

# Mapas mentais

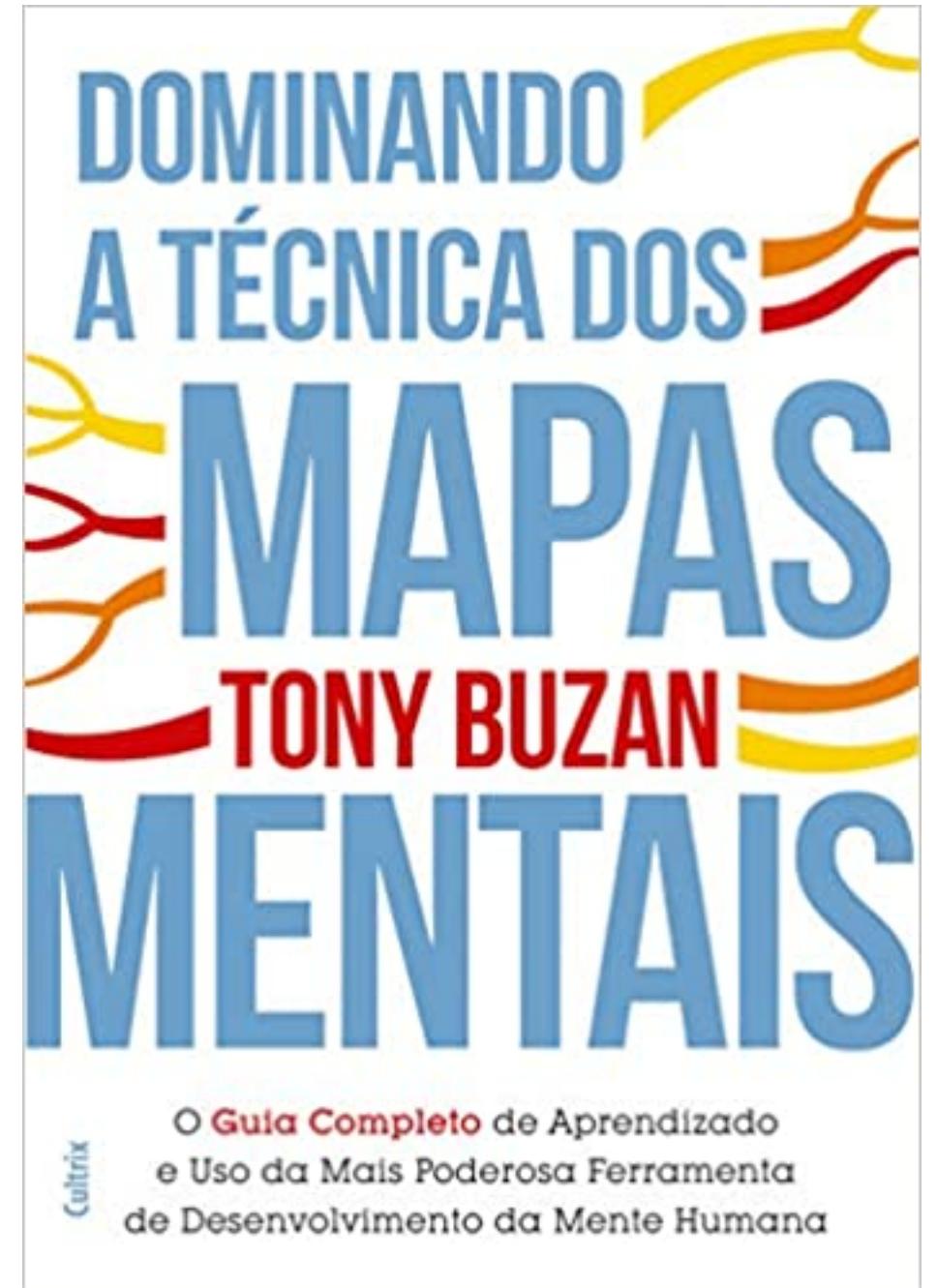
---

**Profa. Dra. Raquel Minardi**  
Departamento de Ciência da Computação  
Universidade Federal de Minas Gerais

“

Mapa mental é um tipo de **diagrama**, sistematizado pelo psicólogo inglês Tony Buzan, voltado para a **gestão de informações**, de conhecimento e de capital intelectual; para a **compreensão e solução de problemas**; na **memorização e aprendizado**; na criação de manuais, livros e palestras; como **ferramenta de brainstorming** e no auxílio da gestão estratégica de uma empresa ou negócio.

—Wikipedia



► Para **criar um mapa mental**, você precisará de:

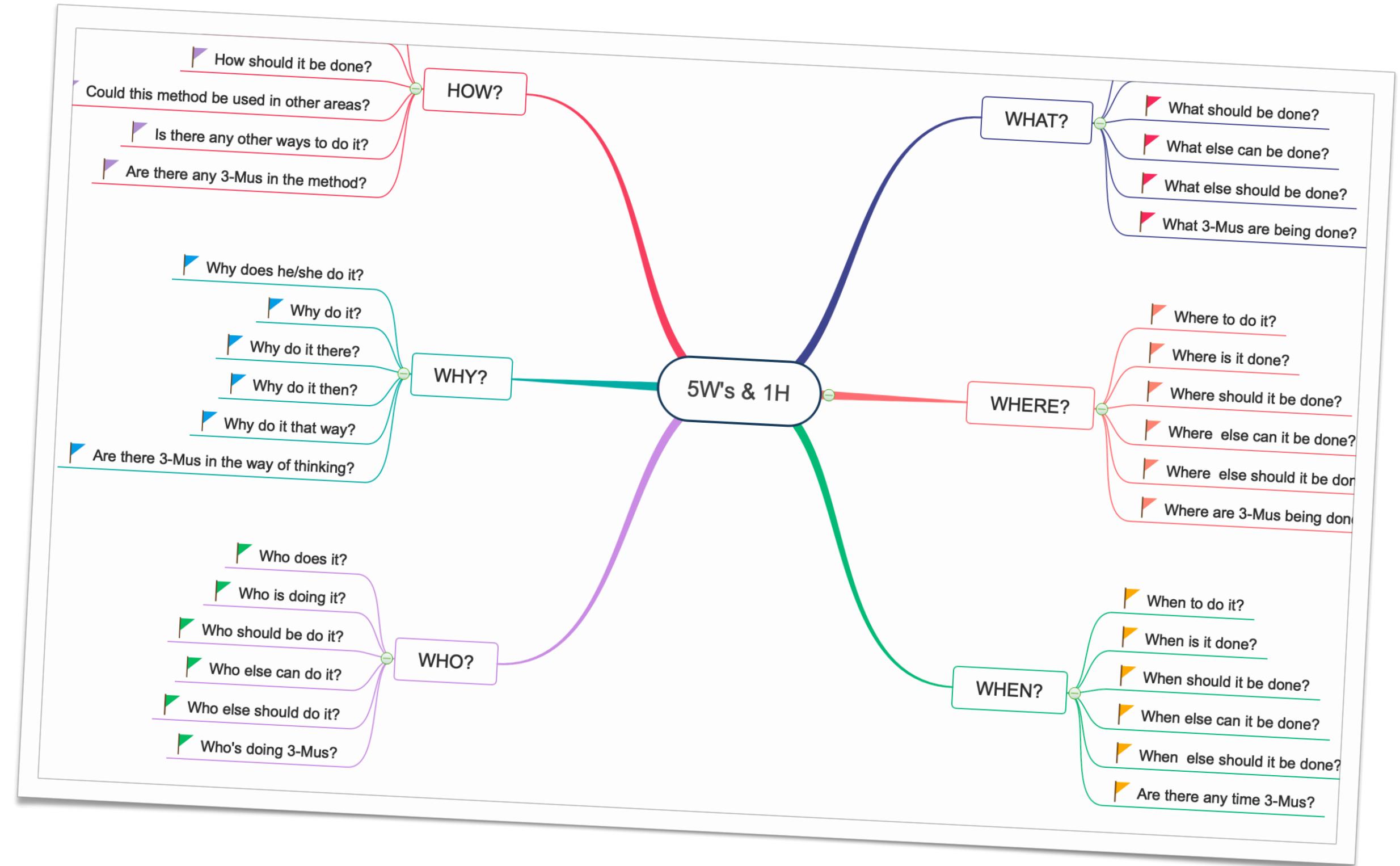
- Uma folha grande de papel em branco
- Um conjunto de canetas ou lápis coloridos

ou

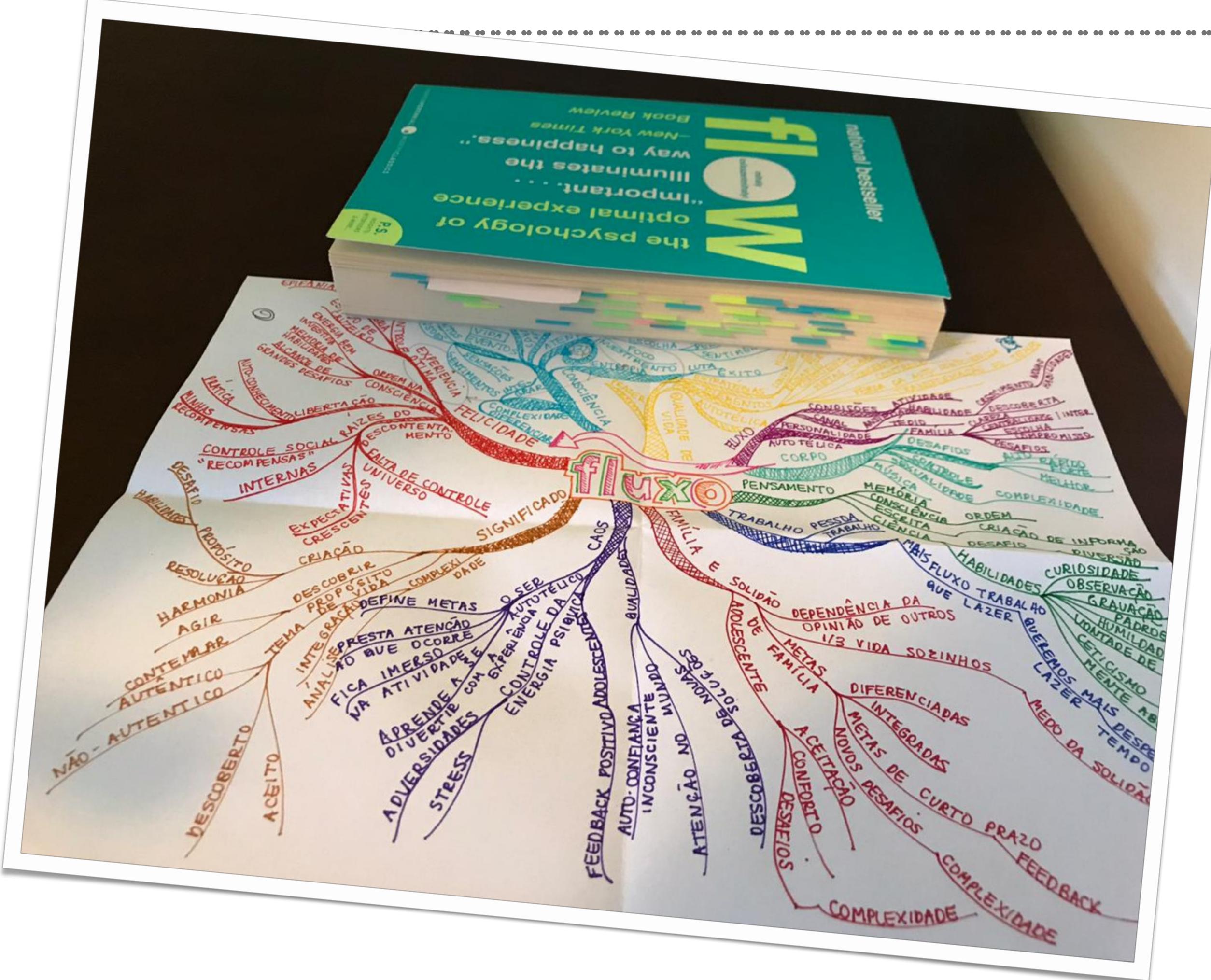
- Um programa para criação de mapas mentais

# UM MAPA MENTAL TEM AS SEGUINTE CARACTERÍSTICAS

- Uma **imagem central**: não é preciso ter habilidades artísticas.
- **Ramificações grossas que se irradiam** da imagem central: representam os temas fundamentais relacionados ao assunto principal e cada uma delas será representada em uma cor diferente. Delas, brotam ramificações secundárias.
- **Uma única imagem ou palavra-chave** é colocada em cada ramificação.



# UM MAPA MENTAL TEM AS SEGUINTE VANTAGENS



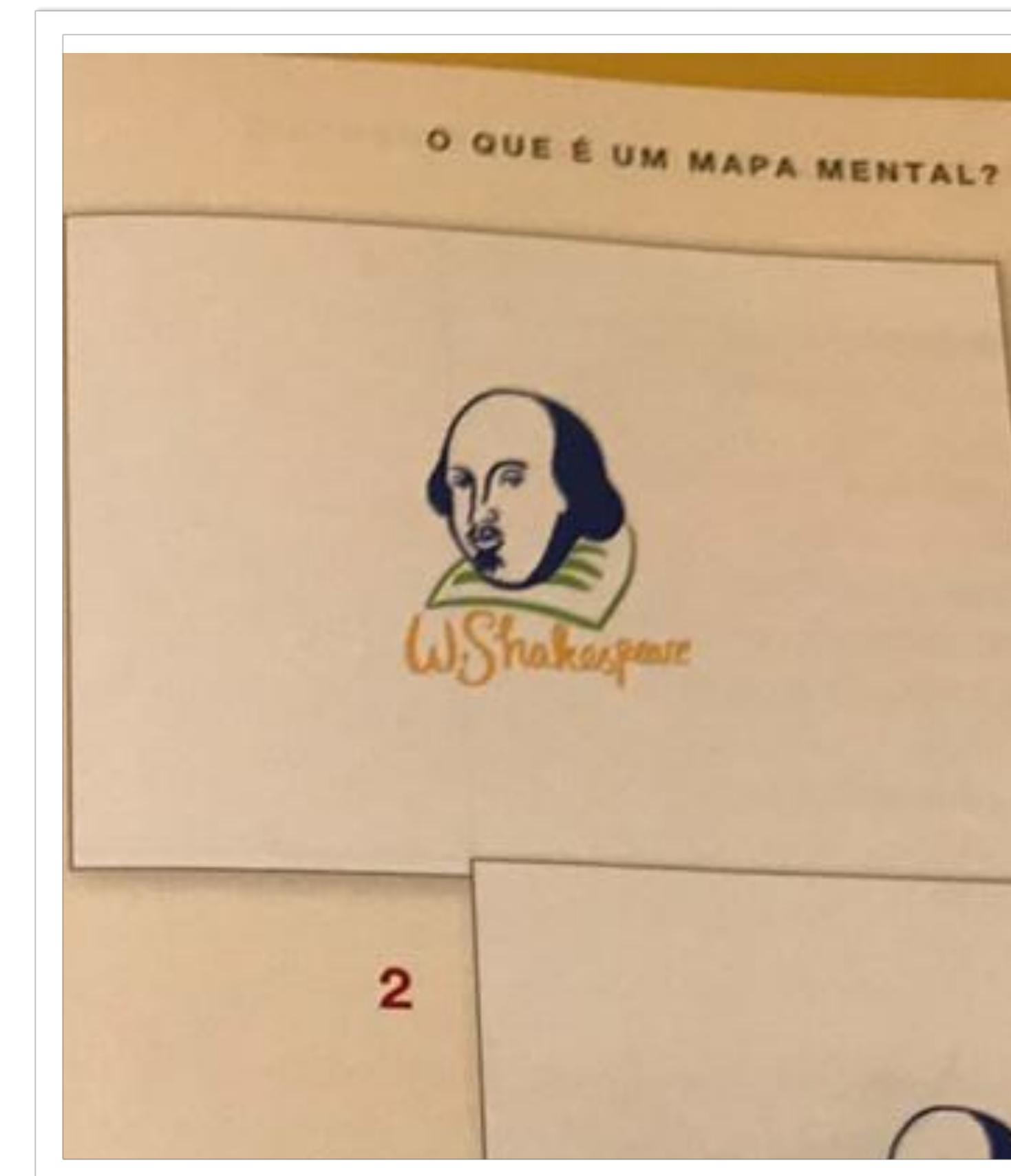
- Capturam nossa **atenção**
- Melhorar a nossa **compreensão**
- Aumentam a **motivação**
- Estimulam a **comunicação**
- Intensificam o processamento e **armazenamento de imagens pela mente**

# PASSOS PARA ELABORAR UM MAPA MENTAL

---

# 1

- Separe uma **folha** e a use na **horizontal**.
- Use pelo menos **3 cores** para desenhar a **imagem central** que representa o tema que você está resumindo.
- Pode ser uma palavra mas tente escrevê-la em destaque, colorida, tridimensional.



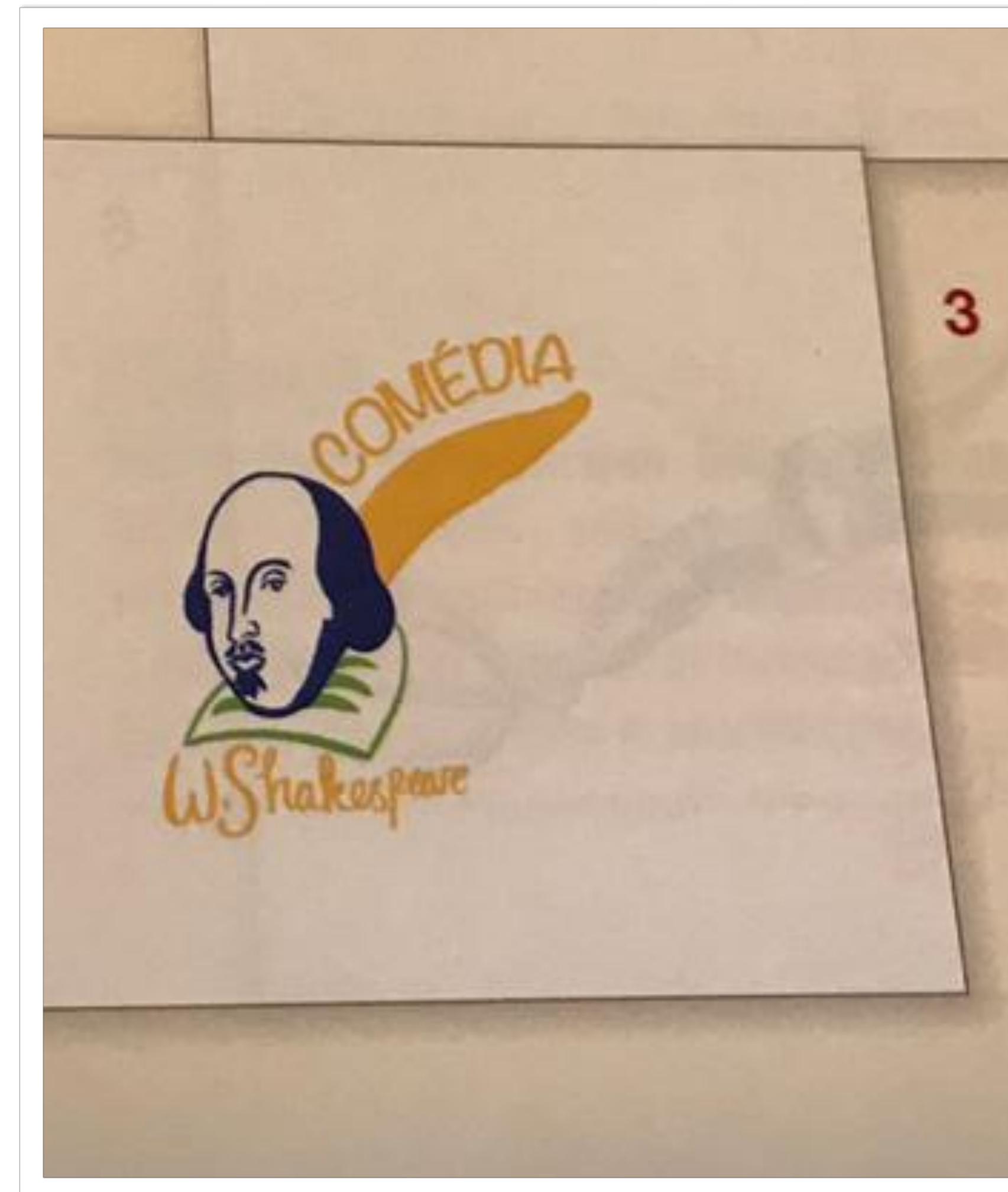
# 2

- Escolha uma cor e desenhe uma **ramificação bem grossa** que nasce na imagem central.
- A grossura da ramificação simboliza sua **importância**.
- A ramificação deve se **curvar organicamente**, pois isso será envolvente no aspecto visual e mais interessante para seu cérebro, fazendo você memorizar as informações contidas naquela ramificação.



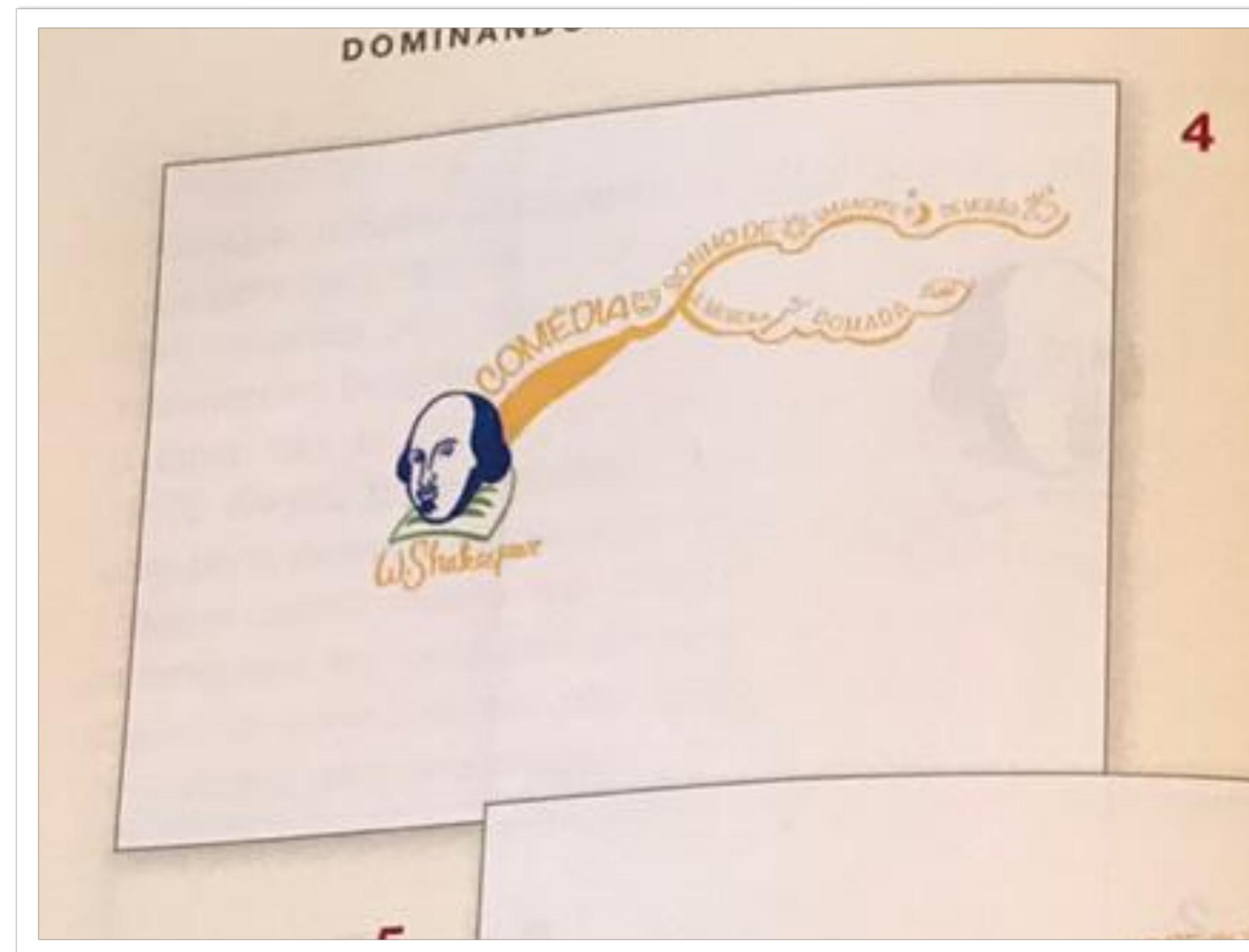
# 3

- Rotule a ramificação com uma única palavra em letras maiúsculas.



# 4

- Faça brotar **ramificações secundárias** e depois **terciárias** que se irradiem a partir dos brotos secundários.
- Escreva as **palavras-chave**, desenhe símbolos ou use uma combinação das duas coisas em todas as ramificações.



# 5

- Escolha outra cor e crie a **próxima ramificação principal**.
- **Continue acrescentando ramificações** até ter umas cinco ou ser com que trabalhar.



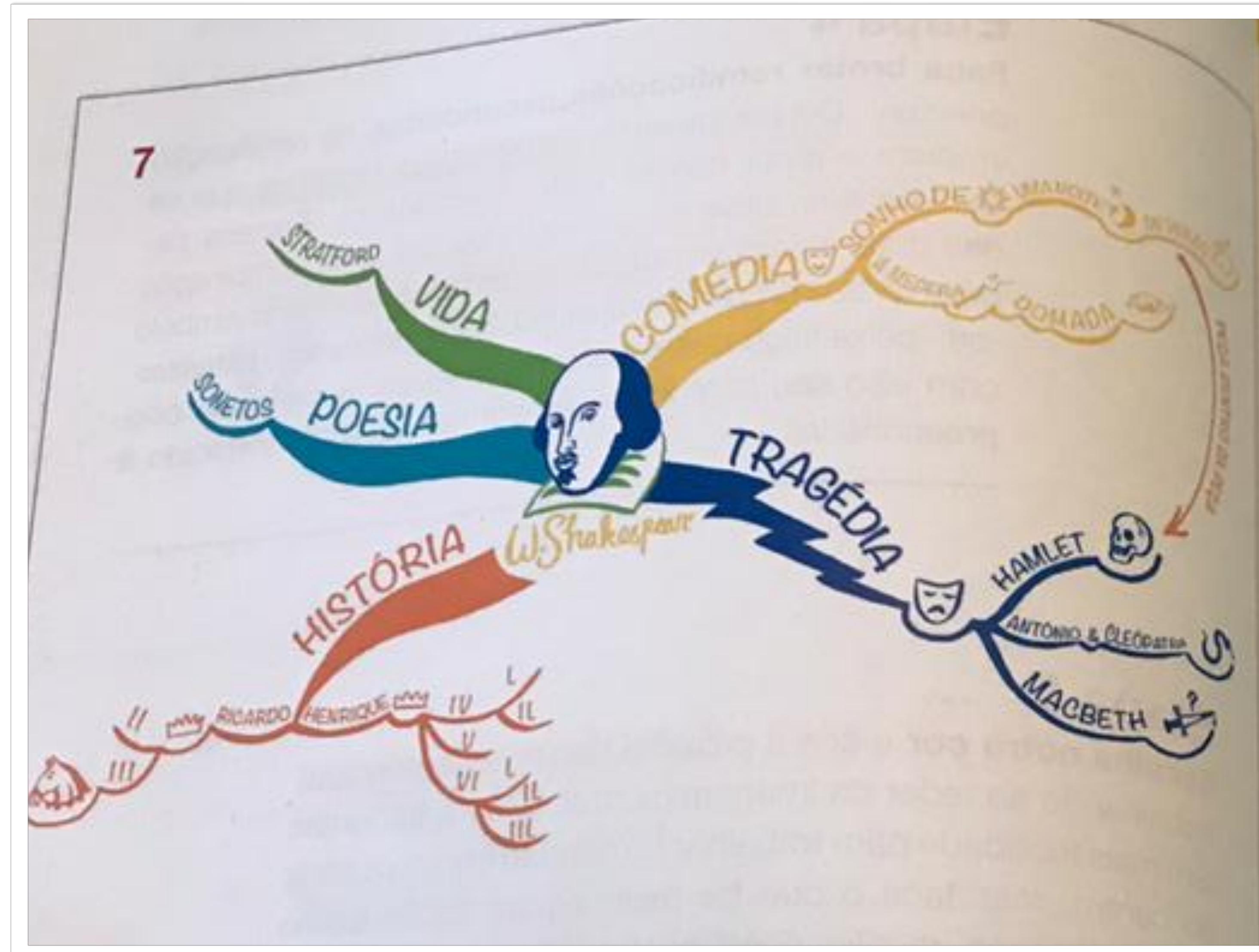
# 6

- Vá **preenchendo** suas ramificações com **palavras-chave** a medida que lhe ocorrerem idéias e associações.



# 7

- Se quiser, acrescente setas, linhas curvas e ligações entre as ramificações principais a fim de reforçar as conexões entre elas.



**Profa. Dra. Raquel Minardi**

# **VISUALIZAÇÃO DE DADOS**

Departamento de Ciência da Computação  
Universidade Federal de Minas Gerais

# Motivação e apresentação do curso

---

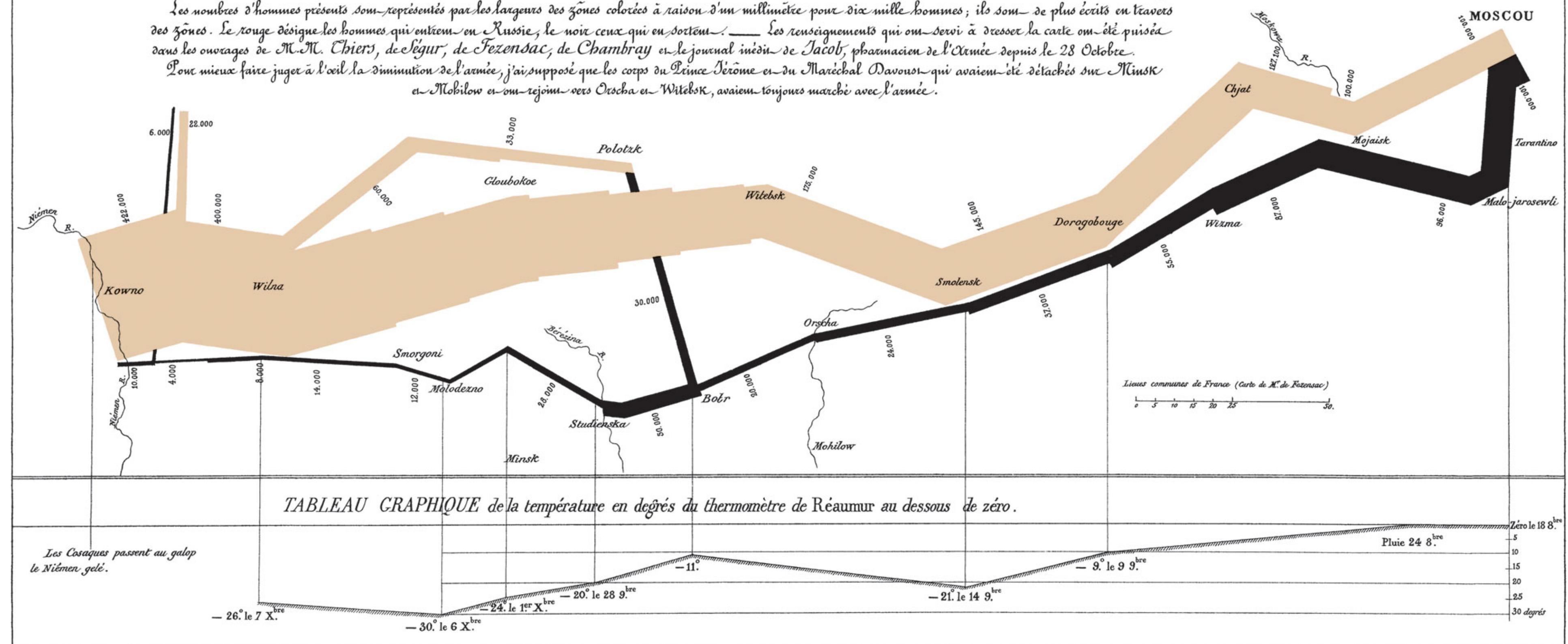
**Profa. Dra. Raquel Minardi**  
Departamento de Ciência da Computação  
Universidade Federal de Minas Gerais

# Carte Figurative des pertes successives en hommes de l'Armée Française dans la Campagne de Russie 1812-1813.

Dressée par M. Minard, Inspecteur Général des Ponts et Chaussees en retraite Paris, le 20 Novembre 1869.

Les nombres d'hommes présents sont représentés par les largeurs des zones colorées à raison d'un millimètre pour dix mille hommes; ils sont de plus écrits en travers des zones. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. — Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M. M. Chiers, de Séjourné, de Fezensac, de Chambray et le journal médical de Jacob, pharmacien de l'Armée depuis le 28 Octobre.

Pour mieux faire juger à l'œil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davout, qui avaient été détachés sur Minsk et Mohilow en se rejoignant vers Orscha et Wilebsk, avaient toujours marché avec l'armée.

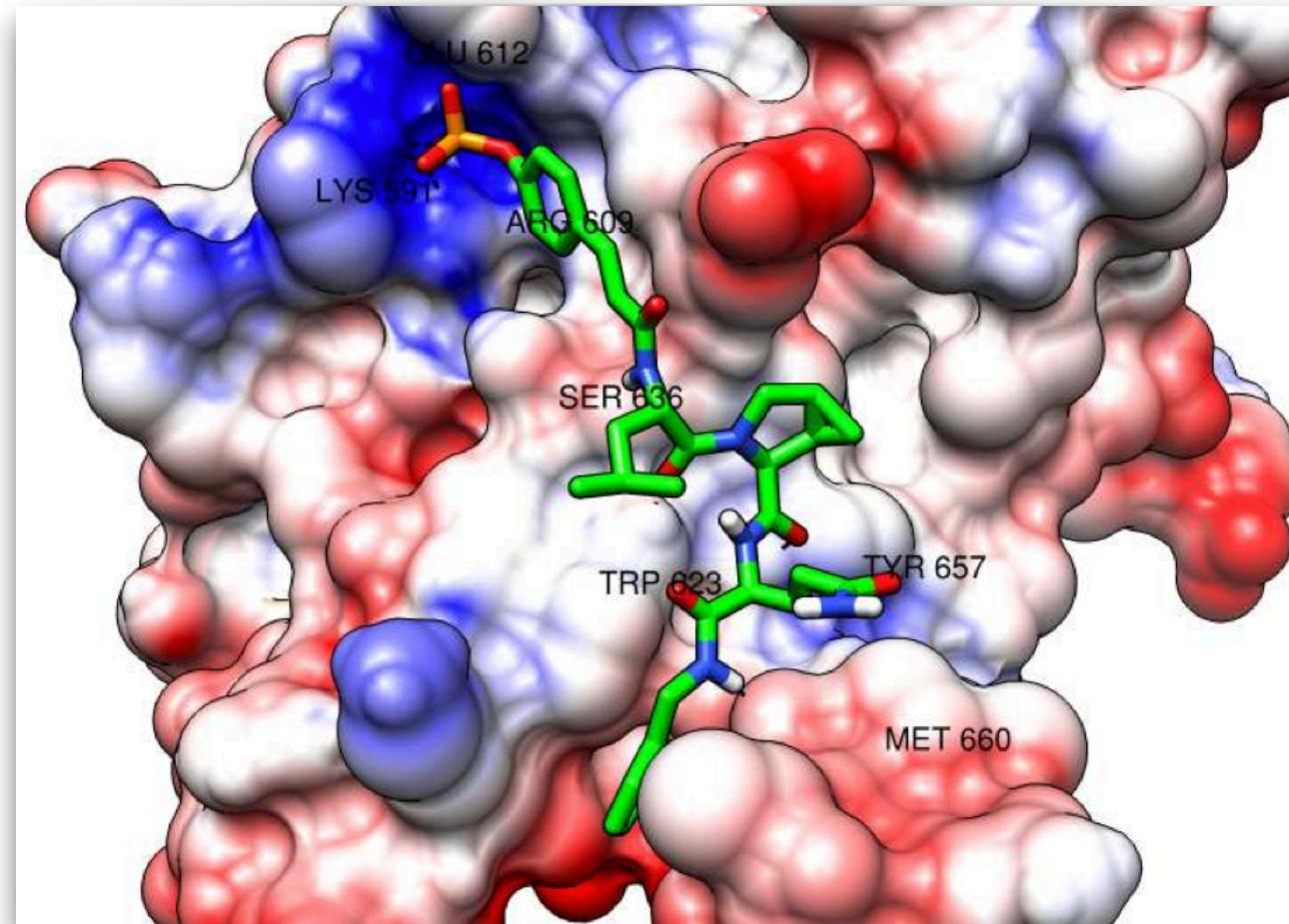


Charles Minard, campanha de 1812, 1869

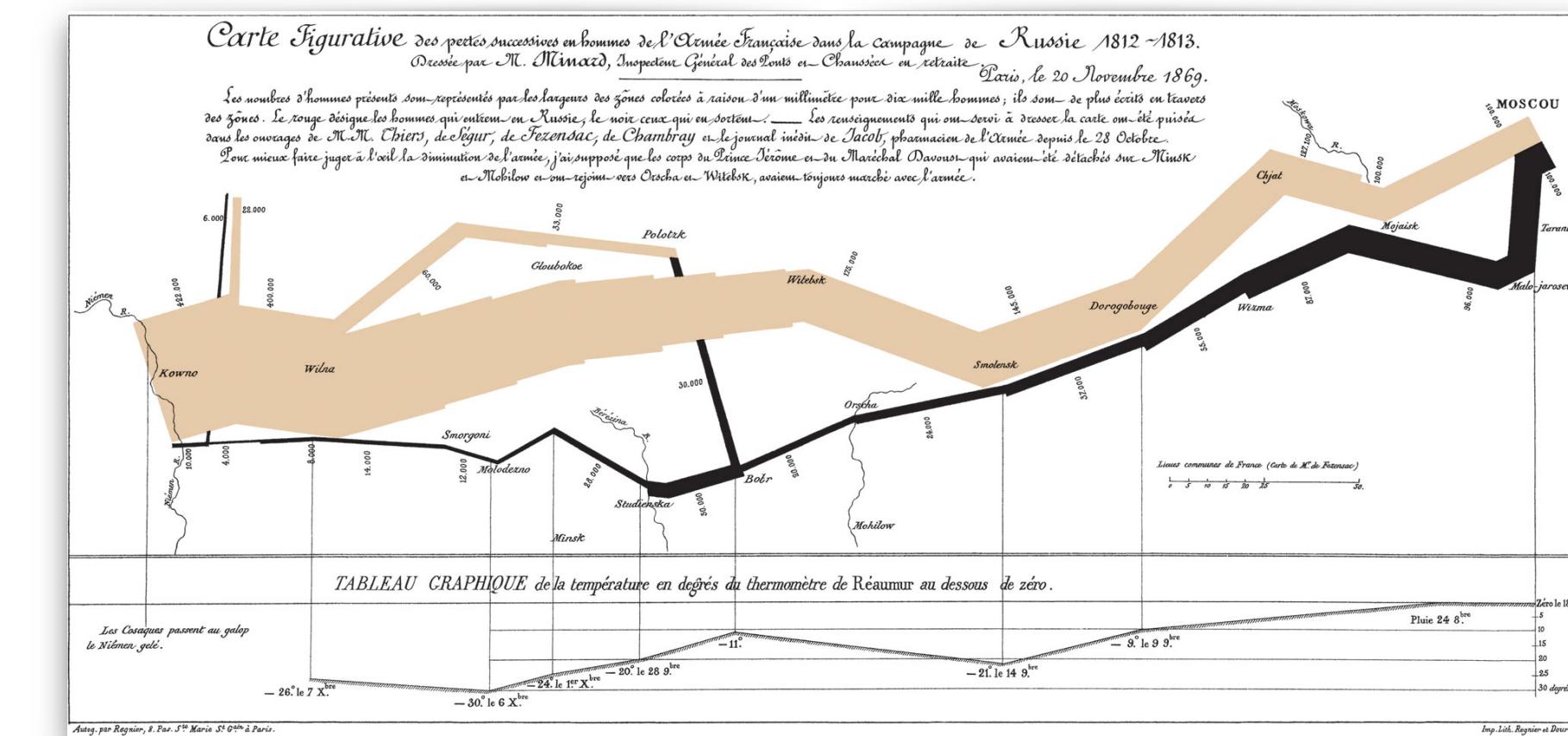
# O que é **visualização de dados**?

É um “termo guarda-chuva” que envolve a **visualização de informação** e a **visualização científica**

A **visualização científica** ilustra objetos que tem um correspondente físico



A **visualização de informação** (ou de dados) representa objetos mais abstratos



**McCafé**

**blended**

Item	Calories	Price
Real Fruit Smoothies	210-350 Cal.	0.00 0.00 0.00
mango pineapple   strawberry banana   wild berry		
Frappé	450-680 Cal.	0.00 0.00 0.00
mocha   caramel		
espresso & chocolate	with whole or nonfat milk	
Mocha	240-400 Cal.	0.00 0.00 0.00
Caramel Mocha	200-360 Cal.	0.00 0.00 0.00
Iced Mocha	230-390 Cal.	0.00 0.00 0.00
Iced Caramel Mocha	200-380 Cal.	0.00 0.00 0.00
Latte	80-330 Cal.	0.00 0.00 0.00
Iced Latte	40-230 Cal.	0.00 0.00 0.00
Hot Chocolate	250-460 Cal.	0.00 0.00 0.00

**brewed**

Item	Calories	Price
Premium Roast Coffee	0 Cal.	0.00 0.00 0.00
Iced Coffee	90-280 Cal.	0.00 0.00 0.00
made with cream		
choose a flavor	caramel   hazelnut   vanilla   sugar-free vanilla	

**LIMITED TIME**

**new FRAPPE CHOCOLATE CHIP**

Size	Calories	Price
sm.	0.00	0.00
med.	0.00	0.00

**DAILY DOUBLE**

Item	Calories	Price
Soft Drink	0.00	0.00 0.00
Iced Tea	0 Cal.	0.00 0.00 0.00
Orange Juice	150   190 Cal.	0.00 0.00 0.00
Milk	low-fat white 100 Cal.   fat-free chocolate 130 Cal.	0.00 0.00
Dasani Bottled Water	0 Cal.	0.00 0.00

**sweet treats**

Item	Calories	Price
Chocolate Shake	710 Cal.	0.00
Oreo' McFlurry™	340 Cal.	0.00
McFlurry™	340-430 Cal.	0.00
Sundae	280-340 Cal.	0.00
Soft Baked Cookies	150-160 Cal. ea.	0.00
Pies	250 Cal. ea.	0.00
Cone	170 Cal.	0.00
fries	230 Cal. 380 Cal. 500 Cal.	0.00 0.00 0.00

**extra value menu**

**20 PC. CHICKEN McNUGGETS**

**0 0 0**

**extra value menu**

**DAILY DOUBLE**

**0 0 0**

**med. meal**

**0 0 0**

**sandwich**

**extra value menu**

**dollar menu \$1**

**lunch & dinner**

**McDouble®** 390 Cal.

**Daily Double®** 440 Cal.

**20 Pc. Chicken McNuggets®** 0.00

**serves 2** 470 Cal./ea.

**Angus Snack Wrap®** 390-430 Cal.

**Angus Snack Wrap®** 390-430 Cal.

**Chicken Snack Wrap®** 230-350 Cal.

**Double Cheeseburger** 440 Cal.

**Iced Coffee** med. 90-200 Cal.

**McFlurry™** snack size 340-430 Cal.

**Local Option**

**Local Option**

**breakfast** (available during breakfast hours only)

**Sausage Burrito** 300 Cal.

**Sausage McMuffin®** 370 Cal.

**Sausage Biscuit** 430 Cal.

**Hash Browns** 150 Cal./ea.

**Coffee** sm. 0 Cal.

**Local Option**

**Julius by paul frank.**

**LIMITED TIME**

**new SPICY CHICKEN McBites®**

**POPEM WHILE THEY'RE HOT!**

**0 0 0**

**regular size meal**

**Choose your FAVORITES UNDER 400 CALORIES**

**200** calories or under each

**400** calories or under each

**300** calories or under each

**100** calories or under each

**200** Iced Coffee (reg.)

**170** Vanilla Cone

**150** Parfait

**150** Hash Browns

**130** Chocolate Milk

**290** Southwest Salad

**380** Medium Fries

**380** Premium Chicken Sandwich

**380** Filet-O-Fish®

**350** Snack Wrap®

**340** OREO® McFlurry®

**240** Caramel Mocha

**230** Small Fries

**20** Side Salad

**fruit & oatmeal**

**Fruit & Maple Oatmeal** 290 Cal.

**Fruit & Walnuts** 210 Cal.

**happy meal**

**includes apple slices & kids fries**

**4 Pc. Chicken McNuggets®** 305 Cal. 0.00

**Hamburger** 365 Cal. 0.00

**Cheeseburger** 415 Cal. 0.00

**mighty kids meal®** (includes apple slices & small fries)

**6 Pc. Chicken McNuggets®** 525 Cal. 0.00

**McDouble®** 635 Cal. 0.00

**add McNuggets® Sauces** 30-110 Cal. ea.

**choose a drink for your meal** (additional charges may apply)

1% Low-Fat White Milk +100 Cal. Kids Soft Drink +0-120 Cal. Fat Free Chocolate Milk +130 Cal. Small Soft Drink +0-160 Cal. 100% Apple Juice +100 Cal. Mighty Kids Meal only

**With every Happy Meal or Mighty Kids Meal purchase, a donation is made to Ronald McDonald House Charities®**

**McDonald's®** A registered trademark of McDonald's Corporation. ©2008 McDonald's System of Restaurants, Inc. All rights reserved. "McDonald's" and "McDonald's Golden Arches" are trademarks of McDonald's Corporation.

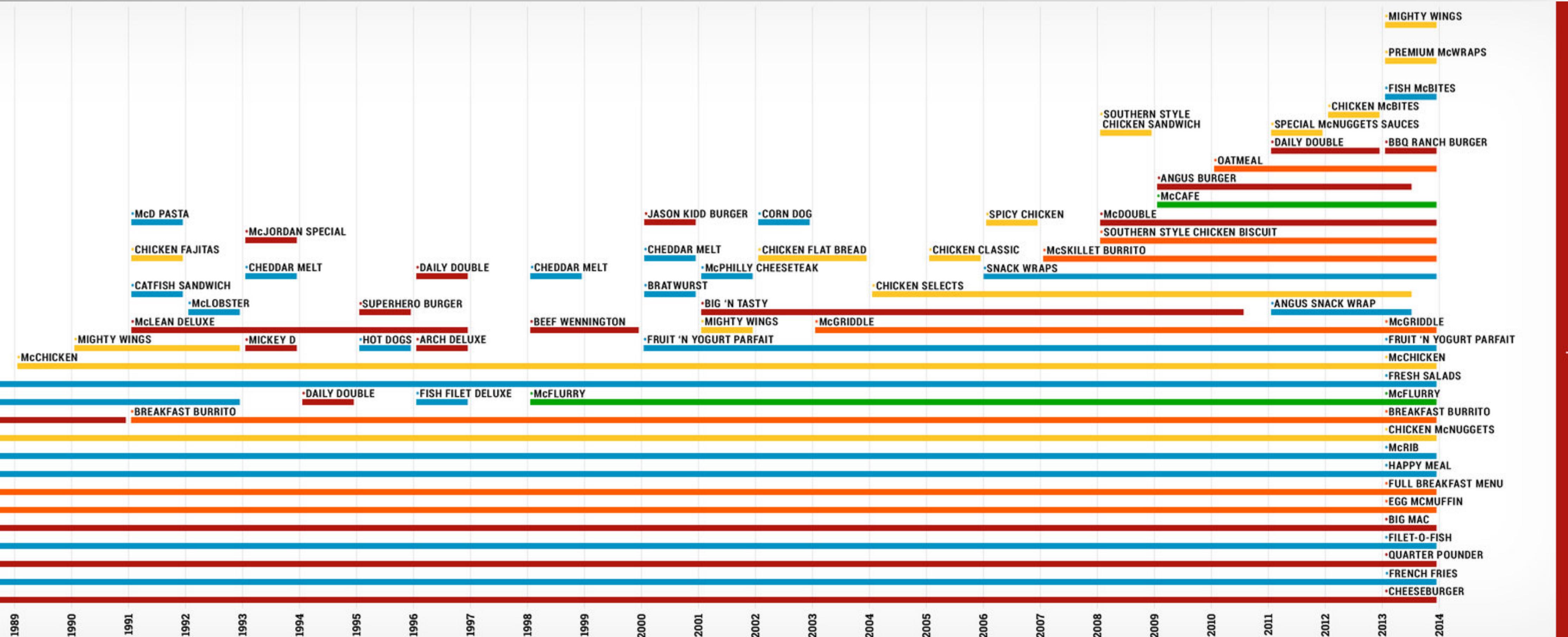
**pairfection!**

**extra value meals**

include medium fries and soft drink  
large fries and soft drink +... add 120-230 Cal.

Meal	Calories	Price
1 Big Mac	550 Cal.	0.00 0.00 meal
2 Quarter Pounder™ with cheese	900-1140 Cal.	0.00 0.00 meal
3 Double Quarter Pounder™ with cheese	1130-1370 Cal.	0.00 0.00 meal
4 2 Cheeseburgers	600 Cal.	0.00 0.00 meal
5 Ranch BLT	980-1220 Cal.	0.00 0.00 meal
6 Club	1000-1240 Cal.	0.00 0.00 meal
7 Classic	350-510 Cal.	0.00 0.00 meal
8 Southern Style Chicken	800-1040 Cal.	0.00 0.00 meal
9 Chicken Selects®	760-1000 Cal.	0.00 0.00 3 pc. meal
10 10 Pc. Chicken McNuggets®	470 Cal.	0.00 0.00 meal
11 Filet-O-Fish®	760-1000 Cal.	0.00 0.00 meal
12 Angus Deluxe	750 Cal.	0.00 0.00 meal
13 Angus Bacon & Cheese	1170-1410 Cal.	0.00 0.00 meal
14 Angus Mushroom & Swiss	1150-1390 Cal.	0.00 0.00 meal
salads		
Southwest	290-450 Cal.	0.00
Bacon Ranch	230-390 Cal.	0.00
Caesar	190-350 Cal.	0.00

**grilled/crispy chicken**

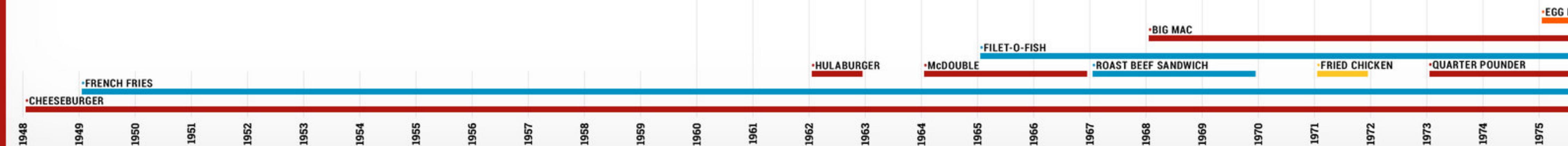


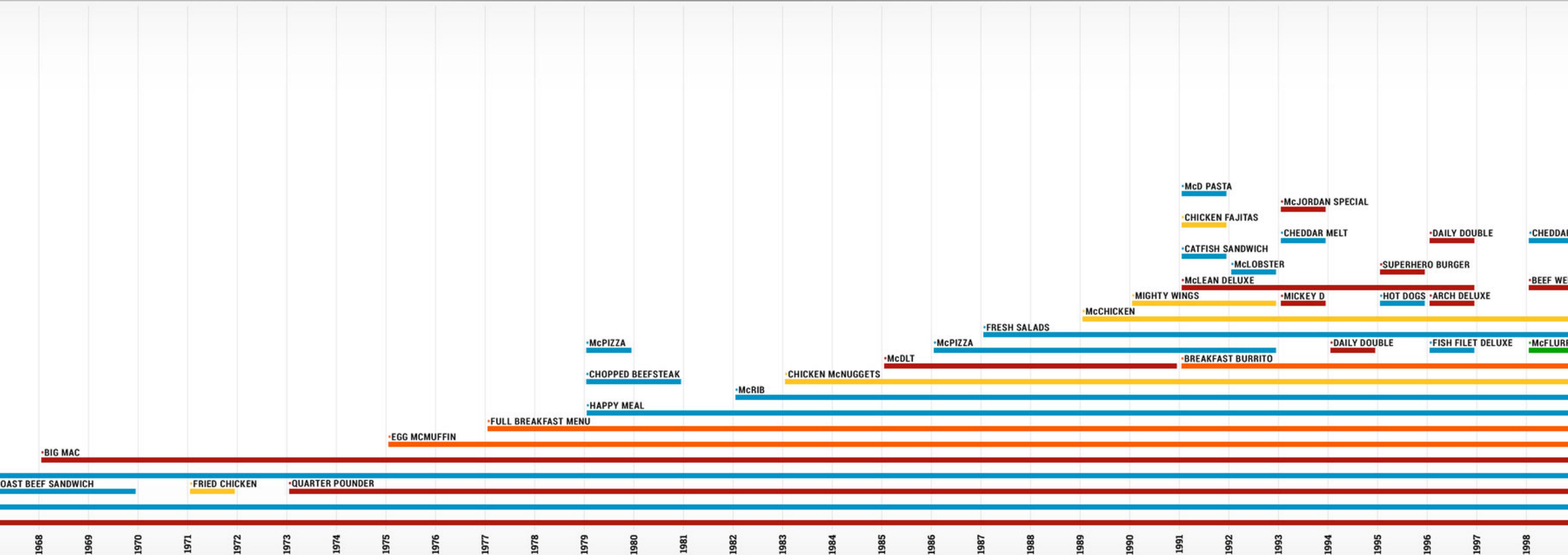


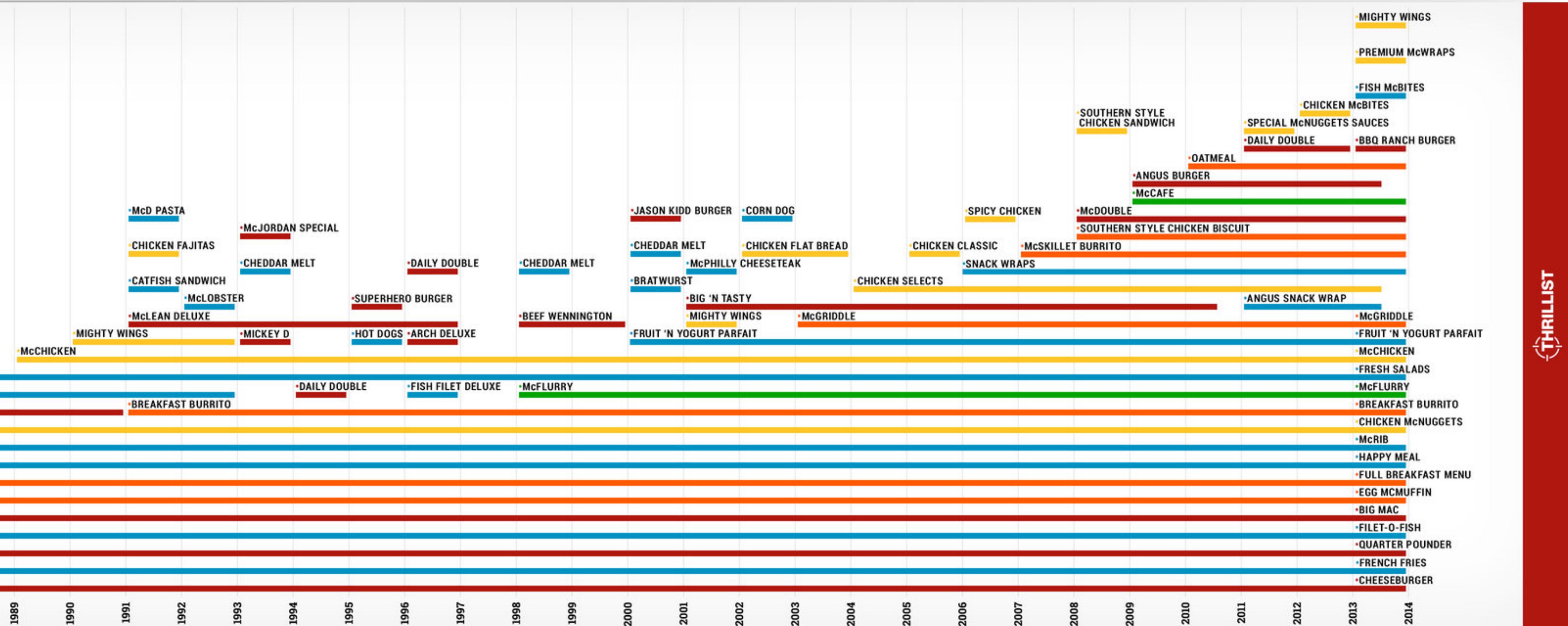
# THE HISTORICAL TIMELINE OF THE McDONALD'S MENU

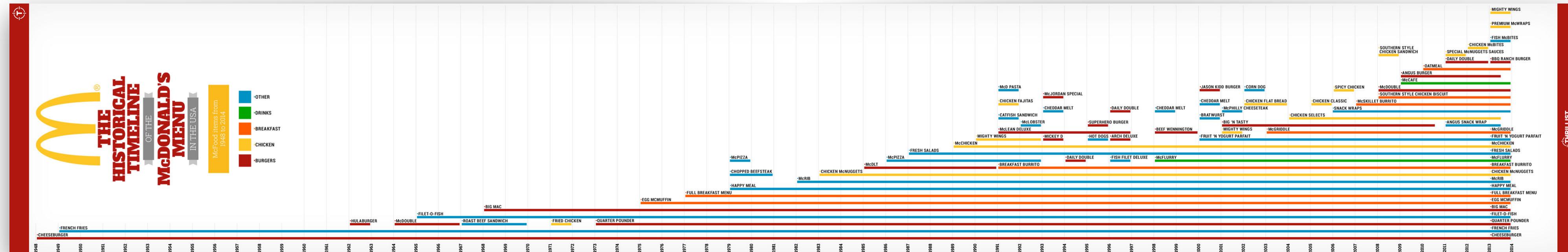
IN THE USA  
McFood items from  
1948 to 2014

- OTHER
- DRINKS
- BREAKFAST
- CHICKEN
- BURGERS









Essa é uma **técnica de visualização** chamada *timeline* ou linha do tempo

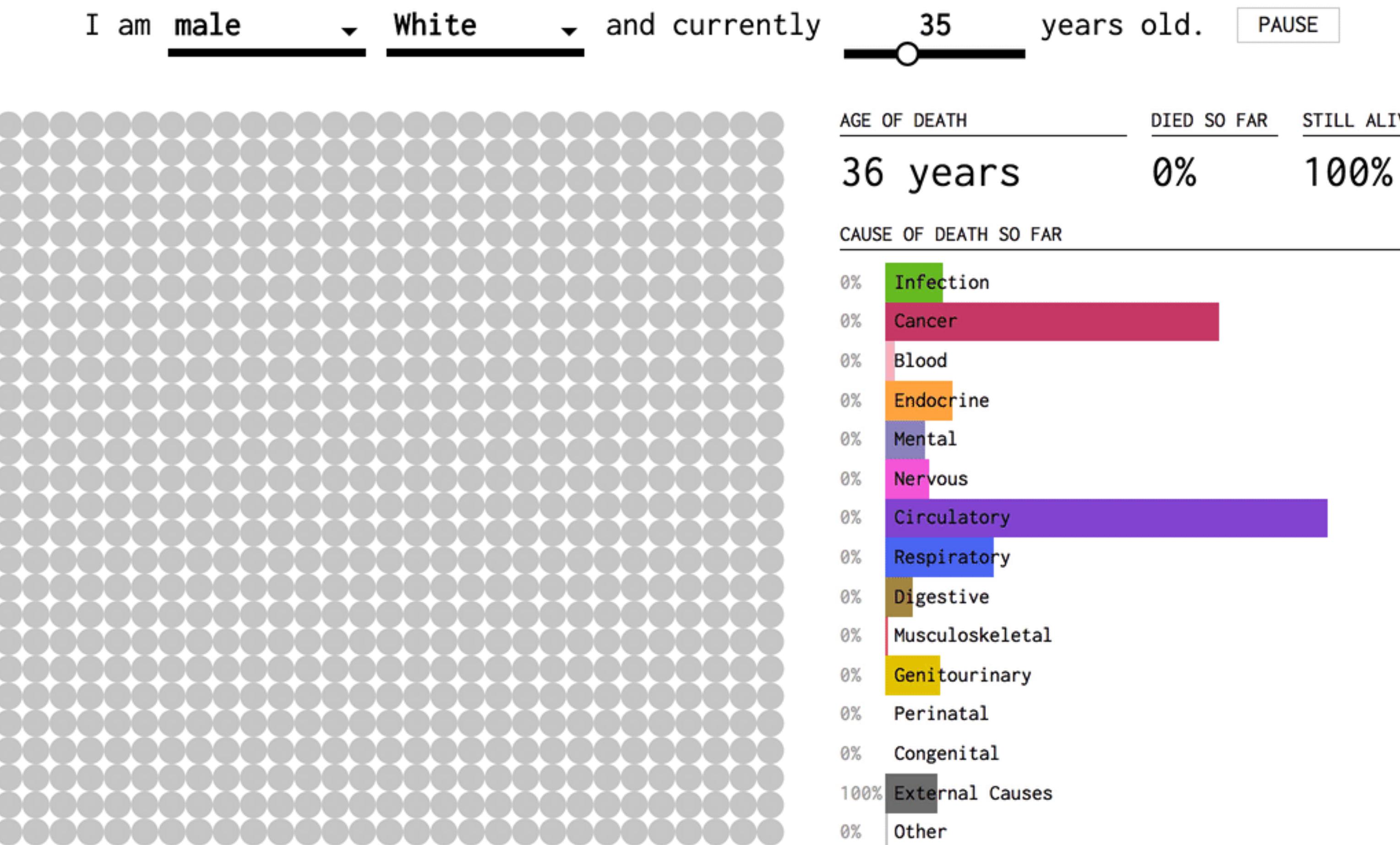
Ela ilustra um **relacionamento quantitativo** intrínseco aos dados que é denominado **série temporal**

A **análise quantitativa** envolve relacionamentos entre valores que, por sua vez, envolve a busca por **padrões visuais** nos dados e o uso de **técnicas** de análise particulares

Os principais tipos de relacionamentos são:

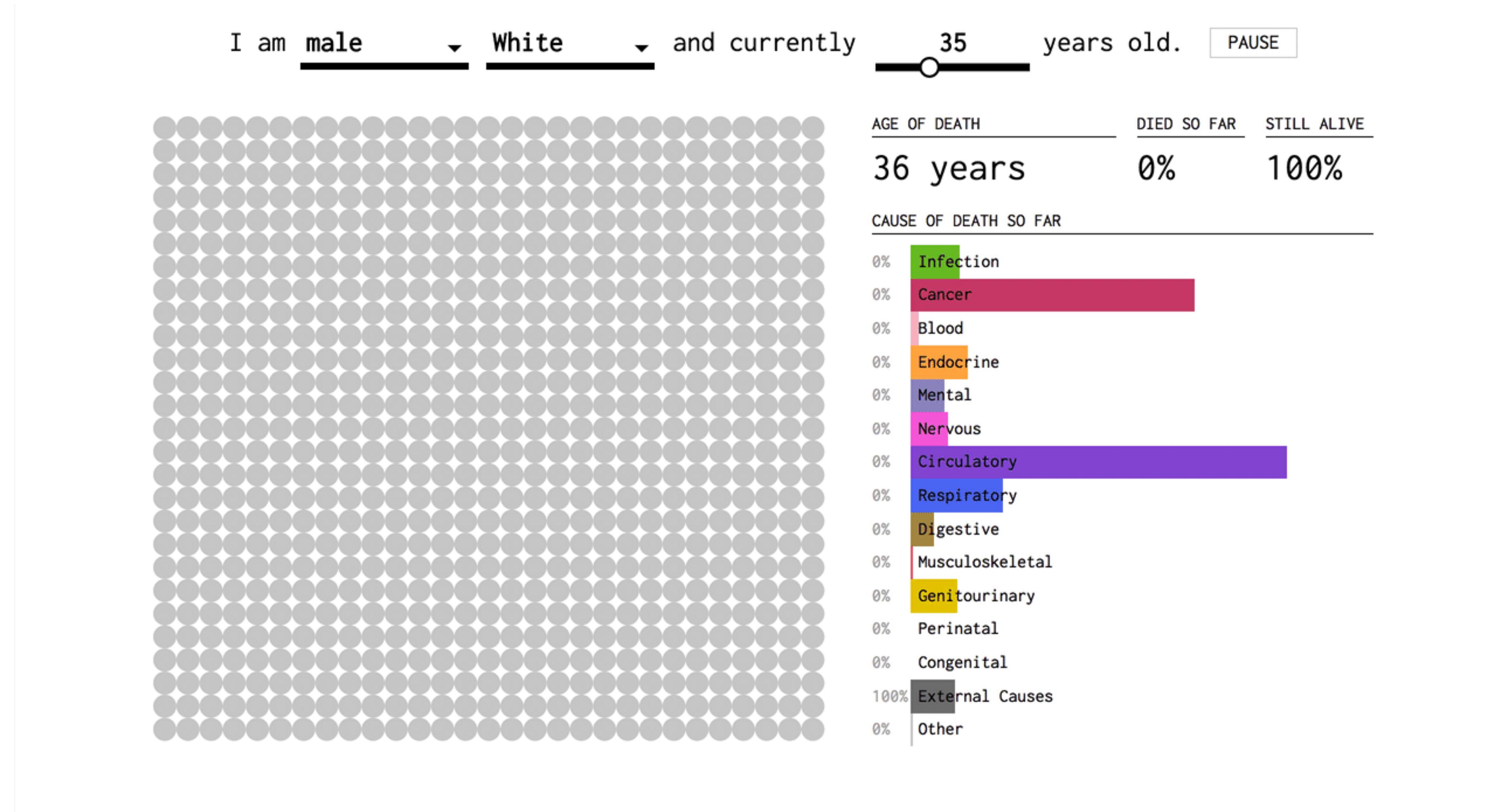
- Séries temporais
- Ranking
- Parte-todo
- Desvio
- Distribuição
- Correlação
- Multivariados

Essa visualização é uma composição entre um diagrama e um **histograma** que representa o relacionamento quantitativo **parte-todo**

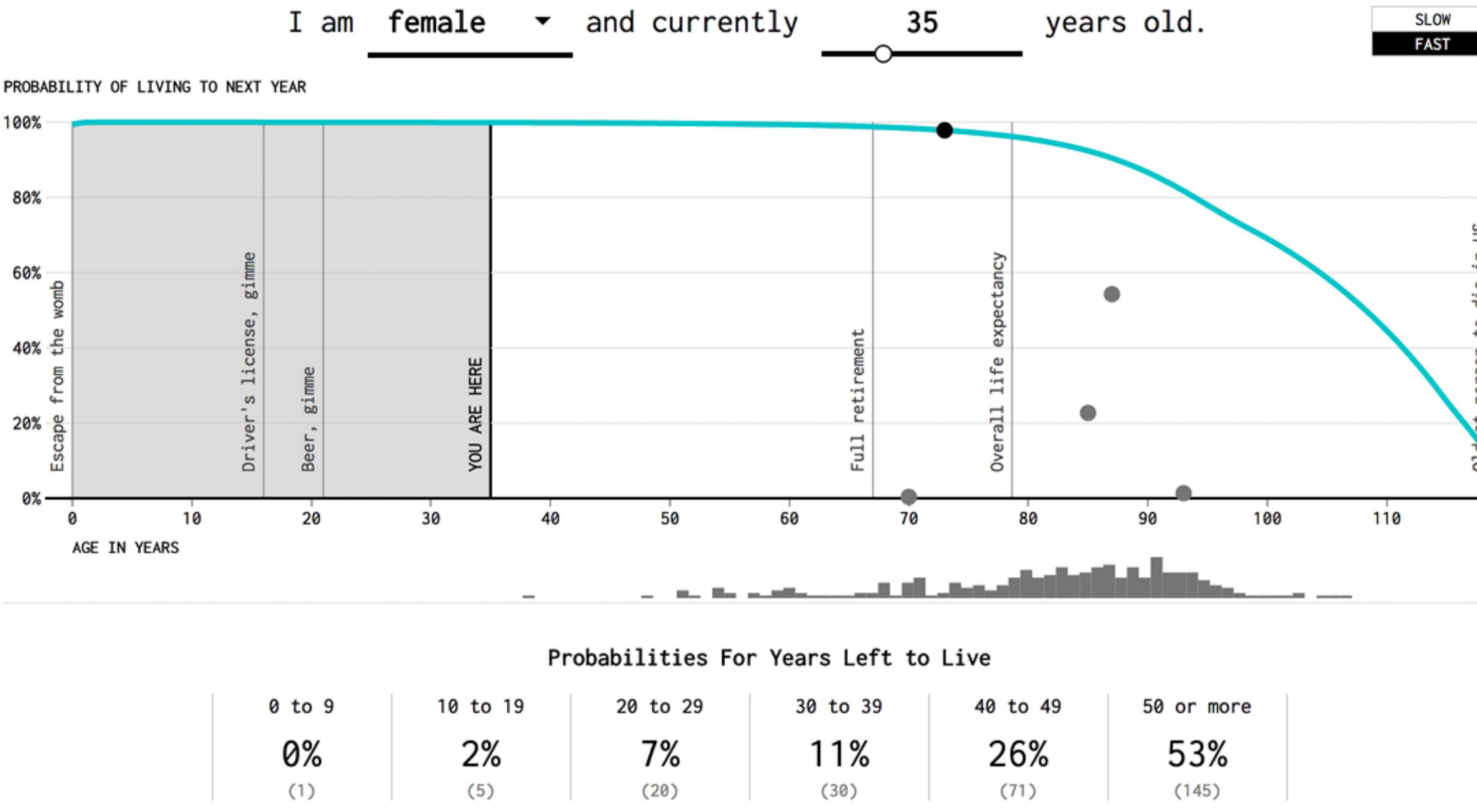


Nathan Yau

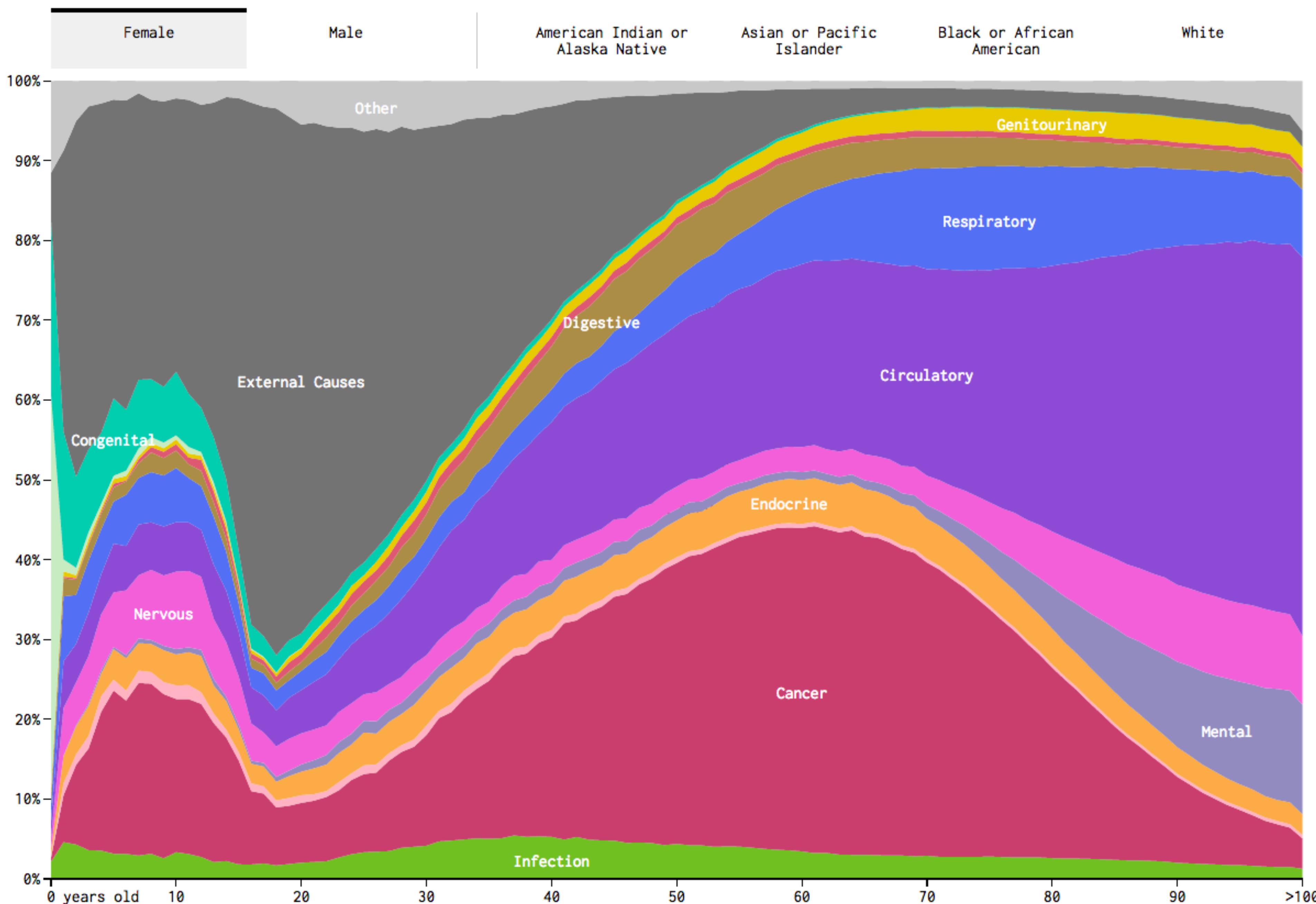
O que você é capaz de **aprender ou descobrir** através dessa visualização?



Essa visualização é um **gráfico de linhas** que representa o relacionamento quantitativo **distribuição**



E com essa visualização, o que podemos aprender?



Essa visualização ilustra os relacionamentos quantitativos **série temporal** e **parte-todo** e a técnica é o **gráfico de áreas empilhadas**

Note que esse tipo de visualização traz mais **densidade de informação**, tem maior **potencial analítico** e nos permite **aprender mais**

“

*“Over history, visual abstractions have been developed to aid thinking...*

*What information visualization is really about is external cognition, that is, how resources outside the mind can be used to boost the cognitive capabilities of the mind.”*

-Stuart Card

in Visualization: Perception for Design

Visualizações só serão **efetivas** se operarem sobre princípios que **respeitem como a percepção visual e a cognição funcionam.**

# A Field Guide to Red and Blue America

Randy Yeip,  
Stuart A. Thompson e  
Will Welch

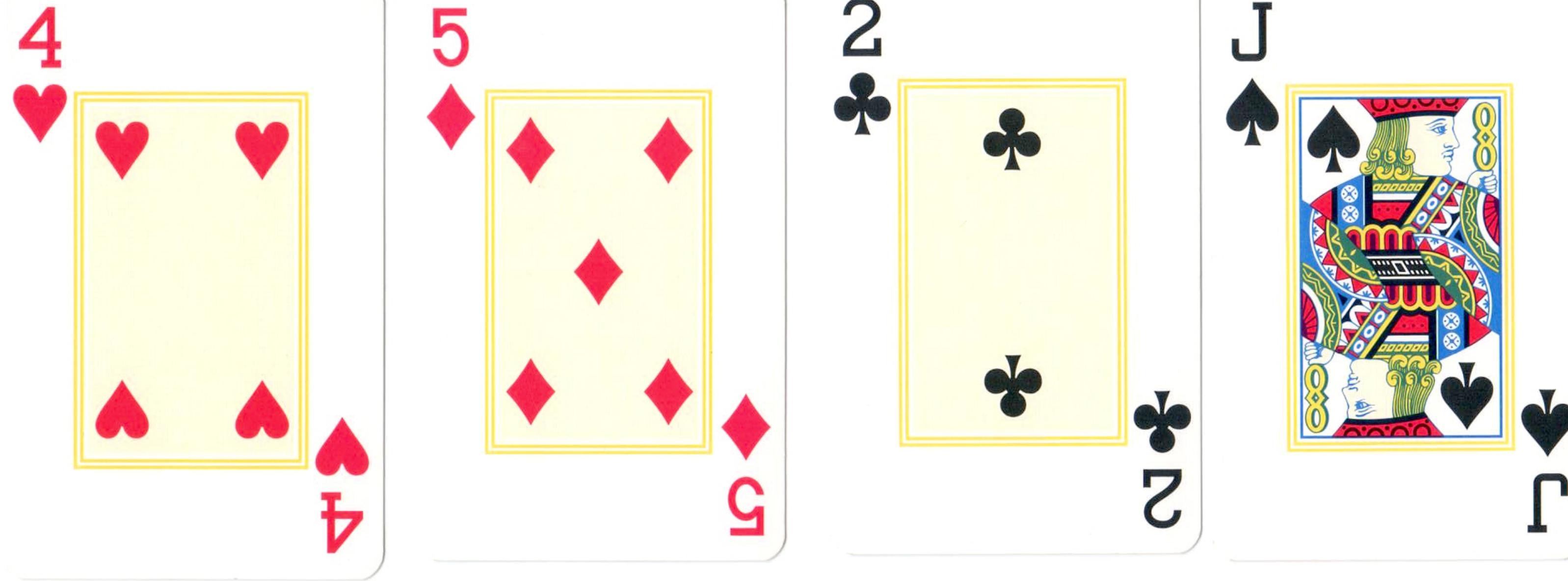


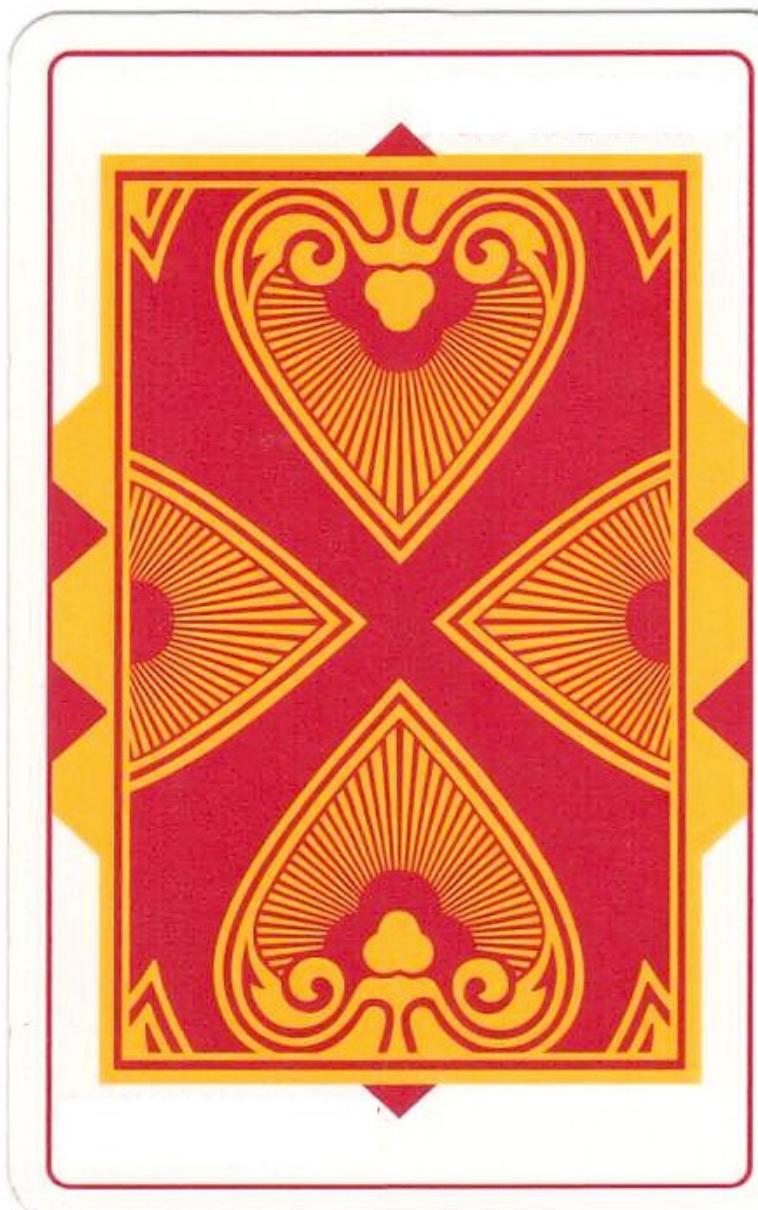
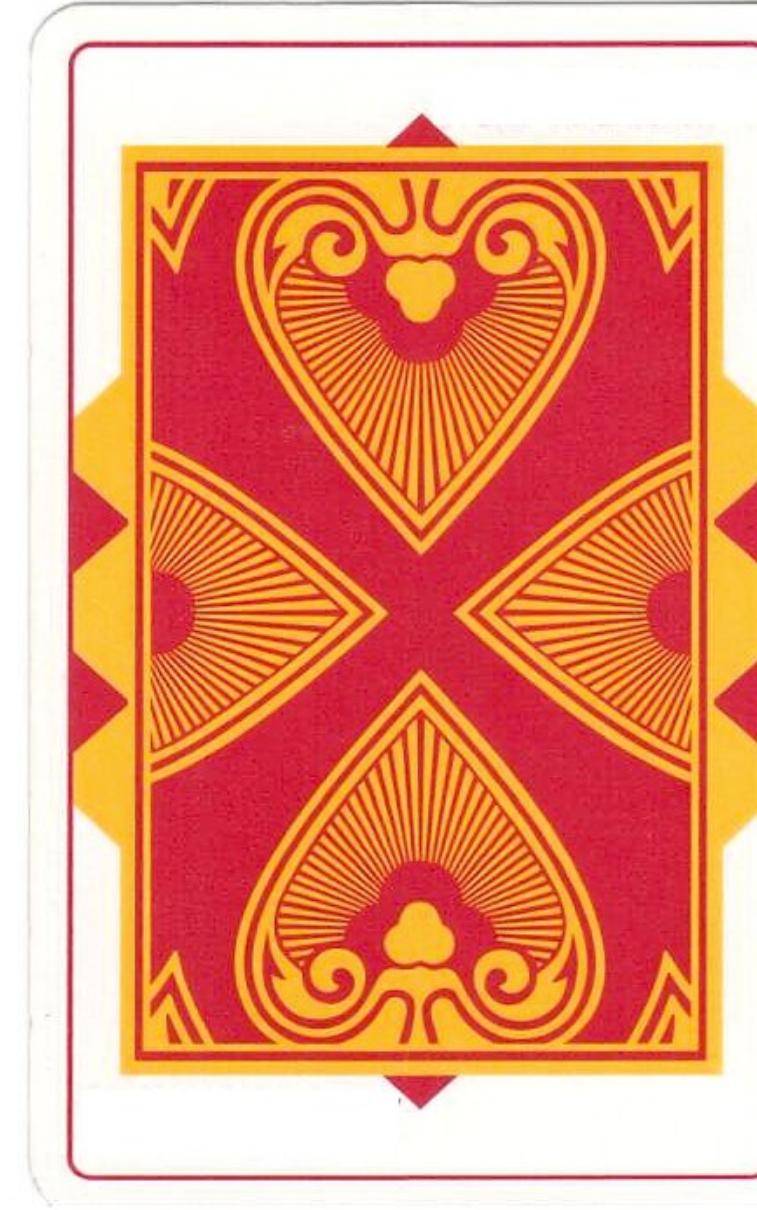
Esse gráfico usa o padrão visual denominado **pequenos múltiplos** (Edward Tufte) para compor um diagrama composto por **histogramas** que retratam **séries temporais**

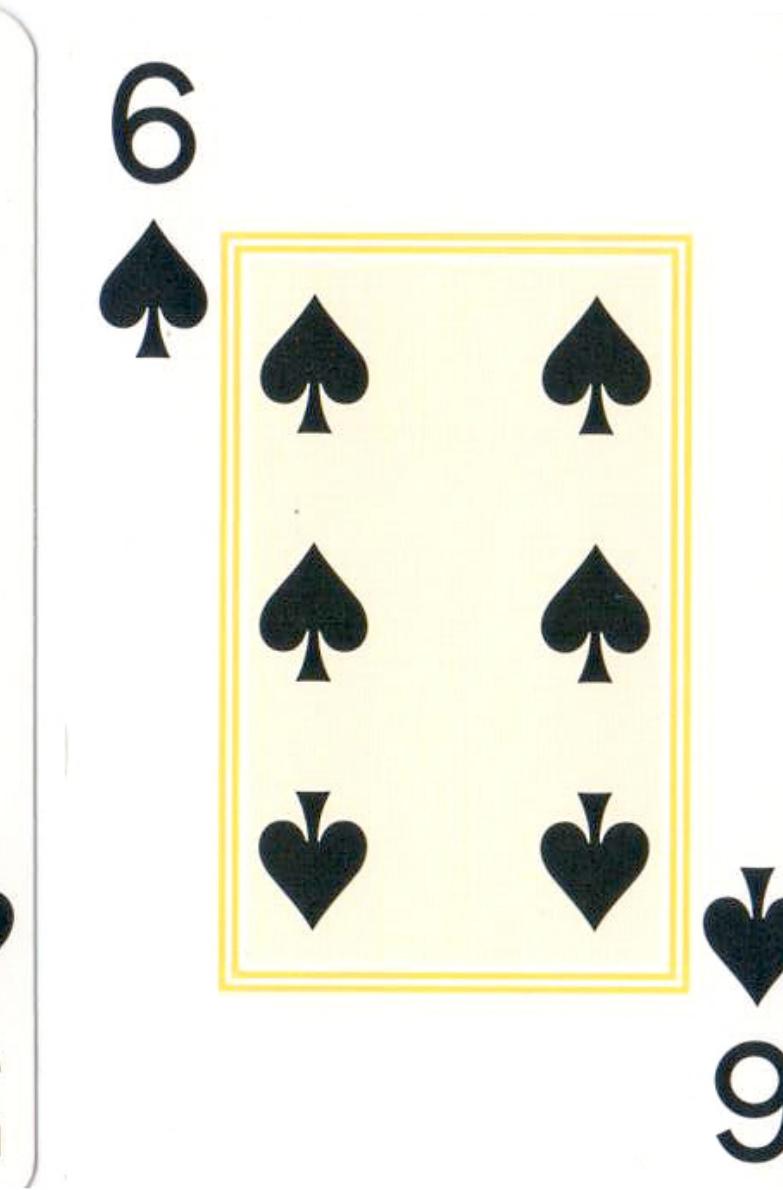
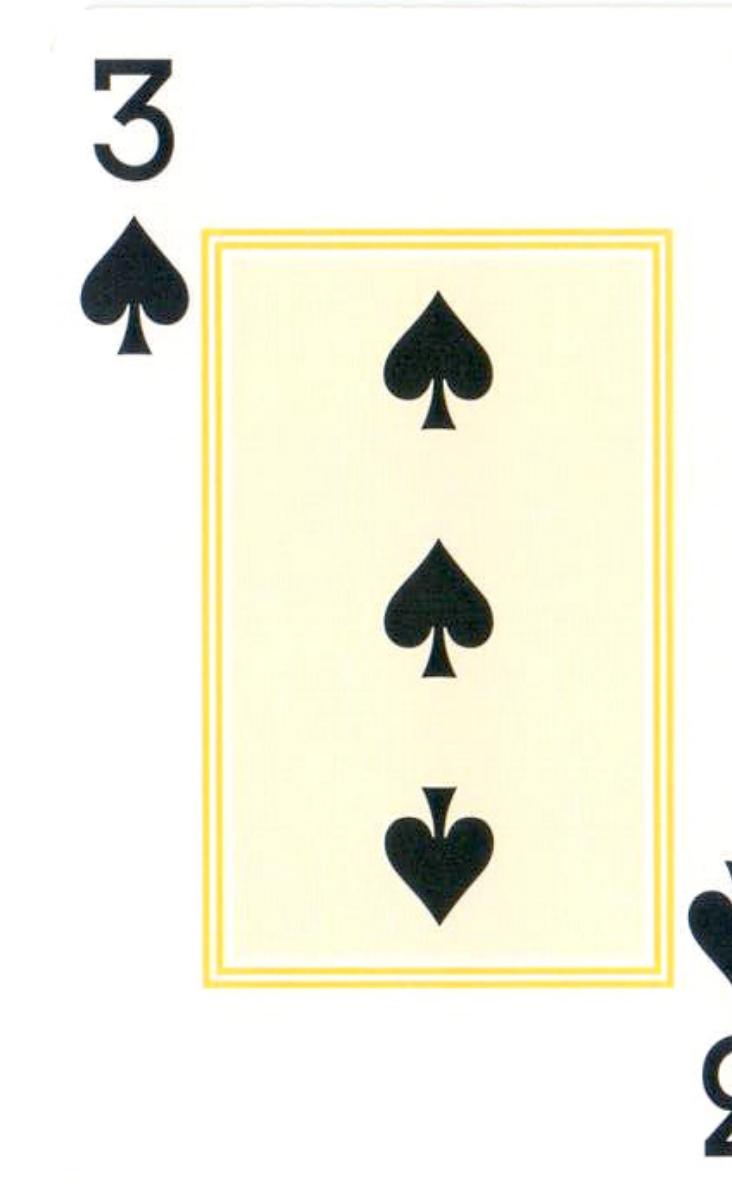
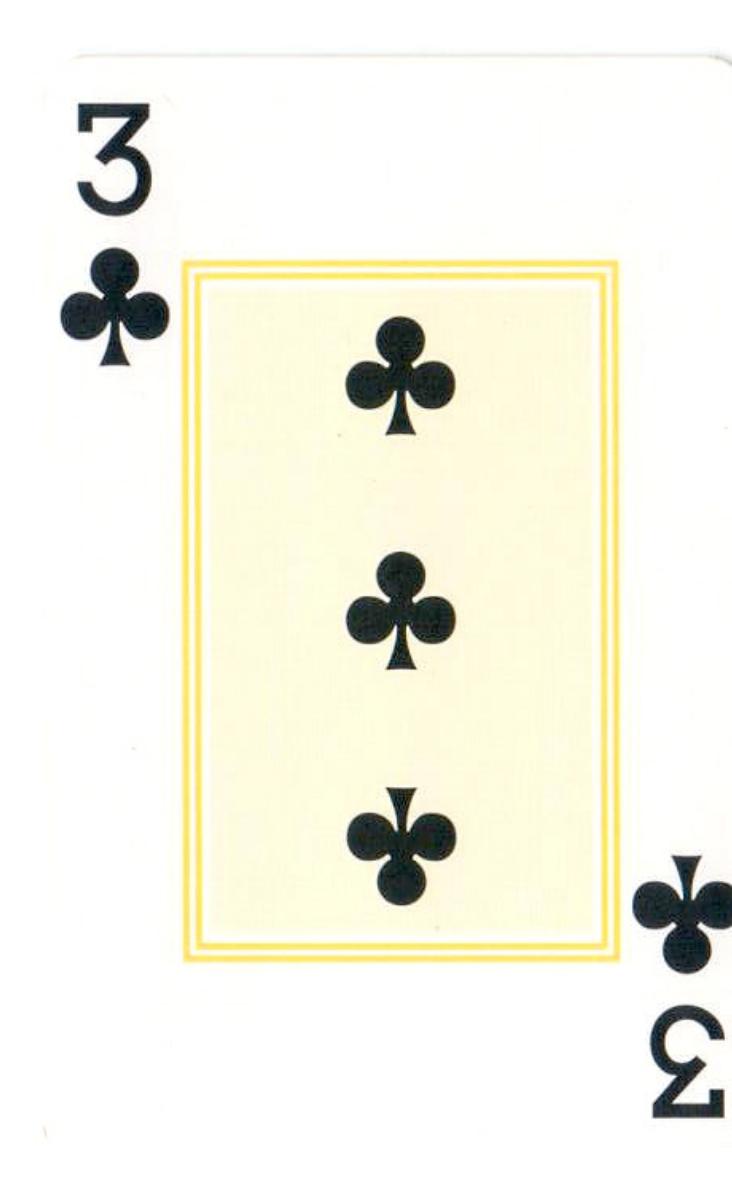
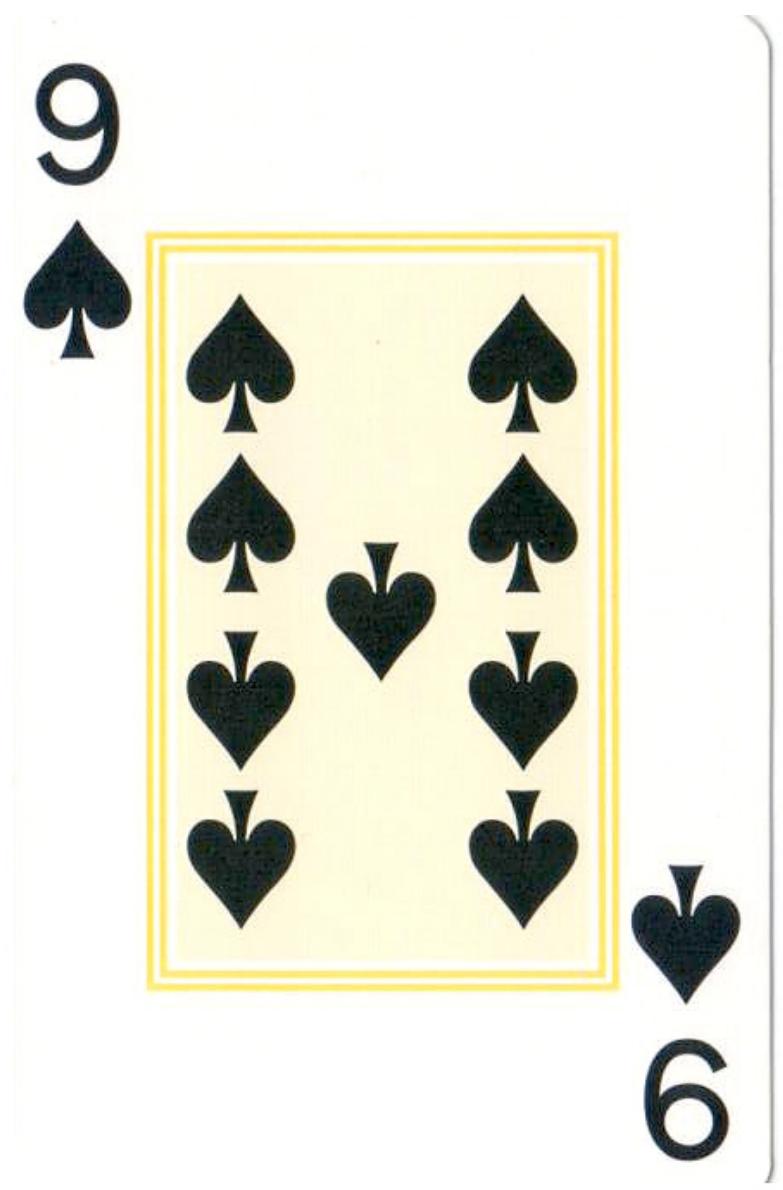
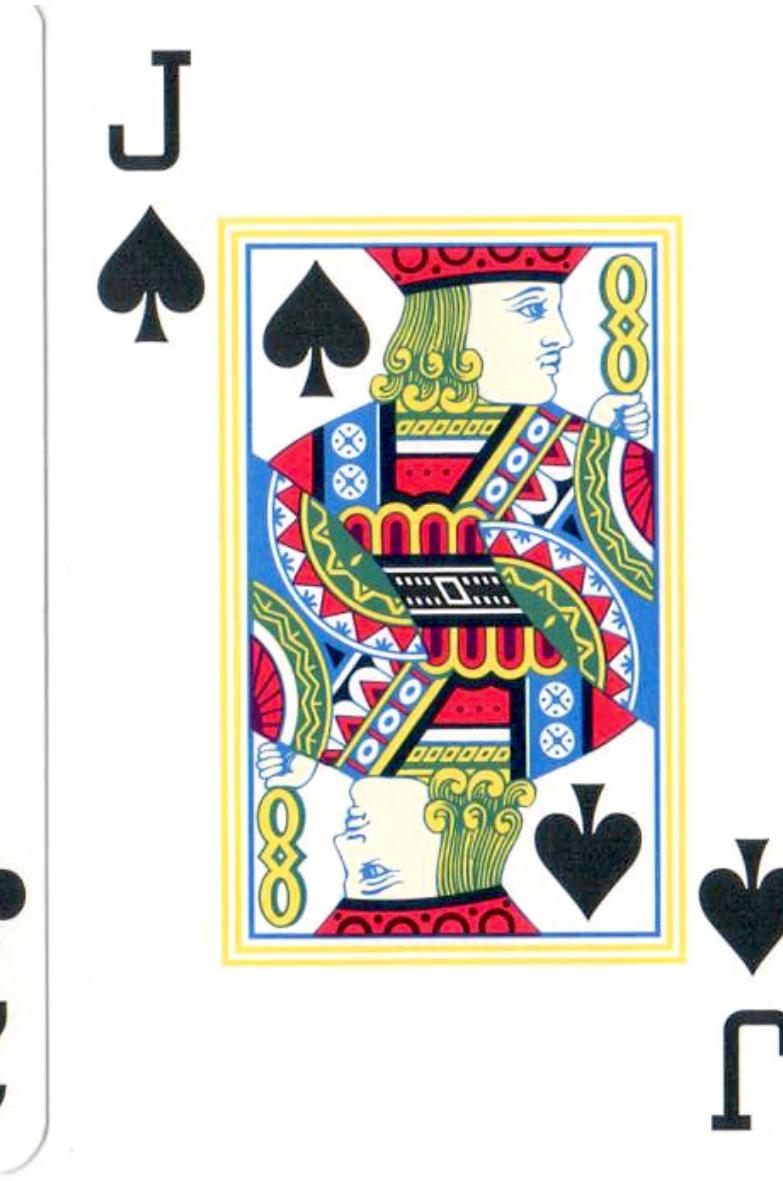
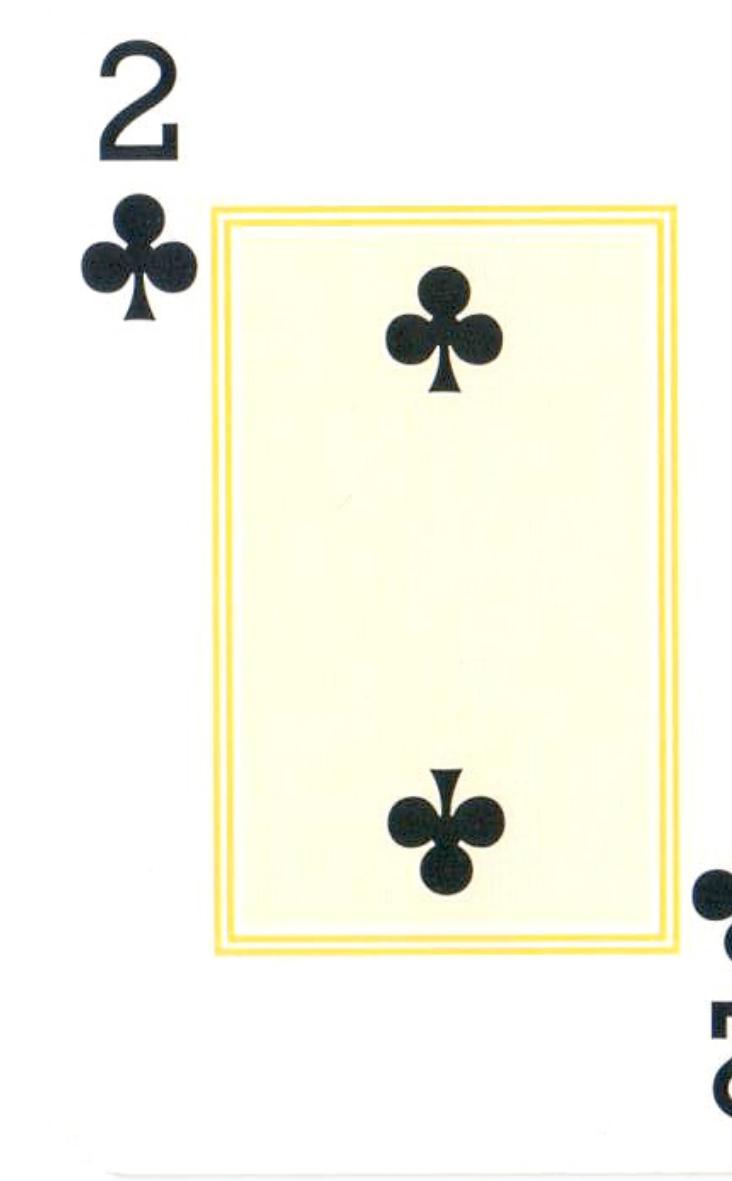
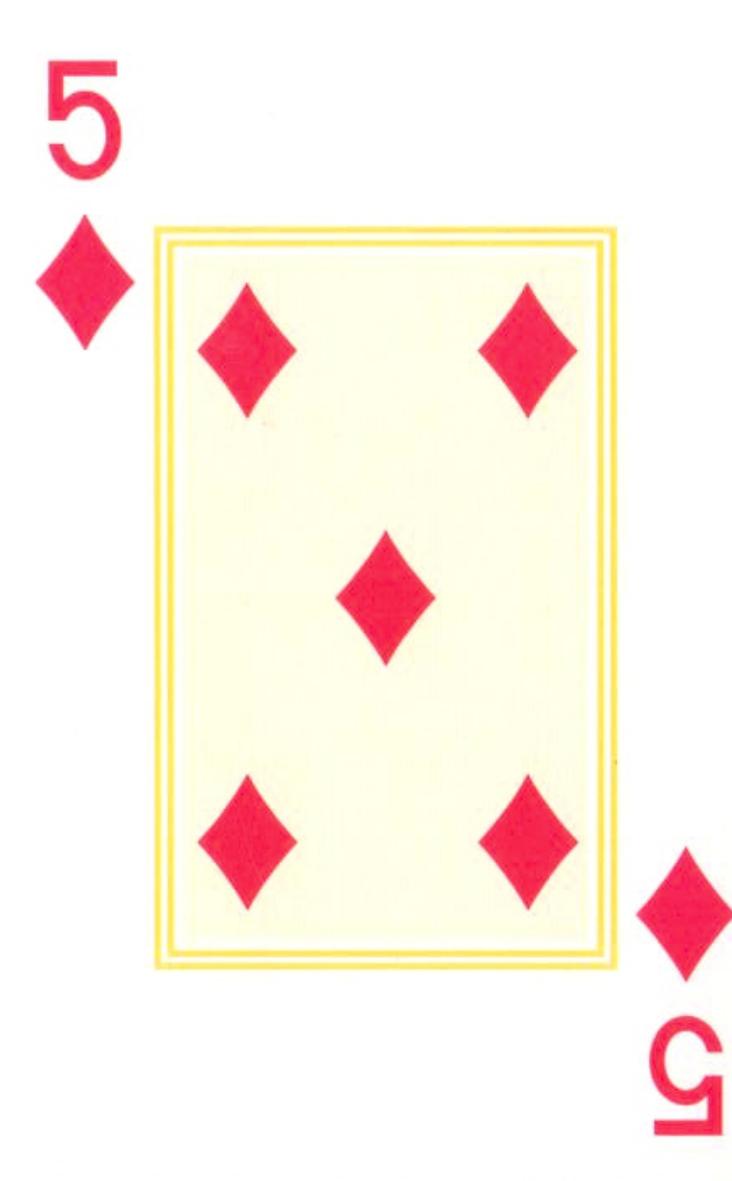
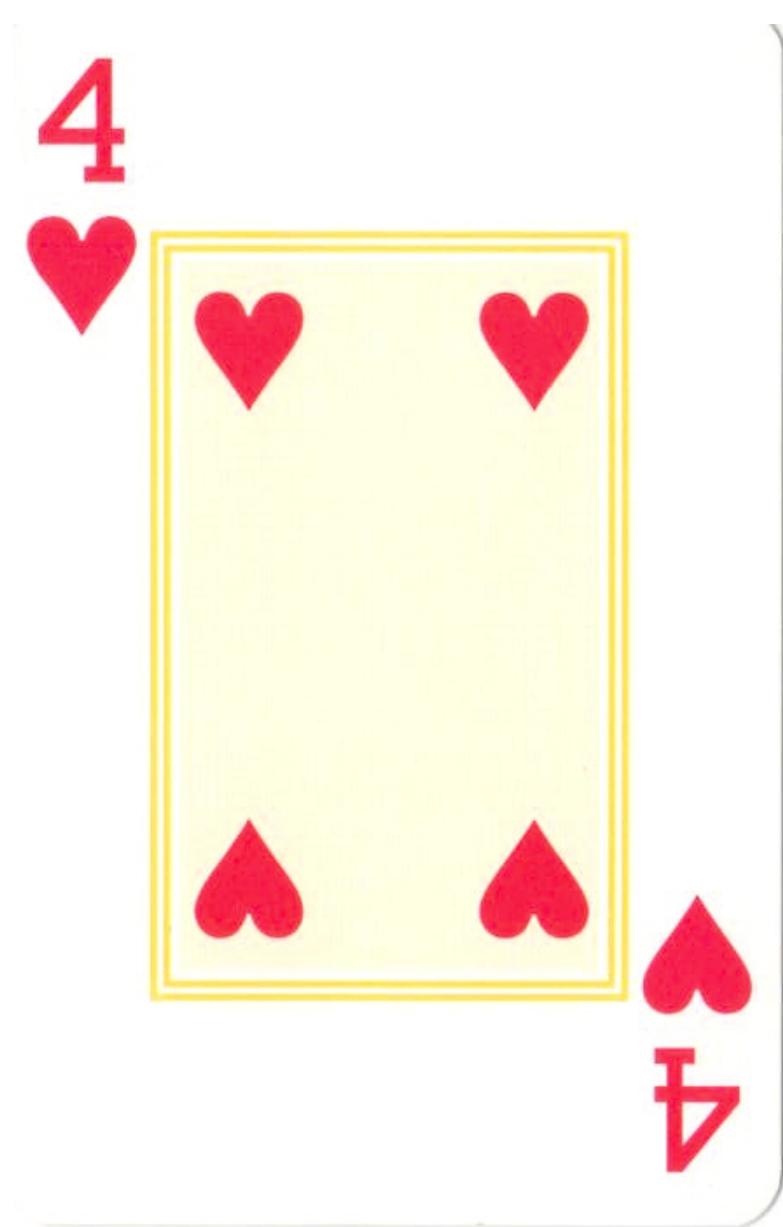
O diagrama é organizado usando uma posição relacionada à posição geográfica

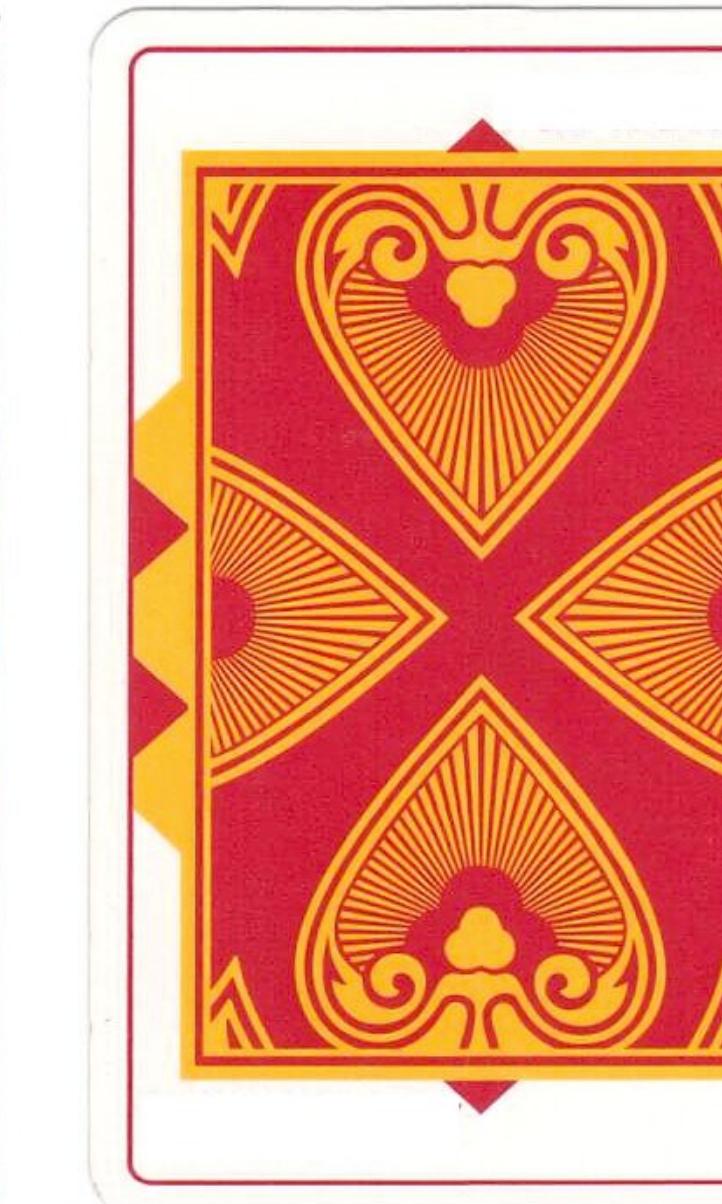
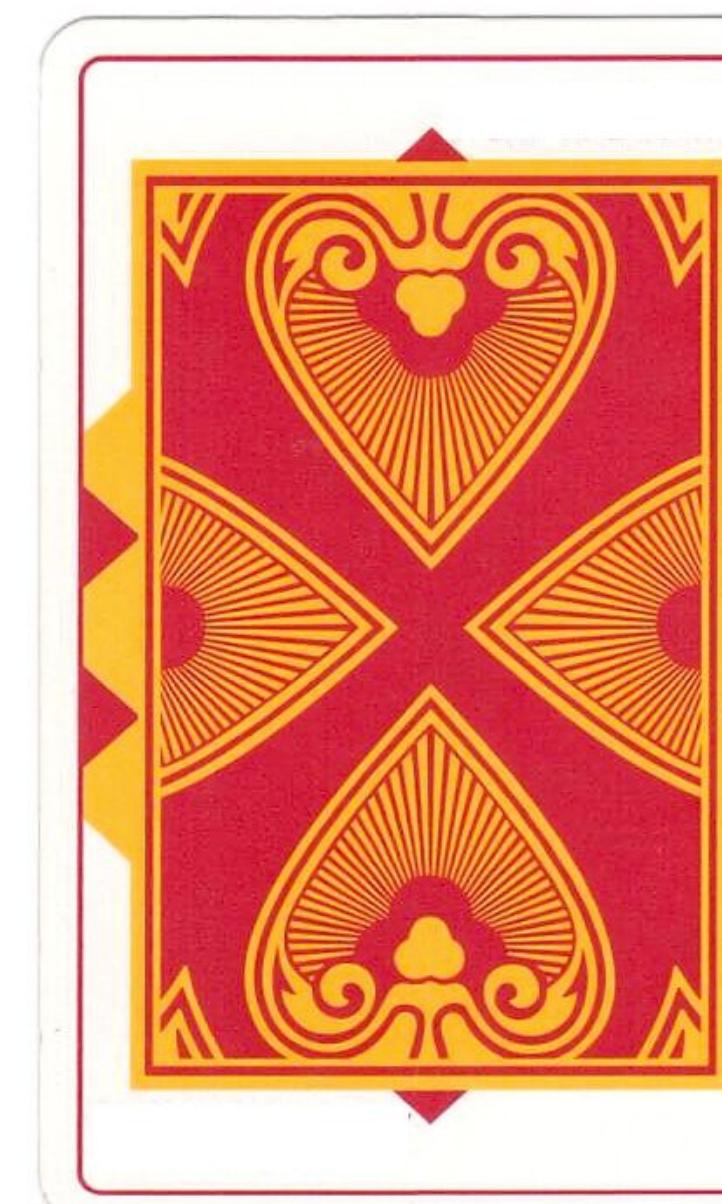
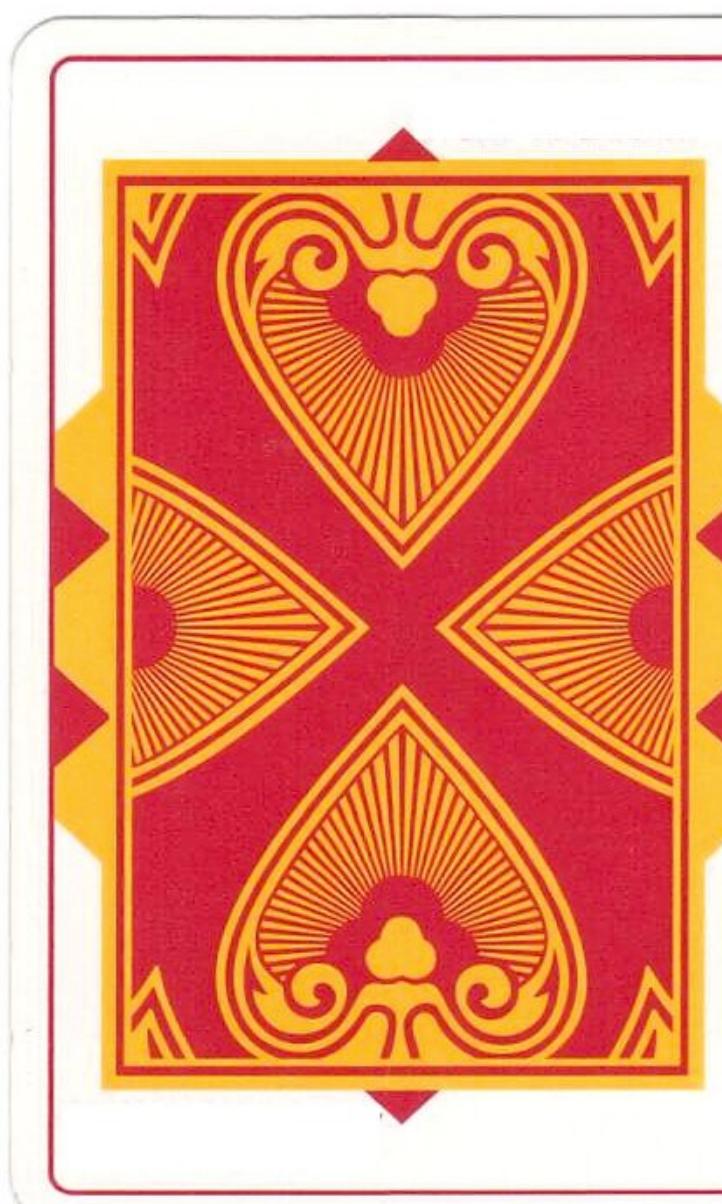
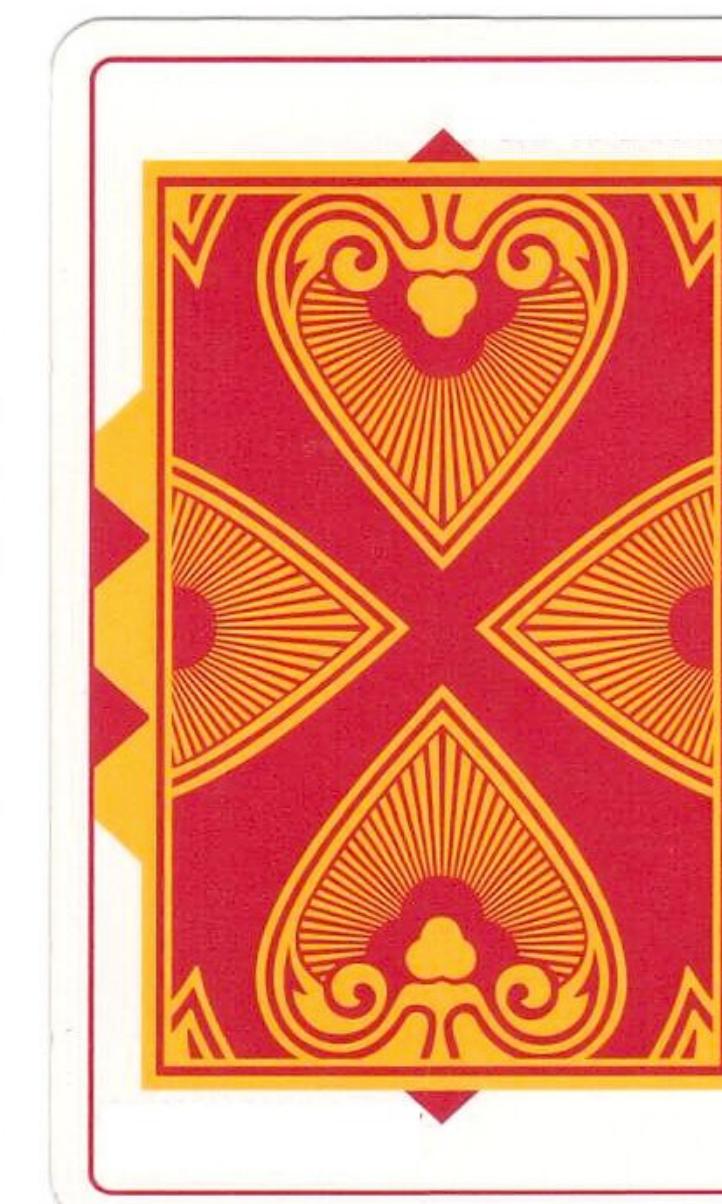
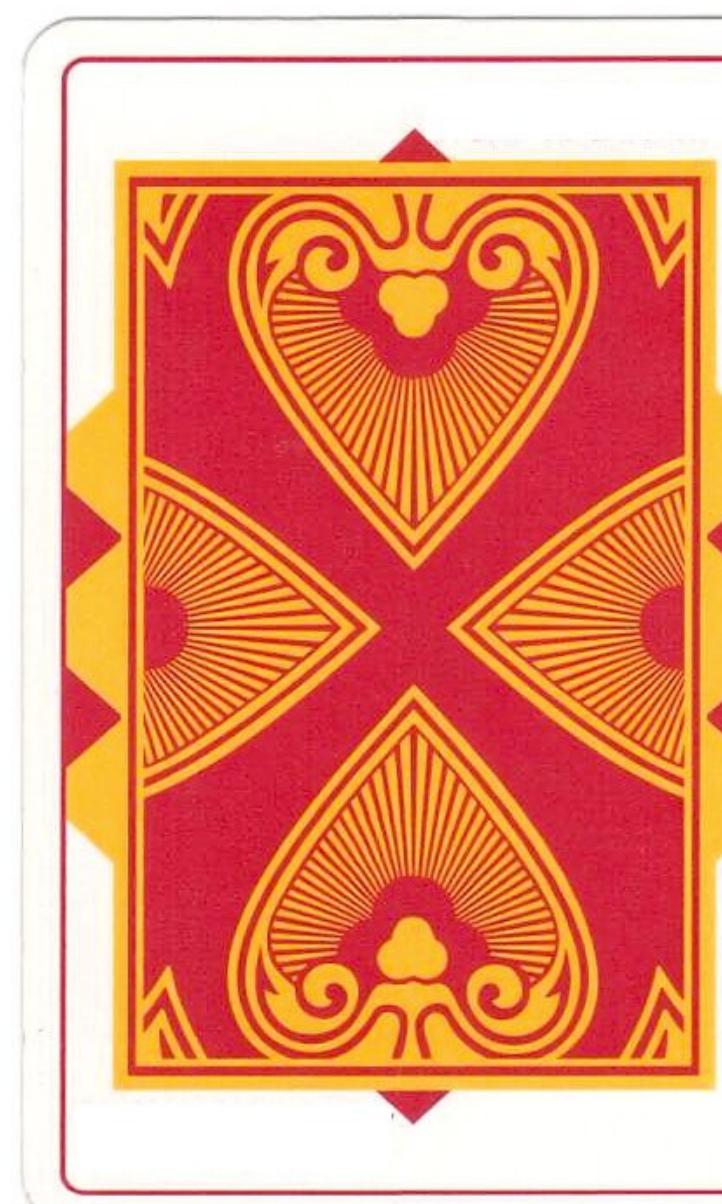
**O que** essa organização nos permite aprender?

**Por que** essa visualização potencializa nossas capacidades de cognição?









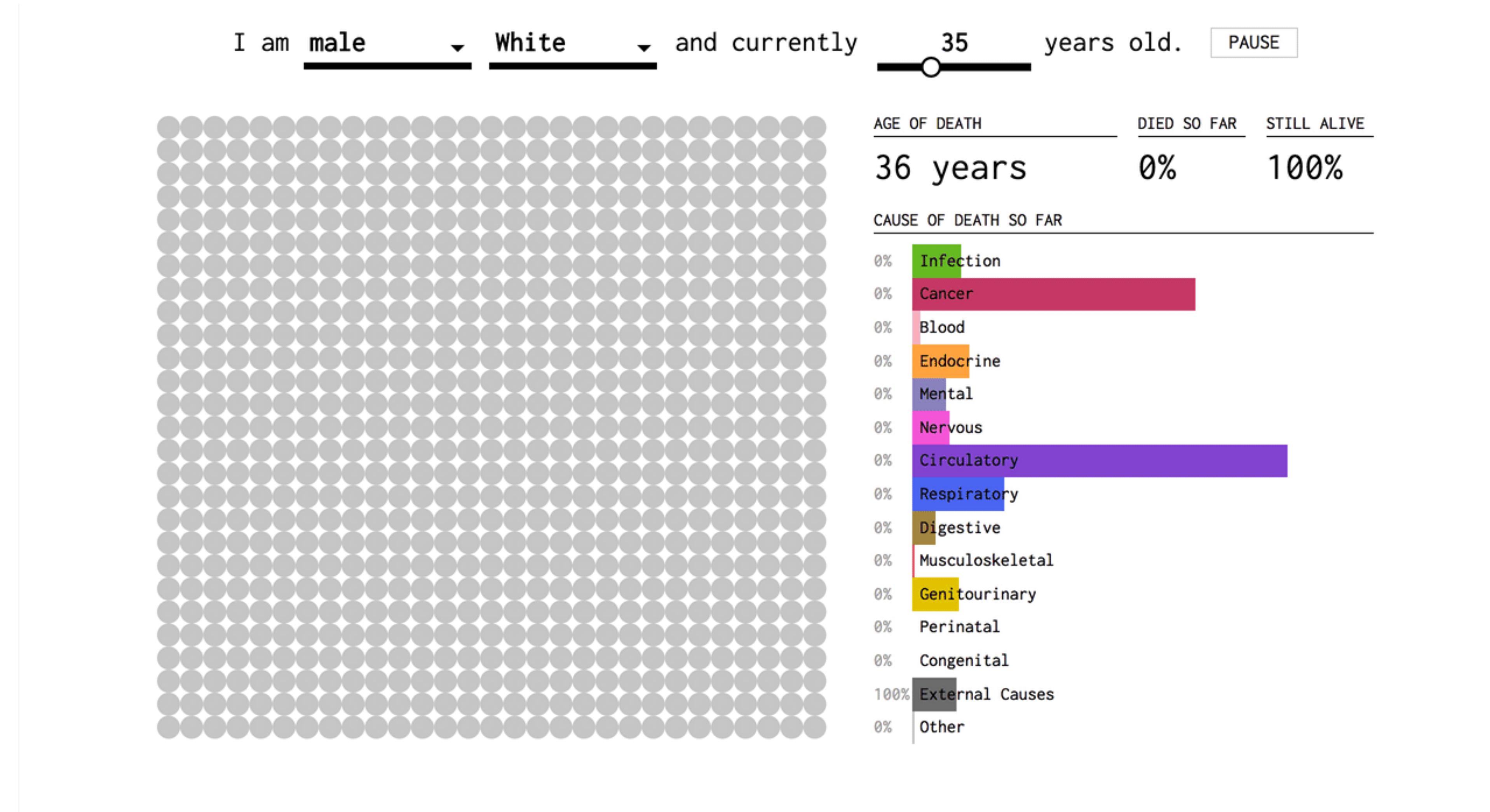
A **memória** tem **papel fundamental** na cognição,  
mas nossa memória de trabalho é **extremamente limitada**

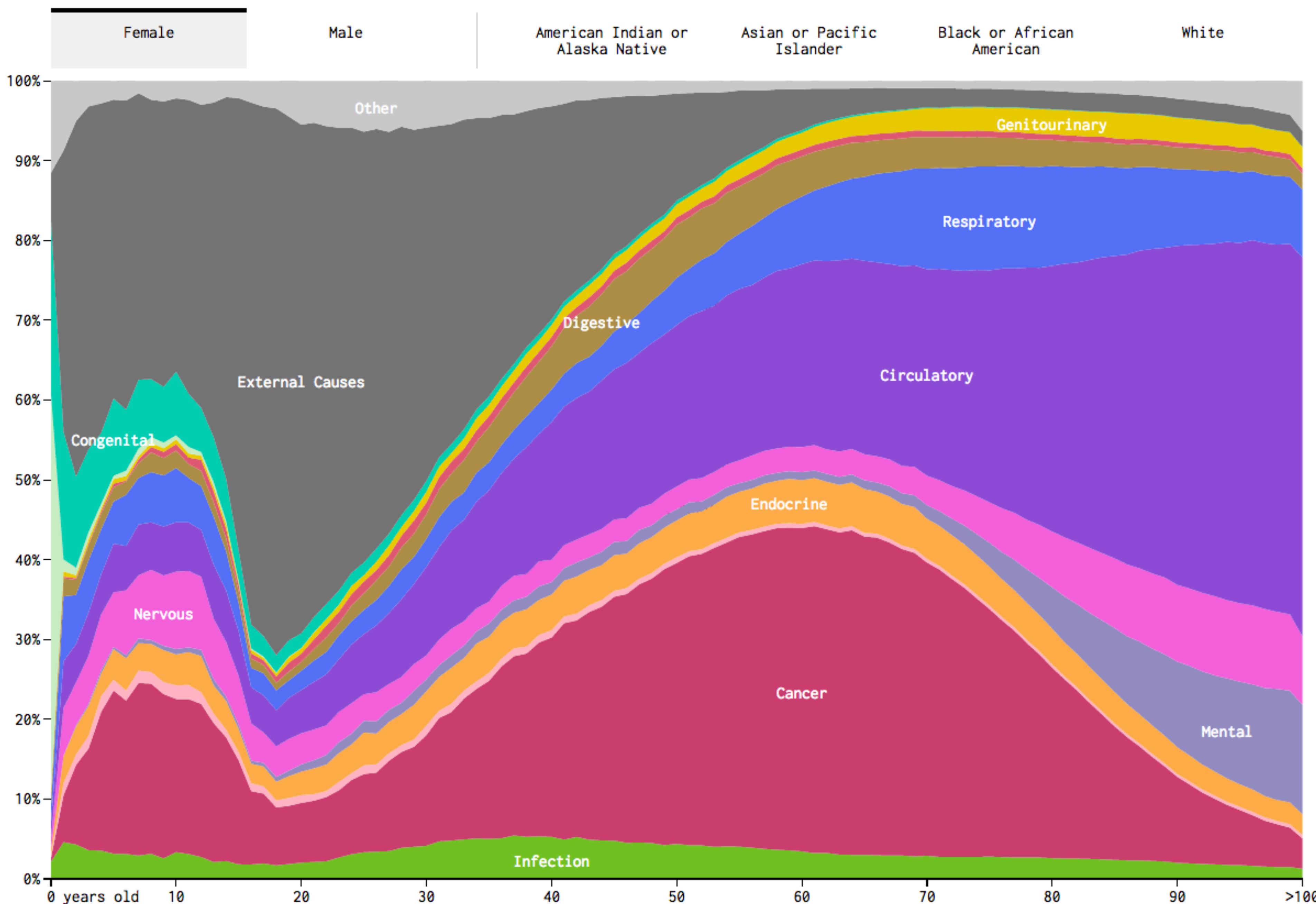
# A Field Guide to Red and Blue America

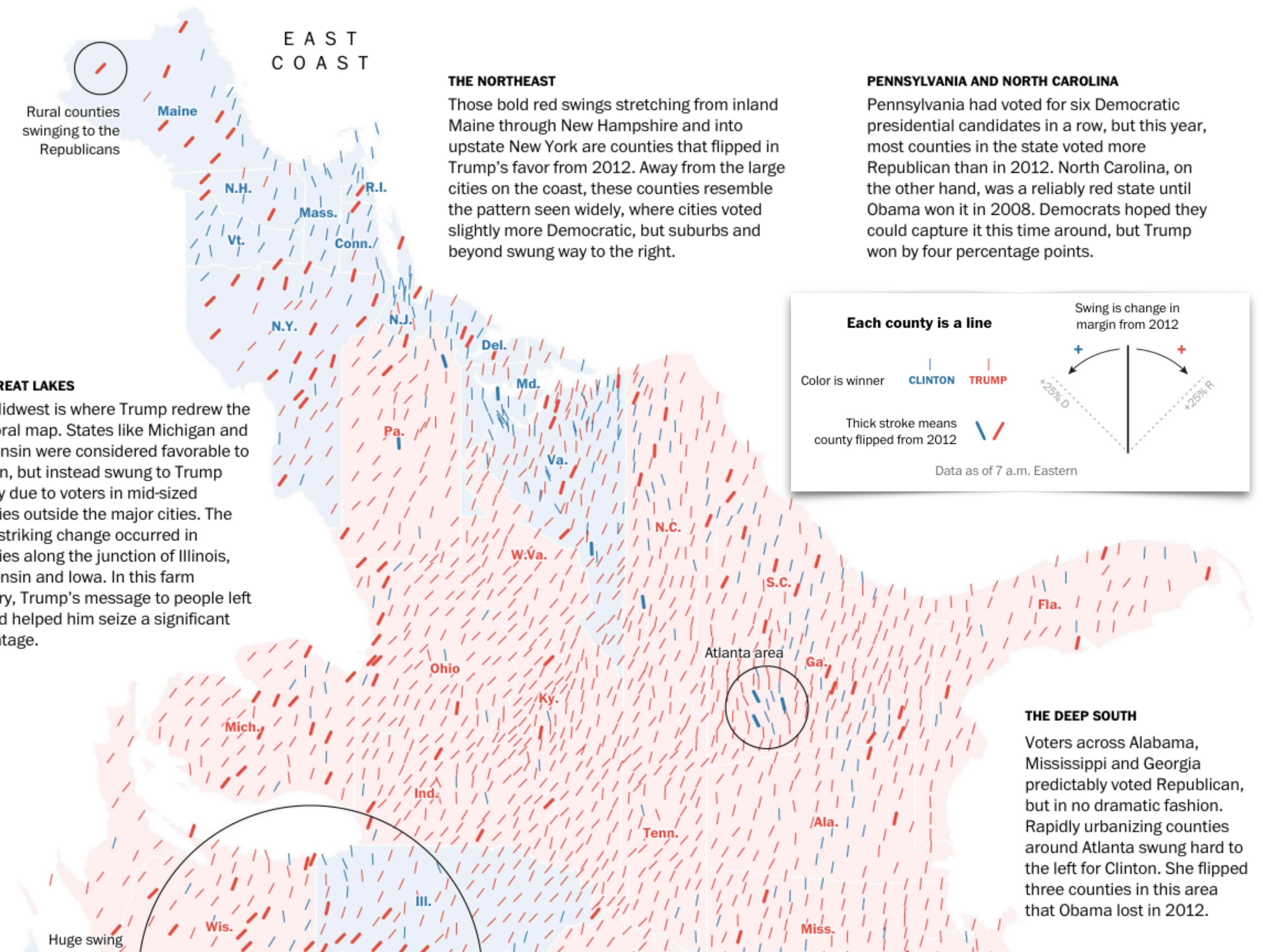
Randy Yeip,  
Stuart A. Thompson e  
Will Welch



É devido à limitação da nossa memória que animações não são tão efetivas no processo analítico







# Princípio da densidade de dados de Tufte

## Data Density in Graphical Practice

The numbers that go into a graphic can be organized into a data matrix of observations by variables. Taking into account the size of the graphic in relation to the amount of data displayed yields the *data density*:

$$\text{data density of a graphic} = \frac{\text{number of entries in data matrix}}{\text{area of data graphic}}$$

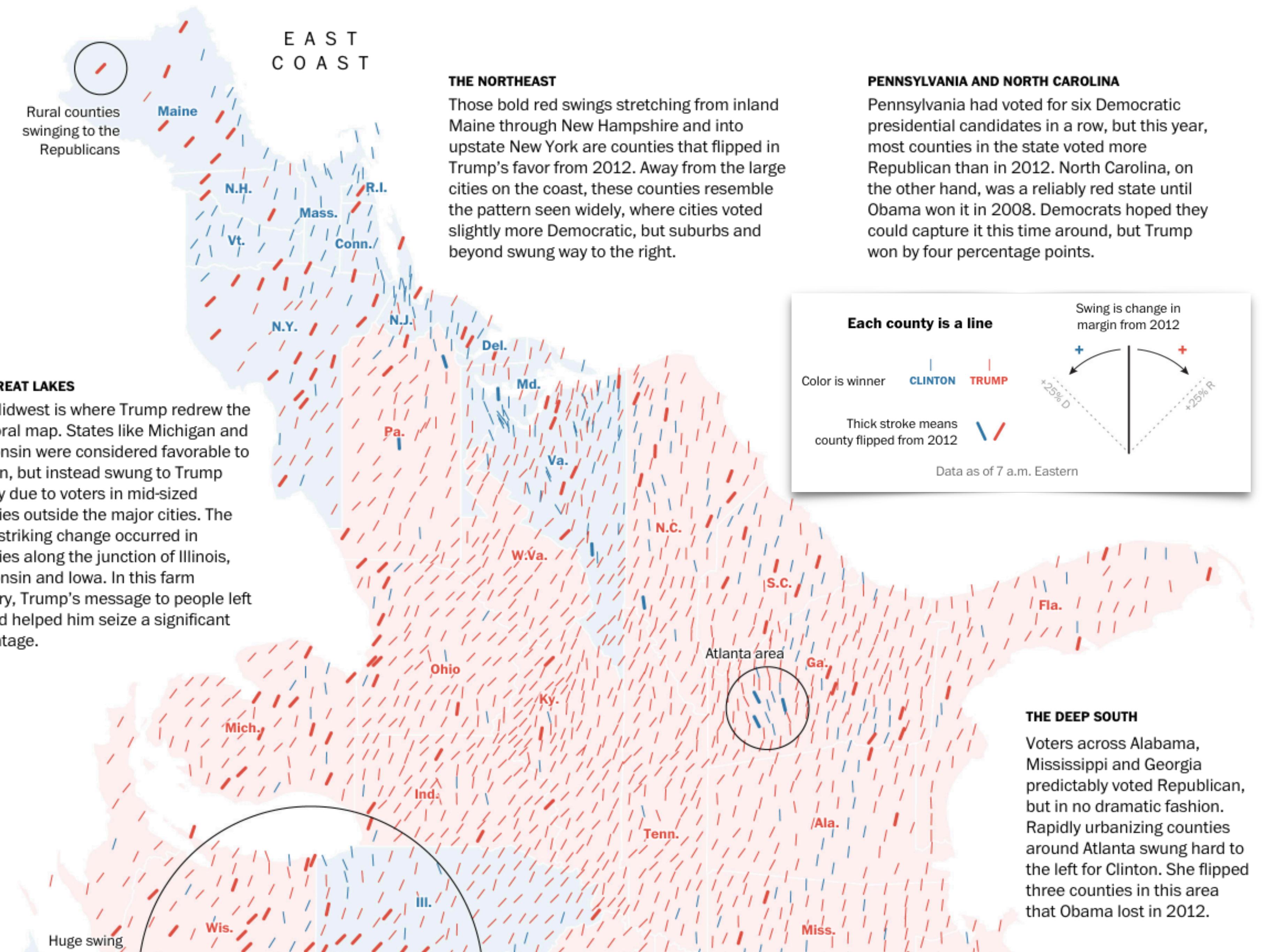
Data matrices and data densities vary enormously in practice. At one extreme, this overwrought display (originally printed in five colors) presents a data matrix of four entries, the names and the numbers for the two bars on the right. The left bar is merely the total of the other two. The graph covers 26.5 square inches (171 square centimeters), resulting in a data density of .15 numbers per square inch (.02 numbers per square centimeter), which is thin indeed.

Edward Tufte  
Visual display of quantitative information

Data-rich designs give a context and credibility to statistical evidence. Low-information designs are suspect: what is left out, what is hidden, why are we shown so little? High-density graphics help us to compare parts of the data by displaying much information within the view of the eye: we look at one page at a time and the more on the page, the more effective and comparative our eye can be.<sup>5</sup> The principle, then, is:

Maximize data density and the size of the data matrix, within reason.

High-information graphics must be designed with special care. As the volume of data increases, data measures must shrink (smaller dots for scatters, thinner lines for busy time-series). The clutter of

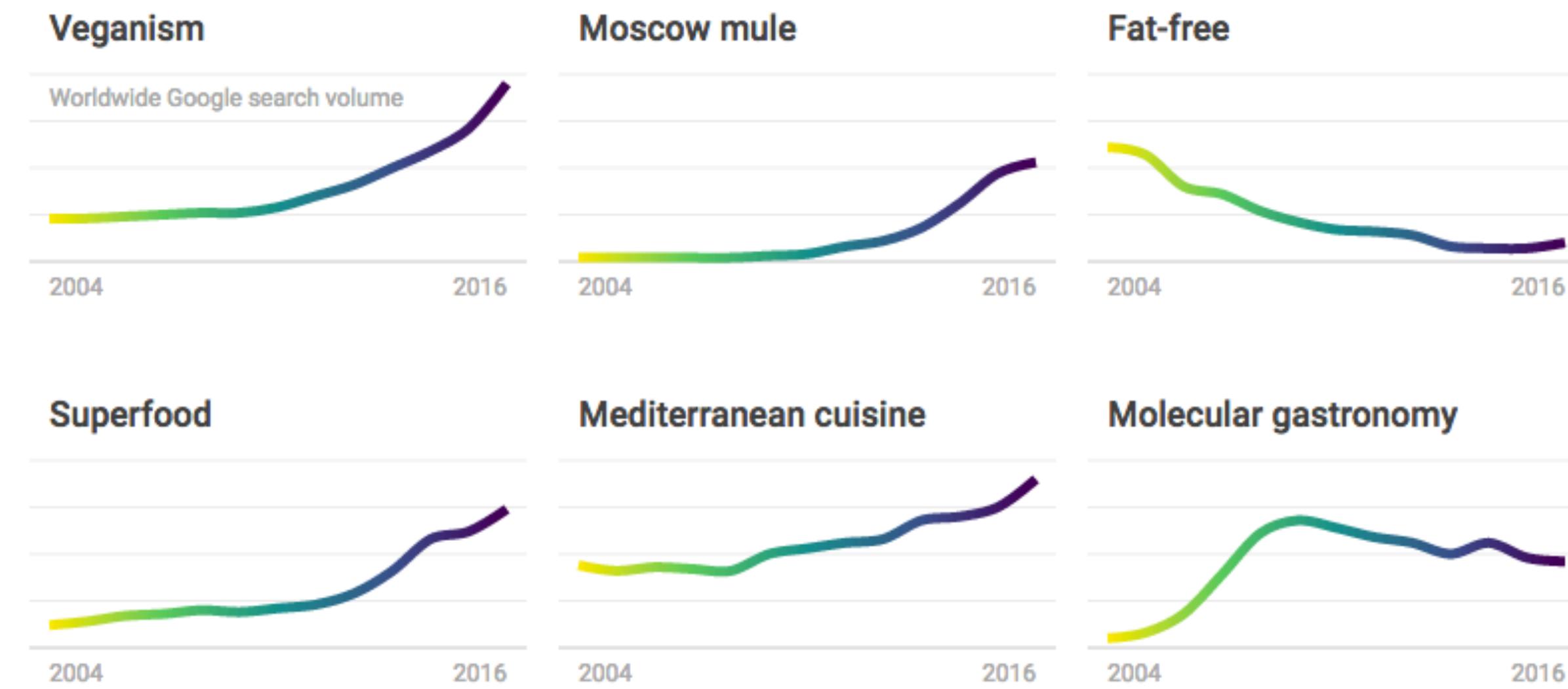


Nesse curso, vamos entender porque esse tipo de visualização, apesar de **complexa e densa** é **eficaz**

Entender o funcionamento dos sistemas de **percepção** e **cognição** humanos será fundamental para nos ajudar a projetar visualizações eficazes

**How do we search for food?  
Google search interest can reveal  
key food trends over the years.**

From the rise and fall of recipes over diets  
and drinks to cooking trends and regional  
cuisines.



<http://rhythm-of-food.net/>

But it turns out that we can learn even more from analyzing search patterns.

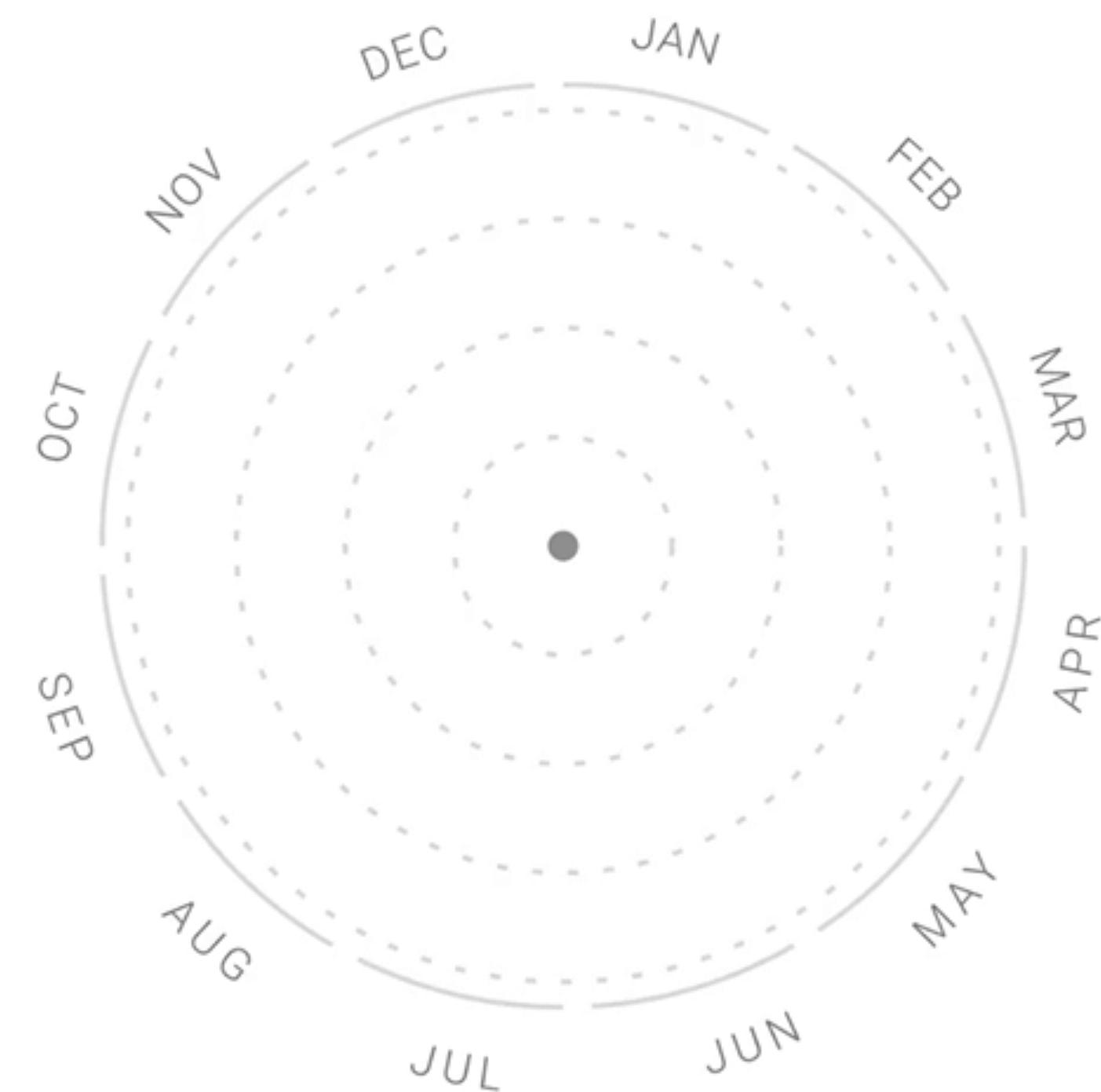
We collected weekly **Google Trends** data for hundreds of dishes and ingredients, over twelve years

Searches for  
**Apricot**

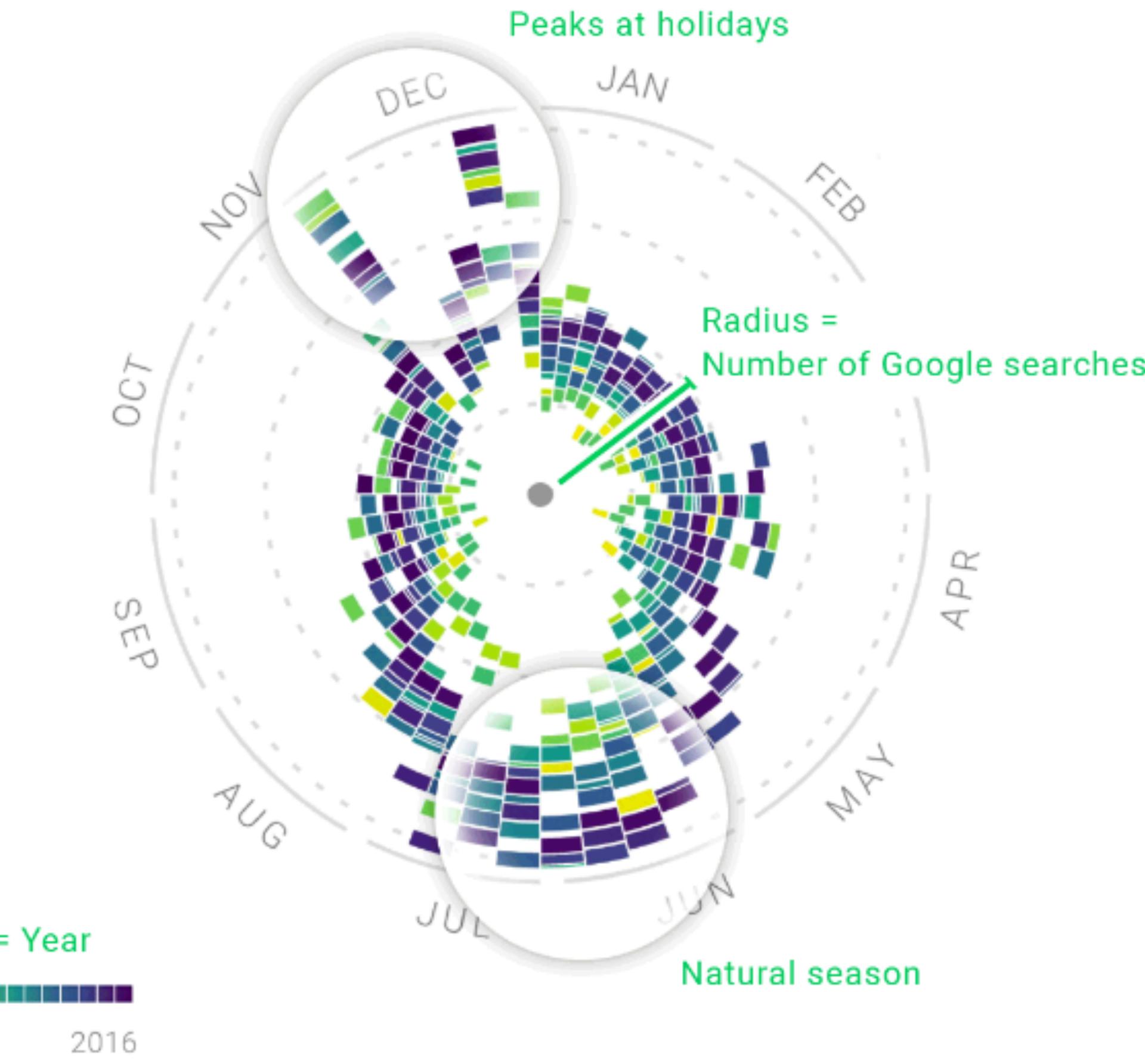


and plotted the results on a **year clock**

and plotted the results on a **year clock**

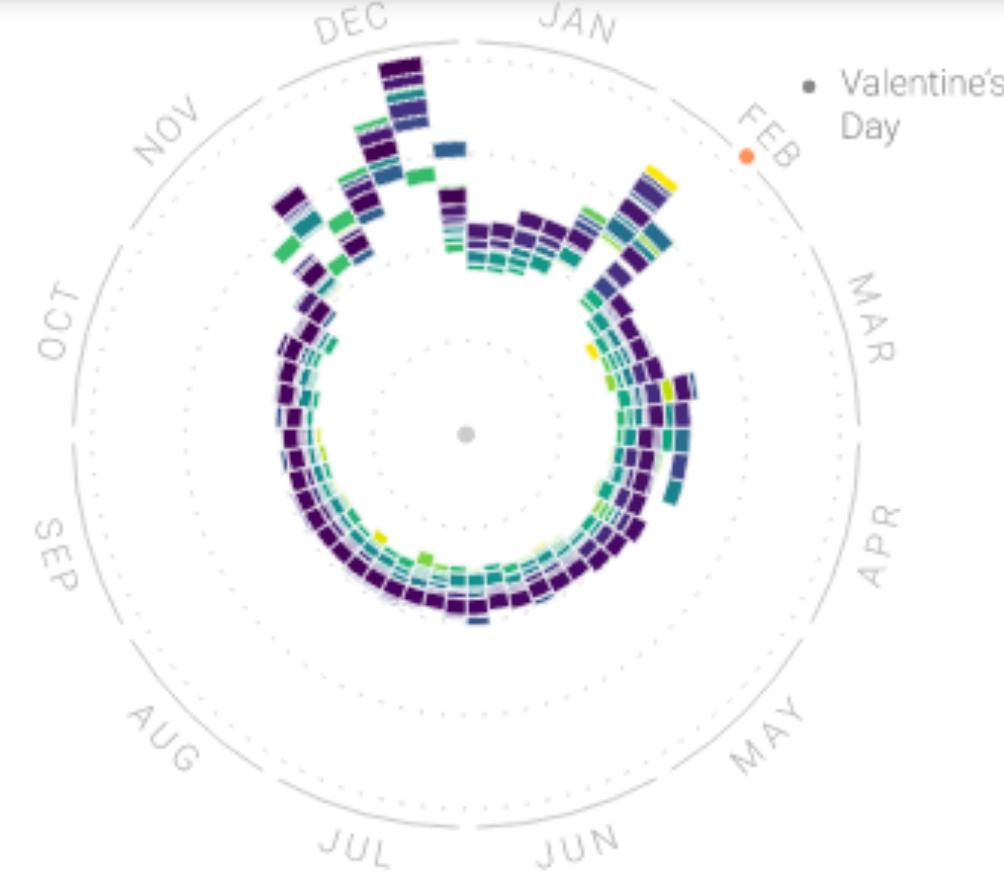


to investigate the seasons and  
**rhythm of food** around the world.

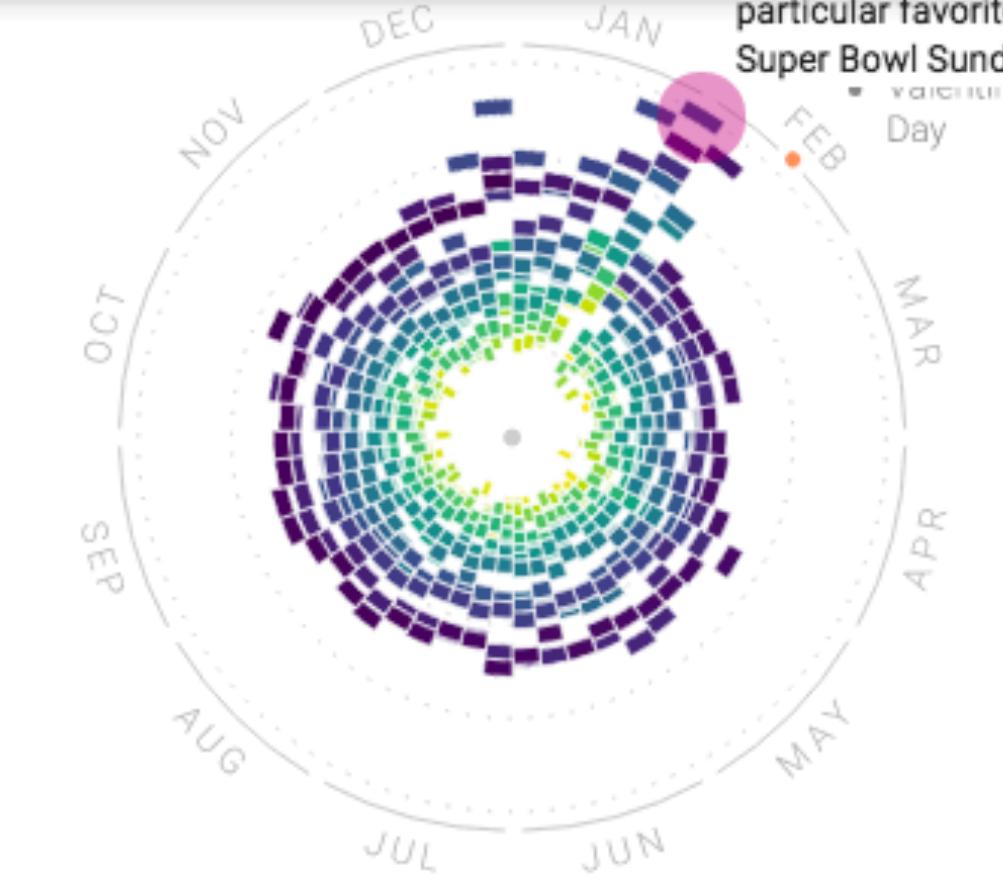


APRICOT

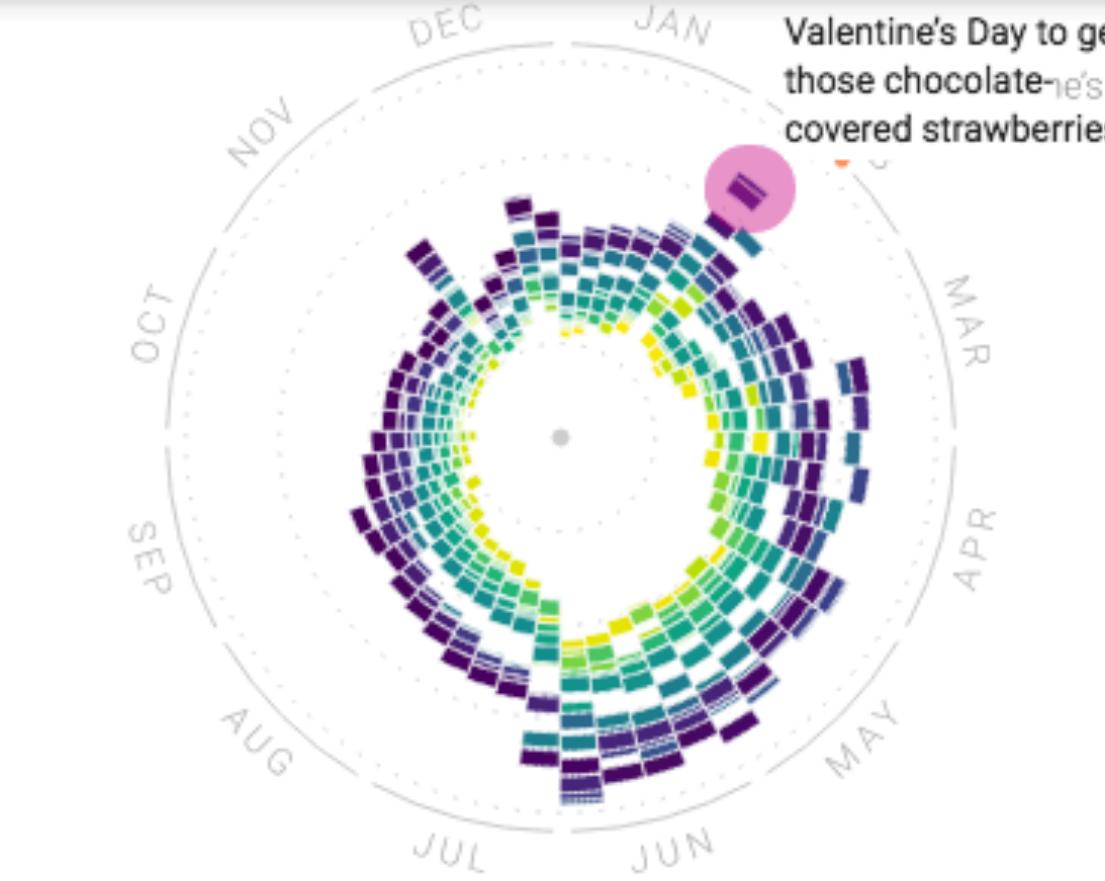
January	February	March	April	May	June	July	August	September	October	November	December
---------	----------	-------	-------	-----	------	------	--------	-----------	---------	----------	----------



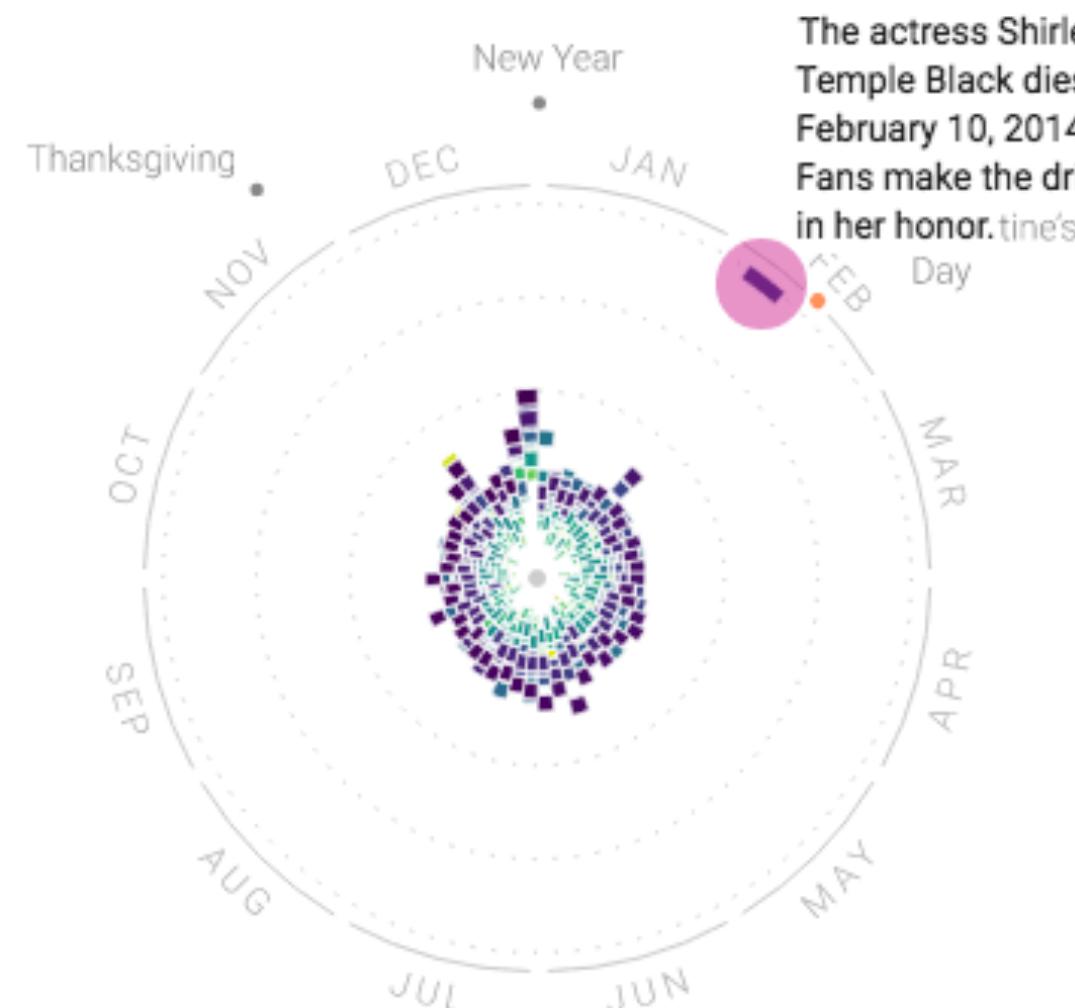
CHOCOLATE



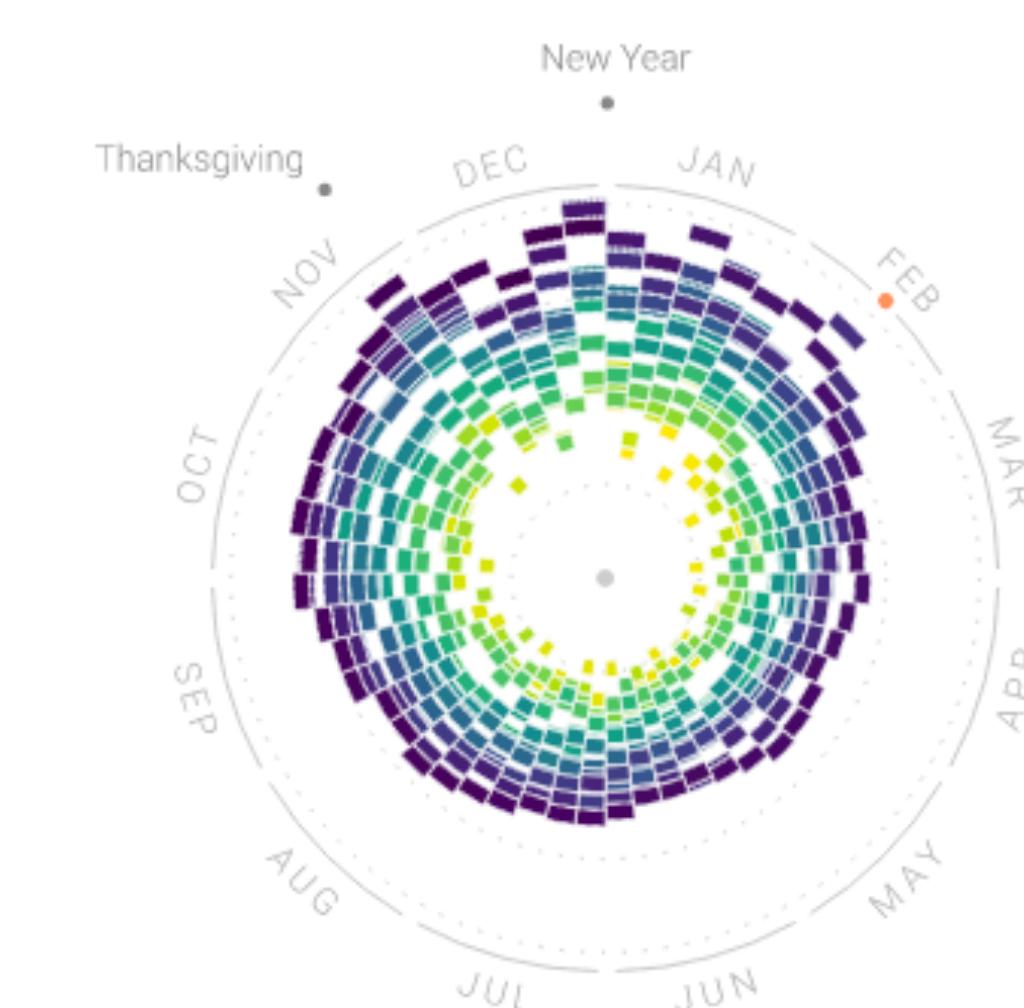
NACHOS



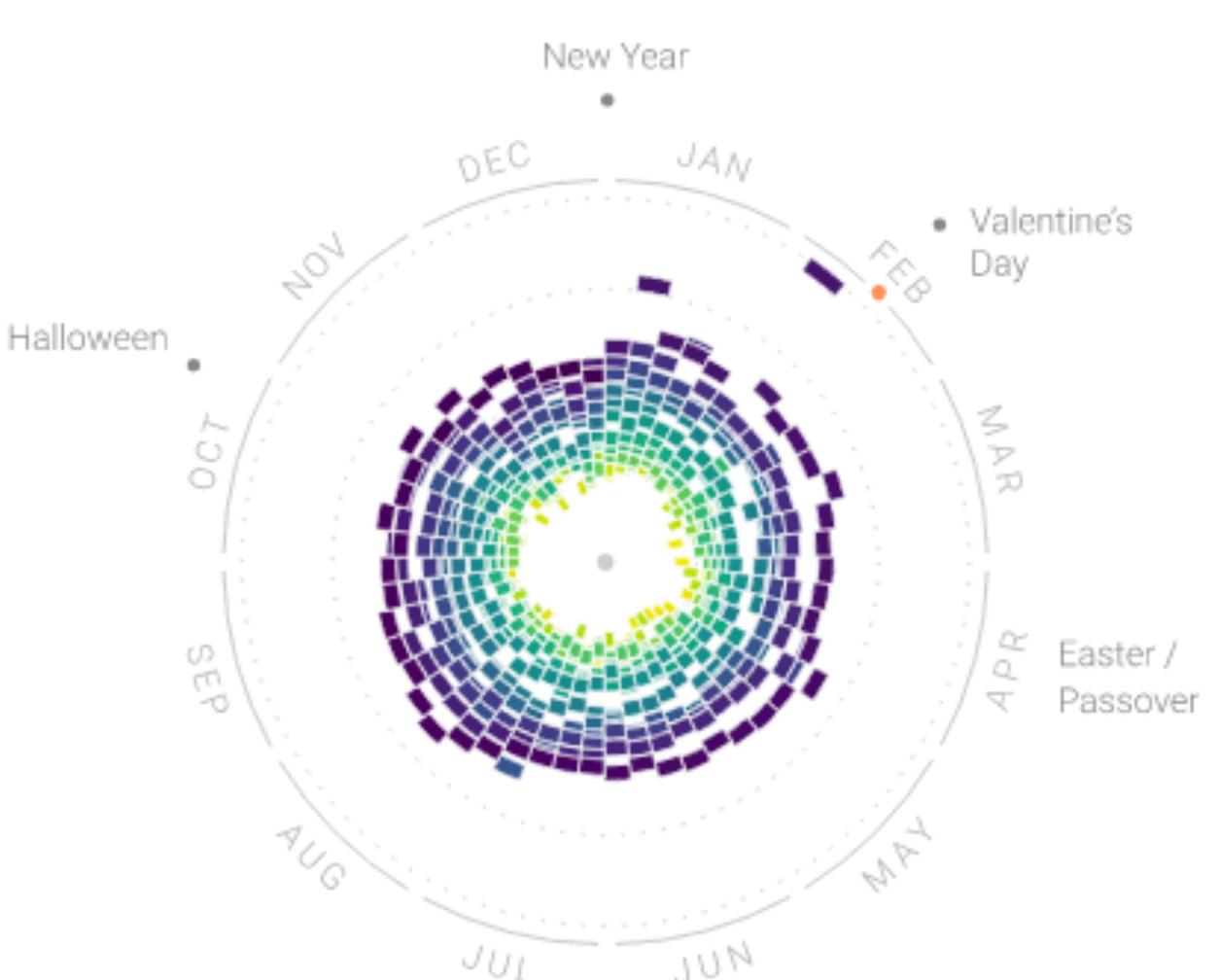
STRAWBERRY



SHIRLEY TEMPLE



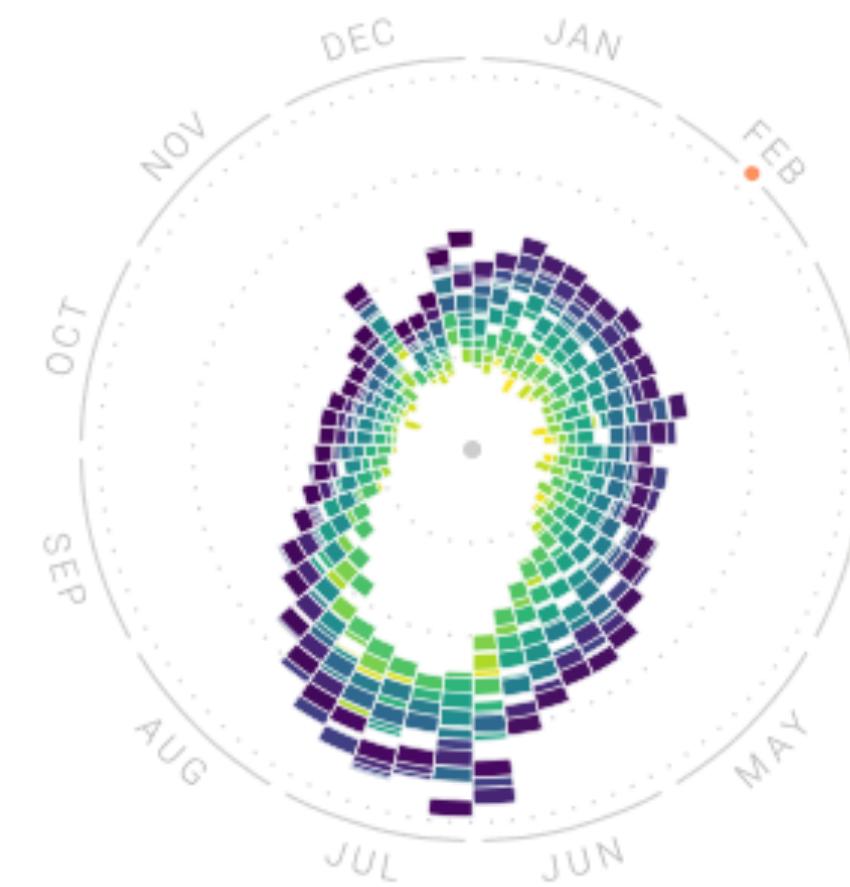
DUMPLING



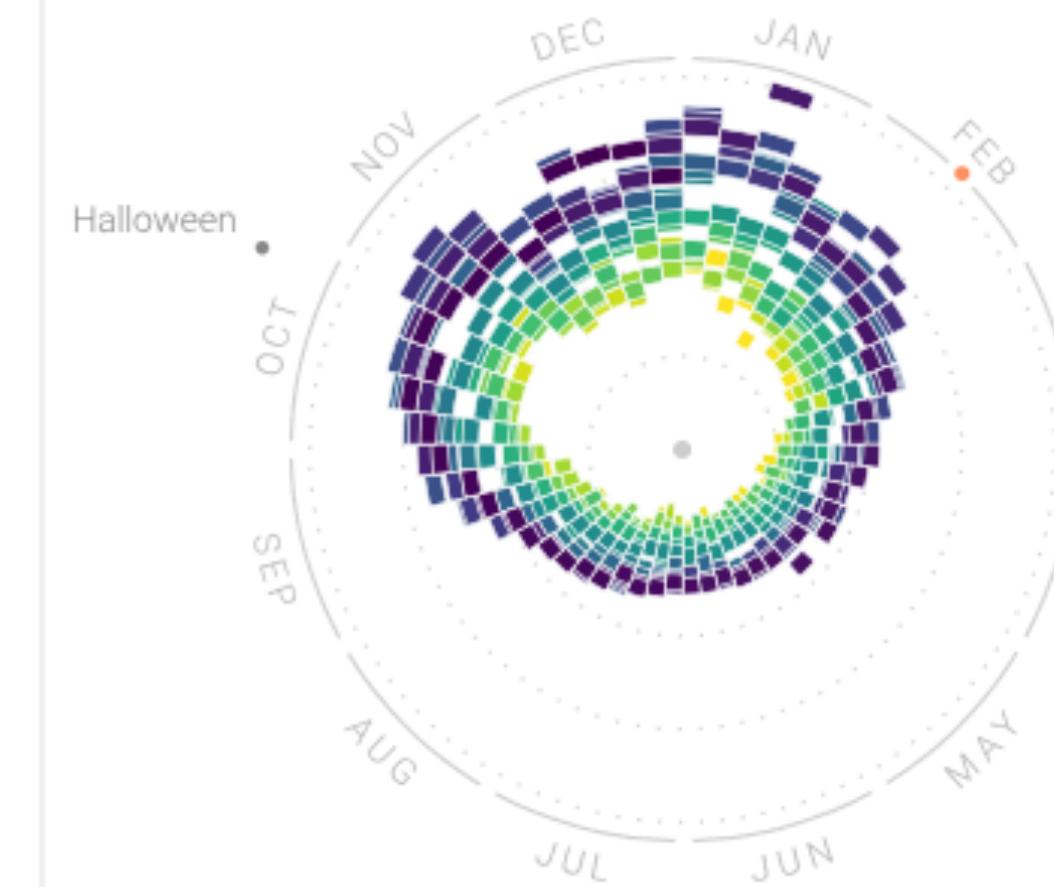
BURRITO

# What are the most common patterns?

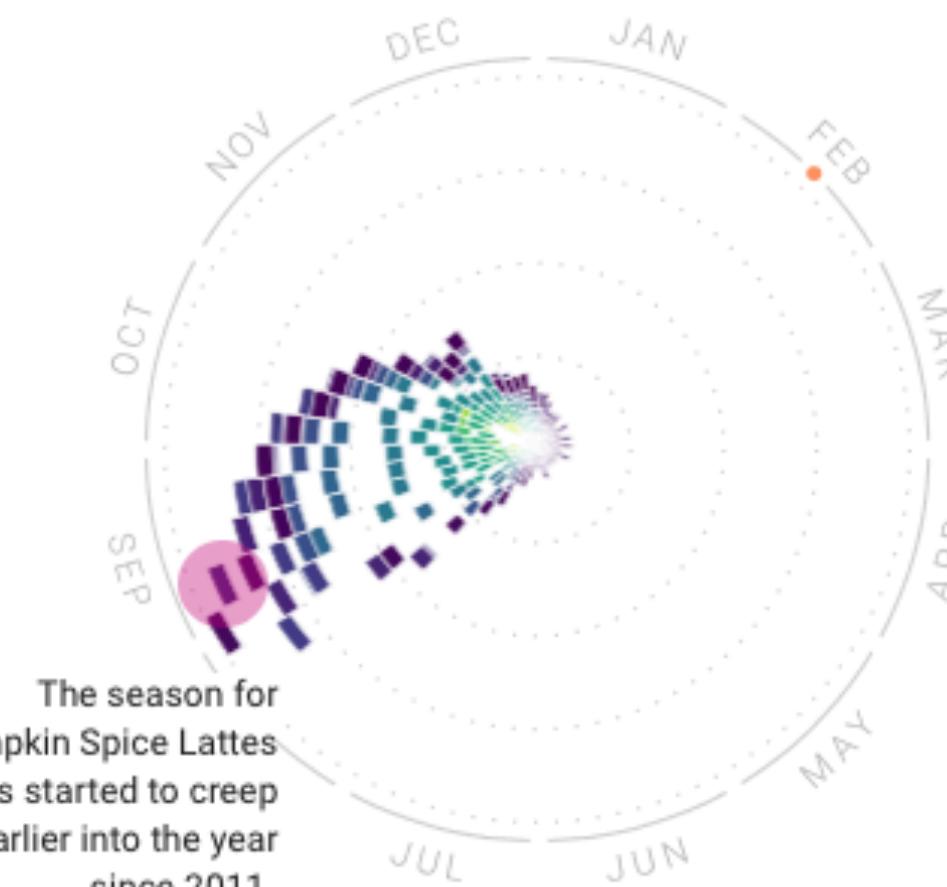
Some items fade in and out of a natural season.



