

# Lesson 2: Designing Graphs to Enlighten

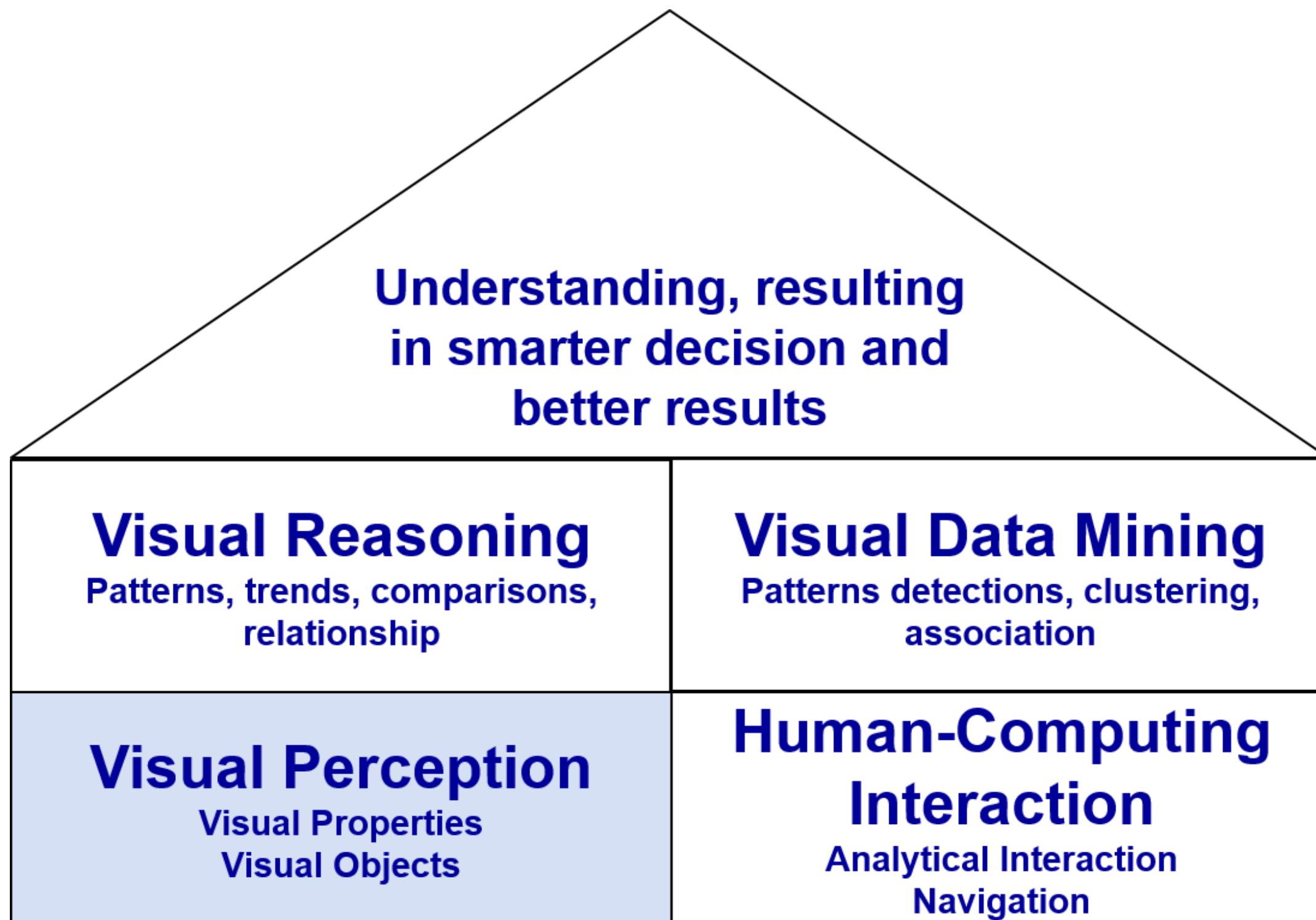
Dr. Kam Tin Seong  
Assoc. Professor of Information Systems (Practice)  
School of Computing and Information Systems,  
Singapore Management University

23 Jan 2023

# What will you learn from this lesson?

- Data visualisation design process
- The Devil is in the data
- Human perception and information processing
- Components of a graph
- Principles and practical guides for data visualisation design
- Data visualisation critics framework

# Building Block of Visual Analytics



# DataVis design process

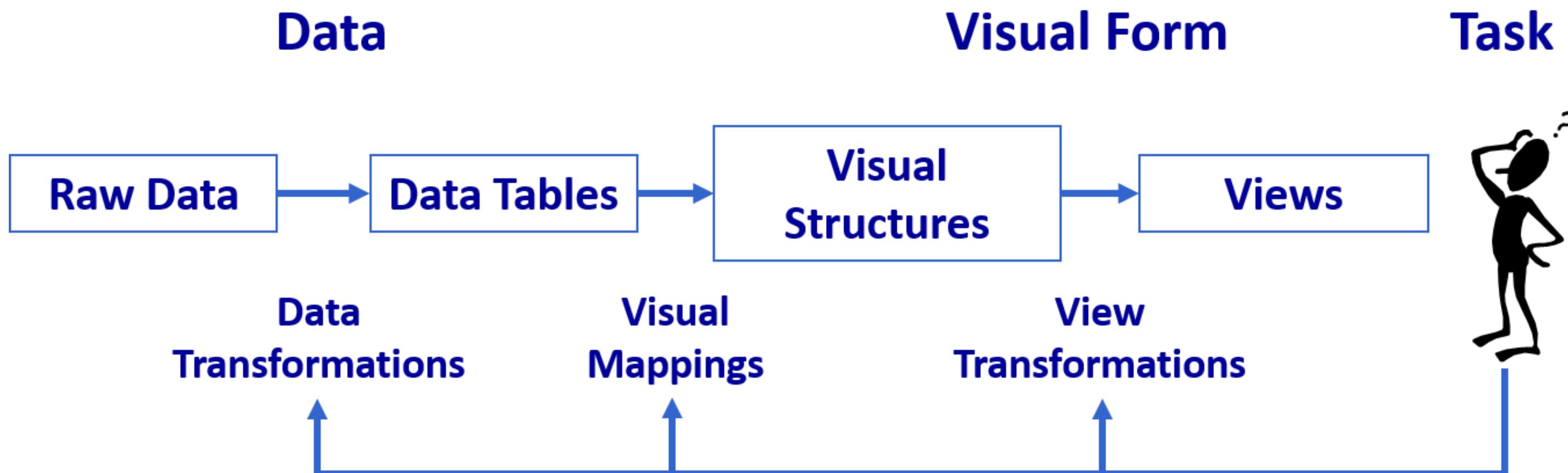
- Start with a clear message
- Select the right data
- Use the right visualisations
- Express and explain
- Review and seek feedback from experts and casual readers

# DataViz design process: Start with a clear message

**Data Management**  
**Select data source**  
**Clean data**  
**Categorise data**  
**Moderate data**

**Visualisation**  
**Information design**  
**Visual encoding**  
**Interface design**

**Visual Analytics**  
**Observations**  
**Hypothesis**  
**Evidence (+/-)**  
**Summarise**  
**Communicate**

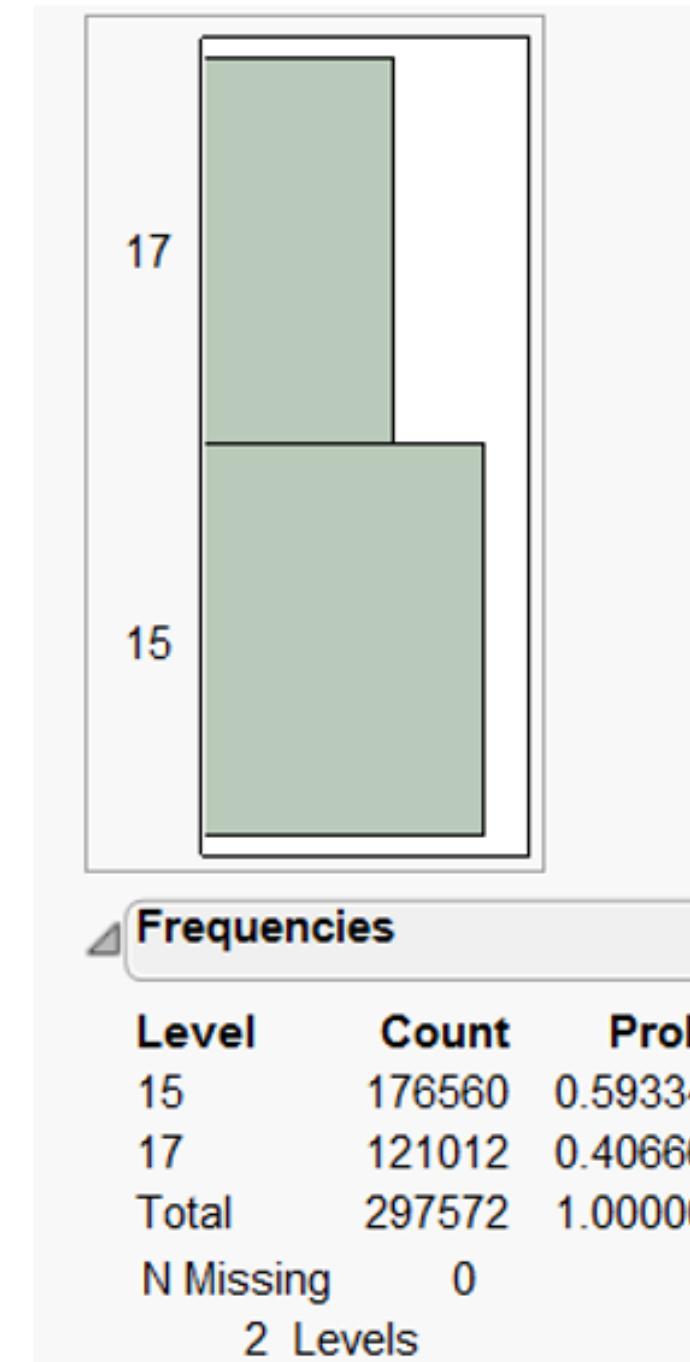
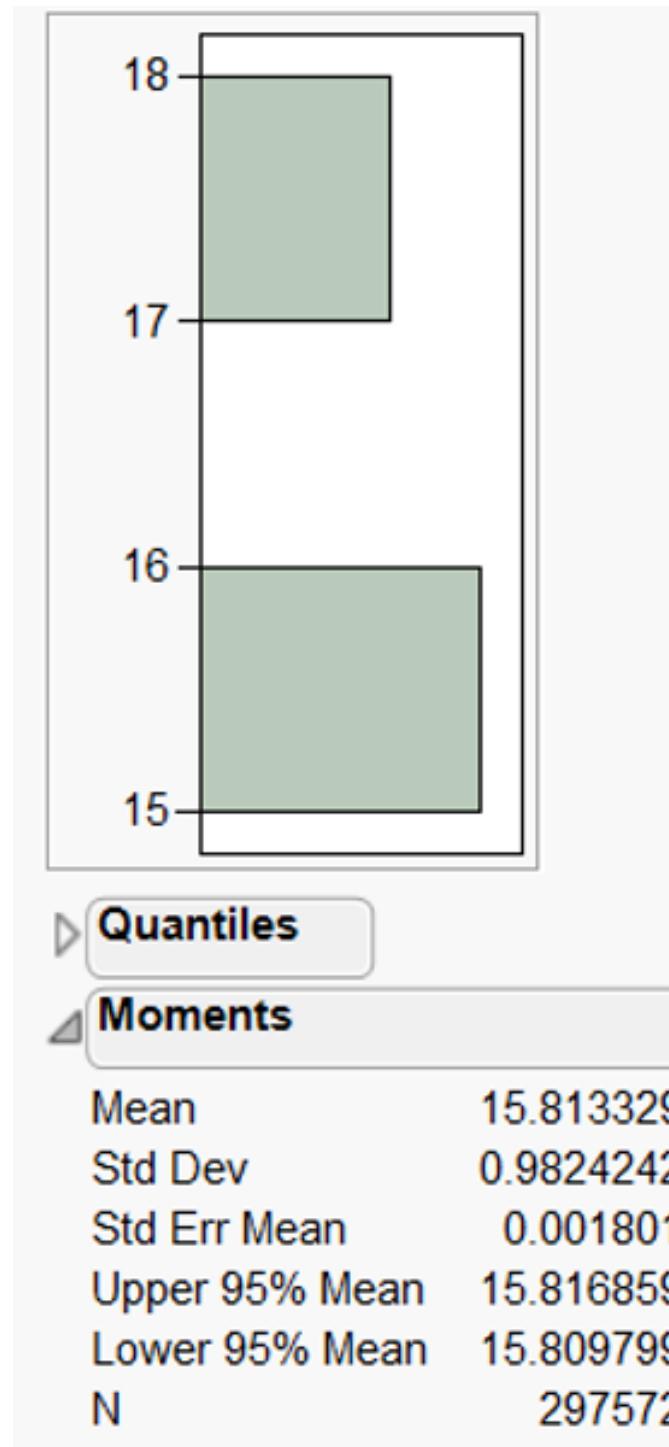


# The Devil is in the Data



# Numbers Worth Knowing

Not all numerical data are continuous!



# Visualising the Right Data

- There are more than one way to present the data.

## FIVE DATA VARIATIONS TO CONSIDER

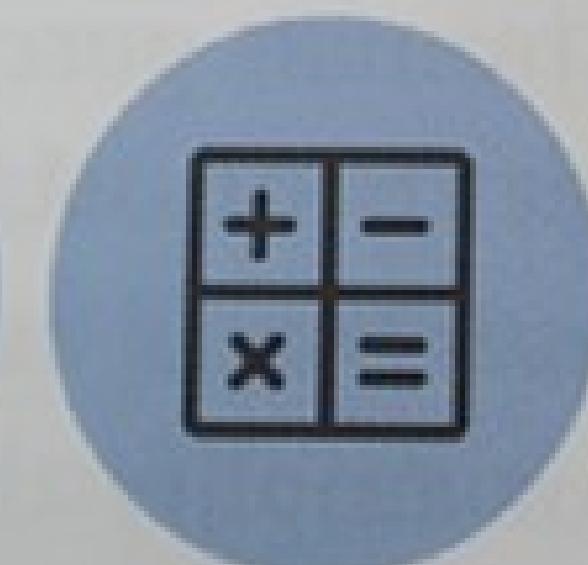
Totals



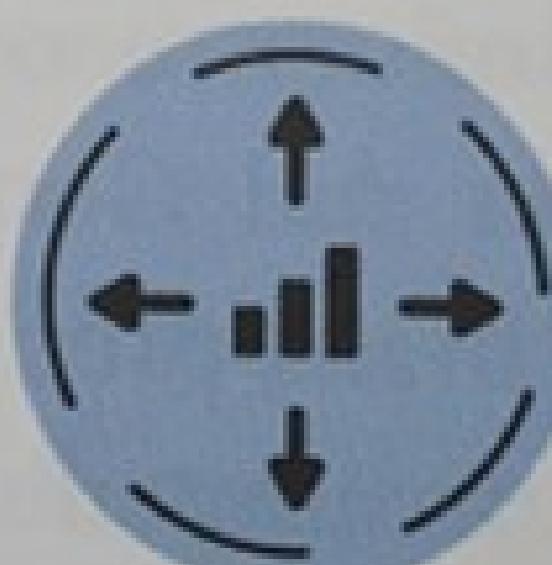
% Change



Calculated Metric (Ratio)



Added Context

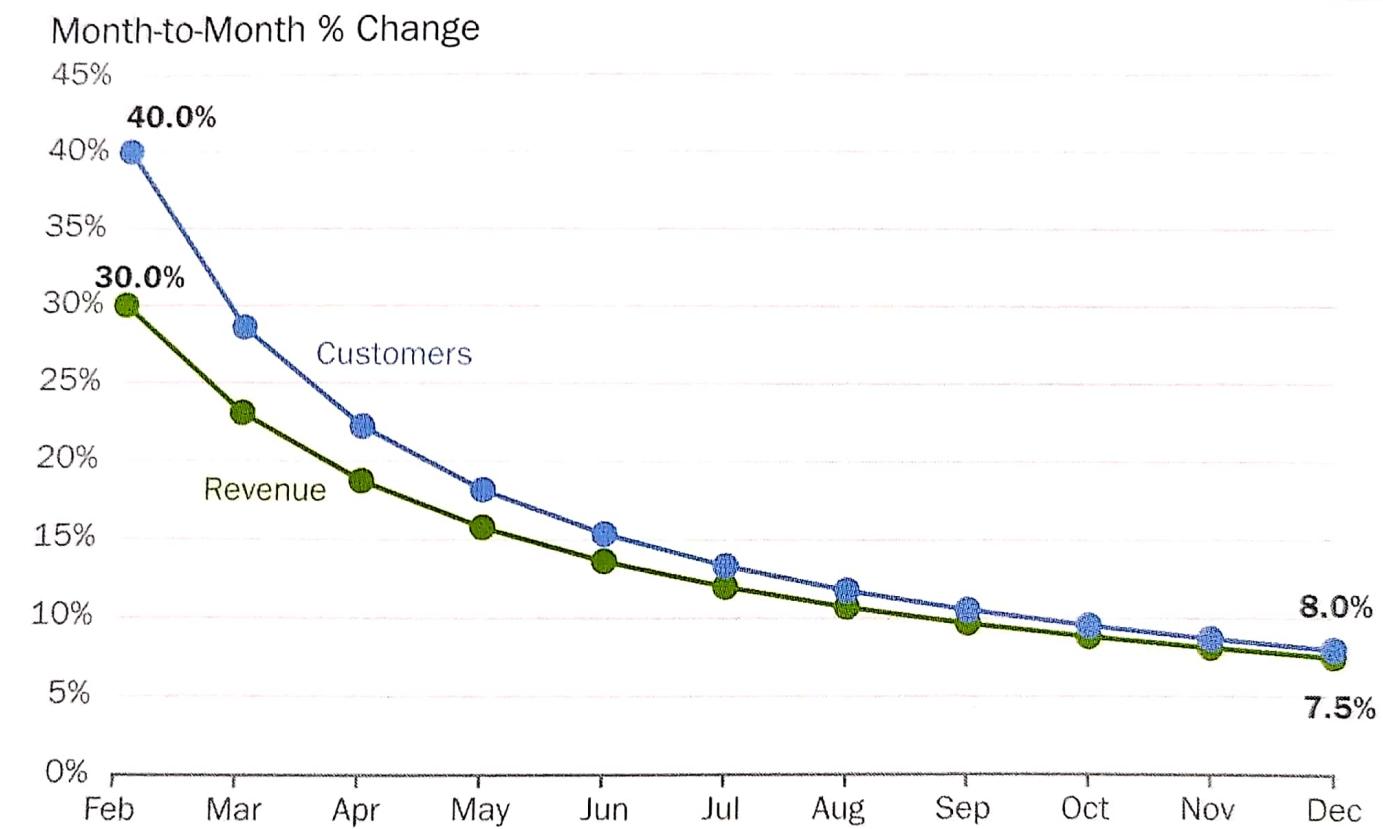
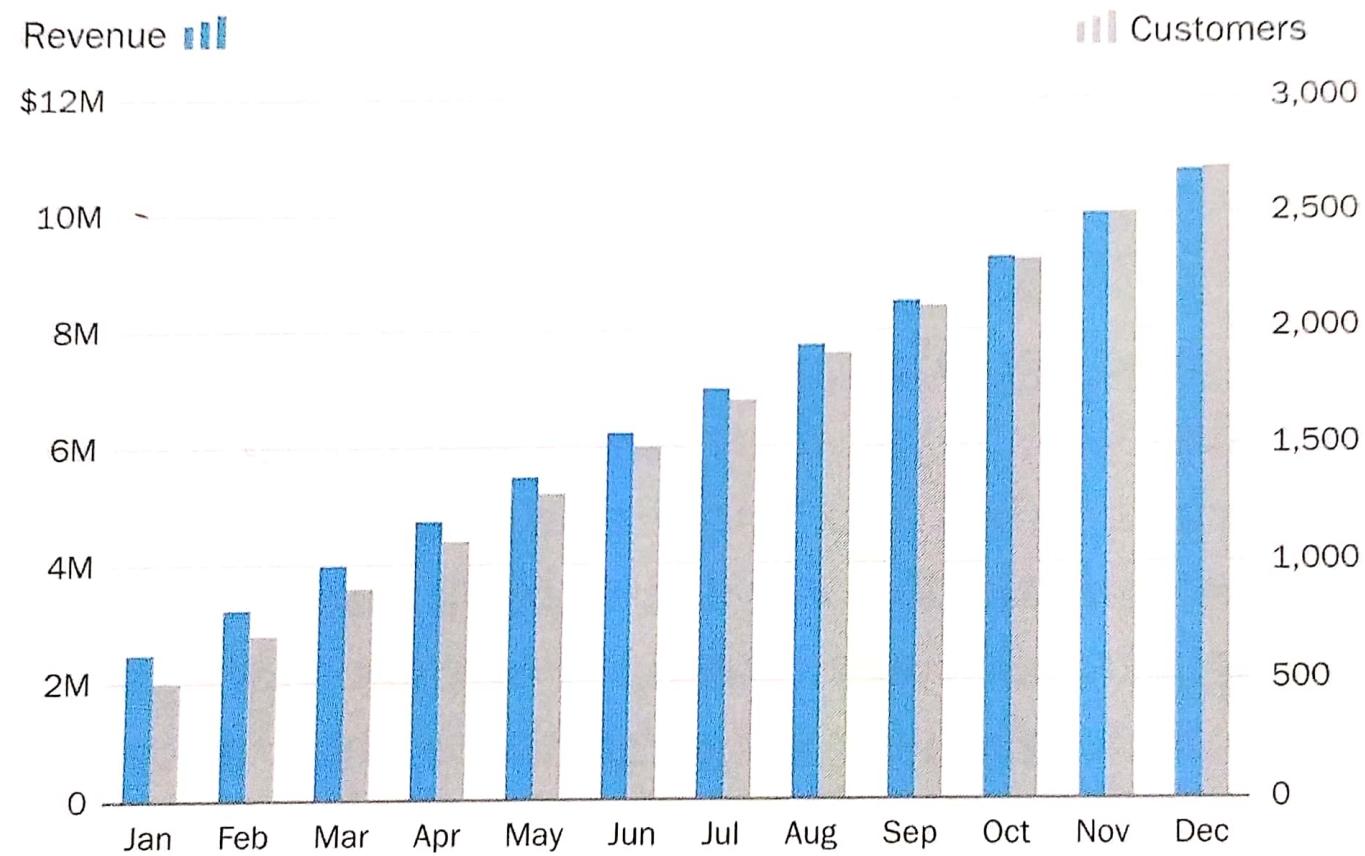


Variance



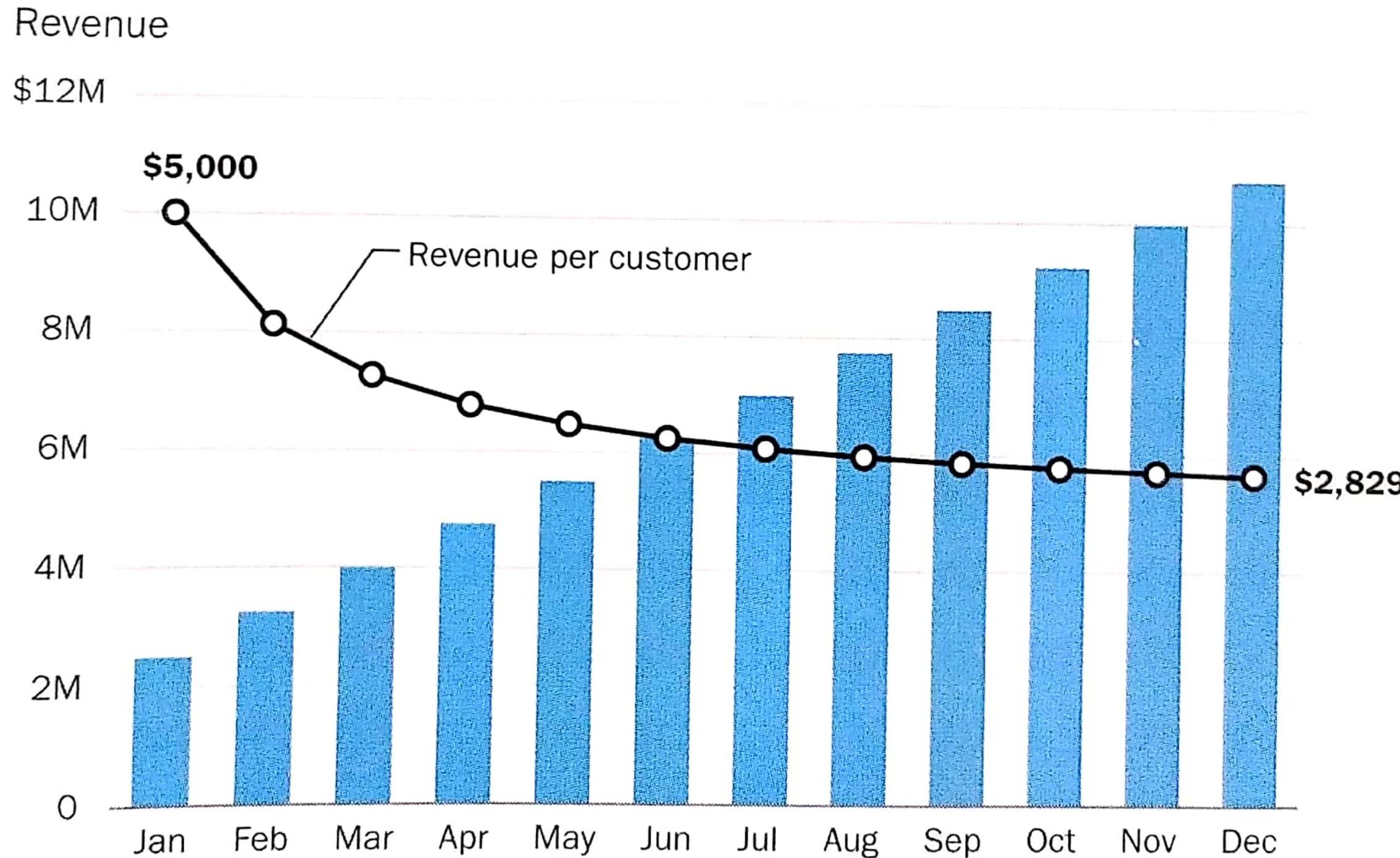
# Visualising the Right Data

- Derived values reveal more interesting patterns than absolute values.

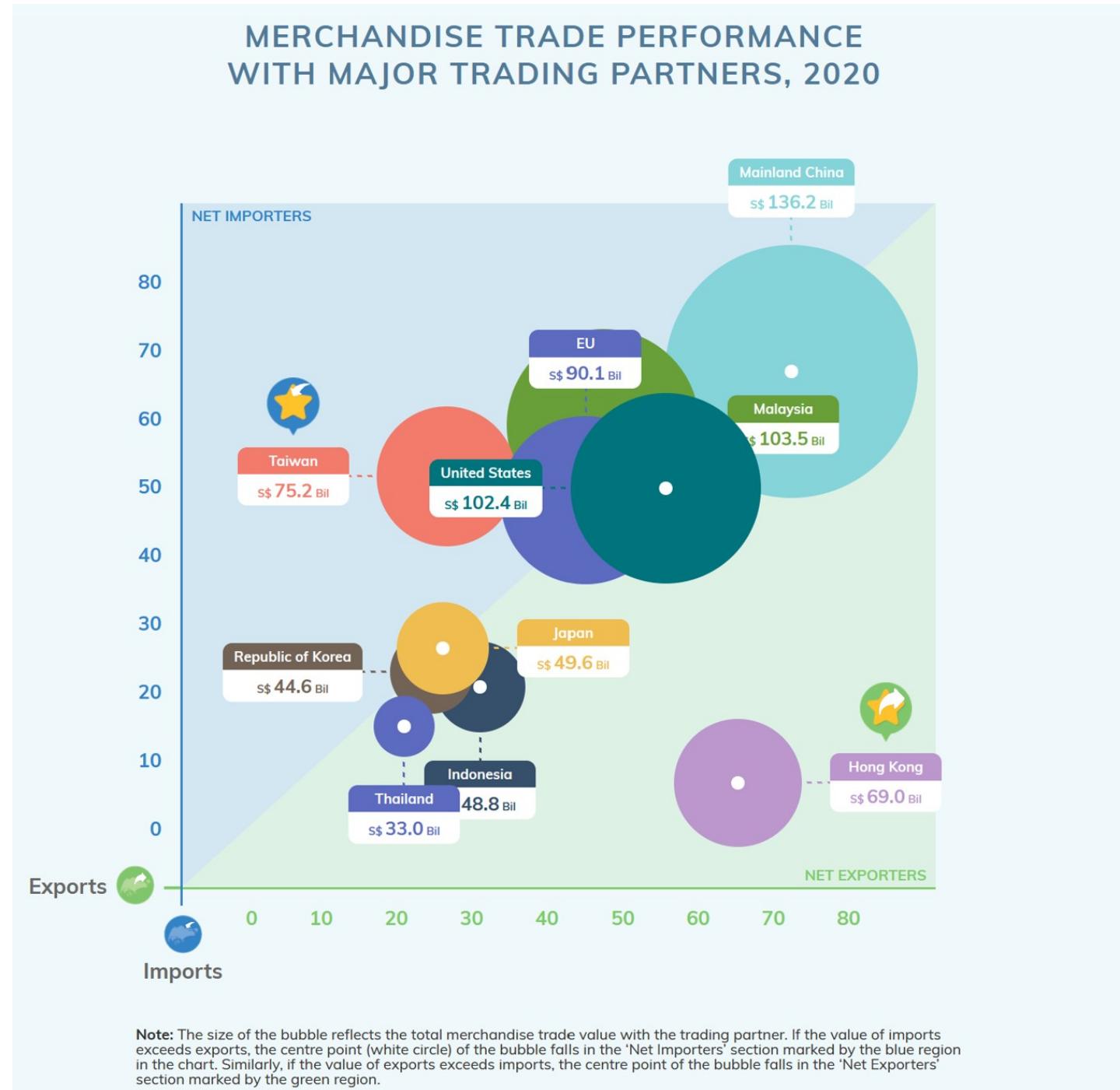


# Visualising the Right Data

- Derived values reveal more interesting patterns than absolute values.



