01

# Course Overview

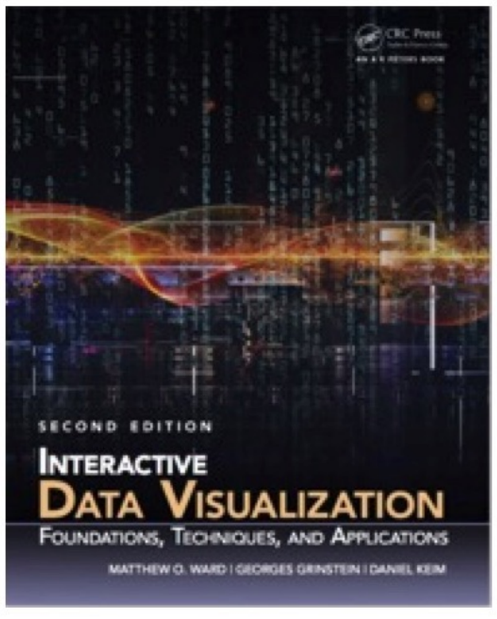
# Notice

- **Author**
  - **João Moura Pires (jmp@fct.unl.pt)**

# Bibliography

- **Many examples are extracted and adapted from:**


**Visualization Analysis & Design**
Tamara Munzner
2015,
ISBN: 9781466508910
ISBN (e-Book): 9781498707763

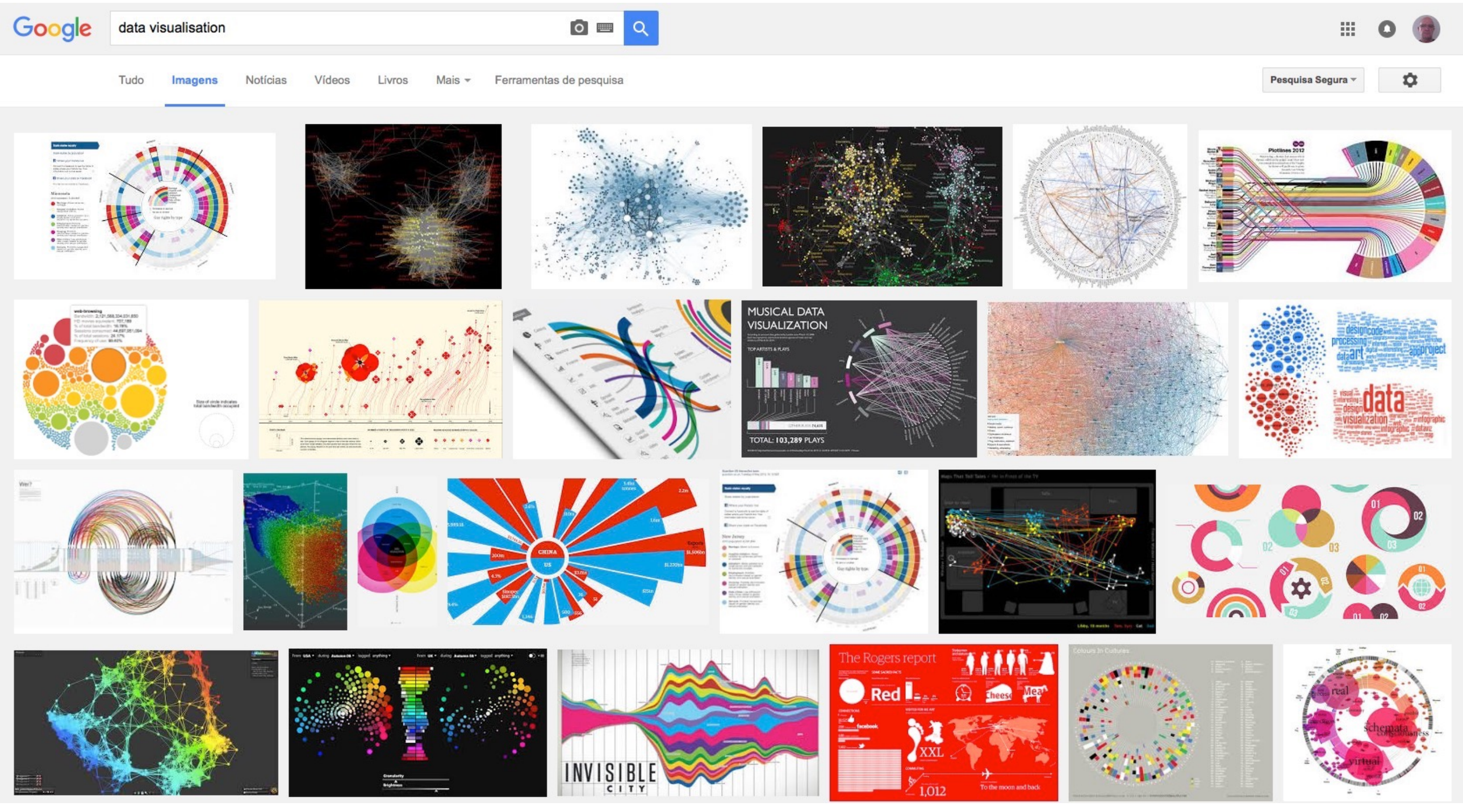



**Interactive Data Visualization: Foundations, Techniques, and Applications**
Matthew O. Ward, Georges Grinstein, Daniel Keim
2015, 2nd Edition
ISBN: 9781482257373
ISBN (e-Book): 9781482257397

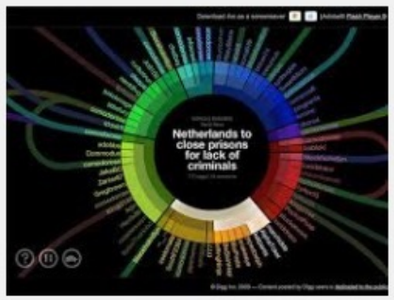
# Table of Contents

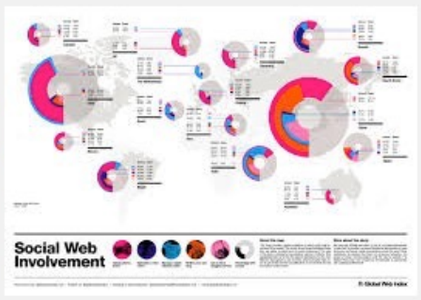
# Ask google for Data Vis (images)

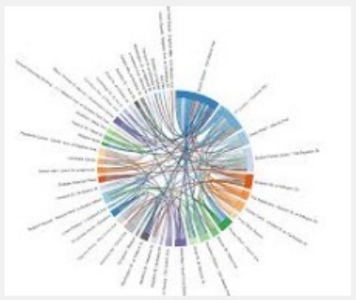
# Ask google for Data Vis (images)



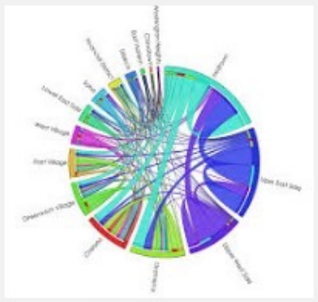
# Ask google for Data Vis (images)

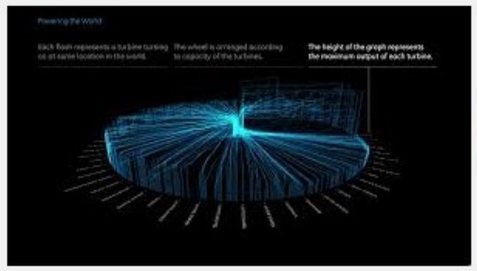



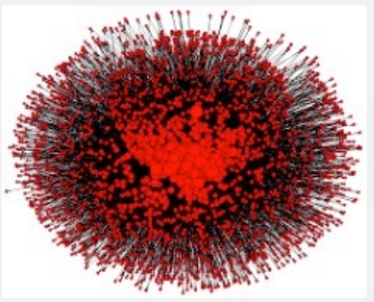
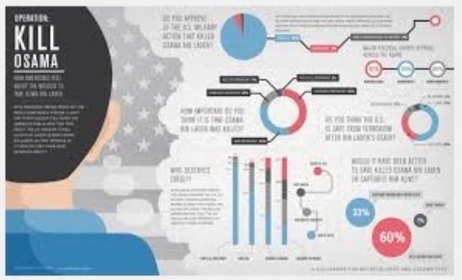


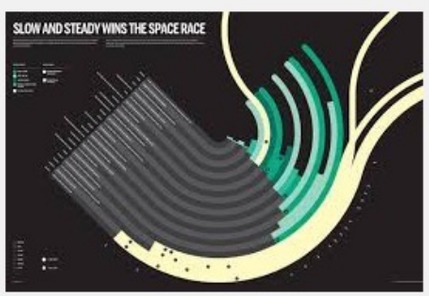


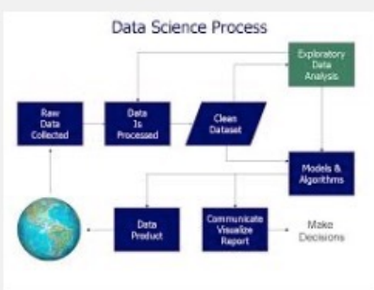




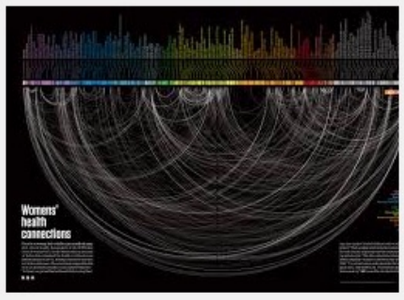



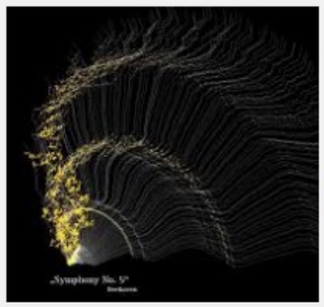




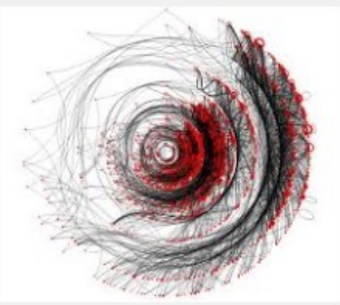
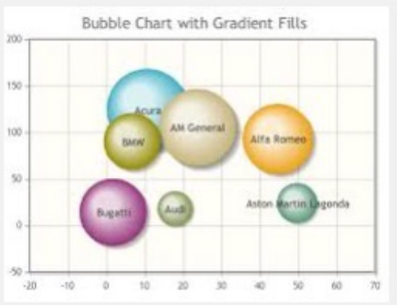



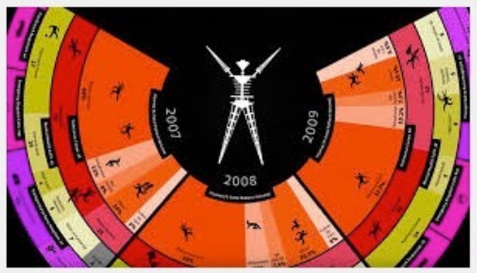


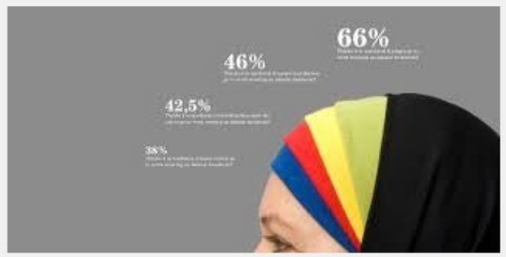


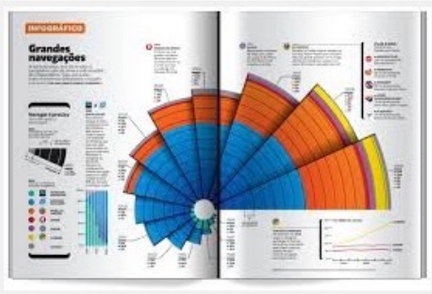


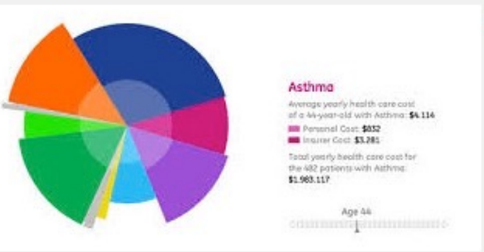



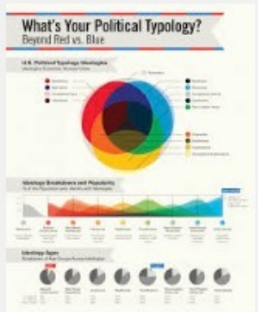
# Ask google for Data Vis (images)

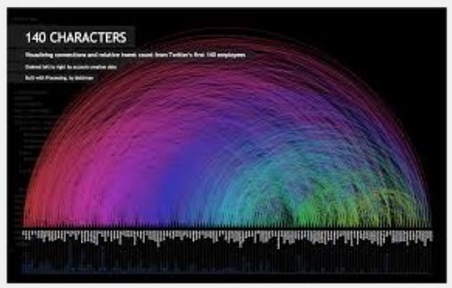



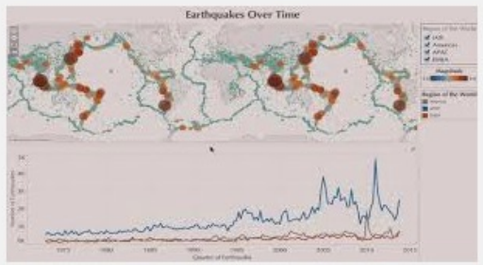



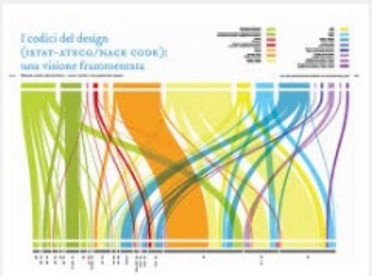


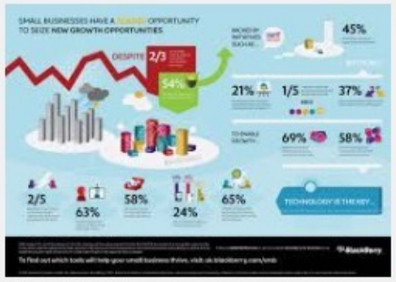


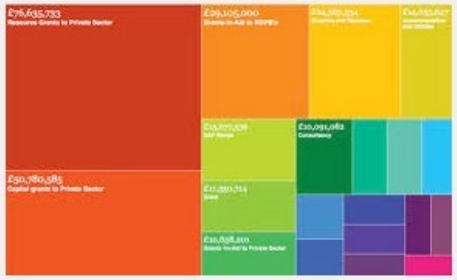



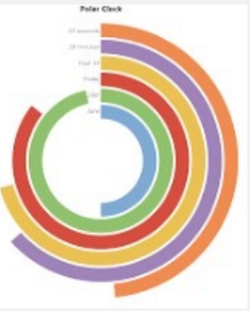


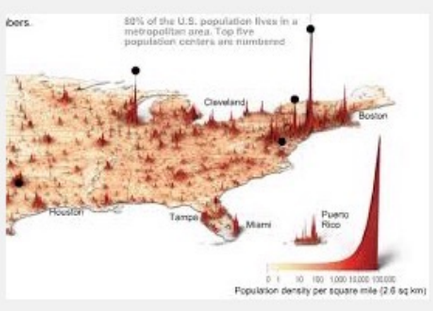


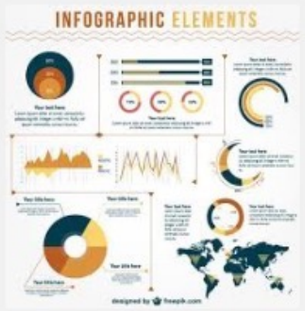



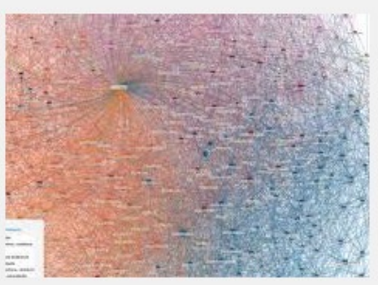
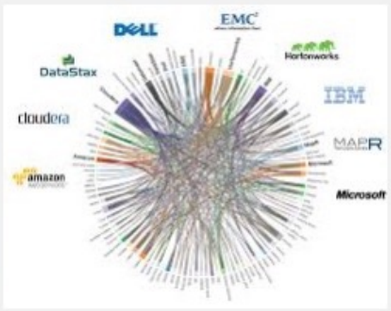




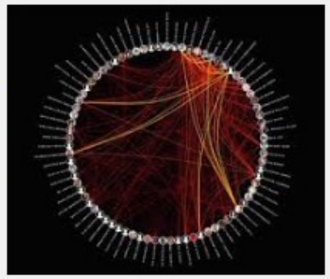
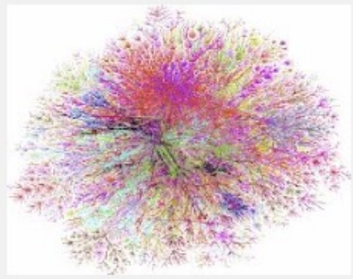
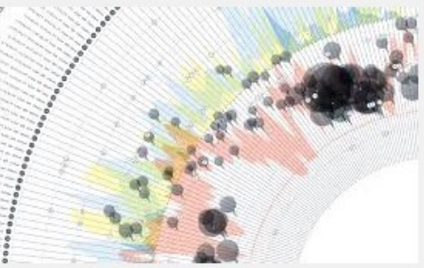
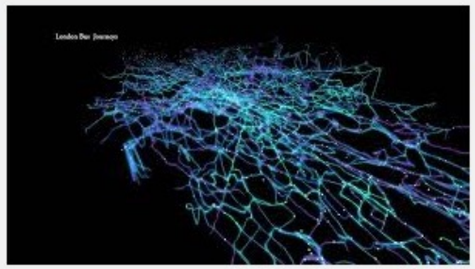


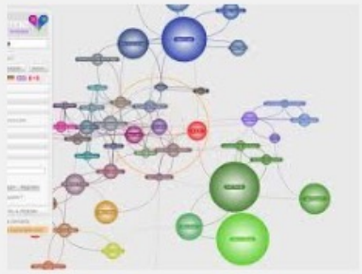
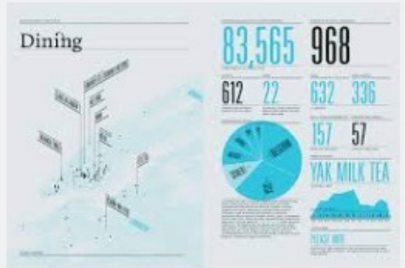



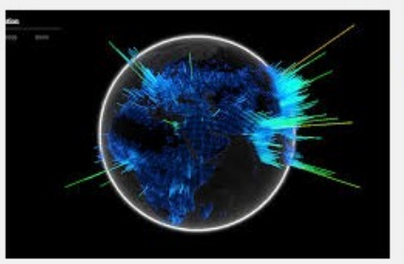



# Ask google for Data Vis (images)

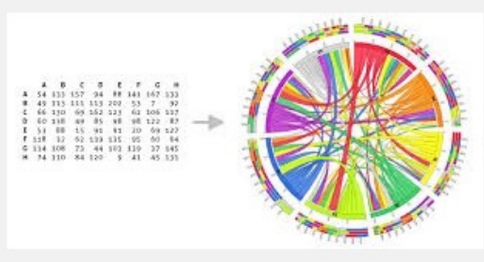
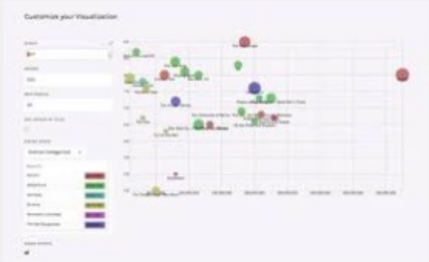
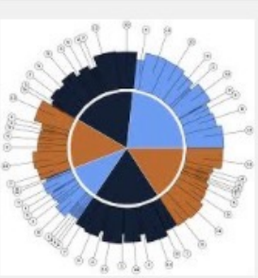
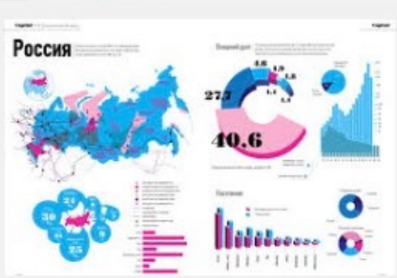


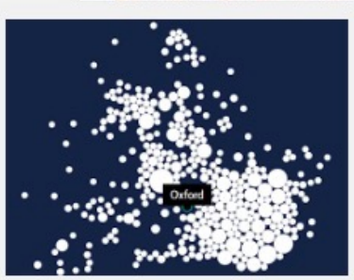





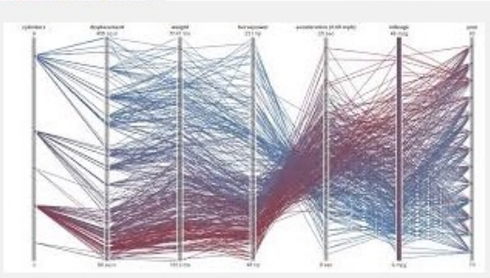



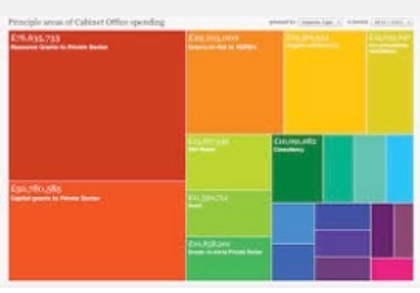


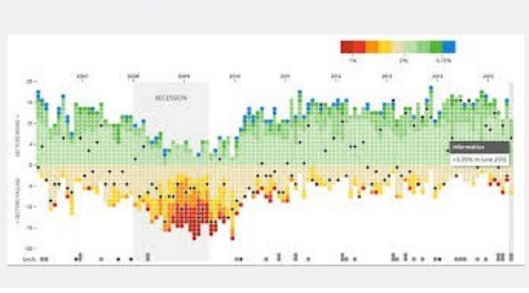


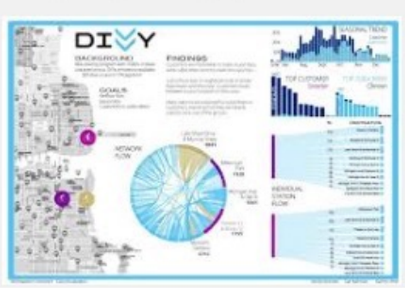












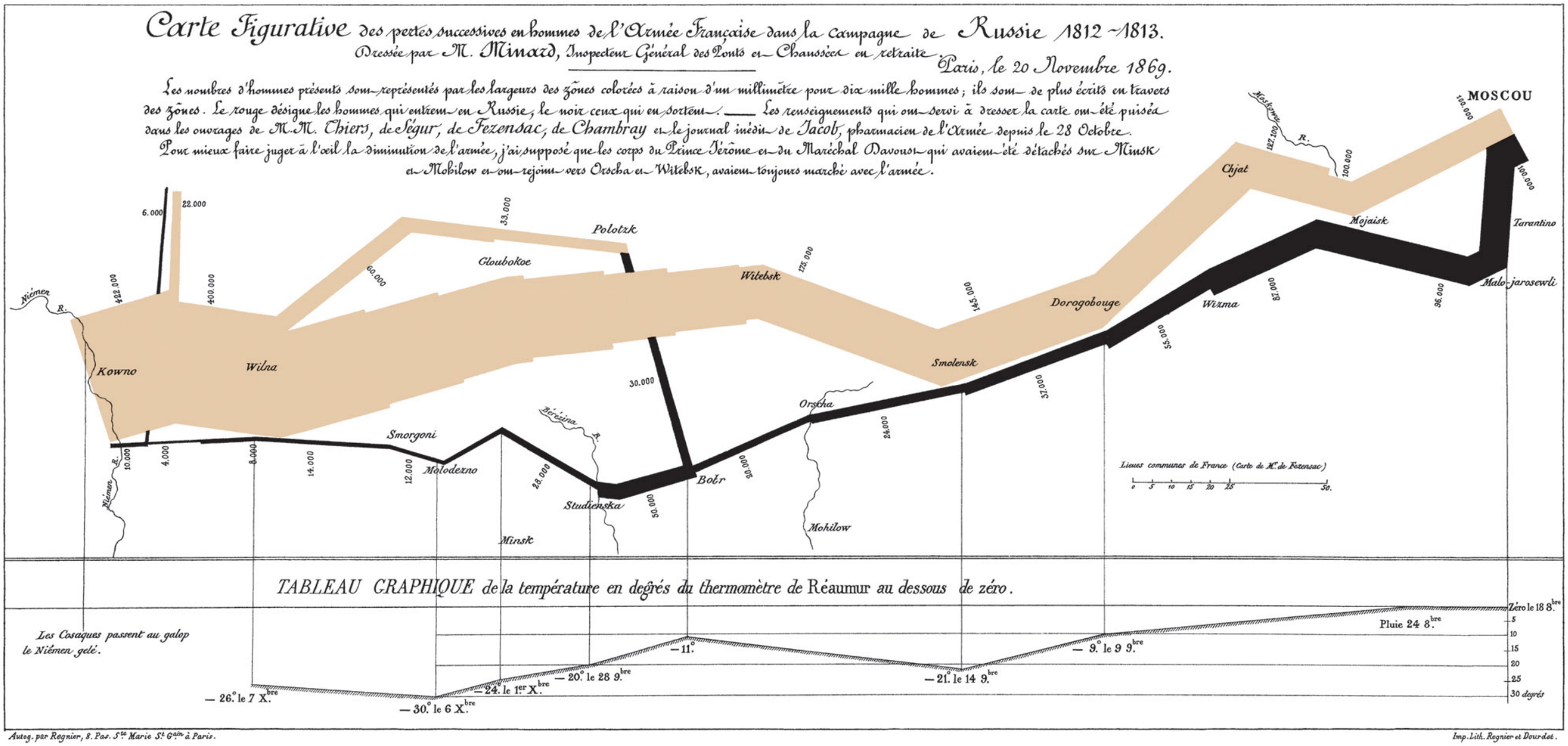





# Landmarks of (Data) Visualization

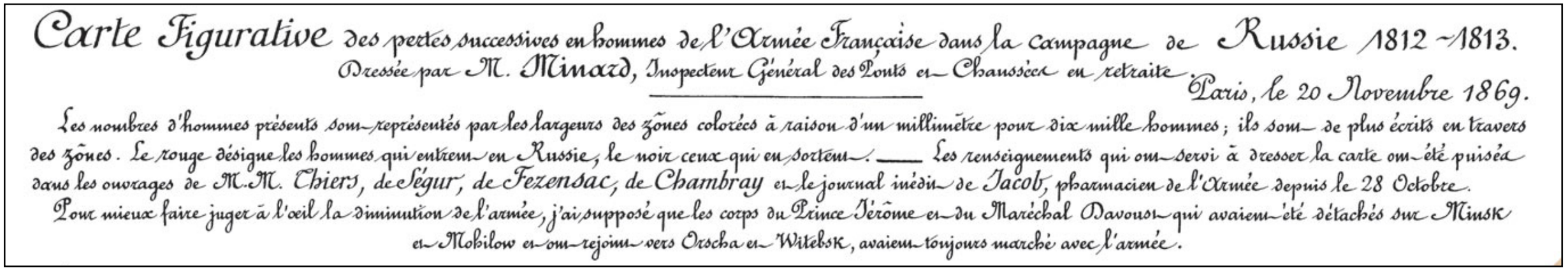
# Landmarks in Visualization

- **Charles Minard's map of Napoleon's disastrous Russian campaign of 1812.**



# Landmarks in Visualization

- **Charles Minard's map of Napoleon's disastrous Russian campaign of 1812**


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

Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite.

Paris, le 20 Novembre 1869.

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

Pour mieux faire juger à l'oeil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davoust qui avaient été détachés sur Minsk et Mobilow et ont rejoint vers Orscha et Witebsk, avaient toujours marché avec l'armée.

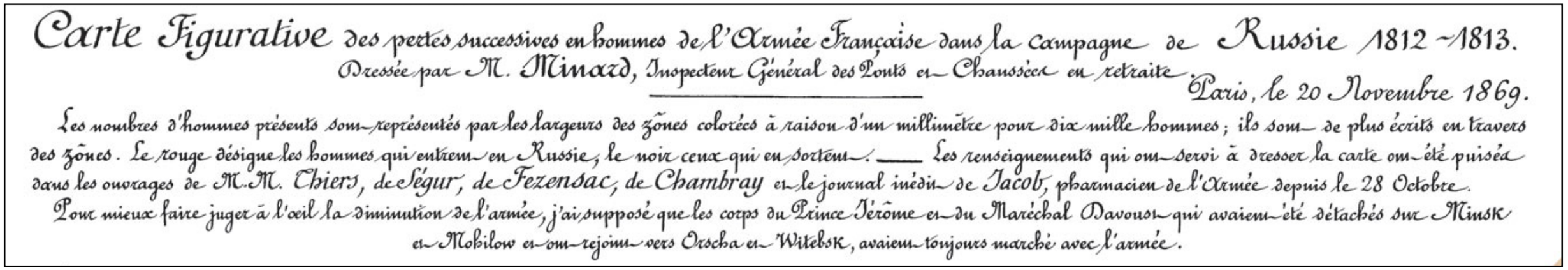
- **Carte Figurative des pertes successives en hommes de l'armée Française dans la campagne de Russie 1812-1813**

- **Le nombre d'hommes présents sont représentés par les largeurs des zones colorés à raison d'un millimètre pour dix mille hommes au travers des zones.**
- **Le "rouge" désignent des hommes qui entrem en Russie, le noir ceux qui en sorte**

- **Les renseignements qui on servit ….. [References]**

# Landmarks in Visualization

- **Charles Minard's map of Napoleon's disastrous Russian campaign of 1812**


Carte Figurative des pertes successives en hommes de l'Armée Française dans la campagne de Russie 1812-1813.
Dressée par M. Minard, Inspecteur Général des Ponts et Chaussées en retraite.
Paris, le 20 Novembre 1869.
Les nombres d'hommes présents sont représentés par les largeurs des zônes colorées à raison d'un millimètre pour dix mille hommes ; ils sont de plus écrits en travers des zônes. Le rouge désigne les hommes qui entrent en Russie, le noir ceux qui en sortent. ___ Les renseignements qui ont servi à dresser la carte ont été puisés dans les ouvrages de M.M. Thiers, de Ségur, de Fezensac, de Chambray et le journal inédit de Jacob, pharmacien de l'Armée depuis le 28 Octobre.
Pour mieux faire juger à l'oeil la diminution de l'armée, j'ai supposé que les corps du Prince Jérôme et du Maréchal Davoust qui avaient été détachés sur Minsk et Mobilow et ont rejoint vers Orscha et Witebsk, avaient toujours marché avec l'armée.

