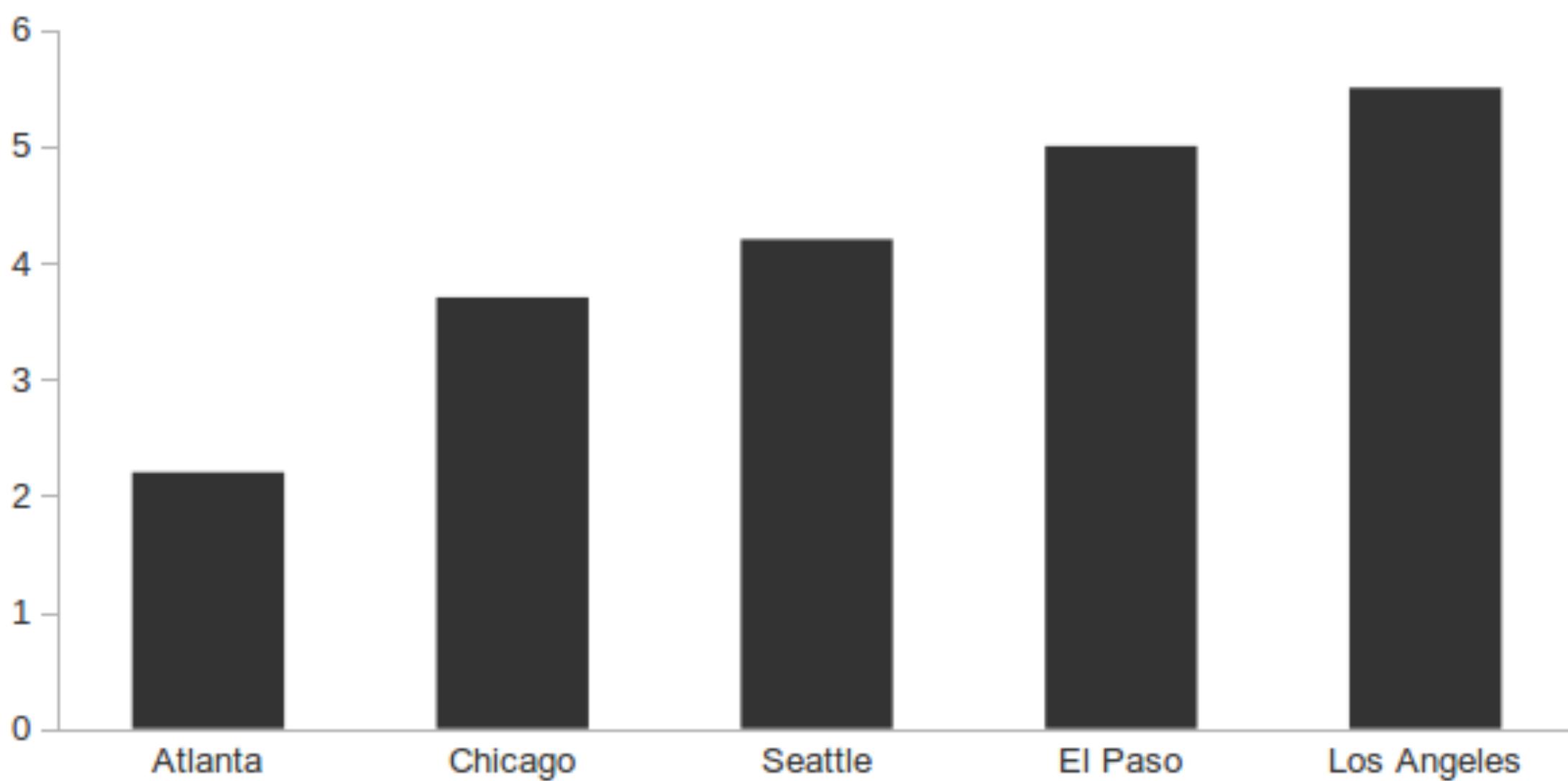


Distribuições

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

- Análise de como um conjunto de valores está distribuído em um intervalo
- ou ainda
- Contrastar como um ou mais conjunto de valores se distribuem

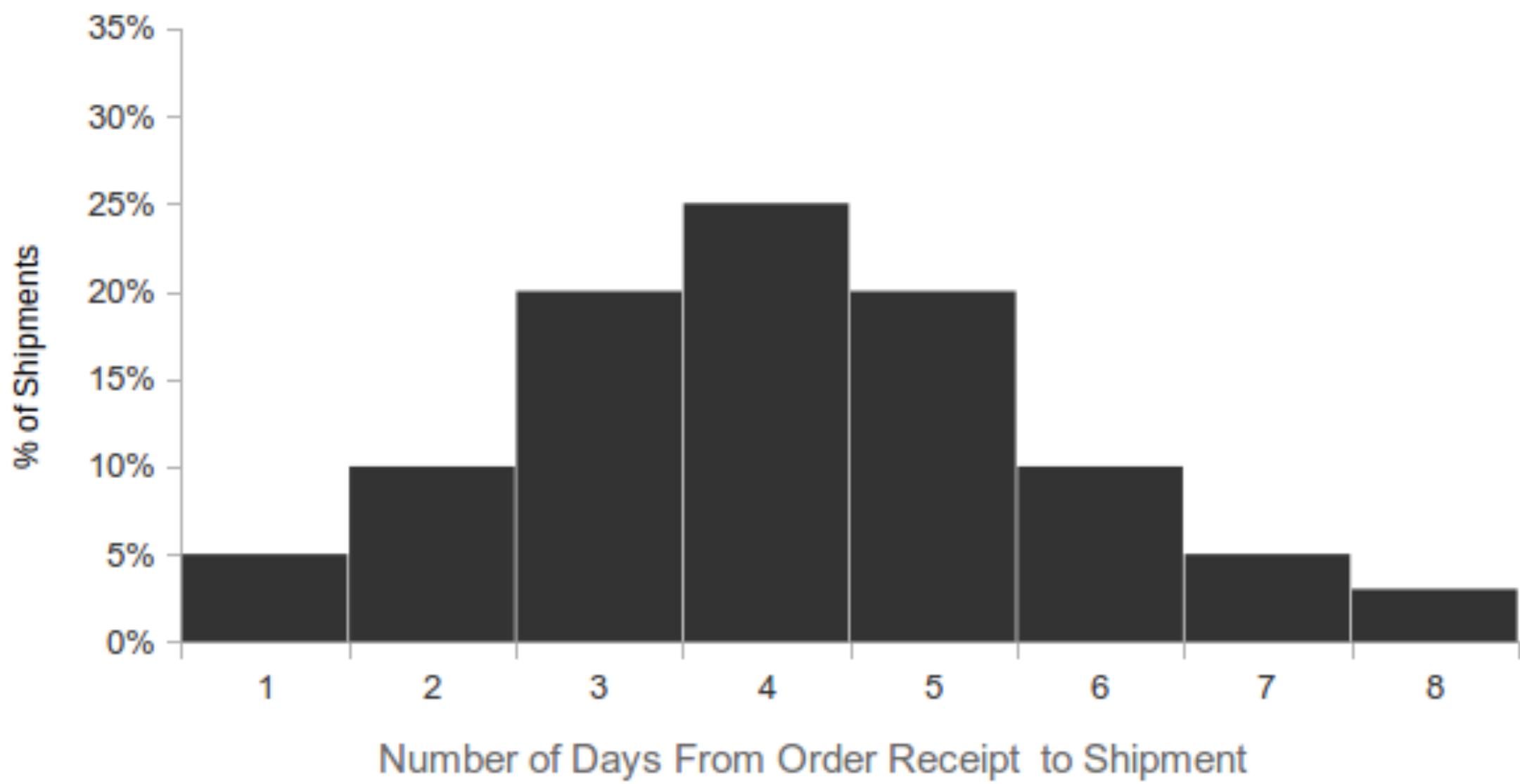
Average Days to Ship Orders



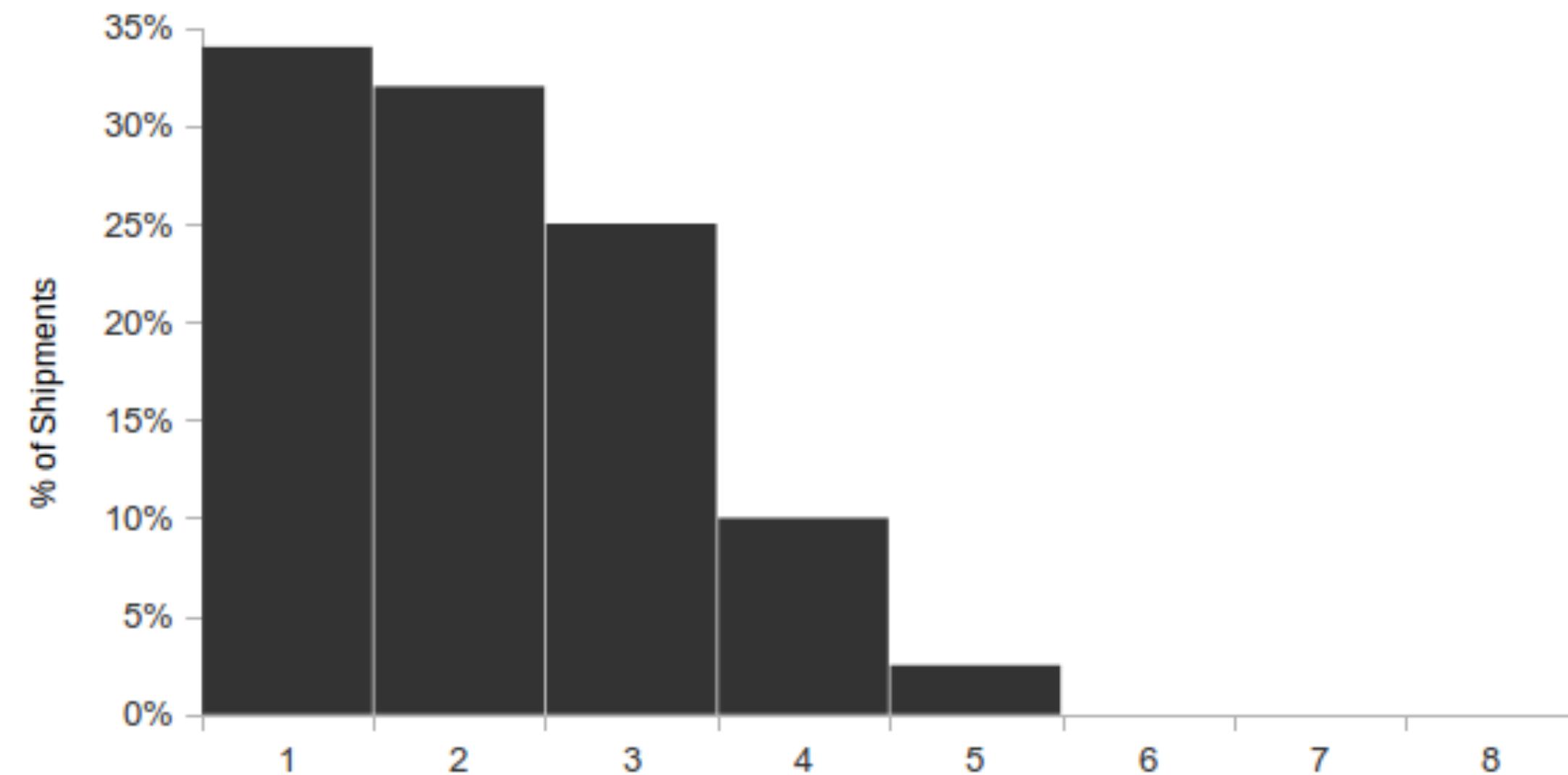
1 4 8

1 1 1 2 2 2 3 3 3 3 4 4 5 5 5 6 7 7 8 8

Seattle



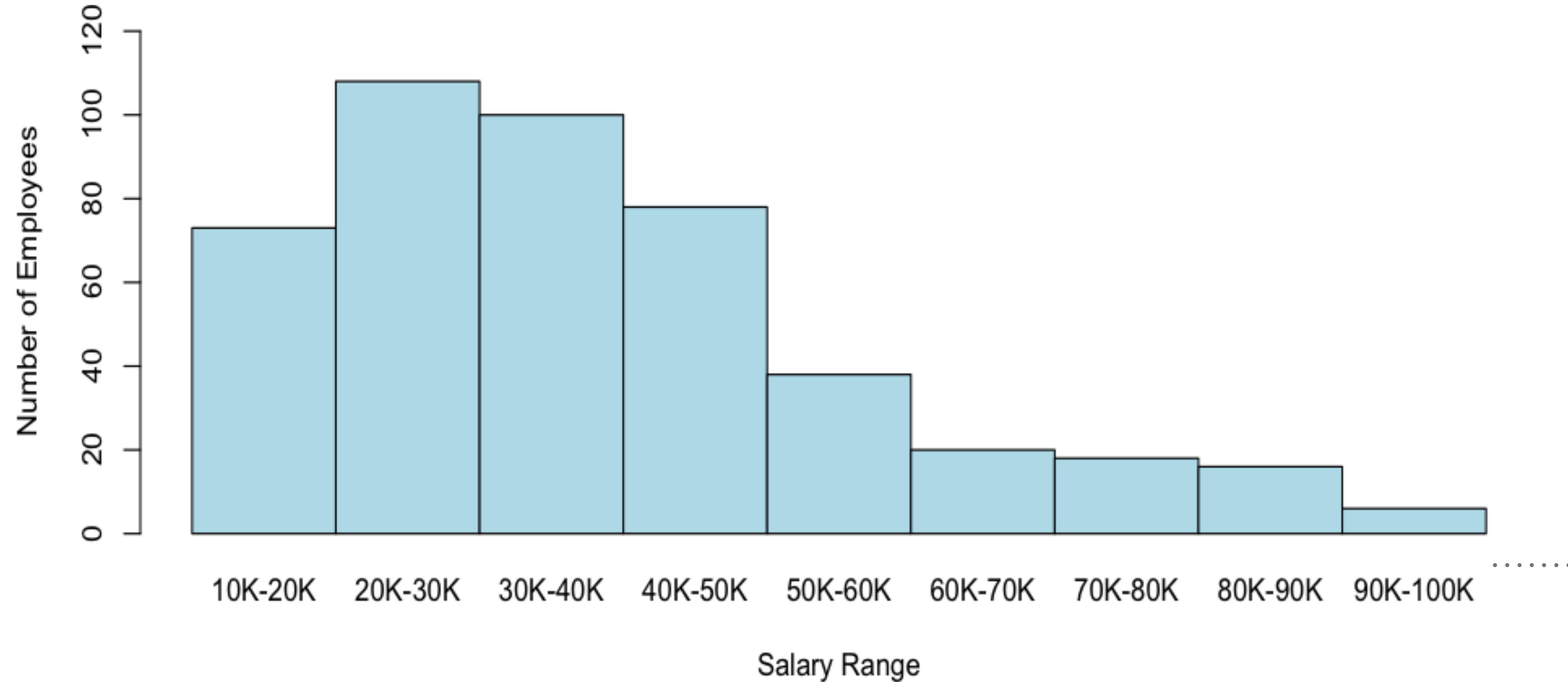
Atlanta



CARACTERÍSTICAS VISUAIS DE DISTRIBUIÇÕES

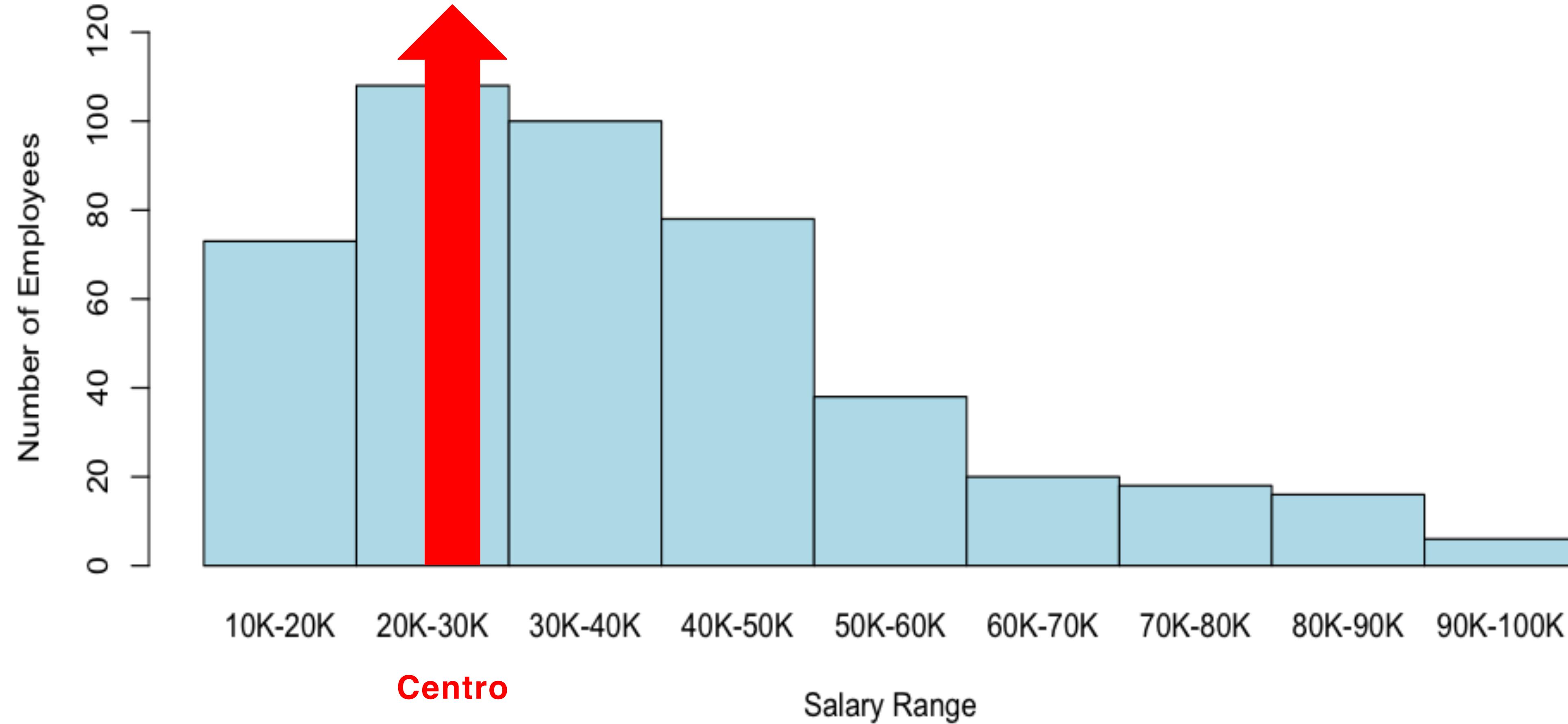
CARACTERÍSTICAS VISUAIS DE DISTRIBUIÇÕES

- Espalhamento
- Centro
- Forma

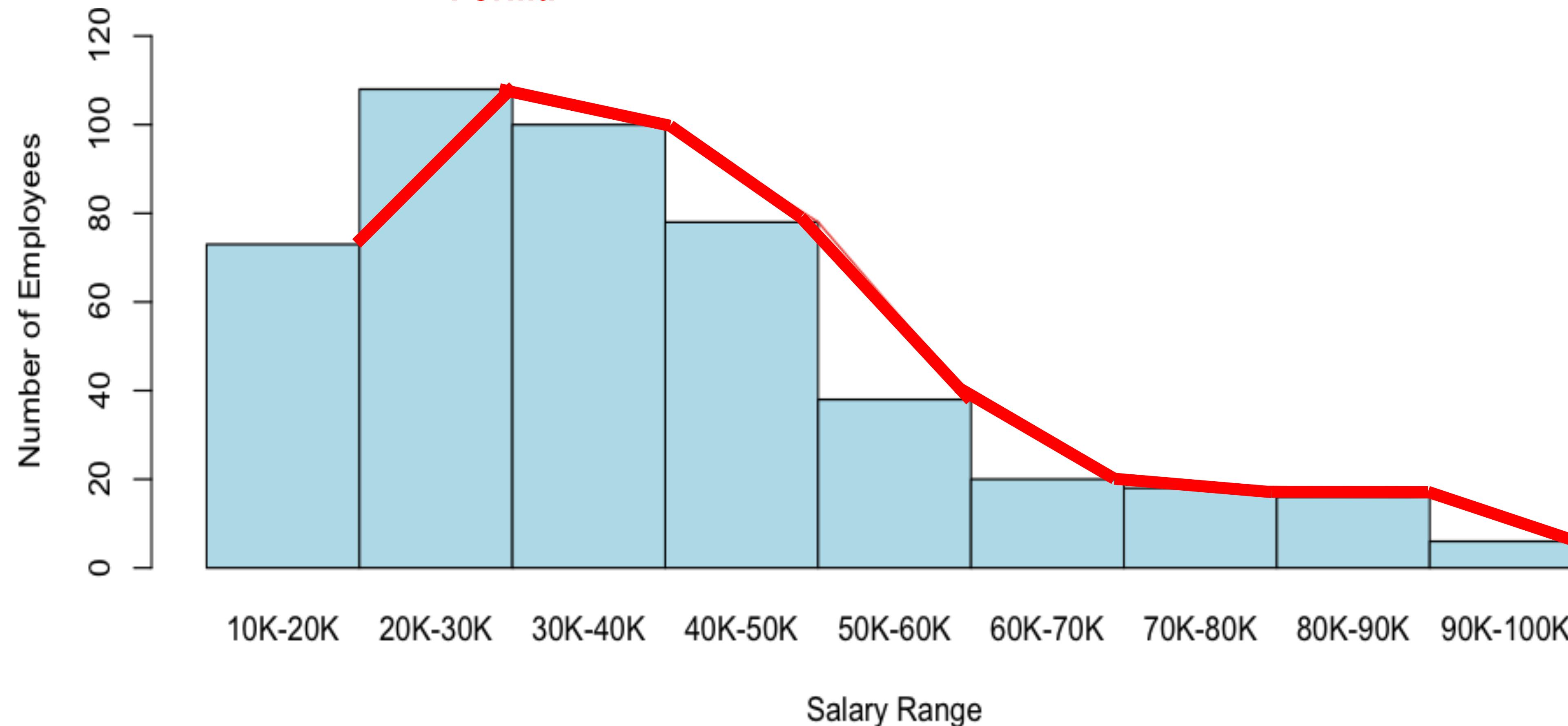


Espalhamento





Forma

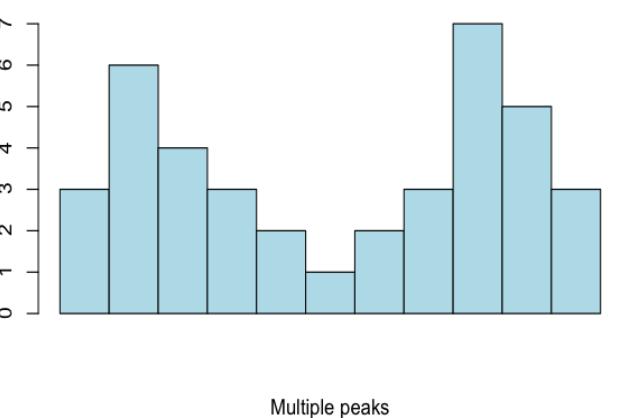
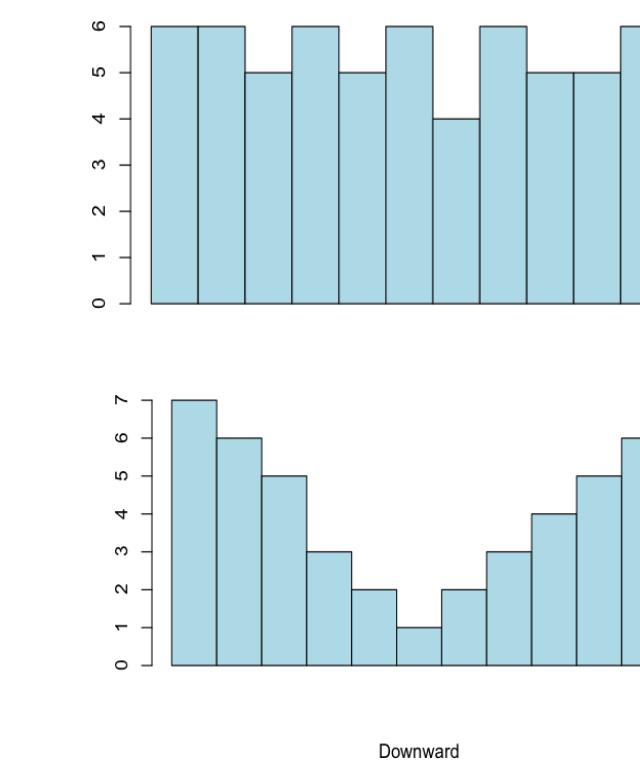
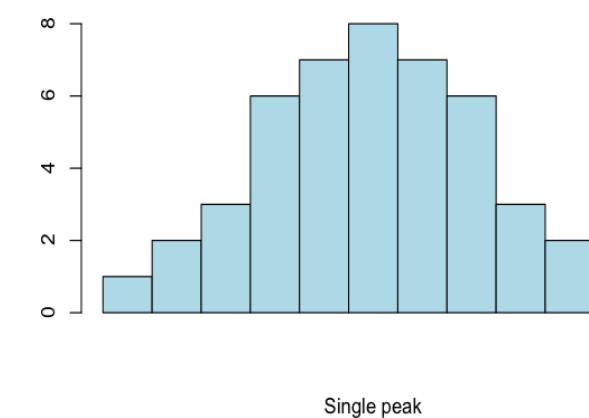
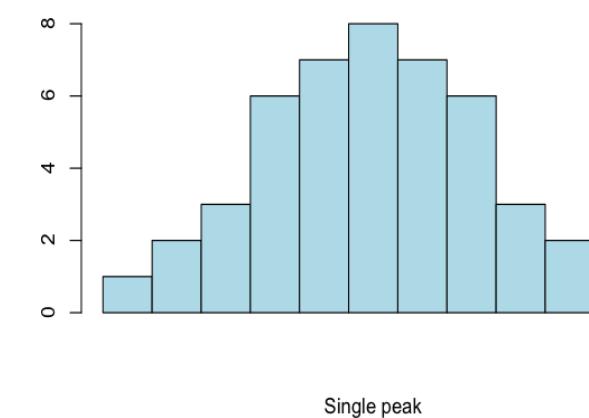
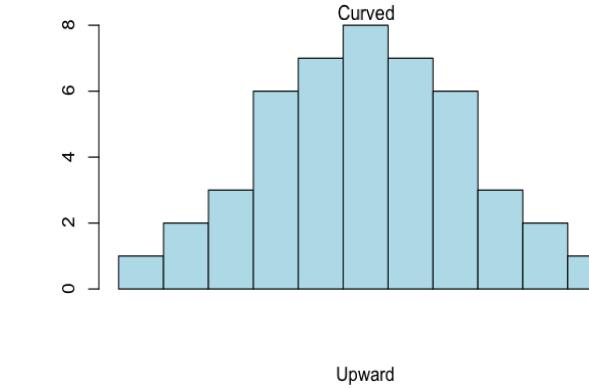
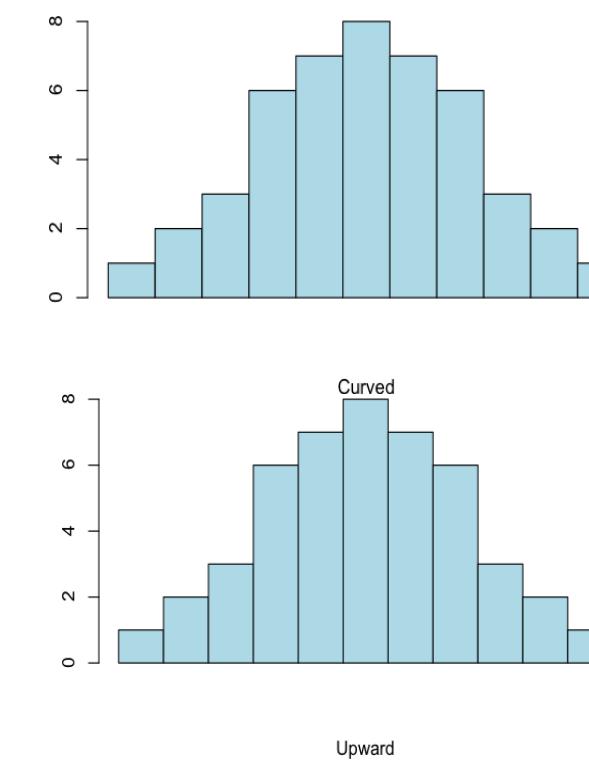
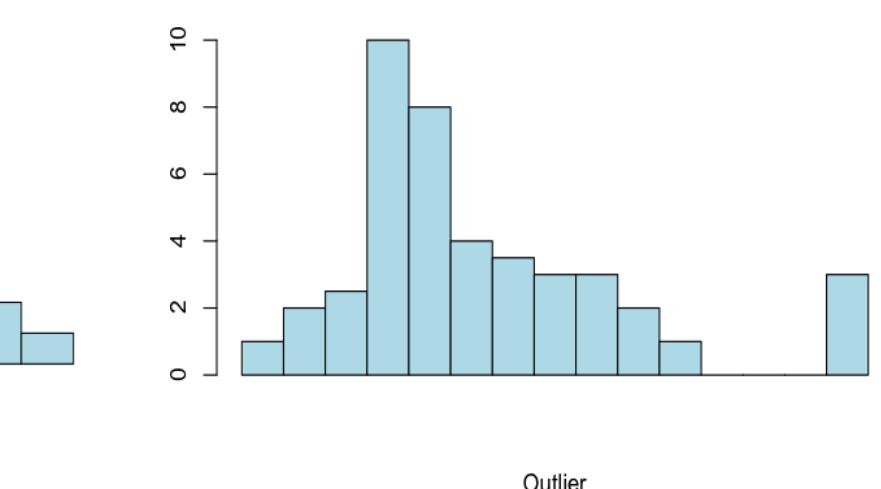
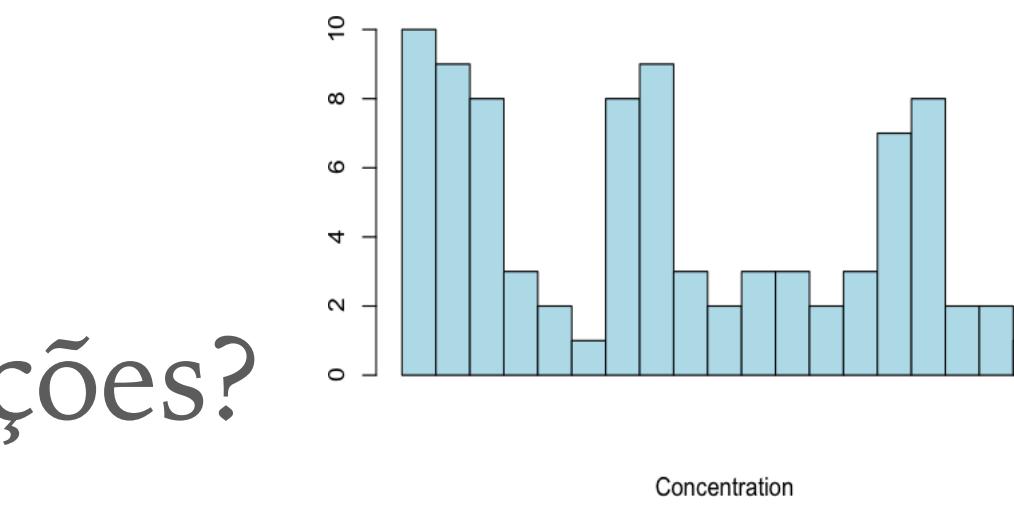


PADRÕES ANALÍTICOS

PADRÕES ANALÍTICOS

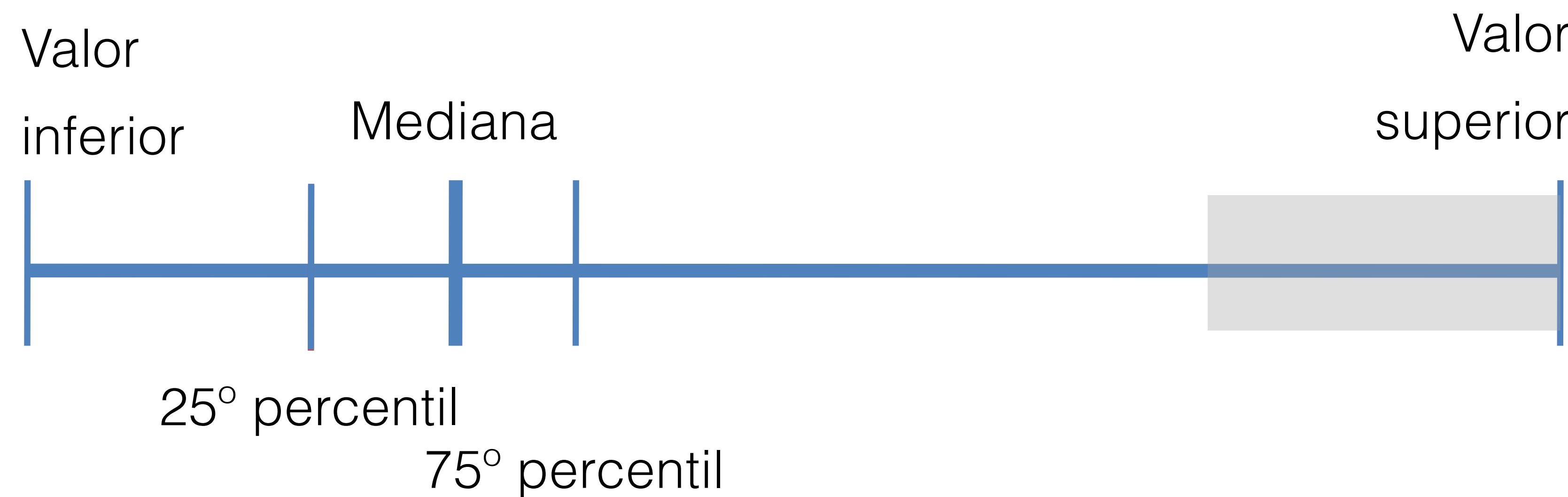
- Curva ou plana?
- Se curva, crescente ou decrescente?
- Se curva crescente, um ou vários picos?

- Se apenas um pico, simétrica ou enviesada?
- Concentrações?
- Lacunas?



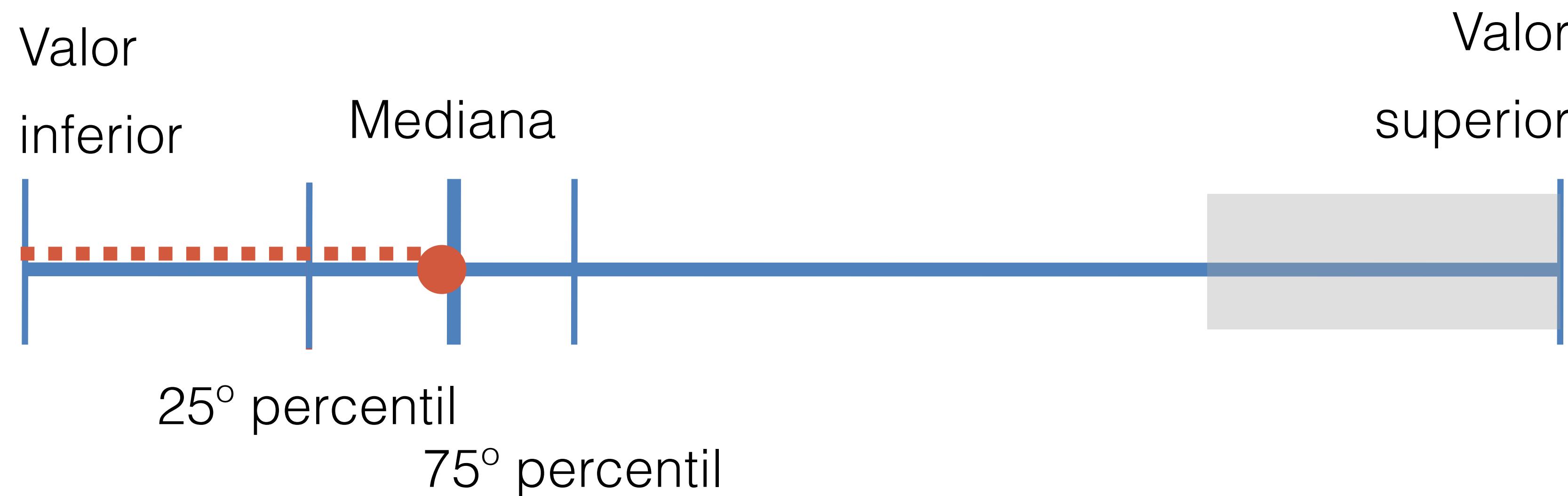
Regra simples para identificação de exceções

ponto médio entre 25º e 75º percentil multiplicado por 1,5 e subtraído e somado dos mesmos



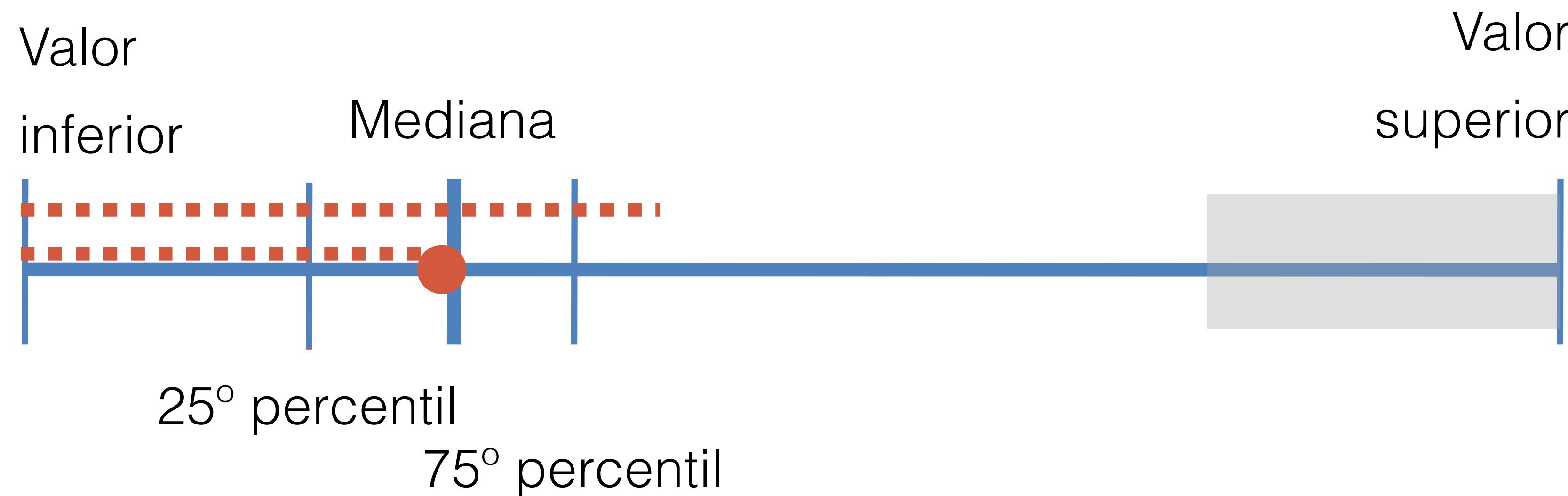
Regra simples para identificação de exceções

ponto médio entre 25º e 75º percentil multiplicado por 1,5 e subtraído e somado dos mesmos



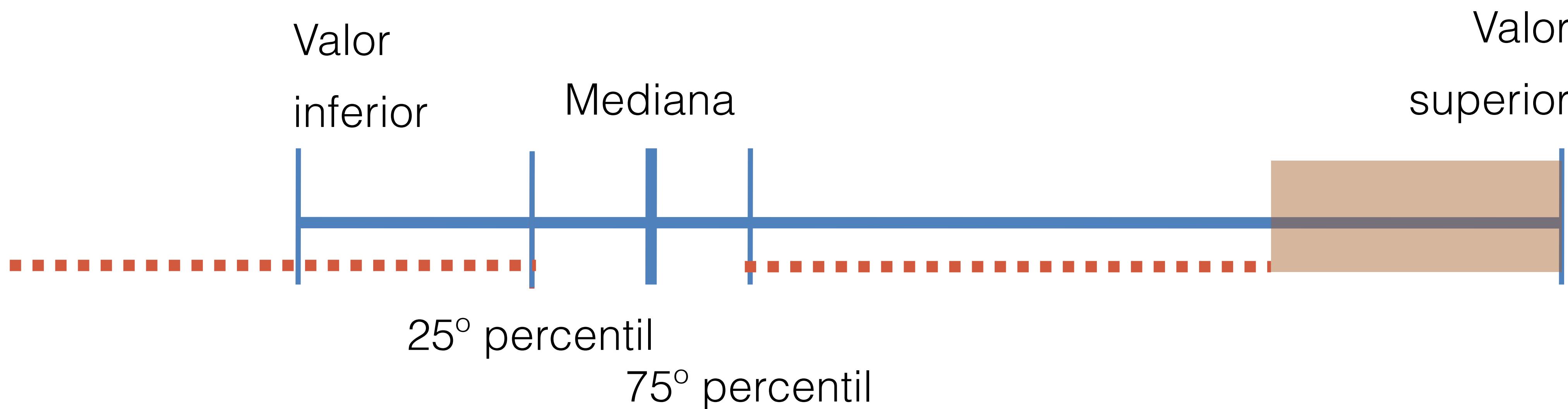
Regra simples para identificação de exceções

ponto médio entre 25º e 75º percentil multiplicado por 1,5 e subtraído e somado dos mesmos



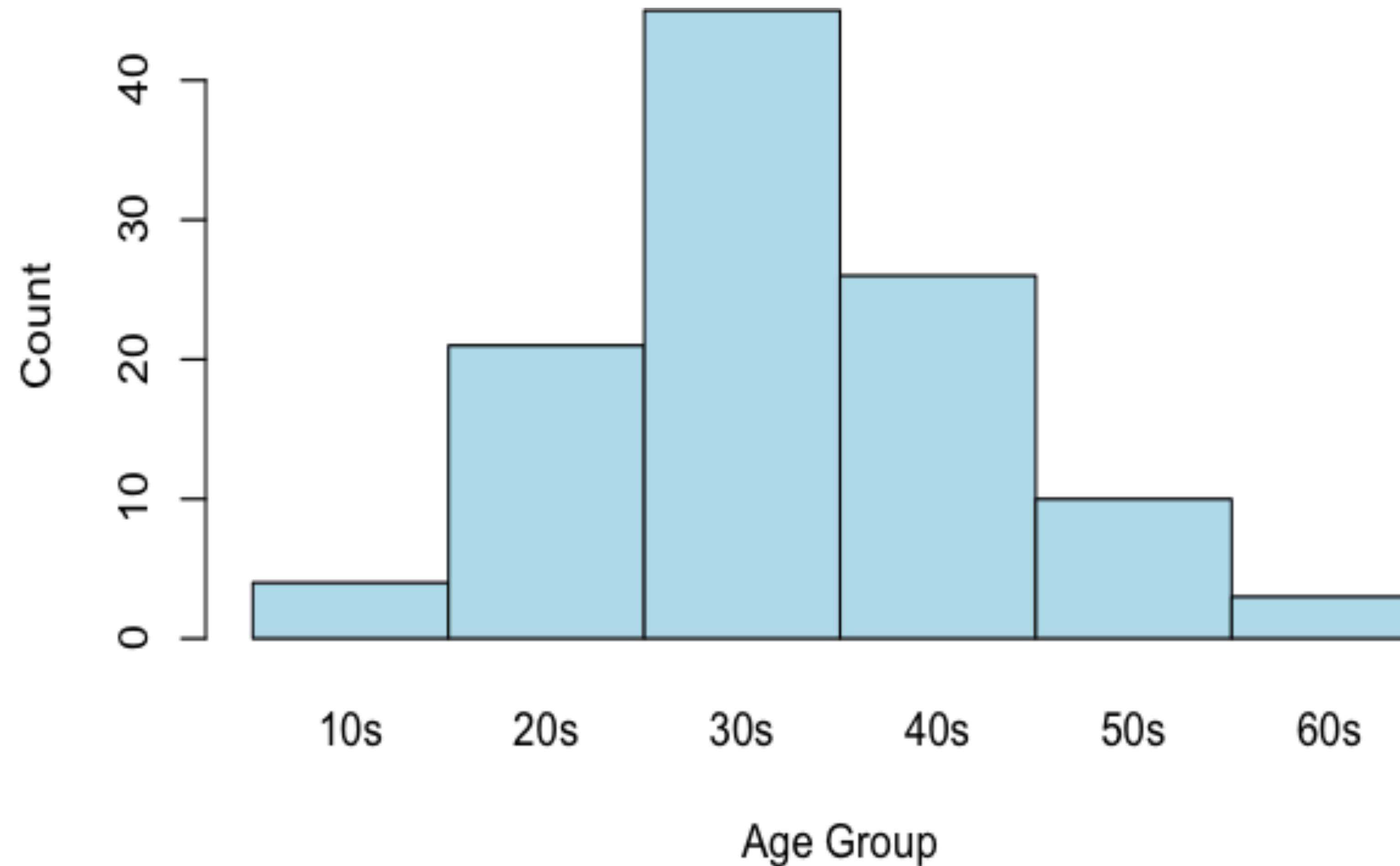
Regra simples para identificação de exceções

ponto médio entre 25º e 75º percentil multiplicado por 1,5 e subtraído e somado dos mesmos



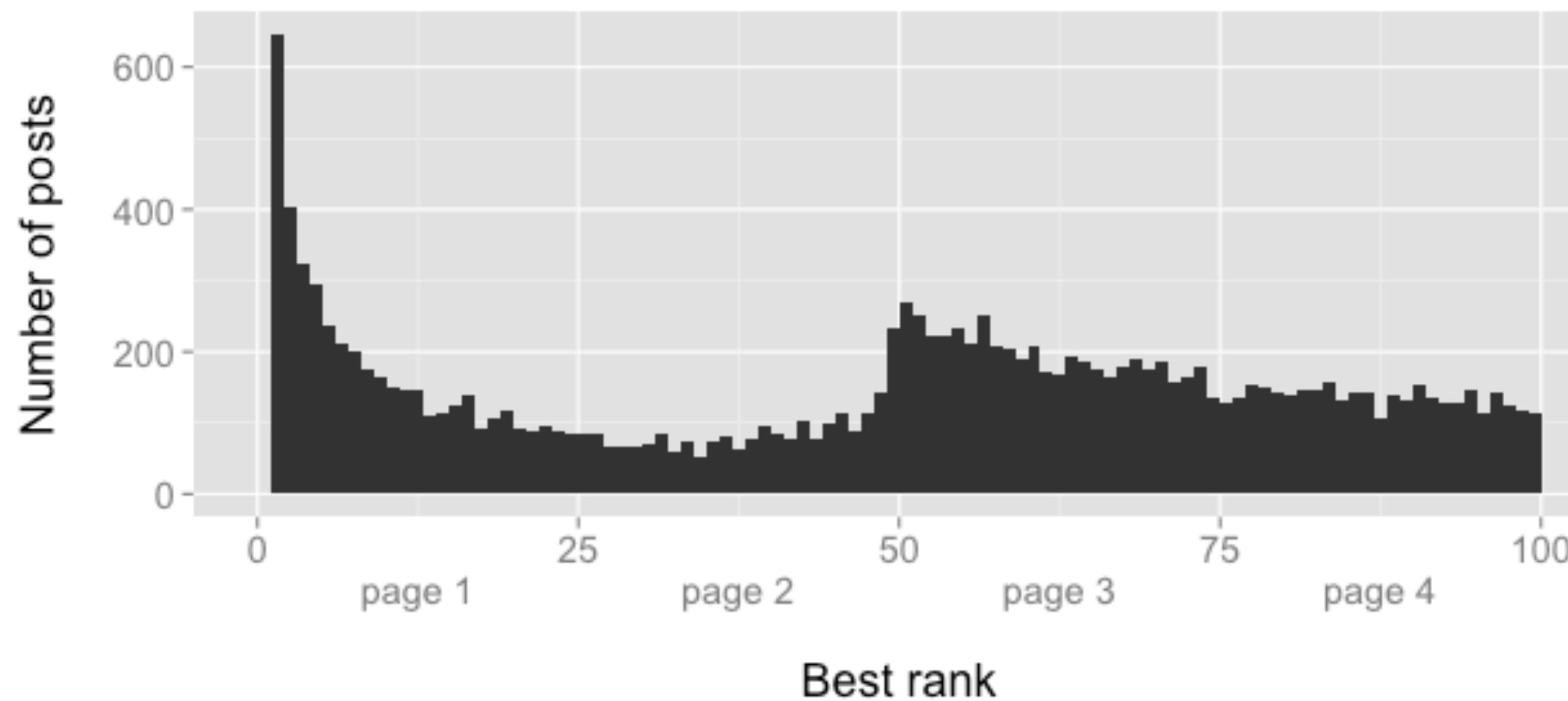
REPRESENTAÇÕES VISUAIS

HISTOGRAMA



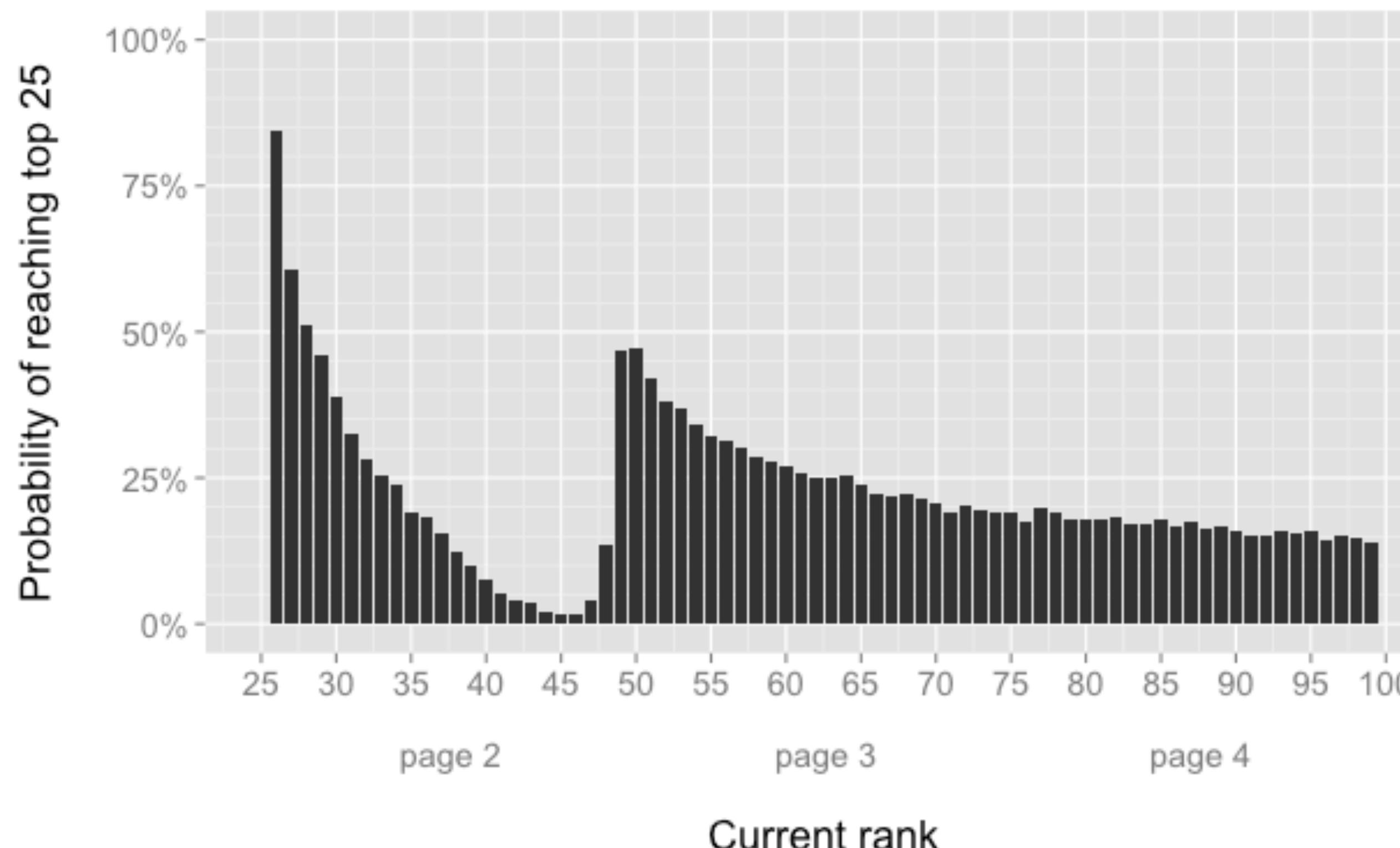
HISTOGRAMA

Distribution of best rank for 14979 posts observed in reddit top 100 between Sep 18 and Oct 31, 2014

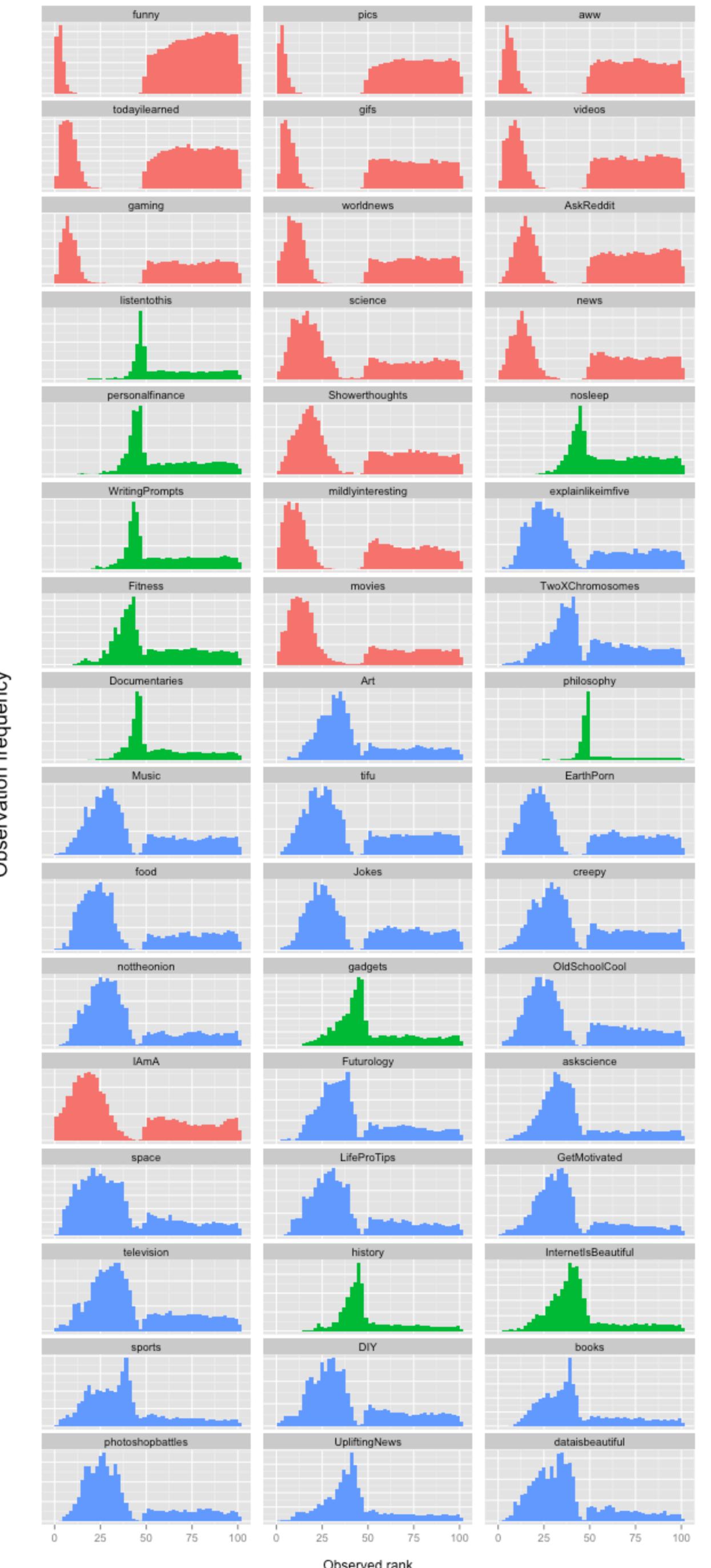


HISTOGRAMA

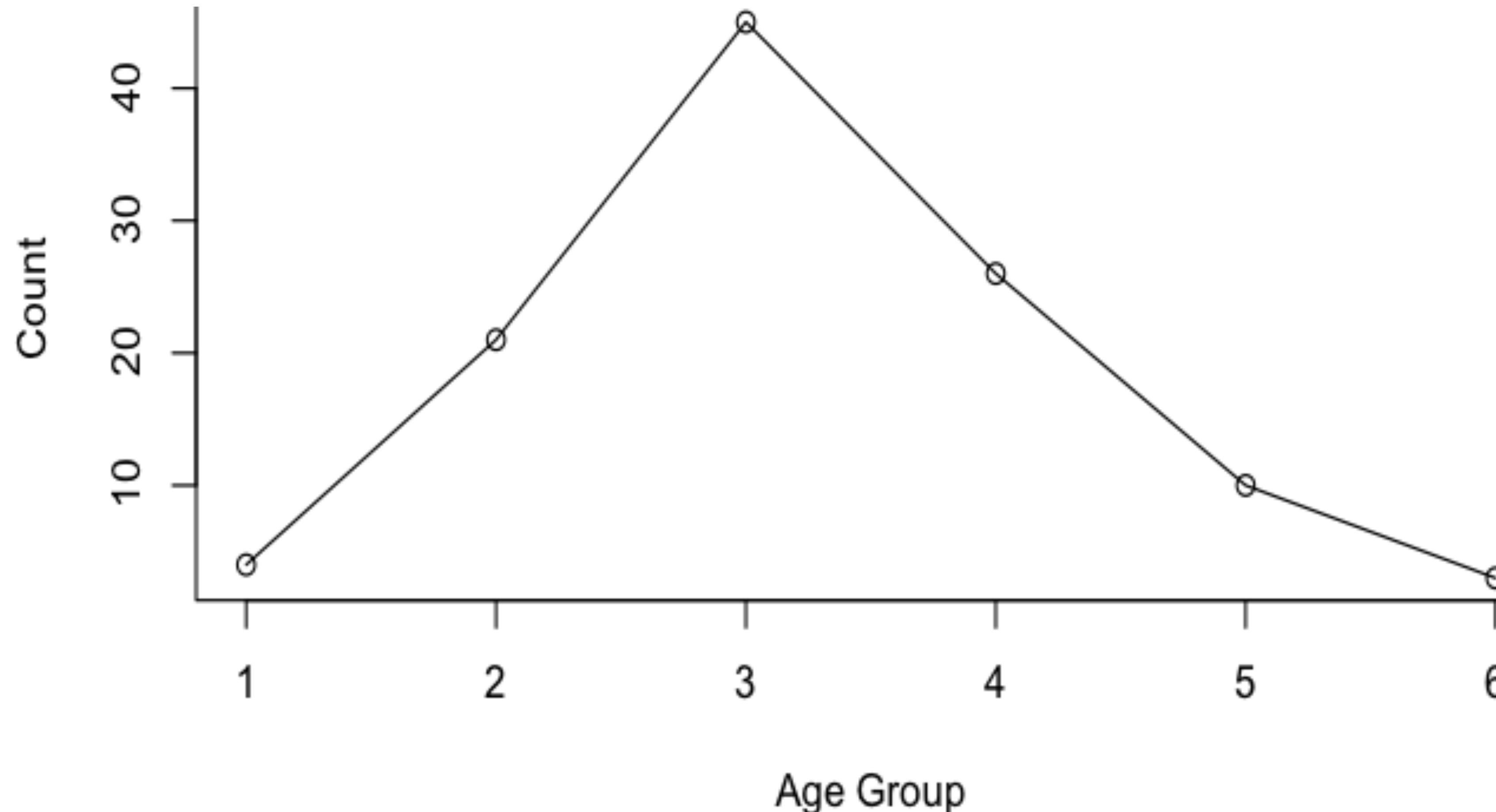
Conditional probability of a reddit post reaching the logged-out homepage, given that it's in the top 100 and its rank is not declining



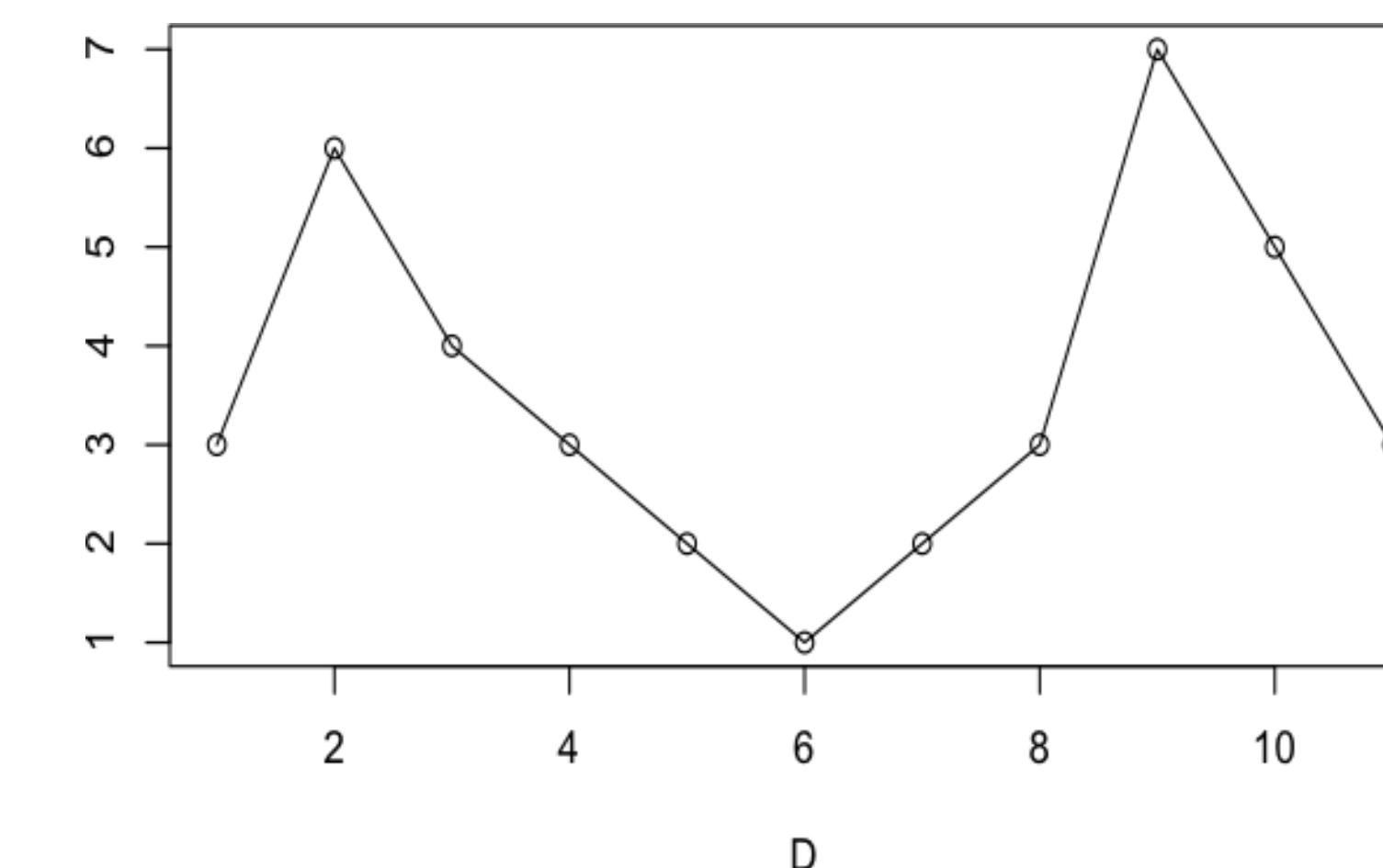
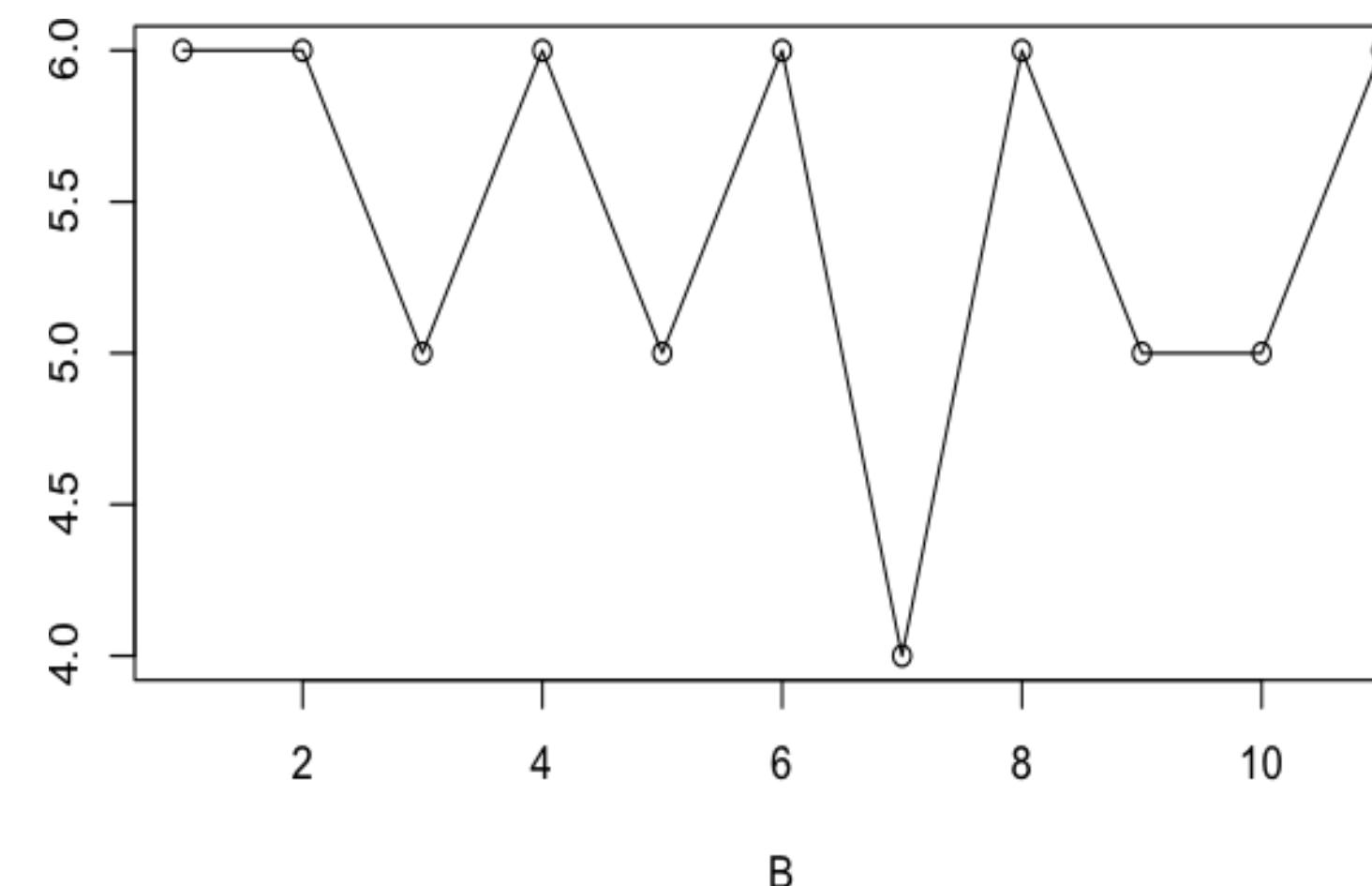
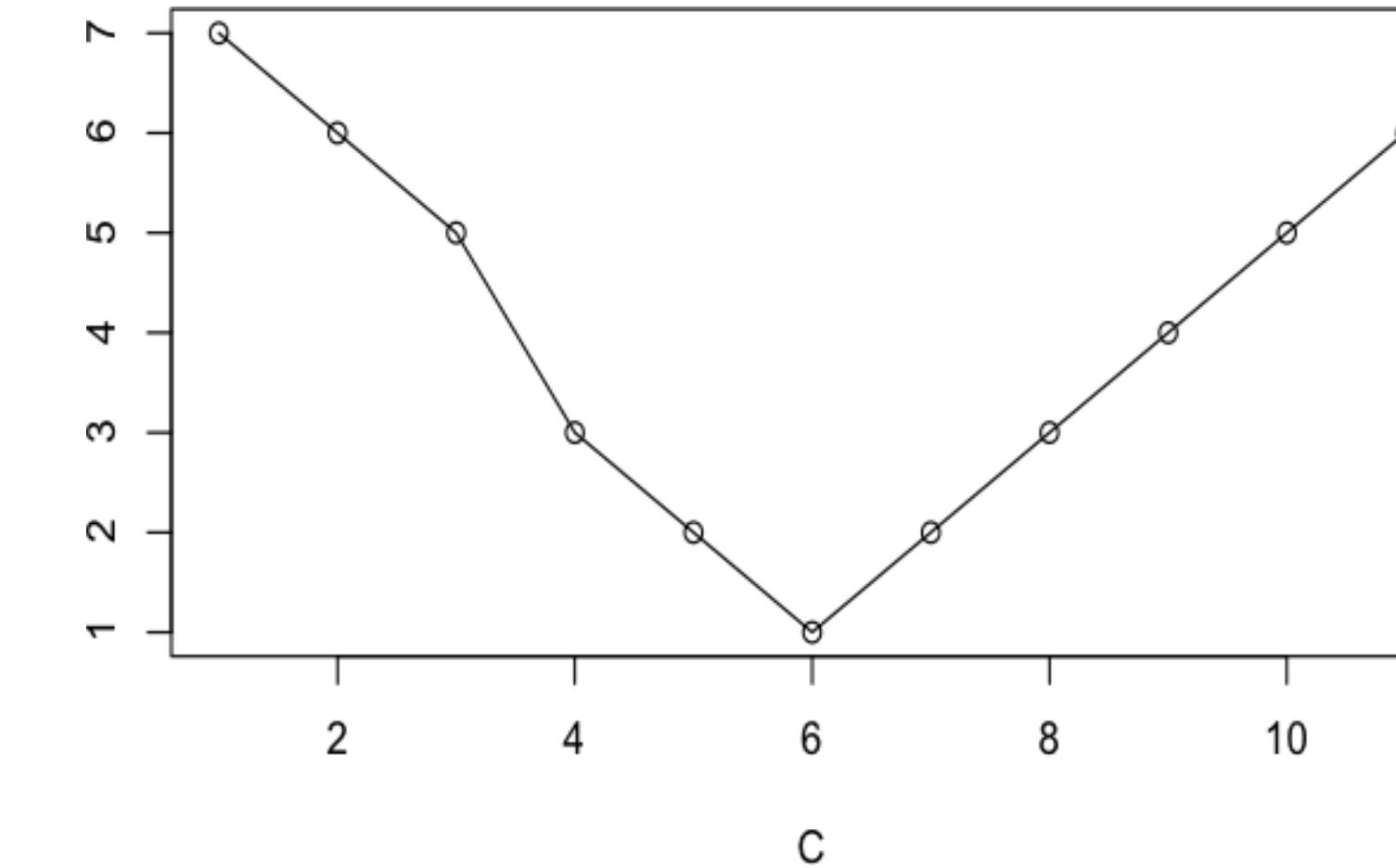
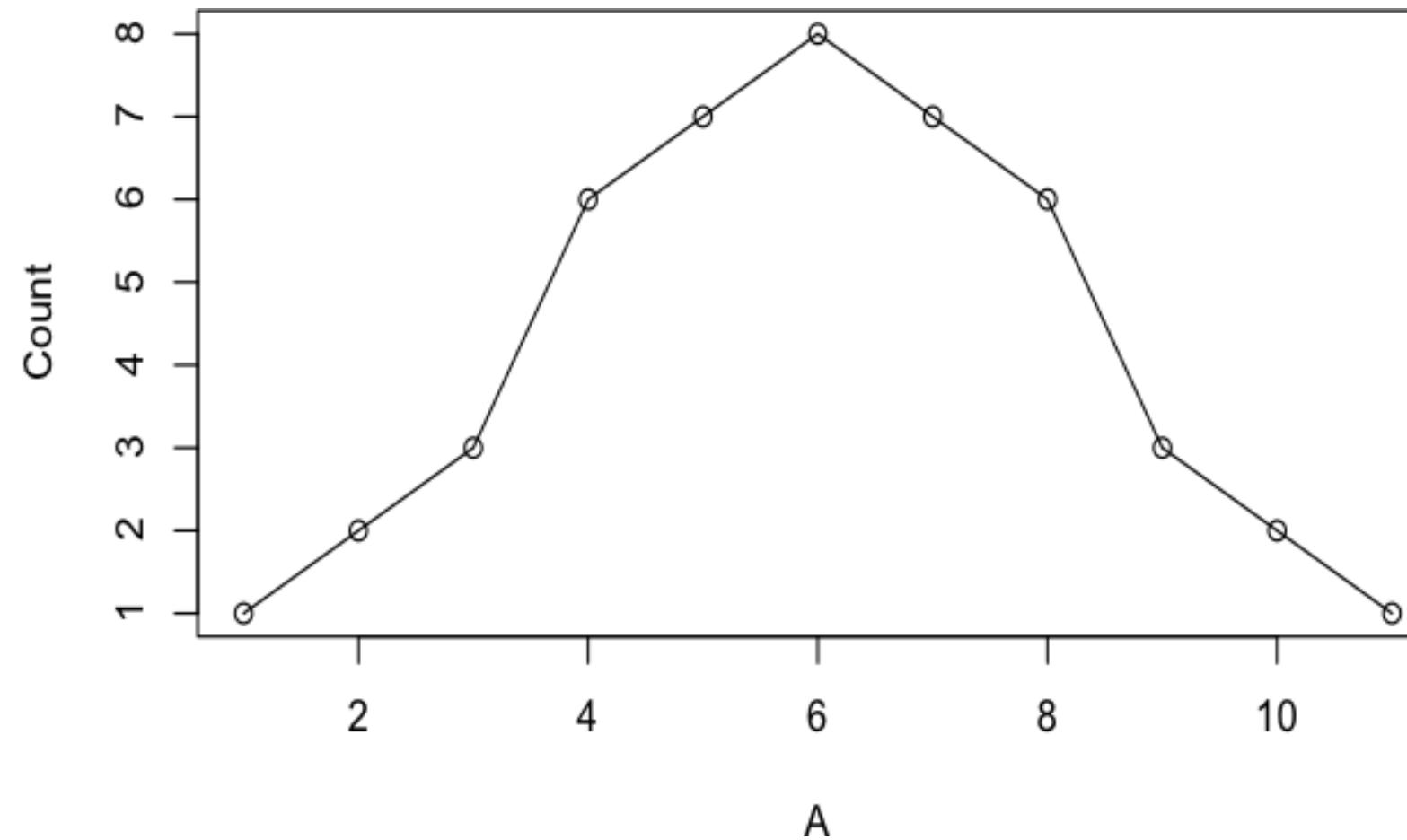
Observed ranks by subreddit



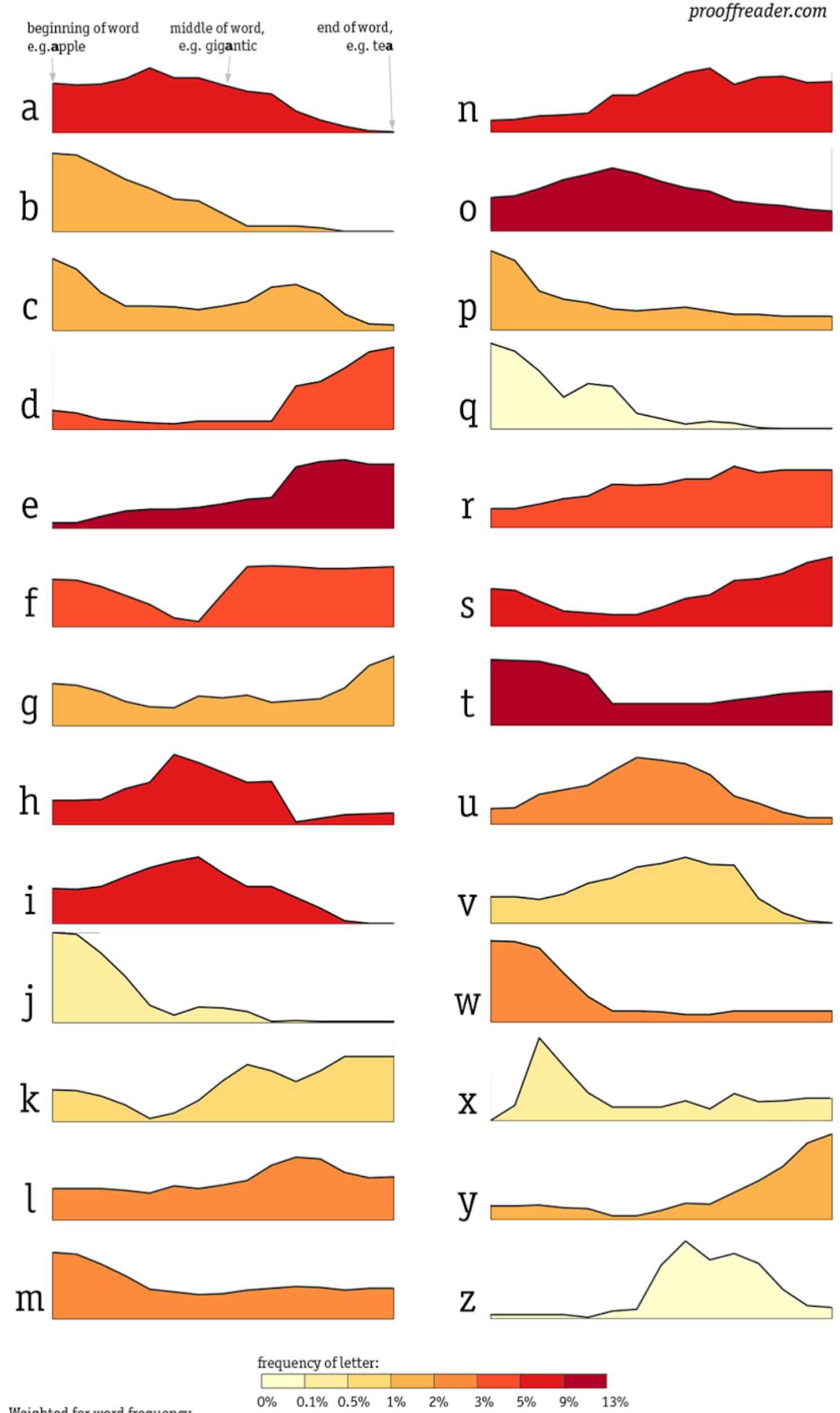
POLÍGONO DE FREQUÊNCIA



MÚLTIPLOS POLÍGONO DE FREQUÊNCIA



Distribution of English letters toward beginning, middle and end of words



STEM-AND-LEAF (RAMO E FOLHAS)

1. 8 9 9 9

2. 0 0 1 1 3 3 3 4 4 4 5 6 6 6 6 6 6 7 7 8 8 9 9 9 9 9

3. 0 0 0 0 1 1 2 2 2 2 2 2 2 3 3 3 3 3 3 4 4 4 5 5 5 6 7 8 9

4. 0 0 0 0 1 1 2 2 3 3 4 4 5 5 5 6 6 6 7 8 9

5. 0 0 1 3 5 7 9

6. 1 4 8

Usar no máximo 2 valores no stem e 3 valores no leaf

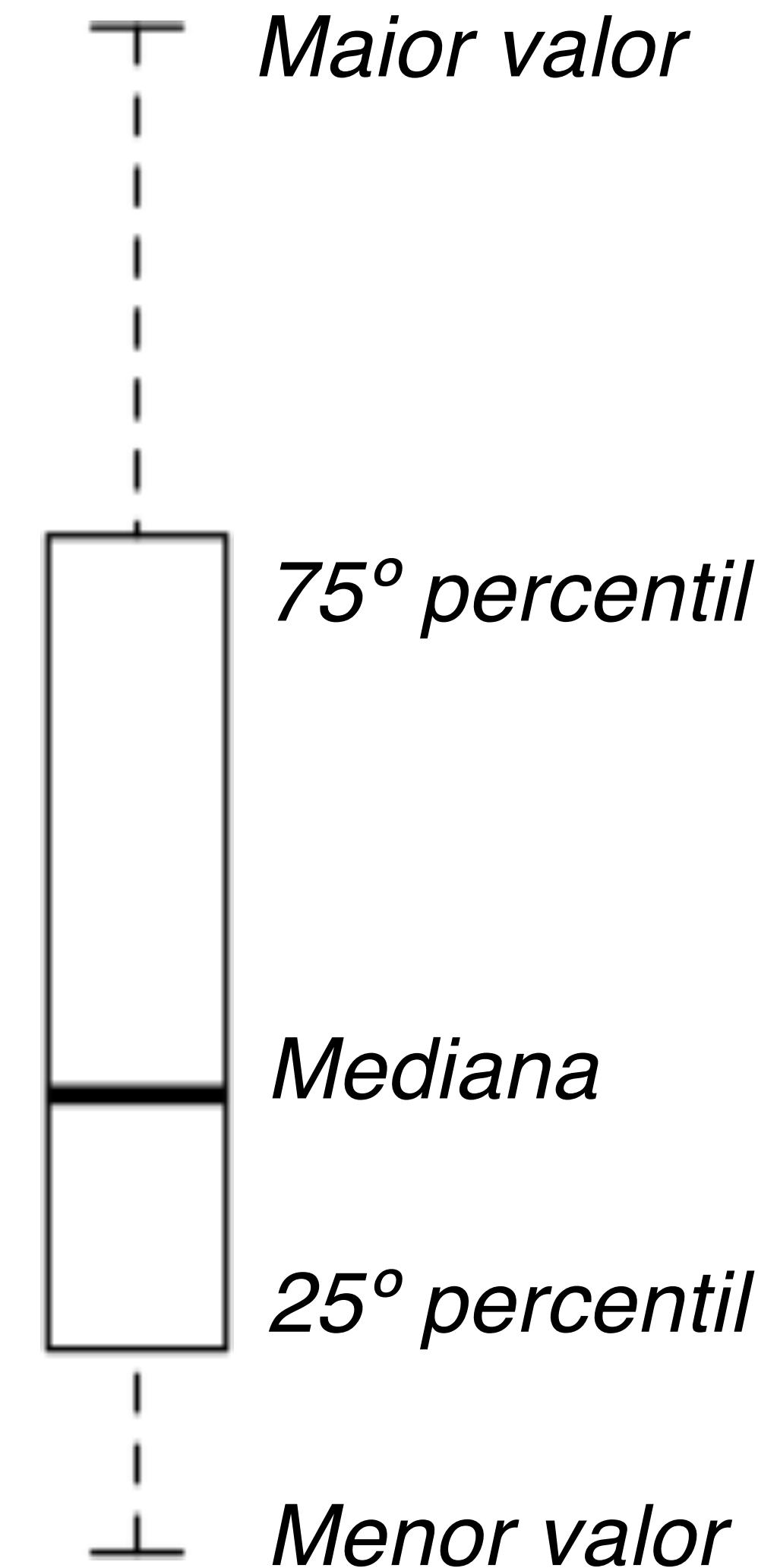
STEM-AND-LEAF (RAMO E FOLHAS)



BOX PLOT

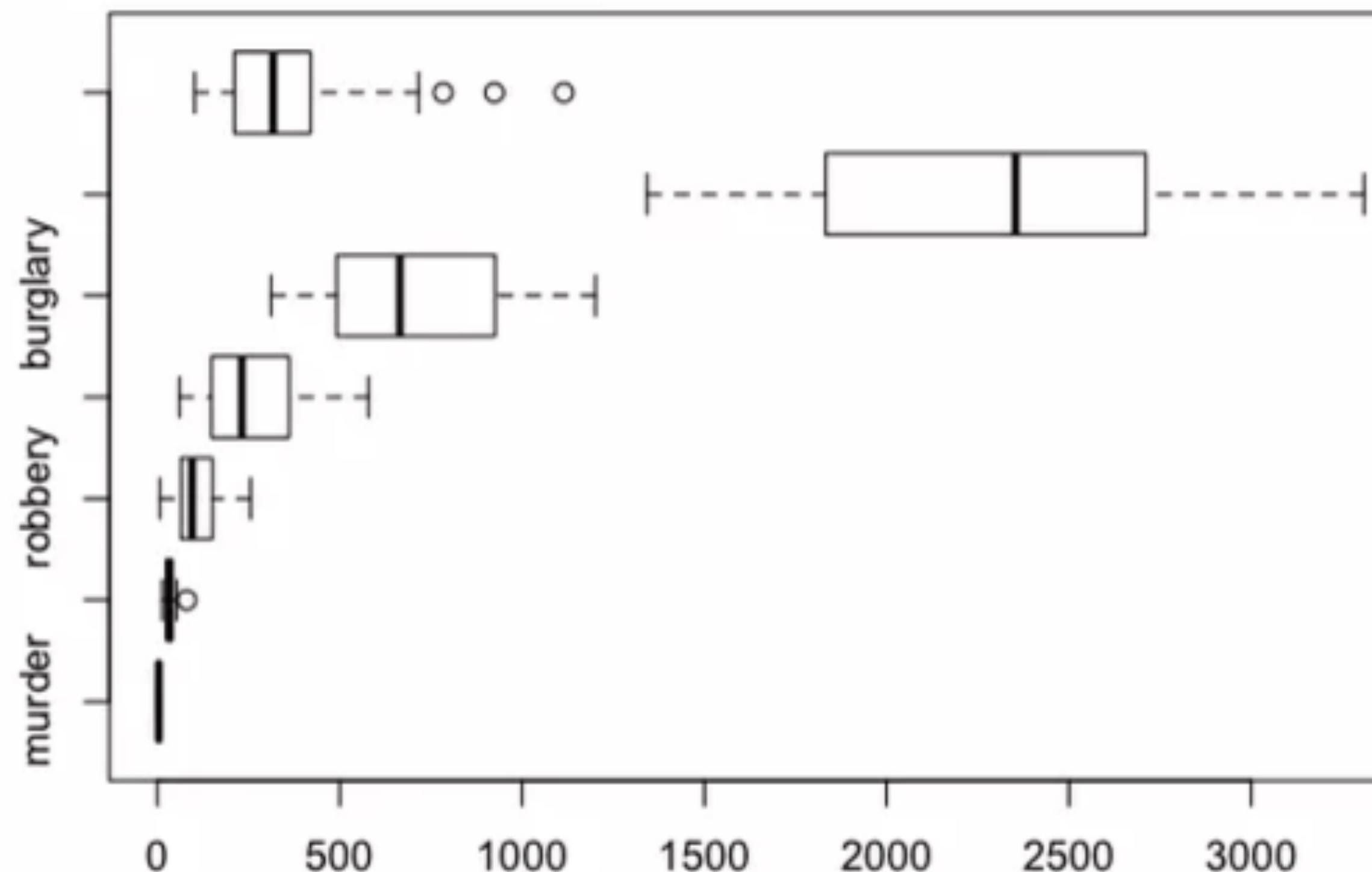
*Espalhamento
de 100%
dos valores*

*Espalhamento
de 50%
dos valores*

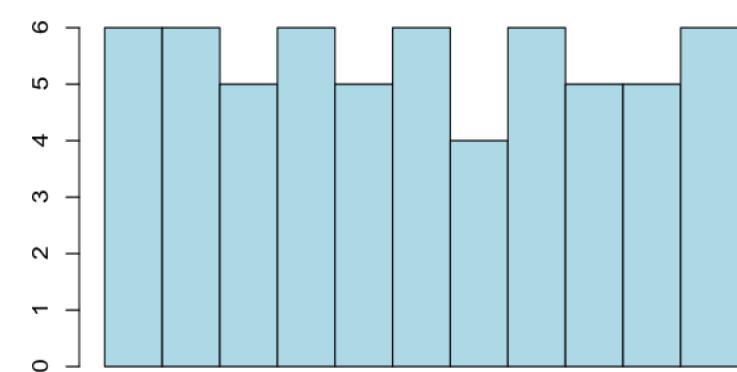
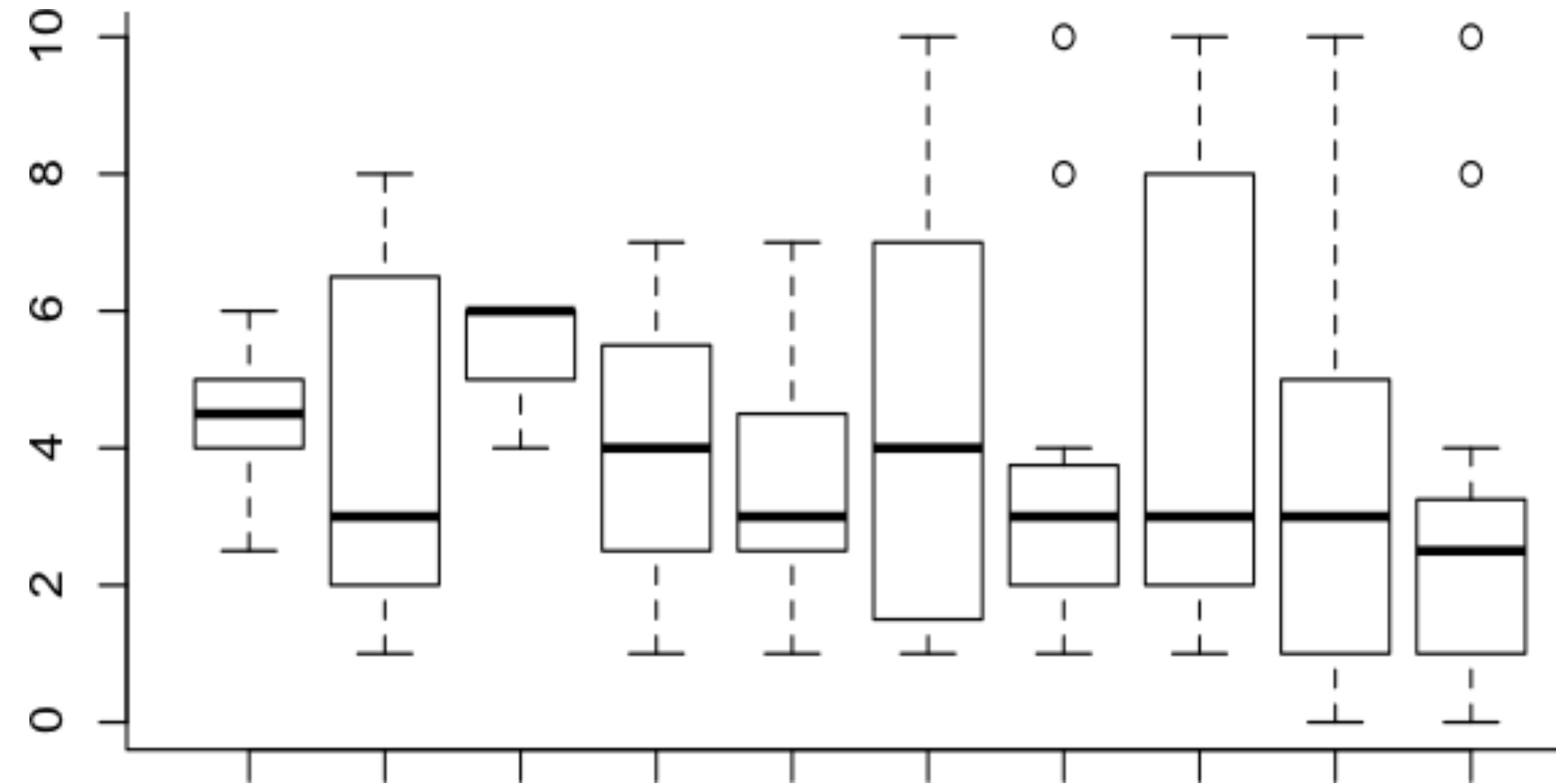


BOX PLOT

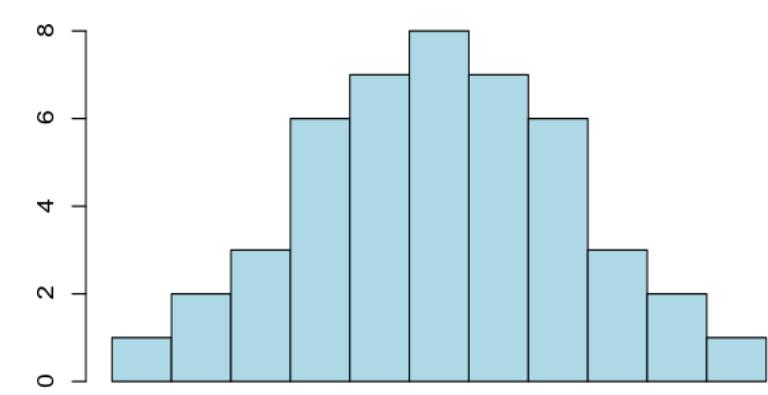
Crime Rates in US



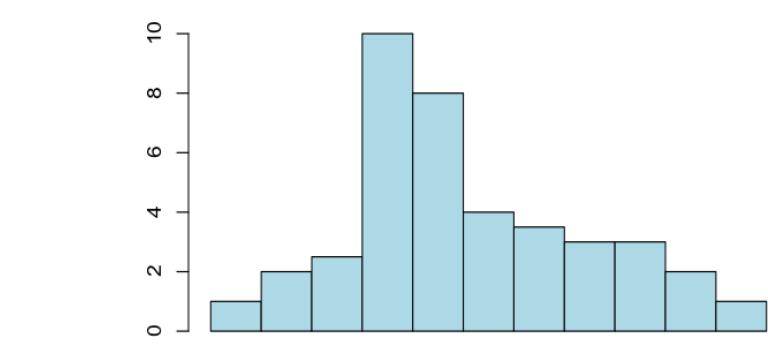
BOX PLOT



3°



4°



7°

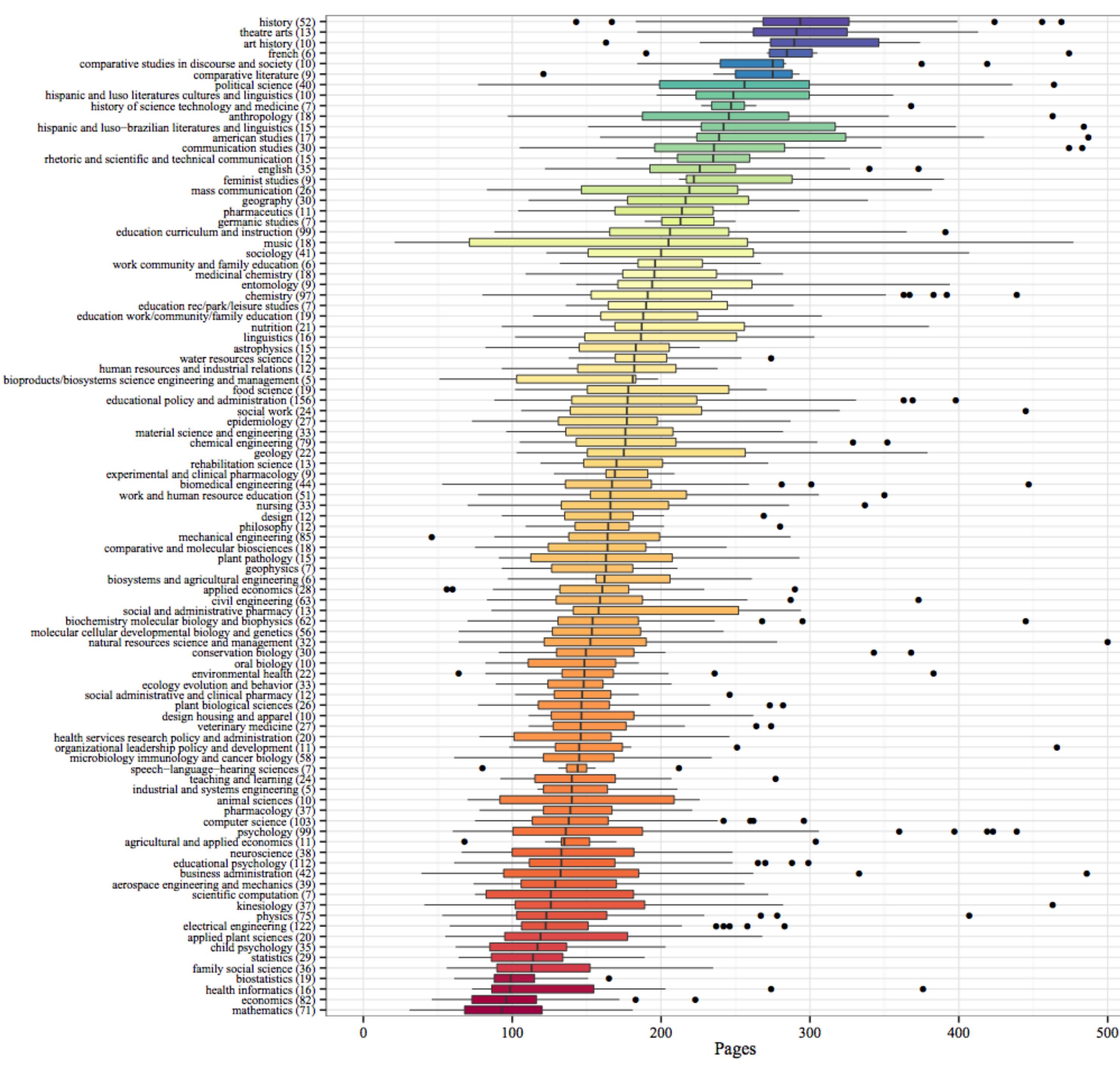


Fig: Summary of page lengths of doctoral dissertations by major, sorted and color-coded by median. Boxes represent the median, 25th and 75th percentiles, 1.5 times the interquartile range as whiskers, and outliers beyond the whiskers. Number of records for each major are in parentheses.

Type Thesis Dissertation

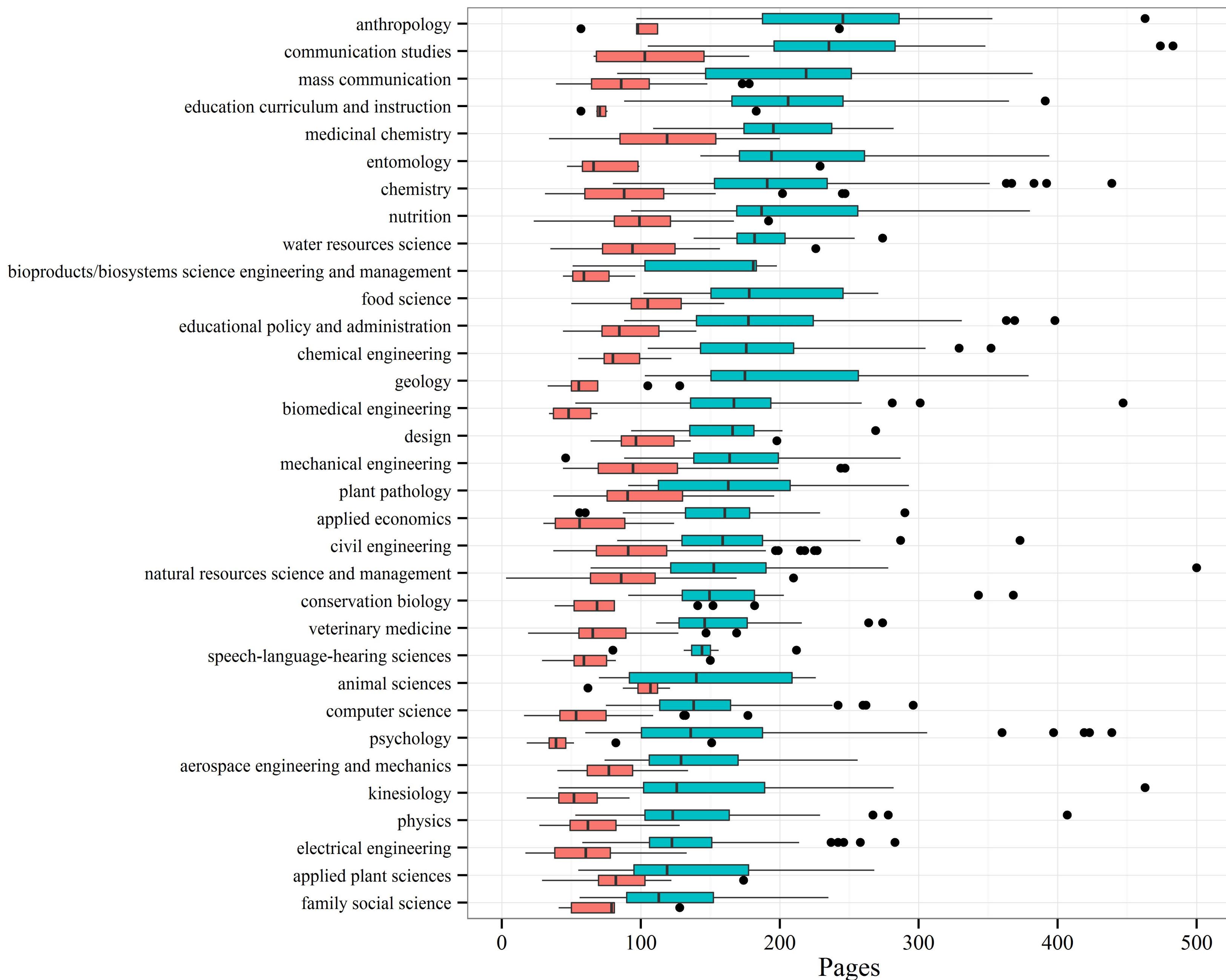
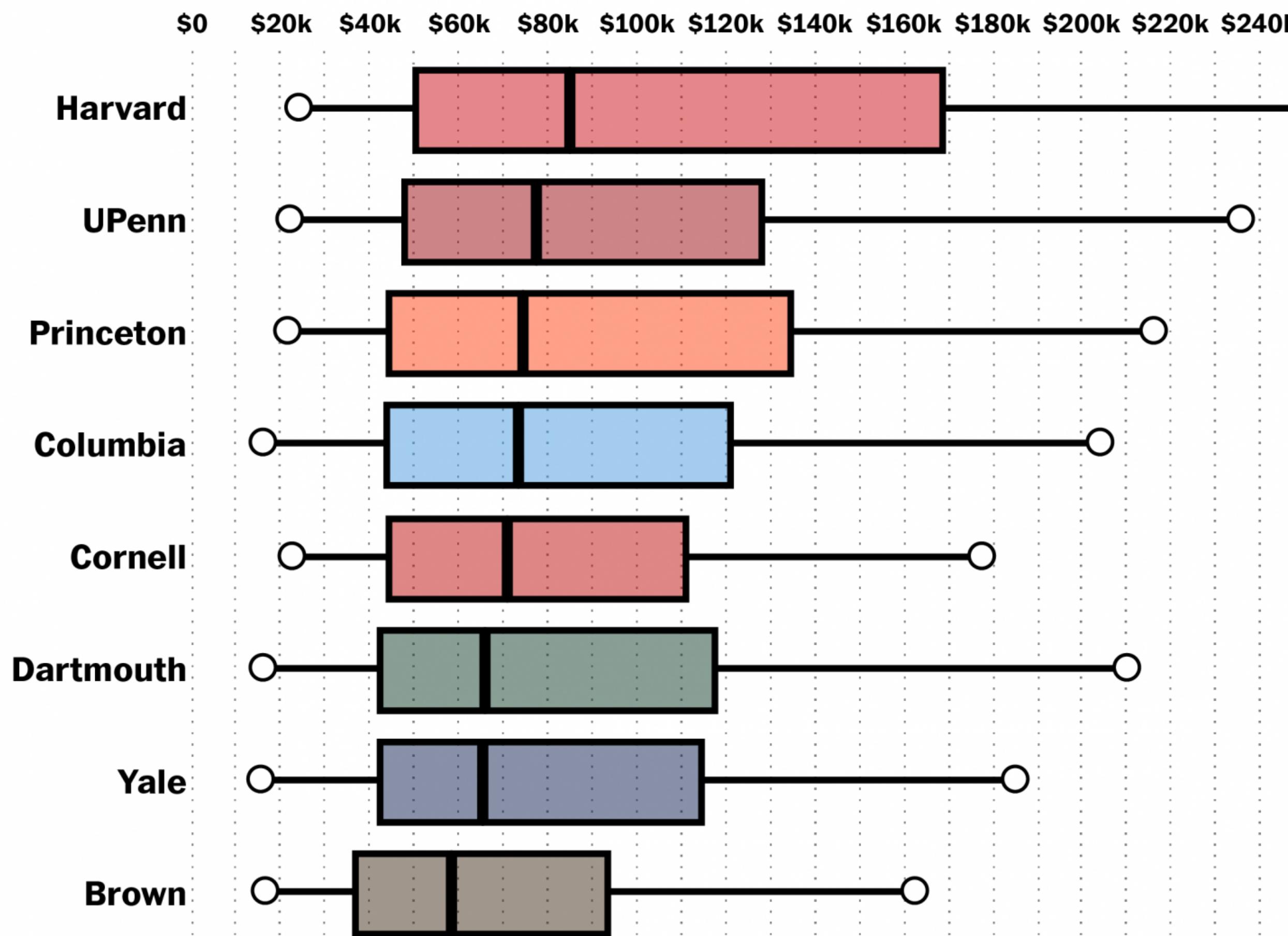


Fig: Summary of page lengths of doctoral dissertations by major, sorted and color-coded by median. Boxes represent the median, 25th and 75th percentiles, 1.5 times the interquartile range as whiskers, and outliers beyond the whiskers. Number of records for each major are in parentheses.

Ranking the Ivies

Annual earnings distributions, 10 years after starting school

How to read



Ivy League educations pay off

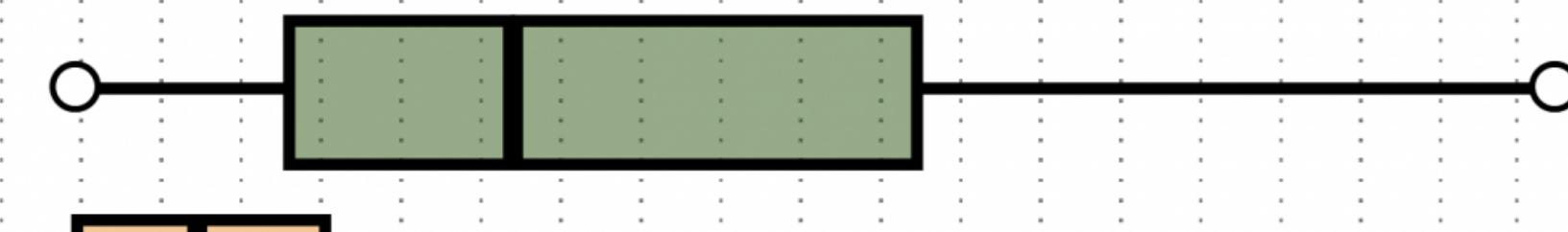
Annual earnings distributions, 10 years after starting school

How to read

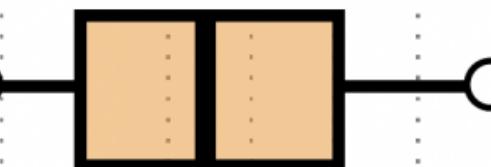


\$0 \$20k \$40k \$60k \$80k \$100k \$120k \$140k \$160k \$180k \$200k \$220k \$240k

Ivy League



All schools



WAPO.ST/WONKBLOG

Source: U.S. Dept. of Education College Scorecard

Red Sox Roller Coaster

Actual vs. projected number of wins, 1999-2014

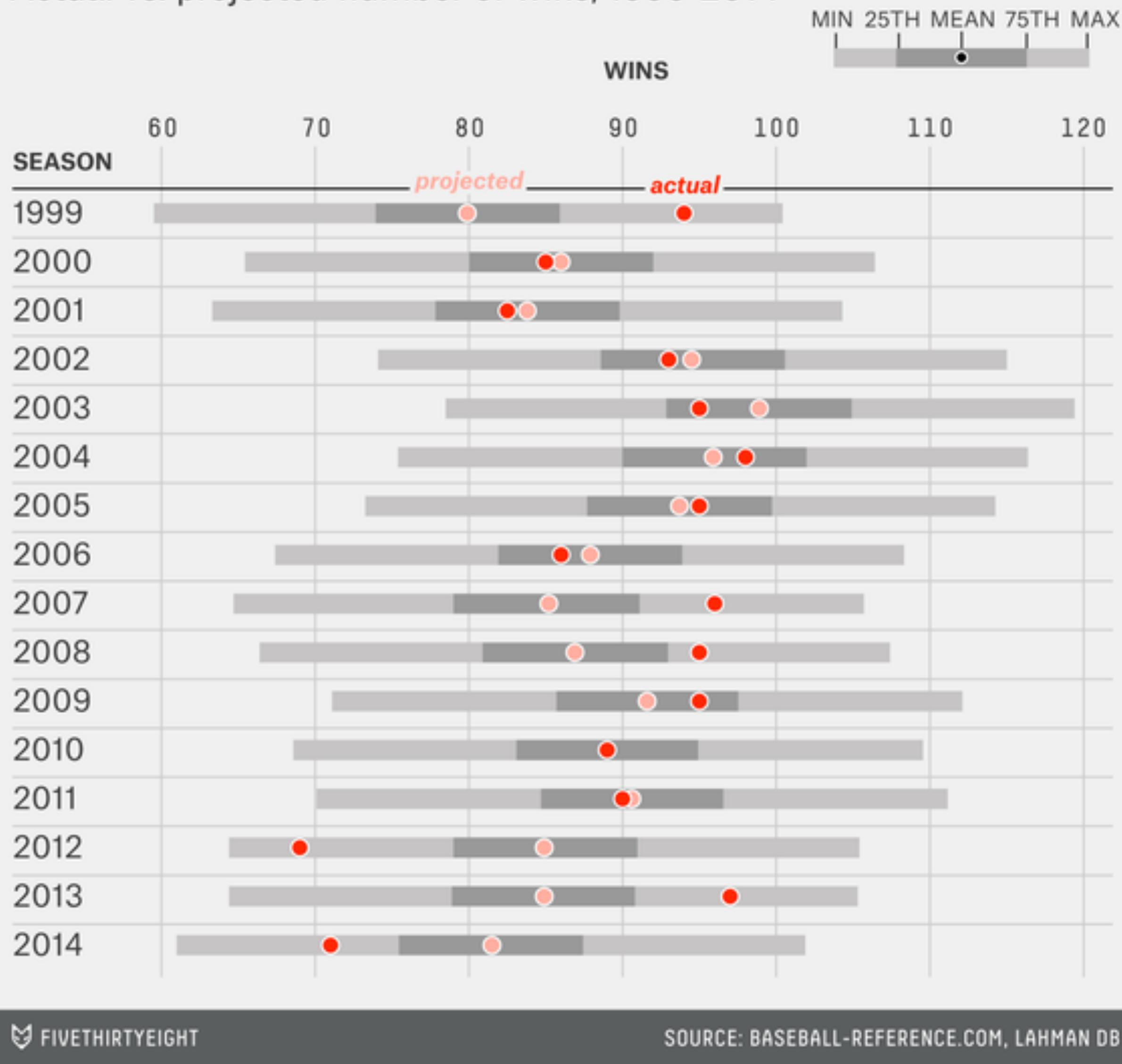
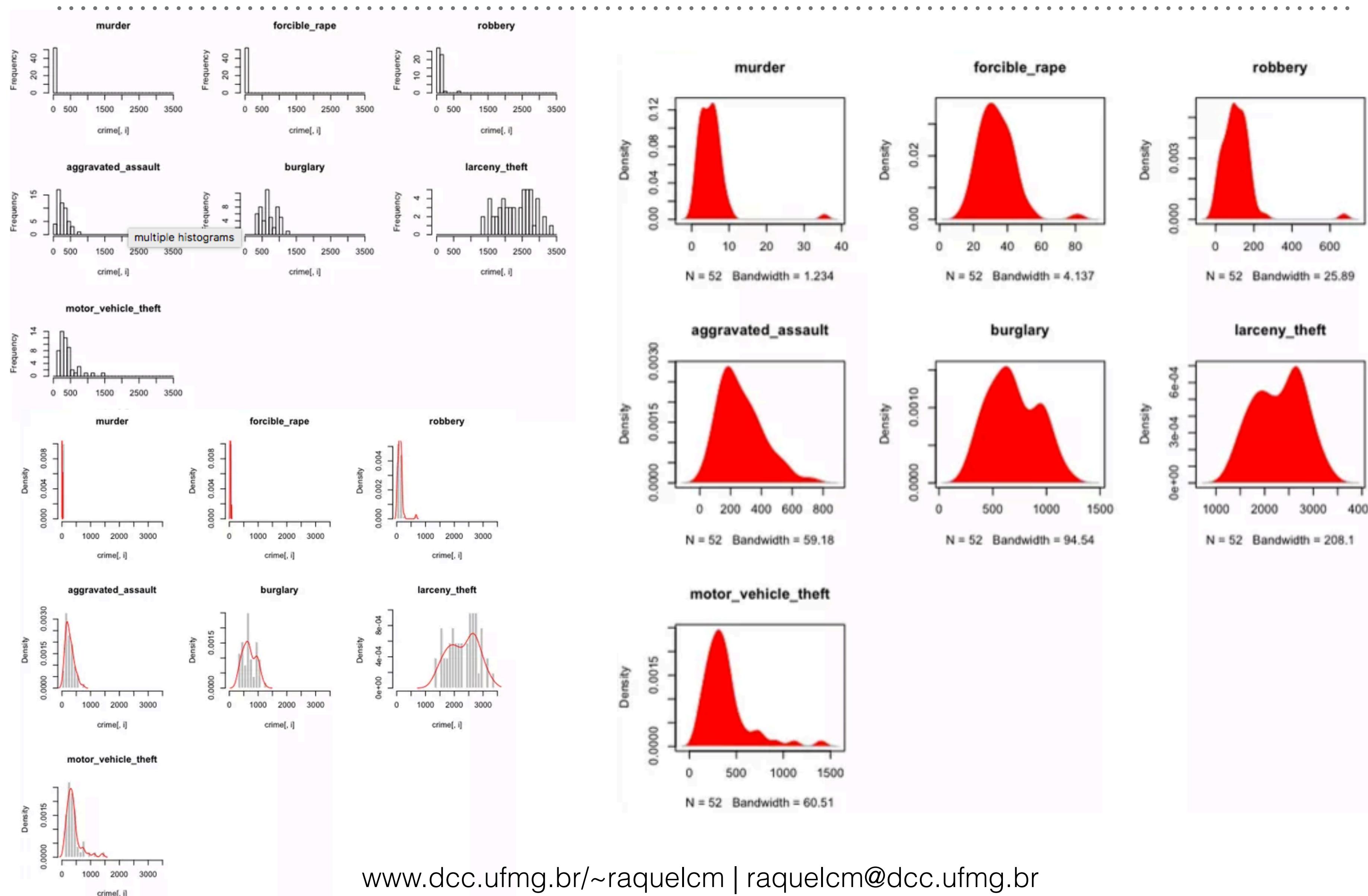


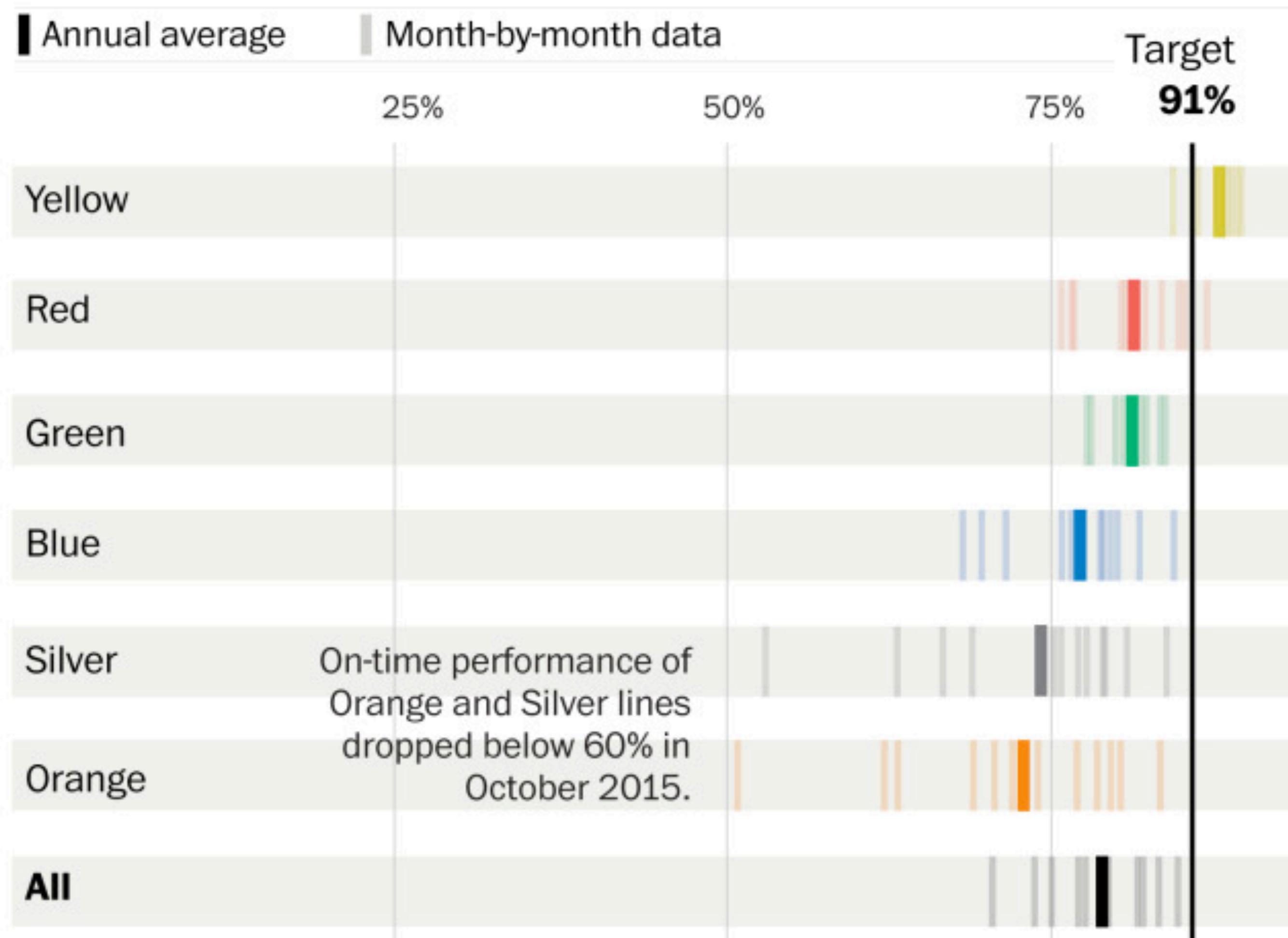
GRÁFICO DE DENSIDADE



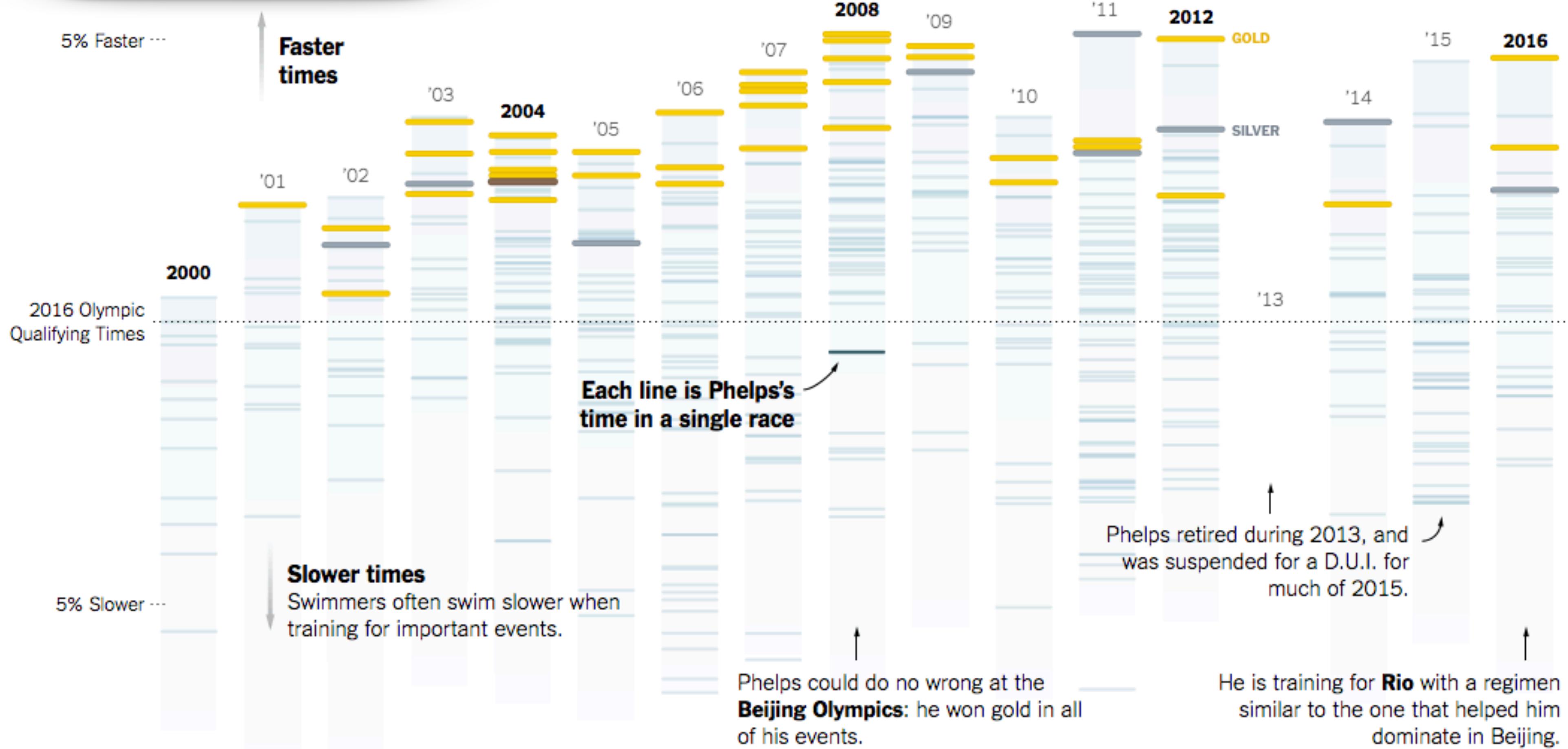
RUG PLOT OU BAR CODE

The system has performance issues

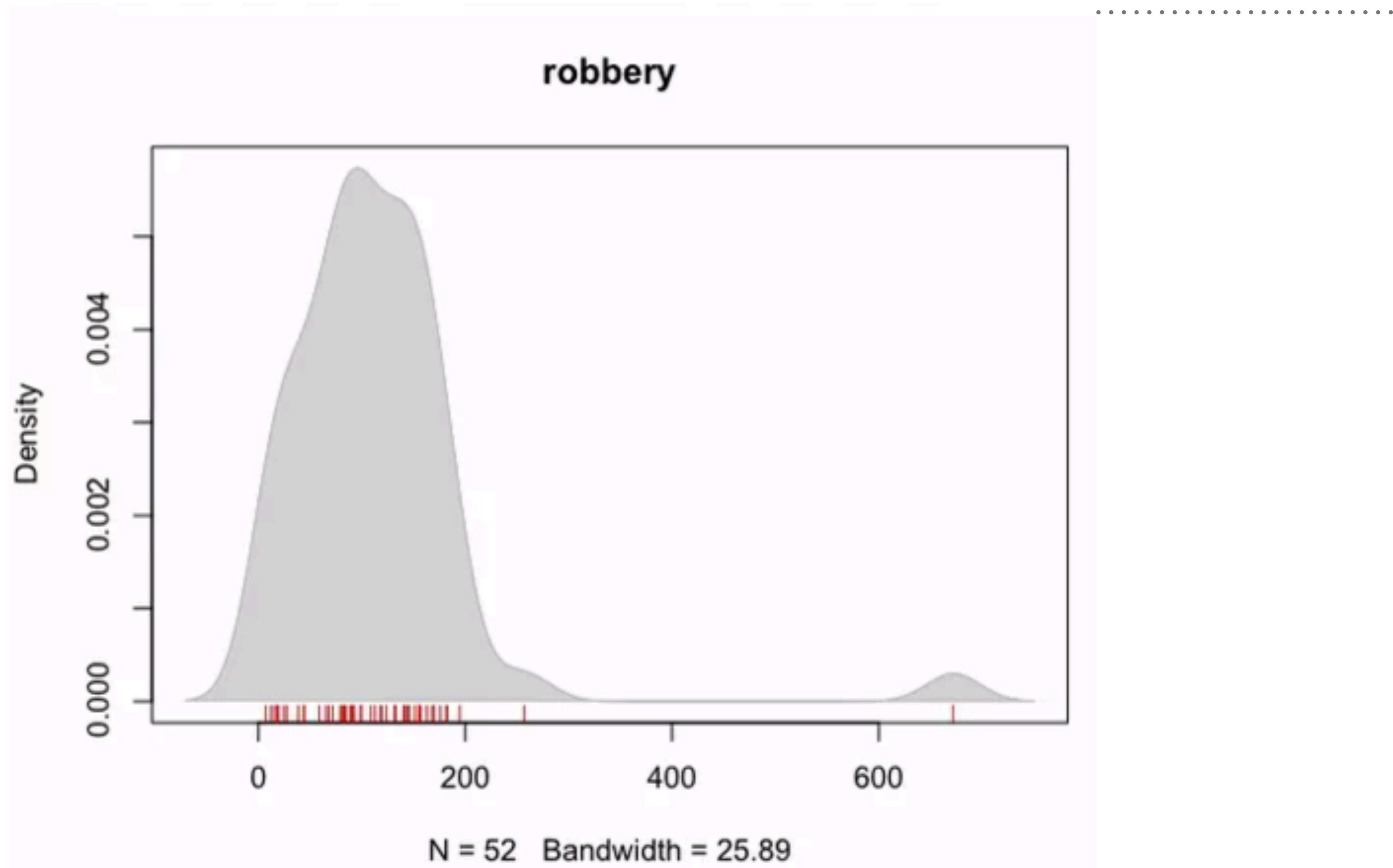
On-time performance by line in 2015



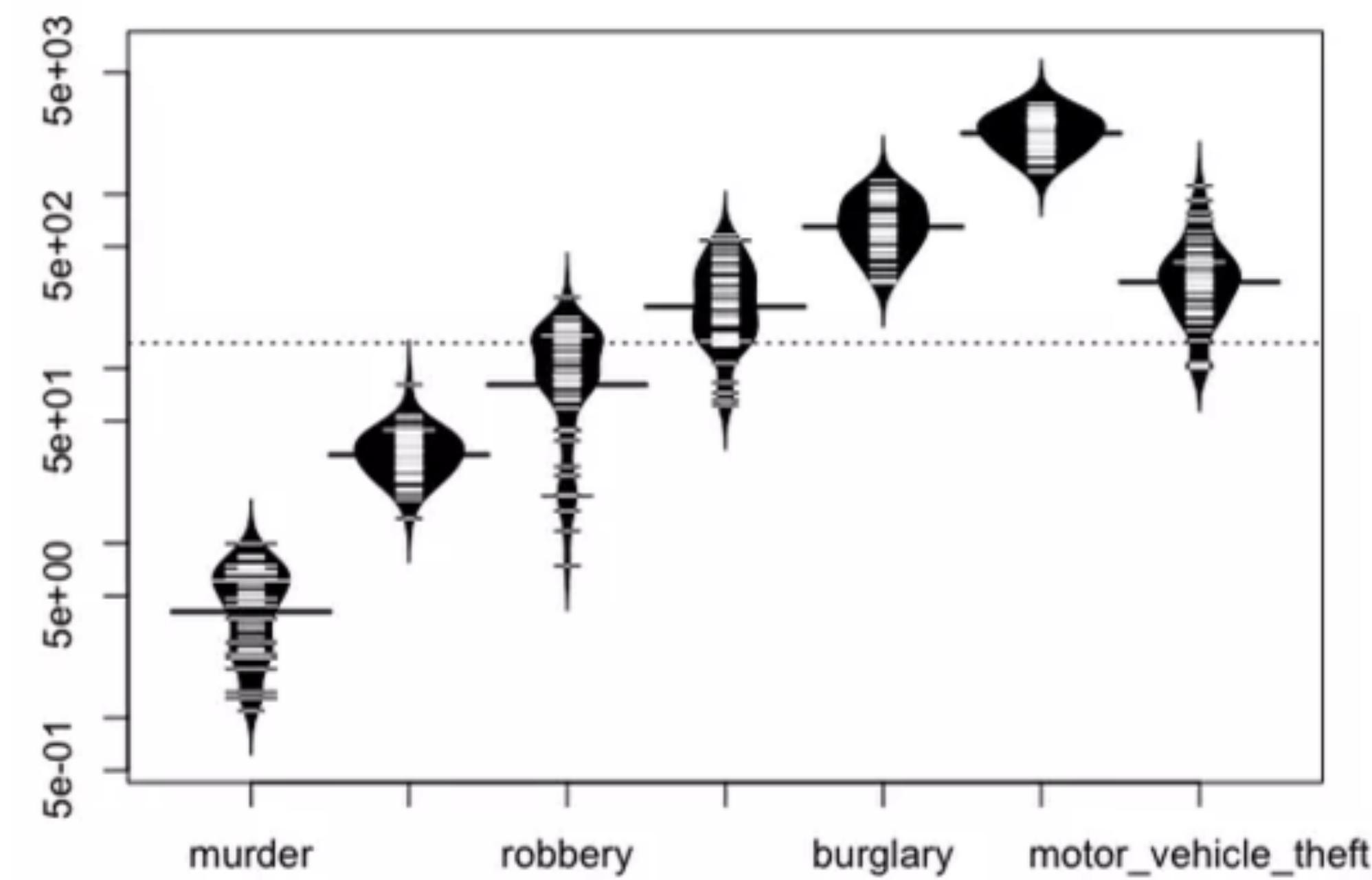
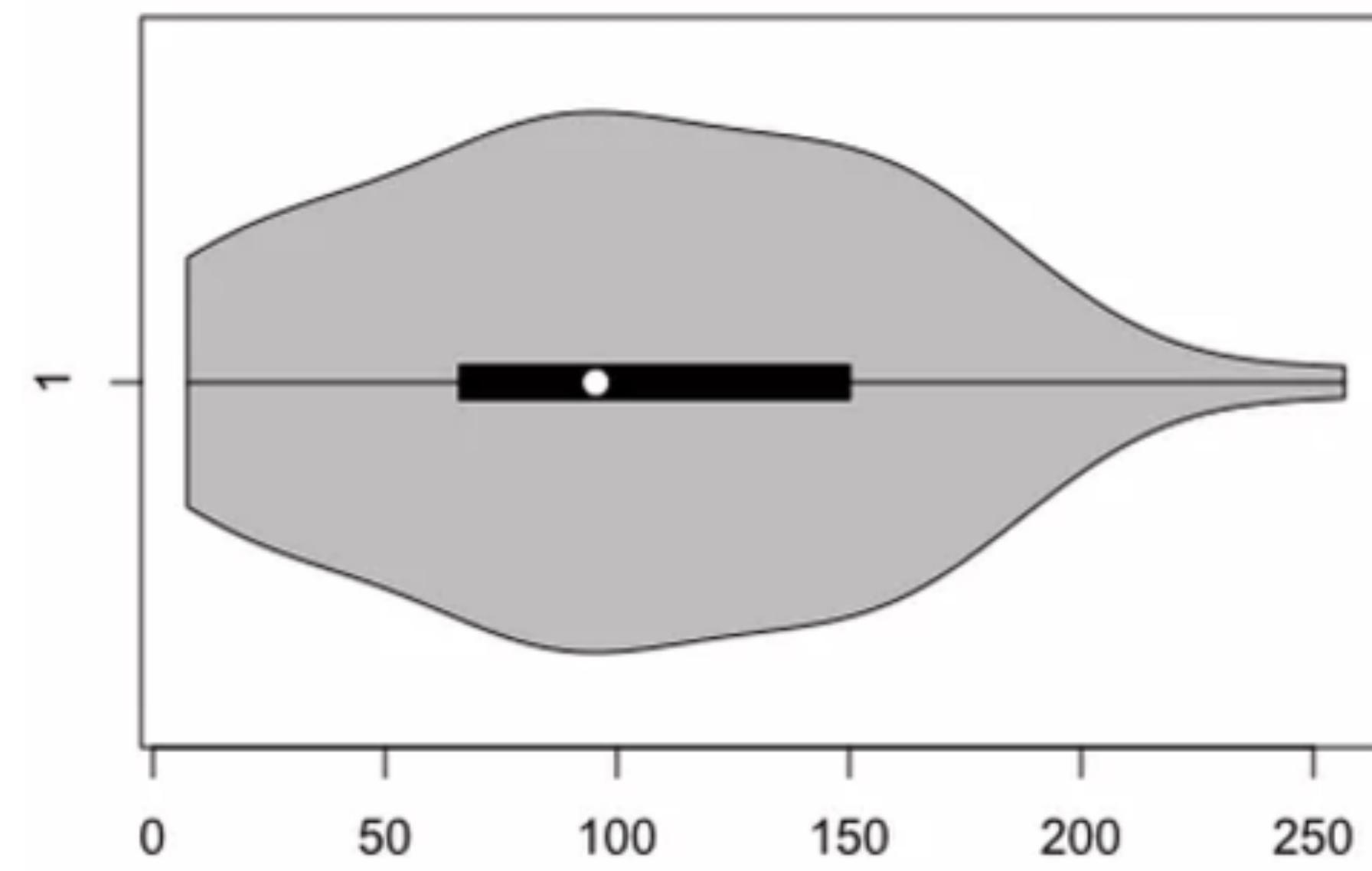
RUG PLOT OU BAR CODE



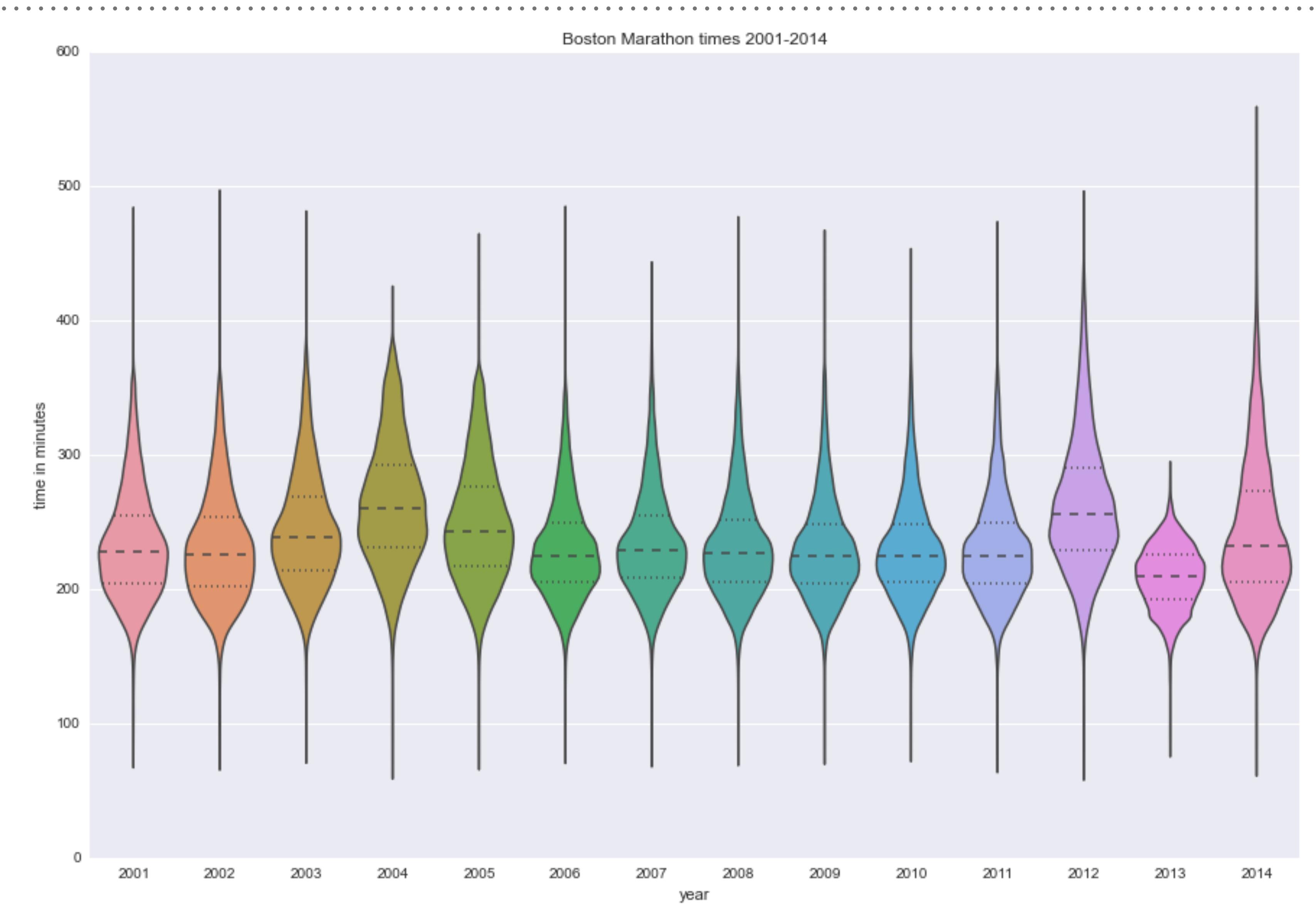
RUG PLOT OU BAR CODE



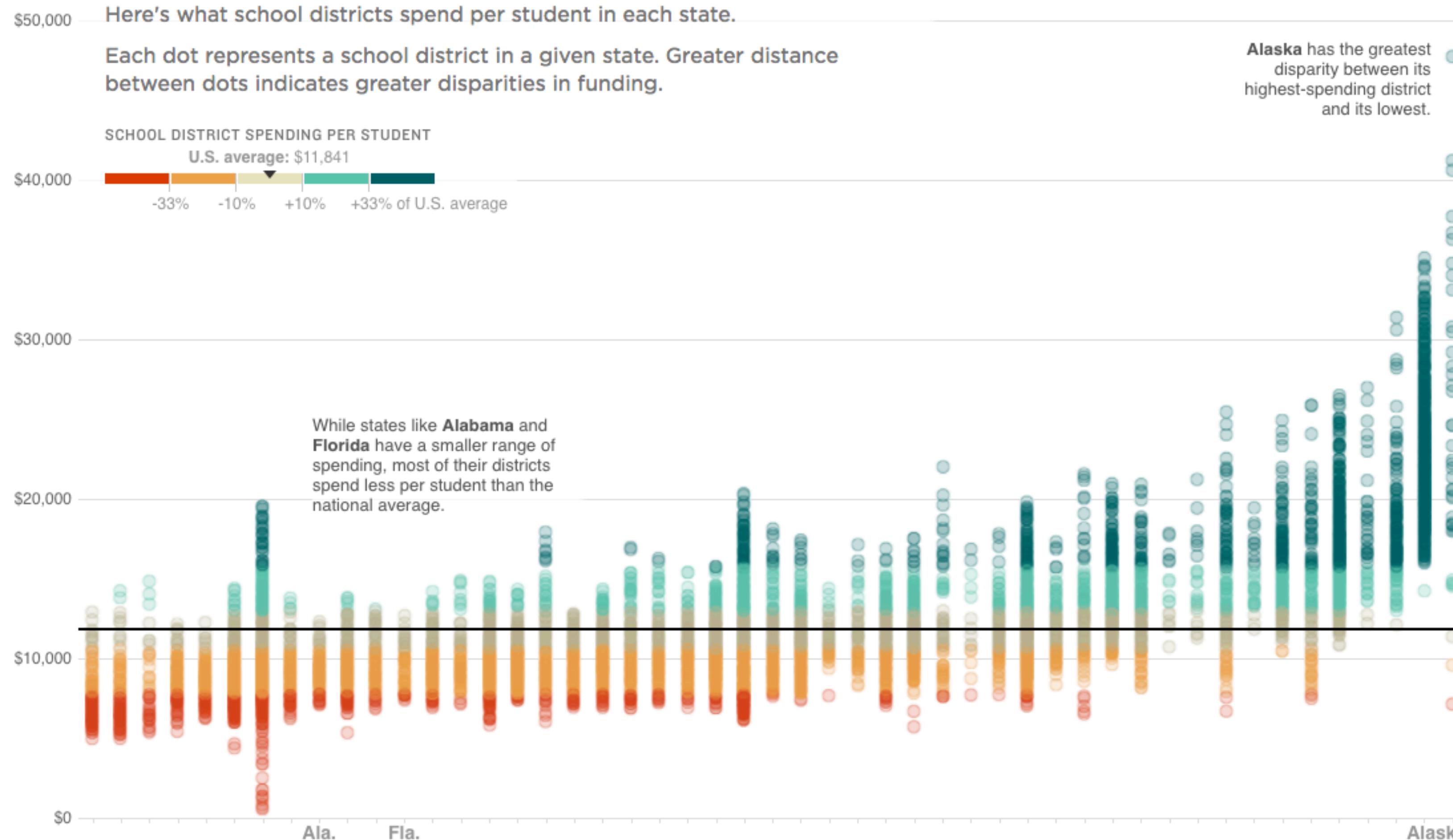
VIOLIN / BEAN PLOT



VIOLIN / BEAN PLOT



STRIP PLOT



Notes

This [Education Week](#) analysis of federal and state data excludes extreme outliers as well as districts with fewer than 200 students. Hawaii and Washington, D.C., are excluded because each has only one school district.

Source: [Education Week](#) analysis of federal and state data.

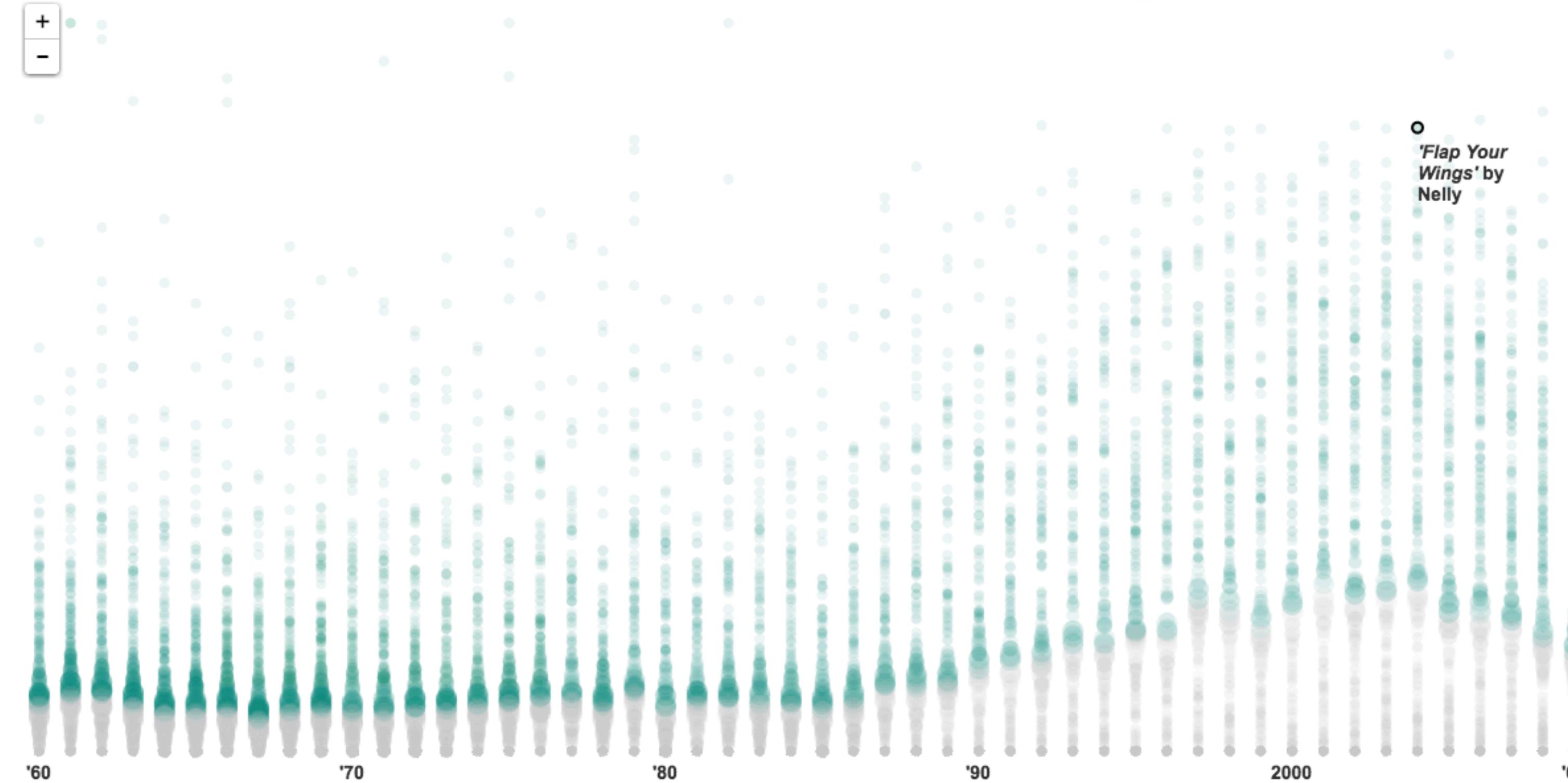
Credit: Katie Park, Alyson Hurt and Lisa Charlotte Rost/NPR

STRIP PLOT

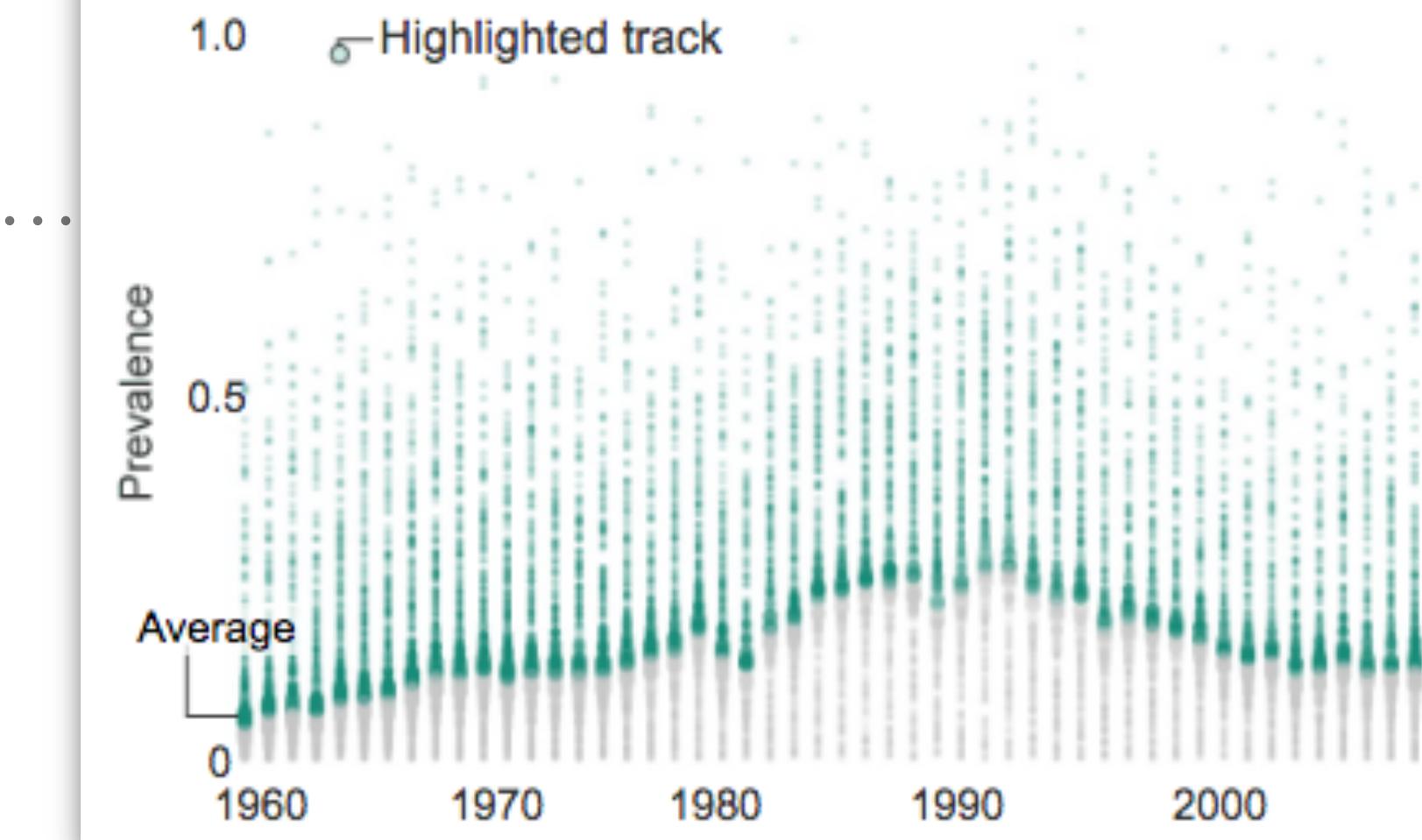
Hip-hop revolution ignites

Ramps up with the rise of rap and hip-hop in the late '80s and into the '90s, with artists like [Busta Rhymes](#), [Ludacris](#) and [Snoop Dogg](#).

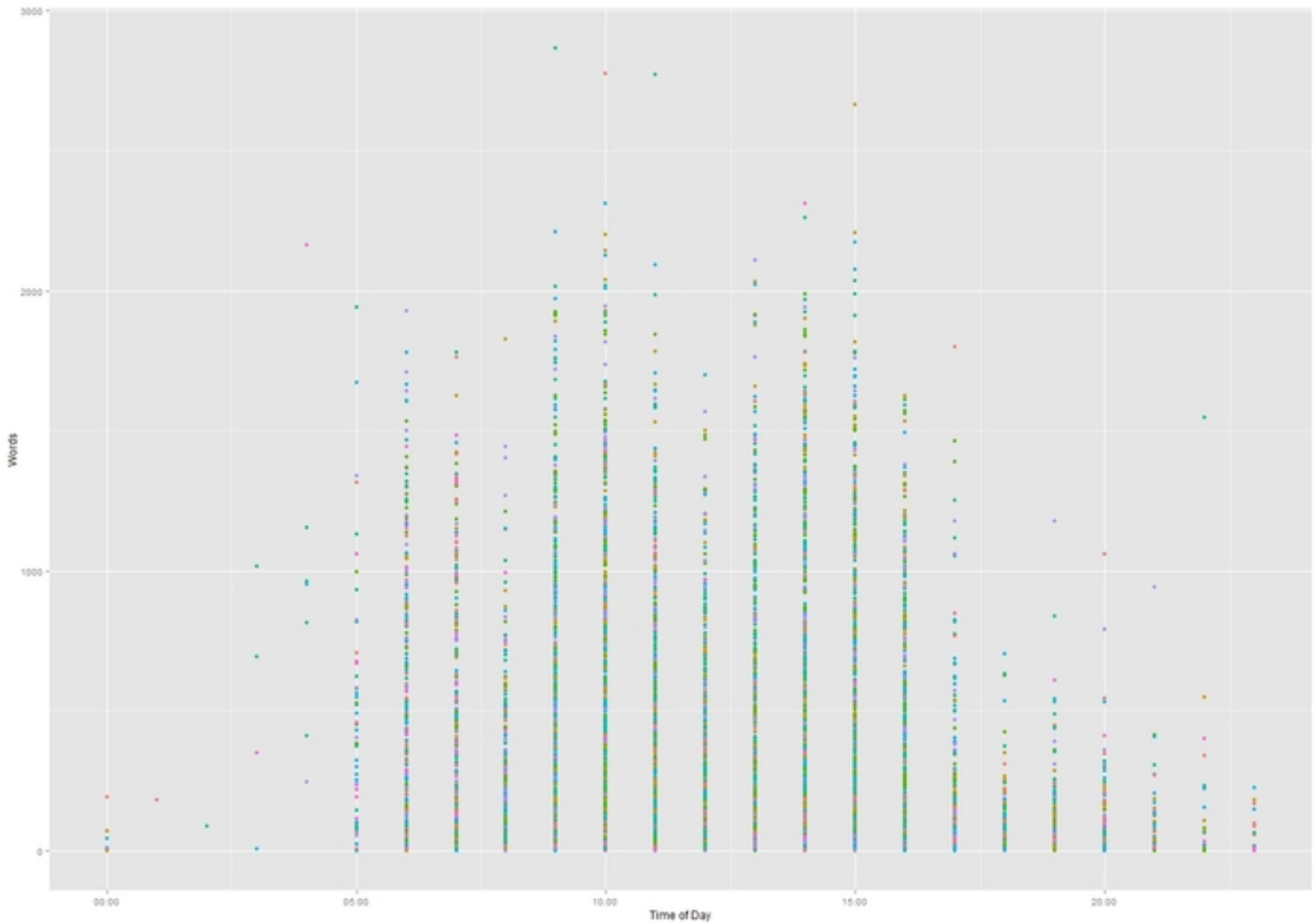
Elements: Energetic, speech, bright



Reading this chart:

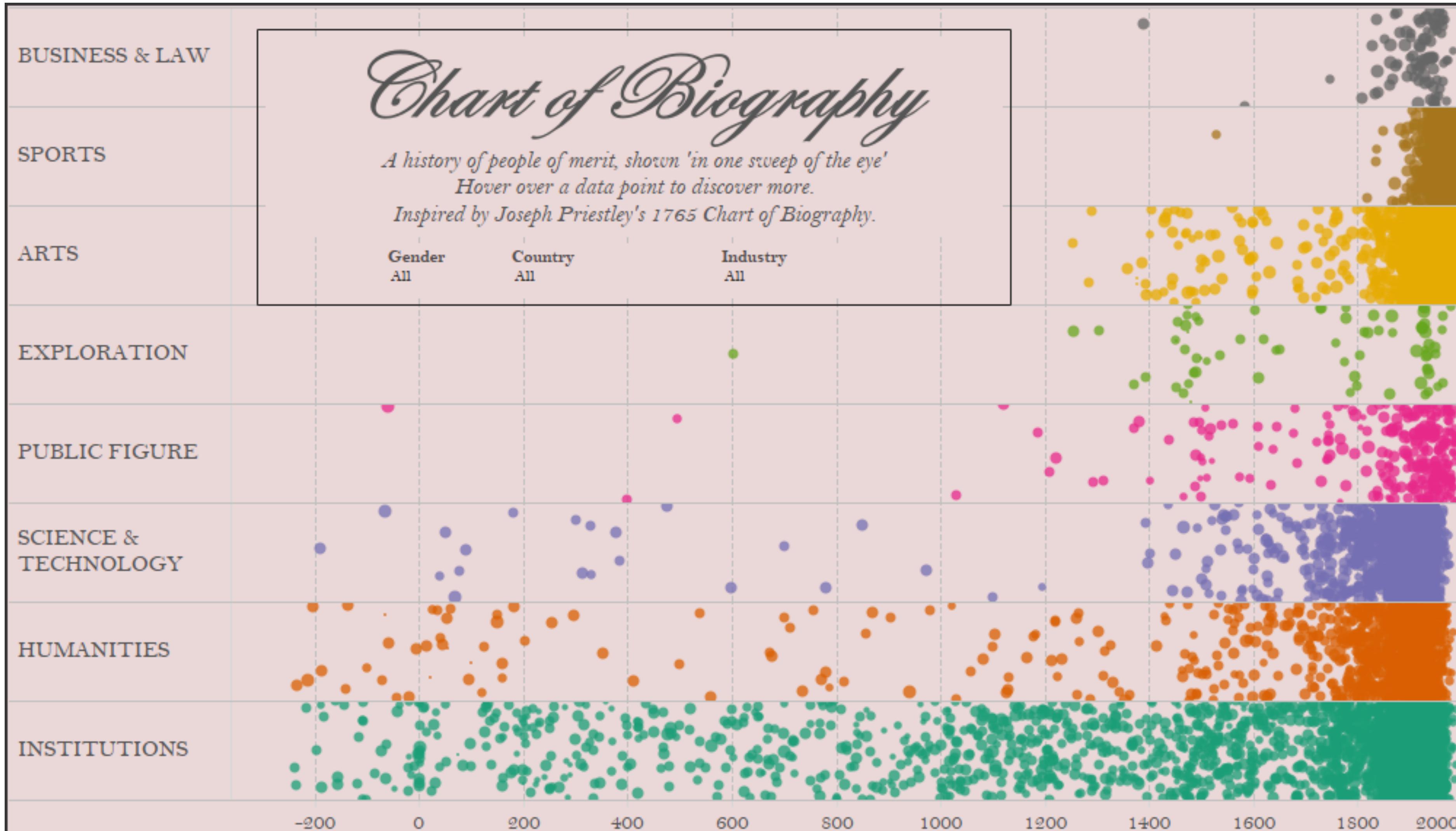


STRIP PLOT

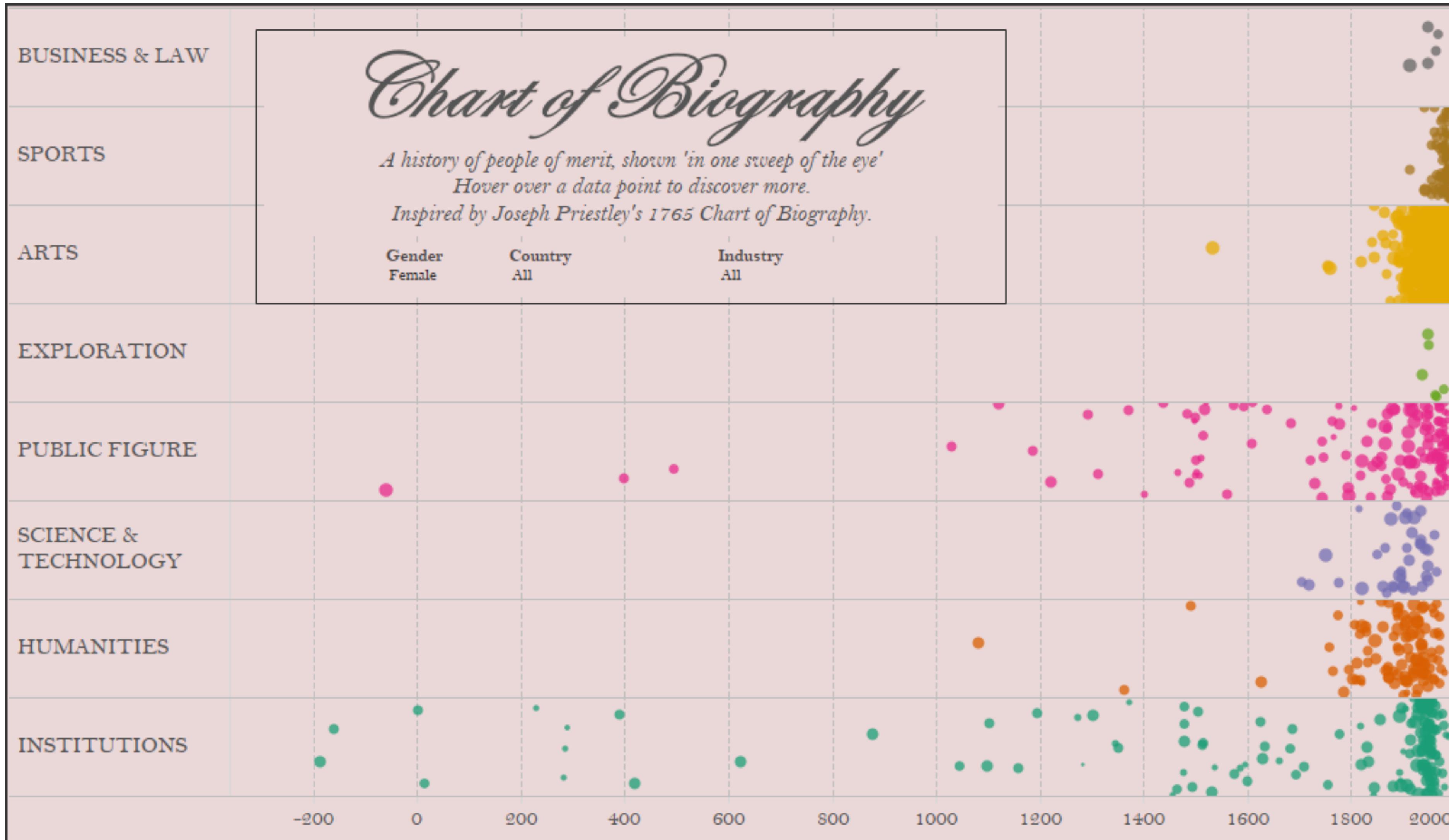


JITTERED STRIP PLOT

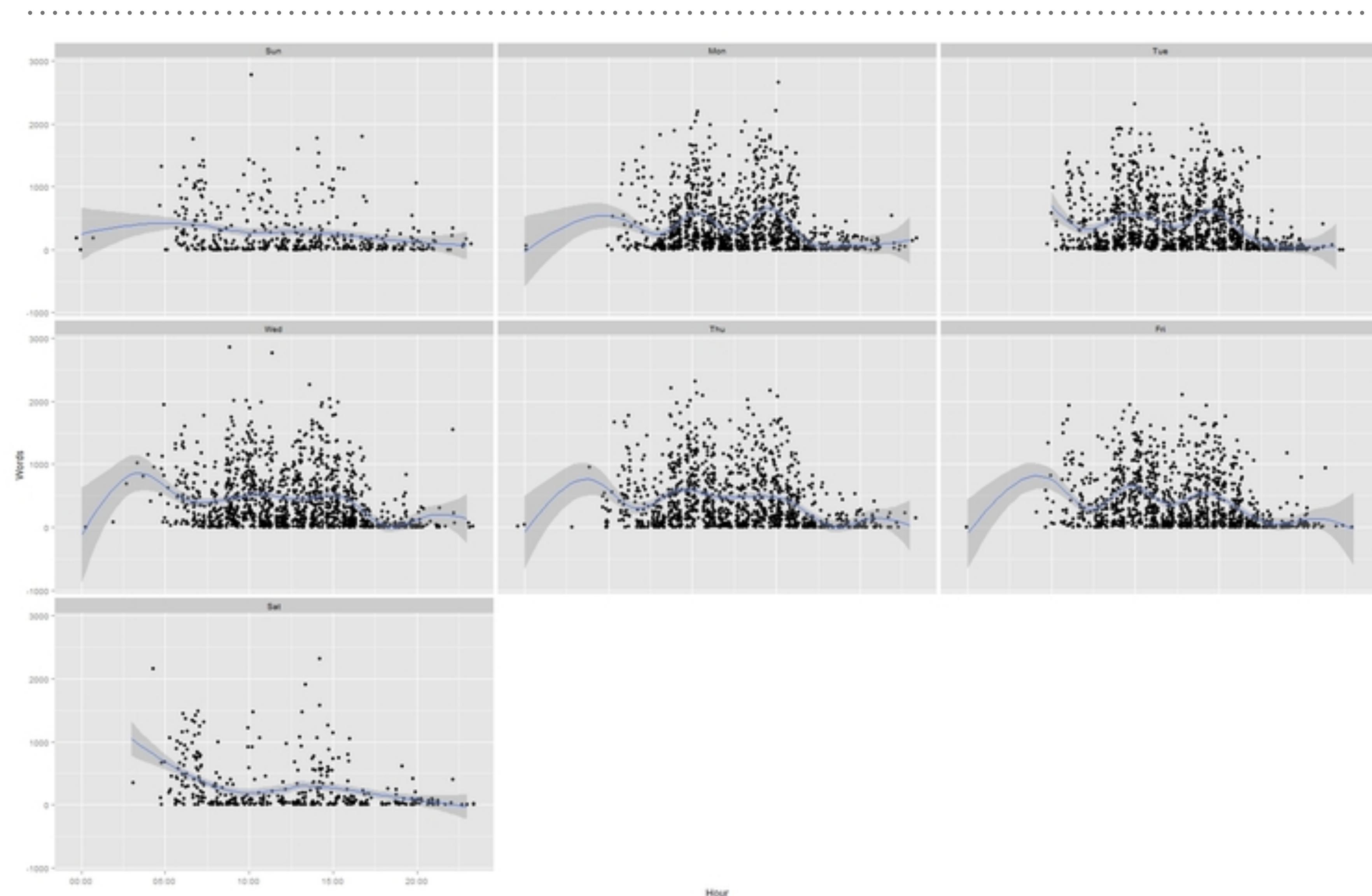
.....



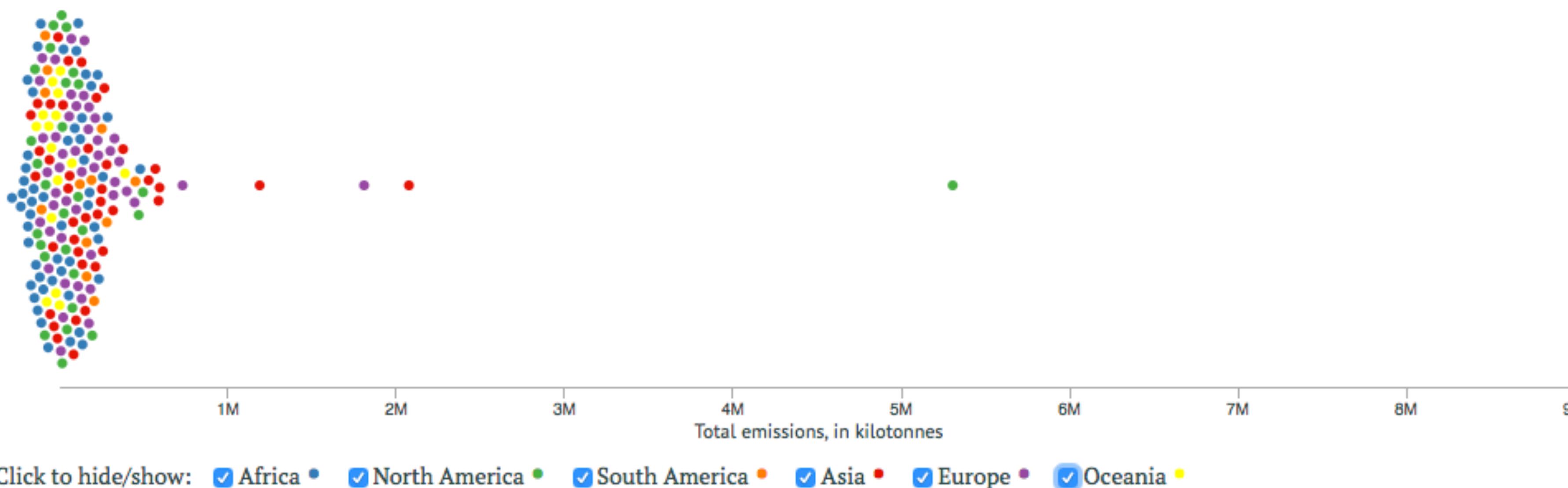
JITTERED STRIP PLOT



JITTERED STRIP PLOT



BEE SWARM PLOT

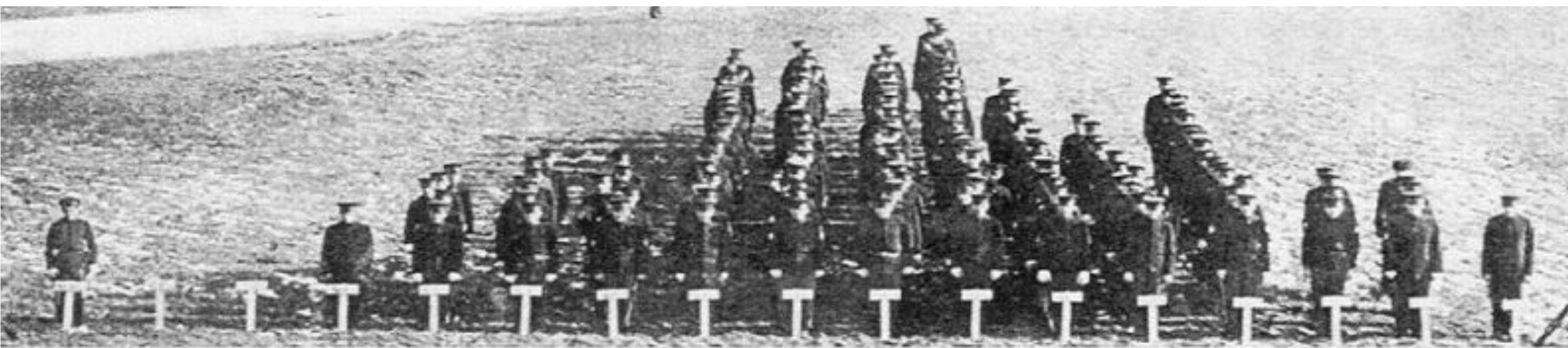


BEE SWARM PLOT

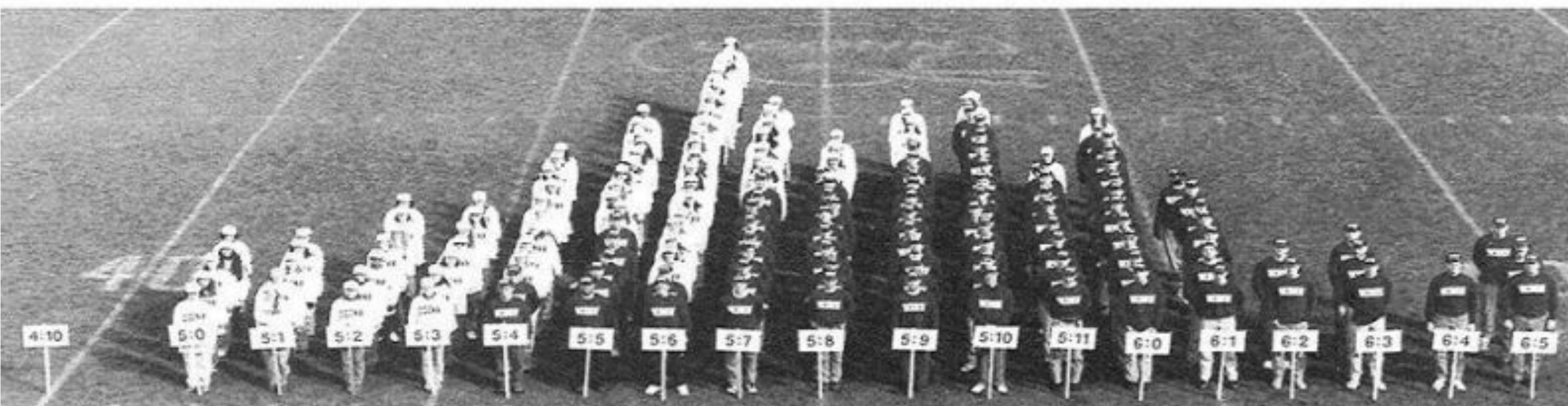
SELECT YEAR 1960 1980 2000 2014



UNIT PLOT



4:10 4:11 5:0 5:1 5:2 5:3 5:4 5:5 5:6 5:7 5:8 5:9 5:10 5:11 6:0 6:1 6:2

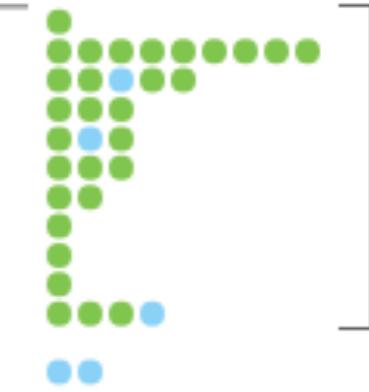


4:10 5:0 5:1 5:2 5:3 5:4 5:5 5:6 5:7 5:8 5:9 5:10 5:11 6:0 6:1 6:2 6:3 6:4 6:5

UNIT PLOT

Net finish time

2 hours —
12 minutes



Of the top 32 runners overall,
30 came from Ethiopia or
Kenya. The two others were
from Japan and North Korea.

2 hours —
30 minutes



Kim Hye-gyong of North Korea
broke the domination of
Africans by winning the
women's marathon

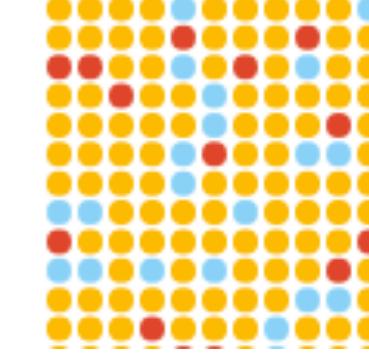
2 hours —
45 minutes



3 hours —



3 hours —
15 minutes



In the long run

(2015 revision)

By Cedric Sam

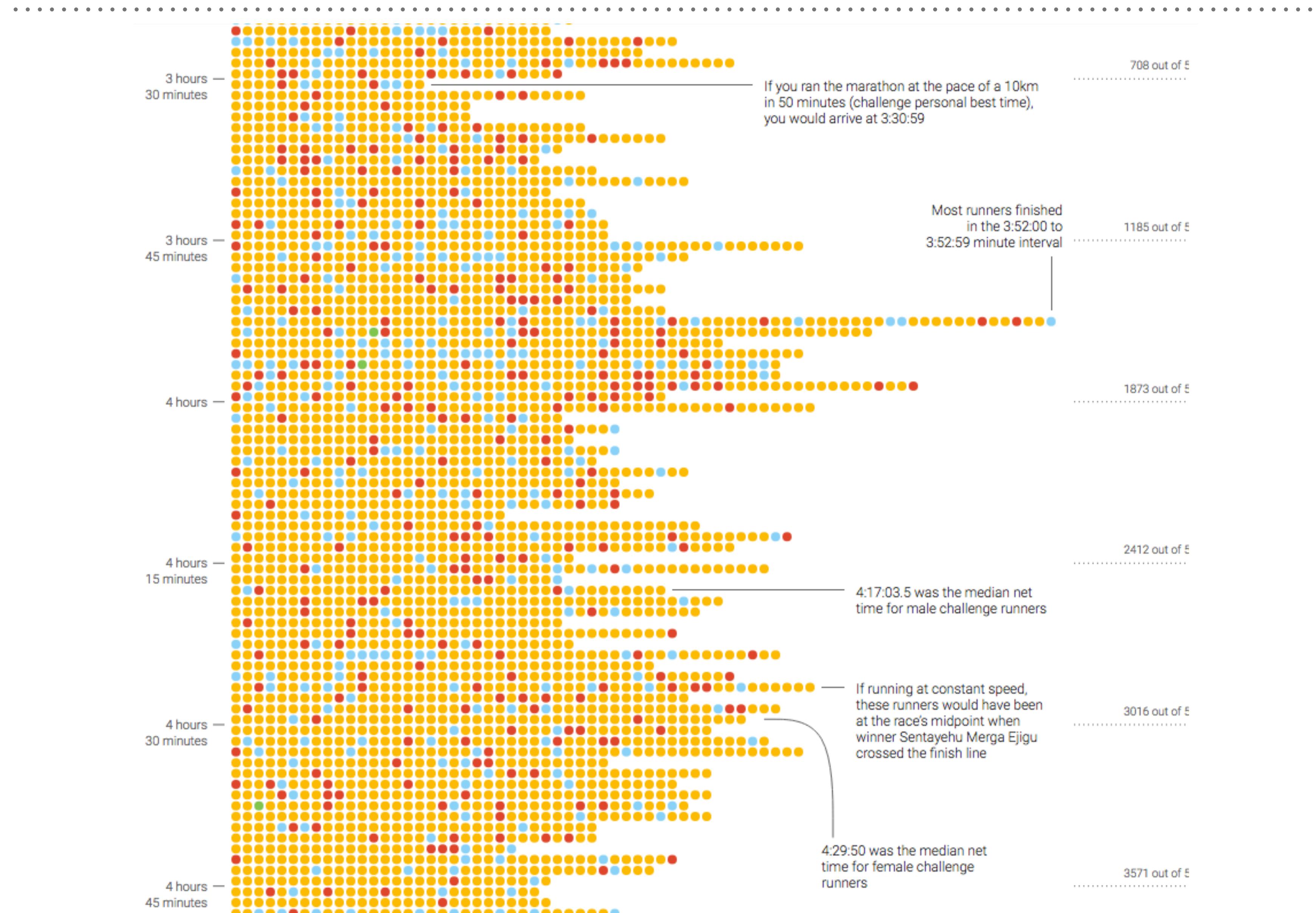
This year's Standard Chartered Hong Kong Marathon was won on Sunday morning by Ethiopia's Sentayehu Merga Ejigu in 2:12:59 (net time), almost 2 minutes better than the best score in 2014. In the women's competition, Kim Hye-gyong of North Korea broke the domination of Africans by winning in 2:31:46. We reprise our 2014 graphic and again trace the milestones of the marathon challenge, the toughest among the event's races, from the firing gun to the last second. (Published On 2015-01-26)

[By gender](#) [By origin](#) [By category](#)

● Hong Kong & Macau ● China ● African countries ● Other

311 finishing at this mark out of 5
finishing the race under six hours

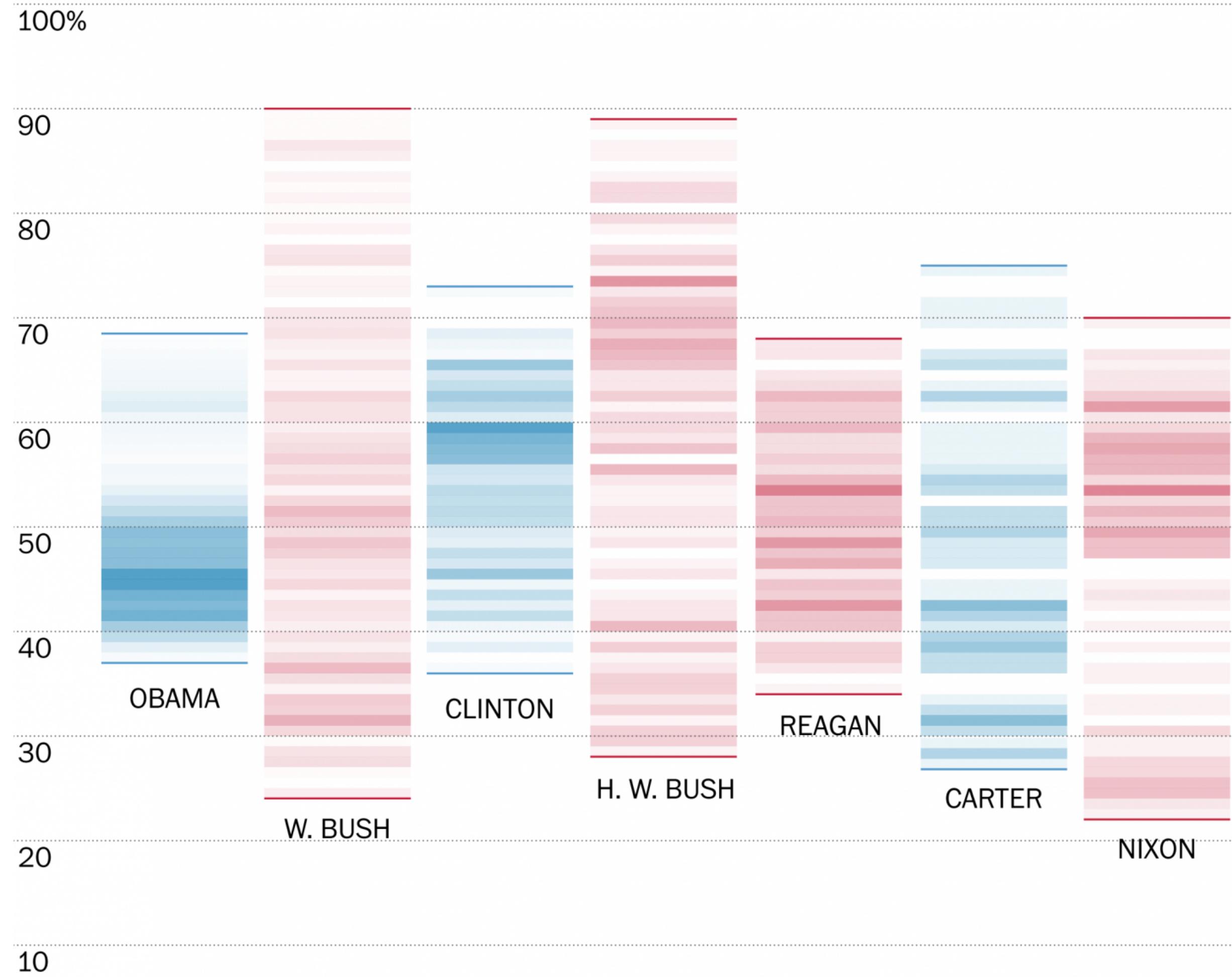
UNIT PLOT



MAPA DE CALOR

Presidential approval frequency

According to polling from Gallup.