BLUEBERRY



STEW

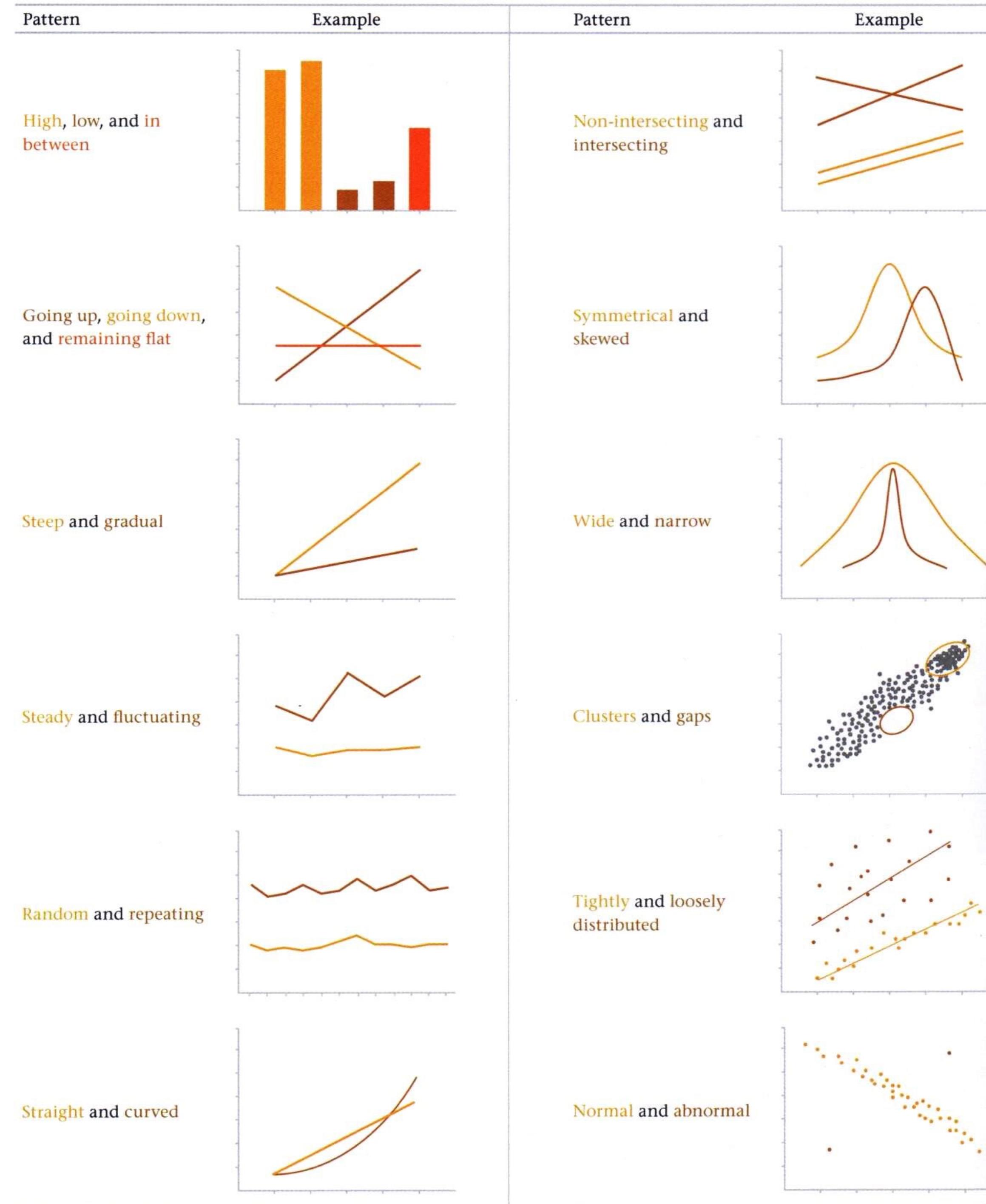


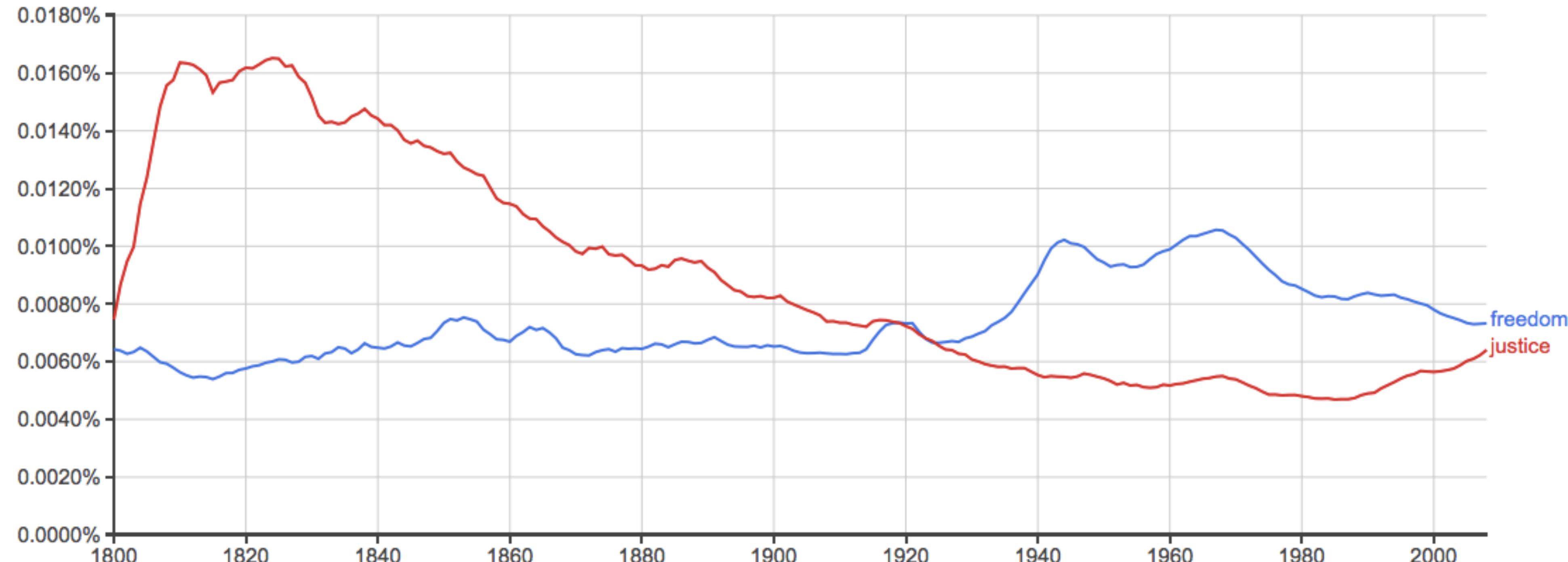
PUMPKIN SPICE LATTE

# Padrões visuais

Como padrões visuais particulares podem ser úteis em determinados problemas?

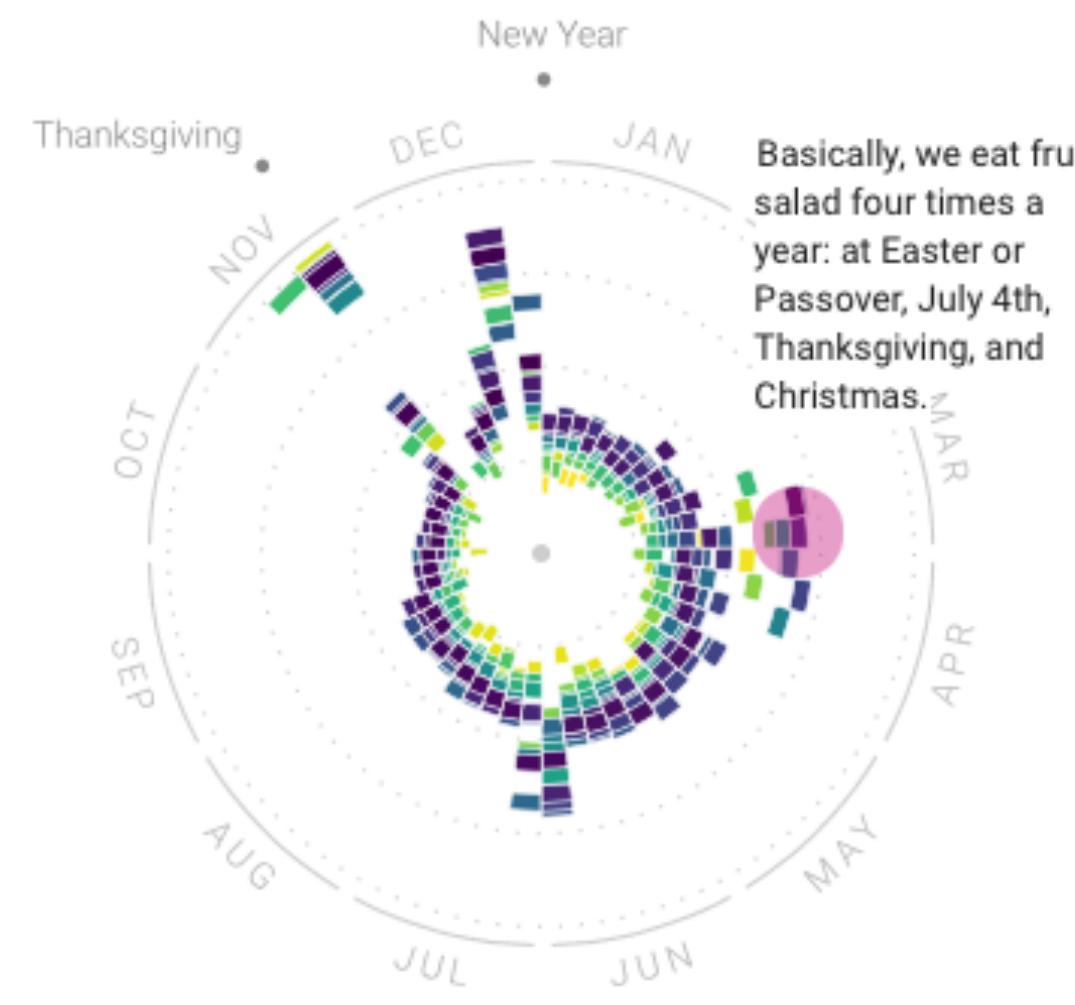
Quais as técnicas capazes de evidenciar certos tipos de padrões mais efetivamente?



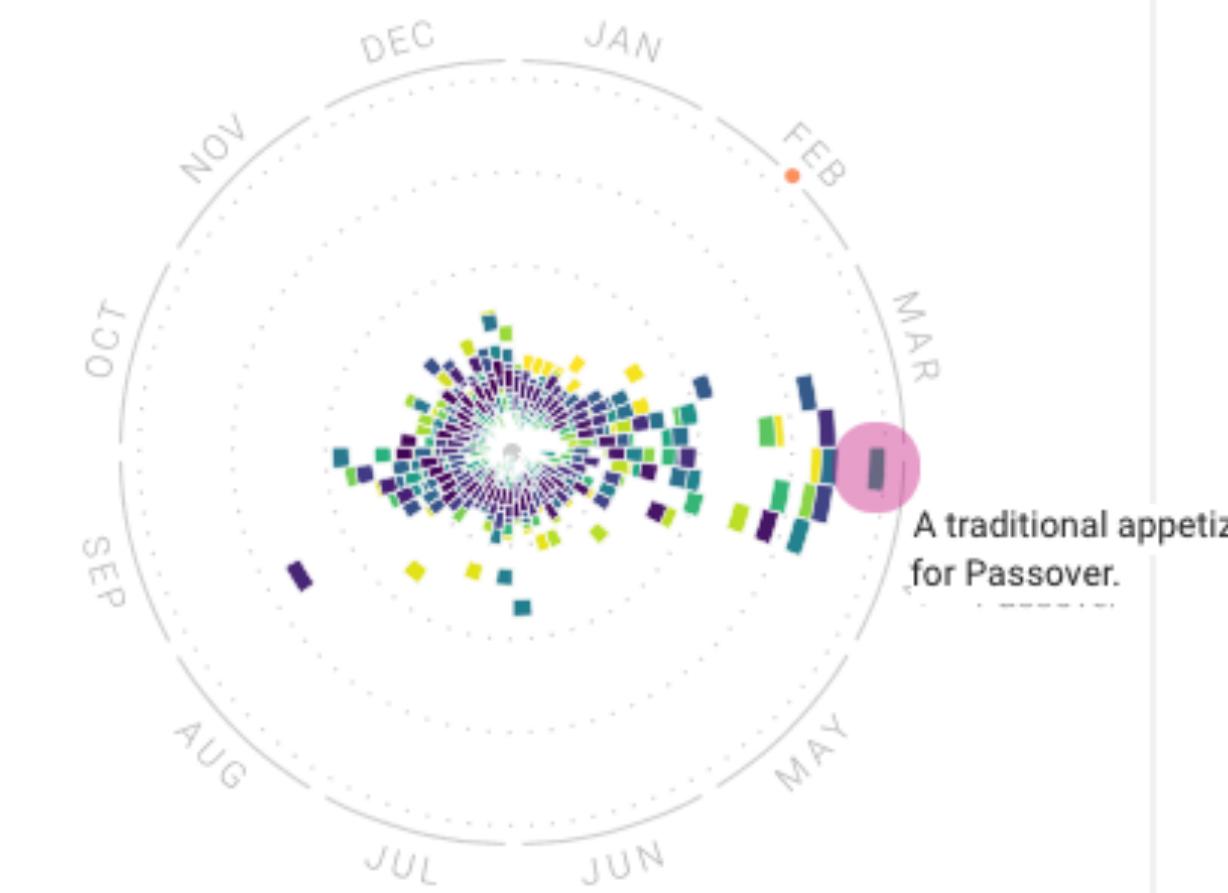


Google NGram Viewer

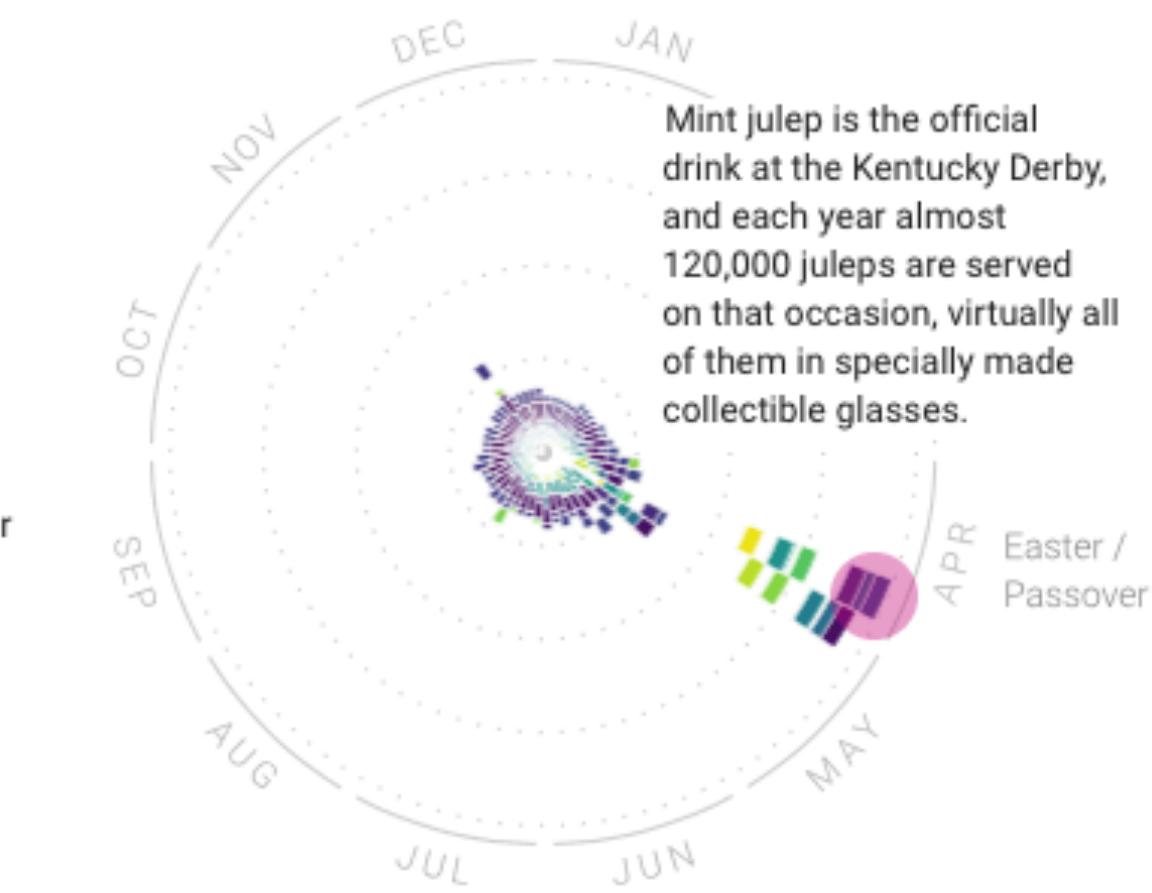
Others peak at holidays and special events.



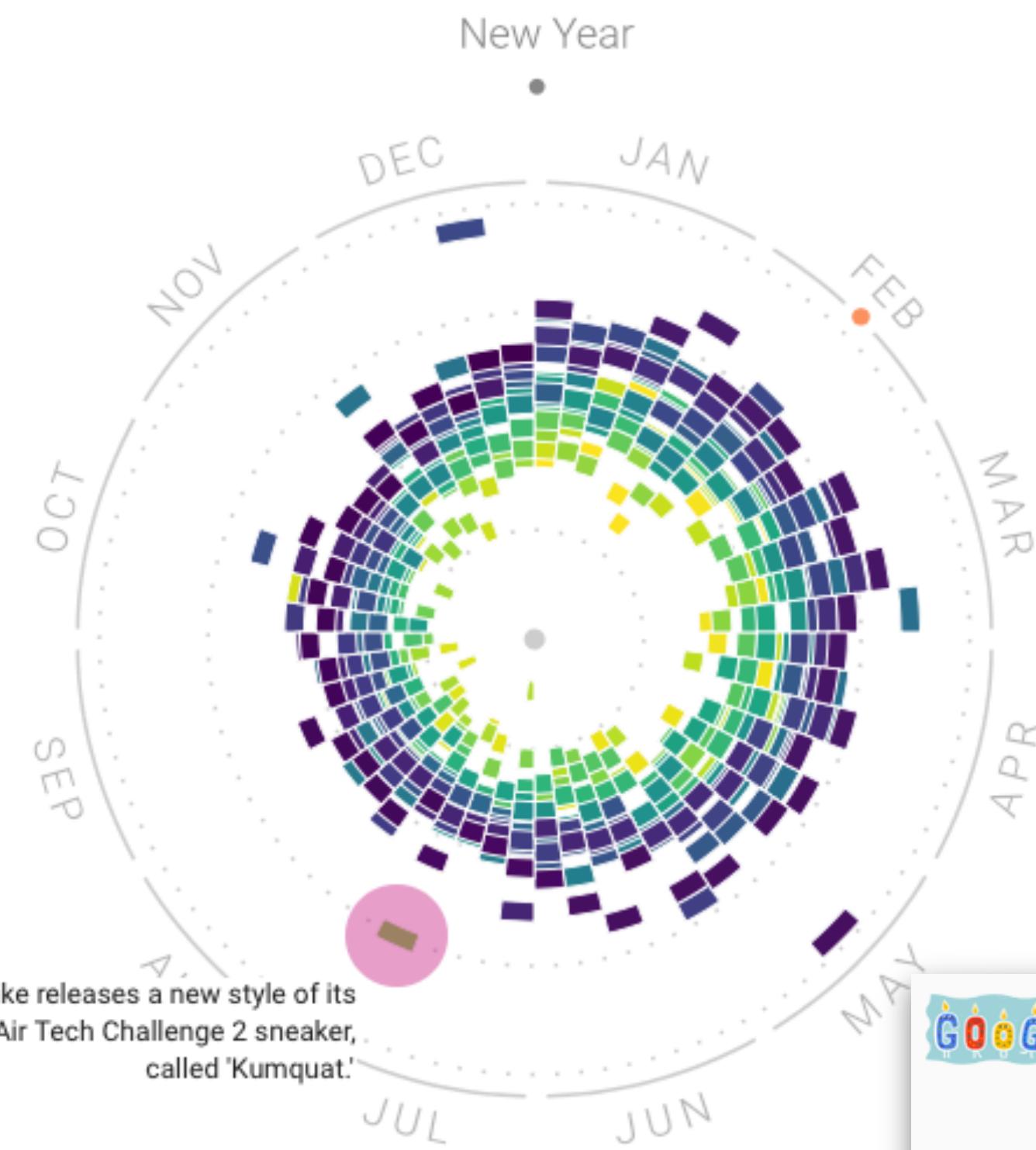
FRUIT SALAD



GEFILTE FISH



MINT JULEP



# KUMQUAT



## Fortunella

Fruto

A Citrus japonica, conhecida pelos nomes comuns de quincã ou cunquate, ou ainda xinxim, é uma pequena fruta cítrica da família das rutáceas. [Wikipédia](#)

### Informação Nutricional

Laranjas kinkan

Quantidade por 100 gramas

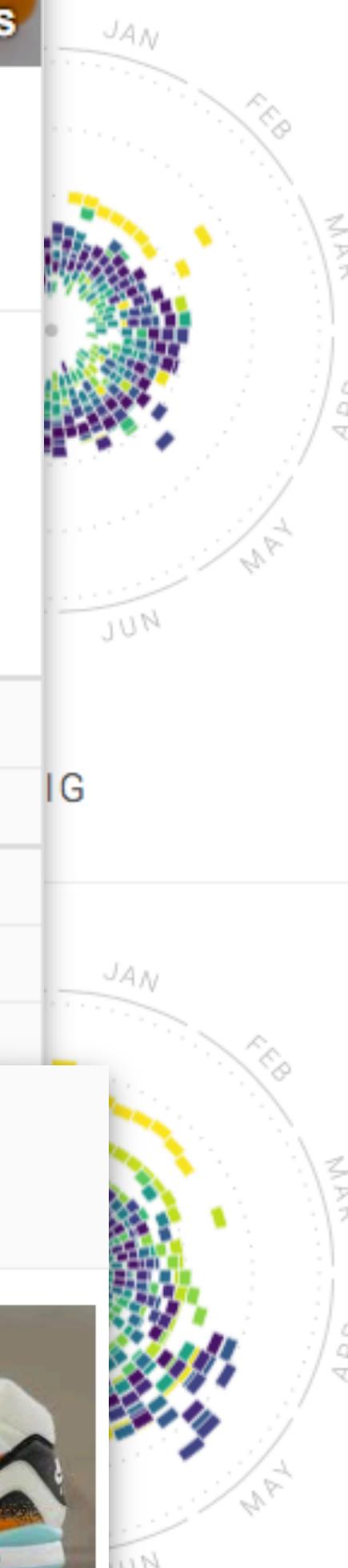
Calorias 71

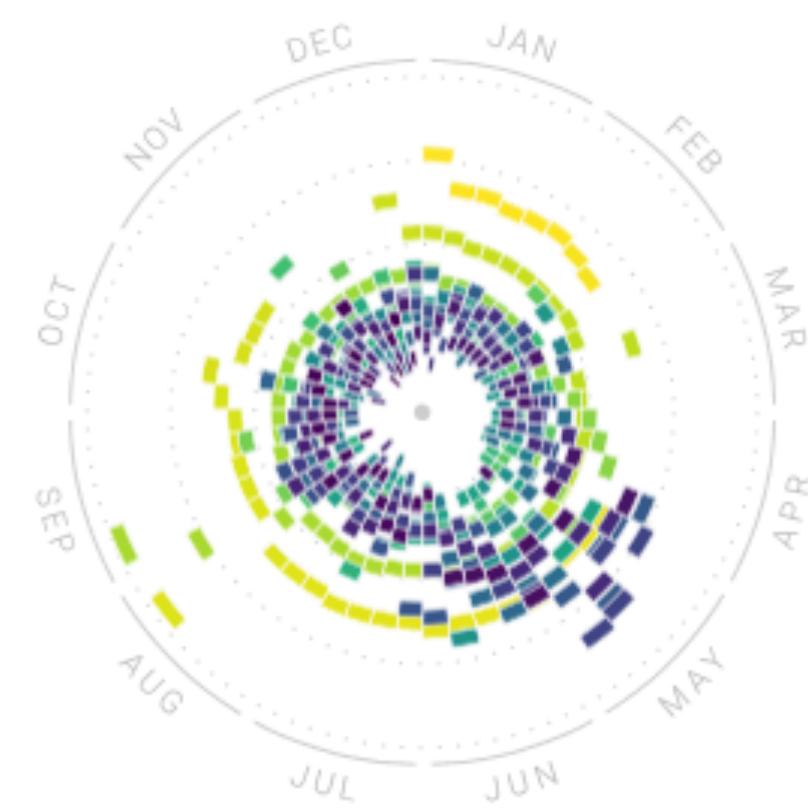
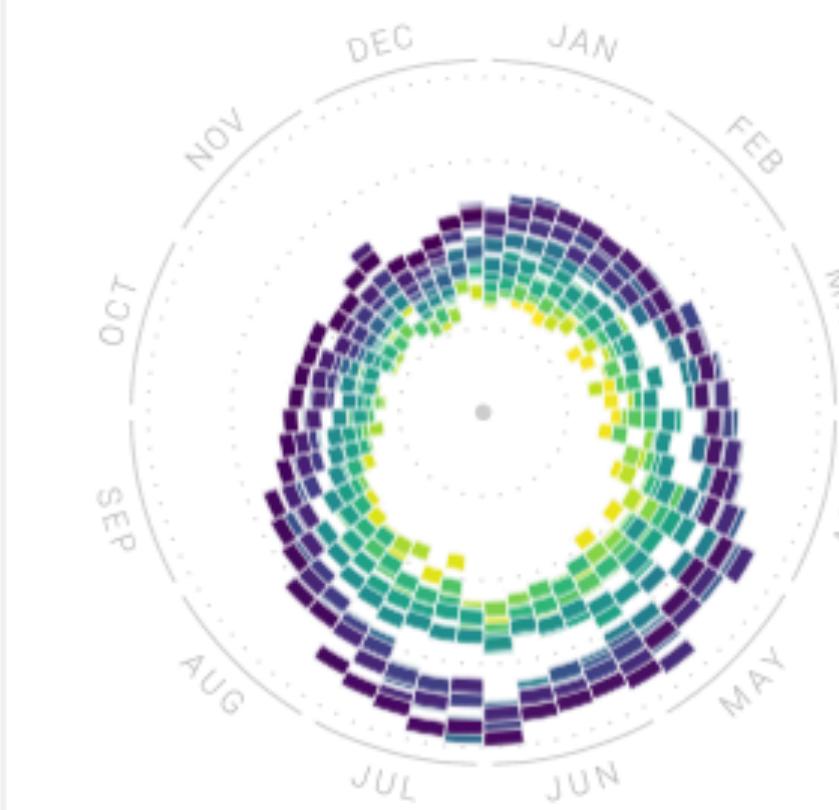
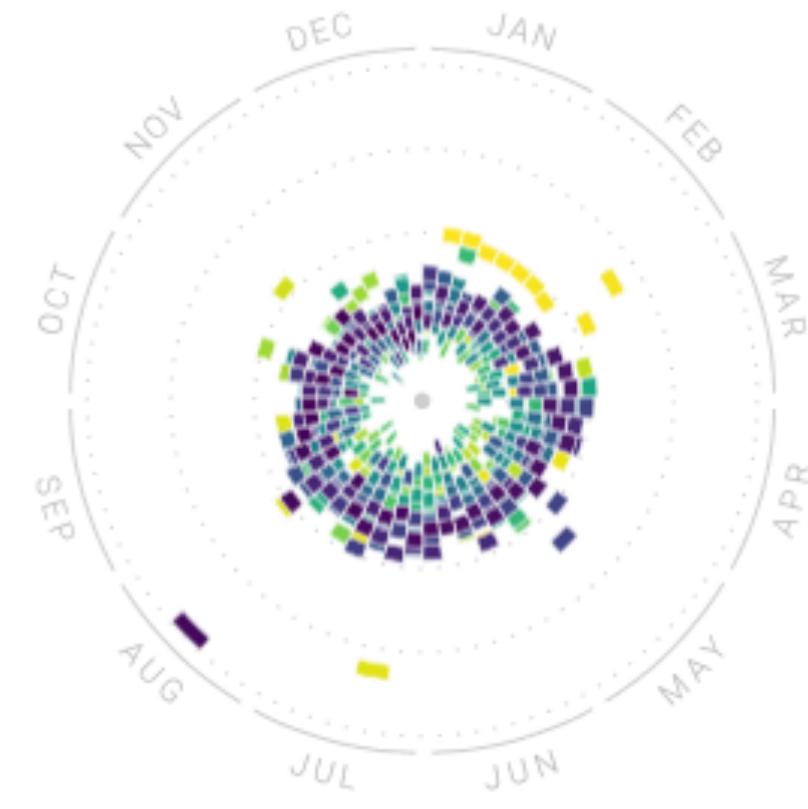
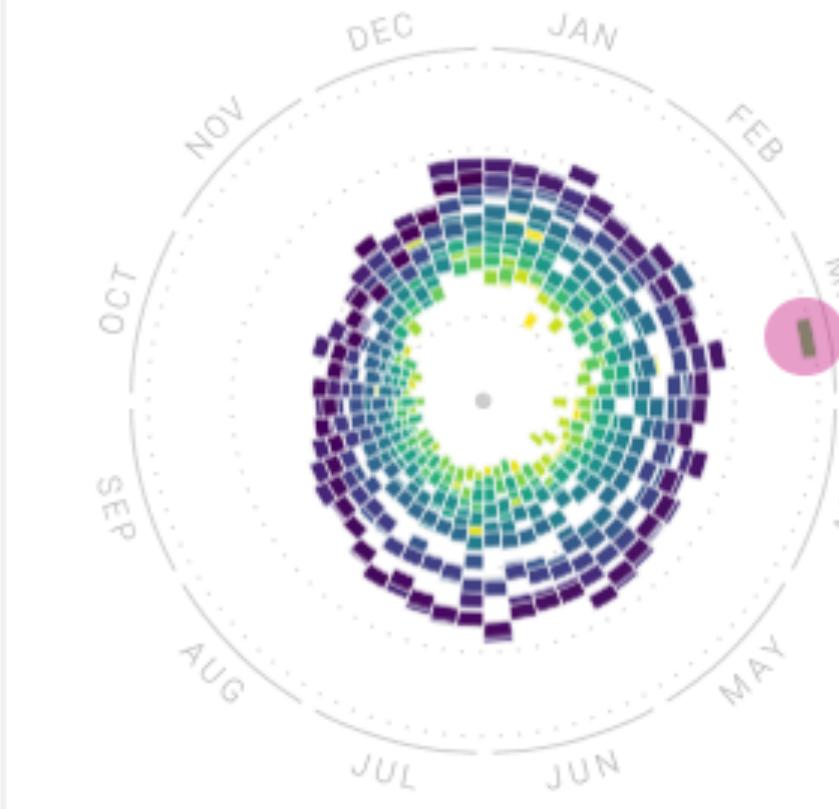
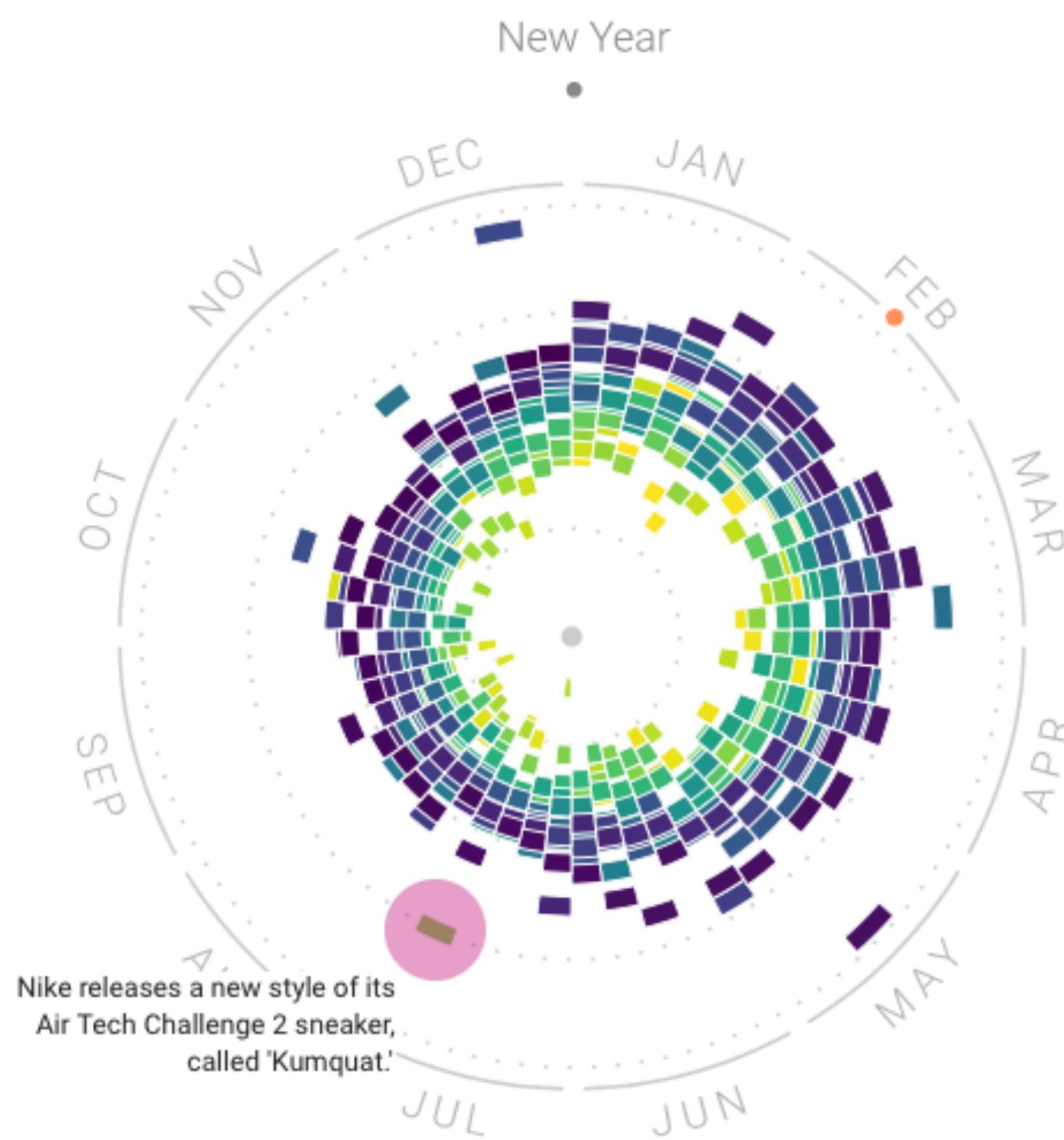
Gorduras Totais 0,9 g

Gorduras Saturadas 0,1 g

Gorduras Poliinsaturadas 0,2 g

IG





# ALGUNS DOS OBJETIVOS DE UMA VISUALIZAÇÃO

---

- Identificar informações ocultas nos dados
- Identificar tendências
- Identificar exceções ou anomalias
- Prever acontecimentos, dadas condições particulares, prevenindo problemas

# PARA QUE TIPO DE DADOS É INTERESSANTE

---

- Volume
- Histórico
- Consistente
- Multivariados
- Atômicos
- Limpos
- Claros
- Estruturados
- Ricamente segmentados
- De origem conhecida

# Percepção visual

---

**Profa. Dra. Raquel Minardi**  
Departamento de Ciência da Computação  
Universidade Federal de Minas Gerais

## Why should we be *interested* in *visualization*?

Because *human visual system is a pattern seeker* of enormous power and subtlety

“

The eye and the cortex of the brain form a *massively parallel processor* that provides the highest-bandwidth channel into human cognitive centers

*Perception and cognition are closely interrelated*, which is the reason why the words ‘understanding’ and ‘seeing’ are synonymous

-Stuart Card  
in Visualization: Perception for Design

However, the *visual system has its own rules*

“

We can *easily see patterns* presented in certain ways, but if they are presented in other ways, they become *invisible*

If we can *understand how perception works*, our knowledge can be translated into *rules for displaying information*

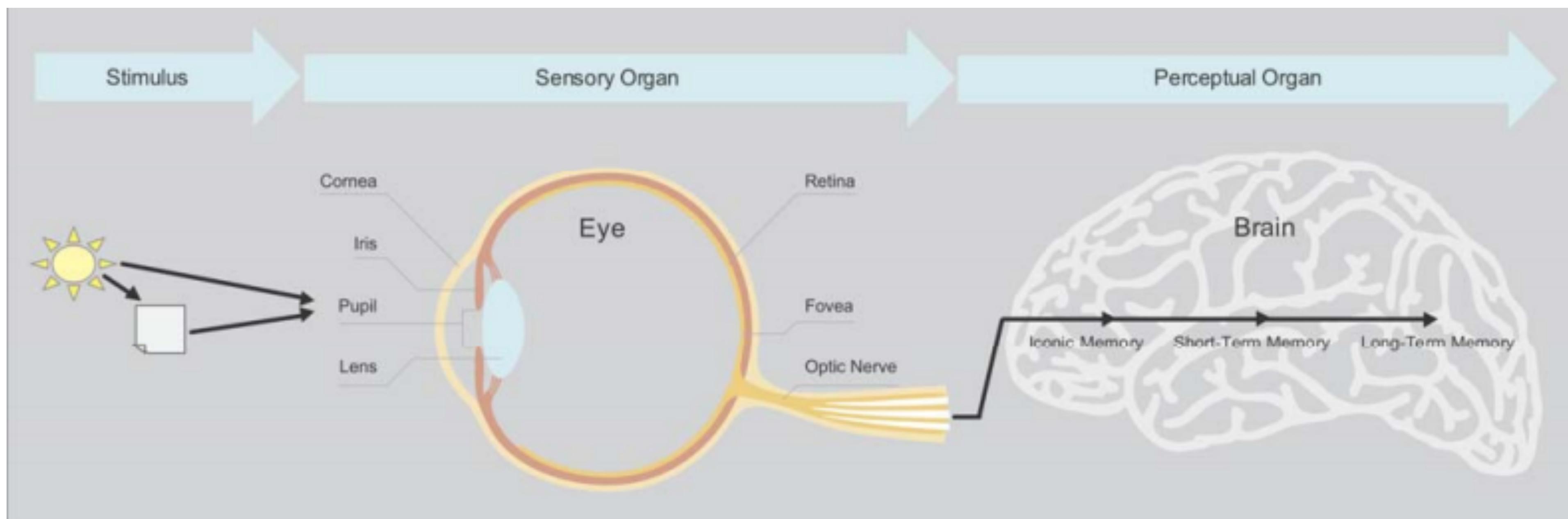
-Stuart Card

in Visualization: Perception for Design

# COMO FUNCIONA



# COMO FUNCIONA



- Nossos olhos são sensíveis à luz que reflete dos objetos
- Nossa percepção de um objeto é construída no nosso cérebro como uma composição de propriedades visuais
- Somos capazes de perceber esta composição de propriedades como um objeto, mas também de distinguir as propriedades individualmente
- Estas propriedades individuais, que nossa visão está preparada para identificar, são: localização, tamanho, área, forma, cor, orientação, entre outras

# COMO FUNCIONA



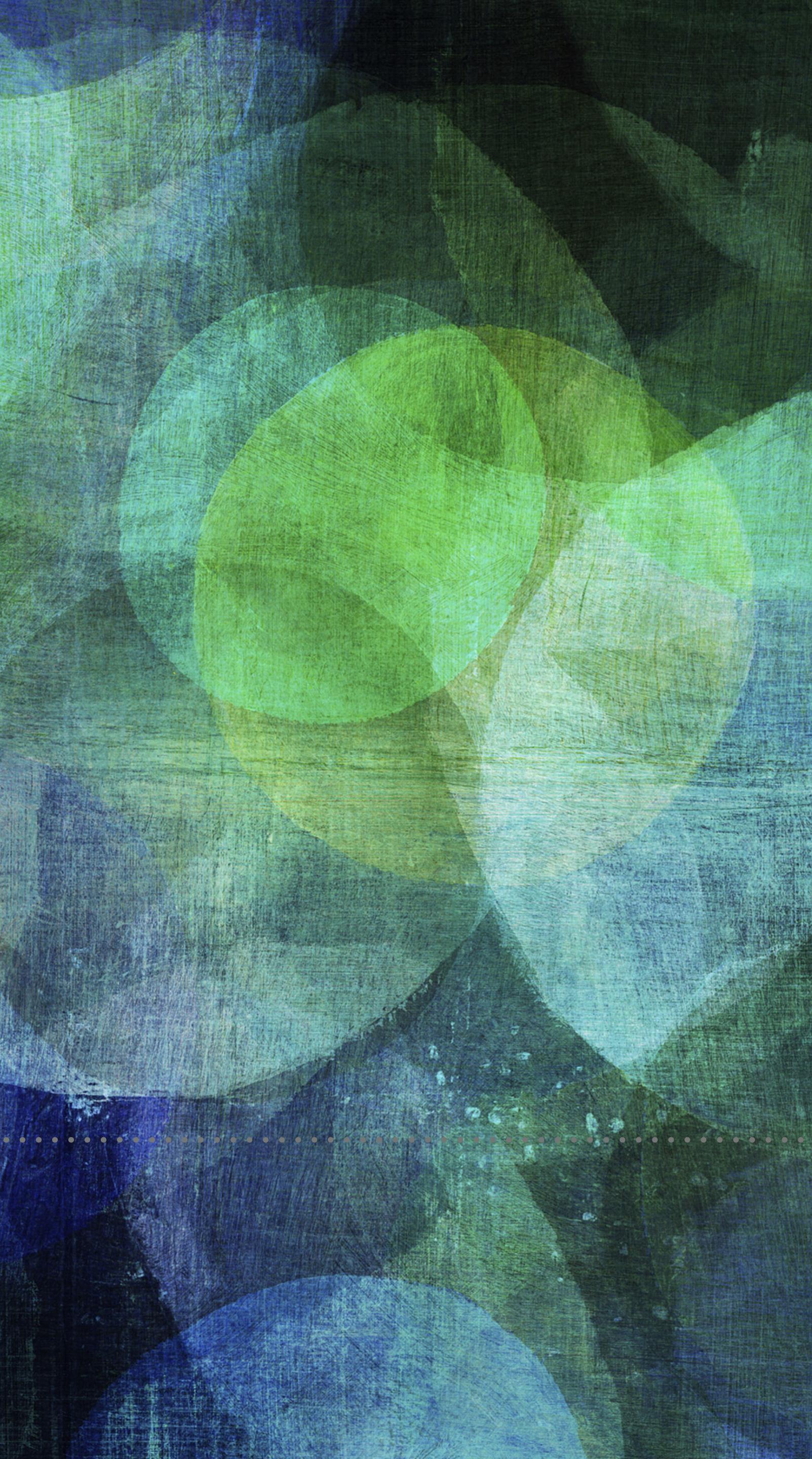
# COMO FUNCIONA

---

- Ao invés de ler valores individualmente, como em tabelas ou em texto, através de representações visuais dos dados podemos perceber e compreender inúmeros valores de uma só vez
- Isto é possível pois nossa visão pode detectar os padrões formados pelos atributos visuais

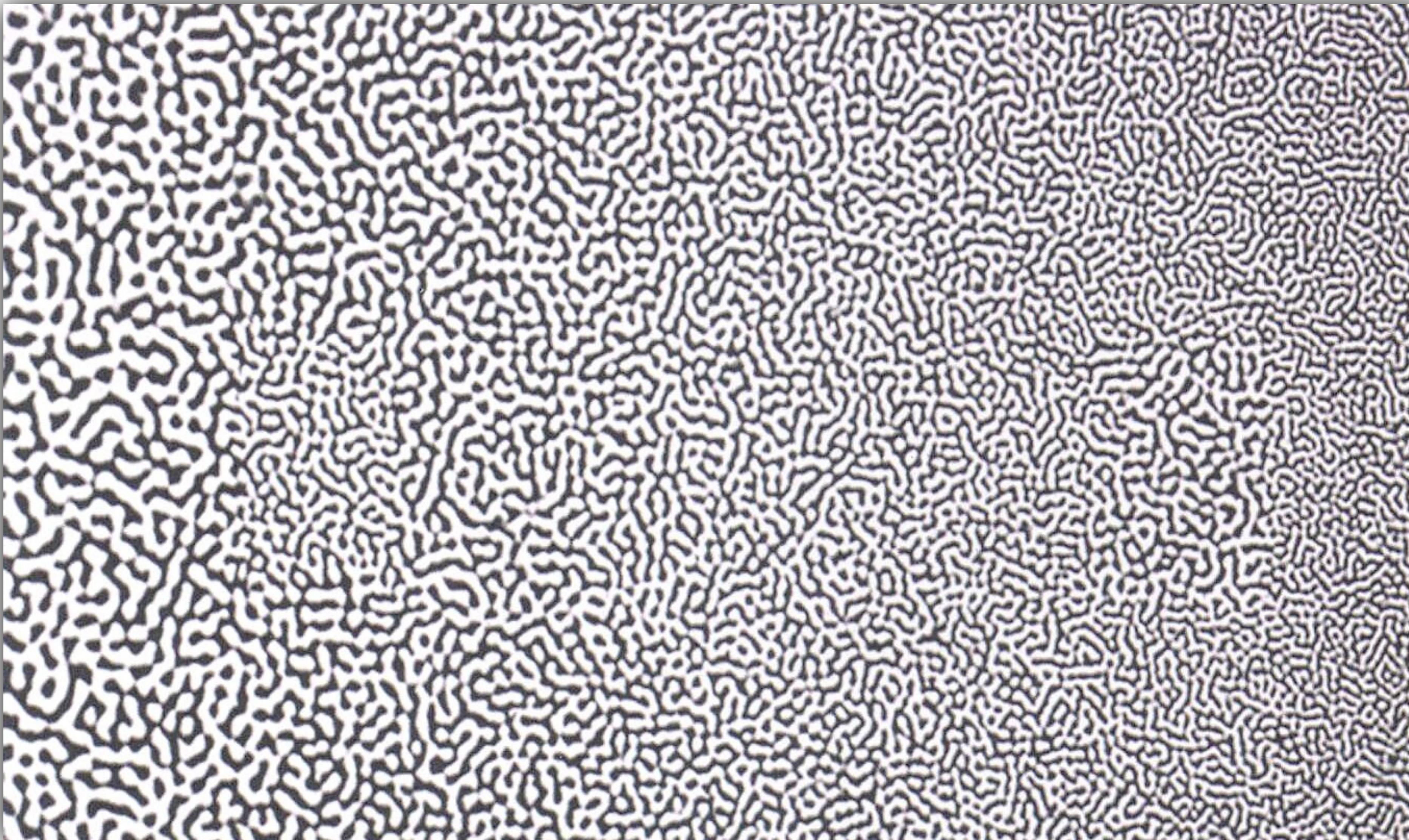
# USANDO CONHECIMENTOS SOBRE A PERCEPÇÃO VISUAL NA CRIAÇÃO DE VISUALIZAÇÕES EFETIVAS

---



# FATO #1

---



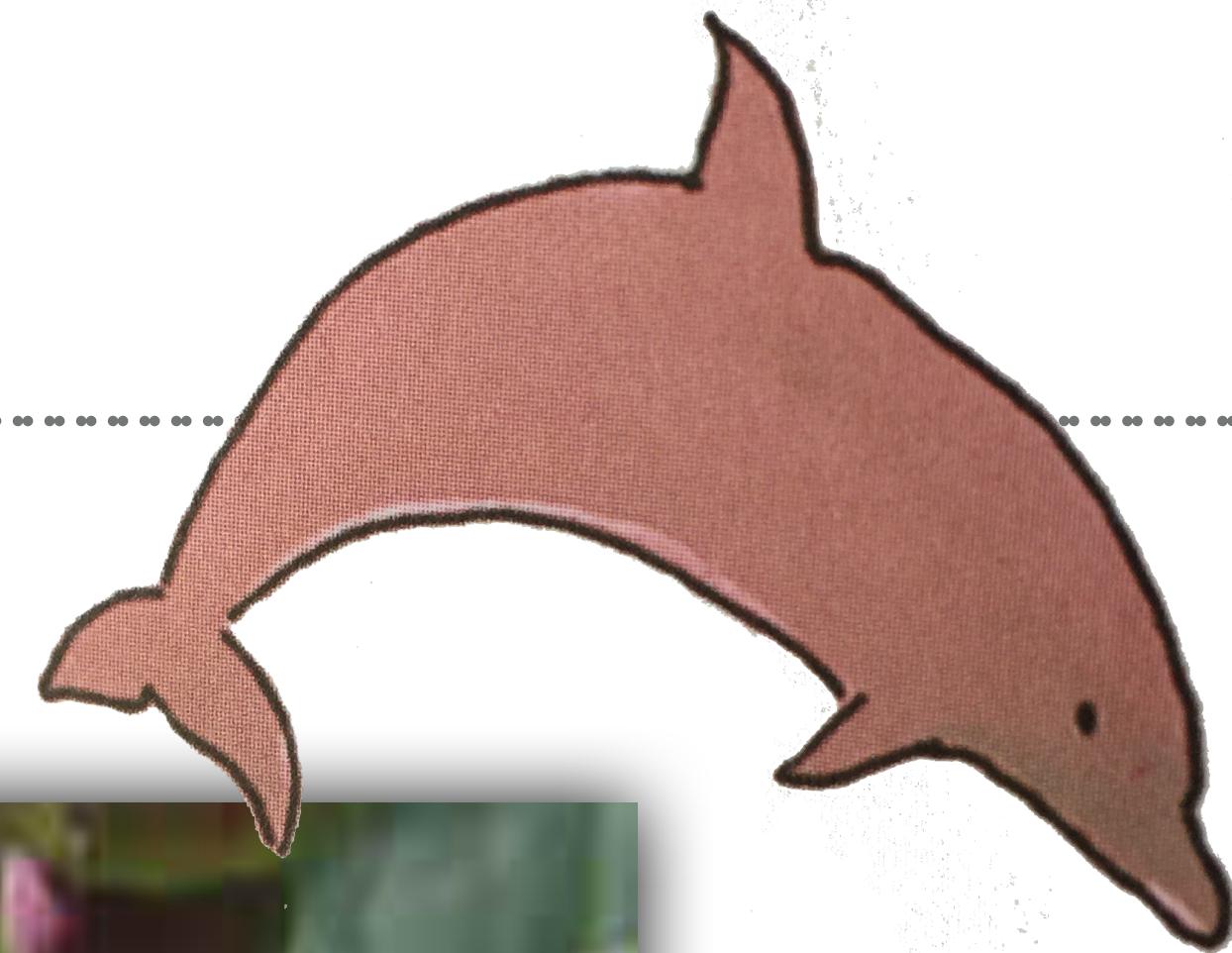
# FATO #1

---

- Não prestamos atenção a tudo o que vemos. A percepção visual é seletiva. Comumente, nossa atenção dirige-se aos contrastes
- Quando construímos visualizações, o que é interessante e potencialmente cheio de significado deve se distinguir do que não é

## FATO #2

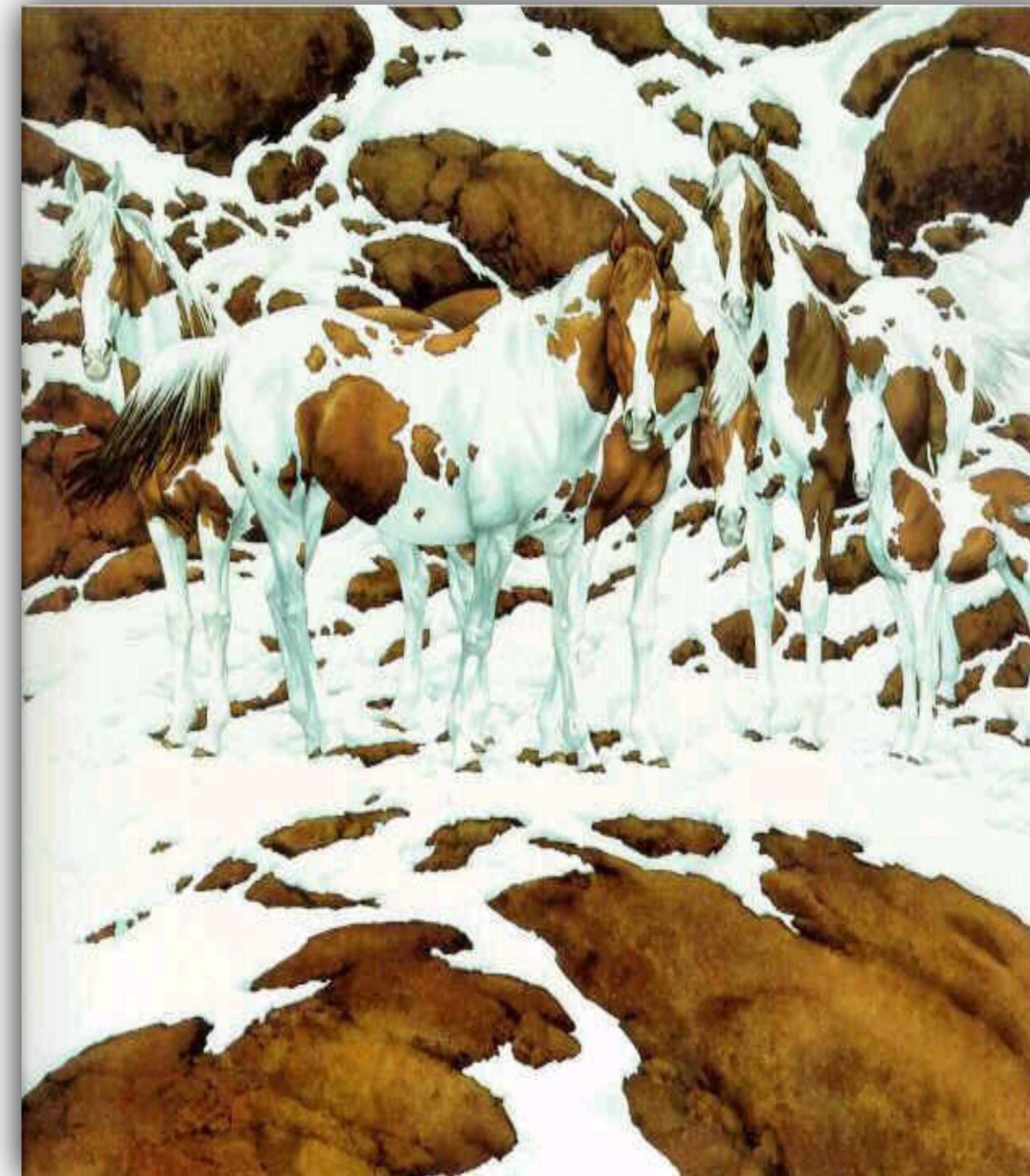
---



## FATO #2

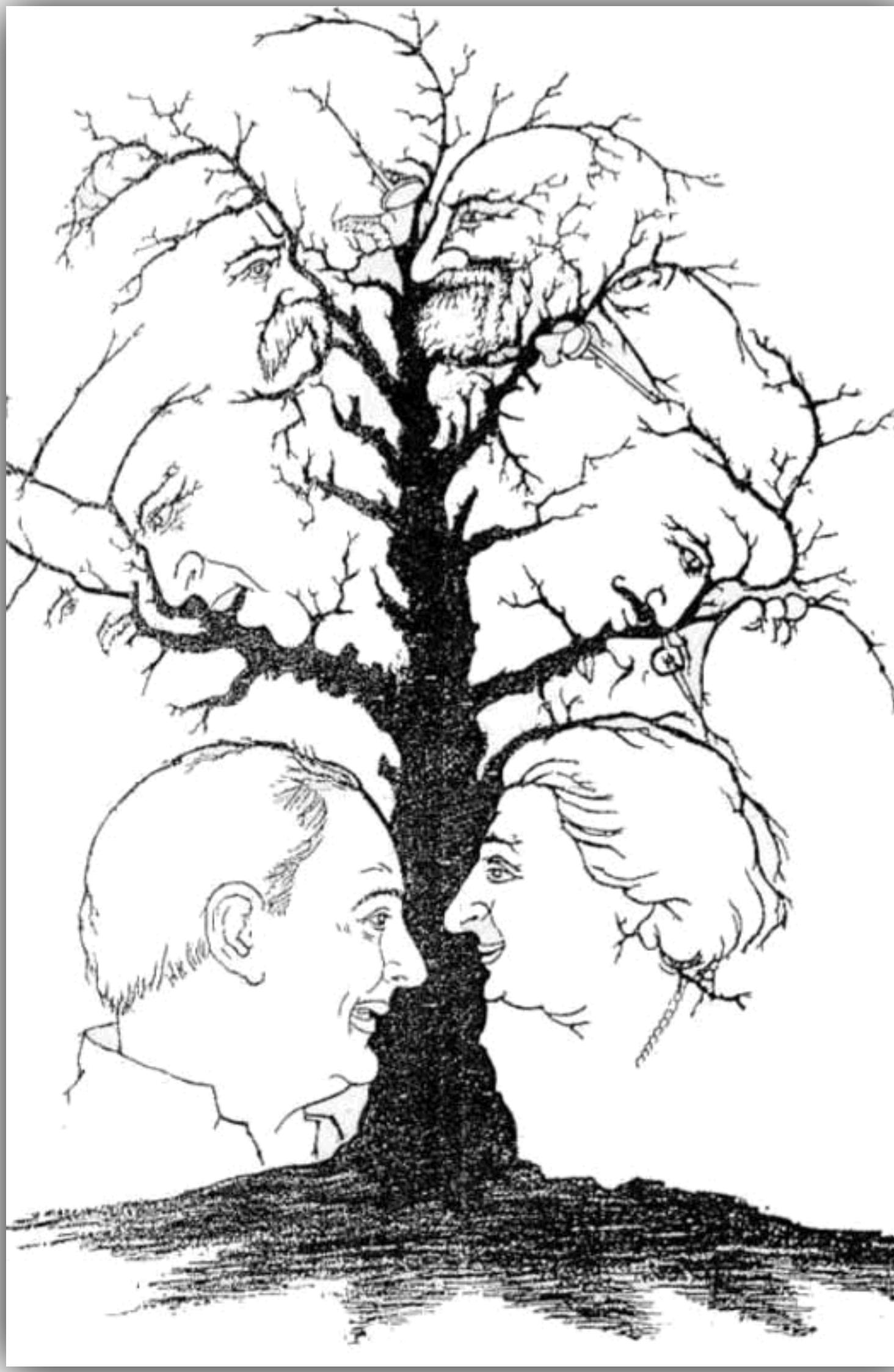


## FATO #2



# FATO #2

---



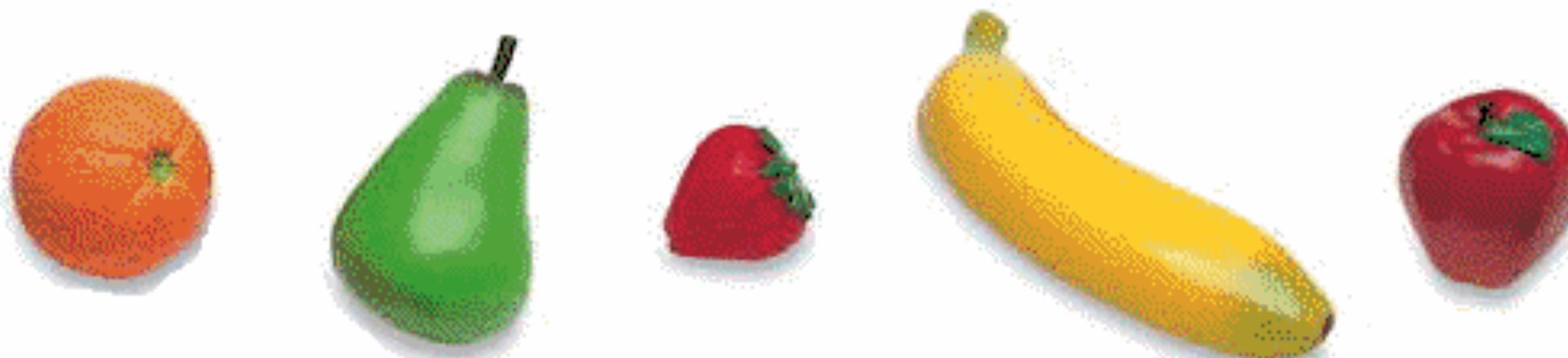
## FATO #2

---

- Nosso sistema de percepção é sensível a padrões familiares, ou seja, nós somos capazes de identificar padrões conhecidos e esperados

# FATO #3

---

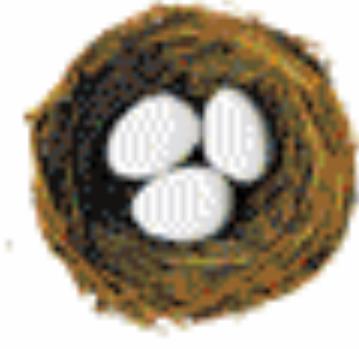


# FATO #3

---

# FATO #3

---



# FATO #3

---

# FATO #3

---

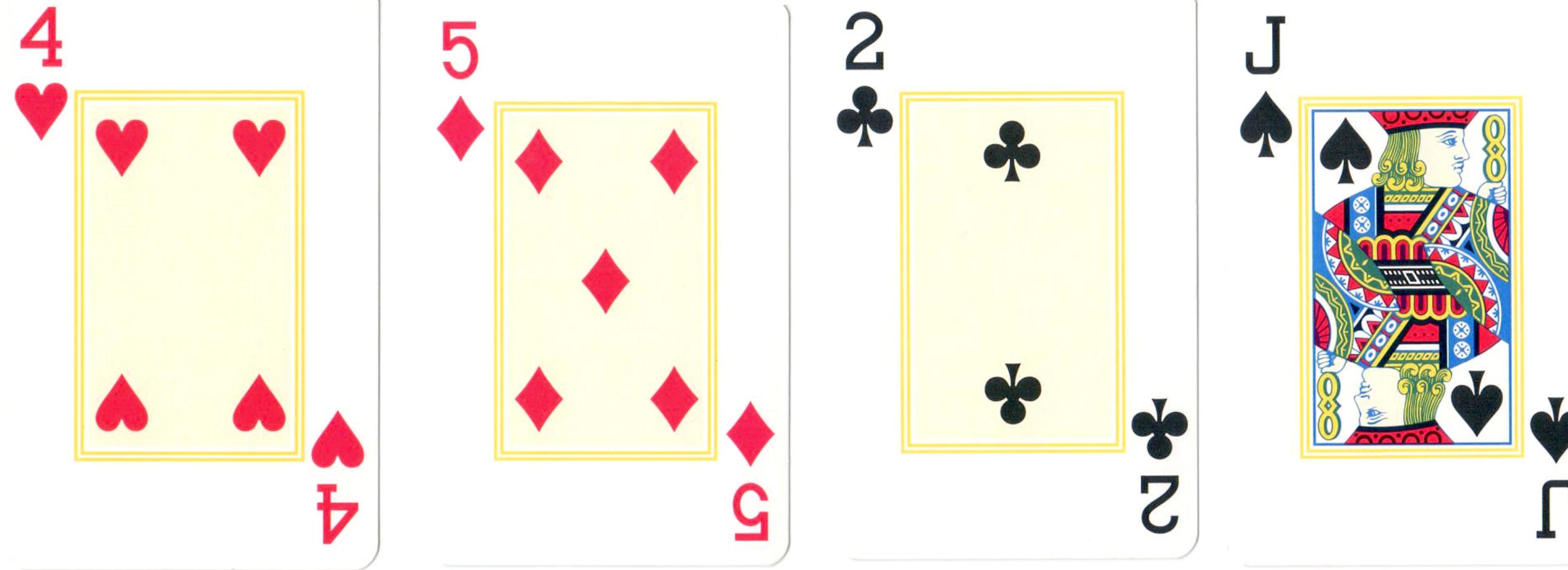


# FATO #3

---

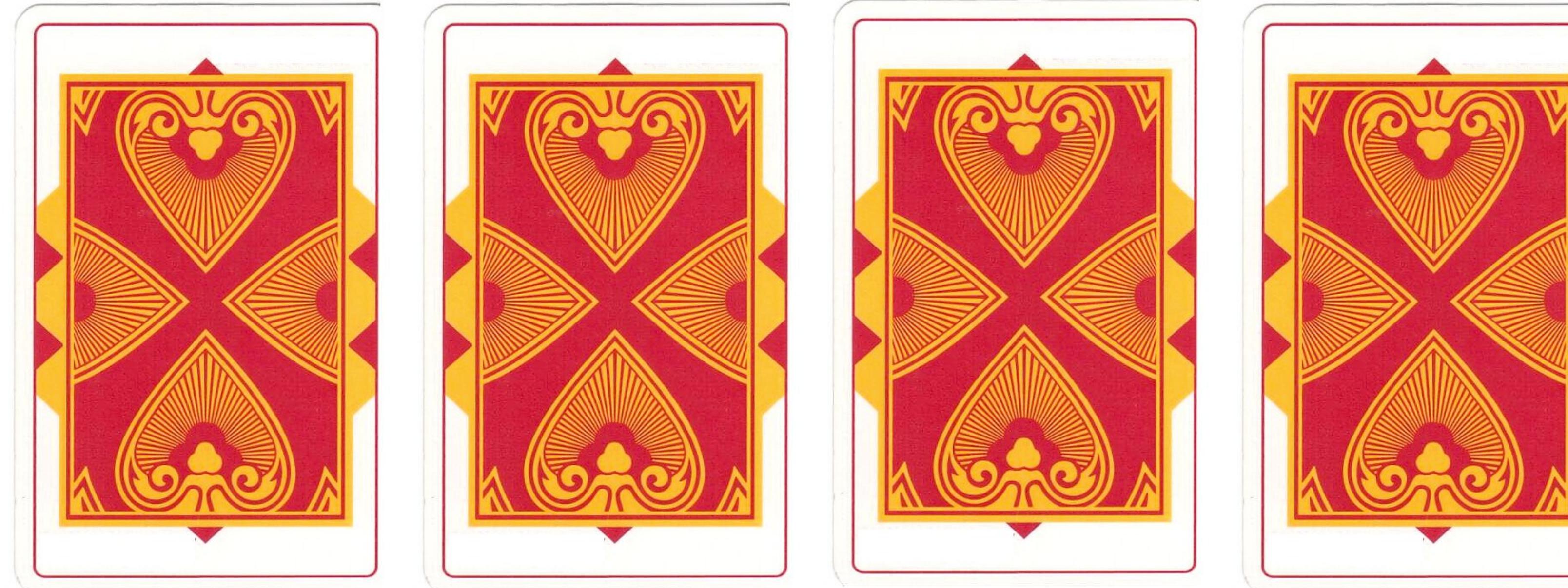
# FATO #3

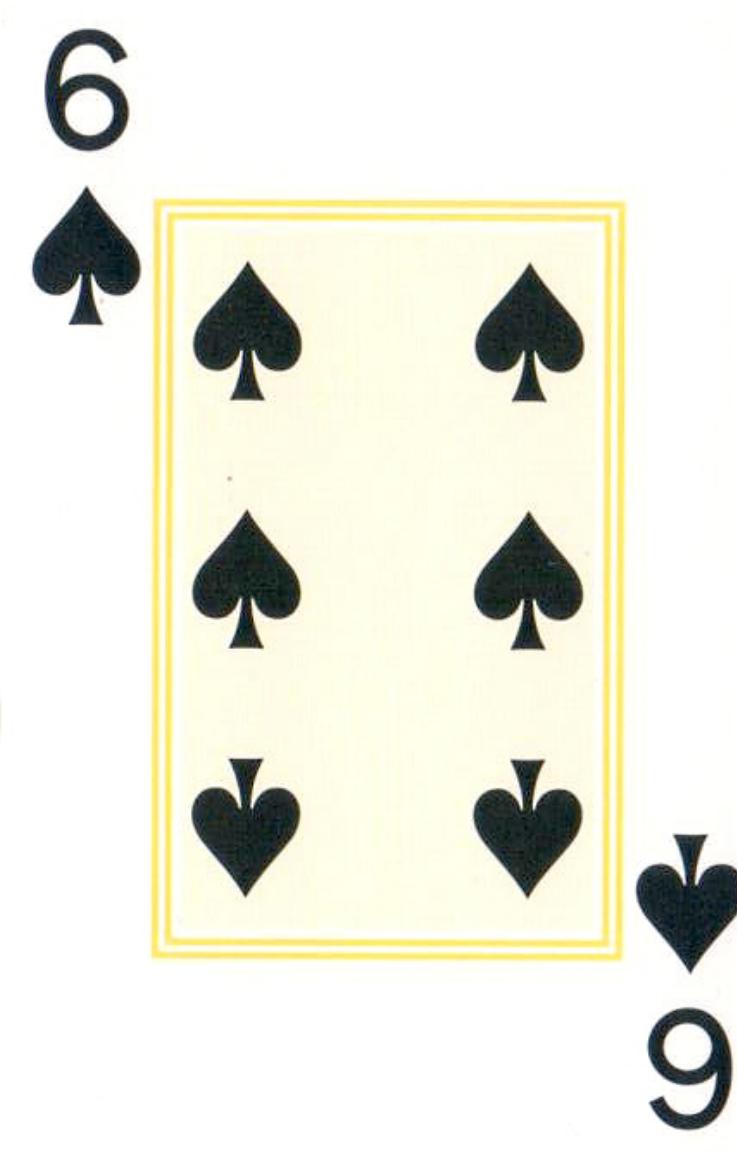
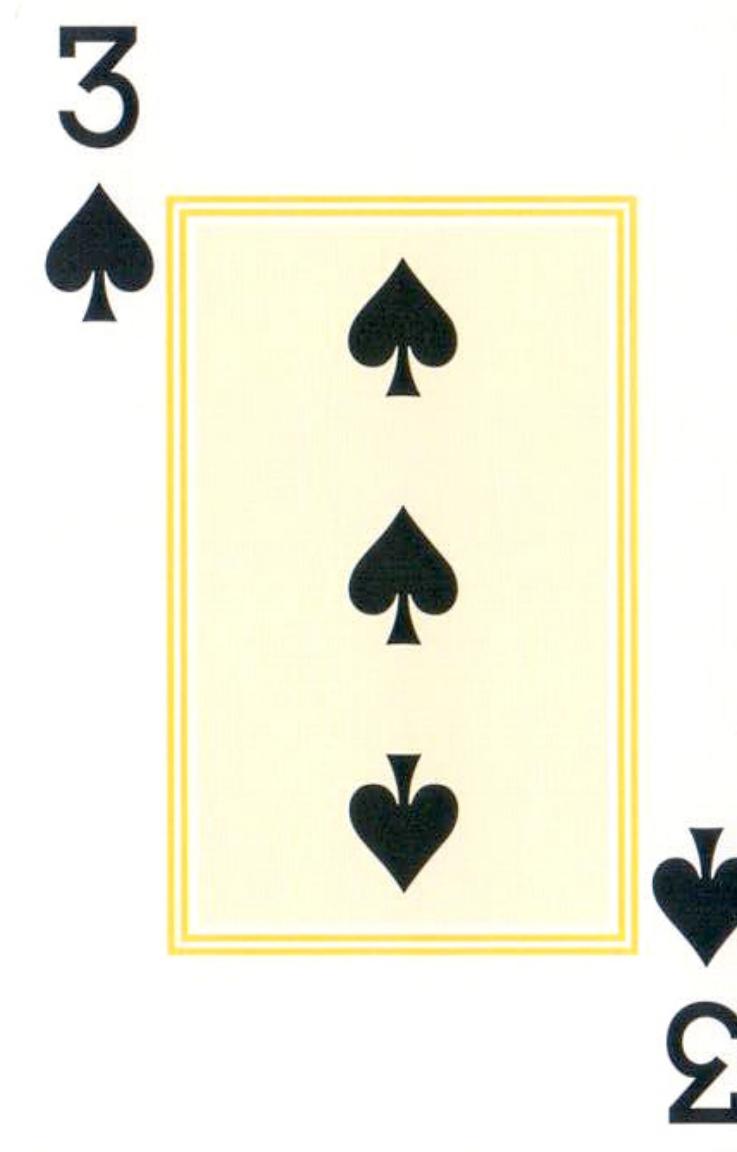
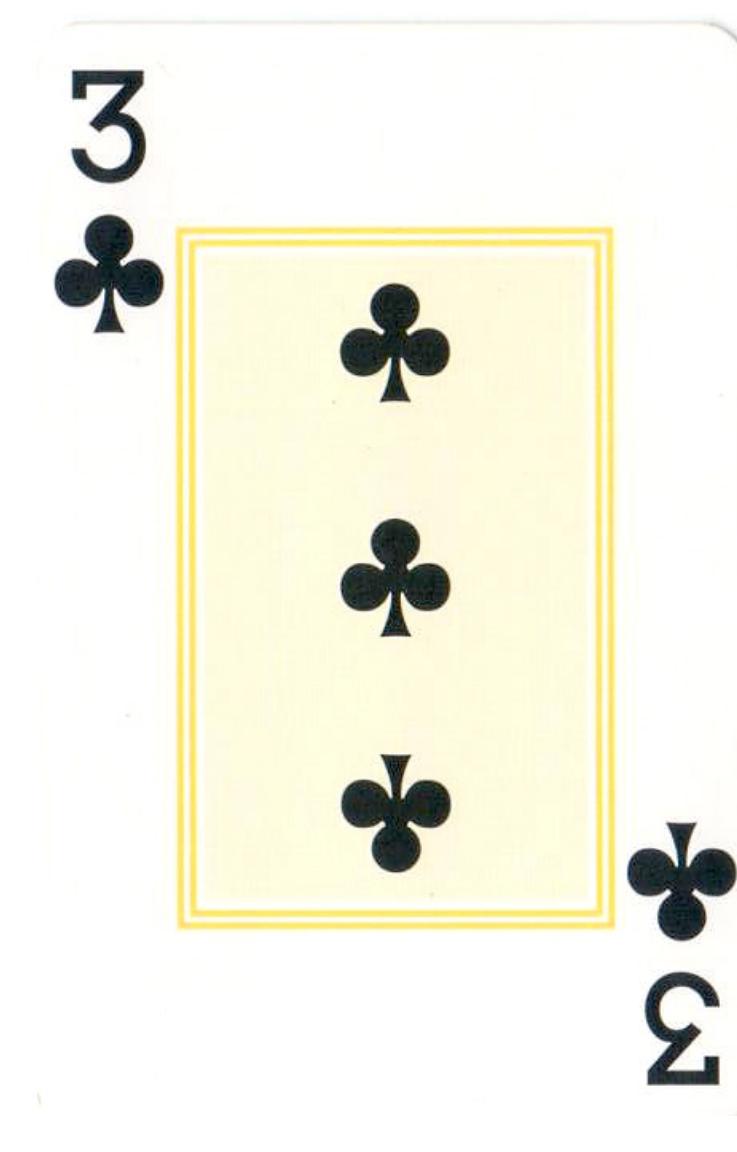
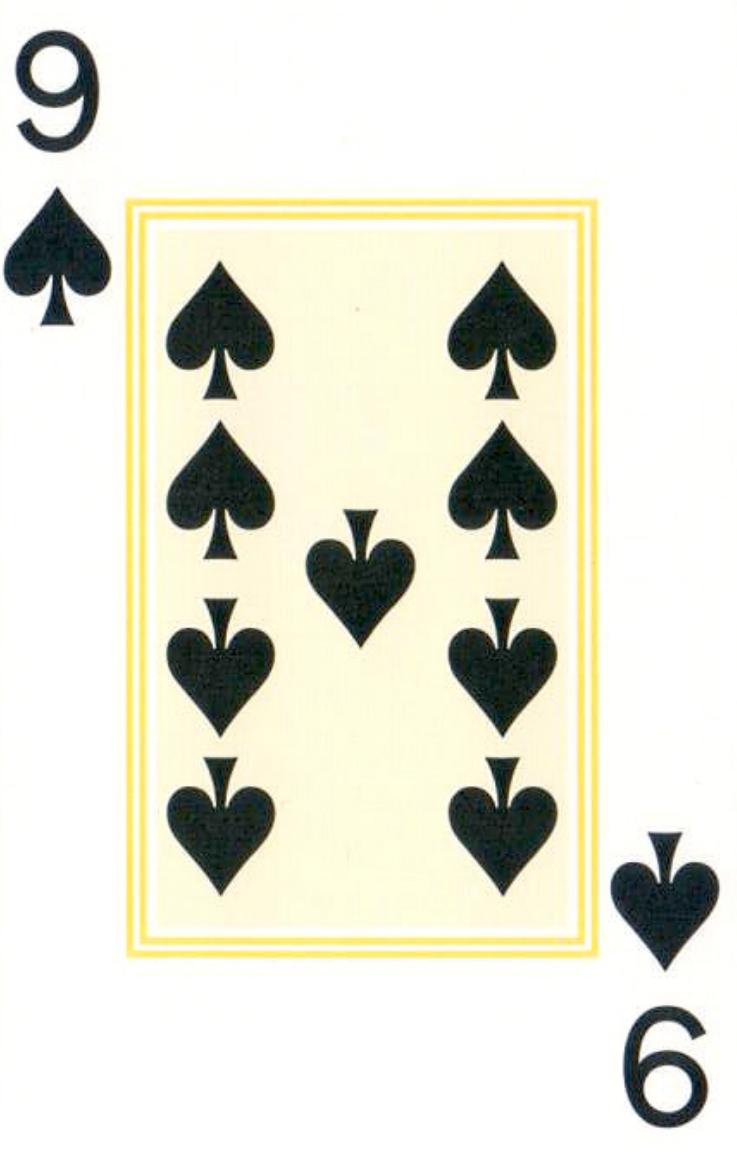
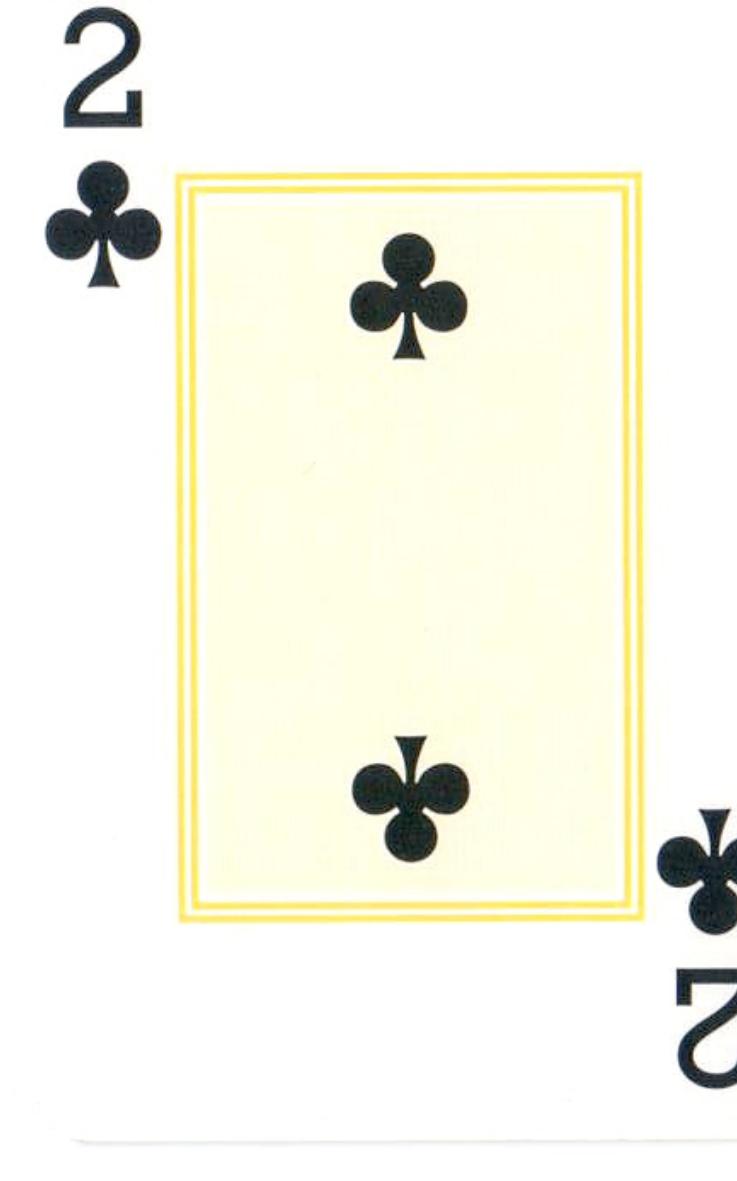
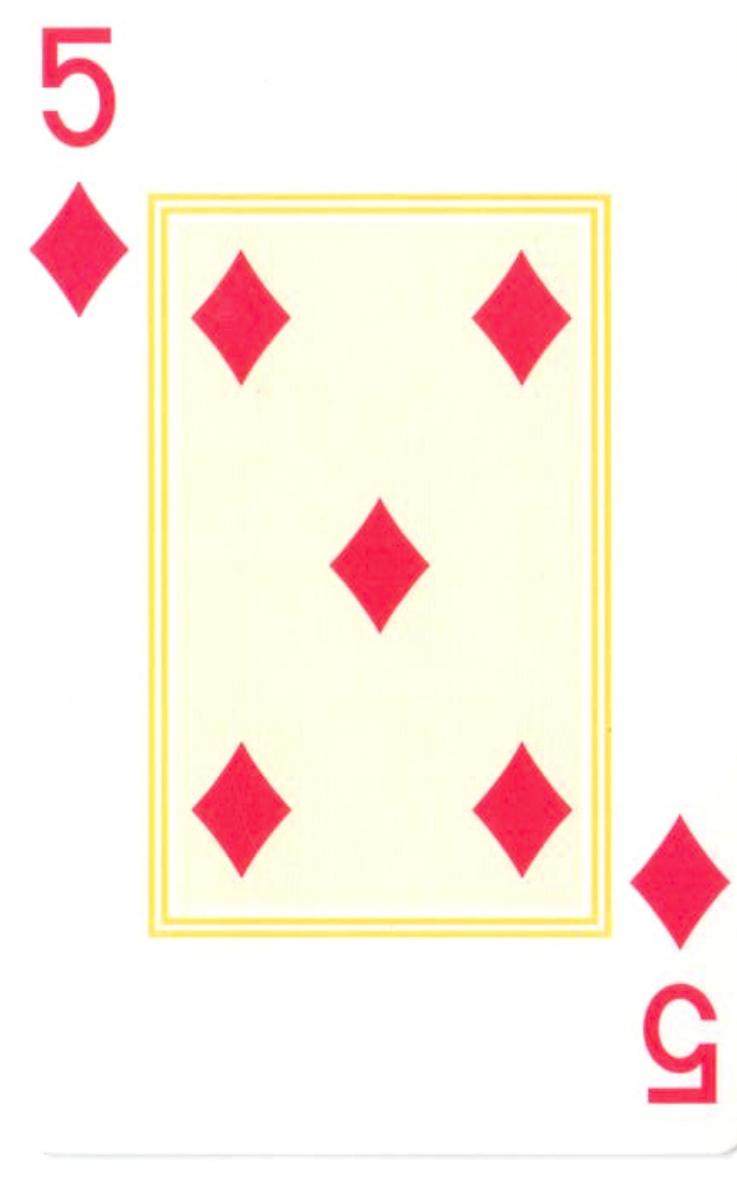
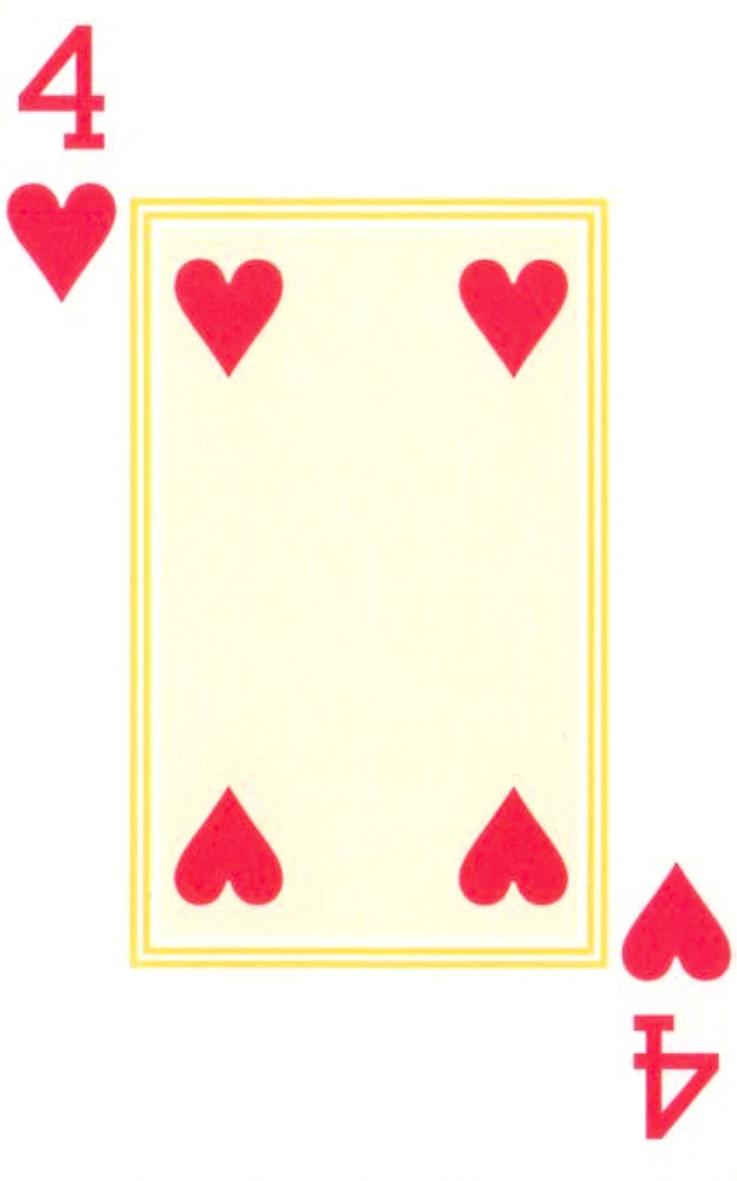
---

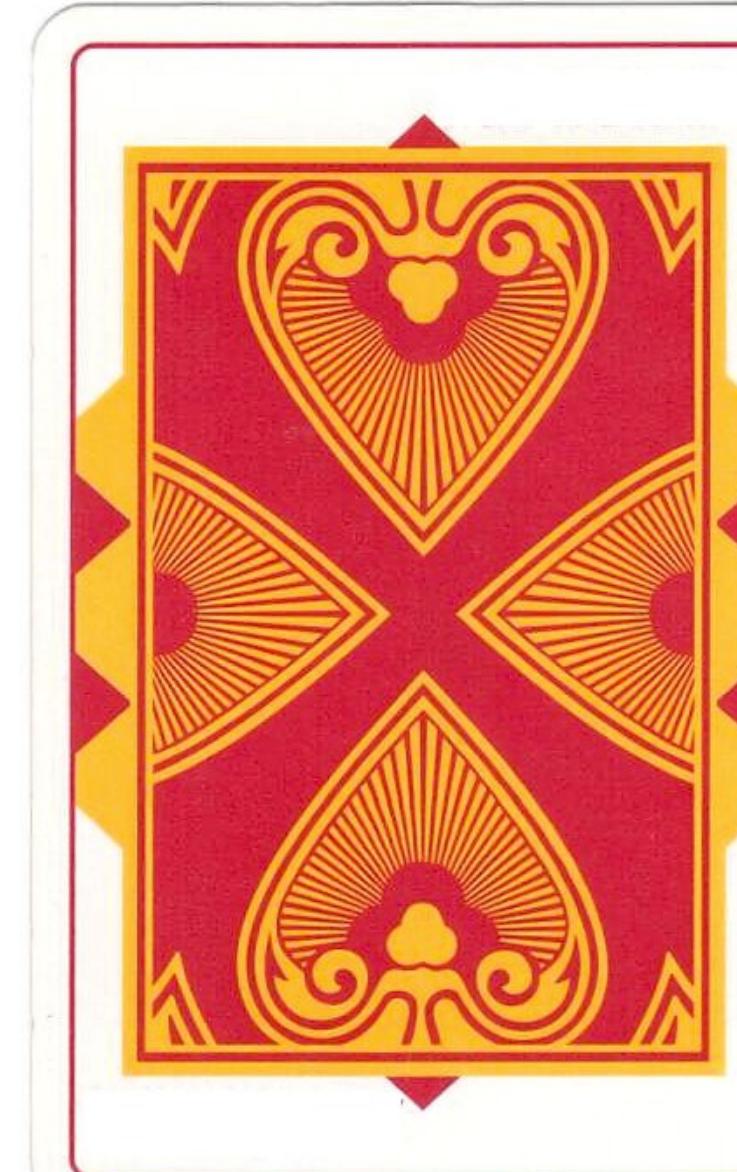
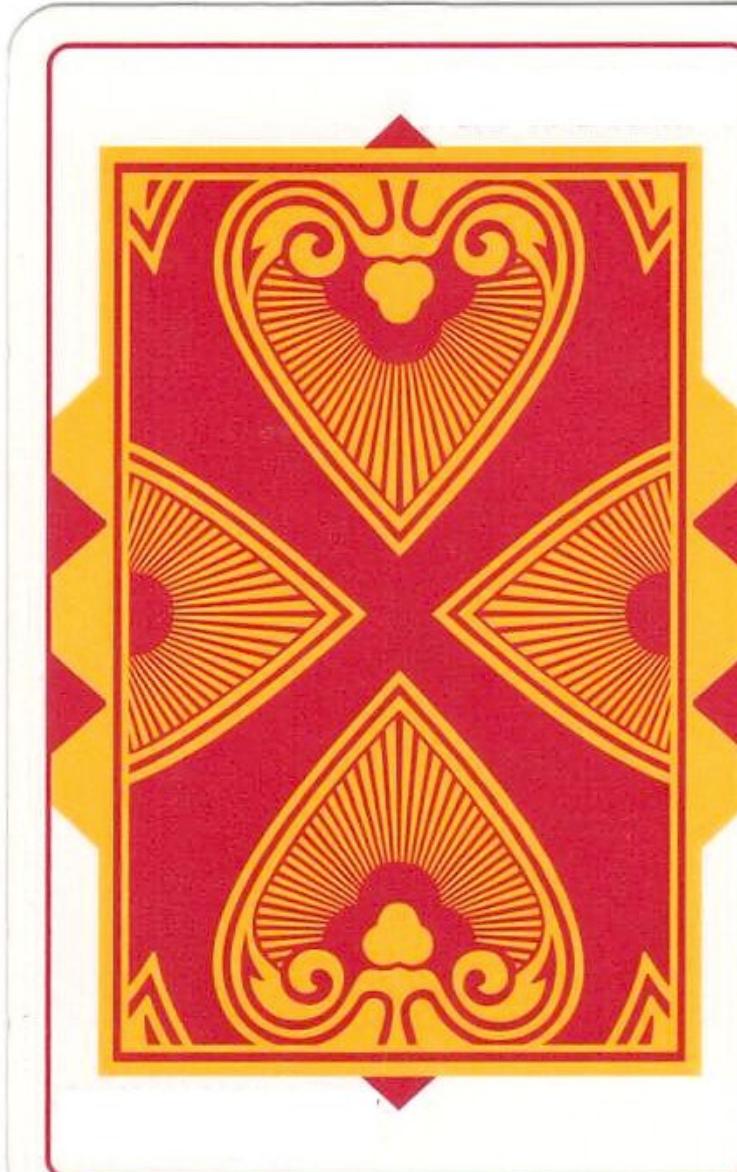
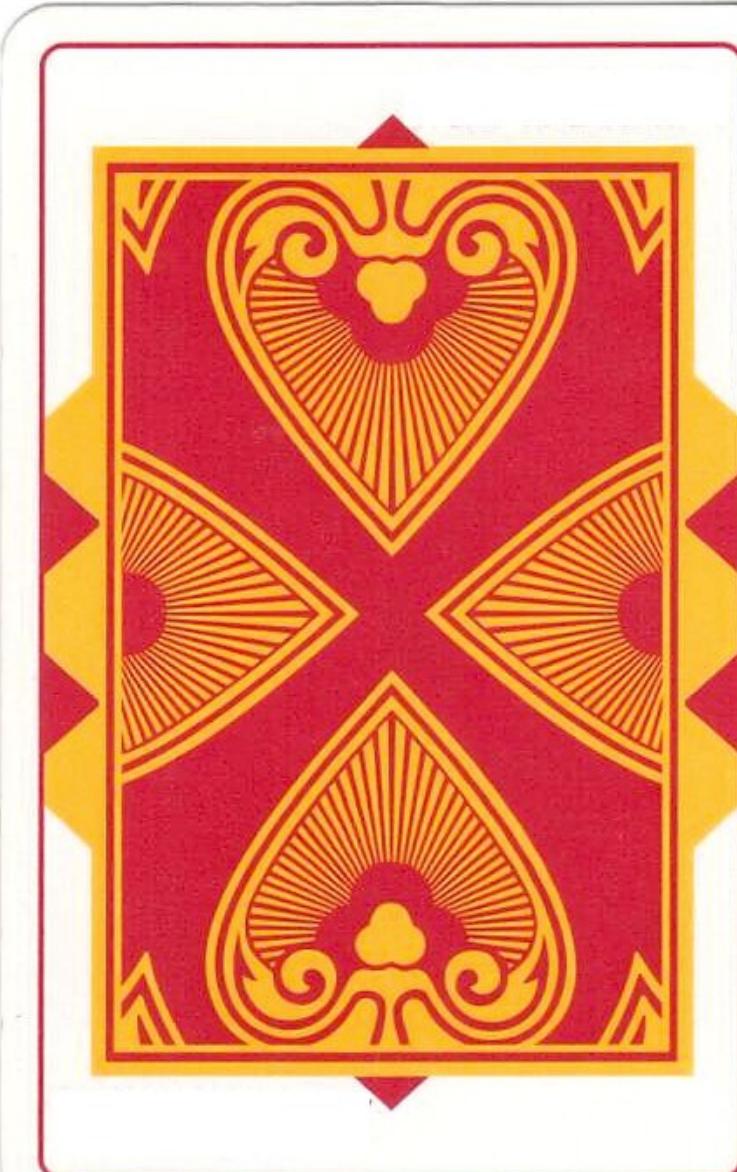
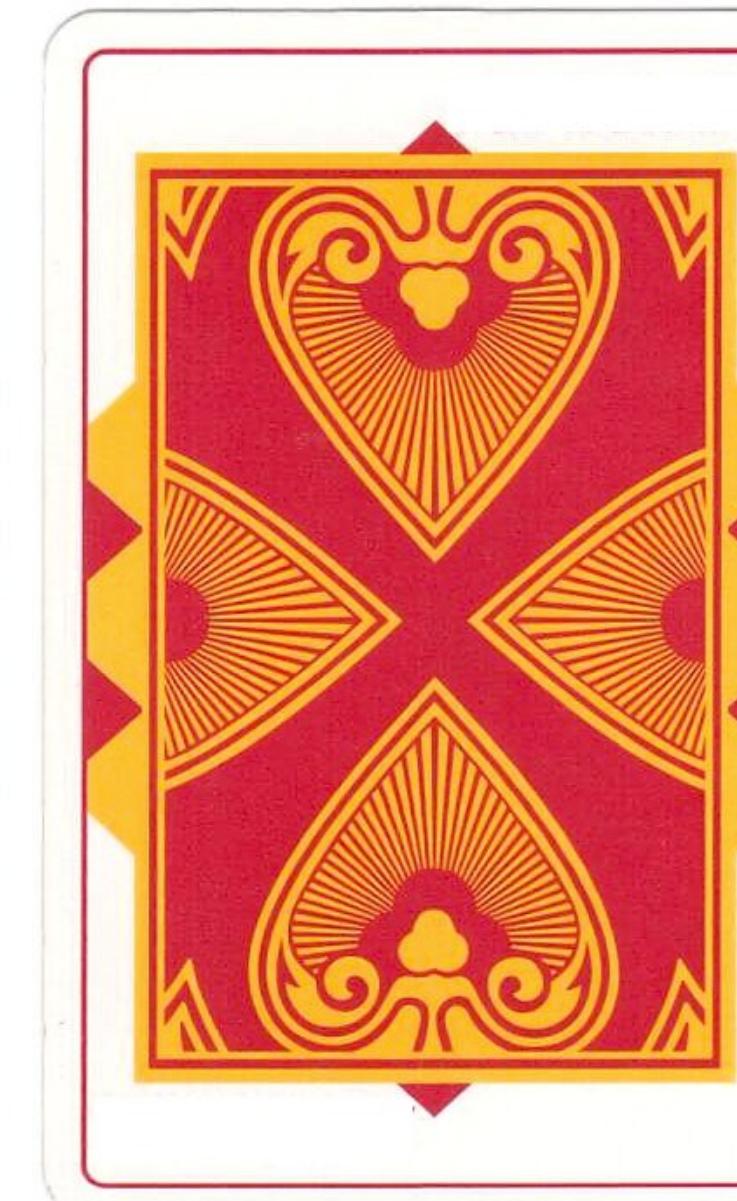
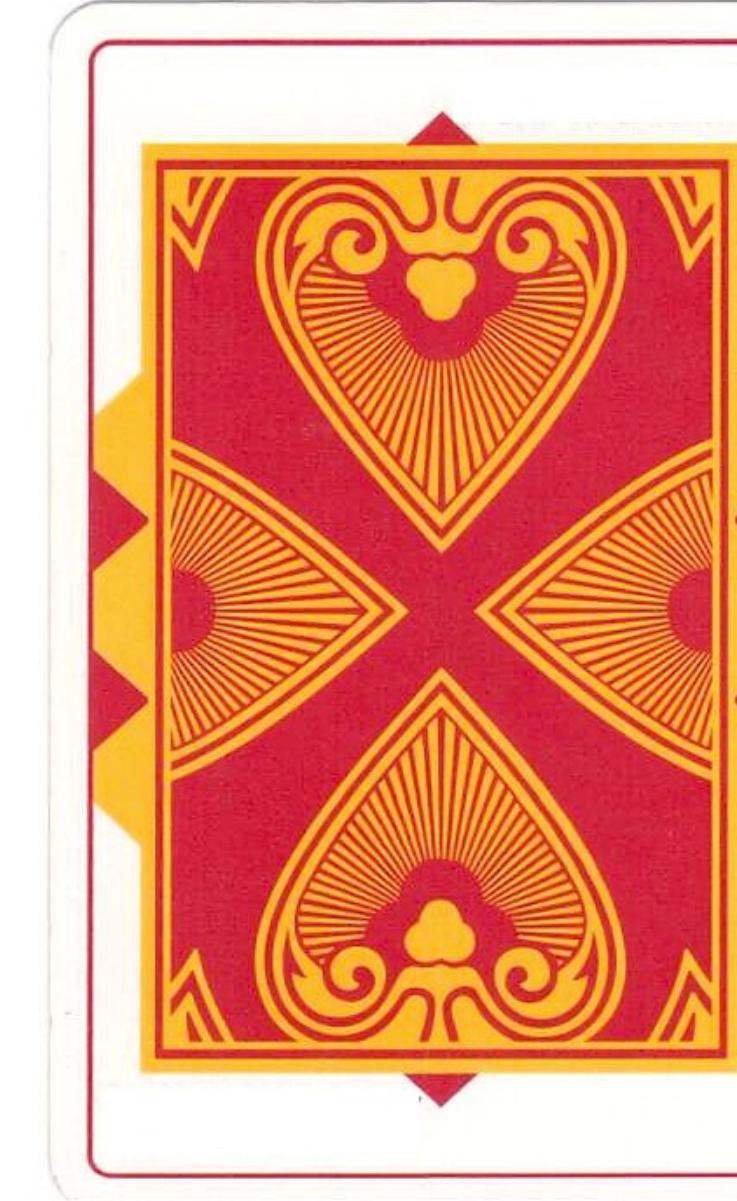
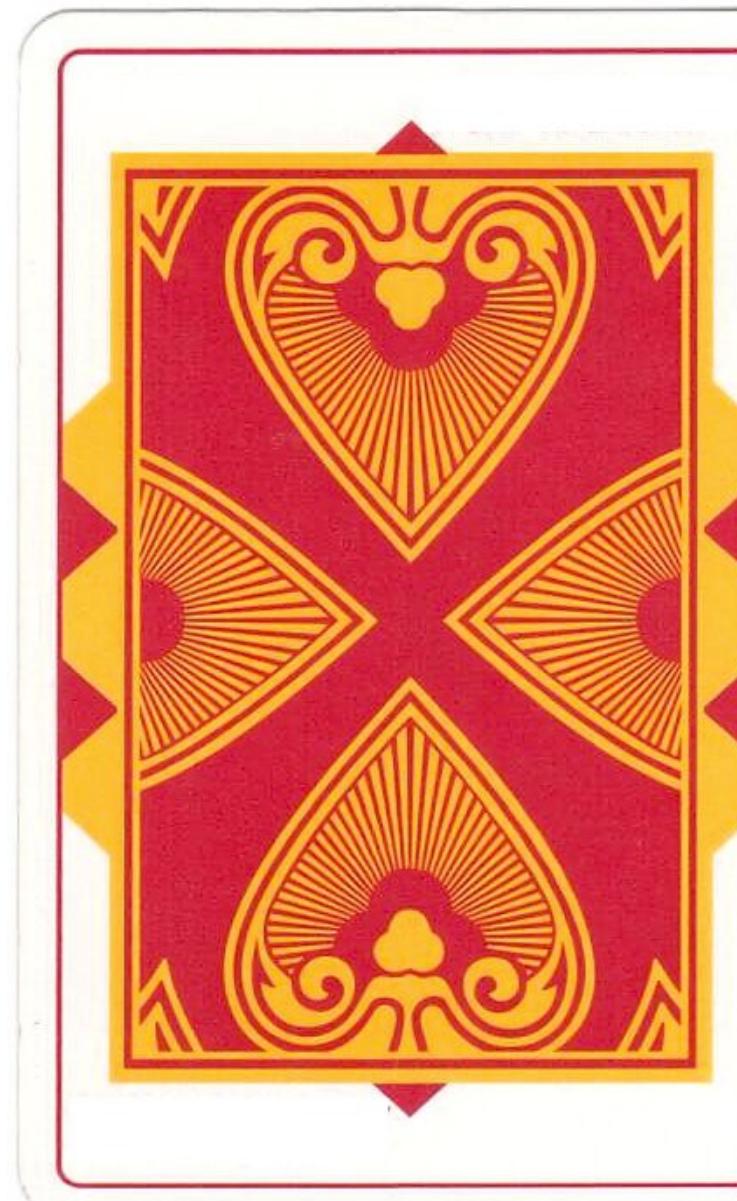
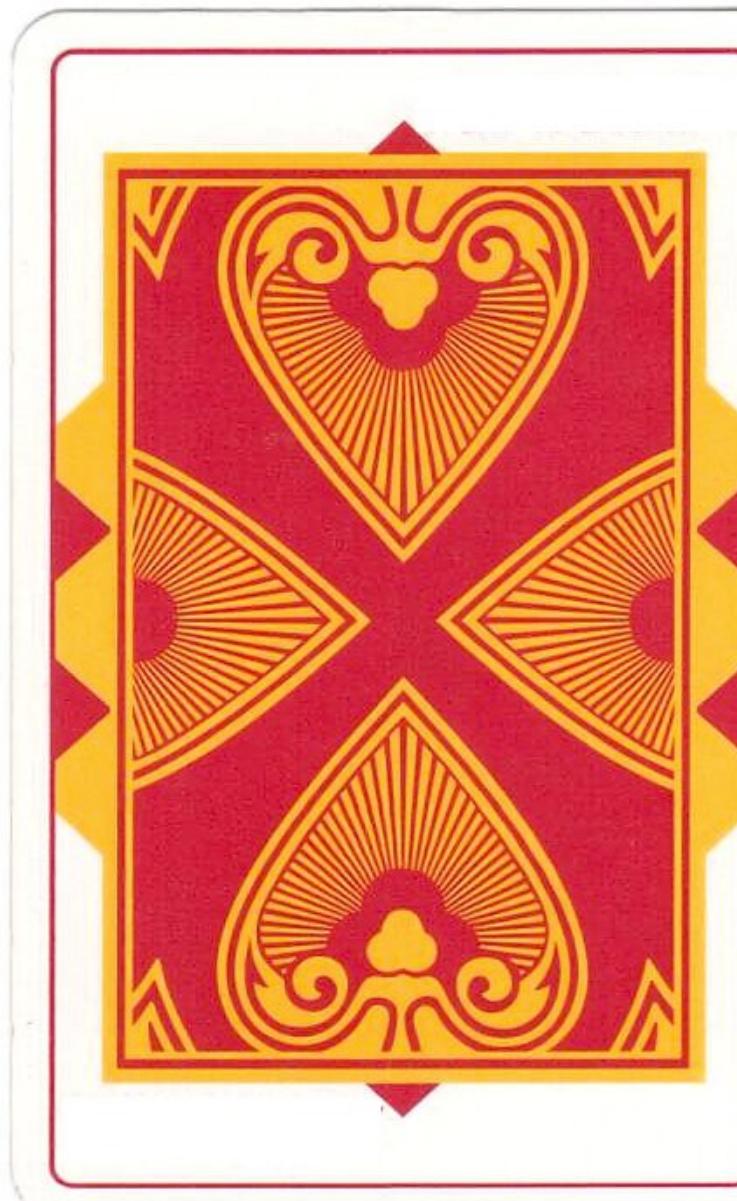


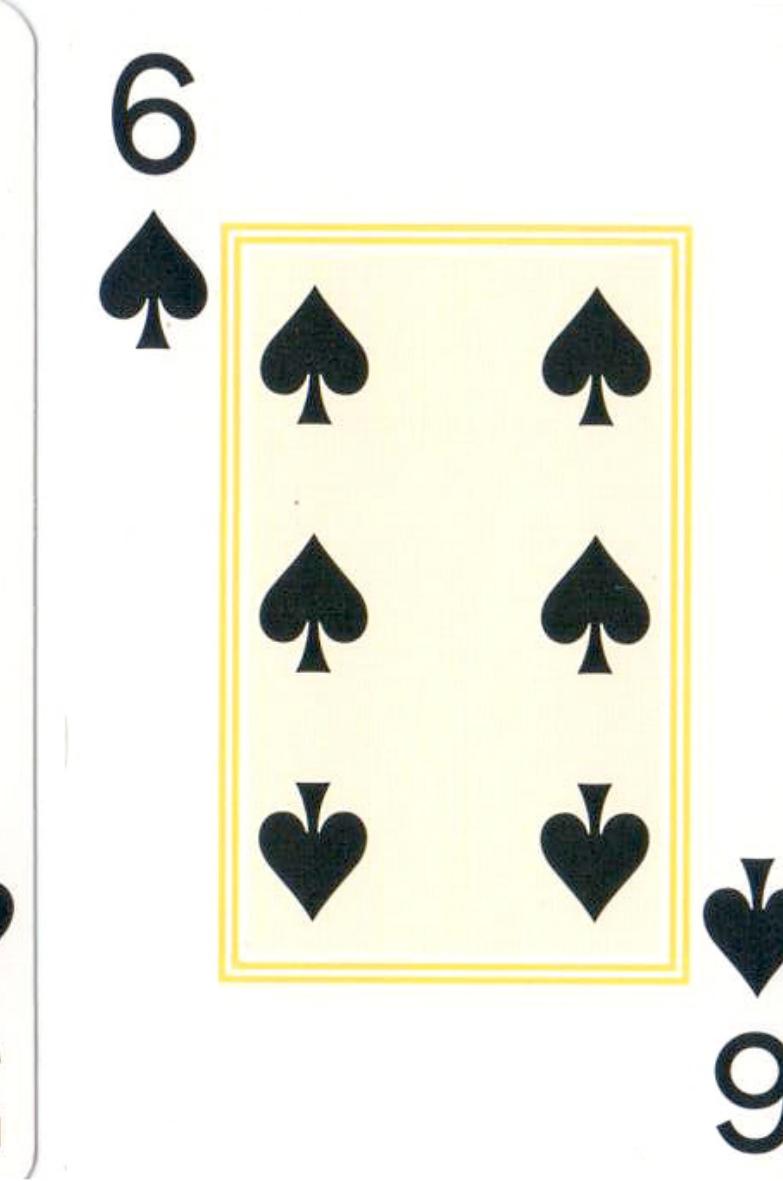
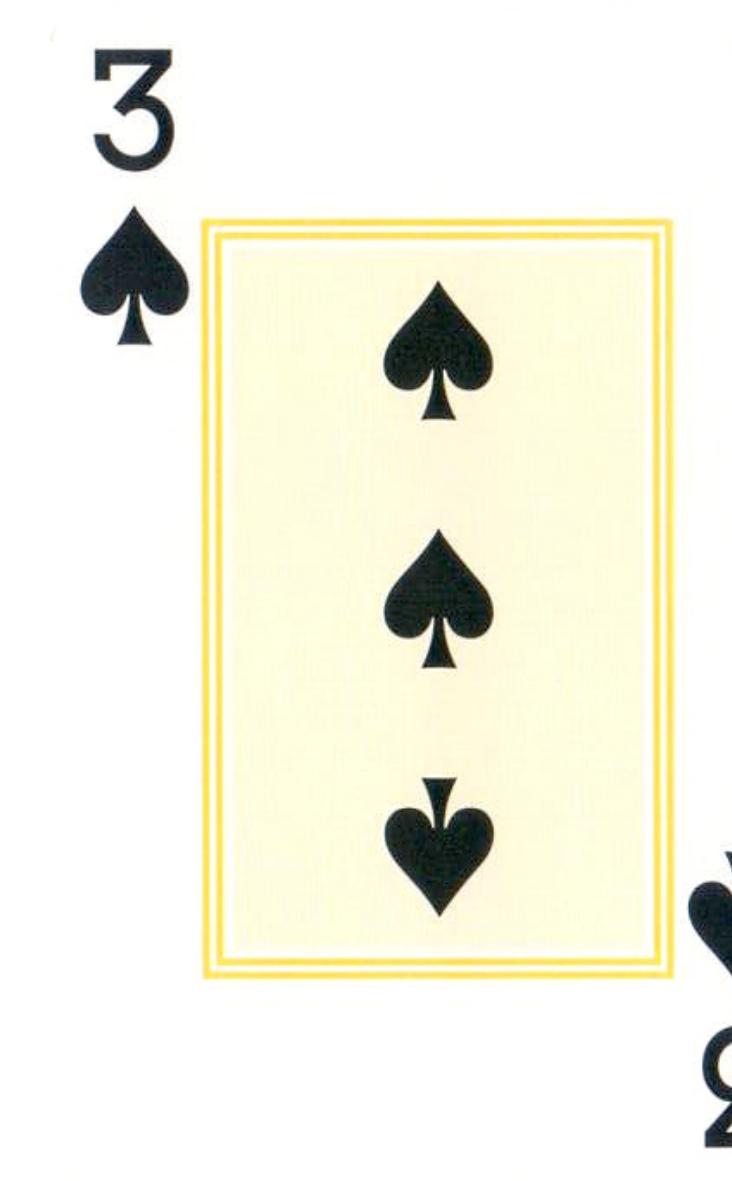
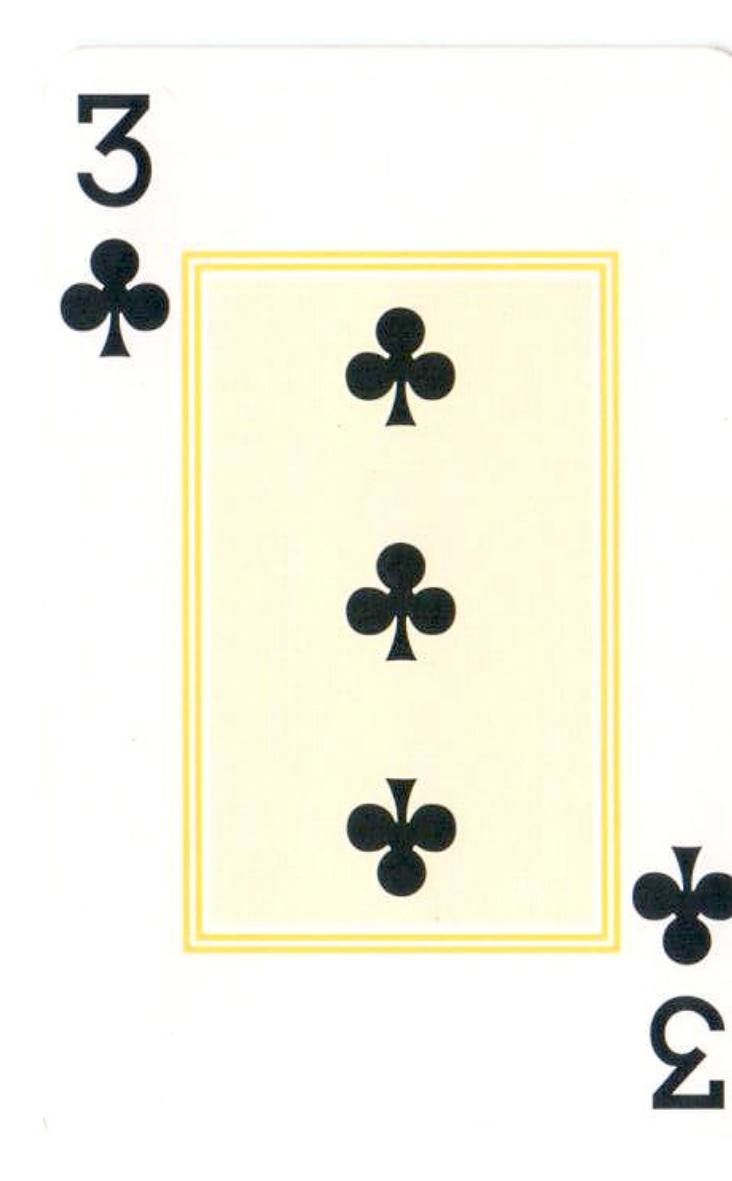
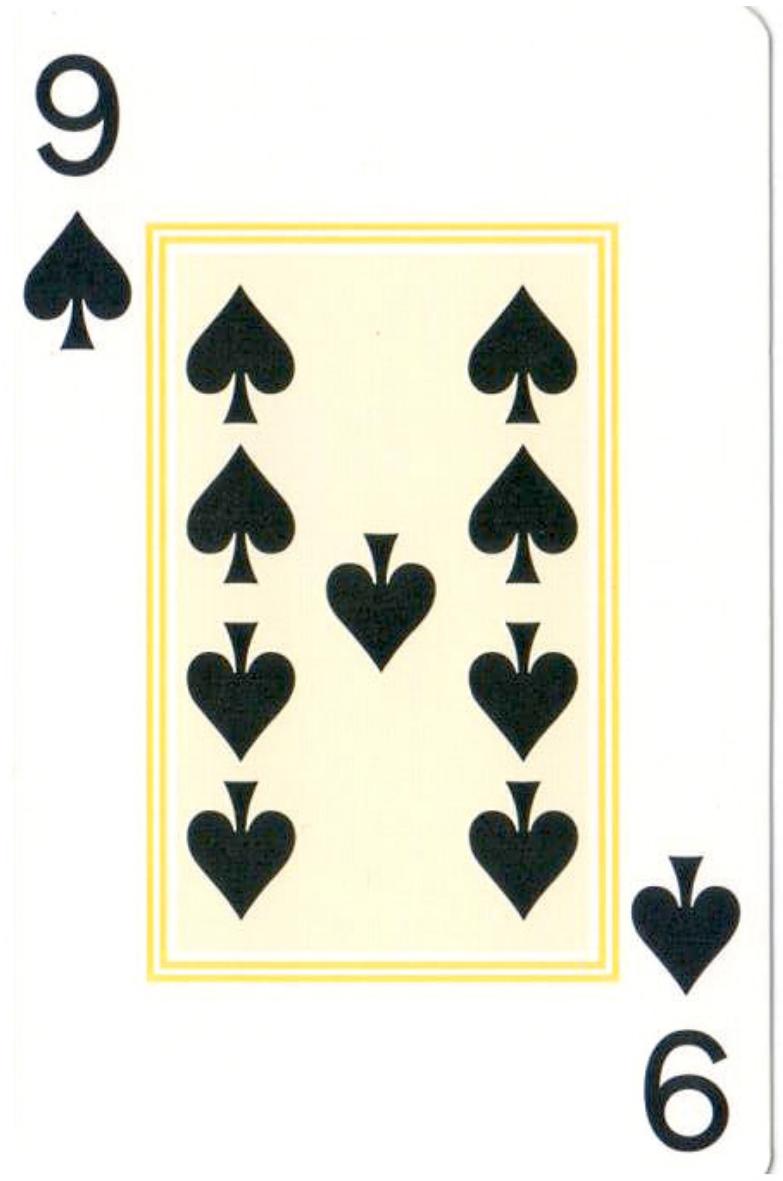
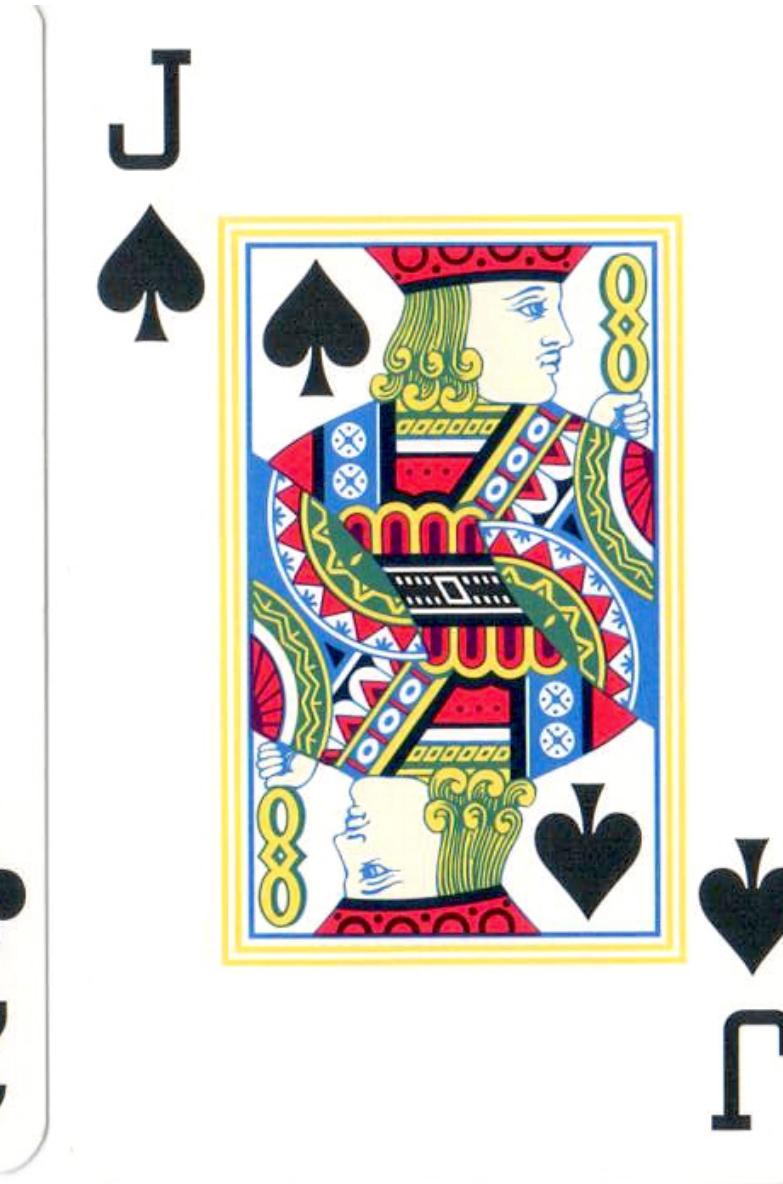
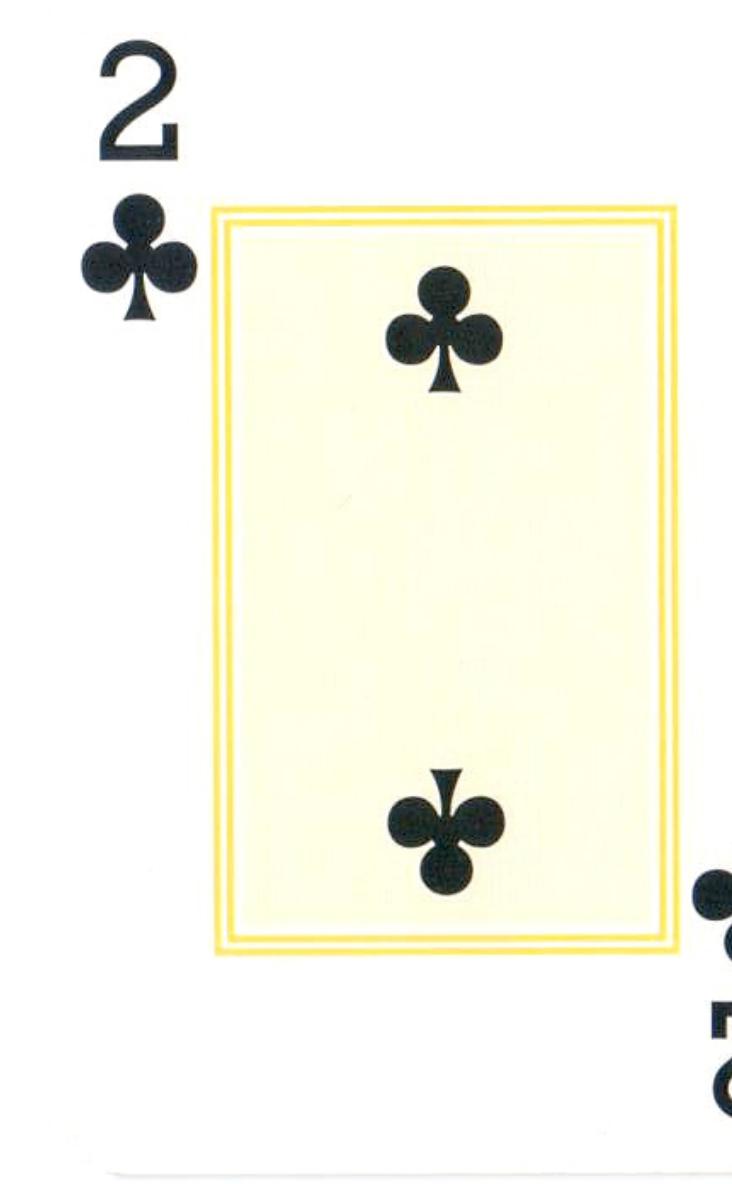
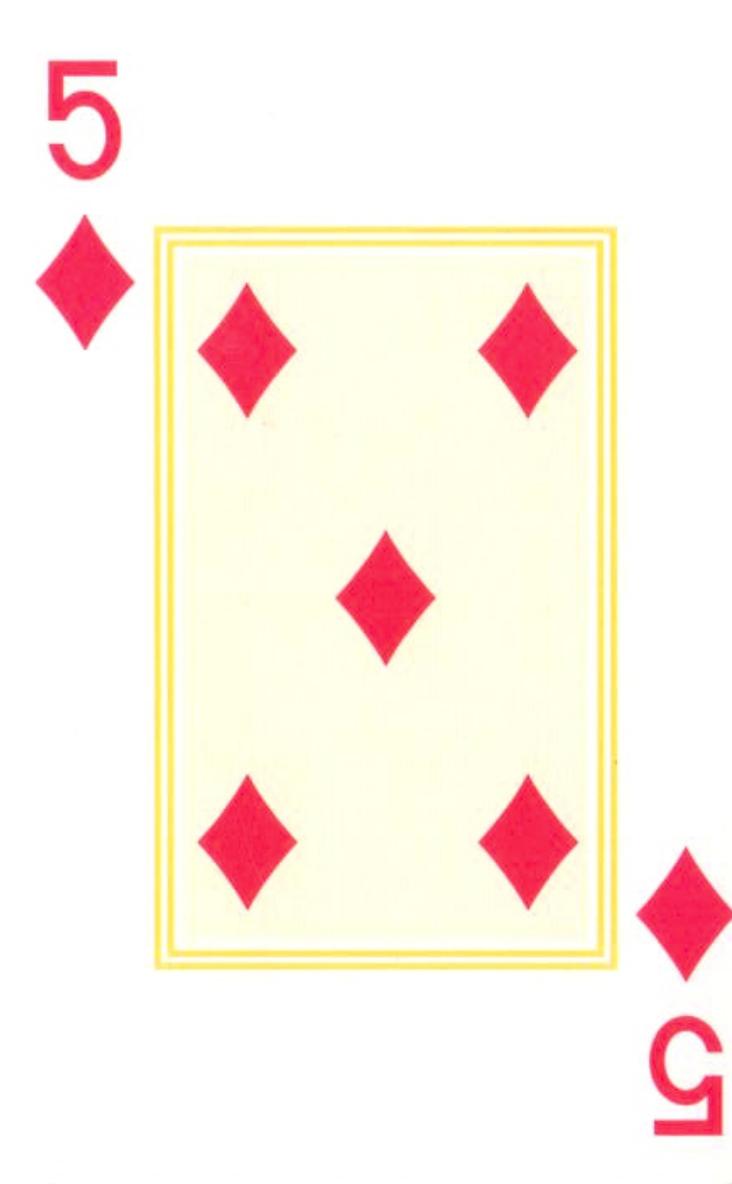
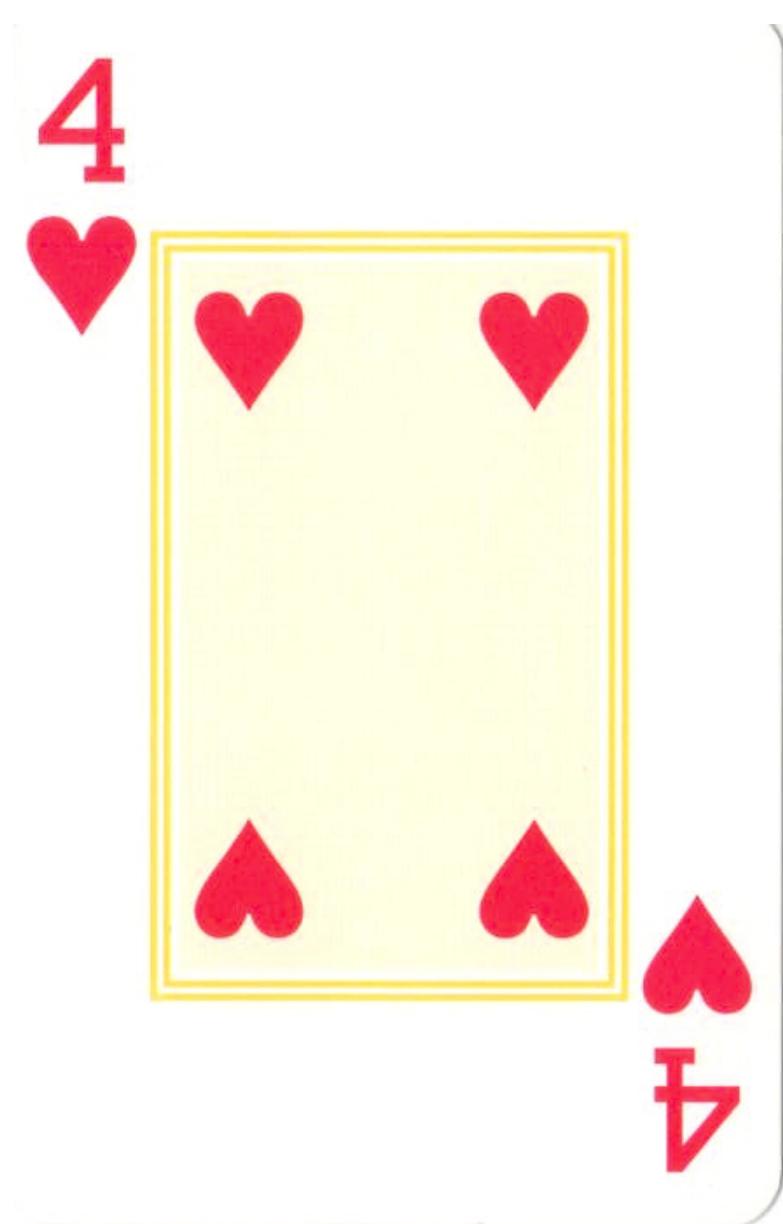
# FATO #3

---









## FATO #3

---

- A memória tem papel fundamental na cognição, mas a memória de trabalho é extremamente limitada

# FATO #3

---



# FATO #3

---

# FATO #3

---

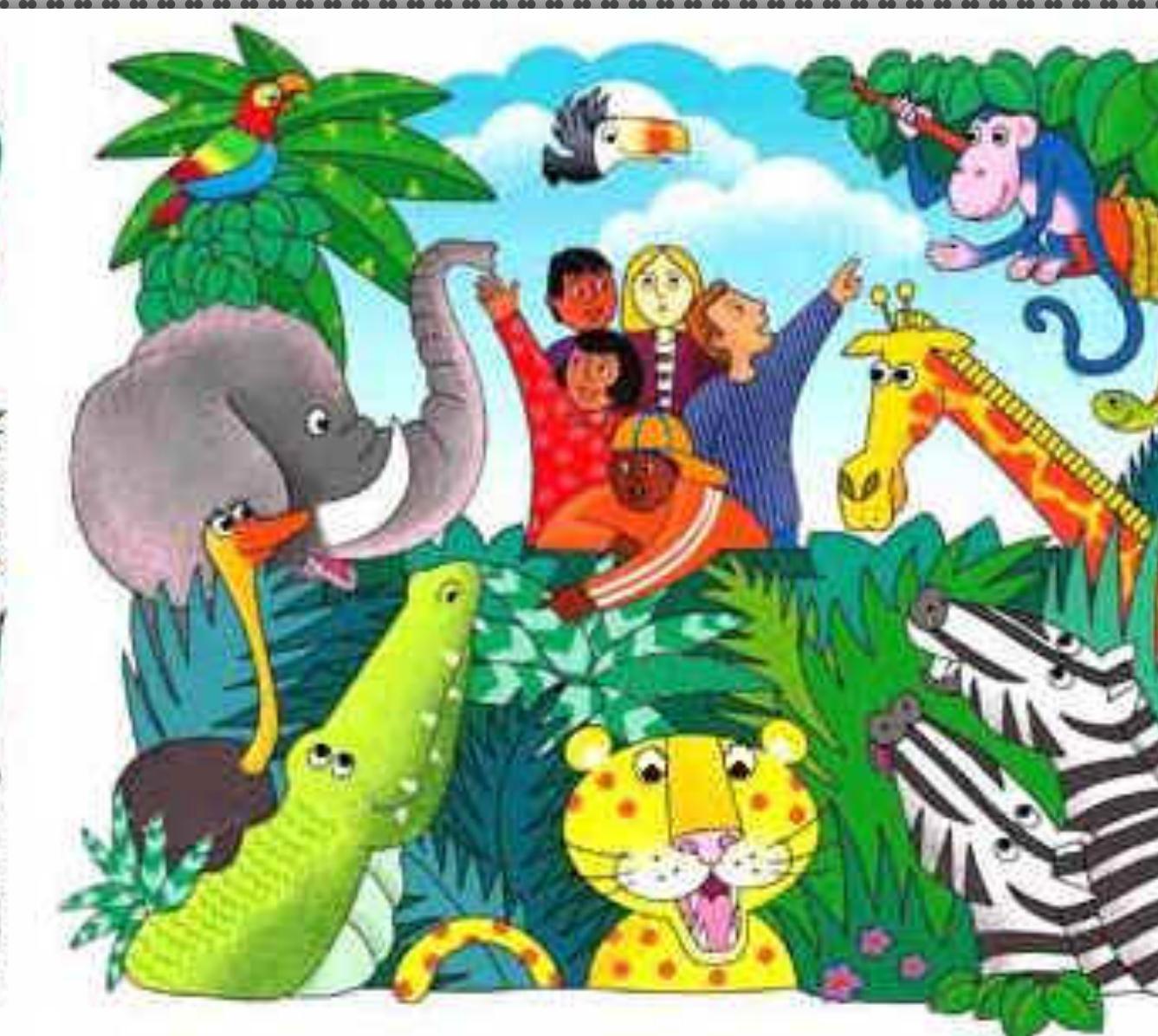
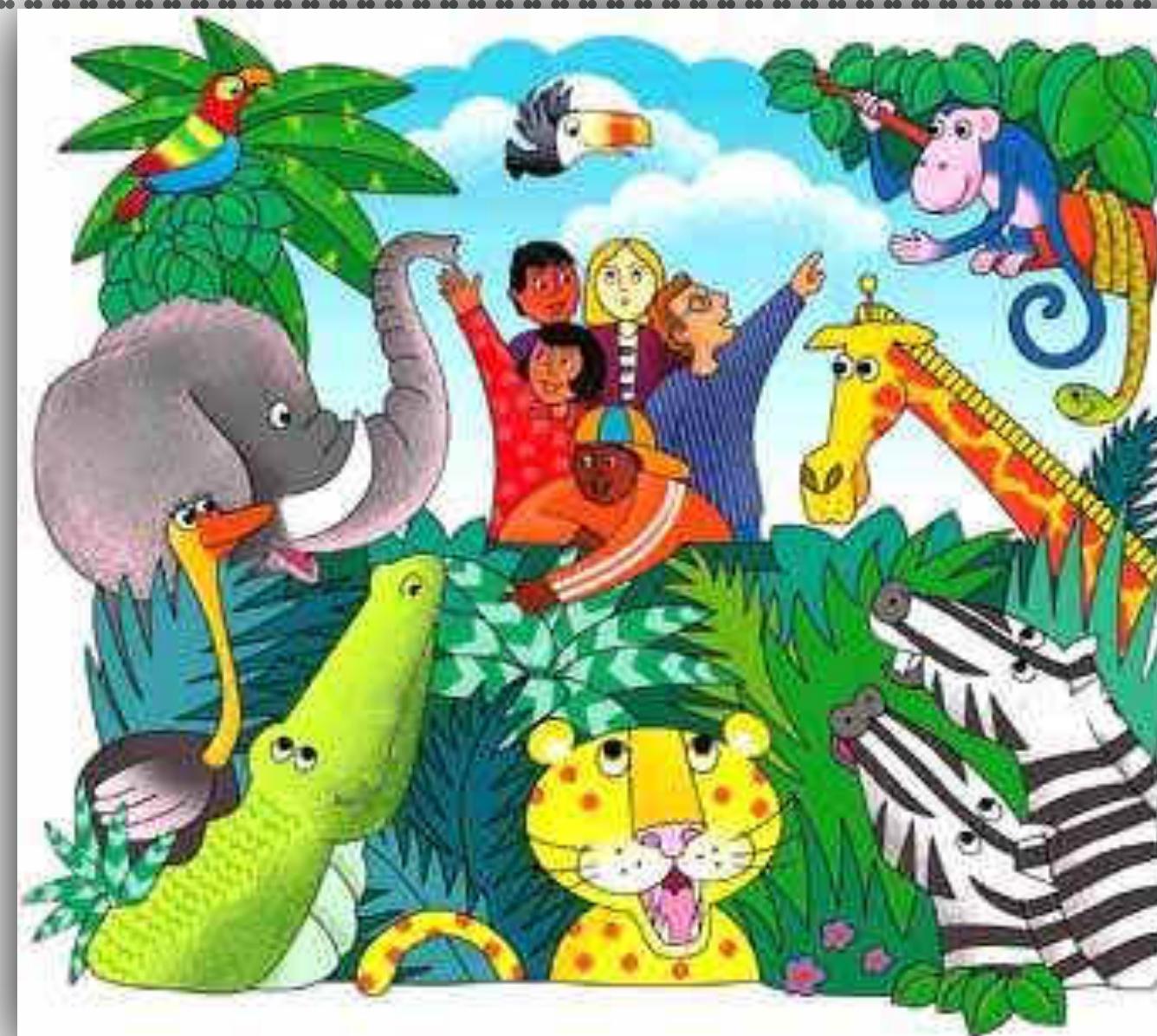


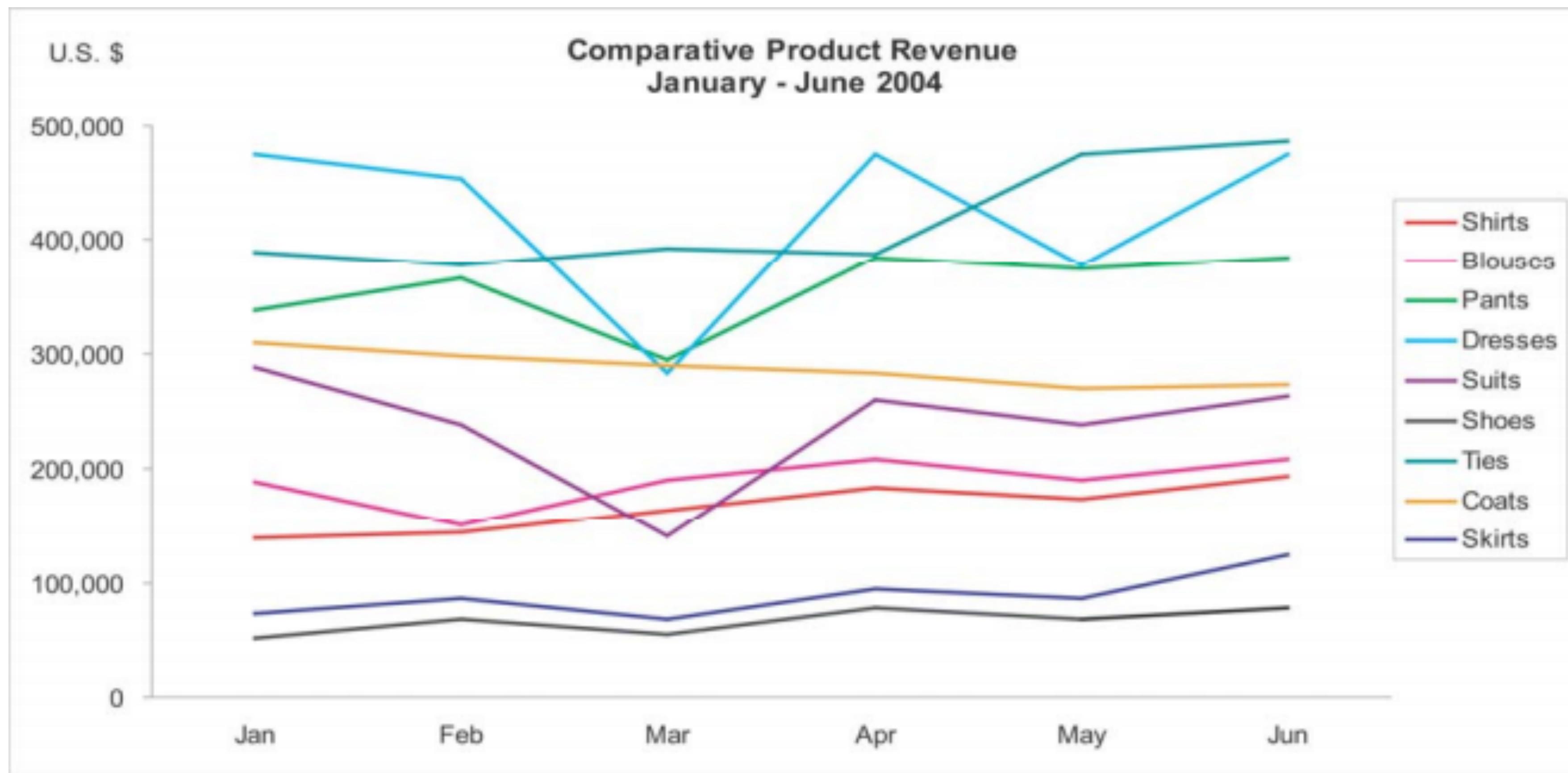
# FATO #3

---



## FATO #3





**Figure 2:** Example of a graph that exceeds the limits of short-term memory.

“

“Over history, *visual abstractions* have been developed to  
*aid thinking...*

What information visualization is really about is external cognition, that is, *how resources outside the mind can be used to boost the cognitive capabilities of the mind.*”

-Stuart Card

in Visualization: Perception for Design

Visualizações só serão **efetivas** se operarem sobre princípios que  
**respeitem como a percepção visual e a cognição funcionam.**

# Processamento pré-atentivo

---

**Profa. Dra. Raquel Minardi**  
Departamento de Ciência da Computação  
Universidade Federal de Minas Gerais

12768679489326456584791209193021483490386  
24814001480912808401209475283758237503407  
67465748572308402394083590235803275904376  
49679024376043765096730964036753067034760  
37603760934706734096709347609430697039462  
09765902347306047307603476034076034650967

127686794893264~~56~~584791209193021483490386  
2481400148091280840120947~~5~~2837~~5~~8237~~5~~03407  
6746~~5~~748~~5~~72308402394083~~5~~9023~~5~~80327~~5~~904376  
4967902437604376~~5~~0967309640367~~5~~3067034760  
37603760934706734096709347609430697039462  
0976~~5~~9023473060473076034760340760346~~5~~0967

# PROCESSAMENTO PRÉ-ATENTIVO

---

- Devemos codificar os símbolos de tal forma que eles possam ser visualmente identificados mesmo após uma breve exposição
- Certas formas e cores se destacam dos seus contextos

# PROCESSAMENTO PRÉ-ATENTIVO

---

- O mecanismo que explica este destaque é chamado **processamento pré-atentivo** (ocorre logicamente antes da atenção consciente)
- Processamento pré-atentivo determina quais objetos visuais são oferecidos à nossa atenção

# ATRIBUTOS PRÉ-ATENTIVOS

---

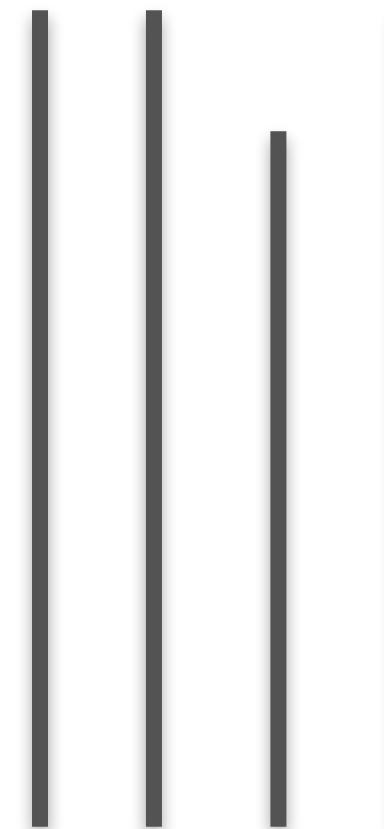
- São atributos percebidos anteriormente à atenção consciente
- Estes atributos são usados com diferentes propósitos em visualizações
- Muitas vezes são usados para diferenciar conjuntos de objetos visuais para diferentes grupos categóricos

# ATRIBUTOS PRÉ-ATENTIVOS

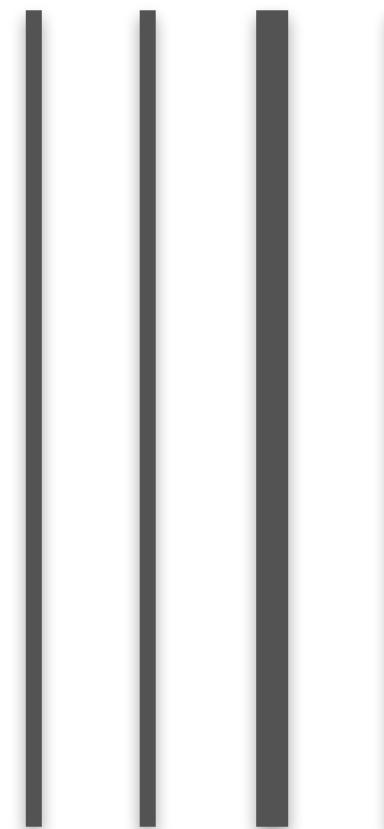
---

## FORMAS

Comprimento



Espessura



Orientação

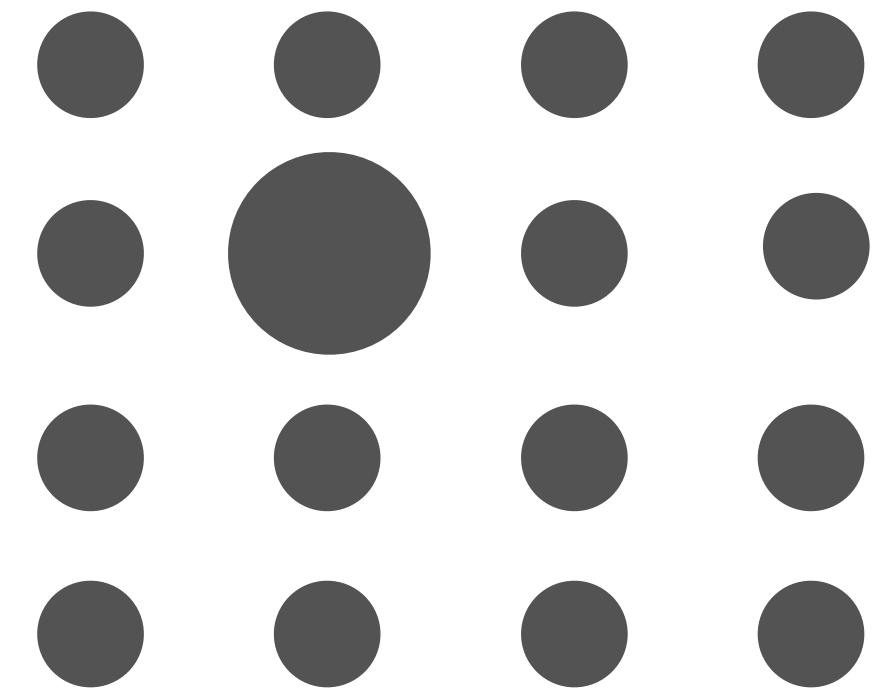


# ATRIBUTOS PRÉ-ATENTIVOS

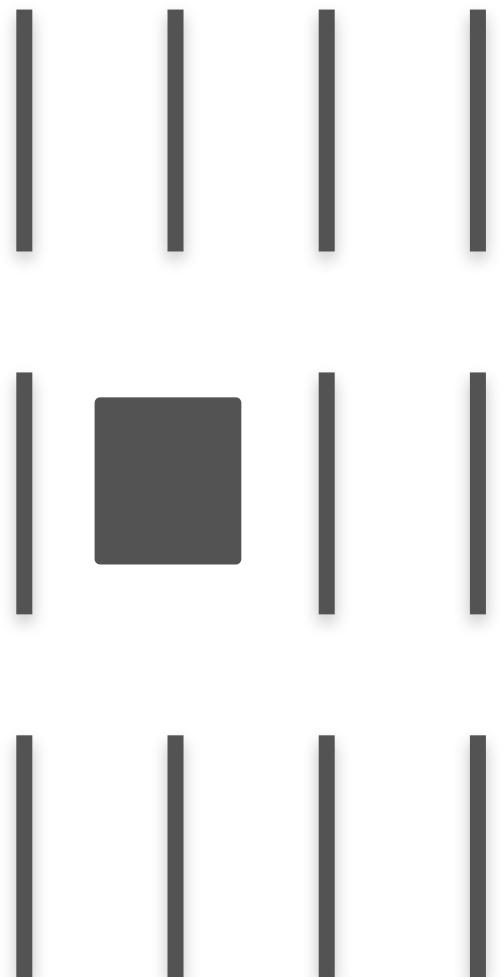
---

## FORMAS

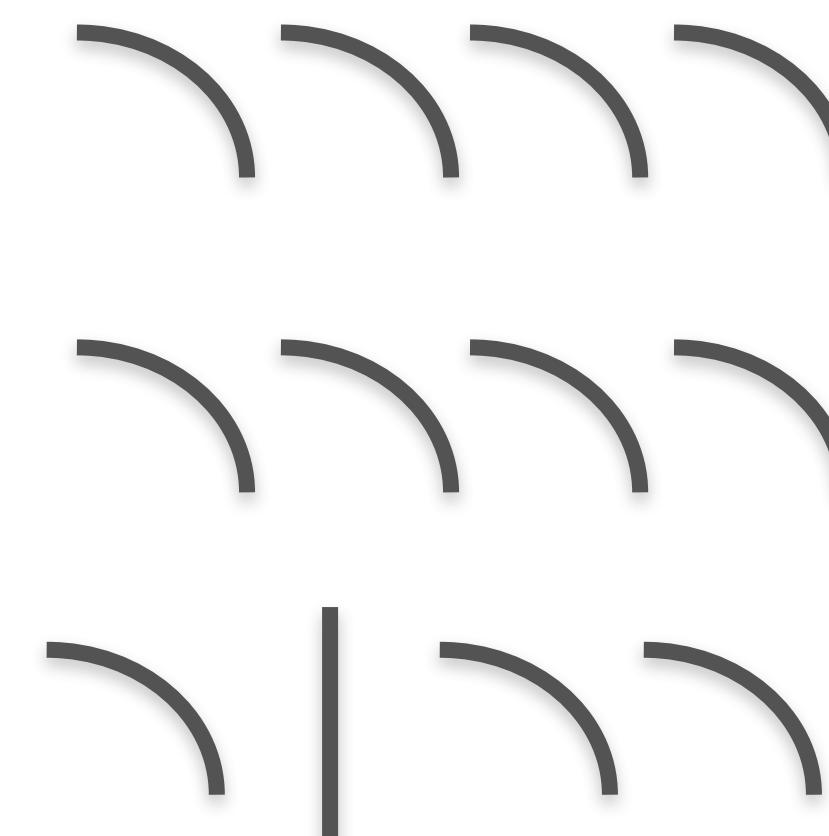
Tamanho



Forma



Curvatura

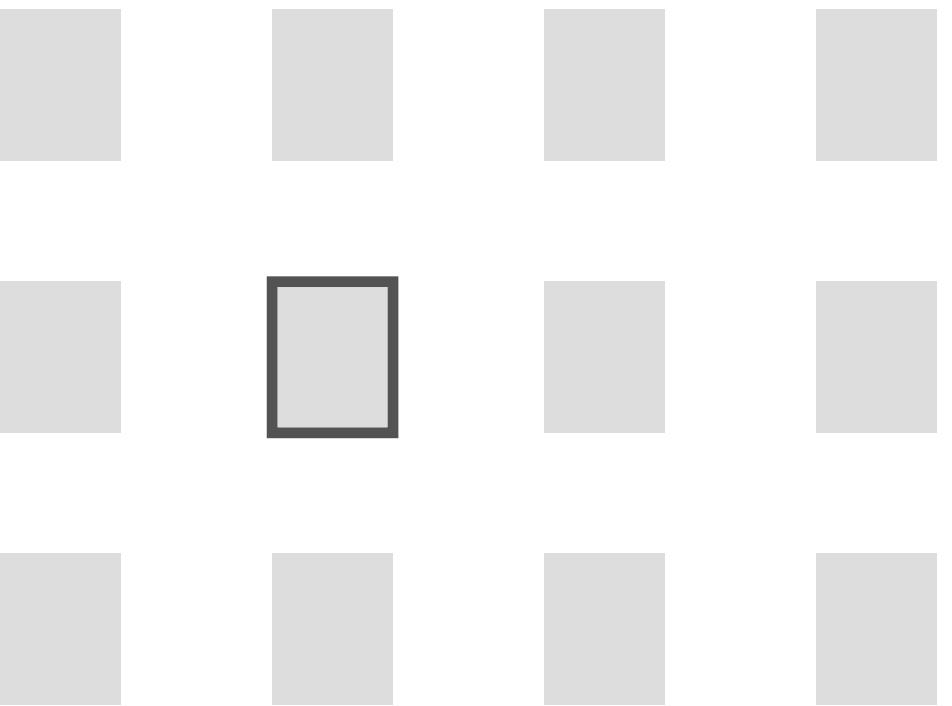


# ATRIBUTOS PRÉ-ATENTIVOS

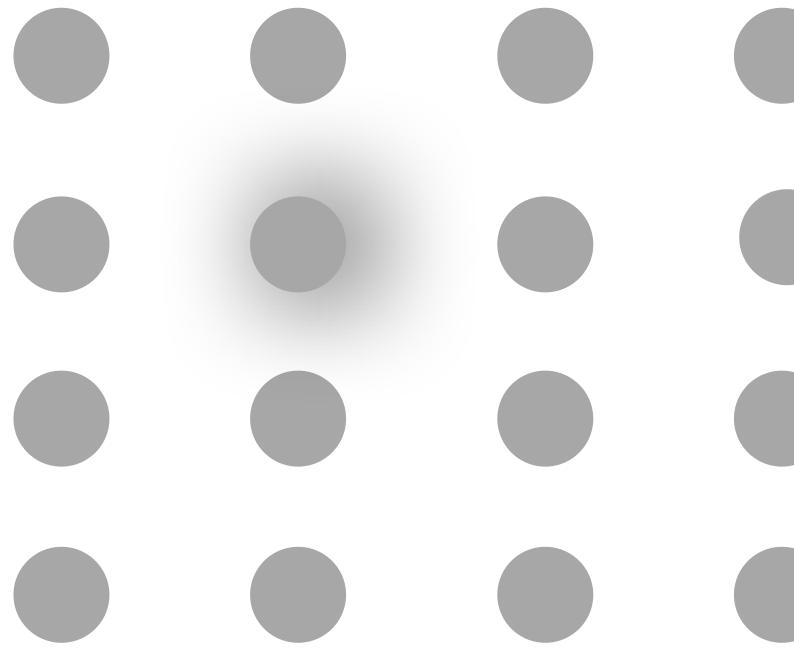
---

## FORMAS

Contorno



Sombra

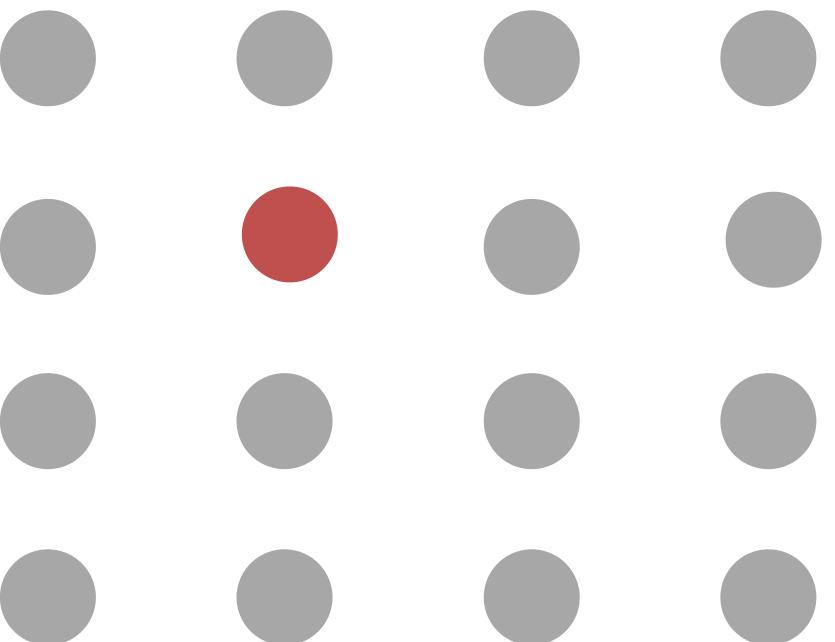


# ATRIBUTOS PRÉ-ATENTIVOS

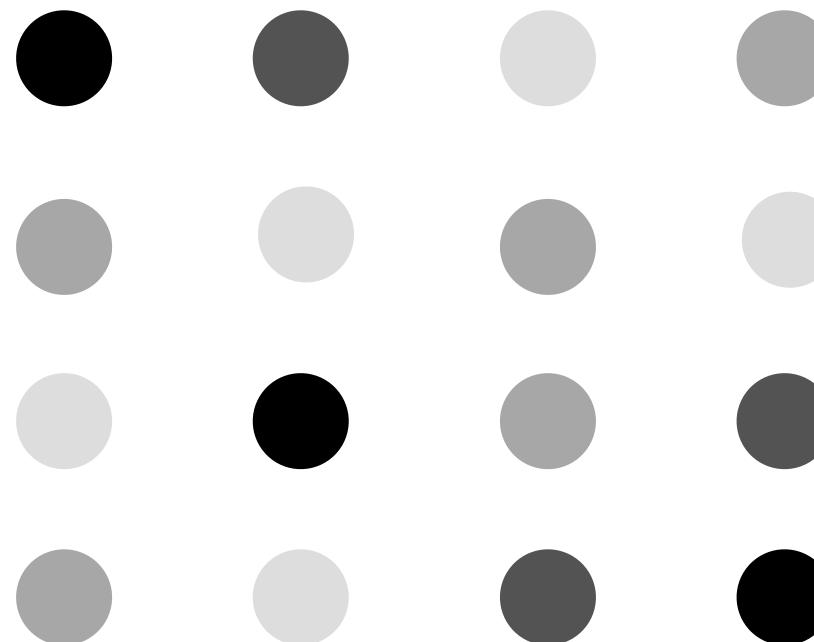
---

## CORES

Cor (matiz/hue)



Intensidade

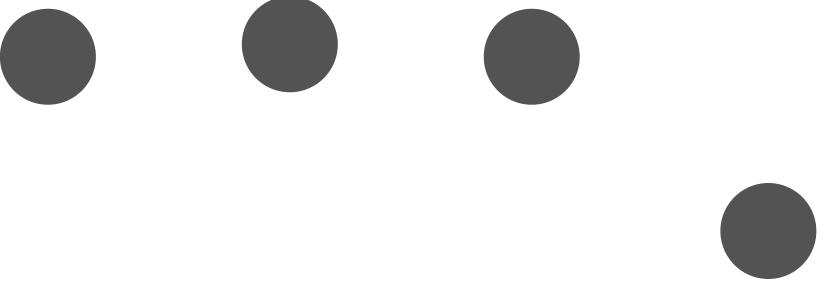


# ATRIBUTOS PRÉ-ATENTIVOS

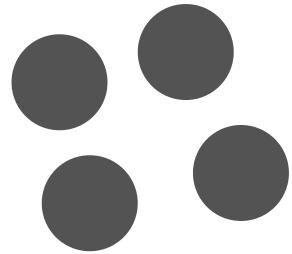
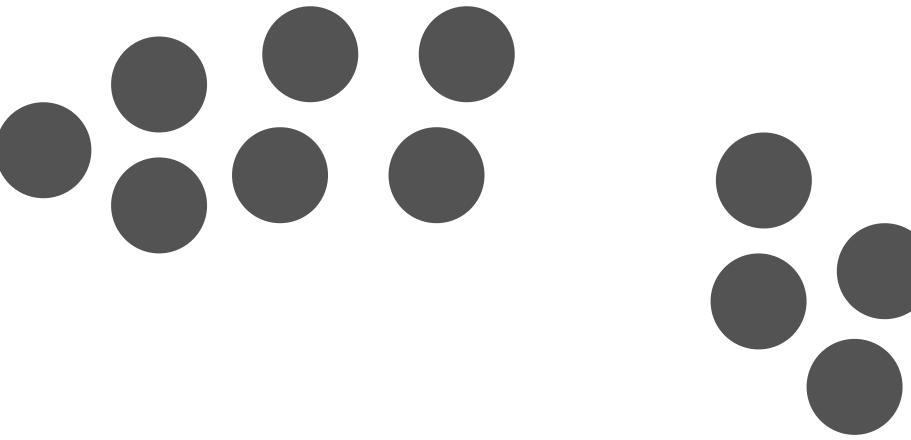
---

POSIÇÃO

2D



Agrupamento

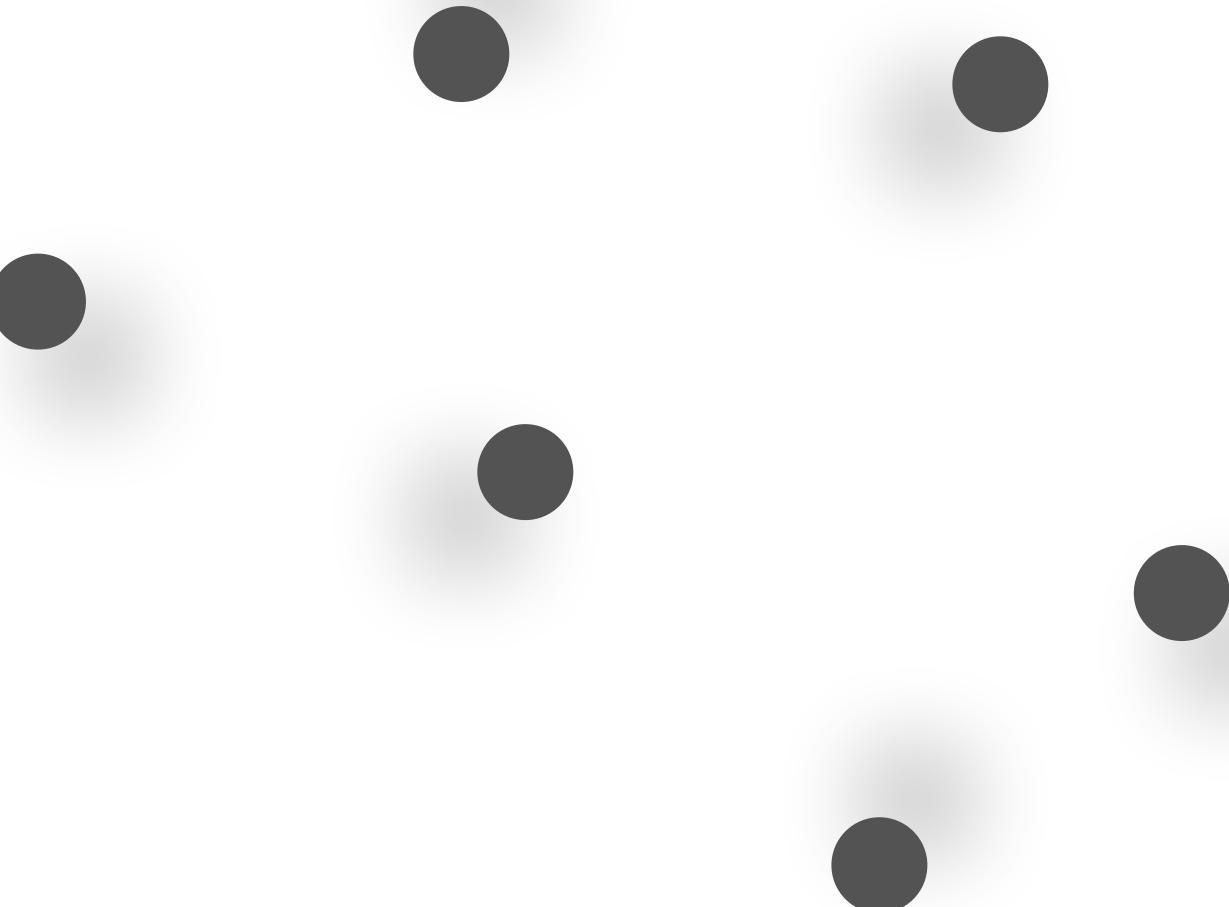


# ATRIBUTOS PRÉ-ATENTIVOS

---

## MOVIMENTO

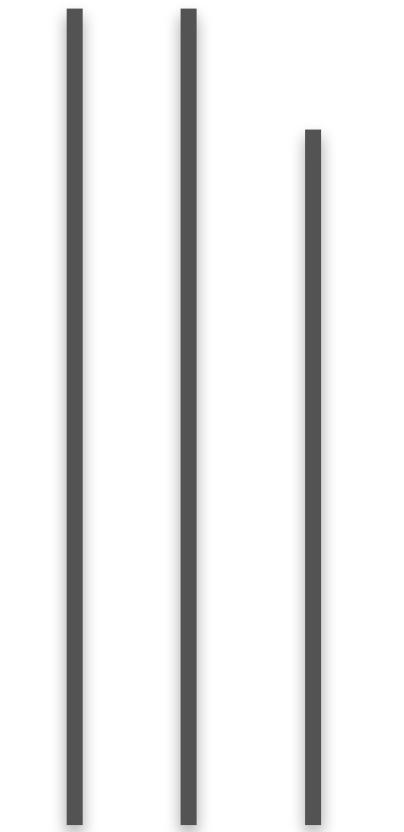
Direção



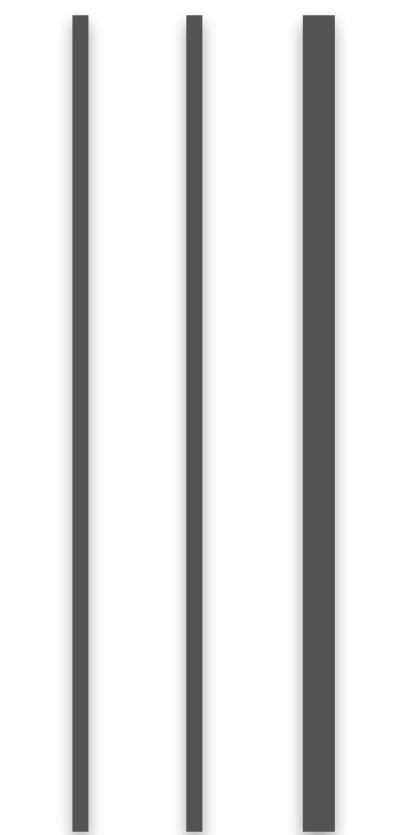
# PERCEPÇÃO DE QUANTIDADES

---

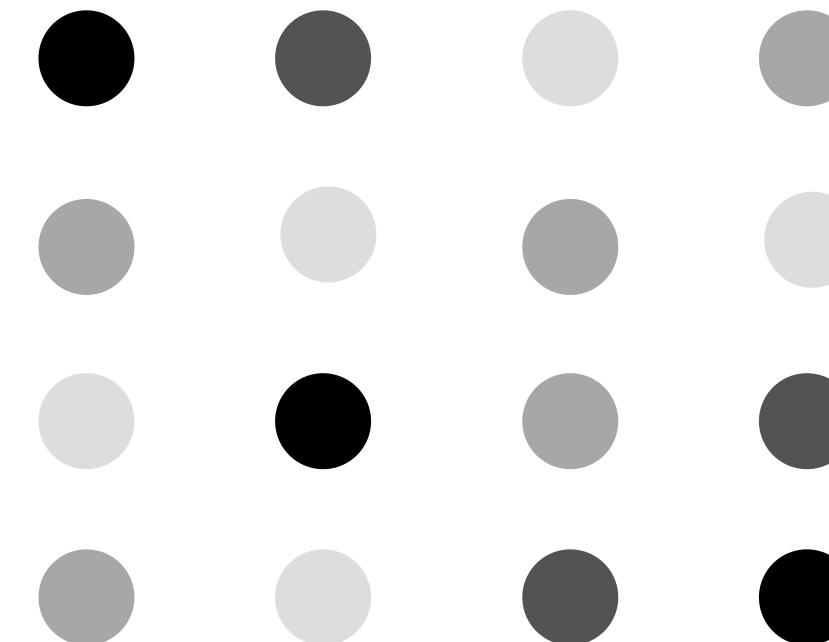
Comprimento



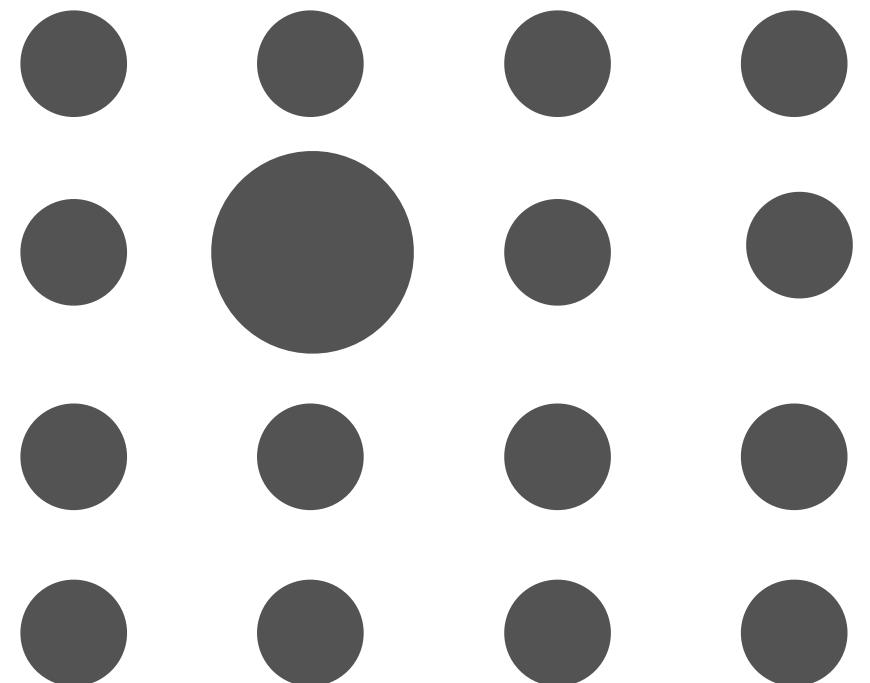
Espessura



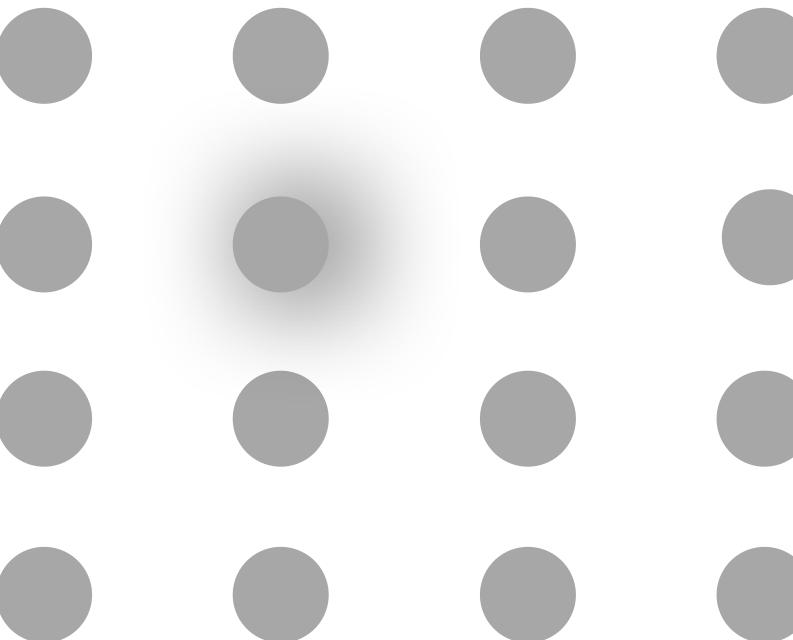
Intensidade



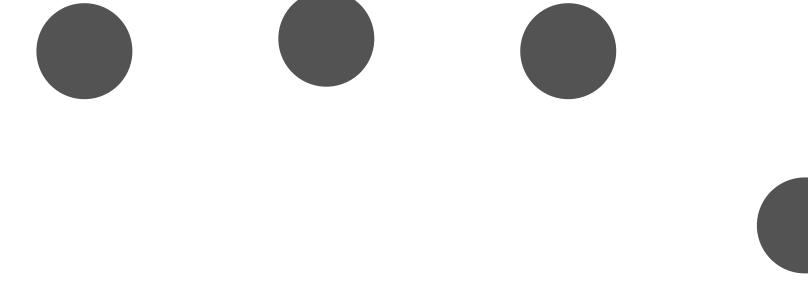
Tamanho



Sombra



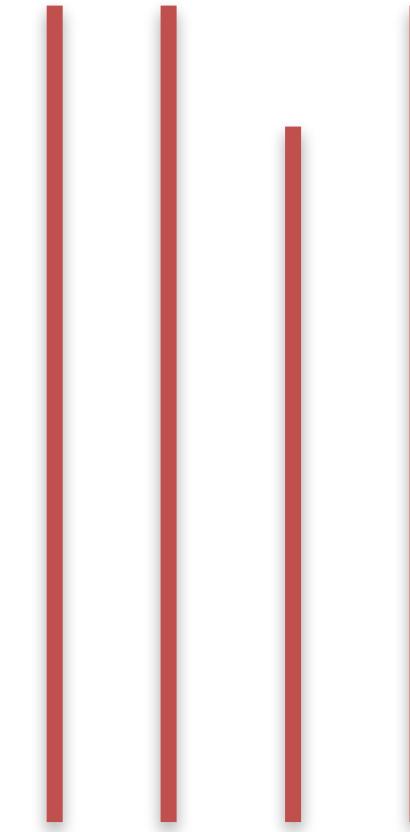
Posição 2D



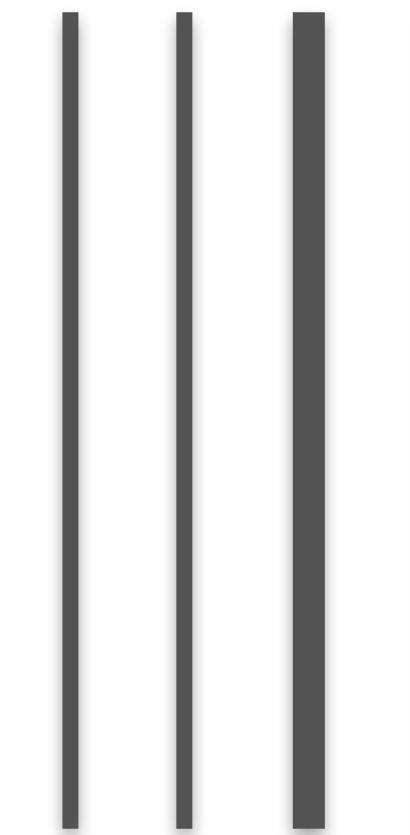
# PERCEPÇÃO DE QUANTIDADES

---

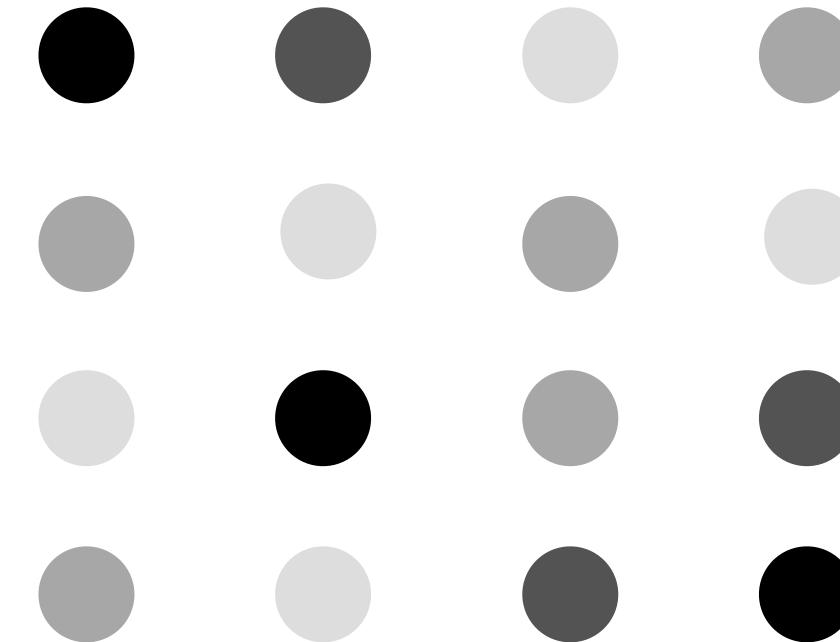
**Comprimento**



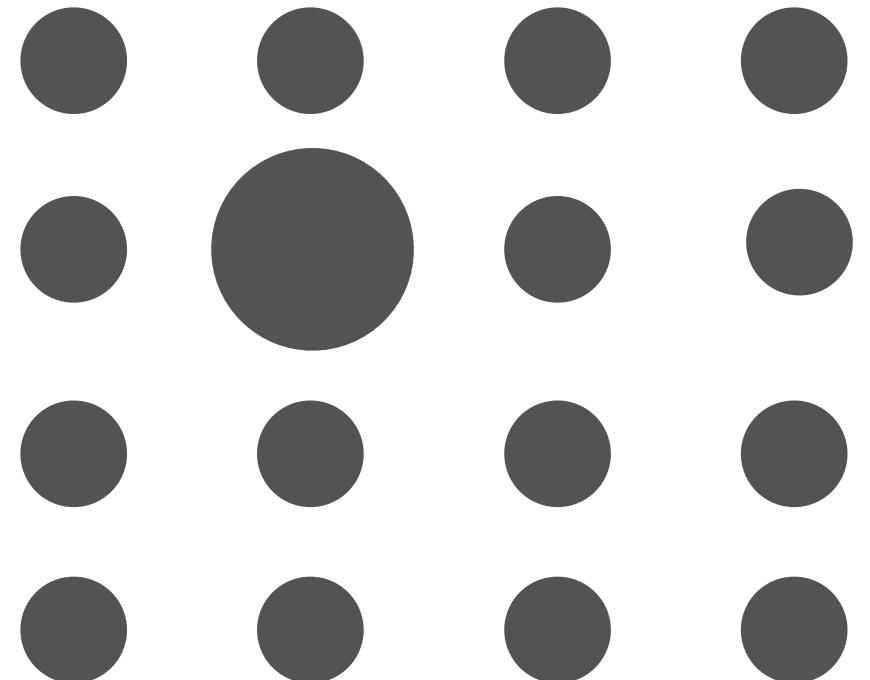
Espessura



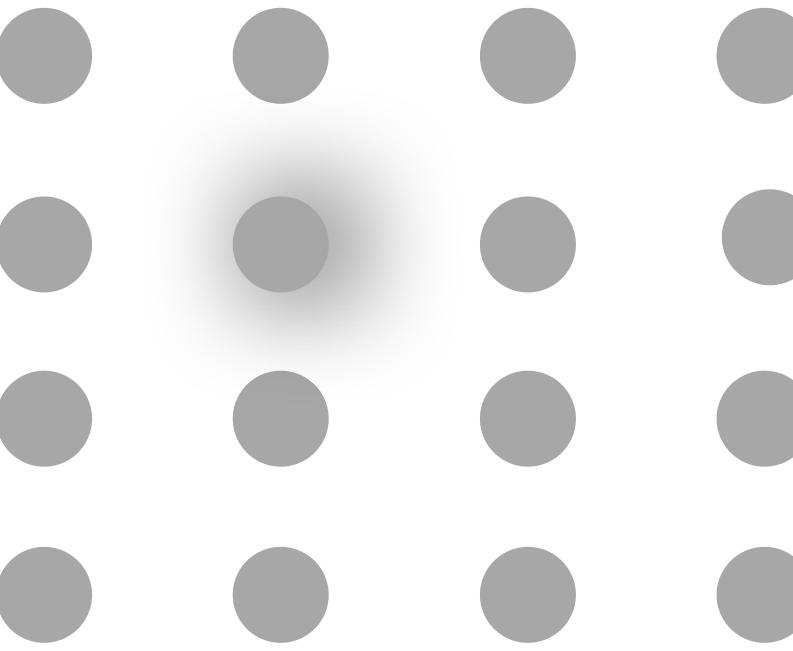
Intensidade



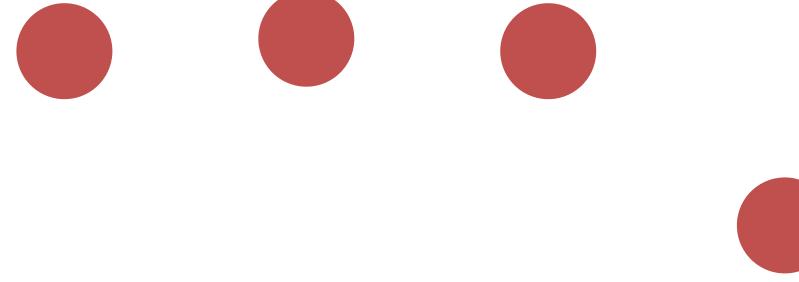
Tamanho



Sombra



**Posição 2D**

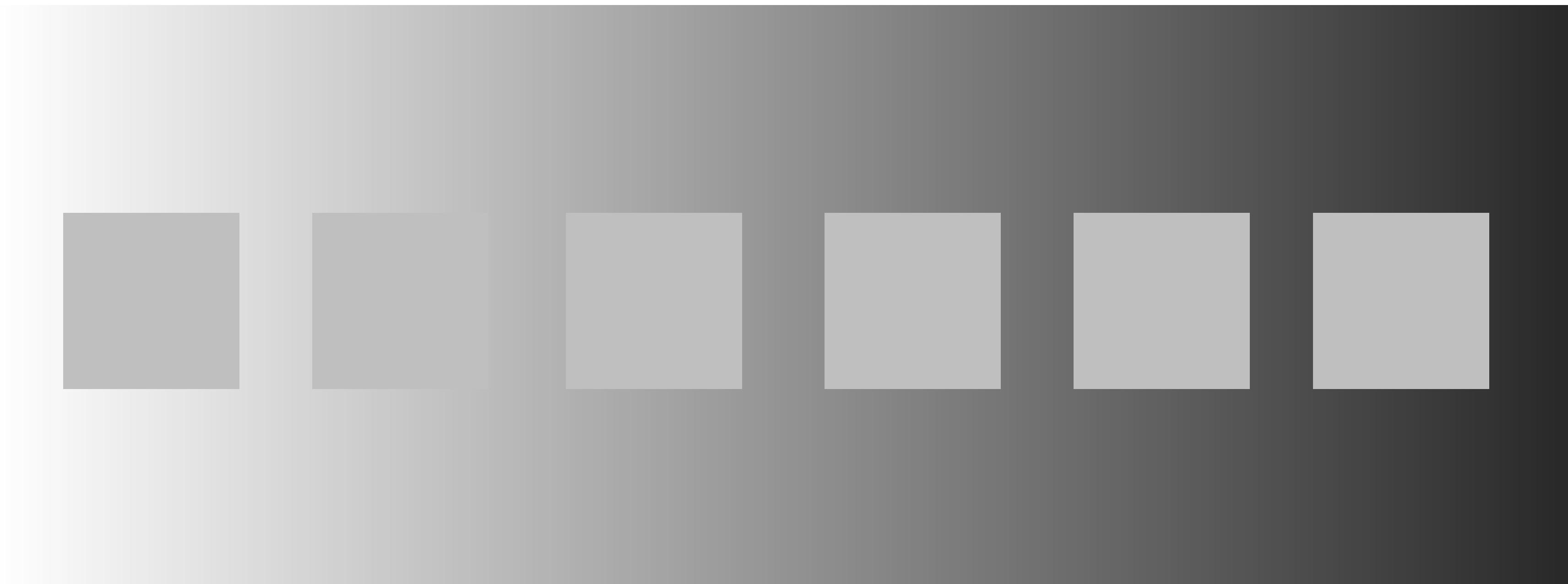


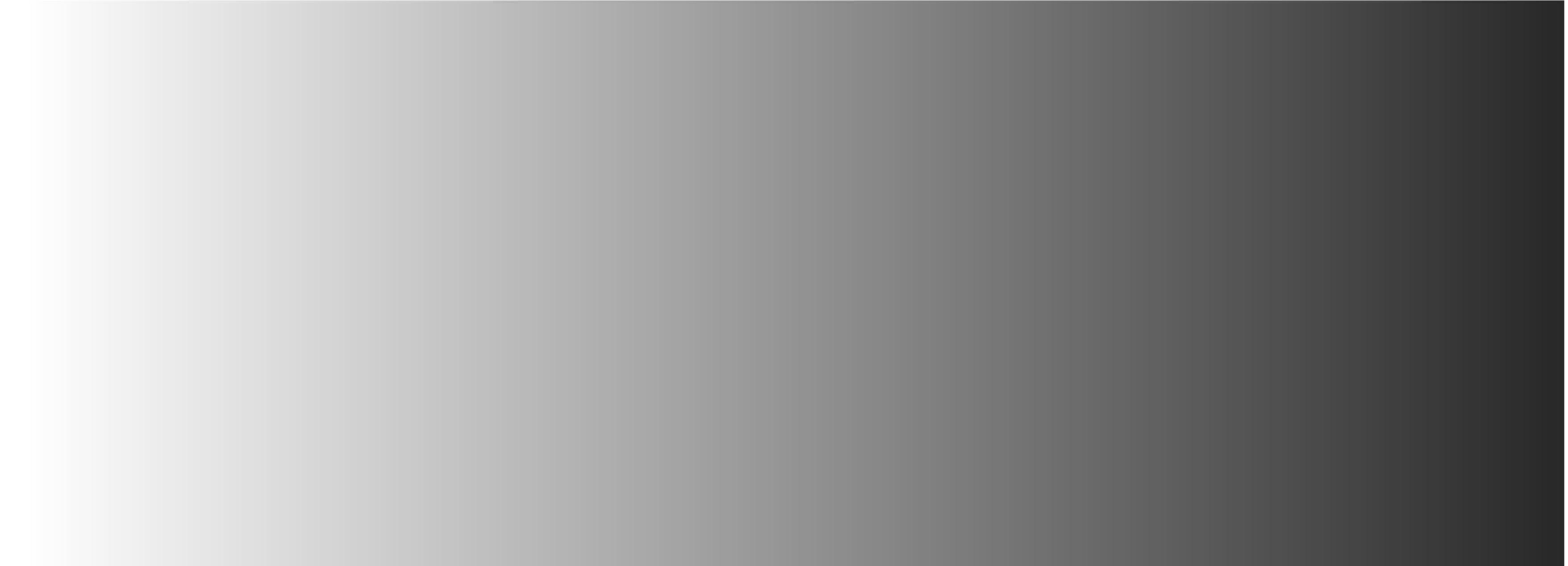
**MUITO PRECISA**

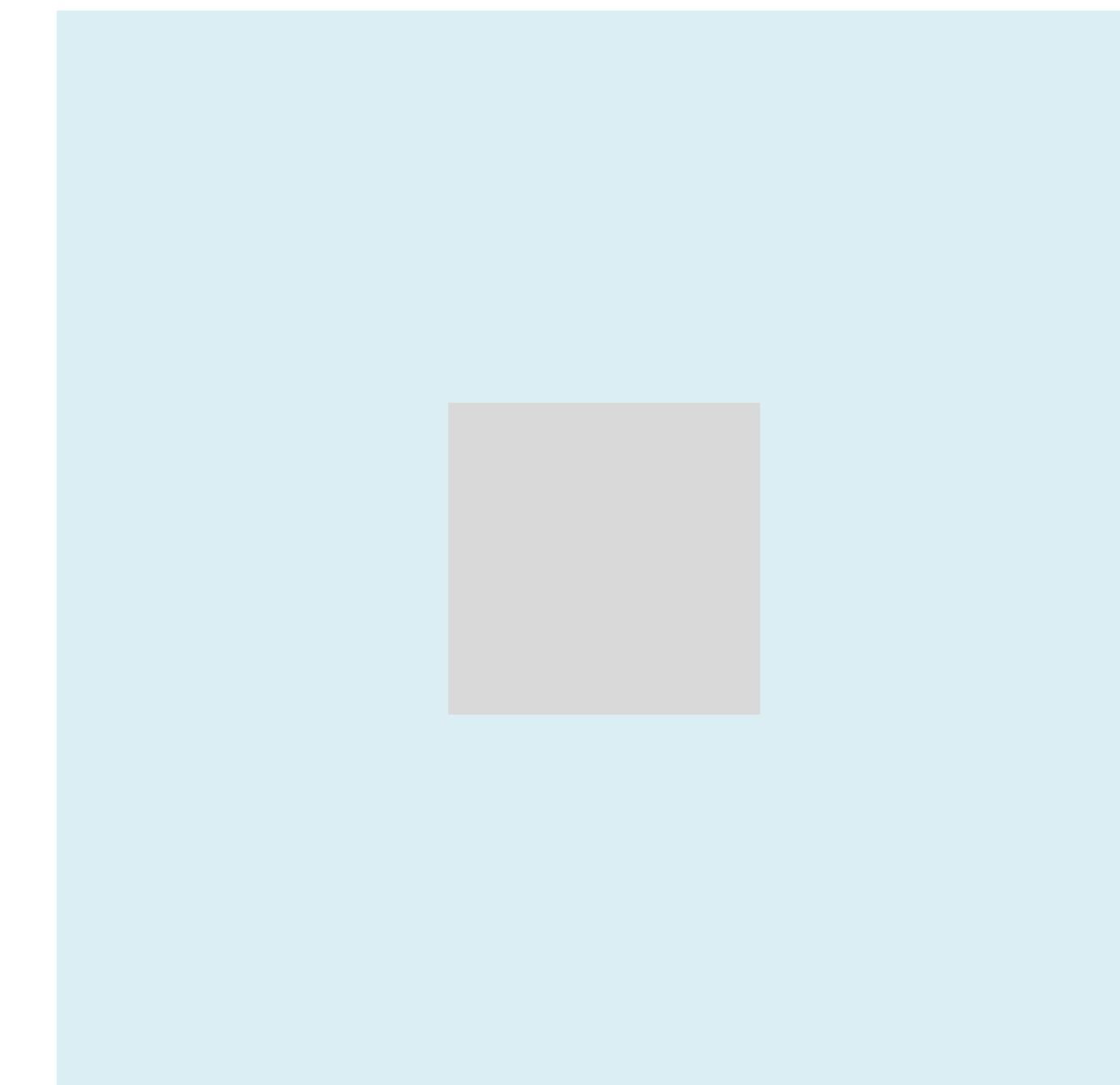
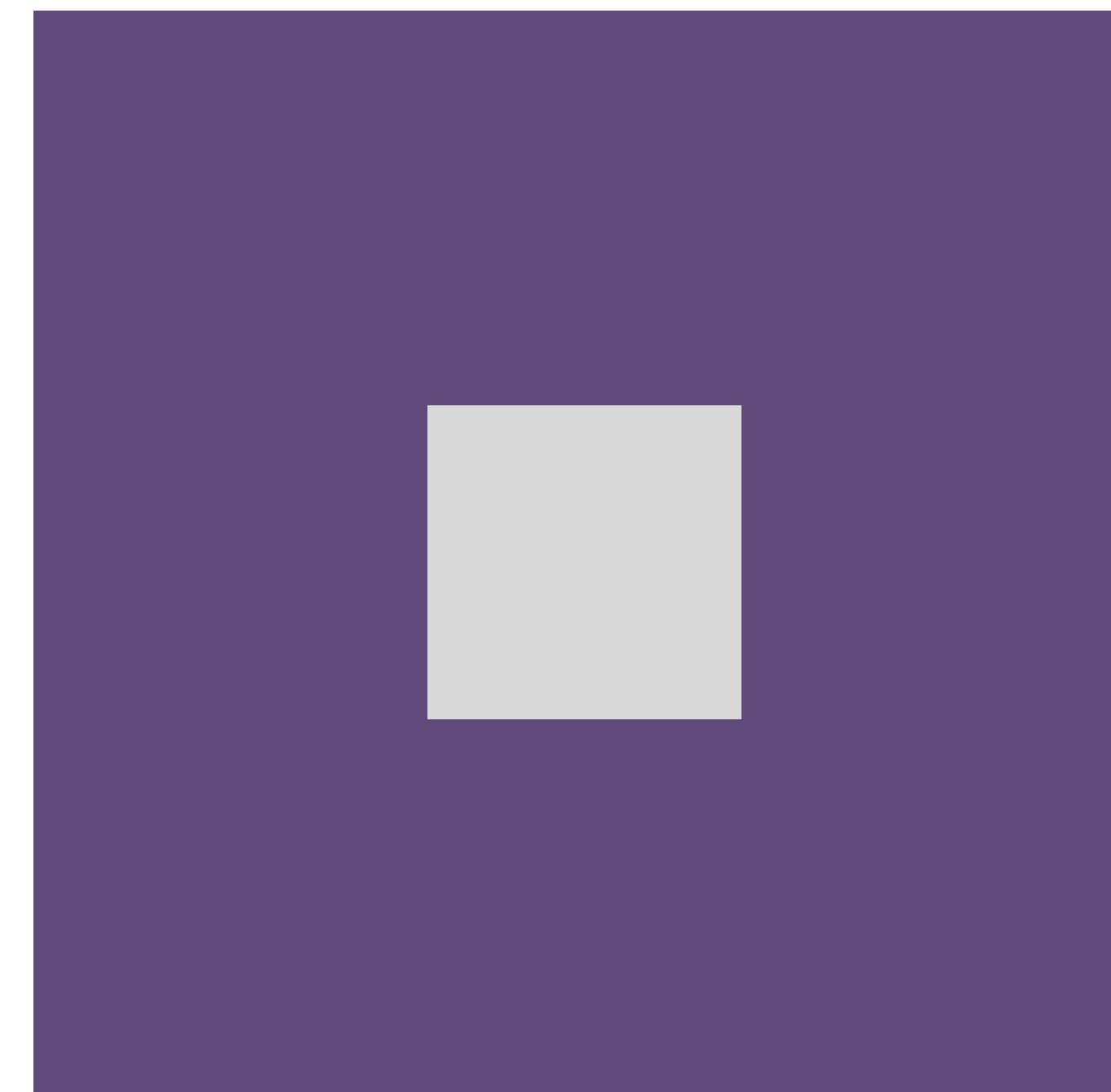
# Percepção e contexto

---

**Profa. Dra. Raquel Minardi**  
Departamento de Ciência da Computação  
Universidade Federal de Minas Gerais









# A PERCEPÇÃO E O CONTEXTO

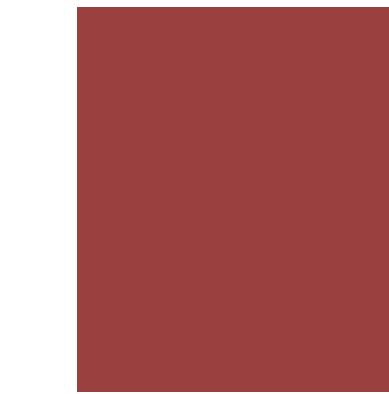
---

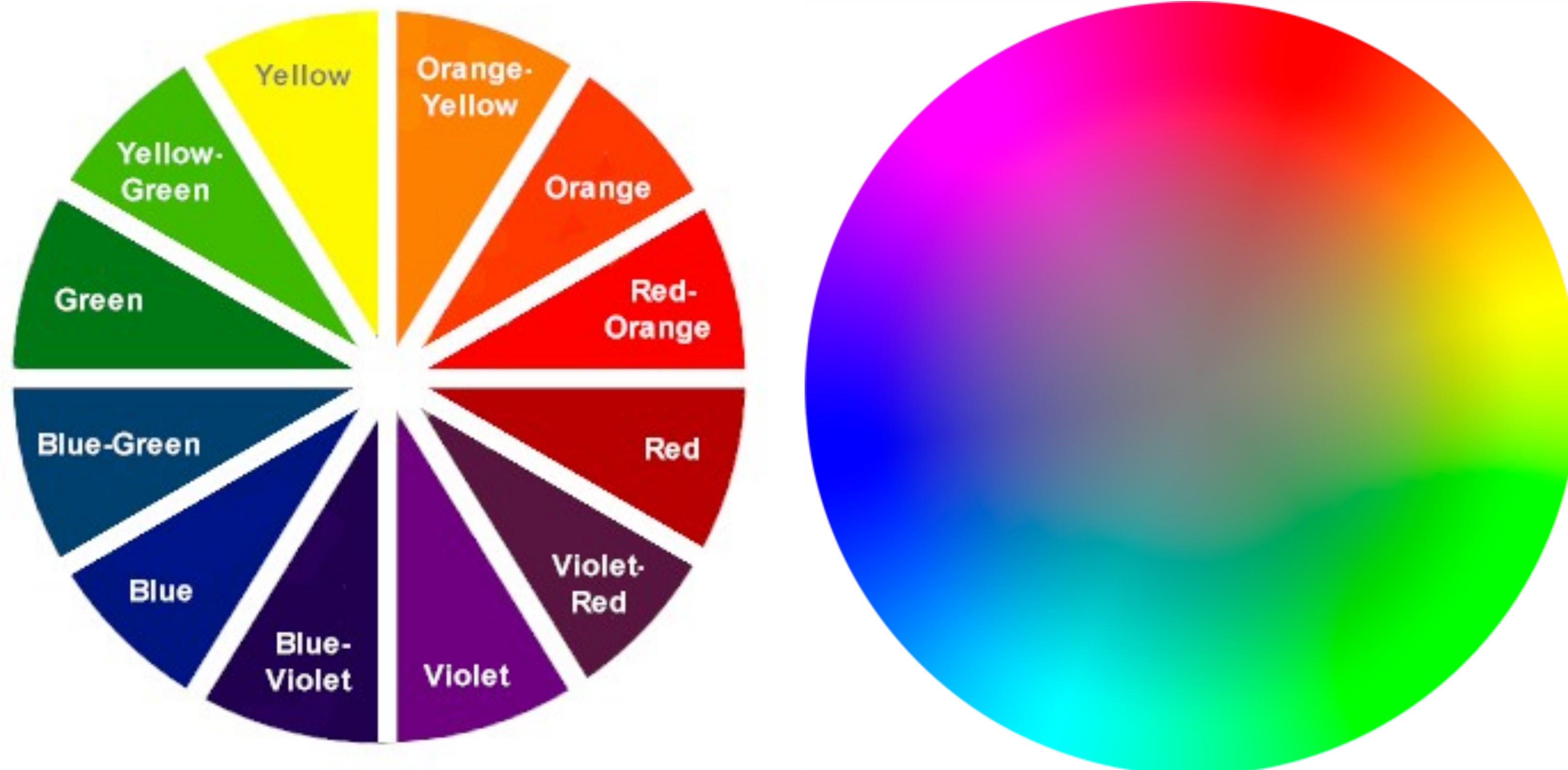
- A percepção visual **captura diferenças** e não valores absolutos, ou seja, é dependente do contexto

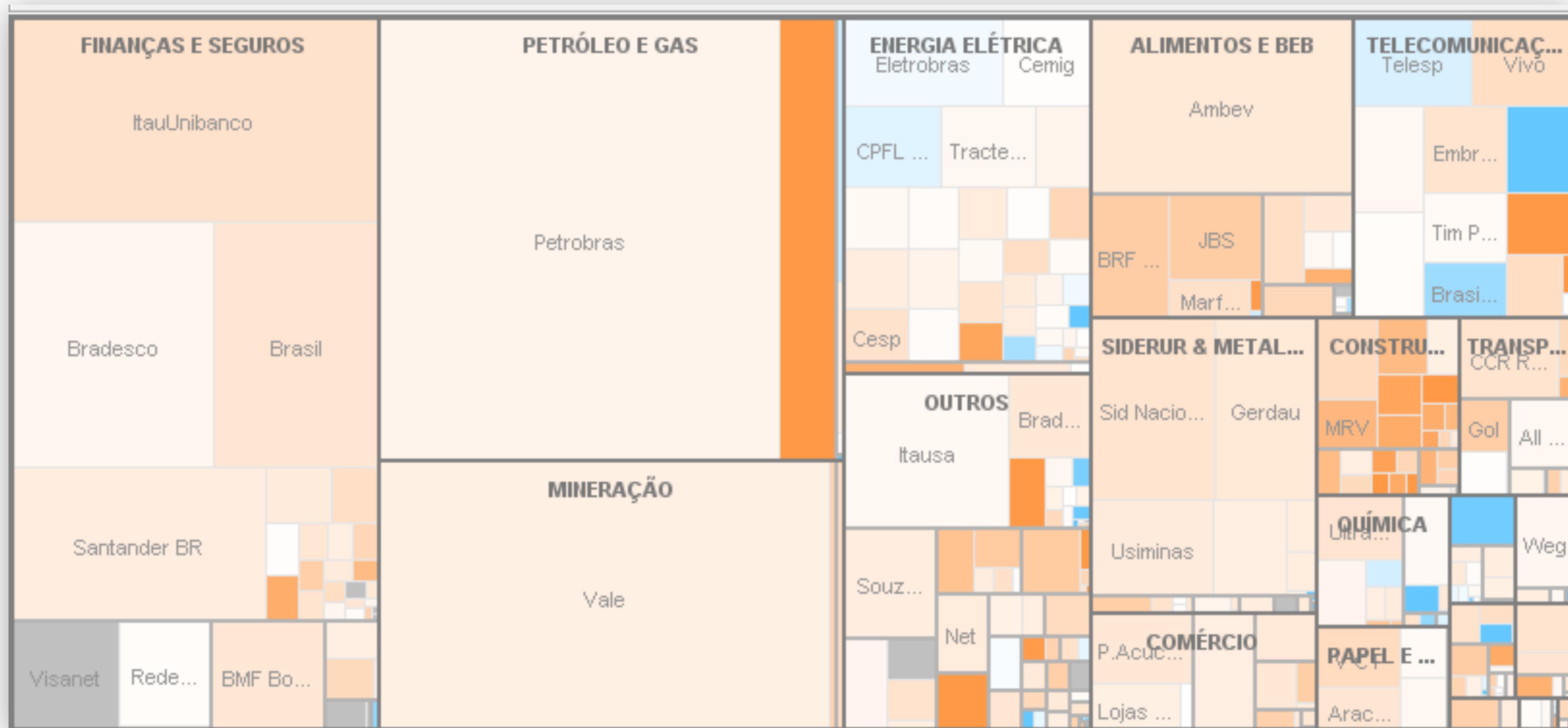
# CORES

---

- Quando desejamos usar cores para diferenciar categorias, devemos escolher valores dos atributos que variam significativamente







# DISTRAÇÕES

---

- É fácil distinguir um gavião no céu no meio de vários pombos
- Não é tão fácil distinguir um gavião se há diversos tipos de pássaros voando juntos no céu
- Portanto, símbolos pré-atentivos se tornam menos distintos quando o número de tipos de símbolos aumenta

# **Limitações de memória**

---

**Profa. Dra. Raquel Minardi**

Departamento de Ciência da Computação

Universidade Federal de Minas Gerais

# TIPOS DE MEMÓRIA

---

- **Memória de longo prazo:** onde a informação é armazenada permanentemente, disponível para recuperação quando necessária
- **Memória de trabalho:** armazena a informação brevemente enquanto está sendo processada

# MEMÓRIA DE TRABALHO

---

- A informação entra na memória de trabalho de três formas: através dos sentidos, através de nossa imaginação ou da memória de longo prazo
  
- A informação fica na memória de trabalho enquanto estamos pensando nela
  
- Uma vez que deixamos de pensar nela, ela reside ali por poucos segundos
  
- Se pensamos nela suficientemente, ela vai para a memória de longo prazo, senão ela é apagada

# MEMÓRIA DE TRABALHO VISUAL

---

- A memória de trabalho possui compartimentos para diferentes tipos de informação, por exemplo, texto ou imagem
  
- Nossa memória visual de trabalho tem em torno de 3 compartimentos
  
- Quanto de informação cabe em um compartimento?
  
