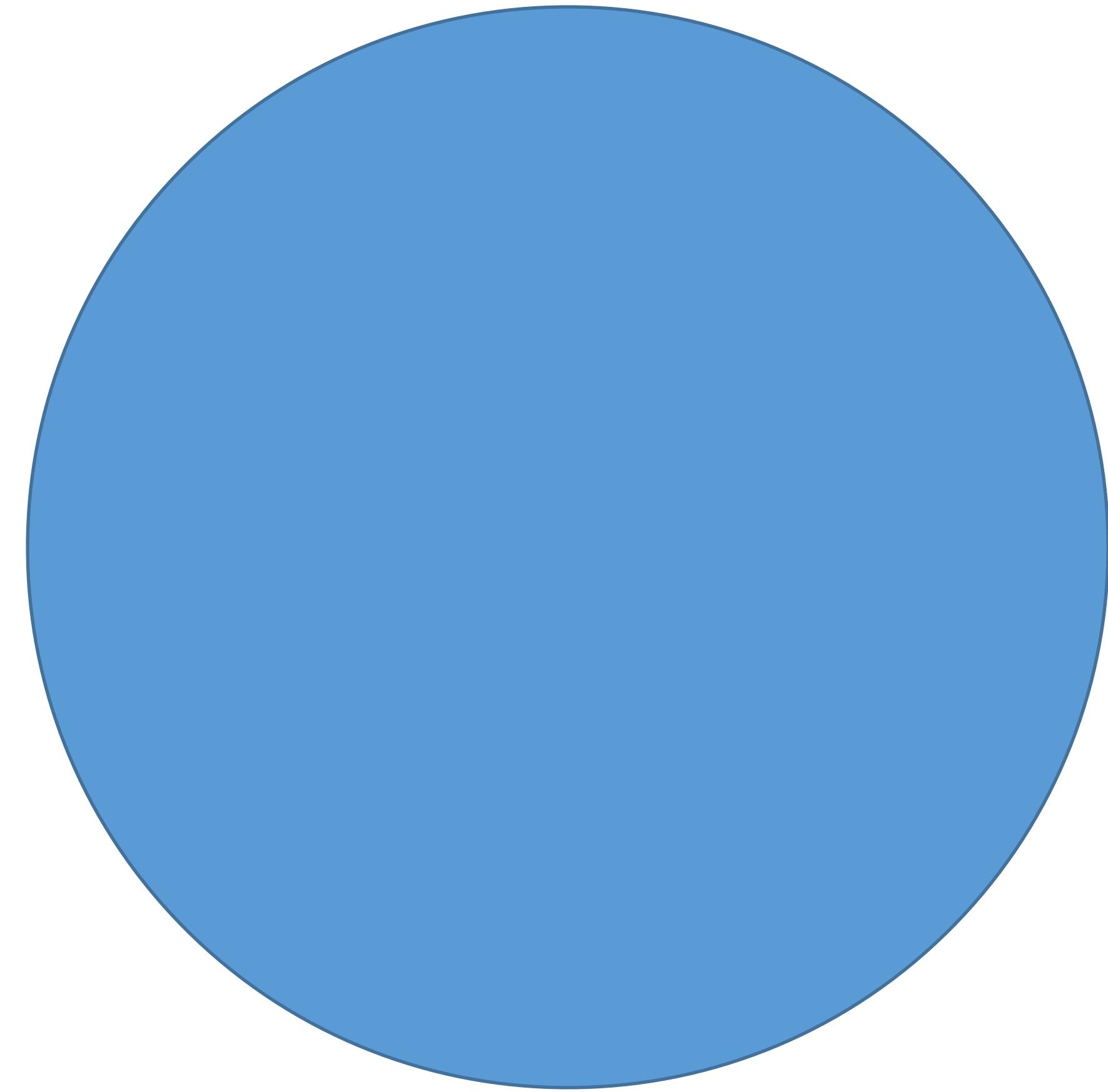
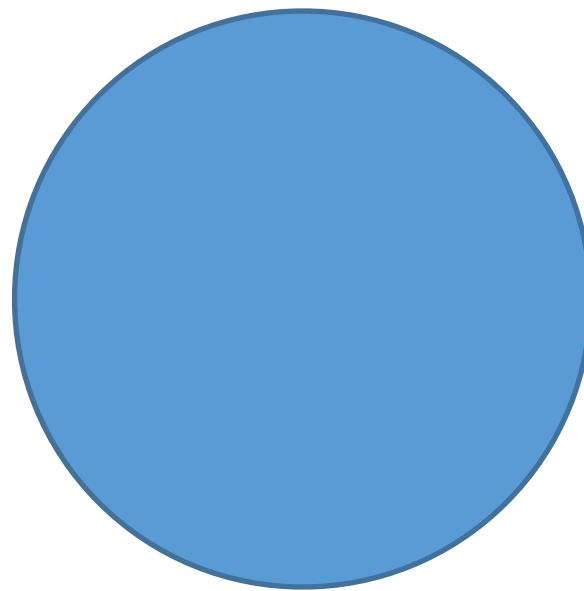
→ Volume



Percepção e Cognição Humana



Percepção e Cognição Humana



Percepção e Cognição Humana



Percepção e Cognição Humana

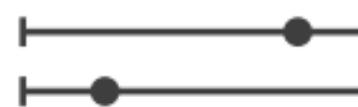


Percepção e Cognição Humana

Channels: Expressiveness Types and Effectiveness Ranks

④ Magnitude Channels: Ordered Attributes

Position on common scale



Position on unaligned scale



Length (1D size)



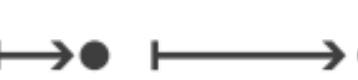
Tilt/angle



Area (2D size)



Depth (3D position)



Color luminance



Color saturation



Curvature



Volume (3D size)



④ Identity Channels: Categorical Attributes

Spatial region



Color hue



Motion



Shape



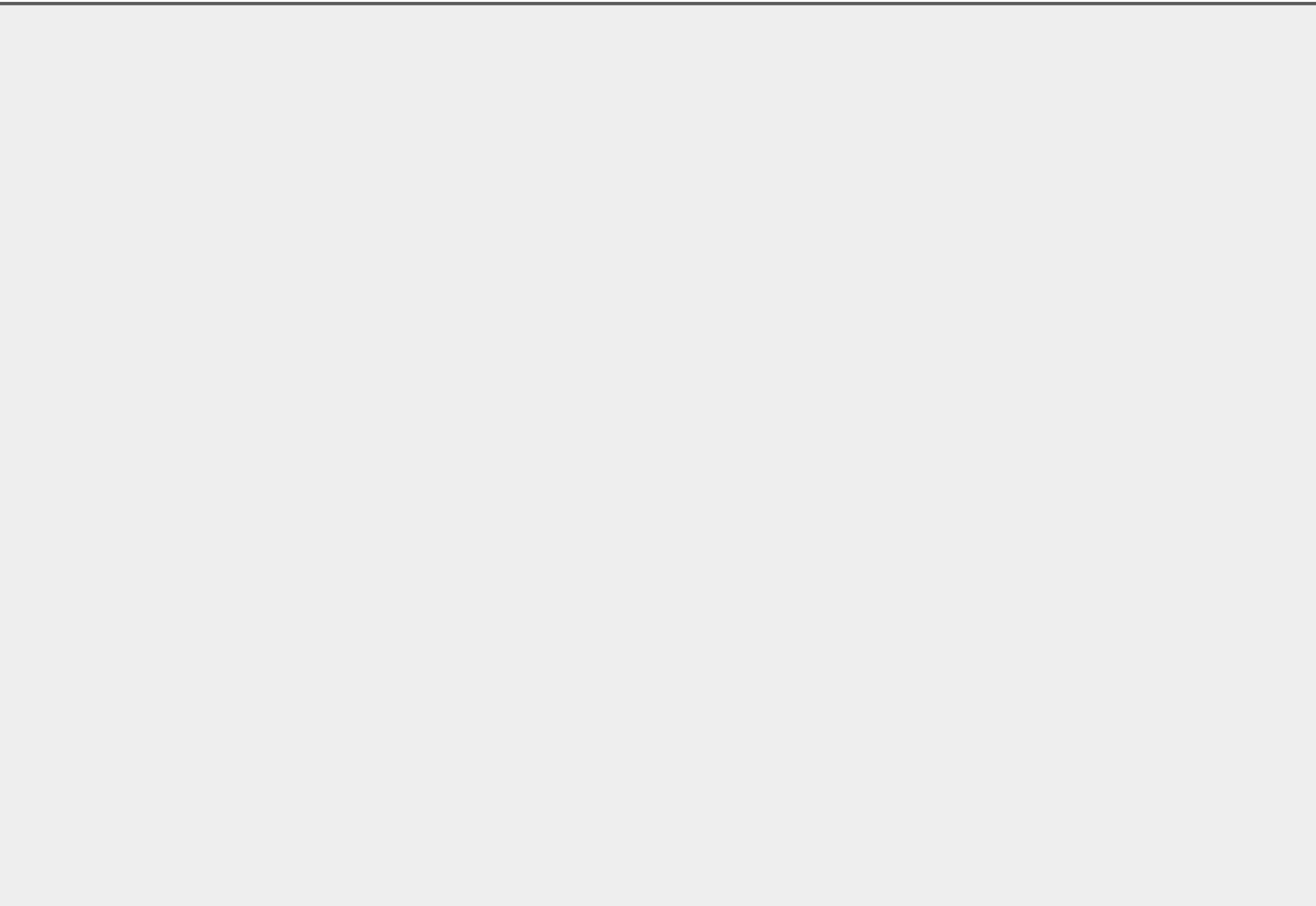
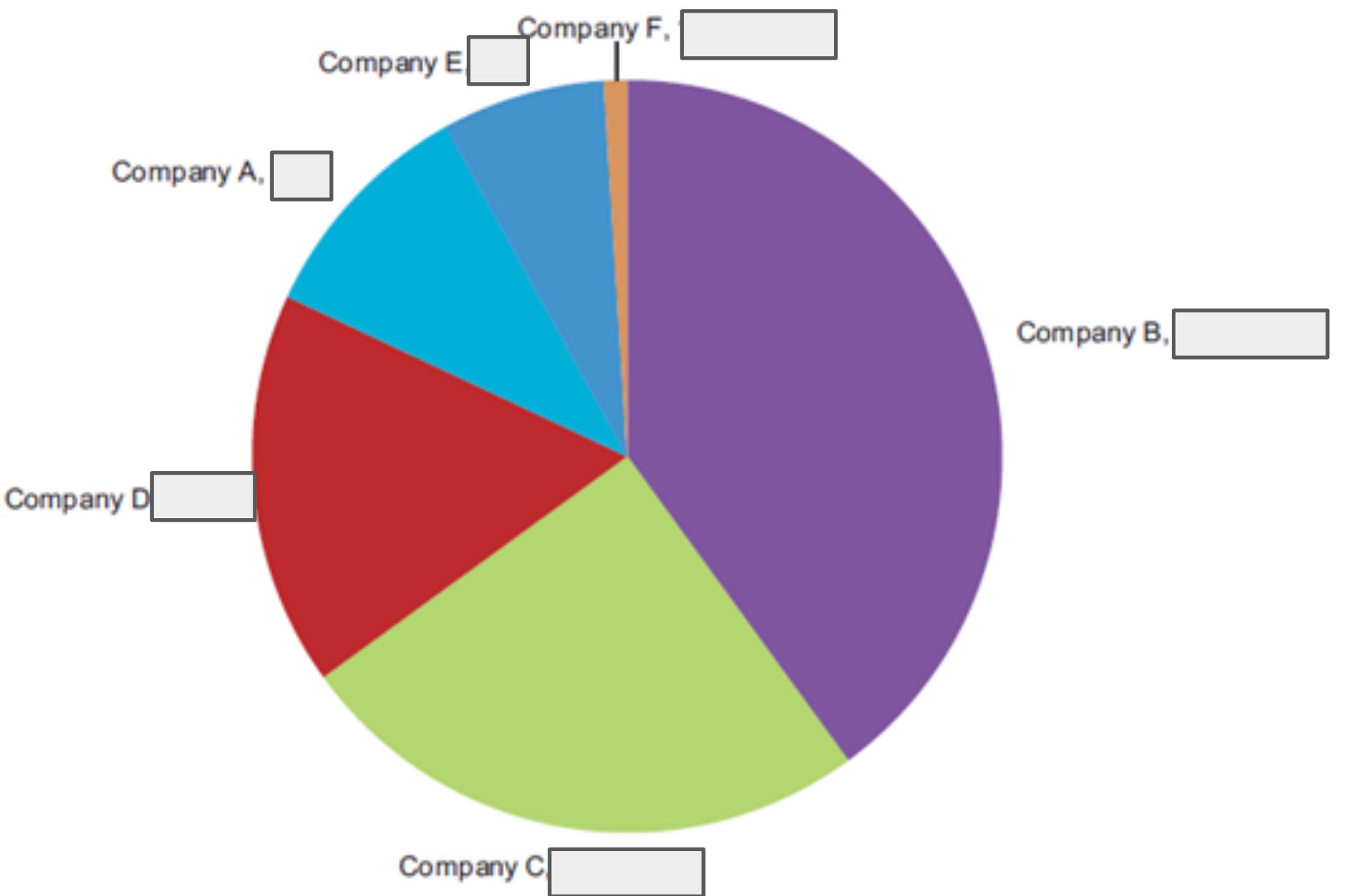
▲ Most