MAPA DE CALOR

Utilization of Squares by Chess Masters

Alexander Alekhine

	Playing as White								Playing as Black							
8	0.12	0.12	0.16	0.19	0.13	0.11	0.08	0.07	0.17	0.42	0.75	0.95	0.79	1.35	1.04	0.22
7	0.26	0.27	0.27	0.35	0.38	0.28	0.17	0.18	0.13	0.55	0.58	1.29	1.47	0.56	0.54	0.22
6	0.27	0.32	0.48	0.60	0.50	0.58	0.38	0.29	0.65	0.90	1.77	1.26	1.48	2.15	0.96	0.63
5	0.36	0.77	0.72	1.36	1.30	0.77	0.96	0.42	0.74	0.76	1.46	1.86	1.56	1.17	0.73	0.65
4	0.72	0.58	1.52	2.04	1.80	1.24	0.67	0.64	0.40	0.88	0.73	1.07	1.04	0.65	0.71	0.39
3	0.48	0.81	1.71	1.21	1.32	1.95	0.84	0.52	0.23	0.36	0.63	0.46	0.43	0.43	0.28	0.26
2	0.10	0.21	0.62	1.01	1.02	0.43	0.40	0.18	0.22	0.32	0.30	0.33	0.26	0.24	0.21	0.16
1	0.16	0.34	0.74	0.94	0.69	1.16	0.95	0.22	0.12	0.12	0.16	0.22	0.18	0.15	0.08	0.06
	a	b	c	d	e	f	g	h	a	b	c	d	e	f	g	h

Viswanathan Anand

	Playing as White								Playing as Black							
8	0.15	0.11	0.14	0.22	0.14	0.10	0.07	0.07	0.23	0.52	0.83	0.90	0.78	1.47	1.12	0.23
7	0.21	0.28	0.26	0.27	0.26	0.22	0.19	0.14	0.16	0.57	0.66	1.44	1.37	0.45	0.69	0.25
6	0.29	0.34	0.51	0.49	0.40	0.48	0.28	0.25	0.77	0.91	1.62	1.14	1.49	2.17	0.91	0.61
5	0.39	0.92	0.59	1.21	1.10	0.59	0.85	0.39	0.74	0.83	1.55	1.86	1.35	0.92	0.58	0.50
4	0.88	0.62	1.30	2.10	1.79	1.02	0.73	0.68	0.57	1.09	1.14	1.36	2.11	0.88	0.64	
3	0.23	0.36	0.63	0.46	0.43	0.43	0.28	0.26	0.25	0.32	0.55	0.35	0.29	0.33	0.19	0.12
2	0.17	0.27	0.61	1.35	1.17	0.49	0.51	0.27	0.19	0.24	0.22	0.15	0.15	0.14	0.09	
1	0.19	0.50	0.77	0.99	0.88	1.34	1.01	0.26	0.12	0.10	0.16	0.18	0.11	0.09	0.06	0.05
	a	b	c	d	e	f	g	h	a	b	c	d	e	f	g	h

José Raúl Capablanca

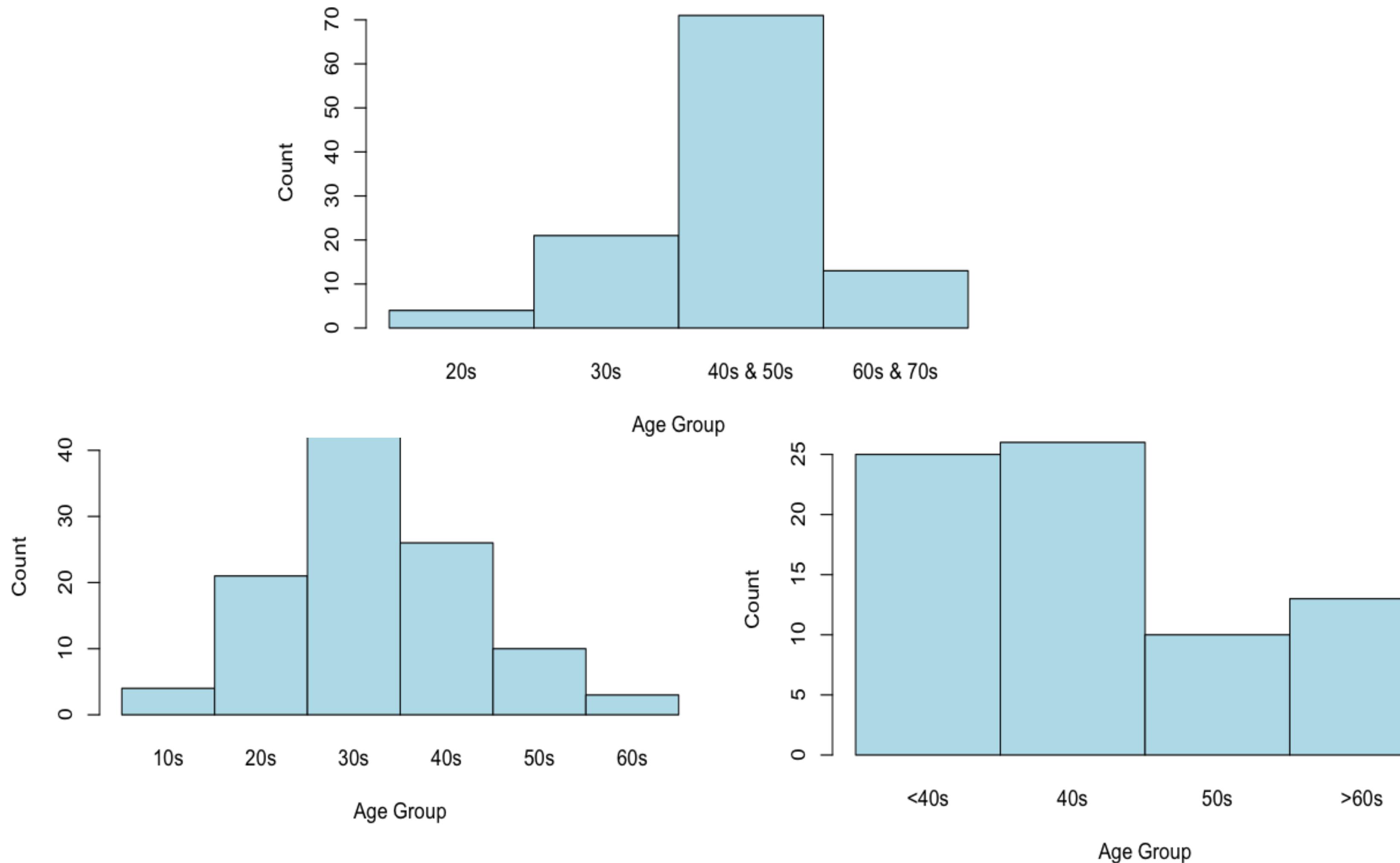
	Playing as White								Playing as Black							
8	0.09	0.09	0.15	0.25	0.20	0.12	0.03	0.03	0.18	0.37	0.72	0.88	0.81	1.28	1.04	0.18
7	0.21	0.27	0.25	0.32	0.47	0.19	0.19	0.11	0.10	0.54	0.57	1.49	1.57	0.47	0.45	0.19
6	0.23	0.30	0.54	0.46	0.48	0.56	0.34	0.22	0.66	0.94	1.69	1.14	1.46	2.16	0.82	0.58
5	0.32	0.80	0.84	1.19	1.12	0.68	1.05	0.38	0.75	0.78	1.44	1.95	1.34	0.94	0.61	0.64
4	0.32	0.83	0.92	1.27	1.82	1.58	1.05	0.76	0.69	0.32	0.83	0.92	1.02	1.10	0.54	0.32
3	0.51	0.89	1.70	2.27	2.03	0.89	0.60	0.24	0.21	0.34	0.59	0.40	0.29	0.45	0.24	0.18
2	0.09	0.27	0.62	1.04	1.05	0.44	0.59	0.24	0.19	0.19	0.26	0.39	0.23	0.15	0.23	0.12
1	0.13	0.34	0.82	0.79	0.76	1.24	1.02	0.13	0.11	0.10	0.17	0.19	0.19	0.10	0.06	0.05
	a	b	c	d	e	f	g	h	a	b	c	d	e	f	g	h

Magnus Carlsen

	Playing as White								Playing as Black							
8	0.14	0.16	0.18	0.27	0.19	0.15	0.08	0.10	0.27	0.51	0.96	1.02	1.58	1.21	0.25	
7	0.24	0.33	0.30	0.39	0.30	0.28	0.21	0.17	0.18	0.62	0.62	1.50	1.48	0.60	0.90	0.26
6	0.34	0.38	0.60	0.61	0.51	0.55	0.34	0.28	0.94	1.00	1.67	1.28	1.55	2.30	1.07	0.60
5	0.47	0.89	0.80	1.40	1.28	0.77	0.88	0.44	0.85	0.88	1.65	1.88	1.55	1.12	0.71	0.62
4	0.88	0.75														

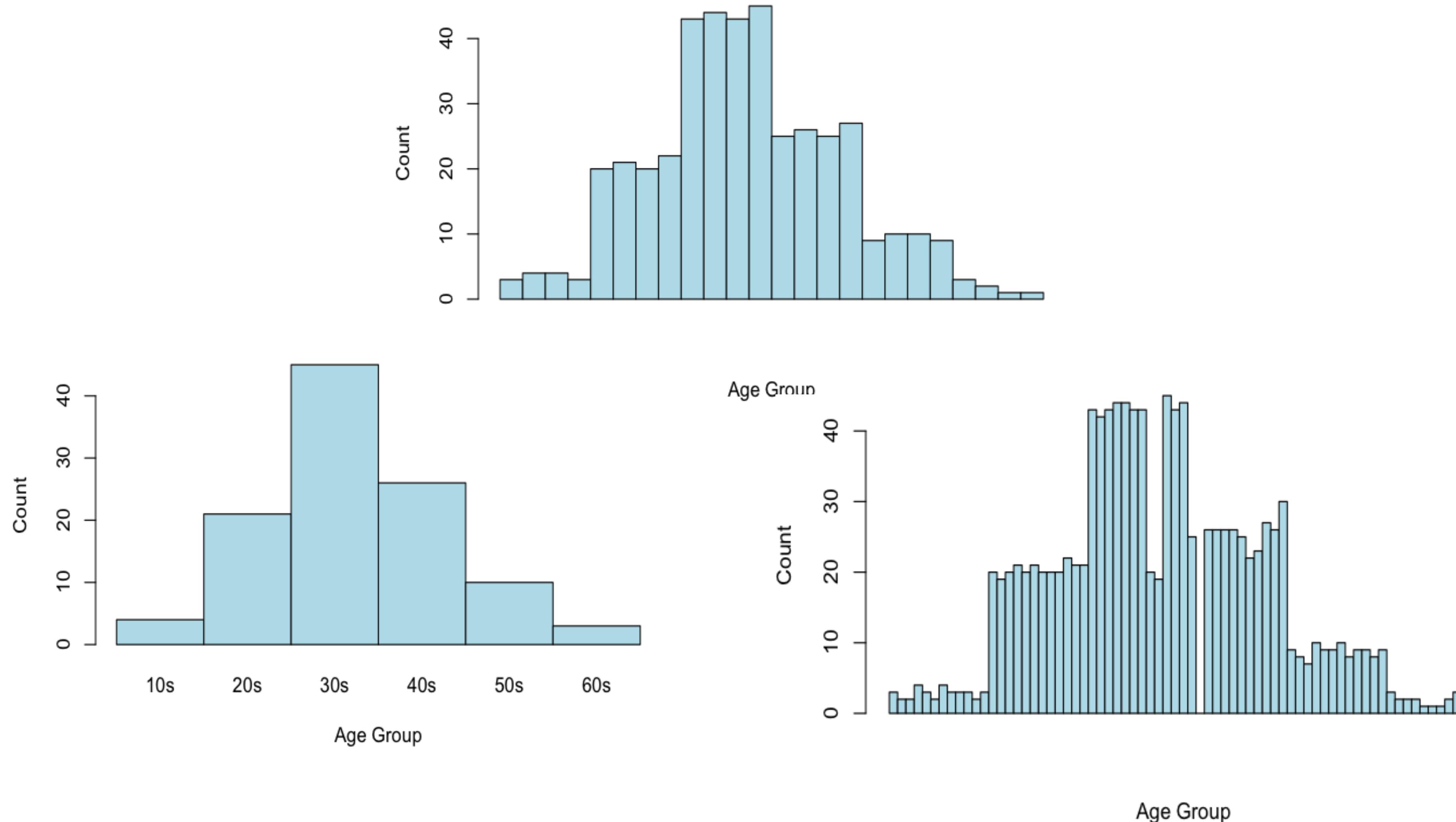
BOAS PRÁTICAS

MANTENHA OS INTERVALOS CONSISTENTES



ESCOLHA O TAMANHO APROPRIADO PARA OS INTERVALOS

.....



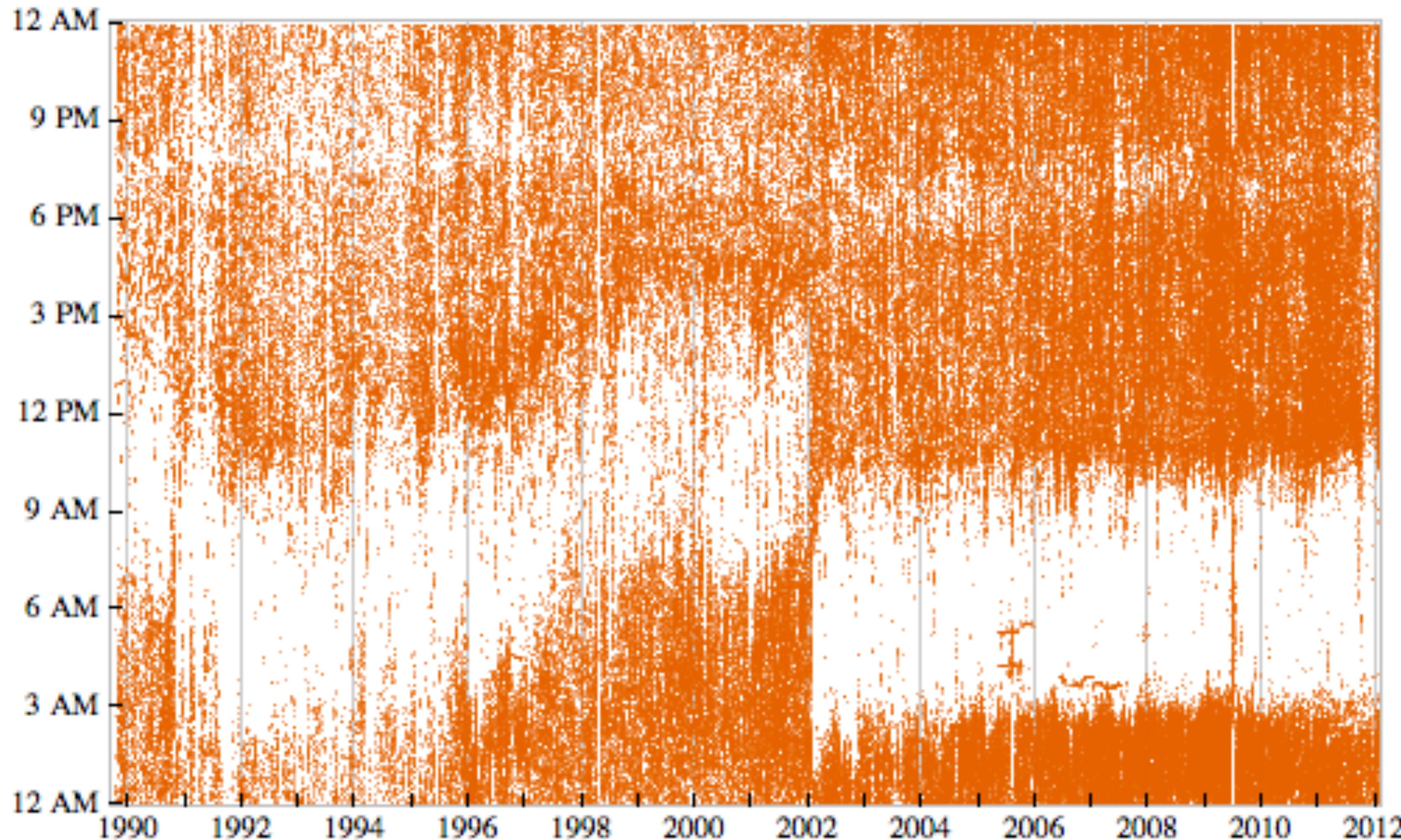
USE METRICAS RESISTENTES A EXCEÇÕES

- A mediana e os percentis são medidas “resistentes” a exceções
- A média e o desvio padrão não o são

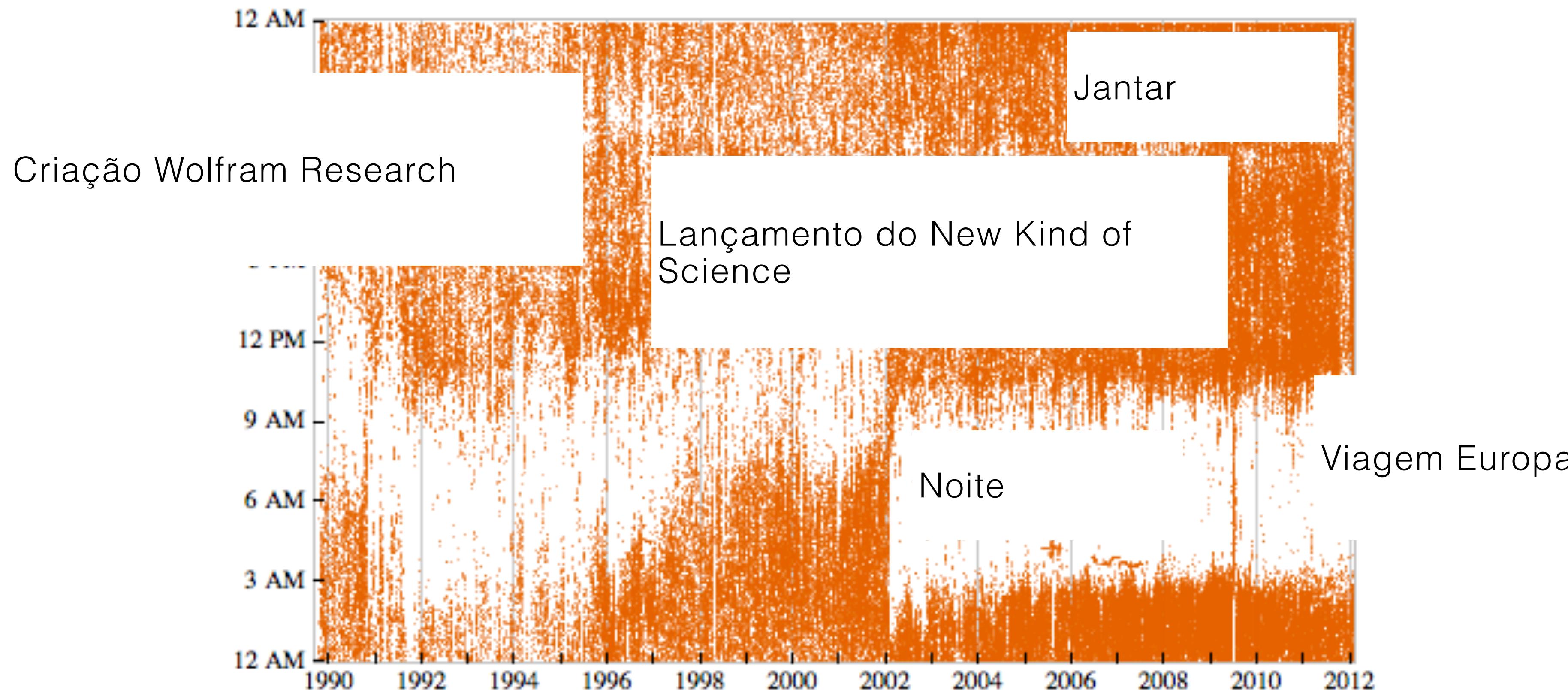
UM EXEMPLO

Personal Analytics

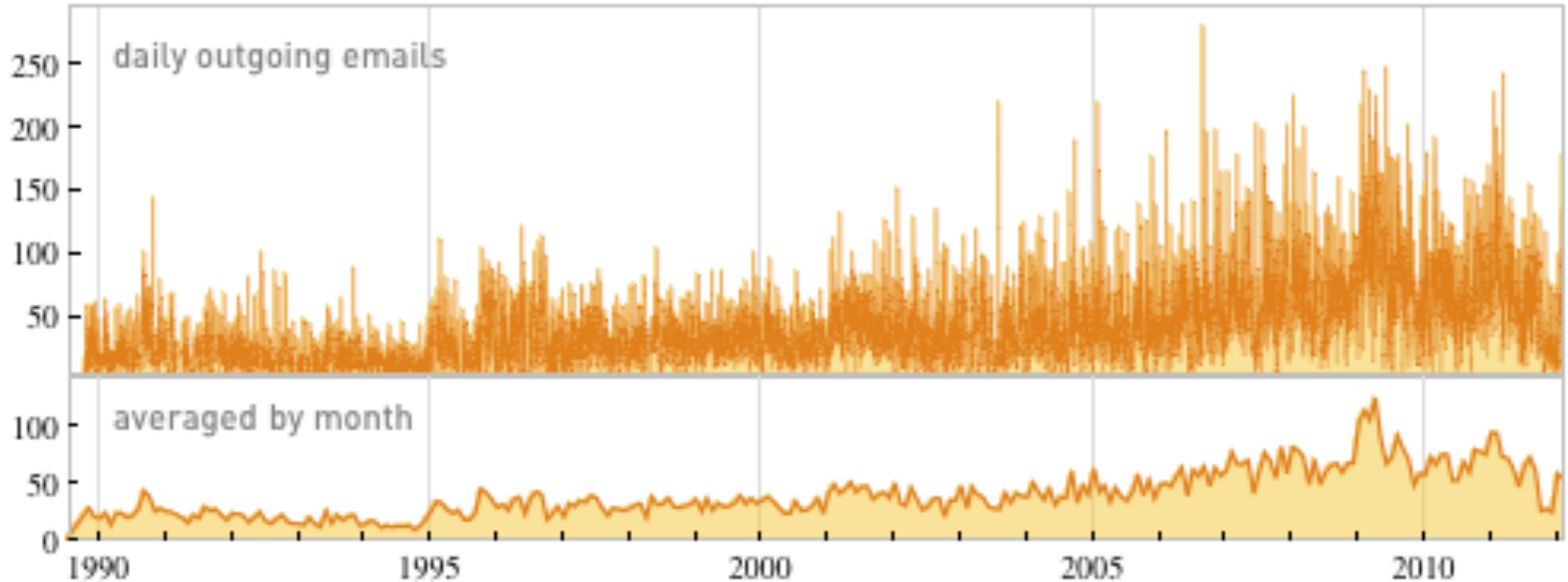
Pontos representam os cerca 300.000 **e-mails enviados** desde 1989



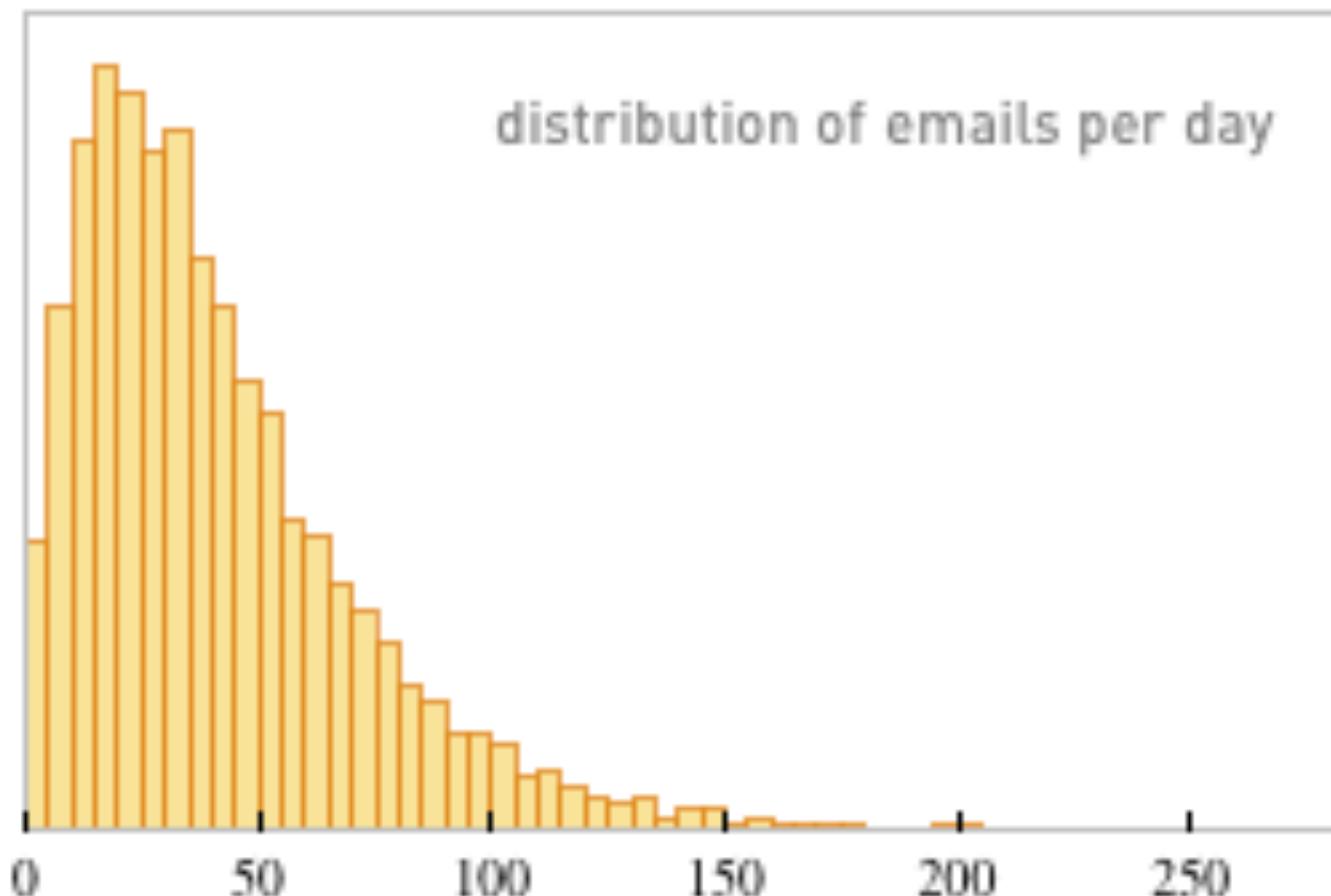
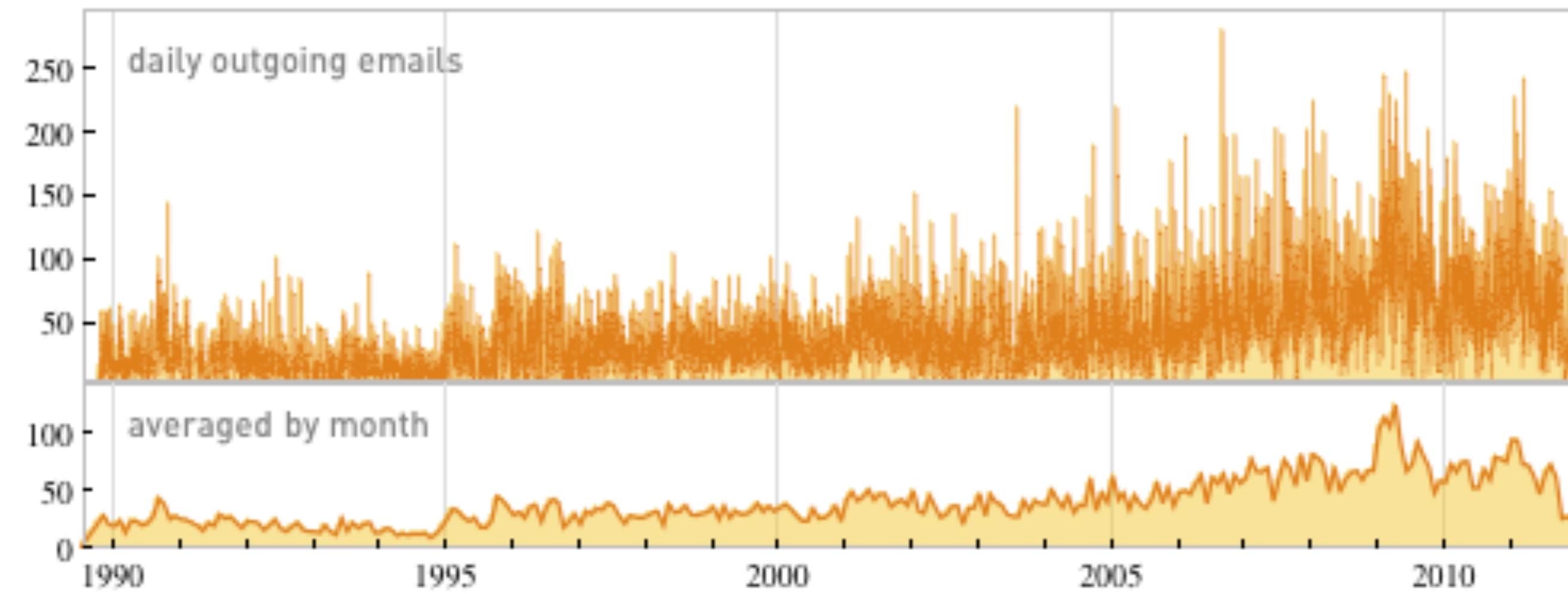
Pontos representam os cerca 300.000 **e-mails enviados** desde 1989



Quantidade de **e-mails enviados** ao longo do tempo e
sua média mensal



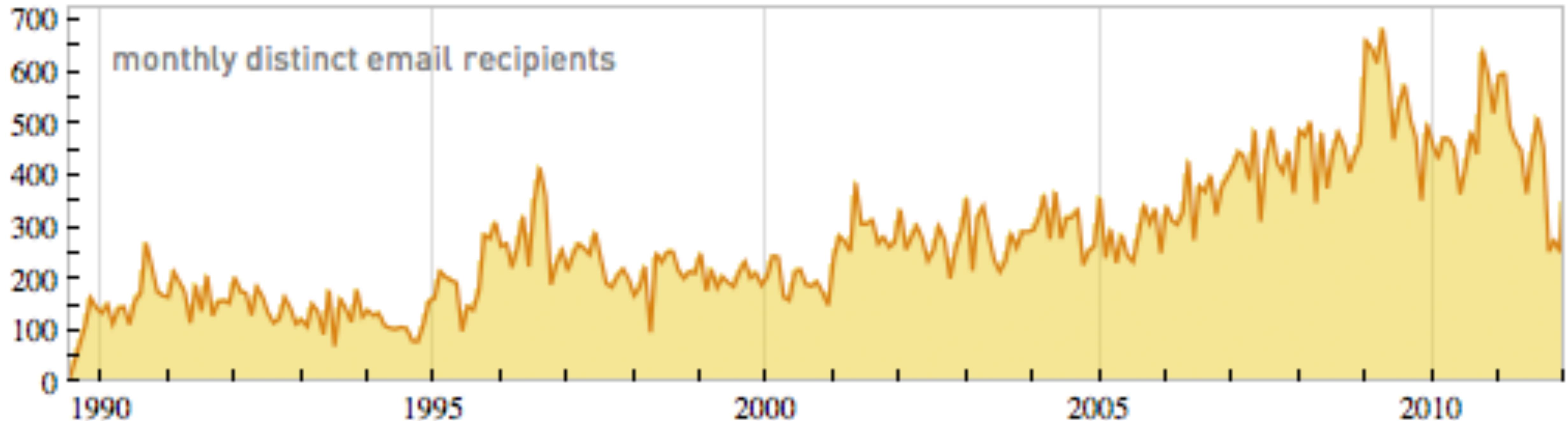
Quantidade de **e-mails enviados** ao longo do tempo e sua média mensal



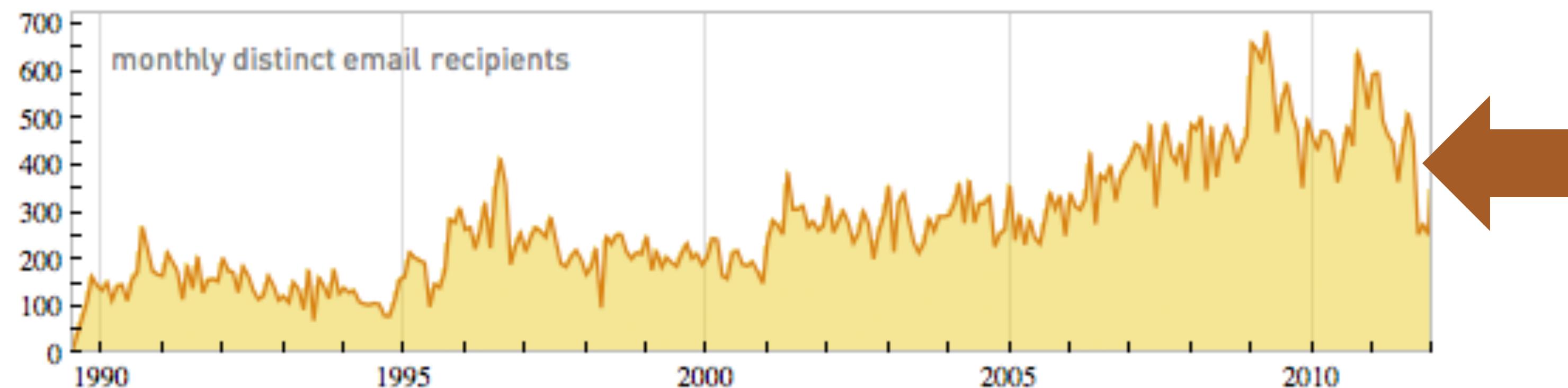
Mudança de perspectiva:
distribuição?

A cauda parece seguir uma lei de
potência

Número de **diferentes destinatários** por dia

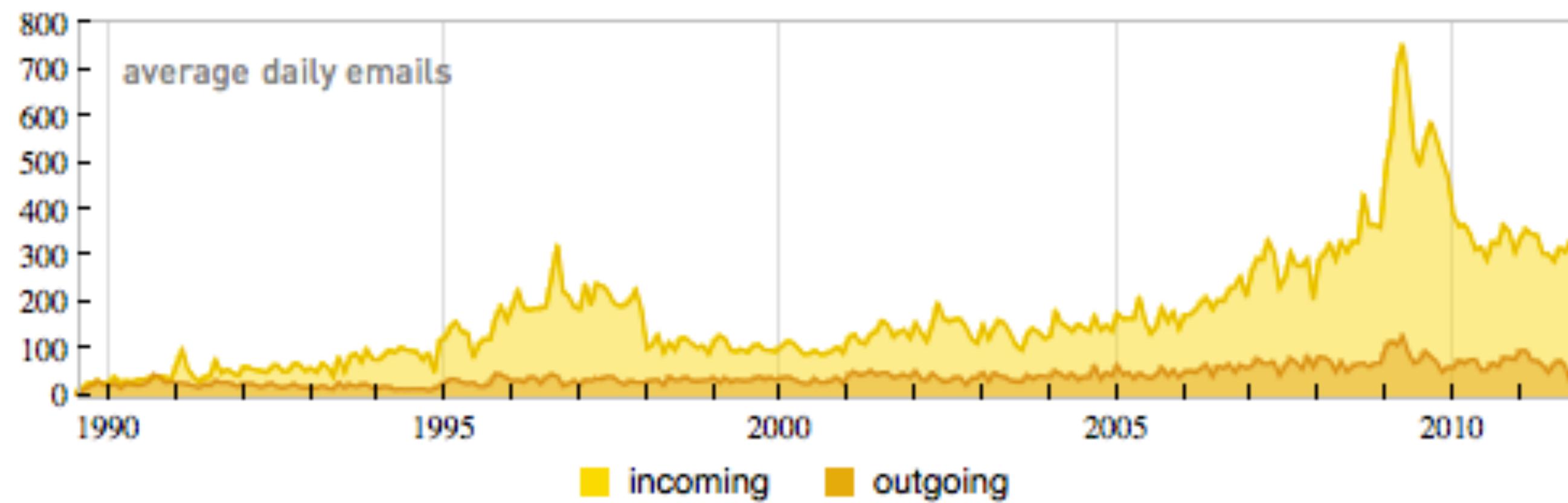


Número de **diferentes destinatários** por dia



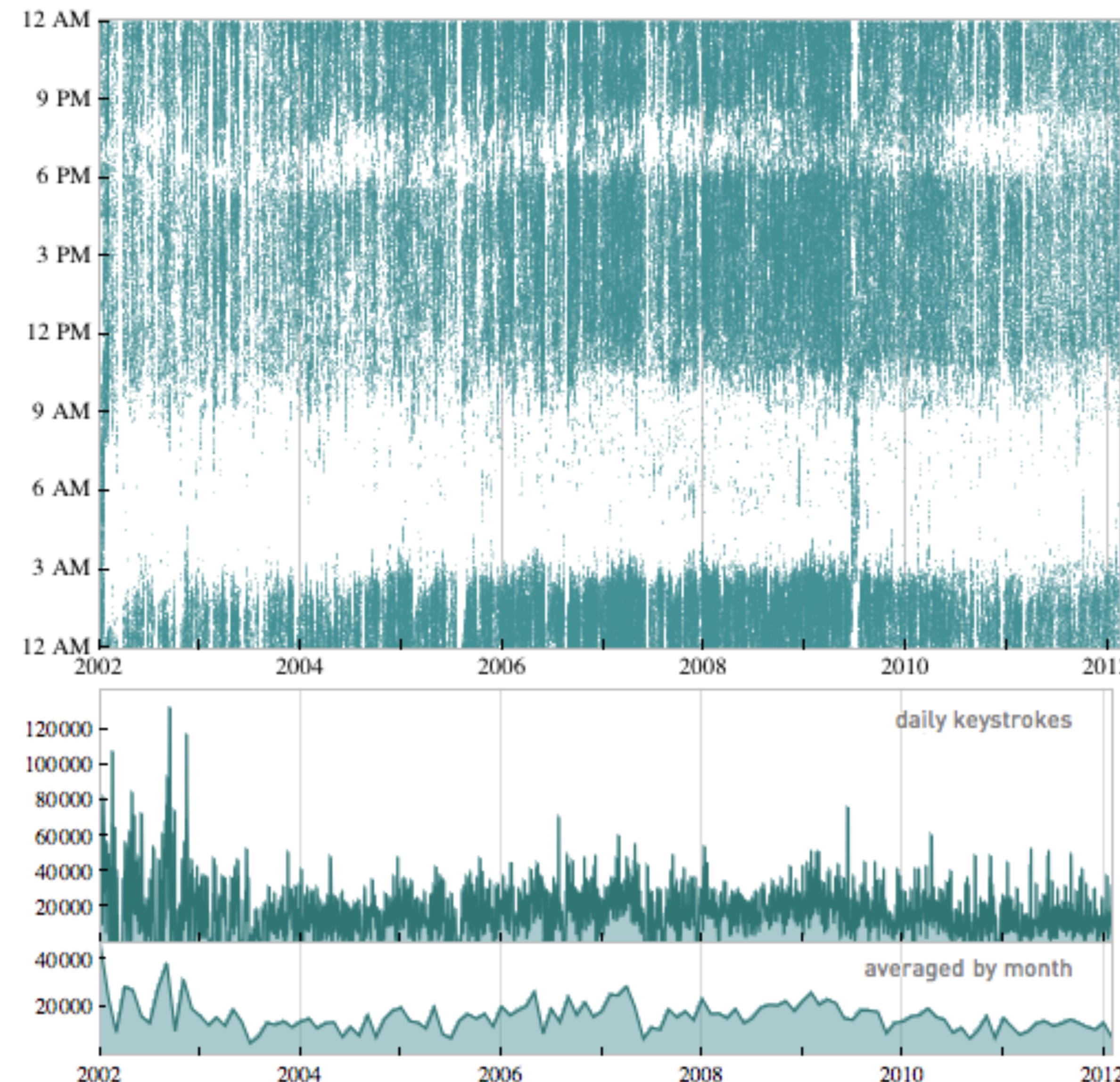
Picos coincidem com inícios de projetos, quando
esta normalmente envolvido com um número
maior de pessoas

Número de **diferentes remetentes** por dia

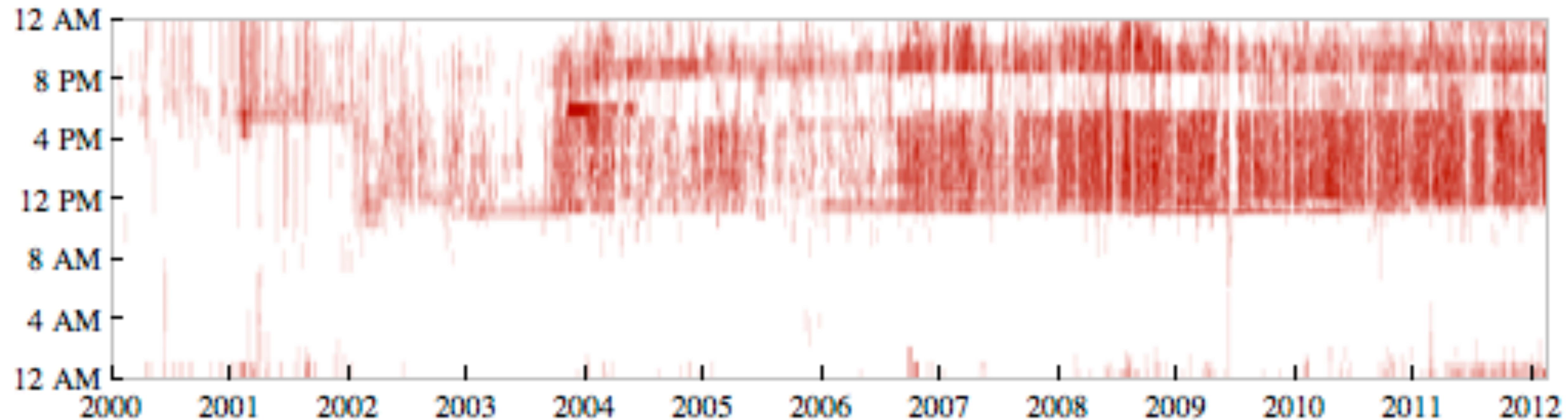


Picos coincidem com lançamentos de
projetos: Mathematica 3 e Wolfram|
Alpha

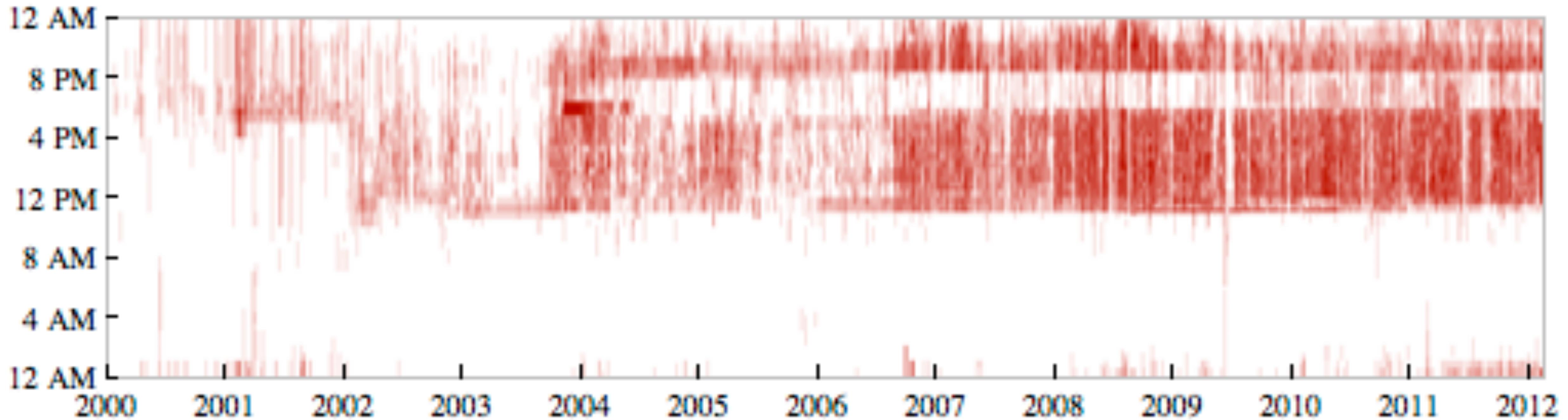
Cerca de 100.000.000 de **teclas** que digitou



Eventos na agenda



Eventos na agenda

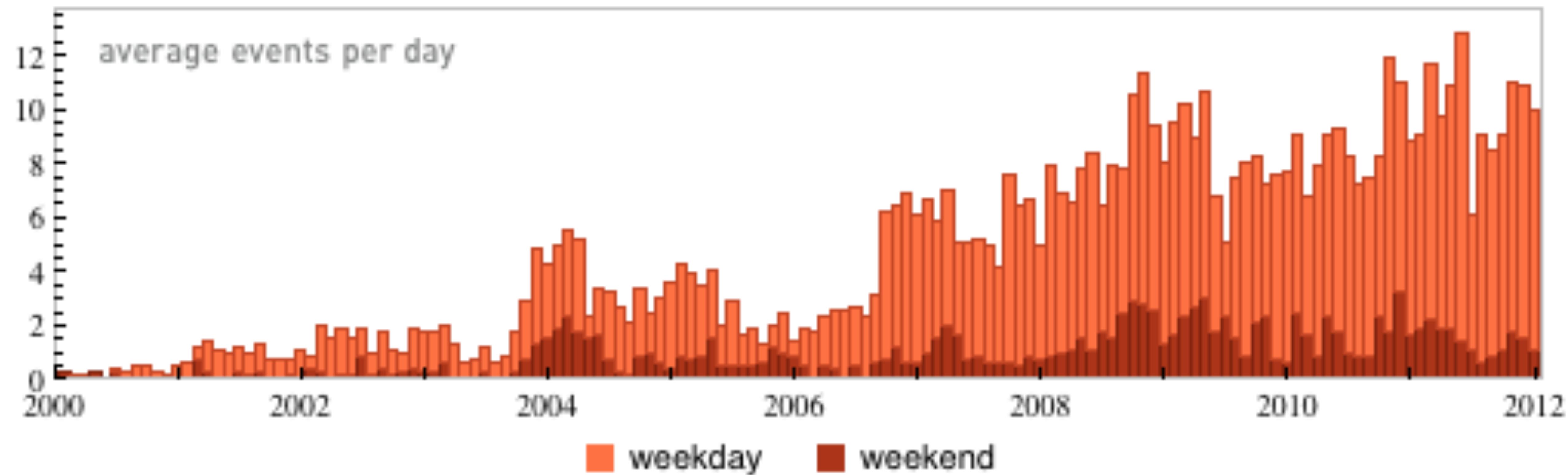


Antes de 2002, trabalhava sozinho na escrita do NKS

Depois o engajamento na empresa fez com que tivesse mais compromissos

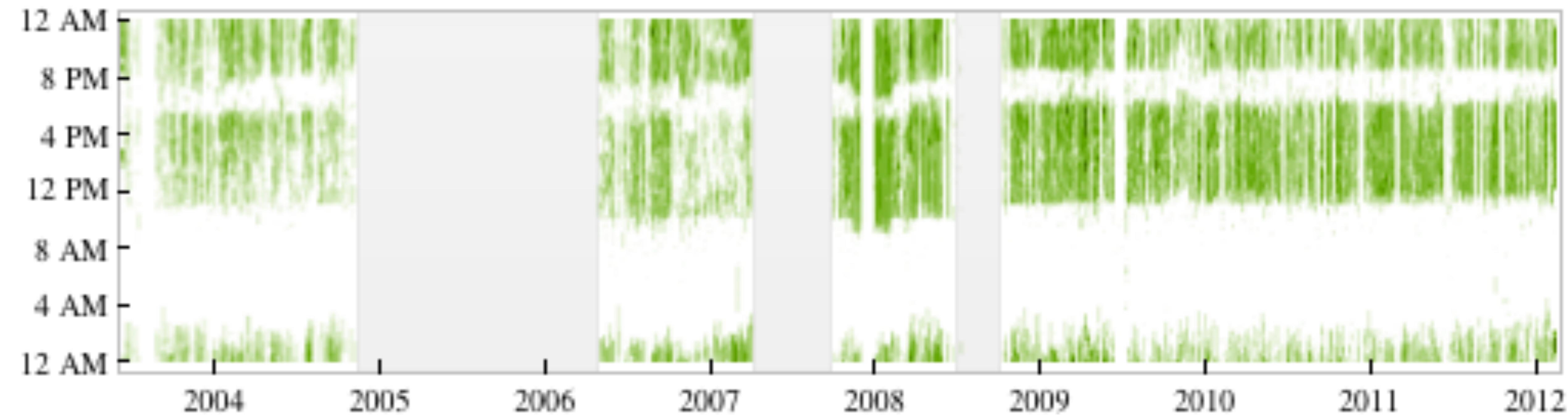
Observe a faixa do período de jantares

Distribuição do número de **eventos** por dia



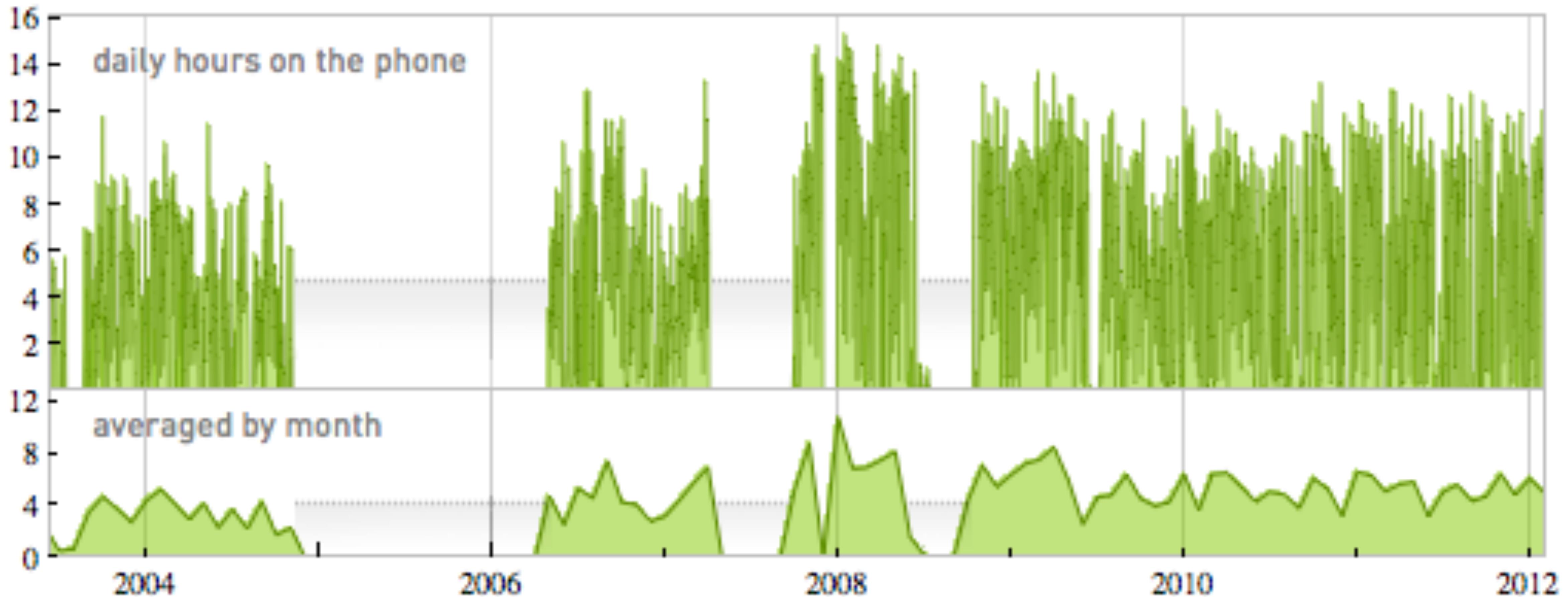
Mudança de paradigma: começou a trabalhar melhor “em público”. Antes ele era um CEO “remoto”

Chamadas telefônicas ao longo dos dias

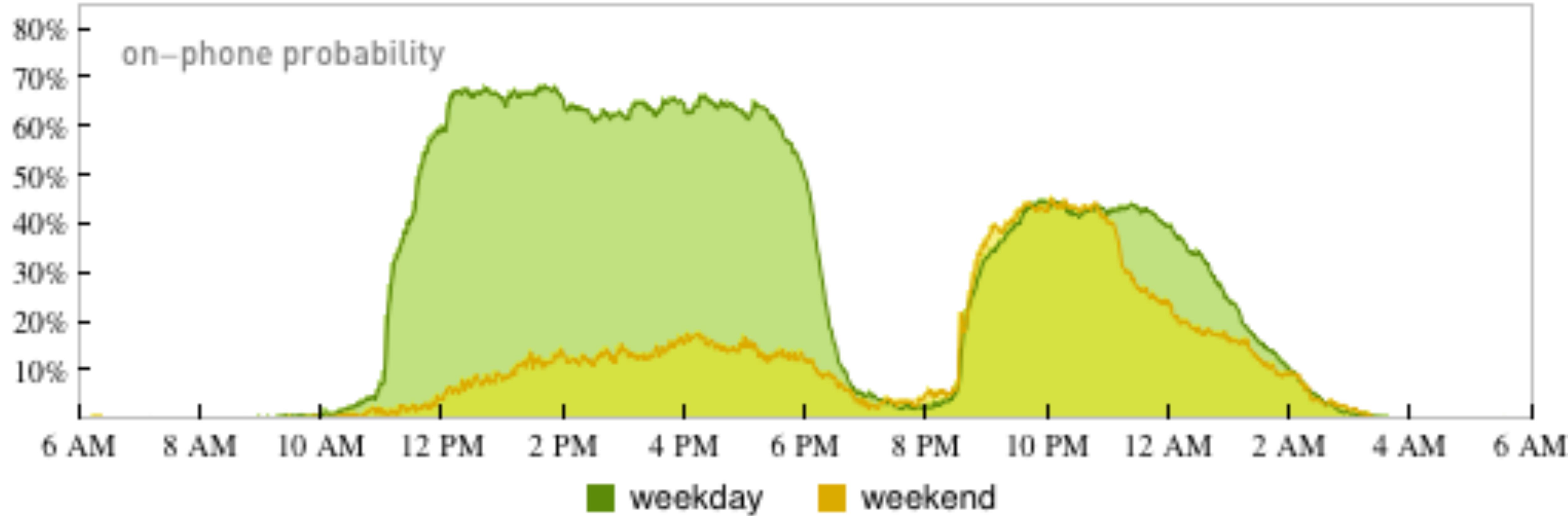


Faixas em cinza indicam ausência de dados

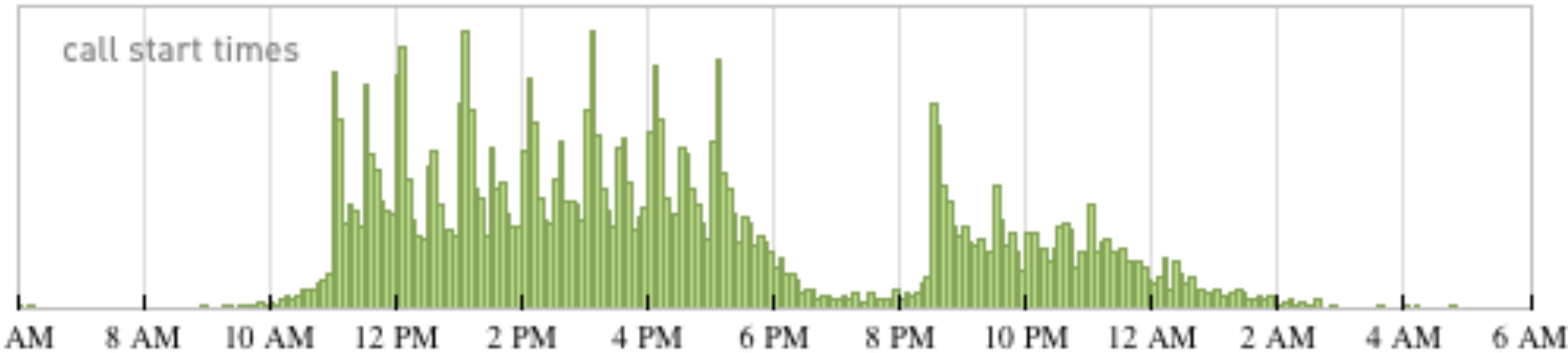
Sim, ele passa muitas horas do dia ao telefone!!!



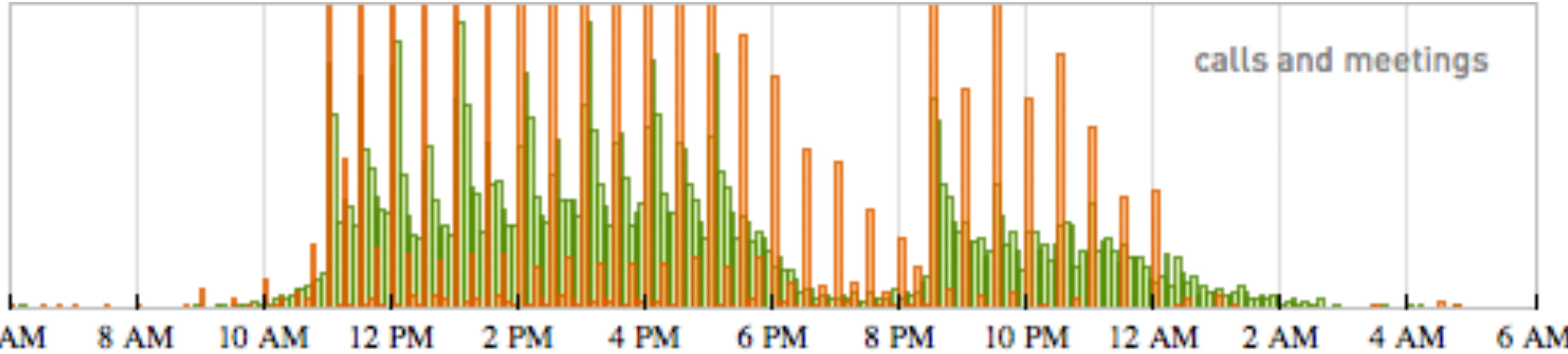
Veja a probabilidade de ele estar ao telefone nas
diversas horas do dia!



Número de chamadas que se iniciam ao longo do dia



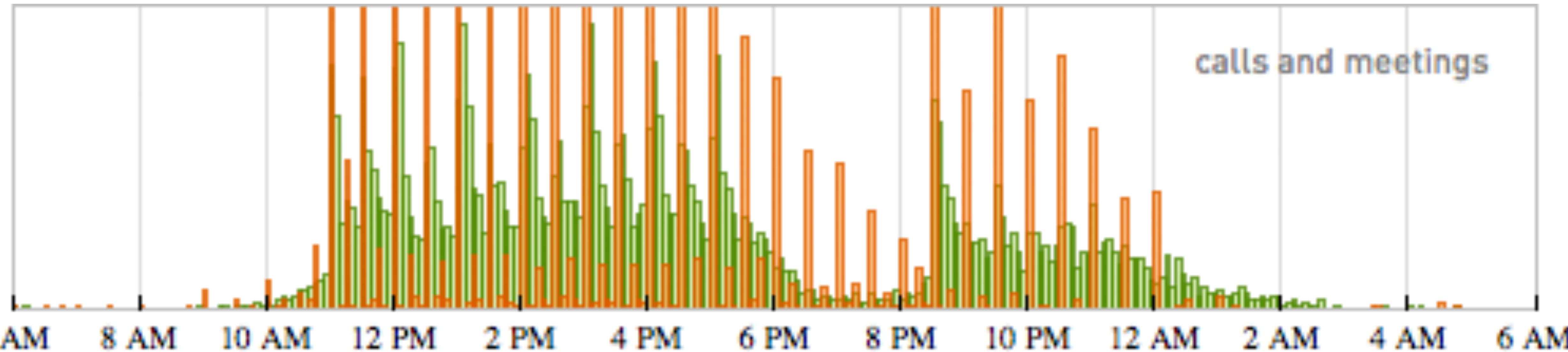
Número de chamadas que se iniciam ao longo do dia



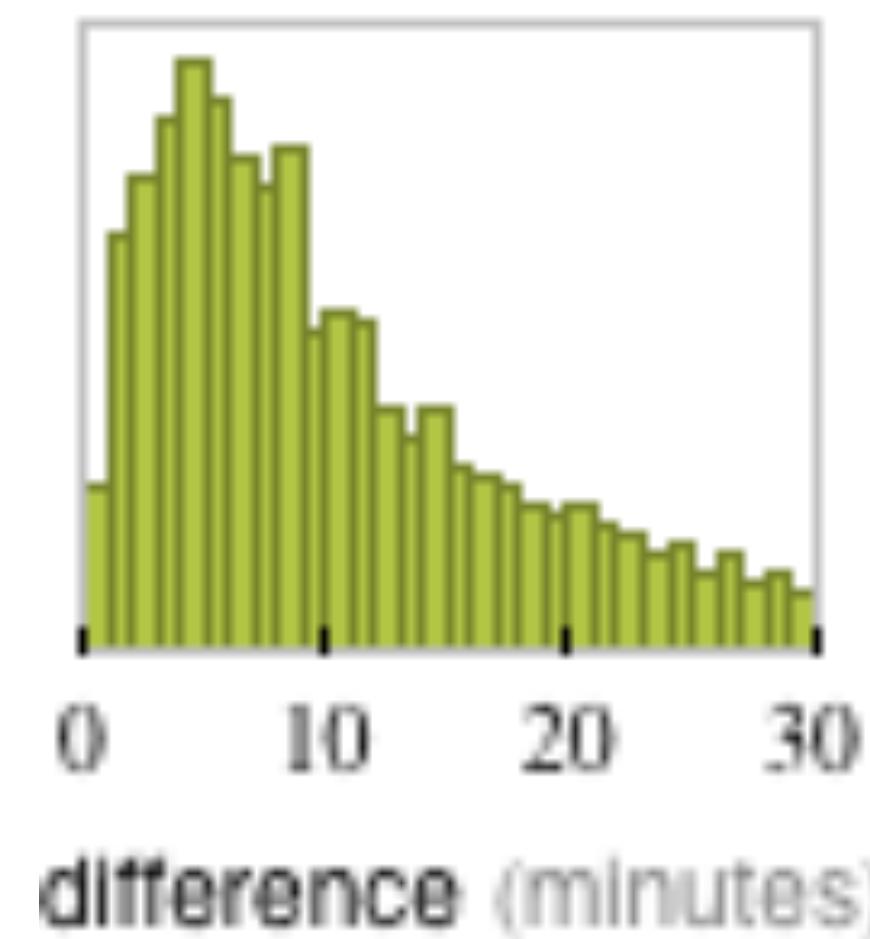
Número de chamadas e eventos na agenda

Há alguma correlação?

Por que?

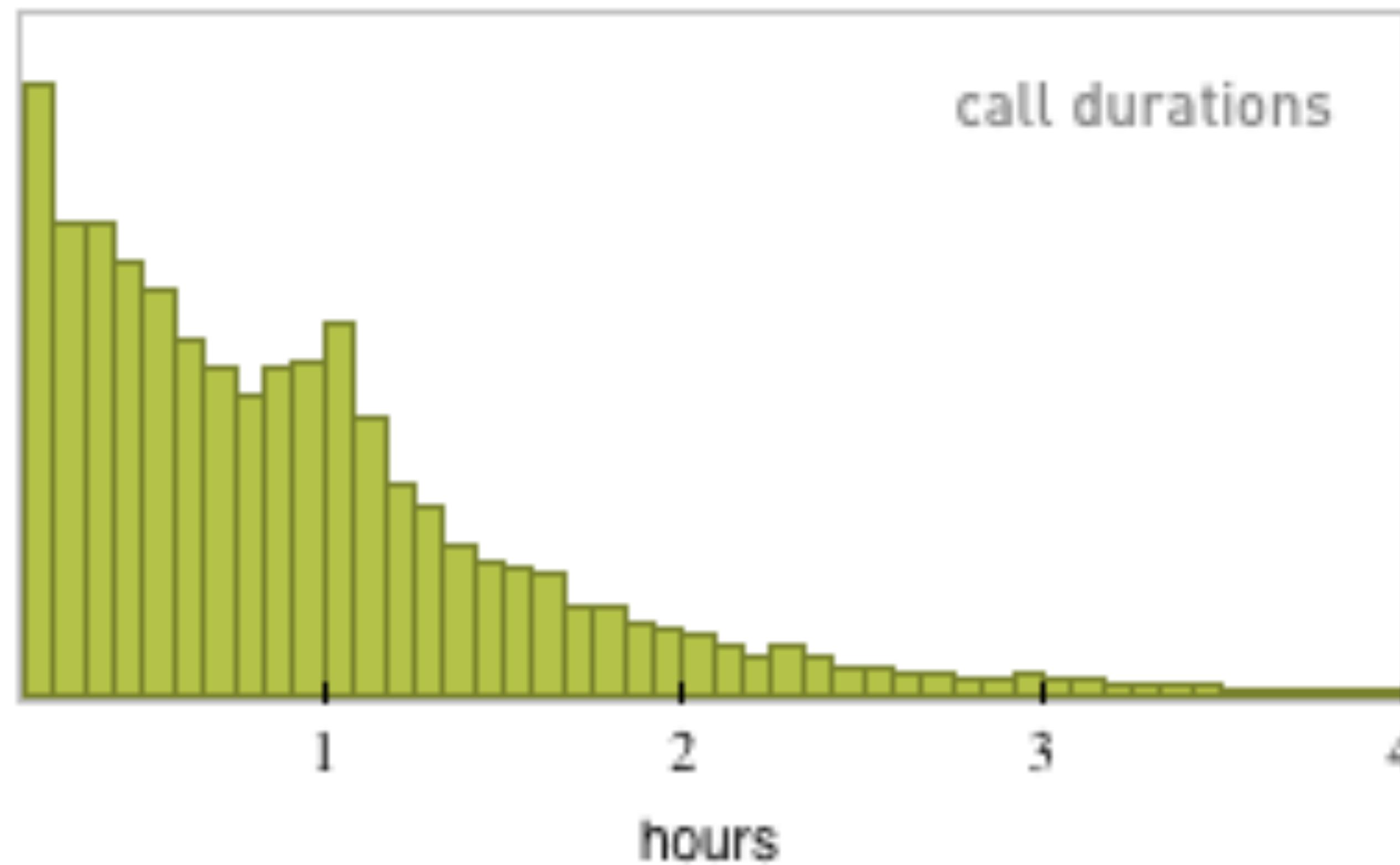


Reuniões começam com 2 minutos de **atraso**

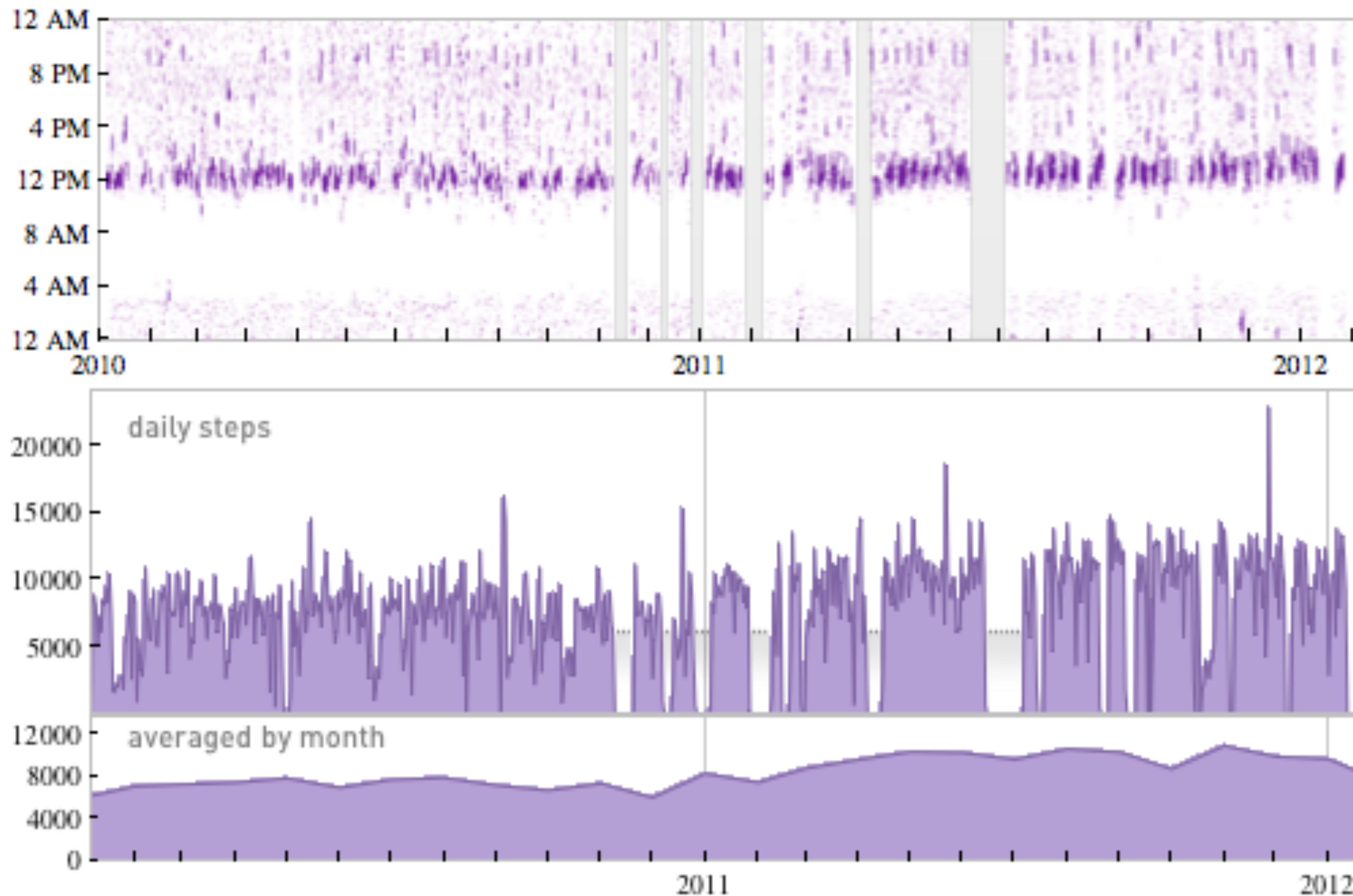


Distribuição das durações das chamadas

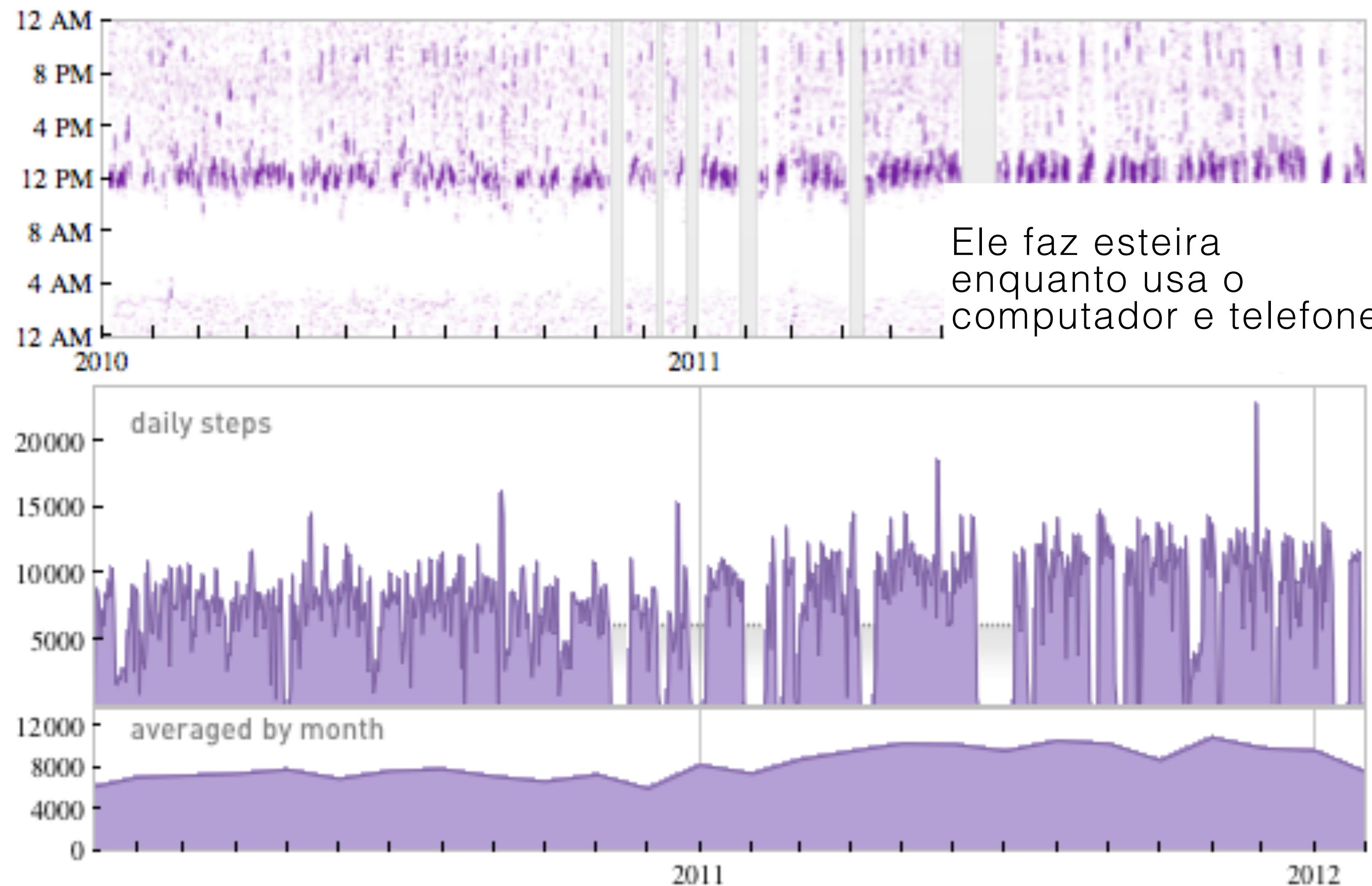
Por que há um pico em 1 hora?

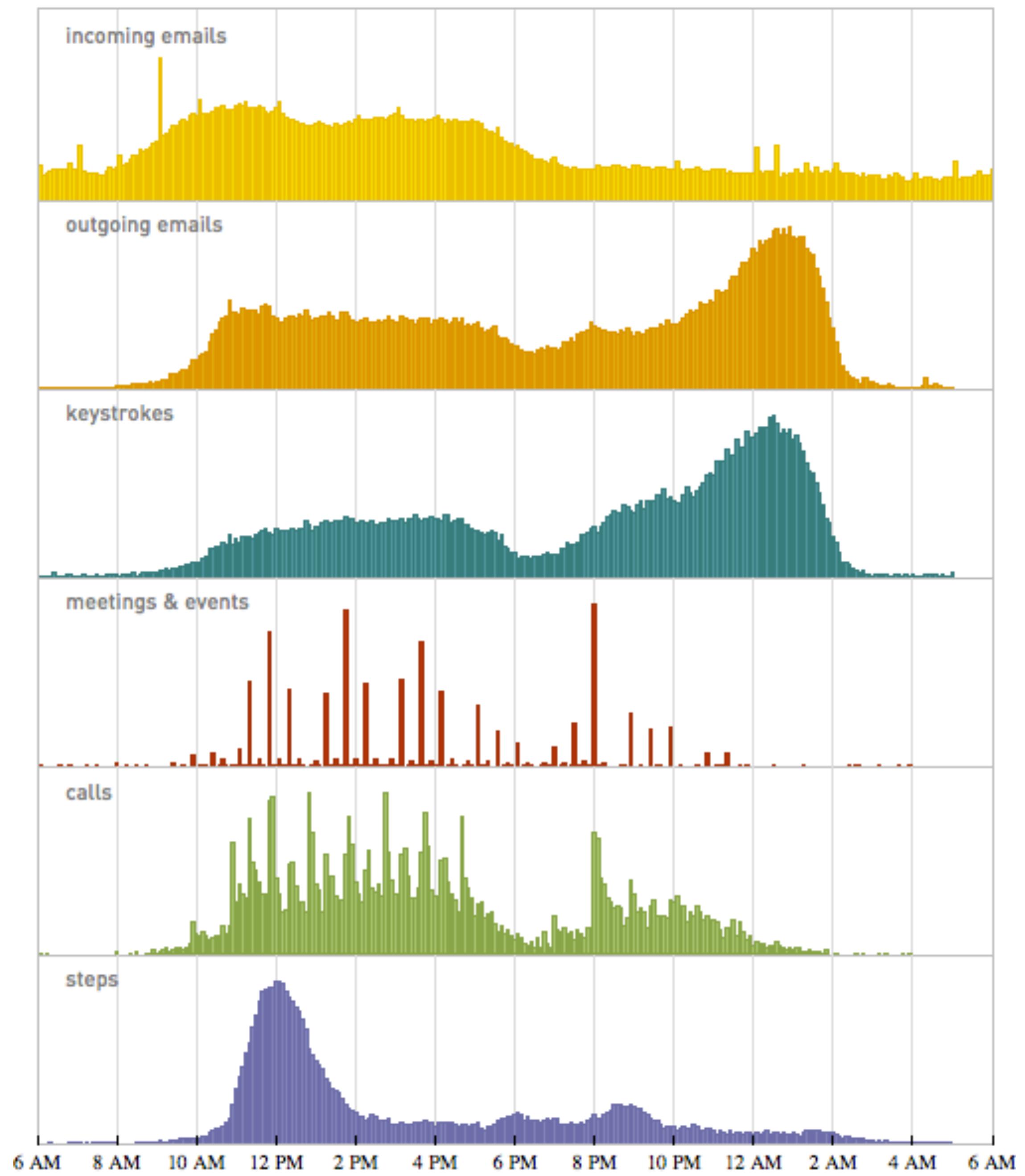


Ele contou até quantos **passos** deu



Ele contou até quantos **passos** deu

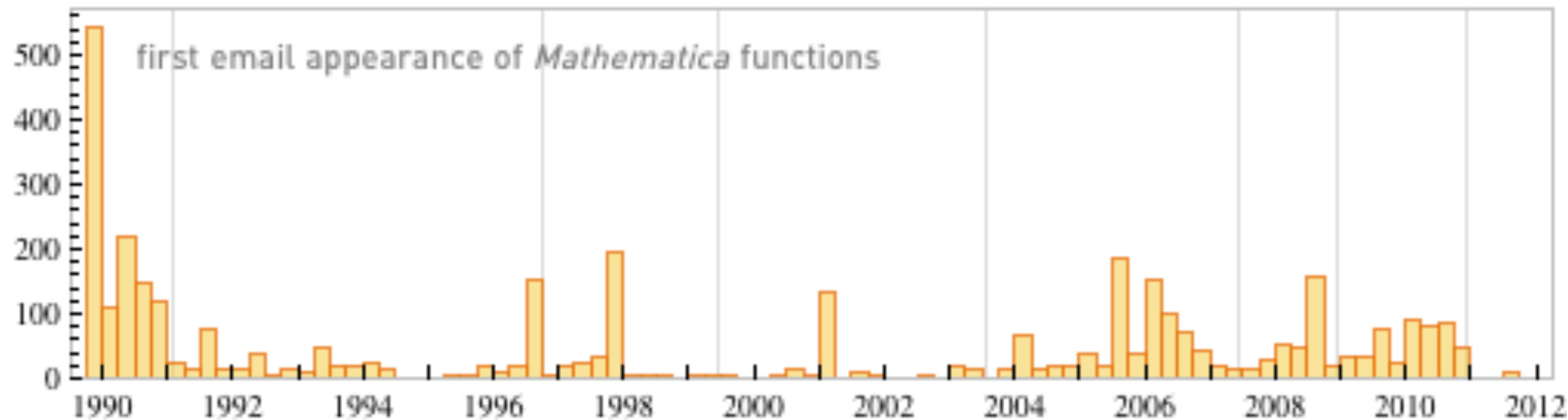




É possível correlacionar os **hábitos** com os **momentos criativos?**

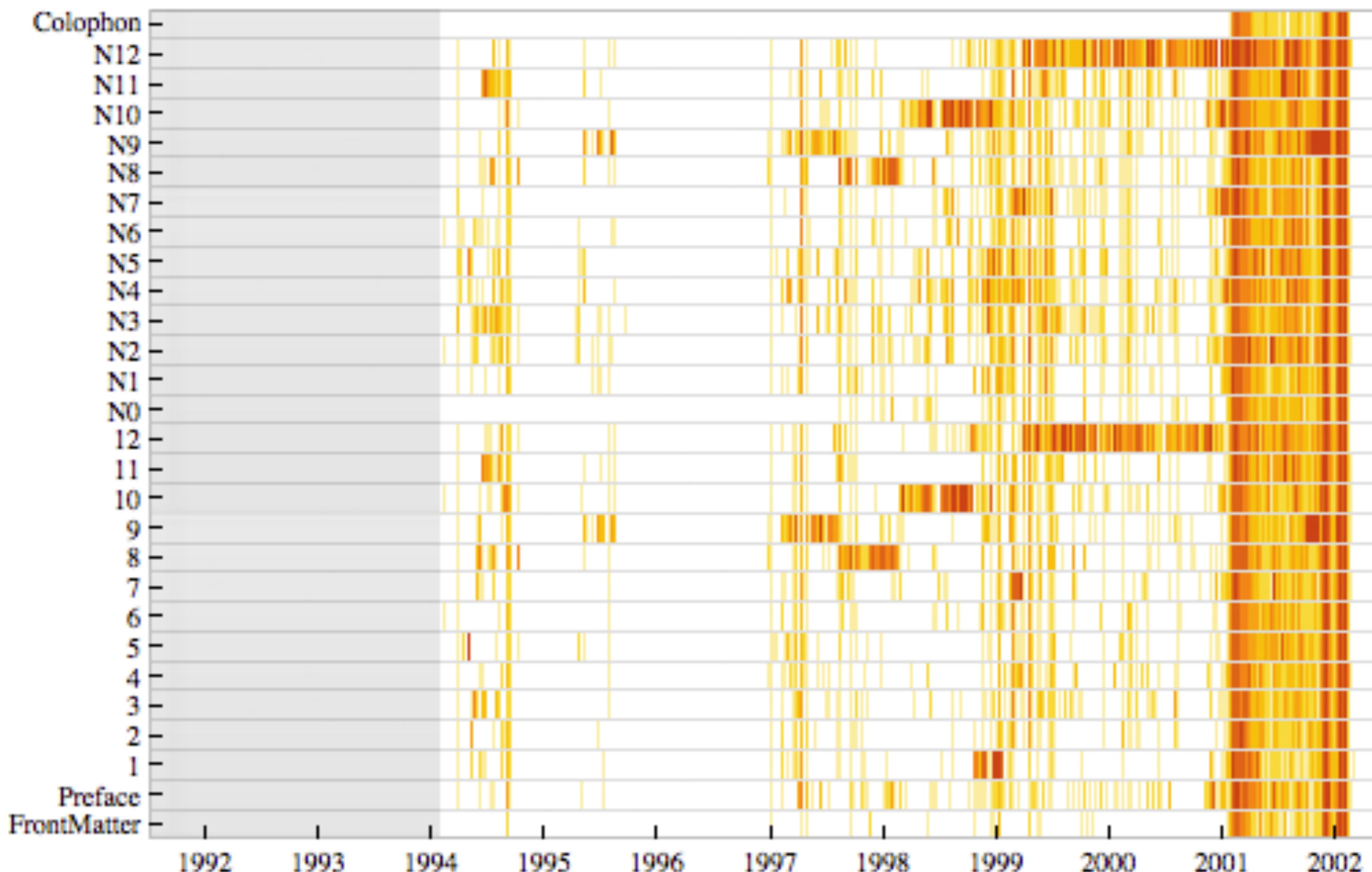
Analisou as “**novas**” **palavras** que apareciam em seus e-mails

Cerca de 33.000 diferentes palavras na última década



Tentativa de visualização do progresso criativo

Modificações no texto de NKS



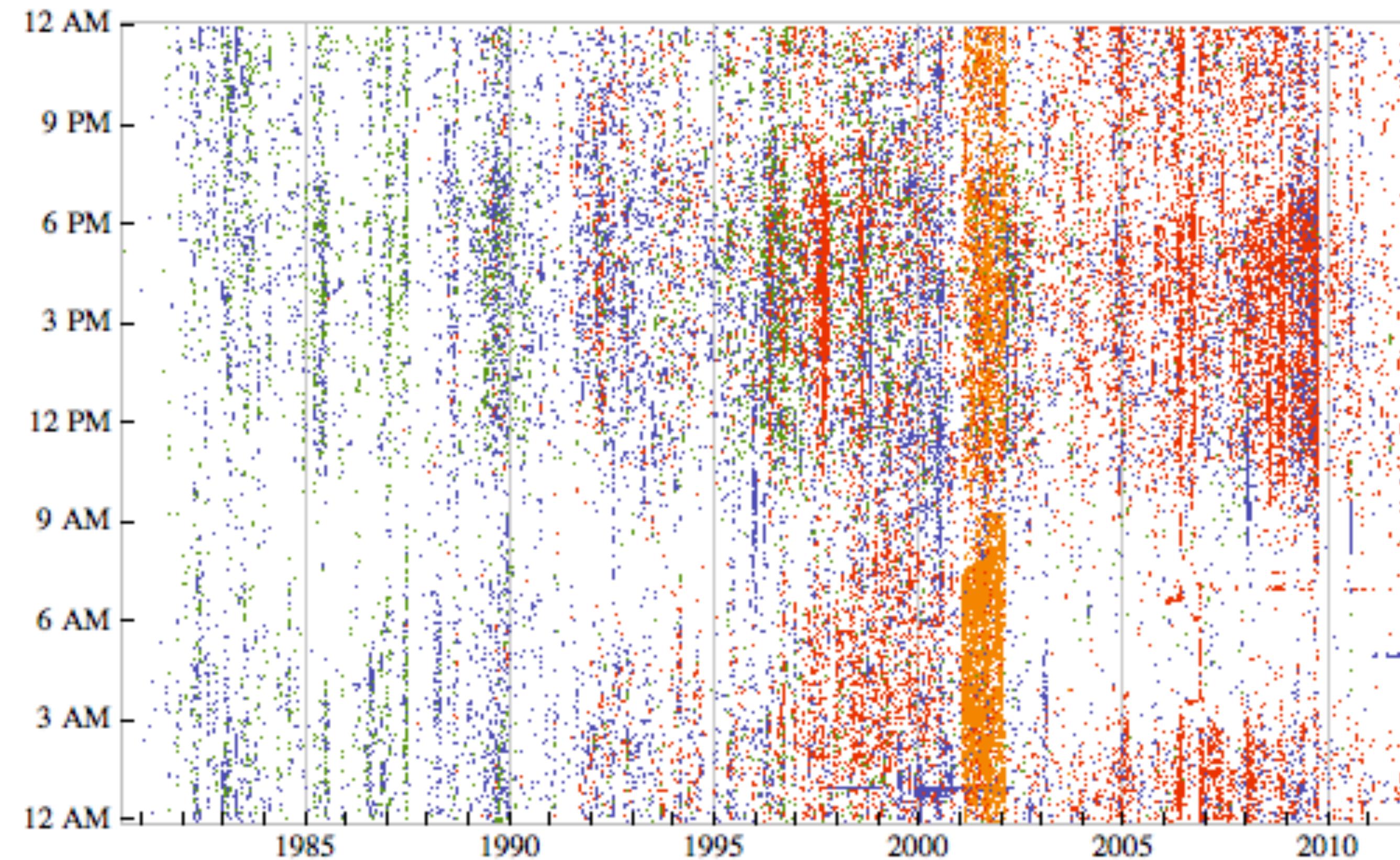
Modificações de todos os seus **arquivos**

Azul: texto

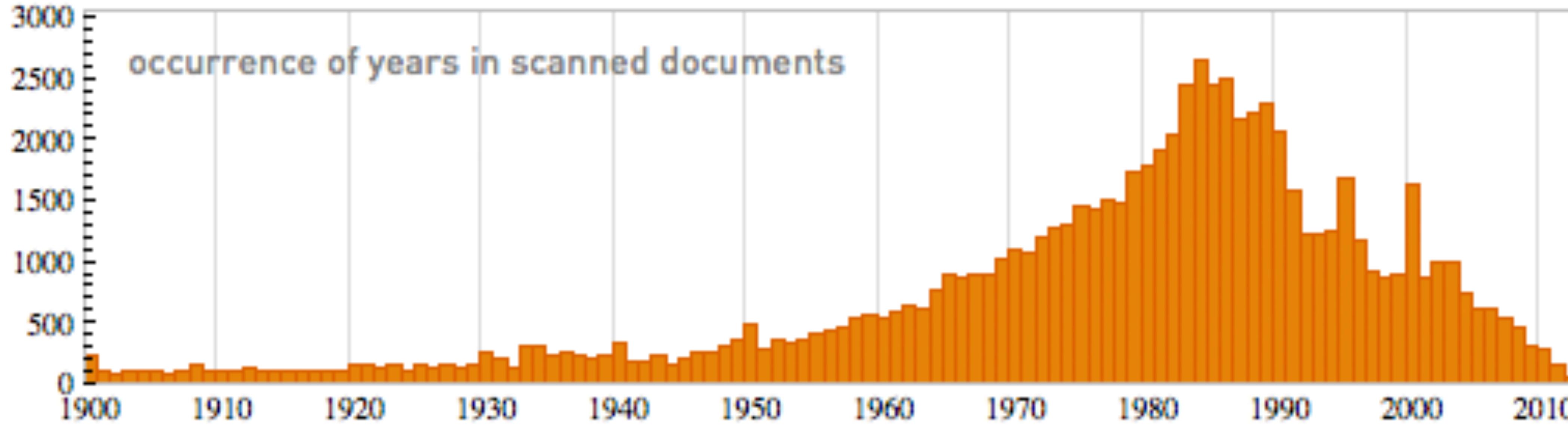
Verde: códigos em C

Vermelho: Mathematica

Laranja: leiautes



Ocorrência de **anos em documentos em papel, digitalizados**



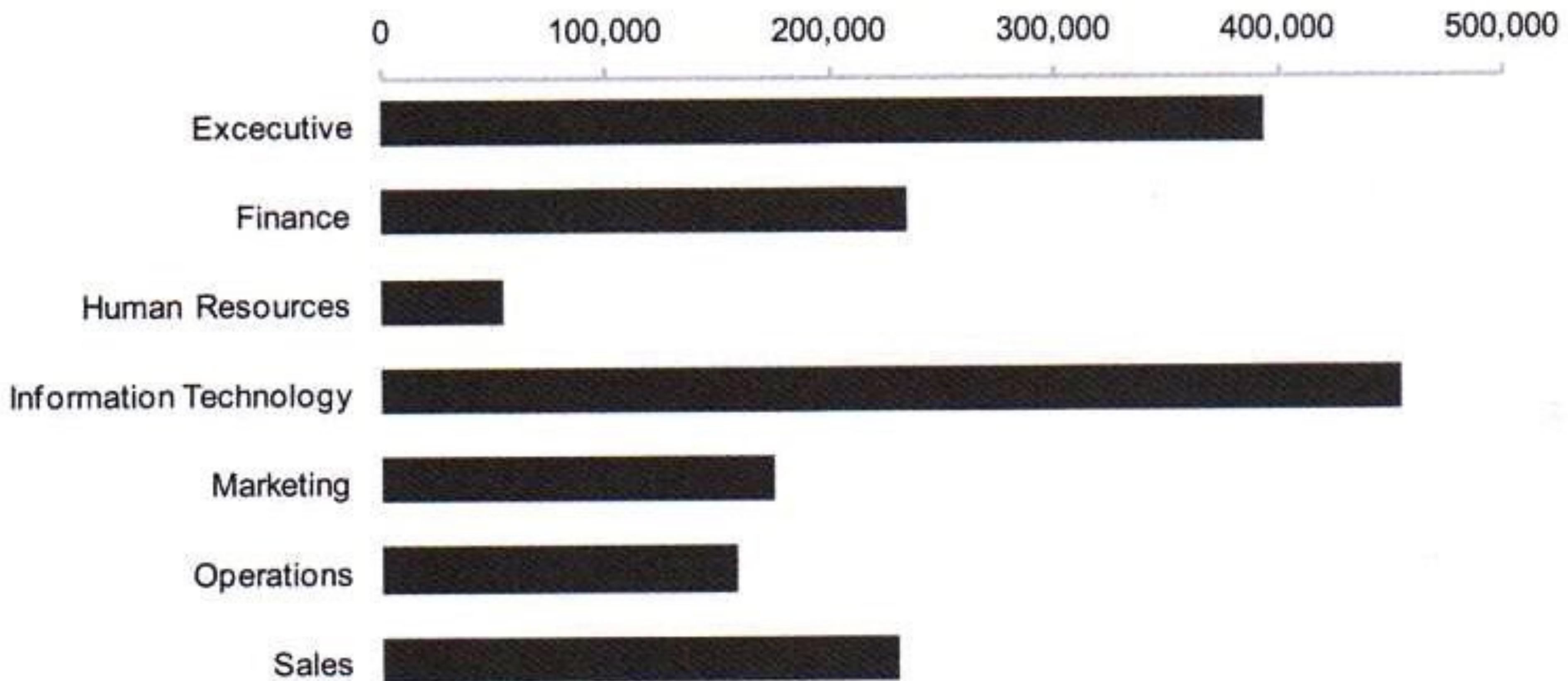
A propósito, ele ainda tem muitos dados não analisados: seu genoma completo, rastreamento via GPS, monitoramento de mudanças de sala, registros médicos...

<http://blog.stephenwolfram.com/2012/03/the-personal-analytics-of-my-life/>

Ranking e Parte-todo

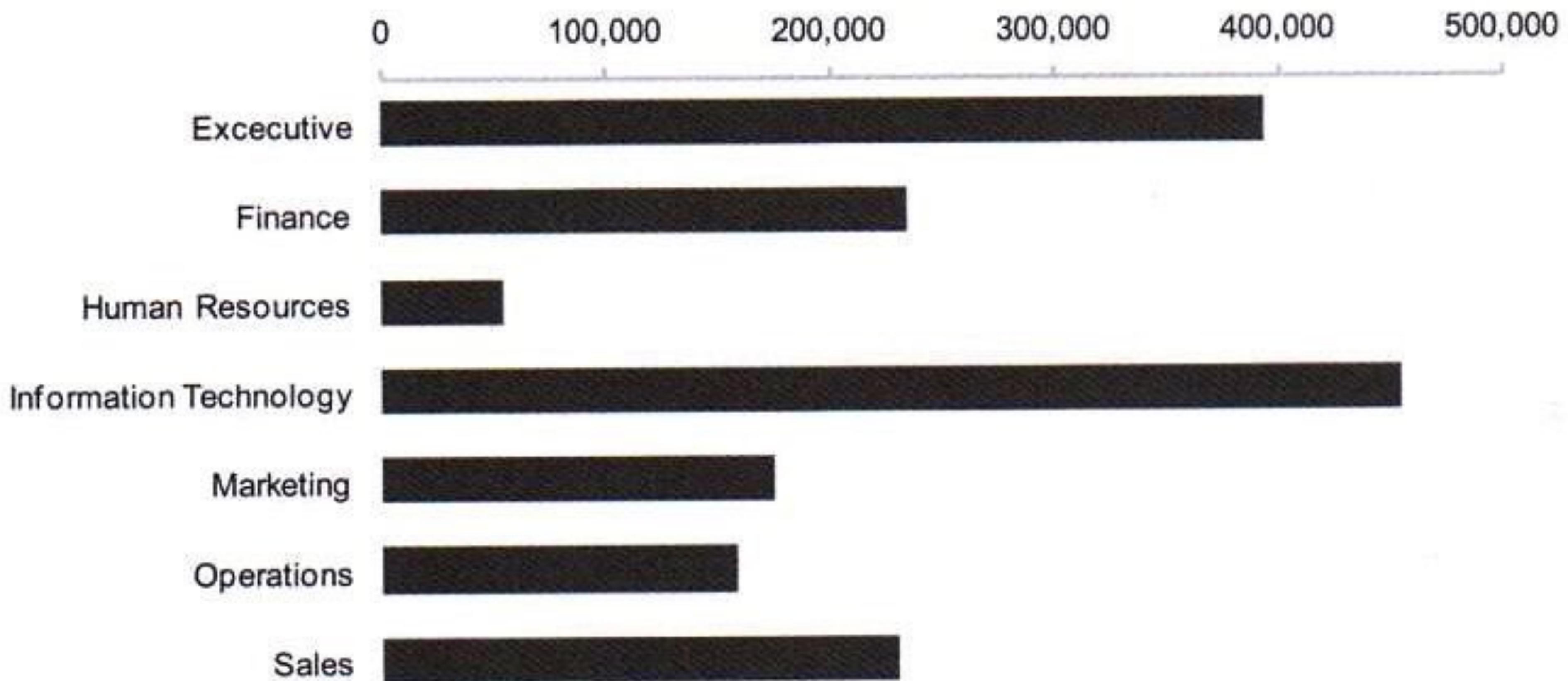
Profa. Dra. Raquel Minardi
Departamento de Ciência da Computação
Universidade Federal de Minas Gerais

**January 2008 Expenses
(USD)**



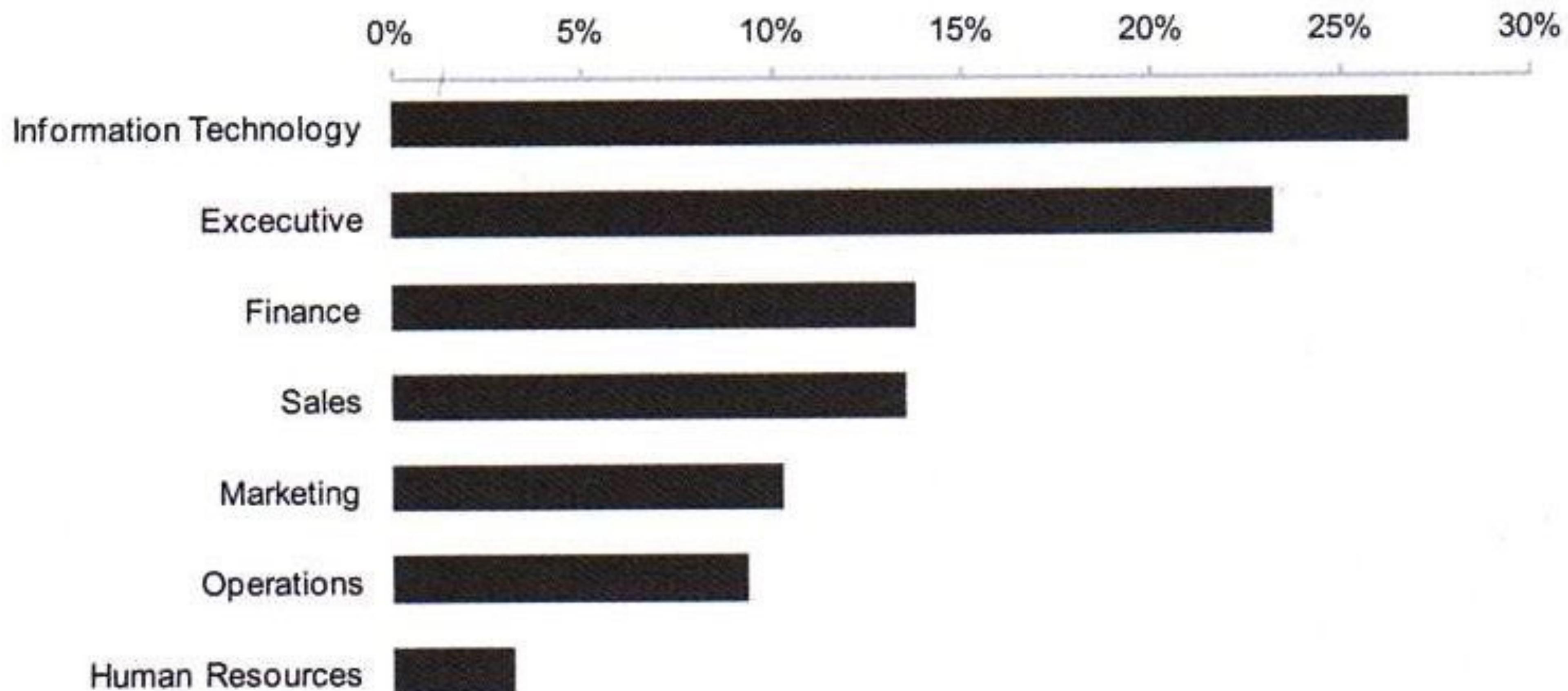
As despesas dos vários departamentos de uma empresa representam partes de um todo

January 2008 Expenses
(USD)



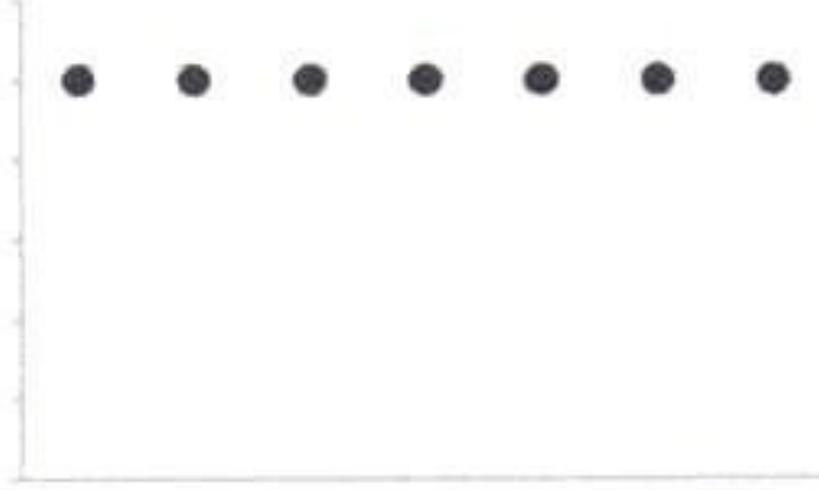
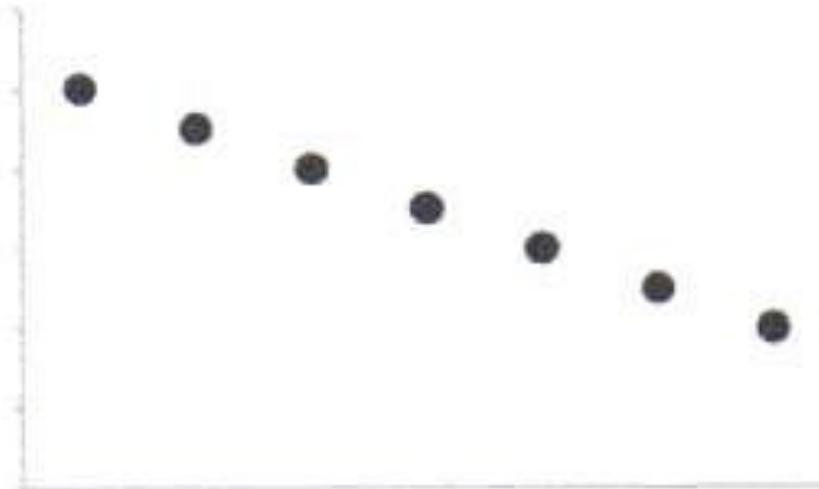
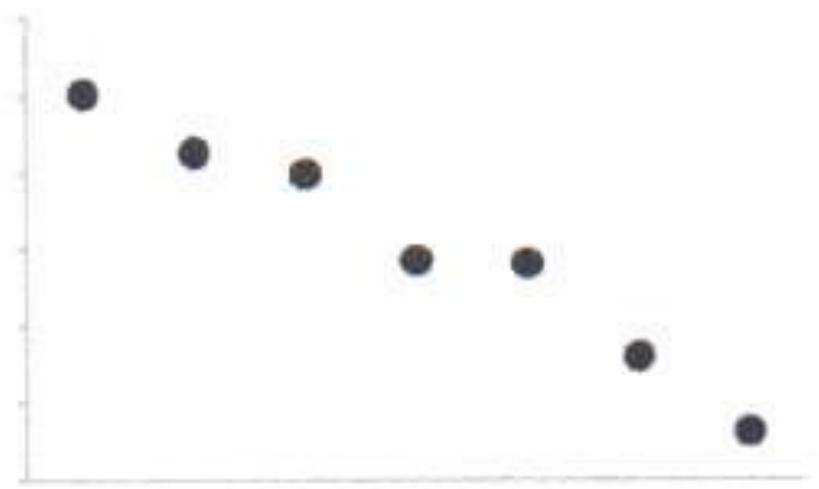
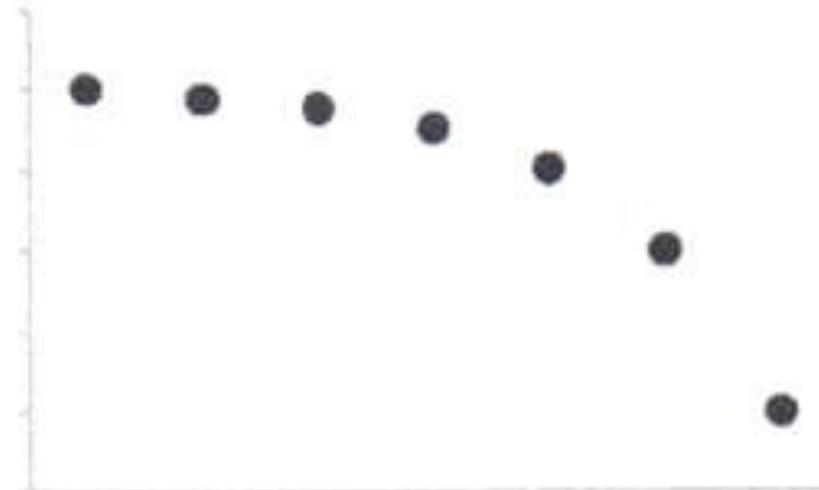
A **ordenação alfabética** e a **unidade** usada para medição
dificultam a visualização deste tipo de relacionamento

January 2008 Expenses



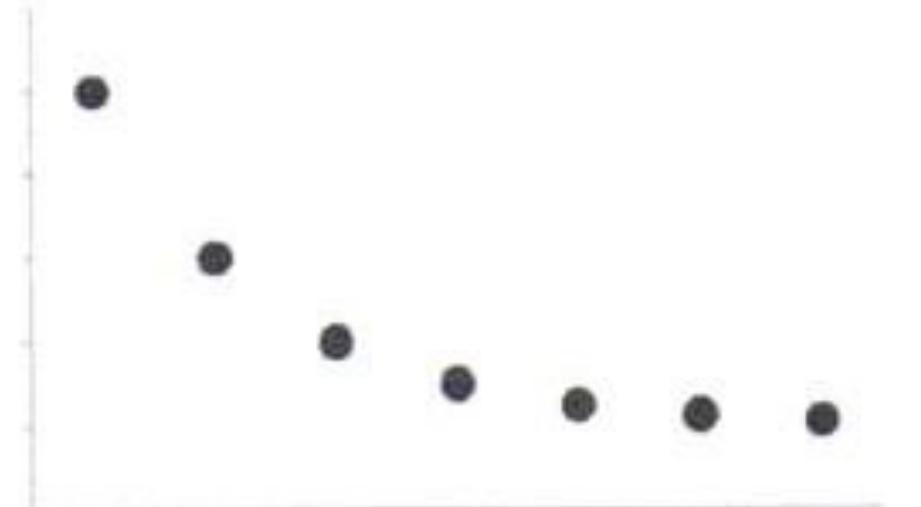
A **ordenação decrescente** e o uso de percentuais evidenciam o **relacionamento parte-todo**

PADRÕES ANALÍTICOS

Pattern	Description	Visual Example
Uniform	All values are roughly the same.	
Uniformly different	Differences from one value to the next decrease by roughly the same amount.	
Non-uniformly different	Differences from one value to the next vary significantly.	
Increasingly different	Differences from one value to the next increase.	

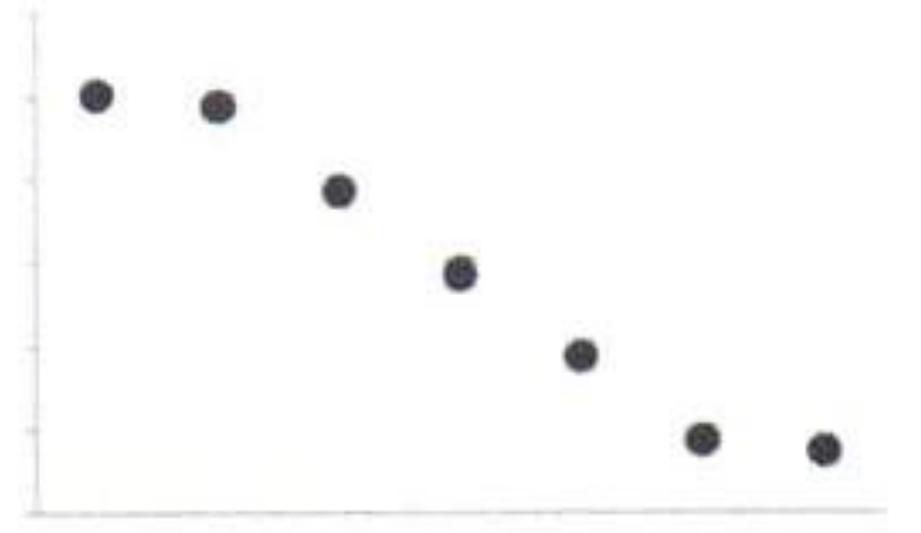
Decreasingly different

Differences from one value to the next decrease.



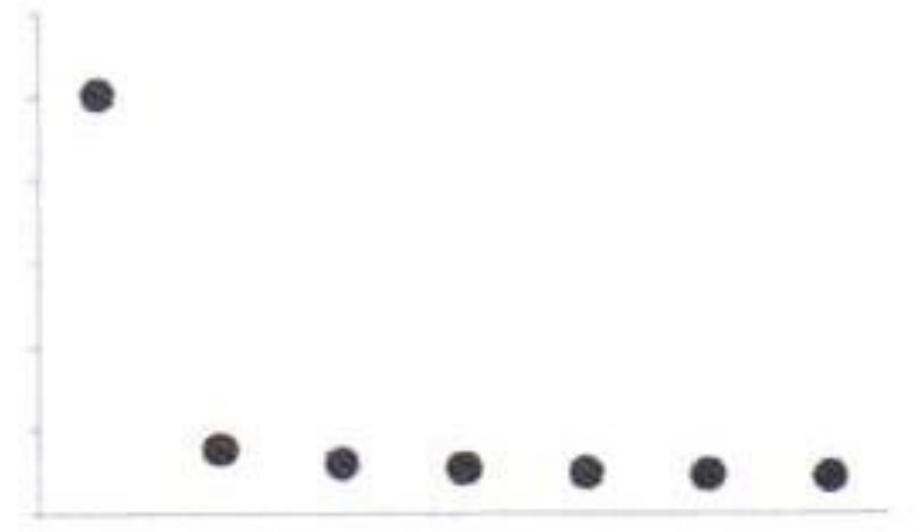
Alternating differences

Differences from one value to the next begin small and then shift to large and finally shift back again to small.



Exceptional

One or more values are extraordinarily different from the rest.

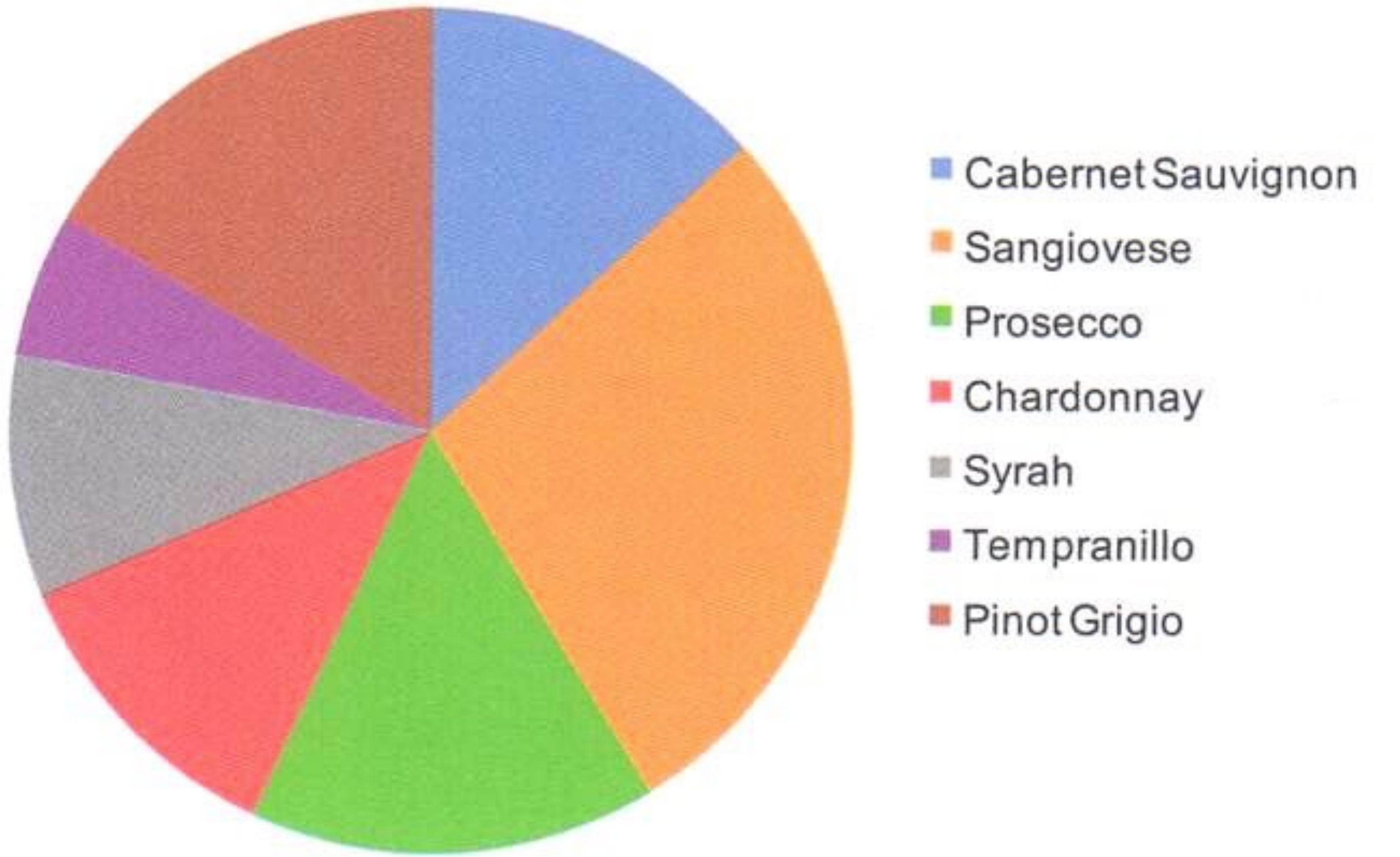


REPRESENTAÇÕES VISUAIS

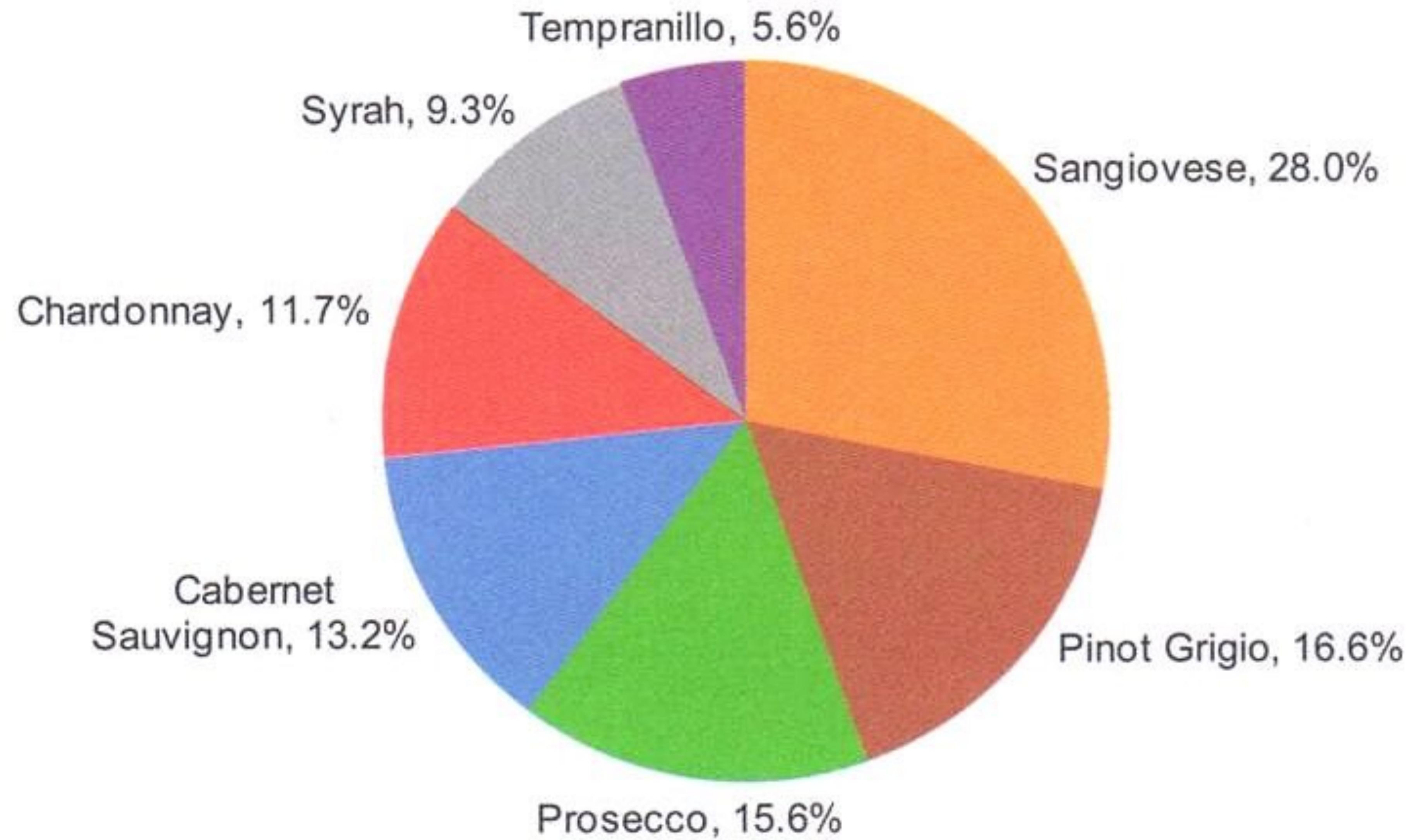


Quanto trabalho para ordenar as fatias?

E para obter a diferença entre os percentuais representados?



O uso de legendas em separado é especialmente difícil pois força o deslocamento dos olhos entre gráfico e legenda

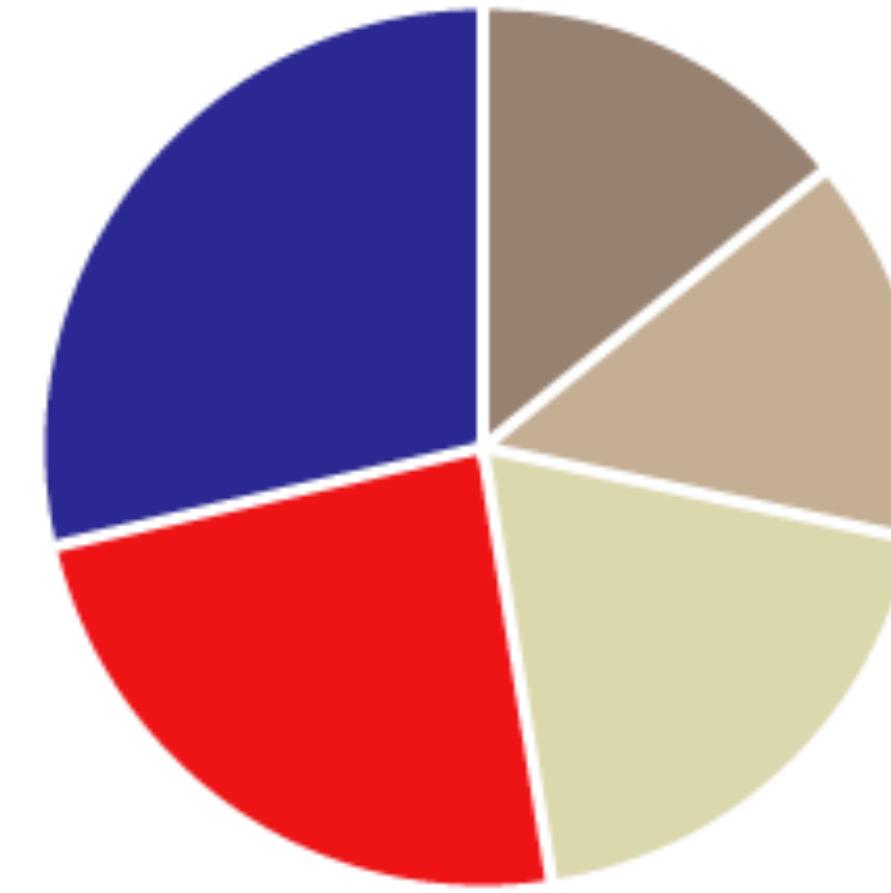


Melhor assim?

Wine	Percent
Sangiovese	28.0%
Pinot Grigio	16.6%
Prosecco	15.6%
Cabernet Sauvignon	13.2%
Chardonnay	11.7%
Syrah	9.3%
Tempranillo	5.6%
Total	100.0%

Ou assim?

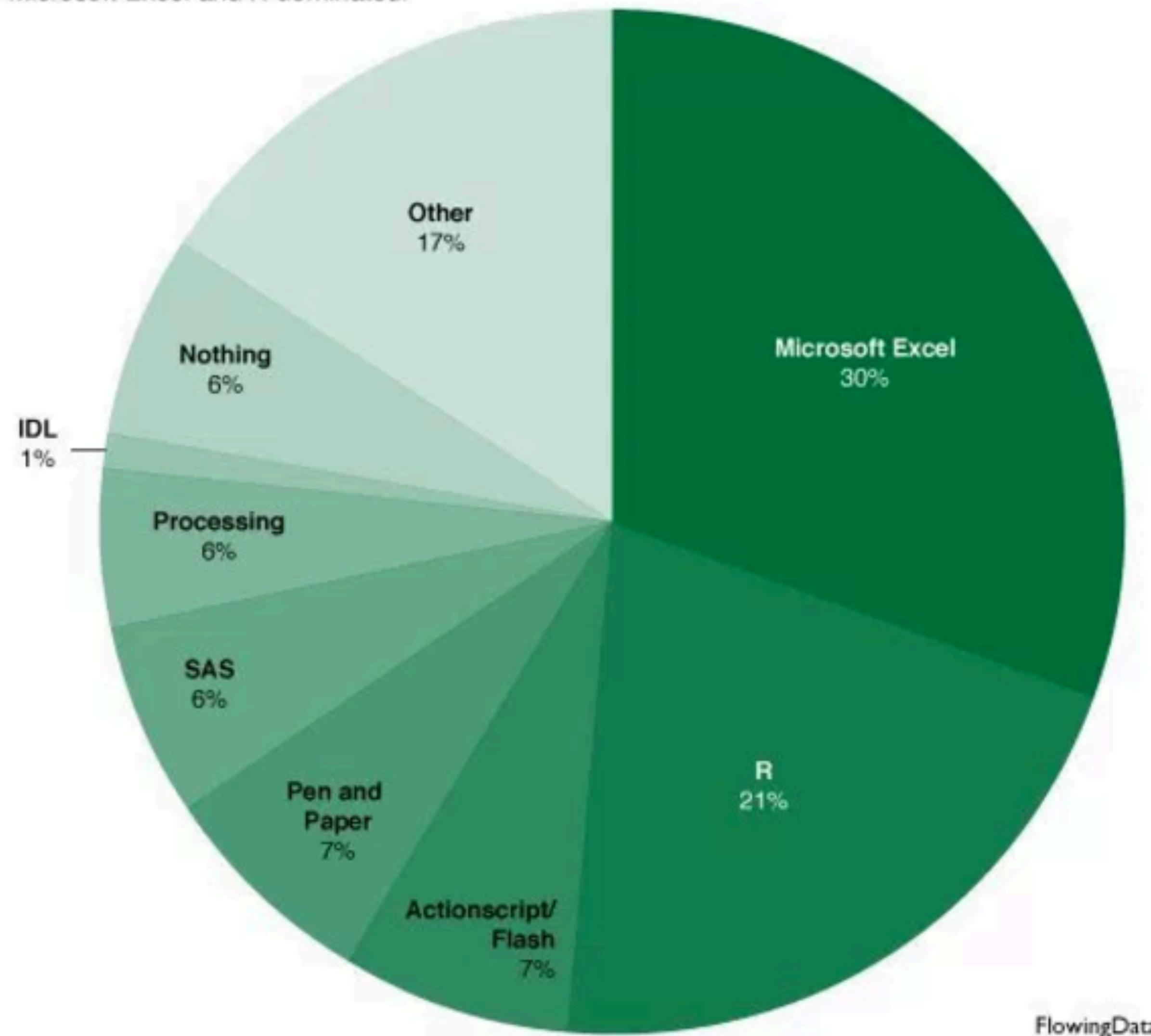
PIE CHART (GRÁFICO DE PIZZA)



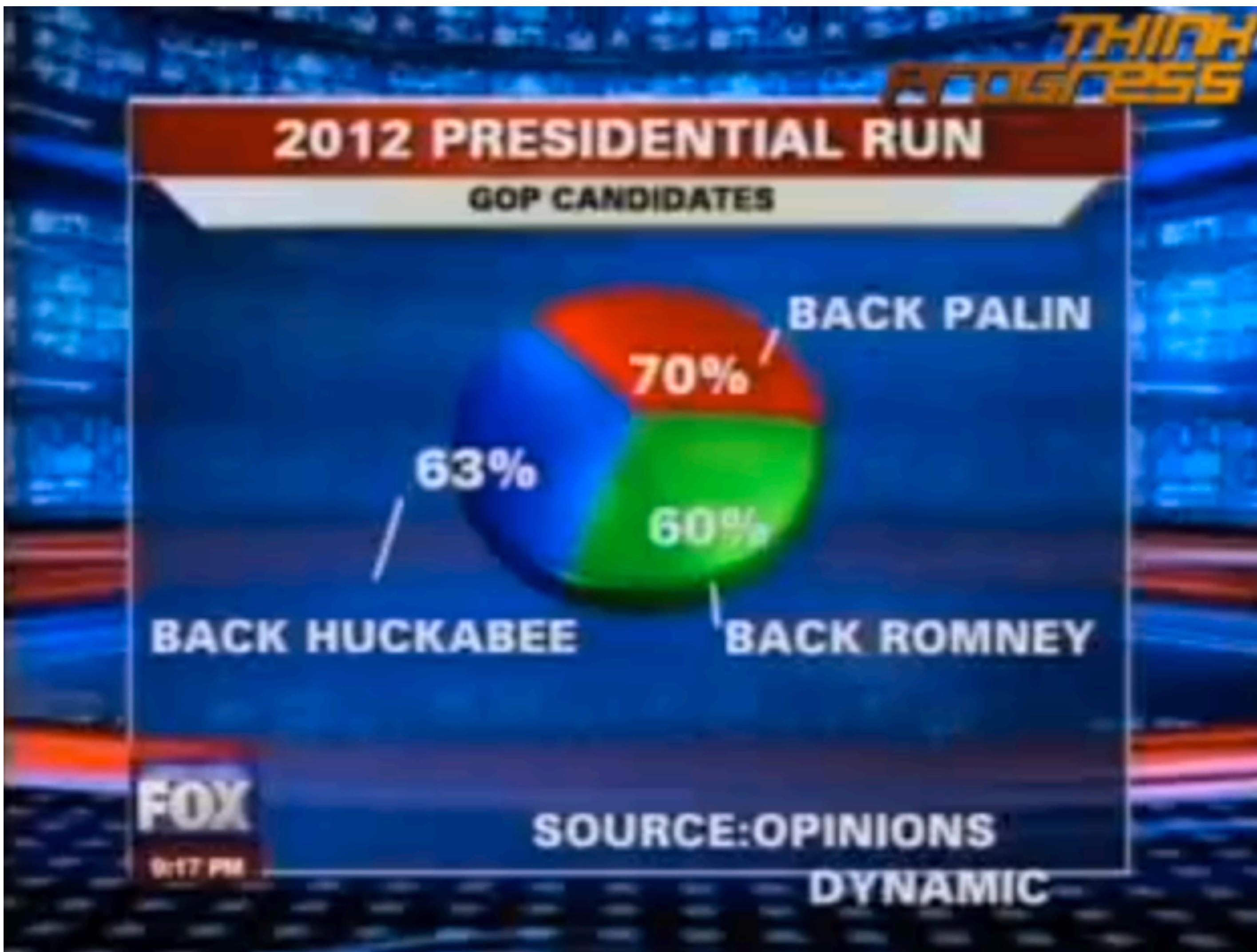
- Não conseguimos estimar precisamente os valores de áreas
- Não podem ser usados para gráficos com muitas séries
- Ocupam muito espaço (necessidade de legenda)
- Não são eficazes para uso comparativo com outros gráficos de pizza

Visualization and Analysis Tools of Choice

A recent FlowingData poll asked readers what they used to visualize and analyze data. Microsoft Excel and R dominated.



FlowingData



Save the Pies for Dessert

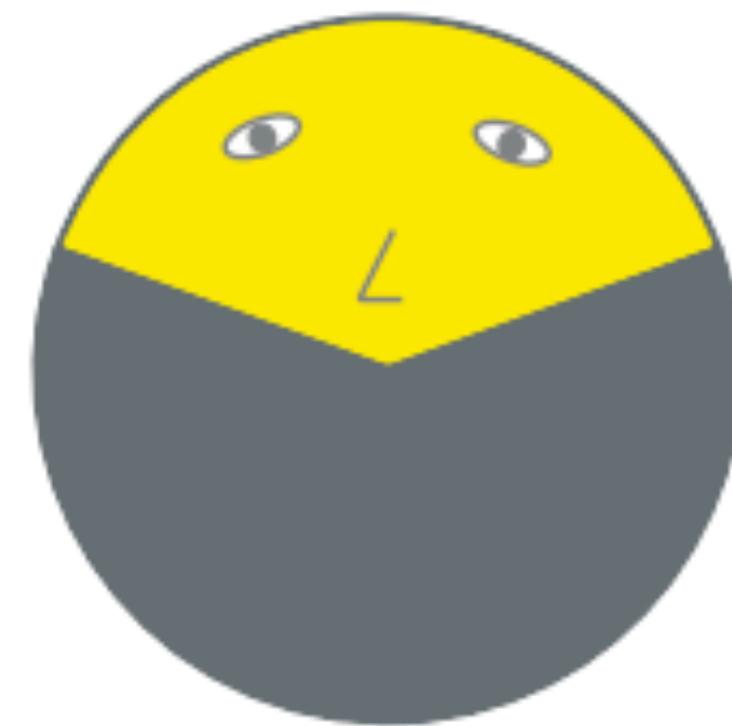
Stephen Few, Perceptual Edge

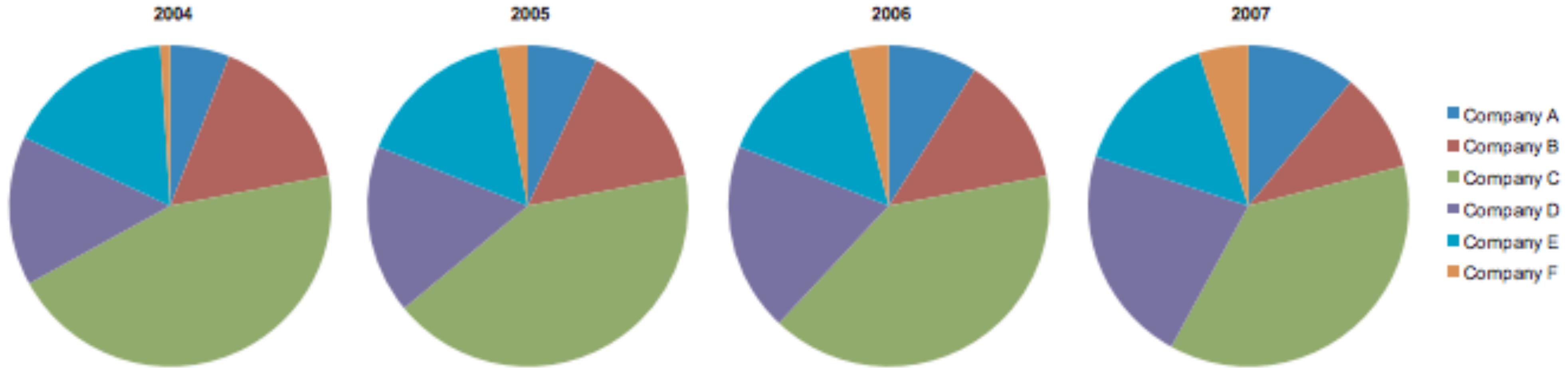
Visual Business Intelligence Newsletter

August 2007

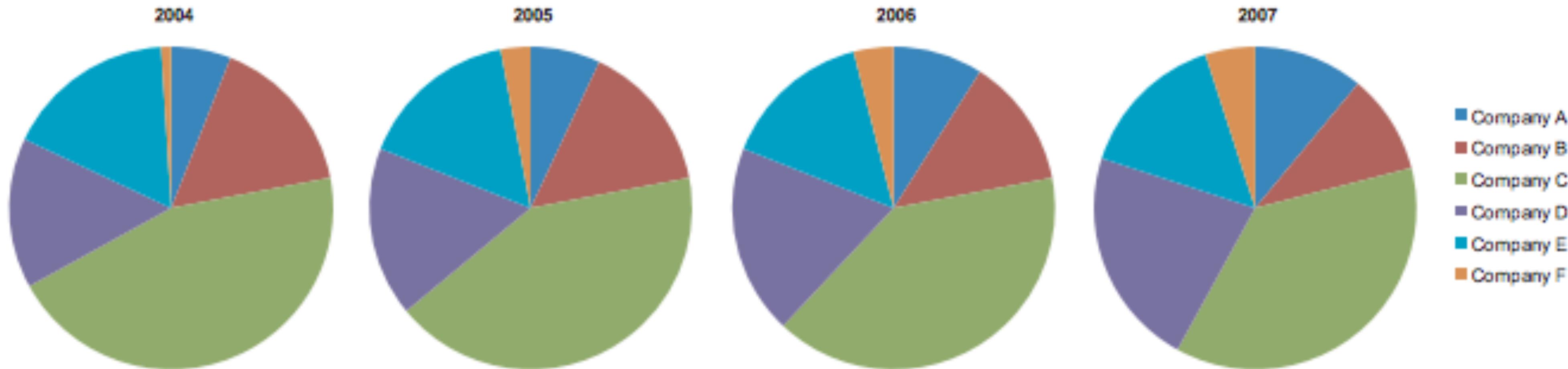
Not long ago I received an email from a colleague who keeps watch on business intelligence vendors and rates their products. She was puzzled that a particular product that I happen to like did not support pie charts, a feature that she assumed was basic and indispensable. Because of previous discussions between us, when I pointed out ineffective graphing practices that are popular in many BI products, she wondered if there might also be a problem with pie charts. Could this vendor's omission of pie charts be intentional and justified? I explained that this was indeed the case, and praised the vendor's design team for their good sense.

Here sits the friendly pie chart:

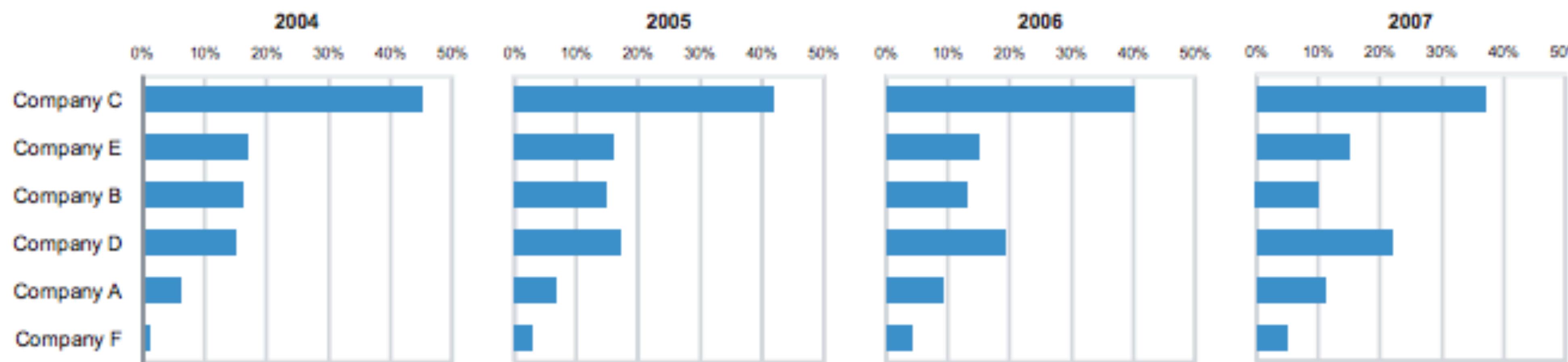




Try to follow the changes of these various companies and how they compare to one another through time. It is nearly impossible. Notice how easily you can do it, however, using the following display:



Try to follow the changes of these various companies and how they compare to one another through time. It is nearly impossible. Notice how easily you can do it, however, using the following display:



U.S. \$

20,000

16,000

12,000

8,000

4,000

0

Total Company Revenues by Year

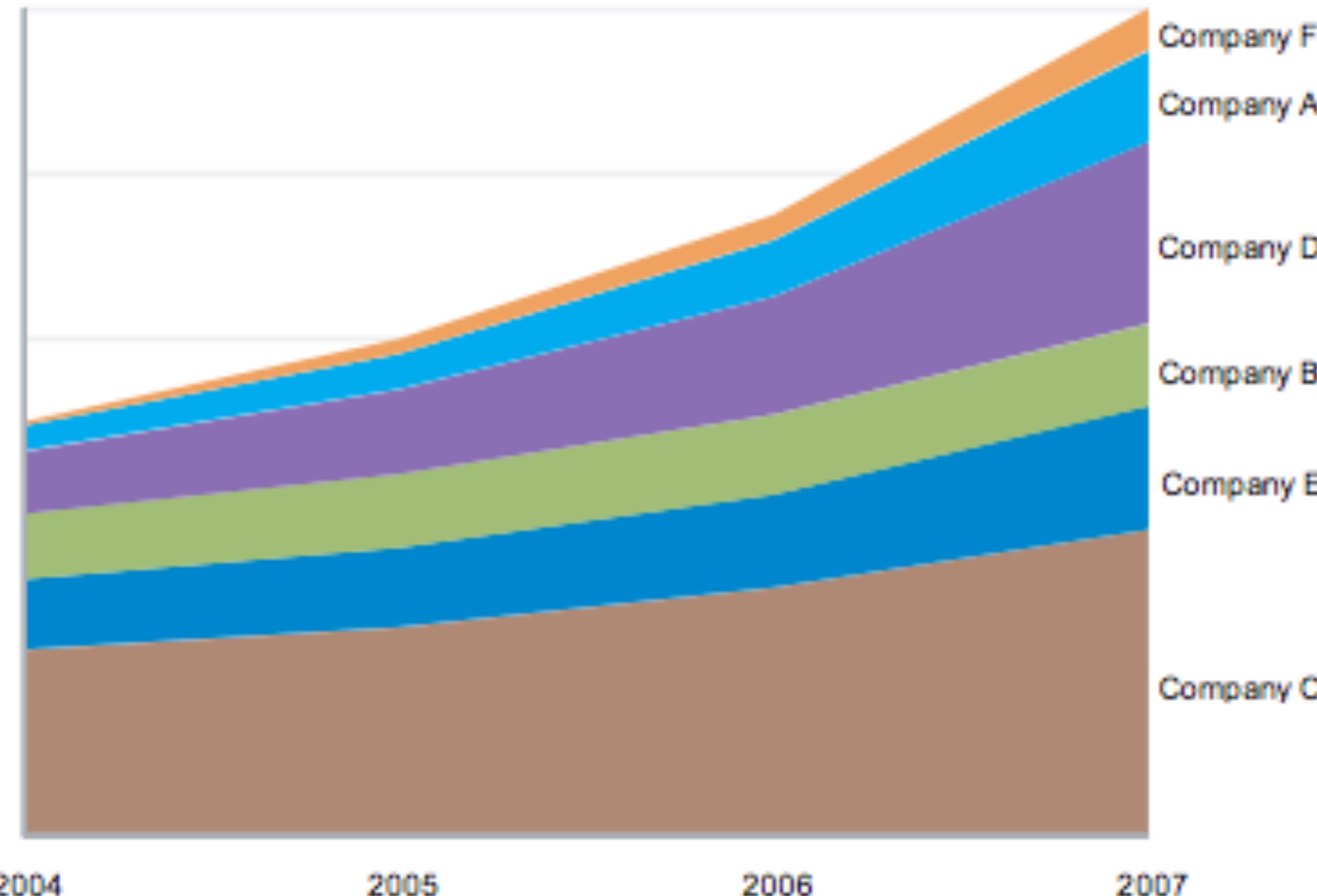


GRÁFICO DE ÁREAS EMPILHADAS

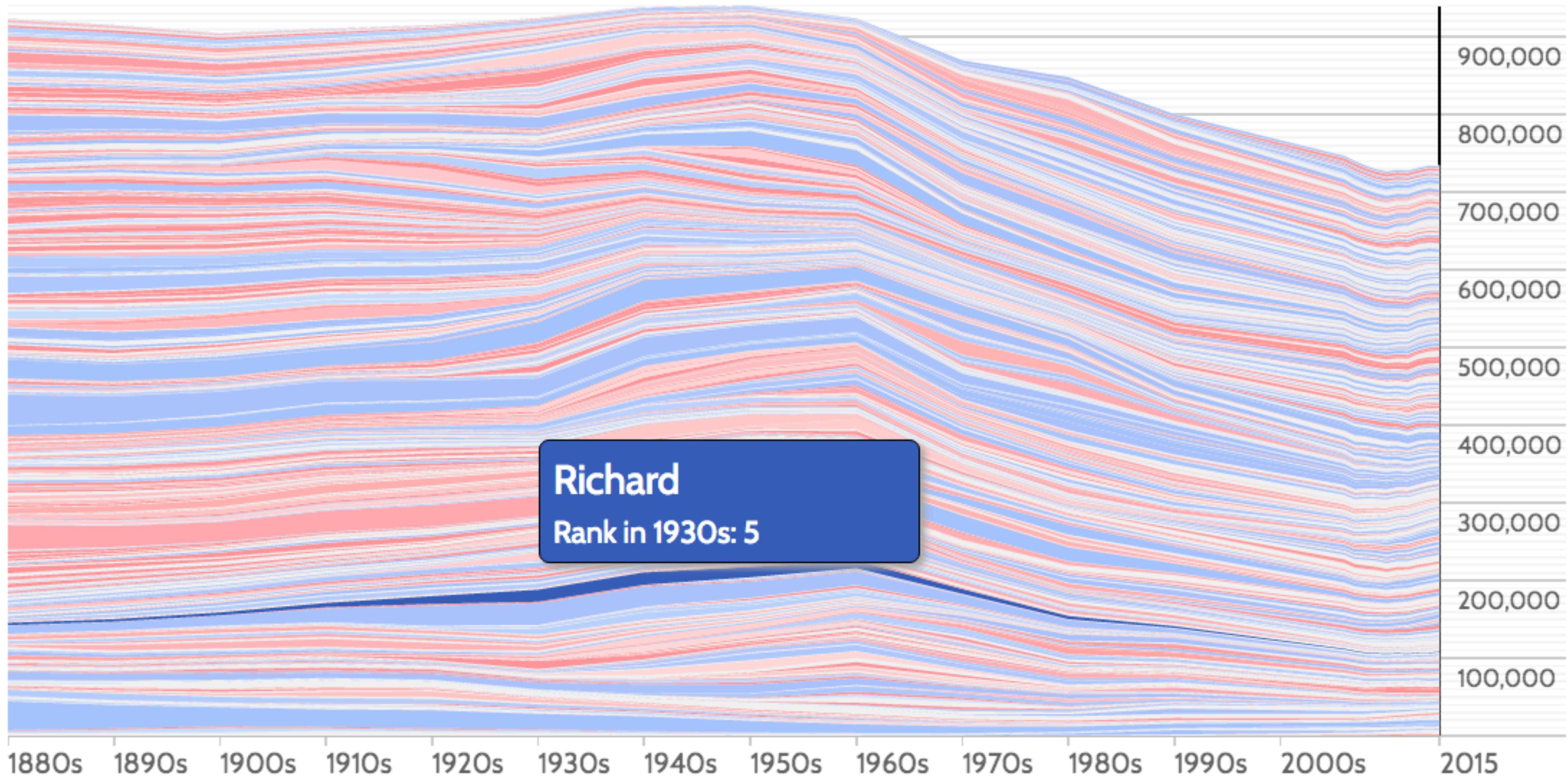
Baby Name >

Both Boys Girls

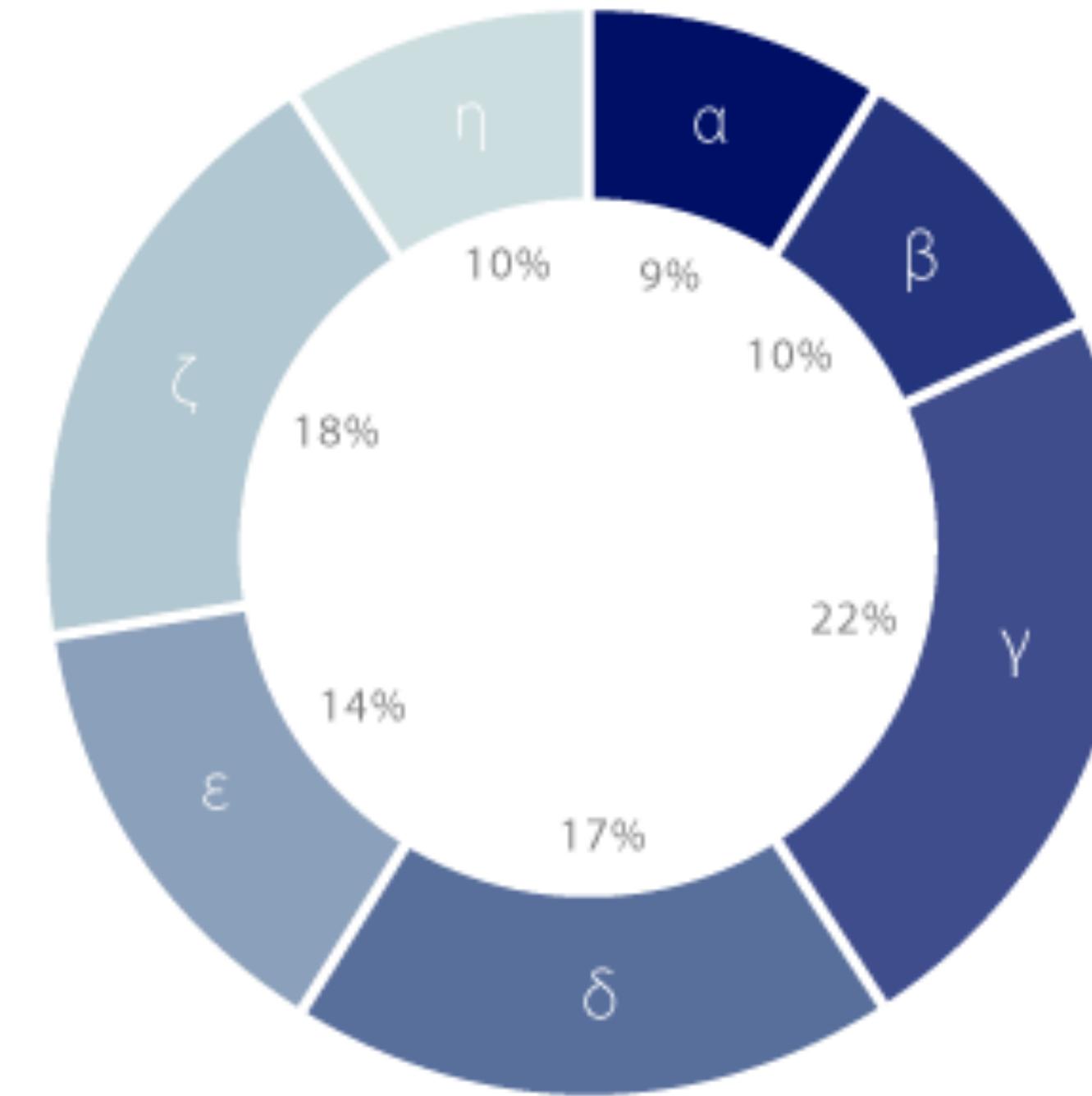
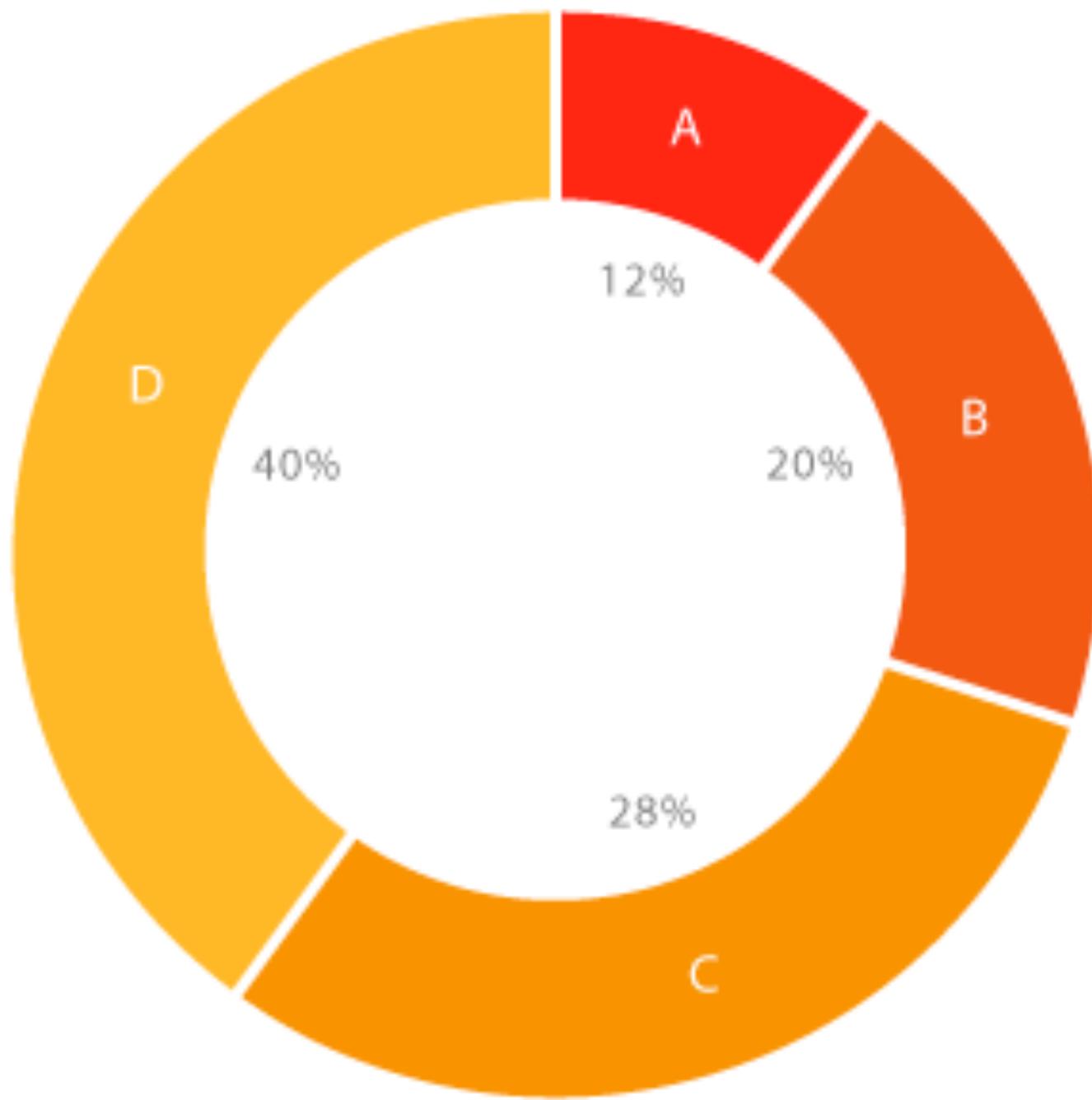
boys	1000	500	100	25	1
girls	1000	500	100	25	1

Current rank:

per million births

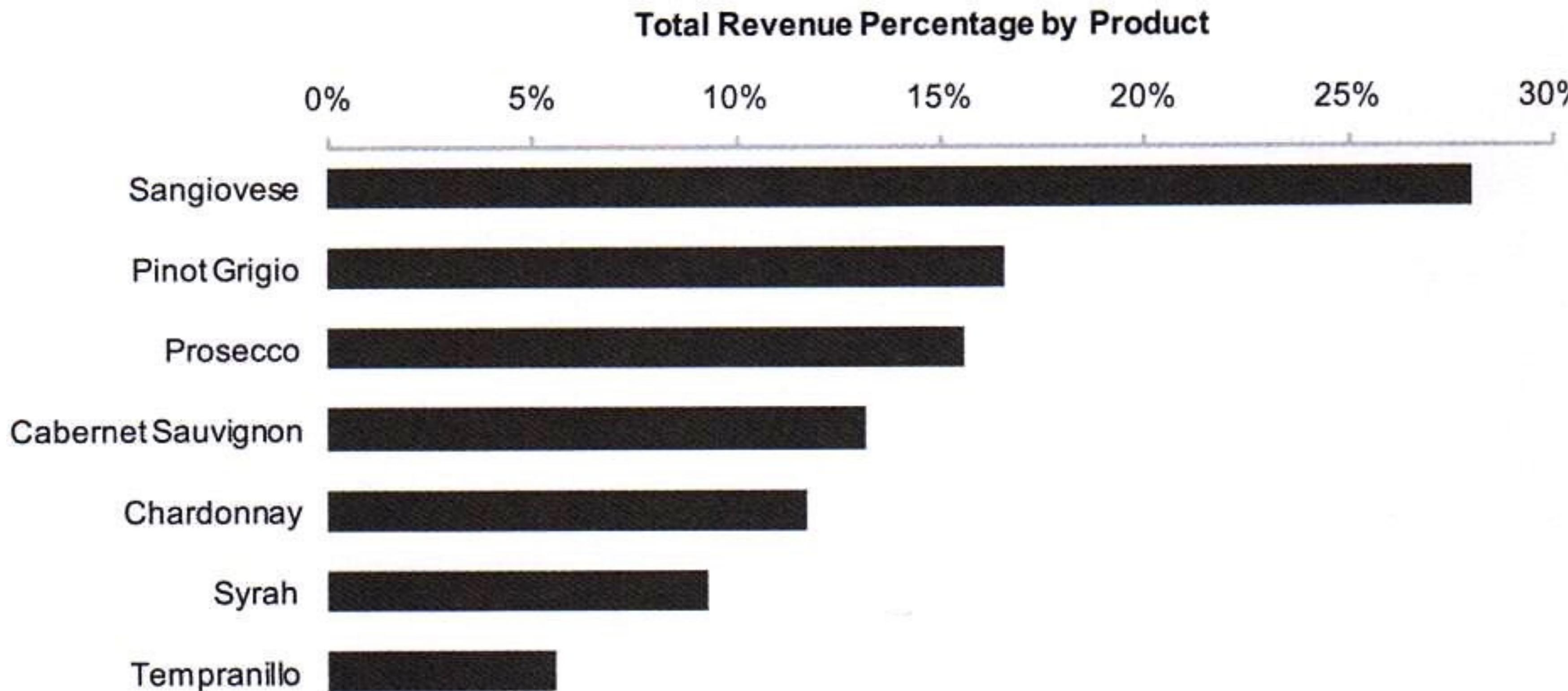


DOUGHNUT CHART (GRÁFICO DE ROSCA)



- Essencialmente o mesmo que um gráfico de pizza
- Há quem defende que eles são melhores pois o foco está na comparação dos comprimentos dos arcos e não nas áreas dos setores
- Uma vantagem é ocupar menos espaço e a possibilidade de se colocar valores ou texto no centro

GRÁFICO DE BARRAS

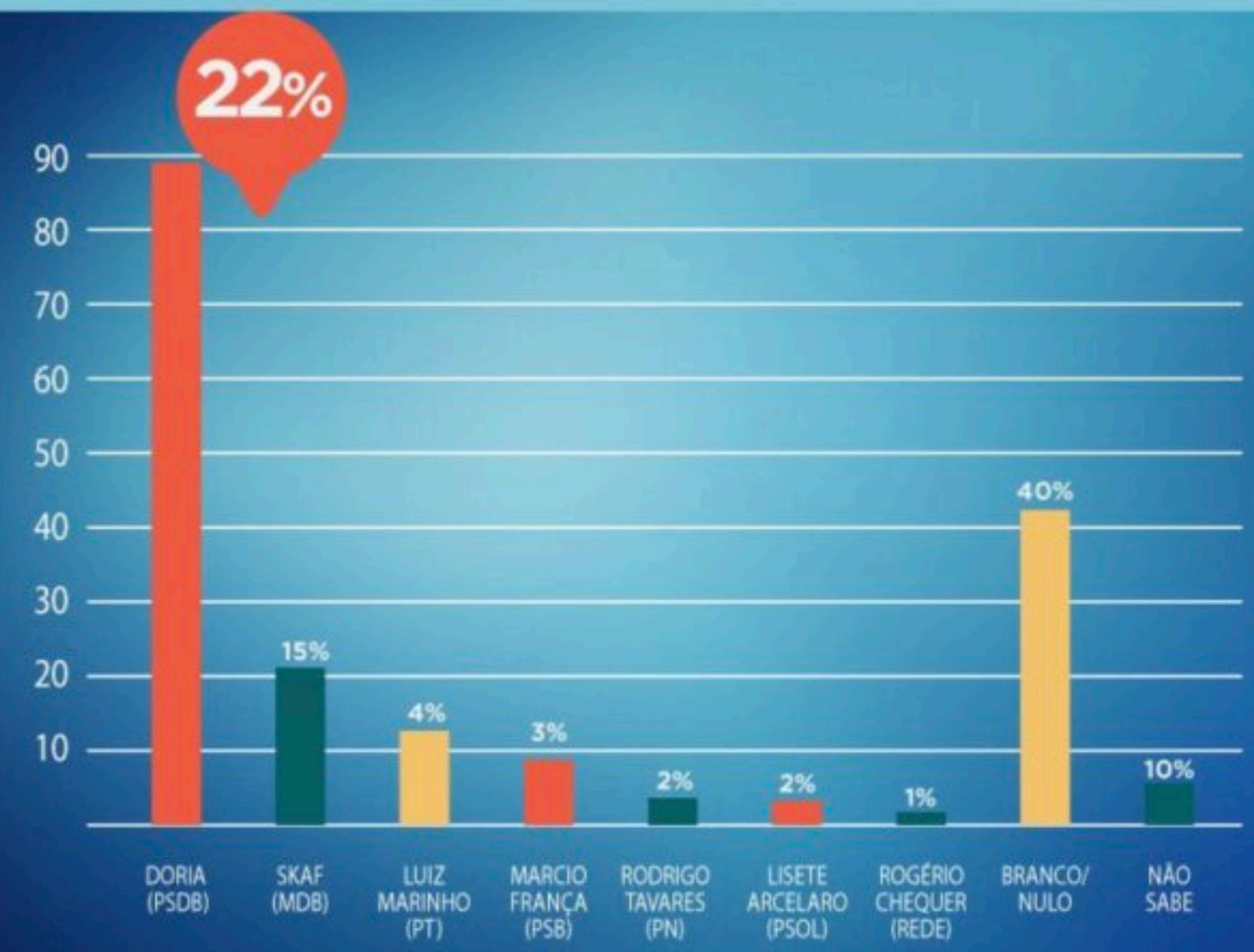


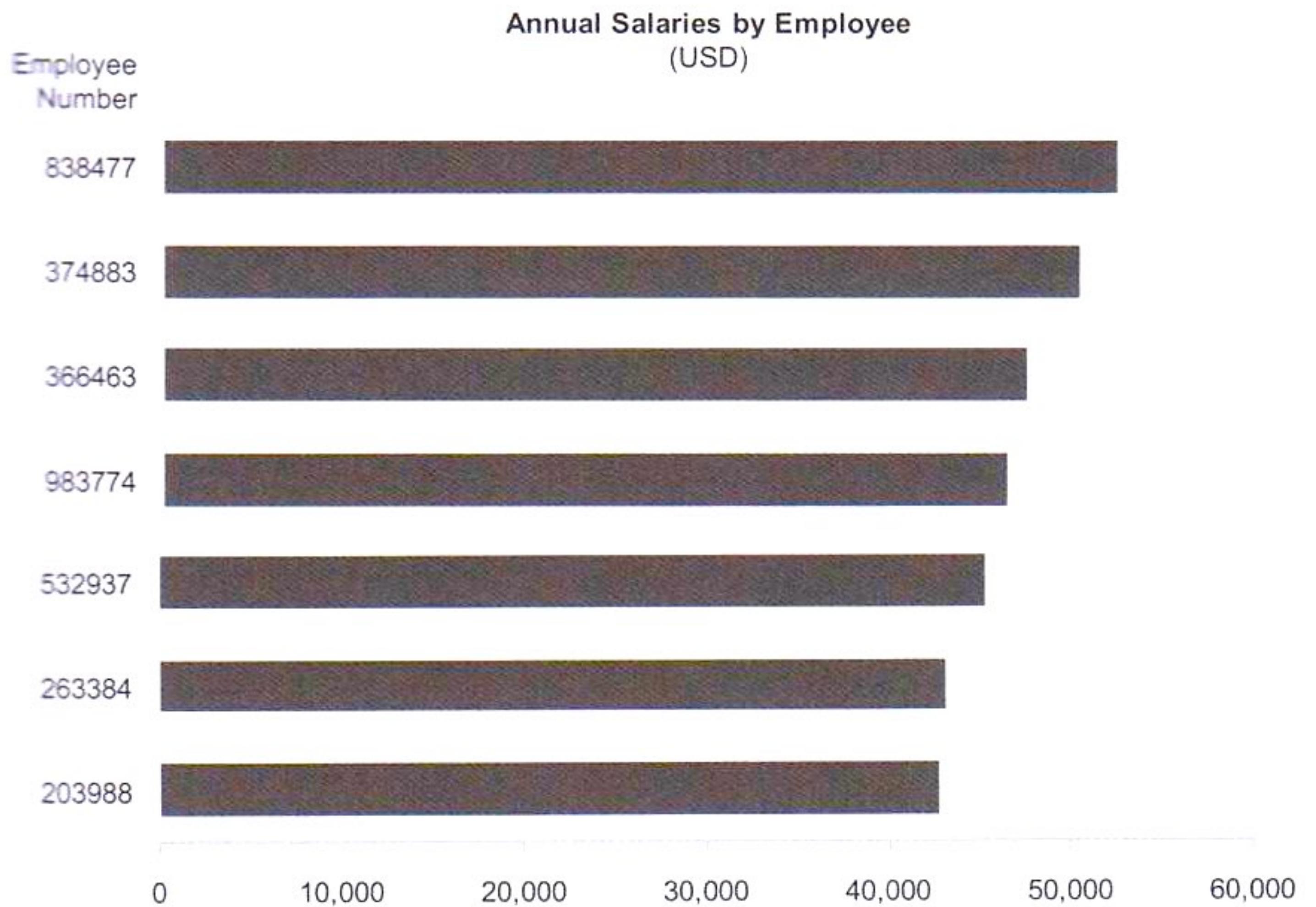
- Gráficos de barra são muito mais efetivos para representar relacionamentos do tipo ranking e parte-todo



DORIA LIDEREA PESQUISA

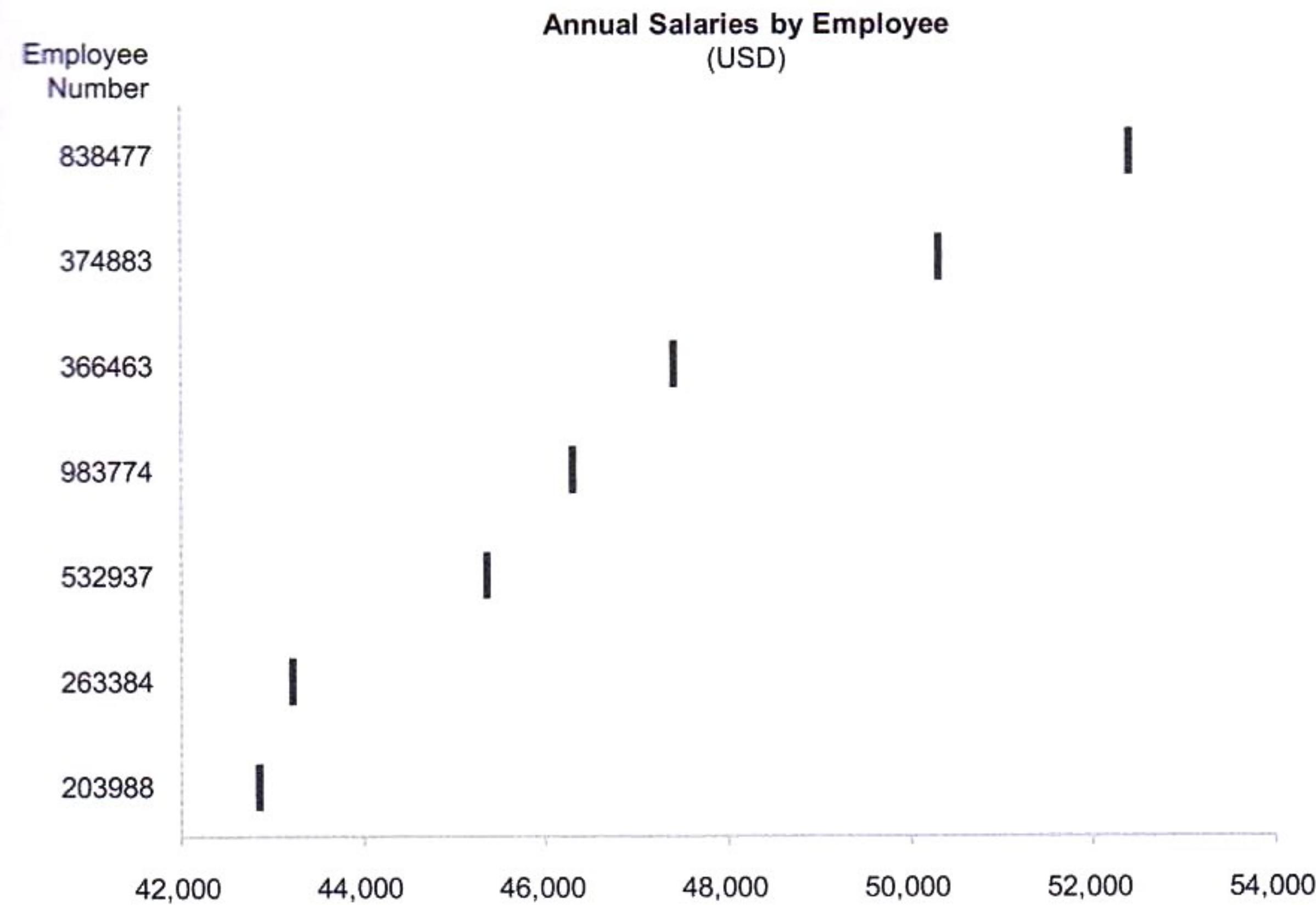
APONTA IBOPE





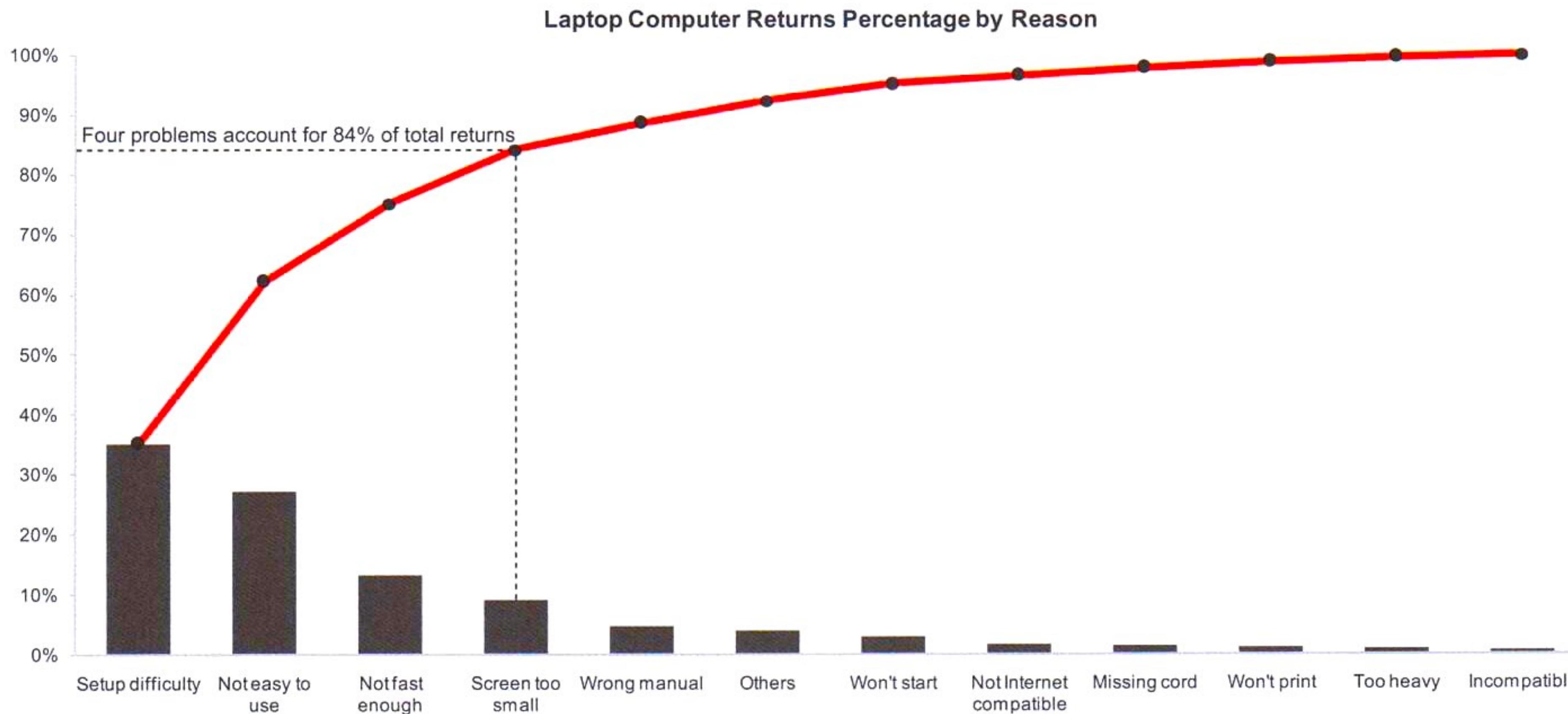
Quando as barras tem **comprimentos parecidos**, são menos efetivas na estimativa das diferenças

GRÁFICO DE PONTOS



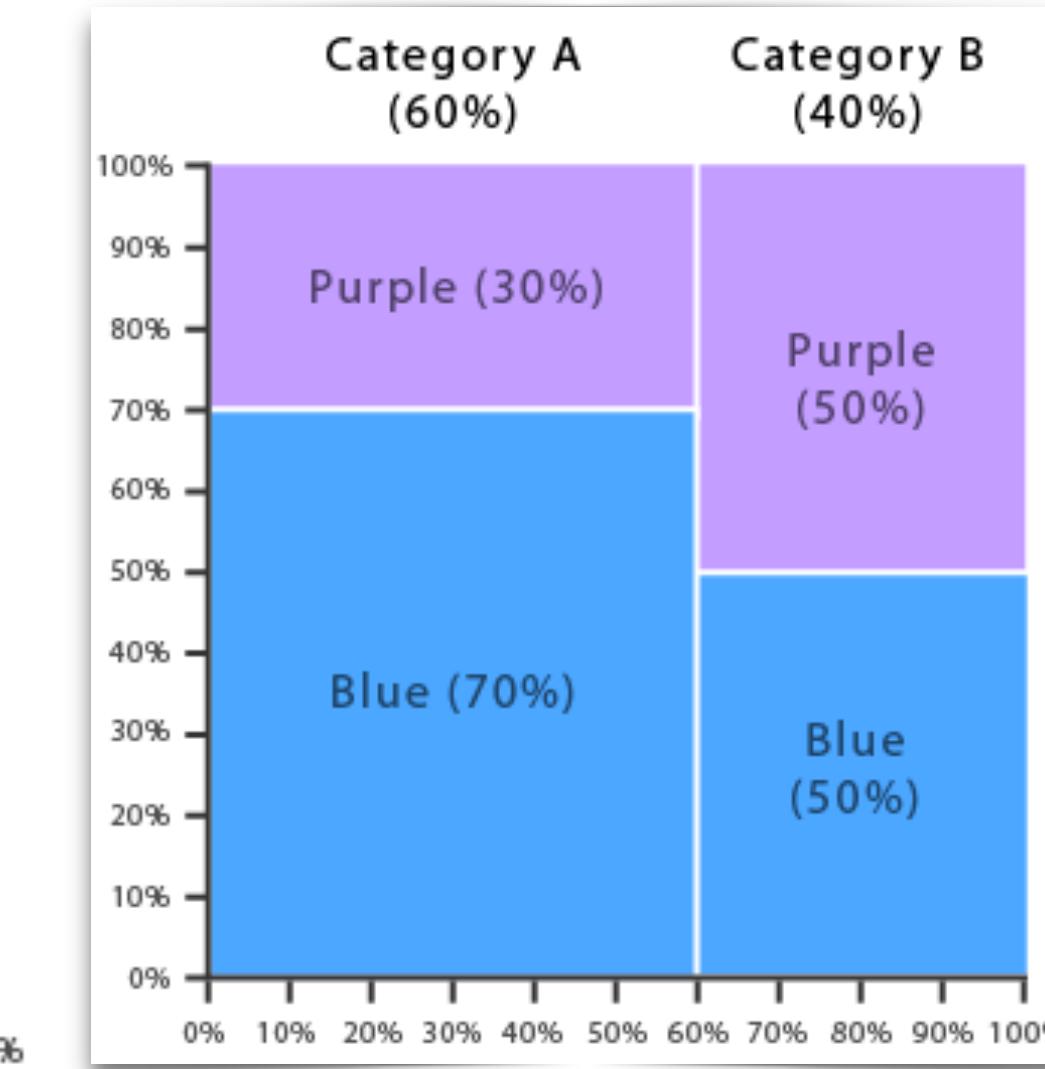
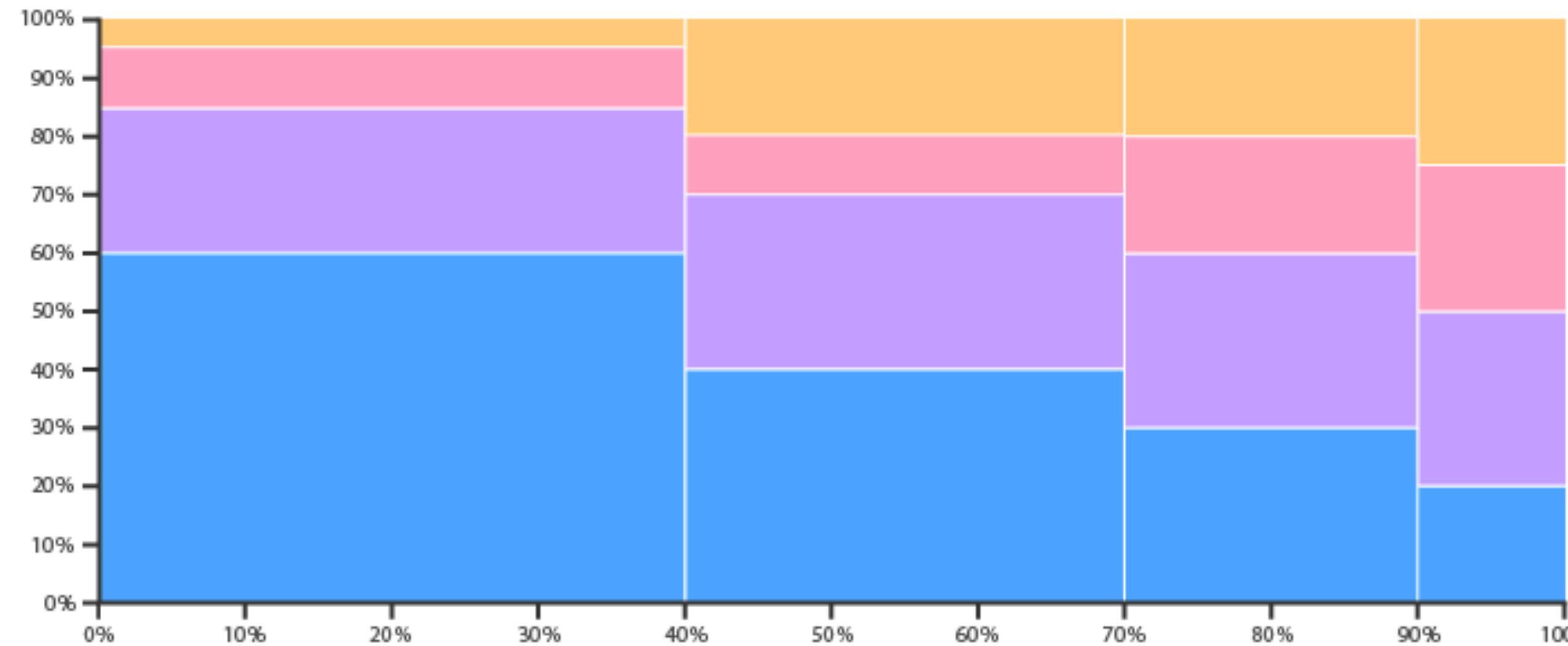
- Nesses casos, o uso de gráficos de pontos facilita essa diferenciação visto que podemos usar eixos que não iniciam de zero

GRÁFICO OU DIAGRAMA DE PARETO



- Barras ordenadas representando o ranking e linha mostrando valores acumulados representando assim a parte do todo

MARIMEKKO CHART (MOSAIC PLOT)

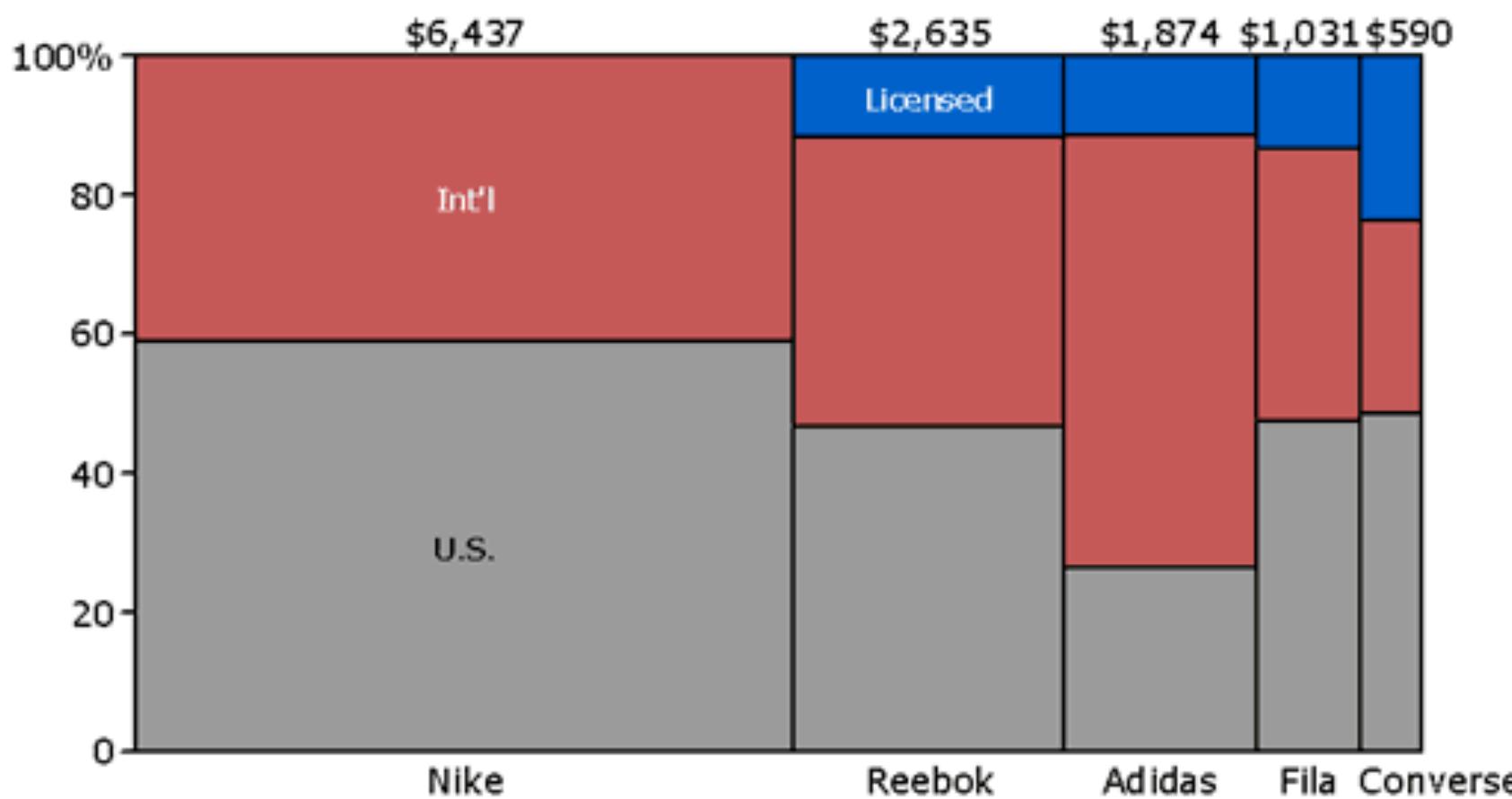


- Codifica duas variáveis quantitativas: uma na largura e outra na altura
- De difícil leitura, especialmente quando há muitos segmentos
- De difícil comparação já que os segmentos não se apoiam na mesma base

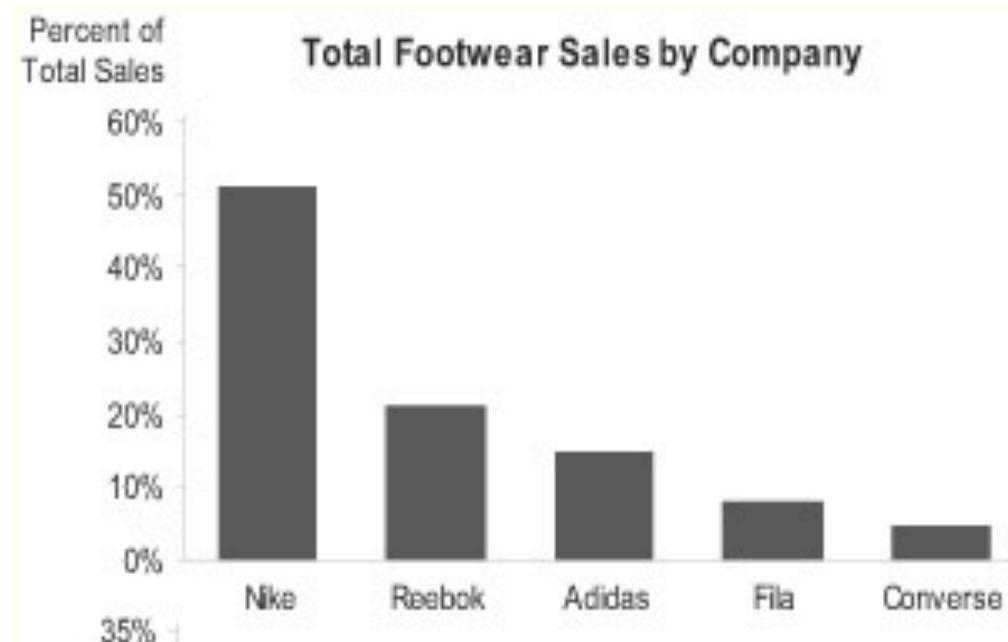
Marimekko Chart

Nike dominates its top four competitors with a mix of U.S. and international sales.