- **Figurative map of the successive losses of men of the French army in the Russian campaign 1812-1813**

- **The number of men present is represented by the widths of the colored areas at the rate of one millimeter for every ten thousand men across the areas.**
- **The "red" designate men who enter Russia, the black those who come out.**

- **The information we served… .. [References]**

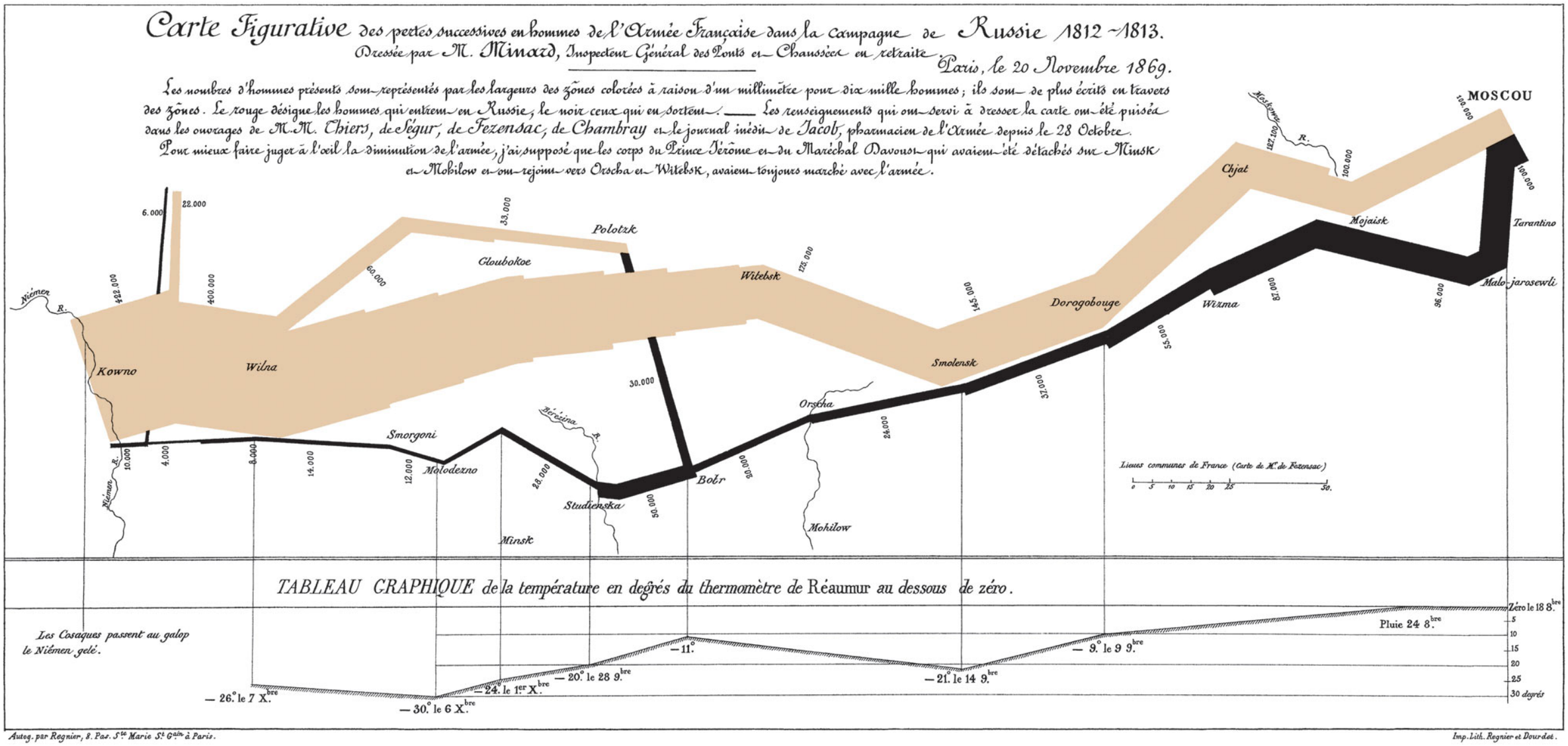
Estónia
Letónia
Lituânia
Bielorrússia
Saaremaa
Pärnu
Tartu
Narva
Gatchina
Гатчина
Veliky Novgorod
Великий Новгород
Pskov
Псков
Cherepovets
Череповец
Rybinsk
Рыбинск
Cēsis
Riga
Jūrmala
Tver
Тверь
Moscovo
Москва
Podolsk
Подольск
Vitebsk
Віцебск
Smolensk
Смоленск
Kaunas
Vilnius
Tula
Тула
Minsk
Мінск
Mahilyow
Маґілёў
Labkovičy
Лабковічы
Briansk
Брянск
Oriol
Орёл
Białystok
Gomel
Гомель
Kursk
Курск
Voronej
Воронеж
Chernigov
Чернігів
Staryy Oskol
Старый Оскол
Lublin
Sumy

December

October

# Landmarks in Visualization

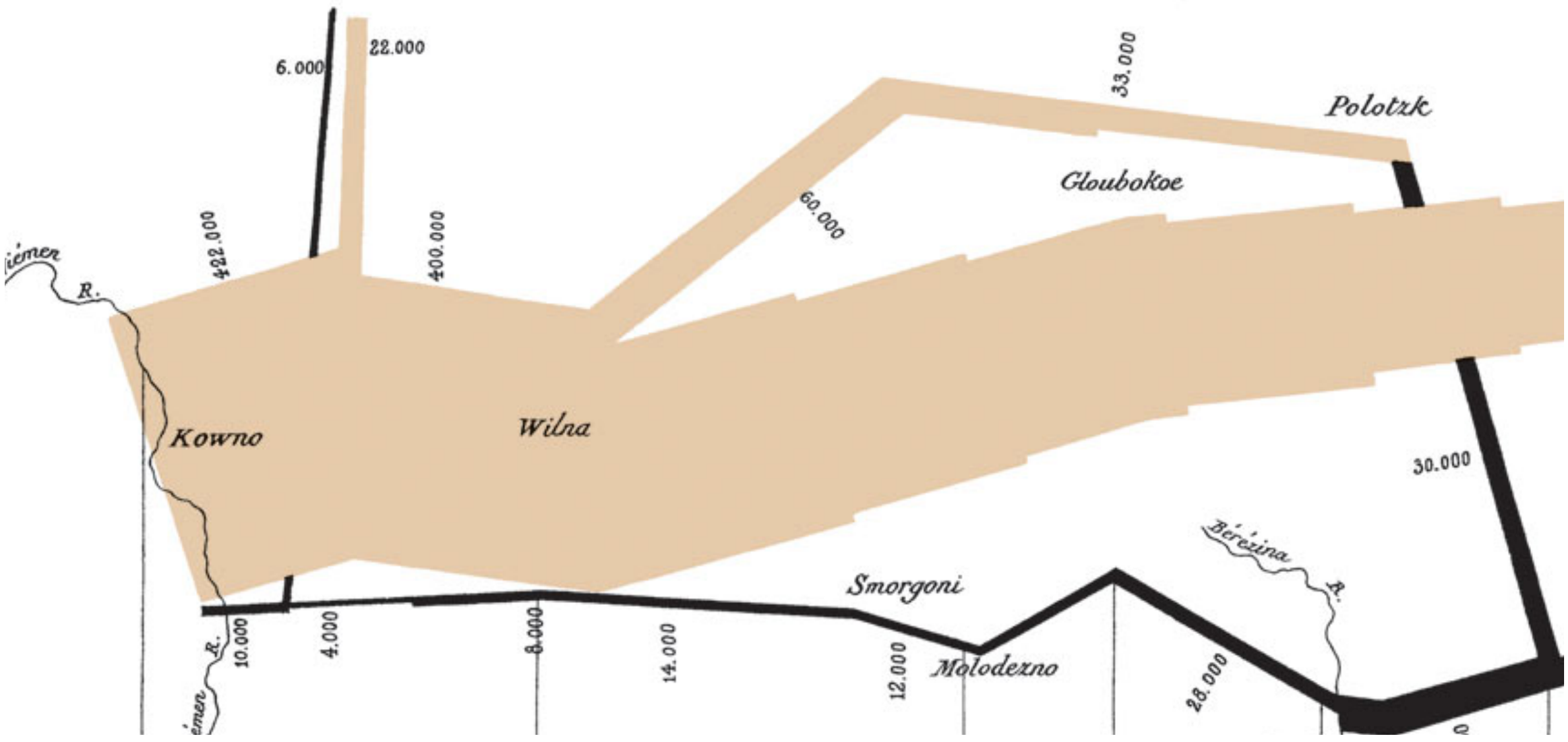
- **Charles Minard's map of Napoleon's disastrous Russian campaign of 1812.**




**Six different sets of data**: **geography**, the army's **course**, the army's **direction**; the **number of soldiers** remaining; **temperature**; **time**.

# Landmarks in Visualization

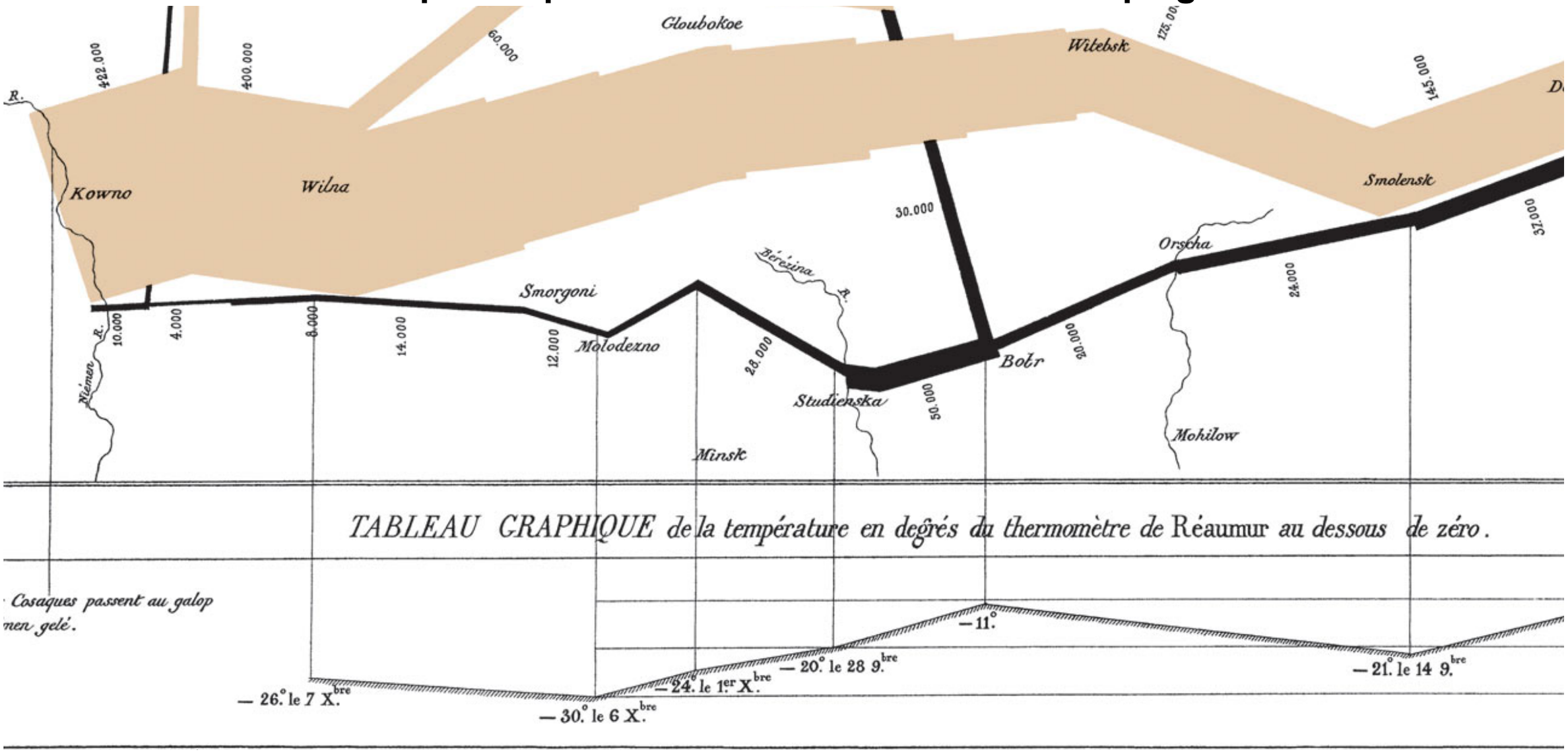
- **Charles Minard's map of Napoleon's disastrous Russian campaign of 1812.**




**Six different sets of data**: **geography**, the army's **course**, the army's **direction**; the **number of soldiers** remaining; **temperature**; **time**.

# Landmarks in Visualization

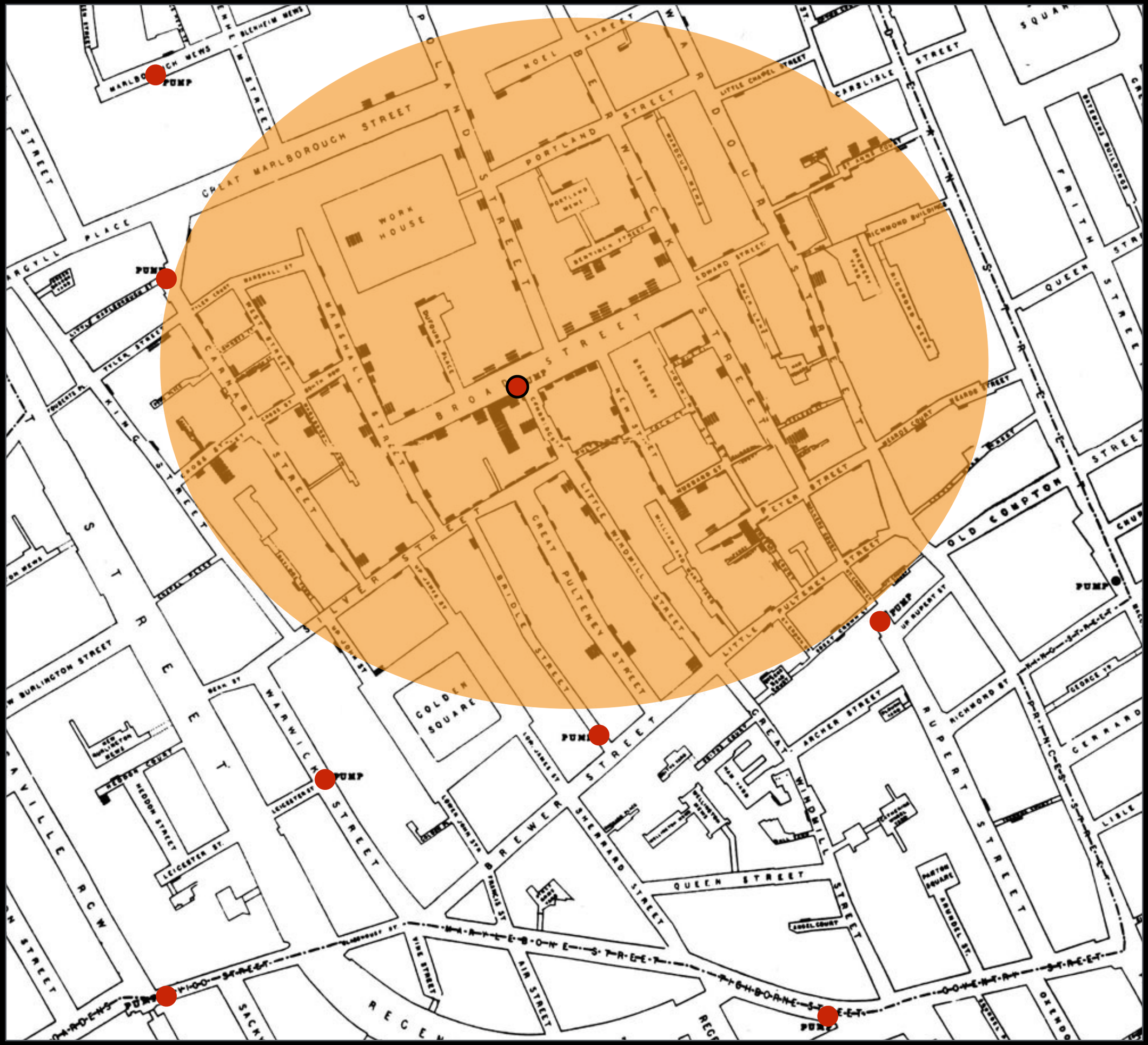
- **Charles Minard's map of Napoleon's disastrous Russian campaign of 1812.**




**Six different sets of data**: **geography**, the army's **course**, the army's **direction**; the **number of soldiers** remaining; **temperature**; **time**.

# Landmarks in Visualization

- **Broad Street cholera outbreak, (John Snow, 1854)**

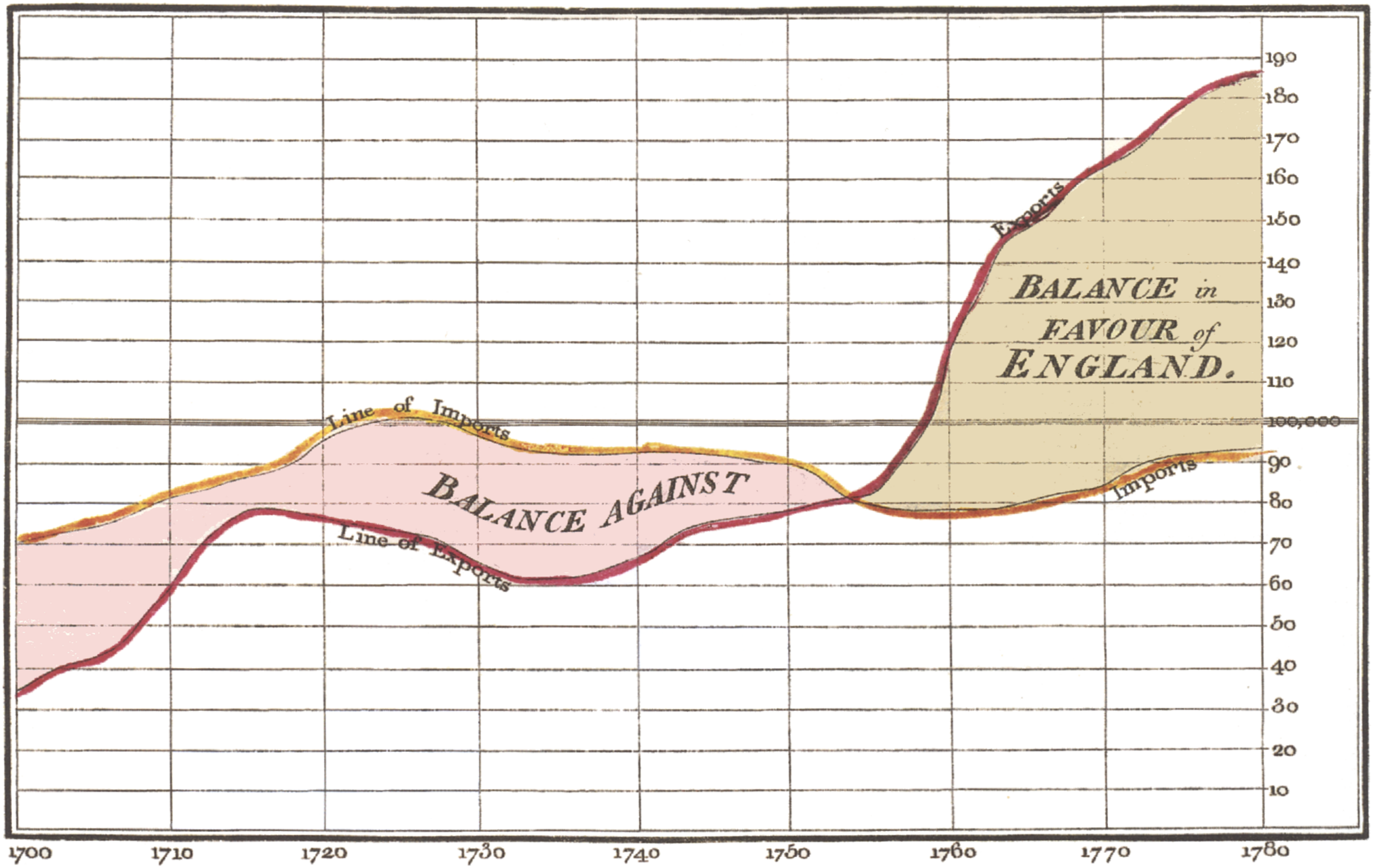

GREAT MARLBOROUGH STREET
PORTLAND STREET
BROAD STREET
WORK HOUSE
GOLDEN SQUARE
QUEEN STREET
PUMP

# Early visualizations

- **Recommended readings:**
  - **(Matthew O. Ward et all, 2010) pages 6 - 15.**
  - **See the suggested links at the end.**

- **Some “Landmarks”:**
  - **Broad Street cholera outbreak,  by John Snow, 1854.**
  - **Charles Minard's map of Napoleon's disastrous Russian campaign of 1812.**
  - **William Playfair founder of graphical methods of statistics.**

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




*The Bottom line is divided into Years, the Right hand line into £10,000 each.*

*Published as the Act directs, 1st May 1786. by Wm Playfair*

*Neele sculpt 352, Strand, London.*

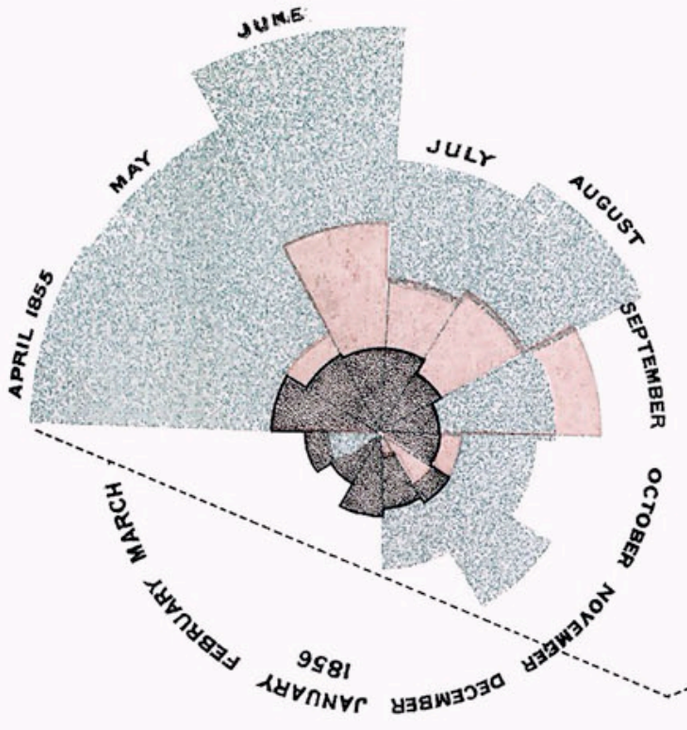
# Early visualizations

- **Recommended readings:**
  - **(Matthew O. Ward et all, 2010) pages 6 - 15.**
  - **See the suggested links at the end.**

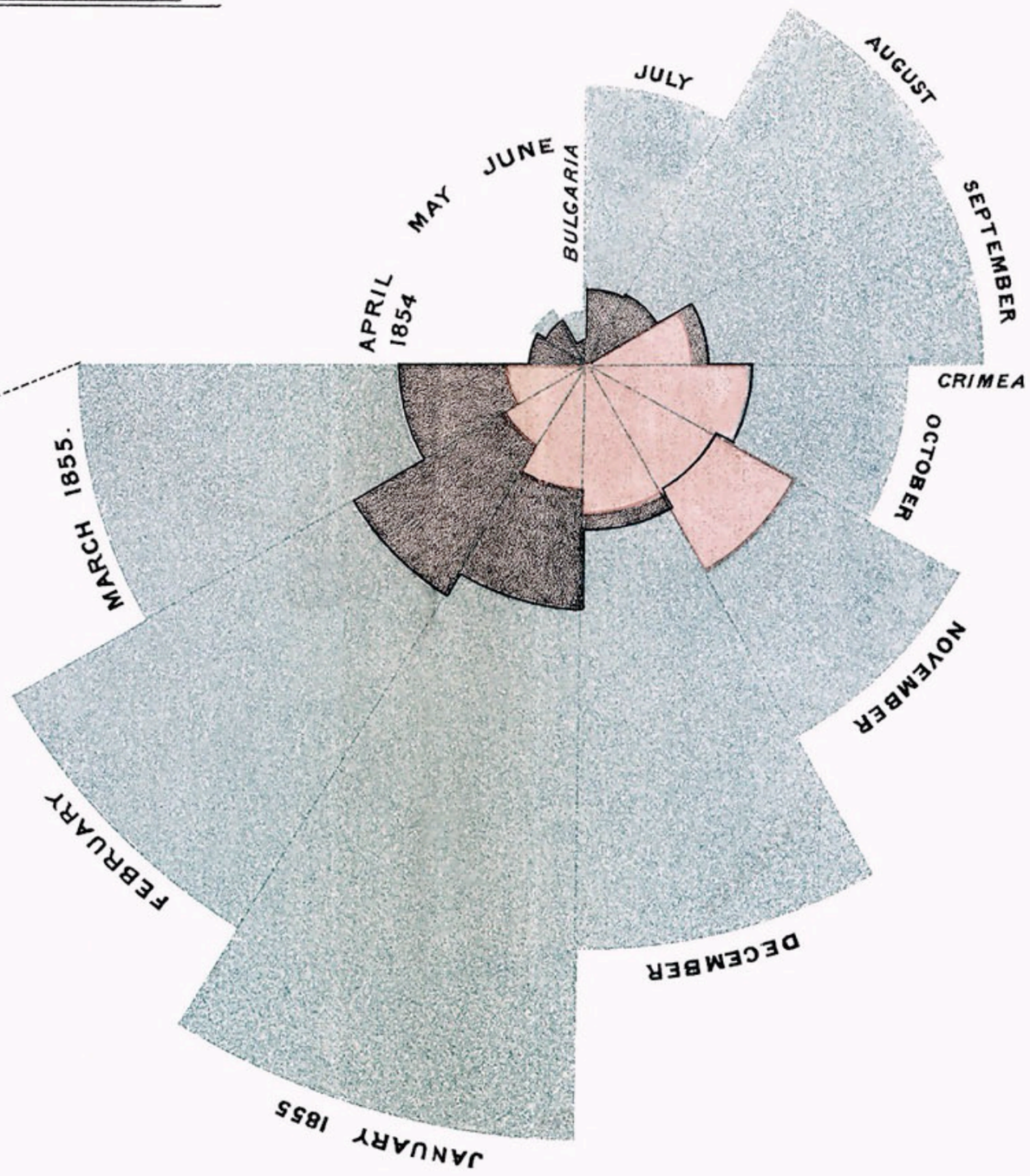
- **Some “Landmarks”:**
  - **Broad Street cholera outbreak,  by John Snow, 1854.**
  - **Charles Minard's map of Napoleon's disastrous Russian campaign of 1812.**
  - **William Playfair founder of graphical methods of statistics.**
  - **Florence Nightingale’s “Rose” representing the causes of mortality in the army.**

# DIAGRAM of the CAUSES of MORTALITY
## IN THE ARMY IN THE EAST.

2.
APRIL 1855 to MARCH 1856.




1.
APRIL 1854 to MARCH 1855.




*The Areas of the blue, red, & black wedges are each measured from the centre as the common vertex.*

*The blue wedges measured from the centre of the circle represent area for area the deaths from Preventible or Mitigable Zymotic diseases; the red wedges measured from the centre the deaths from wounds; & the black wedges measured from the centre the deaths from all other causes.*

*The black line across the red triangle in Nov.r 1854 marks the boundary of the deaths from all other causes during the month.*

*In October 1854, & April 1855, the black area coincides with the red; in January & February 1856, the blue coincides with the black.*

*The entire areas may be compared by following the blue, the red & the black lines enclosing them.*

# Early visualizations

- **Recommended readings:**
  - **(Matthew O. Ward et all, 2010) pages 6 - 15.**
  - **See the suggested links at the end.**

- **Some “Landmarks”:**
  - **Broad Street cholera outbreak,  by John Snow, 1854.**
  - **Charles Minard's map of Napoleon's disastrous Russian campaign of 1812.**
  - **William Playfair founder of graphical methods of statistics.**
  - **Florence Nightingale’s “Rose” representing the causes of mortality in the army.**
  - **Periodic Table: Dmitri Mendeleev (1869); Julius Lothar Meyer (1870).**

# What is (Data) Visualization?

# Modes of Visualization

## Modes of Visualization

**John C. Hart**

### Interactive Visualization

- Used for discovery
- Intended for a single investigator or collaborators
- Rerenders based on input
- Prototype quality

# Modes of Visualization

## Modes of Visualization

**John C. Hart**

### Presentation Visualization

- Used for communication
- Intended for large group or mass audience
- Does not support user input
- Highly polished

# Modes of Visualization

## Modes of Visualization

**John C. Hart**

### Interactive Visualization

- Used for discovery
- Intended for a single investigator or collaborators
- Rerenders based on input
- Prototype quality

### Presentation Visualization

- Used for communication
- Intended for large group or mass audience
- Does not support user input
- Highly polished

### Interactive Storytelling

- Presentations via interative webpages

6

# Modes of Visualization

## Modes of Visualization

**John C. Hart**

| Visualization Mode | User Interaction | Graphics Rendering | Target | Medium |
|---|---|---|---|---|

# Modes of Visualization

## Modes of Visualization

**John C. Hart**

| Visualization Mode | User Interaction | Graphics Rendering | Target | Medium |
|---|---|---|---|---|
| **Presentation Visualization** | User only observes | Precomputed rendering | Colleagues, mass audience | Slide shows, video |

# Modes of Visualization

## Modes of Visualization

**John C. Hart**

| Visualization Mode | User Interaction | Graphics Rendering | Target | Medium |
|---|---|---|---|---|
| **Interactive Visualization** | User controls everything, including dataset | Real-time rendering | Individual or collaborators | Software or internet |

# Modes of Visualization

## Modes of Visualization

**John C. Hart**

| Visualization Mode | User Interaction | Graphics Rendering | Target | Medium |
|---|---|---|---|---|
| **Interactive Storytelling** | User can filter or inspect details of preset datasets | Real-time rendering | Mass audience | Internet or kiosk |

# Modes of Visualization

## Modes of Visualization

**John C. Hart**

| Visualization Mode | User Interaction | Graphics Rendering | Target | Medium |
|---|---|---|---|---|
| **Interactive Visualization** | User controls everything, including dataset | Real-time rendering | Individual or collaborators | Software or internet |
| **Interactive Storytelling** | User can filter or inspect details of preset datasets | Real-time rendering | Mass audience | Internet or kiosk |
| **Presentation Visualization** | User only observes | Precomputed rendering | Colleagues, mass audience | Slide shows, video |