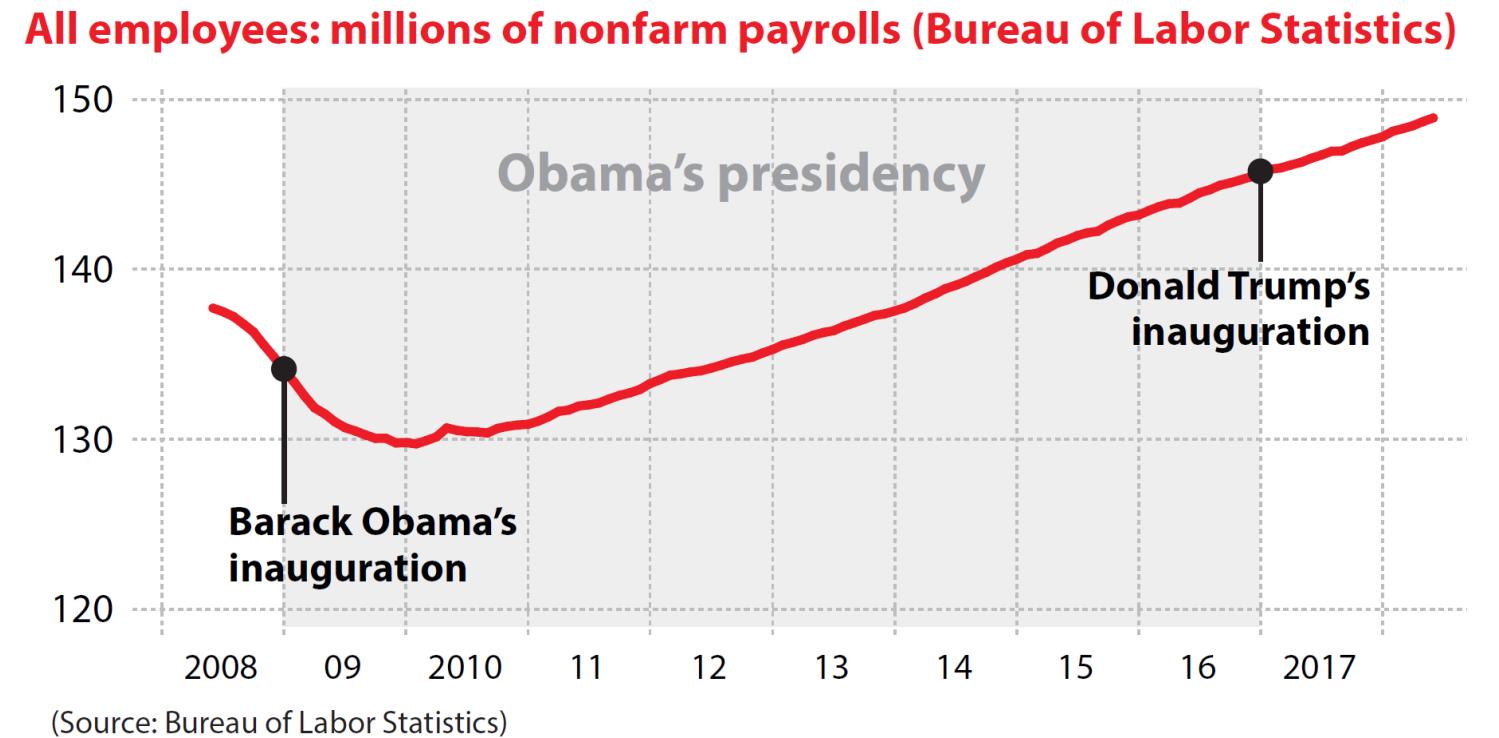
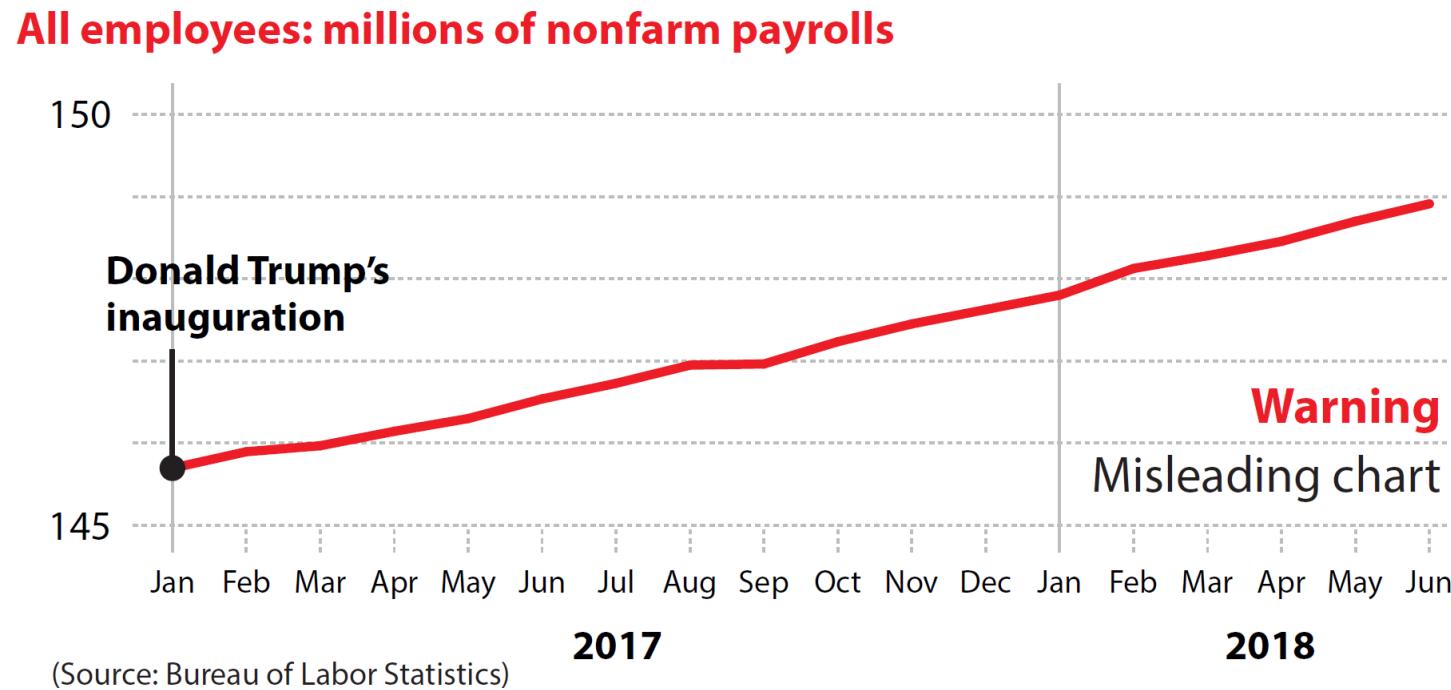
# Graphical Integrity: Show Me the Truth



Source: Department of Statistics

# Graphical Integrity: Show Me the Truth

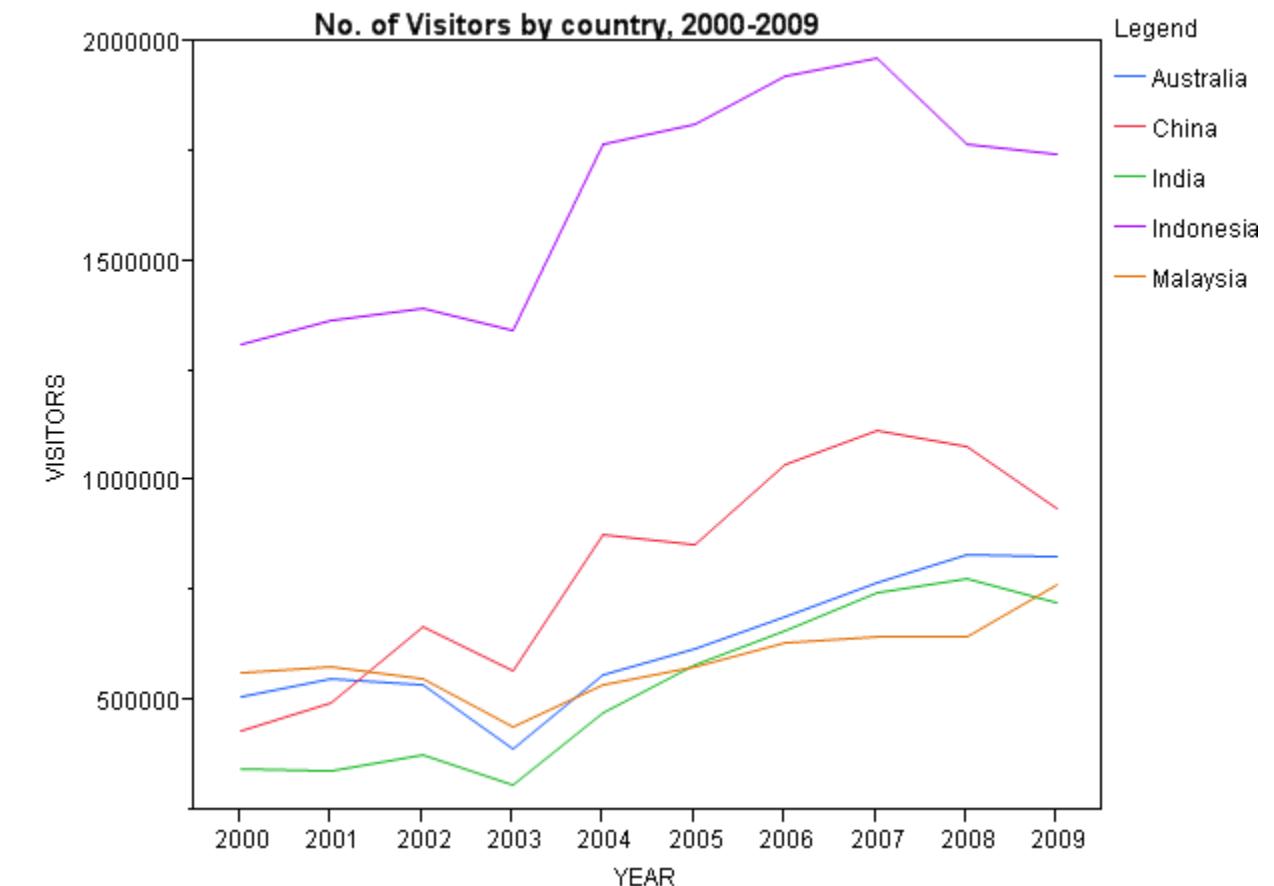
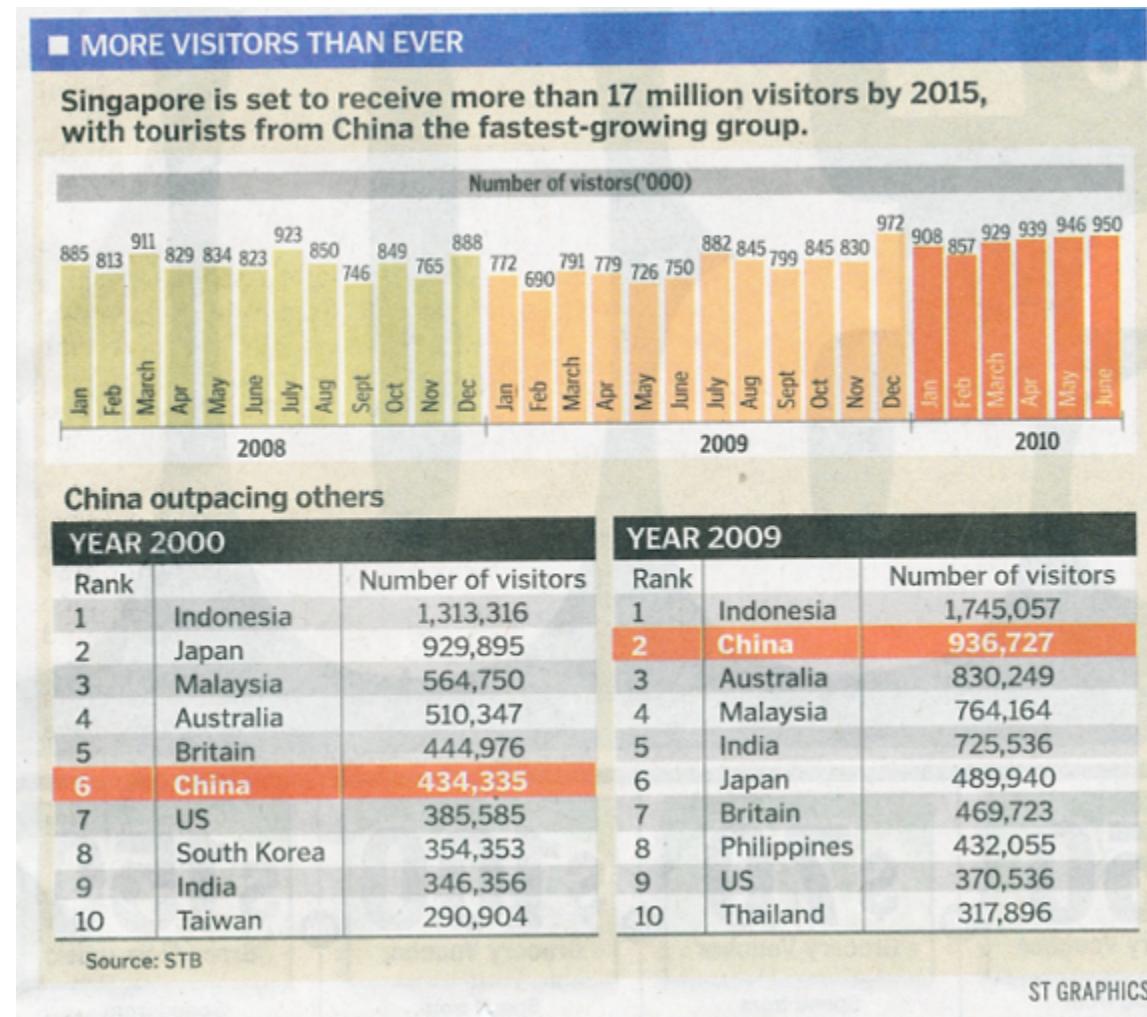
## Snapshot can be misleading!



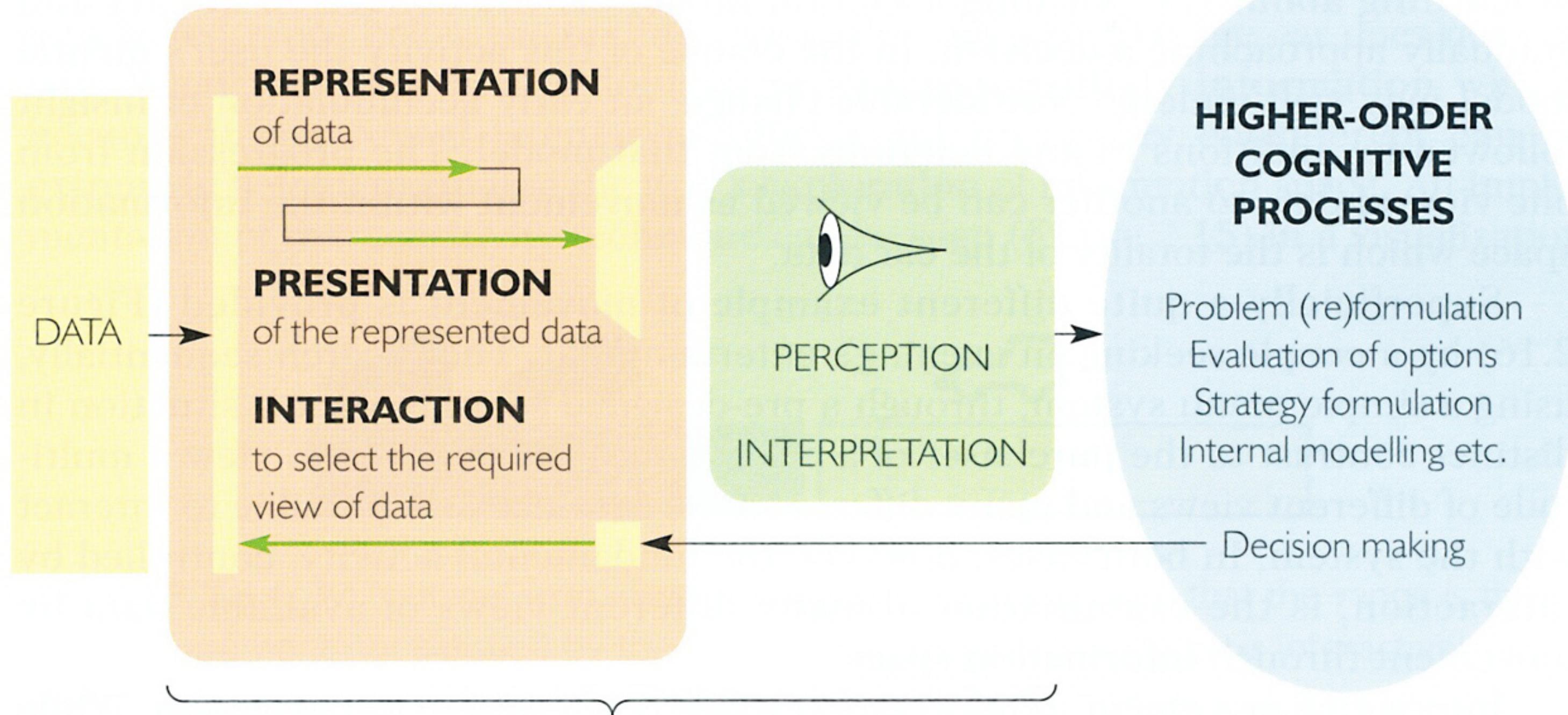
Source: Cairo, Alberto (2019) **How Charts Lie**, W.W. Norton & Company, USA. pg 168.

# Graphical Integrity: Show Me the Truth

Do not miss-out what had happened in between.



# Human Perception and Information Processing



# Pre-attentive Processing

- A limited set of visual properties are processed preattentively (without need for focusing attention).
- This is important for design of visualizations
  - What can be perceived immediately?
  - Which properties are good discriminators?
  - What can mislead viewers?

# How Visual Sensing Works?

## *(i)* Fact 1:

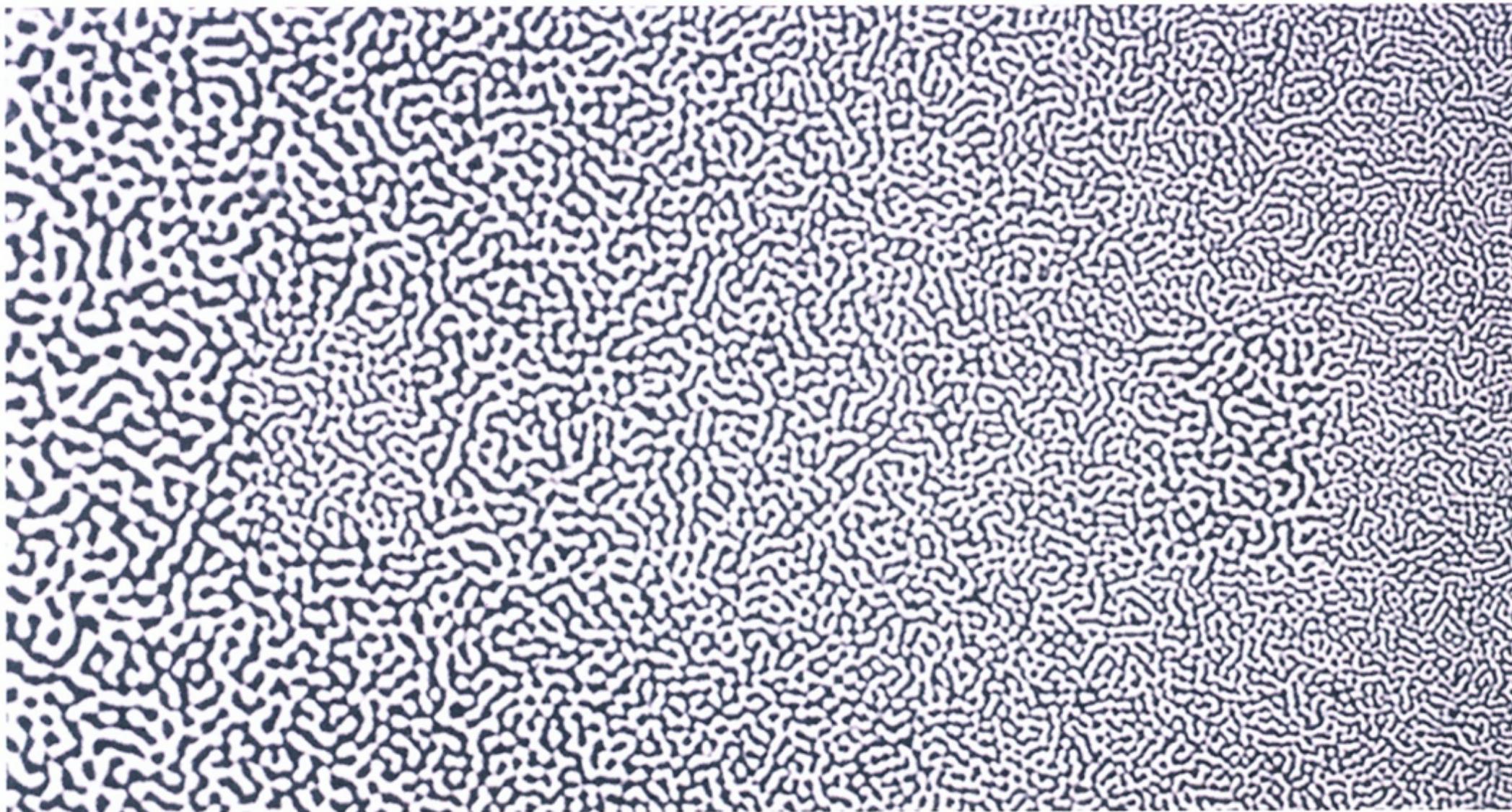
We see what we know and expect.



# How Visual Sensing Works?

## Fact 2:

We do not attend to everything we see.



# How Visual Sensing Works?

## *(i)* Fact 3:

We don't remember everything we see.



# How Visual Sensing Works?

## *(i)* Fact 3:

We don't remember everything we see.



# How Visual Sensing Works?

How Many 3's?

**1281768756138976546984506985604982826762  
9809858458224509856458945098450980943585  
90910302099059595772564675050678904567  
8845789809821677654876364908560912949686**

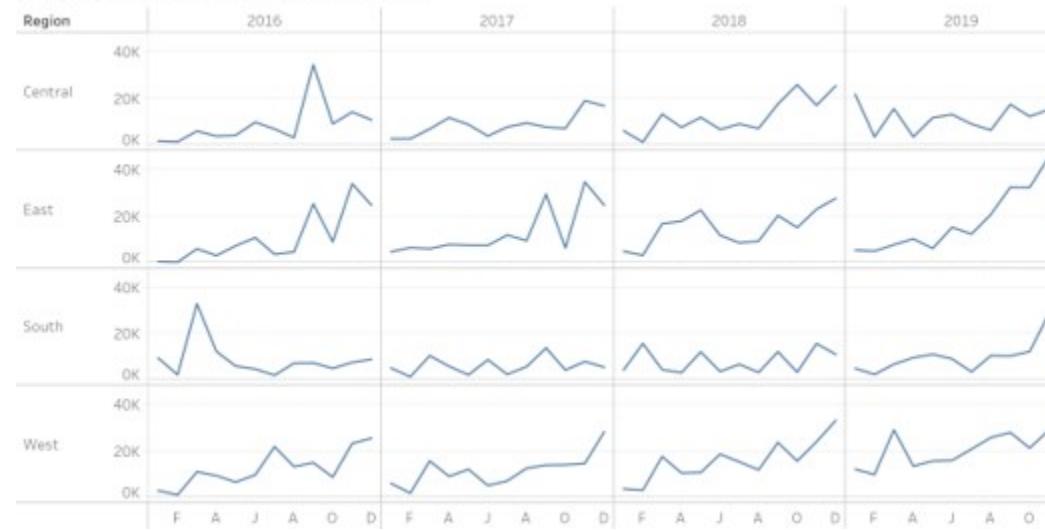
# How Visual Sensing Works?

Now you see them!

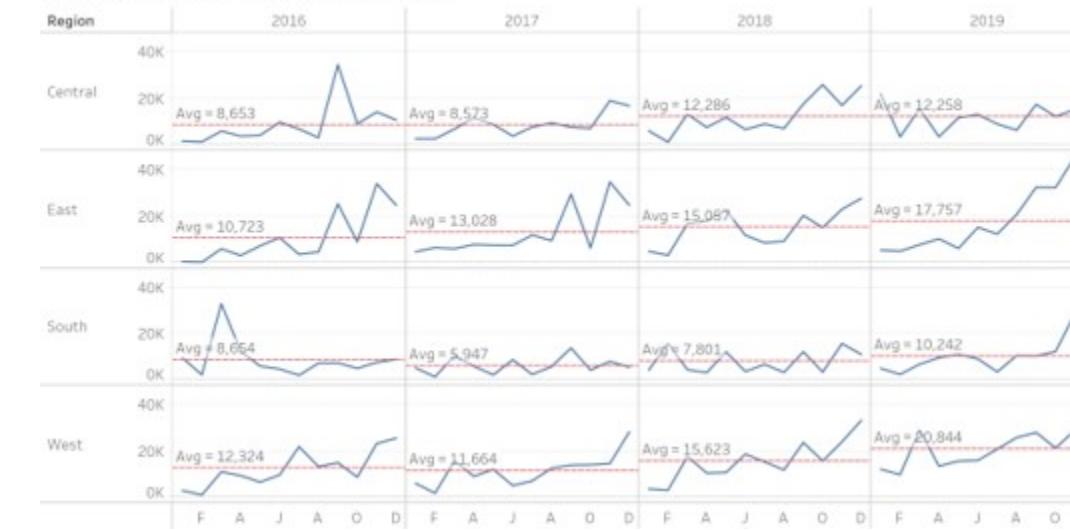
12817687561**3**8976546984506985604982826762  
980985845822450985645894509845098094**3**585  
90910**3**02099059595772564675050678904567  
8845789809821677654876**3**64908560912949686

# Application of pre-attentive principle in data visualisation design

Monthly sales by region, 2016-2019

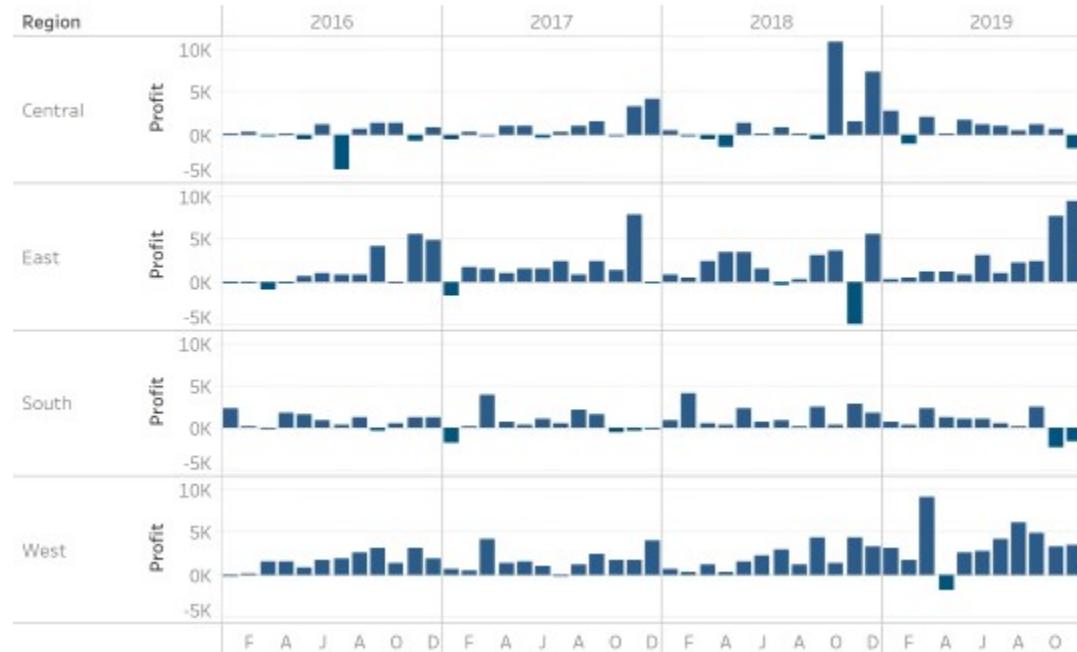


Monthly sales by region, 2016-2019

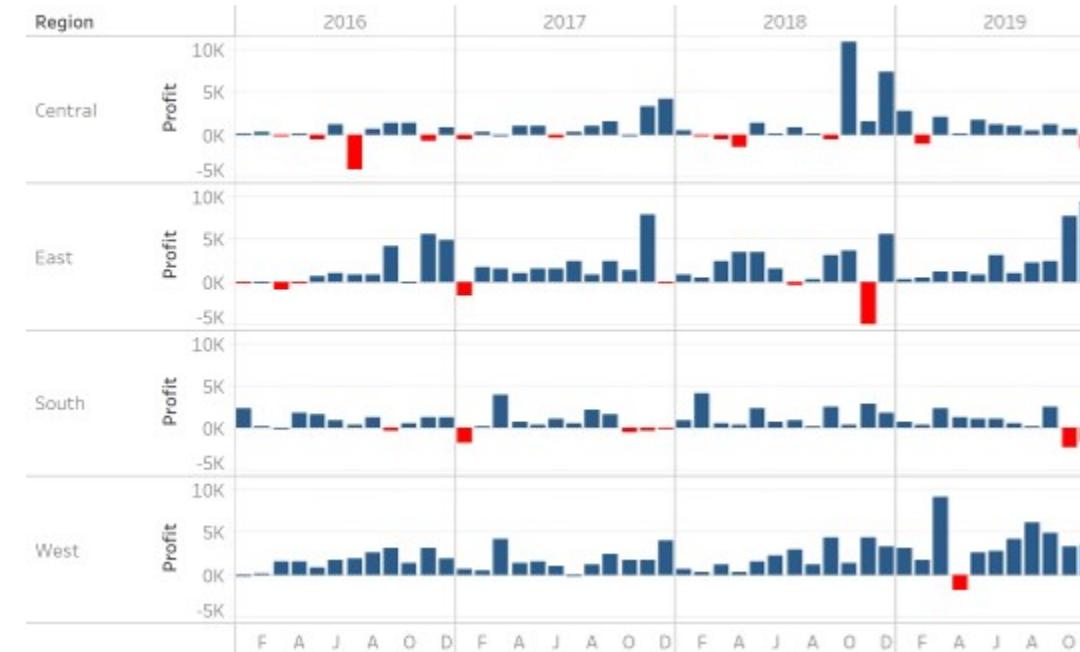


# Application of pre-attentive principle in data visualisation design

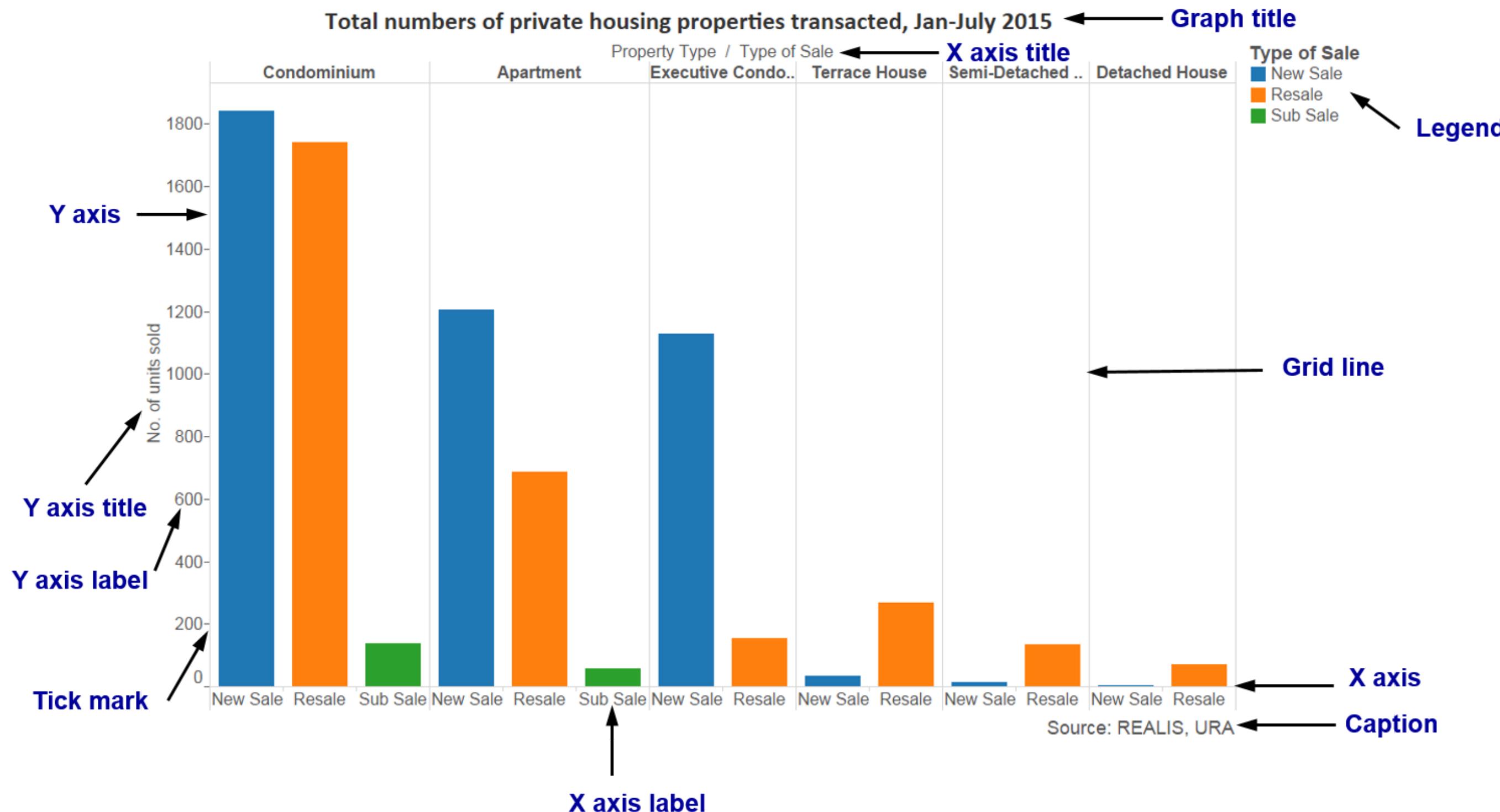
Monthly profit by Region, 2016-2019



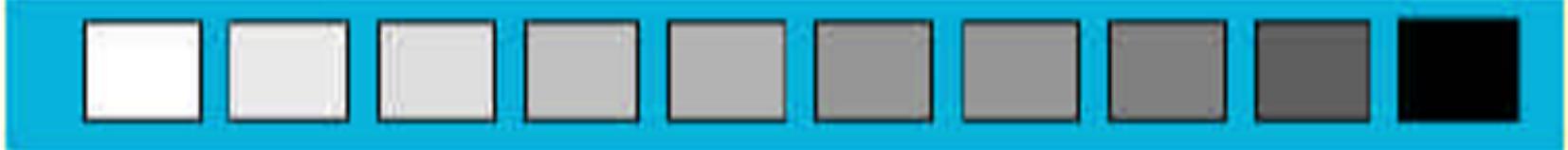
Monthly profit by Region, 2016-2019



# Components of a graph



# Bertin's Semiology of graphics

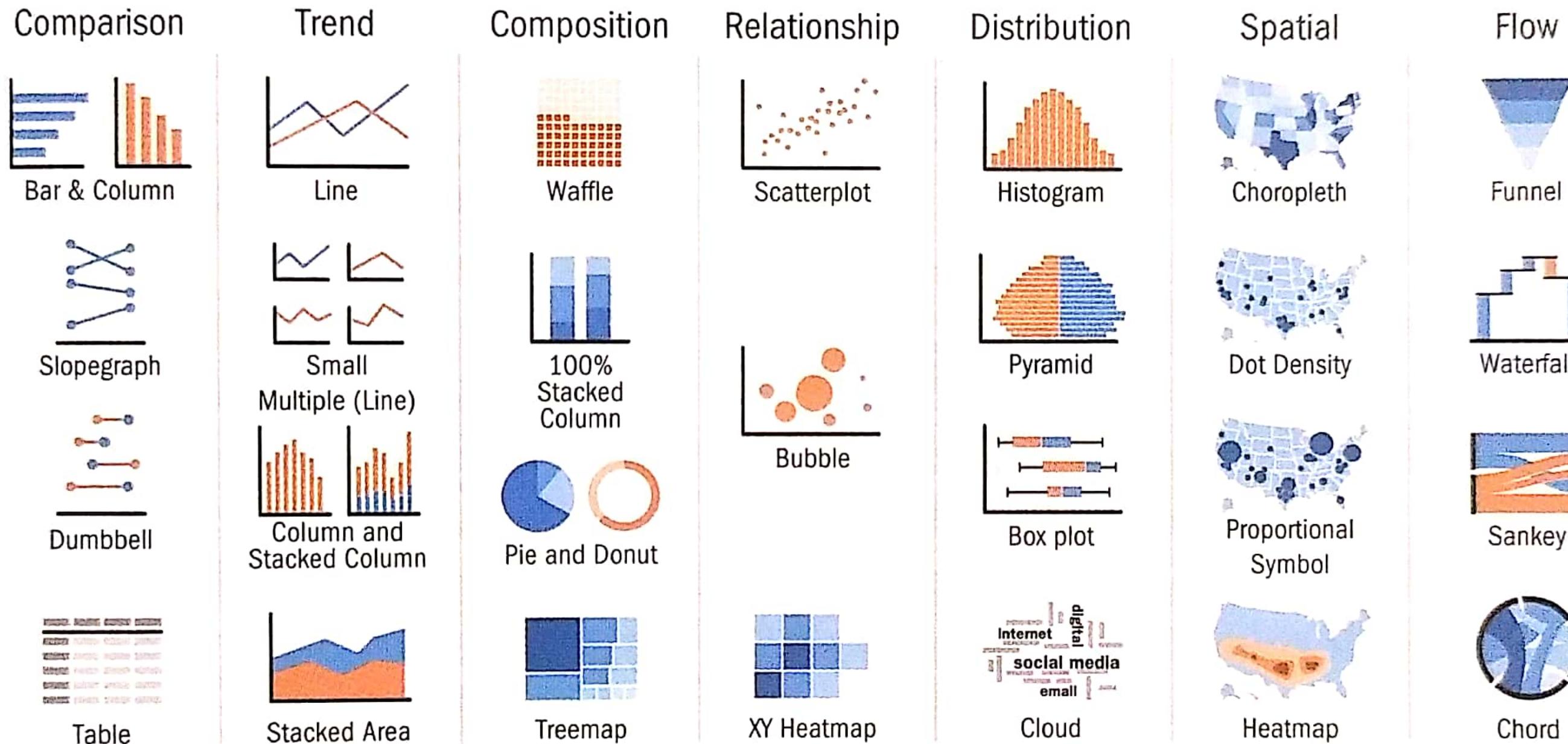
Bertin's Original Visual Variables	
<b>Position</b> changes in the x, y location	
<b>Size</b> change in length, area or repetition	
<b>Shape</b> infinite number of shapes	
<b>Value</b> changes from light to dark	
<b>Colour</b> changes in hue at a given value	
<b>Orientation</b> changes in alignment	
<b>Texture</b> variation in 'grain'	