- Um chunk ou bloco de informação

# MEMÓRIA DE TRABALHO VISUAL

---

- Esta é uma das grandes vantagens da visualização para exploração e análise de dados
- Quando a informação é exibida de forma visual, fazendo uso de padrões visuais, mais informação é contida em um chunk, possibilitando o pensamento simultâneo a respeito de grandes quantidades de dados e aumentando nossa capacidade de identificar padrões complexos

“

The process of grouping simple concepts into more complex ones is called chunking. A chunk can be almost anything:  
a mental representation of an object;  
a plan;  
a group of objects;  
a method for achieving some goal.

The process of becoming an expert in a particular domain is largely one of creating effective high-level concepts or chunks.

-Stuart Card

in Visualization: Perception for Design

# “ARMAZENAMENTO EXTERNO”

---

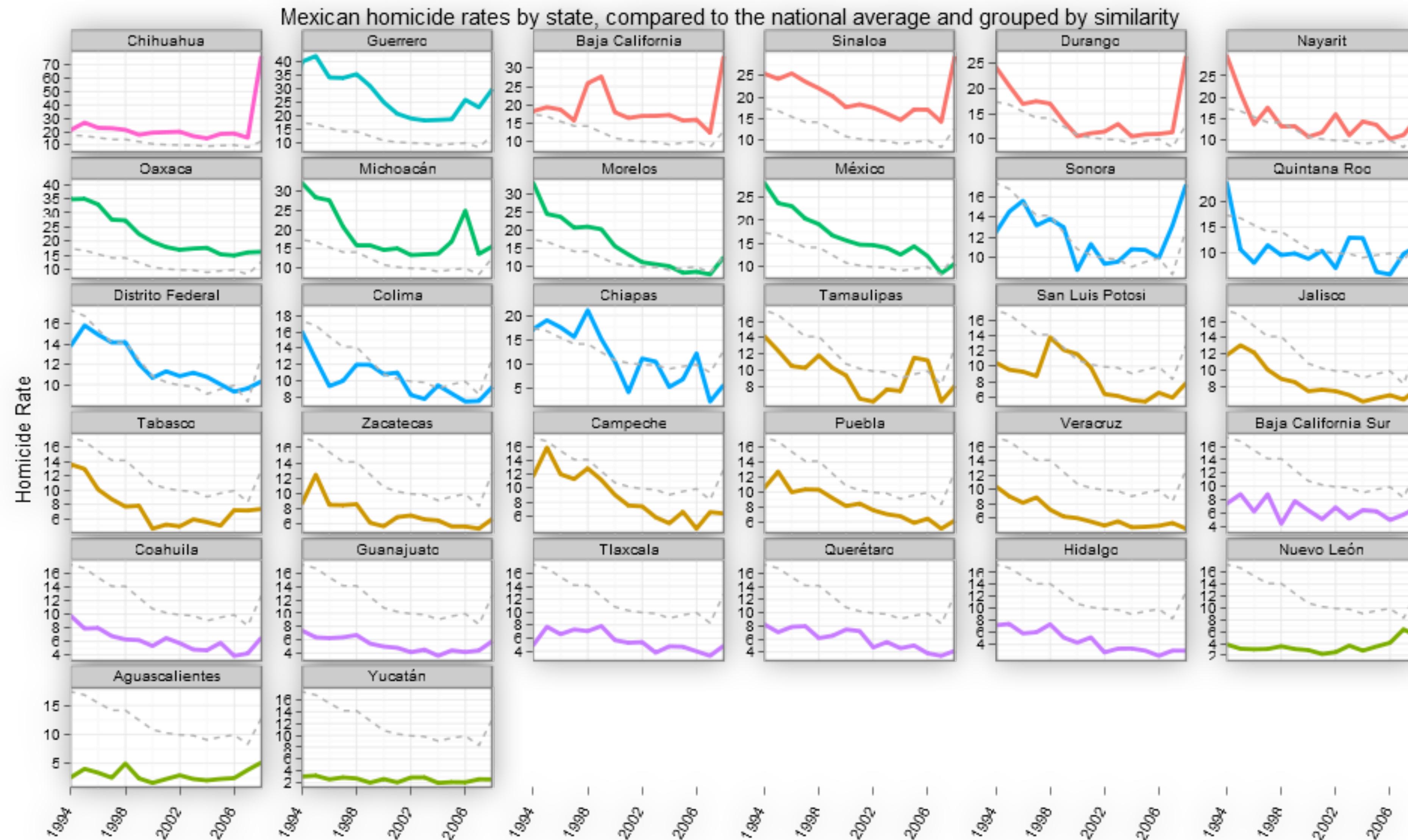
- Mesmo que nossa memória “cresça” com a utilização de representações visuais na representação de informação, a memória ainda assim é limitada
- Assim, é desejável que tenhamos um meio externo à memória para armazenar os dados e informações em exploração
- Uso da própria tela do computador ou **página de papel** para exibir informações que podem ser carregadas e descarregadas rapidamente da memória de trabalho visual para facilitar a cognição

# PEQUENOS MÚLTIPLOS (SMALL MULTIPLES)

- É um padrão de projeto clássico, proposto por Edward Tufte, que consiste na mesma representação exibida múltiplas vezes (diferentes valores ou variáveis) e possível de ser acessada visualmente facilmente
- Comparativos
- Multivariados
- Alta densidade
- Interpretação eficiente
- Narrativos visto que representam mudanças de relacionamentos entre variáveis a cada diferente índice



# PEQUEÑOS MULTIPLOS (SMALL MULTIPLES)



# **Outras leis**

---

**Profa. Dra. Raquel Minardi**

Departamento de Ciência da Computação

Universidade Federal de Minas Gerais

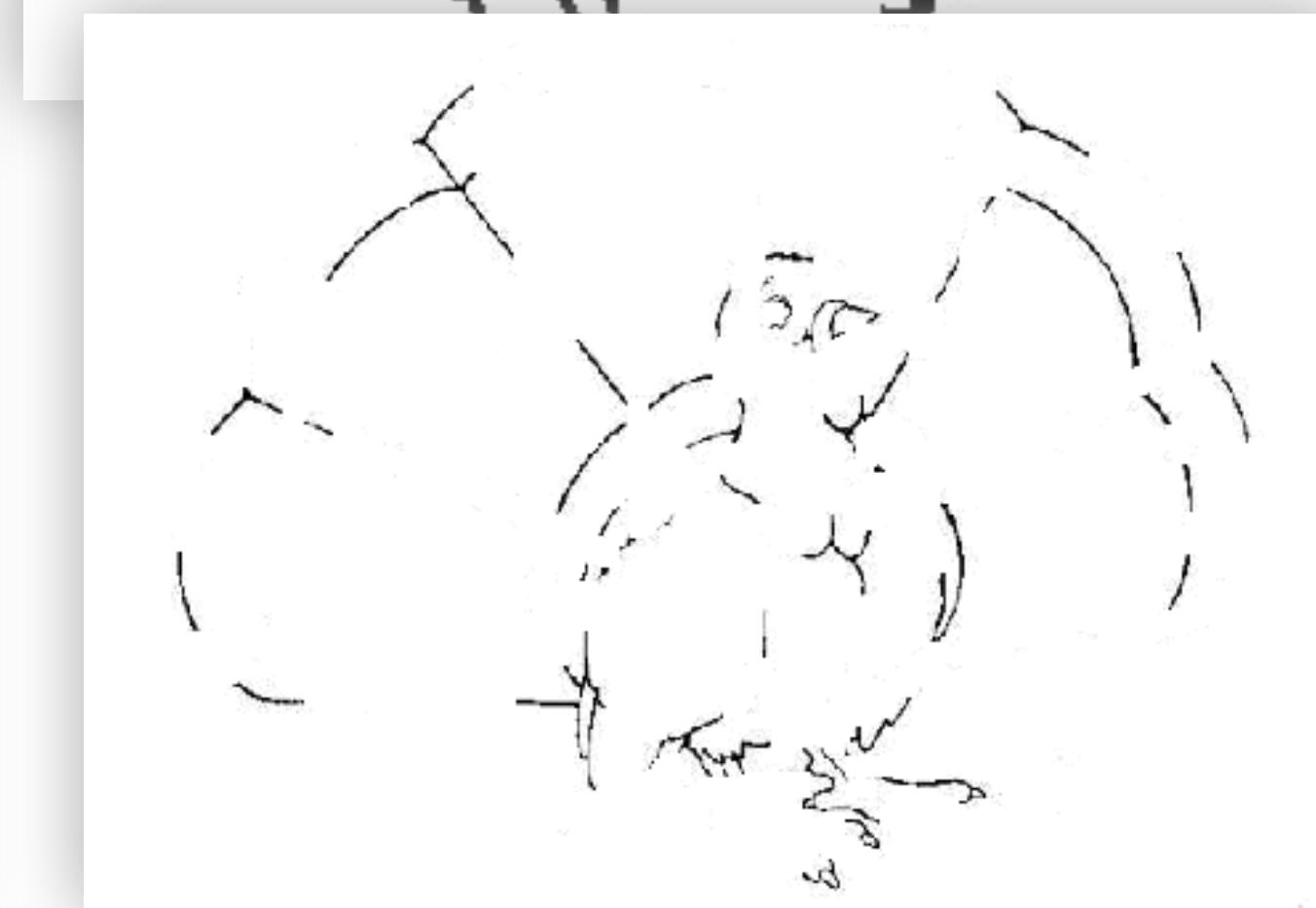
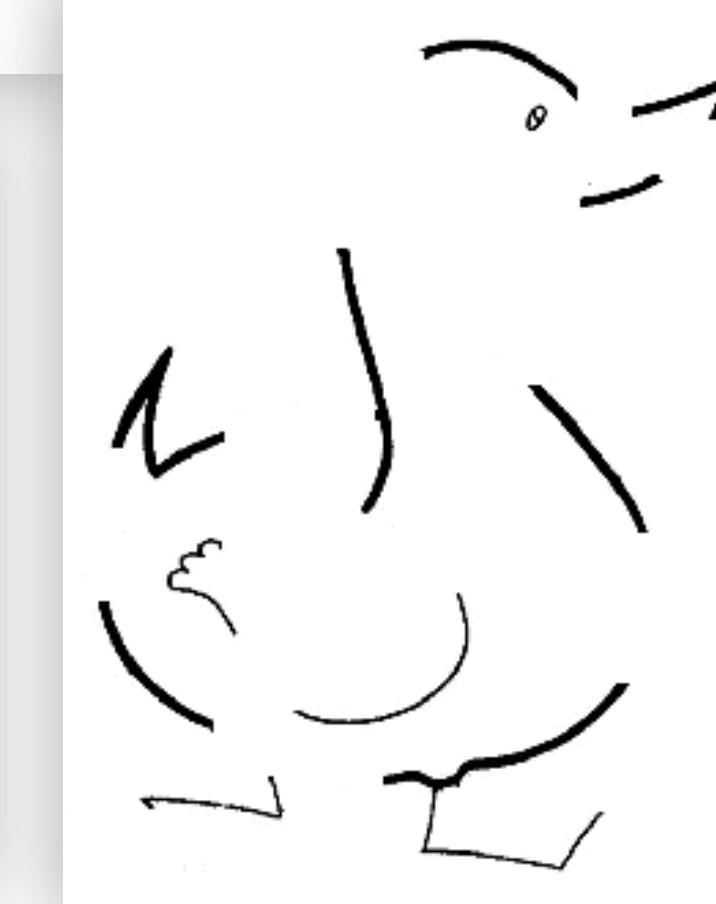
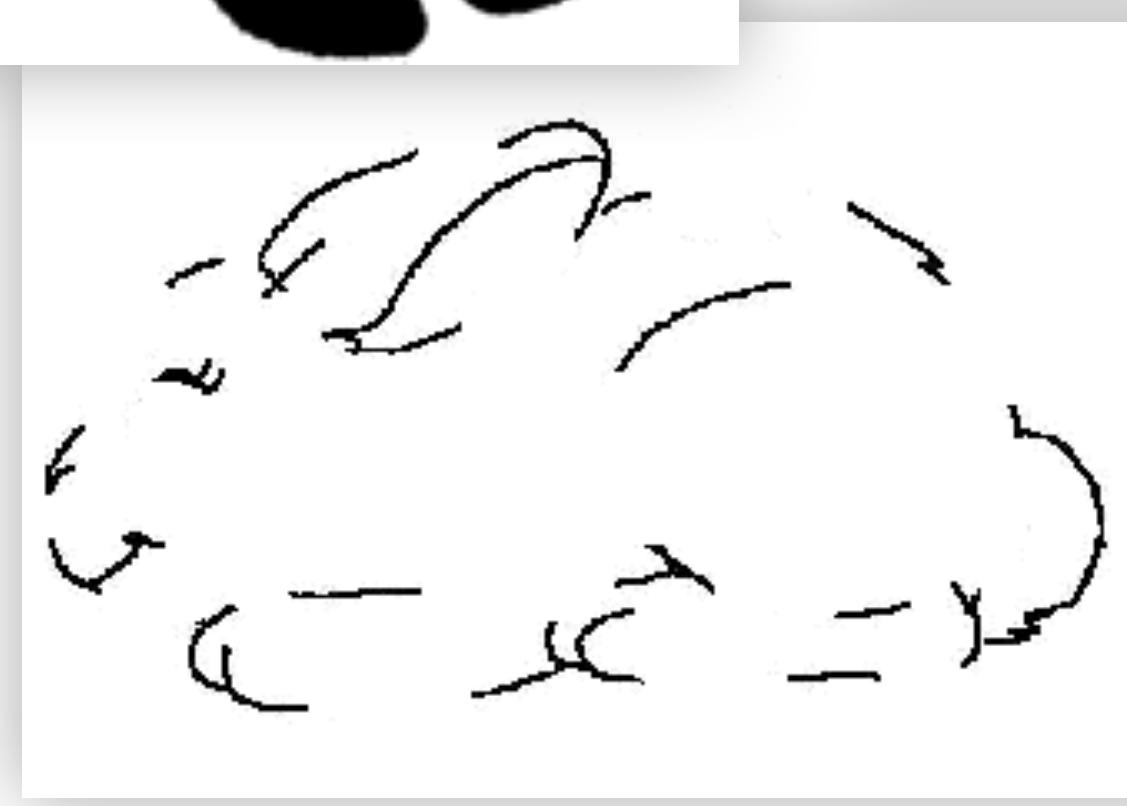
# PRINCÍPIOS GESTALT

---

- Gestalt significa essência ou forma de um objeto
- Estudos da psicologia que tiveram início na década de 20 sobre a nossa capacidade visual de reconhecimento de figuras ao invés de simples coleções de linhas e formas

# LEI DO FECHAMENTO

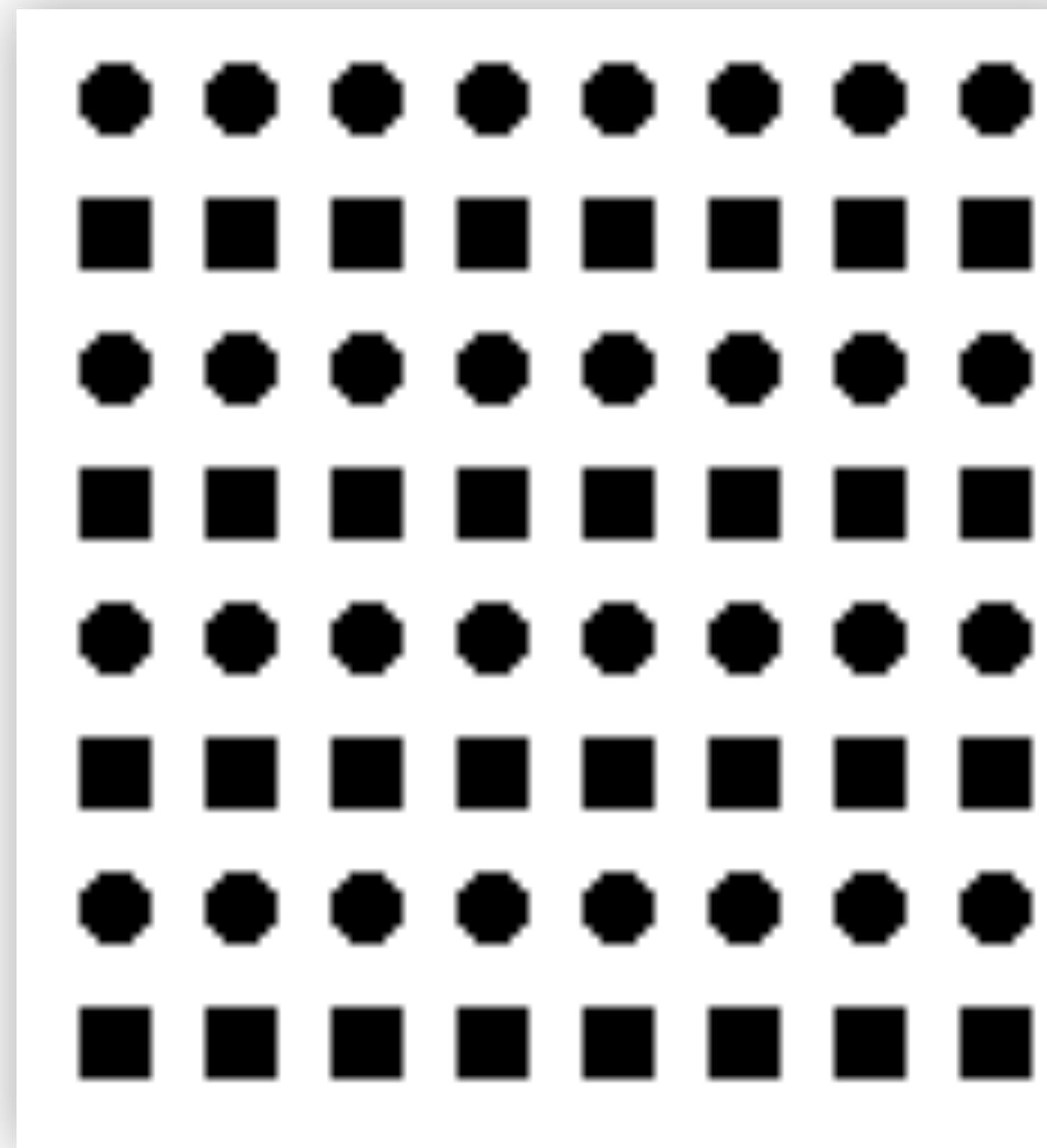
- A mente pode utilizar elementos da sua experiência mesmo sem a percepção de elementos visuais para completar uma figura



# LEI DA SIMILARIDADE

---

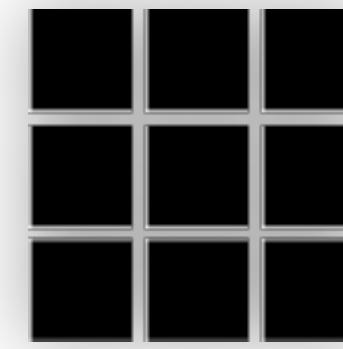
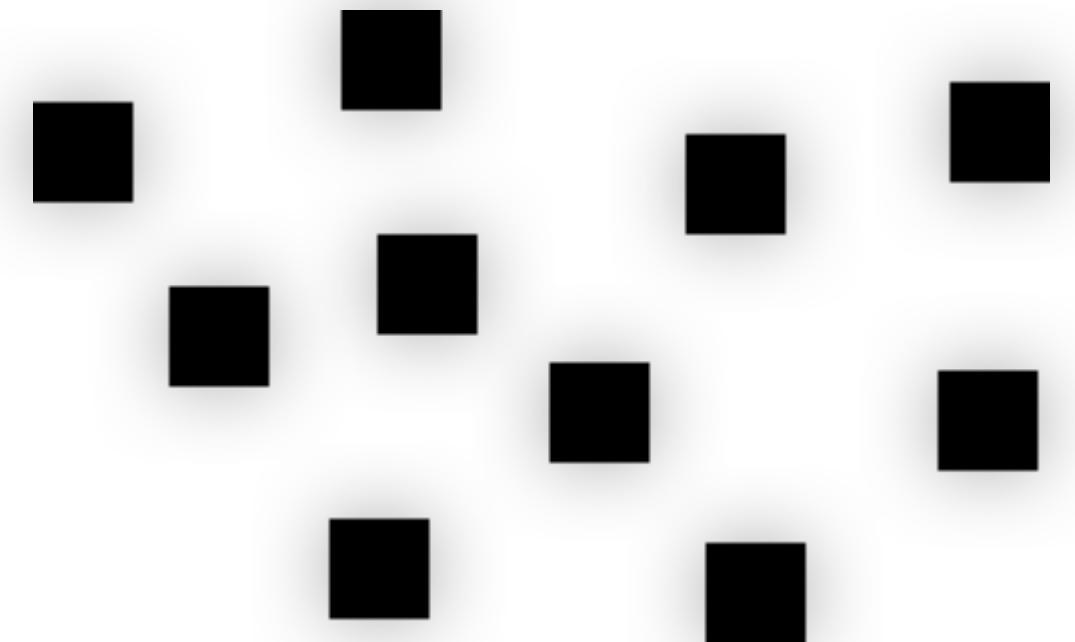
- A mente agrupa elementos similares em entidades coletivas
- Esta similaridade pode ser baseada em forma, cores, tamanho ou brilho



# LEI DA PROXIMIDADE

---

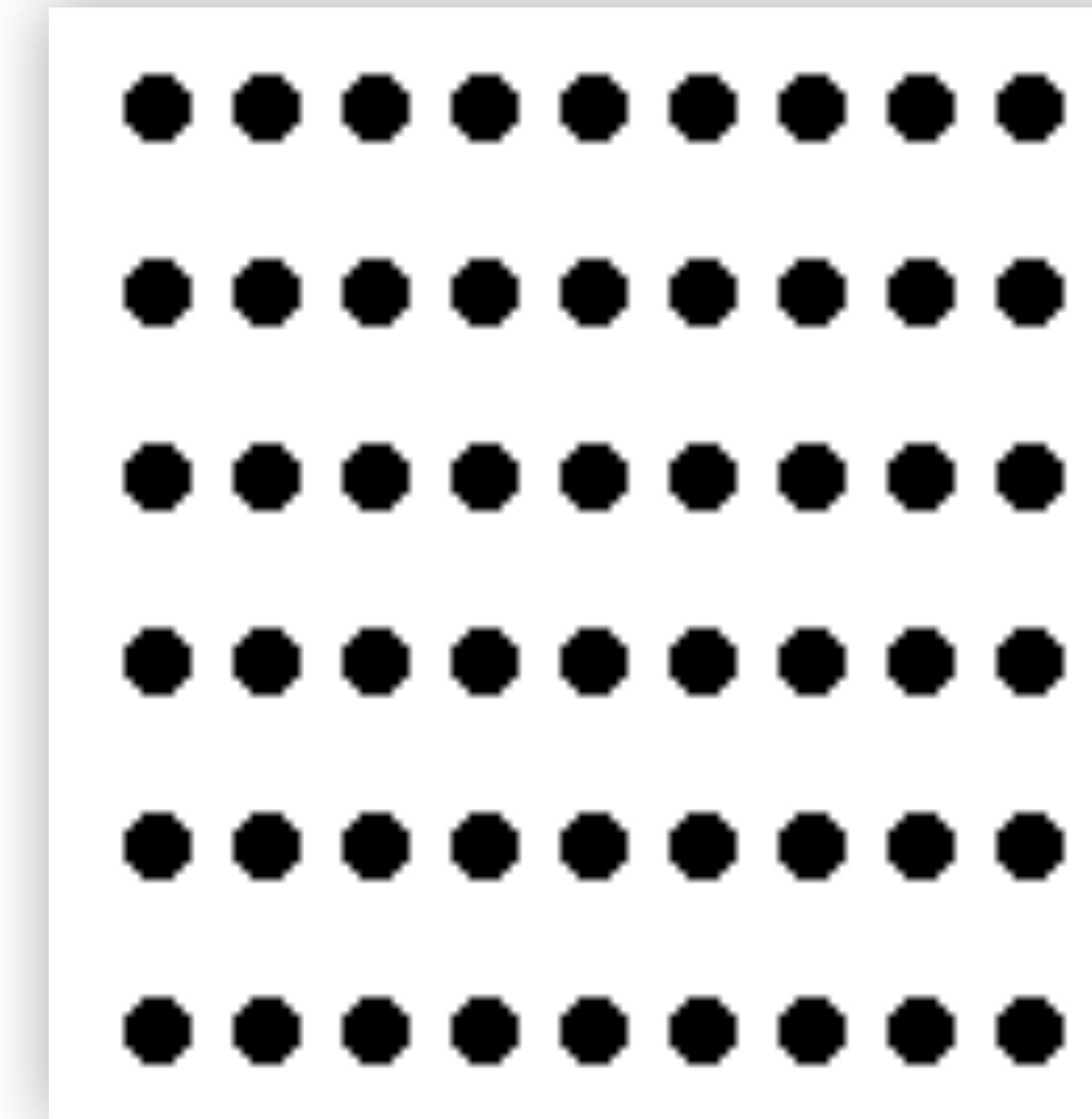
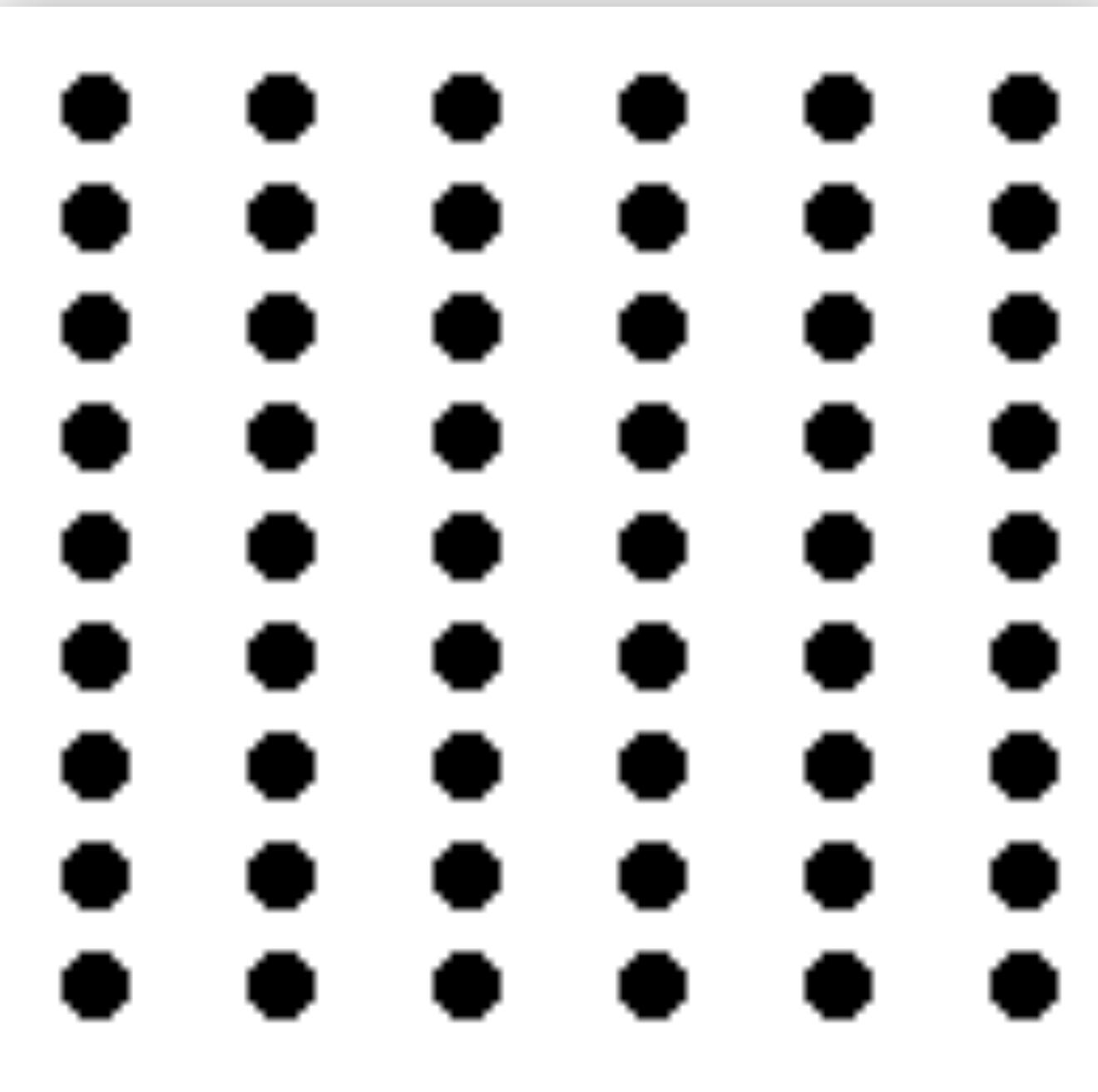
- Proximidade espacial ou temporal de elementos induz a mente a percebê-los coletivamente



# LEI DA PROXIMIDADE

---

- Características mais próximas são mais rapidamente associadas



# LEI DA PROXIMIDADE

---

- O que você vê?
- Pares de linhas próximas ou pares de linhas distantes?



# **Conceitos fundamentais sobre tabelas e gráficos**

---

**Profa. Dra. Raquel Minardi**  
Departamento de Ciência da Computação  
Universidade Federal de Minas Gerais

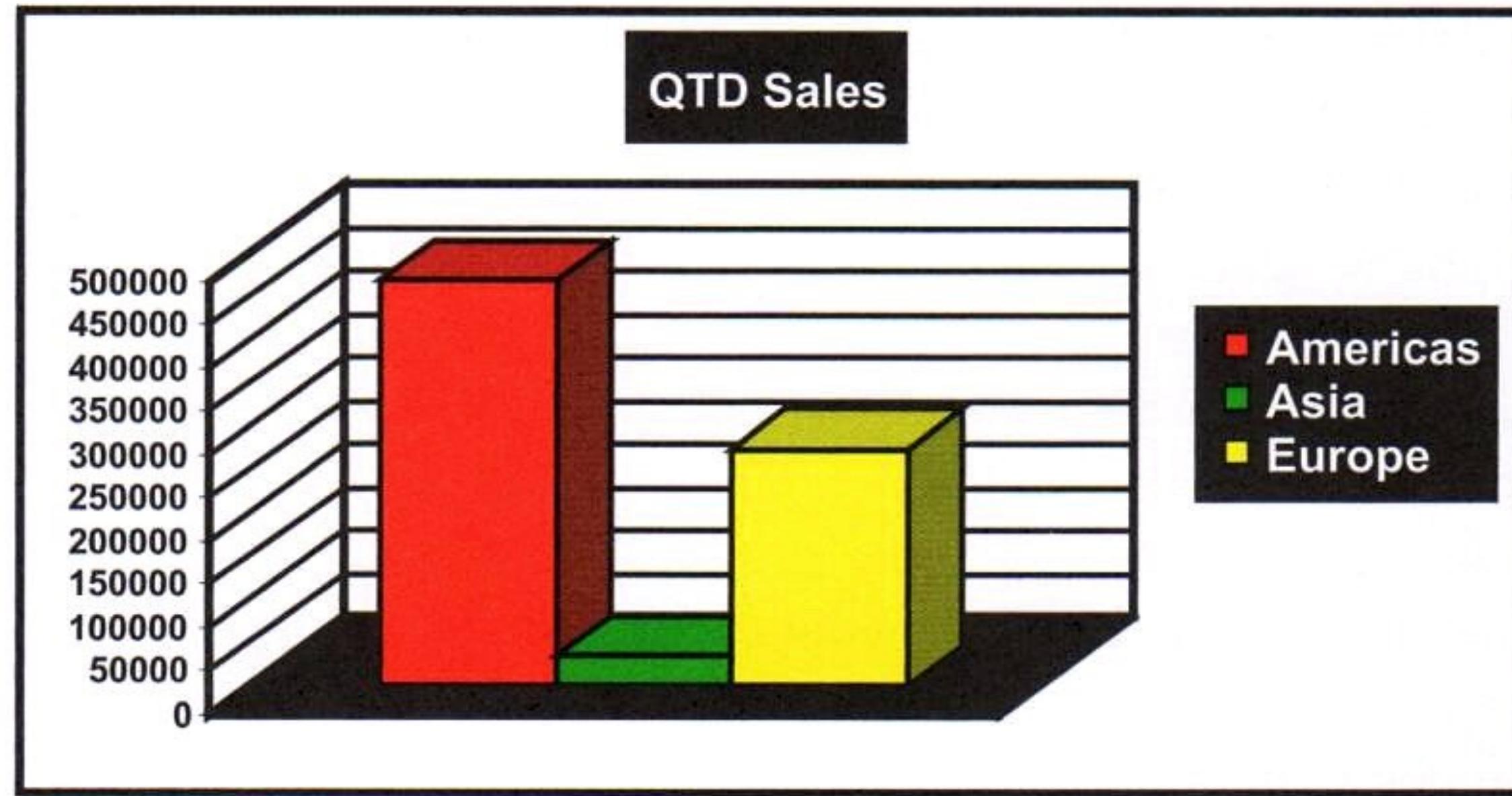


FIGURE 1.1 This is a typical example of a poorly designed business graph. Notice the attempt at artistic flair in the use of color, 3-D, and shading of the vertical bars.

- Unidades de medida?
- Data?
- Como estas medidas se comparam aos planos?
- Como estas medidas se comparam a outros períodos?
  
- Faltam no gráfico informações importantes sobre o contexto

## 2003 Q1-to-Date Regional Sales

March 15, 2003

	Sales (U.S. \$)	Percent of Total Sales	Current Percent of Qtr Plan	Projected Sales (U.S. \$)	Qtr End Projected Percent of Qtr Plan
Americas	469,384	60%	85%	586,730	107%
Europe	273,854	35%	91%	353,272	118%
Asia	34,847	5%	50%	43,210	62%
	\$778,085	100%	85%	\$983,212	107%

Note: To date, 83% of the quarter has elapsed.

FIGURE 1.2 This table contains all of the information that was contained in the graph in *Figure 1.1*, plus much, much more, without becoming complicated.

## SlicersDicers Sales Compared to Other Product Sales

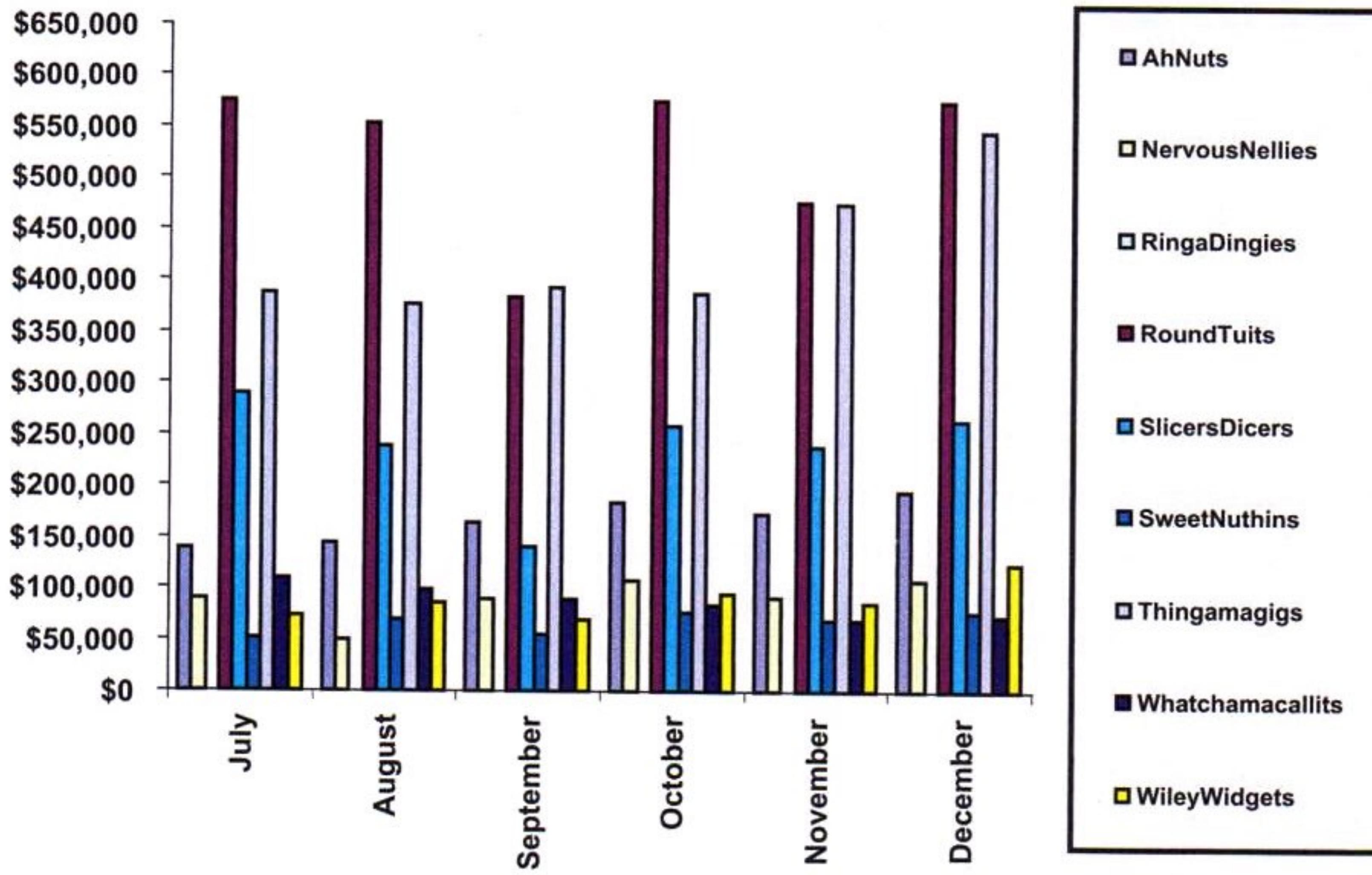


FIGURE 1.3 This is a typical example of a poorly designed vertical bar graph.



## Sales of SlicersDicers Compared to Sales of Other Products

July - December, 2006

(SlicersDicers' sales are displayed as black reference lines of 100%; the gray lines represent the average monthly sales for July through December.)

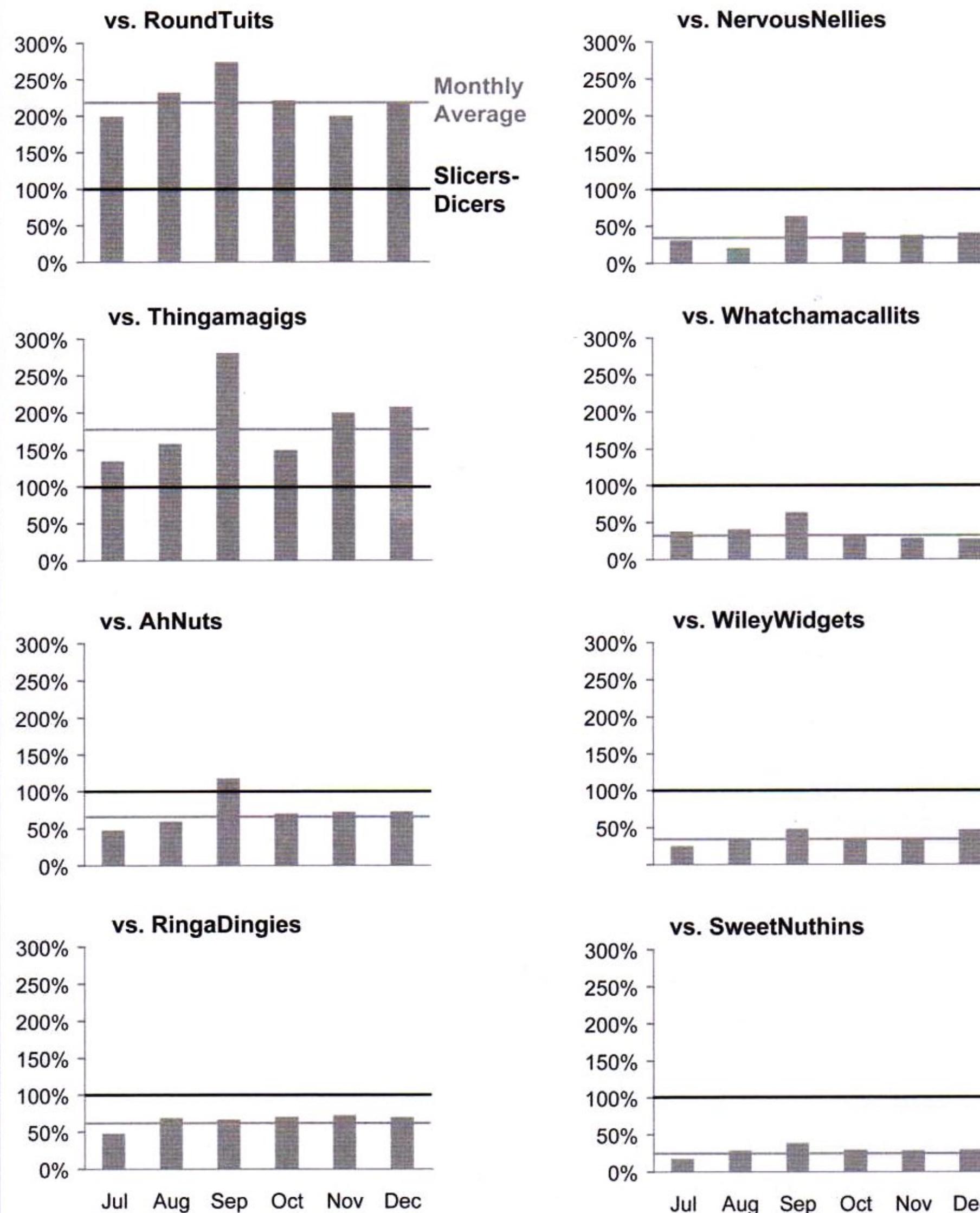


FIGURE 1.4 This is a series of related graphs, each designed to compare the sales of the SlicersDicers product to those of a different product.

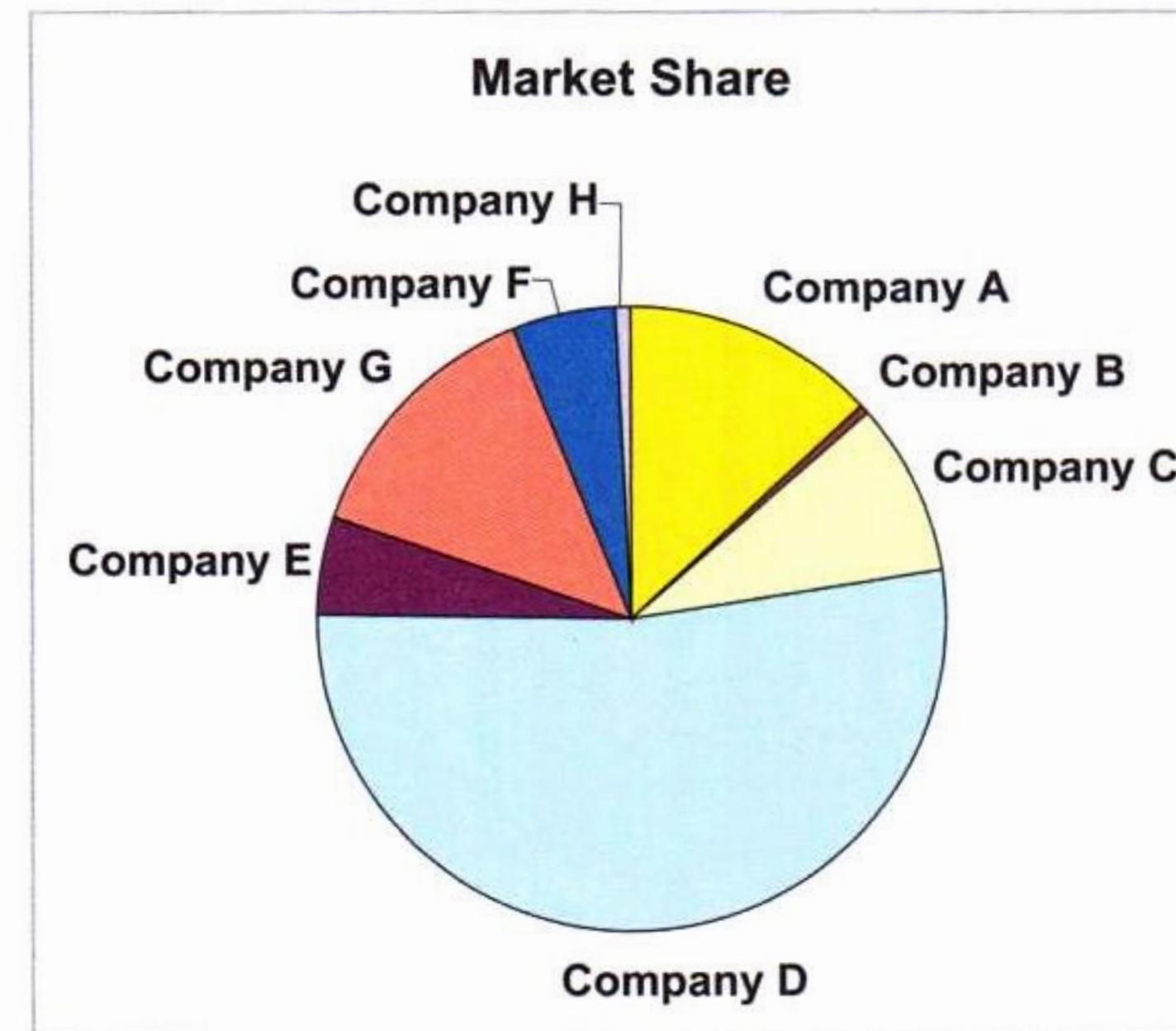


FIGURE 1.5 This is an example of an ineffective style of graph: a pie chart.

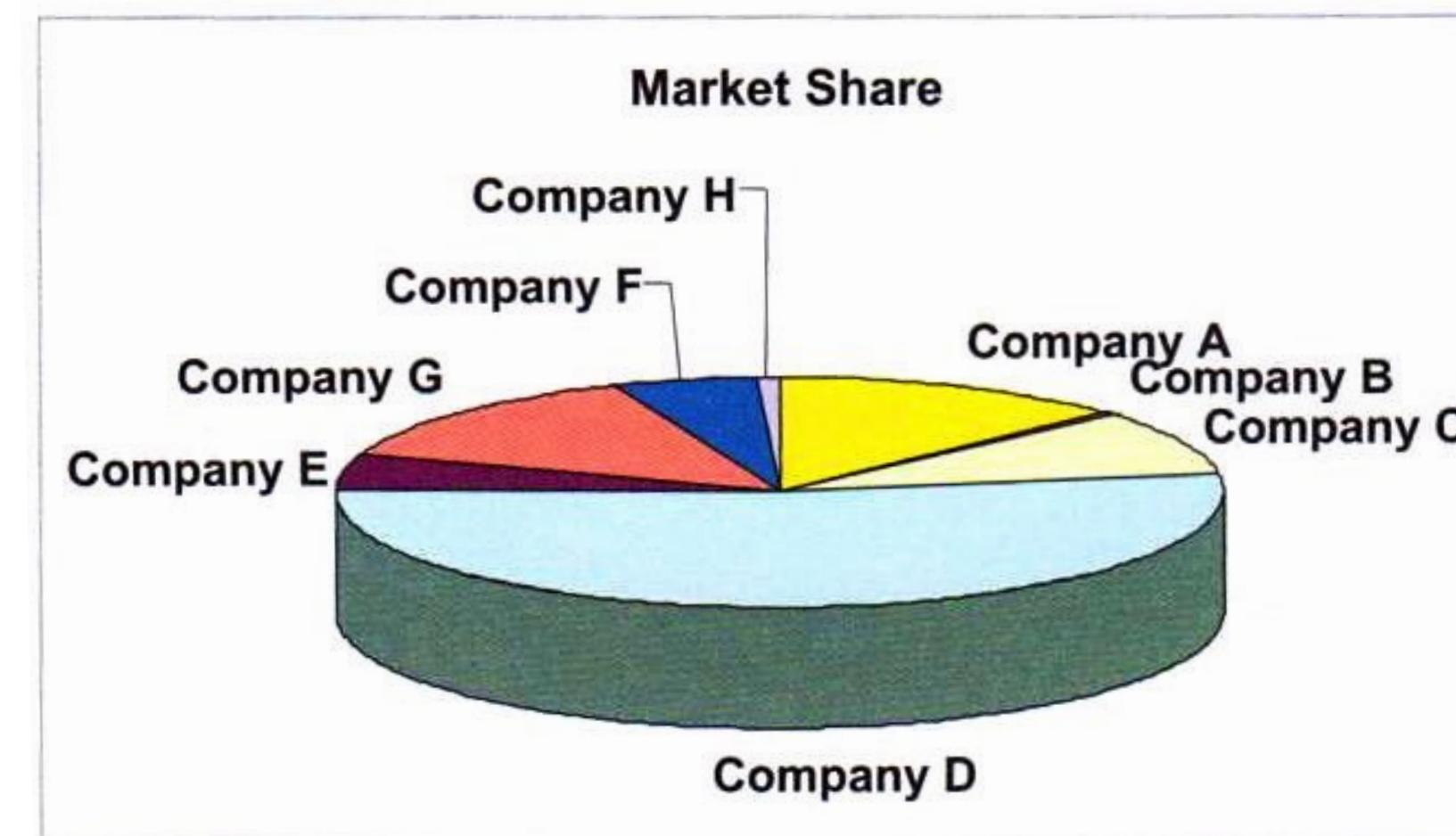


FIGURE 1.6 This is a slight variation of Figure 1.5, this time incorporating a 3-D perspective. This design is even less effective.

**A terceira dimensão auxilia na visualização?**

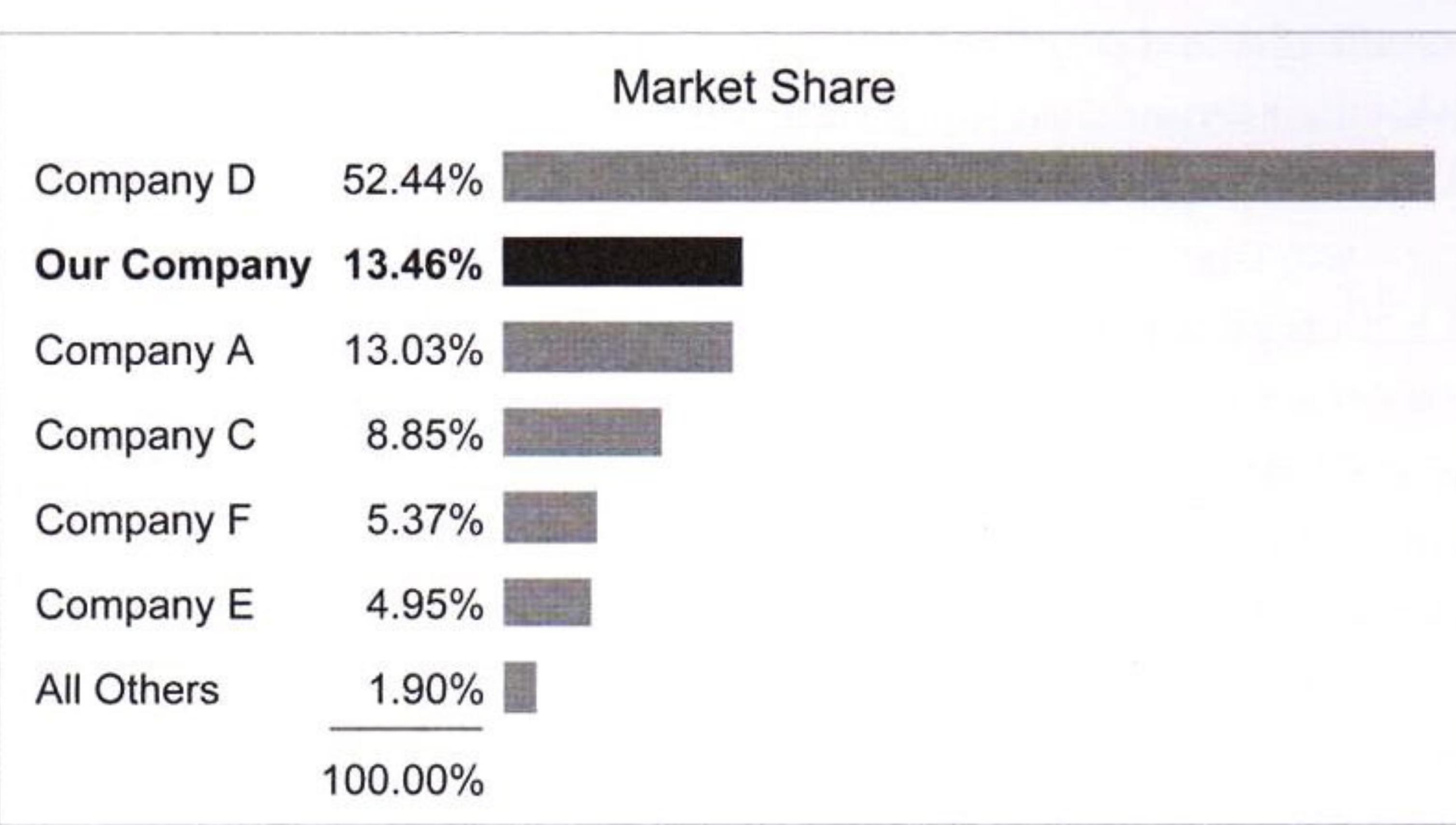


FIGURE 1.7 Percentages can be displayed in a number of ways. This example uses a traditional bar graph (a.k.a. horizontal bar) to do the job.

Existe alguma **vantagem** em se utilizar **gráficos de pizza**?

# PROPÓSITO DE TABELAS E GRÁFICOS

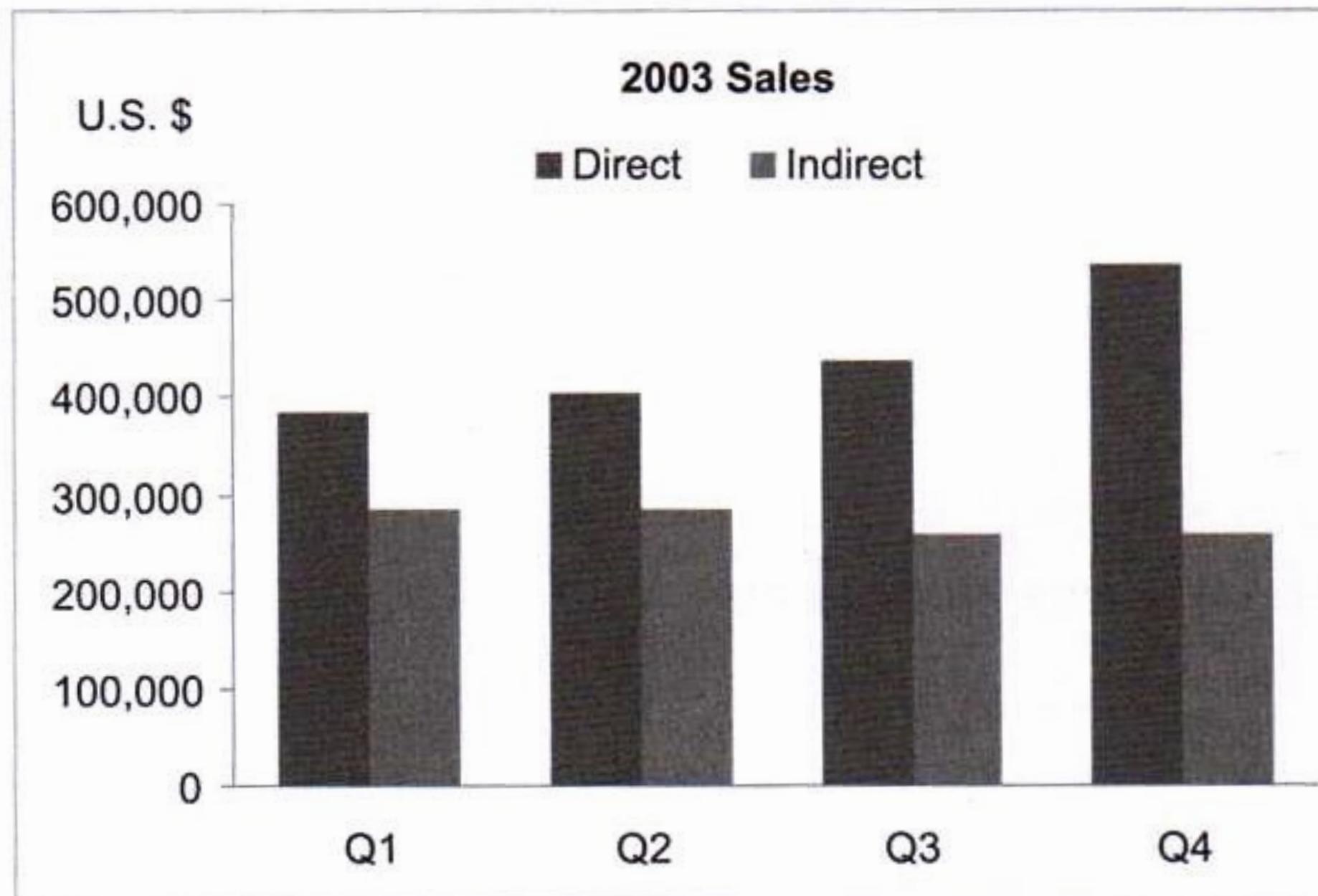
---

- Ao contrário do que dizem: “os dados **não** falam por si só”
  
- Os números têm importantes histórias a contar e cabe a nós comunicá-los de forma efetiva
  
- O propósito do uso de tabelas e gráficos é a **comunicação de dados de forma efetiva**

# INFORMAÇÃO QUANTITATIVA

---

- Dois tipos de dados:
  - Quantitativos: medem coisas
  - Categóricos: subdividem coisas medidas em grupos úteis
- Têm diferentes papéis em tabelas e gráficos e normalmente são estruturados e exibidos de formas distintas



# EXEMPLOS DE RELACIONAMENTOS QUANTITATIVOS

---

- Unidades de um produto por região geográfica
- Lucros por intervalo de tempo
- Despesas de departamentos por intervalo de tempo
- Valor gasto em marketing relacionado ao número de pedidos recebidos
- Preço de produtos relacionados ao número de unidades vendidas

# RELACIONAMENTOS QUANTITATIVOS

---

- Informações quantitativas envolvem dois tipos de dados:
  - Quantitativos
  - Categóricos
  
- Informações quantitativas sempre envolvem relacionamentos:
  - Simples entre valores quantitativos e categóricos
  - Mais complexos, entre conjuntos de valores quantitativos

# RELACIONAMENTOS QUANTITATIVOS

---

- **Quantitativo/categórico:**
  - Unidades de um produto por região geográfica
  - Lucros por intervalo de tempo
  - Despesas de departamentos por intervalo de tempo
  
- **Quantitativo/quantitativo:**
  - Valor gasto em marketing relacionado ao número de pedidos recebidos
  - Preço de produtos relacionados ao número de unidades vendidas

# RELACIONAMENTOS ENVOLVENDO CATEGORIAS

---

- **Nominais:** sub-divisões de uma categoria sem ordem intrínseca (norte, oeste, leste, etc)
- **Ordinais:** sub-divisões de uma categoria com ordenação precisa (primeiro, segundo, pequeno, médio, etc)
- **Intervalos:** sub-divisões que consistem em uma série de intervalos sequenciais que dividem um conjunto de valores quantitativos em intervalos (faixas e renda)
- **Hierárquicos:** relacionamento que envolve múltiplas categorias que estão relacionadas em diferentes níveis de organização (instituto, departamento, cursos)

# RELACIONAMENTOS ENVOLVENDO QUANTIDADES

---

- **Ranking:** sub-divisões com ordenação
- **Razão:** dois valores são comparados pela divisão de um pelo outro
- **Correlação:** dois conjuntos de valores pareados são comparados para determinar se seu crescimento ou decrescimento está relacionado

# MEDIDAS IMPORTANTES

---

## MÉDIA

- Média é um número que representa um valor central de um conjunto de valores
- **Média:** soma de todos os valores divididos pelo número de valores
- **Mediana:** dado um conjunto de valores ordenados, o valor intermediário
- **Moda:** valor mais comum
- **Valor intermediário:** valor intermediário de um intervalo

# MEDIDAS IMPORTANTES

---

## PERCENTIL

- Percentil: o  $p$ -ésimo percentil equivale ao valor que ocupa a posição dada pelo valor inteiro mais próximo a  $k=p(n-1) / 100$ , onde  $n$  é o número de valores ordenados
  
- A mediana é o 50º percentil

# MEDIDAS IMPORTANTES

Rank	Cargo	Salário anual
1	Vice-presidente	R\$475.000
2	Gerente	R\$165.000
3	Gerente	R\$165.000
4	Analista	R\$65.000
5	Artista gráfico	R\$64.000
6	Analista	R\$63.000
7	Artista gráfico	R\$62.000
8	Escritor	R\$54.000
9	Escritor	R\$52.000
10	Assistente	R43.000
11	Assistente	R\$39.000
12	Estagiário	R\$28.000
13	Estagiário	R\$25.000
MÉDIA		?????

**Média:**

**Mediana:**

**Moda:**

**Valor intermediário:**

**25º percentil:**

**75º percentil:**

# MEDIDAS IMPORTANTES

Rank	Cargo	Salário anual
1	Vice-presidente	R\$475.000
2	Gerente	R\$165.000
3	Gerente	R\$165.000
4	Analista	R\$65.000
5	Artista gráfico	R\$64.000
6	Analista	R\$63.000
7	Artista gráfico	R\$62.000
8	Escritor	R\$54.000
9	Escritor	R\$52.000
10	Assistente	R43.000
11	Assistente	R\$39.000
12	Estagiário	R\$28.000
13	Estagiário	R\$25.000
MÉDIA		?????

**Média:** R\$100.000

**Mediana:** R\$62.000

**Moda:** R\$165.000

**Valor intermediário:** R\$250.000

**25º percentil:** R\$39.000

**75º percentil:** R\$64.000

# MEDIDAS IMPORTANTES

---

## DISTRIBUIÇÃO

- Medidas do grau de variação de um conjunto de valores
- **Amplitude do intervalo:** diferença entre o maior valor do intervalo pelo menor
- **Desvio padrão:** mede como o conjunto de valores varia de acordo com a média

# MEDIDAS IMPORTANTES

---

Pedido	Transportadora 1	Transportadora 2	Transportadora 3
1	3	1	1
2	3	1	5
3	3	1	5
4	4	3	5
5	4	3	5
6	4	4	5
7	5	5	5
8	5	5	5
9	5	5	5
10	5	6	5
11	5	7	5
12	5	10	10

**Média:**

4,25; 4,25; 5

**Mediana:**

4.5, 4.5; 5

**Amplitude:**

2; 9; 9

**Desvio padrão:**

0,87; 2,70; 1,93

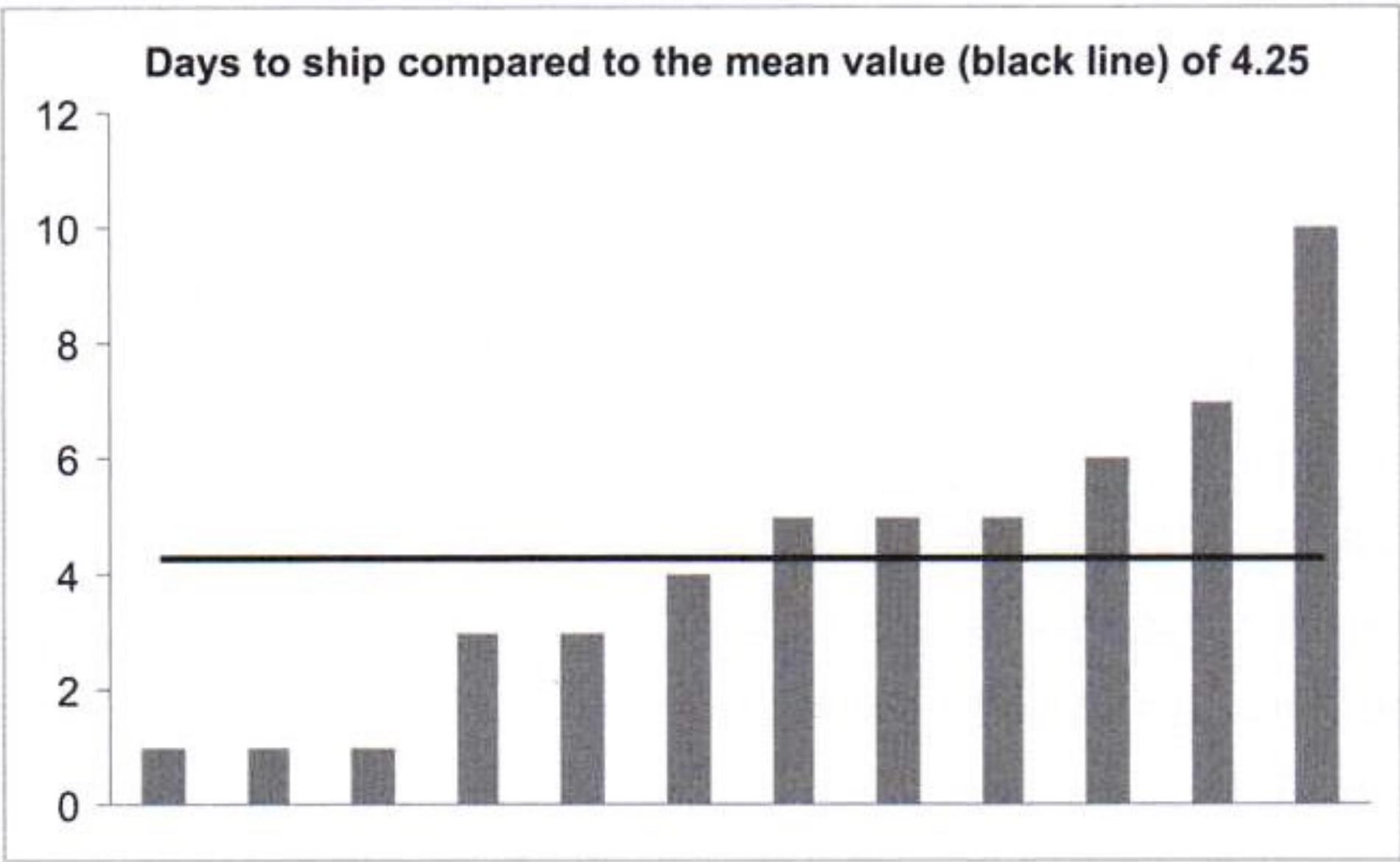


FIGURE 2.14 This graph shows a simple way to visualize the days it took Warehouse B to ship each of the 12 orders compared to the mean value of 4.25 days.

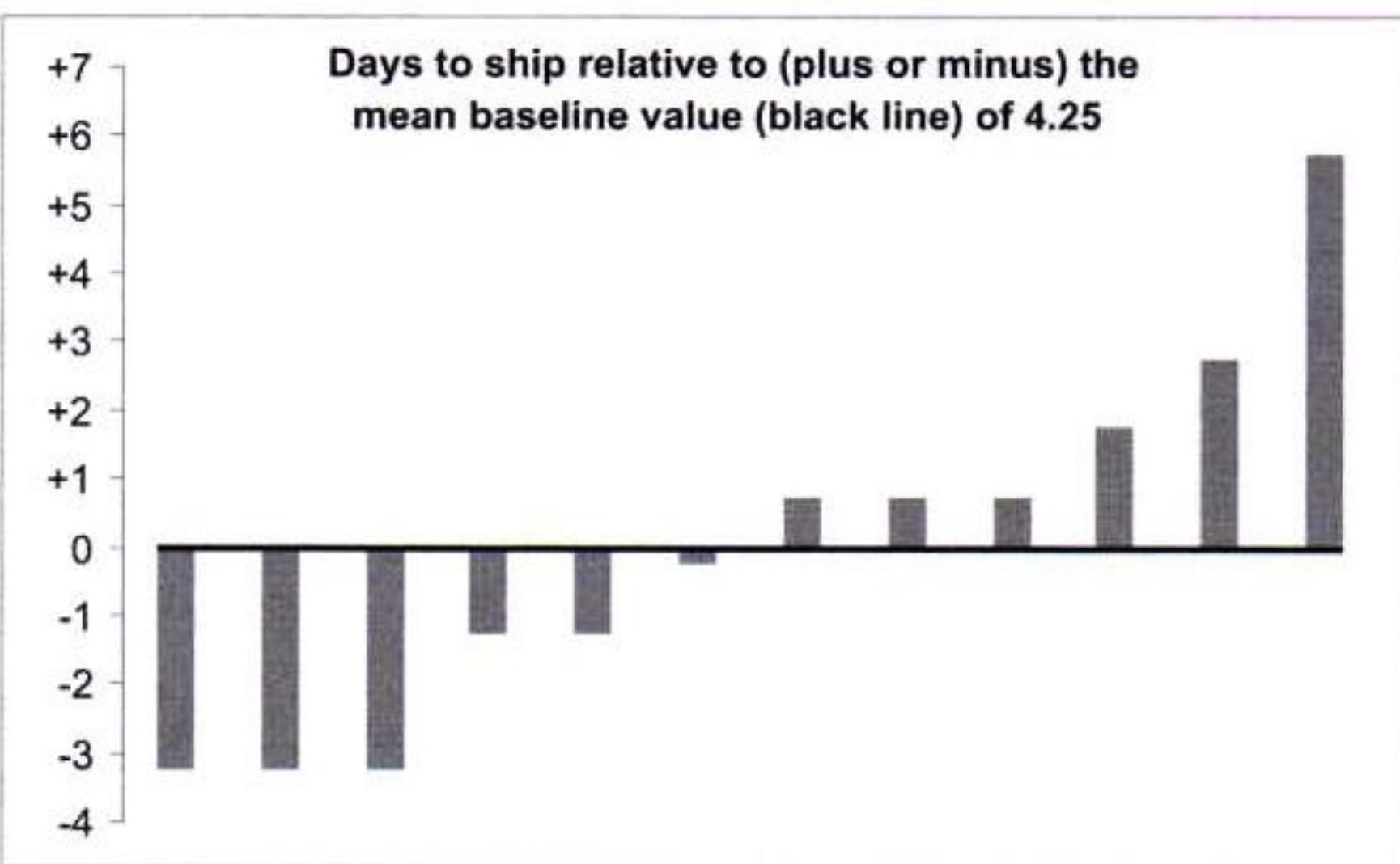


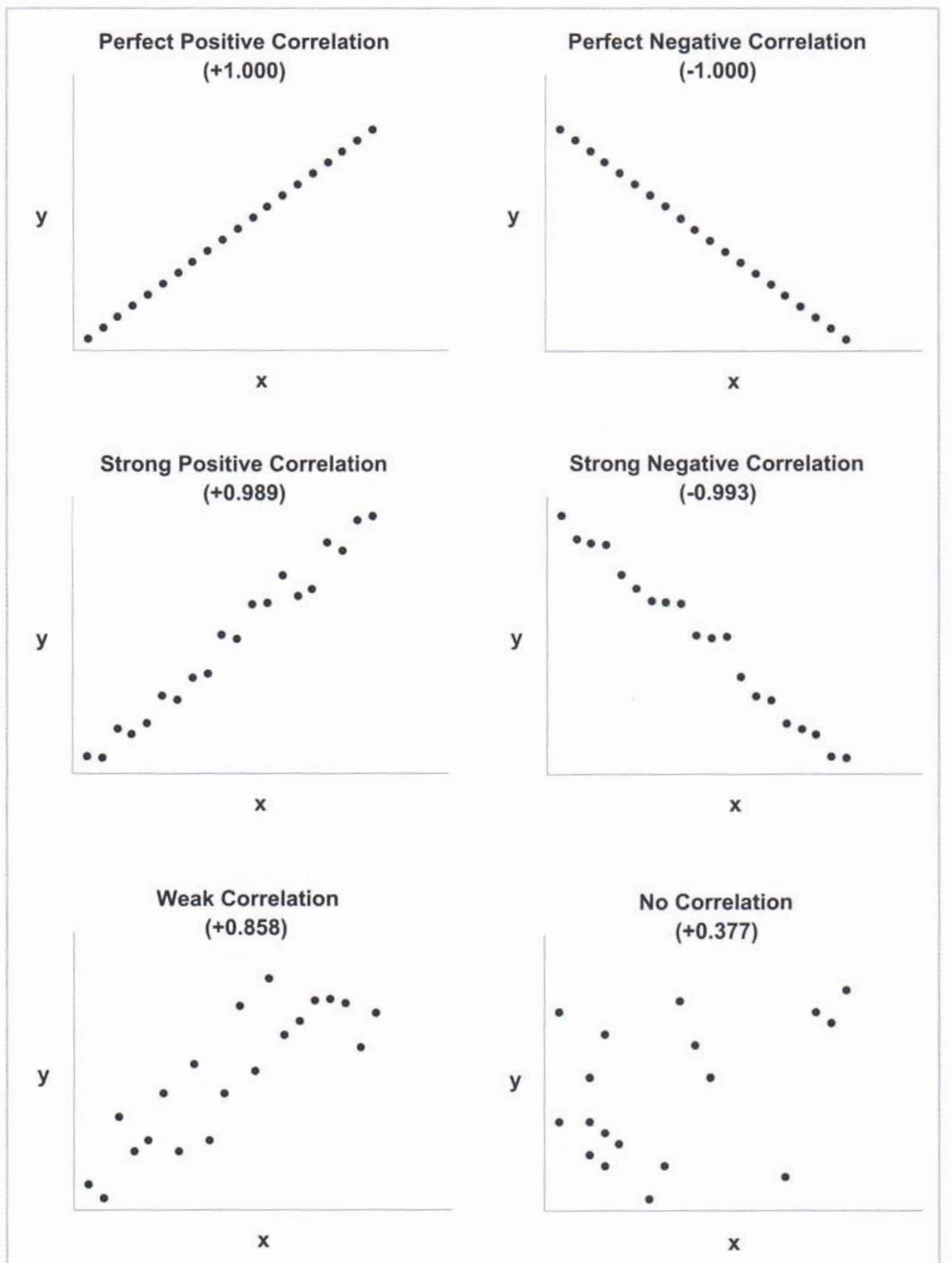
FIGURE 2.15 This graph displays the days it took to ship the individual orders relative to the mean.

# MEDIDAS IMPORTANTES

---

## CORRELAÇÃO

- **Coeficiente de correlação linear:** mede a direção (positiva ou negativa) e o grau (forte ou fraco) de relacionamento linear entre dois conjuntos pareados de valores



# MEDIDAS IMPORTANTES

---

## PROPORÇÃO

- Fração: 6/8
- Taxa: 0,75
- Porcentagem: 75%

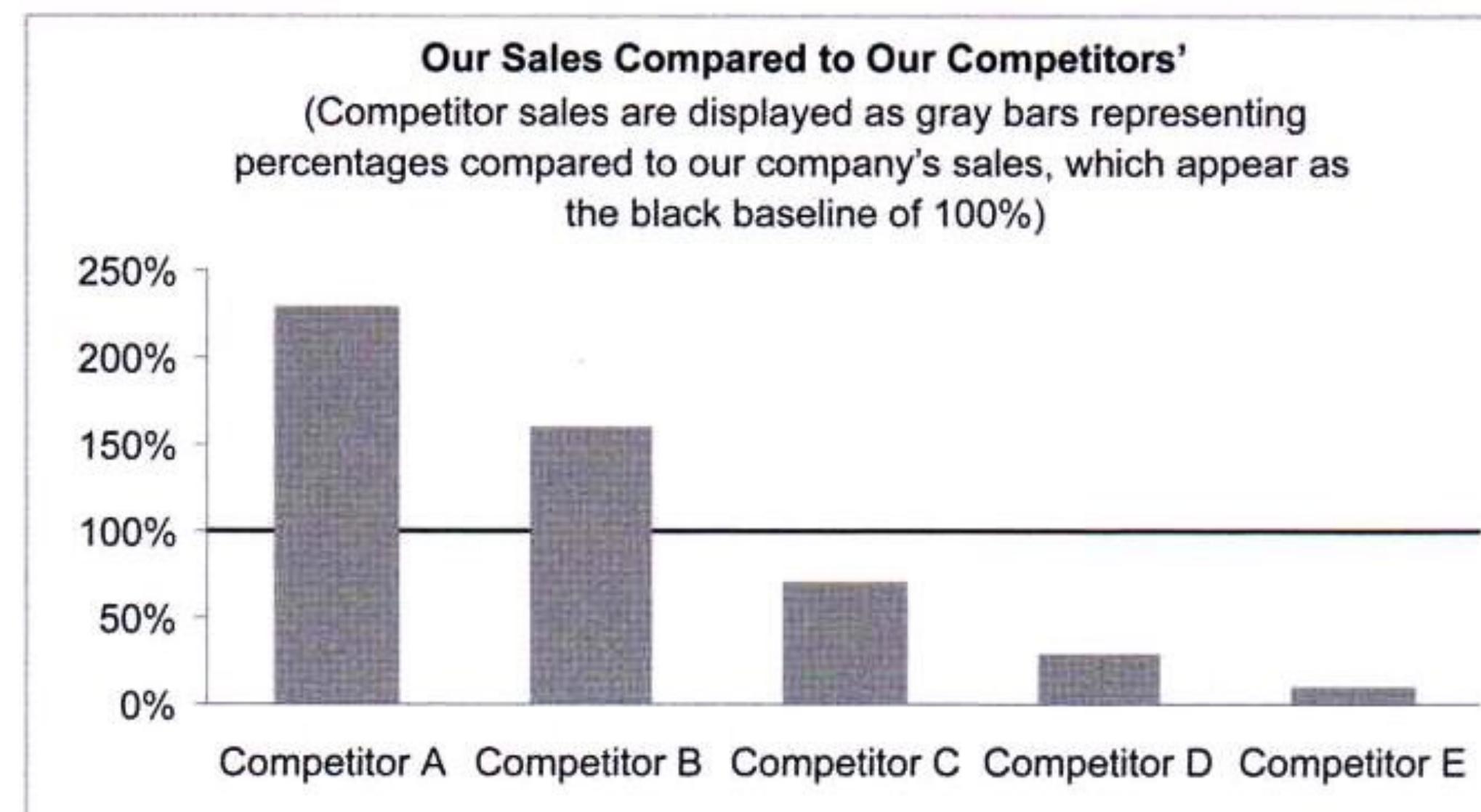


FIGURE 2.21 This graph includes a baseline of 100% for the primary set of values, making it easy to see how the other values, also expressed as percentages, differ.

# MEDIDAS IMPORTANTES

---

## MOEDA

- Ajustes sempre com base na inflação
- Conversão de múltiplas moedas na mesma moeda

# CONCEITOS FUNDAMENTAIS DE TABELAS E GRÁFICOS

---

# TABELAS

---

- Estruturas para organização e exibição de informação:
  - Dados organizados em linhas e colunas
  - Dados codificados como texto
- Apesar do uso de *grids*, o que caracteriza as tabelas é a organização dos dados em linhas e colunas
- Usadas por muitos séculos

# QUANDO USAR TABELAS

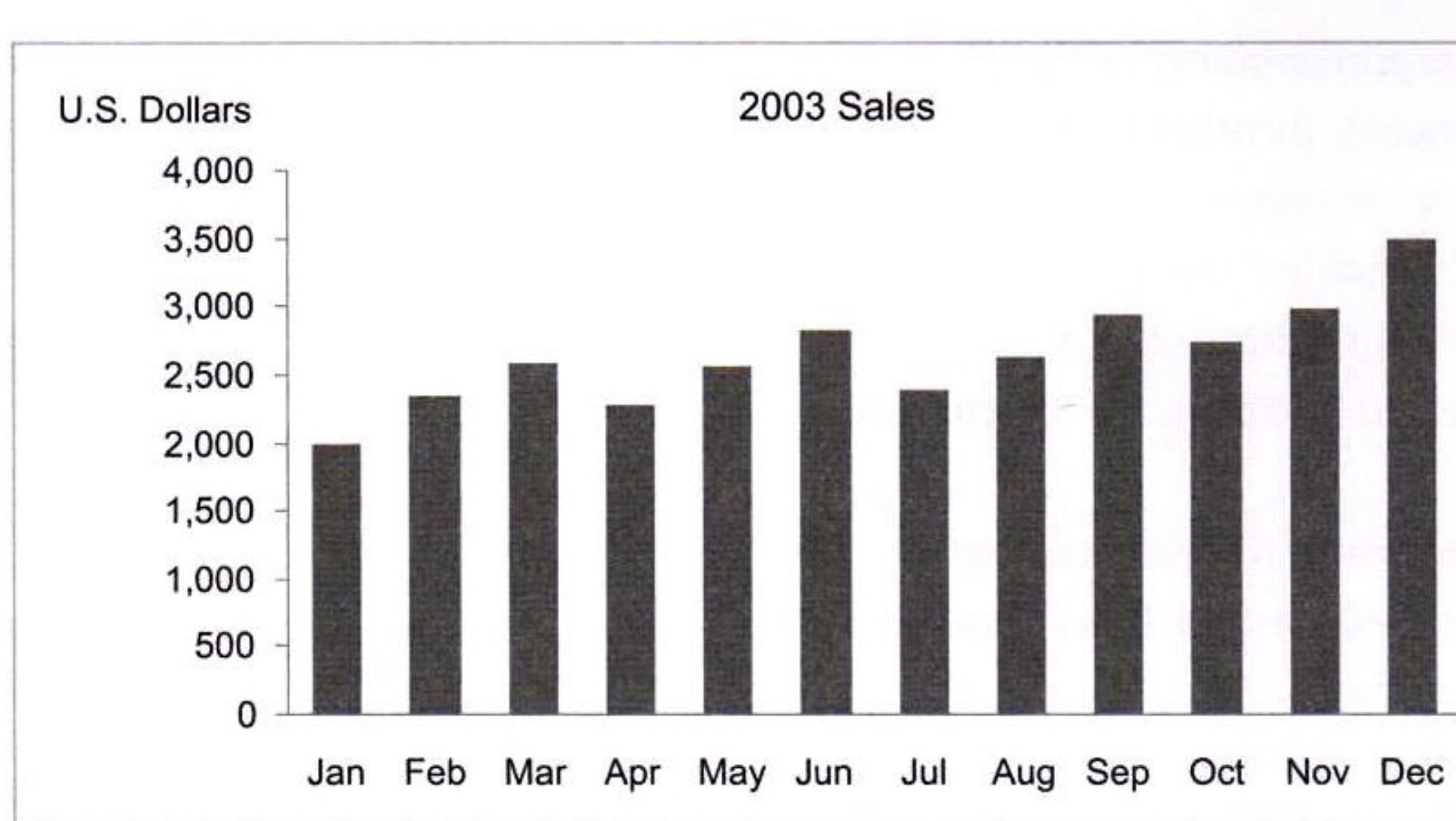
---

- Documento será utilizado para **consulta de valores individuais**
- Será utilizada para **comparação de valores individuais**
- **Valores precisos** são necessários
- A informação quantitativa envolve **mais de uma unidade de medida**

# GRÁFICOS

---

- Métodos para exibição de informação quantitativa nos quais:
  - os valores são exibidos em uma área delineada por um ou mais eixos
  - os valores são codificados por **objetos visuais** posicionados entre os eixos
  - os eixos provêm **escalas** (quantitativas ou categóricas) usadas para



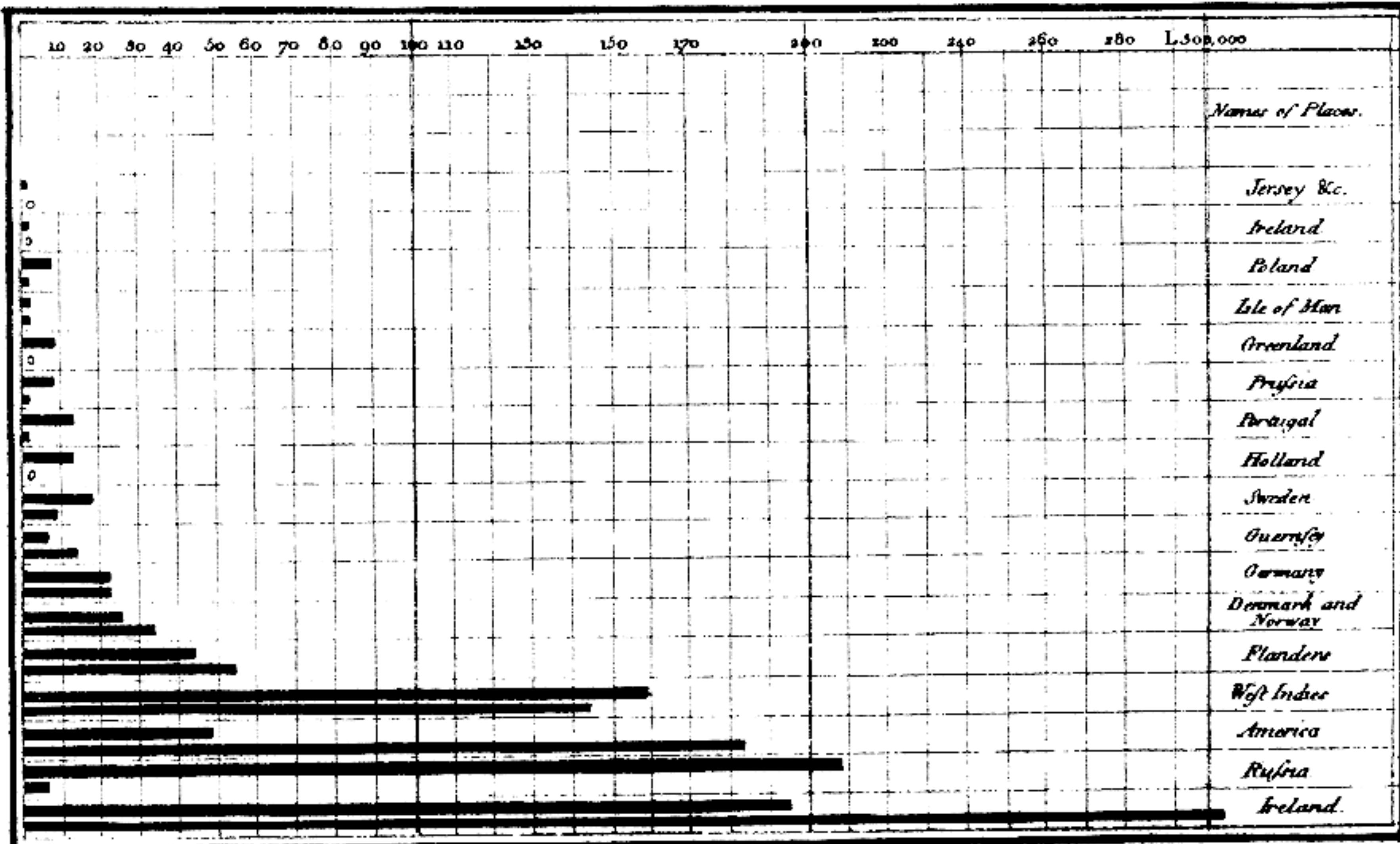
L A  
G E O M E T R I E.  
LIVRE PREMIER.

*Des problemes qu'on peut construire sans y employer que des cercles & des lignes droites.*

O u s les Problemes de Geometrie se peuvent facilement reduire a tels termes, qu'il n'est besoin par aprés que de connoître la longeur de quelques lignes droites, pour les construire.

Et comme toute l'Arithmetique n'est composée, que Comme de quatre ou cinq operations, qui sont l'Addition, la le calcul Soustraction, la Multiplication, la Division, & l'Extraction d'Arithmeti- des racines, qu'on peut prendre pour vne espece que se rapporte de Division : Ainsi n'at'on autre chose a faire en Geometrie aux opérations de touchant les lignes qu'on cherche, pour les preparer Geome- a estre connuës, que leur en adiouster d'autres, ou trie. en oster, Oubien en ayant vne, que ie nommeray l'vnité pour la rapporter d'autant mieux aux nombres, & qui peut ordinairement estre prise a discretion, puis en ayant encore deux autres, en trouuer vne quatriesme, qui soit à l'vne de ces deux, comme l'autre est à l'vnité, ce qui est le mesme que la Multiplication ; oubien en trouuer vne quatriesme, qui soit à l'vne de ces deux, comme l'vnité

Exports and Imports of SCOTLAND to and from different parts for one Year from Christmas 1780 to Christmas 1781.

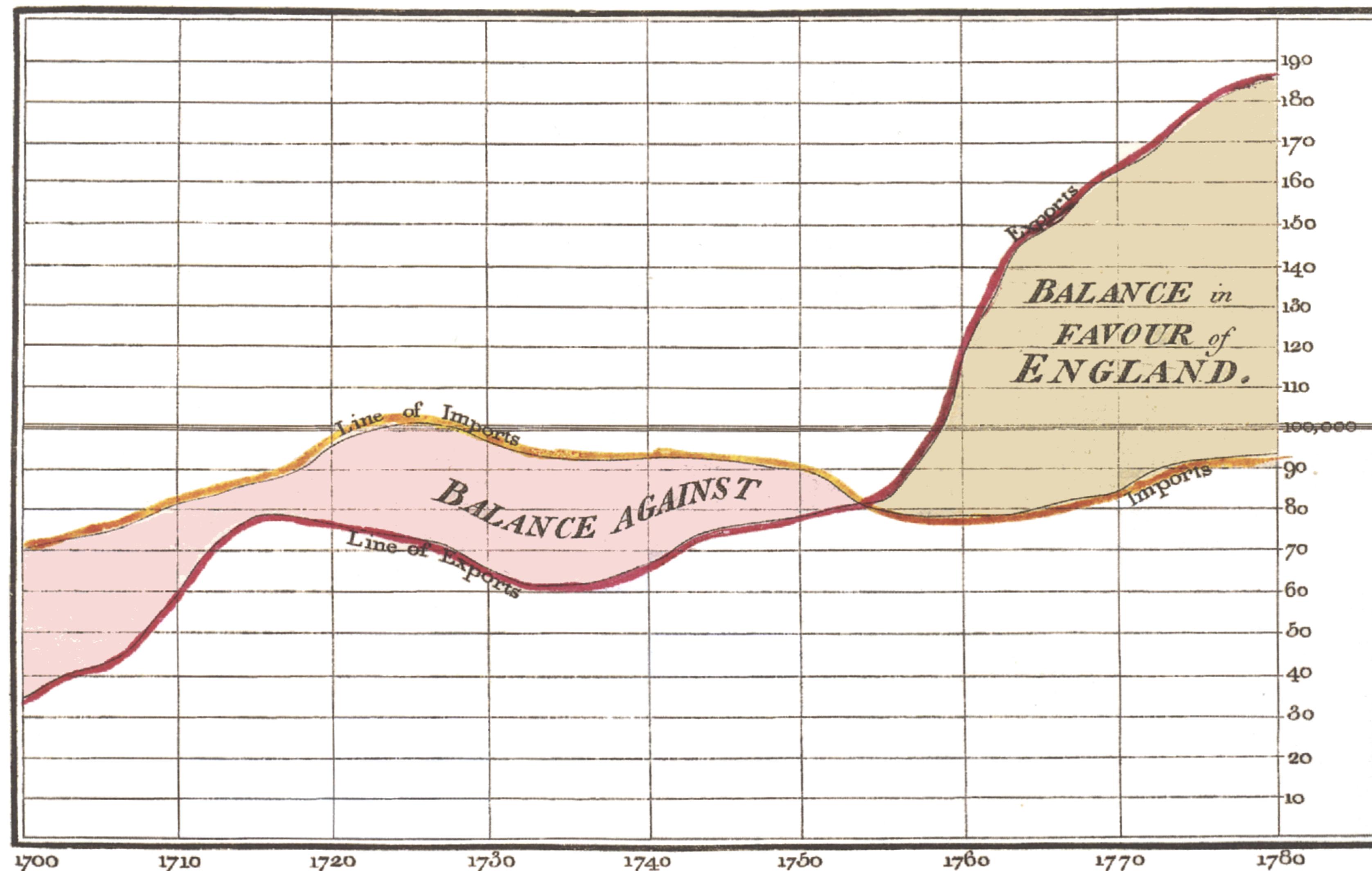


The upright divisions are Ten Thousand Pounds each. The Black Lines are Exports the Ribbed lines Imports.

Published at the Adelphi June 7<sup>th</sup> 1781 by W<sup>m</sup> Playfair

No. 100, 332 Strand, London.

Exports and Imports to and from DENMARK & NORWAY from 1700 to 1780.



*The Bottom line is divided into Years, the Right hand line into £10,000 each.*  
Published as the Act directs, 1<sup>st</sup> May 1786, by W<sup>m</sup> Playfair  
Neale sculpt<sup>r</sup> 352, Strand, London.

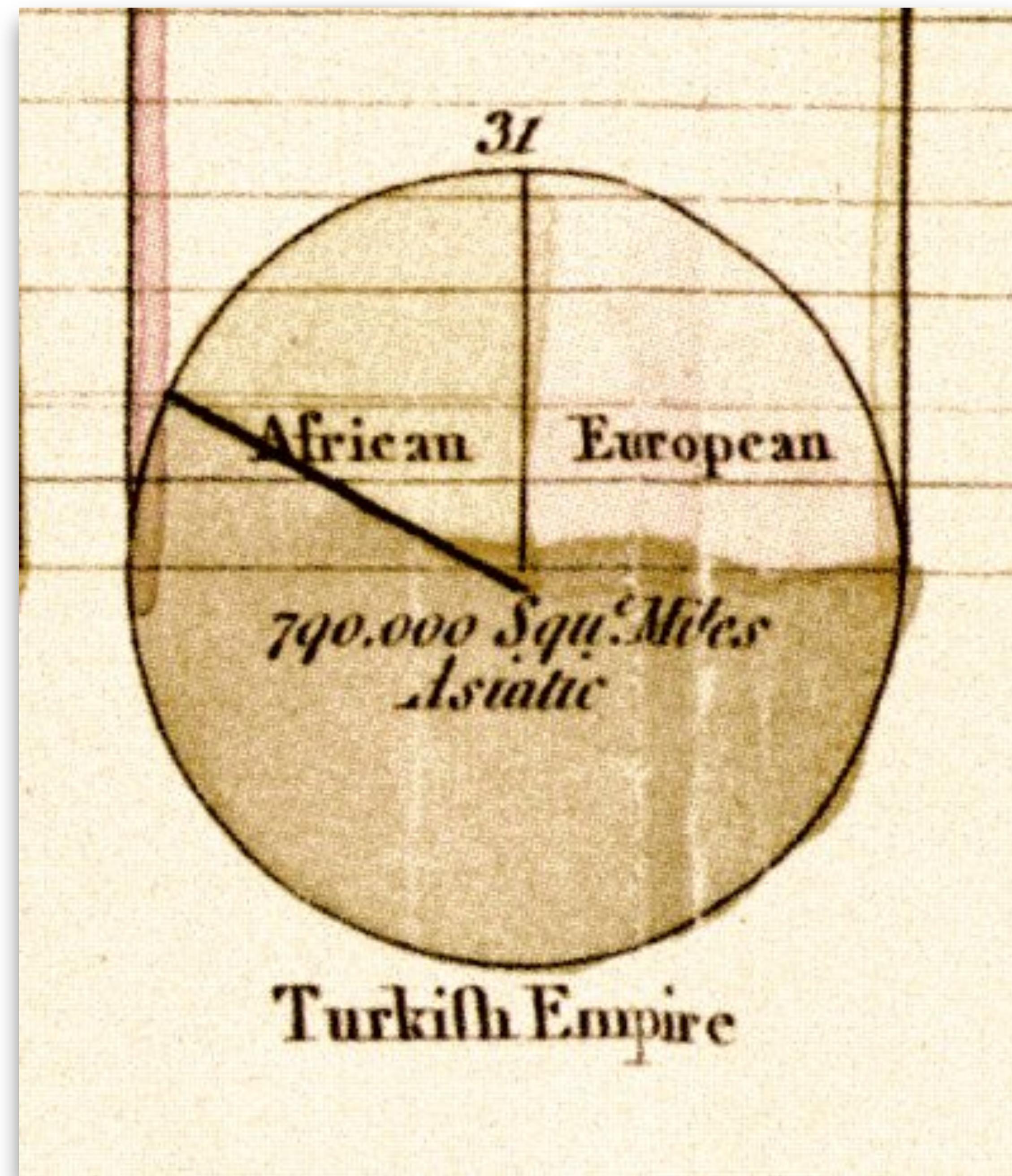


Gráfico de pizza, William Playfair, 1801

# GRAFICOS

---

- Gráficos revelam mais que uma coleção de dados, revelam também sua **forma**
  
- Permitem a **detecção** de inúmeros pontos de interesse
  
- Permitem a **comparação** de valores, identificação de **similaridades e dissimilaridades**
  
- Mostram **relacionamentos** entre diversos valores quantitativos dando a eles formas

# GRAFICOS

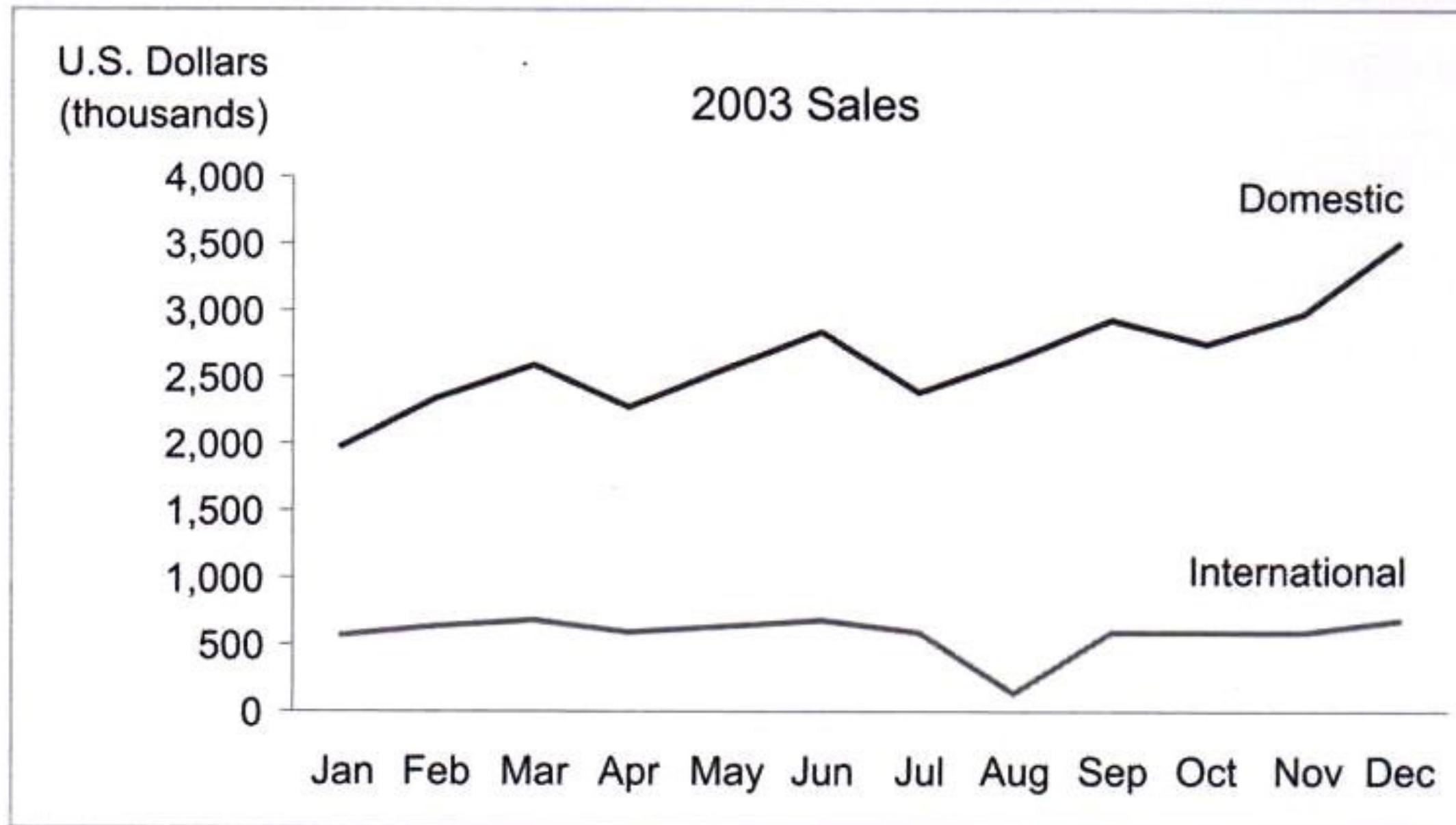


FIGURE 3.9 This fairly typical line graph is an example of the shapes and patterns in quantitative data that graphs make visible.

- A mensagem está contida na forma dos valores
- O objetivo é revelar relacionamentos entre múltiplos valores

# Tabelas

---

**Profa. Dra. Raquel Minardi**  
Departamento de Ciência da Computação  
Universidade Federal de Minas Gerais

# ONDE USAMOS TABELAS?

---

- Relatórios
- Artigos
- Software

# ESTRUTURANDO OS DADOS

---

# ORGANIZANDO OS DADOS

---

- Como organizar os dados para que eles **contem uma história?**
- As subdivisões categóricas devem ser organizadas em **linhas ou colunas?**
- Os dados deveriam ser separados em **subgrupos com quebras entre eles?**
- Os dados deveriam ser **ordenados de alguma forma particular?**

Product	Regions			
	North	East	South	West
Product 01	94	152	174	87
Product 02	122	198	226	113
Product 03	101	164	188	94
Product 04	142	230	263	131
Product 05	132	214	244	122
Product 06	174	282	323	161
Product 07	401	648	742	371
Product 08	281	454	519	260
Product 09	112	182	208	104
Product 10	584	944	1,081	540
Product 11	543	878	1,005	502
Product 12	163	263	301	151
Product 13	489	790	904	452
Product 14	327	529	606	303
Product 15	295	476	545	273
Total	3,960	6,403	7,330	3,665

Arranjo bidimensional

Region	Product	Units Sold
Region 01	Product 01	152
	Product 02	198
	Product 03	164
	Product 04	230
	Product 05	214
	Product 06	282
	Product 07	648
	Product 08	454
	Product 09	182
	Product 10	944
	Product 11	878
	Product 12	263
	Product 13	790
	Product 14	529
	Product 15	476
Region 02	Product 01	443
	Product 02	133
	Product 03	399

Arranjo unidimensional

# COLUMNAS OU LINHAS?

---

- Quantas subdivisões cada categoria contém?
  - Se são poucas, você tem a opção de organizá-las em colunas
  - Se são muitas, organizá-las em uma única coluna
- Qual o número máximo de **caracteres** por subdivisão?
  - Se alguma das subdivisões possui texto longo, provavelmente será necessário o uso do arranjo unidimensional
- Alguma das subdivisões envolvem **séries temporais** ou **rankings**?
  - Séries temporais devem ser organizadas em colunas
  - Ranking devem ser organizados em linhas funciona melhor

	1992				1993	
Region	Q1	Q2	Q3	Q4	Q1	Q2
North	393	473	539	639	439	538
East	326	393	447	530	364	447
South	401	483	550	652	448	549
West	538	647	737	874	601	736
Total	1,658	1,996	2,274	2,696	1,852	2,270

Year	Qtr	Region					Total
		North	East	South	West		
1992	1	393	326	401	538		1,658
	2	473	393	483	647		1,996
	3	539	447	550	737		2,273
	4	639	530	652	874		2,695
1993	1	439	364	448	601		1,852
	2	538	447	549	736		2,270

Rank	Product	Sales (U.S. \$)
1	Product J	1,939,993
2	Product E	1,784,794
3	Product G	1,642,010
4	Product A	1,510,649
5	Product D	1,389,797
6	Product C	1,278,614
7	Product B	1,176,324
8	Product H	1,082,219
9	Product F	995,641
10	Product I	915,990

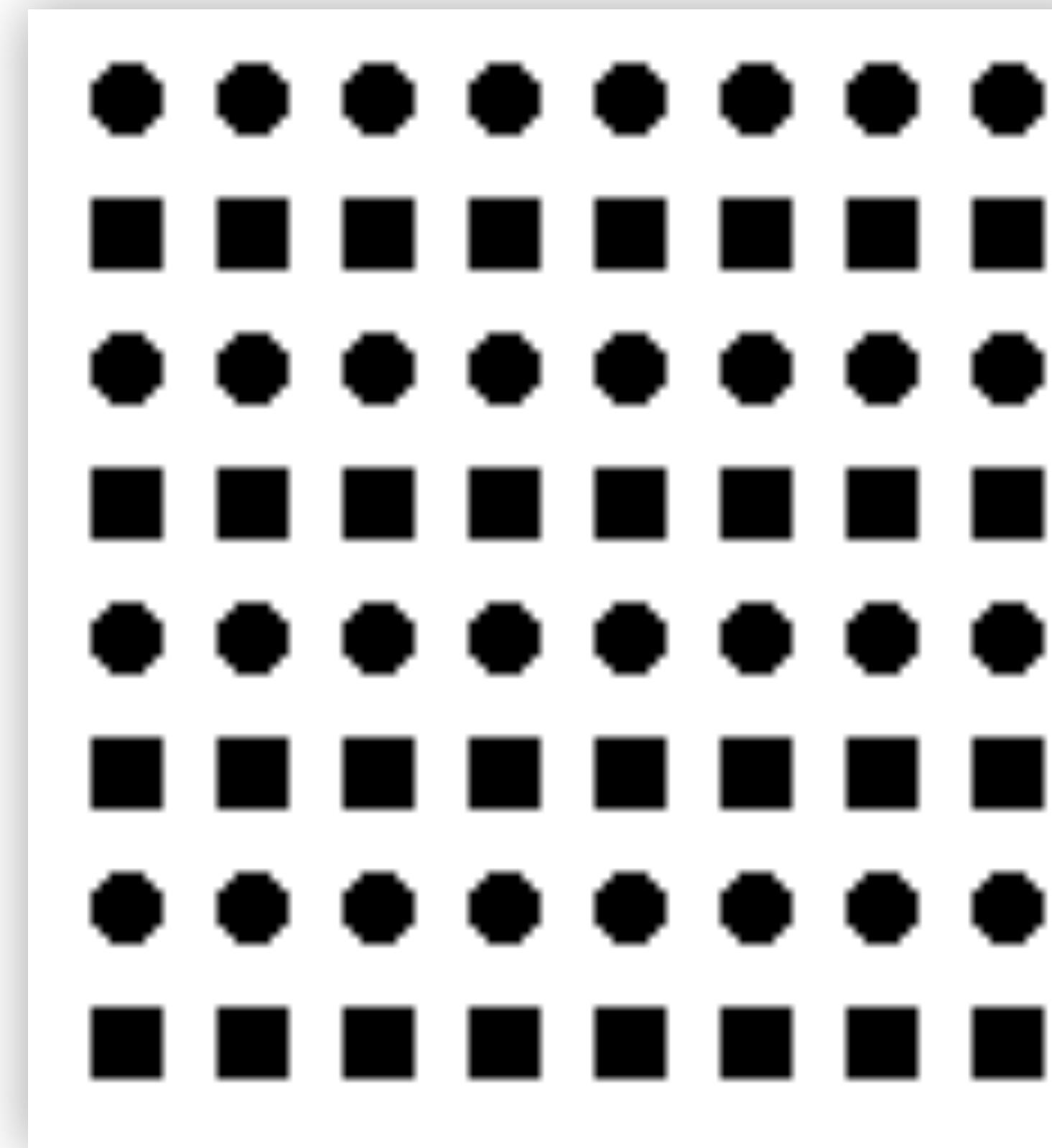
Product	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Product 01	93,993	84,773	88,833	95,838	93,874	83,994	84,759	92,738	93,728	93,972	93,772	99,837
Product 02	87,413	78,839	82,615	89,129	87,303	78,114	78,826	86,246	87,167	87,394	87,208	92,848
Product 03	90,036	81,204	85,093	91,803	89,922	80,458	81,191	88,834	89,782	90,016	89,824	95,634
Product 04	92,737	83,640	87,646	94,557	92,620	82,872	83,626	91,499	92,476	92,716	92,519	98,503
Product 05	86,245	77,785	81,511	87,938	86,136	77,071	77,773	85,094	86,002	86,226	86,043	91,608
Product 06	88,833	80,119	83,956	90,576	88,720	79,383	80,106	87,647	88,582	88,813	88,624	94,356
Product 07	82,614	74,511	78,079	84,236	82,510	73,826	74,498	81,511	82,382	82,596	82,420	87,751
Product 08	85,093	76,746	80,421	86,763	84,985	76,041	76,733	83,957	84,853	85,074	84,893	90,384
Product 09	87,646	79,048	82,834	89,366	87,535	78,322	79,035	86,475	87,399	87,626	87,440	93,095
Product 10	90,275	81,420	85,319	92,047	90,161	80,672	81,406	89,070	90,021	90,255	90,063	95,888

Product	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Product 01	93,993	84,773	88,833	95,838	93,874	83,994	84,759	92,738	93,728	93,972	93,772	99,837
Product 02	87,413	78,839	82,615	89,129	87,303	78,114	78,826	86,246	87,167	87,394	87,208	92,848
Product 03	90,036	81,204	85,093	91,803	89,922	80,458	81,191	88,834	89,782	90,016	89,824	95,634
Product 04	92,737	83,640	87,646	94,557	92,620	82,872	83,626	91,499	92,476	92,716	92,519	98,503
Product 05	86,245	77,785	81,511	87,938	86,136	77,071	77,773	85,094	86,002	86,226	86,043	91,608
Product 06	88,833	80,119	83,956	90,576	88,720	79,383	80,106	87,647	88,582	88,813	88,624	94,356
Product 07	82,614	74,511	78,079	84,236	82,510	73,826	74,498	81,511	82,382	82,596	82,420	87,751
Product 08	85,093	76,746	80,421	86,763	84,985	76,041	76,733	83,957	84,853	85,074	84,893	90,384
Product 09	87,646	79,048	82,834	89,366	87,535	78,322	79,035	86,475	87,399	87,626	87,440	93,095
Product 10	90,275	81,420	85,319	92,047	90,161	80,672	81,406	89,070	90,021	90,255	90,063	95,888

# LEI DA SIMILARIDADE

---

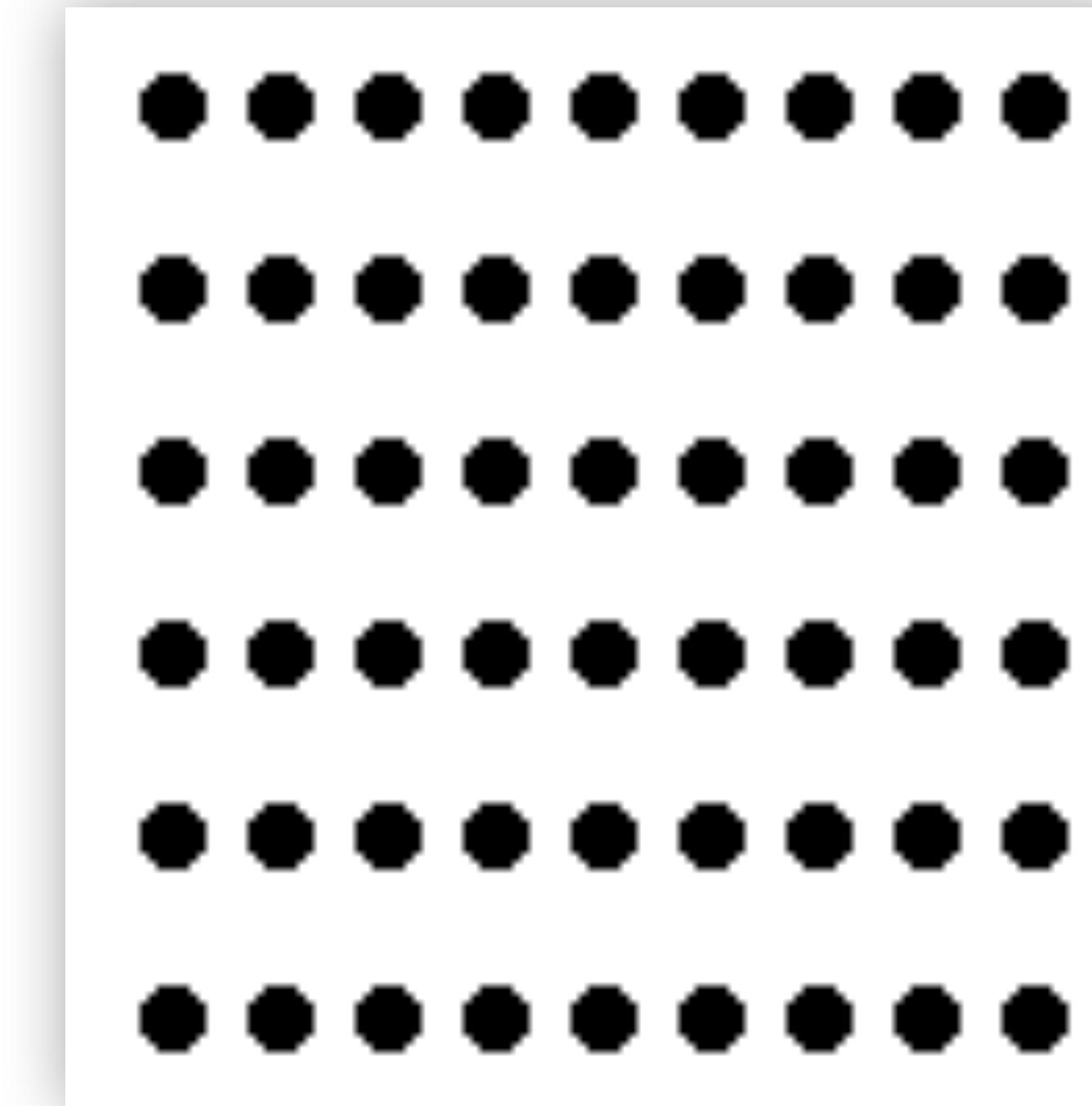
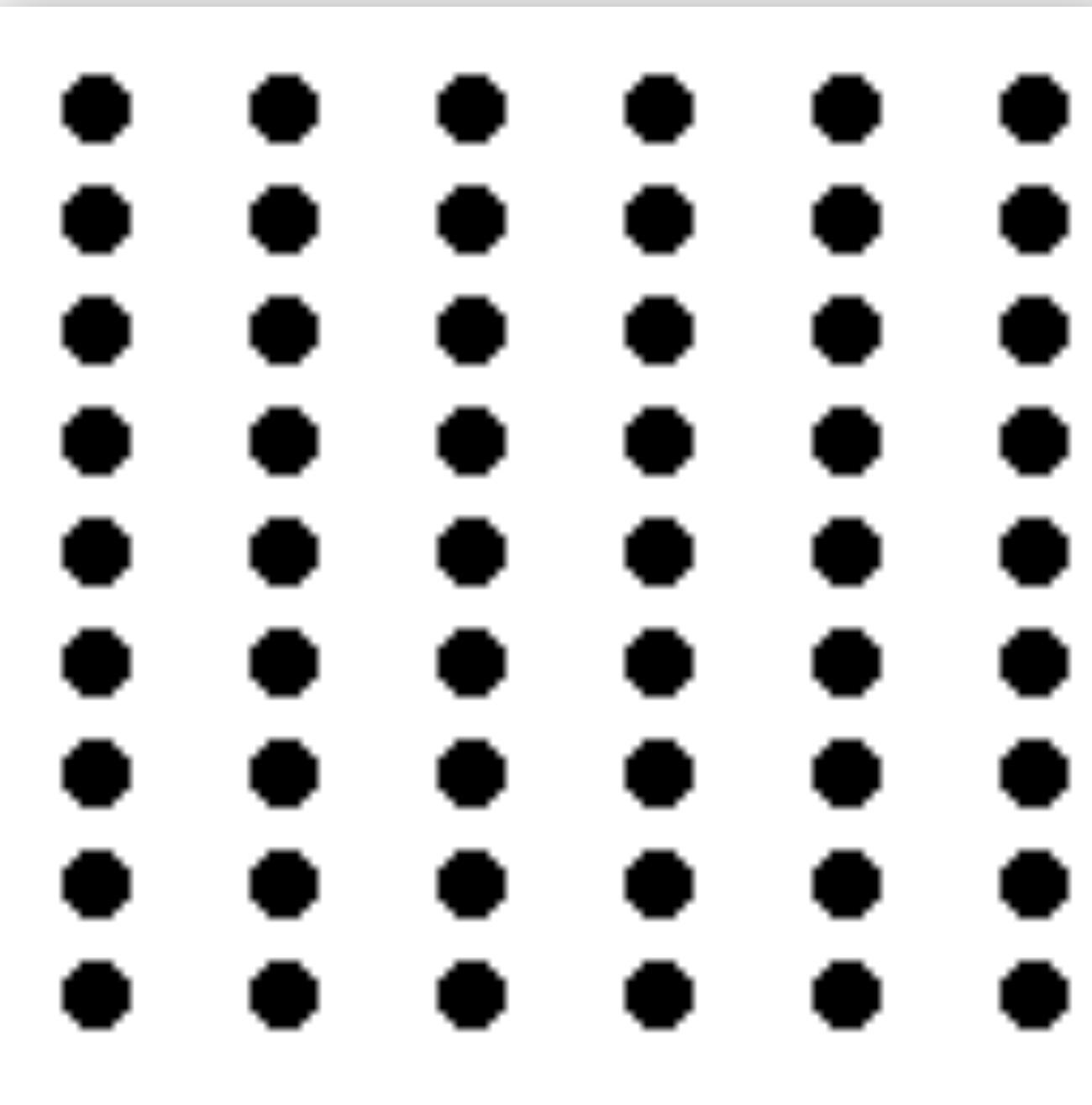
- A mente agrupa elementos similares em entidades coletivas
- Esta similaridade pode ser baseada em forma, cores, tamanho ou brilho



# LEI DA PROXIMIDADE

---

- Características mais próximas são mais rapidamente associadas



# USO DO ESPAÇO EM BRANCO

---

- A quantidade de espaços em branco pode ser manipulada para direcionamento dos olhos do usuário entre linhas e colunas
- Se você deseja que o usuário avalie os dados coluna a coluna, privilegie o espaço entre colunas em detrimento do espaço entre linhas
- Se você deseja que o usuário avalie os dados linha a linha, faça o oposto
  
- Se não há espaço suficiente, você deve usar outro método

Product	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Product 01	93,993	84,773	88,833	95,838	93,874	83,994	84,759	92,738	93,728	93,972	93,772	99,837
Product 02	87,413	78,839	82,615	89,129	87,303	78,114	78,826	86,246	87,167	87,394	87,208	92,848
Product 03	90,036	81,204	85,093	91,803	89,922	80,458	81,191	88,834	89,782	90,016	89,824	95,634
Product 04	92,737	83,640	87,646	94,557	92,620	82,872	83,626	91,499	92,476	92,716	92,519	98,503
Product 05	86,245	77,785	81,511	87,938	86,136	77,071	77,773	85,094	86,002	86,226	86,043	91,608
Product 06	88,833	80,119	83,956	90,576	88,720	79,383	80,106	87,647	88,582	88,813	88,624	94,356
Product 07	82,614	74,511	78,079	84,236	82,510	73,826	74,498	81,511	82,382	82,596	82,420	87,751
Product 08	85,093	76,746	80,421	86,763	84,985	76,041	76,733	83,957	84,853	85,074	84,893	90,384
Product 09	87,646	79,048	82,834	89,366	87,535	78,322	79,035	86,475	87,399	87,626	87,440	93,095
Product 10	90,275	81,420	85,319	92,047	90,161	80,672	81,406	89,070	90,021	90,255	90,063	95,888

Product	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Product 01	93,993	84,773	88,833	95,838	93,874	83,994	84,759	92,738	93,728	93,972	93,772	99,837
Product 02	87,413	78,839	82,615	89,129	87,303	78,114	78,826	86,246	87,167	87,394	87,208	92,848
Product 03	90,036	81,204	85,093	91,803	89,922	80,458	81,191	88,834	89,782	90,016	89,824	95,634
Product 04	92,737	83,640	87,646	94,557	92,620	82,872	83,626	91,499	92,476	92,716	92,519	98,503
Product 05	86,245	77,785	81,511	87,938	86,136	77,071	77,773	85,094	86,002	86,226	86,043	91,608
Product 06	88,833	80,119	83,956	90,576	88,720	79,383	80,106	87,647	88,582	88,813	88,624	94,356
Product 07	82,614	74,511	78,079	84,236	82,510	73,826	74,498	81,511	82,382	82,596	82,420	87,751
Product 08	85,093	76,746	80,421	86,763	84,985	76,041	76,733	83,957	84,853	85,074	84,893	90,384
Product 09	87,646	79,048	82,834	89,366	87,535	78,322	79,035	86,475	87,399	87,626	87,440	93,095
Product 10	90,275	81,420	85,319	92,047	90,161	80,672	81,406	89,070	90,021	90,255	90,063	95,888

Product	Jan	Feb	Mar	Apr	May	Jun
Product 01	93,993	84,773	88,833	95,838	93,874	83,994
Product 02	87,413	78,839	82,615	89,129	87,303	78,114
Product 03	90,036	81,204	85,093	91,803	89,922	80,458
Product 04	92,737	83,640	87,646	94,557	92,620	82,872
Product 05	83,733	75,520	79,137	85,377	83,627	74,826
Total	447,913	403,976	423,323	456,705	447,346	400,264

Product	Jan	Feb	Mar	Apr	May	Jun
Product 01	93,993	84,773	88,833	95,838	93,874	83,994
Product 02	87,413	78,839	82,615	89,129	87,303	78,114
Product 03	90,036	81,204	85,093	91,803	89,922	80,458
Product 04	92,737	83,640	87,646	94,557	92,620	82,872
Product 05	86,245	77,785	81,511	87,938	86,136	77,071
Total	450,425	406,241	425,697	459,266	449,854	402,508

Product	Jan	Feb	Mar	Apr	May	Jun
Product 01	93,993	84,773	88,833	95,838	93,874	83,994
Product 02	87,413	78,839	82,615	89,129	87,303	78,114
Product 03	90,036	81,204	85,093	91,803	89,922	80,458
Product 04	92,737	83,640	87,646	94,557	92,620	82,872
Product 05	83,733	75,520	79,137	85,377	83,627	74,826
Total	447,913	403,976	423,323	456,705	447,346	400,264

Product	Jan	Feb	Mar	Apr	May	Jun
Product 01	93,993	84,773	88,833	95,838	93,874	83,994
Product 02	87,413	78,839	82,615	89,129	87,303	78,114
Product 03	90,036	81,204	85,093	91,803	89,922	80,458
Product 04	92,737	83,640	87,646	94,557	92,620	82,872
Product 05	86,245	77,785	81,511	87,938	86,136	77,071
Total	450,425	406,241	425,697	459,266	449,854	402,508

Product	Jan	Feb	Mar	Apr	May	Jun
Product 01	93,993	84,773	88,833	95,838	93,874	83,994
Product 02	87,413	78,839	82,615	89,129	87,303	78,114
Product 03	90,036	81,204	85,093	91,803	89,922	80,458
Product 04	92,737	83,640	87,646	94,557	92,620	82,872
Product 05	83,733	75,520	79,137	85,377	83,627	74,826
Total	447,913	403,976	423,323	456,705	447,346	400,264

Product	Jan	Feb	Mar	Apr	May	Jun
Product 01	93,993	84,773	88,833	95,838	93,874	83,994
Product 02	87,413	78,839	82,615	89,129	87,303	78,114
Product 03	90,036	81,204	85,093	91,803	89,922	80,458
Product 04	92,737	83,640	87,646	94,557	92,620	82,872
Product 05	83,733	75,520	79,137	85,377	83,627	74,826
Total	447,913	403,976	423,323	456,705	447,346	400,264

# GRIDS E LINHAS

---

- Podem ser usados para:
  - Delimitar linhas e colunas
  - Agrupar subconjuntos de dados
  - Ressaltar subconjuntos de dados
- No exemplo do *slide* anterior, nada mais é necessário além da linha separando o cabeçalho e o rodapé
- Linhas delimitadoras da tabela como um todo também são desnecessárias, exceto quando a tabela está cercada de outros objetos e não há espaço suficiente para usar espaço em branco para destacá-la

Product	Jan	Feb	Mar	Apr	May	Jun	Total
Product 01	93,993	84,773	88,833	95,838	93,874	83,994	541,305
Product 02	87,413	78,839	82,615	89,129	87,303	78,114	503,414
Product 03	90,036	81,204	85,093	91,803	89,922	80,458	518,516
Product 04	92,737	83,640	87,646	94,557	92,620	82,872	534,072
Product 05	83,733	75,520	79,137	85,377	83,627	74,826	482,220
Total	447,913	403,976	423,323	456,705	447,346	400,264	2,579,526

Quando usar grids e linhas,  
**certifique-se de que as linhas são finas e claras**

Product	Jan	Feb	Mar	Apr	May	Jun	Total
Product 01	93,993	84,773	88,833	95,838	93,874	83,994	541,305
Product 02	87,413	78,839	82,615	89,129	87,303	78,114	503,414
Product 03	90,036	81,204	85,093	91,803	89,922	80,458	518,516
Product 04	92,737	83,640	87,646	94,557	92,620	82,872	534,072
Product 05	83,733	75,520	79,137	85,377	83,627	74,826	482,220
Total	447,913	403,976	423,323	456,705	447,346	400,264	2,579,526

Product	Jan	Feb	Mar	Apr	May	Jun	Total
Product 01	93,993	84,773	88,833	95,838	93,874	83,994	541,305
Product 02	87,413	78,839	82,615	89,129	87,303	78,114	503,414
Product 03	90,036	81,204	85,093	91,803	89,922	80,458	518,516
Product 04	92,737	83,640	87,646	94,557	92,620	82,872	534,072
Product 05	83,733	75,520	79,137	85,377	83,627	74,826	482,220
Total	447,913	403,976	423,323	456,705	447,346	400,264	2,579,526

# CORES E PREENCHIMENTO

---

- Quando espaço em branco não pode ser usado para delinear colunas ou linhas efetivamente, o uso de **cores suaves** para preenchimento do fundo é mais interessante que o uso de grids

Product	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Product 01	93,993	84,773	88,833	95,838	93,874	83,994	84,759	92,738	93,728	93,972	93,772	99,837
Product 02	87,413	78,839	82,615	89,129	87,303	78,114	78,826	86,246	87,167	87,394	87,208	92,848
Product 03	90,036	81,204	85,093	91,803	89,922	80,458	81,191	88,834	89,782	90,016	89,824	95,634
Product 04	92,737	83,640	87,646	94,557	92,620	82,872	83,626	91,499	92,476	92,716	92,519	98,503
Product 05	86,245	77,785	81,511	87,938	86,136	77,071	77,773	85,094	86,002	86,226	86,043	91,608
Product 06	88,833	80,119	83,956	90,576	88,720	79,383	80,106	87,647	88,582	88,813	88,624	94,356
Product 07	82,614	74,511	78,079	84,236	82,510	73,826	74,498	81,511	82,382	82,596	82,420	87,751
Product 08	85,093	76,746	80,421	86,763	84,985	76,041	76,733	83,957	84,853	85,074	84,893	90,384
Product 09	87,646	79,048	82,834	89,366	87,535	78,322	79,035	86,475	87,399	87,626	87,440	93,095
Product 10	90,275	81,420	85,319	92,047	90,161	80,672	81,406	89,070	90,021	90,255	90,063	95,888

Product	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Product 01	93,993	84,773	88,833	95,838	93,874	83,994	84,759	92,738	93,728	93,972	93,772	99,837
Product 02	557,378	502,704	526,780	568,319	556,673	498,084	502,621	549,936	555,807	557,254	556,068	592,033
Product 03	574,100	517,785	542,583	585,369	573,373	513,027	517,699	566,434	572,481	573,972	572,750	609,794
Product 04	304,273	274,426	287,569	310,246	303,888	271,904	274,381	300,210	303,415	304,205	303,557	323,191
Product 05	30,427	27,443	28,757	31,025	30,389	27,190	27,438	30,021	30,342	30,420	30,356	32,319
Product 06	305,186	275,249	288,432	311,176	304,799	272,720	275,204	301,111	304,325	305,118	304,468	324,161
Product 07	6,104	5,505	5,769	6,224	6,096	5,454	5,504	6,022	6,087	6,102	6,089	6,483
Product 08	61,220	55,215	57,859	62,422	61,143	54,708	55,206	60,403	61,048	61,207	61,076	65,027
Product 09	50,813	45,828	48,023	51,810	50,748	45,407	45,821	50,134	50,670	50,801	50,693	53,972
Product 10	52,337	47,203	49,464	53,365	52,271	46,770	47,196	51,638	52,190	52,326	52,214	55,591

# CORES E PREENCHIMENTO

---

Product	Jan	Feb	Mar	Apr	May	Jun	Total
Product 01	93,993	84,773	88,833	95,838	93,874	83,994	541,305
Product 02	87,413	78,839	82,615	89,129	87,303	78,114	503,414
Product 03	90,036	81,204	85,093	91,803	89,922	80,458	518,516
Product 04	92,737	83,640	87,646	94,557	92,620	82,872	534,072
Product 05	83,733	75,520	79,137	85,377	83,627	74,826	482,220
Total	447,913	403,976	423,323	456,705	447,346	400,264	2,579,526

Product	Jan	Feb	Mar	Apr	May	Jun	Total
Product 01	93,993	84,773	88,833	95,838	93,874	83,994	541,305
Product 02	87,413	78,839	82,615	89,129	87,303	78,114	503,414
Product 03	90,036	81,204	85,093	91,803	89,922	80,458	518,516
Product 04	92,737	83,640	87,646	94,557	92,620	82,872	534,072
Product 05	83,733	75,520	79,137	85,377	83,627	74,826	482,220
Total	447,913	403,976	423,323	456,705	447,346	400,264	2,579,526

# GRUPOS E QUEBRAS

---

- Usados quando:
  - Os dados não cabem em simples tabelas
  - Há necessidade de direcionar o usuário a examinar grupos de informações separados do resto
  - Pedaços menores de informação são mais fáceis de serem analisados
  - Inclusão de valores que summarizam os dados

# GRUPOS E QUEBRAS

---

Country: USA		Region: North											
Product		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Product 01		93,993	84,773	88,833	95,838	93,874	83,994	84,759	92,738	93,728	93,972	93,772	99,837
Product 02		87,413	78,839	82,615	89,129	87,303	78,114	78,826	86,246	87,167	87,394	87,208	92,848
Product 03		90,036	81,204	85,093	91,803	89,922	80,458	81,191	88,834	89,782	90,016	89,824	95,634
Product 04		92,737	83,640	87,646	94,557	92,620	82,872	83,626	91,499	92,476	92,716	92,519	98,503
Product 05		86,245	77,785	81,511	87,938	86,136	77,071	77,773	85,094	86,002	86,226	86,043	91,608
Product 06		88,833	80,119	83,956	90,576	88,720	79,383	80,106	87,647	88,582	88,813	88,624	94,356
Product 07		82,614	74,511	78,079	84,236	82,510	73,826	74,498	81,511	82,382	82,596	82,420	87,751
Product 08		85,093	76,746	80,421	86,763	84,985	76,041	76,733	83,957	84,853	85,074	84,893	90,384
Product 09		87,646	79,048	82,834	89,366	87,535	78,322	79,035	86,475	87,399	87,626	87,440	93,095
Product 10		90,275	81,420	85,319	92,047	90,161	80,672	81,406	89,070	90,021	90,255	90,063	95,888
Total		\$884,886	\$798,085	\$836,307	\$902,255	\$883,765	\$790,751	\$797,953	\$873,070	\$882,391	\$884,688	\$882,805	\$939,903
Country: USA		Region: East											
Product		Jan	Feb	Mar	Apr	May	Jun	Jul	Aug	Sep	Oct	Nov	Dec
Product 01		93,993	84,773	88,833	95,838	93,874	83,994	84,759	92,738	93,728	93,972	93,772	99,837
Product 02		87,413	78,839	82,615	89,129	87,303	78,114	78,826	86,246	87,167	87,394	87,208	92,848
Product 03		90,036	81,204	85,093	91,803	89,922	80,458	81,191	88,834	89,782	90,016	89,824	95,634

# GRUPOS E QUEBRAS

---

- Use espaço vertical para separação entre os grupos
- Repita os cabeçalhos das tabelas em cada novo grupo
- Mantenha a estrutura da tabela (mesmas colunas, na mesma ordem...) em todos os grupos
- Quando desejar que cada grupo seja avaliado separadamente, separe-os em páginas

# SEQUENCIAMENTO DAS COLUNAS

---

- Se há um relacionamento hierárquico entre conjuntos de subdivisões categóricas, as mesmas devem ser sequenciadas da esquerda para a direita
- Valores quantitativos derivados de outros valores quantitativos presentes na tabela, devem ser colocados à direita das colunas das quais derivam

Product	Units Sold	Actual Revenue	% of Total	Fcst Revenue	% of Fcst
Product A	938	187,600	47%	175,000	107%
Product B	1,093	114,765	28%	130,000	88%
Product C	3,882	62,112	15%	50,000	124%
Product D	873	36,666	9%	40,000	92%
Product E	72	2,088	1%	50,000	4%
Total	6,858	\$403,231	100%	\$445,000	91%

# SEQUENCIAMENTO DAS COLUNAS

---

- Colunas contendo valores que serão comparados pelos usuários devem ser colocadas lado a lado

Product	Units Sold	Last	This	Fcst	Planned
		Year's Revenue	Year's Revenue		
Product A	938	159,497	187,600	175,000	160,000
Product B	1,093	123,007	114,765	130,000	125,000
Product C	3,882	45,384	62,112	53,000	50,000
Product D	873	41,003	36,666	38,000	40,000
Product E	72	2,485	2,088	4,000	5,000
Total	6,858	\$371,376	\$403,231	\$400,000	\$380,000

# SEQUENCIAMENTO DOS DADOS

---

- Números e datas têm ordenações naturais bastante significativas
- Quando se trata de nomes, a ordenação alfabética nos auxilia a encontrar valores desejados em grandes tabelas
- Porém, em alguns casos pode-se criar outra ordenação que faça mais sentido

# ORIENTAÇÃO DO TEXTO

---

orientação

Orientação

Orientação

orientação

Orientação

orientação

# ORIENTAÇÃO DO TEXTO

---

- Não há representação melhor que a da linguagem natural da esquerda para a direita
- Há raríssimos casos nos quais esta orientação deva ser sacrificada e normalmente estes casos acontecem em gráficos e não em tabelas

# ALINHAMENTO

---

- Números que representem valores quantitativos devem ser alinhados à direita

Sales	Sales	Sales
93,883.39	93,883.39	93,883.39
5,693,762.32	5,693,762.32	5,693,762.32
483.84	483.84	483.84
674,663.39	674,663.39	674,663.39
548.93	548.93	548.93
3,847.33	3,847.33	3,847.33
\$6,467,189.20	\$6,467,189.20	\$6,467,189.20

# ALINHAMENTO

---

- Datas devem ser alinhadas à esquerda usando um formato que mantenha constantes o número de dígitos

Date	Date	Date	Date
12/17/02	12/17/02	12/17/02	12/17/02
1/2/03	01/02/03	1/2/03	01/02/03
1/17/03	01/17/03	1/17/03	01/17/03
2/9/03	02/09/03	2/9/03	02/09/03
10/29/03	10/29/03	10/29/03	10/29/03
12/1/03	12/01/03	12/1/03	12/01/03
1/1/03	01/01/03	1/1/03	01/01/03

# ALINHAMENTO

---

- Texto que não expressa nem números nem datas, devem ser alinhados à esquerda

Product Code	Product Name	Product Code	Product Name	Product Code	Product Name
A1838	2-Door Sport	A1838	2-Door Sport	A1838	2-Door Sport
A89	4-Door Sport	A89	4-Door Sport	A89	4-Door Sport
J98488	2-Door Luxury	J98488	2-Door Luxury	J98488	2-Door Luxury
J3883	4-Door Luxury	J3883	4-Door Luxury	J3883	4-Door Luxury
K9288	2-Door Truck	K9288	2-Door Truck	K9288	2-Door Truck
K38733	4-Door Truck	K38733	4-Door Truck	K38733	4-Door Truck

# ALINHAMENTO

---

- Quando o texto tem tamanho fixo e é pequeno, uma exceção para balancear melhor o espaço em branco é alinhar ao centro

Cust Code	Preferred?	Cust Code	Preferred?
193847394	Y	193847394	Y
109388484	N	109388484	N
187466463	N	187466463	N
174563553	N	174563553	N
175357736	Y	175357736	Y
167374565	Y	167374565	Y

# ALINHAMENTO

---

- Alinhar o cabeçalho juntamente com o conteúdo
- Spinner headers devem ser alinhados ao centro mostrando que se referem ao conjunto de colunas

Product	Regions				Total
	North	East	South	West	
Product A	94	152	174	87	507
Product B	122	198	226	113	659
Product C	101	164	188	94	547
Product D	142	230	263	131	766
Product E	132	214	244	122	712
Product F	304	491	562	281	1,638
Total	895	1,448	1,657	829	4,829

# NUMERO E FORMATO

---

- Não incluir informação desnecessária
- Expressar a data de forma familiar aos usuários
- Alinhar os dados de forma consistente

# NUMERO E FORMATO

---

- Colocar . (ou ,) a cada três dígitos (Ex.: 1.393.033 ao invés de 1393033): quebra o número em blocos (*chunks*) menores
- Truncar a exibição de três dígitos para milhar, milhão, bilhão, etc., sempre que possível e indicar isto através de título ou cabeçalho: reduz a quantidade de informação a ser processada
- Usar o sinal de negativos ou parênteses para indicação de números negativos
- Colocar o sinal de % imediatamente a esquerda de todo valor percentual

# DATAS E FORMATO

---

- Expresse meses no formato de dois dígitos (01, 02, ...) ou três caracteres (jan, fev, ...)
- Expresse dias usando o formato de dois dígitos
- Exclua porções da data que não sejam necessárias no contexto

# DATAS E FORMATO

---

- Atenção: formatos de datas variam entre países
  - Nos Estados unidos, usa-se o mês antes do dia
  - Na Europa usa-se o dia antes do mês
- Explicitar a organização utilizada

Order Date (mm/dd/yy)
12/17/02
01/02/03
01/17/03

# PRECISAO DE NUMEROS

---

- A precisão de um número ou data é o seu grau de detalhe
- O nível de precisão não deve exceder o necessário para se atingir os objetivos de comunicação e as necessidades dos usuários

Region	Revenue	% of Total	% of Total	% of Total	% of Total
Americas	636,663,663	40%	39.8%	39.82%	39.816%
Europe	443,874,773	28%	27.8%	27.76%	27.759%
Asia	399,393,993	25%	25.0%	24.98%	24.978%
Australia	99,838,333	6%	6.2%	6.24%	6.244%
Middle East	10,399,383	1%	0.7%	0.65%	0.650%
Africa	7,939,949	1%	0.5%	0.50%	0.497%
Total	\$1,598,110,094	100%	100.0%	99.94%	99.944%

# PRECISAO DAS DATAS

---

- A mensagem requer precisão de dia, semana, quinzena, mês, estação, ano?

# FONTES

---

- Devem ser legíveis
- As mesmas fontes devem ser mantidas em toda a tabela
- Fontes são classificadas geralmente em com e sem serifas
- A serifas é uma linha ou curva, localizada em extremos das letras

Boa legibilidade		Difícil legibilidade	
Com serifa	Sem serifa	Com serifa	Sem serifa
Times new roman	Arial	Stencil	<b>Britanic bold</b>
Palatino	Verdana	Baskerville	Papyrus
Courier	Tahoma	<i>Monotype corsiva</i>	PT sans narrow

# ENFASE E COR

---

- Texto pode ser agrupado ou destacado através do uso de negrito e itálico

# SUMARIZANDO VALORES

---

- Tabelas são meios ideais para combinar detalhes e sumários
  - Somas
  - Médias
  - Número de ocorrências

Product	Jan	Feb	Mar	Apr	May	Jun	Total
Product 01	93,993	84,773	88,833	95,838	93,874	83,994	\$541,305
Product 02	87,413	78,839	82,615	89,129	87,303	78,114	\$503,414
Product 03	90,036	81,204	85,093	91,803	89,922	80,458	\$518,516
Product 04	92,737	83,640	87,646	94,557	92,620	82,872	\$534,072
Product 05	86,245	77,785	81,511	87,938	86,136	77,071	\$496,687
Product 06	88,833	80,119	83,956	90,576	88,720	79,383	\$511,587
Product 07	82,614	74,511	78,079	84,236	82,510	73,826	\$475,776
Product 08	85,093	76,746	80,421	86,763	84,985	76,041	\$490,049
Product 09	87,646	79,048	82,834	89,366	87,535	78,322	\$504,751
Product 10	90,275	81,420	85,319	92,047	90,161	80,672	\$519,893
<b>Total</b>	<b>\$884,886</b>	<b>\$798,085</b>	<b>\$836,307</b>	<b>\$902,255</b>	<b>\$883,765</b>	<b>\$790,751</b>	<b>\$5,096,049</b>

### **Region: North**

Product	Jan	Feb	Mar	Apr	May	Jun
Product 01	93,993	84,773	88,833	95,838	93,874	83,994
Product 02	87,413	78,839	82,615	89,129	87,303	78,114
Product 03	90,036	81,204	85,093	91,803	89,922	80,458
Product 04	92,737	83,640	87,646	94,557	92,620	82,872
Product 05	86,245	77,785	81,511	87,938	86,136	77,071
Product 06	88,833	80,119	83,956	90,576	88,720	79,383
Product 07	82,614	74,511	78,079	84,236	82,510	73,826
Product 08	85,093	76,746	80,421	86,763	84,985	76,041
Product 09	87,646	79,048	82,834	89,366	87,535	78,322
Product 10	90,275	81,420	85,319	92,047	90,161	80,672
<b>Total</b>	<b>\$884,886</b>	<b>\$798,085</b>	<b>\$836,307</b>	<b>\$902,255</b>	<b>\$883,765</b>	<b>\$790,751</b>

### **Region: East**

Product	Jan	Feb	Mar	Apr	May	Jun
Product 01	93,993	84,773	88,833	95,838	93,874	83,994
Product 02	87,413	78,839	82,615	89,129	87,303	78,114
Product 03	90,036	81,204	85,093	91,803	89,922	80,458

Product	Jan	Feb	Mar	Q1 Total	Apr	May	Jun	Q2 Total
Product 01	93,993	84,773	88,833	267,599	95,838	93,874	83,994	273,706
Product 02	87,413	78,839	82,615	248,867	89,129	87,303	78,114	254,547
Product 03	90,036	81,204	85,093	256,333	91,803	89,922	80,458	262,183
Product 04	92,737	83,640	87,646	264,023	94,557	92,620	82,872	270,048
Product 05	86,245	77,785	81,511	245,541	87,938	86,136	77,071	251,145
Product 06	88,833	80,119	83,956	252,908	90,576	88,720	79,383	258,679
Product 07	82,614	74,511	78,079	235,204	84,236	82,510	73,826	240,572
Product 08	85,093	76,746	80,421	242,260	86,763	84,985	76,041	247,789
Product 09	87,646	79,048	82,834	249,528	89,366	87,535	78,322	255,223
Product 10	90,275	81,420	85,319	257,014	92,047	90,161	80,672	262,879
Total	\$884,886	\$798,085	\$836,307	\$2,519,278	\$902,255	\$883,765	\$790,751	\$2,576,771

As colunas ou linhas de sumarização não devem ser confundidas com as de dados

Portanto, devem ser destacadas com espaços adicionais ou cor de fundo clara

# CABEÇALHOS OU RODAPES

---

- Quando a summarização é a coisa mais importante da mensagem, colocá-la no cabeçalho

<b>Region: North</b>	<b>\$1,568,586</b>	<b>\$1,414,719</b>	<b>\$1,482,474</b>	<b>\$1,599,376</b>	<b>\$1,566,600</b>	<b>\$1,401,719</b>
<b>Salesperson</b>	<b>Jan</b>	<b>Feb</b>	<b>Mar</b>	<b>Apr</b>	<b>May</b>	<b>Jun</b>
Abrams, S	93,993	84,773	88,833	95,838	93,874	83,994
Benson, J	87,413	78,839	82,615	89,129	87,303	78,114
James, R	86,245	77,785	81,511	87,938	86,136	77,071
Wilson, O	86,704	78,199	81,944	88,406	86,594	77,481
Yao, J	89,305	80,545	84,403	91,058	89,192	79,805

<b>Region: North</b>	<b>Total</b>	<b>\$1,568,586</b>	<b>\$1,414,719</b>	<b>\$1,482,474</b>	<b>\$1,599,376</b>	<b>\$1,566,600</b>	<b>\$1,401,719</b>
<b>Salesperson</b>	<b>Jan</b>	<b>Feb</b>	<b>Mar</b>	<b>Apr</b>	<b>May</b>	<b>Jun</b>	
Abrams, S	93,993	84,773	88,833	95,838	93,874	83,994	
Benson, J	87,413	78,839	82,615	89,129	87,303	78,114	
James, R	86,245	77,785	81,511	87,938	86,136	77,071	
Wilson, O	86,704	78,199	81,944	88,406	86,594	77,481	
Yao, J	89,305	80,545	84,403	91,058	89,192	79,805	

# TABELAS EM MULTIPLAS PAGINAS

---

- Quando tabelas ocupam várias páginas, repetir o cabeçalho

State	Order Date	Sales Volume	Sales Revenue
	05/04/03	354	176,646
	05/05/03	456	227,544
	05/06/03	556	277,444
	05/07/03	598	298,402
	05/08/03	553	275,947
	05/09/03	465	232,035
	05/10/03	434	216,566
	05/11/03	676	337,324
	05/12/03	589	293,911
	05/13/03	688	343,312
	05/14/03	701	349,799
	05/15/03	501	249,999
	05/16/03	556	277,444
	05/17/03	623	310,877
	05/18/03	456	227,544
	05/19/03	563	280,937
	05/20/03	367	183,133
	05/21/03	356	177,644

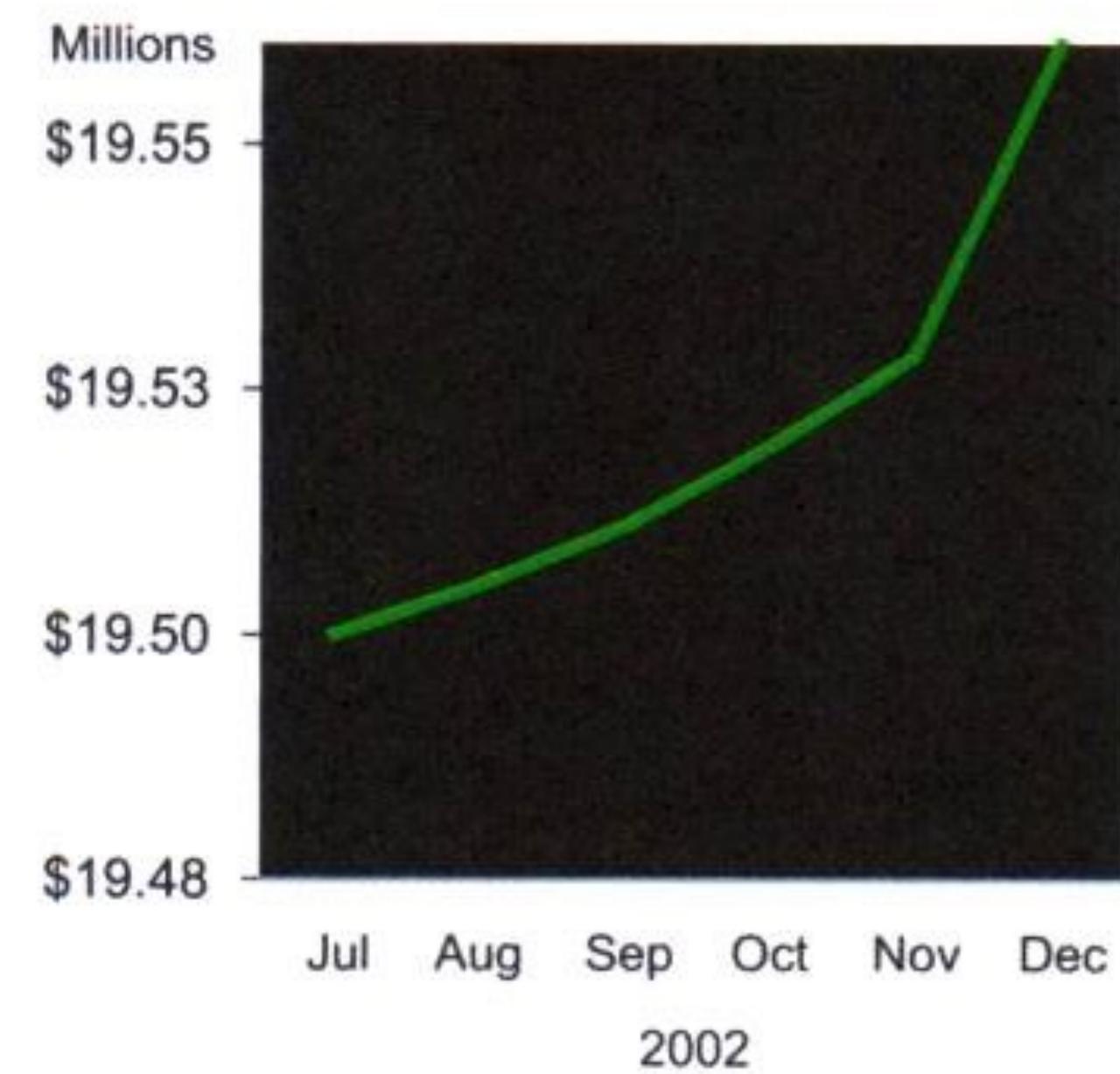
State	Order Date	Sales Volume	Sales Revenue
Alabama	05/01/03	432	215,568
Alabama	05/02/03	534	266,466
Alabama	05/03/03	466	232,534
Alabama	05/04/03	354	176,646
Alabama	05/05/03	456	227,544
Alabama	05/08/03	553	275,947
Alabama	05/09/03	465	232,035
Alabama	05/12/03	589	293,911
Alabama	05/15/03	501	249,999
Alabama	05/16/03	556	277,444
Alabama	05/17/03	623	310,877
Alabama	05/19/03	563	280,937
Alabama	05/22/03	675	336,825
Alabama	05/23/03	702	350,298
Alabama	05/24/03	658	328,342
Alabama	05/26/03	798	398,202
Alabama	05/29/03	801	399,699
Alabama	05/30/03	735	366,765
Alabama	05/31/03	802	400,198
Arkansas	05/01/03	201	100,299
Arkansas	05/02/03	247	123,253
Arkansas	05/03/03	245	122,255
Arkansas	05/04/03	277	138,223
Arkansas	05/05/03	203	101,297

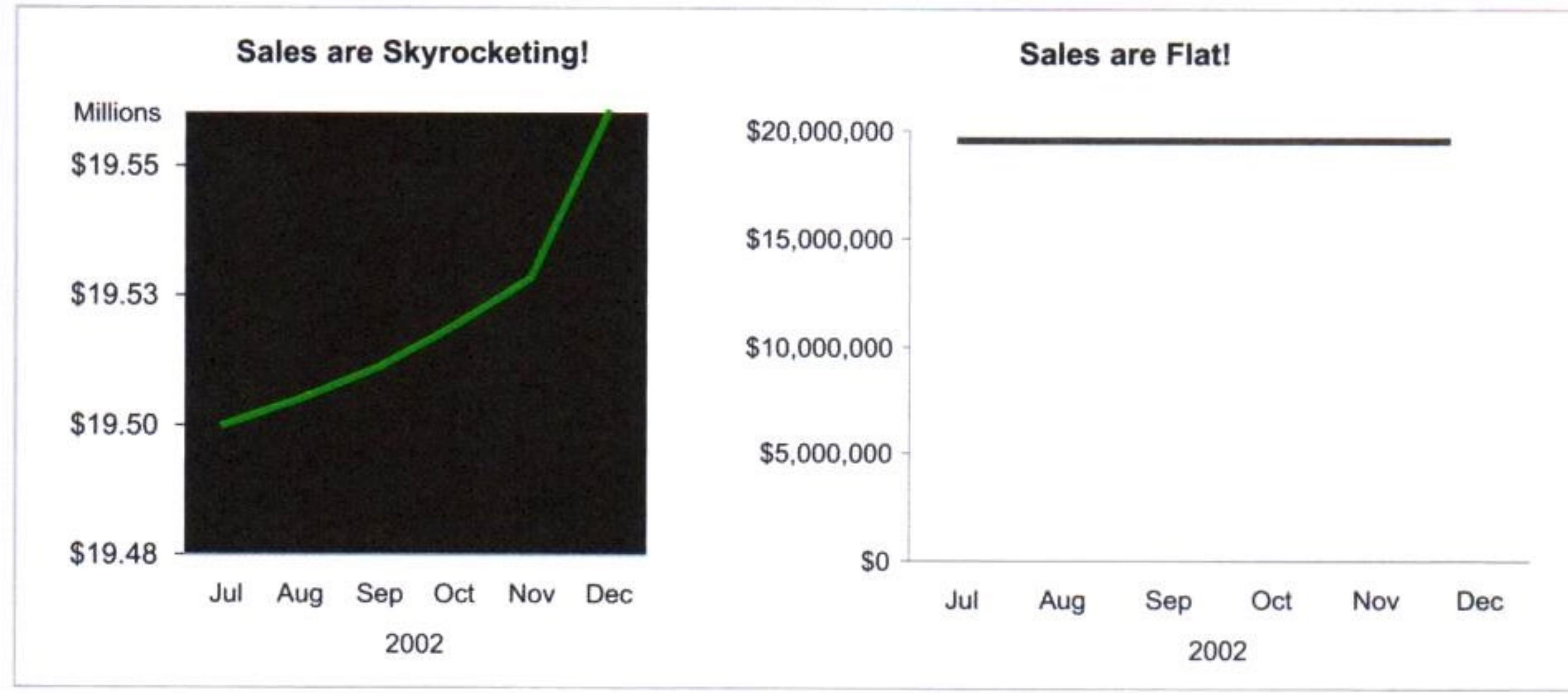
# Gráficos

---

**Profa. Dra. Raquel Minardi**  
Departamento de Ciência da Computação  
Universidade Federal de Minas Gerais

### Sales are Skyrocketing!



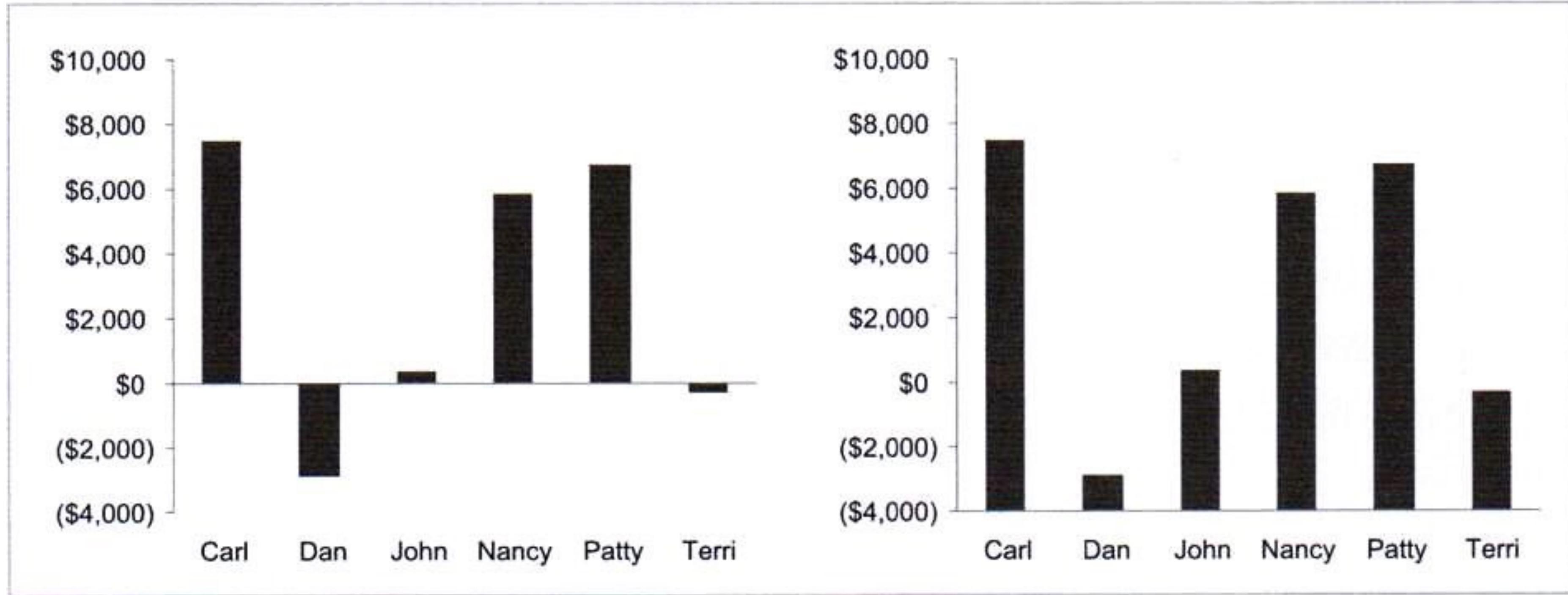


Um aumento de 1% parece um aumento de 200%

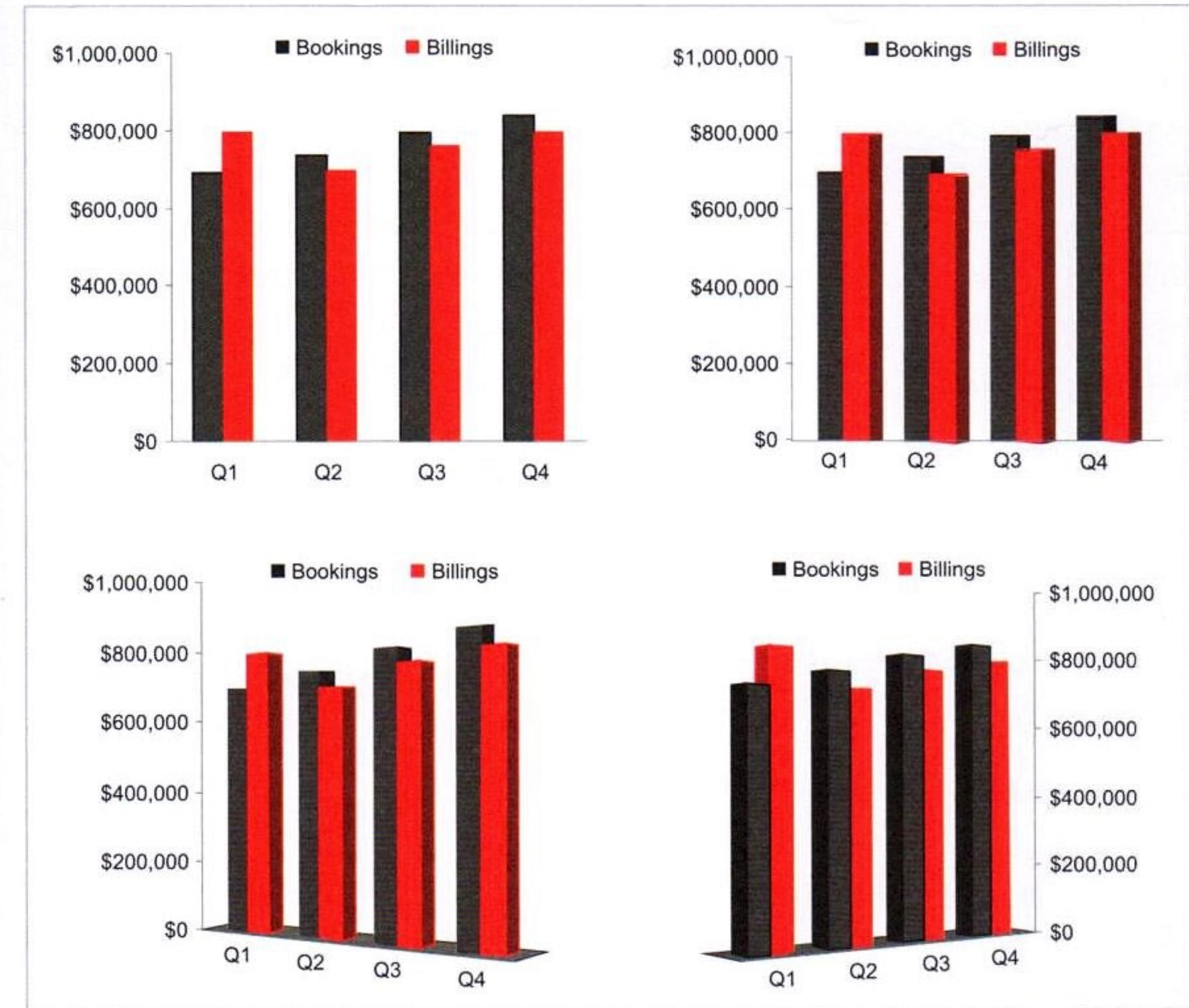
Quais aspectos do gráfico foram usados para causar este efeito?