# What is the Goal of Data Visualization?

**The (ultimate) goal of DV**

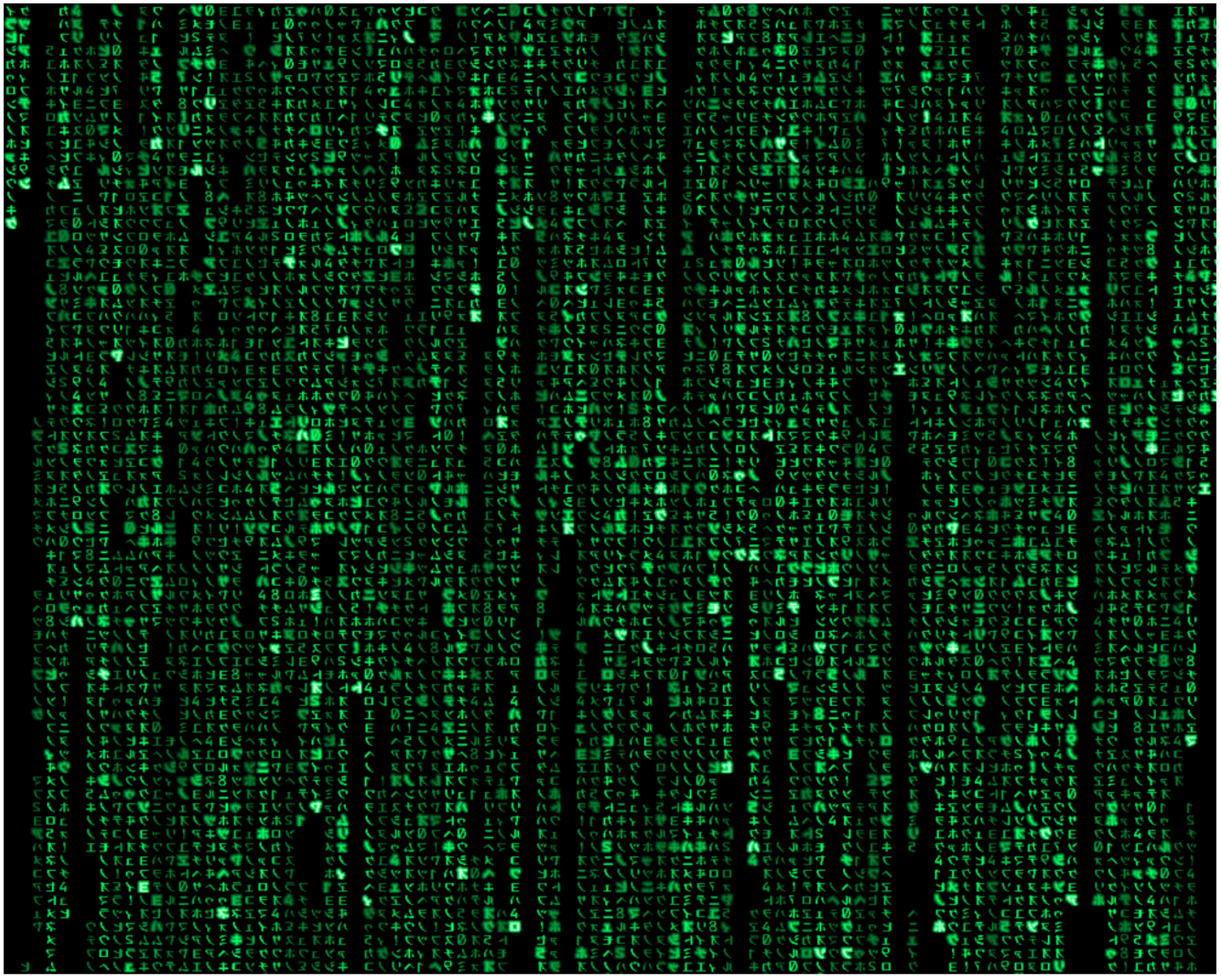
**“Data visualization is not just about seeing data !**

**Is about UNDERSTANDING data,**

**and being able to make decisions based on the data”**

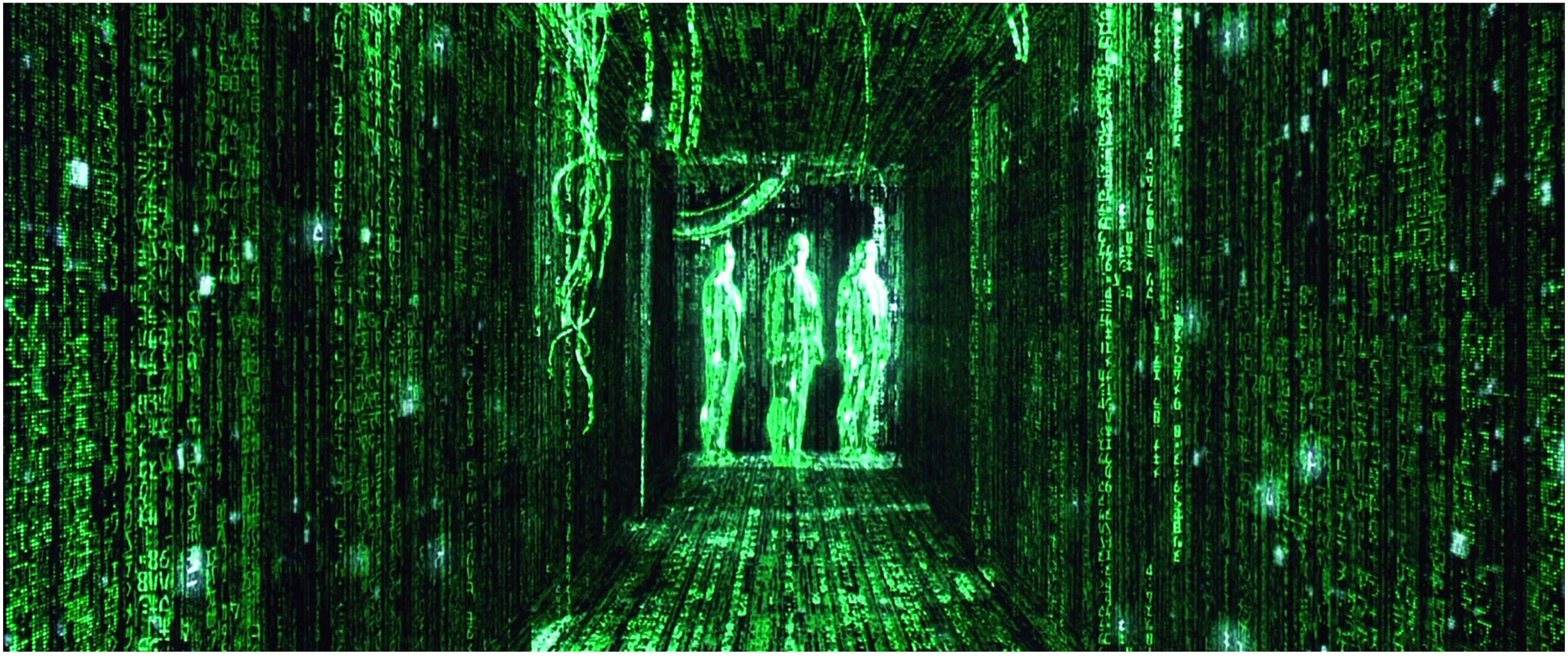
**by John C. Hart**

# What is the Goal of Data Visualization?



# What is the Goal of Data Visualization?

**“Data visualization is not just about seeing data !**



# What is the Goal of Data Visualization?

**“Data visualization is not just about seeing data !**

**Is about UNDERSTANDING data,**

**and being able to make decisions based on the data”**

**by John C. Hart**

# The concerns of Data Visualization field

- **Visualization provides an alternative or a supplement for textual or verbal information**
- **(in many situations) Visualization provides a richer description of information than the word-based counterpart !**

---

- **Why?**
- **In what kinds of situations are visualizations effective?**
- **What type of information can and cannot be visualized?**
- **How many different ways are there to show the same data? Which ones are best for particular circumstances?**
- **Why should we study visualization?**

# Why (data) visualization is important?

# Why (data) visualization is important?

**Anscombe's Quartet: Raw Data**

| 1 | | 2 | | 3 | | 4 | |
|---|---|---|---|---|---|---|---|
| X | Y | X | Y | X | Y | X | Y |
| 10.0 | 8.04 | 10.0 | 9.14 | 10.0 | 7.46 | 8.0 | 6.58 |
| 8.0 | 6.95 | 8.0 | 8.14 | 8.0 | 6.77 | 8.0 | 5.76 |
| 13.0 | 7.58 | 13.0 | 8.74 | 13.0 | 12.74 | 8.0 | 7.71 |
| 9.0 | 8.81 | 9.0 | 8.77 | 9.0 | 7.11 | 8.0 | 8.84 |
| 11.0 | 8.33 | 11.0 | 9.26 | 11.0 | 7.81 | 8.0 | 8.47 |
| 14.0 | 9.96 | 14.0 | 8.10 | 14.0 | 8.84 | 8.0 | 7.04 |
| 6.0 | 7.24 | 6.0 | 6.13 | 6.0 | 6.08 | 8.0 | 5.25 |
| 4.0 | 4.26 | 4.0 | 3.10 | 4.0 | 5.39 | 19.0 | 12.50 |
| 12.0 | 10.84 | 12.0 | 9.13 | 12.0 | 8.15 | 8.0 | 5.56 |
| 7.0 | 4.82 | 7.0 | 7.26 | 7.0 | 6.42 | 8.0 | 7.91 |
| 5.0 | 5.68 | 5.0 | 4.74 | 5.0 | 5.73 | 8.0 | 6.89 |

# Why (data) visualization is important?

**Anscombe's Quartet: Raw Data**

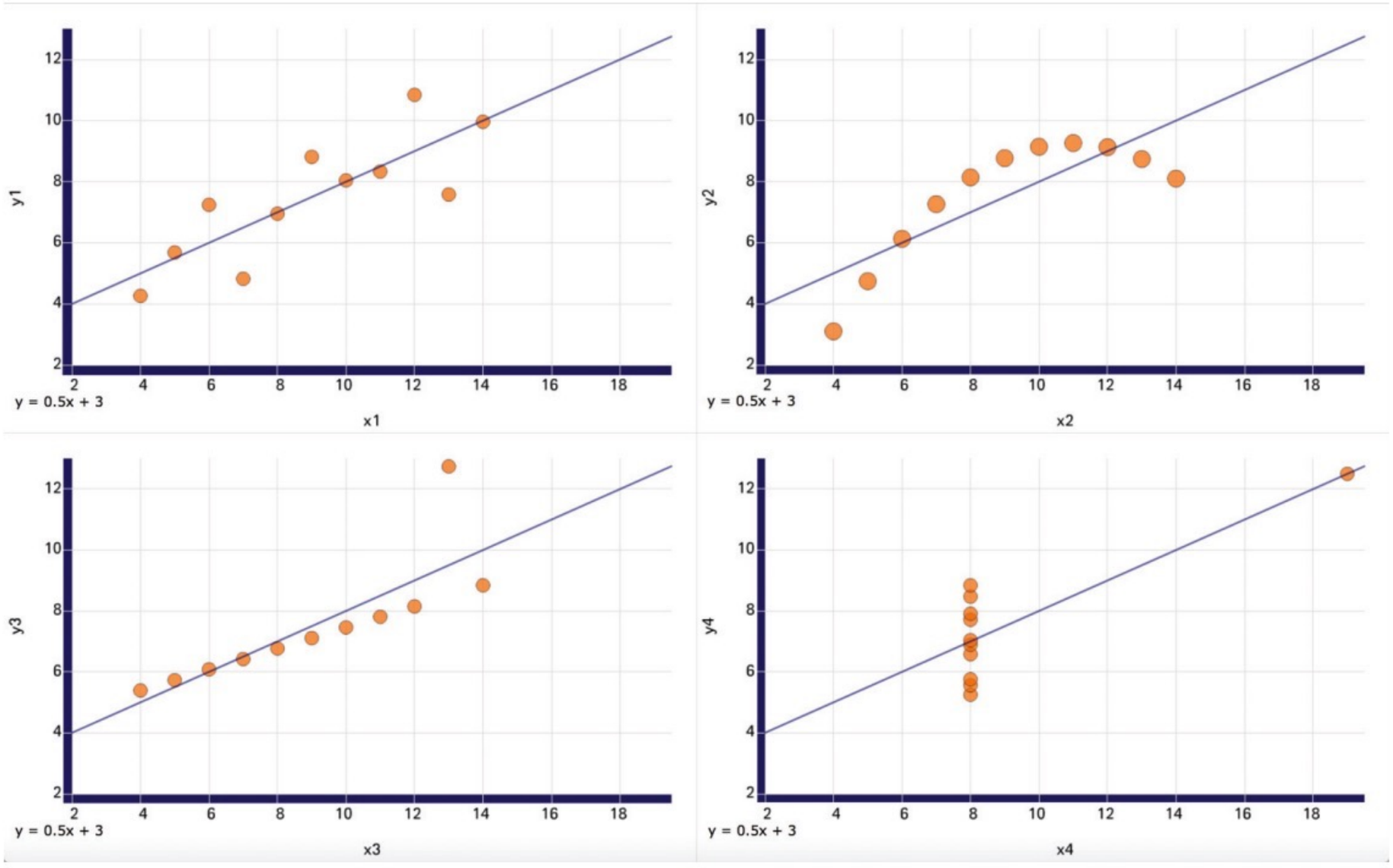
| 1 | | 2 | | 3 | | 4 | |
|---|---|---|---|---|---|---|---|
| X | Y | X | Y | X | Y | X | Y |
| 10.0 | 8.04 | 10.0 | 9.14 | 10.0 | 7.46 | 8.0 | 6.58 |
| 8.0 | 6.95 | 8.0 | 8.14 | 8.0 | 6.77 | 8.0 | 5.76 |
| 13.0 | 7.58 | 13.0 | 8.74 | 13.0 | 12.74 | 8.0 | 7.71 |
| 9.0 | 8.81 | 9.0 | 8.77 | 9.0 | 7.11 | 8.0 | 8.84 |
| 11.0 | 8.33 | 11.0 | 9.26 | 11.0 | 7.81 | 8.0 | 8.47 |
| 14.0 | 9.96 | 14.0 | 8.10 | 14.0 | 8.84 | 8.0 | 7.04 |
| 6.0 | 7.24 | 6.0 | 6.13 | 6.0 | 6.08 | 8.0 | 5.25 |
| 4.0 | 4.26 | 4.0 | 3.10 | 4.0 | 5.39 | 19.0 | 12.50 |
| 12.0 | 10.84 | 12.0 | 9.13 | 12.0 | 8.15 | 8.0 | 5.56 |
| 7.0 | 4.82 | 7.0 | 7.26 | 7.0 | 6.42 | 8.0 | 7.91 |
| 5.0 | 5.68 | 5.0 | 4.74 | 5.0 | 5.73 | 8.0 | 6.89 |
| Mean | 9.0 | 7.5 | 9.0 | 7.5 | 9.0 | 7.5 | 9.0 | 7.5 |
| Variance | 10.0 | 3.75 | 10.0 | 3.75 | 10.0 | 3.75 | 10.0 | 3.75 |
| Correlation | 0.816 | | 0.816 | | 0.816 | | 0.816 | |

**4 datasets**
**2 variables, 11 rows**

| Property | Value |
|---|---|
| Mean of X | 9 |
| Variance of X | 11 |
| Mean of Y | 7.5 |
| Variance of Y | 4.1 |
| Correlation | 0.816 |
| Linear Regression | y = 3.0 + 0.5x |

**F. J. Anscombe (1973)**

# Why (data) visualization is important?




**4 datasets**
**2 variables, 11 rows**

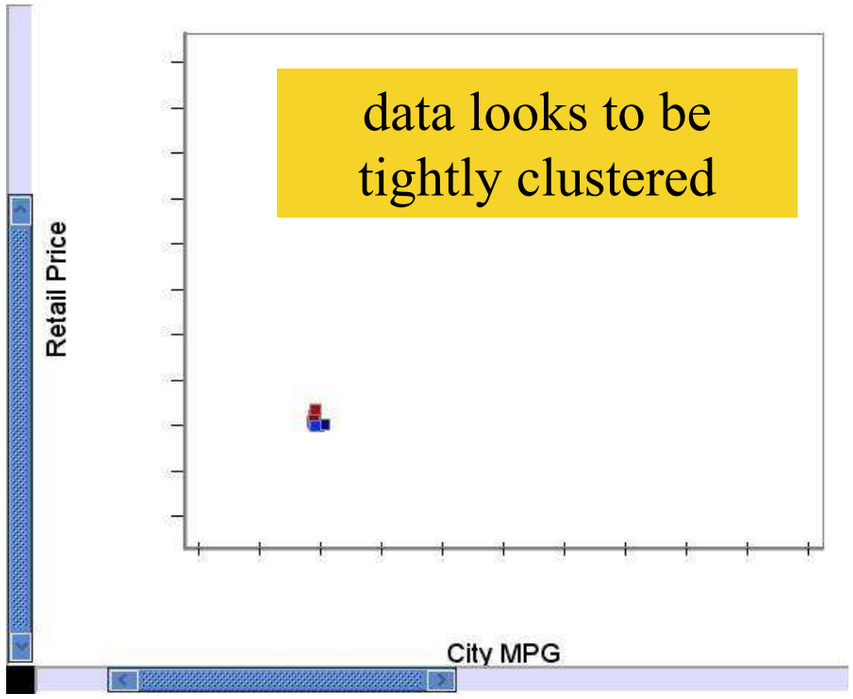
| Property | Value |
|---|---|
| Mean of X | 9 |
| Variance of X | 11 |
| Mean of Y | 7.5 |
| Variance of Y | 4.1 |
| Correlation | 0.816 |
| Linear Regression | y = 3.0 + 0.5x |

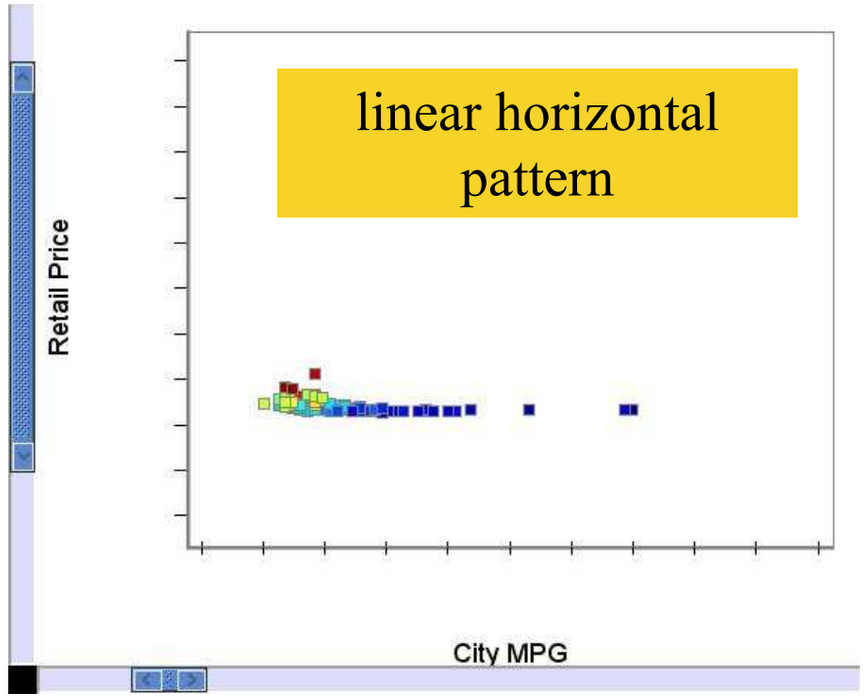
**F. J. Anscombe (1973)**

**Statistic is not enough !** **Data Vis is not enough !**

# Why (data) visualization is important?

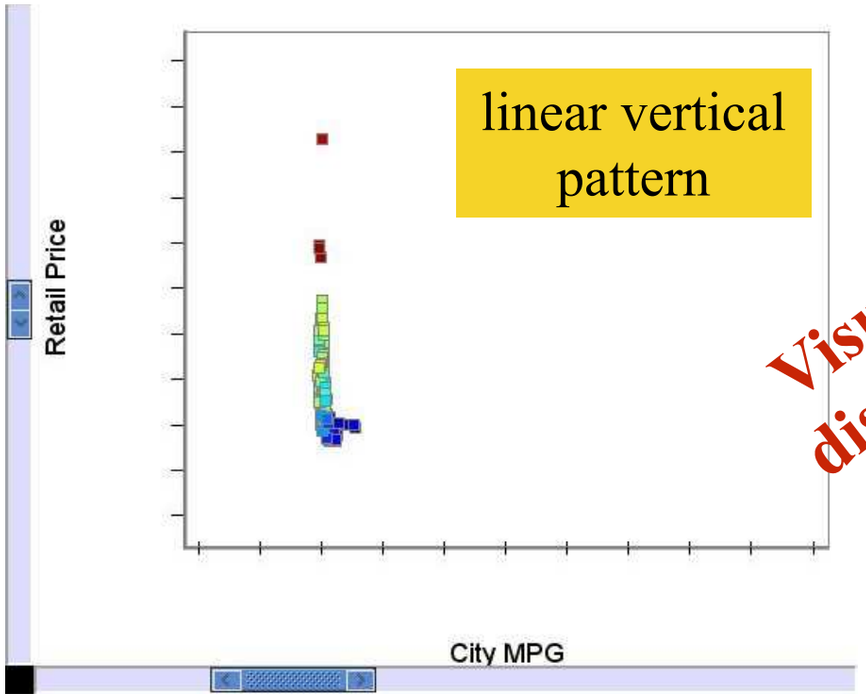
- **What is the effect of the presentation of the data on the decision making process?**
- **Can the presentation of data impact the decision?**
- **Can we say which presentations are better or more influential than others?**



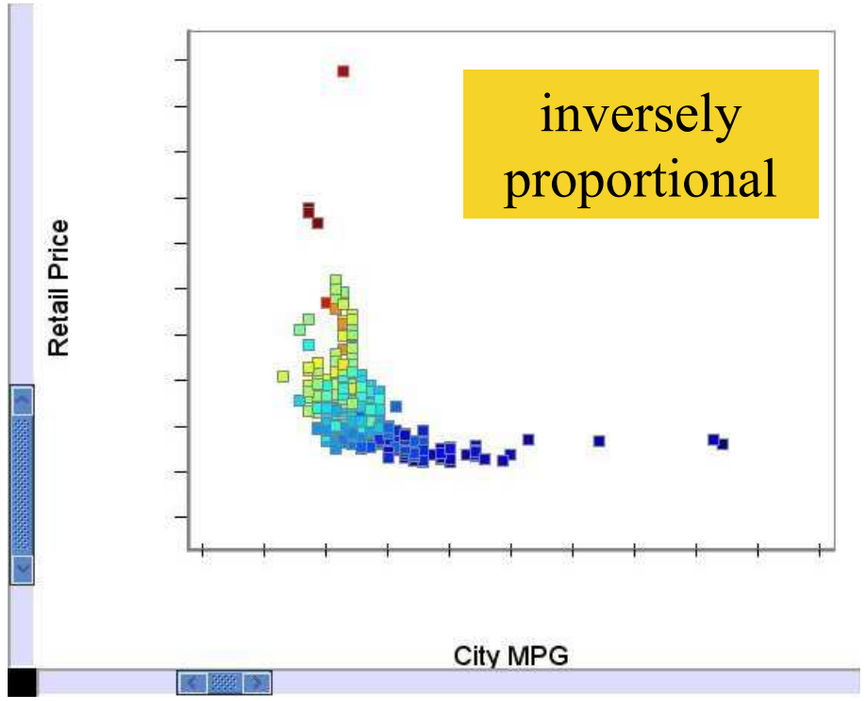



# Why (data) visualization is important?

- **What is the effect of the presentation of the data on the decision making process?**
- **Can the presentation of data impact the decision?**
- **Can we say which presentations are better or more influential than others?**




**Visualization can distort the "truth"**



# Why (data) visualization is important?

- **What is the the role of human preferences and training in the visualization?**

> . Linda S. Elting, James M. Walker, Charles G. Martin, Scott B. Cantor, and Edward B. Rubenstein. "**Influence of Data Display Formats on Decisions** to Stop Clinical Trials." British Medical Journal 318 (1999)

- **Hypothetical clinical trial:**
  - **Two treatments: 50 patients with conventional and 60 with investigational**
  - **Two populations:  65 with good prognosis and 45 with bad prognosis**
  - **Two outcomes for each treatment: Response (positive) vs Fail**
  - **34 clinicians**
  - **If a clinician sees that one treatment is better than the other, then he should stop the clinical treatment**

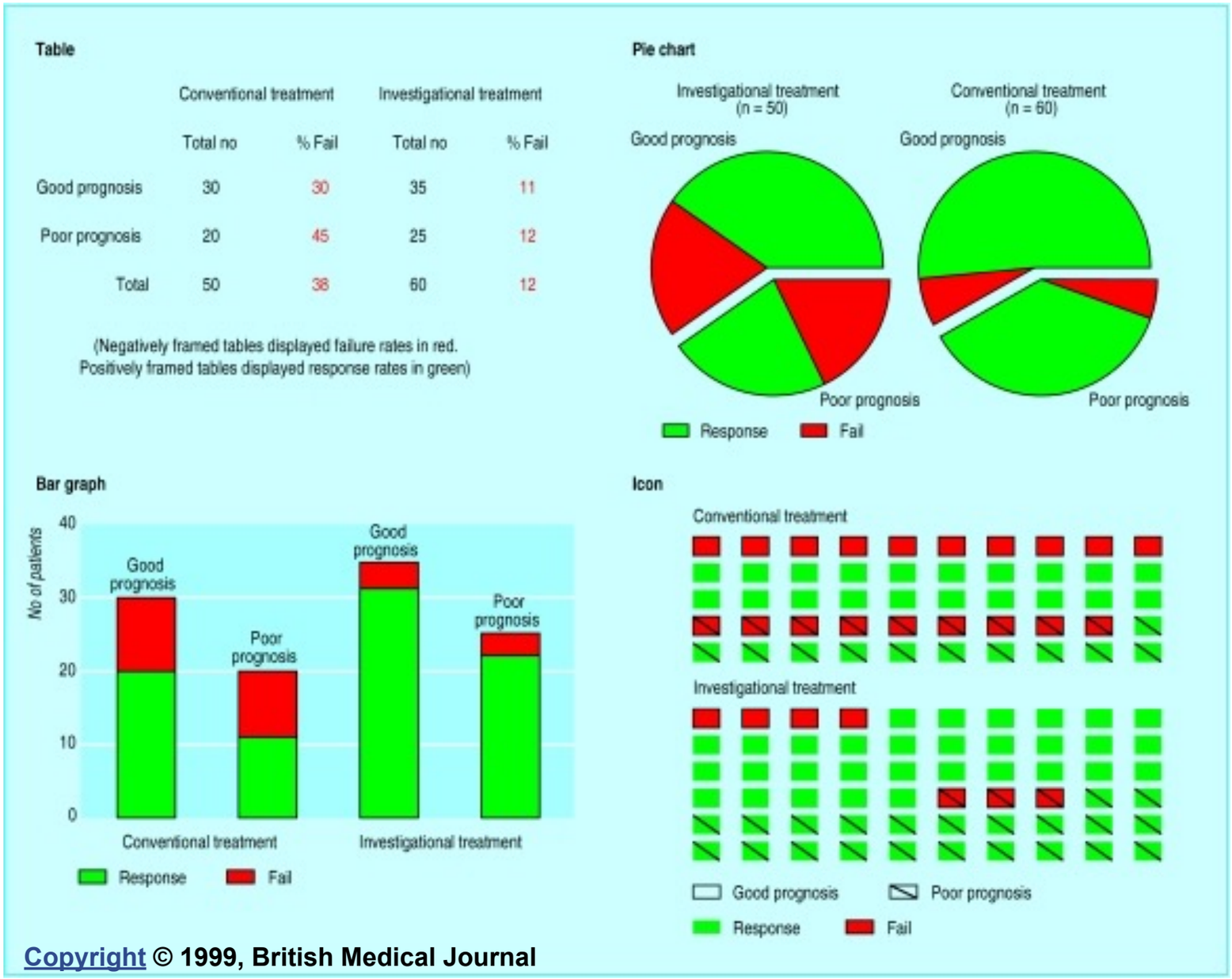
# Why (data) visualization is important?

- **What is the the role of human preferences and training in the visualization?**

**4 visualizations:**

Table
Pie Chart
Bar Graph
Icon

Green - Response
Red - Fail

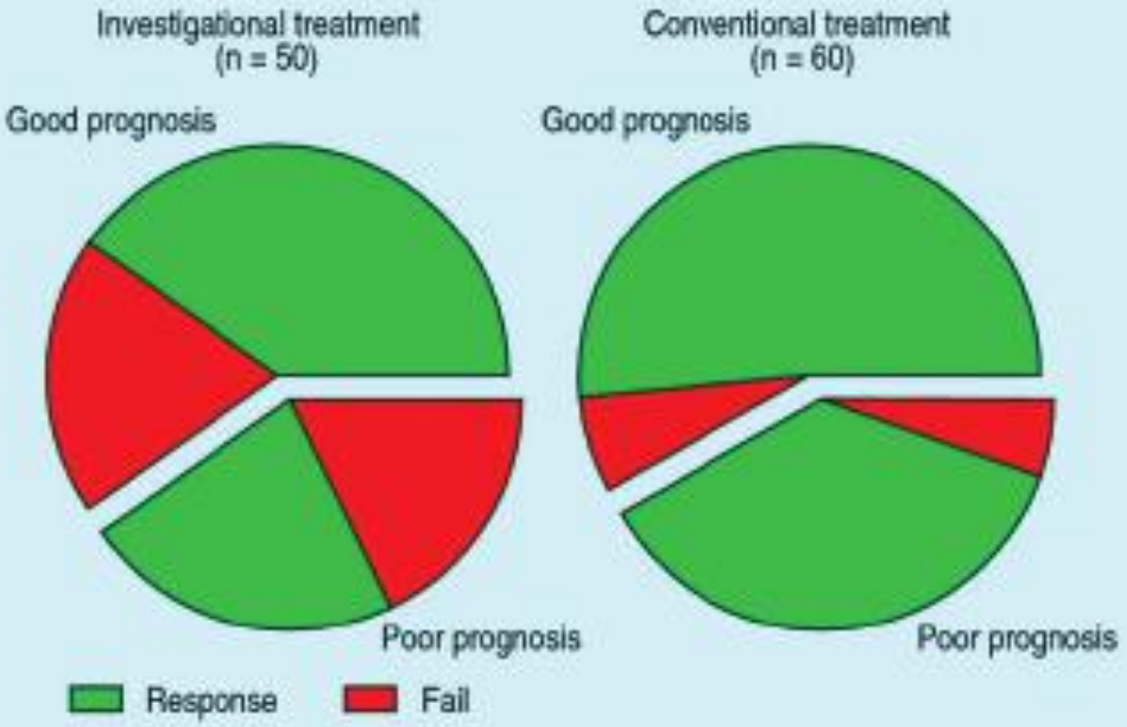


**Table**

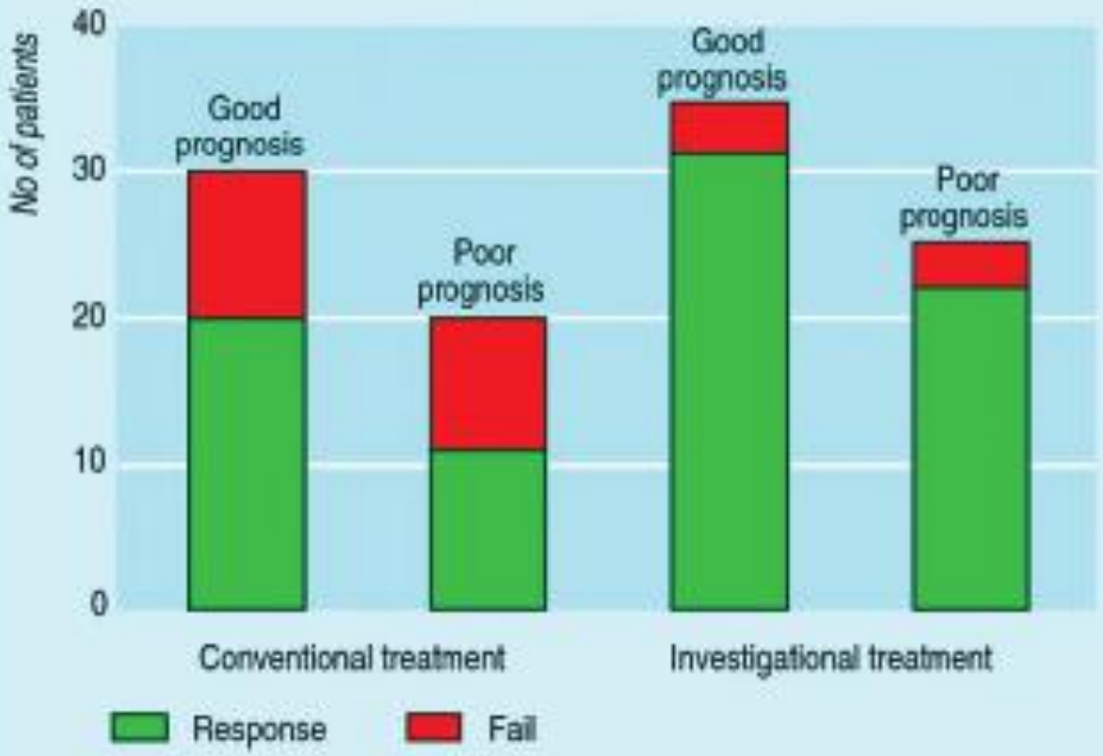
| | Conventional treatment | | Investigational treatment | |
|---|---|---|---|---|
| | Total no | % Fail | Total no | % Fail |
| Good prognosis | 30 | 30 | 35 | 11 |
| Poor prognosis | 20 | 45 | 25 | 12 |
| Total | 50 | 38 | 60 | 12 |

(Negatively framed tables displayed failure rates in red.
Positively framed tables displayed response rates in green)

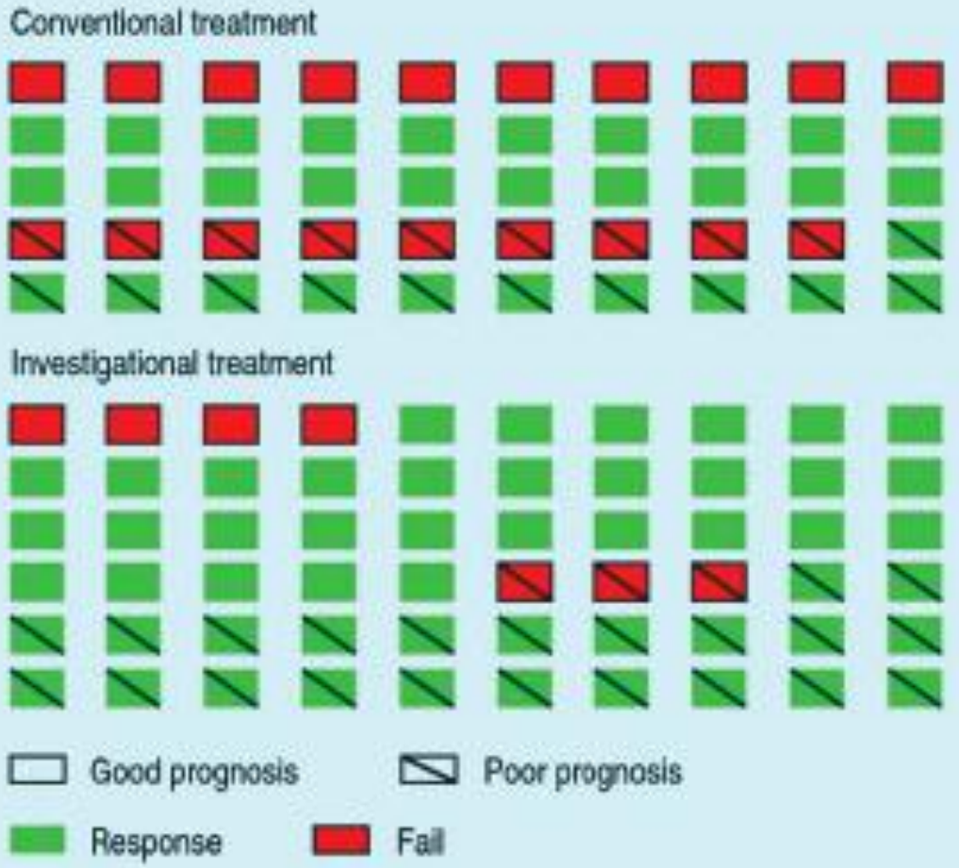
**Pie chart**




**Bar graph**




**Icon**



# Why (data) visualization is important?

- **What is the the role of human preferences and training in the visualization?**
  - **PARTICIPANTS** were shown tables, pie charts, bar graphs, and icon displays containing hypothetical data from a clinical trial and were asked to decide whether to continue the trial or stop for an unplanned statistical analysis.
  - **MAIN MEASURE : Percentage of accurate decisions with each type of display**
  - **RESULTS:**
    - **More correct decisions were made with icon displays** (82%) than with tables (68%), pie charts (56%), and bar graphs (43%).
    - Most (21) clinicians  preferred the table; Several were contemptuous of the icon display.