Effectiveness

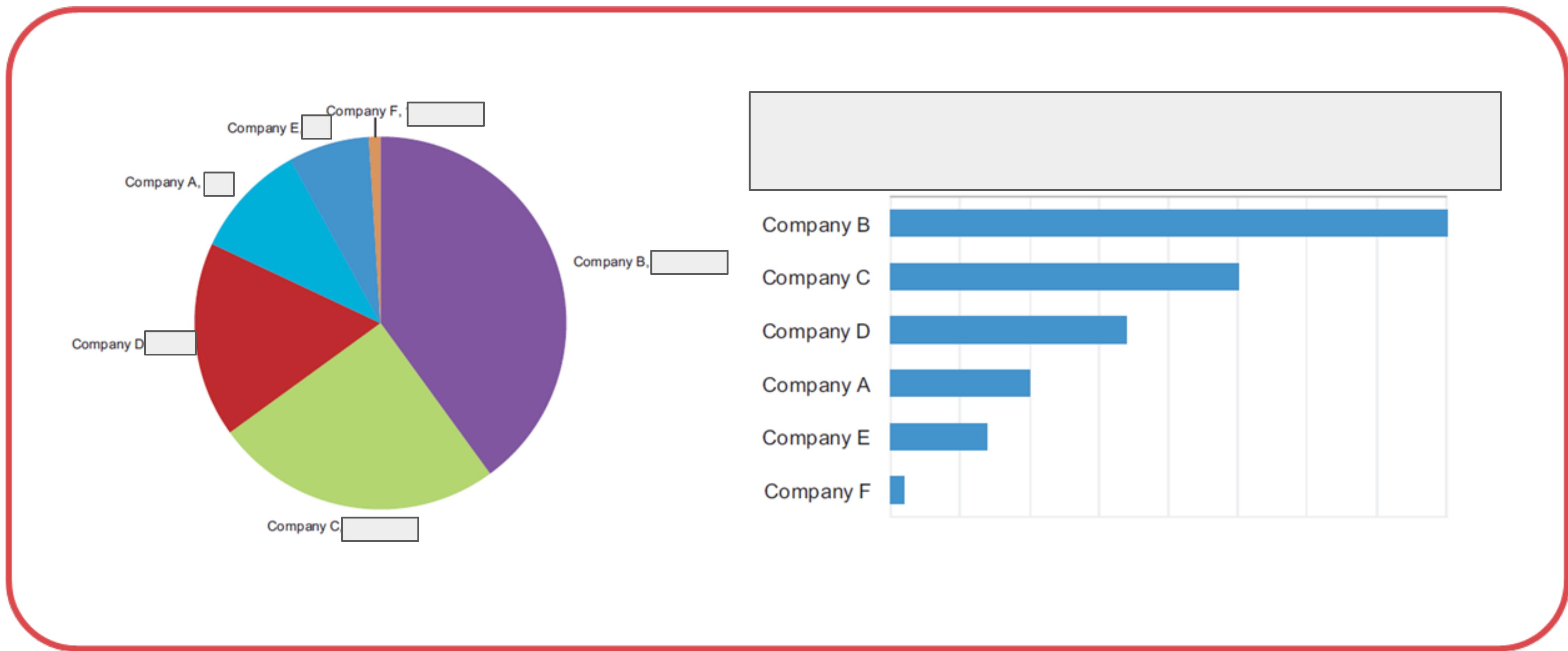
Same

Least

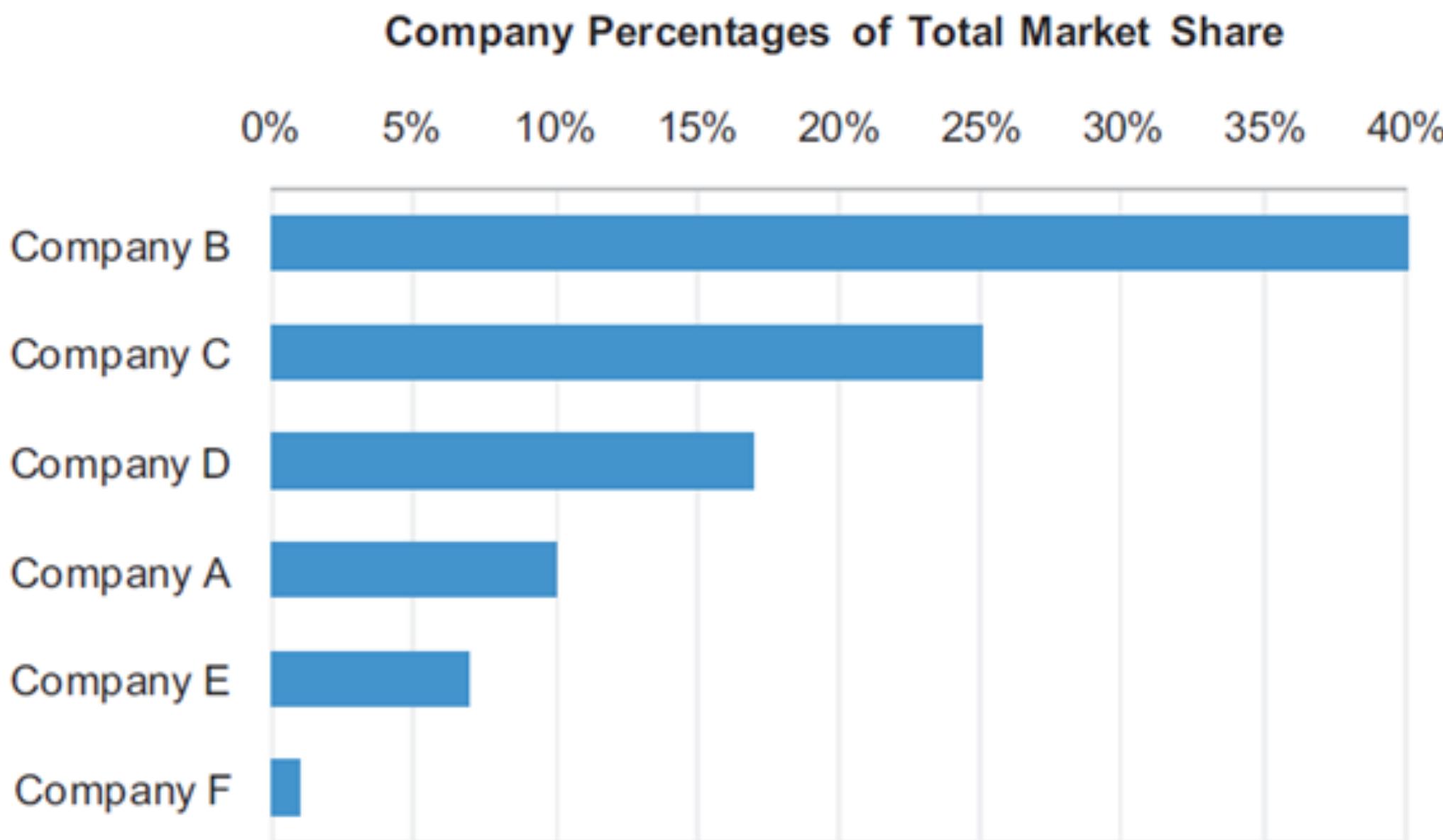
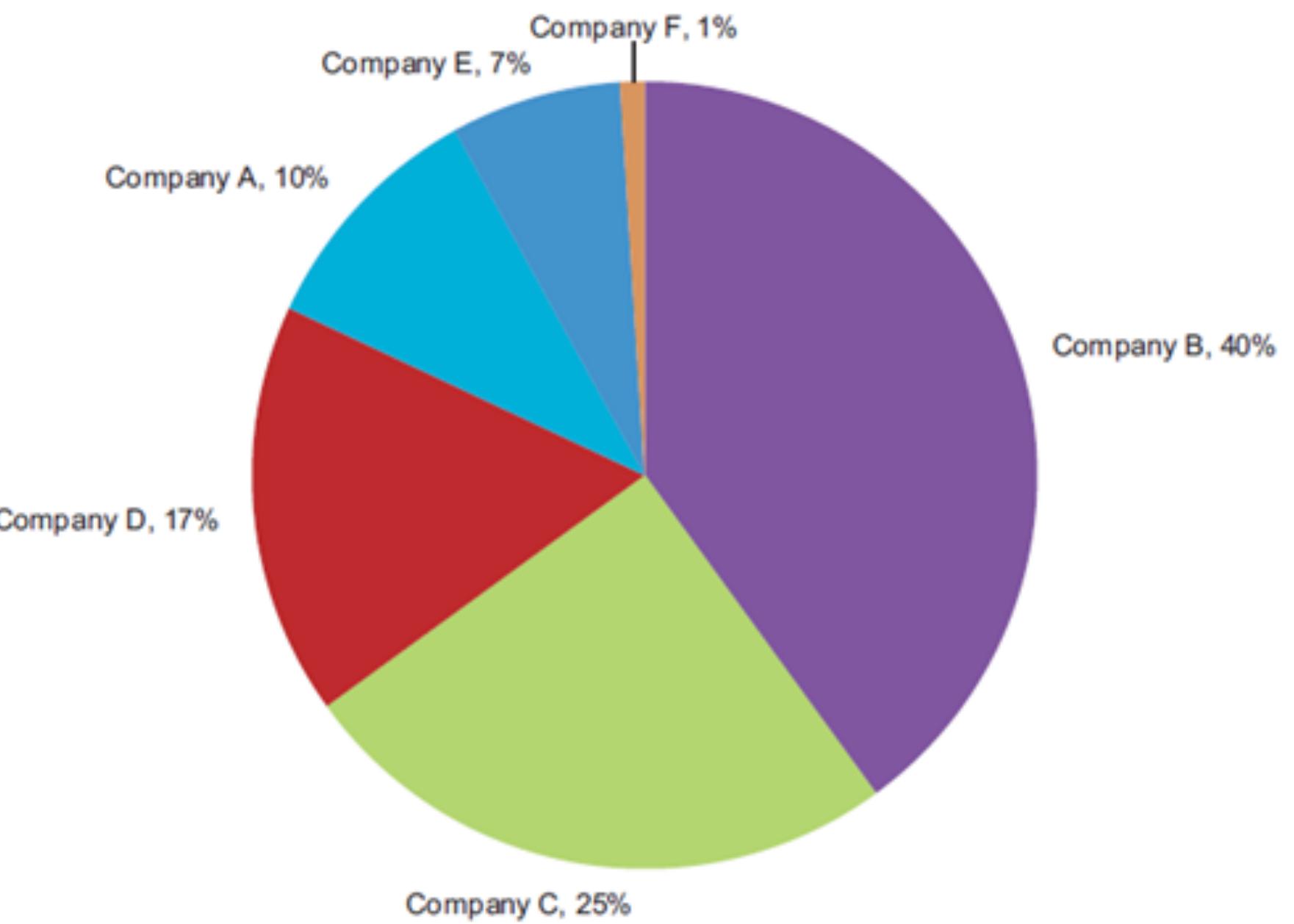
Percepção e Cognição Humana



Percepção e Cognição Humana



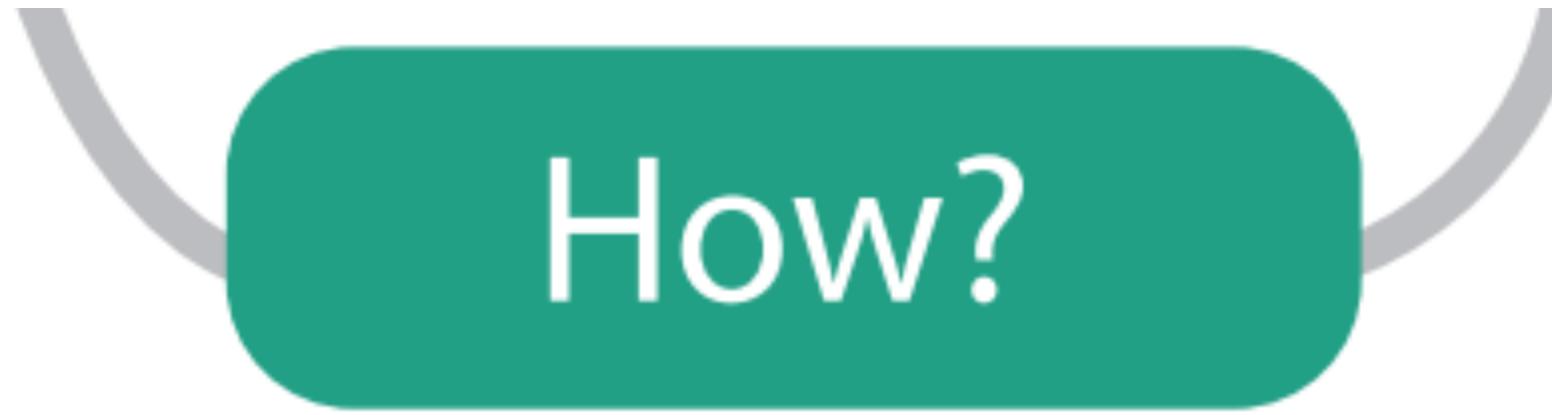
Percepção e Cognição Humana



Cenário Recorrente

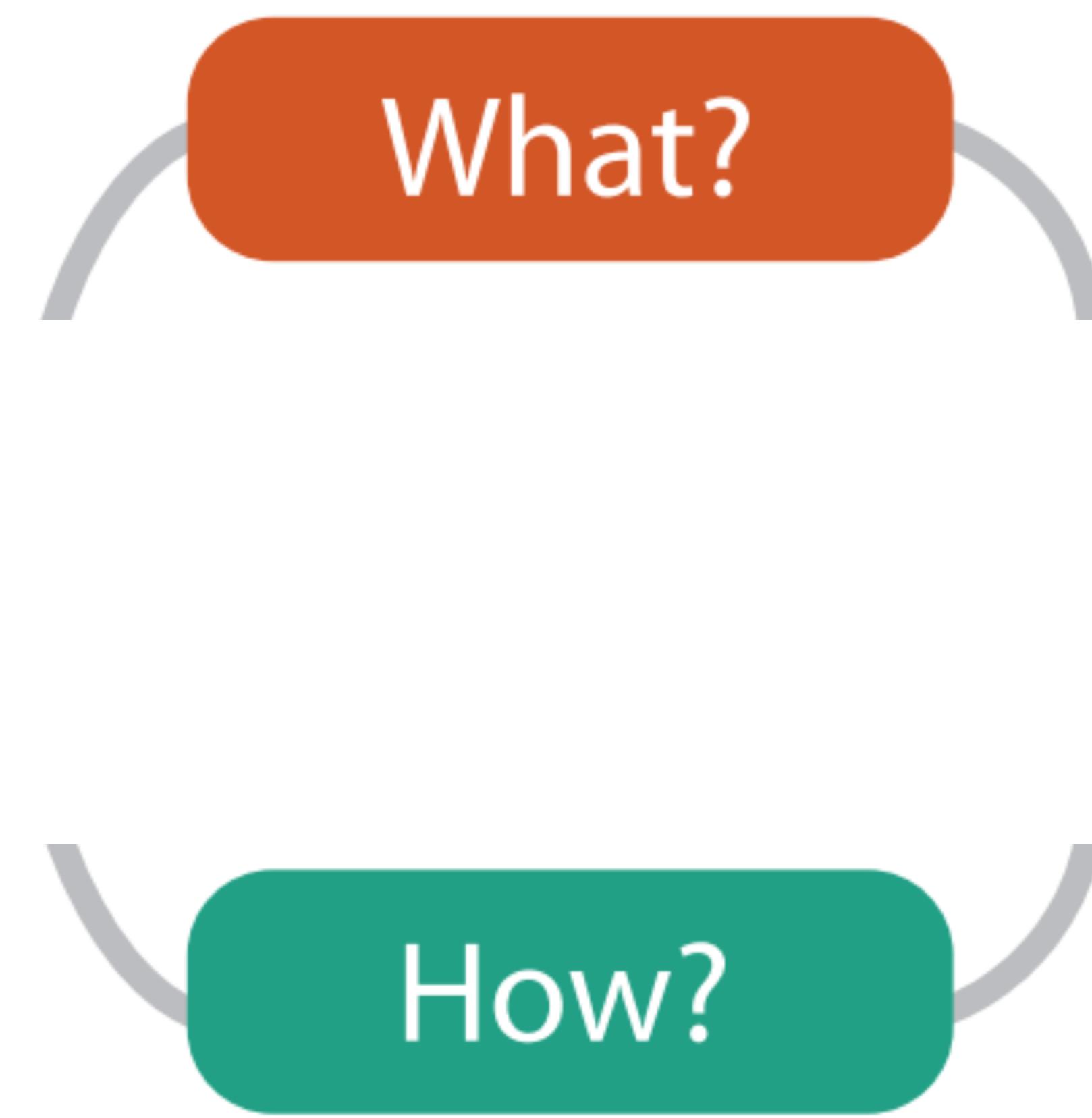


Como Atacar esse Problema?