Footwear Sales in \$ Million



Total Footwear Sales by Company



Total Footwear Sales by Channel

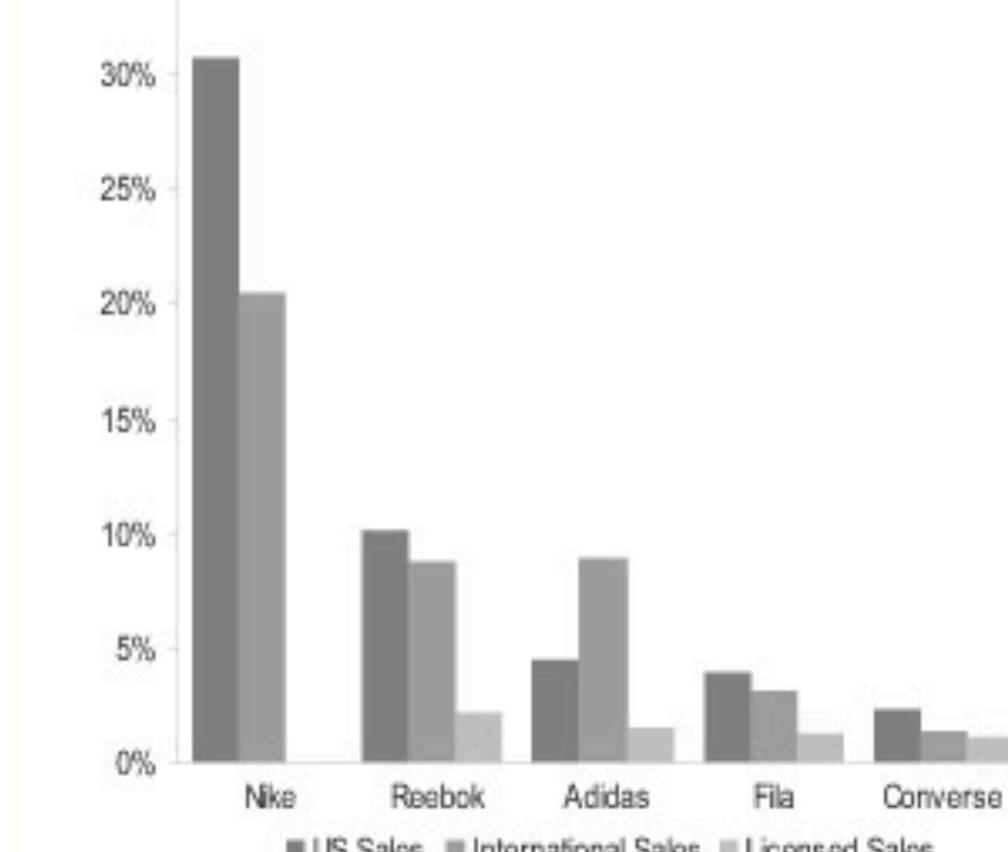
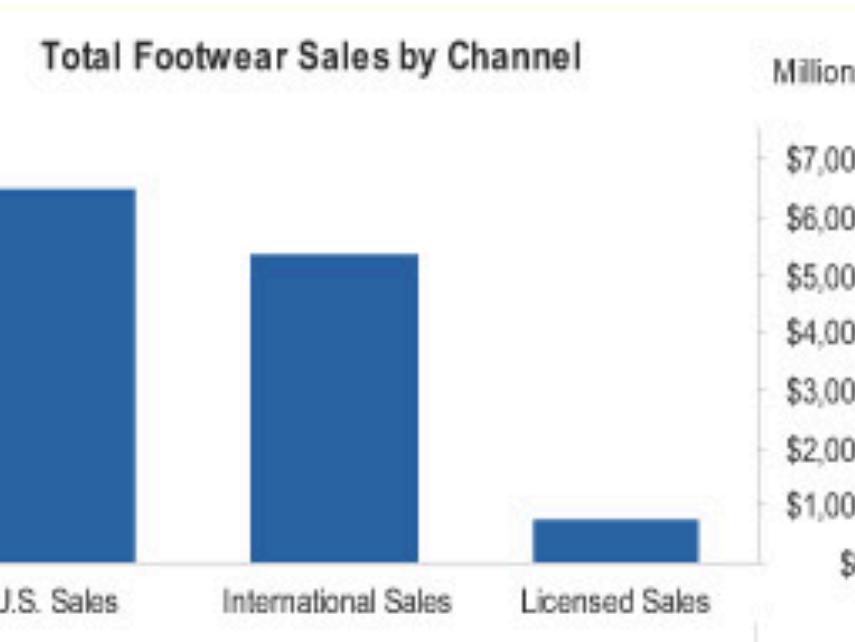
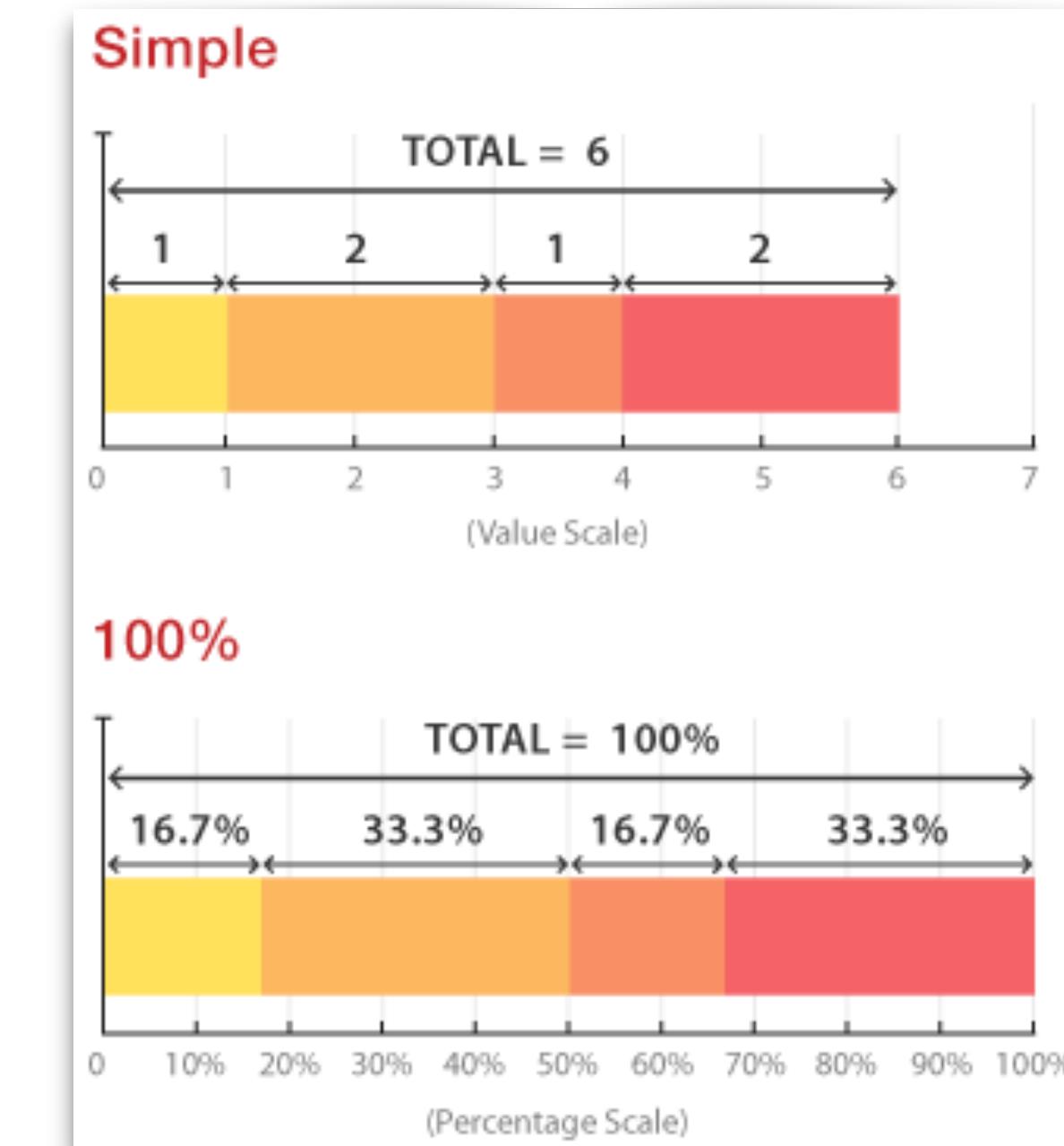
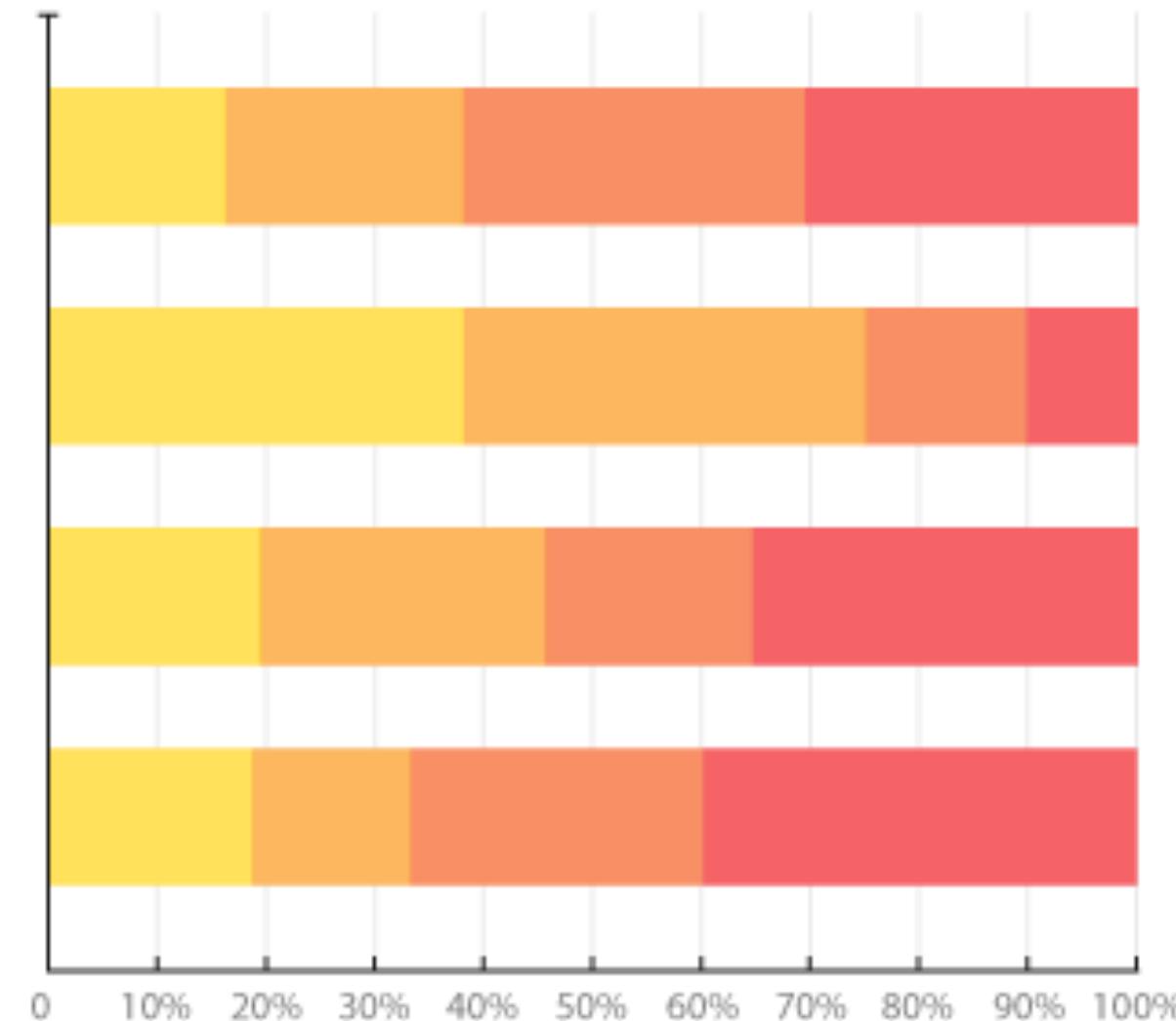
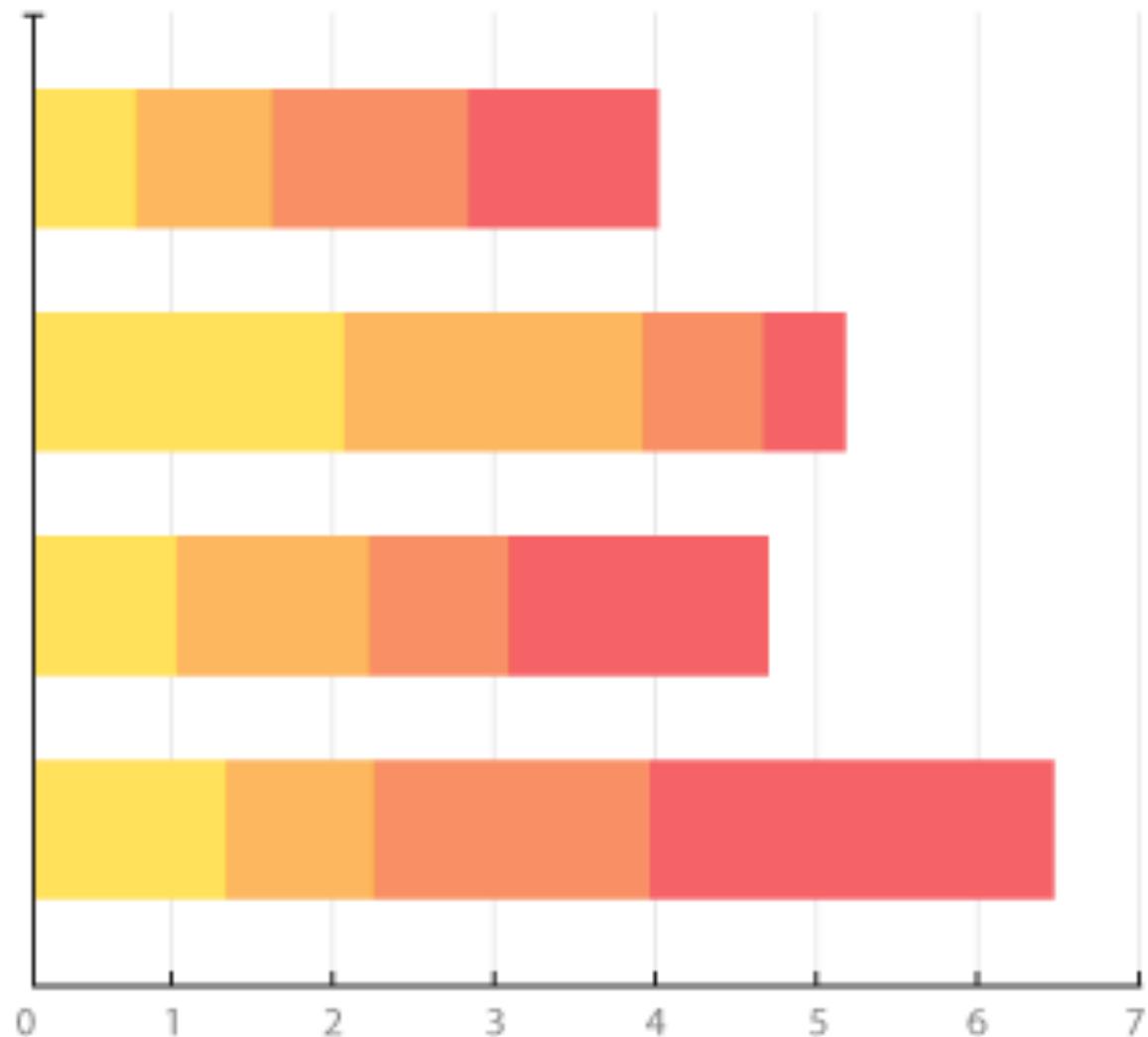
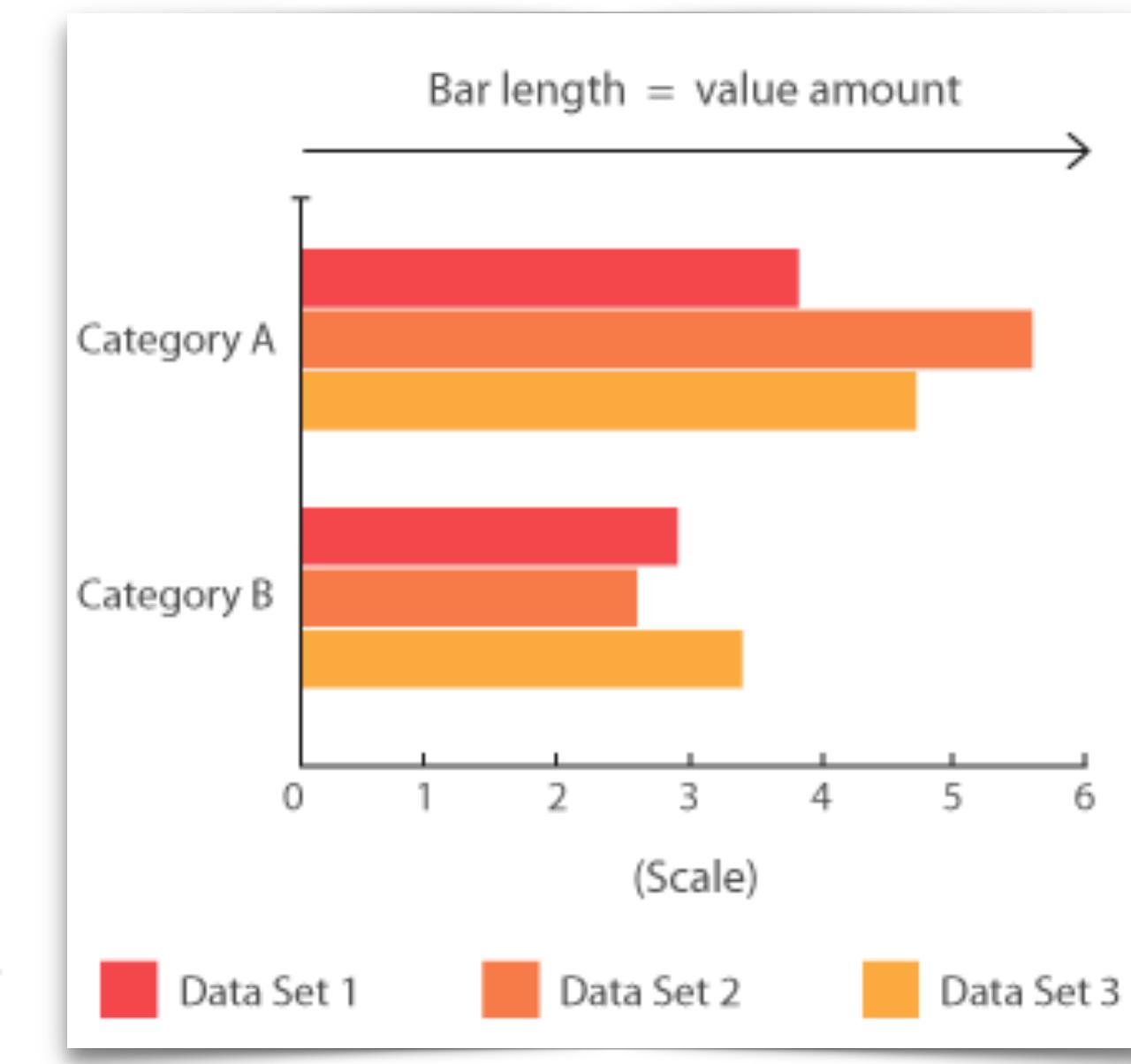
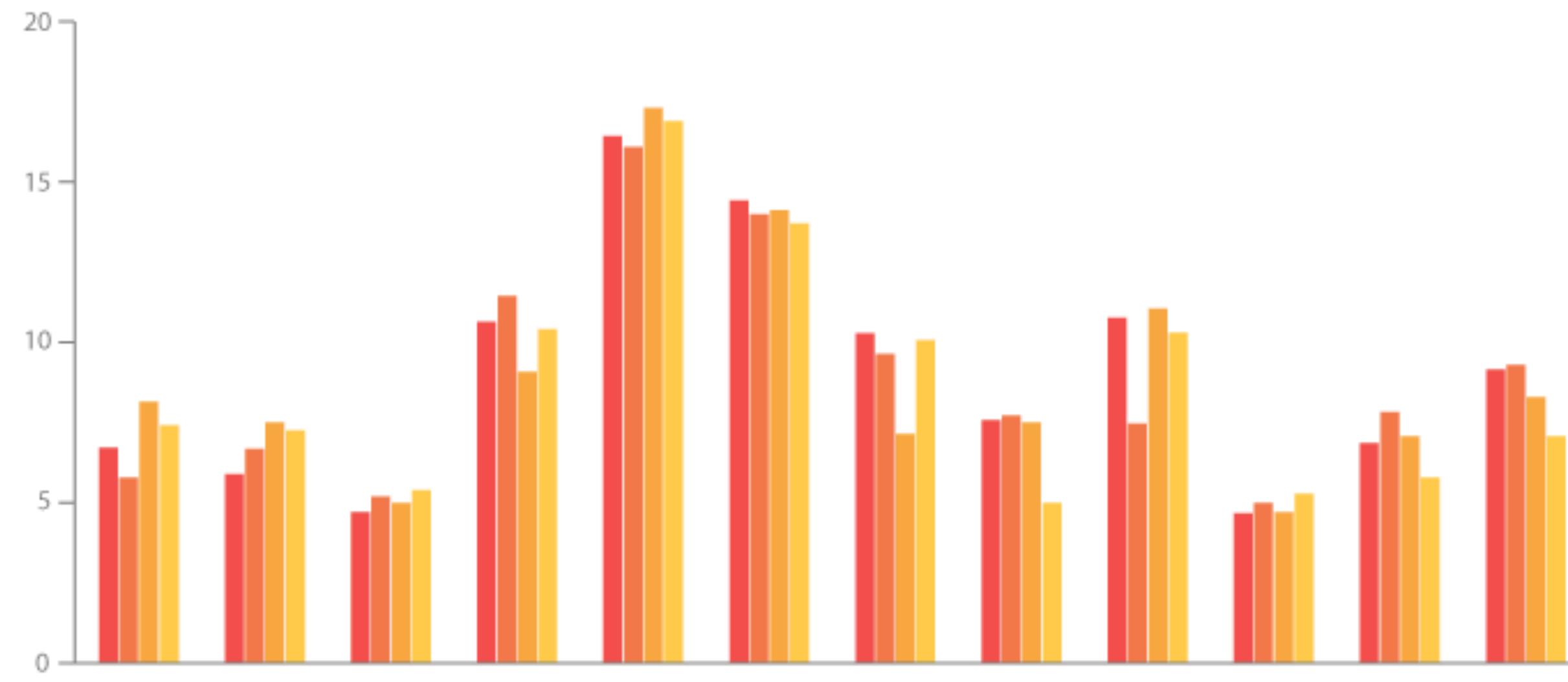


GRÁFICO DE BARRAS EMPILHADAS



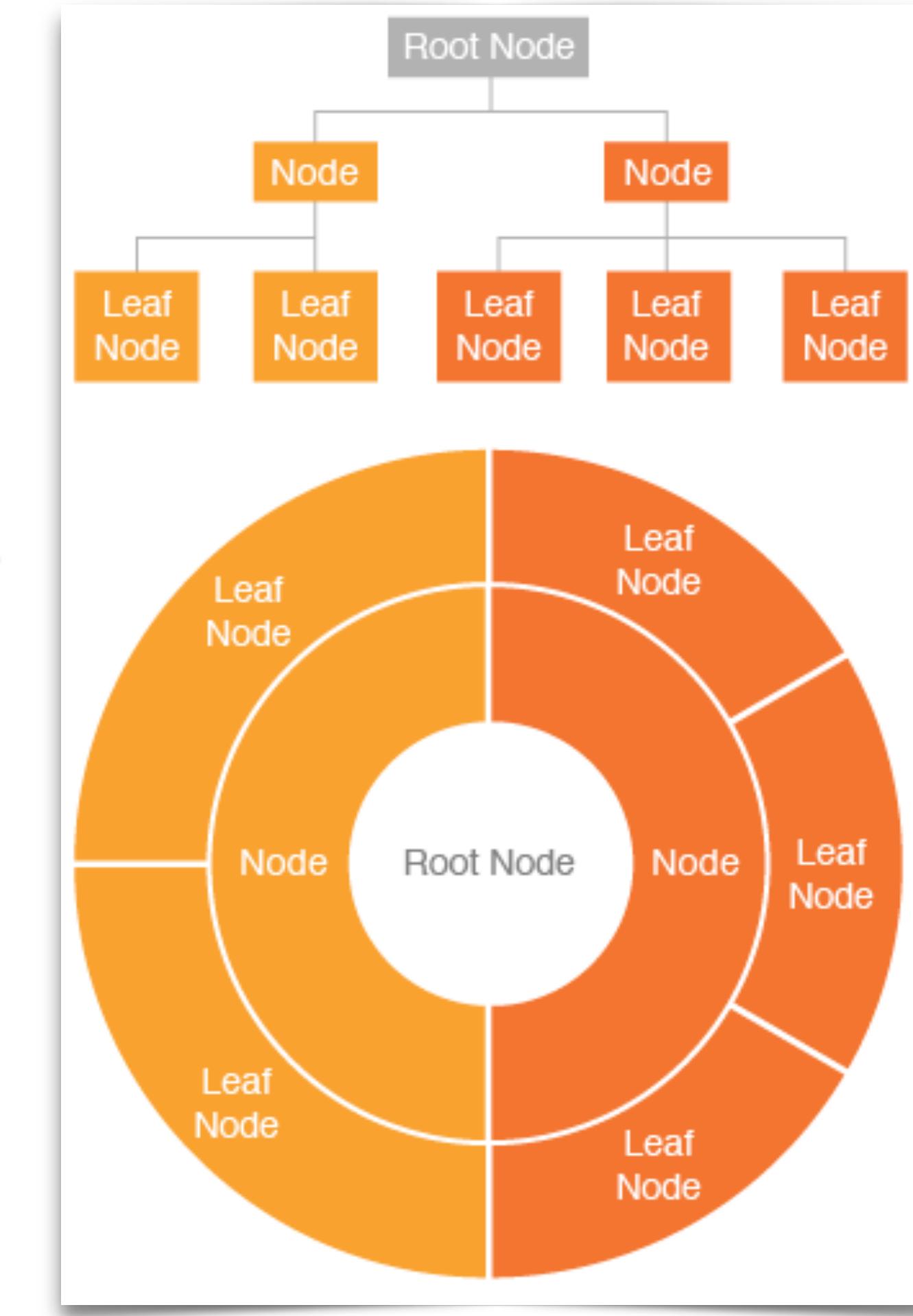
- Quanto mais categorias, mais difícil a leitura
- A comparação entre os segmentos também é dificultada pois não se apoiam sobre a mesma base

GRÁFICO DE BARRAS AGRUPADO



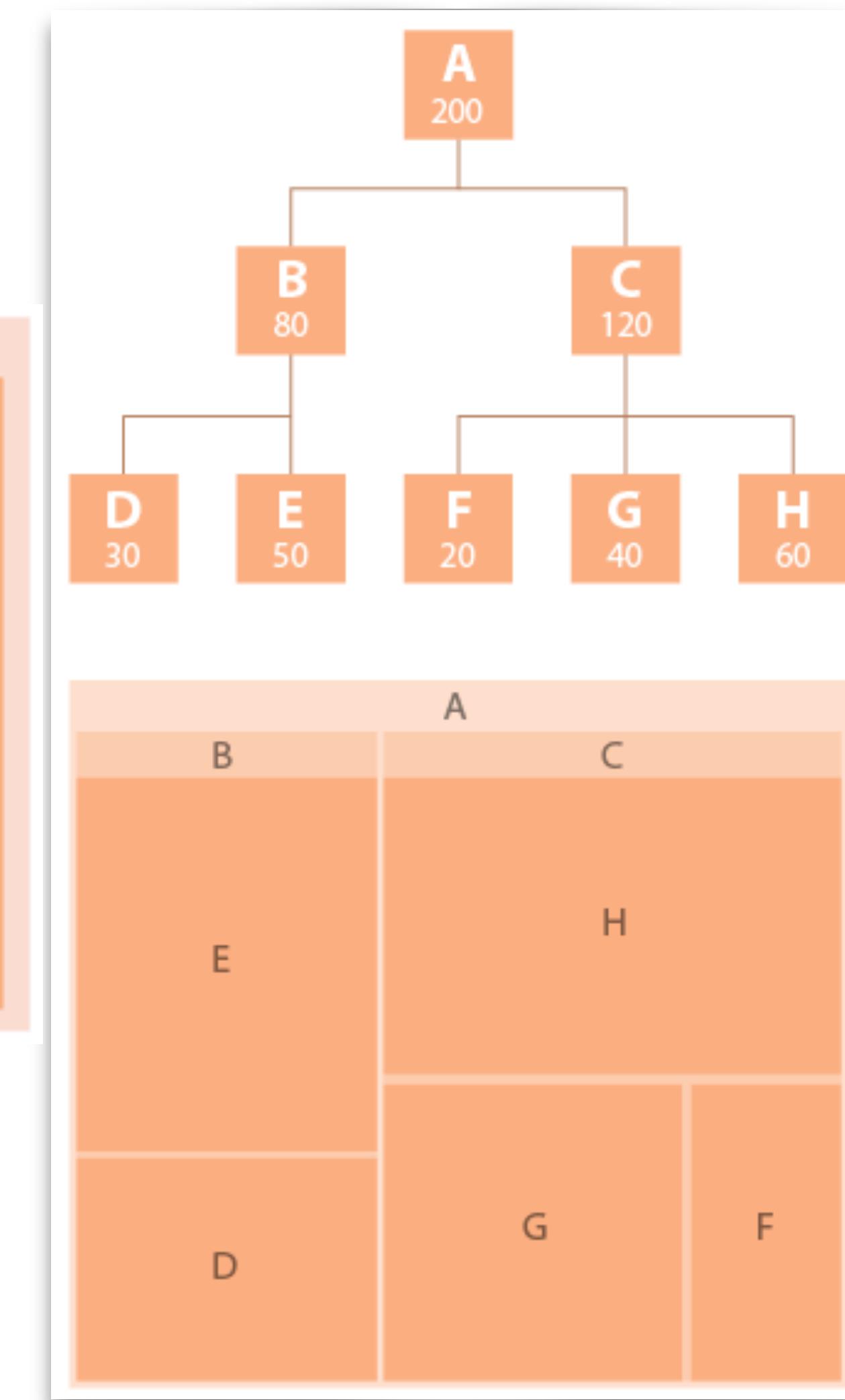
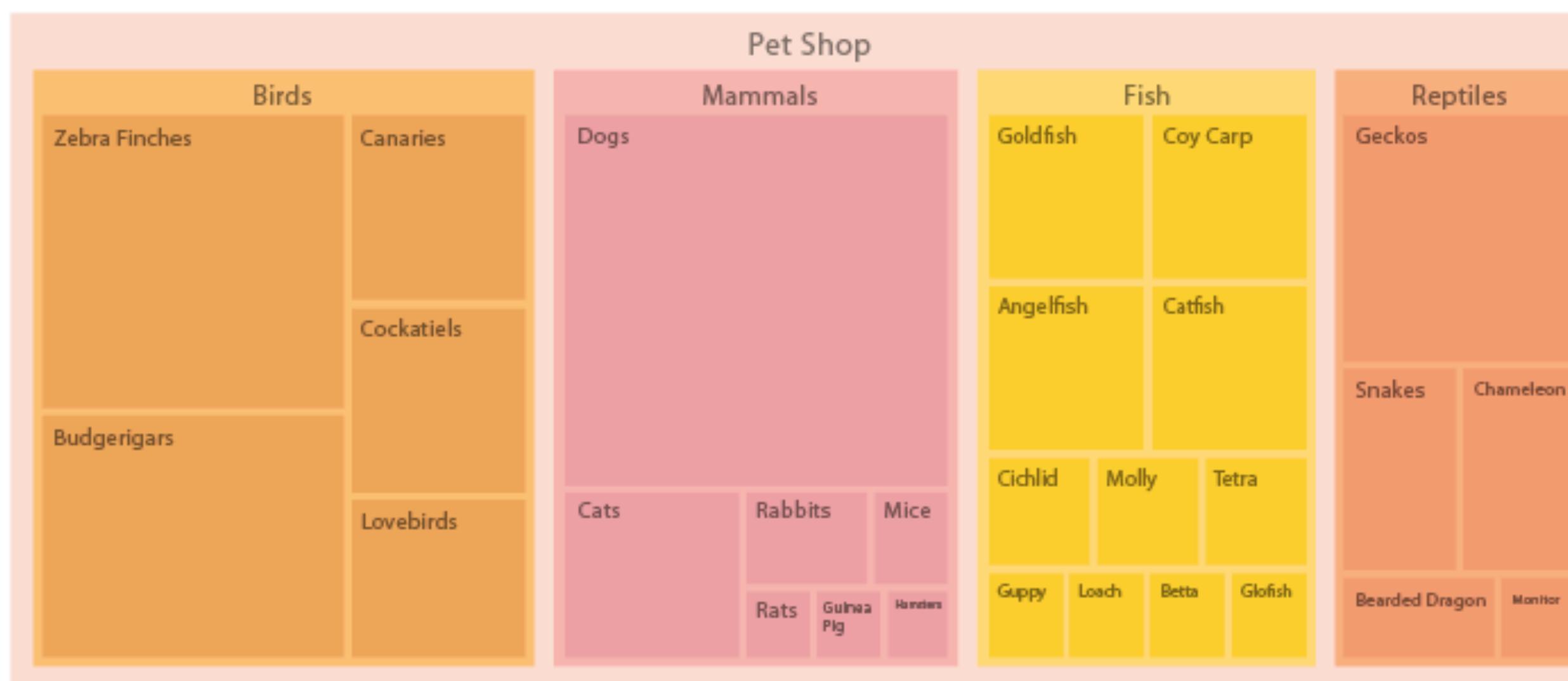
- Uma alternativa seria o gráfico de barras agrupado usando uma escala percentual
- Não passa tão bem a noção de parte-todo

SUNBURST

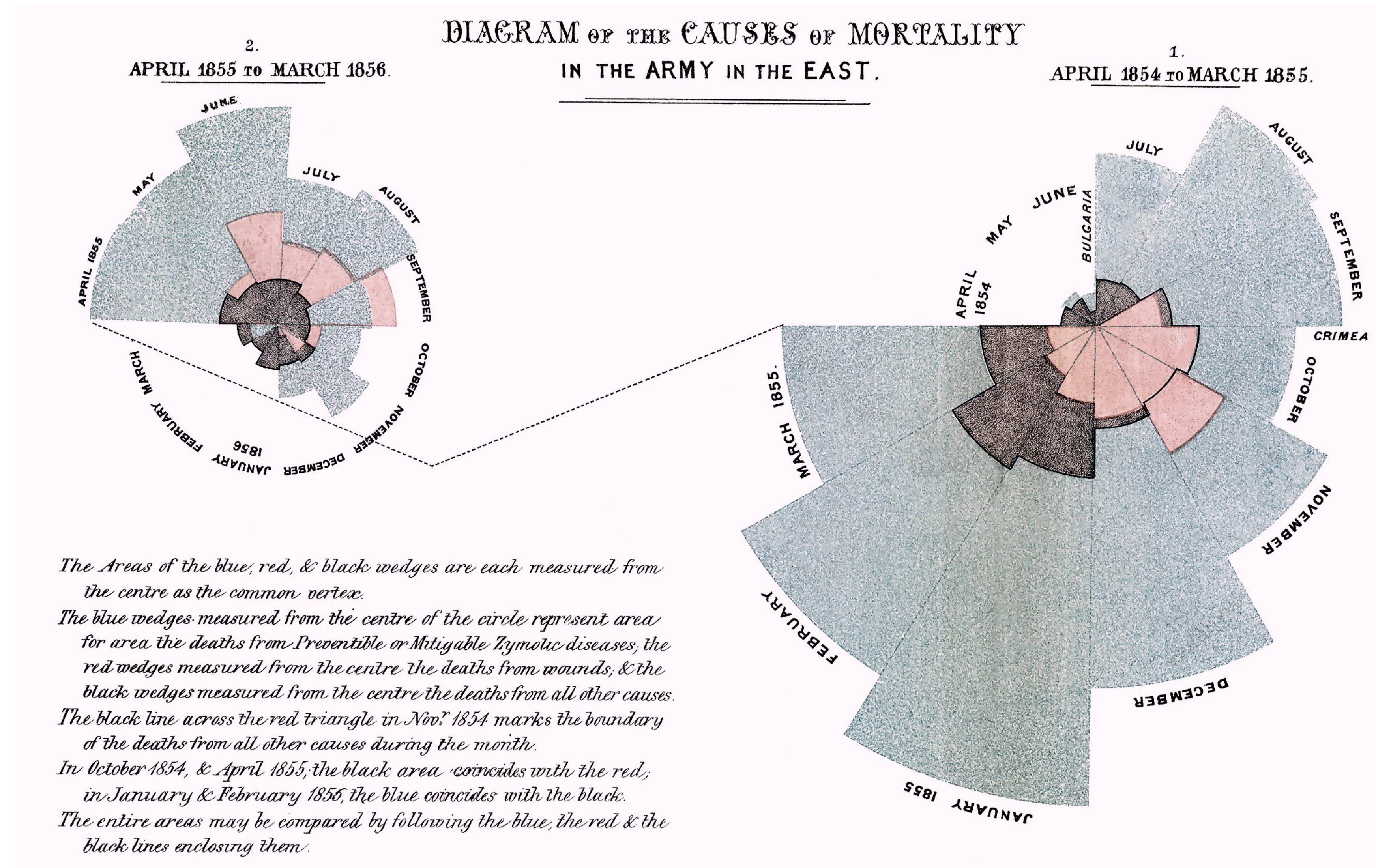


- Conhecido por vários outros nomes: *ring chart, multi-level pie chart, belt chart, radial treemap*

TREEMAP



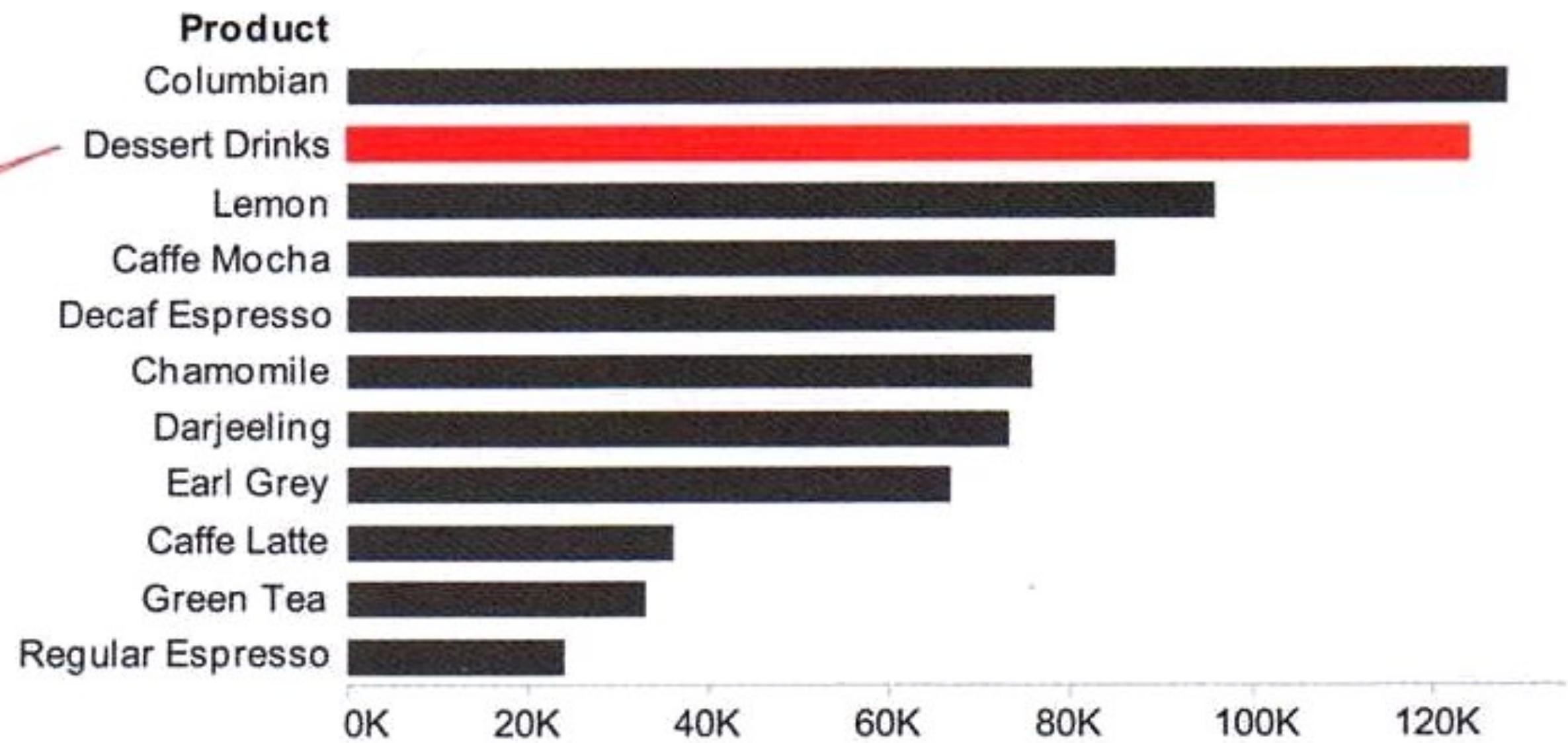
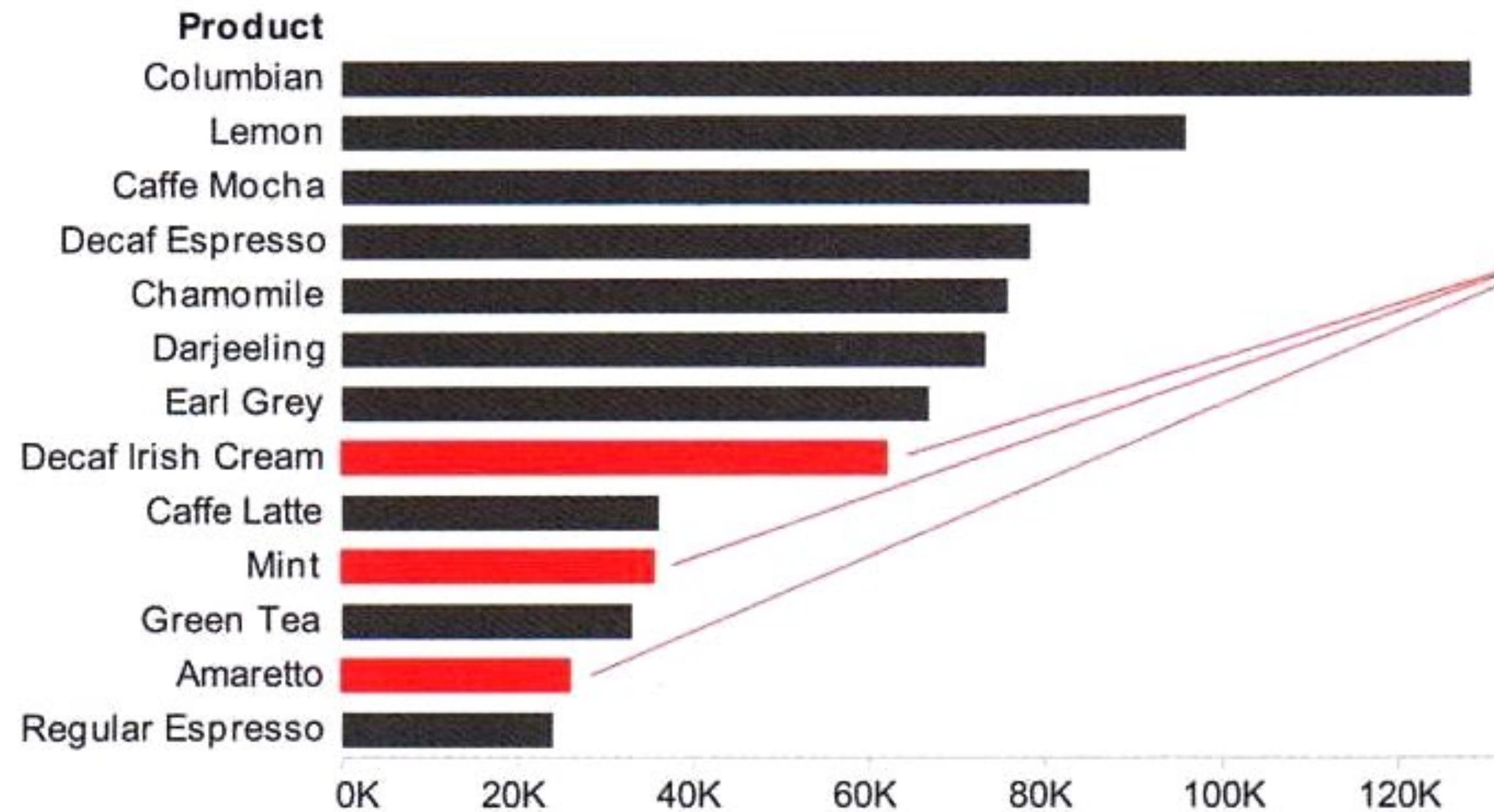
NIGHTINGALE



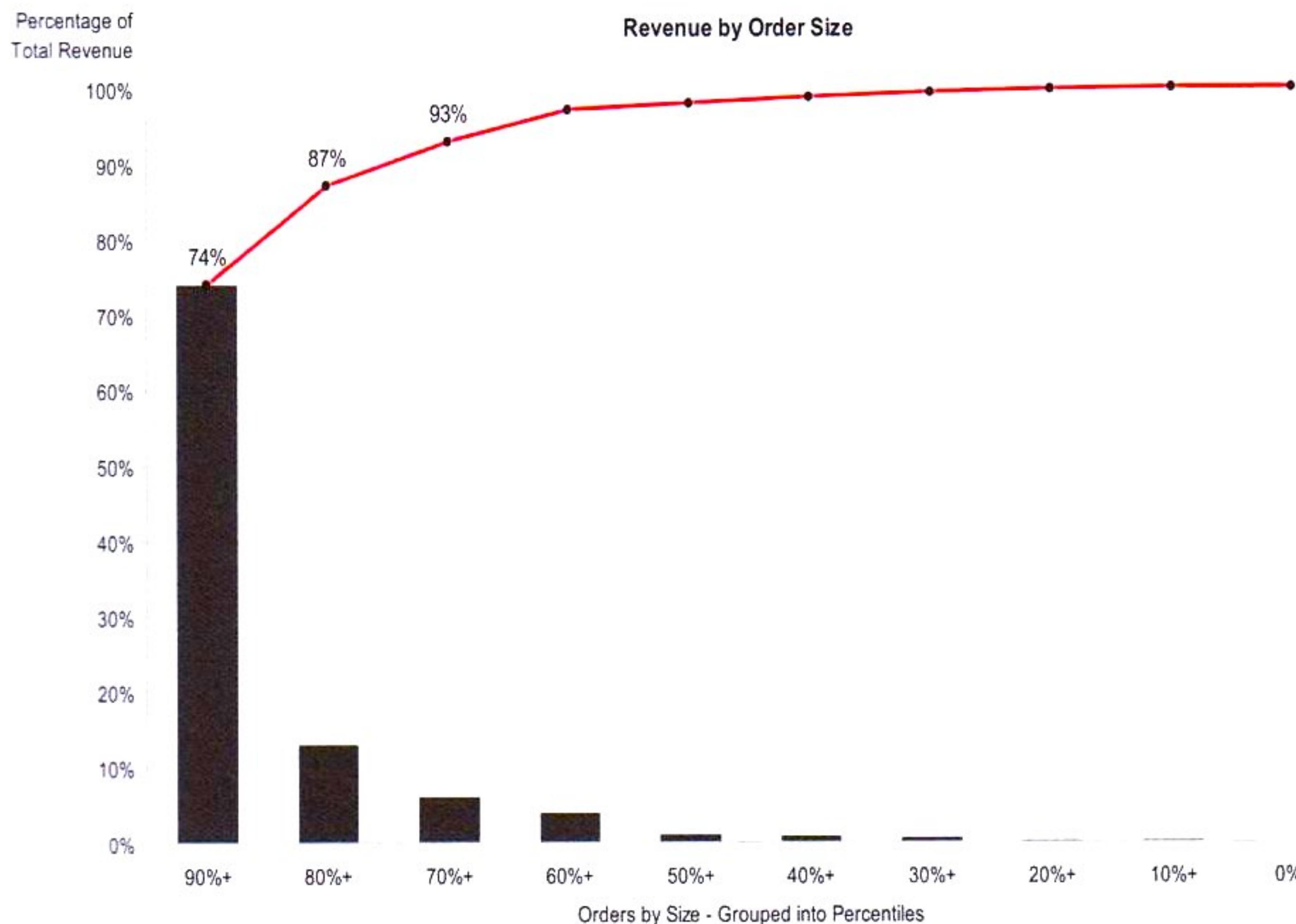
BOAS PRÁTICAS

PERMITA A CATEGORIZAÇÃO DE ITENS DE UMA MANEIRA AD-HOC

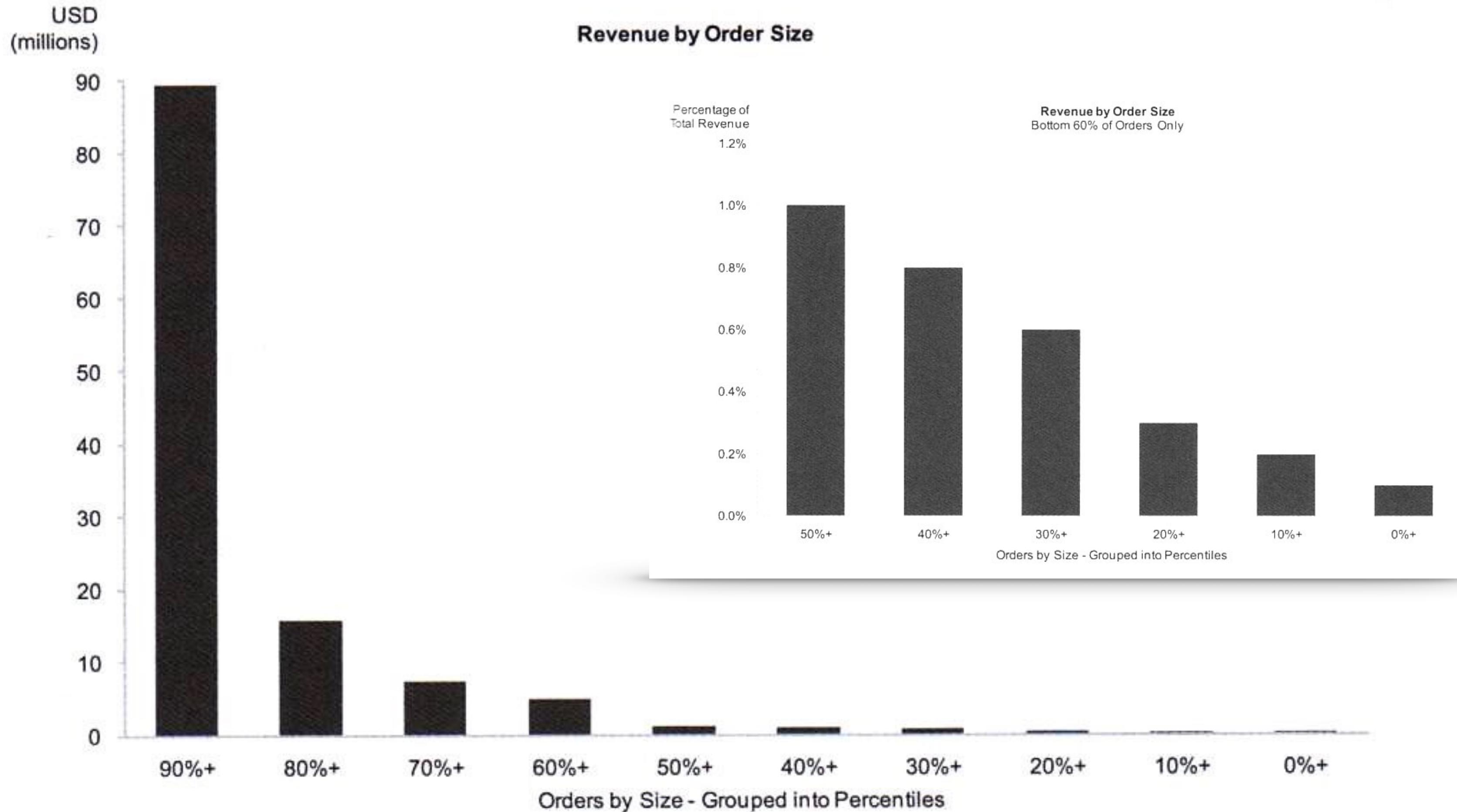
.....



USE GRÁFICOS DE PARETO COM PERCENTIS

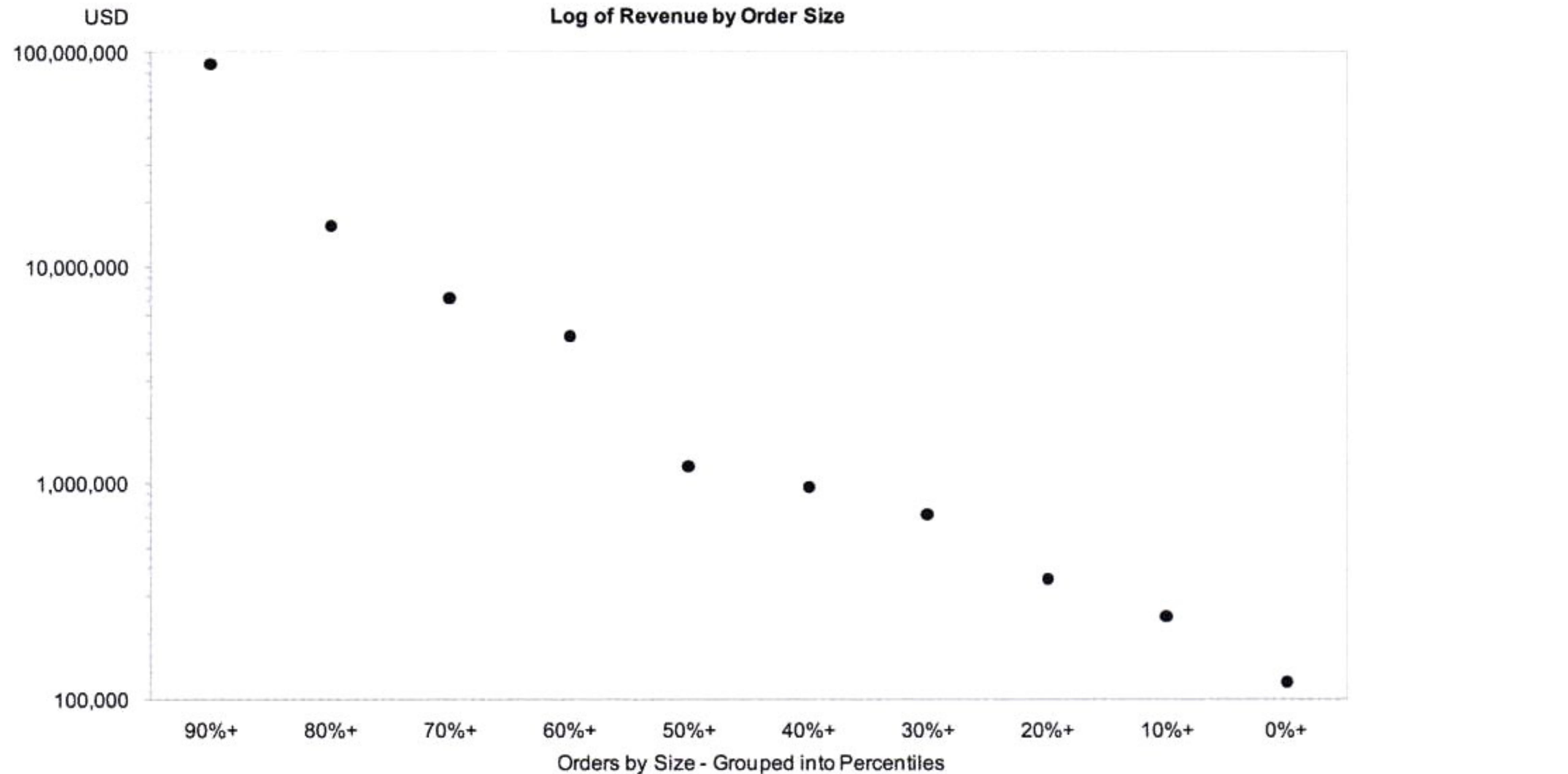


TESTE DIFERENTES ESCALAS

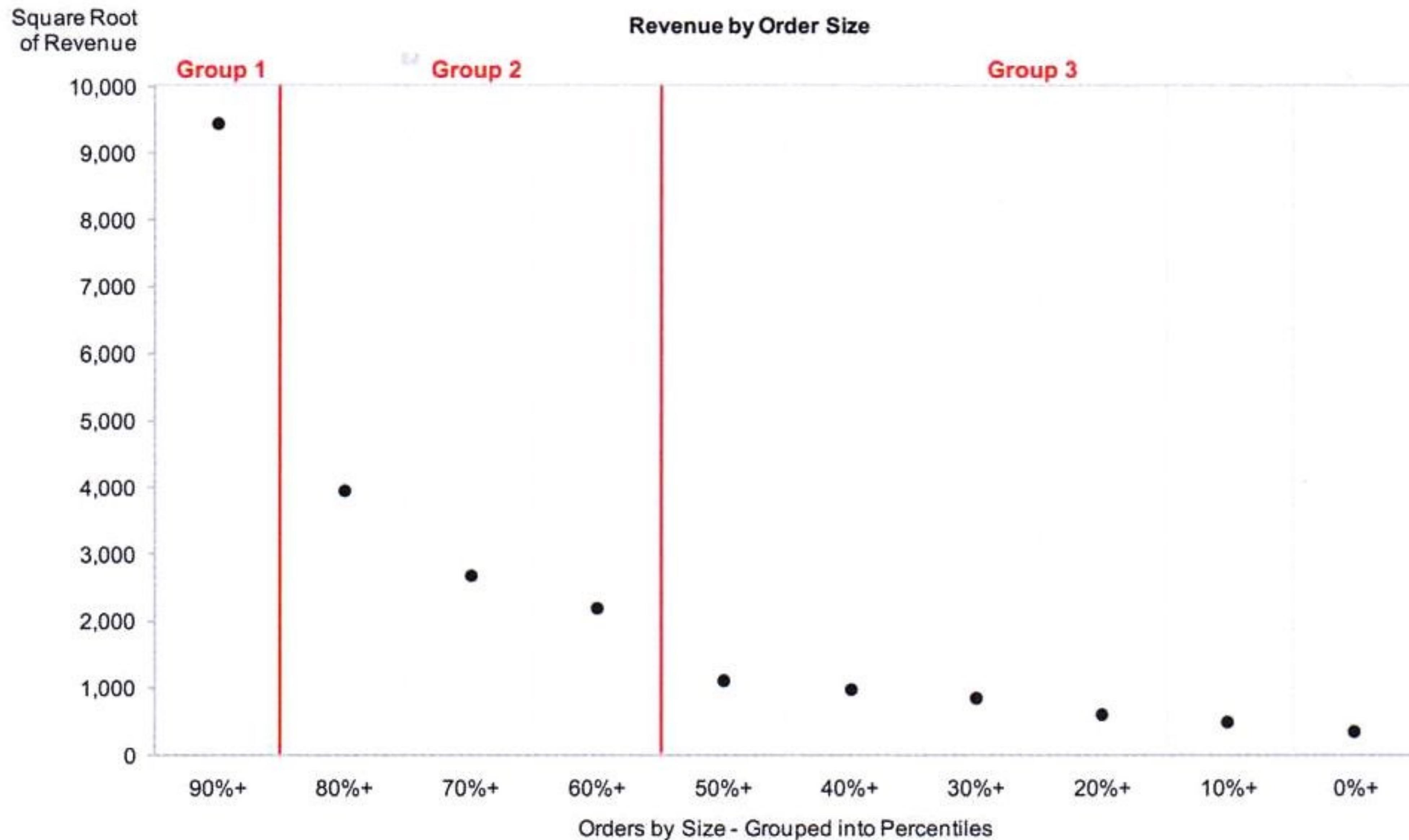


Exiba as barras com tamanho muito diferentes em gráficos separados

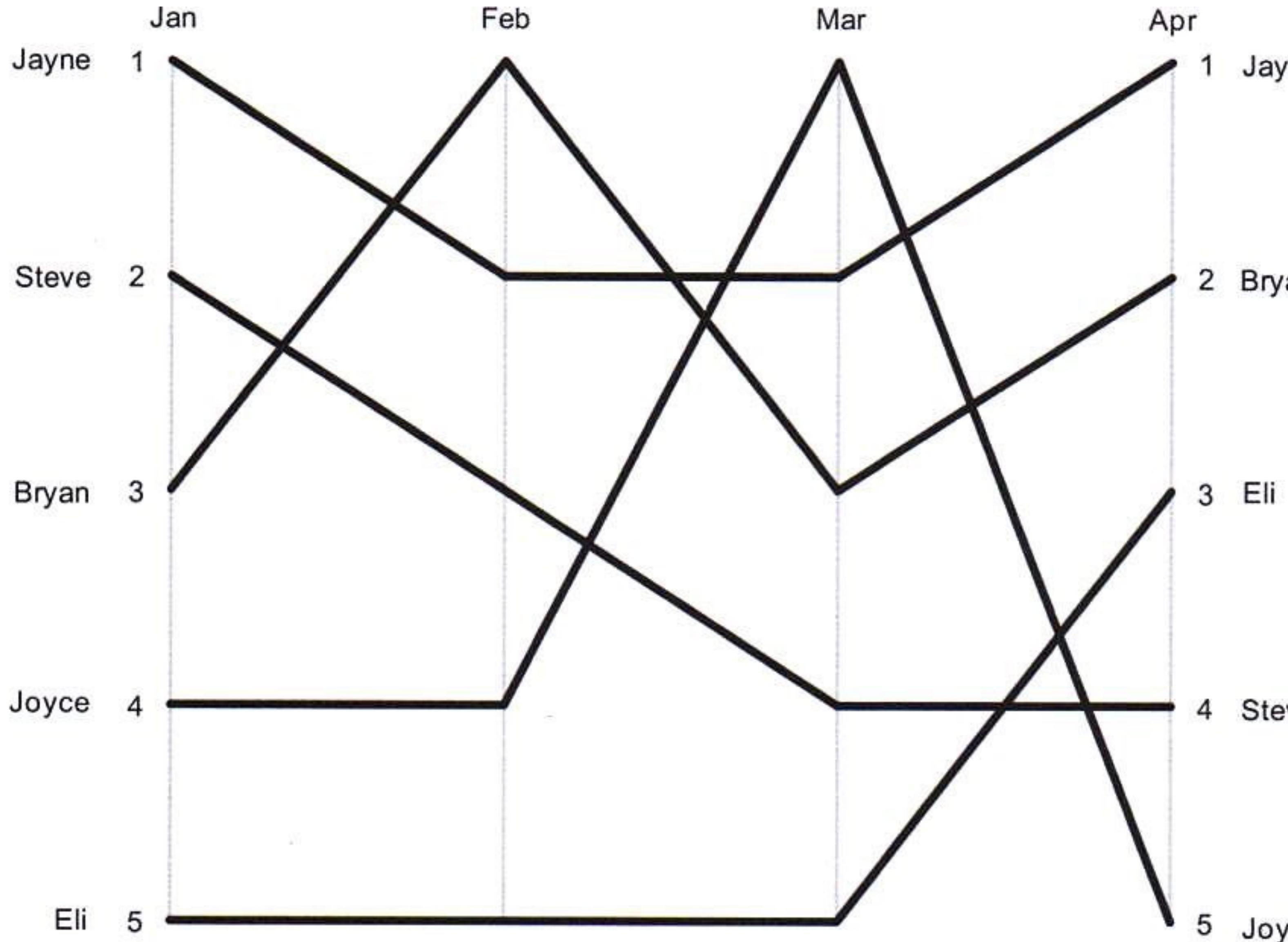
TESTE DIFERENTES ESCALAS - LOGARÍTMICA



TESTE DIFERENTES ESCALAS - QUADRÁTICA



USE GRÁFICOS DE LINHAS PARA VISUALIZAÇÃO DE MUDANÇAS NOS RANKINGS AO LONGO DO TEMPO



USE GRÁFICOS DE LINHAS PARA VISUALIZAÇÃO DE MUDANÇAS NOS RANKINGS AO LONGO DO TEMPO

Division 2

19 X-Press 3

20 City 3

21 St Ives

22 Max'm Entropy

23 Champs

24 Cantabs 5

25 Champs 2

26 X-Press 4

27 Isle of Ely

28 Cantabs 6

29 Simoco 2

30 '99 5

31 Champs 3

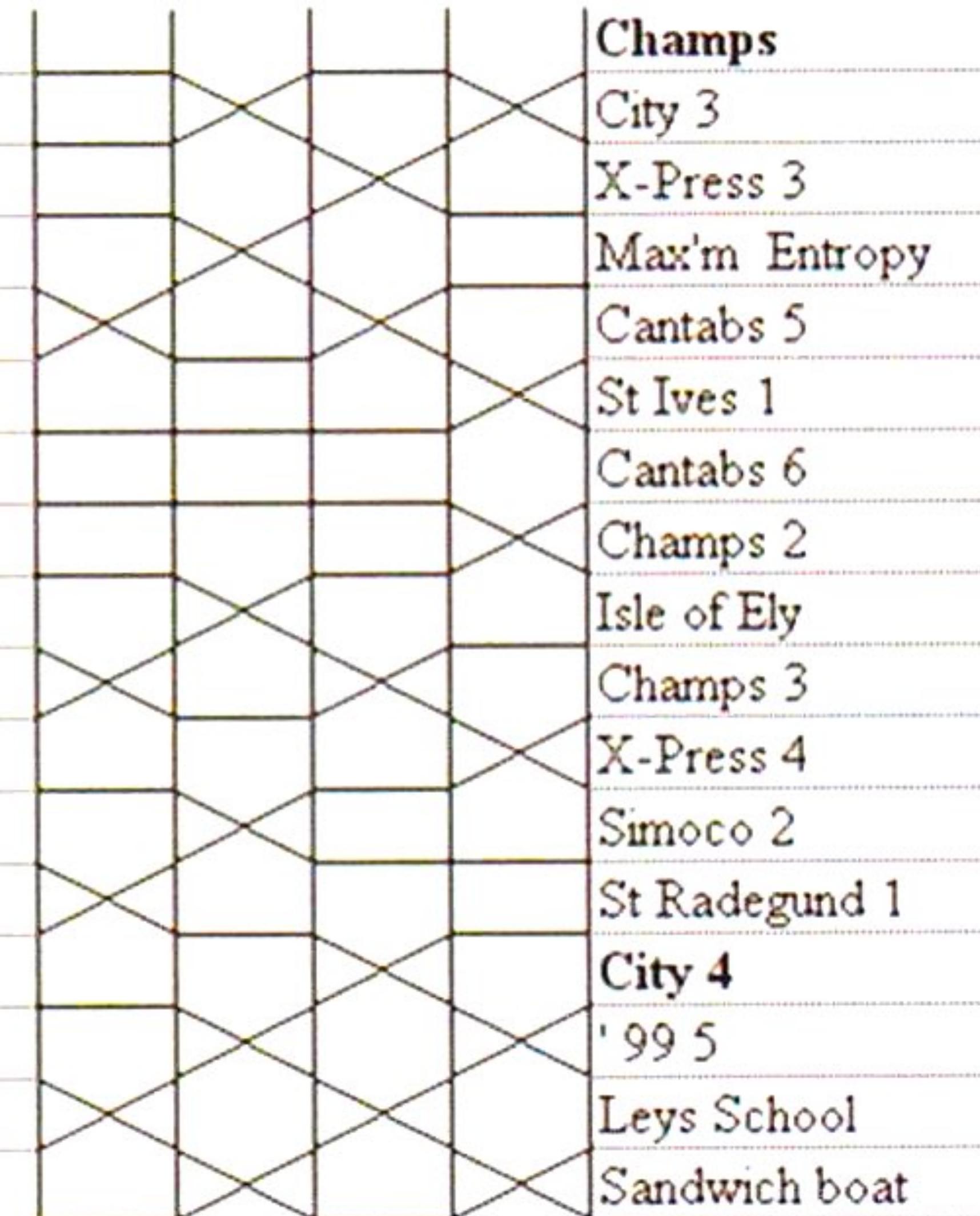
32 Camb Veterans

33 Cantabs 7

34 St Radegund

35 Sandwich boat

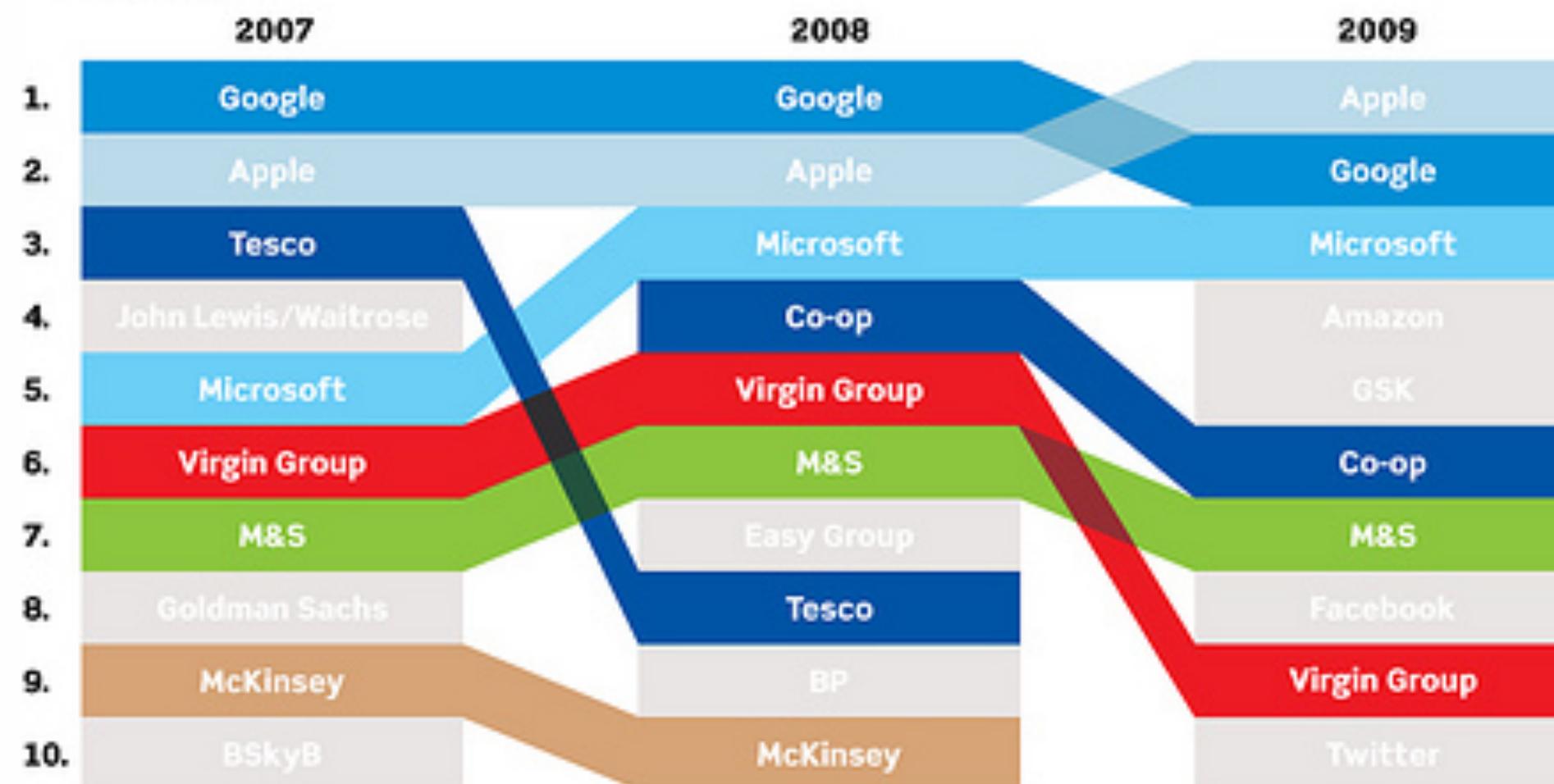
Tue Wed Thur Fri



USE GRÁFICOS DE LINHAS PARA VISUALIZAÇÃO DE MUDANÇAS NOS RANKINGS AO LONGO DO TEMPO

MOST HIGHLY-REGARDED BRANDS BY UK'S PROMINENT LEADERS

>BUSINESSES



>NON-PROFIT ORGANISATIONS

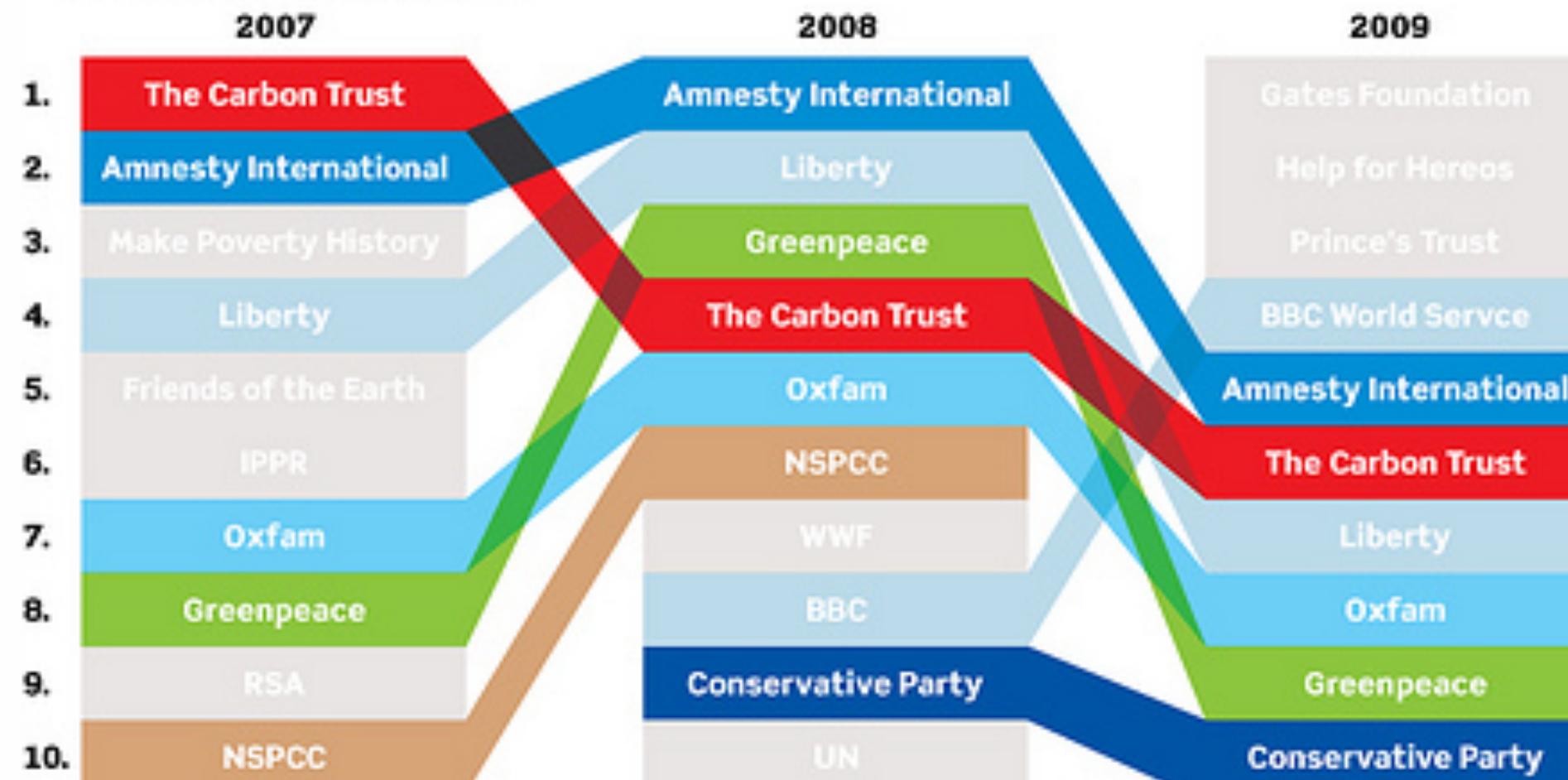
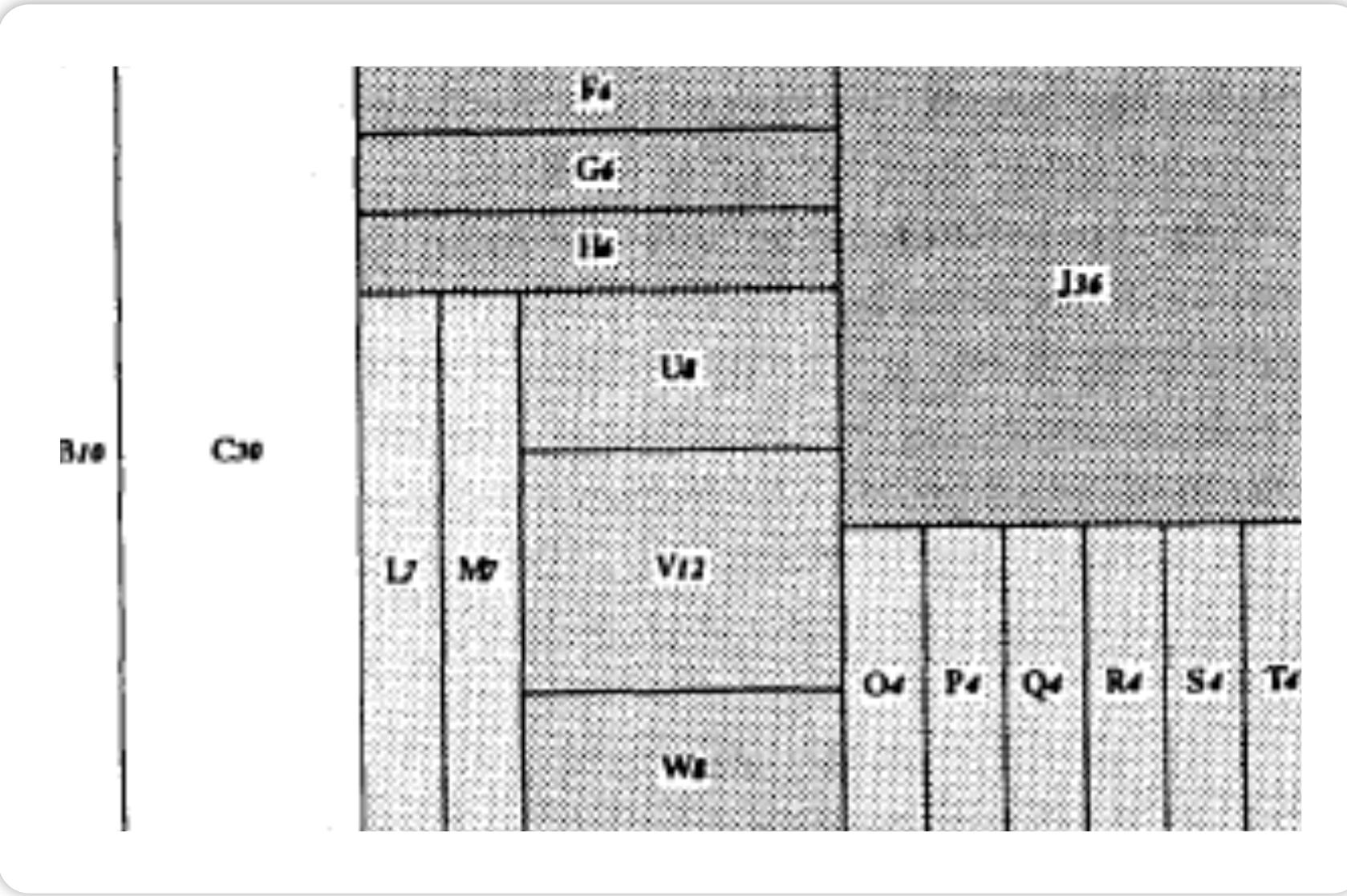


Chart showing the top ten brands' standing over the last three years

TREE-MAPS: A SPACE FILLING APPROACH TO THE VISUALIZATION OF HIERARCHICAL INFORMATION STRUCTURES

B. Johnson e B. Schneiderman
Proceedings of the 2nd Conference on Visualization
1991



ALGORITMOS

TREEMAPS

- Motivação
 - Visualizar estruturas de diretórios e seus tamanhos
- Objetivos
 - Utilização eficiente do espaço
 - Interatividade
 - Compreensão
 - Estética

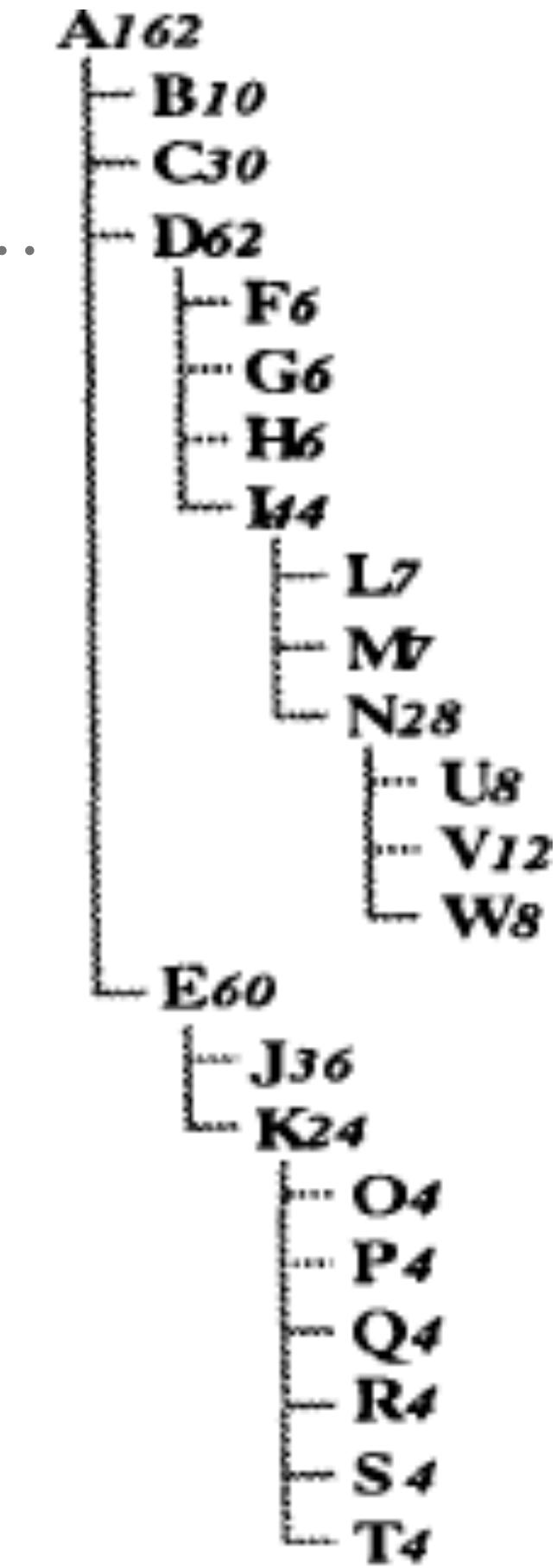


Figure 1. Outline

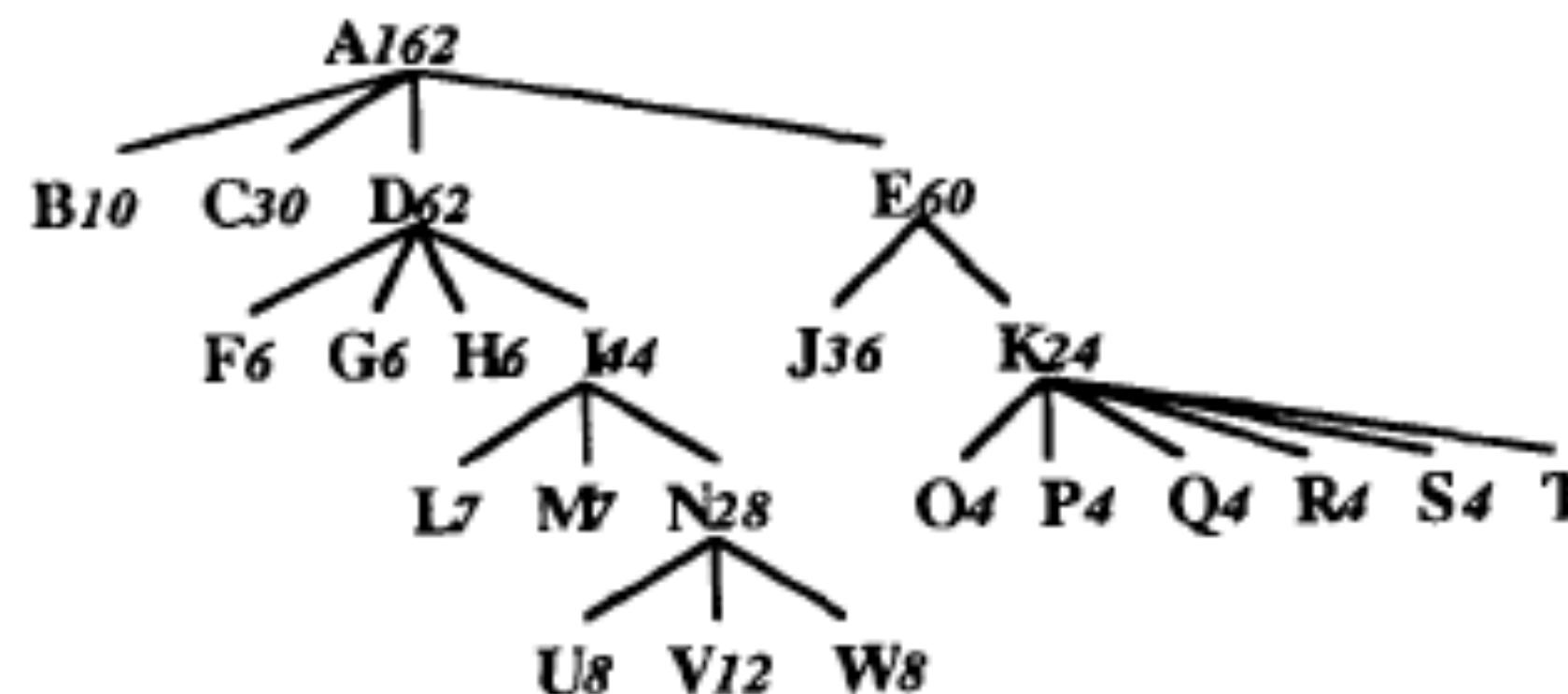


Figure 2. Tree Diagram

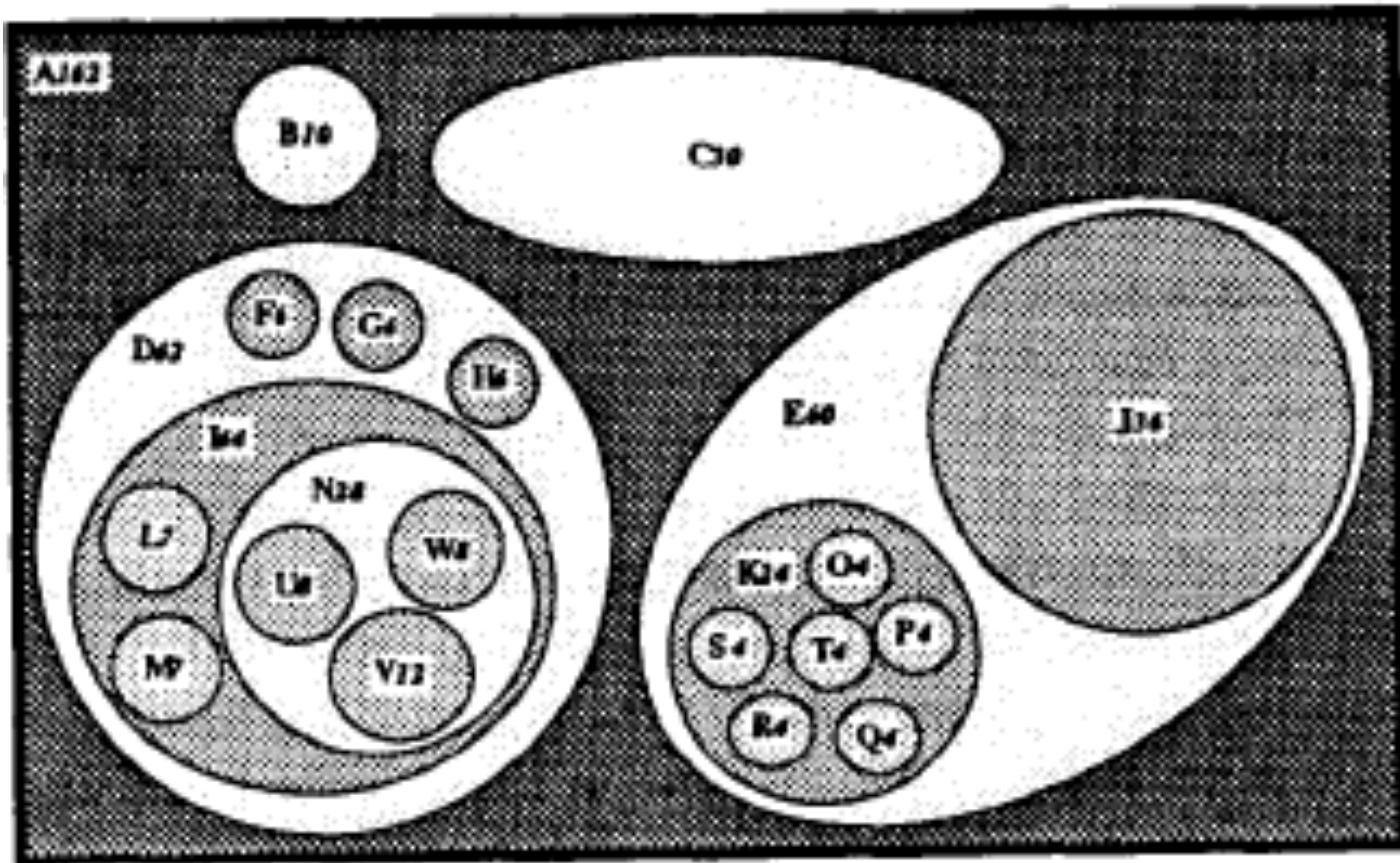


Figure 3. Venn Diagram

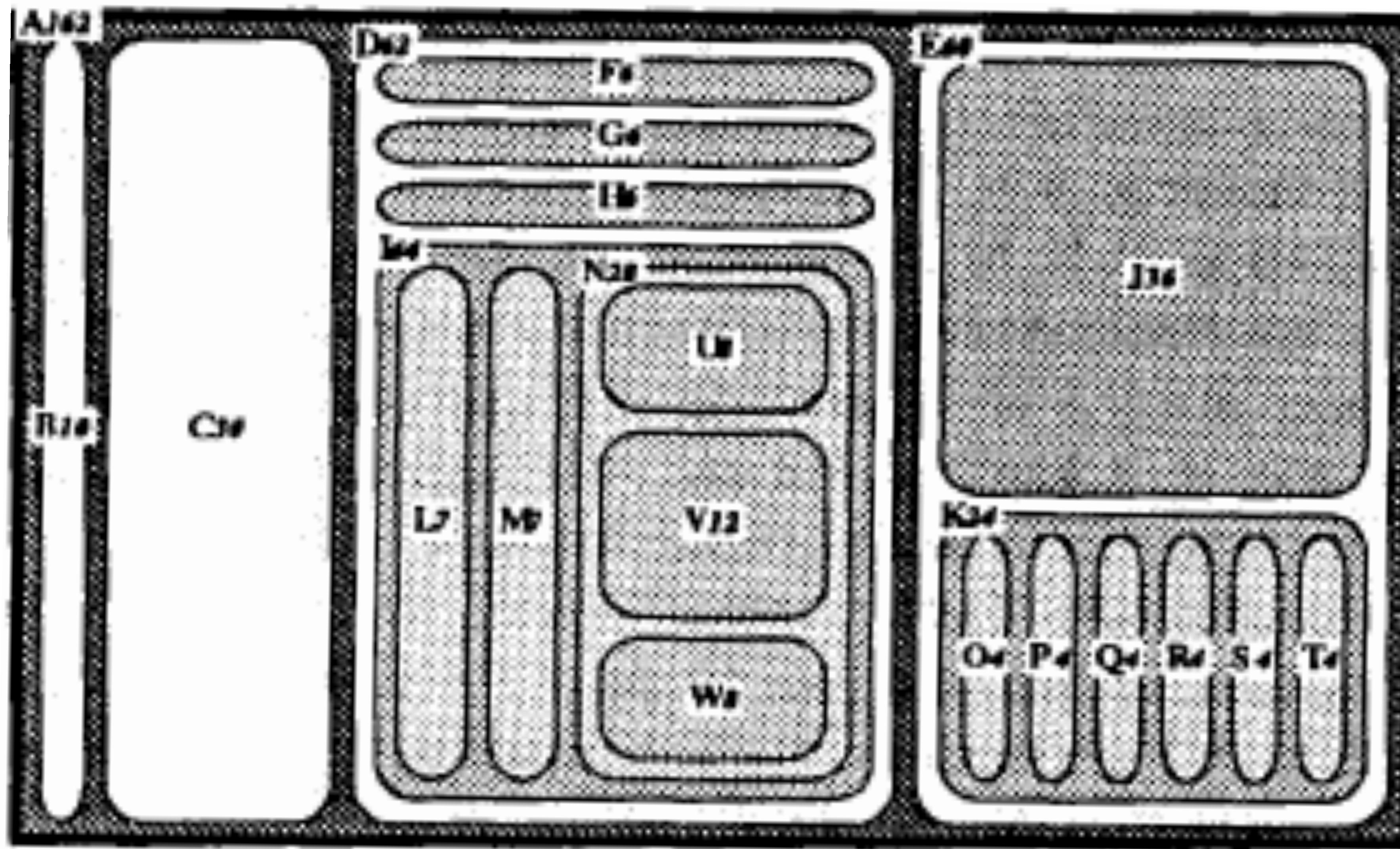


Figure 4. Nested Tree-Map

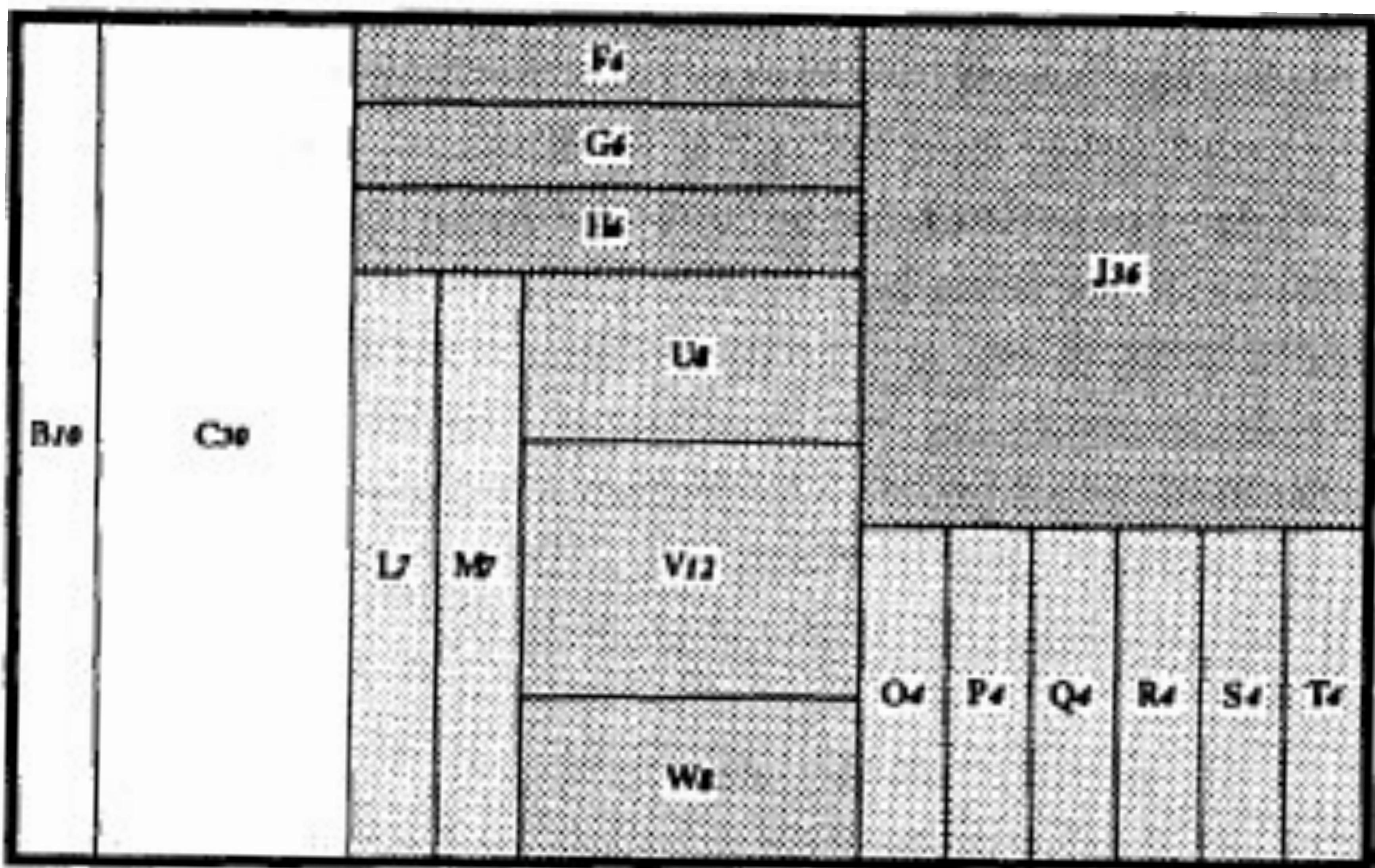


Figure 5. Tree-Map

REPRESENTAÇÃO VISUAL

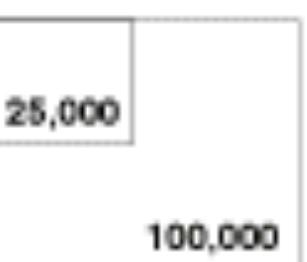
- Cada nó folha é representado por um retângulo
 - A área de cada retângulo é proporcional ao peso do nó que representa
 - Se o nó I é ancestral do nó II, então o retângulo que representa o nó I contém área de nó II
 - Os retângulos de dois nós possuem interseção se, e somente se, um é ancestral do outro
 - O peso de um nó tem valor igual à soma de seus filhos

Truck Sales Slip, Tripping Up Chrysler

Over the past few years, Chrysler executives said they were following the lead of Toyota and Honda, focusing on vehicles that met the needs of their customers. But as American consumers turned away from large trucks and S.U.V.'s in 2006, Chrysler continued to churn out big vehicles, which are now sitting unsold at dealerships across the country.

READING THE CHART

Boxes are scaled proportionally according to number of cars sold in 2006



Change in sale

Below

-10%

-2%

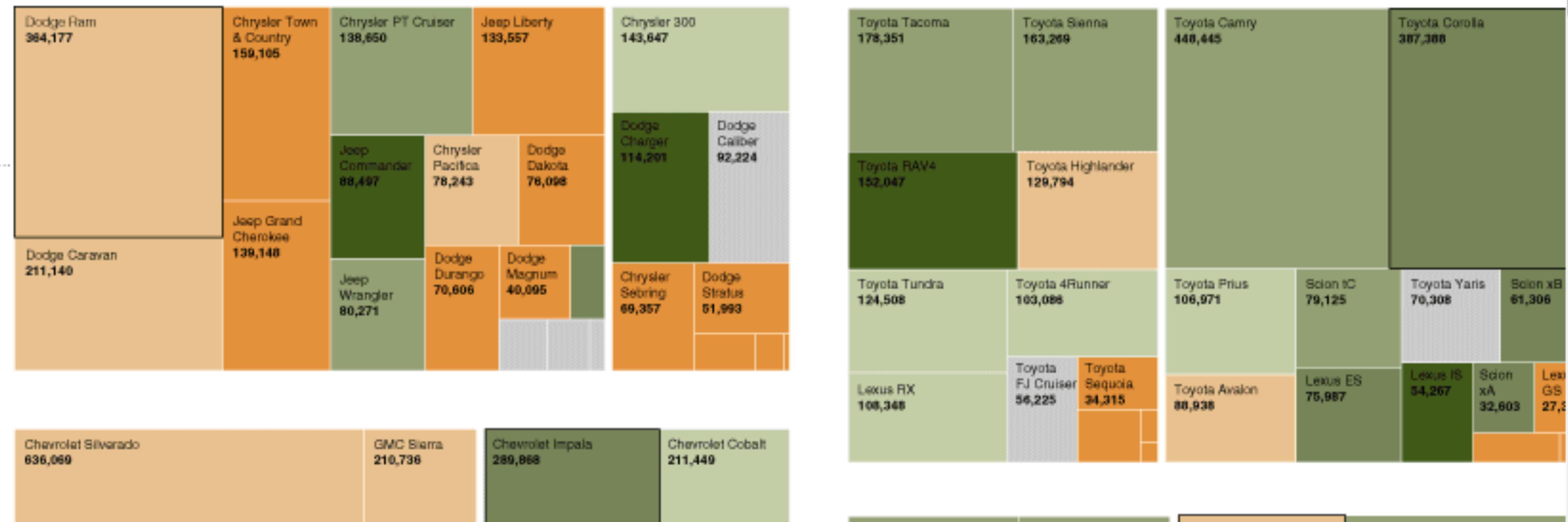
SALES CHANGE
'05 TO '06

Chrysler Group	-7.0%
Trucks/vans/S.U.V.'s	1.6 million
Cars	0.5 million

Pickups, minivans and S.U.V.'s made up 76 percent of Chrysler's sales, which left it vulnerable when consumers shifted to cars.

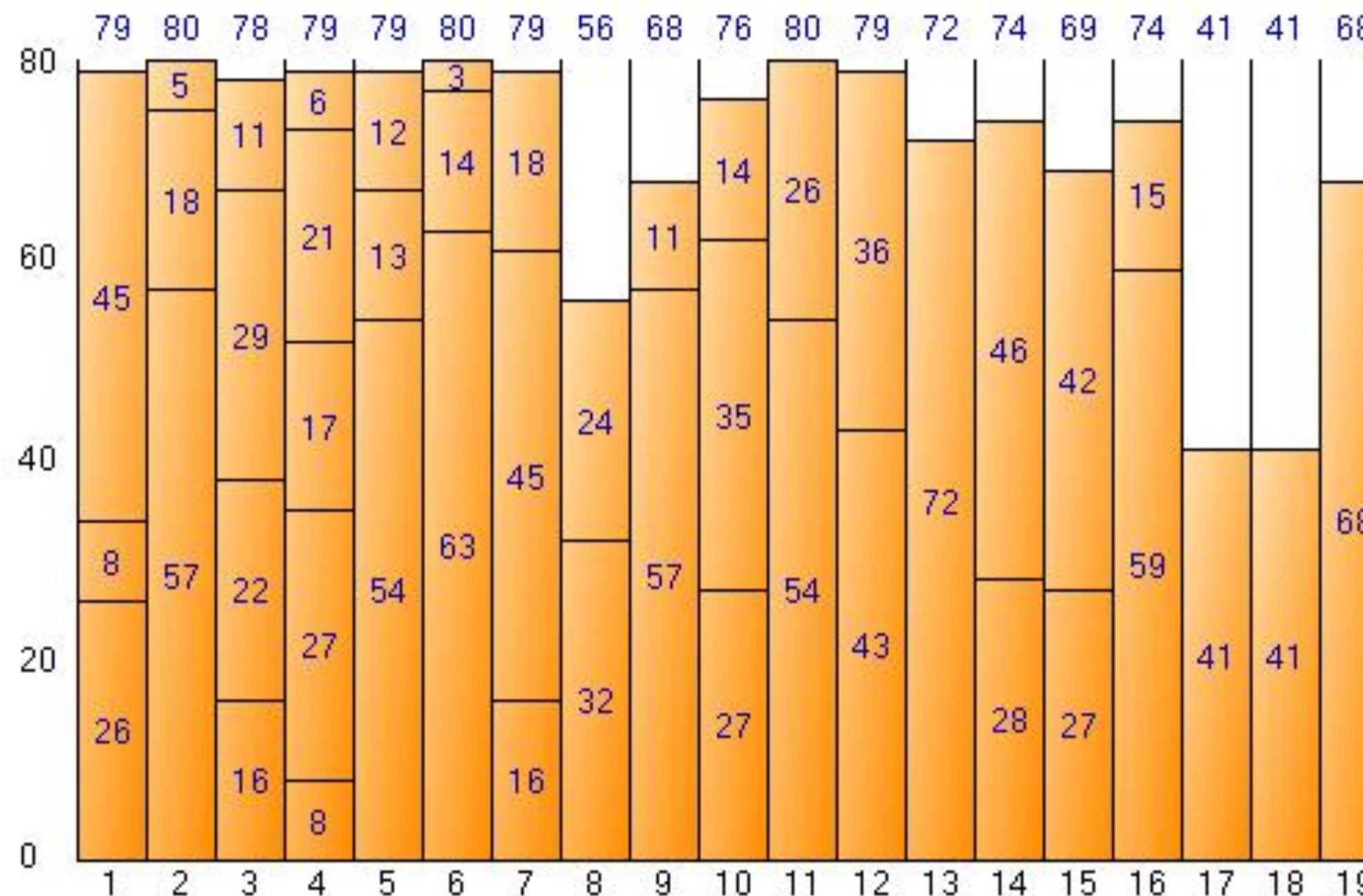


General Motors	-8.7%
Trucks/vans/S.U.V.'s	2.5 million
Cars	1.6 million



Variação do problema bin packing que é NP-difícil

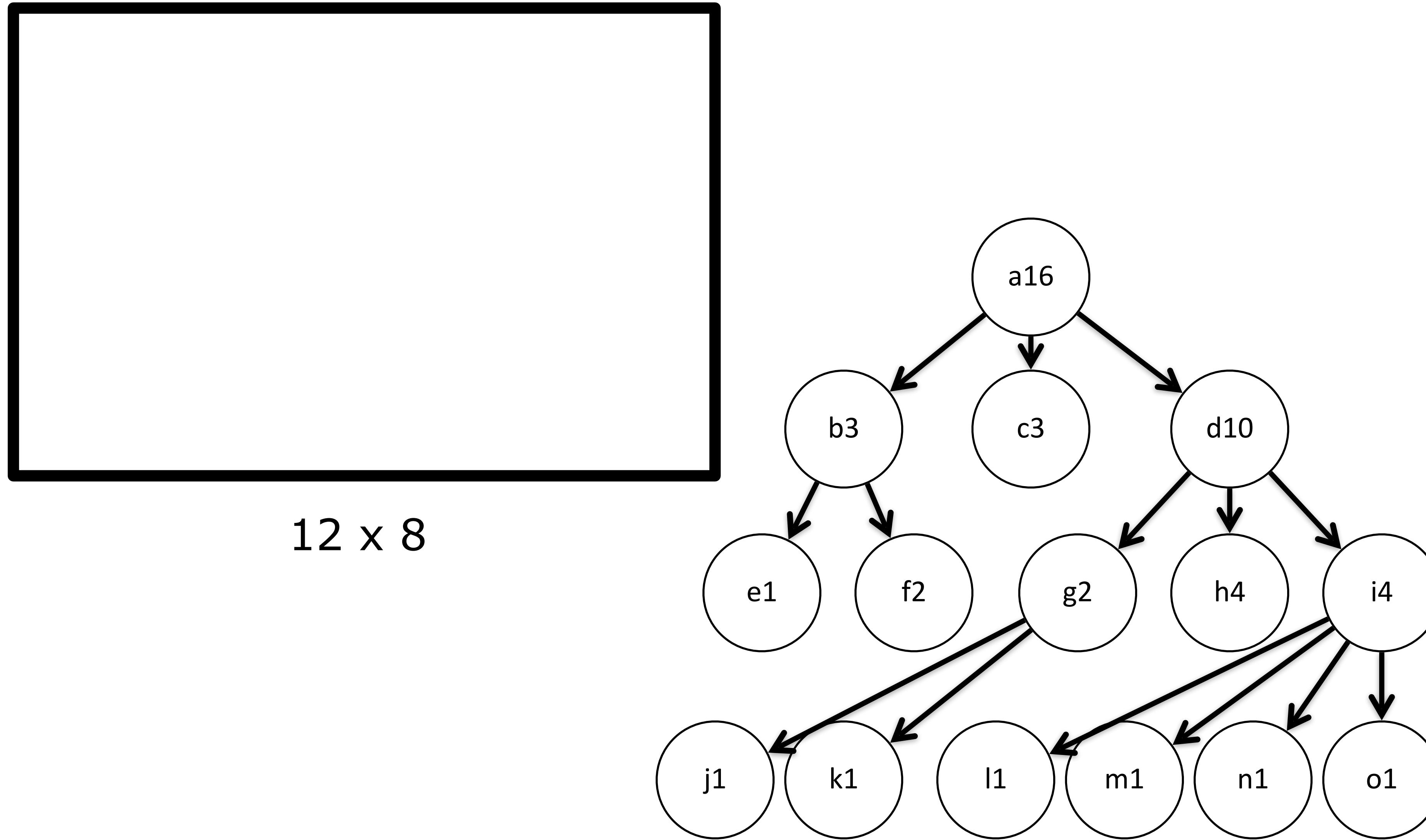
Objetos de diferentes volumes devem ser colocados em um número finito de compartimentos de capacidade V minimizando o número de compartimentos usados



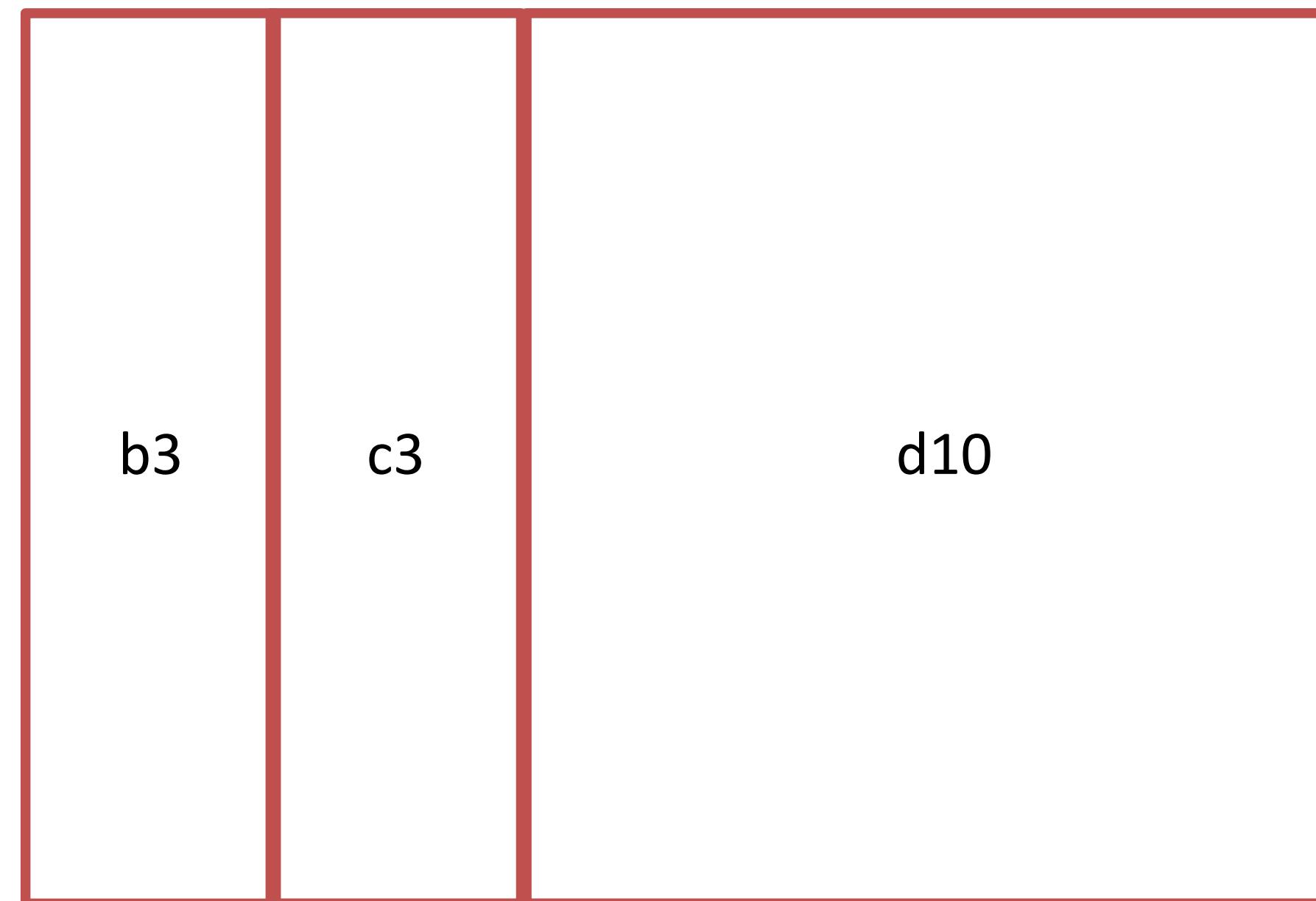
ALGORITMO

- Heurística *slice-and-dice* que pode ser aplicada através do caminhamento em pré-ordem em tempo $O(n)$

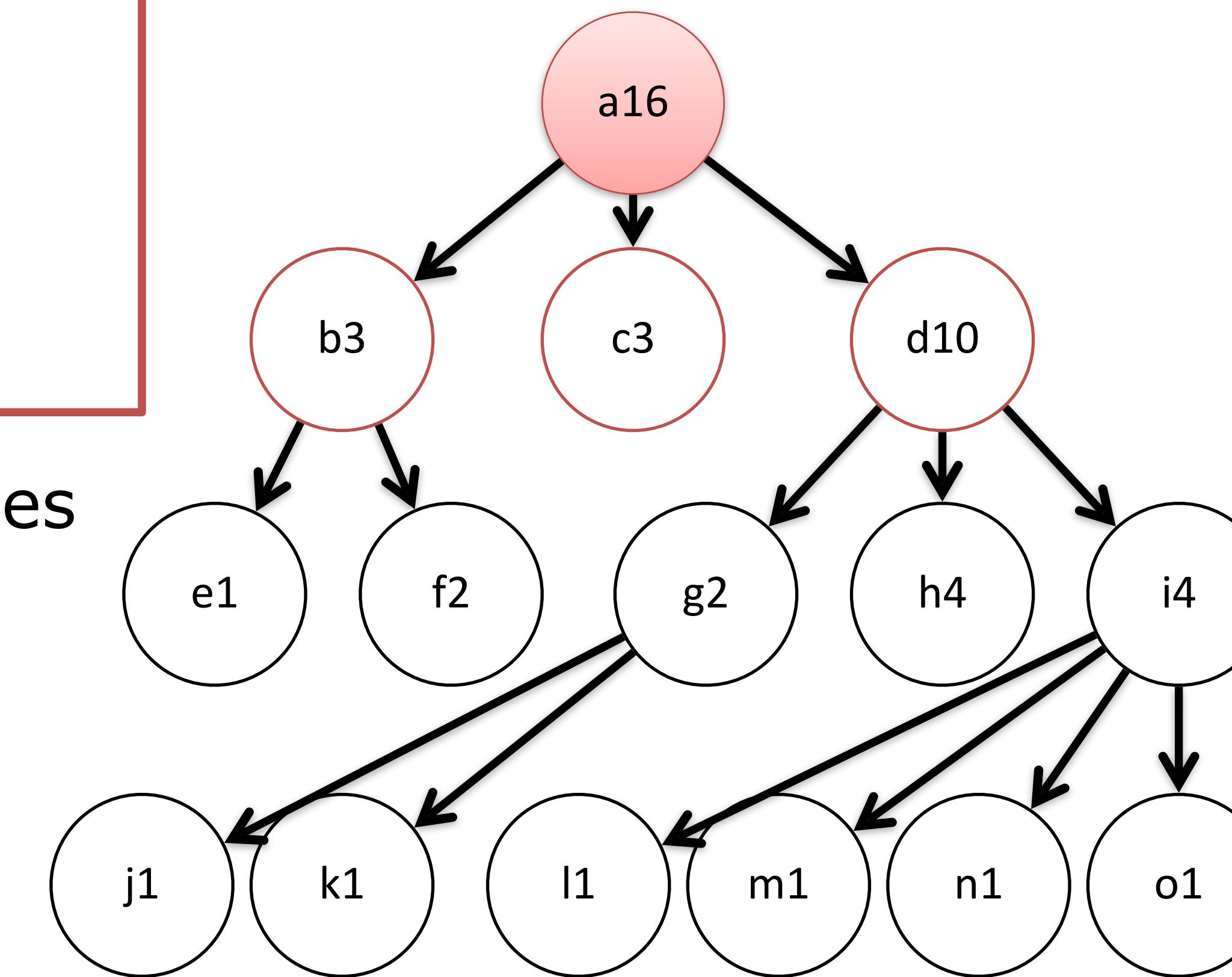
slice-and-dice



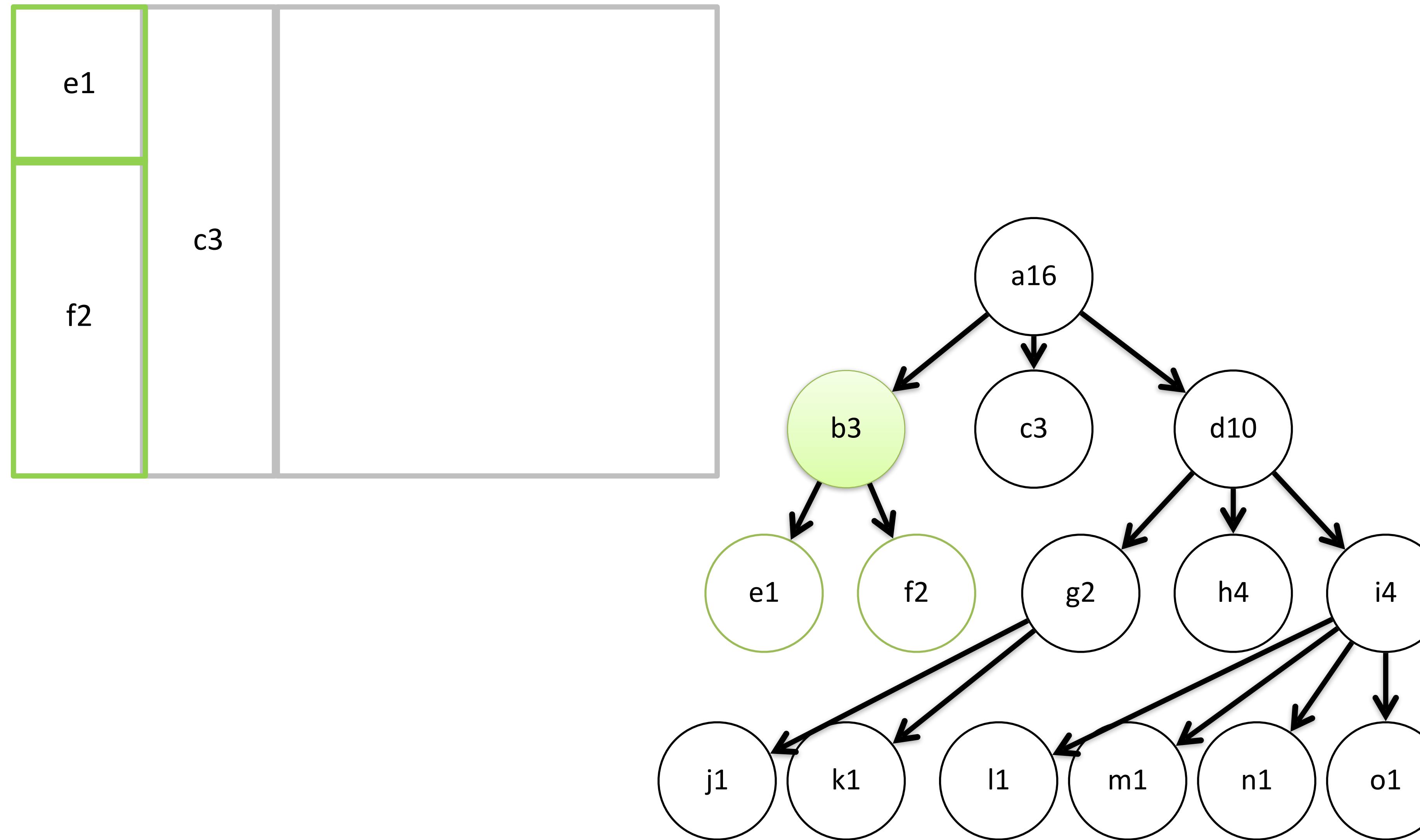
slice-and-dice



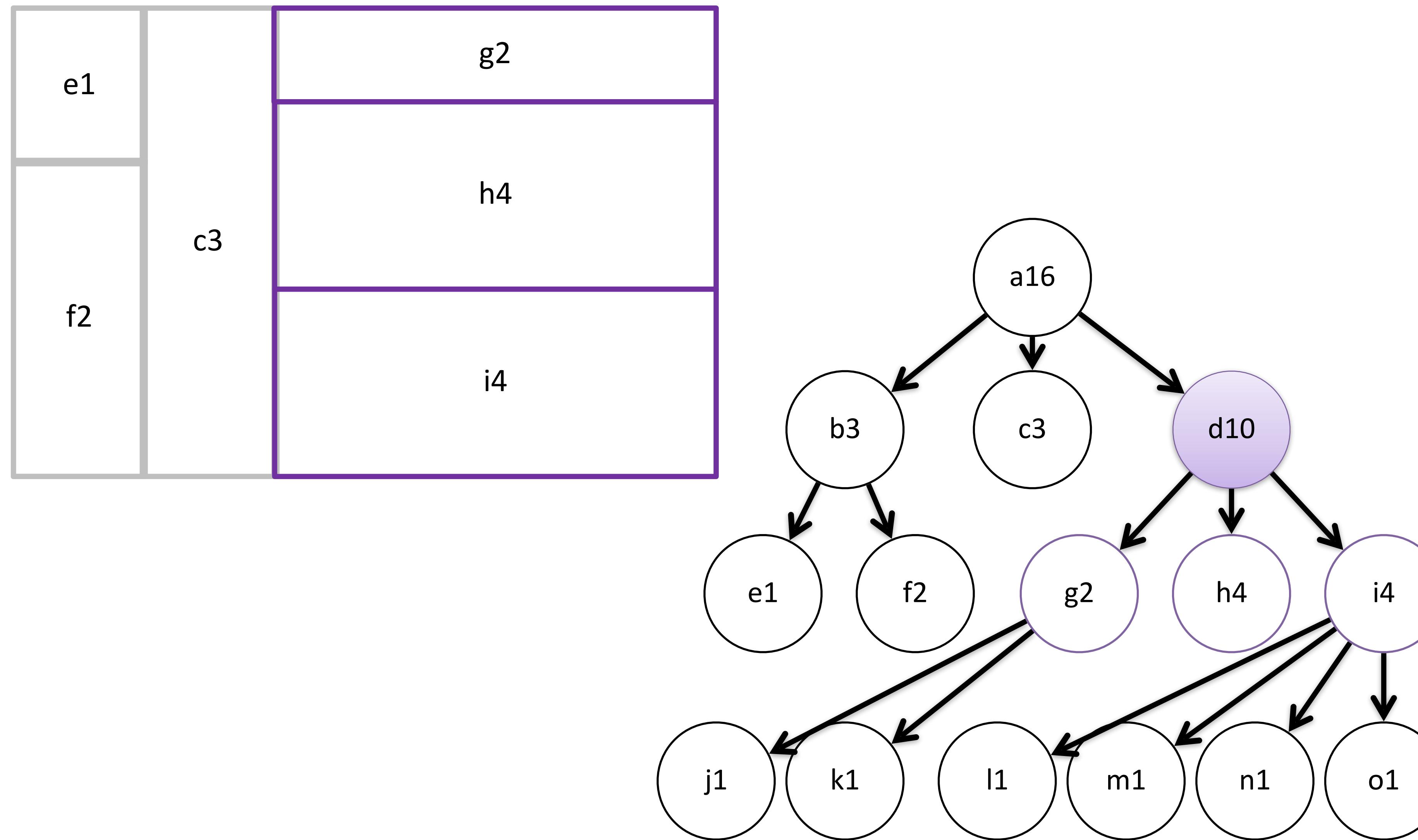
$12 \times 8 = 96$ para 16 unidades



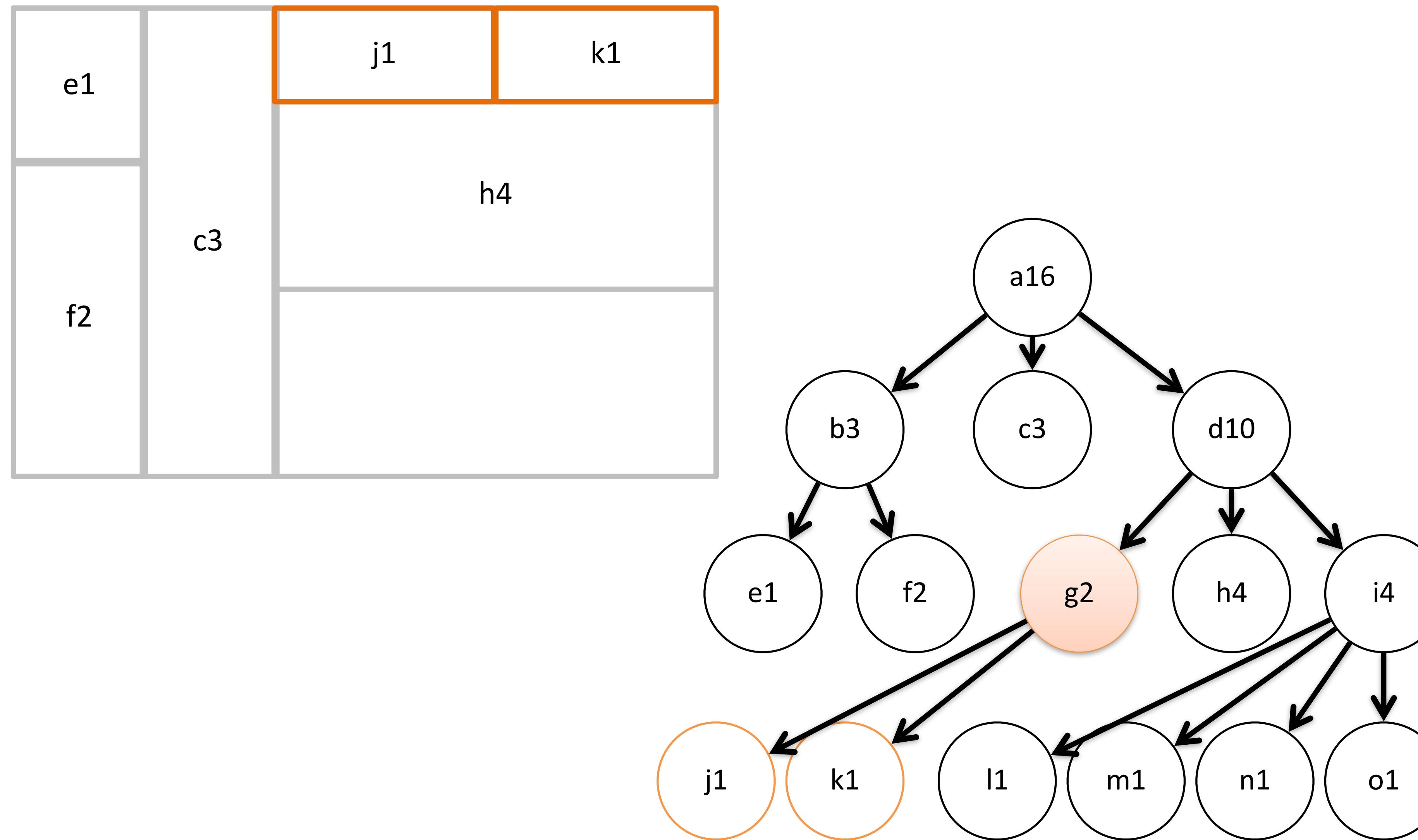
slice-and-dice



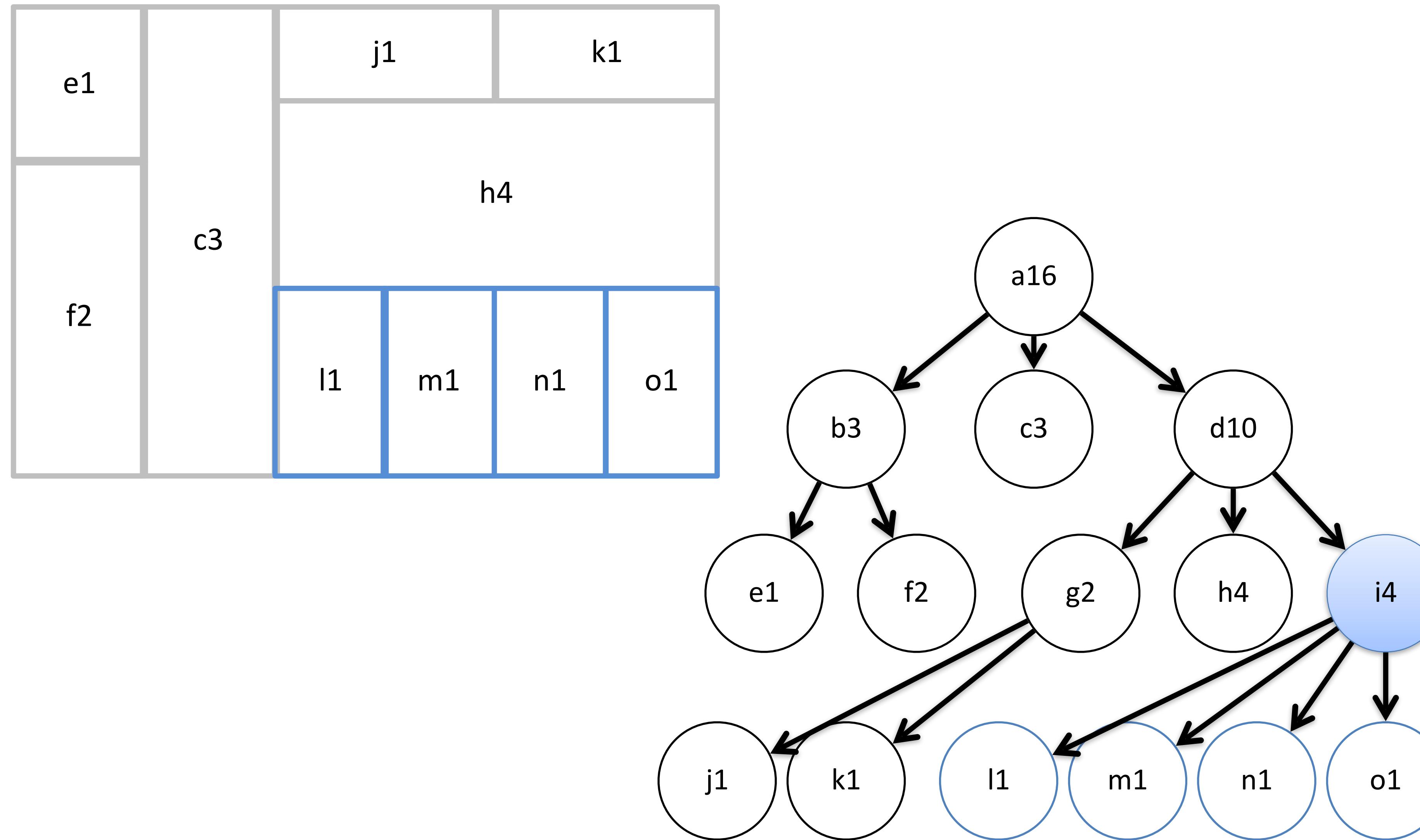
slice-and-dice



slice-and-dice



slice-and-dice



Desvantagens?

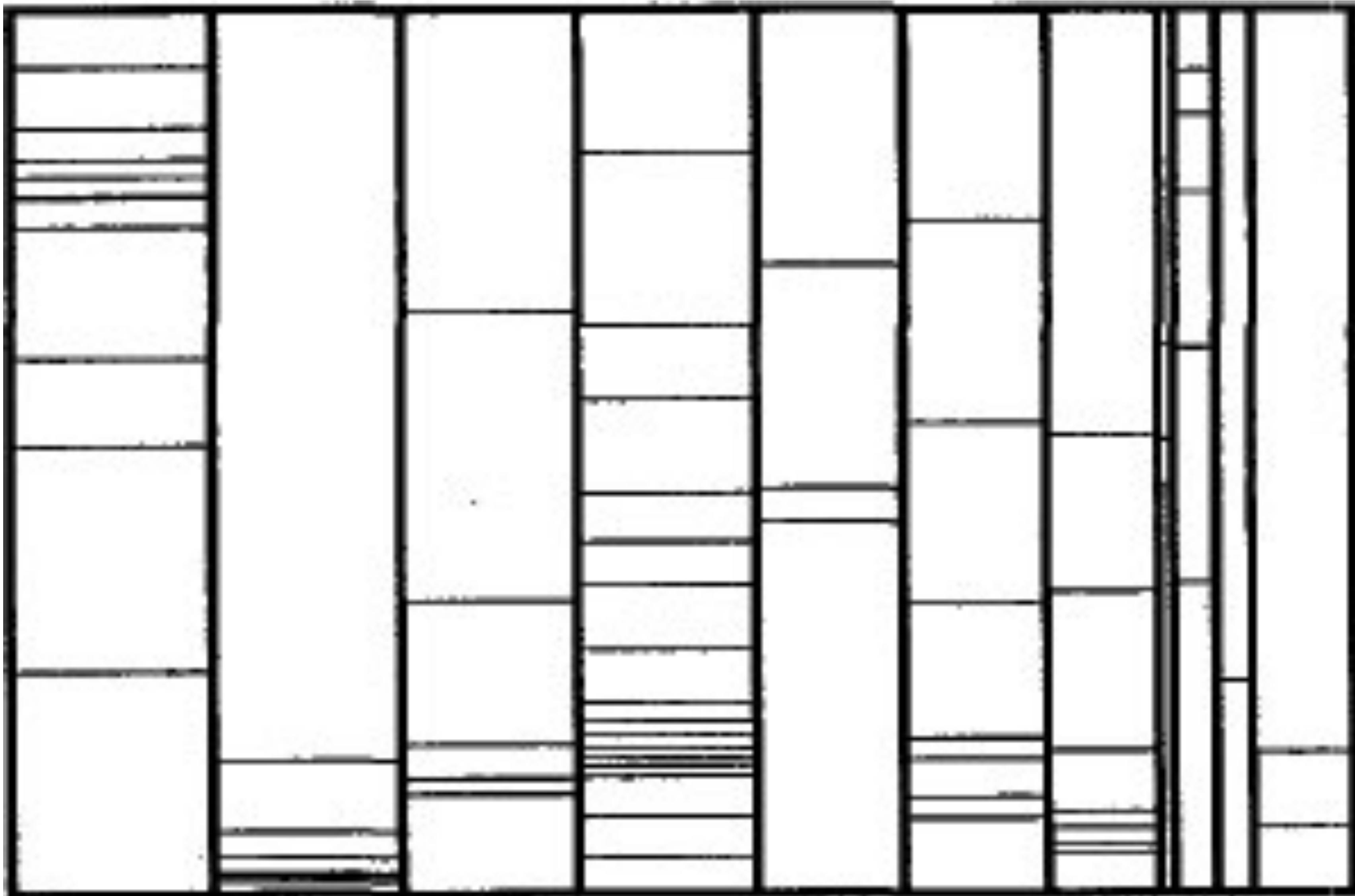
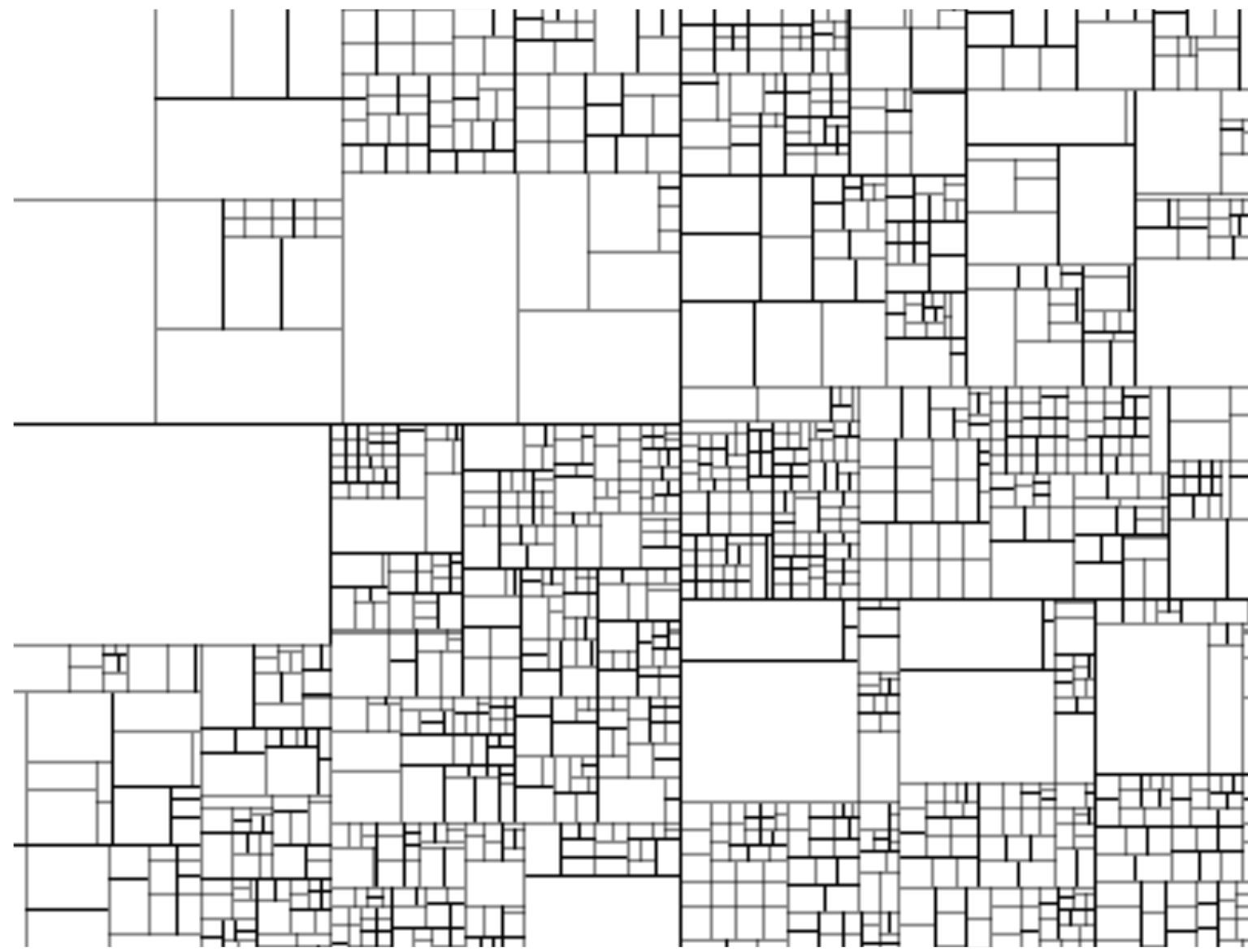


Figure 1: Typical slice-and-dice layout

SQUARIFIED TREEMAPS

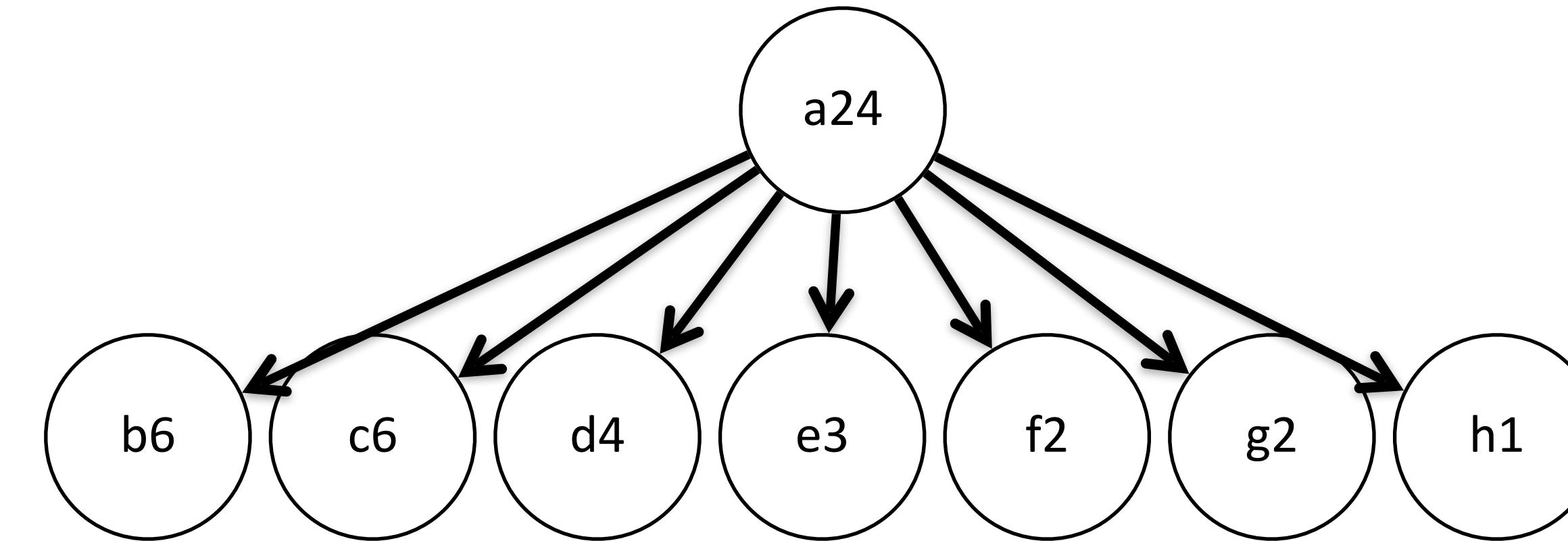
M. Bruls, K. Huizing e J.J. van Wijk
*Eurographics / IEEE Transactions on Visualization and
Computer Graphics*
2000



ALGORITMO

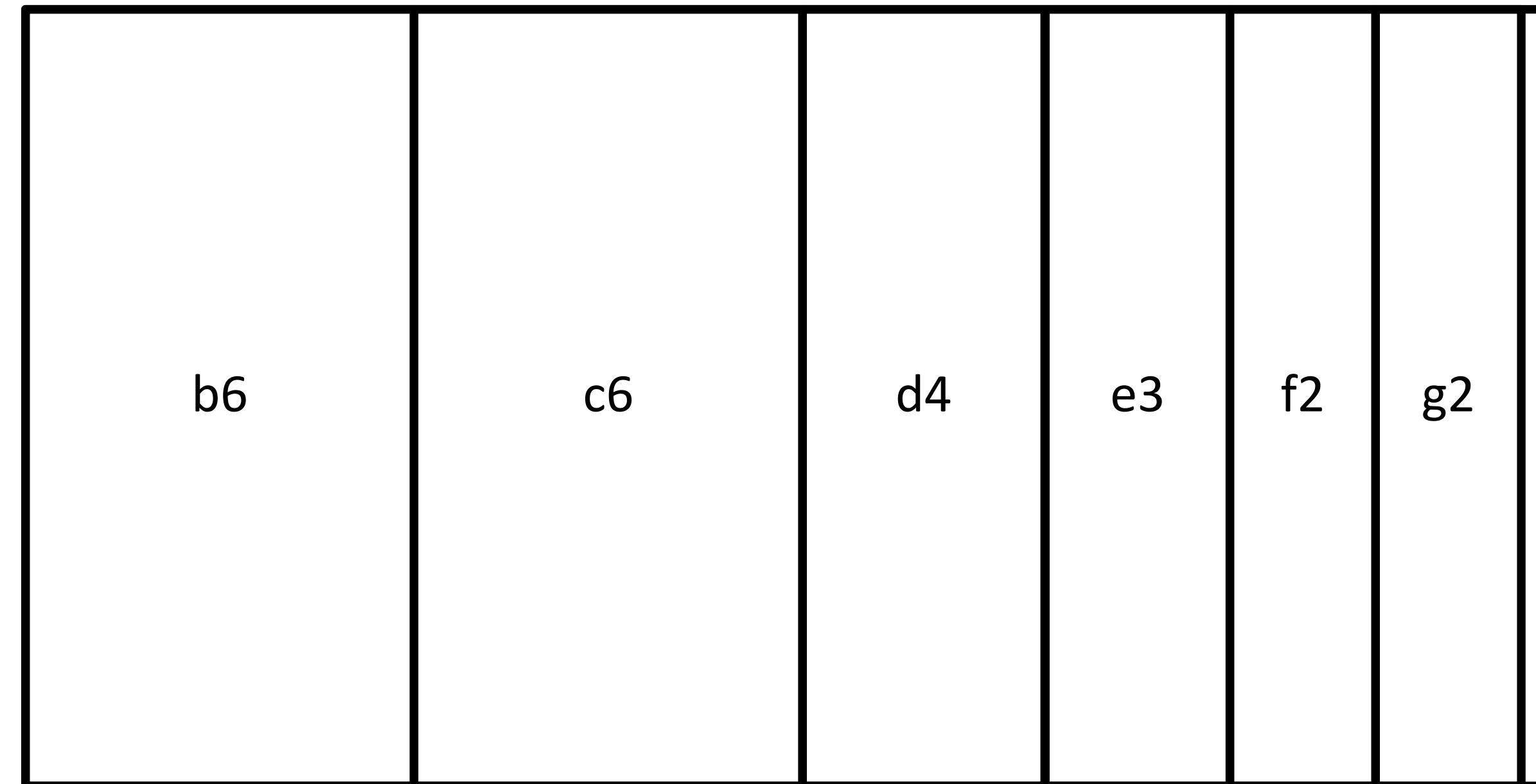
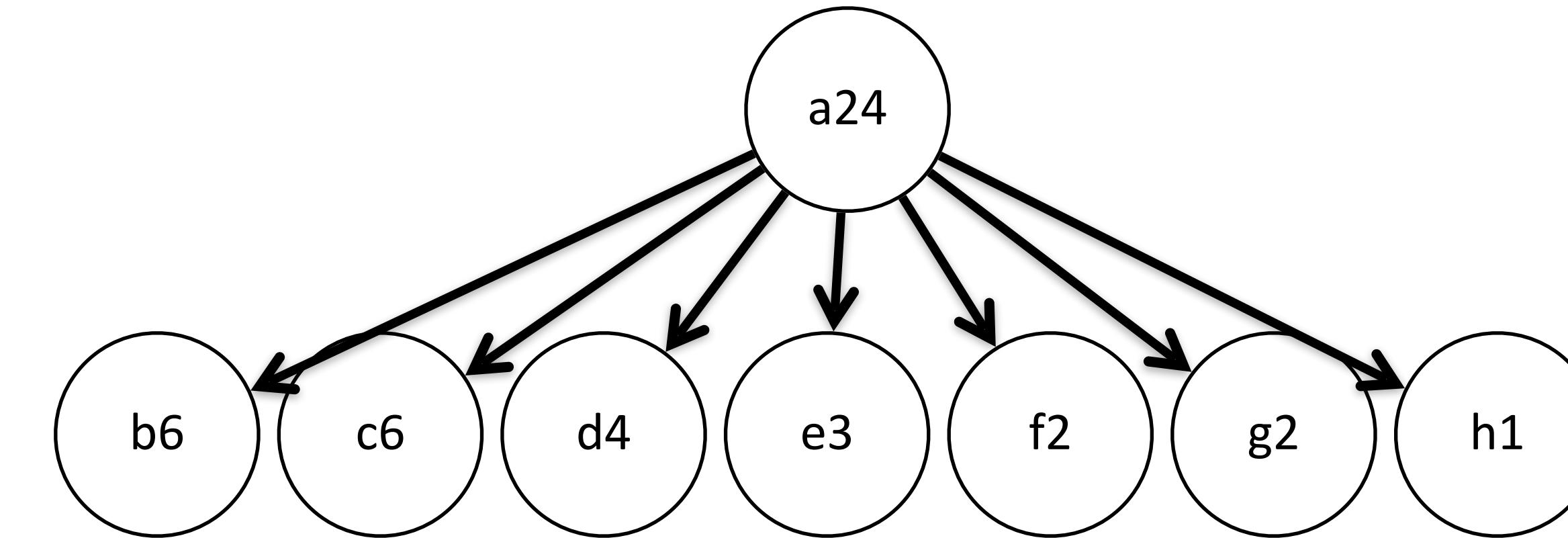
- O objetivo do artigo é produzir *Treemaps* com razão de aspecto próxima de 1
- Mais fácil de se estimar a área de quadrados que de retângulos
- Mais fácil interação uma vez que é mais fácil clicar em quadrados

squarified



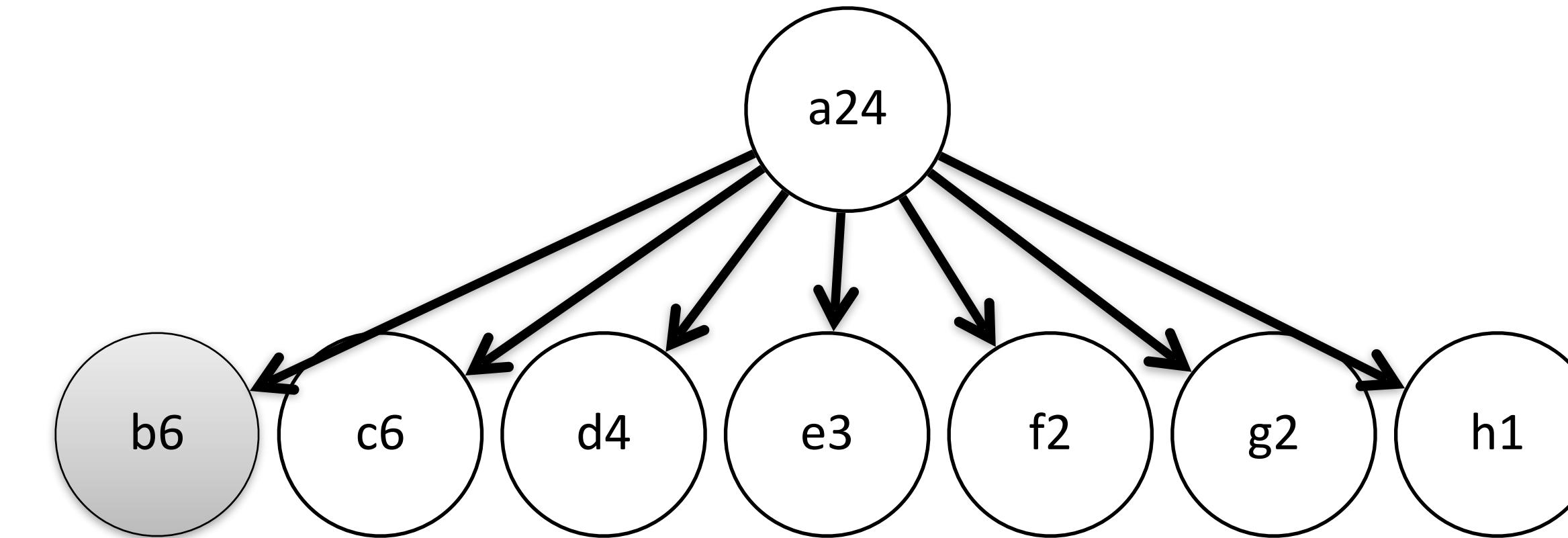
6 x 4

squarified



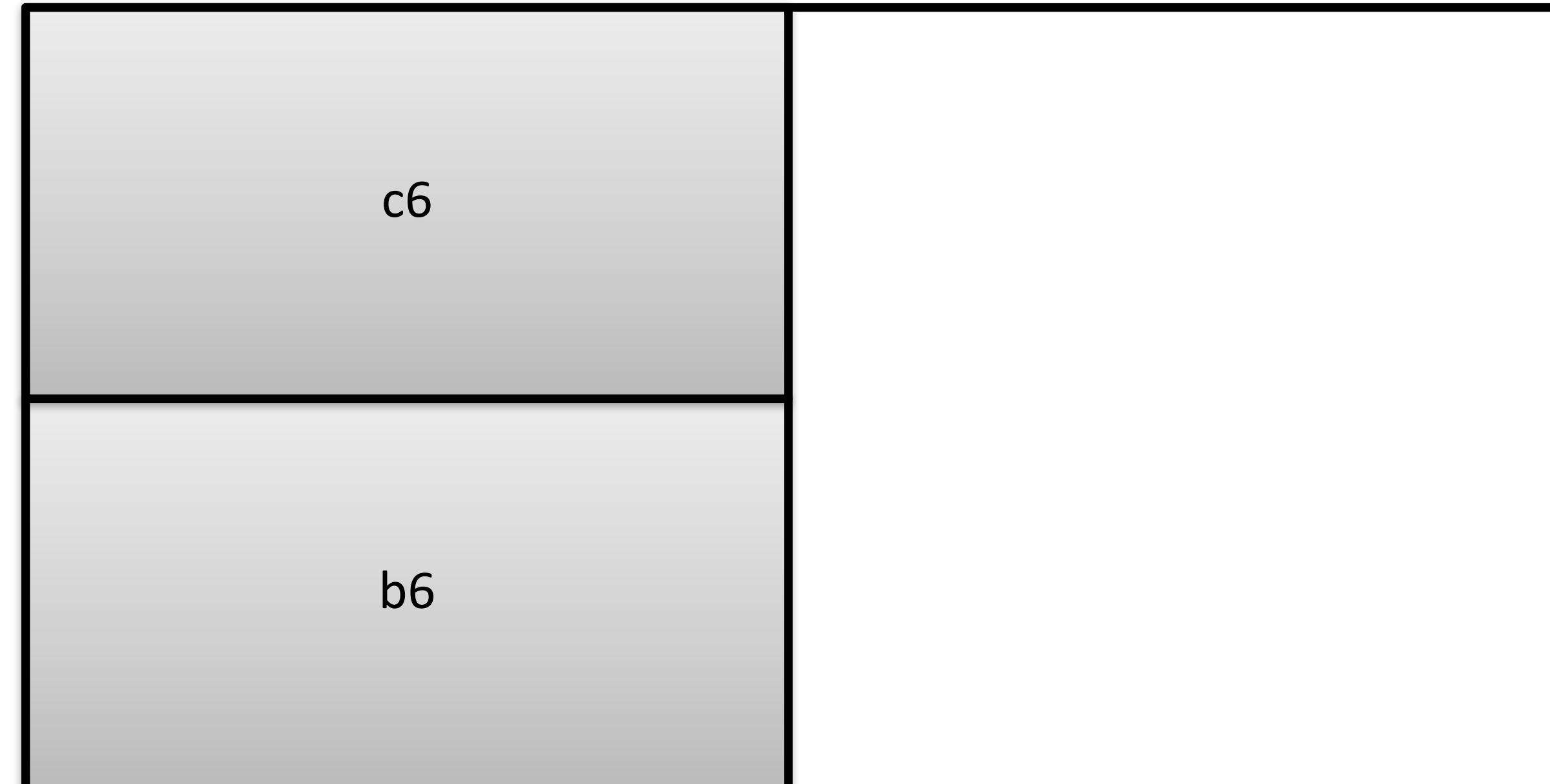
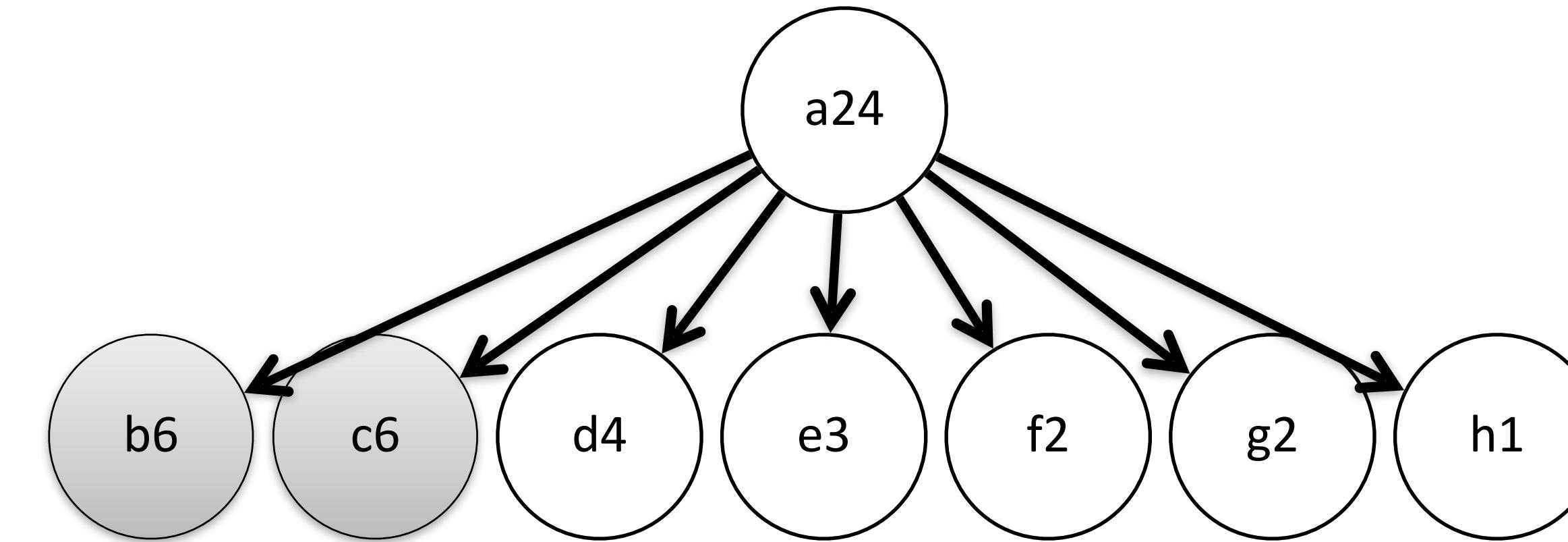
$$\text{Max}(\text{aspect ratio}) = \text{Max}(\text{width}/\text{height}, \text{height}/\text{width}) = 16$$

squarified



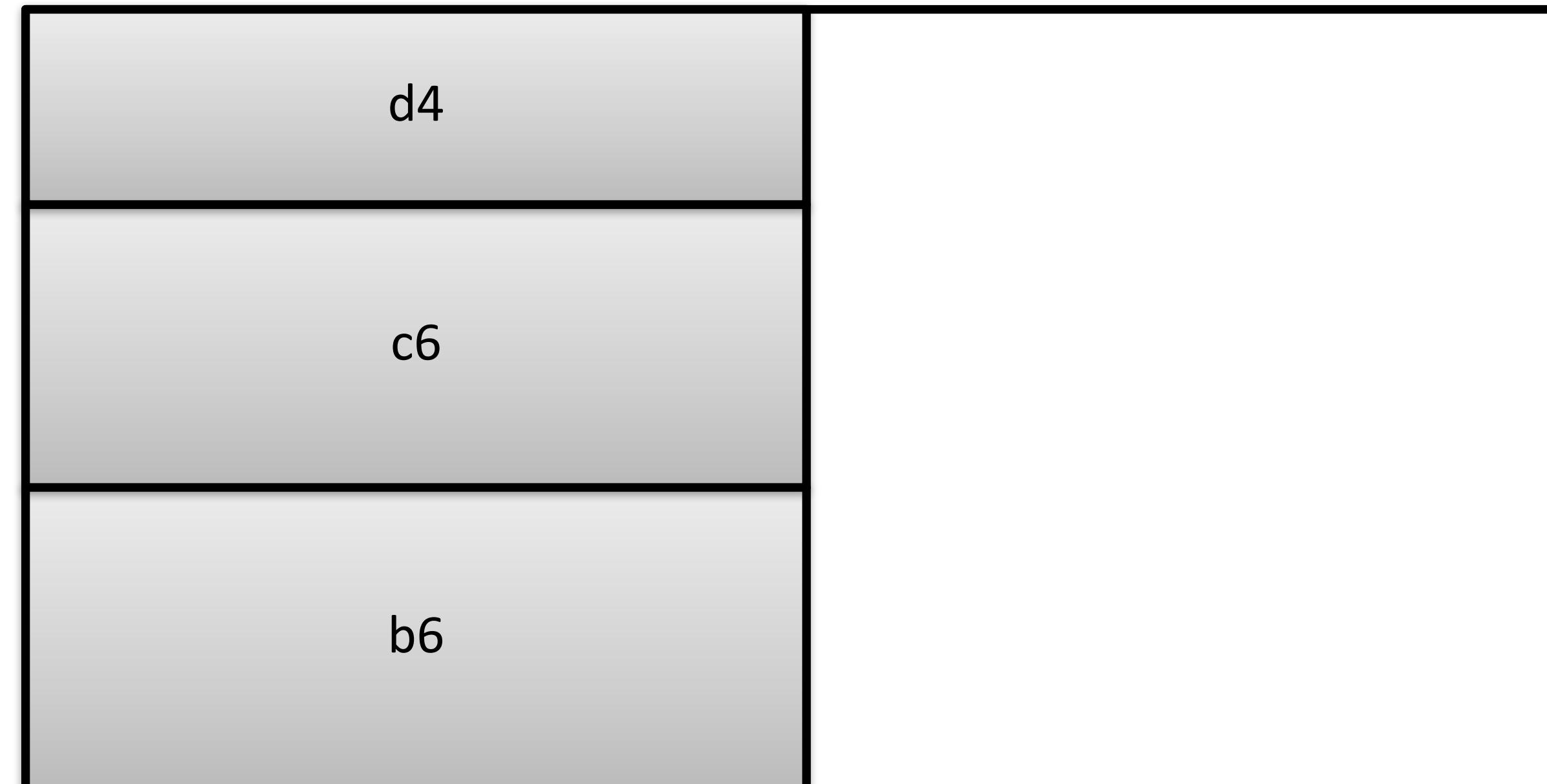
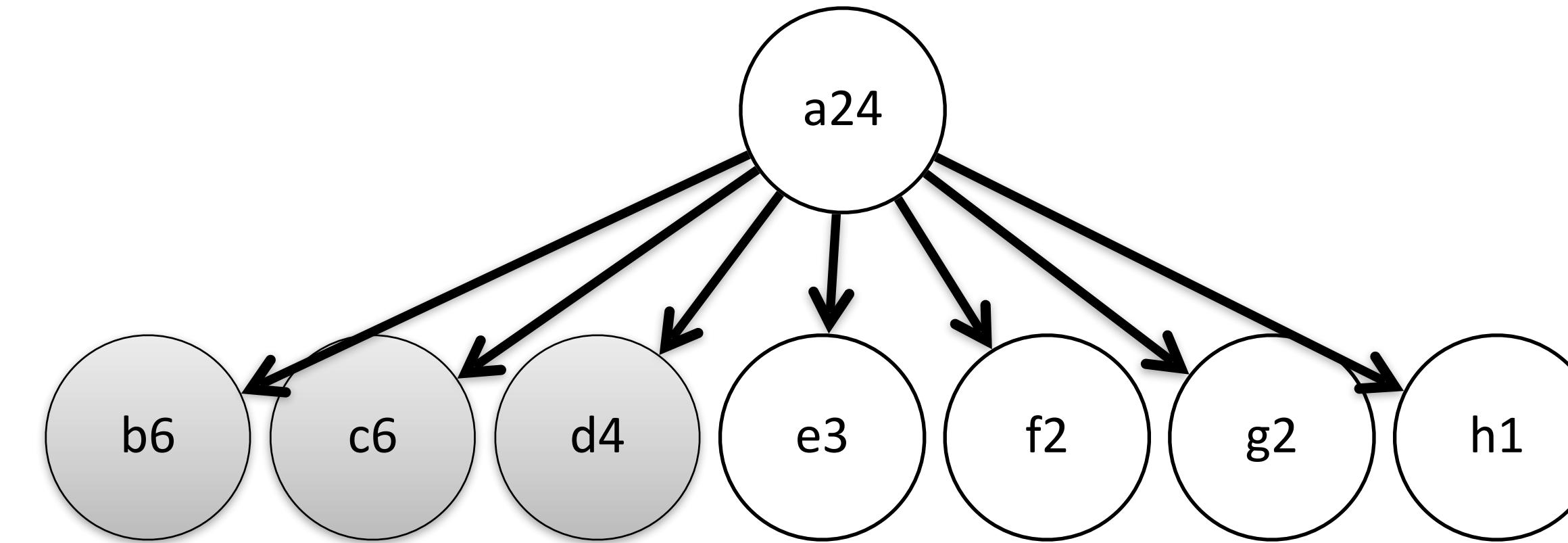
$$\text{Max(aspect ratio)} = 8/3 = 2,66$$

squarified



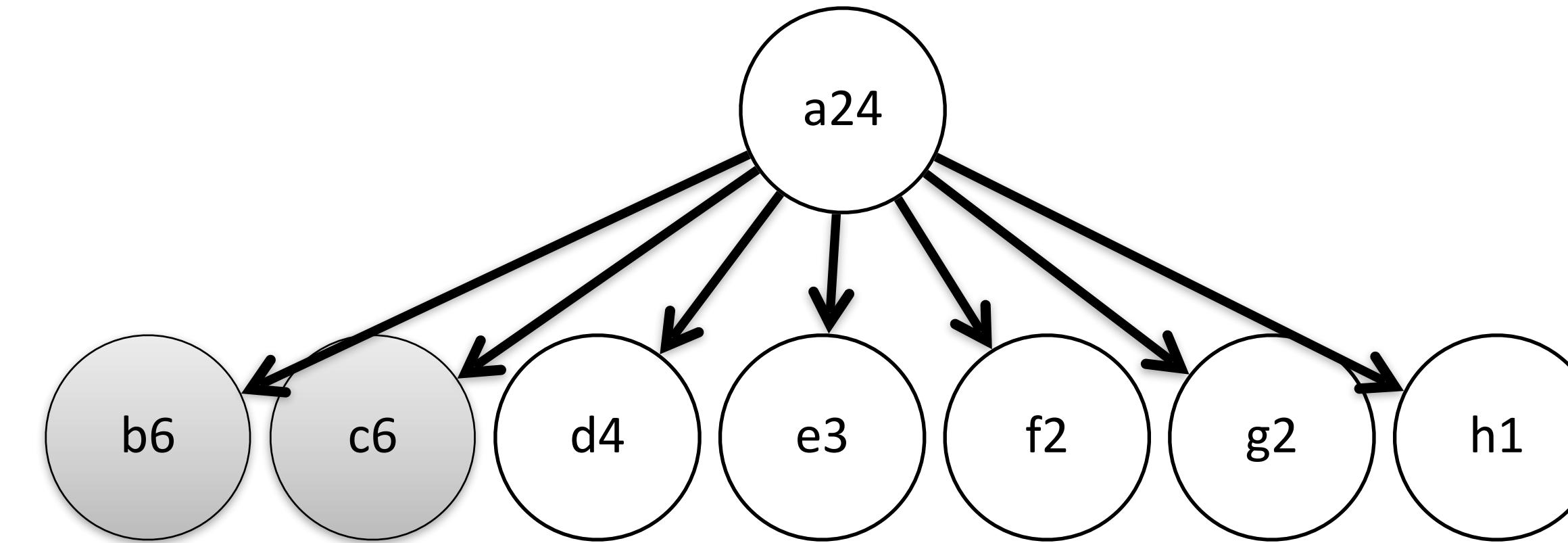
$$\text{Max}(\text{aspect ratio}) = 3/2 = 1,5$$

squarified



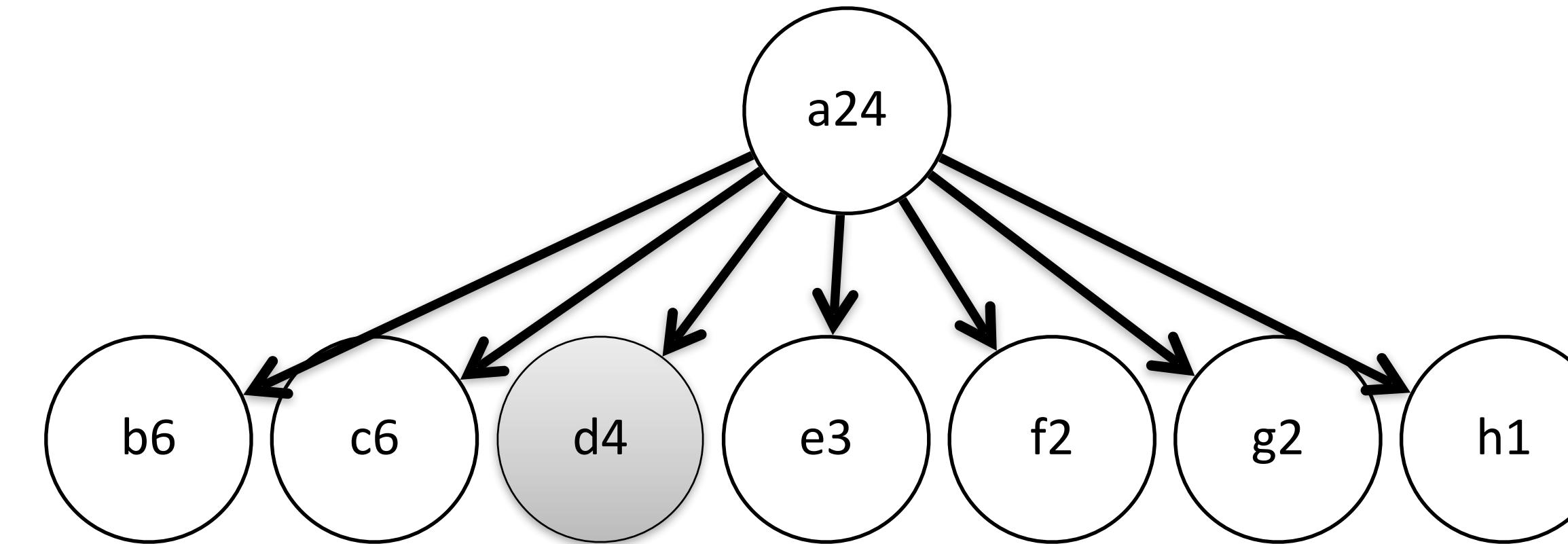
$$\text{Max(aspect ratio)} = 8/2 = 4$$

squarified



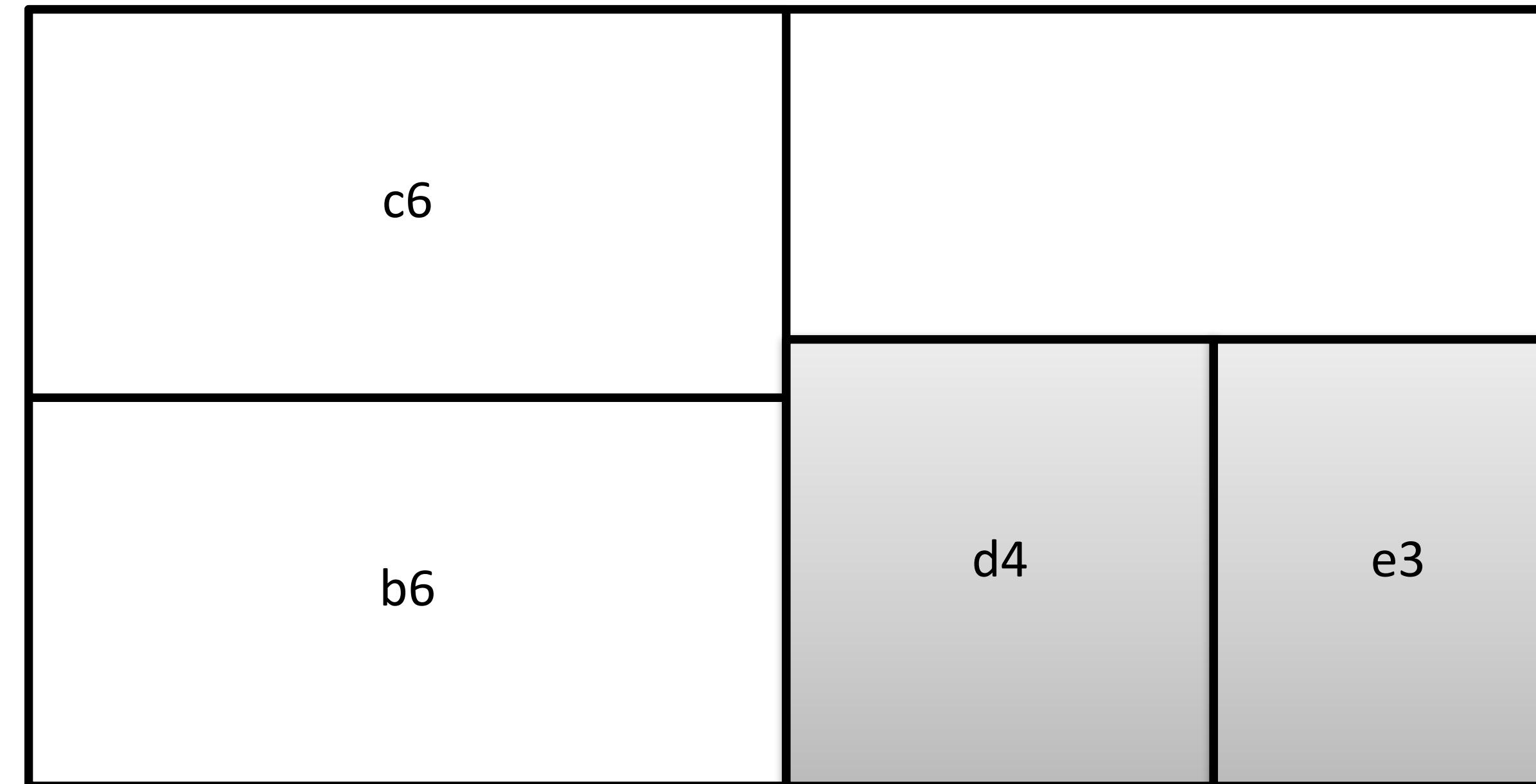
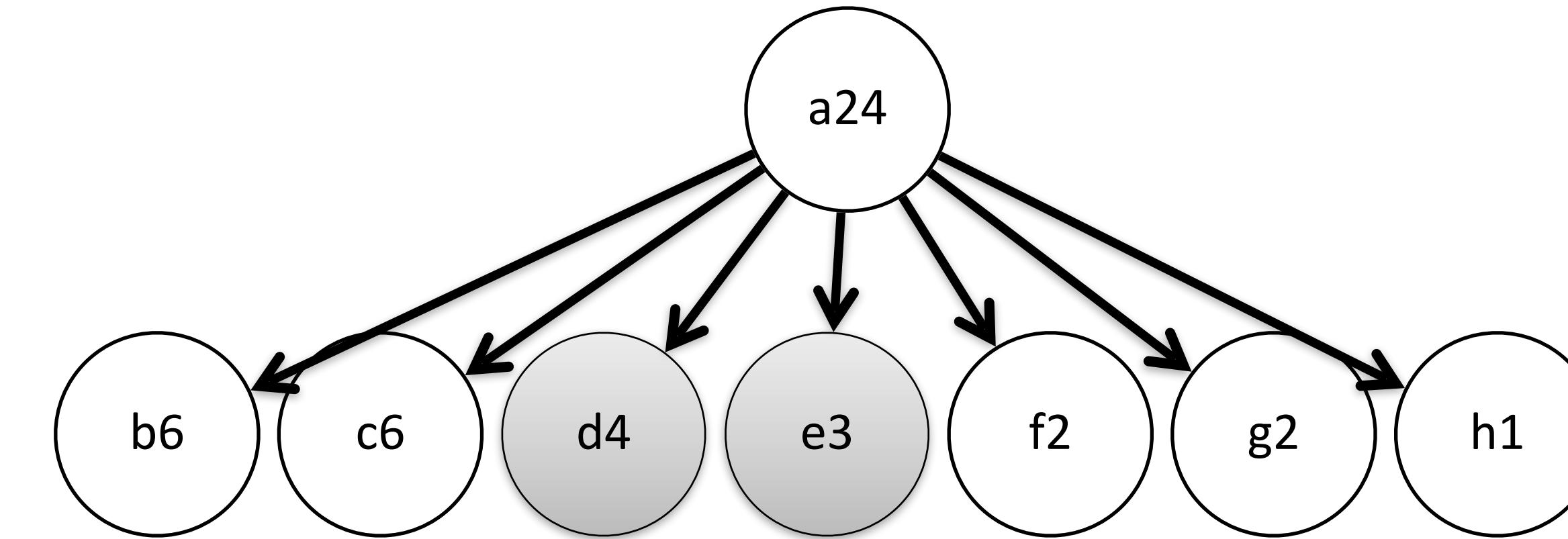
$$\text{Max}(\text{aspect ratio}) = 3/2 = 1,5$$