**The visualization is key in presenting data but the user preferences are very involved**

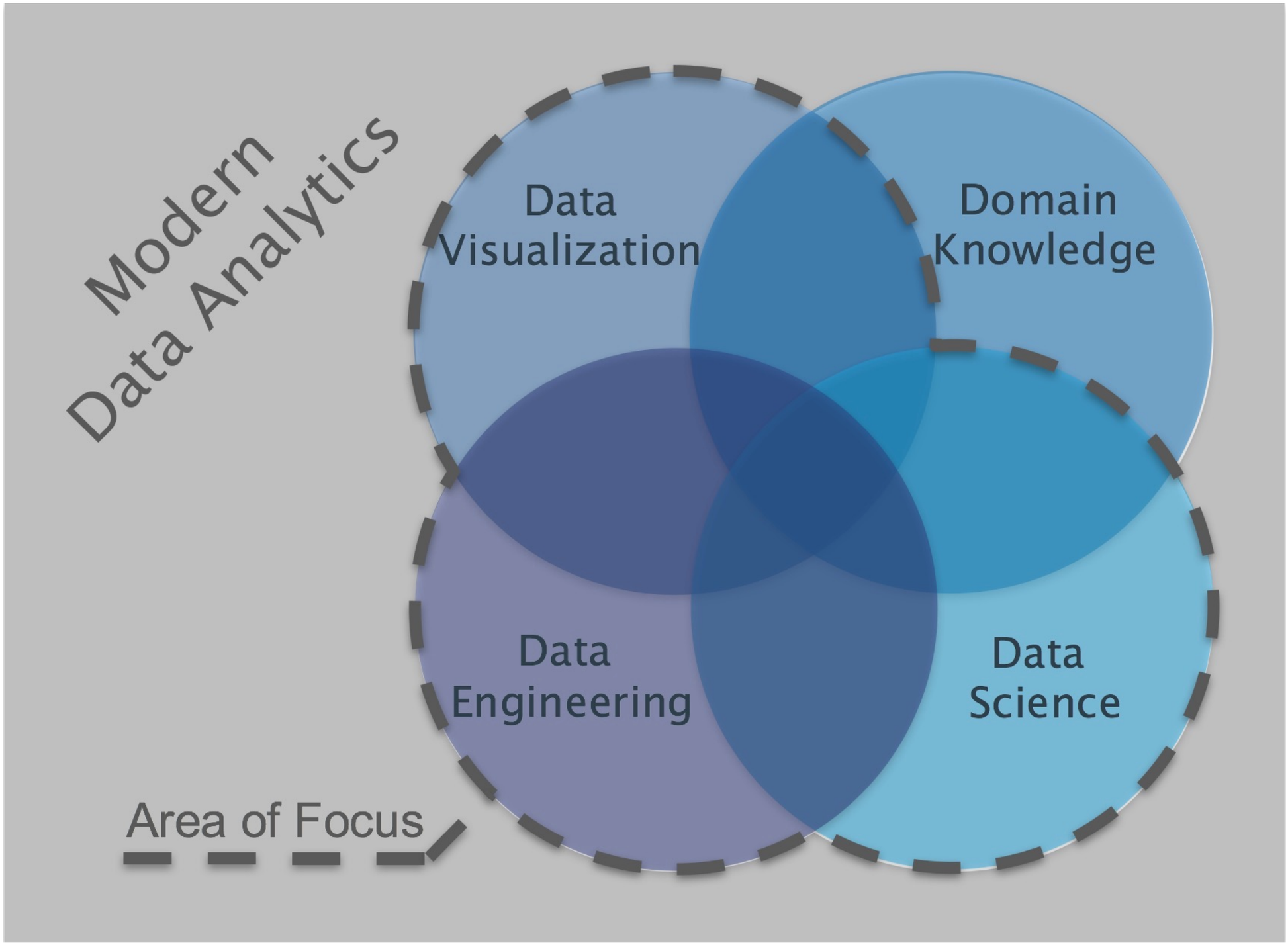
# Why (data) visualization is important?

- **Nowadays**
  - **Generating a lot of data and information**
  - **Need to process such information**
  - **Need to communicate increasing levels of information**

**Visualization is a cornerstone of modern knowledge discovery tools. Applications often include one or more visualizations to provide different views of data to describe some patterns or structures.**

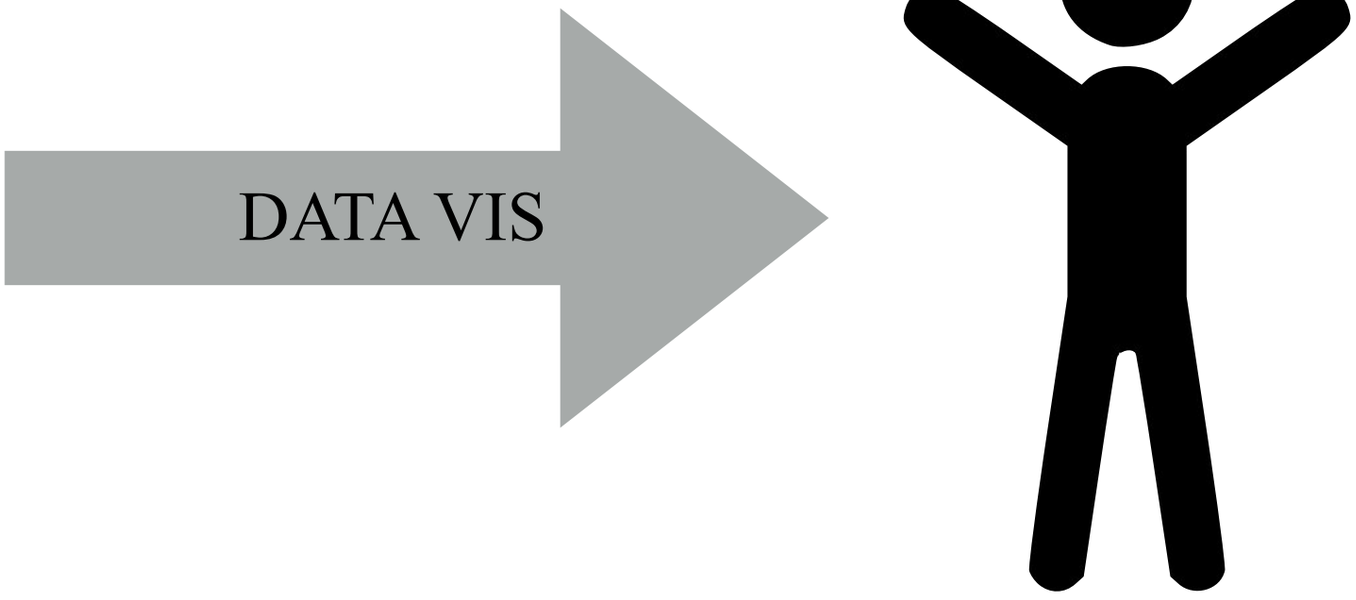
**We need to communicate information to people in an efficient and effective manner.**

# Why (data) visualization is important?



# **When** (data) visualization is important?

**DATA**



# **When** (data) visualization is not important?

# Whys from Tamara Munzner

- **Why have a human in the decision-making loop?**
- **Why have a computer in the loop?**
- **Why use an external representation?**
- **Why depend on vision?**
- **Why show the data in detail?**
- **Why use interactivity?**

# Whys from Tamara Munzner

- **Why is the vis idiom design space huge?**
- **Why focus on tasks?**
- **Why are most designs ineffective?**
- **Why care about effectiveness?**
- **Why is validation difficult?**
- **Why are there resource limitations?**
- **Why analyze vis?**

# Why the **study** of data visualization is important?

- **Tamara Munzner, 2015**

## 1.11 Why Are Most Designs Ineffective?

The most fundamental reason that vis design is a difficult enterprise is that the vast majority of the possibilities in the design space will be ineffective for any specific usage context. In some cases, a possible design is a poor match with the properties of the human perceptual and cognitive systems. In other cases, the design would be comprehensible by a human in some other setting, but it's a bad match with the intended task. Only a very small number of possibilities are in the set of reasonable choices, and of those only an even smaller fraction are excellent choices. Randomly choosing possibilities is a bad idea because the odds of finding a very good solution are very low.

# Why the **study** of data visualization is important?

- **Tamara Munzner, 2015**

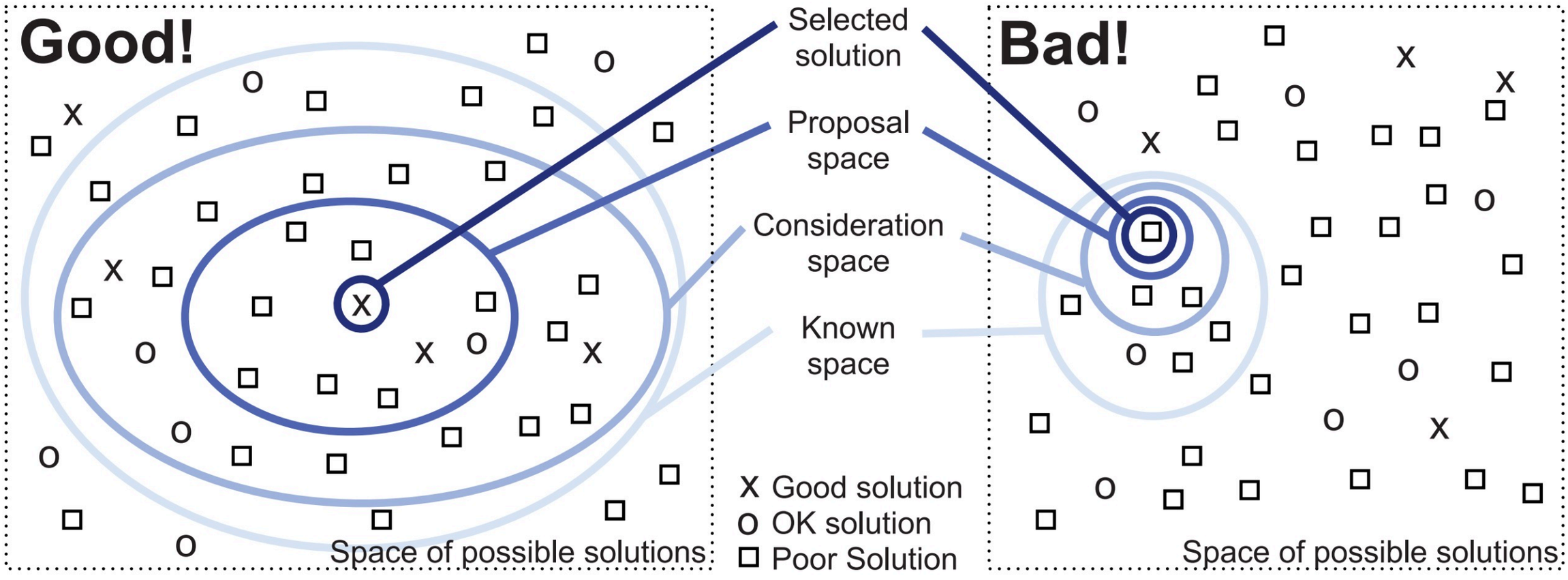
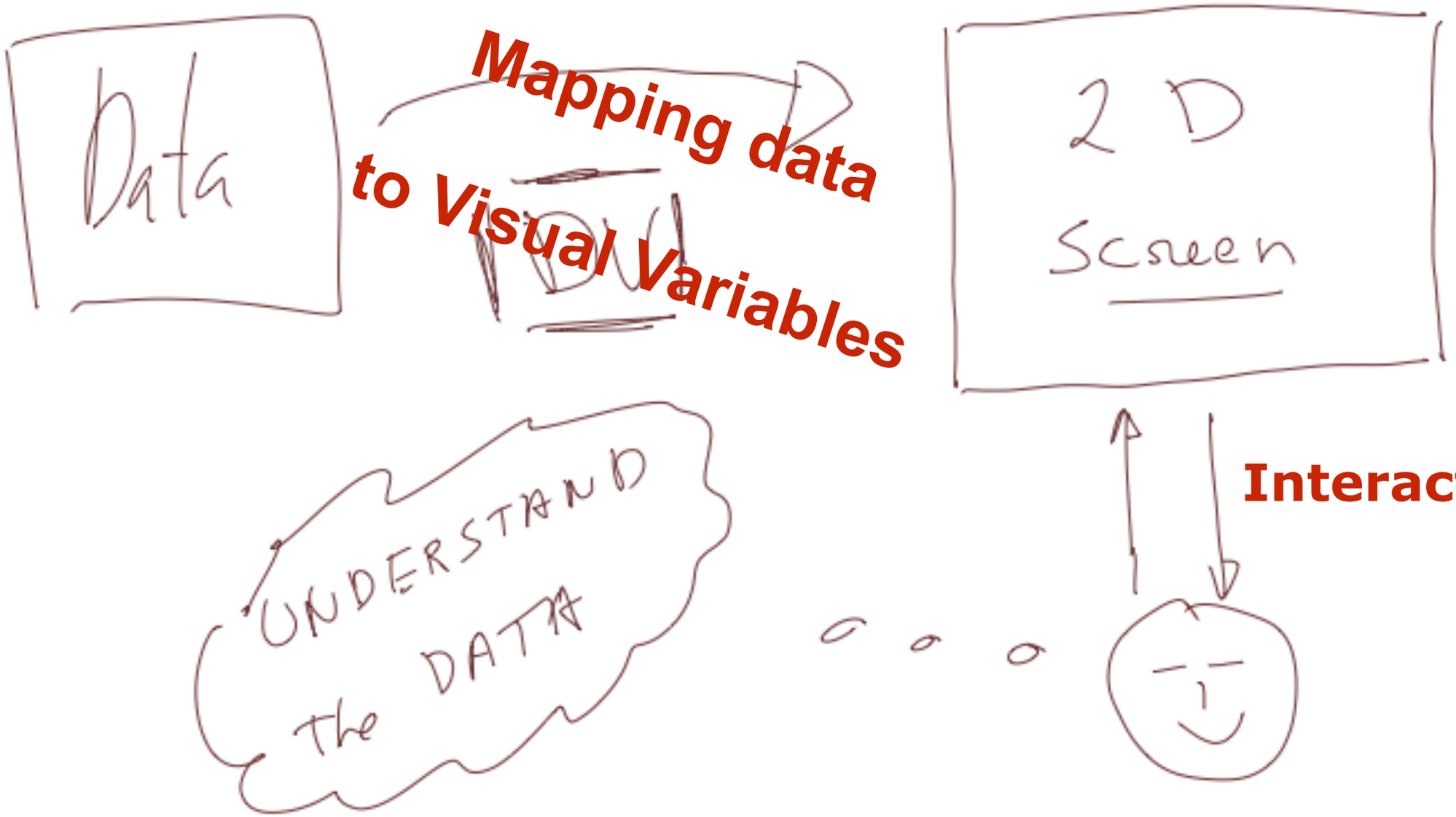


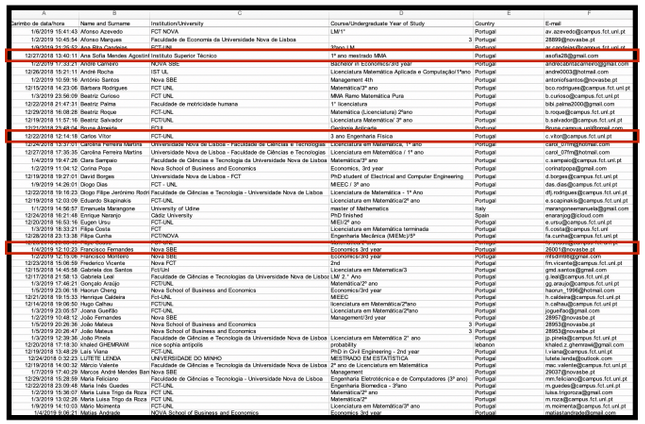

**Figure 1.5.** A search space metaphor for vis design.

# What is the core idea of Data Visualization?

# What is the core idea of Interactive Data Visualization?



# What is the core idea of Interactive Data Visualization?

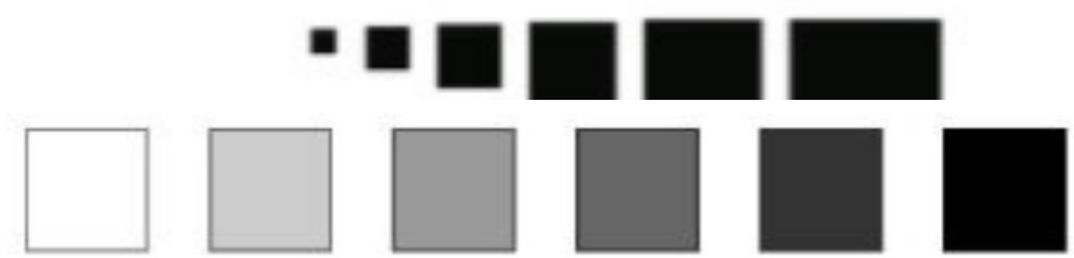


What mark?

Where?

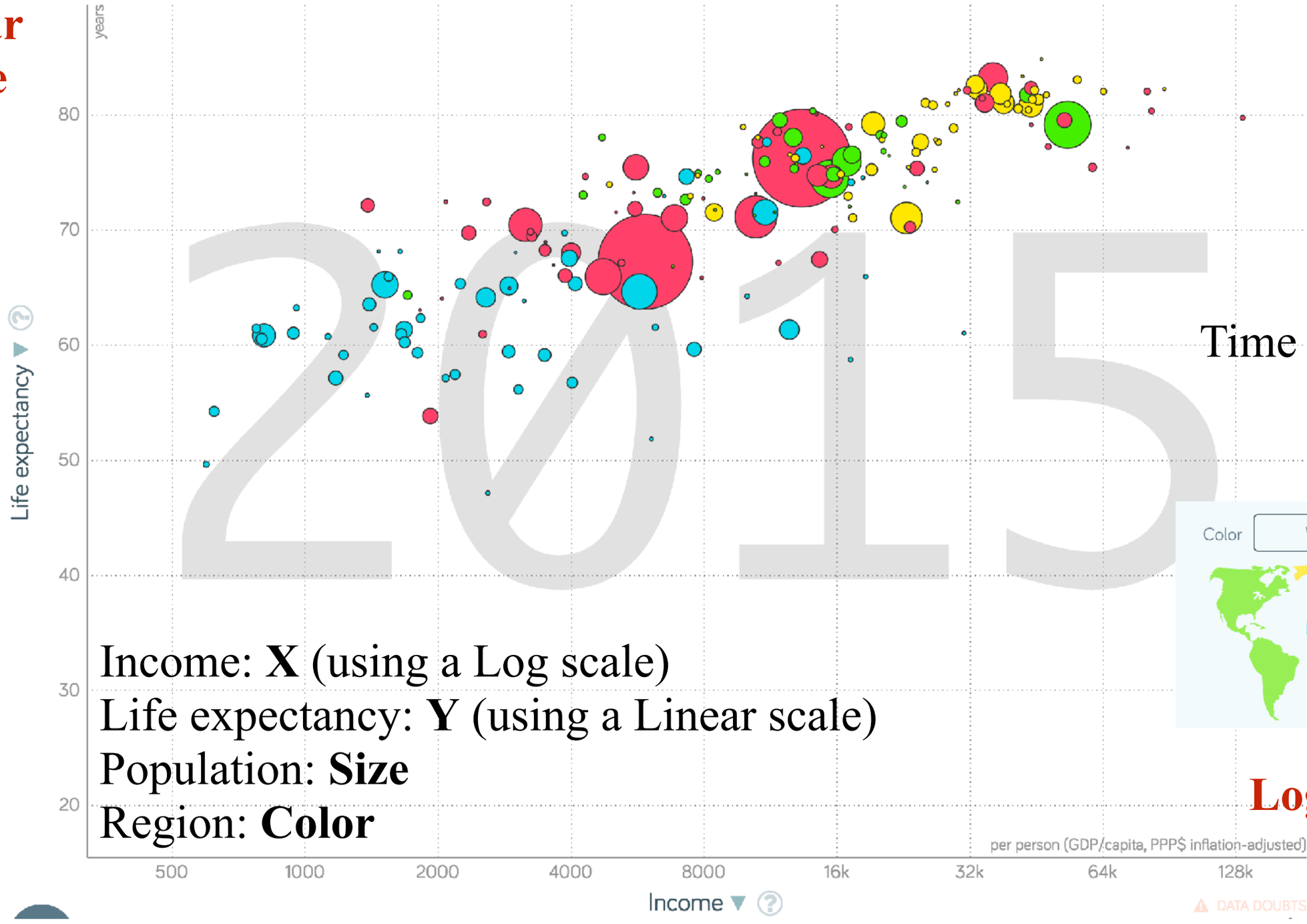
# What is the core idea of Interactive Data Visualization?

- **eight visual variables:**
  - **position,**
  - **shape,**
  - **size,**
  - **brightness,**
  - **color,**
  - **orientation,**
  - **texture,**
  - **motion**



# Eight visual variables

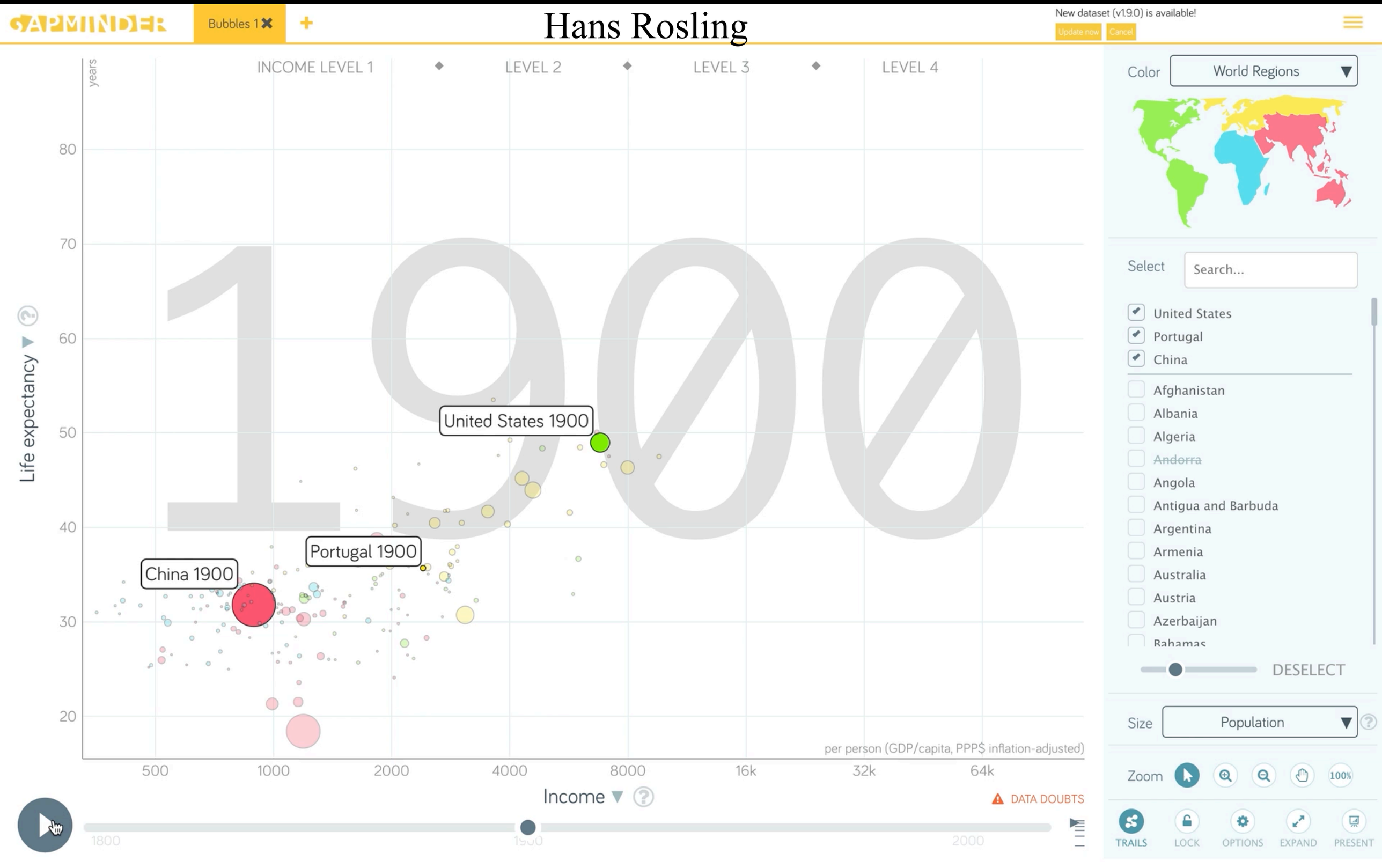
**Linear Scale**




Time (motion)

**Log Scale**

# Hans Rosling



# The Visual Variables

# Eight visual variables

- **eight visual variables:**
  - **position,**
  - **shape,**
  - **size,**
  - **brightness,**
  - **color,**
  - **orientation,**
  - **texture,**
  - **motion**

It is important to remember that the result will **be an image** that is to be interpreted by the human visual system

# Eight visual variables

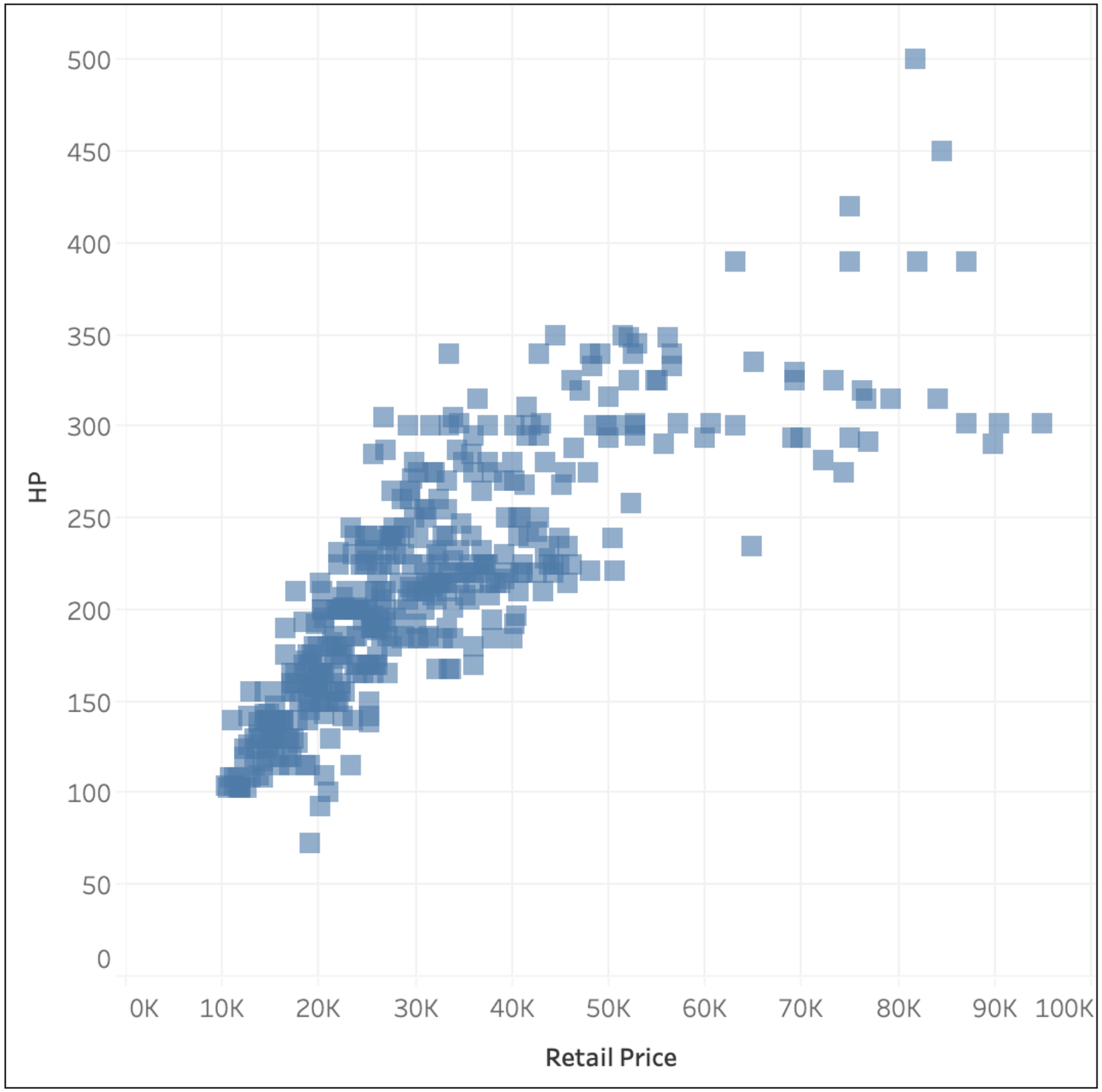
- **eight visual variables:**
  - **position,**
  - **shape,**
  - **size,**
  - **brightness,**
  - **color,**
  - **orientation,**
  - **texture,**
  - **motion**