- A escala no eixo y não inicia do 0 fazendo com que pequenas elevações pareçam significativas
- A área do gráfico tem altura maior ou igual a largura, fazendo com que pequenas elevações pareçam maiores
- A linha é verde, cor que reforça a positividade da mensagem
- A legenda com o termo “**Millions**”

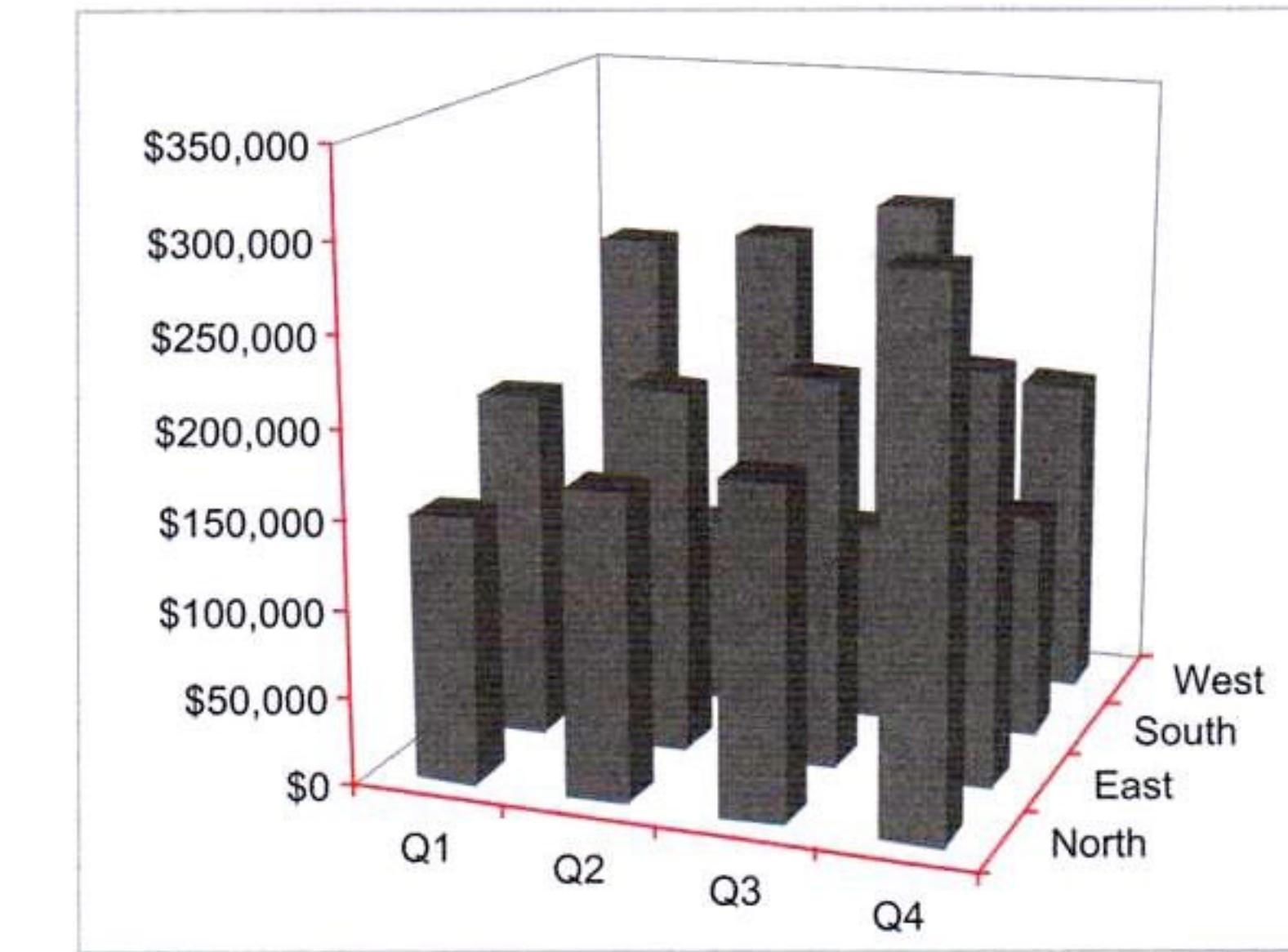
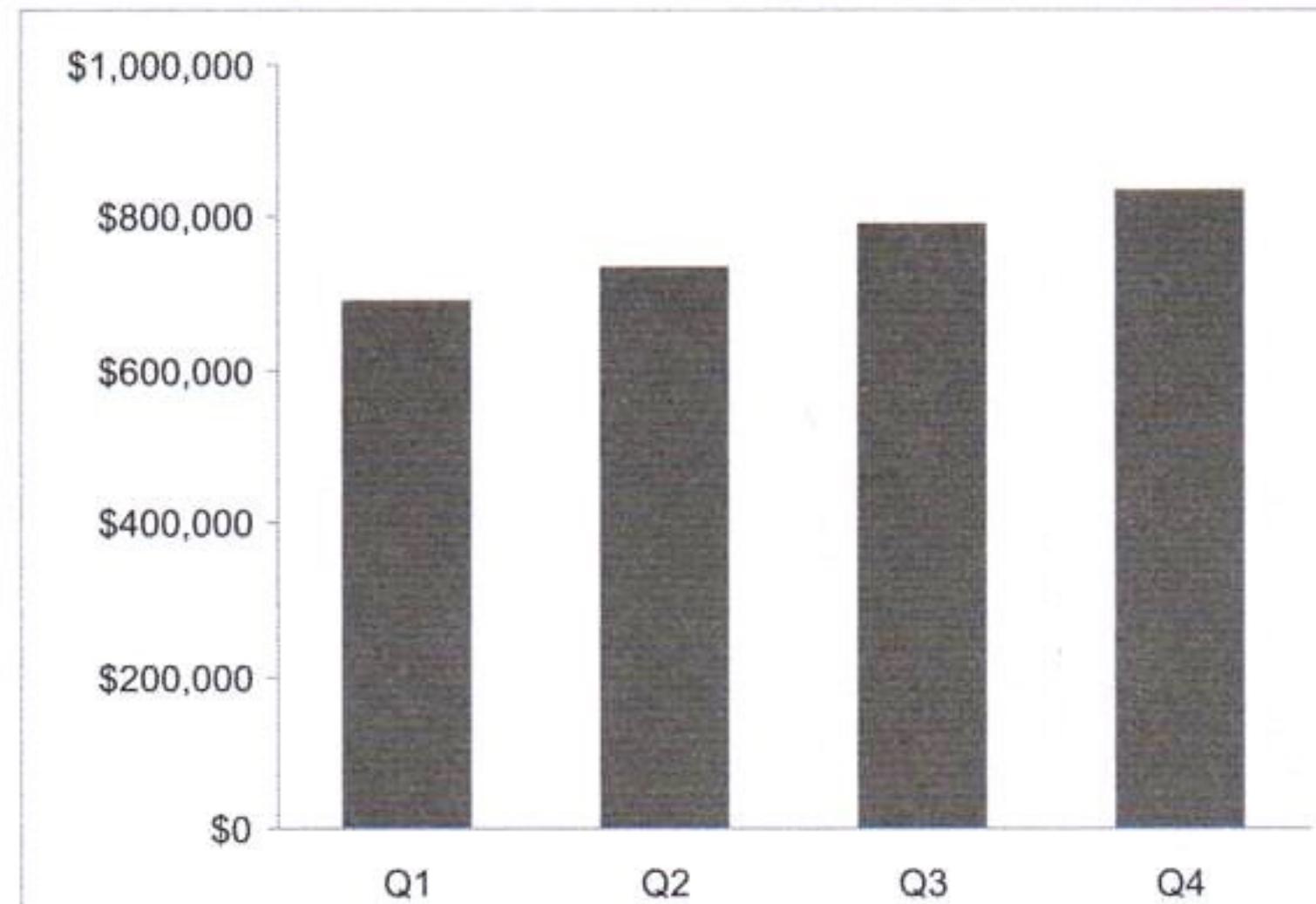
- Faça com que a diferença de tamanho entre as marcações nos eixos correspondam às diferenças nos valores que elas representam
- Inclua o valor zero nas escalas quantitativas ou alerte o usuário quando não o fizer
- Atenção especial para os gráficos de barras: tamanho das barras codificam a informação quantitativa e deve ser mantido



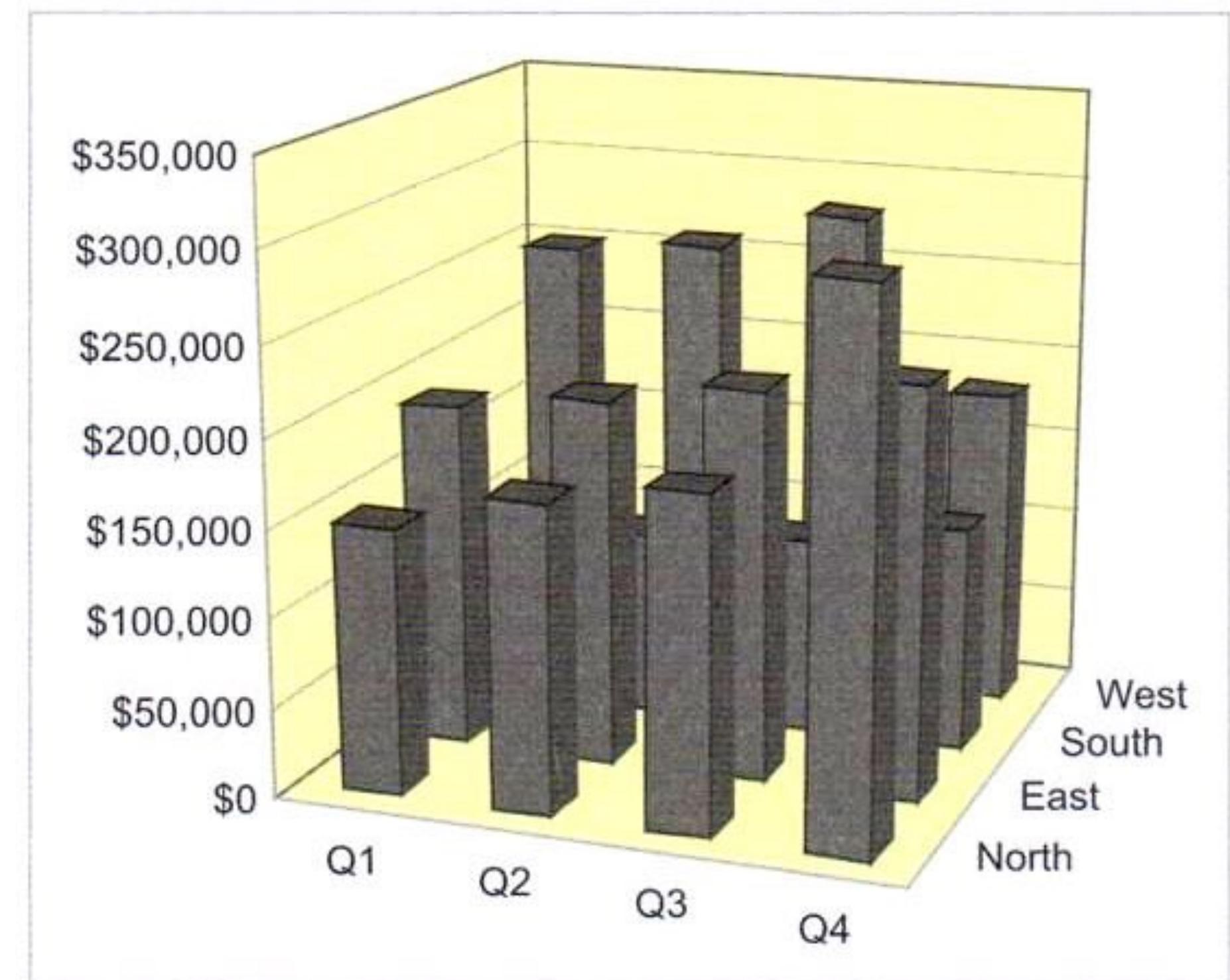
Representação confusa

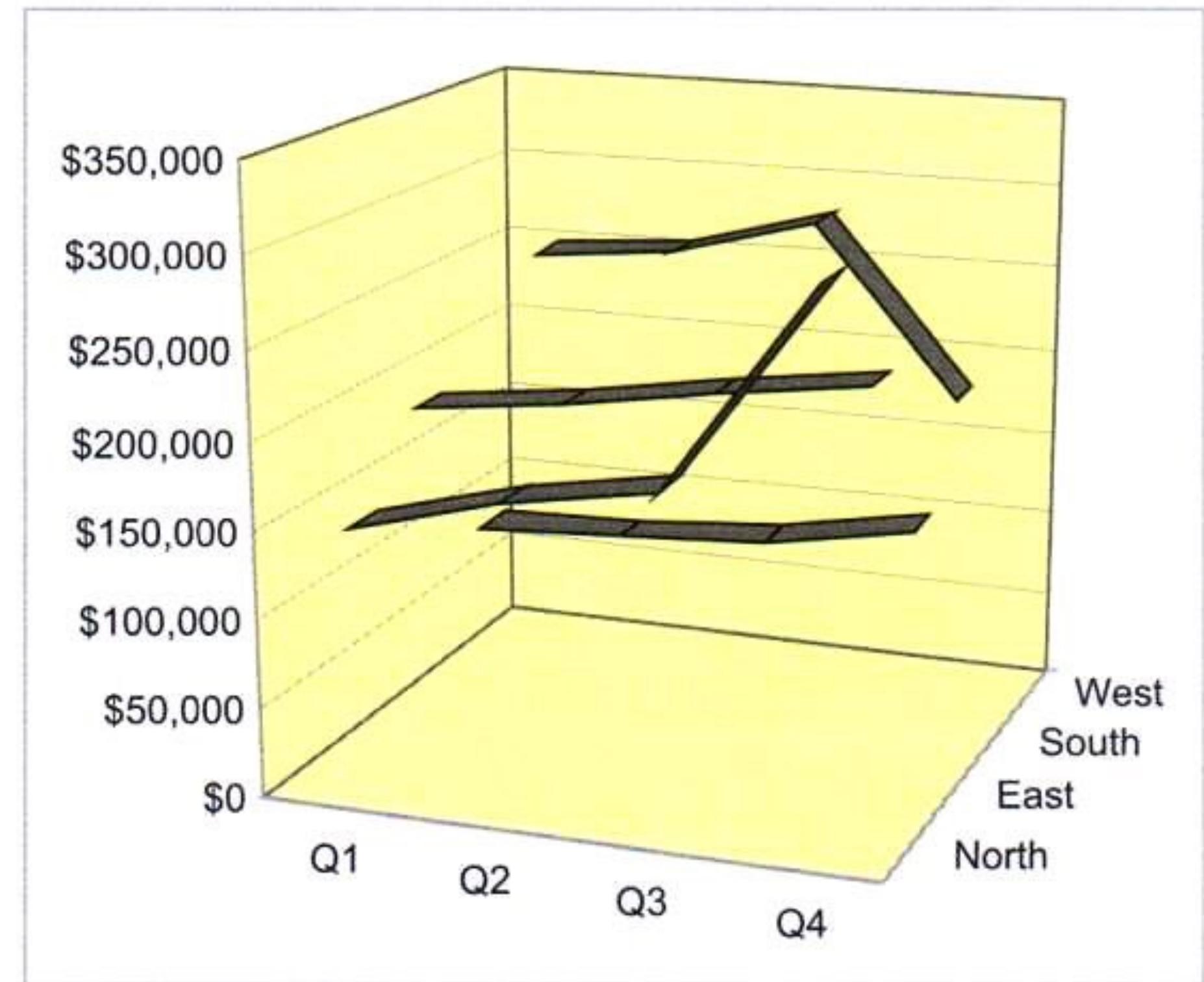


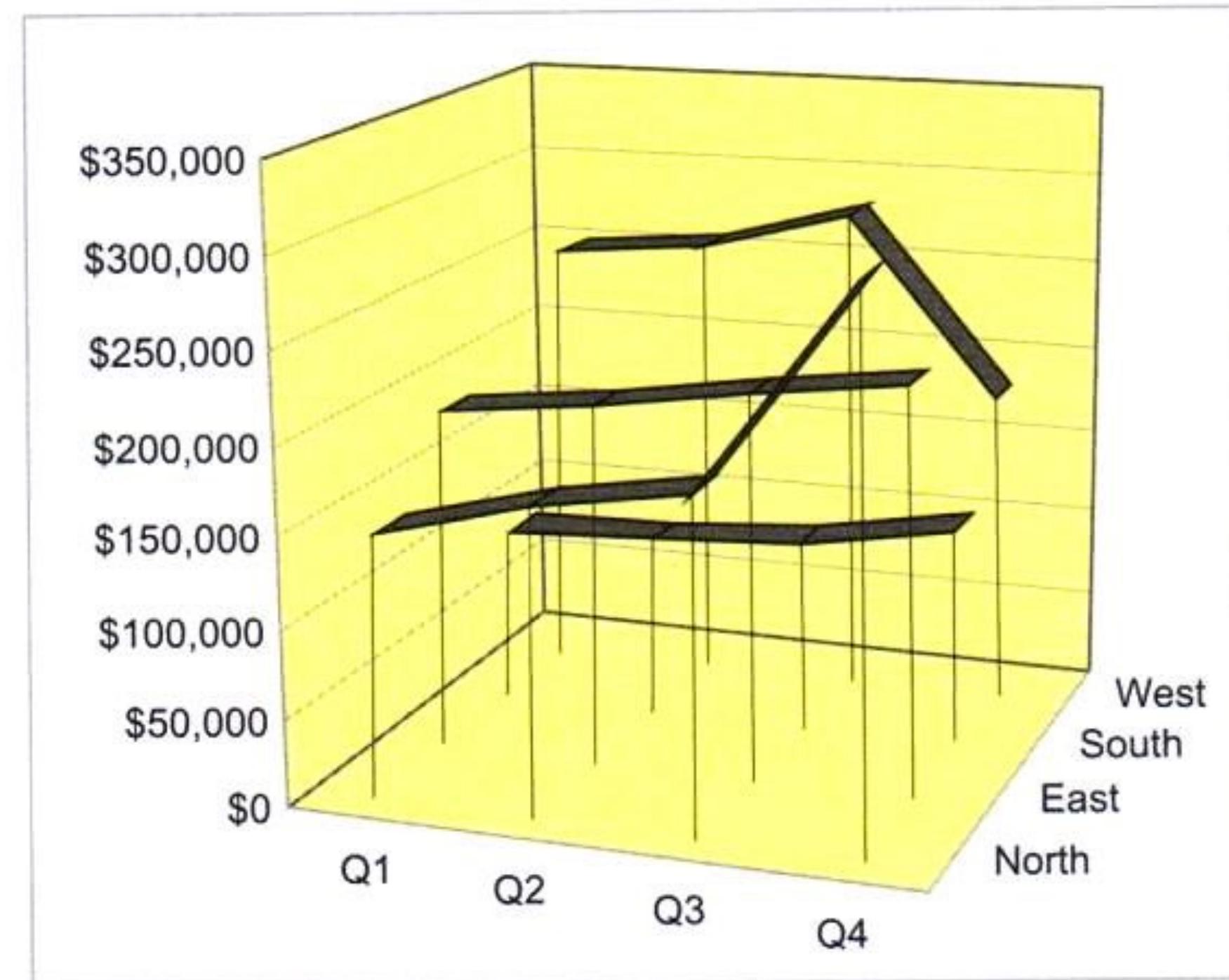
- A adição de uma terceira dimensão adiciona não adiciona informação qualquer: princípio *data-ink ratio* de Tufte
- Adiciona-se conteúdo visual que os usuários precisam processar resultando em perda de tempo e esforço



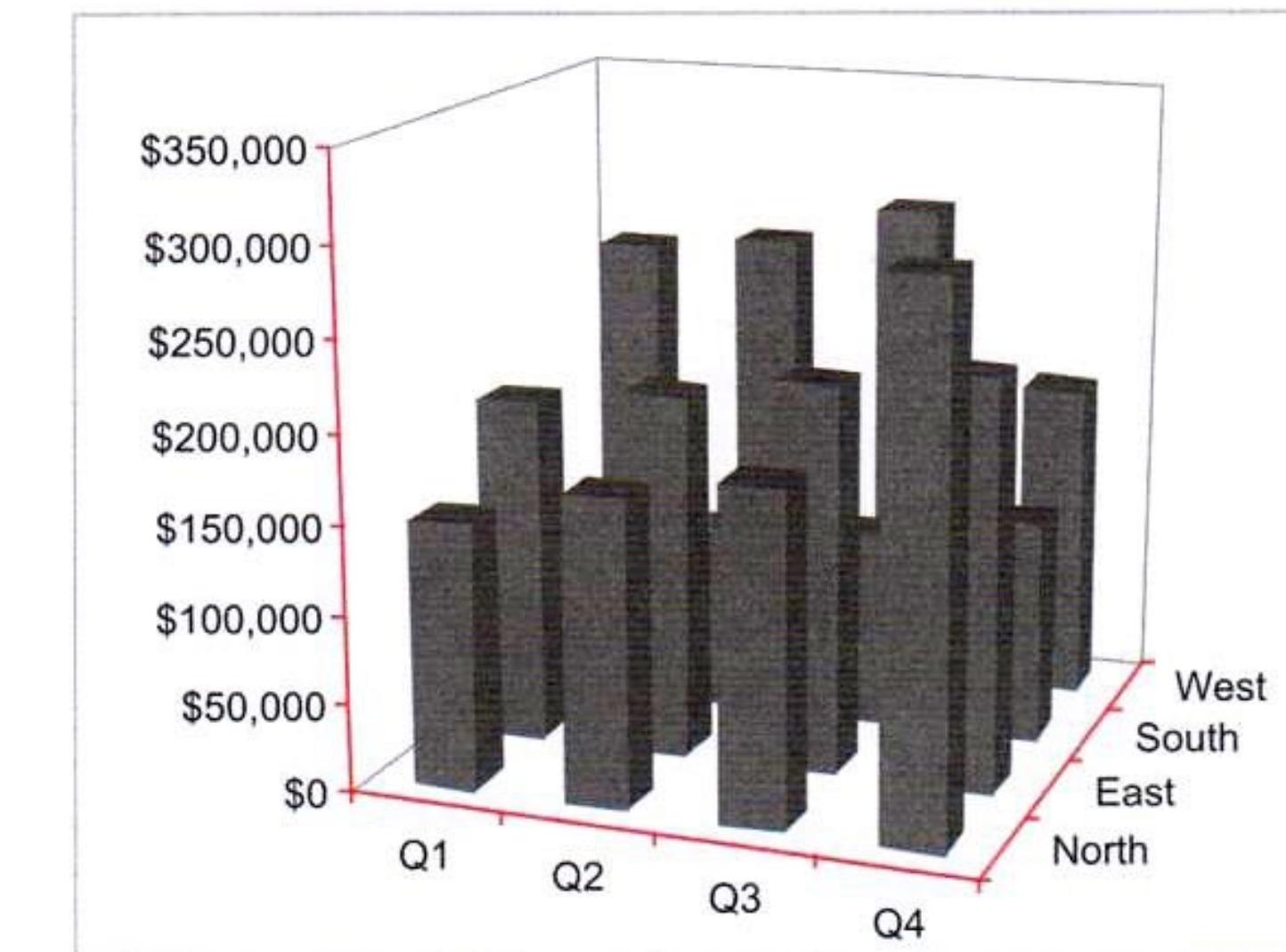
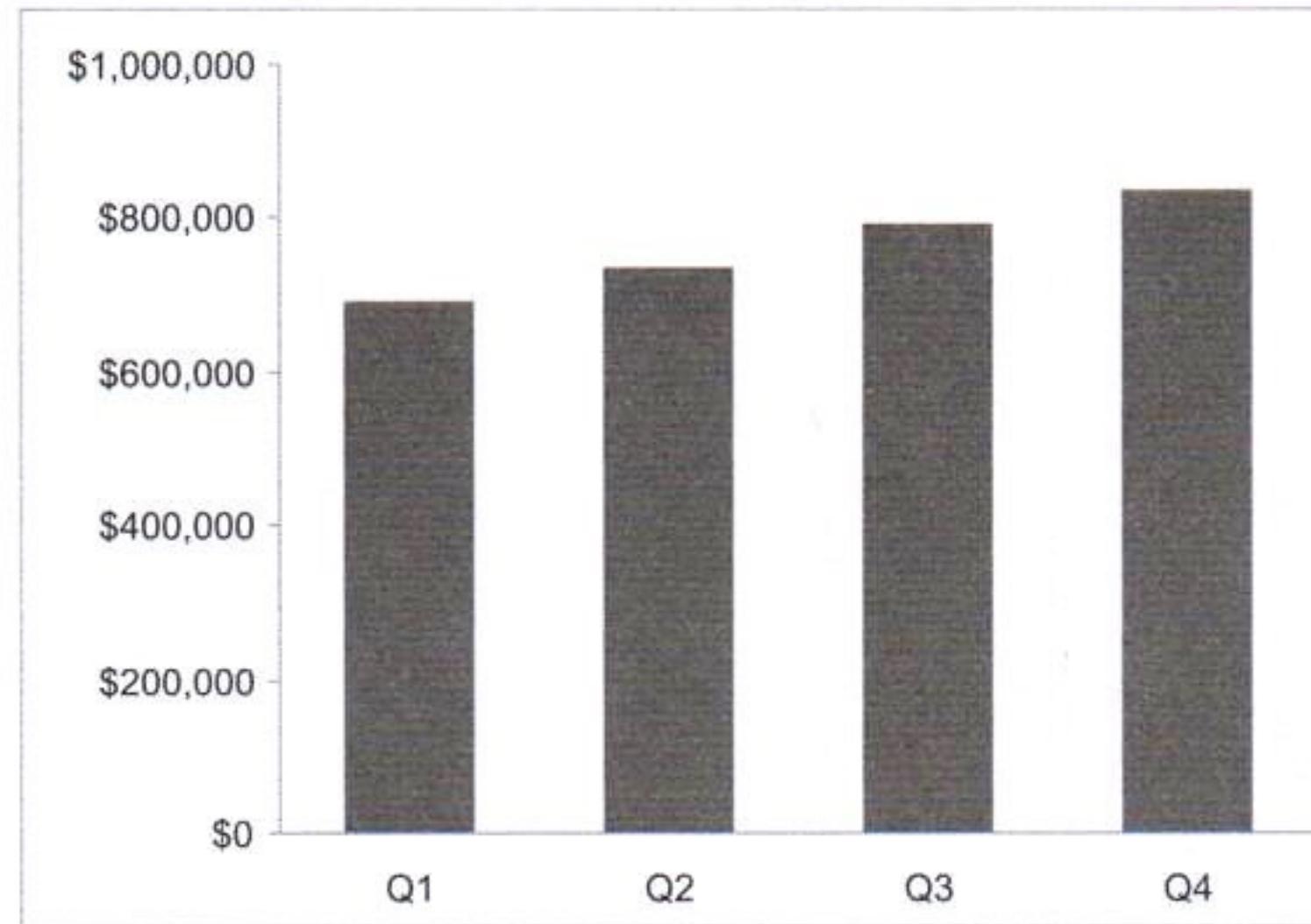
Em teoria, um gráfico tridimensional pode exibir dados com três variáveis



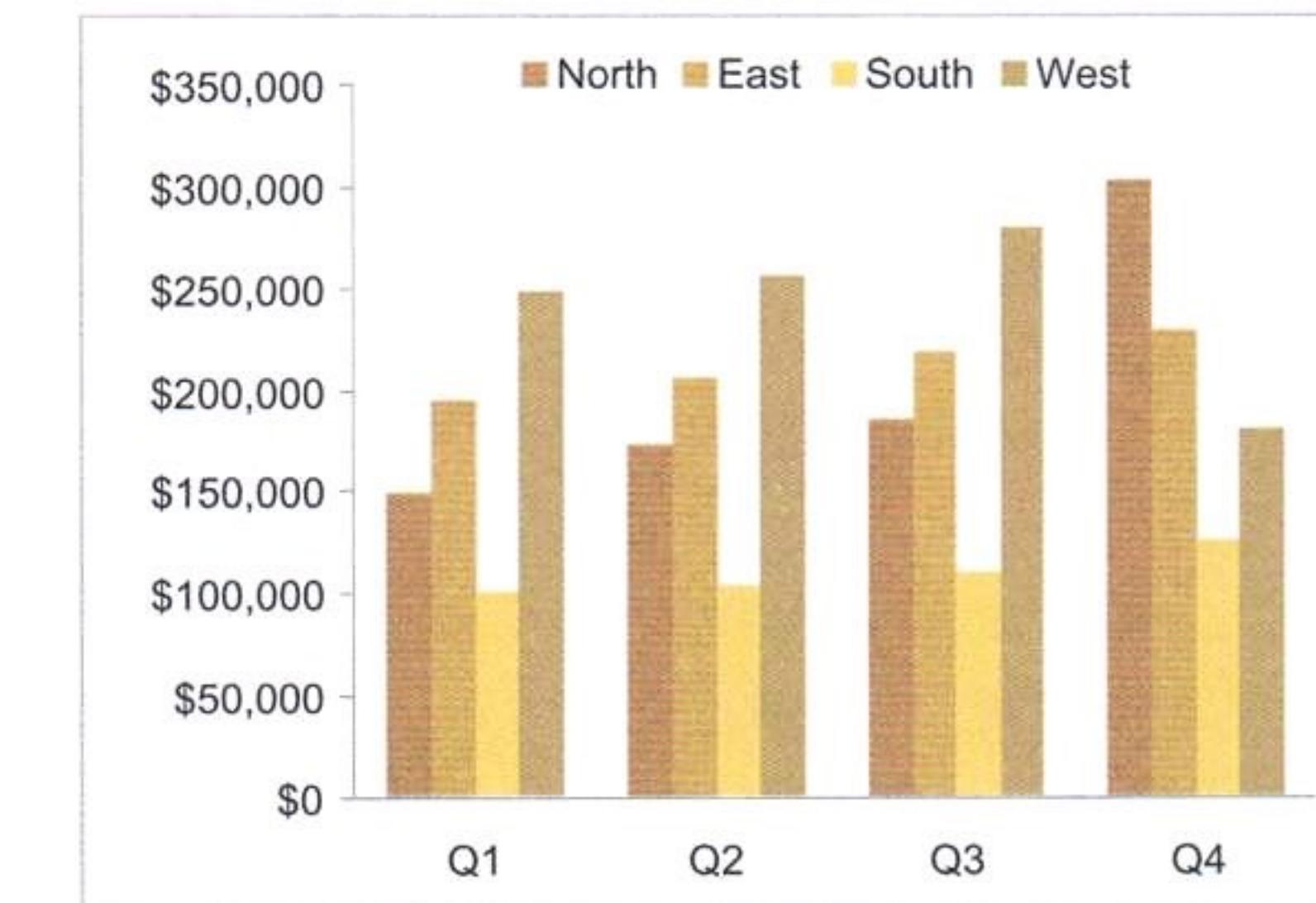




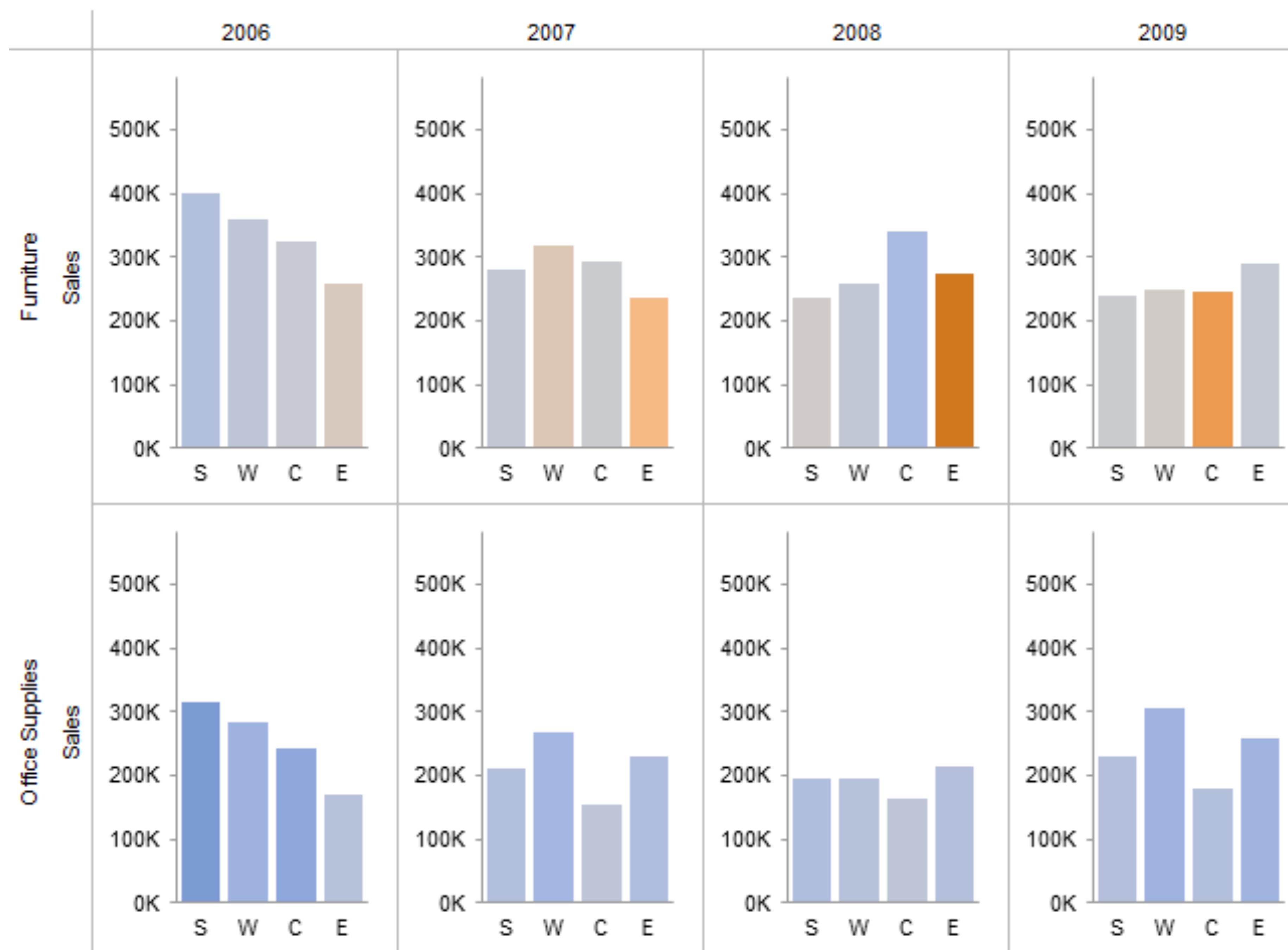
*Uso de drop lines*

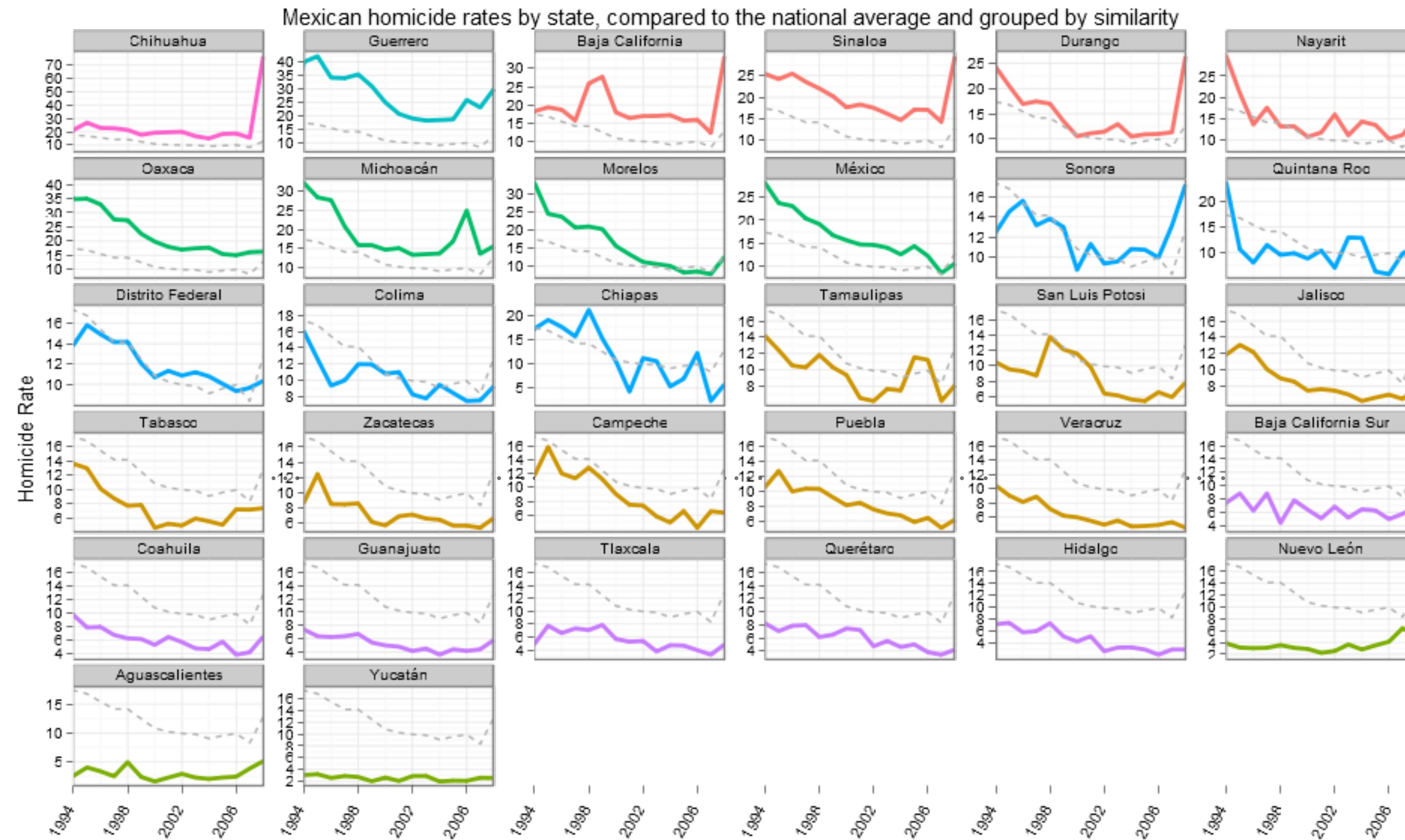


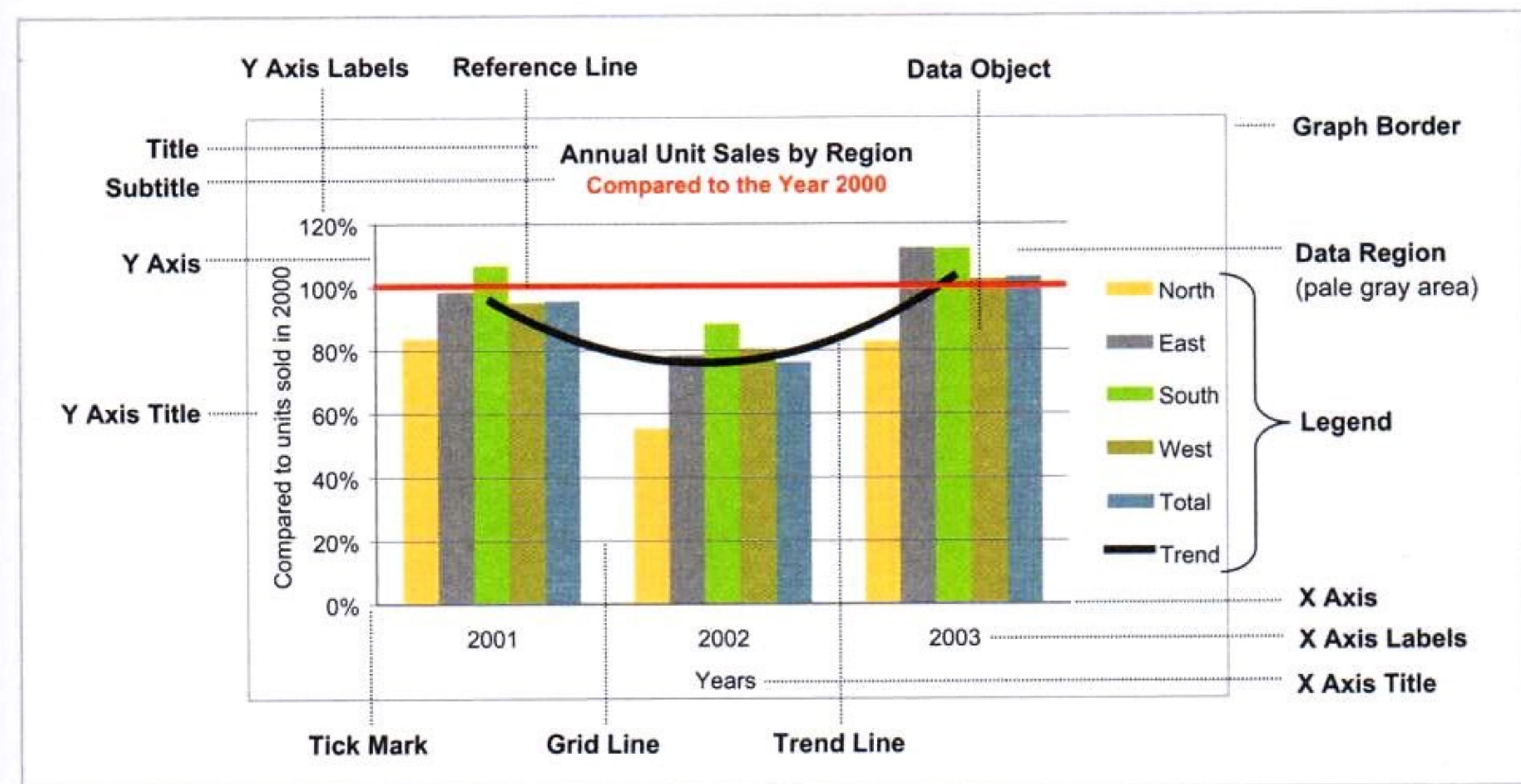
A melhor opção acaba não usando a terceira dimensão



E se houverem mais dimensões?  
Como resolver este problema?



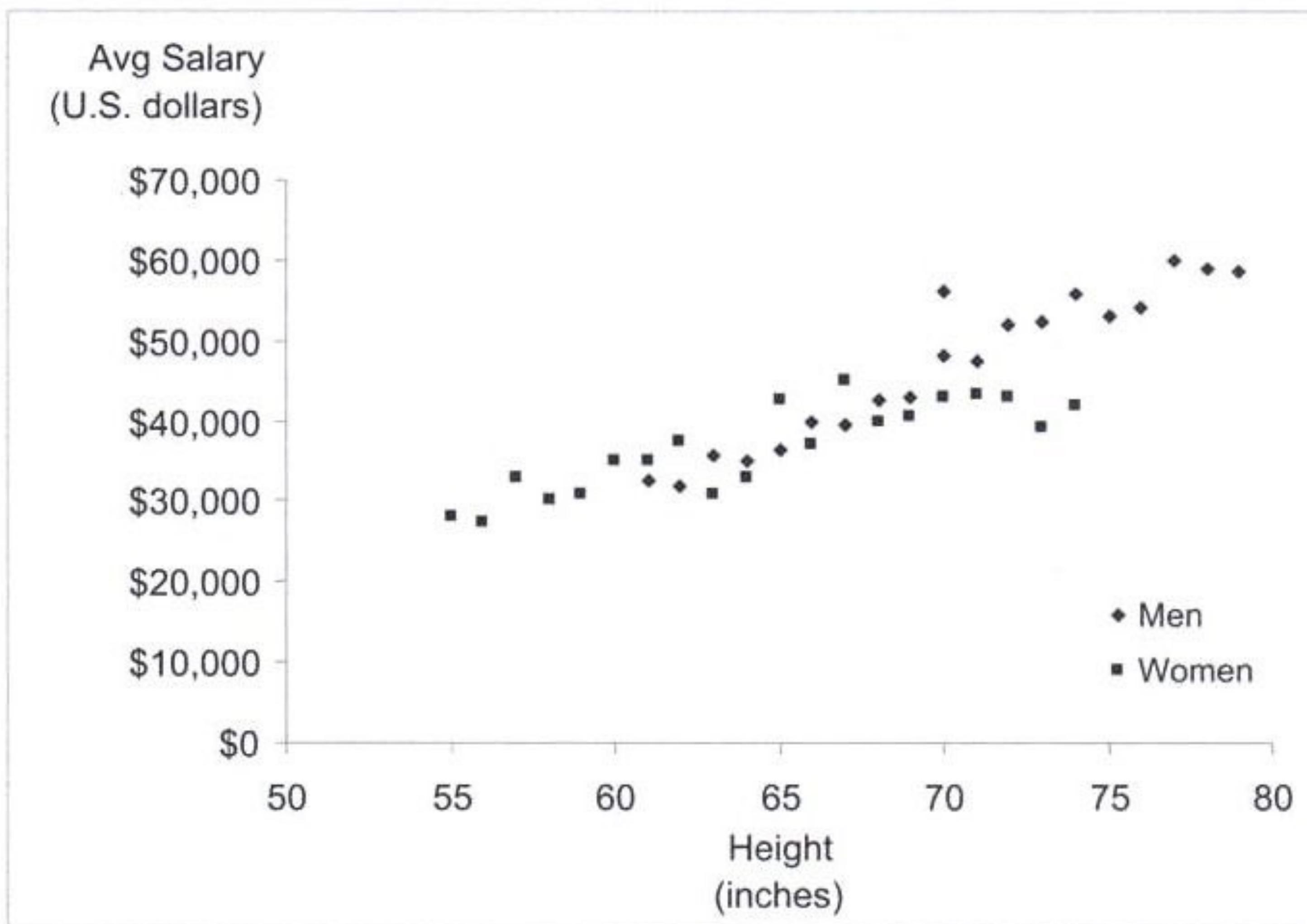


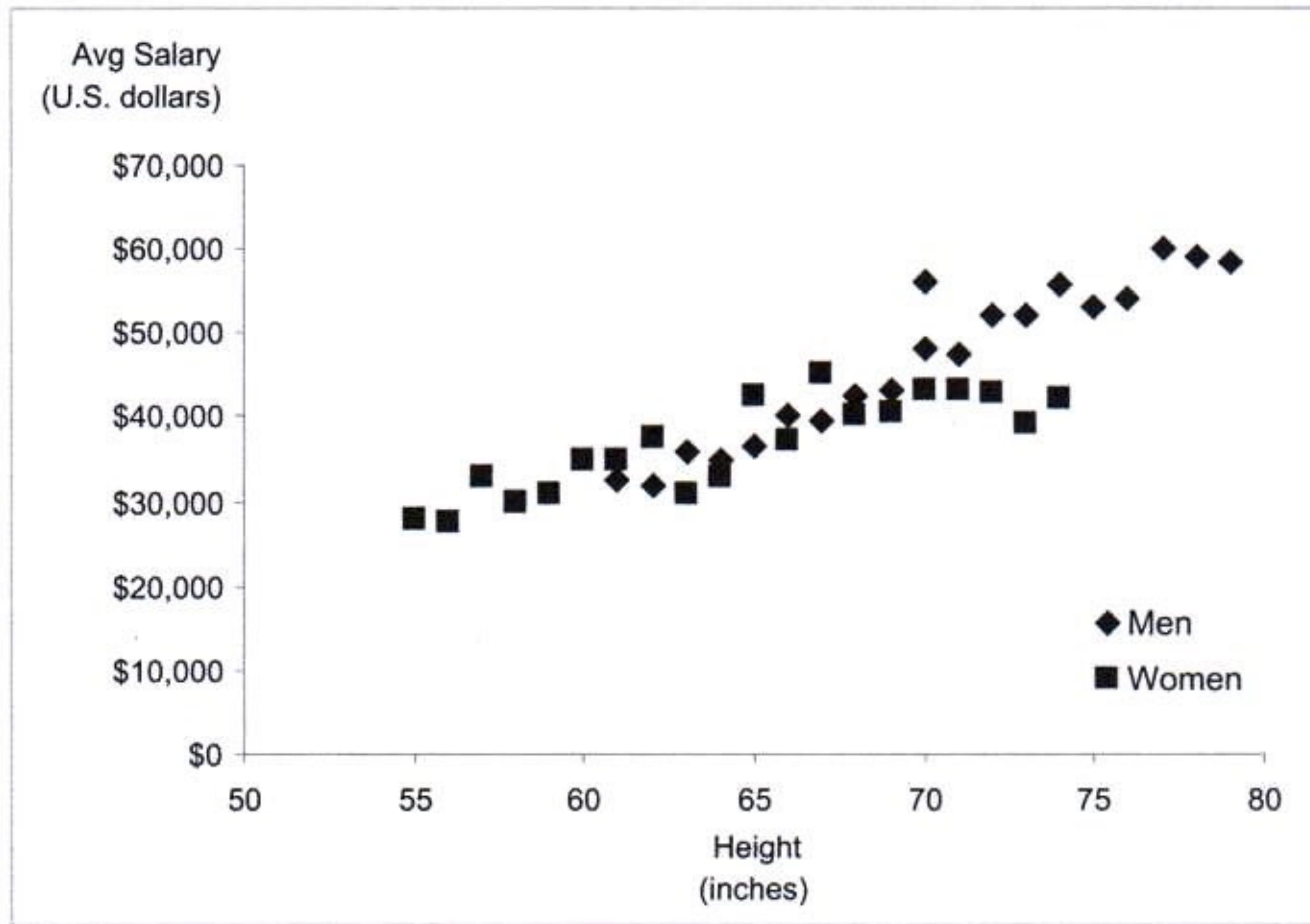


# PONTOS

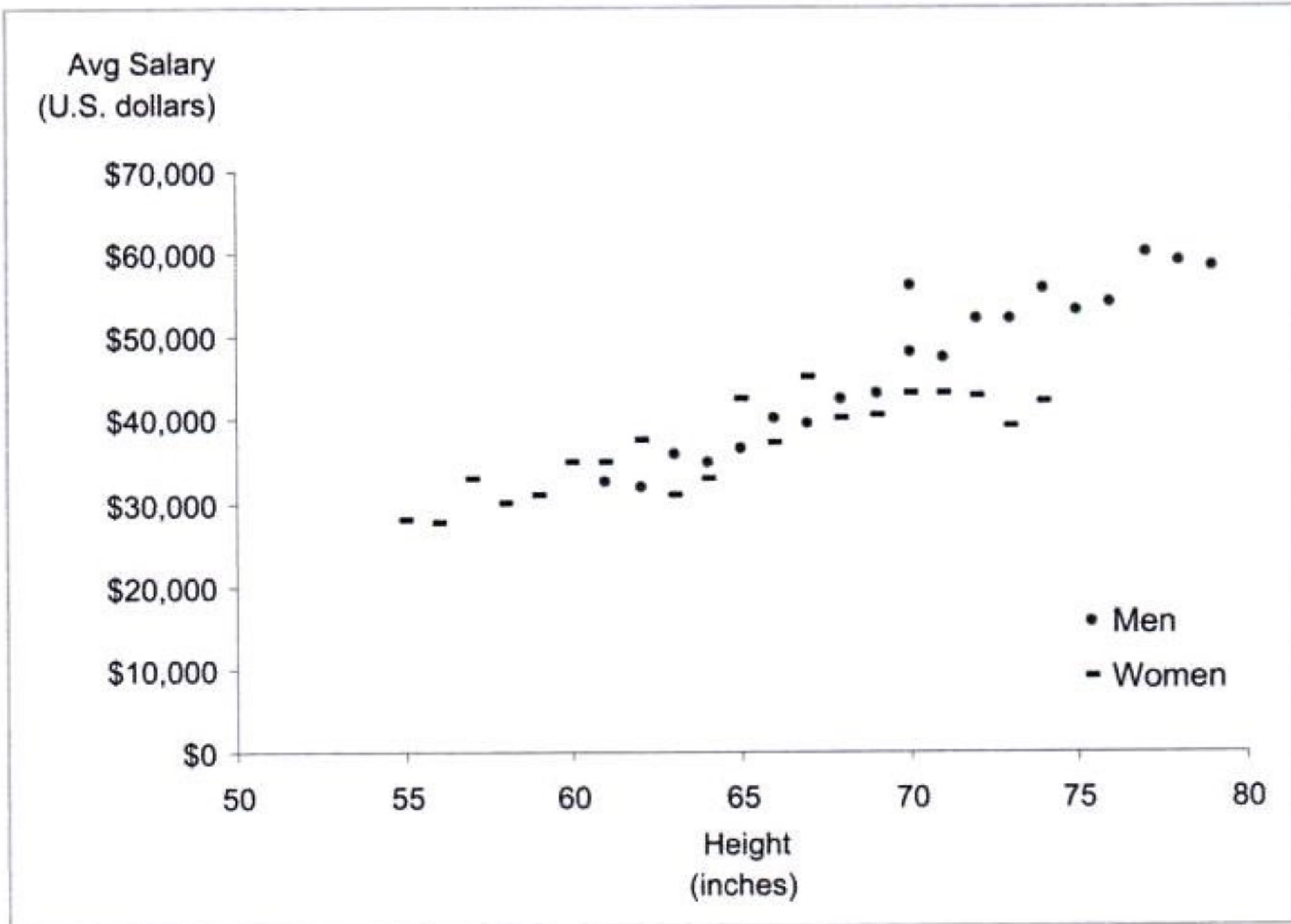
---

- Faça com que os pontos sejam distintamente visíveis

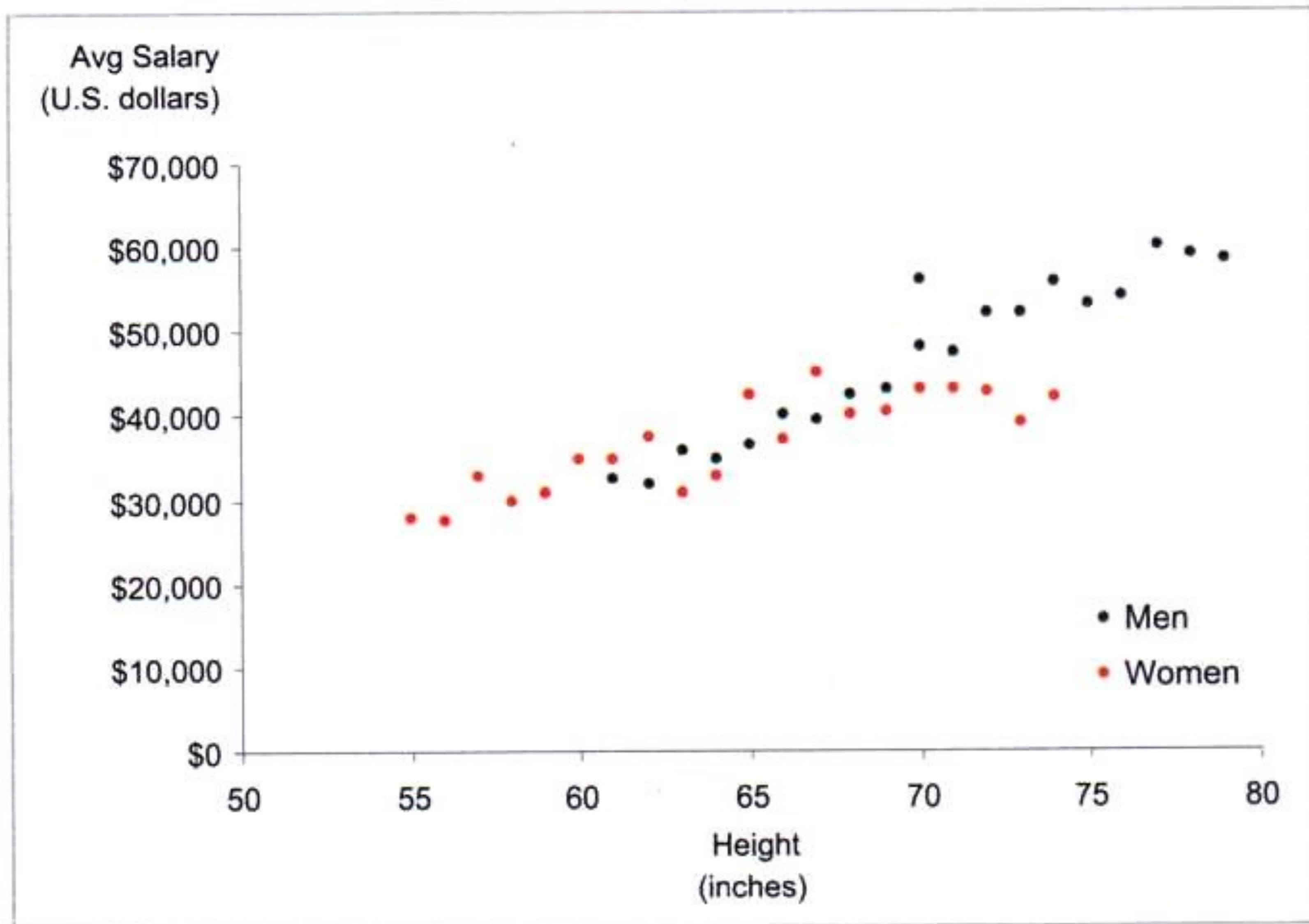




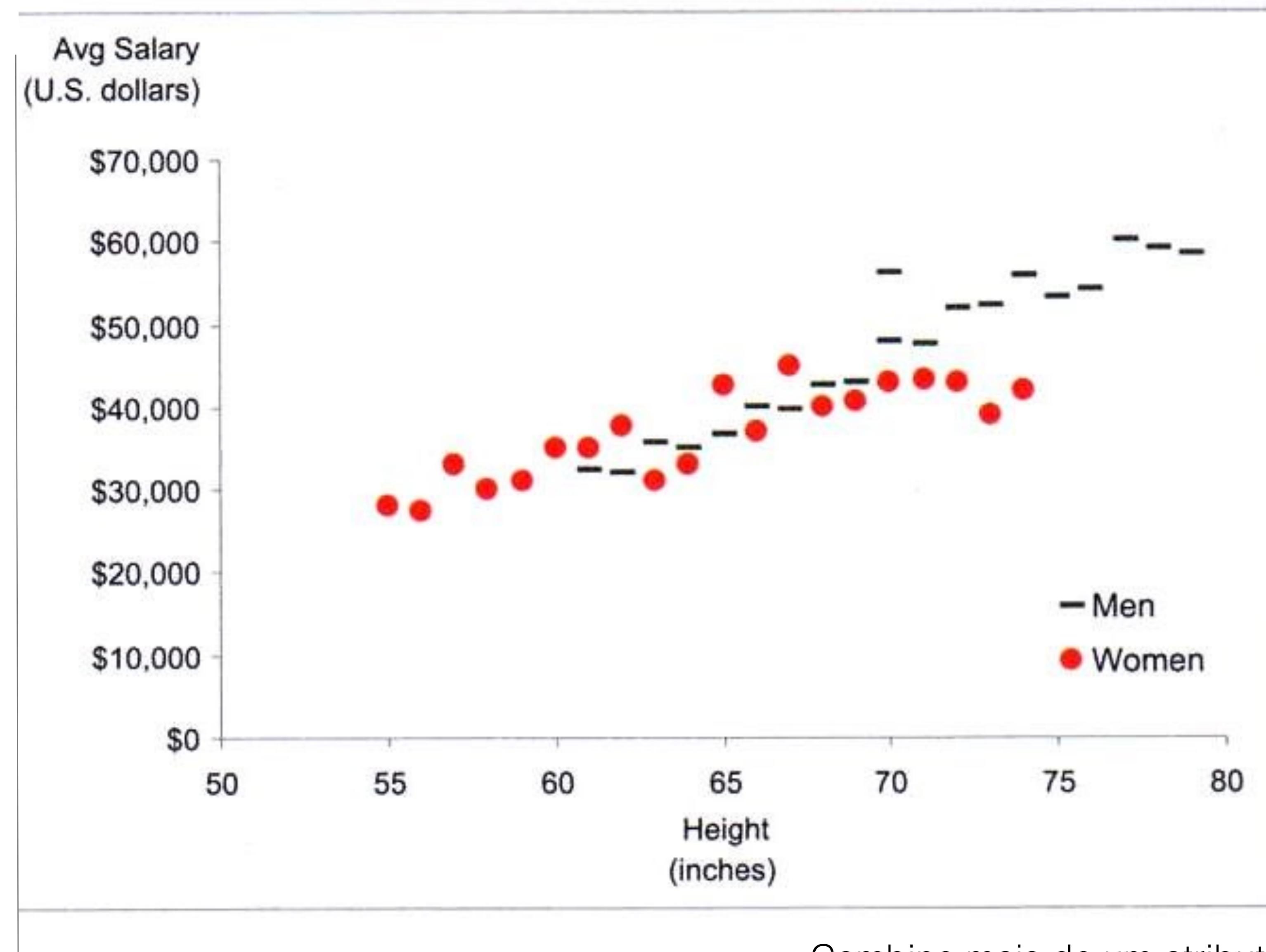
Aumente o tamanho dos pontos



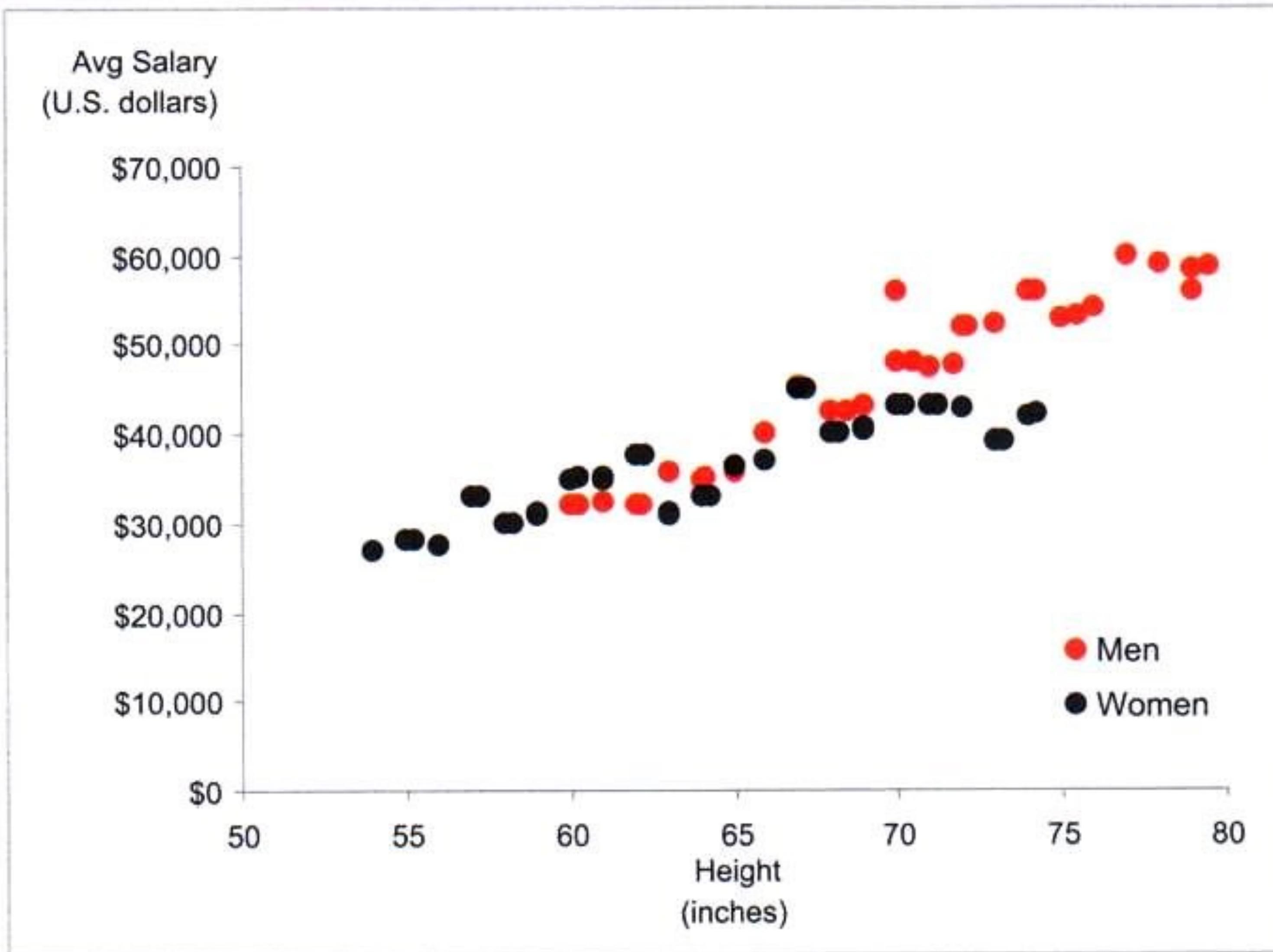
Use formas mais visualmente distintas

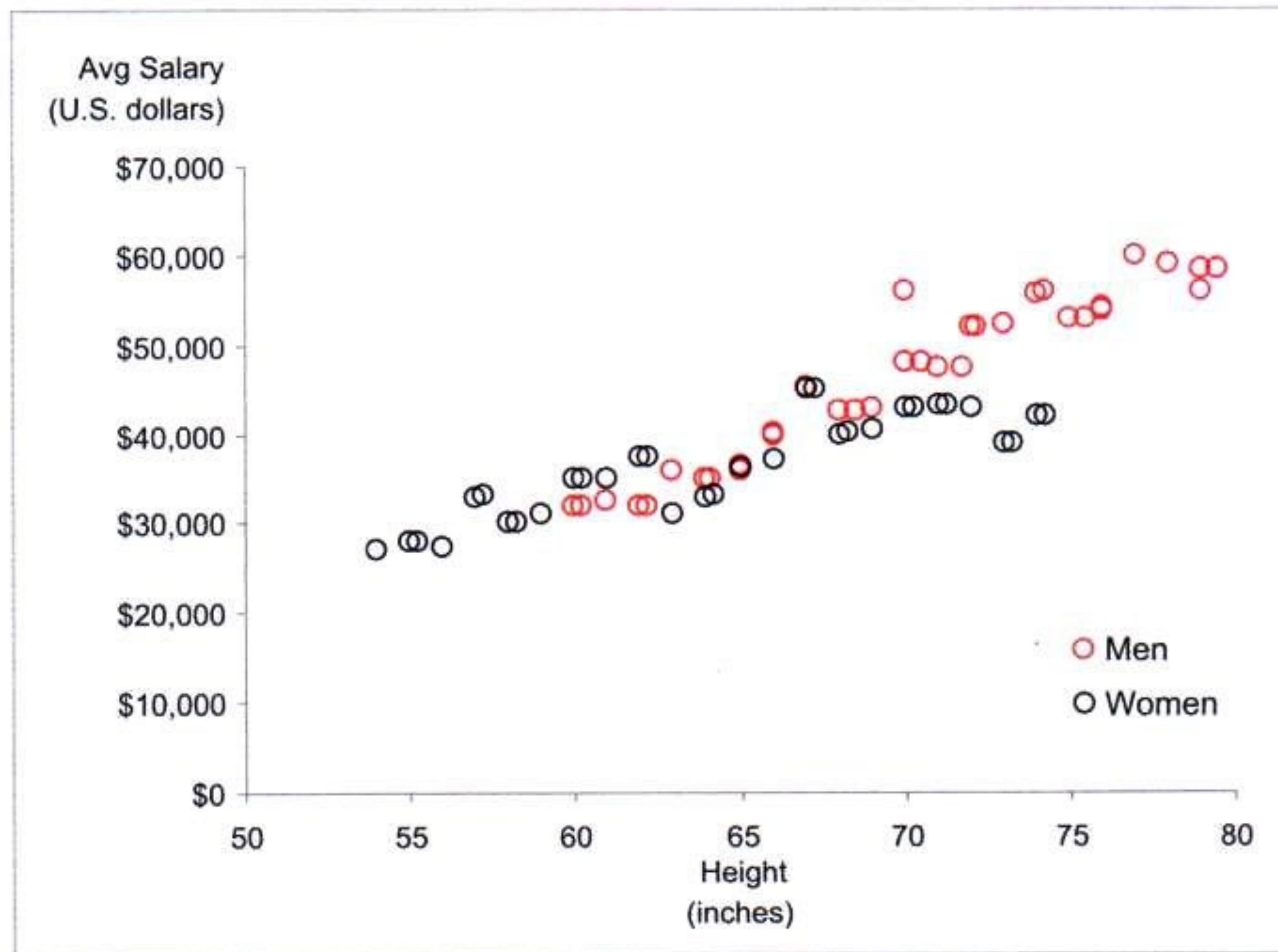


Use cores diferentes

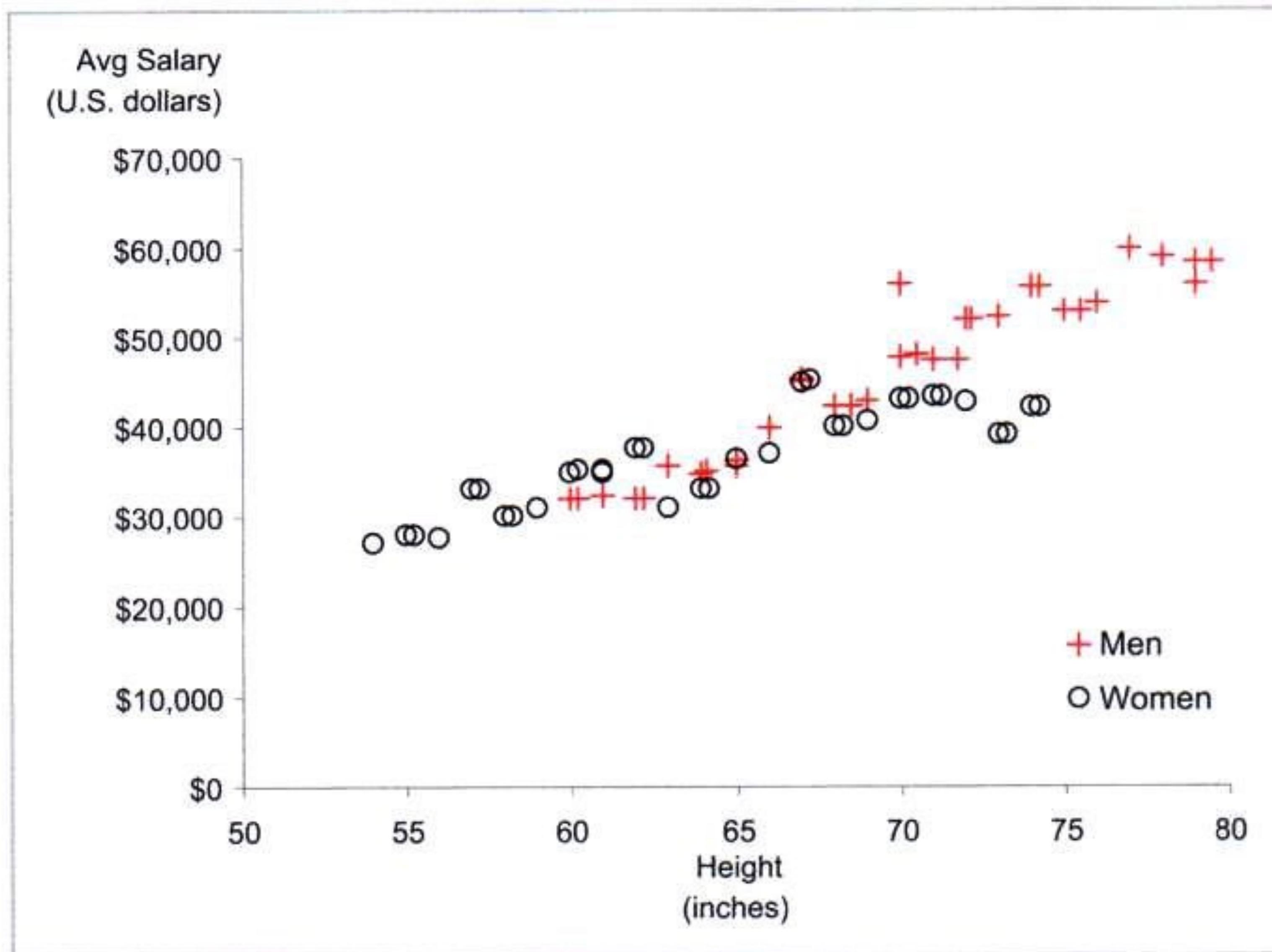


Combine mais de um atributo

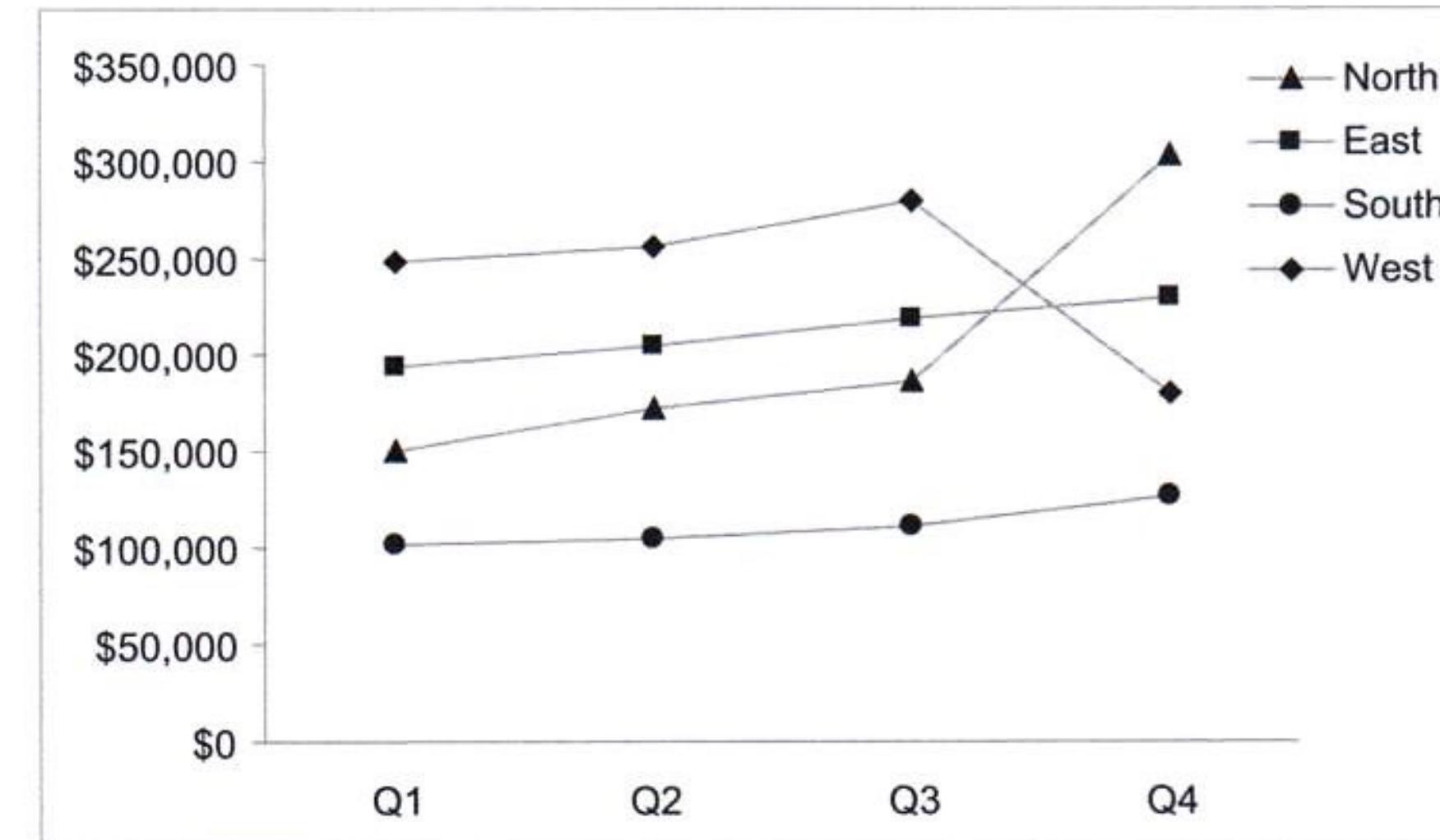
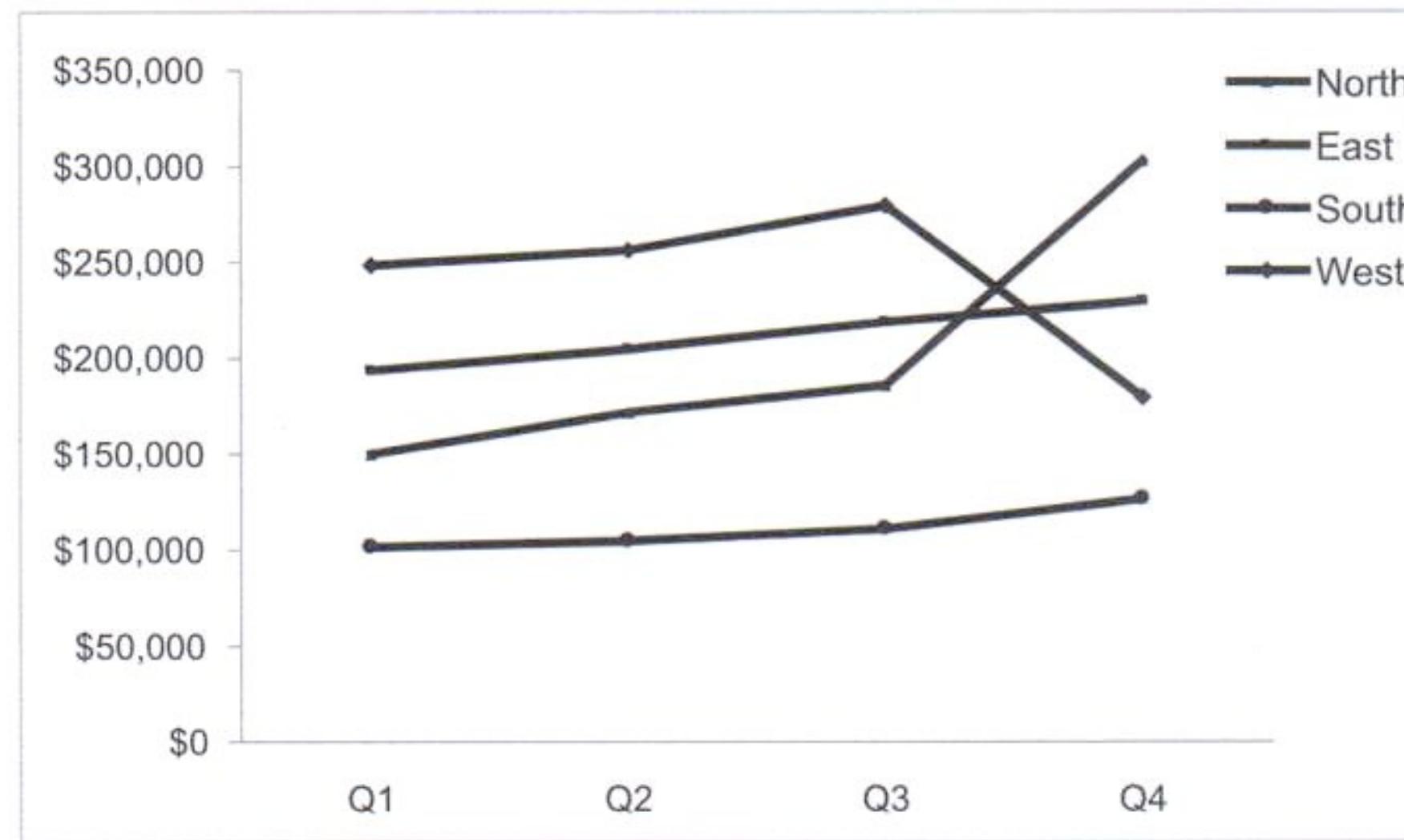




Use pontos sem preenchimento



E use objetos mais distintos

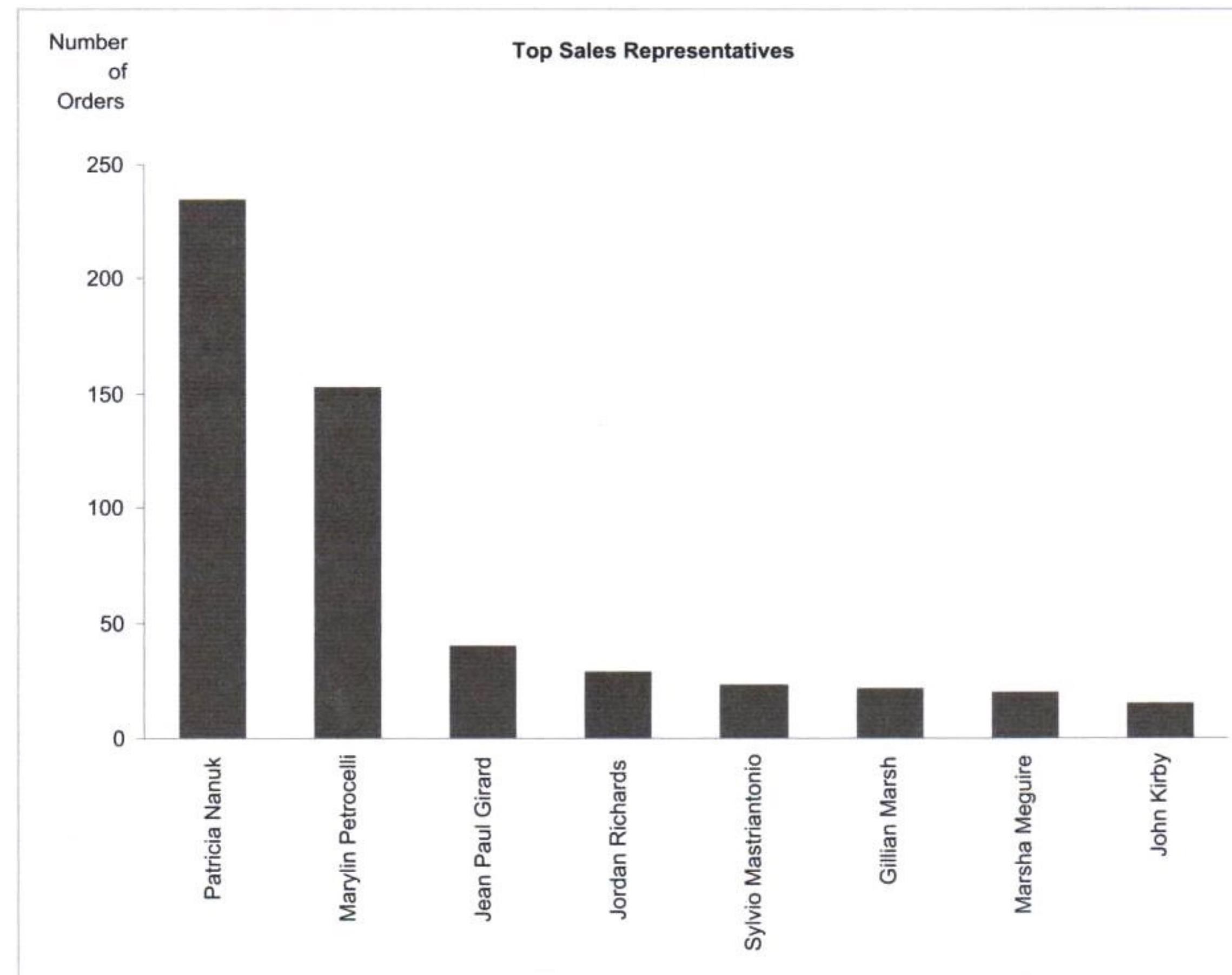
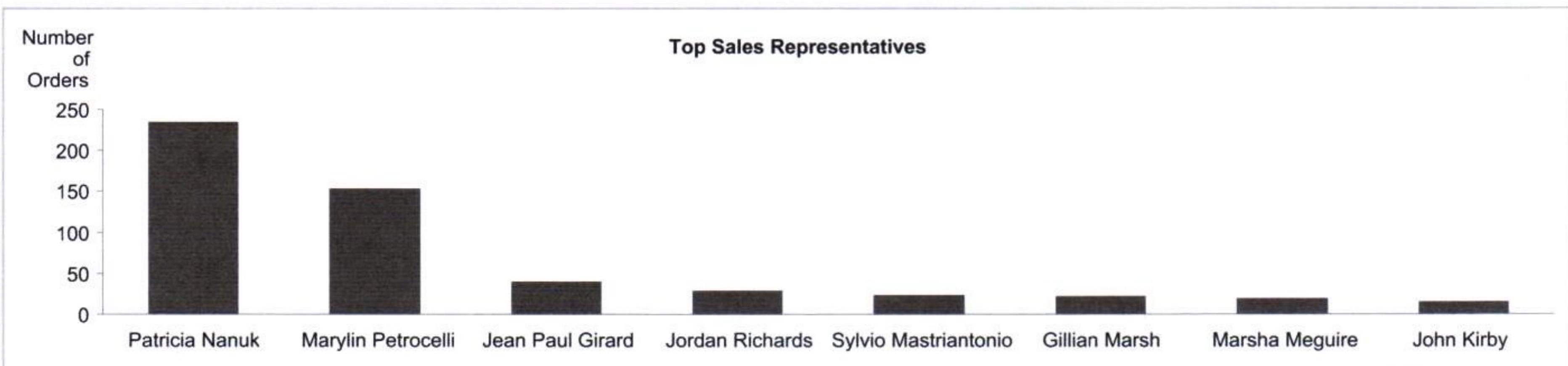


- Aumente os pontos e escolha objetos mais distintos
- Faça os pontos se destacarem com relação às linhas

# BARRAS

---

- Vários atributos das barras merecem atenção:
  - Orientação
  - Proximidade
  - Preenchimento
  - Bordas
  - Bases



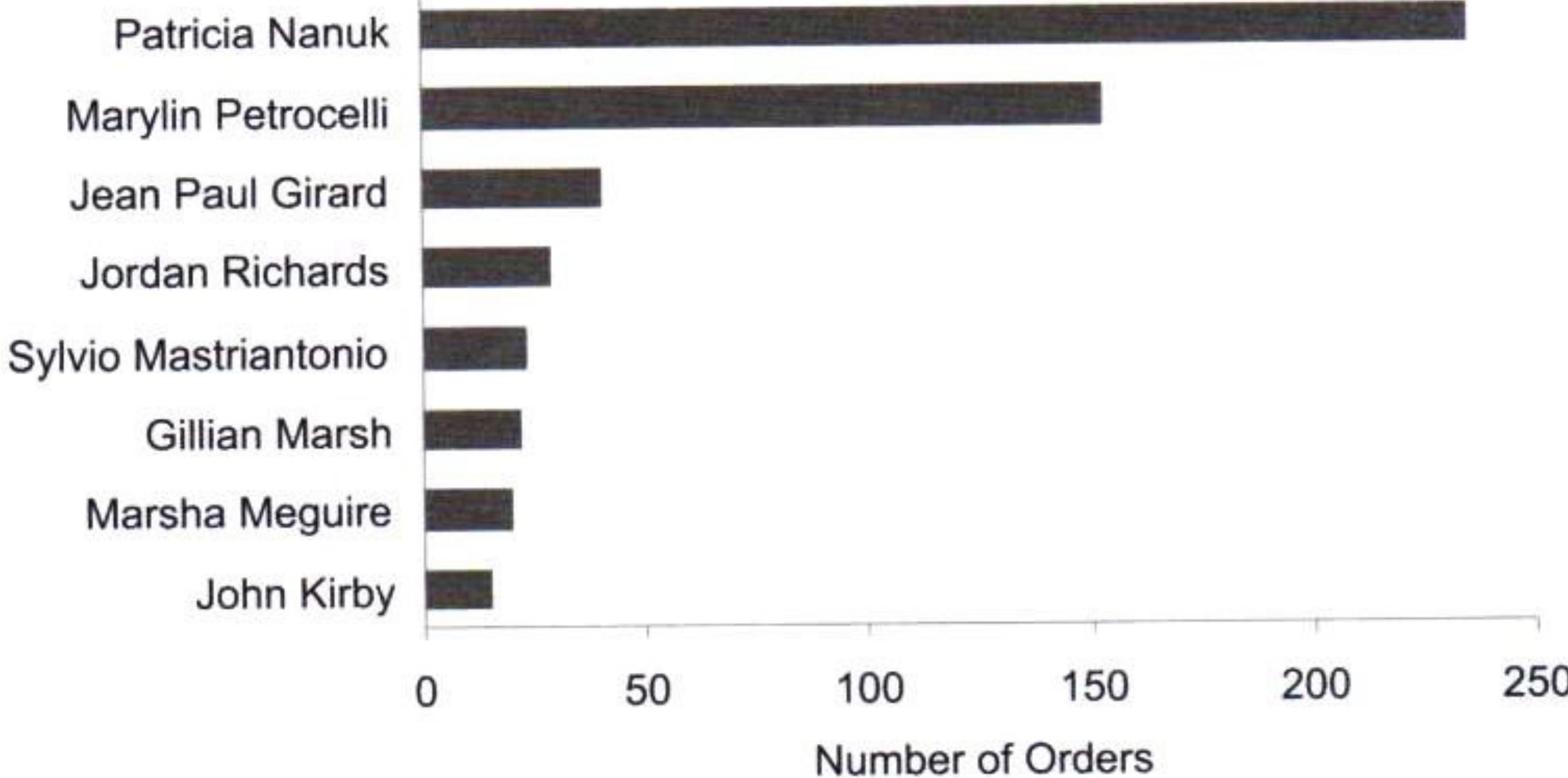
# BARRAS

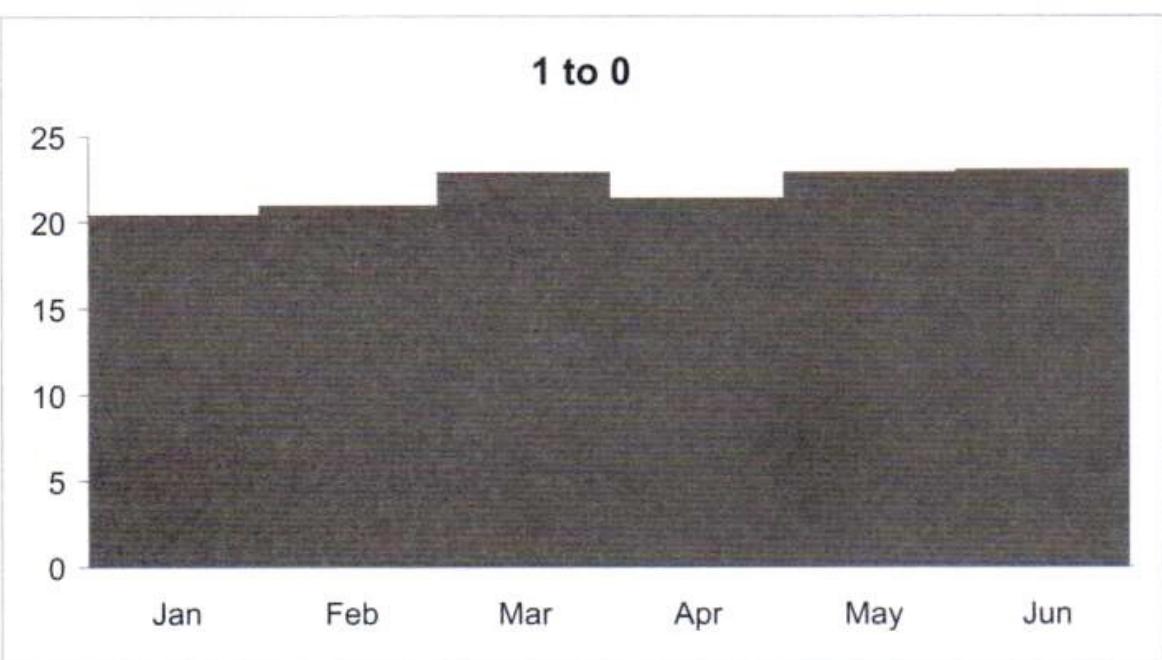
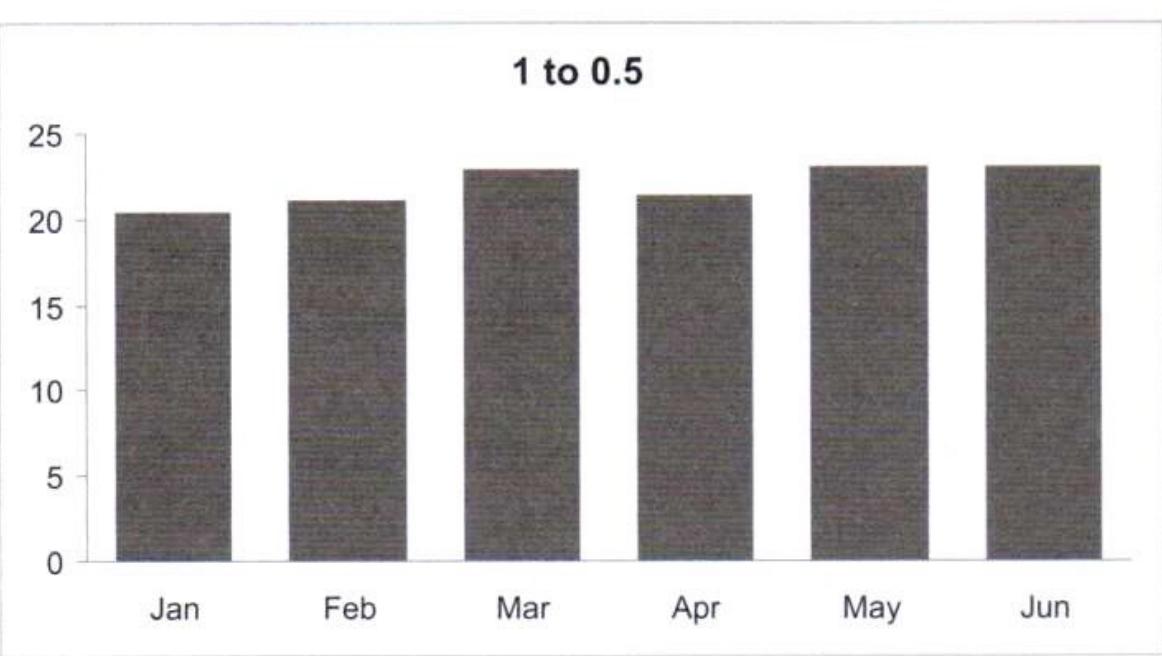
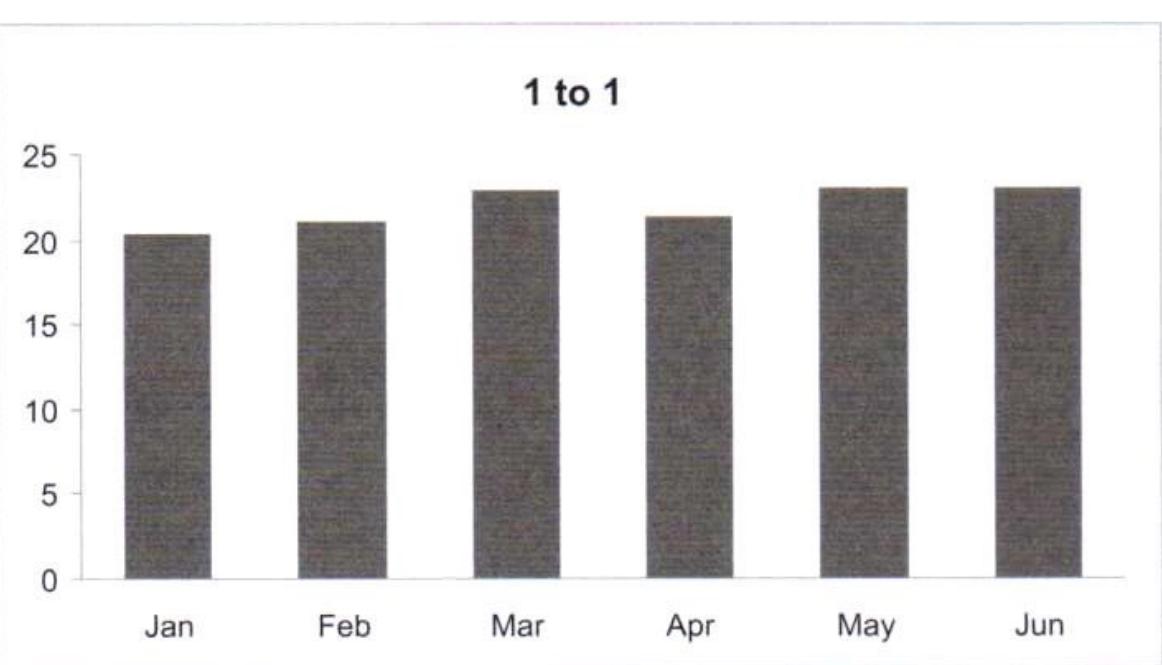
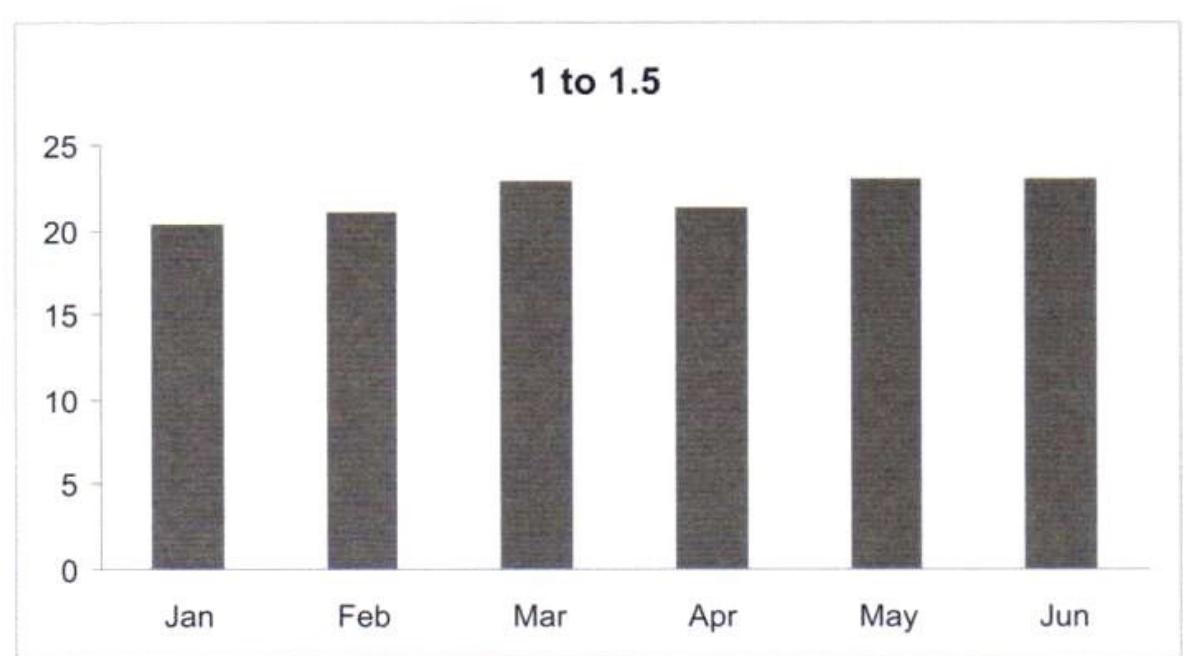
---

## ORIENTAÇÃO

- Barras horizontais são a melhor escolha quando:
  - O gráfico mostra relacionamento de *ranking*
  - Os rótulos das categorias não cabem na versão vertical

### Top Sales Representatives



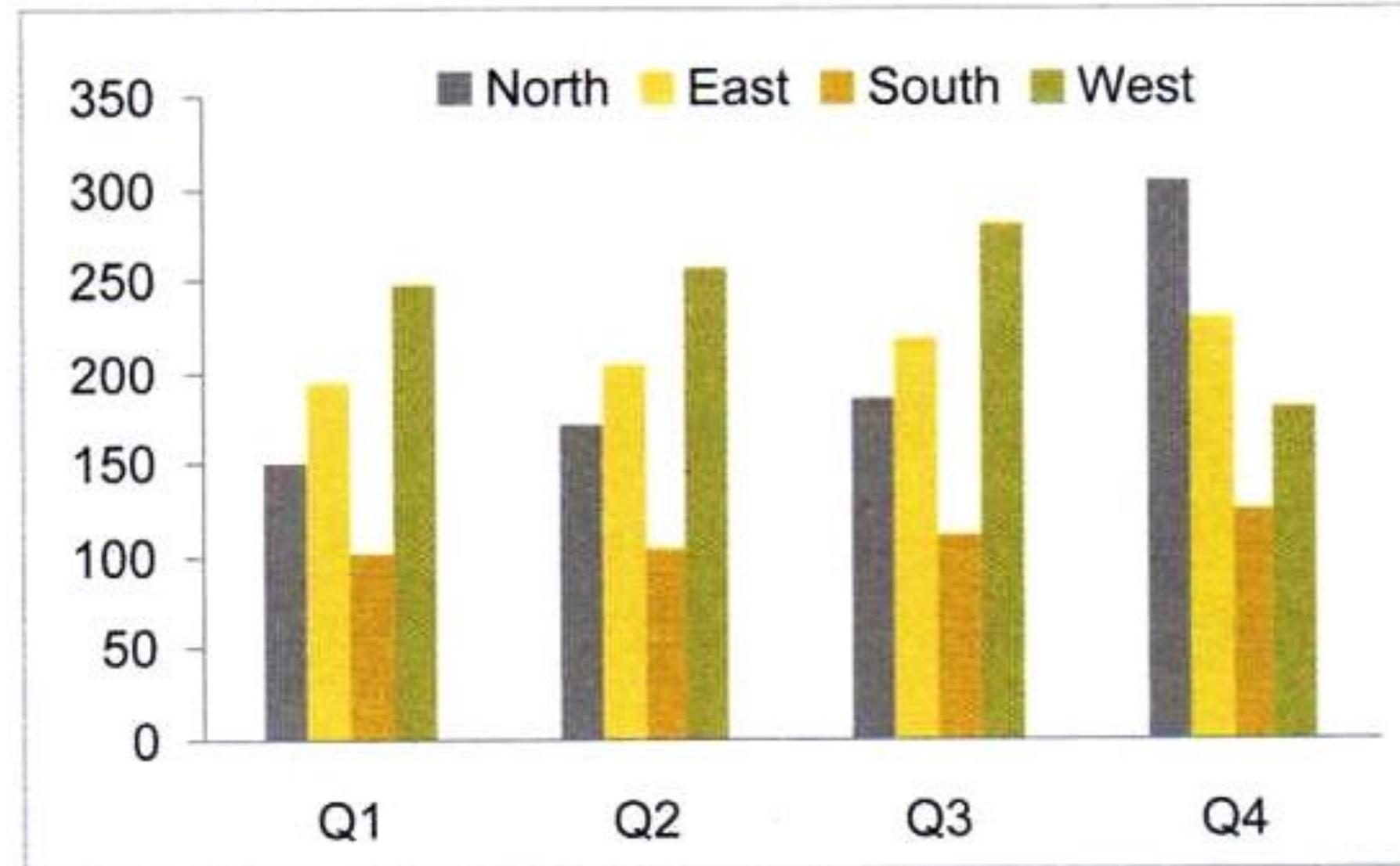


# BARRAS

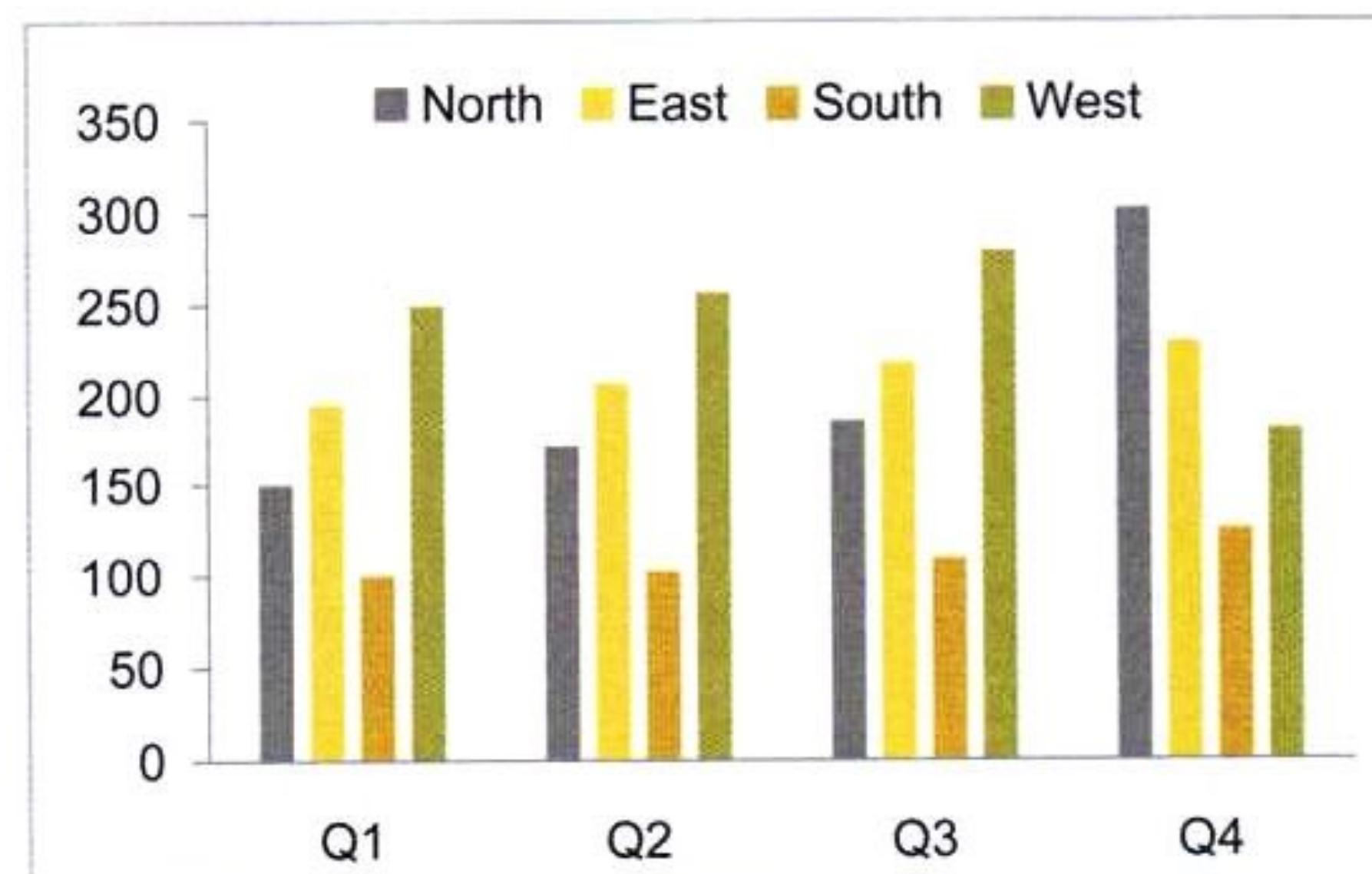
---

## PROXIMIDADE

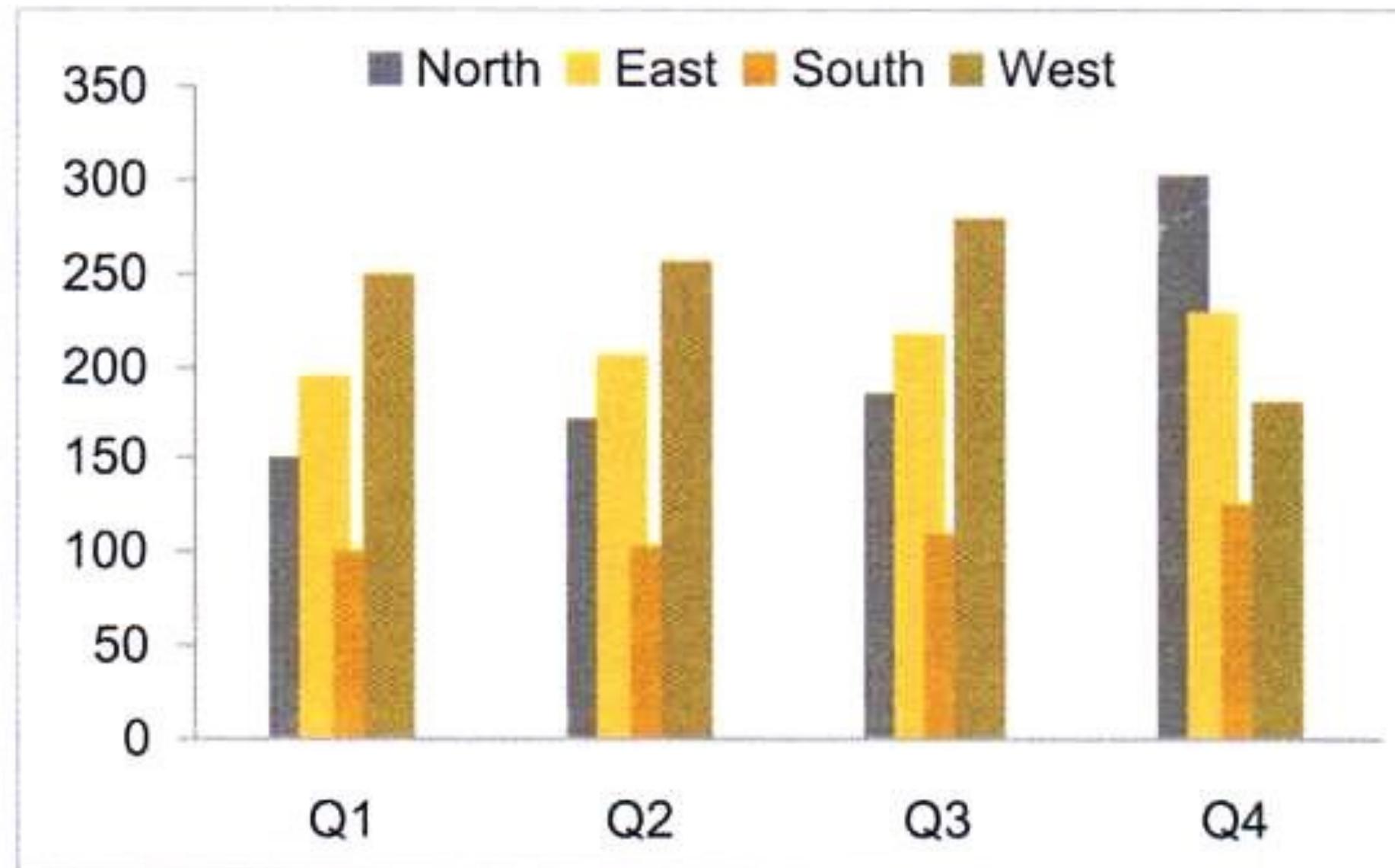
- Manter a razão entre a largura das barras e o espaçamento entre 1:1,5 e 1:0,5



Uma exceção ocorre quando as barras representam subdivisões categóricas (barras agrupadas)



Neste caso, o espaço não é necessário



Nunca sobreponha barras...



...exceto quando  
quiser analisar  
correlações

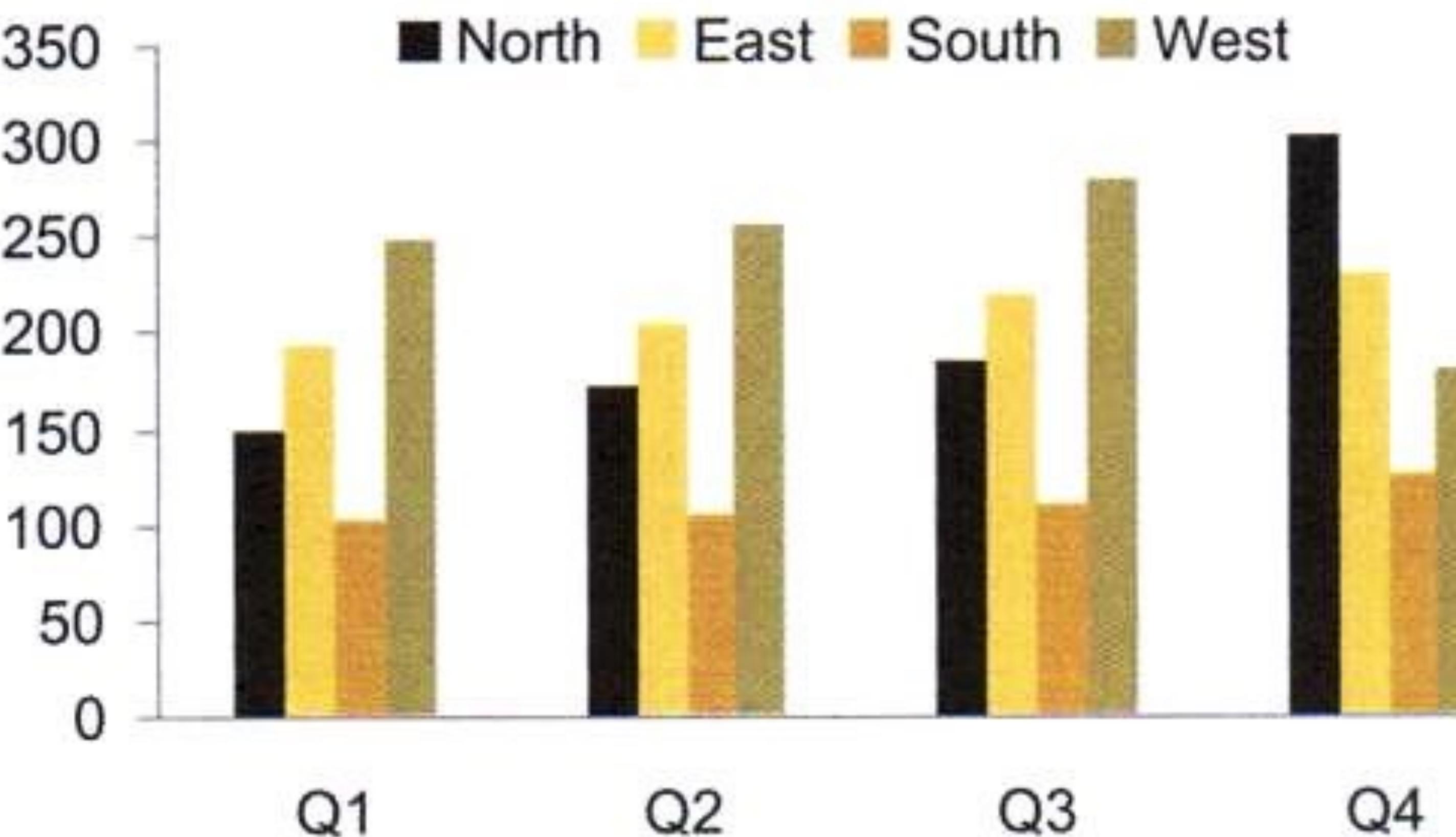
# BARRAS

---

## PREENCHIMENTO

- Preenchimento
- Evite usar padrões pois eles comumente provocam efeitos visuais indesejados
- Use cores de preenchimento claramente distintas
- Use cores balanceadas em termos de intensidade para dados que tenham a mesma importância
- Use cores mais intensas apenas quando desejar realçar valores particulares

## **North Region Exhibits Greatest Growth!**



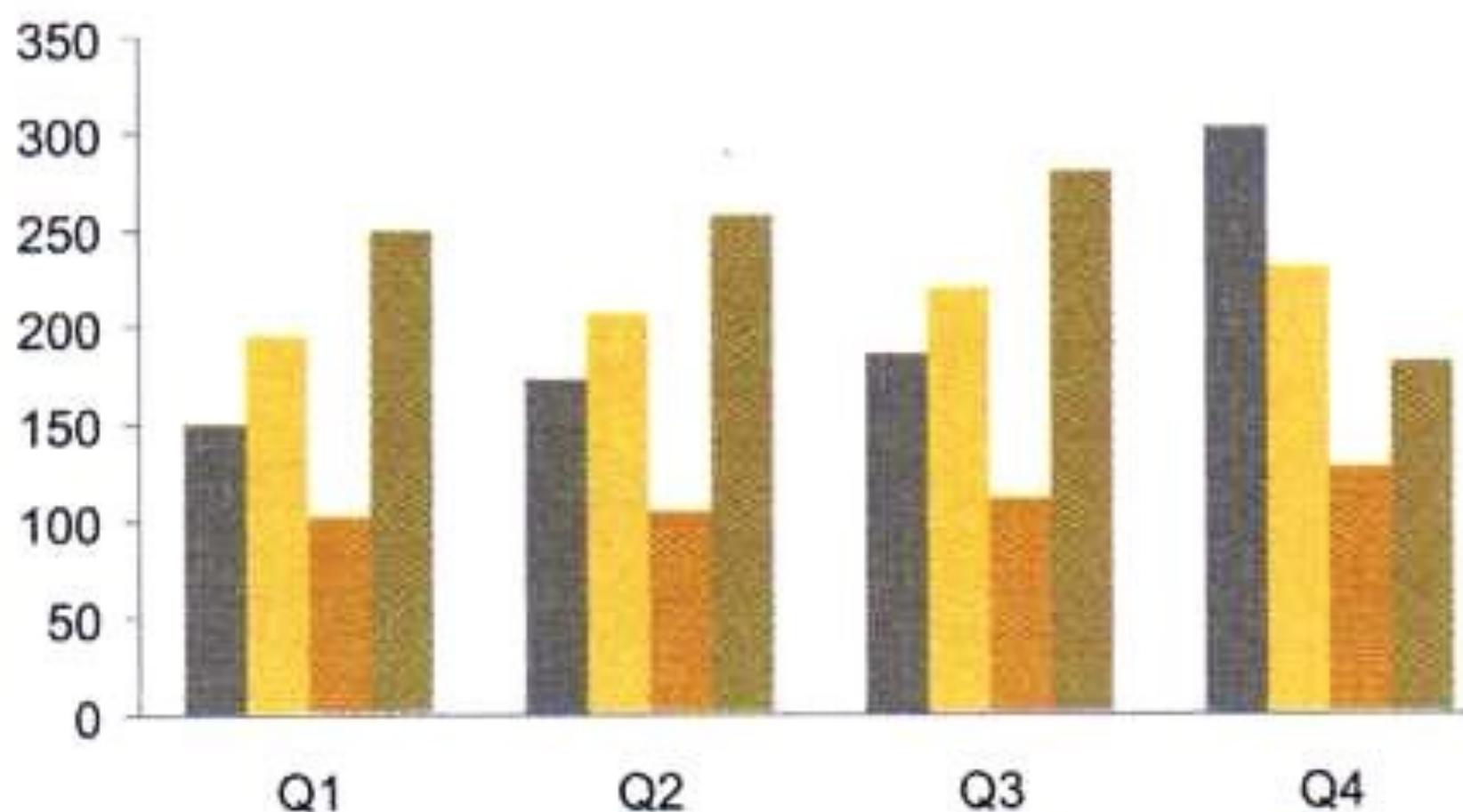
# BARRAS

---

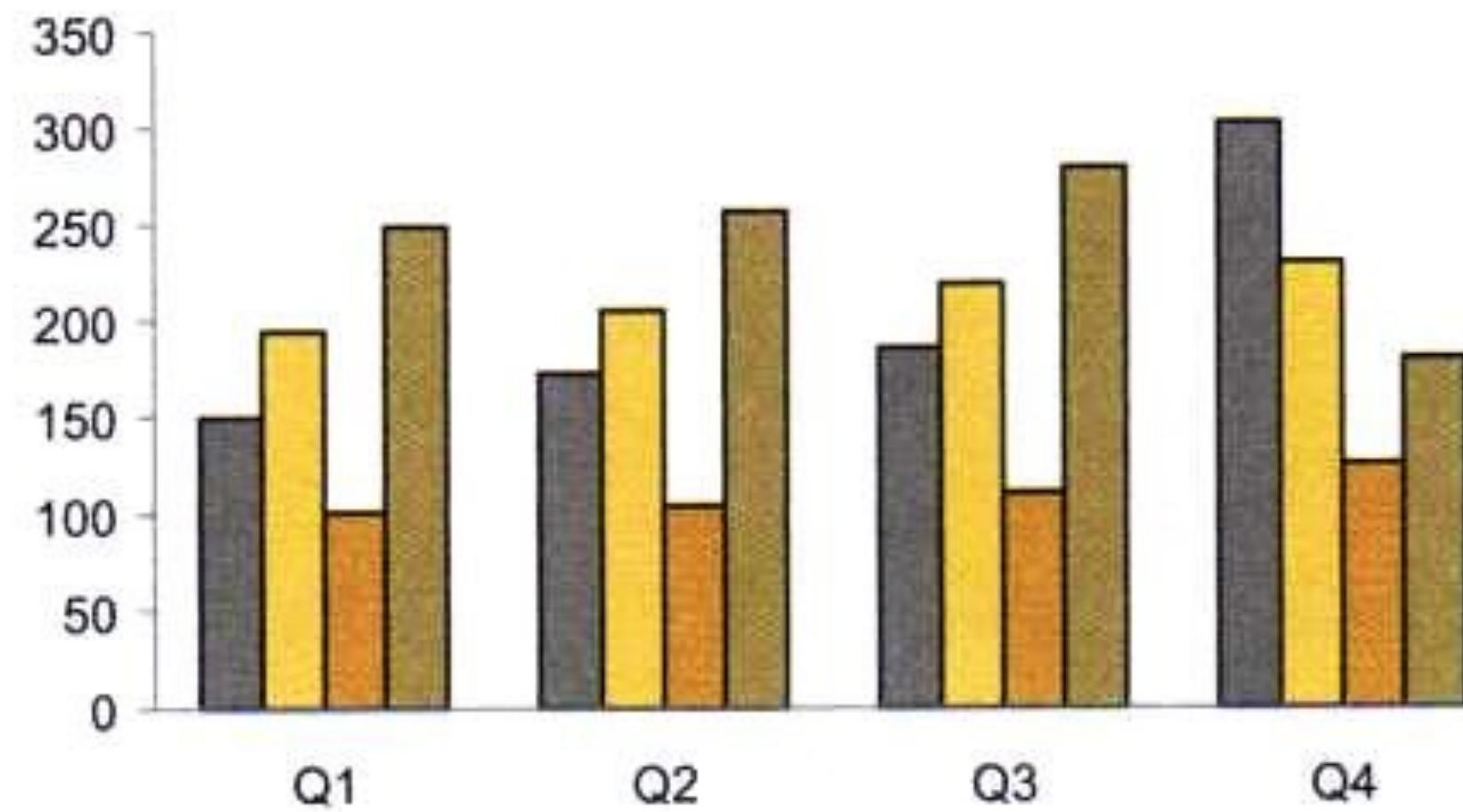
## BORDAS

- Componente visual que não adiciona informação
- Útil apenas quando a cor da barra pode ser confundida com a cor do fundo
- Útil para destacar valores

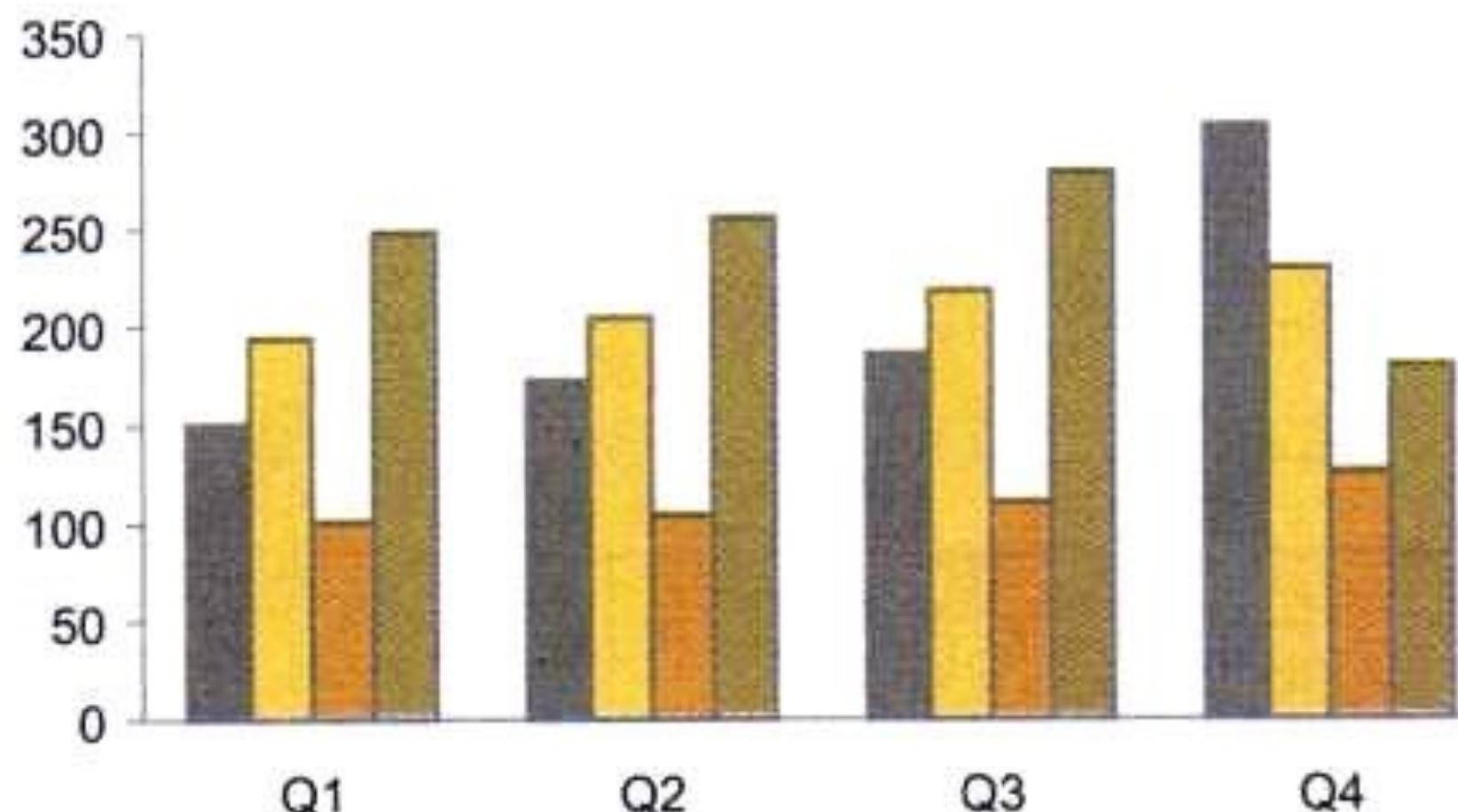
**No bar borders**



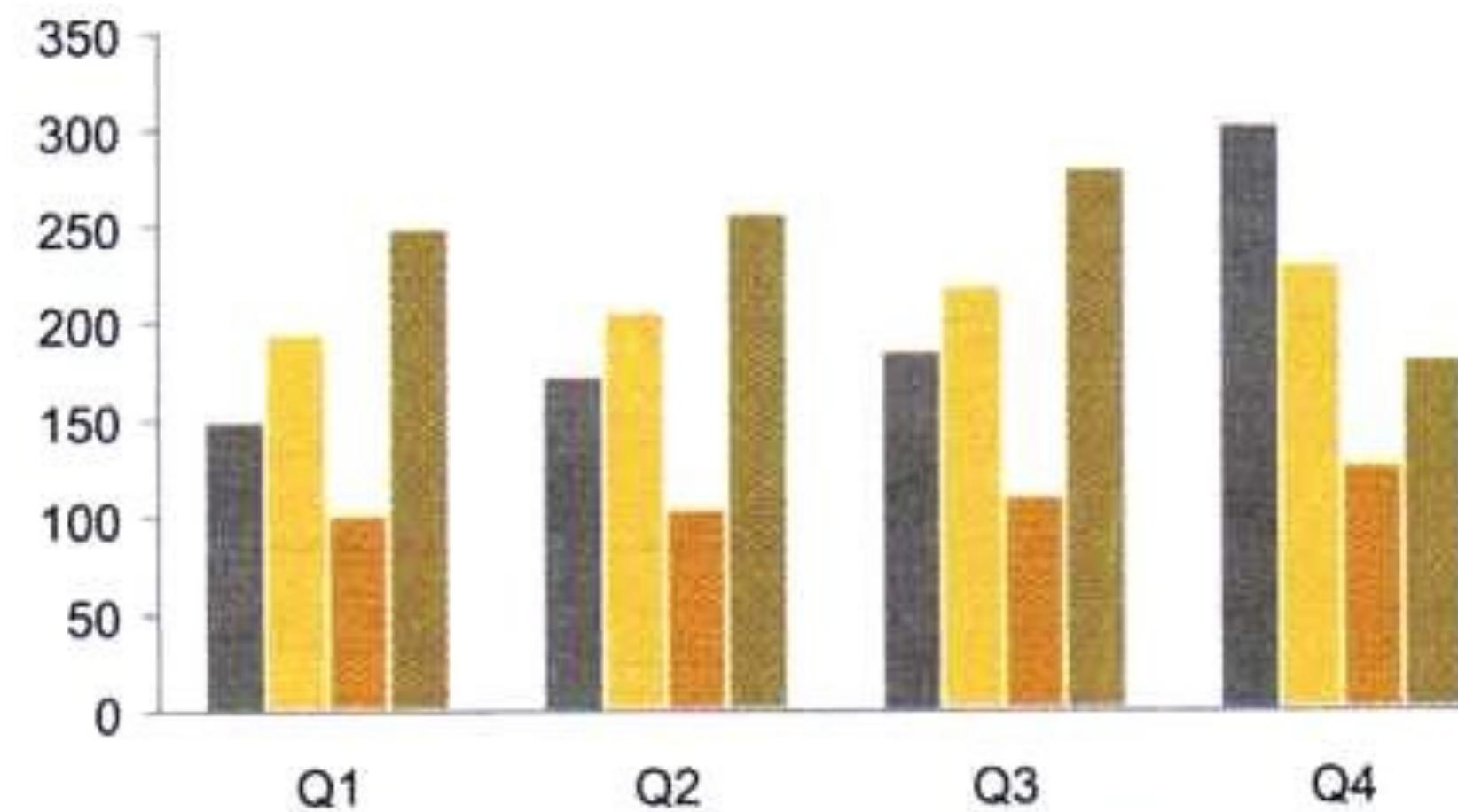
**Black bar borders**

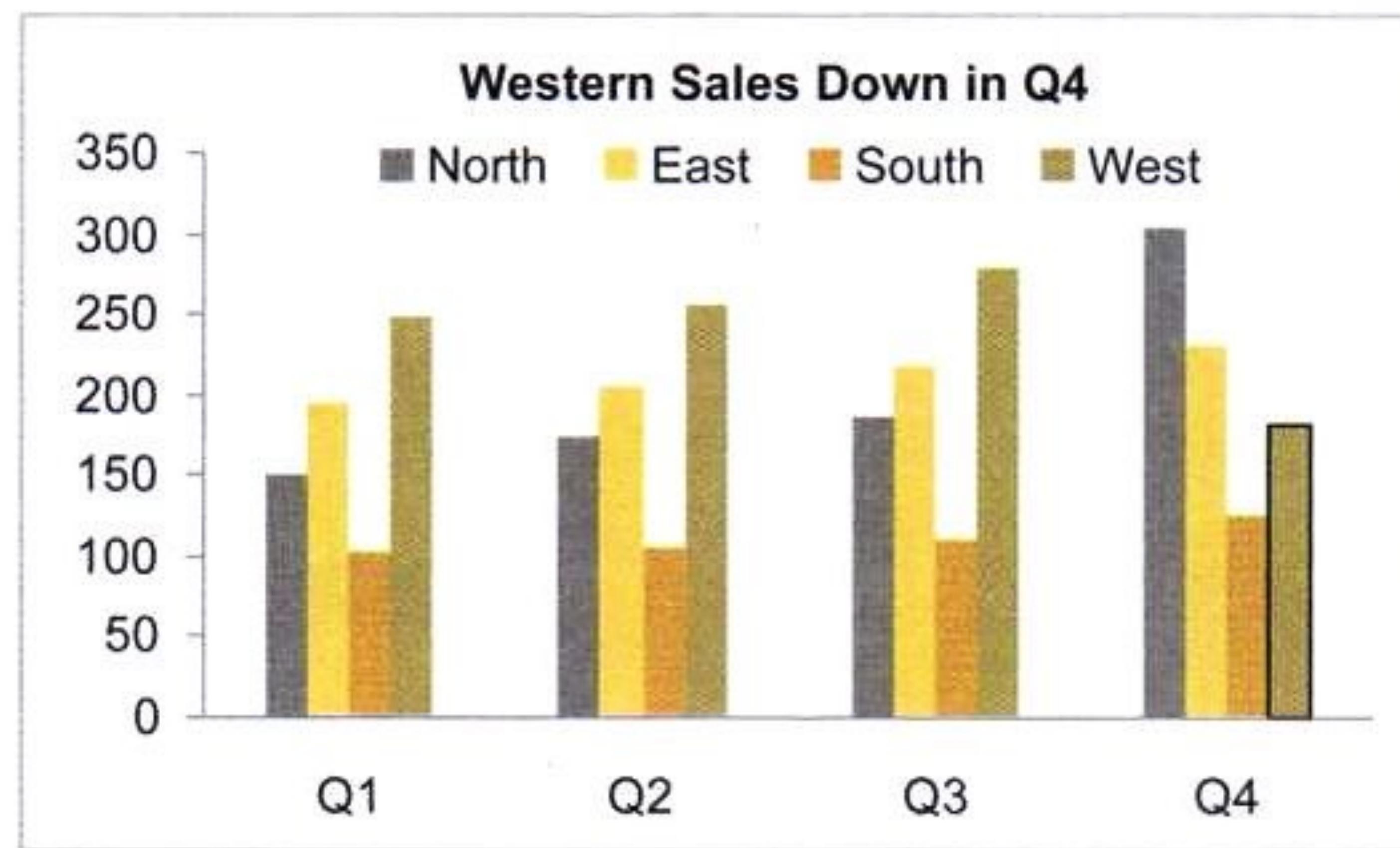


**Gray bar borders**



**White bar borders**



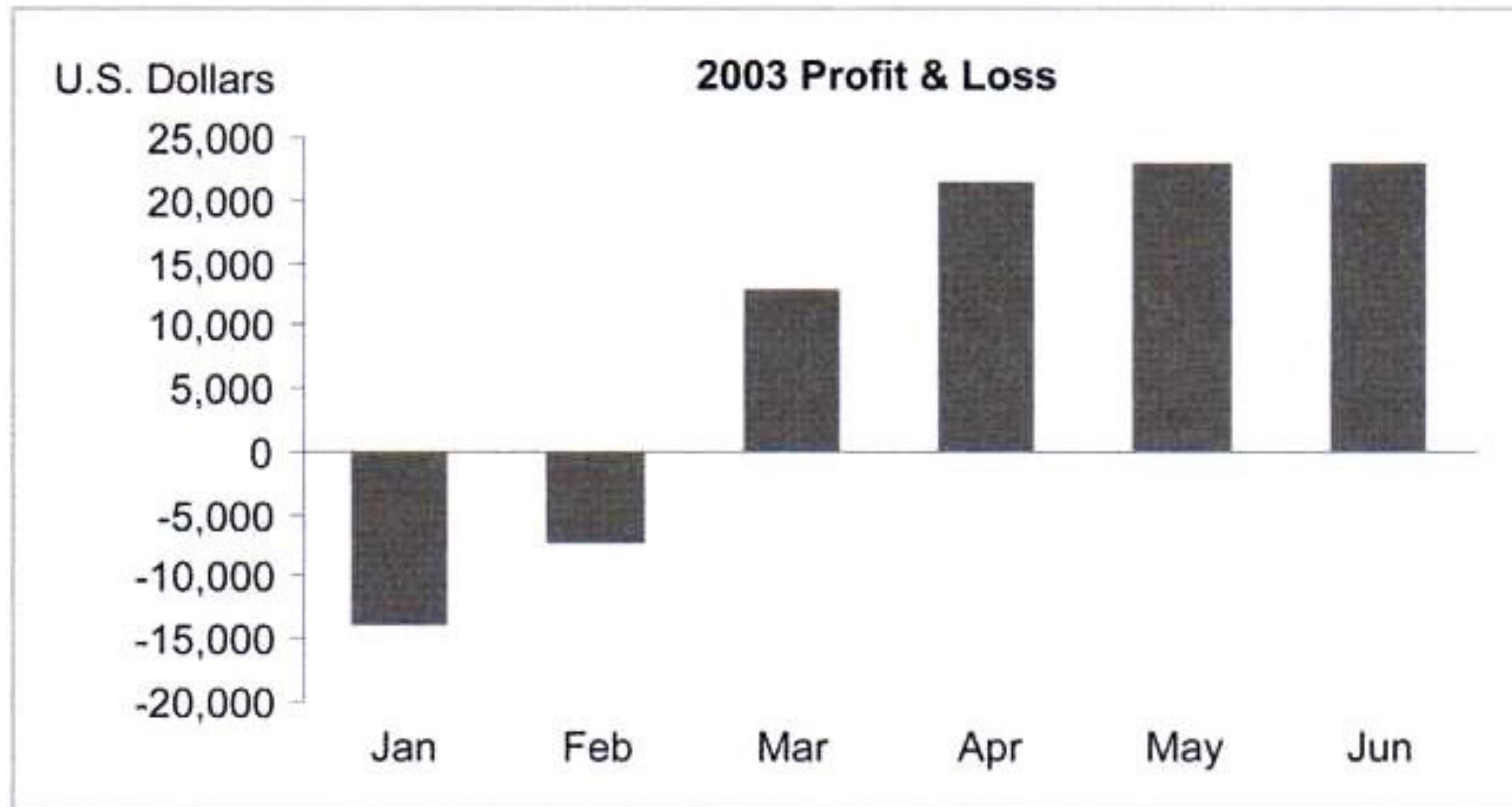


# BARRAS

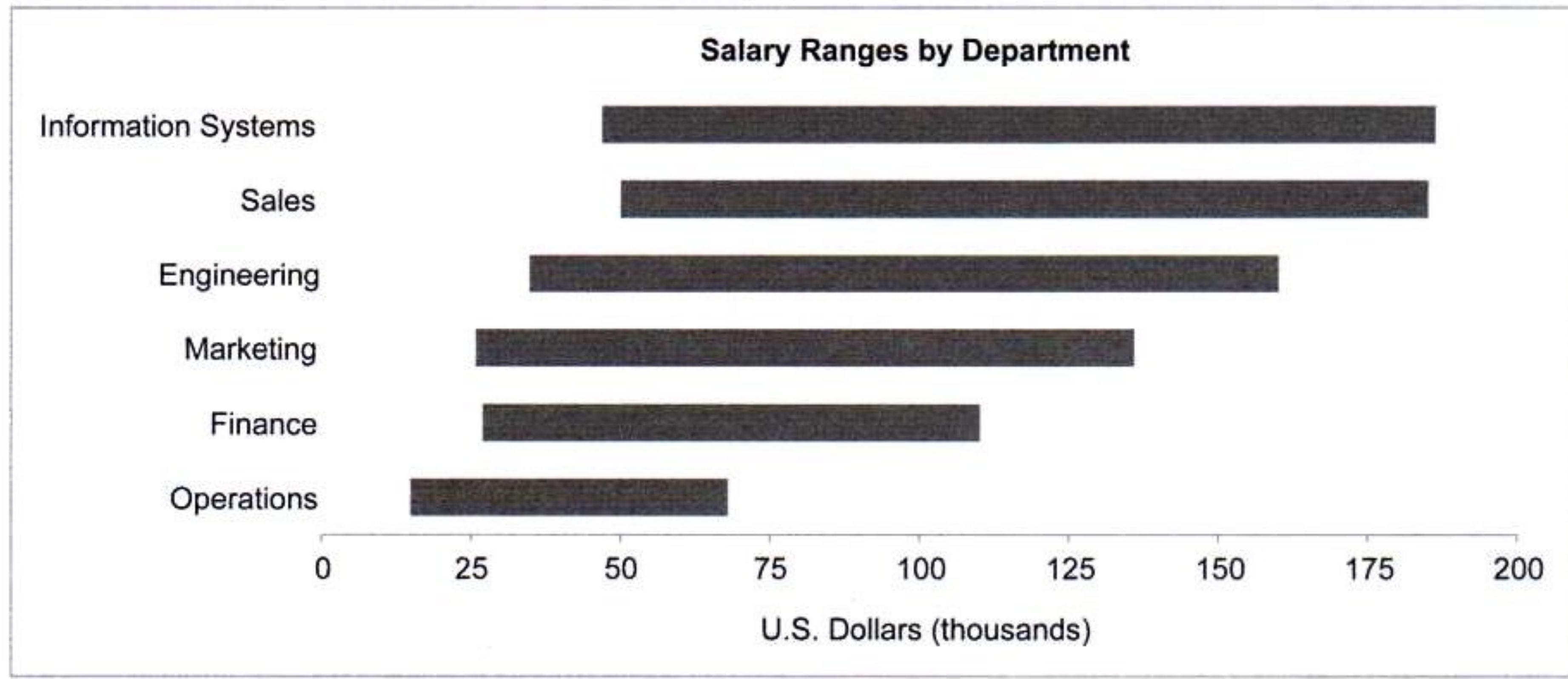
---

## BASES

- Onde devem começar as barras e por que?



Sempre de 0 uma vez que seu comprimento indica quantidade



Exceto quando indicam uma **faixa de valores**

Neste caso, seu **comprimento indica a amplitude do intervalo**

U.S. Dollars

### Minimum, Maximum, and Median Salaries by Department

150,000

125,000

100,000

75,000

50,000

25,000

Operations

Finance

Marketing

Engineering

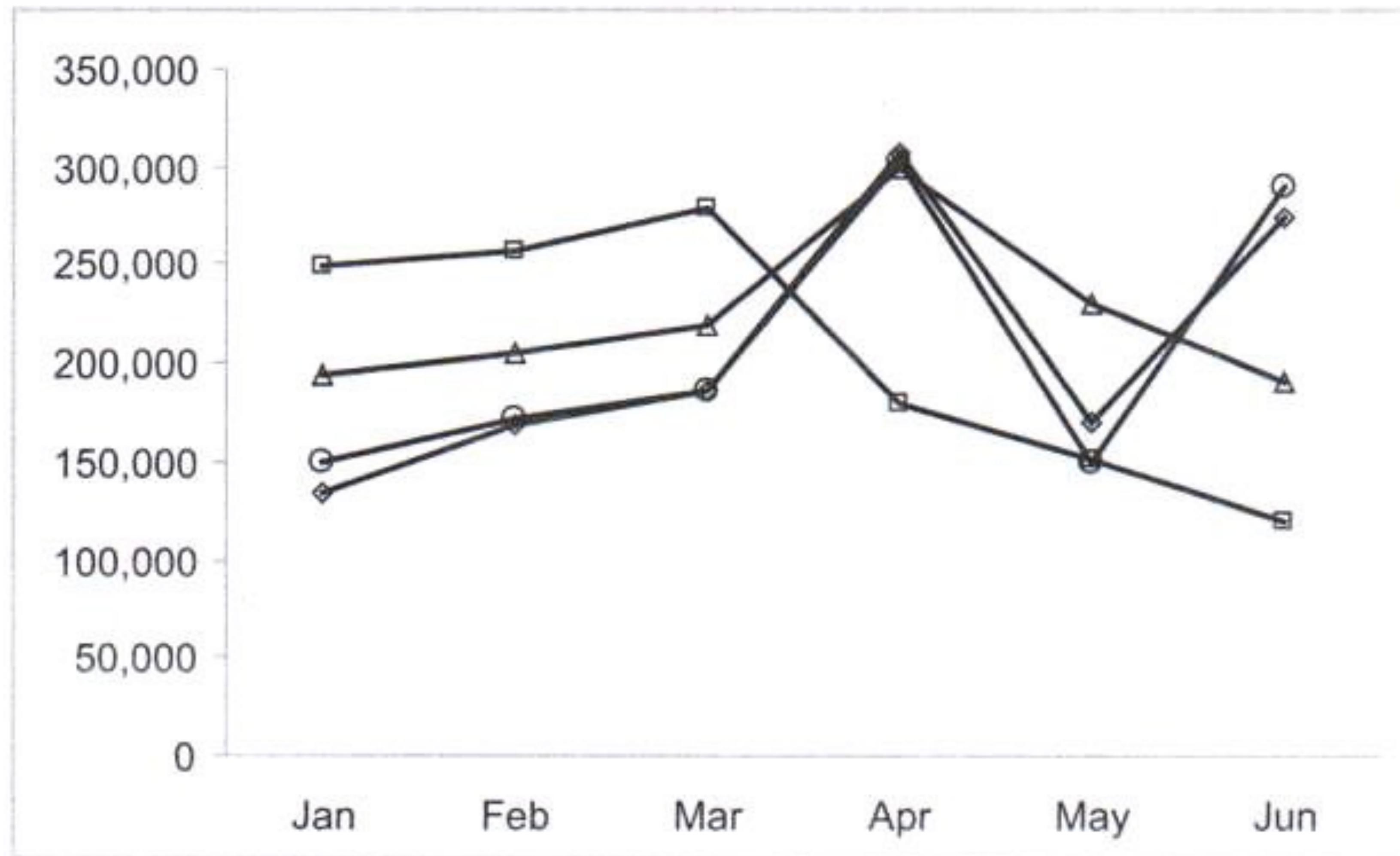
Sales

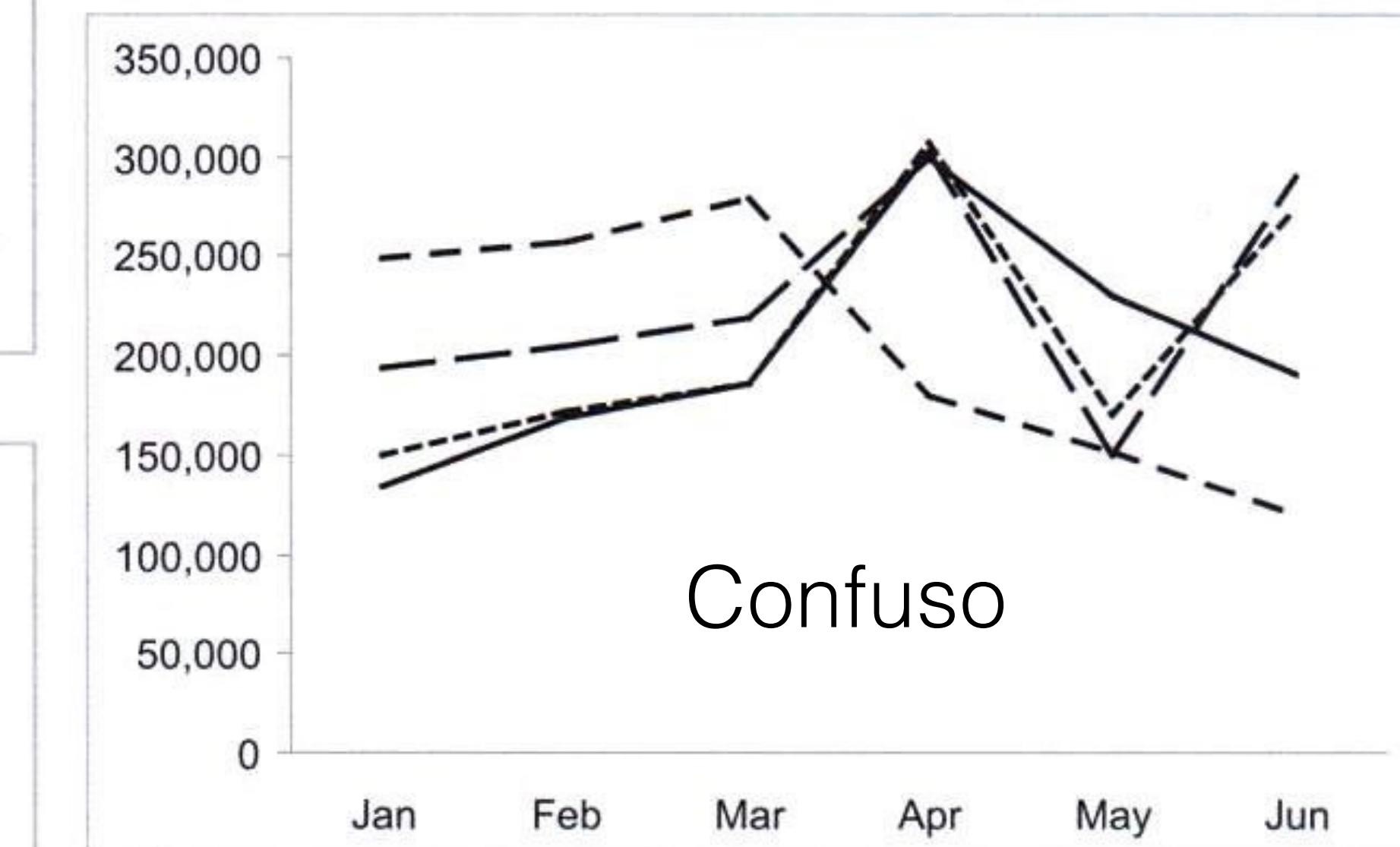
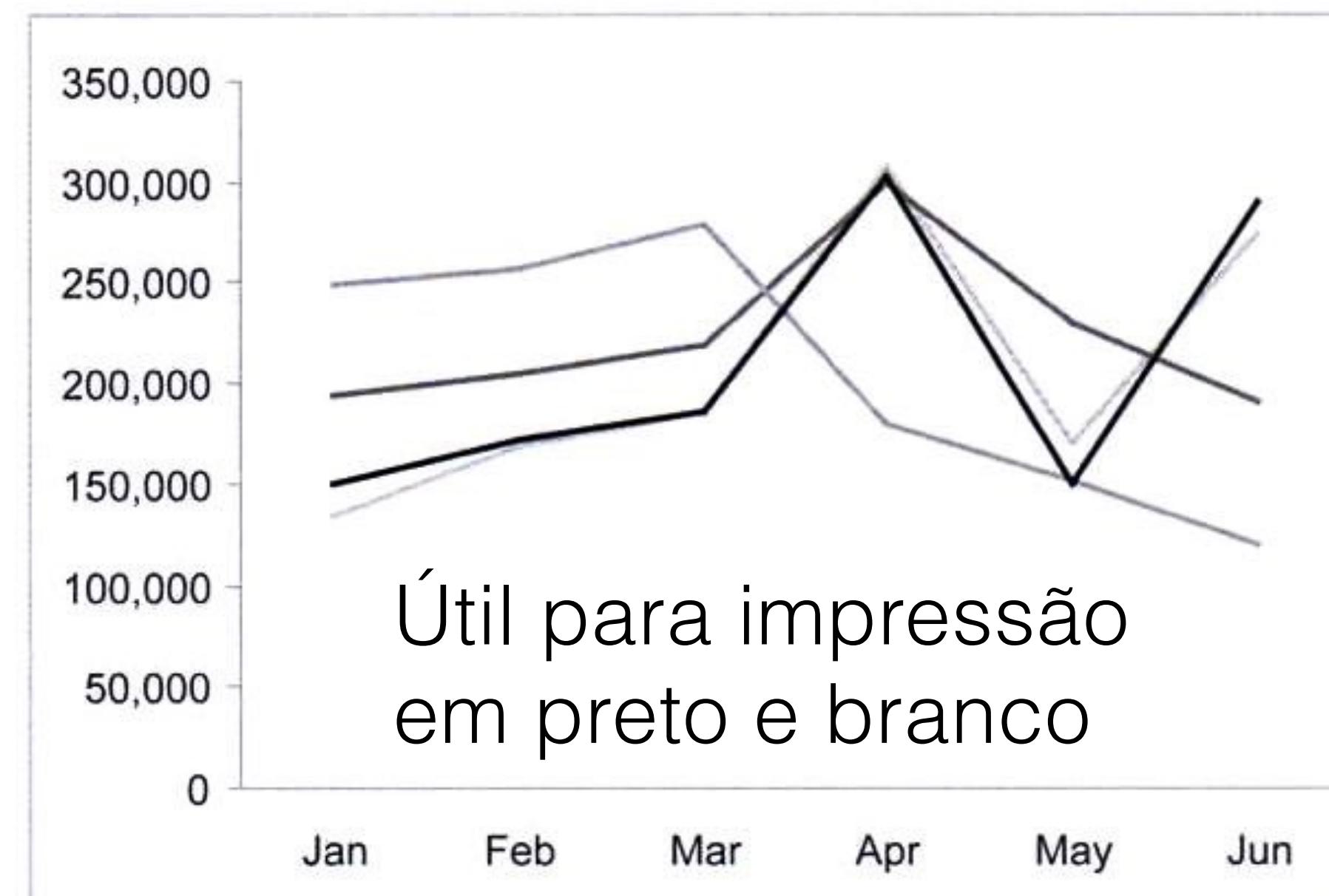
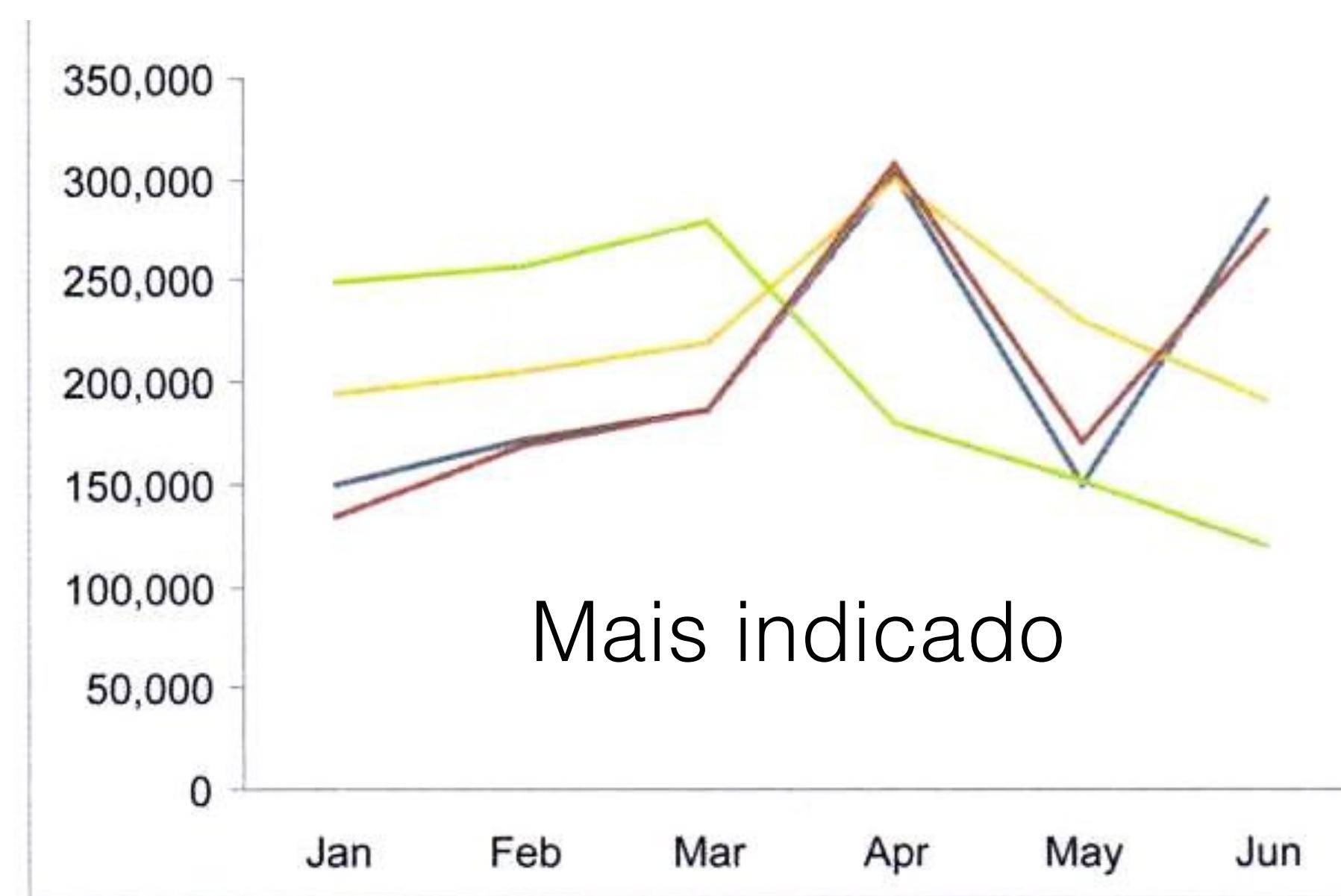
Information  
Systems

# LINHAS

---

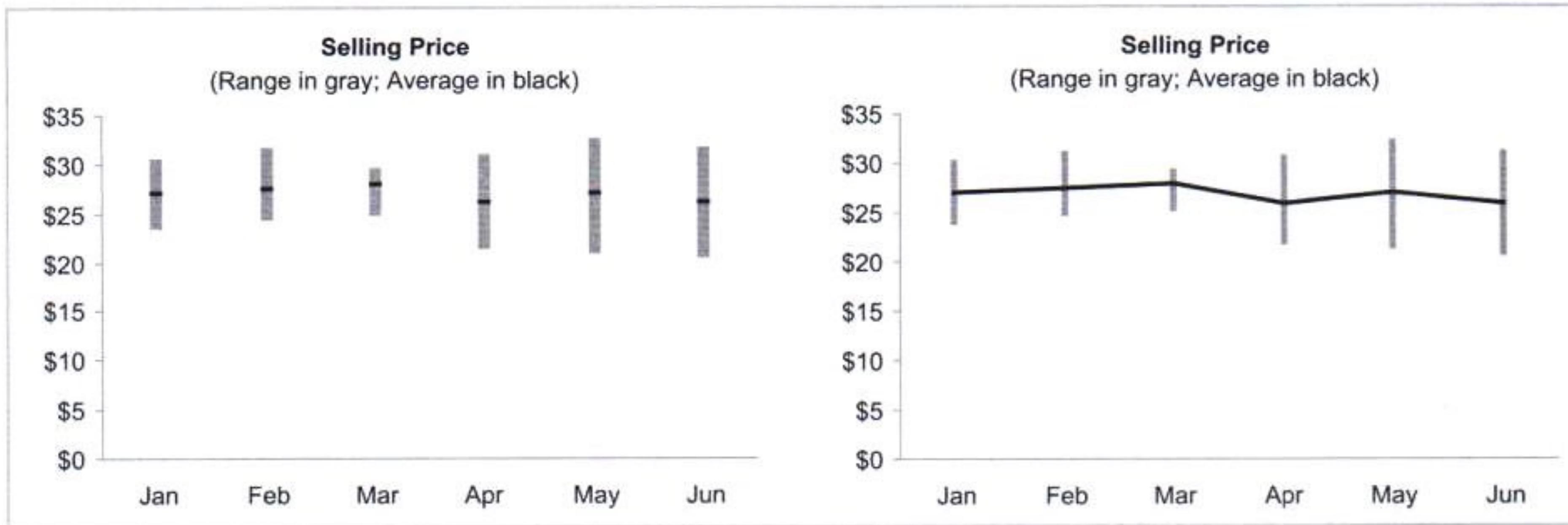
- Há basicamente quatro categorias:
  - Linha padrão
  - Linha max-min
  - Linha de tendência
  - Linha de referência

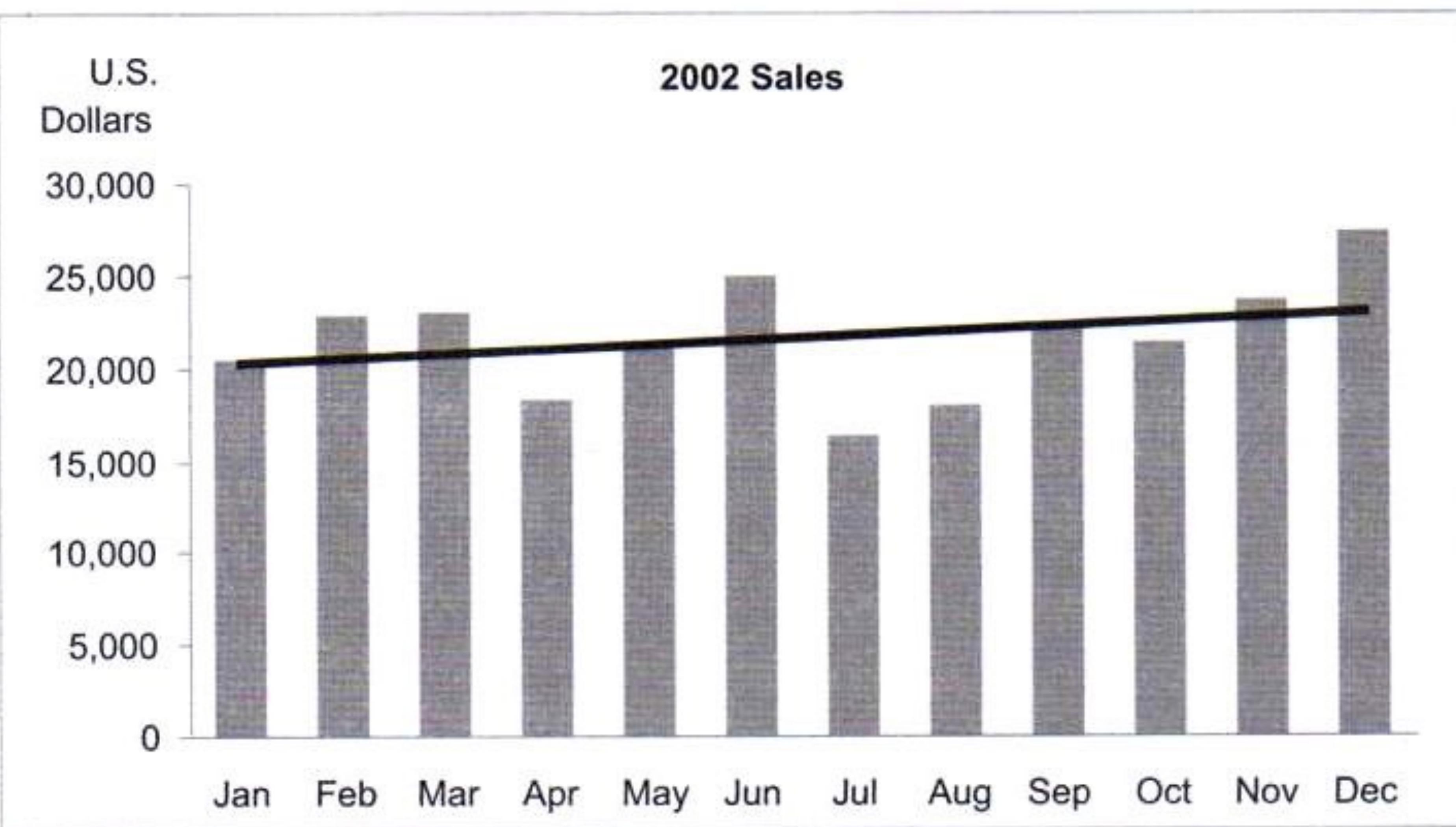


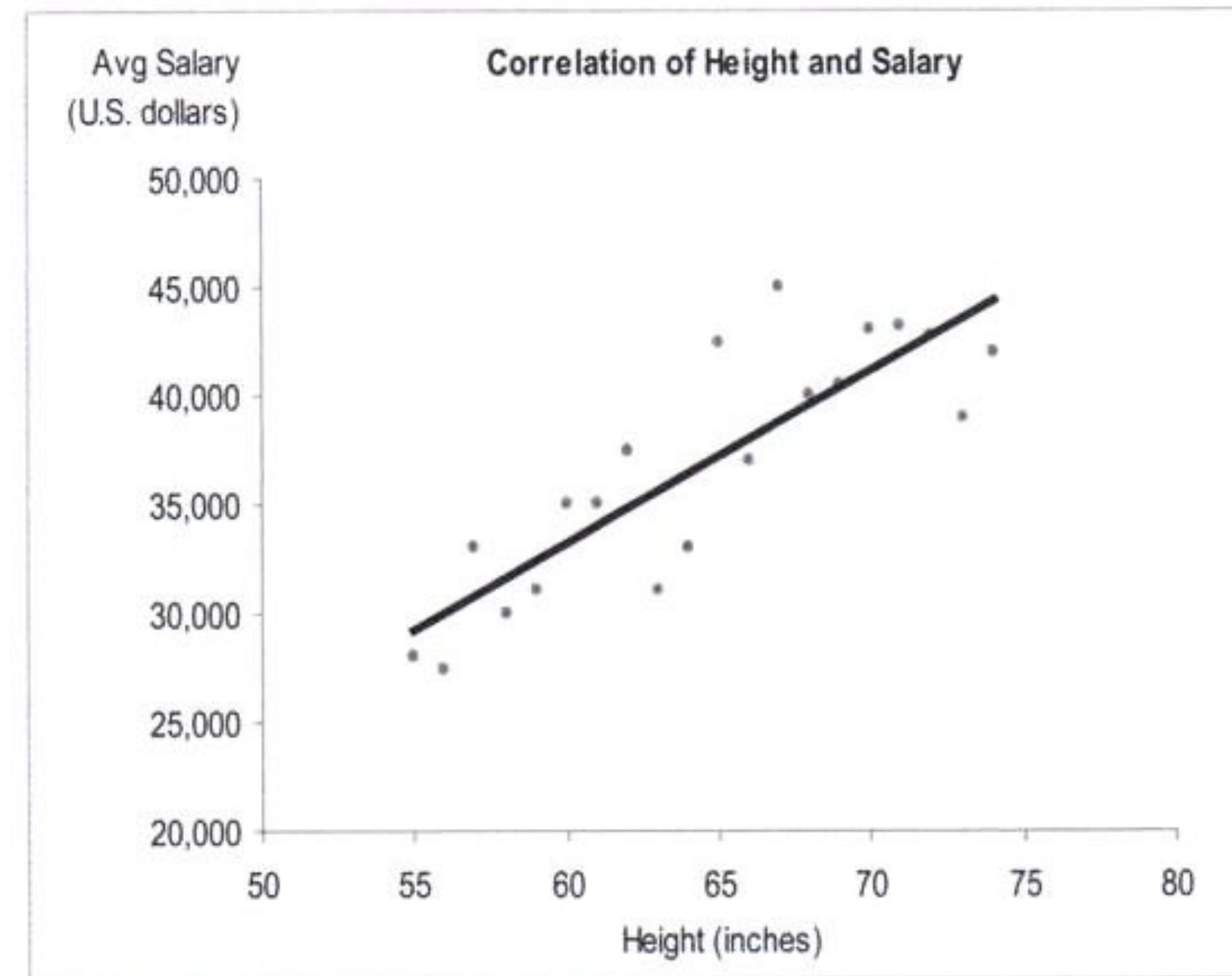


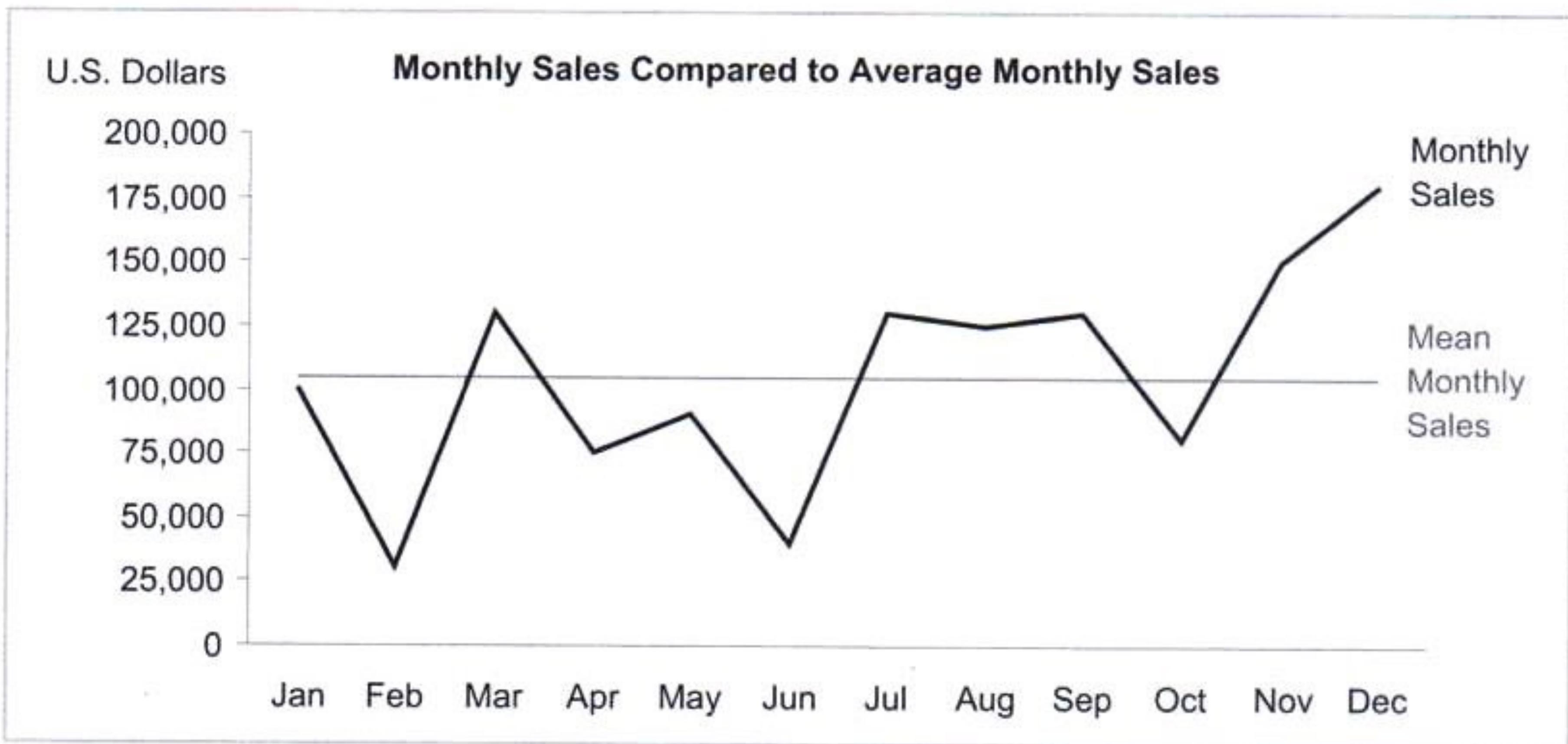


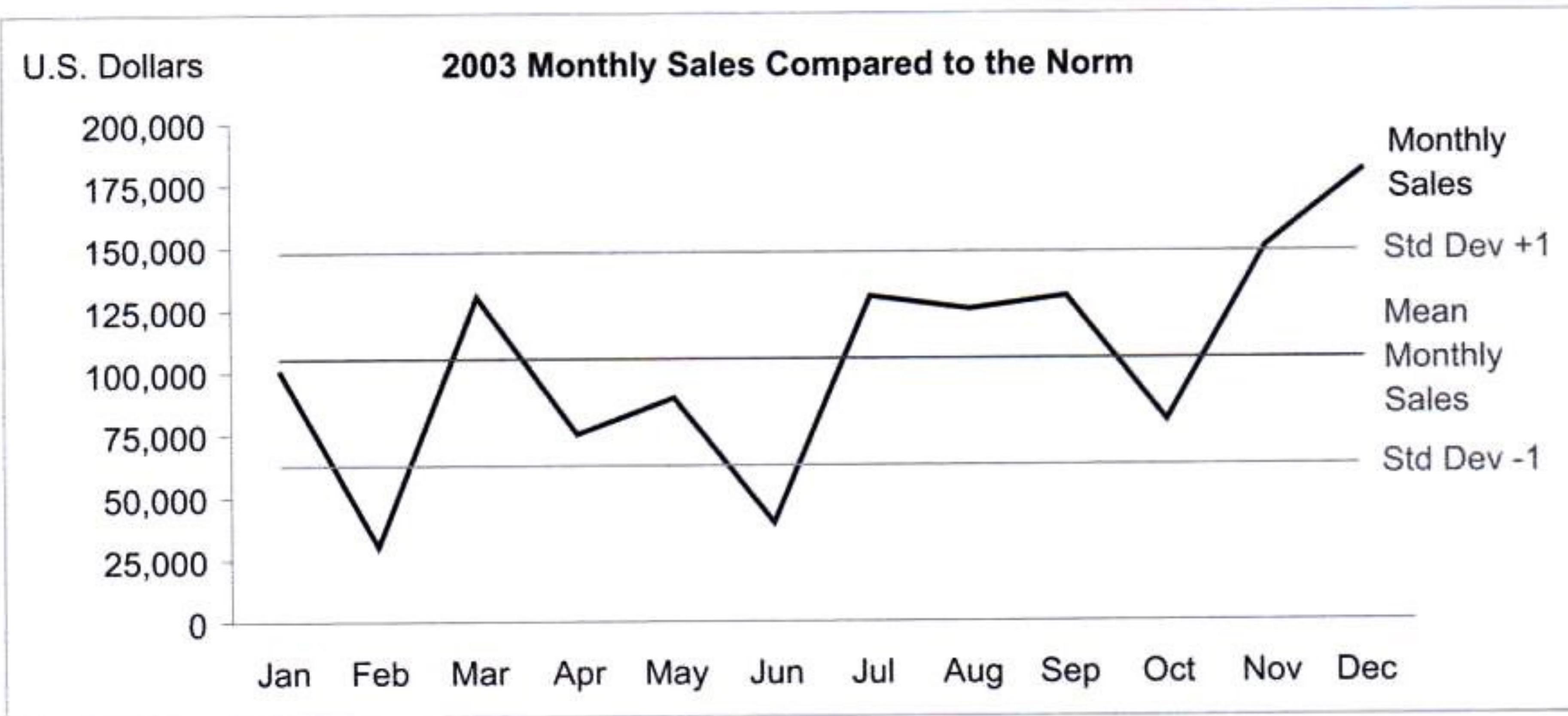






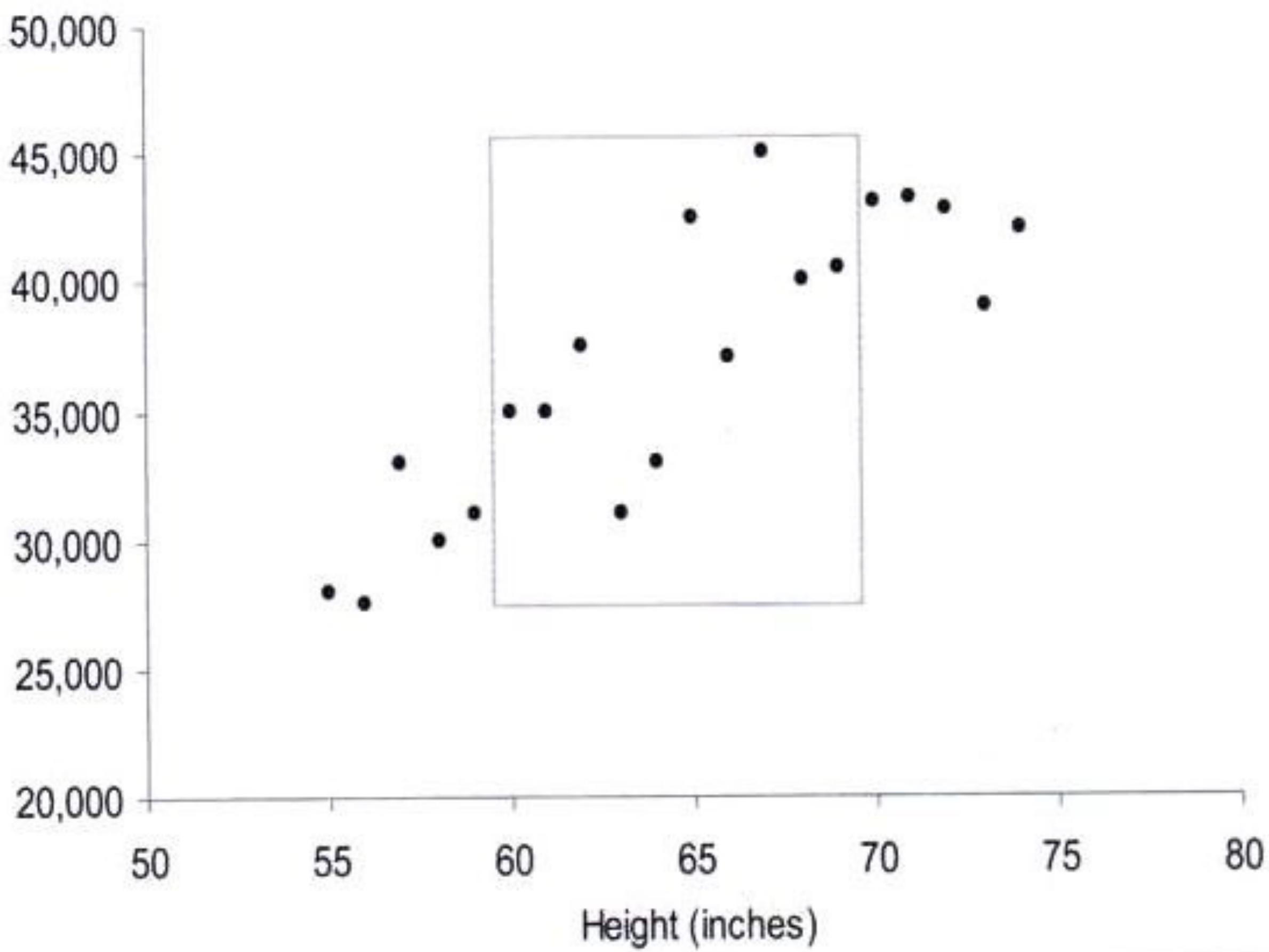




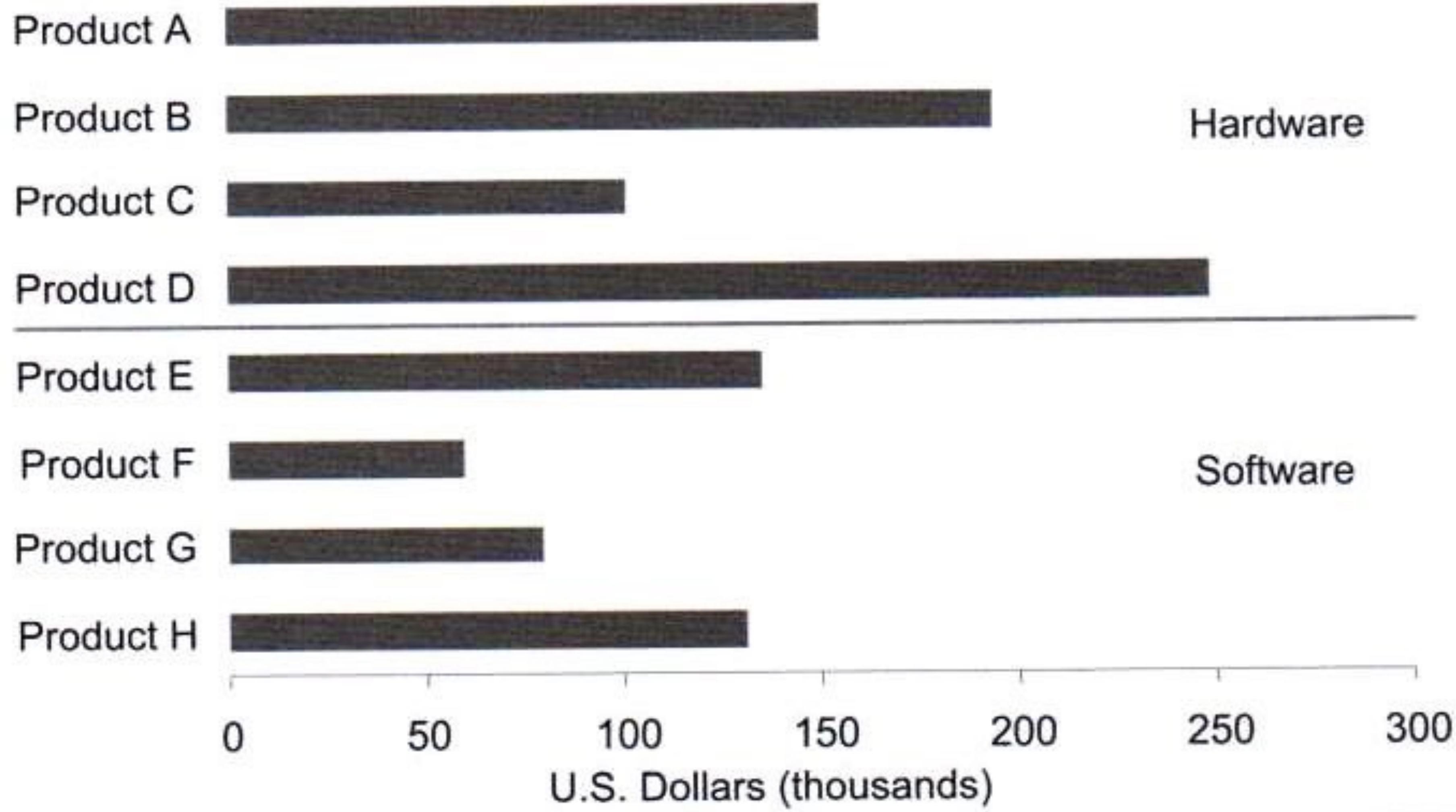


Avg Salary  
(US dollars)

**Correlation of Male Heights and Salaries**  
(Gray rectangle displays 1 standard deviation from the mean)

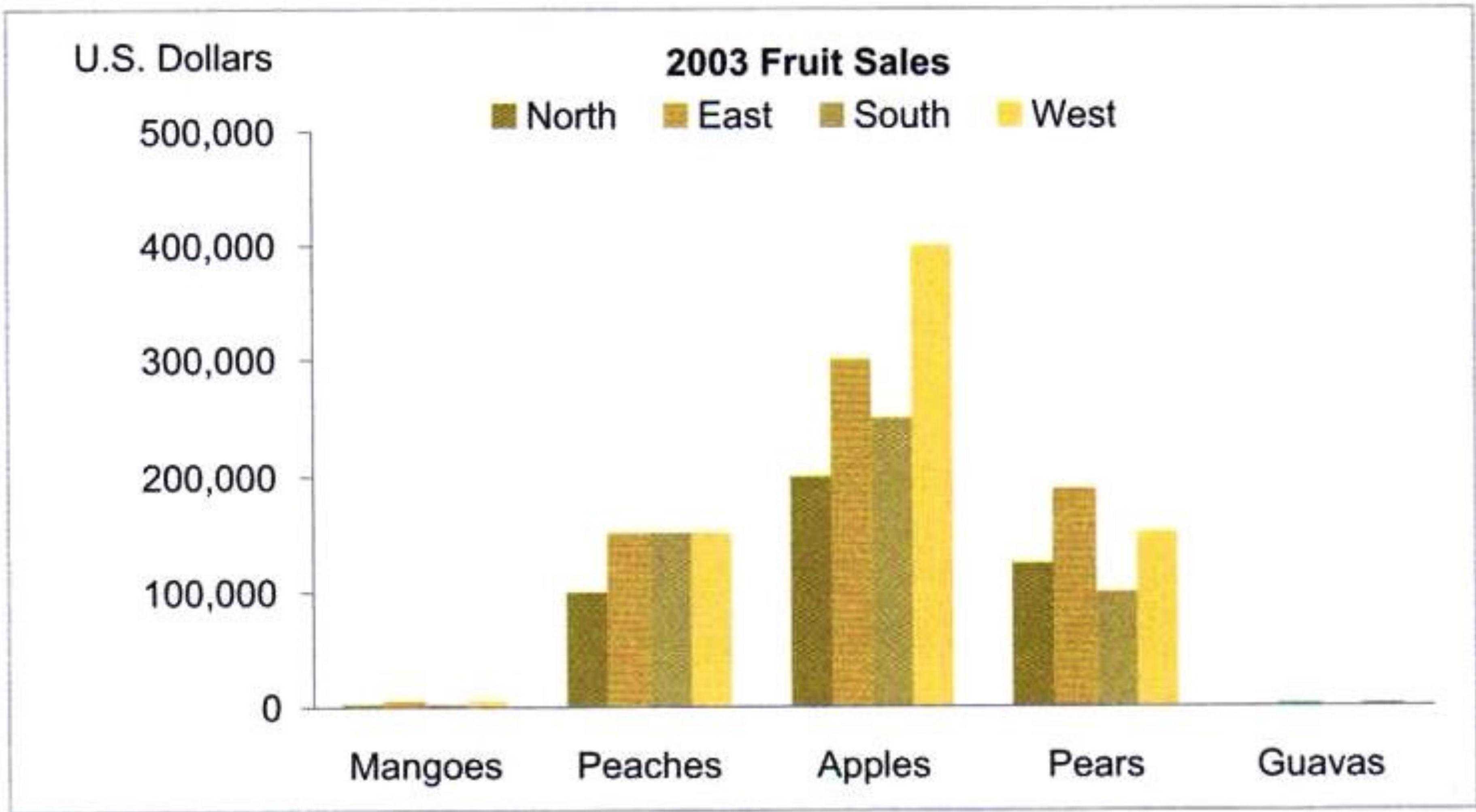


## 2005 Sales





Útil para marcar ocorrência de eventos

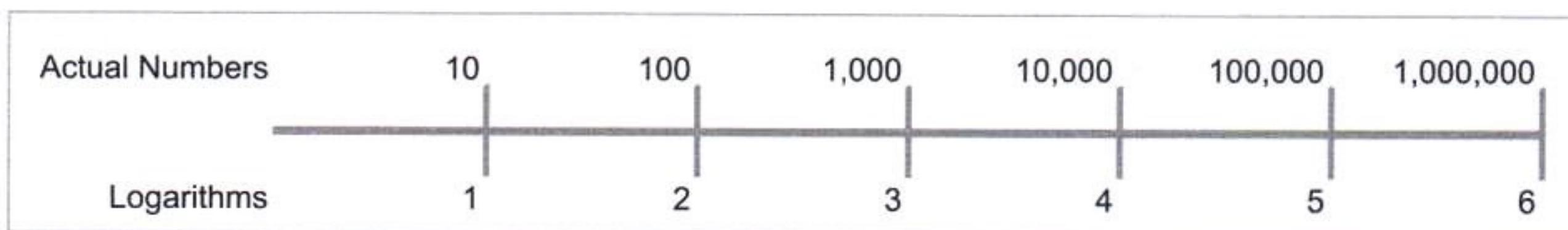


Qual o problema com esta representação?