squarified



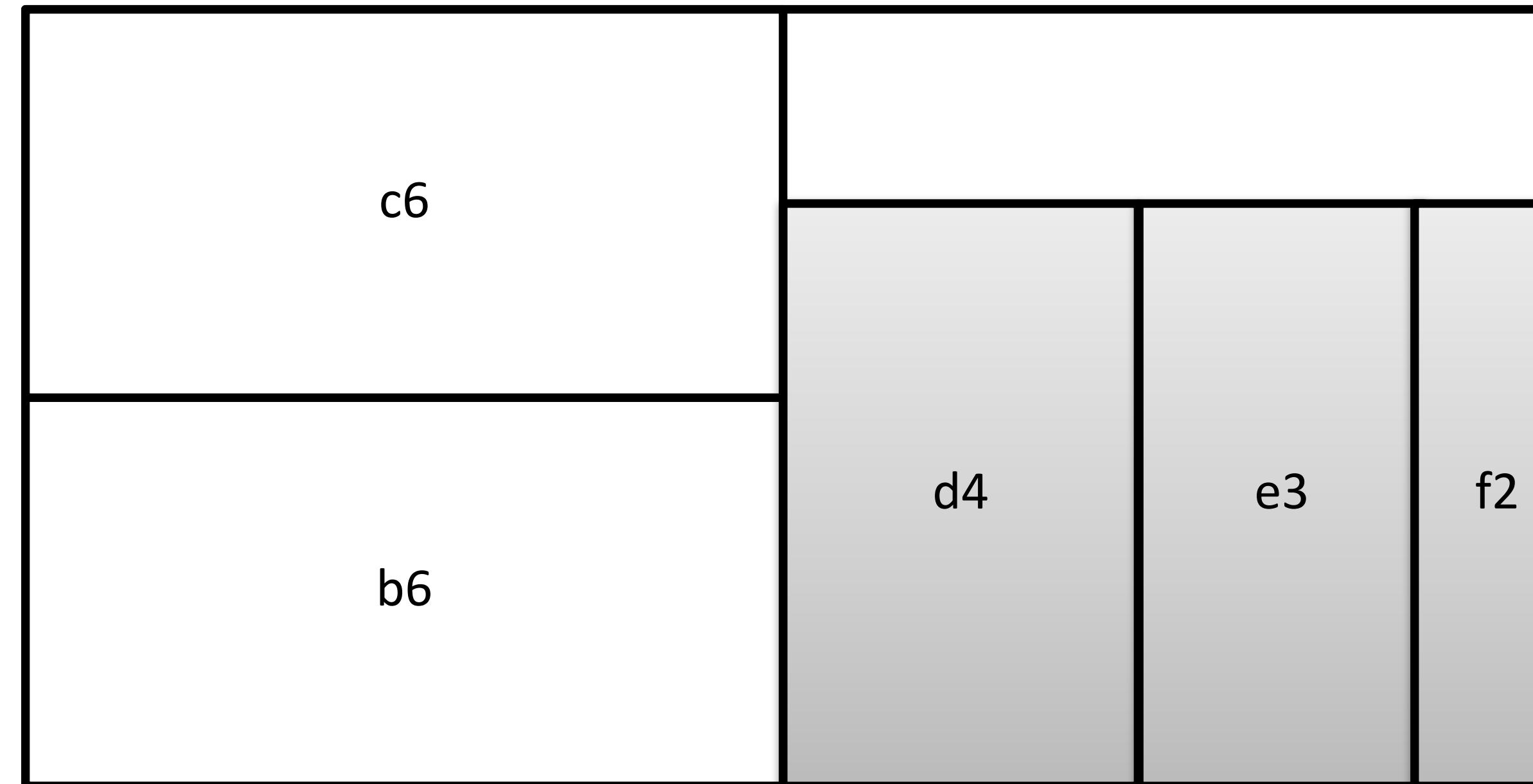
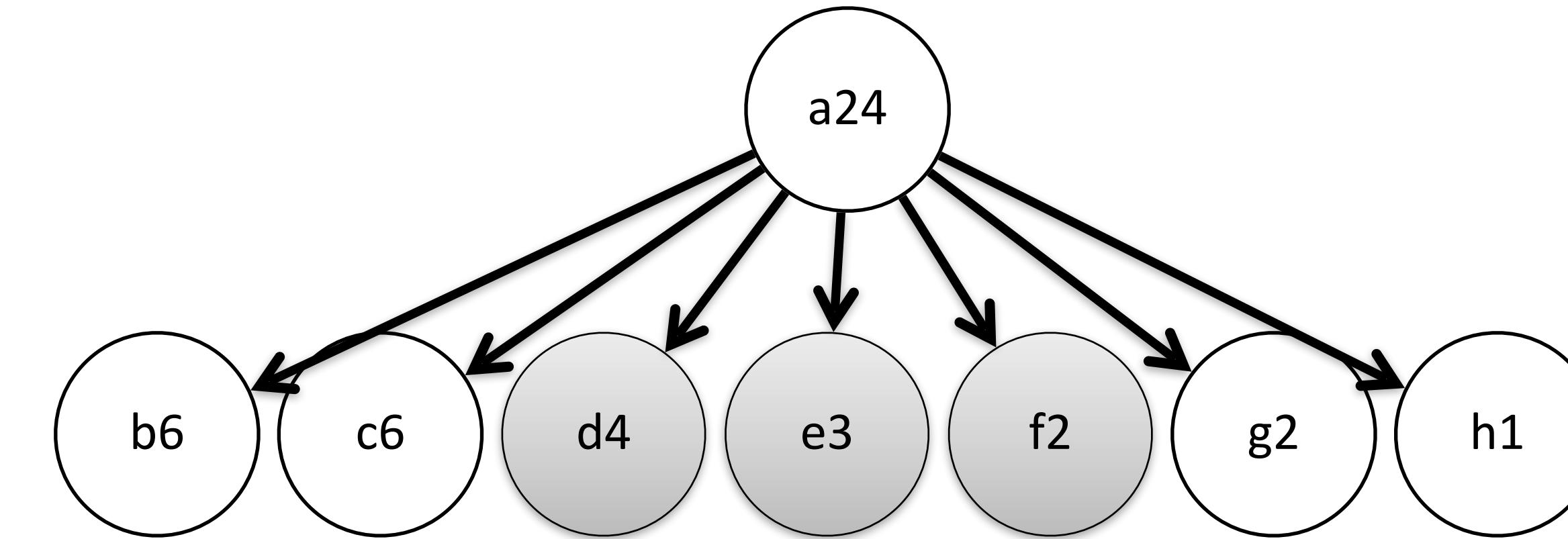
$$\text{Max(aspect ratio)} = 9/4 = 2,25$$

squarified



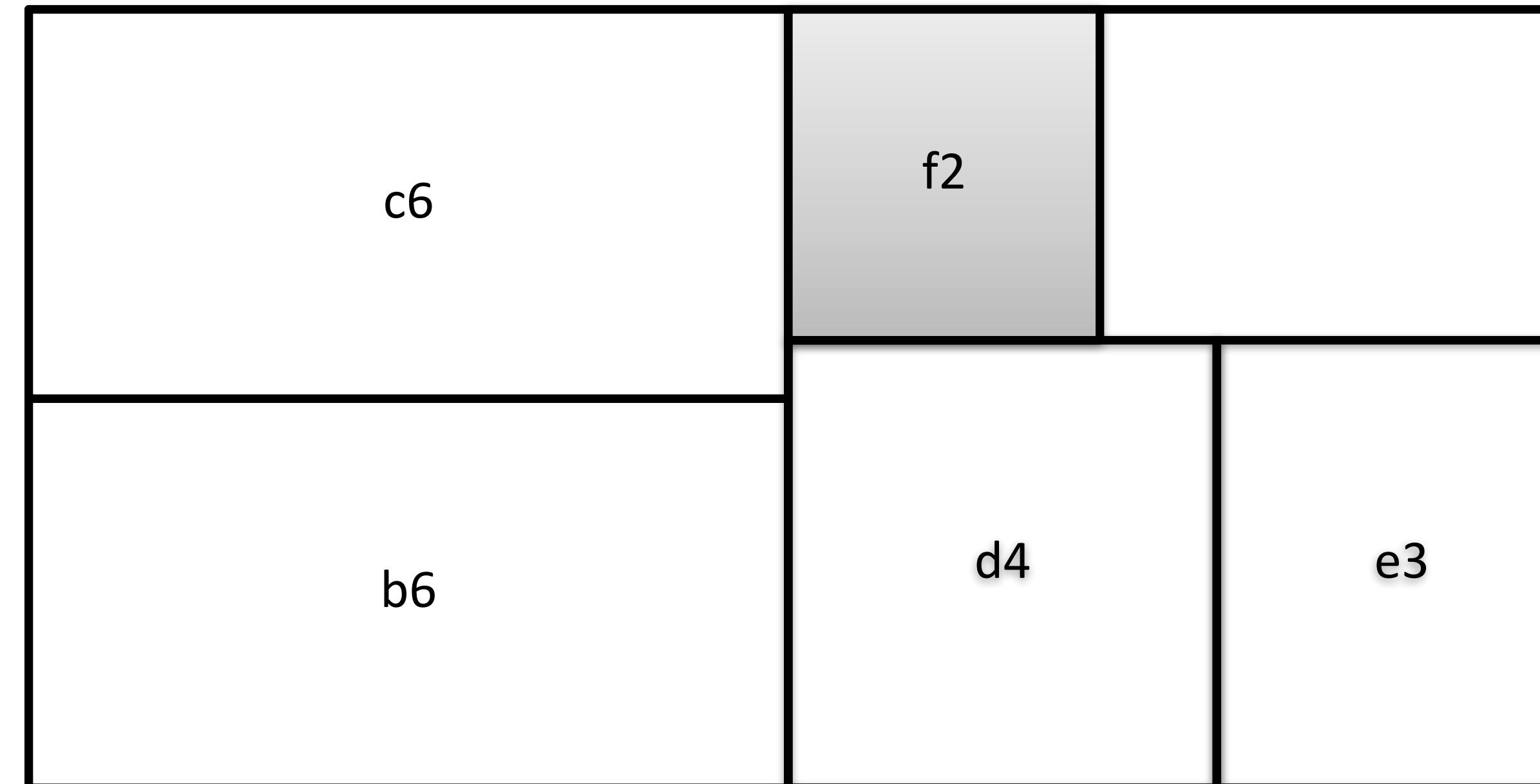
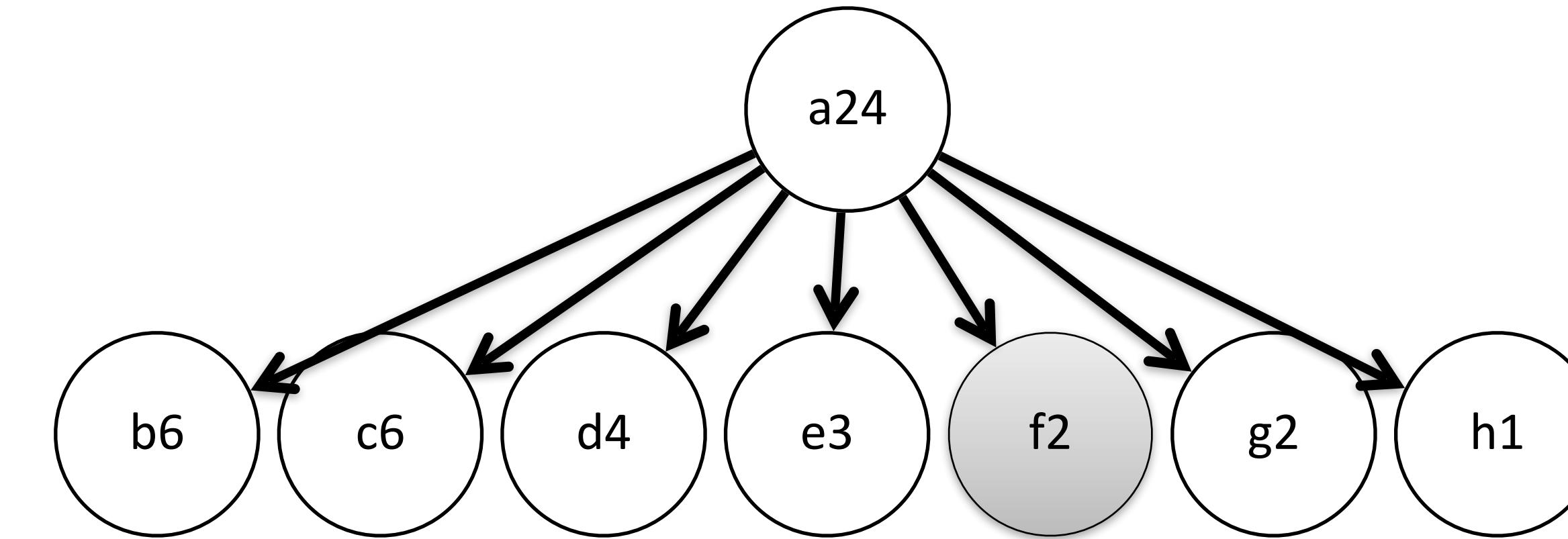
$$\text{Max(aspect ratio)} = 49/27 = 1,81$$

squarified



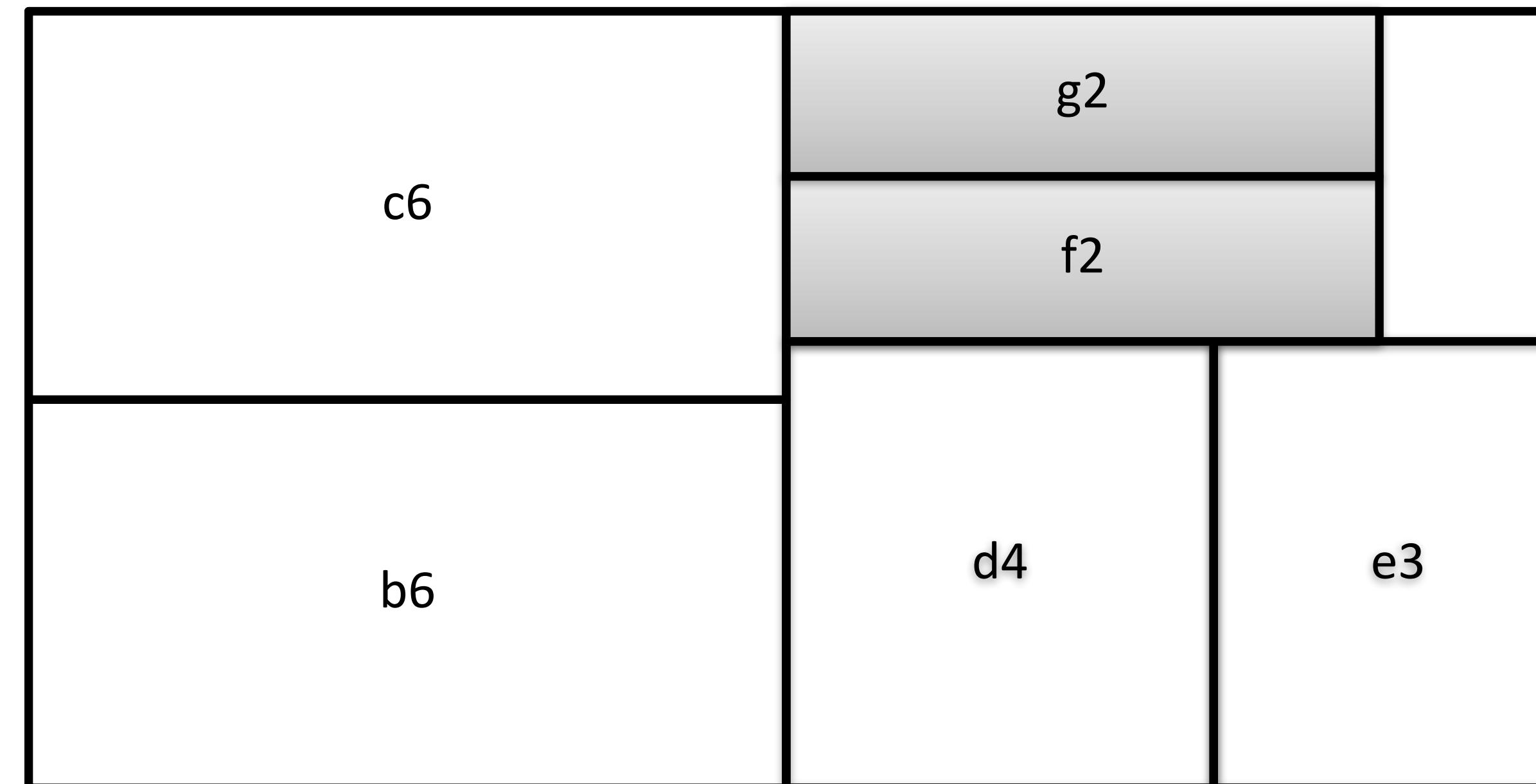
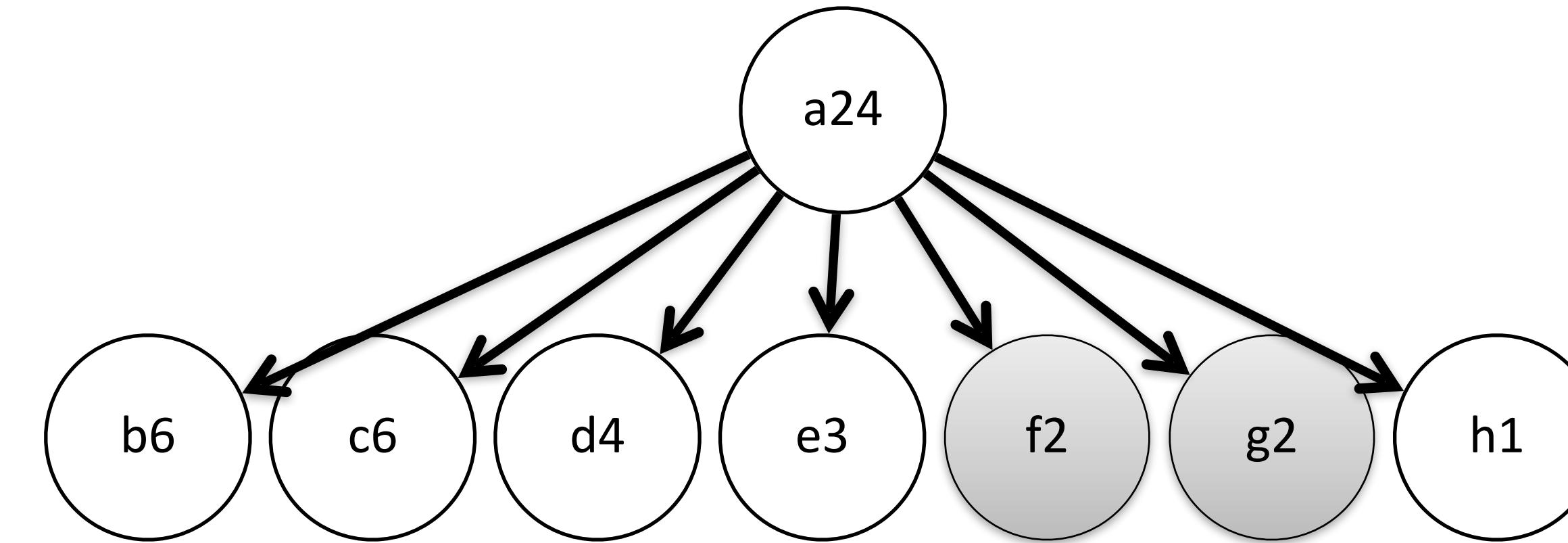
$$\text{Max}(\text{aspect ratio}) = 9/2 = 4,5$$

squarified



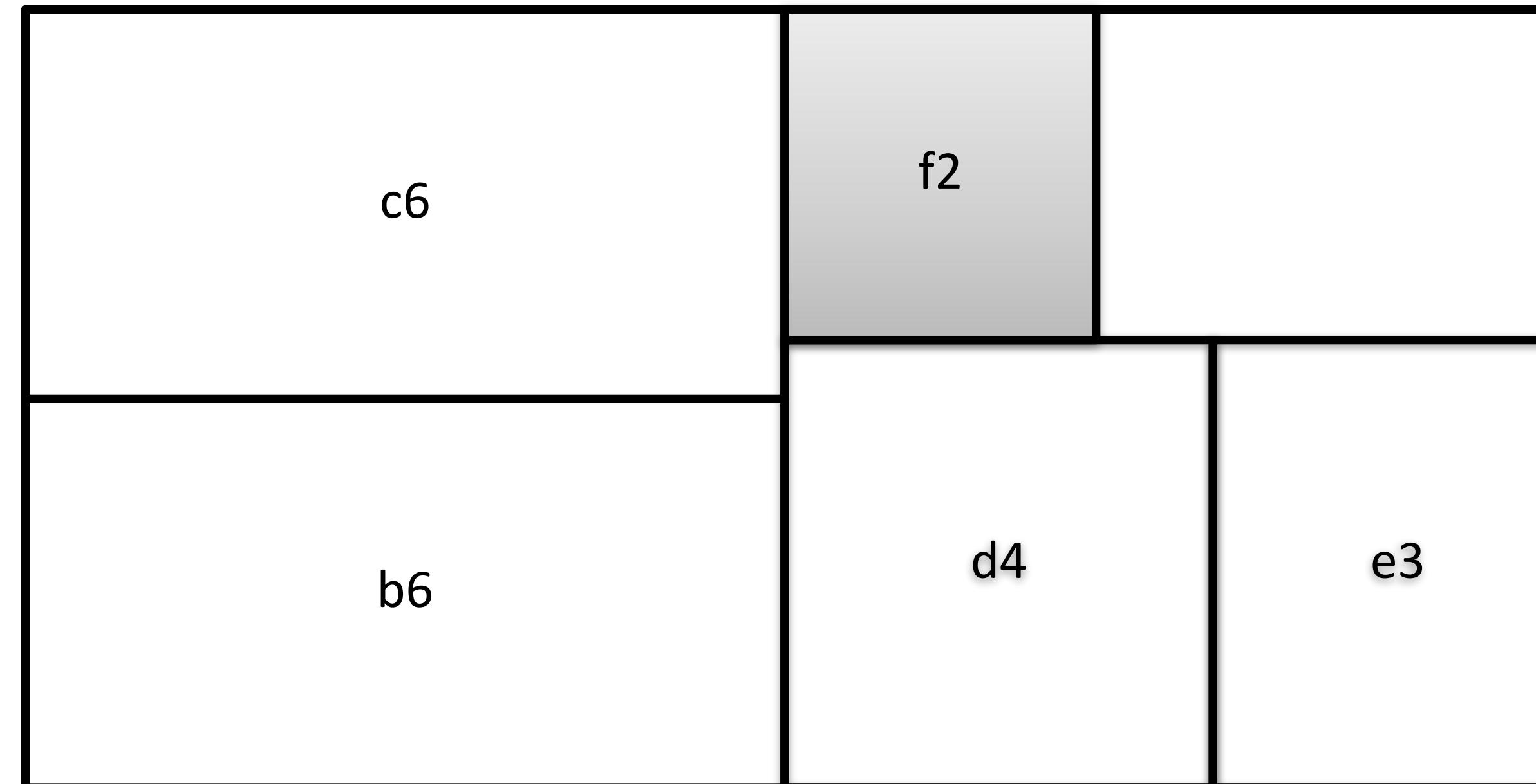
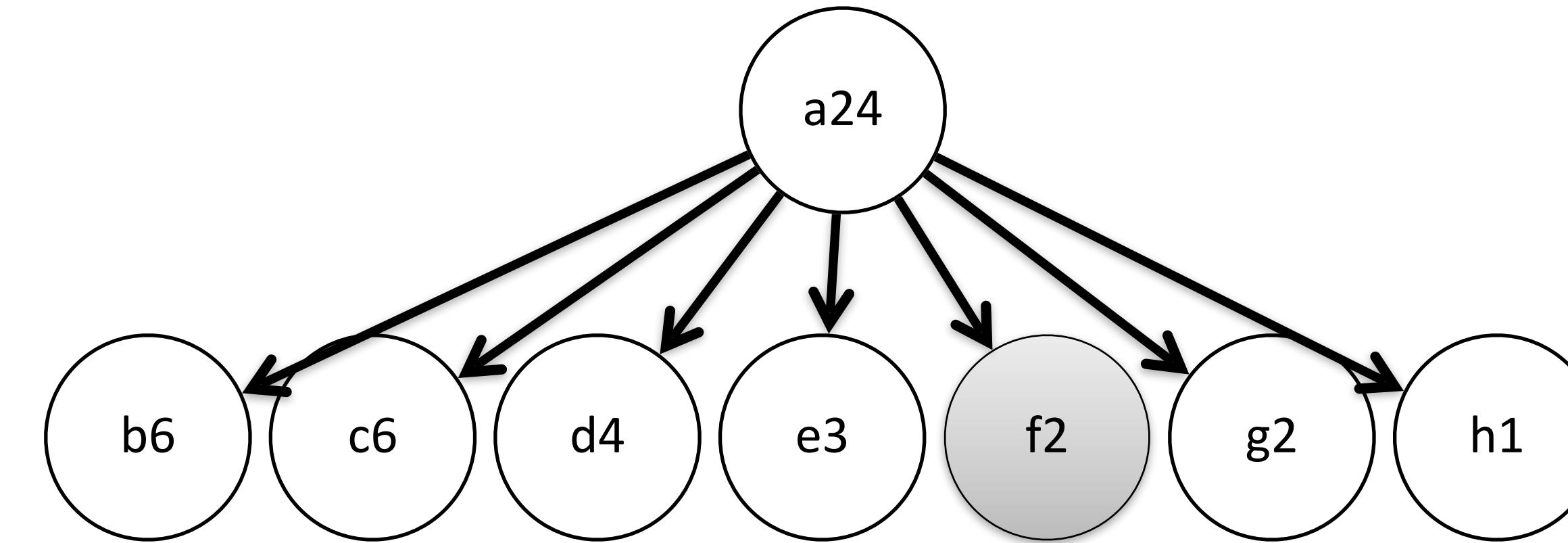
$$\text{Max(aspect ratio)} = 25/18 = 1,39$$

squarified



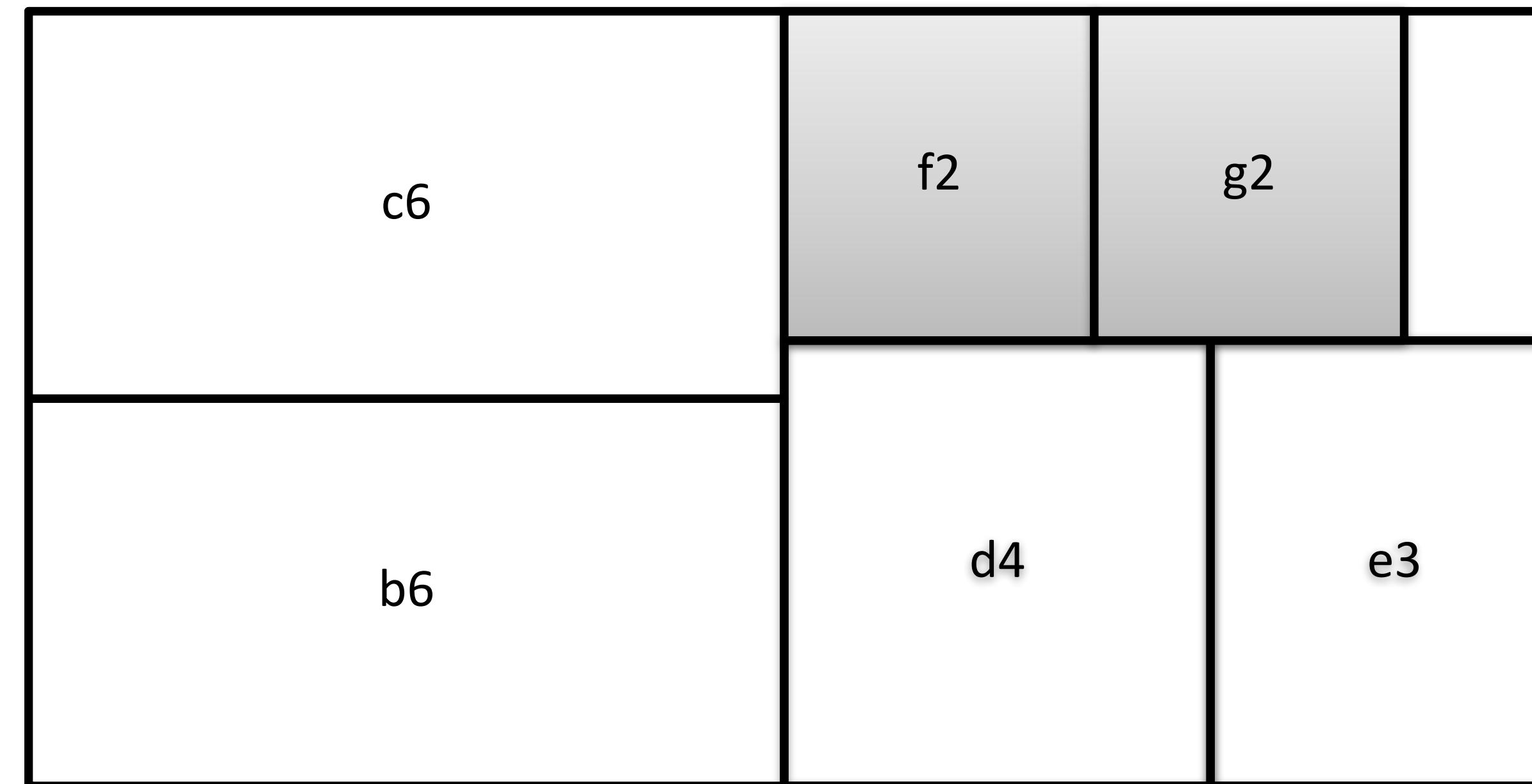
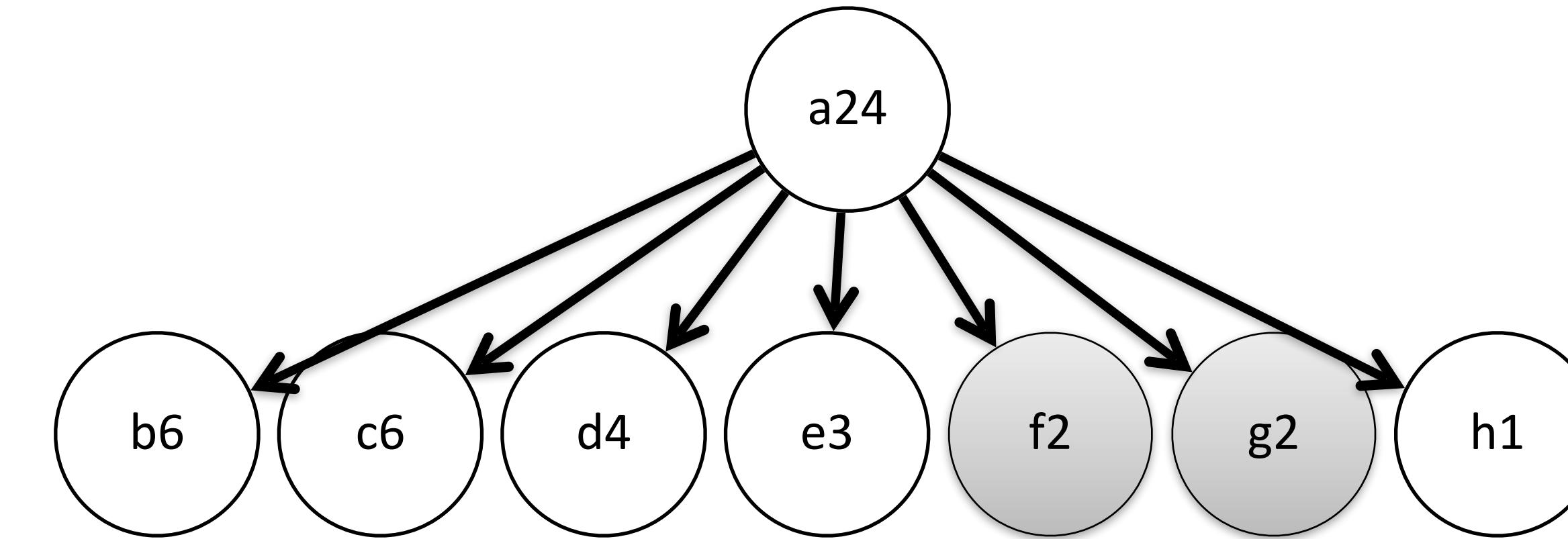
$$\text{Max(aspect ratio)} = 144/50 = 2,88$$

squarified



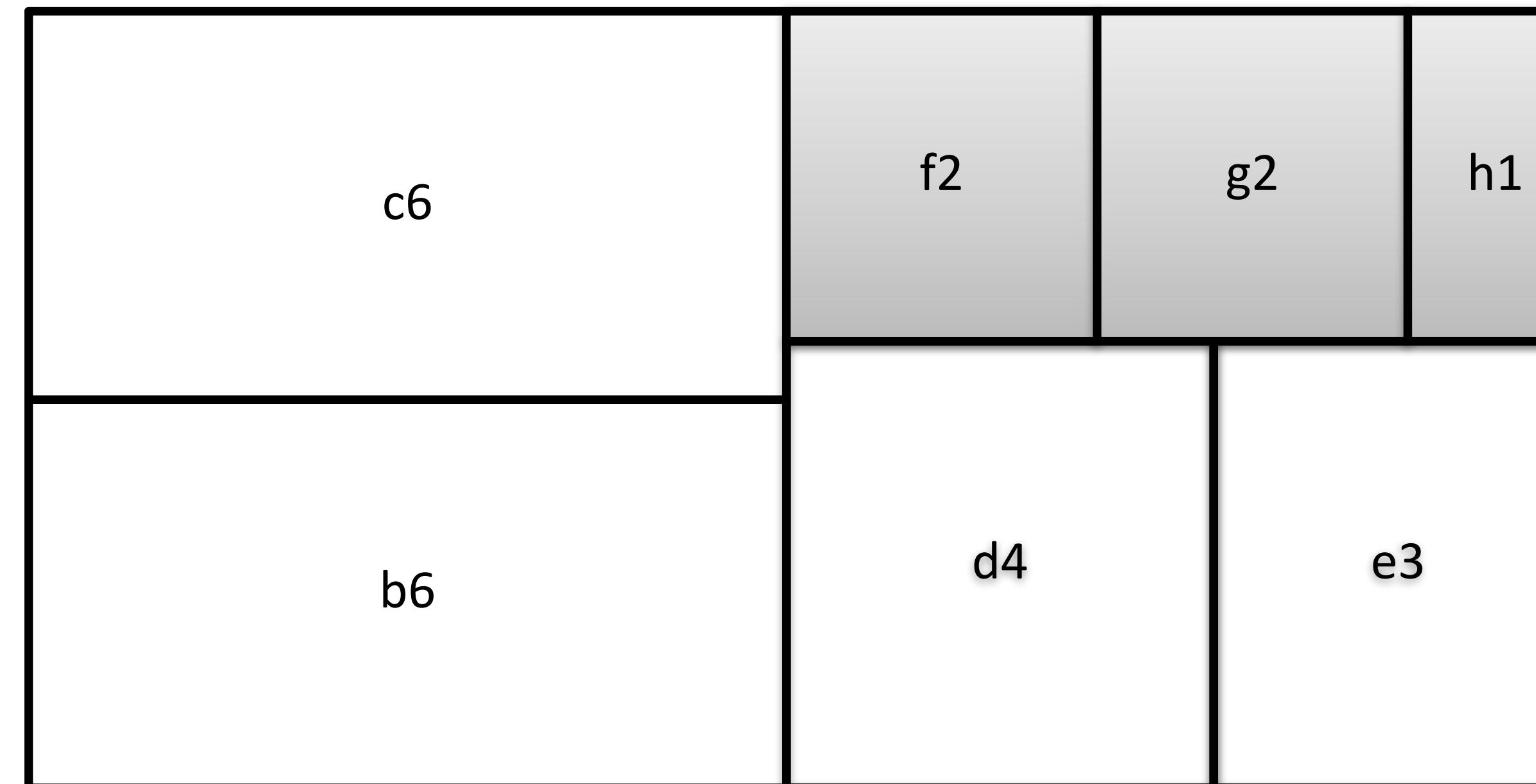
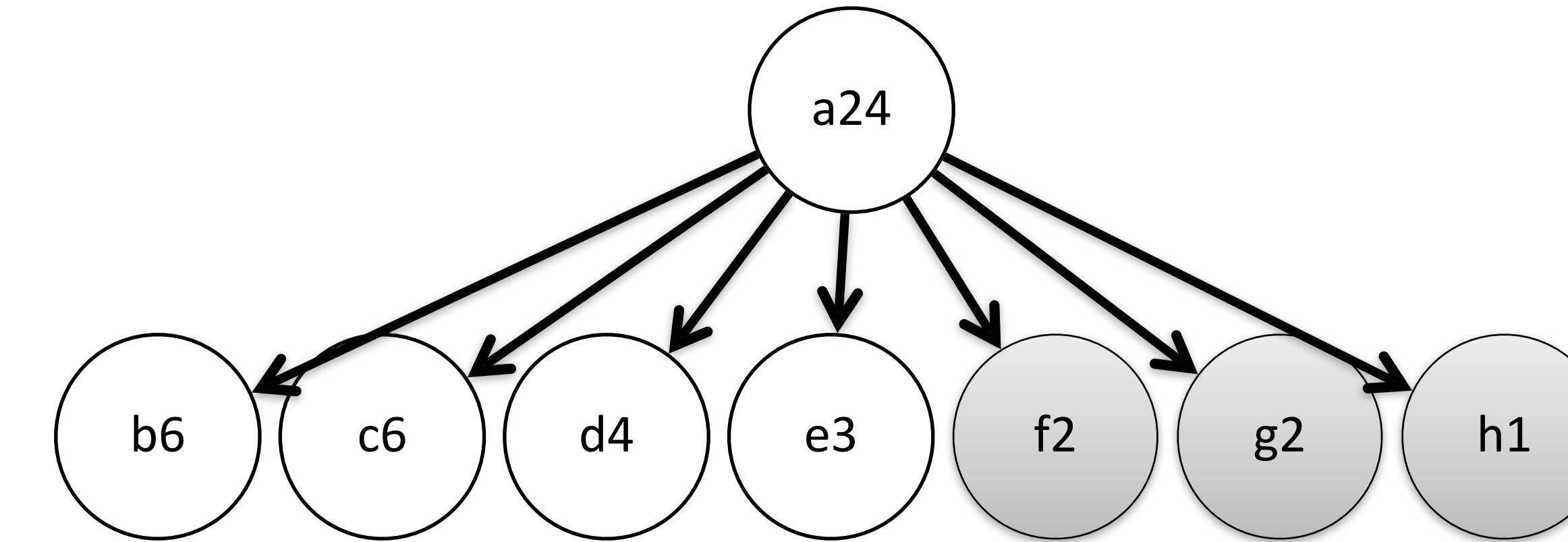
$$\text{Max(aspect ratio)} = 25/18 = 1,39$$

squarified



$$\text{Max(aspect ratio)} = 25/18 = 1,39$$

squarified



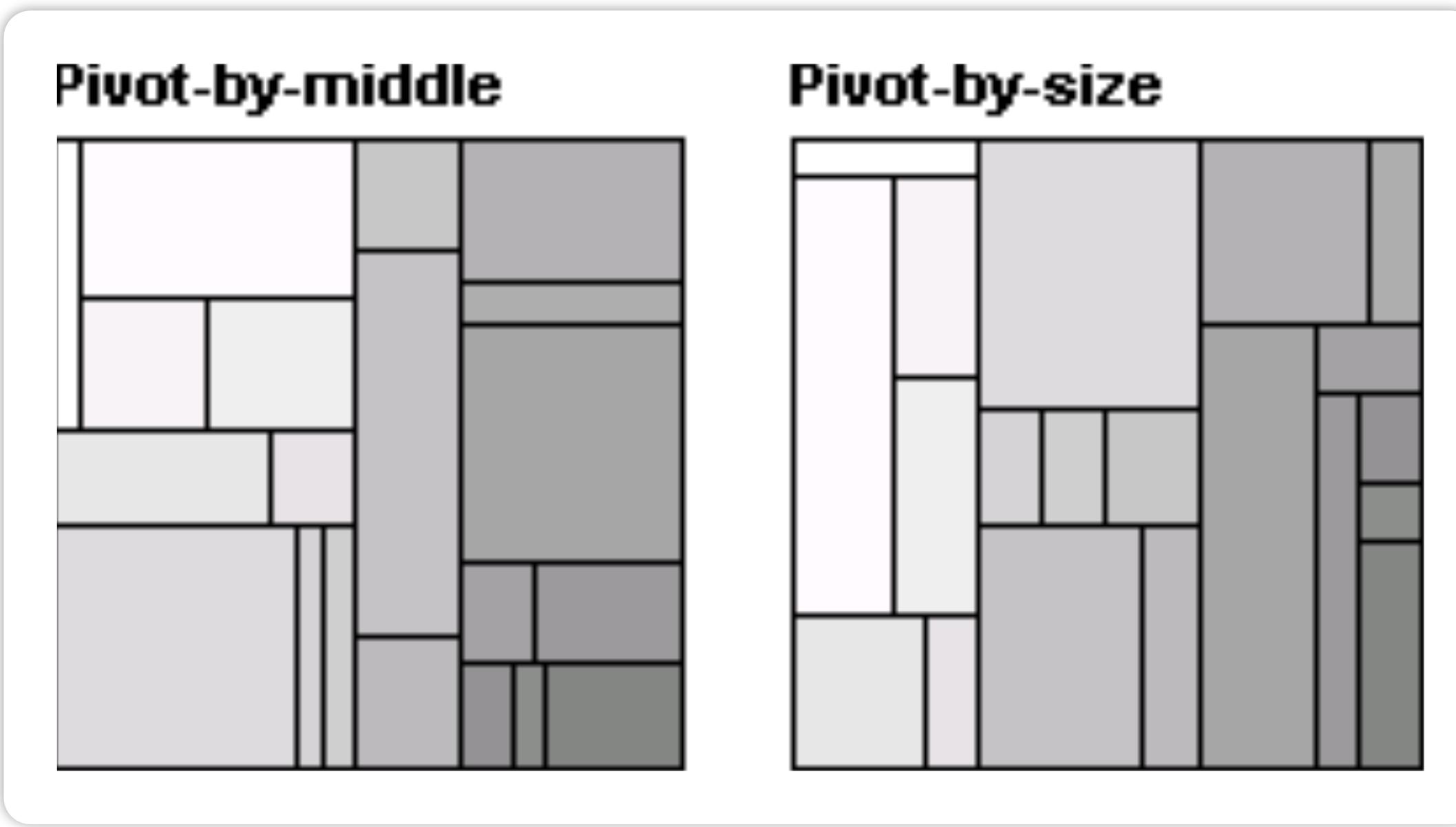
$$\text{Max(aspect ratio)} = 25/9 = 2,78$$

Desvantagens?

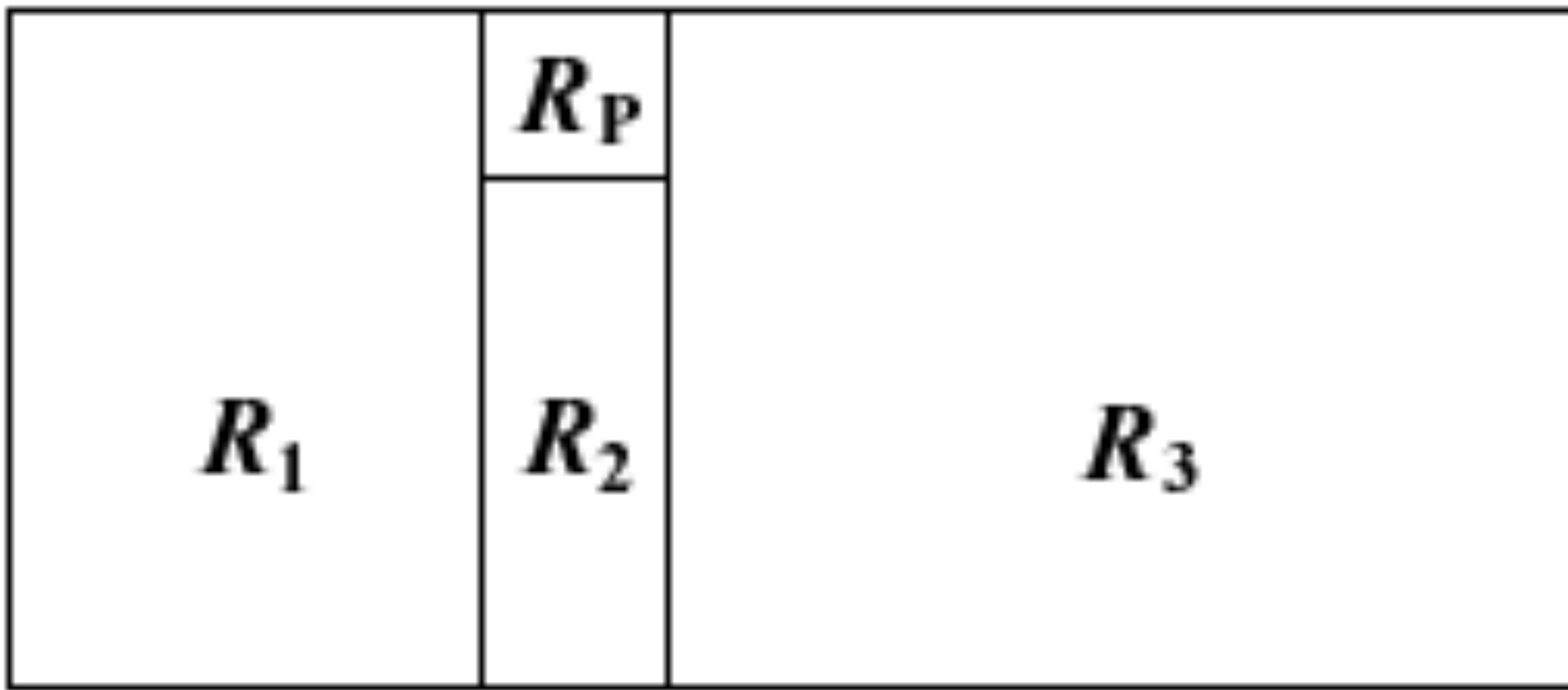
Pequenas mudanças nos dados podem provocar **mudanças radicais no desenho**

Perde a ordem da estrutura da árvore

ORDERED TREEMAP LAYOUTS

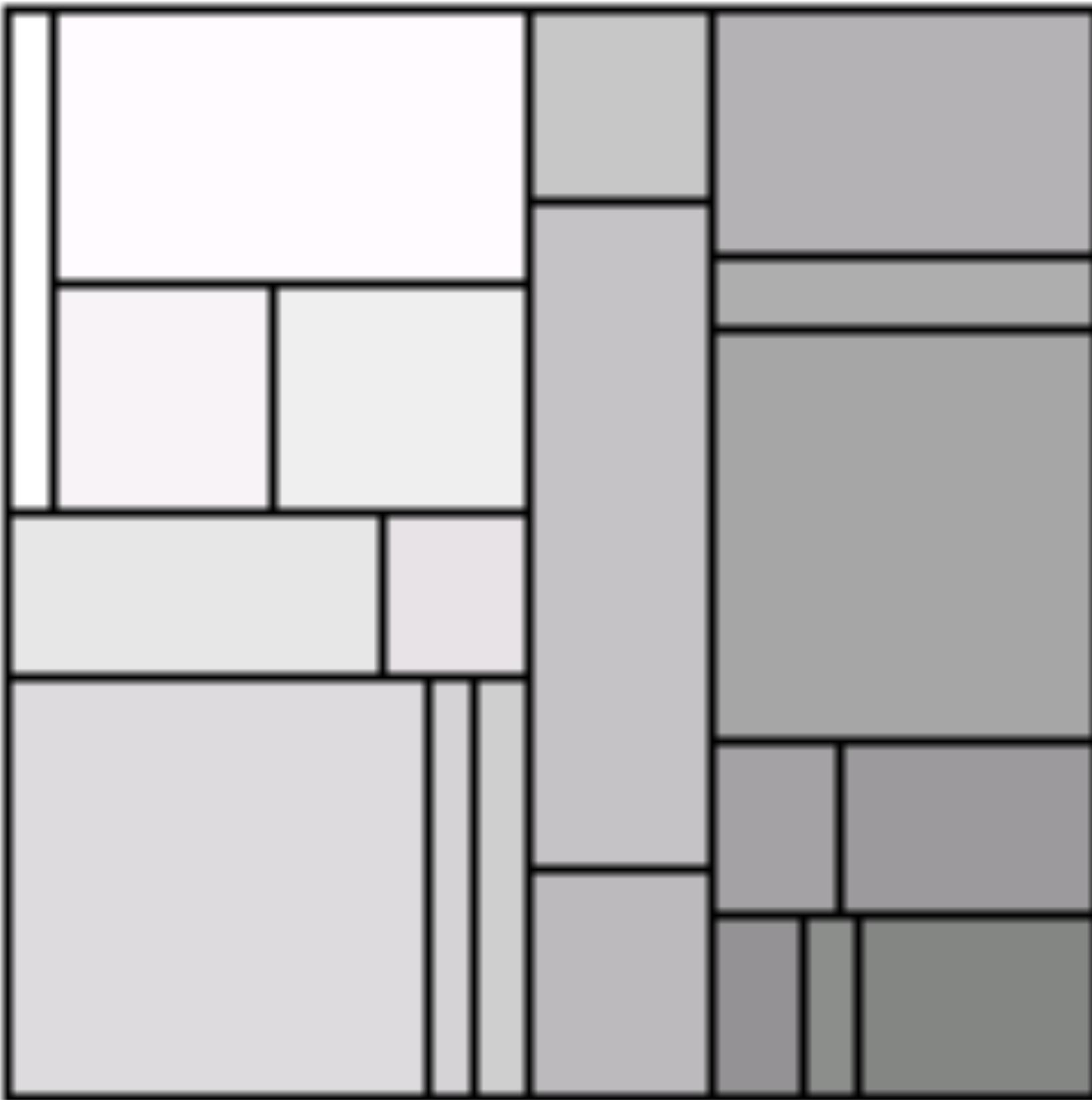


*B. Schneiderman e M. Wattenberg
IEEE Information Visualization
2001*

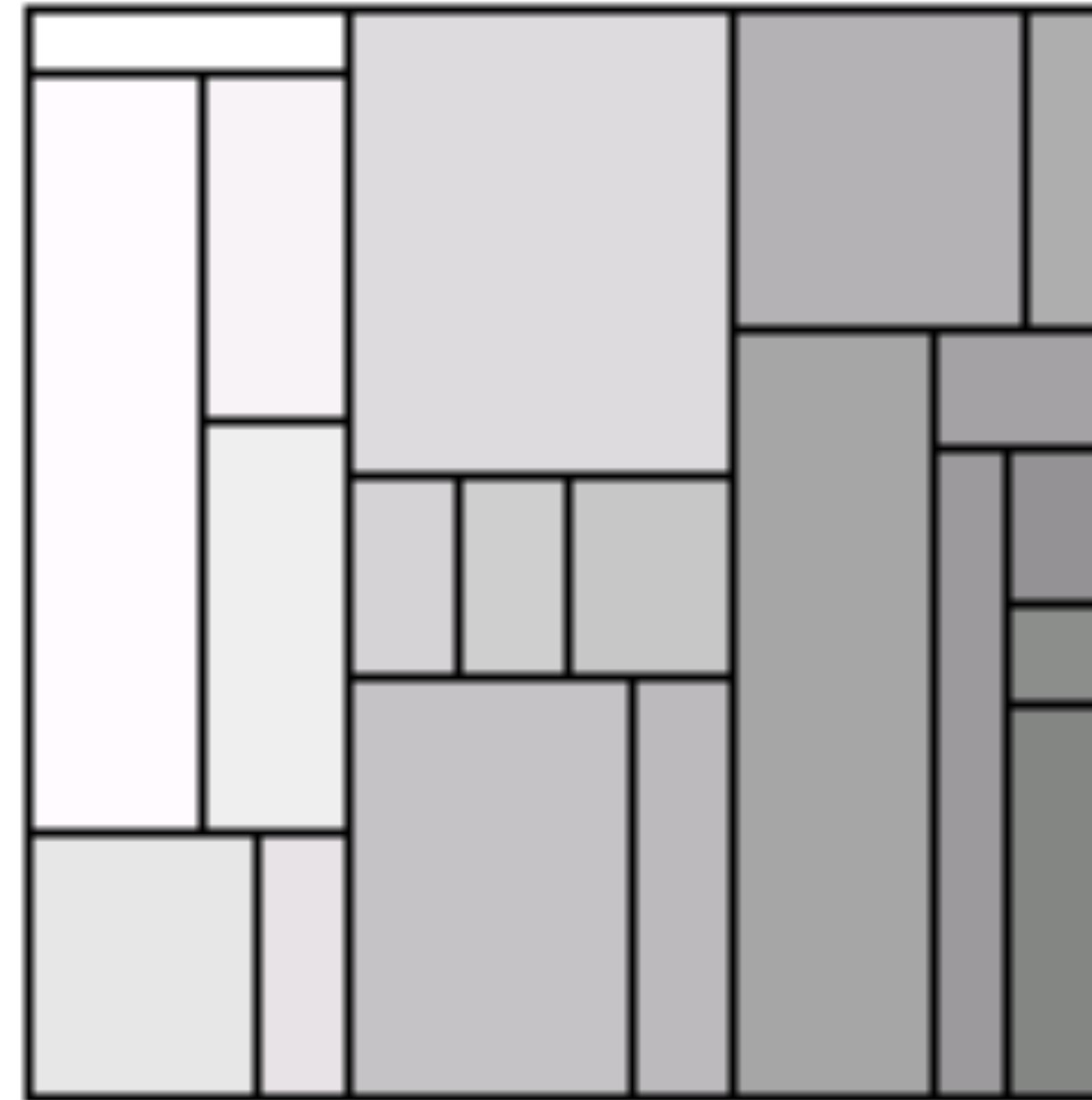


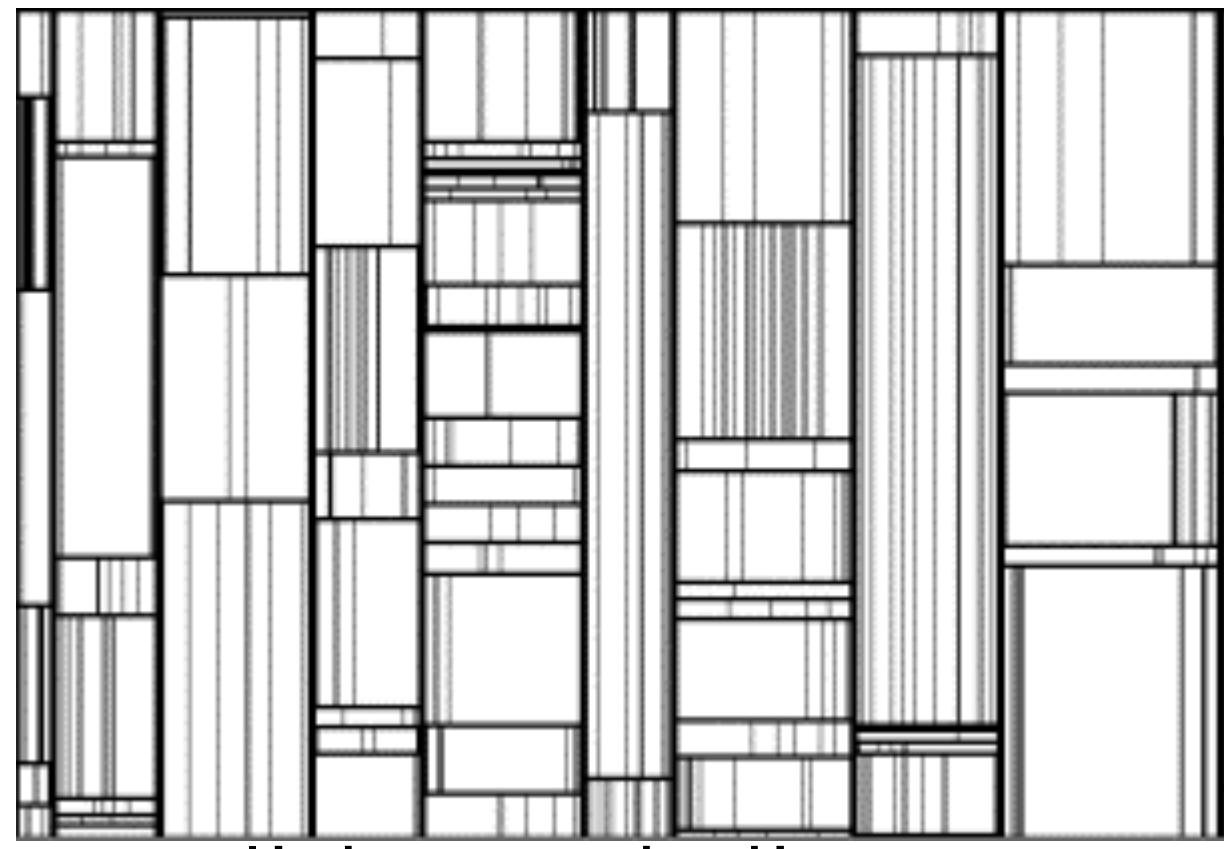
1. Escolha um pivot (maior área, valor intermediário, etc)
2. Posicione o pivot na maior base, criando os retângulos R_1 , R_2 e R_3
3. Divida os itens em listas L_1 , L_2 e L_3 a serem colocadas em R_1 , R_2 e R_3 . L_1 contém os itens anteriores ao pivot e L_2 e L_3 , os posteriores. L_1 e L_3 podem ser vazias.
4. Posicione L_1 , L_2 e L_3 recursivamente com o mesmo algoritmo

Pivot-by-middle

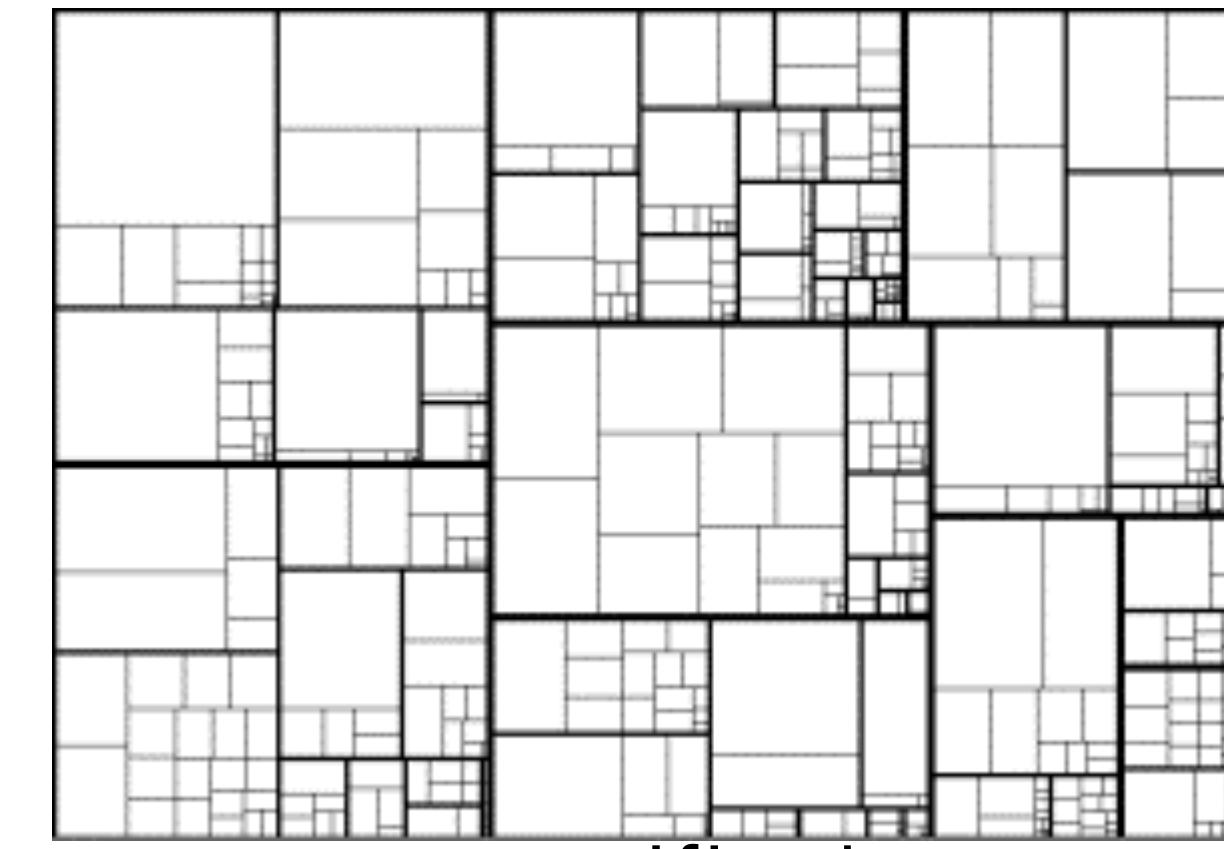


Pivot-by-size





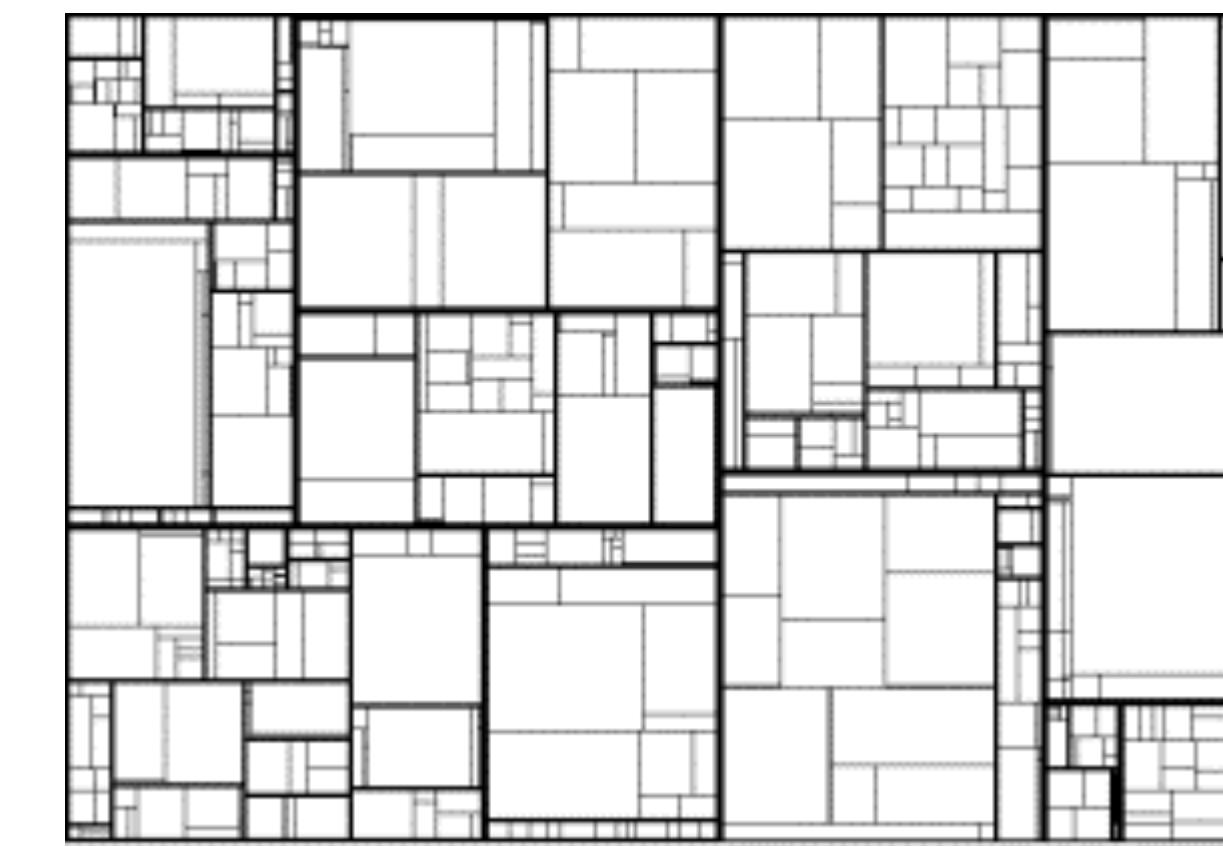
slide-and-dice



squarified



pivot by middle



pivot by size

Dynamic treemap layout comparison

- Martin Wattenberg, w@bewitched.com
- Ben Bederson, (University of Maryland, [Human-Computer Interaction Lab](#))

[Continue](#)

[Start over](#)

Type of change:

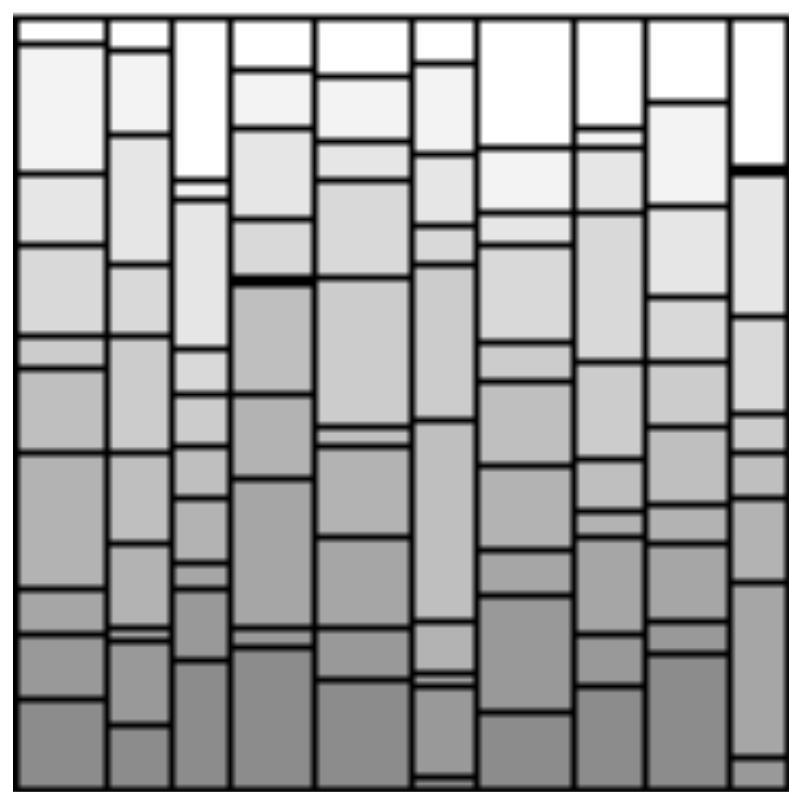
Color by order

10 x 10 ▾

Random walk

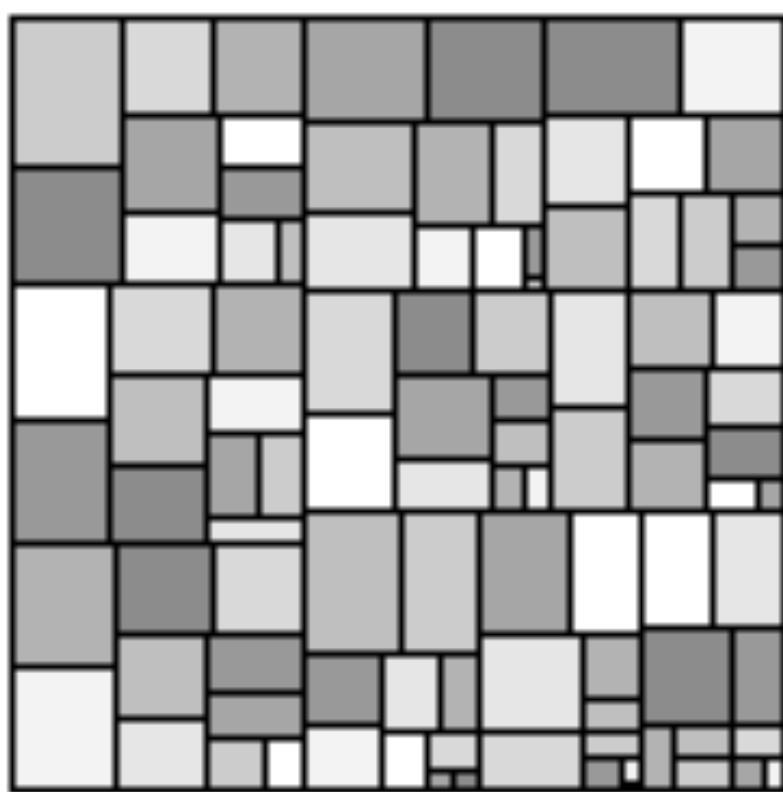
Sine waves

Slice-and-dice



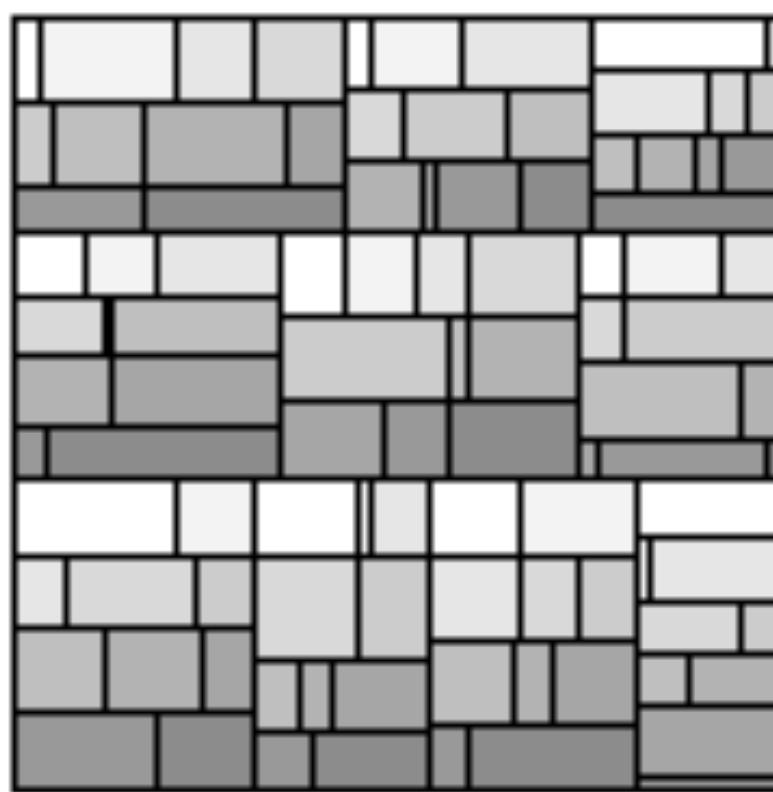
Avg. Aspect = 2.39

Squarified



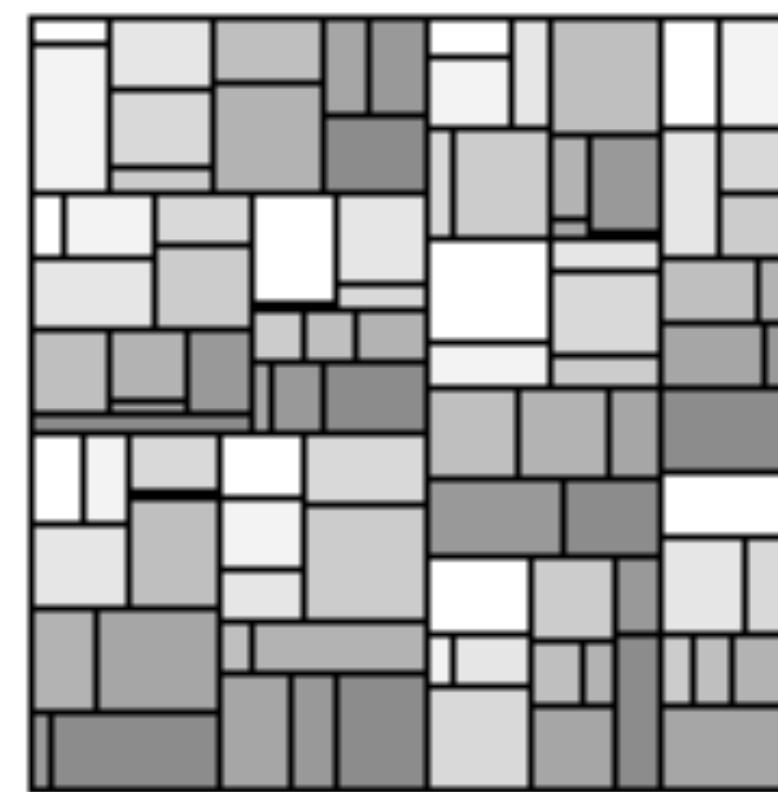
Avg. Aspect = 1.57

StripTreemap



Avg. Aspect = 2.06

Pivot by Split Size

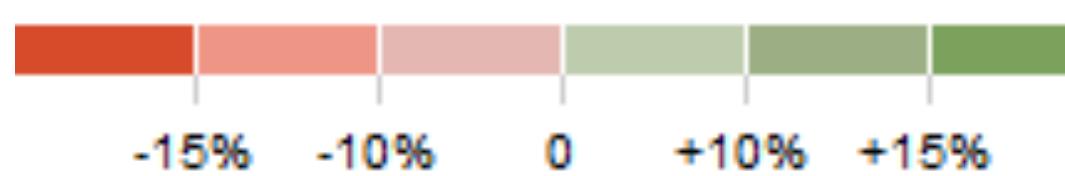


Avg. Aspect = 2.3

Obama's 2012 Budget Proposal: How \$3.7 Trillion is Spent

Explore every nook and cranny of President Obama's budget proposal.

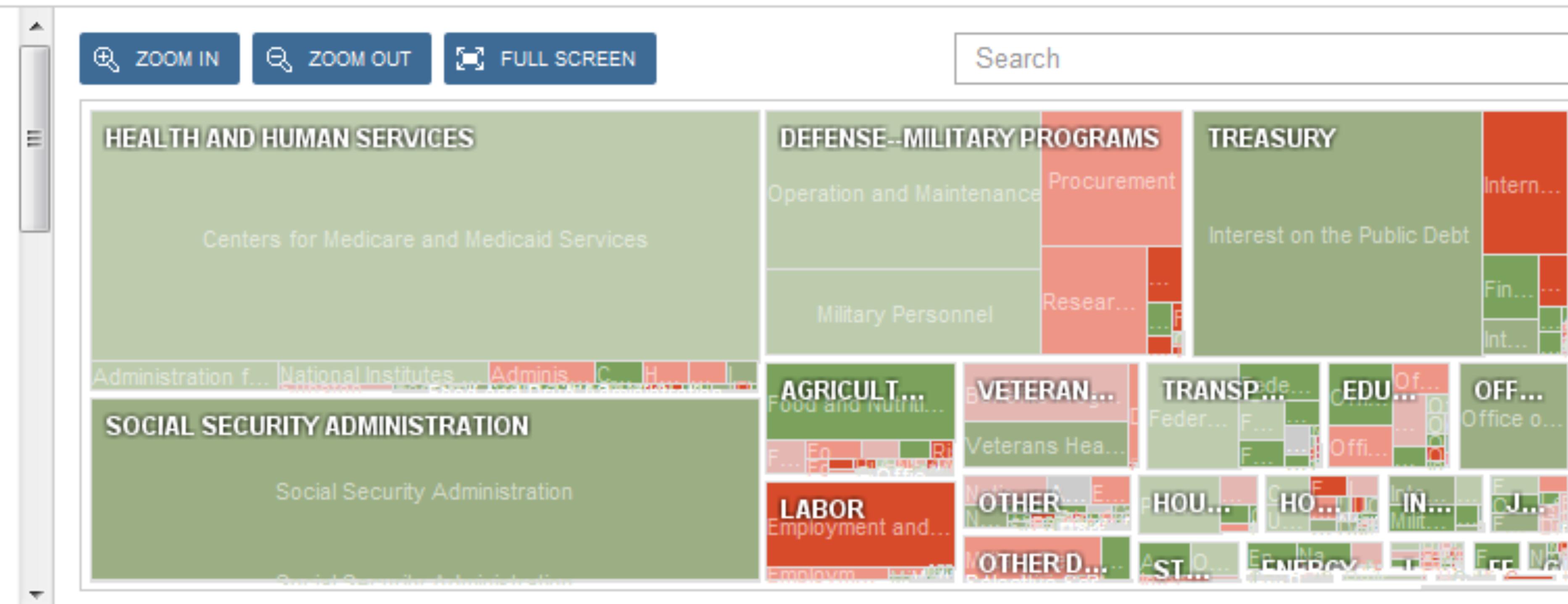
Rectangles are sized according to the proposed spending. Color shows severity of cut or increase from 2010.



The president has proposed a five-year freeze of discretionary spending, excluding national security spending. This type of spending accounts for about one fifth of all government spending.

Isolate discretionary spending

Mandatory spending, which includes entitlement programs like Medicare and Social Security, is expected to continue to



Published: February 14, 2011 | By SHAN CARTER and AMANDA COX | Source: Office of Management and Budget



TWITTER



 FACEBOOK

Cuidado com variações perigosas

Qual a vantagem de se usar células de Voronoi?

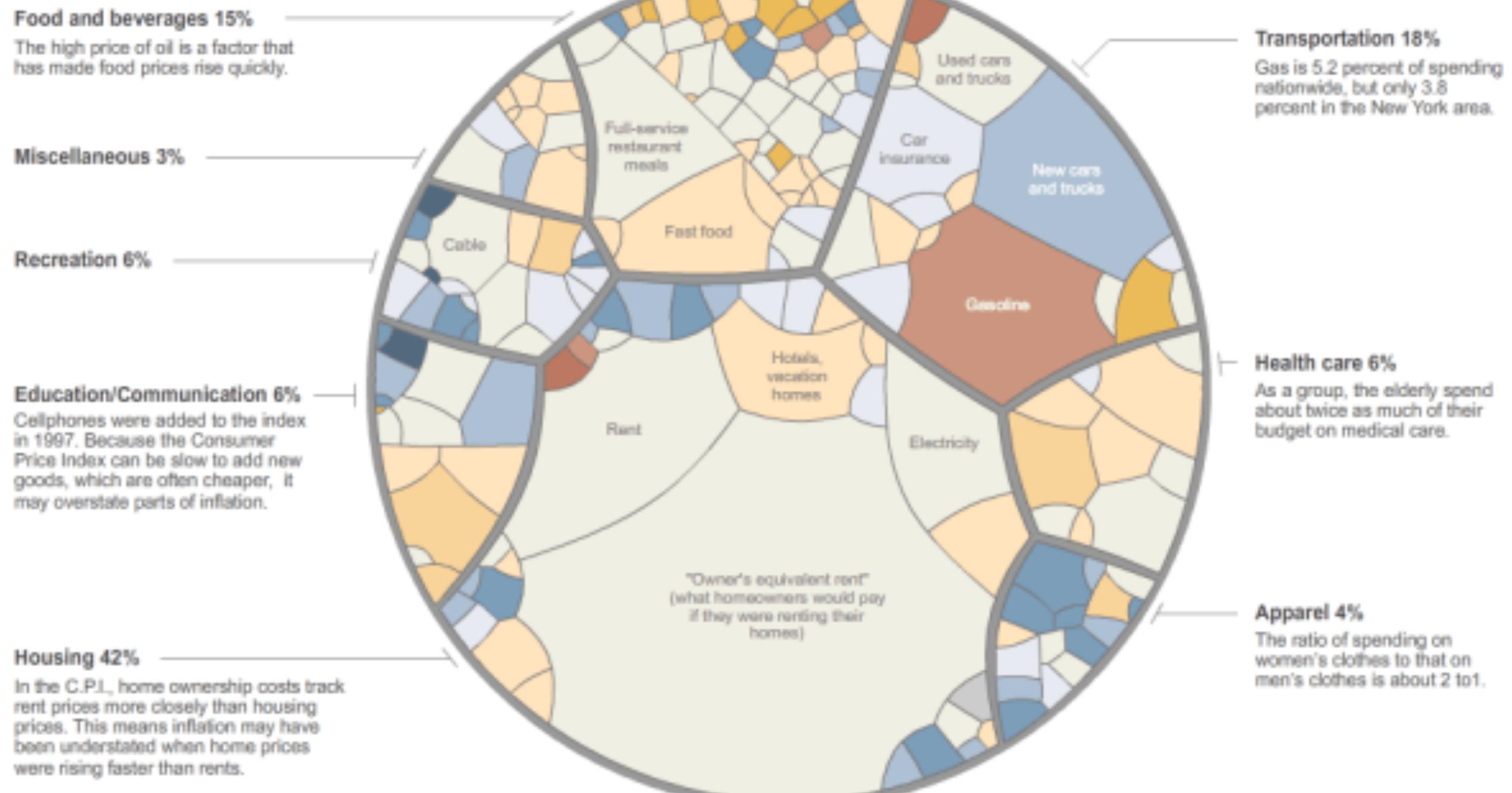
An Average Consumer's Spending

Each shape below represents how much the average American spends in different categories.
Larger shapes make up a larger part of spending.

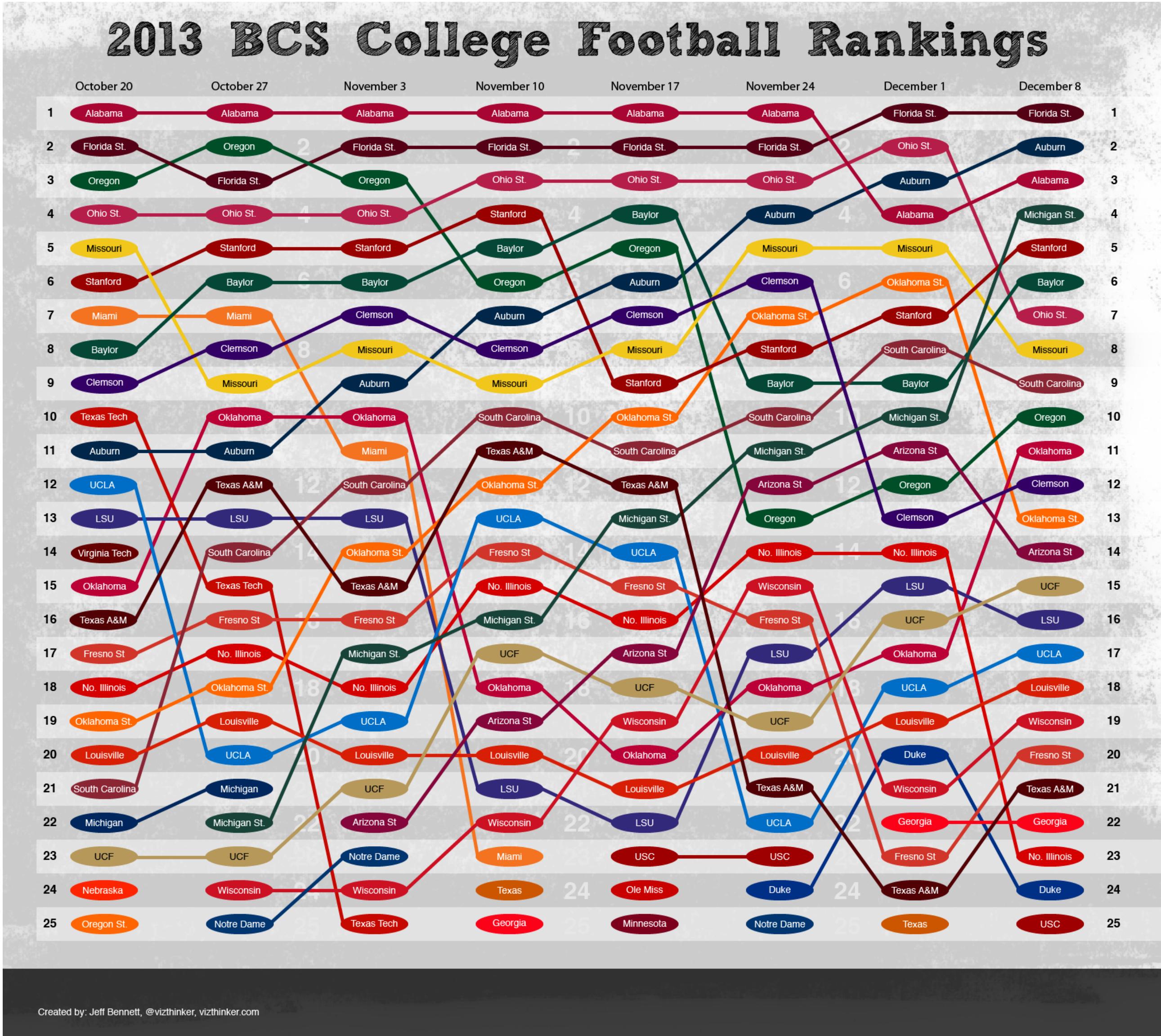
Color shows change in prices from March 2007 to March 2008



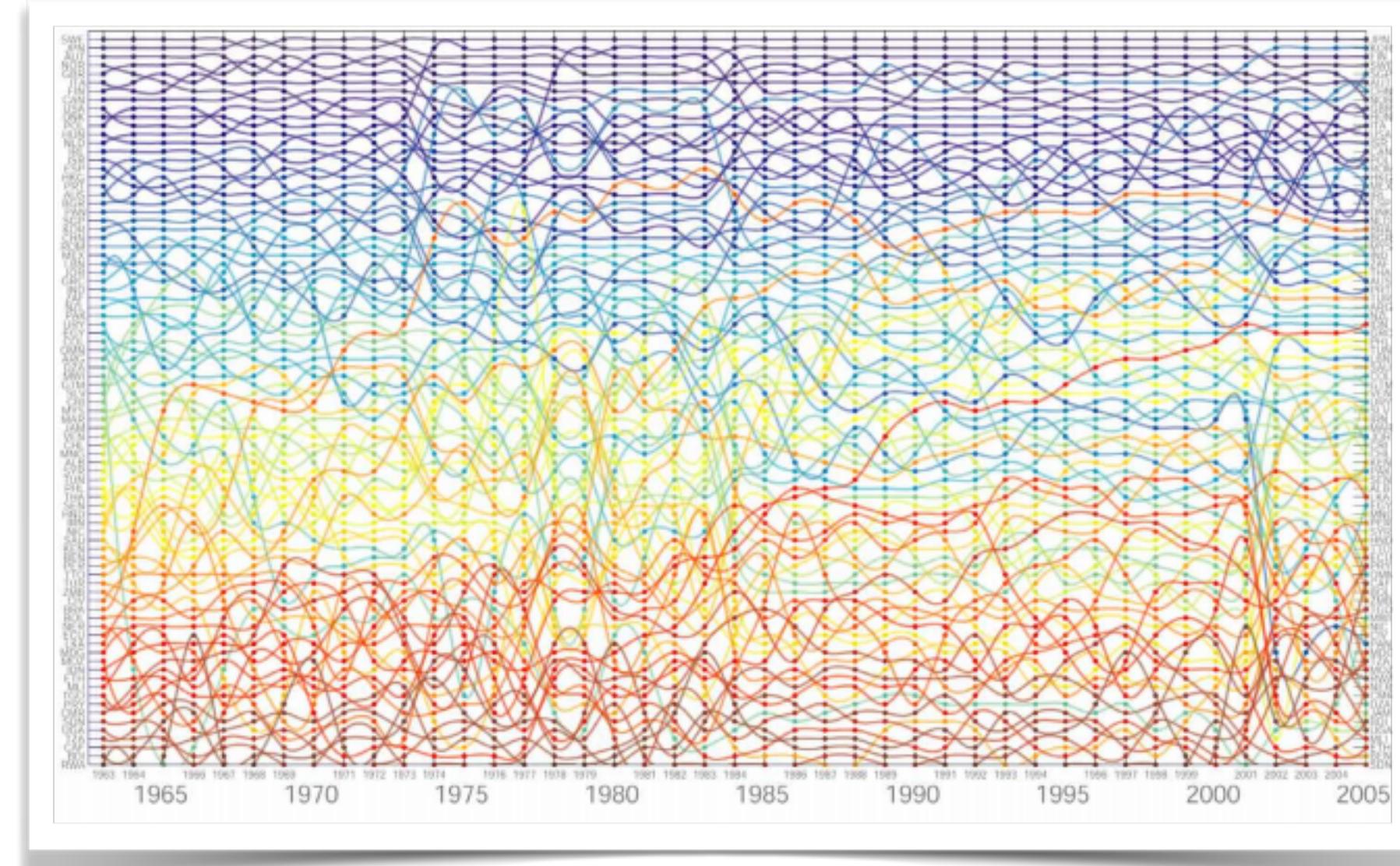
ZOOM IN ZOOM OUT



BUMPCHART



THE BUILDING BLOCKS OF ECONOMY COMPLEXITY



C.A. Hidalgo e R. Hausman
*Proceedings of the National Academy of
Sciences*
2009

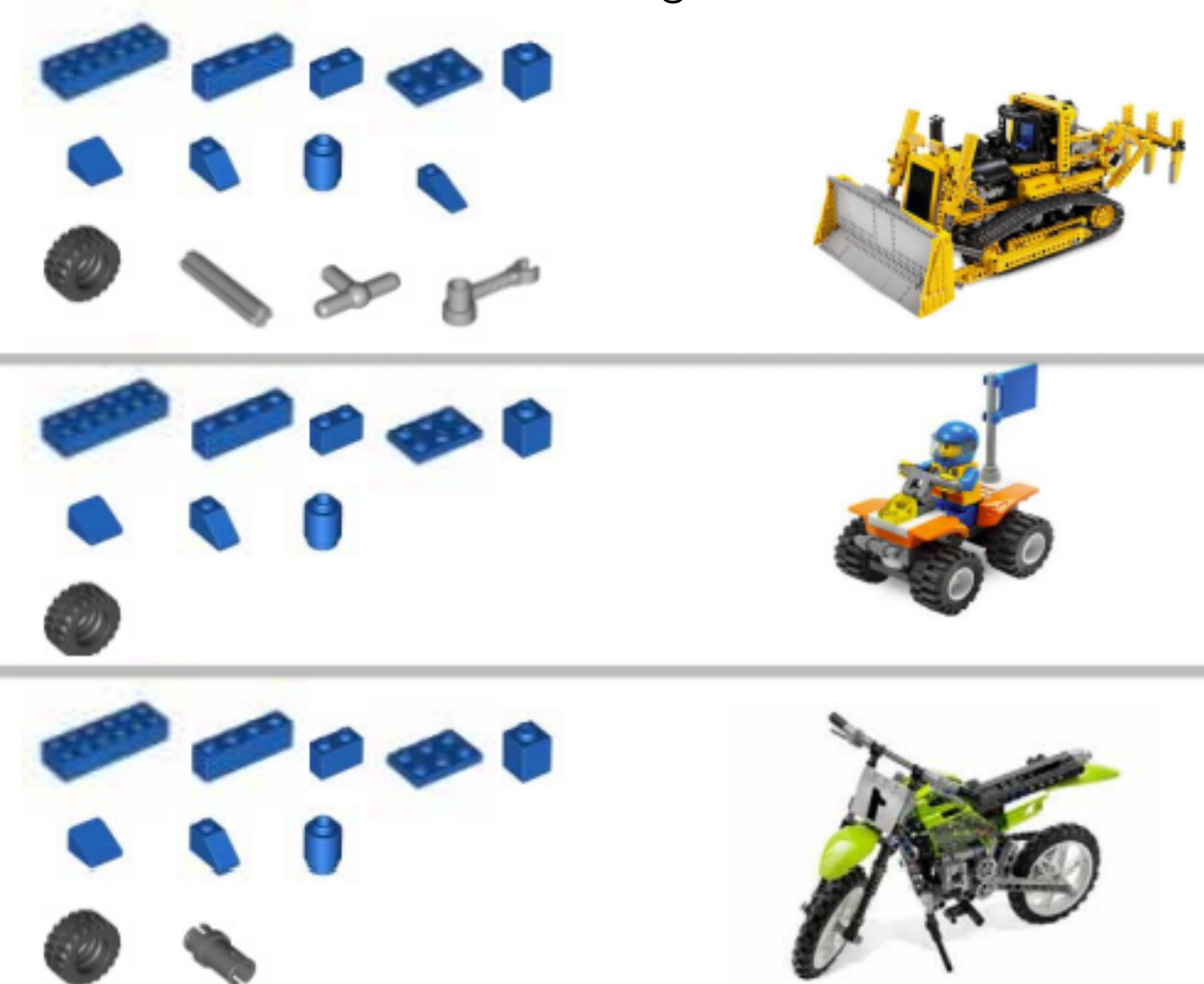
“

“Assim, o mercador ou comerciante, movido apenas pelo seu próprio interesse egoísta, é levado por uma mão invisível a promover algo que nunca fez parte do interesse dele: o bem-estar da sociedade”

-Adam Smith

Teórico do liberalismo econômico e
estudioso da causa da riqueza das nações

Segundo Adam Smith, a riqueza das nações está associada à divisão de trabalho e a especialização e a complexidade que emergem da economia



1.- A Country



2.- with a great diversity of Legos (capabilities)



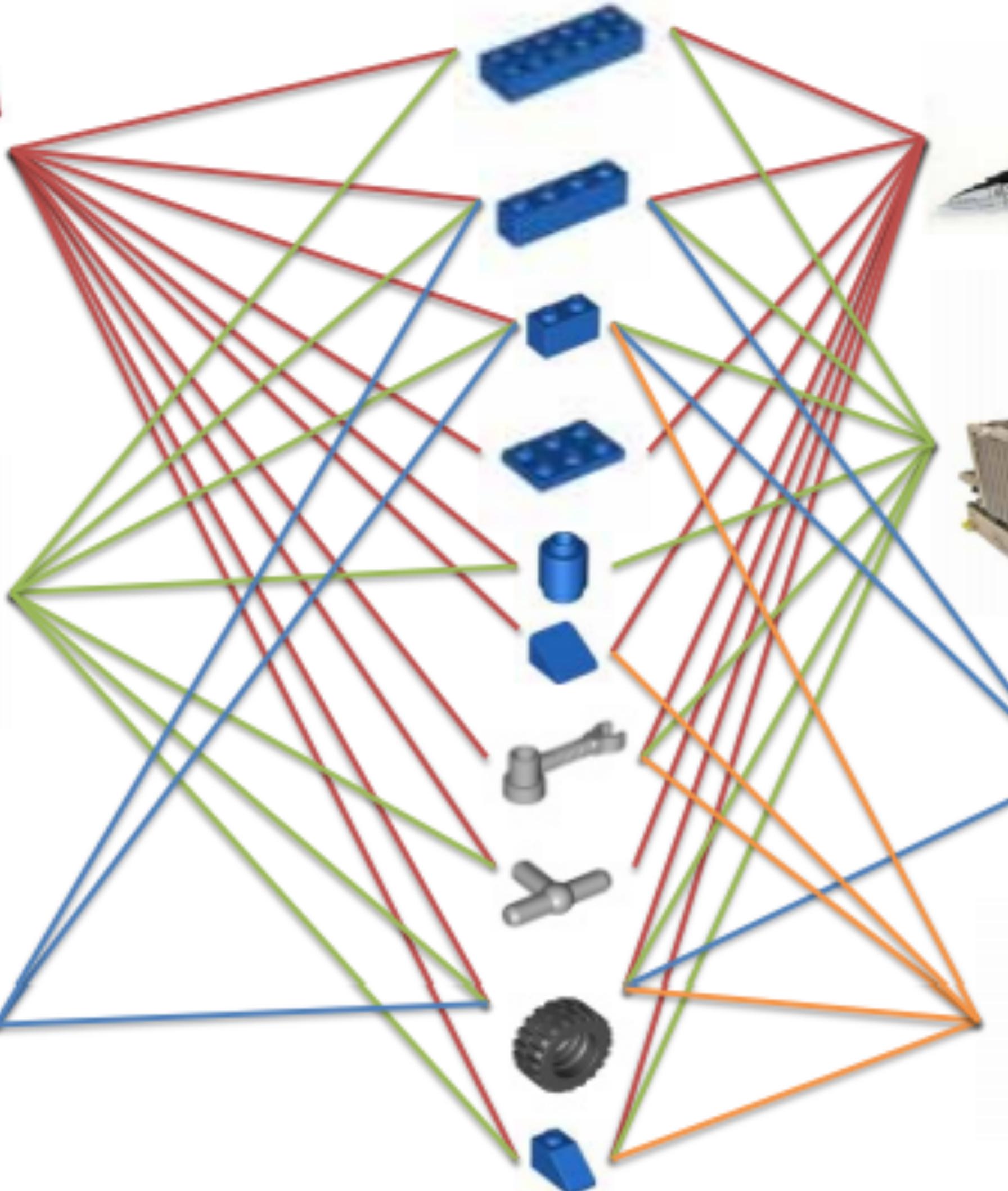
3.- Can make many products



Countries



Capabilities



Products



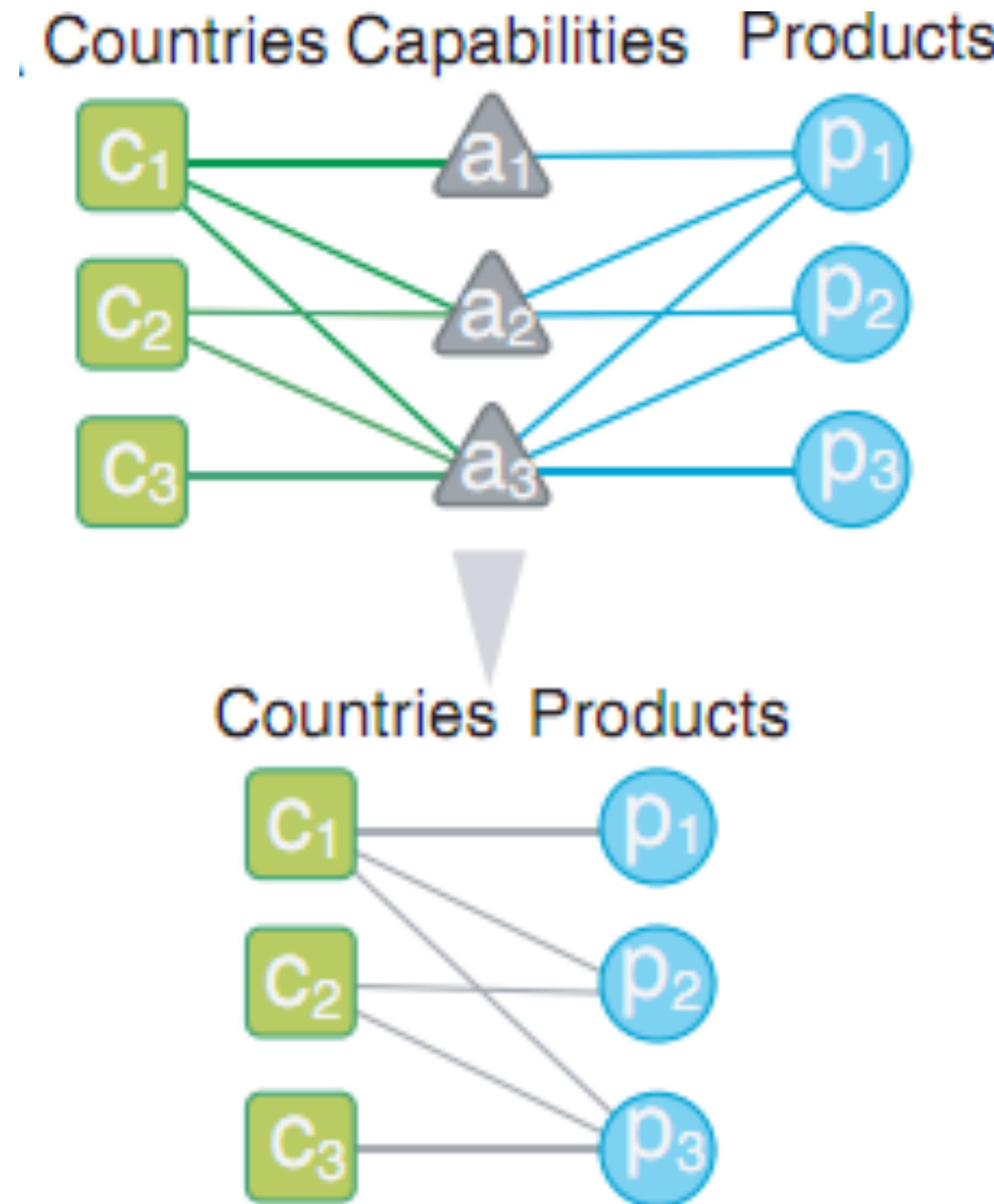
Countries



Products



Um grafo bipartido é um grafo cujos nós podem ser separados em dois conjuntos, ou partições, sendo que todas as arestas conectam nós em partições diferentes





In Lego Language



In Network Language

$$k_{e,0} = \sum_p M_{ep} \quad , \quad k_{p,0} = \sum_e M_{cp}$$
$$k_{e,n} = \frac{1}{k_{e,0}} \sum_p M_{ep} k_{p,n-1} \quad , \quad k_{p,n} = \frac{1}{k_{p,0}} \sum_e M_{cp} k_{e,n-1}$$

(Simple)
Few Capabilities



(Complex)
Diverse Set
of Capabilities



$$k_{c,0} = \sum_p M_{cp} , \quad k_{p,0} = \sum_c M_{cp}$$

$$k_{c,n} = \frac{1}{k_{c,0}} \sum_p M_{cp} k_{p,n-1} , \quad k_{p,n} = \frac{1}{k_{p,0}} \sum_c M_{cp} k_{c,n-1}$$

Reflexão	Tipo	Descrição
Kc,0	Diversificação	Número de produtos exportados por c
Kp,0	Ubiquidade	Números de países exportadores de p
Kc,1	Ubiquidade	Ubiquidade média dos produtos exportados por c
Kp,1	Diversificação	Diversificação média dos países exportadores de p
kc,2	Diversificação	Diversificação média dos países com cestas de
Kp,2	Ubiquidade	Ubiquidade média dos produtos exportados pelos países exportadores de p

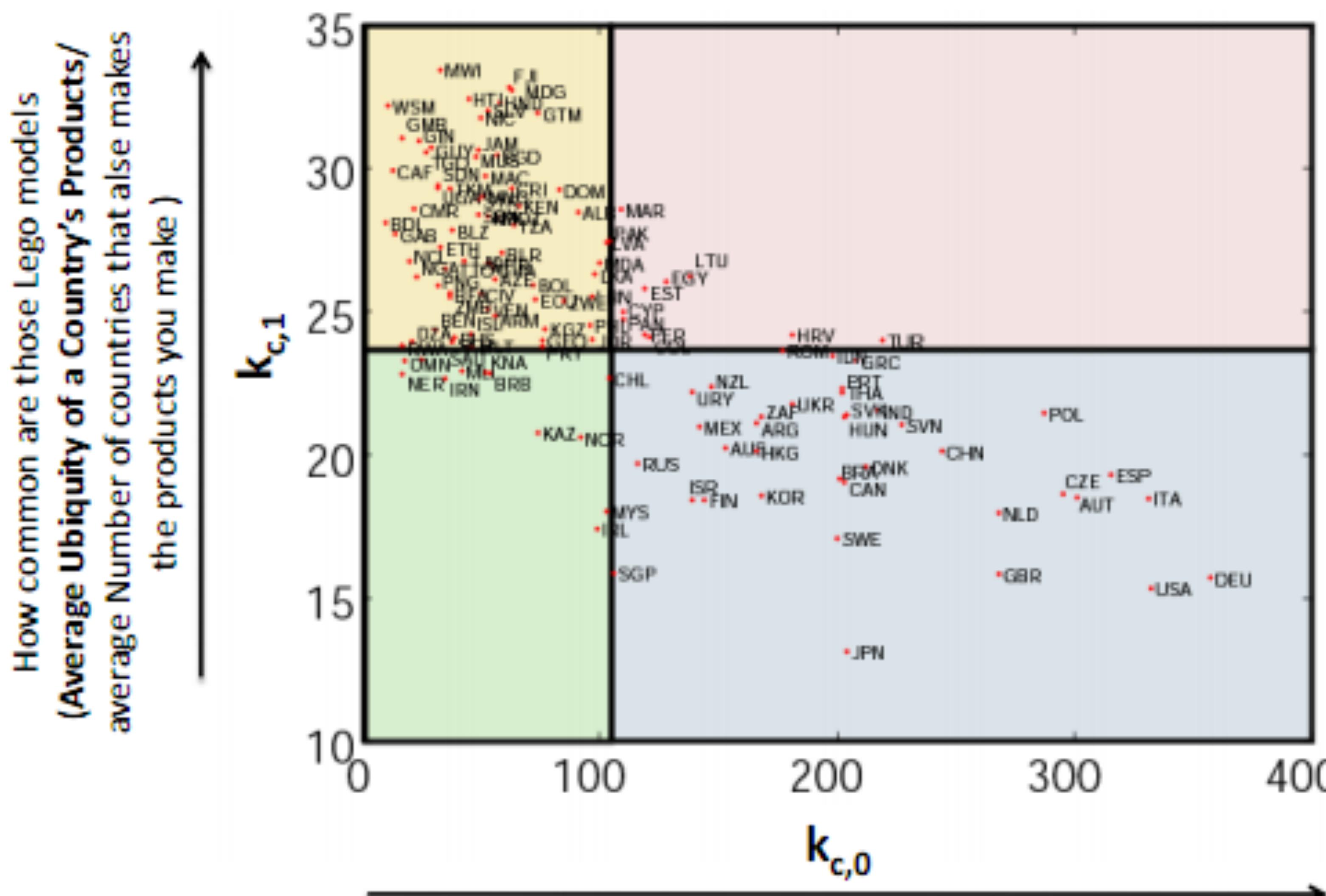
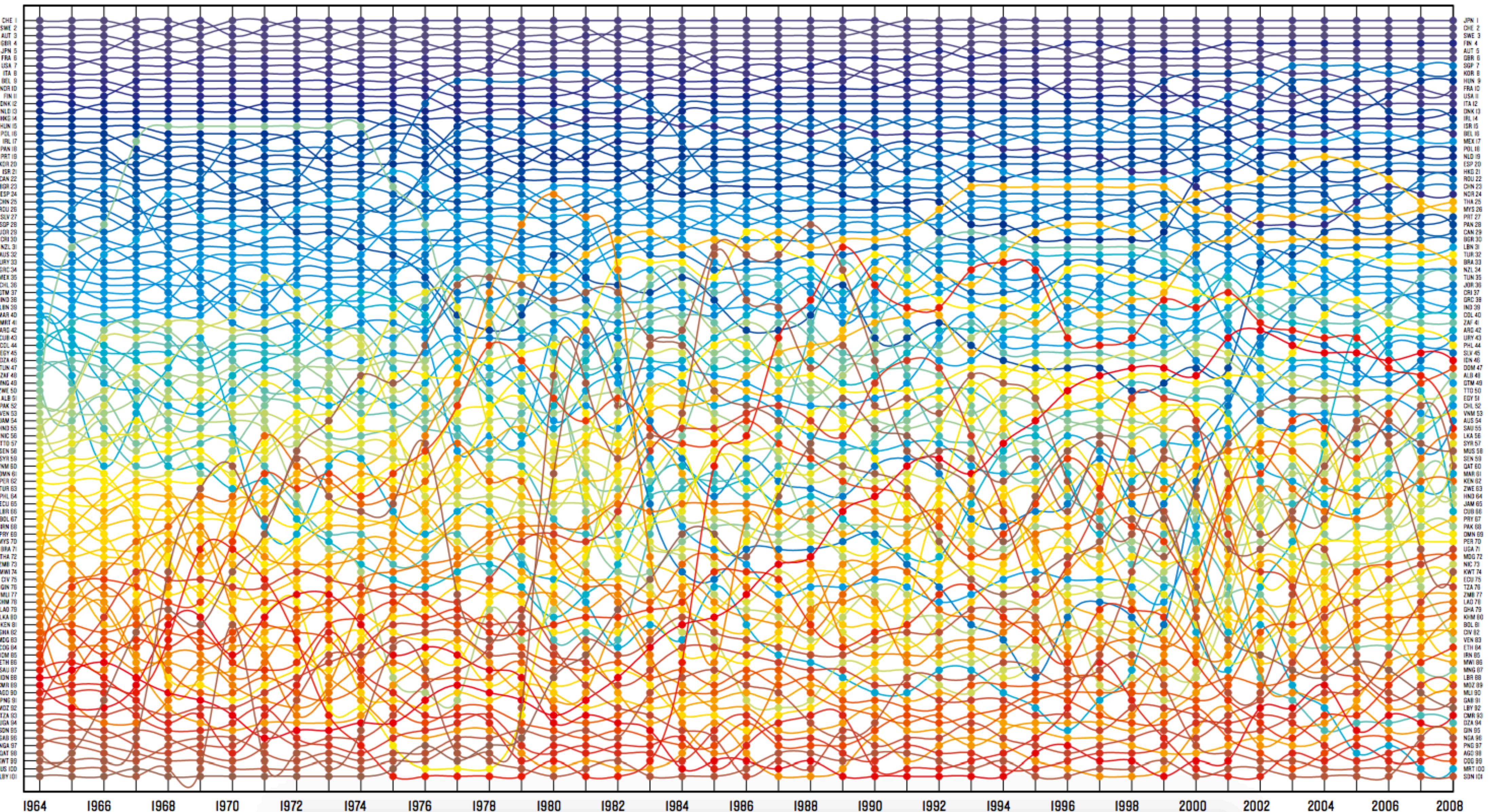


FIGURE 4.1:

► Evolution of the ranking of countries based on ECI between 1964 and 2008. Please see pages 352-353 for a larger version.



Correlação

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

- Análise de como variáveis quantitativas se relacionam e como uma afeta as outras
- Envolve a comparação de variáveis para verificar se a variação em um conjunto de valores interfere no outro sistematicamente e em que grau

1. Uma variável influencia na outra
2. As variáveis não se influenciam
3. As variáveis não se influenciam mutuamente mas são conectadas por uma terceira variável
4. As variáveis se correlacionam aparentemente mas devido à amostragem insuficiente ou enviesada

EXEMPLO DO CASO 3

- Pode-se mostrar uma forte correlação entre a taxa de mortalidade e a temperatura da cidade
- Pessoas mais velhas buscam viver em cidades mais quentes

EXEMPLO DO CASO 4

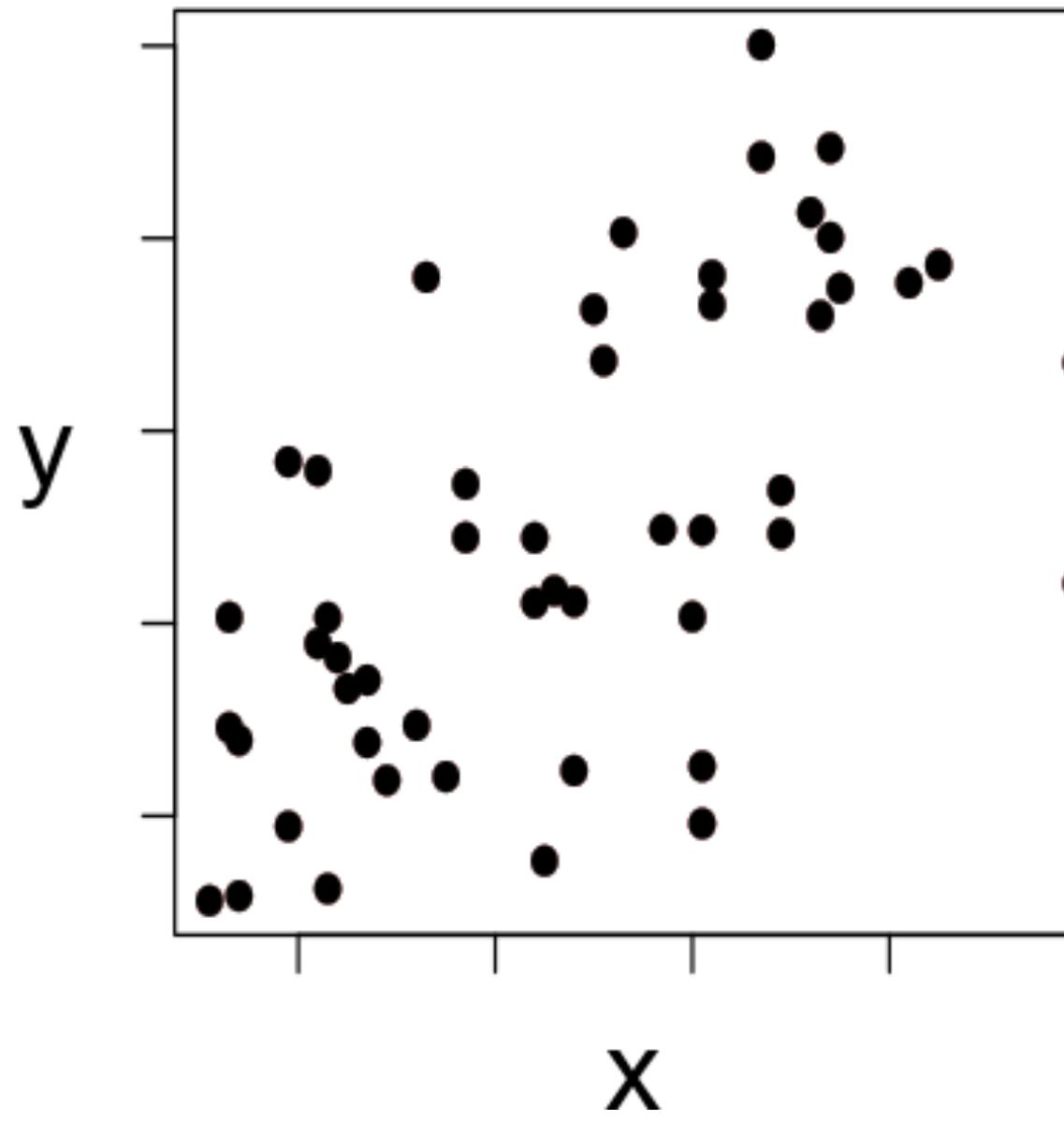
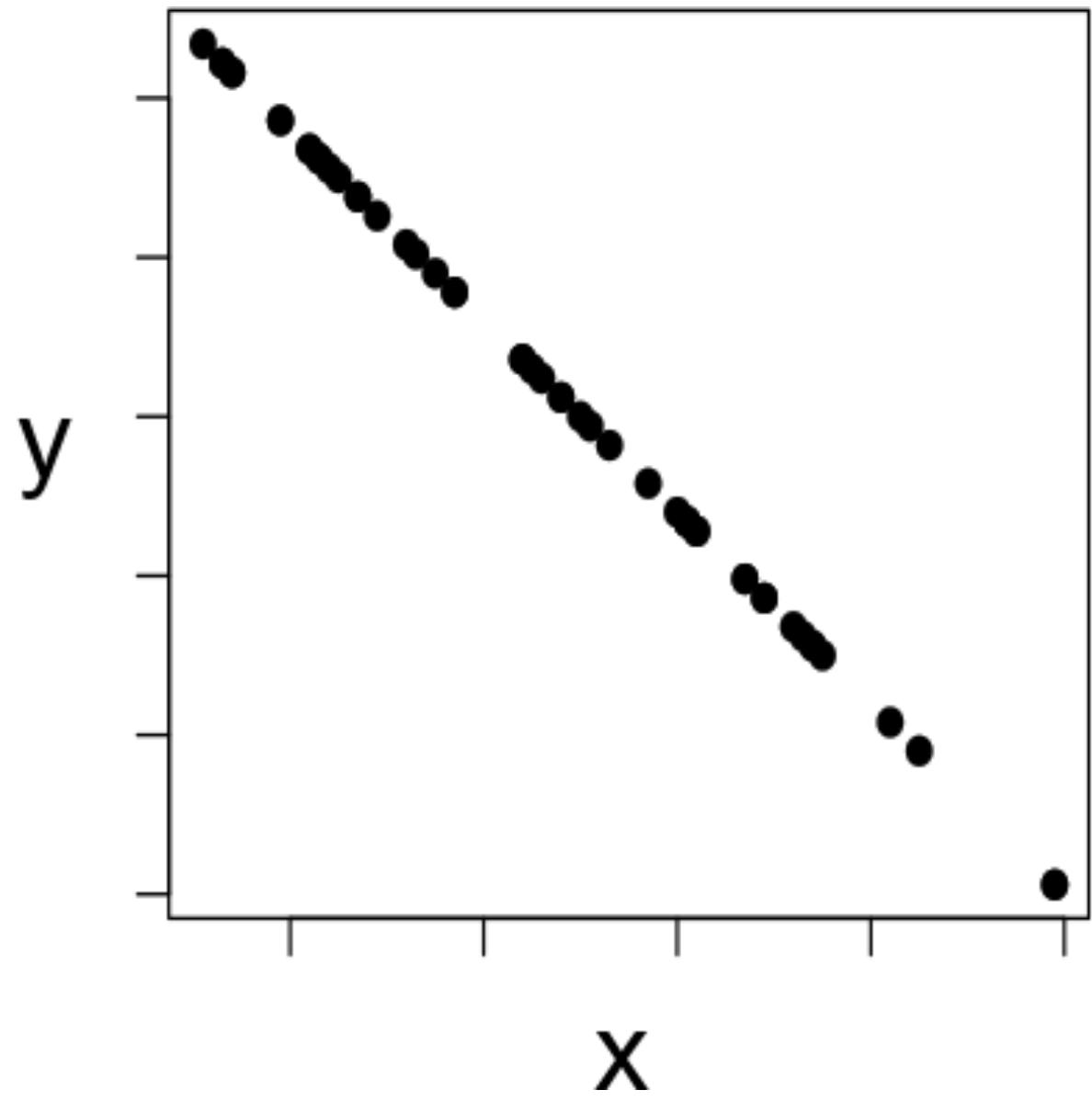
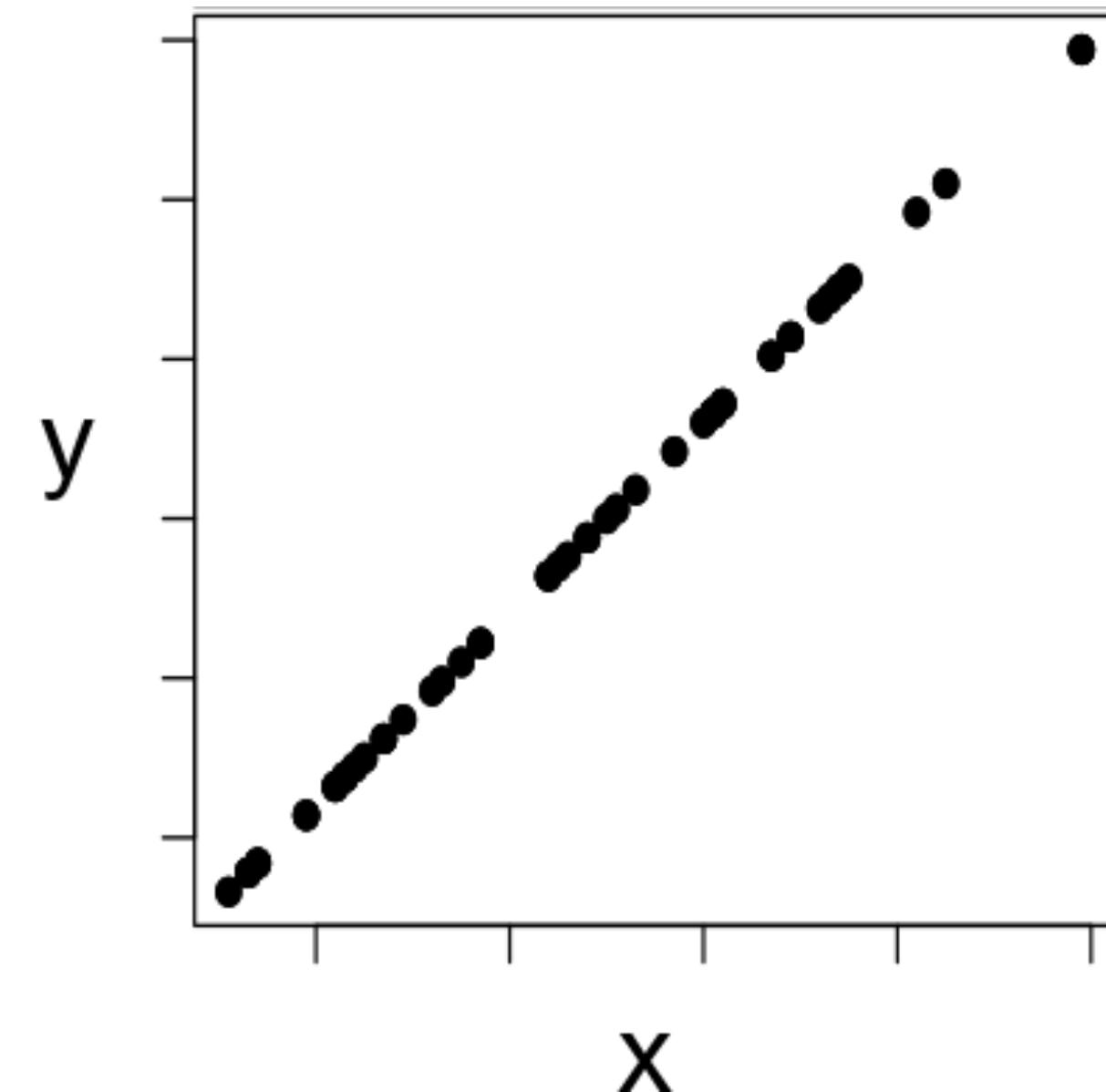
- Pode-se mostrar uma forte correlação entre salários de professores e o número de carros novos vendidos
- Ambas as variáveis são afetadas pelo estado da economia

PADRÕES ANALÍTICOS

PADRÕES ANALÍTICOS

- Direção
- Grau
- Forma

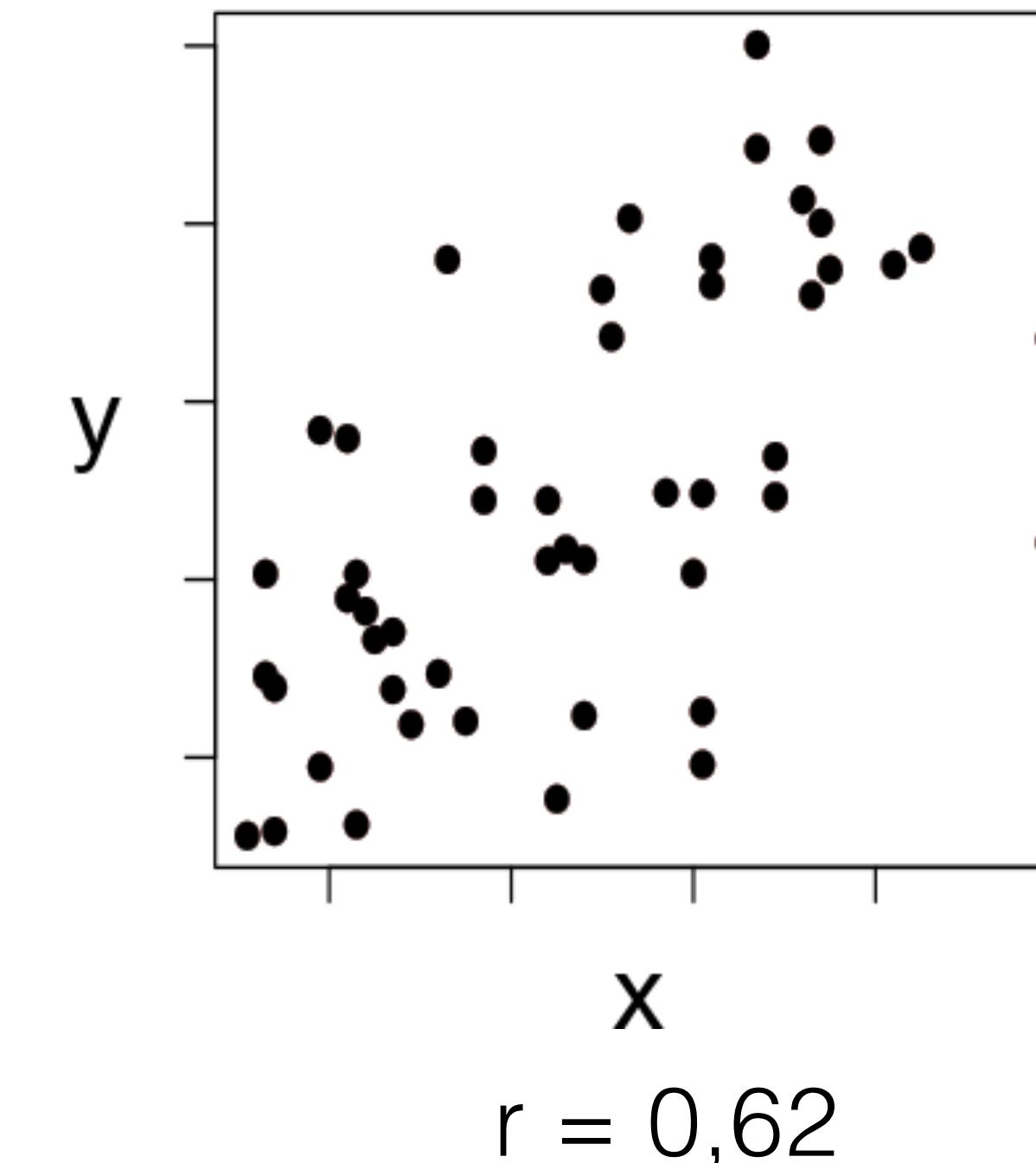
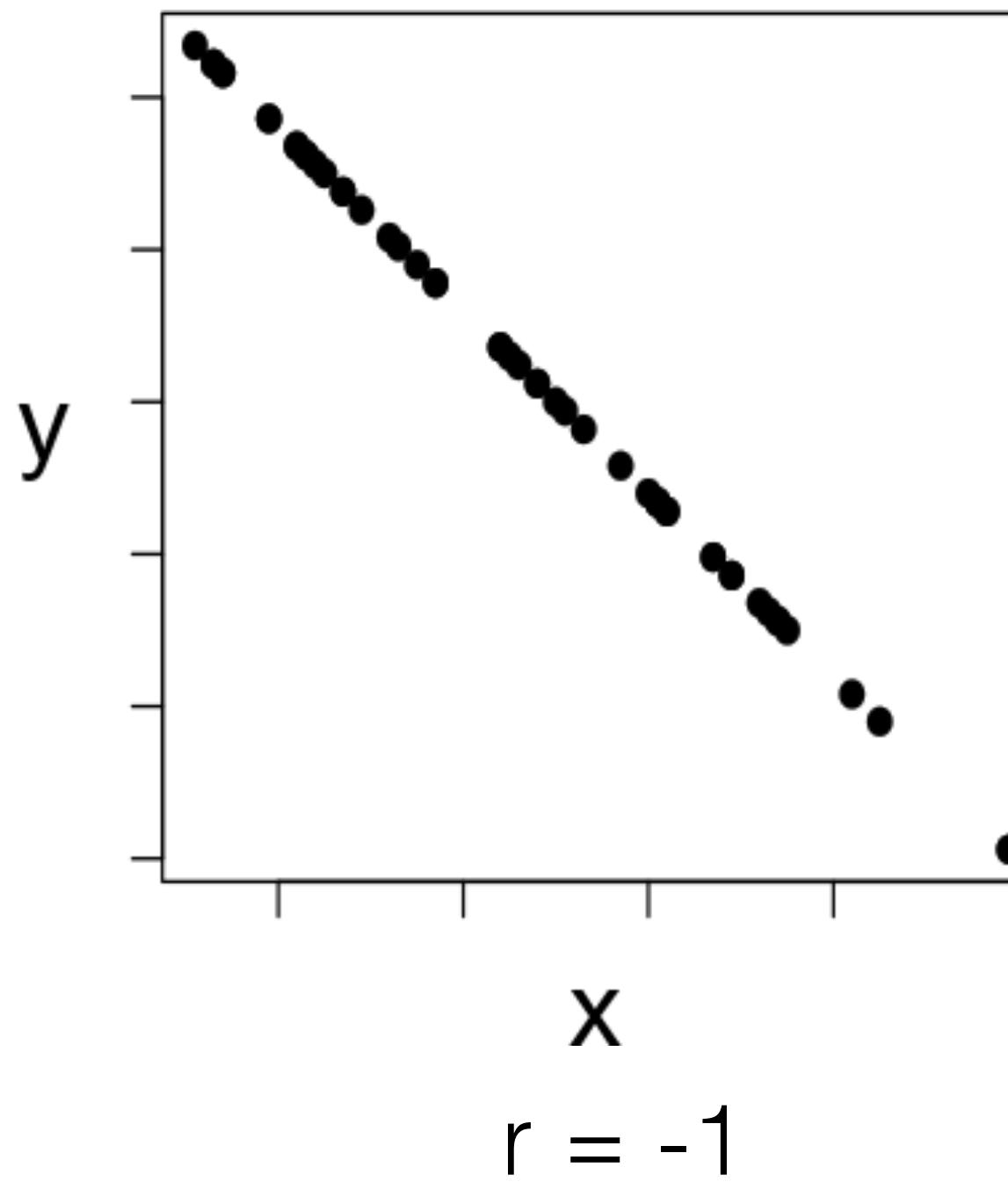
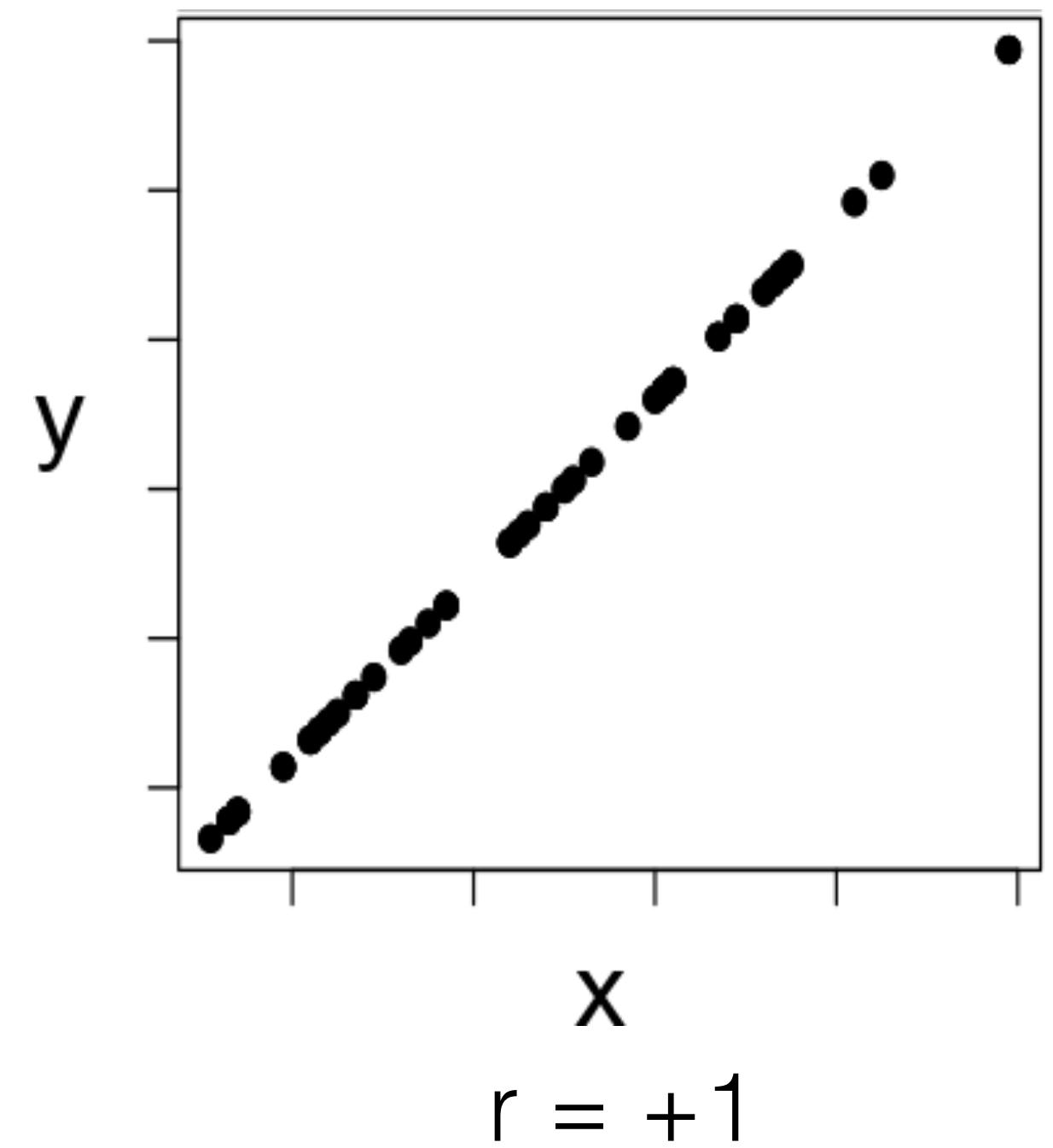
PADRÕES ANALÍTICOS



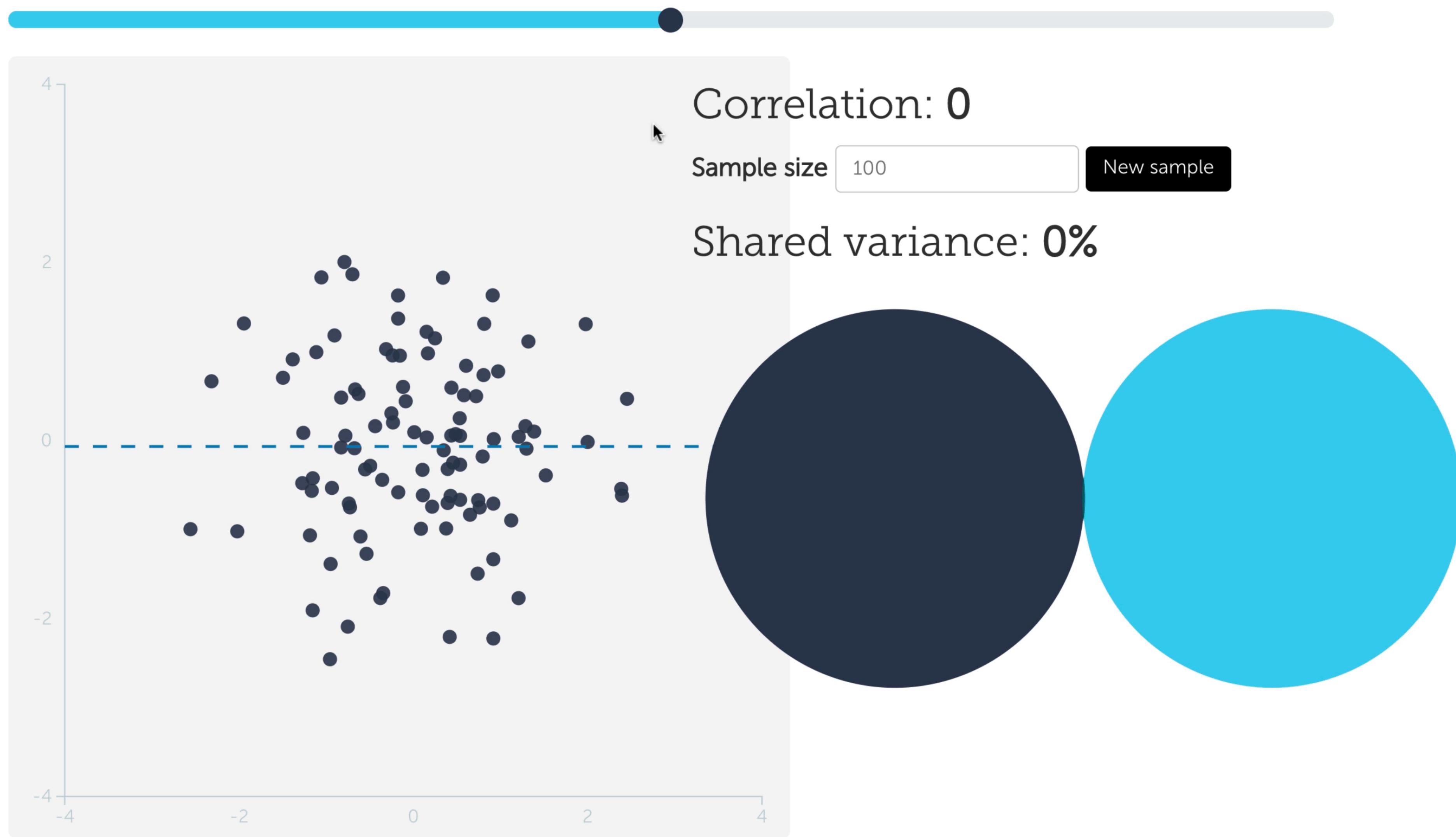
COEFICIENTE DE CORRELAÇÃO LINEAR

- O coeficiente de correlação linear r descreve a direção e o grau da correlação
 - $r = +1$: correlação forte e positiva
 - $r = -1$: correlação forte e negativa
 - $r = 0$: sem correlação

COEFICIENTE DE CORRELAÇÃO LINEAR



Slide me



COEFICIENTE DE DETERMINAÇÃO

- O coeficiente de determinação r^2 é sempre positivo não indicando a direção da correlação mas é útil pois retrata um percentual
- $r^2 = 0,986$ indica que 98,6% da mudança da variável dependente pode ser determinada pela variável independente

ANScombe's Quartet

Data set	1-3	1	2	3	4	
Variable	x	y	y	y	x	y
Obs. no.	i :					
1 :	10.0	8.04	9.14	7.46	8.0	6.58
2 :	8.0	6.95	8.14	6.77	8.0	5.76
3 :	13.0	7.58	8.74	12.74	8.0	7.71
4 :	9.0	8.81	8.77	7.11	8.0	8.84
5 :	11.0	8.33	9.26	7.81	8.0	8.47
6 :	14.0	9.96	8.10	8.84	8.0	7.04
7 :	6.0	7.24	6.13	6.08	8.0	5.25
8 :	4.0	4.26	3.10	5.39	19.0	12.50
9 :	12.0	10.84	9.13	8.15	8.0	5.56
10 :	7.0	4.82	7.26	6.42	8.0	7.91
11 :	5.0	5.68	4.74	5.73	8.0	6.89

TABLE. Four data sets, each comprising 11 (x, y) pairs.

Each of the four data sets yields the same standard output from a typical regression program, namely

Number of observations (n) = 11

Mean of the x 's (\bar{x}) = 9.0

Mean of the y 's (\bar{y}) = 7.5

Regression coefficient (b_1) of y on x = 0.5

Equation of regression line: $y = 3 + 0.5x$

Sum of squares of $x - \bar{x}$ = 110.0

Regression sum of squares = 27.50 (1 d.f.)

Residual sum of squares of y = 13.75 (9 d.f.)

Estimated standard error of b_1 = 0.118

Multiple R^2 = 0.667

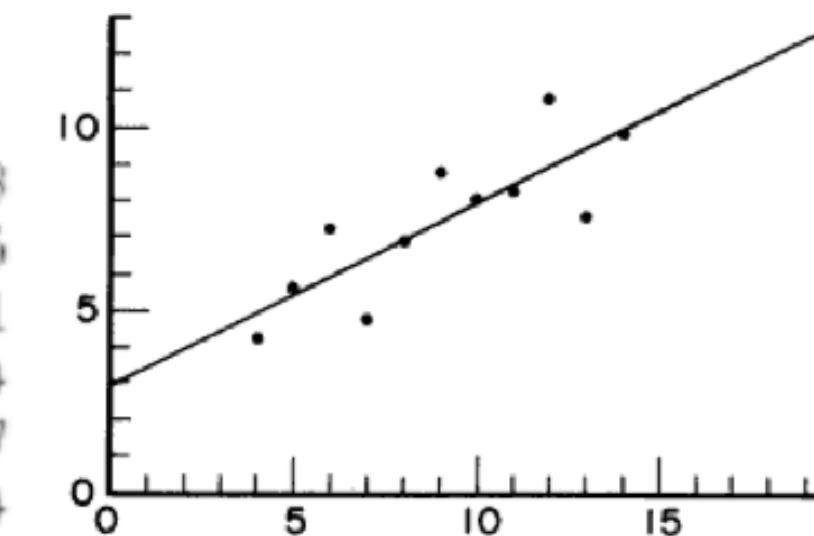


Figure 1

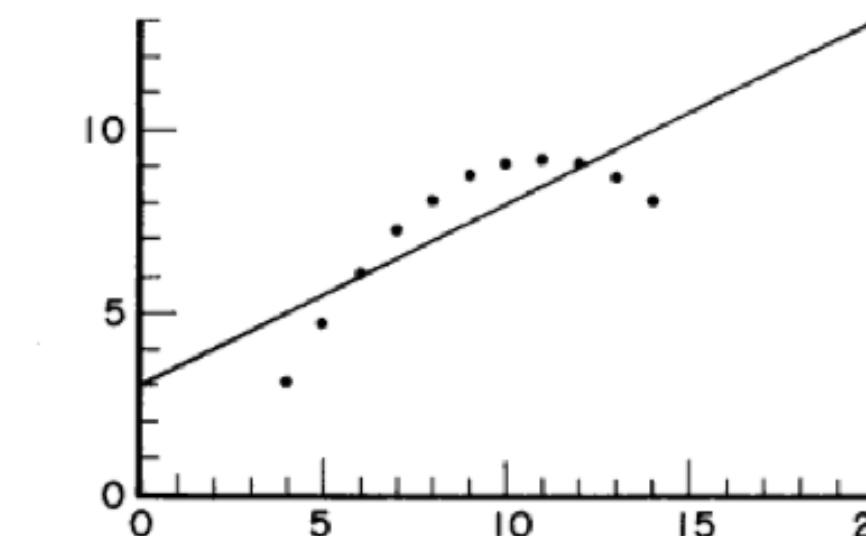


Figure 2

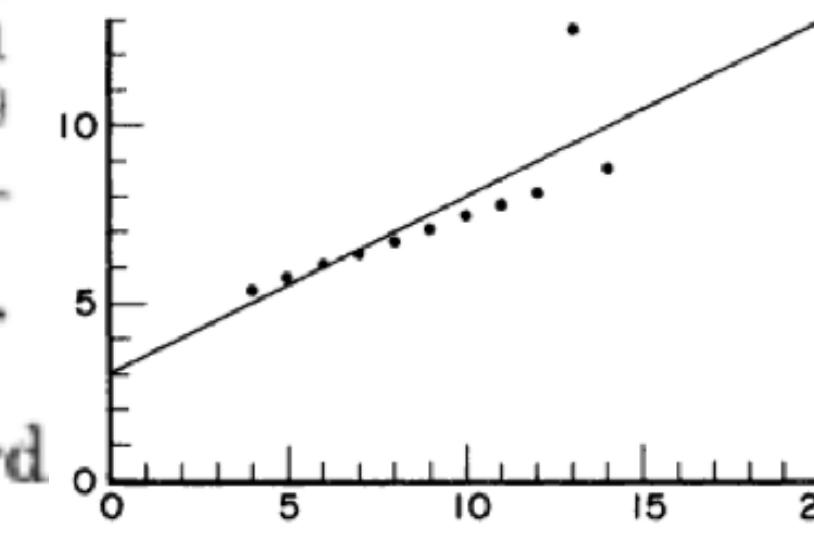


Figure 3

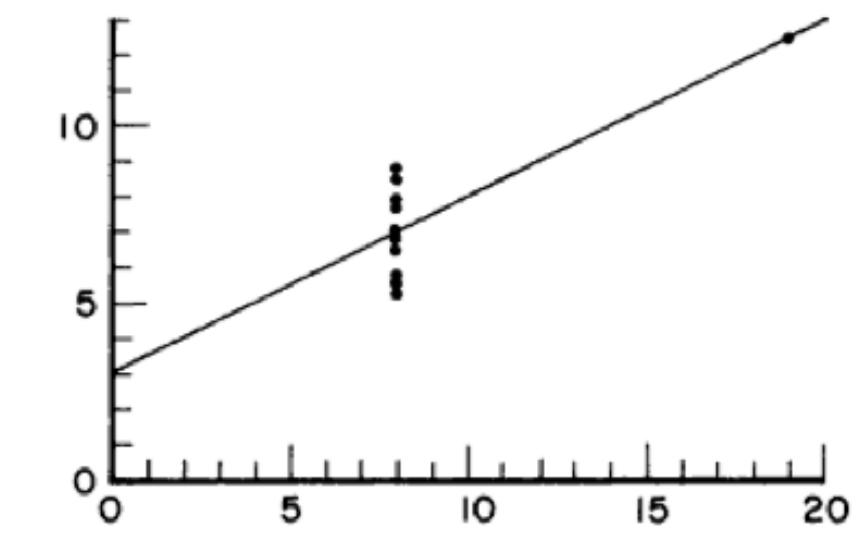
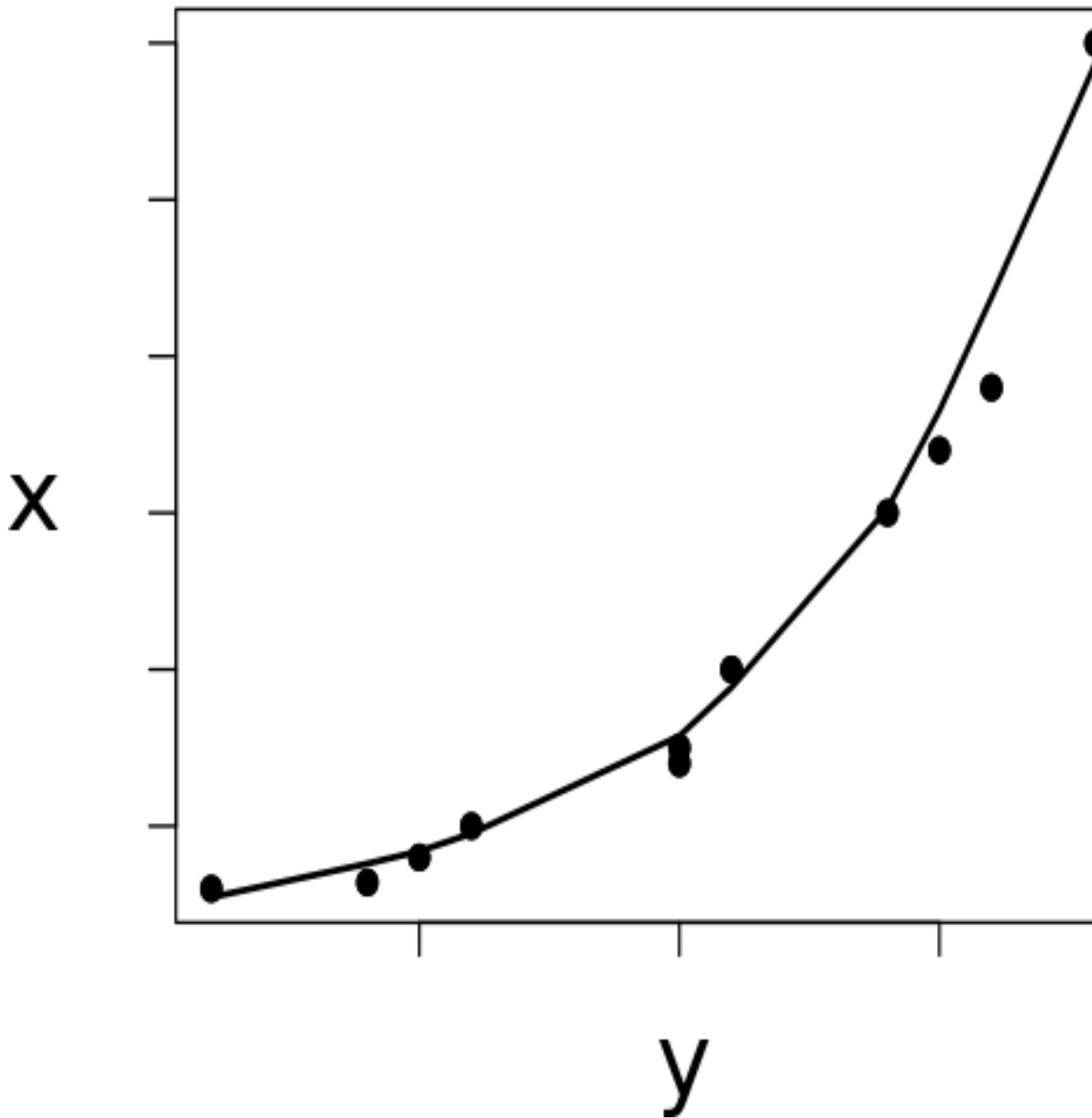
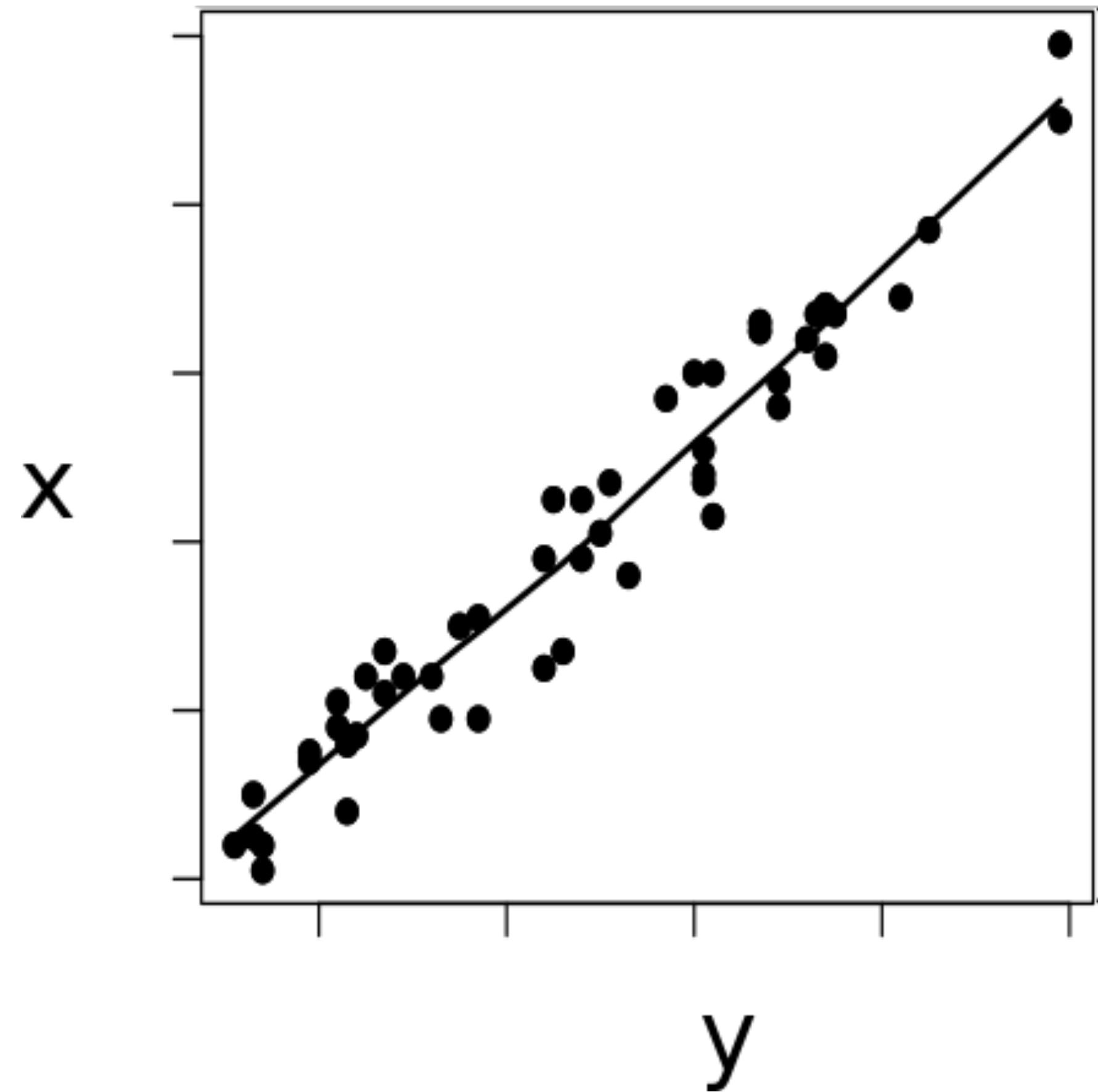


Figure 4

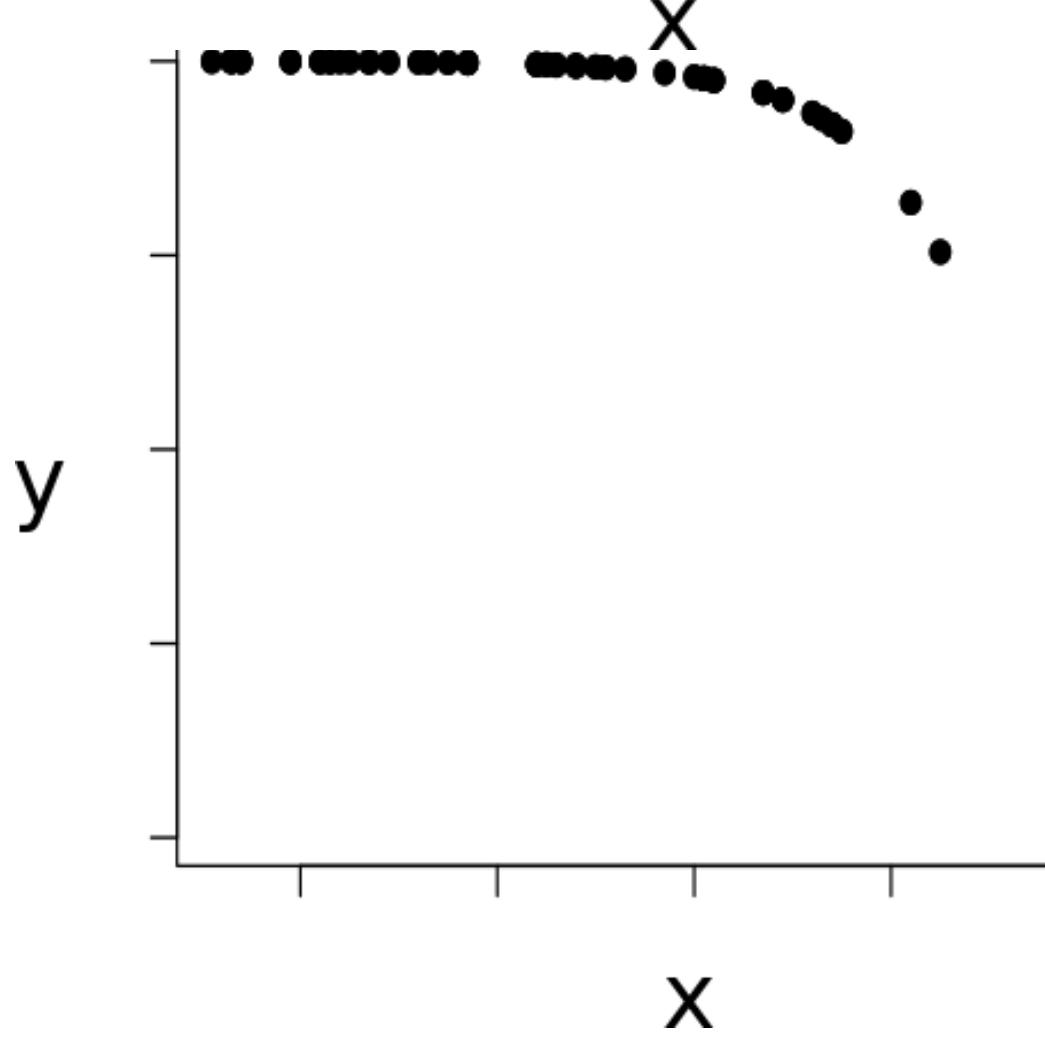
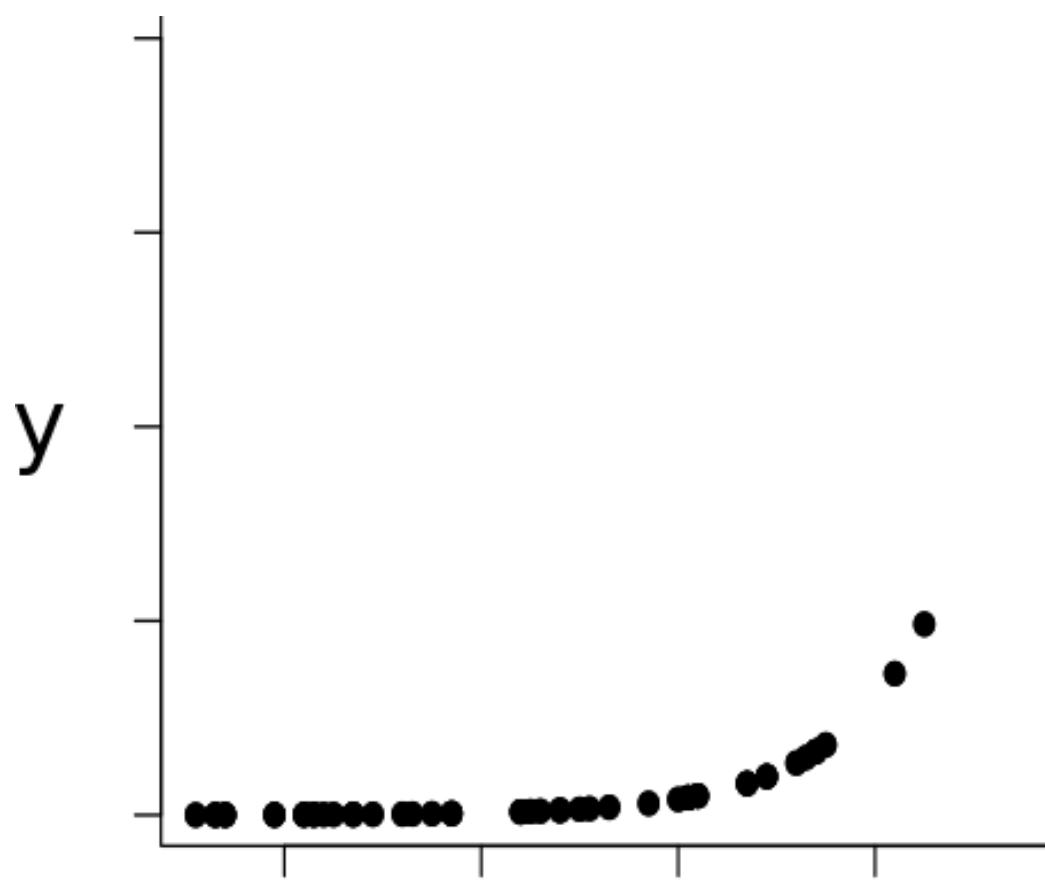
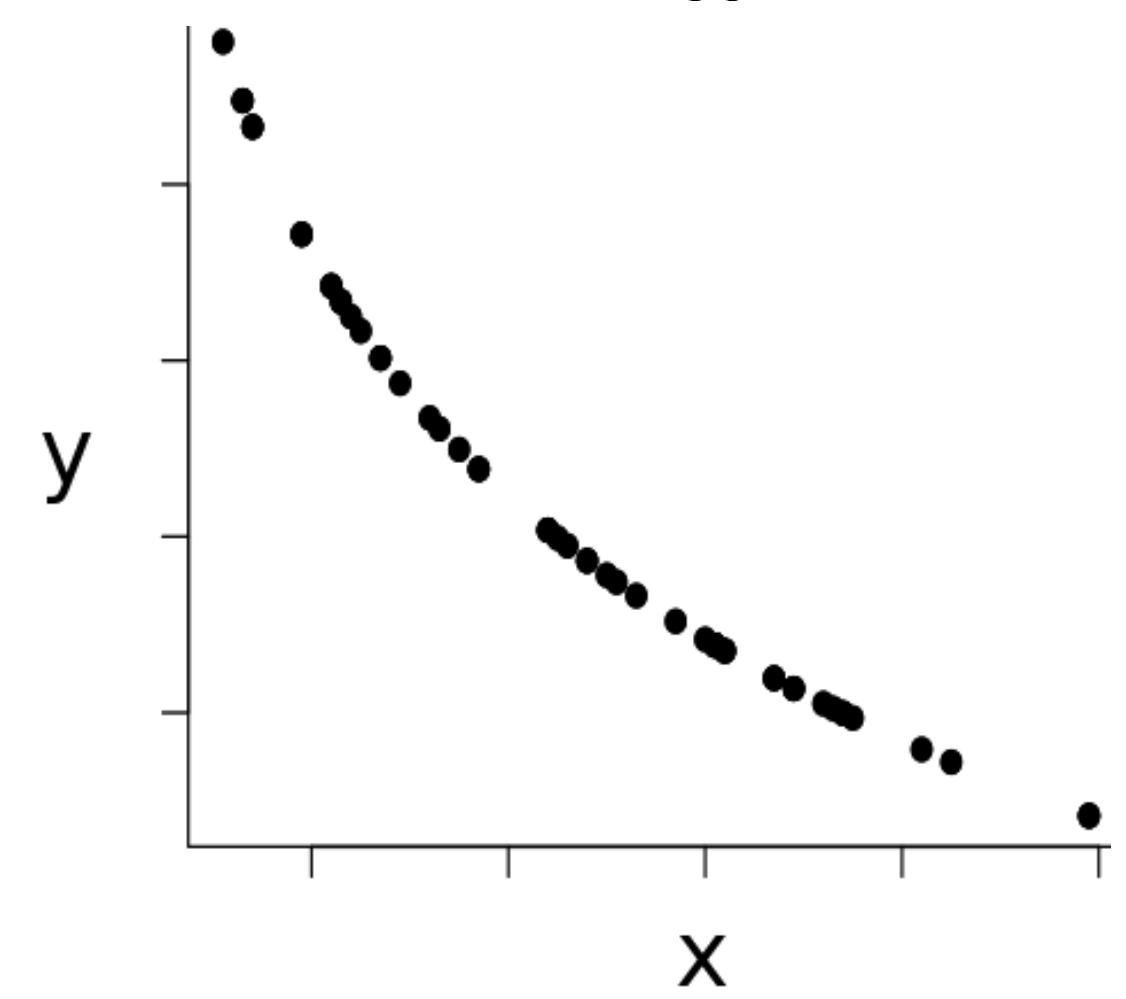
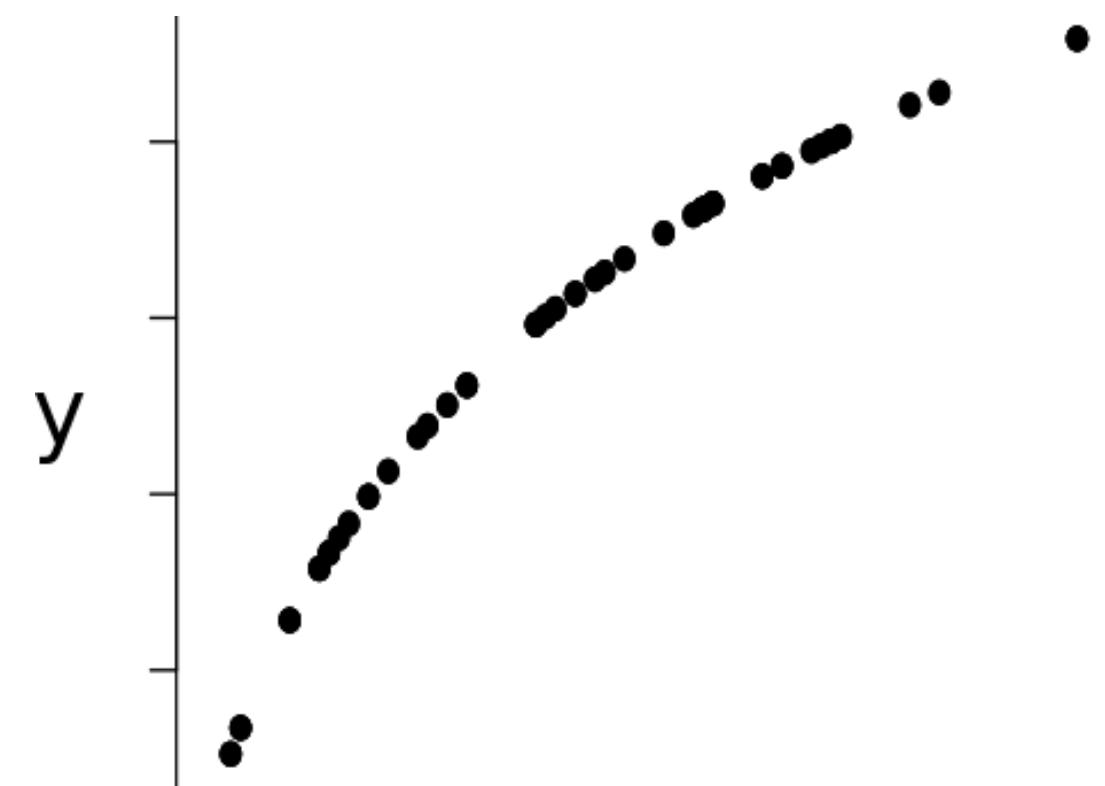
FORMA

Reta ou curva?



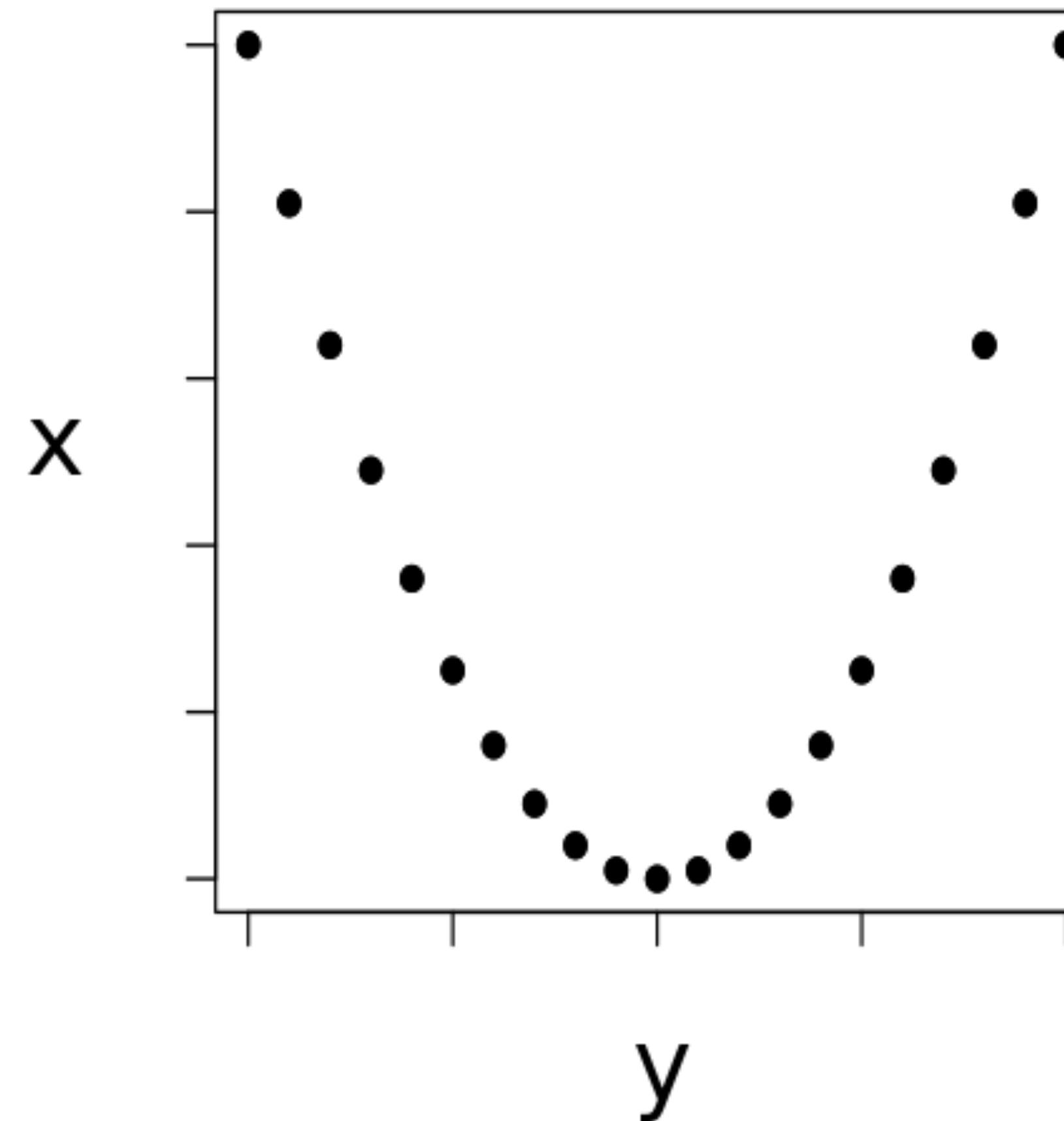
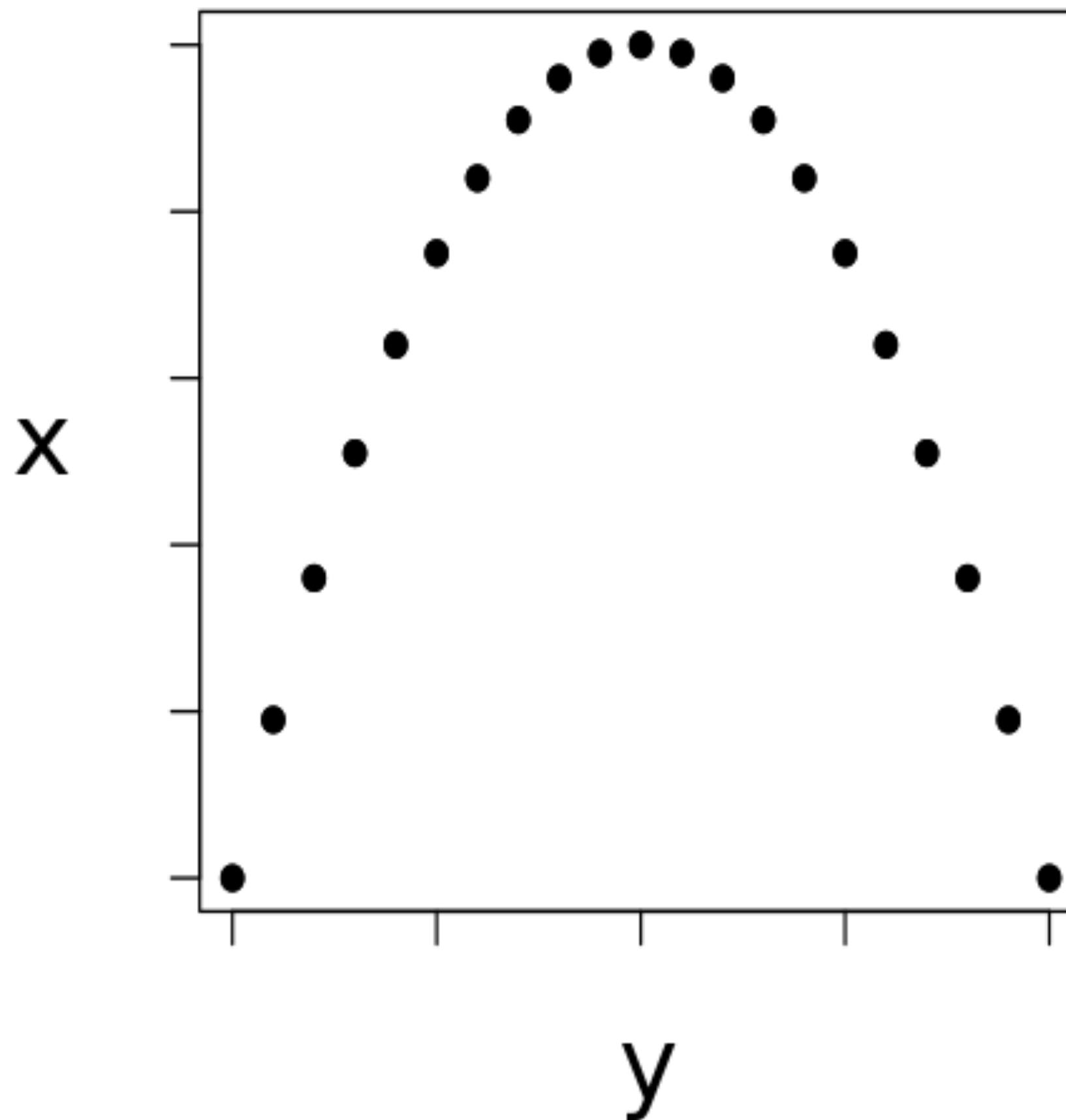
FORMA

Logarítmica ou exponencial?