Jacques Bertin

# Choosing the Right Visualisation

## MAJOR CHART TYPE CATEGORIES FOR BUSINESS PROFESSIONALS



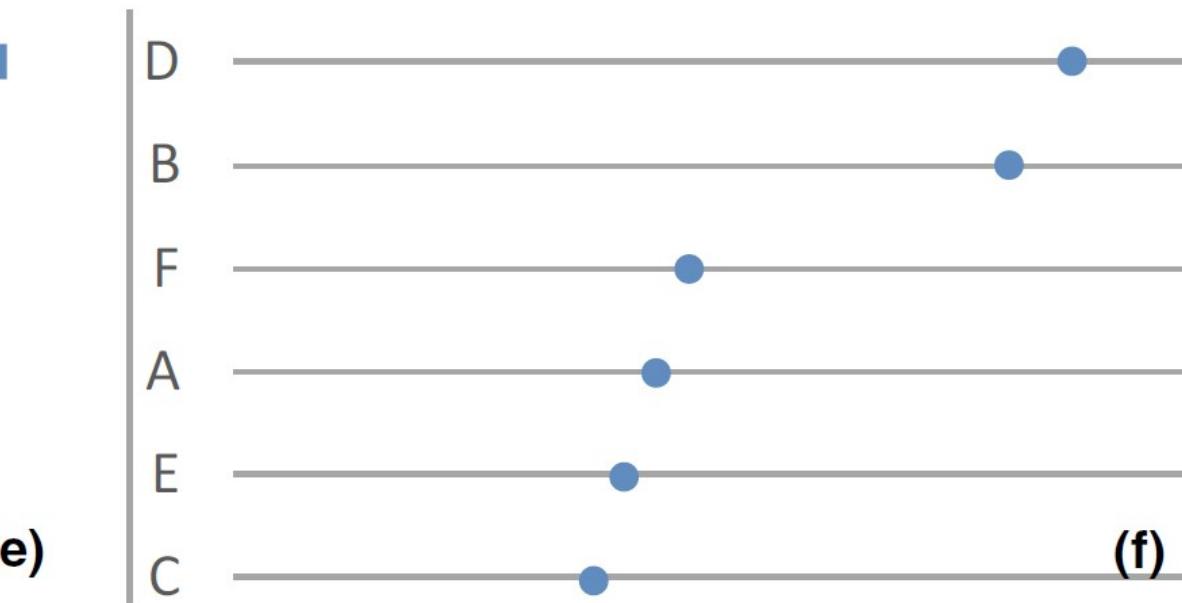
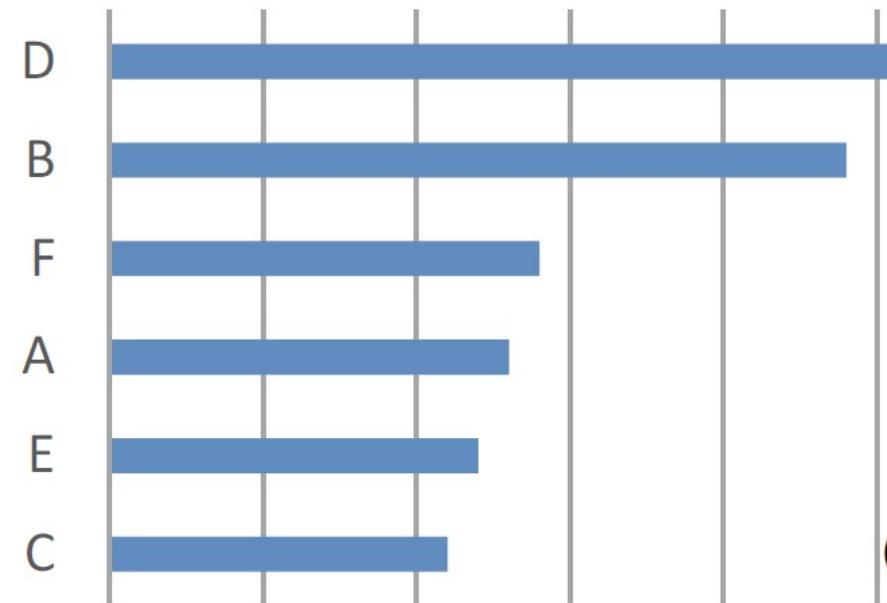
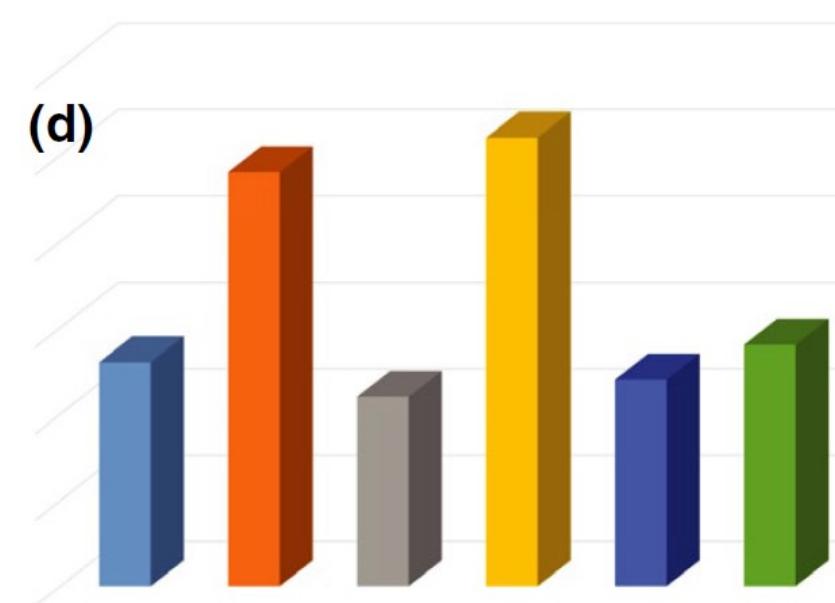
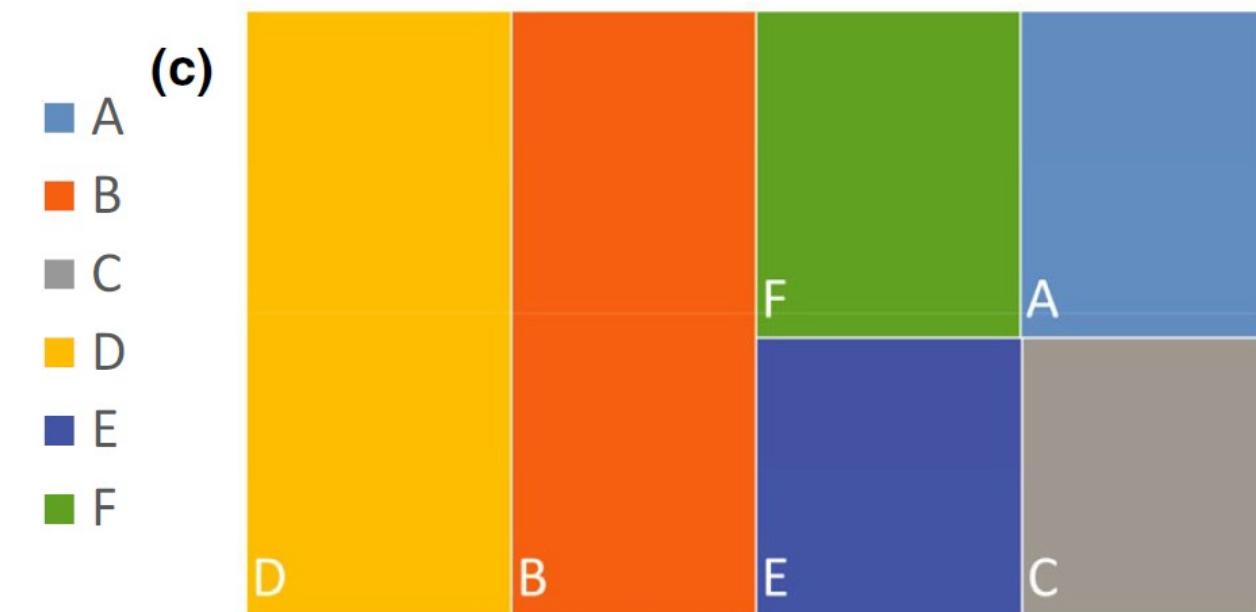
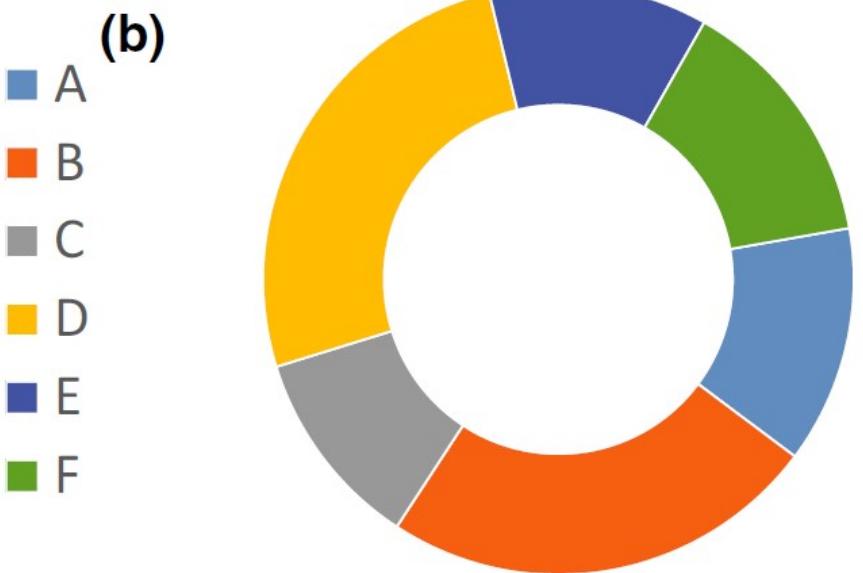
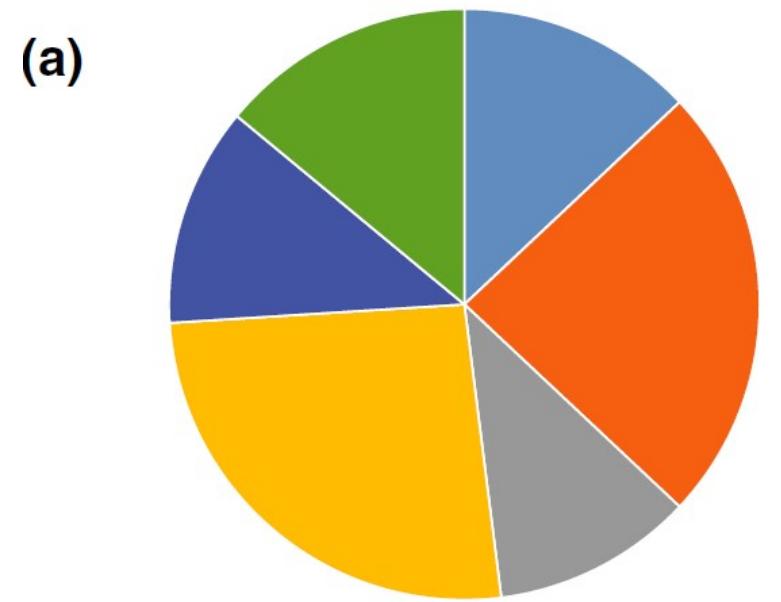
# Design principles for effective visual presentation

- Guides for Encoding Values in Graph
- JunkCharts
- Practical Guides for Using Colour in Charts
- Data-ink

# Design principles for effective visual presentation

- Guides for Encoding Values in Graph
- JunkCharts
- Practical Guides for Using Colour in Charts
- Data-ink

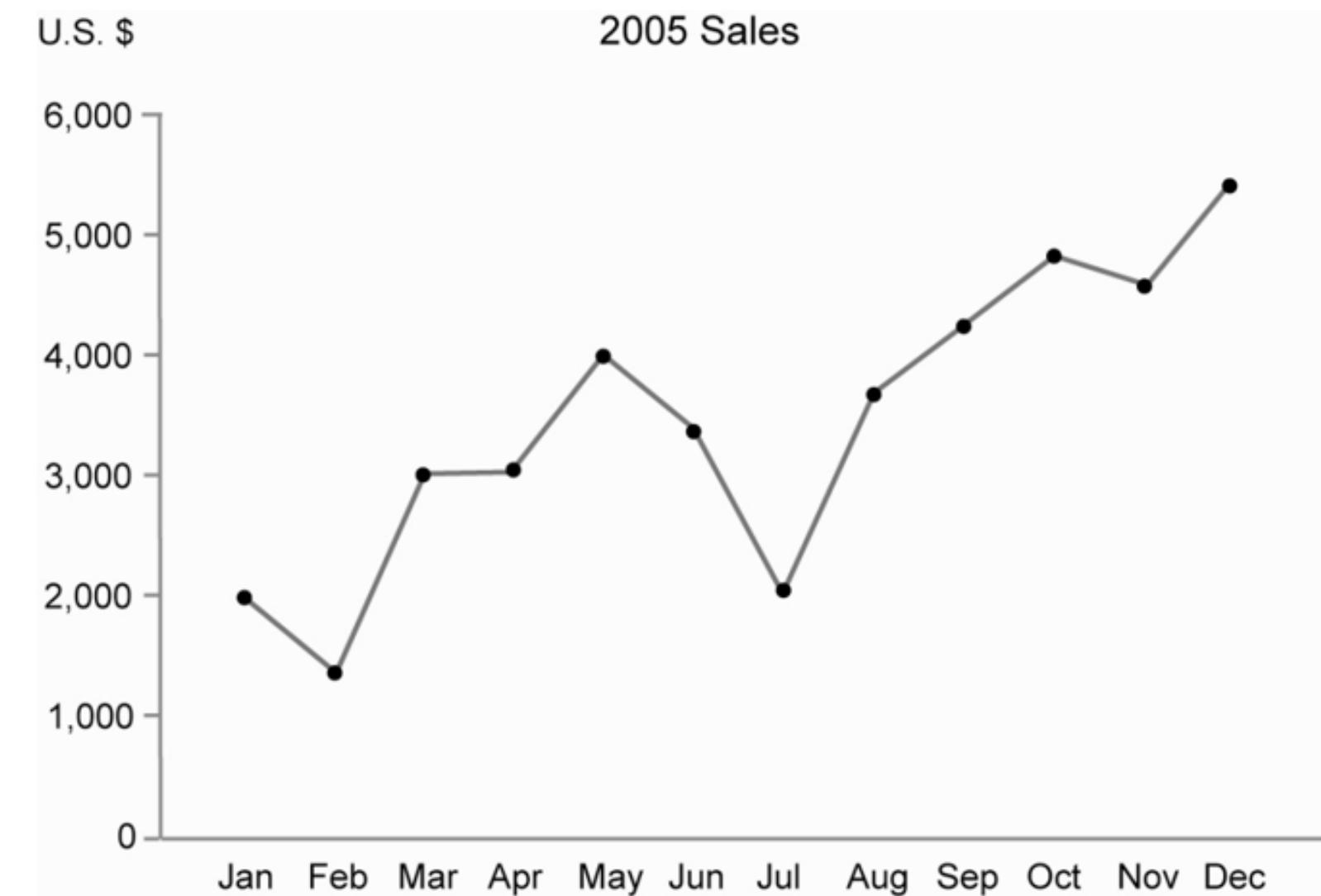
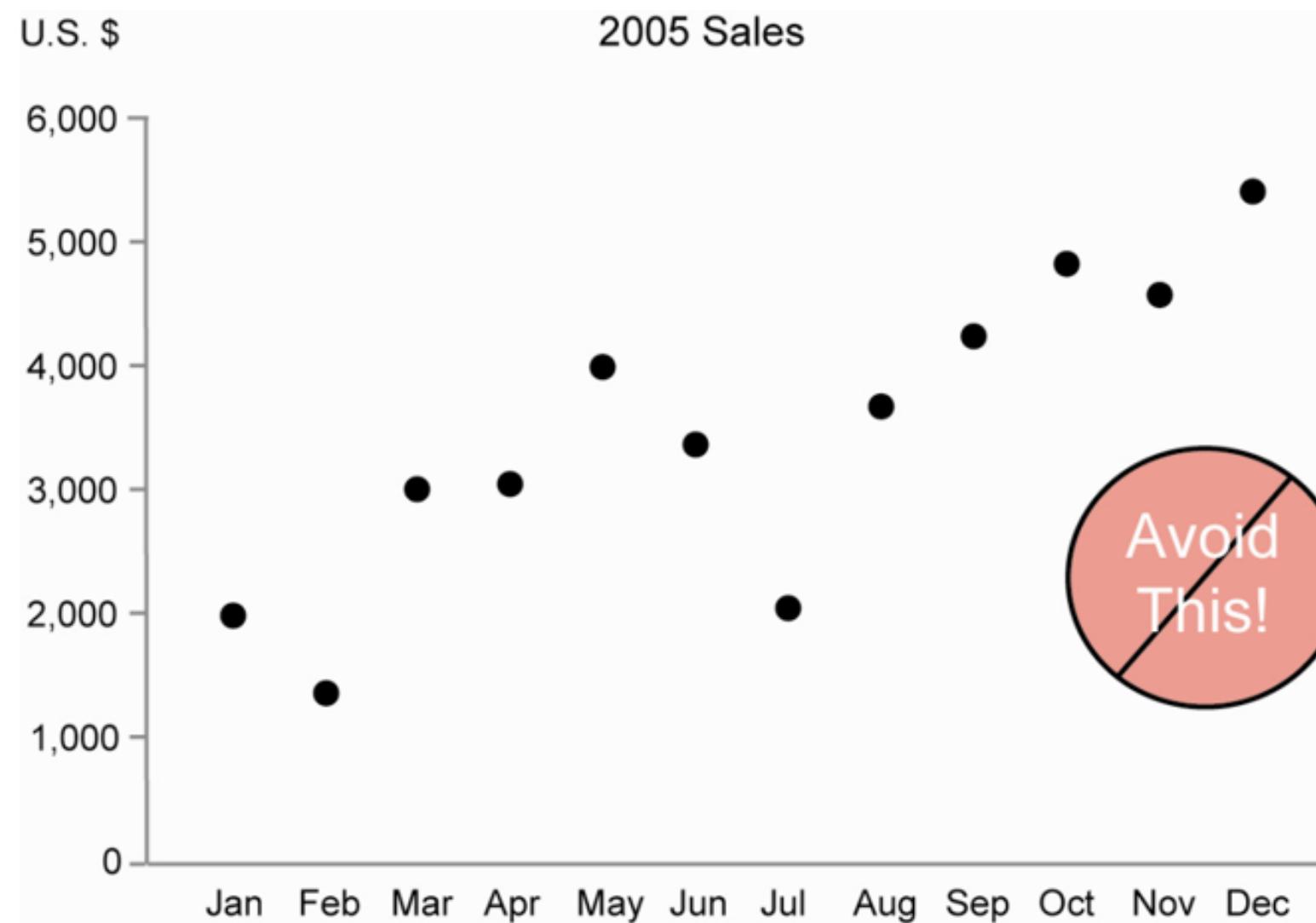
# General Guide: Show the Data Clearly



# Guides for Encoding Values in Graph

## Guide 1:

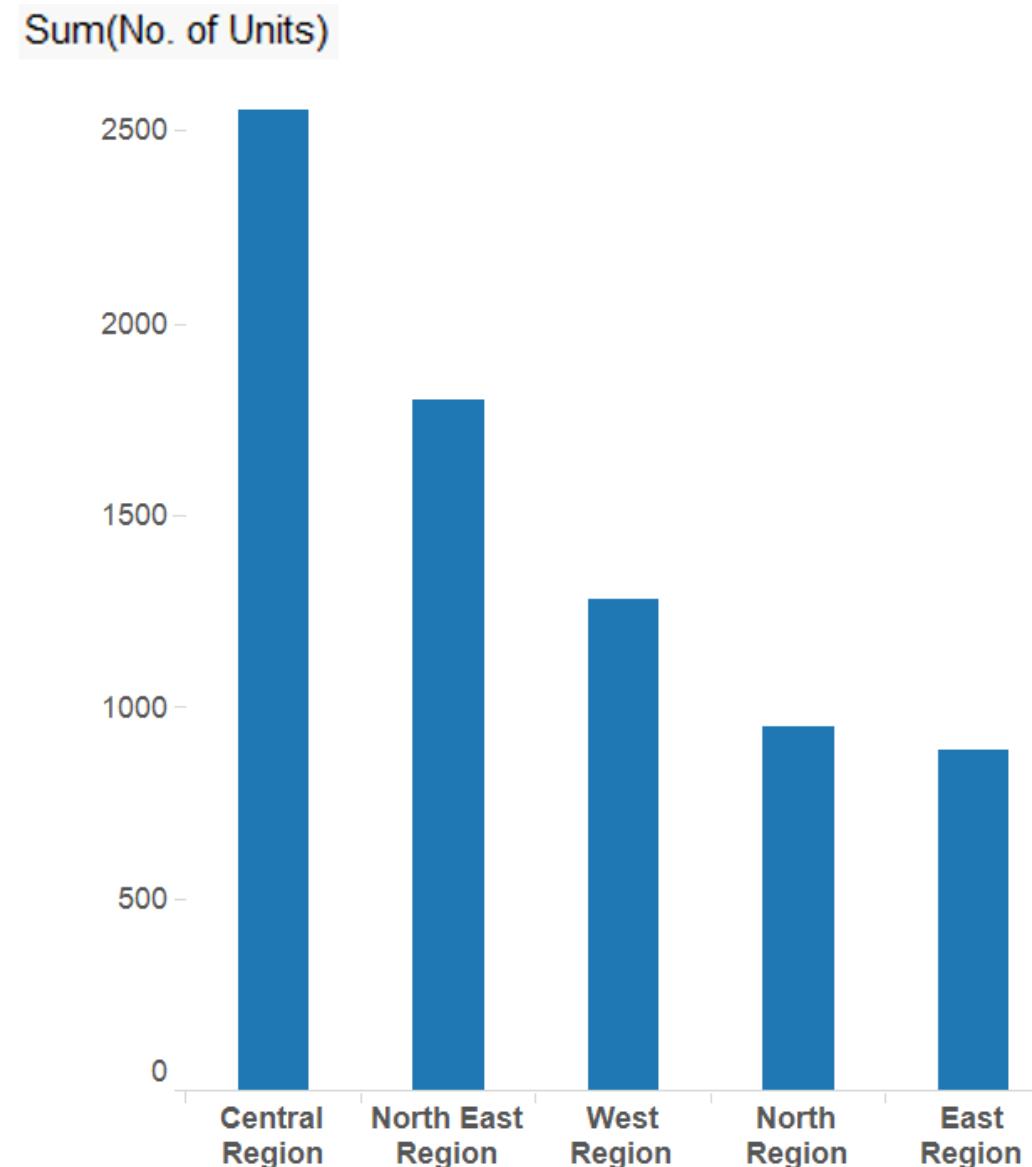
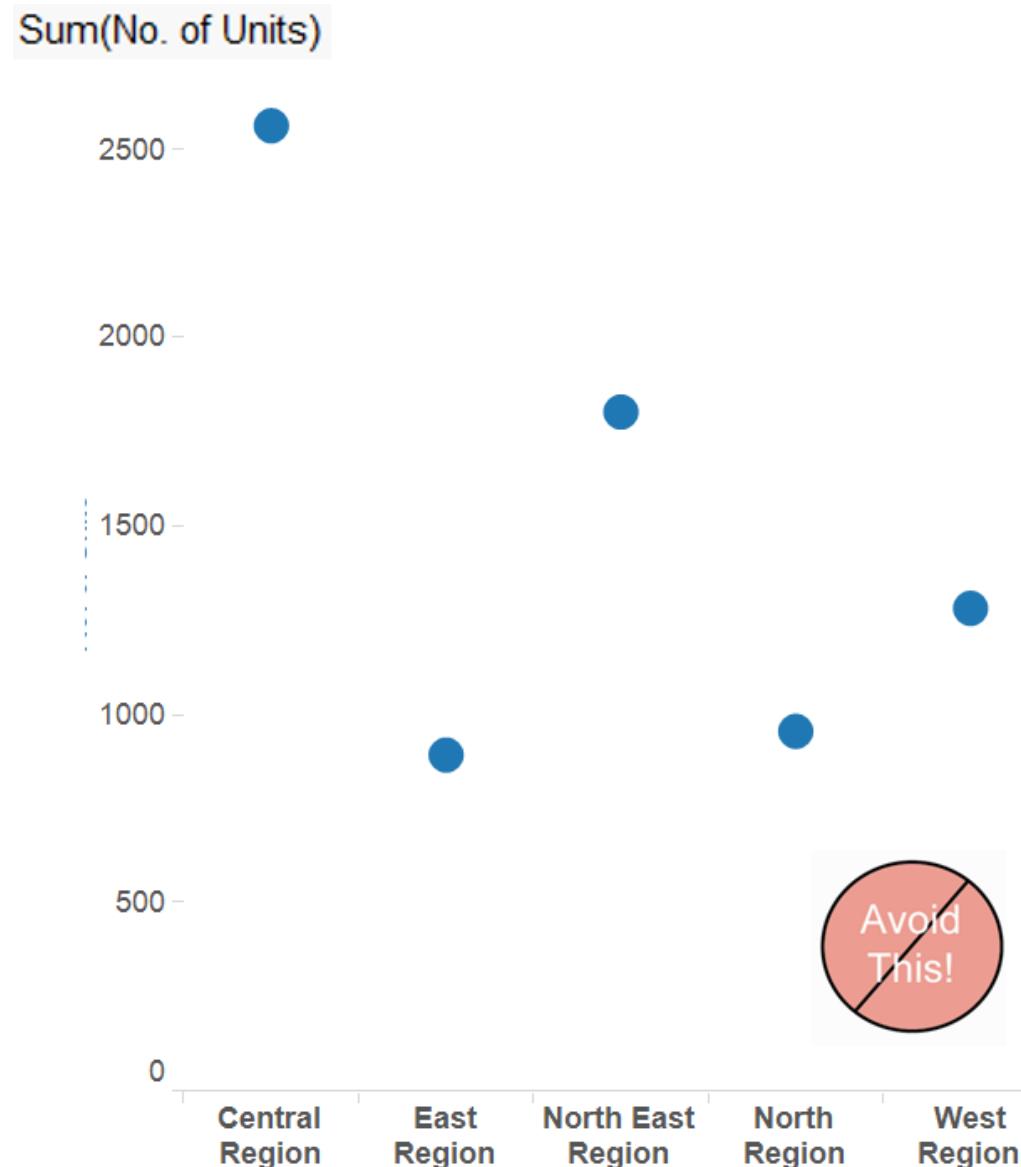
Avoid using point alone to display time-series data.