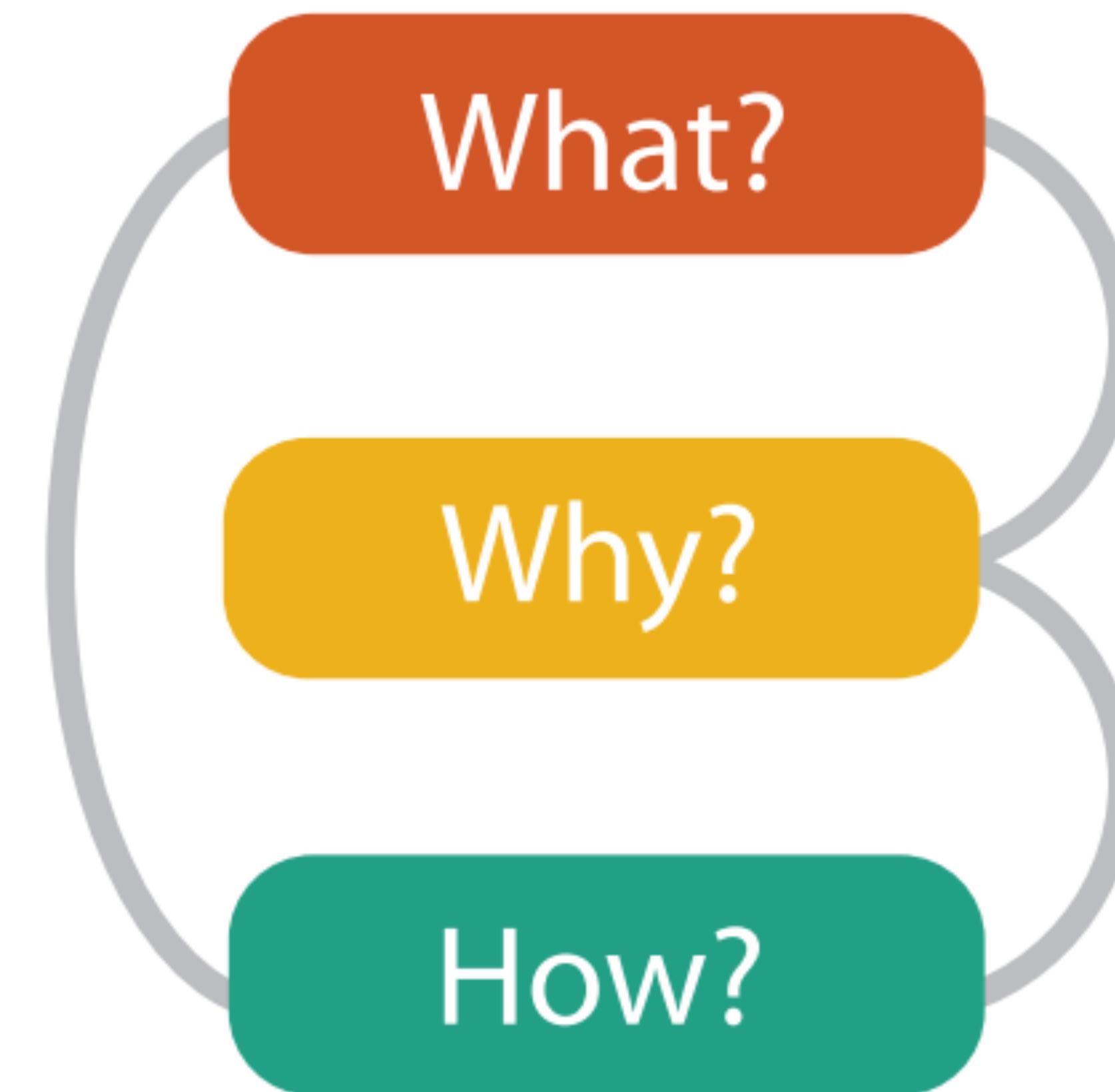


How?

Como Atacar esse Problema?



Como Atacar esse Problema?

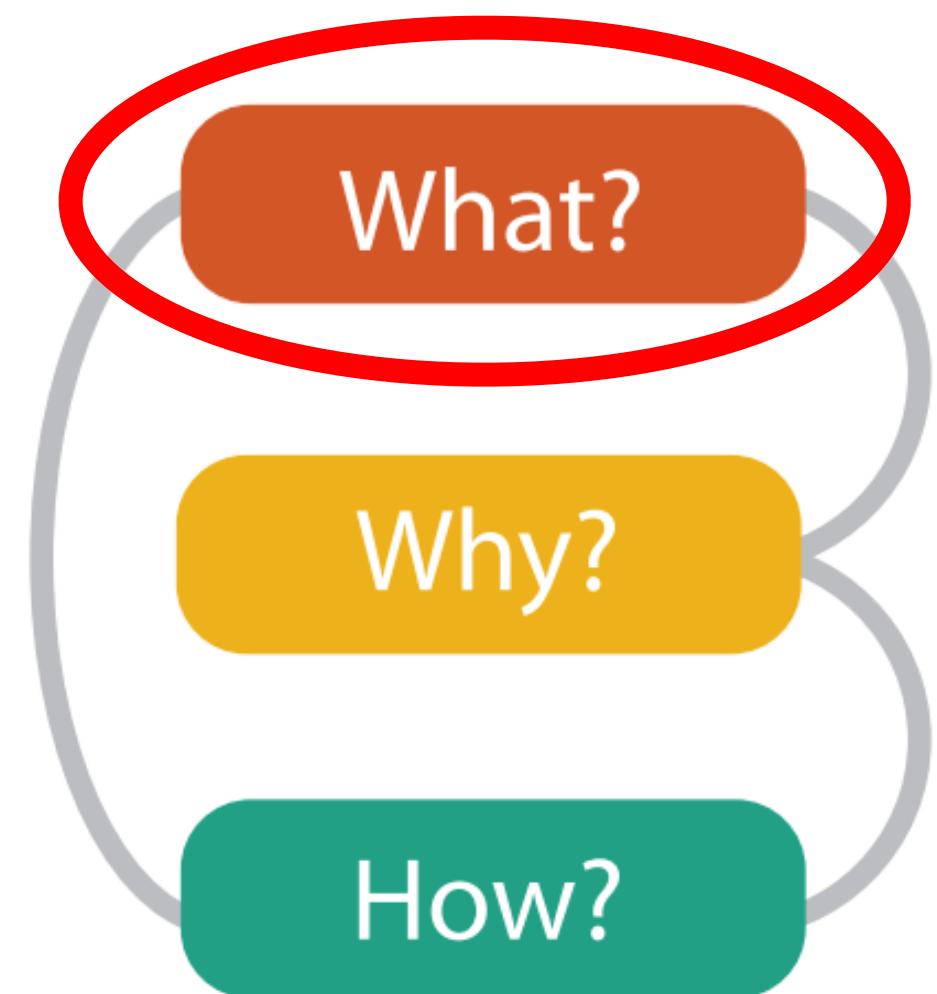


Exemplo 1

Cenário Hipotético

- Imagine que uma determinada empresa gostaria de entender como seus funcionários estão usando a infra-estrutura da empresa
- Para tanto dados sobre o número de empregados dentro do prédio é monitorado hora a hora por um ano inteiro
- O que você poderia fazer com esses dados? Como você faria uma análise? Como você usaria visualização?

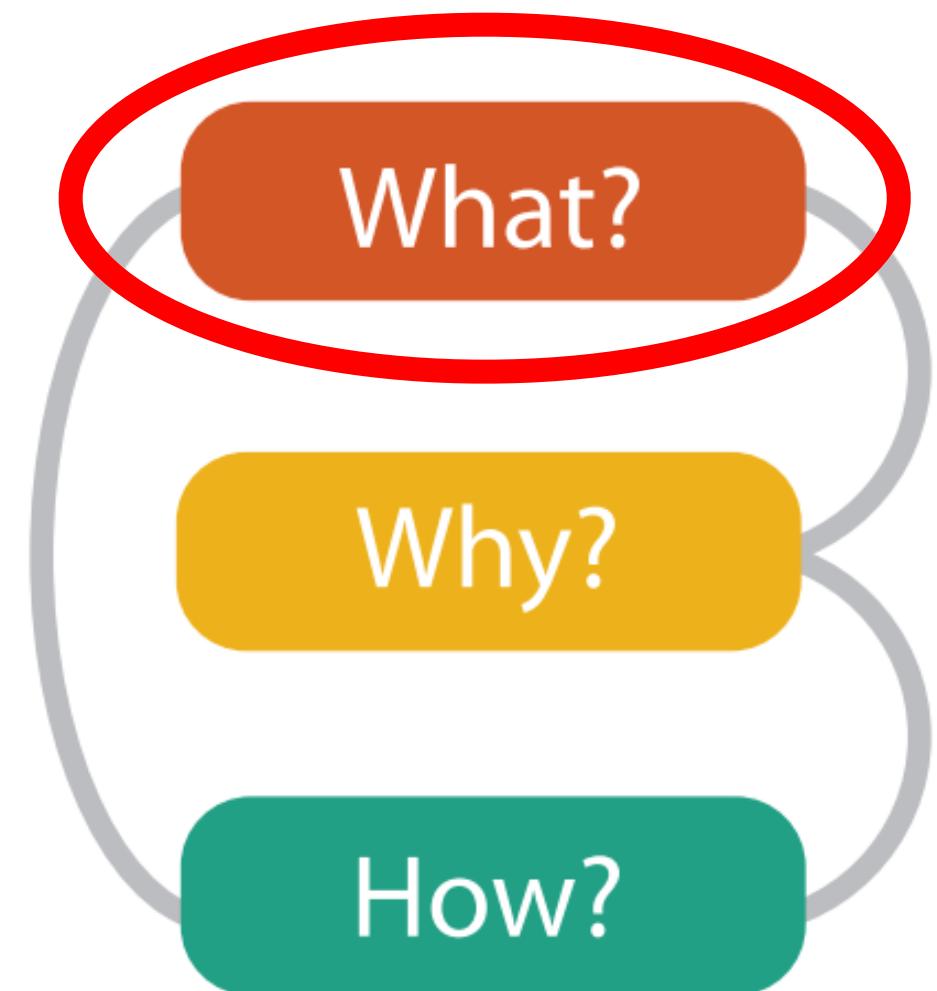
Primeiro Passo: Entender os Dados (O quê)