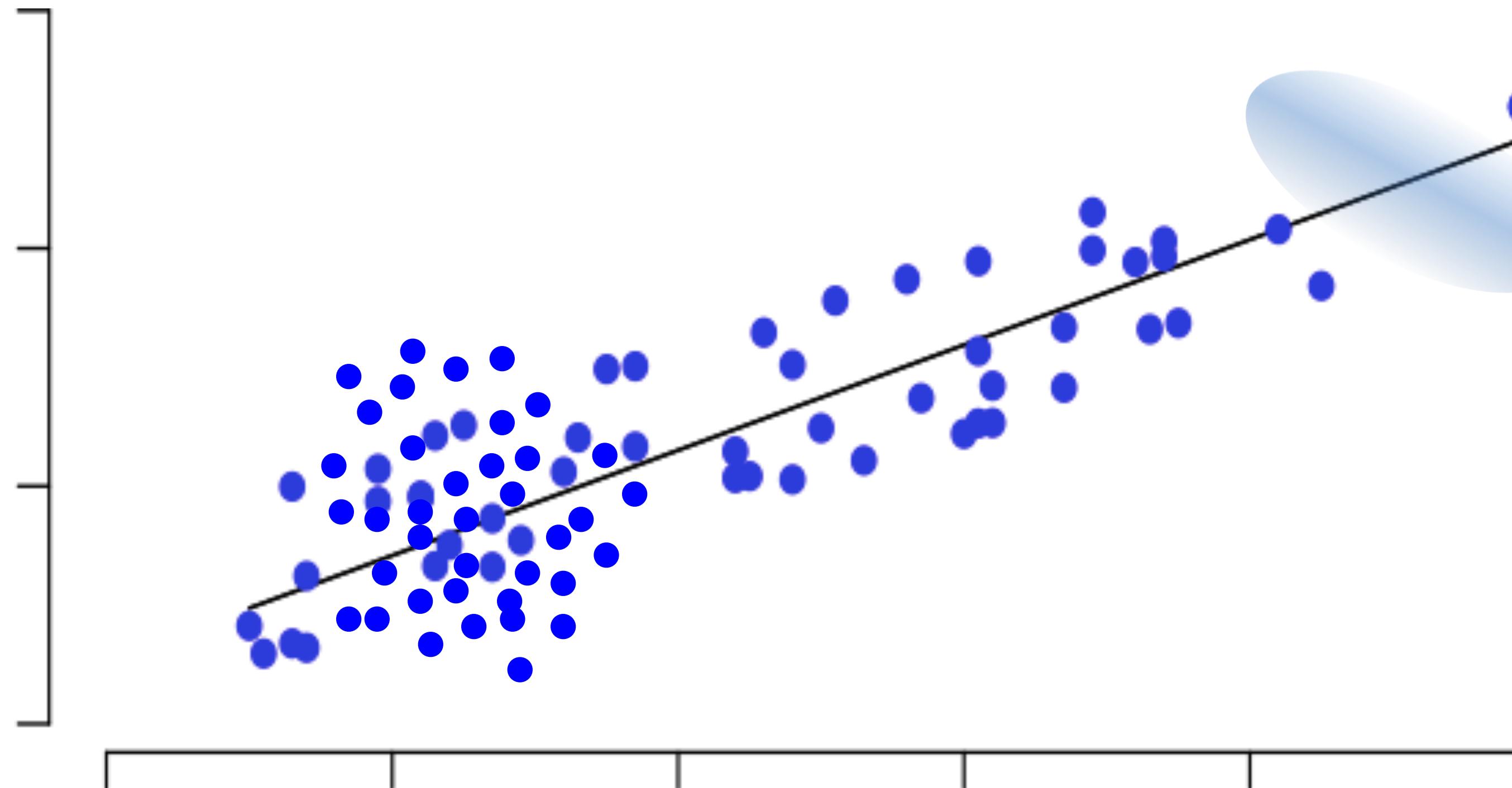


FORMA

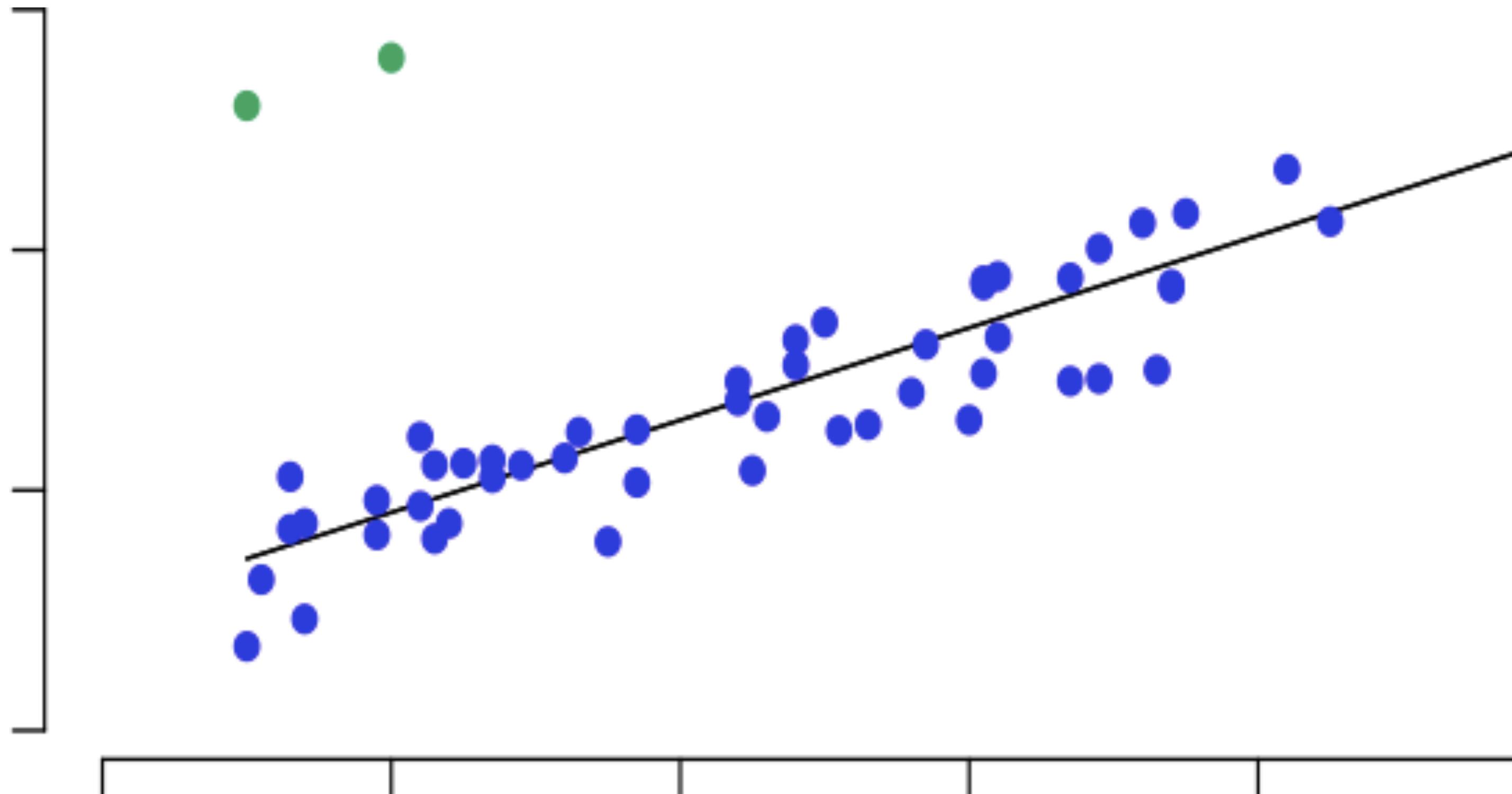
Certas correlações podem ser não monotônicas



LACUNAS E CONCENTRAÇÕES



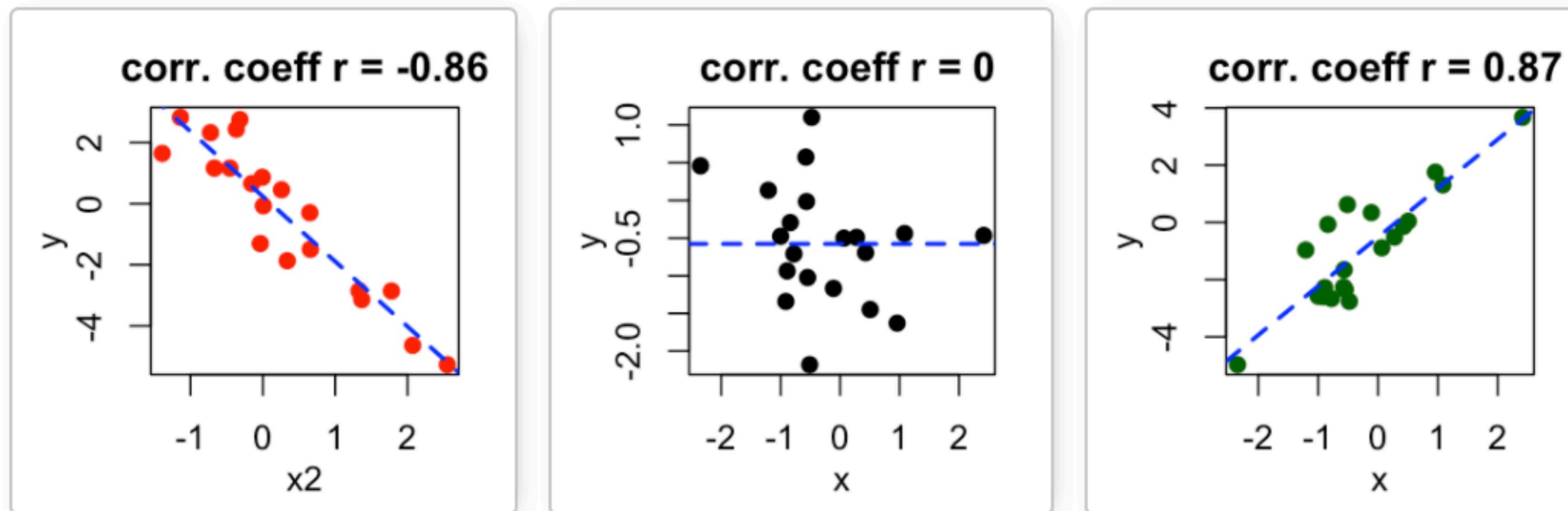
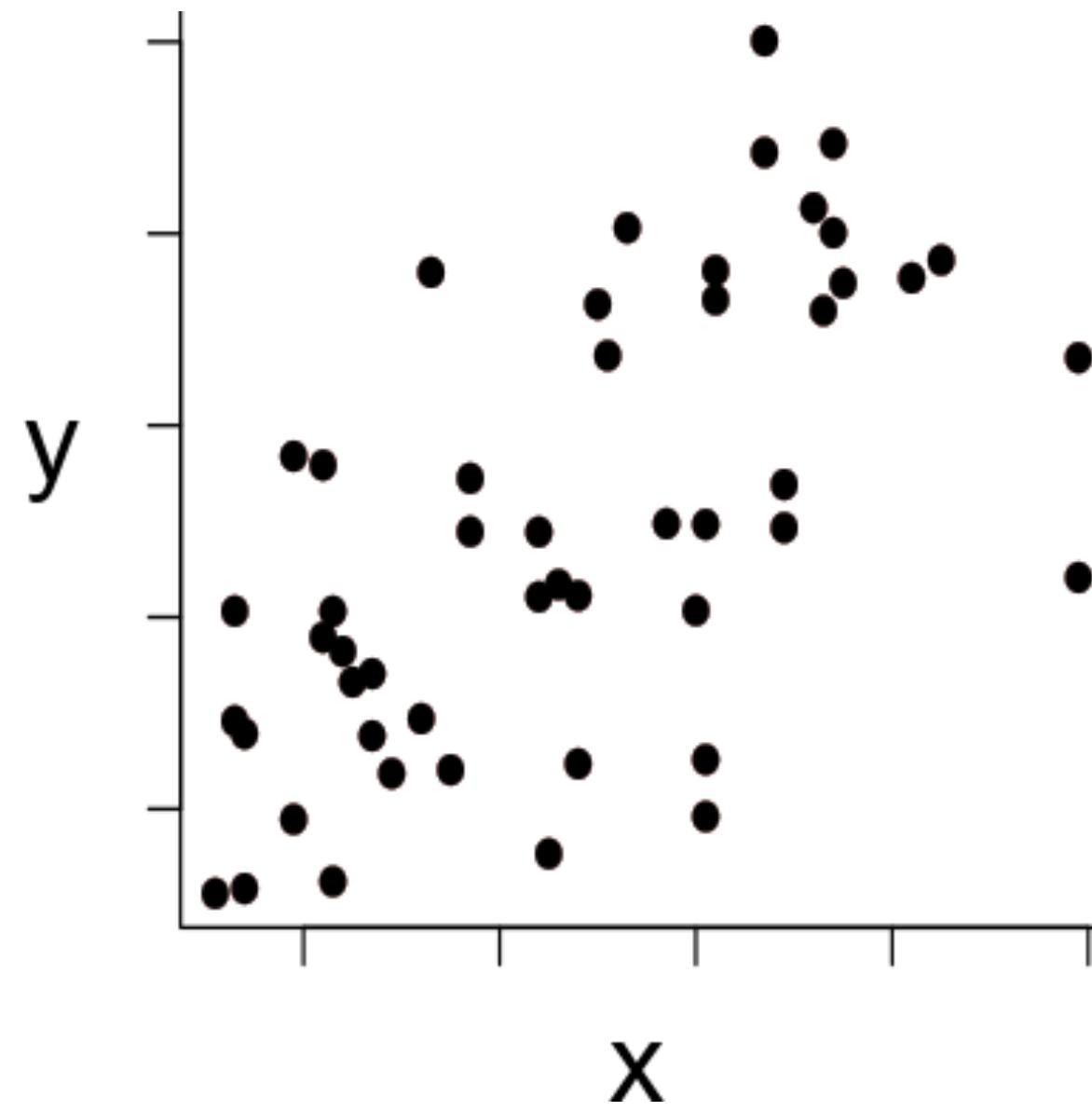
EXCEÇÕES

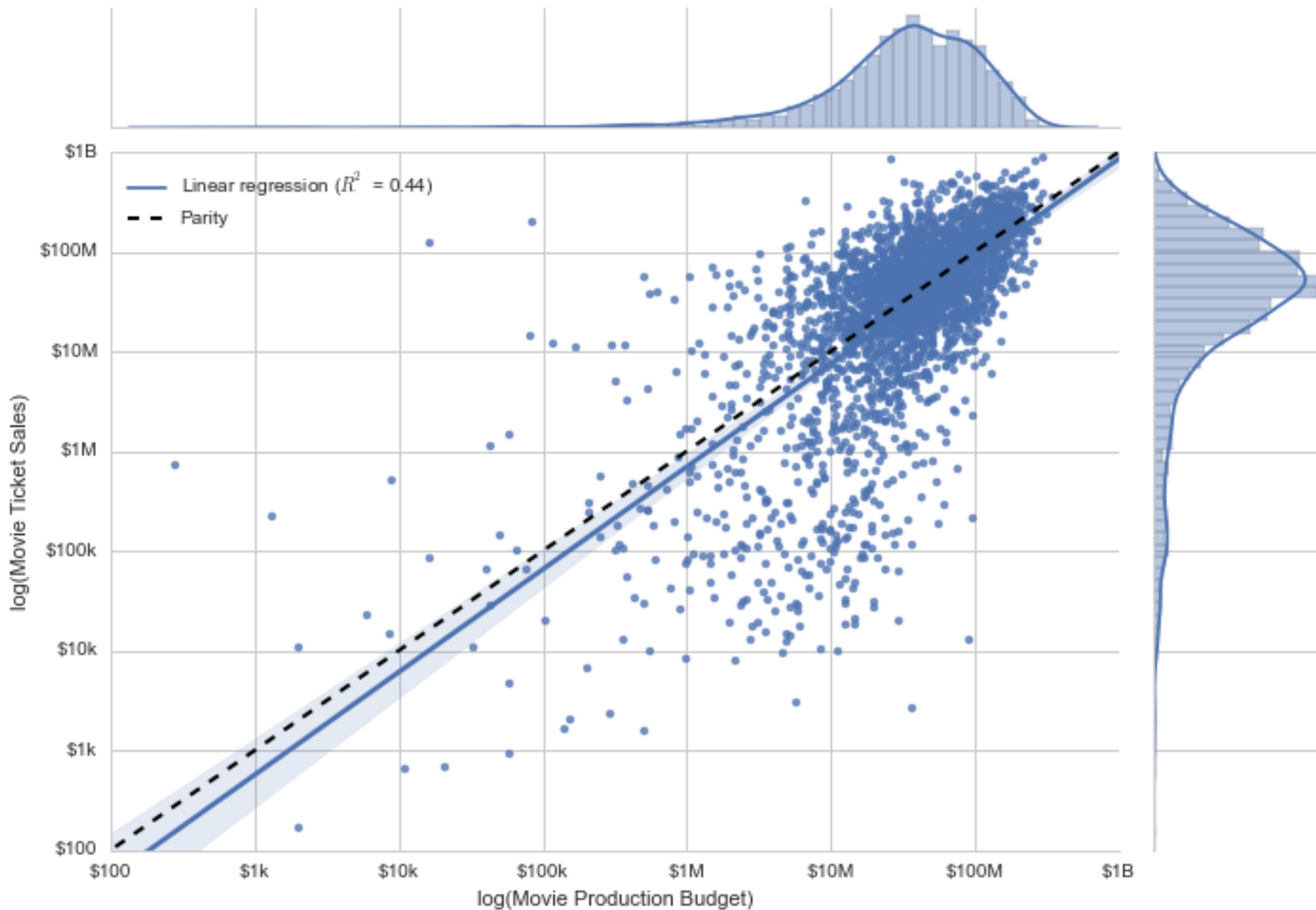


TÉCNICAS DE VISUALIZAÇÃO

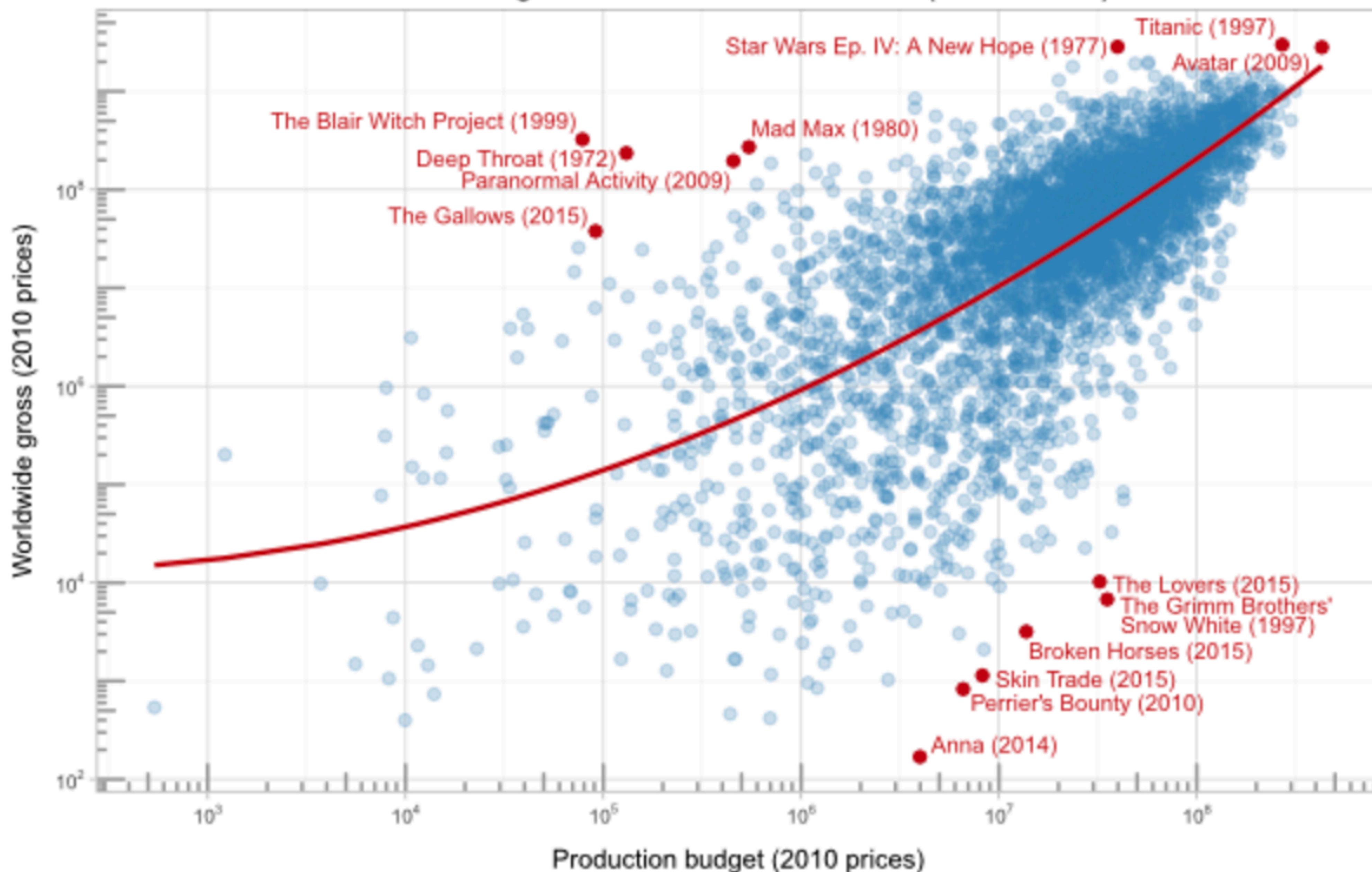
GRÁFICO DE DISPERSÃO

Úteis para analisar a correlação entre duas variáveis



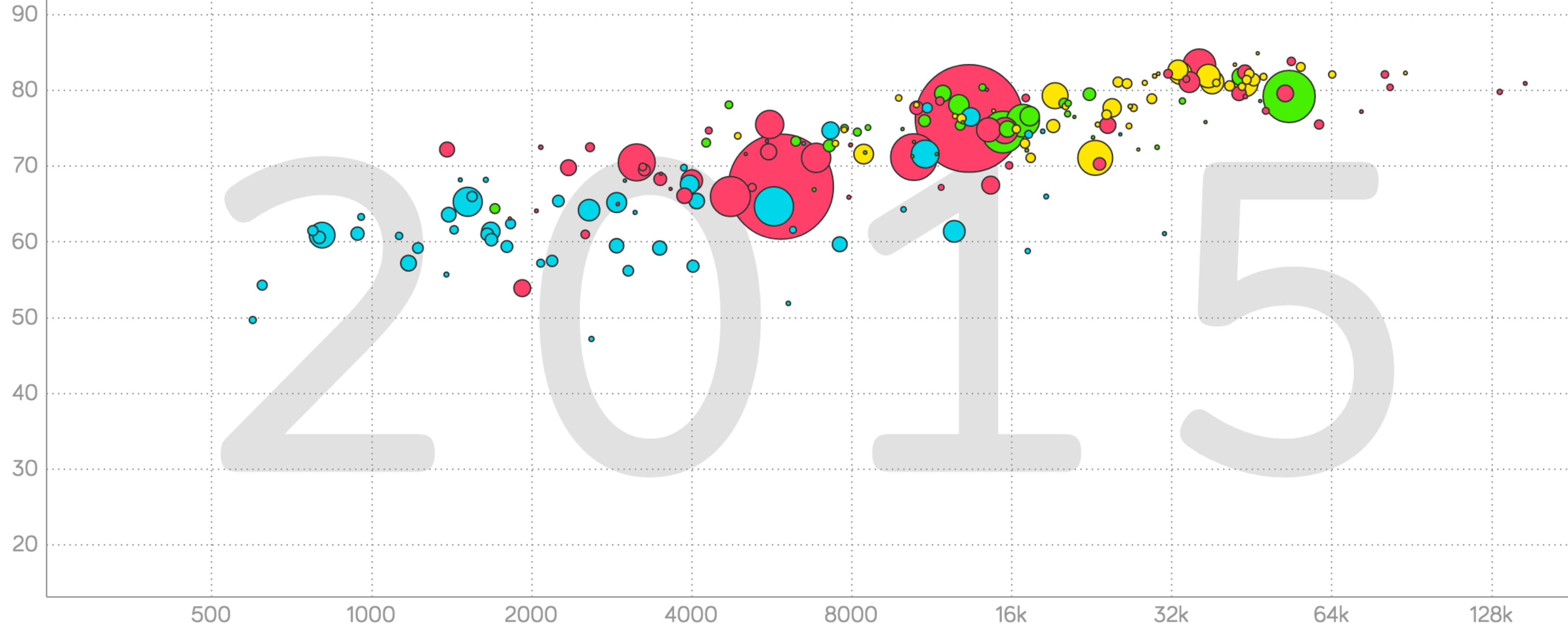


Movie budgets and box office success (1955–2015)



MOTION CHART

Life expectancy, years



Income per person, GDP/capita in \$/year adjusted for inflation & prices

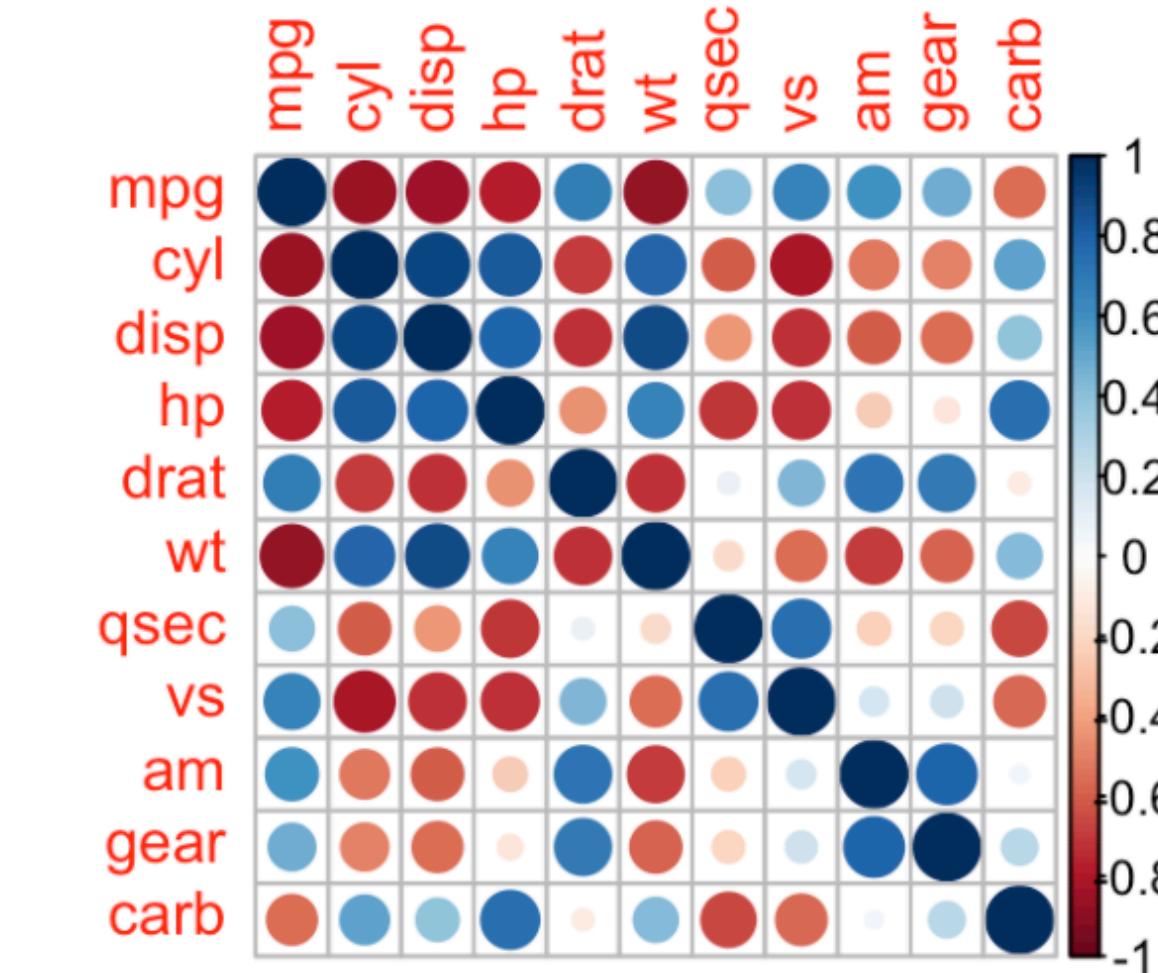
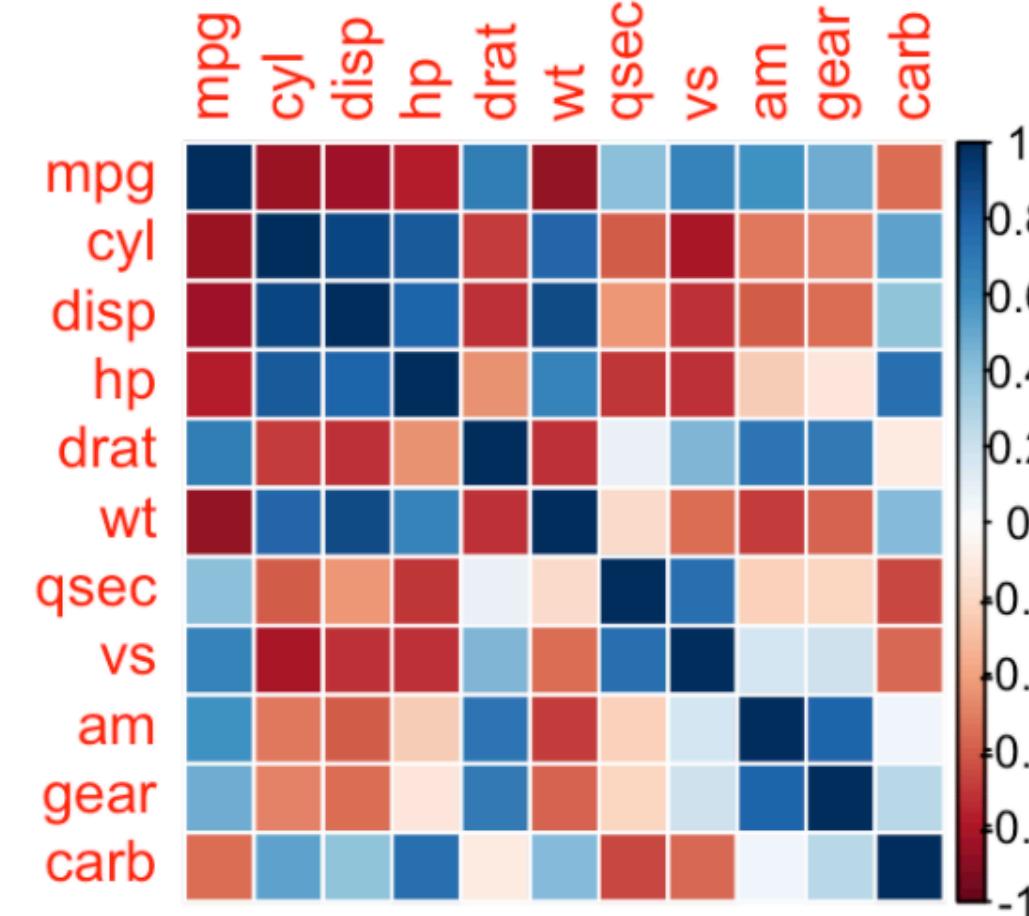


⚠ DATA DOUBTS

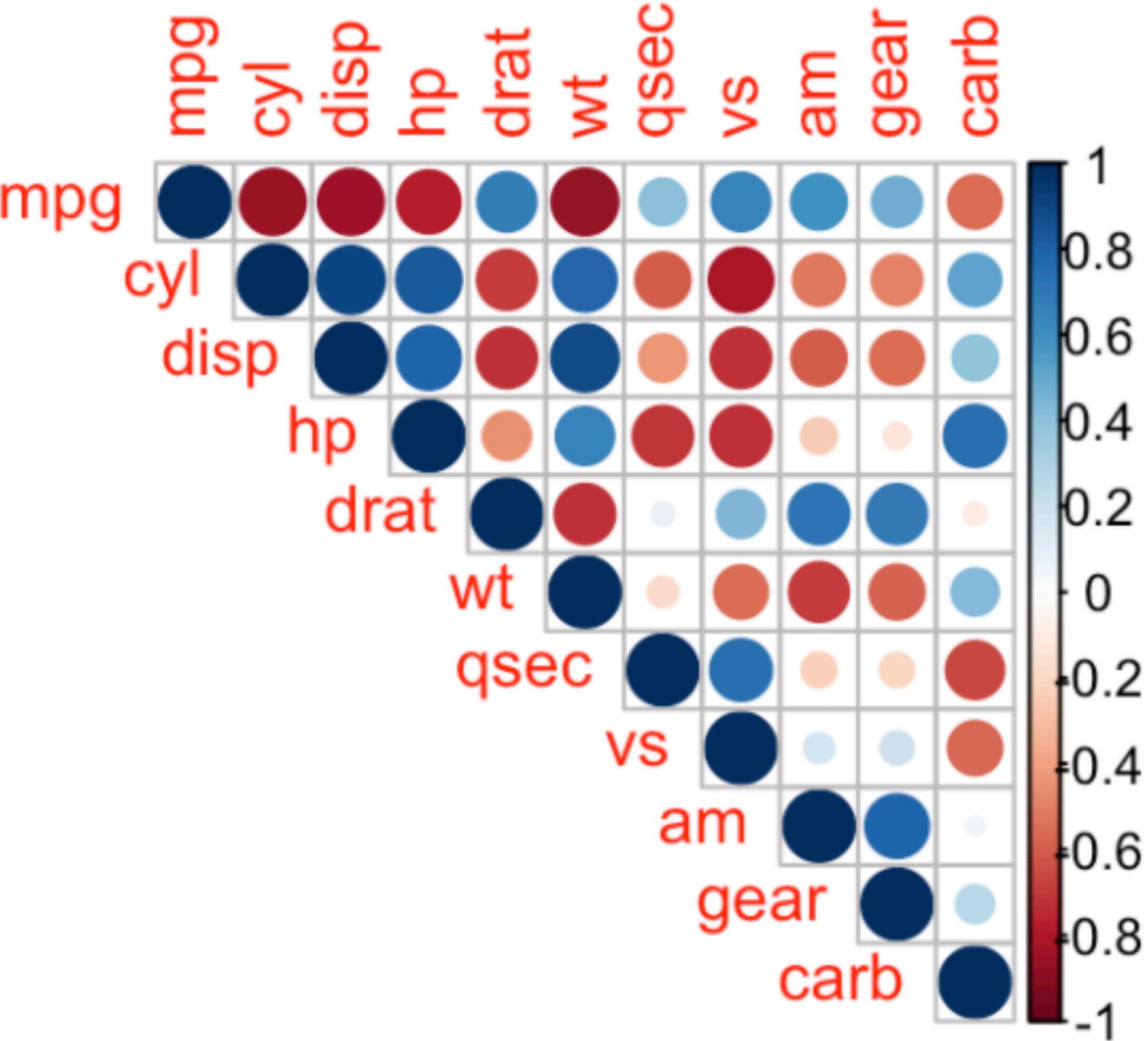
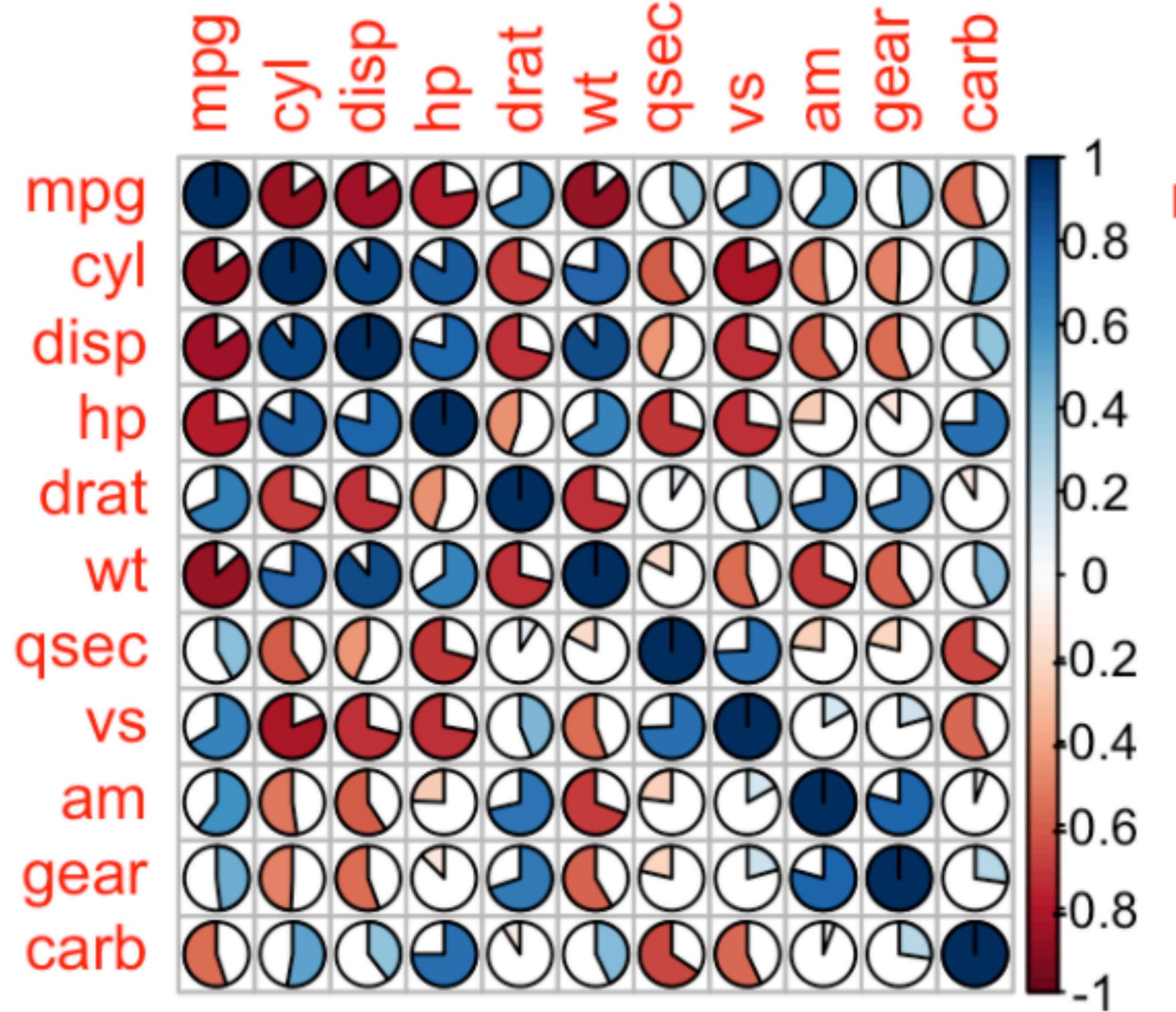


MATRIZES DE CORRELAÇÃO

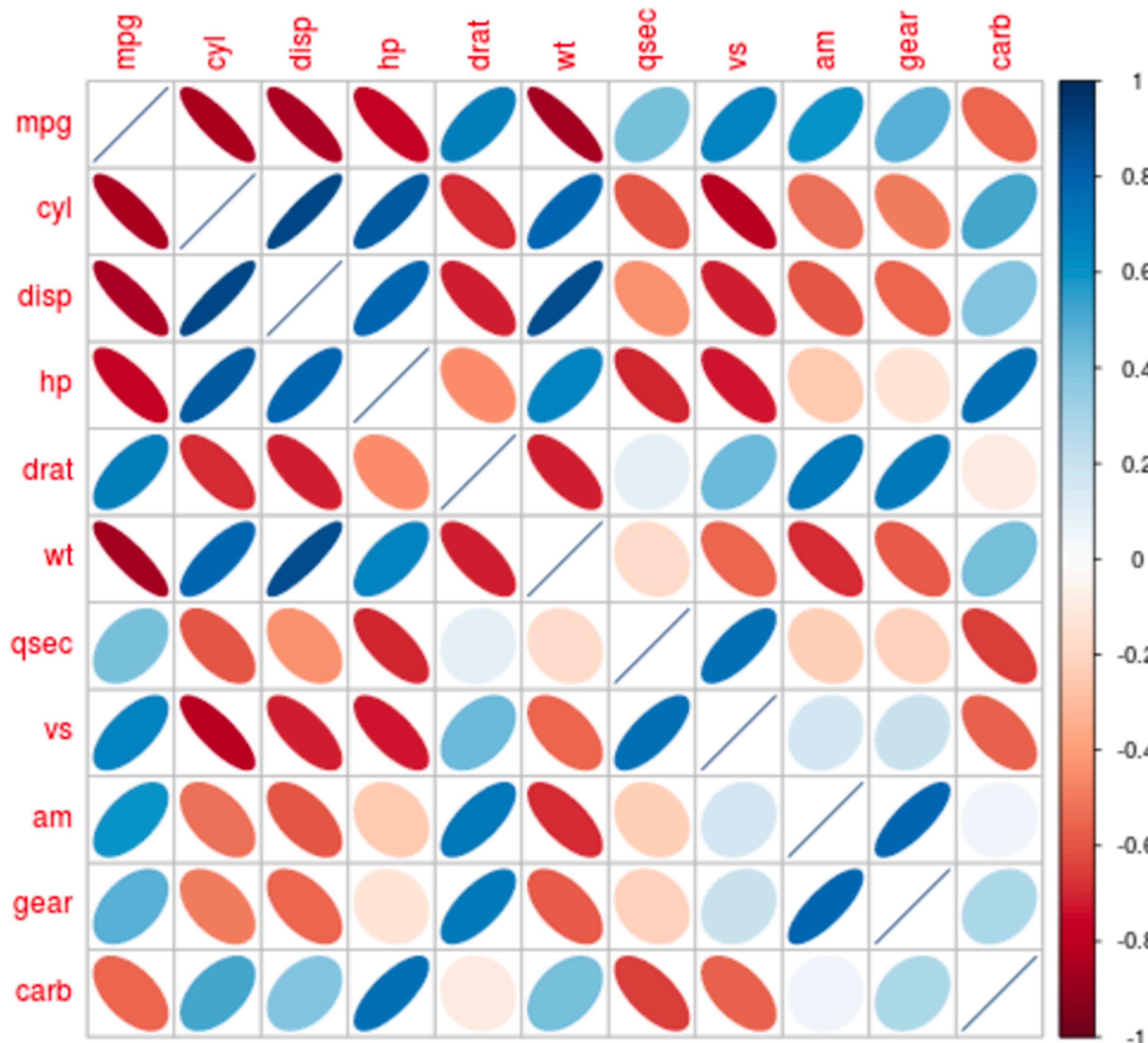
	mpg	cyl	disp	hp	drat	wt	qsec	vs	am	gear	carb
mpg	1	-0.85	-0.85	-0.78	0.68	-0.87	0.42	0.66	0.6	0.48	-0.55
cyl	-0.85	1	0.9	0.83	-0.7	0.78	-0.59	-0.81	-0.52	-0.49	0.53
disp	-0.85	0.9	1	0.79	-0.71	0.89	-0.43	-0.71	-0.59	-0.56	0.39
hp	-0.78	0.83	0.79	1	-0.45	0.66	-0.71	-0.72	-0.24	-0.13	0.75
drat	0.68	-0.7	-0.71	-0.45	1	-0.71	0.09	0.44	0.71	0.7	-0.09
wt	-0.87	0.78	0.89	0.66	-0.71	1	-0.17	-0.55	-0.69	-0.58	0.43
qsec	0.42	-0.59	-0.43	-0.71	0.09	-0.17	1	0.74	-0.23	-0.21	-0.66
vs	0.66	-0.81	-0.71	-0.72	0.44	-0.55	0.74	1	0.17	0.21	-0.57
am	0.6	-0.52	-0.59	-0.24	0.71	-0.69	-0.23	0.17	1	0.79	0.06
gear	0.48	-0.49	-0.56	-0.13	0.7	-0.58	-0.21	0.21	0.79	1	0.27
carb	-0.55	0.53	0.39	0.75	-0.09	0.43	-0.66	-0.57	0.06	0.27	1



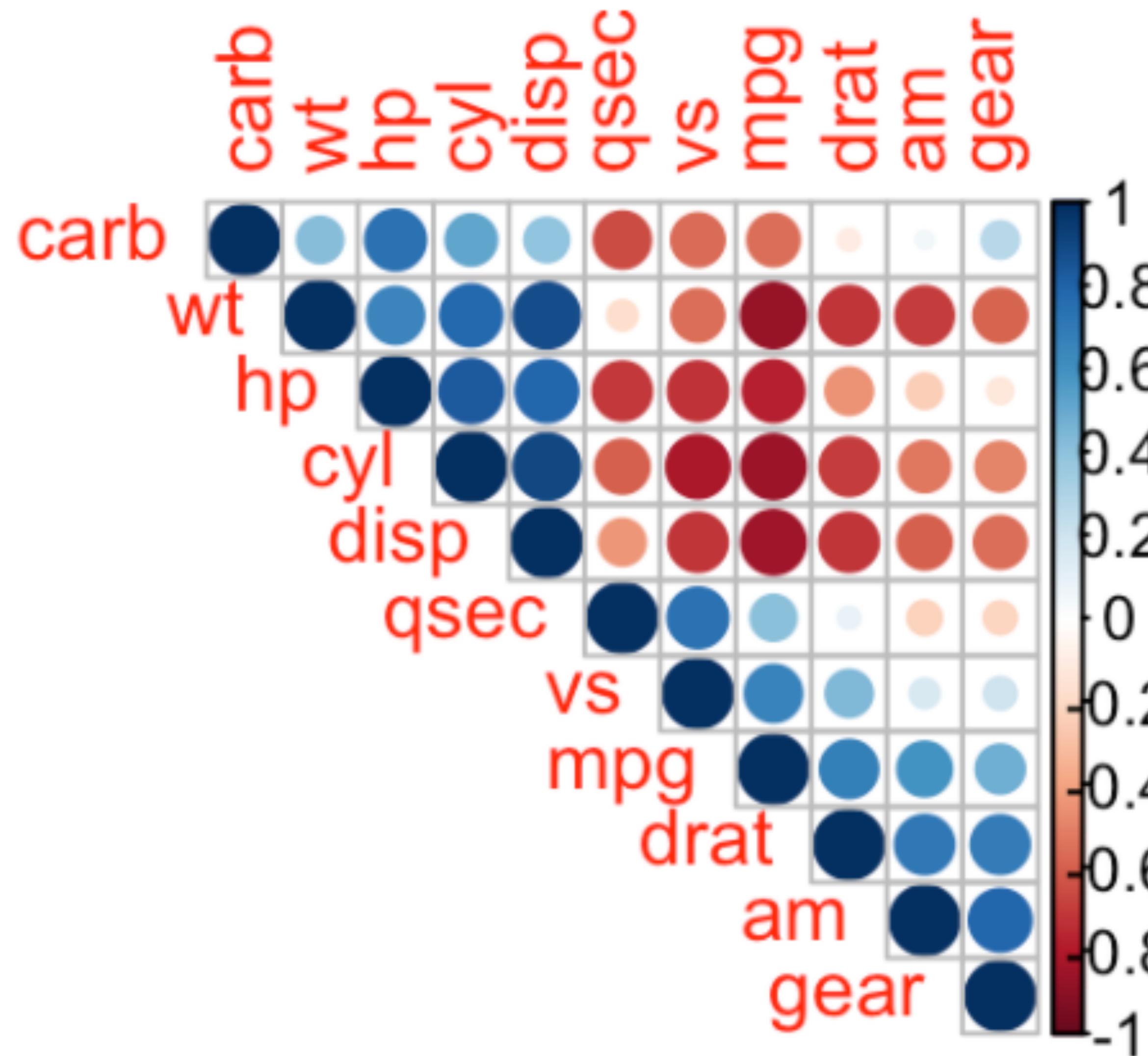
MATRIZES DE CORRELAÇÃO - VARIAÇÕES



MATRIZES DE CORRELAÇÃO - VARIACÕES

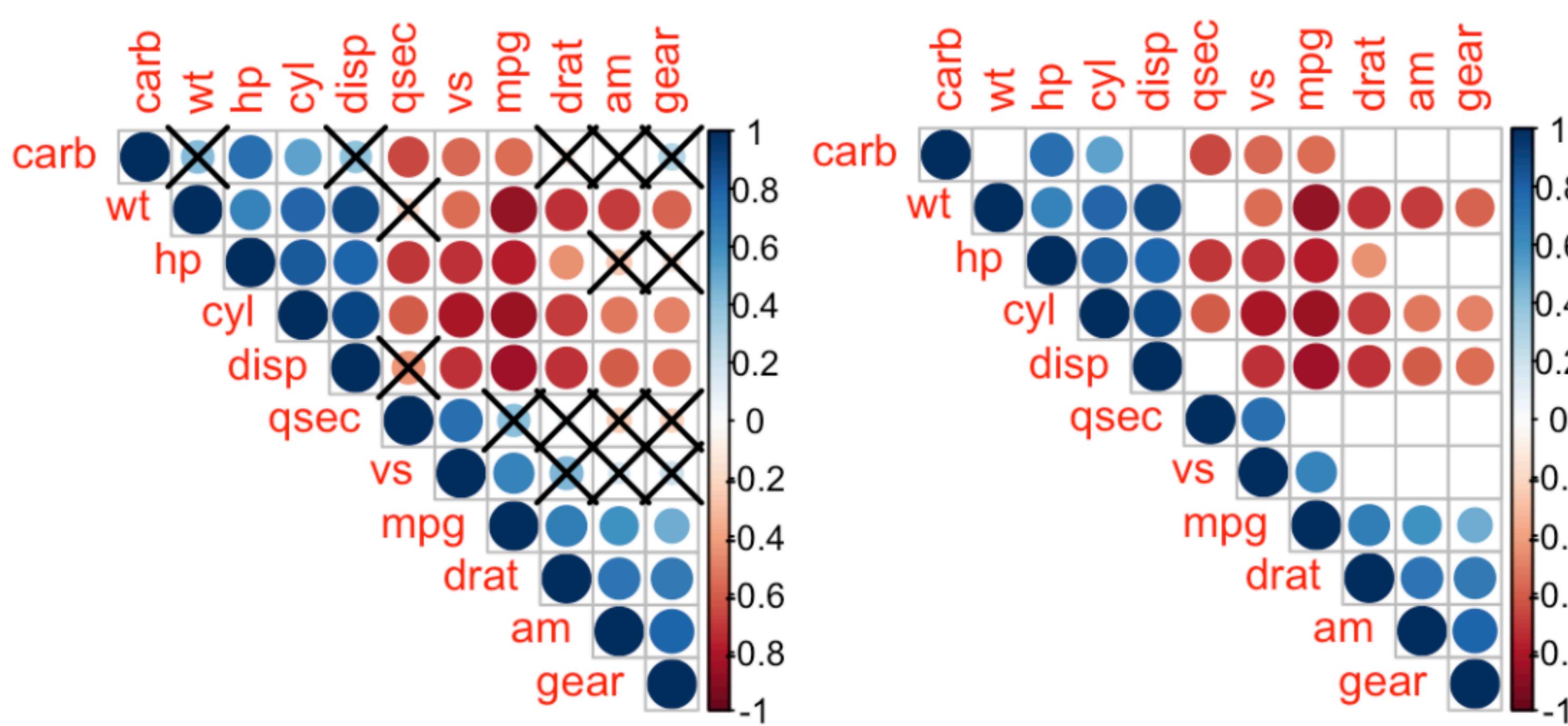


MATRIZES DE CORRELAÇÃO - AGRUPAMENTO DE COLUNAS

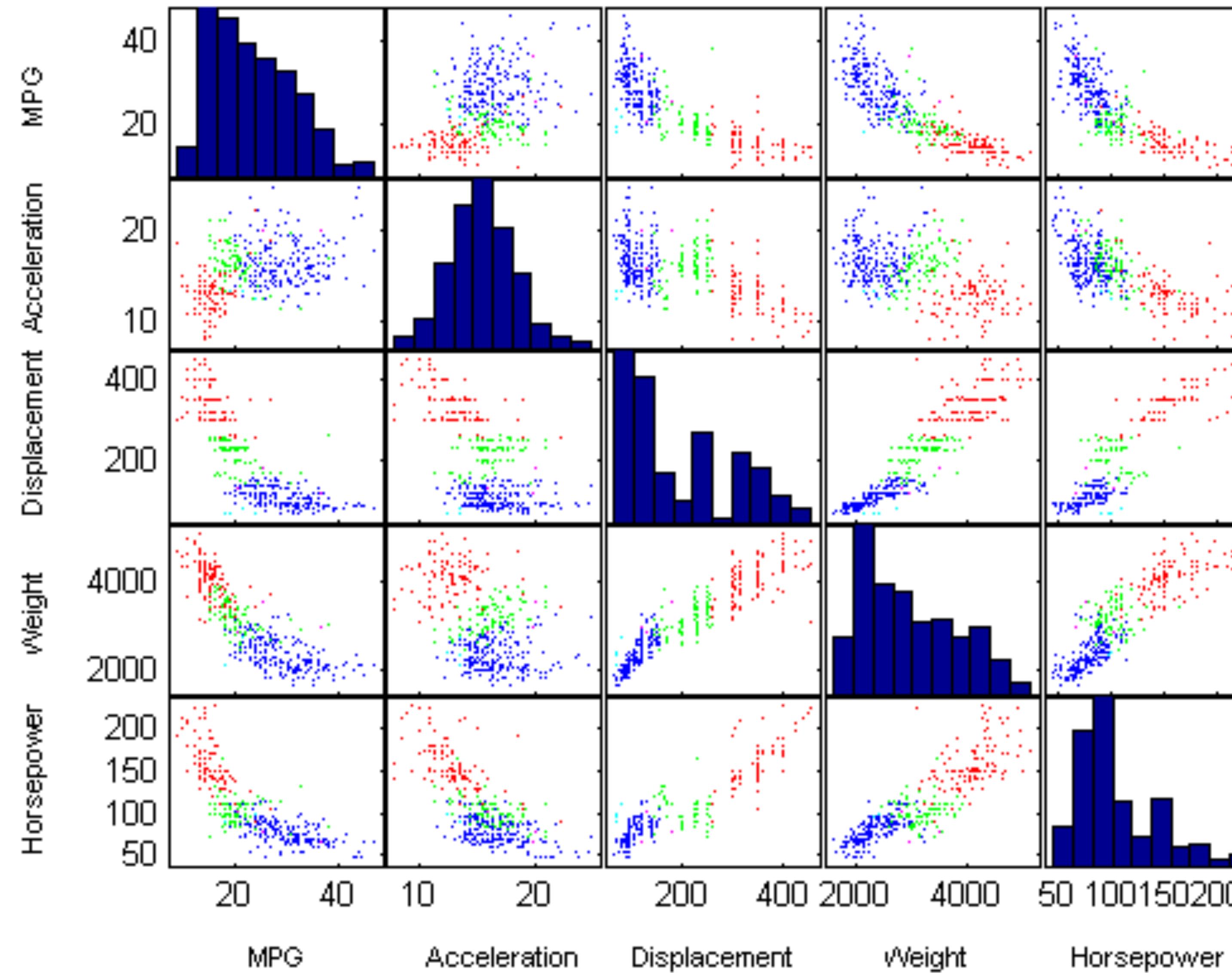


MATRIZES DE CORRELAÇÃO - NÍVEL DE SIGNIFICÂNCIA

	mpg	cyl	disp	hp	drat
mpg	0.000e+00	6.113e-10	9.380e-10	1.788e-07	1.776e-05
cyl	6.113e-10	0.000e+00	1.803e-12	3.478e-09	8.245e-06
disp	9.380e-10	1.803e-12	0.000e+00	7.143e-08	5.282e-06
hp	1.788e-07	3.478e-09	7.143e-08	0.000e+00	9.989e-03
drat	1.776e-05	8.245e-06	5.282e-06	9.989e-03	0.000e+00
wt	1.294e-10	1.218e-07	1.222e-11	4.146e-05	4.784e-06



MATRIZES DE GRÁFICOS DE DISPERSÃO



FilmStrips

Movies, money and metadata, visualised.

Choose which data to view with the panel on the right.

Use the mouse-wheel to zoom in and out; click-and-drag to scroll.

Move your cursor over a strip to see more information.

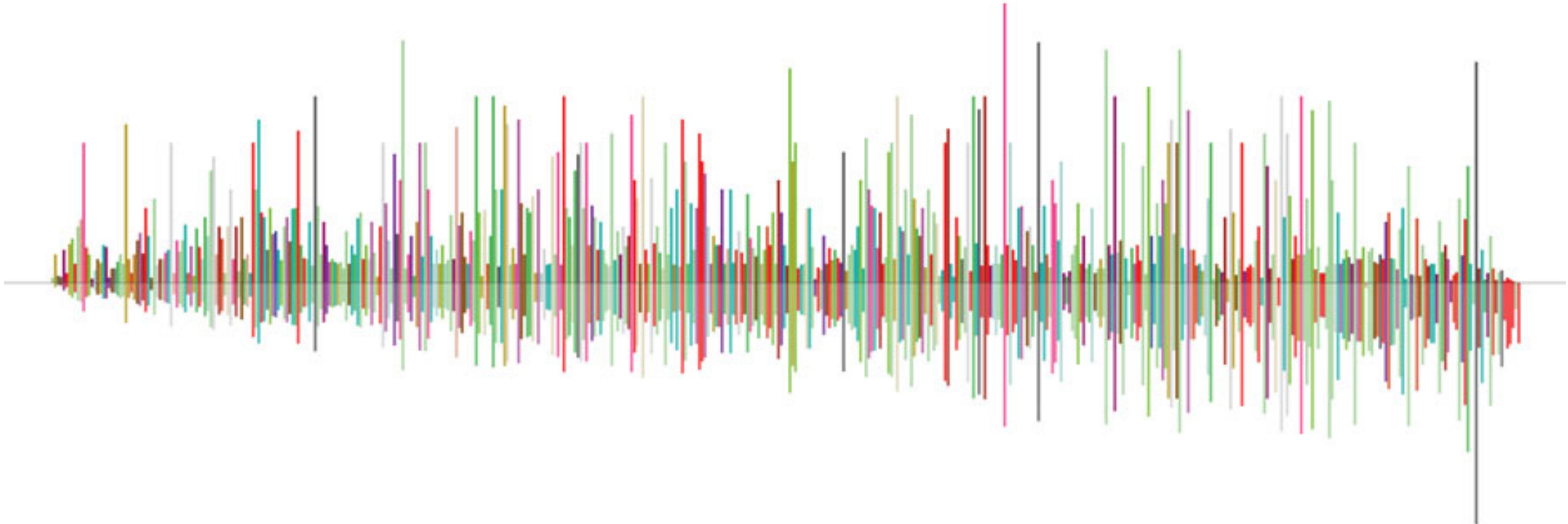
Data by Information is Beautiful Visualisation by Tom Evans

What data would you like to see?

Above the line: Production Budget ▾

Below the line: Worldwide Gross ▾

Sort by: Profitability ▾



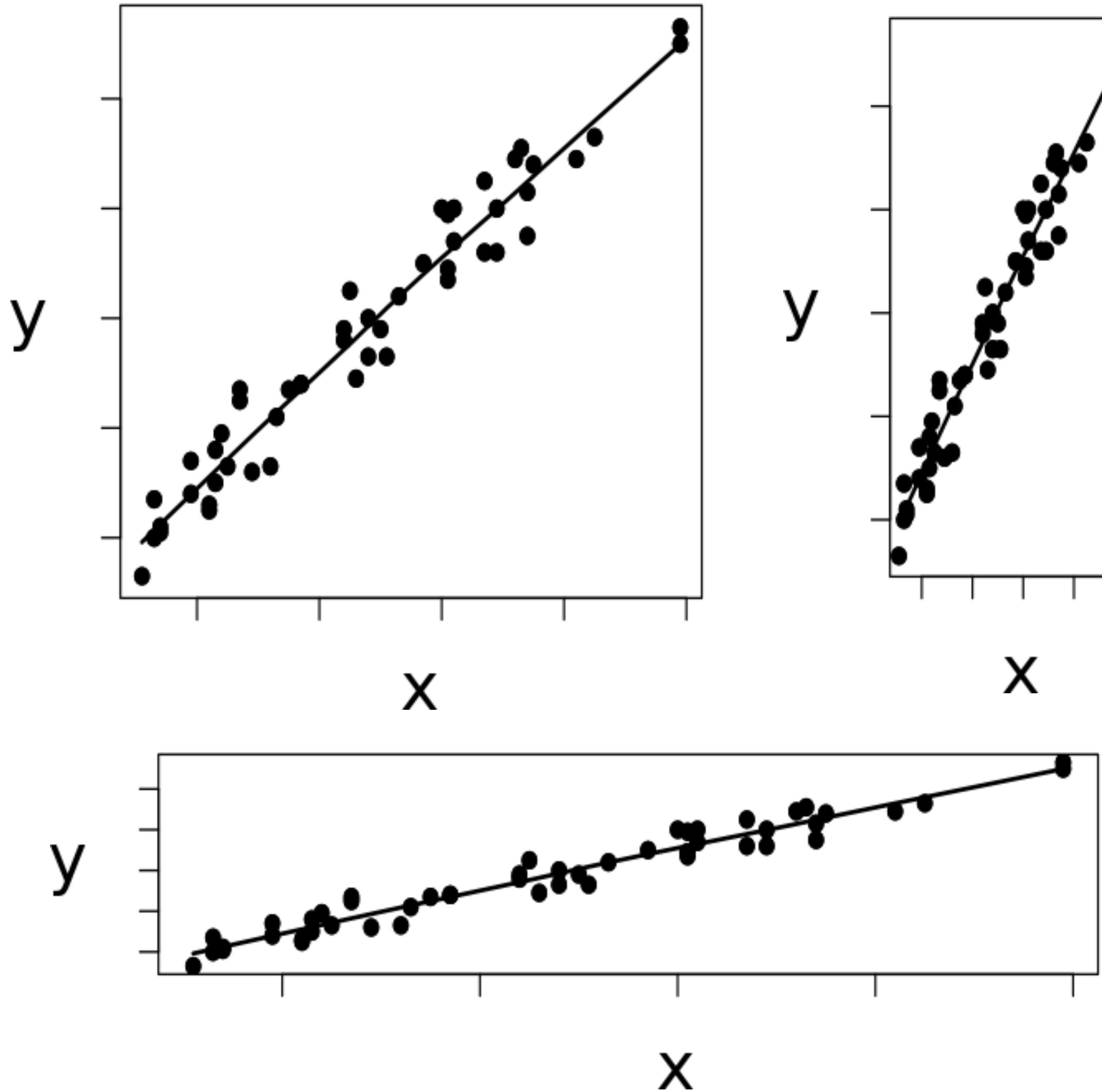
Story Type key

BOAS PRÁTICAS

.....

EQUILIBRE A RAZÃO DE ASPECTO

- Em gráficos de pontos para análise de correlações, use sempre a razão do aspecto 1

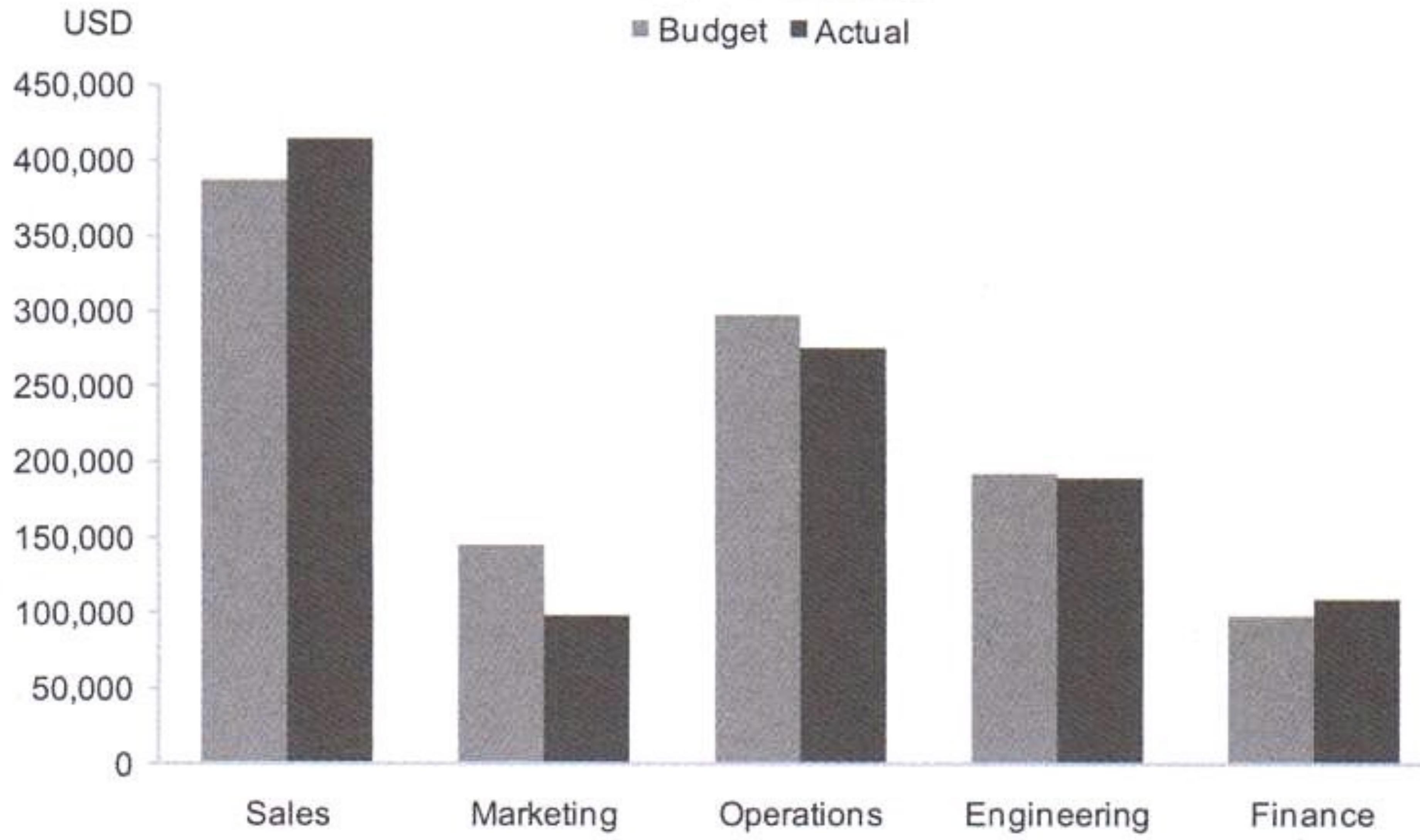


Desvio

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

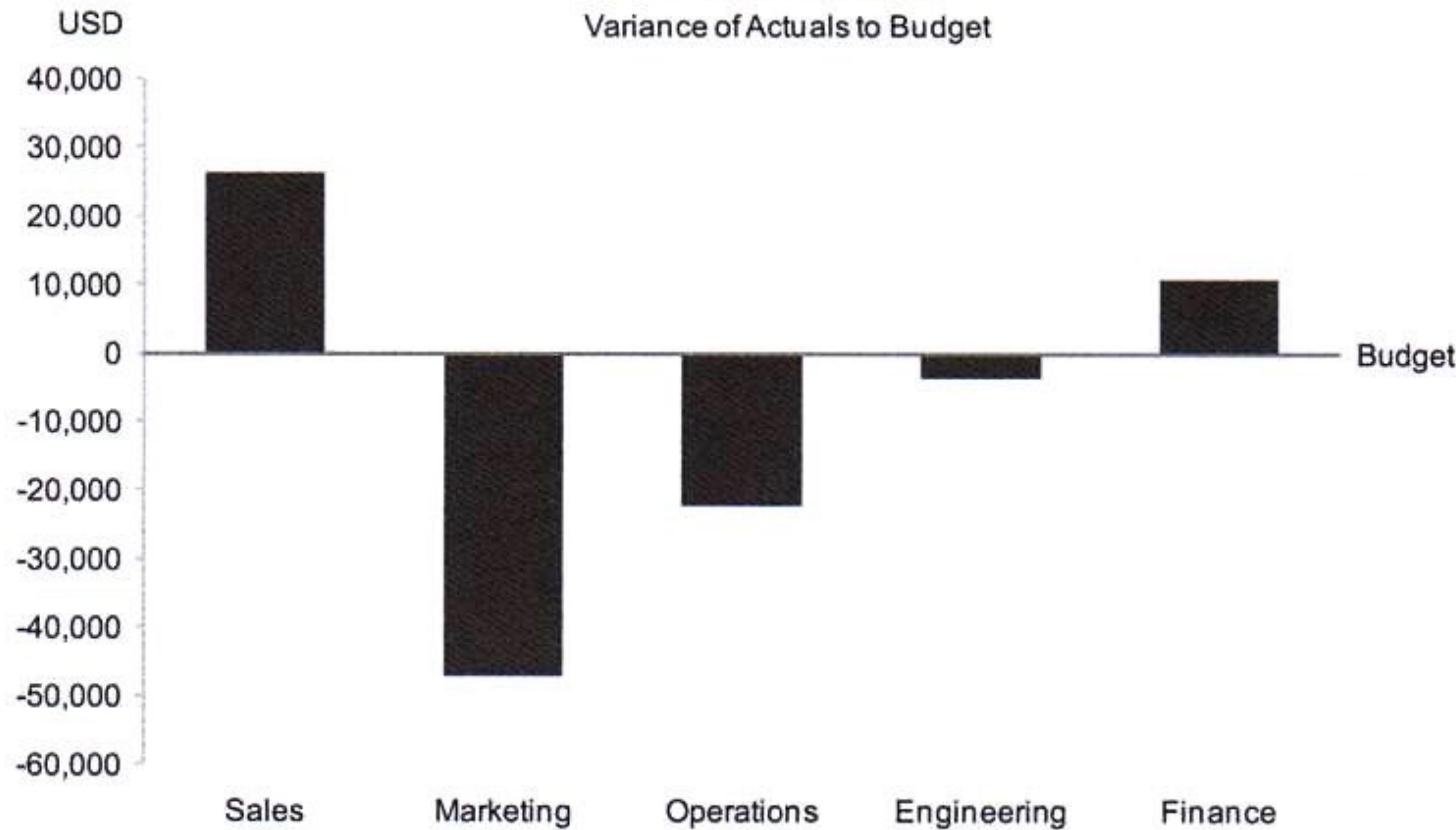
- A análise de como um ou mais valores **desviam** de um **conjunto de referência** é a análise de desvio

2008 YTD Expenses



Exige que realizemos cálculos

2008 YTD Expenses
Variance of Actuals to Budget

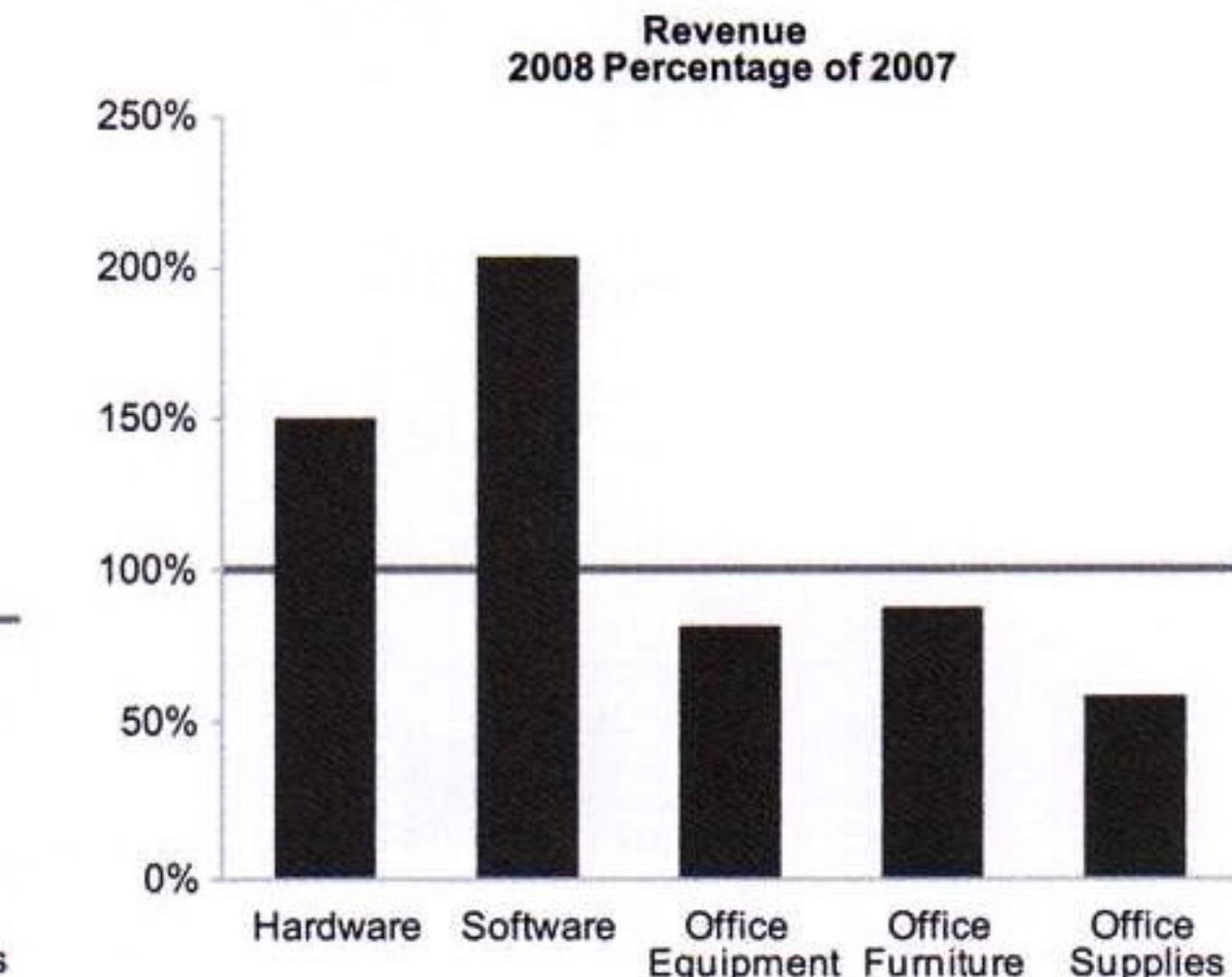
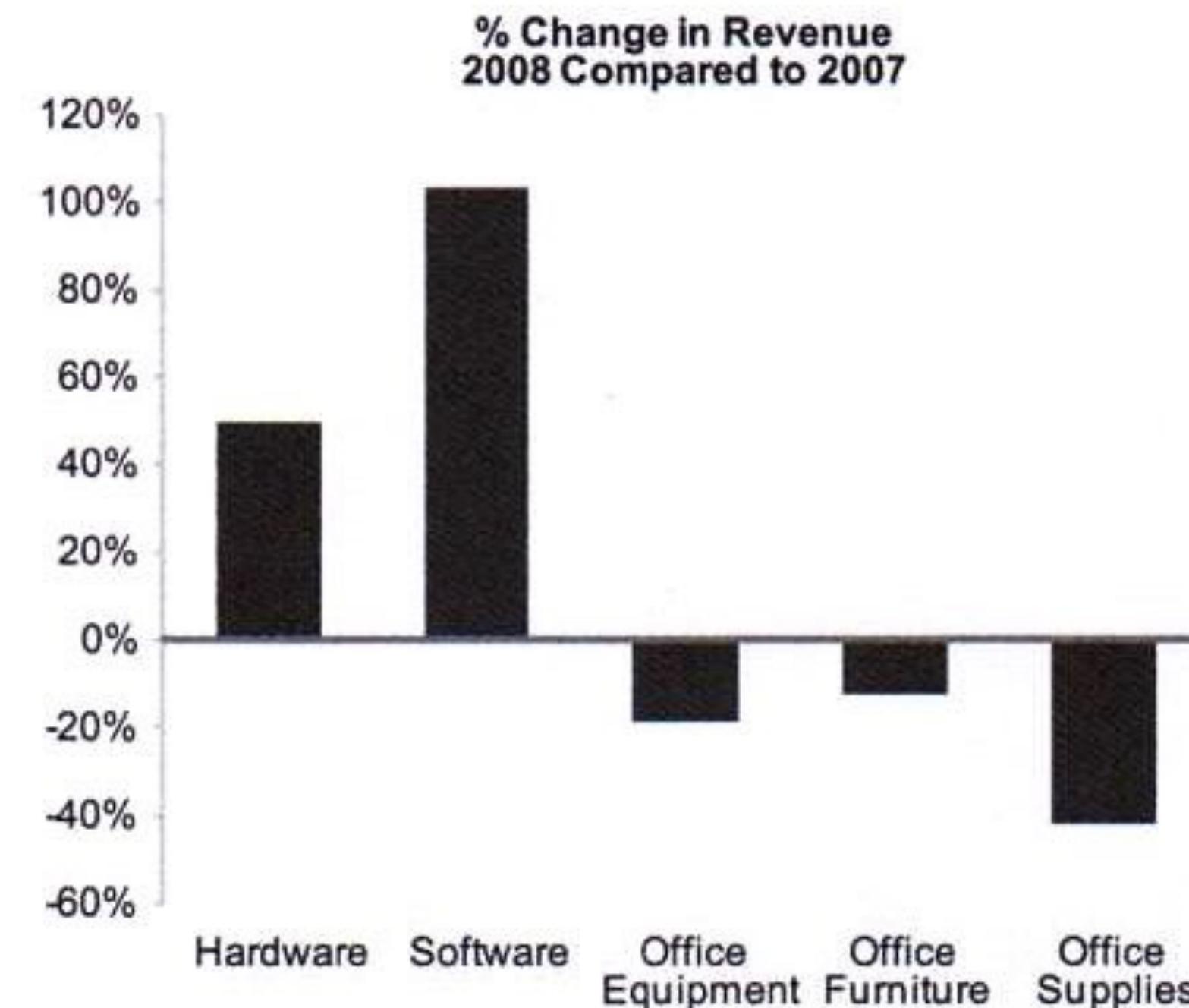
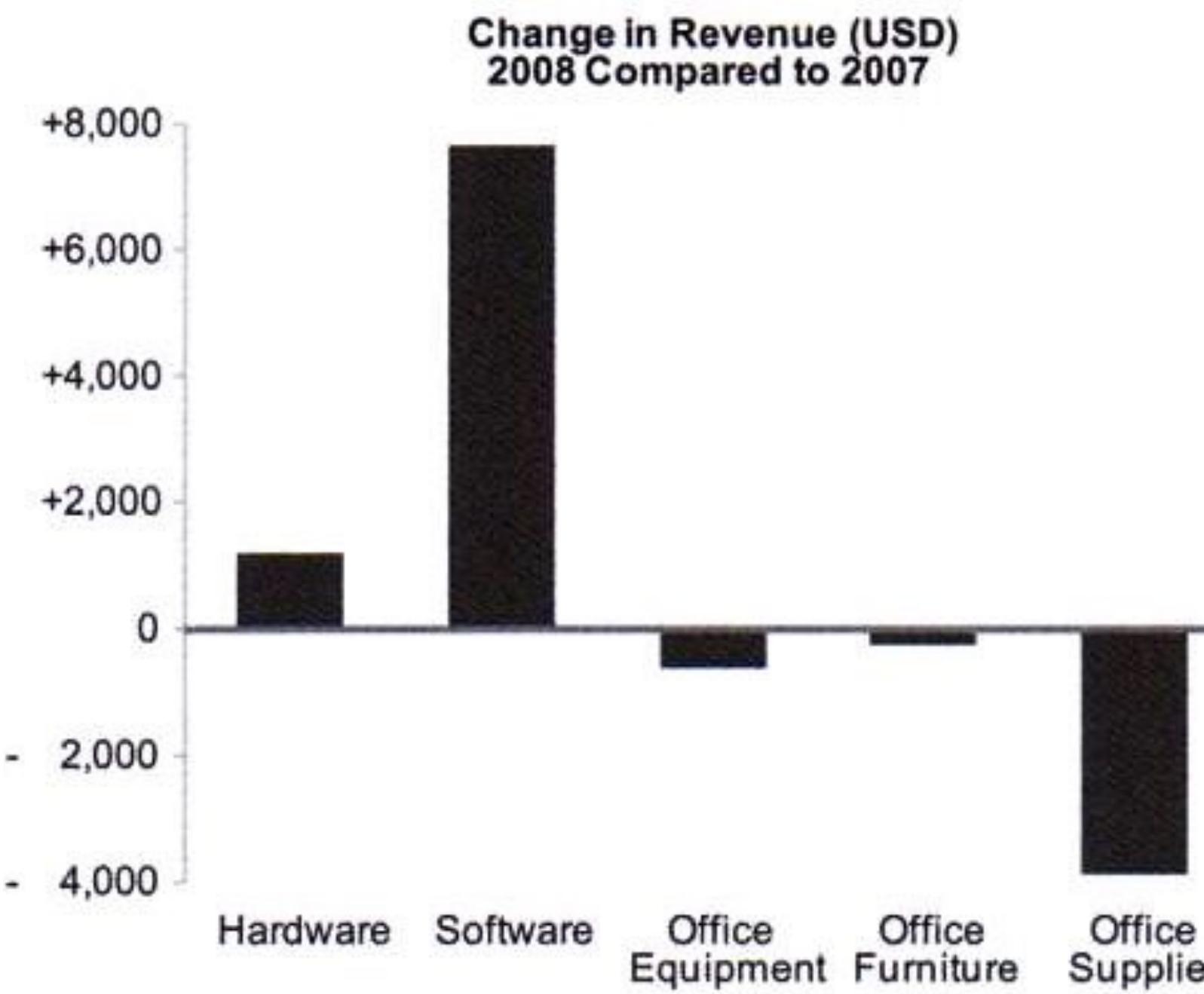


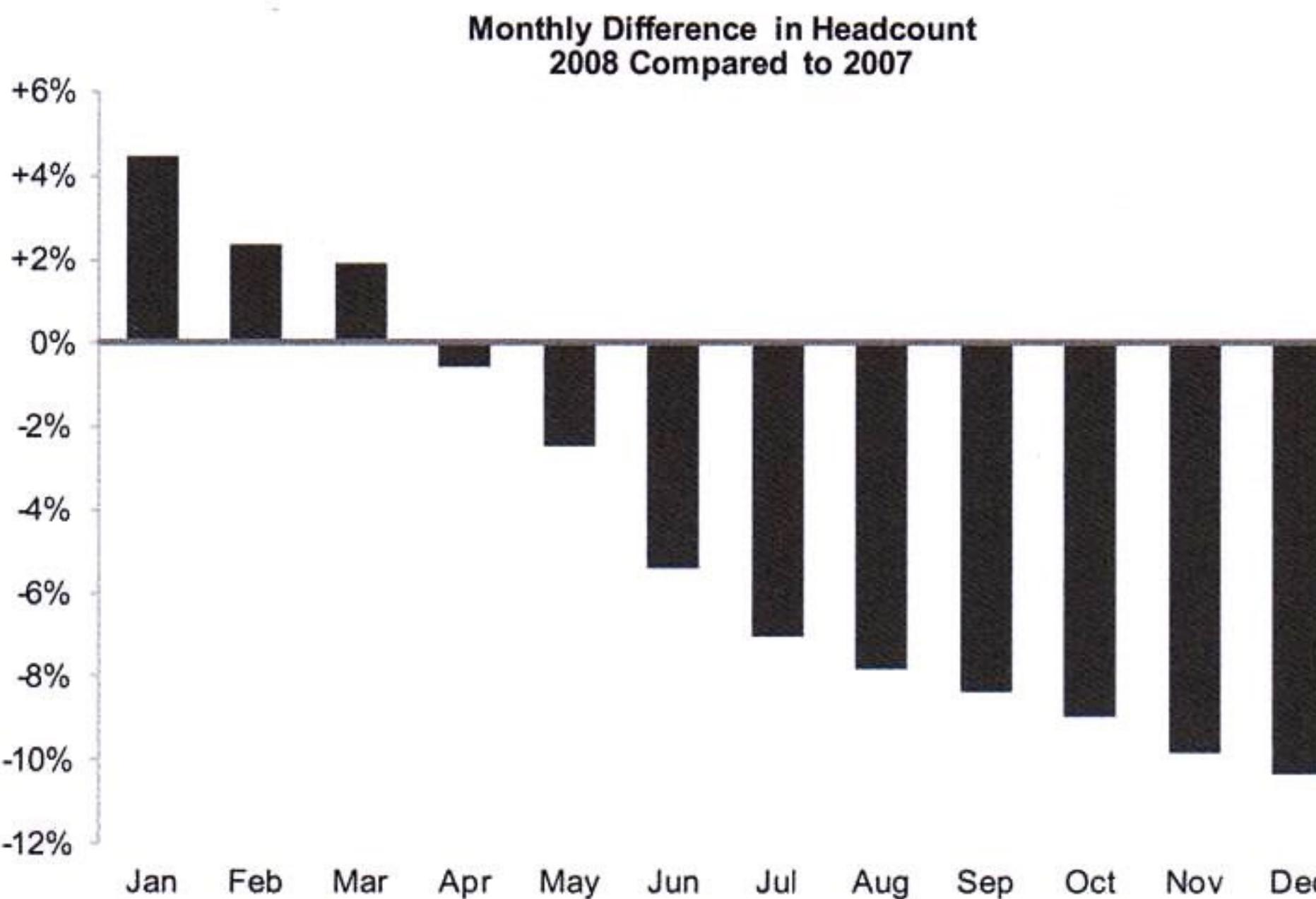
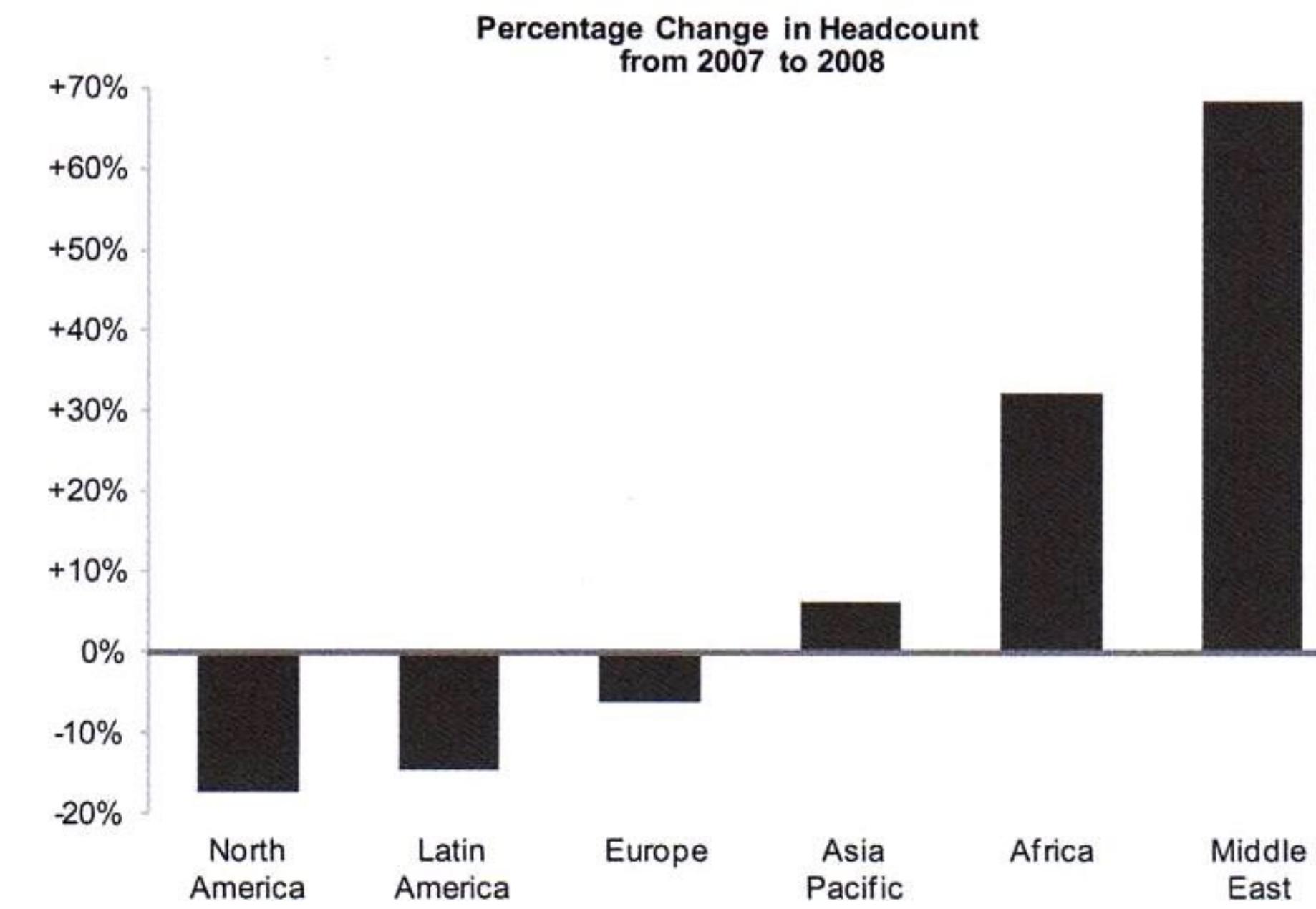
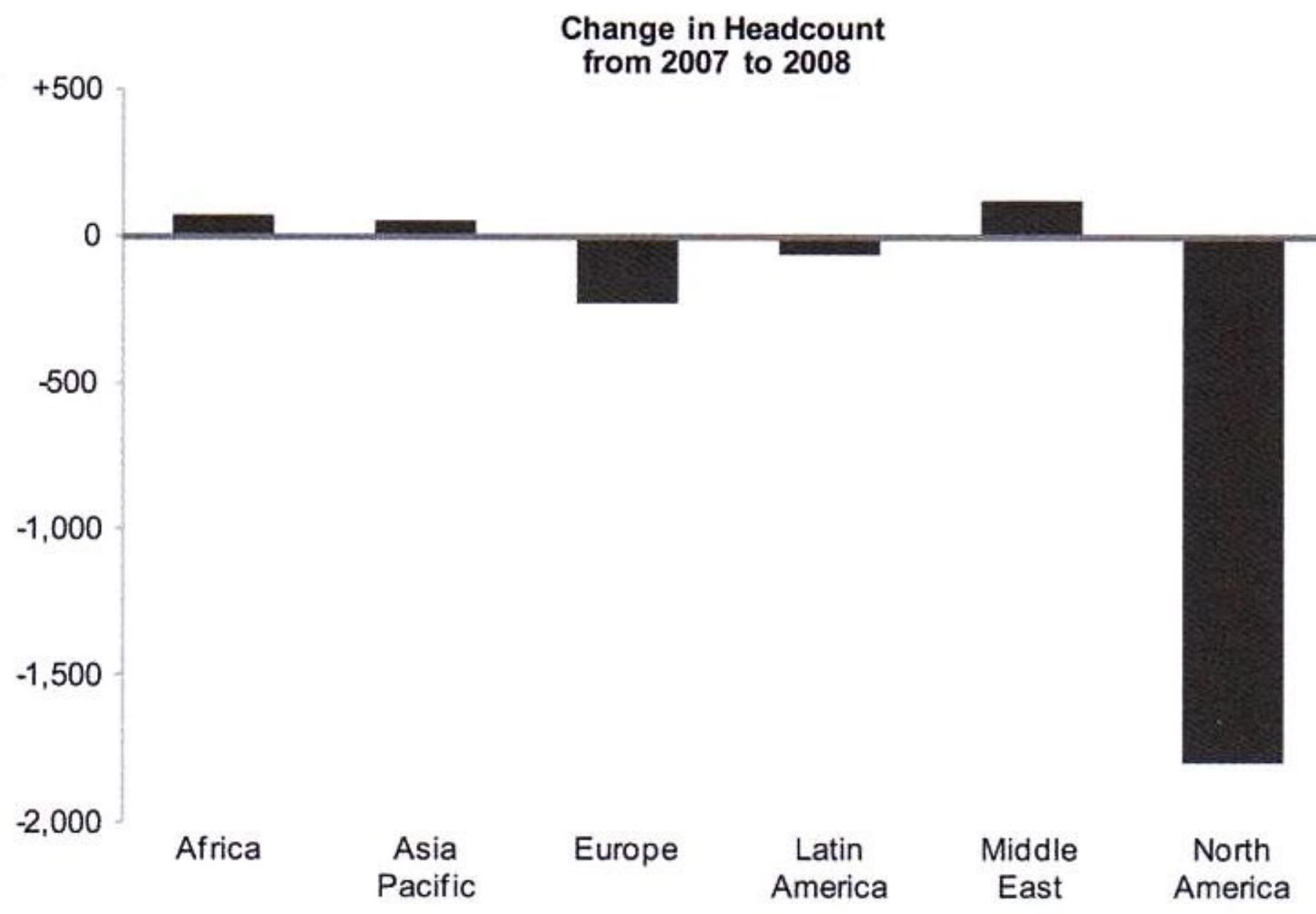
Diferenças explícitas

- Alguns exemplos de análises de desvio envolvem comparações com:
 - Alvo atual ou futuro
 - Mesmo valor no passado
 - Período imediatamente anterior
 - Norma ou padrão
 - Outros itens da mesma categoria

REPRESENTAÇÕES VISUAIS

As melhores representações para análise de desvios são os gráficos de barras e linhas

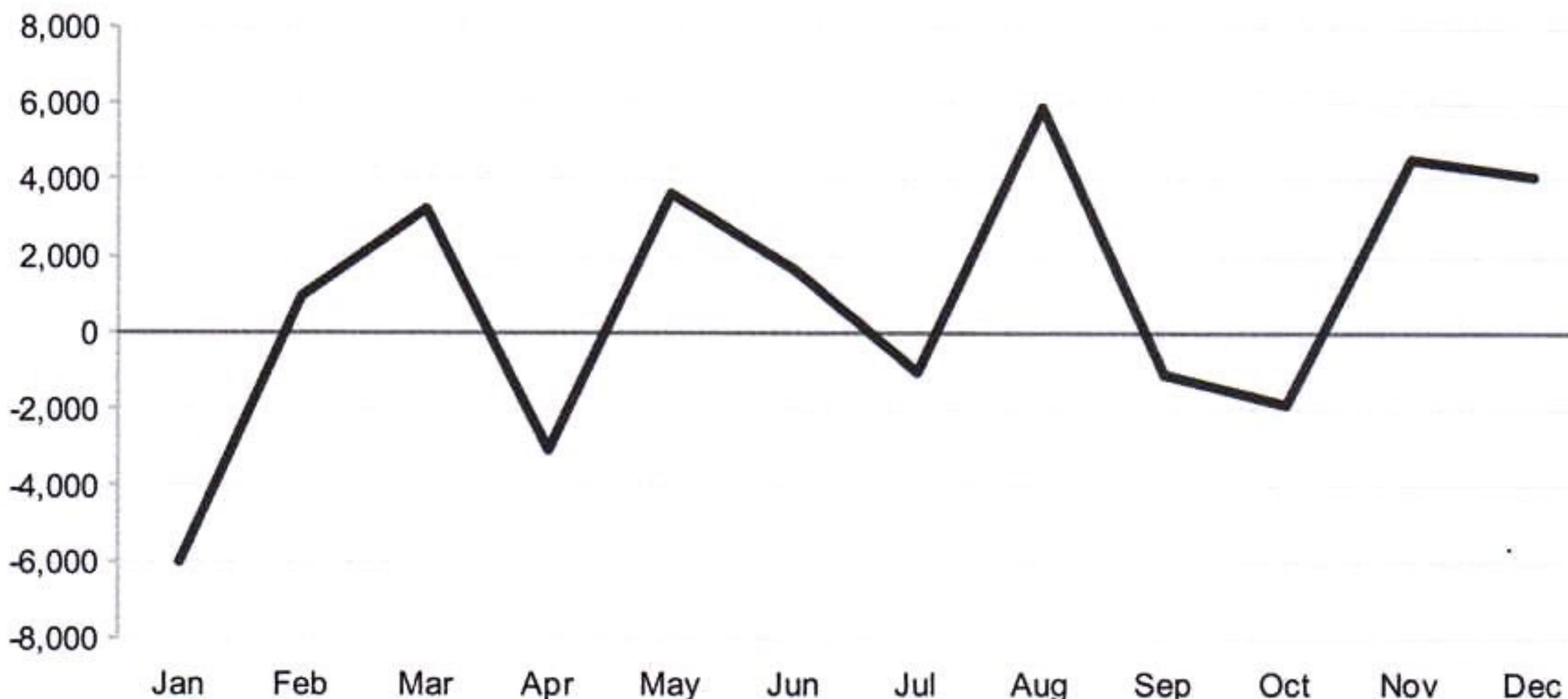
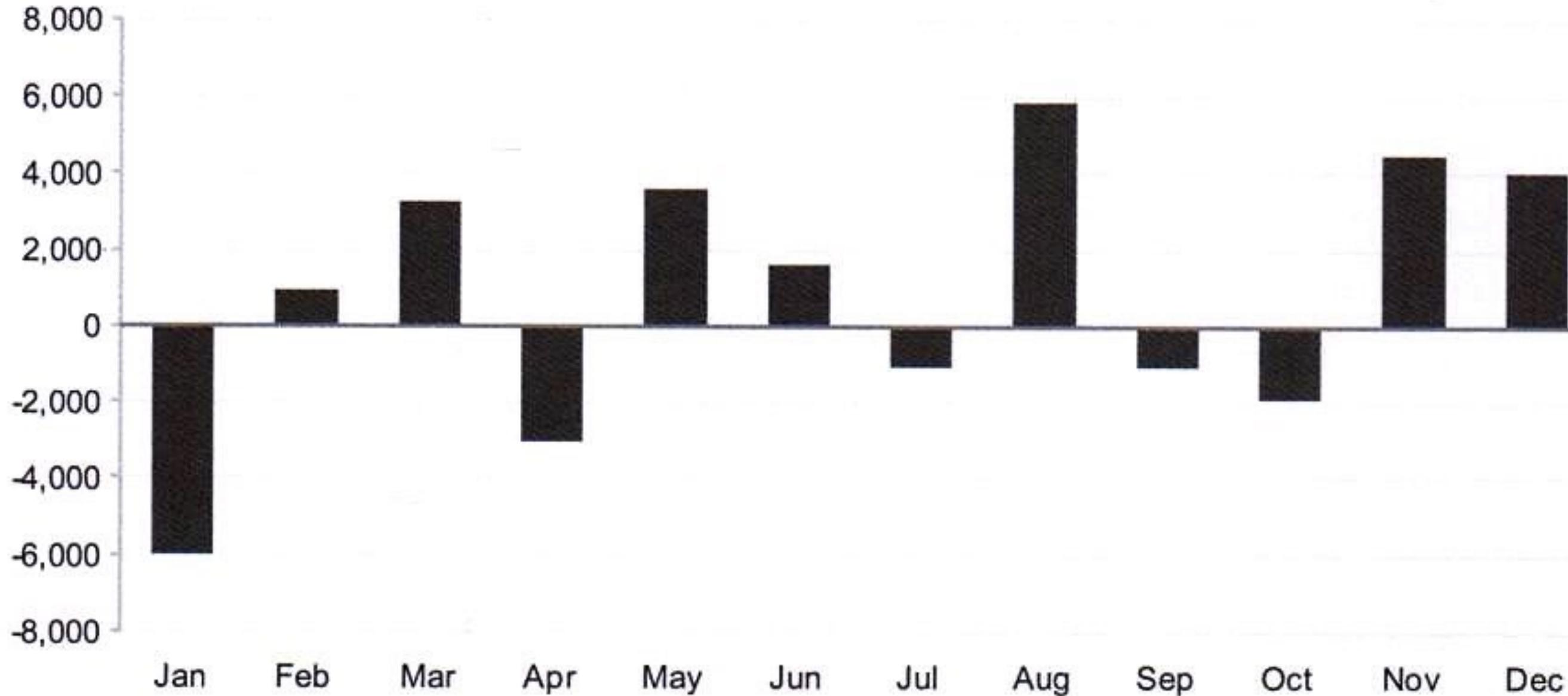




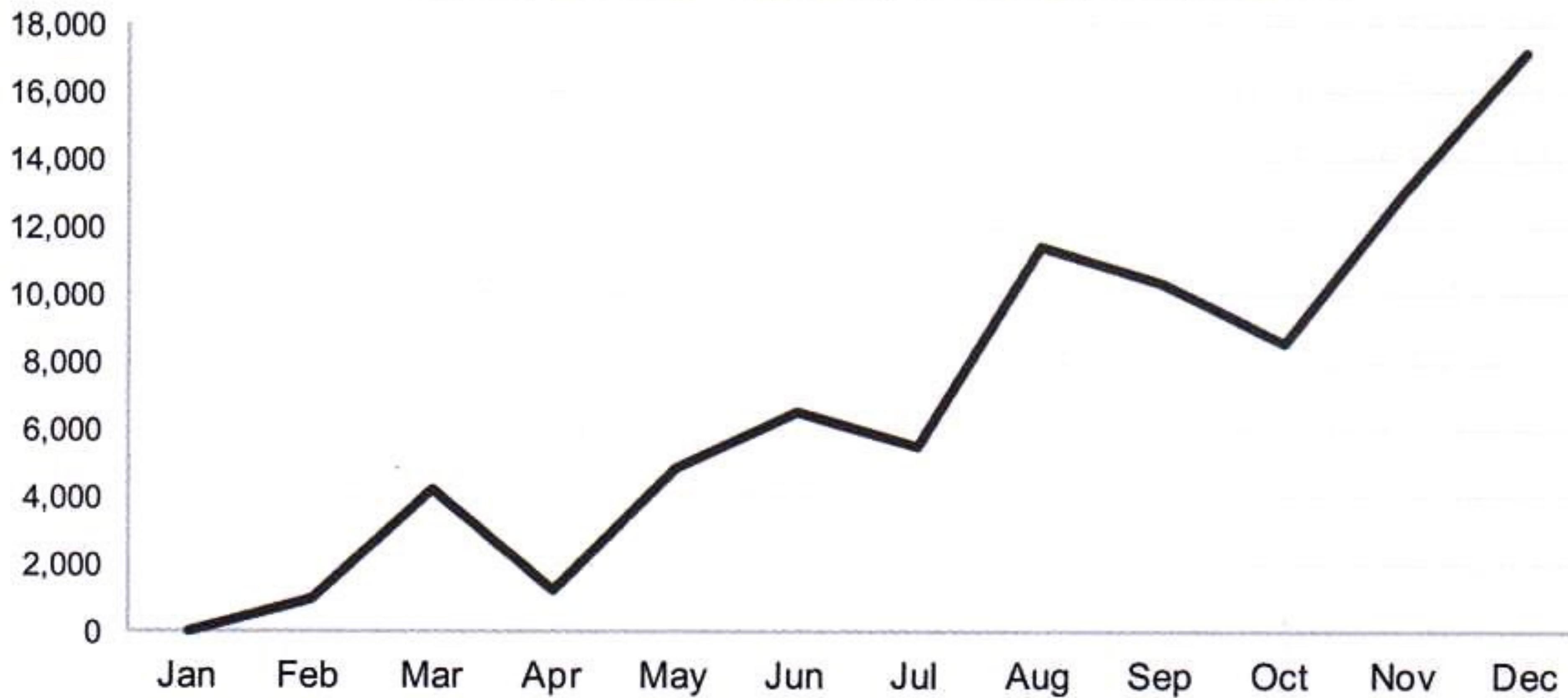
Barras são excelentes para
mostrar os desvios
individualmente

Mas não mostram claramente
as mudanças no tempo

Revenue Difference in U.S. Dollars Compared to Last Year's Average



Monthly Revenue Compared to January in U.S. Dollars

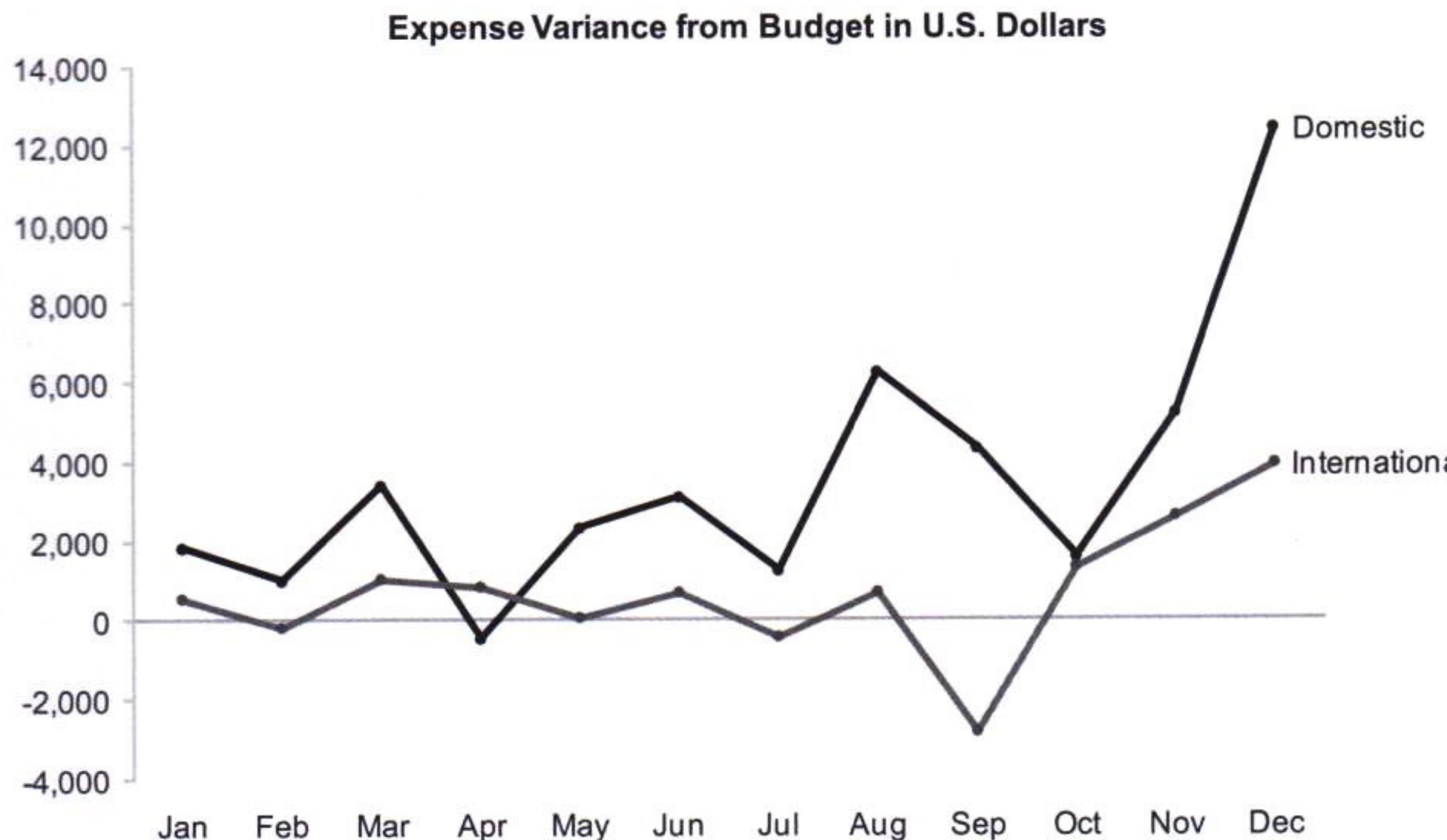


BOAS PRÁTICAS

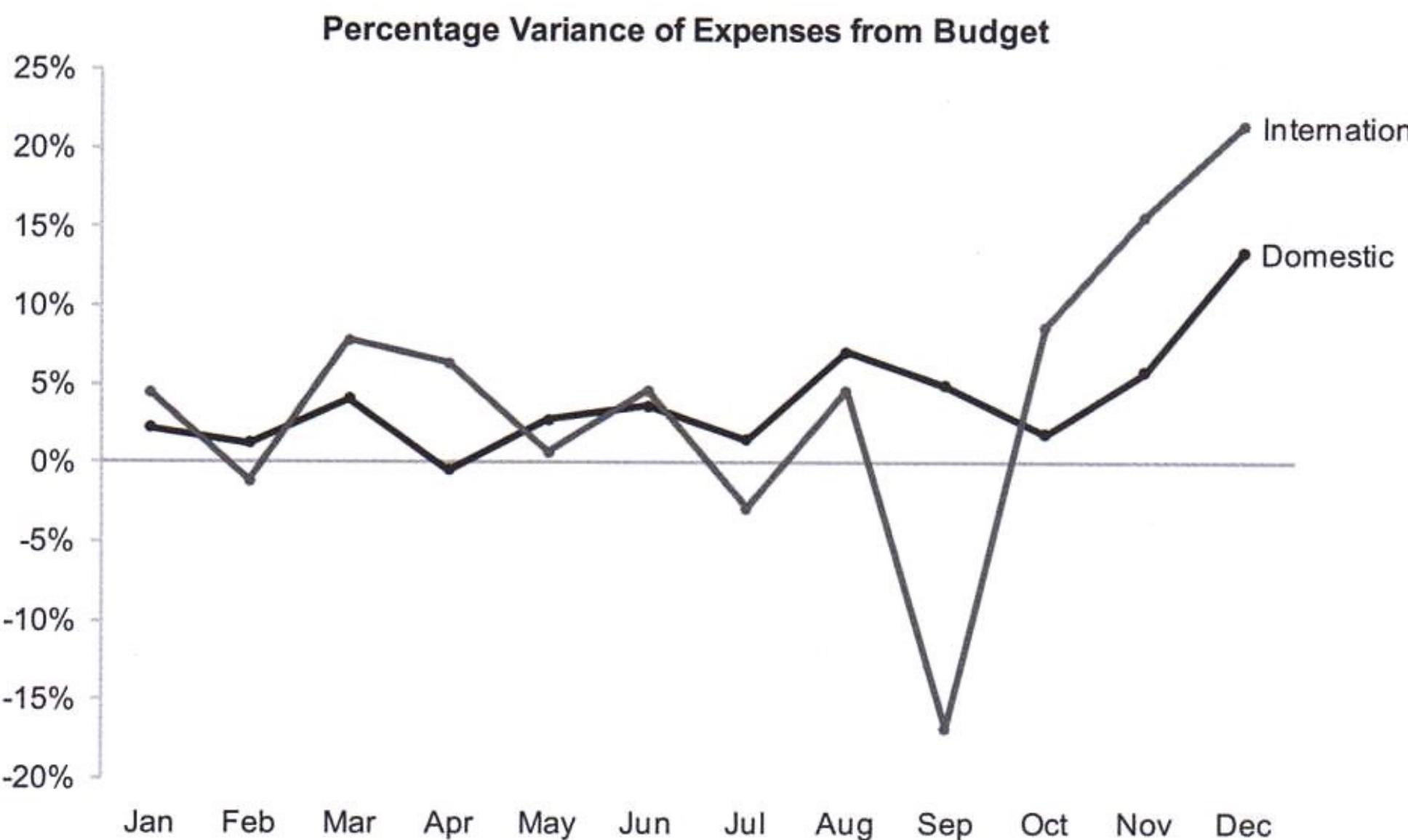
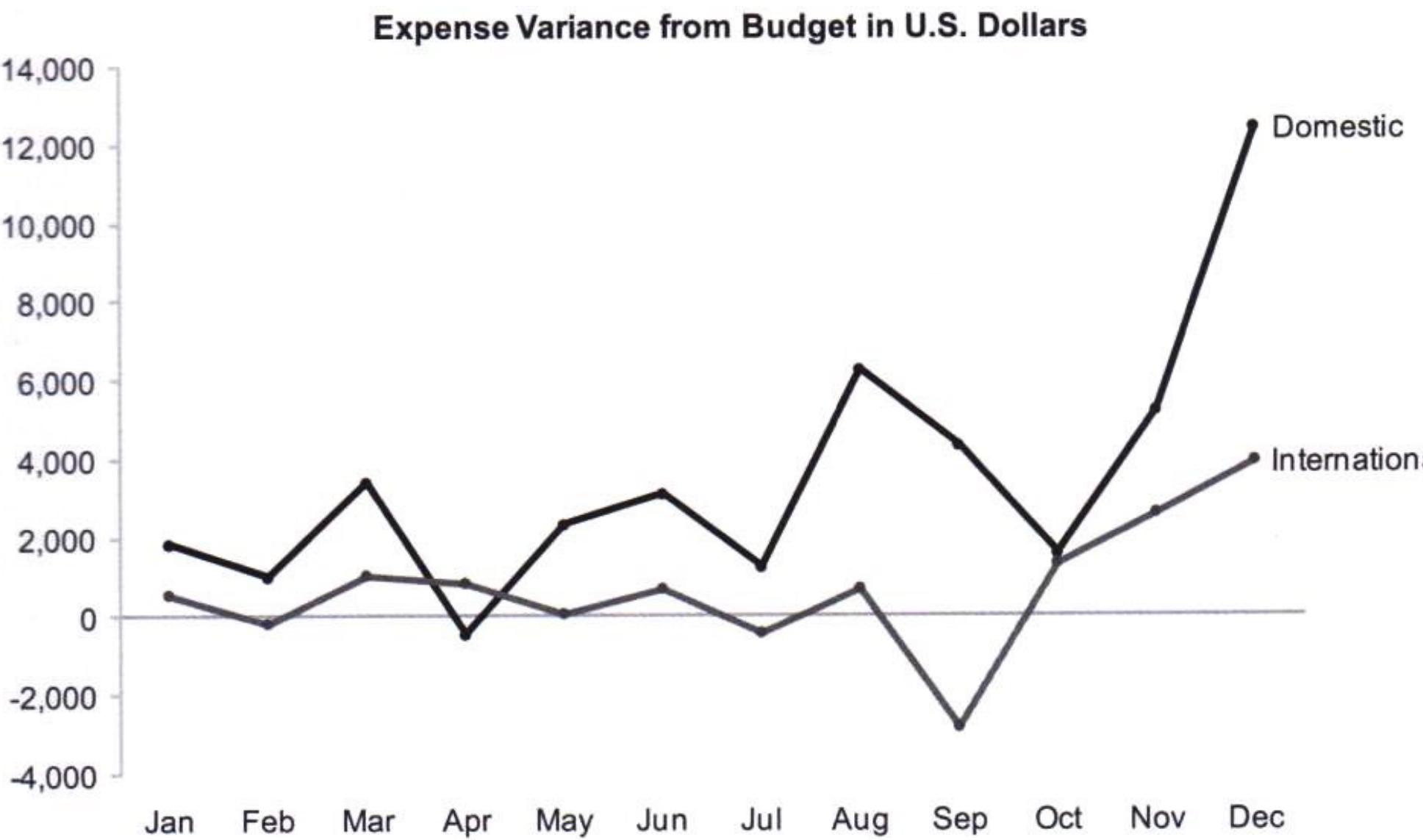
.....

EXPRESSE OS DESVIOS COMO PERCENTUAIS

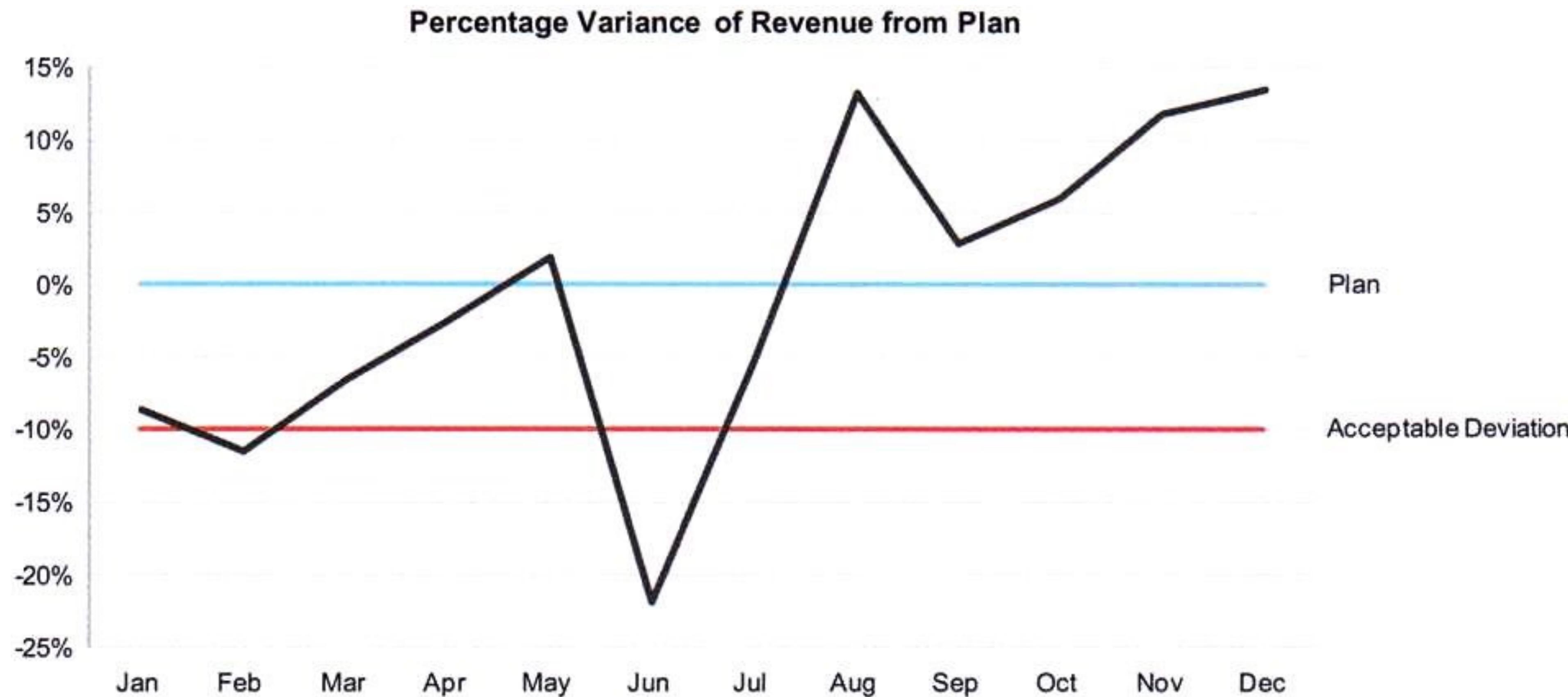
- Em alguns casos, expressar os desvios como percentuais traz uma visão completamente diferente dos dados



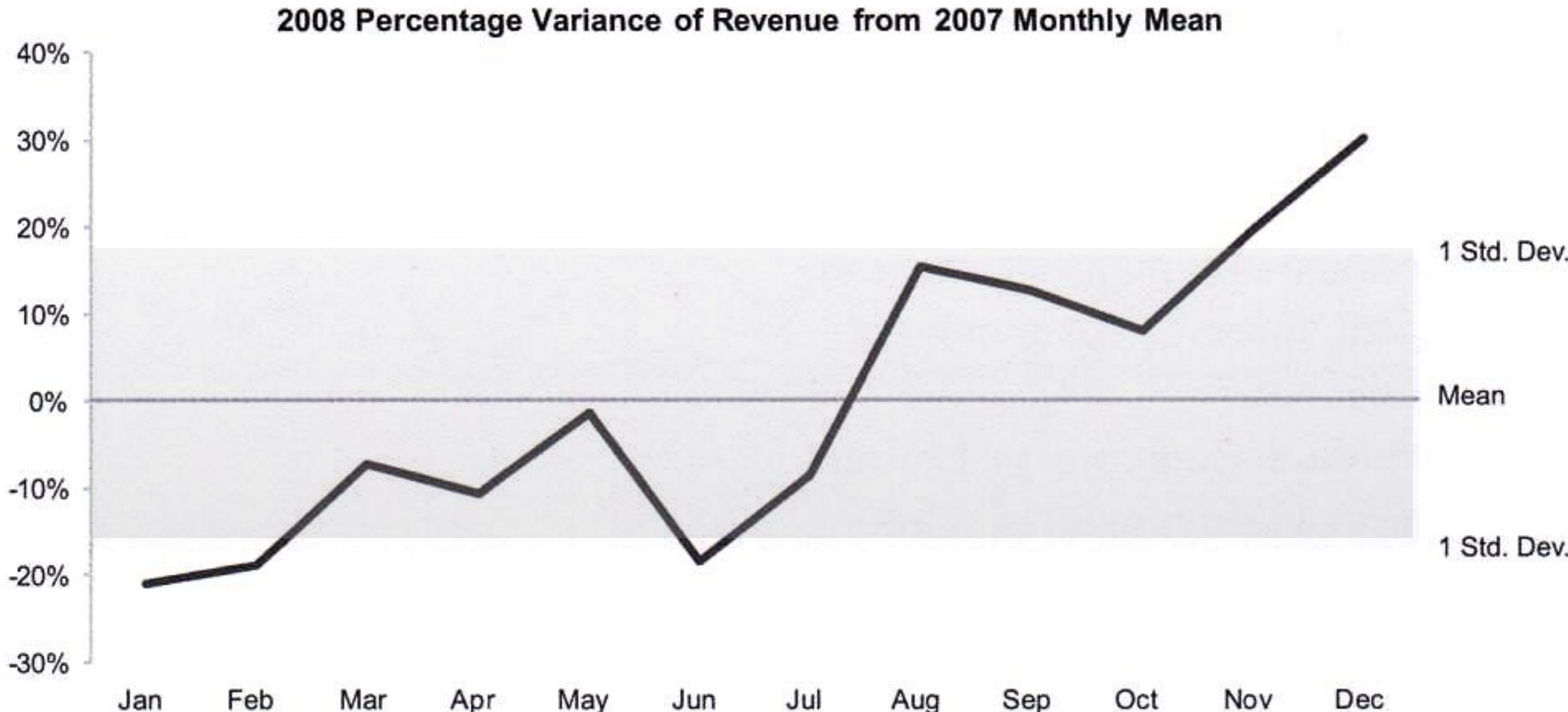
EXPRESSE OS DESVIOS COMO PERCENTUAIS



COMPARE OS DESVIOS COM VALORES DE REFERÊNCIA



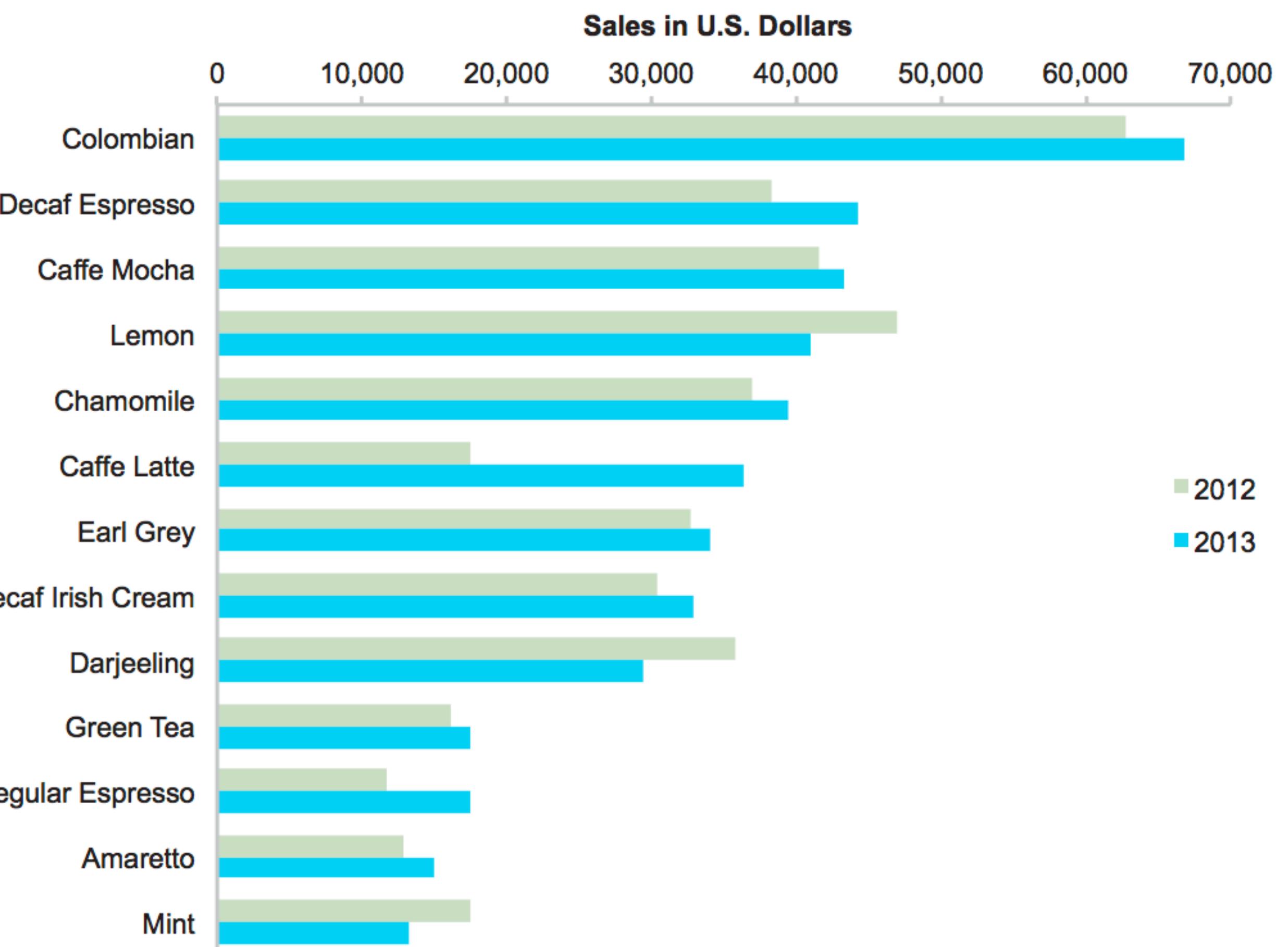
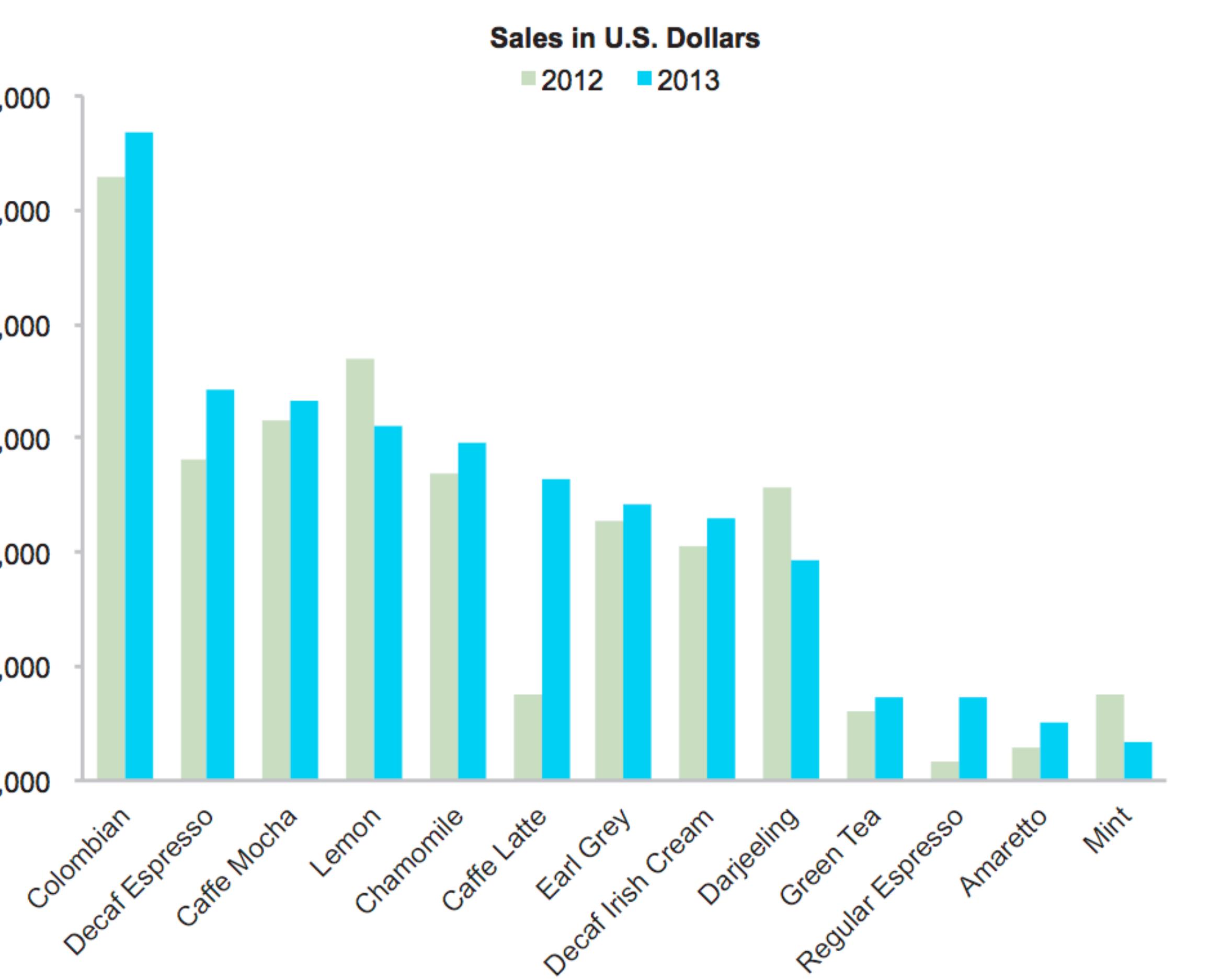
COMPARE OS DESVIOS COM VALORES DE REFERÊNCIA

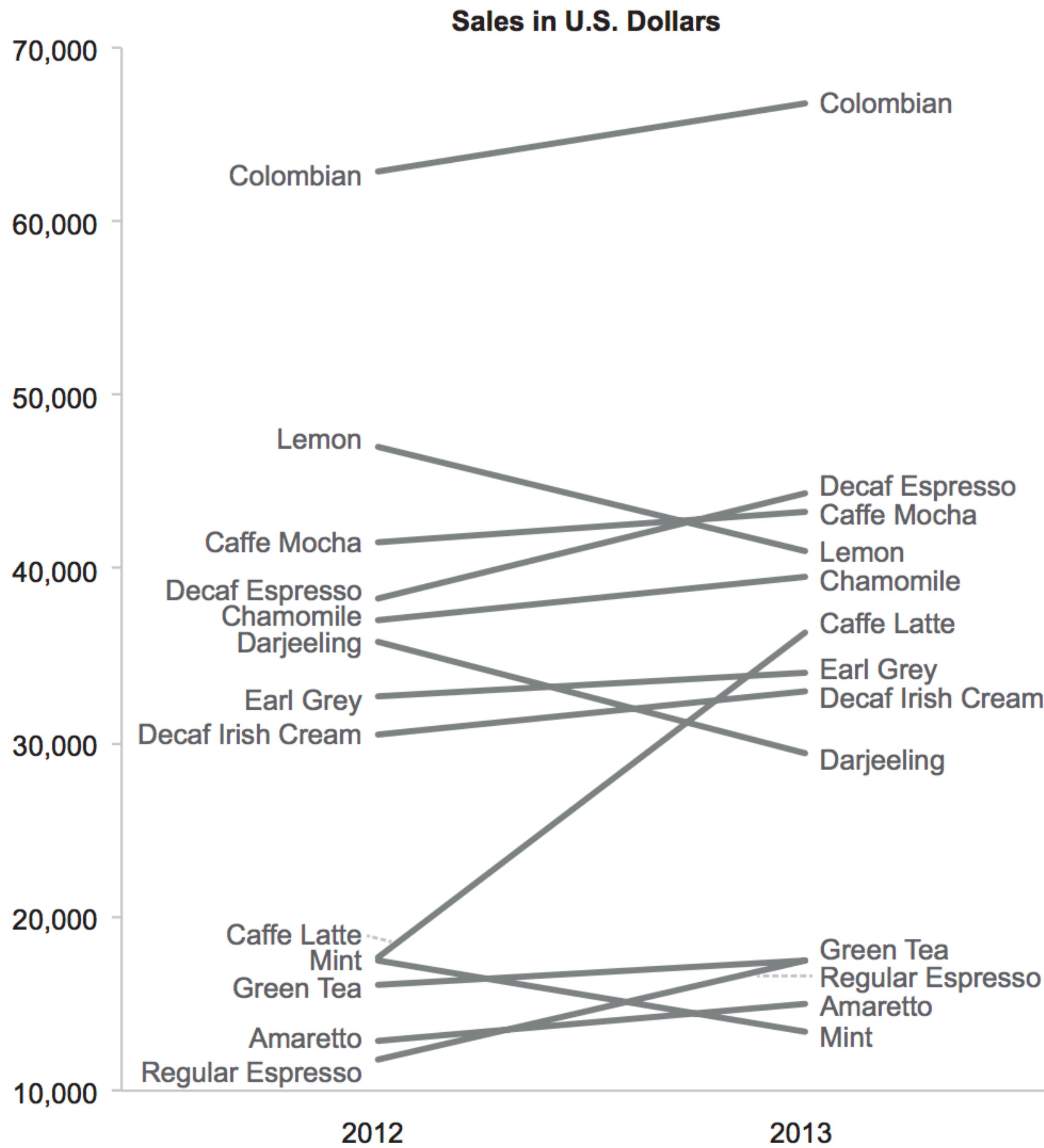


Use cores de fundo para destacar referência como por exemplo
um desvio padrão abaixo ou acima

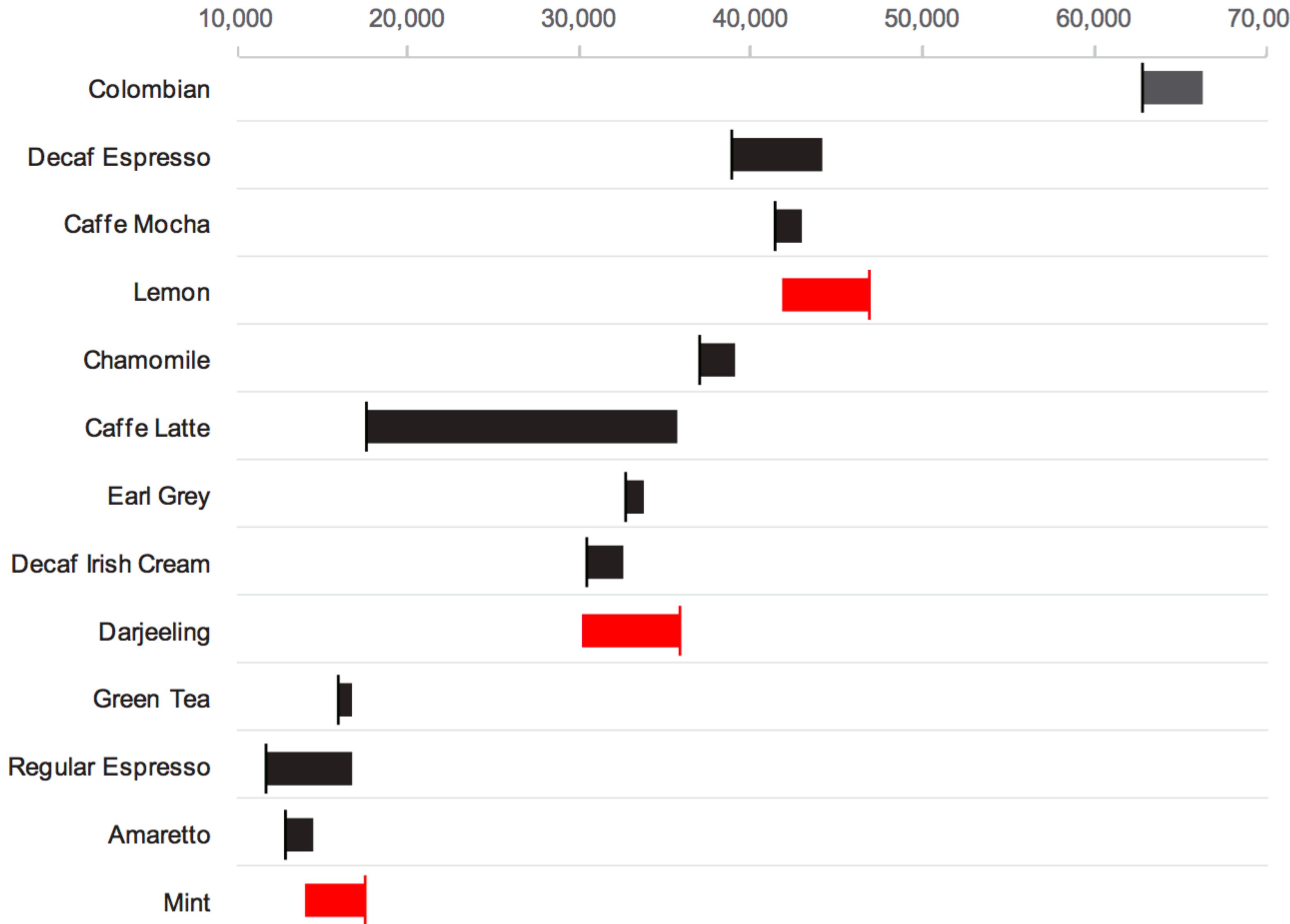
EXEMPLO PRÁTICO: MOSTRANDO DESVIOS SOBRE DOIS PONTOS NO TEMPO

*Displaying Change Between Two Points in
Time. Stephen Few, Perceptual Edge Visual
Business Intelligence Newsletter April/May/
June 2014*

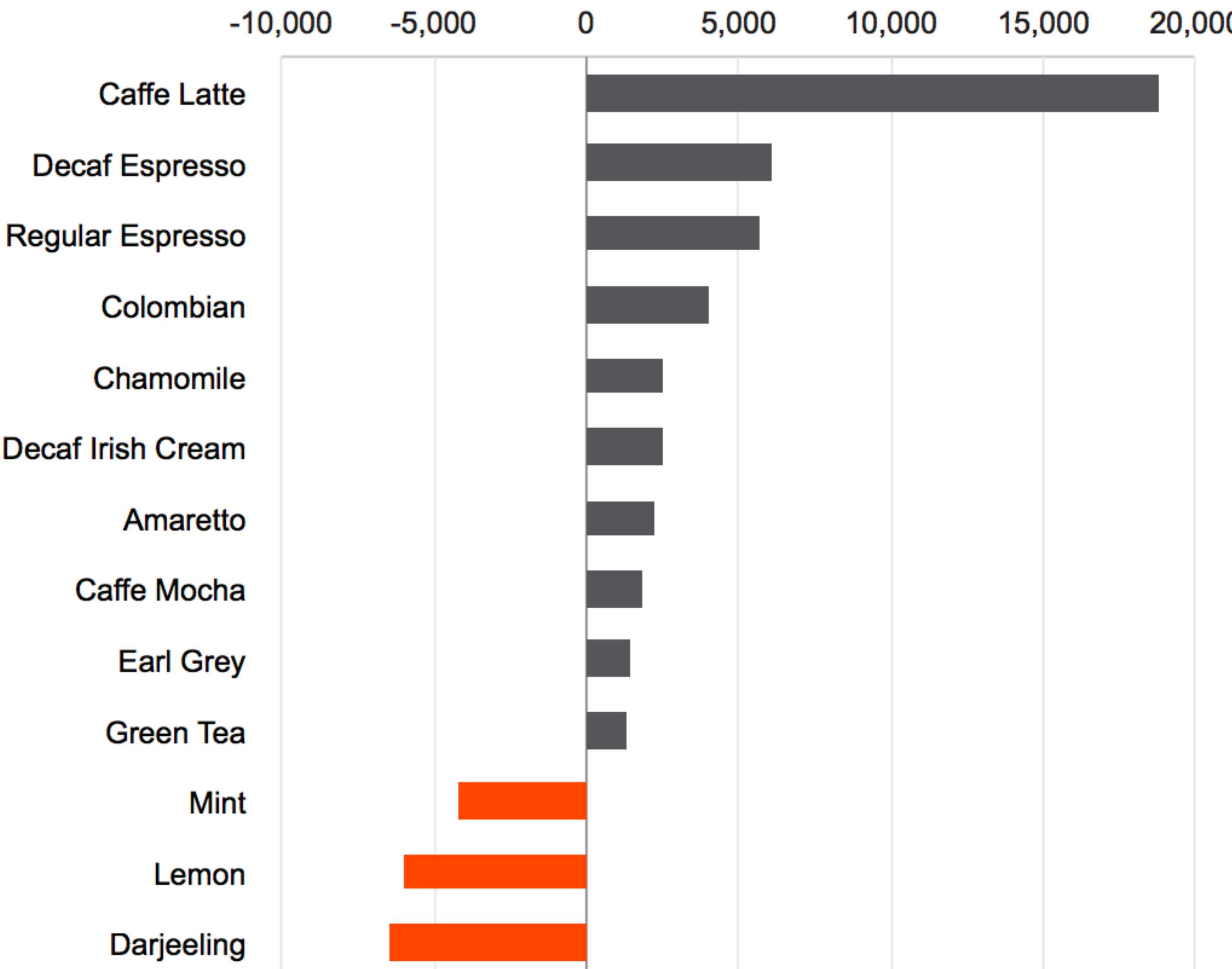




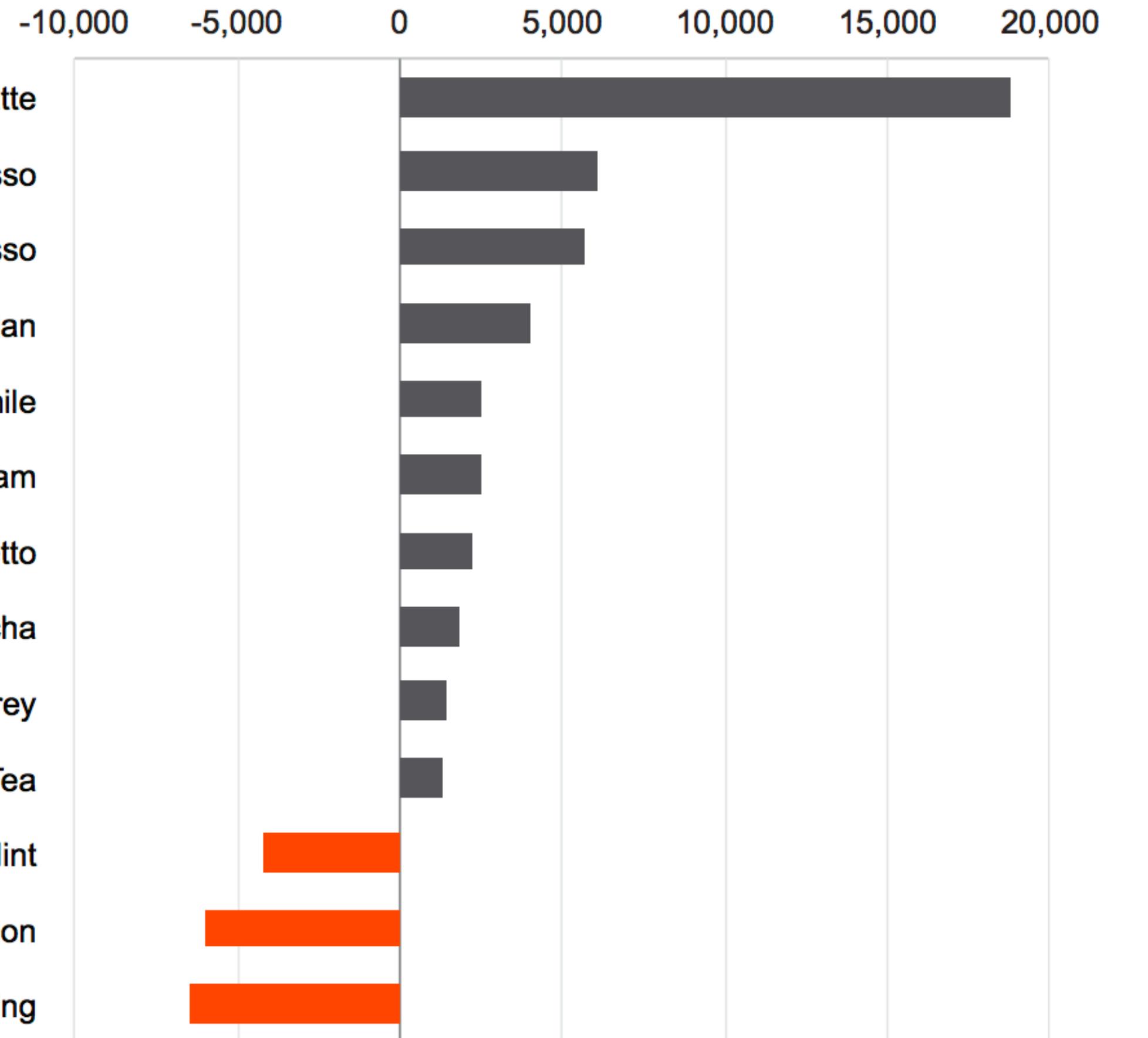
Change in Product Revenues from 2012 to 2013 (USD)



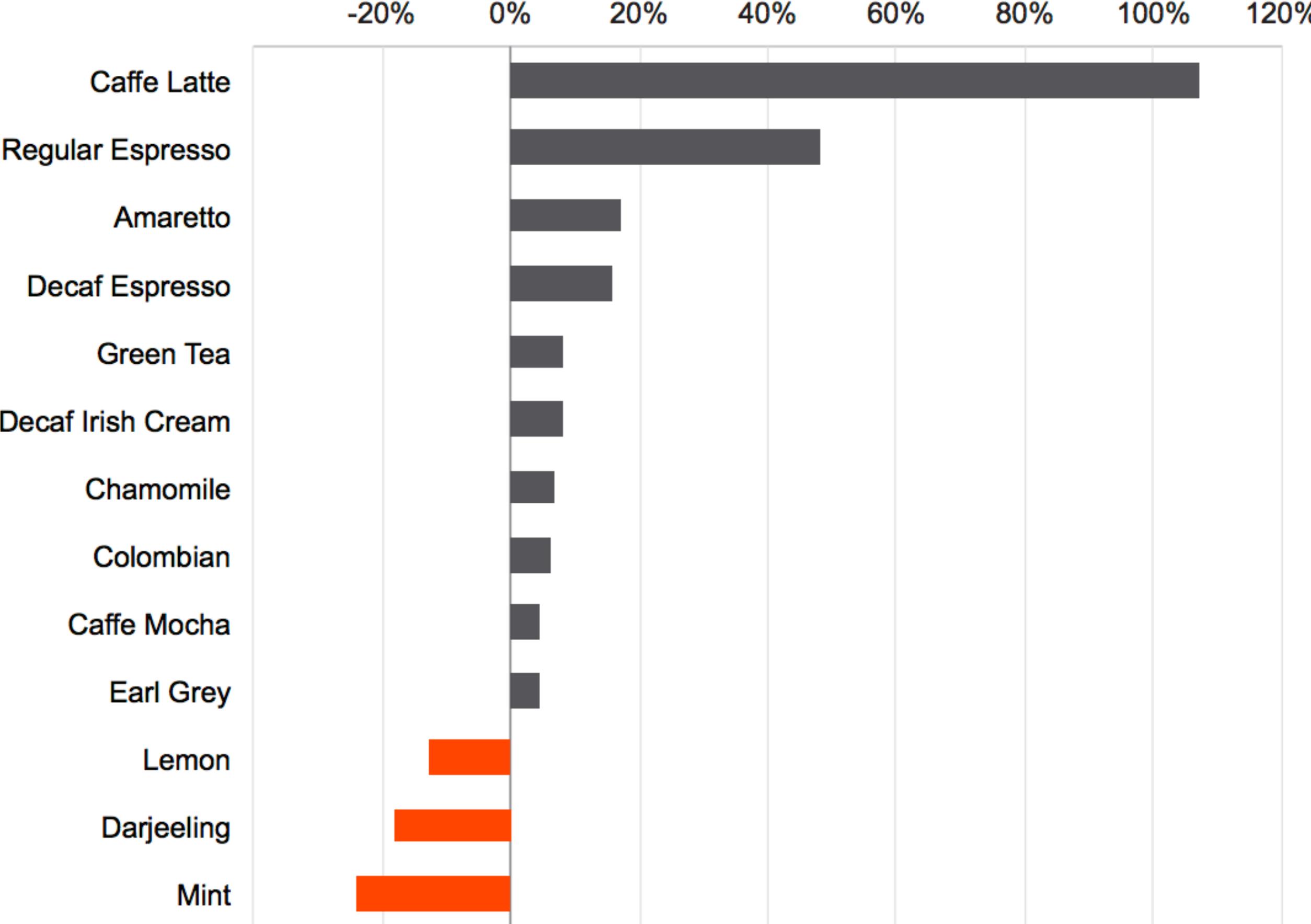
Change in Product Revenues from 2012 to 2013 (USD)



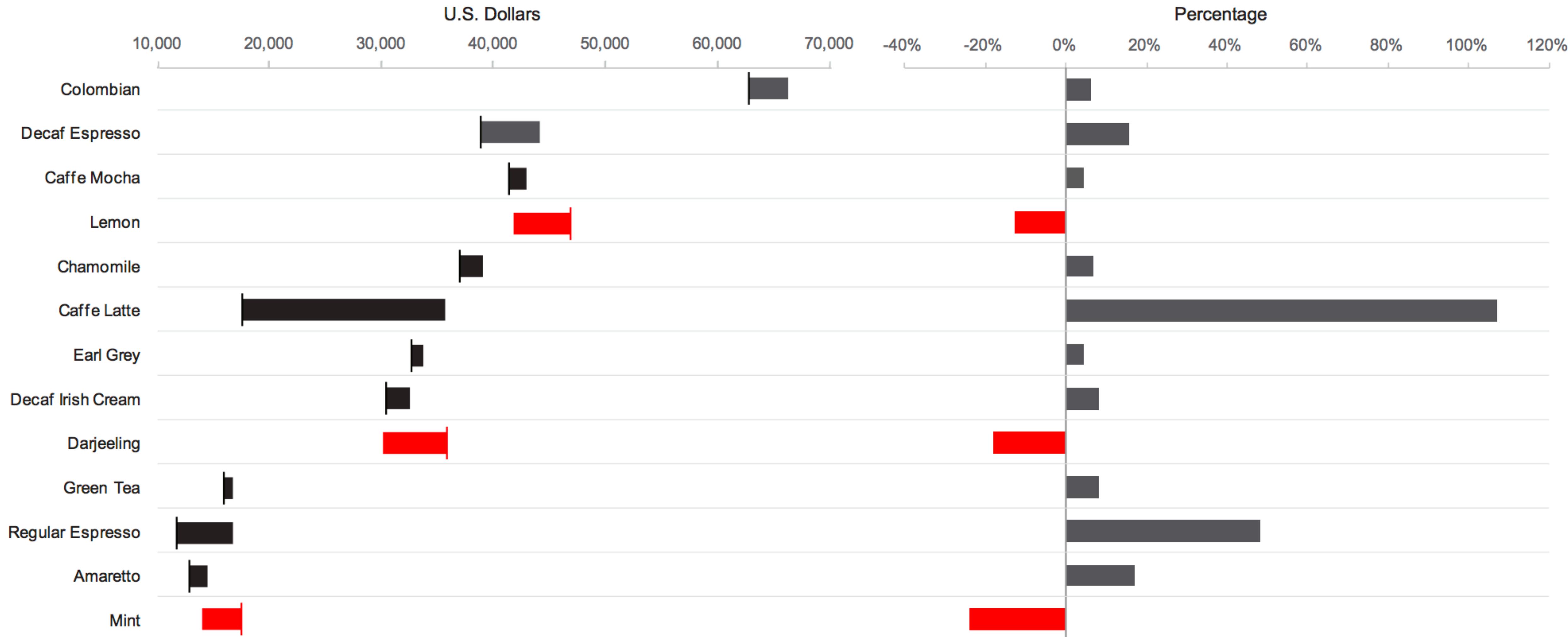
Change in Product Revenues from 2012 to 2013 (USD)



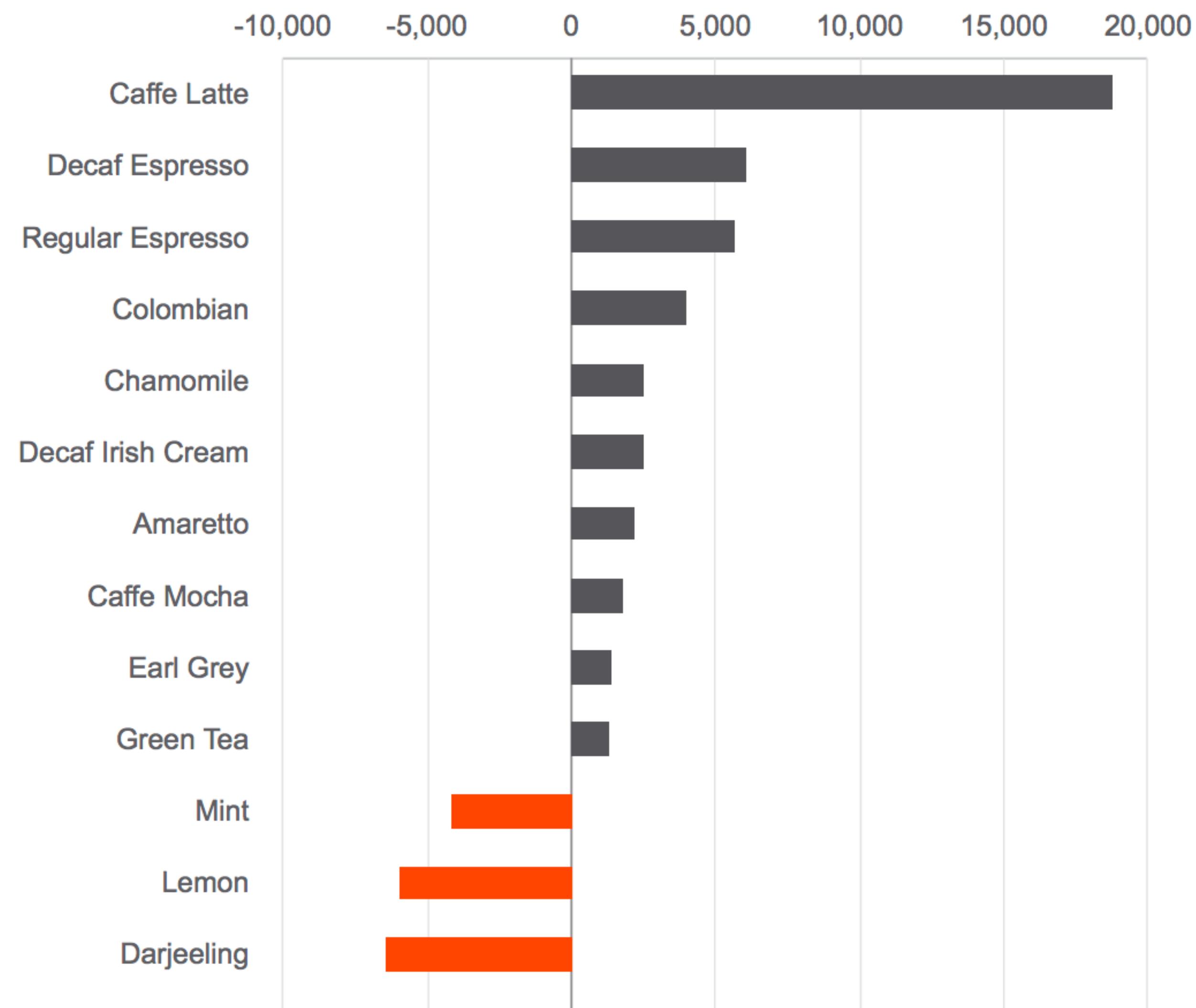
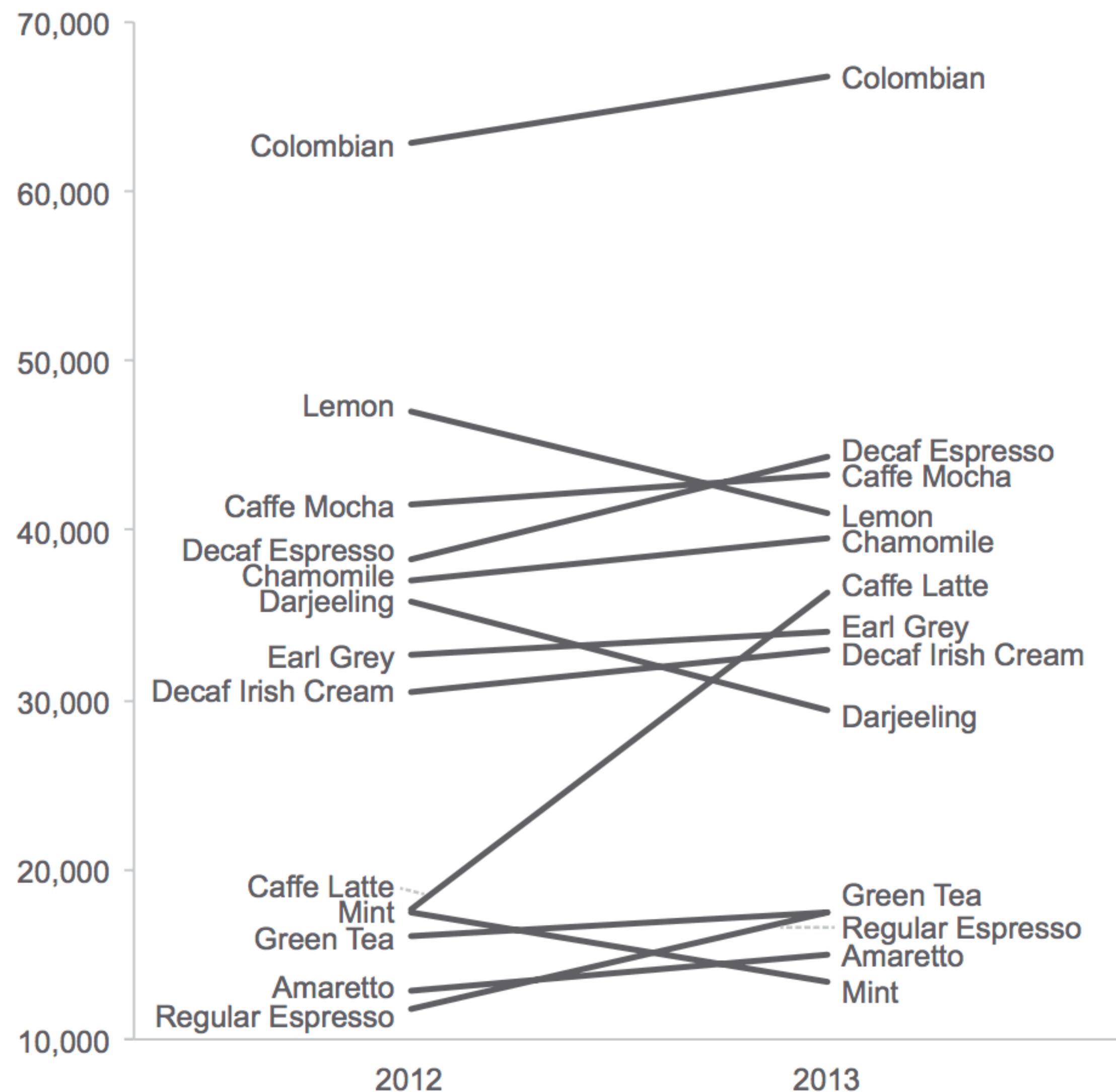
Percentage Change in Product Revenues from 2012 to 2013 (%)



Change in Product Revenues from 2012 to 2013



Change in Product Revenues from 2012 to 2013 (USD)



Feature	Line Graph	Range Bar Graph	Deviation Bar Graph
Easy to compare magnitudes of change	Good	Satisfactory	Excellent
Easy to compare directions of change	Good	Excellent	Excellent
Easy to compare values at the same point in time	Excellent	Good	N/A
Easy to compare rates of change	Good ¹	Poor	Excellent ²
Easy to spot changes in rank	Excellent	Satisfactory	N/A

¹ Quantitative scale must be switched from linear to logarithmic

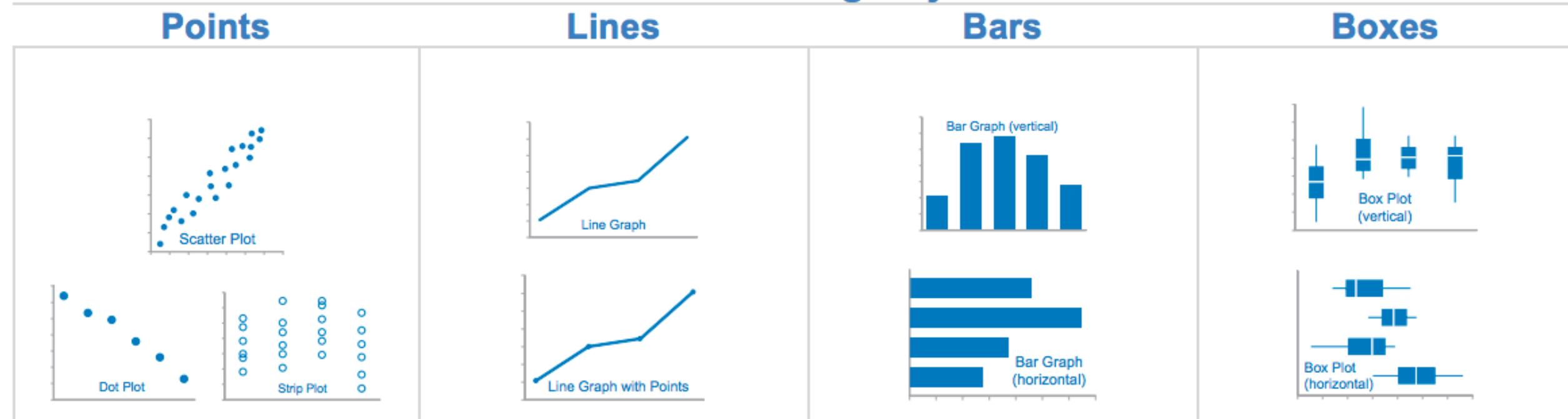
² Quantitative scale must directly express the rate of change

FINALIZAMOS OS RELACIONAMENTOS QUANTITATIVOS CLÁSSICOS

Featured Relationships

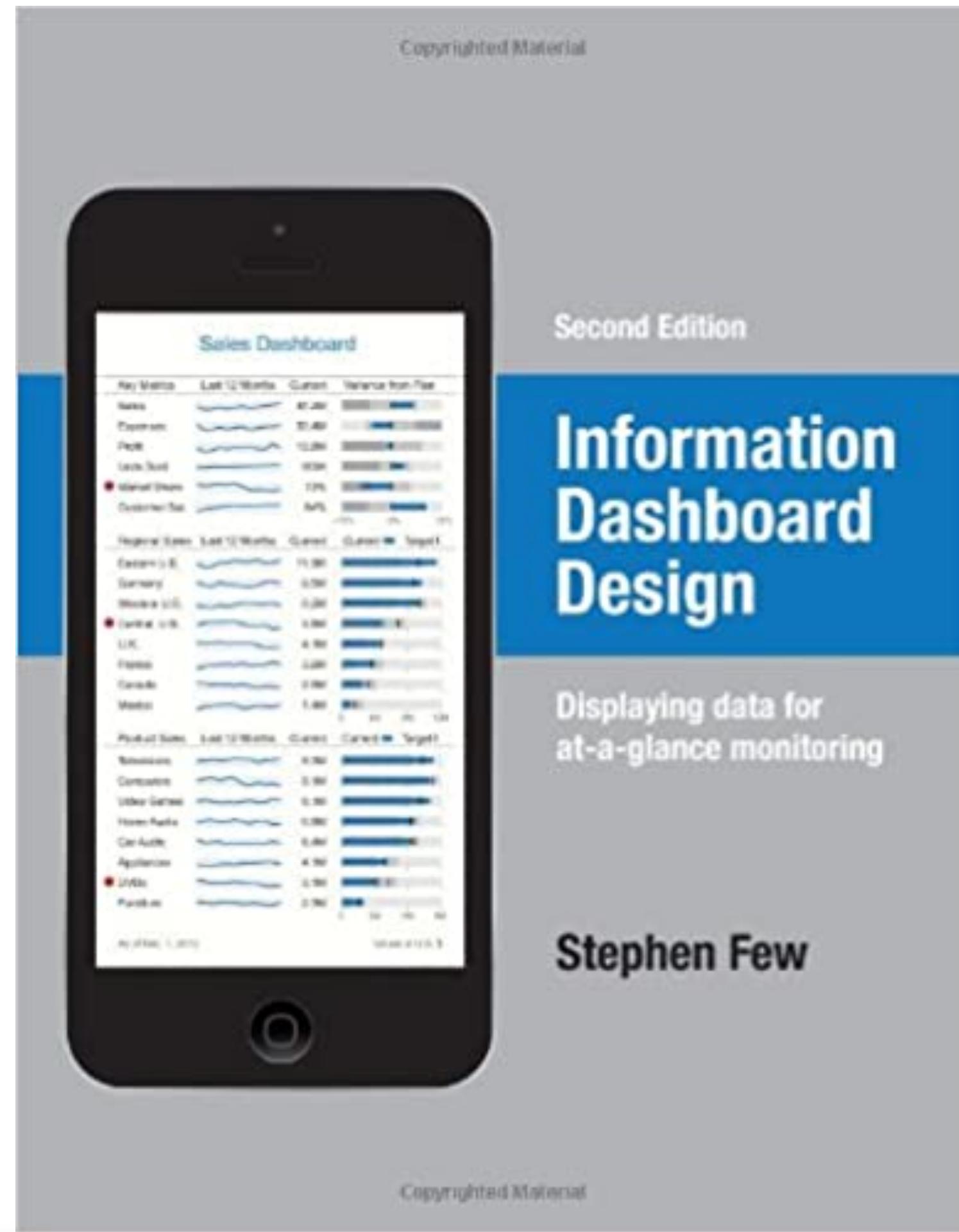
	Points	Lines	Bars	Boxes
Time Series Values display how something changed through time (yearly, monthly, etc.)	Yes (as a <i>dot plot</i> , when you don't have a value for every interval of time)	Yes (to feature overall trends and patterns and to support their comparisons)	Yes (vertical bars only, to feature individual values and to support their comparisons)	Yes (vertical boxes only, to display how a distribution changes through time)
Ranking Values are ordered by size (descending or ascending)	Yes (as a <i>dot plot</i> , especially when the quantitative scale does not begin at zero)	Yes (as a <i>bumps chart</i> , to show how rankings change through time)	Yes	Yes (to display a ranked set of distributions)
Part-to-Whole Values represent parts (proportions) of a whole (for example, regional portions of total sales)	No	Yes (to display how parts of a whole have changed through time)	Yes	No
Deviation The difference between two sets of values (for example, the variance between actual and budgeted expenses)	Yes (as a <i>dot plot</i> , especially when the quantitative scale does not begin at zero)	Yes (when also featuring a time series)	Yes	No
Distribution Counts of values per interval from lowest to highest (for example, counts of people by age intervals of 10 years each)	Yes (as a <i>strip plot</i> , to feature individual values)	Yes (as a <i>frequency polygon</i> , to feature the overall shape of the distribution)	Yes	Yes (when comparing multiple distributions)
Correlation Comparison of two paired sets of values (for example, the heights and weights of several people) to determine if there is a relationship between them	Yes (as a <i>scatter plot</i>)	No	Yes (as a <i>table lens</i> , especially when your audience is not familiar with <i>scatter plots</i>)	No

Value-Encoding Objects



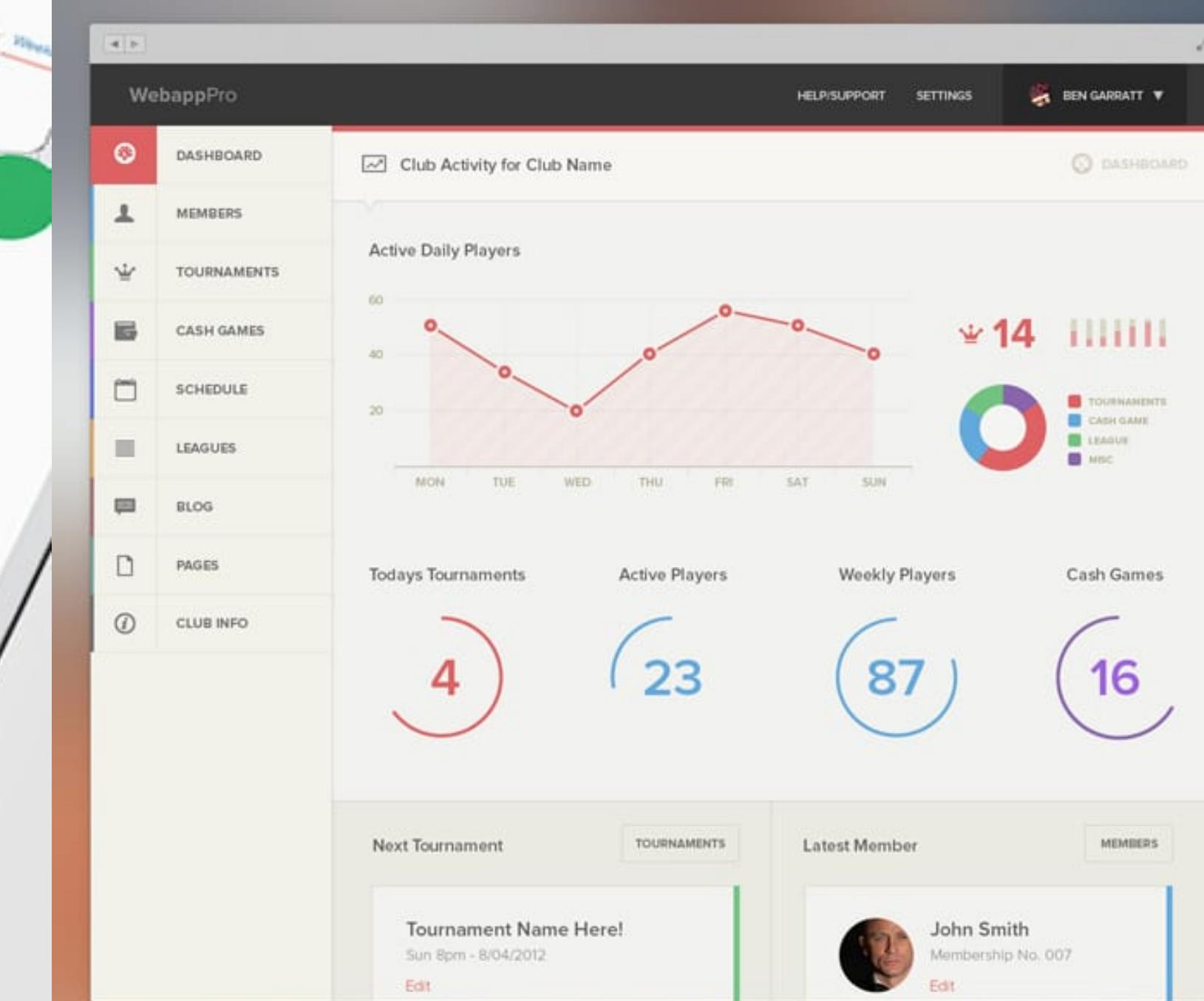
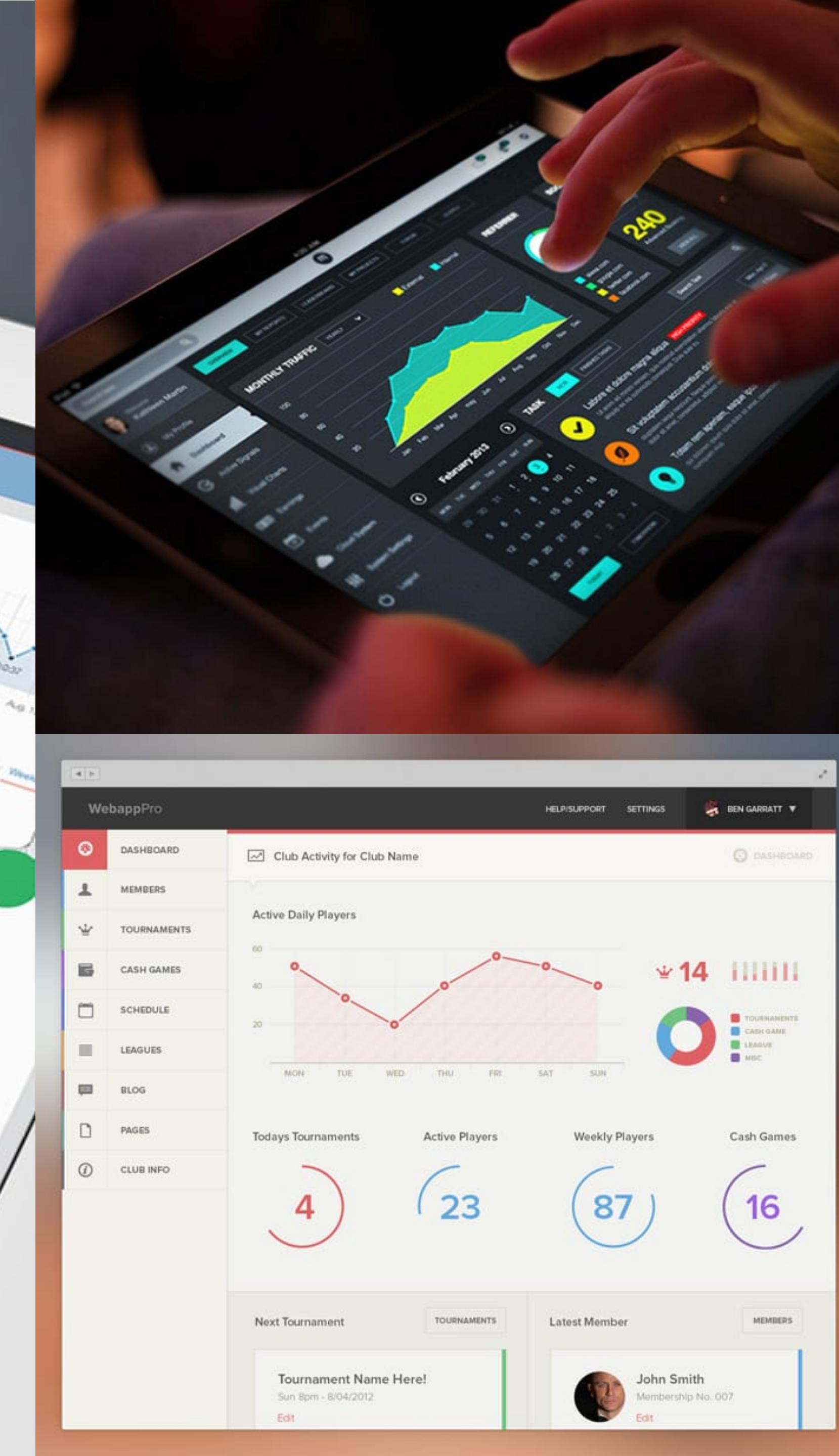
Dashboards

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



Information dashboard design: the effective visual communication of data

Stephen Few



DEFINIÇÃO

Representação visual
das
informações mais importantes necessárias para atingir um ou
mais objetivos
que
cabe inteiramente em uma tela de computador
de forma a poderem
ser monitoradas **simultaneamente**

REPRESENTAÇÃO VISUAL

- A informação é apresentada visualmente através de uma combinação entre texto e gráficos
- A apresentação gráfica pode comunicar os dados com maior eficiência
- Entendimento de como funciona a percepção visual

INFORMAÇÃO PARA ATINGIR METAS

- A resolução de problemas requer comumente uma coleção de dados normalmente oriundos de diversas fontes e de visões complementares

CABE EM UMA ÚNICA TELA

- Toda a informação apresentada deve caber em uma única tela de forma que possa ser analisada simultaneamente
- Não deve ser necessário rolar a tela
- Não deve ser necessário alternar entre telas

- O objetivo é ter toda a informação mais importante prontamente disponível

COMPONENTES (VISUALIZAÇÕES E TEXTO)

- Os componentes de um dashboard devem ser:
 - Pequenos
 - Concisos
 - Claros
 - Intuitivos

PAPEL

- **Estratégico:** apresentação dos principais indicadores para tomada de decisão
- **Analítico:** demanda contexto da informação mais riqueza de comparações, histórico dos dados
- **Operacional:** natureza dinâmica e de tempo real

TIPOS DE DADOS

- Quantitativos
- Não quantitativos

EXEMPLOS DE DADOS QUANTITATIVOS

- **Vendas:** reservas, cobranças, vendas antecipadas, encomendas, preços de vendas
- **Marketing:** market share, sucesso de campanhas, demografia de clientes
- **Financeiro:** renda, despesa, lucro
- **Suporte técnico:** número de chamadas, casos resolvidos, satisfação do cliente, duração de chamadas
- **Indústria:** número de itens produzidos, tempo de produção, número de defeitos
- **Recursos humanos:** satisfação dos empregados, rotatividade, número de posições em aberto, medidas de desempenho
- **Servidores web:** número de visitantes, número de acessos a páginas, tempo de duração das visitas

DADOS QUANTITATIVOS E O TEMPO

- Medidas quantitativas podem ser expressas em diferentes intervalos de tempo:
 - Ano atual até a data de hoje
 - Semana atual até a data de hoje
 - Quinzena atual até a data de hoje
 - Ontem
 - Mês atual até a data de hoje
- O intervalo ideal depende da natureza dos objetivos que o dashboard suporta

A IMPORTÂNCIA DE PERMITIR COMPARAÇÕES

- As métricas podem ser exibidas individualmente mas as comparações enriquecem o cenário:
- Métricas em um ponto passado
 - O mesmo dia no ano passado
- Métricas em algum ponto do passado:
 - O fim do ano passado
- Alvo atual para a métrica:
 - Lucro esperado

A IMPORTÂNCIA DE PERMITIR COMPARAÇÕES

- As métricas podem ser exibidas individualmente mas as comparações enriquecem o cenário:
- Relacionamento com um alvo futuro:
 - Percentual vendido em relação a meta anual
- Predição do valor de uma métrica:
 - Comparação do valor atual com o valor esperado
- Comparação com a norma:
 - Número de dias que são necessários para entregas

A IMPORTÂNCIA DE PERMITIR COMPARAÇÕES

- As métricas podem ser exibidas individualmente mas as comparações enriquecem o cenário:
- Projeções futuras:
 - Com base no crescimento do número de unidades vendidas ao mês, projeção da quantidade de vendas no próximo ano
- Métricas alheias:
 - Comparação com métricas dos concorrentes
- Métricas relacionadas:
 - Número de pedidos comparado ao lucro

- Uma vez que dashboards devem possibilitar a análise dos dados com relativa rapidez, representações visuais de avaliações das comparações são muito úteis
- Medidas em faixa de alerta podem ser indicadas em vermelho

- Há espaço para dados qualitativos em dashboards?
- Que tipo de dados?

EXEMPLOS DE DADOS QUALITATIVOS

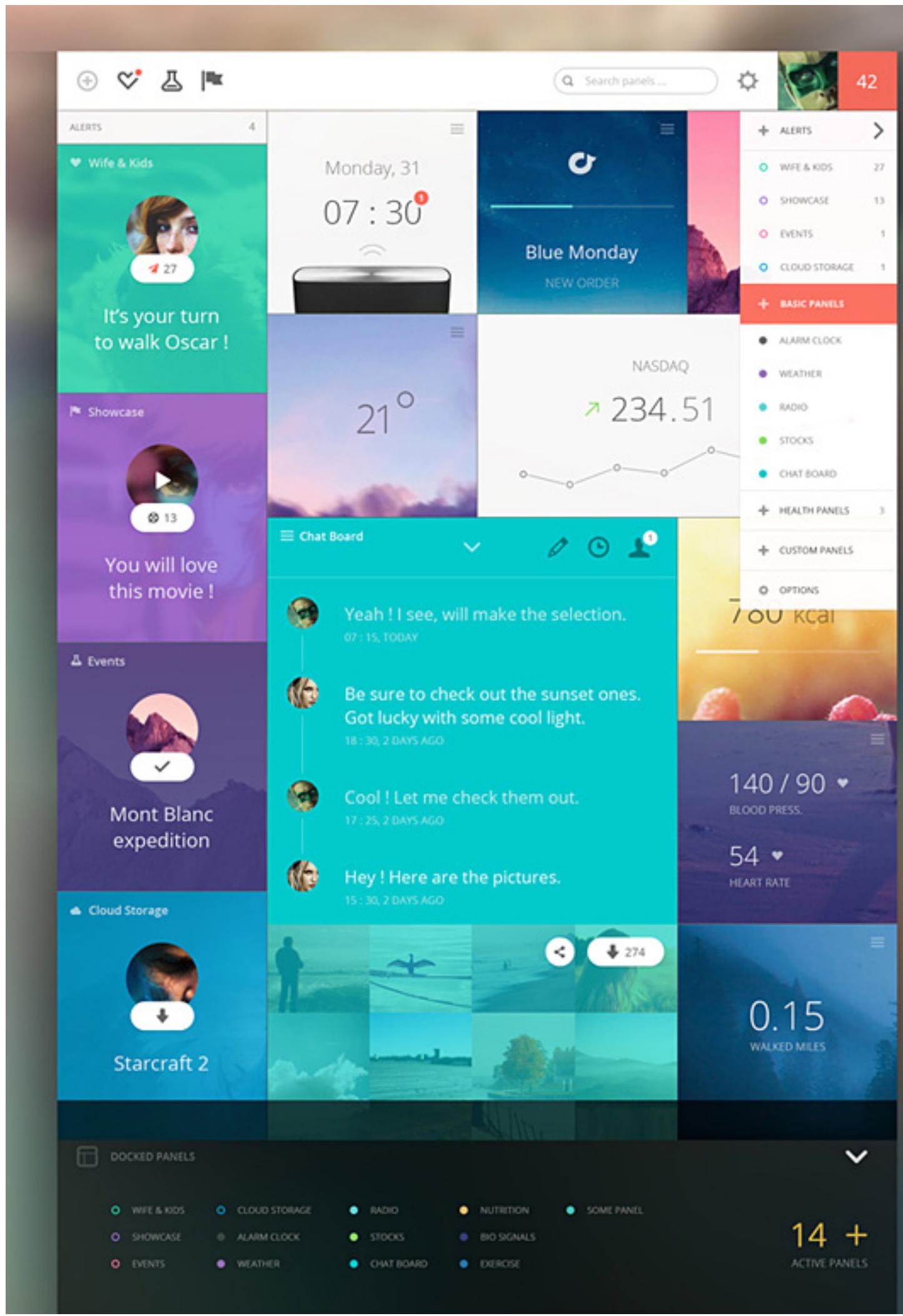
- Os n melhores clientes
- Questões que devem ser investigadas
- Tarefas a serem completadas
- Pessoas que devem ser contatadas
- Datas de entrega
- Pessoas responsáveis por tarefas

ERROS FREQUENTES

DESAFIO FUNDAMENTAL NO DESENHO DE DASHBOARDS

- Apresentar grande **quantidade de informação** em um **pequeno espaço** resultando em uma **visualização facilmente comprehensível**

EXCEDER OS LIMITES DA TELA



[WebApp] Reports Daniel Alelio 3 ? Q

IMO Report Contract Report •
Users CRM MDM Plan Builder Disco

Champs of The Month

Average time of processing one contract

Oct 14 - 20 Oct 21 - 27 Oct 28 - Nov 3 Nov 4 - 10 Nov 11 - 17 Nov 18 - 24

Info 1 Info 2

Month wise contract Processed For 2013
Total Contract Processed : 1000

Info 2
Info 3
Info 1
Info 3

Data Created	Store Name	Channel Name	Manager Name	Location	Status
01/12/13	Store Name	Channel Name	Manager Name	Location	Active
01/12/13	Store Name	Channel Name	Manager Name	Location	Active
01/12/13	Store Name	Channel Name	Manager Name	Location	Active

REFINE MY REPORTS DELIVERY

SELECT REPORTS

- Master Report
- My Report 1
- My Report 2
- My Report 3
- My Report 4

DOWNLOAD OPTIONS

- CSV Download
- Email

Email Options

- PDF CSV
- abc@email.com +/-
- abc@email.com +/-

Schedule

- Schedule Time 11:30 AM
- Frequency Daily

FTP OPTION

- FTP Dump

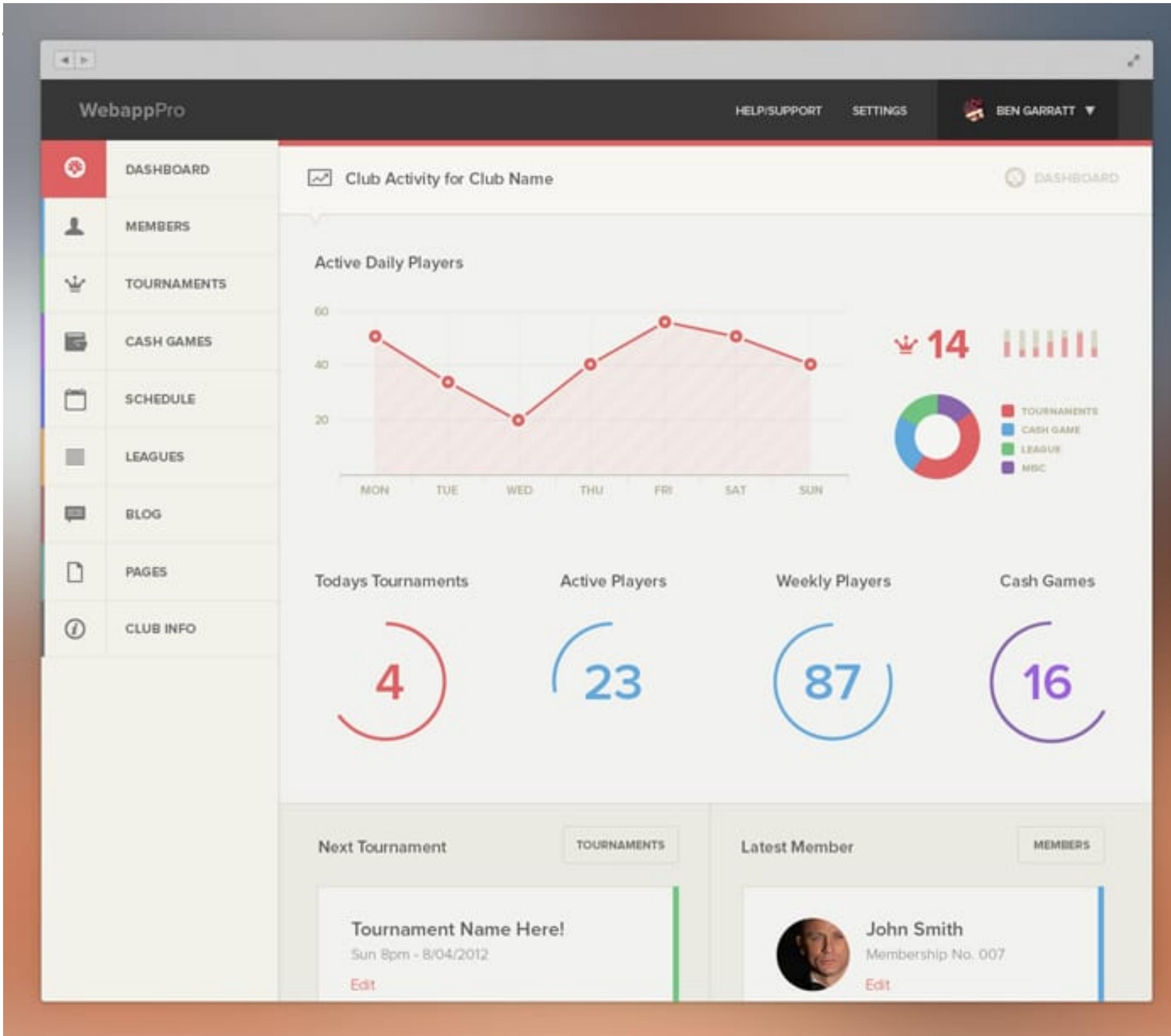
Login
Password
Path

FRAGMENTAR OS DADOS EM MULTIPLAS TELAS

The image displays three separate screenshots illustrating how data can be fragmented across multiple screens or devices:

- Screenshot 1 (Left):** A dashboard for monitoring website traffic. It features a main chart showing traffic over time (15 April - 15 May), three circular KPIs (New Sessions: 80.05%, Bounce Rate: 33.33%, Goals Completed: 65.75%), and sections for Top 5 Locations, Most Viewed Pages, and Age Groups.
- Screenshot 2 (Middle):** A user profile page for "Newton Barley". It includes a sidebar with navigation links (HOME, JOBS, RESUMES, TASKS, CALENDAR, USERS, LOCATIONS, WORKFLOWS, JOB BOARDS, SOCIAL), a summary section with counts for applicants, interviews, and forwards, and a detailed view of a job listing for a Customer Service Representative.
- Screenshot 3 (Right):** A job listing for a Customer Service Representative. It shows a summary section with counts for applicants, interviews, and forwards, a bar chart of applicants per day, and a "JOB DESCRIPTION" section with fields for POSITION TITLE, LOCATION, EMPLOYMENT TYPE, EXPERIENCE, and STATUS. It also includes social sharing buttons for Facebook, Twitter, LinkedIn, and a "Link to this Job" button.

NÃO FORNECER OU FORNECER O CONTEXTO DE FORMA INADEQUADA



myHealth App

Dashboard

Log

Steps

Calories

Distance

Time

Friends

Reports

Search

+

username 2

Steps: 7.2% of 5823

Calories: 31% of 2184

Distance: 52% of 3.5mi

Time: 35% of 1.5hr

Steps taken this week

Day	Steps
Sun	~4500
Mon	~7500
Tues	~4000
Weds	~6000
Thurs	~8500
Fri	~1000
Sat	0

Goal: 44%

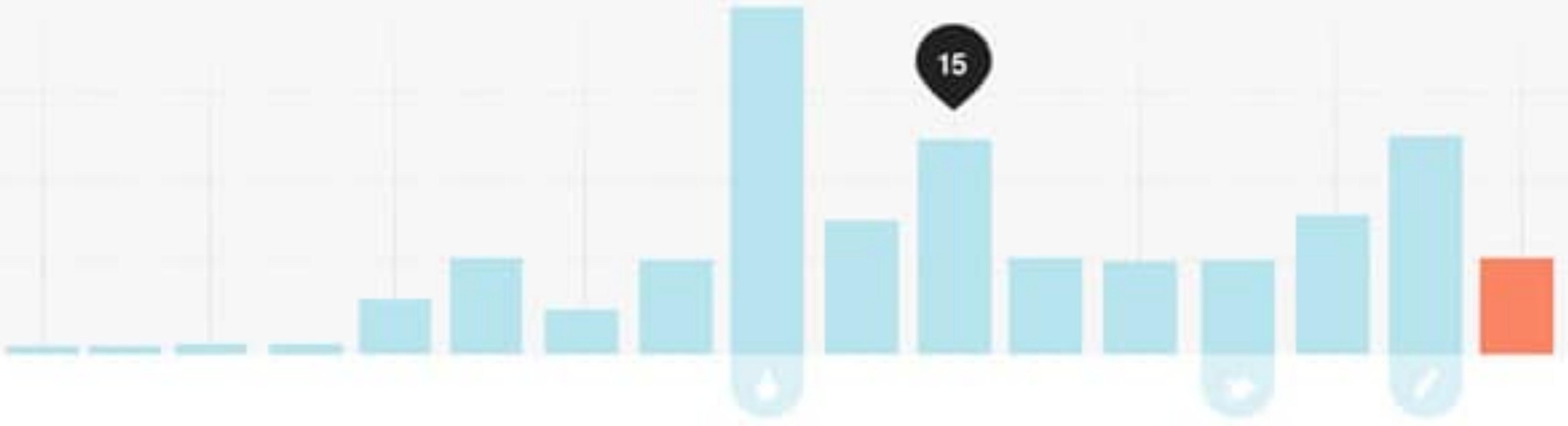
Recent Comments

Tell me, why this strong young colt, foaled in some peaceful valley of Vermont, far removed from all beasts of prey—why is it that upon the sunniest day, if you but shake a fresh buffalo robe behind him, so that

Friends

Anaida Parghel

Lucia Gómez



Stats

Moyers shakers, milk shakers, and quakers.

232

Total Views

12

Total Favorites

4

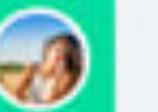
Achivements

232

Total Views



Refilling



Bieżące wartości indeksów

Przeliczenie z dnia: 16/05/2016



Masa Ciała

Ostatni Miesiąc

Tkanka Tłuszcowa

Ostatni Tydzień

- DASHBOARD
- ACCOUNTS
- COMPANIES

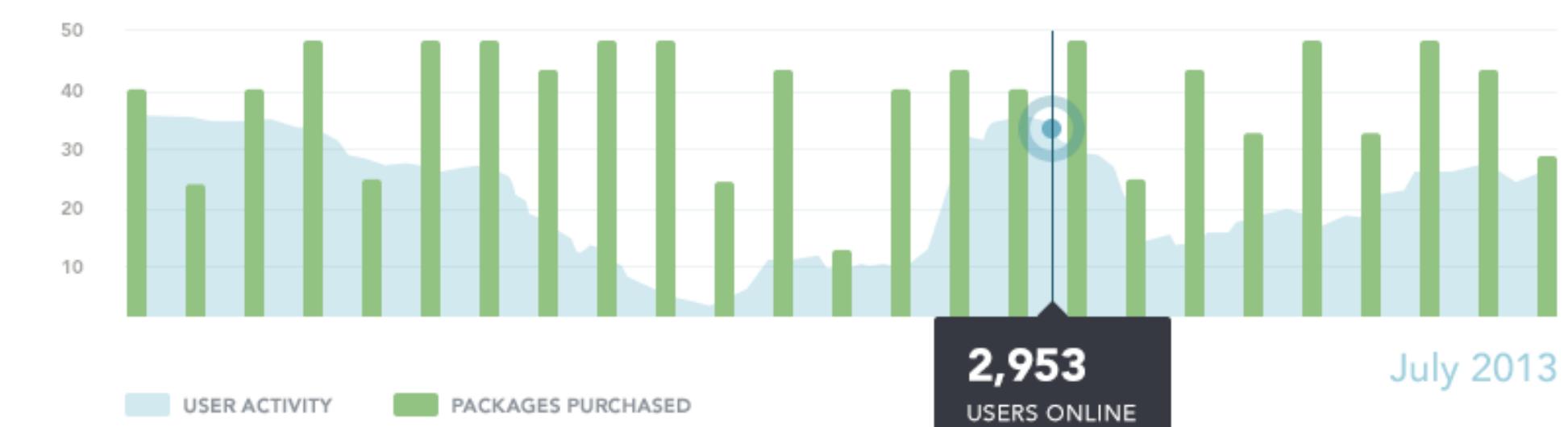
ACCOUNT USERS

- Bruce Wayne
- Lex Luthor
- Thor Odinson
- Sean Cassidy
- Clark Kent

23%
SALES INCREASE

\$5,894
TOTAL SALES

345
PACKAGES SOLD

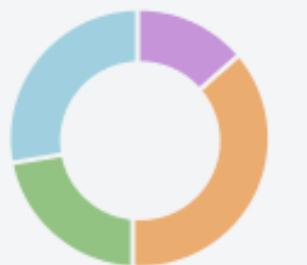


ACCOUNTS USAGE



30 USED 20 AVAILABLE

SALES BY COMPANY



- SHOGUN
- H & M
- MARTIAL CITY
- FORWARD

EXIBIR DETALHES OU PRECISÃO EXCESSIVOS

The screenshot shows a Microsoft Internet Explorer window displaying the Celequest Activity Dashboard. The dashboard includes a navigation tree, a list of active alert messages, and several data visualizations.