# Guides for Encoding Values in Graph

## (i) Guide 2:

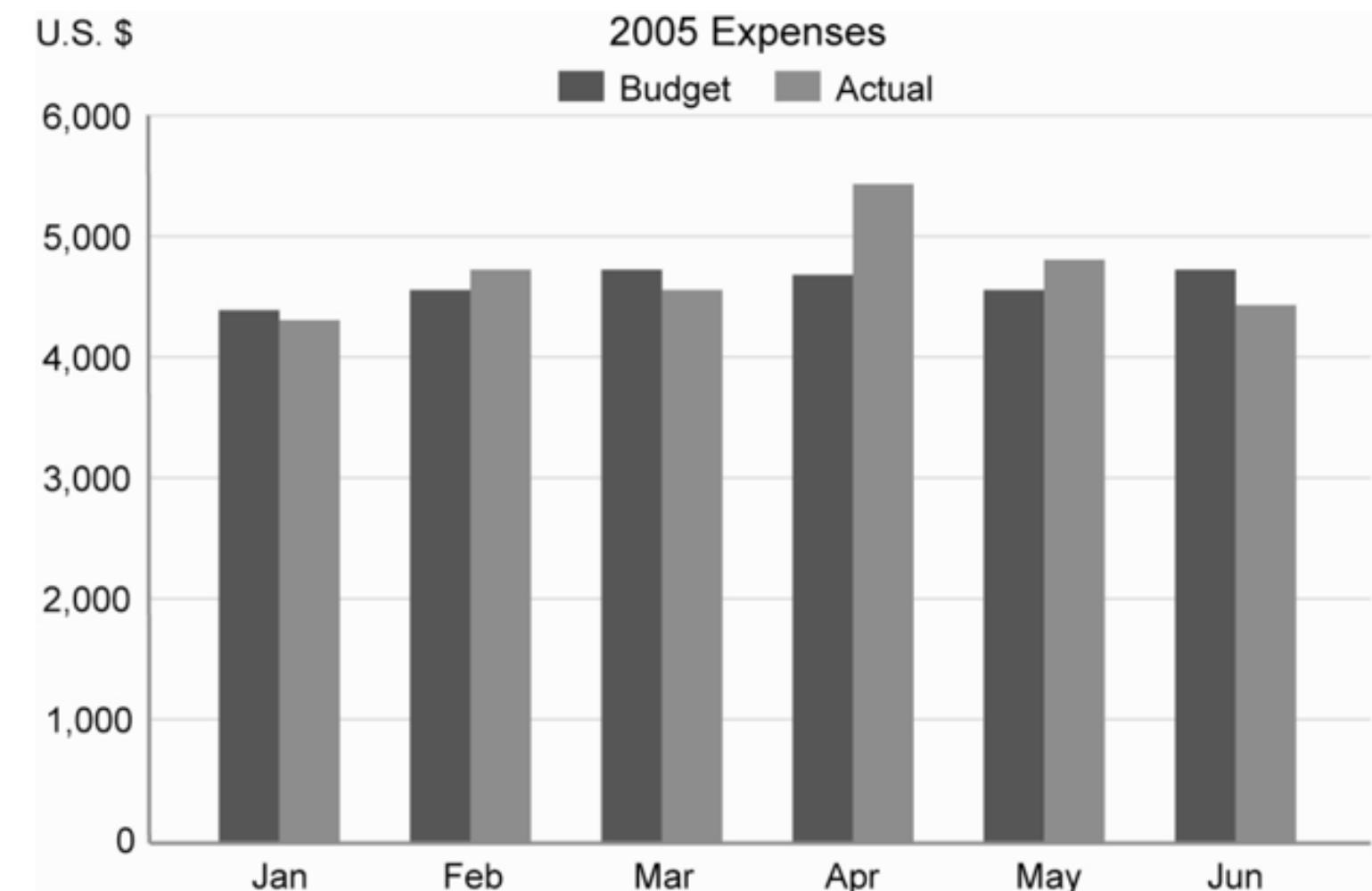
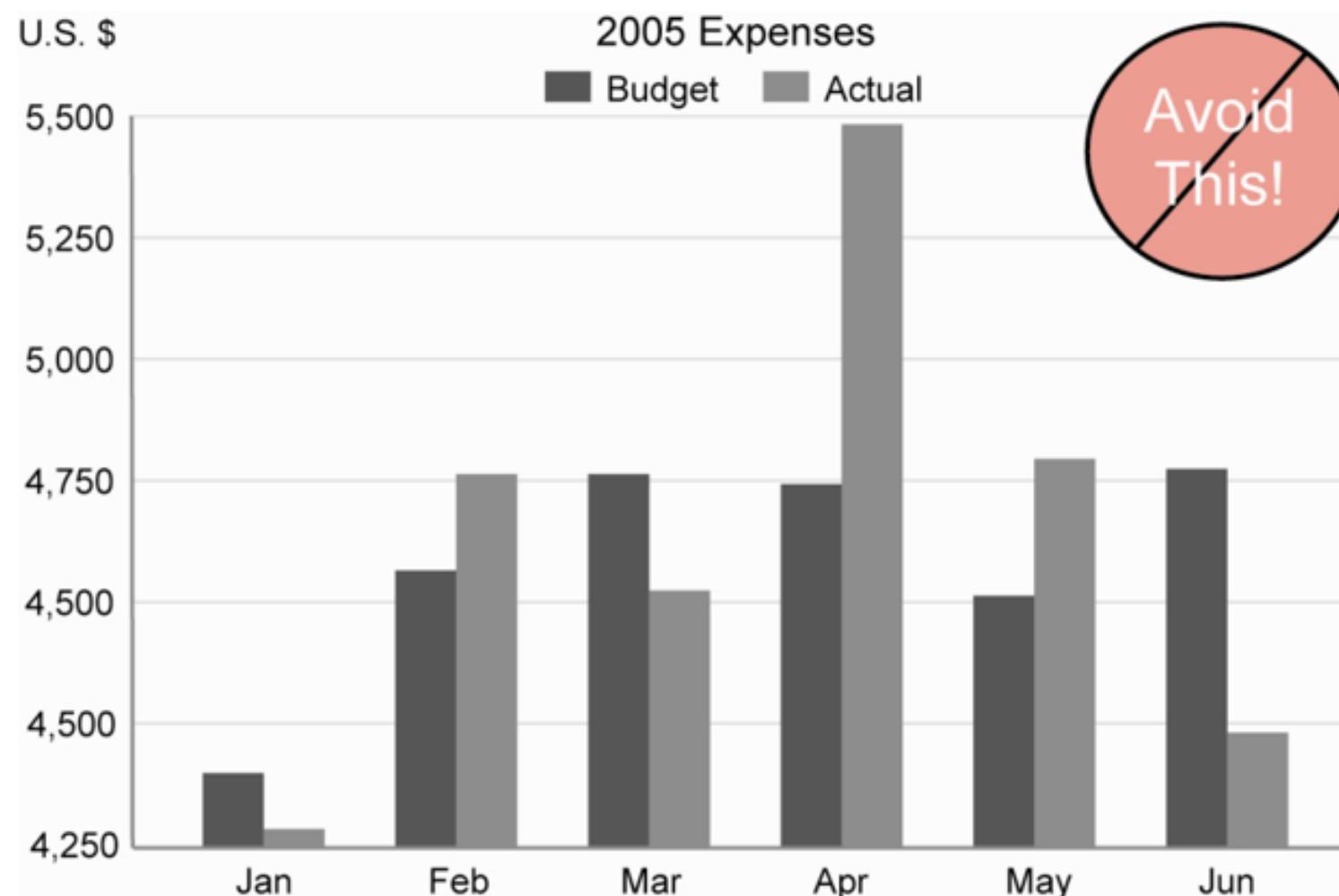
Avoid using points to represent discrete values



# Guides for Encoding Values in Graph

## (i) Guide 3:

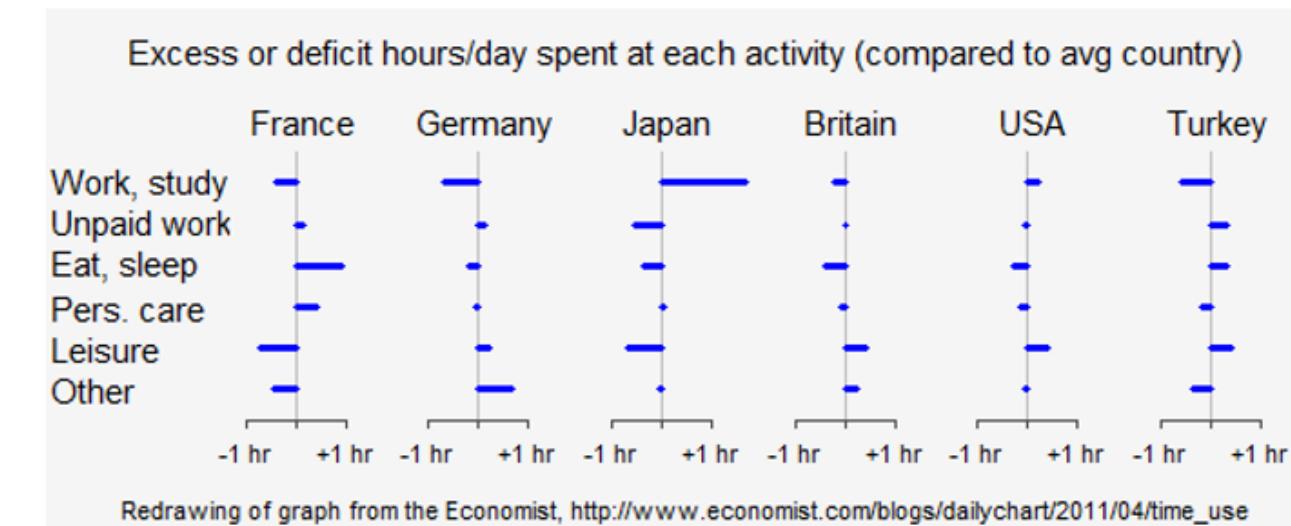
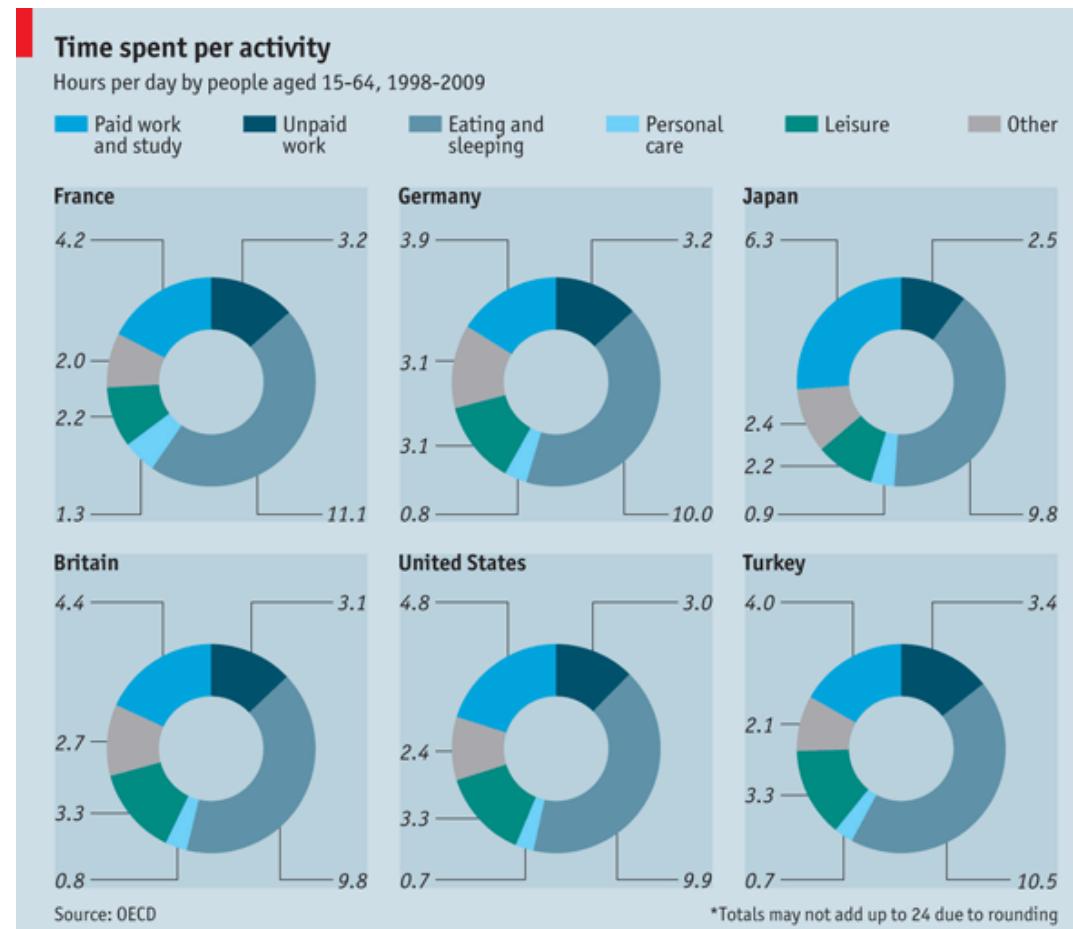
Bars don't work unless the quantitative scale begins at zero



# Guide for Encoding Values in Graph

## (i) Guide 4:

Avoid pie chart if possible because our eyes are not good in reading areas



Source: Time use: A day in the life,  
Apr 19th 2011, 15:00 by The Economist online

Reference: [JunkCharts](#)

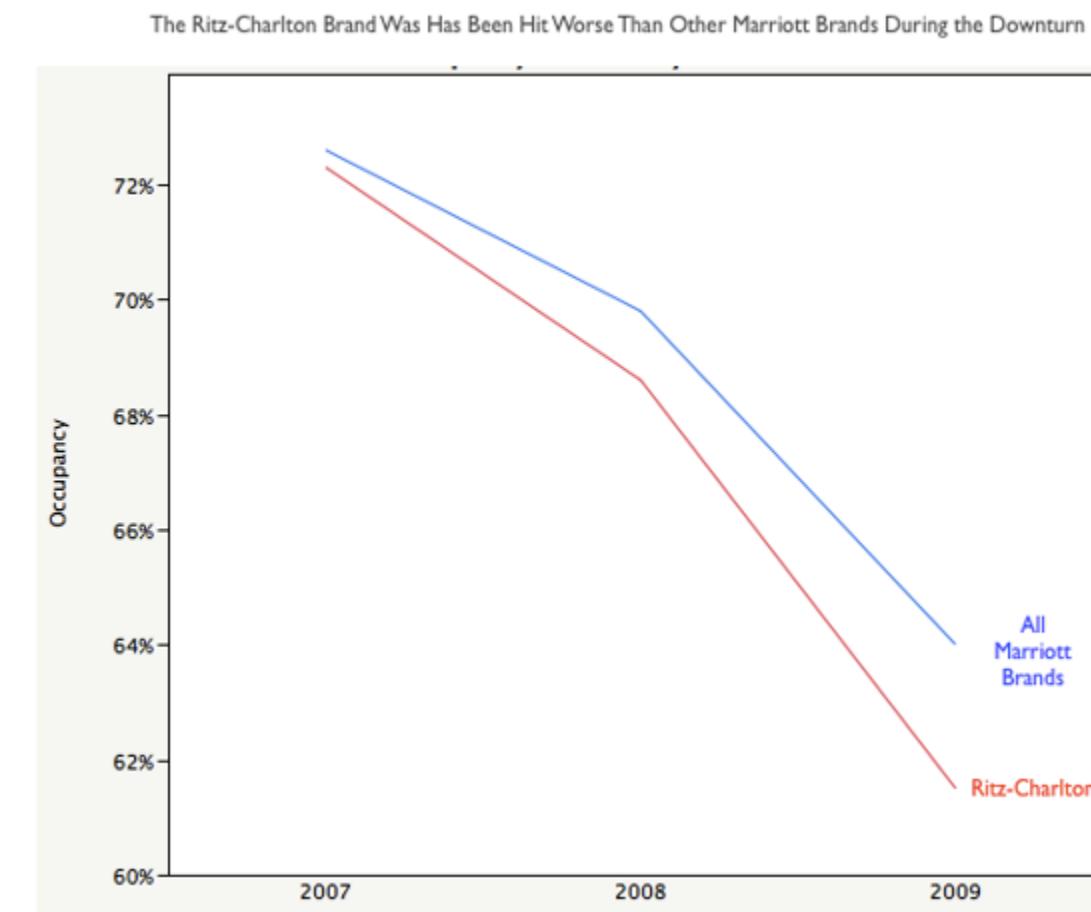
# Guide for Encoding Values in Graph

## (i) Guide 5:

Avoid pie chart if you are comparing changes over time



Avoid This!



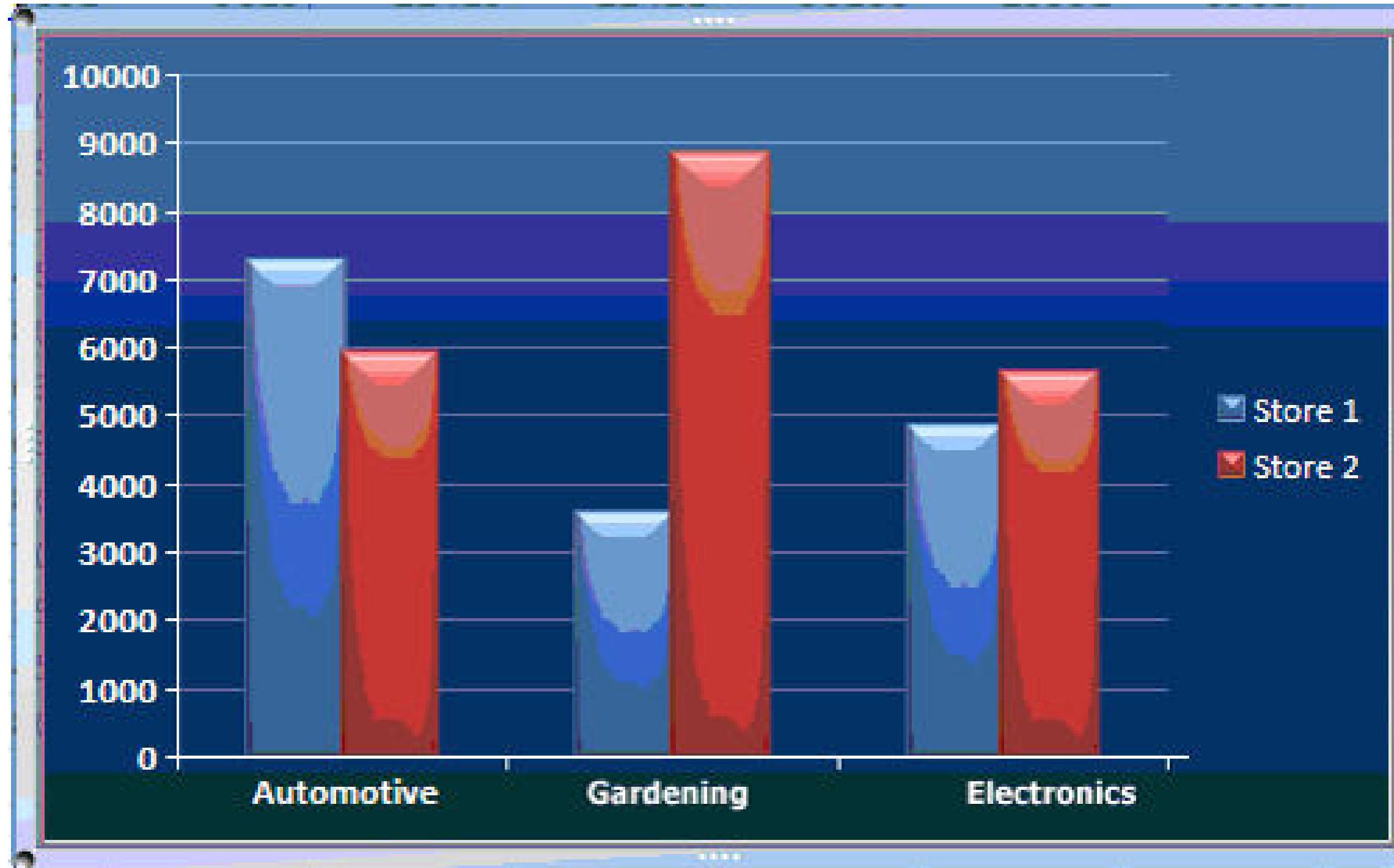
:scale 80%

# Design principles for effective visual presentation

- Guides for Encoding Values in Graph
- **Chartjunk**
- Practical Guides for Using Colour in Charts
- Data-ink

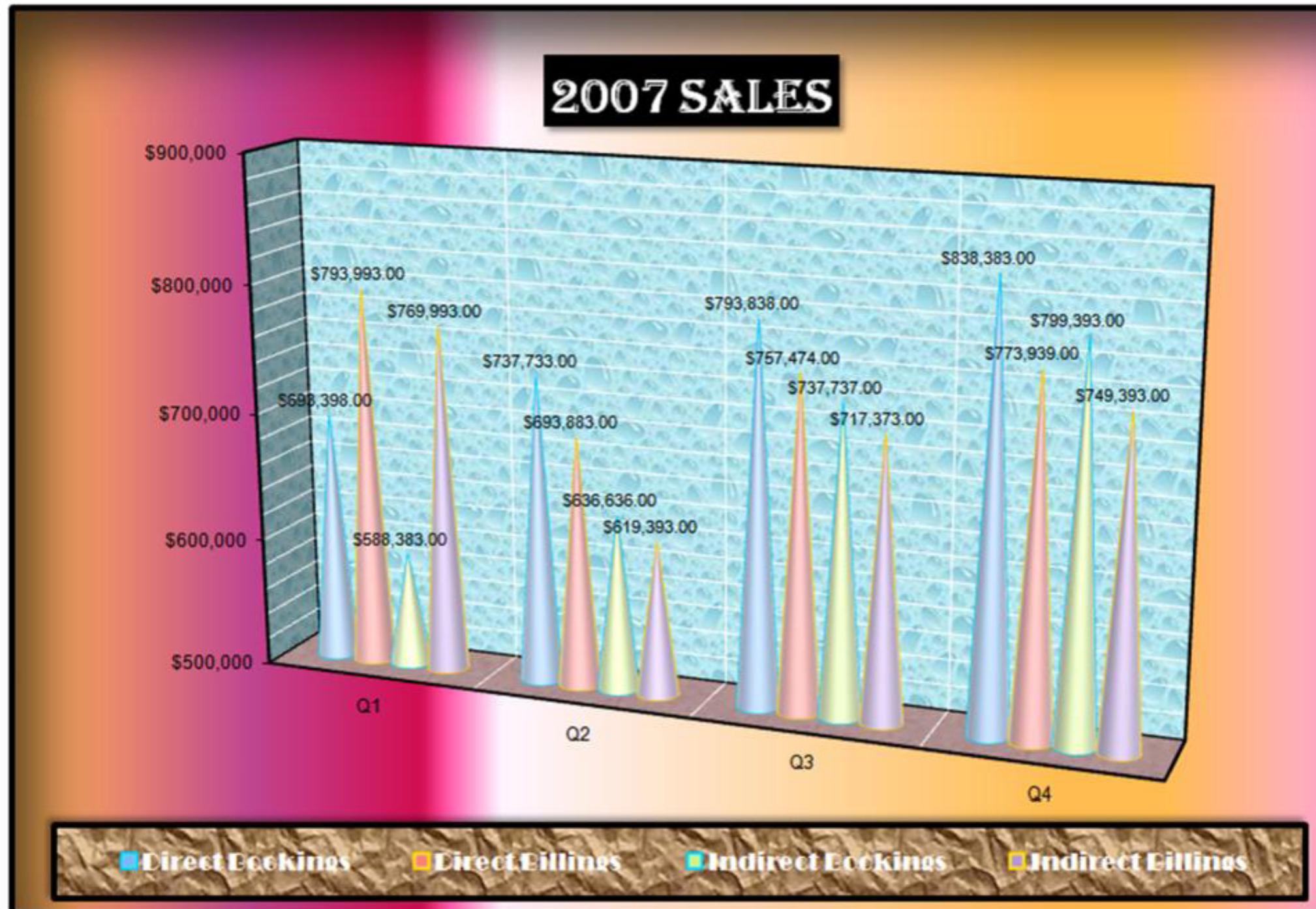
# ChartJunk I

Avoid using unnecessary colour shading for the bar.



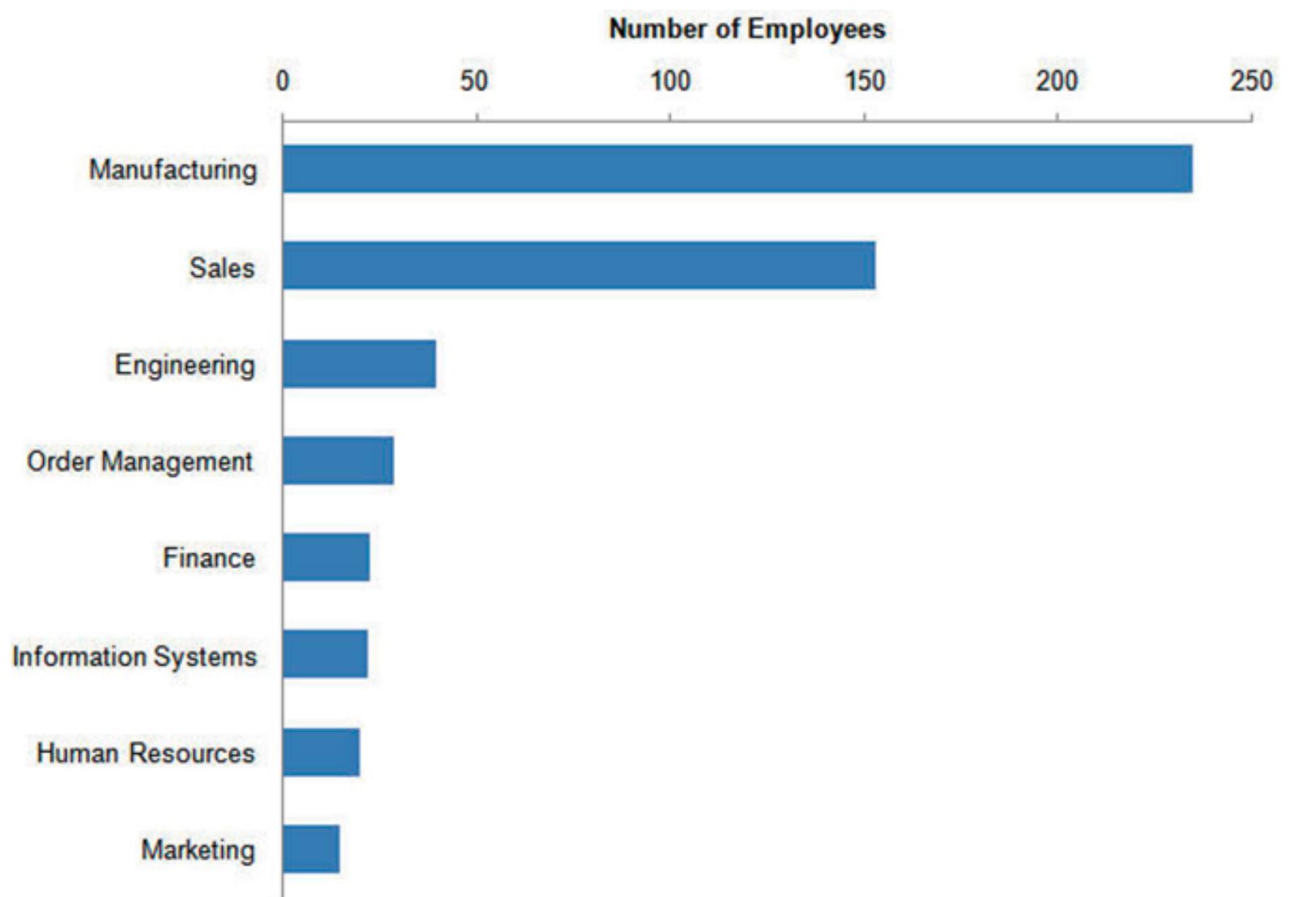
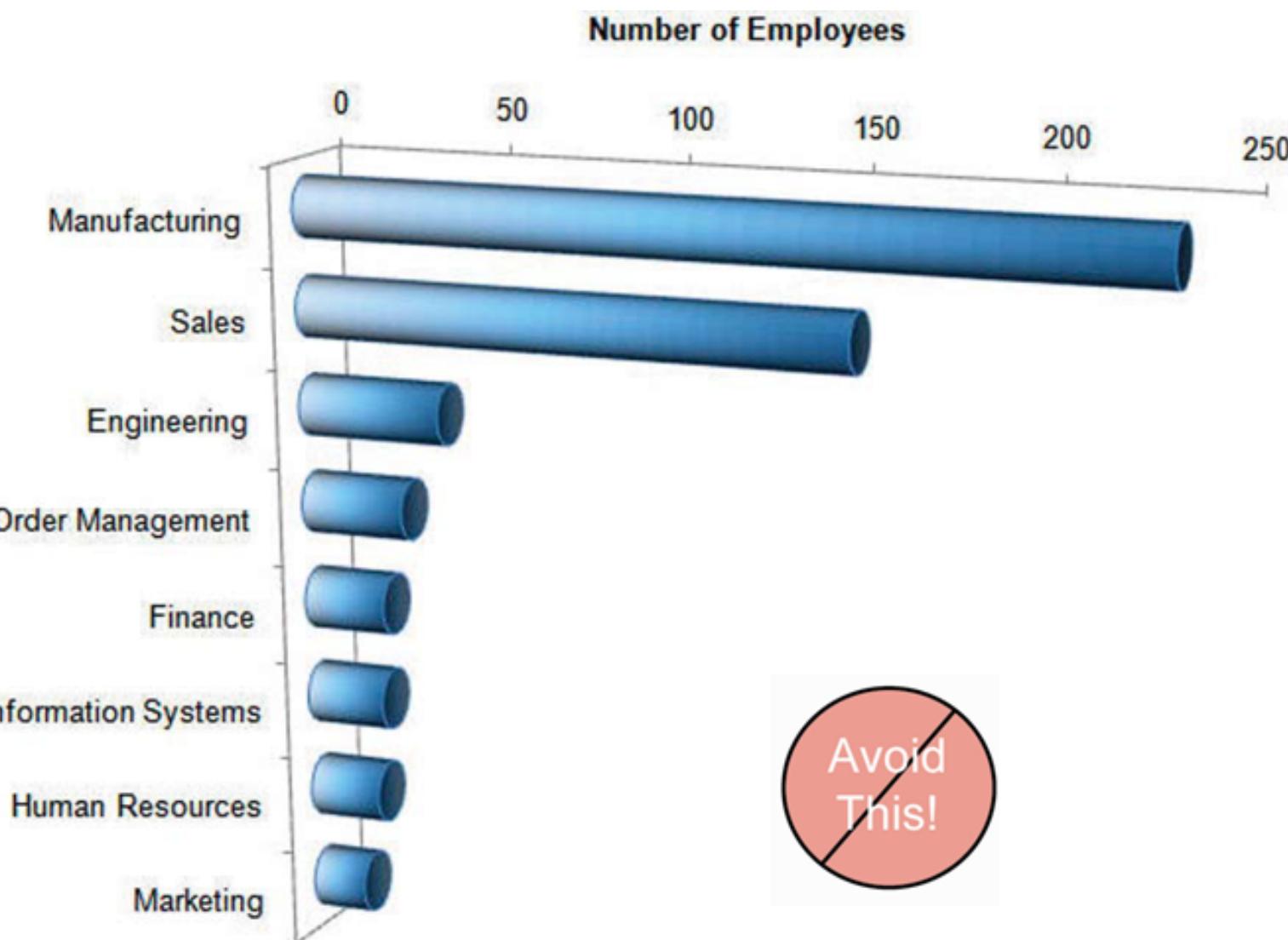
# ChartJunk II

Avoid colourful or wallpaper background.



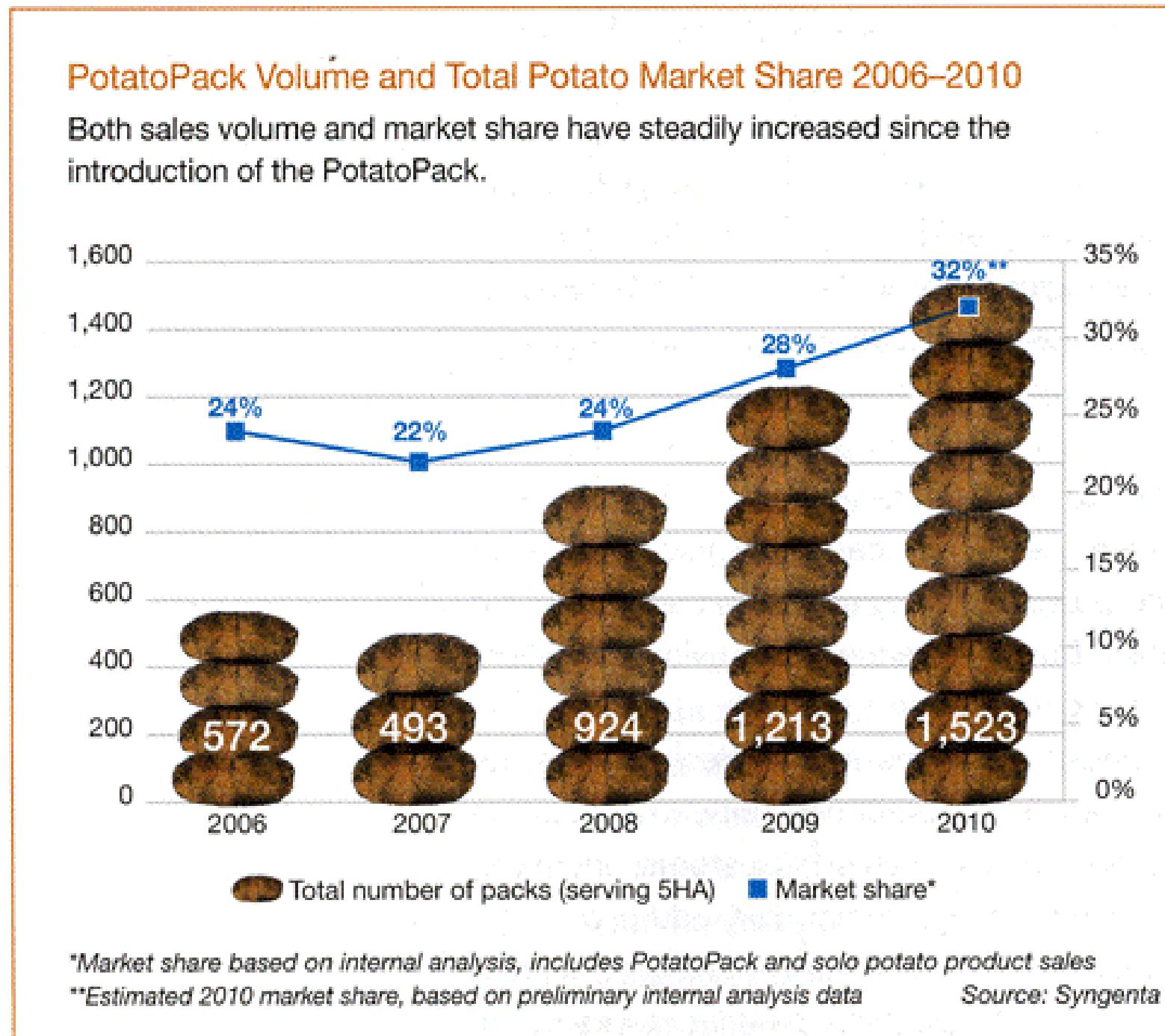
# ChartJunk III

Avoid using 3D effects in graphics.