The first and and **most important visual variable is that of position**

# Eight visual variables: Position

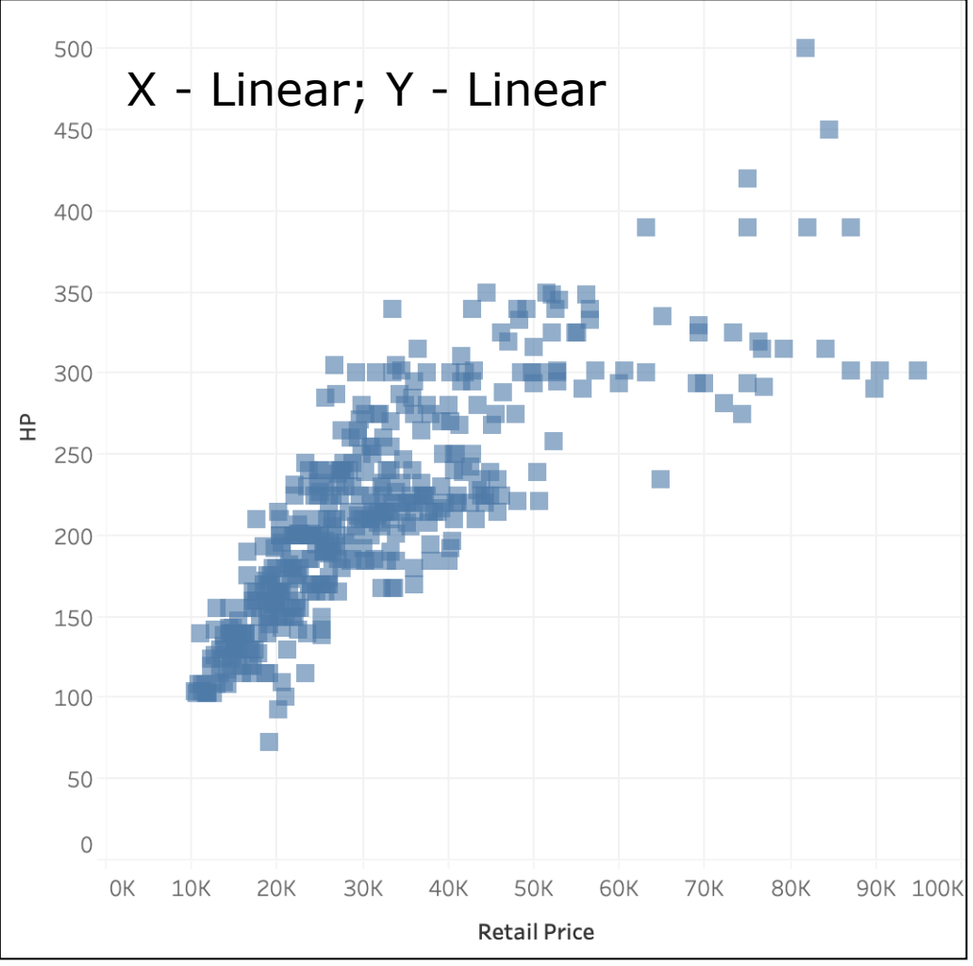
- **Spatial arrangement of graphics is the first step in reading a visualization:**
  - **The maximization of the spread of representational graphics throughout the display space maximizes the amount of information communicated, to some degree.**
  - **Worst case positioning scheme maps all graphics to the exact same position**
  - **Best positioning scheme maps each graphic to unique positions, such that all the graphics can be seen with no overlaps.**

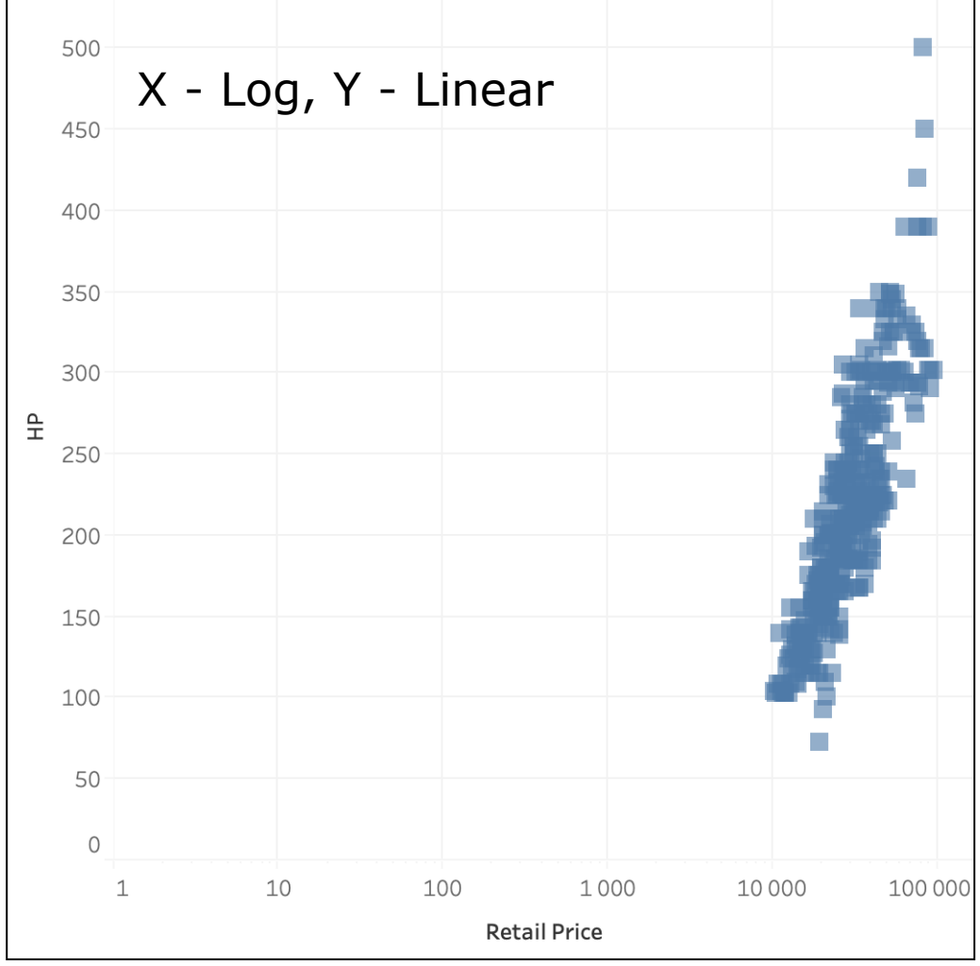
# Eight visual variables: Position - Scales

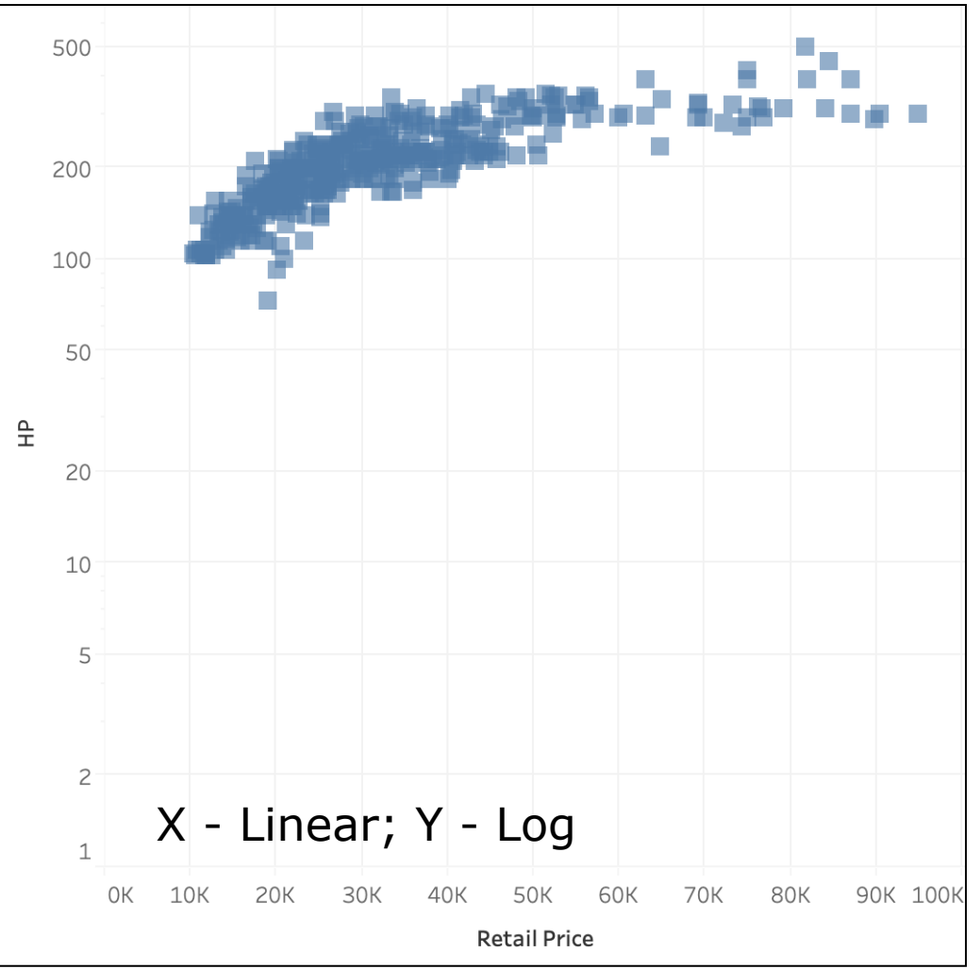


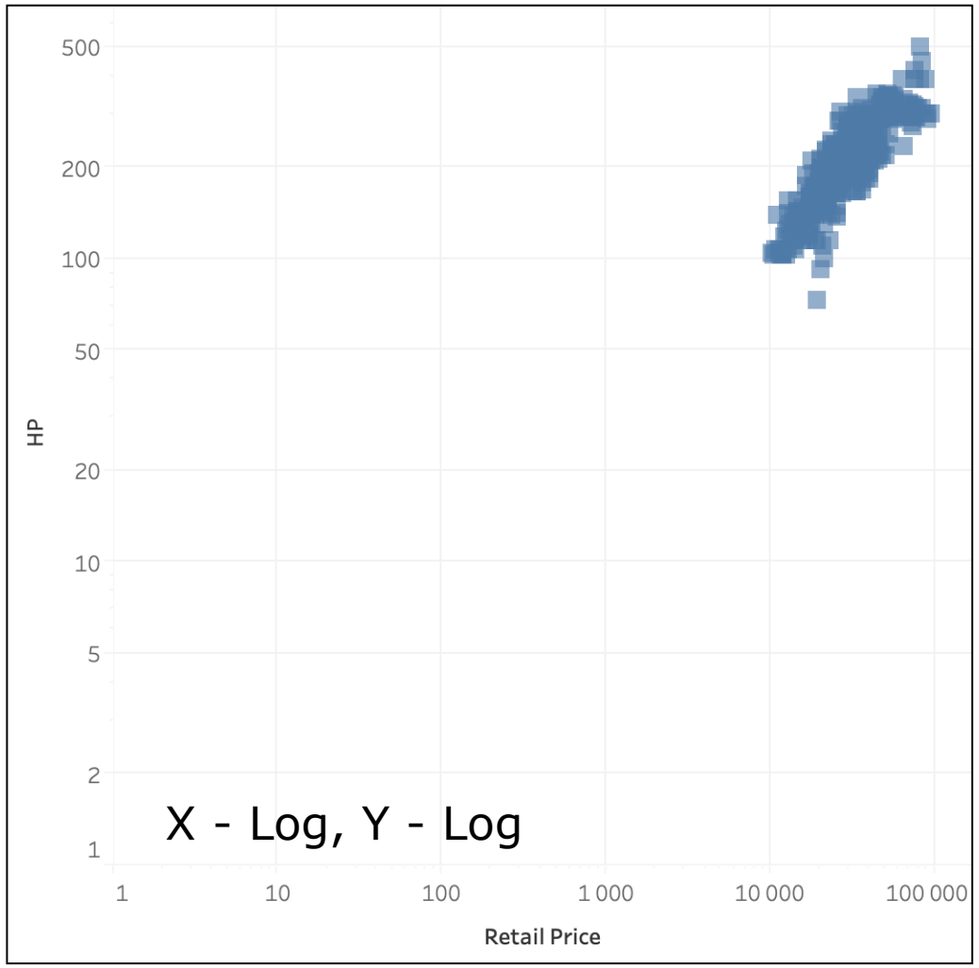

X - Retail Price;
Y - Horse Power

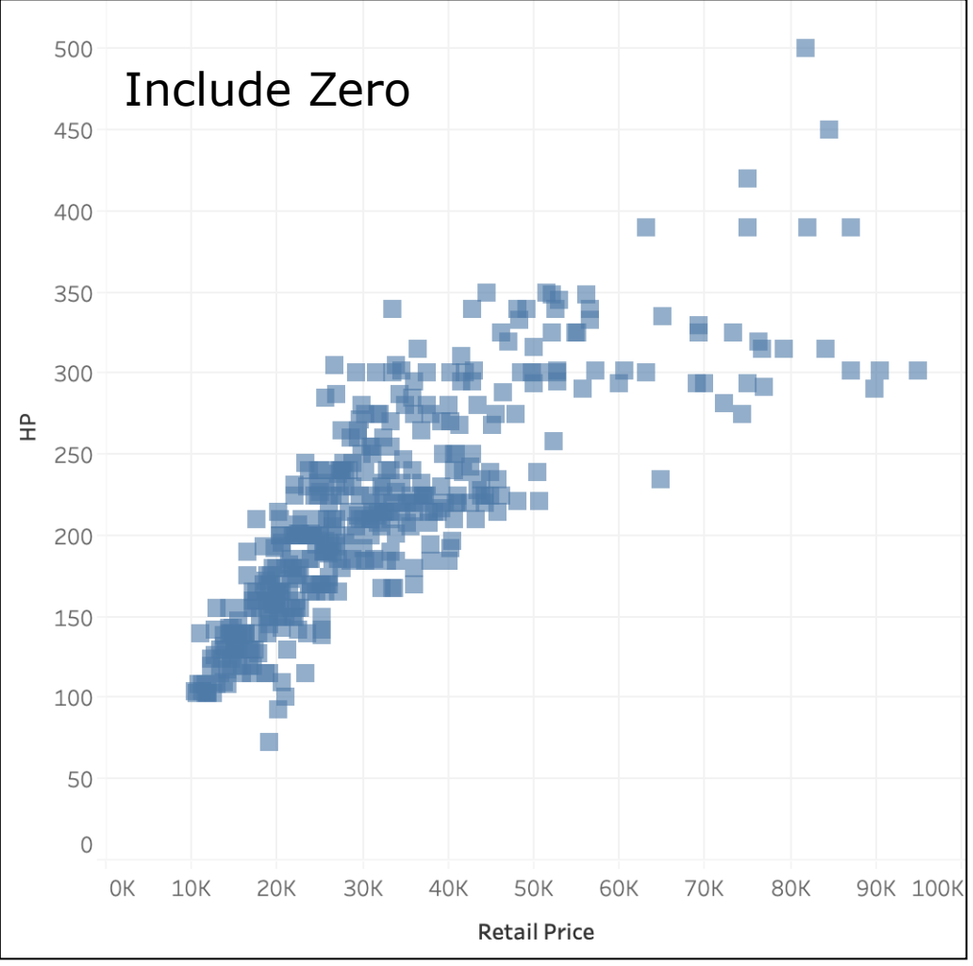
Marks - Small square


X - Linear; Y - Linear
HP
500
450
400
350
300
250
200
150
100
50
0
0K
10K
20K
30K
40K
50K
60K
70K
80K
90K
100K
Retail Price

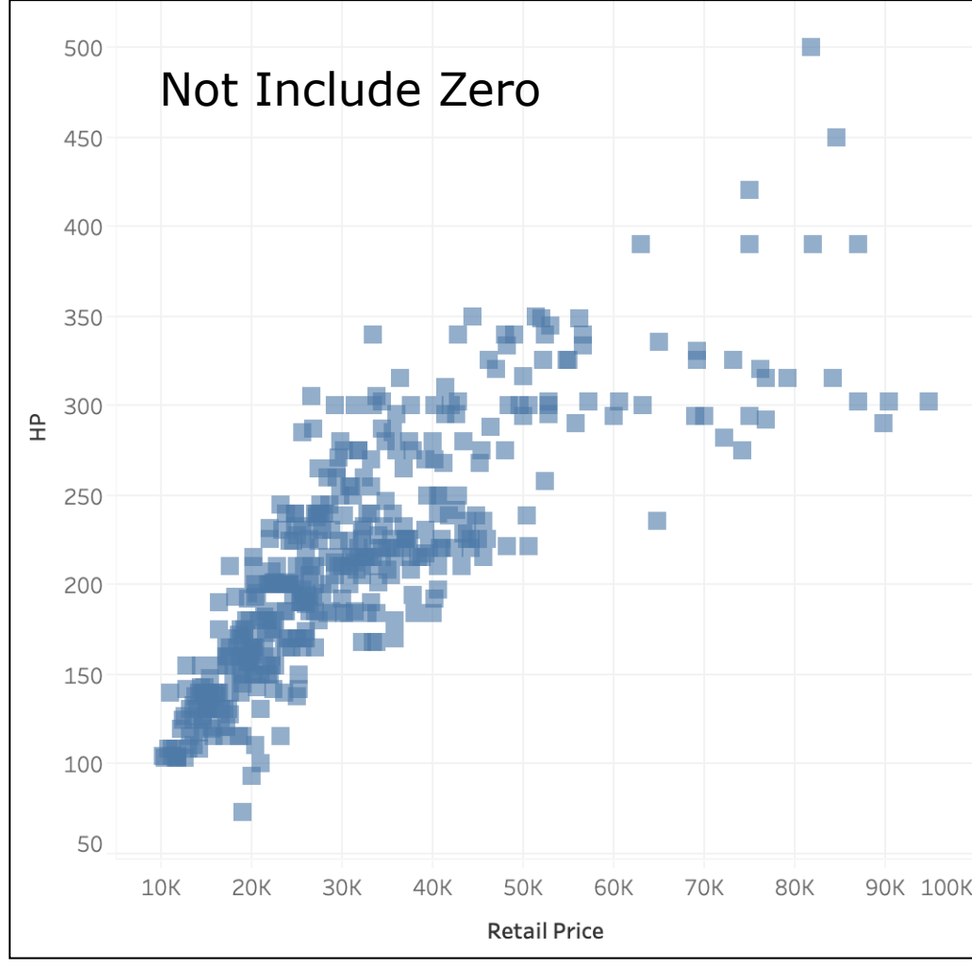

X - Log, Y - Linear
HP
500
450
400
350
300
250
200
150
100
50
0
1
10
100
1 000
10 000
100 000
Retail Price

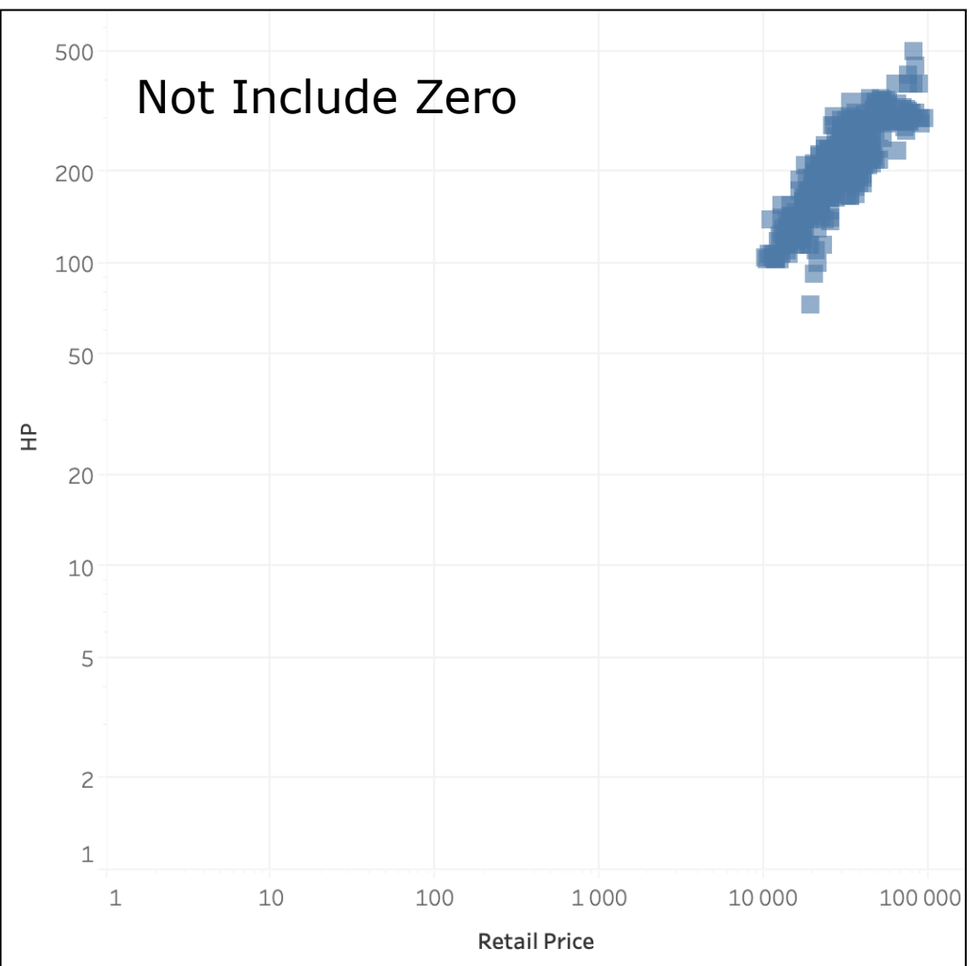

HP
500
200
100
50
20
10
5
2
1
X - Linear; Y - Log
0K
10K
20K
30K
40K
50K
60K
70K
80K
90K
100K
Retail Price


HP
500
200
100
50
20
10
5
2
1
X - Log, Y - Log
1
10
100
1 000
10 000
100 000
Retail Price