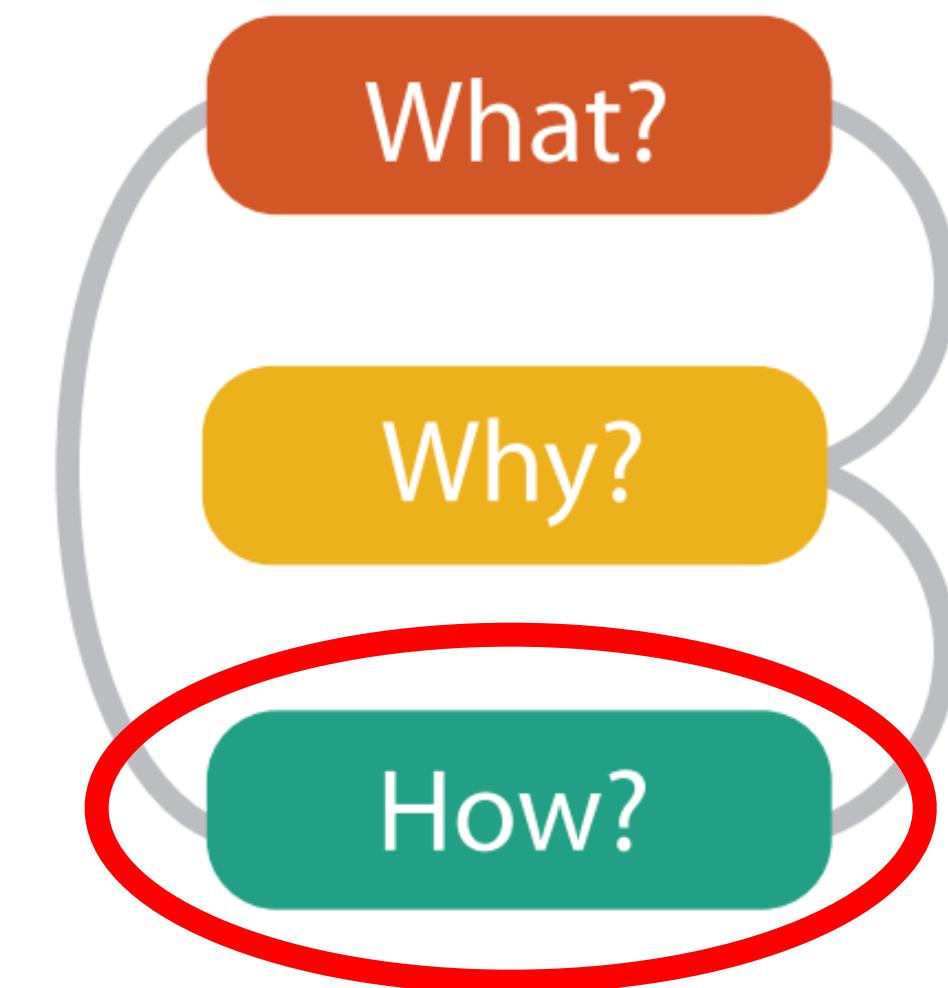
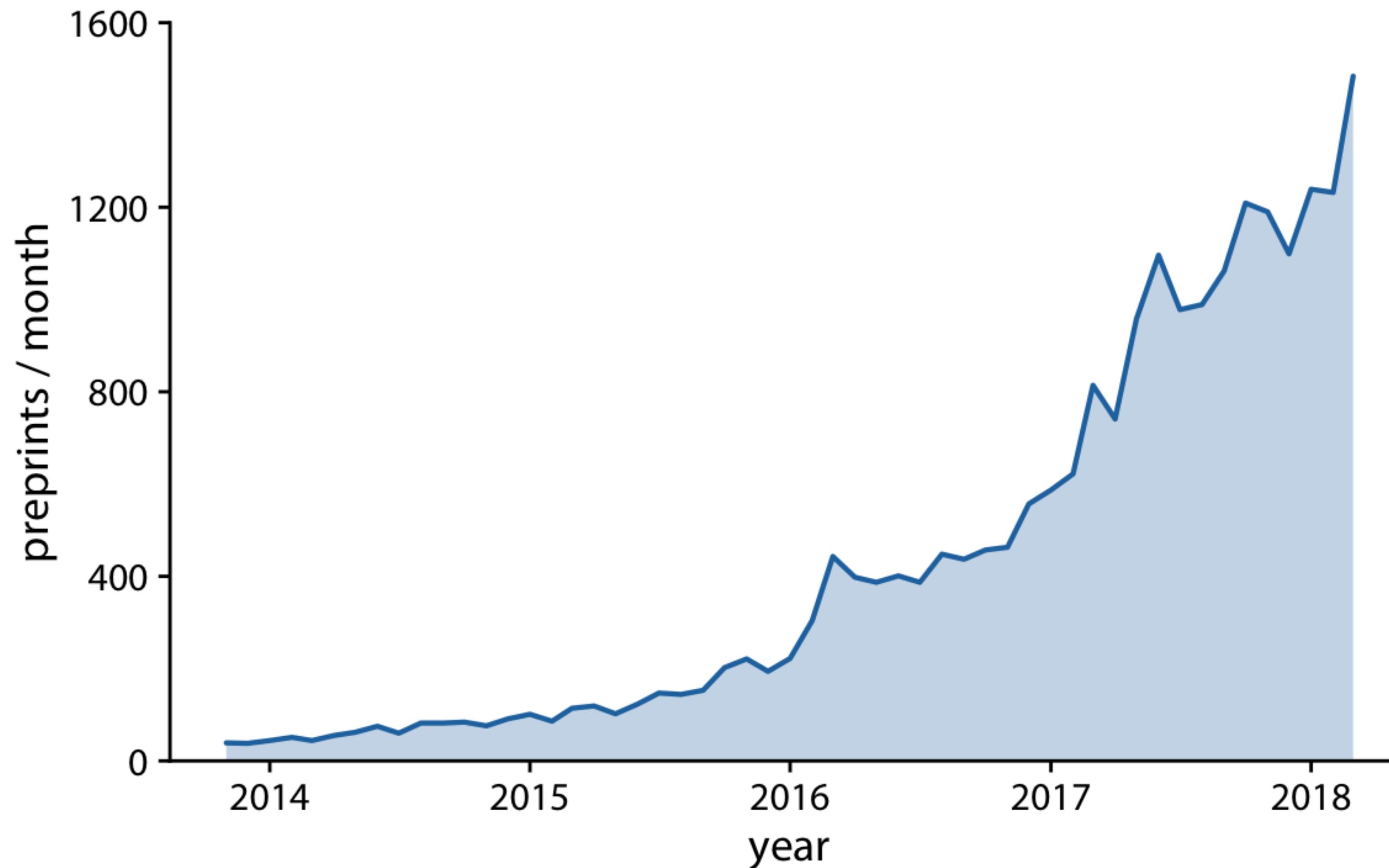
Primeiro Passo: Entender os Dados (O quê)

- Dados são simples

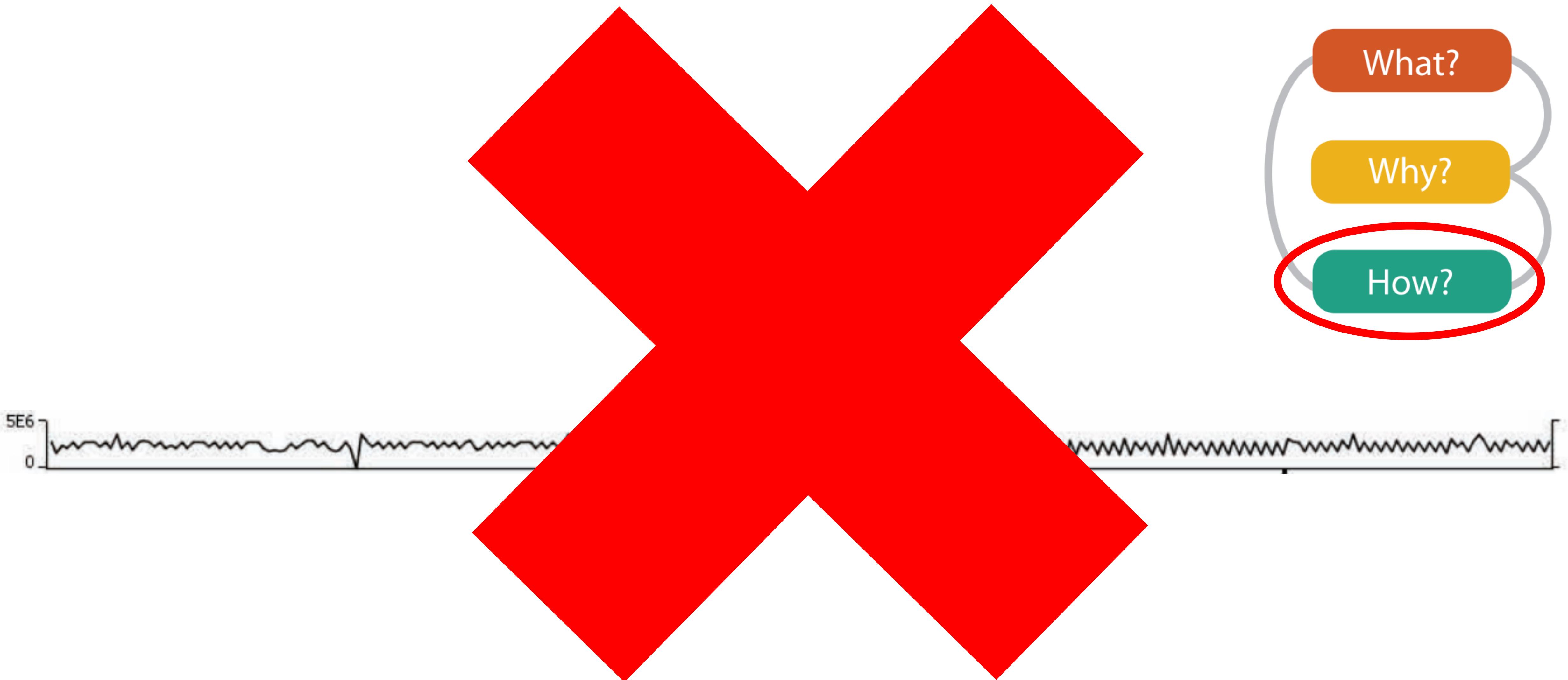
$$T = \{(t_1, v_1), (t_2, v_2), \dots, (t_N, v_N)\}$$



Já sei! Como Visualizar!



Já sei! Como Visualizar!



Segundo Passo: Entender as Tarefas Analíticas

- Que tipo de perguntas podemos fazer sobre esses dados
- Como elas influenciariam a sua visualização de dados?



Alto Nível: Como se dá o uso do prédio pelos funcionários?

Nível Médio: Jornadas de trabalho são de horas como ciclos em dias, então temos que visualizar os dados por dia e tentar investigar o quanto uniforme as jornadas de trabalho são durante o ano

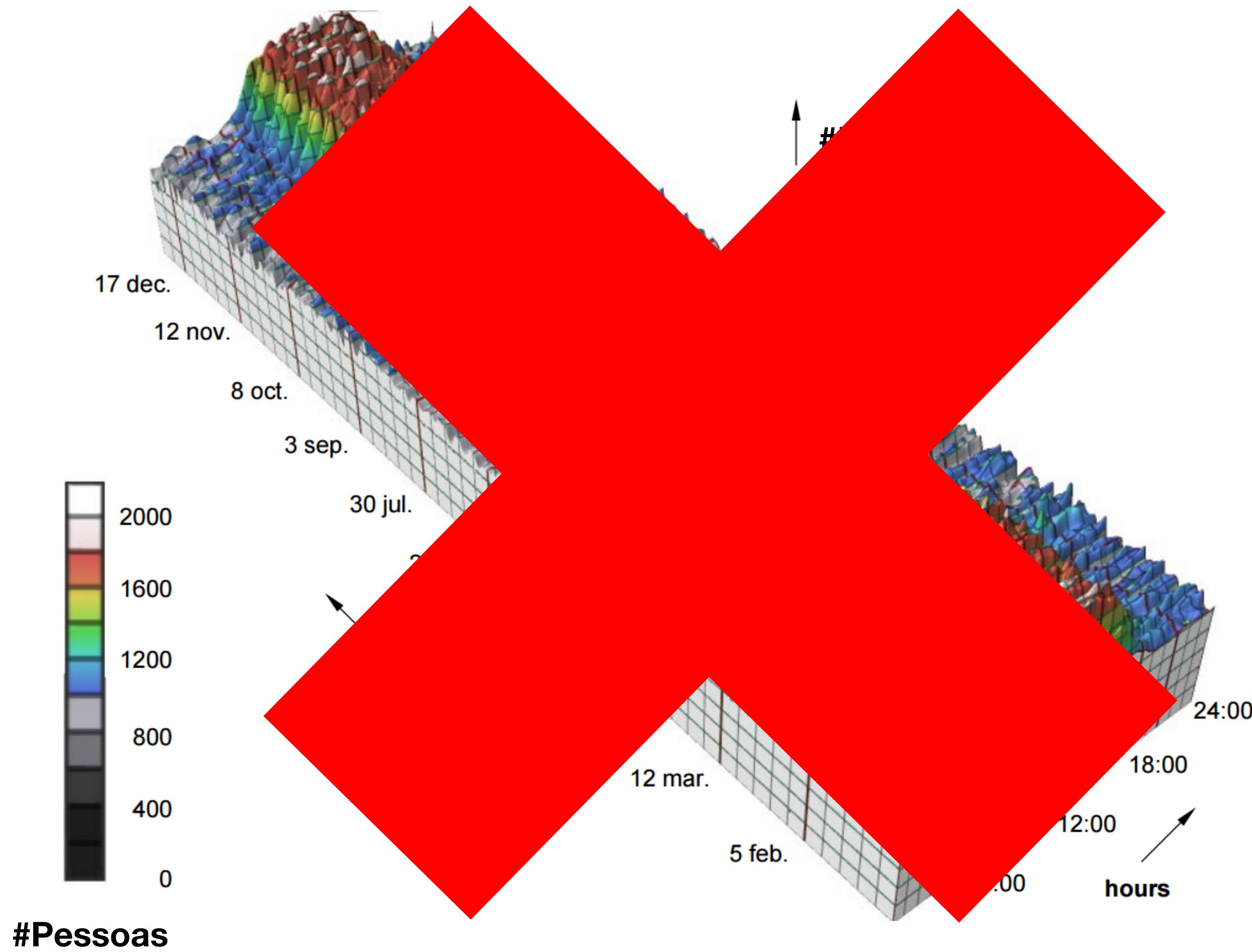
Baixo Nível: Um sistema de visualização para comparar um número considerável (365) séries temporais pequenas (24 pontos)

Rule of Thumb #0: InfoVis Mantra

“Overview first,
zoom and filter,
then details on demand.”

Ben Schneiderman, “The eyes have it: A task by data type taxonomy for information visualization” Visual Languages, 1996

Uma Solução



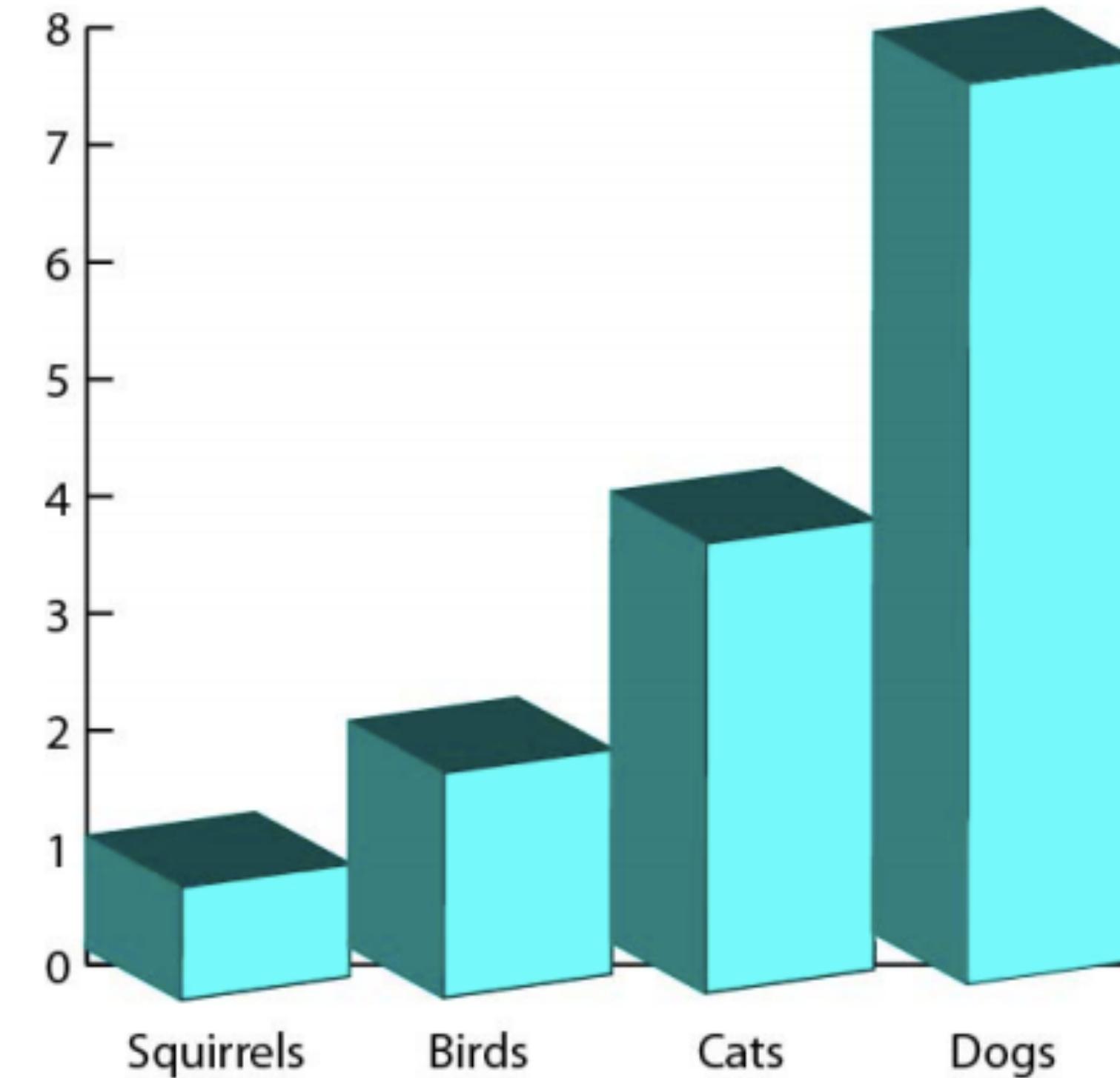
What?