Alerts: A table titled "Active Alert Messages" lists six alerts. The last two alerts, dated 03/15/2004 at 17:10:08 and 17:09:46, are highlighted with a red border.

Subject	Importance	Alert Activated
8/16/2003 Yield Drop in ESS on 60-00...	Normal	03/15/2004 17:10:08
8/16/2003 Yield Drop on 60-0001663 ...	High	03/15/2004 17:10:08
8/13/2003 Yield Drop in ESS on 60-000200...	Normal	03/15/2004 17:10:01
8/13/2003 Critical Component Failure (60-0...	High	03/15/2004 17:10:00
8/13/2003 Impacted Boards for 11-0000040...	High	03/15/2004 17:09:59
8/1/2003 Yield Drop in ESS on 60-000...	Normal	03/15/2004 17:09:46

Data Visualizations: Three charts are displayed: "Board Yield Barchart", "Board Yield Change Barchart", and "Tests Breakdown Pie".

- Board Yield Barchart:** A bar chart comparing yield across different time periods (Yield_1Yr, Yield_3Days, Yield_30Days, Yield_Today) for various product numbers.
- Board Yield Change Barchart:** A bar chart showing yield change over different time periods (Yield_Change_1Day, Yield_Change_1Week, Yield_Change_1Yr).
- Tests Breakdown Pie:** A pie chart showing the distribution of test results by category.

Table Summary: A table titled "Board Yield Table Summary" provides detailed yield data for specific products. The last row, for product 60-0001663-03, is highlighted with a red border.

PRODUCT_NUM	PRODUCT_DESC	YIELD_TODAY	YIELD...	YIELD...	YIELD...	YIELD...	YIELD_CHAN...	YIELD...
40-0000364-05	PCBA,ER05,AP7420	100.000000000	100.0000	100.0000	100.0000	0.0000	0.0000000000	0.0000
60-0000720-01	ASSY,16 PORT CARD,SI,SW12000	89.4308943100	89.6000	98.0535	98.0535	-8.4535	-0.1691056900	-8.4535
60-0001624-06	ASSY,CP,FULL LENGTH	100.000000000	100.0000	99.1549	99.1549	0.8451	0.0000000000	0.8451
60-0001663-03	ASSY, INNER BOX W/MB, SW3600	100.000000000	100.0000	99.1111	99.1111	0.8889	0.0000000000	0.8889

Figure 3-6. This dashboard shows unnecessary detail, such as times expressed to the second and measures expressed to 10 decimal places.

ESCOLHER UMA MÉTRICA RUIM

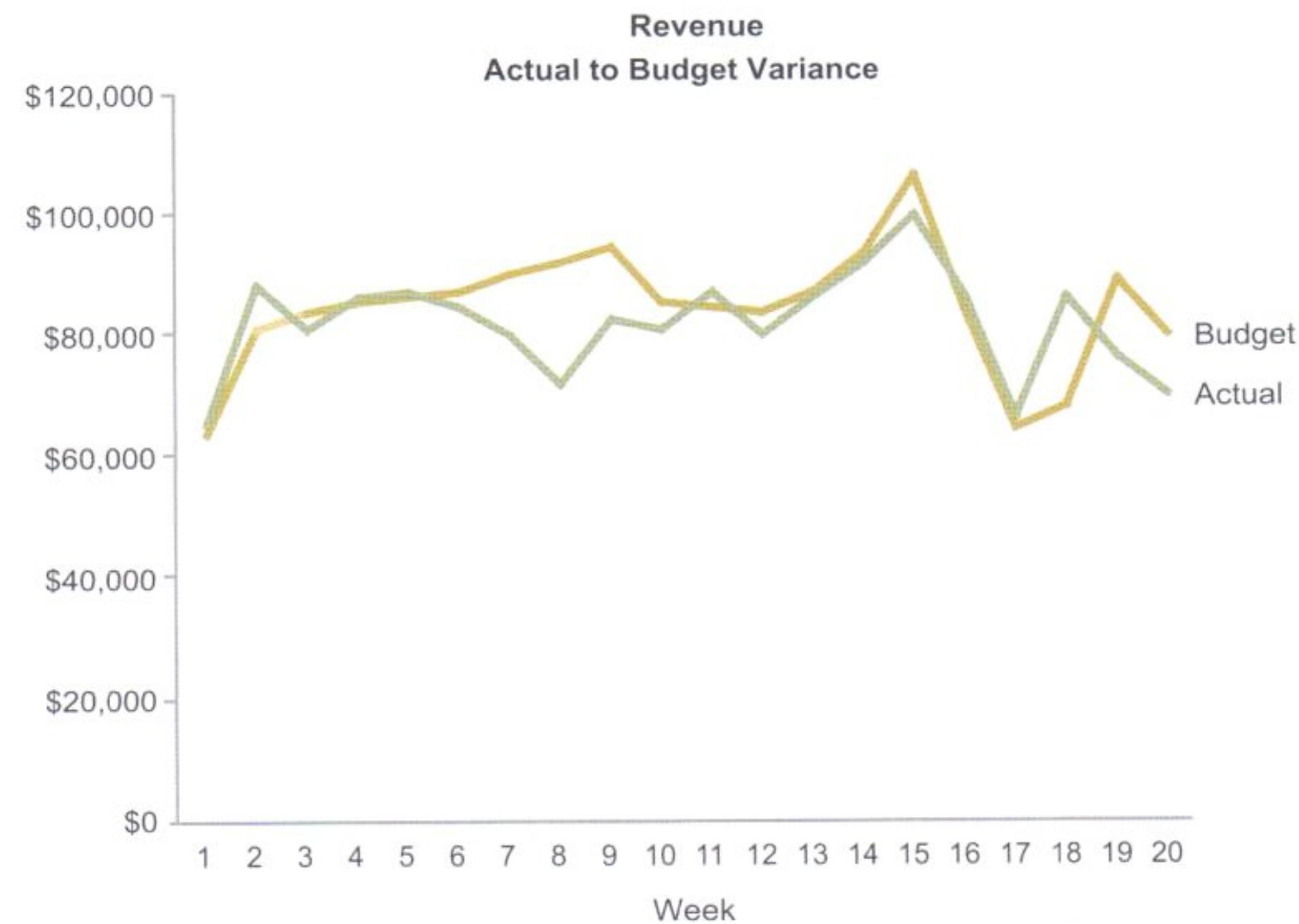


Figure 3-7. This graph illustrates the use of measures that fail to directly express the intended message.

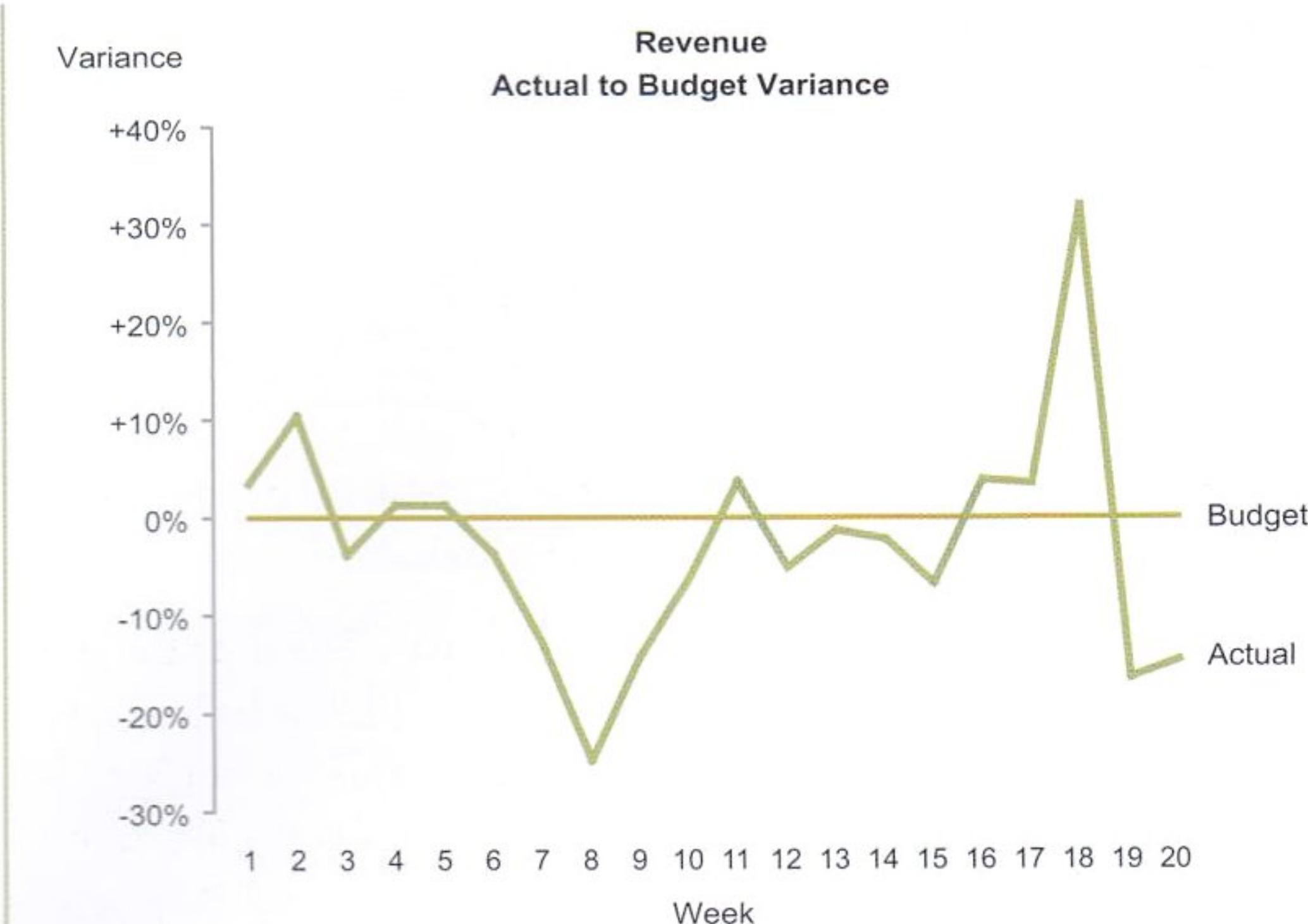


Figure 3-8. This graph is designed to emphasize deviation from a target, which it accomplishes in part by expressing the difference between budgeted and actual revenues using percentages.

USAR REPRESENTAÇÕES VISUAIS INADEQUADAS

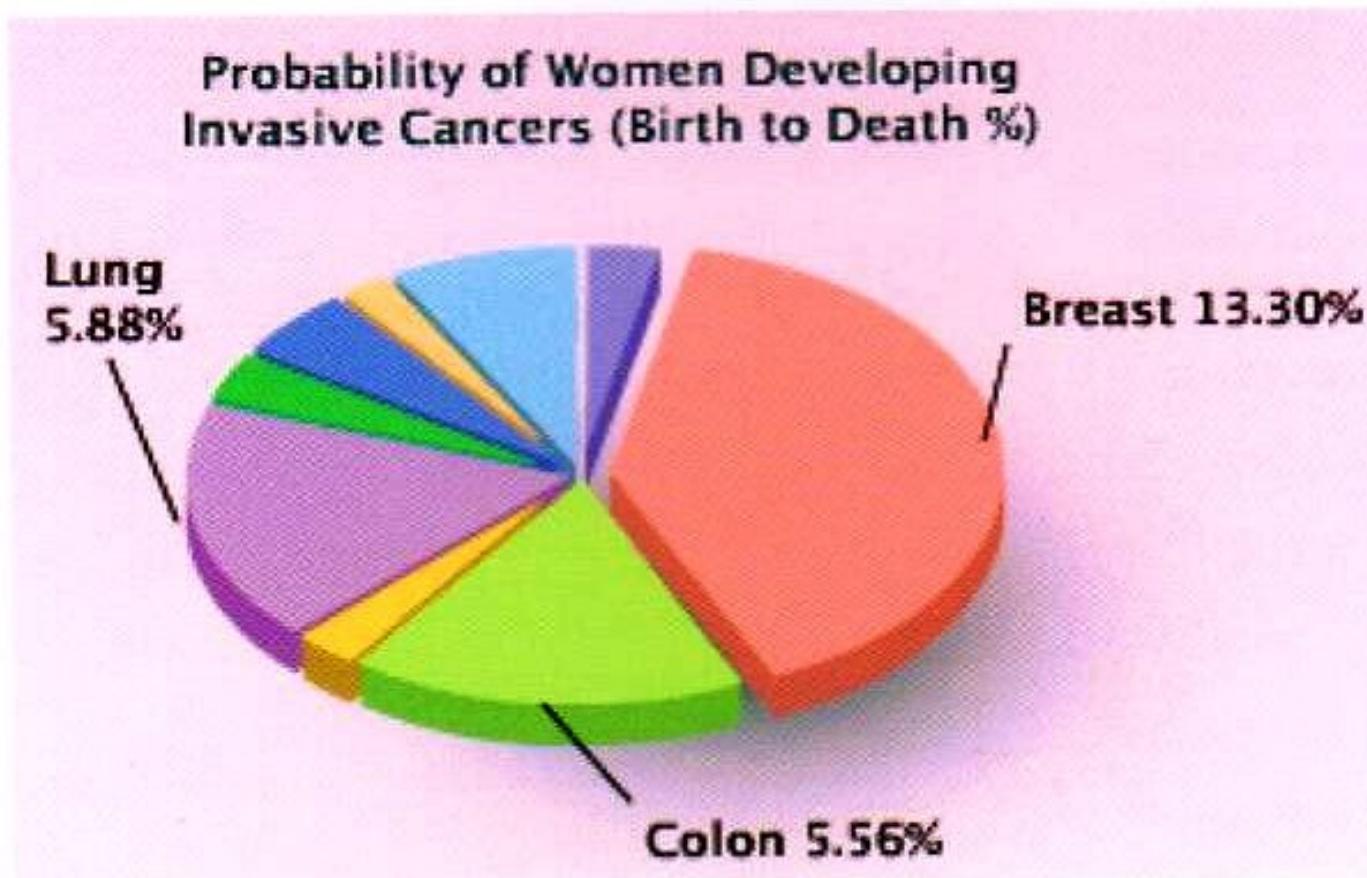


Figure 3-9. This chart illustrates a common problem with pie charts.

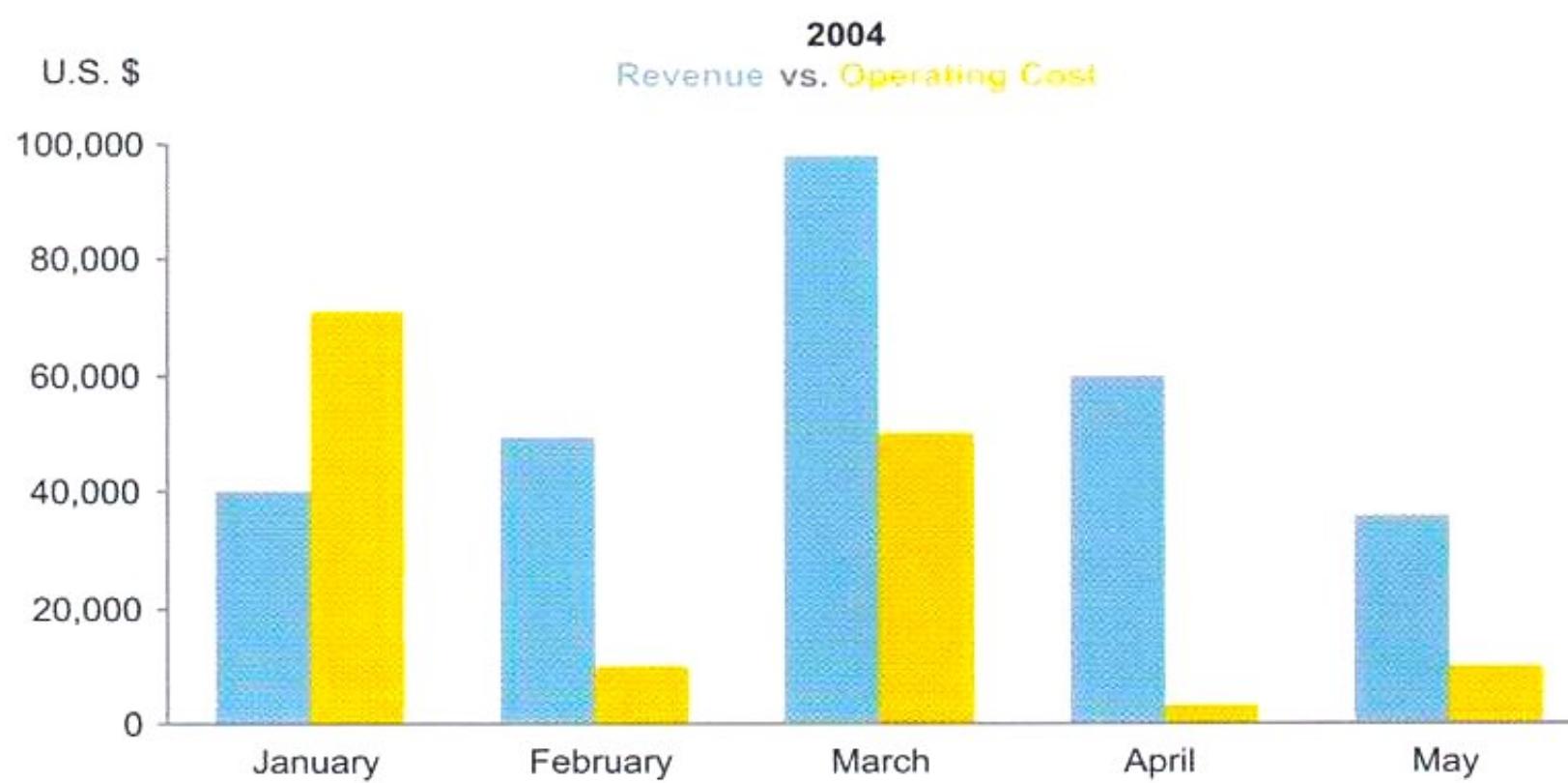


Figure 3-13. This bar graph does a good job of displaying a time series of actual versus budgeted revenue values.



Figure 3-12. This graph uses the two-dimensional area of circles to encode their values, which needlessly obscures the data.

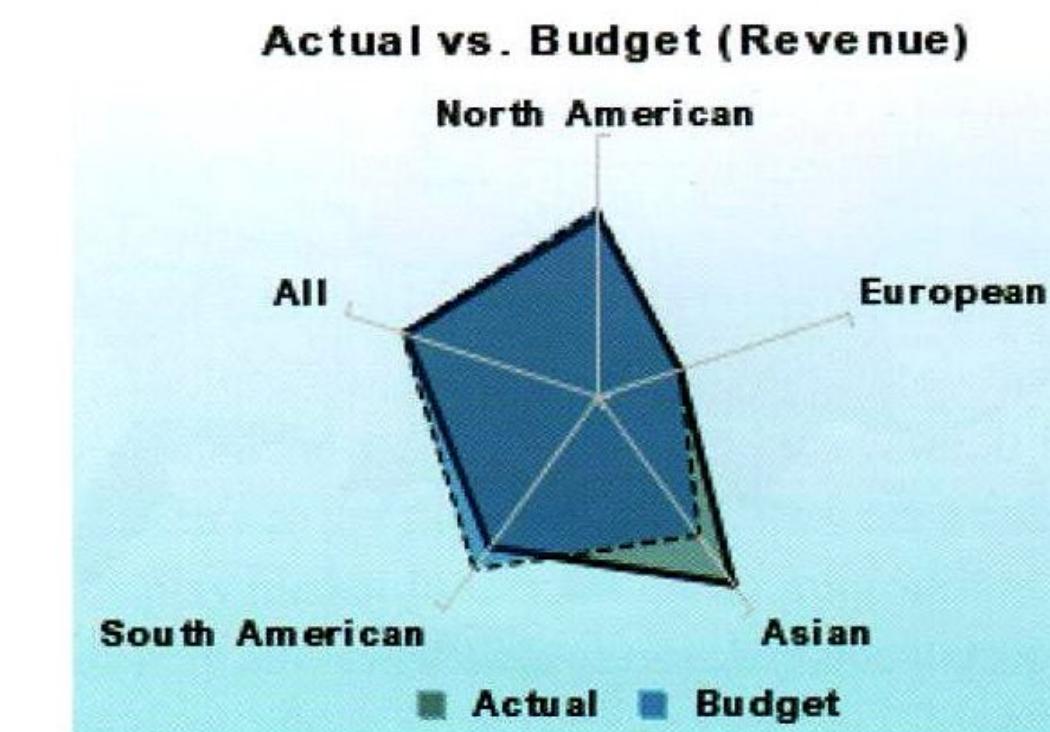


Figure 3-14. This radar graph obscures the straightforward data that it's trying to convey.

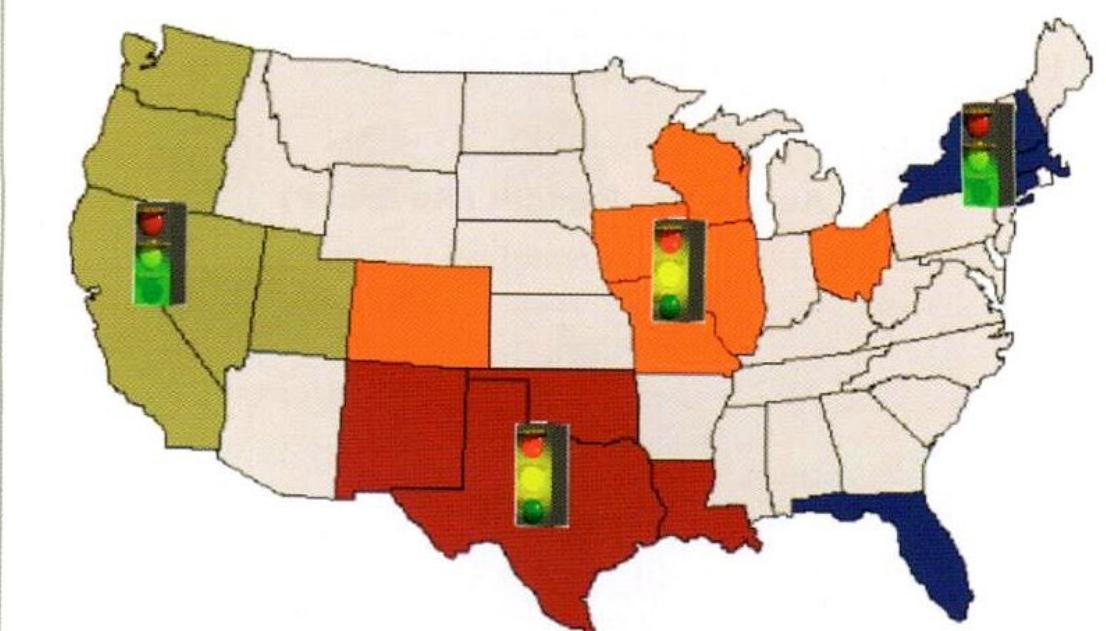


Figure 3-16. This display uselessly encodes quantitative values on a map of the United States.

USO DE REPRESENTAÇÕES VISUAIS POBRES

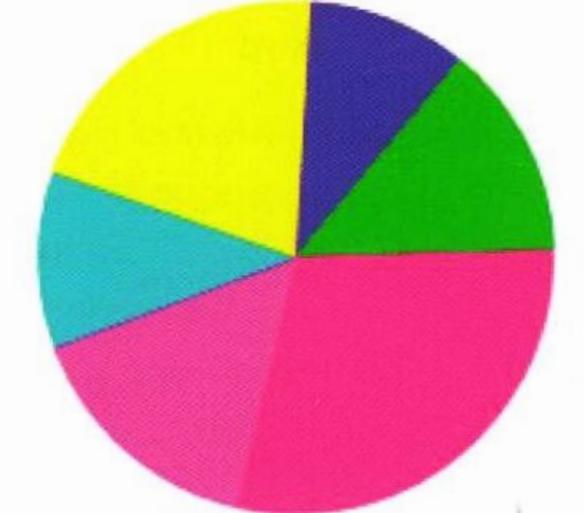


Figure 3-19. This pie chart illustrates several design problems.

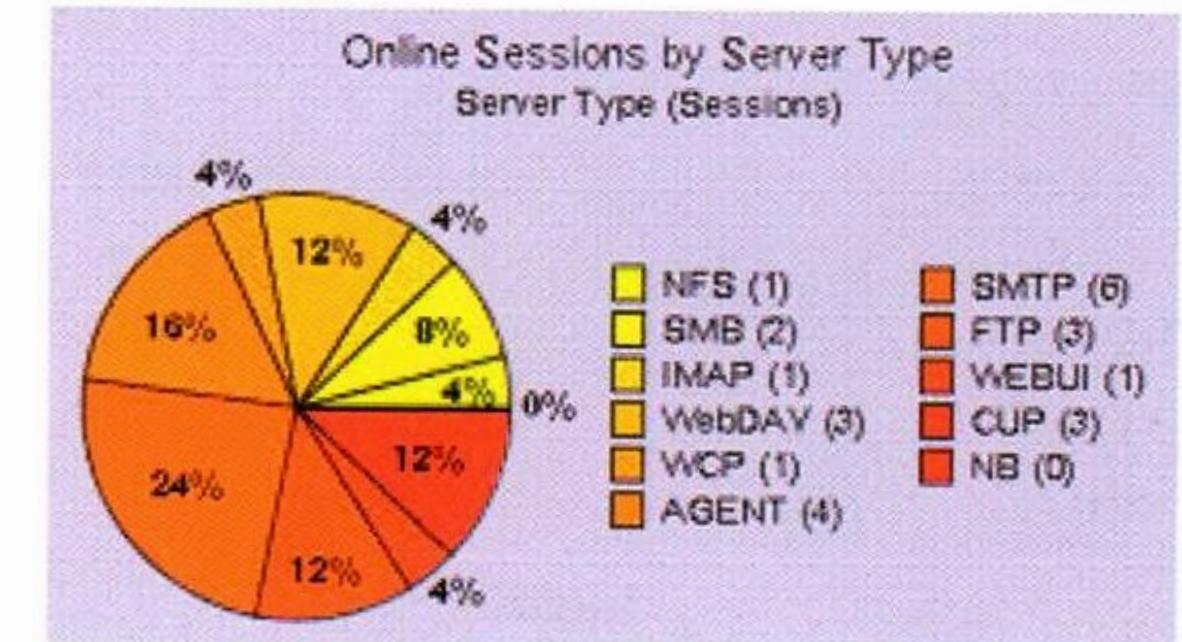
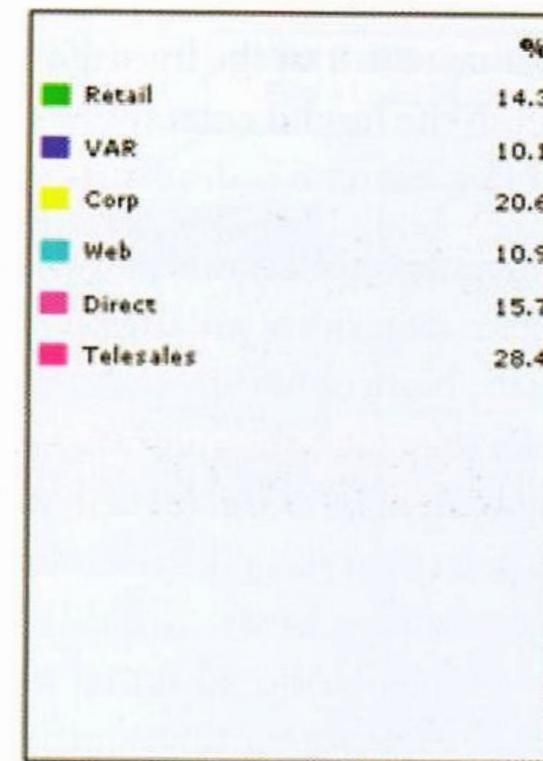


Figure 3-20. This pie chart uses colors for the slices that are too much alike to be clearly distinguished.

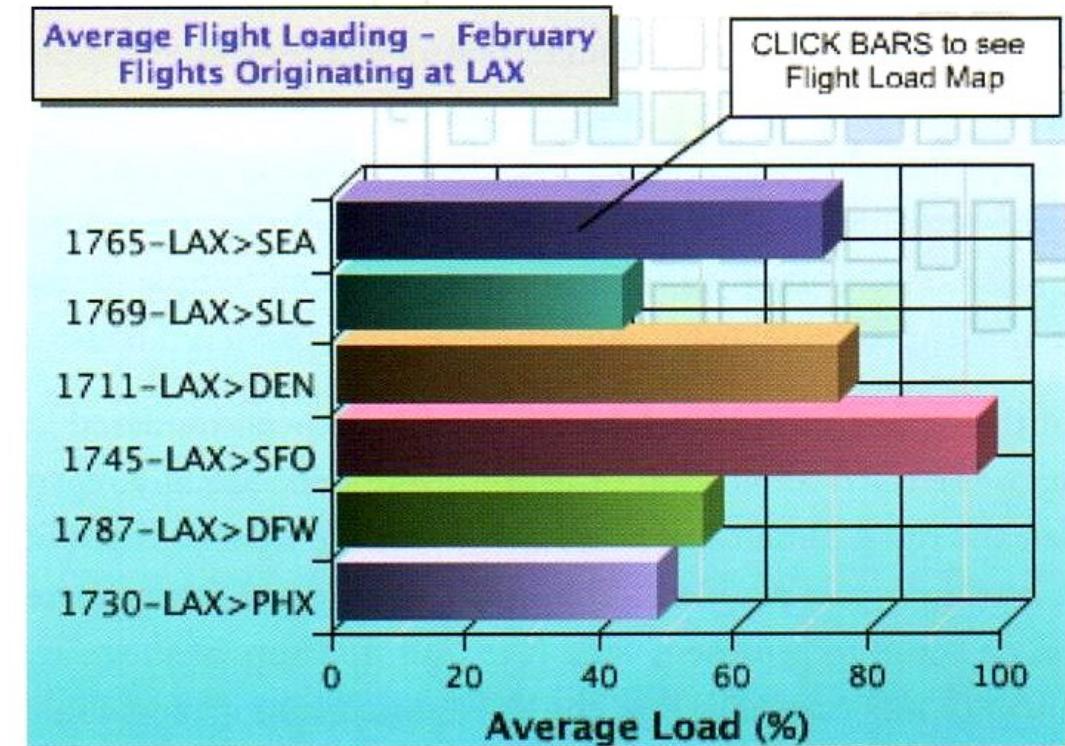


Figure 3-22. This bar graph, found on a dashboard, exhibits several design problems.



Figure 3-21. These dashboard meters have definitely taken the dashboard metaphor too far.

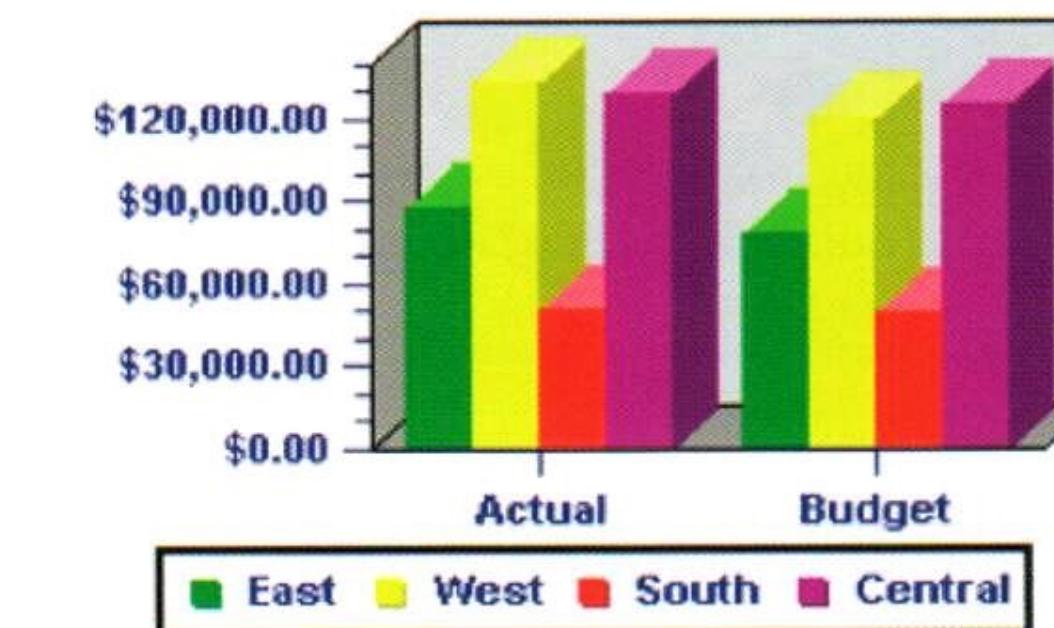


Figure 3-23. This bar graph, found on a dashboard, was poorly designed in a number of ways.

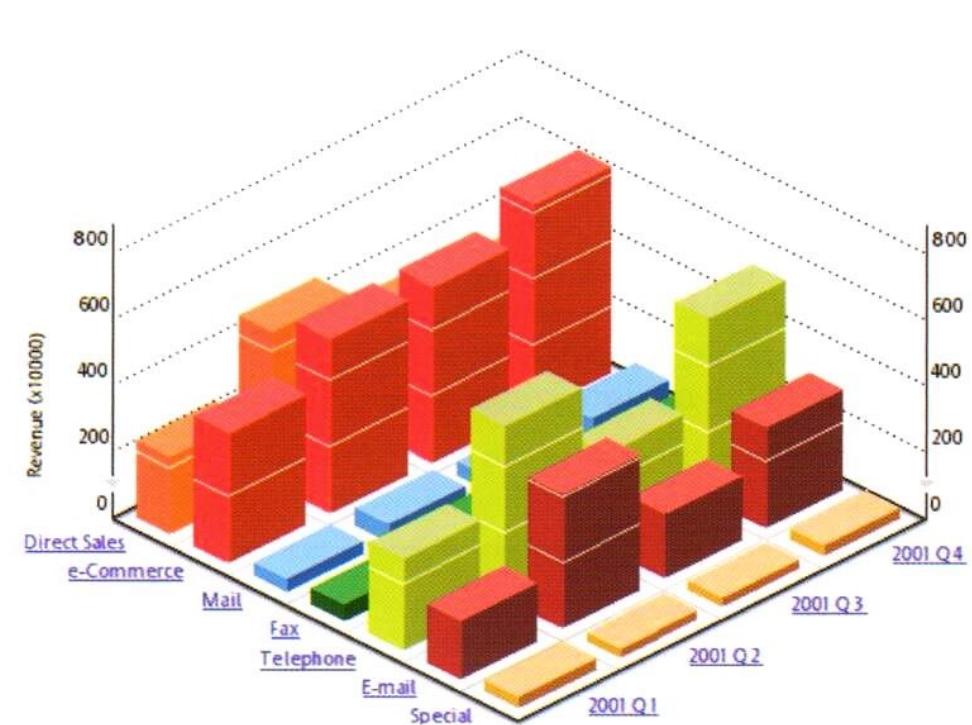


Figure 3-24. This 3-D bar graph illustrates the problem of occlusion.

POSICIONAMENTO POBRE DOS COMPONENTES

- Há várias técnicas para posicionamento dos componentes de forma a
 - Enfatizar elementos mais importantes
 - Permitir a comparação entre dados de componentes de forma eficaz

NAO REALÇAR OS DADOS IMPORTANTES DE FORMA EFETIVA

- É preciso usar atributos pré-atentivos para chamar a atenção do usuário para características importantes ou alarmantes

USO DE OBJETIVOS VISUAIS PURAMENTE DECORATIVOS

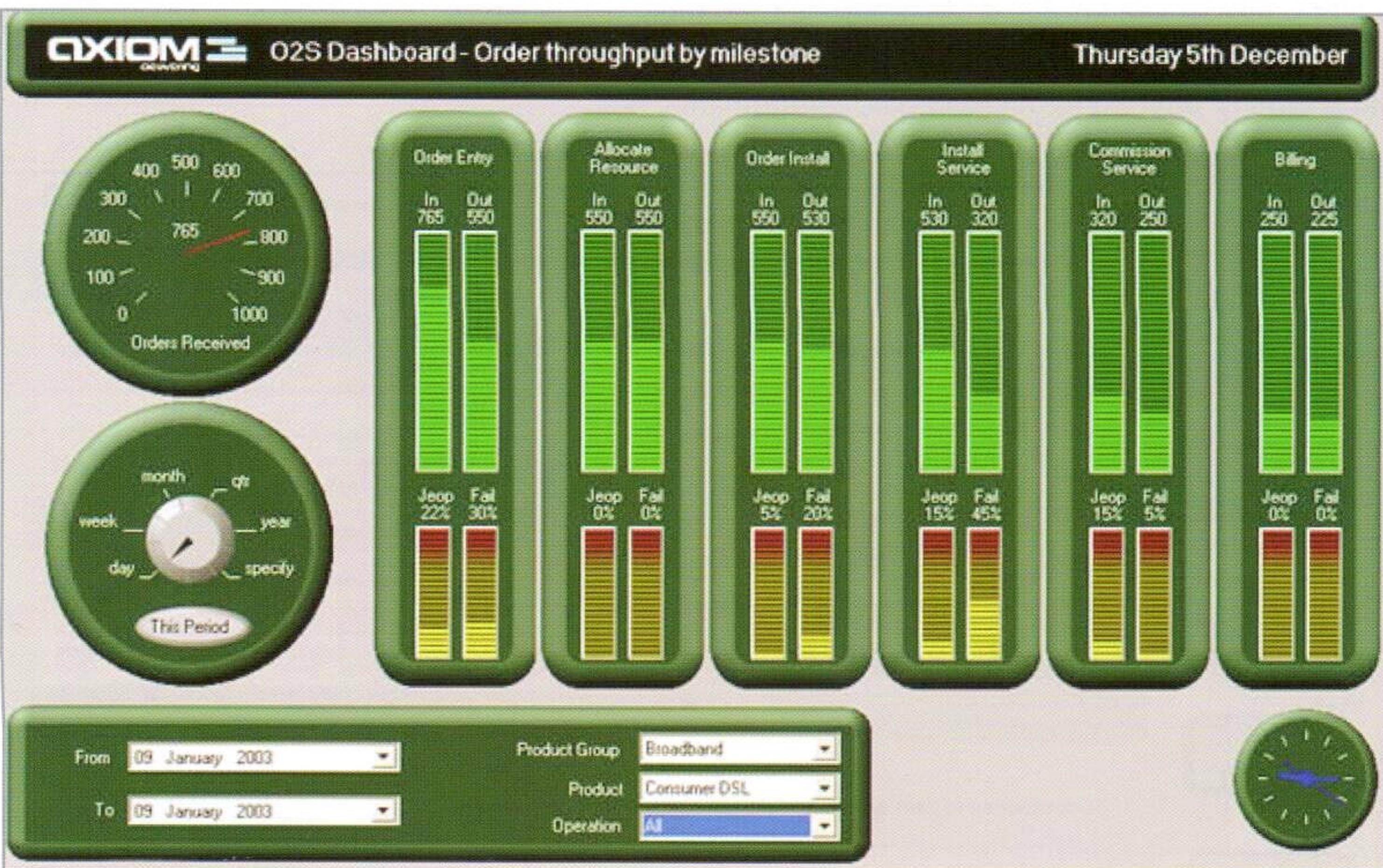


Figure 3-28. This dashboard is trying to look like something that it is not, resulting in useless and distracting decoration.

USO INCORRETO DAS CORES

- Assim como em qualquer visualização, o uso das cores deve ser adequado, especialmente para pessoas daltônicas

DESIGN NAO ATRAENTE

- A satisfação do usuário com relação a beleza de uma ferramenta é atualmente uma característica muito importante

FUNDAMENTOS

.....

LIMITAÇÕES DA MEMÓRIA

- Temos três tipos de memória:
 - Iônica (registro sensorial)
 - Curto prazo (memória de trabalho)
 - Longo prazo
- A memória de trabalho:
 - É temporária
 - Tem capacidade limitada
- O fato de toda a informação necessária estar disponível na tela reduz as limitações de memória

CODIFICAÇÃO DOS DADOS PARA PERCEPÇÃO

- Uso de atributos pré-atentivos
 - Cor
 - Intensidade
 - Posição
 - Forma
 - Movimento

PRINCIPIOS GESTALT

- Fechamento
- Similaridade
- Proximidade

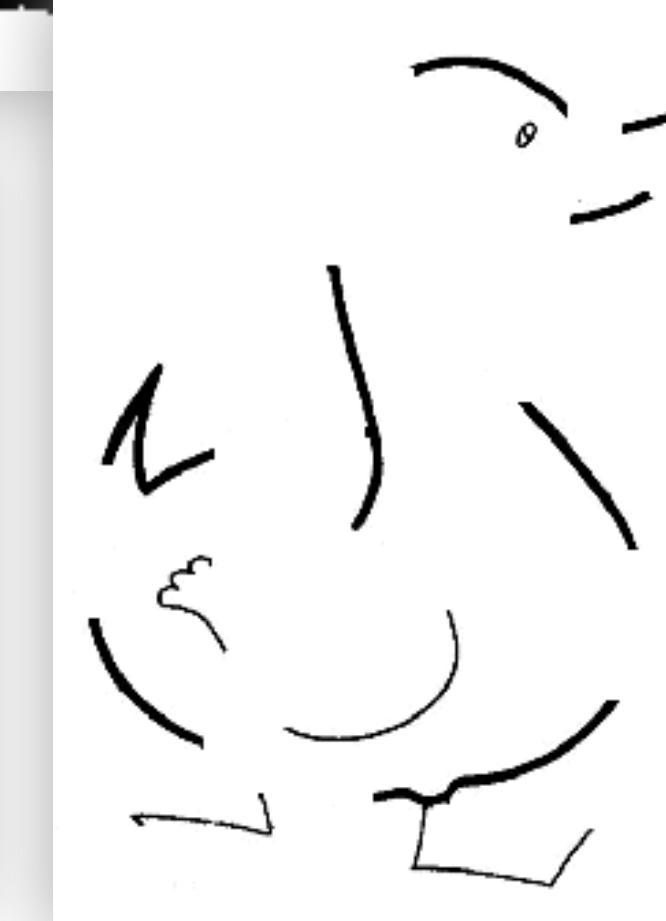
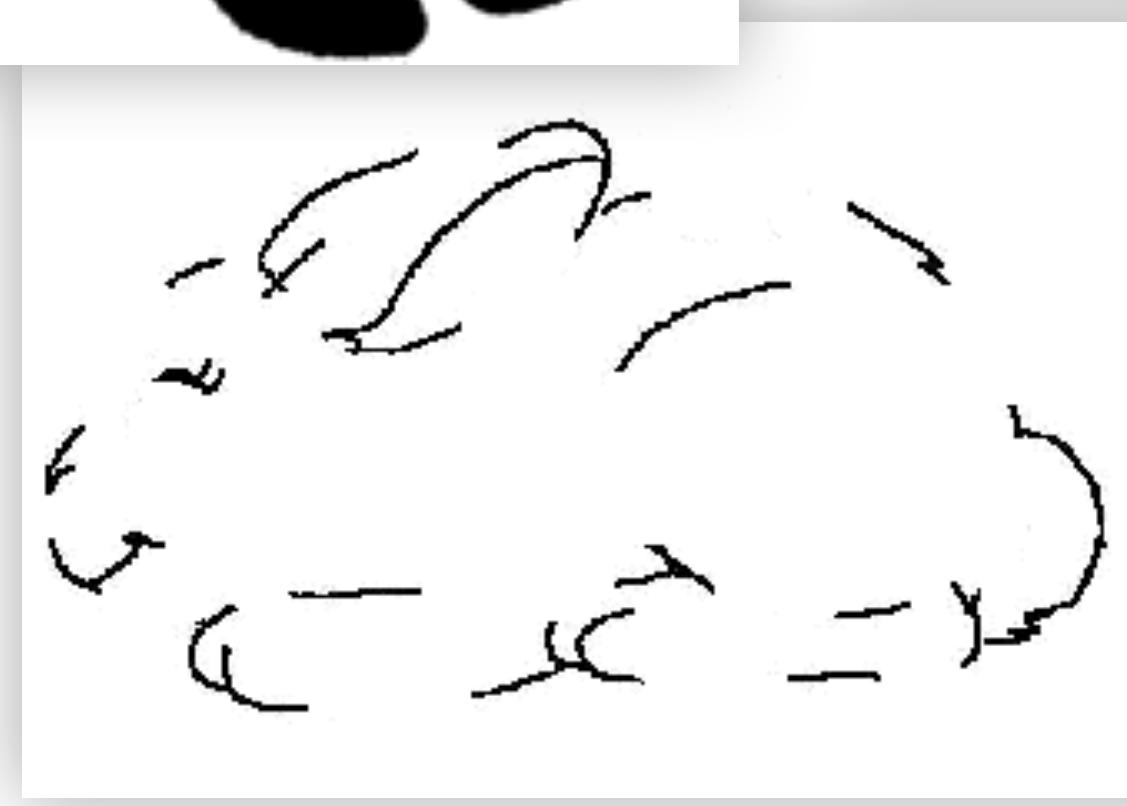
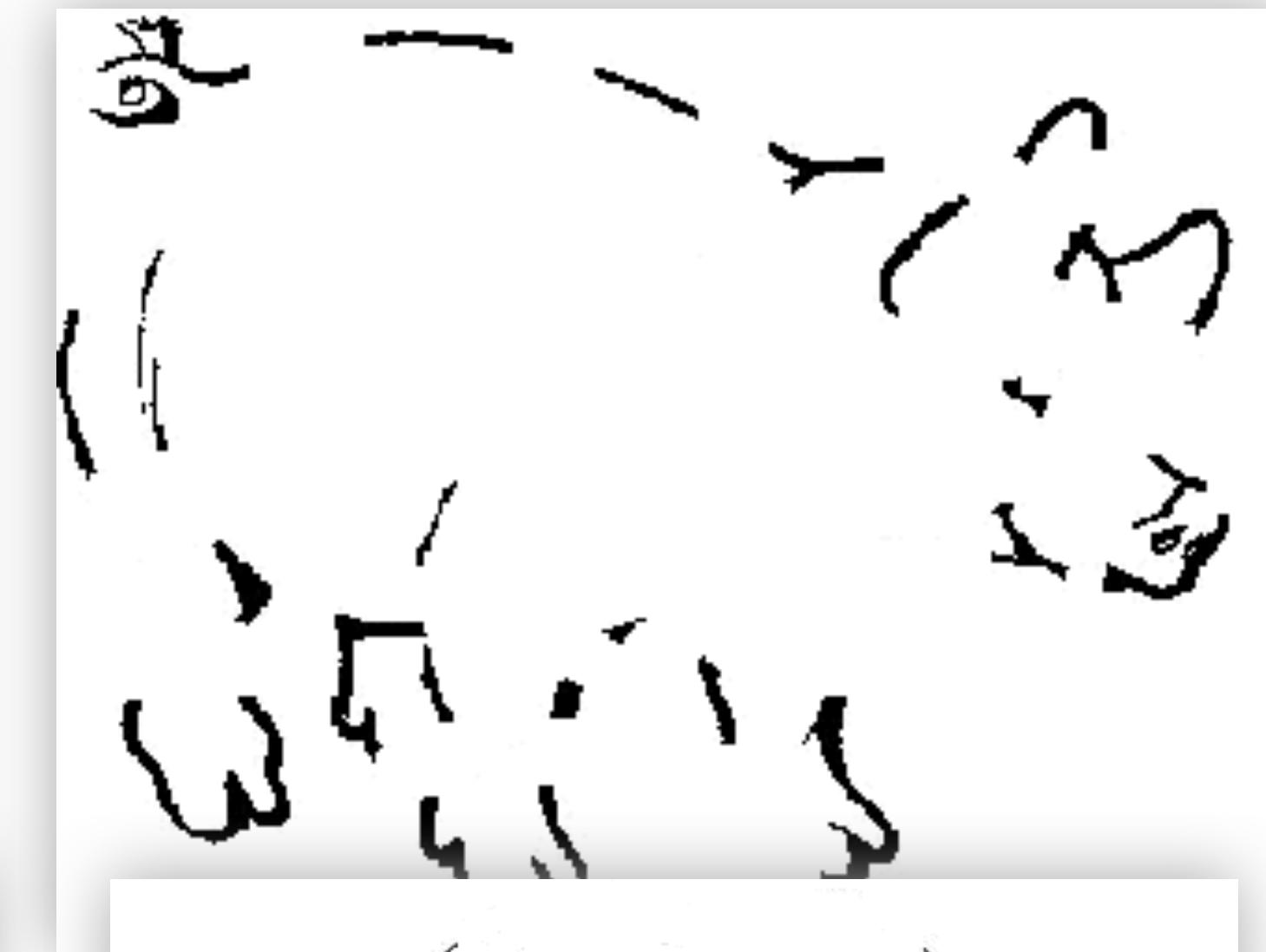
PRINCÍPIOS GESTALT

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

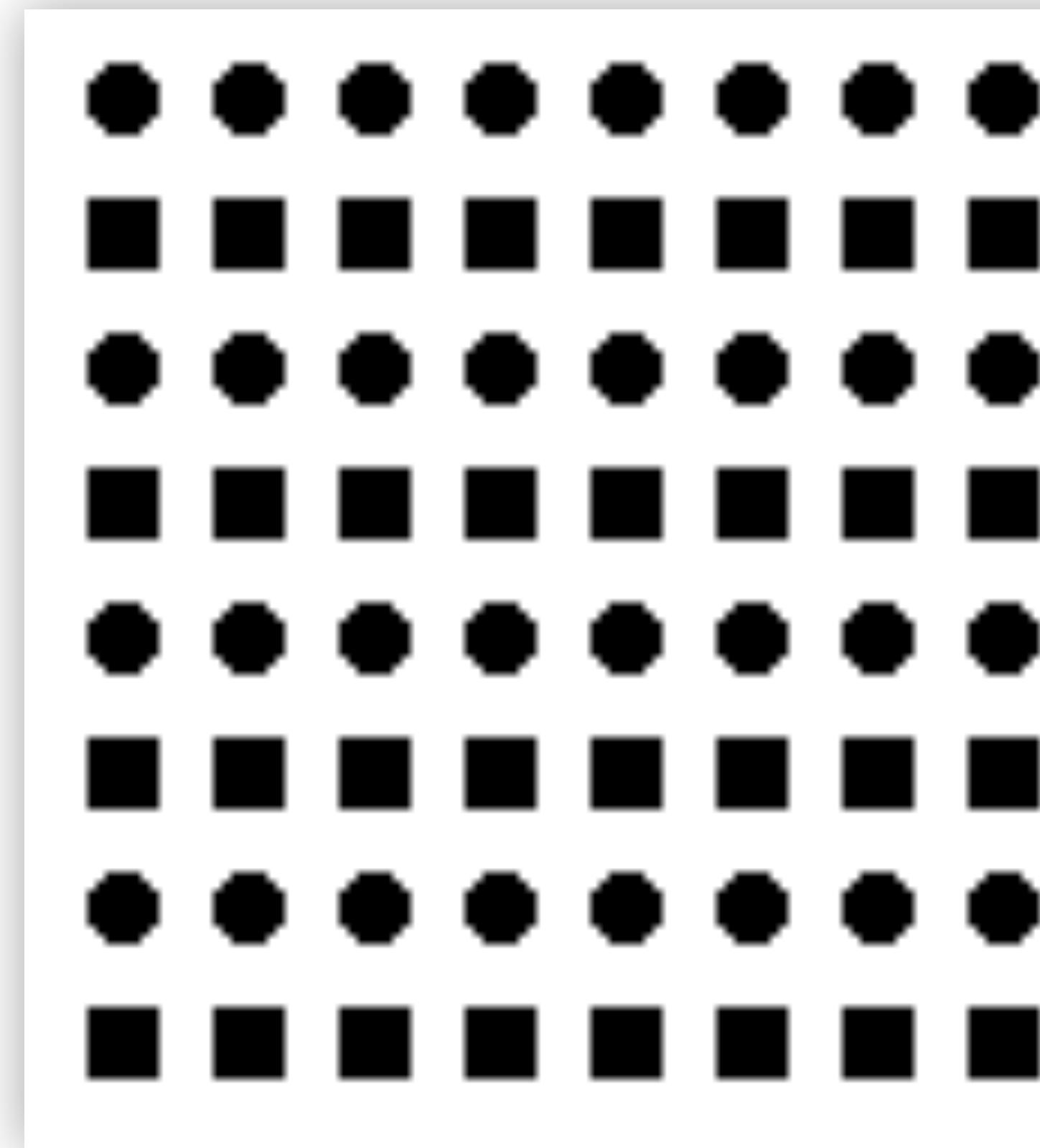
LEI DO FECHAMENTO

- 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



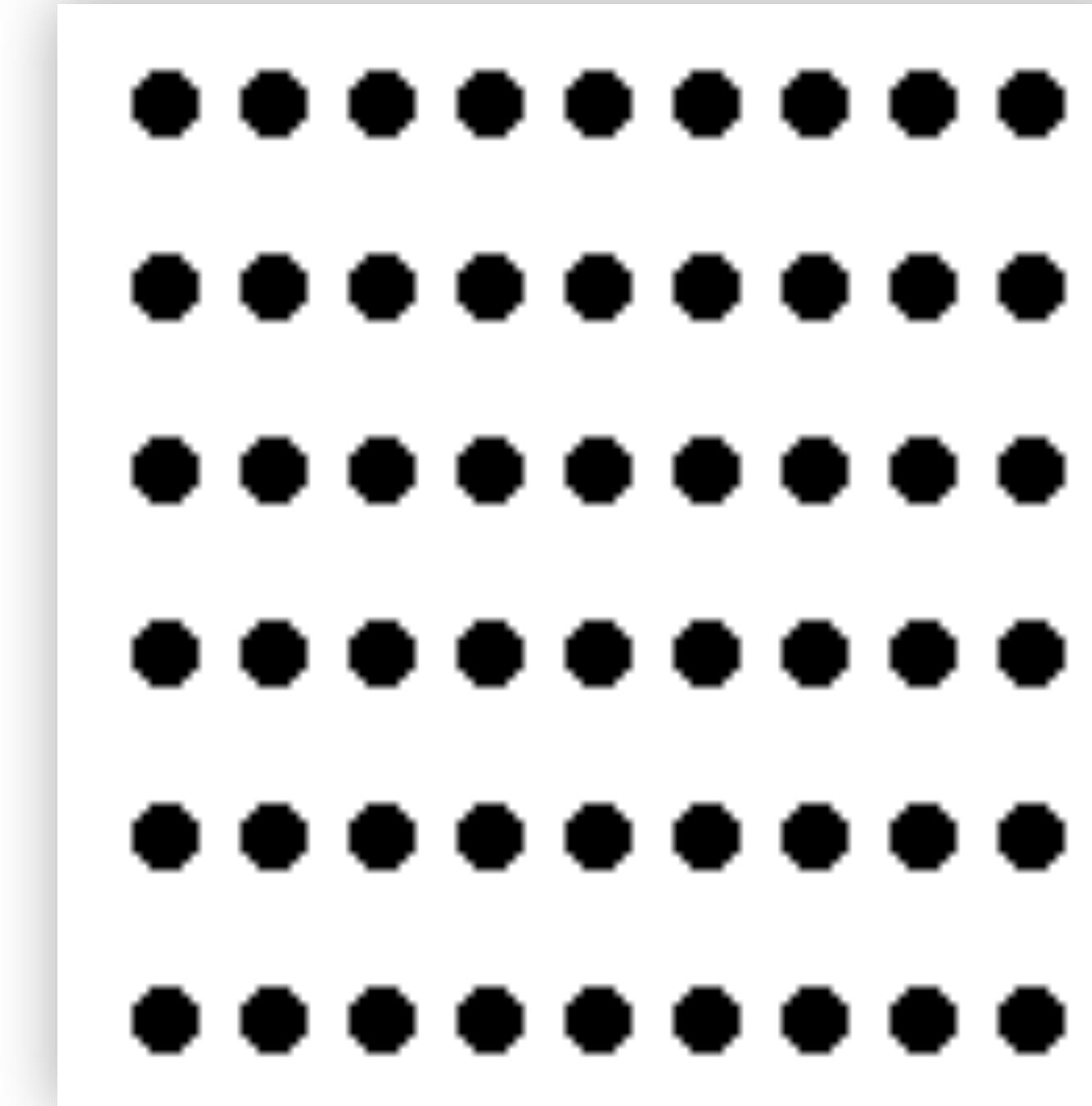
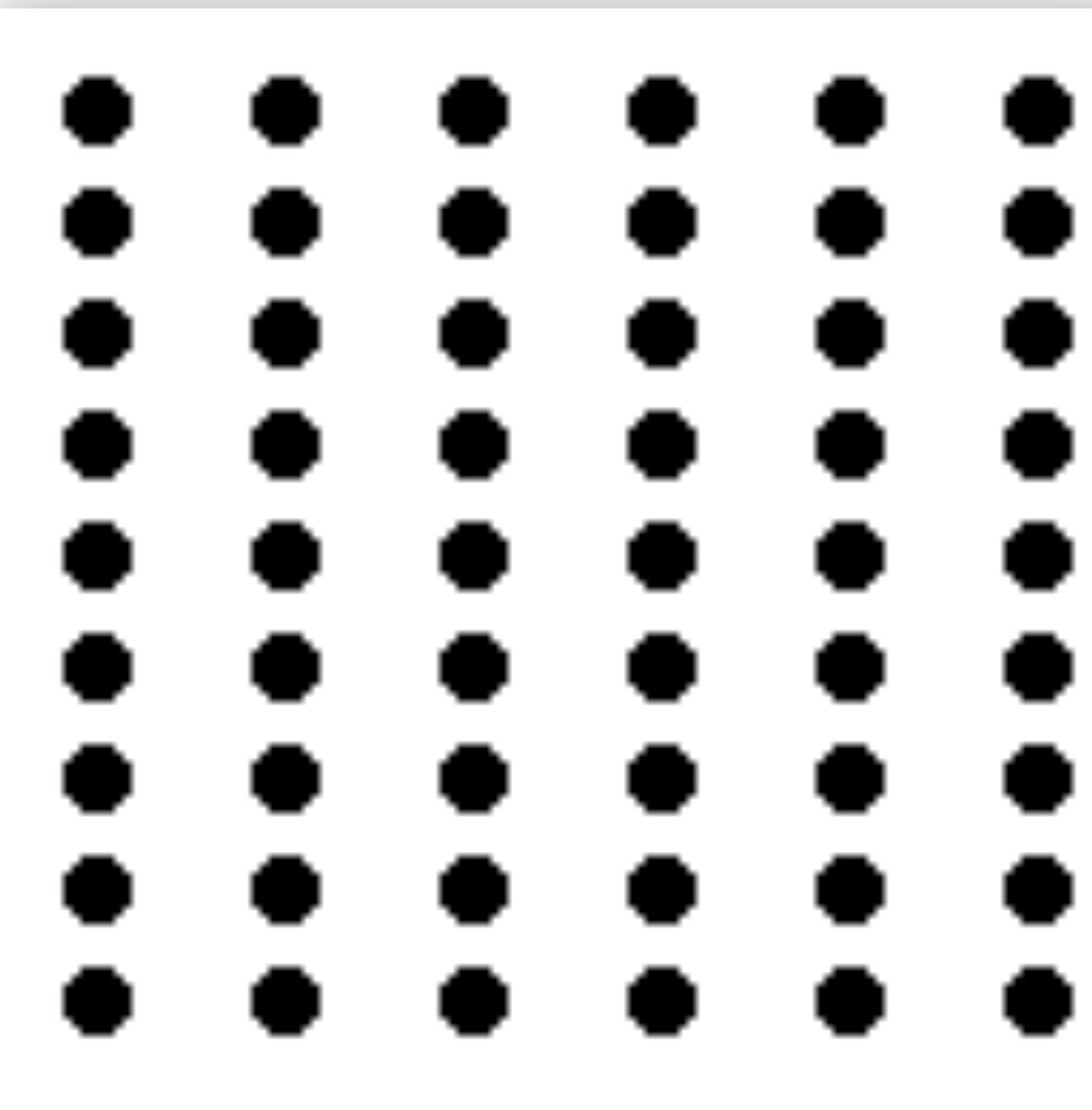
LEI DA PROXIMIDADE

- 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?



PRINCÍPIO DA RAZÃO DATA-INK

- Razão entre os pixels úteis que representam dados e os que não representam dados deve ser maximizada

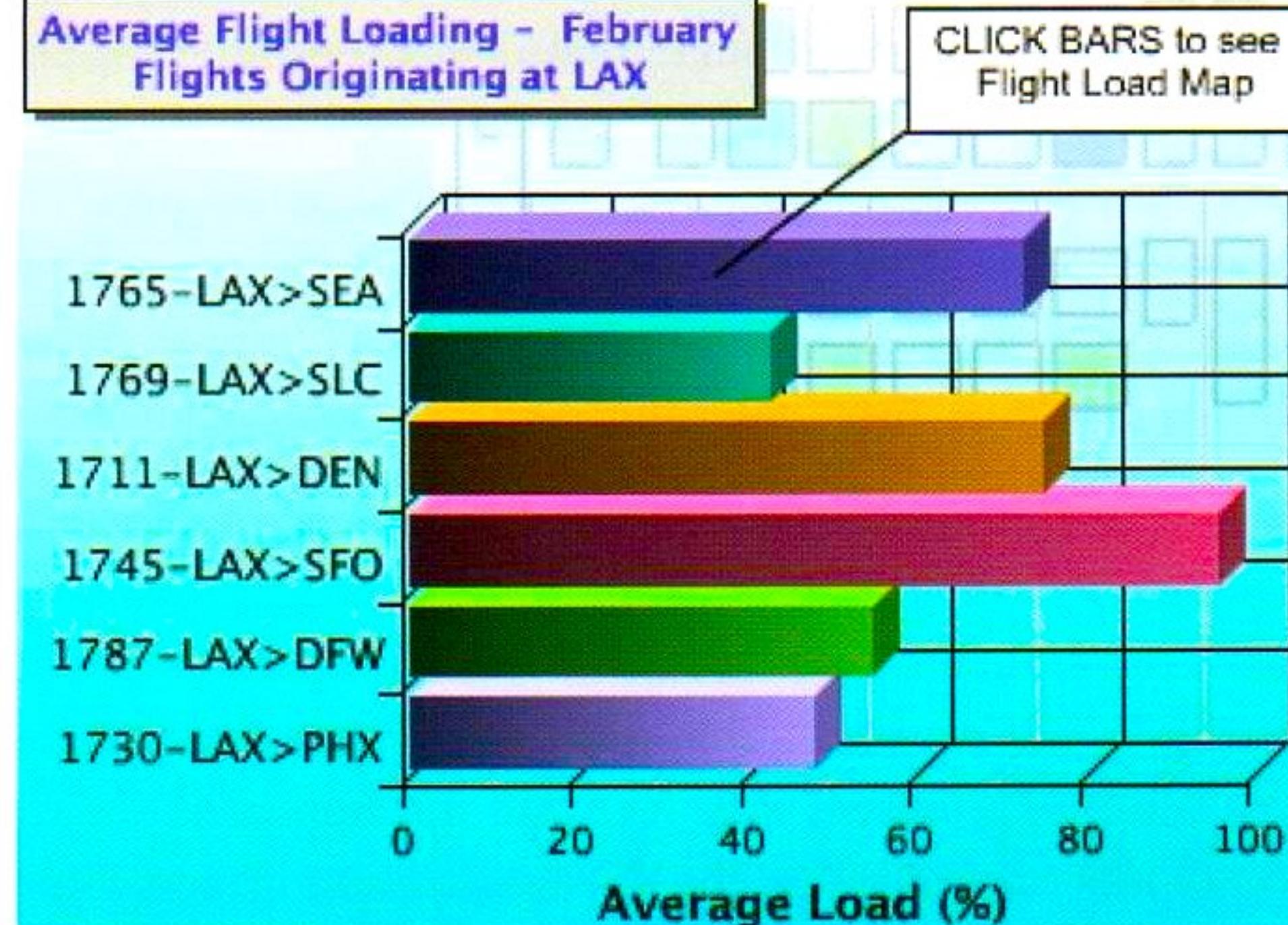


Figure 3-22. This bar graph, found on a dashboard, exhibits several design problems.

PRINCÍPIO DA RAZÃO DATA-INK

1. *Reduza os pixels úteis que não representem dados*

- A. *Elimine os pixels desnecessários*
- B. *Tire a ênfase e regularize os que restarem*

2. *Melhore os pixels que representam dados*

- A. *Elimine os pixels desnecessários*
- B. *Realce os mais importantes*

BIBLIOTECA IDEAL DE REPRESENTAÇÕES

Gráficos

BULLET CHARTS

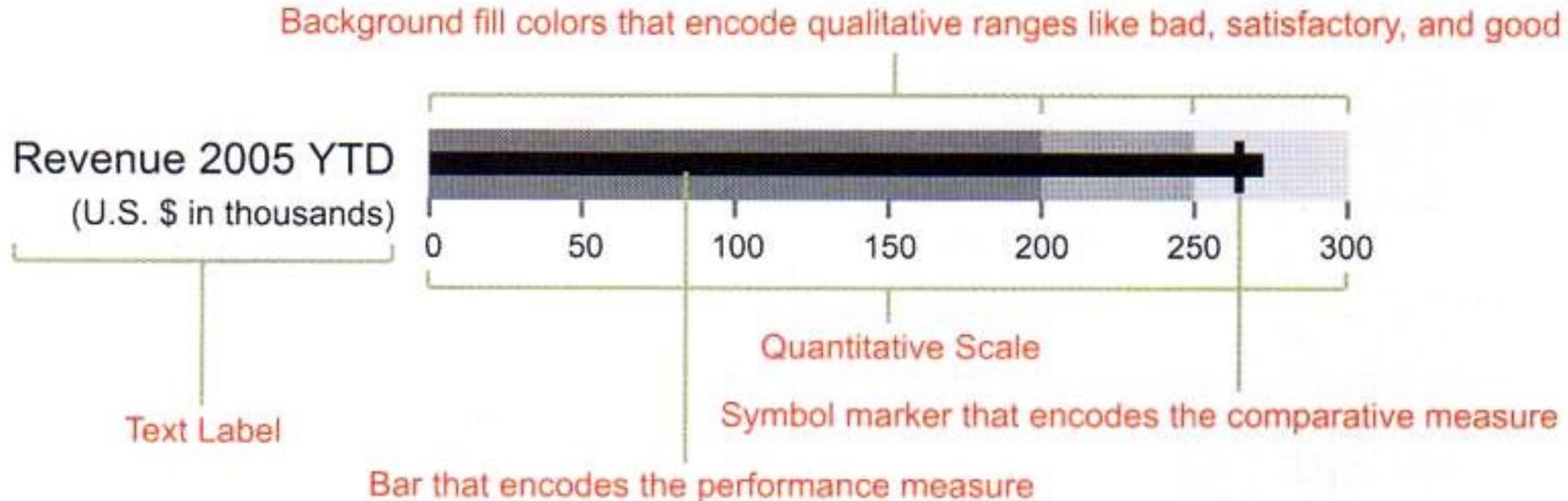


Figure 6-7. A simple bullet graph with each of its components labeled.

2005 YTD

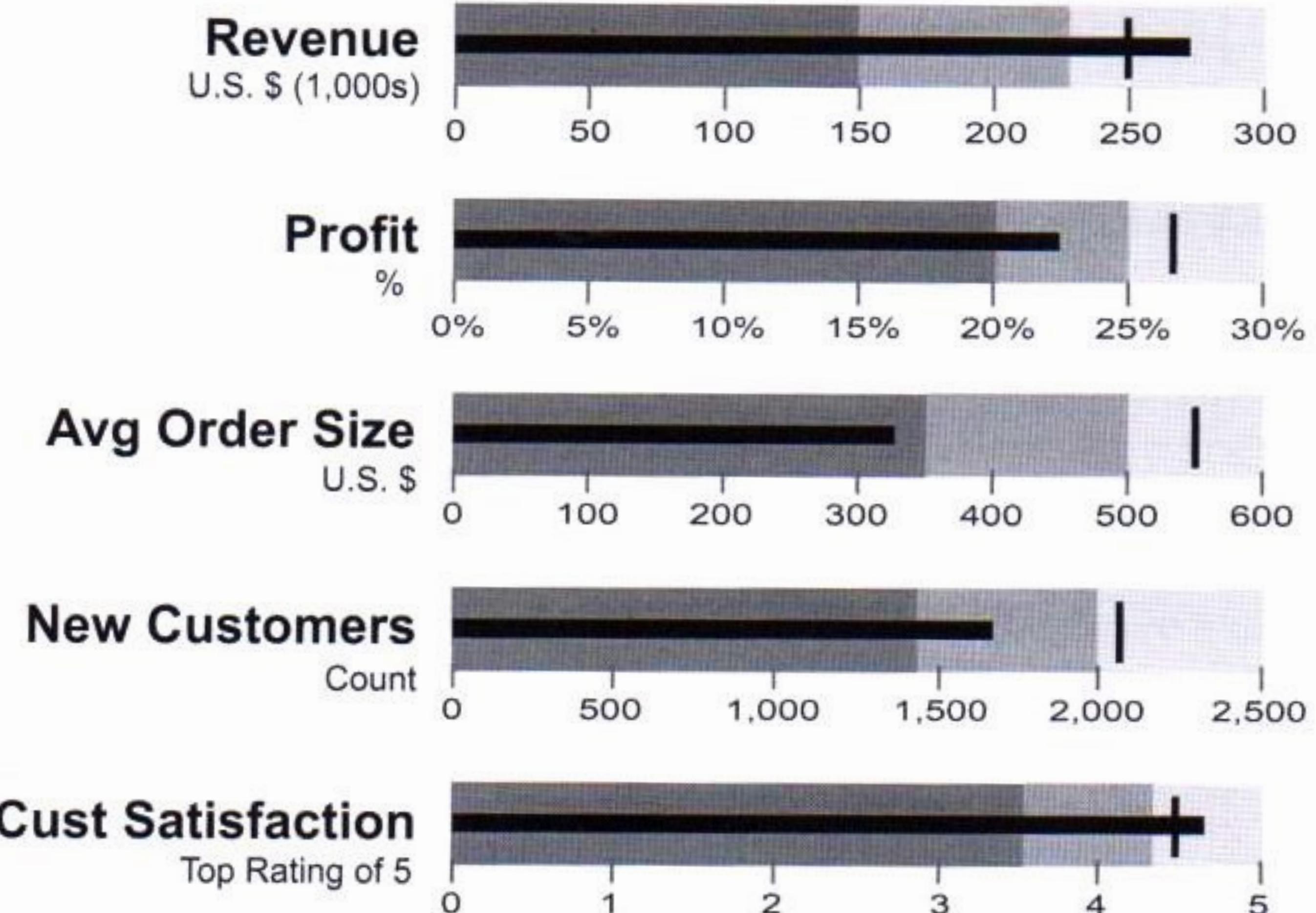


Figure 6-8. A collection of horizontally oriented bullet graphs.

GRÁFICO DE BARRAS

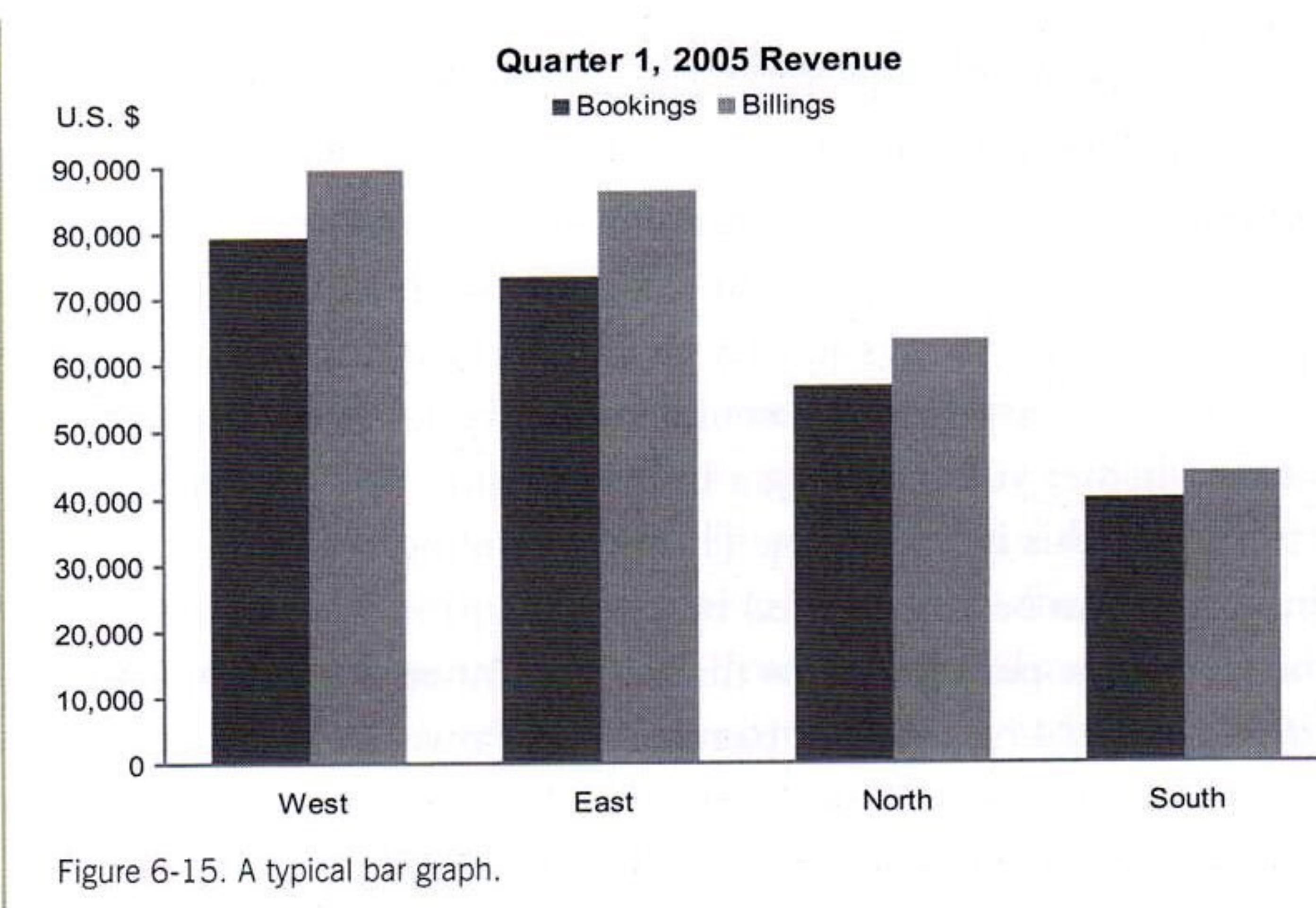


Figure 6-15. A typical bar graph.

GRÁFICO DE BARRAS EMPILHADAS

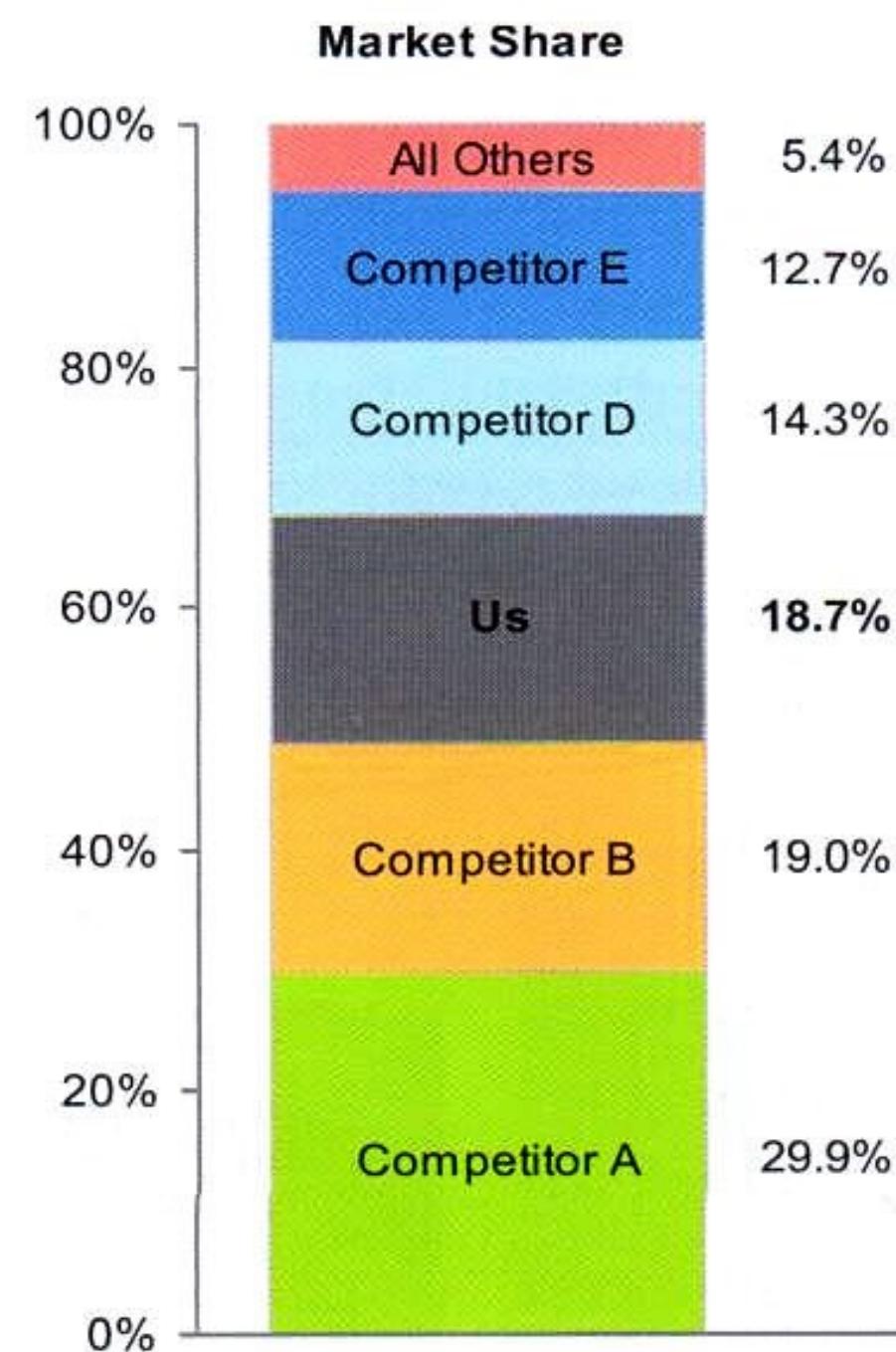


Figure 6-23. A stacked bar graph is not the best way to display a single series of part-to-whole data.

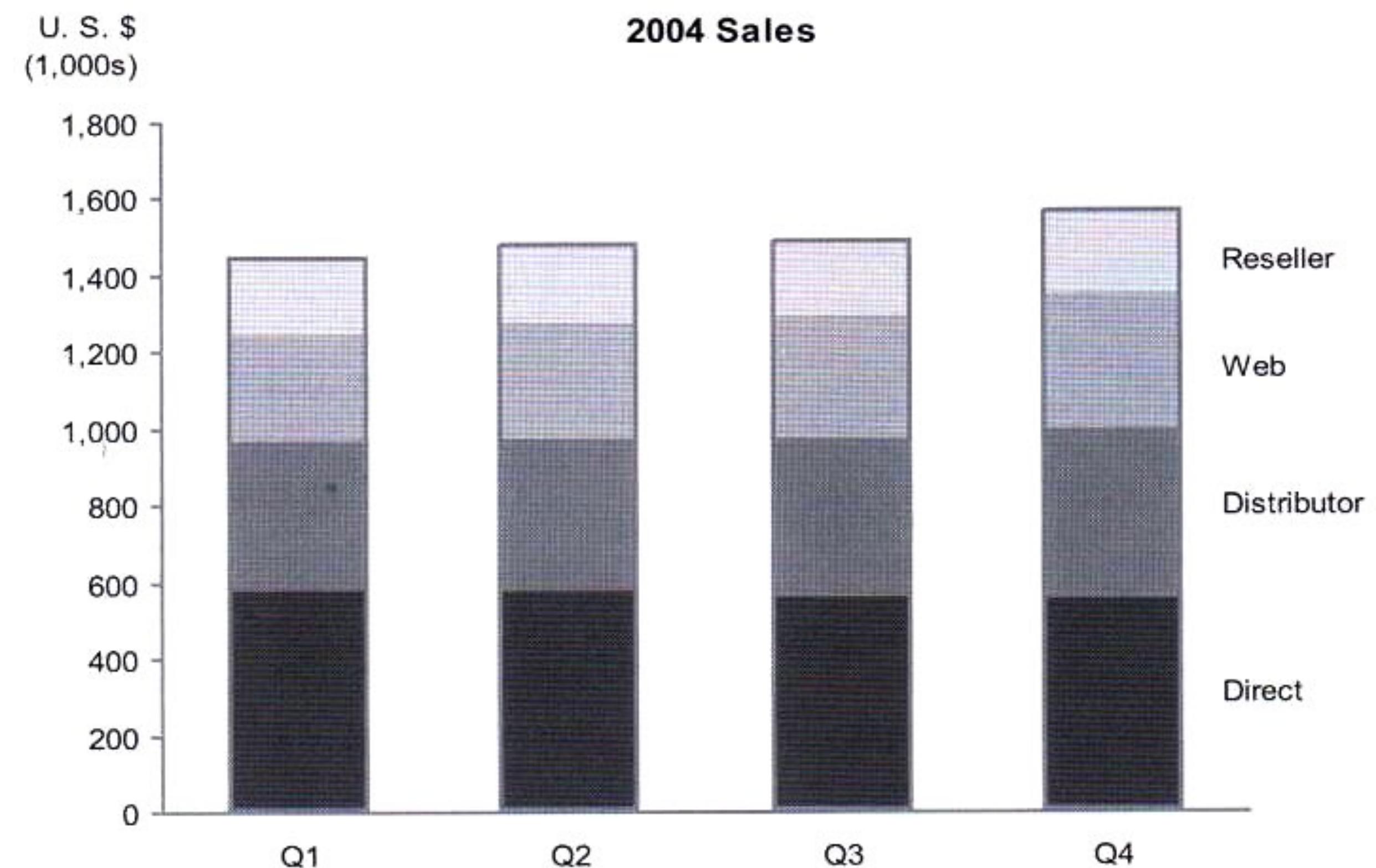


Figure 6-24. The only circumstance when a stacked bar graph is useful is when you must display multiple instances (for example, one for each quarter) of a whole (total sales) and its parts (in this case, per sales channel), with a greater emphasis on the whole than the parts.

GRÁFICO DE BARRAS E LINHA

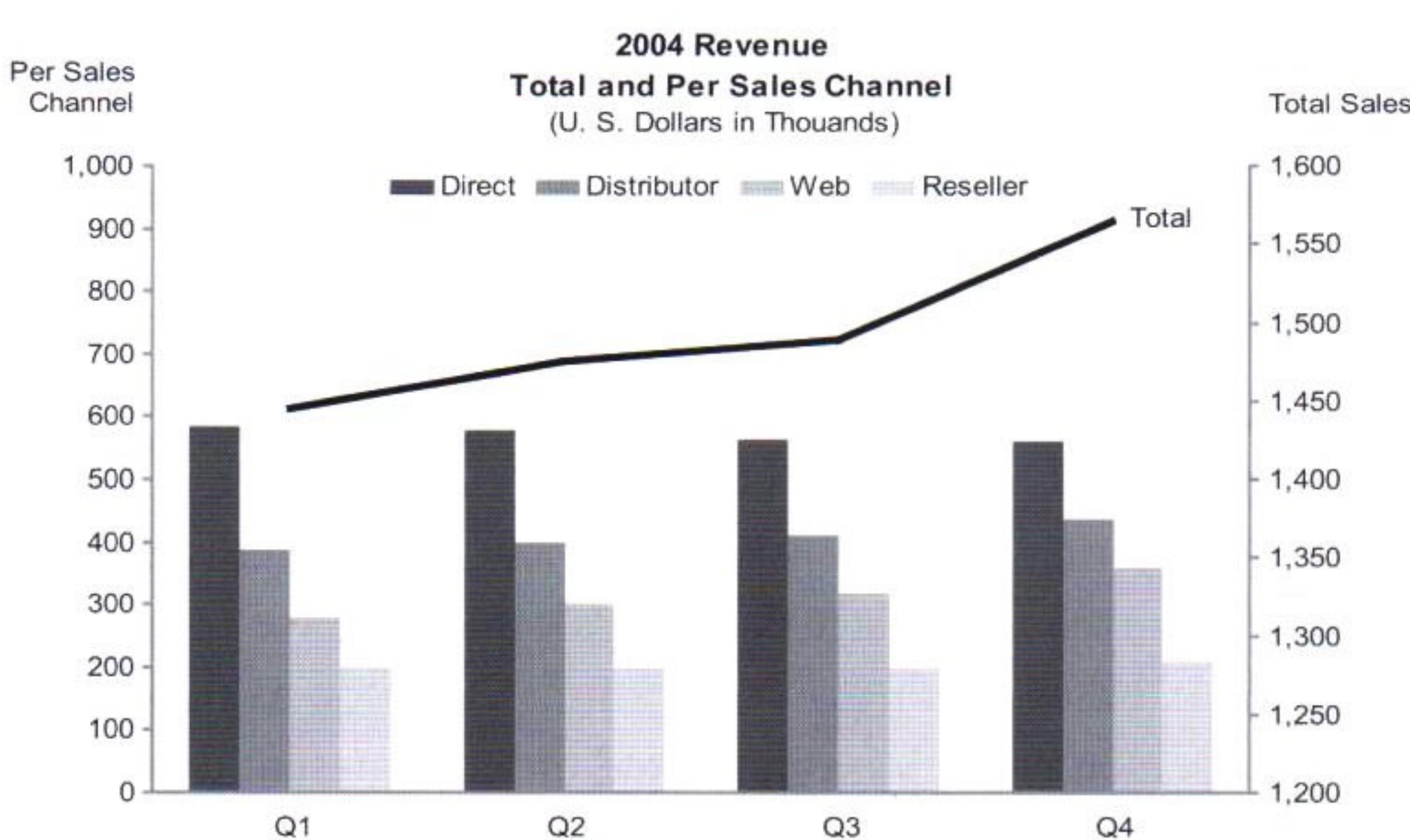


Figure 6-27. Example of a combination bar and line graph that displays quarterly instances of revenue by sales channel, encoded as bars, and total revenue, encoded as a line.

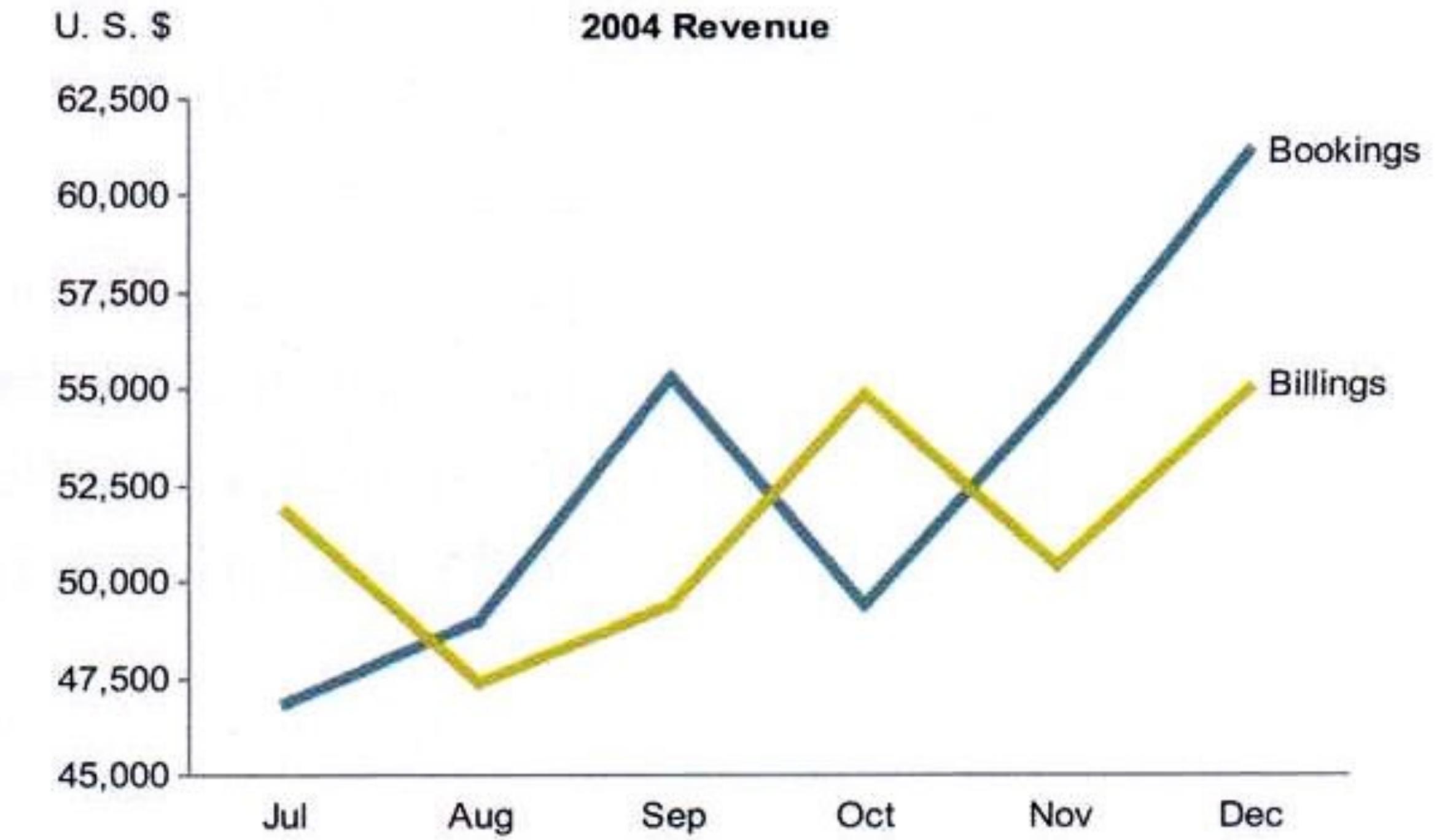
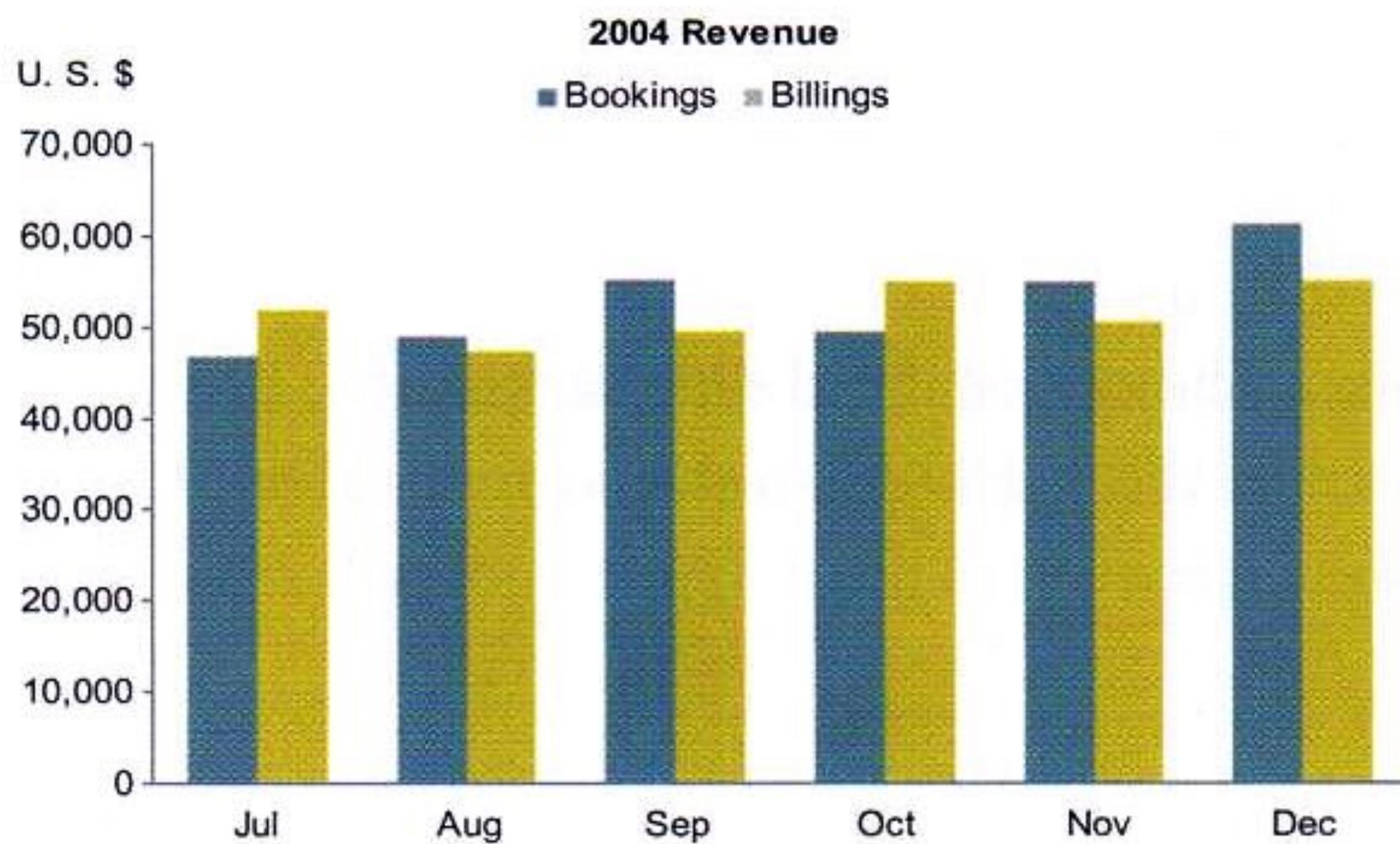


Figure 6-29. Two graphs of the same time-series data: a bar graph on the left and a line graph on the right. Notice how the overall shape of the data is much easier to see in the line graph.

SPARKLINES



\$137,384.28 Checking Balance 05/20/2005

Figure 6-30. A simple sparkline that displays the 12-month history of a checking account balance.

BOXPLOTS

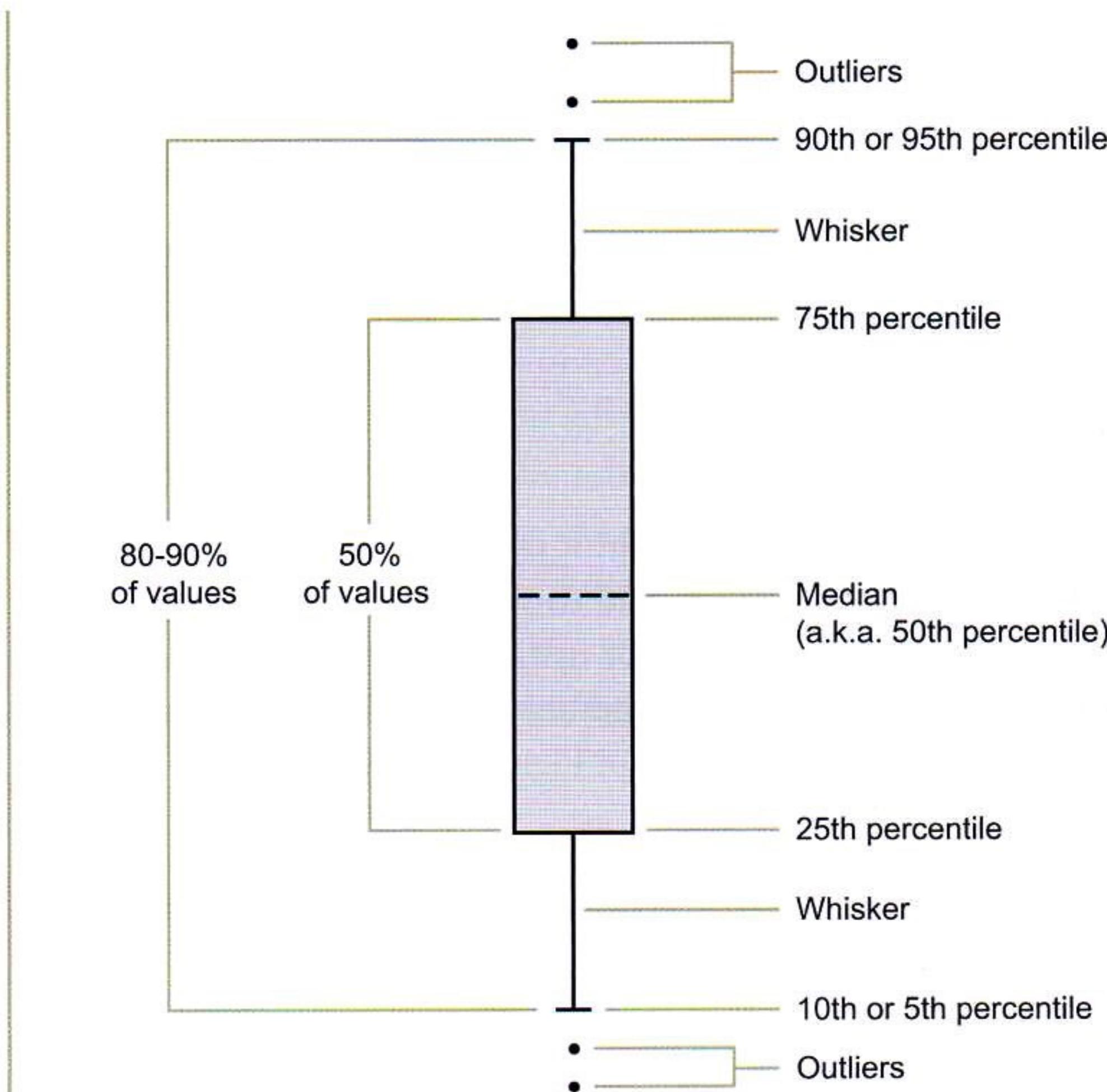


Figure 6-38. An individual box plot with whiskers. Outliers are individual data values that fall outside the range that is defined by the whiskers.

GRÁFICO DE DISPERSÃO

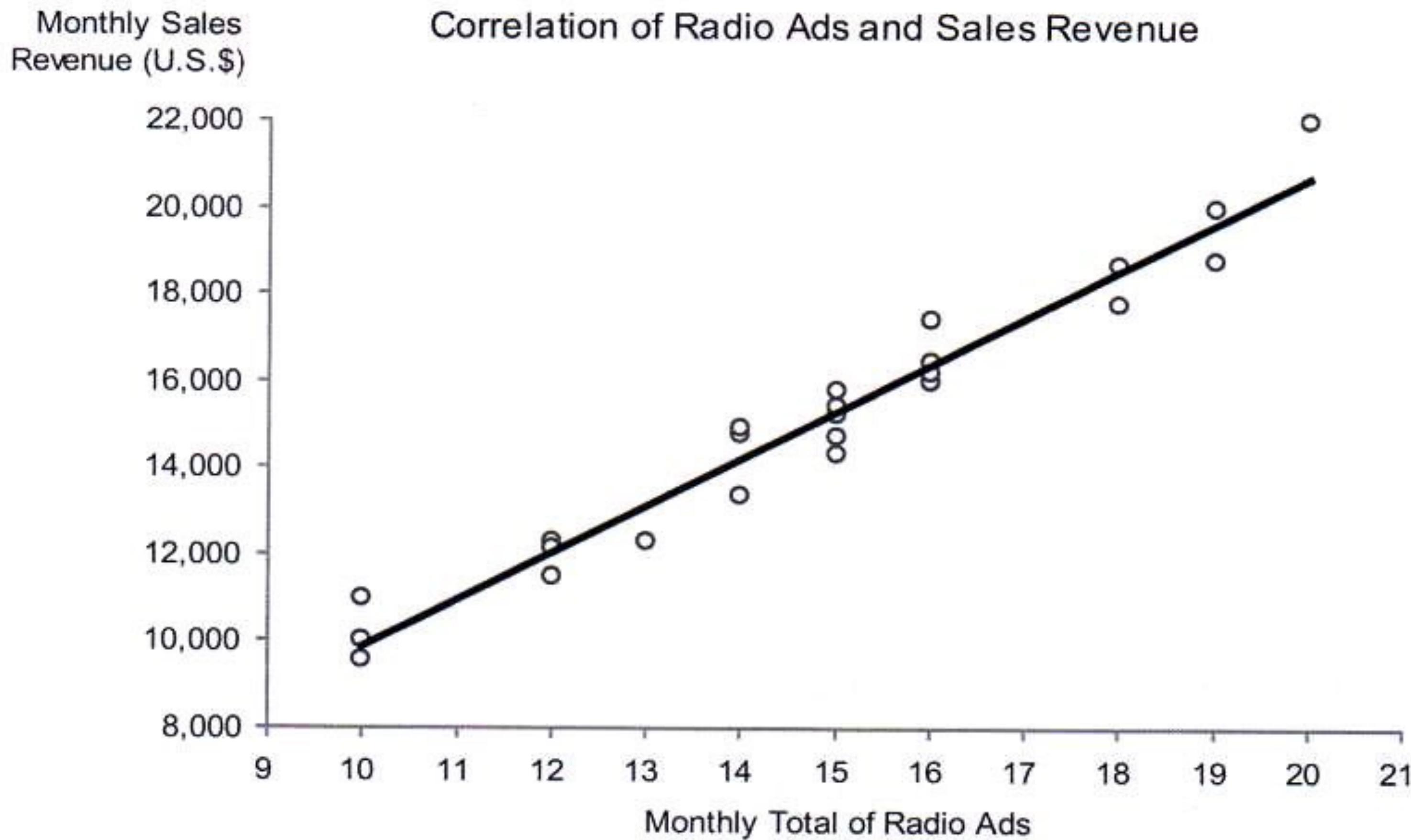


Figure 6-40. This scatter plot displays the correlation between the number of broadcast ads and the amount of sales revenue for 24 months.

TREEMAP

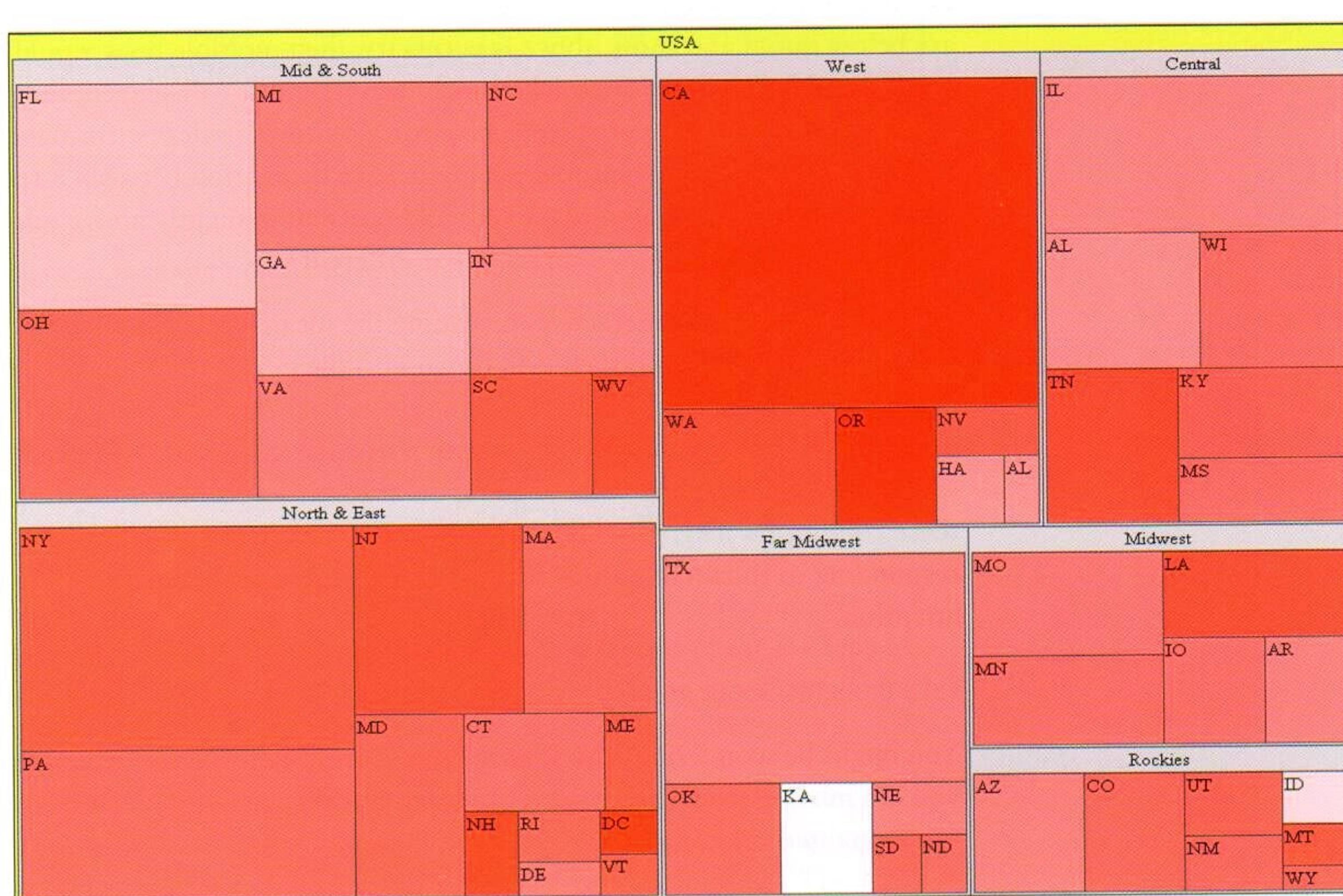
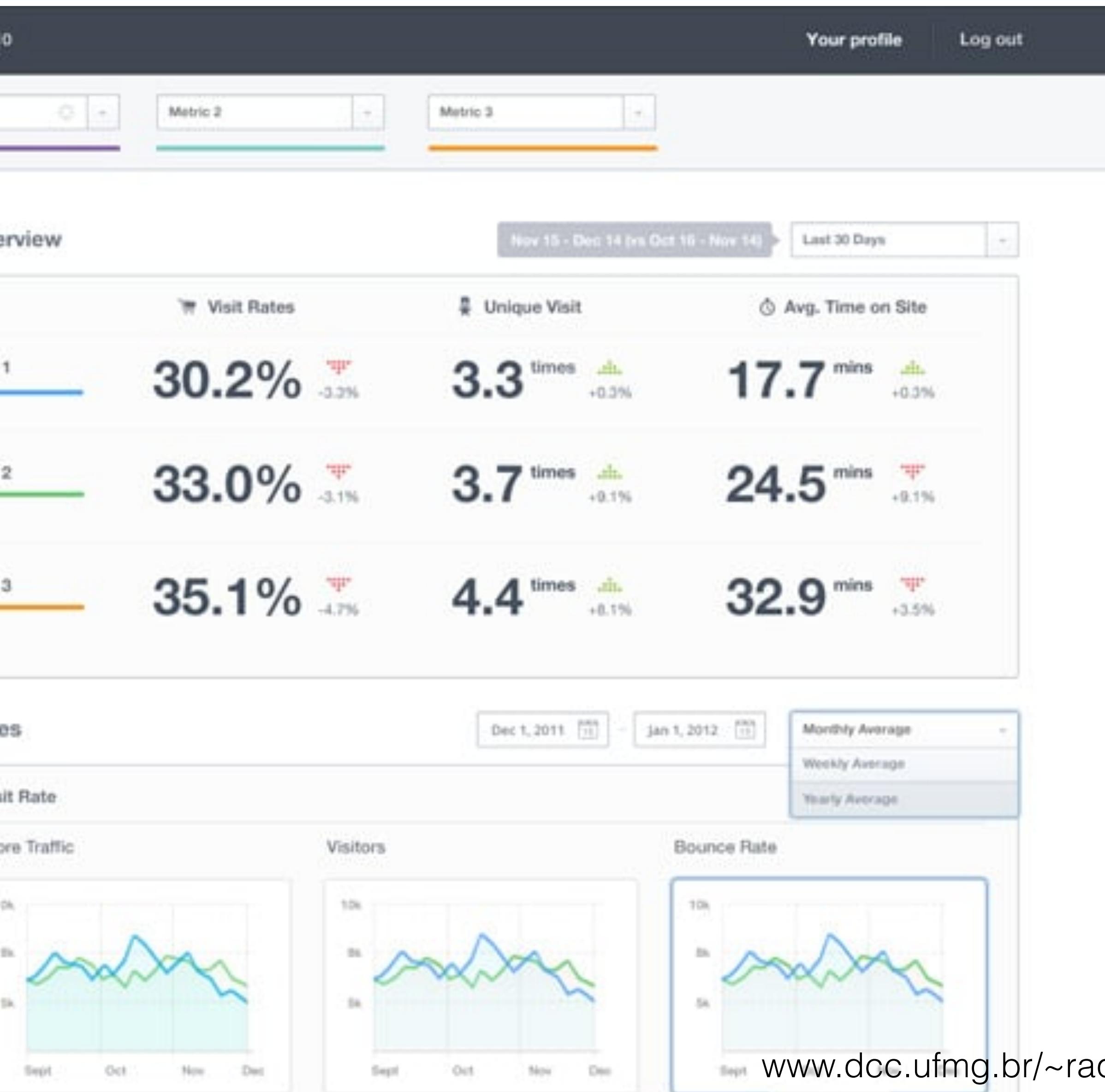


Figure 6-43. This treemap, created using Treemap 4.3 software developed at the University of Maryland's Human-Computer Interaction Lab (HCIL), displays sales data (revenue and percentage of quota) by region.

BIBLIOTECA IDEAL DE REPRESENTAÇÕES

Ícones

ÍCONES



- De alerta
- De altos e baixos
- De ligado e desligado

BIBLIOTECA IDEAL DE REPRESENTAÇÕES

Texto e organizadores

TEXTO E ORGANIZADORES

- Tabelas
- Mapas
- Pequenos múltiplos

TEXTO E ORGANIZAÇÃO

- Organize os grupos de acordo com as funções dos componentes ou seus tipos ou uso
- Posicione proximamente componentes pertencentes ao mesmo grupo
- Delineie os grupos de forma discreta
- Suporte comparações úteis e desestimule comparações sem sentido

ESTUDOS DE CASO

VENDAS

- Renda
- Lucros
- Satisfação dos clientes
- Principais clientes
- Market share

Sales Dashboard

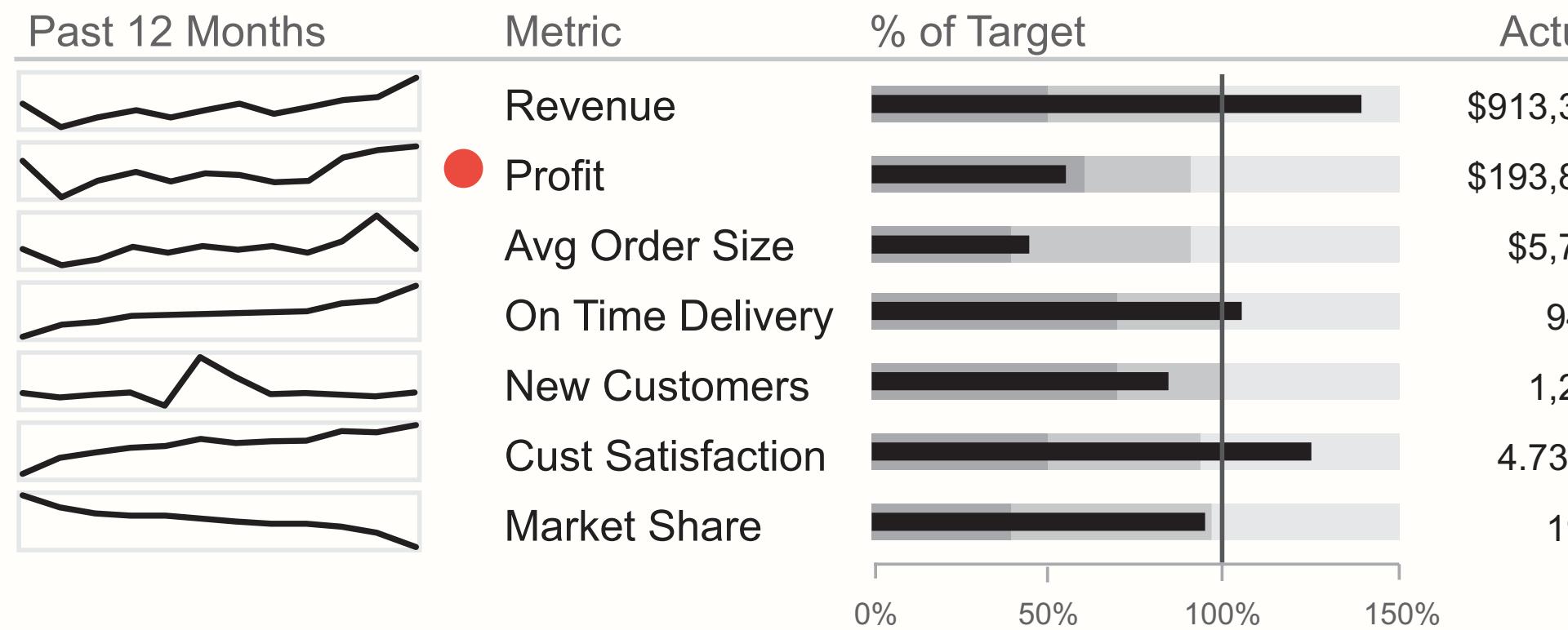
(Data as of December 19, 2012)

Actual | Target

Help

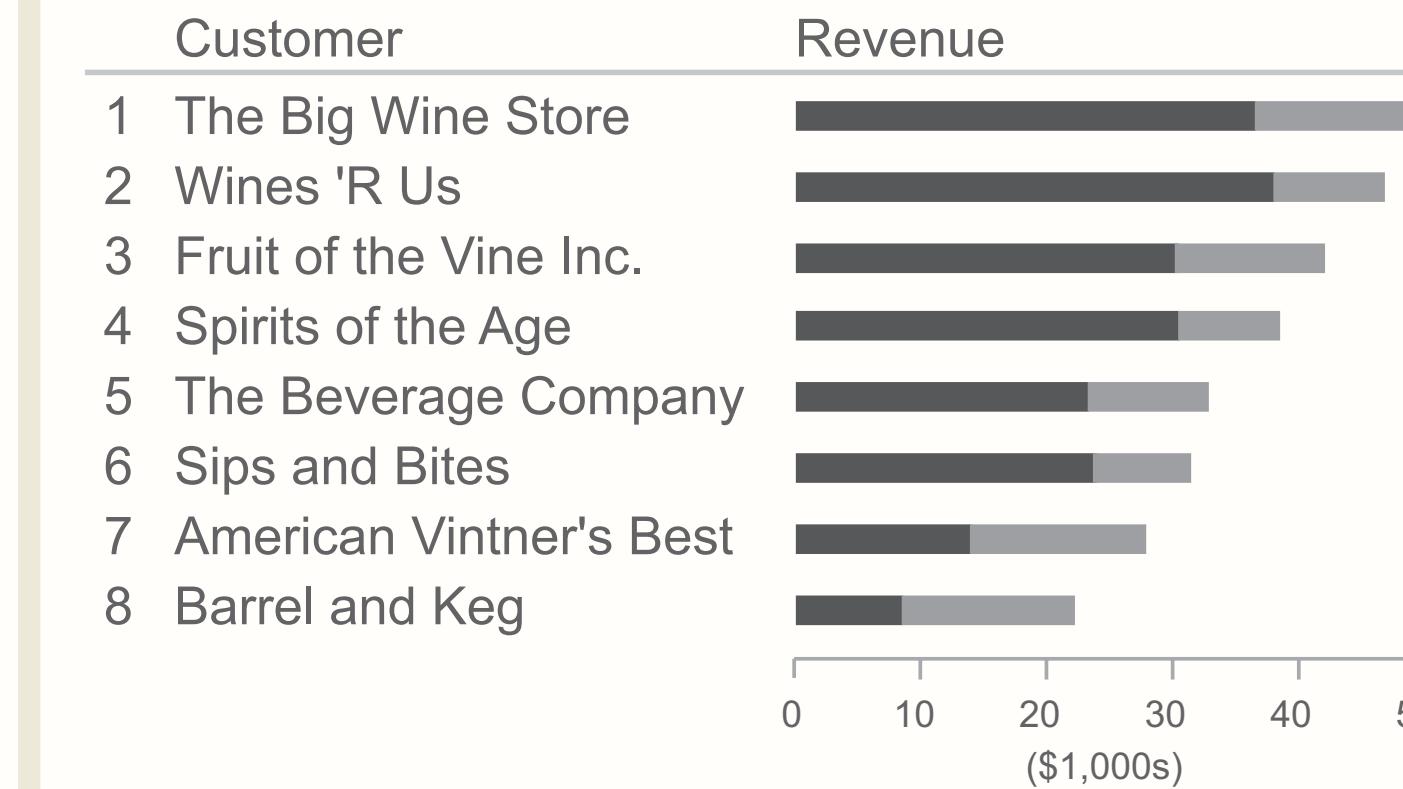
Key Metrics YTD

Poor Satisfactory Good



Top 8 Customers This Quarter

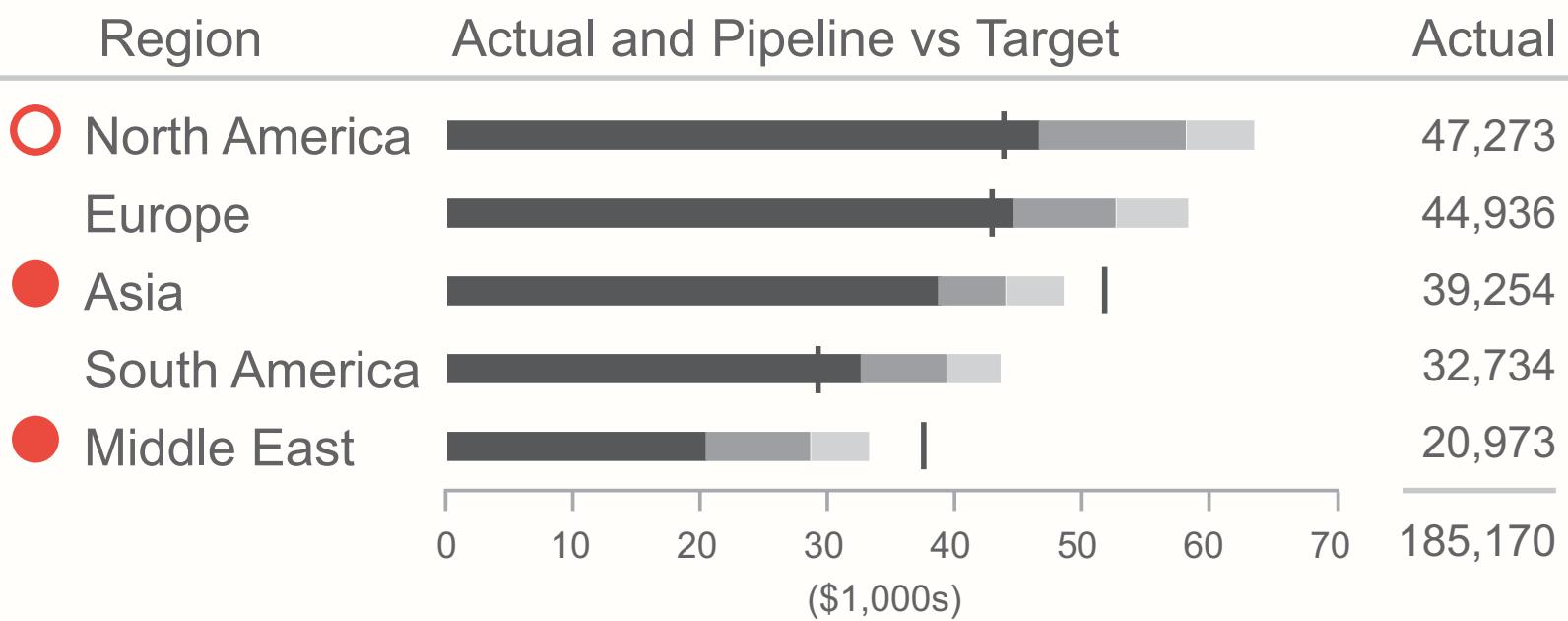
Pipeline



0 10 20 30 40 50
(\$1,000s)

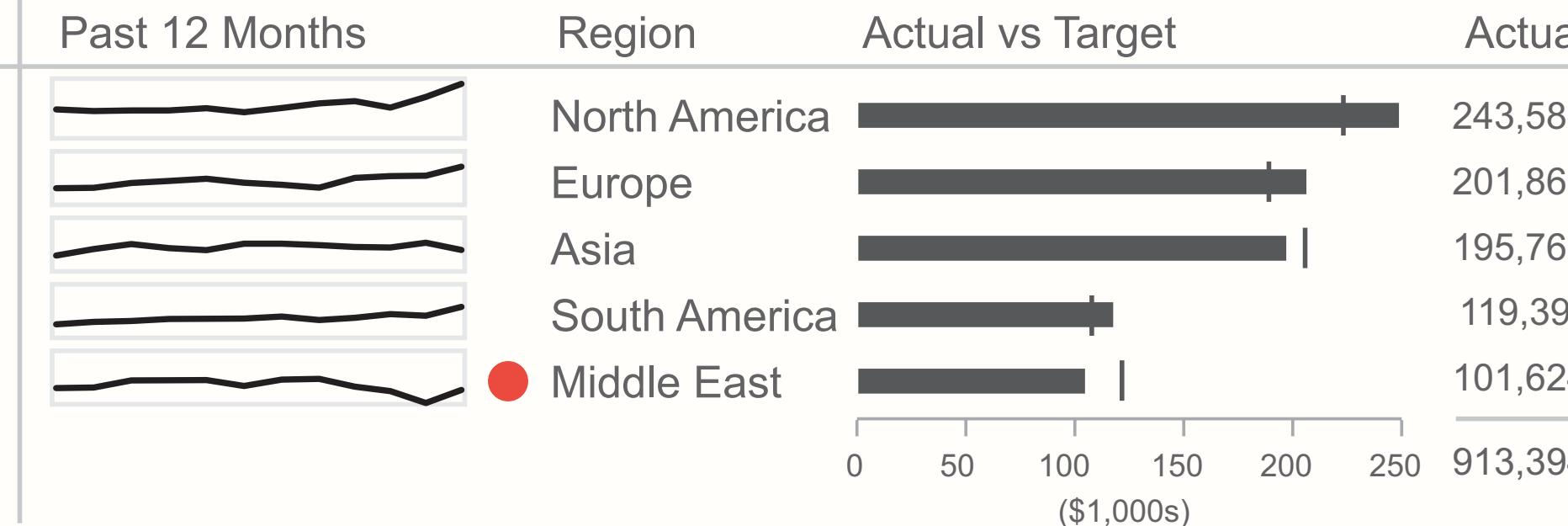
Revenue QTD

Pipeline: 90% 75%



0 10 20 30 40 50 60 70
(\$1,000s)

Revenue YTD



0 50 100 150 200 250
(\$1,000s)

Product Sales YTD



0 5 10 15 20 25
(1,000s)

0 25 50 75 100 125 150
(\$1,000s)

Market Share

1 month ago |



0% 5% 10% 15% 20% 25%

VENDAS: PONTOS FORTES

- Cores discretas
- Posicionamento adequado
- Densidade de informações
- Exibição gráfica e textual
- Espaço em branco usado para separar componentes

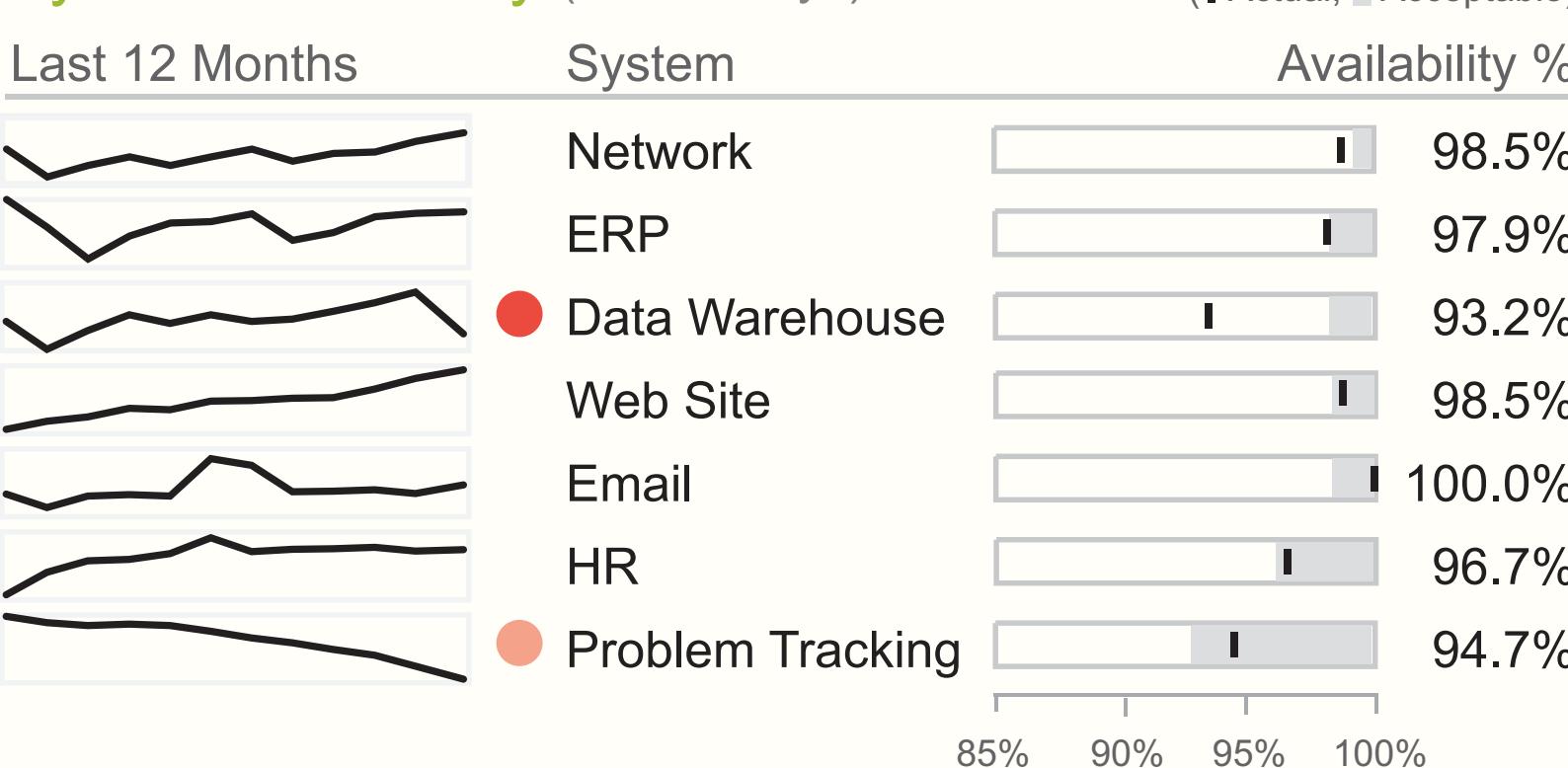
CIO

- Disponibilidade do sistema
- Despesas
- Satisfação dos clientes
- Contagem de problemas graves
- Uso de CPU com relação à capacidade
- Uso de espaço de armazenamento com relação à capacidade
- Tráfego na rede
- Tempo de resposta de aplicações
- Marcos dos projetos
- Principais projetos na fila
- Outros eventos críticos

CIO Dashboard

(As of December 19, 2012)

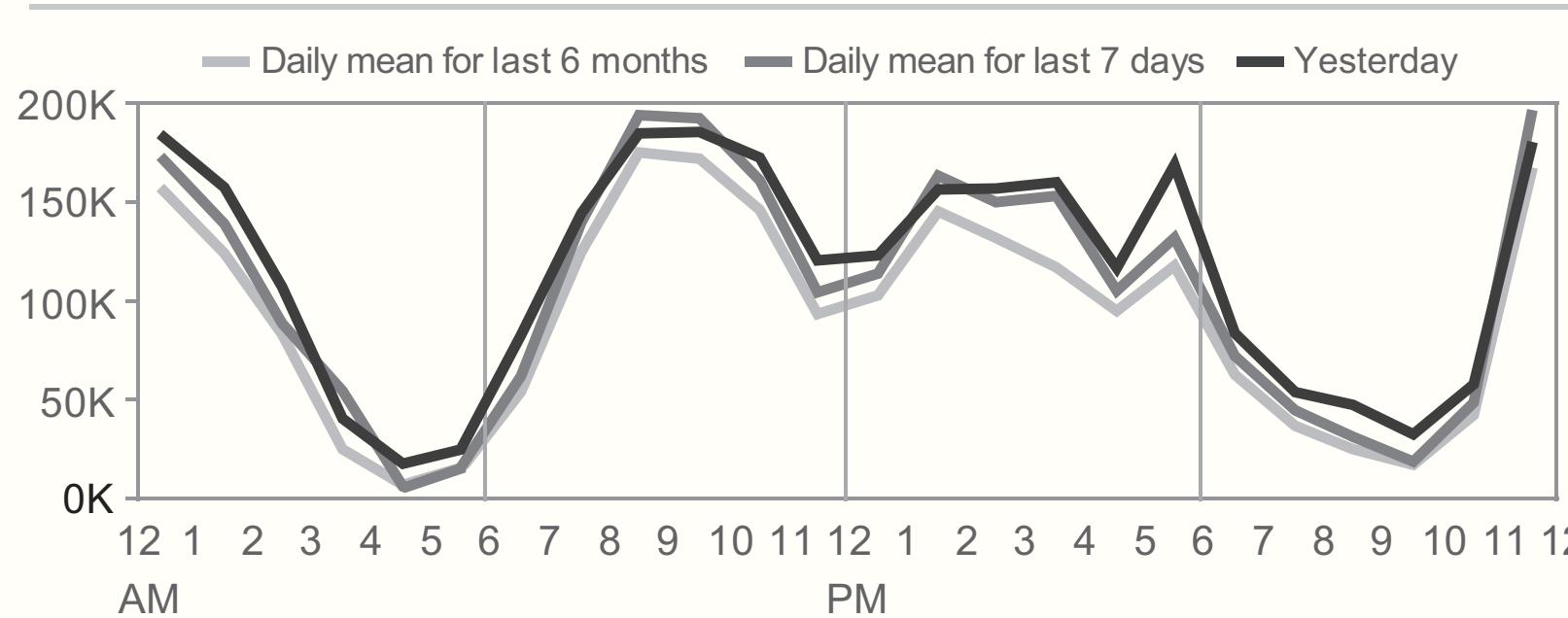
System Availability (last 30 days)



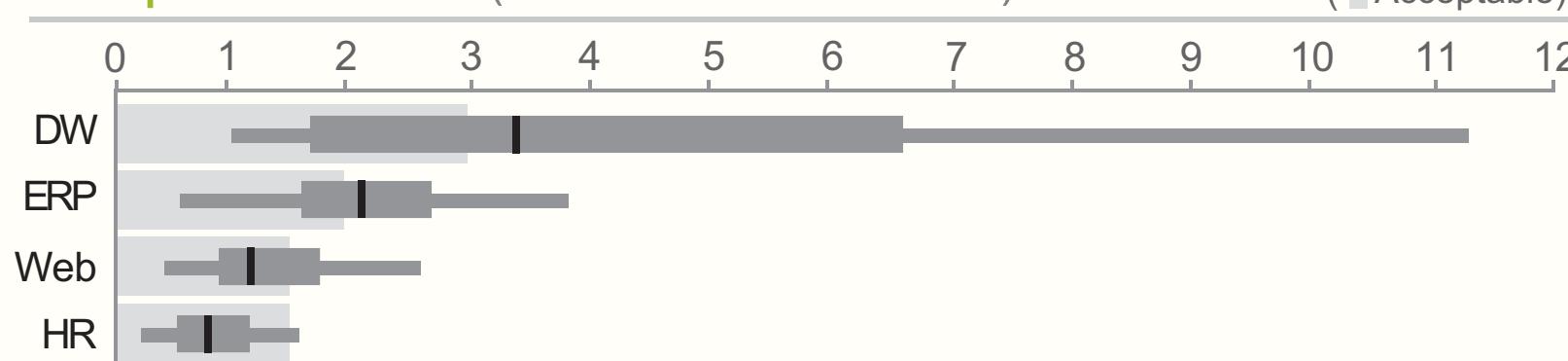
Hardware % of Capacity



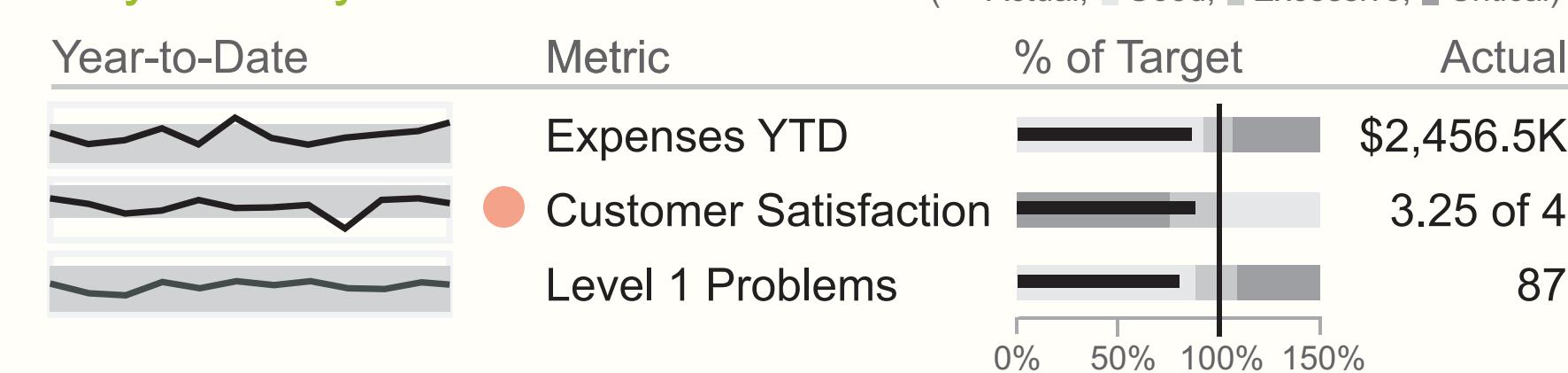
Daily Network Traffic (kilobytes)



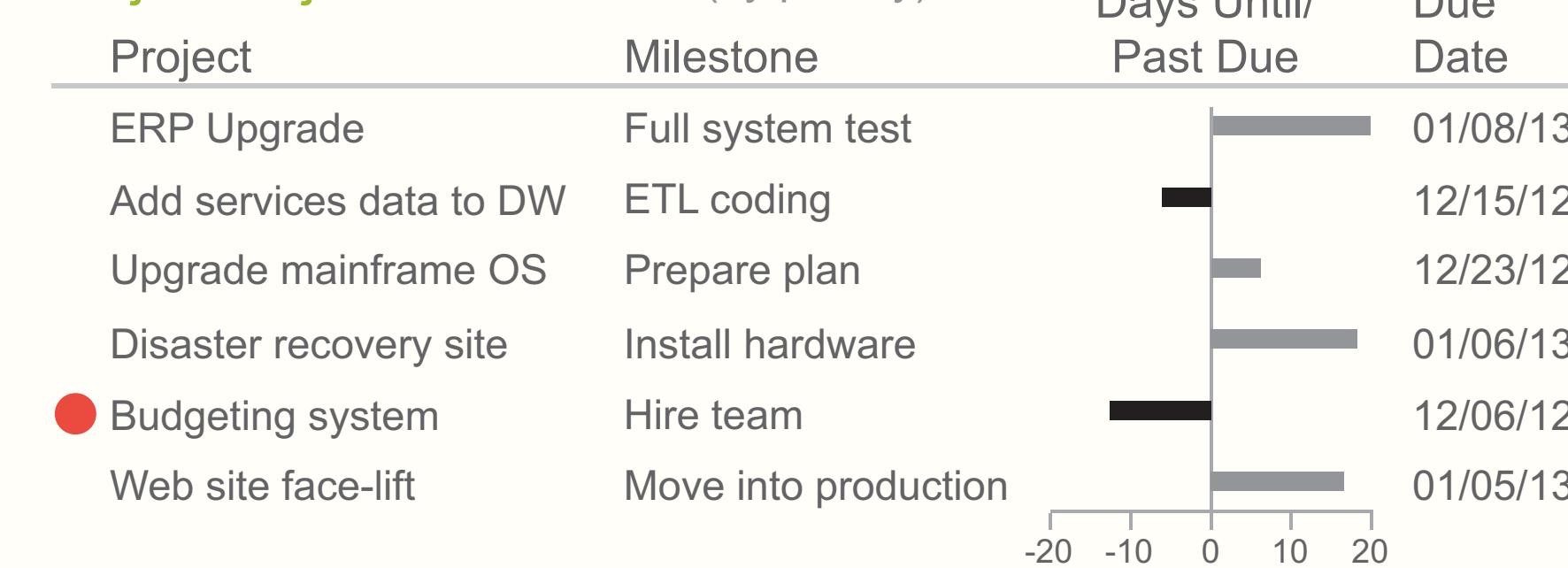
Response Time (distribution in seconds)



Key Non-System Metrics



Major Project Milestones (by priority)



5 Top Projects in the Queue

Project	Status	Funding Approved	Sched. Start
1 Professional services module	Pending available staff	X	02/10/13
2 Upgrade MS Office	Cost-benefit analysis		02/18/13
3 Failover for ERP	Preparing proposal		03/02/13
4 Upgrade data warehouse HW	Evaluating options	X	04/15/13
5 Executive dashboard	Vendor assessment		05/01/13

5 Top Critical Events (next 14 Days)

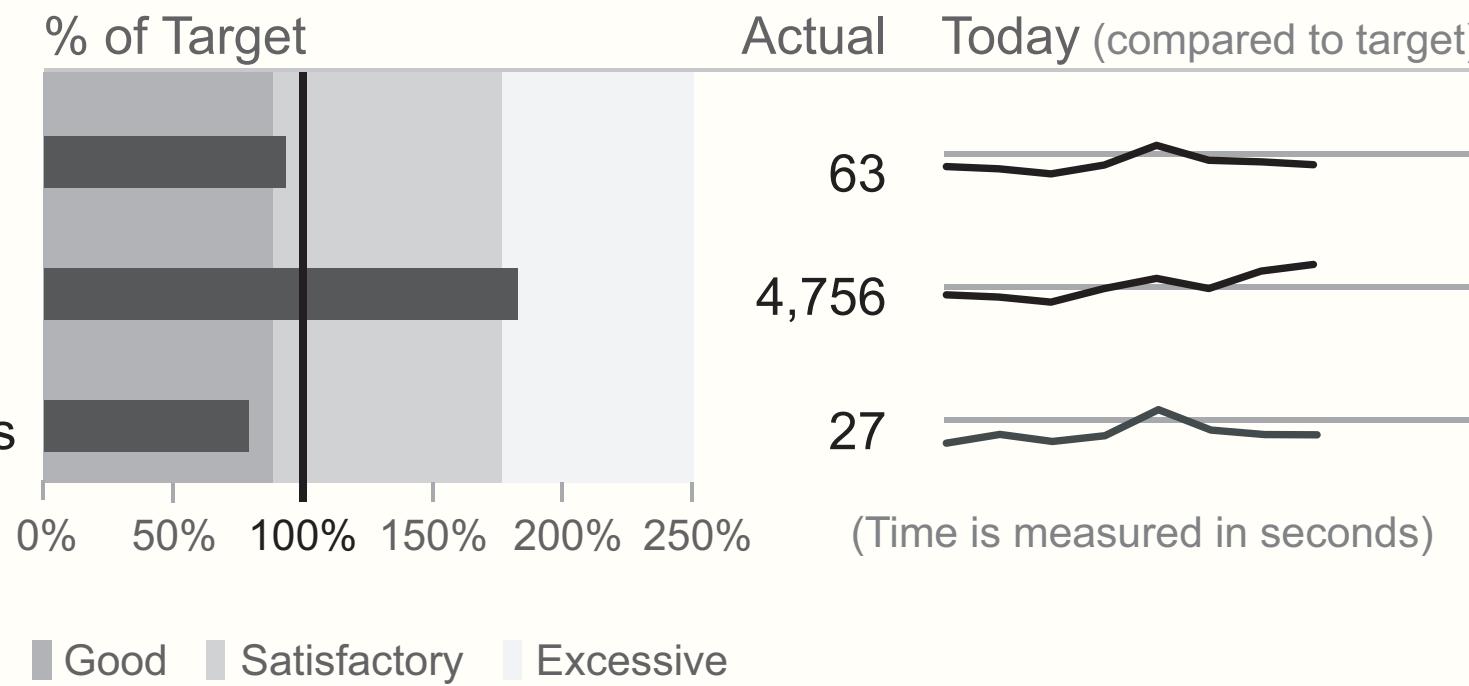
Event	Group Responsible	Date
Full system maintenance outage from 9-11 PM	G. Jones	12/21/12
Present hardware upgrade proposal to CEO	Self/M. Smith	12/22/12
Tom visiting from Asia office	Self	12/23/12
Prepare quarterly financial for public announcement	Self	01/04/13
Present revised information strategy to steering comm.	J. Kane	01/06/13

TELEMARKETING

- Tempo de espera por atendimento
- Duração da chamada
- Número de chamadas abandonadas
- Volume de chamadas
- Número de pedidos

Overall Performance

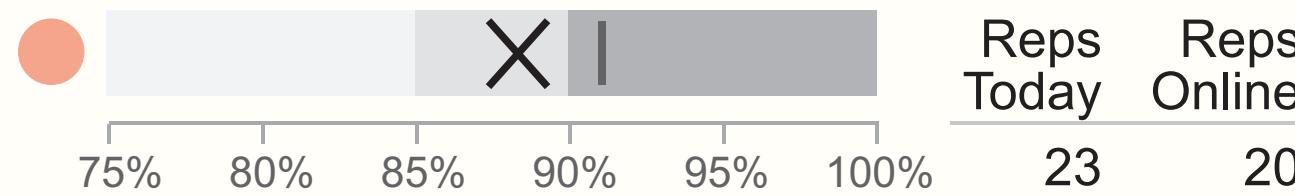
Hold Time



Call Duration

Abandonments

Utilization



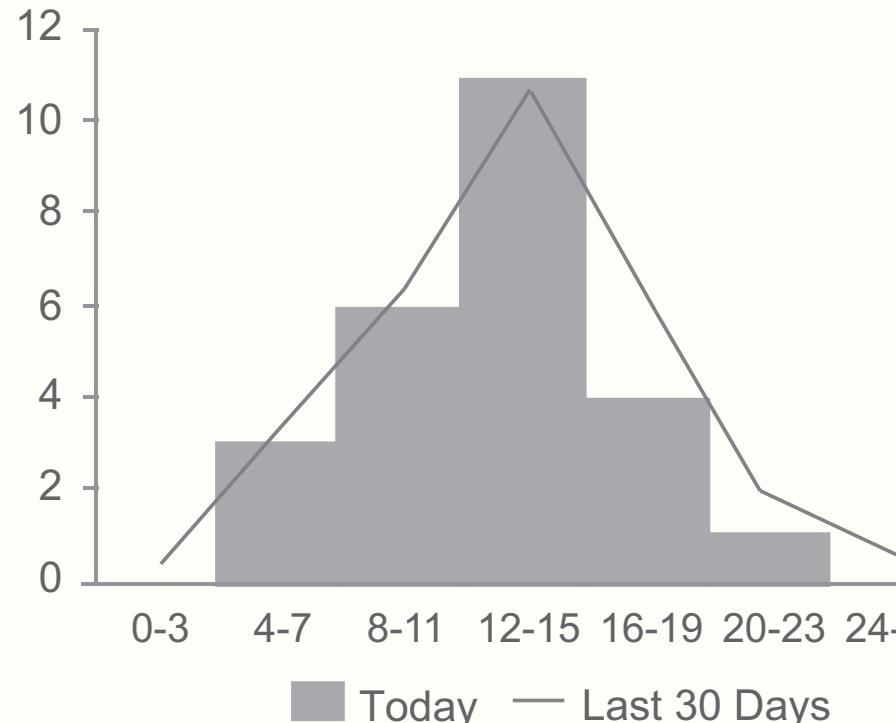
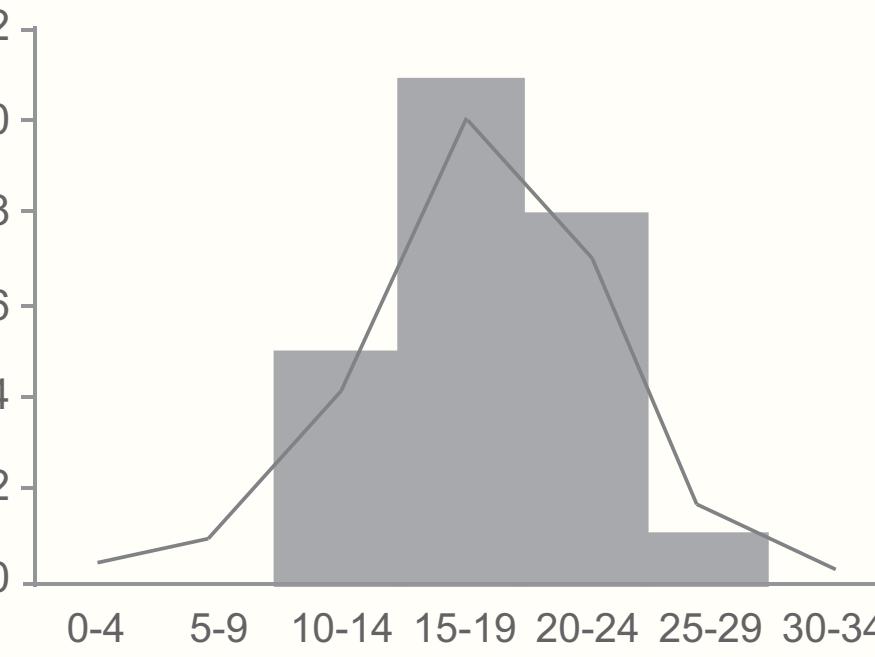
Volume

This Hour Today This Month Per Hour Today

Call Count 373 1,322 25,934

Order Count 234 925 17,834

Reps Mean Hourly Calls per Rep Reps Mean Hourly Orders per Rep



Rep Performance

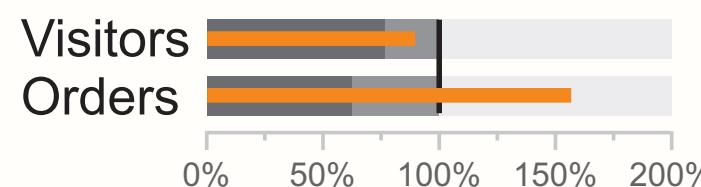
Name	Orders Per Hr	Calls Per Hr	Call Duration (minutes)
Sally Jacobs	5	10	6 9 12 15 18 21
Jake McKinsey	5	13	6 9 12 15 18 21
Valerie Smith	6	12	6 9 12 15 18 21
Rob Wilcox	9	14	6 9 12 15 18 21
Pat Clark	10	14	6 9 12 15 18 21
Brad Simons	10	16	6 9 12 15 18 21
Andrew Newman	11	15	6 9 12 15 18 21
Sharon Bailey	11	16	6 9 12 15 18 21
Tim Barclay	11	17	6 9 12 15 18 21
Joe Jimenez	12	16	6 9 12 15 18 21
- April Chou	12	17	6 9 12 15 18 21
Hank Kata	12	17	6 9 12 15 18 21
Carole Silverstein	13	18	6 9 12 15 18 21
Pete Schuster	13	18	6 9 12 15 18 21
Mary Truman	13	19	6 9 12 15 18 21
- Barry Pierce	14	19	6 9 12 15 18 21
Jon Fisher	14	20	6 9 12 15 18 21
Tamara Jung	14	20	6 9 12 15 18 21
Stu English	15	21	6 9 12 15 18 21
Pete Wiley	15	21	6 9 12 15 18 21
Ned Johnson	16	21	6 9 12 15 18 21
- Joe Lucas	16	22	6 9 12 15 18 21
Randy Forester	17	23	6 9 12 15 18 21

— = Currently offline

ANALISE DE MARKETING

- Número de visitantes (dia, mês, ano)
- Número de pedidos
- Número de visitantes registrados
- Número de visualizações de cada produto
- Produtos exibidos juntos e raramente comprados juntos
- Produtos exibidos separados e comprados juntos
- Referências de sites externos que tenham resultado em mais visitas

This Month Compared to Target

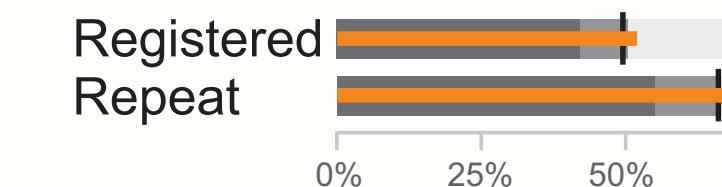


Web Marketing Analysis

Data as of 2:00 PM (PST), April 13, 2012

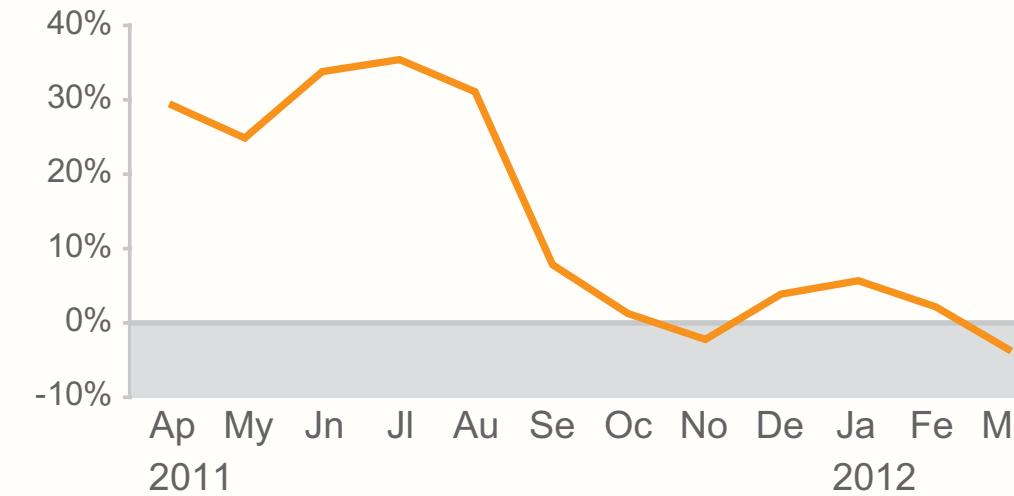
Help

Today Compared to Total Visitors

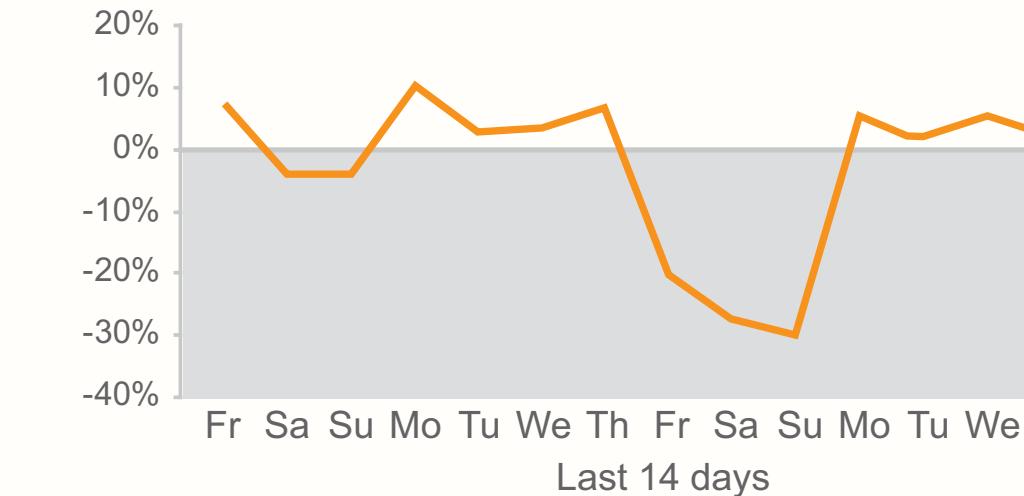


Visitors

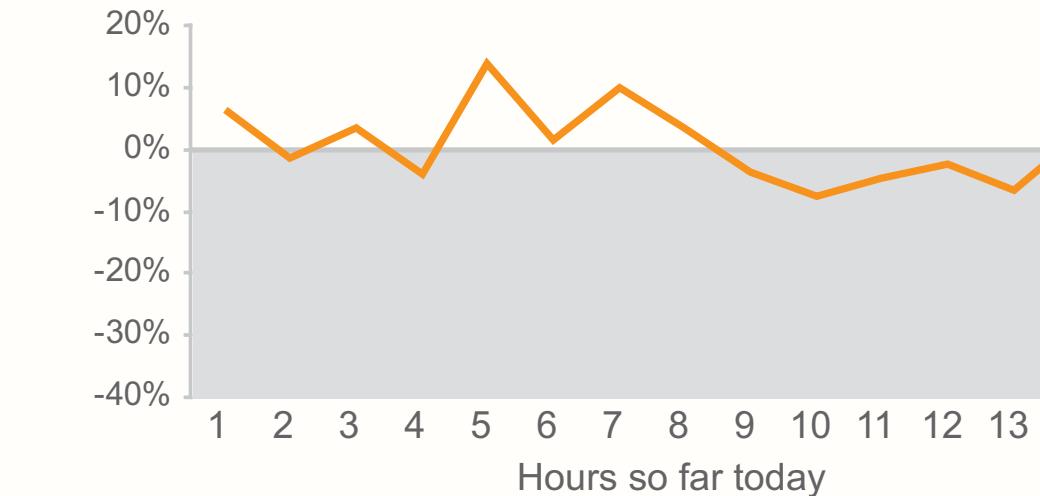
Last 12 months' average daily visitors variance relative to same month in the prior year



Last 14 days' daily visitors variance relative to 13-week average for the same weekday



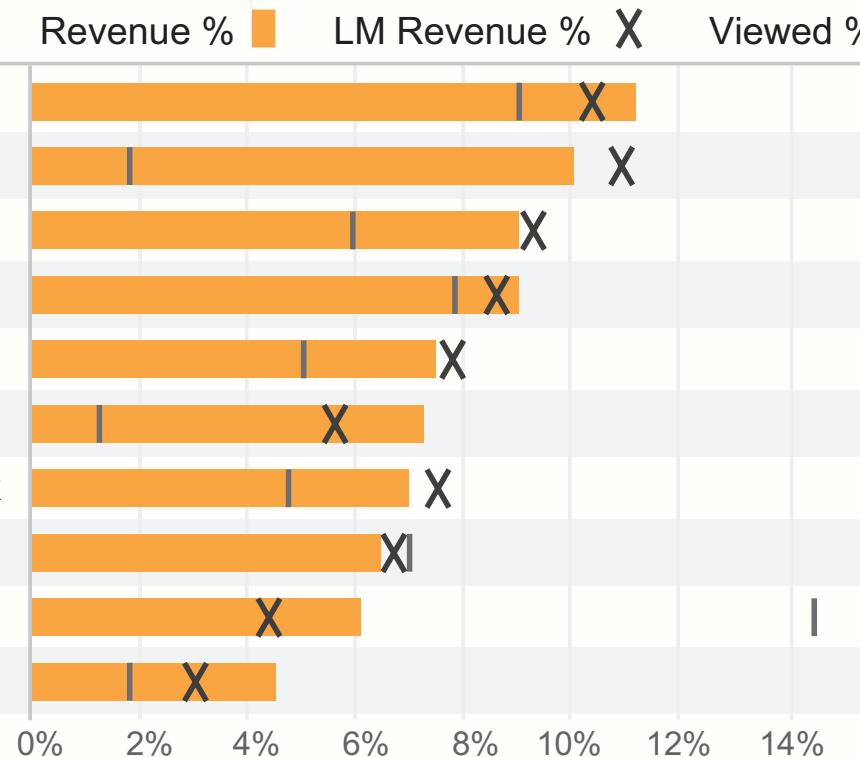
Today's hourly visitors variance relative to 13-week average for same hour



Products

Last 30 Days

Top 10 this Month by Revenue



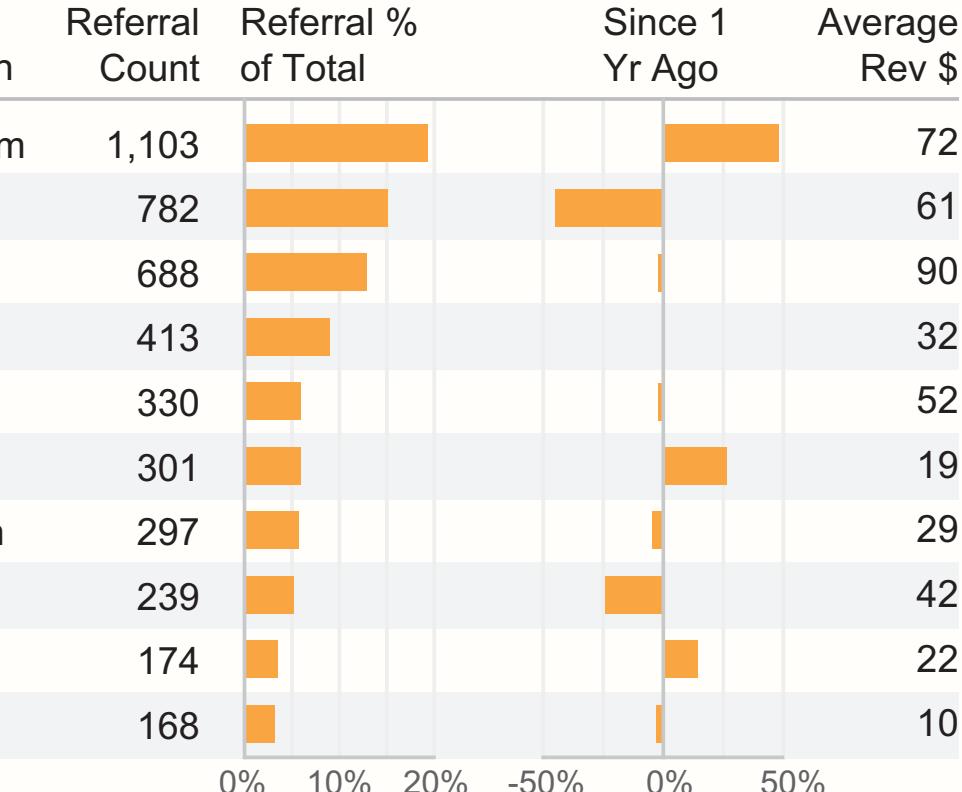
Top 10 products often purchased together but not displayed together



Referral Sites

Last 12 Months

Top 10 Referrers this Month



Top 10 products displayed together but rarely purchased together

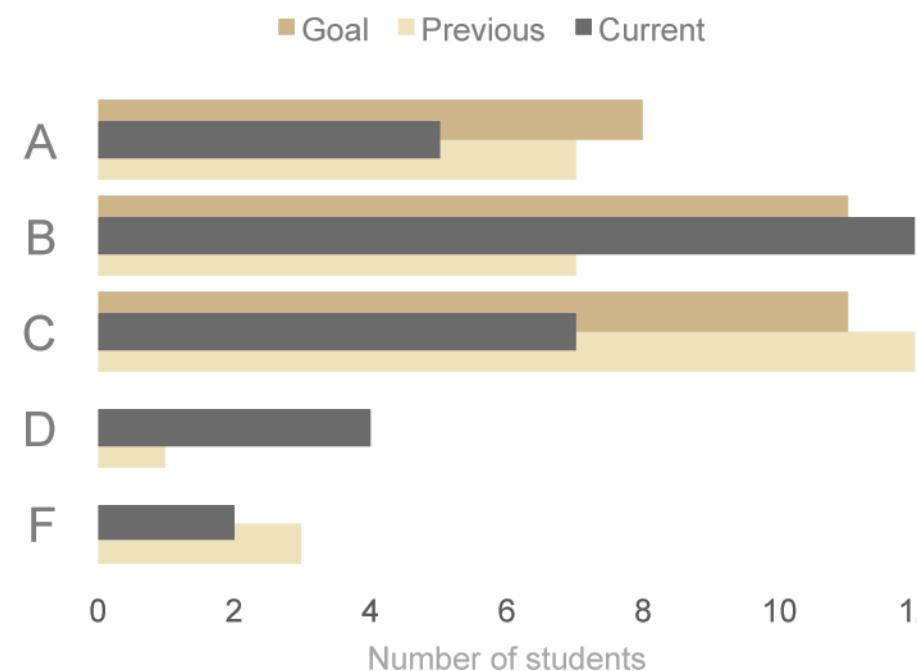


Class Dashboard

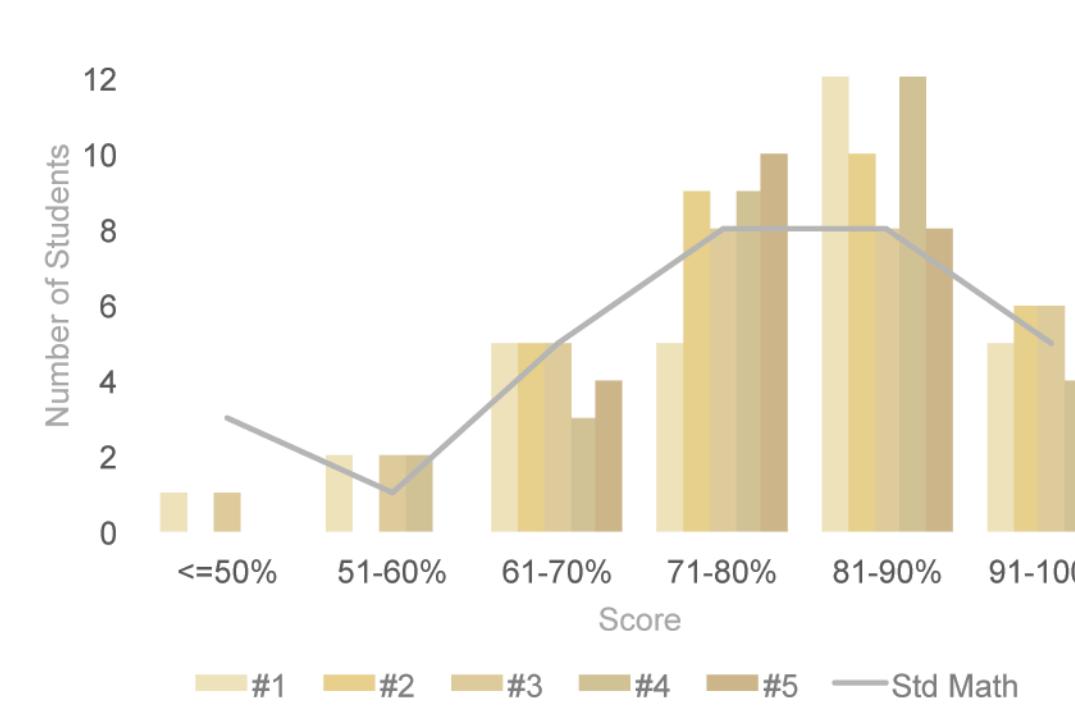
Algebra 1 - 30 students

Tue May 1st, 2012

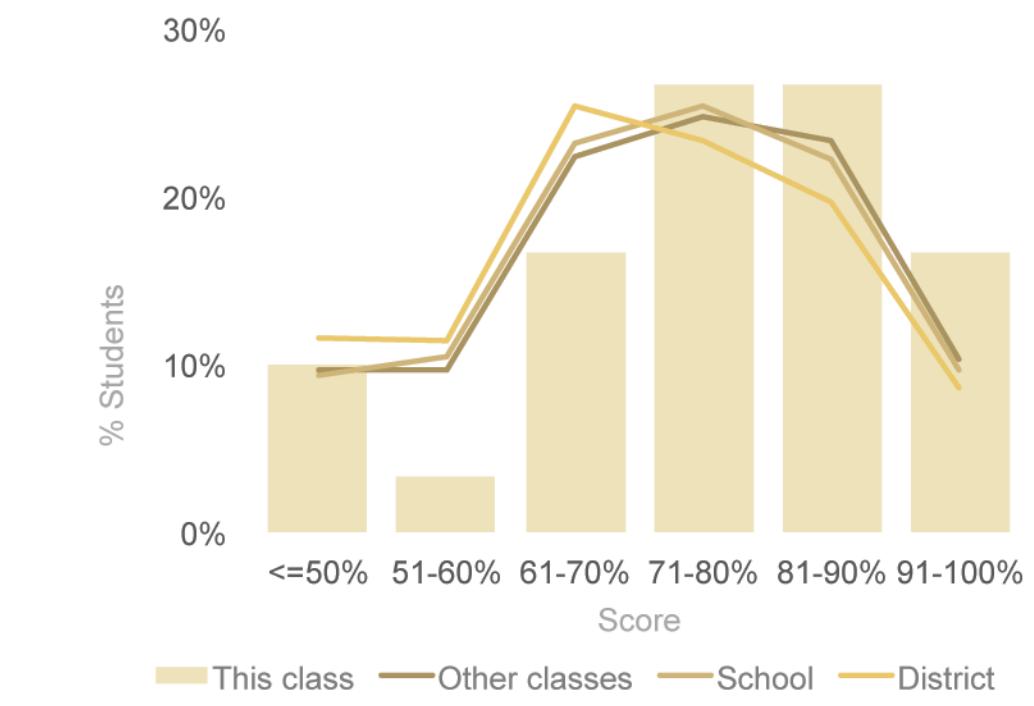
Grades



Assignments



Std Math Assessment



A's

	Current	Goal	Previous
Amala Singh	A	A	A
Hannah Li	A	A	A
Holly Norton	A	A	A
Donald Chase	A	B	A
James Snow	A	A	B

D's and F's

	Current	Goal	Previous
Anthony Harper	D	C	C
Brian Francis	D	C	C
Christopher Murphy	D	C	F
Fiona Reeves	D	C	F
Bae Kim	F	C	C
Frederick Chandler	F	C	C

Highest

Highest	Assignment Scores		Late Completions
	Min - Avg - Max	Average	
Holly Norton	1	98%	0
Donald Chase	1	94%	0
James Snow	1	94%	0
Hannah Li	1	93%	0
Amala Singh	1	92%	0

Lowest

Lowest	Min - Avg - Max	Average	Late Completions
Anthony Harper	1	69%	1
Brian Francis	1	65%	2
Fiona Reeves	1	65%	3
Bae Kim	1	59%	3
Frederick Chandler	1	53%	2

Absences

Student	Jan-Apr	Total
Class	78	78
Brian Francis	9	9
Fiona Reeves	8	8
Bae Kim	6	6
Roshawn Dawson	6	6
James Martin	4	4

Tardies

Student	Jan-Apr	Total
Class	44	44
Bae Kim	8	8
David Chenowith	5	5
Maria Garcia	4	4
Sarah Jameson	4	4
George Smith	3	3

Disciplinary Referrals

Student	This term	Last Term
Class	11	13
Bae Kim	3	0
Frederick Chandler	2	2
Lawrence Parker	1	0
Sarah Jameson	1	0
Fiona Reeves	1	1

Detentions

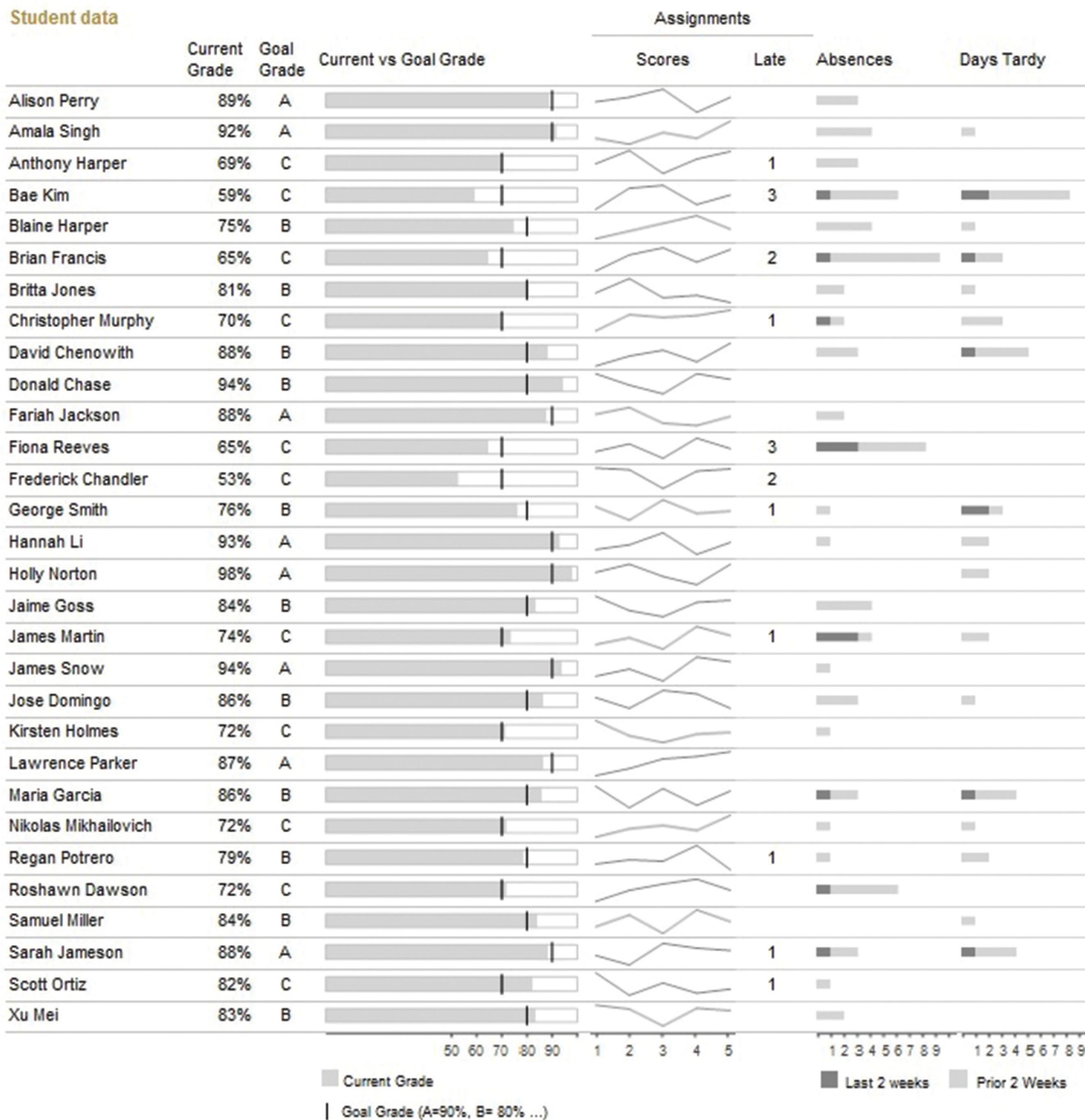
Student	This term	Last Term
Class	6	6
Bae Kim	2	0
Fiona Reeves	1	0
Lawrence Parker	1	0
Sarah Jameson	1	0
Frederick Chandler	1	1

Special Ed Status

Anthony Harper
Fiona Reeves

Low English Language Proficiency
Bae Kim
Nikolas Mikhailovich

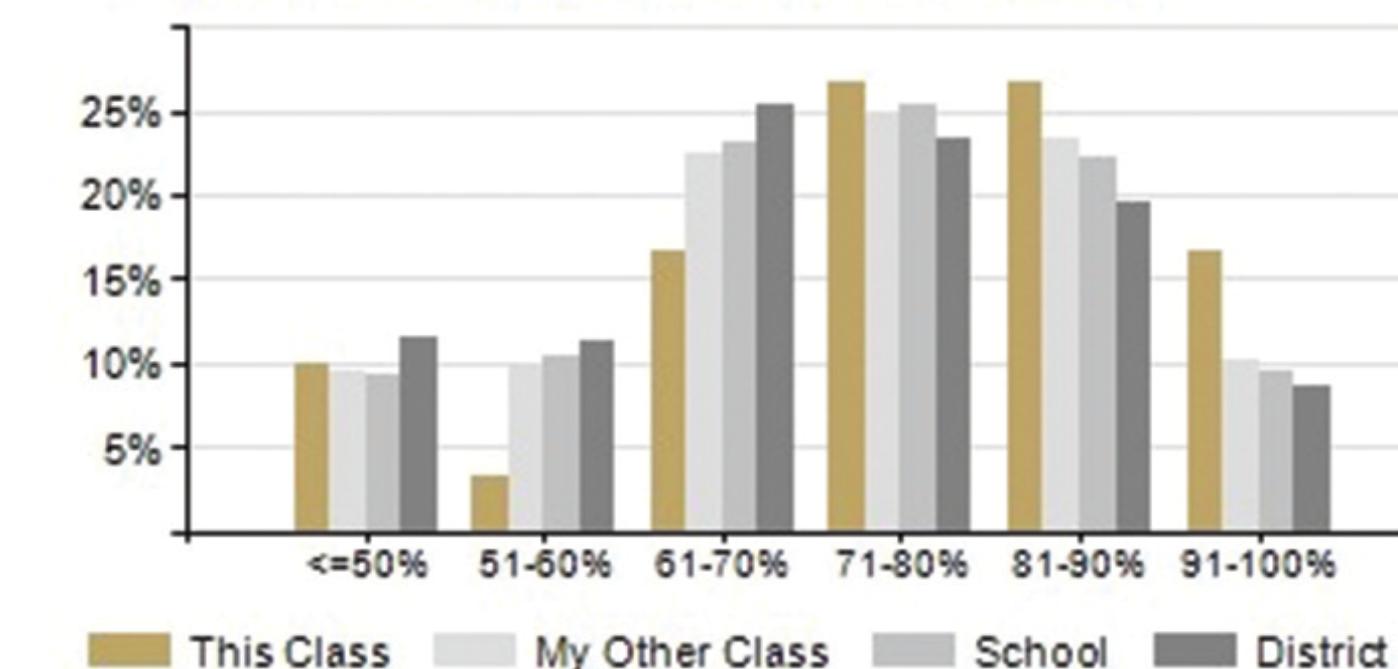
Student data



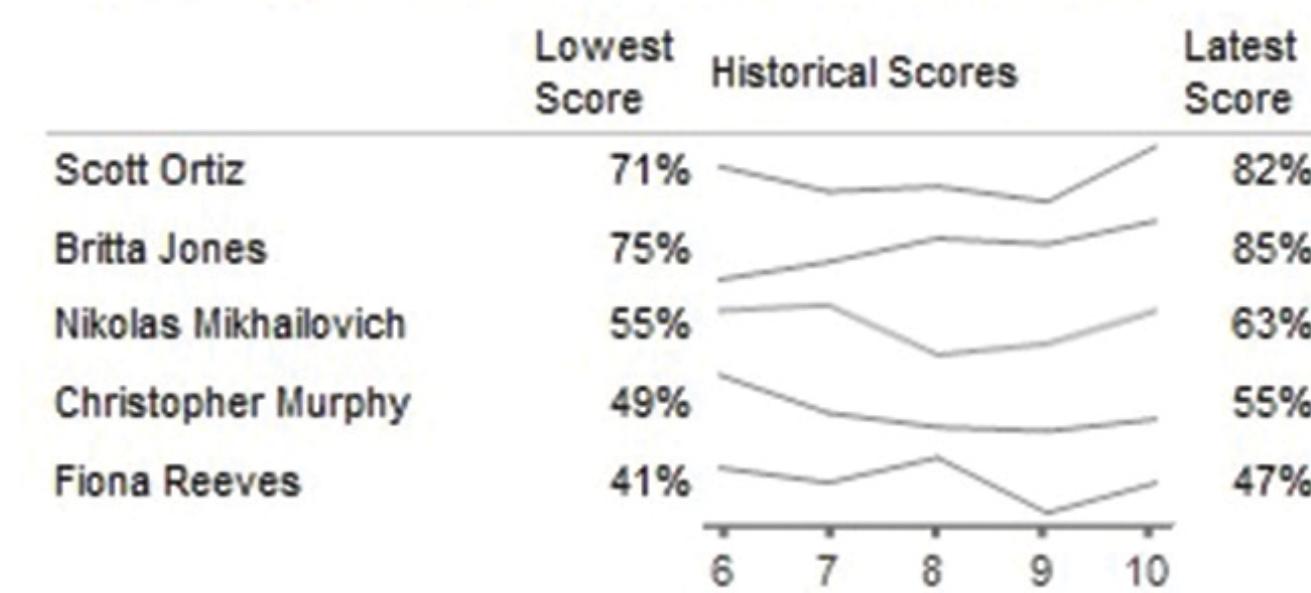
Student Tracking Sheet

Data as of May 1, 2012

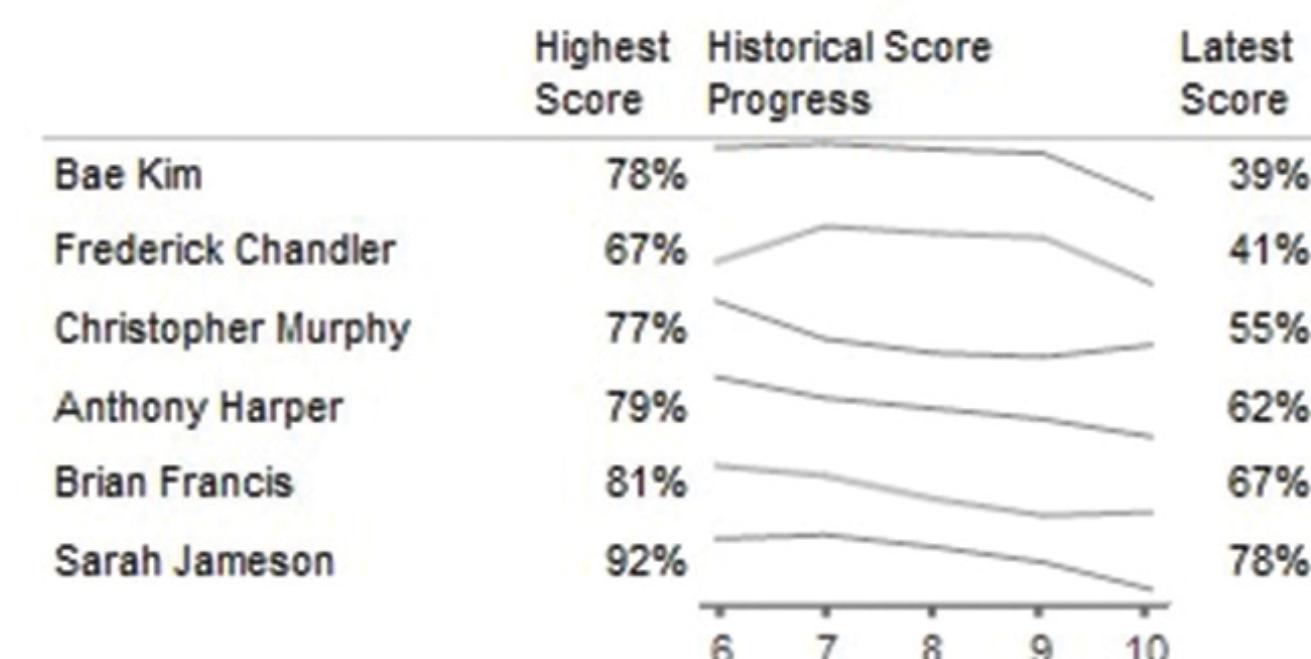
Distribution of Standardized Assessment Scores



Top 5 Standardized Assessment Progress 5 Years



Bottom 5 Standardized Assessment Progress 5 Years



10th Grade Algebra 1

Susan Metcalf
Silver Oaks High School

Tuesday, May 1, 2012
80% of term completed

Student Name	Grade (5 out of 8 assignments)	Late Assign	Tardies	Absences	Detentions
Class Median	B				
Holly Norton	A				
Donald Chase	A				
James Snow	A				
Hannah Li	A				
Amala Singh	A				
Alison Perry	B				
Sarah Jameson	B				
Fariah Jackson	B				
Lawrence Parker	B				
David Chenowith	B				
Jose Domingo	B				
Maria Garcia	B				
Samuel Miller	B				
Jaime Goss	B				
Xu Mei	B				
Scott Ortiz	B				
Britta Jones	B				
Regan Potrero	C				
George Smith	C				
Blaine Harper	C				
James Martin	C				
Nikolas Mikhailovich ^{ESL}	C				
Roshawn Dawson	C				
Kirsten Holmes	C				
Christopher Murphy	D				
Anthony Harper ^{SPED}	D				
Brian Francis	D				
Fiona Reeves ^{SPED}	D				
Bae Kim ^{ESL}	F				
Frederick Chandler	F				

ESL = English as a Second Language
SPED = Special Education

= student's grade goal

Standardized Math Score
Class median 79.0%
School median 74.2%
District median 71.9%

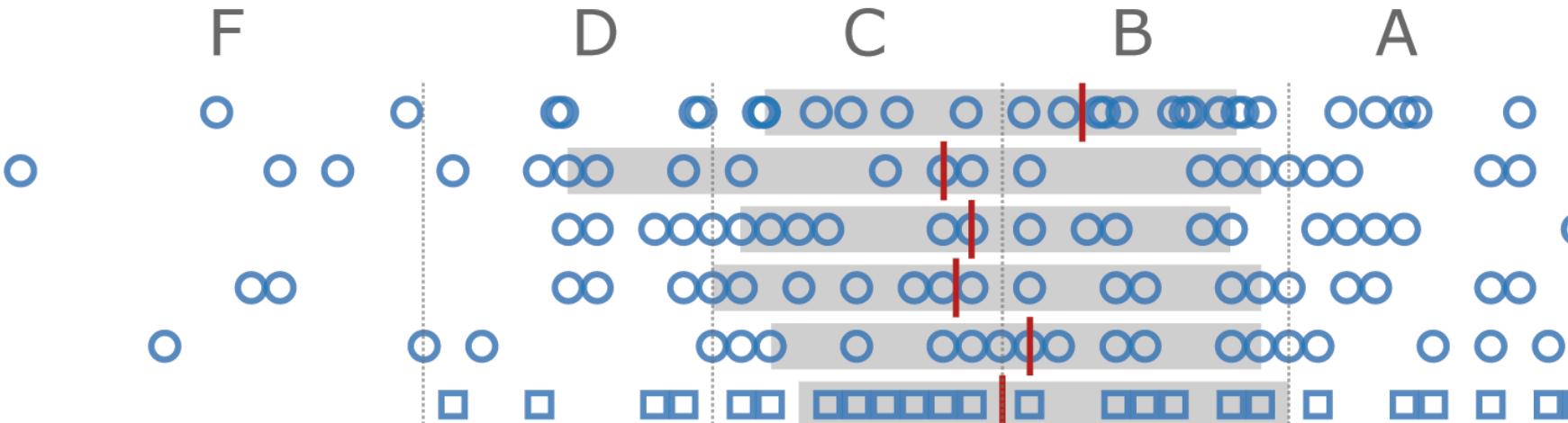
Metcalf

Algebra I

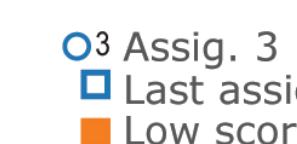
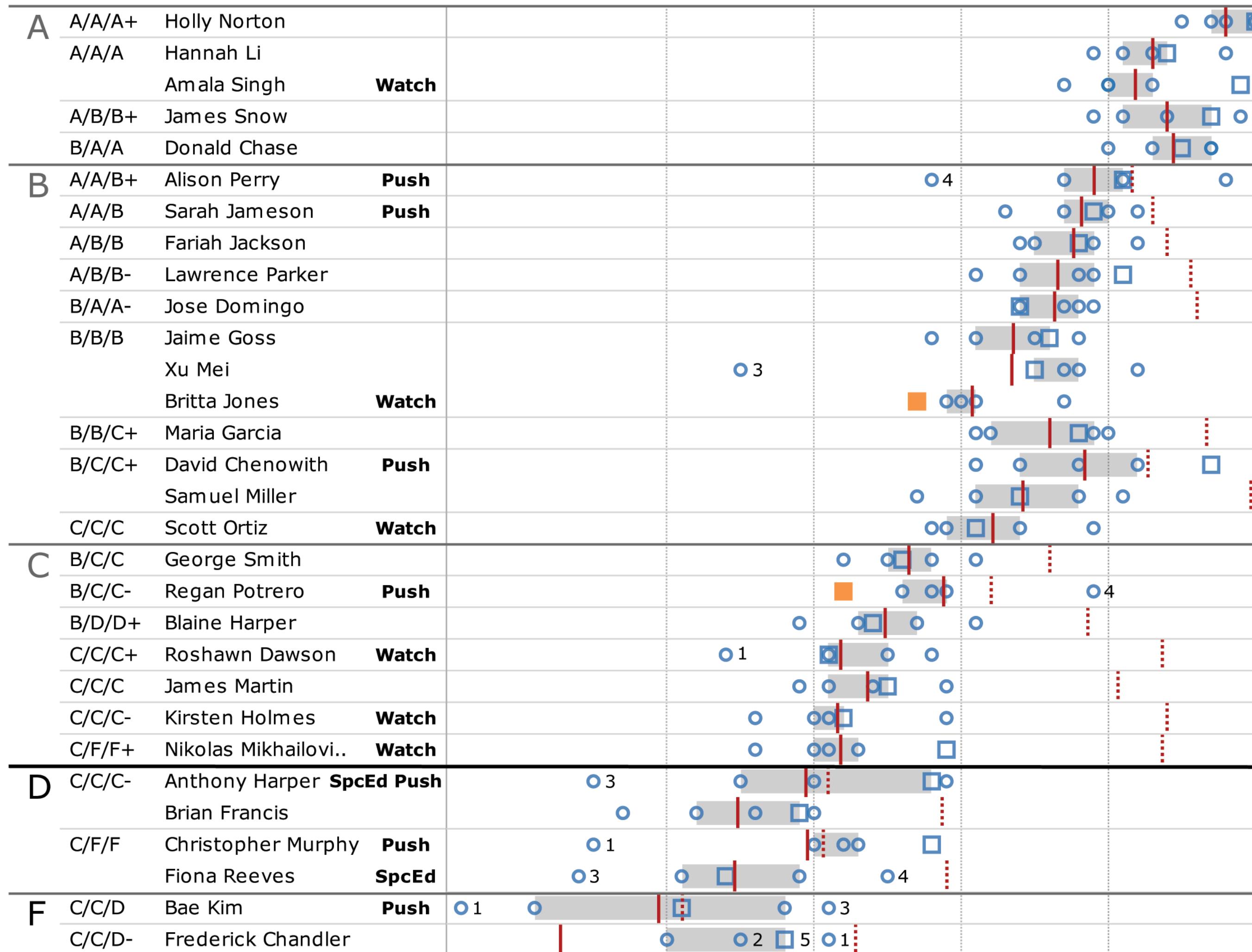
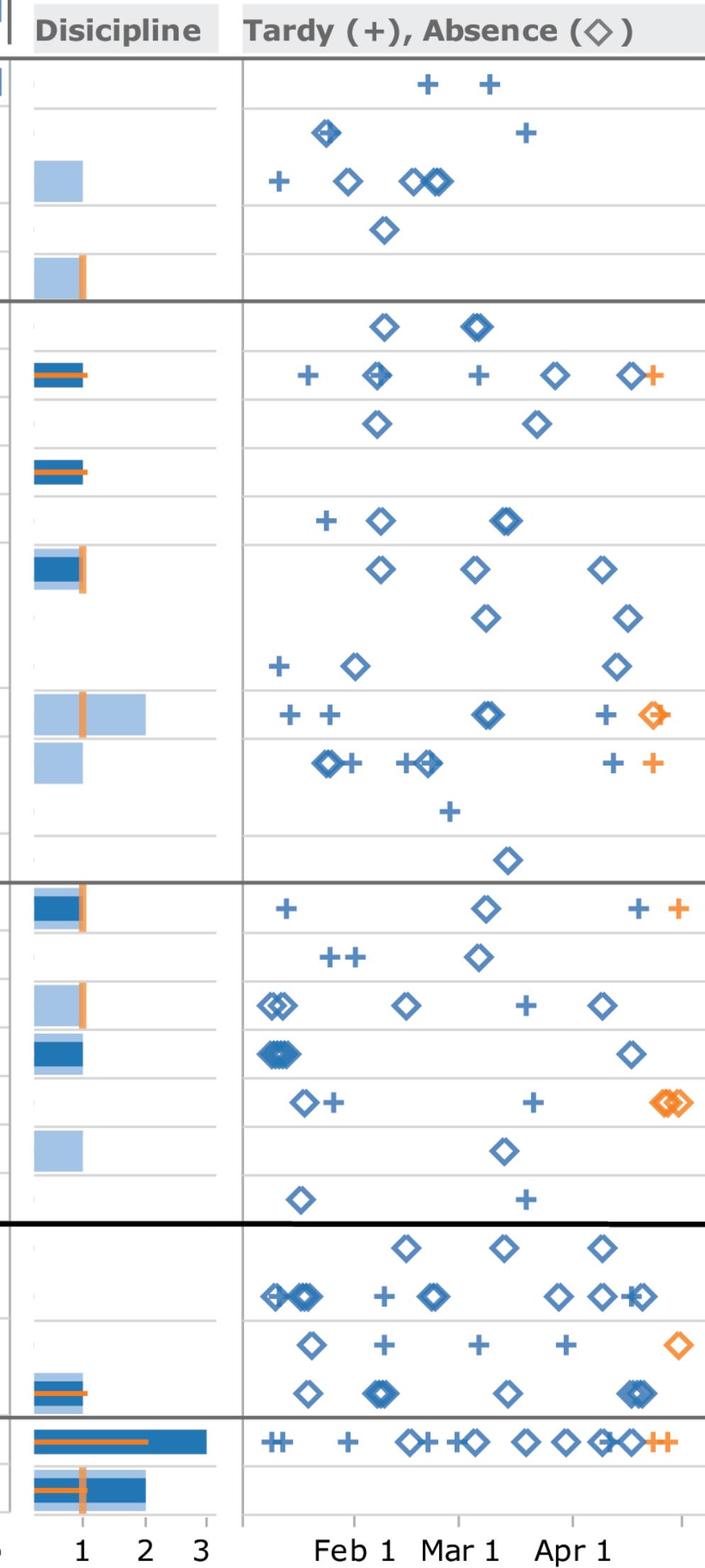
1 May 2012

Class Scores

Current Grade
Assignment 1
Assignment 2
Assignment 3
Assignment 4
Assignment 5



Watch A- / B- / C- students	Amala / Britta, Scott / Roshawn, Kristen, Nicholas
Push B+ / C+ / D+F+ students	Alison, Sarah, David / Regan/ Anthony, Christopher, Bae
Absent in last week	Maria, James, Christopher



Class: Algebra 1
May 1st, 2012

Current grade
Target grade
Previous year

Last 5yrs Standardized
Math Assesments

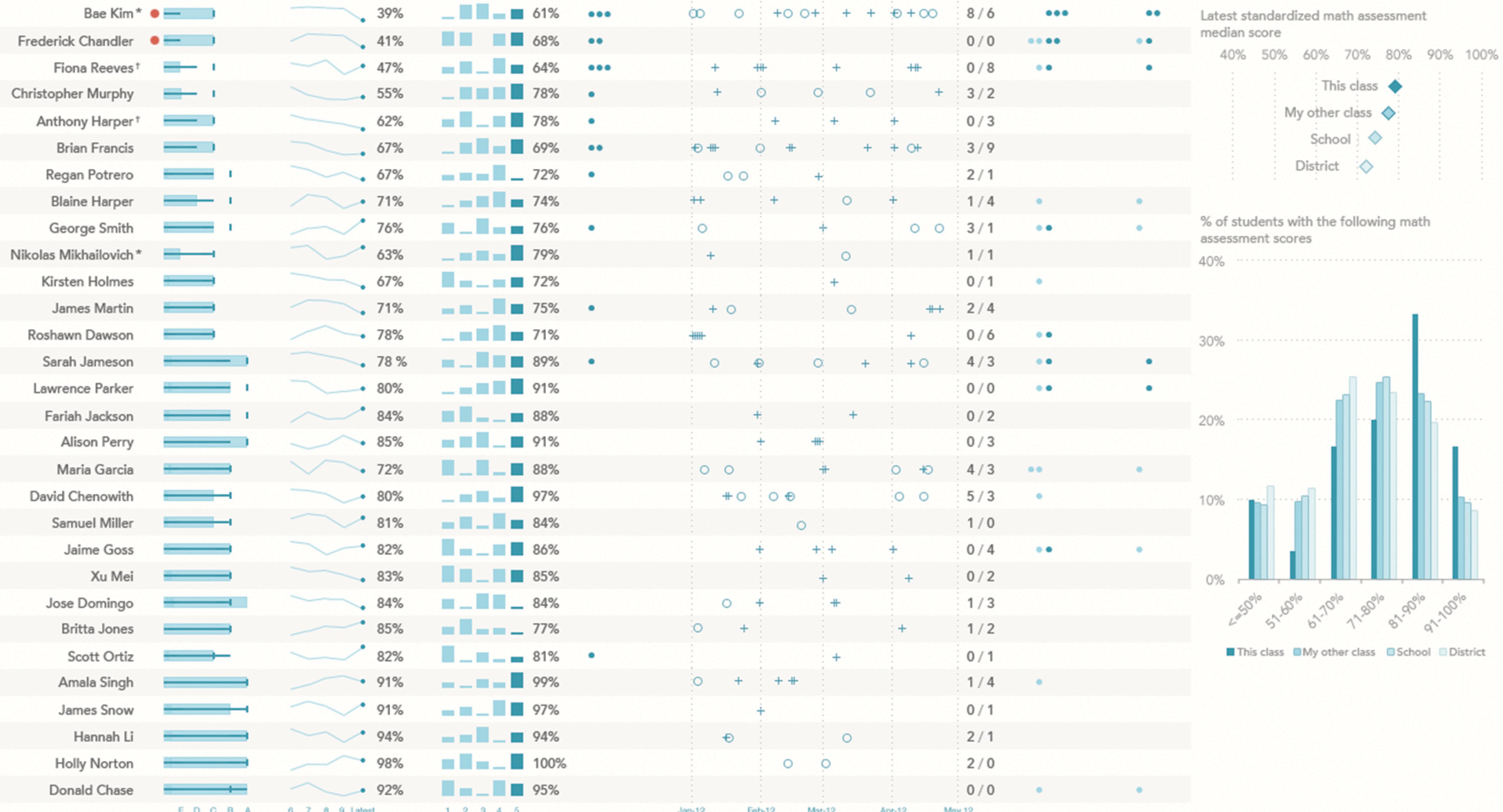
Last 5 assignments
Completed Late

Days tardy / Days absent

Disc. referrals
last term
this term

Detentions
last term
this term

Class comparisions



* No english language proficiency

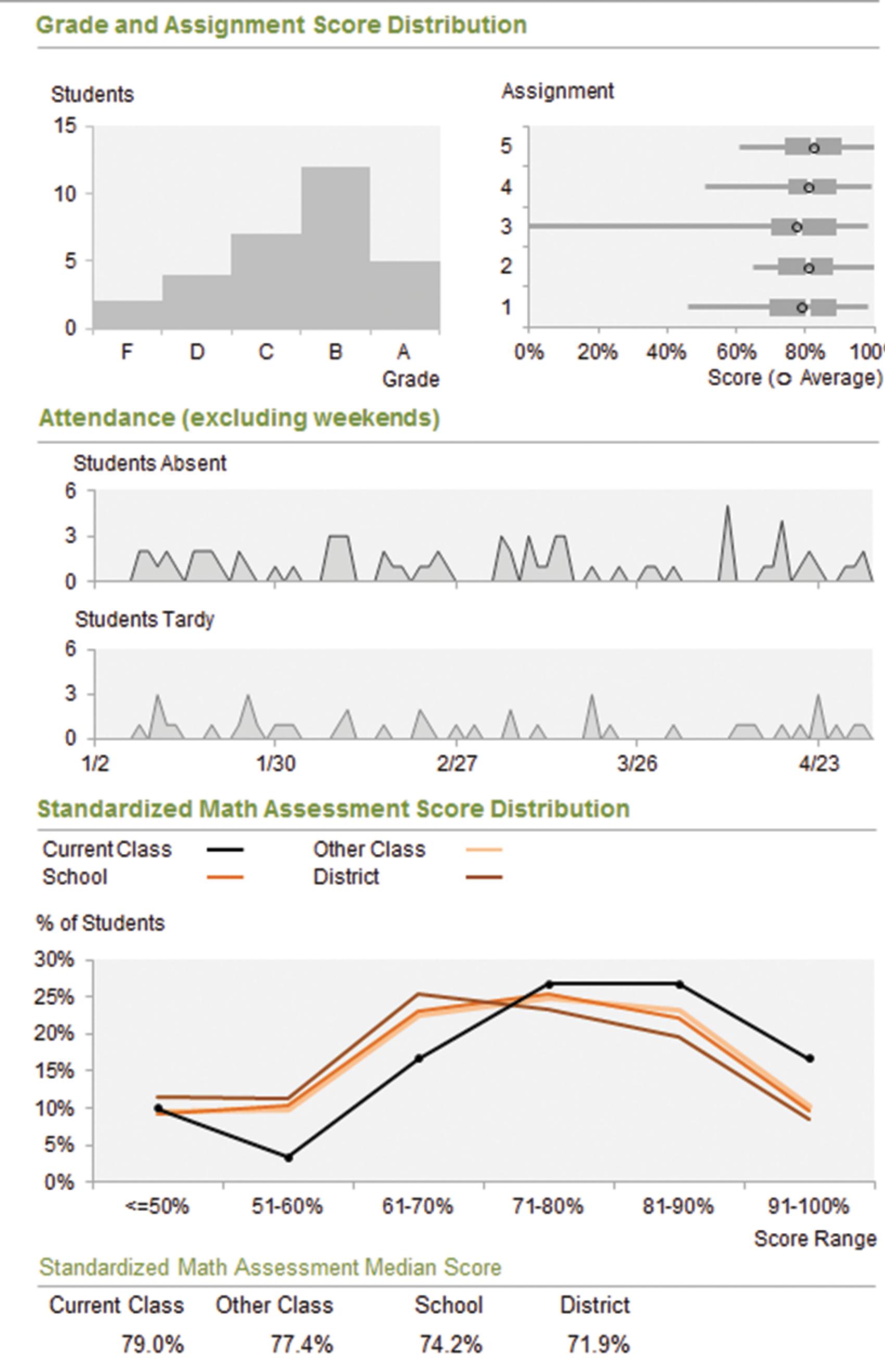
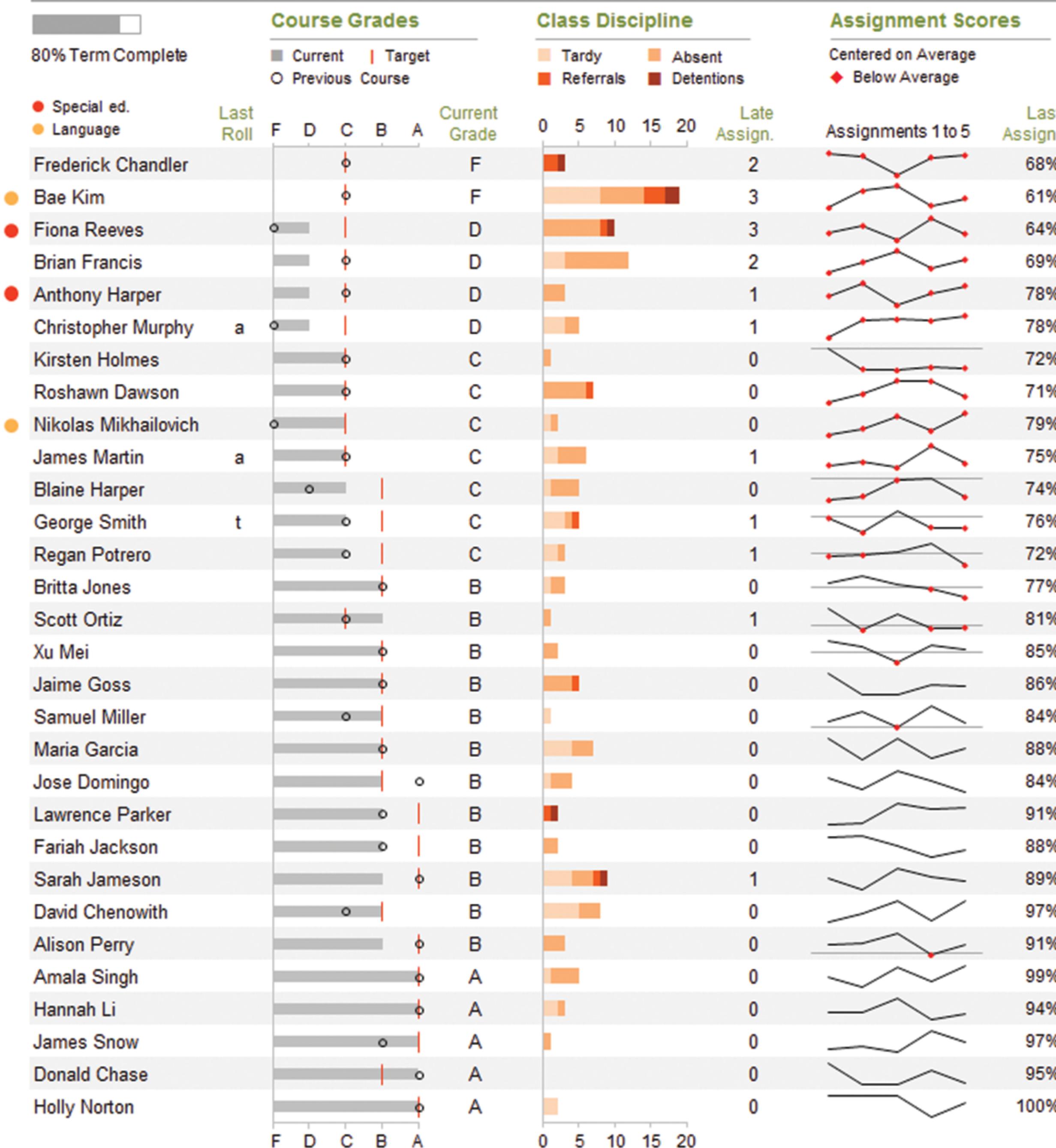
† Special education

Note: Assessment and assignment scores are being expressed as the percentage of points that were earned out of the total points possible.