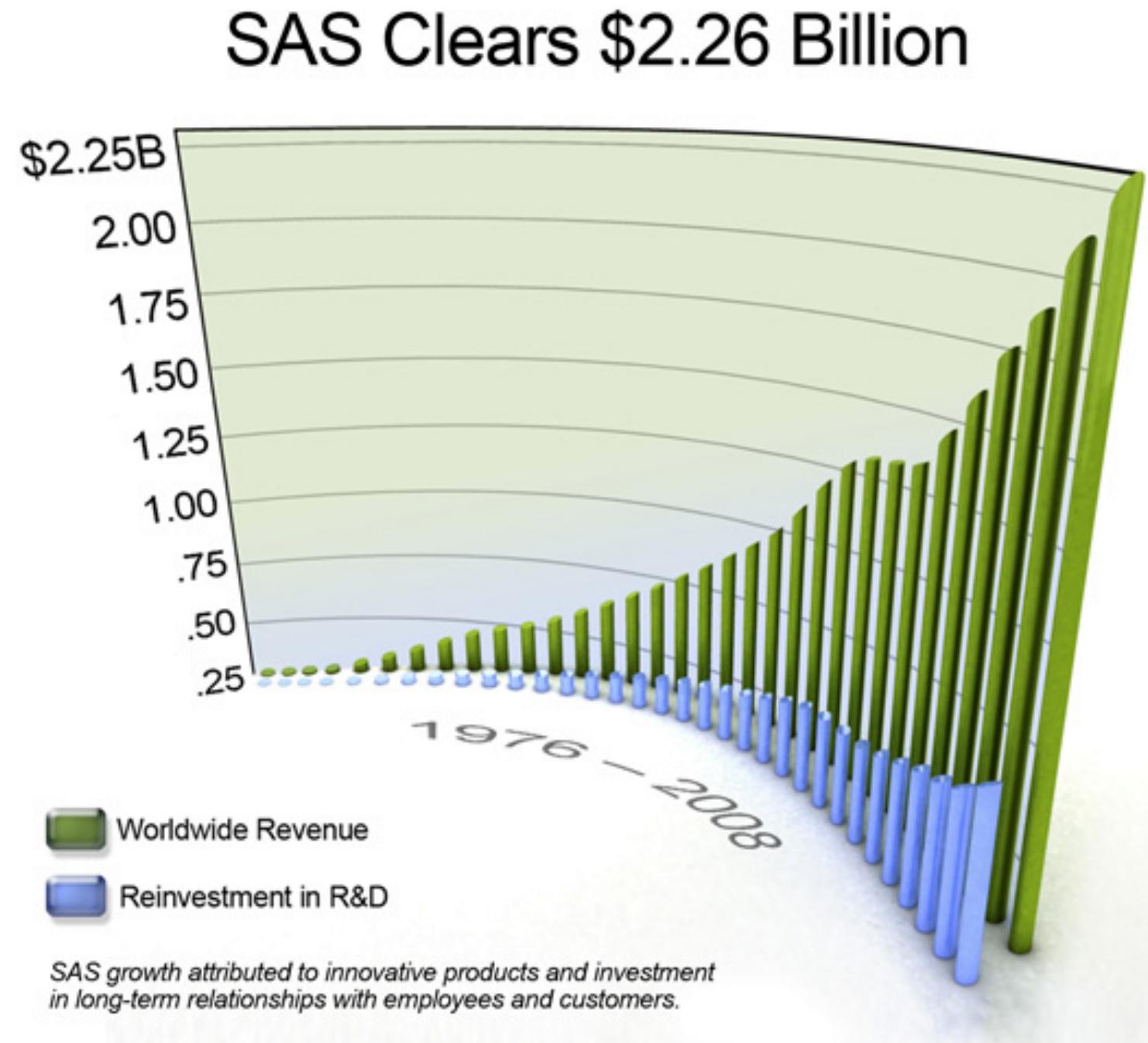
# ChartJunk IV

Avoid using misleading graphical representation.



# ChartJunk V

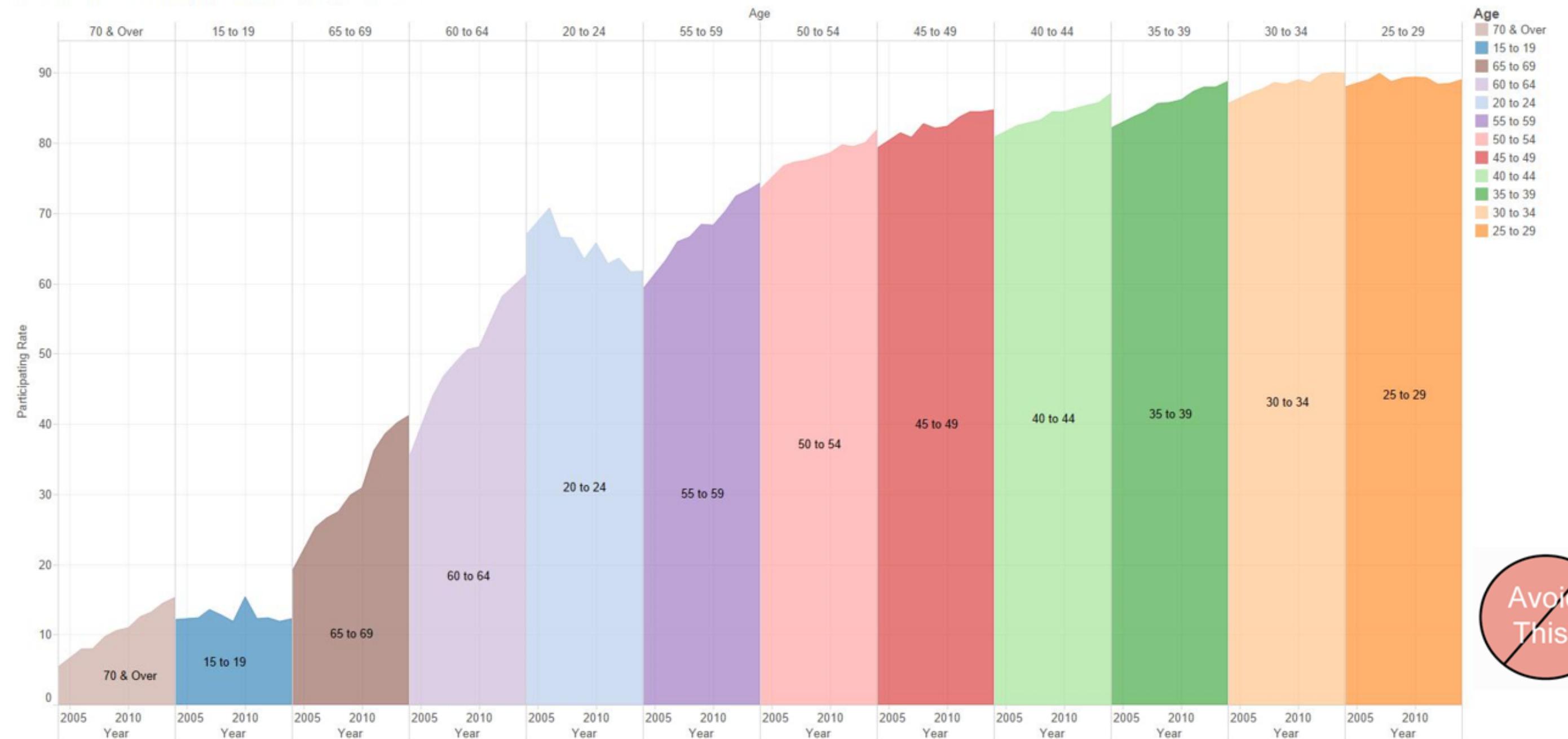
Avoid using artistic design which is difficult to visualise.



# ChartJunk VI

Designing graph to enlighten people – not to entertain them.

Labour Force Participation Rate by Age and Year



The plot of sum of Participating Rate for Year broken down by Age. Color shows details about Age. The marks are labeled by Age. The view is filtered on sum of Participating Rate, Age and Year. The sum of Participating Rate filter keeps non-Null values only. The Age filter excludes Total. The Year filter ranges from 2004 to 2014.

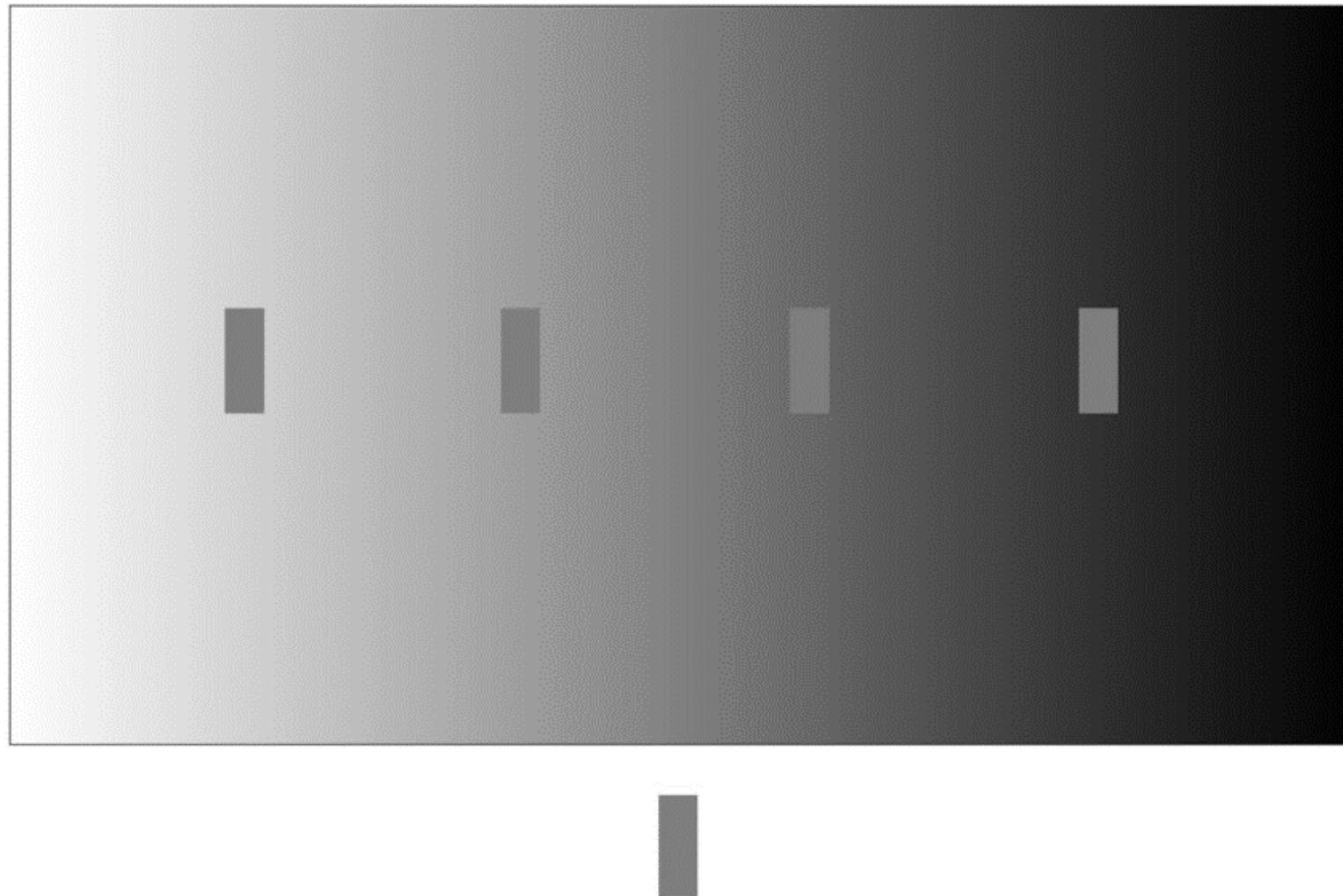
# Design principles for effective visual presentation

- Guides for Encoding Values in Graph
- JunkCharts
- Practical Guides for Using Colour in Charts
- Data-ink

# Practical Guides for Using Colour in Charts

## Guide 1:

If you want different objects of the same colour in a graph to look the same, make sure that the background- the colour that surrounds them – is consistent.



# Practical Guides for Using Colour in Charts

## (i) Guide 2:

If you want objects in a graph to be easily seen, use a background colour that contrasts sufficiently with the object.

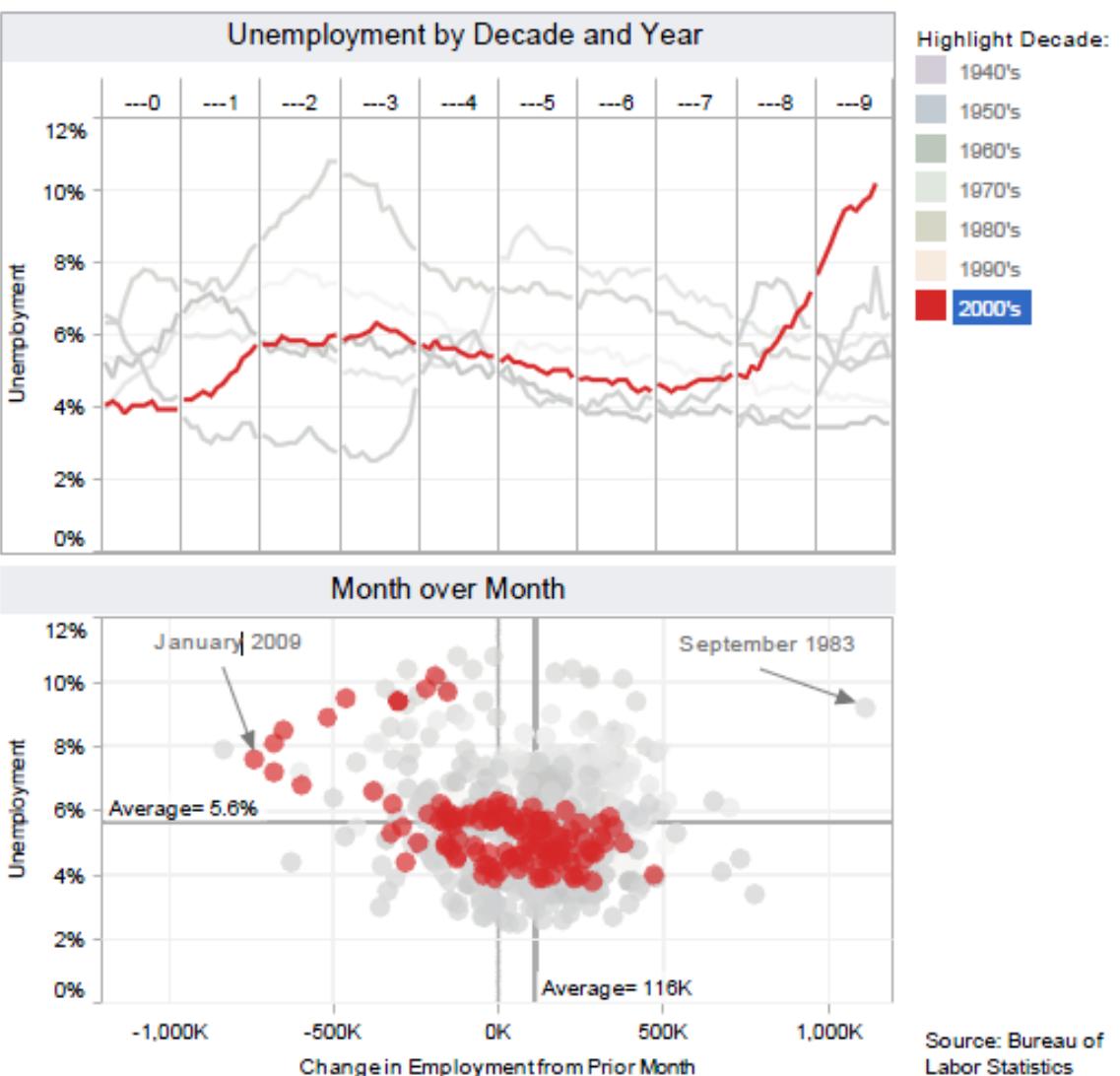
name	age	sex	height	weight	mean weight by age
KATIE	12	F	59	95	99.000
LOUISE	12	F	61	123	99.000
JANE	12	F	55	74	99.000
JACLYN	12	F	66	145	99.000
LILLIE	12	F	52	64	99.000
TIM	12	M	60	84	99.000
JAMES	12	M	61	128	99.000
ROBERT	12	M	51	79	99.000
BARBARA	13	F	60	112	94.714
ALICE	13	F	61	107	94.714
SUSAN	13	F	56	67	94.714
JOHN	13	M	65	98	94.714
JOE	13	M	63	105	94.714
MICHAEL	13	M	58	95	94.714
DAVID	13	M	59	79	94.714
JUDY	14	F	61	81	100.833
ELIZABETH	14	F	62	91	100.833
LESLIE	14	F	65	142	100.833
CAROL	14	F	63	84	100.833
PATTY	14	F	62	85	100.833
FREDERICK	14	M	63	93	100.833
ALFRED	14	M	64	99	100.833
HENRY	14	M	65	119	100.833
LEWIS	14	M	64	92	100.833
EDWARD	14	M	68	112	100.833
CHRIS	14	M	64	99	100.833
JEFFREY	14	M	69	113	100.833
MARY	15	F	62	92	108.286
AMY	15	F	64	112	108.286
ROBERT	15	M	67	128	108.286
WILLIAM	15	M	65	111	108.286

# Practical Guides for Using Colour in Charts

## (i) Guide 3:

Use color only when needed to serve a particular communication goal.

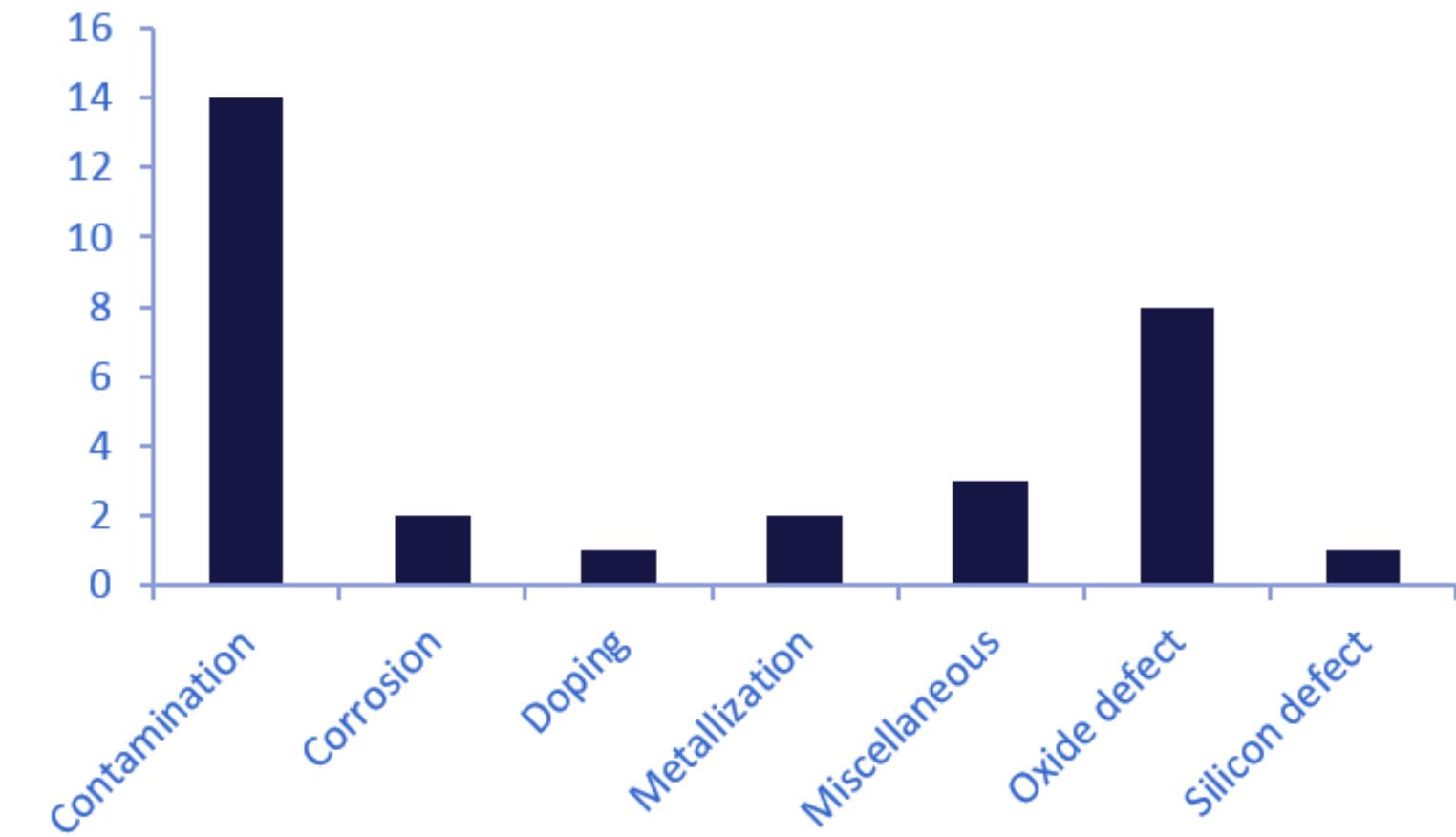
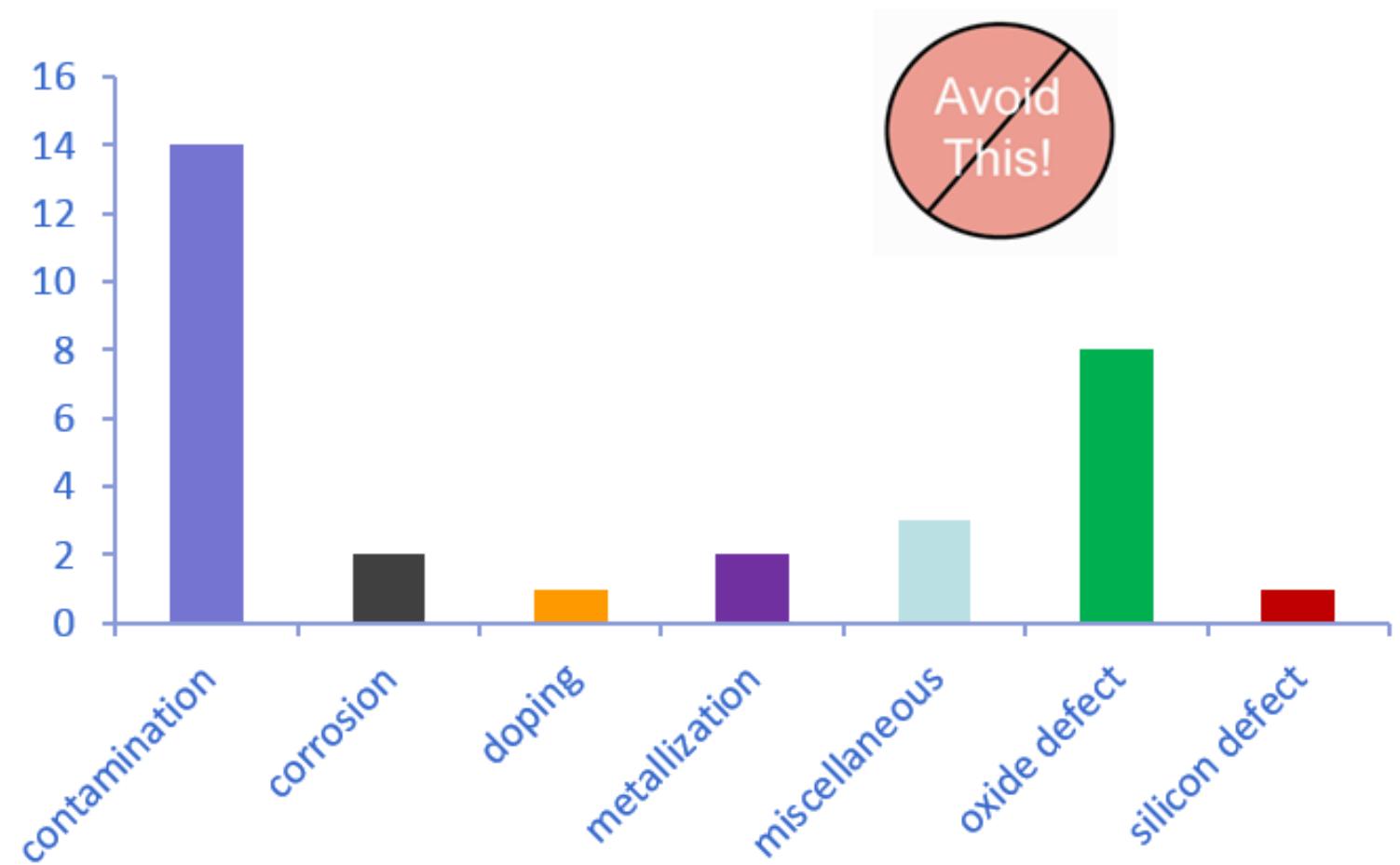
### 60 Years of Unemployment



# Practical Guides for Using Colour in Charts

## (i) Guide 4:

Use different colours when they correspond to differences of meaning in the data.

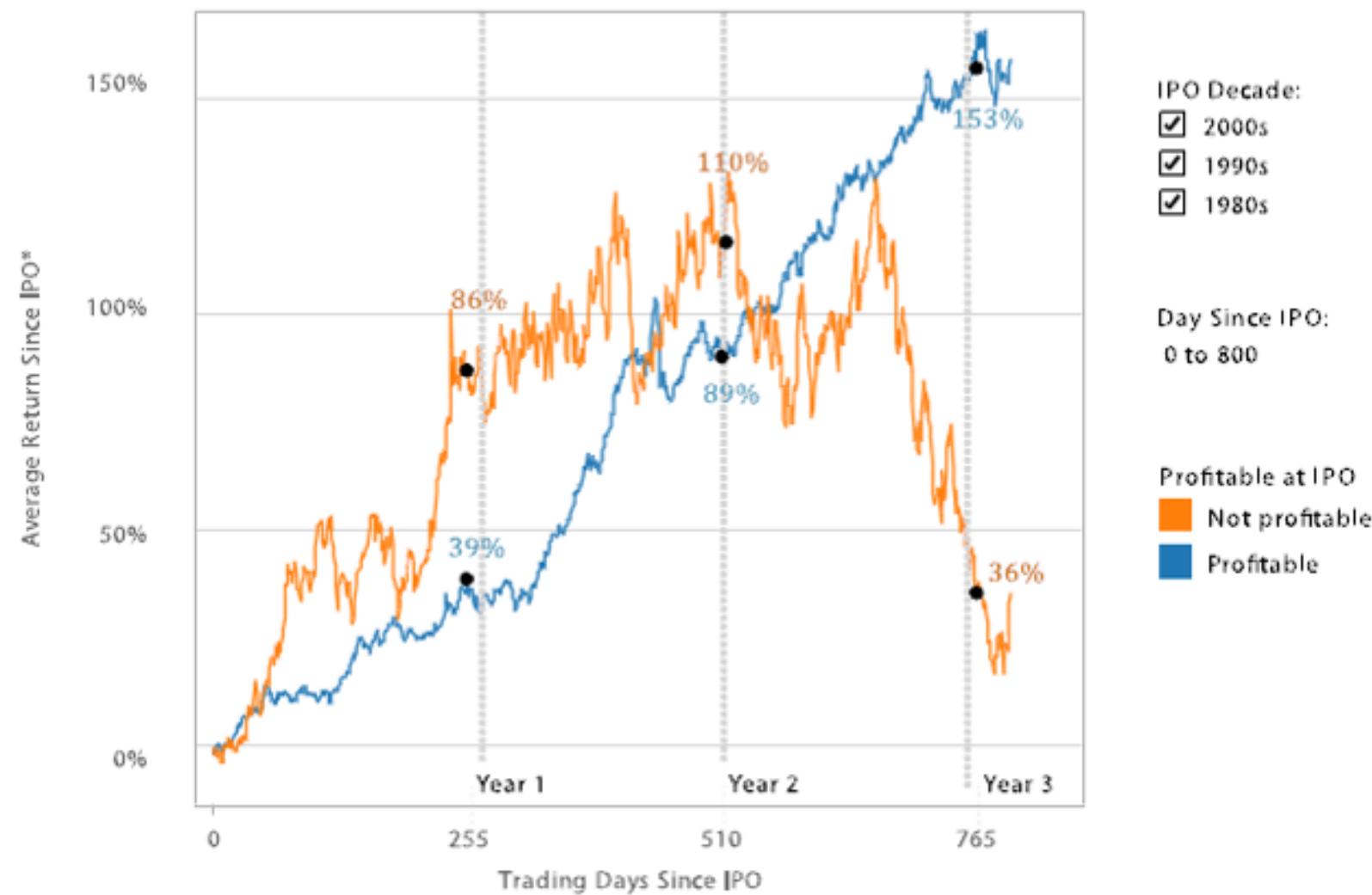


# Practical Guides for Using Colour in Charts

## (i) Guide 5:

Use soft, natural colours to display most information and bright and/or dark colours to highlight information that requires greater attention.

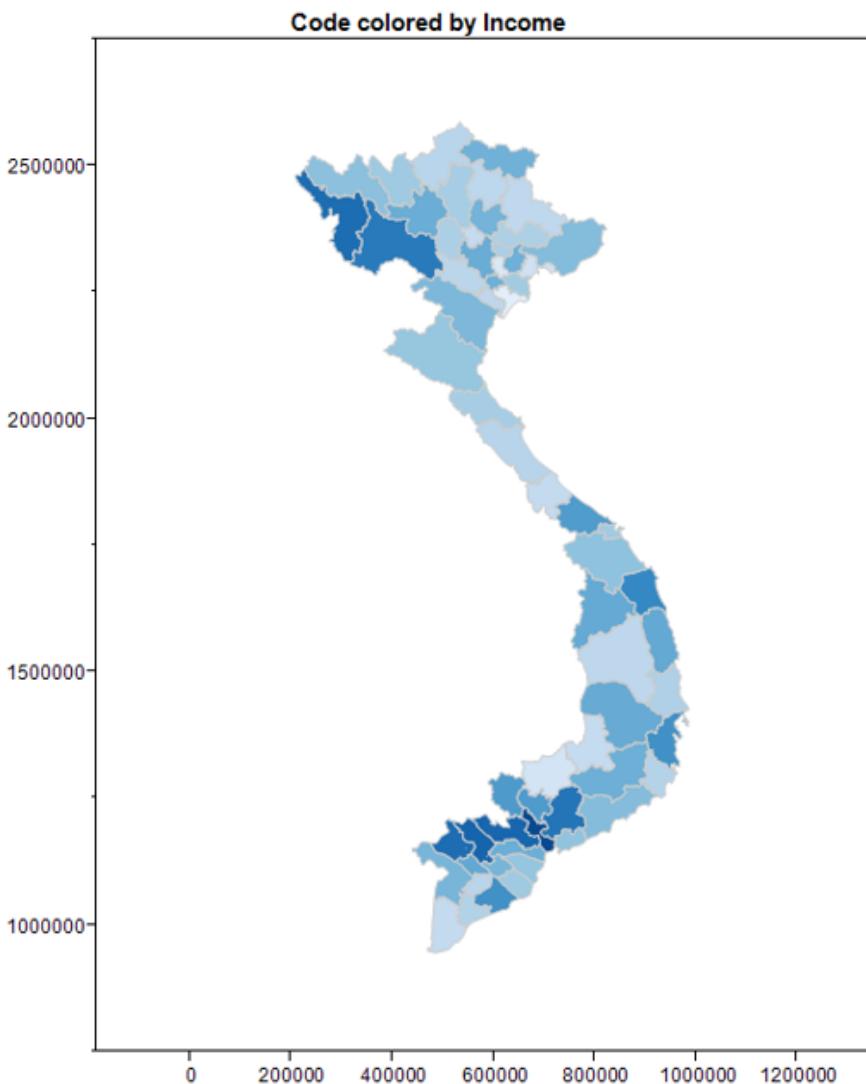
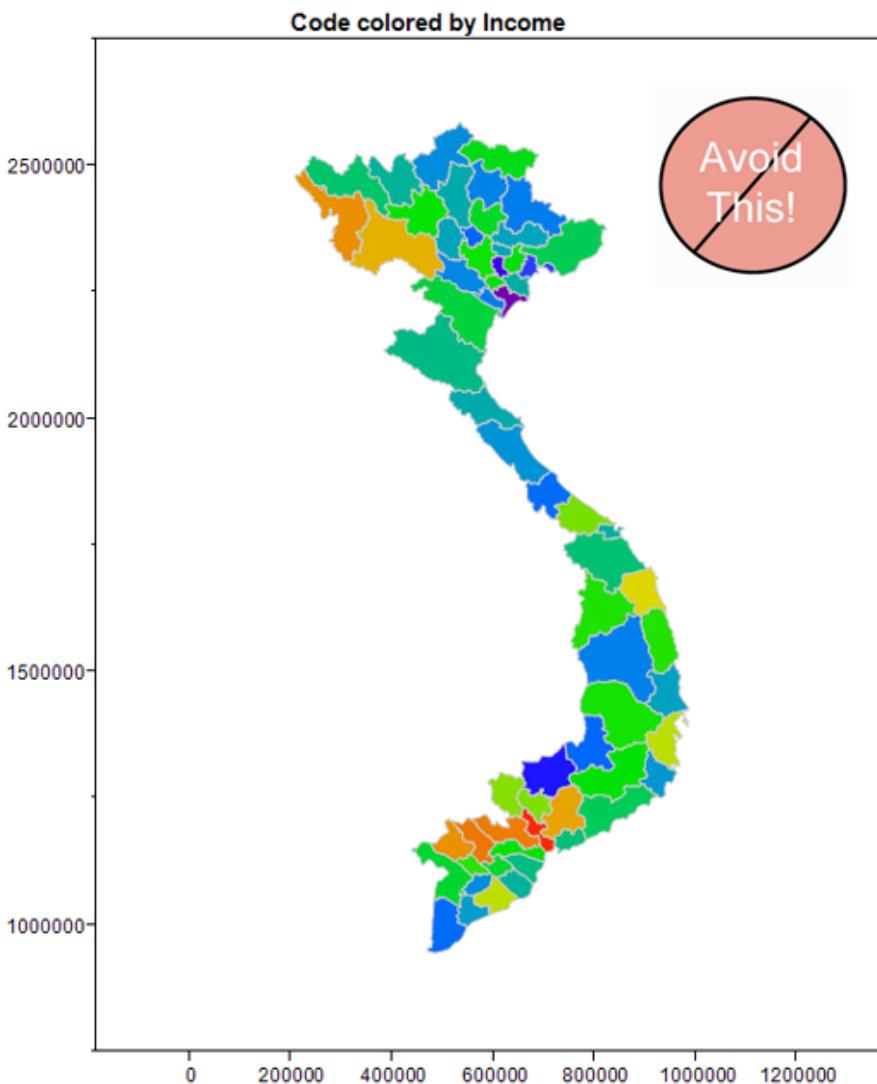
### Profitable vs. Unprofitable IPOs



# Practical Guides for Using Colour in Charts

## ① Guide 6:

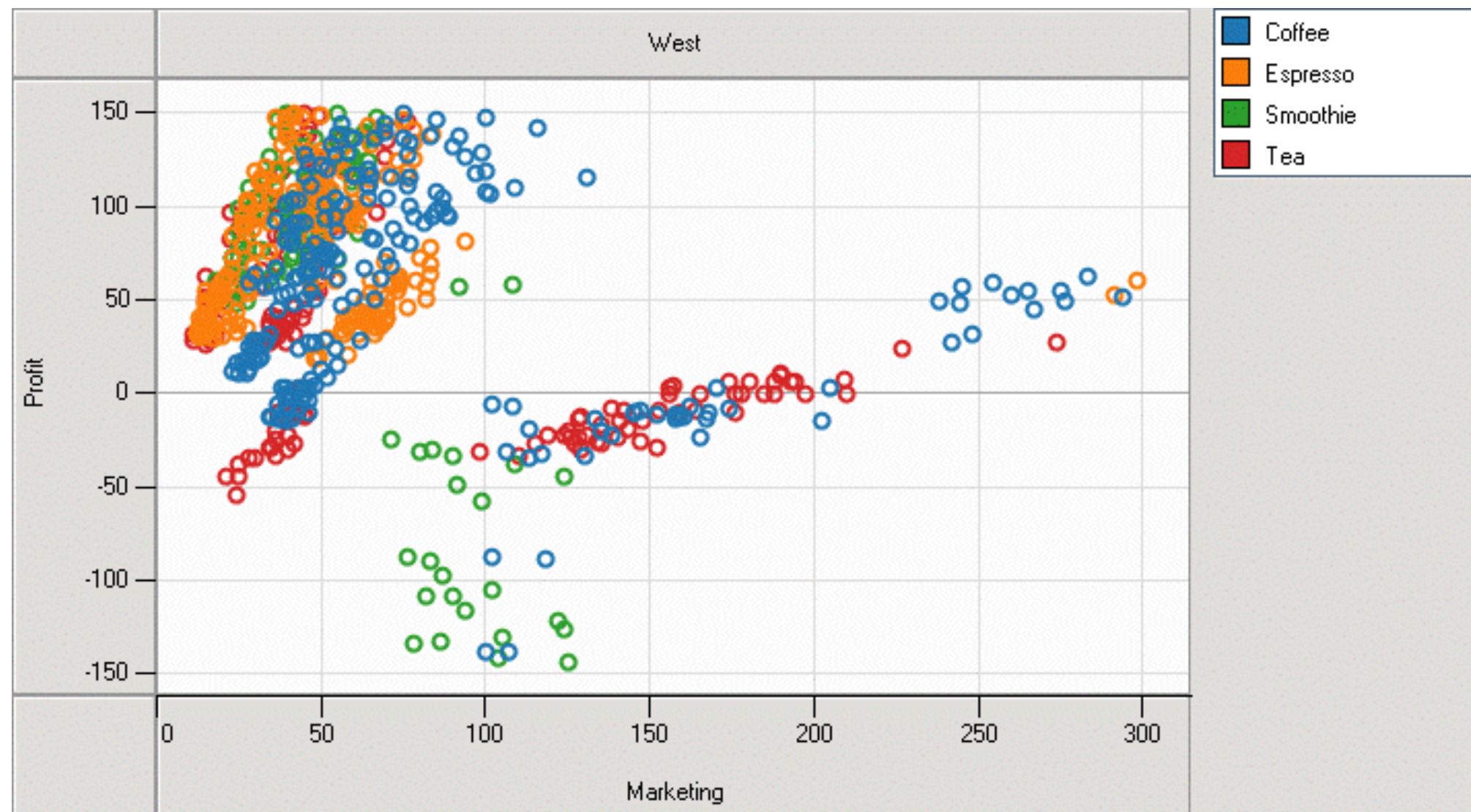
When using colour to encode a sequential range of quantitative values, stick with a single hue (or a small set of closely related hues) and vary intensity from pale colours for low values to increasingly darker and brighter colours for high values.