# ESCALAS

---

- Escala comum
- Escala logarítmica: logaritmo é a potência a qual o número, chamado base, precisa ser elevado para ser igual ao número
- Ex.:  $\log_{10}(1.000.000) = 6$



**Log 10 Scale**  
**(U.S. \$)**

### 2003 Fruit Sales

North    East    South    West

1,000,000

100,000

10,000

1,000

100

10

1

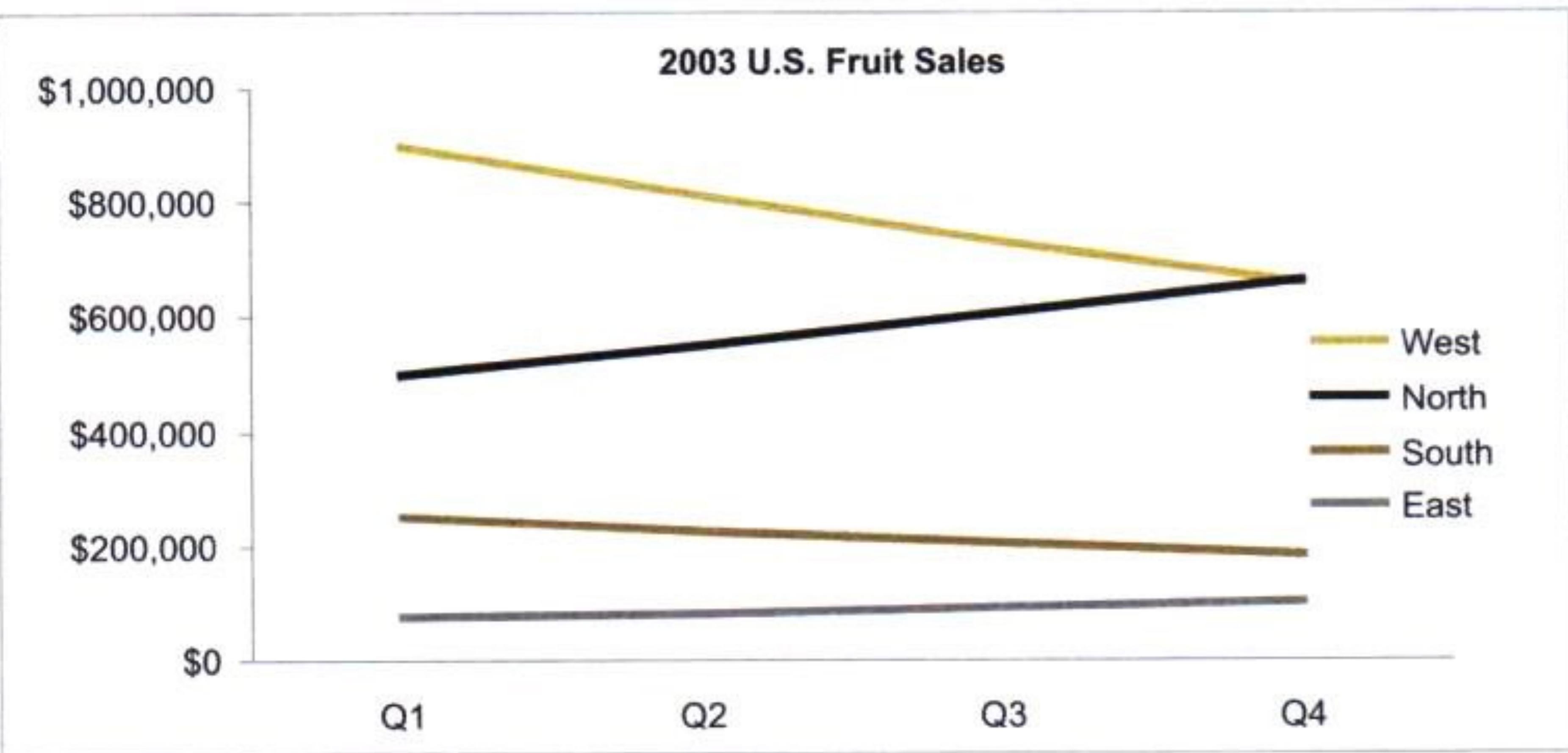
Mangoes

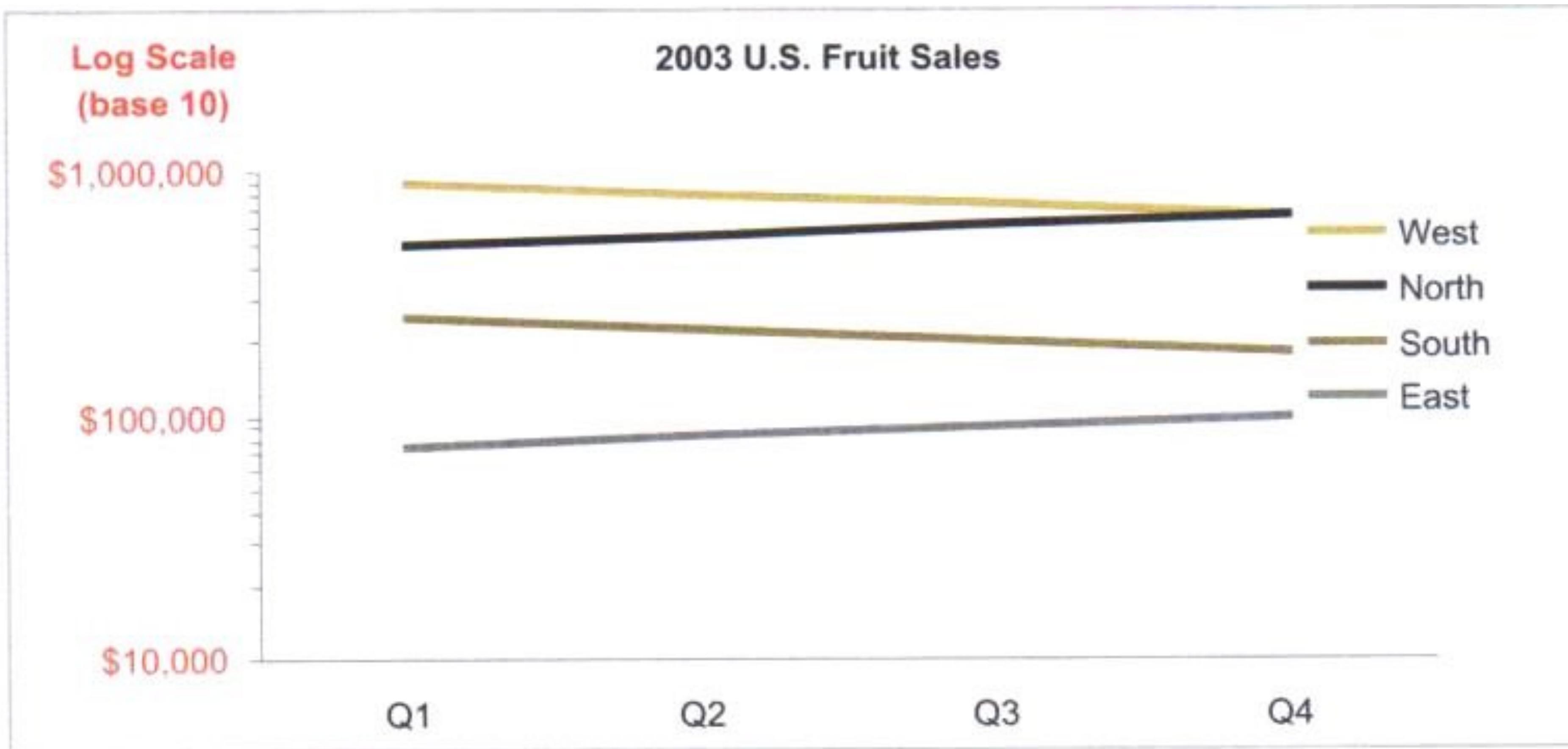
Peaches

Apples

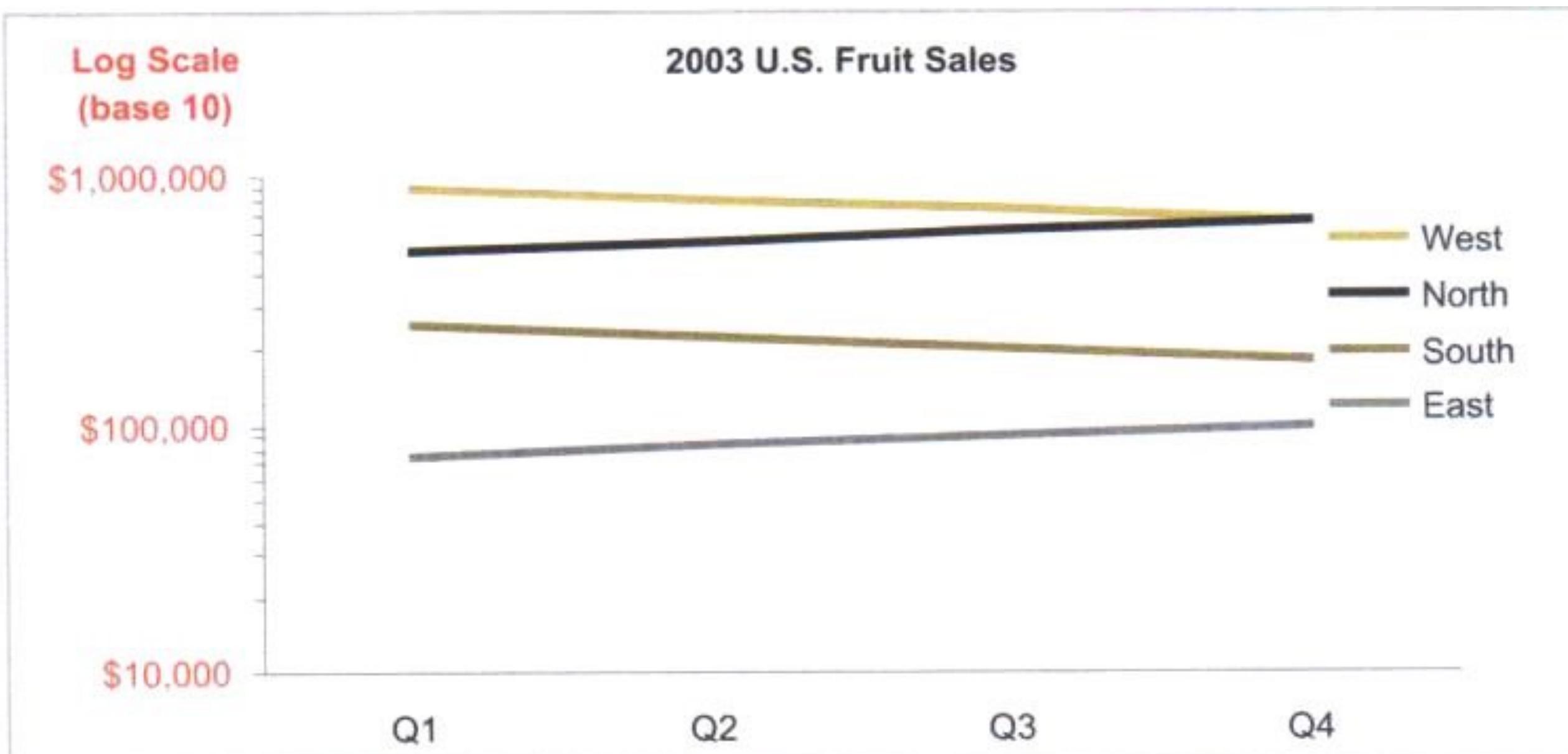
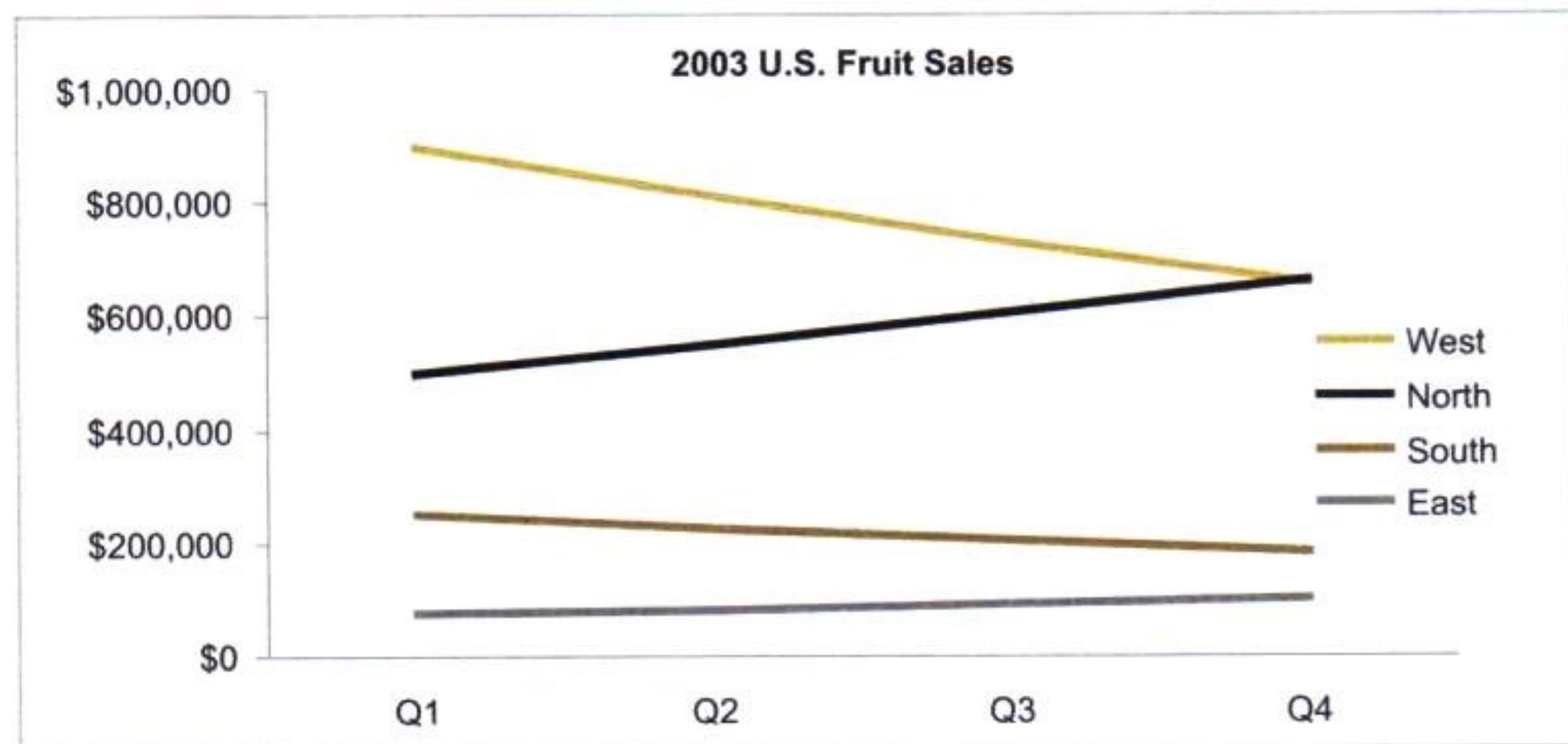
Pears

Guavas





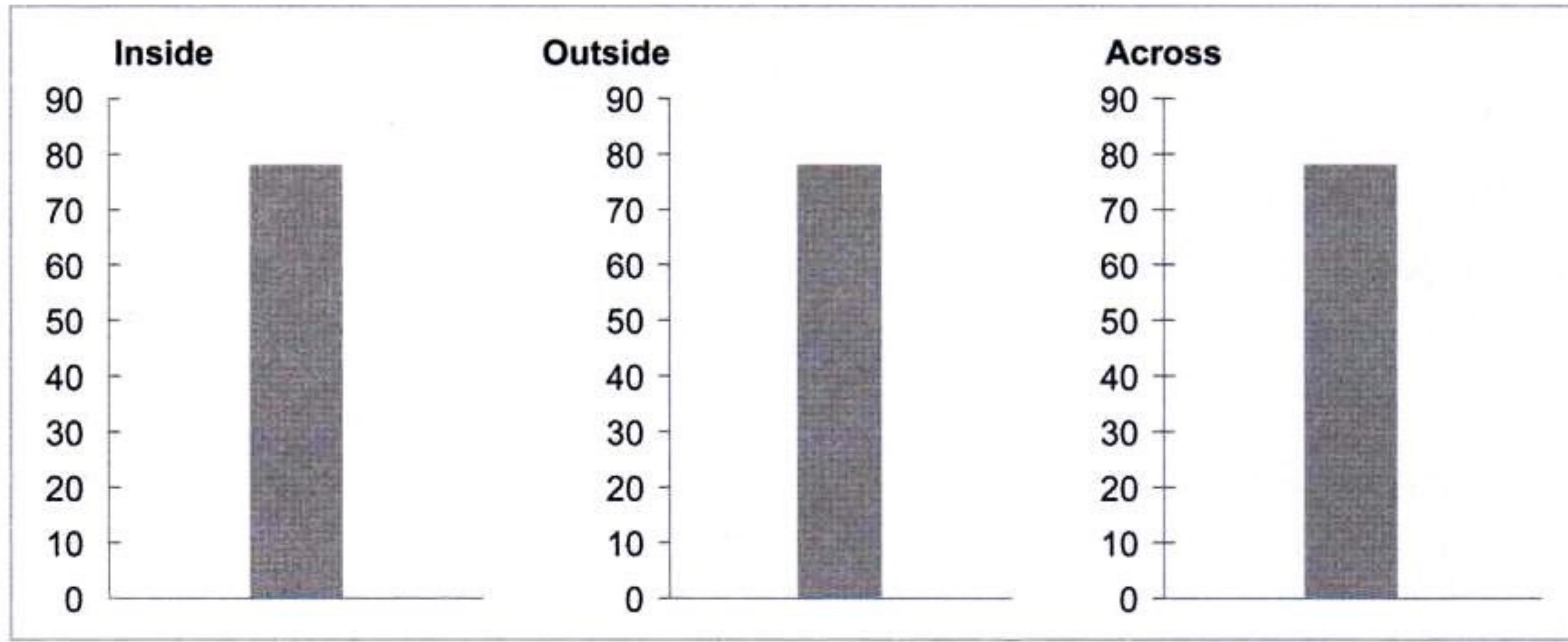
Linhos paralelas indicam mesma taxa de (de)crescimento



# MARCAÇÕES DOS EIXOS

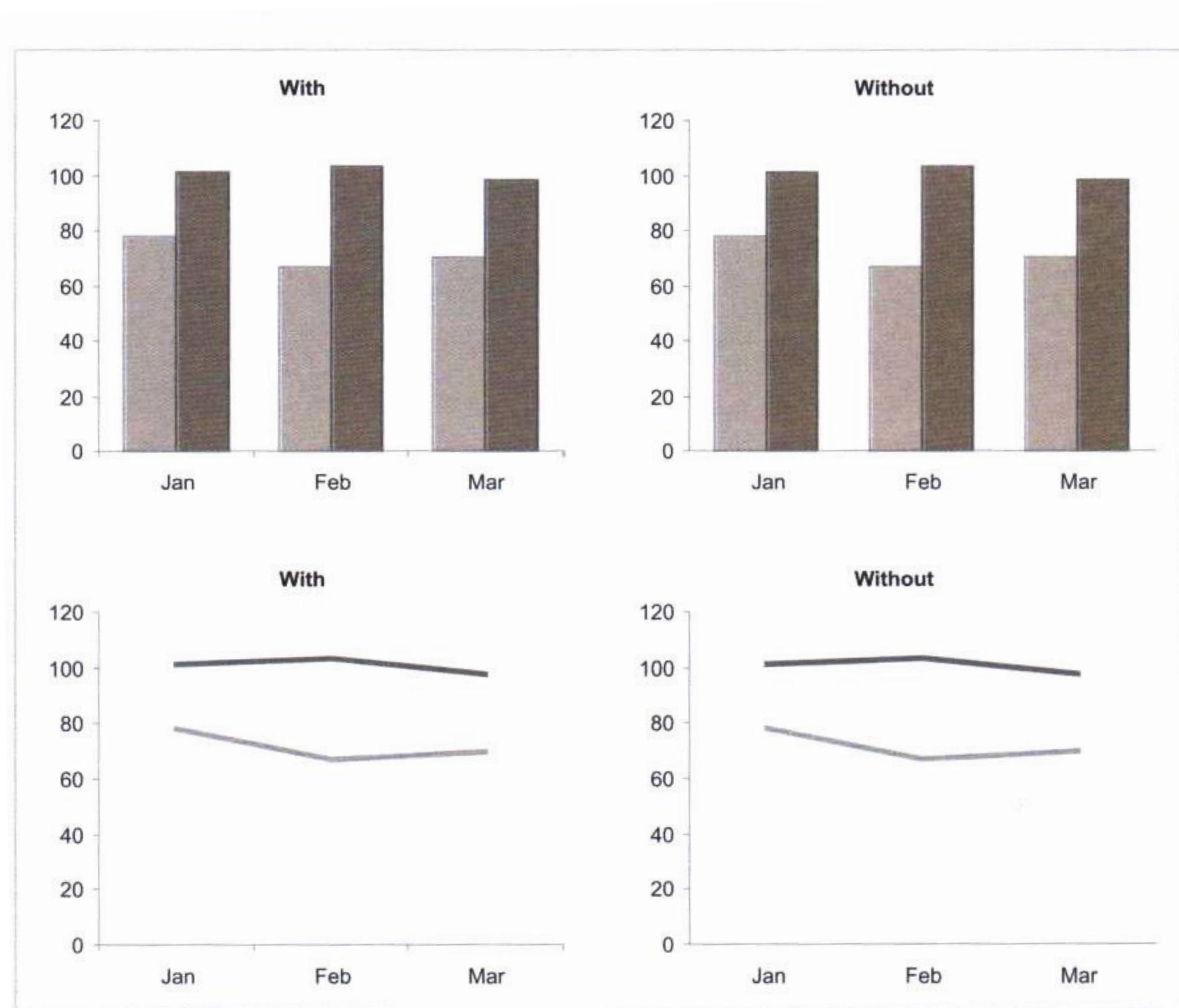
---

- Quão visíveis precisam ser?
- Onde elas deveriam aparecer?
- Quando elas podem ser eliminadas?
- Quantas usar?
- Quais valores deveriam ser exibidos?

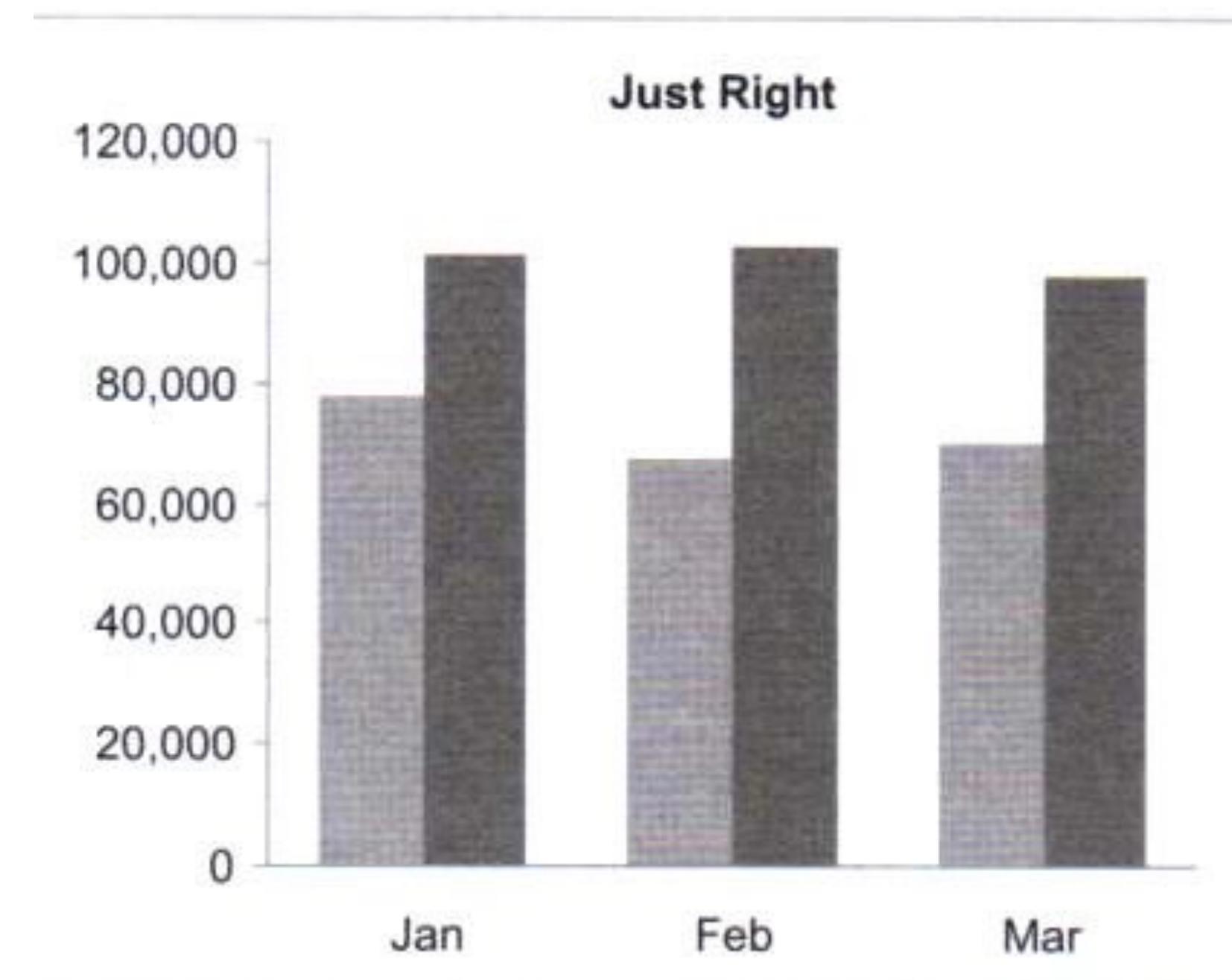
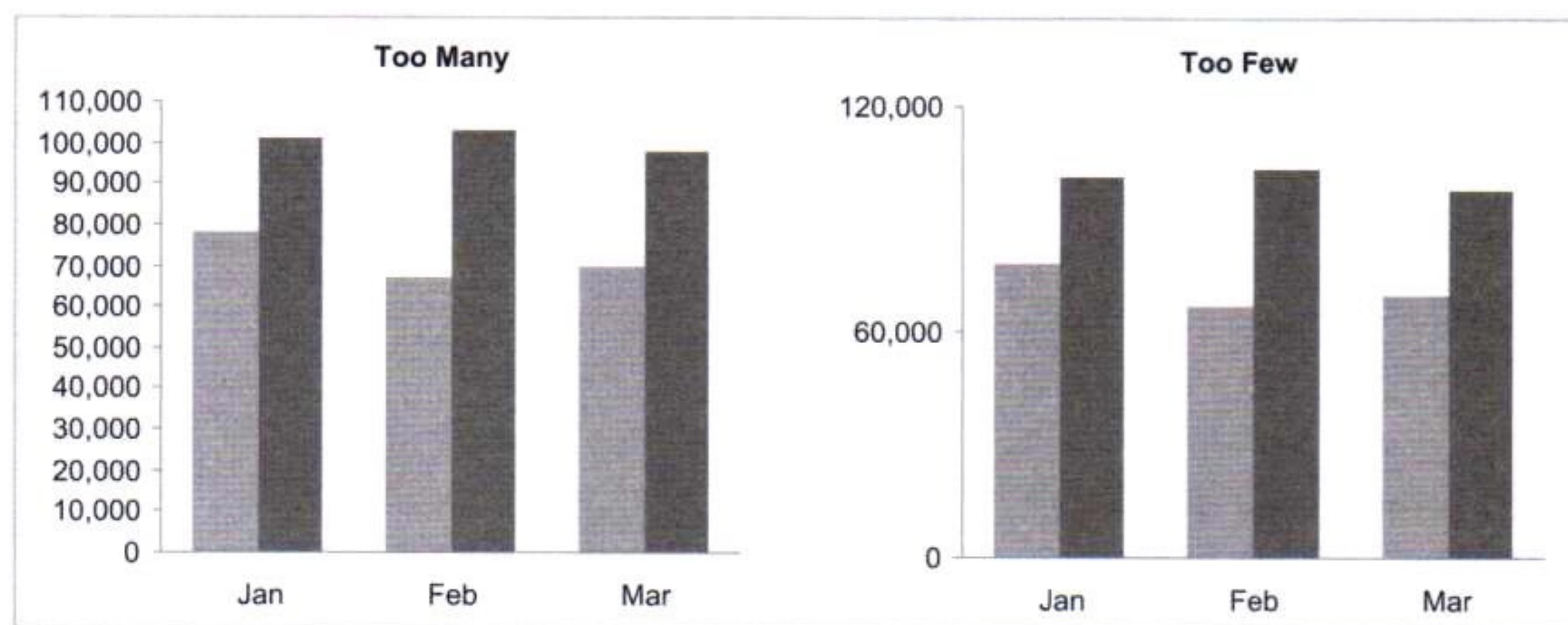


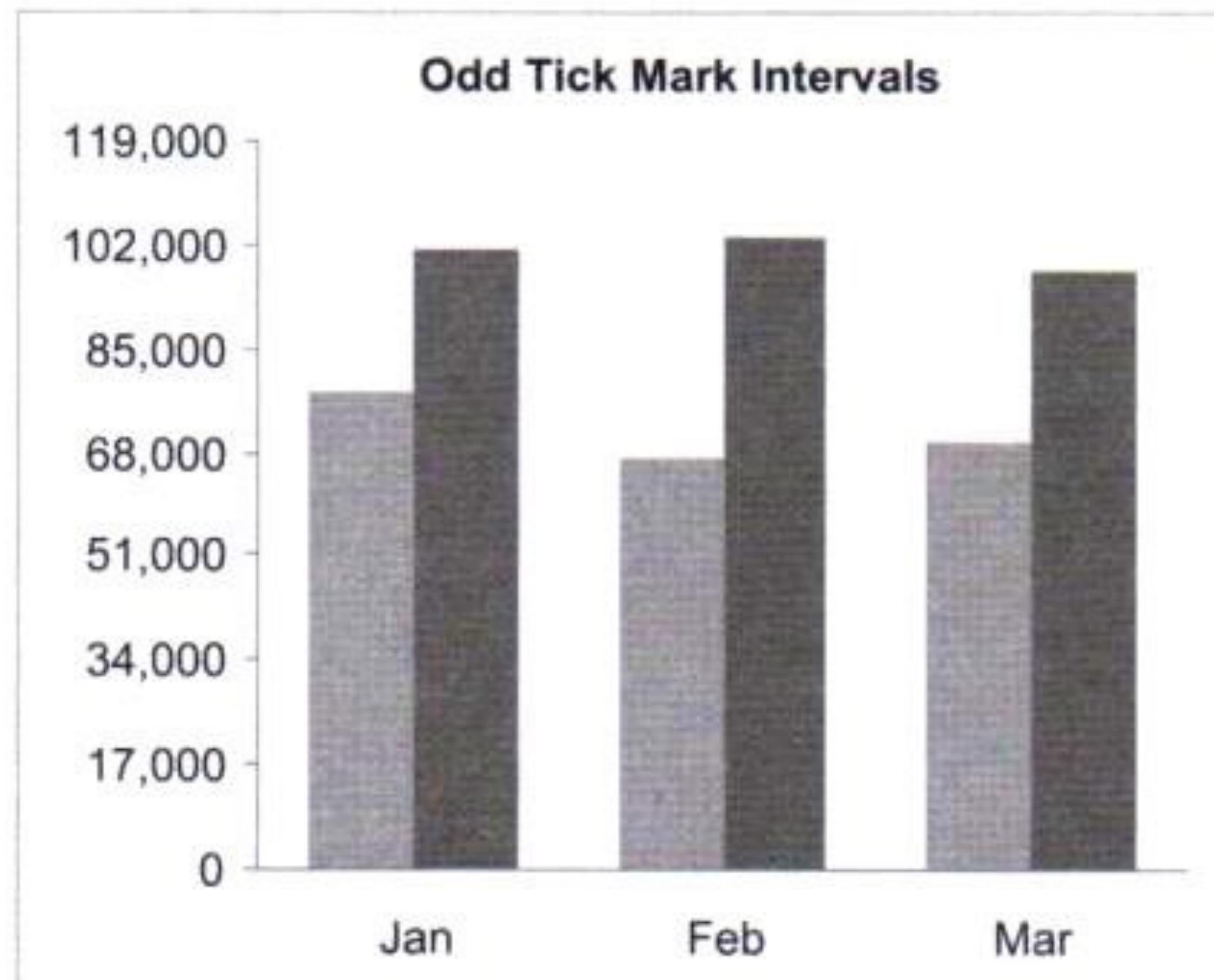
As **marcações** das legendas devem ser discretas quando comparadas aos objetos que representam os dados

Devem ser posicionadas preferencialmente na área externa à área dos dados



Podem ser **eliminadas** quando representarem **dados categóricos**



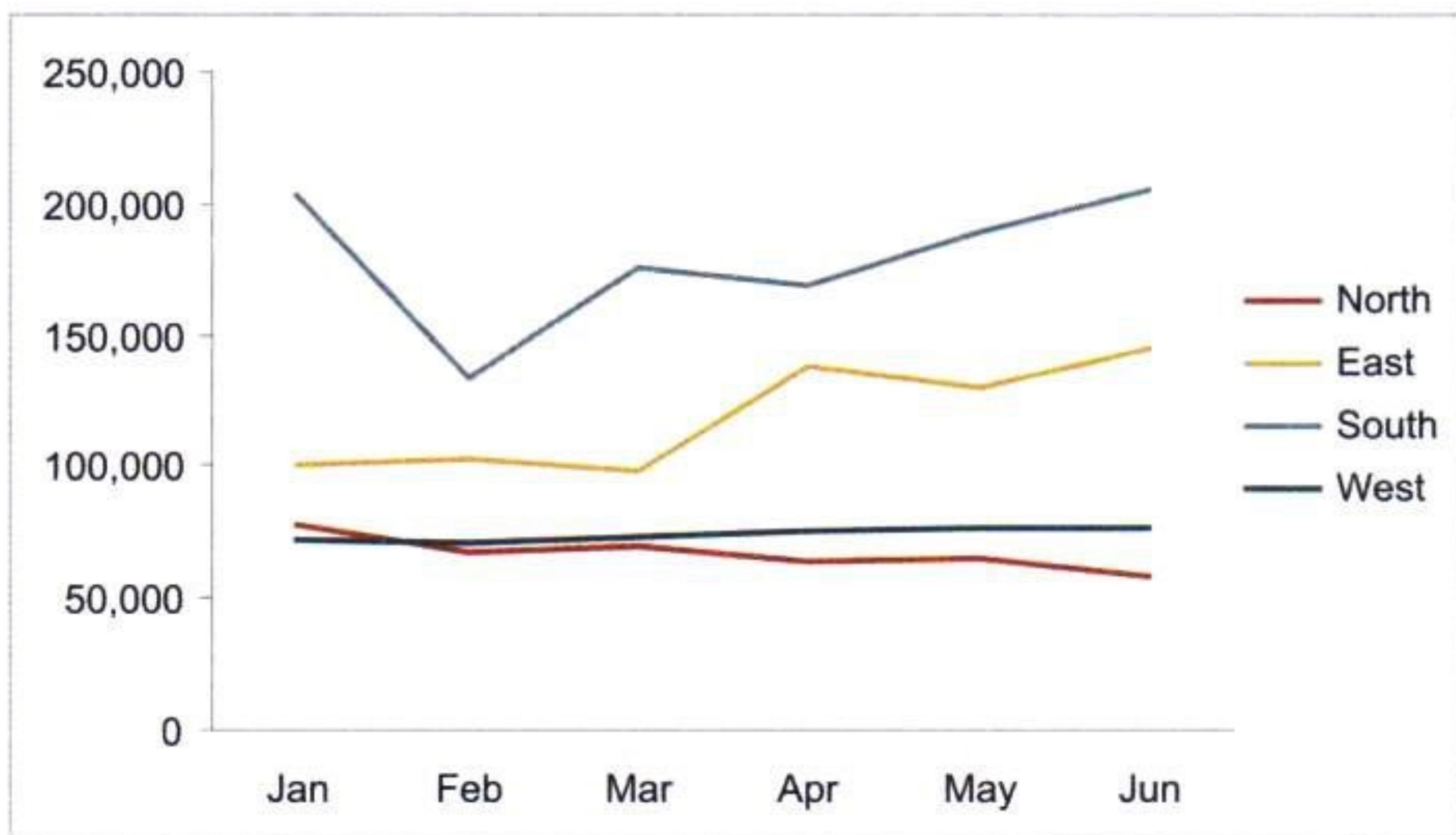


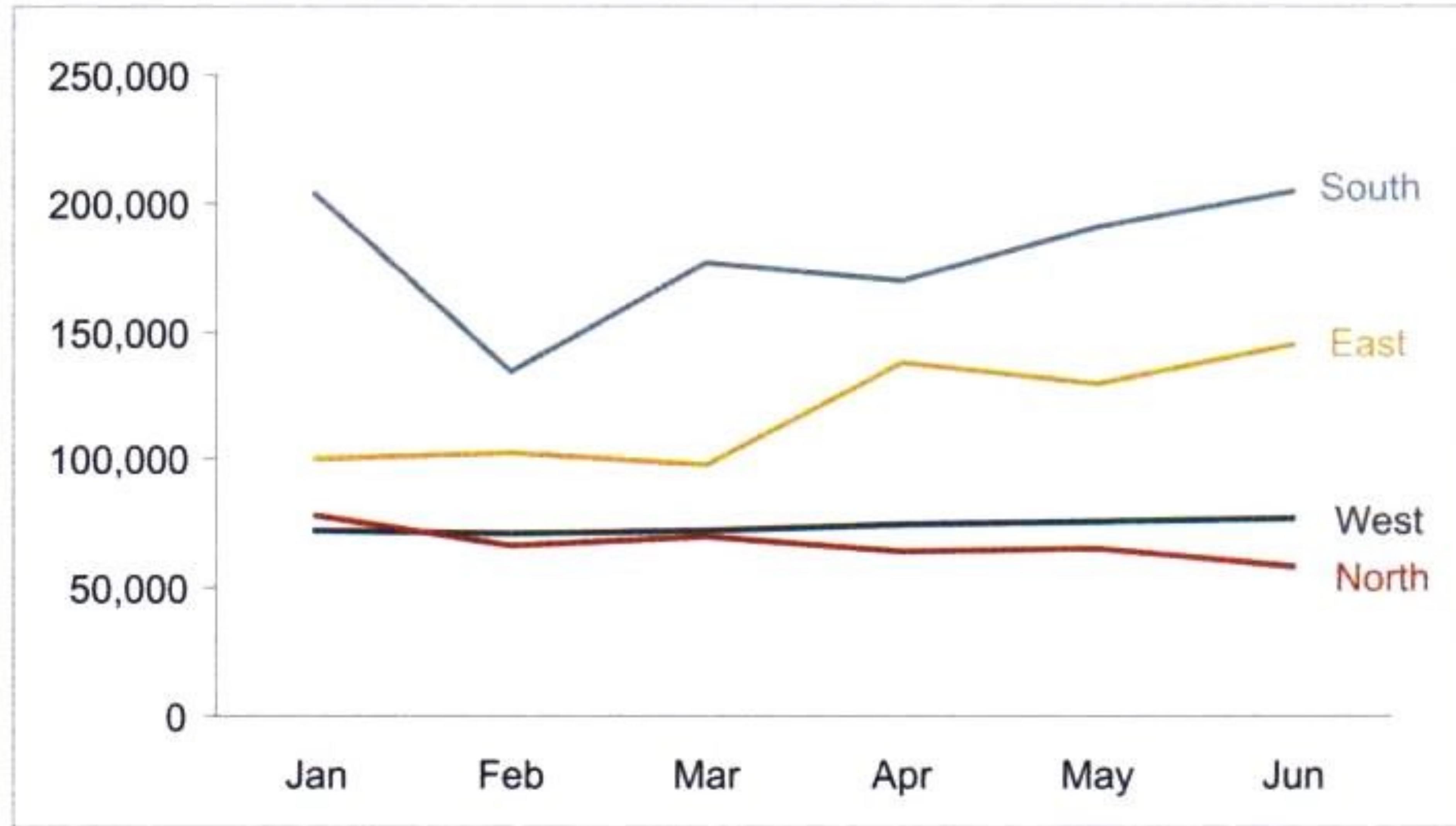
Use valores redondos e pares

# LEGENDAS

---

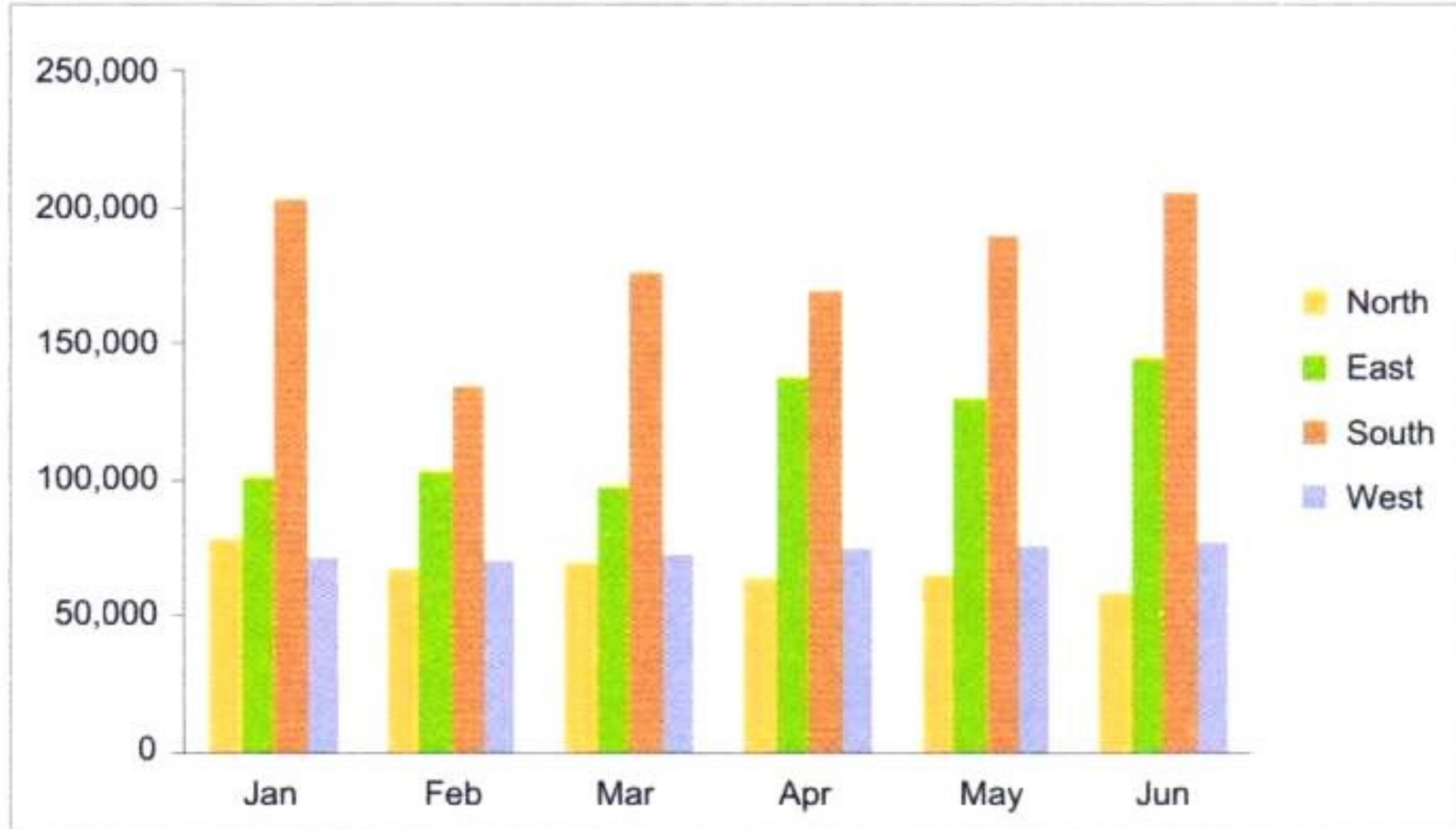
- Quando podem ser eliminadas?
- Onde elas devem ser posicionadas?
- Quão visíveis devem ser?
- Devem ter bordas?
- Onde posicionar os rótulos das legendas?





Legendas podem ser substituídas pela rotulação de linhas

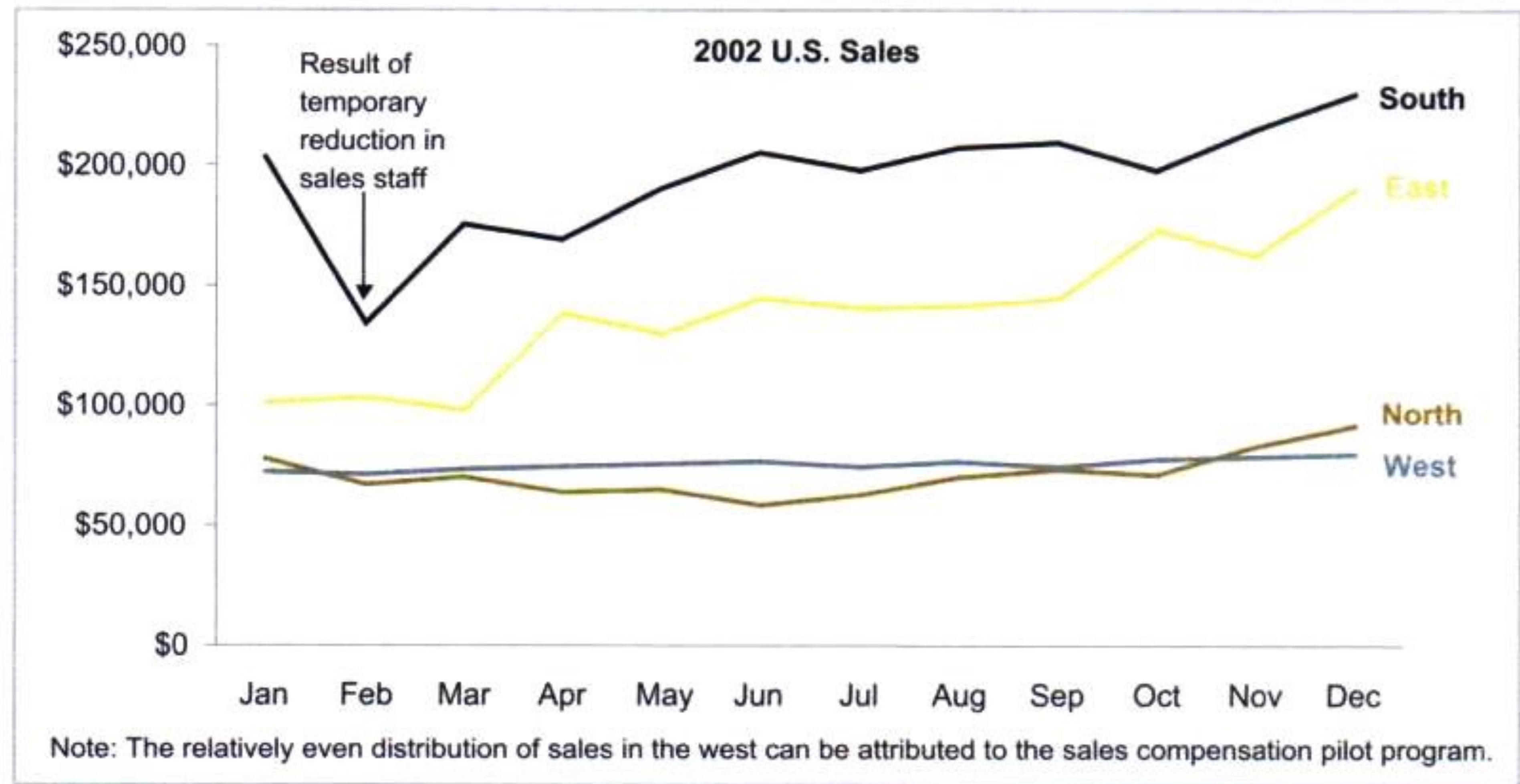
- Legendas devem ser posicionadas preferencialmente fora da área de dados
- Quanto mais próximas estiverem dos dados melhor
- Elas não devem chamar mais atenção que os dados
- Legendas não devem ter bordas, uma vez que bordas chamam atenção para os objetos contornados





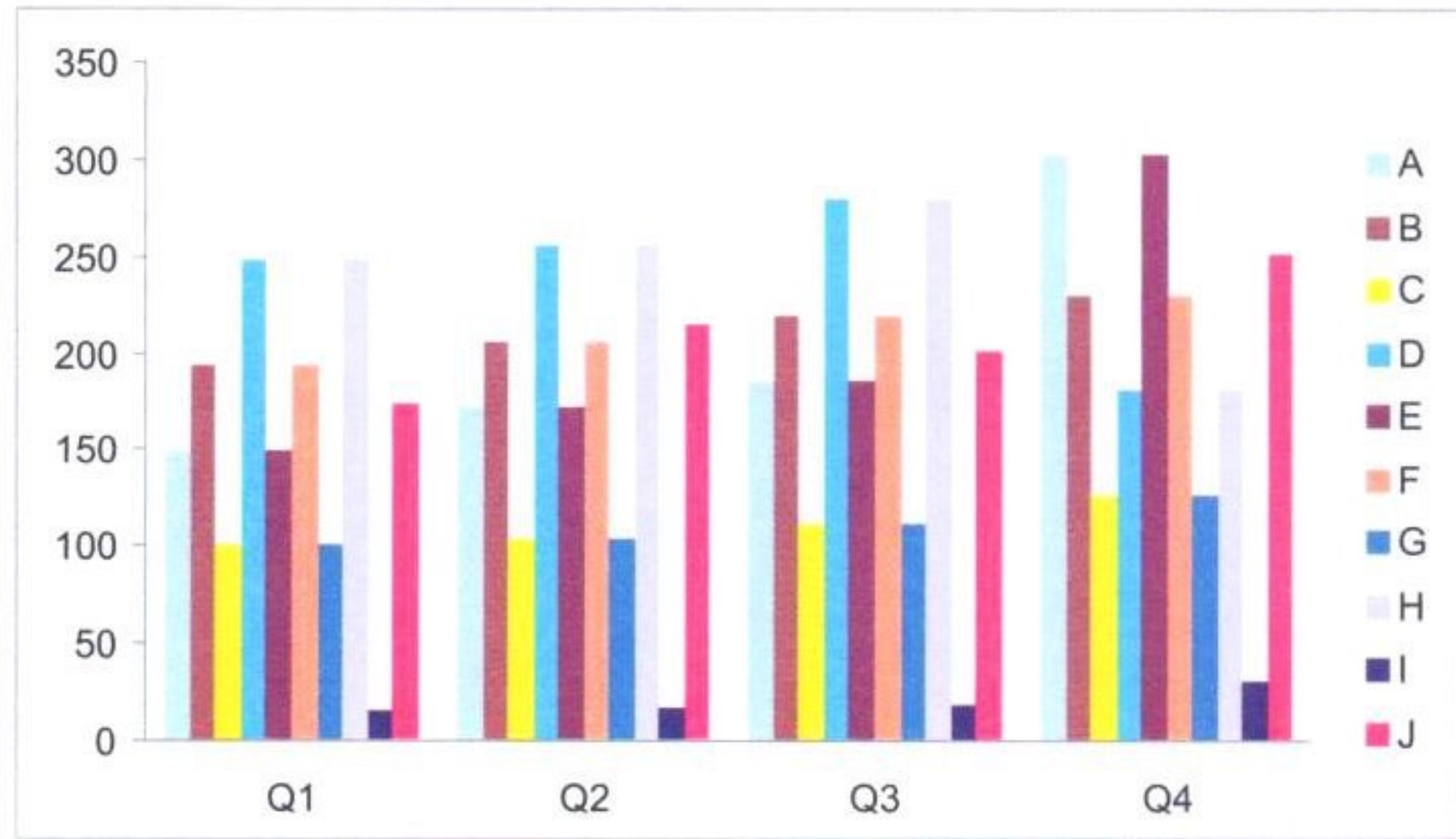


Podem ser posicionadas ao lado do gráfico verticalmente ou mesmo horizontalmente abaixo do título



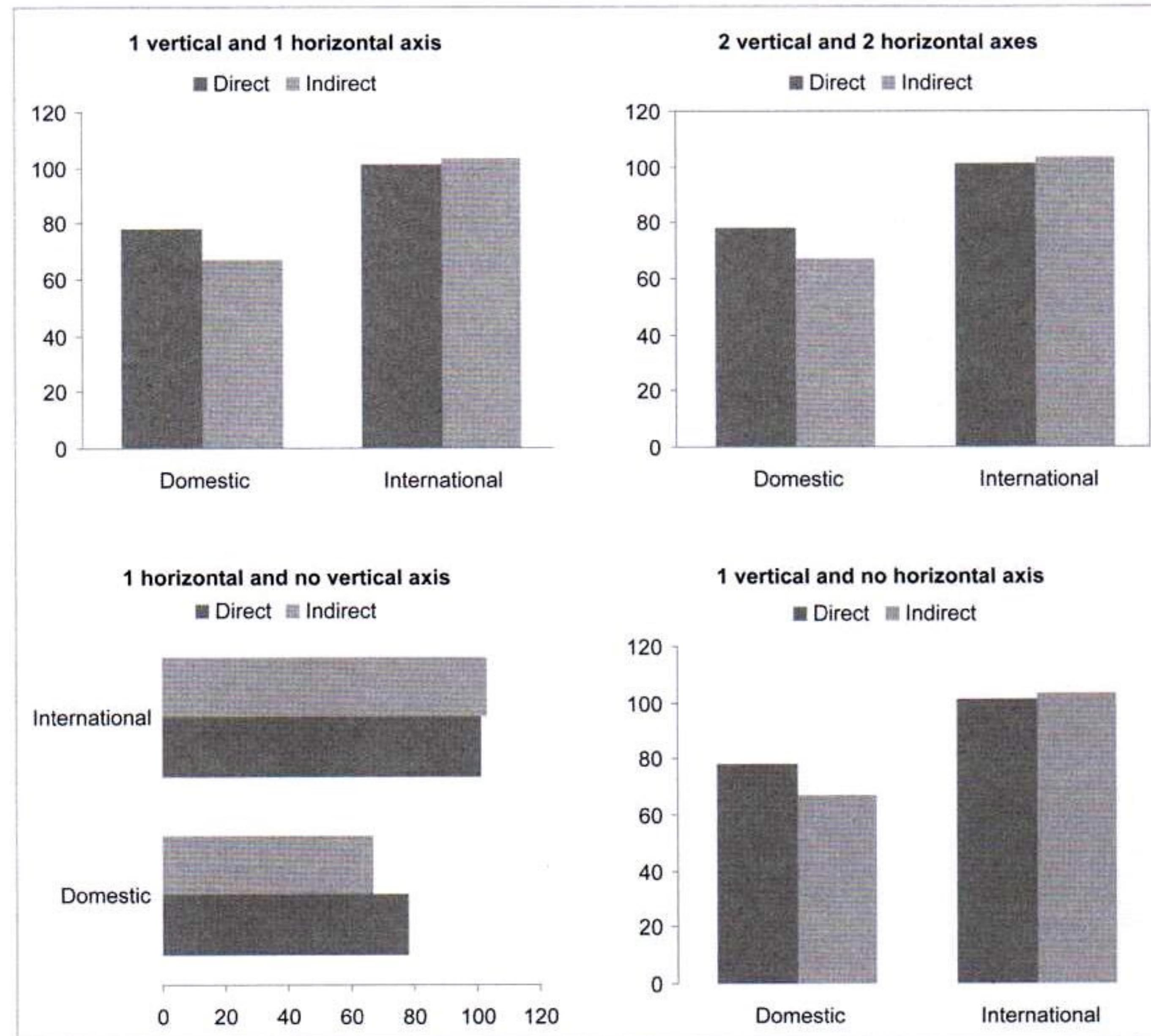
Pode-se colocar **texto adicional** nos gráficos com o objetivo de complementar uma informação

Normalmente, estes textos devem ser posicionados próximos aos objetos que explicam



Devido às nossas limitações de memória de trabalho, o número de **subdivisões categóricas** está **entre 5 e 8**

Isto serve para barras, pontos e linhas



**Borda** é útil  
apenas quando  
é necessário  
separar gráfico  
de texto

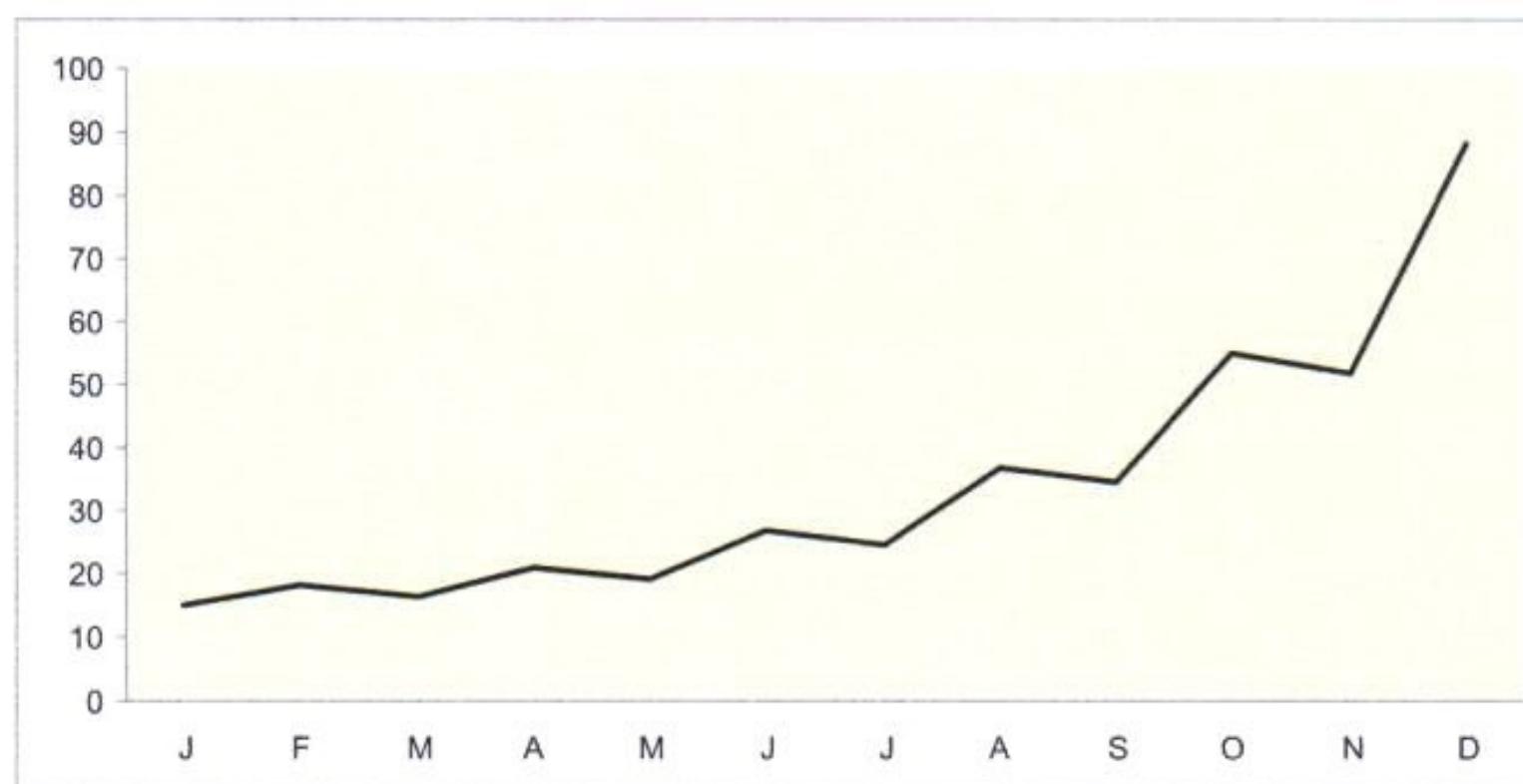


FIGURE 10.65 This graph has an aspect ratio of 1 to 2, or 0.5.

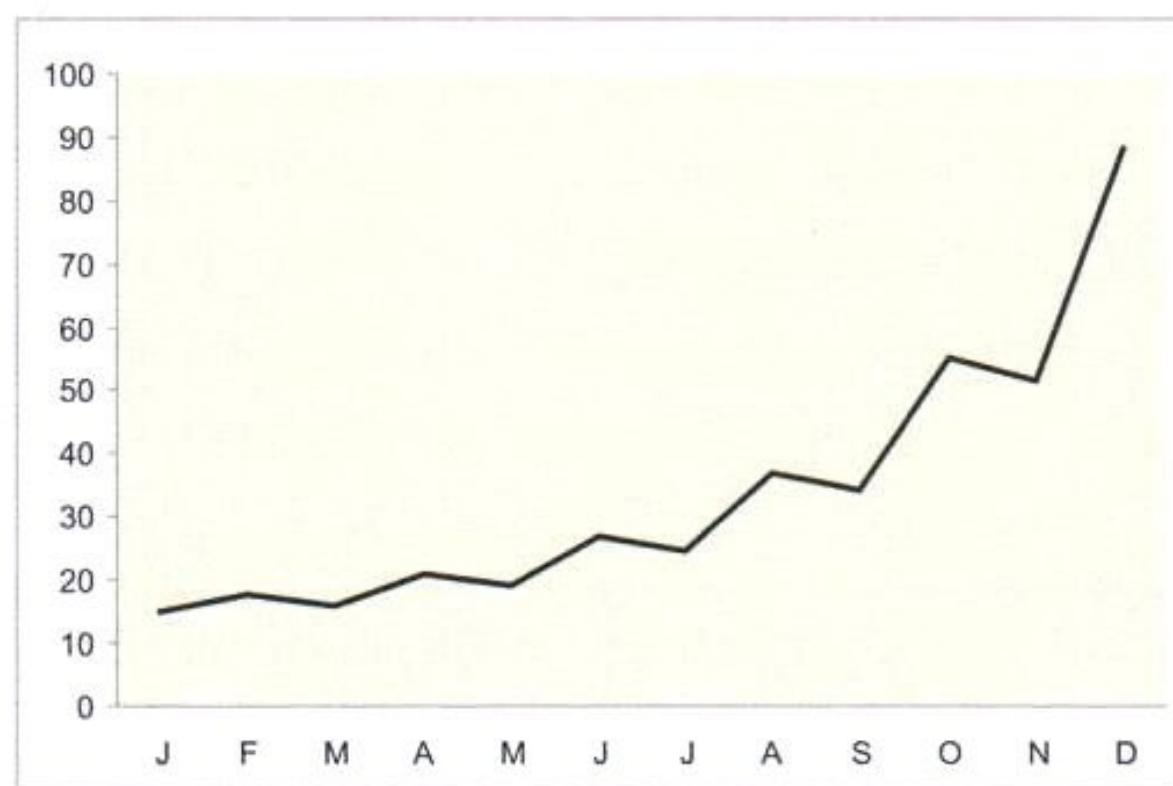


FIGURE 10.66 This graph has an aspect ratio of 1 to 1.5, or 0.67.

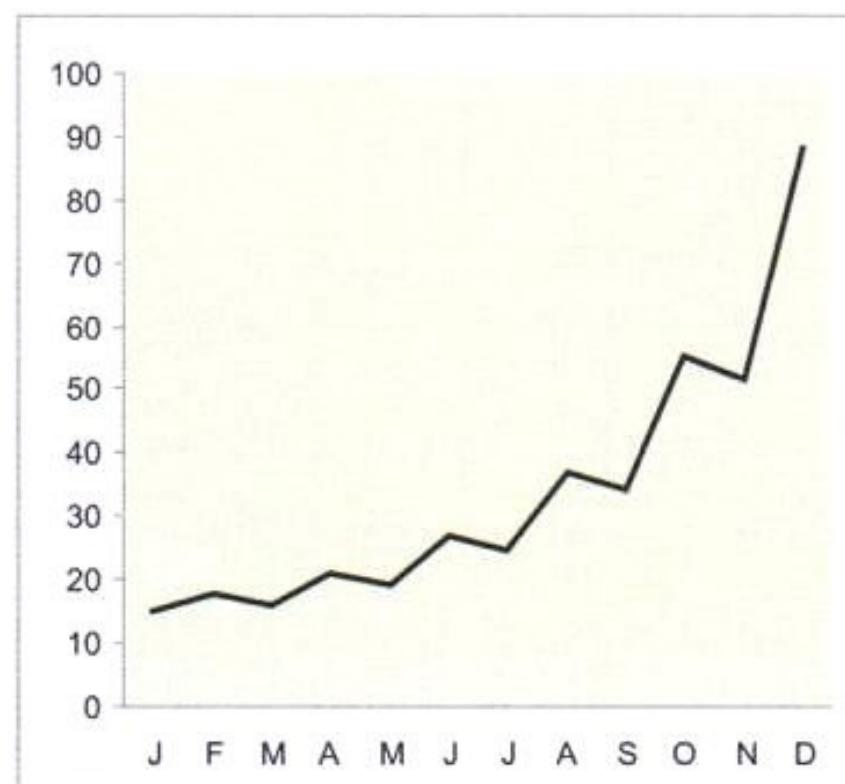


FIGURE 10.67 This graph has an aspect ratio of 1 to 1, or 1.0.

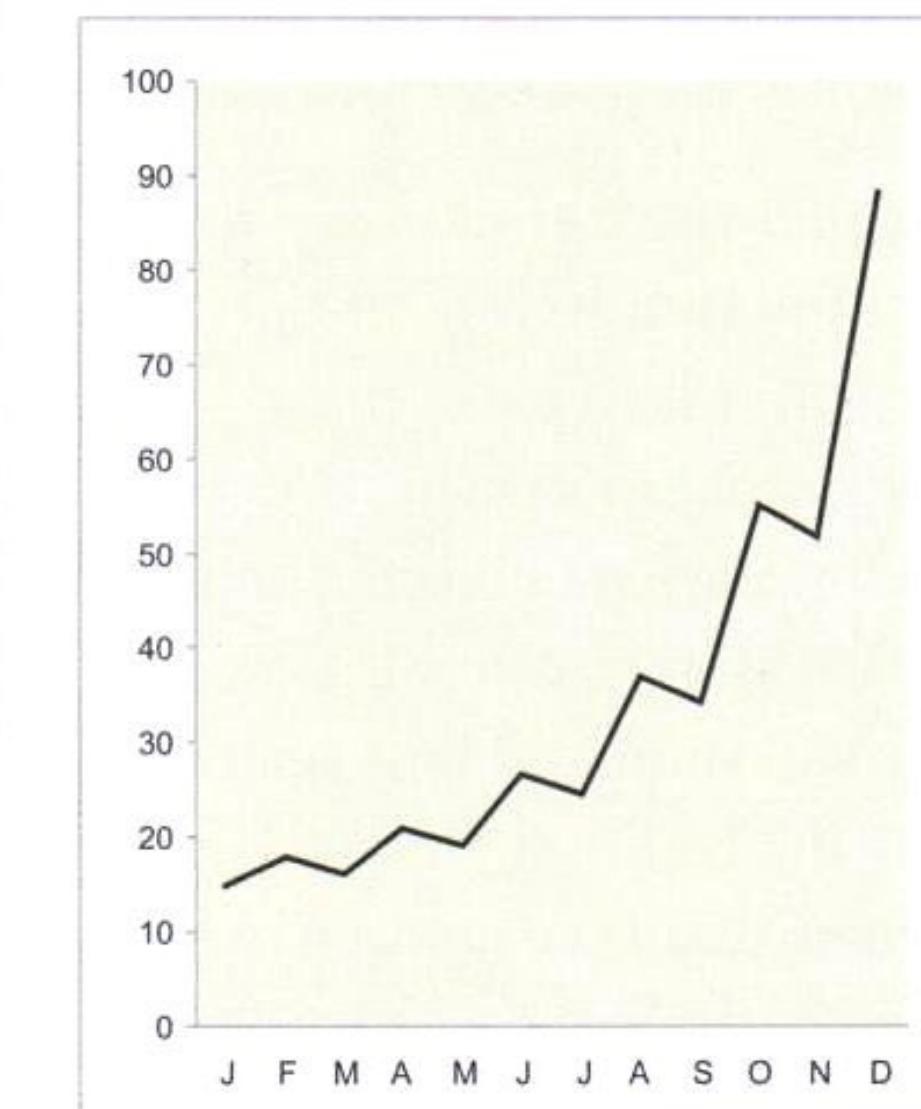


FIGURE 10.68 This graph has an aspect ratio of 1.5 to 1, or 1.5.

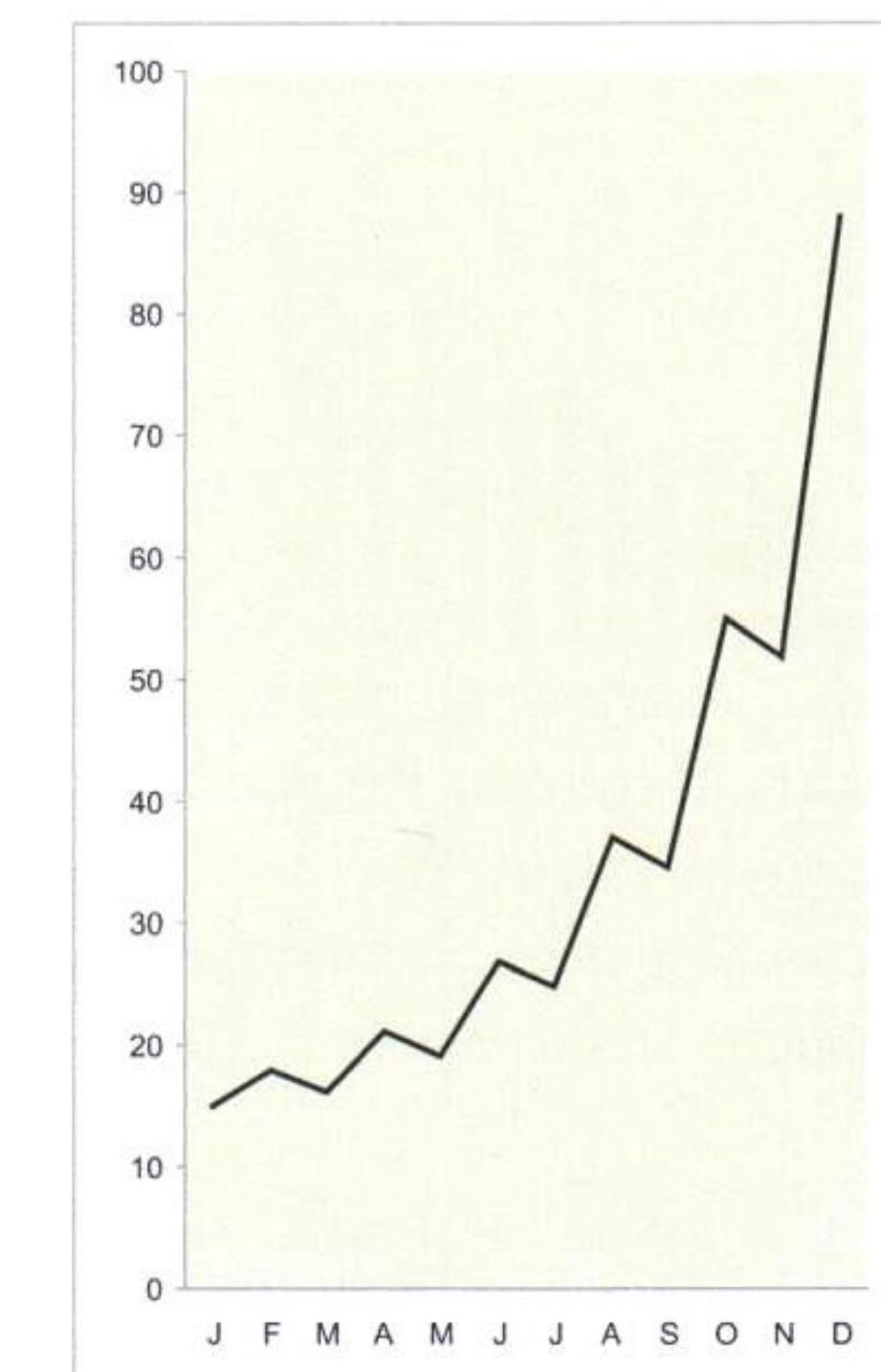
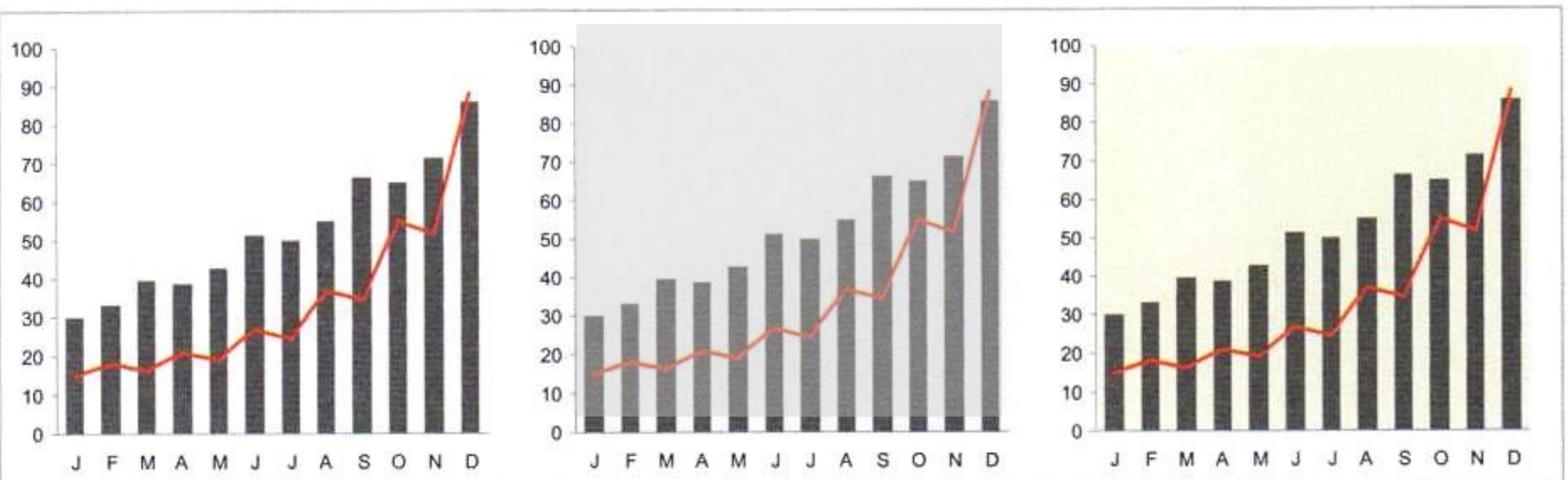


FIGURE 10.69 This graph has an aspect ratio of 2 to 1, or 2.0.

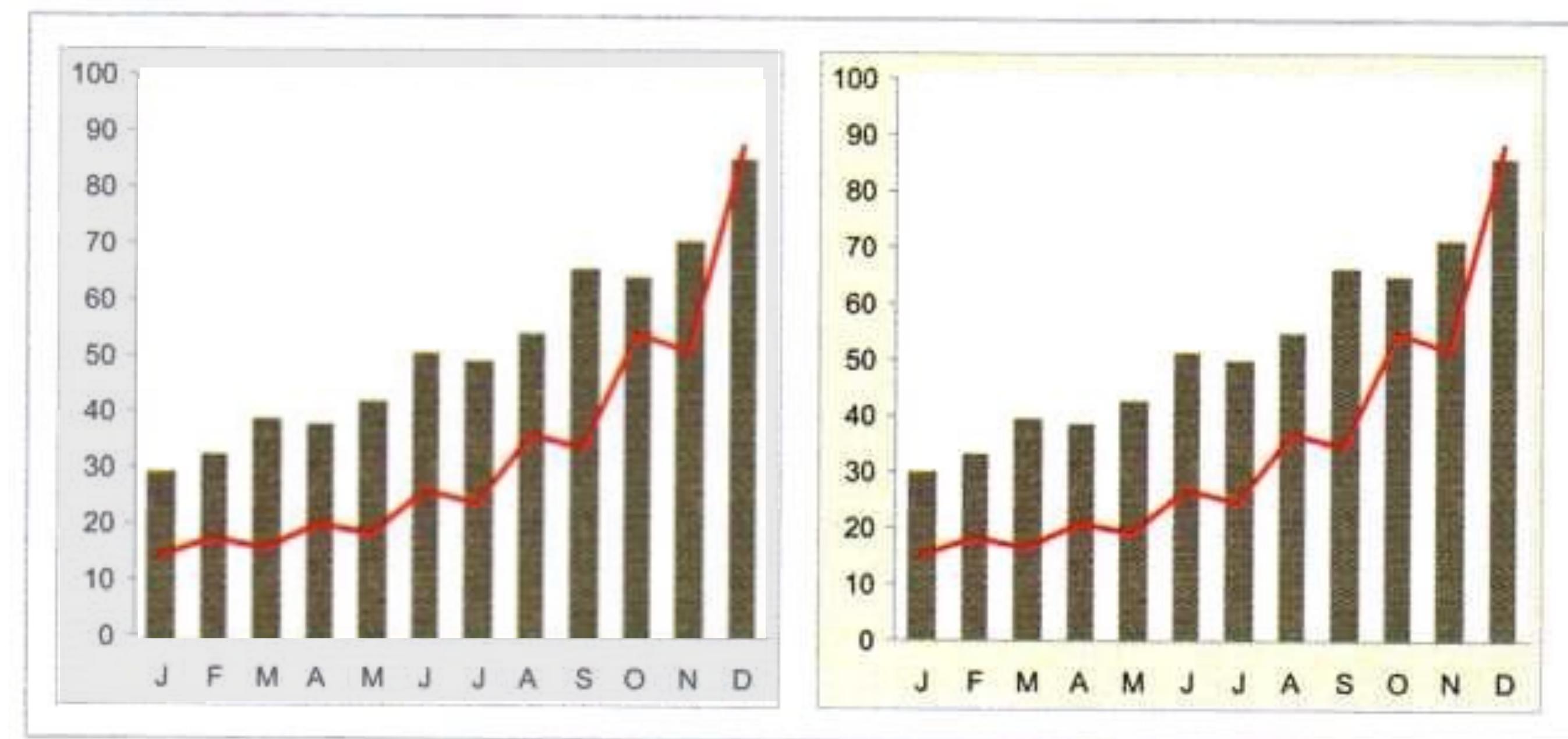
**Nunca manipule a razão do aspecto** para que os gráficos passem uma mensagem errada

**Normalmente os gráficos devem ser mais largos que altos**,  
exceto os *scatter plots* que pedem gráficos com área de dados quadradas



A melhor cor para o **fundo** de um gráfico é o **branco**, podendo também ser substituído por cinza ou amarelo bem suaves

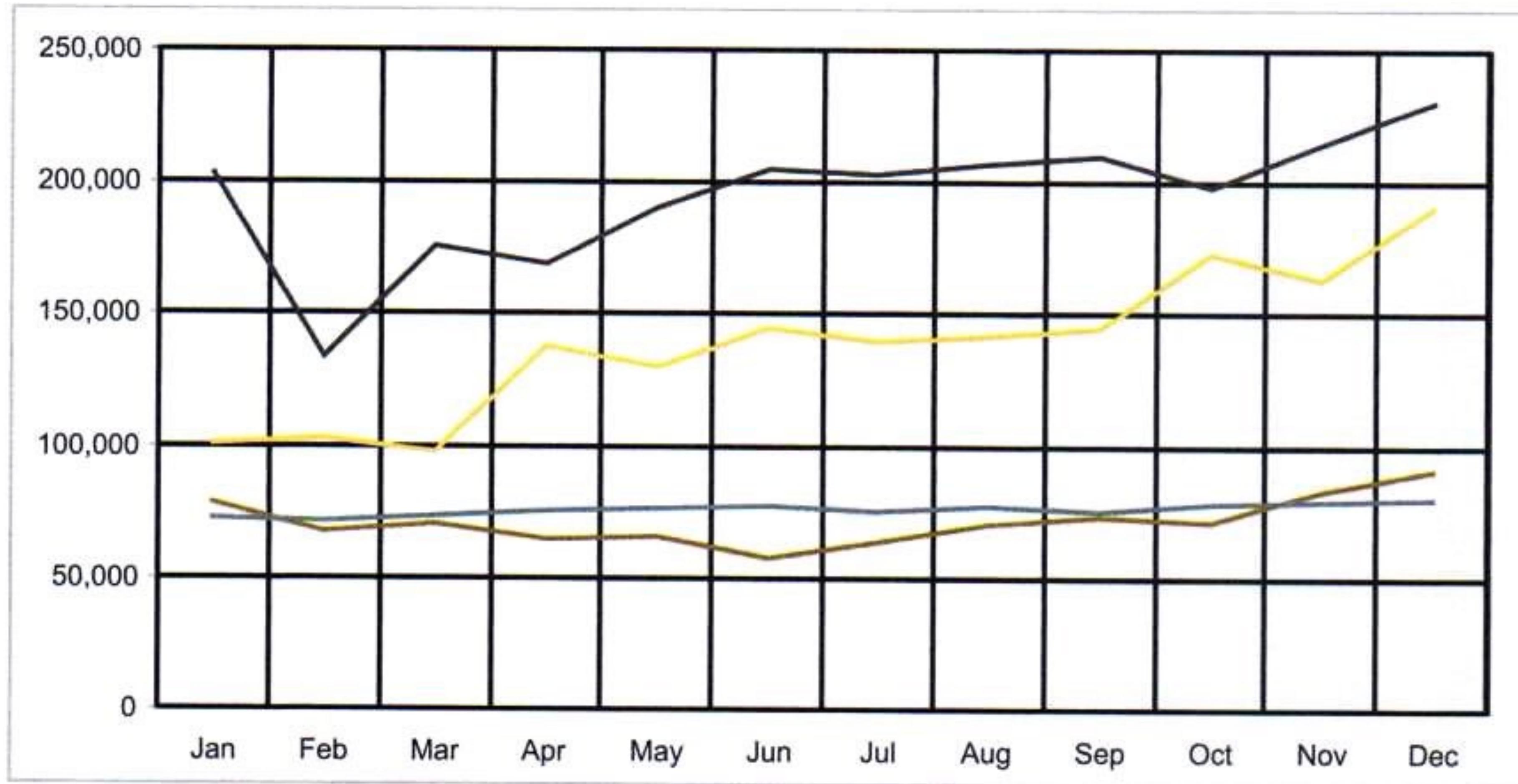
Quando se quer dar destaque à área de dados, pode-se também colocar uma cor de fundo suave ao seu redor



# GRIDS

---

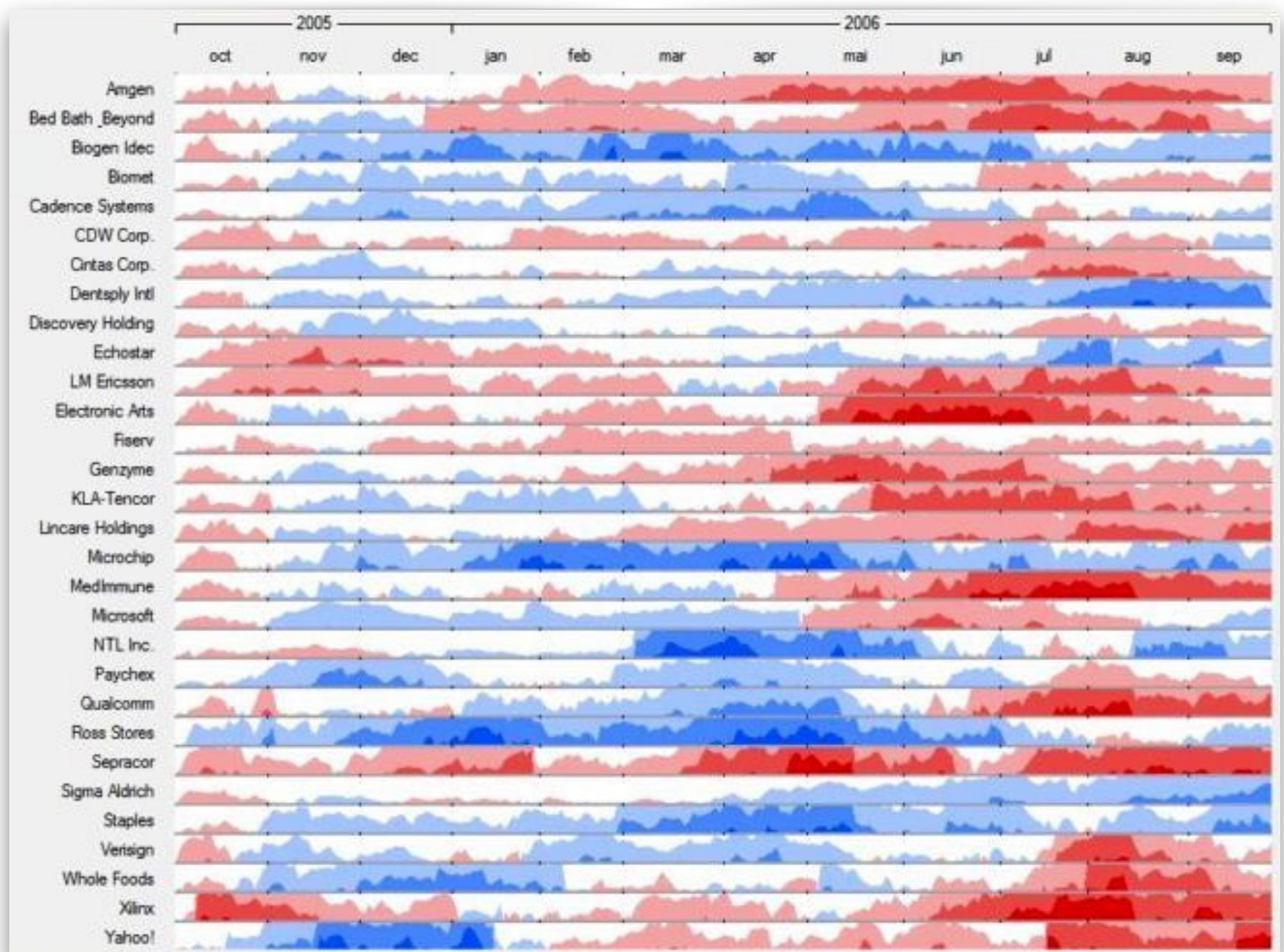
- Grids têm as seguintes funções:
- Melhorar a visibilidade dos valores correspondentes aos objetos visuais
- Facilitar a comparação entre valores
- Melhorar a percepção e comparação de padrões localizados
- Grids não representam dados logo devem ser visualmente secundários



# **Relacionamentos quantitativos, técnicas de visualização e padrões visuais**

---

**Profa. Dra. Raquel Minardi**  
Departamento de Ciência da Computação  
Universidade Federal de Minas Gerais

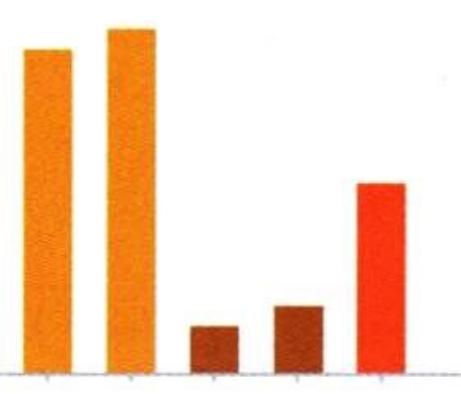
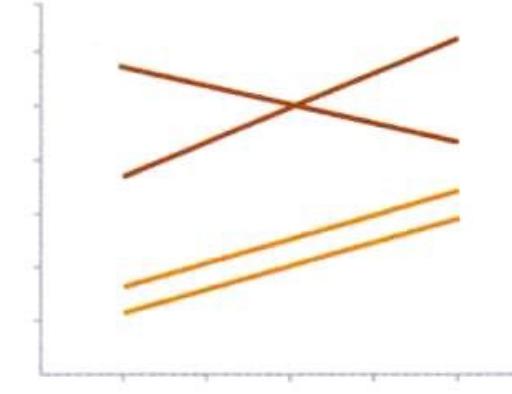
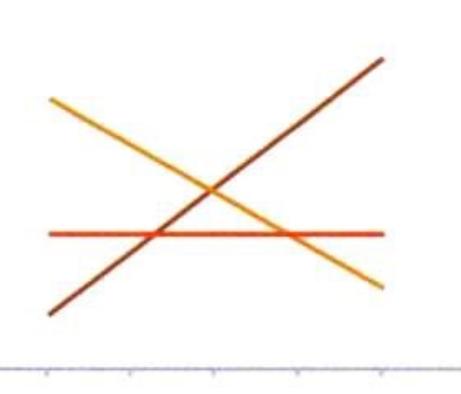
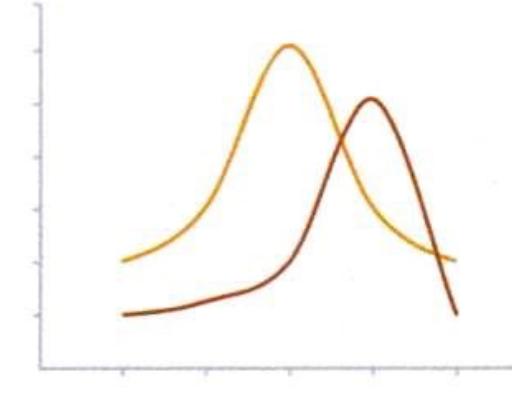
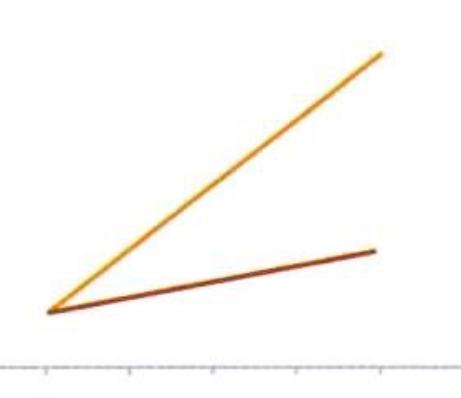
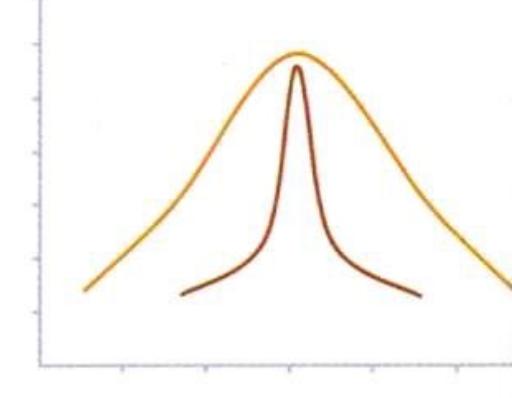
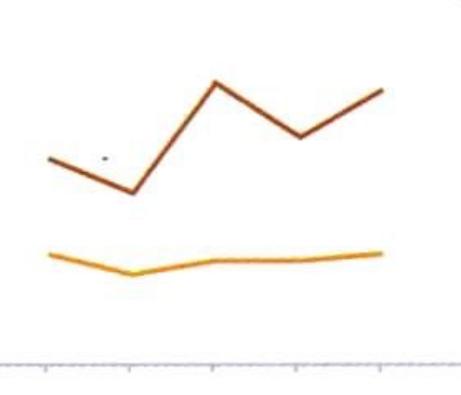
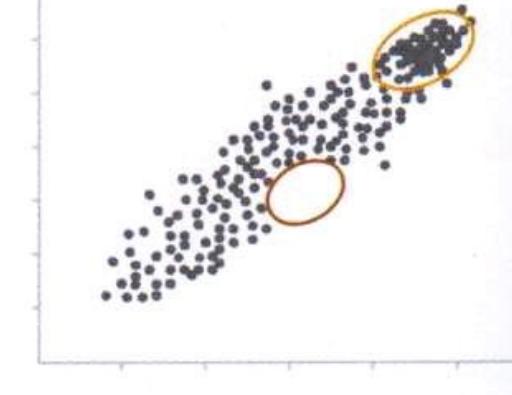
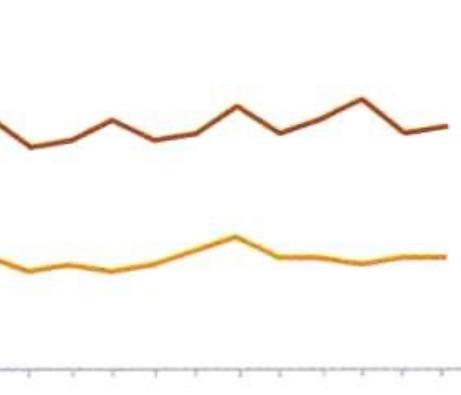
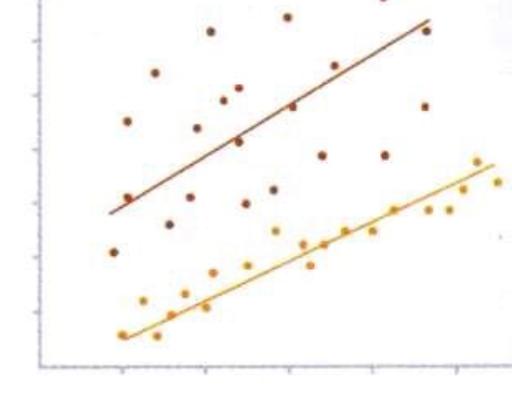
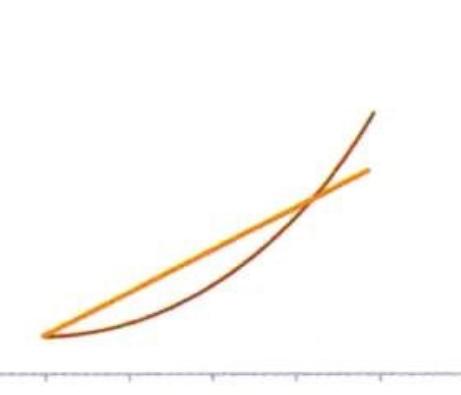
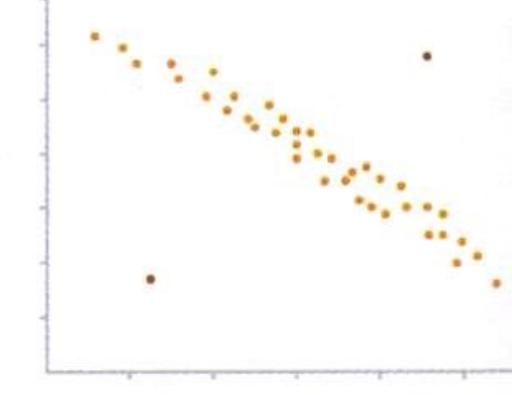


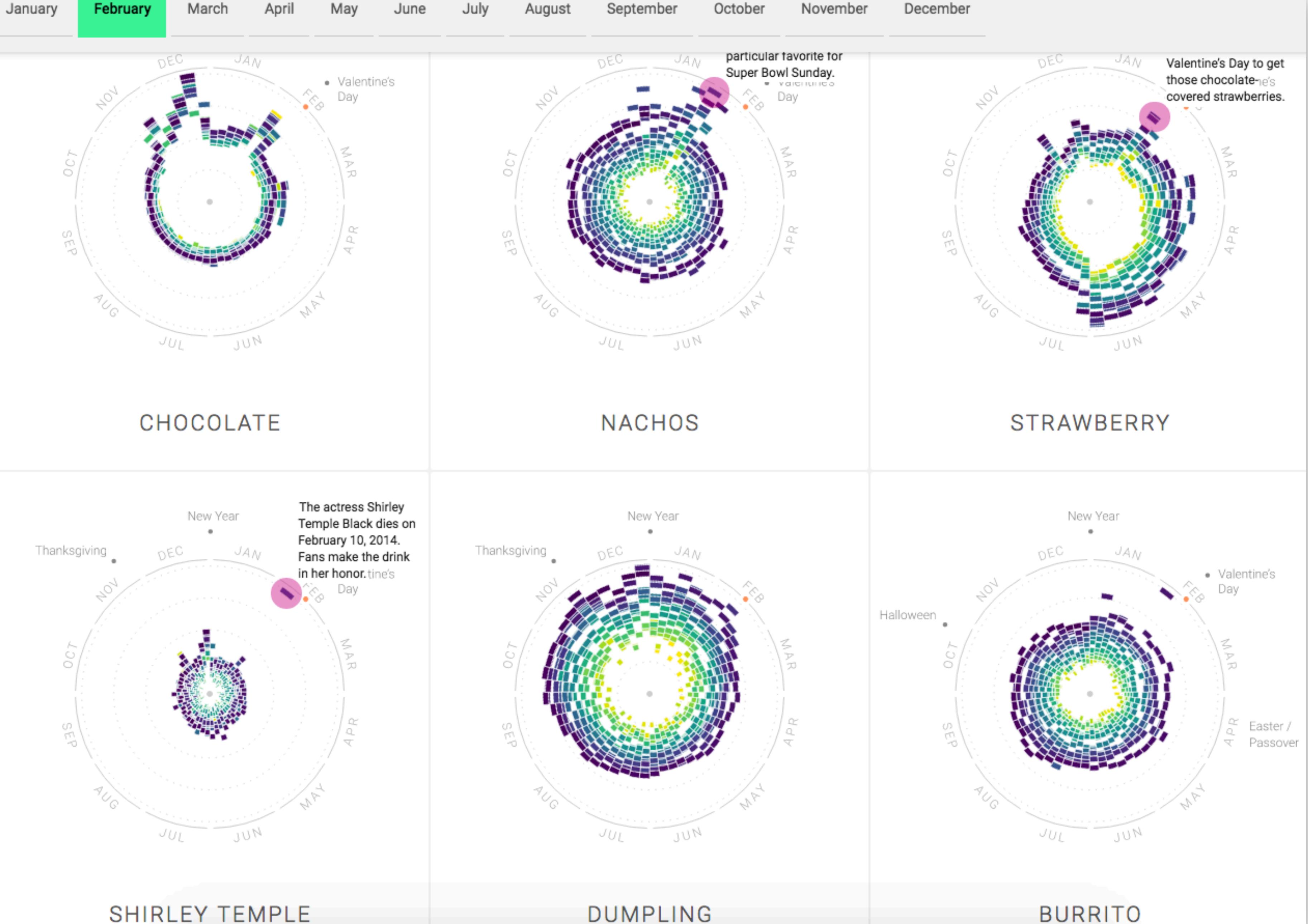
A **análise** envolve **relacionamentos entre valores** que, por sua vez, envolvem a **busca por padrões visuais** nos dados e o uso de **técnicas** de análise particulares

# RELACIONAMENTOS QUANTITATIVOS

---

- Os principais tipos de relacionamentos são:
  - Séries temporais
  - Ranking e parte-todo
  - Desvio
  - Distribuição
  - Correlação
  - Multivariados

Pattern	Example	Pattern	Example
High, low, and in between		Non-intersecting and intersecting	
Going up, going down, and remaining flat		Symmetrical and skewed	
Steep and gradual		Wide and narrow	
Steady and fluctuating		Clusters and gaps	
Random and repeating		Tightly and loosely distributed	
Straight and curved		Normal and abnormal	



- Como padrões visuais particulares podem ser úteis em determinados problemas
- Técnicas de visualização que mostram padrões significativos mais efetivamente
- Técnicas de interação analítica que são úteis e as melhores práticas a seguir

# Séries temporais - Parte I

---

**Profa. Dra. Raquel Minardi**

Departamento de Ciência da Computação

Universidade Federal de Minas Gerais

- Não há relacionamento que receba mais atenção que as séries temporais
- Estima-se que cerca de 75% dos gráficos publicados em jornais internacionais sejam séries temporais
- O tempo apresenta um contexto muito importante para a compreensão dos dados
- “O presente pode ser entendido, o futuro pode se prever na luz do passado”

# PADRÕES ANALÍTICOS

---

- Padrões analíticos úteis quando analisamos dados ao longo do tempo:
  - Tendência
  - Variabilidade
  - Taxa de mudança
  - Covariação
  - Ciclos
  - Exceções

# PADRÕES ANALÍTICOS

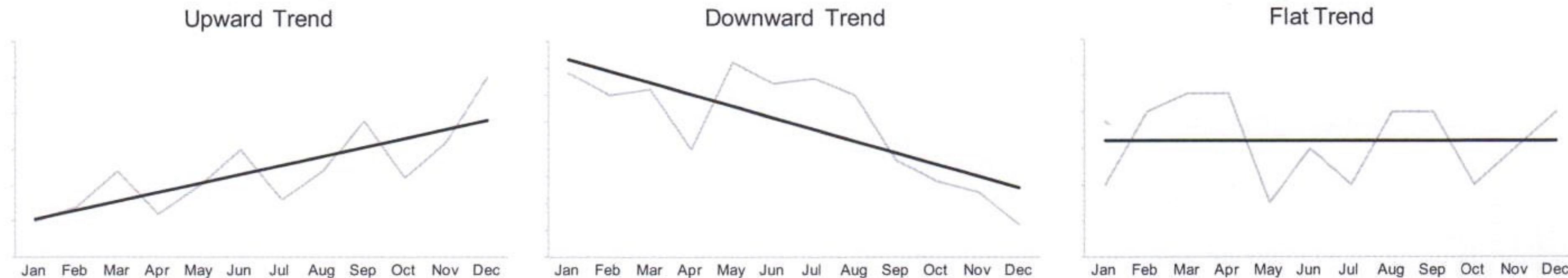
---

# PADRÕES ANALÍTICOS

---

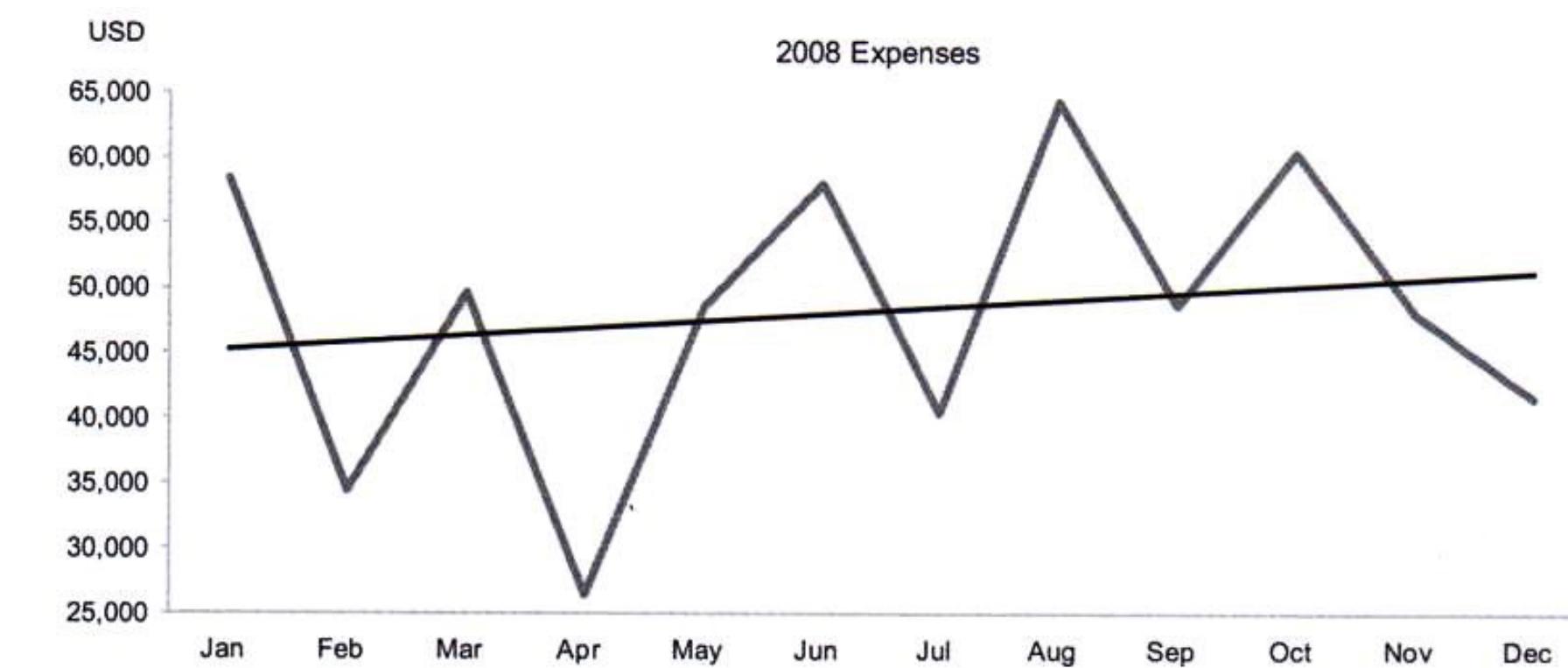
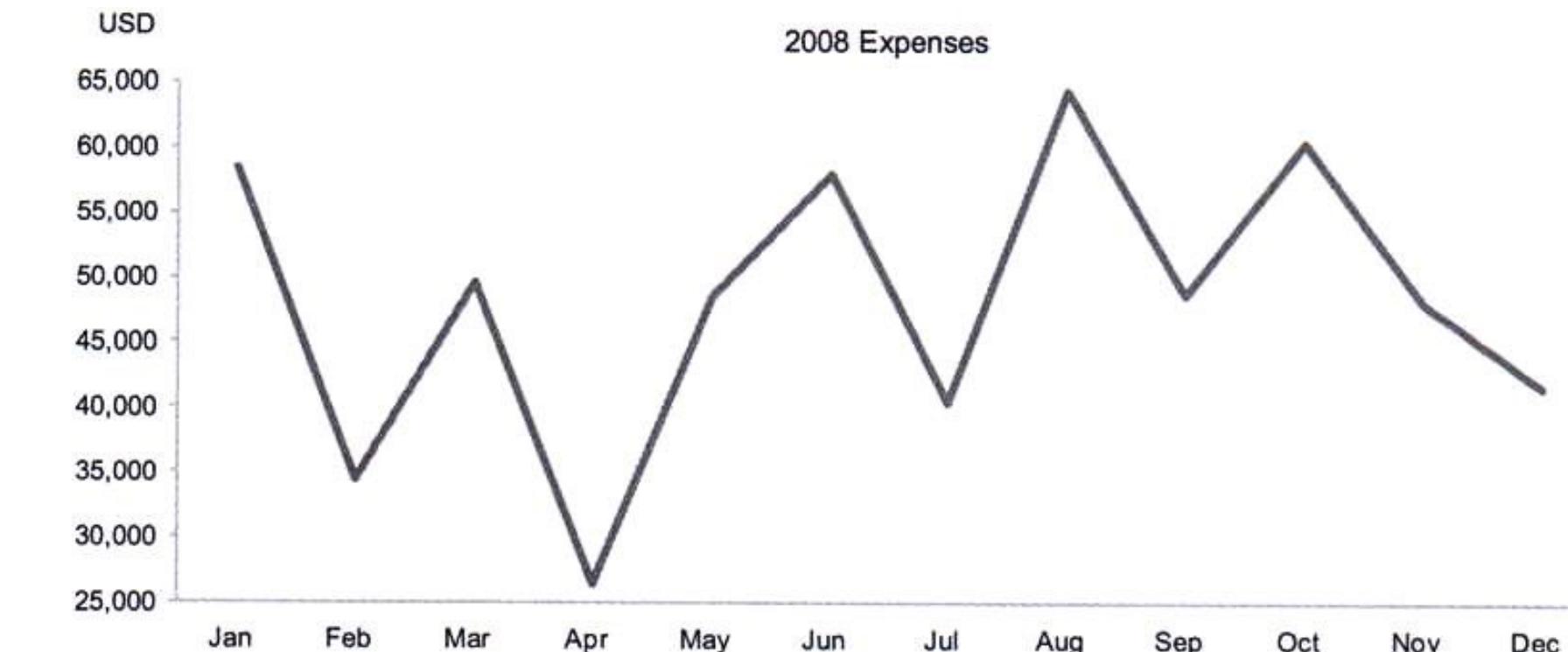
## TENDÊNCIA

- Propensão de uma série de valores em crescer, decrescer ou permanecer estável durante um período particular de tempo



**Gráficos de linha** são úteis na visualização de tendências

Quando há muita variabilidade, **linhas de tendência** são úteis na exibição de tendências



# PADRÕES ANALÍTICOS

---

## VARIABILIDADE

- Variabilidade é o grau médio de variação de um período de tempo para outro



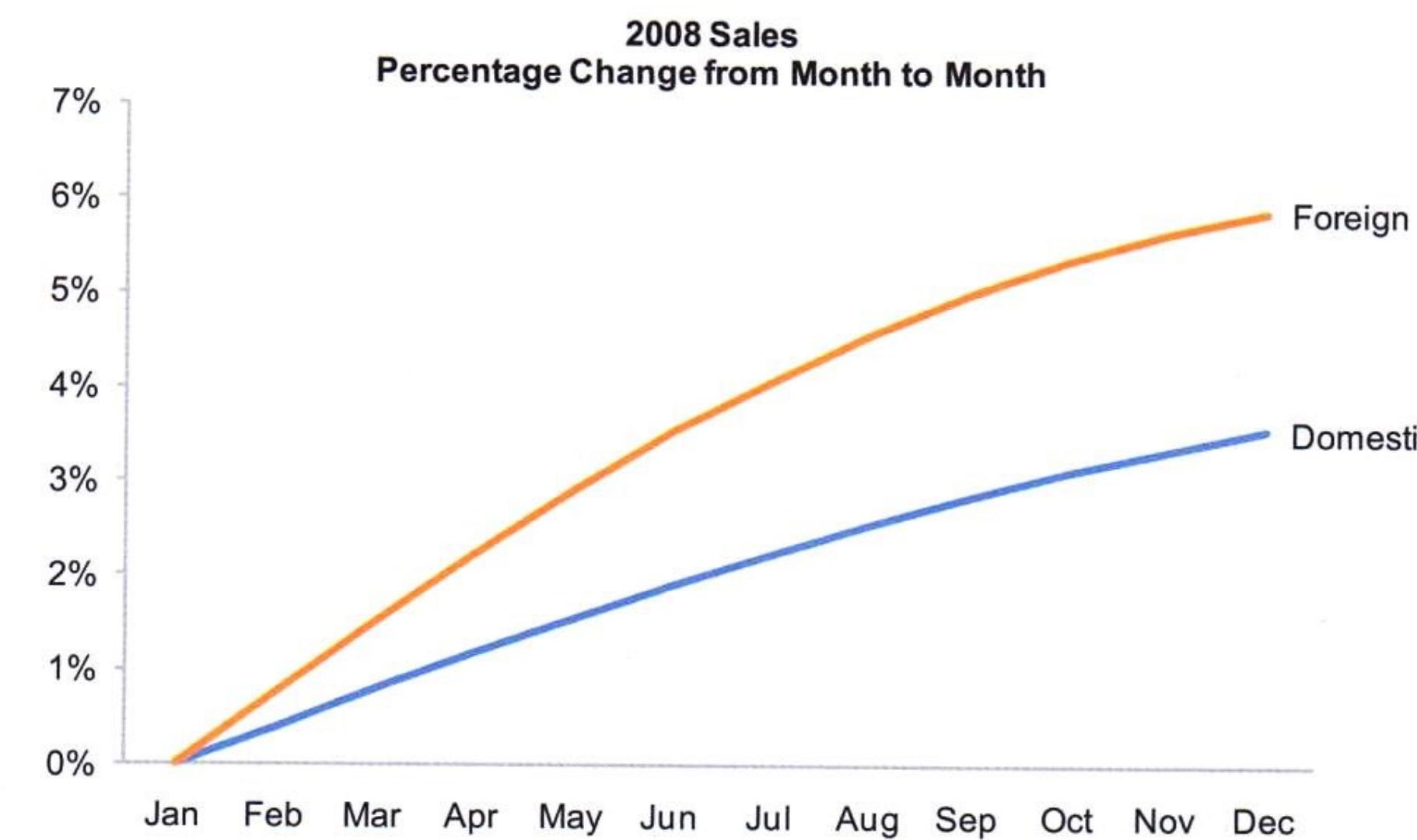
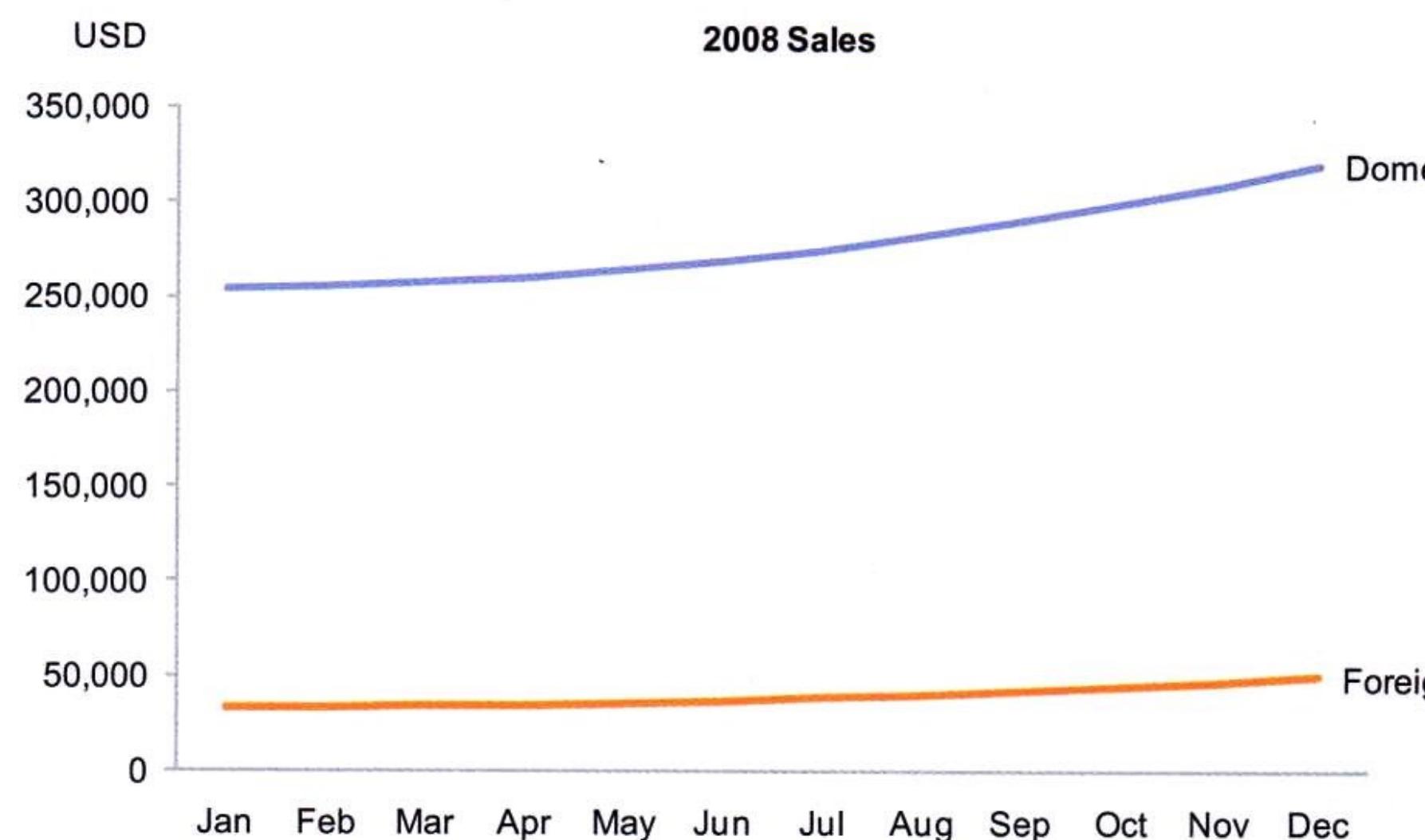


# PADRÕES ANALÍTICOS

---

## TAXA DE MUDANÇA

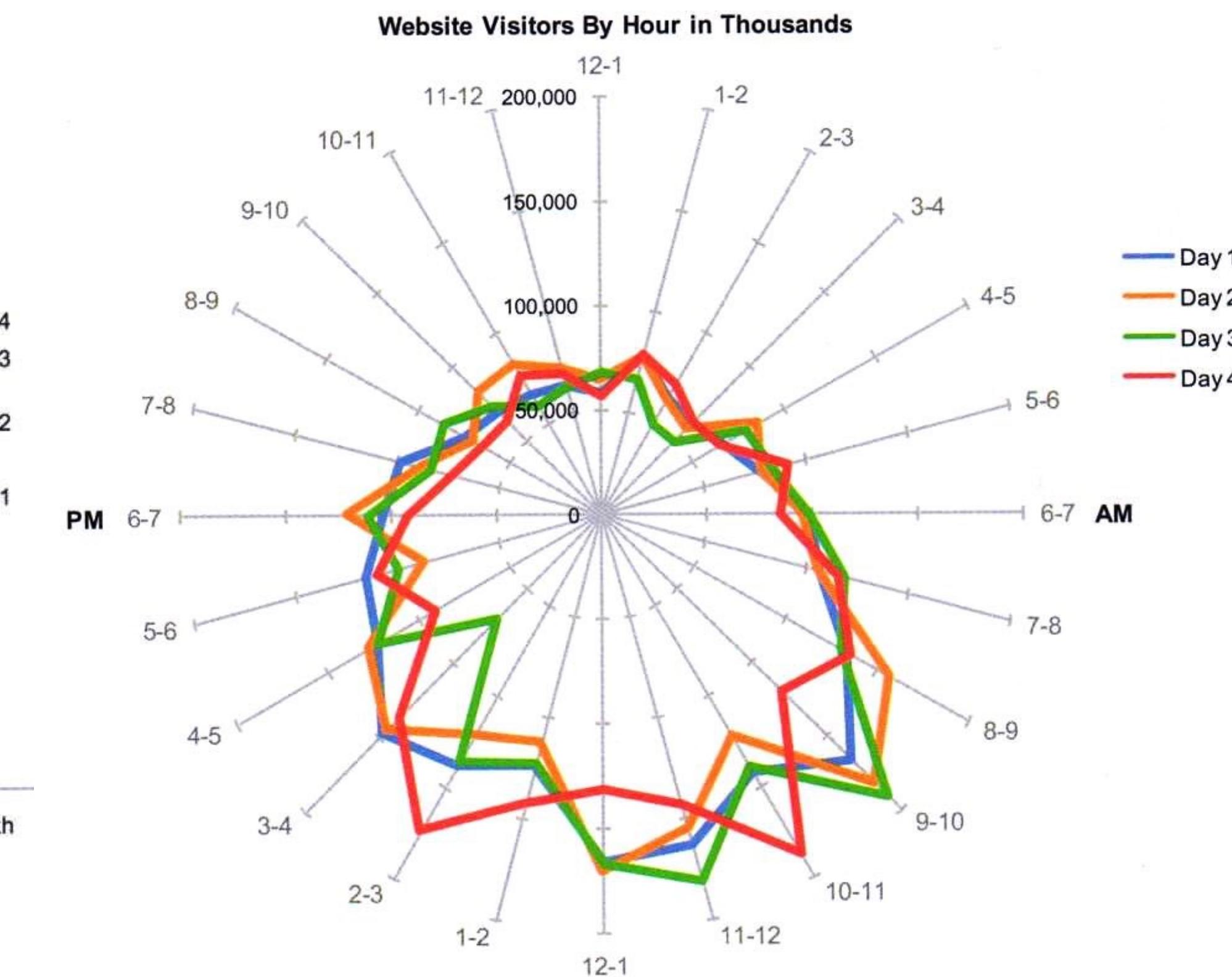
- A taxa de mudança pode ser expressa como o percentual de diferença entre valores referentes a diferentes períodos de tempo



# PADRÕES ANALÍTICOS

# CICLOS

- Ciclos são padrões repetidos em intervalos regulares



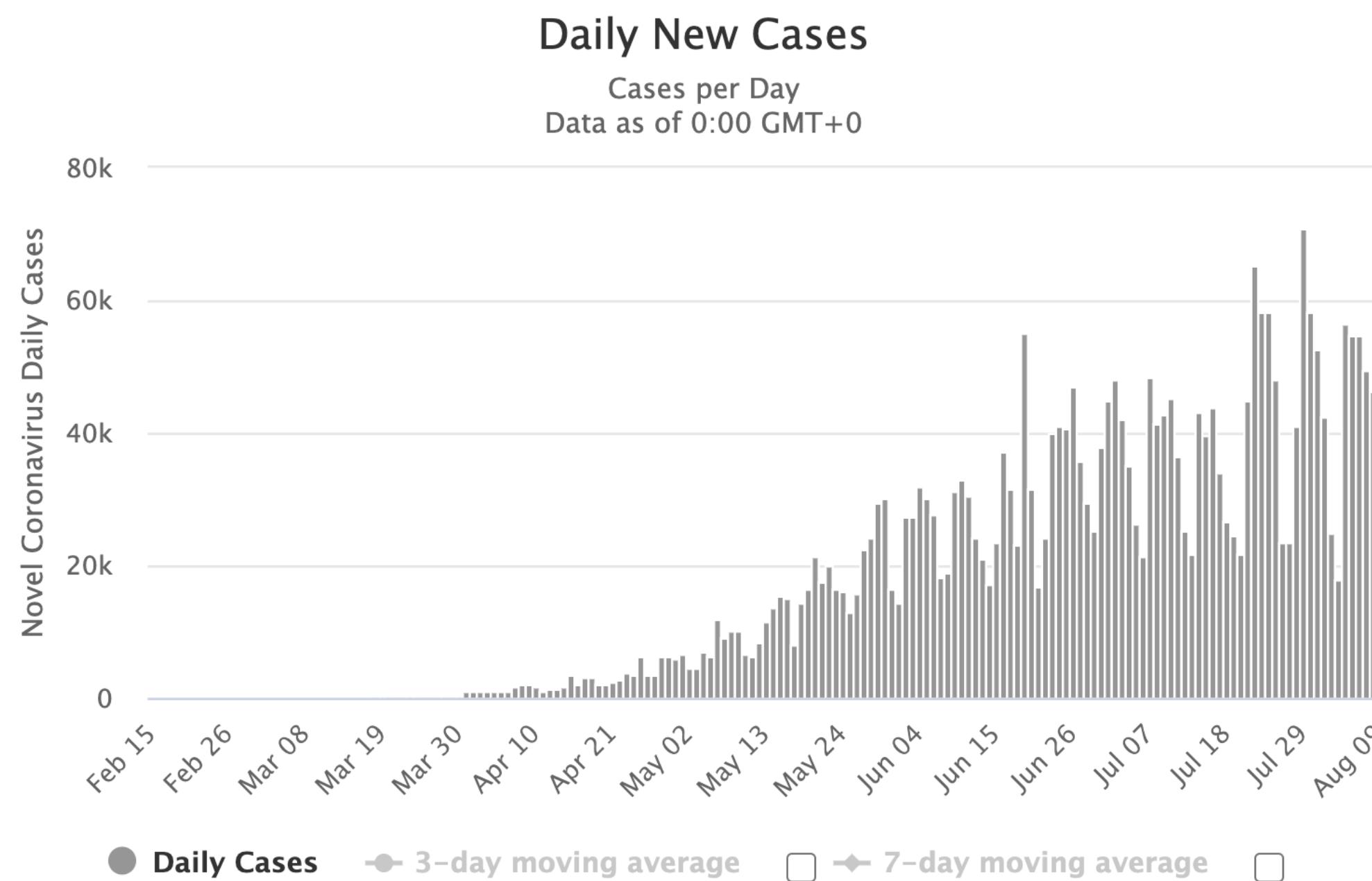
# PADRÕES ANALÍTICOS

---

## CICLOS

- Ciclos são padrões repetidos em intervalos regulares

Daily New Cases in Brazil



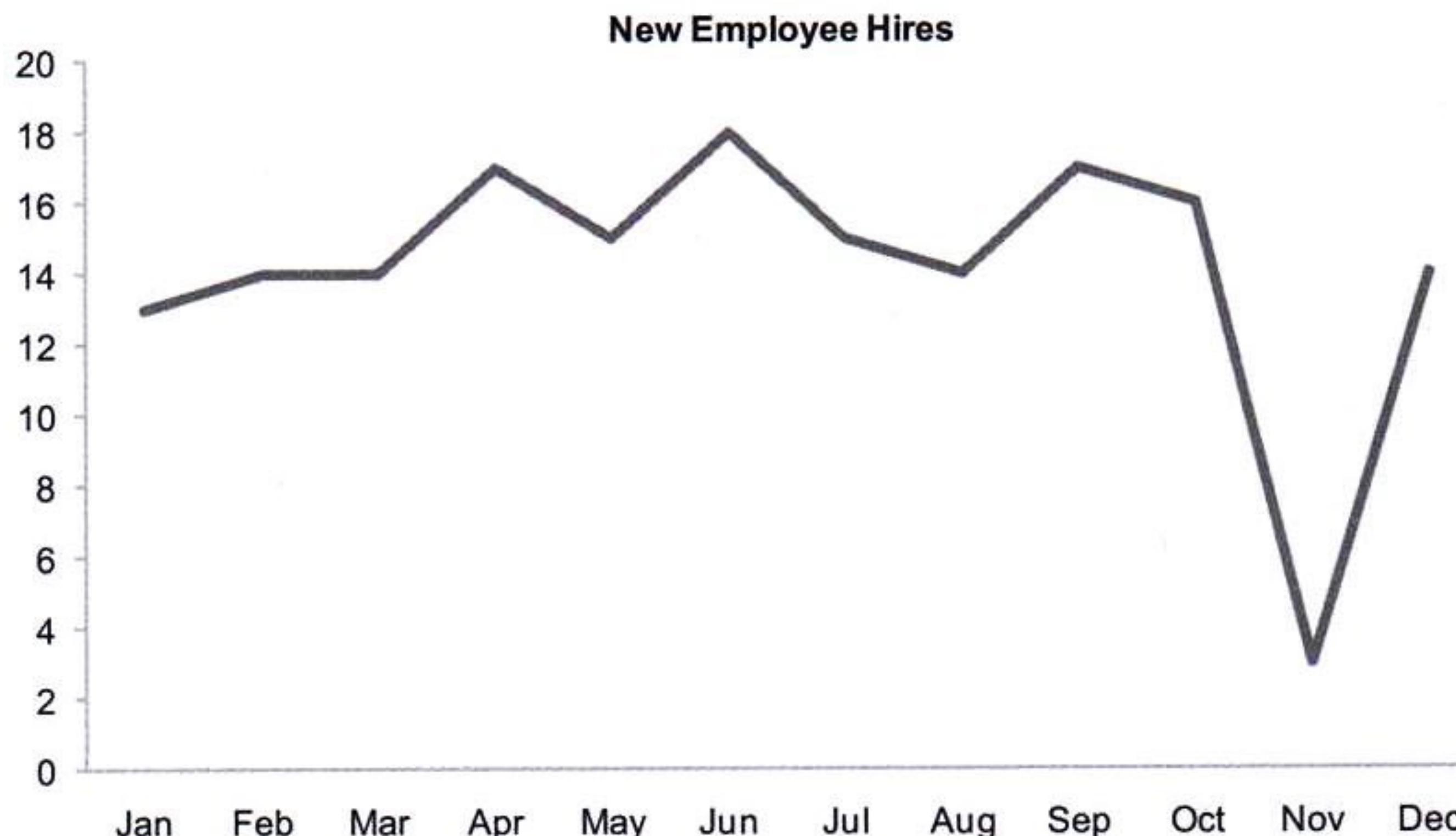
<https://www.worldometers.info/coronavirus/country/brazil/>

# PADRÕES ANALÍTICOS

---

## EXCEÇÕES

- Exceções são valores que saem da faixa de valores normais ou esperados



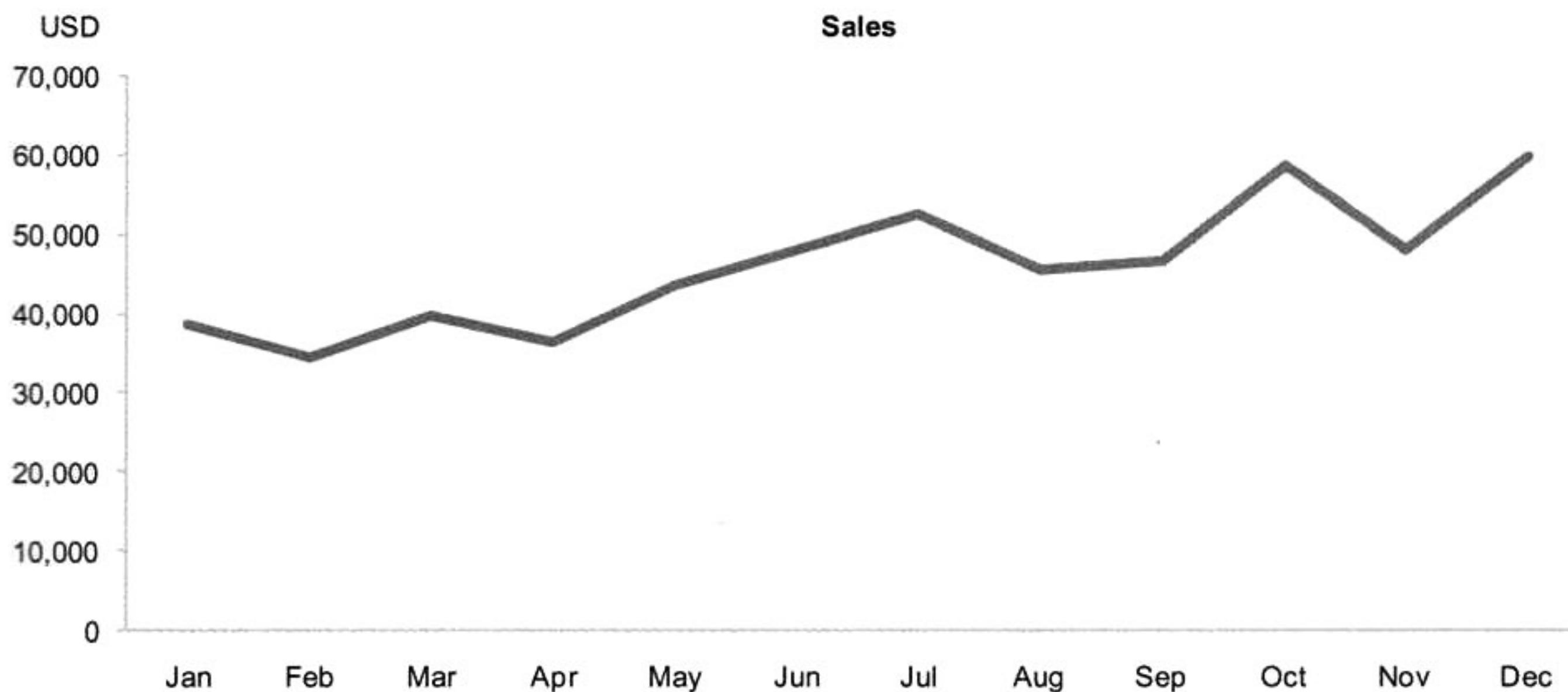
# REPRESENTAÇÕES VISUAIS

---

# GRÁFICO DE LINHAS

---

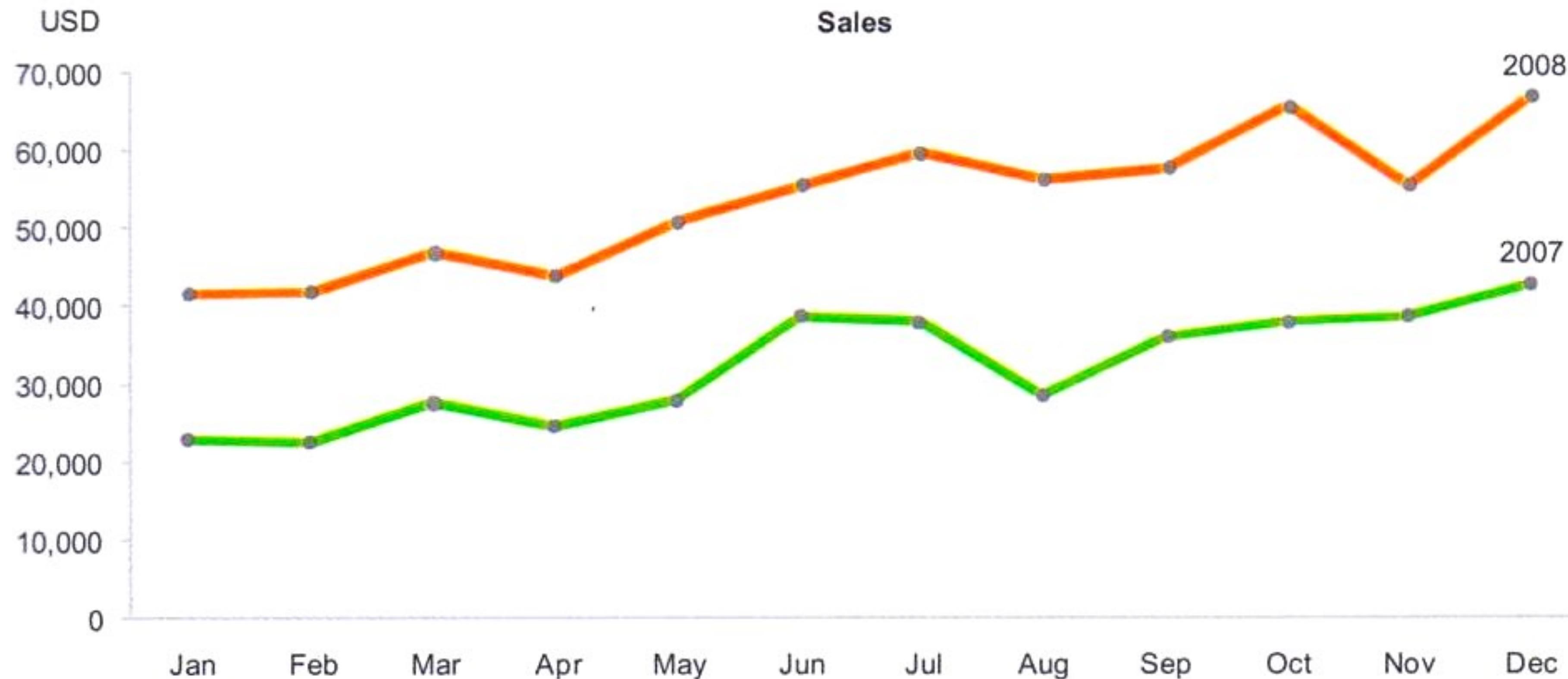
- Quando o objetivo é exibir como valores quantitativos mudam durante períodos contínuos de tempo, gráficos de linha são a opção mais precisa



# GRÁFICO DE LINHAS E PONTOS

---

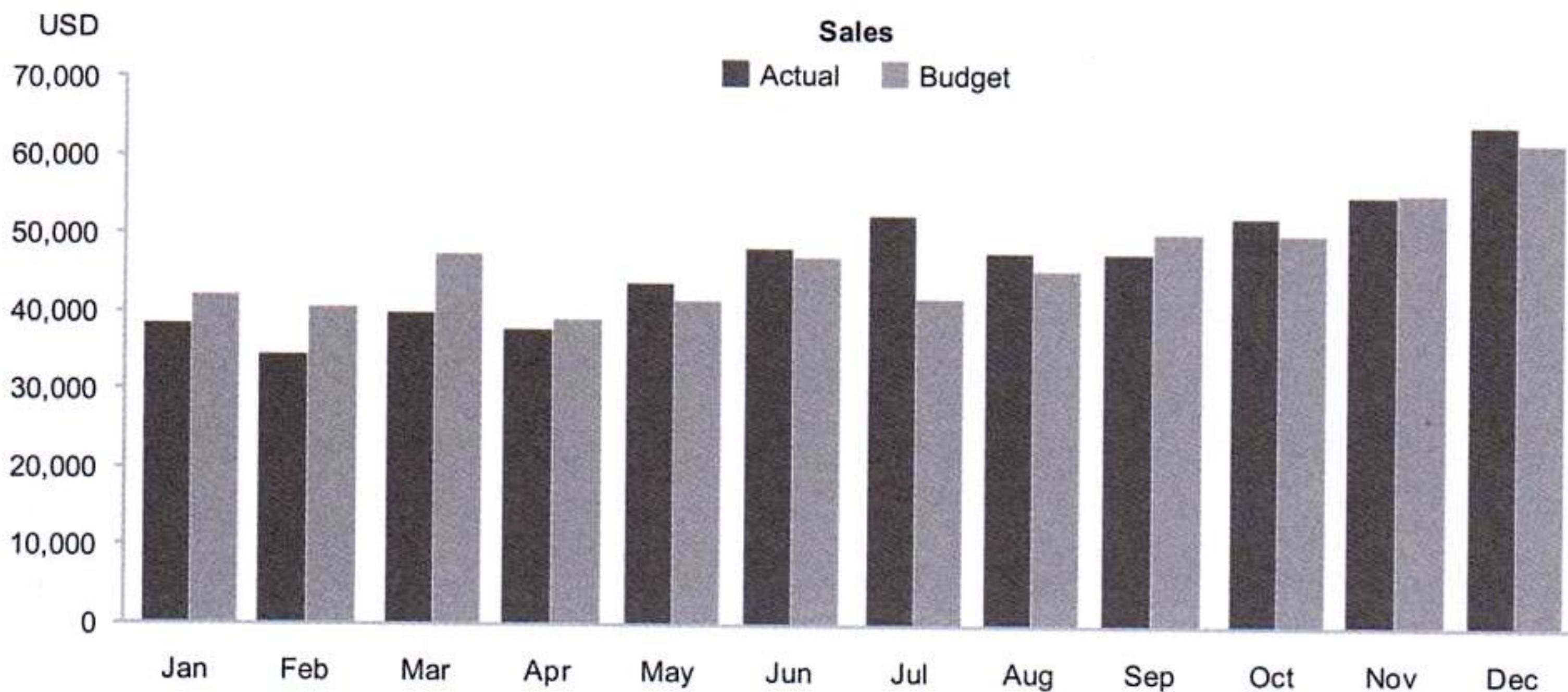
- Quando é necessário comparar valores individuais, usar a combinação de linhas e pontos



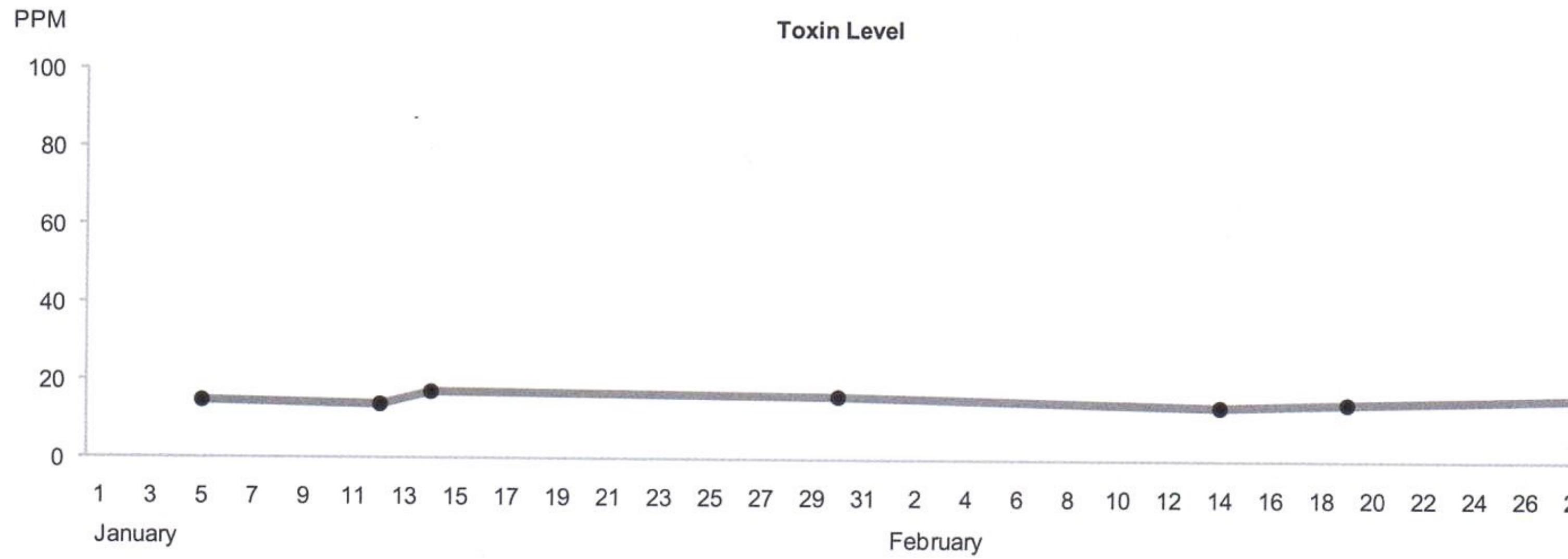
# GRÁFICO DE BARRAS

---

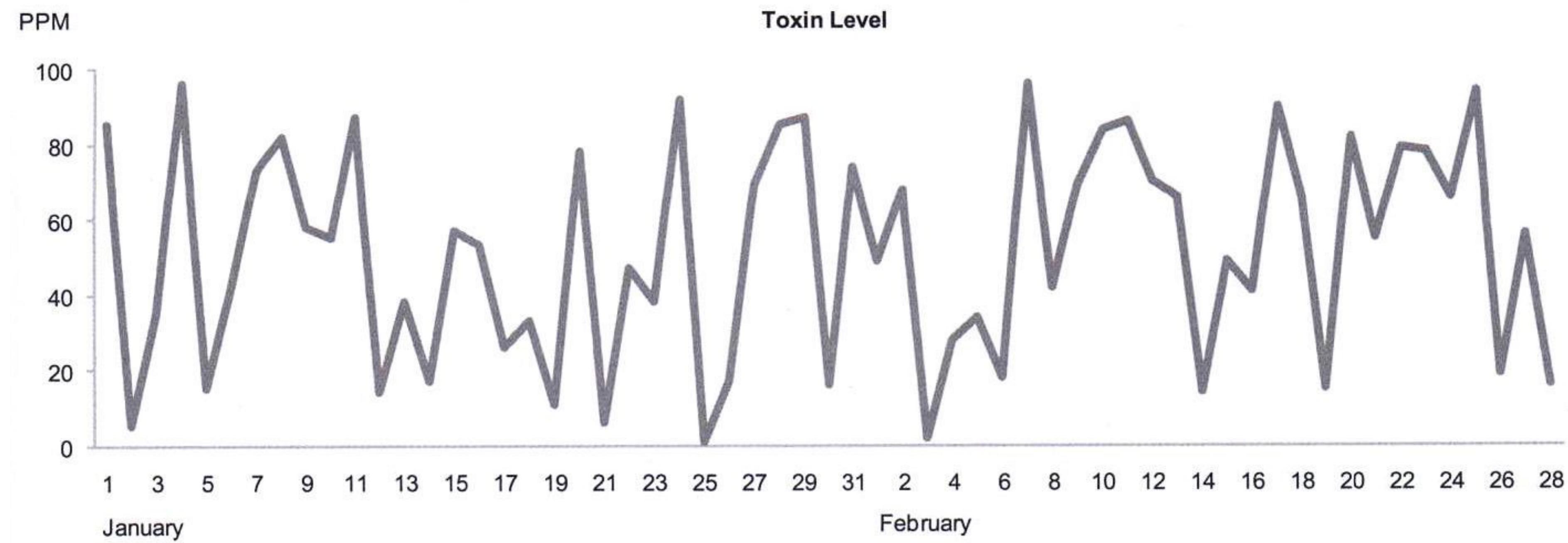
- Os comprimentos das barras permitem a comparação precisa entre os valores que representam mas não permitem a comparação do comportamento geral da série



O que fazer quando os intervalos dos dados não são igualmente espaçados?