Grade 10 Algebra Course

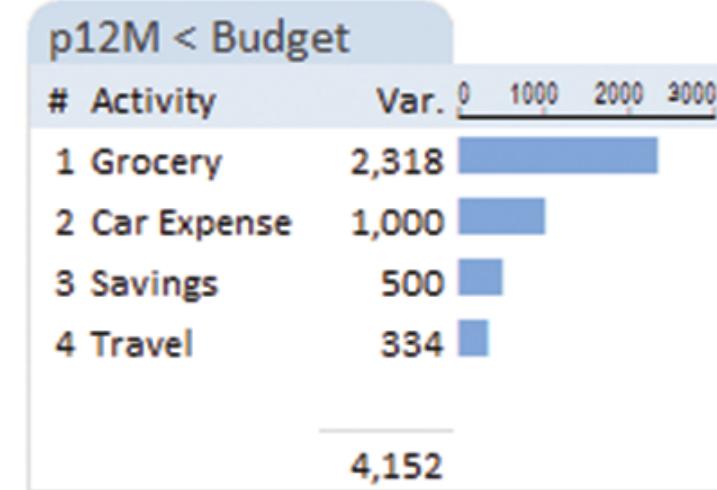
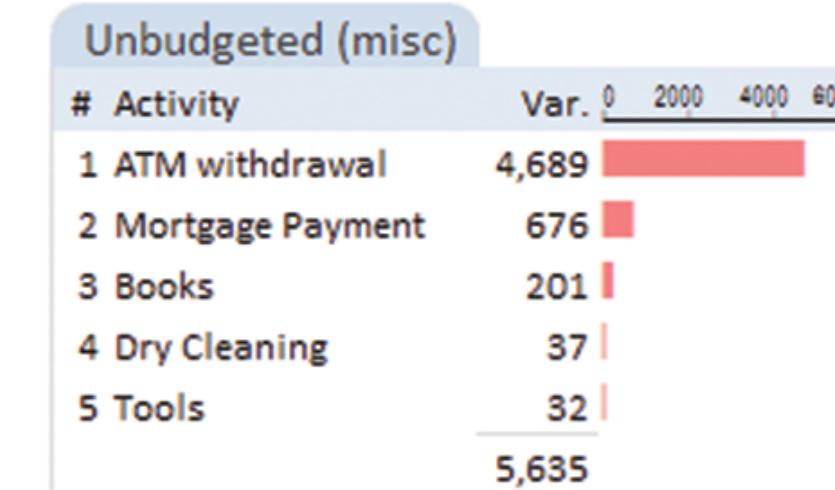
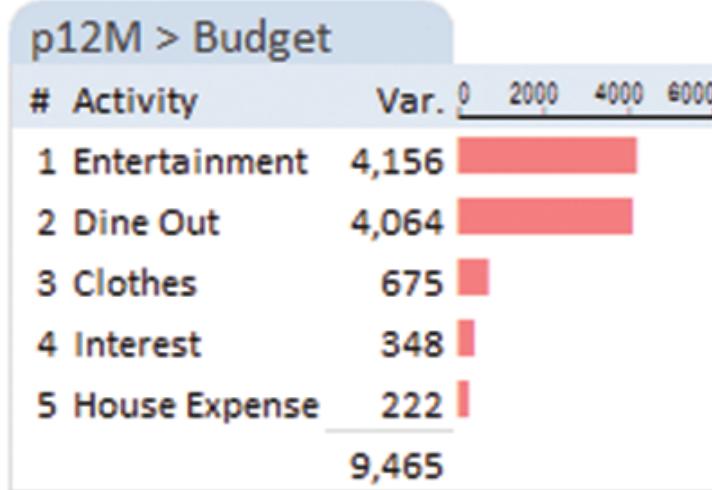
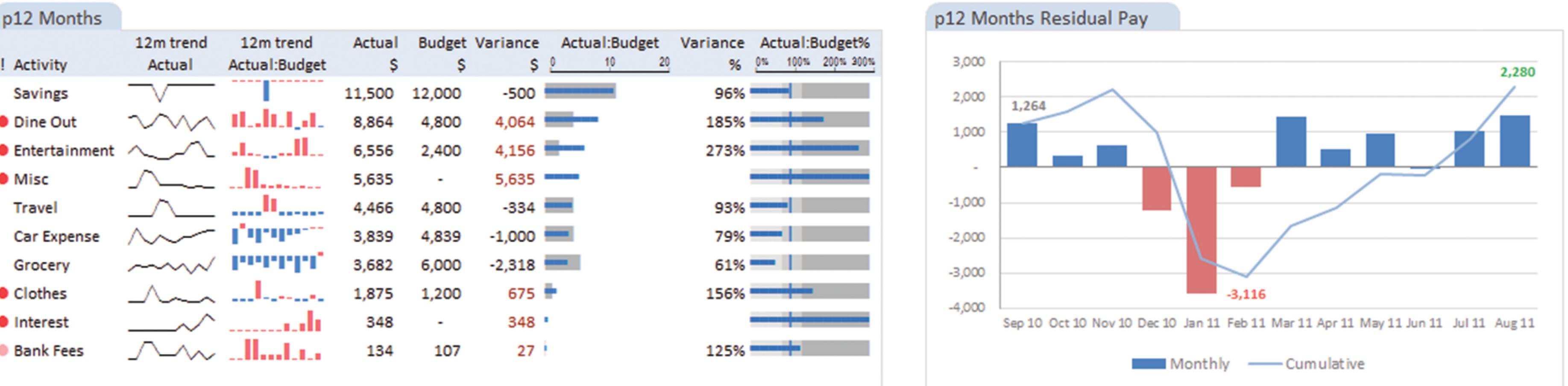
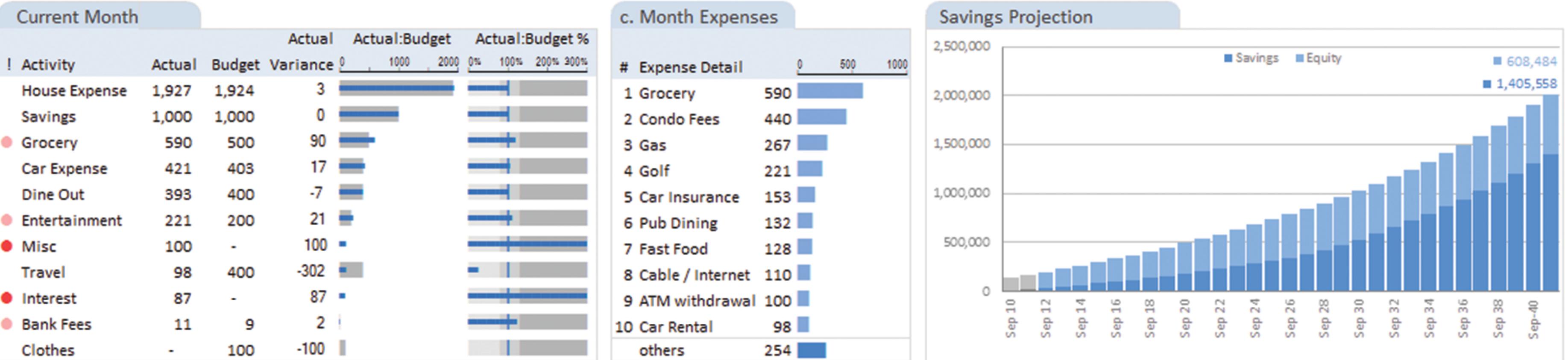
Note: All scores are expressed as percentage of points earned out of the total points possible.

HELP



Personal Finance Dashboard: August 2011

Current Position	
Savings	25,020
Equity	142,720
Mortgage Principal	213,580



Hospital CEO

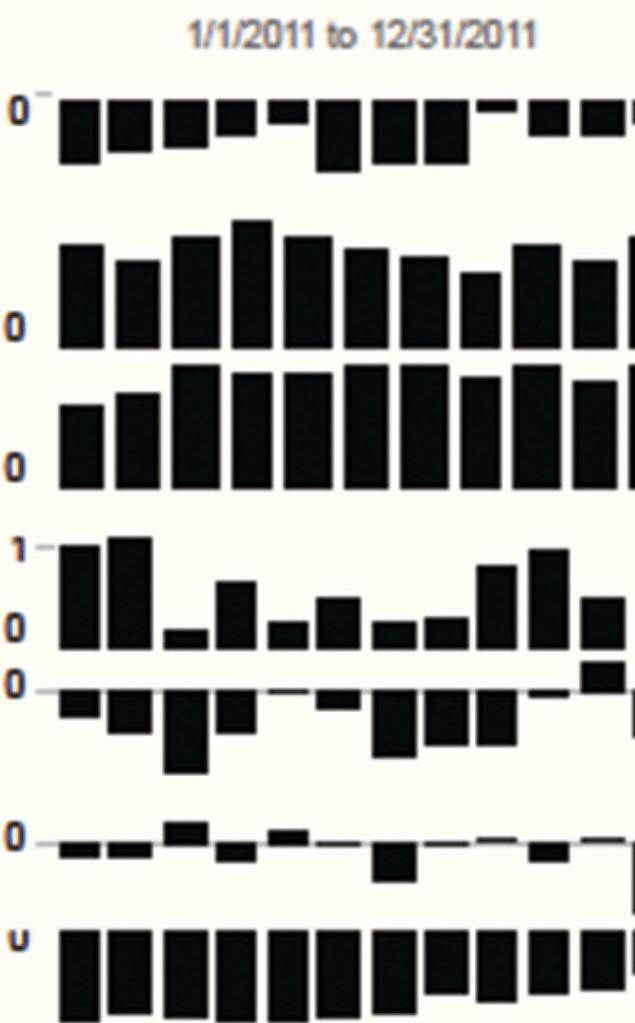
YTD Performance 1/1/2011 to 12/31/2011

One Month Results – Dec 2011

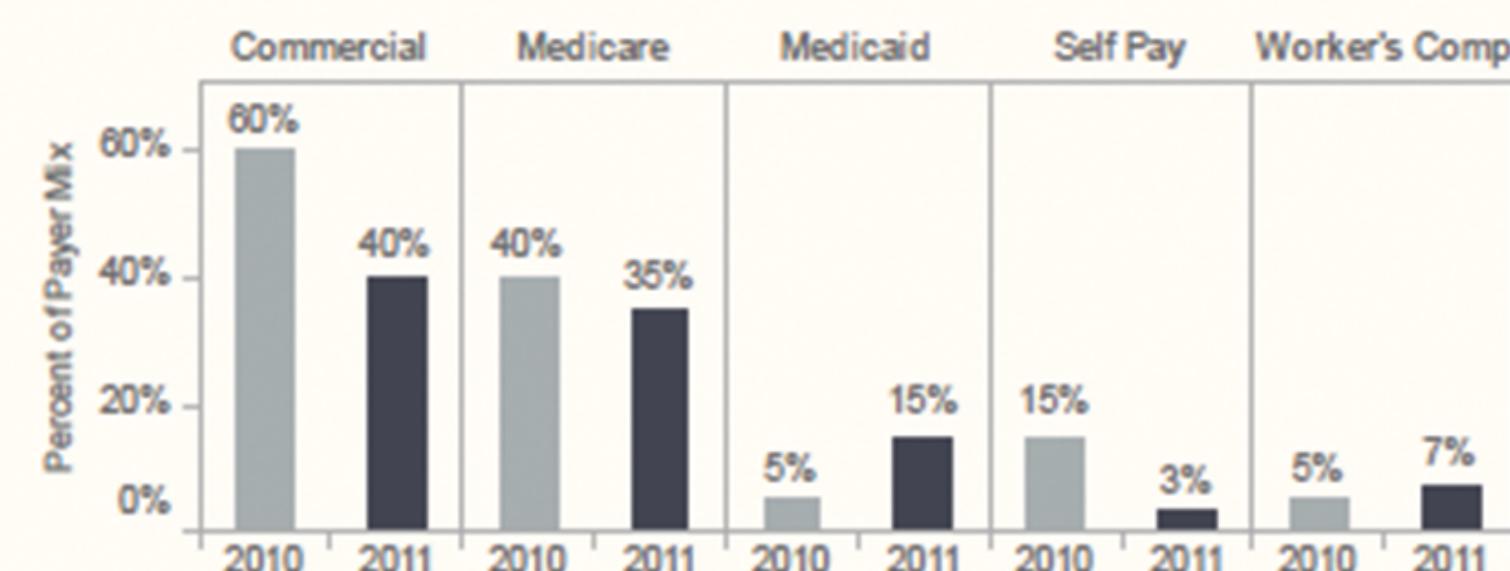
Metric Actual Budget Over/Under

	1/1/2011 to 12/31/2011		
Average Occupancy Rate	75%	85%	-10% ▼
Average LOS (Days)	3.5	2.9	20% ▲
Outpatient Visits	85.4	80.6	6% △
Expenses	\$1.5M	\$1.4M	7% ▲
Revenue	\$1.2M	\$1.4M	-14% ▼
Average Daily Census	254	264	-4% ▼
FTE's	449	465	-3% ▽

Monthly Variance to Budget

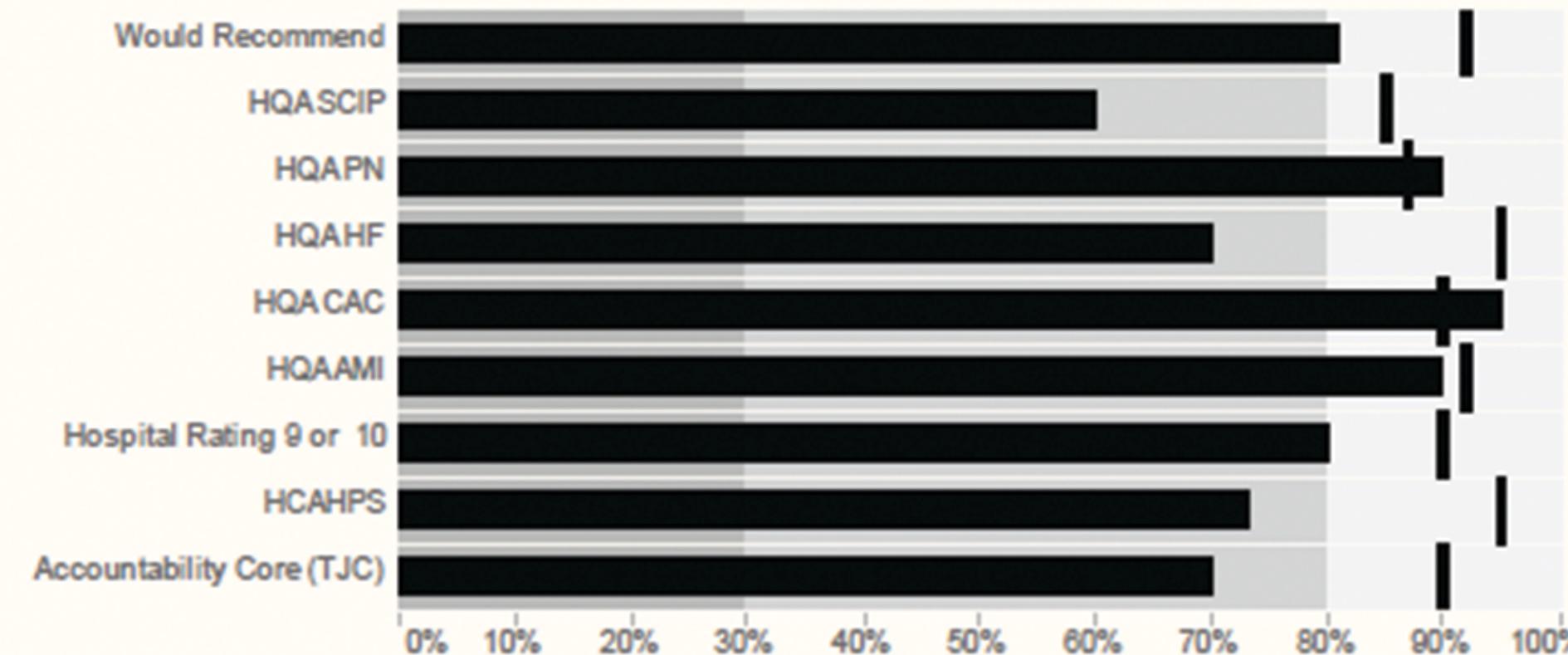


Payer Mix

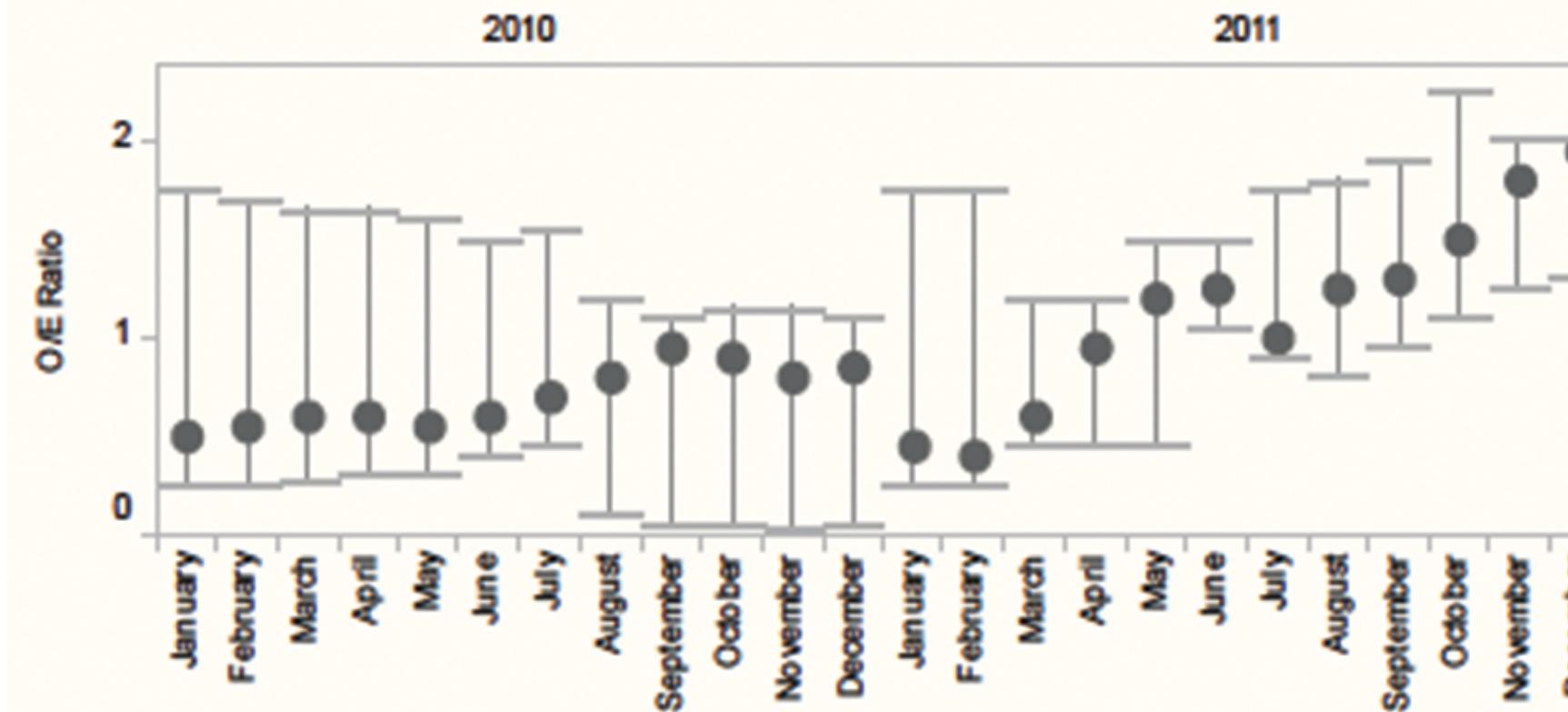


Quality & Patient Satisfaction

■ Actual ━ Target

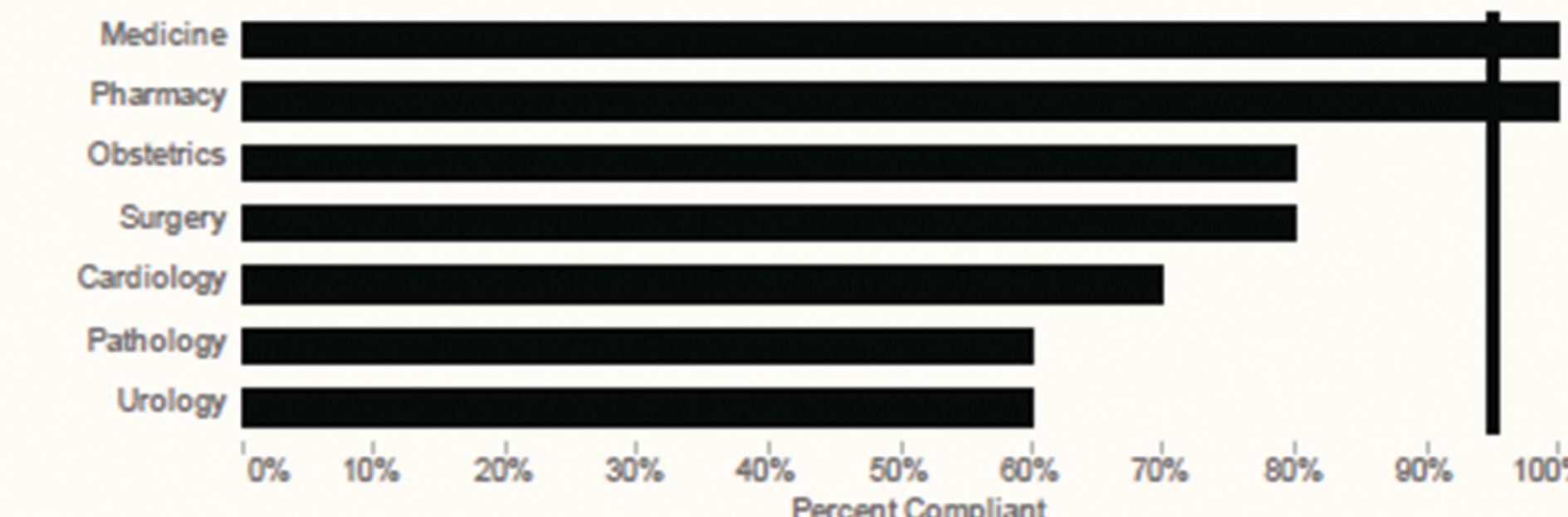


Mortality O/E Ratio



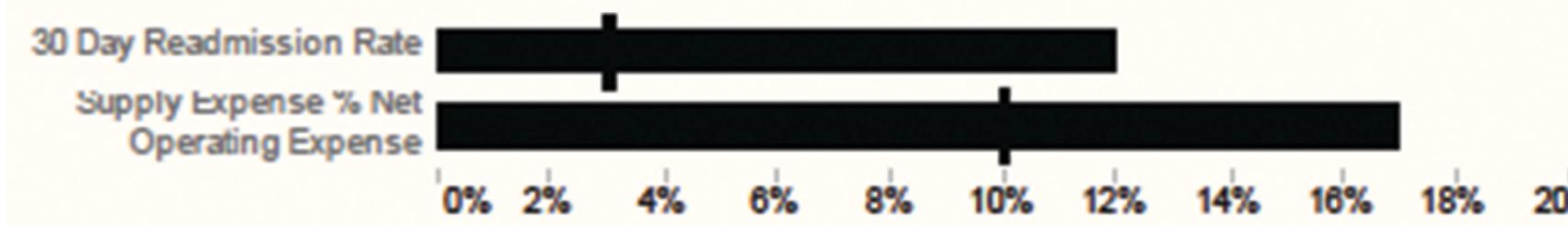
EHR Compliance

■ Compliant ━ Target



Hospital Specific Key Initiatives

■ Current Rate ━ Target



Dados multivariados

Profa. Dra. Raquel Minardi

Departamento de Ciência da Computação

Universidade Federal de Minas Gerais

- O objetivo principal por trás de análises multivariadas são comparações complexas entre múltiplas instâncias sob o ponto de vista de múltiplas variáveis

EXEMPLO

- Comparação entre automóveis
 - Preço
 - Consumo
 - Velocidade
 - Capacidade
 - Custo do seguro
 - Custo de manutenção
 - Taxa de satisfação dos clientes

COMPARAÇÃO ENTRE PERFIS MULTIVARIADOS

- Que tipo de veículo é mais apropriado para cada tipo de consumidor
- Que itens são mais parecidos?
- Que itens são excepcionais?
- Como os itens podem ser combinados de acordo com sua similaridade?

REPRESENTAÇÕES VISUAIS

.....

TIPOS REPRESENTAÇÕES VISUAIS

- Projeções geométricas
- Representações hierárquicas
- Representações iconográficas
- Técnicas orientadas a pixels

PROJEÇÕES GEOMÉTRICAS

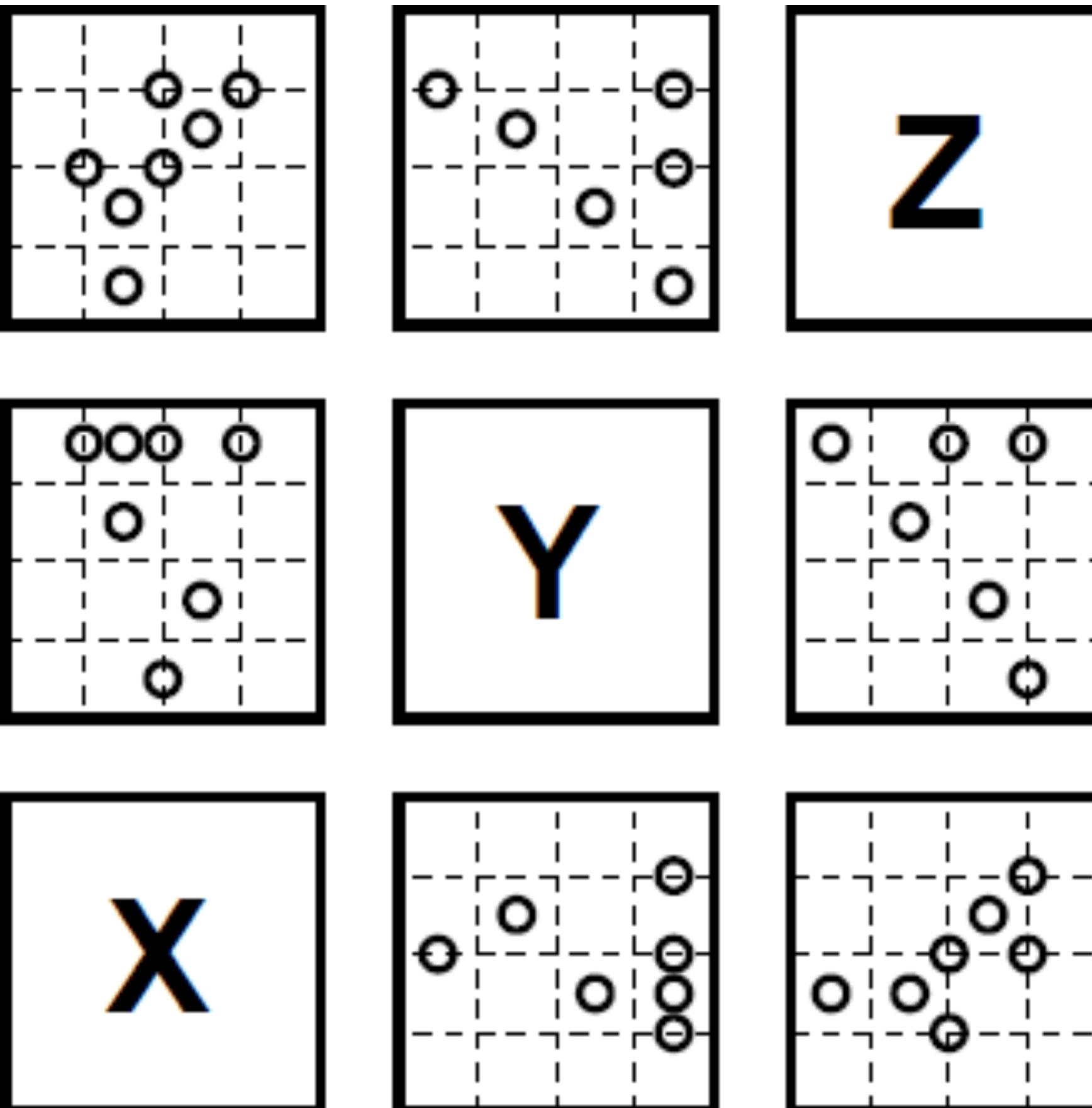
.....

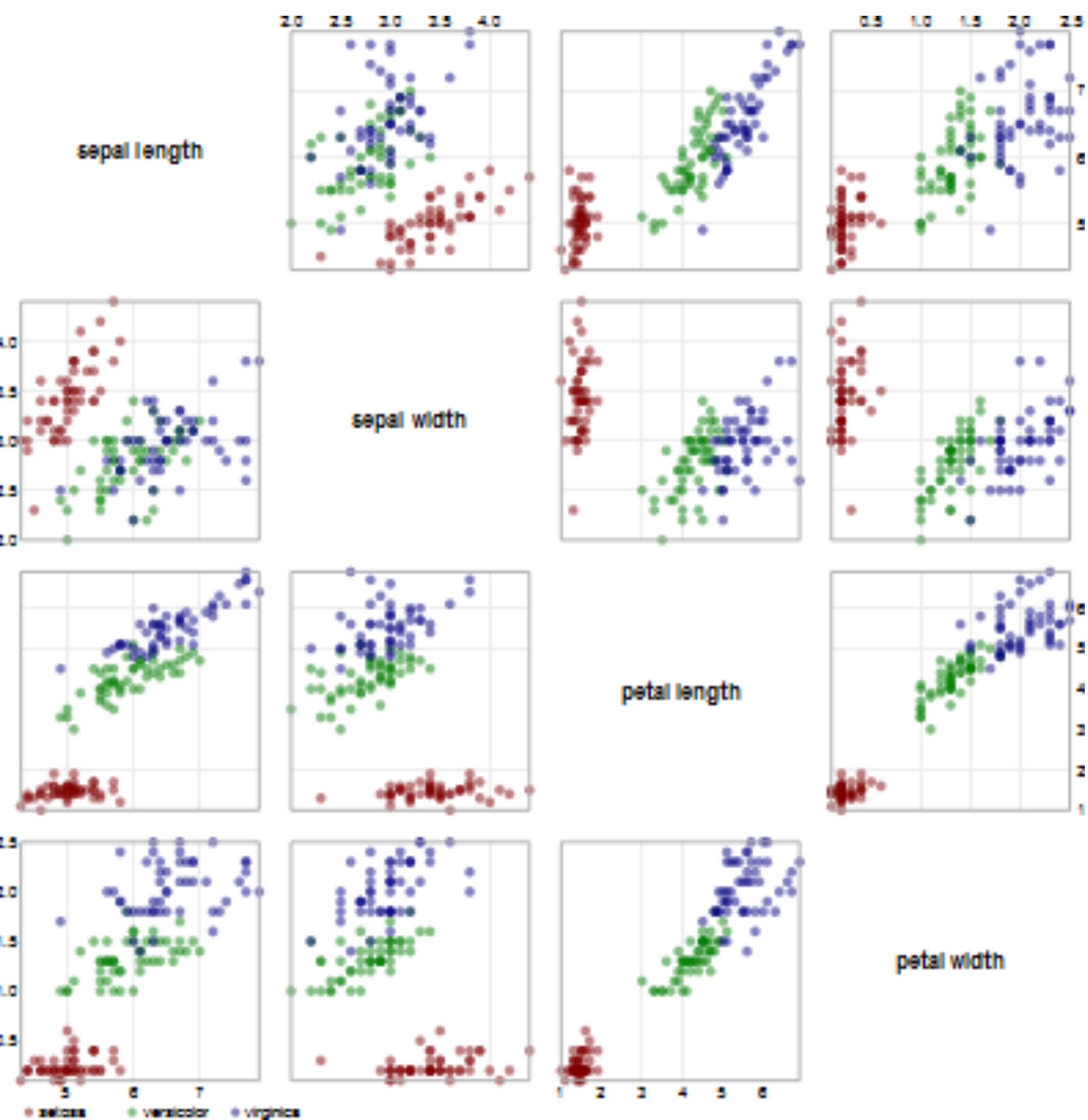
PROJEÇÃO GEOMÉTRICA

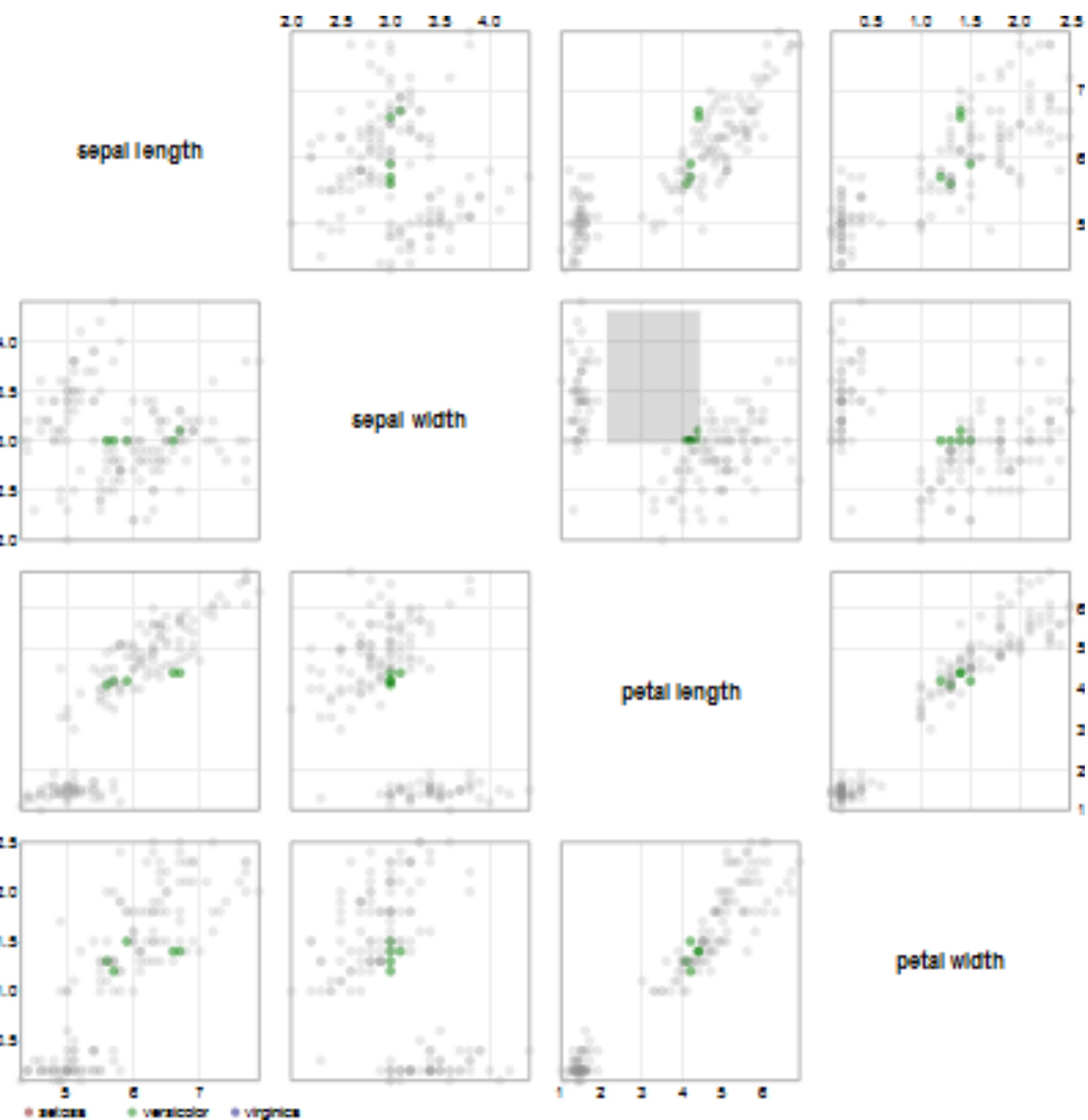
- Busca de projeções 2D para dados multidimensionais
- Úteis na detecção de exceções e ainda de correlações entre diferentes dimensões

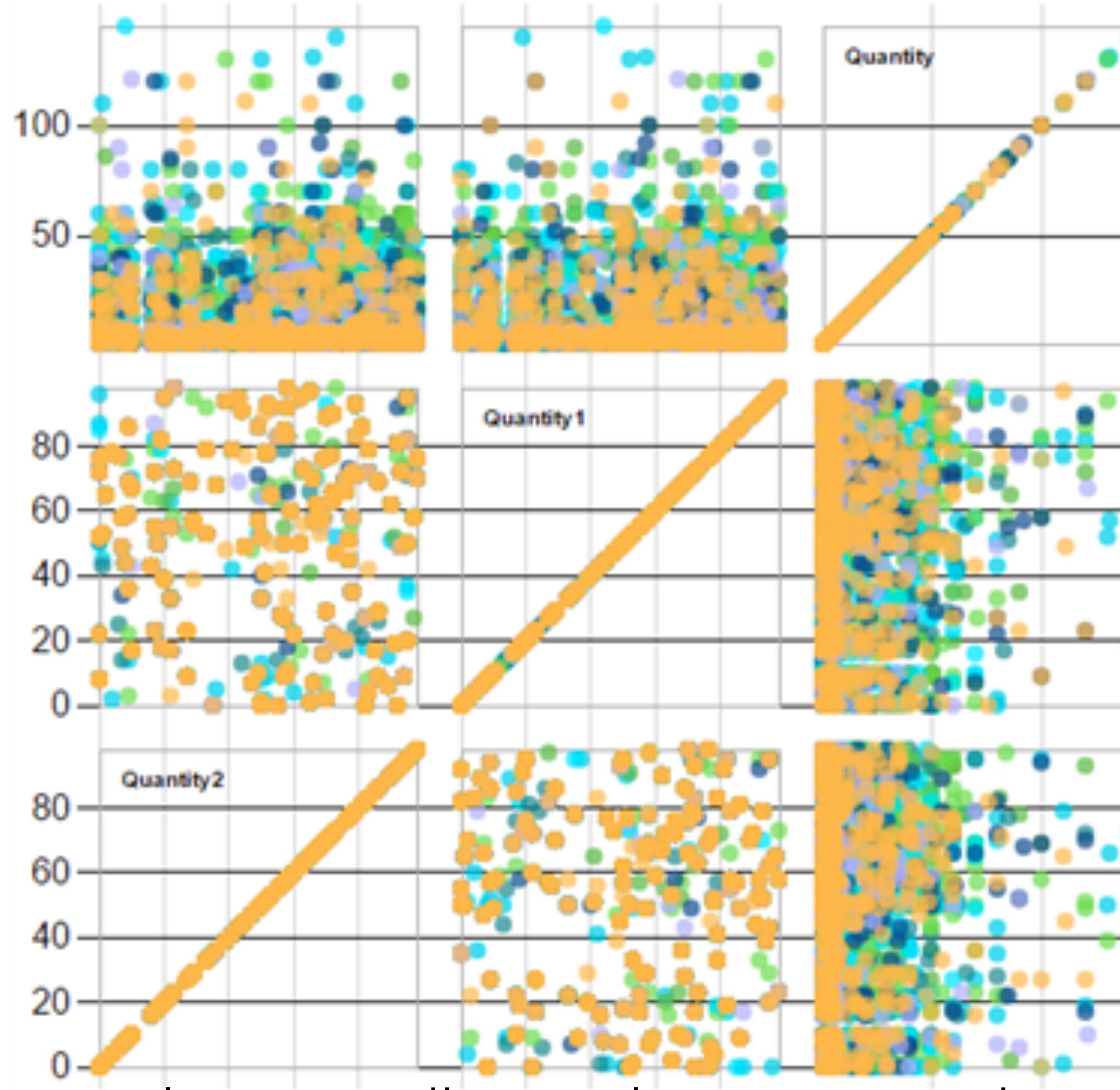
MATRIZES DE GRÁFICOS DE DISPERSÃO

- A matriz de scatterplots é útil para mostrar correlações entre pares de variáveis
- O uso das técnicas de brushing e linking são muitos úteis para a identificação dos itens nos diversos plots



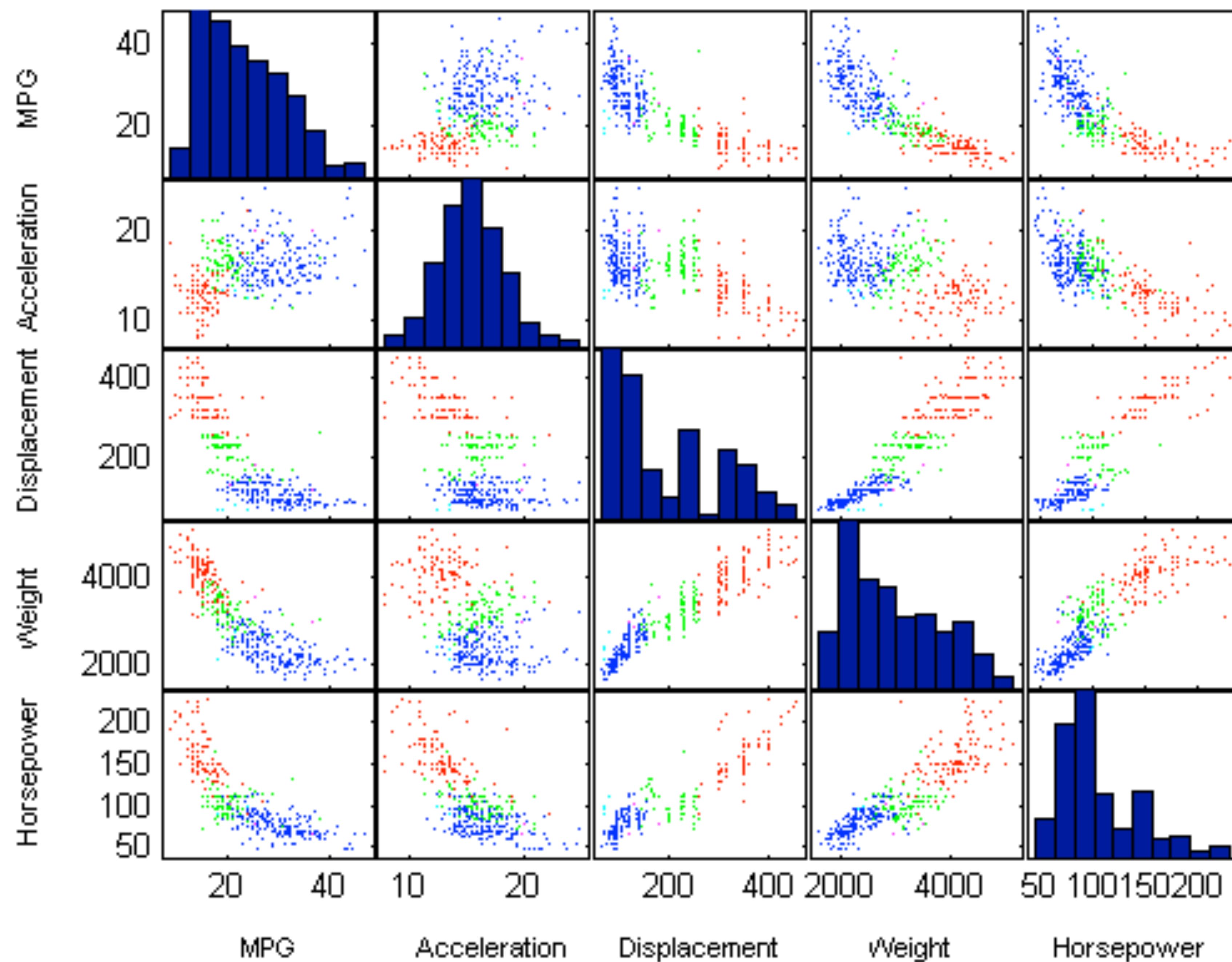




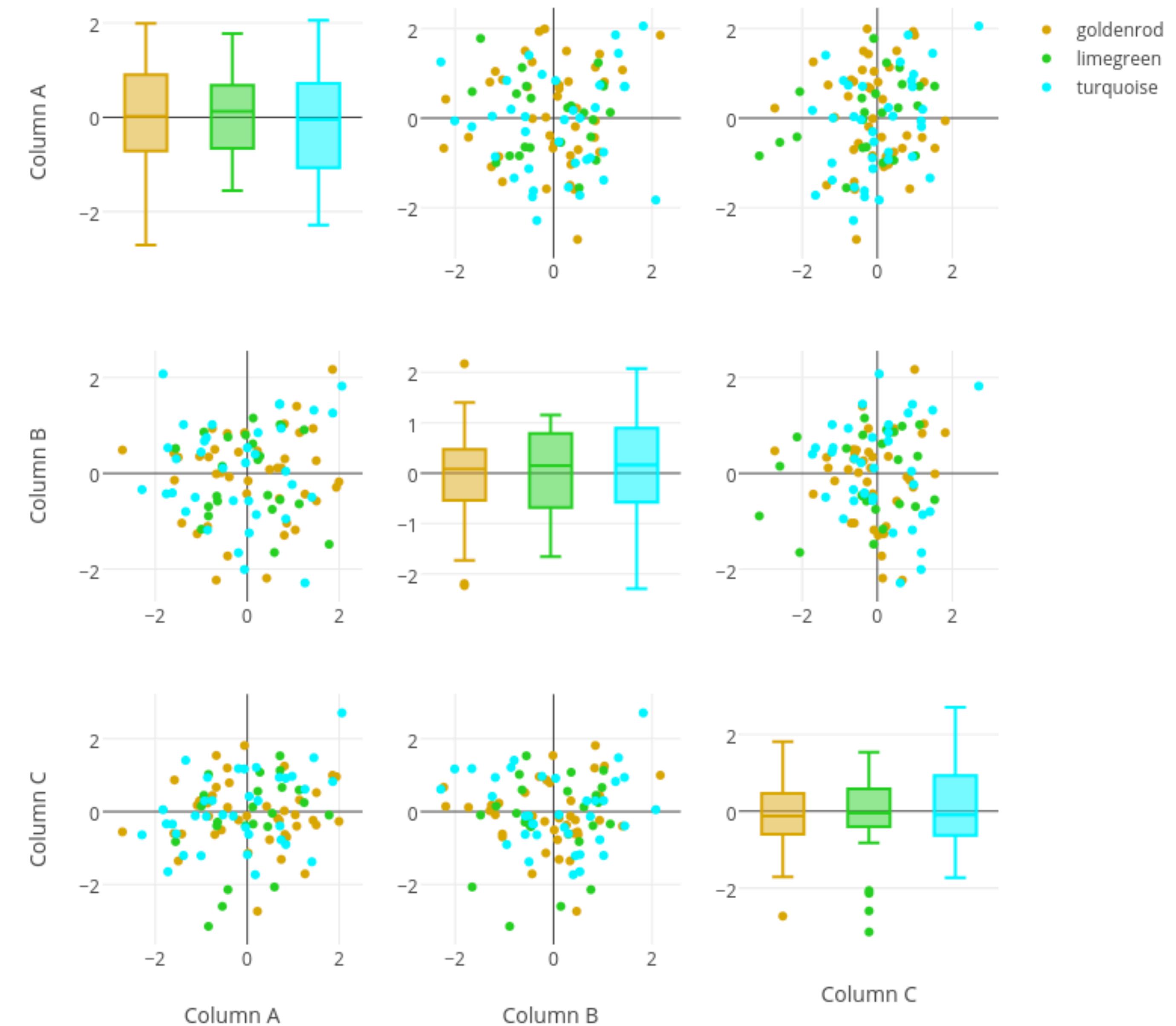


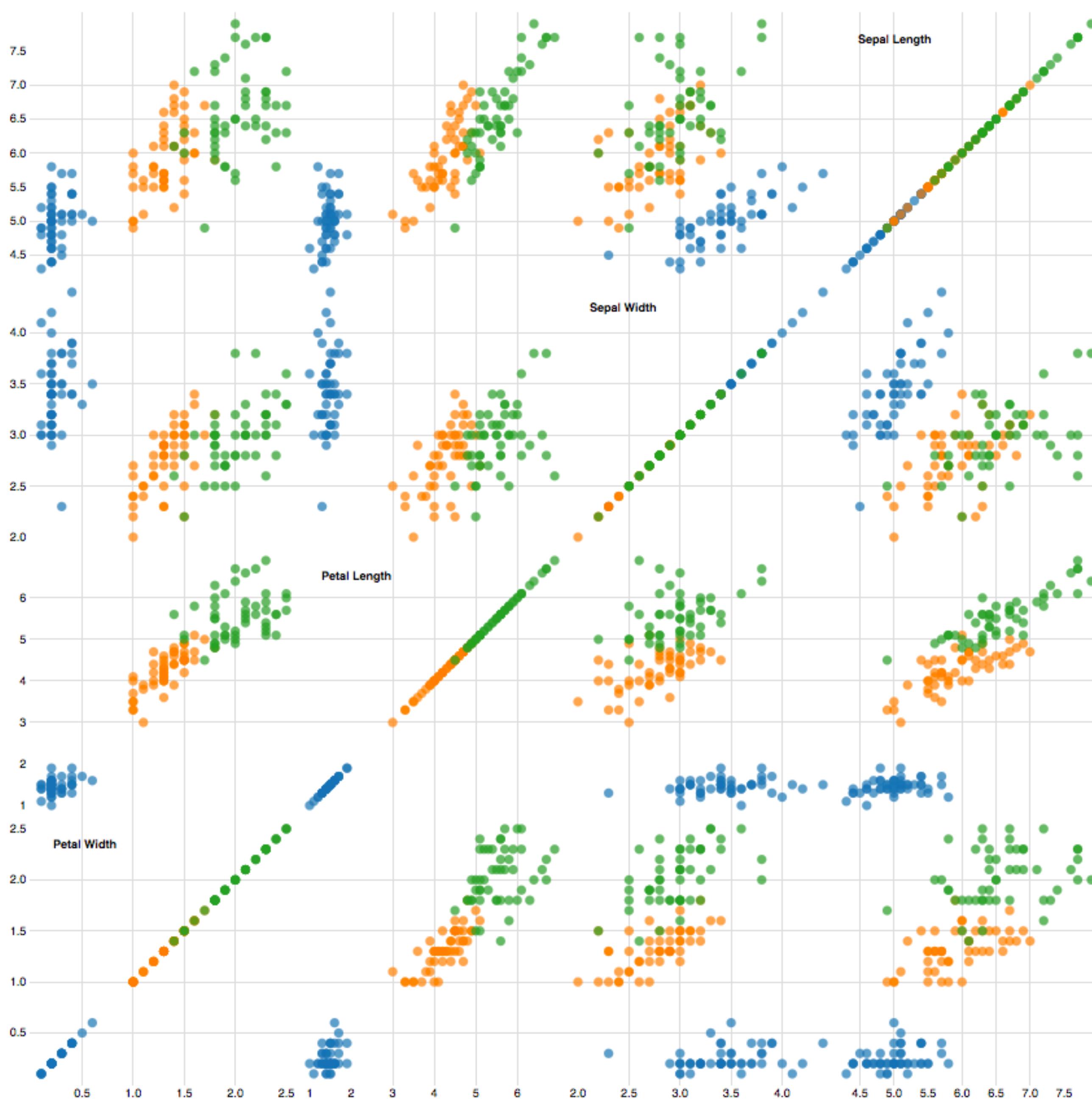
O que colocar na diagonal para aproveitar melhor o espaço com dados?

(princípio Data-ink de Tufte)



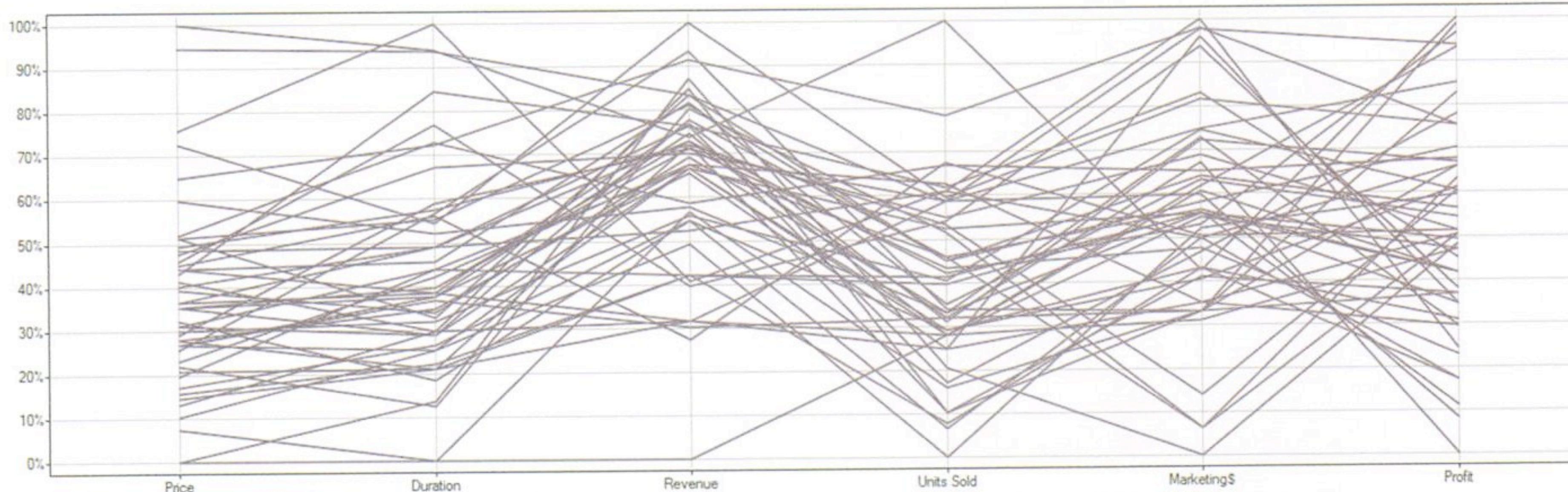
Scatterplot Matrix





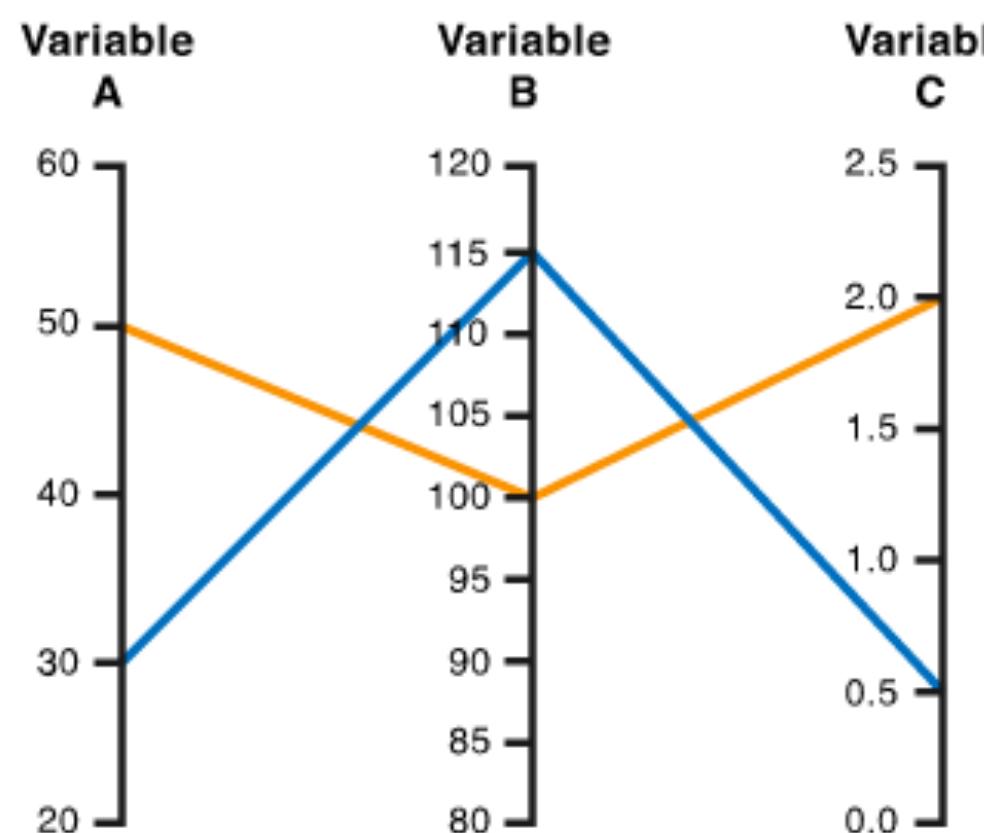
COORDENADAS PARALELAS

- Para representar um conjunto de pontos em um espaço n-dimensional, um plano é desenhado e sobre ele n linhas verticais paralelas e igualmente espaçadas.
- Um ponto no espaço n-dimensional é representado como uma polilinha com vértices nos eixos paralelos
- A posição do vértice no i-ésimo eixo corresponde à i-ésima coordenada do ponto.

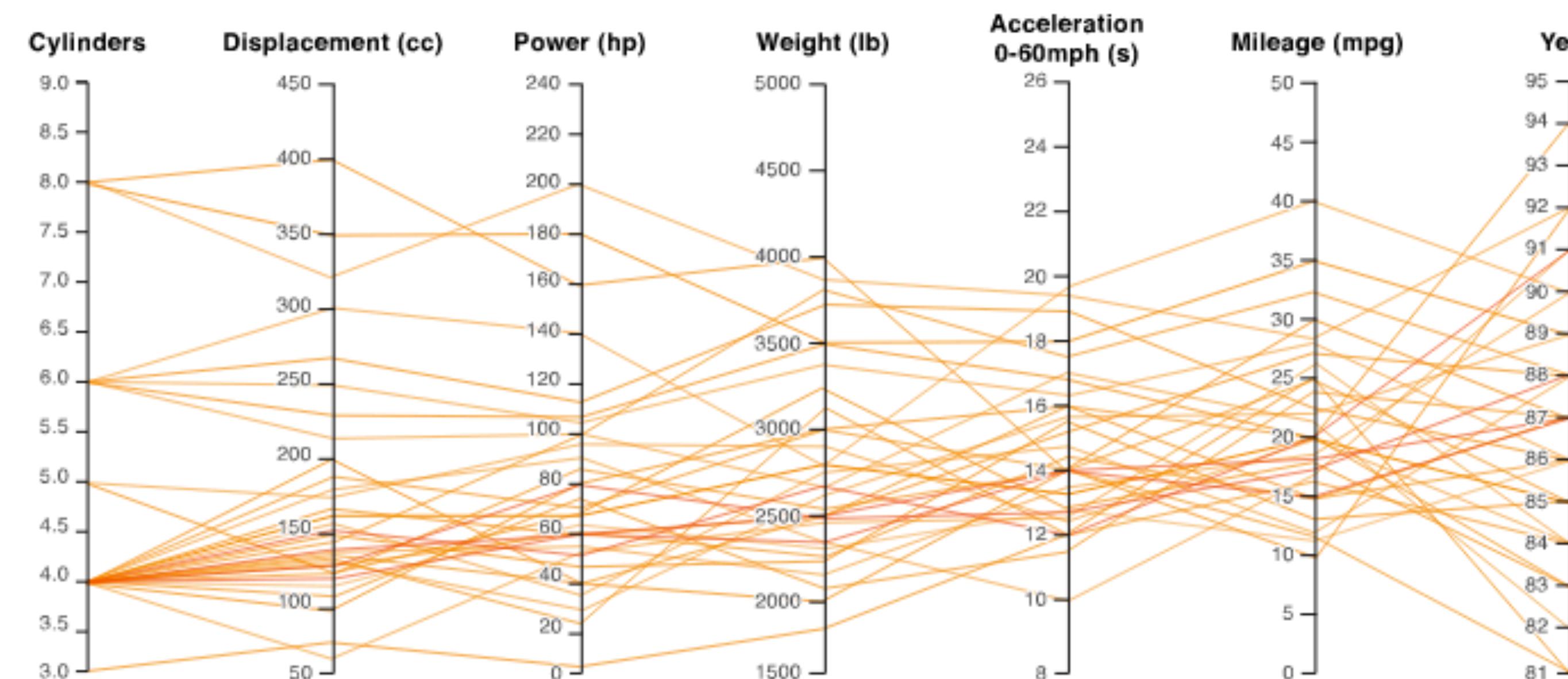


COORDENADAS PARALELAS

- Para representar um conjunto de pontos em um espaço n-dimensional, um plano é desenhado e sobre ele n linhas verticais paralelas e igualmente espaçadas.
- Um ponto no espaço n-dimensional é representado como uma polilinha com vértices nos eixos paralelos
- A posição do vértice no i-ésimo eixo corresponde à i-ésima coordenada do ponto.

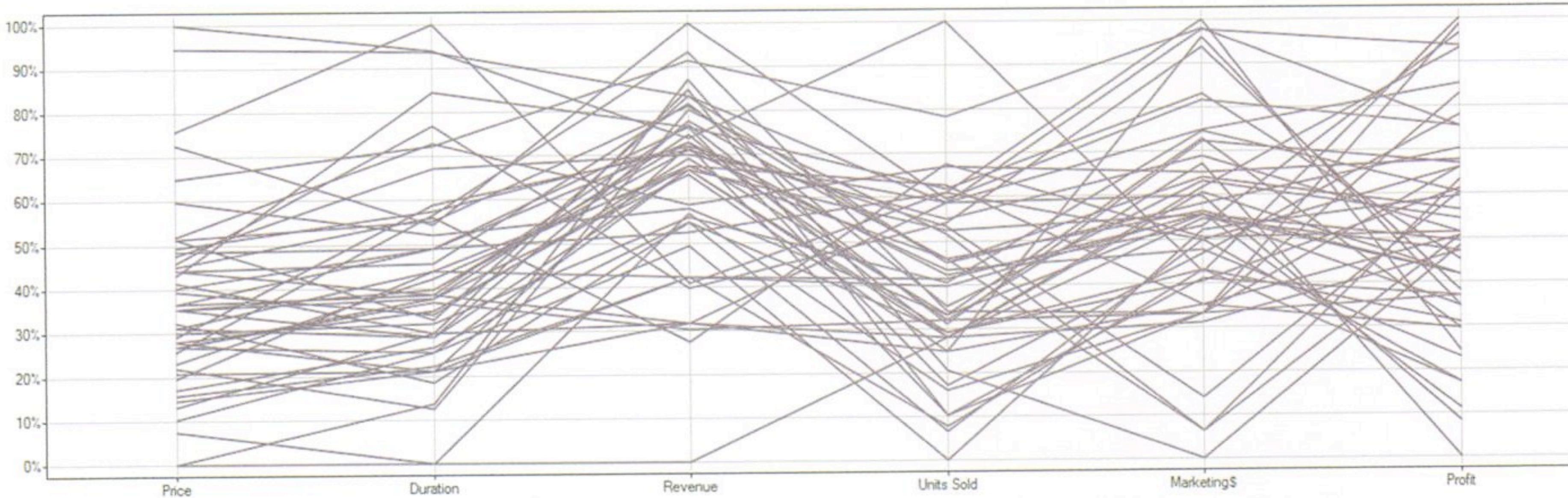


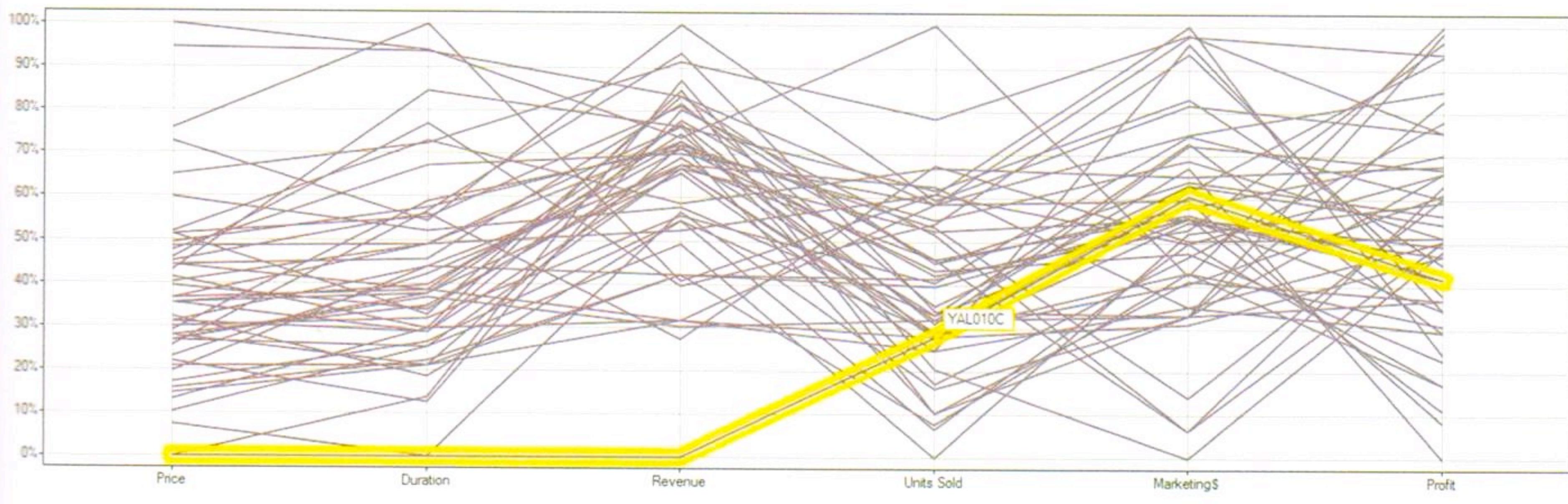
Data			
	Variable A	Variable B	Variable C
Item 1	50	100	2.0
Item 2	30	115	0.5



COORDENADAS PARALELAS

- São gráficos extremamente complexos
- São capazes de mostrar alta densidade de dados
- Apesar de cada coluna representar uma variável diferente, a conexão entre elas revela o padrão multivariado



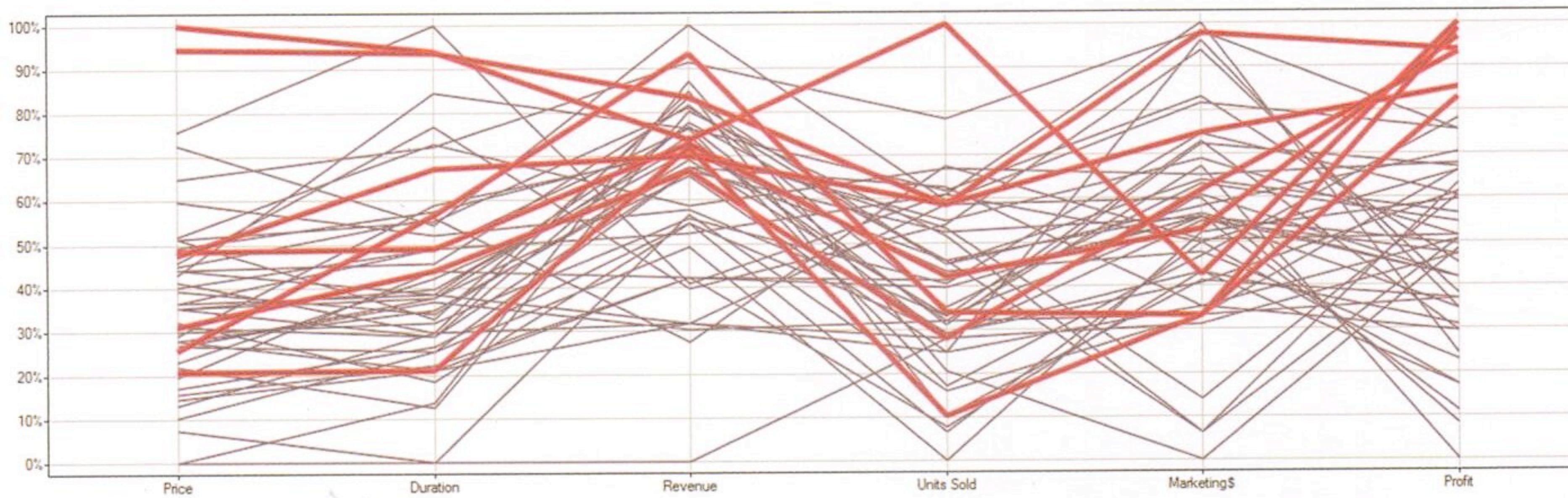


- O produto em destaque tem a menor renda que todos os outros (talvez devido ao baixo preço e ciclo de vida curto)
- Um produto é muito mais vendido que o seu competidor mais próximo
- A maioria dos produtos está no mercado entre 20% e 60% do tempo total avaliado
- Há uma concentração das rendas entre 65% e 85%

- Como se pode perceber, mesmo com o grande volume de curvas, padrões, tendências e exceções podem ser identificados
- Análises mais profundas dependem de interações com a visualização

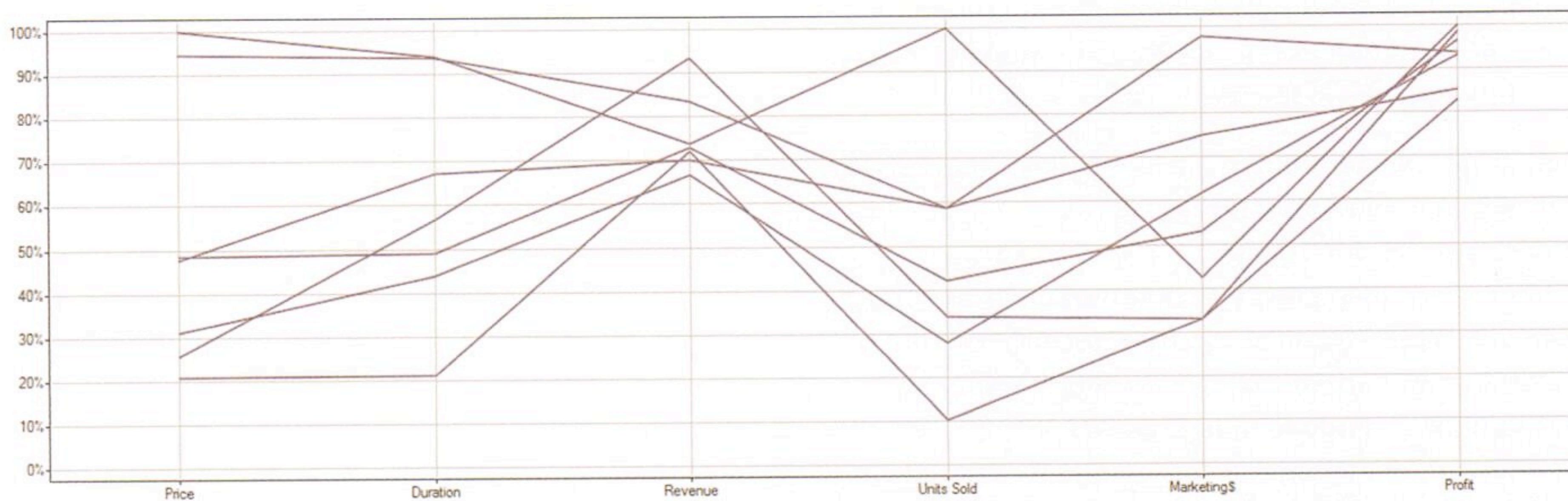
EXEMPLO

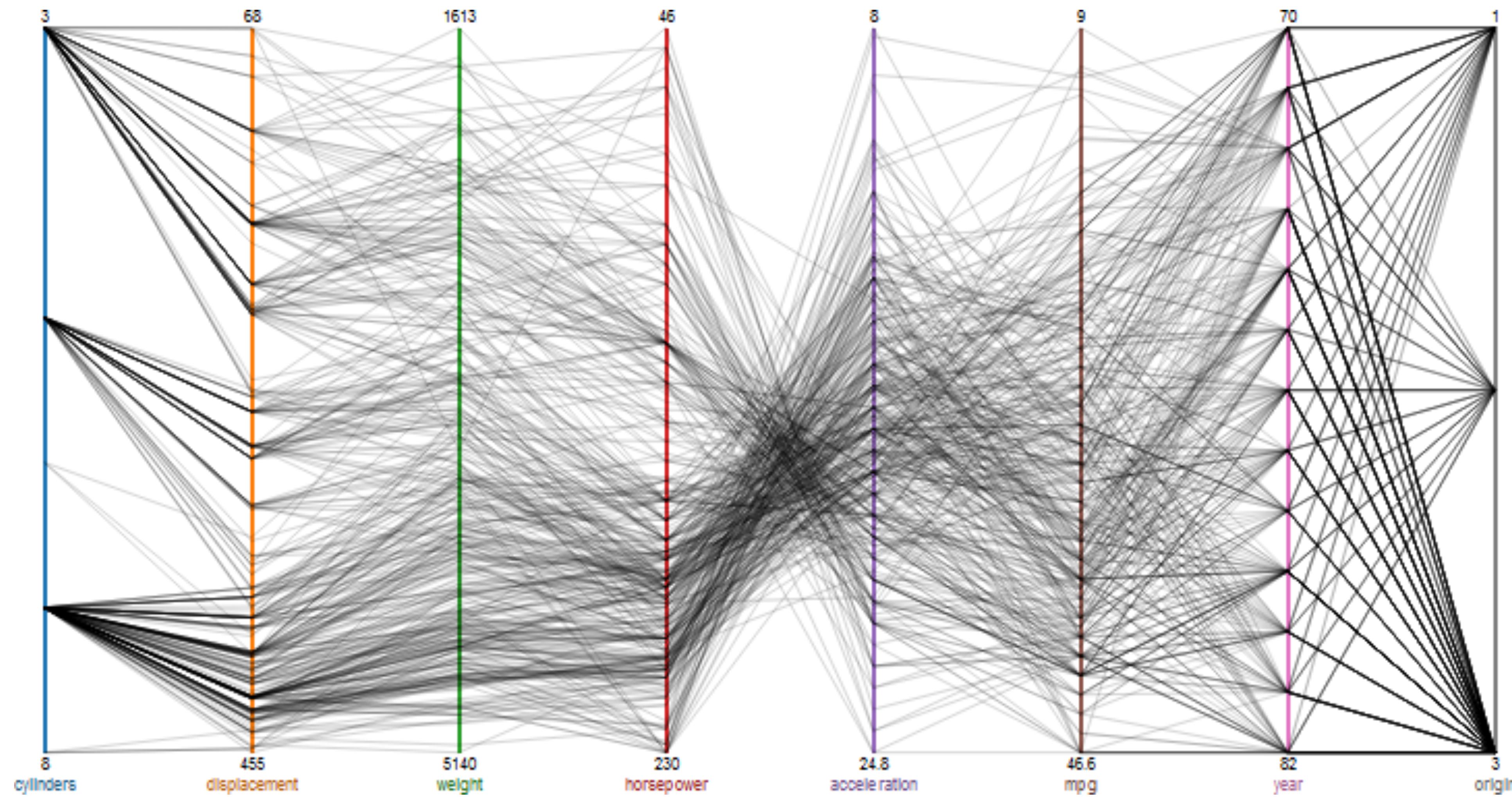
- Selecionar todos os produtos com lucros acima de 80%

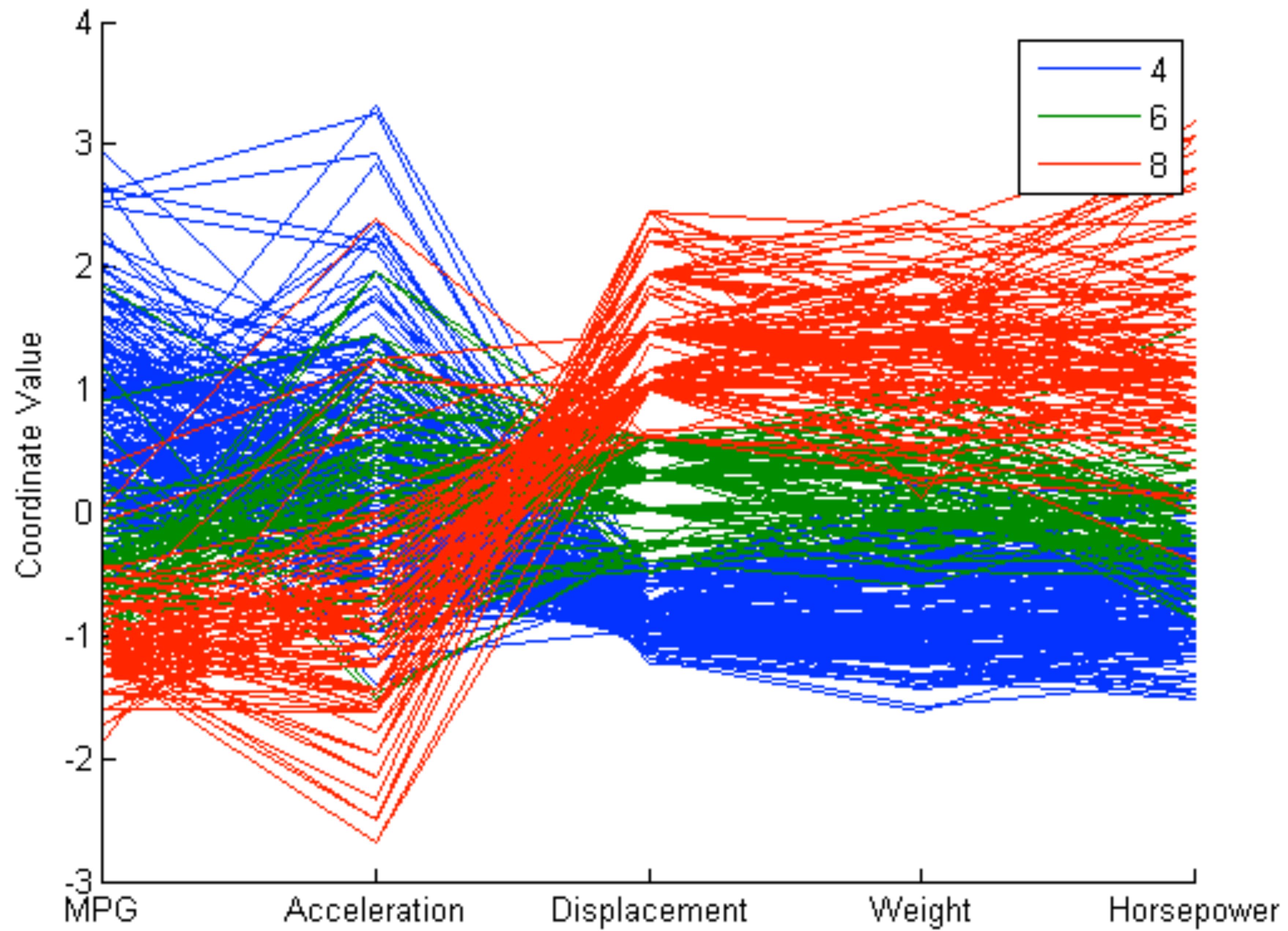


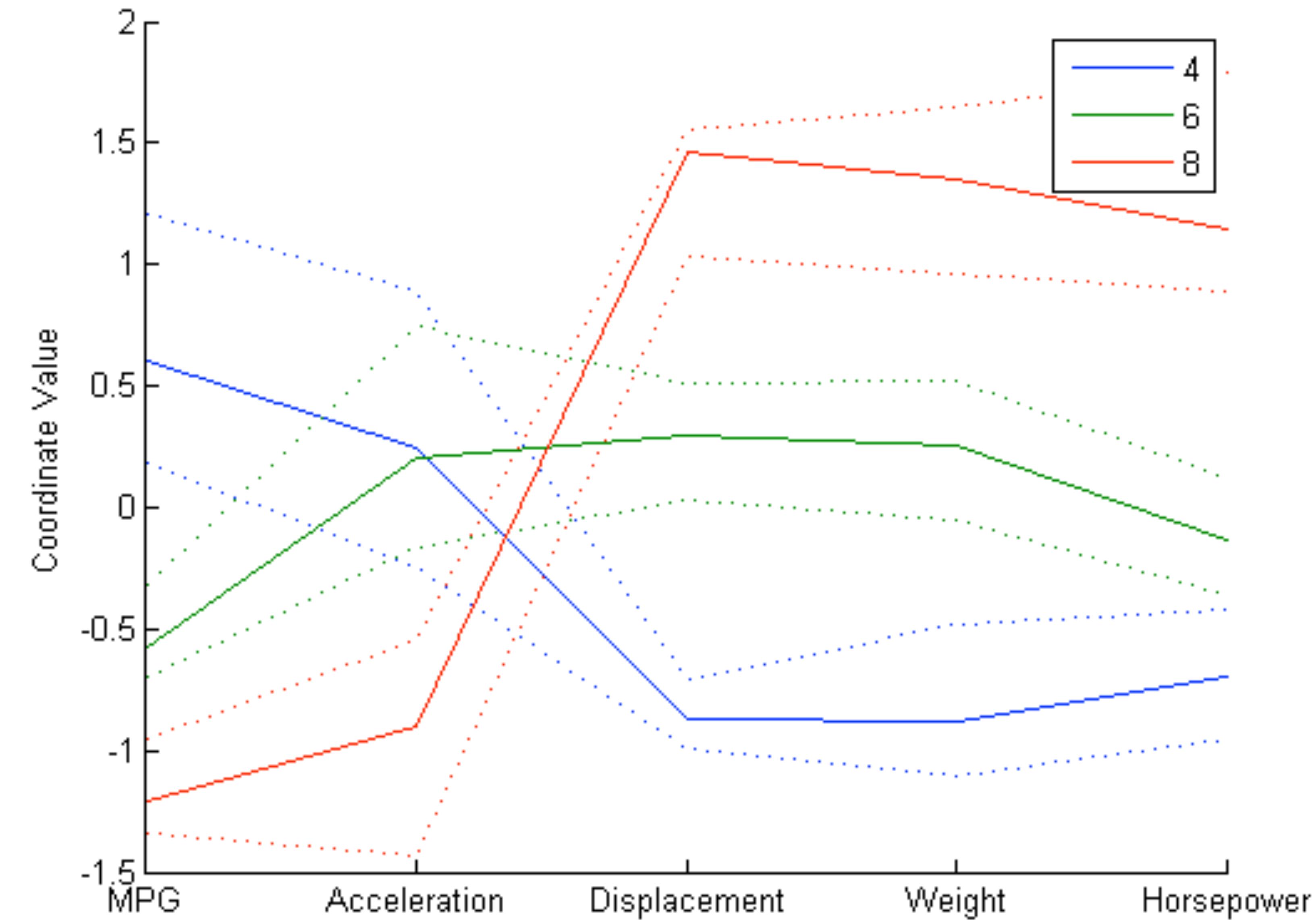
EXEMPLO

- Todos os produtos com alto lucro geram alta renda
- Estes produtos apresentam sempre despesas com marketing acima de 30%







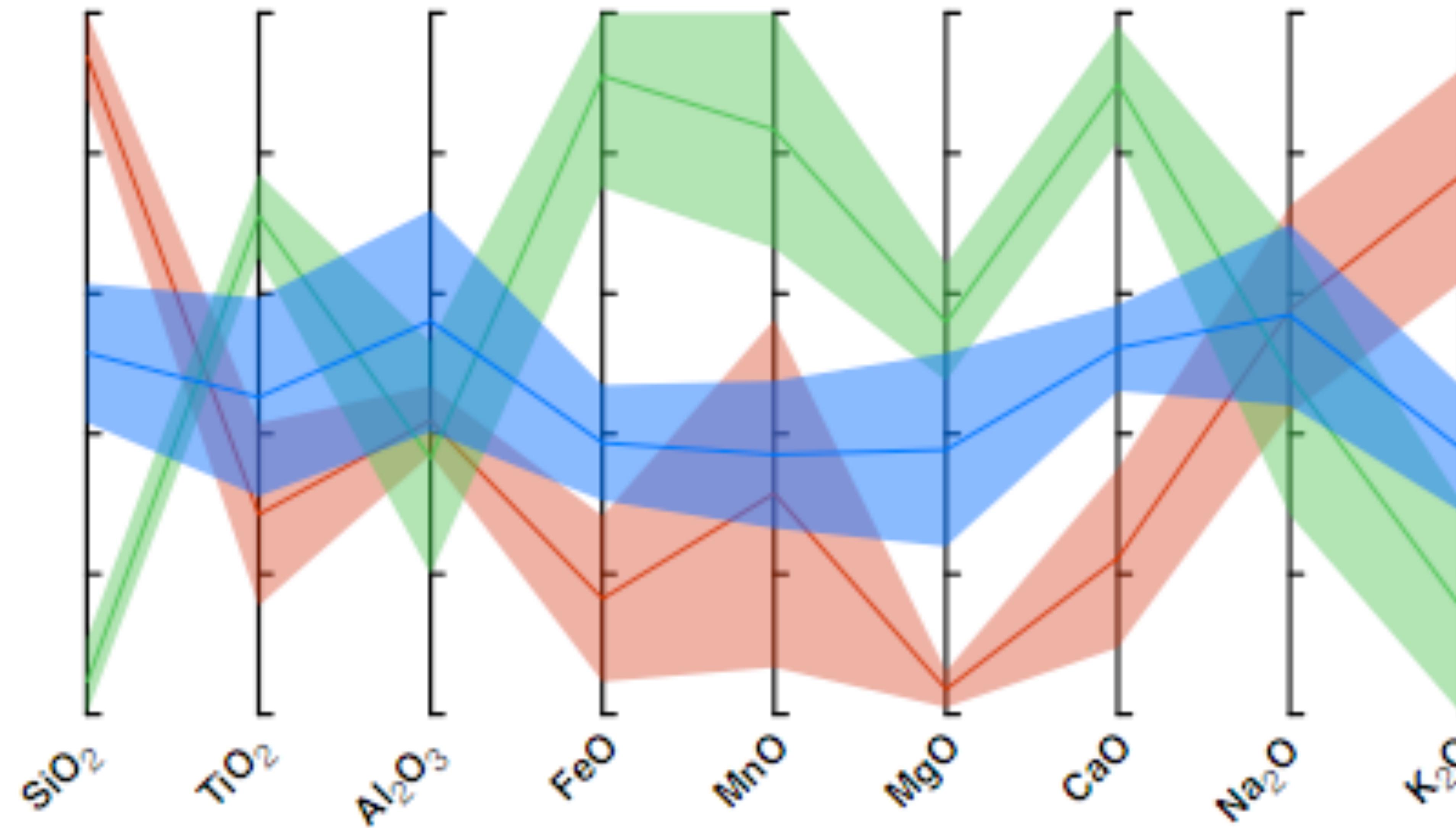


A linha sólida indica a mediana e as linhas pontilhadas, o 25° e o 75° percentil

 Quartz-Rich Samples

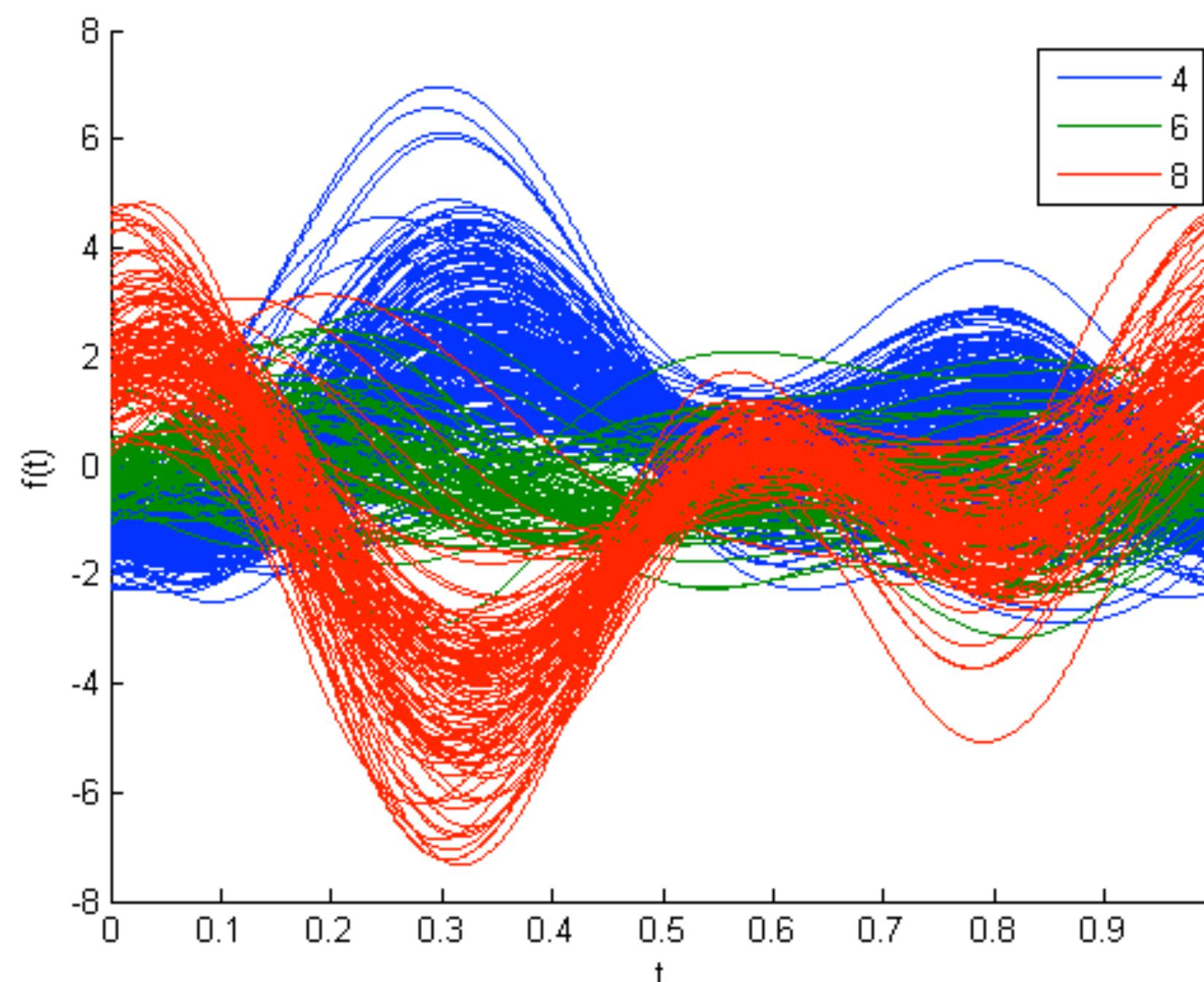
 Silica-Undersaturated Samples

 Intermediate Samples

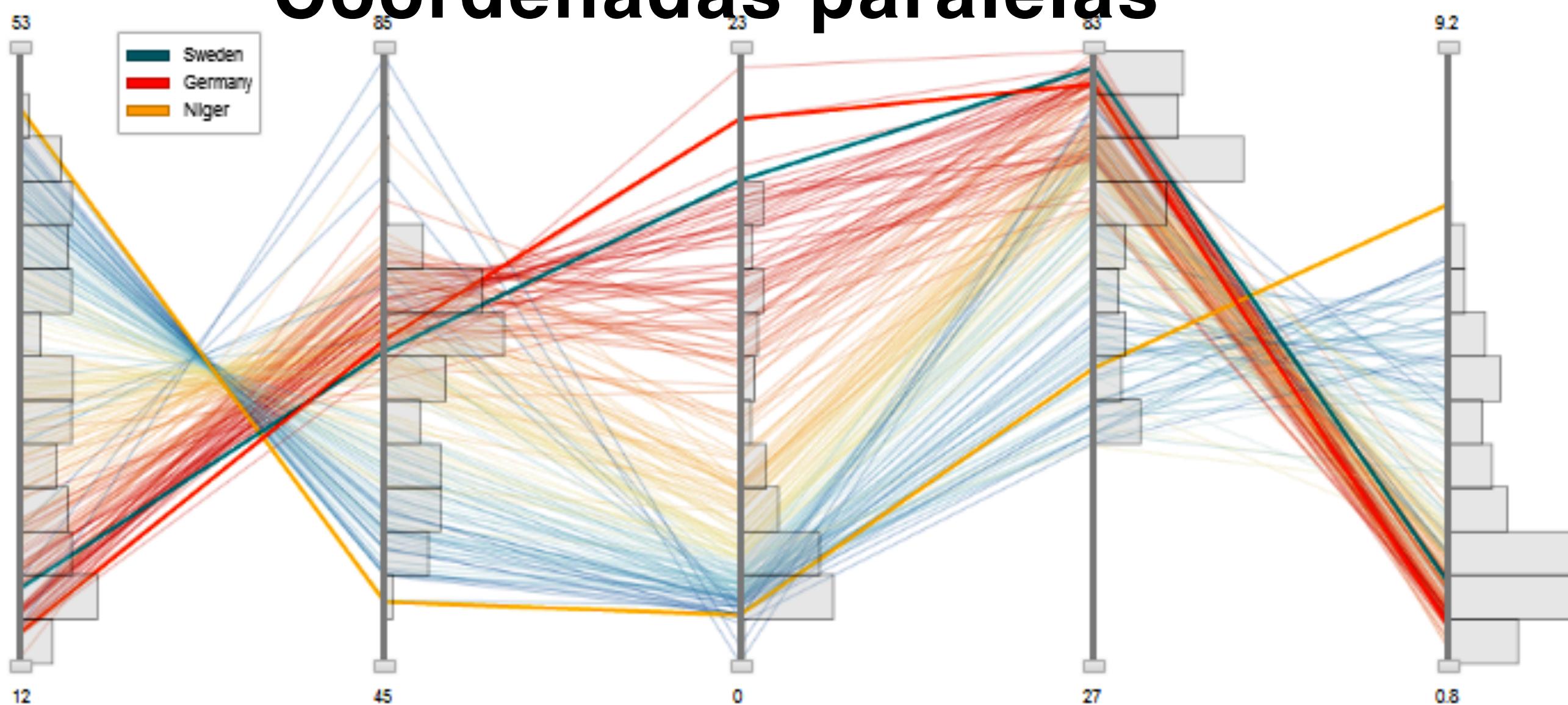


Andrew curves

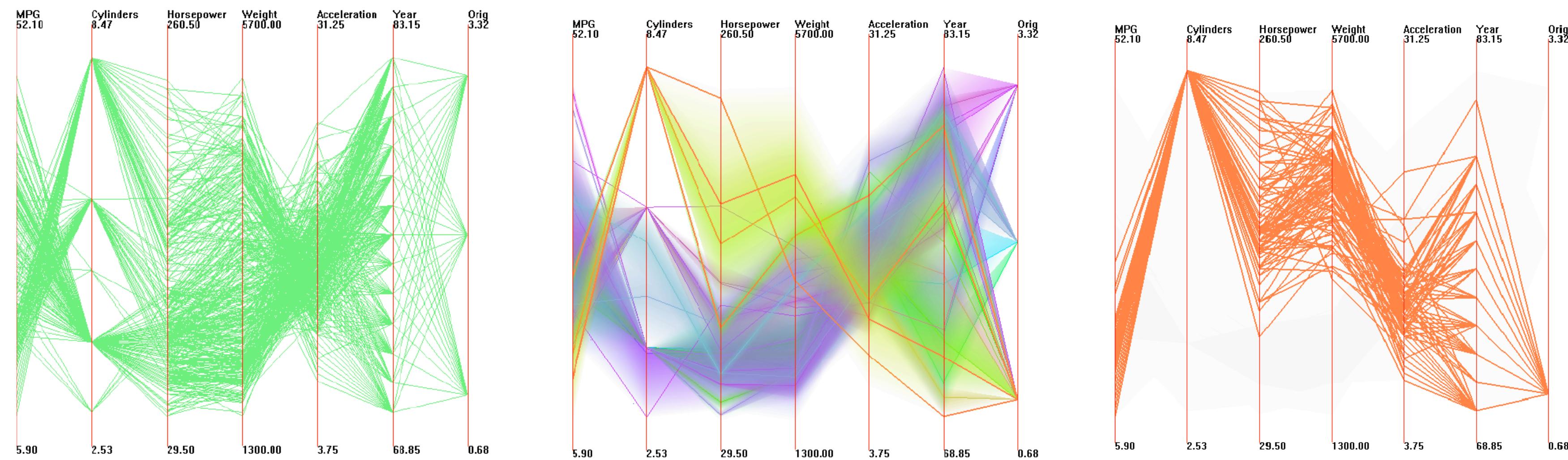
Cada ponto é representado por uma curva obtida por uma transformada de Fourier



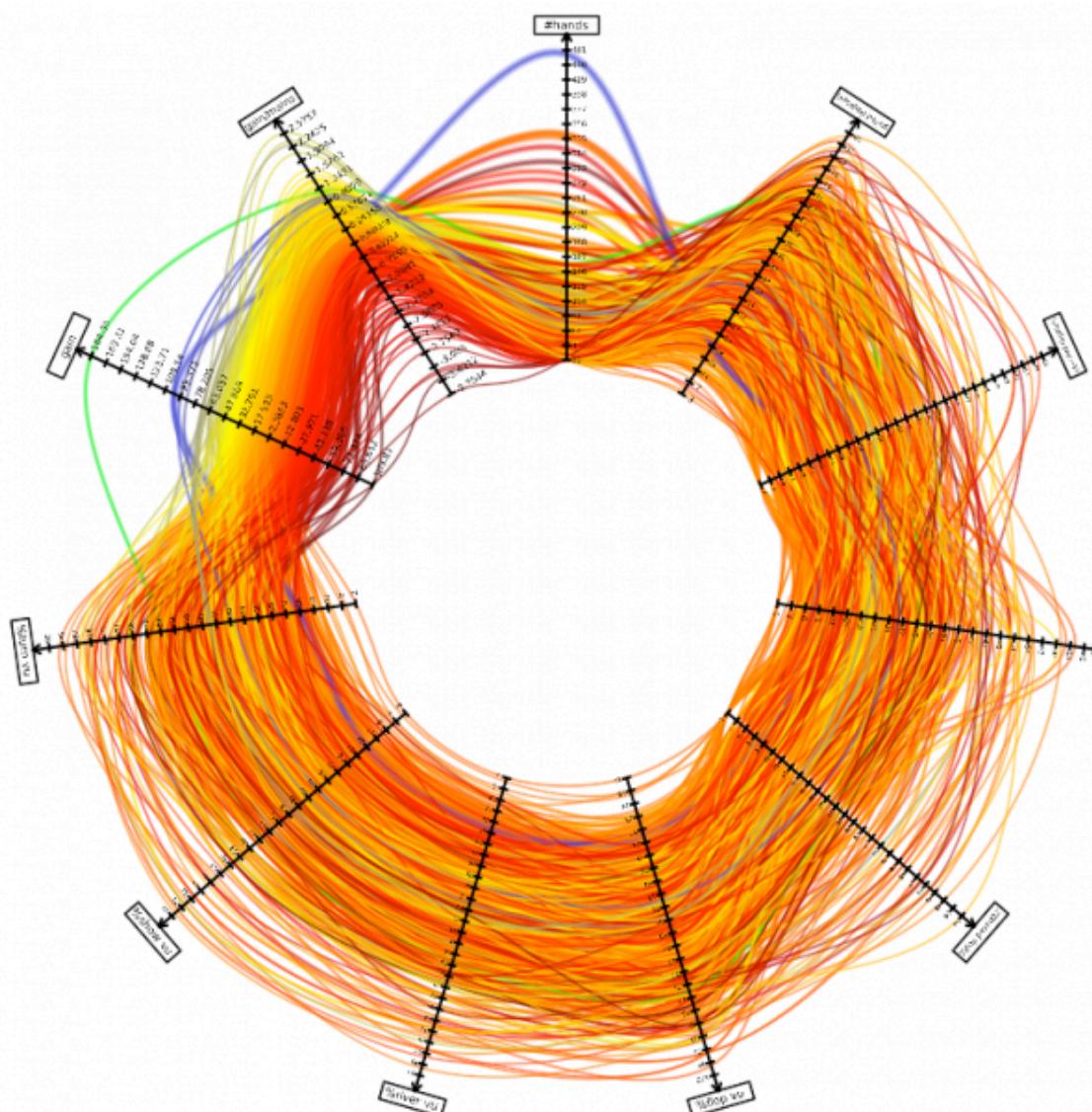
Coordenadas paralelas



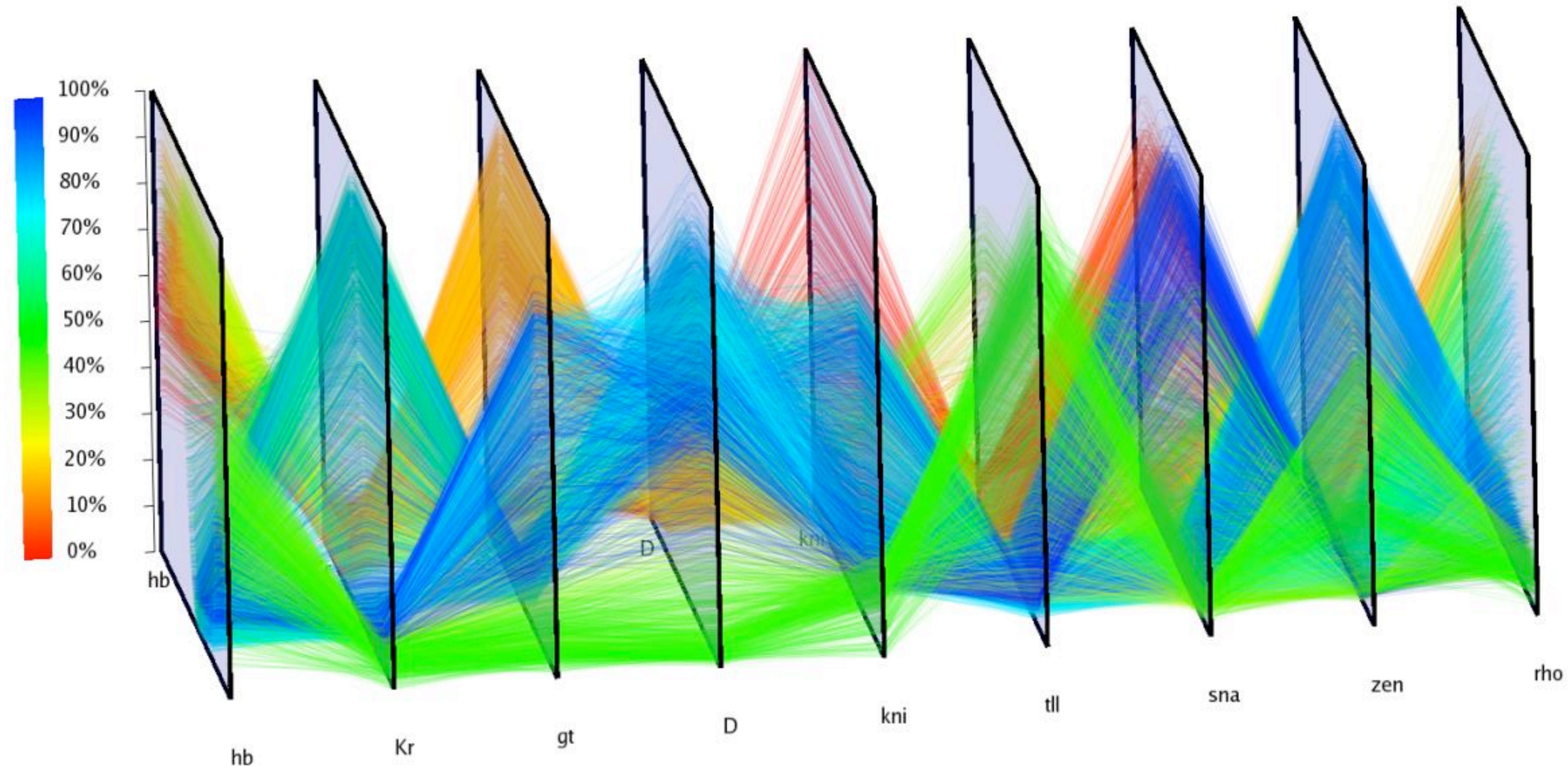
Coordenadas paralelas



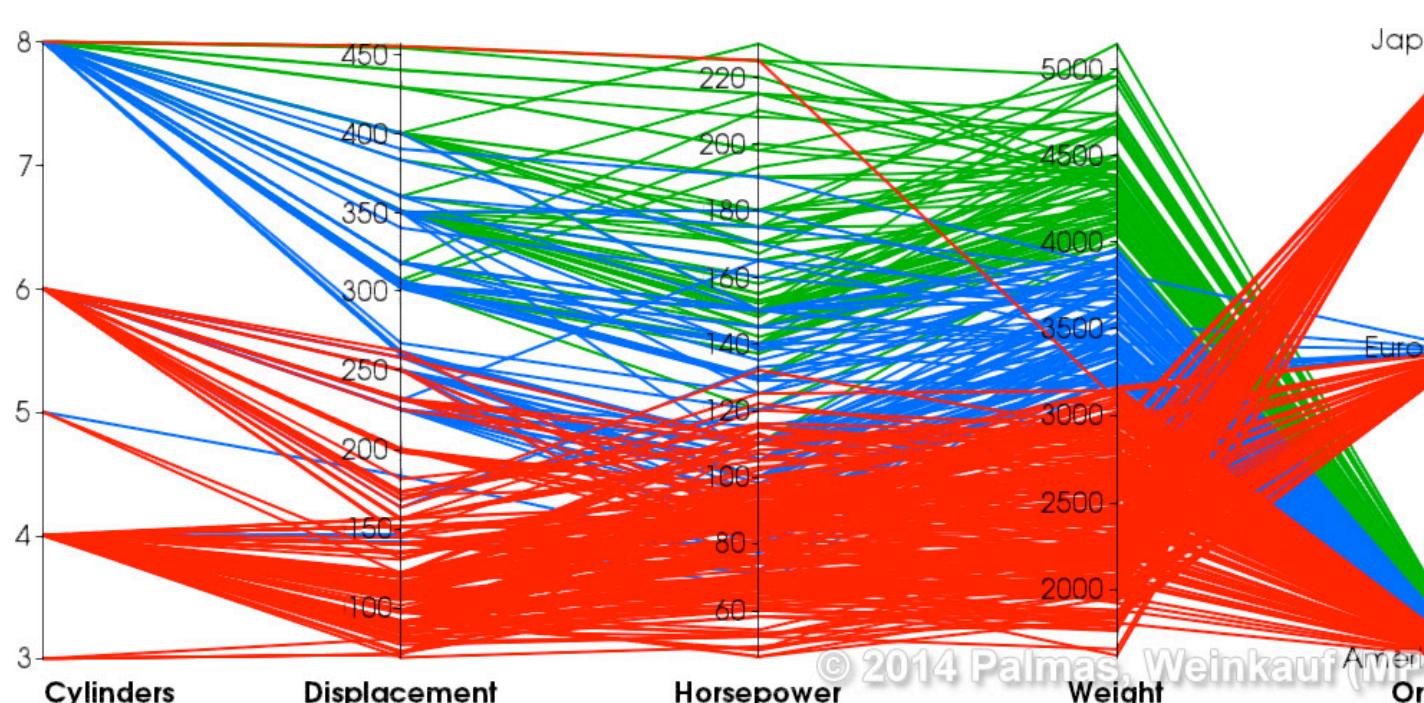
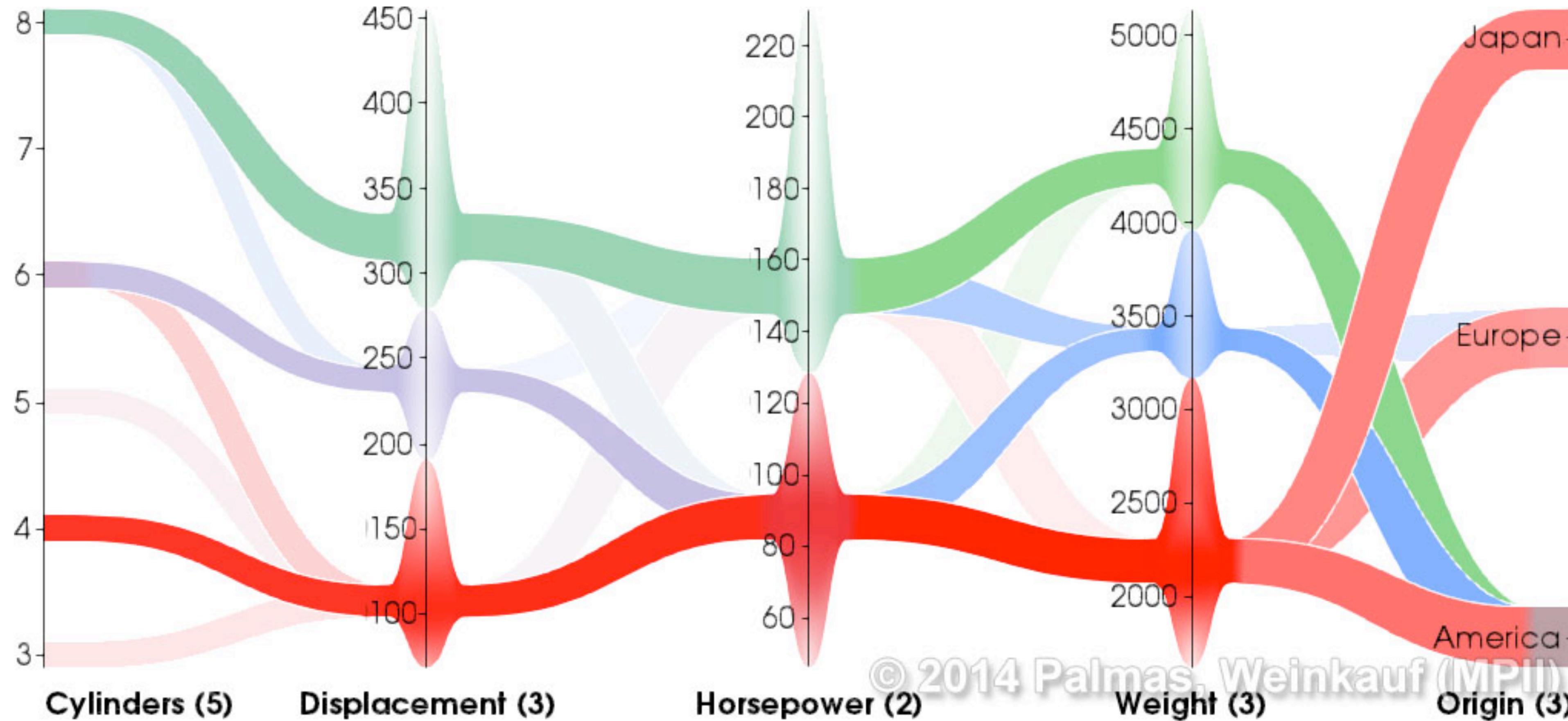
Coordenadas paralelas



Coordenadas paralelas



Coordenadas paralelas com



G. Palmas, M. Bachynskyi, A. Oulasvirta, H.-P. Seidel, T. Weinkauf
An Edge-Bundling Layout for Interactive Parallel Coordinates
Proc. IEEE PacificVis, Yokohama, Japan, March 4 - 7, 2014

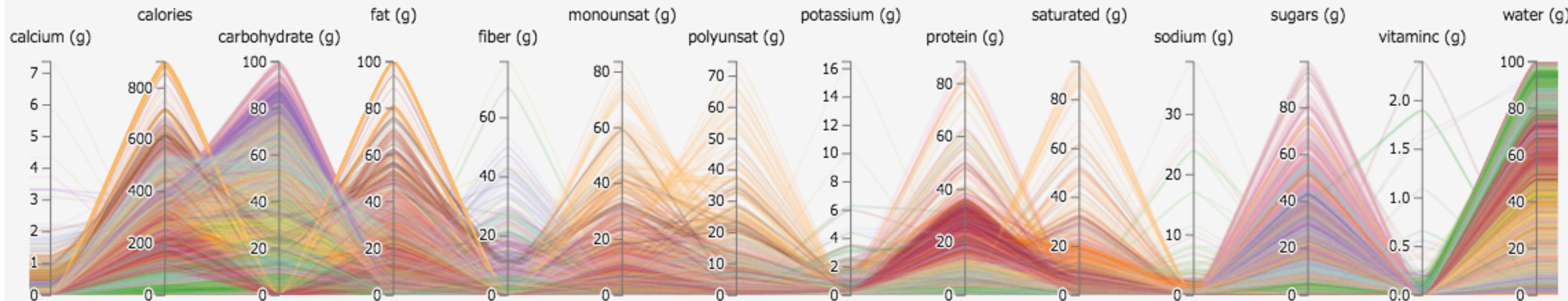
Nutrient Explorer

Keep Exclude Export

7637/7637

Lines at 13.6% opacity.

Hide Ticks Dark



What is this?

A multidimensional explorer of [nutrient data](#) from the [USDA](#).

The parallel coordinates displays the nutrient content of foods in the database across 14 dimensions, colored by food group.

Never heard of parallel coordinates? Read this [tutorial](#).

[Brush](#) the visualization to update other charts on this page.

Let me know what you think on [Reddit](#).

Controls

Brush: Drag vertically along an axis.

Remove Brush: Tap the axis background.

Reorder Axes: Drag a label horizontally.

Invert Axis: Tap an axis label.

Remove Axis: Drag axis label to the left edge.

Credits & License

Adapted from examples by

[Mike Bostock](#) and [Jason Davies](#)

Copyright © 2012, Kai Chang

All rights reserved. Released under the [BSD License](#).

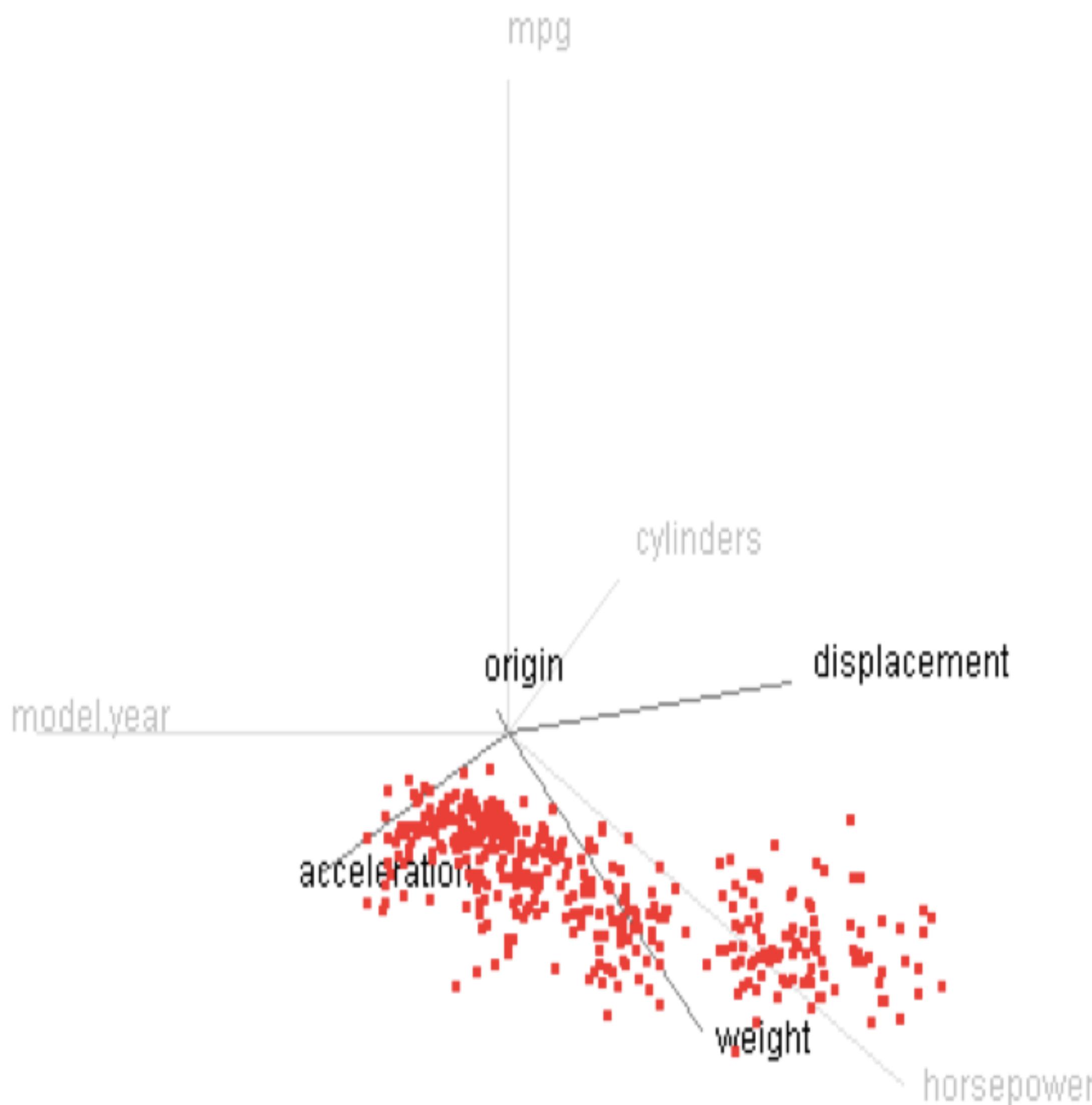
Food Groups

327 Baby Foods
496 Baked Products
619 Beef Products
278 Beverages
403 Breakfast Cereals
183 Cereal Grains and Pasta
239 Dairy and Egg Products
165 Ethnic Foods
365 Fast Foods
215 Fats and Oils
255 Finfish and Shellfish Products
328 Fruits and Fruit Juices
345 Lamb, Veal, and Game Products
365 Legumes and Legume Products
57 Meals, Entrees, and Sidedishes
128 Nut and Seed Products
328 Pork Products
381 Poultry Products
51 Restaurant Foods
234 Sausages and Luncheon Meats
162 Snacks
499 Soups, Sauces, and Gravies
61 Spices and Herbs
341 Sweets

Sample of 25 entries

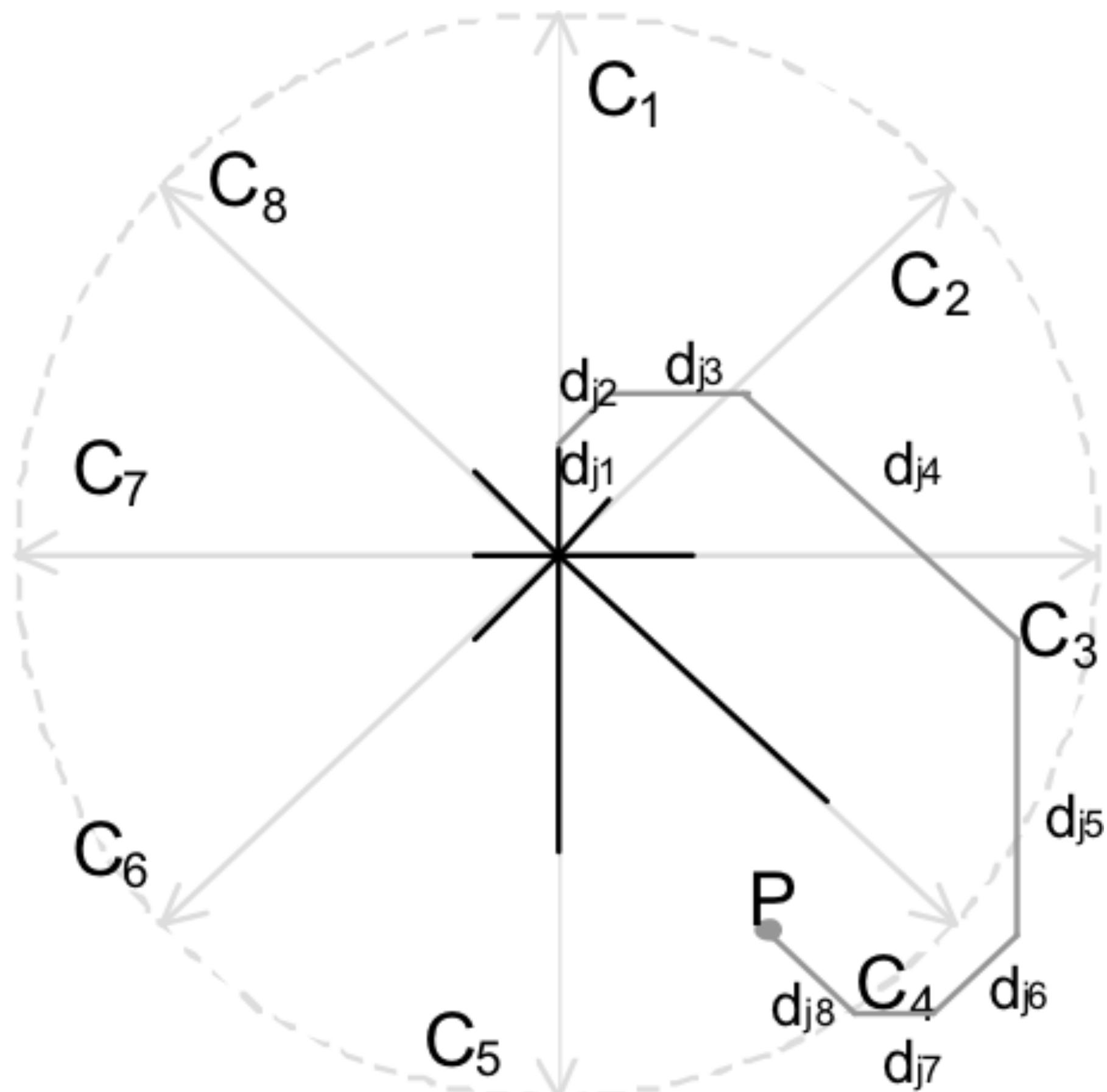
Search Foods...
ARCHWAY Home Style Cookies, Ruth's Oatmeal
Babyfood, dessert, peach cobbler, strained
Bacon and beef sticks
Beans, adzuki, mature seeds, canned, sweetened
Beef, chuck, under blade center steak, boneless, Denver Cut,
Beef, flank, steak, separable lean and fat, trimmed to 0" fat, c
Beef, round, eye of round, roast, separable lean only, trimmed
Candies, MARS SNACKFOOD US, TWIX chocolate fudge cookie
Cereals ready-to-eat, POST, Honeycomb Cereal
Cereals, QUAKER, Creamy Wheat, farina, enriched, prepared
Cookies, oatmeal, prepared from recipe, without raisins
Cowpeas, catjang, mature seeds, cooked, boiled, with salt
Cranberry juice cocktail, frozen concentrate, prepared with wa
Fish, butterfish, raw
Fish, mackerel, Atlantic, raw
Game meat, elk, tenderloin, separable lean only, cooked, broil
Granola bar, soft, milk chocolate coated, peanut butter
KENTUCKY FRIED CHICKEN, Popcorn Chicken
Lamb, domestic, leg, sirloin half, separable lean and fat, trimm
Lamb, variety meats and by-products, spleen, cooked, braised
Nuts, almonds, oil roasted, without salt added
Seeds, sesame seeds, whole, roasted and toasted
Soup, scotch broth, canned, prepared with equal volume wate
Soup, stock, beef, home-prepared

COORDENADAS ESTRELA



- Os eixos são distribuídos radialmente em um círculo separados por ângulos iguais inicialmente
- Pode-se aplicar deformações aos comprimentos dos eixos atribuindo maior ou menor ênfase a estes

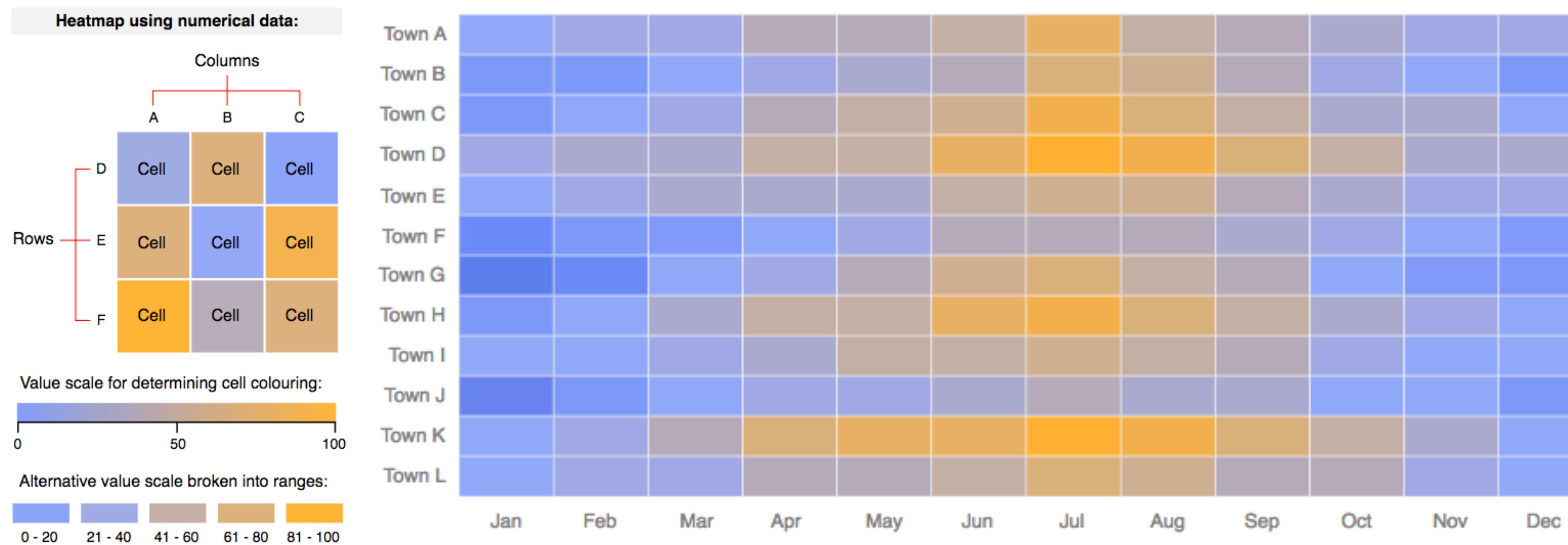
COORDENADAS ESTRELA



- Algoritmo para cálculo das posições dos pontos no novo sistema de coordenadas

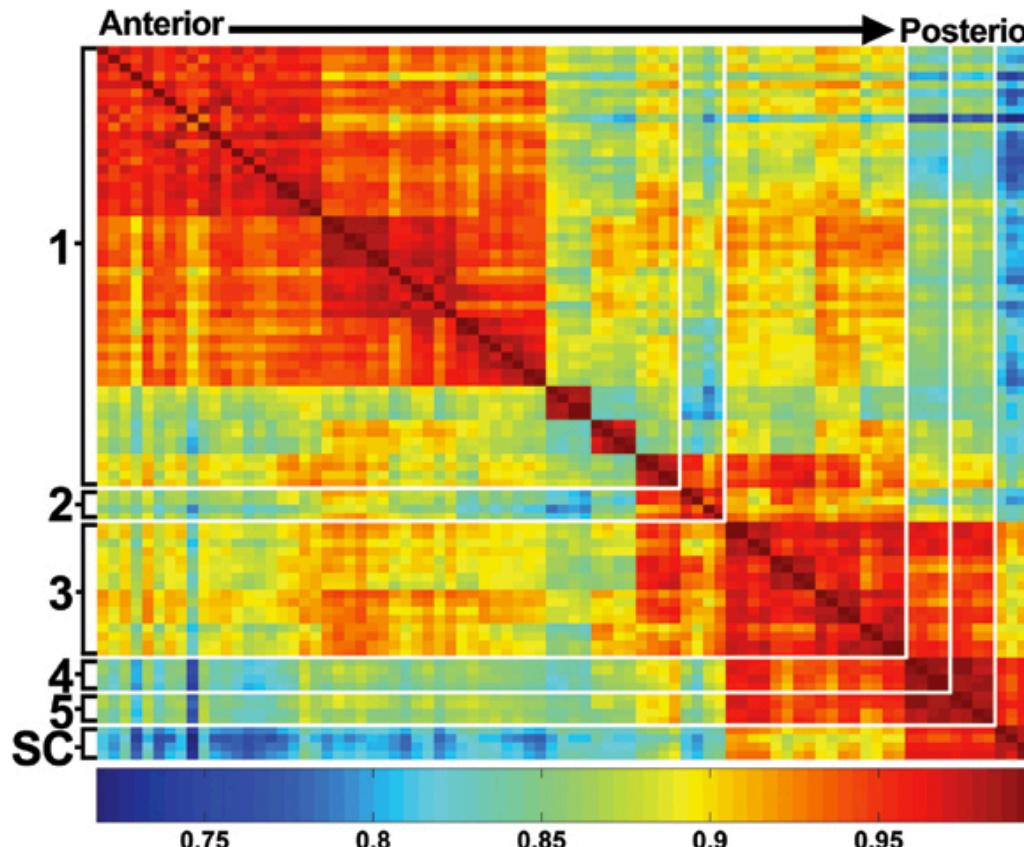
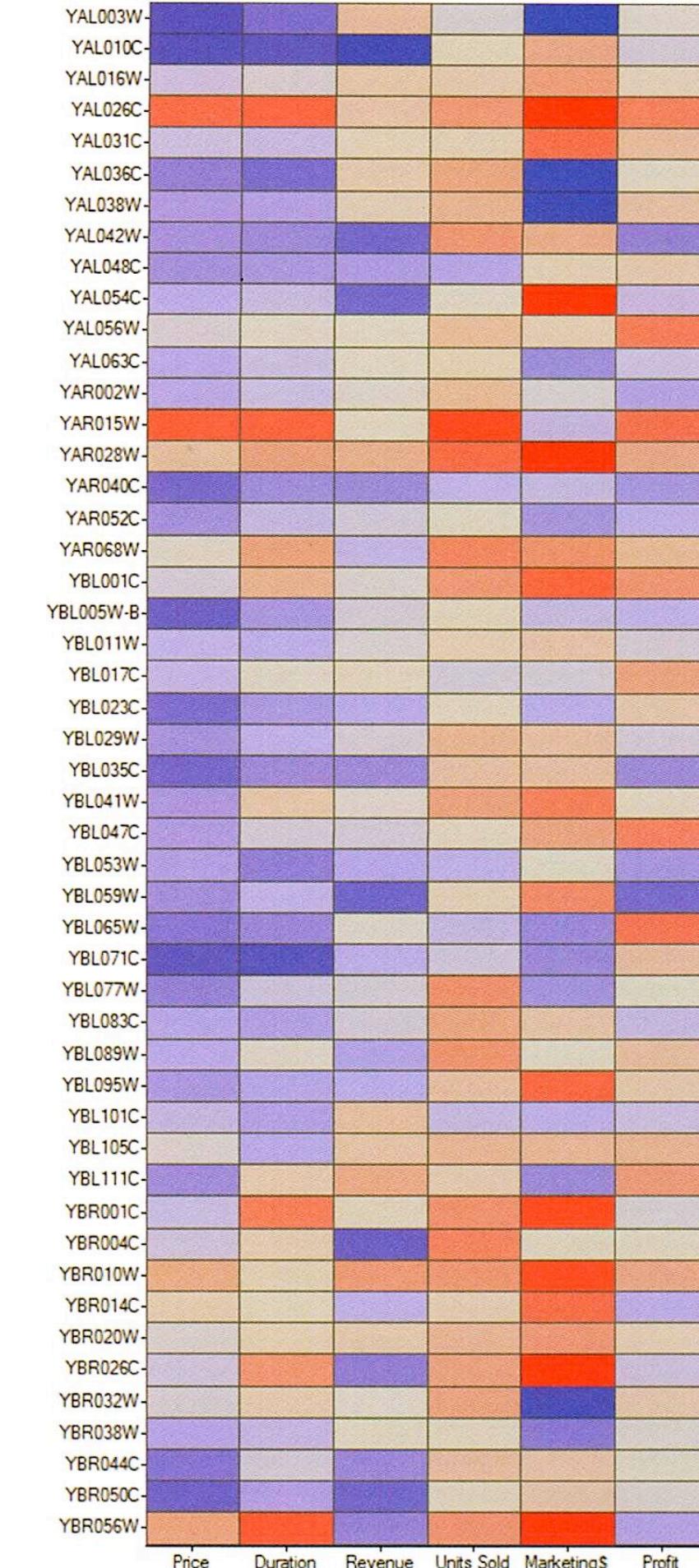
MAPAS DE CALOR

- Usam variações de cores para indicar quantidades
- Diferentes instâncias são representadas em linhas
- Cada coluna representa uma variável
- Células verdes indicam valores acima da média e vermelhos abaixo



PROBLEMA: CORES

- Cerca de 10% dos homens e 1% das mulheres não diferenciam bem entre verde e vermelho
- Preto não é uma boa cor para representar médias
- É difícil distinguir uma grande variedade de cores
- Nossa percepção de quantidades através de cores é imprecisa



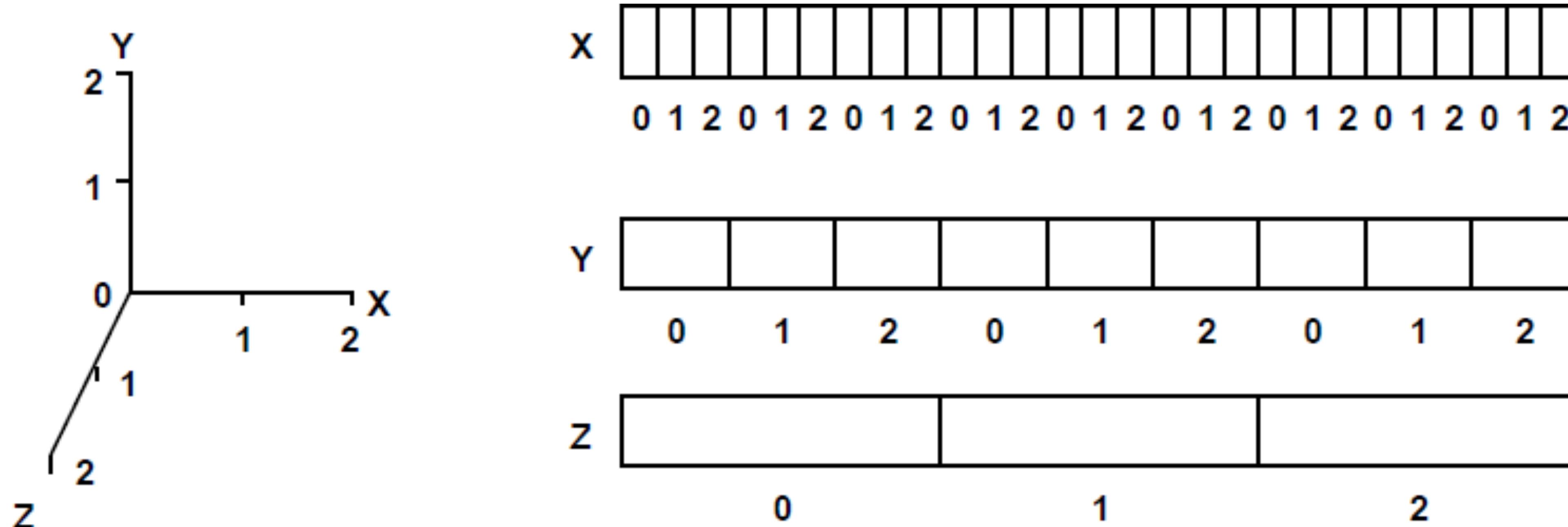
REPRESENTAÇÕES HIERÁRQUICAS

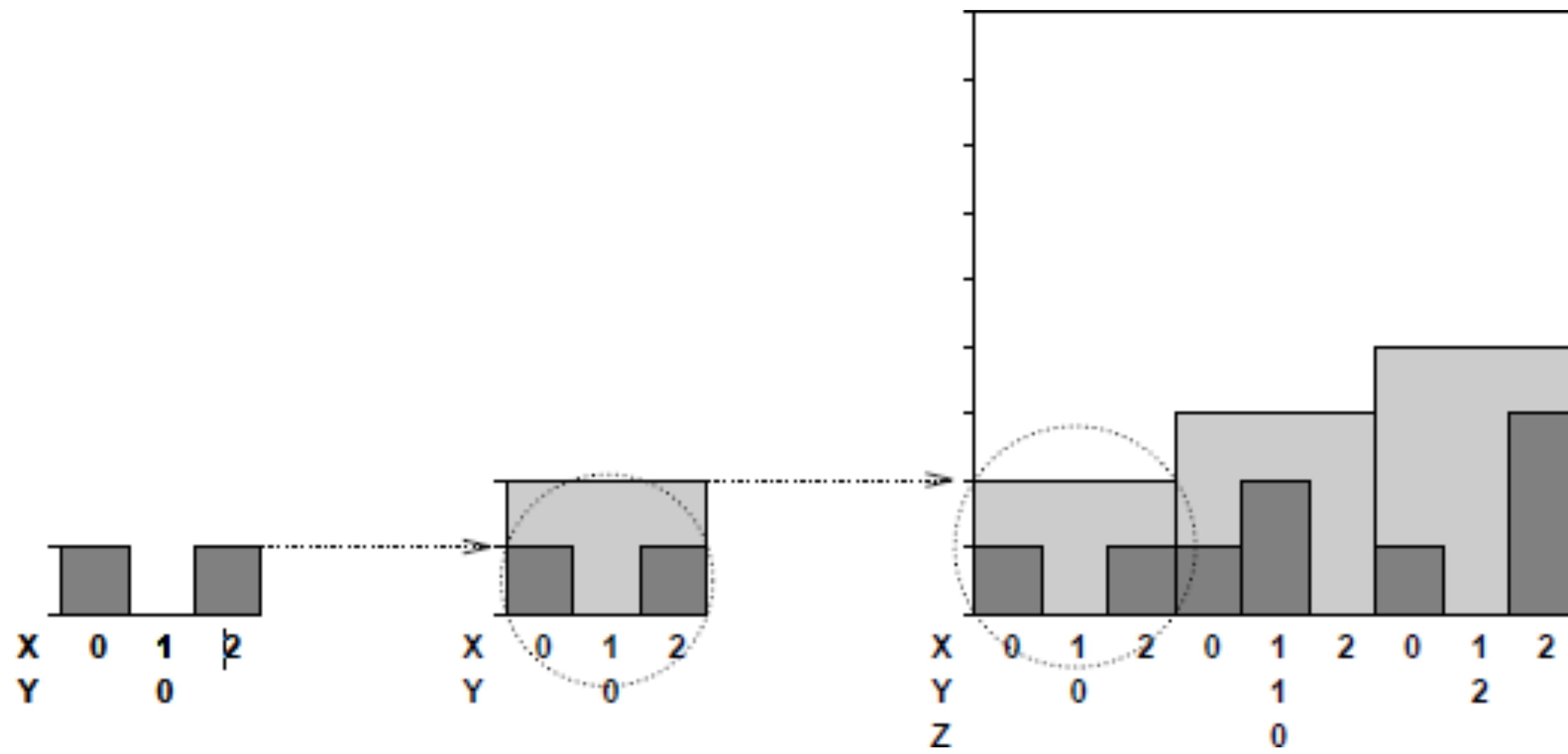
REPRESENTAÇÕES HIERÁRQUICAS

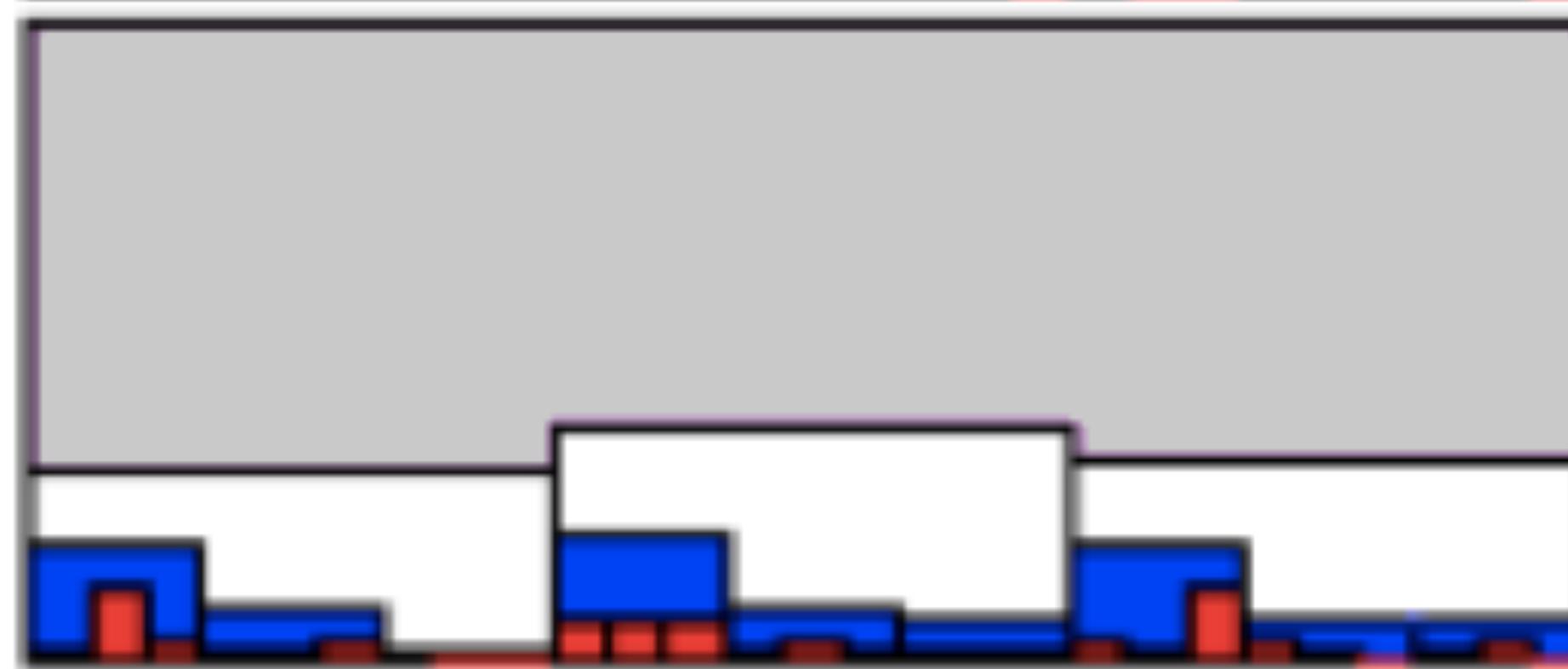
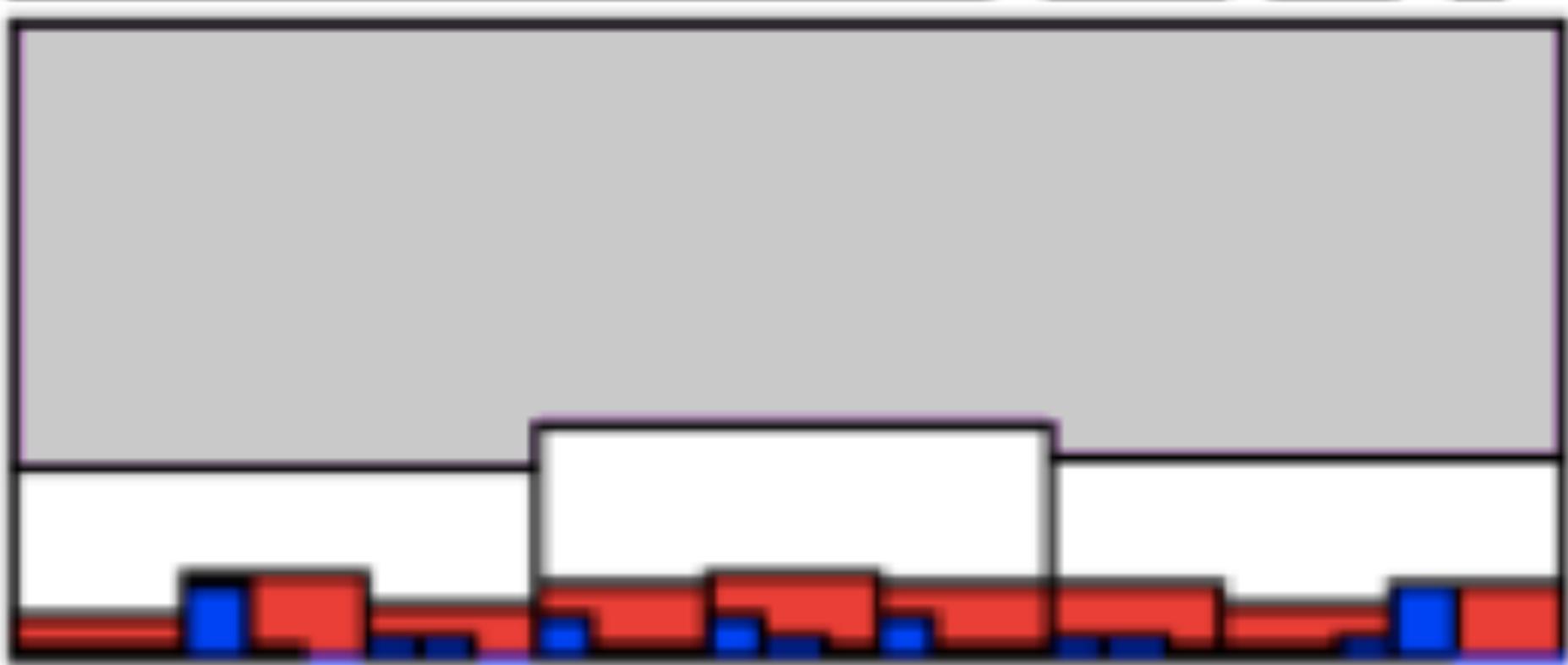
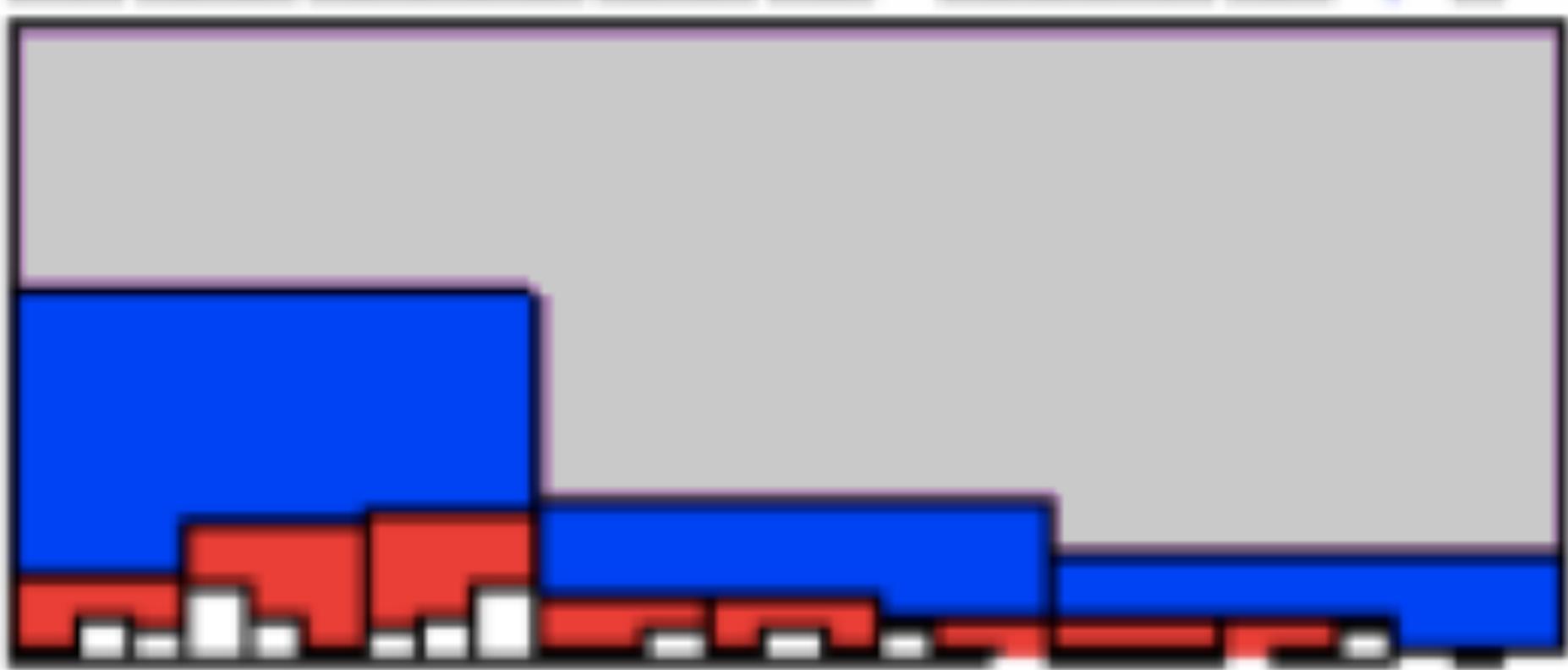
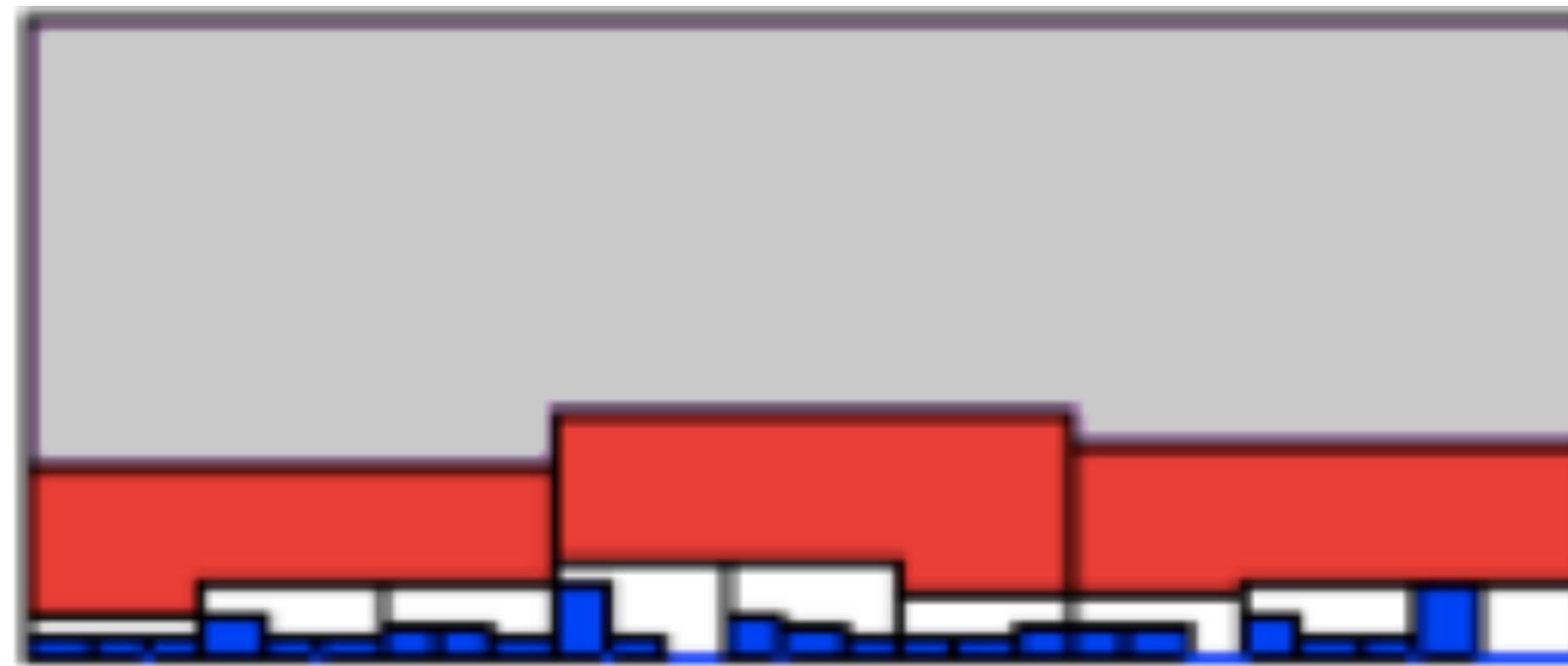
- Sub dividem o espaço e apresentam sub-espacos de modo hierárquico
- Atributos são tratados de forma diferente, portanto geralmente são úteis quando se trata de atributos hierárquicos
- Normalmente, treinamento é necessário para compreensão desse tipo de abordagem

EIXOS HIERÁRQUICOS

- Nos eixos hierárquicos, eixos ortogonais são representados horizontalmente de maneira hierárquica







.....

REPRESENTAÇÕES ICONOGRÁFICAS

REPRESENTAÇÕES ICONOGRÁFICAS

- Consistem no mapeamento de objetos multidimensionais em ícones ou glifos
- As características visuais destes ícones dependem dos atributos

GLIFOS

- São objetos gráficos compostos por um conjunto de atributos visuais

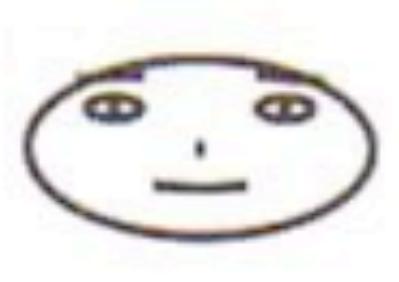
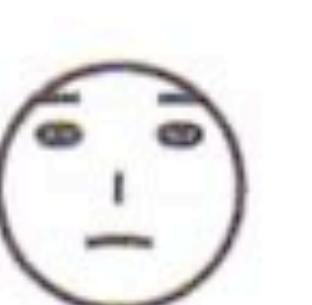
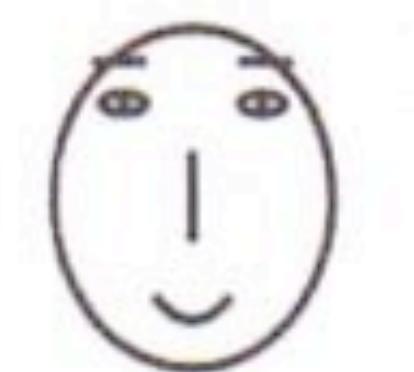


Atributo visual

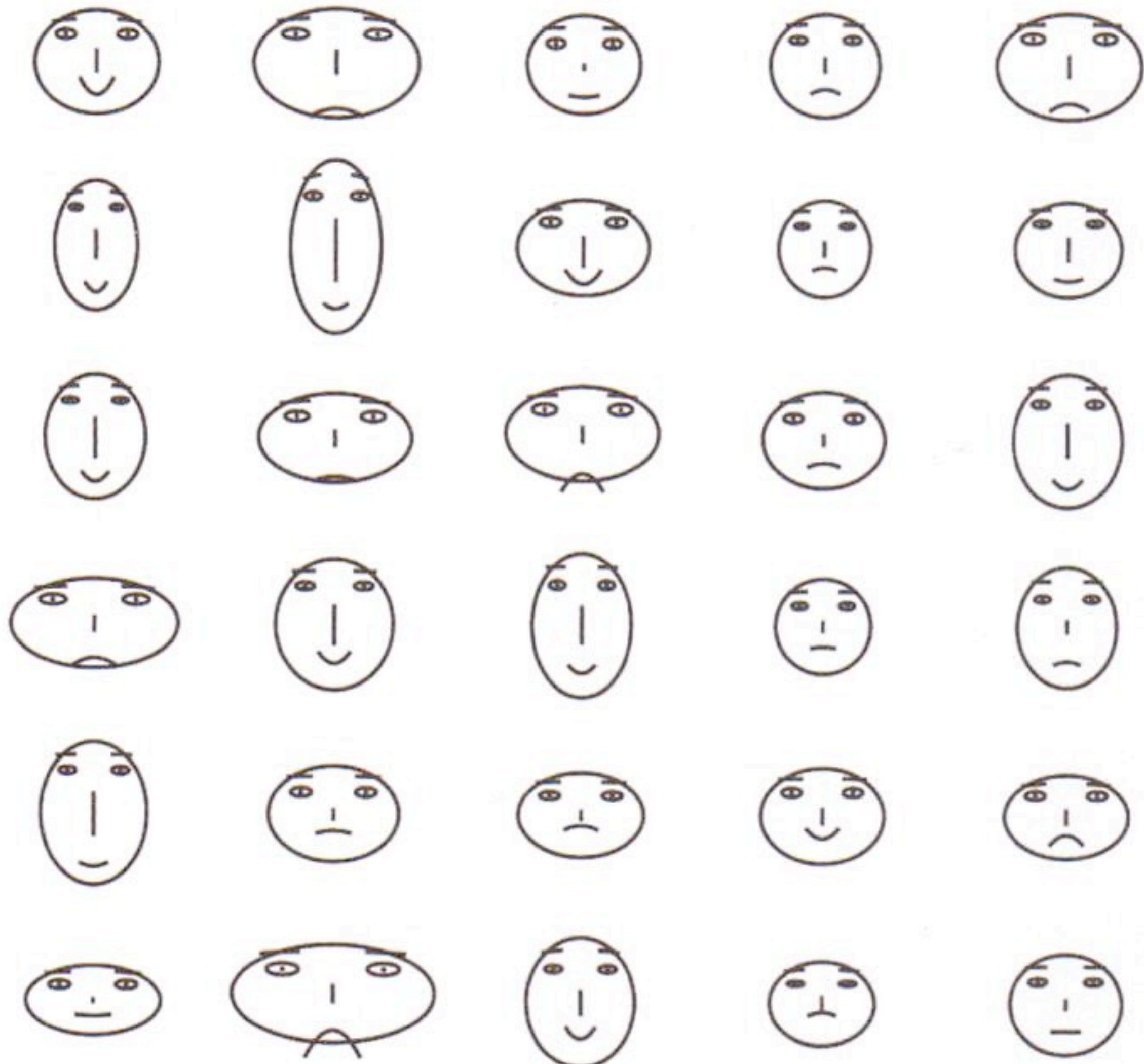
Variável

Cor	Temperatura corporal
Formato da cabeça	Tipo sanguíneo
Largura do tronco	IMC
Posição dos braços	Batimentos cardíacos
Posição das pernas	Taxa de açúcar





ACES DE CHERNOFF



- Faces de Chernoff foram propostas na década de 70 para representação de dados multivariados
- Baseada no fato de que somos capazes de reconhecer similaridades entre faces



Transgender
(general)



Female



Hermaphrodite
(both)



Hermaphrodite
(Androgynie)



Hermaphrodite
(Mercury)



Male

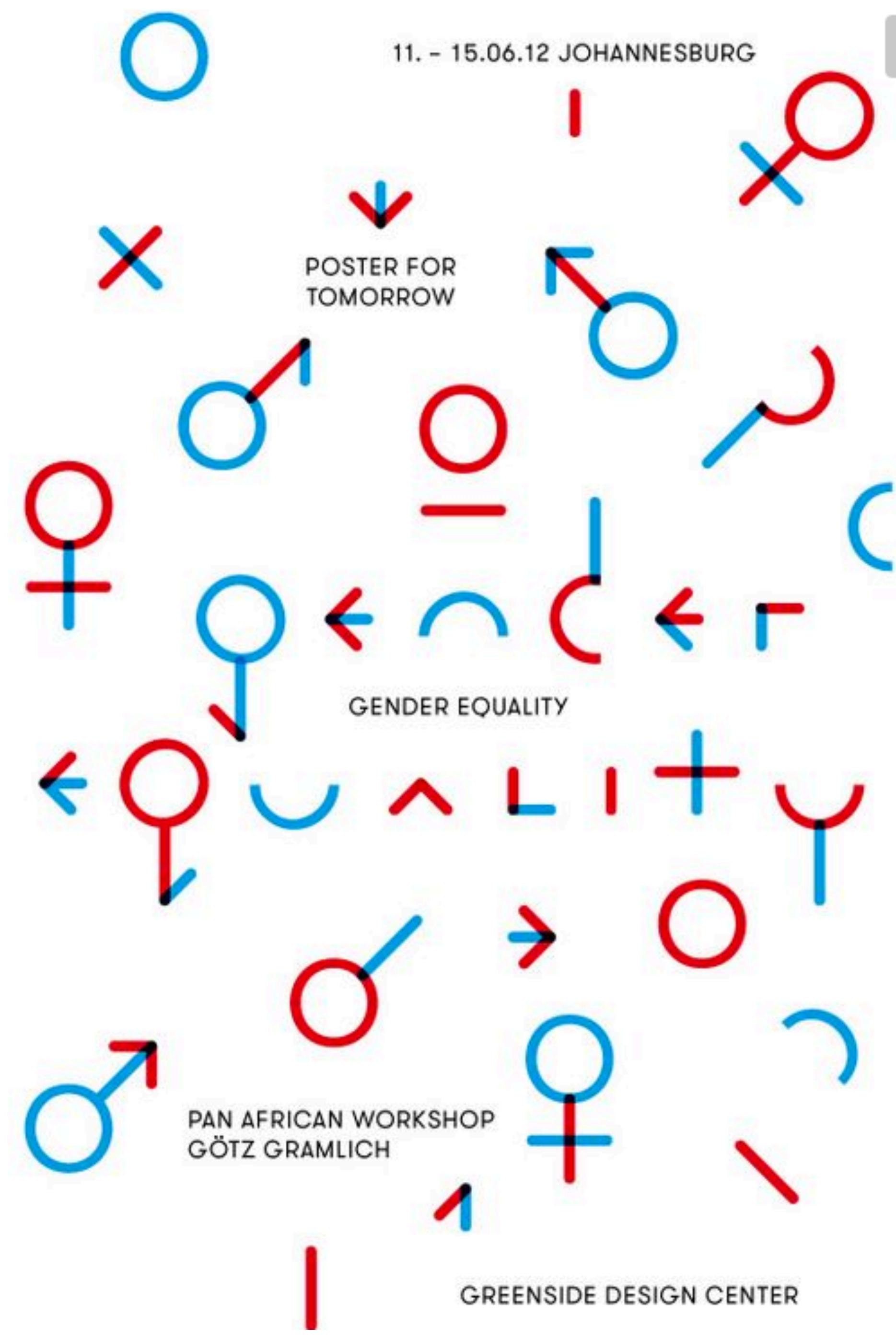


Neutrois
(neutral)



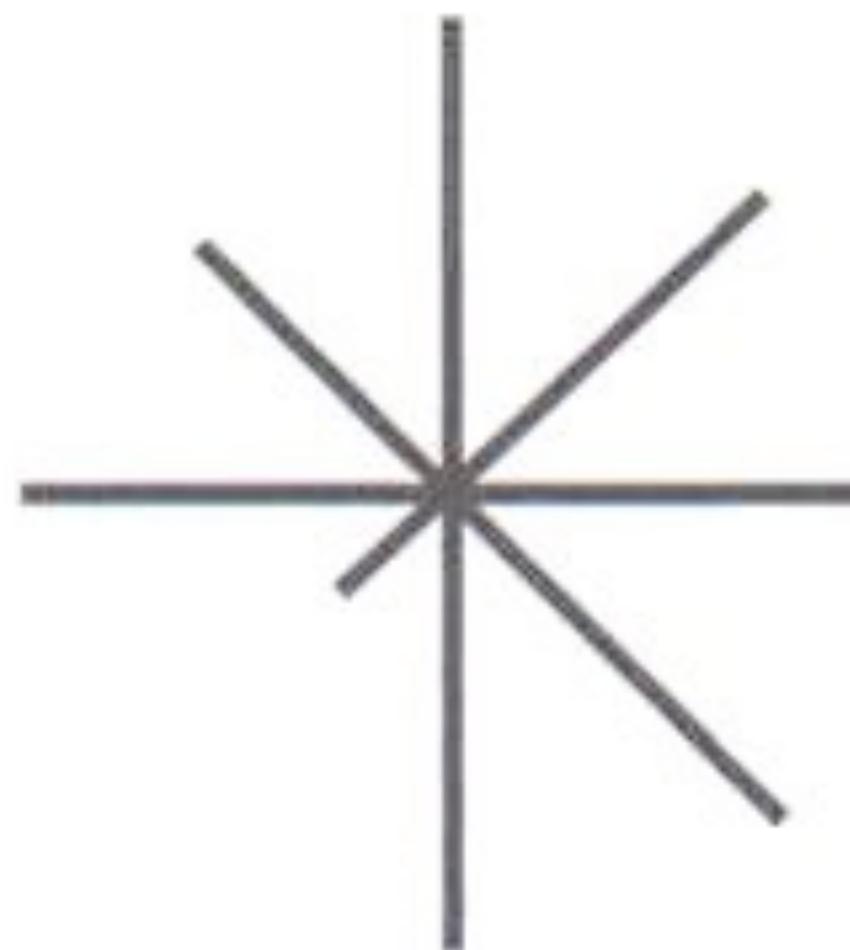
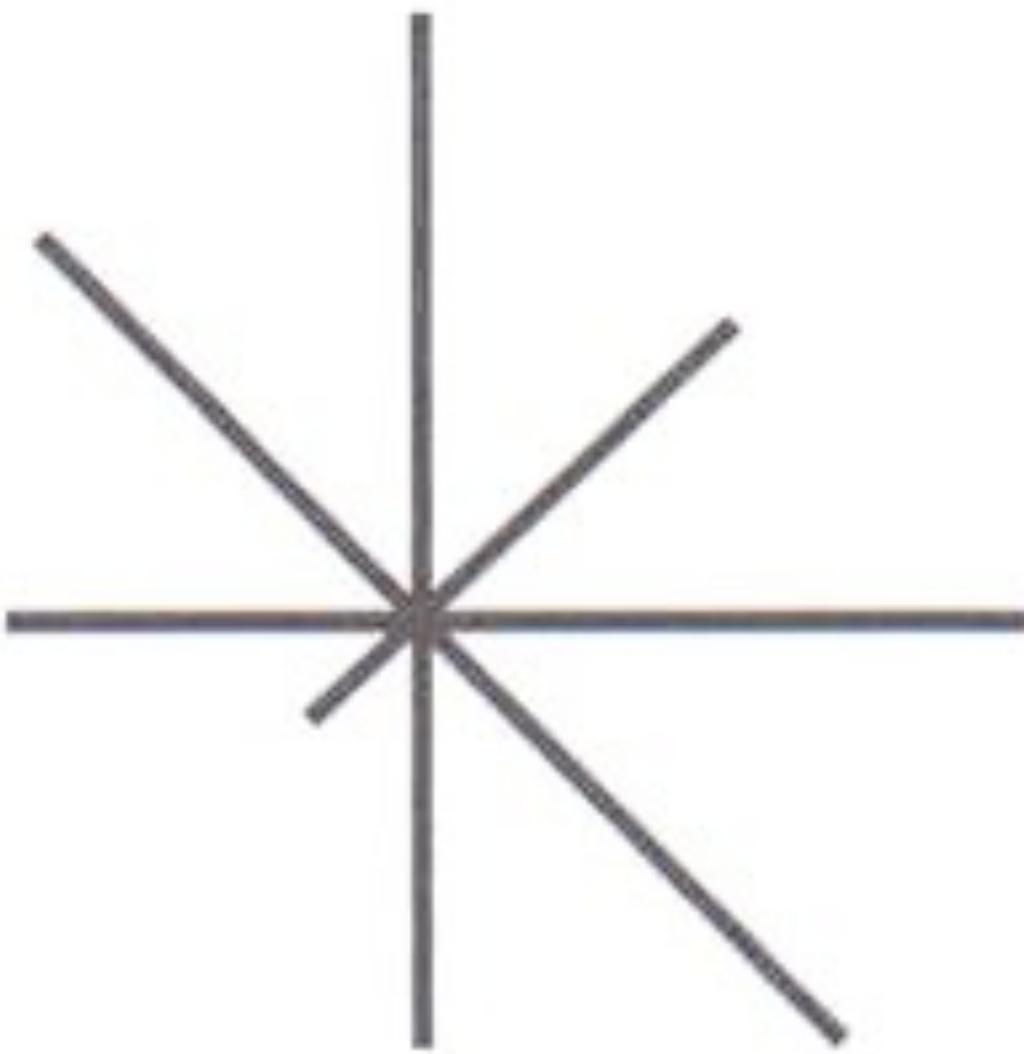
Other

11. - 15.06.12 JOHANNESBURG



WHISKERS

- Whiskers consistem em múltiplas linhas irradiando de um ponto central cada qual representando o valor de uma variável através do comprimento das linhas



Create Your Better Life Index

How do you define a better life? Use our interactive tool to see how your country performs on the topics you feel make for a better life.