# Practical Guides for Using Colour in Charts

## (i) Guide 7:

Non-data components of a graph should be displayed just visibly enough to perform their role, but not more so, for excessive salience could cause them to distract attention from the data.



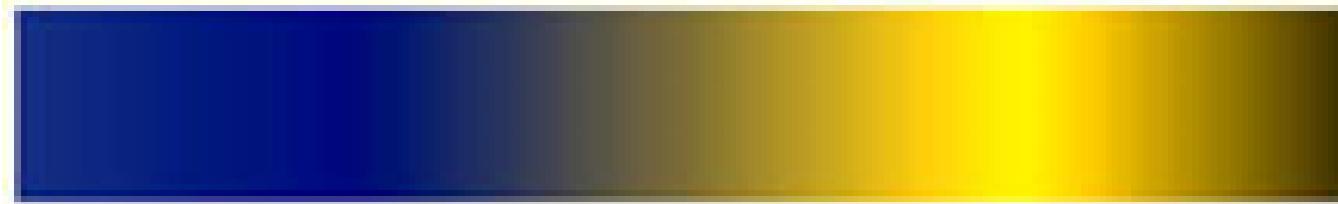
# Practical Guides for Using Colour in Charts

## Guide 8:

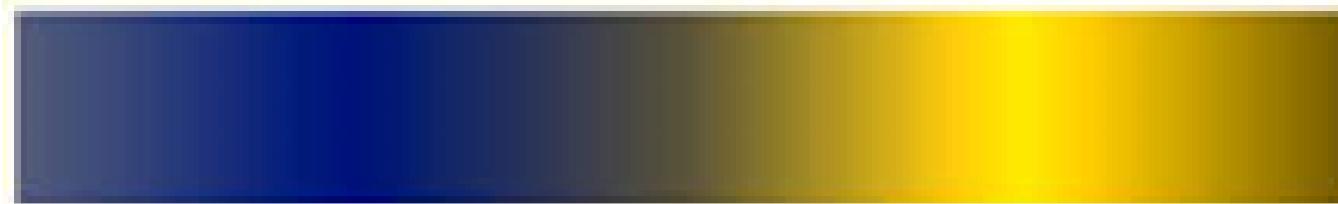
To guarantee that most people who are colourblind can distinguish groups of data that are colour coded, avoid using a combination of red and green in the same display.



**non color blind**



**protanope**  
**(red cone cells defective)**



**deuteranope**  
**(green cone cells defective)**



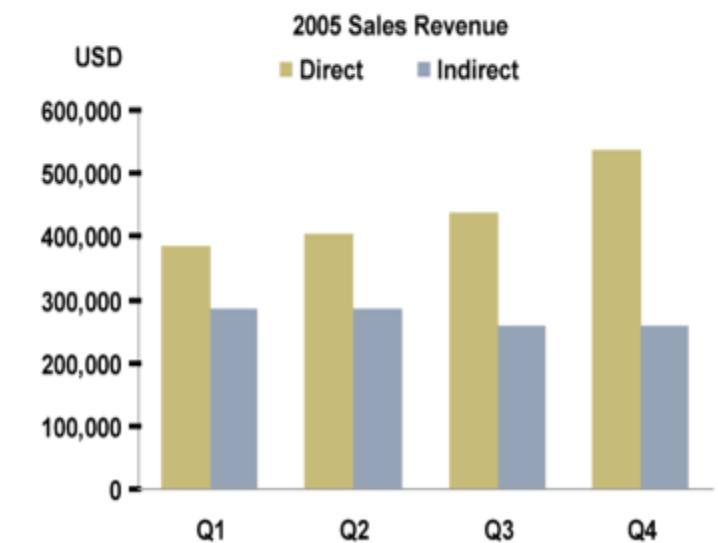
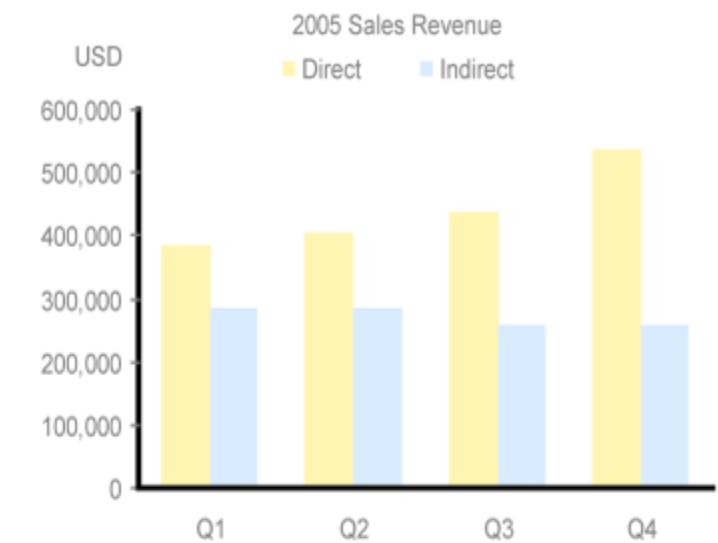
**tritanope**  
**(blue cone cells defective)**

# Design principles for effective visual presentation

- Guides for Encoding Values in Graph
- JunkCharts
- Practical Guides for Using Colour in Charts
- Data-ink

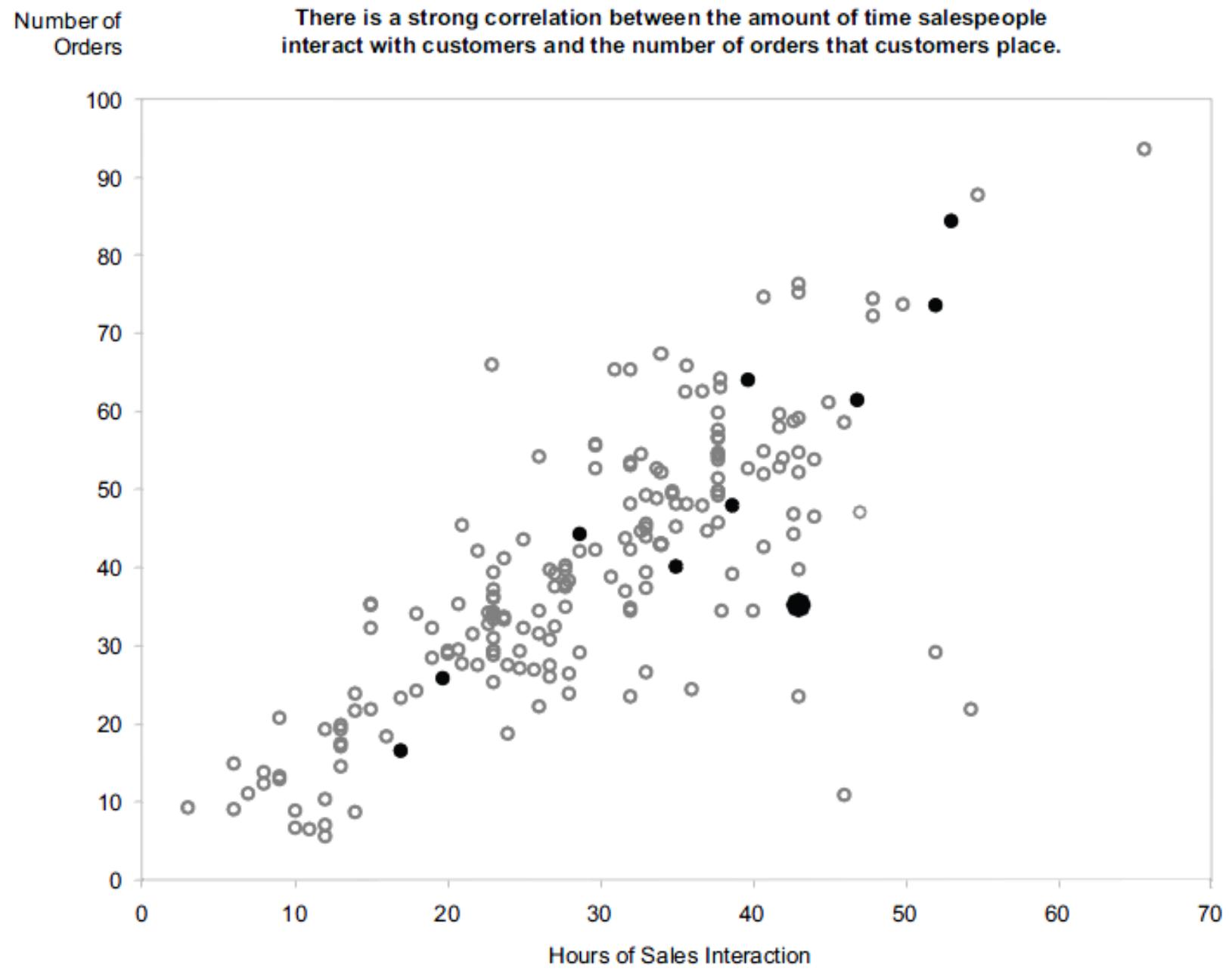
# Data-ink

- Reduce the non data-ink
  - Removed unnecessary non data-ink
  - De-emphasise or regularise the remaining non data-ink
- Enhance the data-ink
  - Remove unnecessary data-ink
  - Emphasise the remaining data-ink



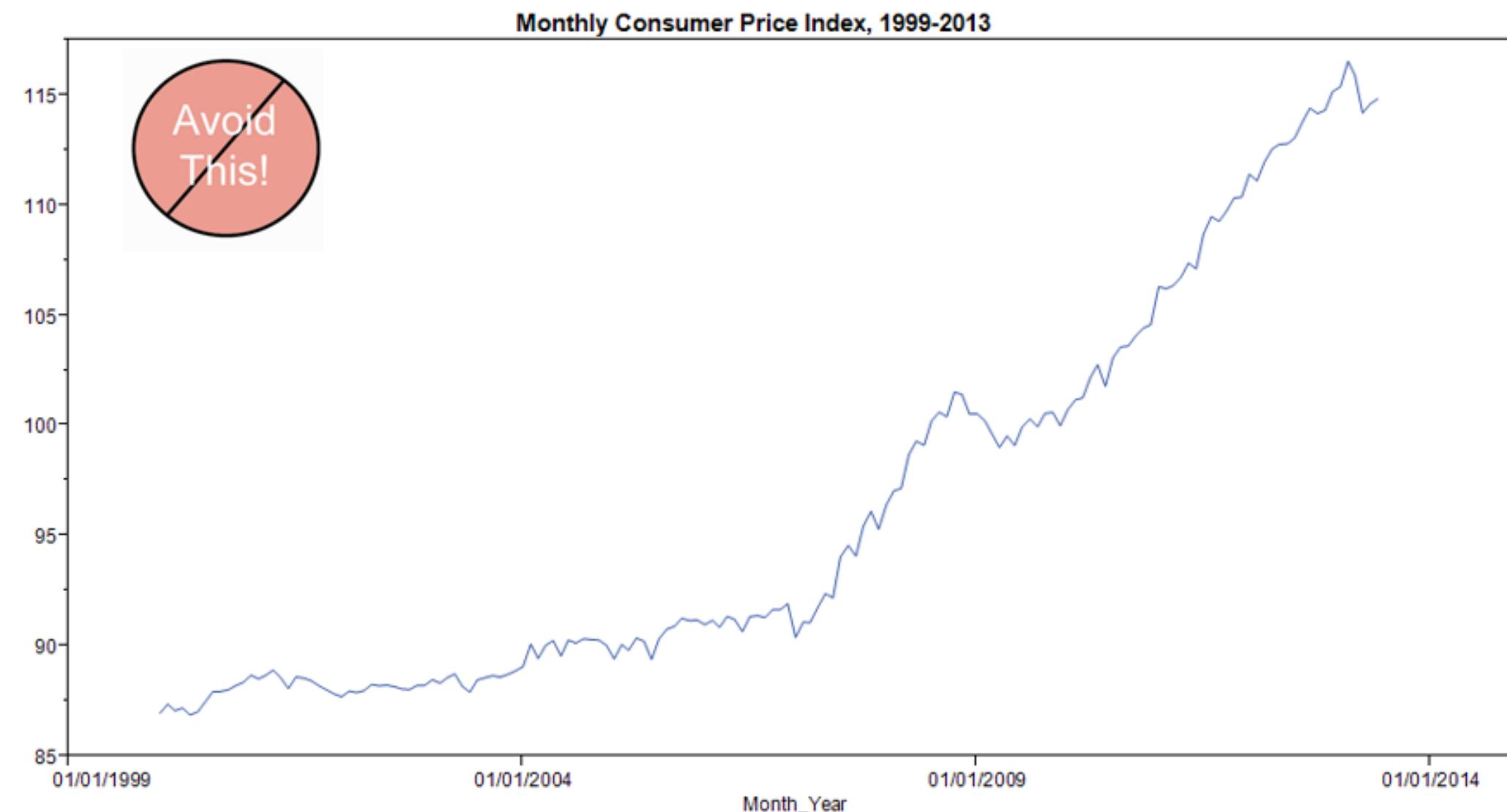
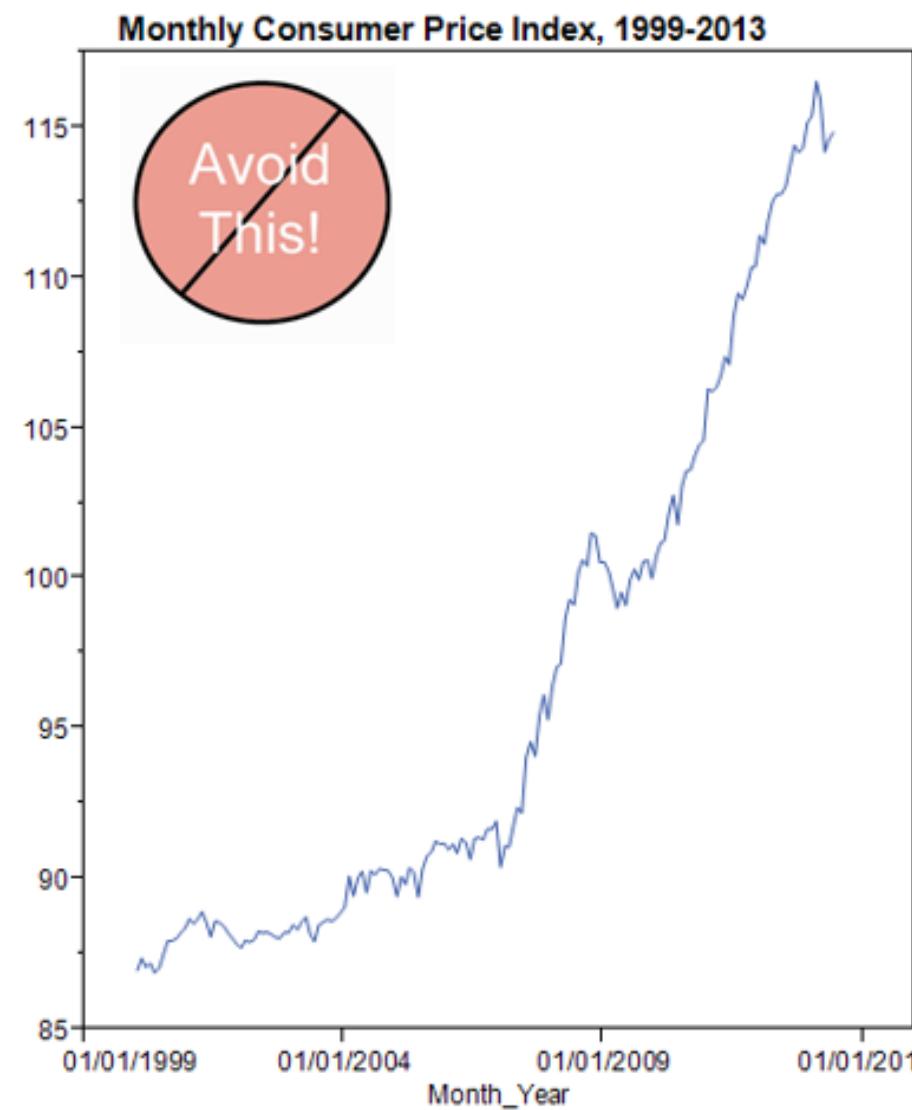
# Practical used of data-ink

## Shouting to emphasize what's interesting



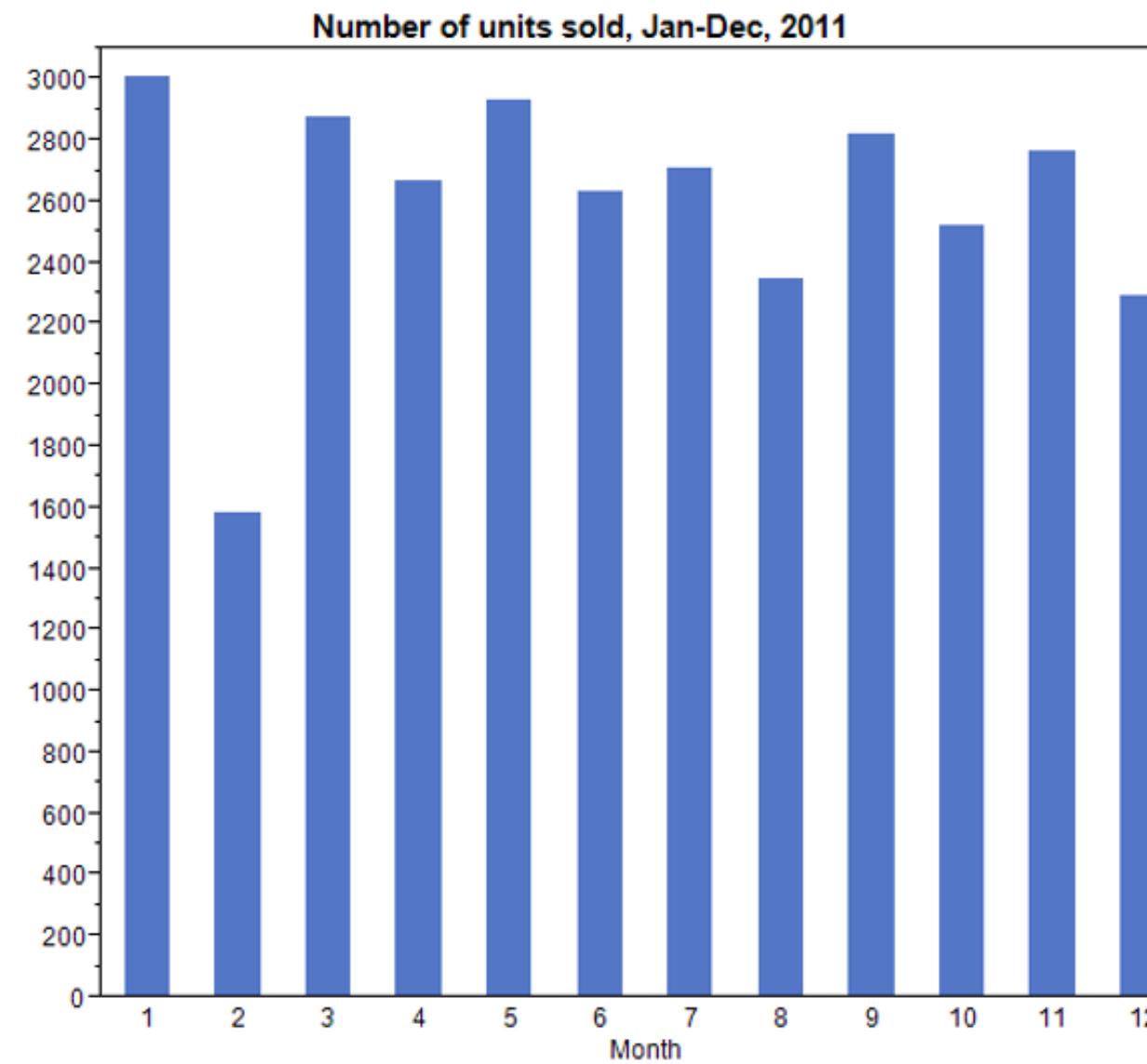
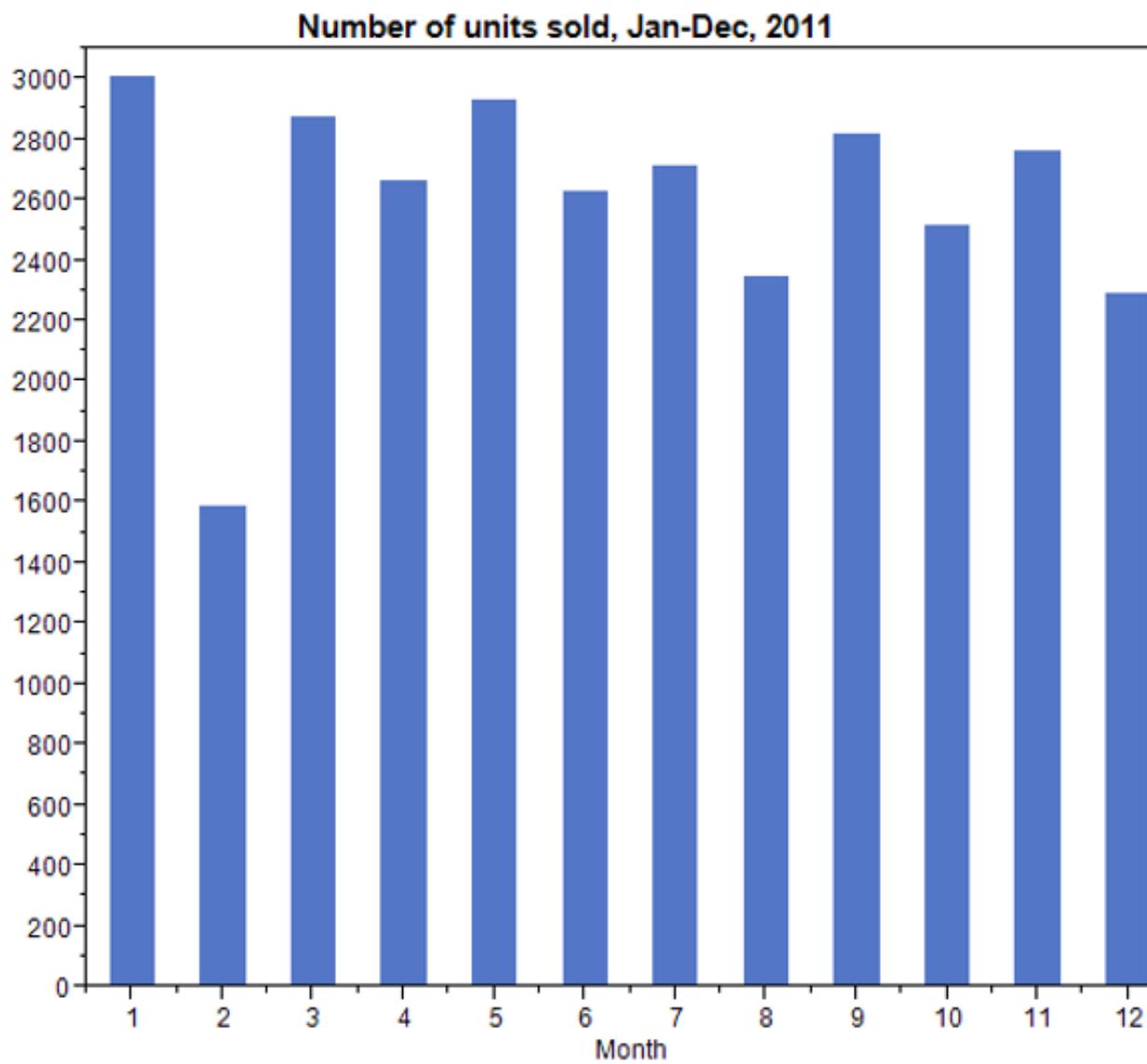
# What should the relative lengths of the axis be?

- Should not manipulate the aspect ratio to intentionally exaggerate or downplay the rate of change.
- Stick to the convention of making your graphs wider than being tall.



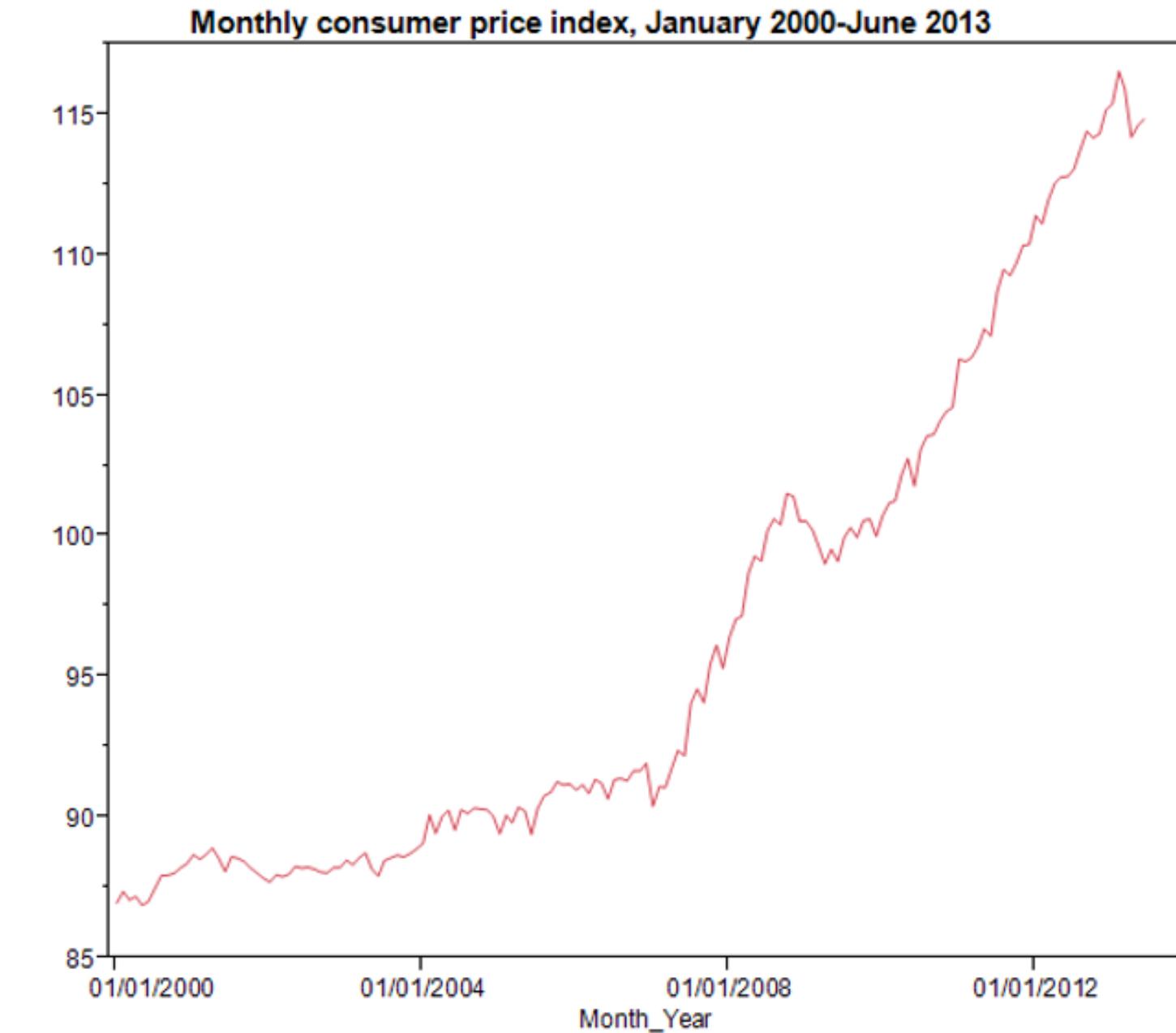
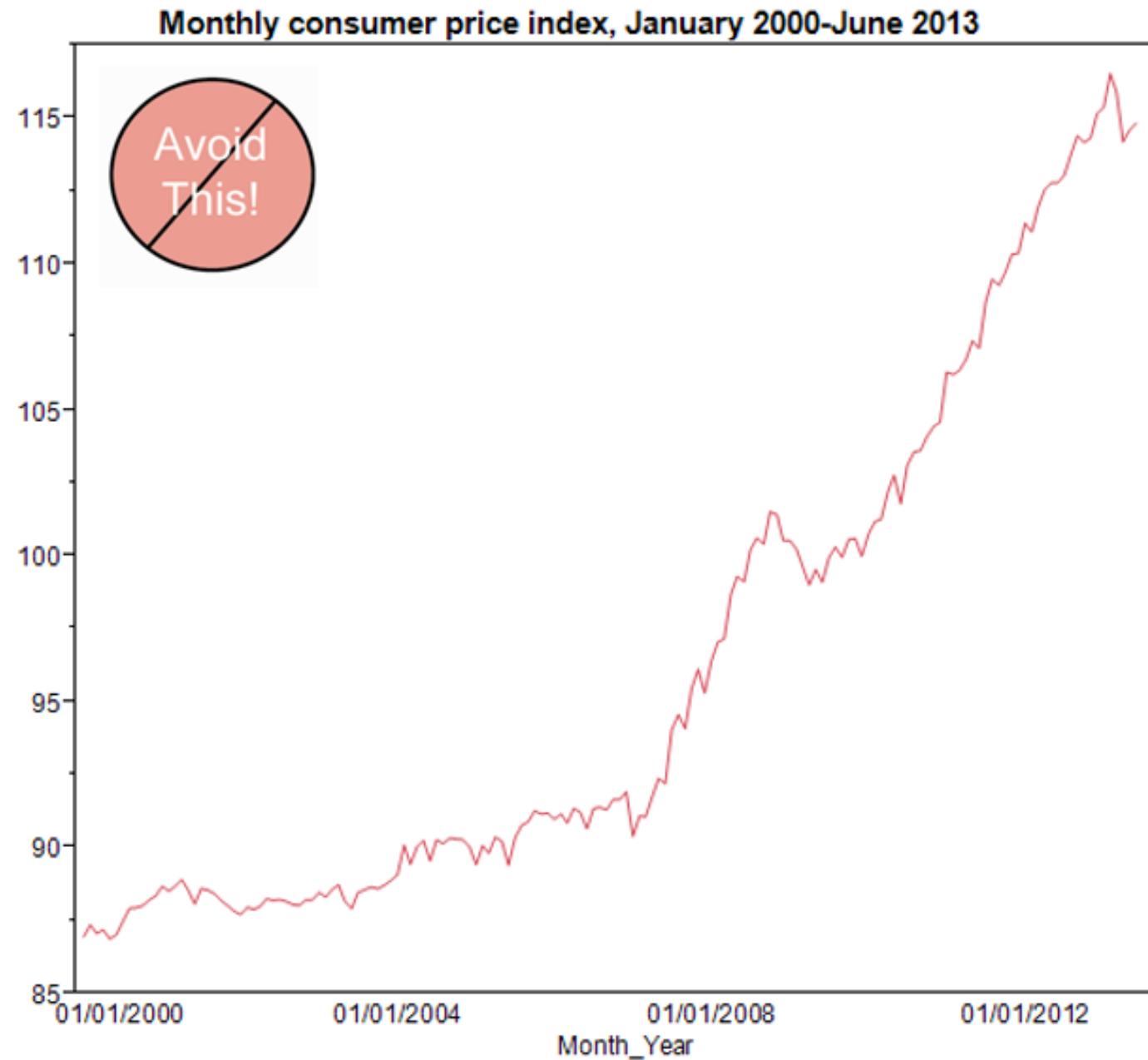
# When can you eliminate tick mark?