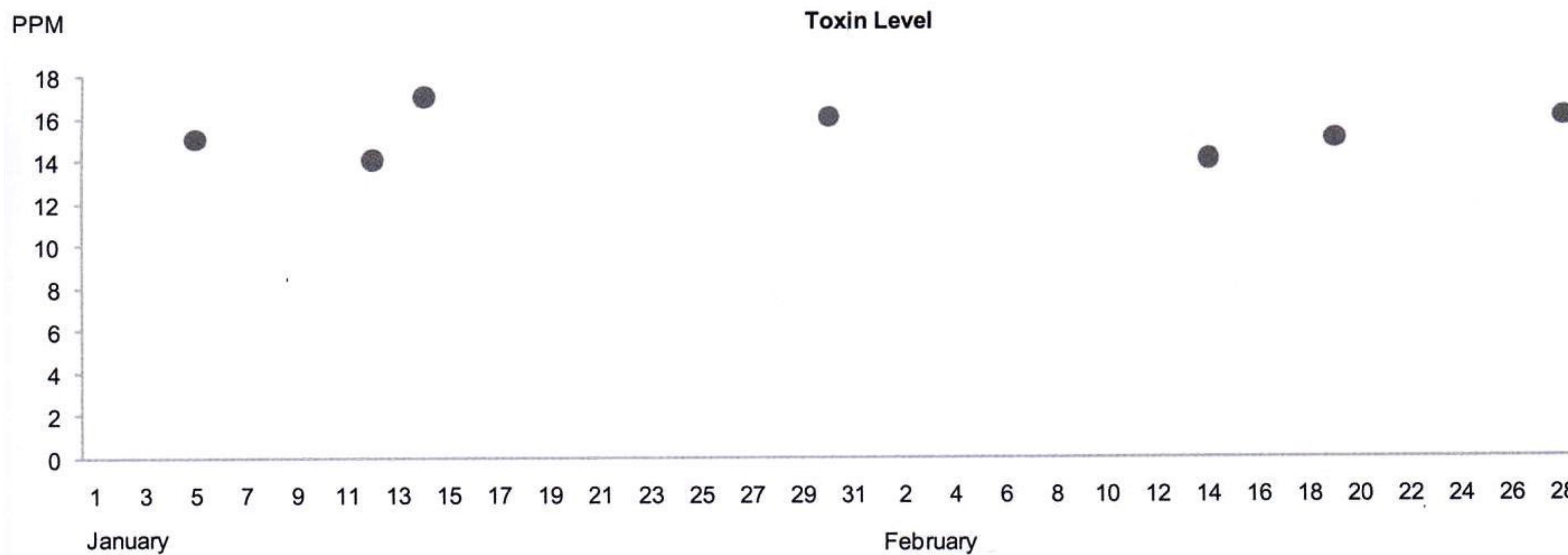
Neste exemplo, as transições suaves  
podem ser irreais

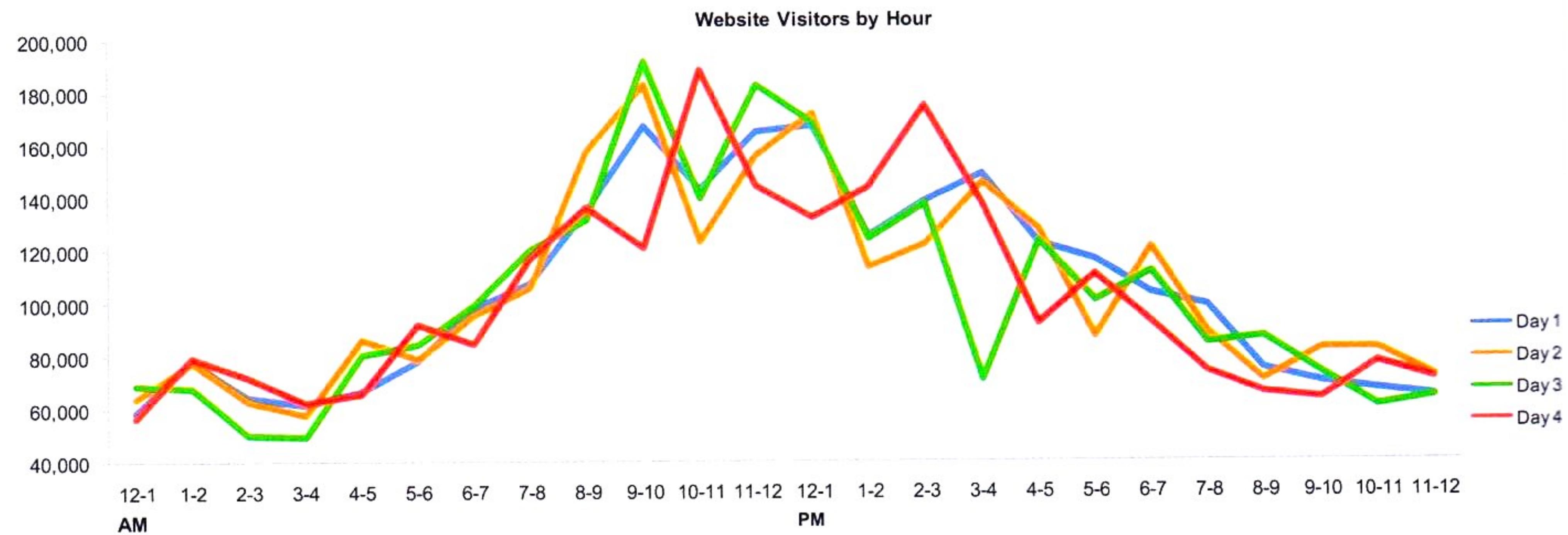


# GRÁFICO DE PONTOS

---

- Os gráficos de pontos desencorajam o pensamento incorreto de que há uma continuidade entre os valores

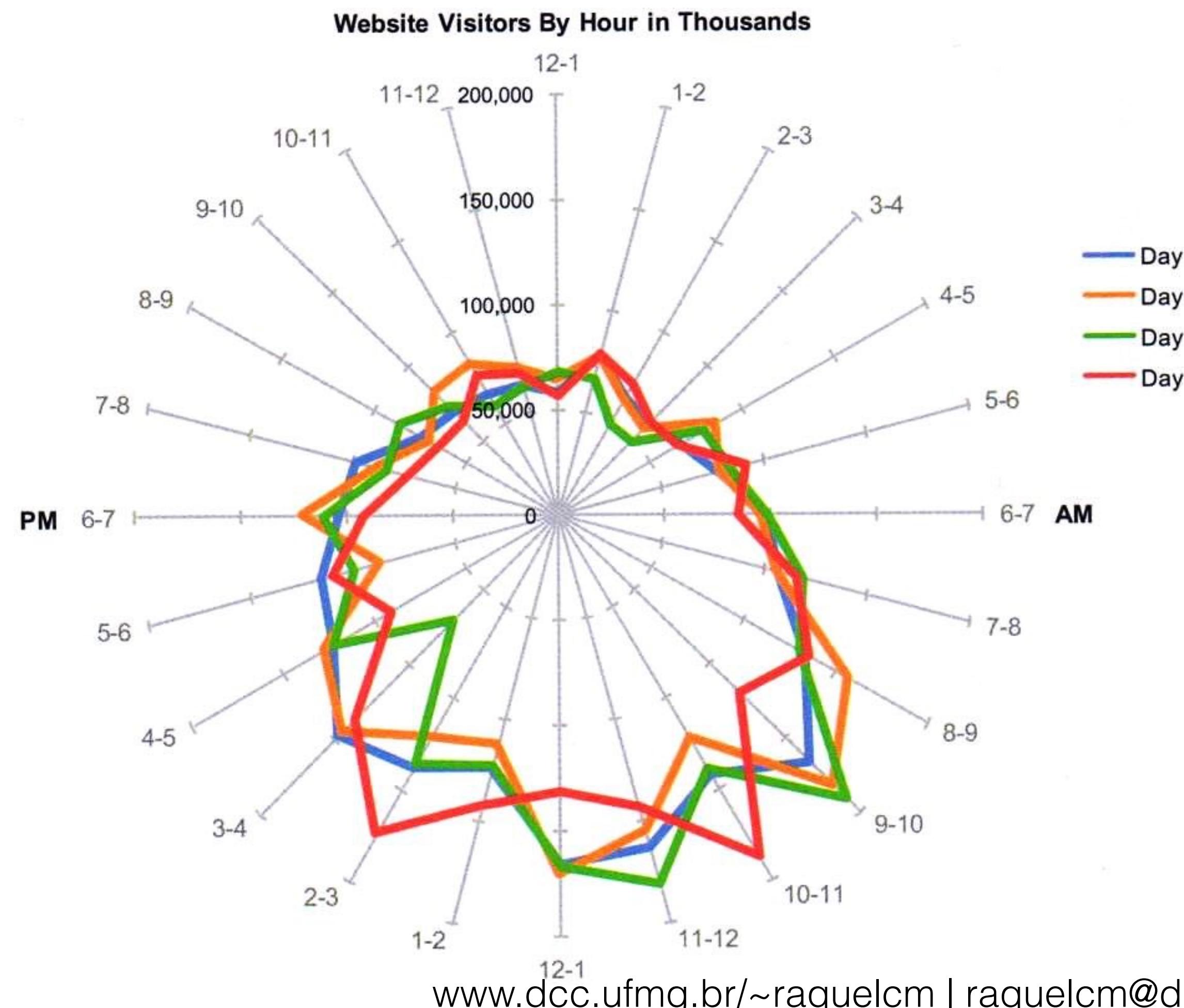


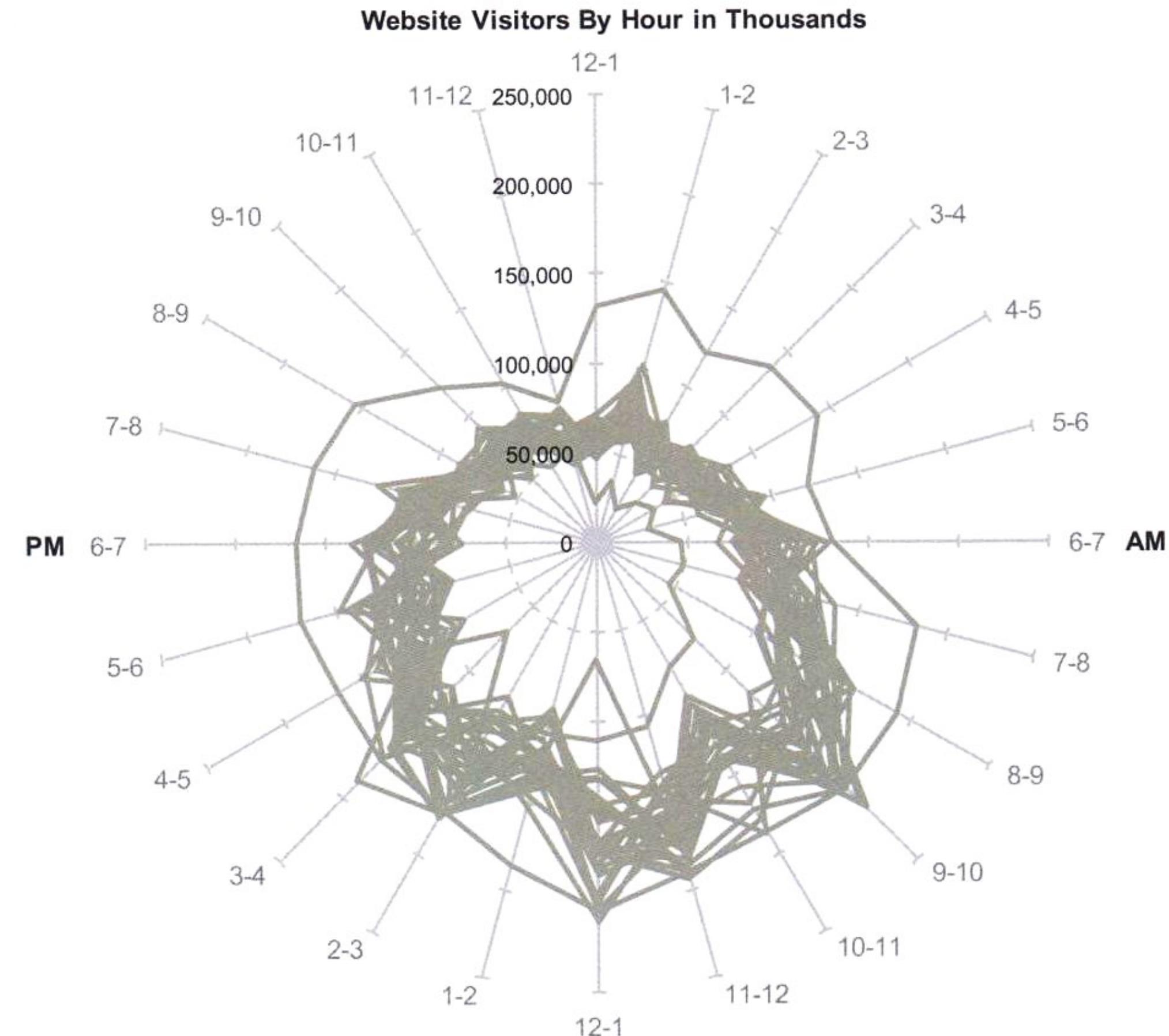


# GRÁFICO DE RADAR

---

- A forma circular dos gráficos de radar podem ser usadas para representar a natureza cíclica dos dados





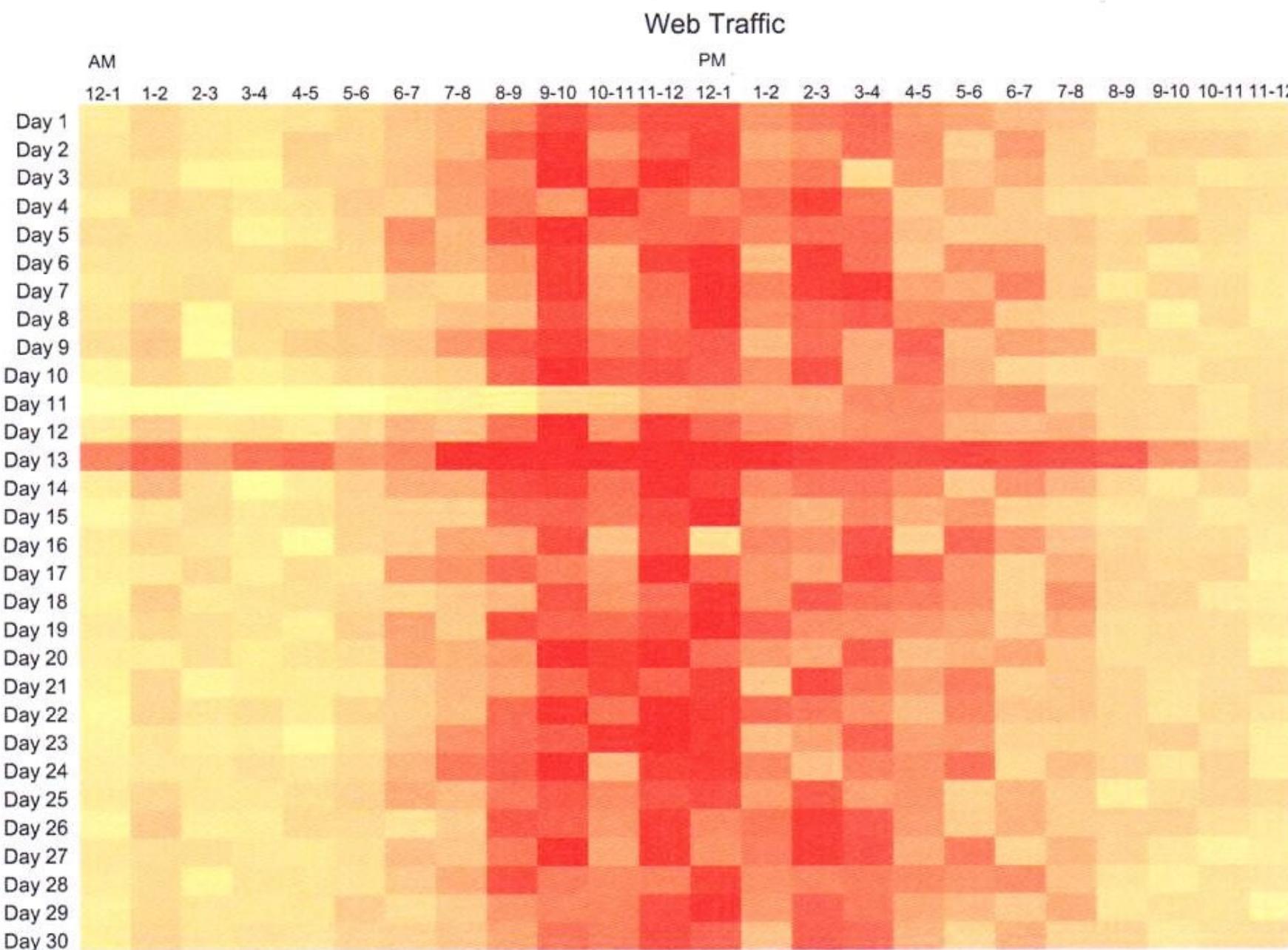
Quando há muitas séries, tanto o gráficos de linhas quanto o de radar se tornam confusos

Ainda assim podem ser úteis na exibição de **tendências** e **exceções**

# MAPAS DE CALOR (HEATMAPS)

---

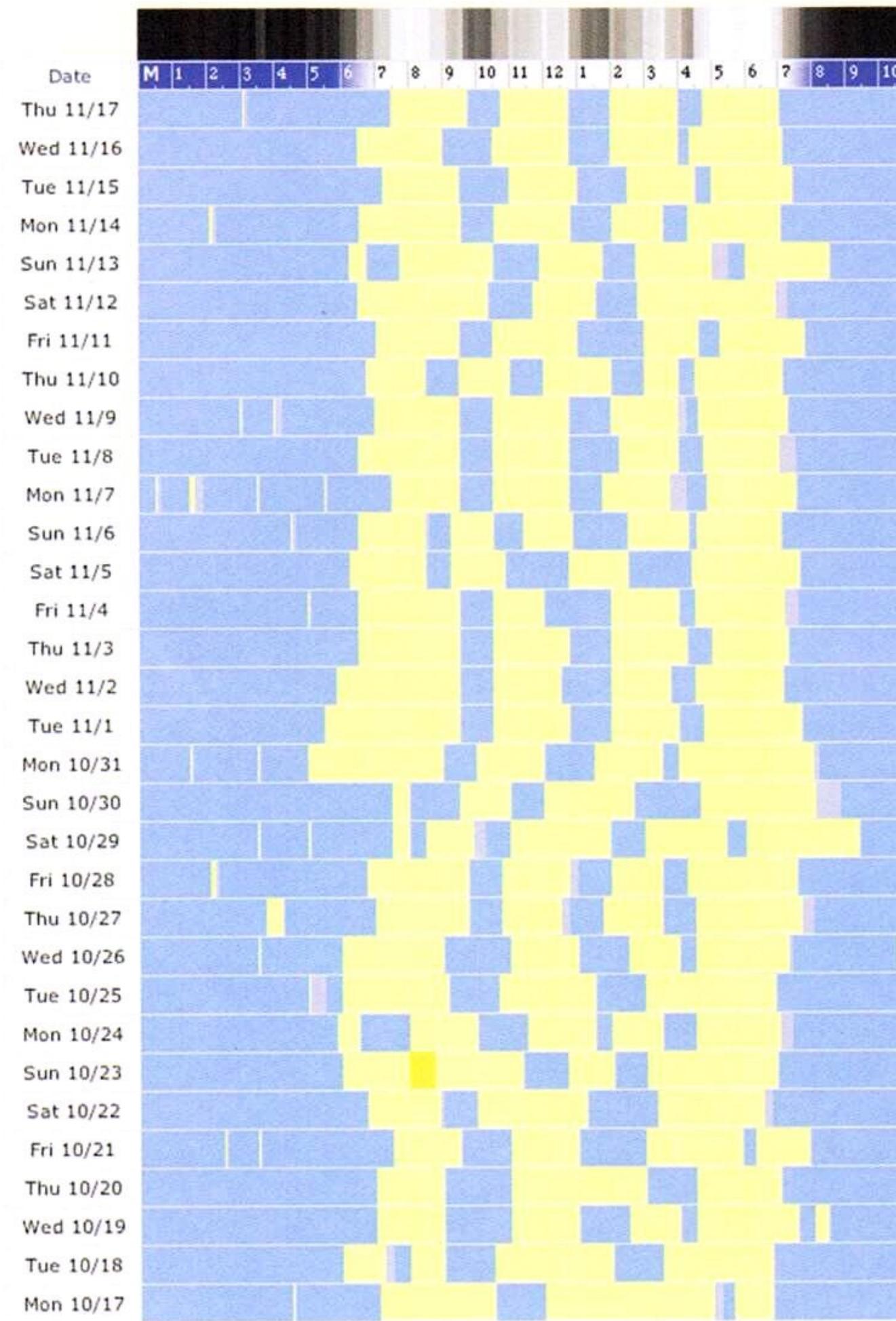
- *Heatmaps* ou mapas de calor são muito úteis na representação de grandes quantidades de dados cílicos sem o problema do *overplotting*
- Tipicamente, usam variações nas cores para representar variações quantitativas



## Sleep Telemetry

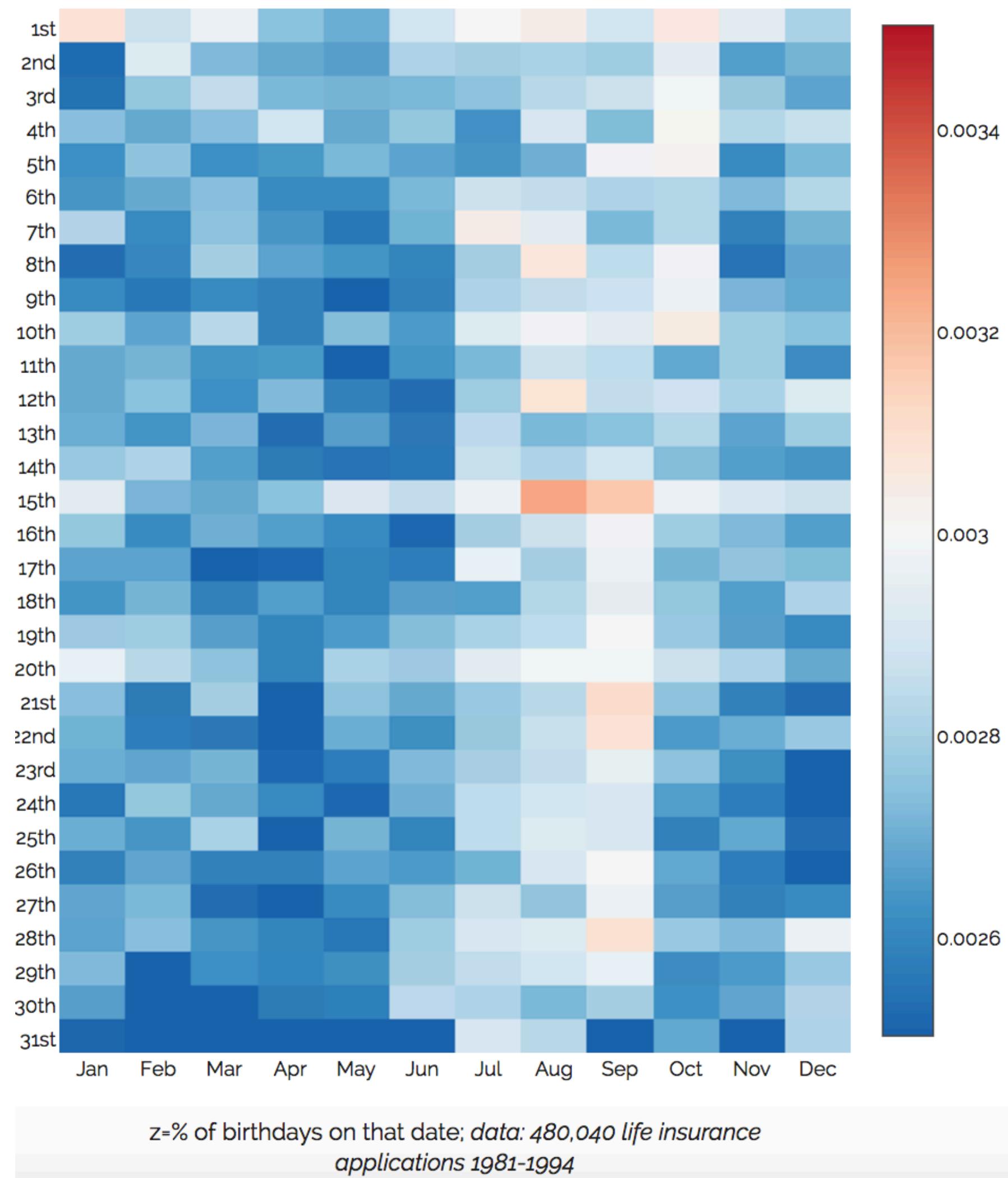
[Summary](#) | [Full chart](#) | [Stats](#) | [S Plot](#) | [Prob](#) | [Averages](#)

Change Age ▾ 6 months old

[Overnight](#) ?

www.trixietracker.com



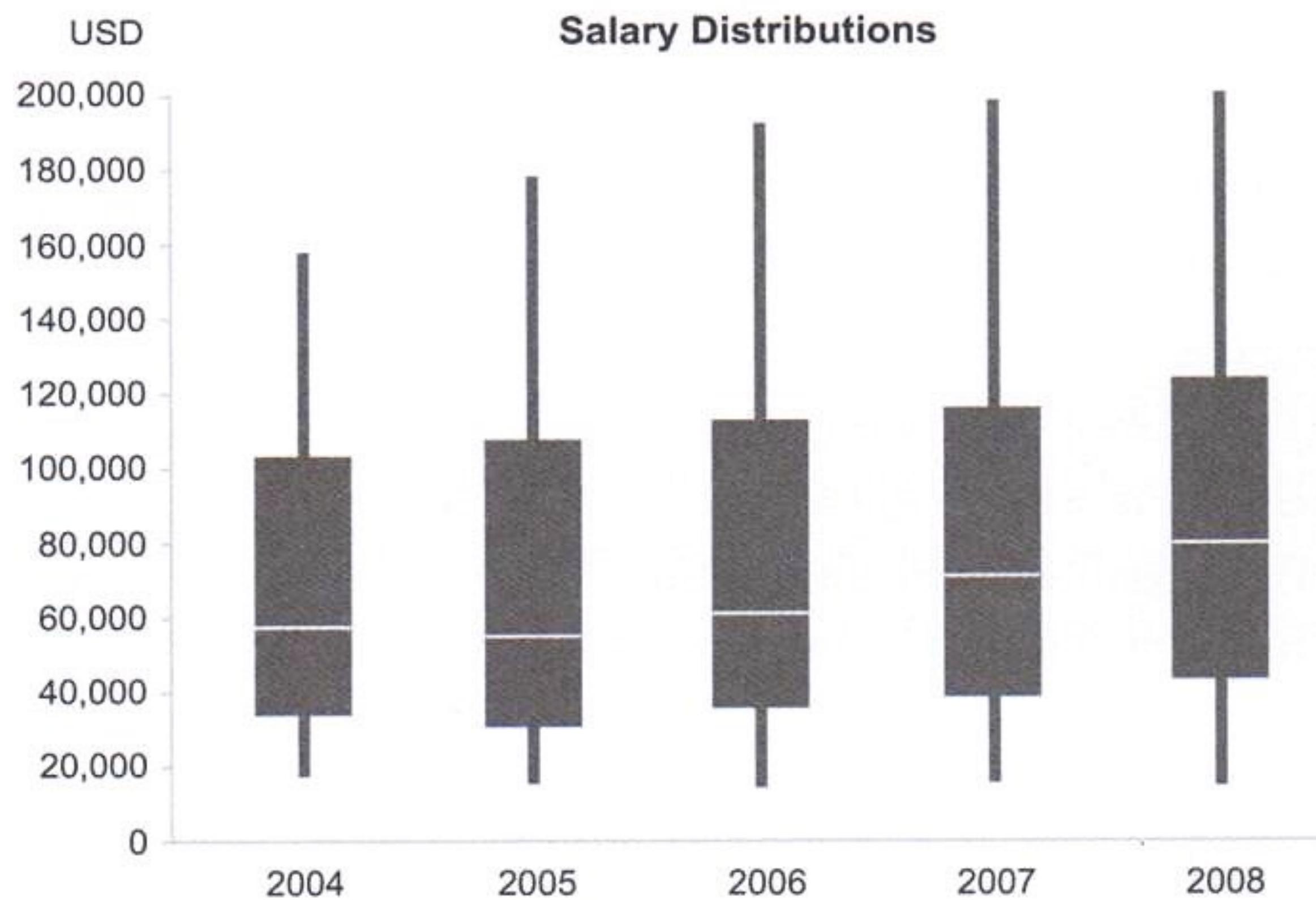


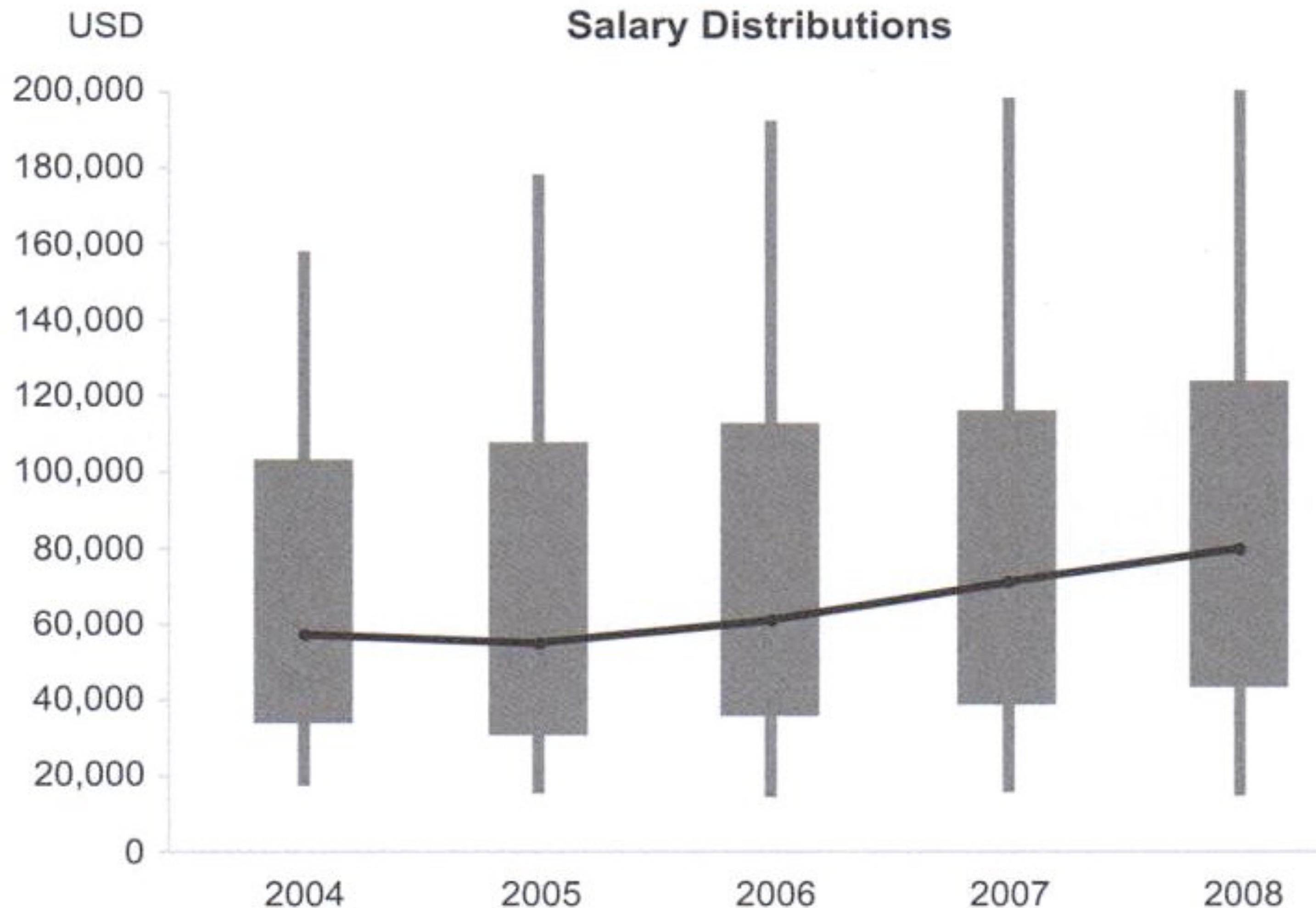
<http://blog.plot.ly/post/117105992082/time-series-graphs-eleven-stunning-ways-you-can>

# BOX PLOTS

---

- Úteis na visualização da distribuição dos valores no tempo



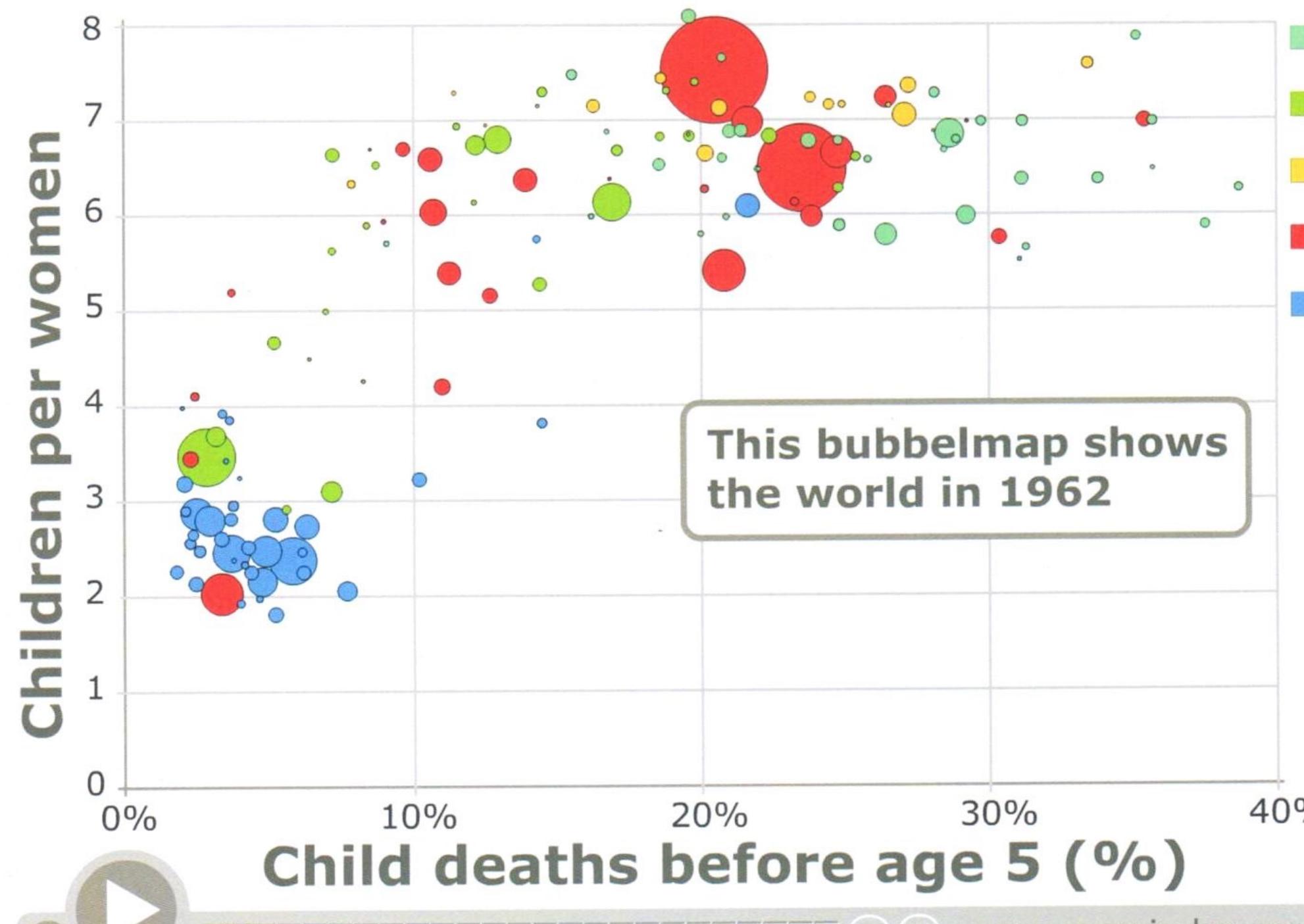


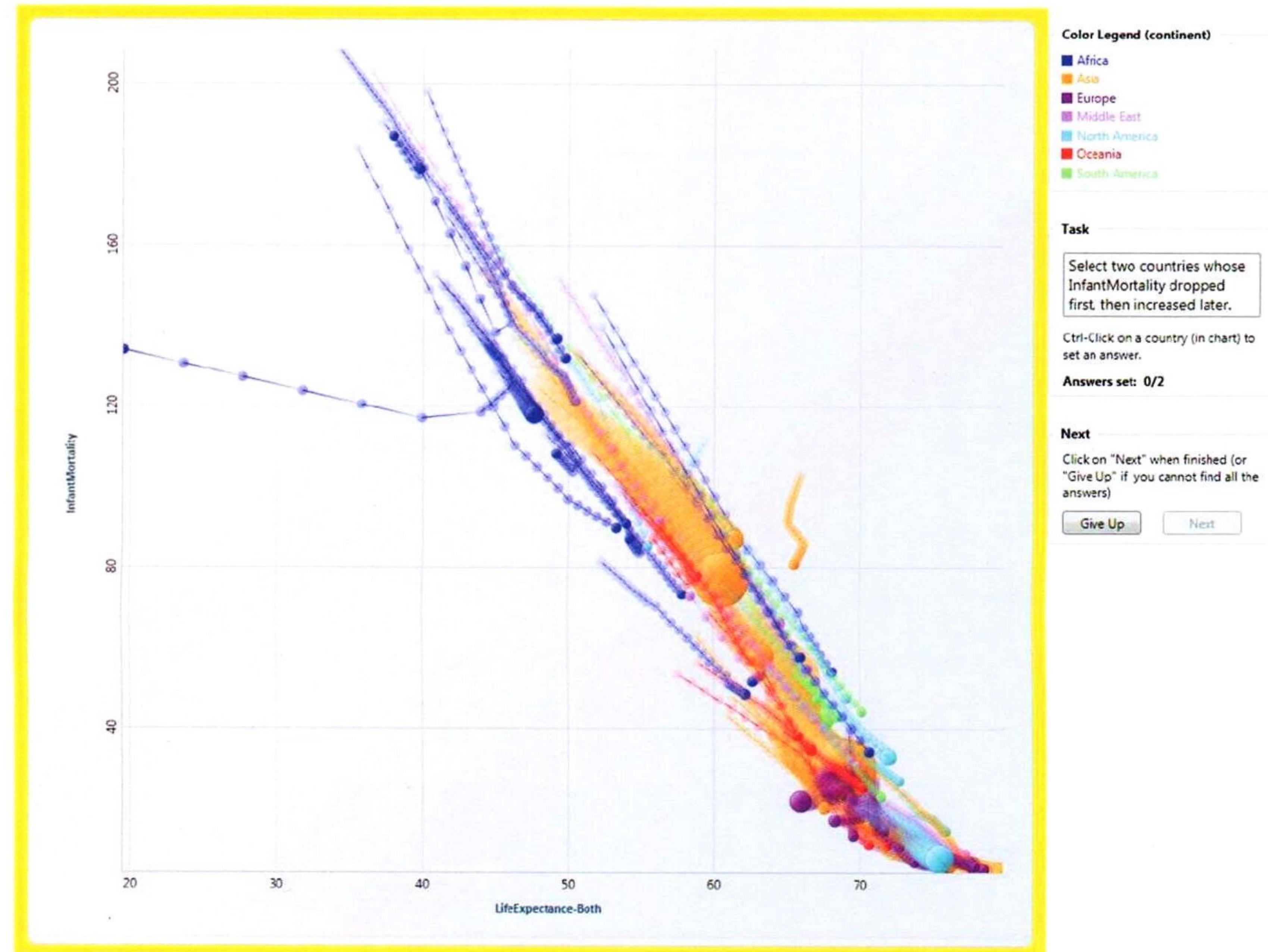
A linha que une as medianas ajuda a destacar a **tendência**

# GRÁFICO DE DISPERSÃO (SCATTERPLOT)

---

- Gráficos de dispersão são excelentes meios de se exibir correlações
- Os animados podem ser usados para mostrar como essas correlações evoluem







G.G. Robertson (2008) *Effectiveness of Animation in Trend Visualization*. IEEE Transactions on Visualization and Computer Graphics, 14(6):1325-1332.

Visualização de **animações** para **tendências** é uma tarefa desafiadora mesmo em apresentações

As pessoas acham atraente mas animações não são uma técnica muito efetiva para séries temporais

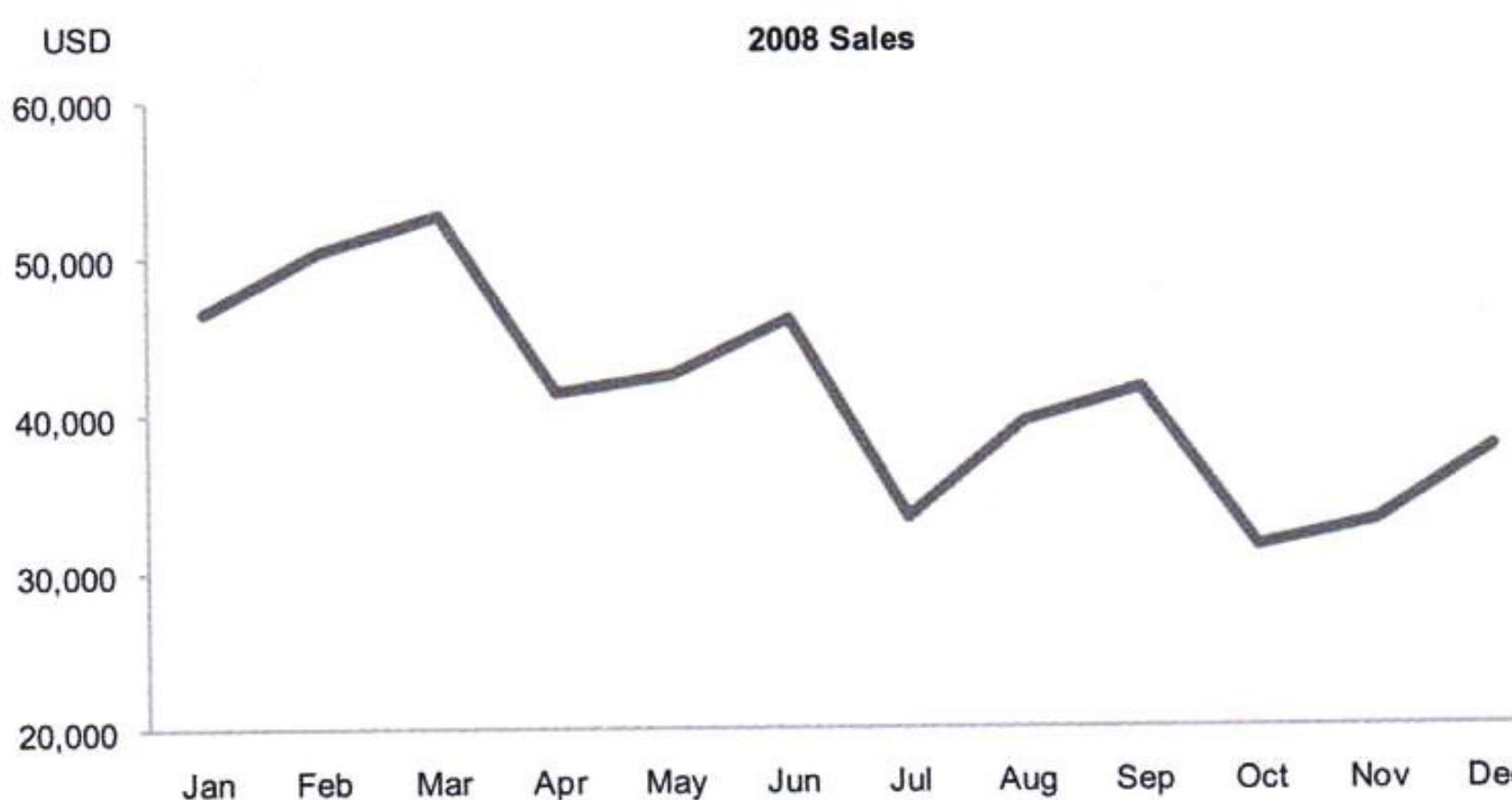
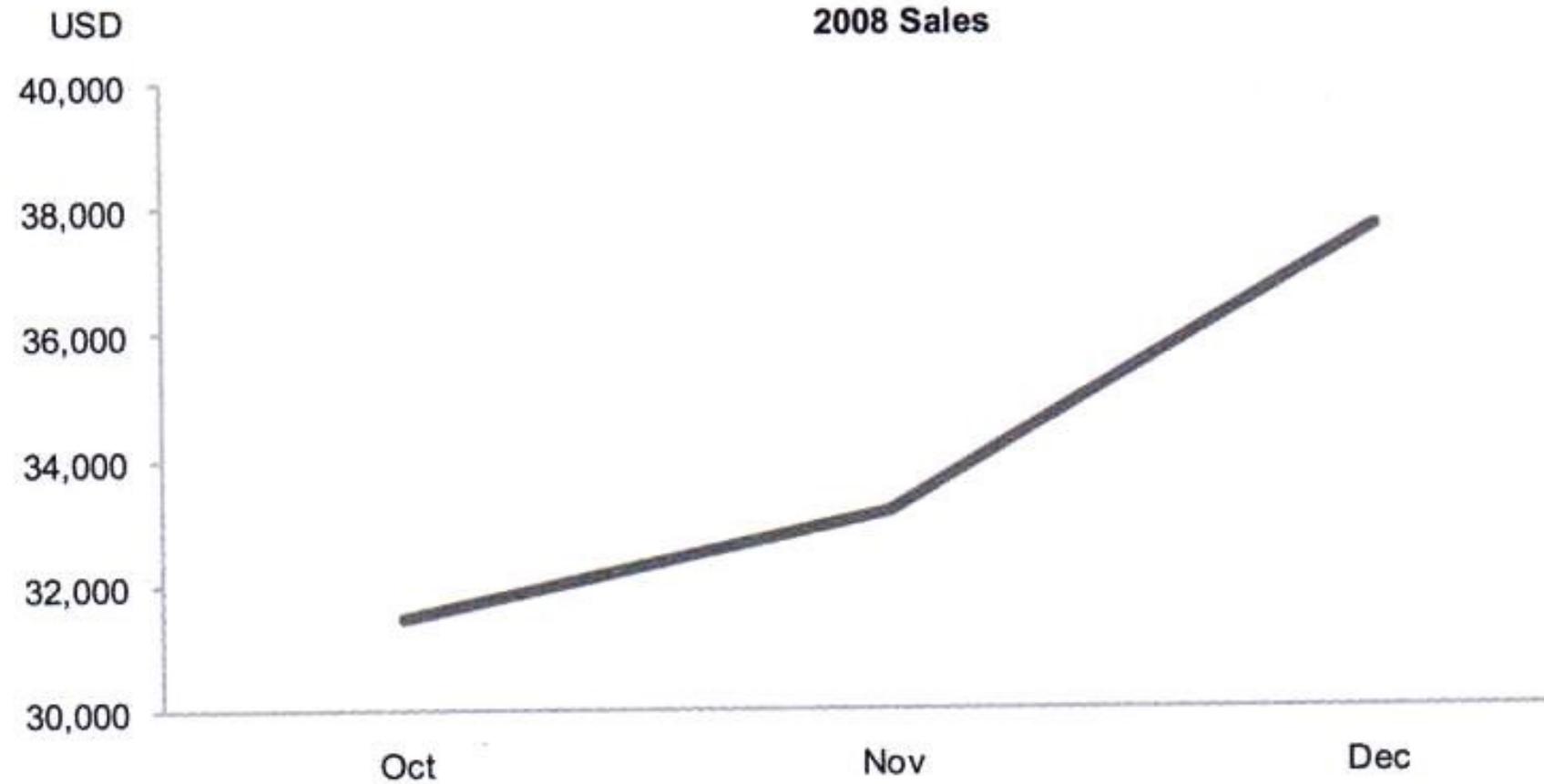
**Pequenos múltiplos** se mostrou a técnica **mais efetiva**

# BOAS PRÁTICAS

---

# ANALISE CADA PERÍODO EM SEU CONTEXTO

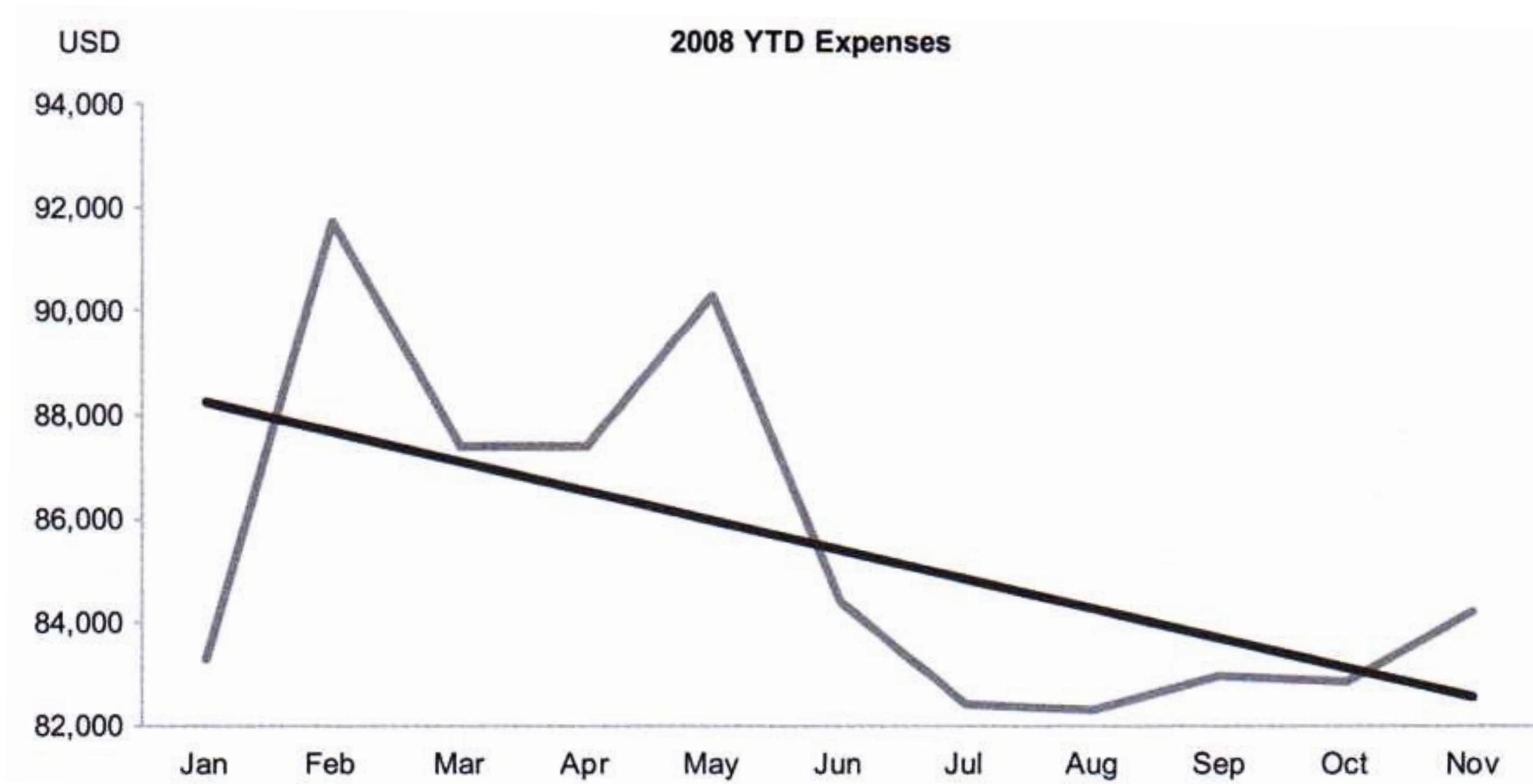
---

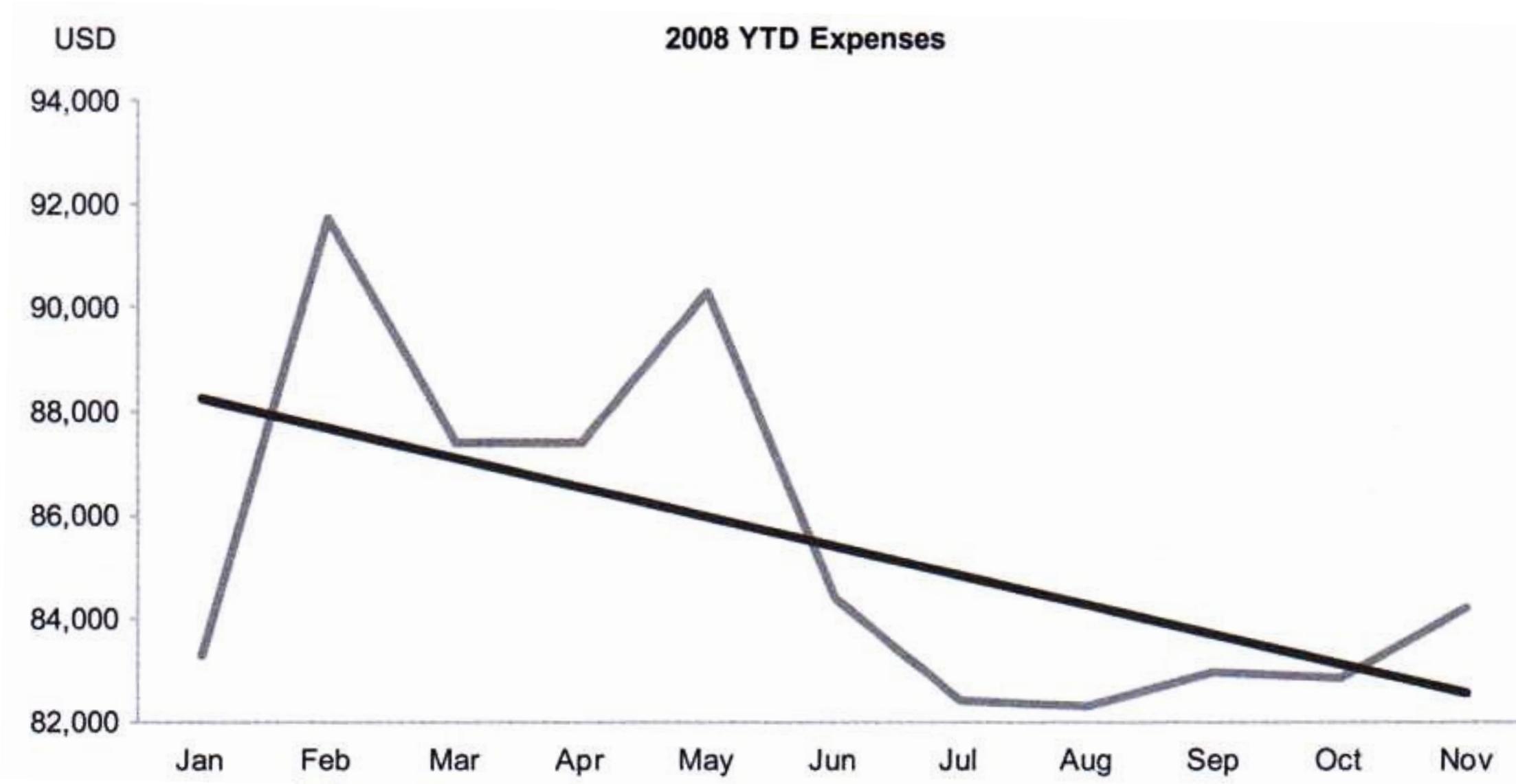


- Quando examinamos pequenos intervalos de tempo, corremos o risco de que os padrões observados pareçam mais representativos do que realmente são

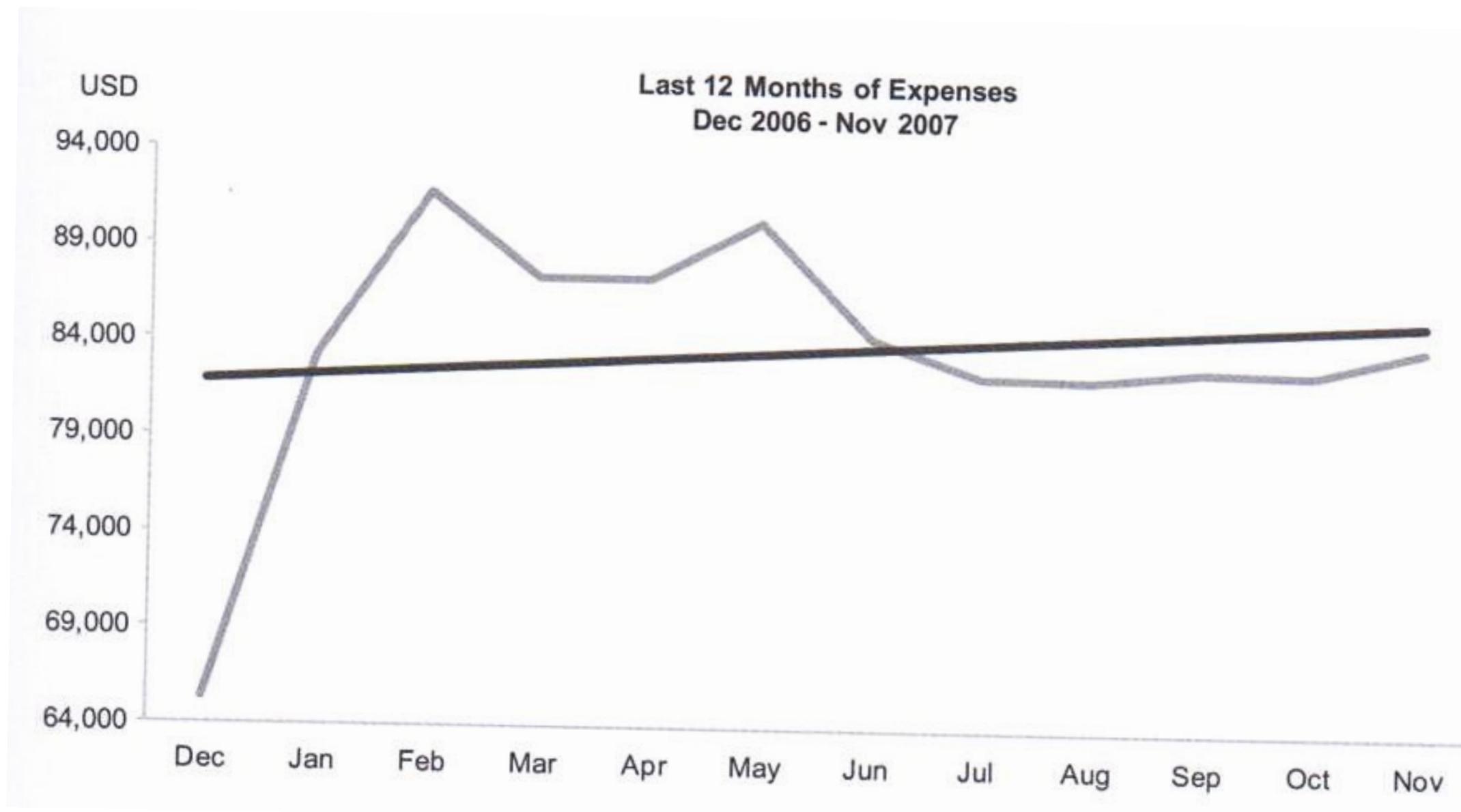
# COMPARE COM VALORES MÉDIOS

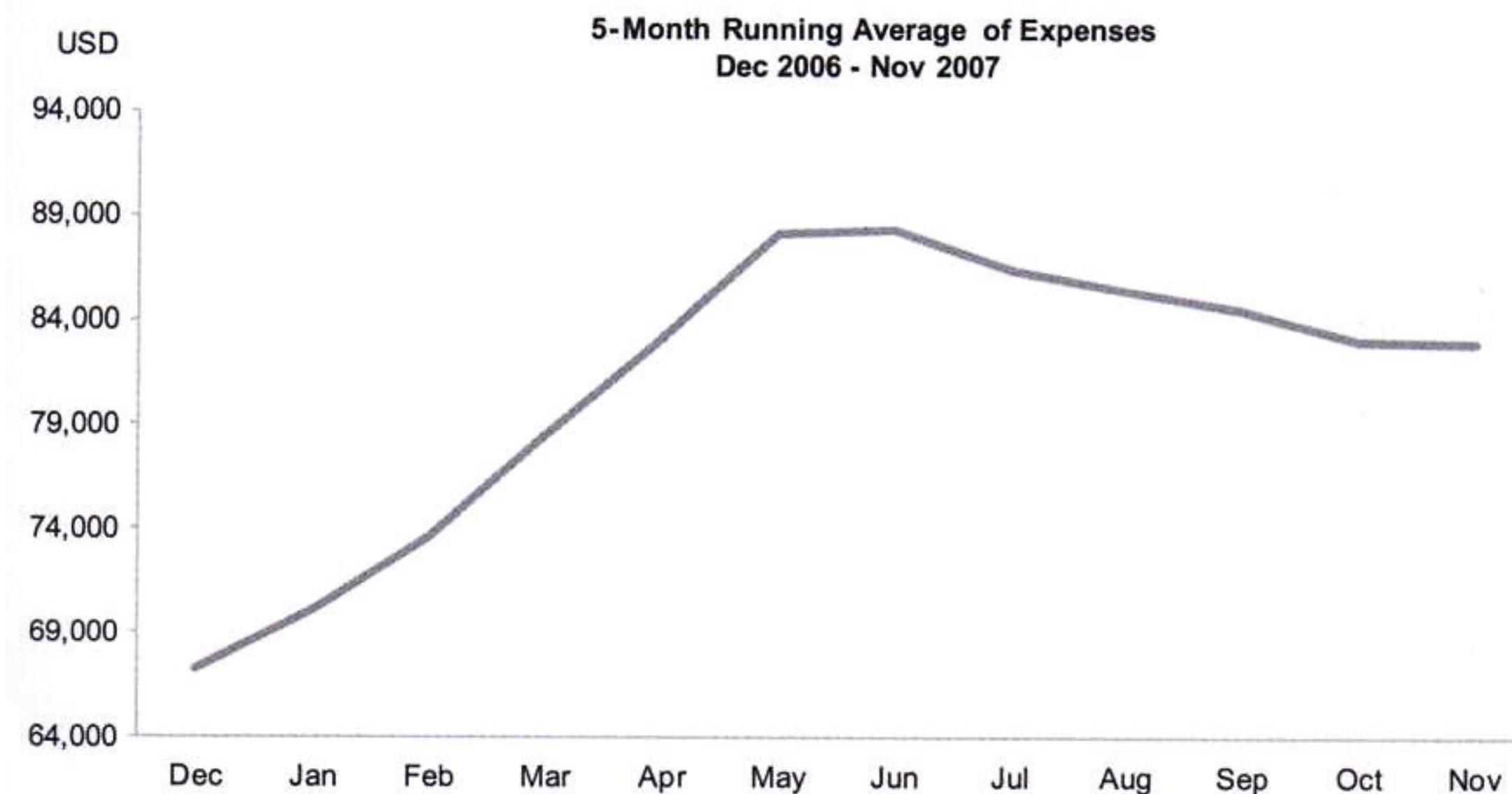
---



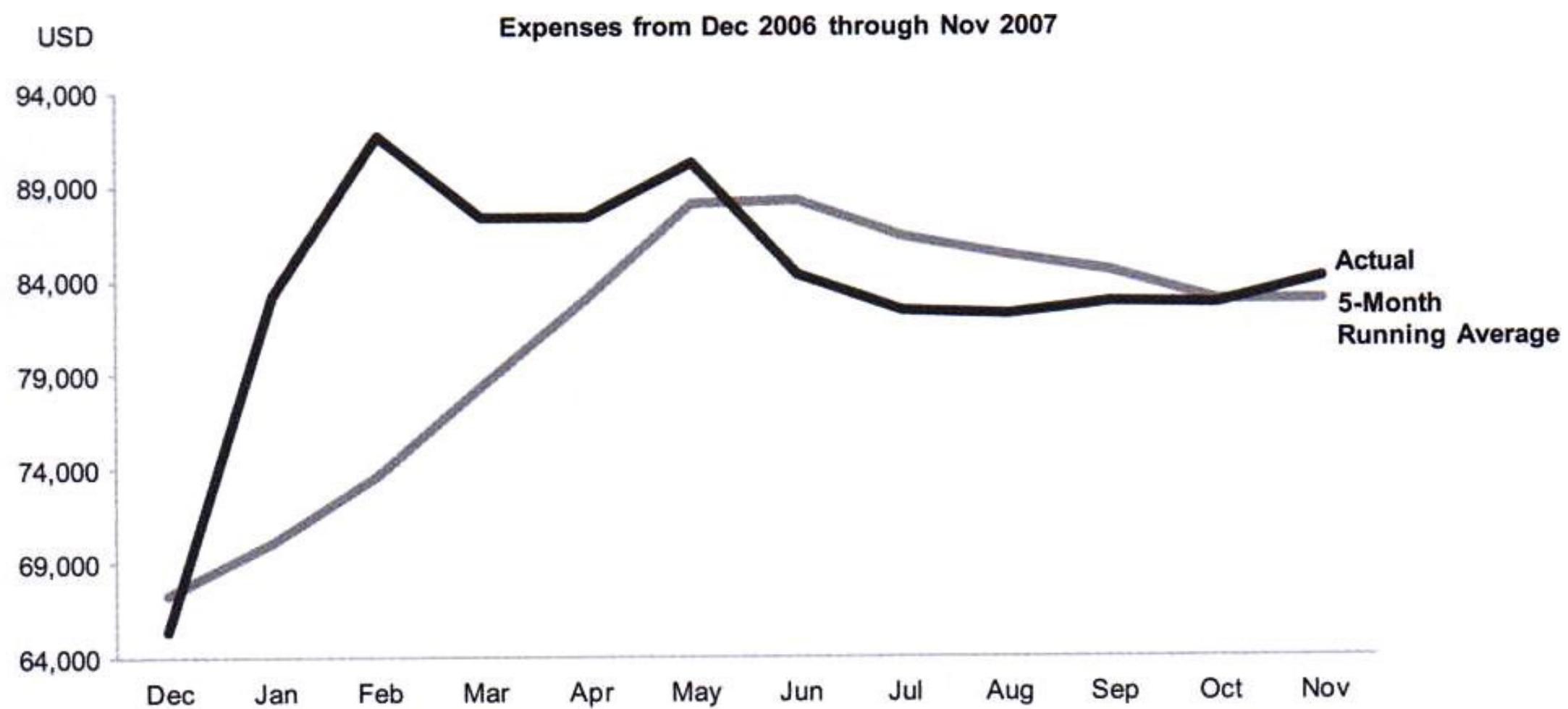


Examine sempre alguns valores fora do intervalo para se certificar de que não está visualizando um intervalo com **tendência enviesada**

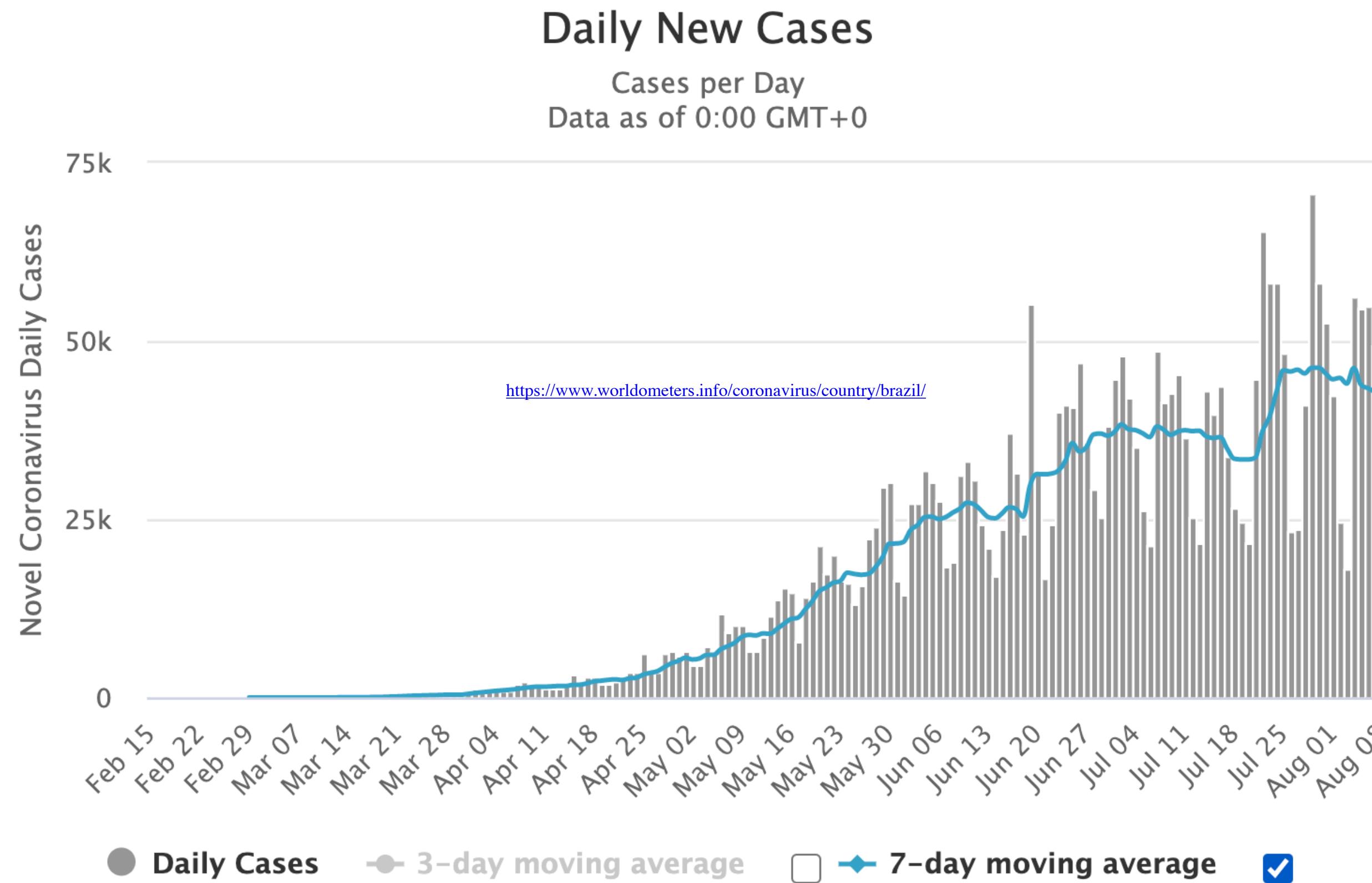




Em alguns casos, pode ser interessante suavizar o gráficos usando a **média móvel** (média de cada valor com os  $n$  valores anteriores)



# Daily New Cases in Brazil

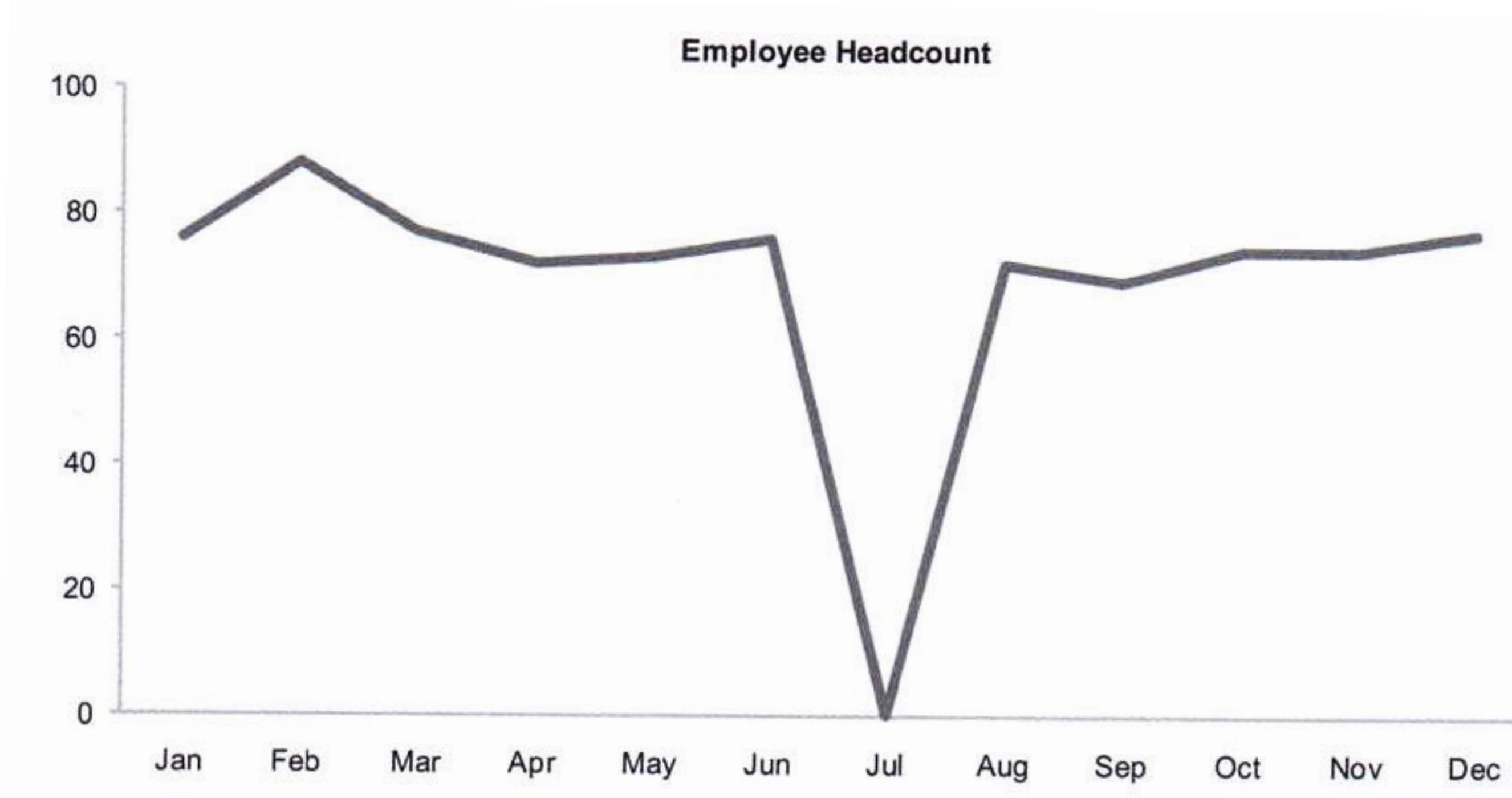


Em alguns casos, pode ser interessante suavizar o gráficos usando a **média móvel** (média de cada valor com os  $n$  valores anteriores)

<https://www.worldometers.info/coronavirus/country/brazil/>

# REPRESENTE ADEQUADAMENTE VALORES AUSENTES

---



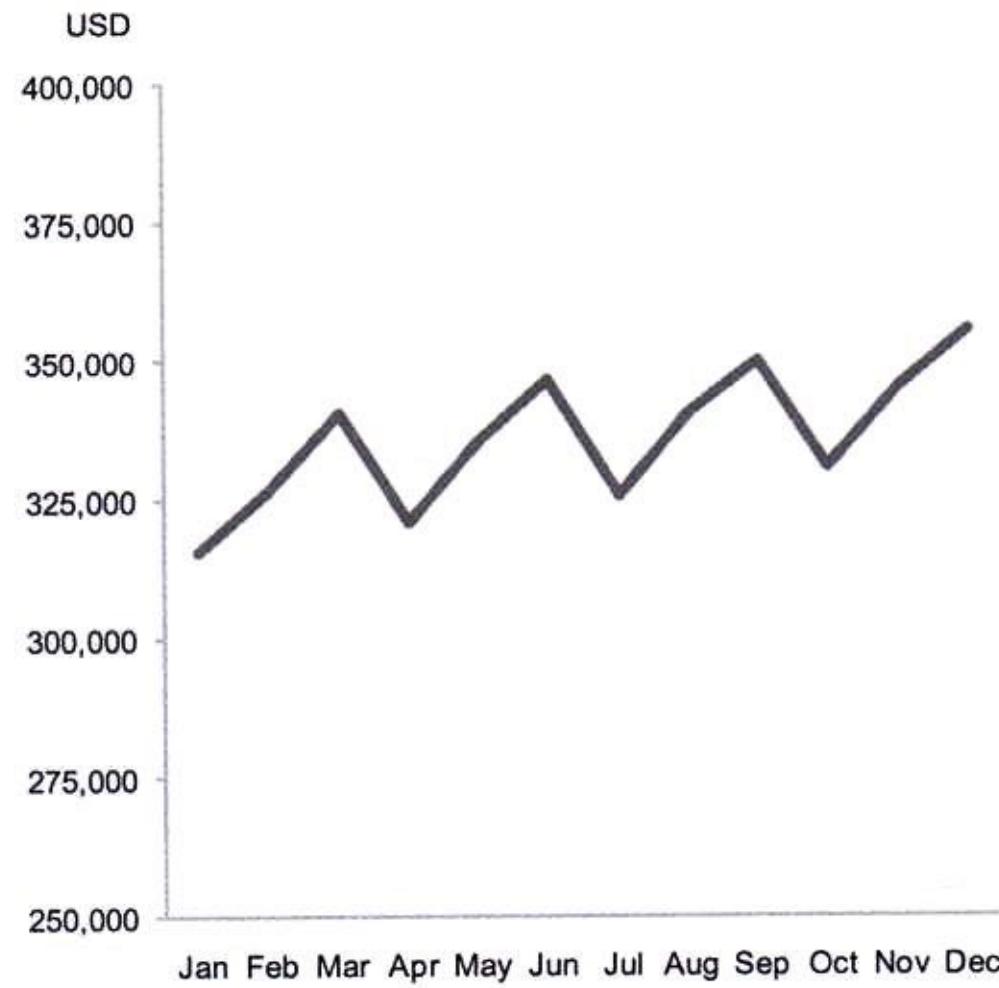
# REPRESENTE ADEQUADAMENTE VALORES AUSENTES

---



# OTIMIZE A RAZÃO DO ASPECTO

---



1:1

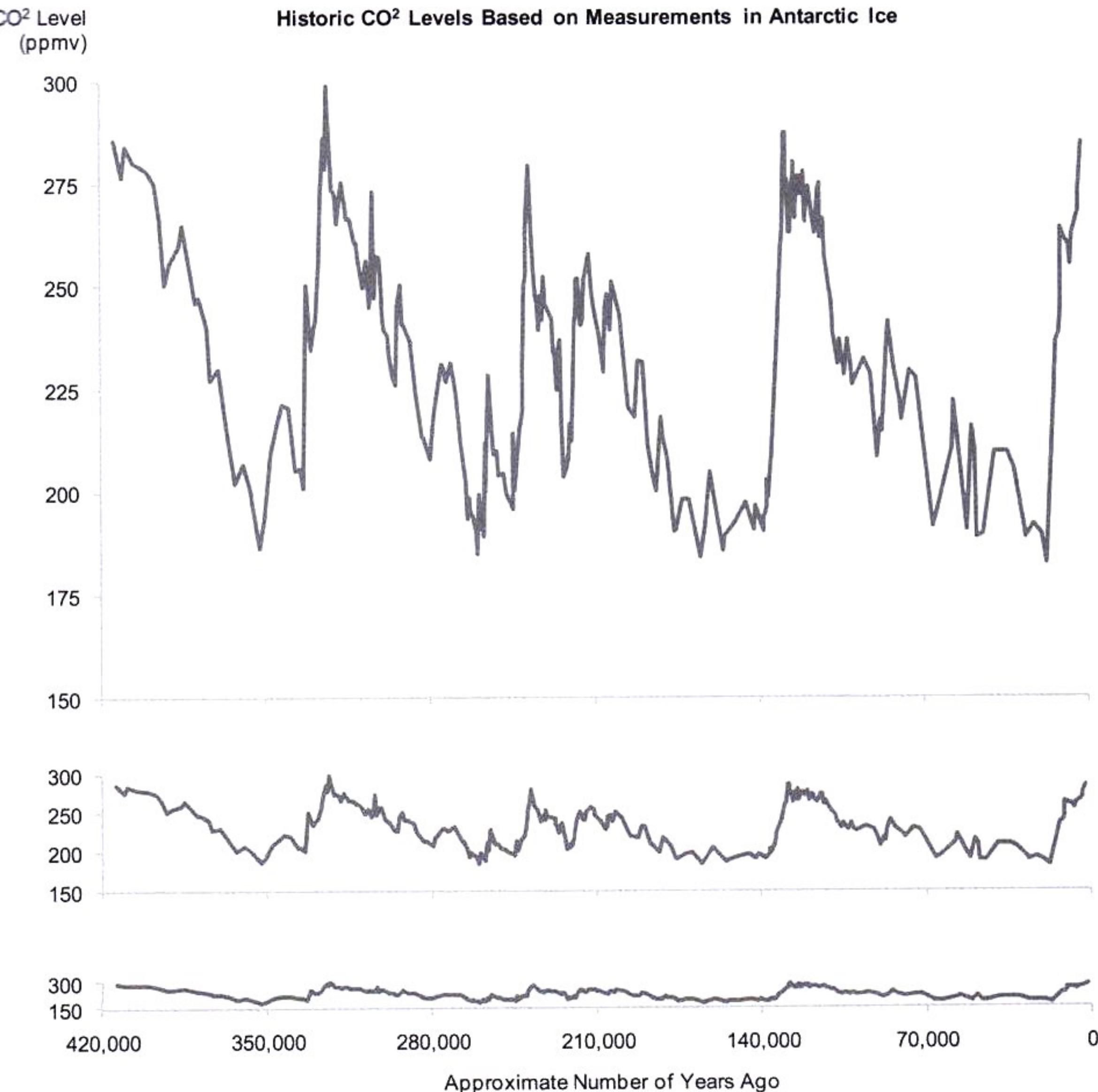
- A razão do aspecto (*aspect ratio*) é a razão da largura pela altura de um gráficos



2:1

# OTIMIZE A RAZÃO DO ASPECTO

---



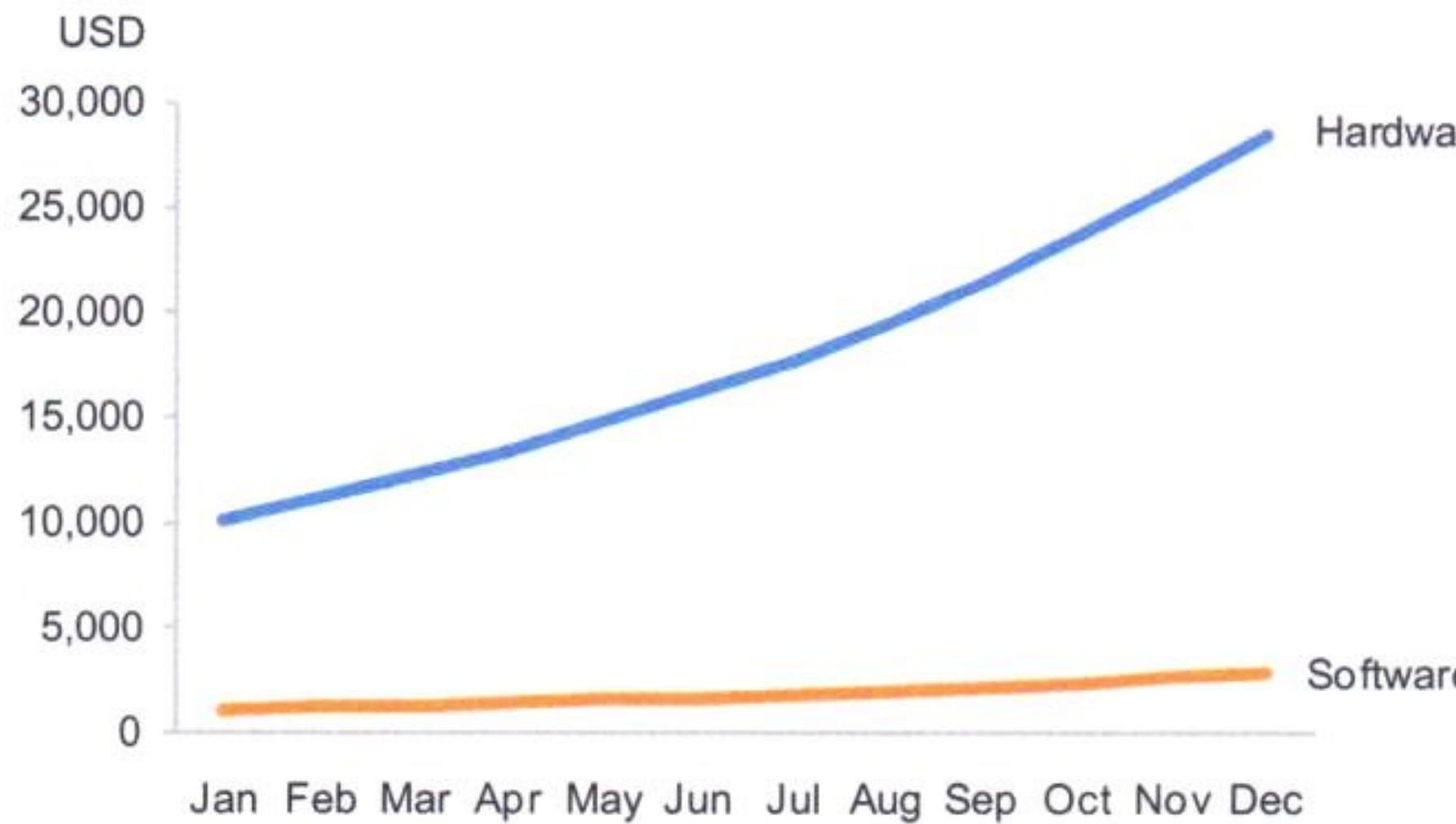
Segundo Cleveland, **45° é uma boa angulação média para os segmentos de reta que compõem gráficos de linhas**

S. Cleveland, The Elements of Graphing Data  
Hobart Press, 1994

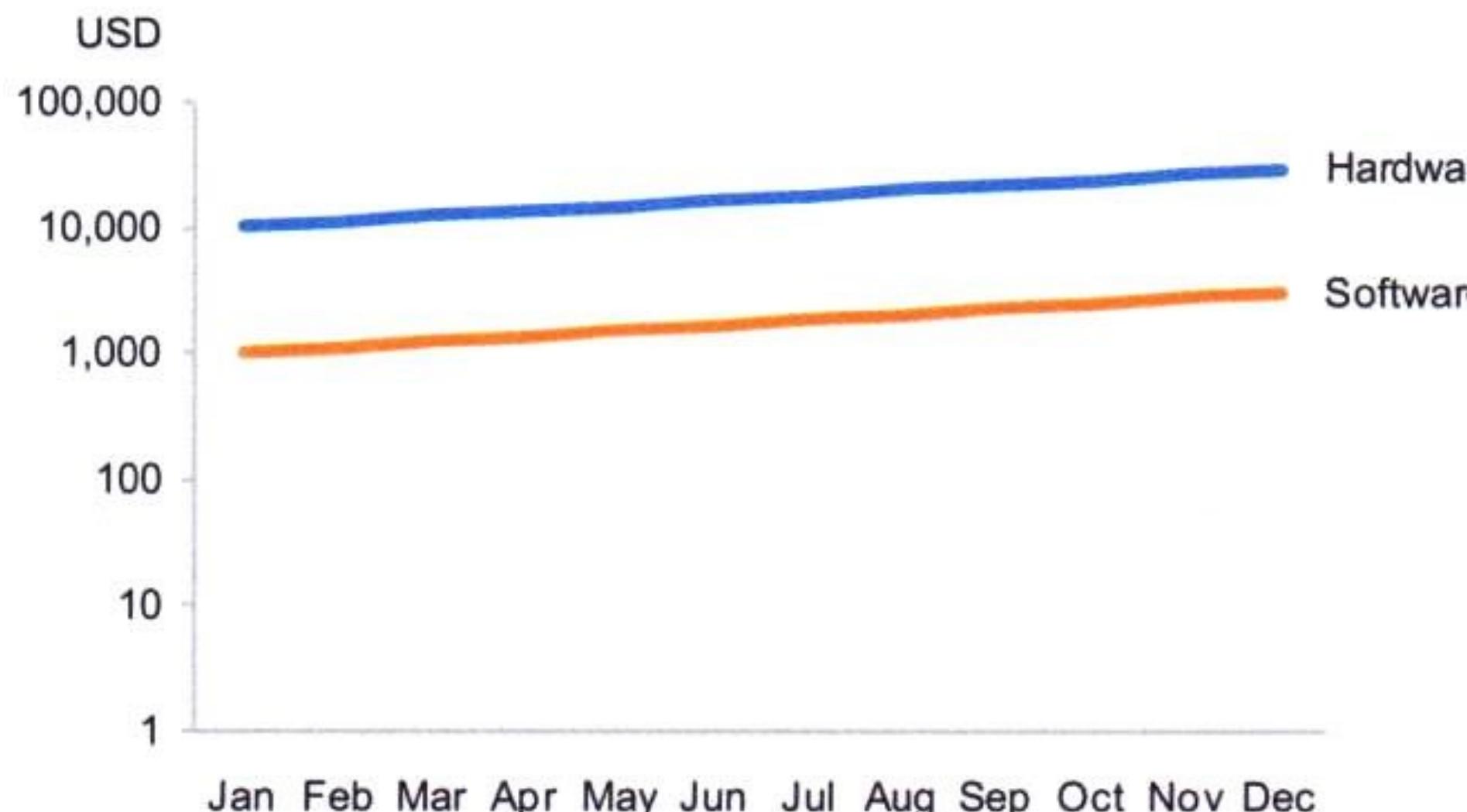
J. Heer e M. Agrawala, *Multi-scale banking to 45°*, IEEE Transactions on visualization and computer graphics, 12:5, 2006.

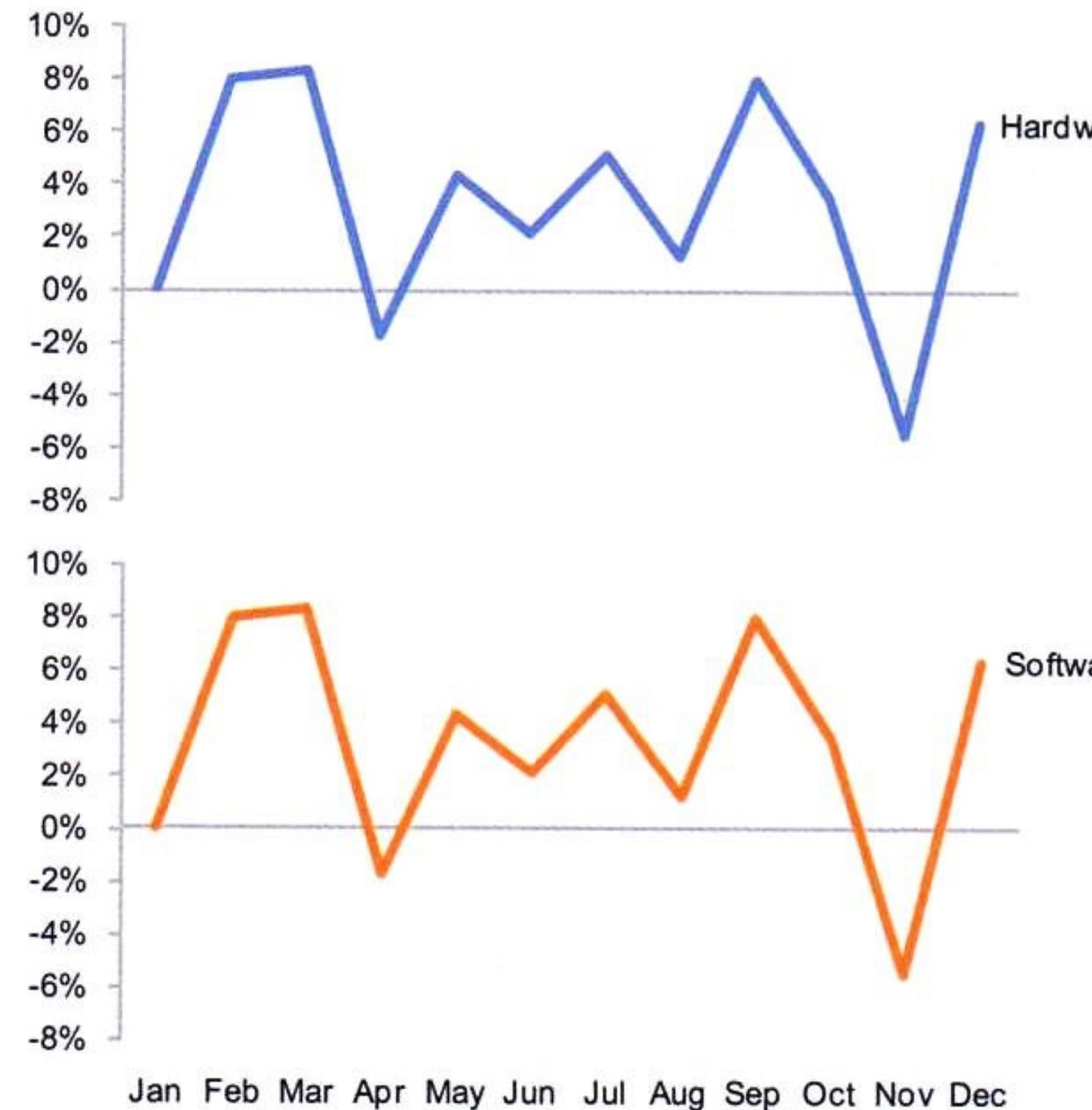
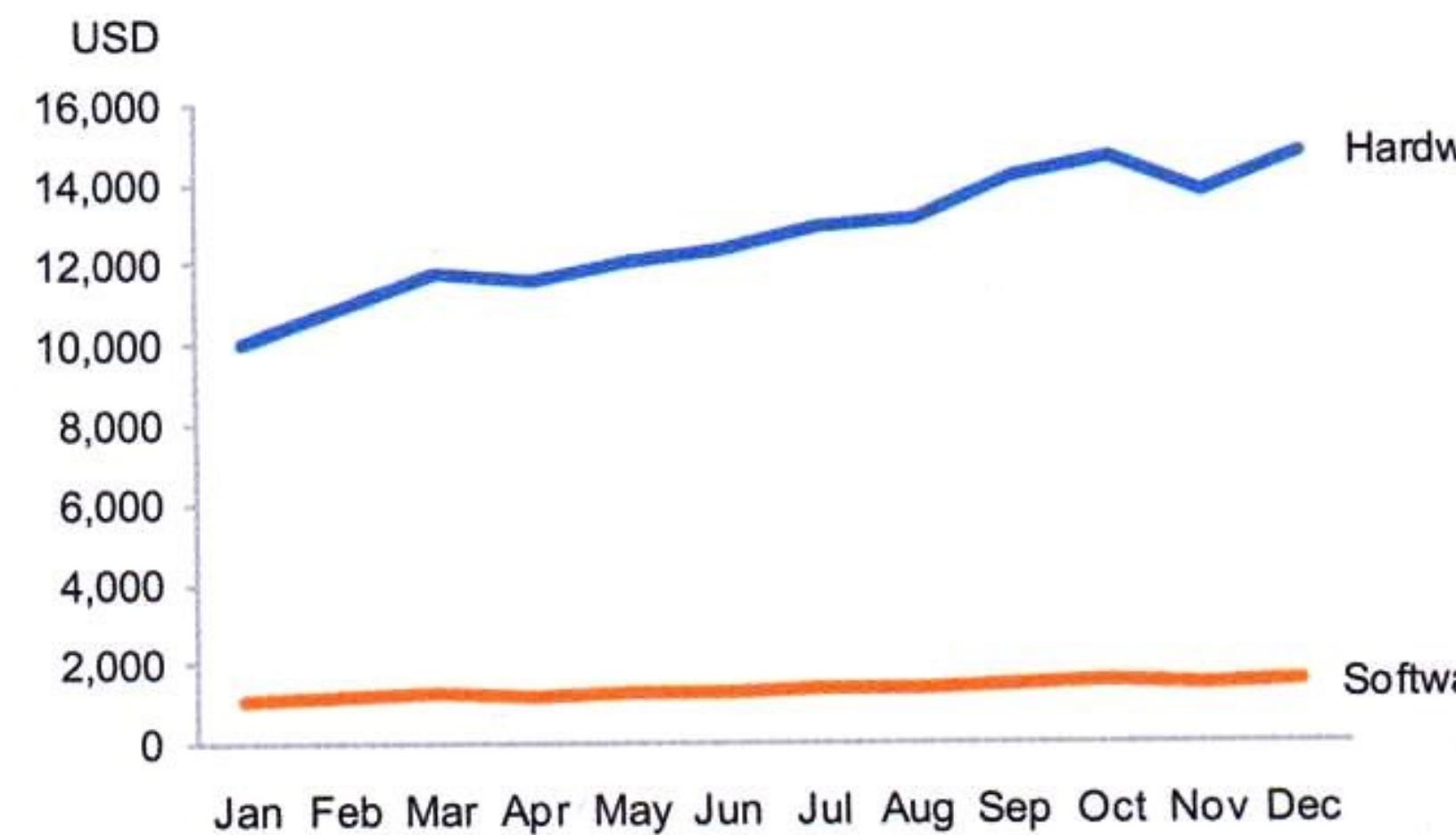
# USE ESCALAS LOGARÍTMICAS PARA COMPARAÇÃO ENTRE TAXAS

---

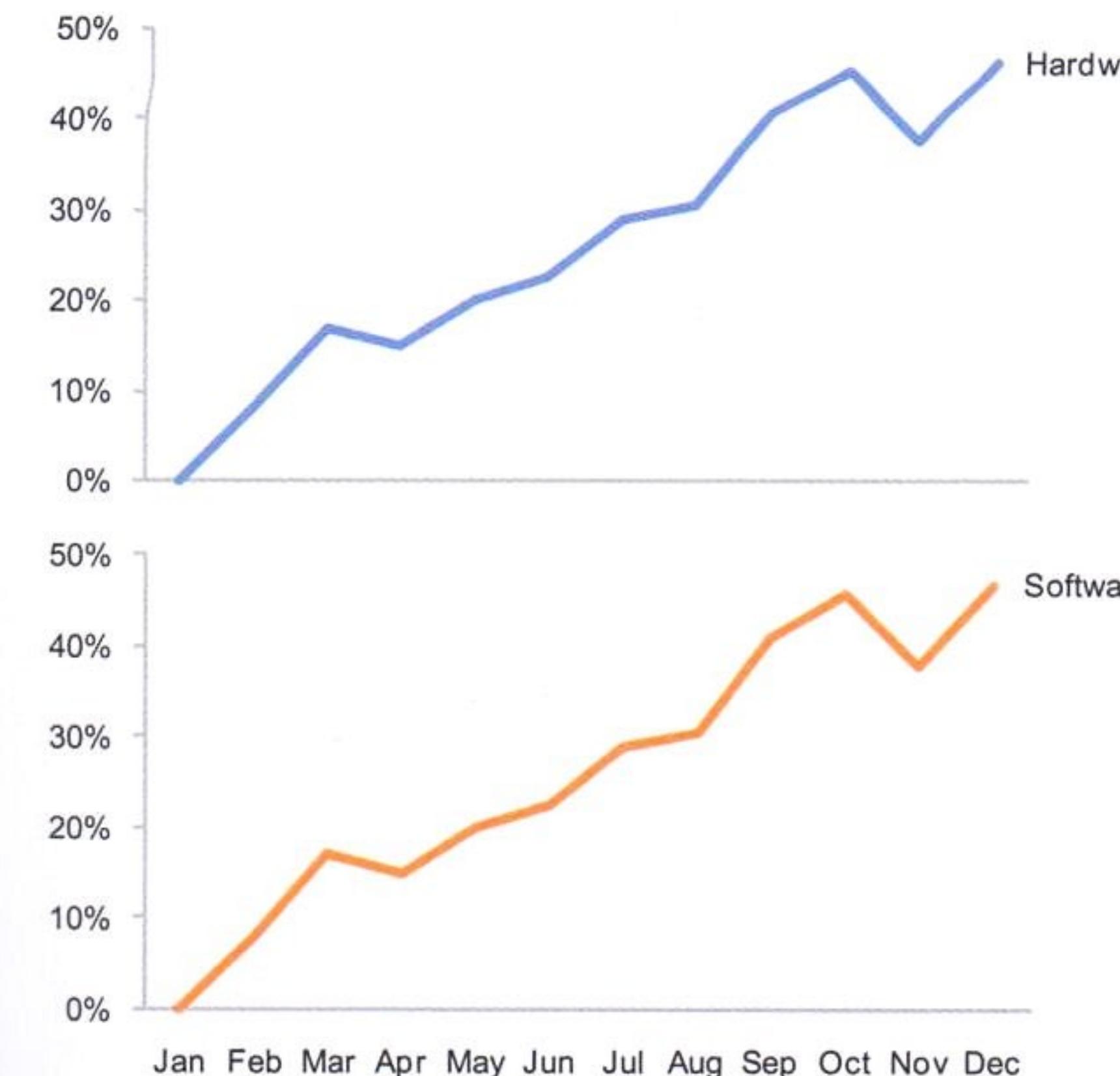
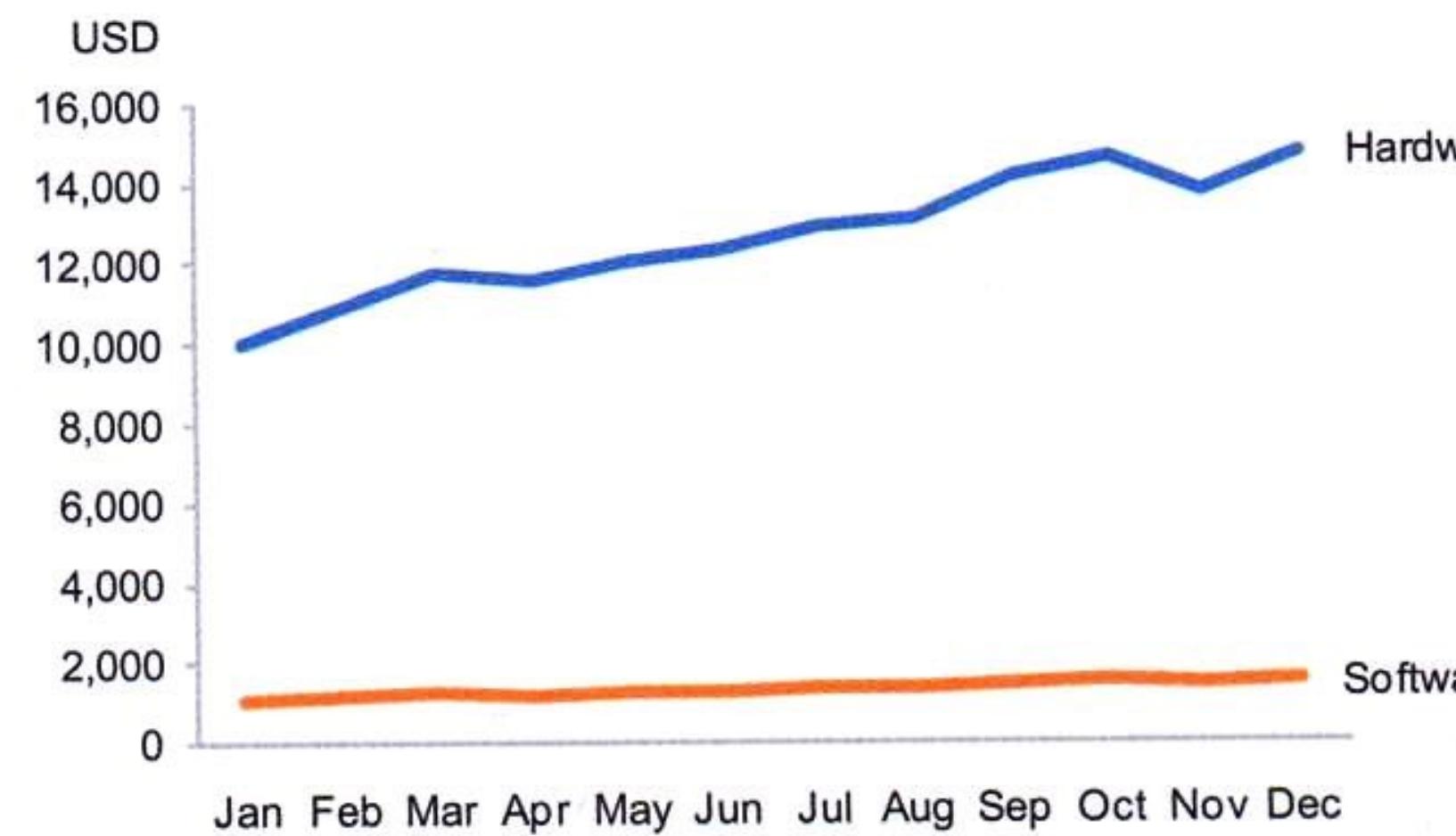


► Neste exemplo, os gráficos apresentam inclinações diferentes em termos dos valores absolutos mas taxas de crescimento próximas





Outra forma de exibição e comparação de taxas é através do **percentual de diferença entre um valor e o próximo**



Ou ainda mostrar o **percentual em relação a um mês específico**, neste exemplo, janeiro

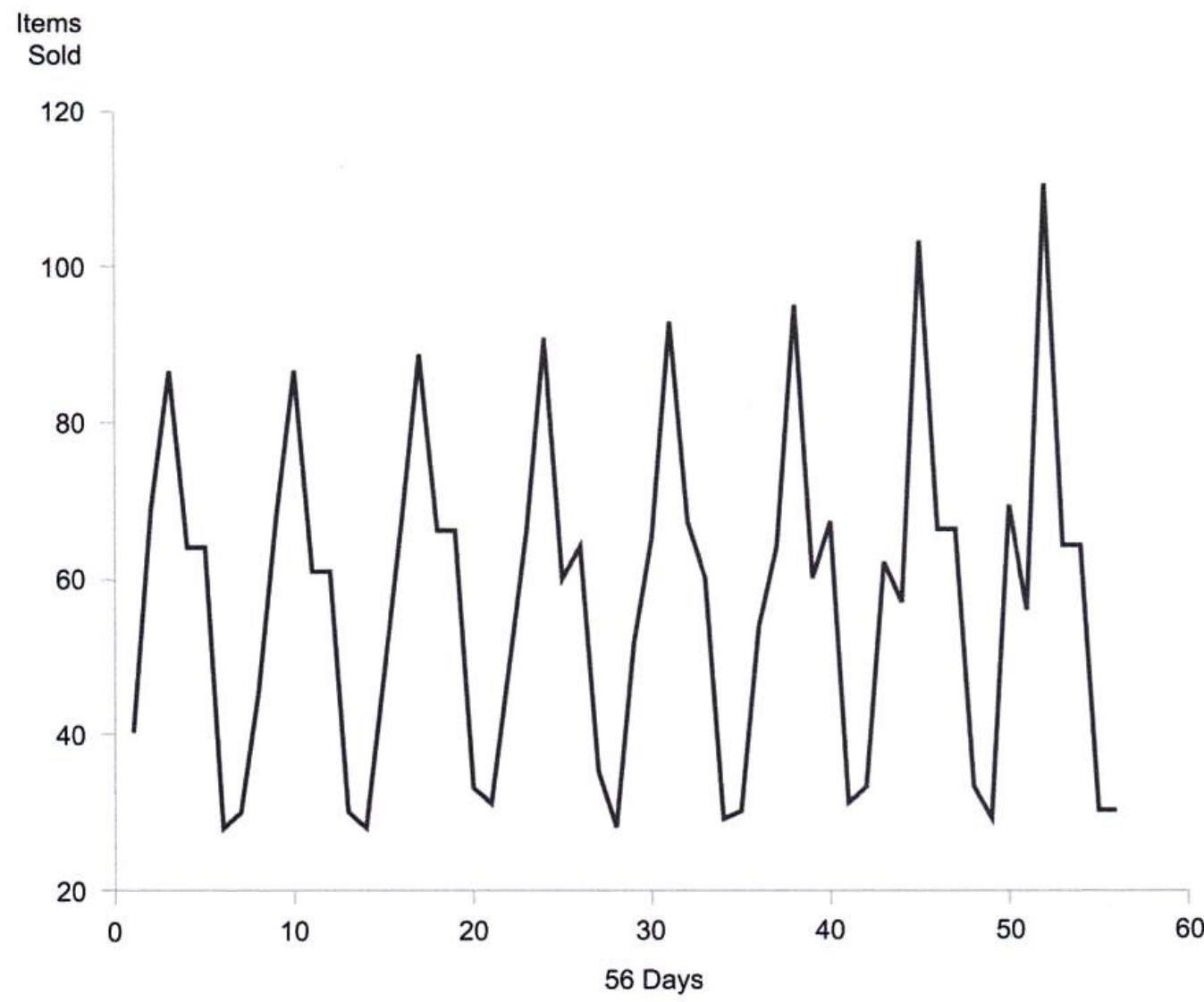
# SOBREPONHA ESCALAS PARA COMPARAR PADRÕES CÍCLICOS

---



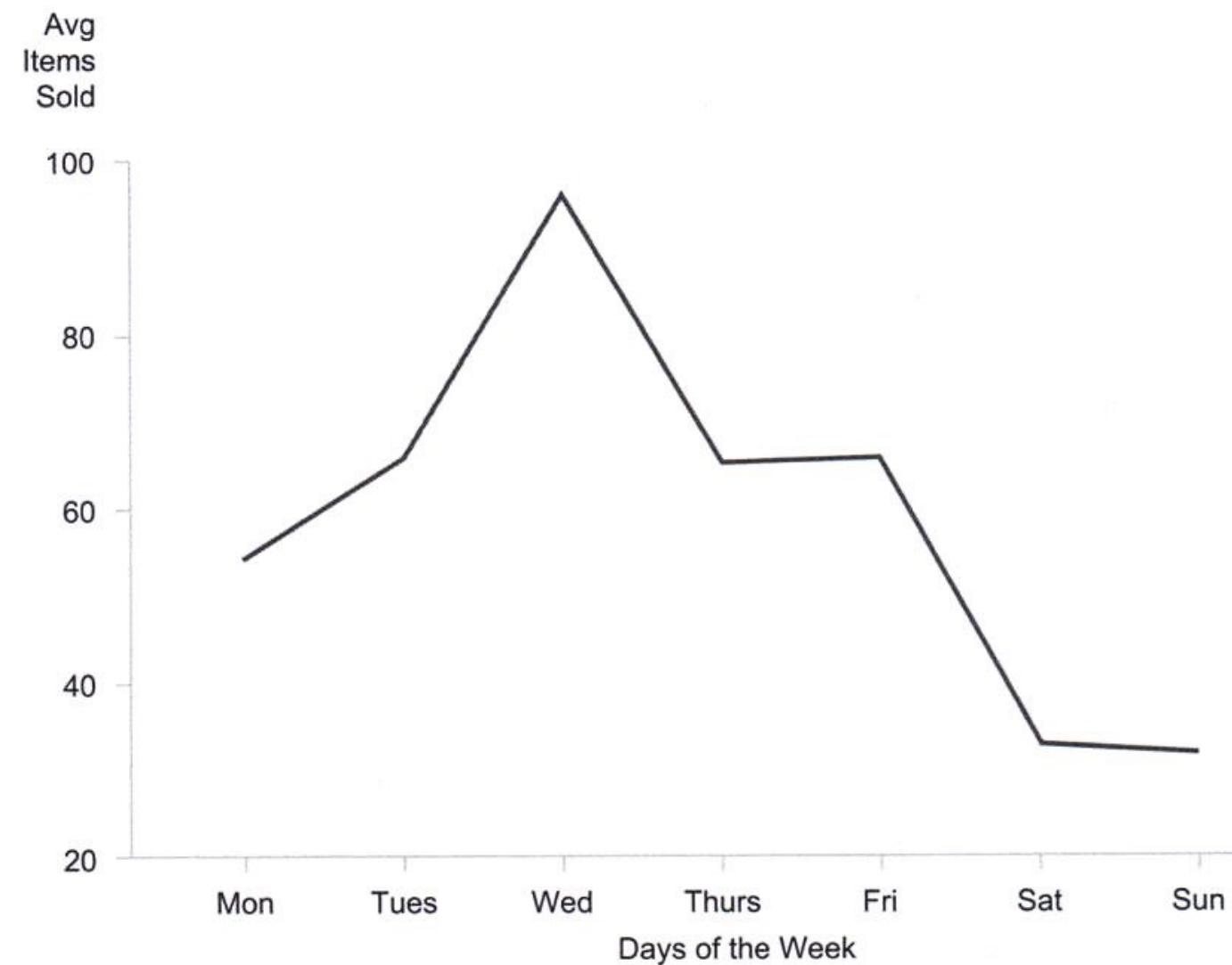
# CYCLEPLOTS

---



- O primeiro gráfico mostra as vendas em 8 semanas

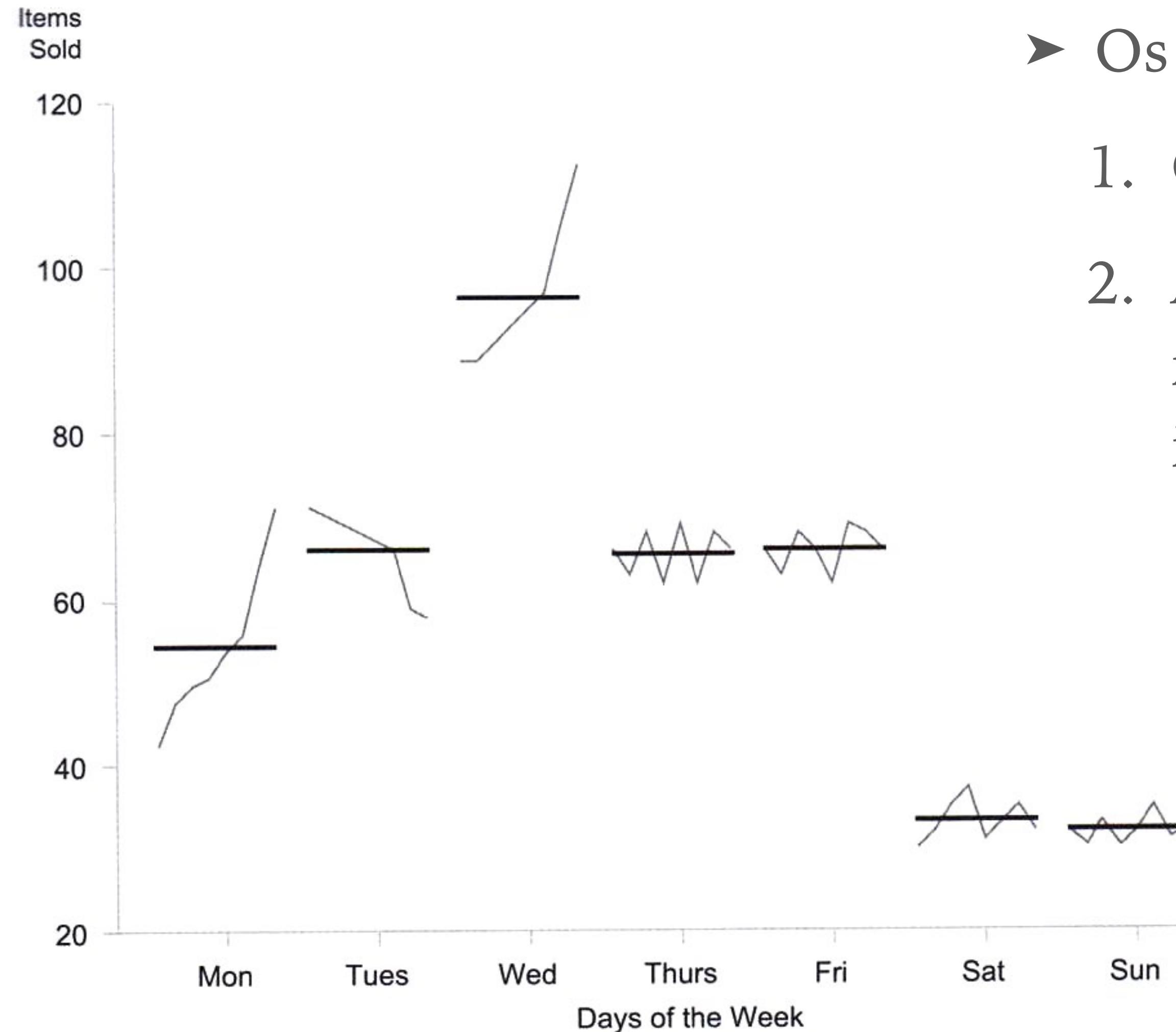
E a variação de semana para semana?



- O segundo, mostra a média dos valores por dia nestas 8 semanas

# CYCLEPLOTS

---

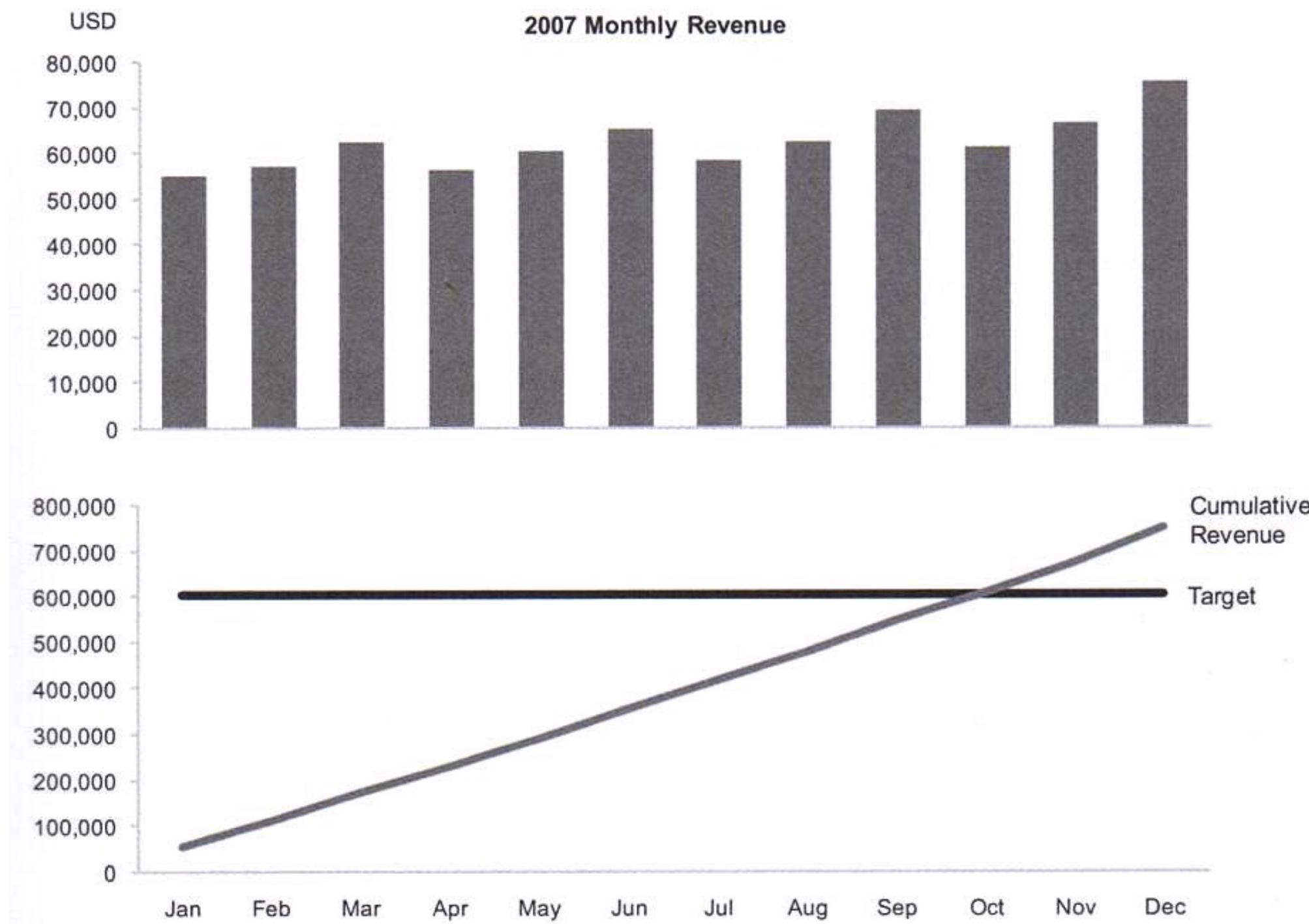


- Os cycle plots mostram:
  1. O padrão no ciclo inteiro
  2. A tendência de cada ponto no ciclo em todo o intervalo de tempo

# COMPARE COM VALORES CUMULATIVOS

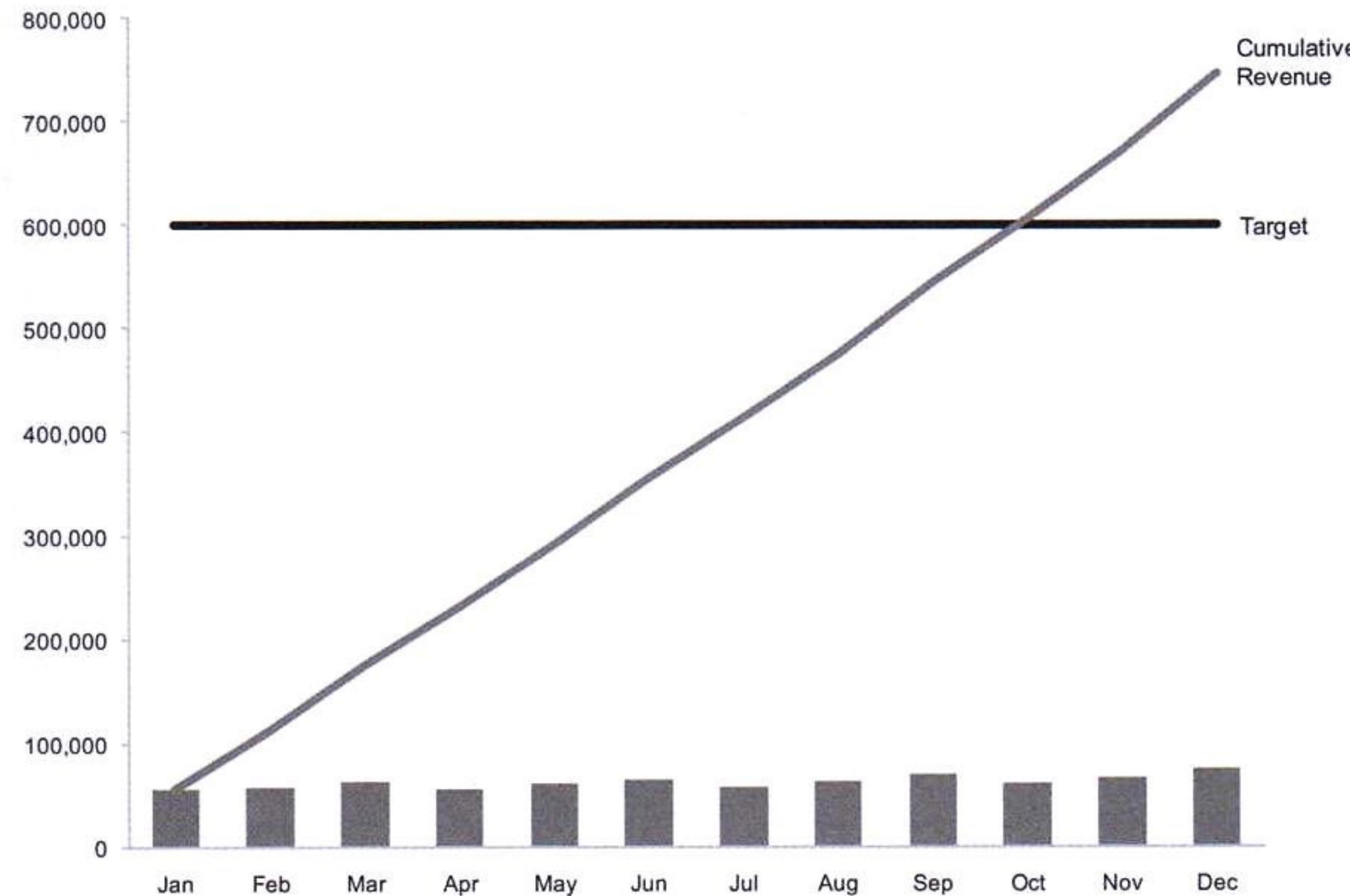
---

- A combinação do gráfico de barras com linhas pode ser usado para mostrar como os valores reais se comparam a valores alvo ou mesmo a médias e valores acumulados

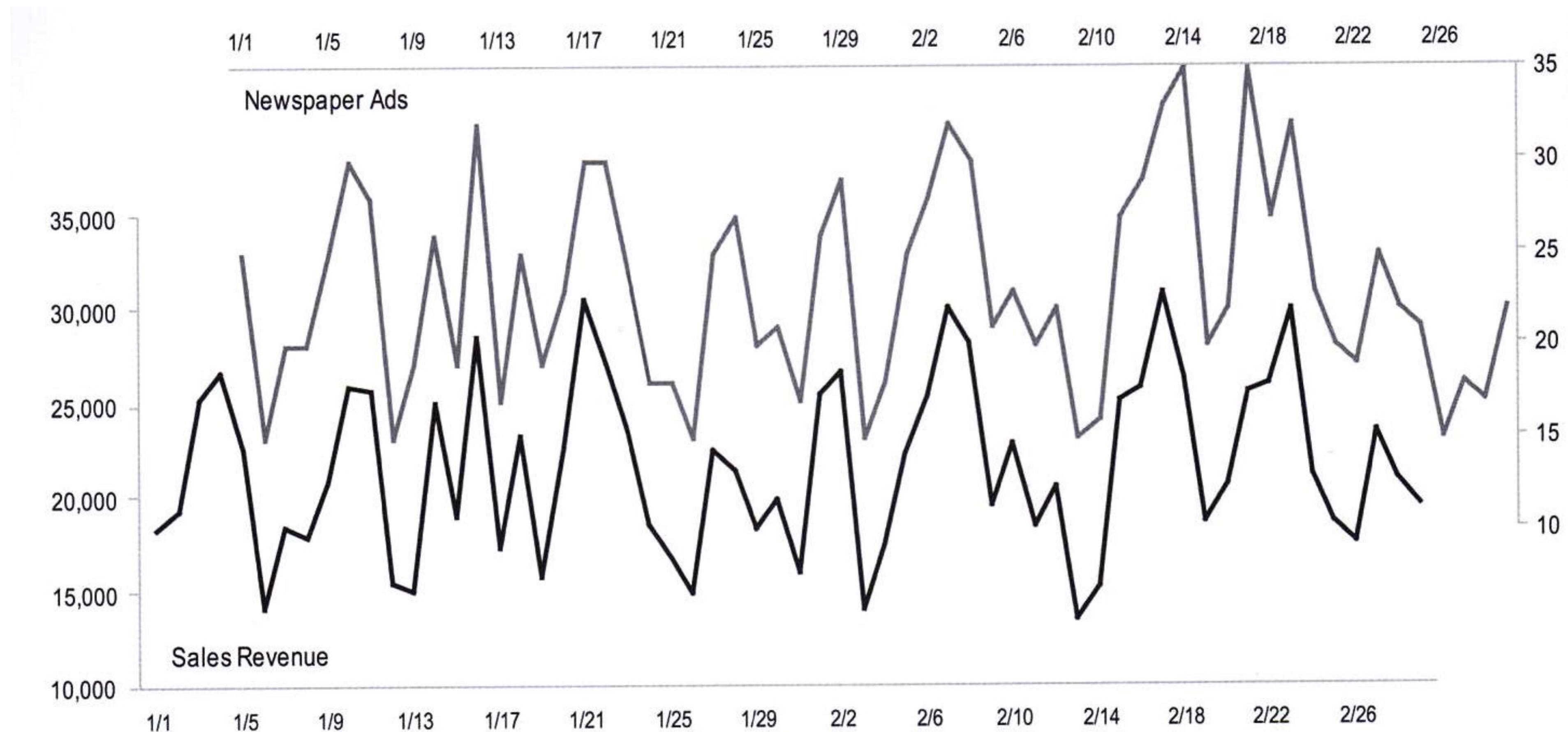


USD

### 2007 Monthly Revenue Compared to YTD Target

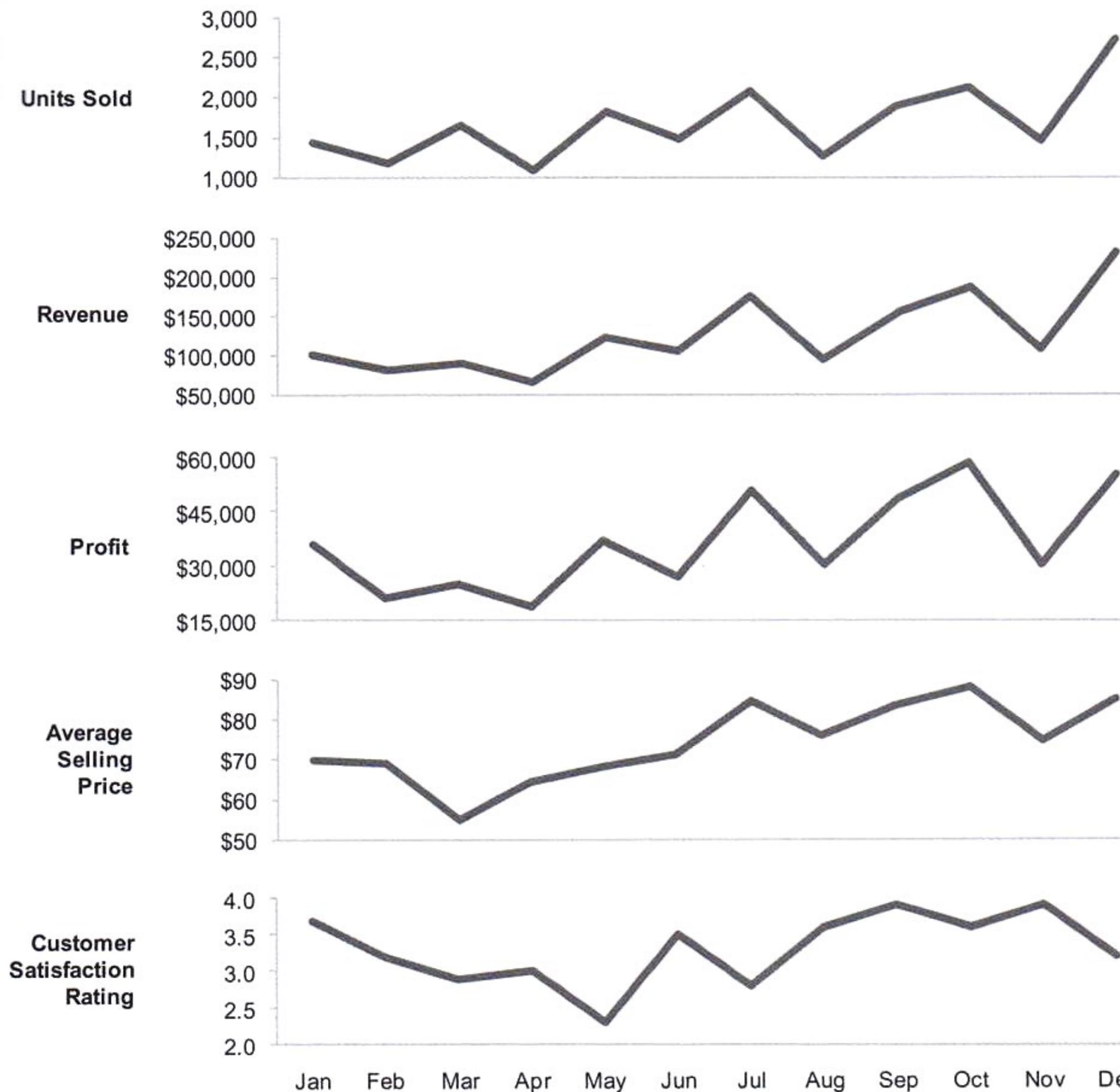


# DESLOQUE OS EIXOS PARA FAVORECER A ANÁLISE DE CORRELAÇÕES



# EMPILHE GRÁFICOS PARA COMPARAR MÚLTIPLAS VARIÁVEIS

---

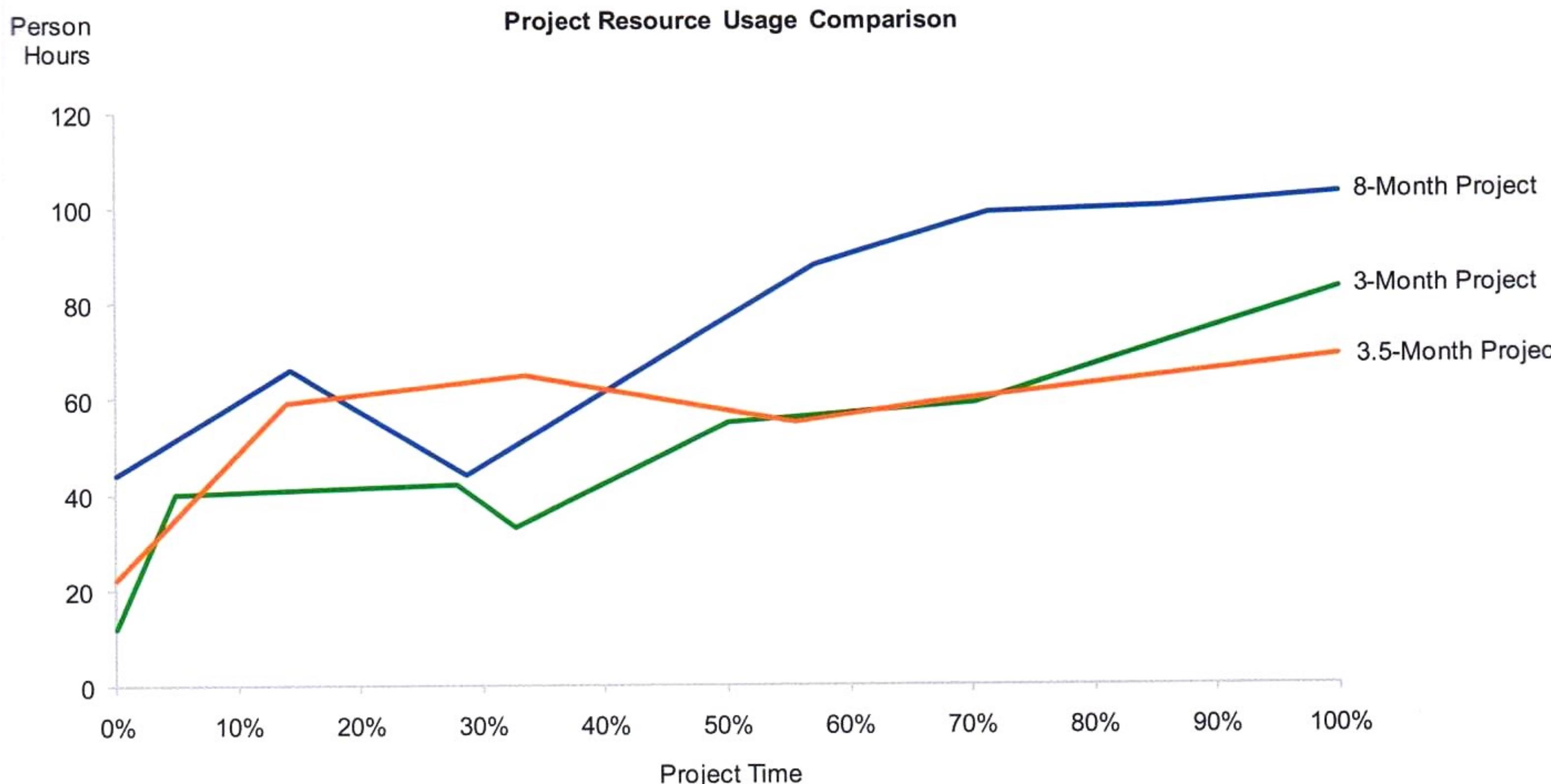


➤ Como as medidas e escalas são diferentes, o empilhamento dos gráficos é útil para permitir a análise de padrões de mudança

# EXPRESSE O TEMPO COMO PERCENTUAL PARA COMPARAÇÃO DE PROCESSOS ASSÍNCRONOS

---

- Quando expressamos o tempo em escala de 0-100%, podemos comparar processos independentemente de quando começaram, de sua duração e término



# Séries temporais - Parte II

---

**Profa. Dra. Raquel Minardi**

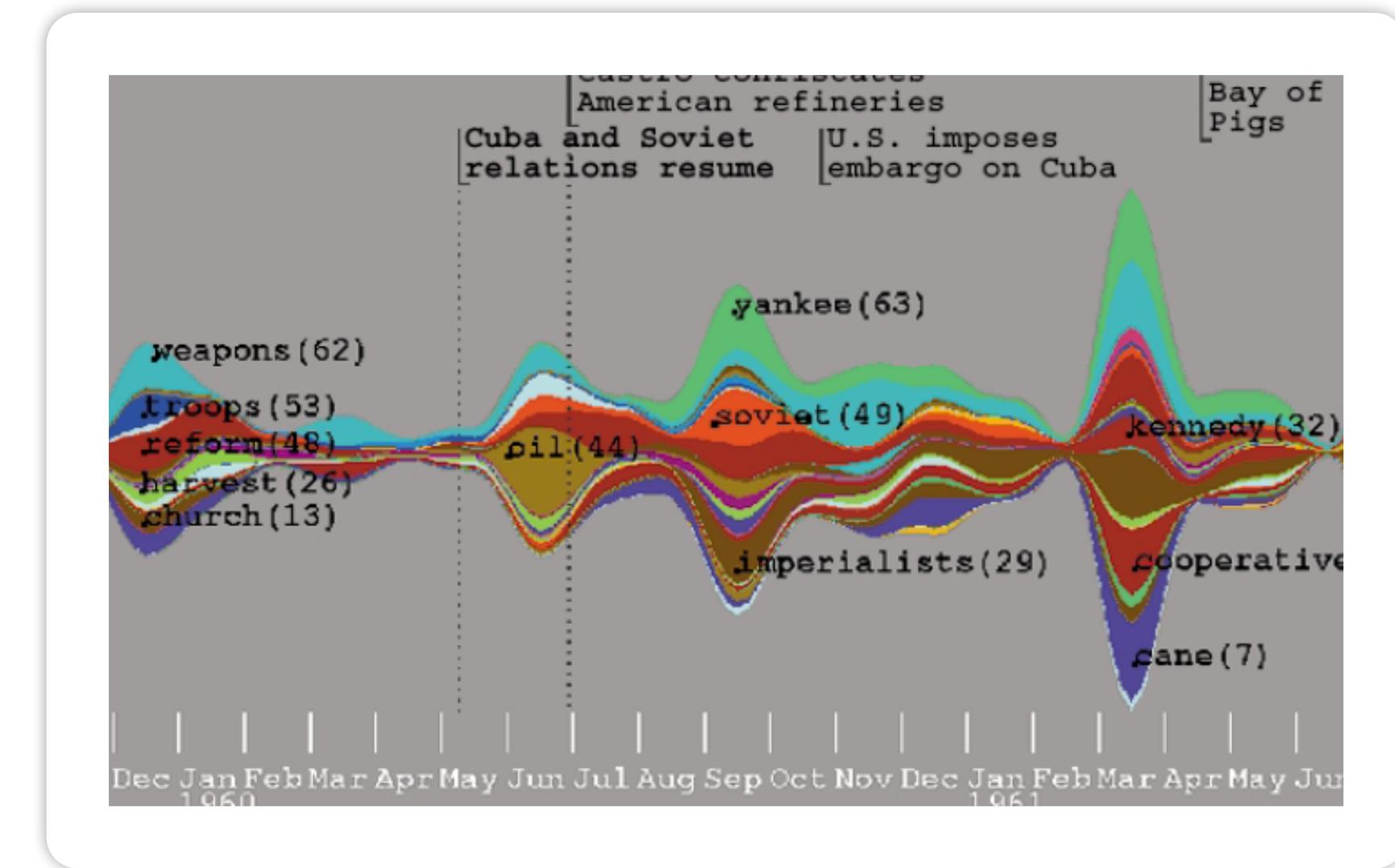
Departamento de Ciência da Computação

Universidade Federal de Minas Gerais

# THEMERIVER: VISUALIZING THEME CHANGES OVER TIME

.....

*S. Havre, B. Hetzler e L. Nowell*  
*IEEE Symposium on Information Visualization*  
*2000*



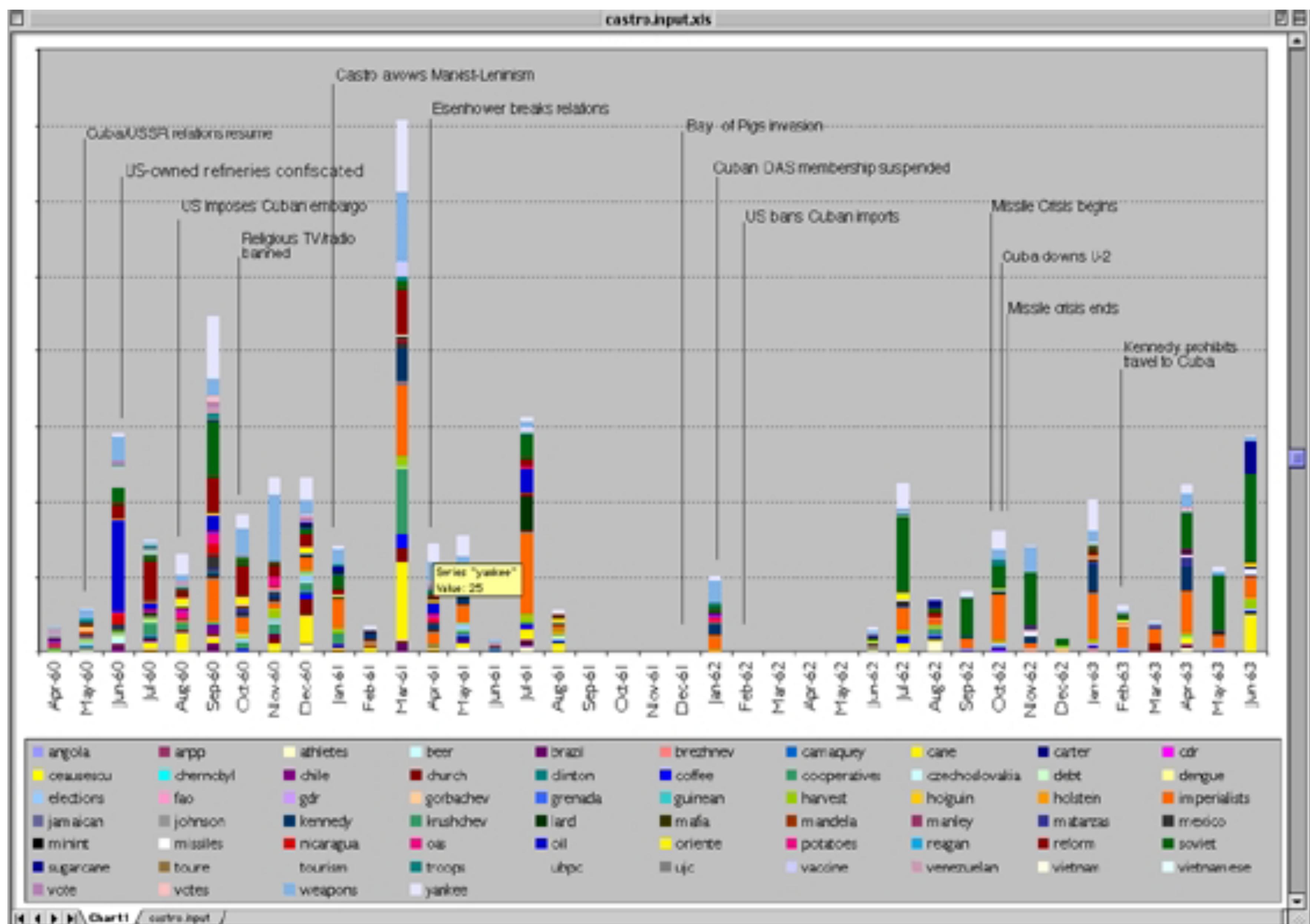


Figure 2: Like ThemeRiver™ in Figure 1, this histogram uses the Castro collection data and depicts changes in thematic content over time.

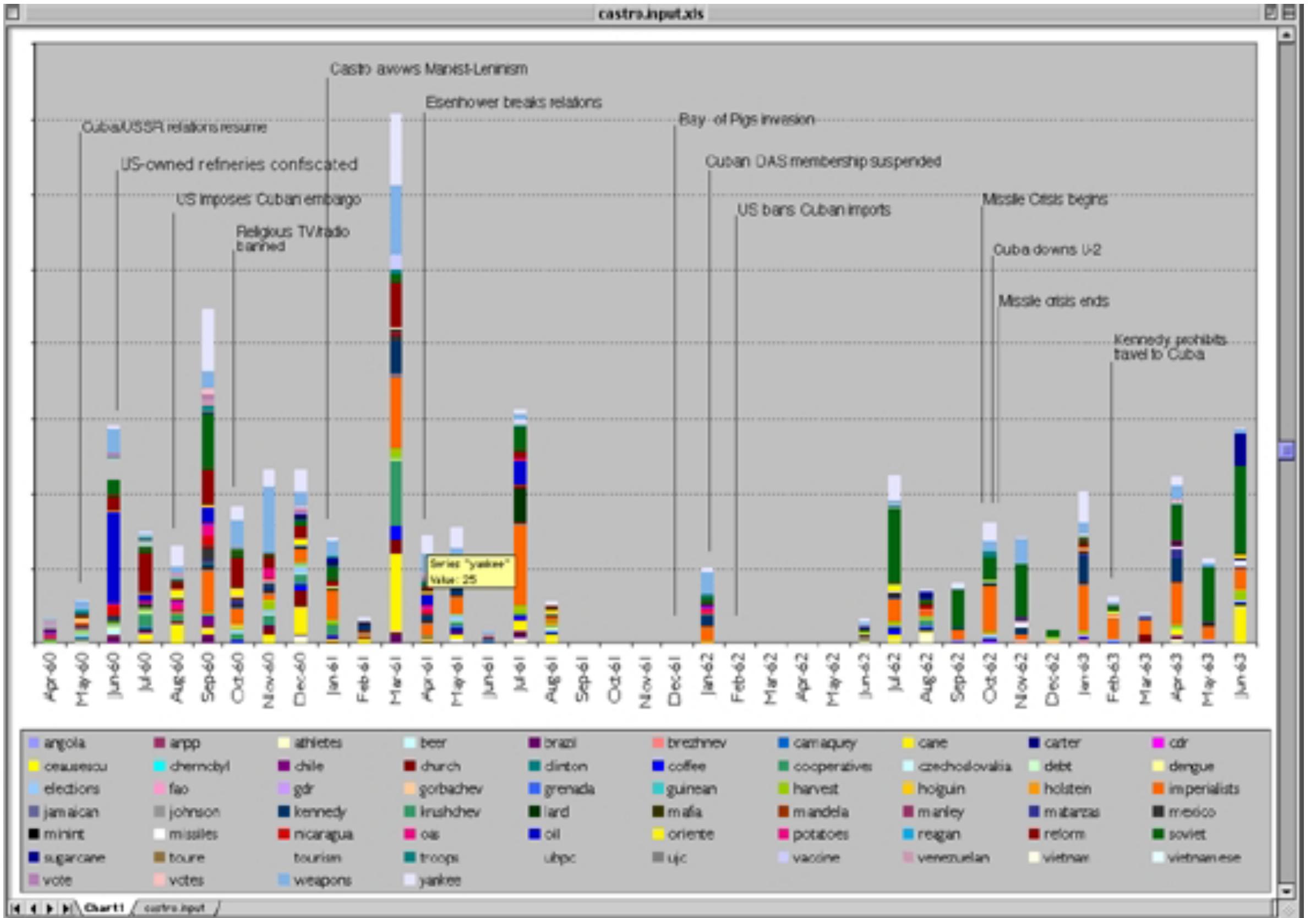


Figure 2: Like ThemeRiver™ in Figure 1, this histogram uses the Castro collection data and depicts changes in thematic content over time.

- Como os temas tratados nos textos de um determinado autor mudam com o tempo?
- Themeriver provê uma visão macro das mudanças temáticas em corpos de documentos na dimensão tempo

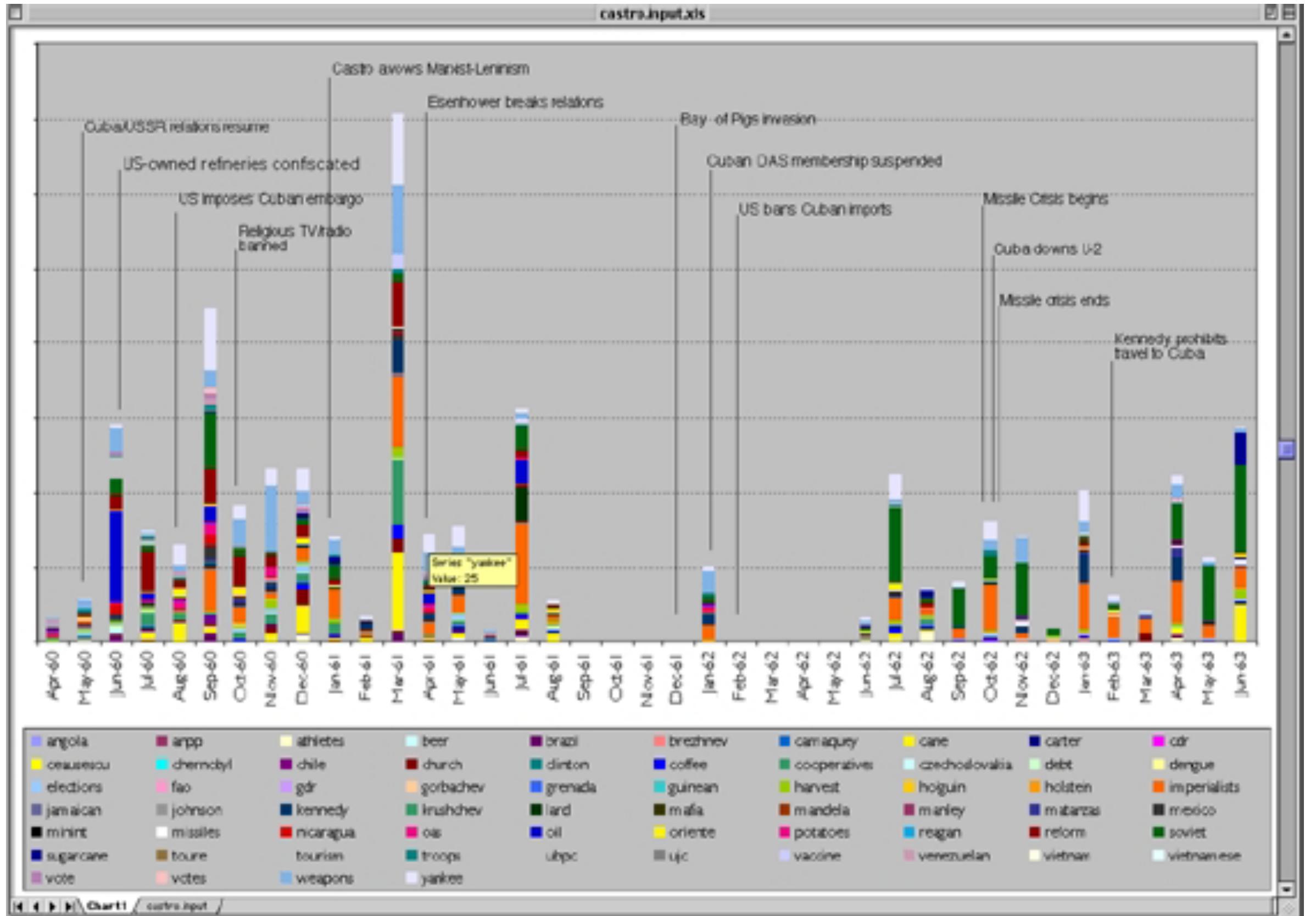
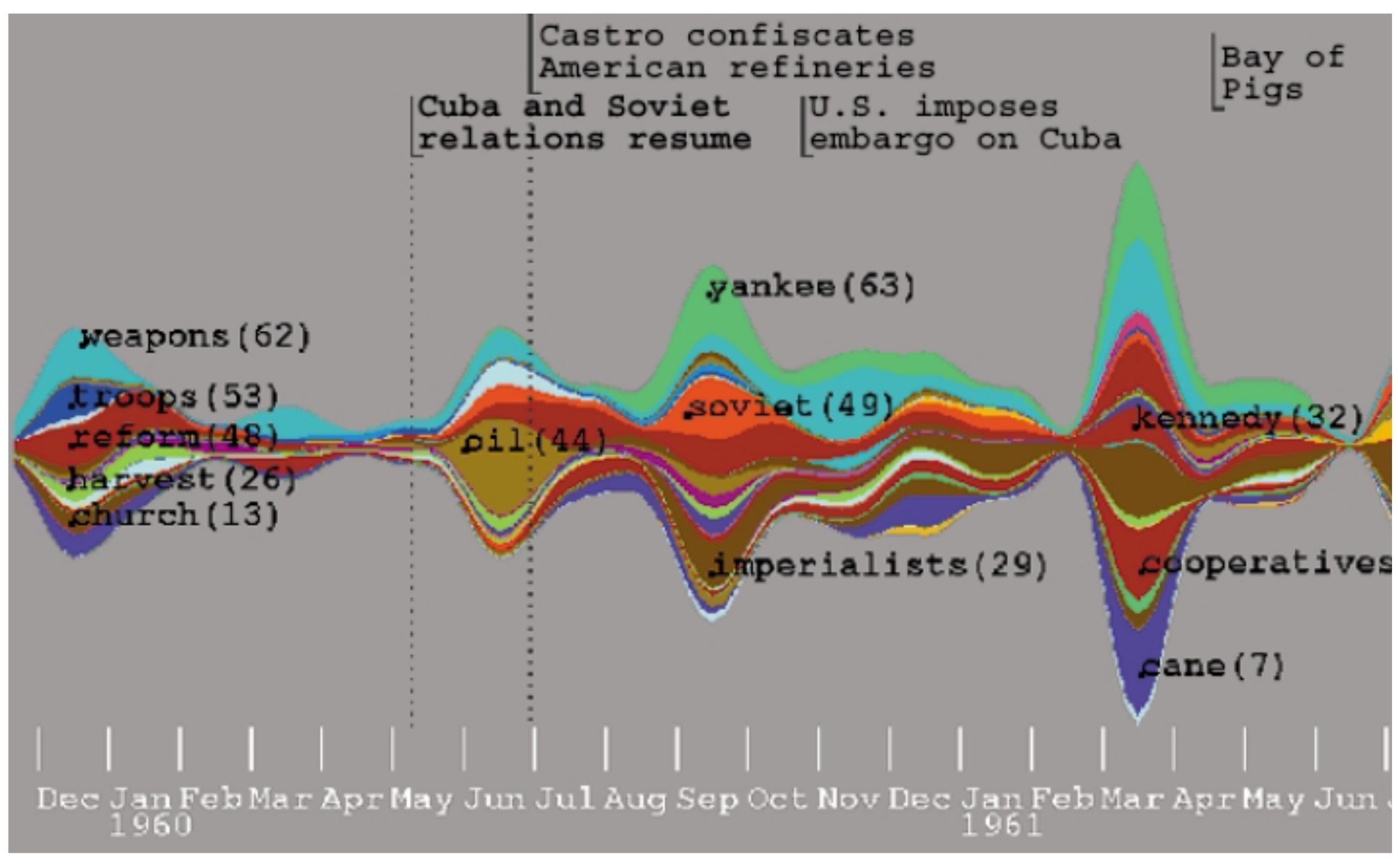


Figure 2: Like ThemeRiver™ in Figure 1, this histogram uses the Castro collection data and depicts changes in thematic content over time.

- Para que se possa responder à pergunta, é preciso associar os histogramas ao longo do tempo
- Como as barras estão ancoradas na sua base, a posição de cada tema pode variar consideravelmente ao longo do tempo, dificultando a identificação e associação dos objetos visuais

# REPRESENTAÇÃO VISUAL

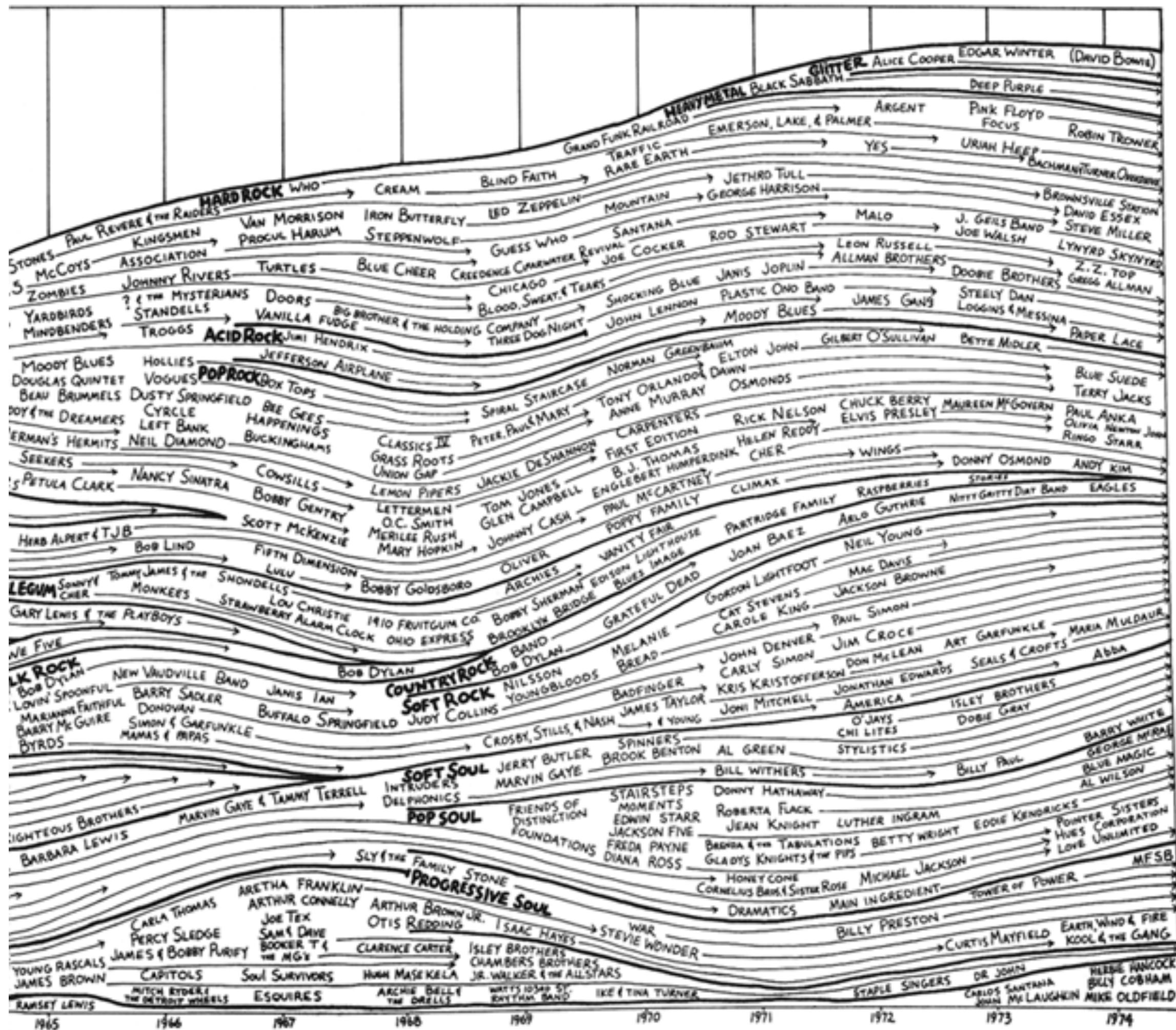
Falas, entrevistas, artigos e outros textos referentes a Fidel Castro



- Assim como no histograma, o *Themeriver* utiliza as espessuras das camadas para representar a frequência dos temas
- Porém, os temas permanecem conectados por curvas contínuas e suaves
- Metáfora: o fluxo do rio representa o fluxo do tempo

- Uso do eixo x representa os intervalos de tempo
- As espessuras das camadas representam a quantidade de menções aos temas

are repeated, as they resurface in fresh currents. The multiple, parallel flows locate music-makers in two dimensions—linking musical parents and offspring from 1955 to 1974, and listing contemporaries for each year.<sup>10</sup> With an intense richness of detail (measuring in at 20% of the typographic density of a telephone book), this nostalgic and engaging chart fascinates many viewers—at least those of a certain age. Also the illustration presents a somewhat divergent perspective on popular music: songs are not merely singles—unique, one-time, *de novo* happenings—rather, music and music-makers share a pattern, a context, a history.



<sup>10</sup> Among the missing are The Weavers, Pete Seeger, Bonnie Raitt, and Lou Reed and The Velvet Underground.

# PONTOS POSITIVOS

---

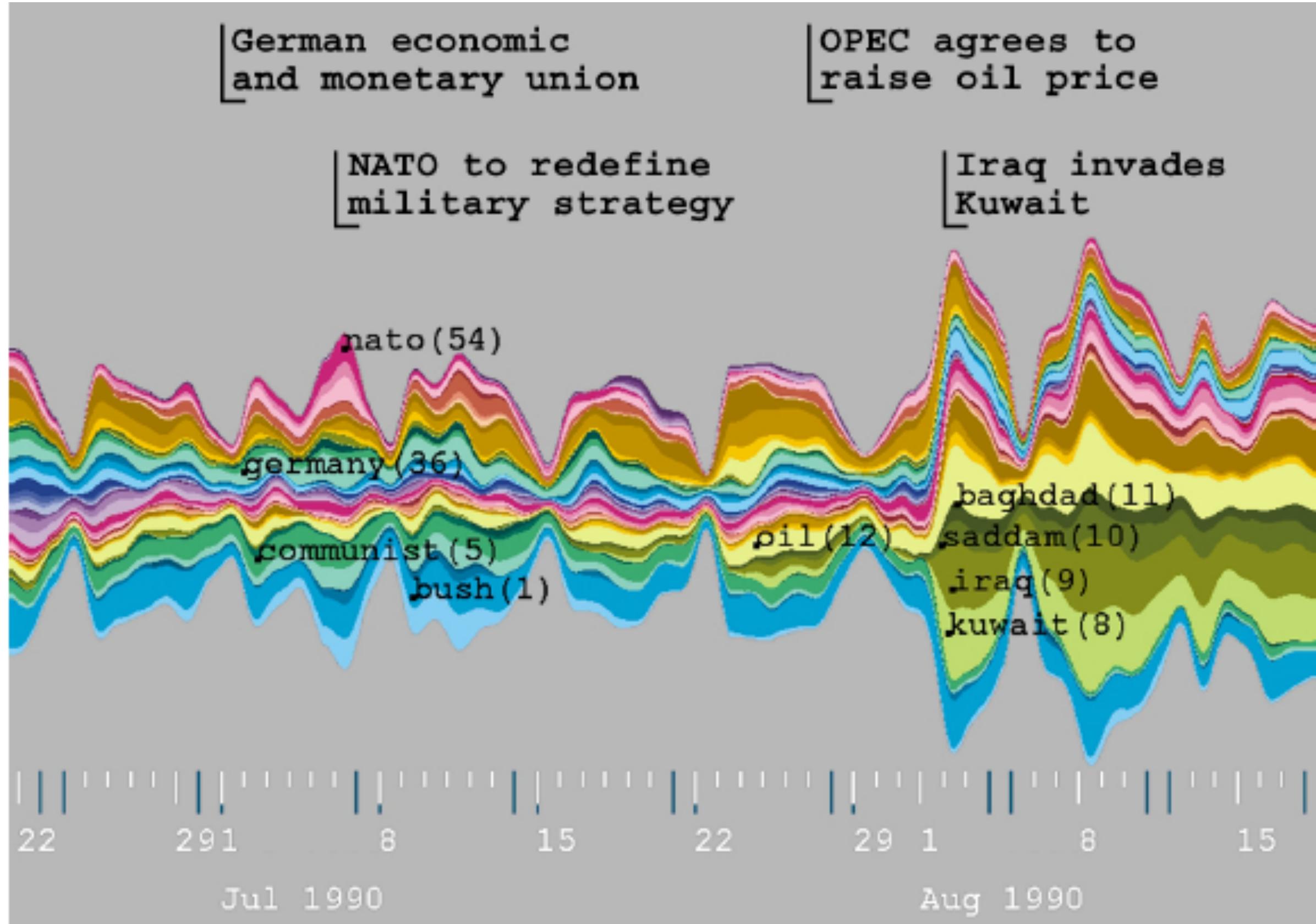


Figure 3: AP data from July - August 1990. A wide current in the river indicates heavy use of a topic, while changes in color distribution correlate to changes in themes.

- Um tema pode aparecer, desaparecer e reaparecer e permanece facilmente reconhecível devido a manutenção da cor e posicionamento
- Exibição de informação de contexto: tempo e eventos importantes

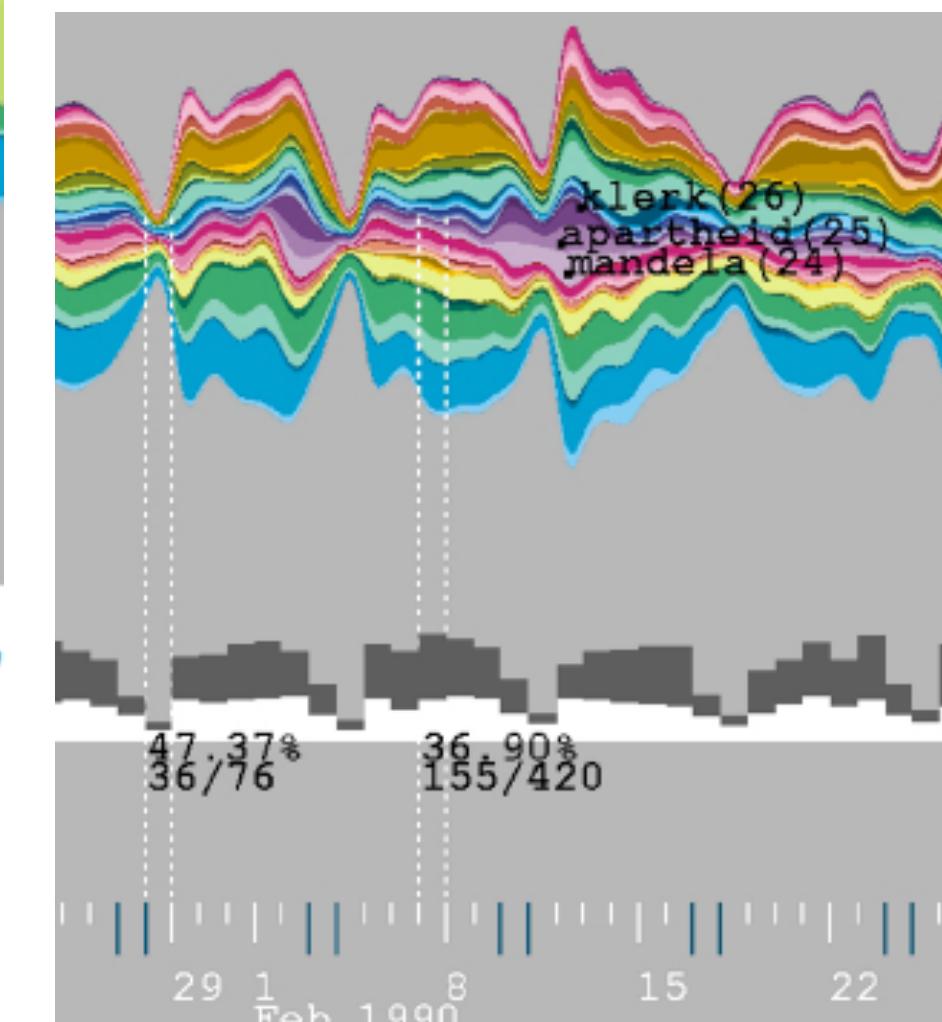
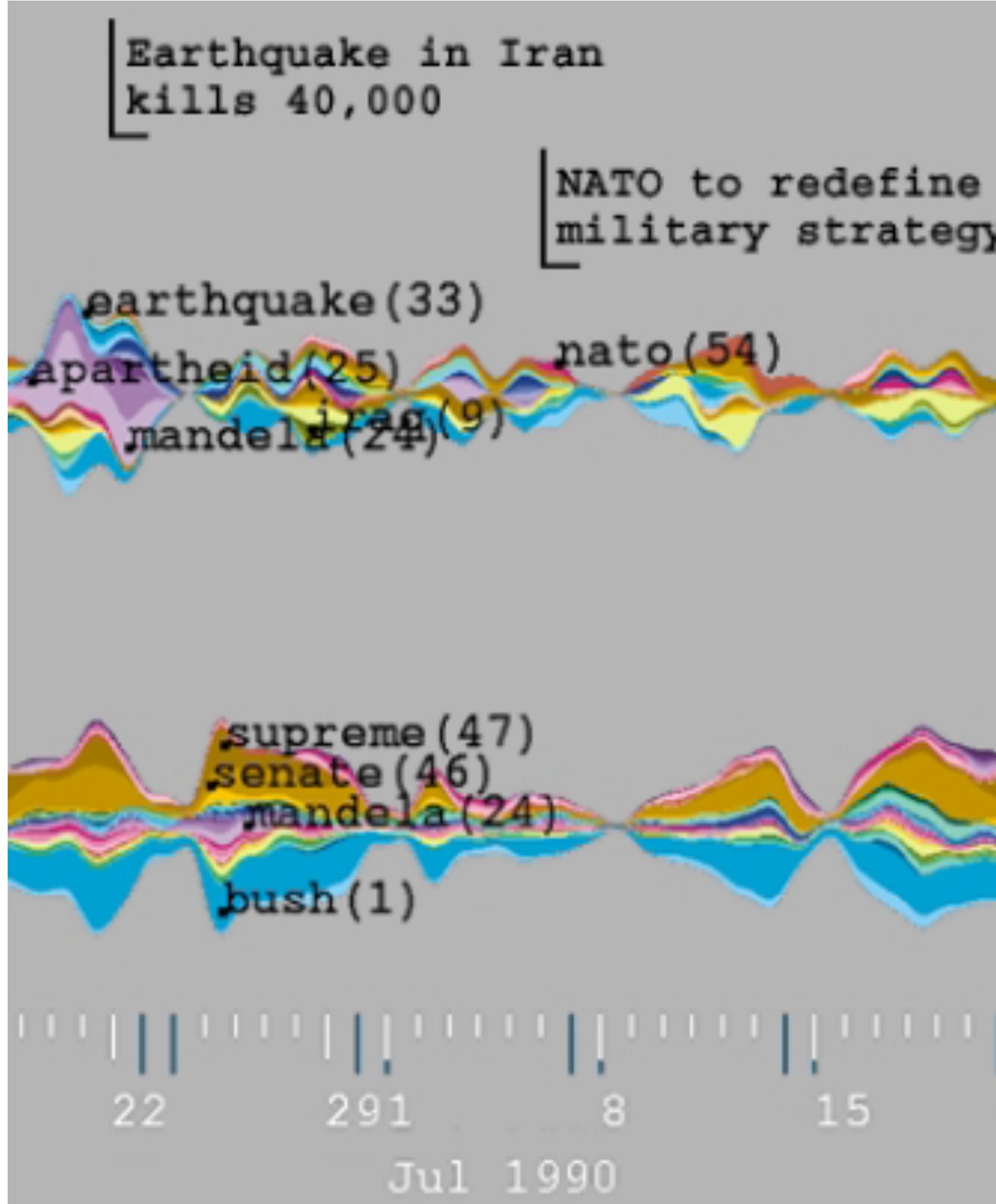


Figure 5: The addition of a histogram to Theme-River™ reveals that news is light on Sundays, not that themes shift.

# LIMITAÇÕES

---



**Figure 6: Parallel rivers let users compare AP data from Washington, D.C. and New York from the same time period.**

- O número de termos que podem ser exibidos é limitado.
- Como exibir um grande número de padrões e ainda assim manter o poder de discriminação entre eles?
- Nossa percepção de cores depende do contexto e como os temas aparecem e desaparecem, é difícil saber as cores que estarão adjacentes
- Uma solução é agrupar os temas em famílias e exibi-los em cores análogas

# AVALIAÇÃO

---

- Os usuários **compreendem a metáfora do rio?**
- Conseguem **identificar temas mais frequentemente discutidos?**
- A visualização é útil no levantamento de **novas questões** sobre os dados?
- Eles interpretam os dados da forma esperada?
- Como a interpretação difere da interpretação do **histograma** que exibe os mesmos dados?

# AVALIAÇÃO

---

- Público
- 2 usuários
- Questões específicas
  - Em julho de 1962, quais eram os temas mais discutidos?
  - Onde um novo tema é introduzido?
- Questões gerais
  - O que te parece interessante? O que você gostaria de explorar?
  - Como você gostaria de modificar ou manipular a visualização?

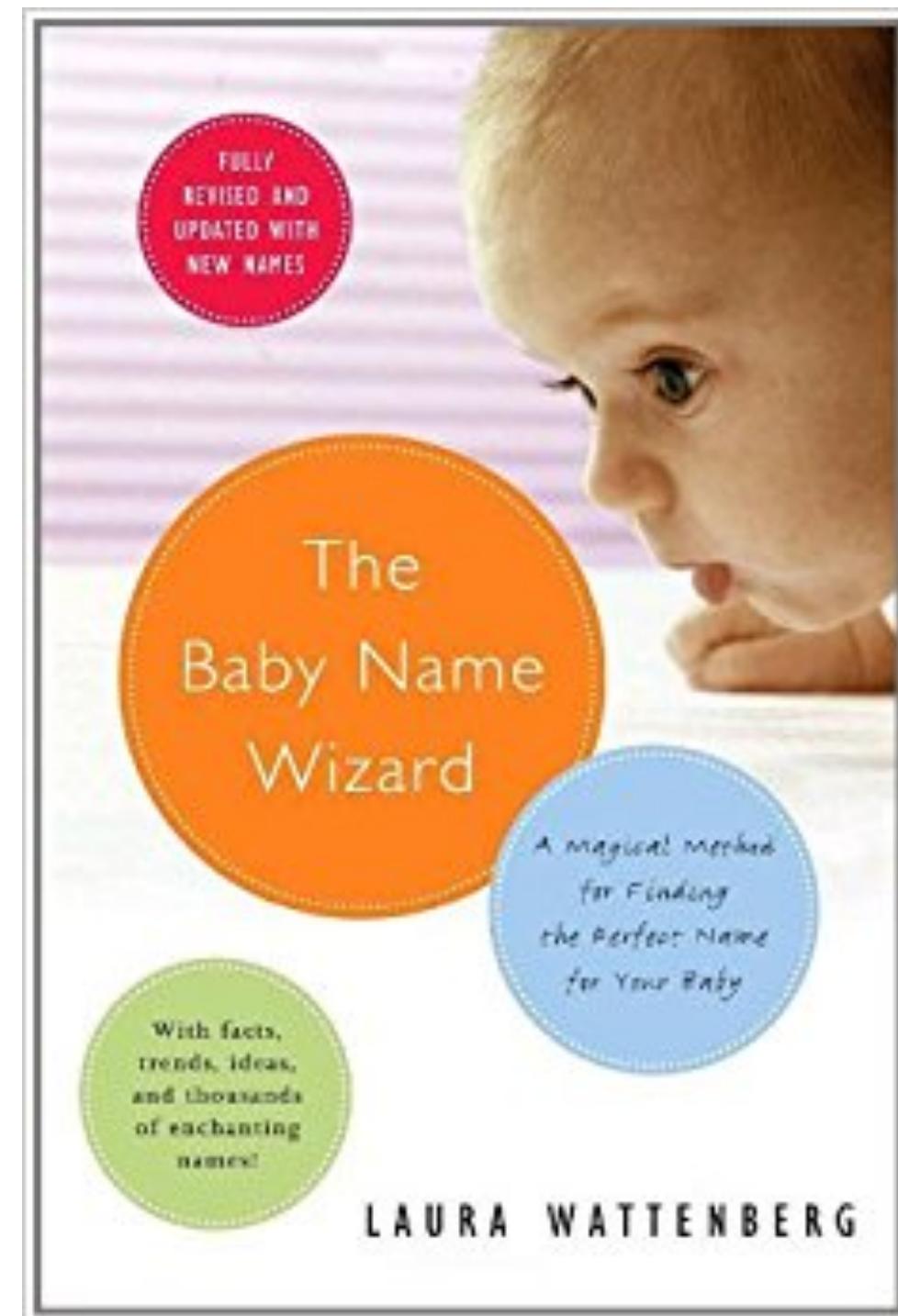
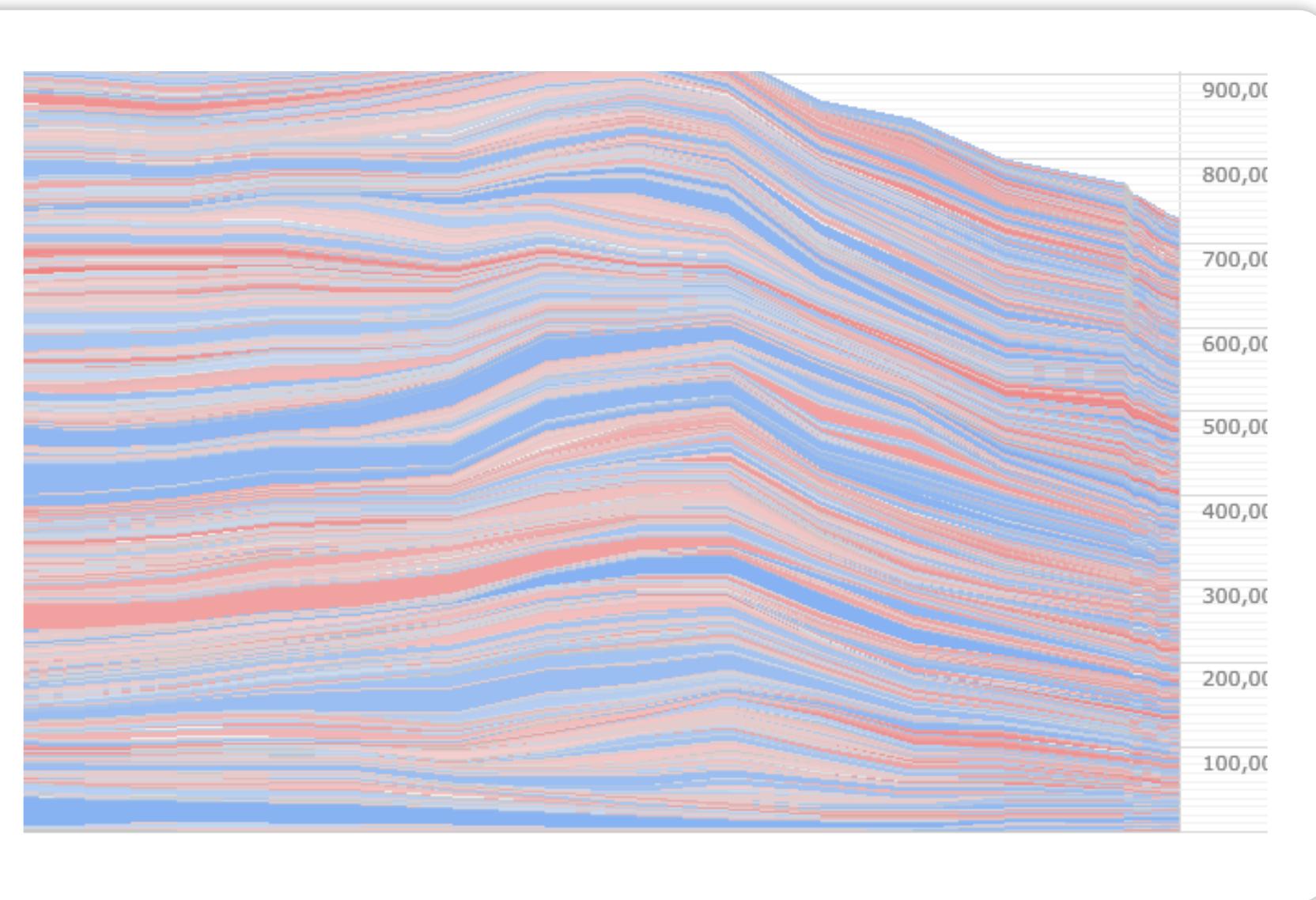
# RESULTADOS DA AVALIAÇÃO

---

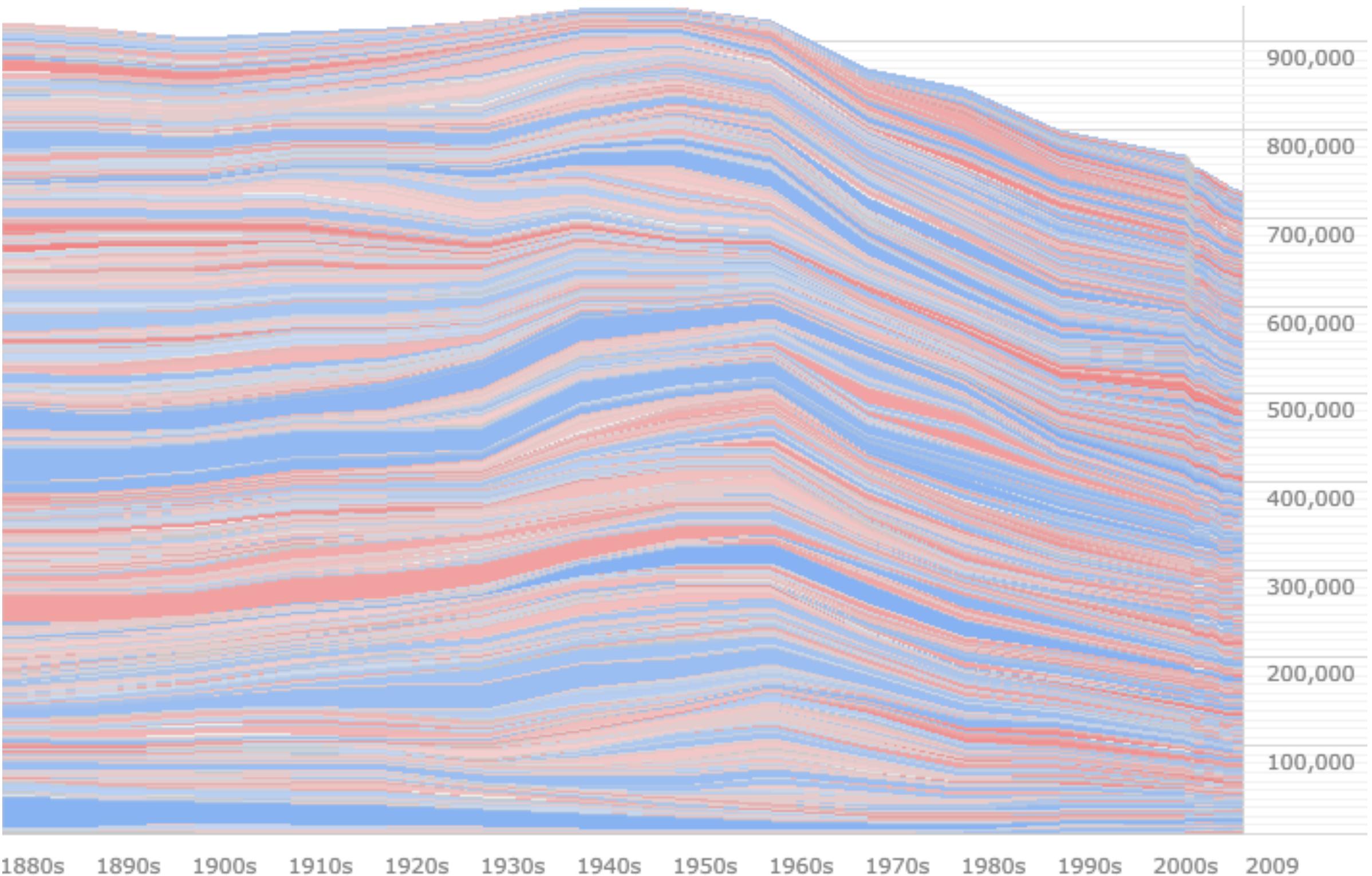
- Usuários não tiveram dificuldades em entender a metáfora
- Foram capazes de identificar temas fortemente representados e de entender o relacionamento entre a espessura da faixa e a representação do tema
- Acharam difícil analisar temas menos frequentes
- Sugeriram possibilitar a visão dos valores dos dados brutos (número e documentos relacionados) através da interação

# BABY NAMES, VISUALIZATION, AND SOCIAL DATA ANALYSIS

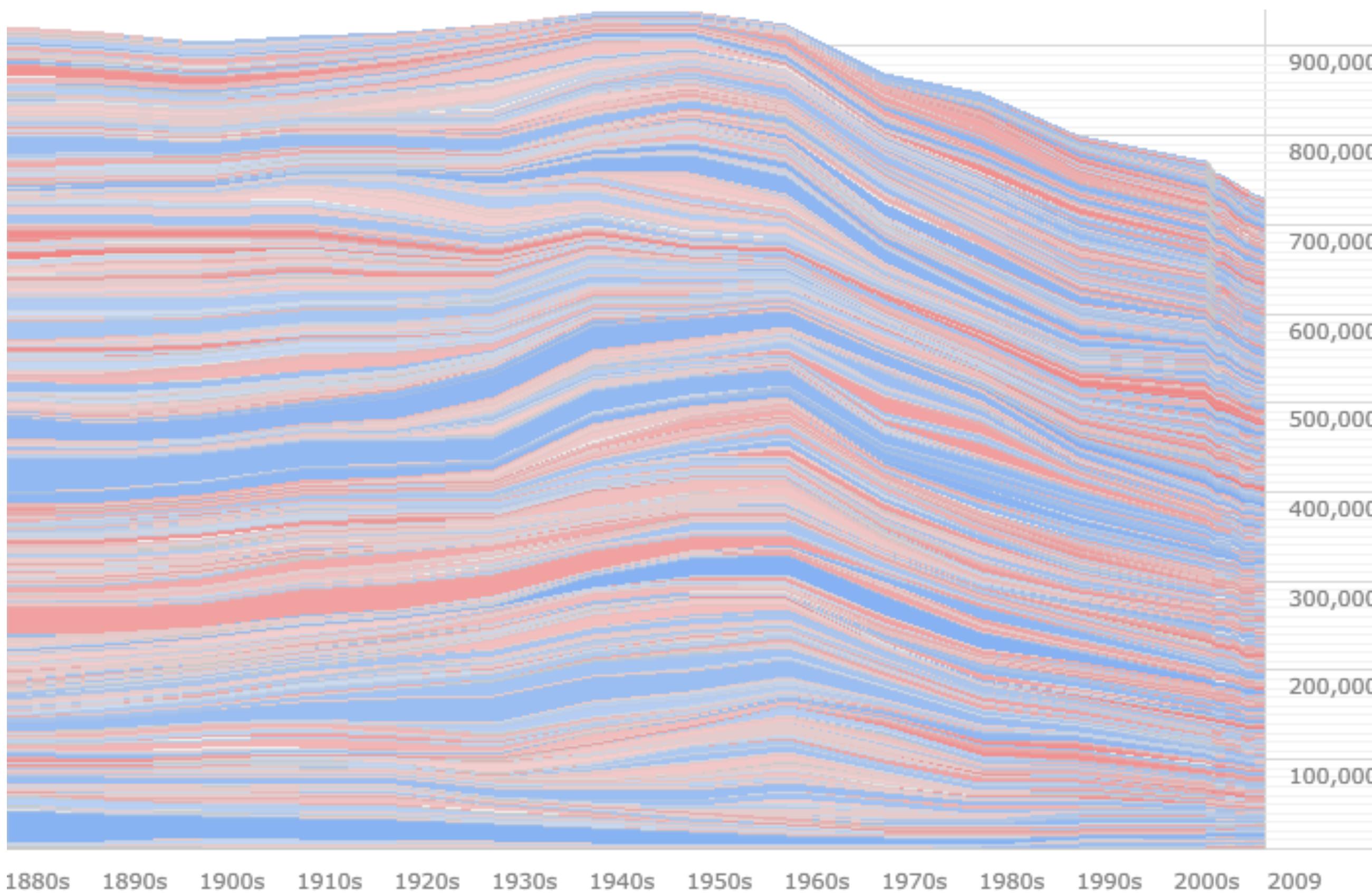
M. Wattenberg  
*IEEE Symposium on Information Visualization  
2005*



- O NameVoyager foi criado para chamar a atenção para o livro *The Baby Name Wizard*
- Ele permite a exploração de dados sobre a popularidade do uso dos nomes no tempo
- 500.000 acessos nas primeiras duas semanas e 10.000 acessos por dia após este período
- Pessoas engajadas com a visualização, gastando tempo considerável na exploração dos dados e na descoberta de fatos sobre os nomes
- 6.000 séries temporais

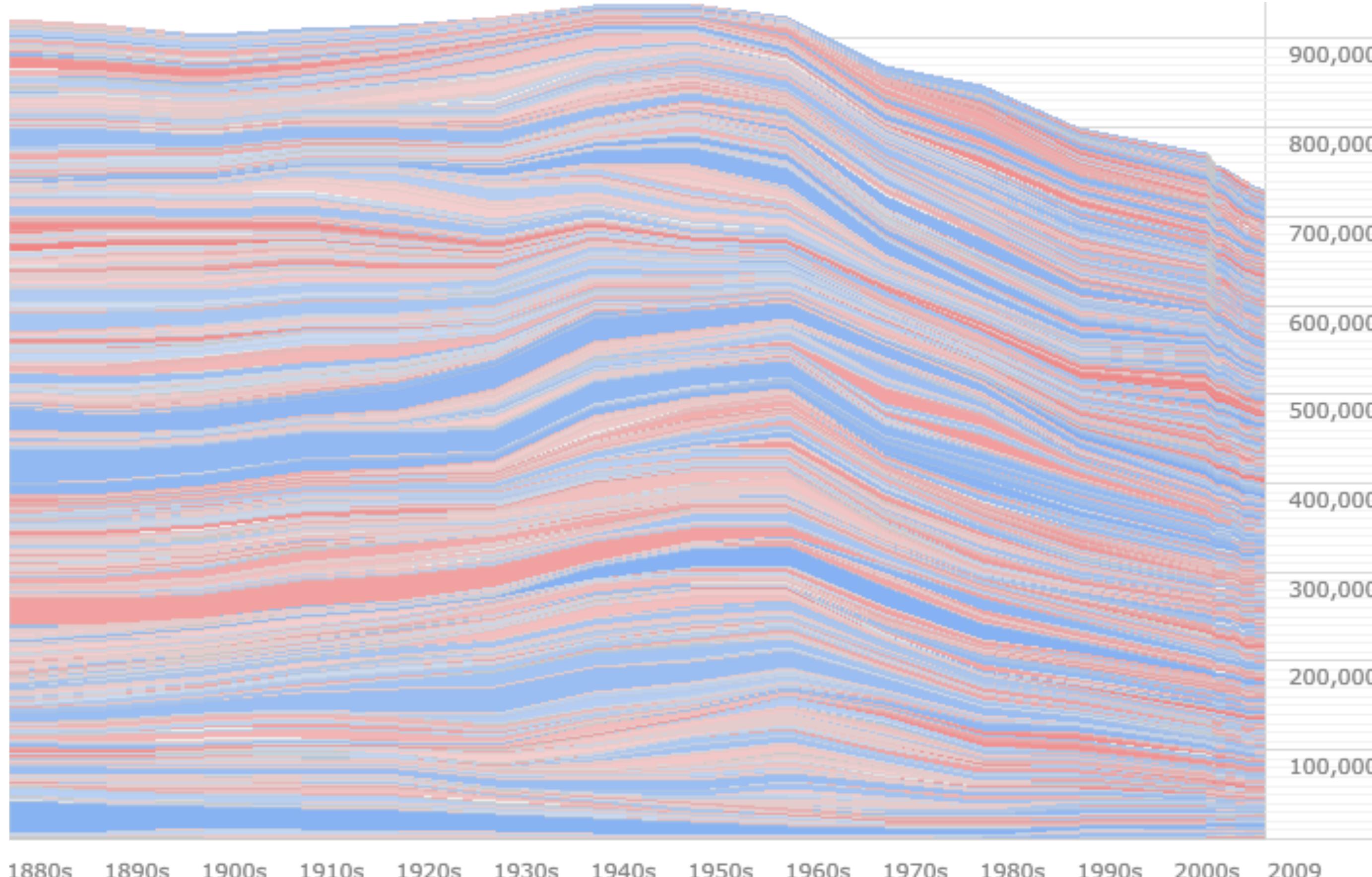


- Dados provenientes do Social Security Administration (SSA) dos Estados Unidos
- Para cada década desde 1.900, a lista dos 1.000 nomes de meninos e meninas mais comuns é publicada
- 6.000 nomes no total



# REPRESENTAÇÃO VISUAL

---



- Gráfico de área empilhada
- Eixo x representa as datas
- Eixo y representa a frequência para os nomes em termos da sua ocorrência a cada um milhão de bebês
- Cada faixa corresponde a um nome e sua largura é proporcional a frequência do nome em um dado instante do tempo

# REPRESENTAÇÃO VISUAL

---

- As cores são tons de azul para meninos e de rosa para meninas
- O brilho contribui no destaque da popularidade de cada nome, sendo as faixas mais escuras referentes a nomes mais populares atualmente
- Nomes populares tendem a ser mais procurados
- O uso do brilho faz que não seja necessário usar bordas nas faixas
  - Desenham apenas as faixas com mais de 2 pixels de largura
  - Na prática, cerca de 200 faixas

# MANTRA DA VISUALIZAÇÃO DE DADOS

“

Overview first, zoom and filter, details  
on demand

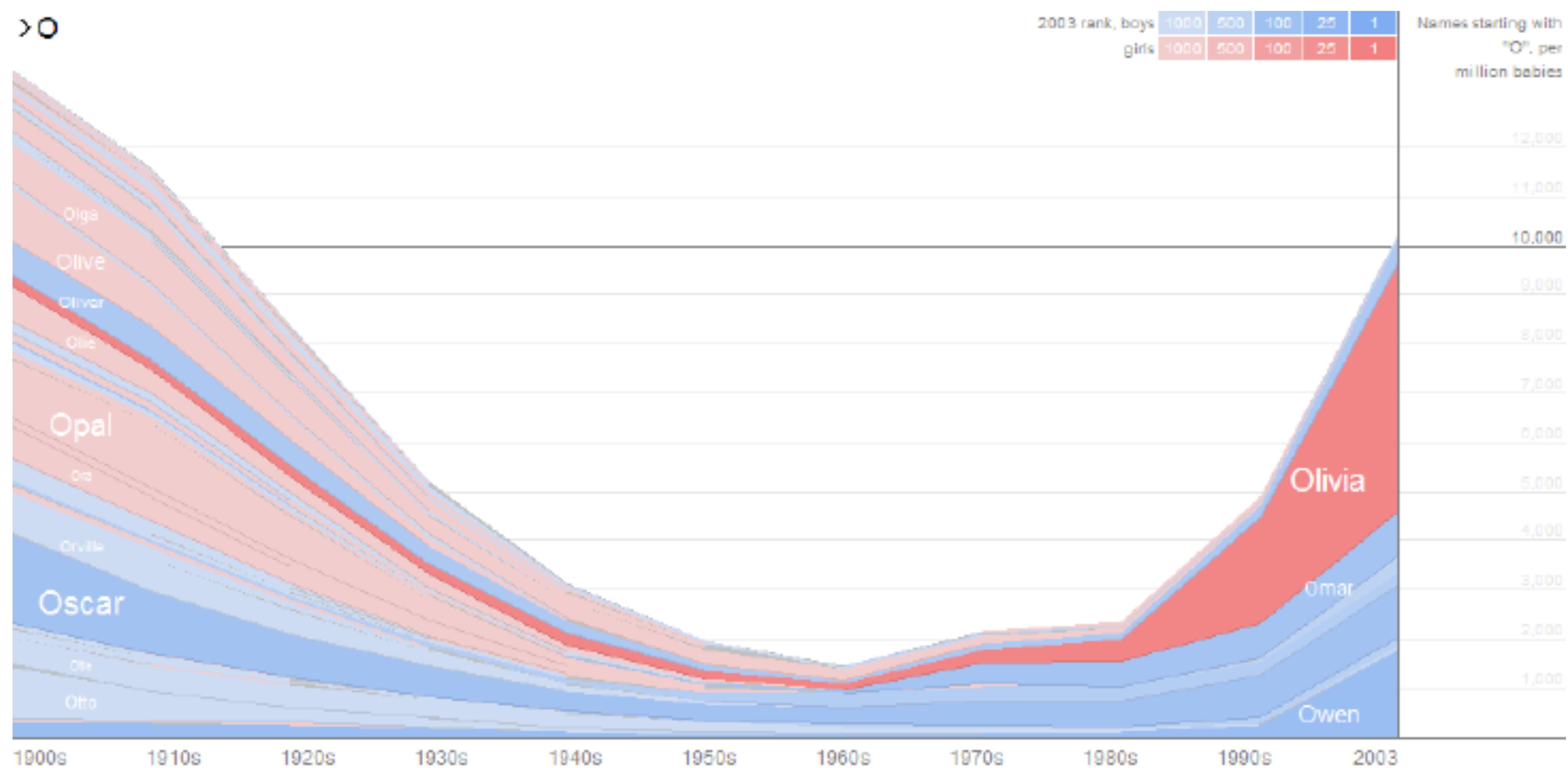
-B. Schneiderman

Shneiderman, Ben. "The eyes have it: A task by data type taxonomy for information visualizations." *Proceedings 1996 IEEE symposium on visual languages*. IEEE, 1996.