Why?

How?

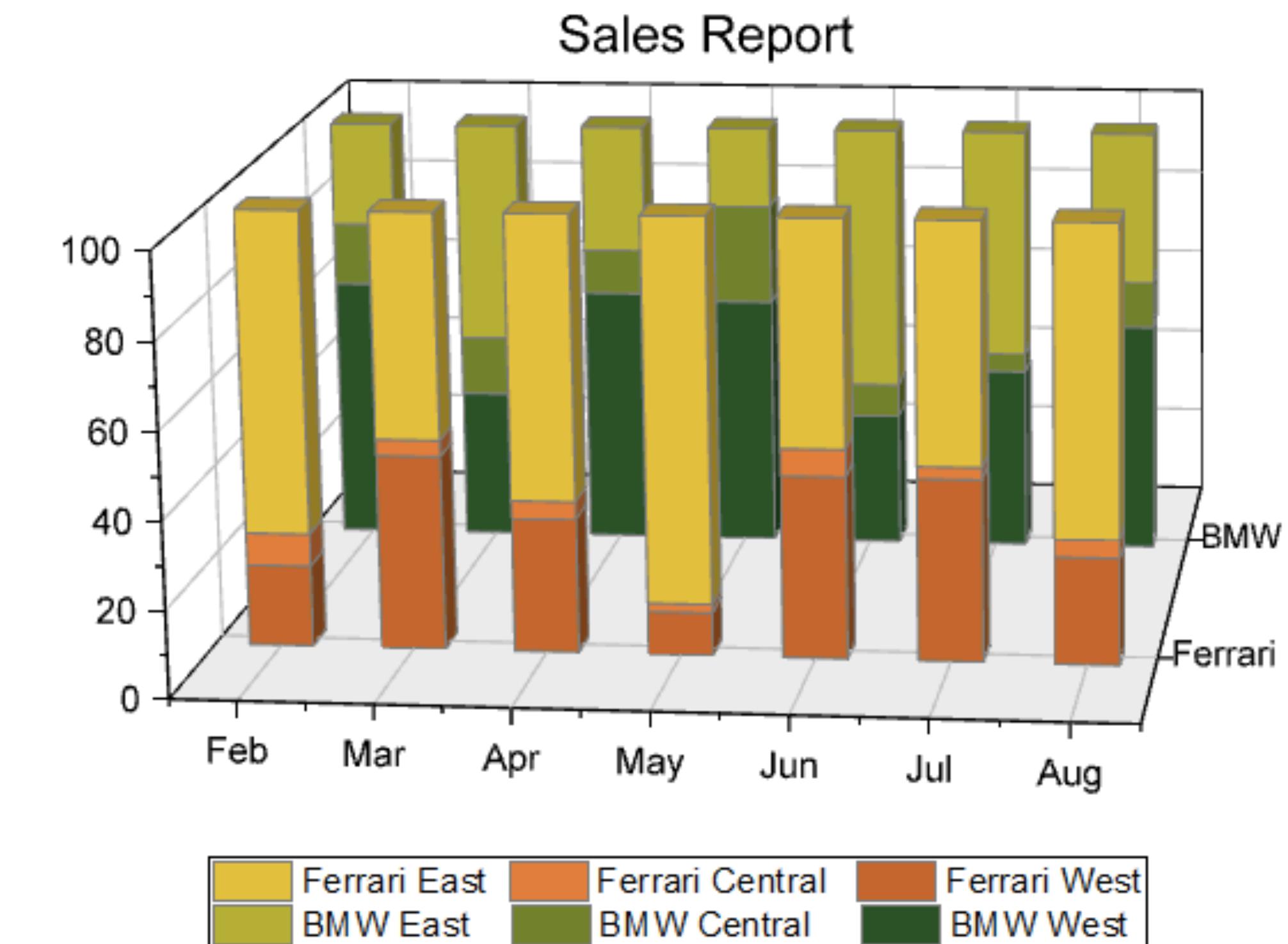
Rule of Thumb #1

- Cautela no Uso do 3D
 - Disparidade da “Profundidade”
 - Navegação e Oclusão
 - Distorção Perspectiva



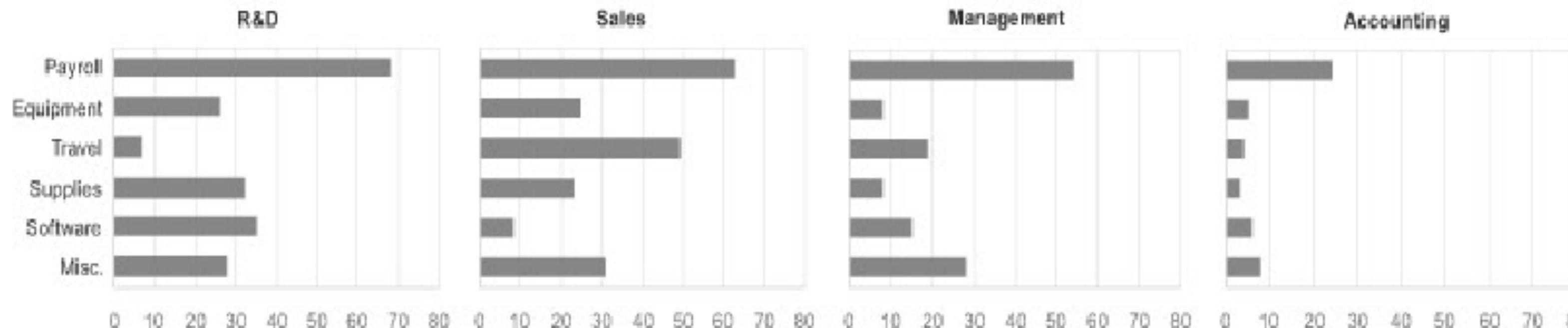
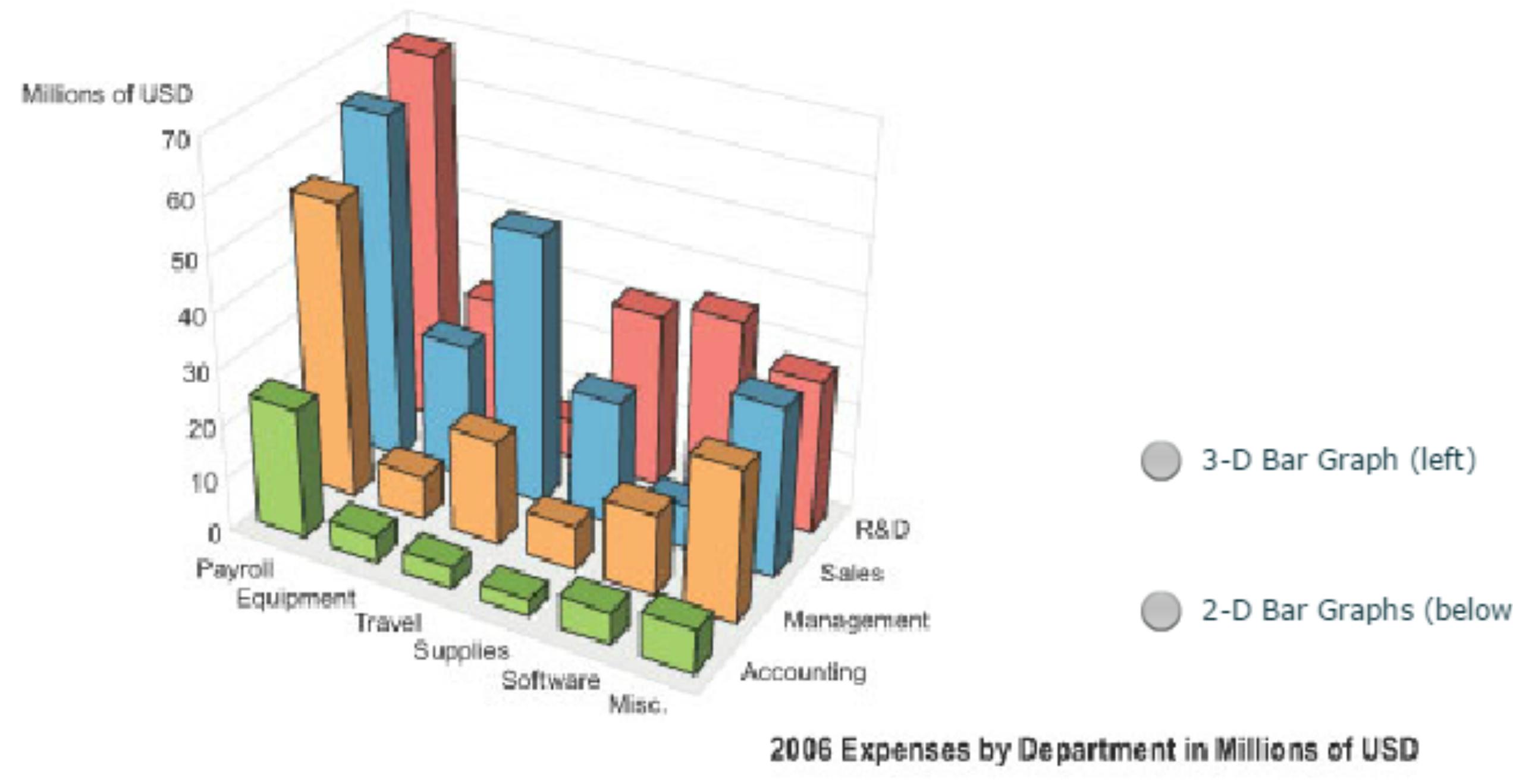
Rule of Thumb #1

- Cautela no Uso do 3D
 - Disparidade da “Profundidade”
 - Navegação e Oclusão
 - Distorção Perspectiva