Tick marks are superfluous on categorical scale.



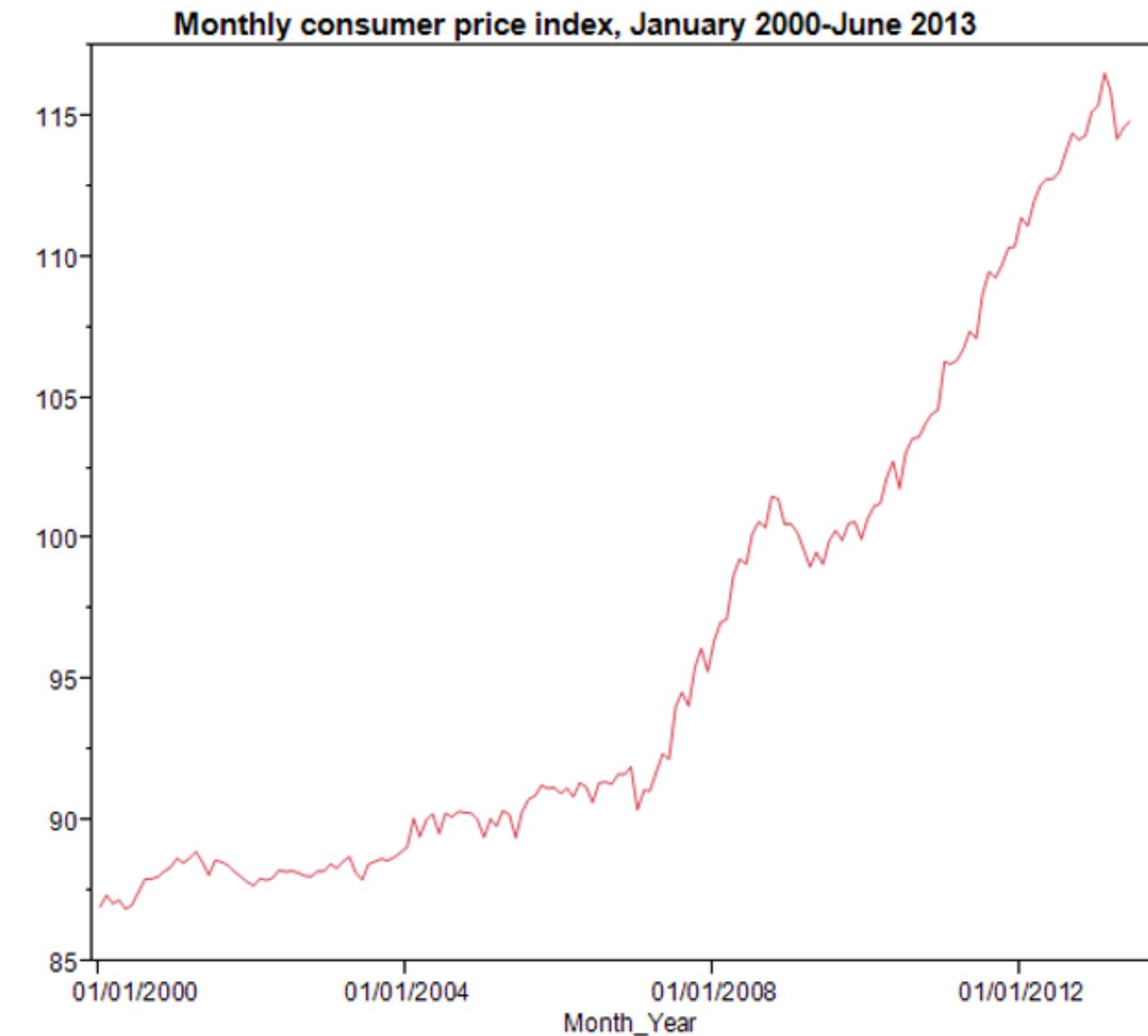
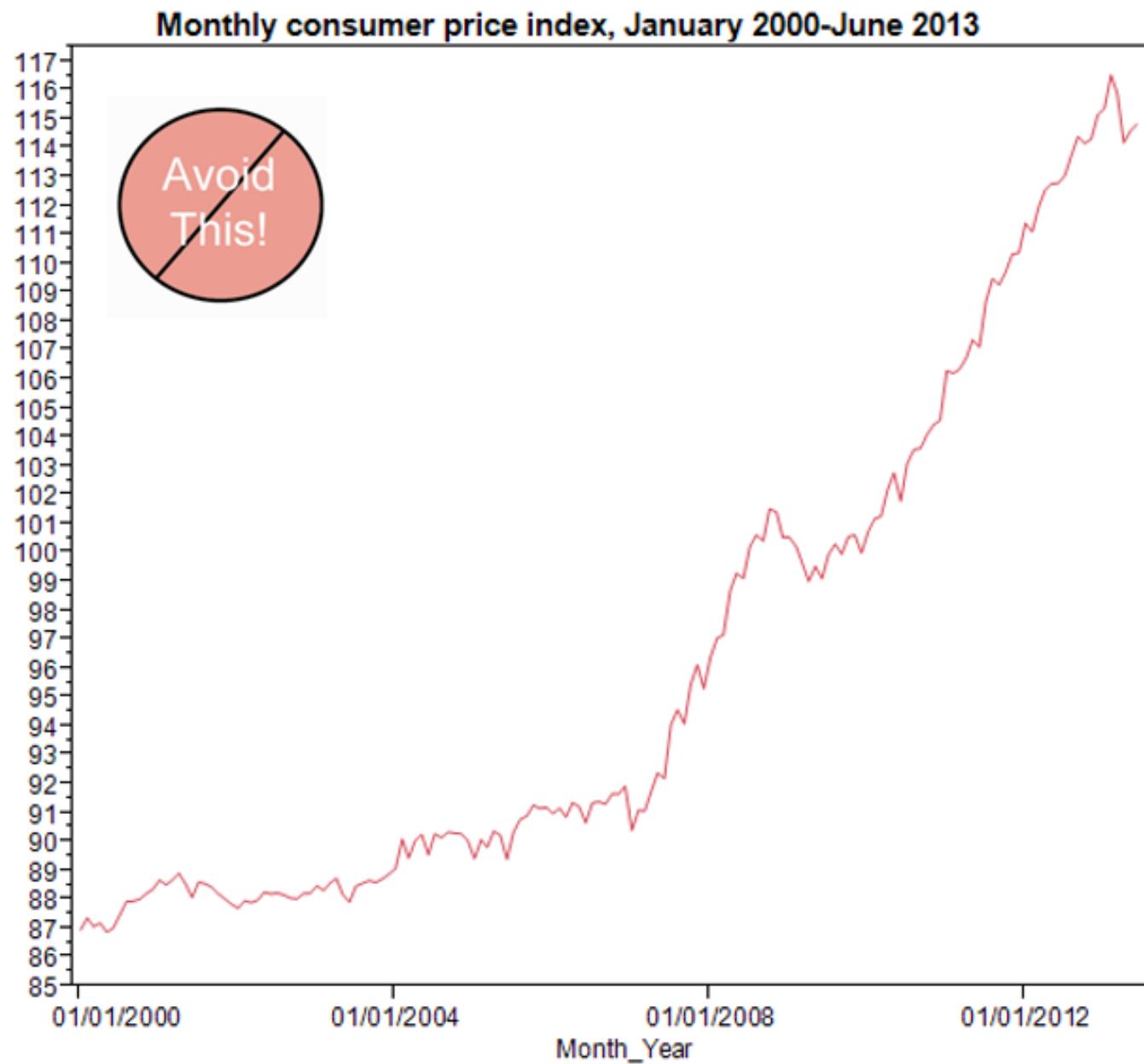
# When you shouldn't eliminate tick mark?

Tick marks are necessary on continuous scale.



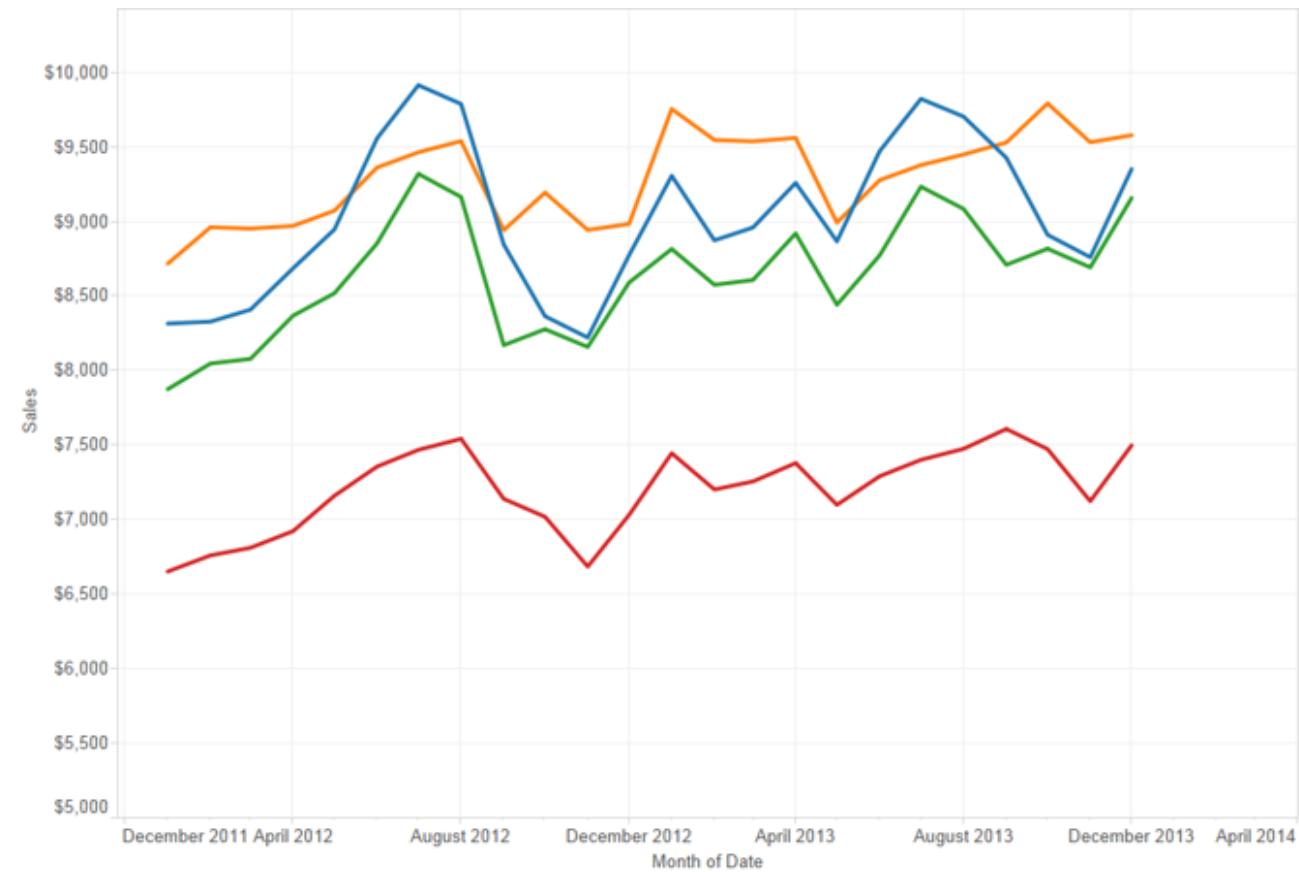
# How many tick marks should you use?

- There is no exact number that works best in all circumstances, and the size of the graph is a factor that must be considered: the longer the scale line, the more tick marks it should contain.

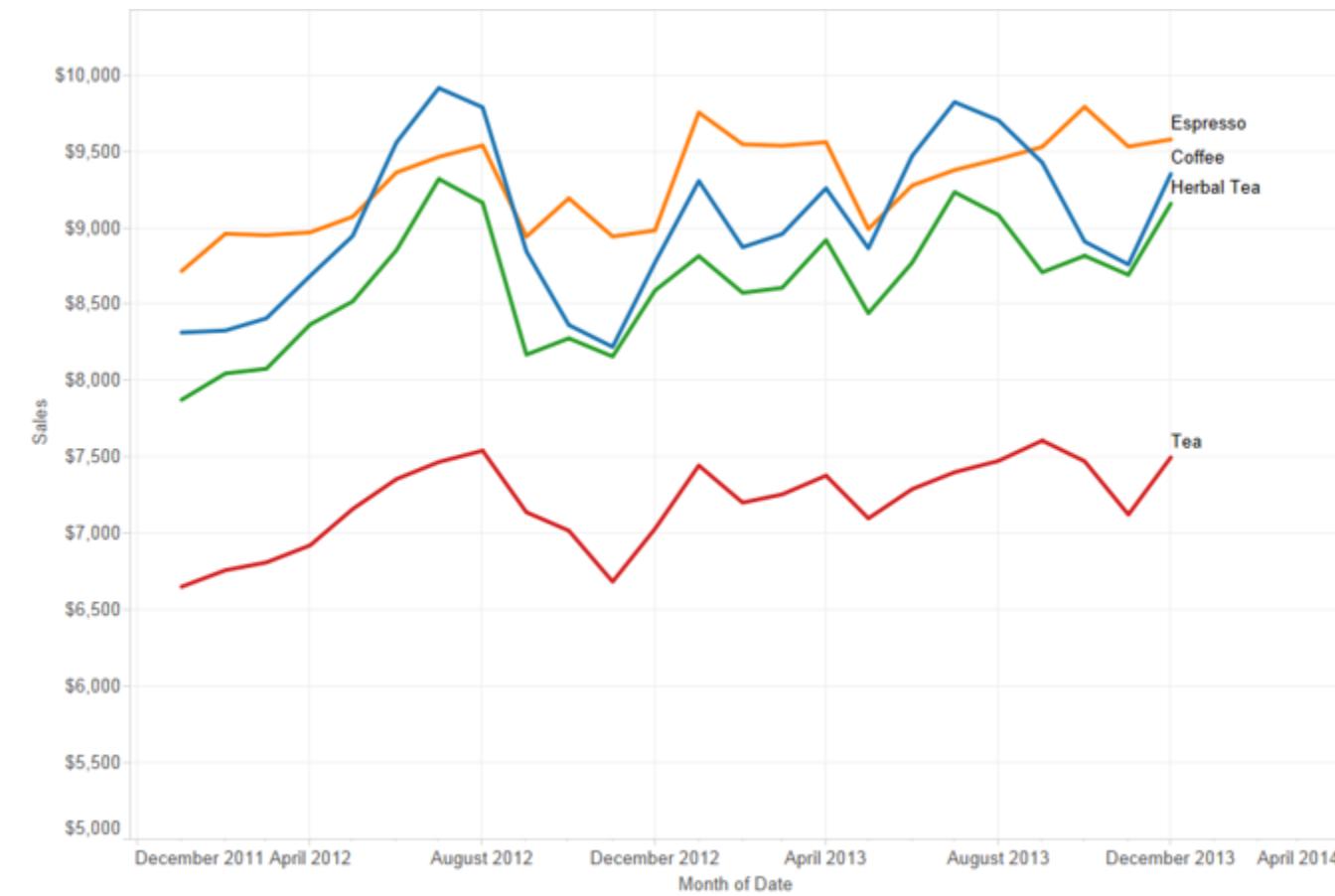


# When can you eliminate legends?

- In this graph, a legend is used to indicate product types.



In this graph, product types are labeled directly.



# An enlightening data visualisation will be incomplete without a well worded title



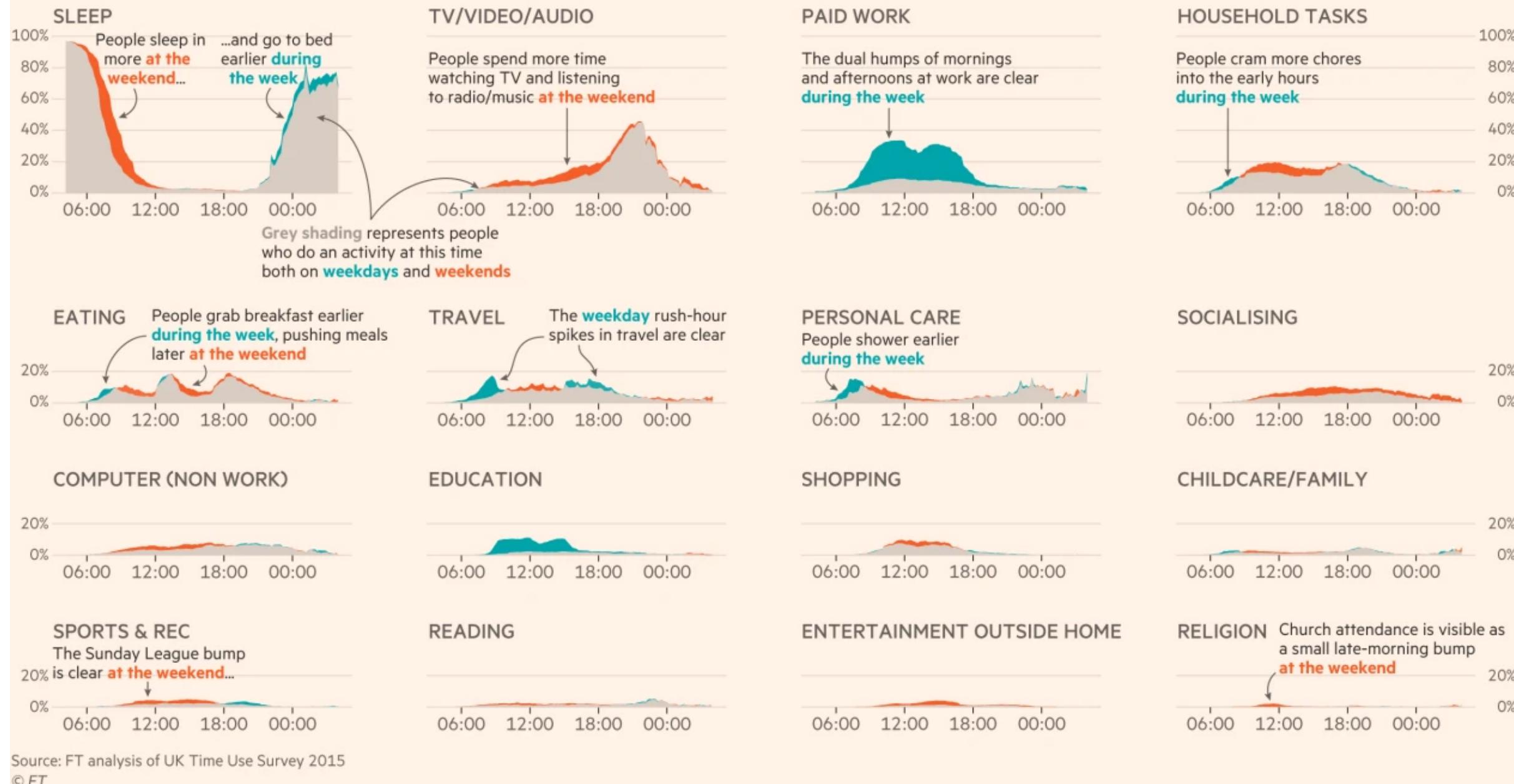
Source: [The truth about weekend working](#), Financial Times, January 23 2020.

- More example can be found [here](#)

# Using Annotation to tell data story

## How Britons spend their time at weekends vs weekdays

Share of people doing specific activities during **weekends** vs **weekdays**, by time of day (%)



Source: [The truth about weekend working](#), Financial Times, January 23 2020.

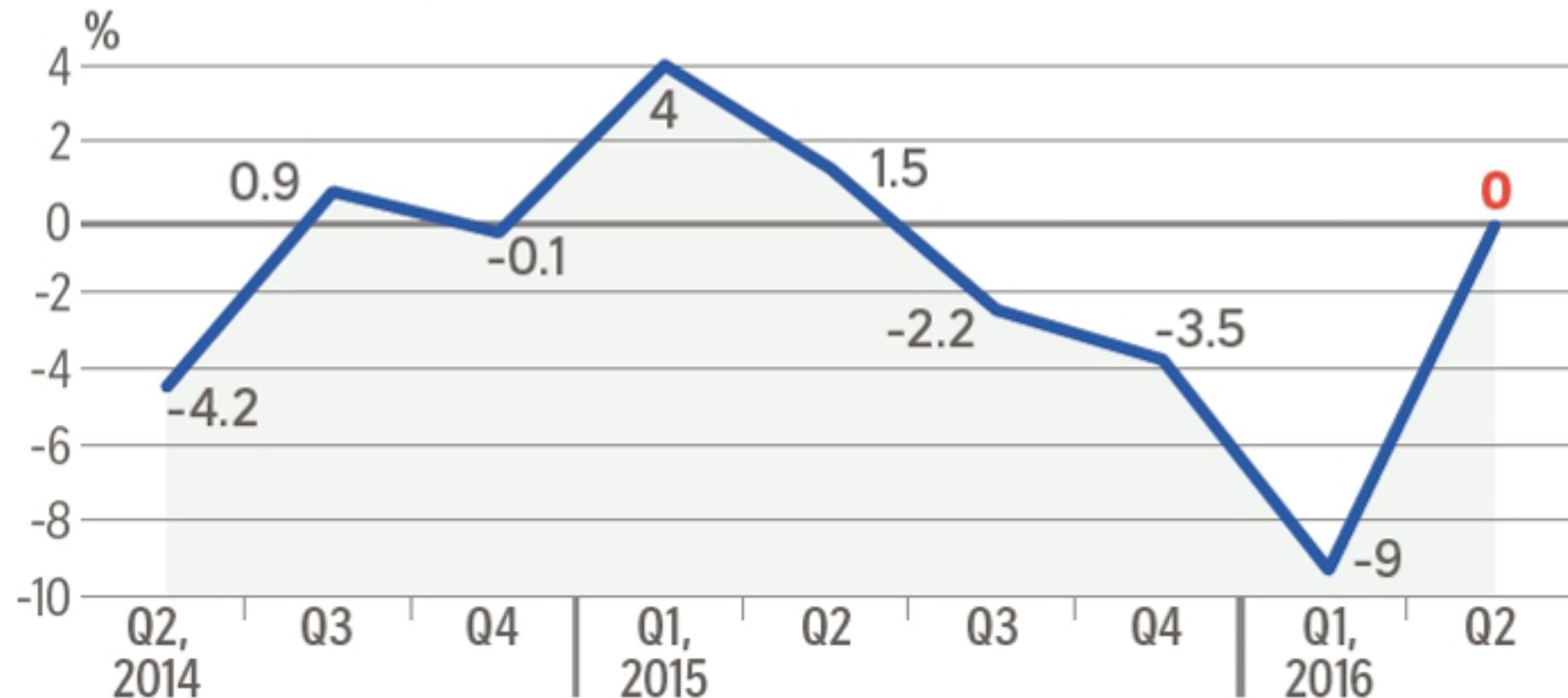
# Graphical Integrity: Show Me the Truth

Don't lie to yourself (or to others) with charts



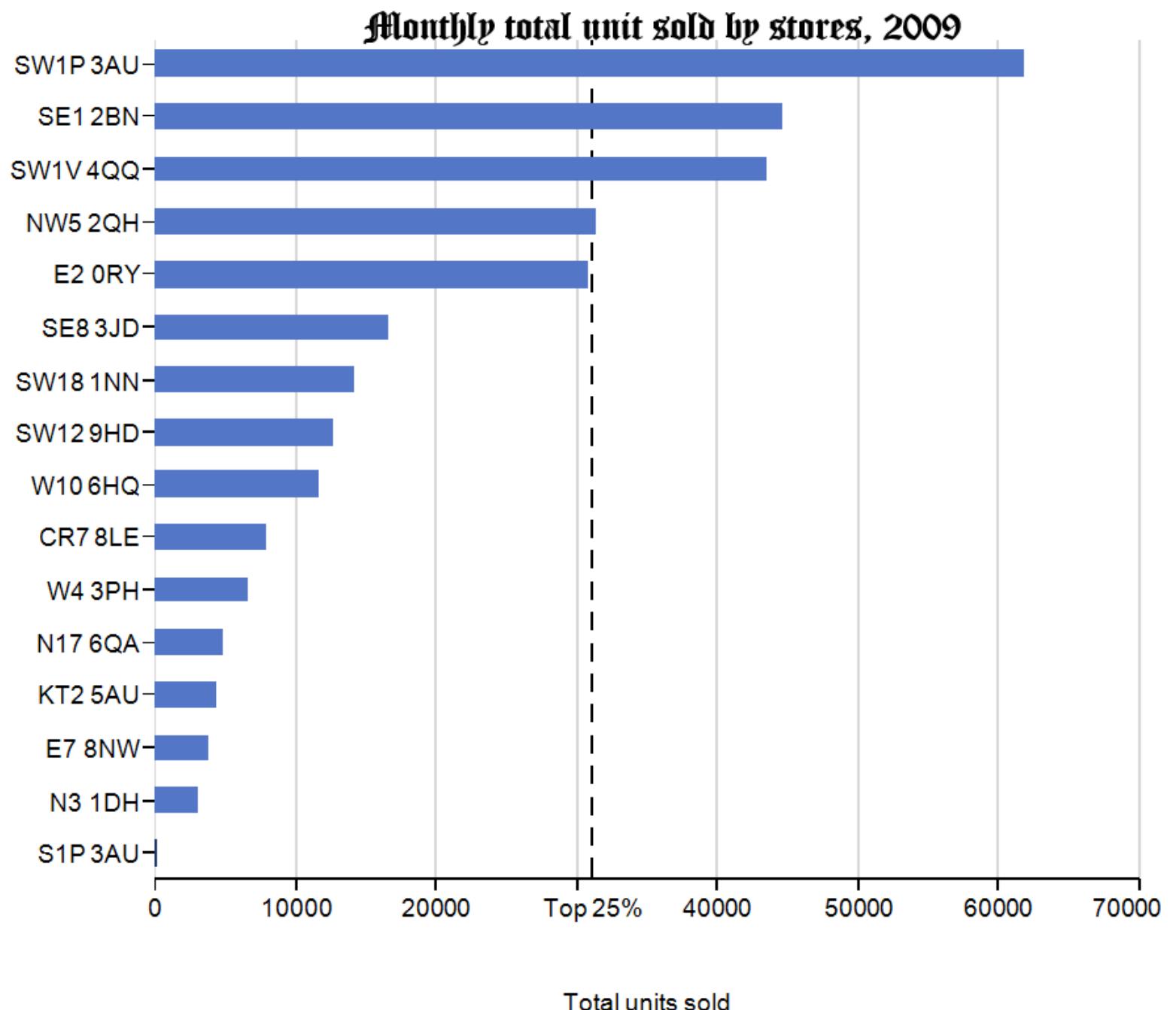
Non-oil domestic exports (Nodx) growth was flat in the second quarter. For the first half this year, **Nodx fell 4.5 per cent** compared with the same period last year.