## INTERATIVIDADE

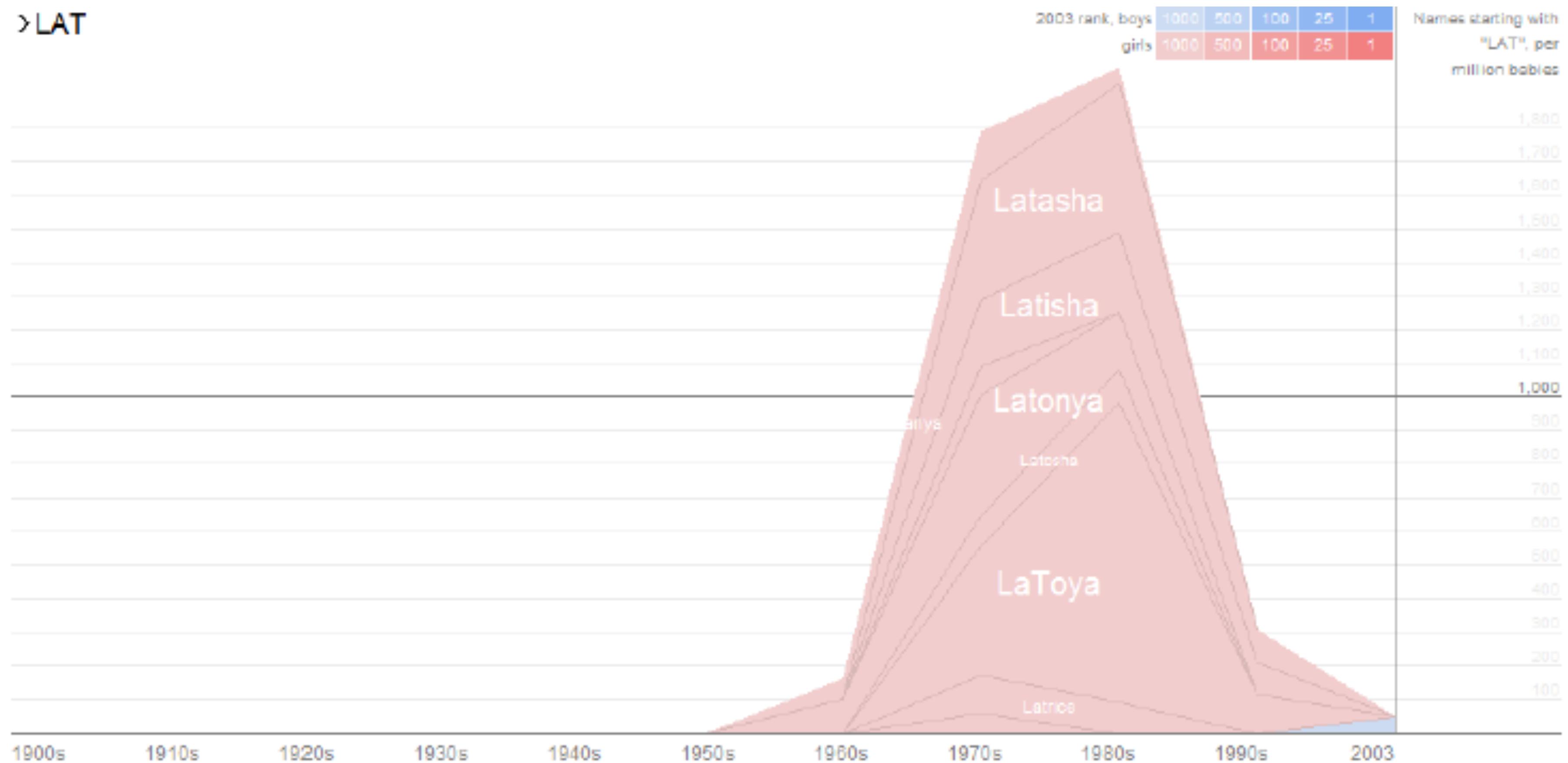
---

- No início o usuário vê os 6.000 nomes
- Pode selecionar por sexo
- Por prefixo (reage a cada nova tecla pressionada)
- Ou mesmo buscar por nomes específicos

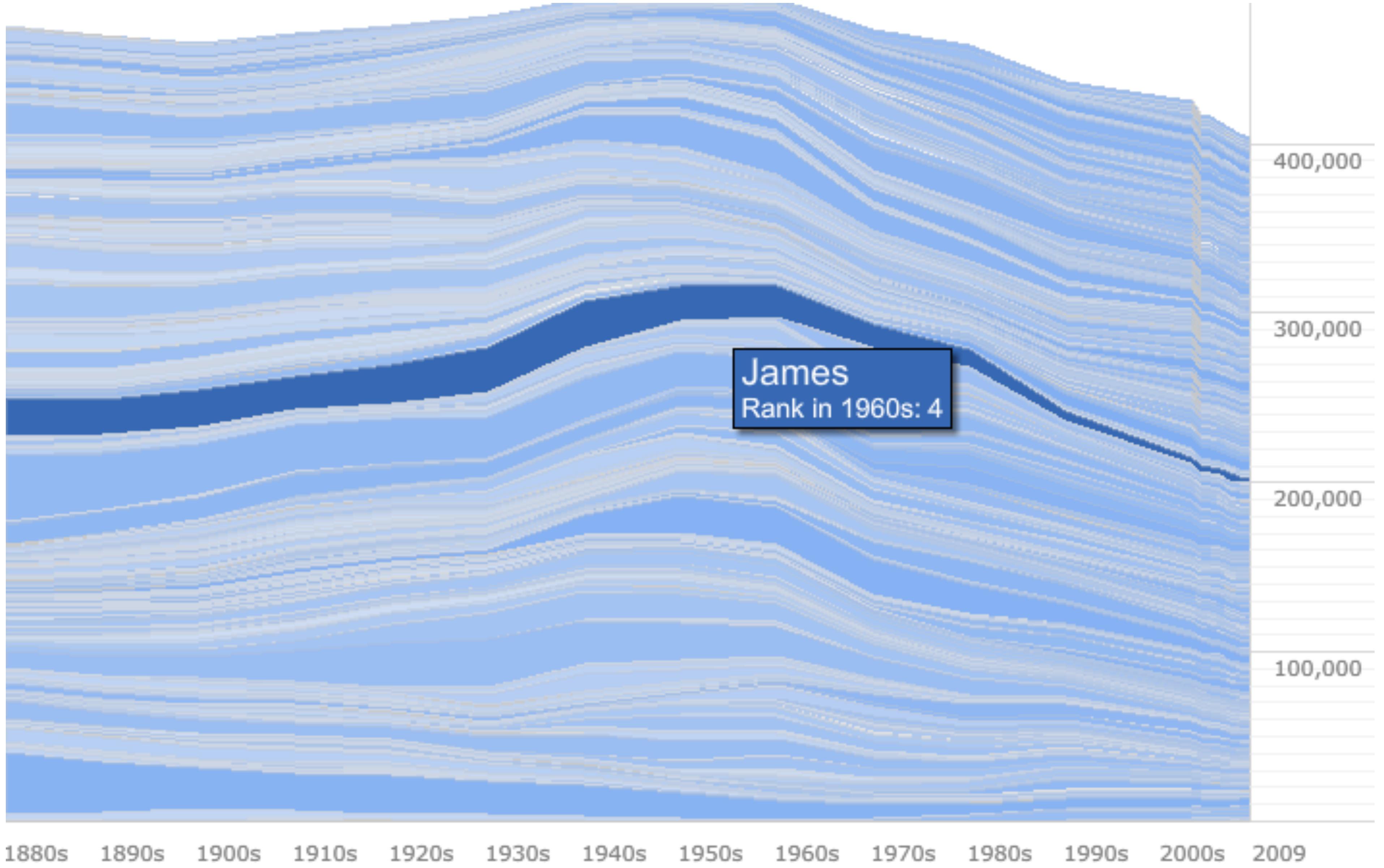


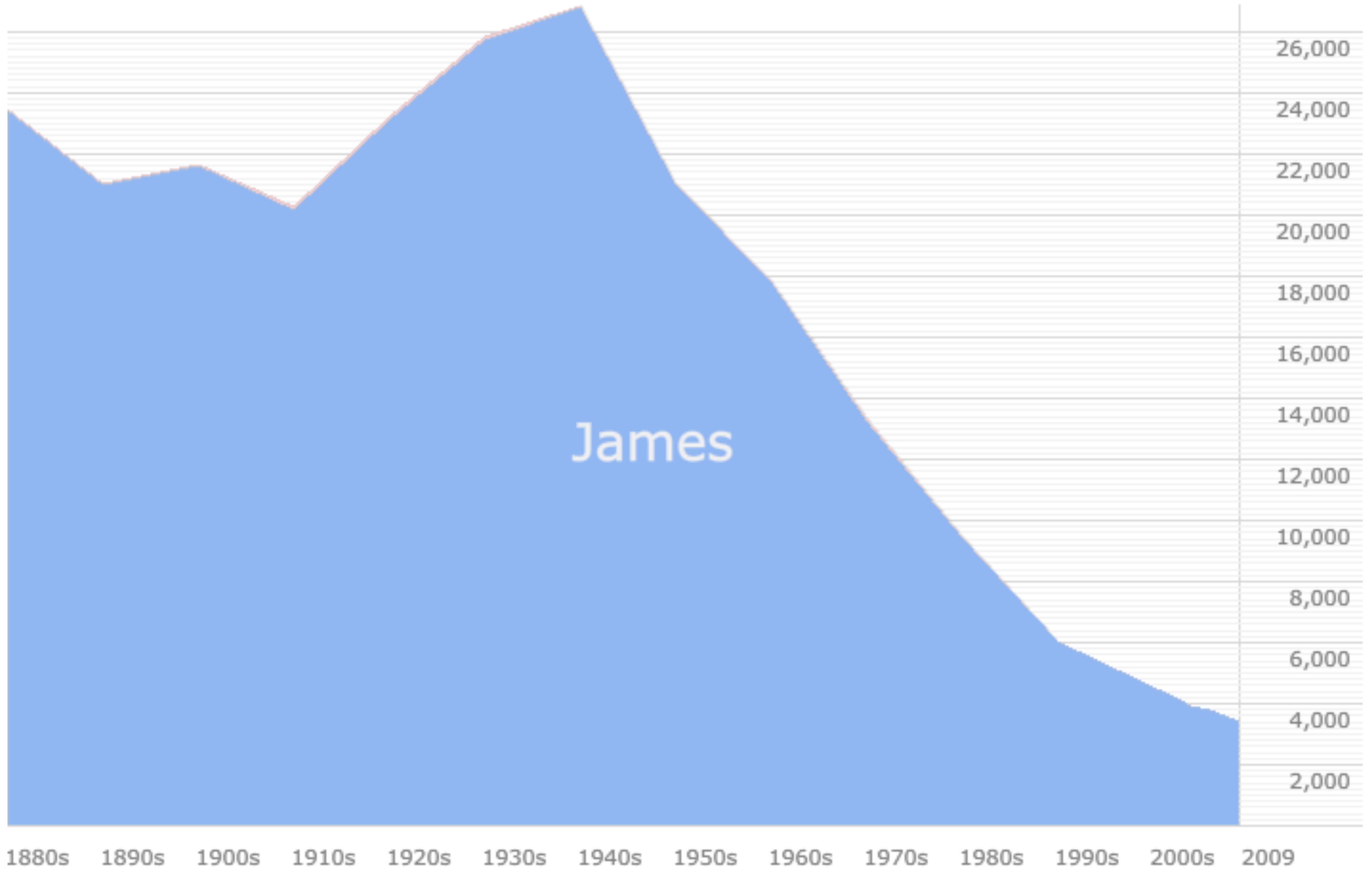
*Figure 2. Names beginning with O*

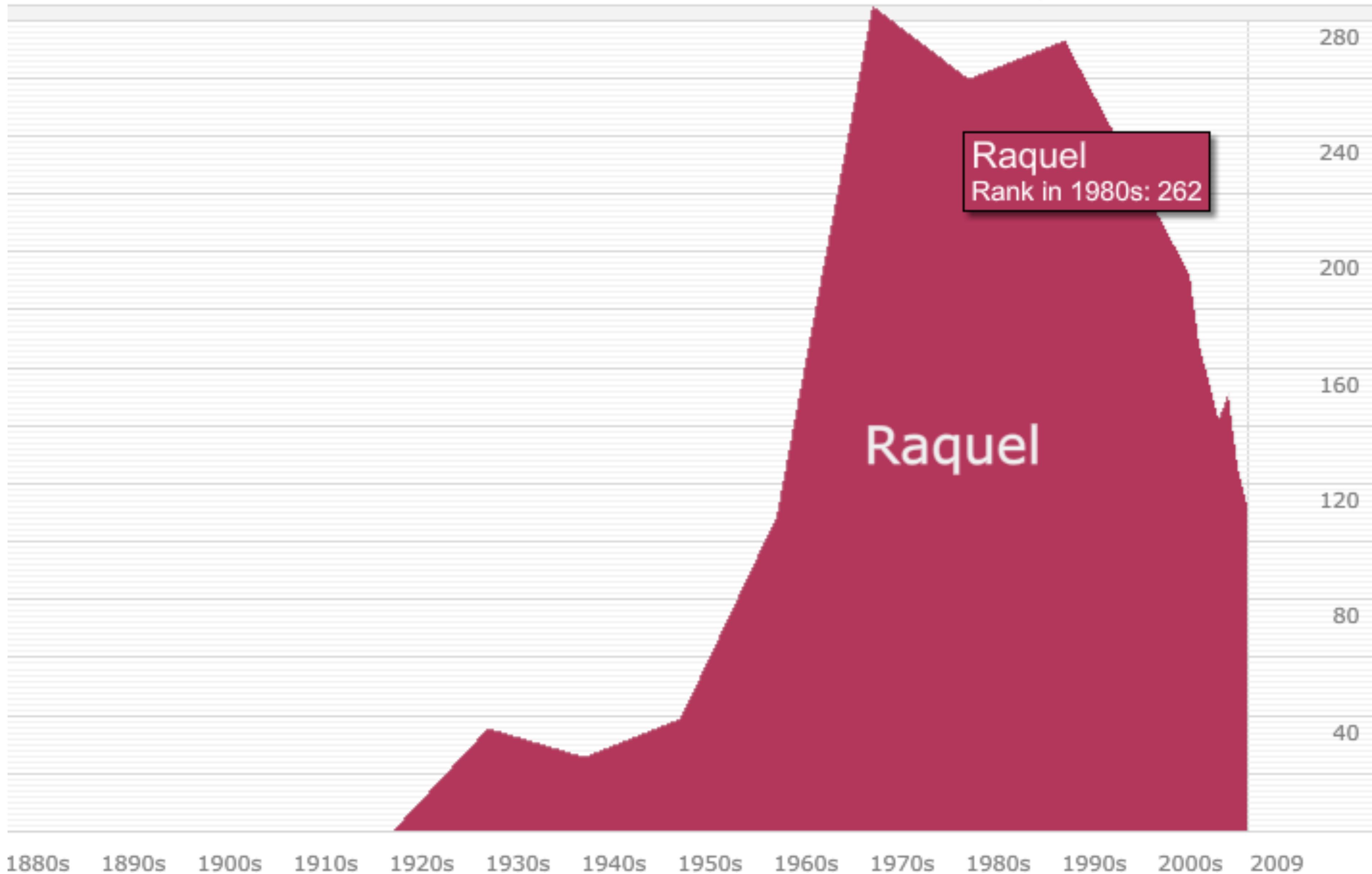
> LAT



*Figure 3. Names Beginning with LAT*







# AVALIAÇÃO

---

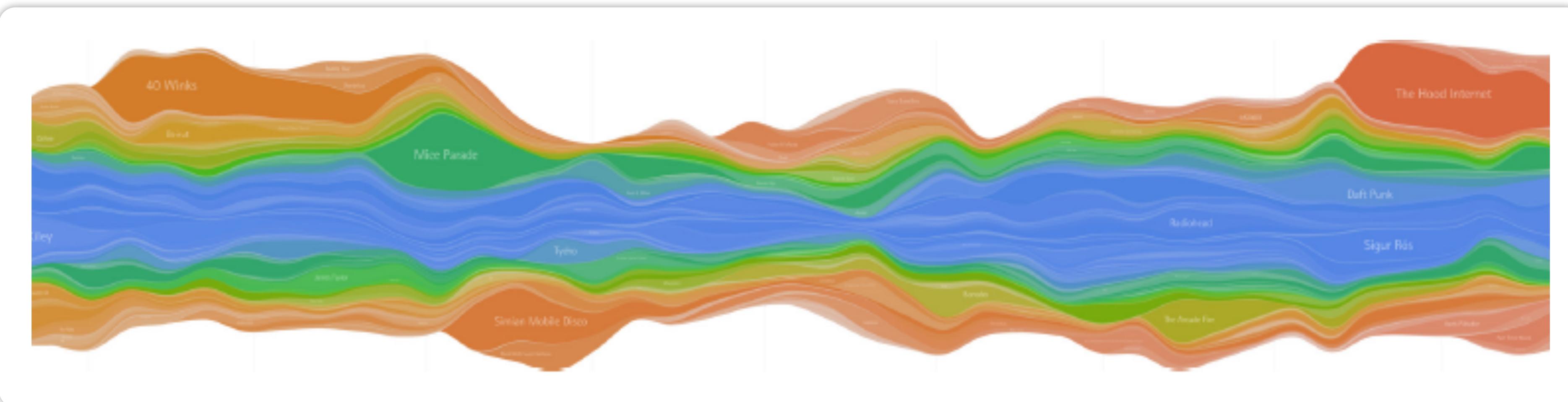
- Os autores avaliam os comentários deixados pelos usuários
- Caráter de atividade social da visualização da visualização c
- *Background* comum
- O uso de animações produz o efeito espectador, ou seja, usa uma interface expressiva

S. Reever et al., *Designing the spectator experience*, Conference for Human-Computer interaction, 2005

- Compartilhamento de descobertas

# STACKED GRAPHS – GEOMETRY & AESTHETICS

*L. Byron e M. Wattenberg*  
*IEEE Symposium on Information Visualization*  
2008



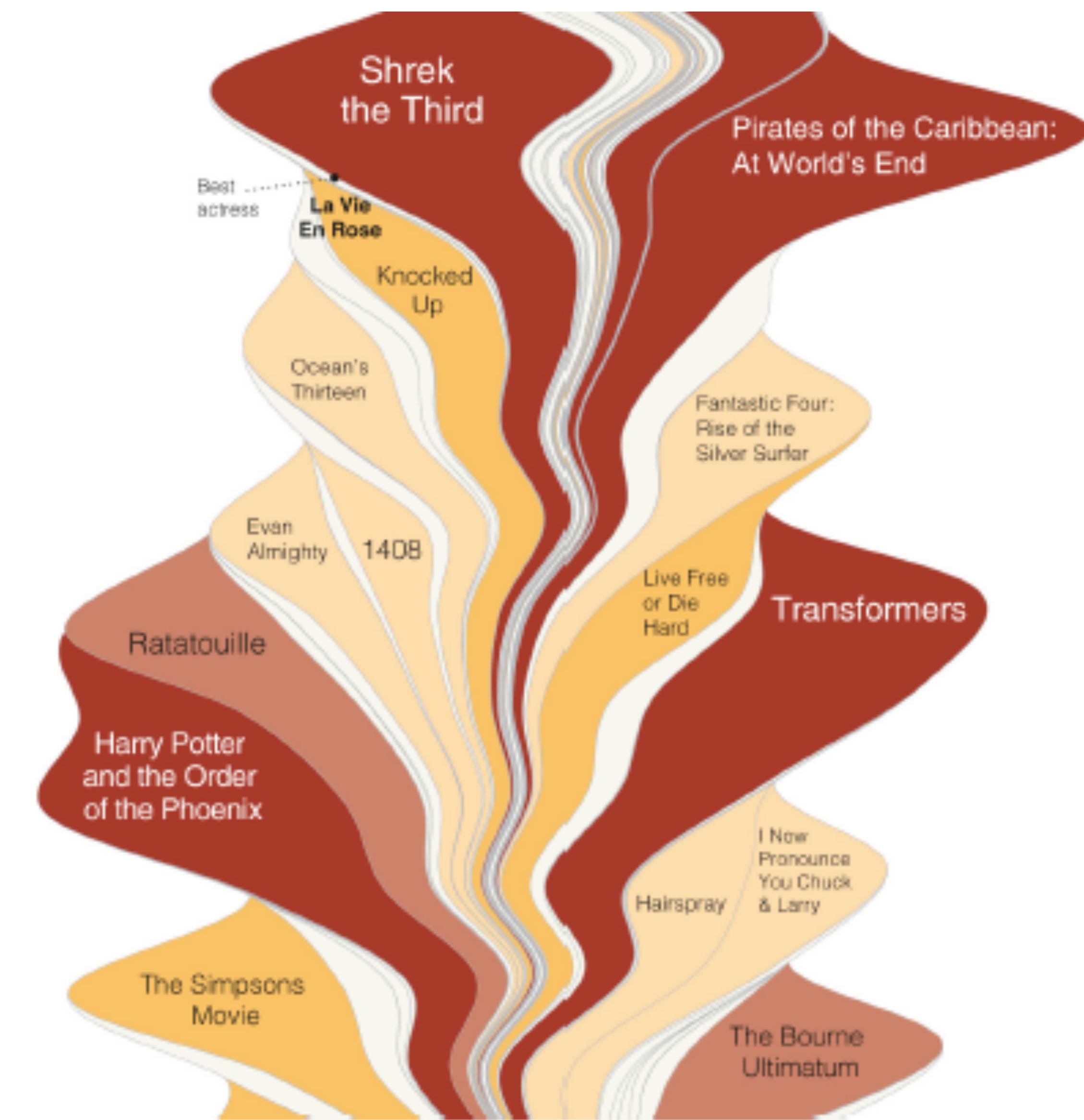
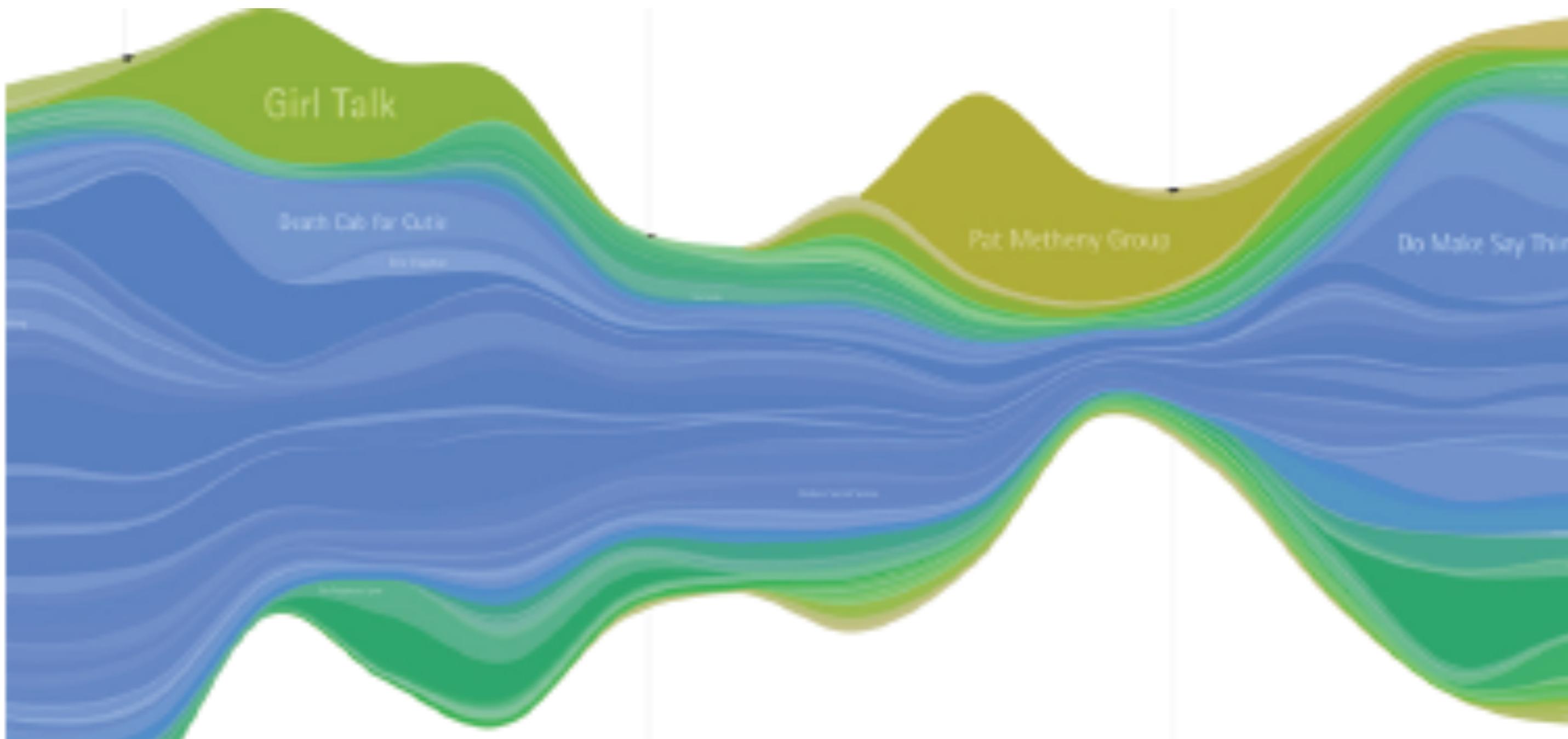


fig 2 – films from the summer of 2007

# OBJETIVO

---

- Criar uma representação que não tivesse aspecto matemático ou científico, mas emocionalmente atraente
- Pessoas associam eventos vividos aos seus hábitos musicais



*fig 1 – section from Listening History of primary author*

# REPRESENTAÇÃO VISUAL

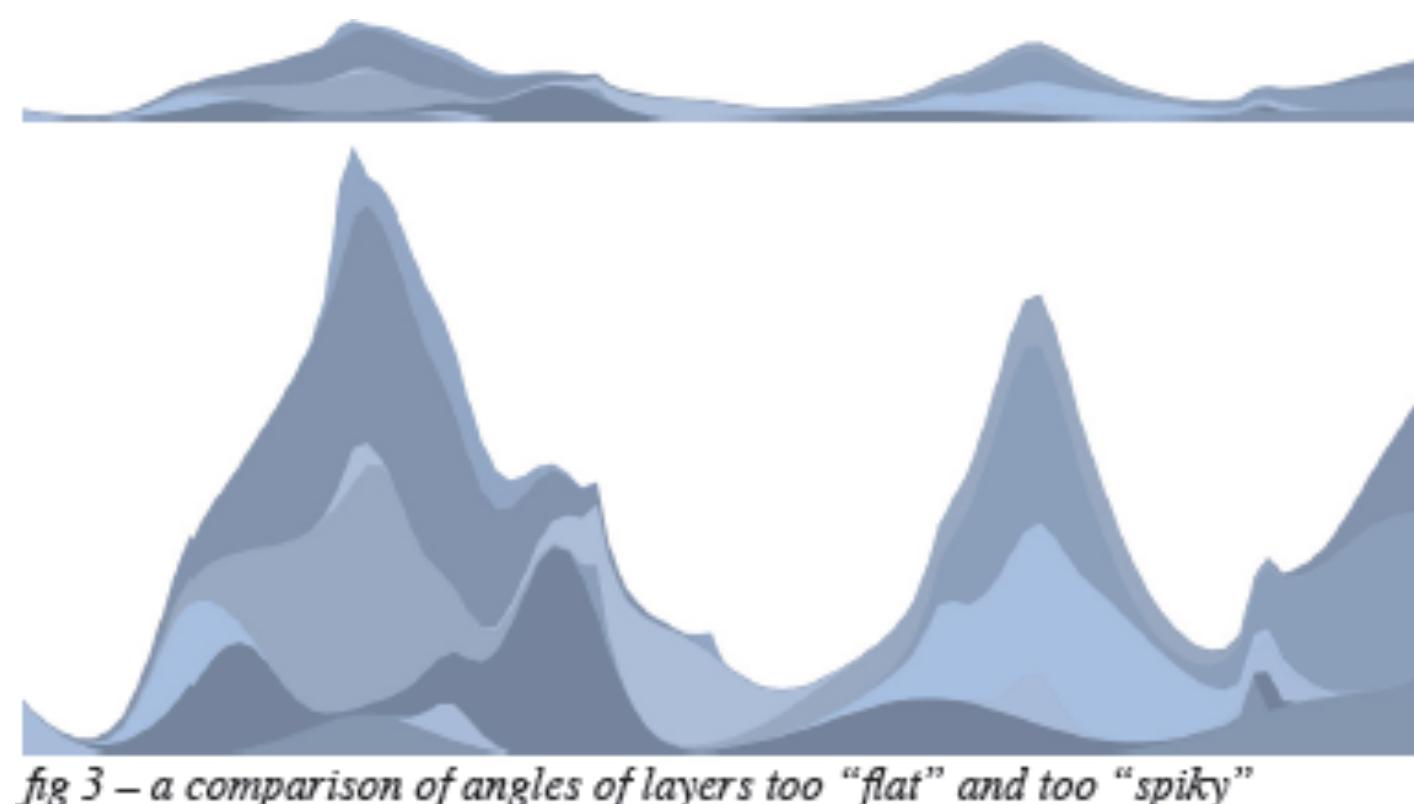
---

- Baseada em gráfico de áreas empilhadas, com ênfase nos aspectos de estética
- Eixo x representa as datas (semanas)
- Cada faixa corresponde a uma série temporal (um artista)
- A espessura da faixa representa o valor quantitativo a ser representado (número de vezes que se ouve uma música do artista)
- A cor representa a data em que um determinado artista foi descoberto pelo usuário
- A saturação indica a popularidade do artista

# CONSIDERAÇÕES ESTÉTICAS: LEGIBILIDADE

---

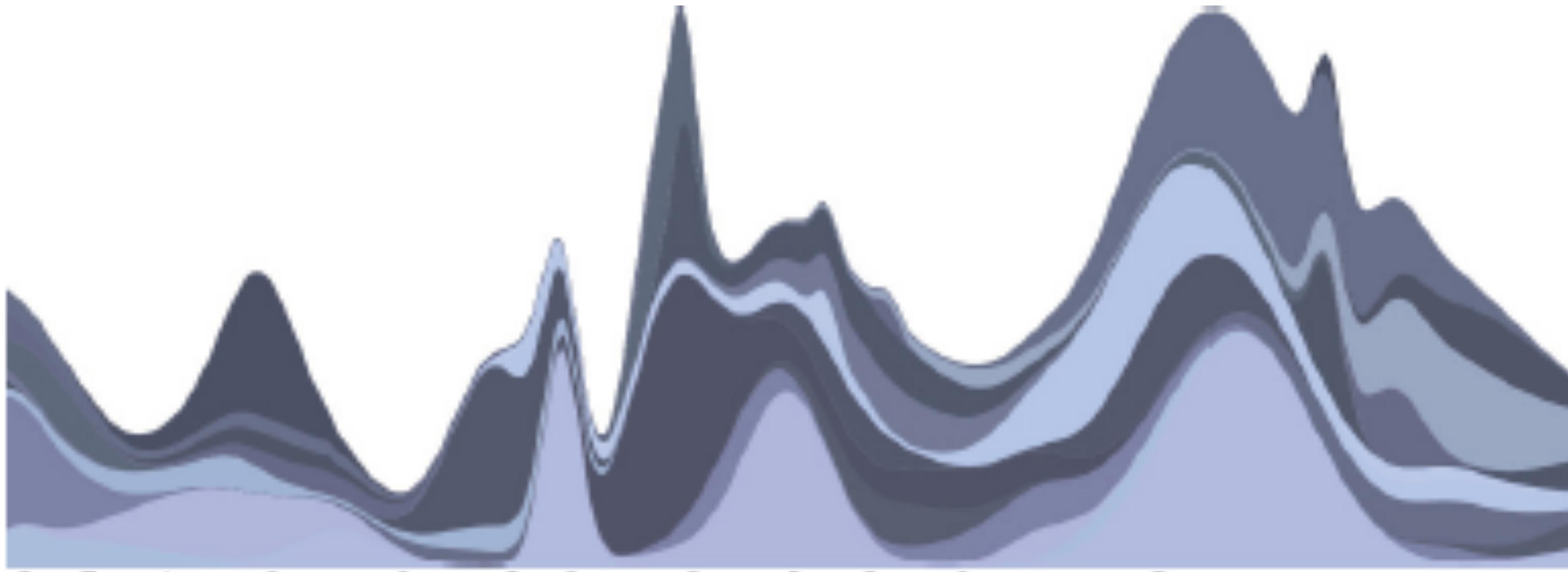
- A. Mudanças em camadas intermediárias provocam distorções em outras camadas ao redor
- B. Duas camadas da mesma espessura mas de inclinações diferentes podem parecer ter espessuras diferentes
- C. É preciso haver um compromisso entre o uso de gráficos muito planos e os cheios de picos (princípios da média de 45° de Cleveland)
- D. É preciso garantir que as camadas possam ser facilmente diferenciáveis



# ALGORITMO

---

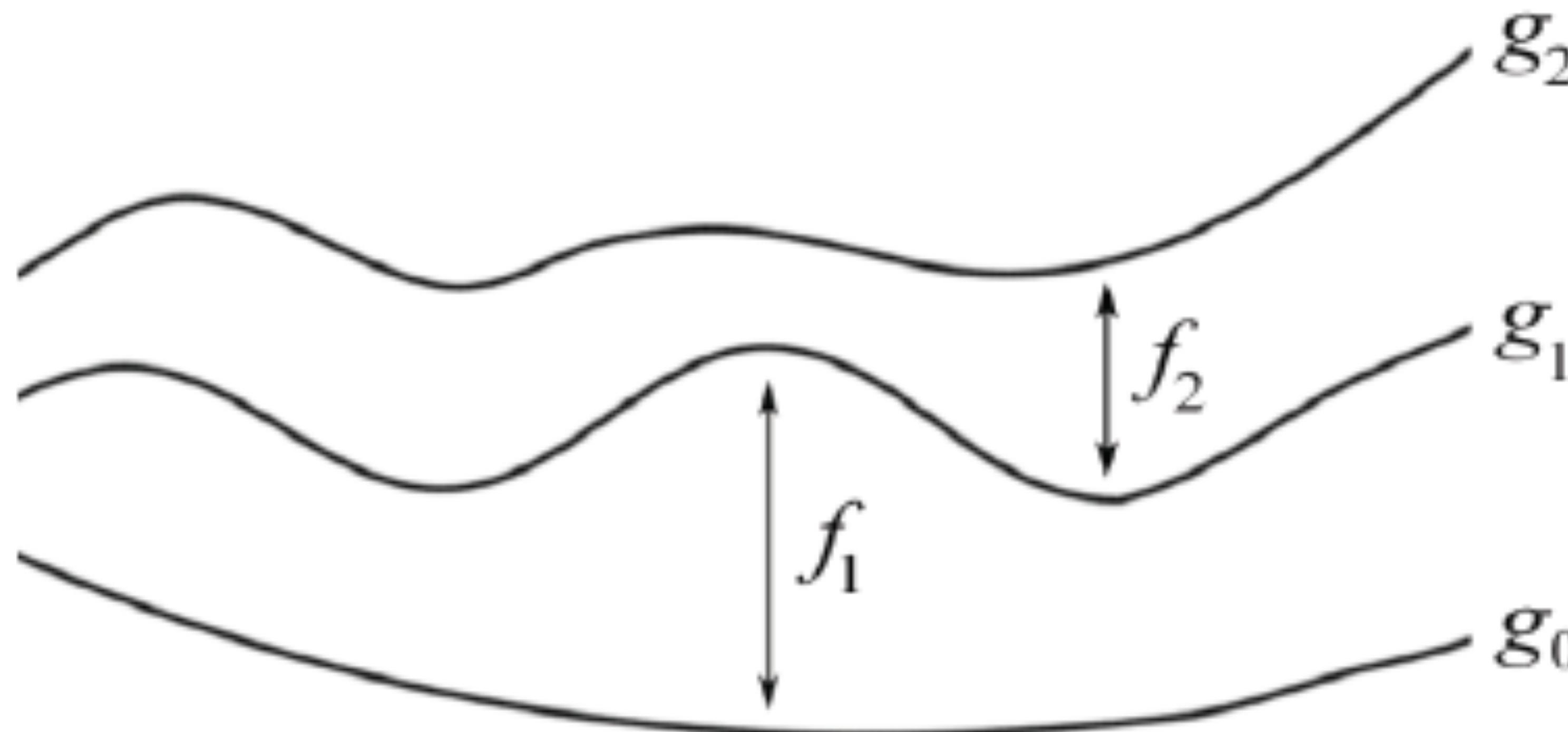
- É preciso considerar propriedades da **silhueta** do gráfico uma vez que ela implica nas inclinações e nas curvaturas de cada camada
- A **ordenação** das camadas é outro aspecto importante podendo levar a visualizações totalmente diferentes
- Os **rótulos** das camadas são essenciais para facilitar a identificação dos dados através dos objetos visuais
- As **cores** distinguem as camadas e podem ainda representar outras dimensões dos dados



*fig 5 – A traditional stacked graph with a baseline  $g_0 = 0$*

A soma das séries é legível mas a legibilidade de séries individuais fica prejudicada

A solução proposta pelo autor do *Themeriver* para este problema é o uso de um layout simétrico em relação a x



*fig 4 – a visual description of stacked graph functions  $f_i$  and  $g_i$  for  $n=2$  as used in this section*

A solução proposta pelo autor do *Themeriver* para este problema é o uso de um layout simétrico em relação a x

$$g_0 + g_n = 0$$

Como

$$g_i = g_0 + \sum_{j=1}^i f_j$$

então

e

$$2g_0 + \sum_{i=1}^n f_i = 0 \quad g_0 = -\frac{1}{2} \sum_{i=1}^n f_i$$

$g_0$  é chamada linha base

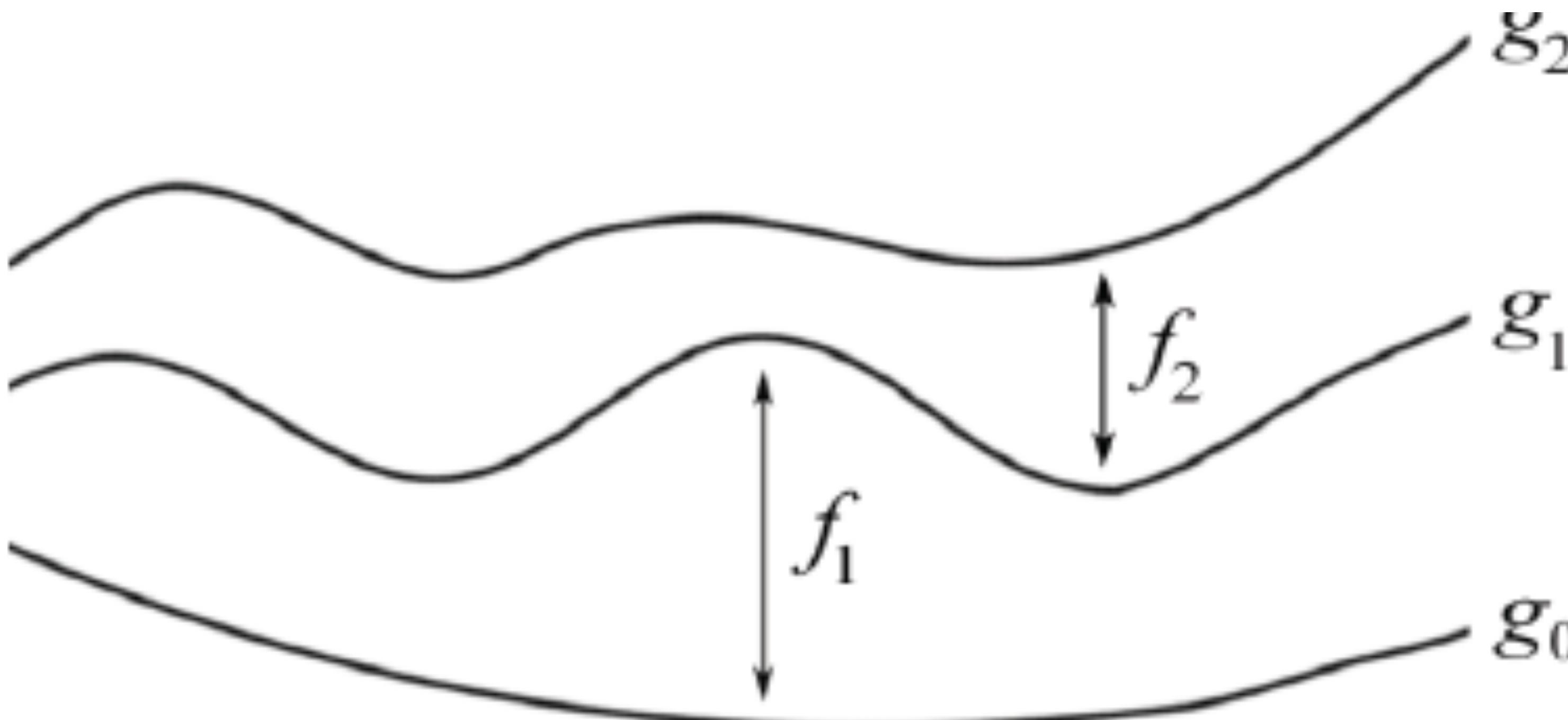


fig 4 – a visual description of stacked graph functions  $f_i$  and  $g_i$  for  $n=2$  as used in this section

Assim o topo do gráfico é dado por

$$g_i = g_0 + \sum_{j=1}^i f_j$$

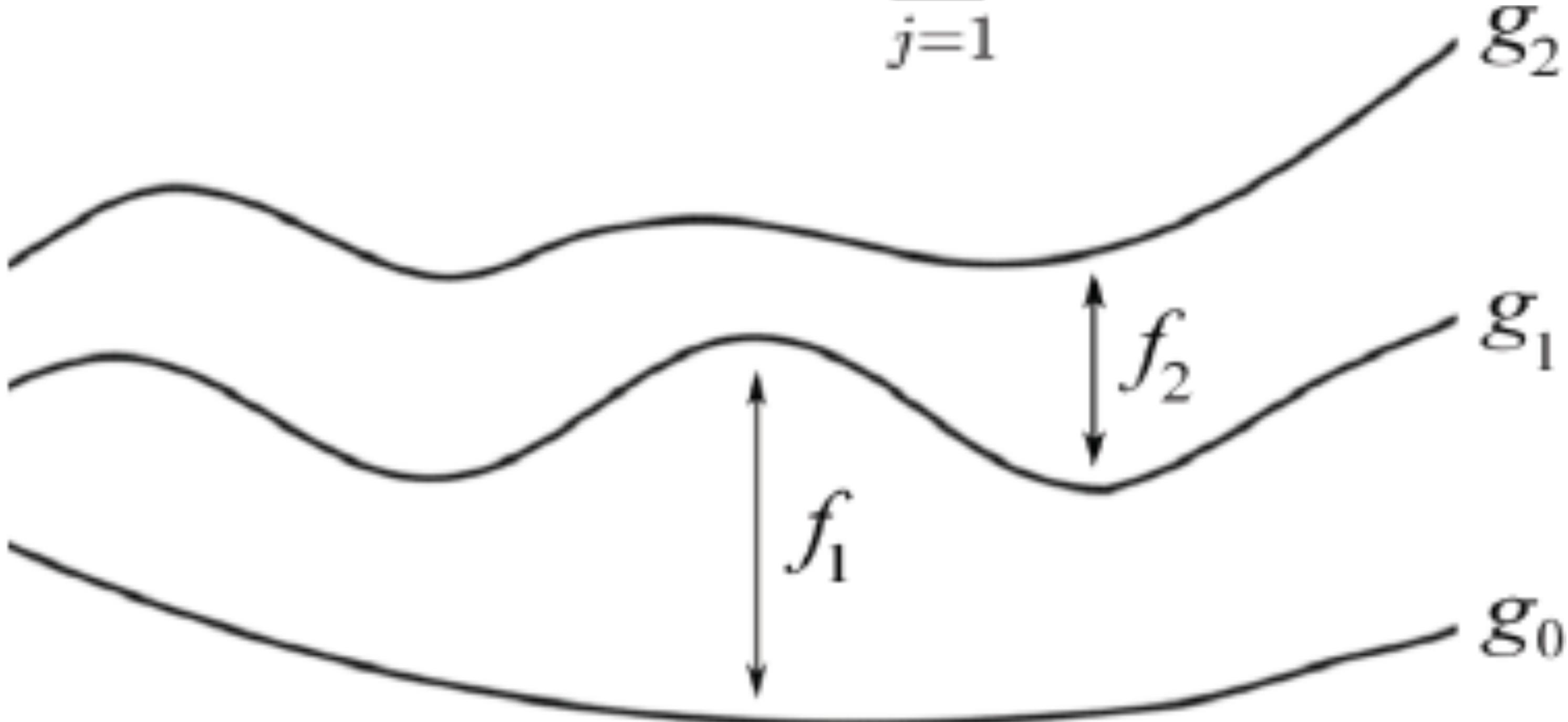
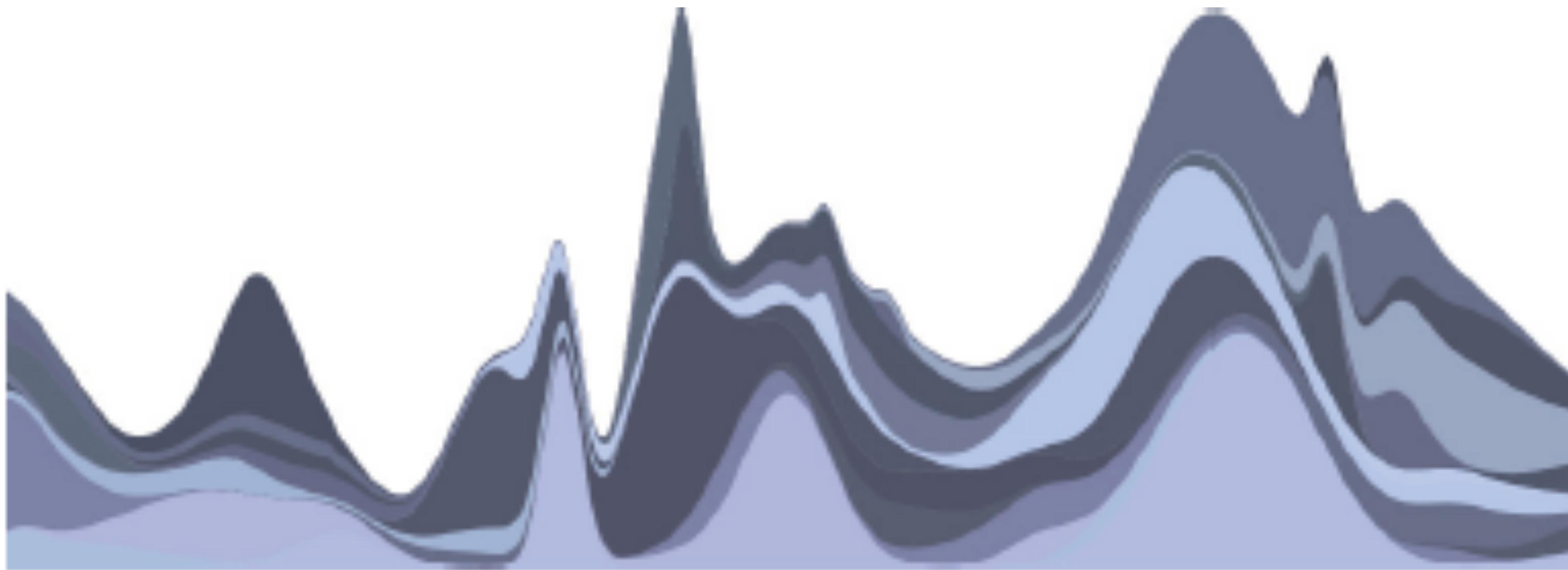


fig 4 – a visual description of stacked graph functions  $f_i$  and  $g_i$  for  $n=2$  as used in this section

Quando  $g_0 = 0$



*fig 5 – A traditional stacked graph with a baseline  $g_0 = 0$*

A soma das séries é legível mas a legibilidade de séries individuais fica prejudicada

É possível mostrar que este procedimento diminui a silhueta da curva deixando a curva mais suave

Os autores propõem **dois novos procedimentos** para melhorar a estética das curvas que consistem em obter  $g_0$  visando:

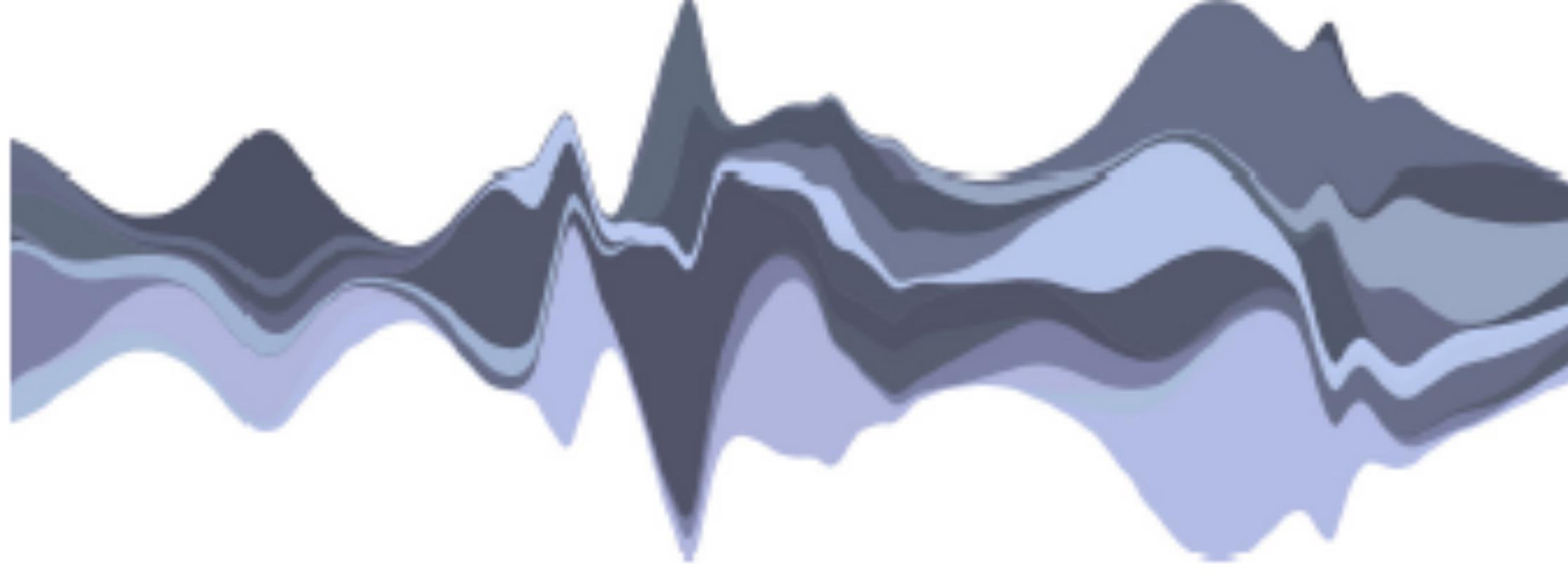
## 1. Minimizar a inclinação em cada valor de $x$

$$wiggle(g_0) = \sum_{i=0}^n g_i'^2 = \sum_{i=0}^n (g_0' + \sum_{j=1}^i f_j')^2$$

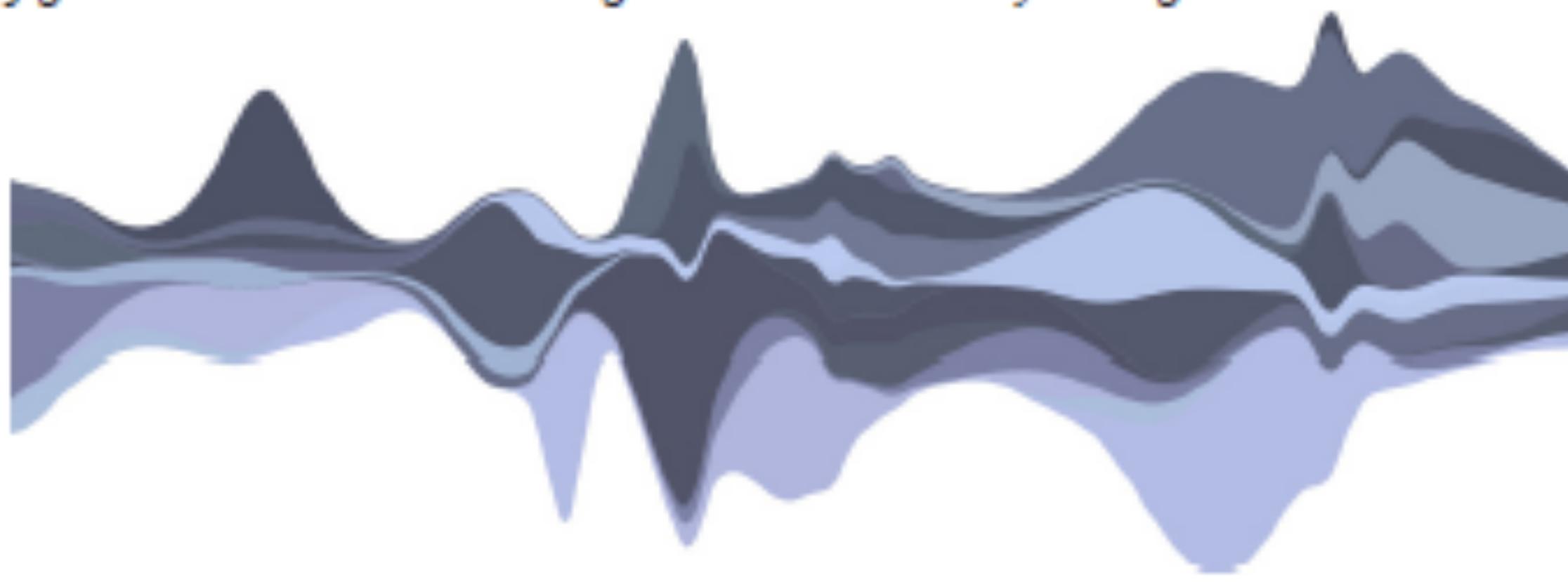
**2. Minimizar as médias dos quadrados das inclinações em pontos médios de cada camada ponderando pela espessura da camada**

$$weighted\_wiggle(g_0) = \sum_{i=1}^n \left( \frac{1}{2}(g'_i + g'_{i-1}) \right)^2 f_i = \sum_{i=1}^n \left( g'_0 + \frac{1}{2}f'_i + \sum_{j=1}^{i-1} f'_j \right)^2 f_i$$

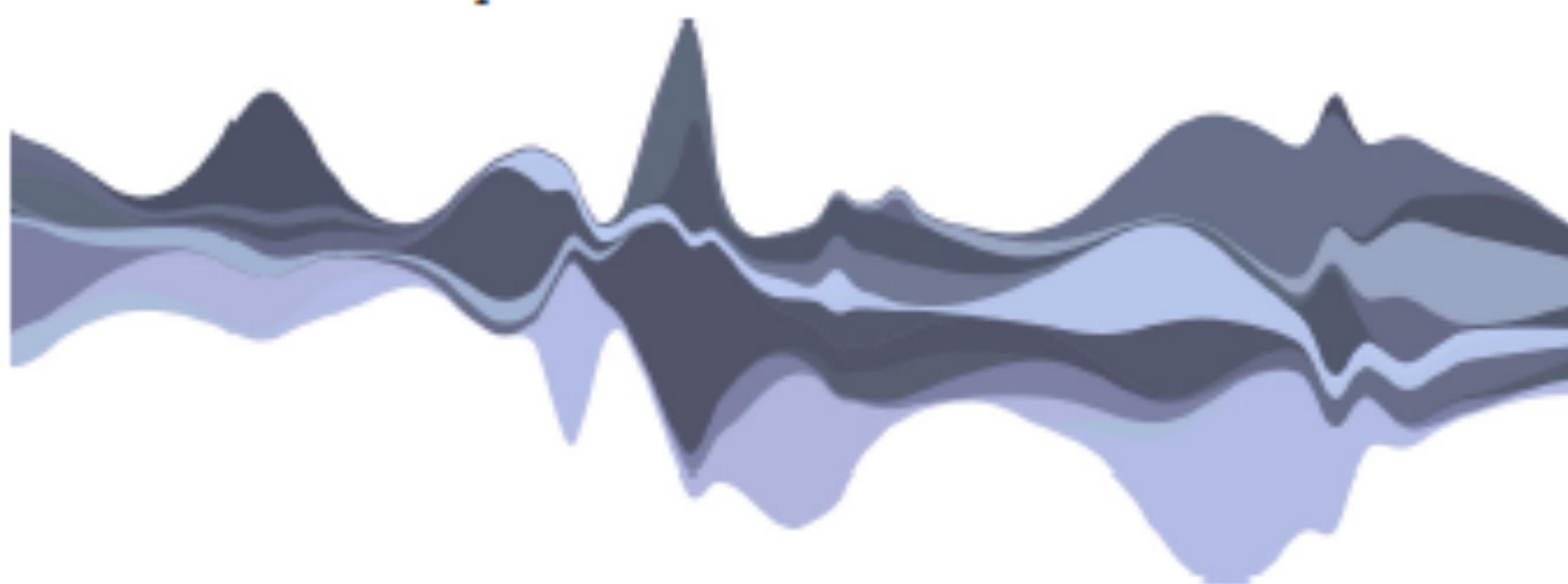
$$g'_0 = -\frac{1}{\sum f_i} \sum_{i=0}^n \left( \frac{1}{2}f'_i + \sum_{j=1}^{i-1} f'_j \right) f_i$$



*fig 6 – the same data set using the ThemeRiver layout algorithm*



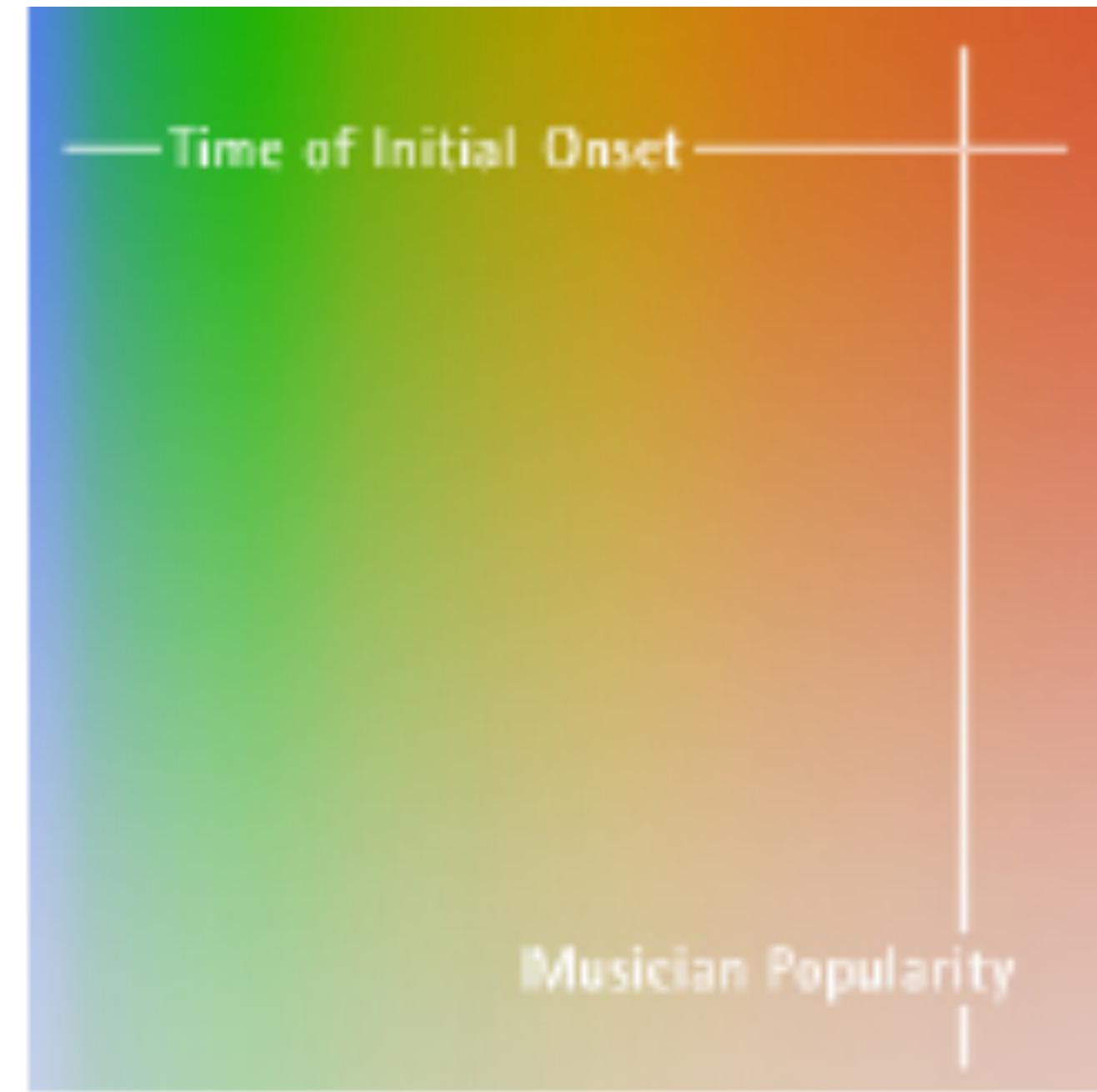
*fig 7 – the same data set optimized to reduce the "wiggle" function, or overall variation in slope*



*fig 8 – the same data set optimized to reduce the "weighted\_wiggle," the algorithm used in Streamgraph*

Paleta de cores gerada com o uso do Photoshop através de **imagens altamente saturadas da natureza**

Não usam todos os matizes existentes mas selecionam cores altamente contrastantes para revelar os diferentes comportamentos das séries

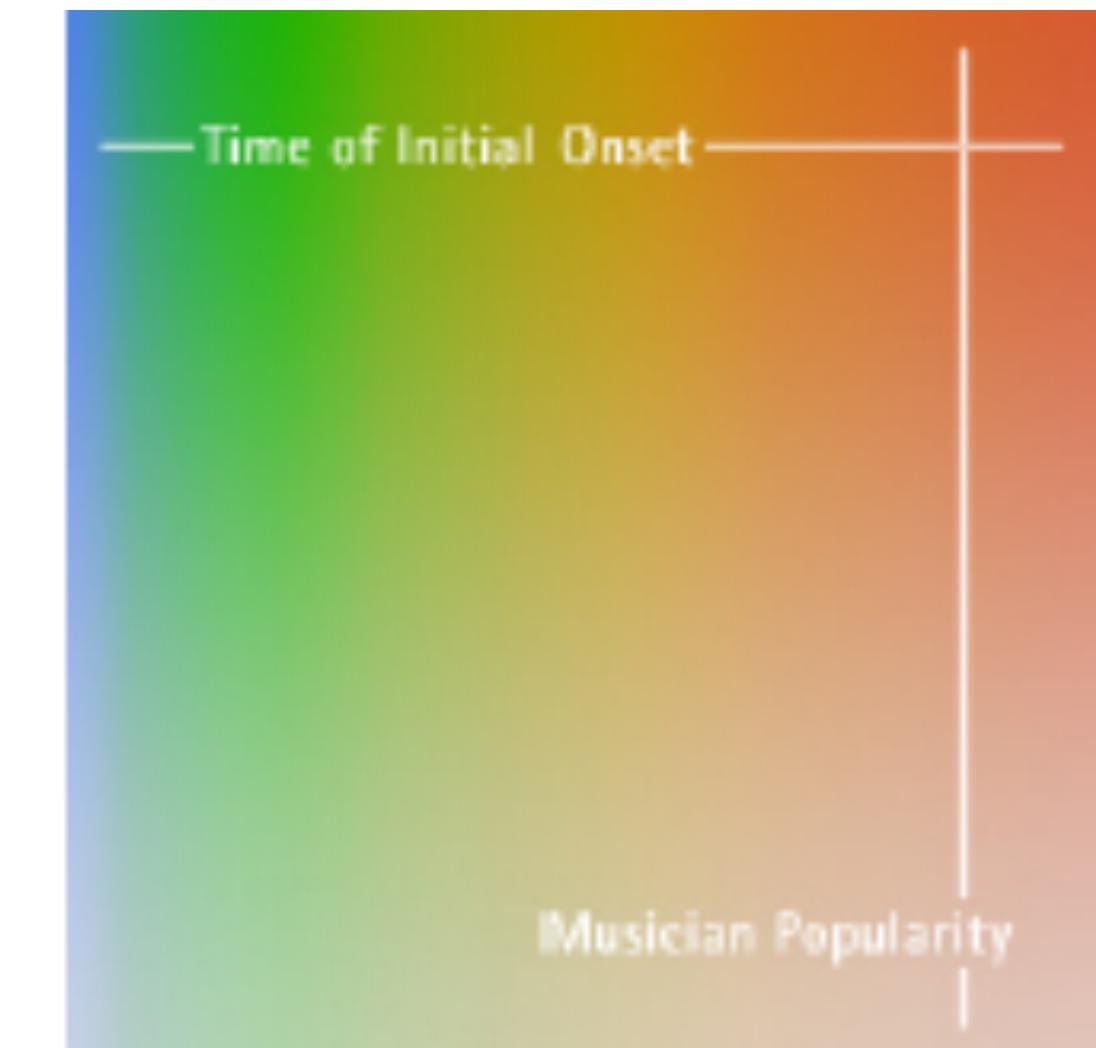


*fig 9 – the 2d color palate used in Listening History*

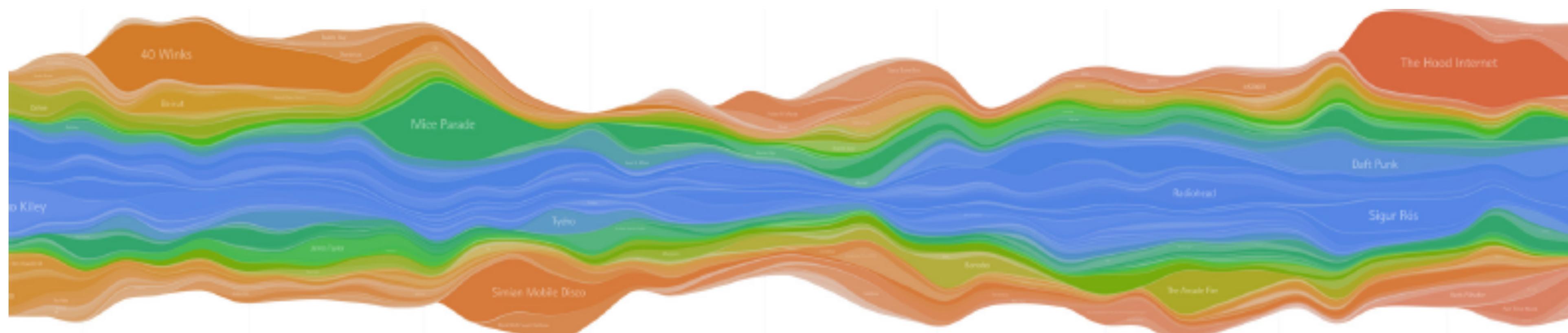
A faixa horizontal varia de cores frias a quentes e reflete a data em que o artista foi descoberto pelo usuário

“Núcleo frio” de artistas conhecidos e as camadas “quentes e novas” das descobertas recentes

A faixa vertical representa a frequência



*fig 9 – the 2d color palate used in Listening History*

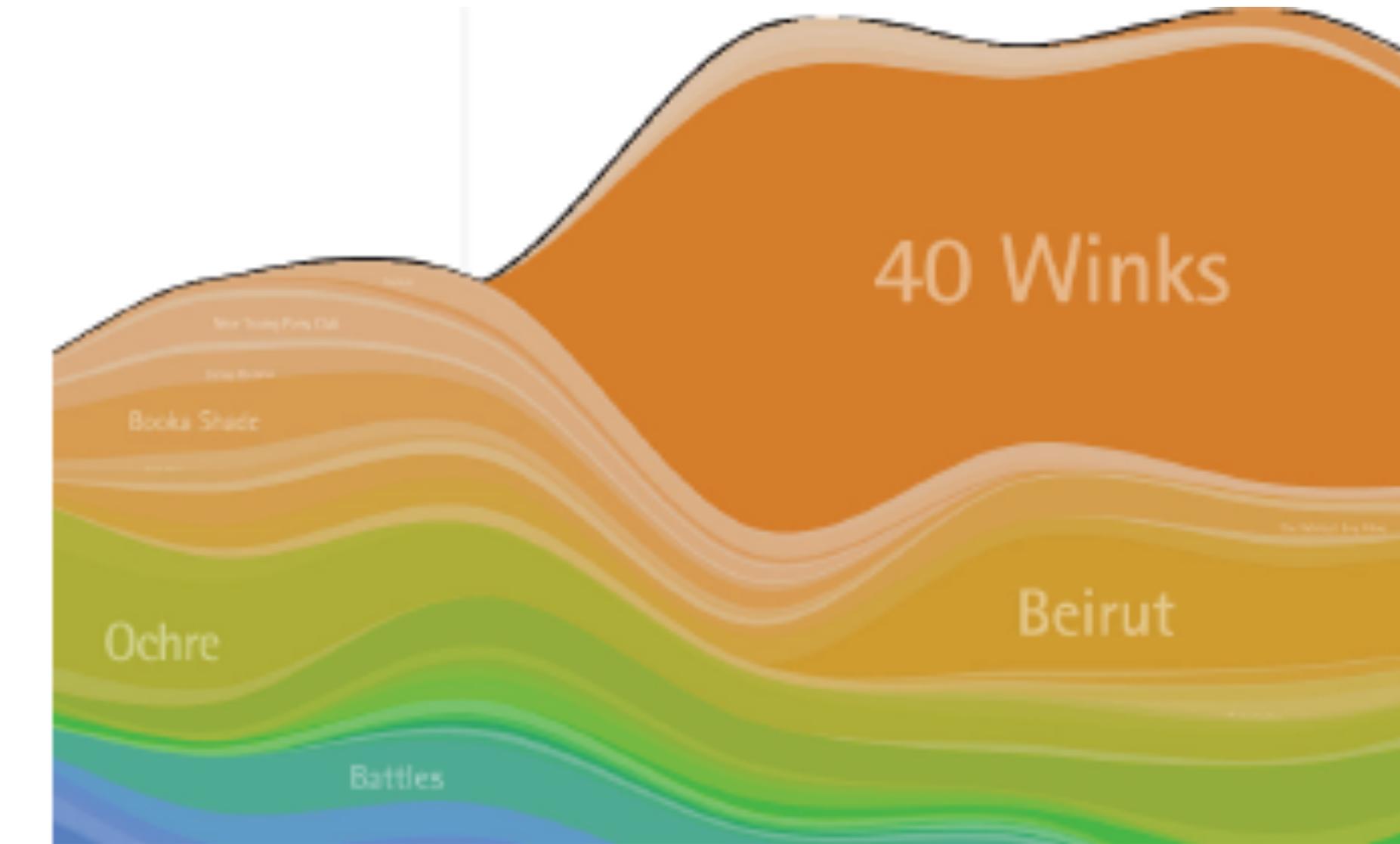


Muitas cores = dificuldade no uso de legendas

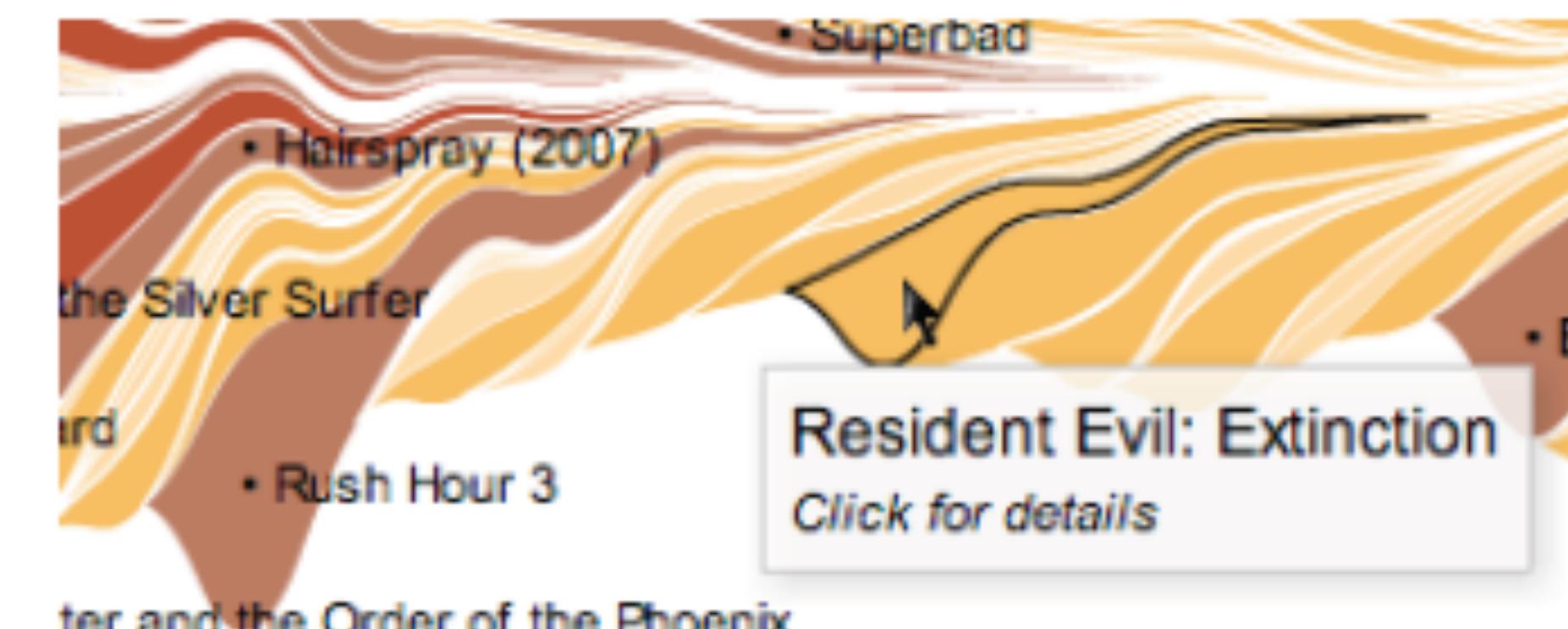
Fonte de cor branca com transparência para conectar o rótulo a camada

Escolha da posição da camada onde poderia ser usada a maior fonte possível usando algoritmo de força bruta (off-line)

Na versão on-line, rótulos nos eventos do mouse



*fig 10 – a detailed look at the labeling strategy of Listening History*



*fig 11 – detail of a roll-over label in the New York Times graphic*

# EVOLUTIONARY COMPUTATION FOR LABEL LAYOUT ON UNUSED SPACE OF STACKED GRAPHS

---

*A. Toledo, K. Sookhanaphibarn, R. Thawonmas e F.  
Rinando*

*ISRN Artificial Intelligence  
2012*

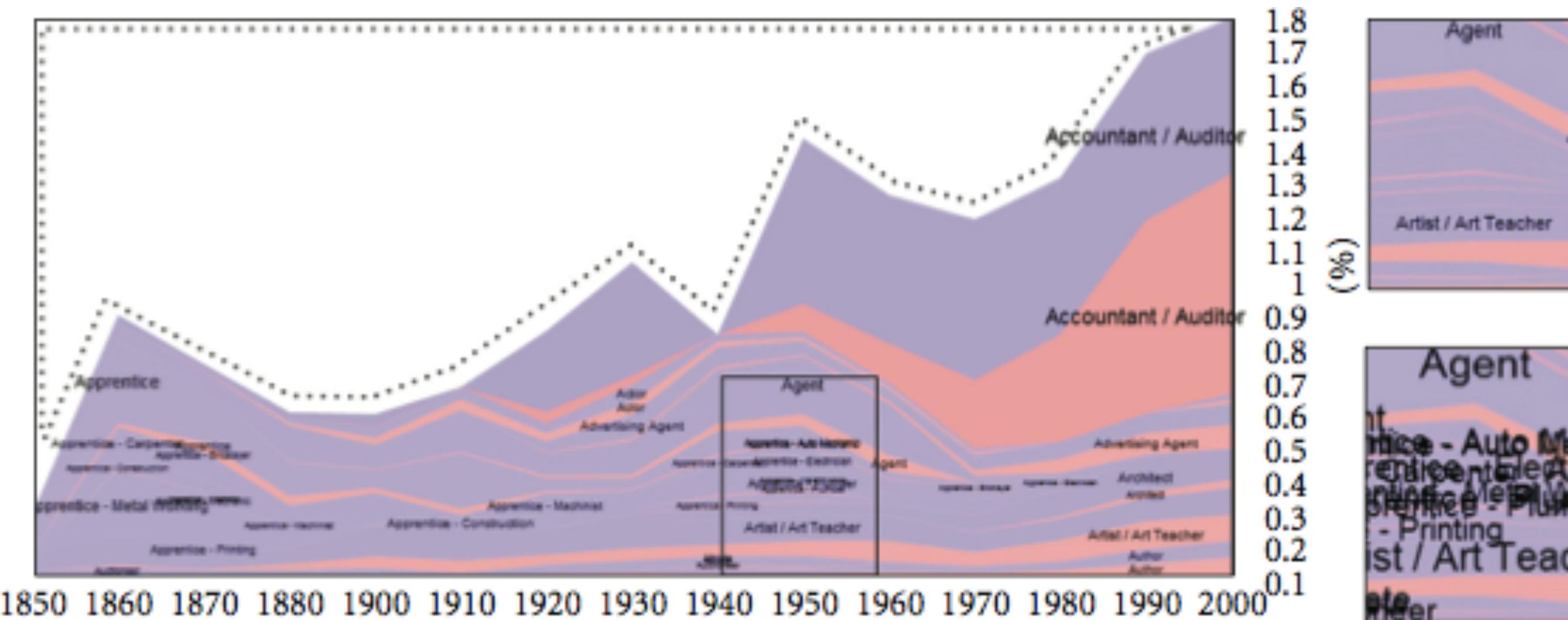


FIGURE 1: Stacked graph view (left) showing 40 job occupation names starting with the prefix “A”, and  $TLV = 0$  pixels. A number of labels are illegible due to either overlapping or small size. The solid line encloses the visualization area, and the dotted line the unused space. At the right, an increased section of the same view (top) with  $TLV = 12$  pixels (some labels are hidden). The same section (bottom) can be seen with  $TLV = 0$  (some labels are overlapped) and with their minimum font size increased three times.

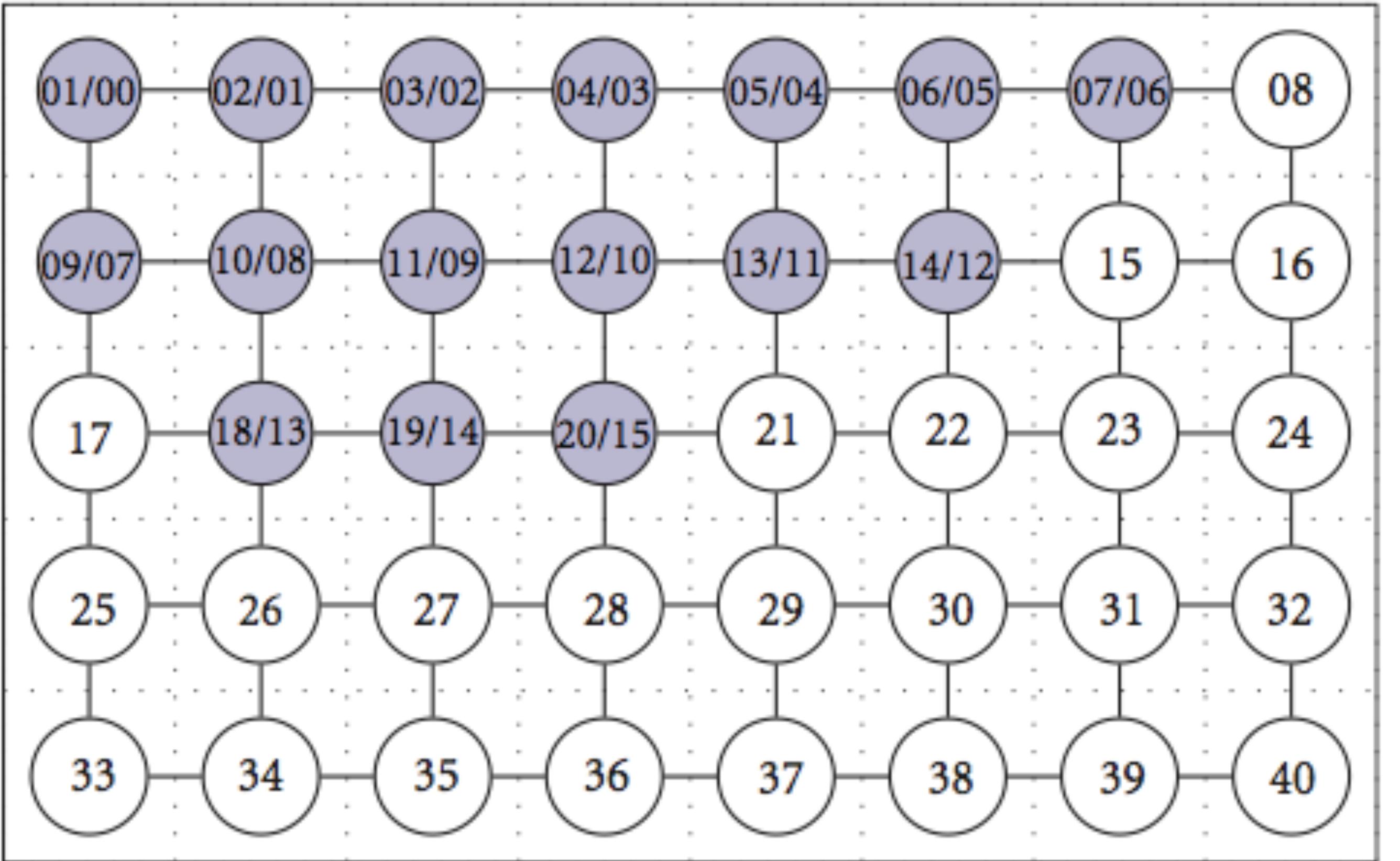


FIGURE 2: Partition of the visualization area of the stacked graph of Figure 1. Partitions in blue color correspond to the candidate partitions for label layout.

Chromosome A: 0000 0110 0010 1000 0111 1001

Chromosome B: 1001 0101 0110 0011 0001 0010

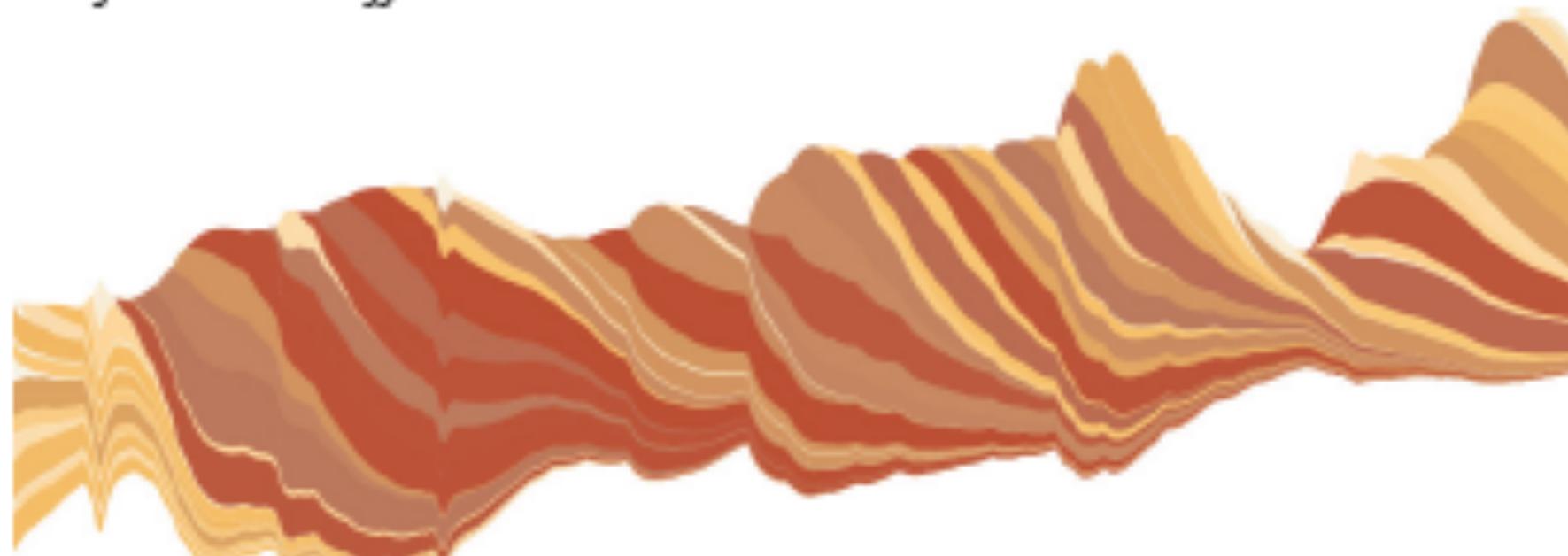
# Ordenação das camadas

Em alguns casos a ordenação natural dos dados é a mais apropriada

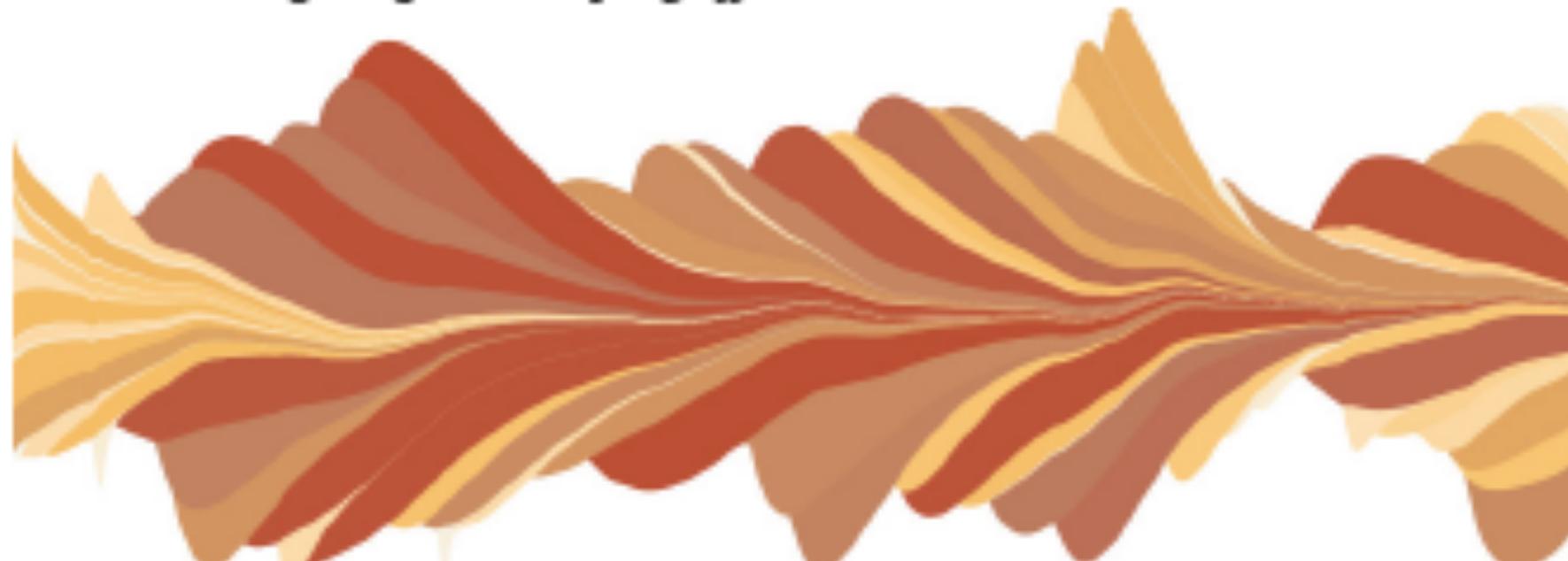
Neste caso, a ordenação influencia na aparência do gráfico e pode ser escolhida com objetivos estéticos



*fig 12 – an unsorted data set, exhibiting the type of “burstiness” apparent in last.fm and box office data sets*



*fig 13 – the same data set, naively sorted in order of “onset time” exhibiting the distracting diagonal striping effect*



*fig 14 – the same data set sorted using the weighted “inside out” strategy to highlight the initial onset of each time series*

# Ordenação das camadas

Camadas **desordenadas**



fig 12 – an unsorted data set, exhibiting the type of “burstiness” apparent in last.fm and box office data sets

Camadas ordenadas

por **onset time**

(descoberta do artista)

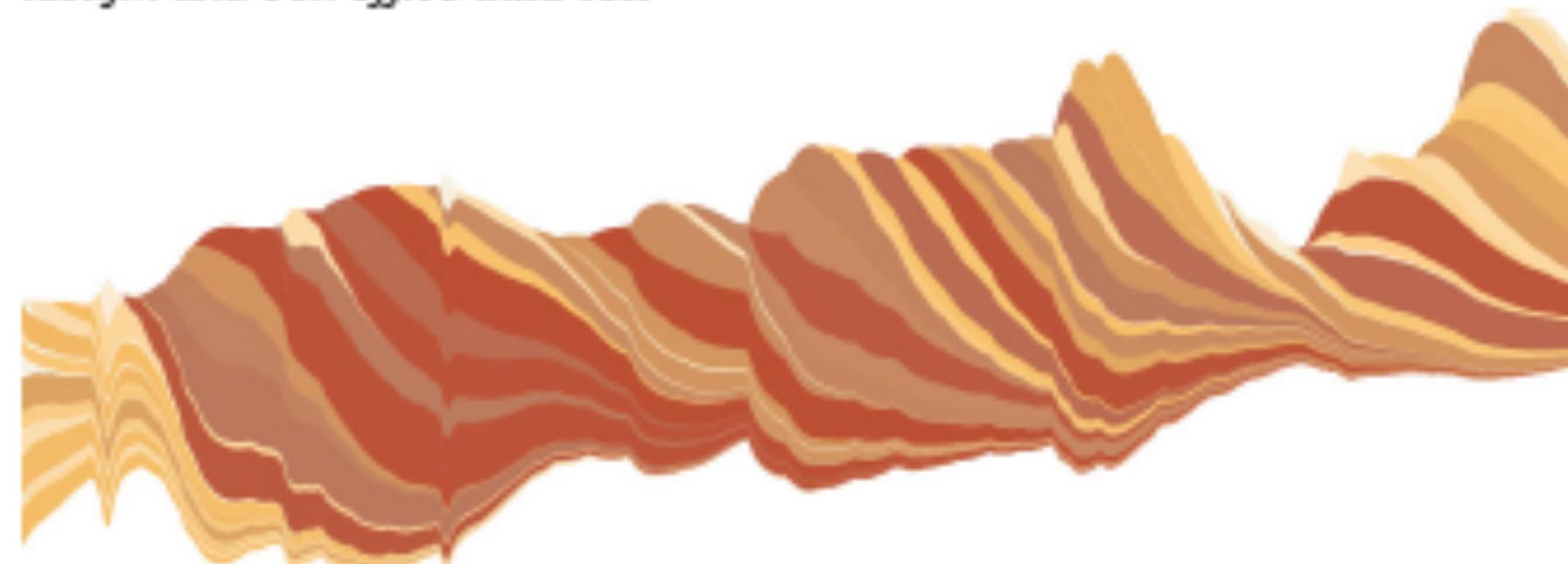


fig 13 – the same data set, naively sorted in order of “onset time” exhibiting the distracting diagonal striping effect

Camadas ordenadas

por **inside-out**

(onset antigo no meio e onset recente em baixo e em cima)



fig 14 – the same data set sorted using the weighted “inside out” strategy to highlight the initial onset of each time series

# Ordenação das camadas

Camadas ordenadas  
por ***inside-out***

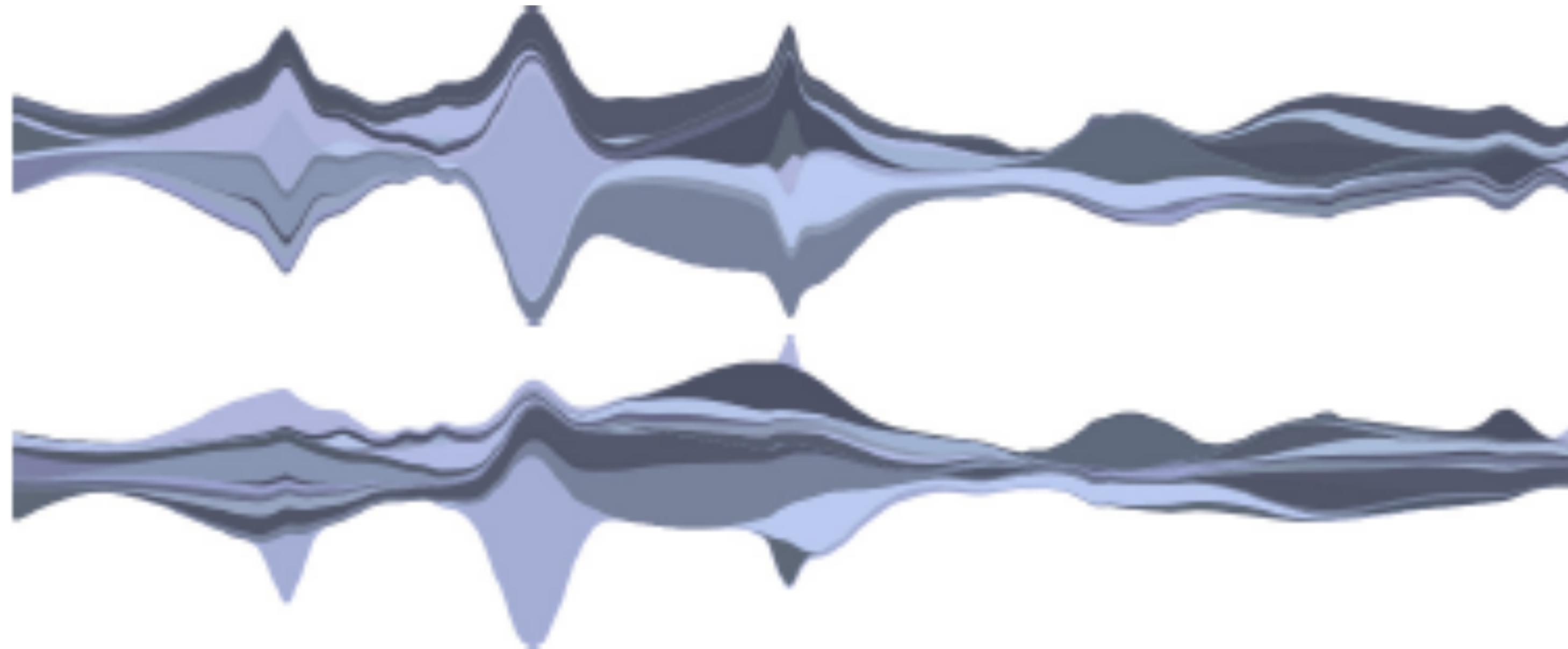
## Solução:

1. Ordenar as camadas por *onset time*
2. Para cada camada, calcular a soma dos valores da série (peso)
3. Insere no início ou fim da lista dependendo se a soma do peso da primeira metade da lista é maior que a metade do peso total.

## Ordenação das camadas

Outra possibilidade é através de uma métrica de **volatilidade**

Séries com menos modificações no centro e séries com mais modificações no exterior



*fig 15 – before and after applying an “inside out” sort using a “measure of volatility” in place of “onset time”*

# Séries temporais - Parte III

---

**Profa. Dra. Raquel Minardi**

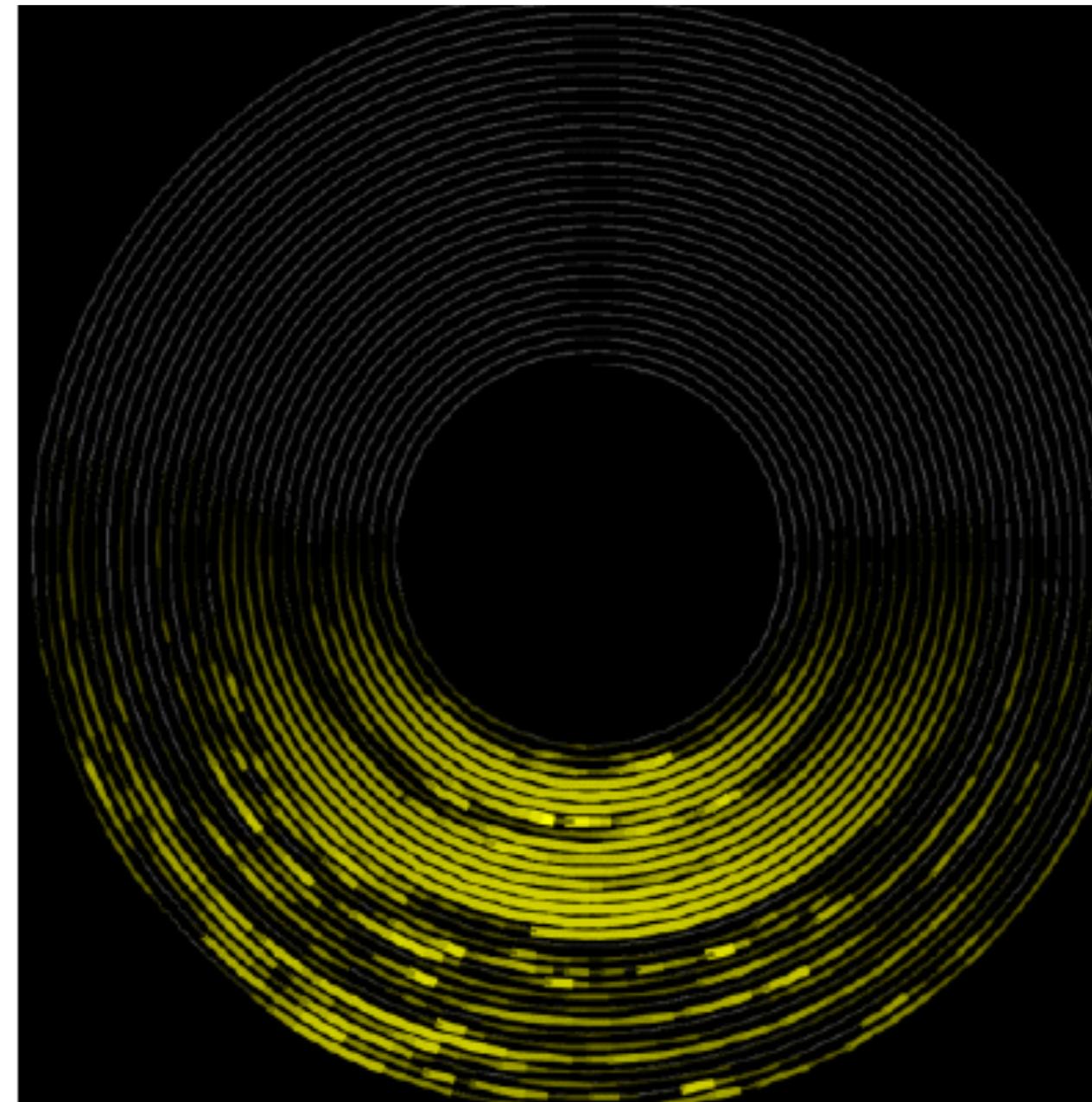
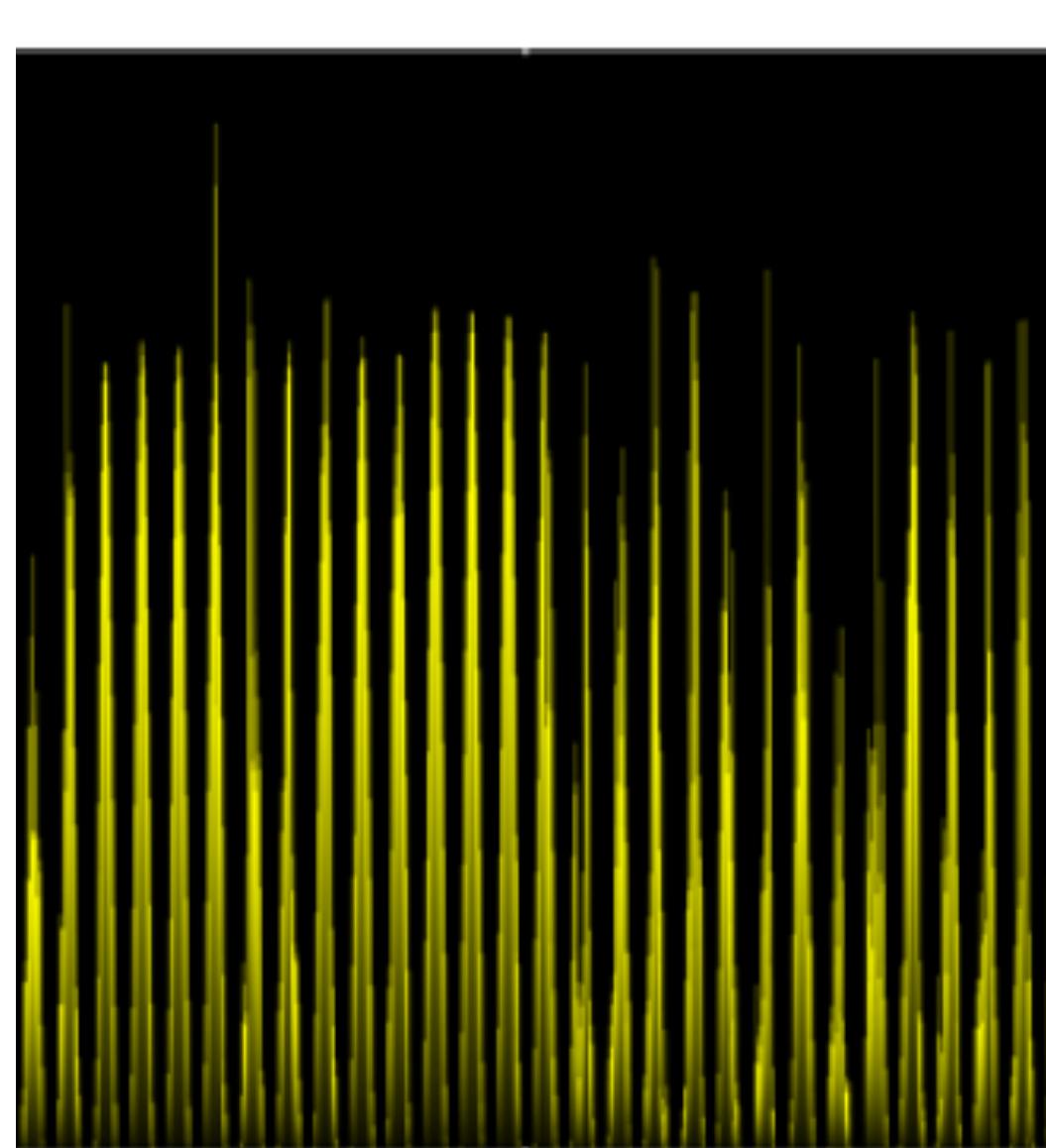
Departamento de Ciência da Computação

Universidade Federal de Minas Gerais

# VISUALIZING TIME-SERIES ON SPIRALS

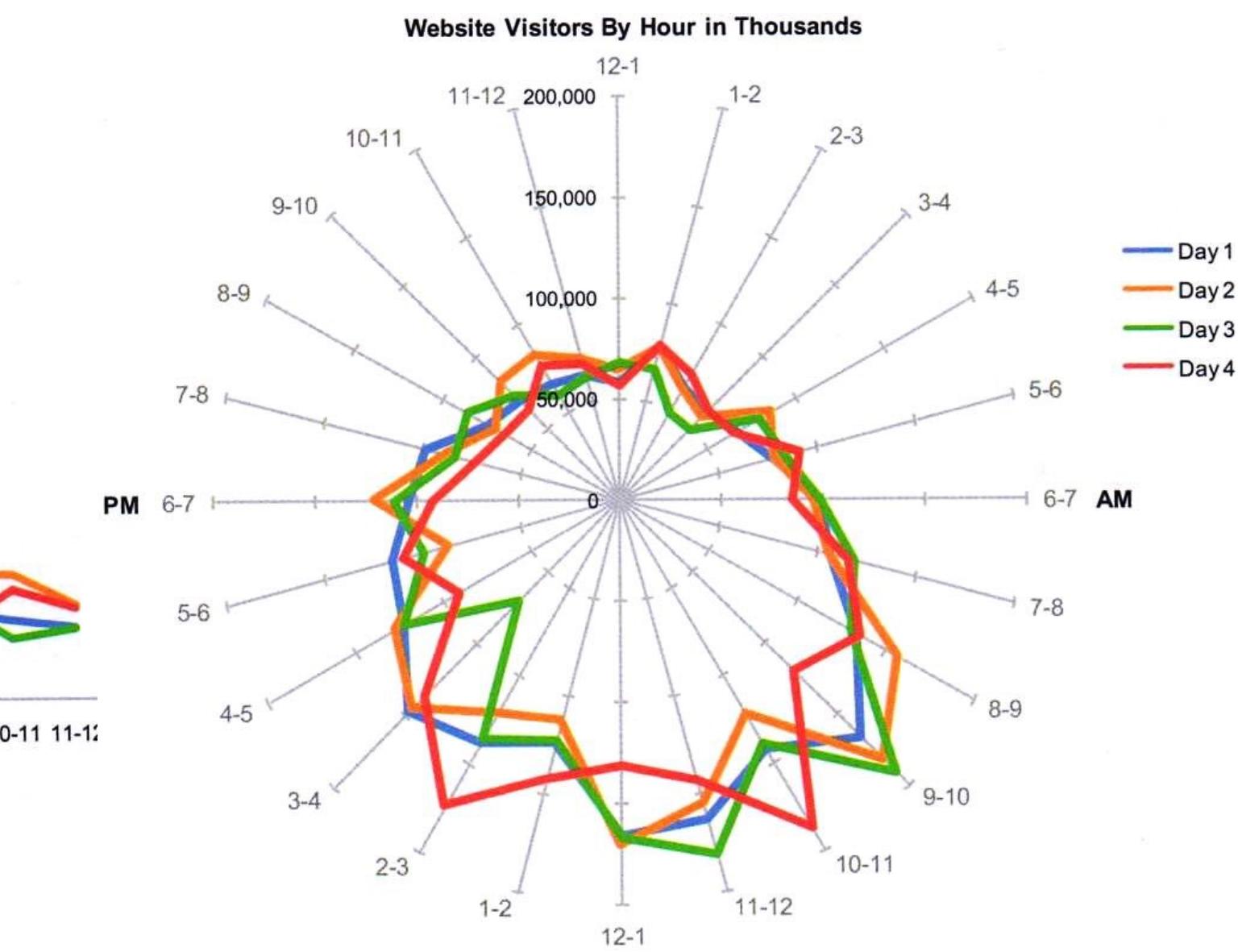
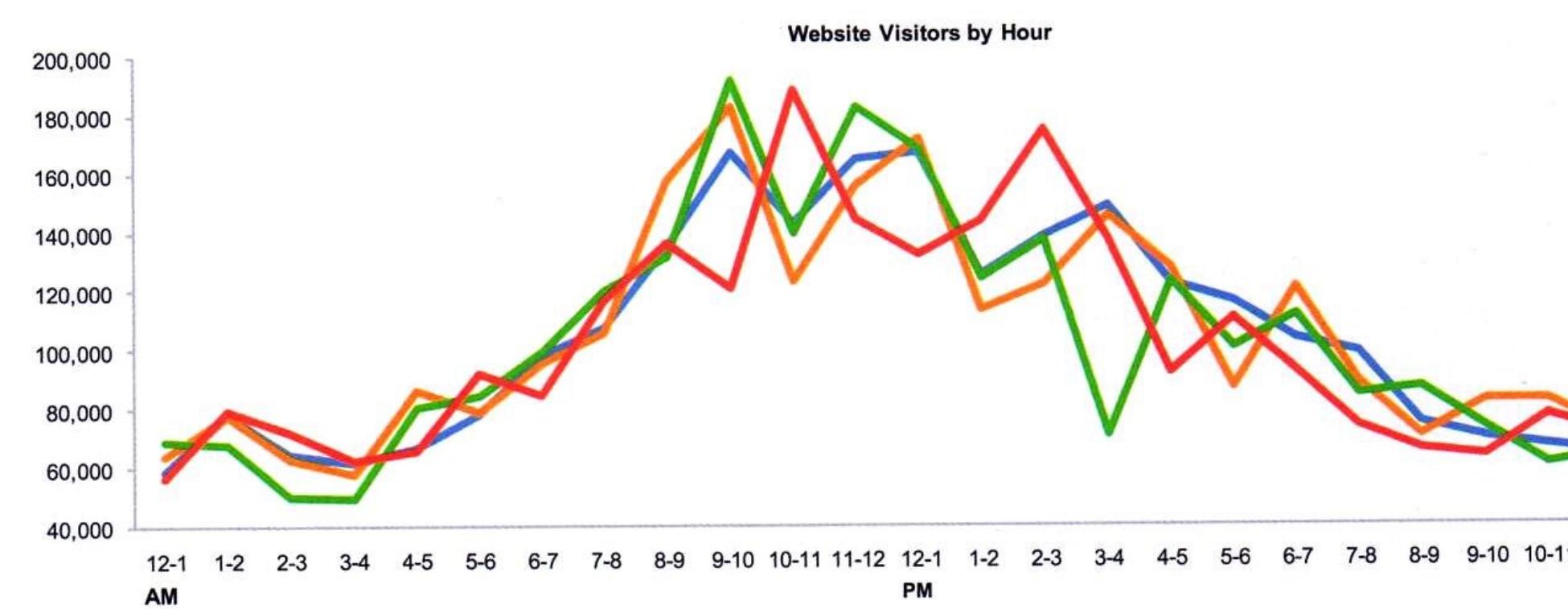
---

*M. Weber, M. Alexa e W. Müller  
IEEE Symposium on Information Visualization  
2001*



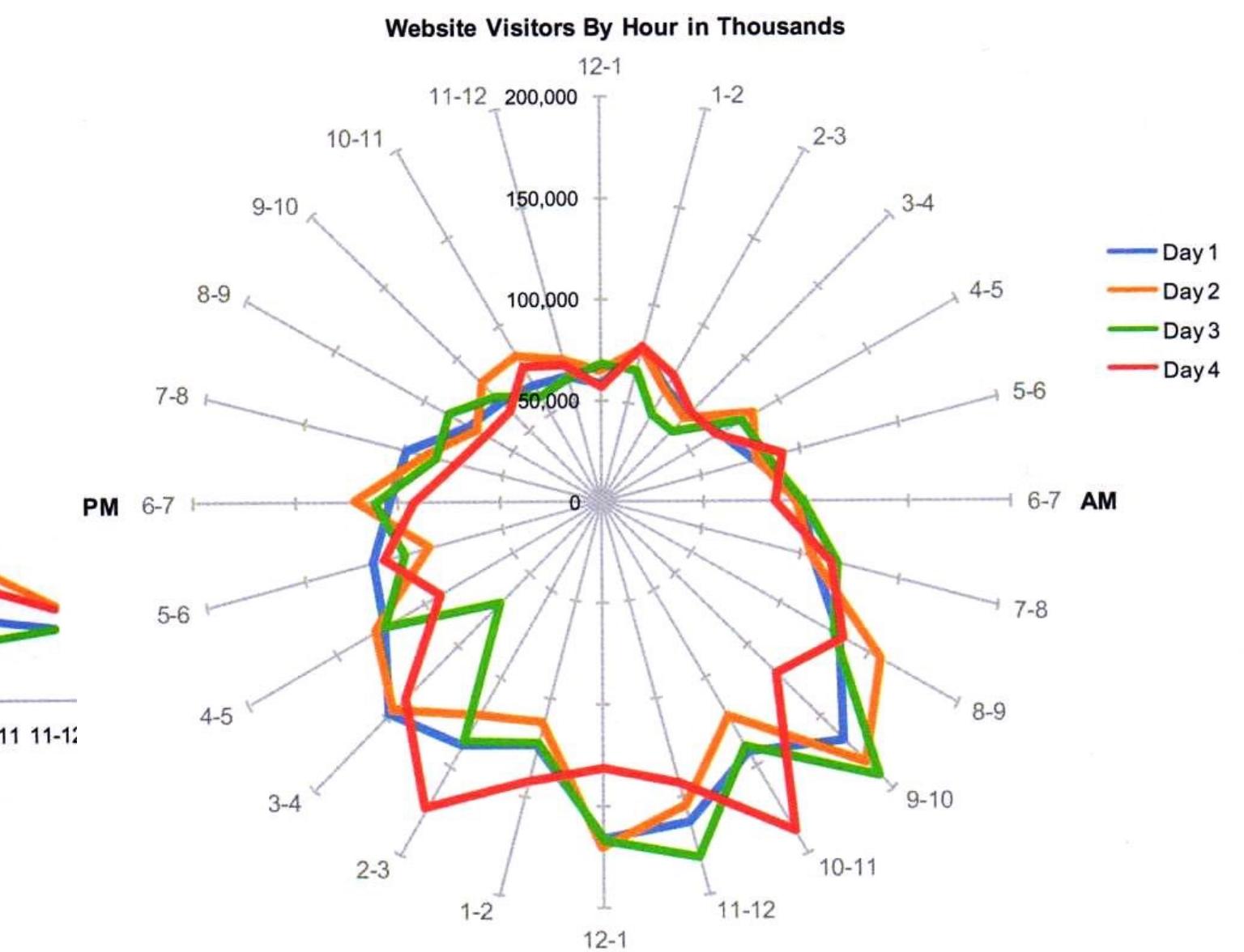
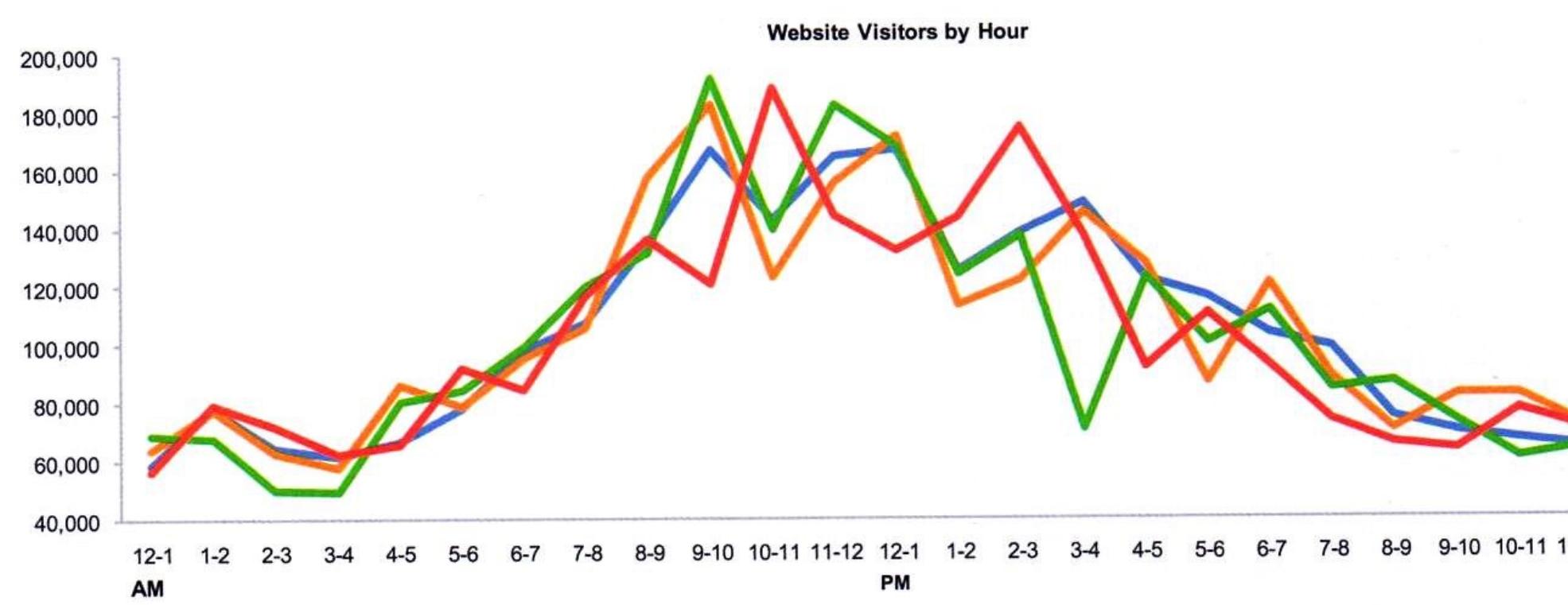
# Gráficos de linhas revelam tendências em relacionamentos temporais

Gráficos em ciclo mapeiam linhas em círculos com comportamento periódico com duração conhecida



# Limitações

Podem mostrar algo entre 2 e 8 diferentes séries  
É preciso conhecer a duração do ciclo



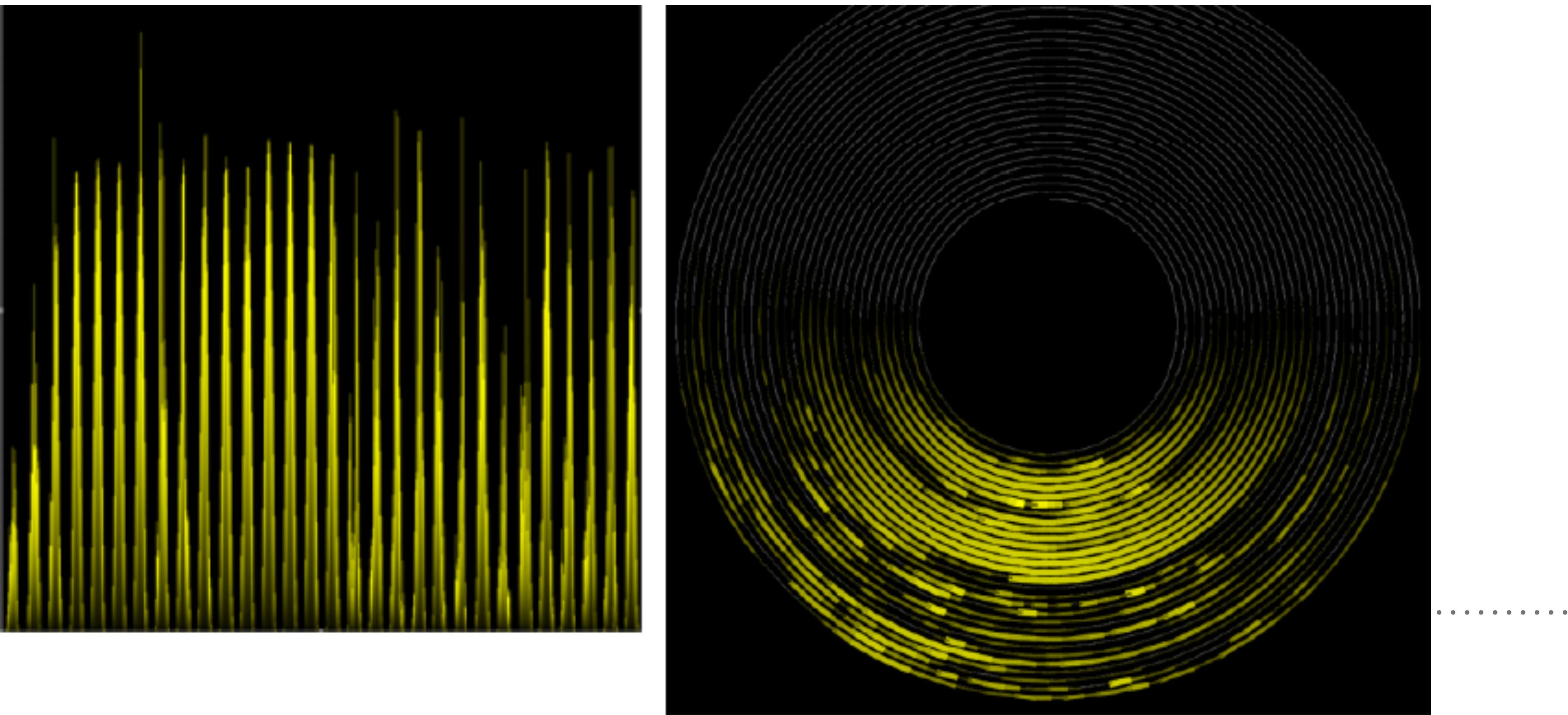
Segundo Weber et al., uma visualização de séries temporais deve:

Possibilitar a **comparação** de dados nominais e quantitativos

Suportar a visualização de **conjuntos de dados volumosos**

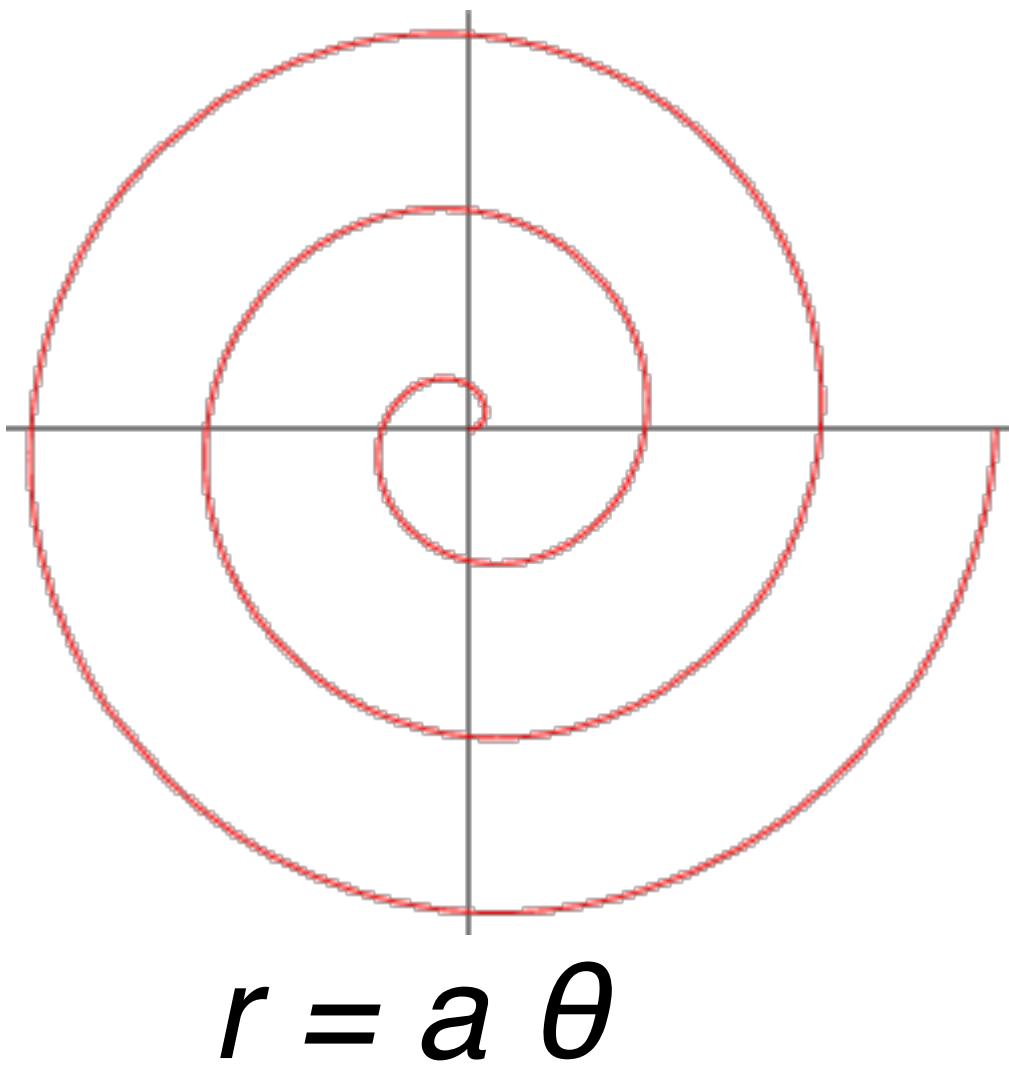
Suportar a **comparação de ciclos** em um conjunto de dados

Auxiliar na **identificação de padrões** de intensidades e das durações dos ciclos



**Figure 1:** Two visualizations of sunshine intensity using about the same screen real estate and the same color coding scheme. In the spiral visualization it is much easier to compare days, to spot cloudy time periods, or to see events like sunrise and sunset.

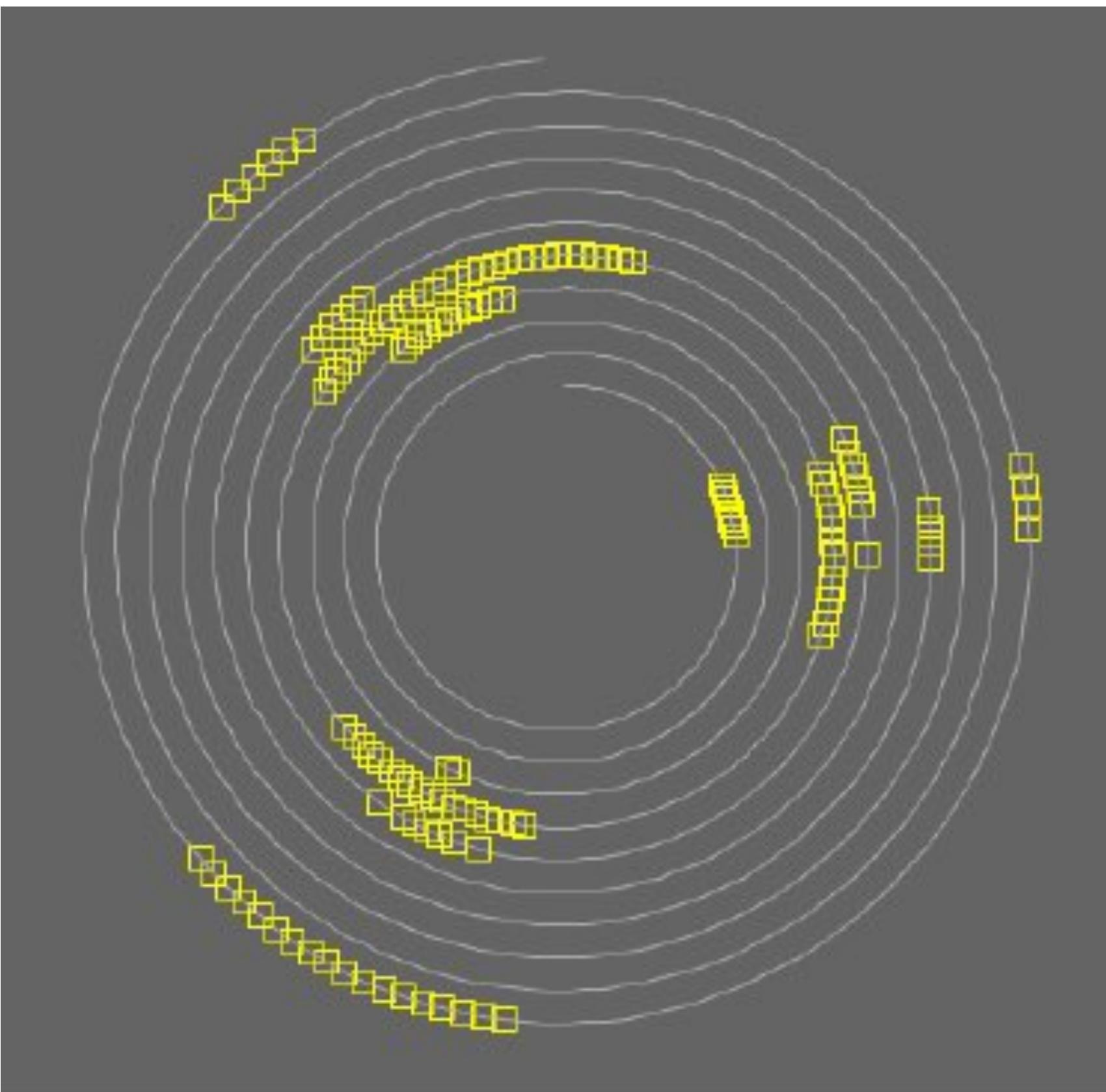
## Espiral de Arquimedes



Dados de diferentes períodos normalmente têm a mesma importância, logo a espiral de Arquimedes é a mais apropriada

# REPRESENTAÇÃO VISUAL

---

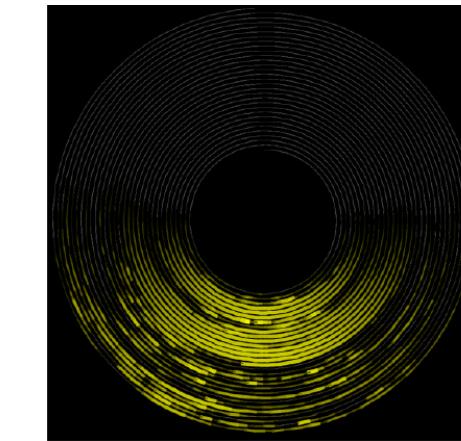


**Figure 2:**Nominal time-series data presented on a Spiral. The periodic behaviour of the underlying process is revealed.

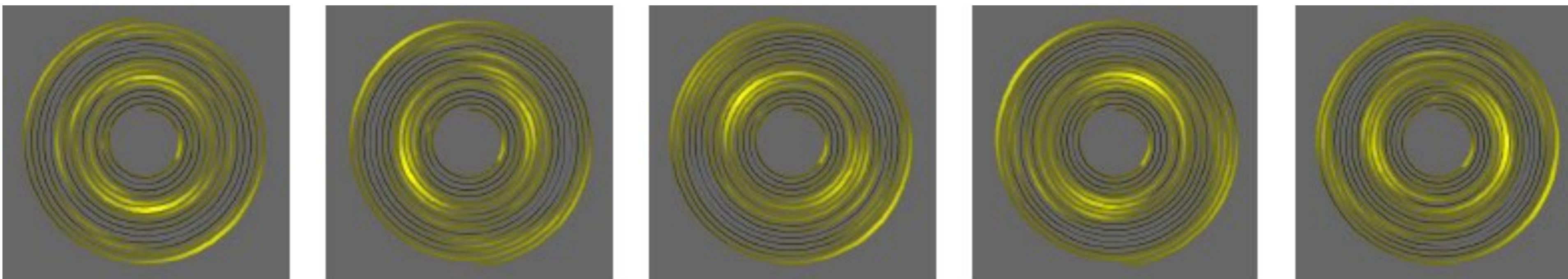
- Cores das linhas
- Estilo das linhas
- Espessura das linhas
- Uso de ícones

# DETECÇÃO DE CICLOS

---



- Em muitos casos, conhecemos os ciclos a serem analisados e a visualização tem grande potencial de evidenciar o comportamento periódicos dos dados
- Quando estes ciclos não são conhecidos, os períodos devem ser inferidos dos próprios dados
- Pode-se usar animações para detectar o comprimento do ciclo

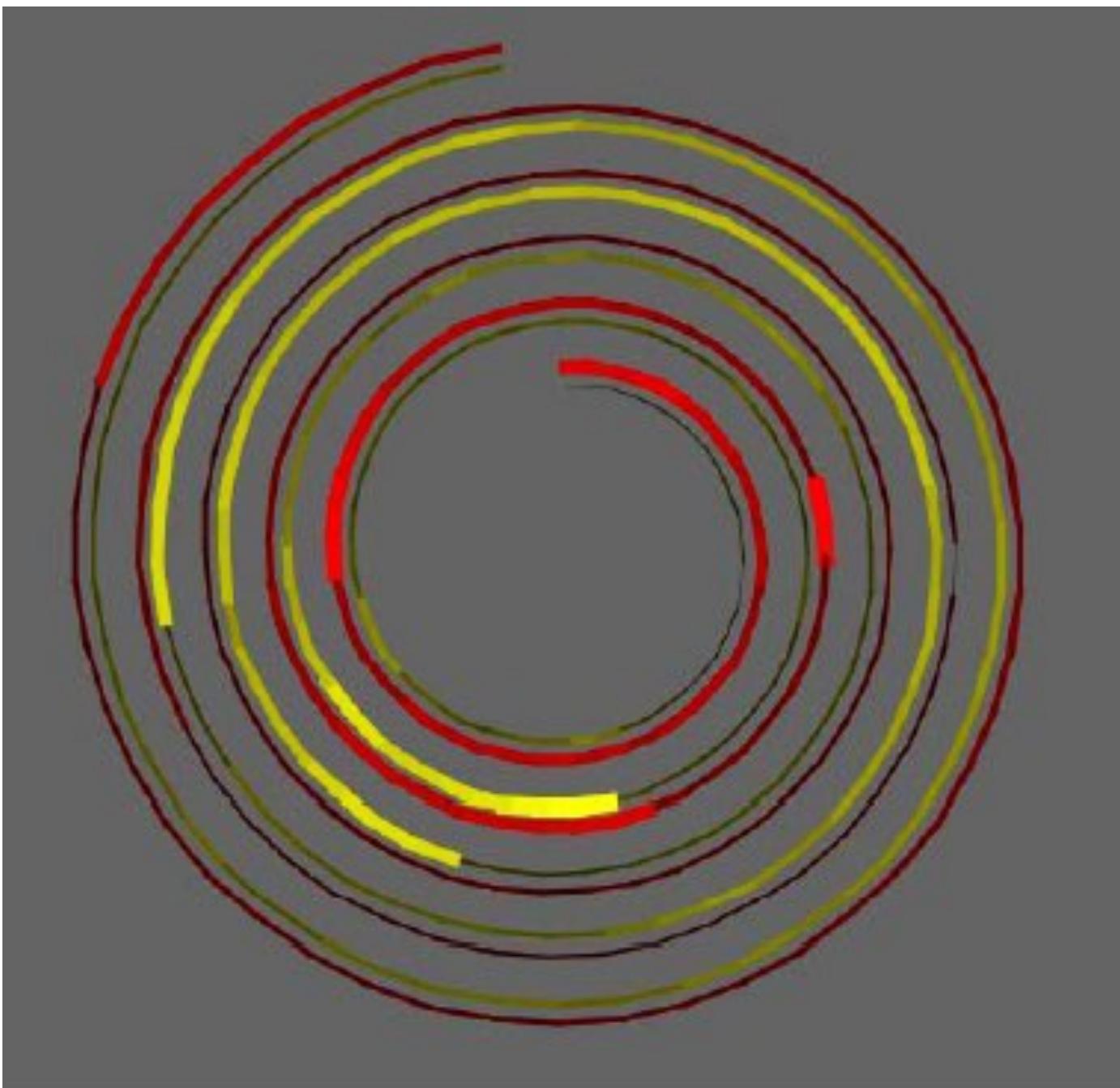


**Figure 3:** Visualizations of the same data with continuously changing cycle length. The period in the data can be found visually, i.e. the visual system is used to detect periodic patterns in the data exploiting the spatial layout on the spiral.

# MULTI-ESPIRAL

---

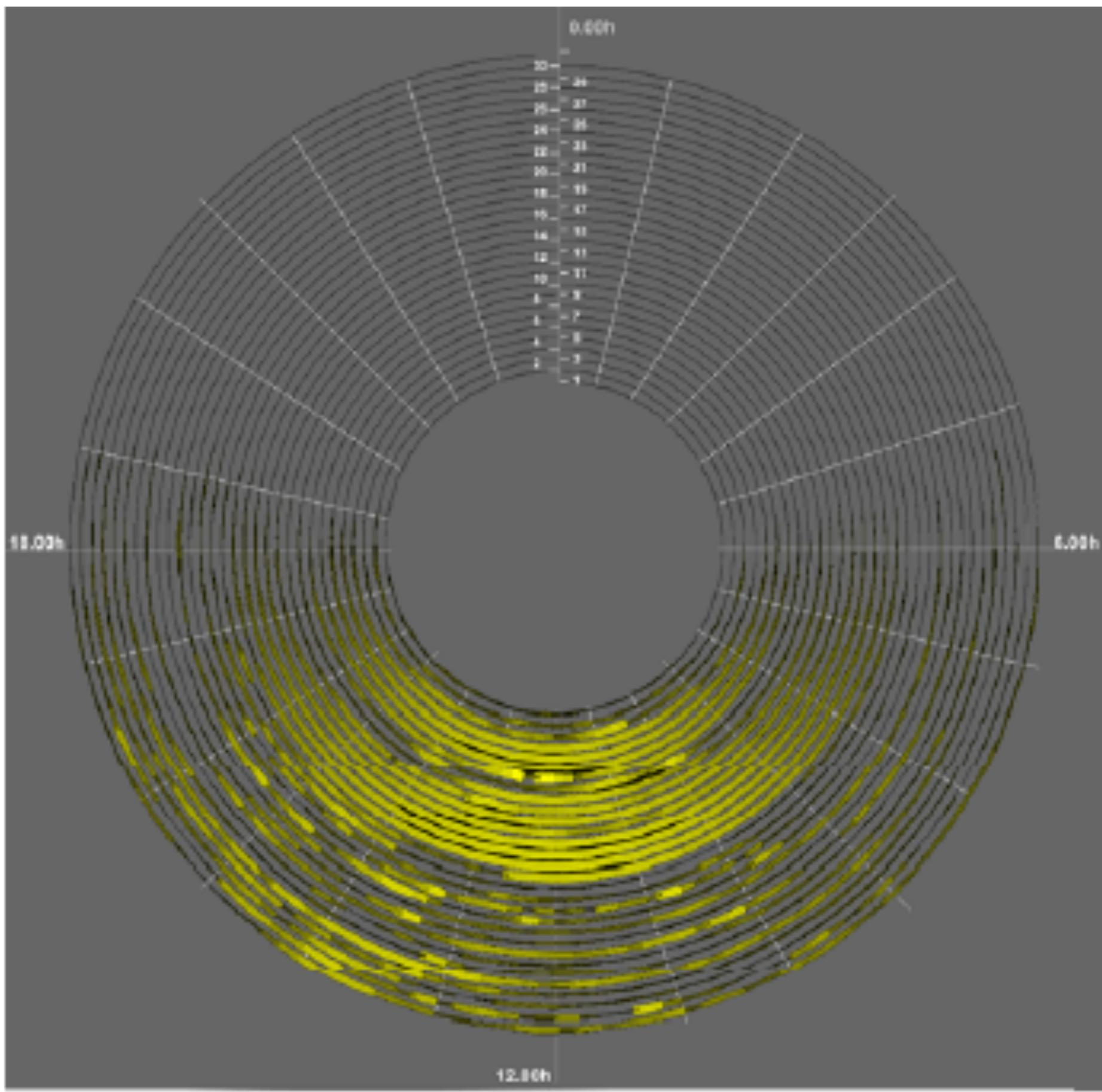
- Uso de diferentes cores para comparação entre séries
- Possível comparar até 8 séries



**Figure 4:** Stock prices of Microsoft (yellow) and Sun Microsystems (red) in five years on parallel spirals.

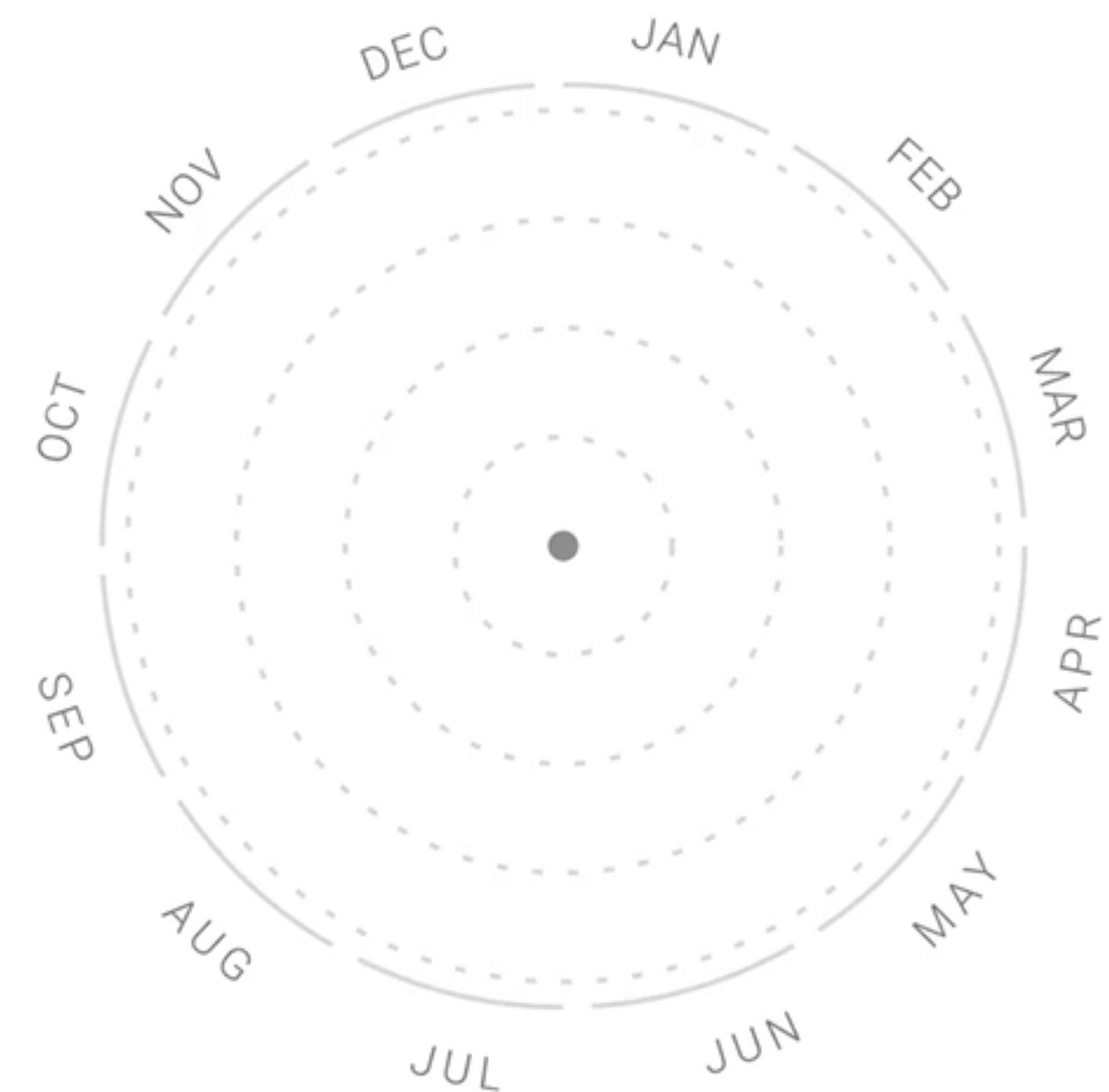
# ESCALAS

---



**Figure 5:**A possible way to add informative scales to the parametric dimensions of a spiral.

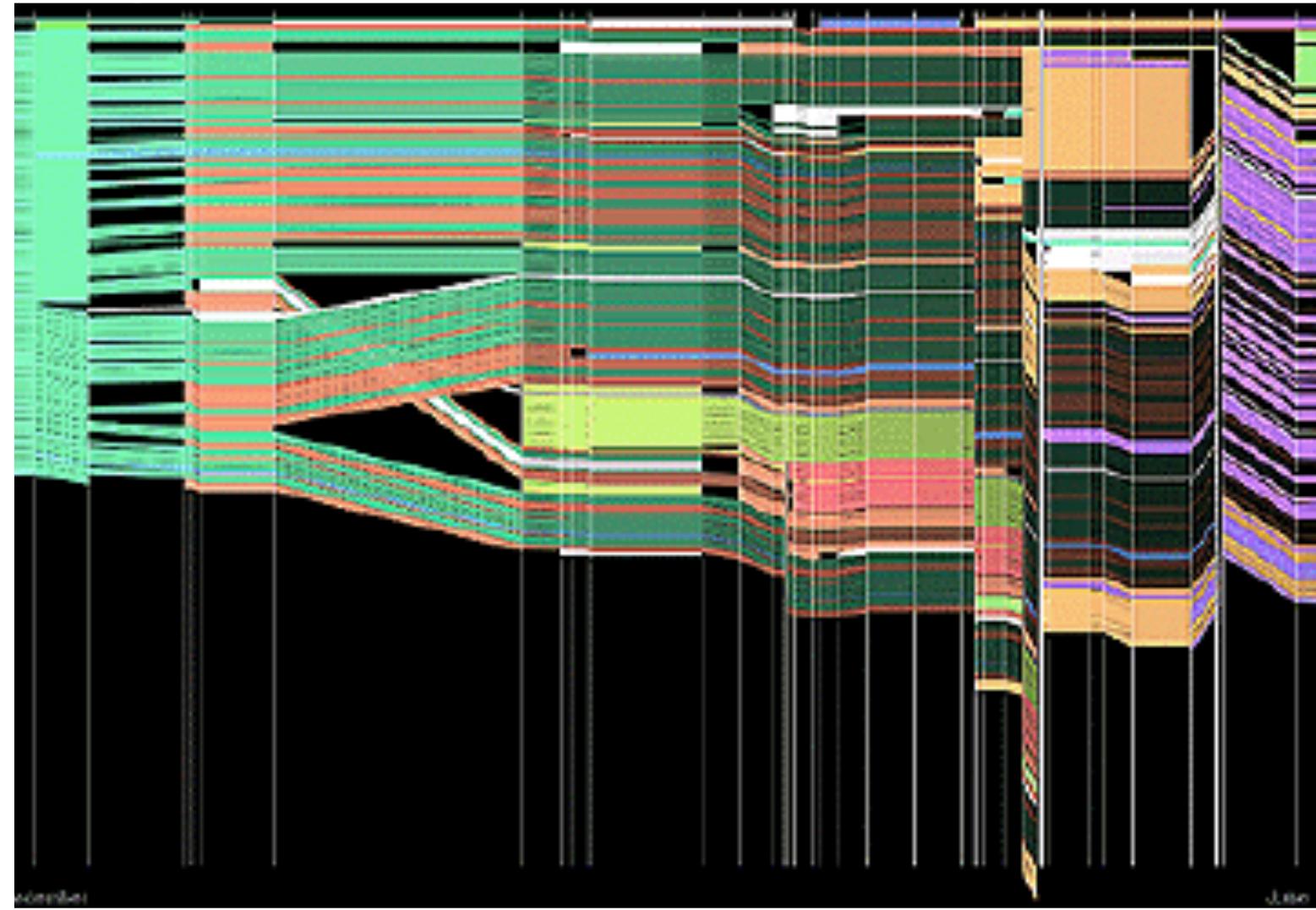
and plotted the results on a **year clock**



# STUDYING COOPERATION AND CONFLICT BETWEEN AUTHORS WITH HISTORY FLOW VISUALIZATIONS

---

*F.B. Viégas, M. Wattenberg e K. Dave  
Conference on Human Factors in Computer Systems  
2004*



# WIKIPEDIA

## English

*The Free Encyclopedia*

3 600 000+ articles

## 日本語

フリー百科事典

741 000+ 記事

## Deutsch

*Die freie Enzyklopädie*

1 209 000+ Artikel

## Español

*La enciclopedia libre*

745 000+ artículos

## Français

*L'encyclopédie libre*

1 085 000+ articles

## Русский

*Свободная энциклопедия*

695 000+ статей

## Italiano

*L'enciclopedia libera*

787 000+ voci

## Português

*A enciclopédia livre*

679 000+ artigos

## Polski

*Wolna encyklopedia*

791 000+ haset

## Nederlands

*De vrije encyclopedie*

679 000+ artikelen



As pessoas estão acostumadas a participar de comunidades *online* nas quais pessoas com perspectivas conflitantes se encontram e discutem

Na Wikipedia, ao contrário de enciclopédias tradicionais, todo o conteúdo pode ser editado por qualquer pessoa a qualquer momento

Alguns acreditam que esta estratégia é vulnerável a erros, desconhecimento e malícia não sendo uma ferramenta de referência confiável

## **Como este sistema aberto e vulnerável pode funcionar?**

**Table 1: 2005 Statistics (SAMPLE05)**

<b>Revision Type</b>	<b>Number</b>	<b>Mean time</b>	<b>Median time</b>
<b>All content</b>	901,242	19.1 days	1113 minutes
<b>Mass delete (MD)</b>	4,848	7.0 days	2.9 minutes
<b>MD obscene</b>	105	0.13 days	2 minutes

A Wikipedia fornece os dados de todas as edições que foram realizadas nos artigos

O *History Flow* é uma visualização que mostra os relacionamentos entre as diversas versões de um mesmo artigo

O algoritmo de casamento de seções do texto é

O. Heckel. *A technique for isolating differences between files.*  
Communications of the ACM, 1978.

A análise exploratória desta visualização revela padrões complexos de cooperação e conflito

# REPRESENTAÇÃO VISUAL

---

- Cada versão do documento é representada como uma linha de revisão com comprimento proporcional ao tamanho do texto



# REPRESENTAÇÃO VISUAL

---

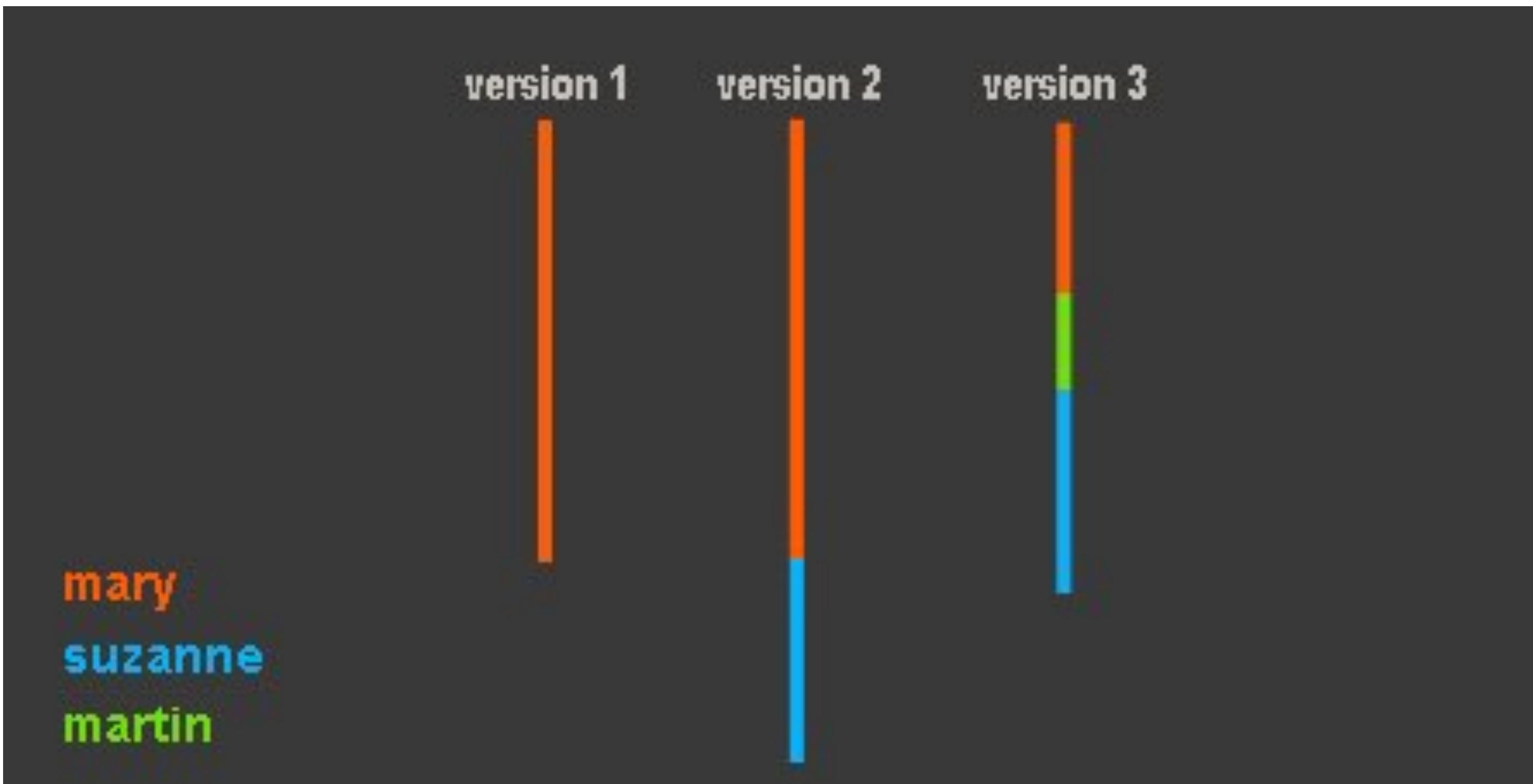
- Os autores são representados por **cores diferentes**
- Seções das linhas de revisão são coloridas de acordo com a cor representativa do autor



# REPRESENTAÇÃO VISUAL

---

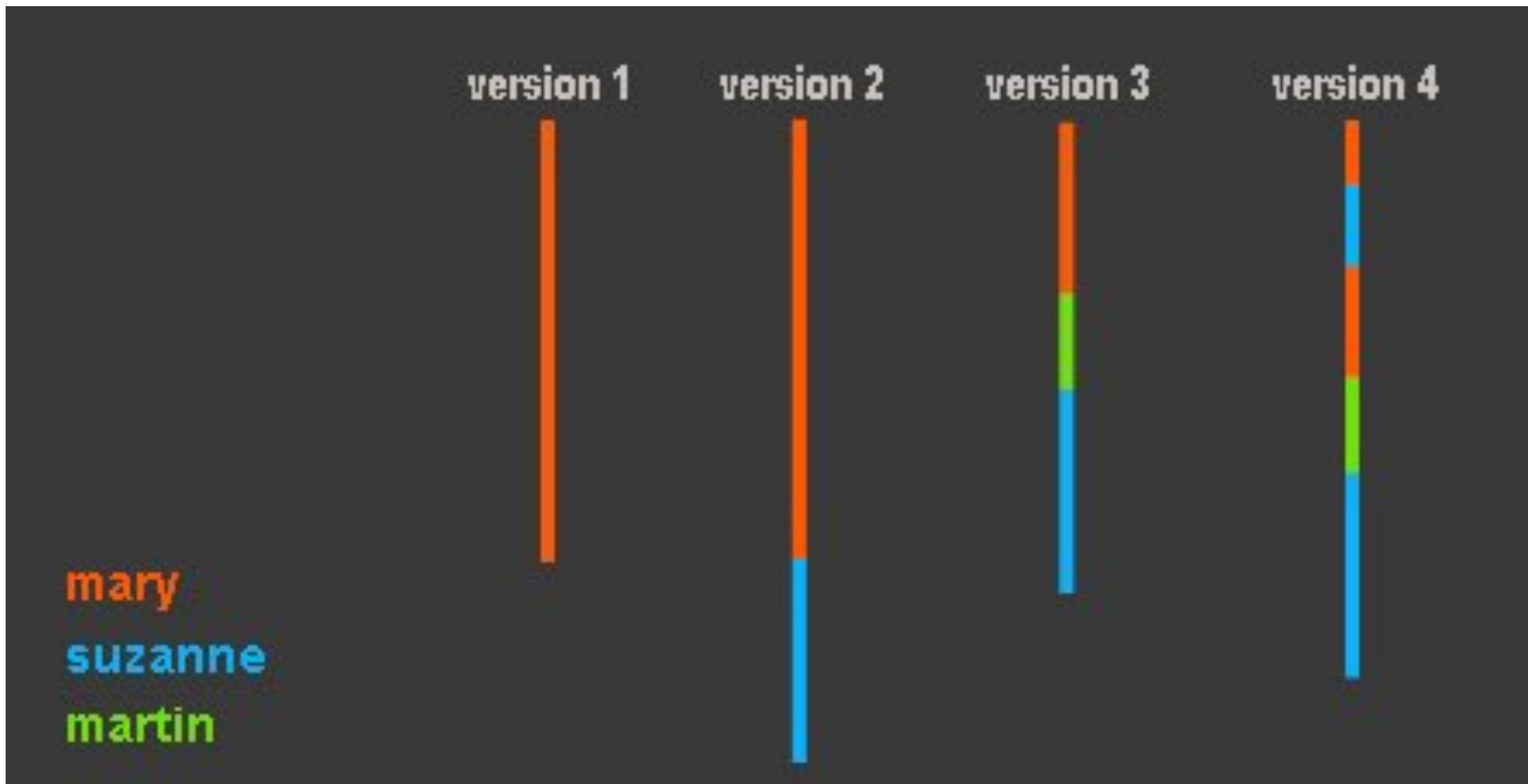
- Os autores são representados por **cores diferentes**
- Seções das linhas de revisão são coloridas de acordo com a cor representativa do autor

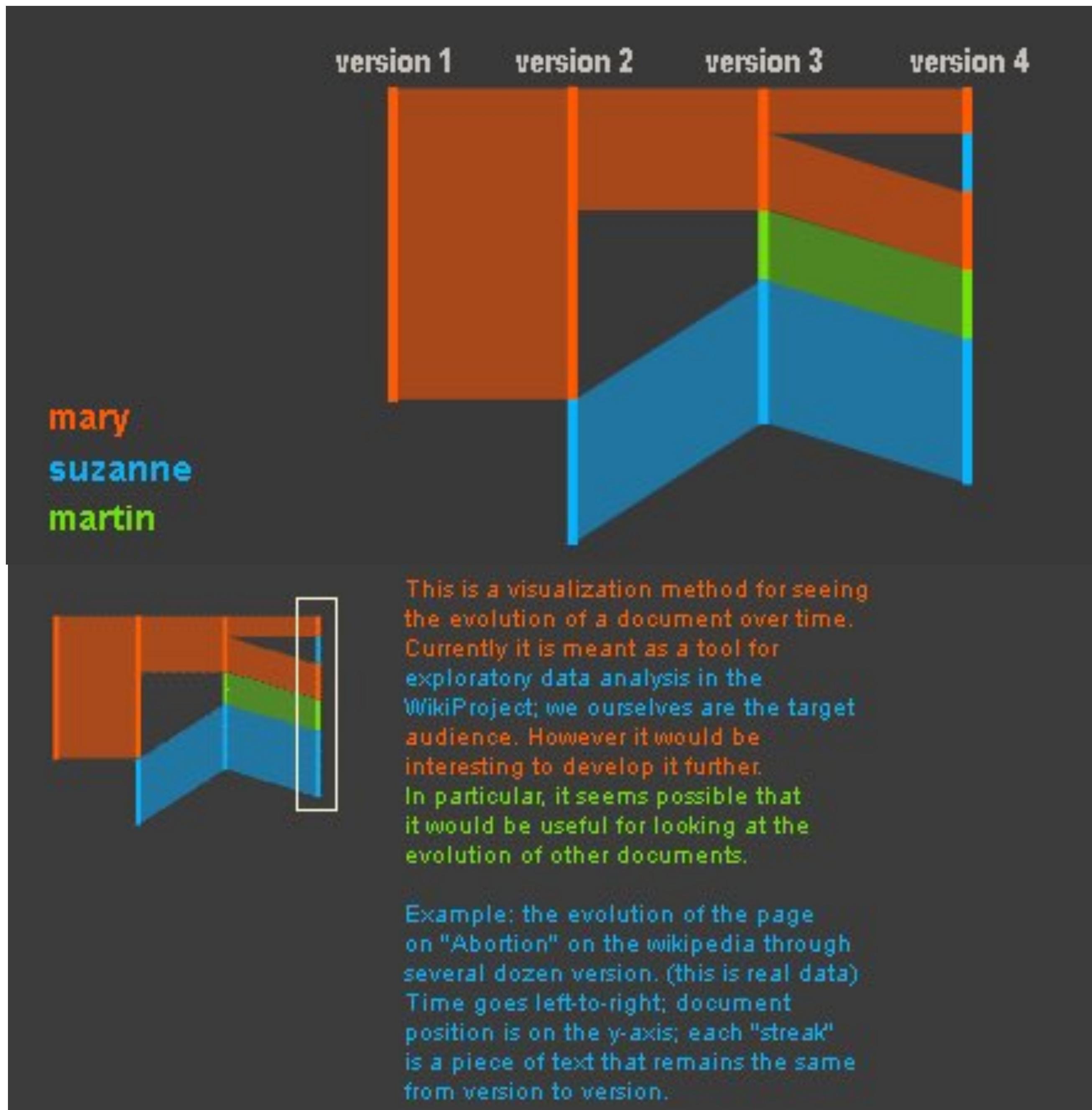


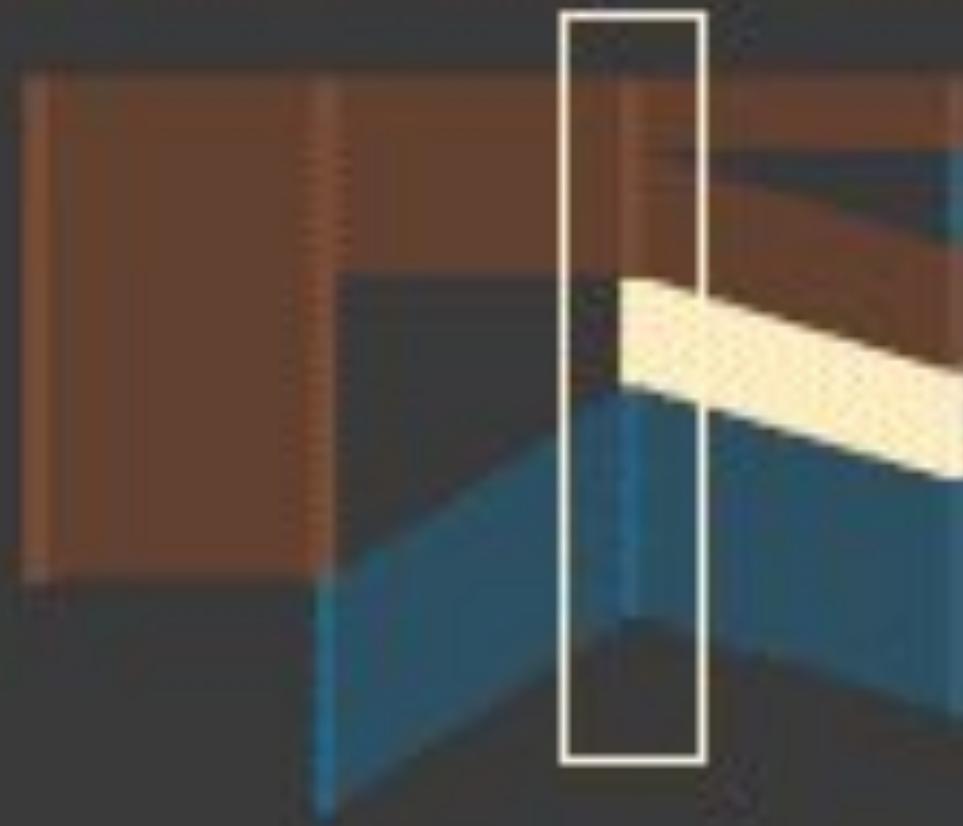
# REPRESENTAÇÃO VISUAL

---

- Os autores são representados por cores diferentes
- Seções das linhas de revisão são coloridas de acordo com a cor representativa do autor





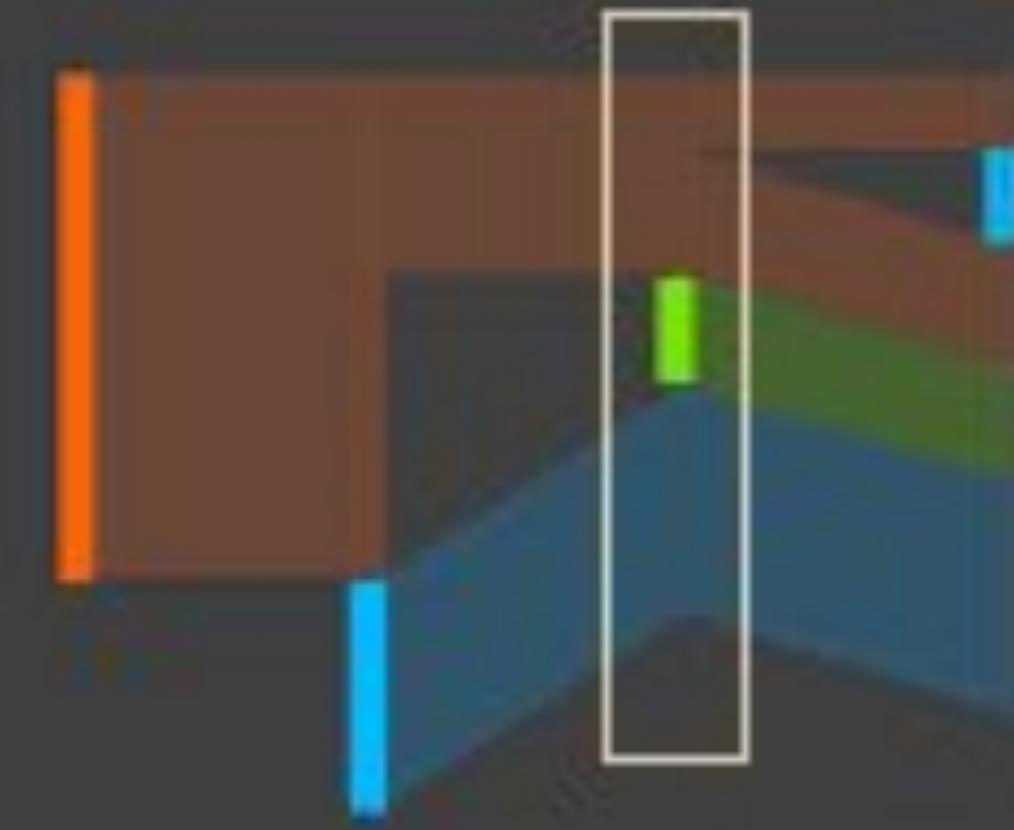


This is a visualization method for seeing the evolution of a document over time. Currently it is meant as a tool for exploratory data analysis in the WikiProject; we ourselves are the target. In particular, it seems possible that it would be useful for looking at the evolution of other documents.

Example: the evolution of the page on "Abortion" on the wikipedia through several dozen version. (this is real data!) Time goes left-to-right, document position is on the y-axis; each "strand" is a piece of text that remains the same from version to version

## Visão por autor

destaca a contribuição por autor ressaltando a persistência de suas contribuições no tempo

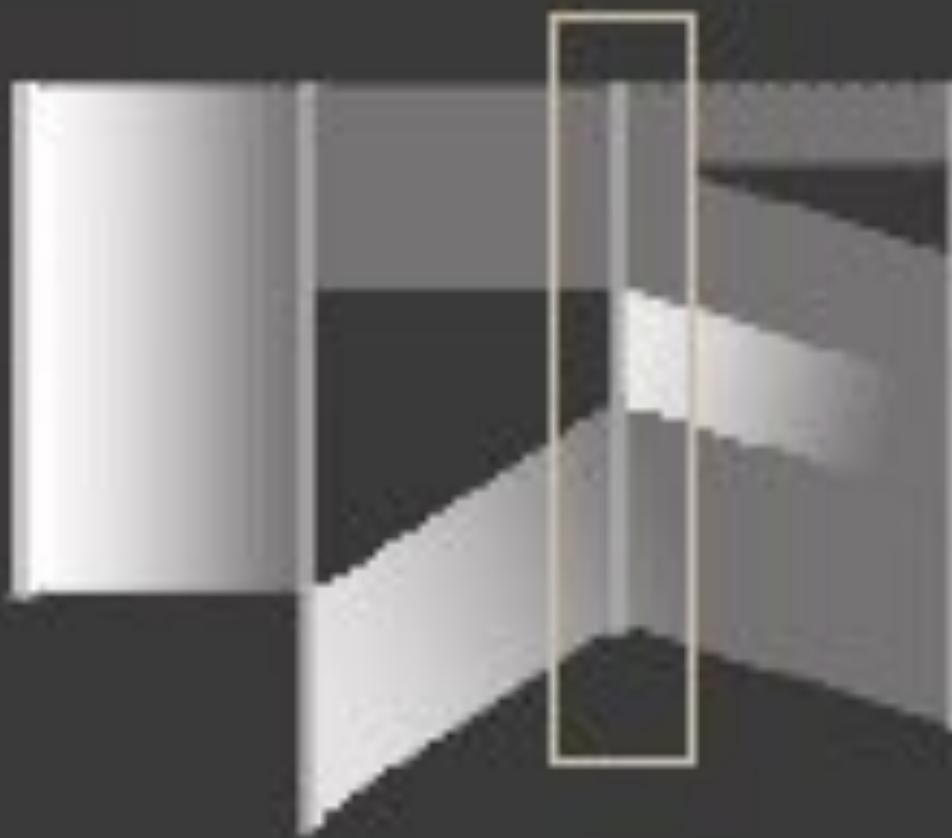


This is a visualization method for seeing the evolution of a document over time. Currently it is meant as a tool for exploratory data analysis in the WikiProject; we ourselves are the target. In particular, it seems possible that it would be useful for looking at the evolution of other documents.

Example: the evolution of the page on "Abortion" on the Wikipedia through several dozen version (this is real data). Time goes left-to-right, document position is on the y-axis; each "stream" is a piece of text that remains the same from version to version.

## Visão por contribuições recentes

ressalta o novo conteúdo em cada nova versão dos artigos.  
Permite identificar os trechos do código mais ativamente editados



This is a visualization method for seeing the evolution of a document over time. Currently it is meant as a tool for exploratory data analysis in the WikiProject; we ourselves are the target. In particular, it seems possible that it would be useful for looking at the evolution of other documents.

Example: the evolution of the page on "Abortion" on the wikipedia through several dozen version. (this is real data)

Time goes left-to-right; document position is on the y-axis; each "streak" is a piece of text that remains the same from version to version.

## Visão por idade

Destaca a persistência das contribuições.