↗ Start with all topics rated equally

or set your own preferences [here](#).



Create Your Better Life Index

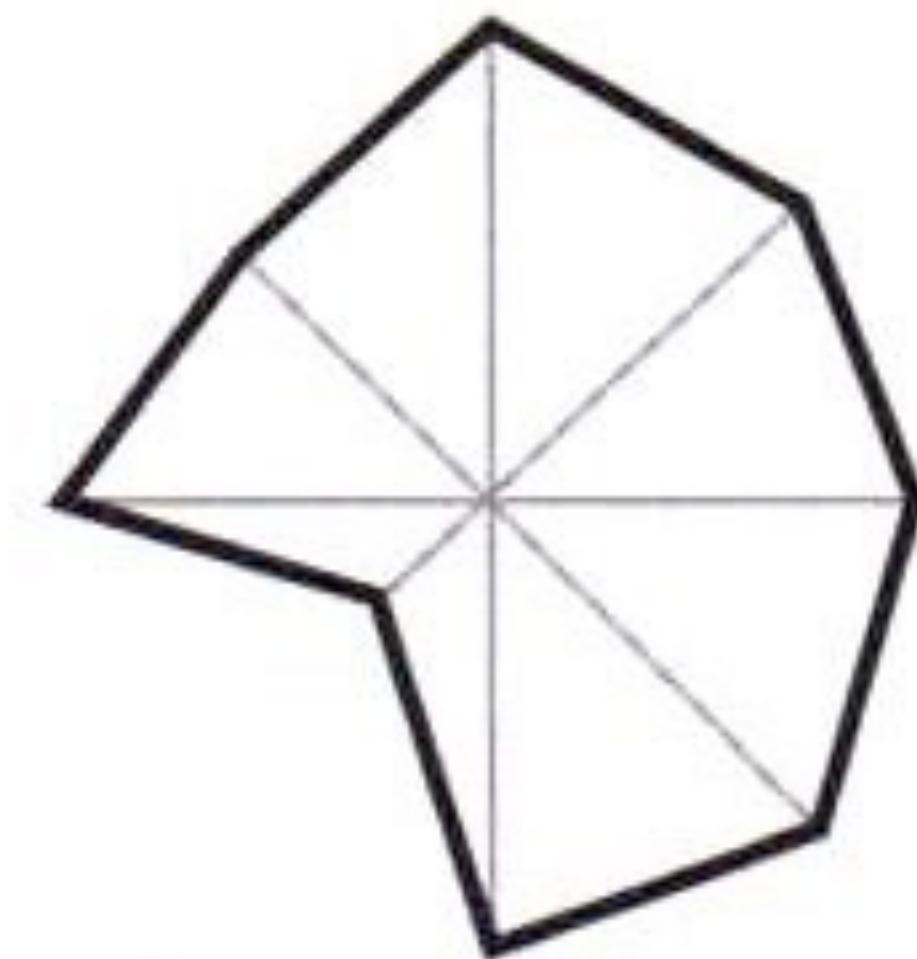
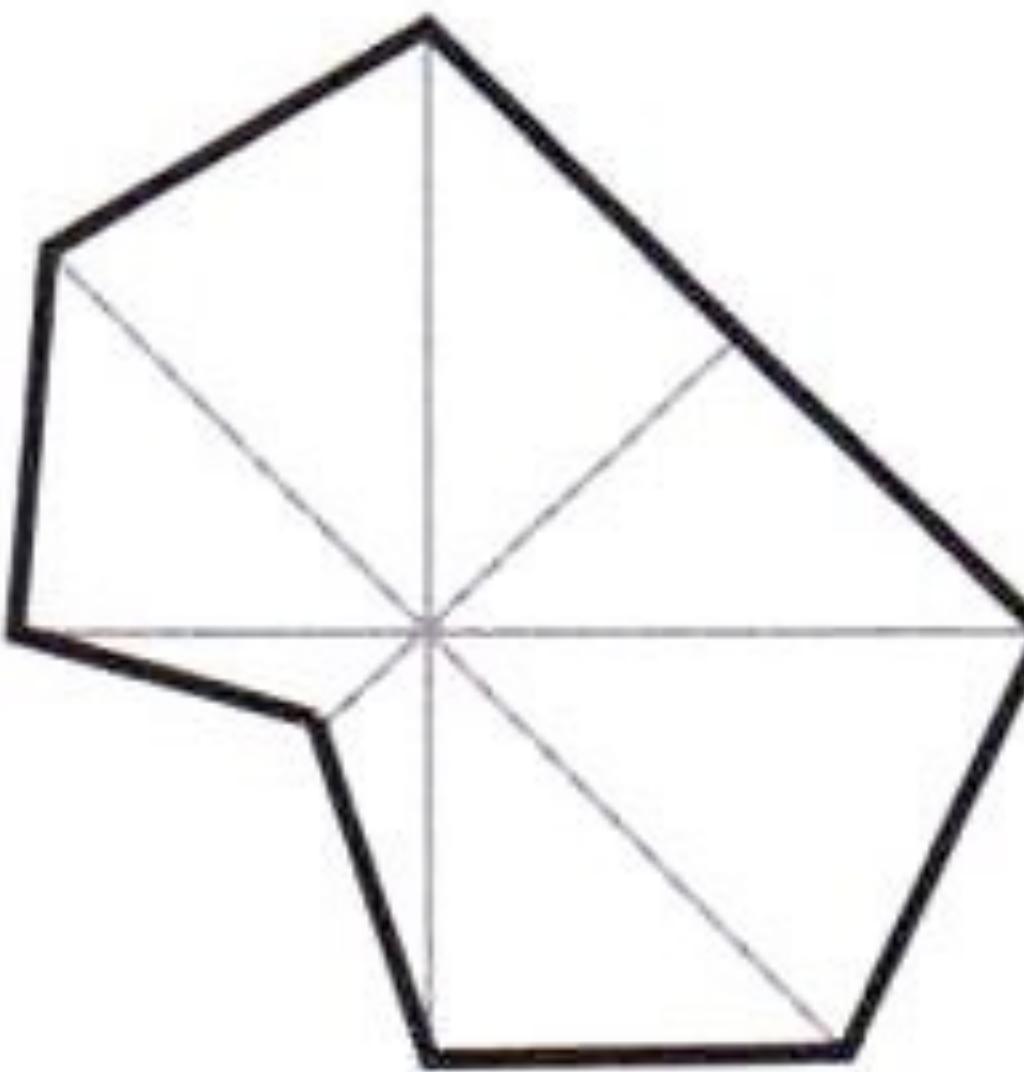
Rate the topics according to their importance to you:

	-	+
Housing	● ● ● ●	● ● ● ● ●
Income	● ● ● ●	● ● ● ● ●
Jobs	● ● ● ●	● ● ● ● ●
Community	● ● ● ●	● ● ● ● ●
Education	● ● ● ●	● ● ● ● ●
Environment	● ● ● ●	● ● ● ● ●
Governance	● ● ● ●	● ● ● ● ●
Health	● ● ● ●	● ● ● ● ●
Life Satisfaction	● ● ● ●	● ● ● ● ●
Safety	● ● ● ●	● ● ● ● ●
Work-Life Balance	● ● ● ●	● ● ● ● ●

Share this index

GLIFOS ESTRELA

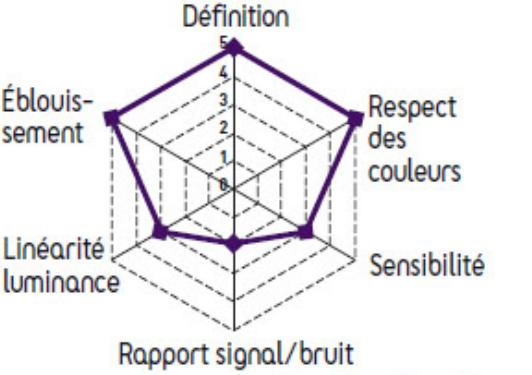
- Glifos estrela ou gráficos de radar são semelhantes aos whiskers



PANASONIC TM-300

Mémoire flash

Tests laboratoire



32 Go



L



optique



STABILISÉ

Zoom x12



10,6 Mpix



3MOS



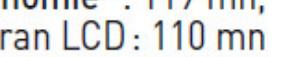
HDMI



USB

Autonomie*: 119 mn,
avec écran LCD: 110 mn

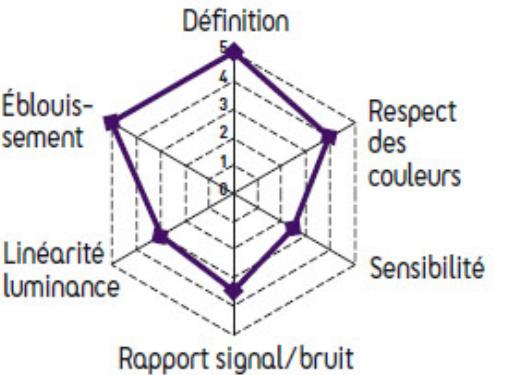
Lisibilité écran



CANON HF S10

Mémoire flash

Tests laboratoire



32 Go



L



optique



STABILISÉ

Zoom x10



8 Mpix



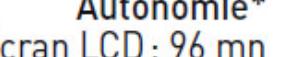
HDMI



USB

Autonomie*: 119 mn,
avec écran LCD: 96 mn

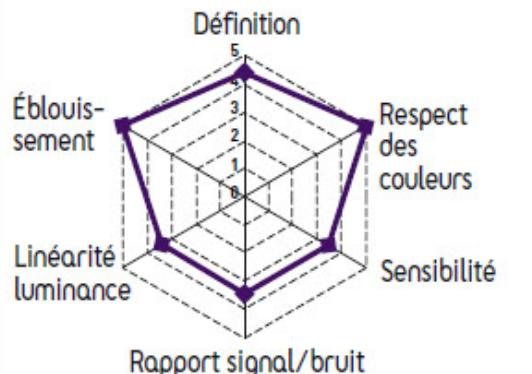
Lisibilité écran



PANASONIC HDC-HS300

Disque Dur

Tests laboratoire



120 Go



L



optique



STABILISÉ

Zoom x12



10,6 Mpix



3MOS



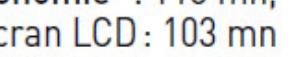
HDMI



USB

Autonomie*: 118 mn,
avec écran LCD: 103 mn

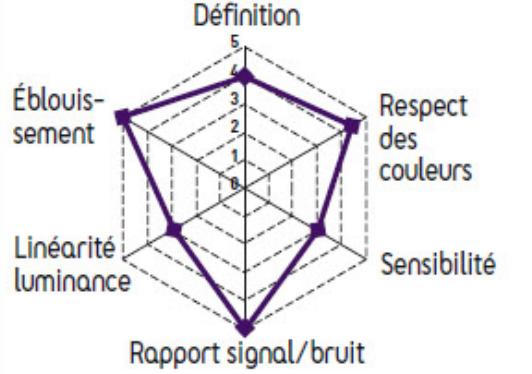
Lisibilité écran



SONY HDR-XR 520

Disque dur

Tests laboratoire



240 Go



L



optique



STABILISÉ

Zoom x12



12 Mpix



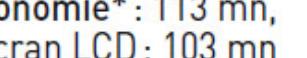
HDMI



USB

Autonomie*: 113 mn,
avec écran LCD: 103 mn

Lisibilité écran



Le jugement de la Fnac: autonomie et mémoire d'éléphant pour le TM300 et ses 32 Go de mémoire flash intégrée. Avec son viseur électronique et sa bague multifonction autour de l'objectif, ce caméscope joue la carte de l'ergonomie. Son originalité se cache du côté des capteurs, puisqu'il est doté d'un capteur triMOS, permettant un meilleur rendu des couleurs. Côté optique, le zoom 12x stabilisé est signé Leica Dicomar. Côté son, on appréciera l'enregistrement en 5.1, avec deux micros, l'un suivant les mouvements de l'optique tandis que l'autre capte les sons d'ambiance. Côté labo : effet de halo parfait.

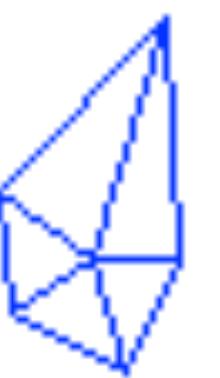
Le jugement de la Fnac: support avec 12 heures d'enregistrement en Full HD sur mémoire interne de 32 Go sont au programme de la HFS 10. Reprenant la même plateforme technique que le HFS 100 dont l'enregistrement en 24 Mbps, le rythme le plus élevé du marché, pour une captation des mouvements plus fluide et une reproduction des détails et des couleurs précise. Le HFS 10 est doté d'un autofocus rapide et d'une optique stabilisée (zoom 10x). Côté connectiques, mini griffe porte-accessoires, entrée micro et prises casque sont au rendez-vous ! Un vrai concentré de technologie.

Le jugement de la Fnac: frère jumeau du Panasonic HDC-TM 300 mais cette fois-ci sur disque dur de 120 Go. Il met l'accent sur une nouvelle gestion des couleurs grâce à trois capteurs triMOS Full HD afin de doubler la quantité de lumière reçue, pour une meilleure qualité d'image. Par ailleurs, afin d'assurer un meilleur rendu des films en Full HD (1080i), le caméscope est doté de la stabilisation optique OIS avancée (4000x/s). Un programme que les amateurs de Panasonic connaissent bien : la fonction IA détermine elle-même le mode scène approprié.

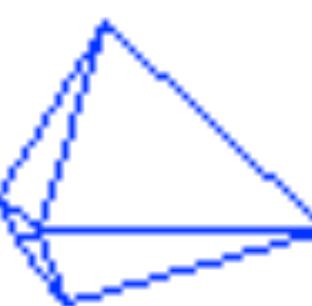
Le jugement de la Fnac: ce modèle HD équipé d'un disque dur de 240 Go est le produit « Star » de la gamme HD Sony. Doté de la fonction hybride pour filmer au format AVCHD sur disque dur ou sur carte, il enregistre jusqu'à 101 heures de film en 1920 x 1080 (full HD), mais aussi en 1440 x 1080 au choix. Reprenant les caractéristiques de la XR 200, il dispose en plus d'une molette de réglages manuels (Mise au point, exposition...). Ses plus : Capteur Exmor R, optique Sony G, un écran LCD de 8,1 cm, viseur en prime ! Existe en version 120 Go sous la référence XR 500.



chevrolet chevelle malibu



buick skylark 320



plymouth satellite



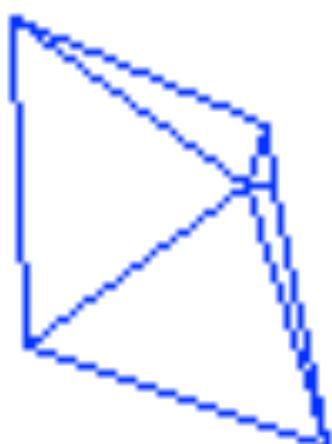
amc rebel sst



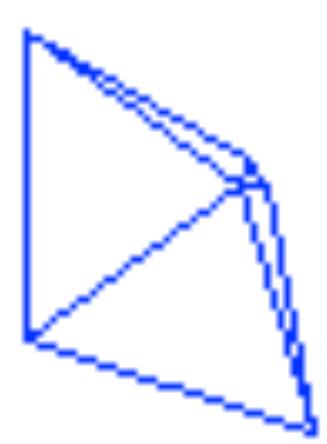
ford torino



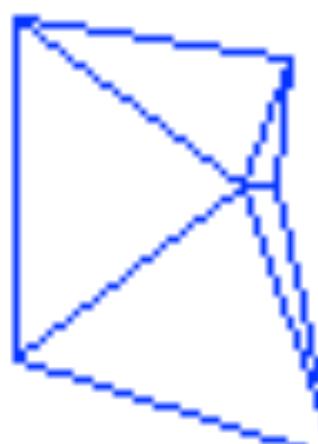
ford galaxie 500



chevrolet impala



plymouth fury iii



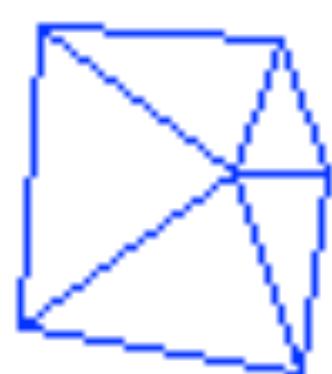
pontiac catalina



chevrolet chevelle malibu buick skylark 320



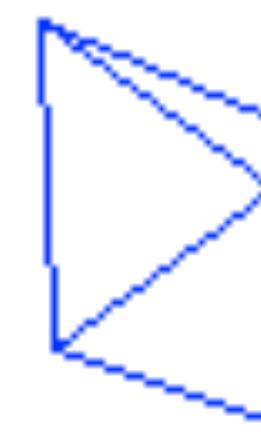
plymouth satellite



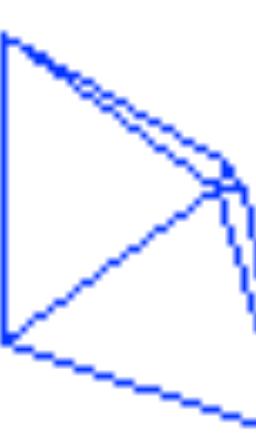
amc rebel sst

ford torino

ford galaxie 500



chevrolet impala



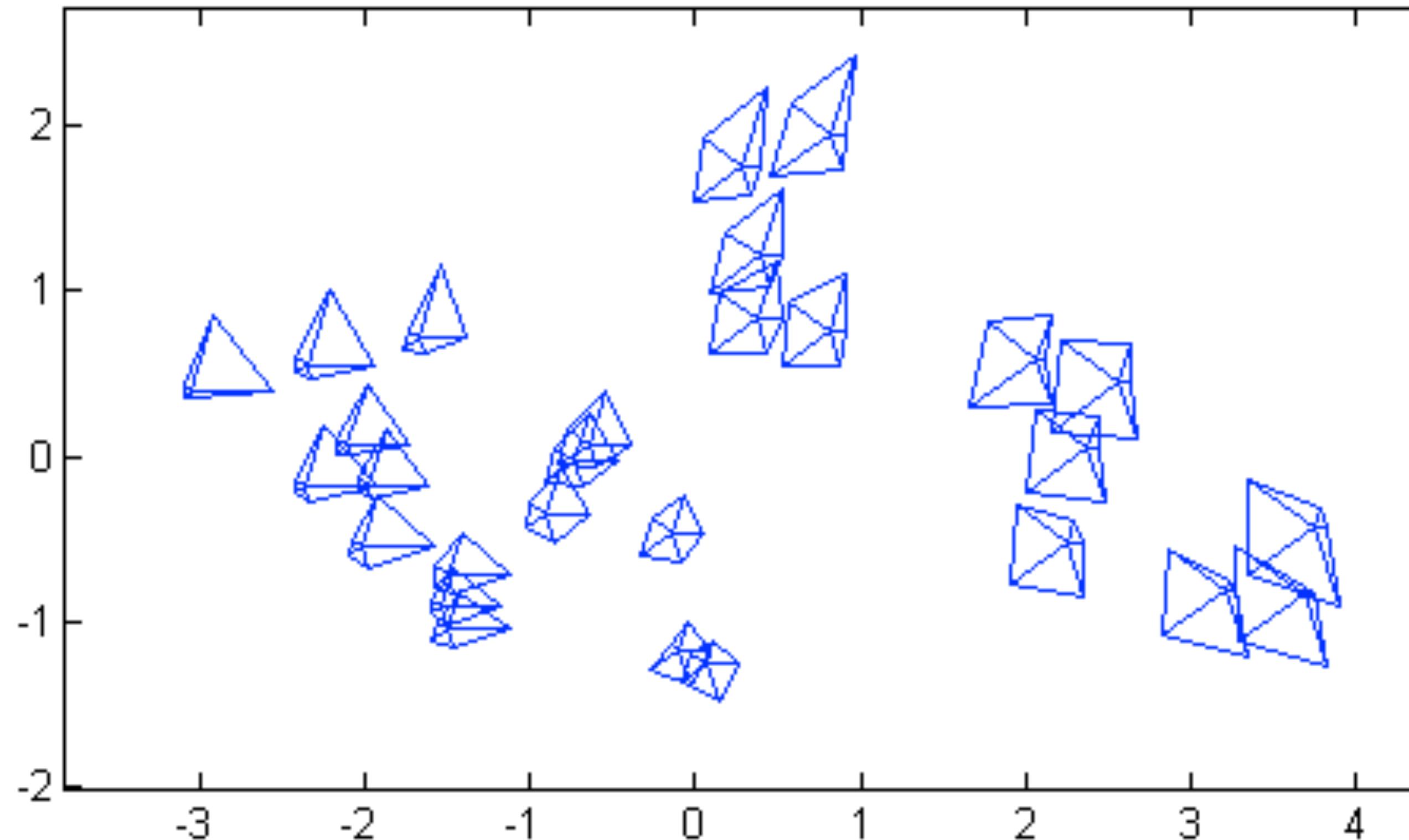
plymouth fury iii



pontiac catalina



1977 Model Year



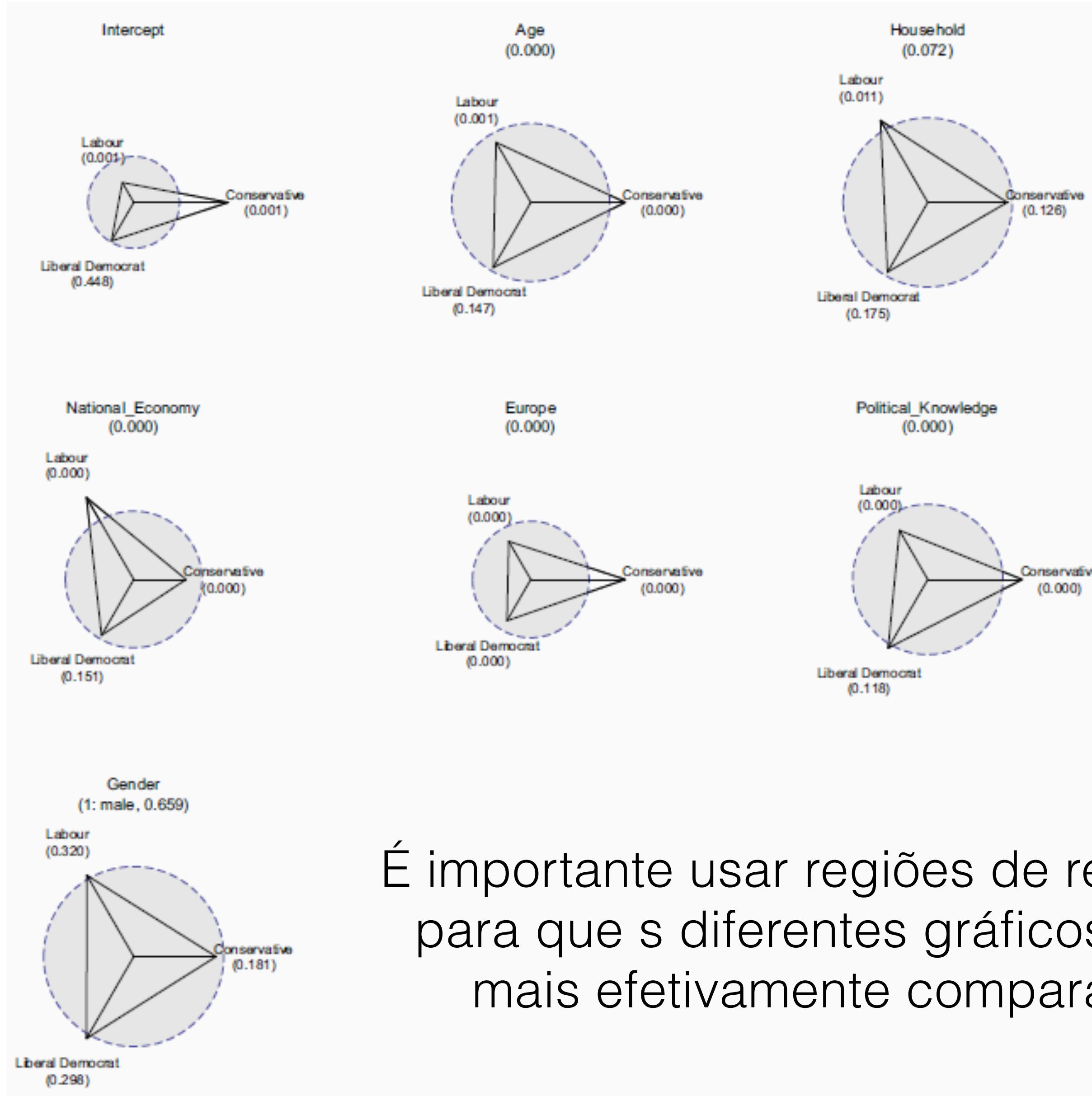
Técnicas de redução de dimensionalidade como:

MDS (Dimensionamento multidimensional)

SVD (Decomposição por valor singular) ou

PCA (Análise de componentes principais)

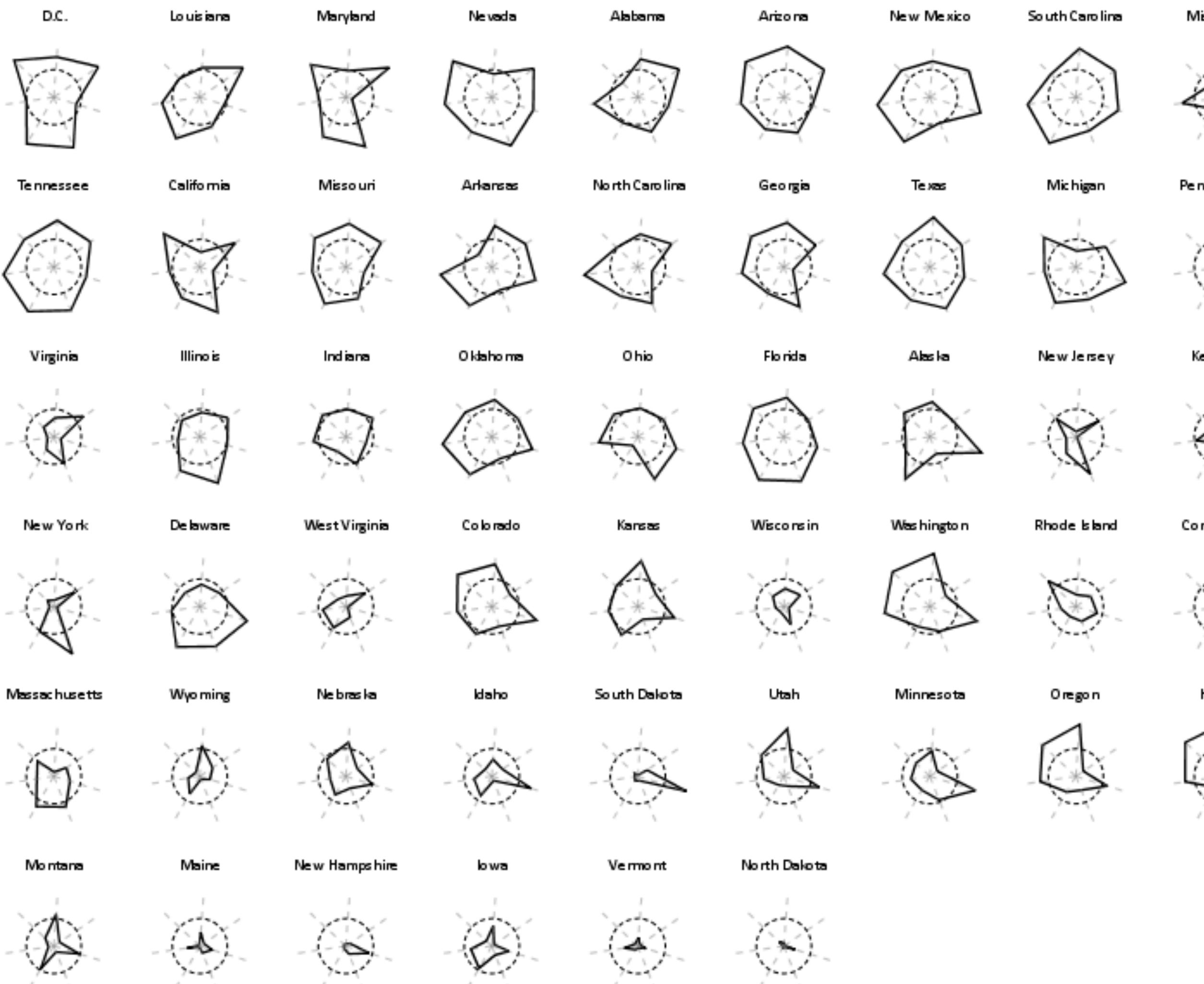
são úteis para desenhar os objetos multivariados em um espaço 2D



É importante usar regiões de referência
para que os diferentes gráficos sejam
mais efetivamente comparáveis

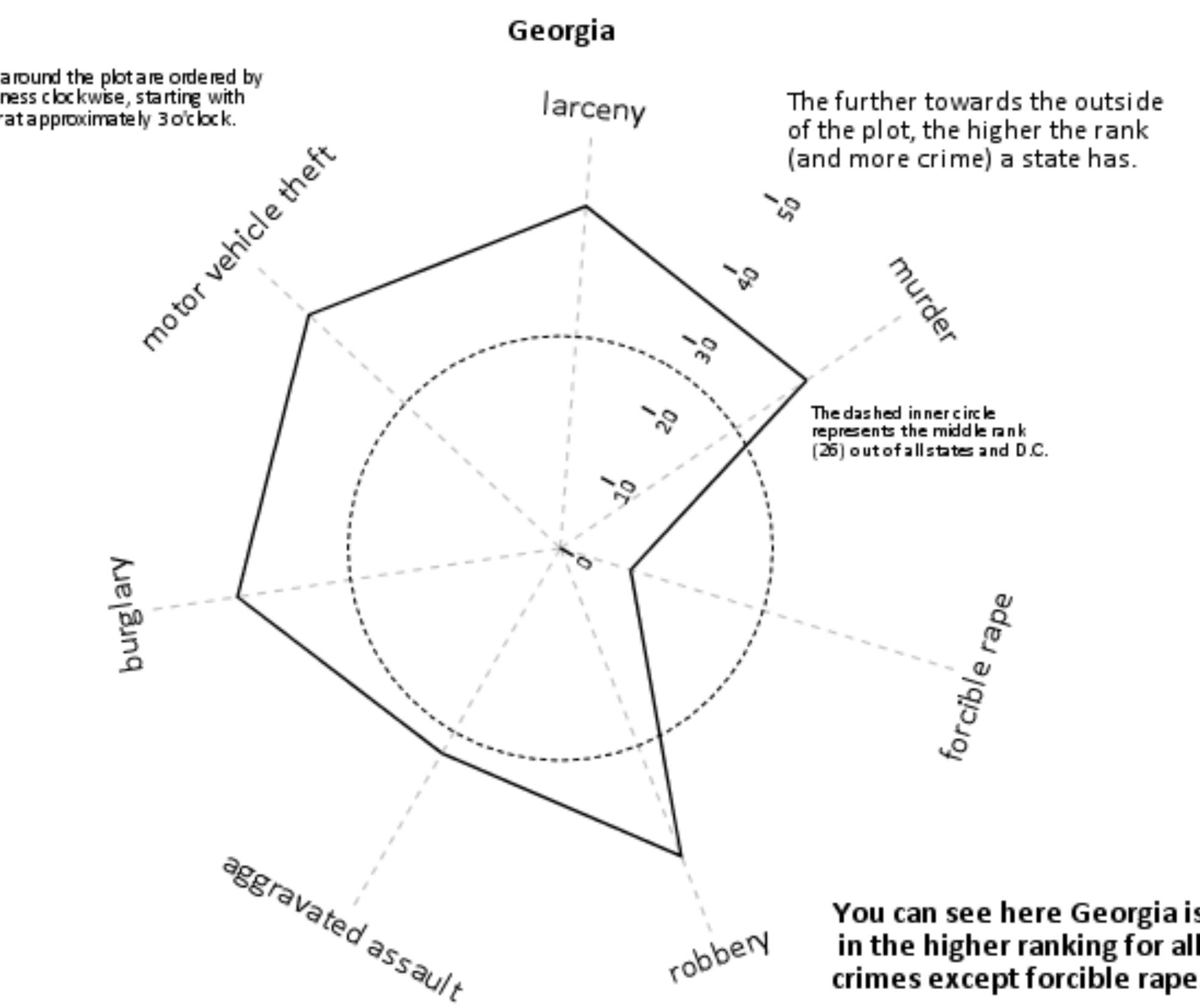
Rankings of Crime Rates for 50 States (and D.C.) Star Plot

States ordered by homicide ranking in plot (left to right)



Annotated Legend for Star Plots

Crimes around the plot are ordered by seriousness clockwise, starting with murder at approximately 3 o'clock.



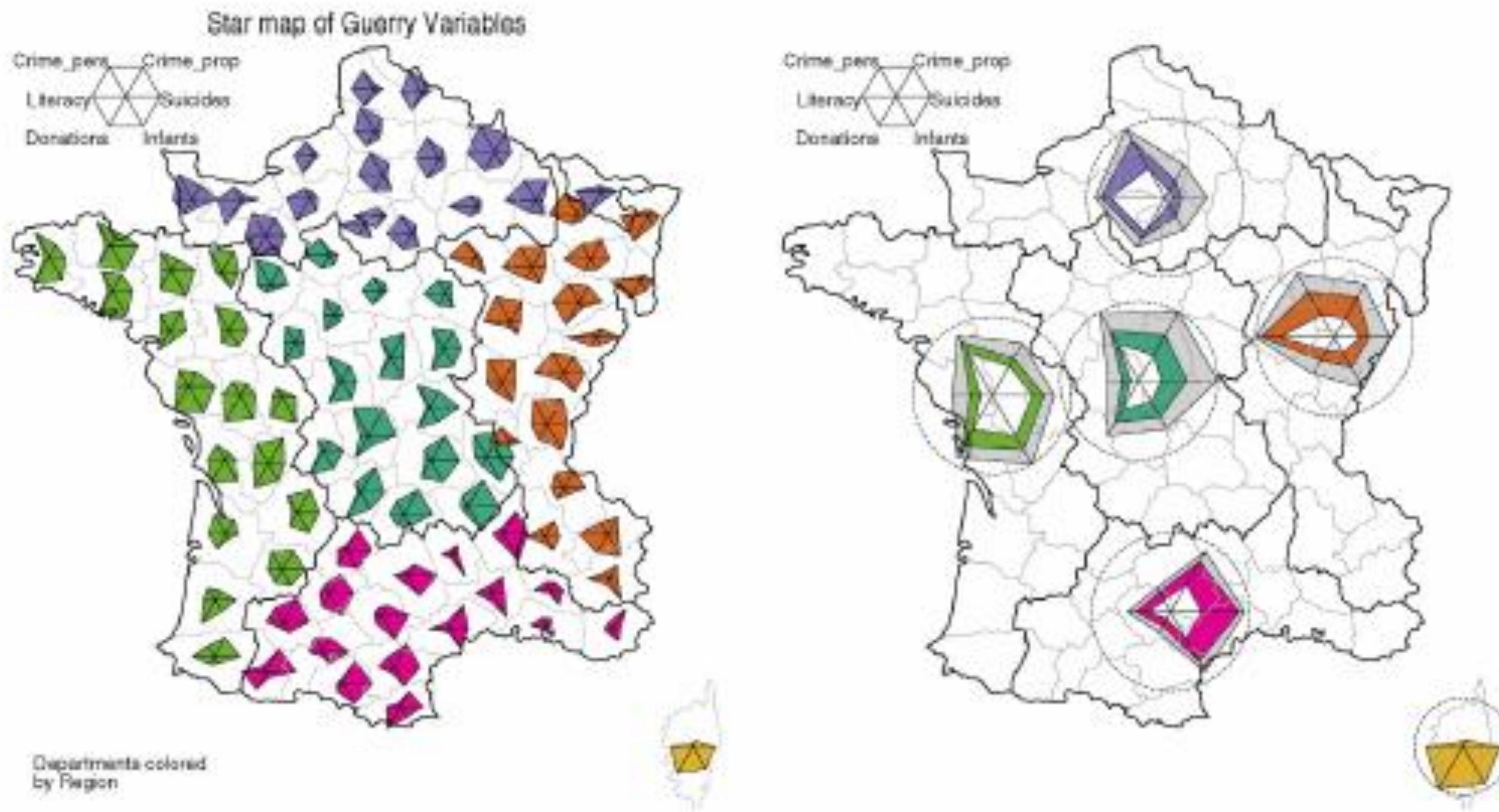
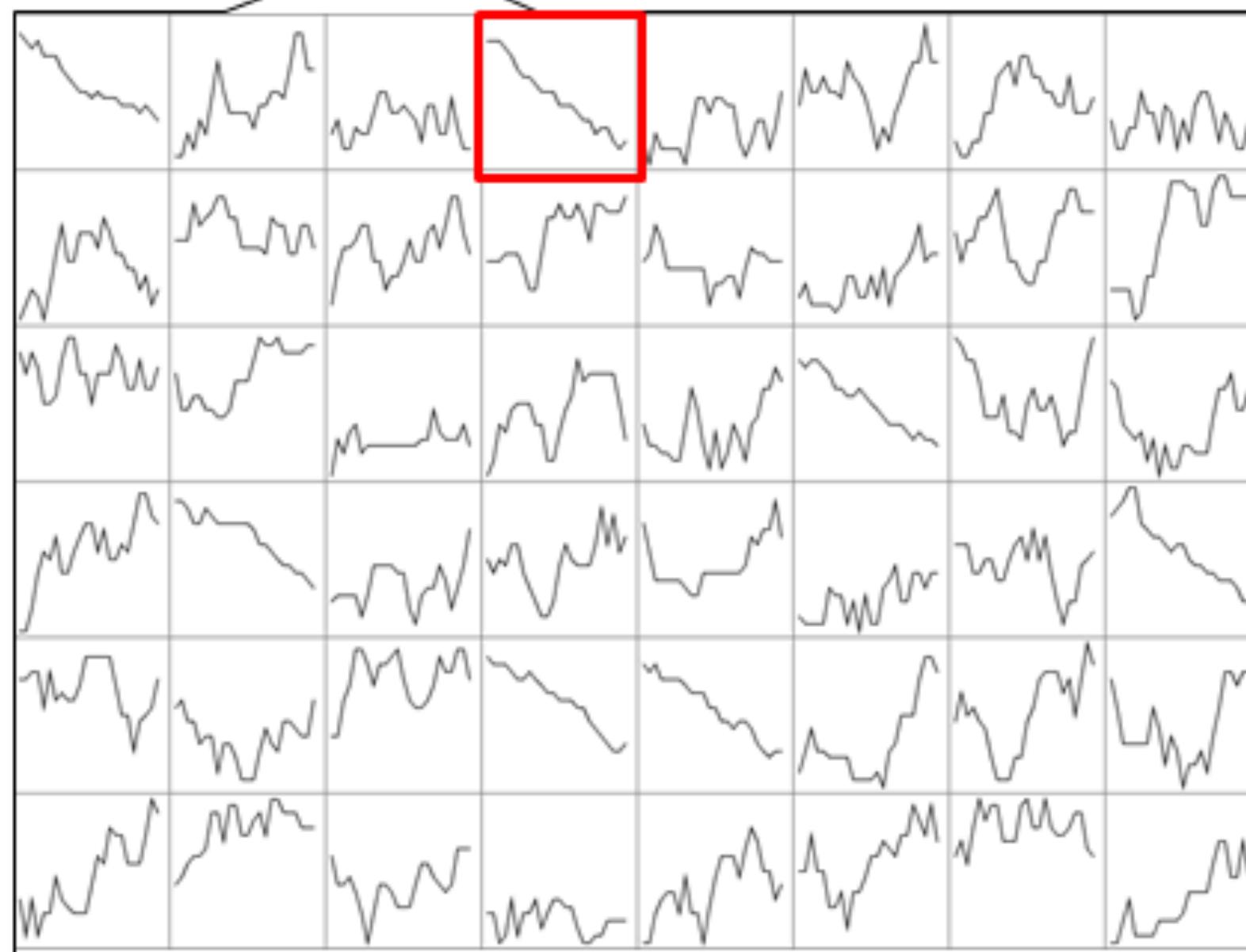
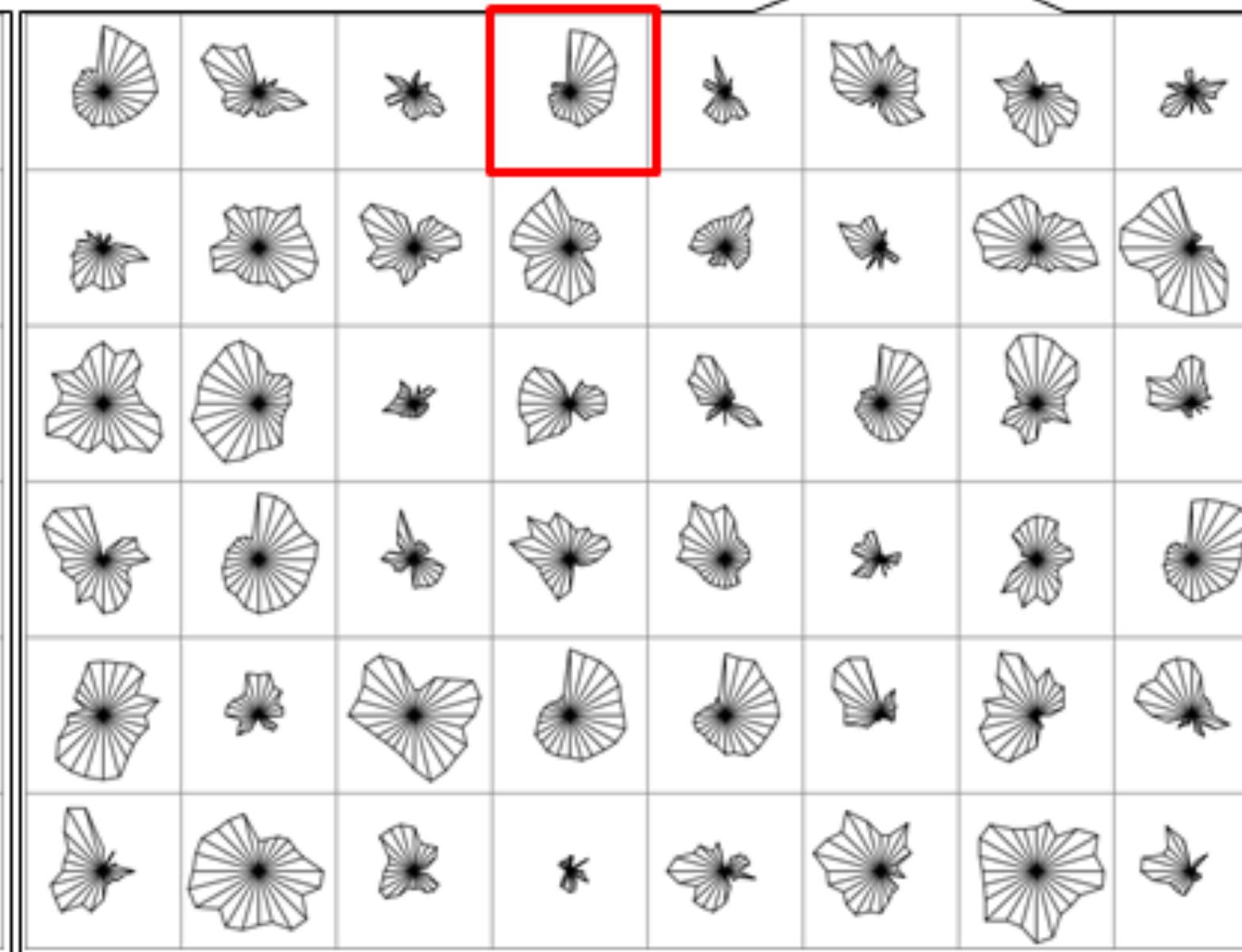


FIG. 17. Star maps of Guerry's data, using rays proportional to the rank of each variable (longer = better). Variables have been ordered according to their angular positions in the biplot (Figure 14). Left: Glyphs for individual départements. Right: Multivariate boxplot glyphs for the medians and quartiles across départements in each region of France.

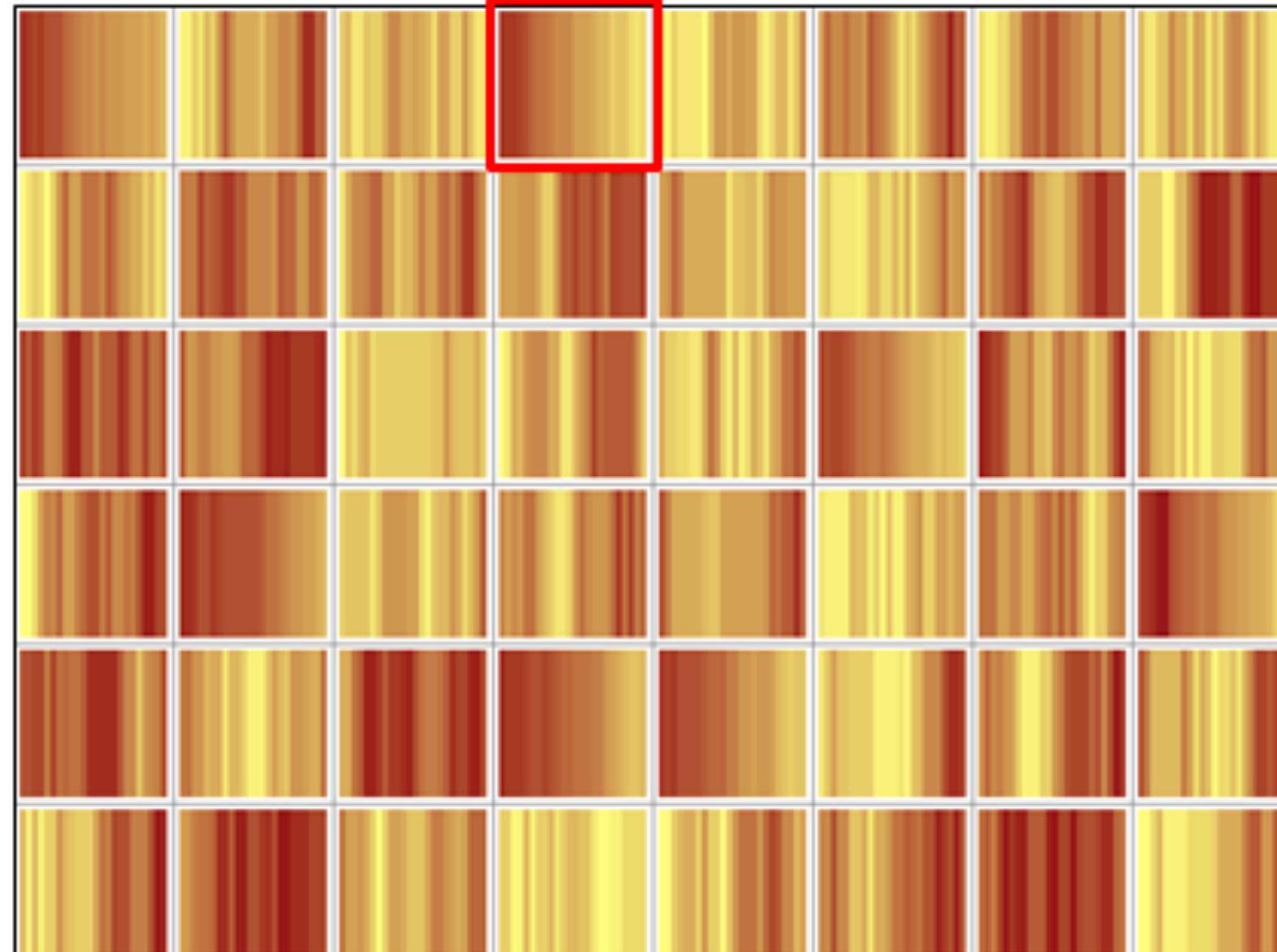
Line



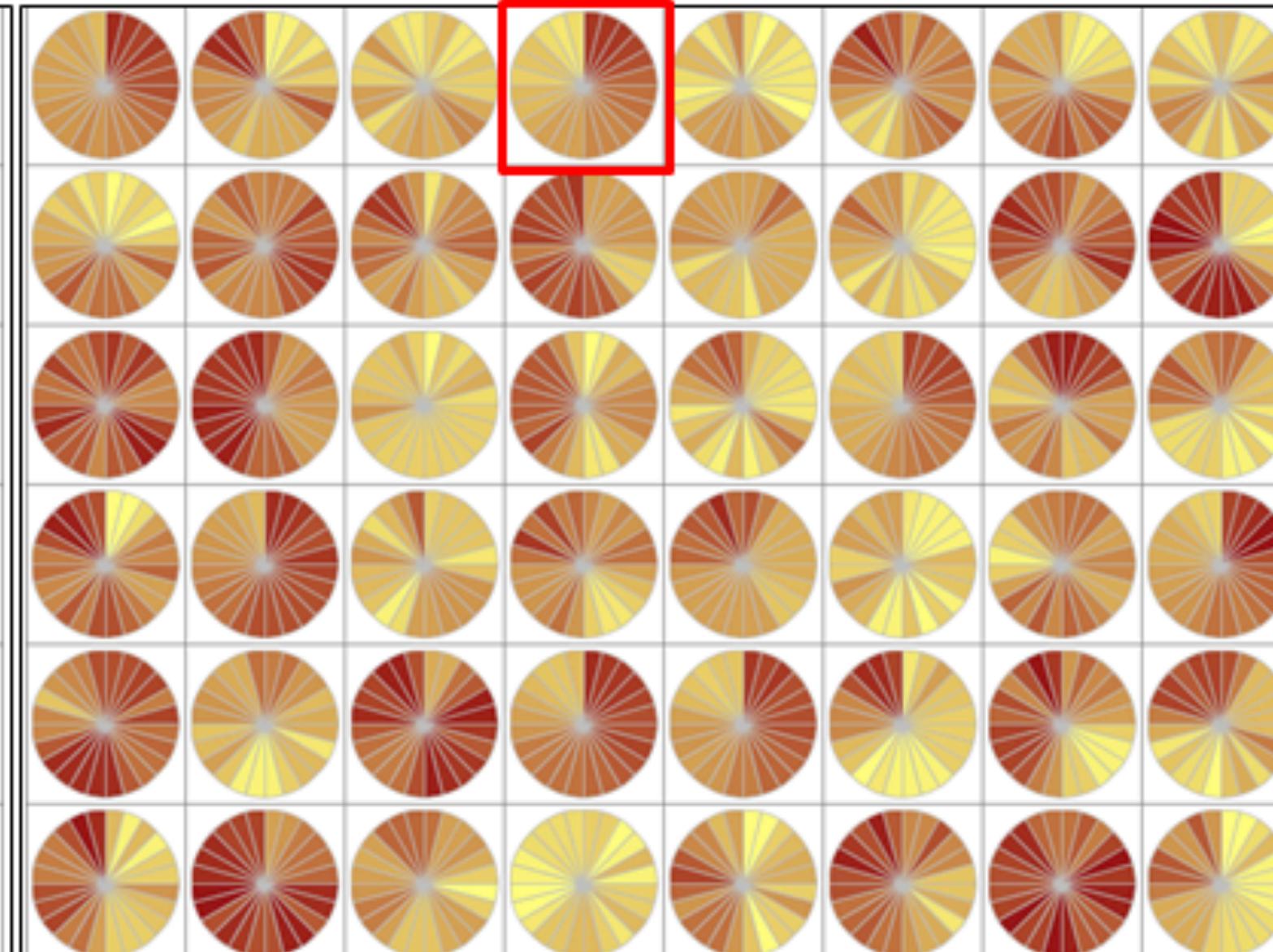
Star



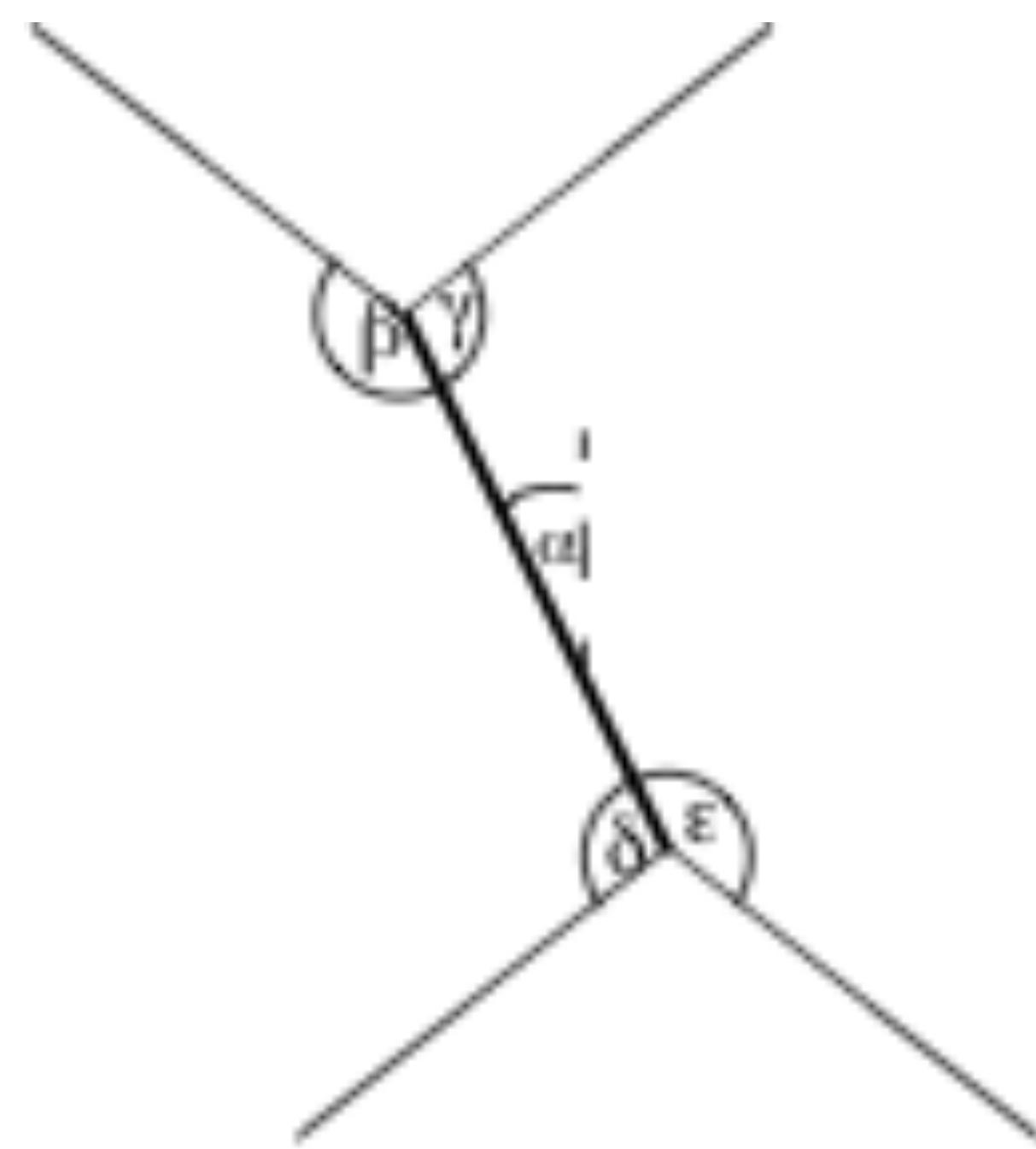
Stripe



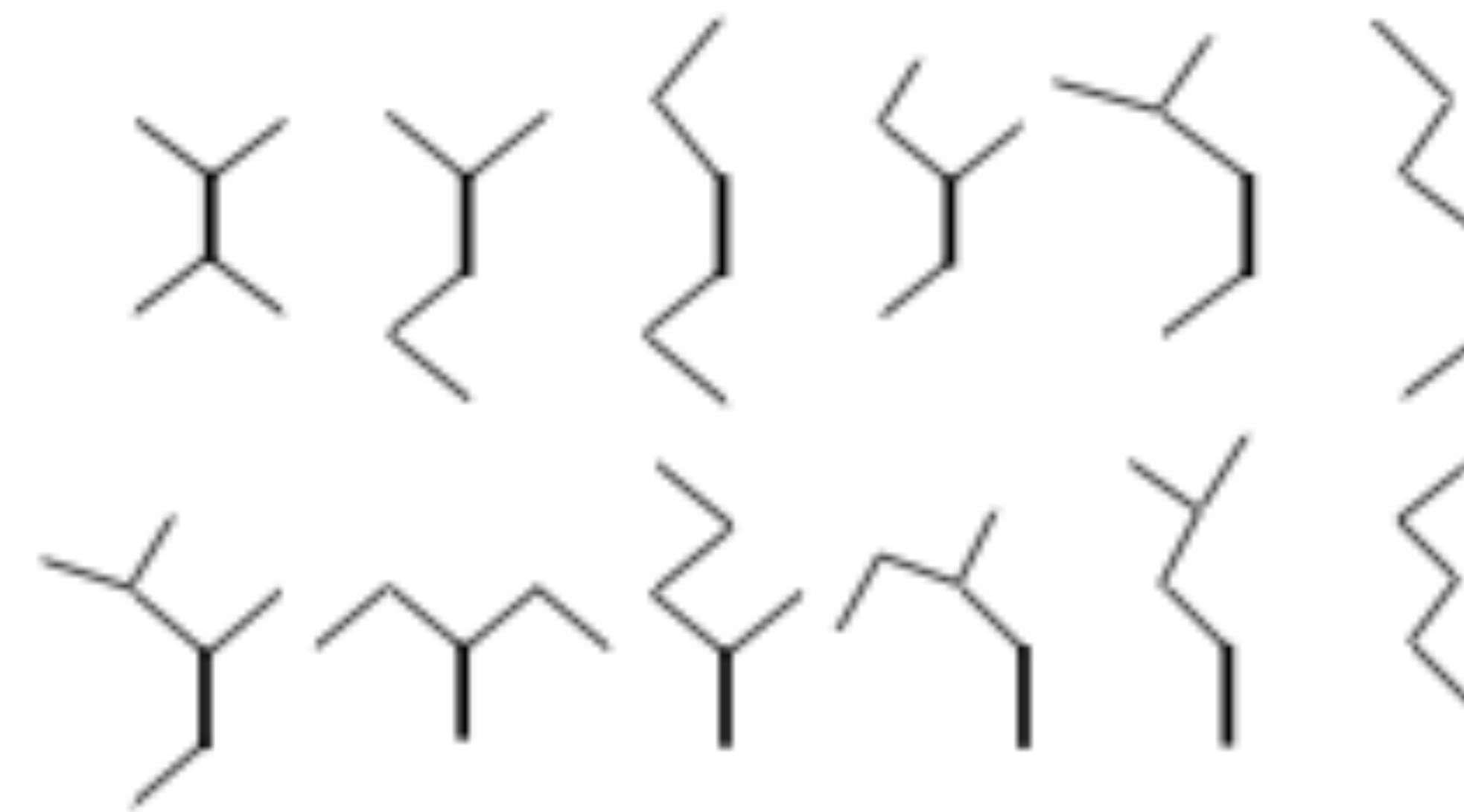
Clock



REPRESENTAÇÕES BASEADAS EM VARETAS



a. Stick Figure Icon



b. A Family of Stick Figures

Figure 5: Stick Figure Visualization Technique

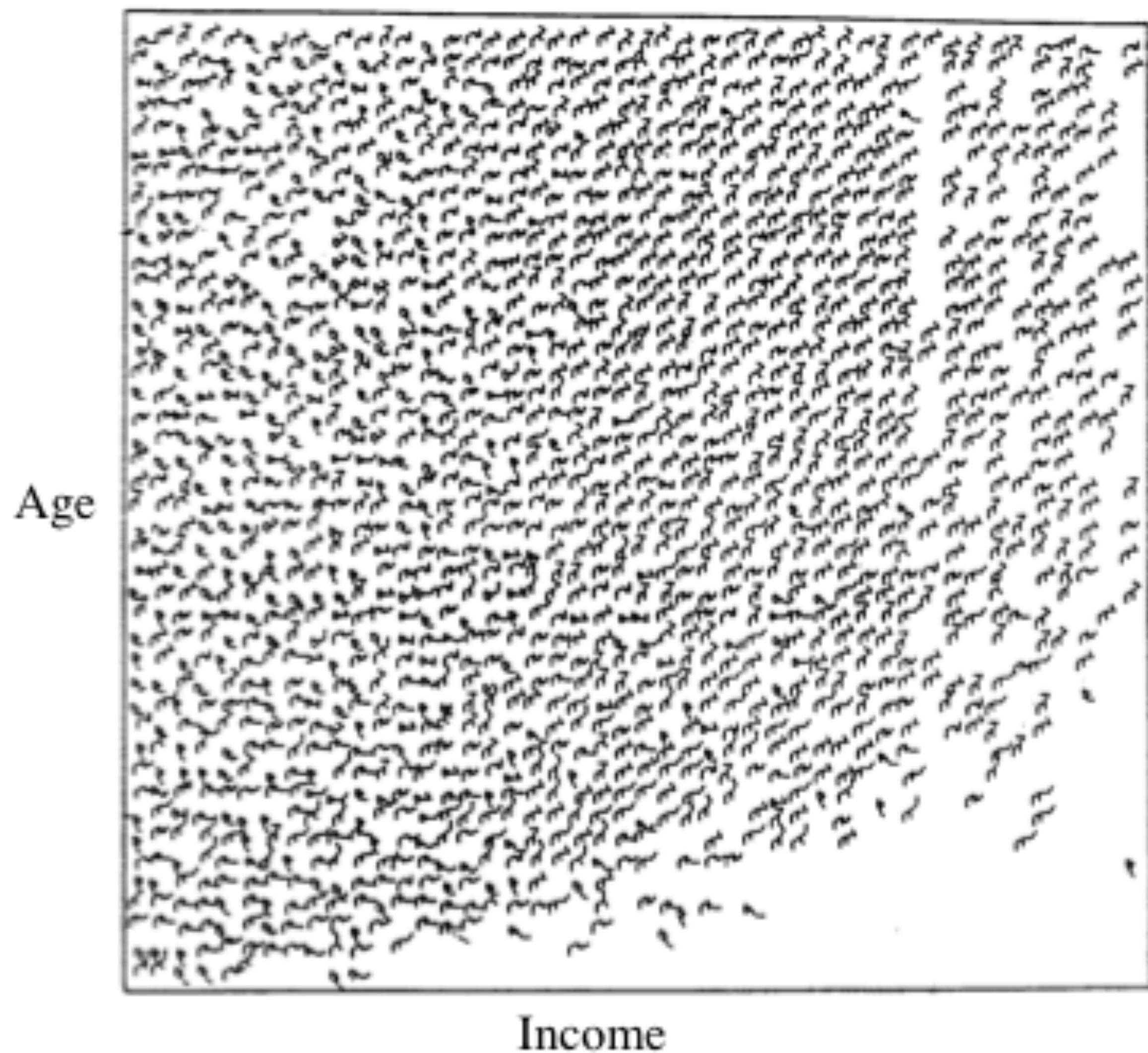


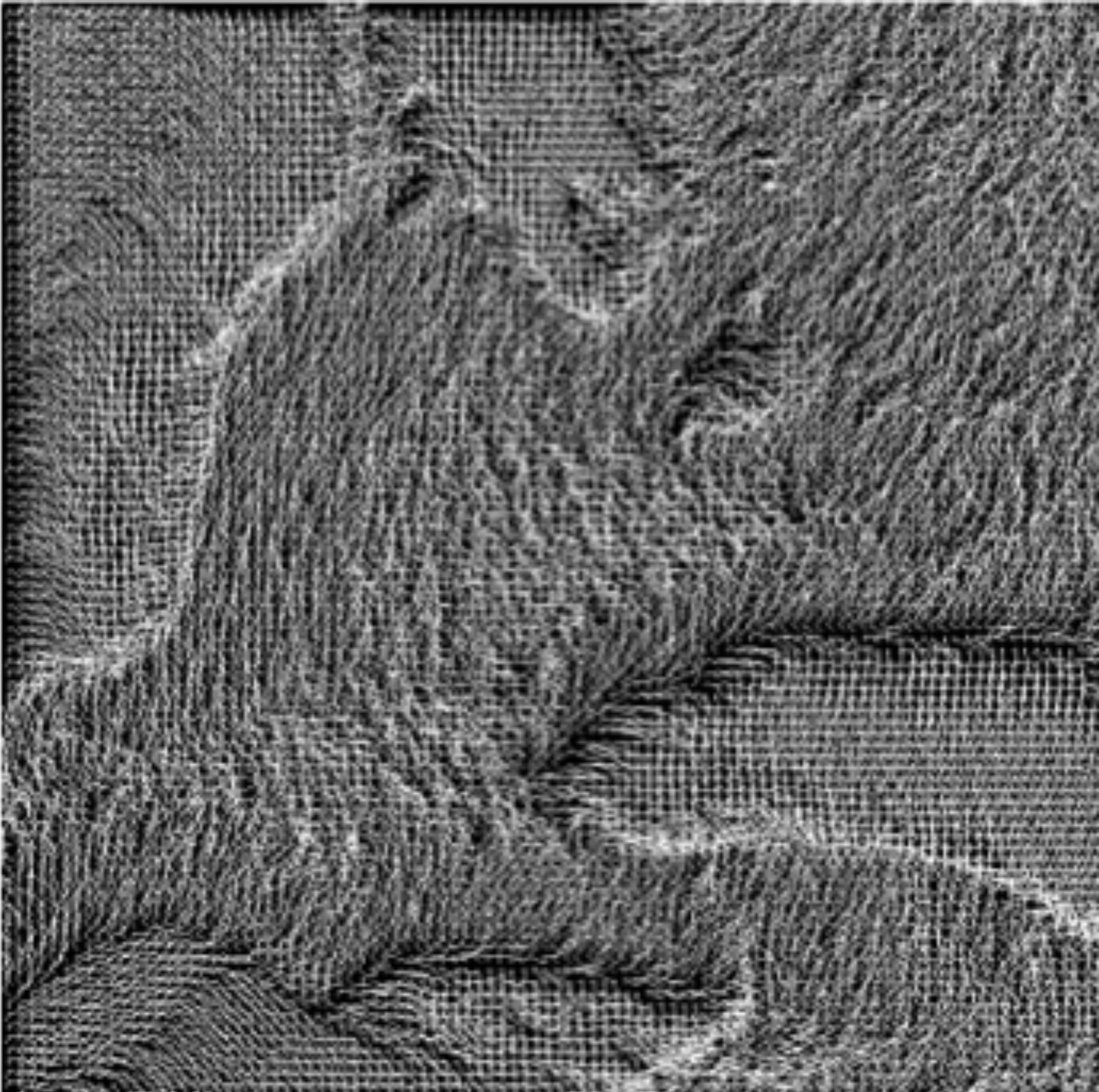
Figure 6: Stick Figure Visualization of Census Data

(used by permission of G. Grinstein, Institute of Visualization and Perception Research,
University of Massachusetts at Lowell; cf. [GPW 89])

(a)



(b)



(c)

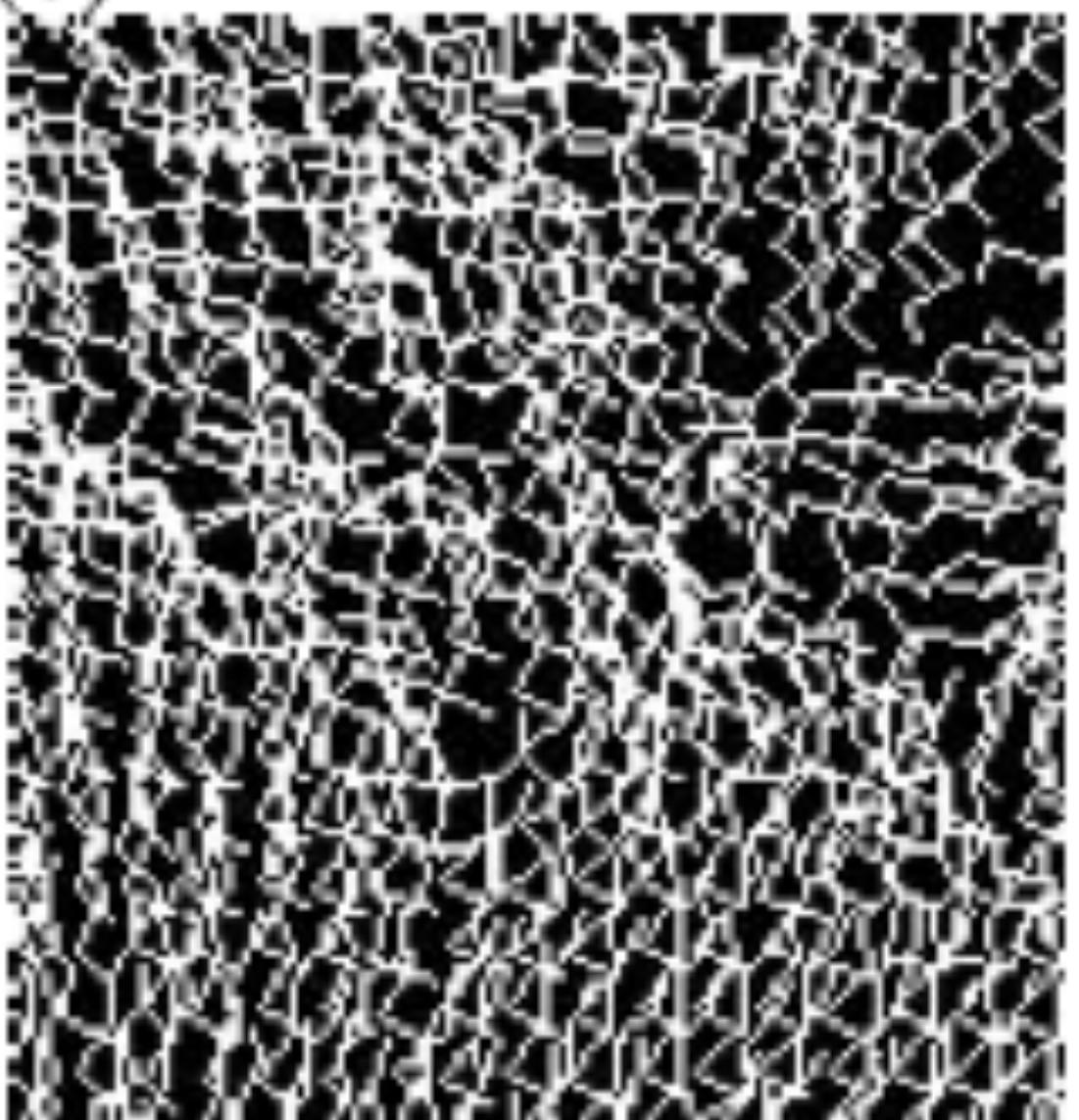
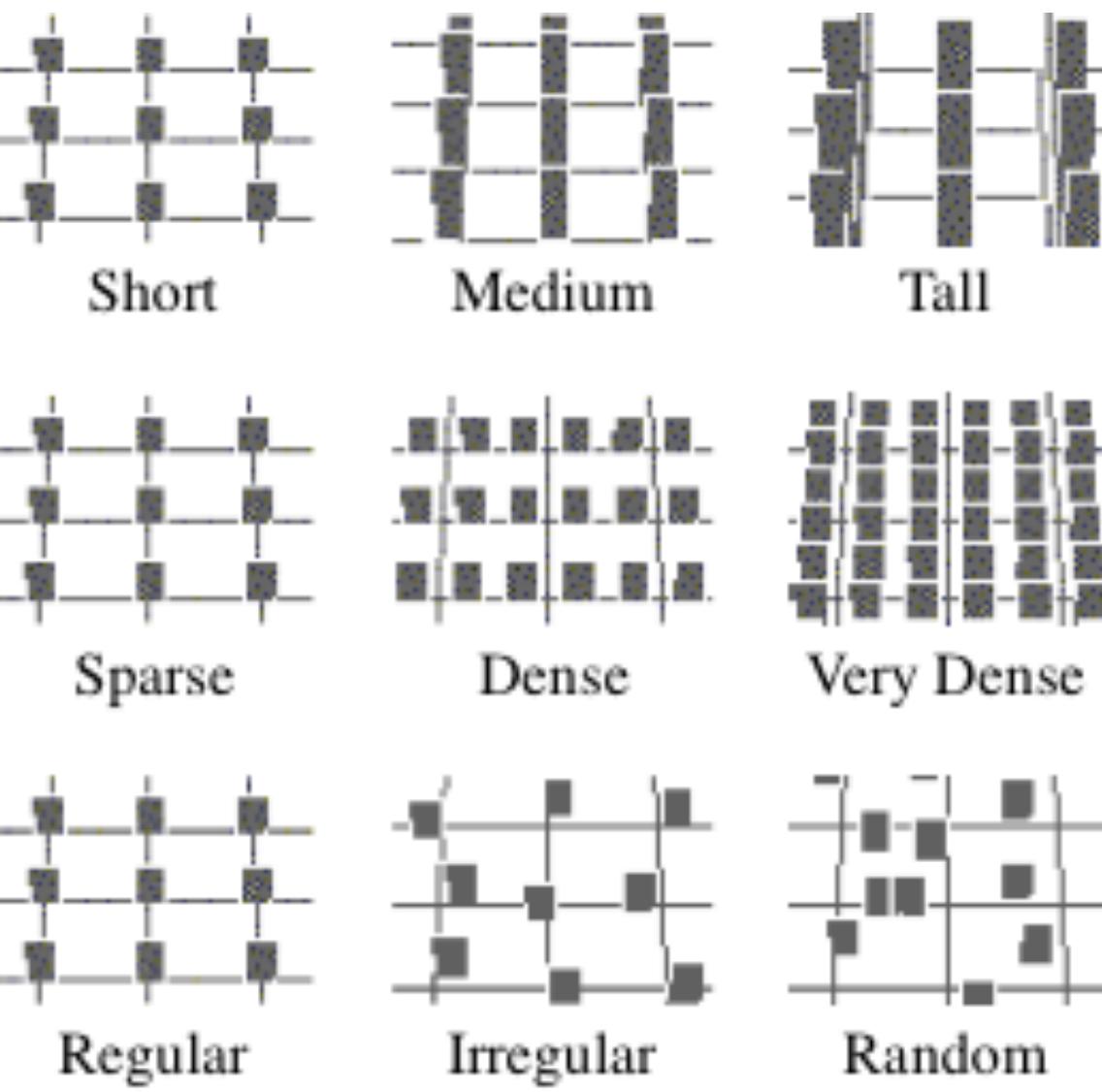
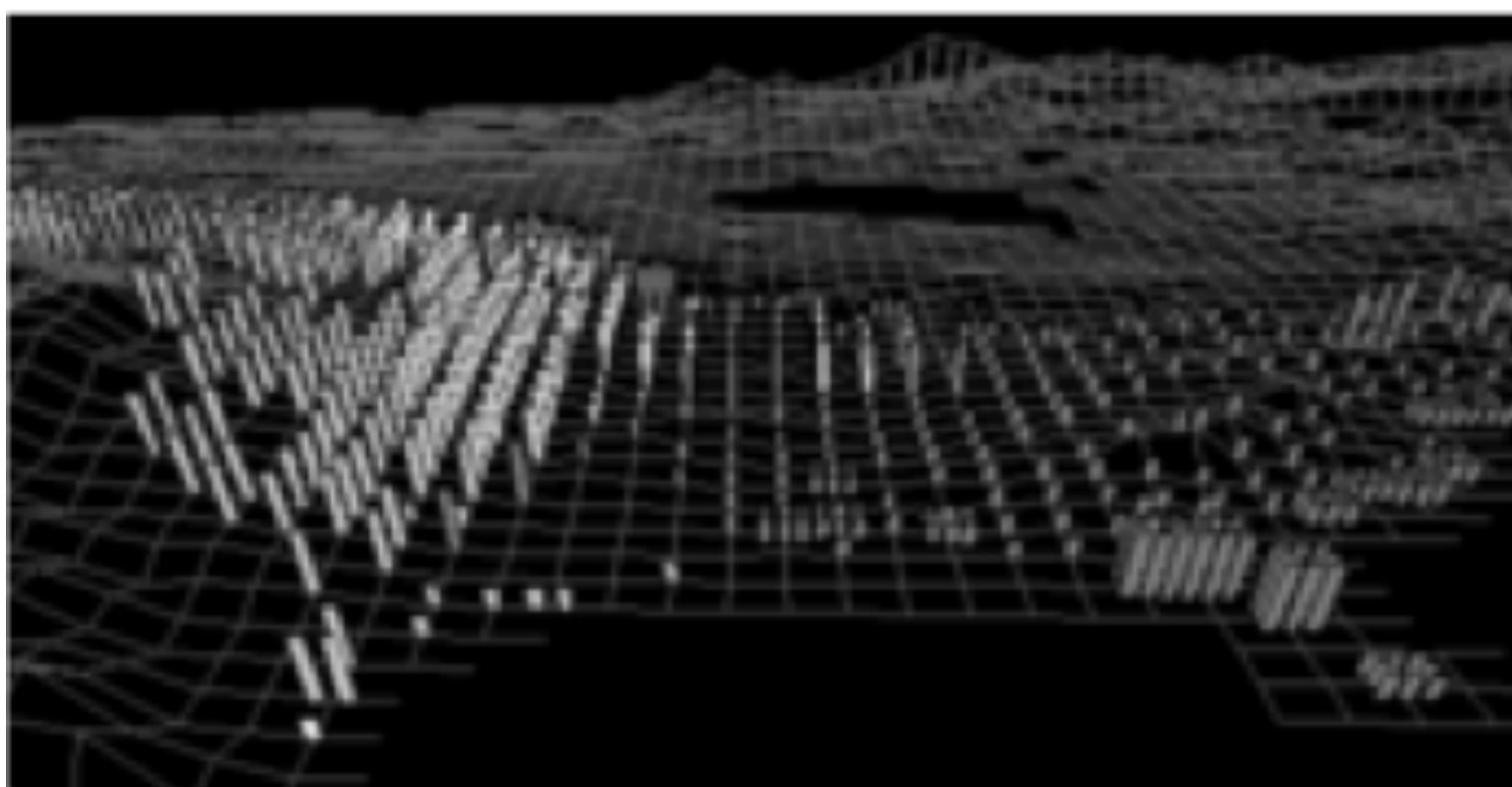


Figure 3.23: (a) Stick figure family [23], (b) 5D image data using stick figures,
(c) Part of (b) in original size [60].

REPRESENTAÇÕES BASEADAS EM CARTÕES



(a)



(b)

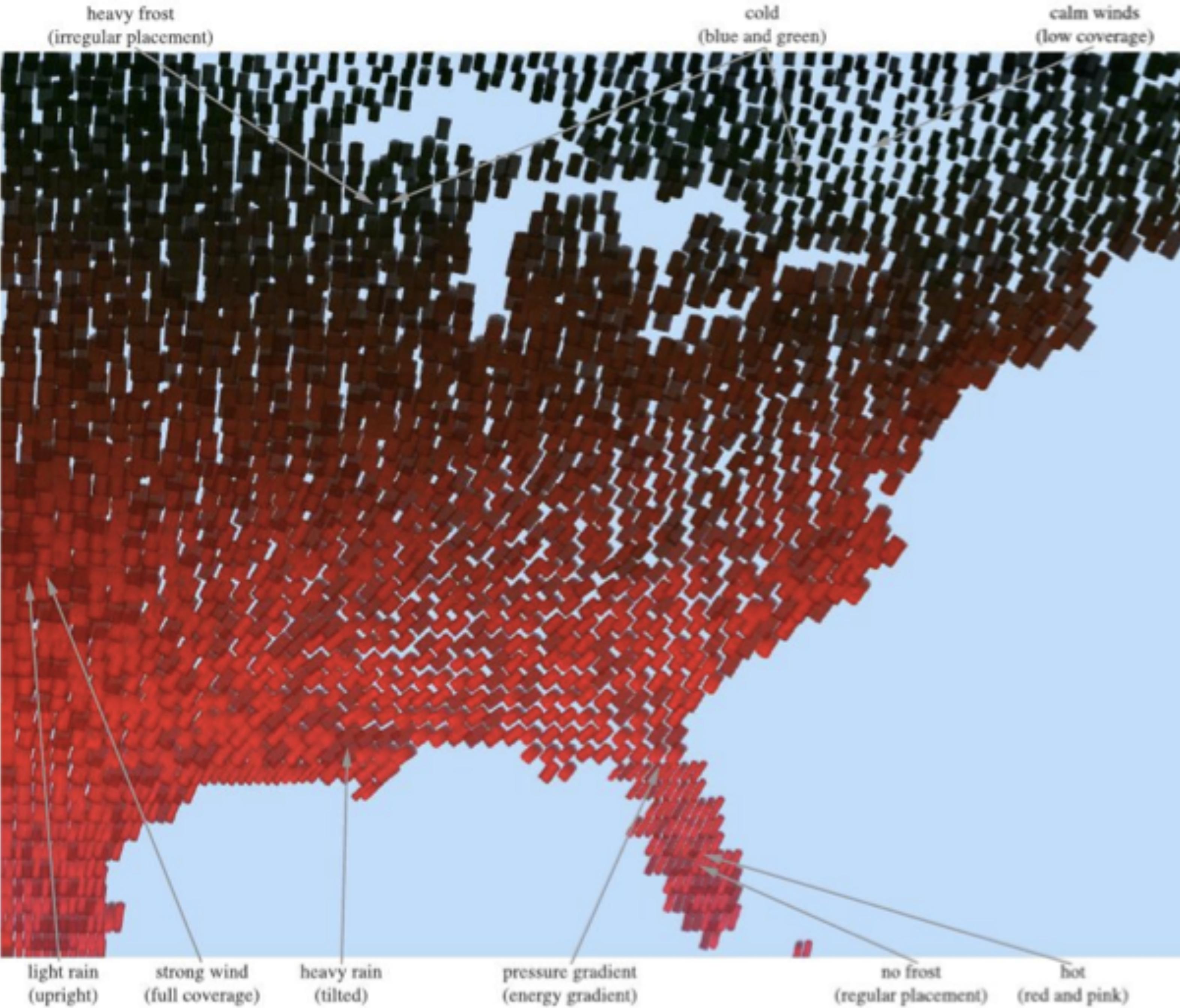


Figure 3.28: Nonphotorealistic visualization of weather conditions [68].

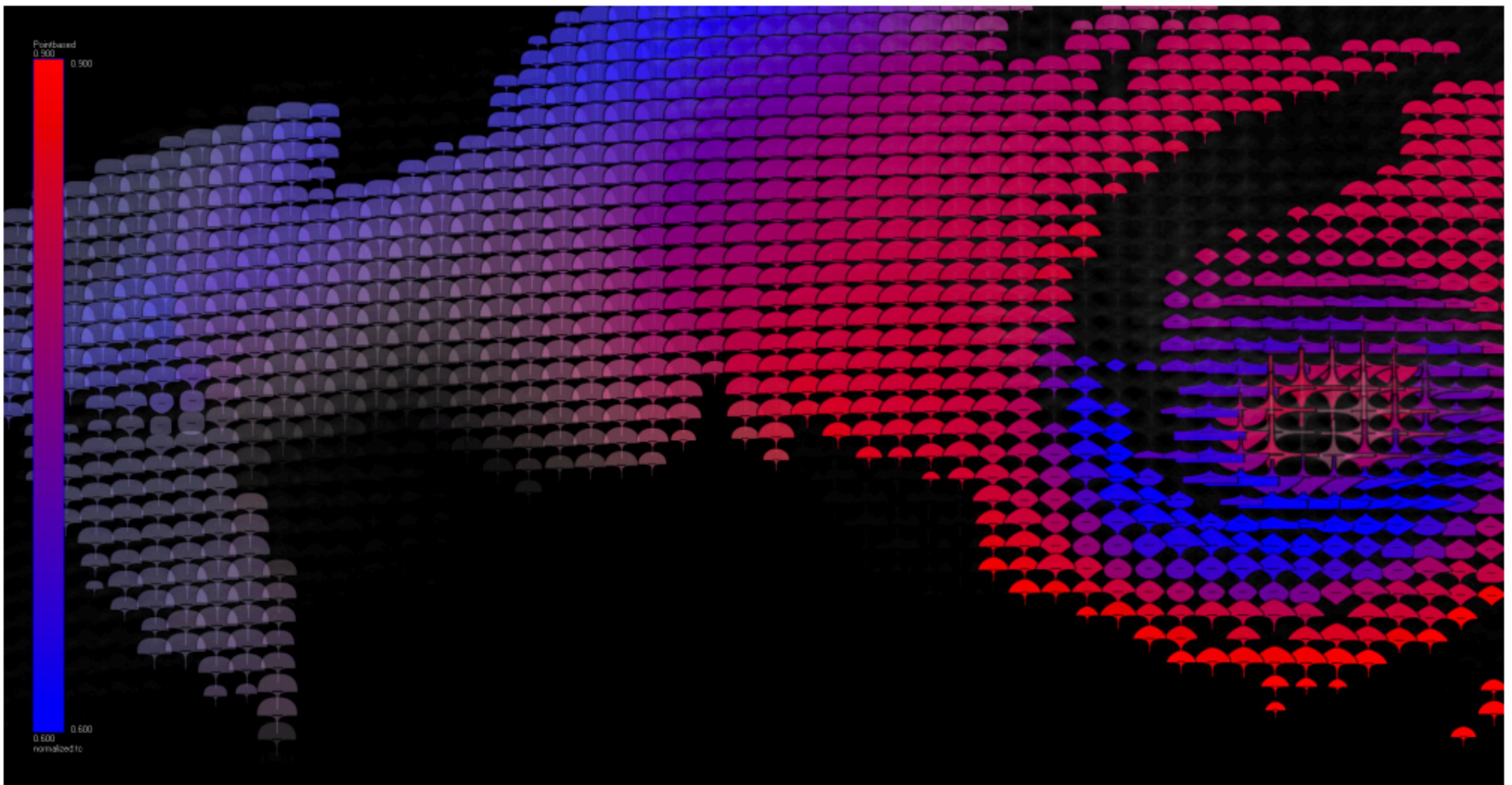


Figure 10: The hurricane Isabel dataset, timestep 12. Color represents temperature, and the amount of clouds is mapped to glyph size. The upper shape represents pressure, and the lower shape precipitation. The visualization depicts the fast moving clouds (specified via brushing), and the eye of the hurricane is visible in the lower right, surrounded by high amounts of precipitation and cold airflow.

Lie, Andreas E., Johannes Kehrer, and Helwig Hauser. "Critical design and realization aspects of glyph-based 3D data visualization." Proceedings of the 25th Spring Conference on Computer Graphics. ACM, 2009.

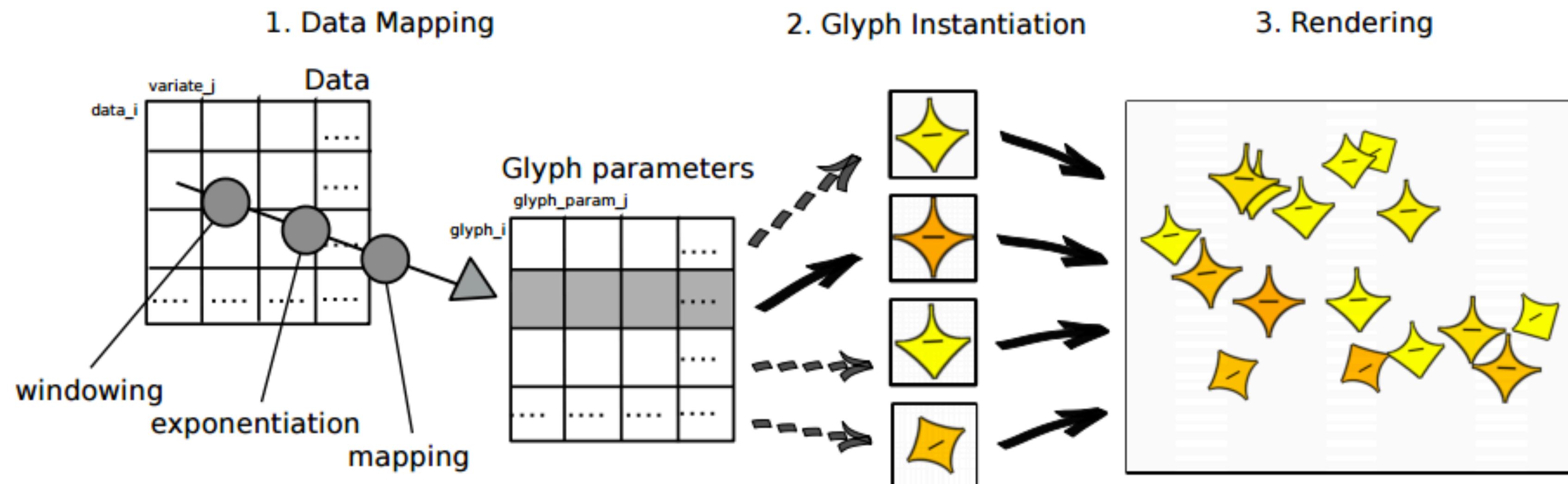


Figure 2: Data variates undergo data mapping stages; windowing, exponentiation and mapping. These values are then used to instantiate the corresponding glyphs, (e.g, determining shape, size) and finally the glyphs are rendered into the context.

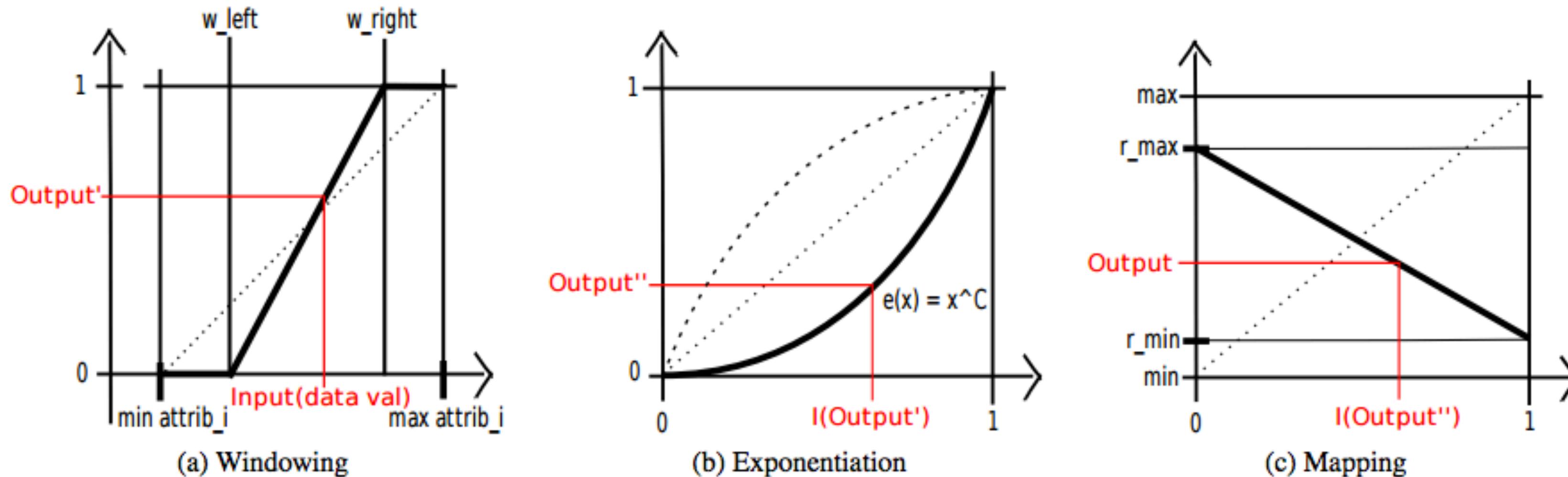
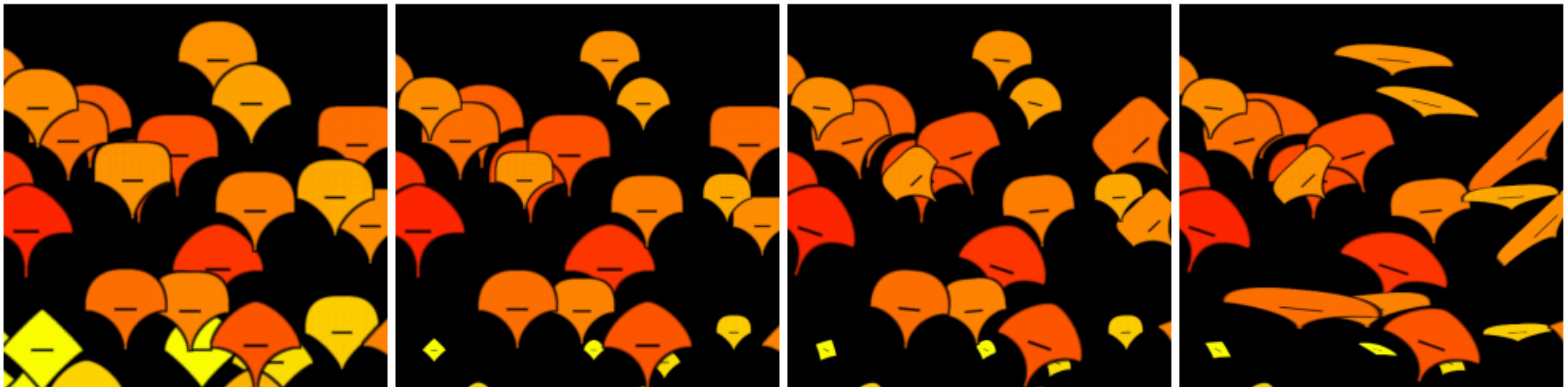


Figure 3: (a) User selects w_{left} and w_{right} which adjusts the data mapping ramp for input datavalues. The dotted line represents the default value for the ramp. (b) User can control the data mapping curve by adjusting the C term in x^C . The middle dotted line represents the default curve, the curve can be bent upwards or downwards. (c) The output range of the datamapping can be adjusted to fit the datavariates. User selects r_{min} and r_{max} to clamp the output range.



(a) Two data attributes are represented as the upper / lower glyph shape (b) Added data attribute to overall glyph size (c) Glyph rotation has been assigned a data attribute as well (d) A data attribute has been assigned to glyph aspect ratio

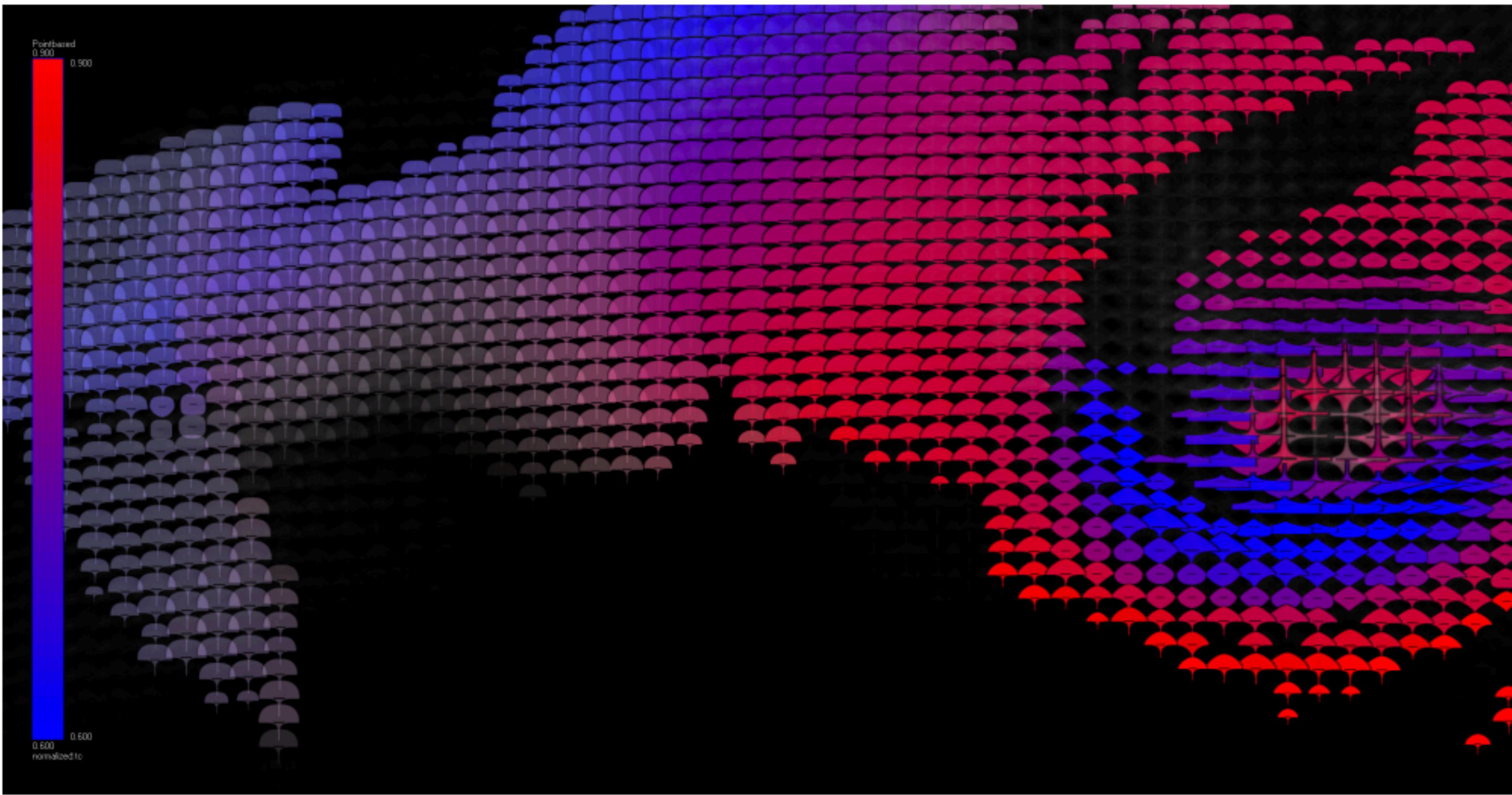


Figure 10: The hurricane Isabel dataset, timestep 12. Color represents temperature, and the amount of clouds is mapped to glyph size. The upper shape represents pressure, and the lower shape precipitation. The visualization depicts the fast moving clouds (specified via brushing), and the eye of the hurricane is visible in the lower right, surrounded by high amounts of precipitation and cold airflow.

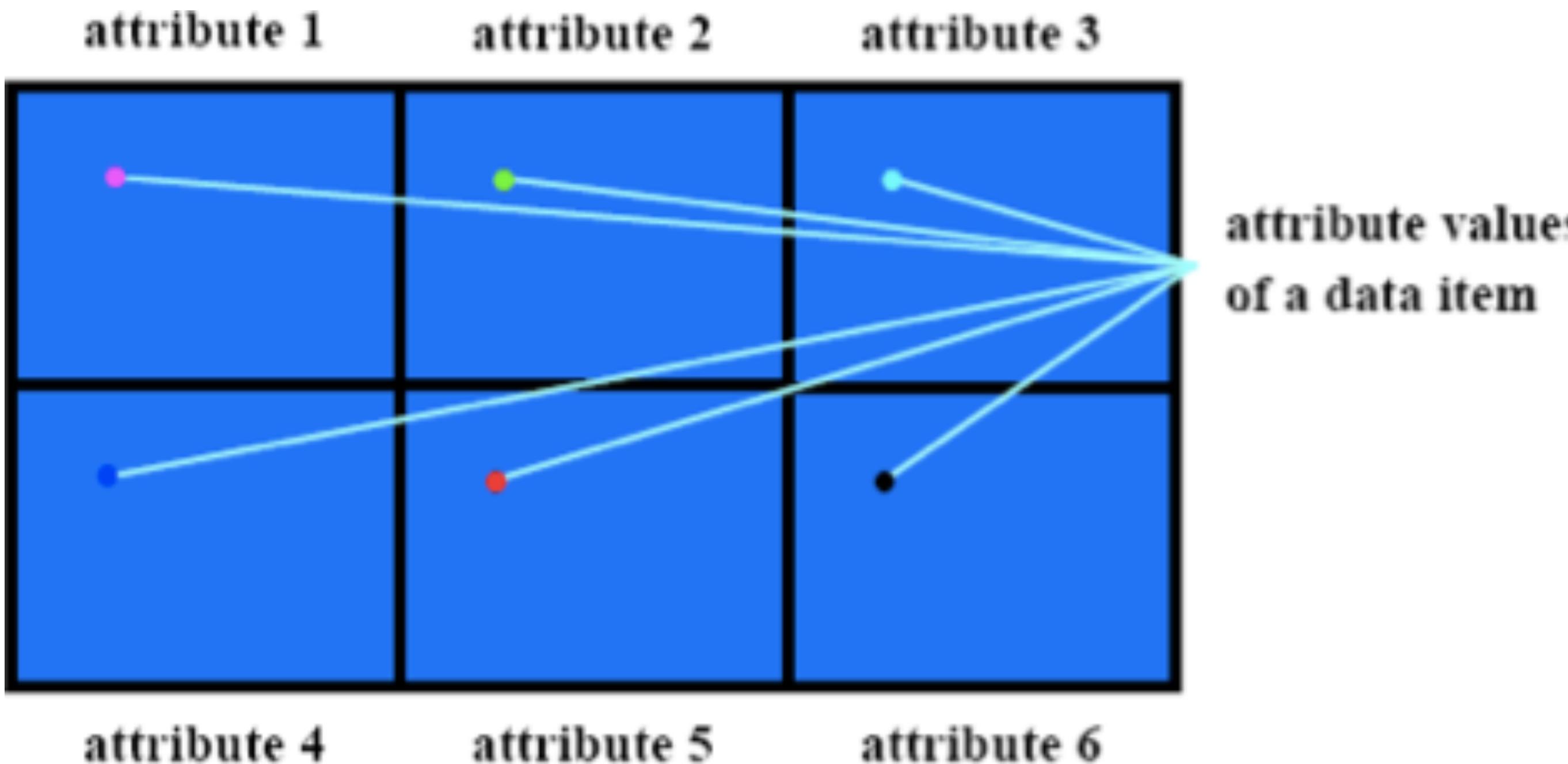
Color	Temperature
Glyph Upper	Pressure
Glyph Lower	Precipitation
Glyph Size	Clouds

Table 2: Glyph property mapping for Hurricane Isabel dataset

TÉCNICAS ORIENTADAS A PIXELS

TÉCNICAS ORIENTADAS A PIXELS

- Consiste na representação de um atributo por um pixel baseado em alguma escala de cores: cada cor indica um valor possível
- Para um conjunto de dados n-dimensional, n pixels coloridos são necessários para representar cada item



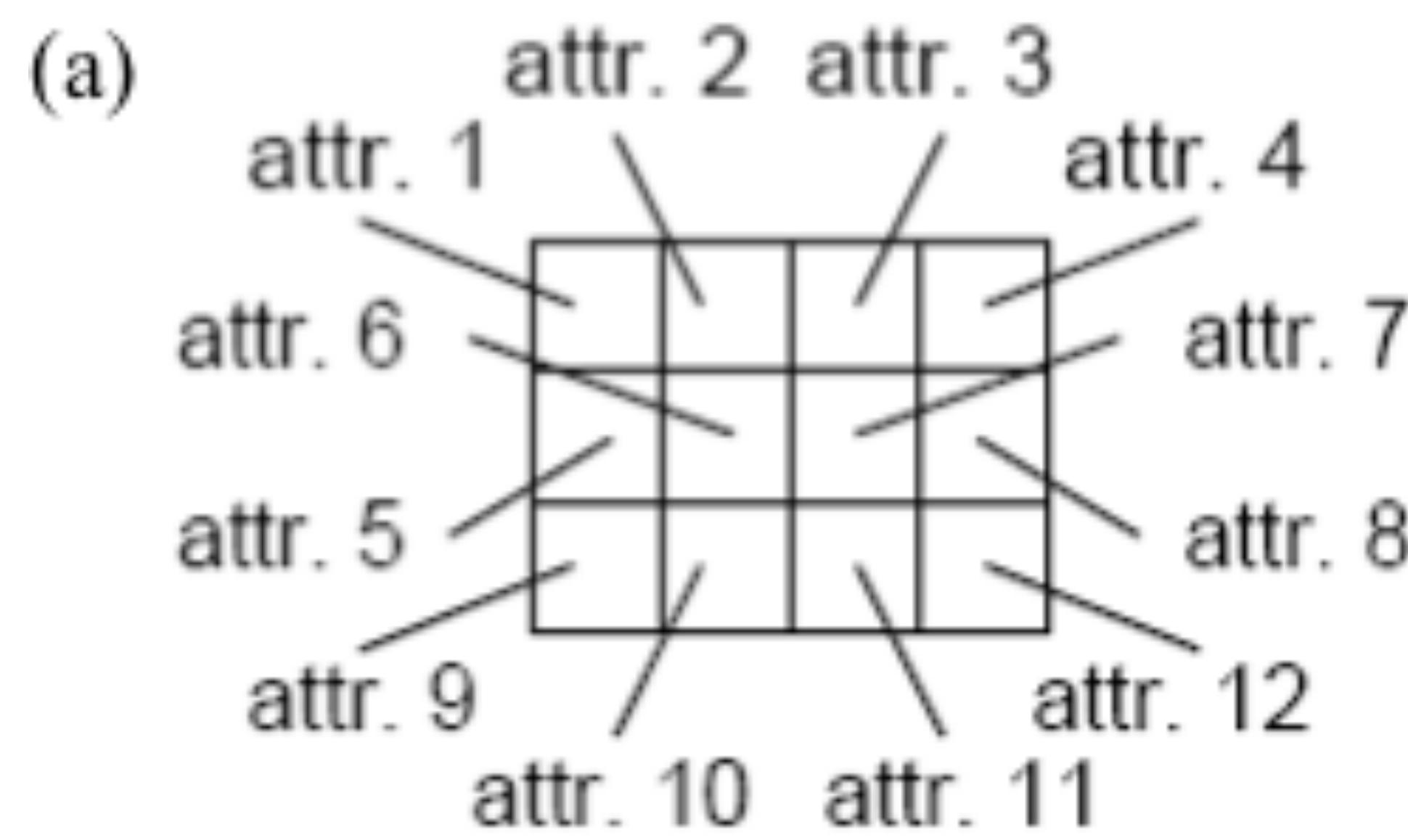


Figure 3.24: (a) Shape coding array [15], (b) An example [61].

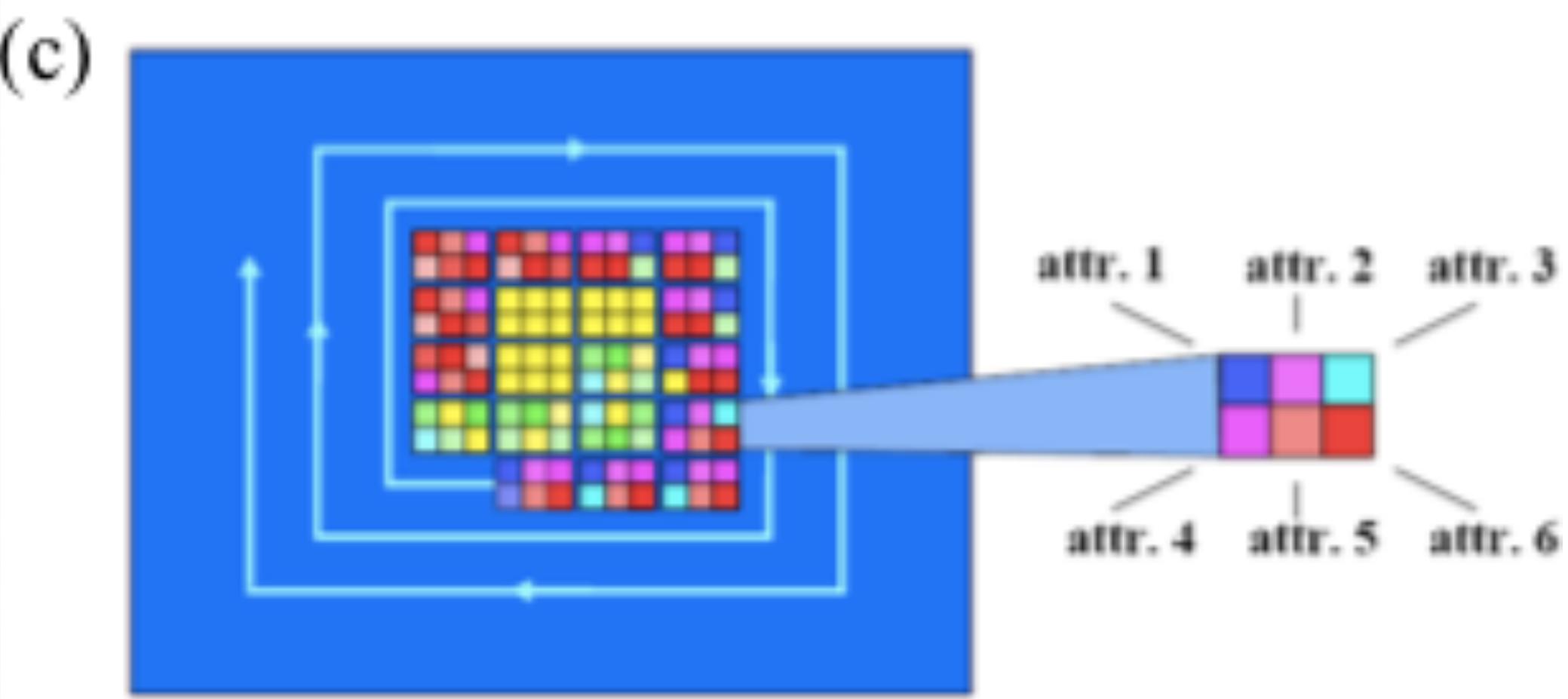
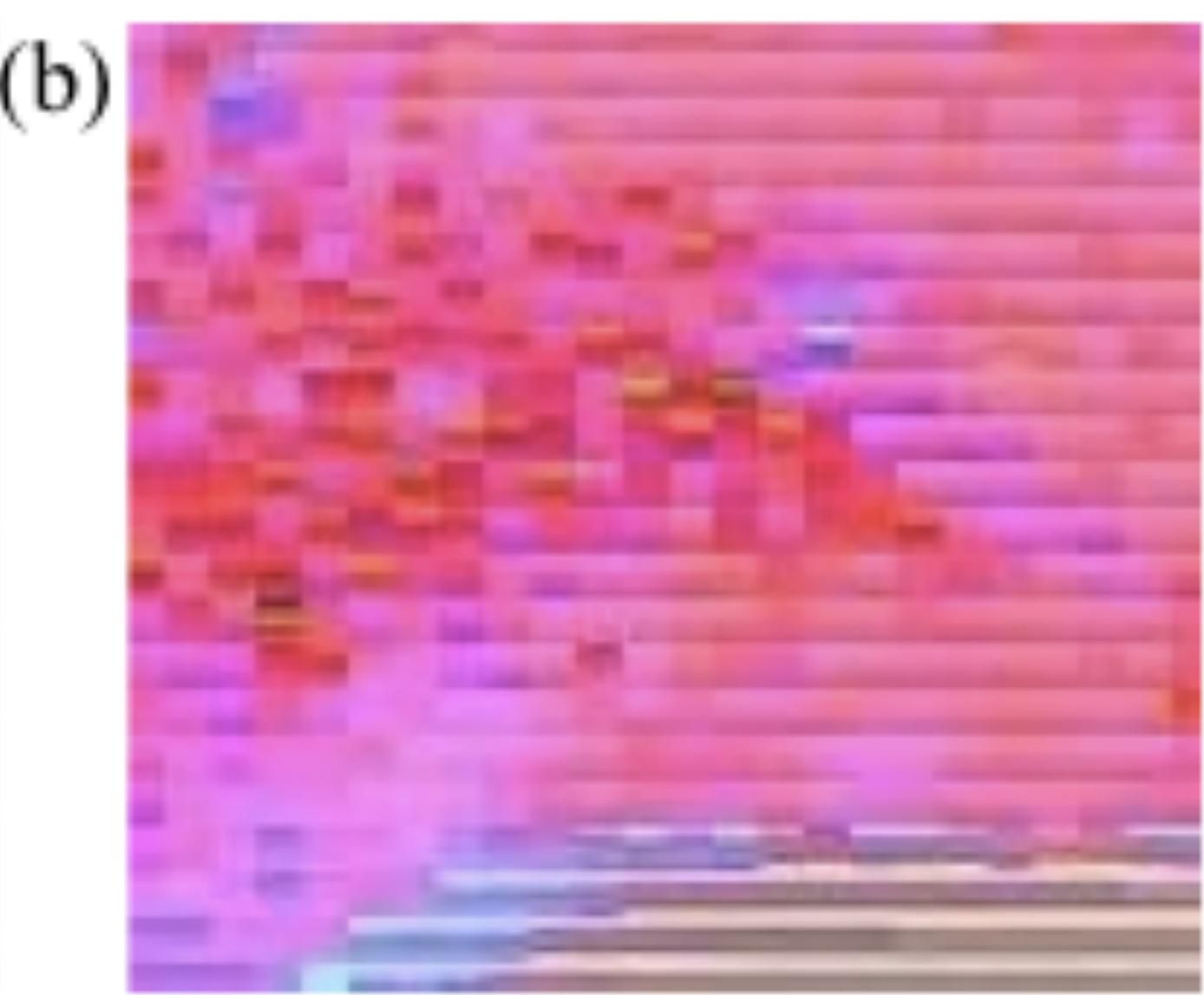
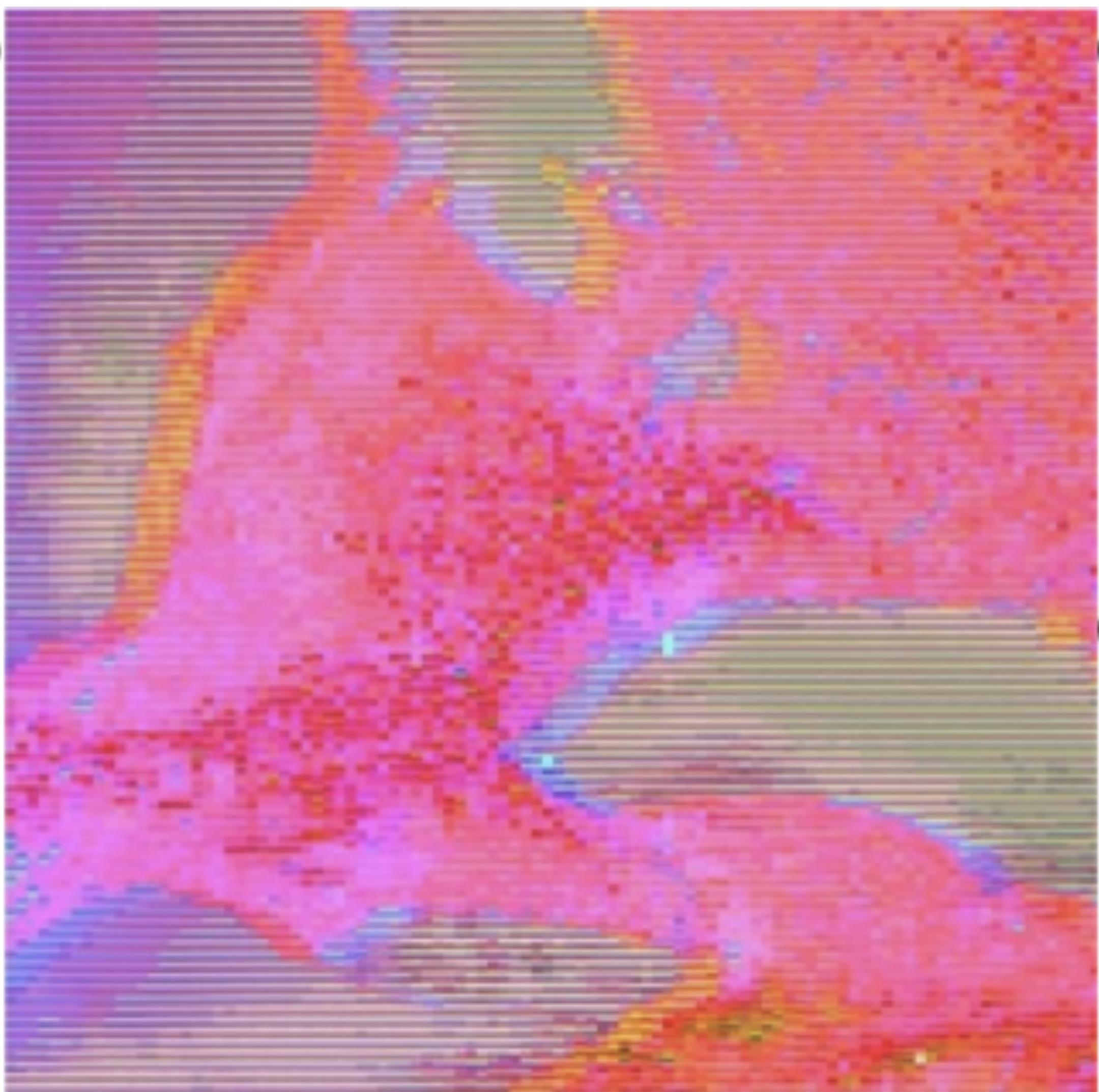
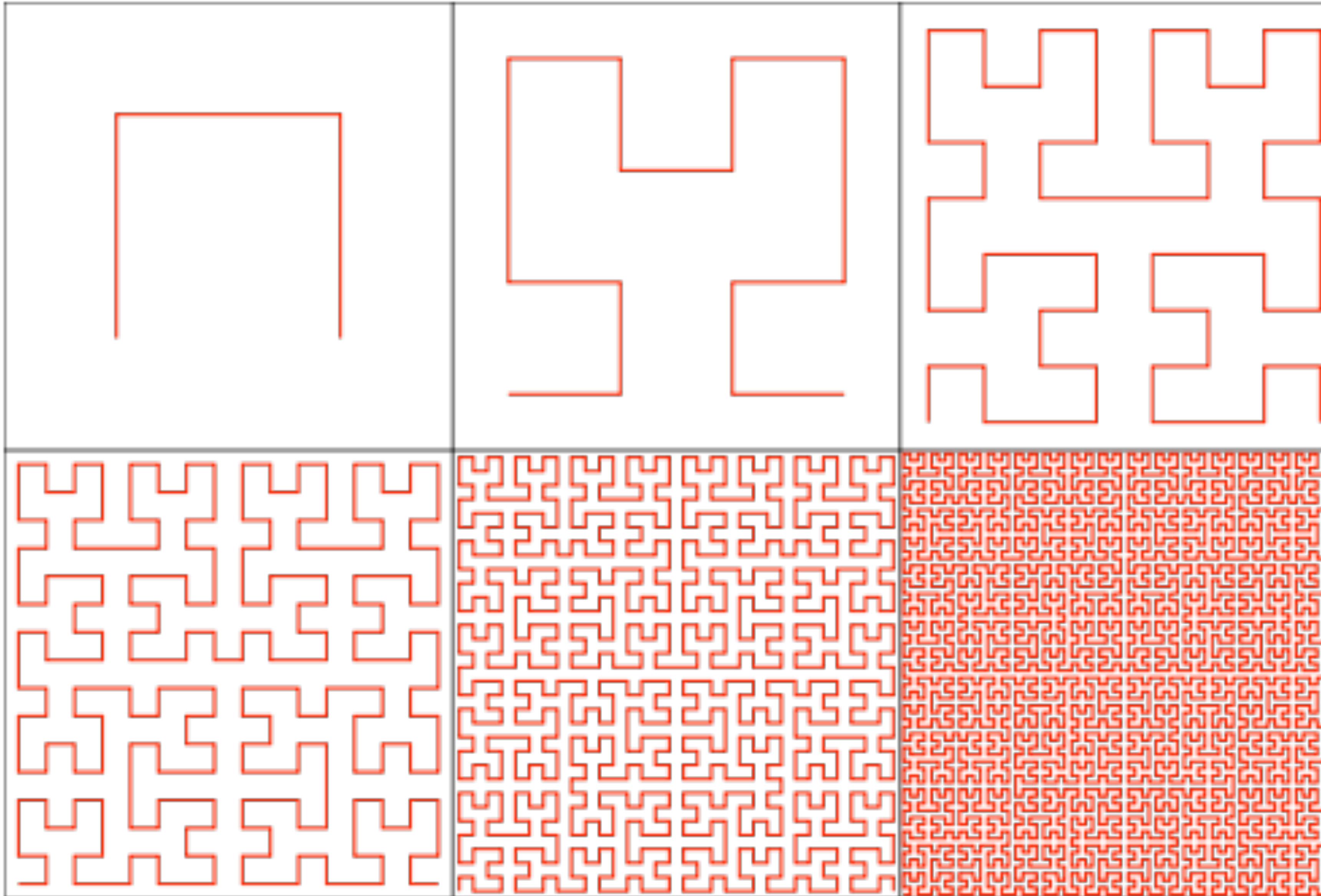


Figure 3.25: (a) 5D image data using color icons, (b) Part of (a) in original size [60],
(c) Color icon scheme [15].

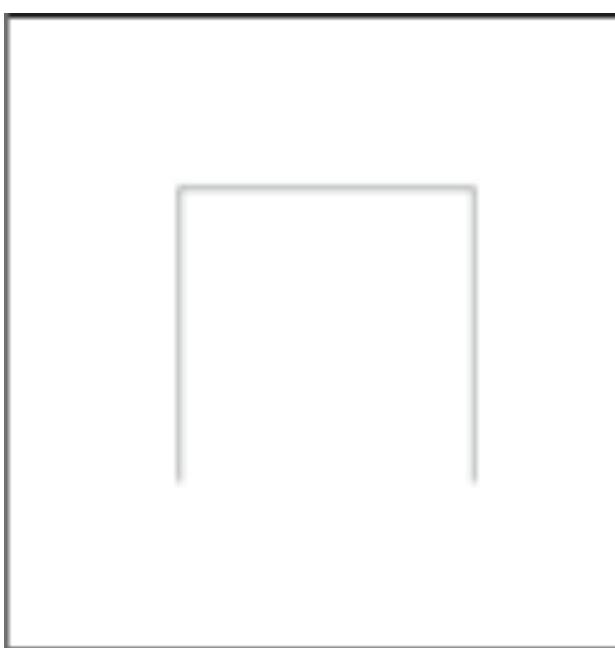
TIPOS DE ABORDAGENS BASEADAS EM PIXELS

- Independentes de consulta:
 - os pixels são coloridos de acordo com os atributos reais de cada item
 - normalmente são interessantes quando algum atributo propicia uma ordenação natural dos itens
- Dependentes de consulta
 - as cores representam a diferença entre os valores da consulta e de cada item representado

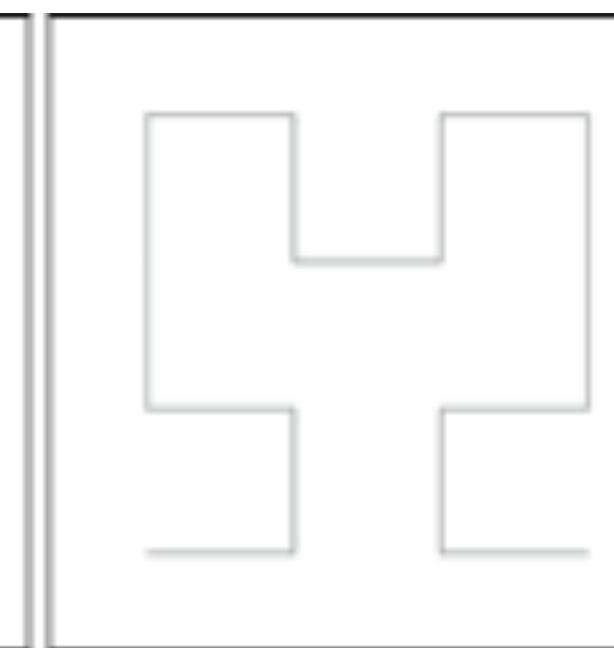
EXEMPLO DE ABORDAGEM INDEPENDENTE DE CONSULTA



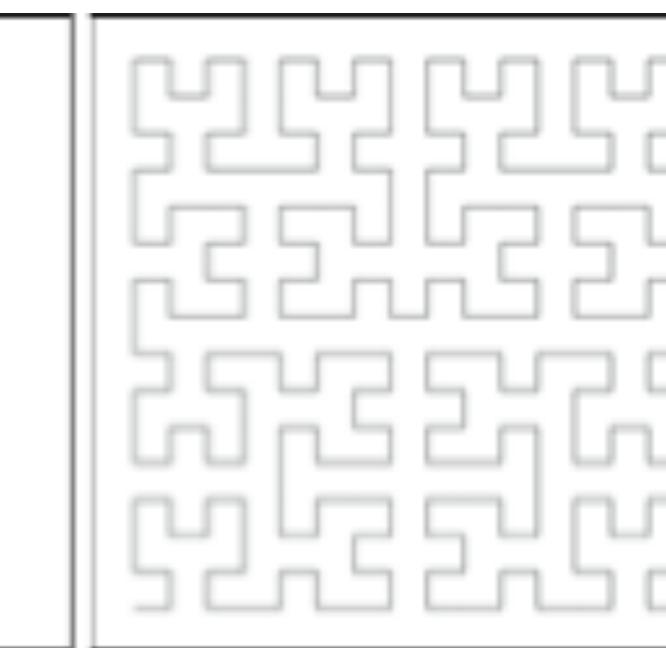
EXEMPLO DE ABORDAGEM INDEPENDENTE DE CONSULTA



(a) Hilbert order 1



(b) Hilbert order 2



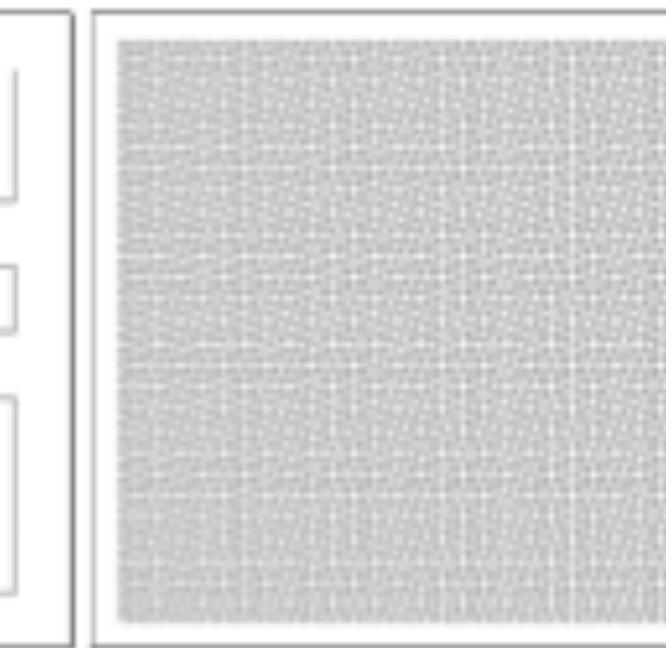
(c) Hilbert order 4



(d) Peano order 1



(e) Peano order 2



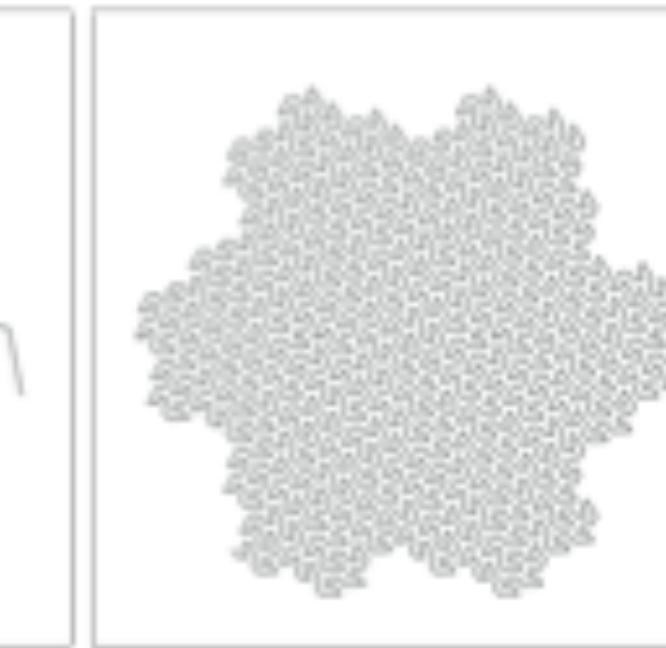
(f) Peano order 4



(g) Gosper order 1



(h) Gosper order 2



(i) Gosper order 4

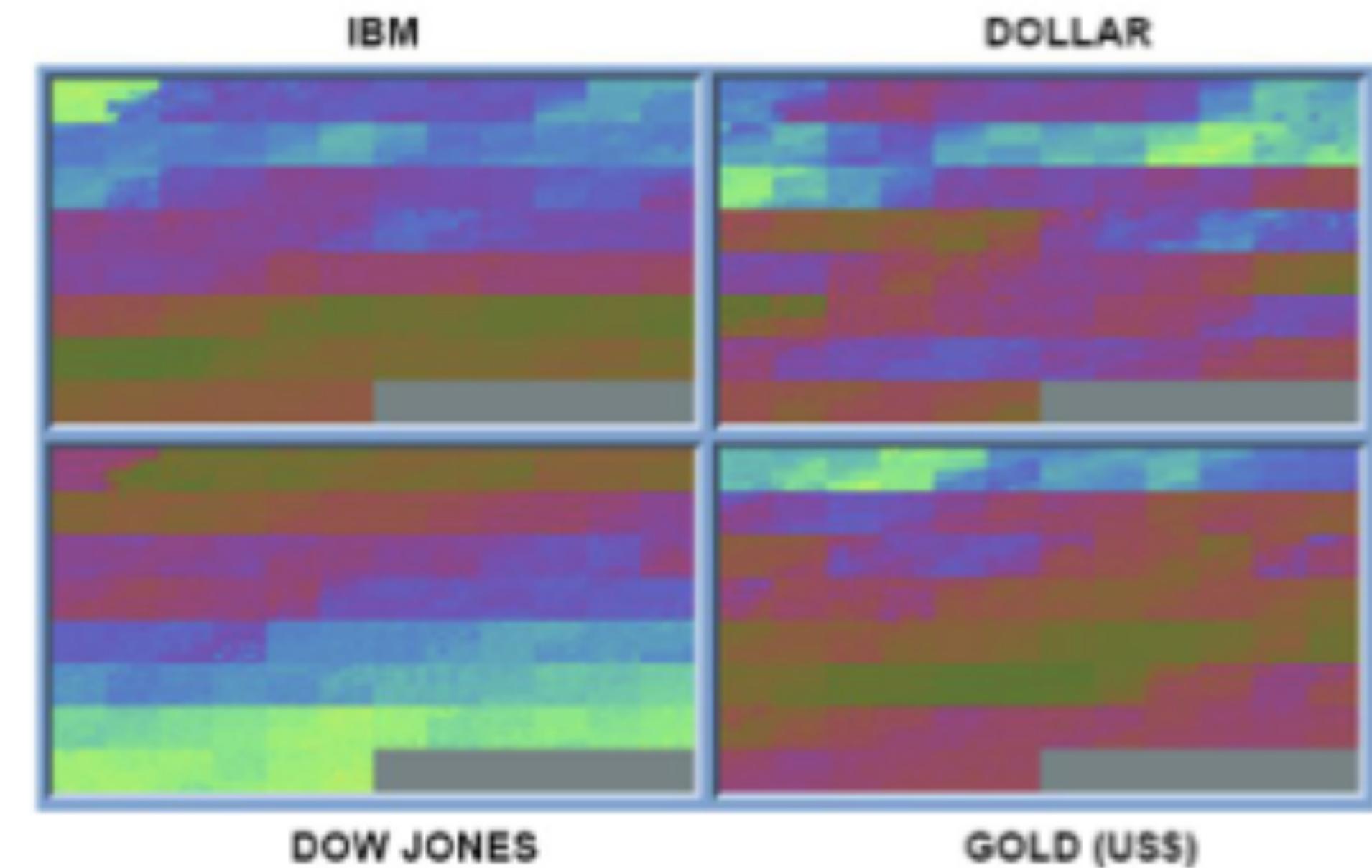
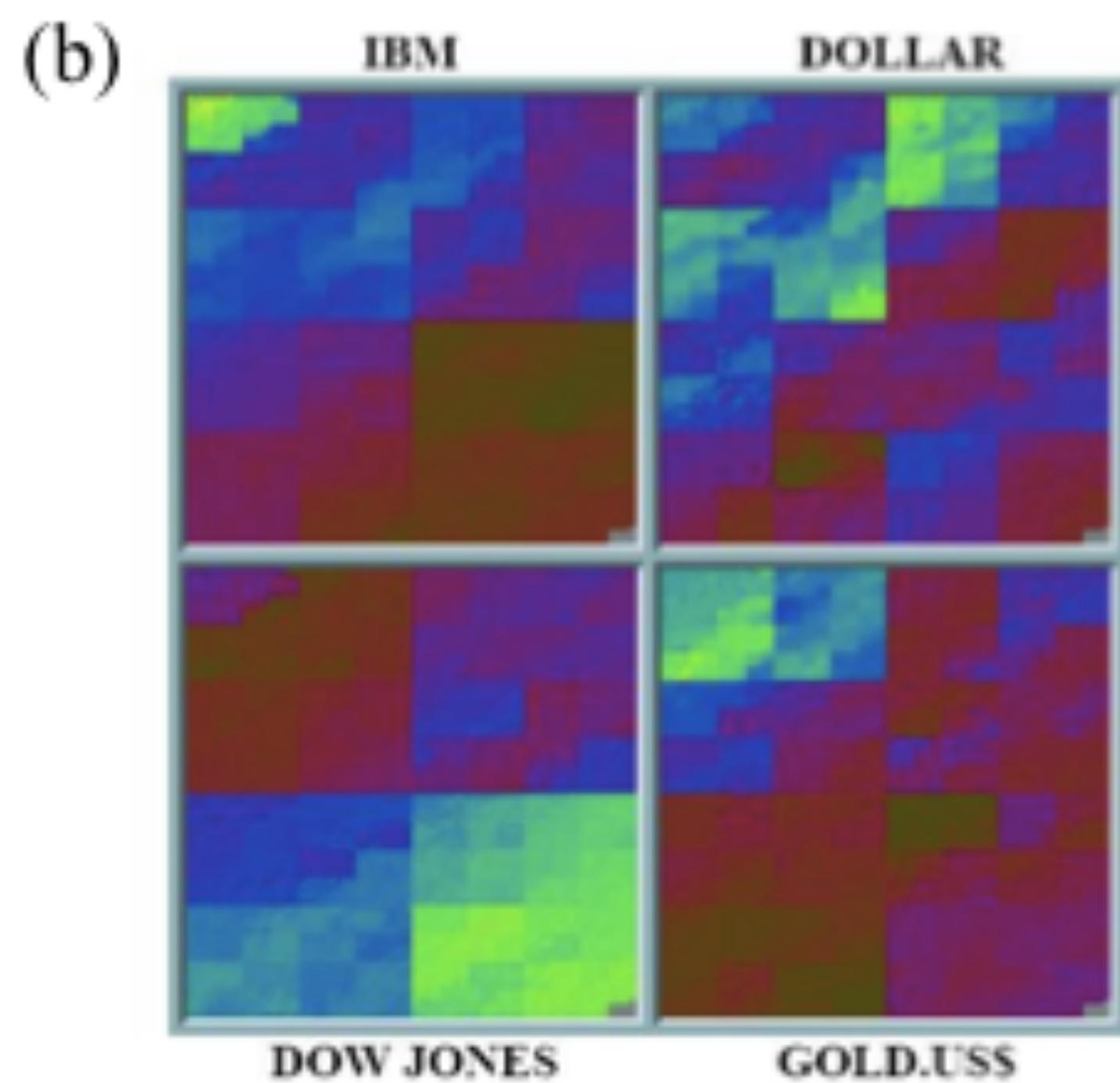
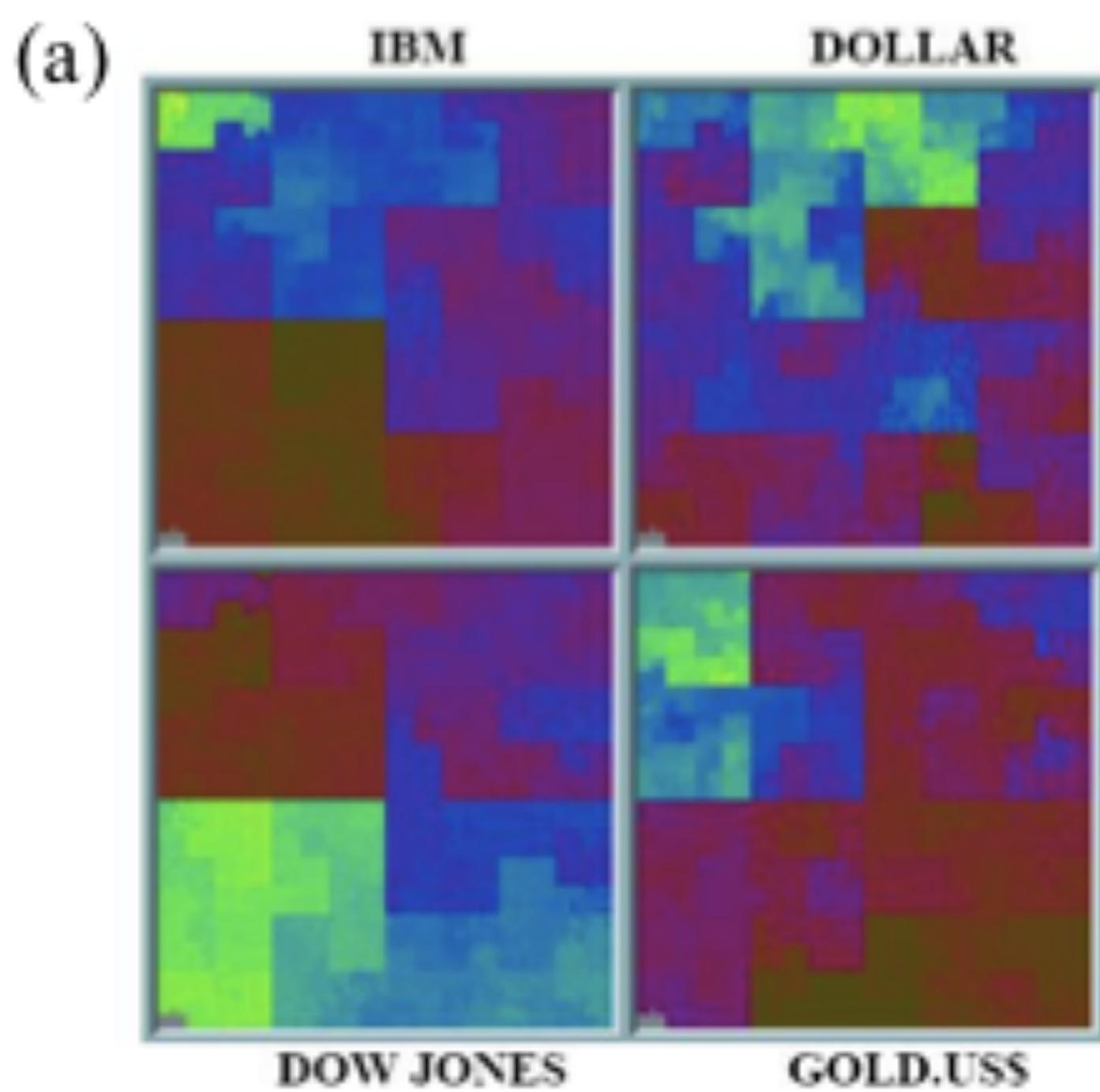
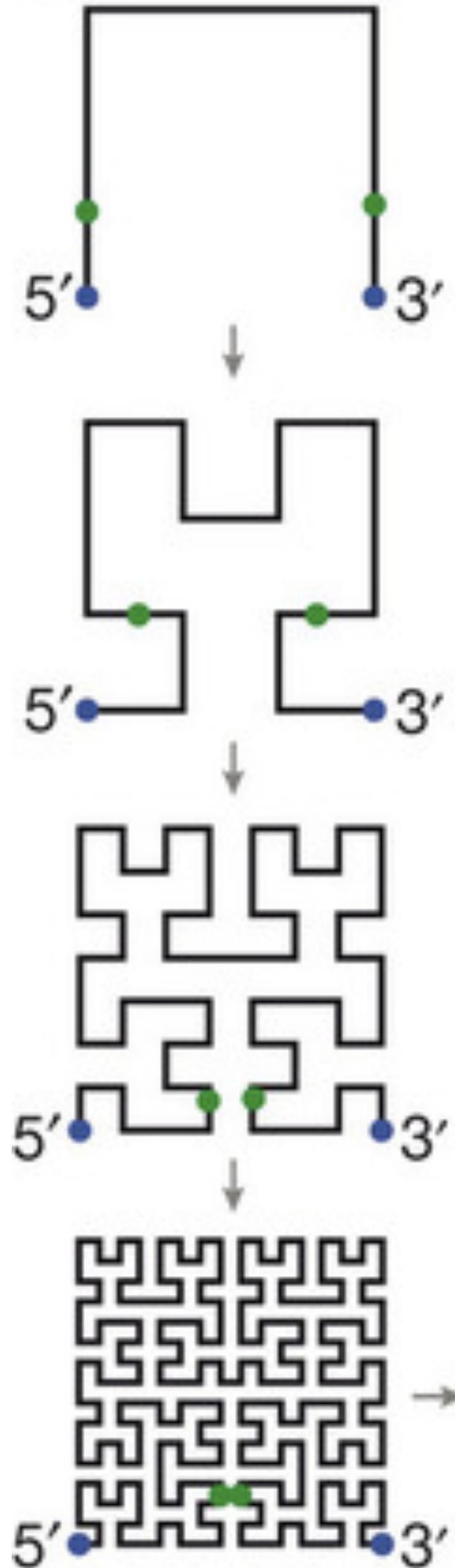
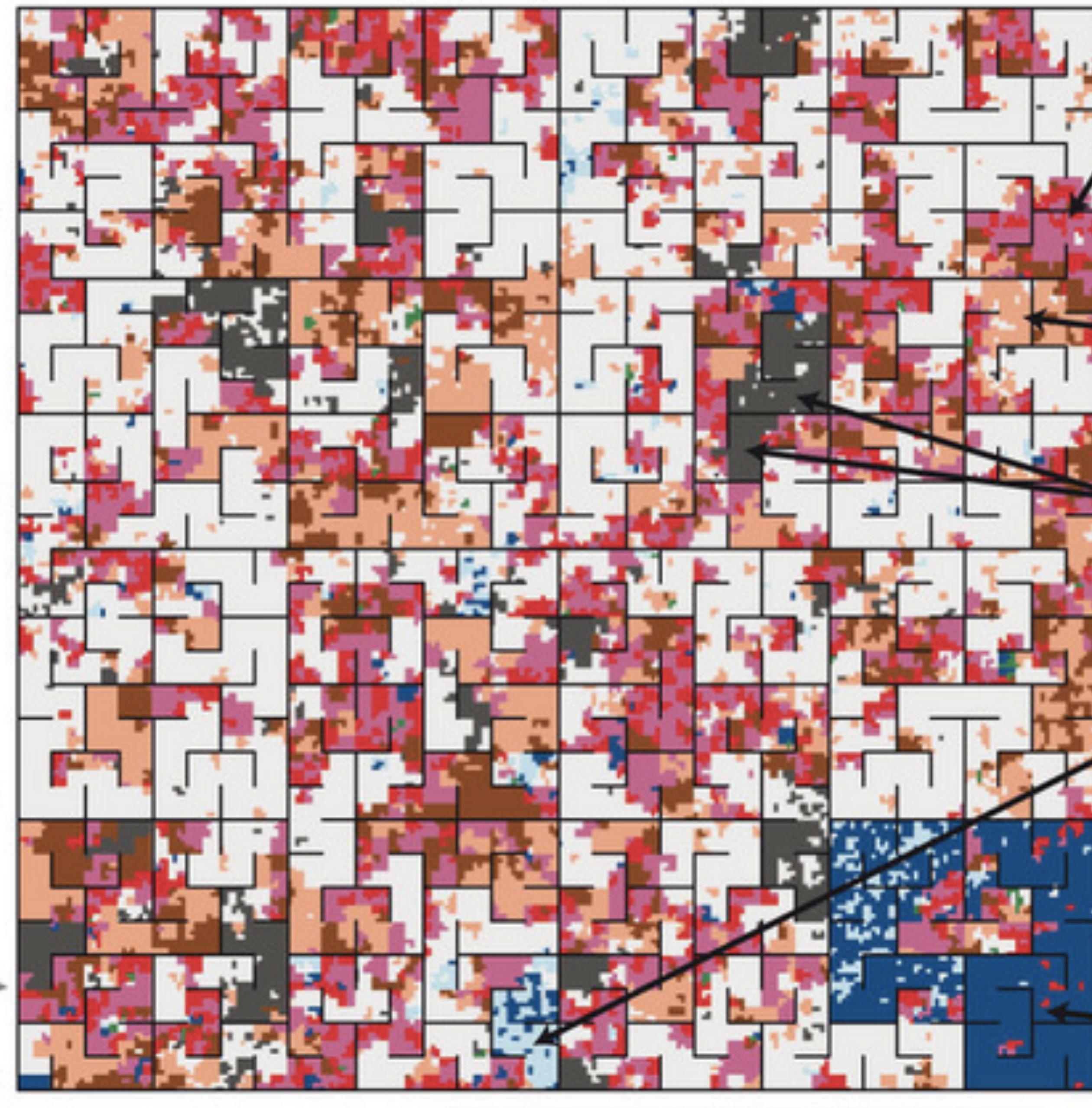


Figure 3.14: (a) Peano-Hibert, (b) Morton or Z-Curve, (c) Recursive Pattern [15].

a**b**

Chromosome 3L



Cluster of small expressed genes
Open chromatin domain
PcG domains
Heterochromatin-like domain
Pericentromeric heterochromatin

Chromatin states: 1 2 3 4 5 6 7 8 9

BG3 - CHR2L



1 2 3 4 5 6 7 8 9

BG3	CHR2L	CHR3R	S2	CHR2L	CHR3R
	CHR2R			CHR2R	CHR4
	CHR3L			CHR3L	CHRX

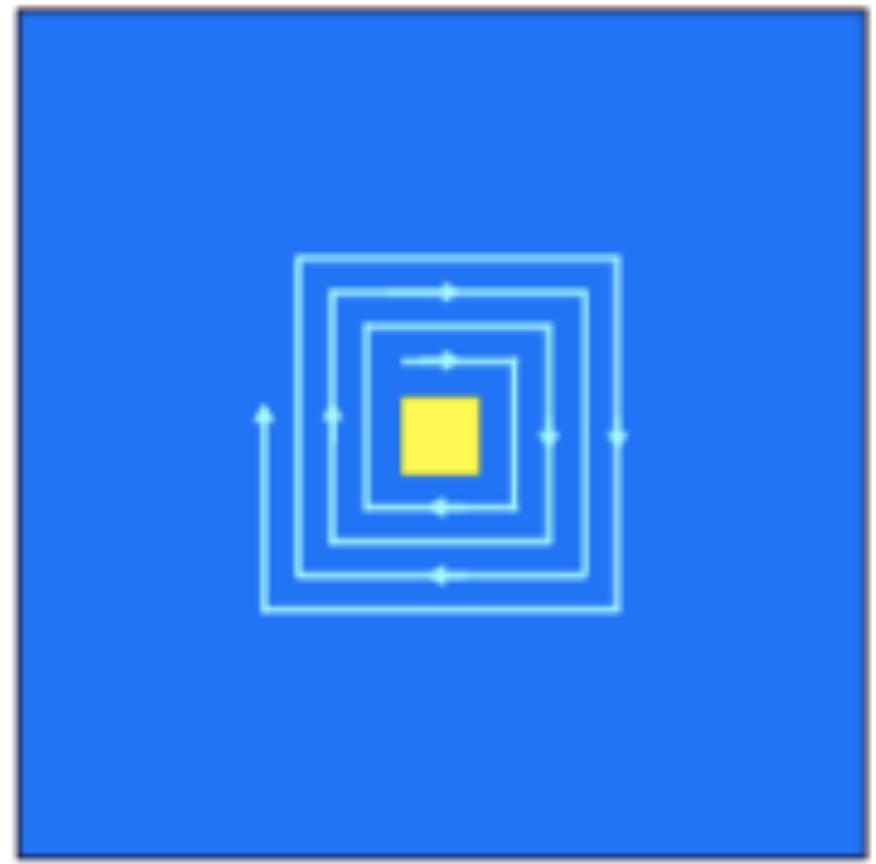


SEARCH: ELF GO

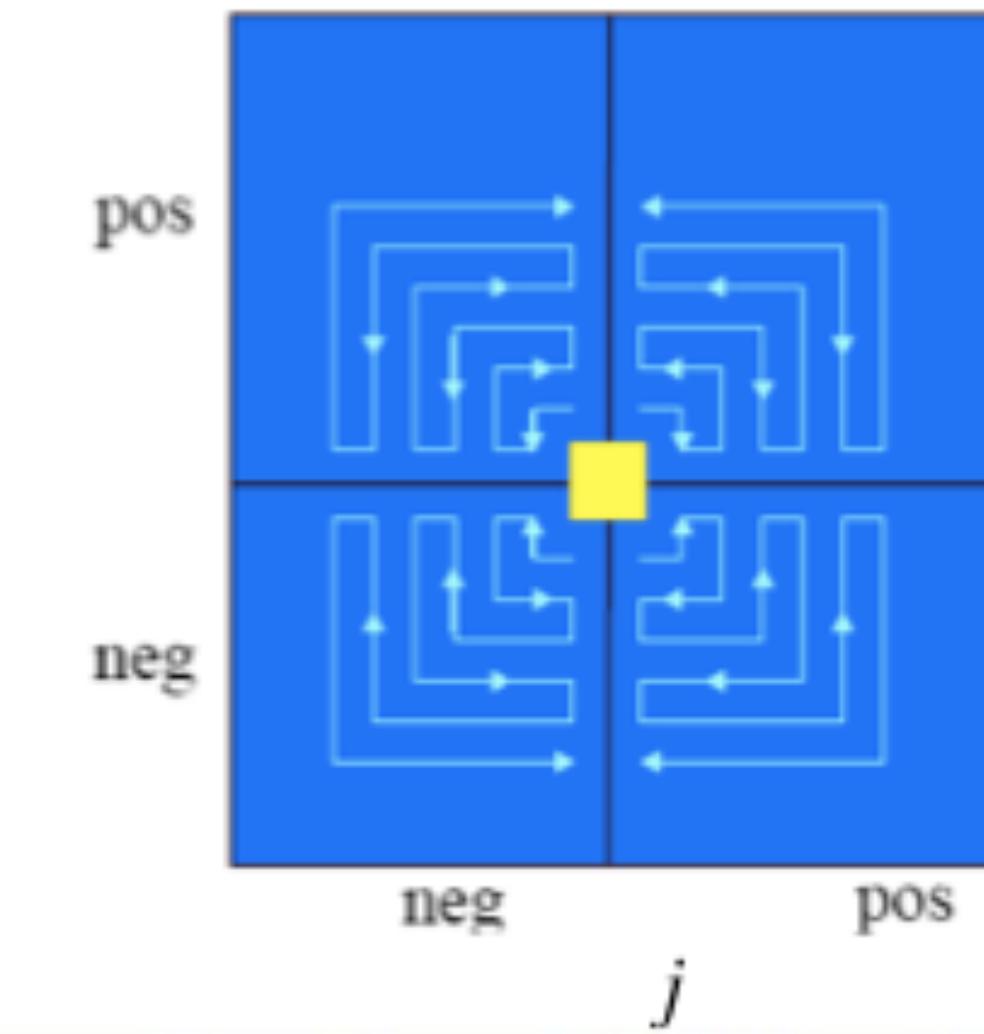
VIEW: BROWSER RNA-SEQ
DHS TSS
REPTIMING

EXEMPLO DE ABORDAGEM DEPENDENTE DE CONSULTA

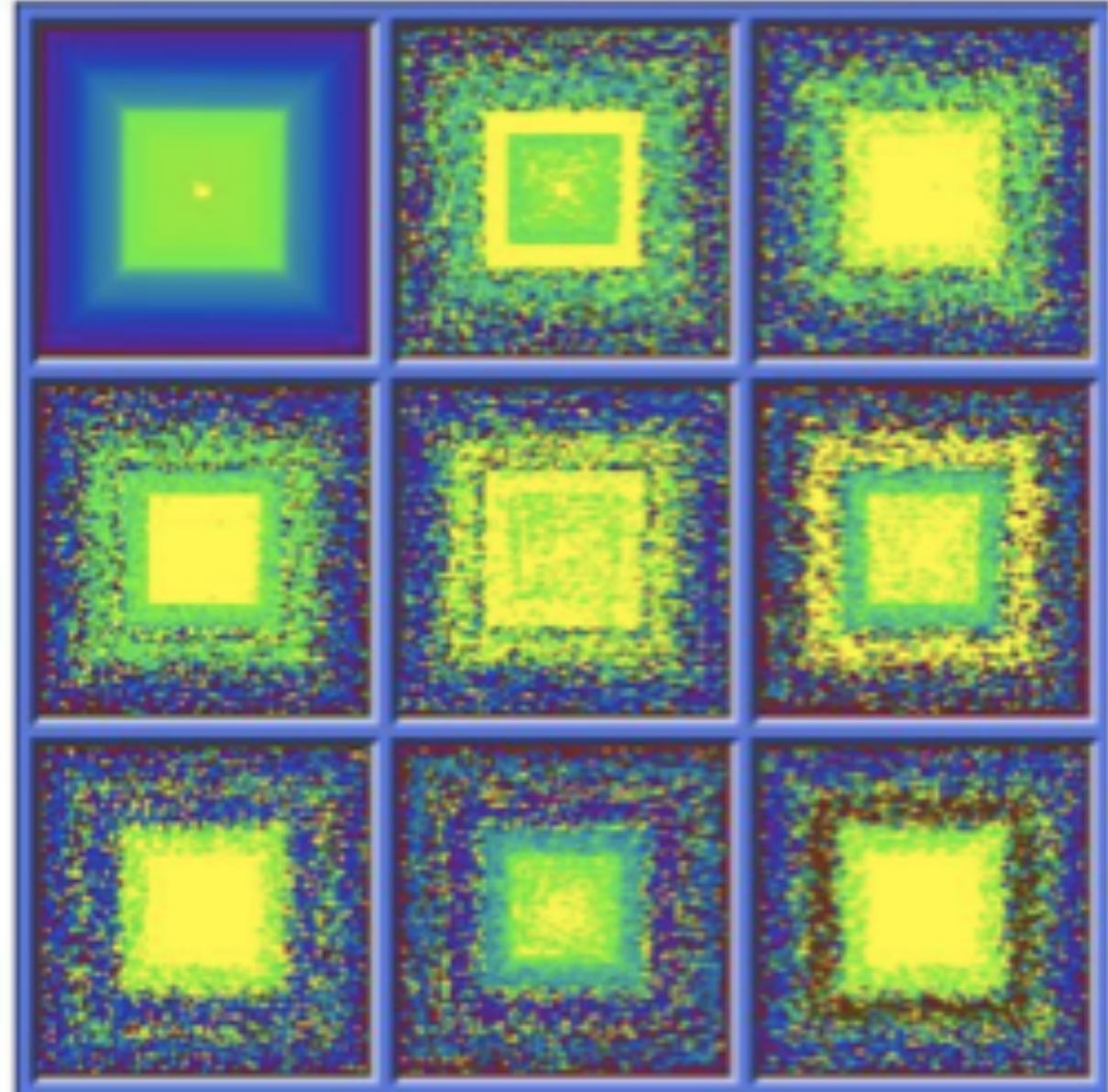
(a)



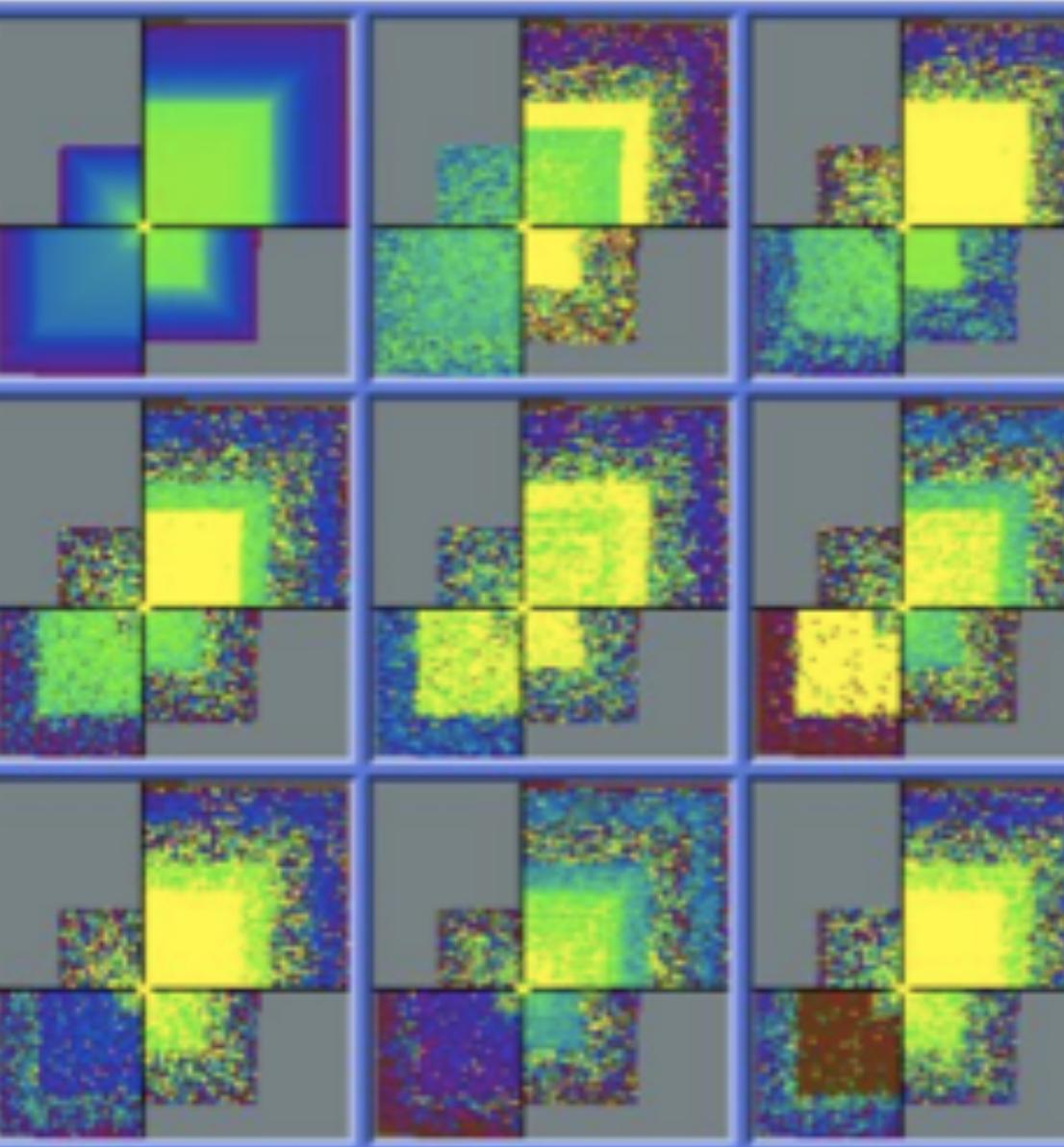
(b)



(c)



(d)



ESPIRAL

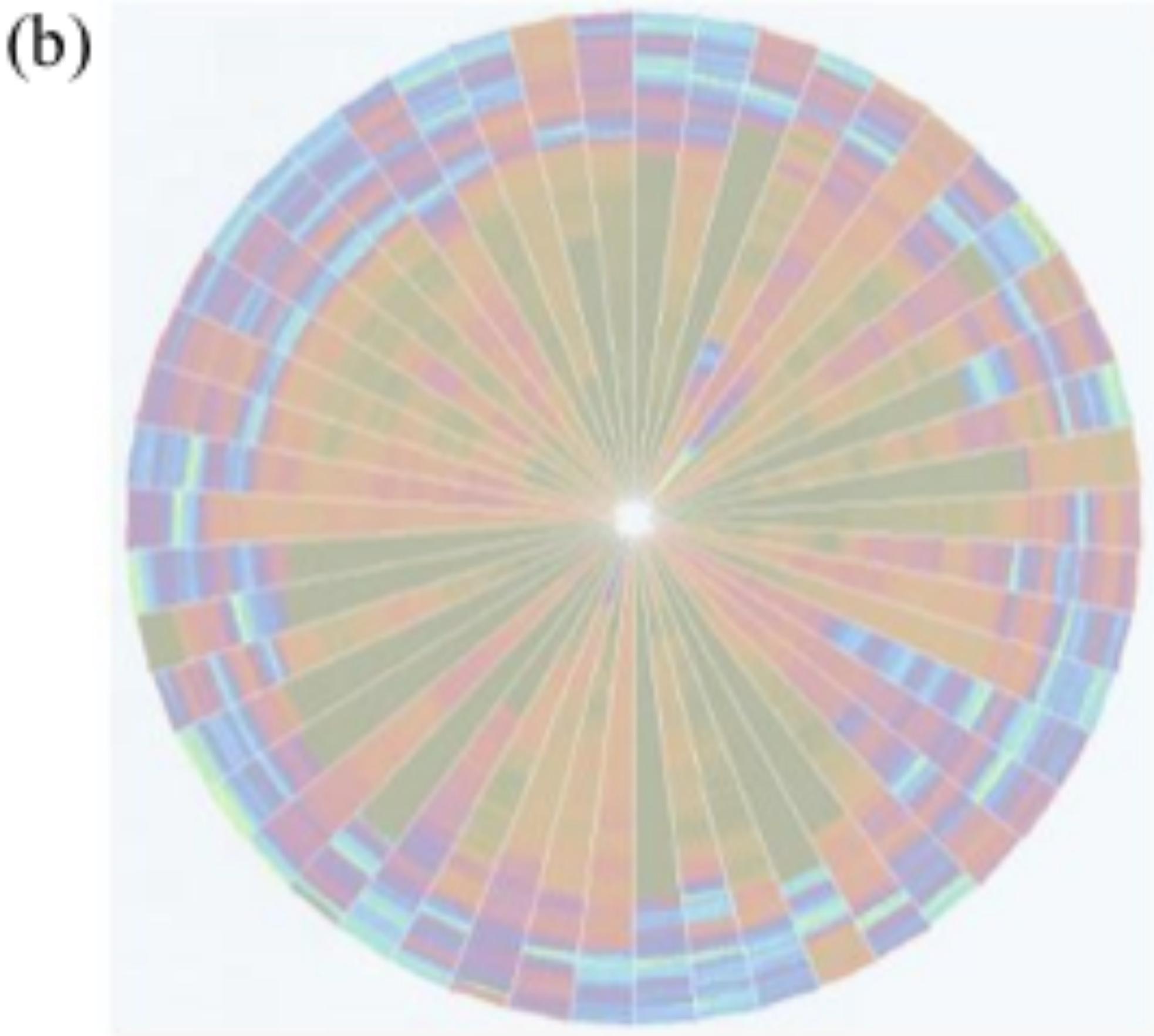
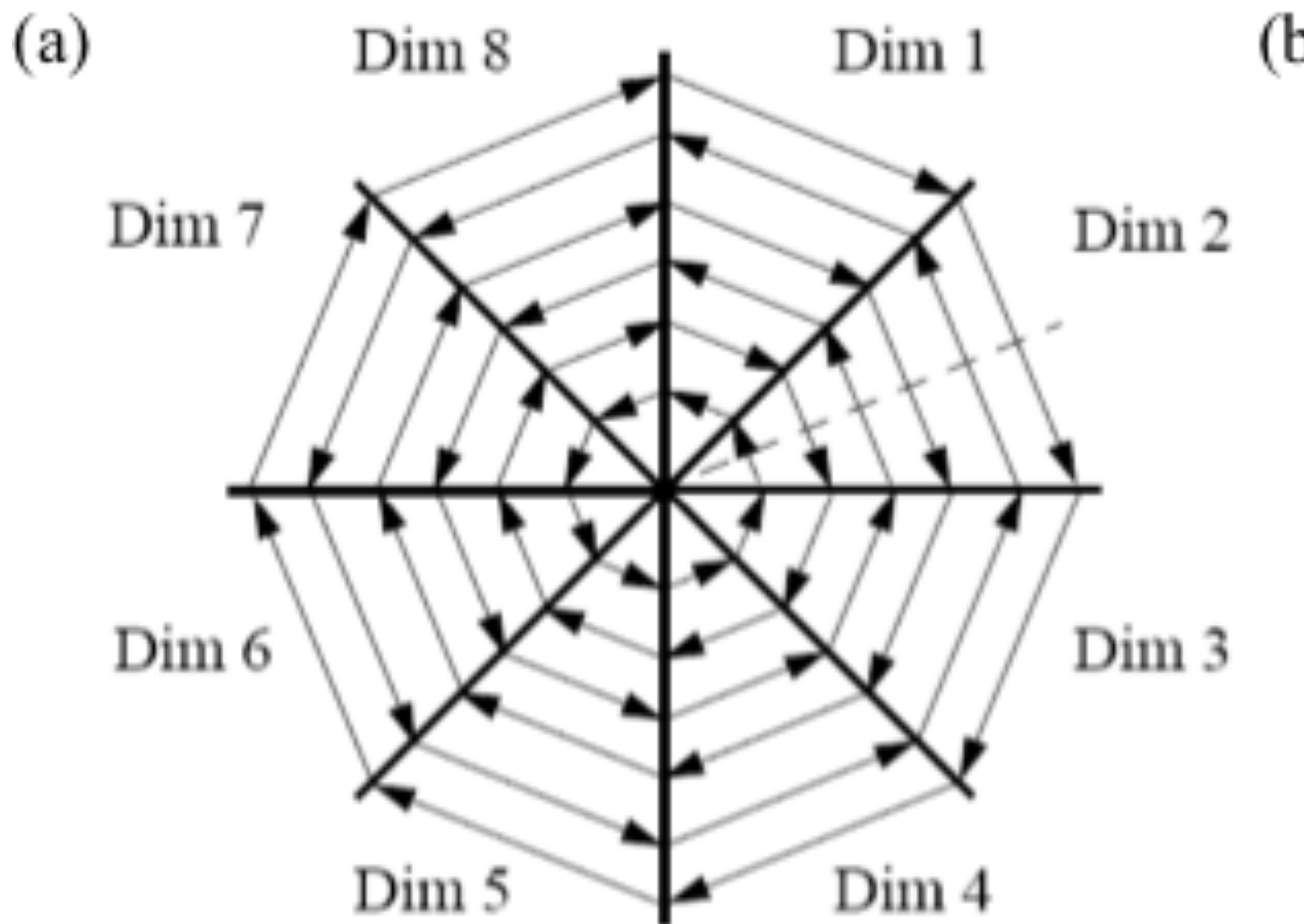


Figure 3.16: (a) Circle segment arrangement for 8-dimensional data [15],
(b) An example of circle segments [7].

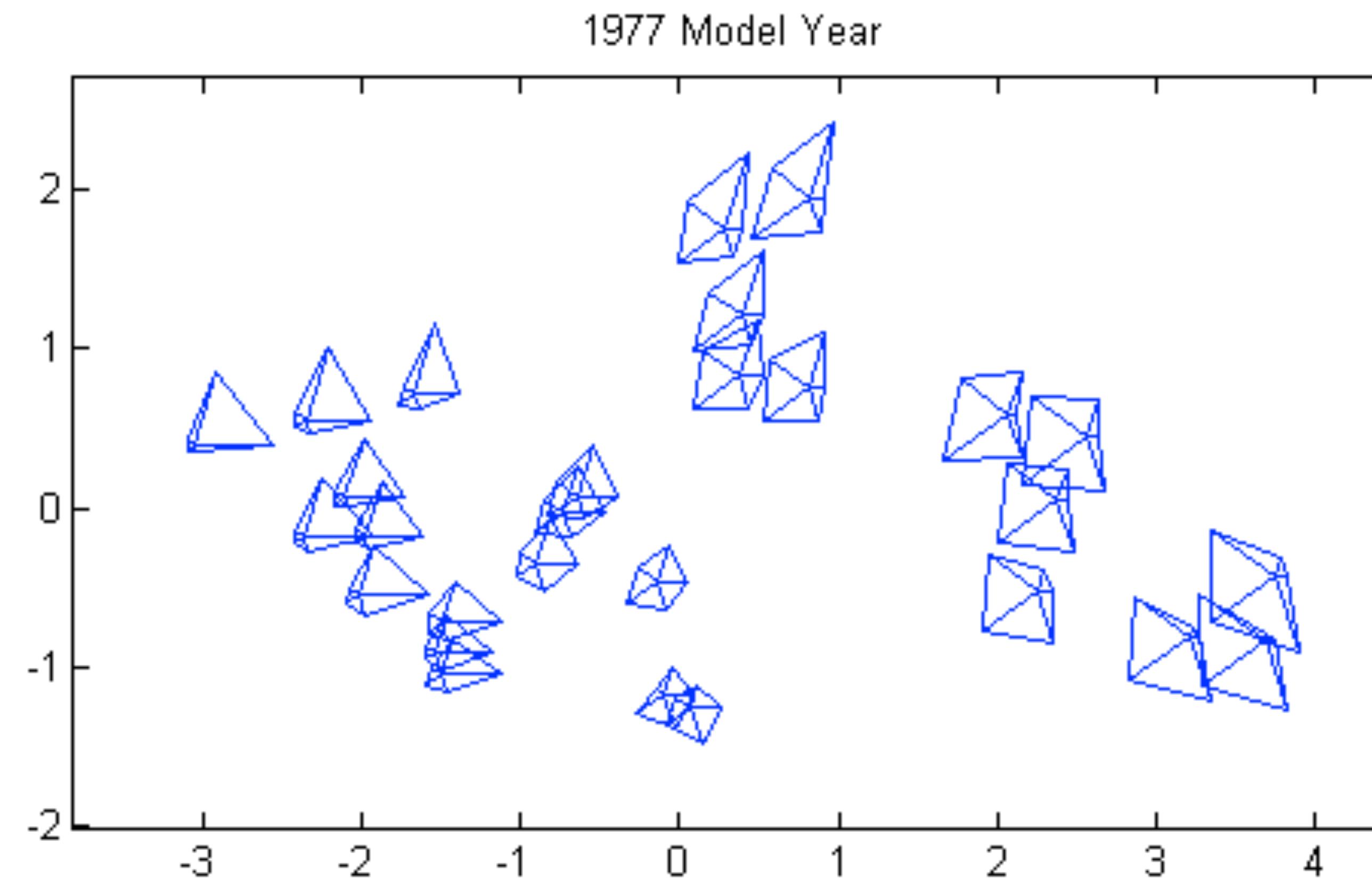
BOAS PRÁTICAS

.....

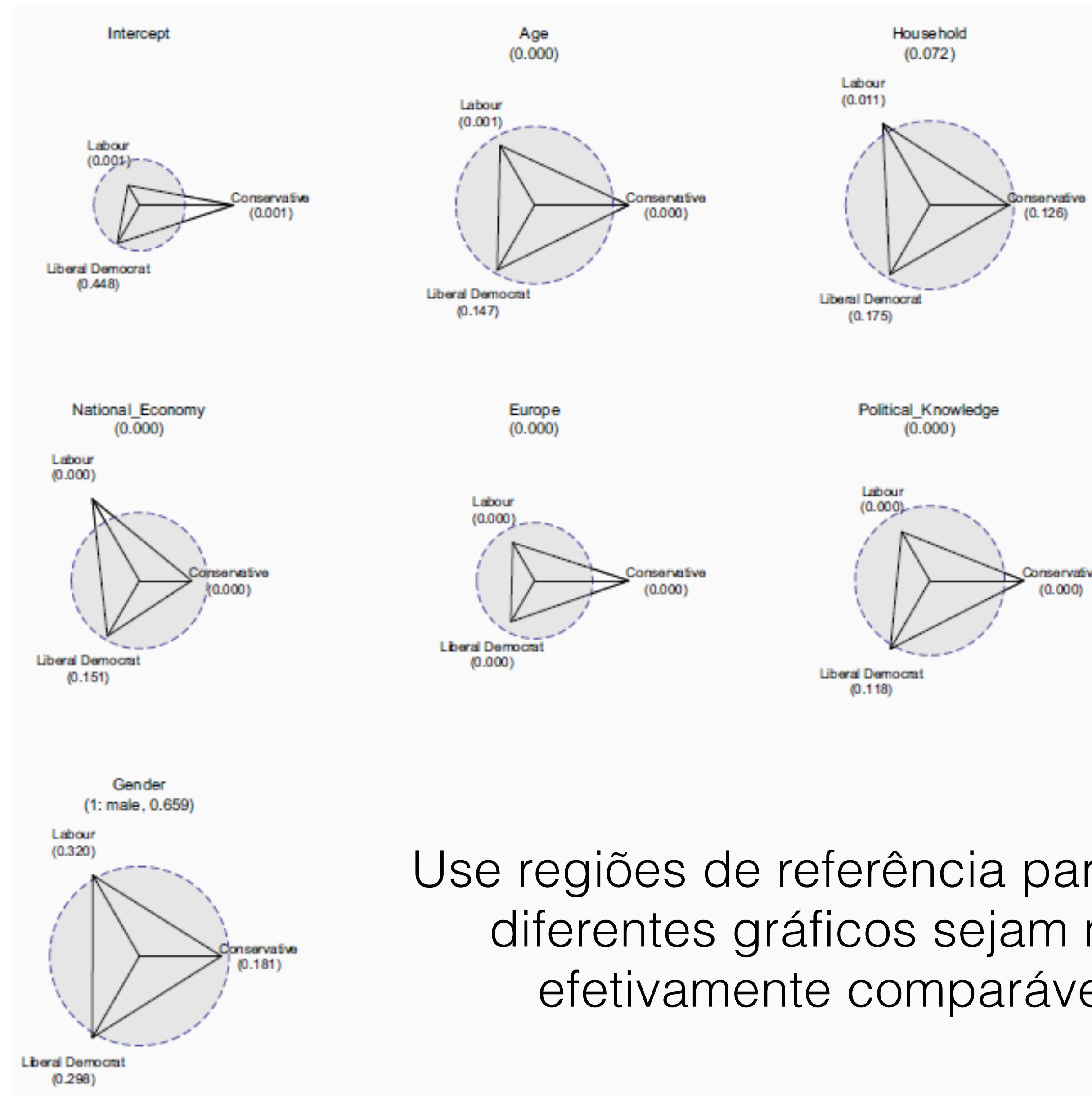
O uso de técnicas de ordenação e agrupamento dos itens por similaridade pode auxiliar na descoberta de padrões interessantes nos dados

Moral Matrix





Técnicas de redução de dimensionalidade como MDS, SVD ou PCA devem ser usadas para posicionar os objetos em espaço 2D preferencialmente



Use regiões de referência para que os diferentes gráficos sejam mais efetivamente comparáveis