Rule of Thumb #1

2006 Expenses by Department



Rule of Thumb #1

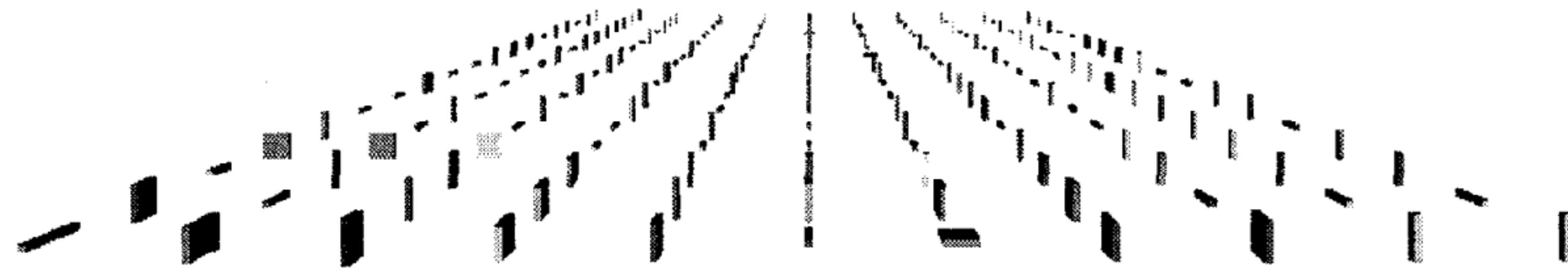
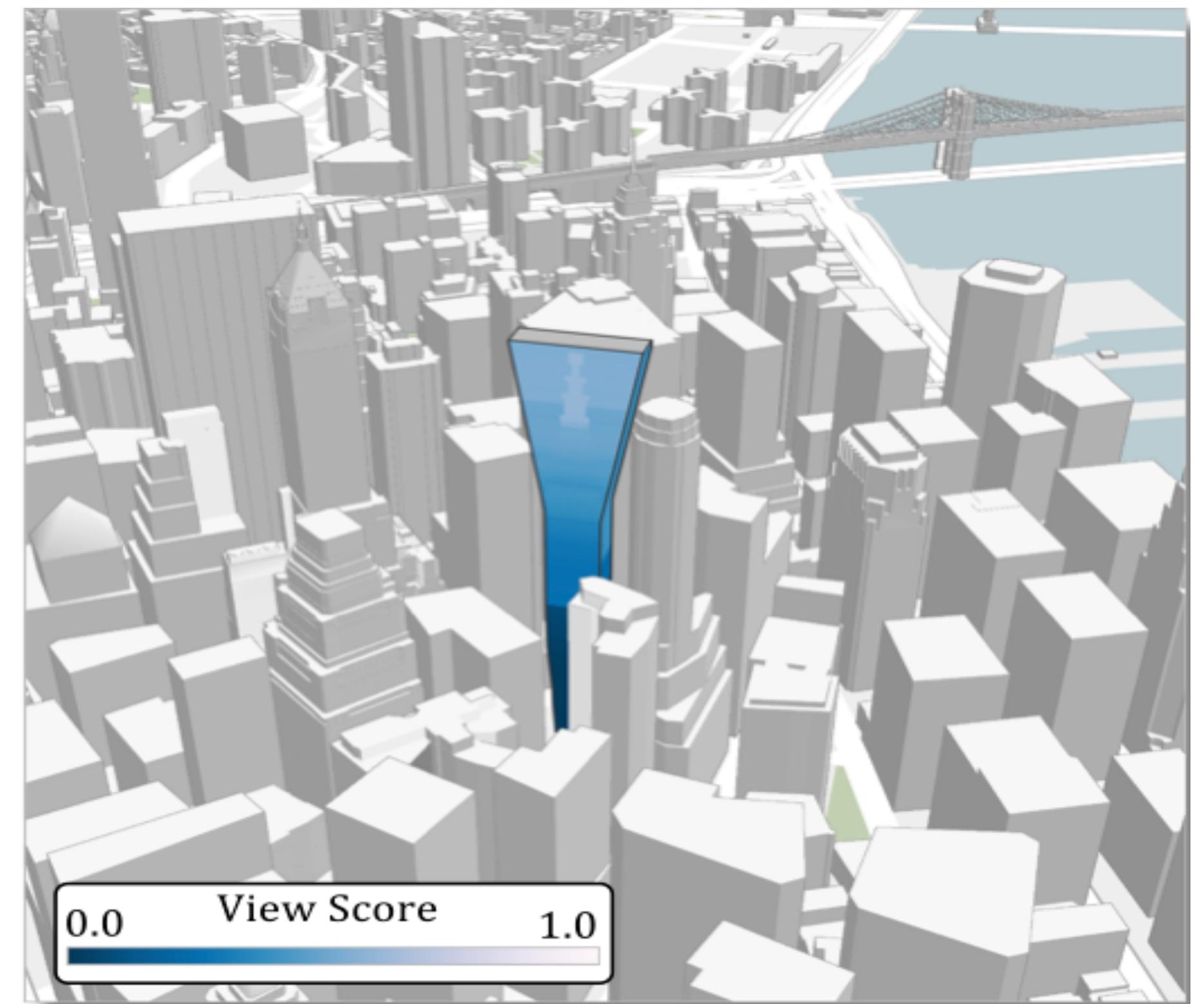
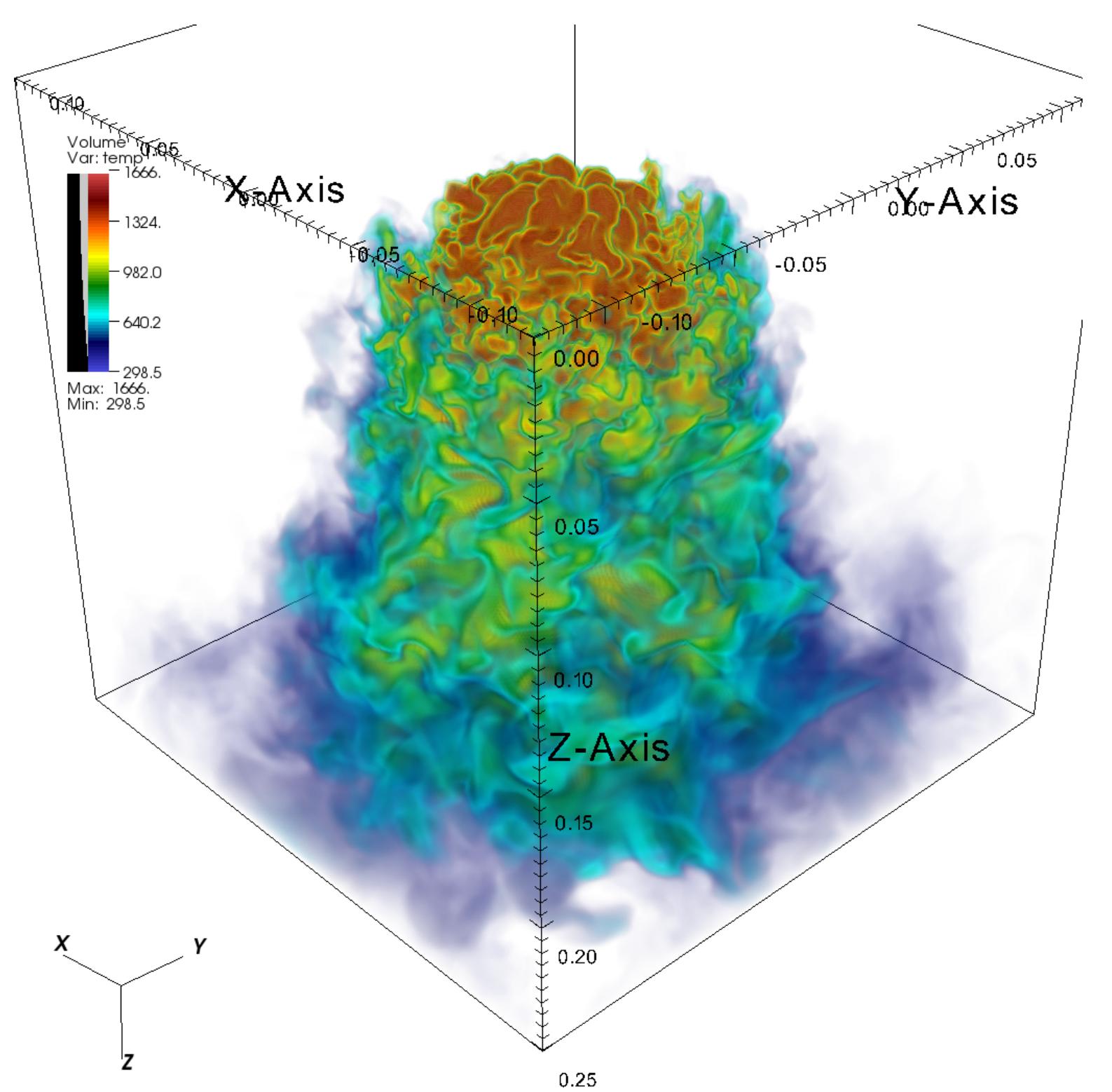


Figure 1: A bird's eye view of a tabular visualization of the results of a text search. Cubes representing retrieved documents are placed from left to right and front to back based on their total number of keyword matches. The width, height and depth of the cubes are matched to the frequency of individual keywords.

Rule of Thumb #1

3D é recomendado para problemas onde forma geométrica é importante

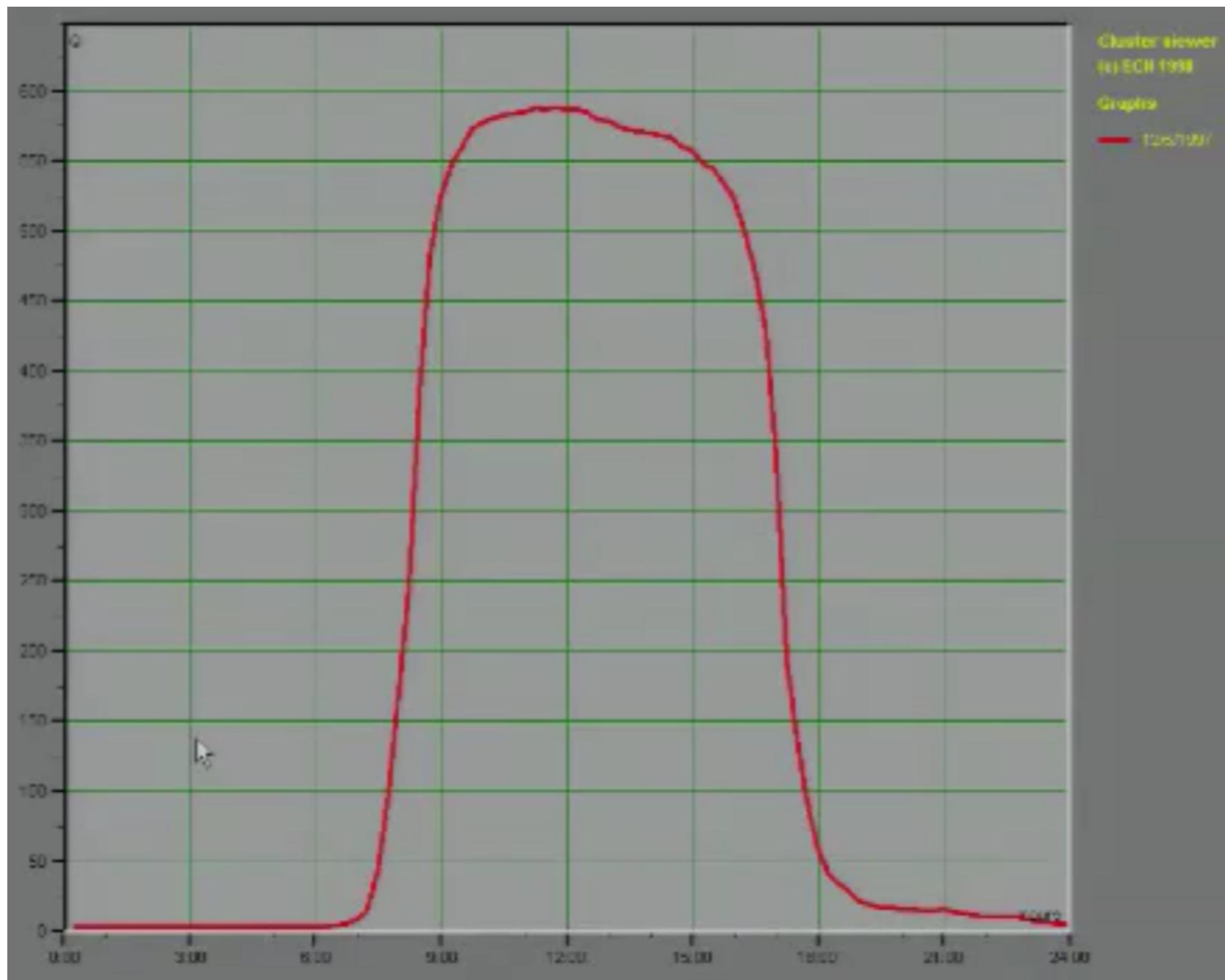
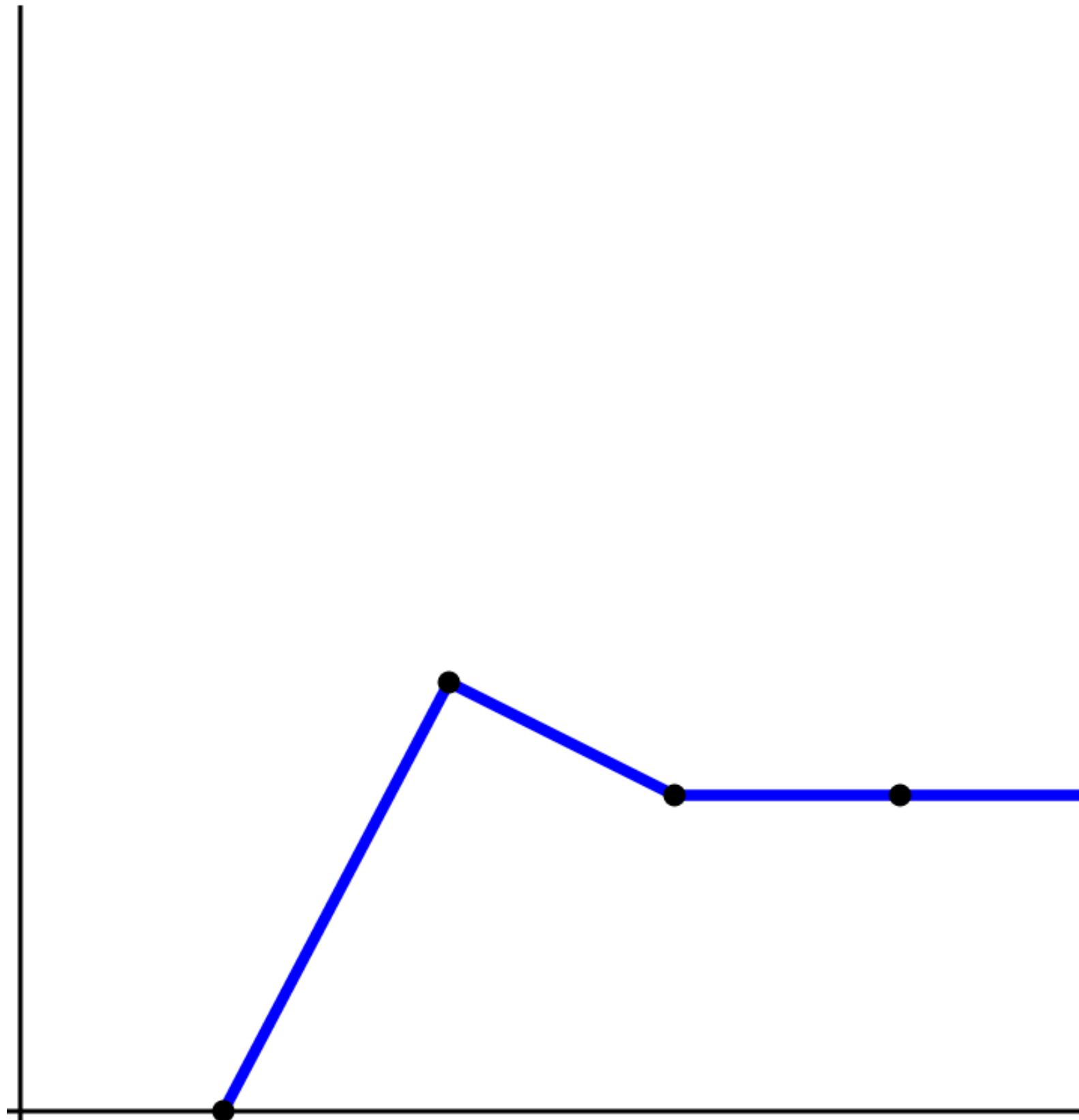


Back to the Drawing Board

What?

Why?

How?