Branco = nova contribuição

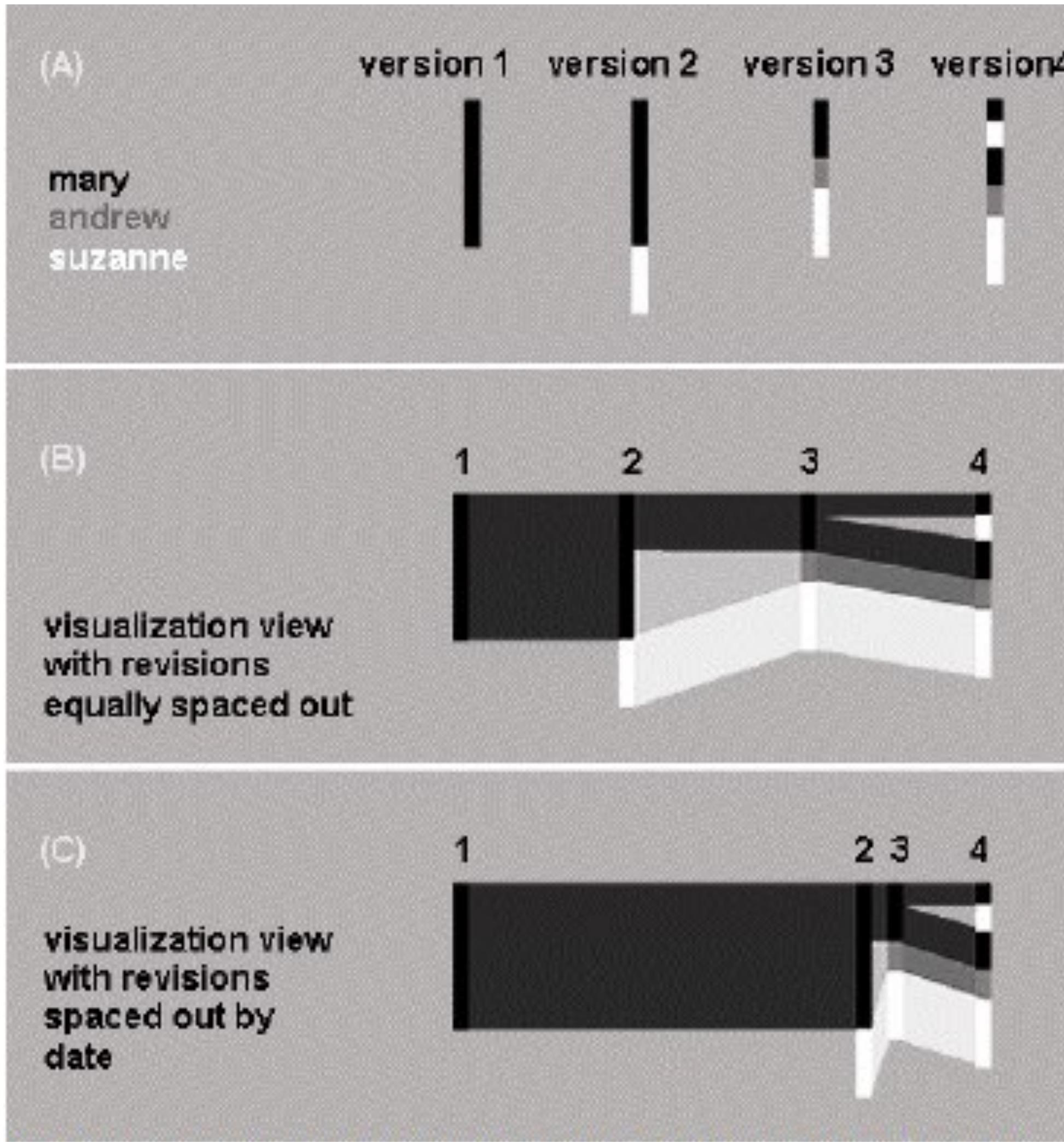
Cinza = contribuição antiga

Há a versão igualmente  
espaçada e outra espaçada  
de acordo com a data de  
edição

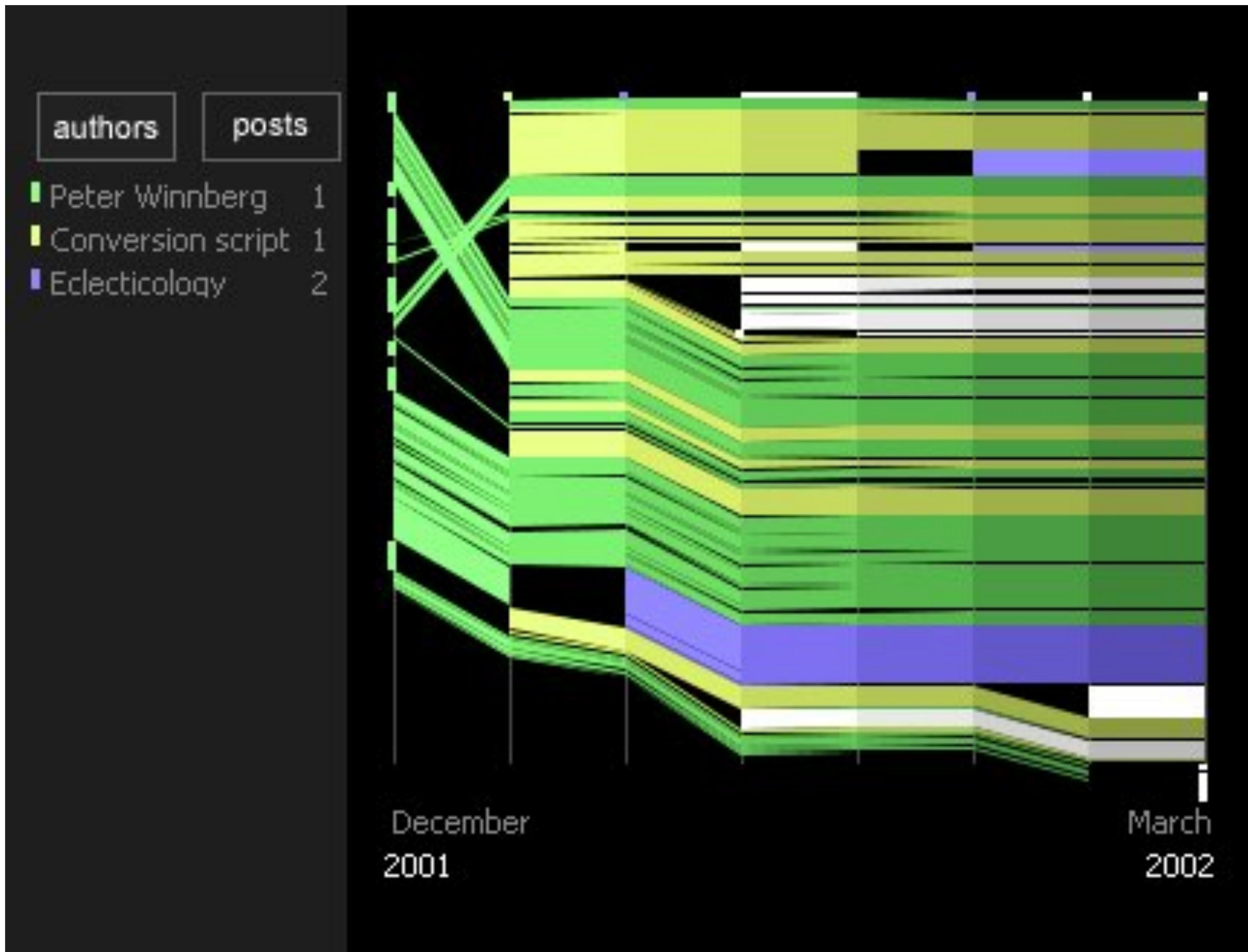
Permite identificar  
eventos de conflito e  
deleção que  
desaparecem rápido

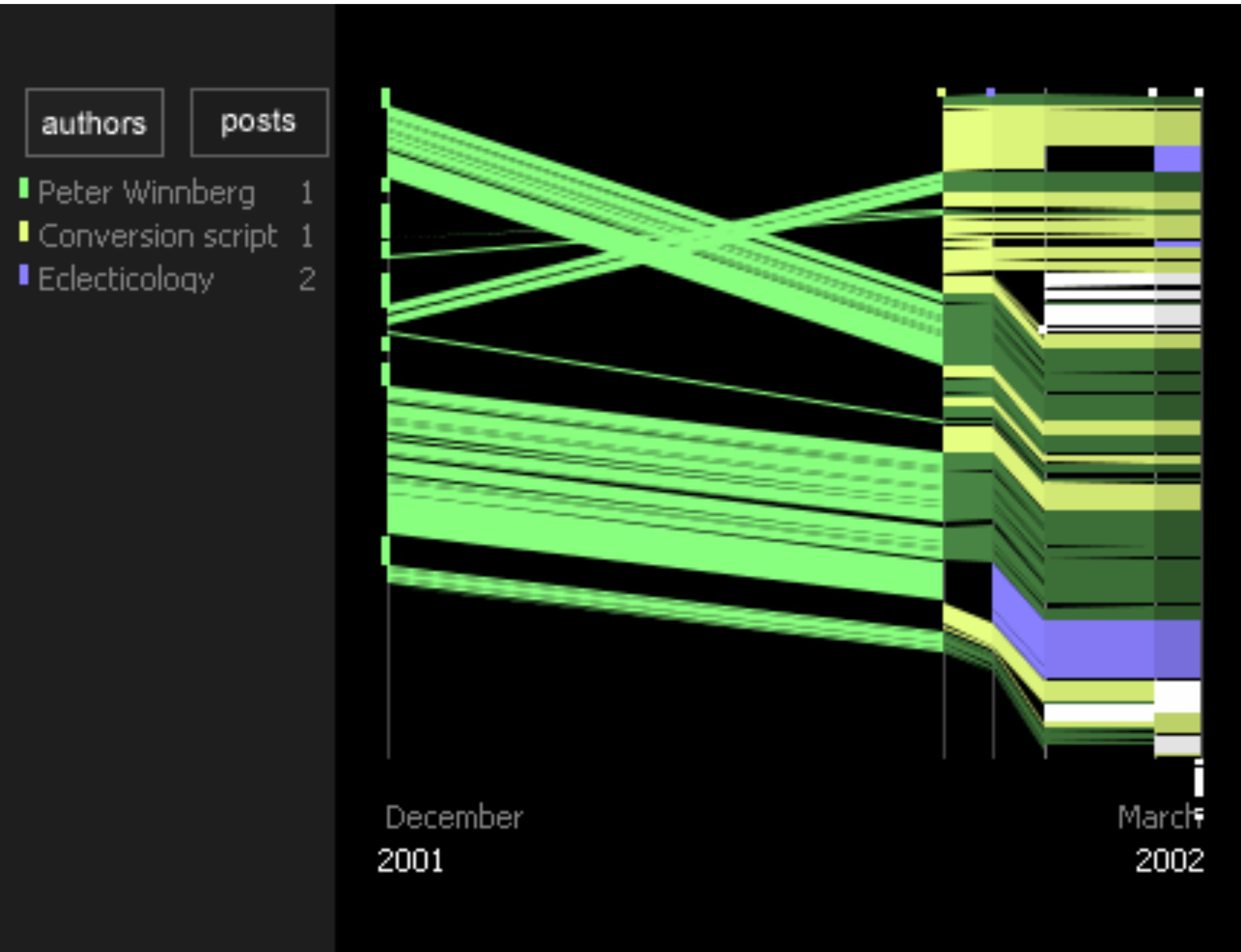
Enfatiza as versões mais  
estáveis

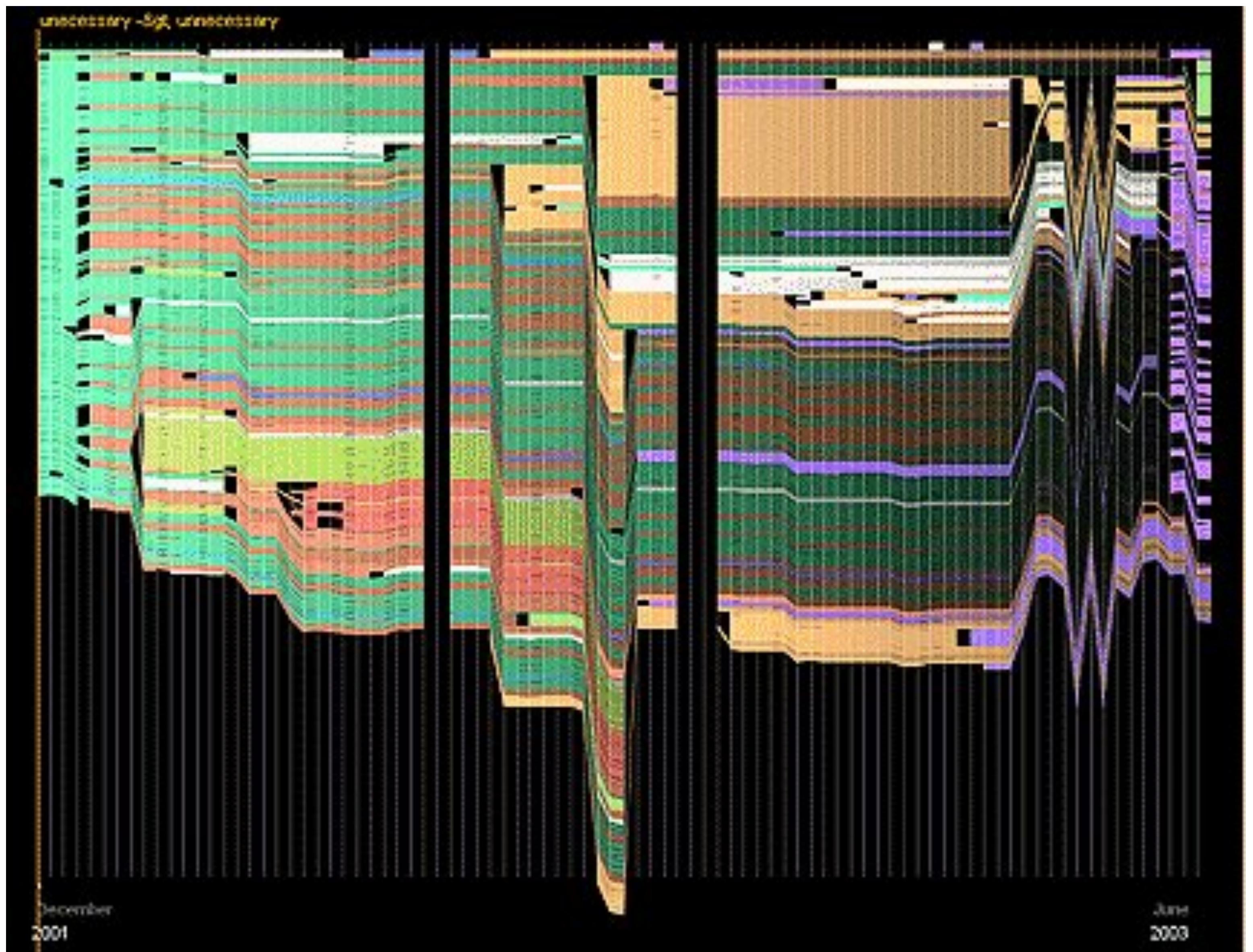
Revela o ritmo das  
colaborações



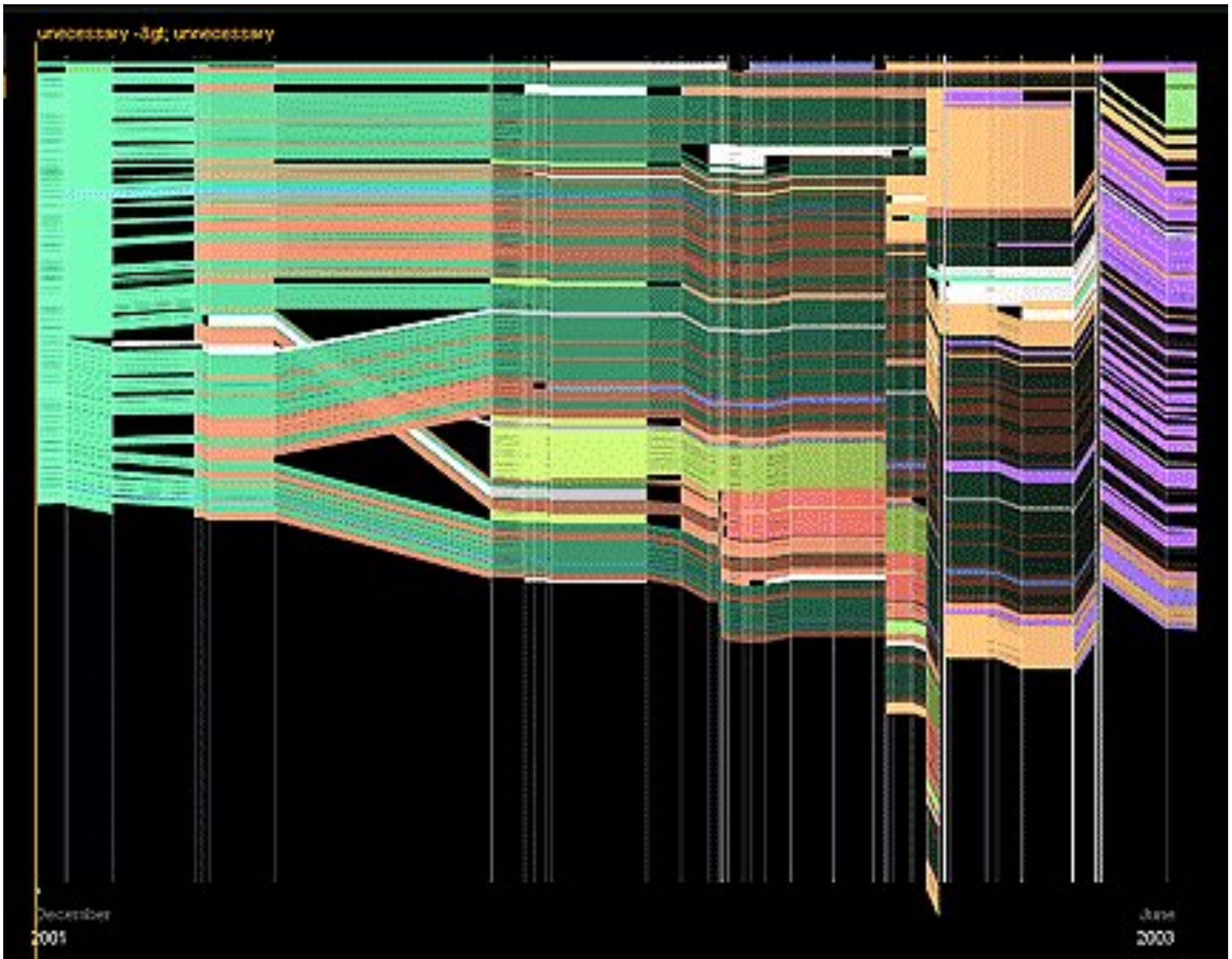
*Fig 2: explanation of history flow's visualization mechanism*







Páginas que tratam de temas **controversos** muitas vezes sofrem **vandalismo**, o que fica evidente com as faixas em preto mostrando deleções



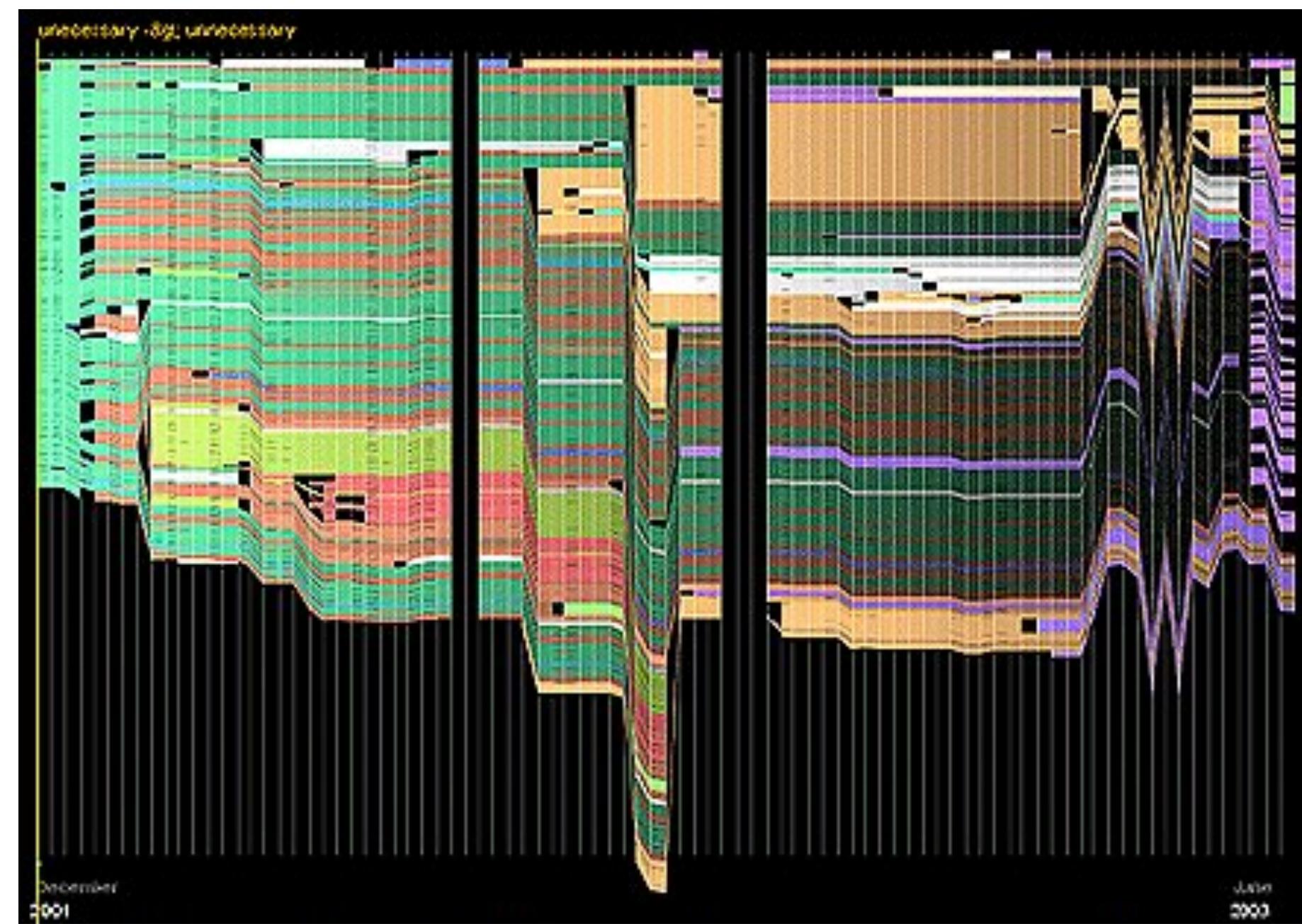


Fig 4: history flow for “Abortion” page, versions equally spaced.

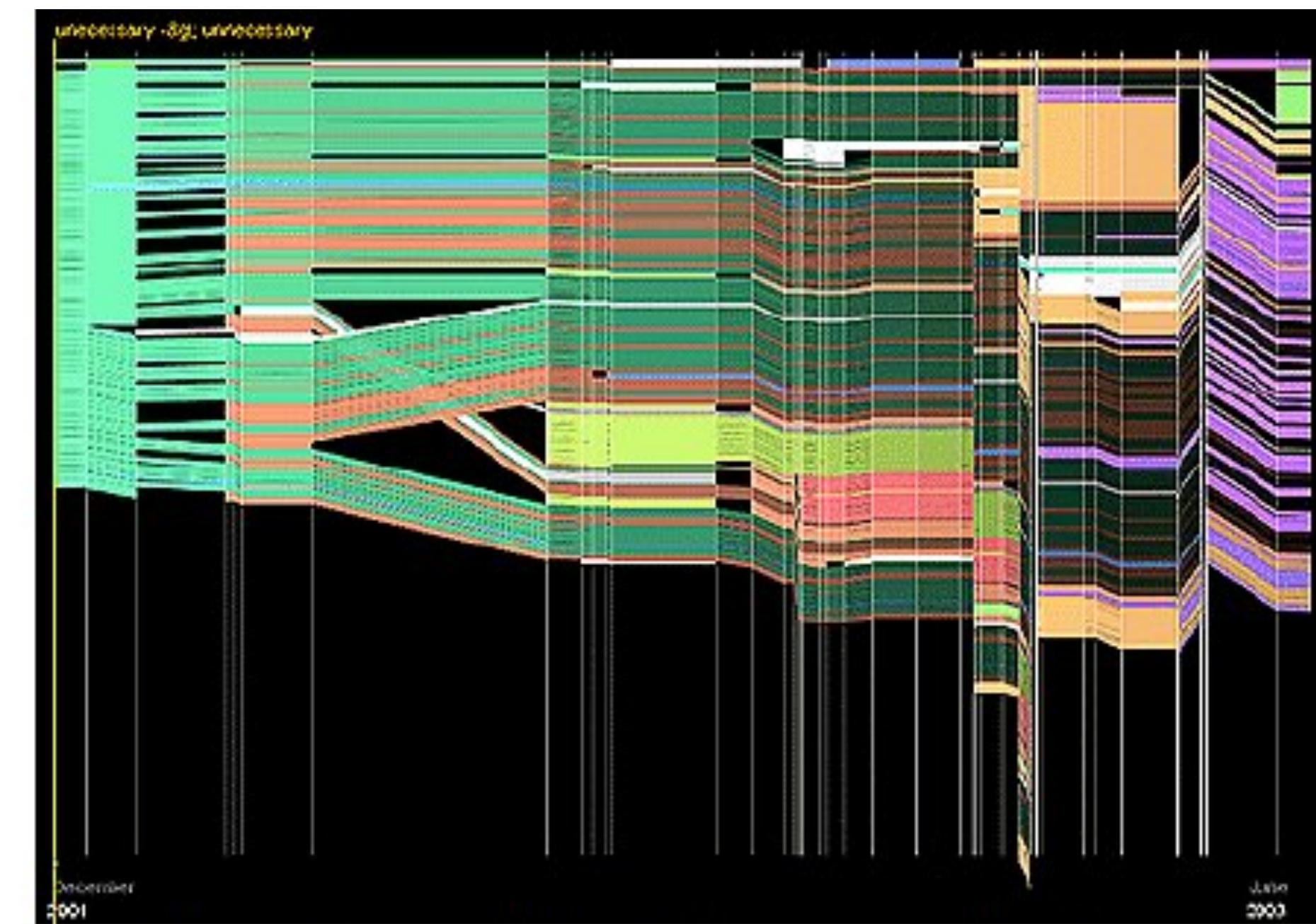
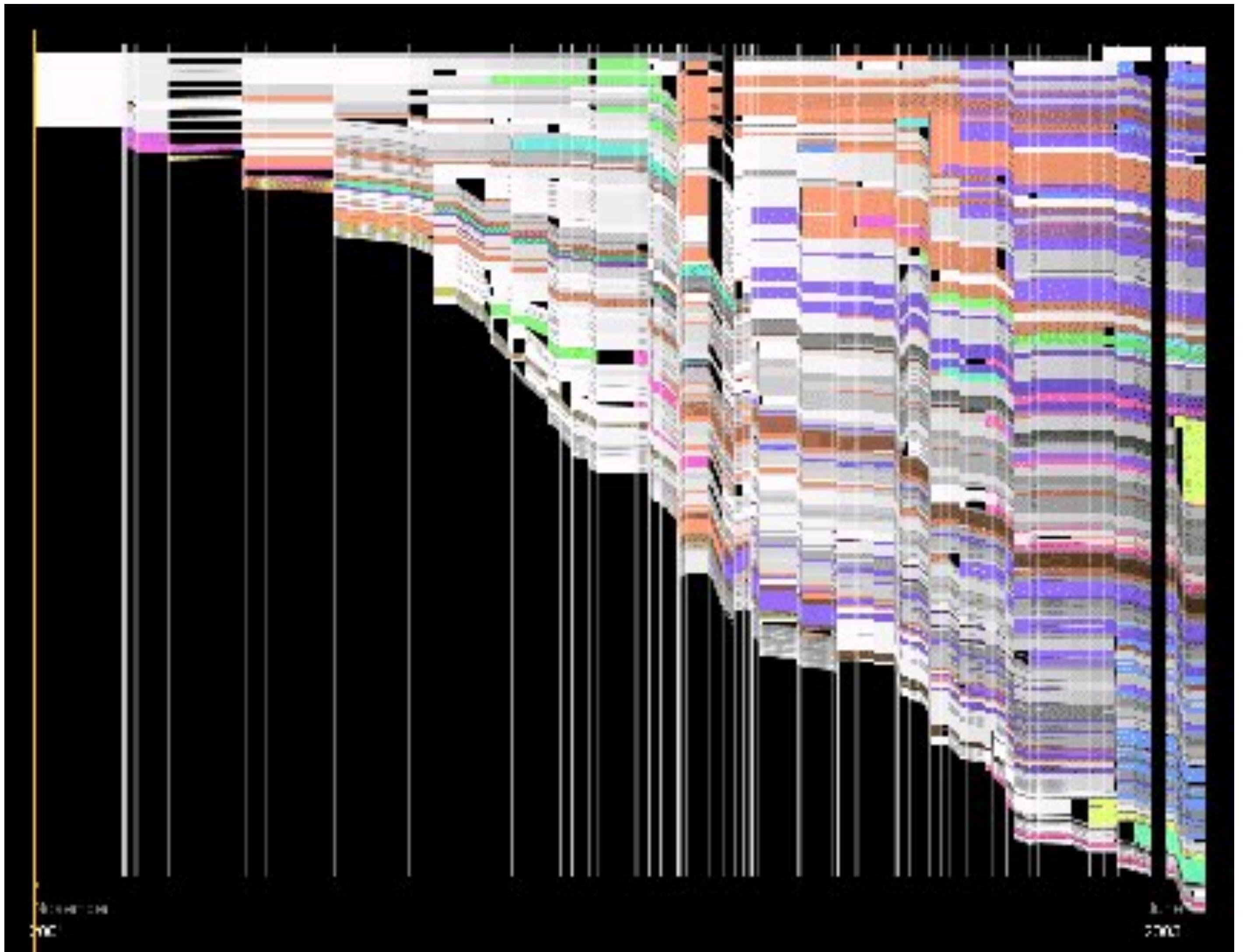


Fig 5: history flow for “Abortion” page, spaced by date

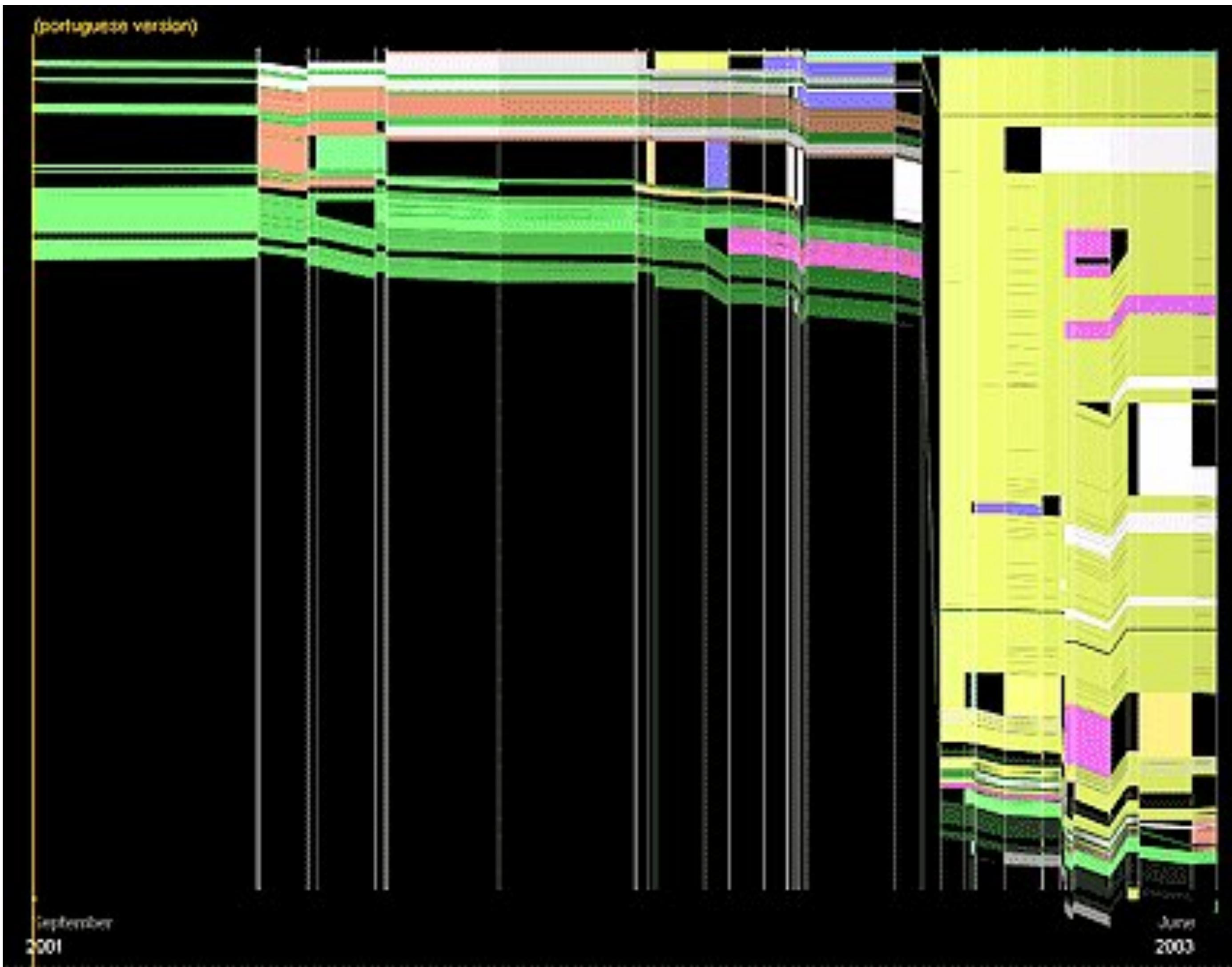


No history flow, pode-se ver o histórico de contribuições por autor e juntamente pode-se ver os autores anônimos branco



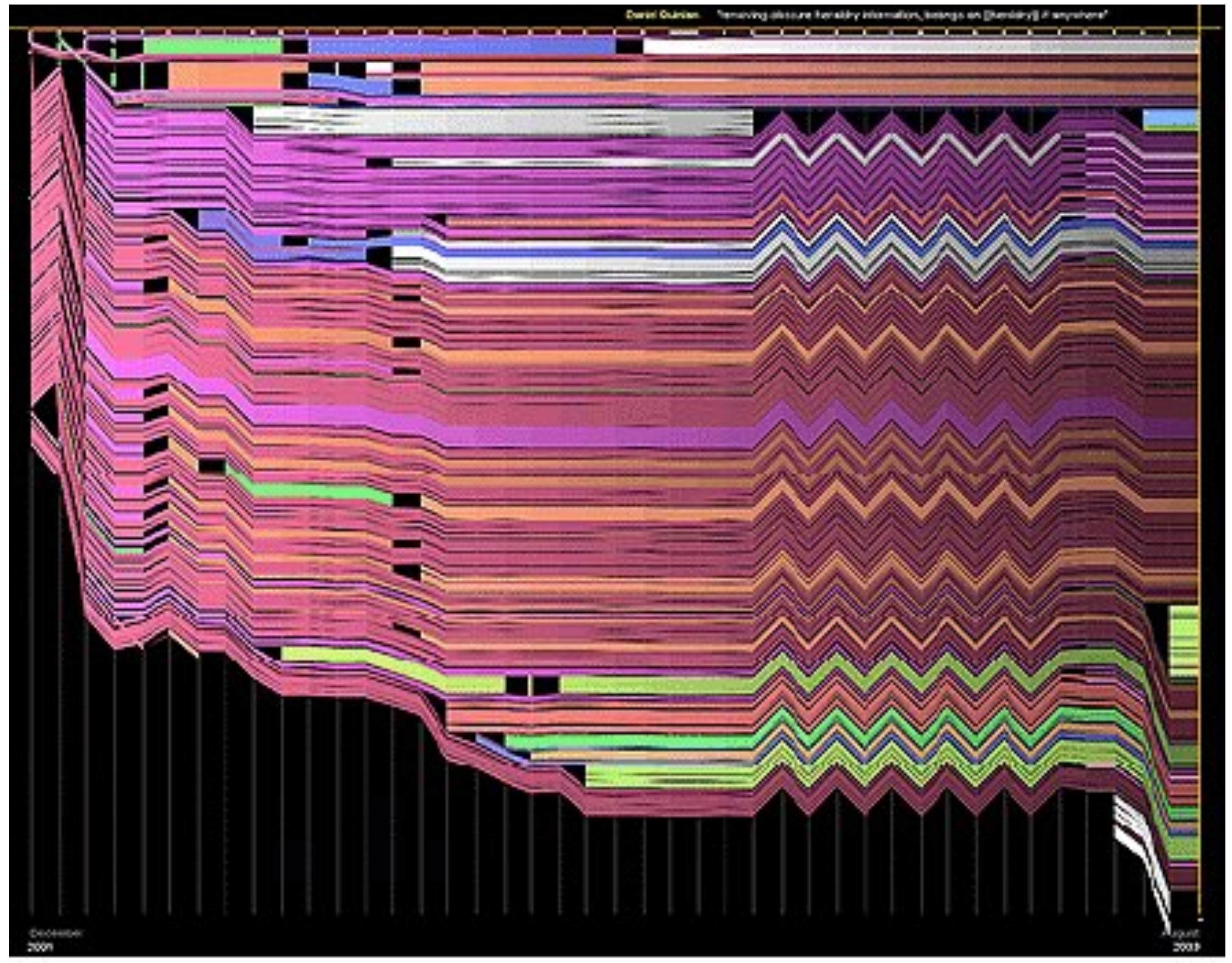
No history flow, pode-se ver o histórico de contribuições por autor e juntamente pode-se ver os autores anônimos branco

Artigo sobre a Microsoft

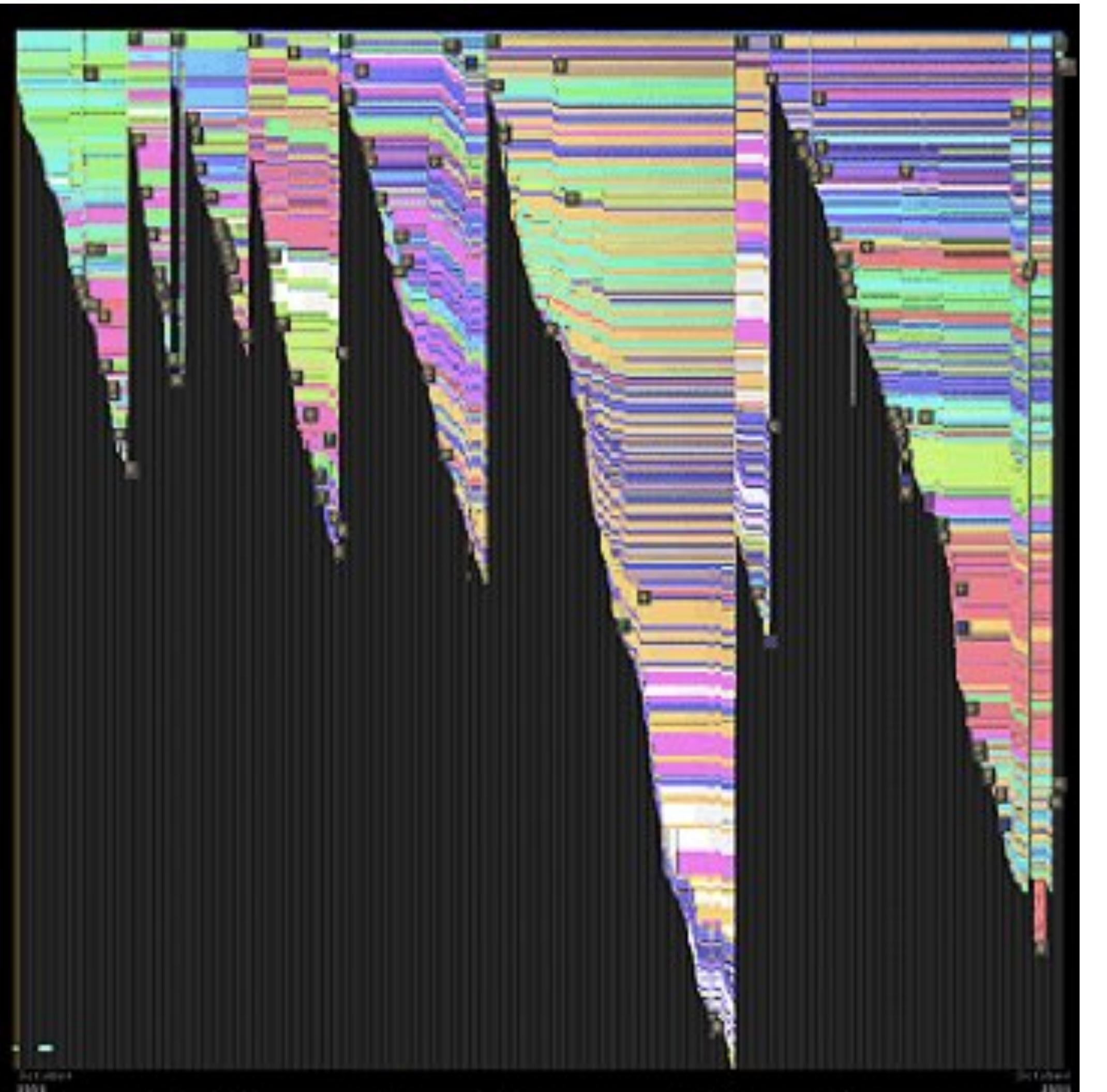


No history flow, pode-se ver o histórico de contribuições por autor e juntamente pode-se ver os autores anônimos branco

Artigo sobre o Brasil



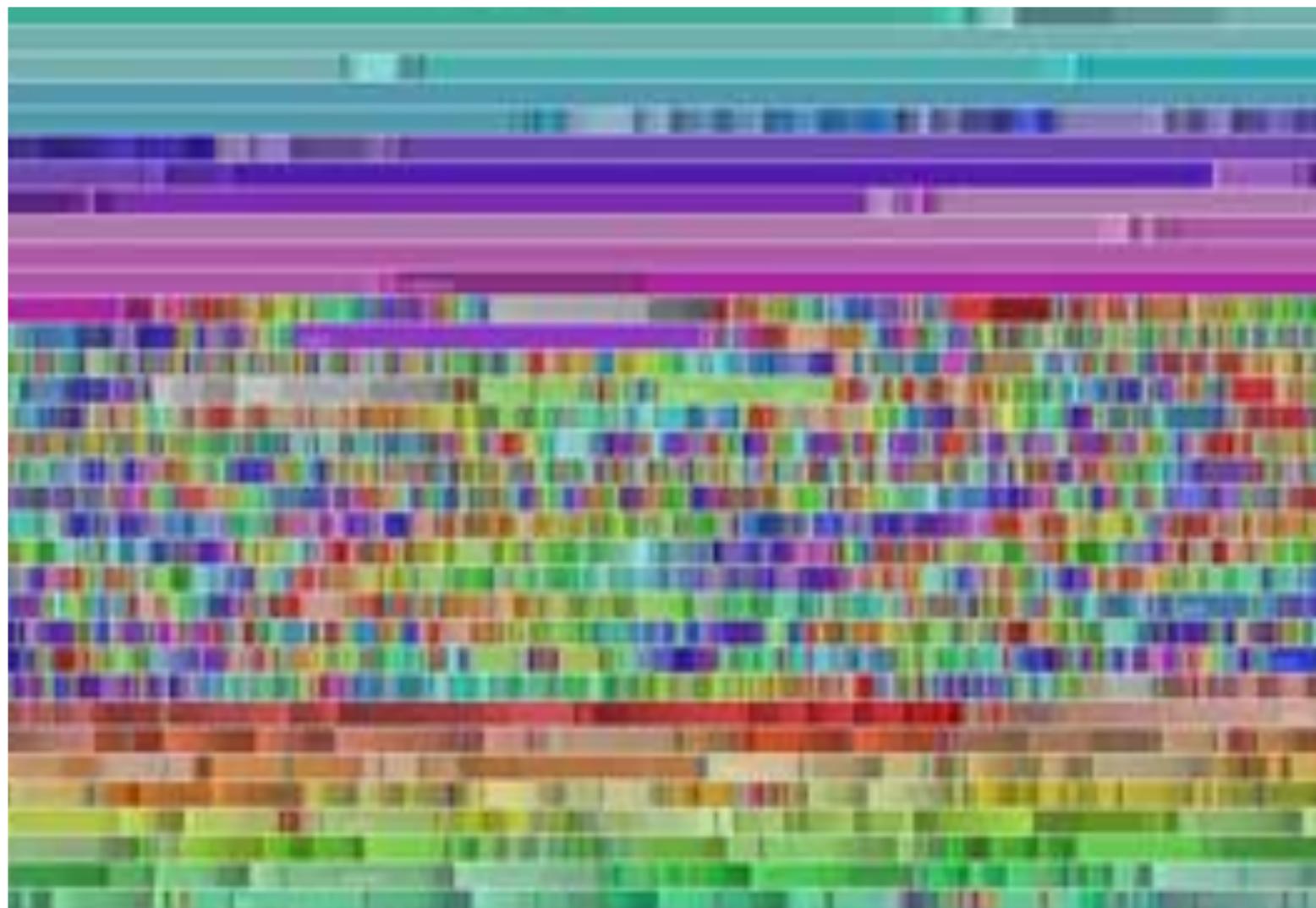
*Fig 6: “Chocolate” page spaced out by number of versions; we can see the zigzag pattern of an edit war.*



**Figure 7:** History flow diagram of the talk page on “God,” showing the cyclical nature of archiving operations. As the increasing length of the page reaches unwieldy proportions, the discussion is archived and the live talk page returns to a more manageable size.

# VISUALIZING ACTIVITY ON WIKIPEDIA WITH CHROMOGRAMS

---



*M. Wattenberg, F. Viégas e K. Hollenbach*  
*Interact*  
*2007*

O objetivo da visualização é analisar os hábitos de edição de  
**usuários administradores da Wikipedia**

Administradores são usuários com permissões especiais para  
deleção e proteção de páginas

Normalmente são usuários ativos, confiáveis e influentes

# DADOS

---

- Históricos das edições de 509 administradores
- Os históricos consistem em uma sequência de eventos com horário de ocorrência
- Os eventos são os títulos dos artigos e comentários opcionais
- Comentários curtos e seguindo alguma padronização

# DESAFIOS

---

- Volume: 100.000 eventos
- Diversos: 80.000 tokens distintos
- Estrutura irregular

# REPRESENTAÇÃO VISUAL

---

- A primeira letra indica a cor
- A segunda letra indica a saturação
- A terceira letra indica o brilho



**a**

Time	Page	Comment
May 21, 9:32 am	Sphere	Add cite
May 22, 10:56 am	Sphere	New Intro
May 22, 1:23 pm	sphere.png	Copyright
May 22, 2:54 pm	Helix	Spell check
May 22, 3:00 pm	Mathematics	Revert
May 24, 11:21 am	Fields Medal	2006 data
May 24, 11:25 am	Talk:Fractal	List of proposed changes
May 24, 11:27 am	Talk:Fractal	List of proposed changes
May 25, 10:13 am	Sphere	Fix tpyo
May 25, 10:23 am	mercator.png	Copyright

**b** Chromogram of comments:  
Timeline View



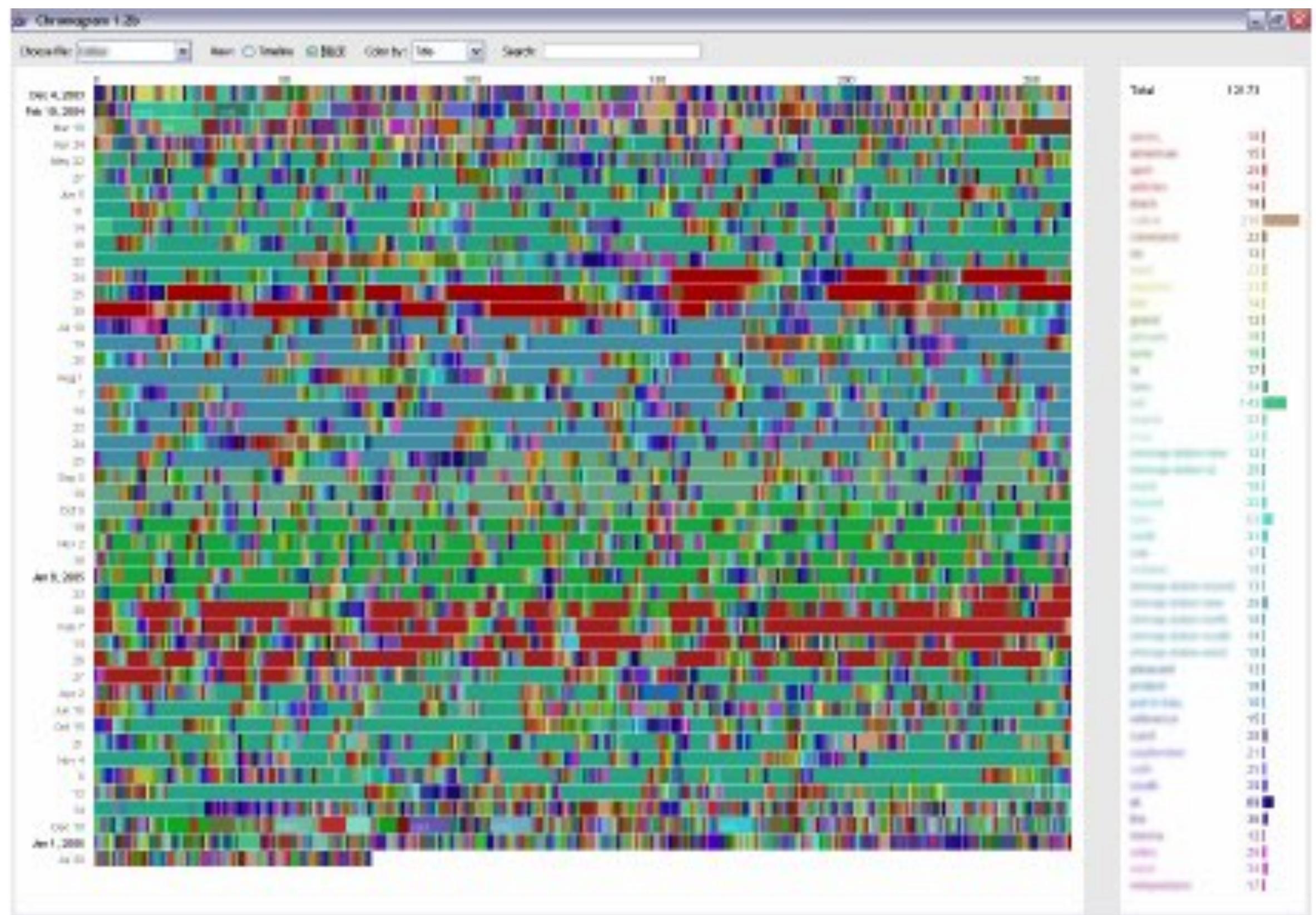
**c** Block View



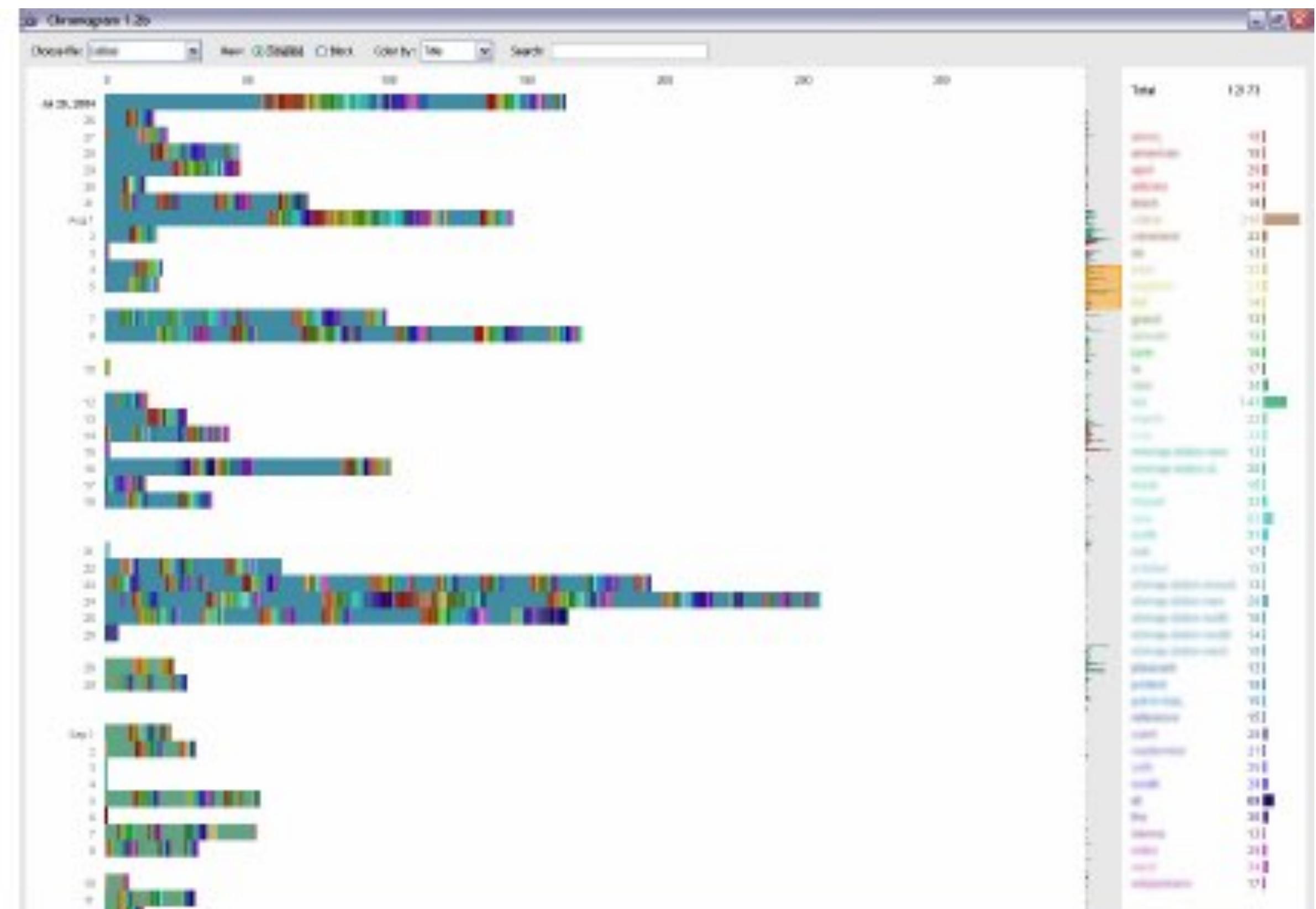
**d** Block View (compressed, no labels)



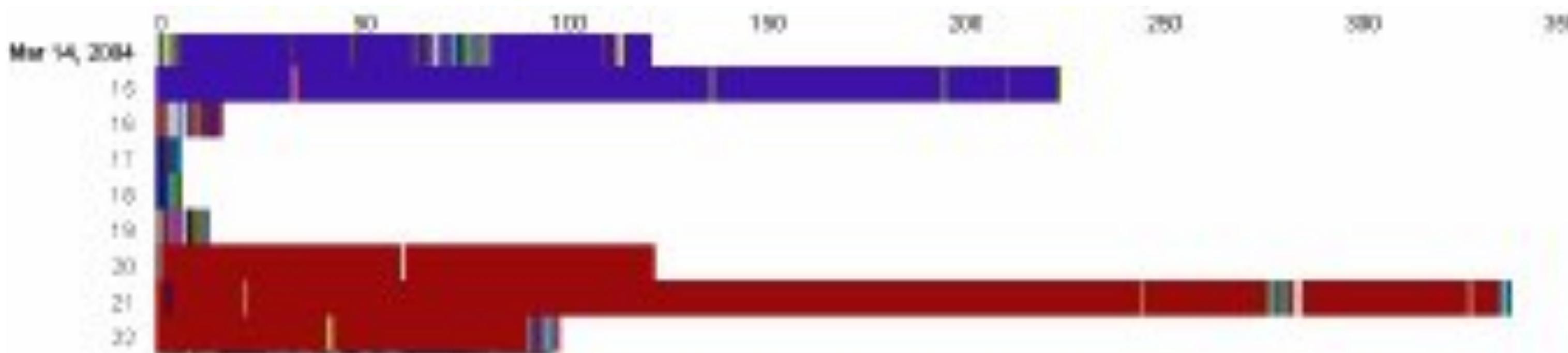
**Fig. 2.** Creating a Chromogram



**Fig. 3.** The Chromogram Application: Block View

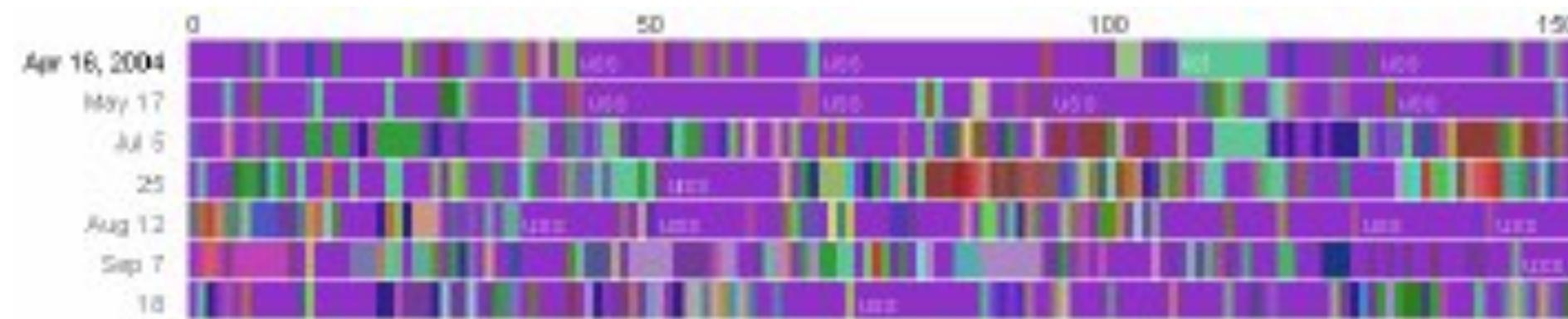


**Fig. 4.** The Chromogram Application: Timeline View, same data as Fig. 3.

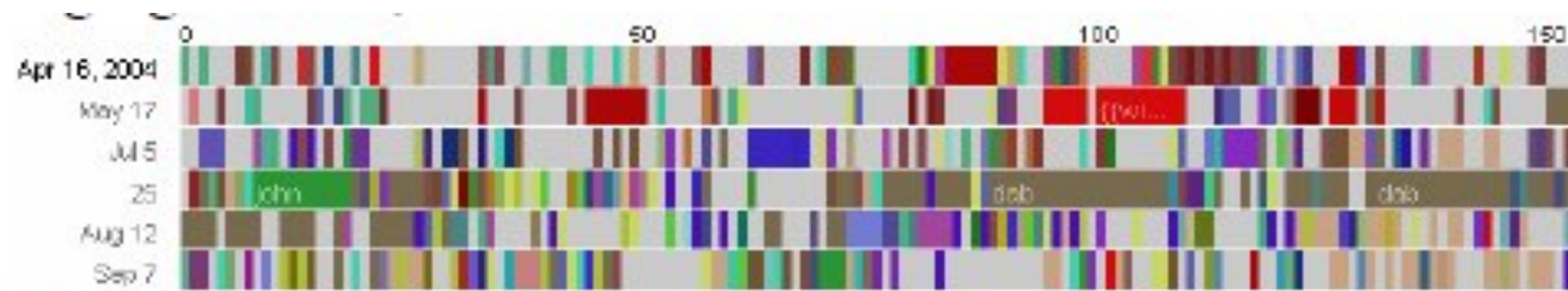


**Fig. 5.** Timeline comment chromogram with activity bursts. Blue: typo removal. Red: adding disambiguation messages.

As atividades acontecem quase sempre em **blocos**

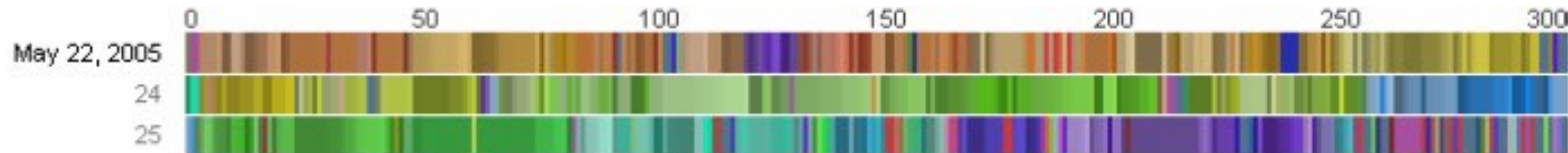


**Fig. 6.** Title chromogram: purple edits relate to U.S. ships, green to lists of ships.



**Fig. 7.** Comment chromogram for same edits as Fig. 6: Highly irregular.

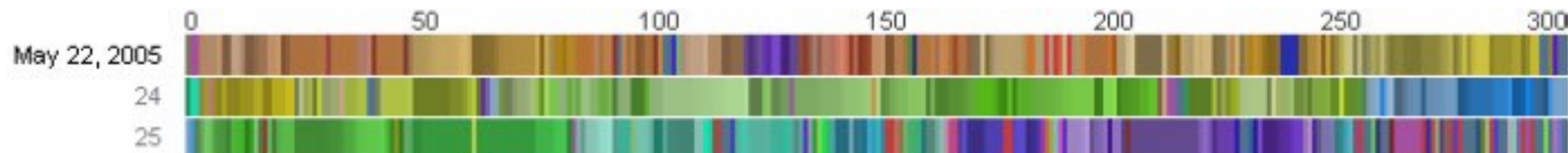
Em vários casos, os administradores editam diferentes artigos alternadamente (roxo e verde). Comumente os artigos relacionam-se ao mesmo tema  
O tipo de atividade é altamente variável



**Fig. 8.** Rainbows: Alphabetical order effects in a title chromogram for 900 edits

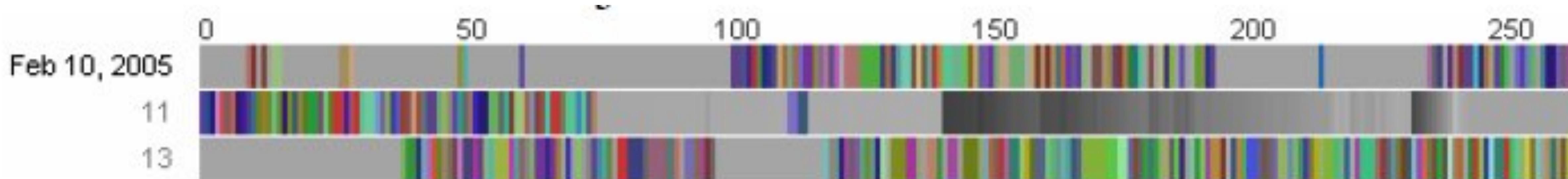
900 edições ao longo de 4 dias?

Alguma ideia do tipo de tarefa realizada?



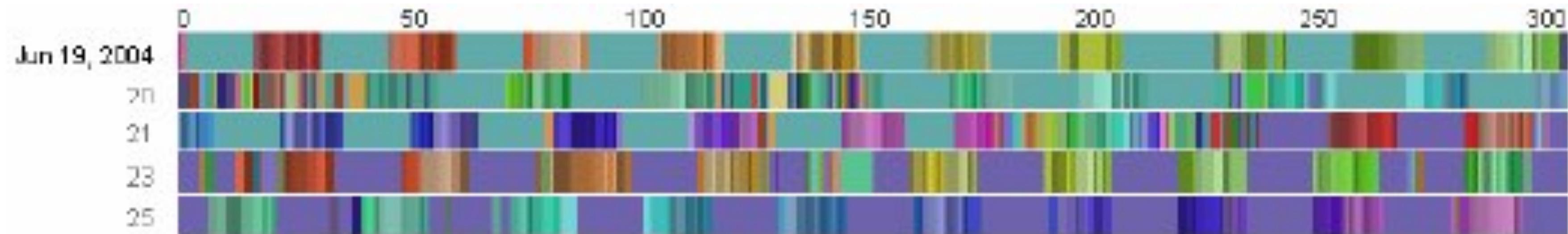
**Fig. 8.** Rainbows: Alphabetical order effects in a title chromogram for 900 edits

Edição de temas em ordem alfabética



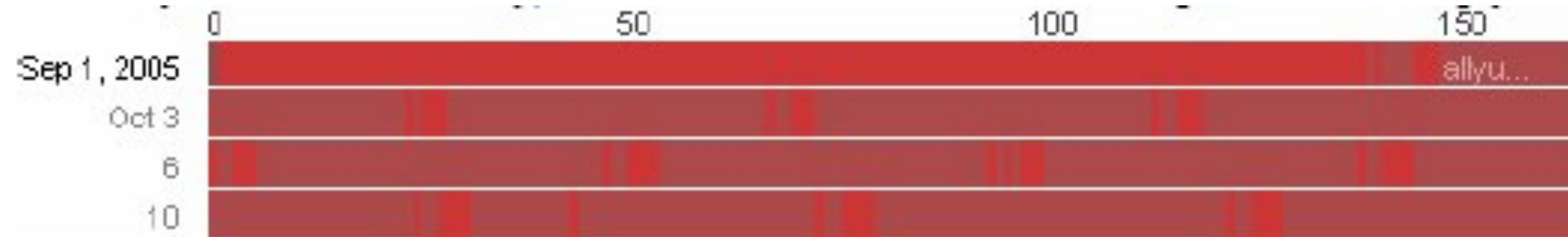
**Fig. 9.** Title chromogram: edits on pages with numerical titles (in gray)

Edição de temas cujos títulos se iniciam com números como  
por exemplo “1922 in Literature”



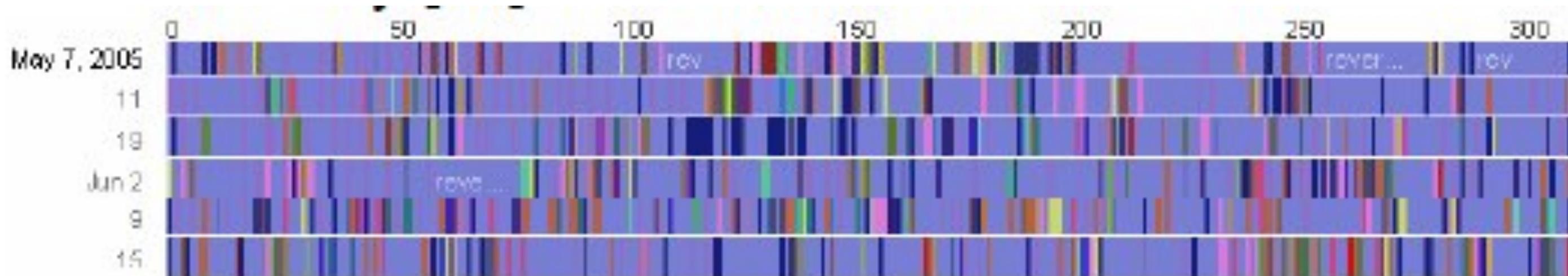
**Fig. 11.** Title chromogram: Switching between images (purple and light blue) and articles (rainbows)

Administradores sistematicamente fazem o upload de várias imagens e depois as inserem no texto

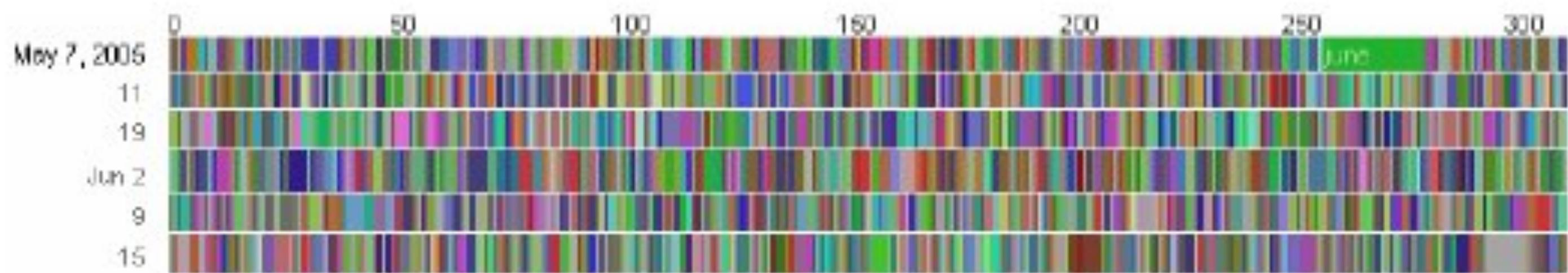


**Fig. 12.** Title chromogram for "AFD bot". Note the small range of colors and regular rhythm of edits.

Administradores comumente escrevem programas para realização de tarefas repetitivas



**Fig. 13.** Comment chromogram: Reverting vandalism and other bad edits



**Fig. 14.** Title chromogram for same edits in Fig. 13.

Reversão de vandalismo em artigos variados

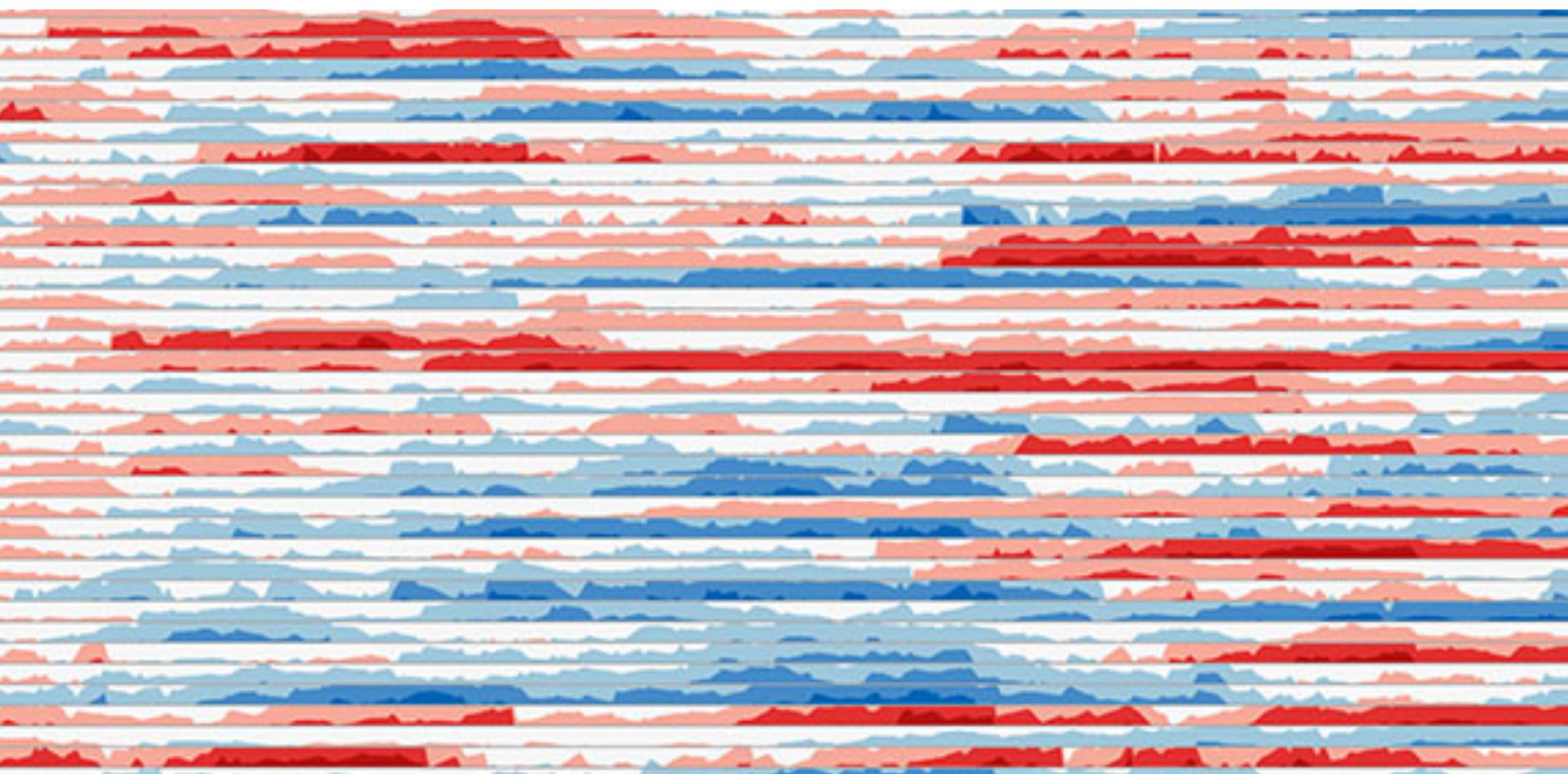
# SIZING THE HORIZON: THE EFFECTS OF CHART SIZE AND LAYERING ON THE GRAPHICAL PERCEPTION OF TIME SERIES VISUALIZATIONS

---

J. Heer and N. Kong e M. Agrawala

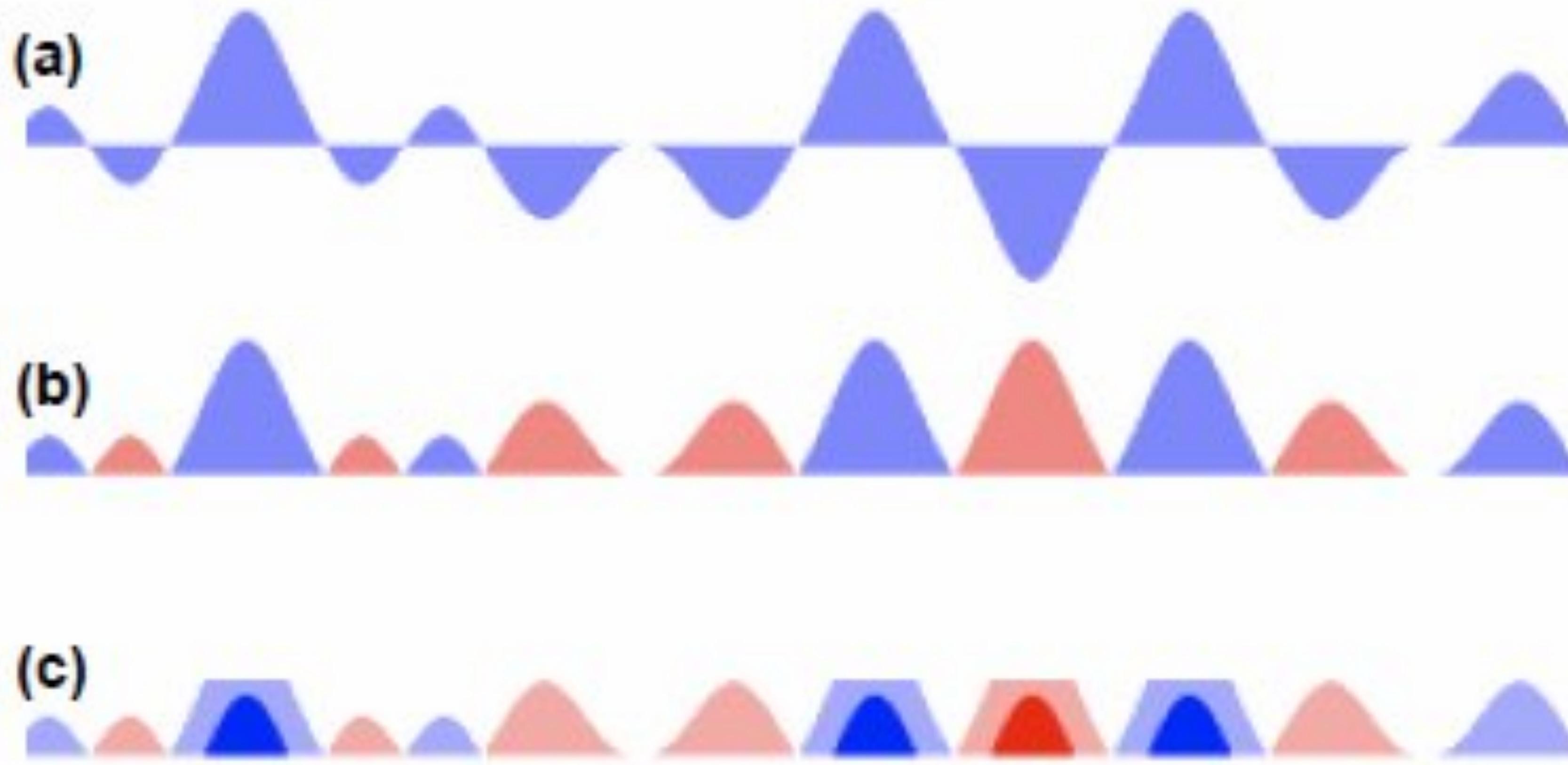
Conference on Human Factors in Computer Systems

2009



O objetivo é a visualização e comparação de múltiplas séries temporais

De acordo com o princípio de Tufte sobre maximizar a densidade de dados

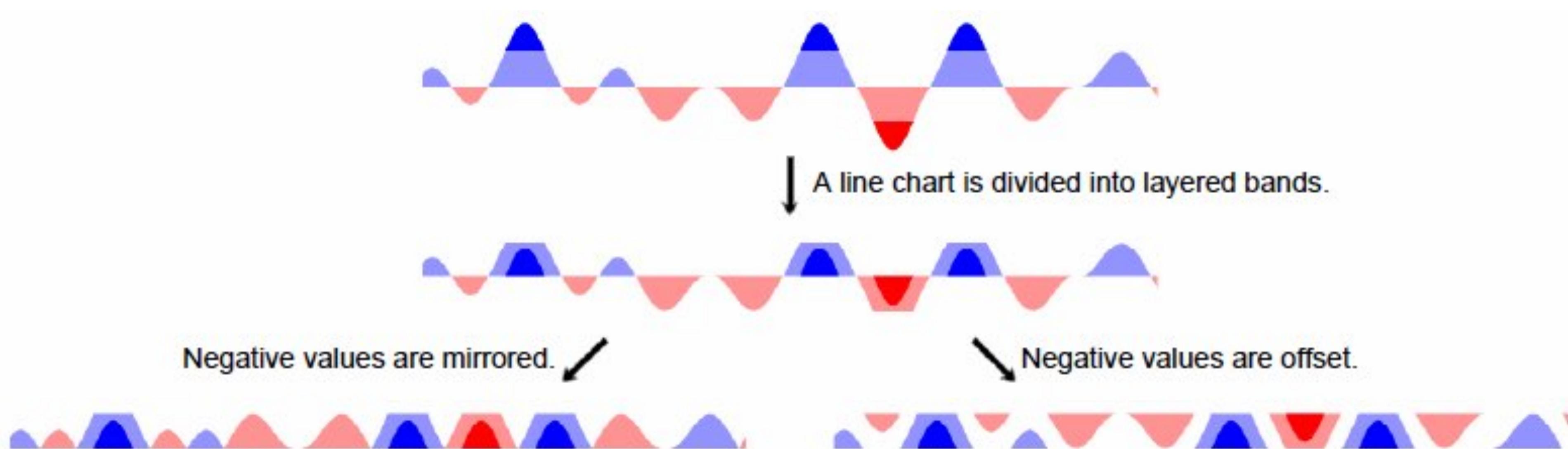


**Figure 1.** (a) **Filled line chart.** Area between data values on line and zero is filled in. (b) “**Mirrored**” chart. Negative values are flipped and colored red, cutting the chart height by half. (c) **2-band horizon graph.** The chart is divided into bands and overlaid, again halving the height.

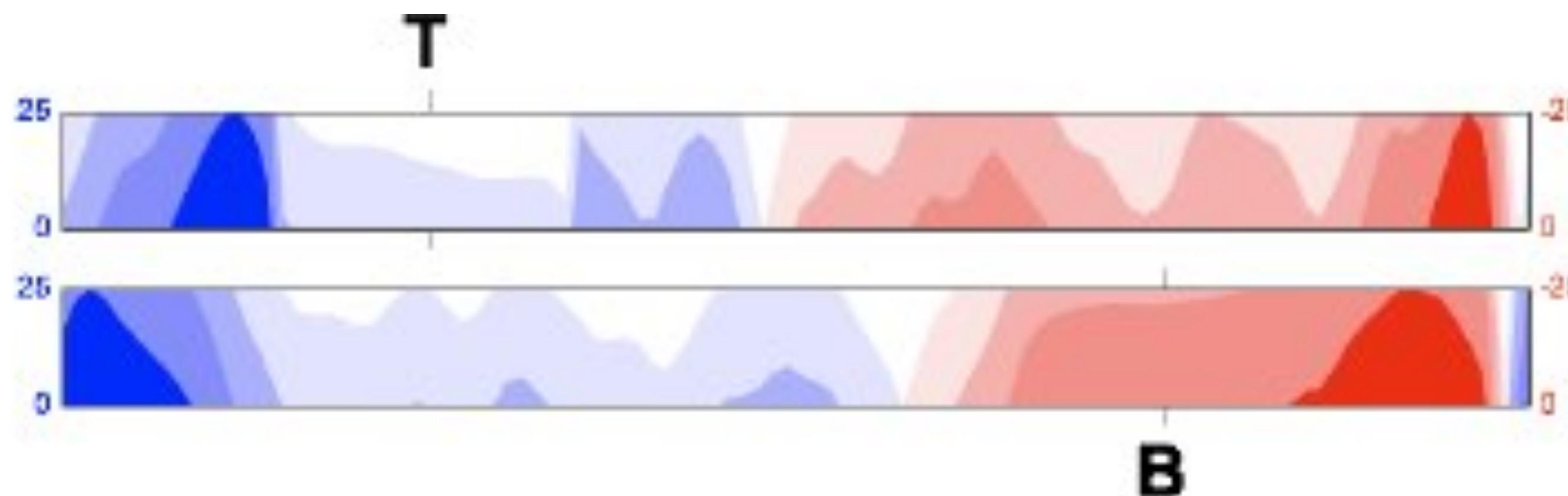
Como o espelhamento e o uso de camadas para representação de séries temporais afeta a habilidade de análise?

Como os valores negativos espelhados são interpretados?

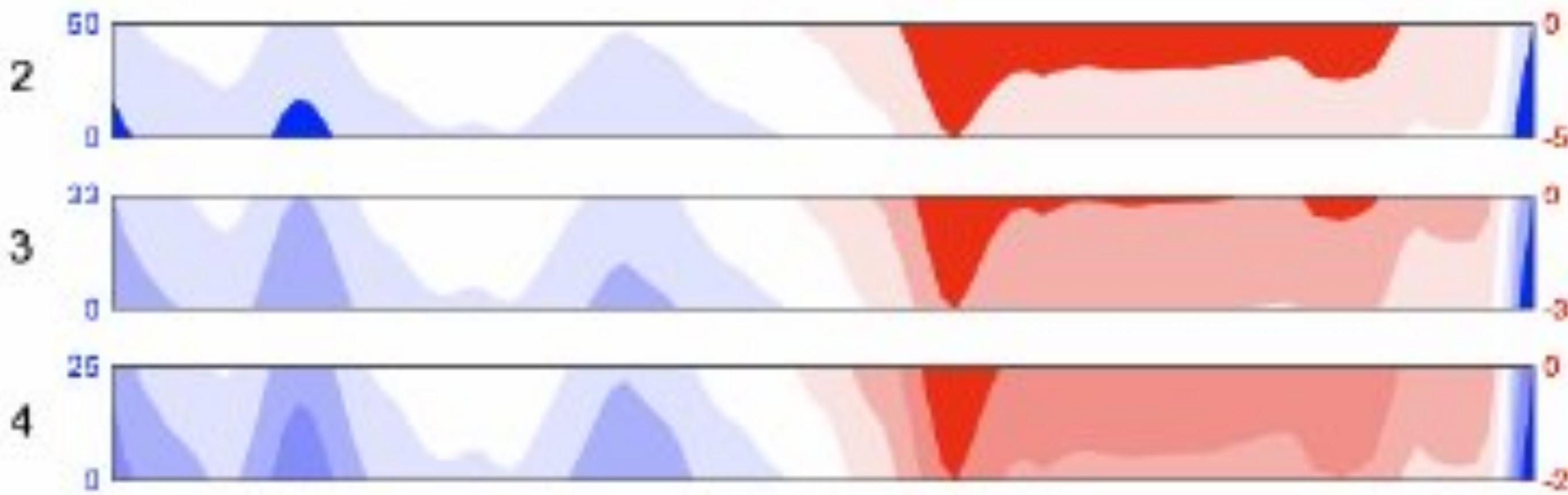
O desempilhamento mental das camadas do gráfico interferem com a estimativa dos valores?



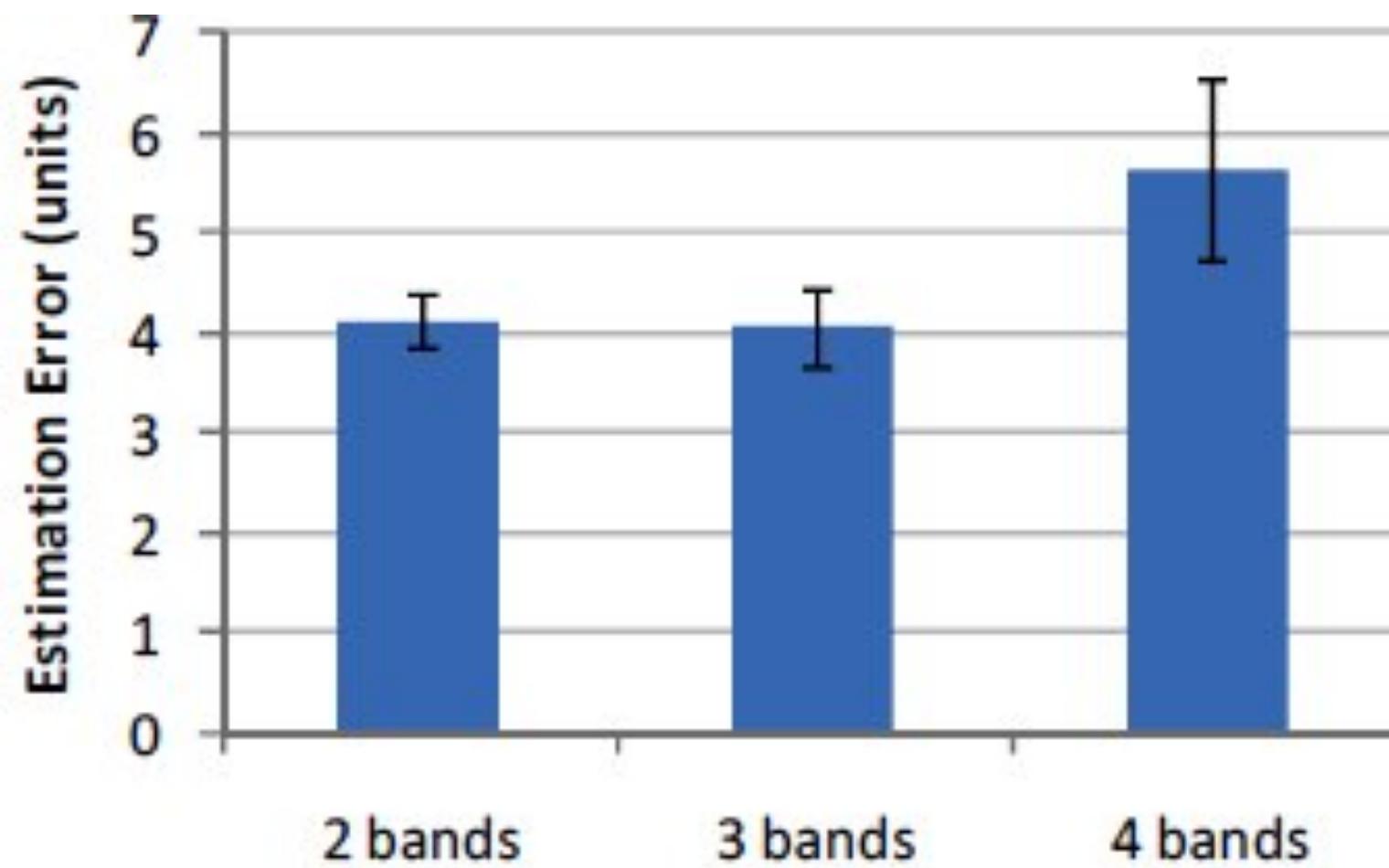
**Figure 2. Horizon graph construction.** A normal line chart is divided into bands defined by uniform value ranges. The bands are then layered to reduce the chart height. Negative values can be mirrored or offset into the same space as positive values.



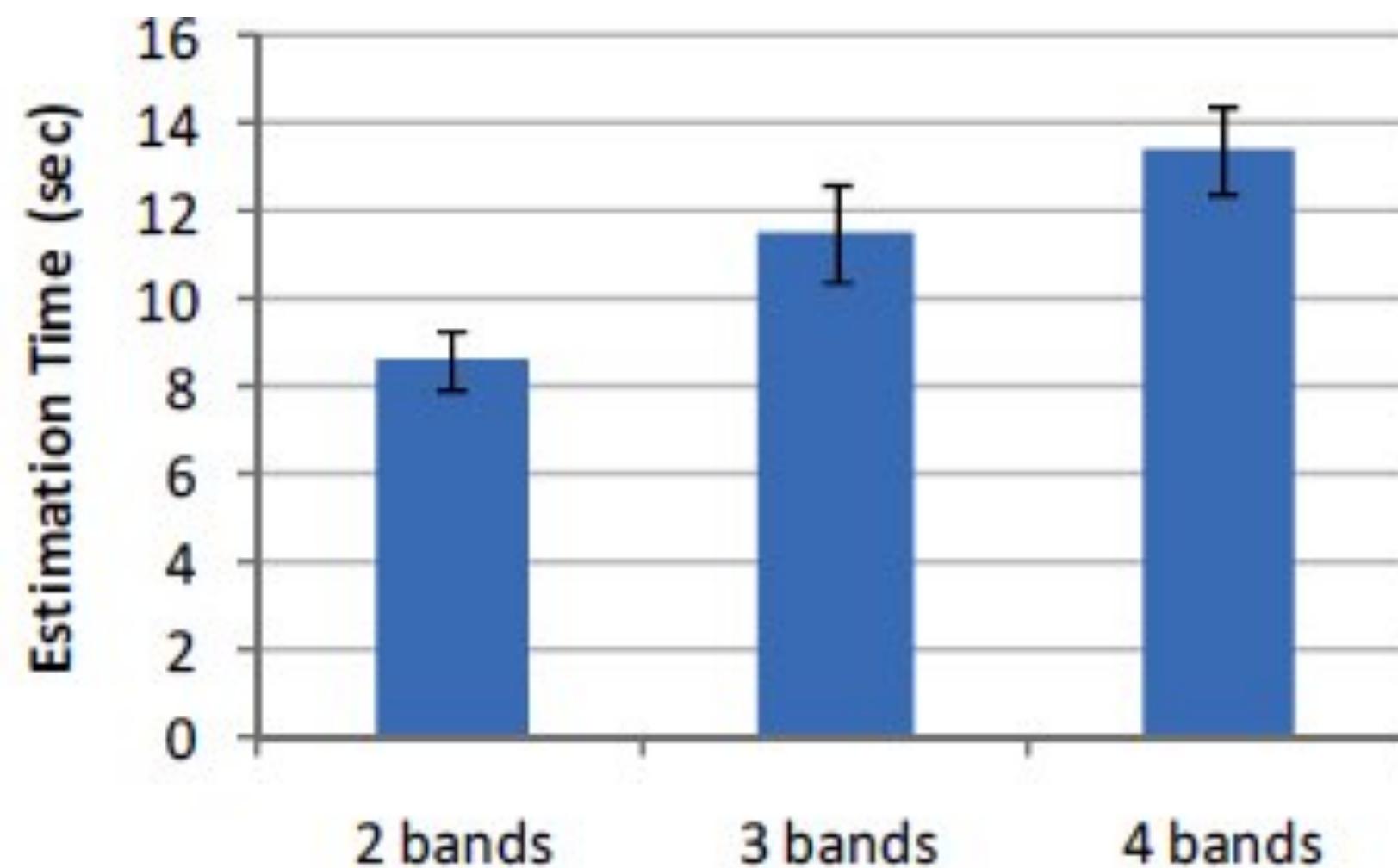
**Figure 3. Example trial with a 4-band mirrored graph.**  
Each band covers 25 values; the total range is [-100, 100].  
Subjects reported if T or B was larger, and by how much.



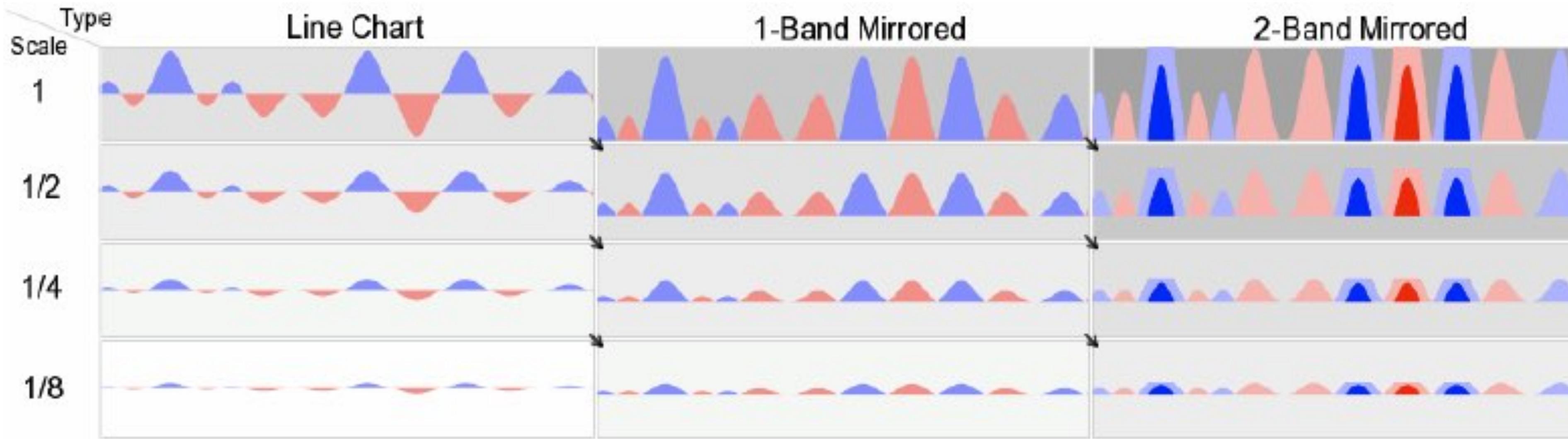
**Figure 4. Offset horizon graphs with 2, 3, and 4 bands.**



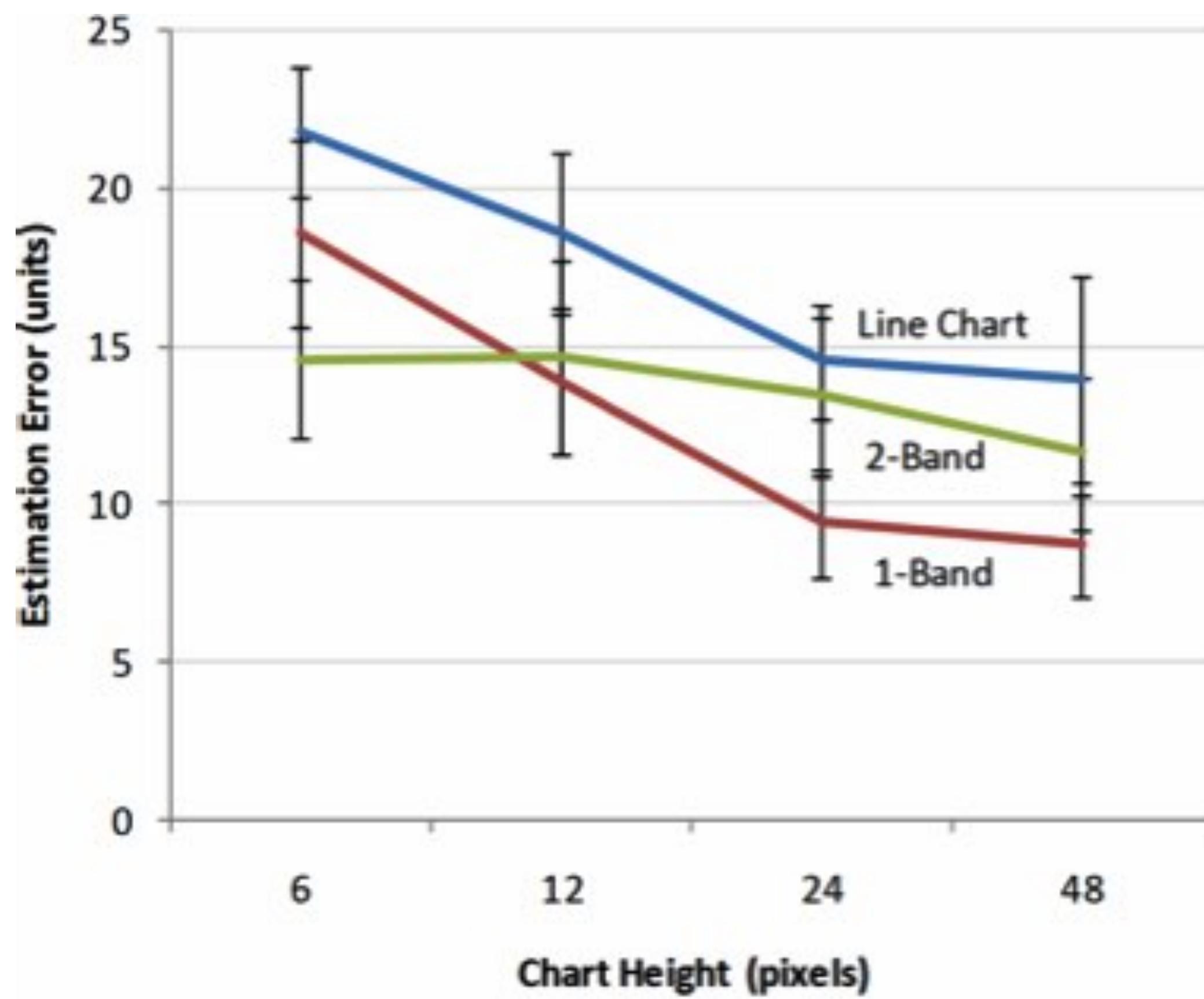
**Figure 5. Estimation Error by Band Count.** 4-band charts have significantly higher error than 2- or 3-band charts.



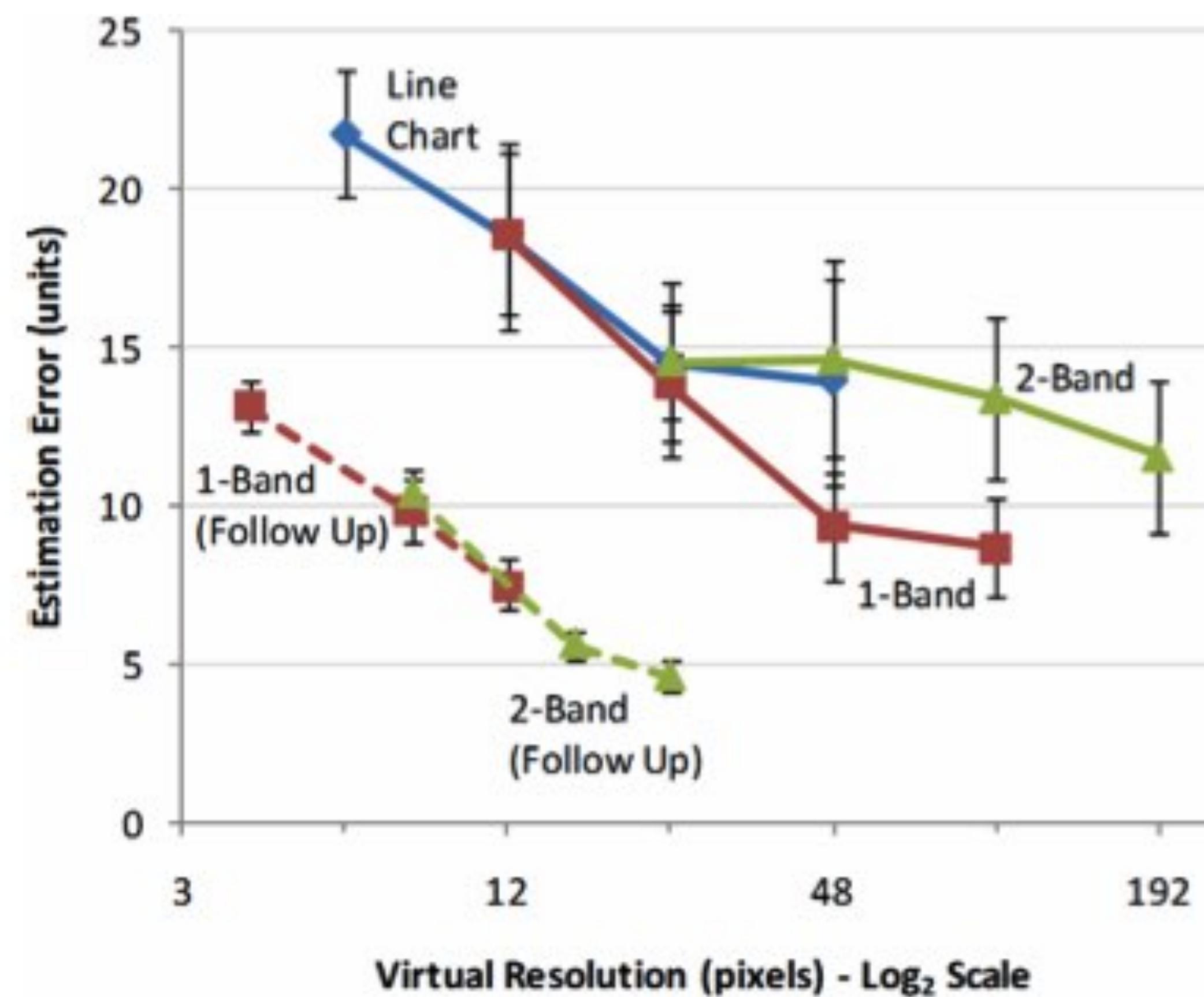
**Figure 6. Estimation Time by Band Count.** Estimation time increases significantly with each additional band.



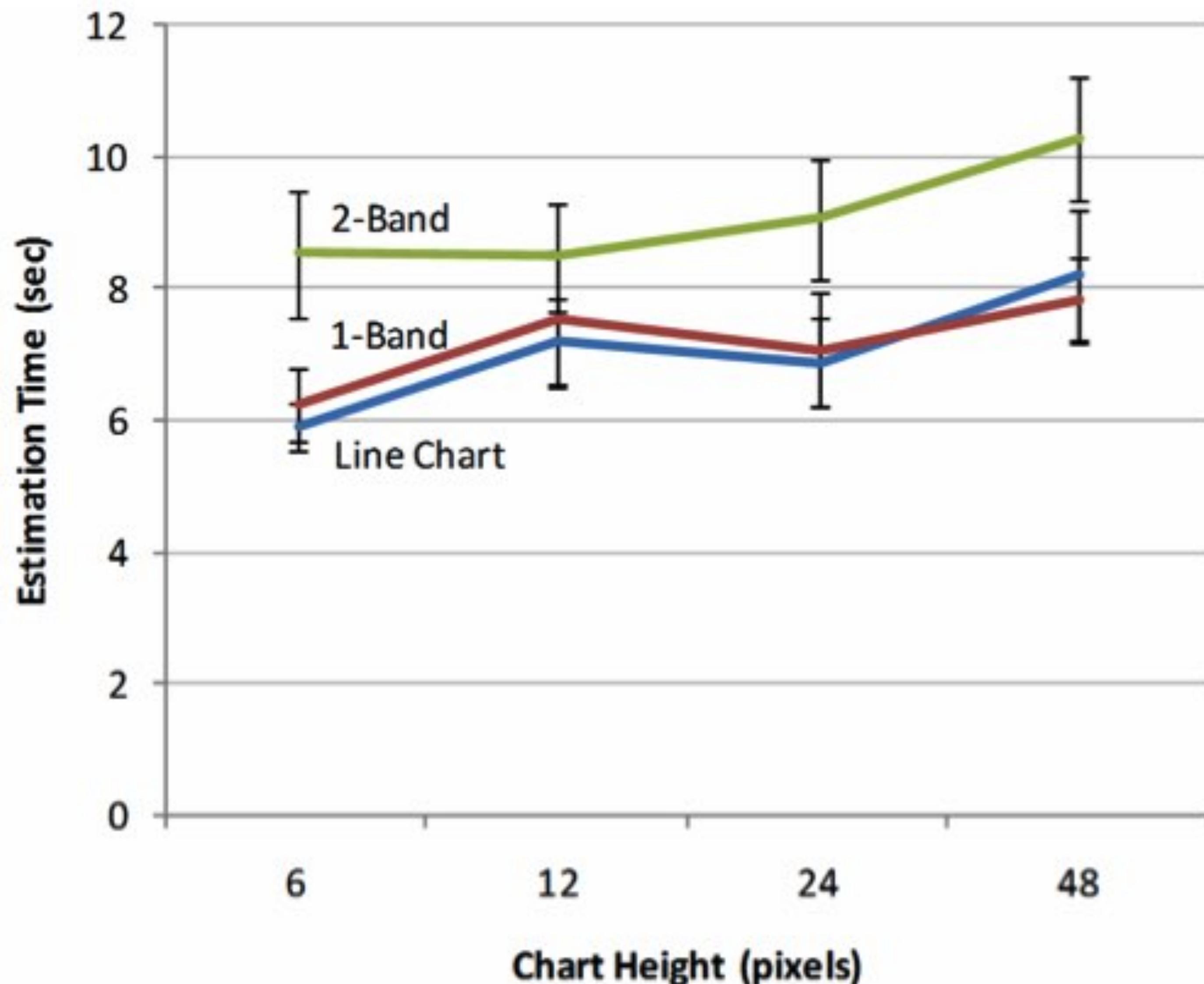
**Figure 7. Chart Type and Scale Conditions in Experiment 2.** We crossed 3 chart types and 4 chart heights. The diagonally adjacent cells indicated by arrows and shading have the same *virtual resolution*: the un-mirrored, un-layered size of the chart.



**Figure 8. Estimation Error by Chart Type and Height.** The 2-band mirror chart crosses the 1-band case at a chart height of 12 pixels (scale factor 1/4).



**Figure 9. Estimation Error by Chart Type and Virtual Resolution.** Error levels hold relatively stable at high virtual resolutions, but increase linearly at smaller resolutions.



**Figure 10. Estimation Time by Chart Type and Height.**  
Line and 1-band mirror charts result in similar estimation times. Both are significantly faster than 2-band charts.

# AVALIAÇÃO

---

- O espelhamento não tem efeito sobre o tempo e a acurácia
- As camadas são benéficas quando o tamanho do gráfico diminui (gráficos com altura menor que 6.8mm tiveram melhores resultados com 2 faixas)
- Não é recomendável usar 4 faixas ou mais: os usuários relataram cansaço para estimar os valores

# VISUALIZAÇÃO DE SÉRIES TEMPORAIS E MINERAÇÃO DE DADOS

---

# PRINCIPAIS TAREFAS

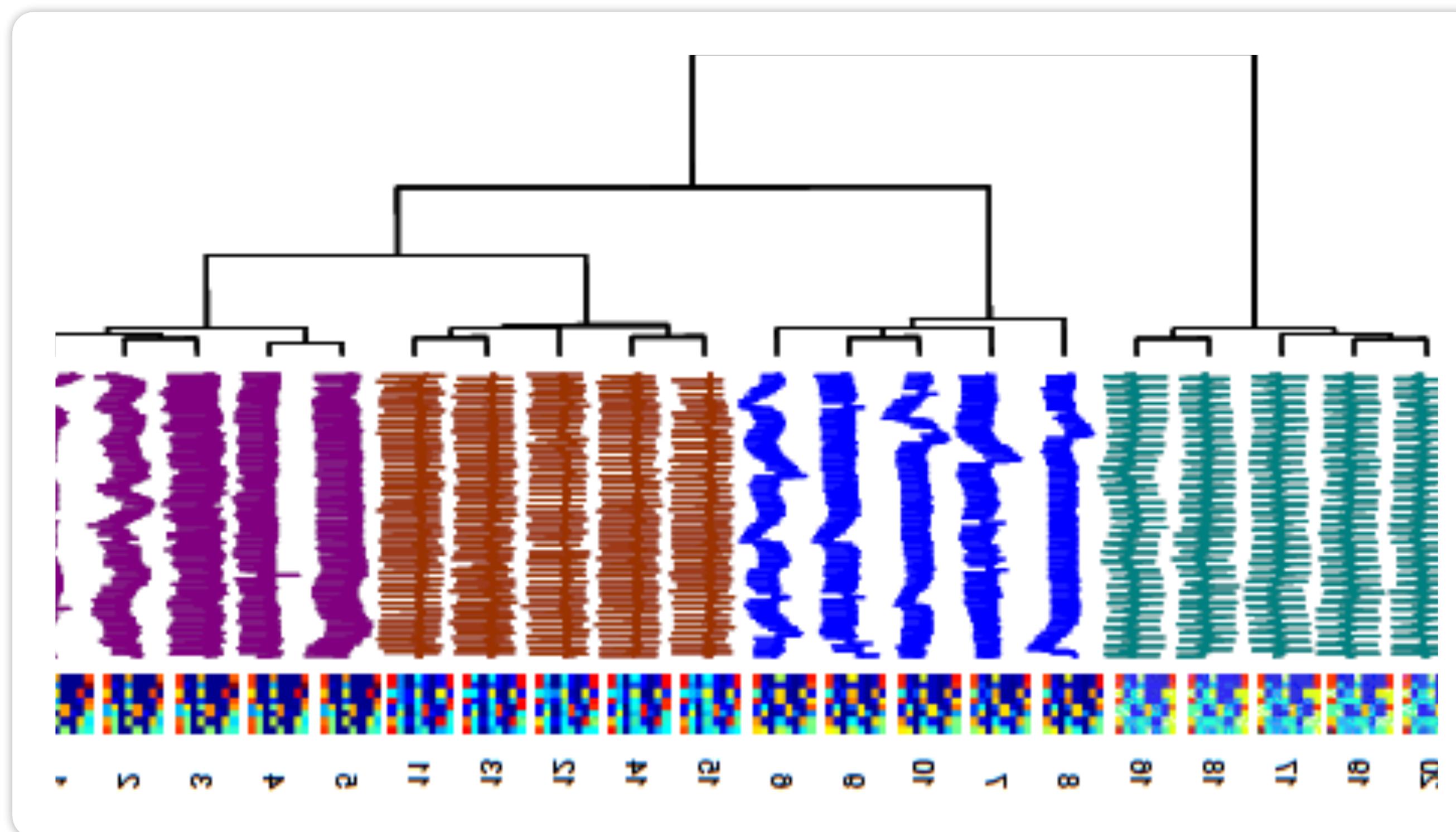
---

- **Agrupamento:** encontrar grupos de séries temporais similares
- **Classificação:** assinalar uma ou mais classes a uma nova série temporal
- **Detecção de anomalias**
  
- **Indexação:** dada uma série temporal e uma métrica de similaridade, encontrar a série temporal mais próxima

# TIME-SERIES BITMAPS: A PRACTICAL VISUALIZATION TOOL FOR WORKING WITH LARGE TIME SERIES DATABASES

---

N. Kuman, N. Lolla, E. Keogh, S. Lonardi e C.  
Hatanamahatana  
SIAM  
2005



# OBJETIVOS

---

- Mineração de dados de séries temporais facilitando a navegação em grandes volumes de séries temporais
- Extração de características de séries temporais e representação visual em bitmaps
- Agrupamento e detecção de regularidades e / ou anomalias em várias séries temporais

# COMO COMPARAR SÉRIES TEMPORAIS

---

- Distância euclidiana no espaço n-dimensional
- Dynamic time warping (DTW)
- Symbolic Aggregate ApproXimation (SAX)

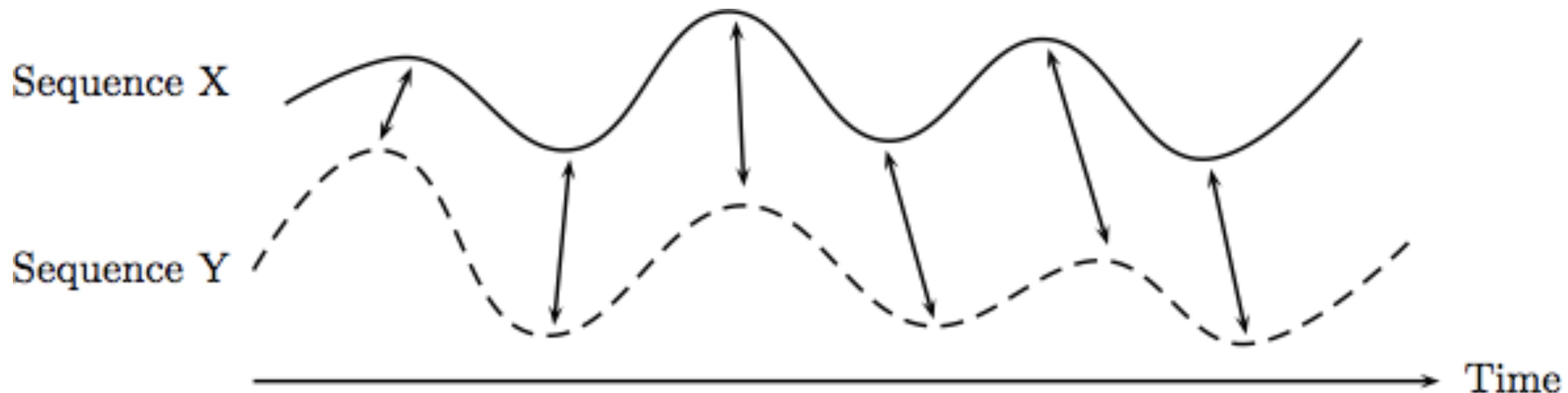
# DYNAMIC TIME WARPING (DTW)

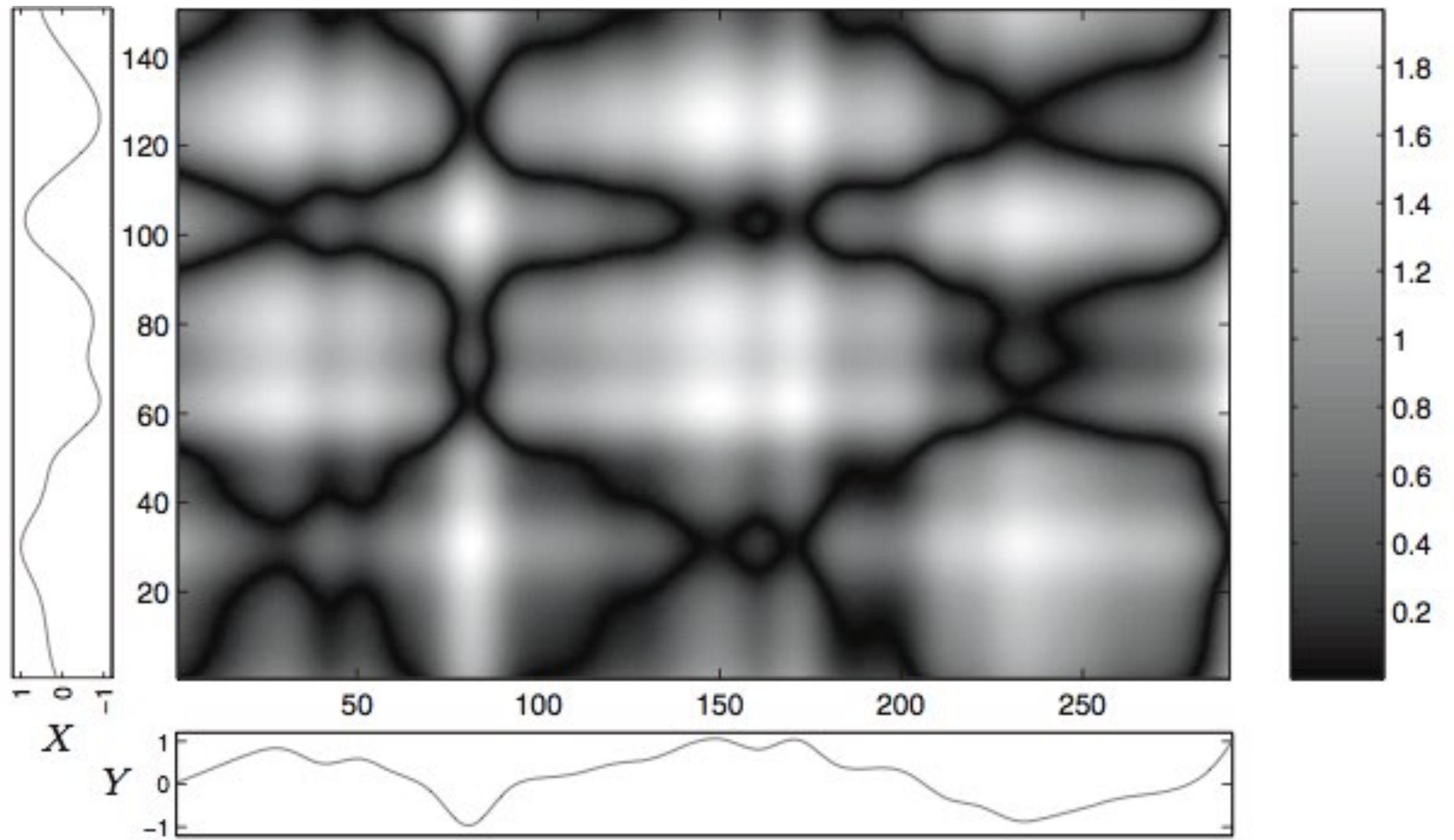
---

*Dynamic time warping* (DTW) é uma técnica comumente usada para alinhamento entre séries temporais

Um par de sequências é distorcido de forma não linear para ser casado

DTW foi usado inicialmente para comparar padrões de fala em sistemas de reconhecimento automático de fala





**Fig. 4.2.** Cost matrix of the two real-valued sequences  $X$  (*vertical axis*) and  $Y$  (*horizontal axis*) using the Manhattan distance (absolute value of the difference) as local cost measure  $c$ . Regions of low cost are indicated by *dark colors* and regions of high cost are indicated by *light colors*

**Algorithm: OPTIMALWARPINGPATH**

**Input:** Accumulated cost matrix  $D$ .

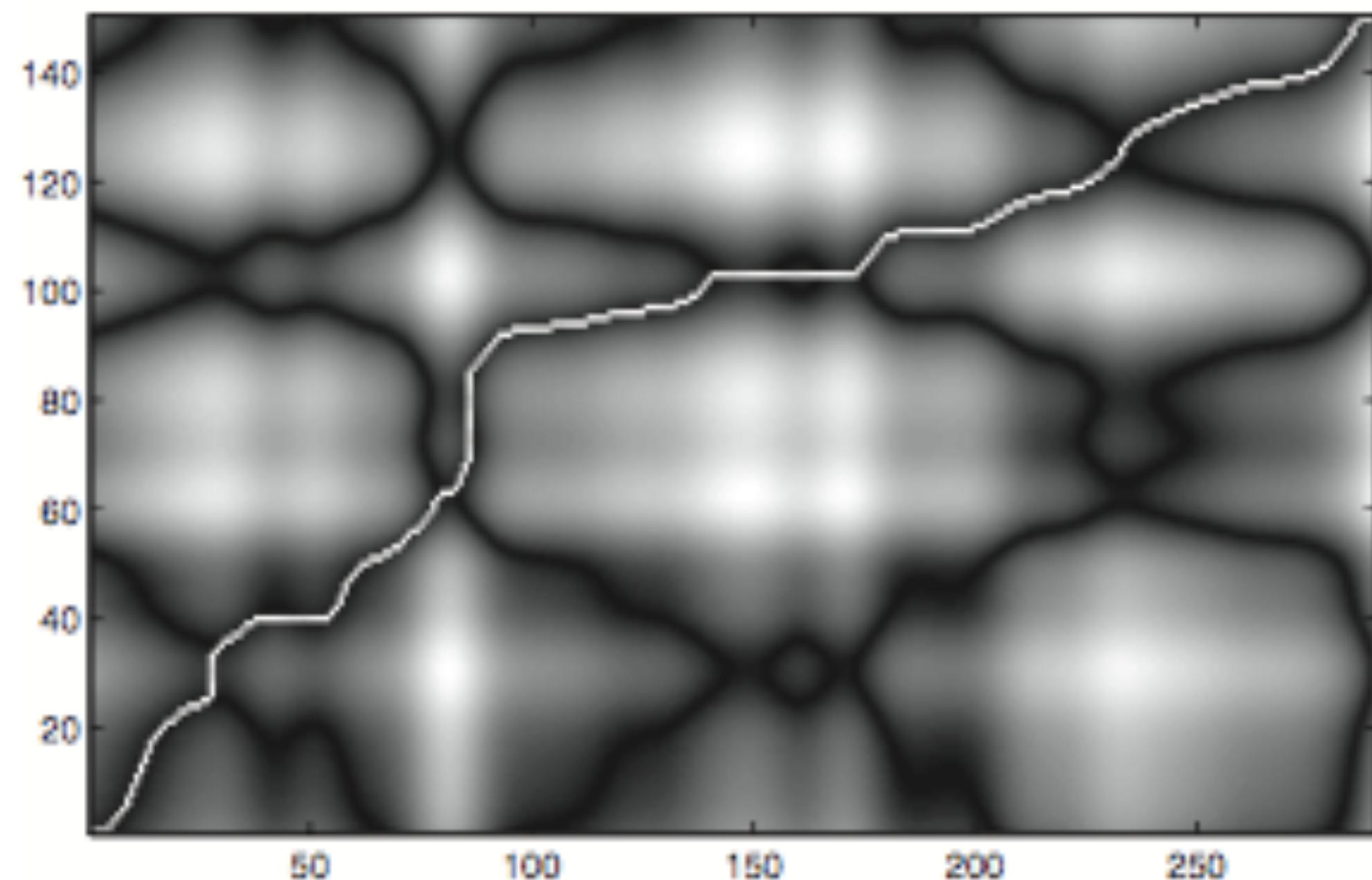
**Output:** Optimal warping path  $p^*$ .

**Procedure:** The optimal path  $p^* = (p_1, \dots, p_L)$  is computed in reverse order of the indices starting with  $p_L = (N, M)$ . Suppose  $p_\ell = (n, m)$  has been computed. In case  $(n, m) = (1, 1)$ , one must have  $\ell = 1$  and we are finished. Otherwise,

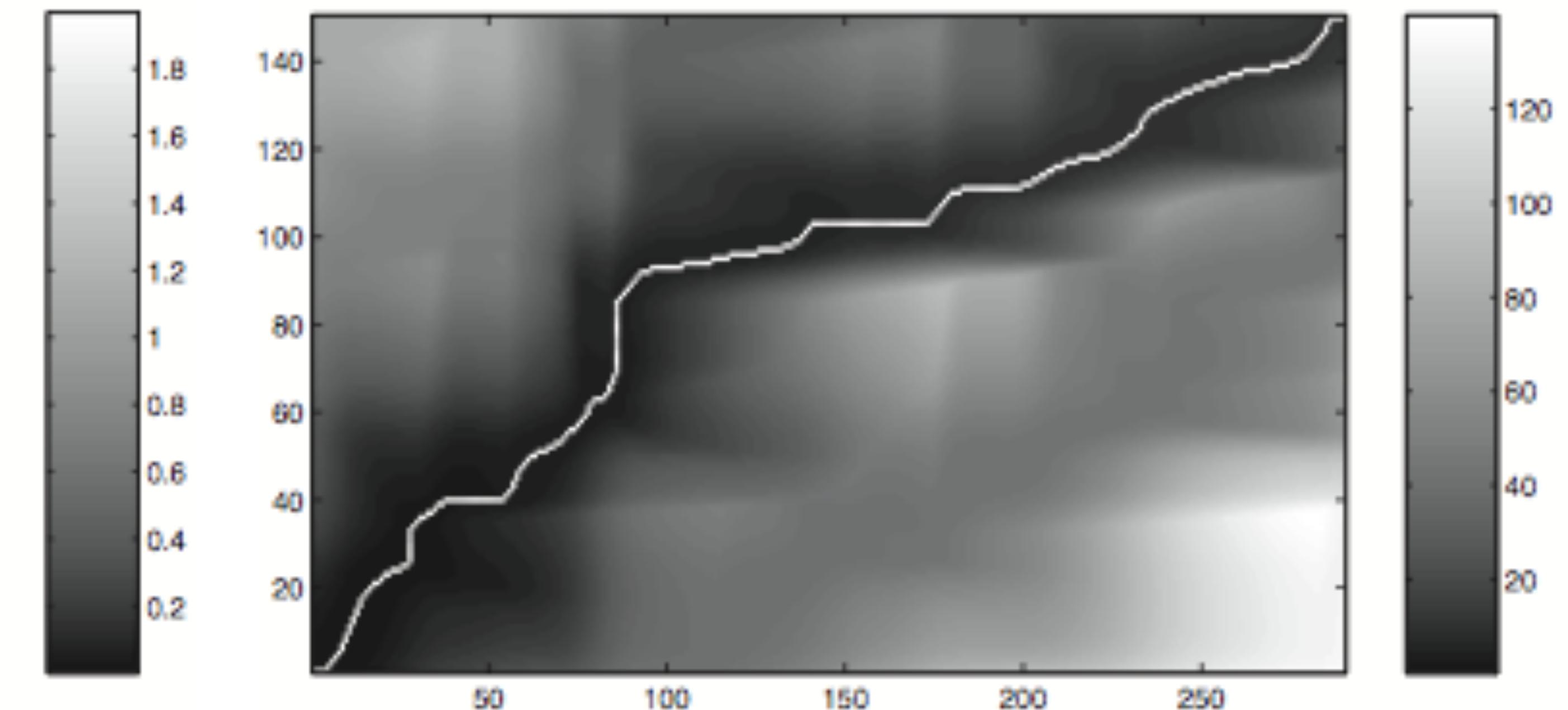
$$p_{\ell-1} := \begin{cases} (1, m-1), & \text{if } n = 1 \\ (n-1, 1), & \text{if } m = 1 \\ \operatorname{argmin}\{D(n-1, m-1), \\ \quad D(n-1, m), D(n, m-1)\}, & \text{otherwise,} \end{cases} \quad (4.6)$$

where we take the lexicographically smallest pair in case “ $\operatorname{argmin}$ ” is not unique.

(a)



(b)



**Fig. 4.4.** (a) Cost matrix  $C$  as in Fig. 4.2 and (b) accumulated cost matrix  $D$  with optimal warping path  $p^*$  (*white line*)

# **SYMBOLIC AGGREGATE APPROXIMATION (SAX)**

---

Consiste em agregar diversos símbolos em uma matriz  $2^q \times 2^q$ , na qual cada posição apresenta uma correspondência 1:1 para a frequência de uma subsequência de tamanho  $q$

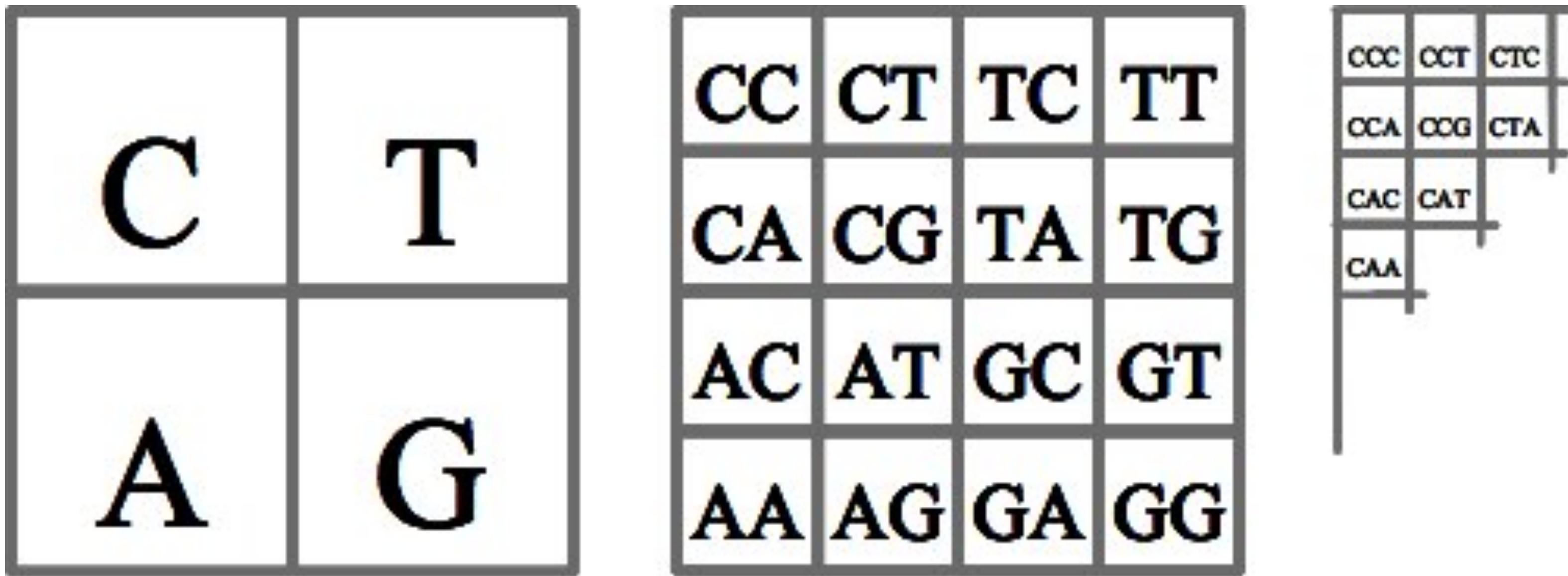
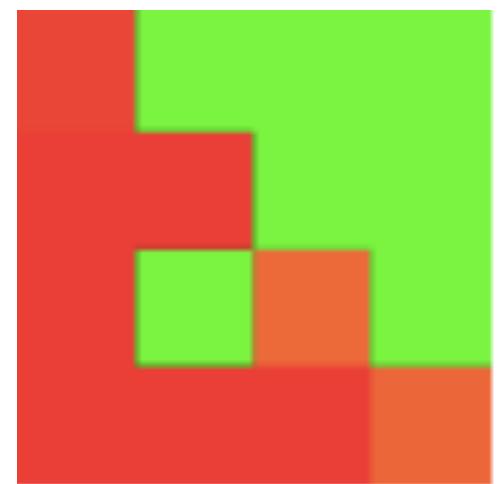
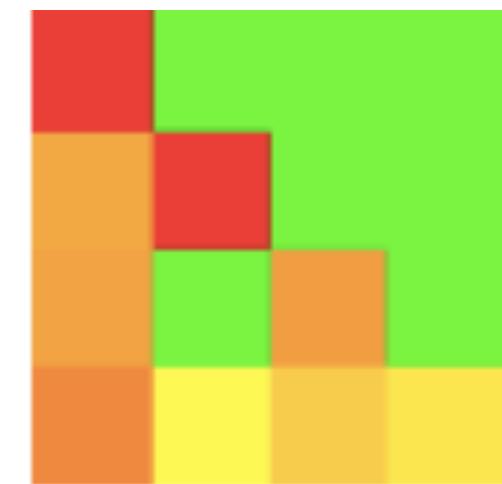


Figure 5: The quad-tree representation of a sequence over the alphabet {A,C,G,T} at different levels of resolution



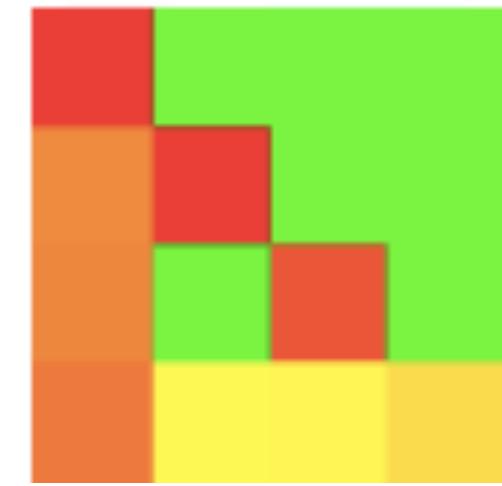
Pan troglodytes



Elephas maximus



Homo sapiens



Loxodonta africana

The gene sequences of mitochondrial DNA of four animals, used to create their own icons using a chaos game representation. Note that *Pan troglodytes* is the familiar Chimpanzee, and *Loxodonta africana* and *Elephas maximus* are the African and Indian Elephants respectively. The file icons show that humans and chimpanzees have similar genomes, as do the African and Indian elephants.

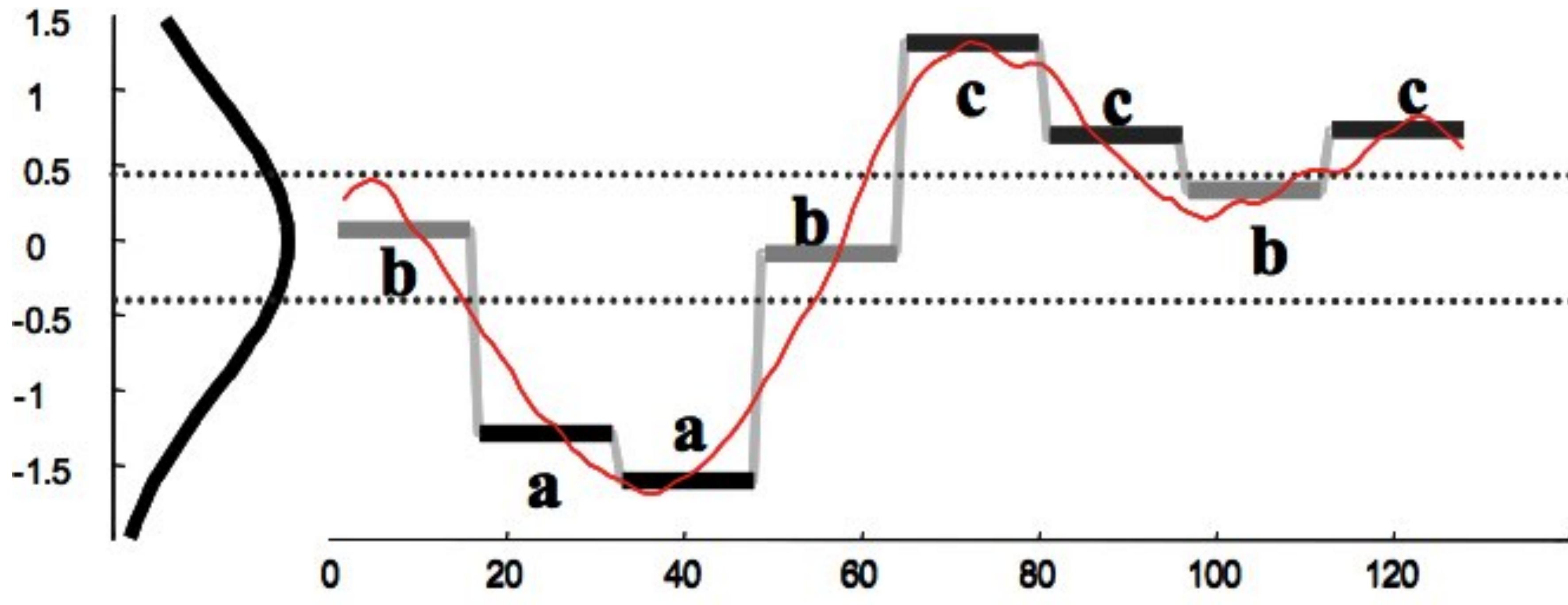


Figure 7: A real valued time series can be converted to the SAX word **baabccbc**. Note that all three possible symbols are approximately equally frequent.

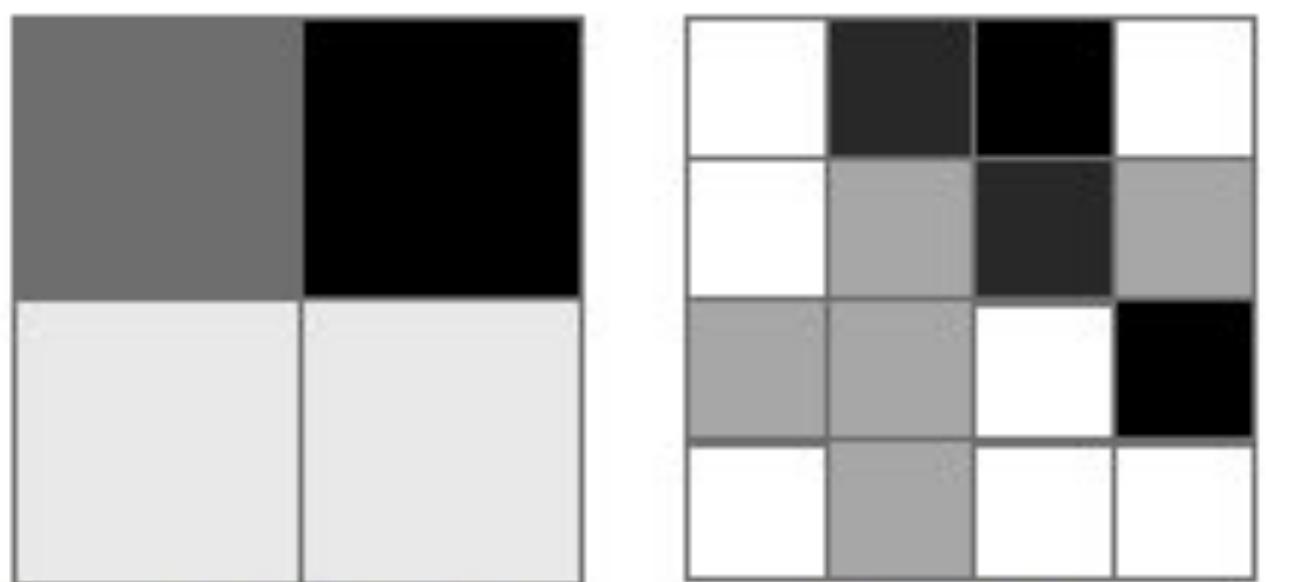
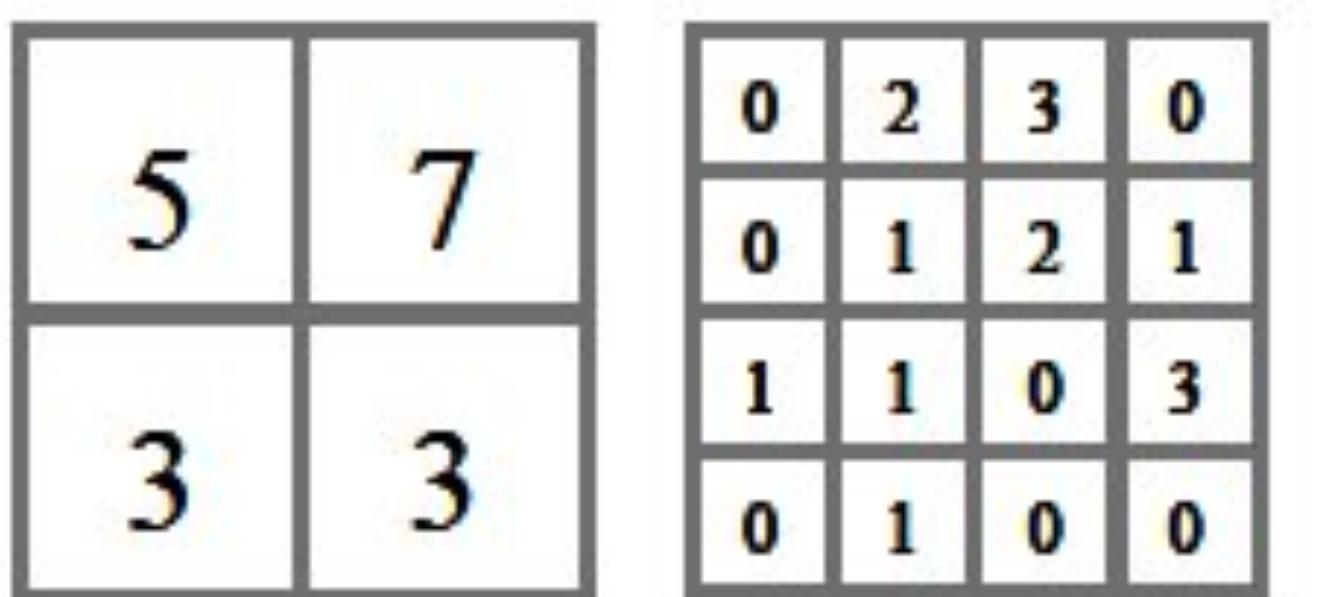
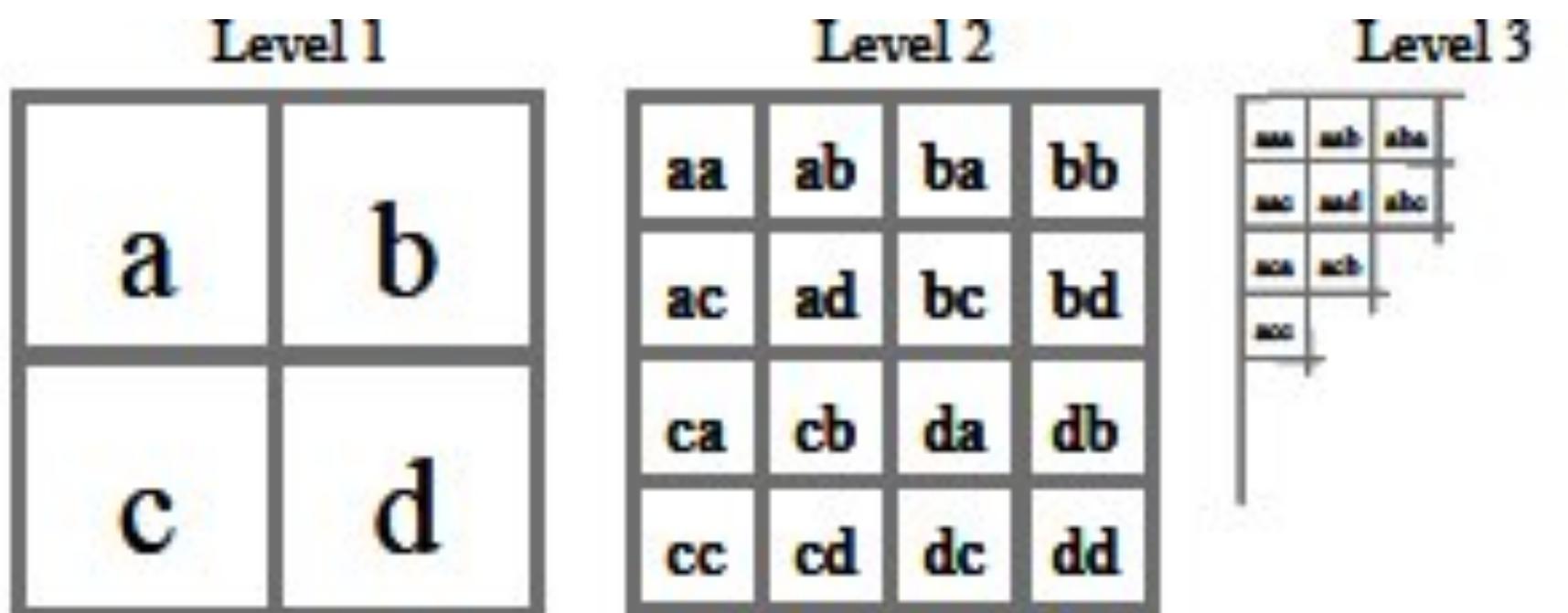
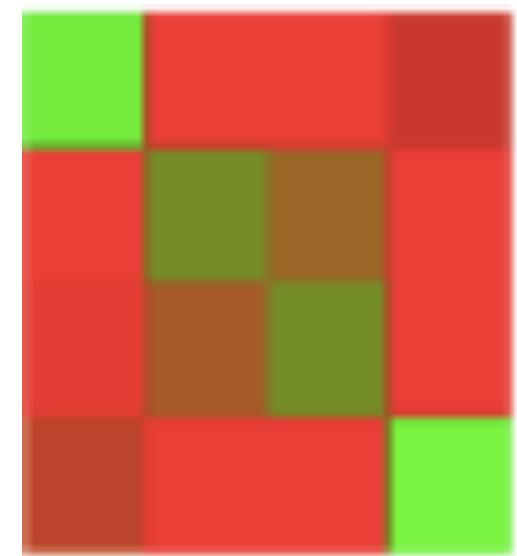


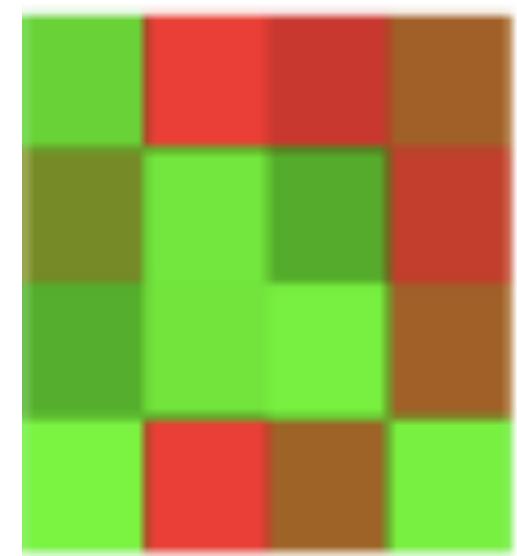
Figure 8: *Top*) The four possible SAX symbols are mapped to four quadrants of a square, and pairs, triplets etc are recursively mapped to finer grids. *Middle*) We can extract counts of symbols from a SAX representation and record them in the grids. *Bottom*) The recorded values can be linearly mapped to colors, thus creating a square bitmap.



example a.dat



example b.dat

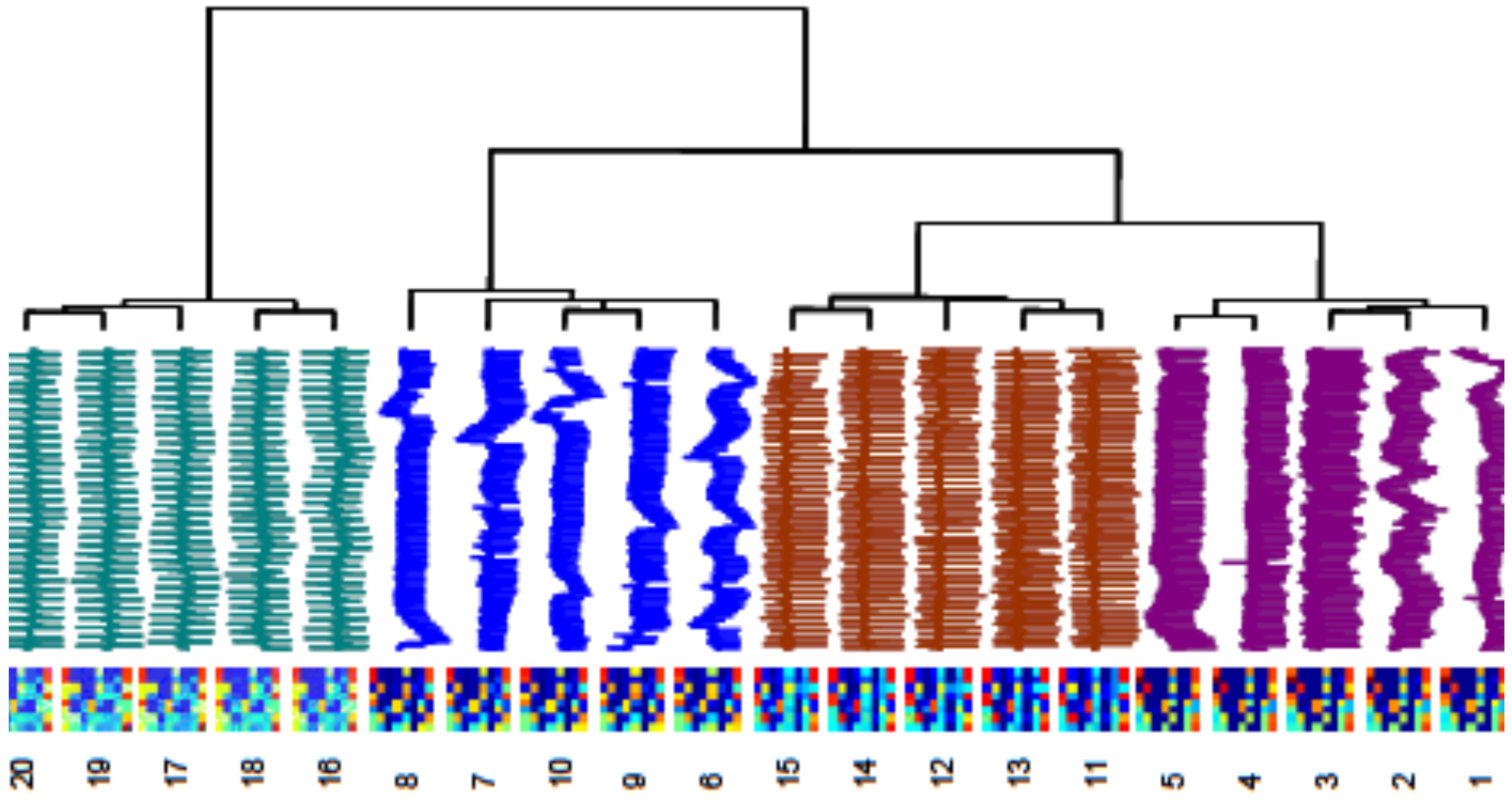


example c.dat



example d.dat

Four time series files represented as time series bitmaps. While they are all examples of EEGs, example\_a.dat is from a normal trace, whereas the others contain examples of spike-wave discharges. The fact that there is some difference between one dataset and all the rest is immediately apparent from a casual inspection of the bitmap representation.



[www.cs.umd.edu/hcil/timesearcher/](http://www.cs.umd.edu/hcil/timesearcher/)

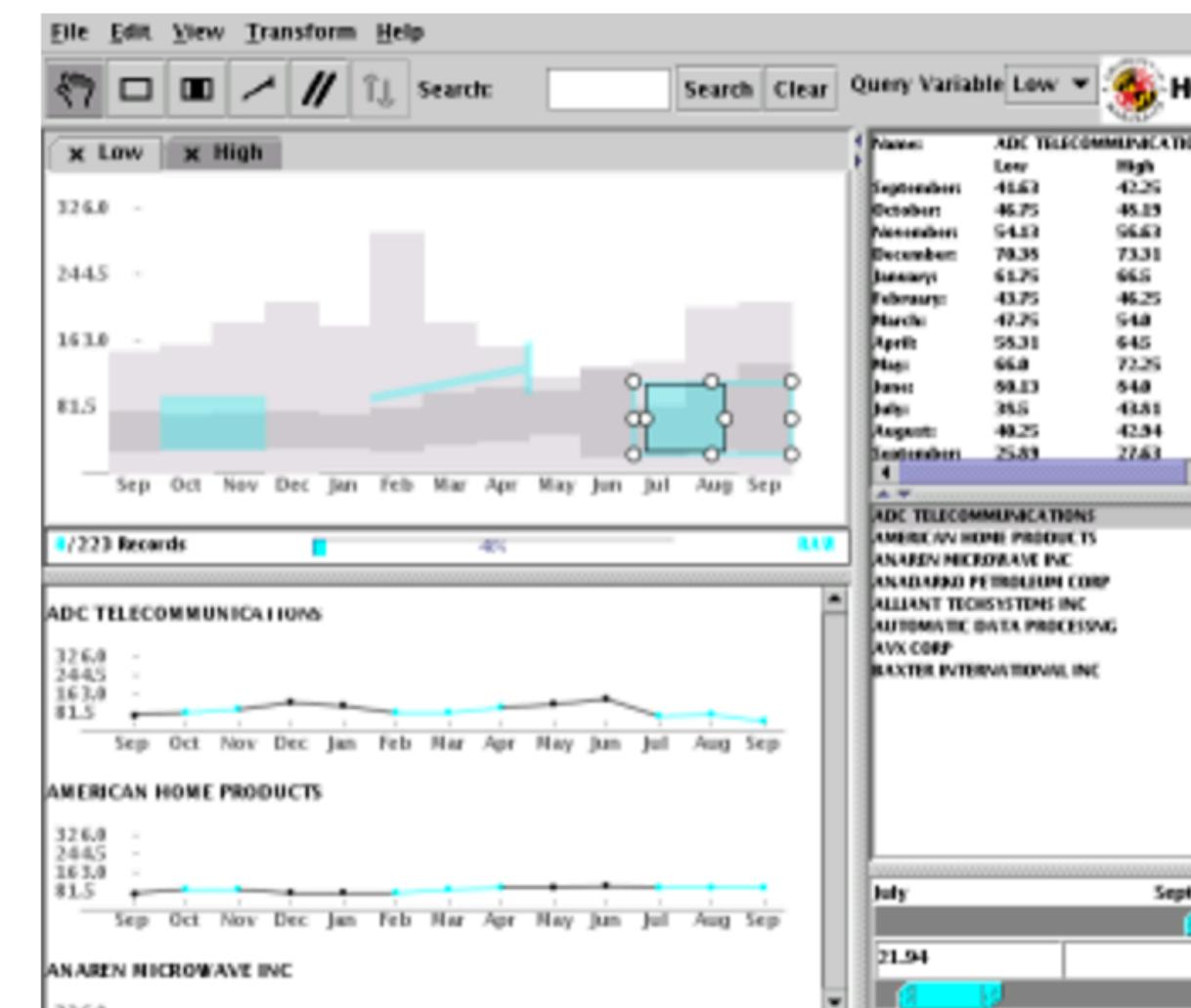


Human-Computer Interaction Lab  
UNIVERSITY OF MARYLAND

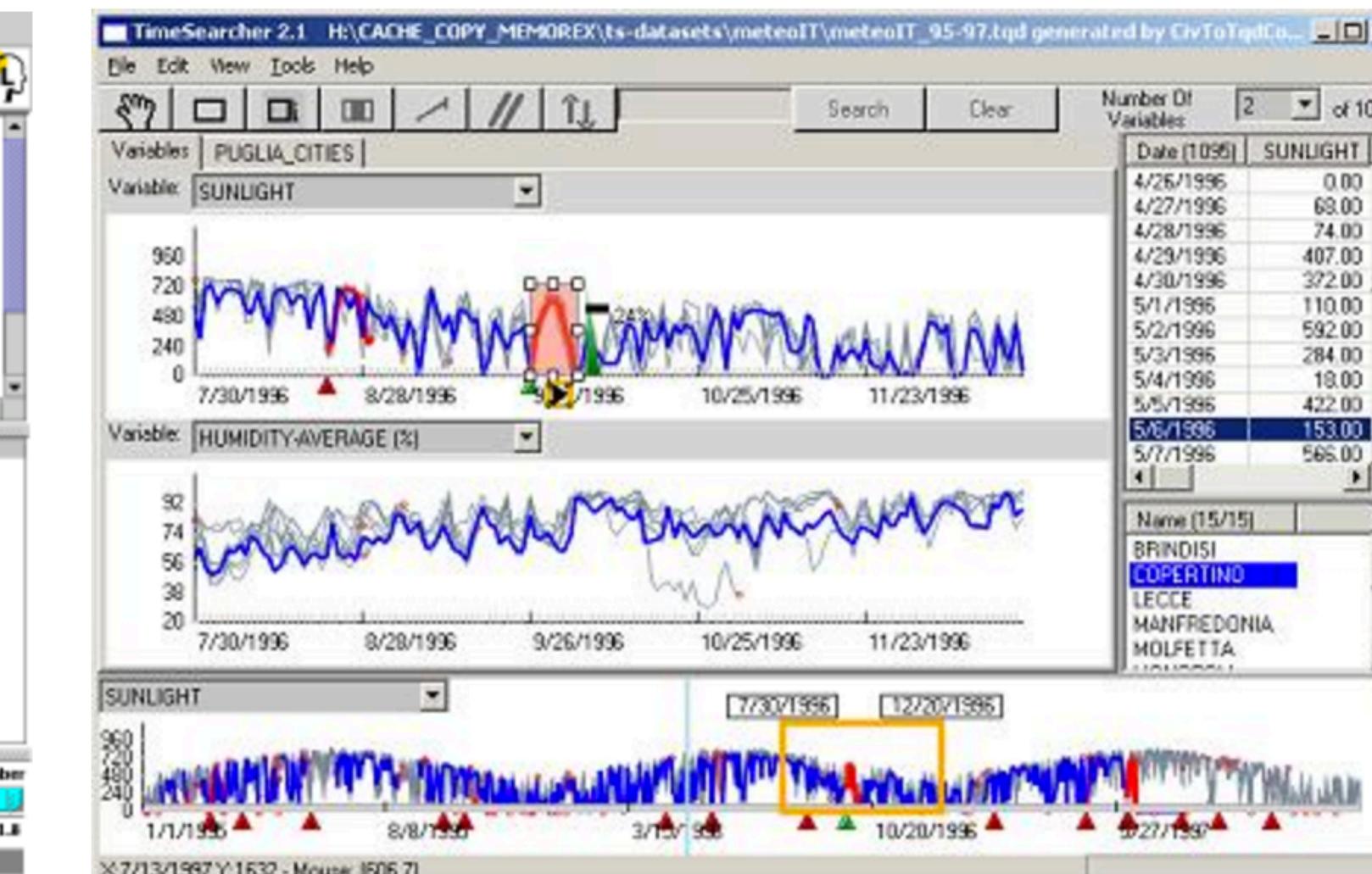
News + Events | About HCIL | People | **Research** | Publications | Academics | Sponsorship

Piccolo

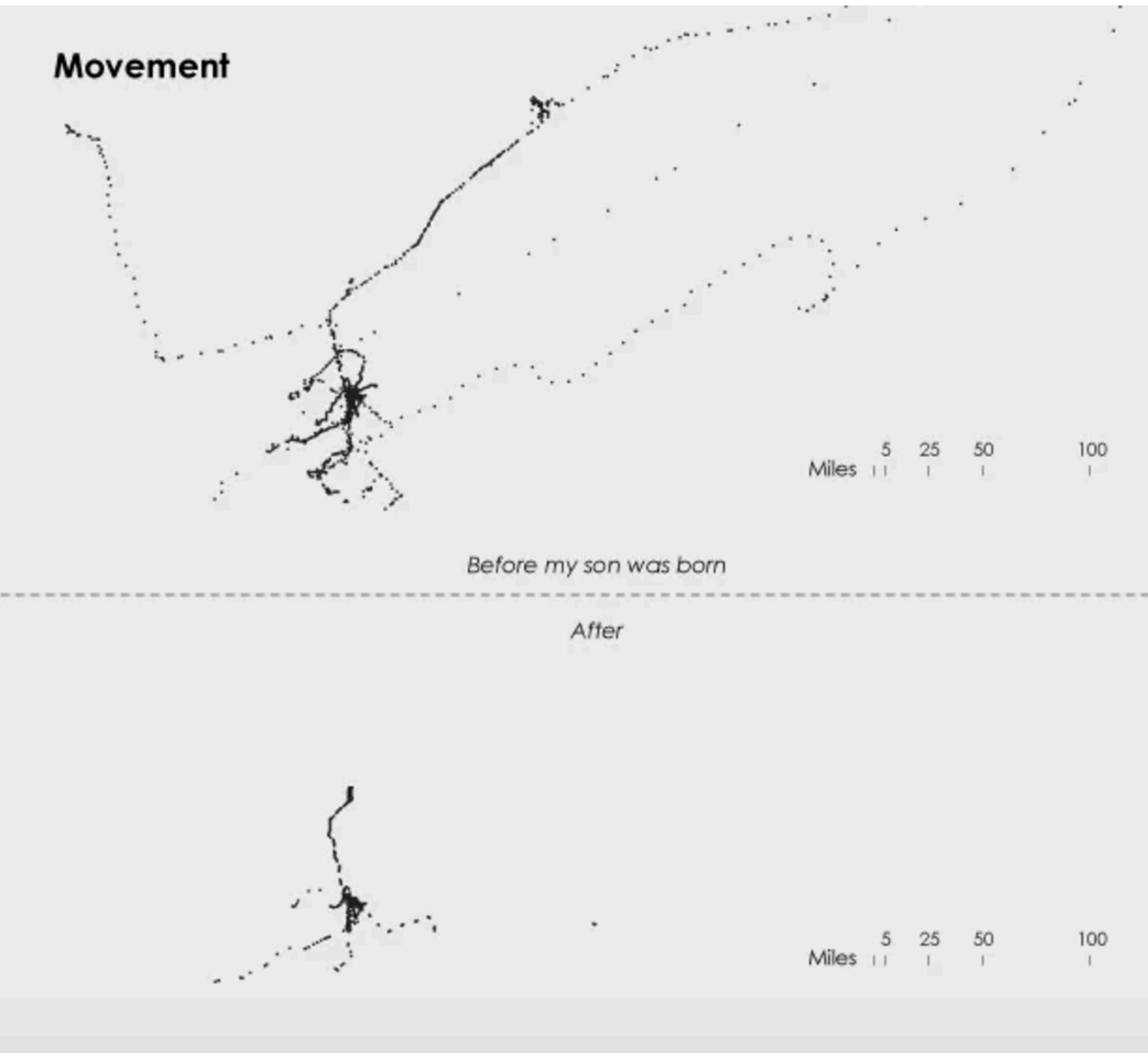
## Visual Exploration of Time-Series Data



[TimeSearcher 1](#)

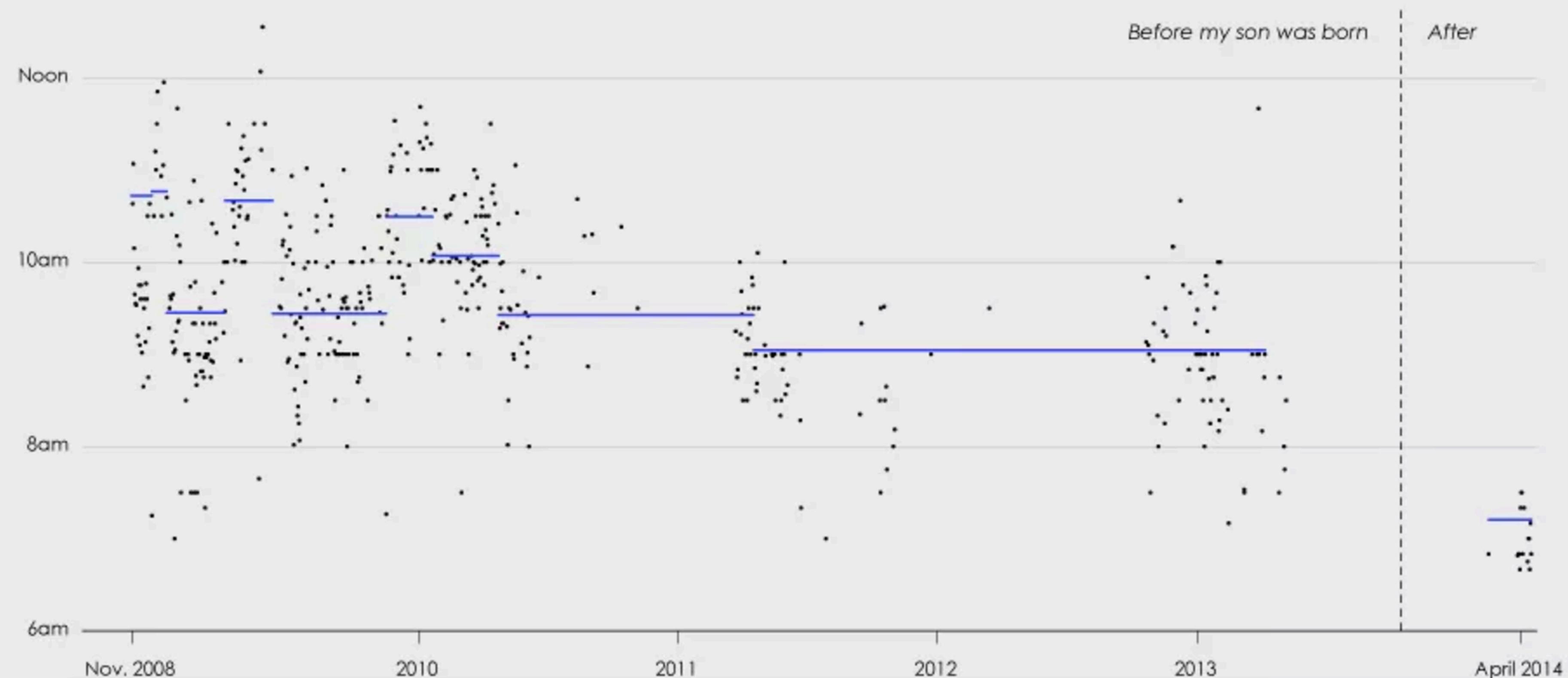


[TimeSearcher 2](#)

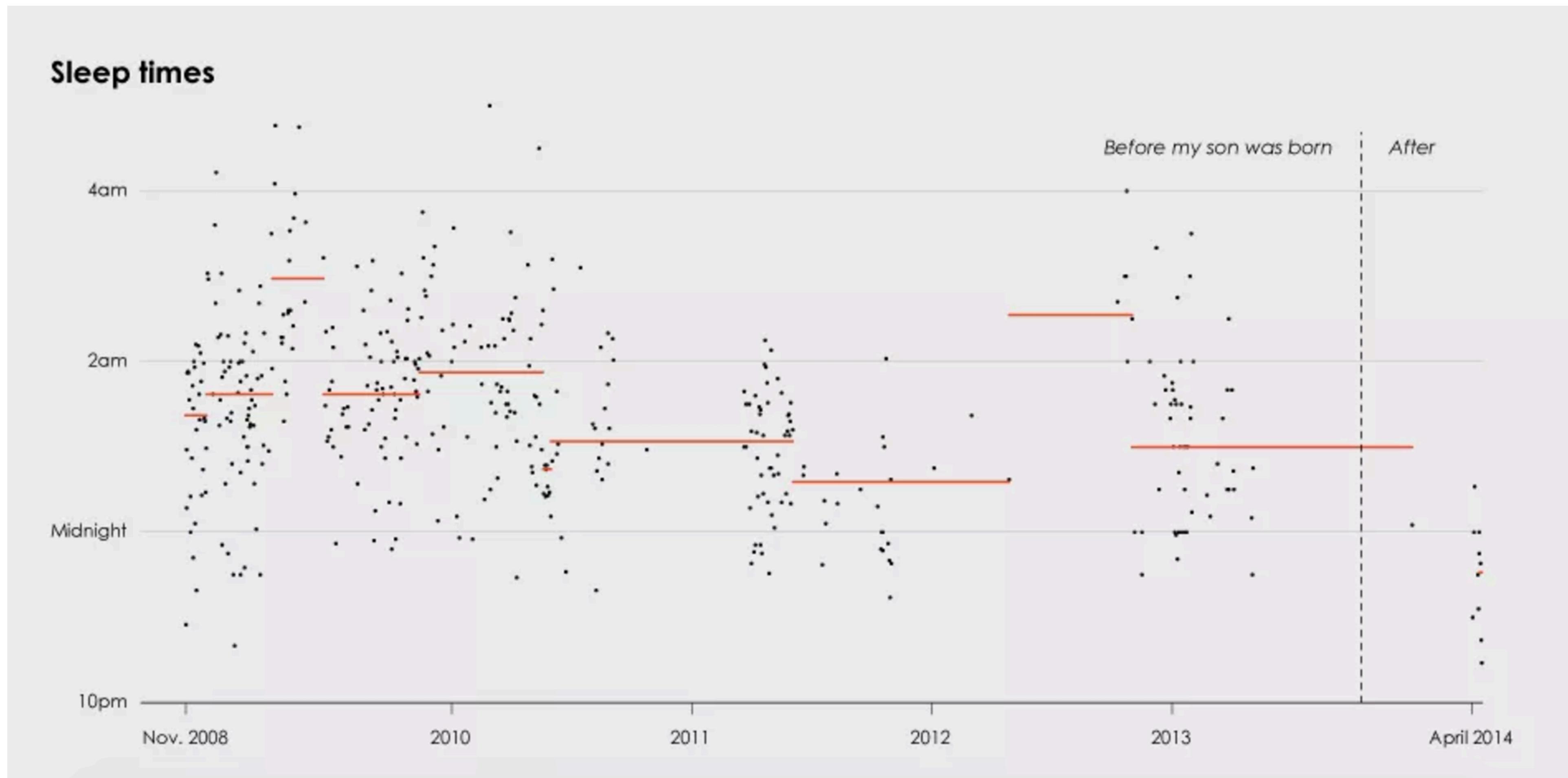


<https://flowingdata.com/2014/04/23/the-difference-a-data-point-makes/>

## Wakeup times



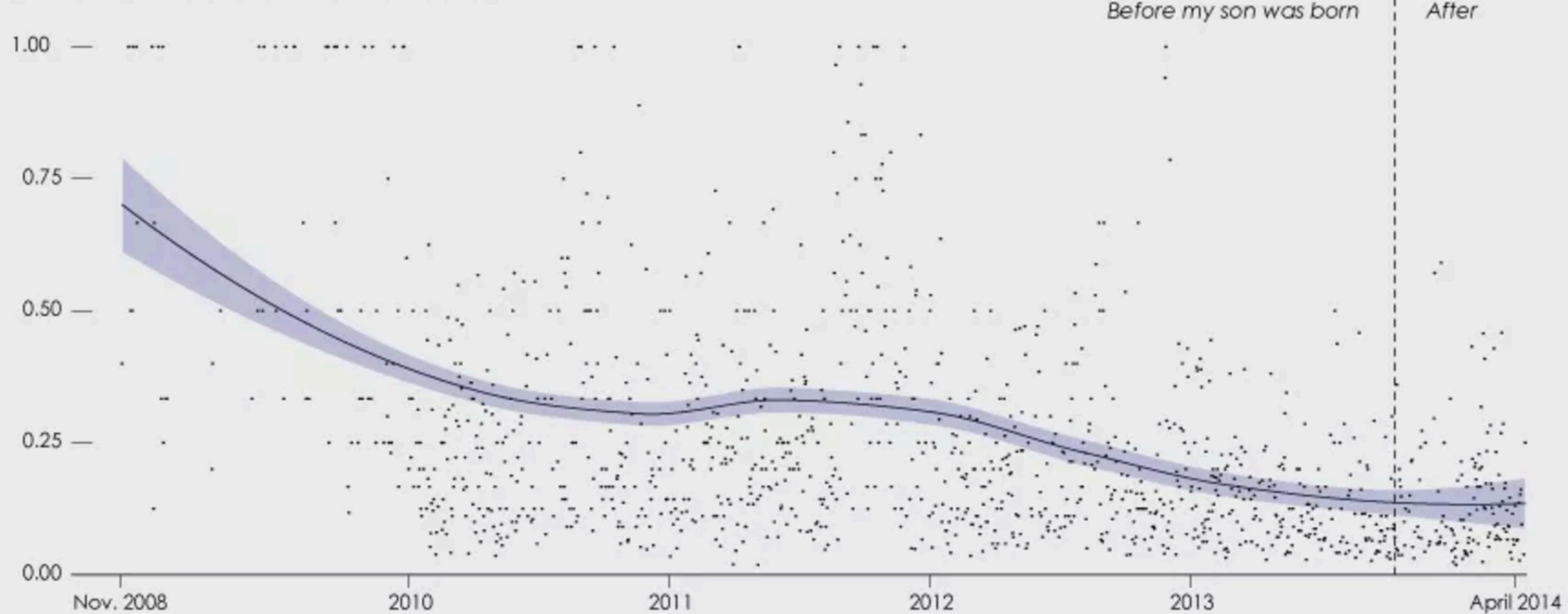
## Sleep times



## Email and sleep



### Ratio of email sent to received



### Email sent per month



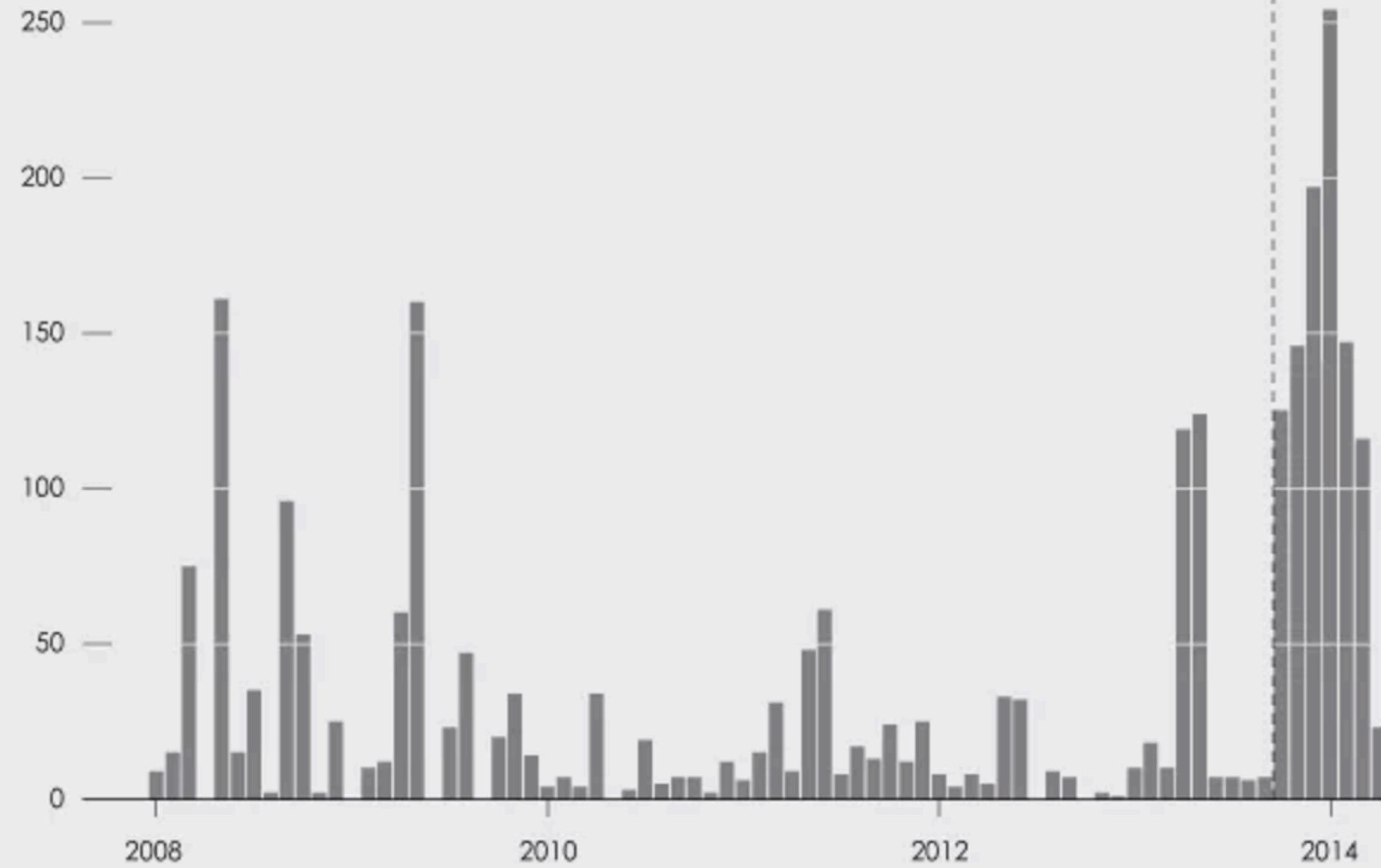
### Email received per month



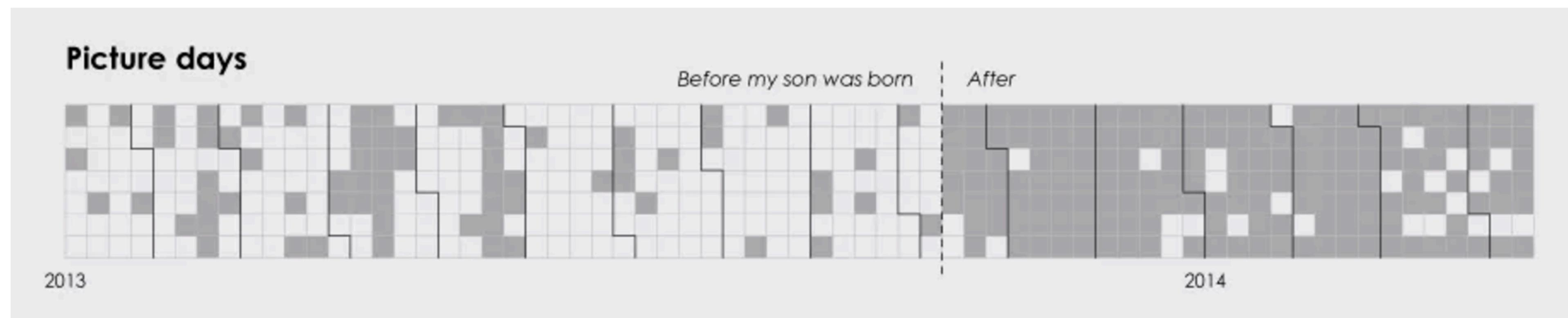
## Monthly photos

*Before my son was born*

*After*



## Picture days



That's my boy. The one on the left. The human one.

## F1 Timeline

Race count

WIDESCREEN

WINNER PODIUM TOP 10 OTHER

Twitter i

0 50 100 150 200 250 300



# The TimeViz Browser

A Visual Survey of Visualization Techniques for Time-Oriented Data

by Christian Tominski and Wolfgang Aigner

# of Techniques: 115

Search:

How to use filters:

- Want:** Show me!
- Indifferent:** I don't care.
- Hide:** I'm not interested!

Data

Frame of Reference

- |  |  |  |
|--|--|--|
|  |  |  |
|  |  |  |

Number of Variables

- |  |  |  |
|--|--|--|
|  |  |  |
|  |  |  |

Time

Arrangement

- |  |  |  |
|--|--|--|
|  |  |  |
|  |  |  |