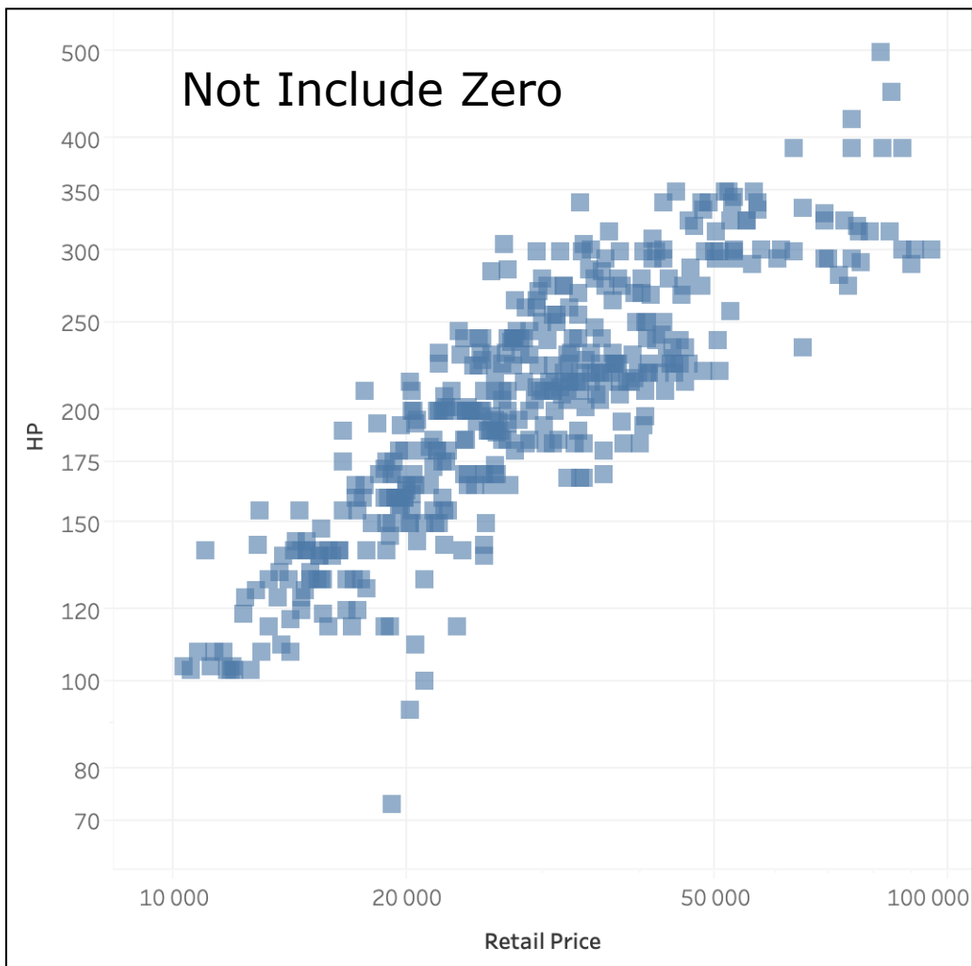



X - Linear; Y - Linear

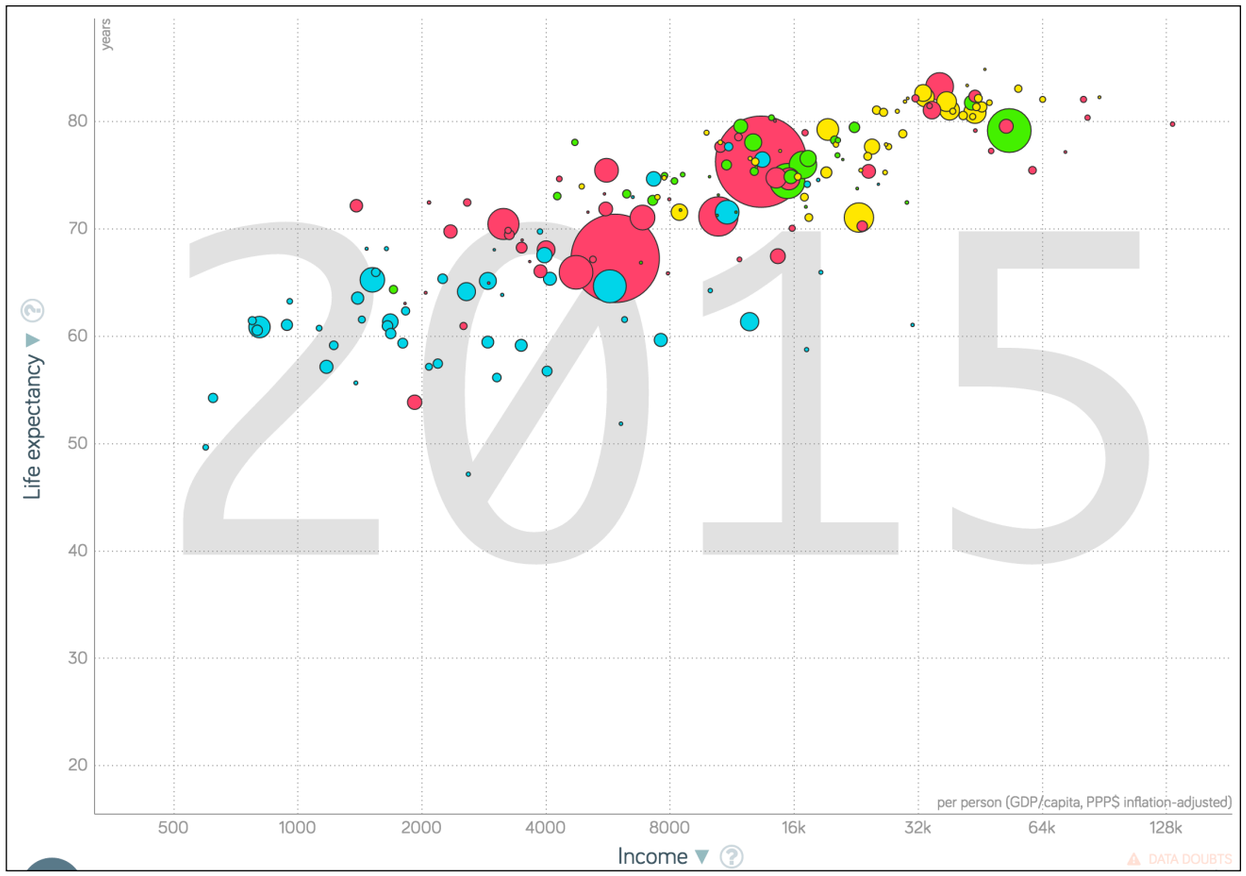


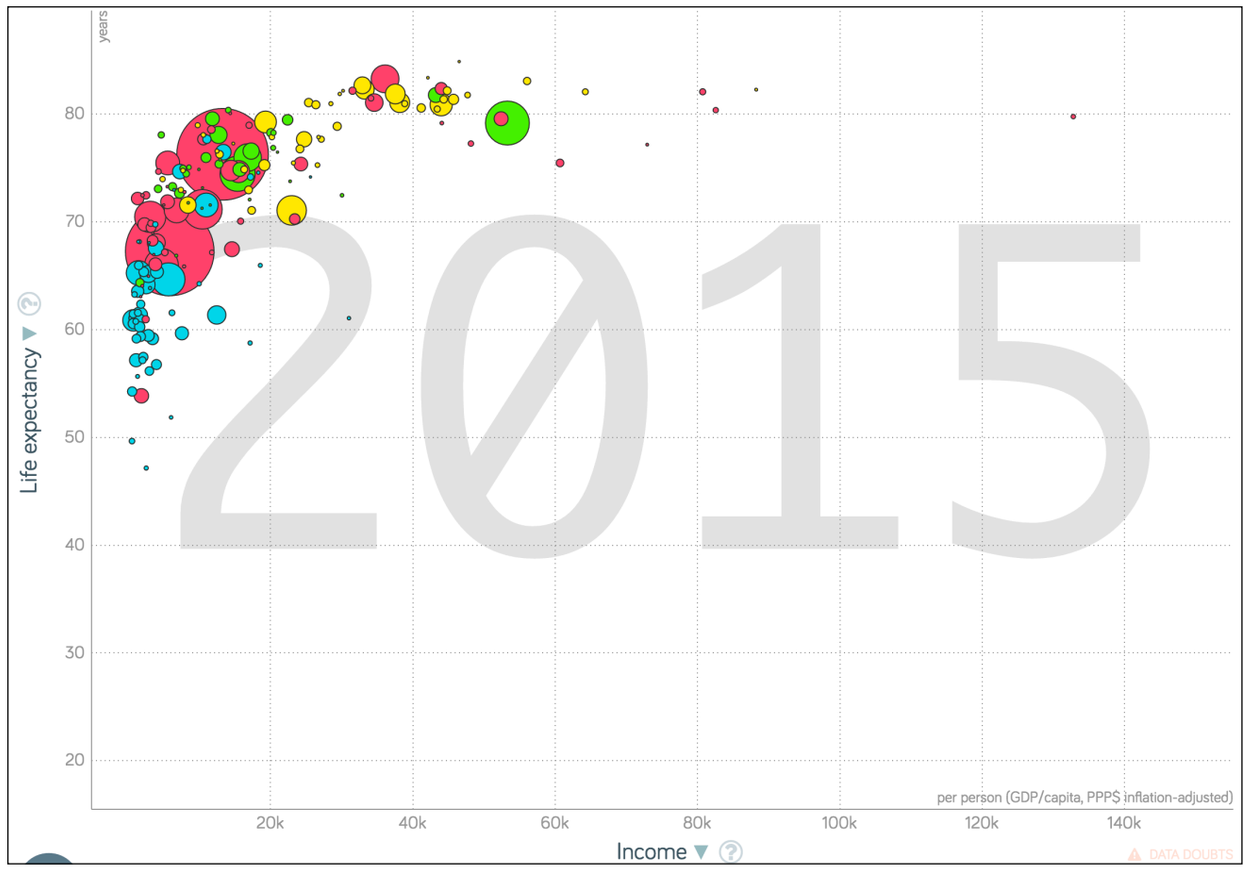

X - Log; Y - Log



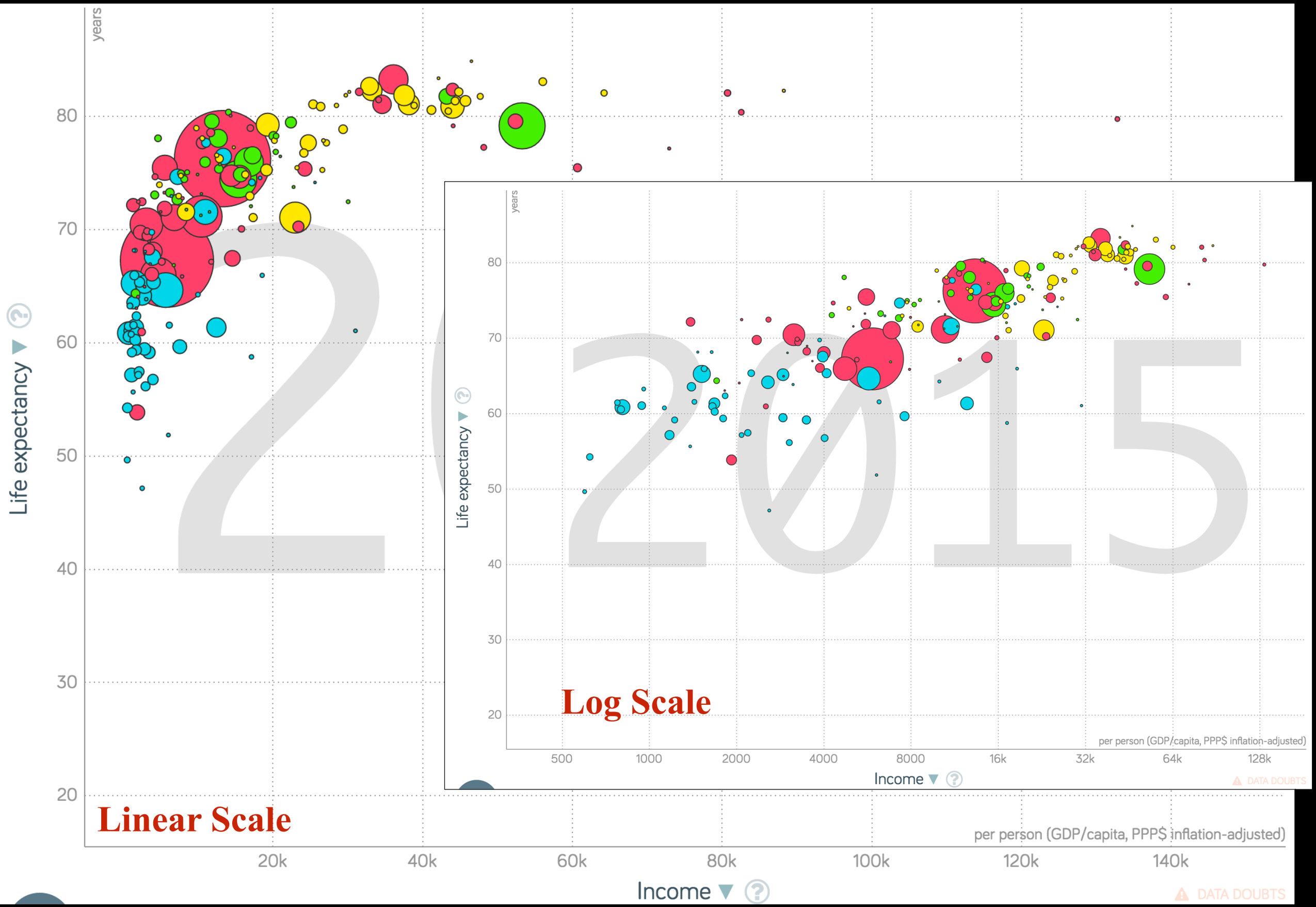
# Eight visual variables: When Log is appropriate




**Log Scale**




**Linear Scale**


years
80
70
60
50
40
30
20
Life expectancy
2015
Linear Scale
20k
40k
60k
80k
100k
120k
140k
per person (GDP/capita, PPP$ inflation-adjusted)
Income
DATA DOUBTS
Log Scale
500
1000
2000
4000
8000
16k
32k
64k
128k

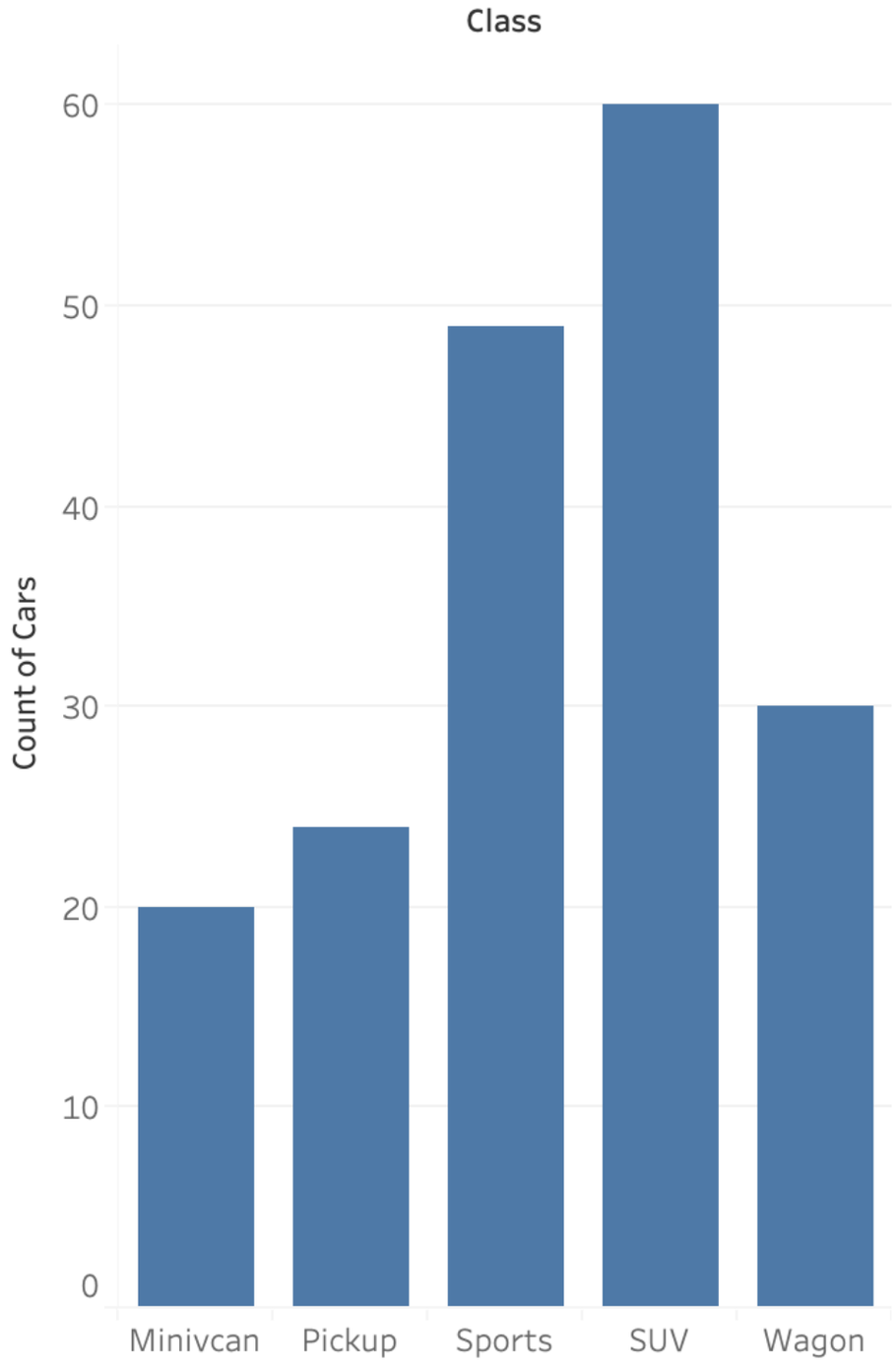
# Eight visual variables: Position - Scales

Y - Count of Cars

X - Class (excluding “Normal”)

Equally spaced;
Sorted alphabetically

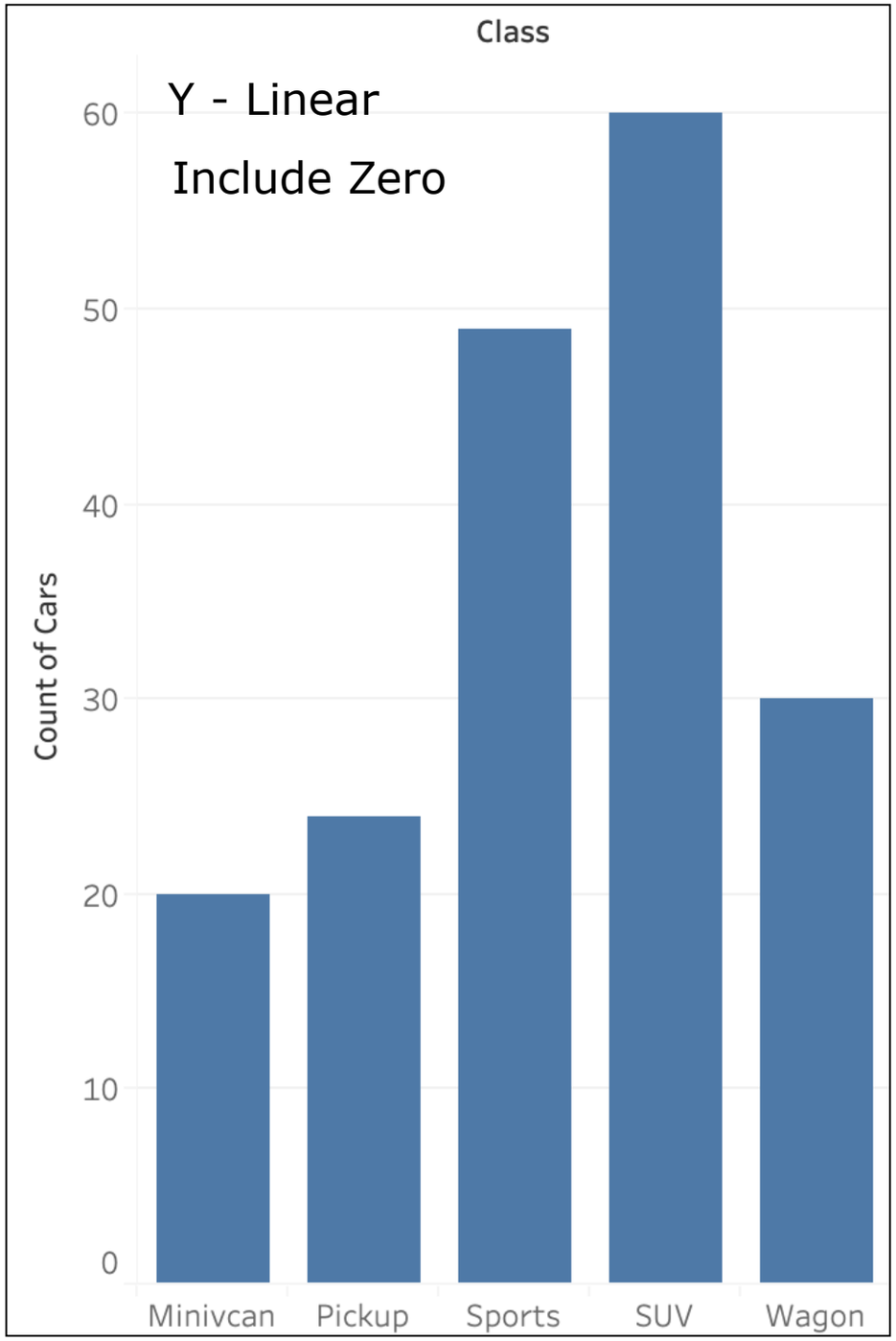
Marks - Bars

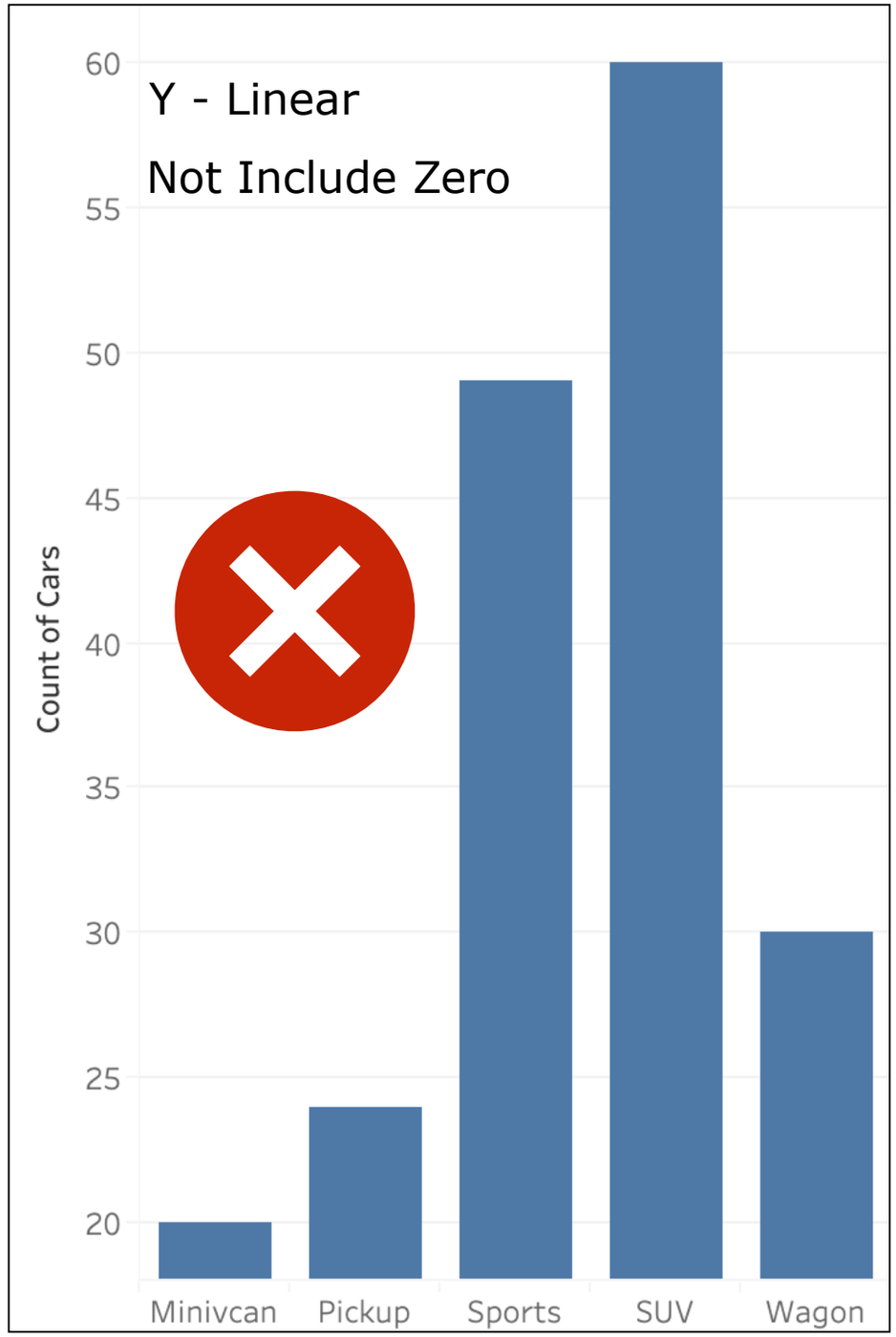


# Eight visual variables: Position - Scales

Y - Count of Cars

X - Class (excluding “Normal”)

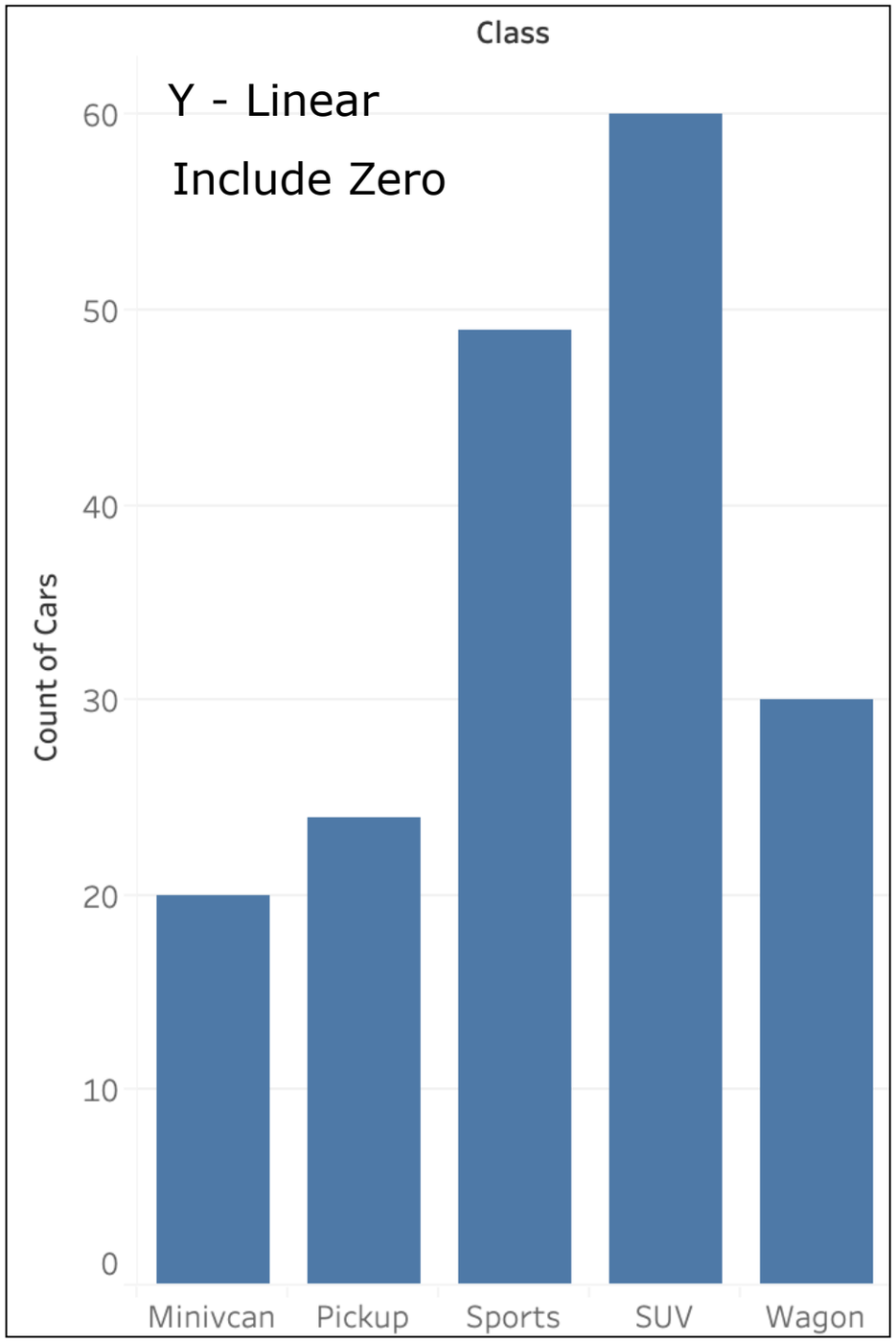


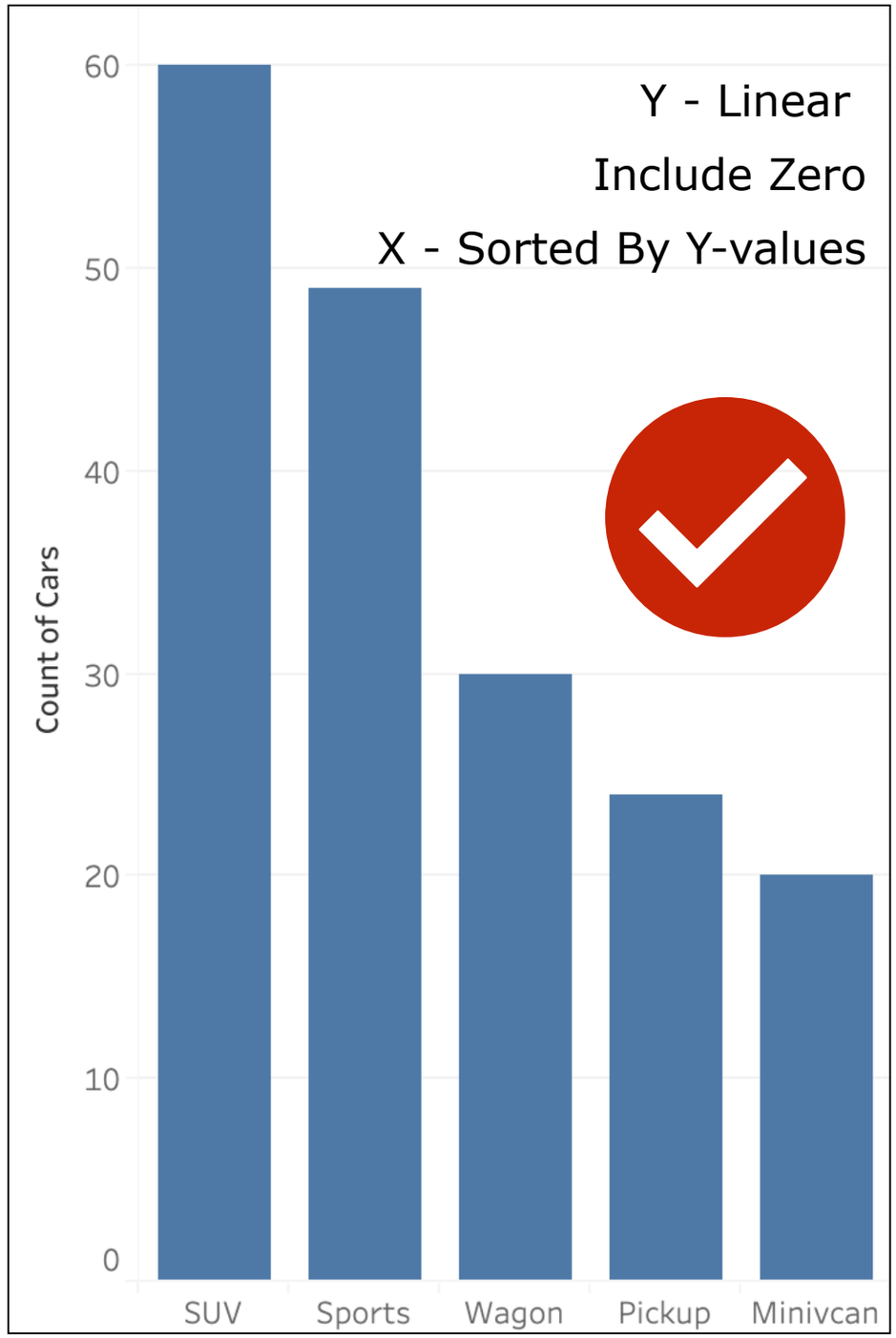



# Eight visual variables: Position - Scales

Y - Count of Cars

X - Class (excluding “Normal”)

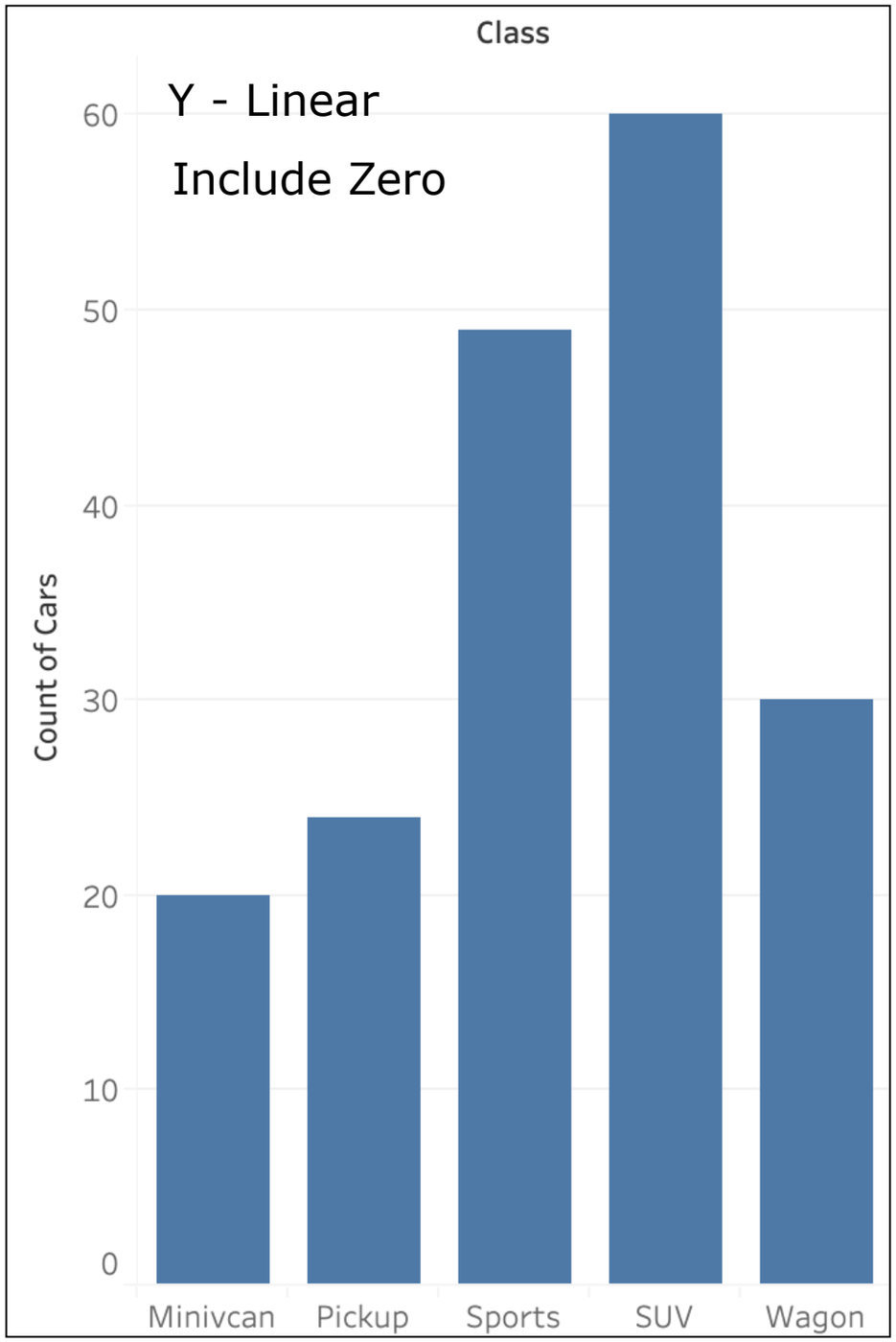


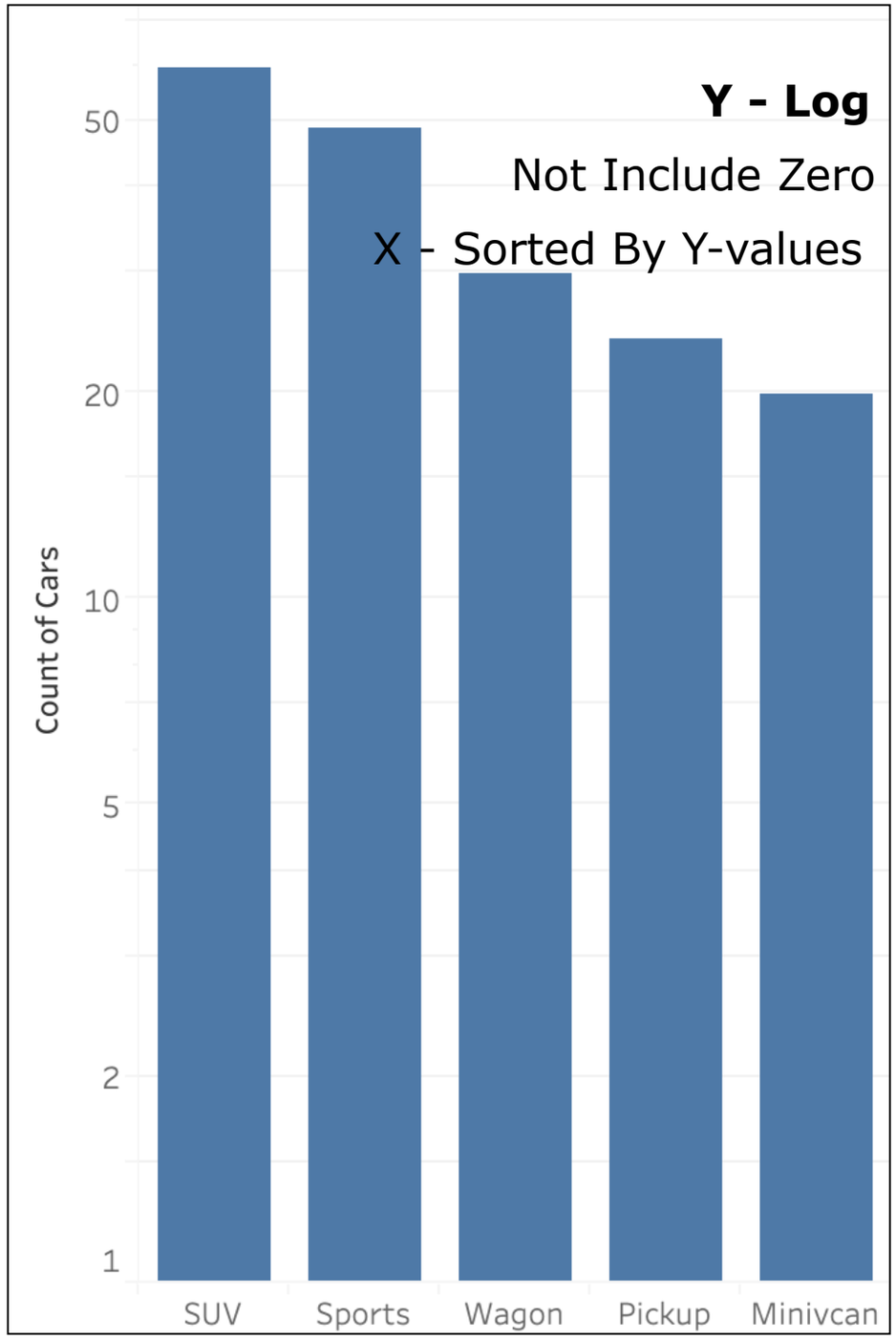



# Eight visual variables: Position - Scales

Y - Count of Cars

X - Class (excluding “Normal”)