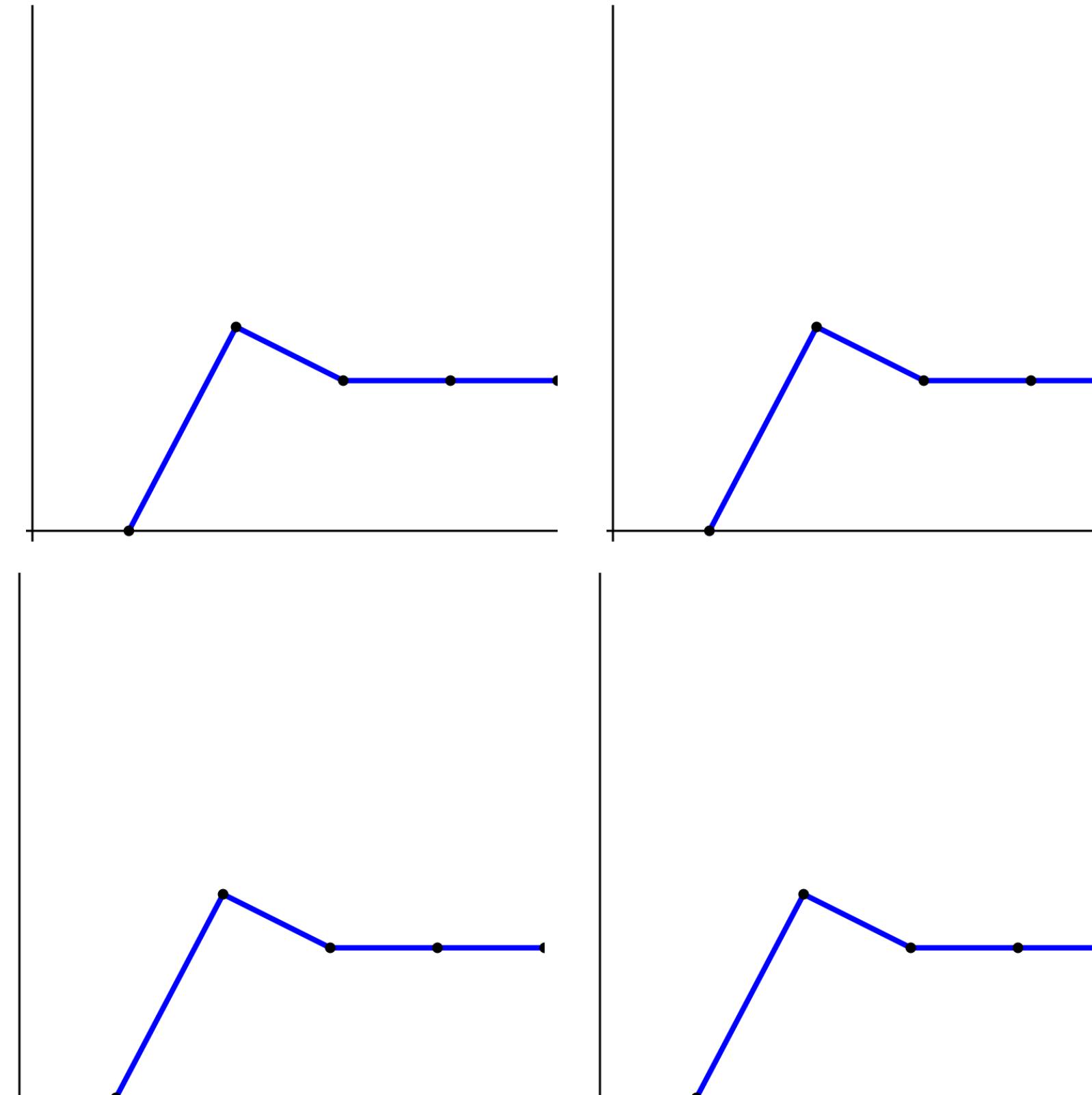
Back to the Drawing Board

- Queremos comparar diversas séries-temporais

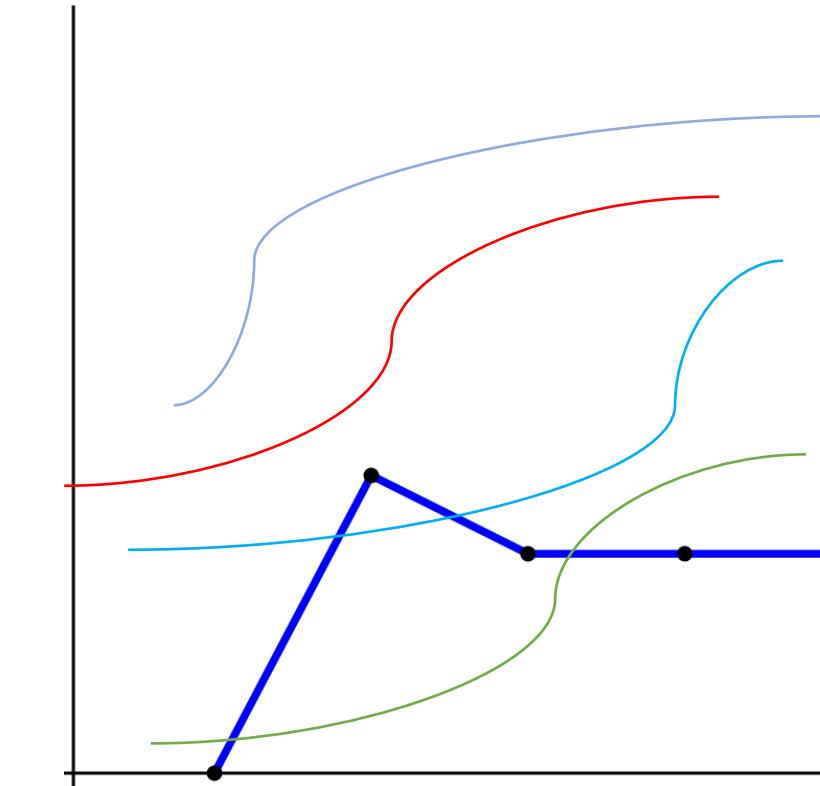
What?

Why?

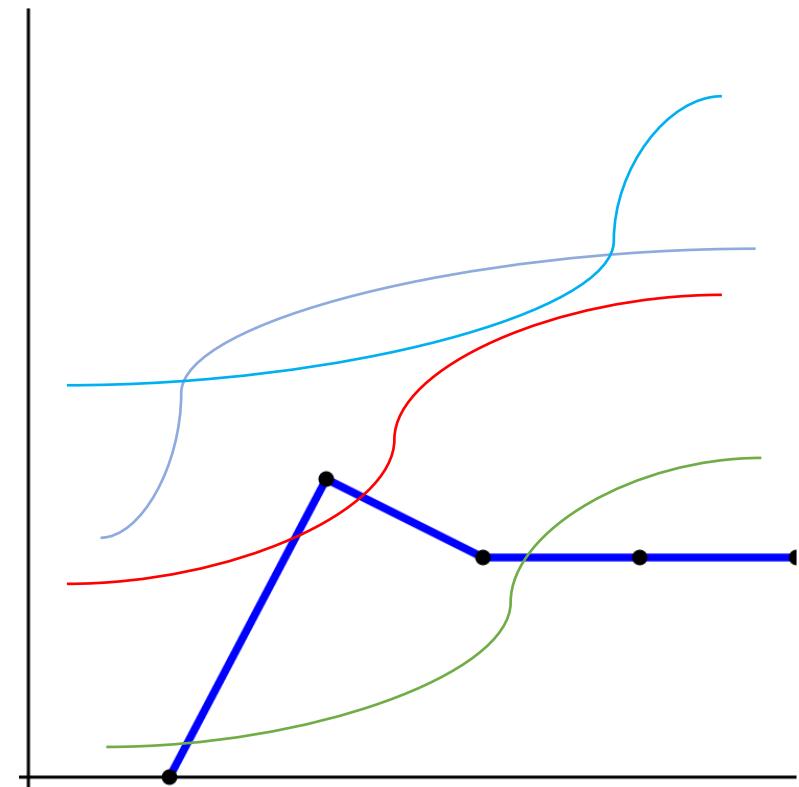
How?