## Stuttering Nodx performance



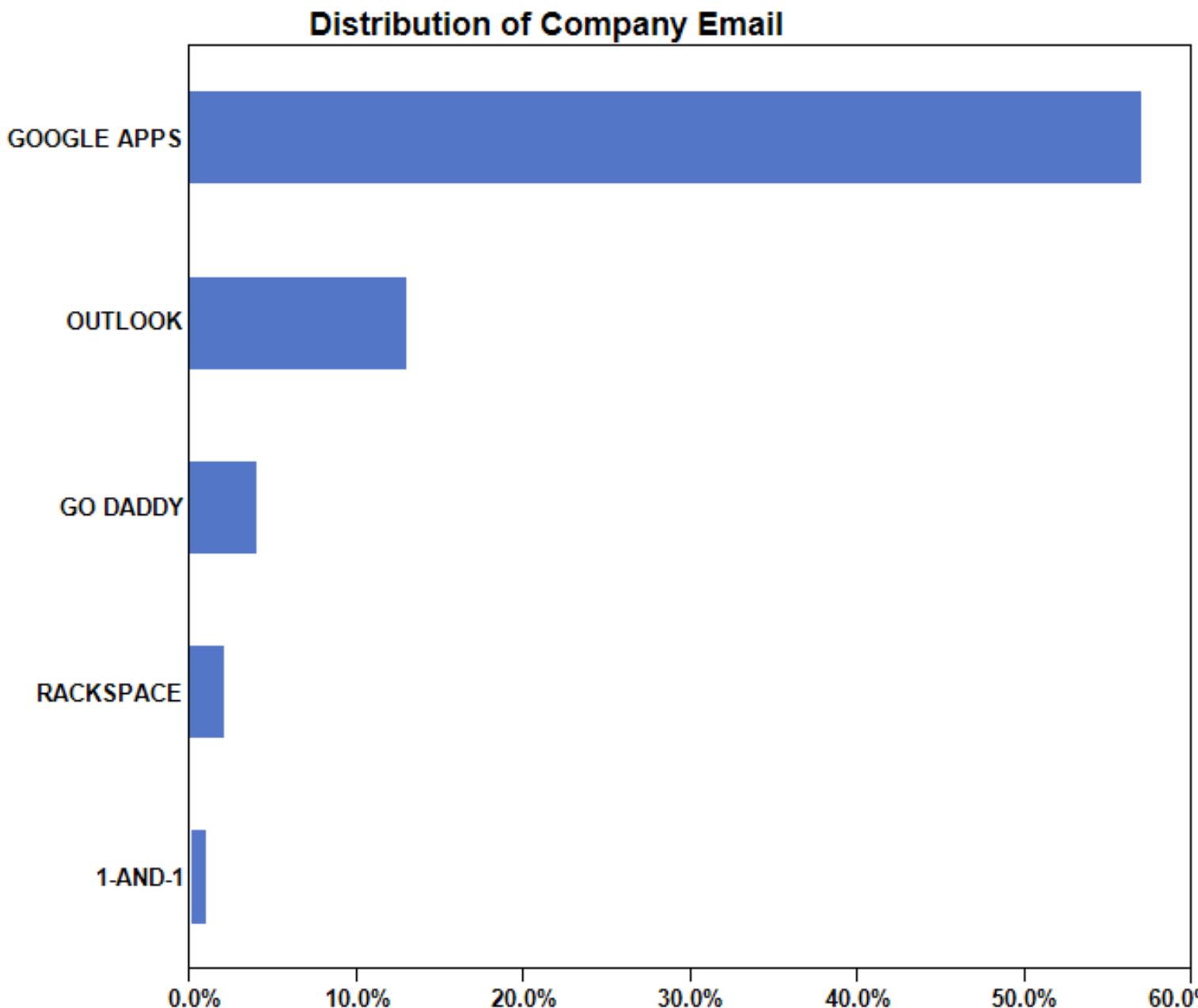
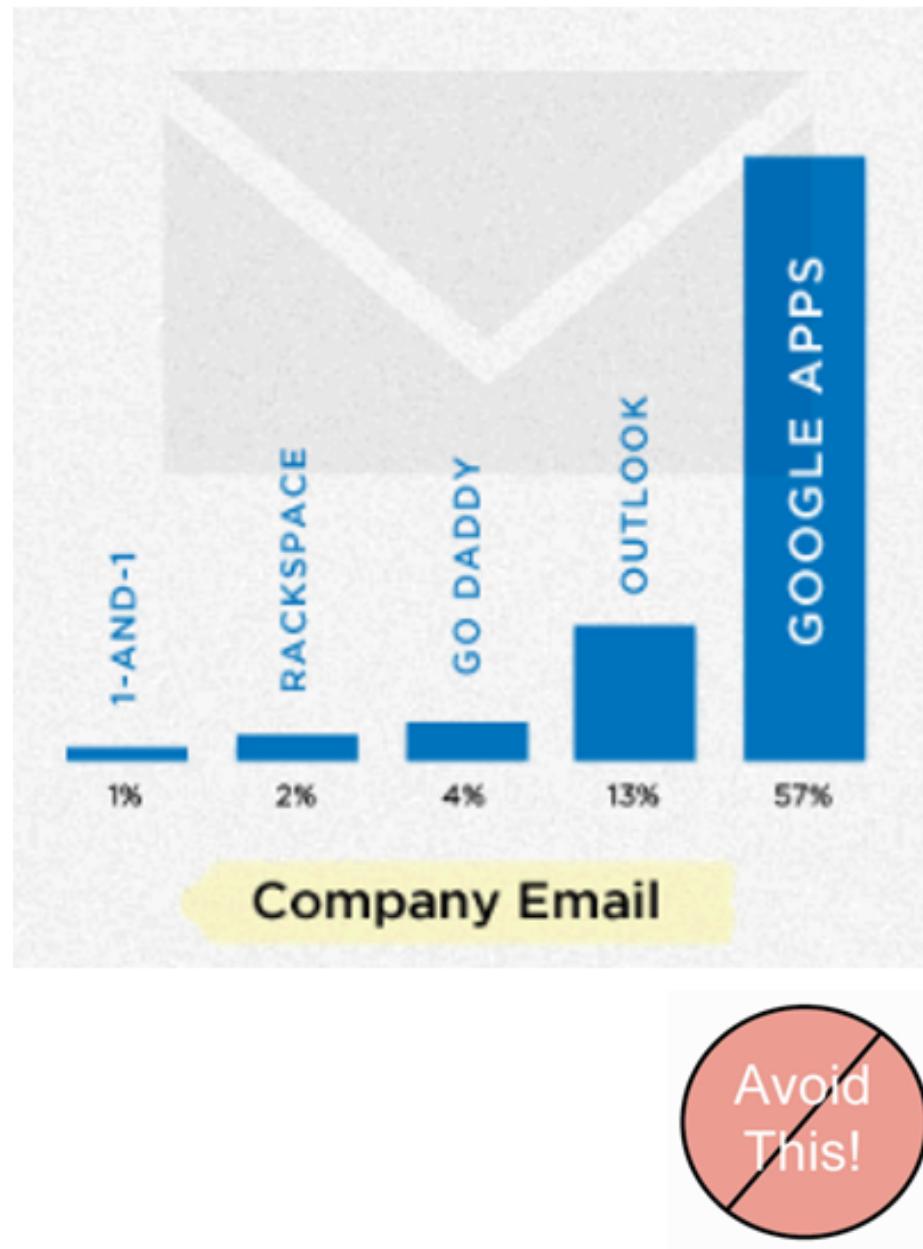
# Graph typography

## Avoid using artistic fonts



# Graph Labeling

Orientation of label should be reader friendly

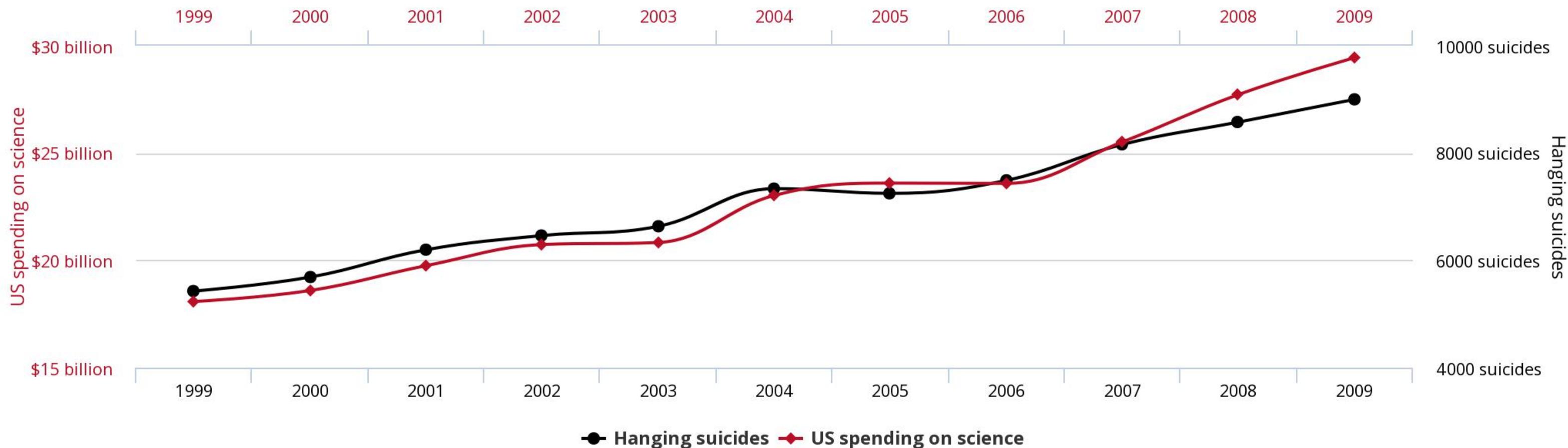


# Three Bugs of Charts Interpretation

- The **Patternicity** bug: We detect interating patterns, regardless of whether or not they are real.
- The **Storytelling** bug: We immediately come up with a coherent explanation for those patterns.
- The **Confirmation** bug: We start seeing all further information we receive, even the one that conflicts with our explanation, in a way that confirmed it. We refuse to give our explanation up, no matter what.

# Three Bugs example

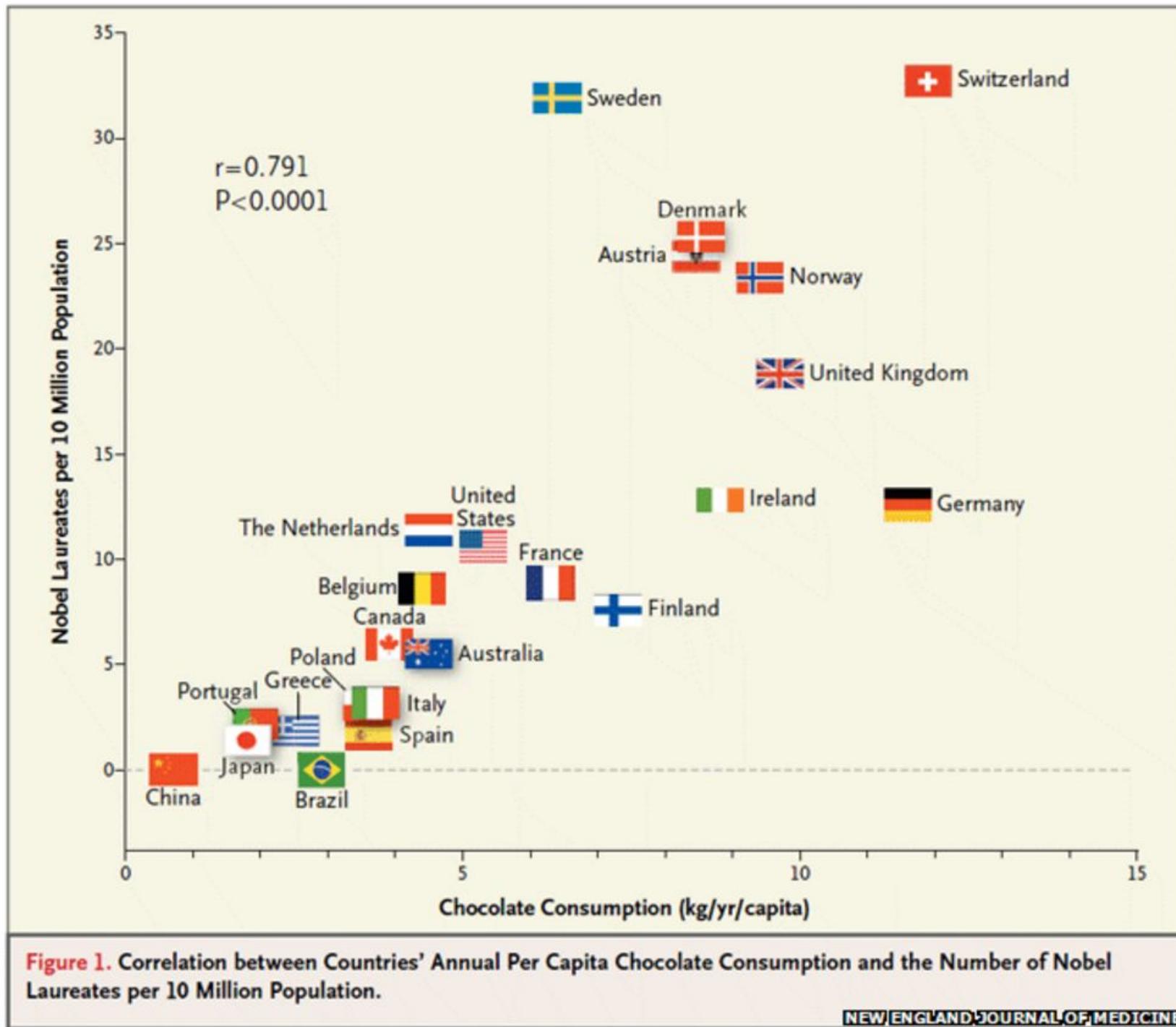
**US spending on science, space, and technology**  
correlates with  
**Suicides by hanging, strangulation and suffocation**



tylervigen.com

Source: [Spurious Correlation](#)

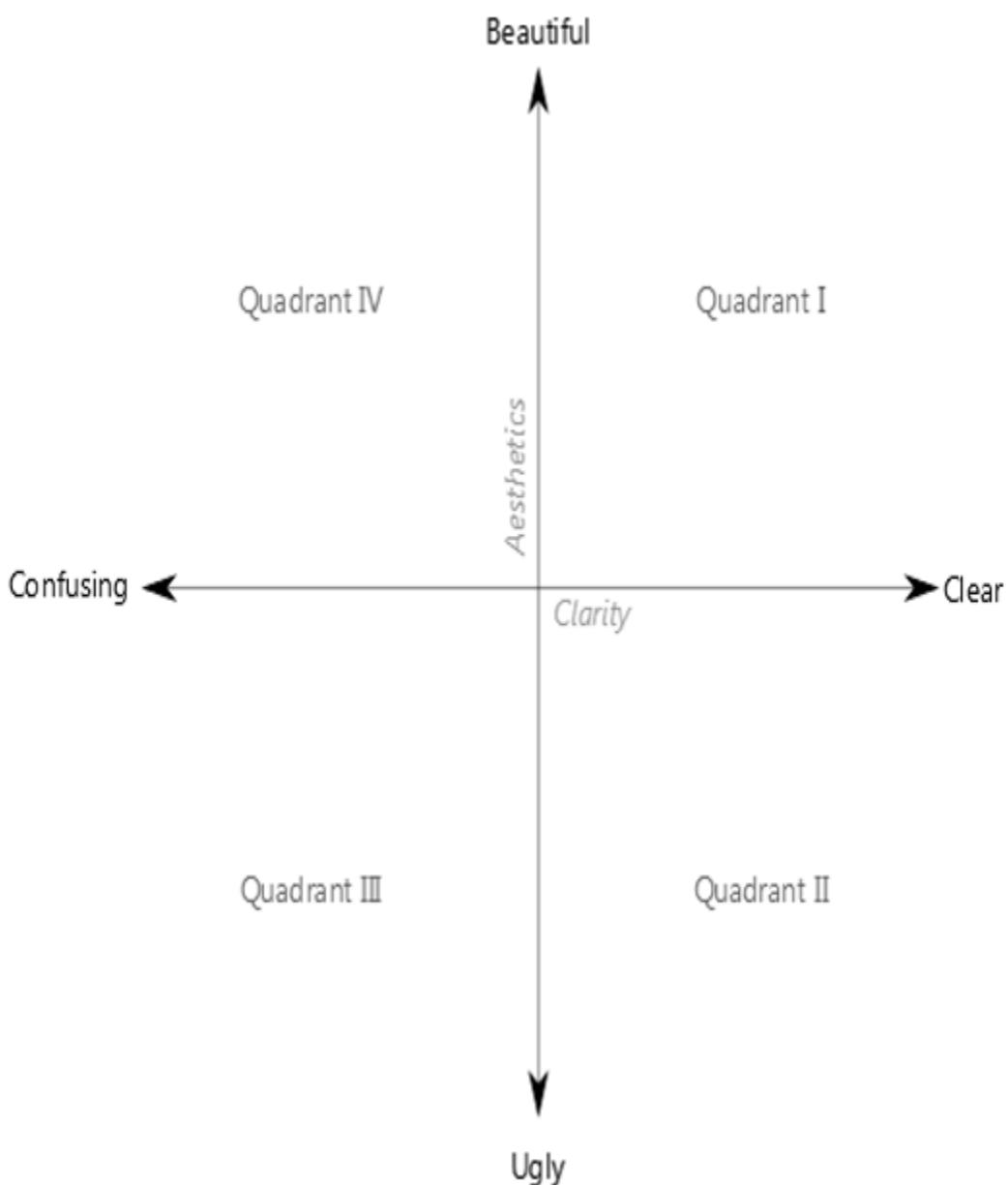
# Stop the Fallacy of Visual Storytelling



Source: Franz H. Messerli (2012) [Chocolate Consumption, Cognitive Function, and Nobel Laureates](#), *The New England Journal of Medicine*.

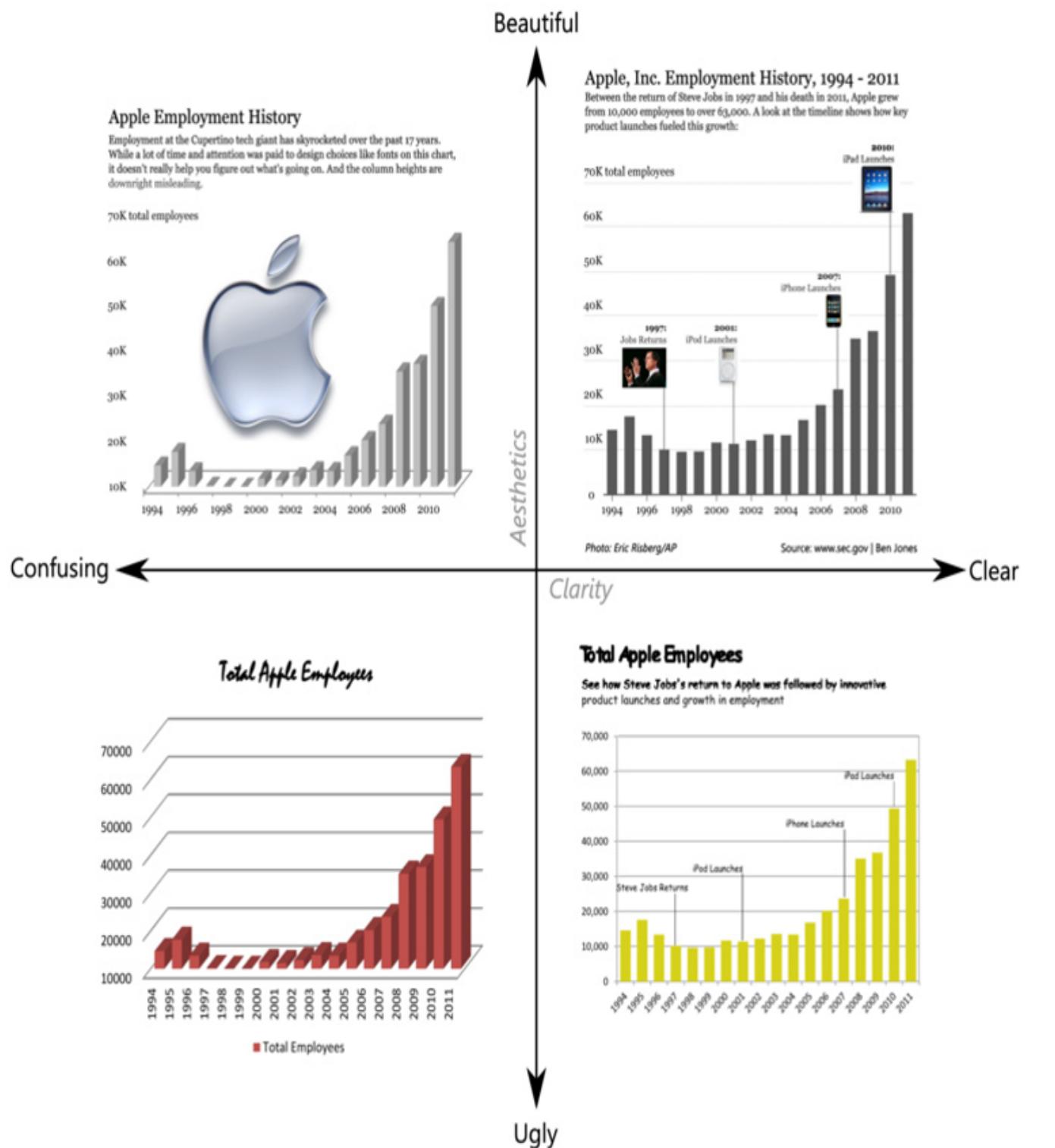
# Data Visualization: Clarity or Aesthetics?

Mapping data visualizations on a Cartesian coordinate system where "clarity" is placed along the horizontal (x) axis and "aesthetics" is placed along the vertical (y) axis provides a framework to gage the objective and subjective merits of a graphic:



Source: [Data Visualisation: Clarity or Aesthetics](#)

# A Tale of Four Quadrants

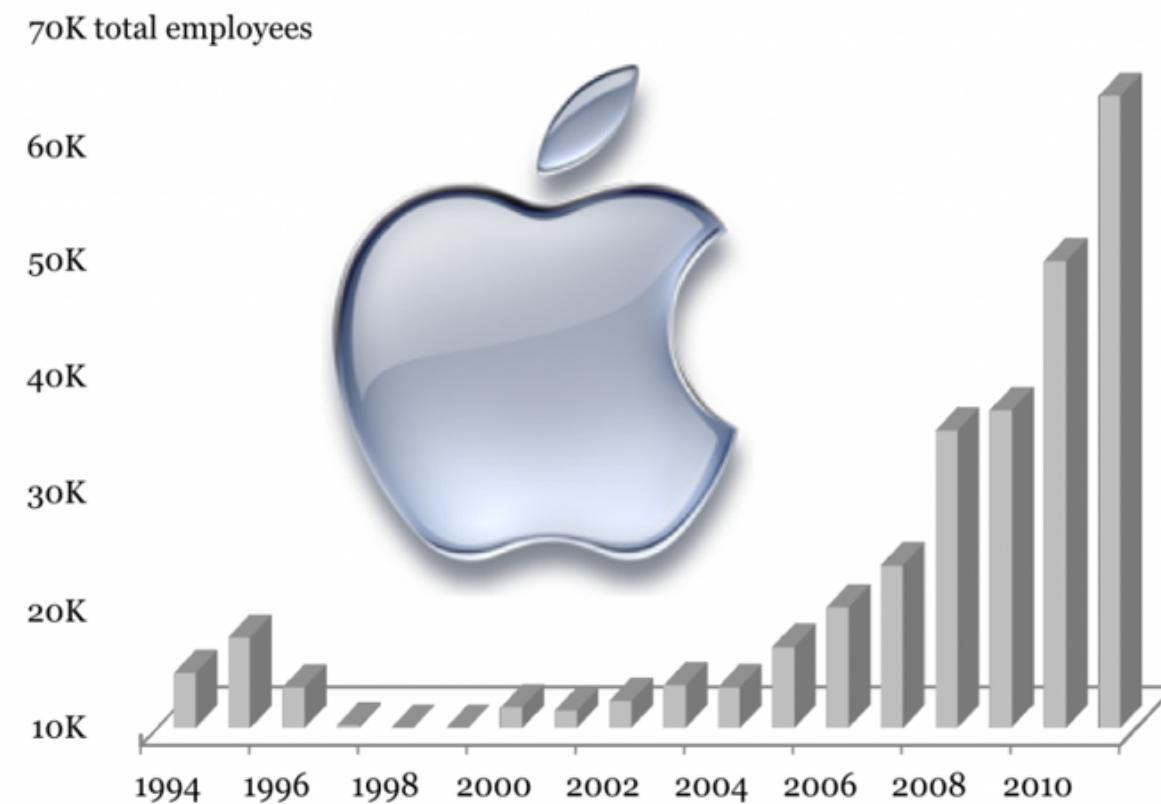


# Quadrant IV – Confusing yet Beautiful

- Why is it “beautiful”?
  - Well placed & aligned title & lead-in.
  - Attention to detail with font selection.
  - Inclusion of image.
- Why is it “confusing”?
  - Y-axis starts at 10K (column height misleading).
  - 3D effect makes it difficult to gage heights.
  - Title & lead-in aren’t helpful.

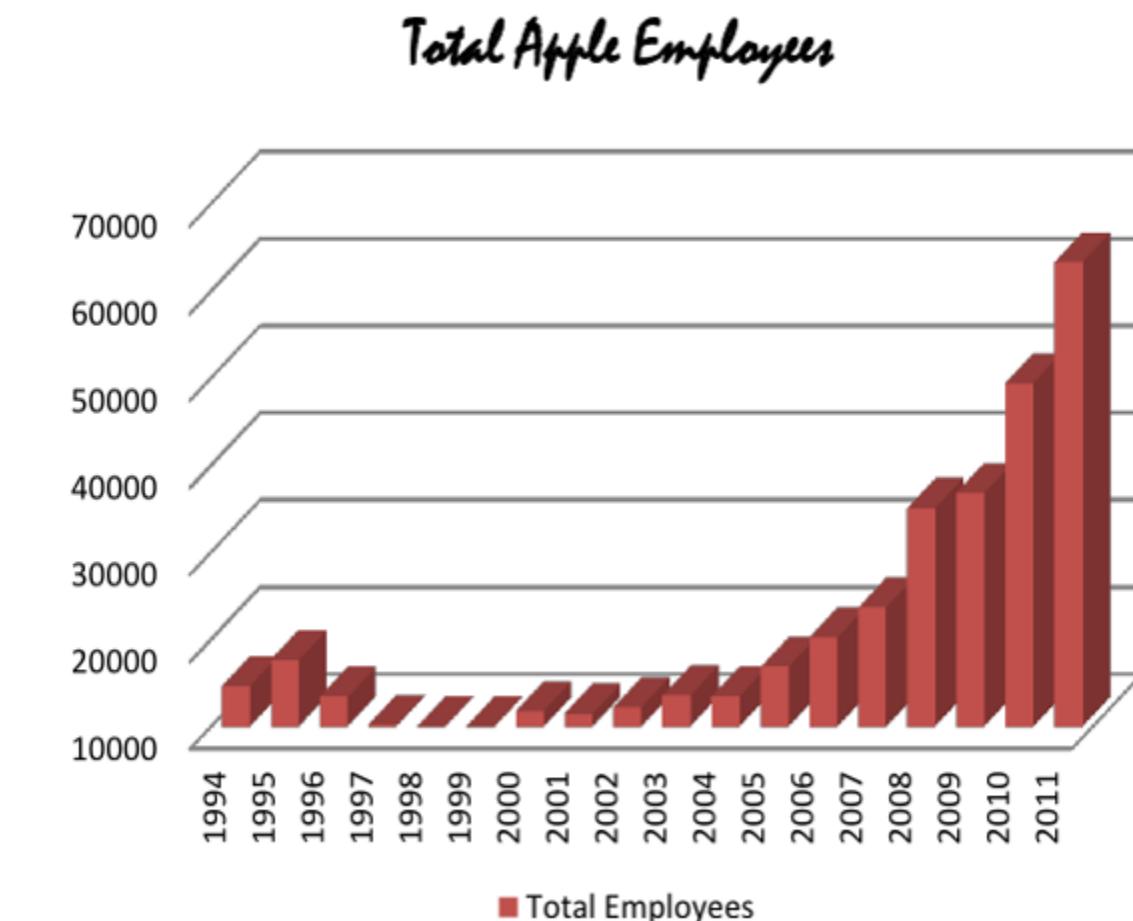
## Apple Employment History

Employment at the Cupertino tech giant has skyrocketed over the past 17 years. While a lot of time and attention was paid to design choices like fonts on this chart, it doesn't really help you figure out what's going on. And the column heights are downright misleading



# Quadrant III – Confusing and Ugly

- Why is it “ugly”?
  - Horrible font & color choice.
  - Grid lines are too dark & distracting.
  - Format of axes (vertical x-axis labels, number format of y-axis).
- Why is it “confusing”?
  - Y-axis starts at 10K (column height misleading).
  - 3D effect makes it difficult to gage heights.
  - No lead-in or call-outs to provide context.

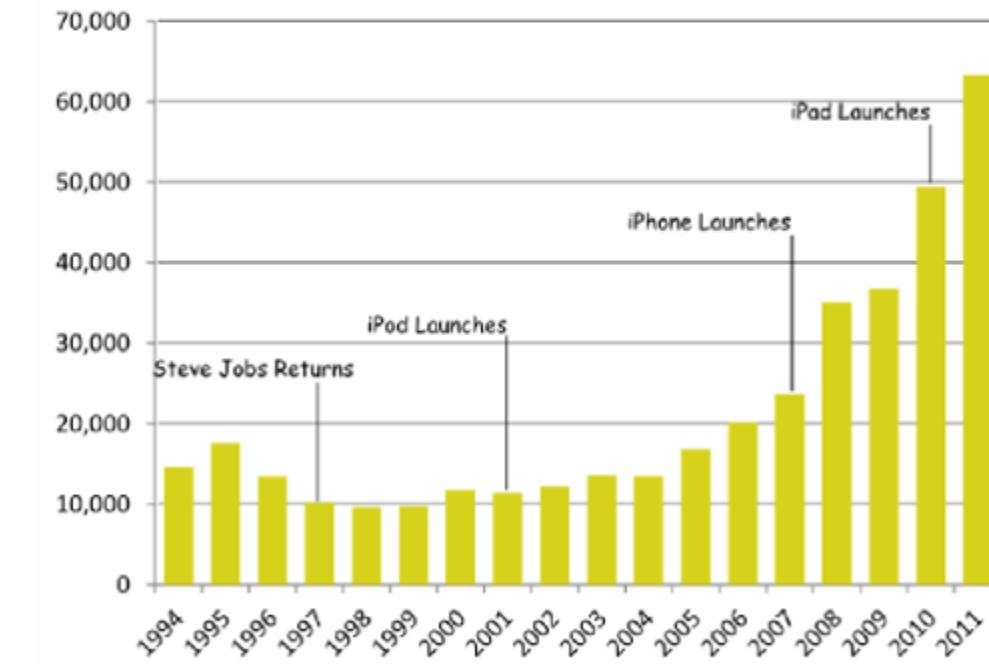


# Quadrant II – Clear but Ugly

- Why is it “ugly”?
  - Poor color (puke yellow?) and font (Comic Sans?) choices.
  - Slightly pixelated – poor attention to image quality detail.
  - Chart details – axis orientation, grid lines, outline.
- Why is it “clear”?
  - The y-axis starts at 0 and the 2D columns are easy to gage.
  - For the first time, we see call-outs of relevant events on the timeline.
  - This time the lead-in paragraph is actually informative.

**Total Apple Employees**

See how Steve Jobs's return to Apple was followed by innovative product launches and growth in employment



# Quadrant I – Clear and Beautiful

- Why is it “beautiful”?
  - Good font & color choices throughout.
  - Soft gridlines don’t distract.
  - All elements well aligned and spaced.
  - High res images are “useful” chartjunk.
- Why is it “clear”?
  - The y-axis starts at 0 and the 2D columns are easy to gage.
  - Call-outs with images aid cognition.
  - Improved title & lead-in verbiage provide further elucidation.
  - For the first time, a photo credit and data source are included.

## Apple, Inc. Employment History, 1994 - 2011

Between the return of Steve Jobs in 1997 and his death in 2011, Apple grew from 10,000 employees to over 63,000. A look at the timeline shows how key product launches fueled this growth:

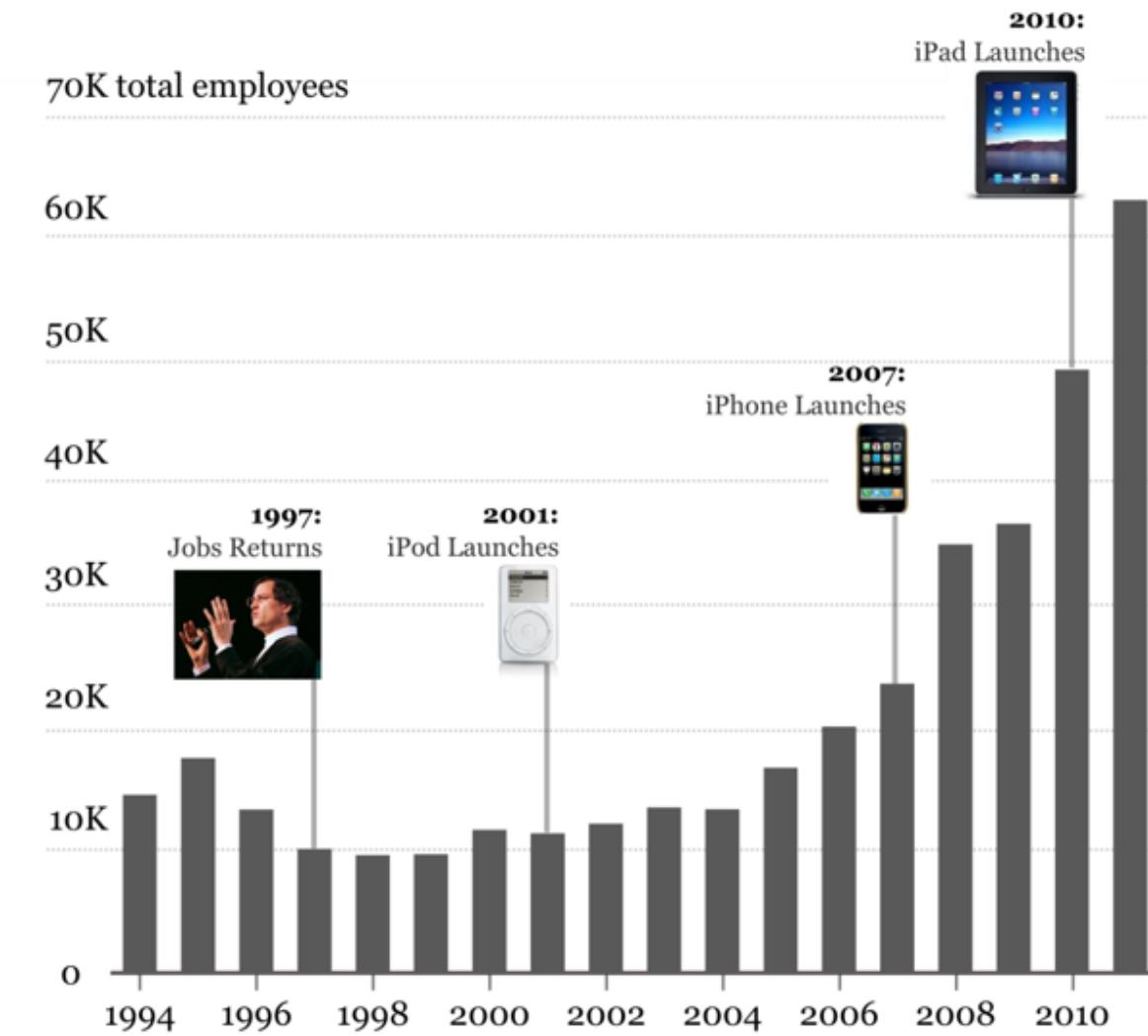


Photo: Eric Risberg/AP

Source: [www.sec.gov](http://www.sec.gov) | Ben Jones

# References

- Claus O. Wilke (2019) [Fundamentals of Data Visualization](#). O'Reilly, USA.
- Few, Stephen (2012) (2nd edition) **Show Me the Numbers: Designing Tables and Graphs to Enlighten**, Analytics Press, Oakland, USA
- Cairo, Alberto (2019) **How Charts Lie**, W.W. Norton & Company, USA.
- Robbins, Naomi B. (2005) **Creating More Effective Graphs**, John Wiley & Sons, New Jersey, USA
- Wong, Dona M. (2010) **The Wall Street Journal Guide to Information Graphics**, W. W. Norton & Company, Inc. New York.
- Tufte, Edward (2nd Edition) **The Visual Display of Quantitative Information**, Graphics Press LLC, Connecticut, USA.

# Highly recommended blog

- [The Functional Art](#)
- [Junk Charts](#)
- [Perceptual Edge](#)
- [EagerEyes](#)
- [Statistical Graphics and more](#)
- [Visualizing data](#)
- [Visualizing Economics](#)