# Eight visual variables: Position - Splitting the views

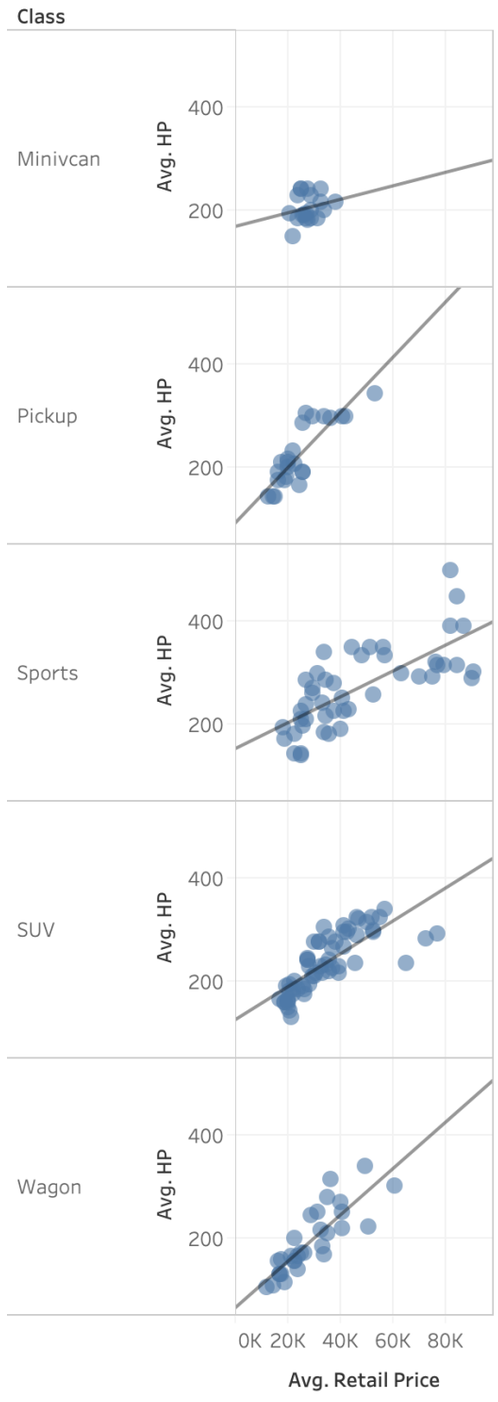

X - Average of Retail Price

Y - Average of HP

**X - Split views by Class**

Decorate with a trend line

# Eight visual variables: Position - Splitting the views



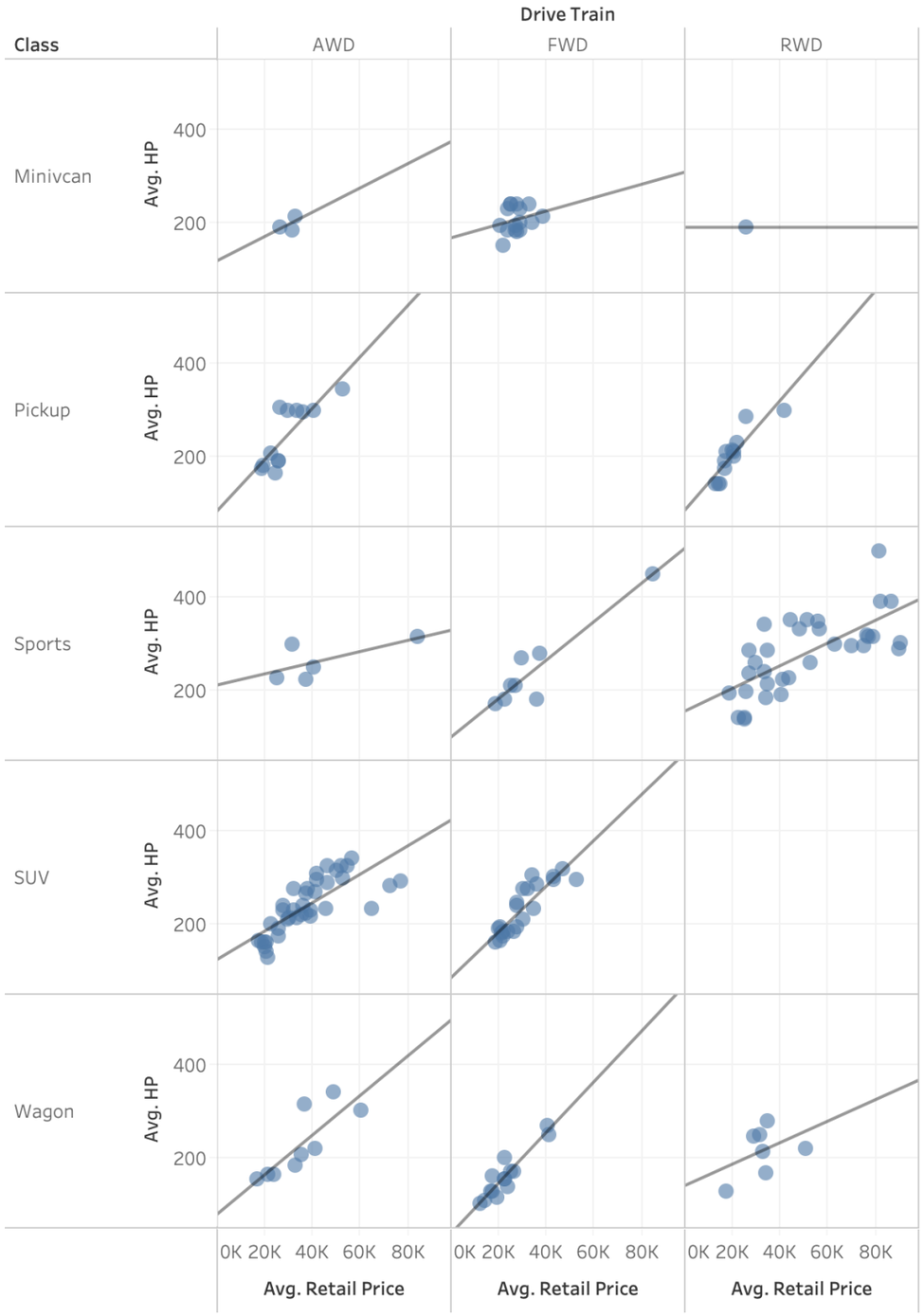

X - Average of Retail Price

Y - Average of HP

**Y - Split views by Class**

Decorate with a trend line

# Eight visual variables: Position - Splitting the views



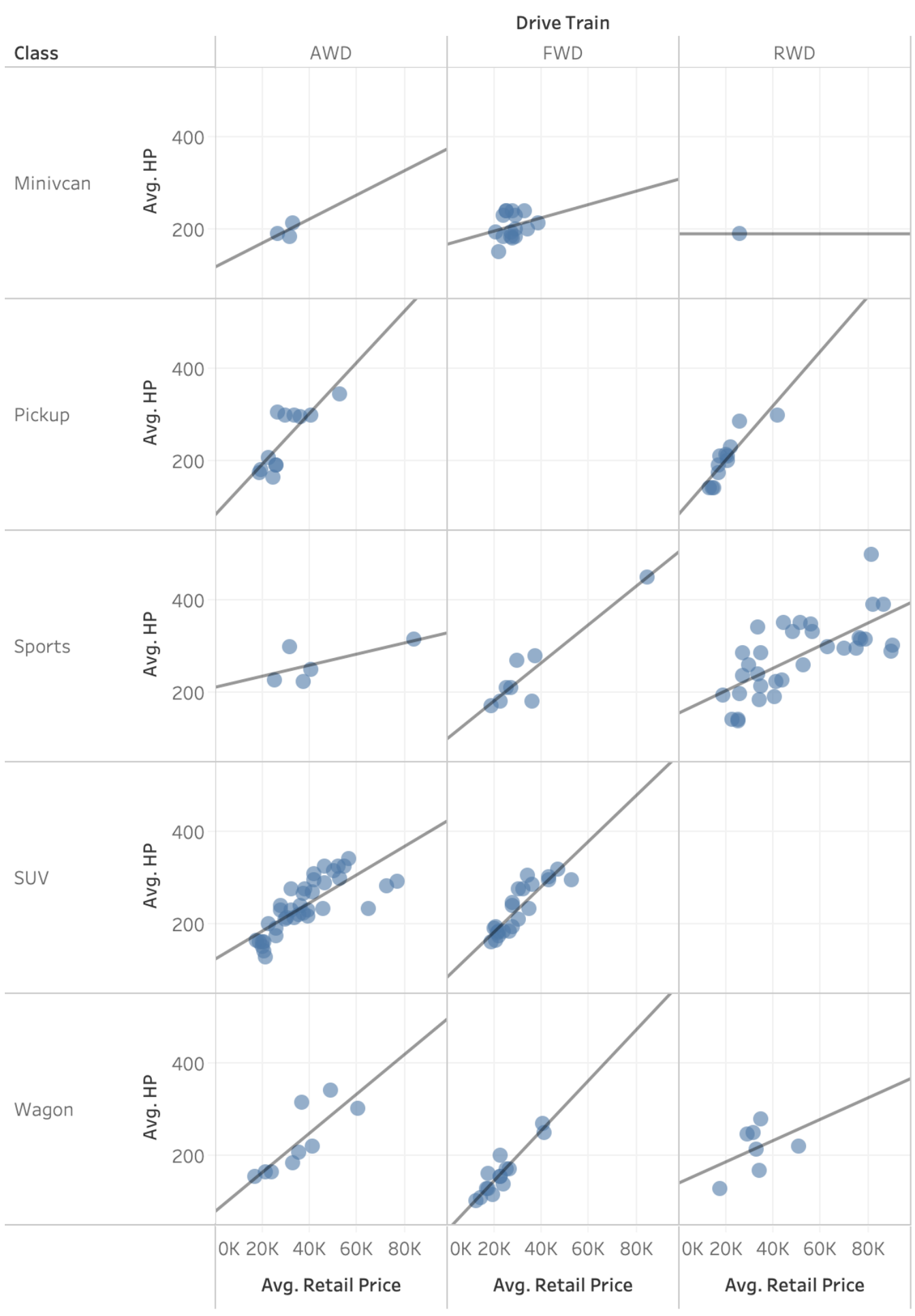

X - Average of Retail Price

Y - Average of HP

**X - Split views by Drive Train**

**Y - Split views by Class**

Decorate with a trend line




X - Average of Retail Price

Y - Average of HP

**X - Split views by Drive Train**

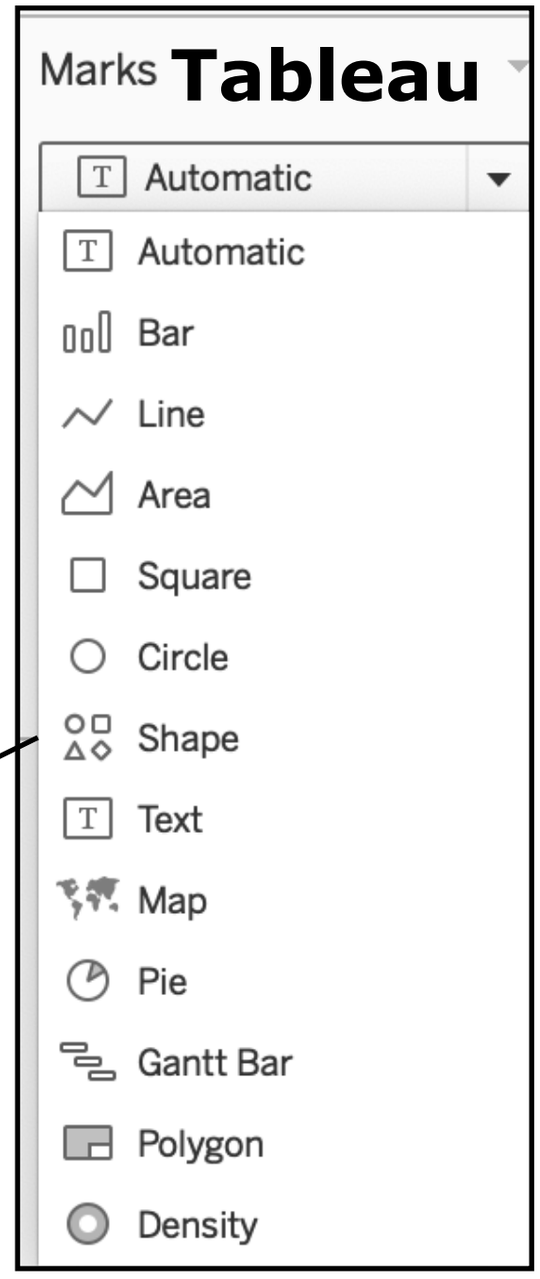
**Y - Split views by Class**

Decorate with a trend line

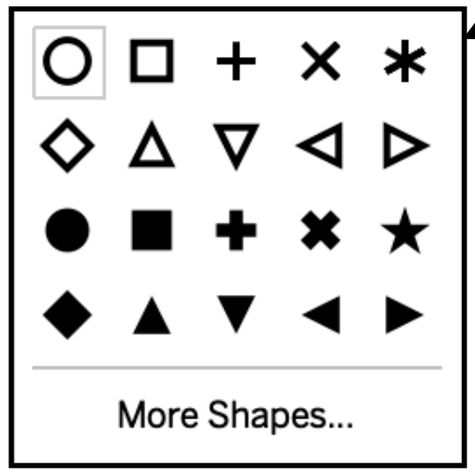
# Eight visual variables: Mark (or shape)

- **The second visual variable is the mark or shape: points, lines, areas, volumes, and their compositions.**
- **Marks are graphic primitives that represent data:**

Several examples of different marks or glyphs that can be used.




- When using marks, it is important to consider how well one mark can be differentiated from other marks



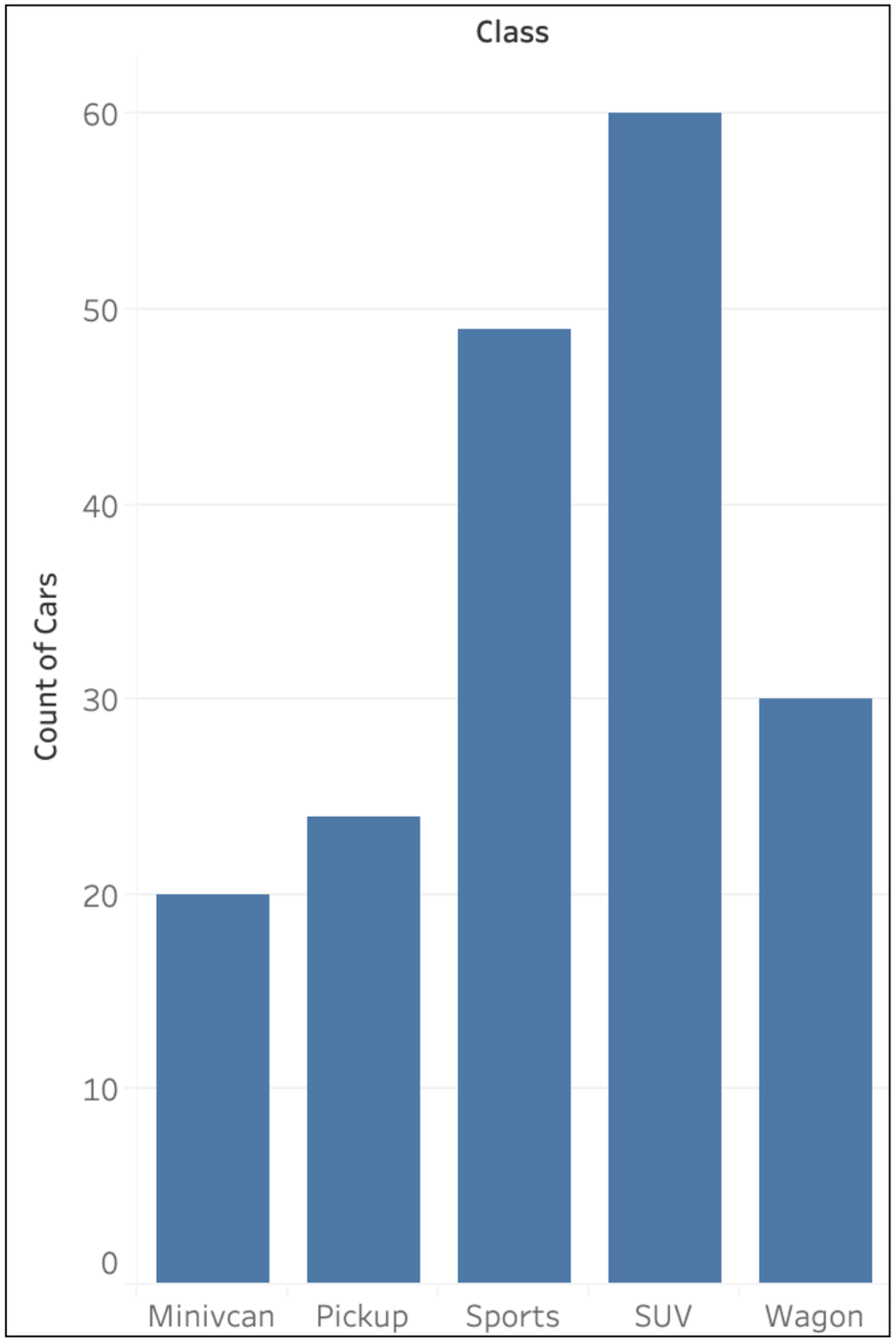
# Eight visual variables: Mark (or shape)

Y - Count of Cars

X - Class (excluding “Normal”)

Equally spaced;
Sorted alphabetically

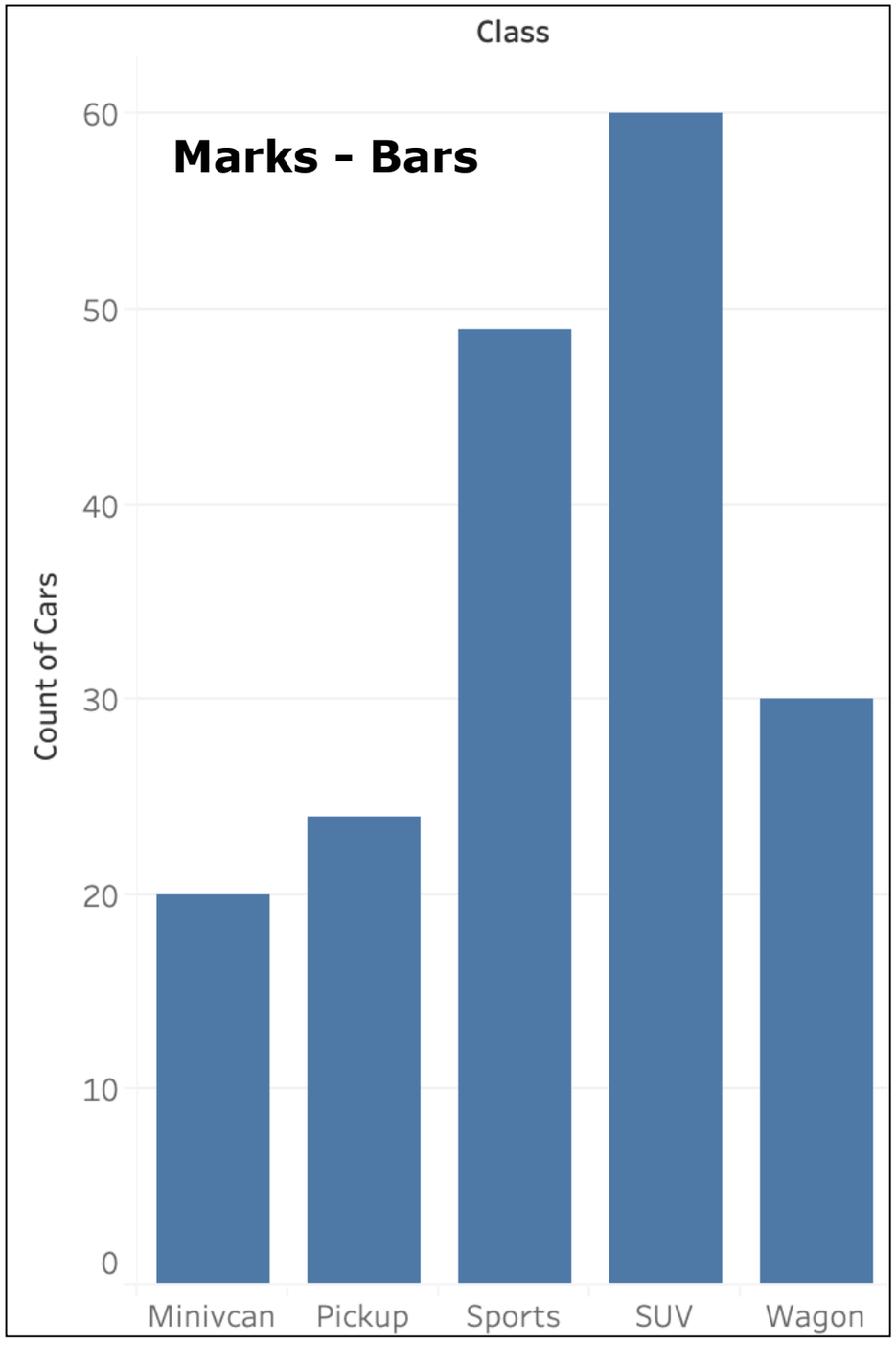
**Marks - Bars**

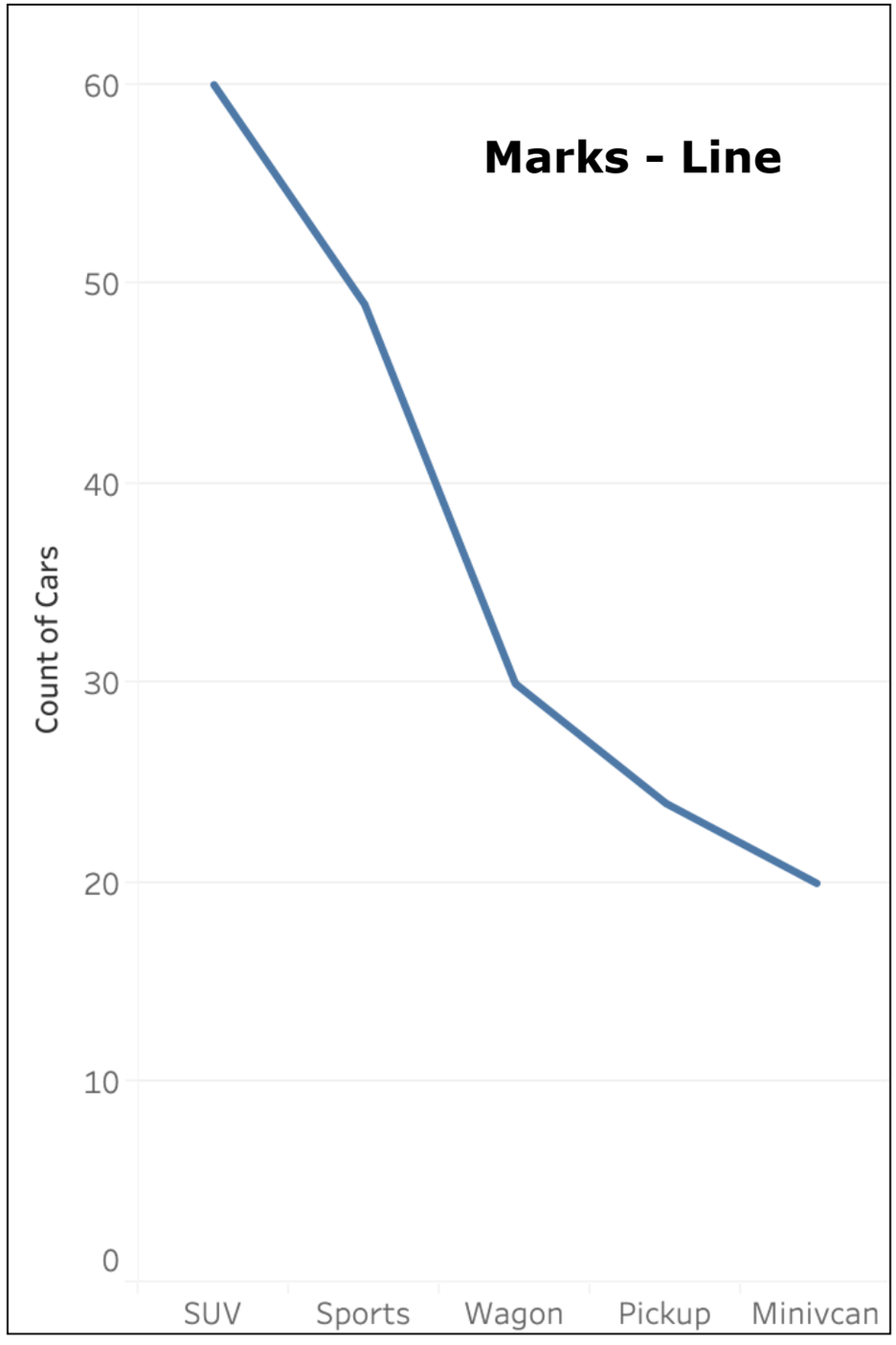


# Eight visual variables: Mark (or shape)

Y - Count of Cars

X - Class (excluding “Normal”)

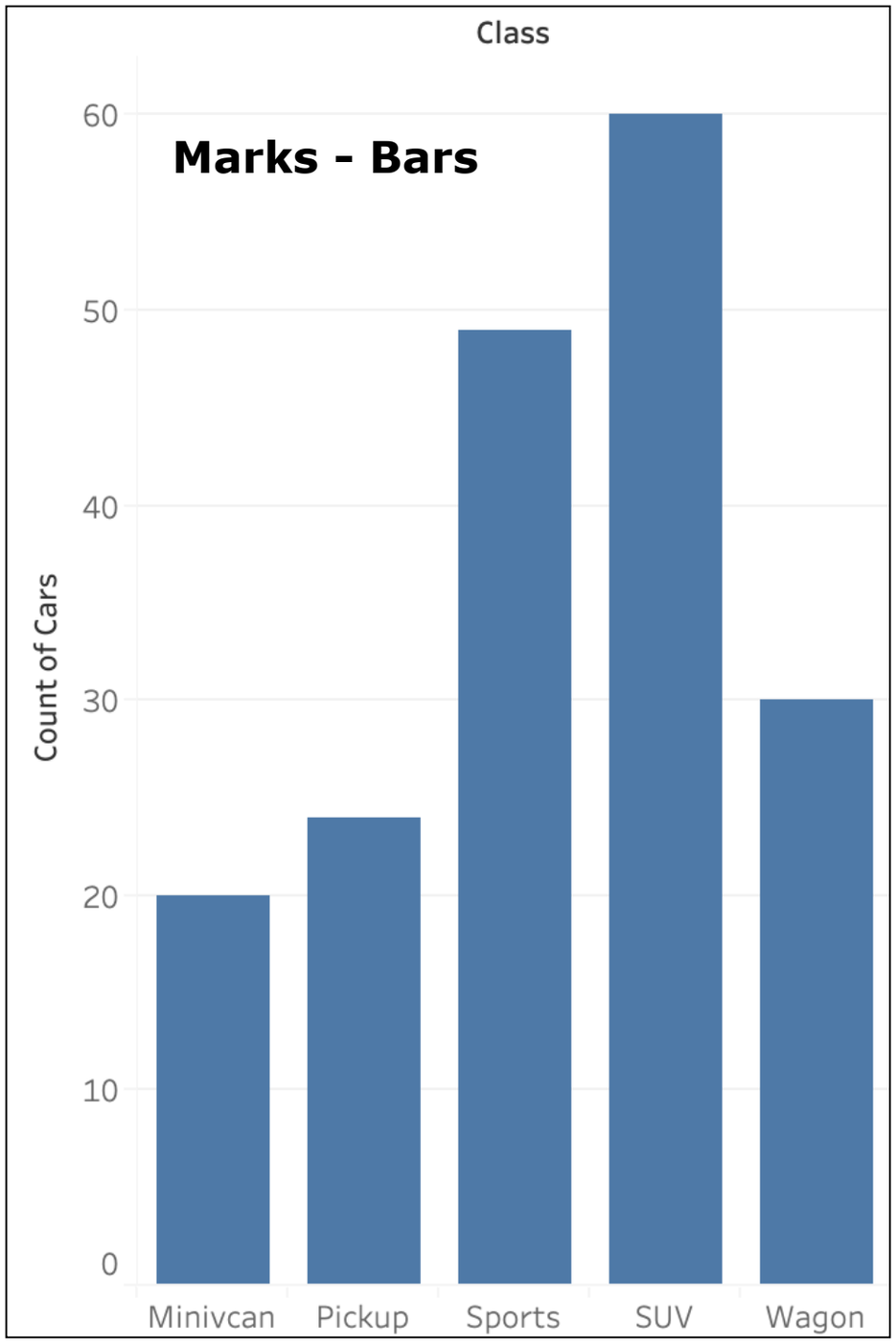


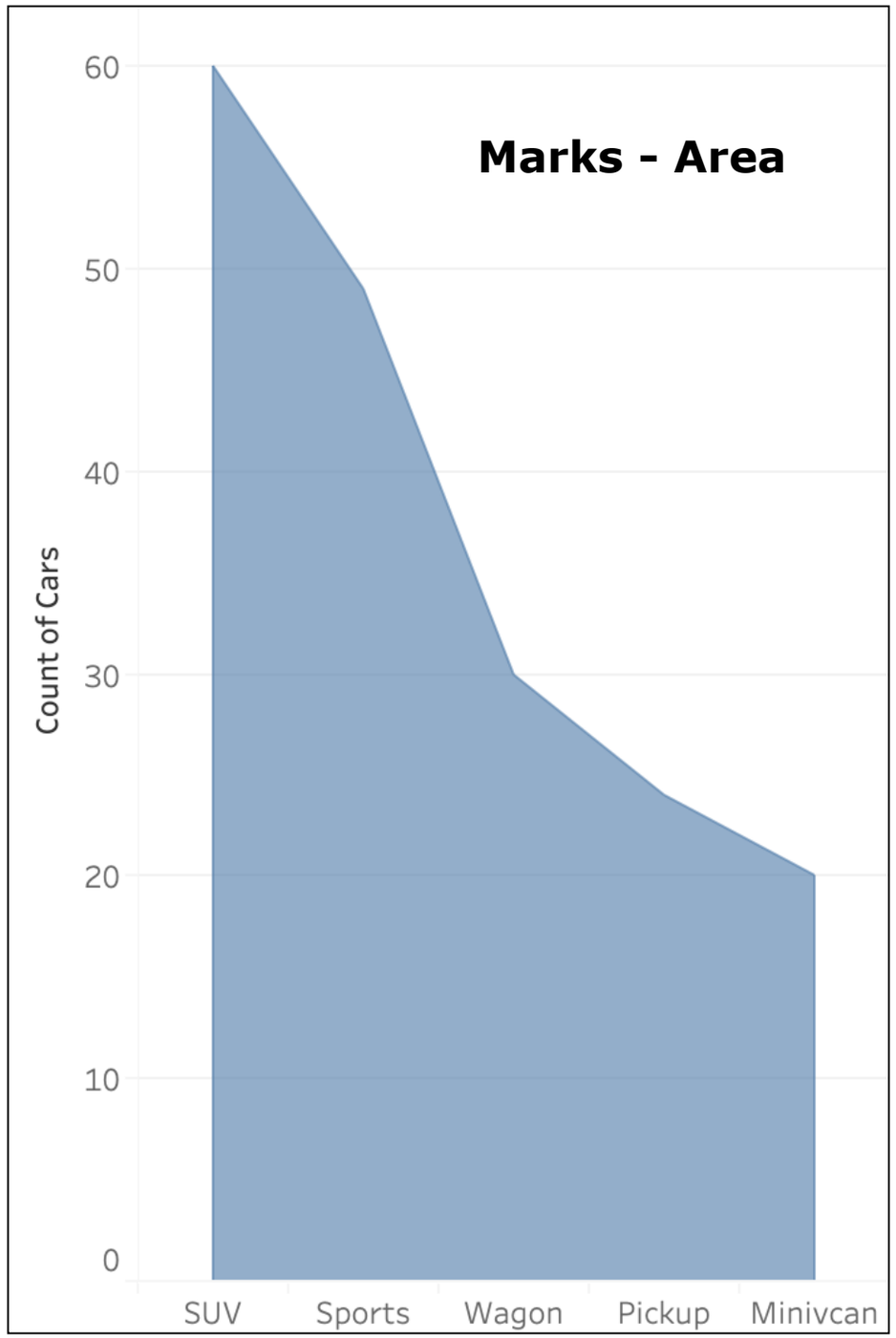



# Eight visual variables: Mark (or shape)

Y - Count of Cars

X - Class (excluding “Normal”)