Justaposição



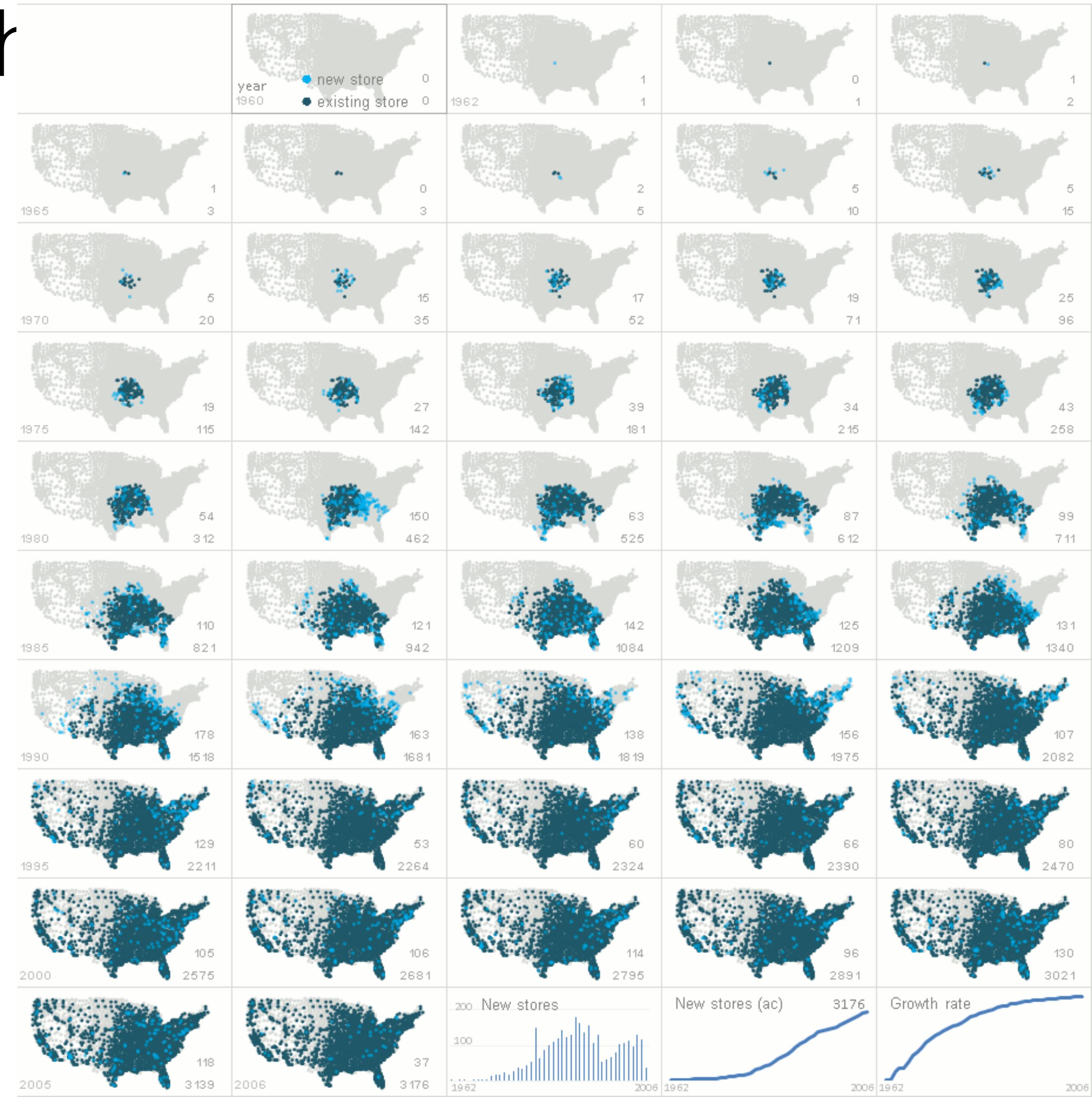
Mesmo Plot



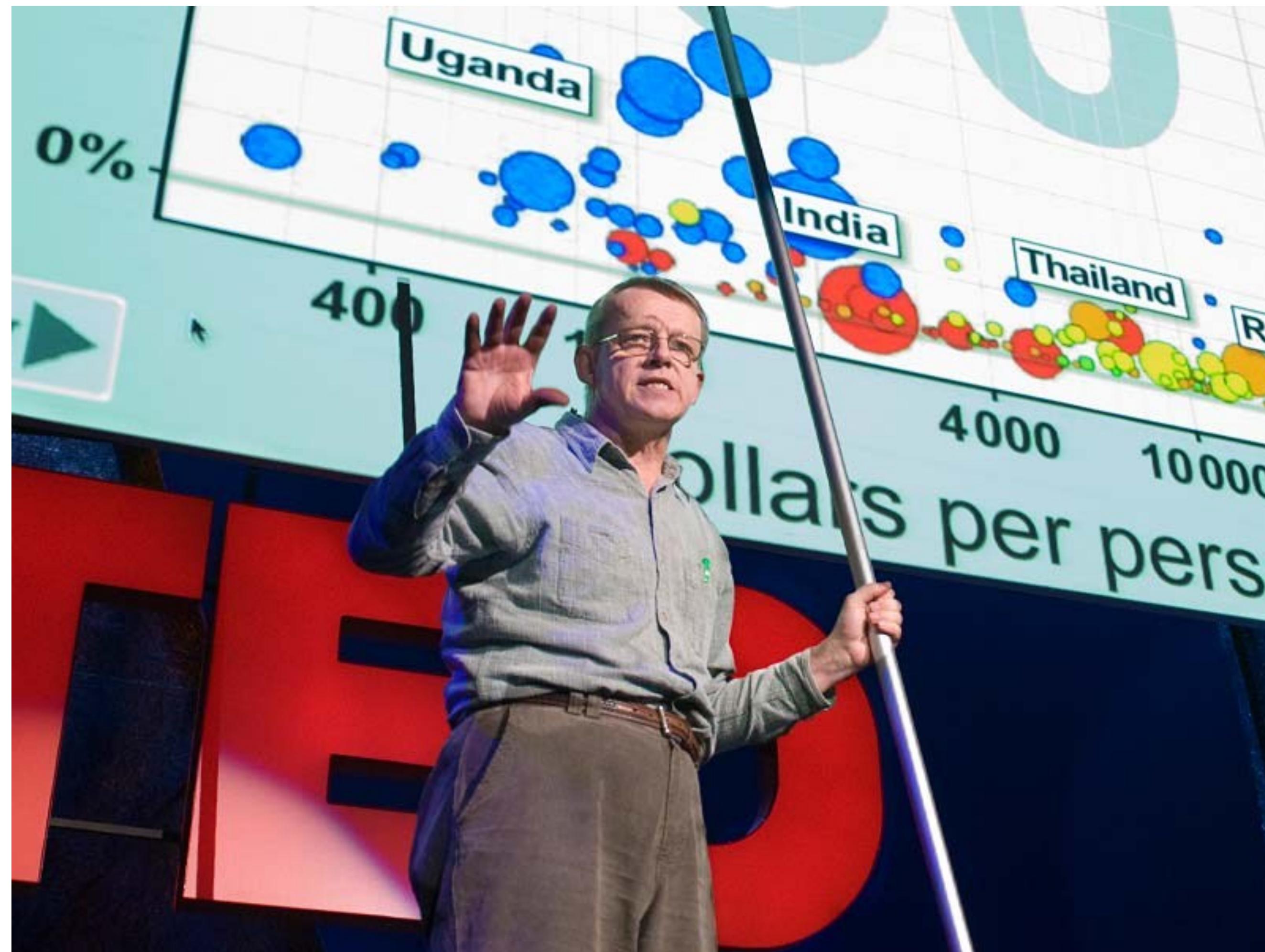
Animações

Rule of Three Memória

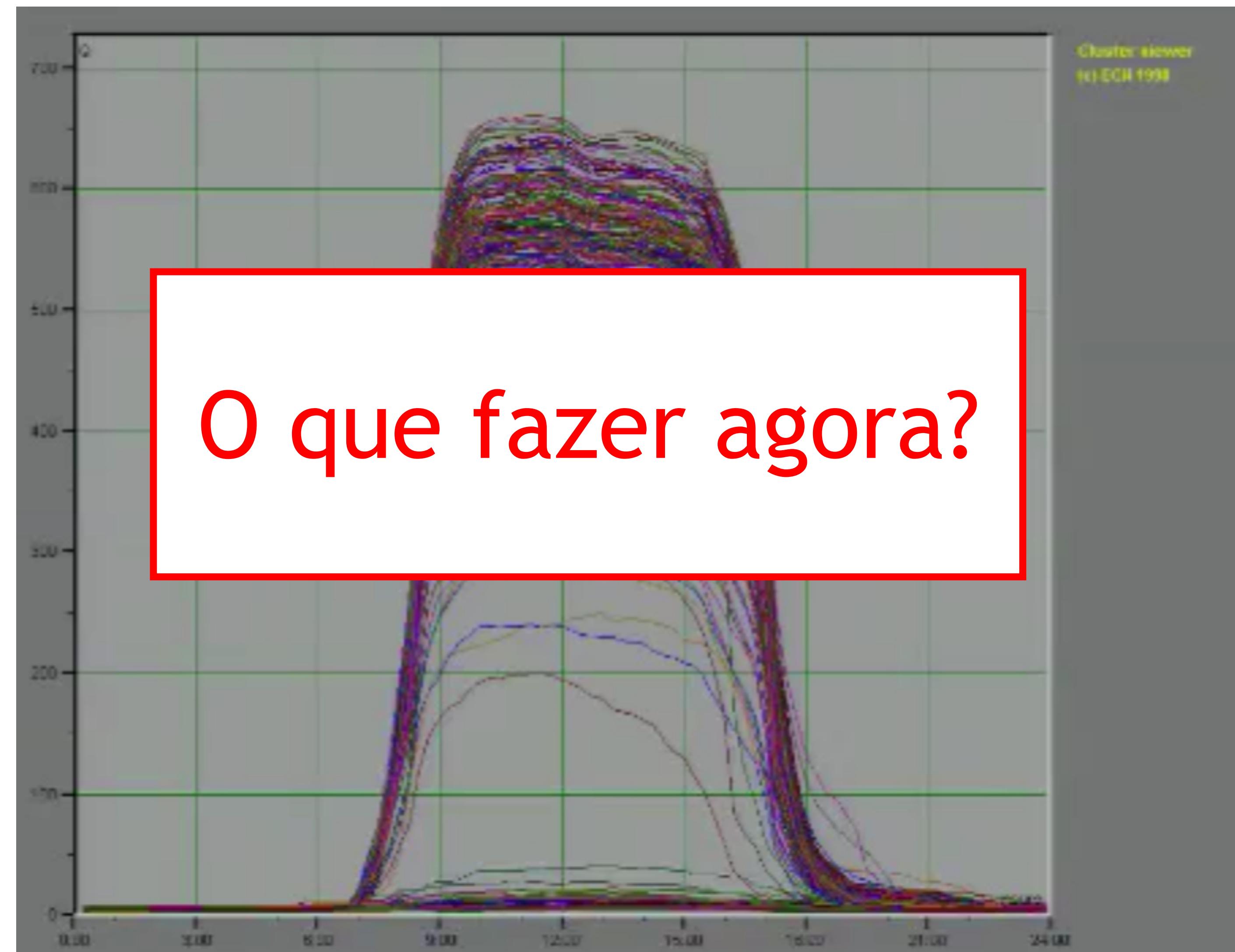
Growth of Walmart



Rule of Thumb #2: Visão Preferível a Memória



Back to the Drawing Board



What?

Why?

How?

Ideia: Sumarizar Dados

- Clustering
 - Agrupar séries-temporais semelhantes
 - Clustering Hierárquico

What?

Why?

How?

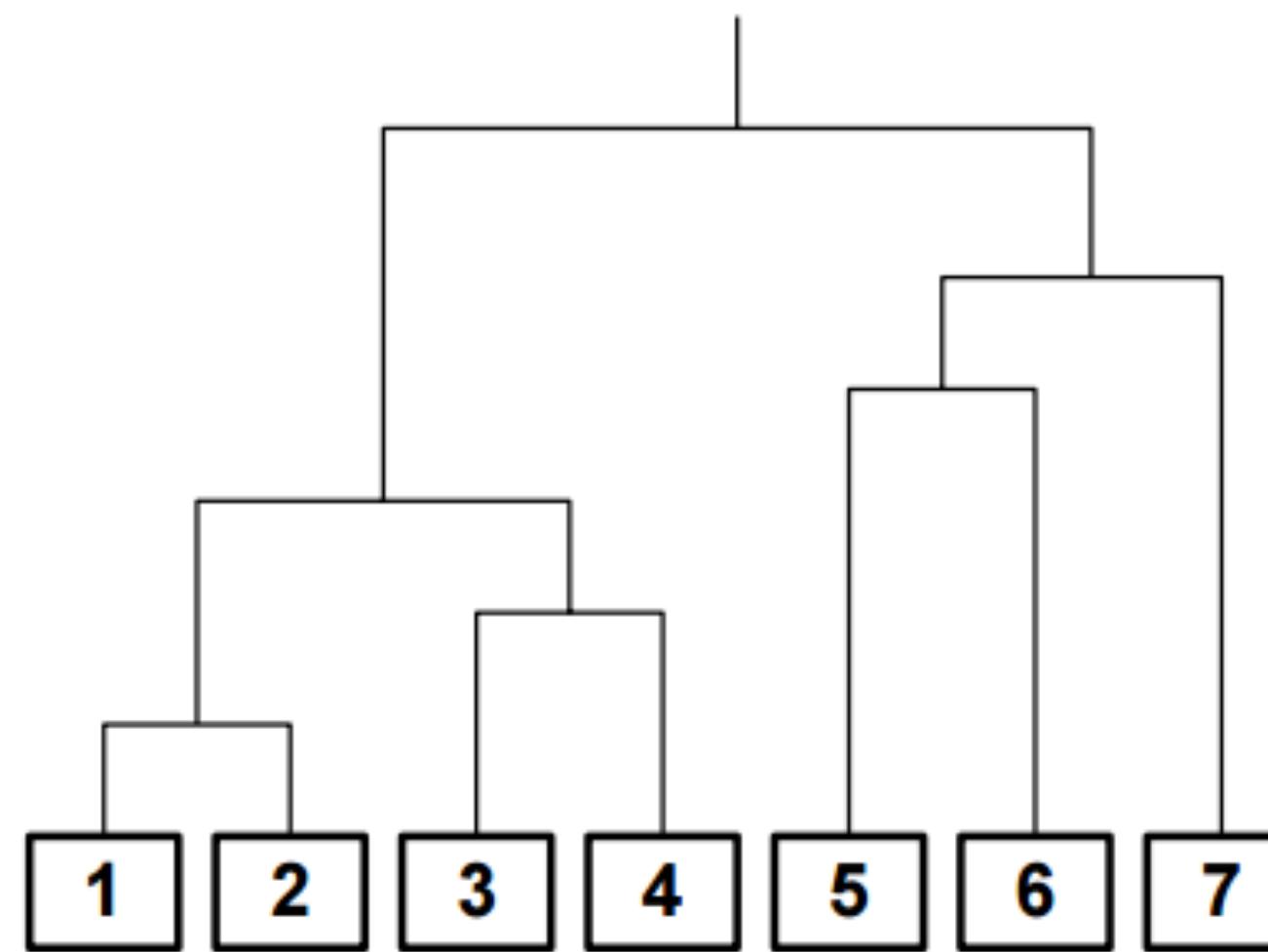


Figure 2. Dendrogram

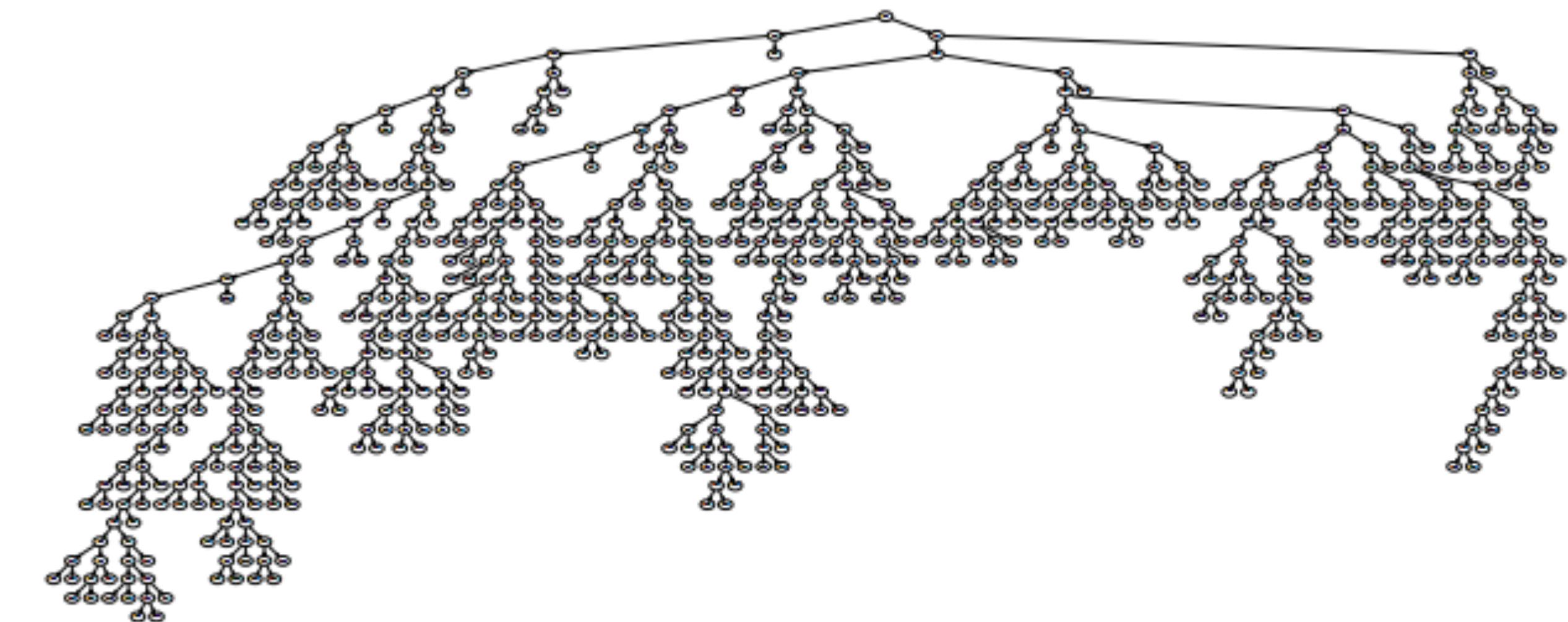
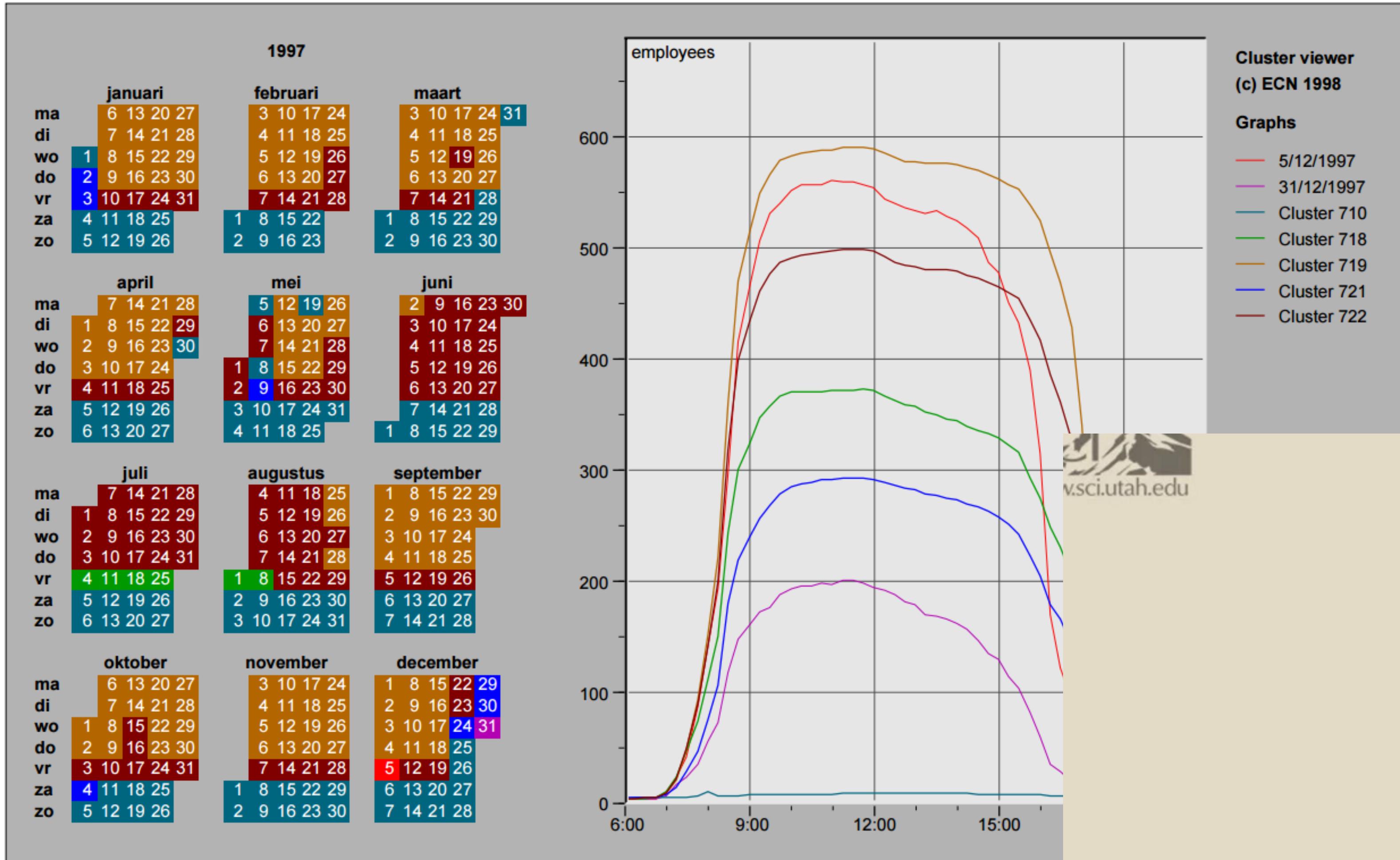


Figure 3. Full clustering tree

Solução



Visual Analytics

Vis + Stats + ML é uma boa ideia!