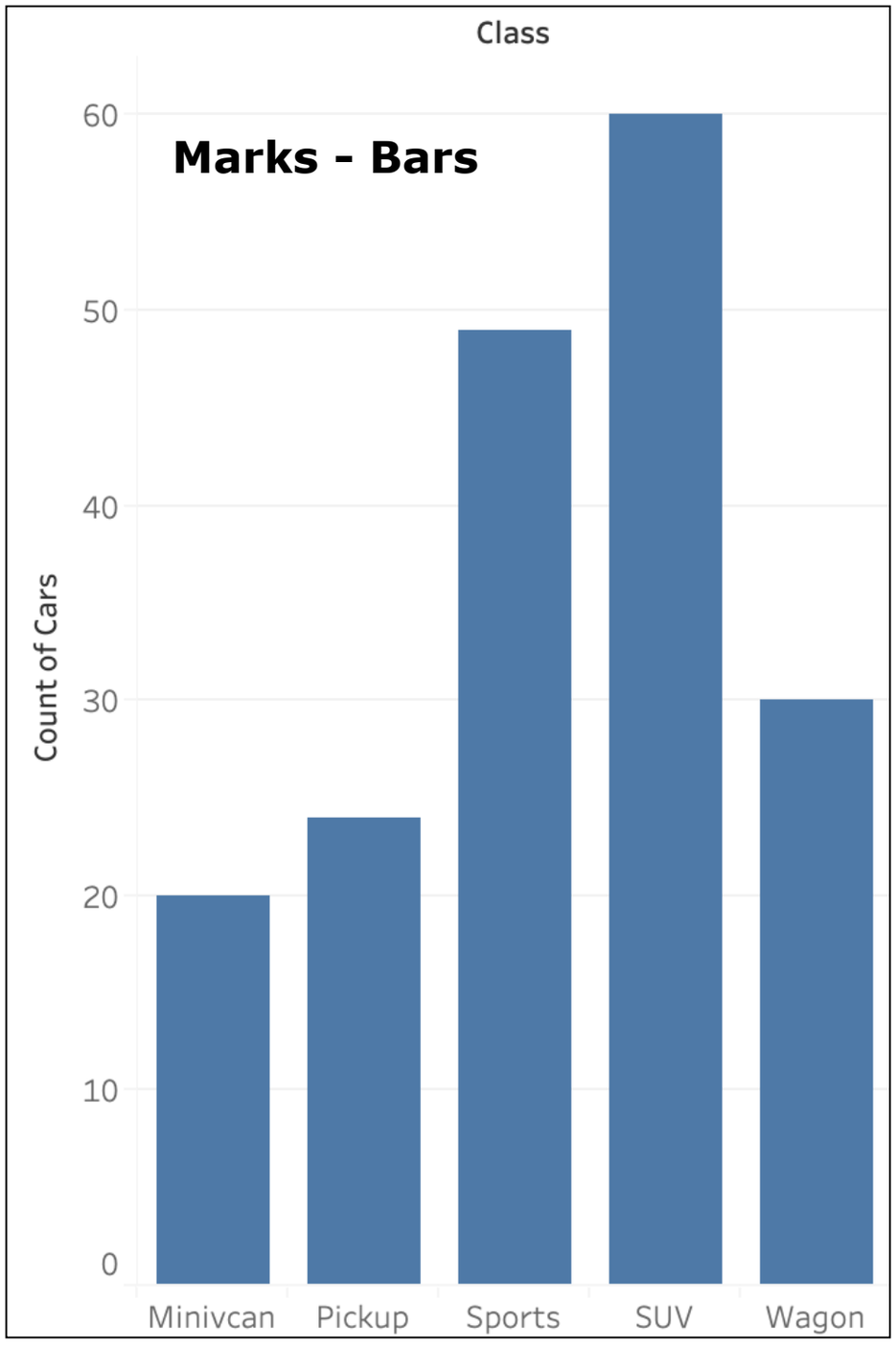


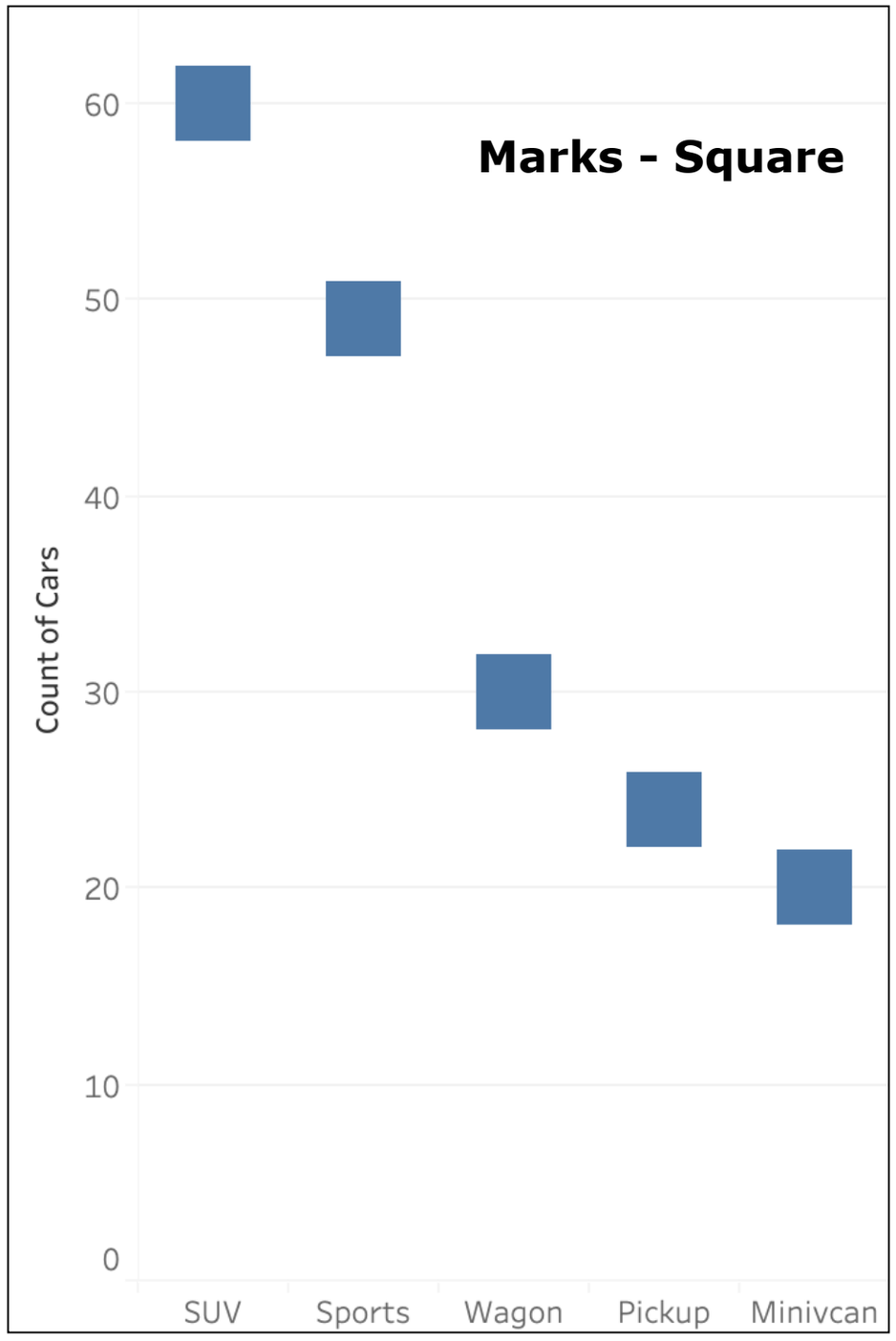



# Eight visual variables: Mark (or shape)

Y - Count of Cars

X - Class (excluding “Normal”)

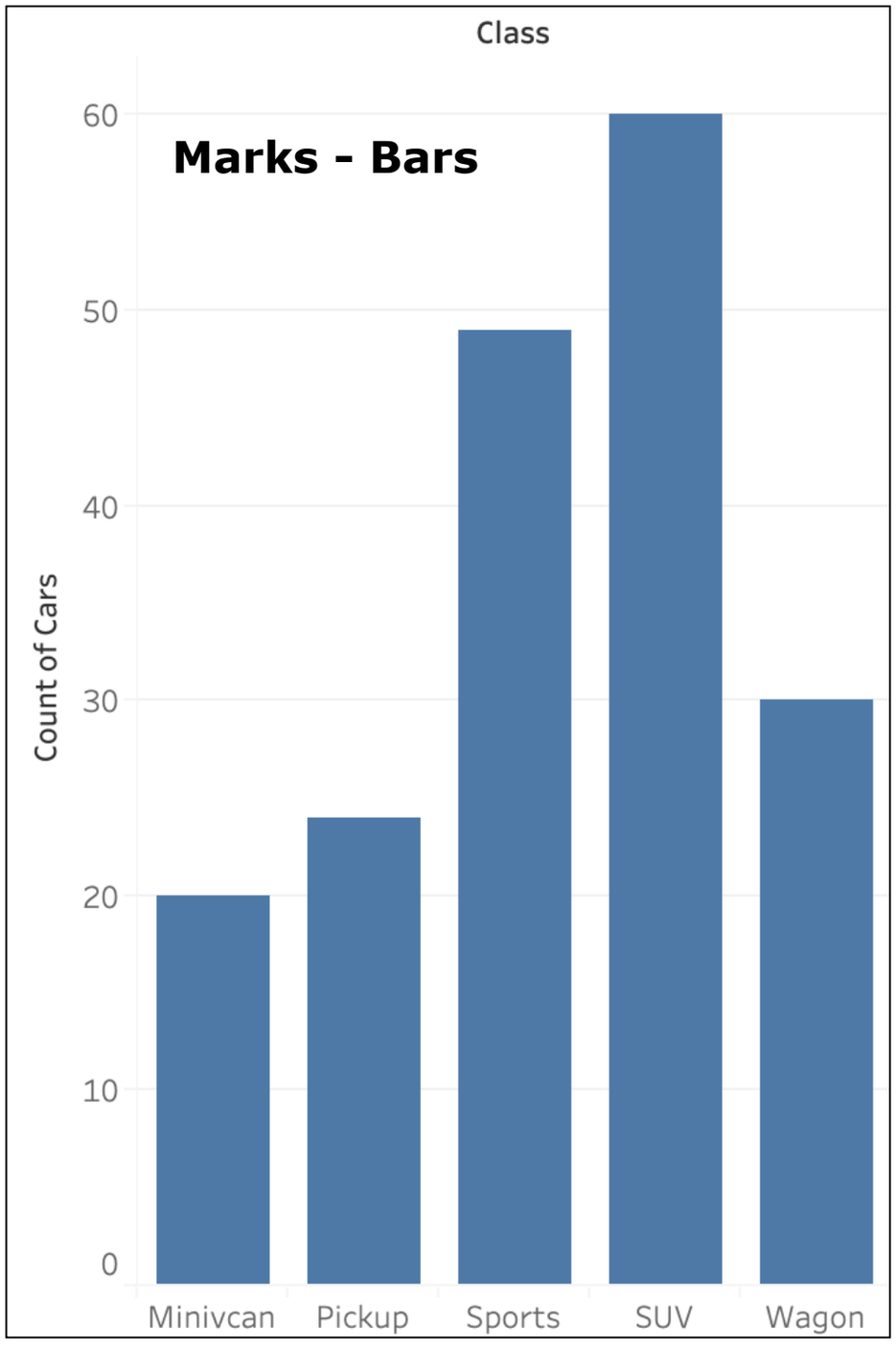


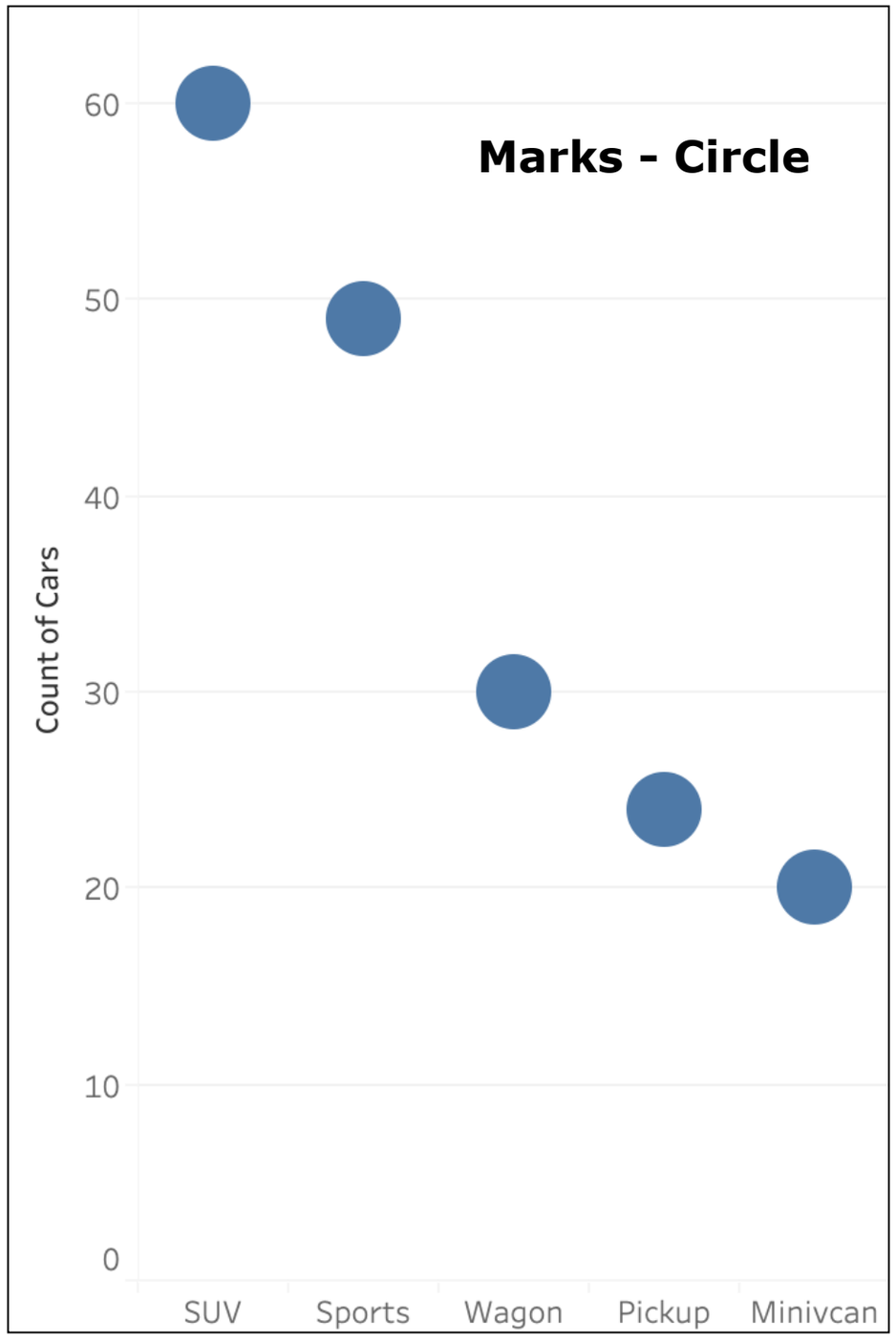



# Eight visual variables: Mark (or shape)

Y - Count of Cars

X - Class (excluding “Normal”)

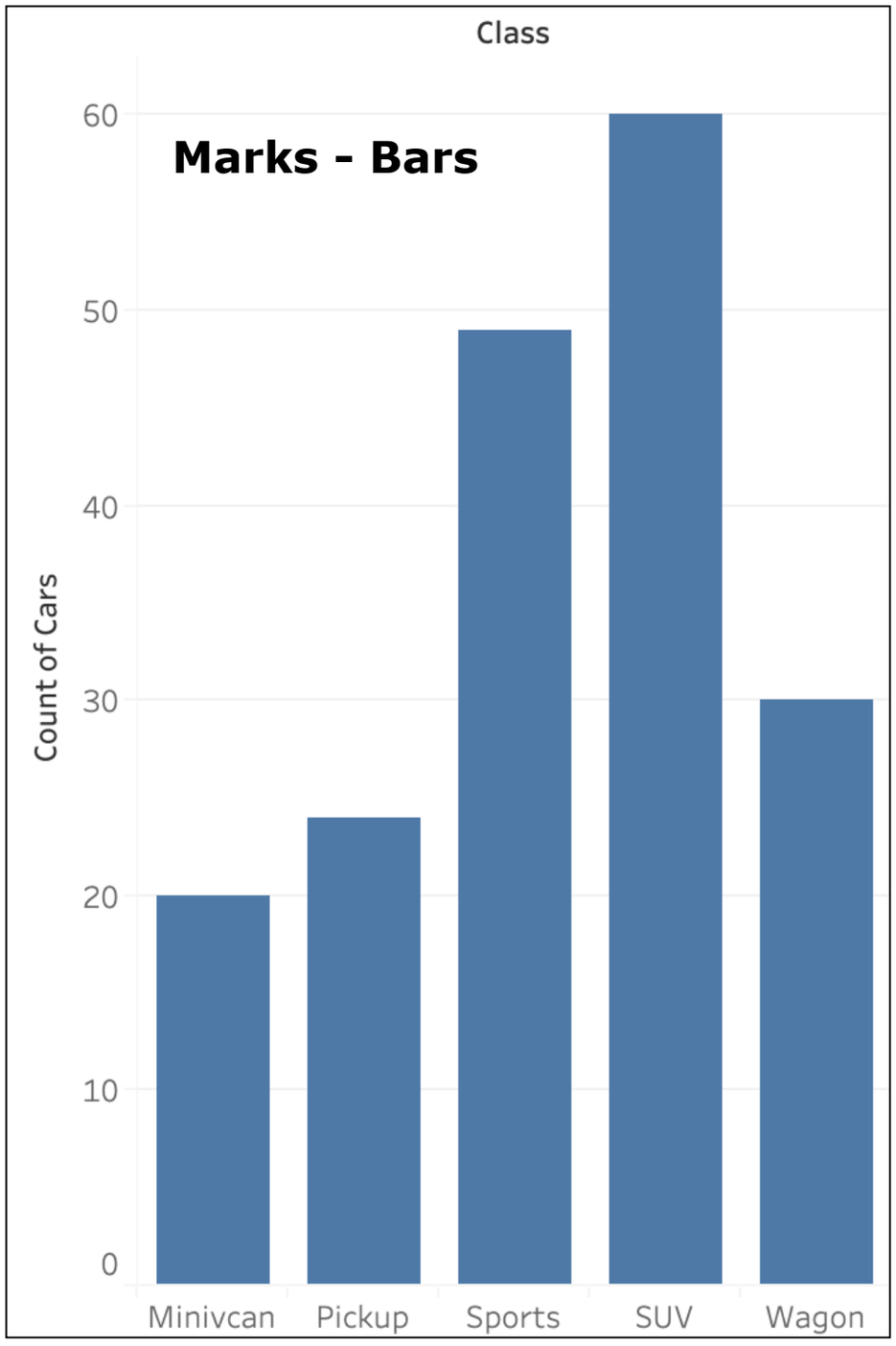


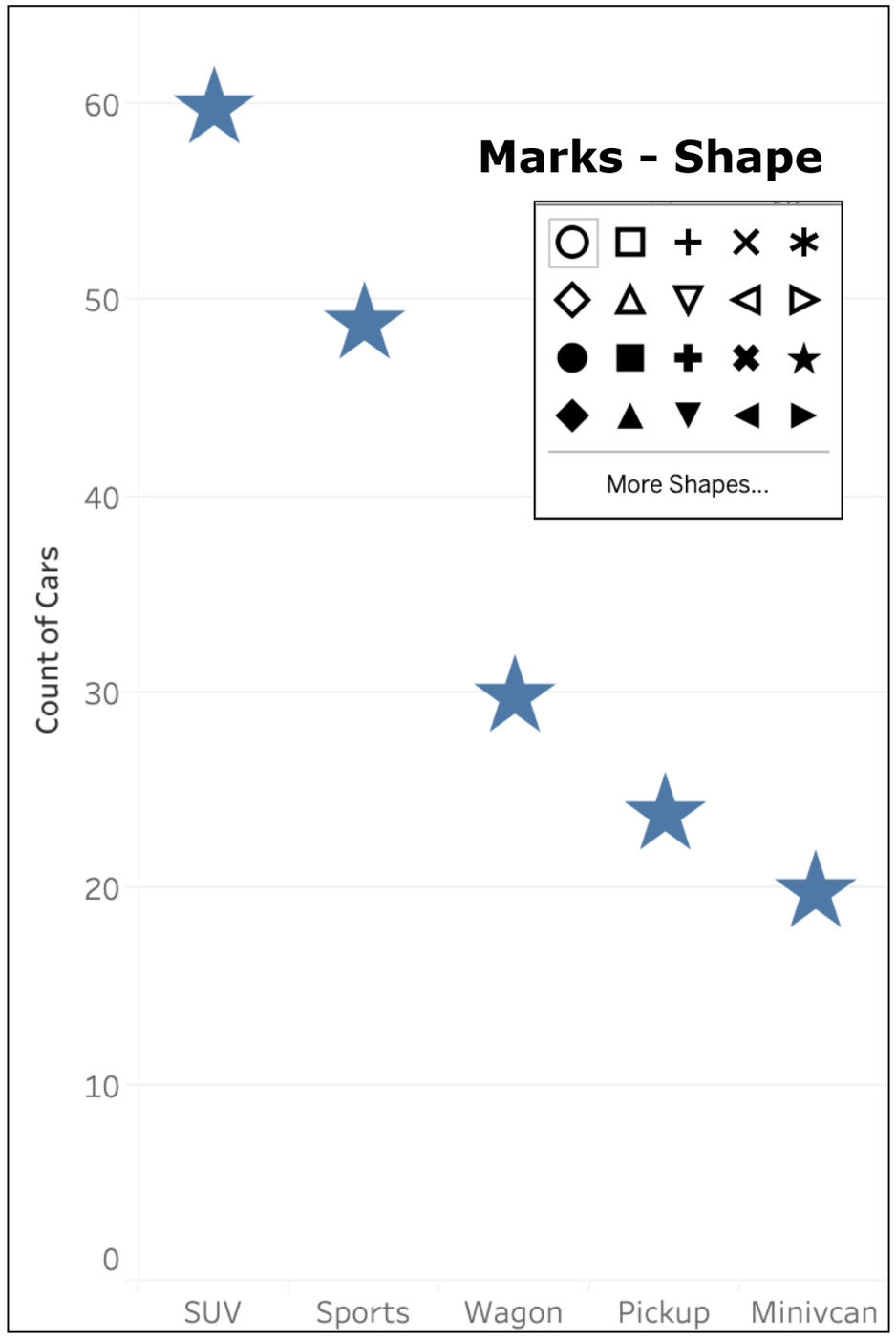



# Eight visual variables: Mark (or shape)

Y - Count of Cars

X - Class (excluding “Normal”)

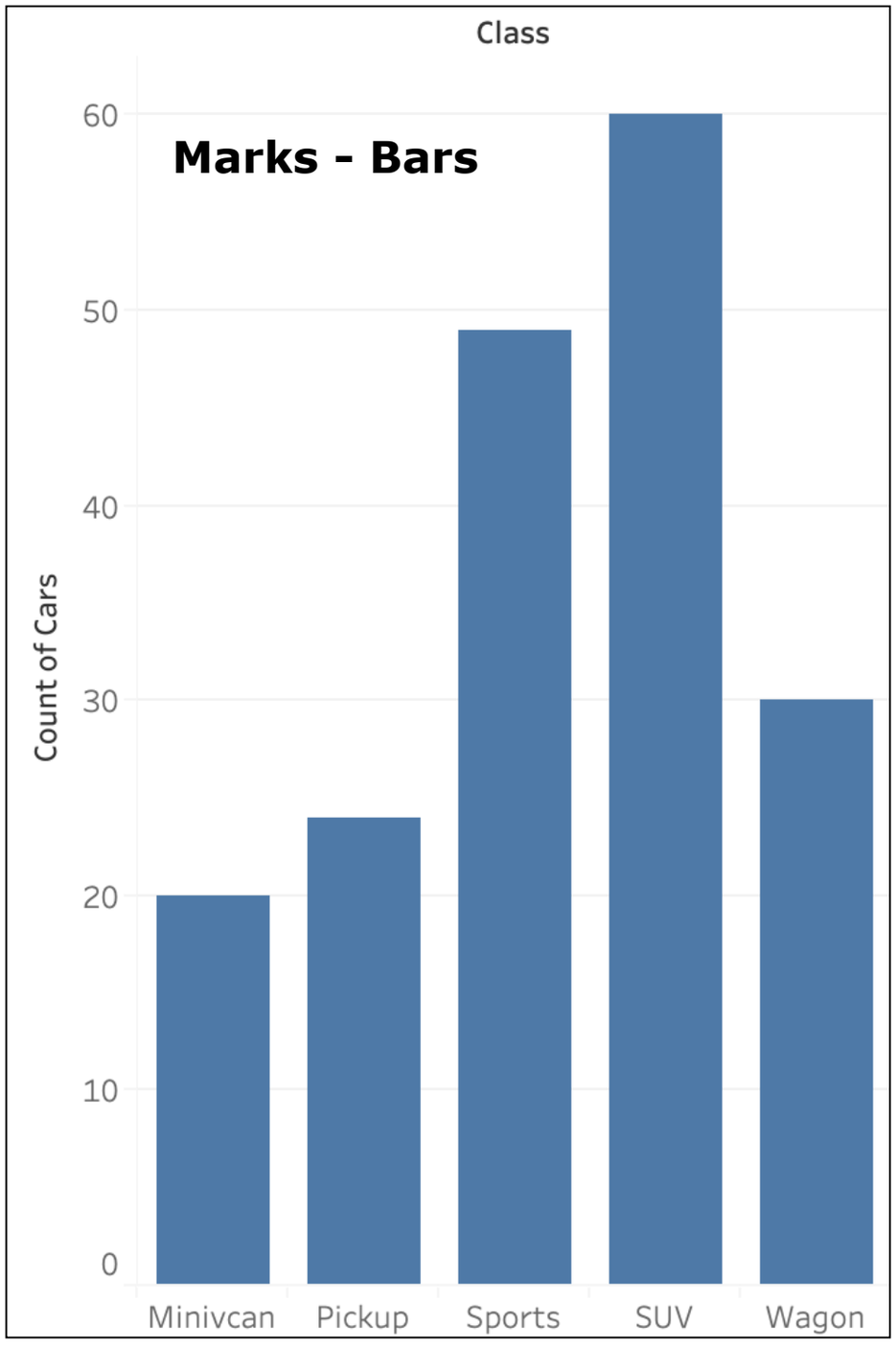


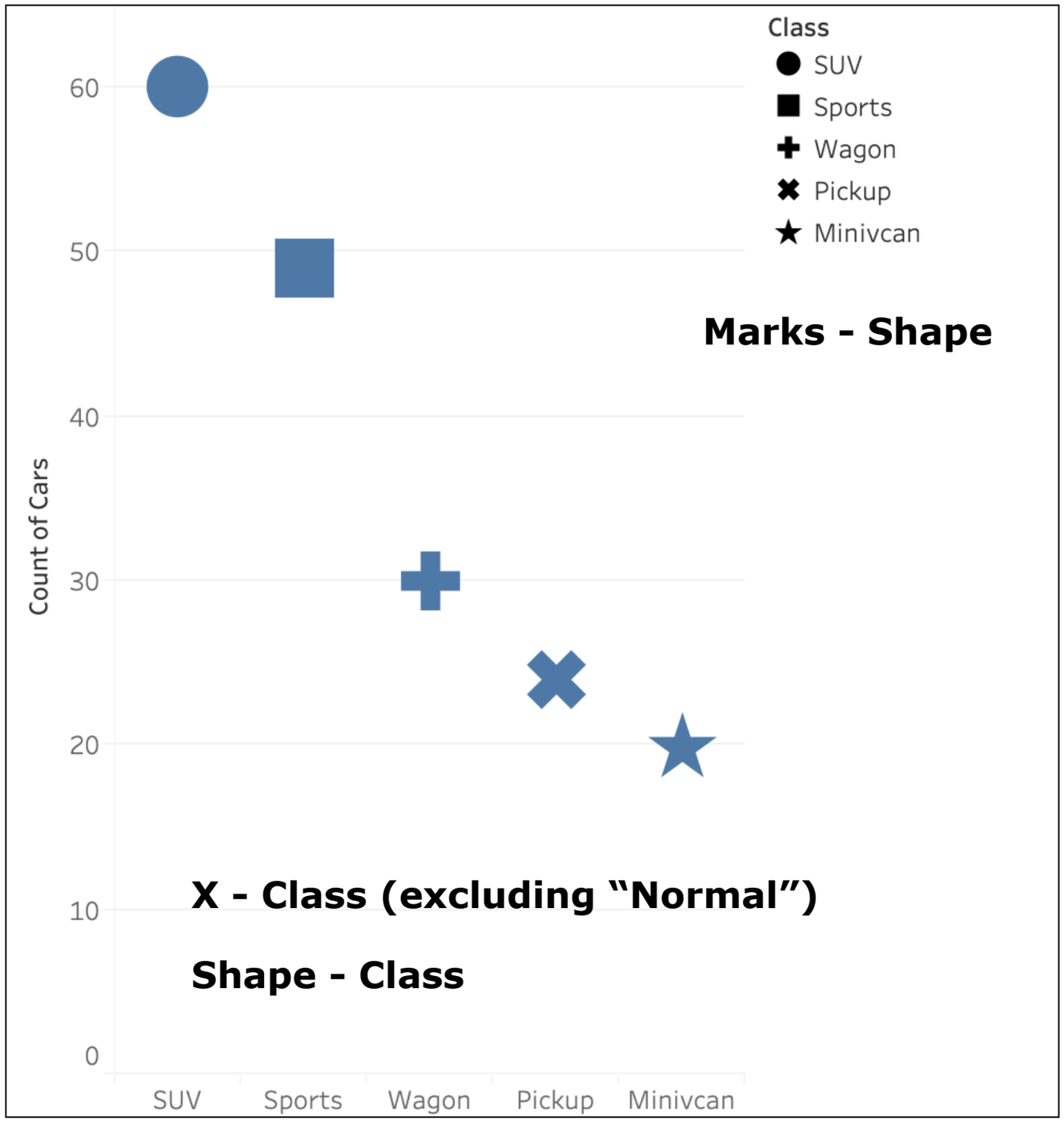



# Eight visual variables: Mark (or shape)

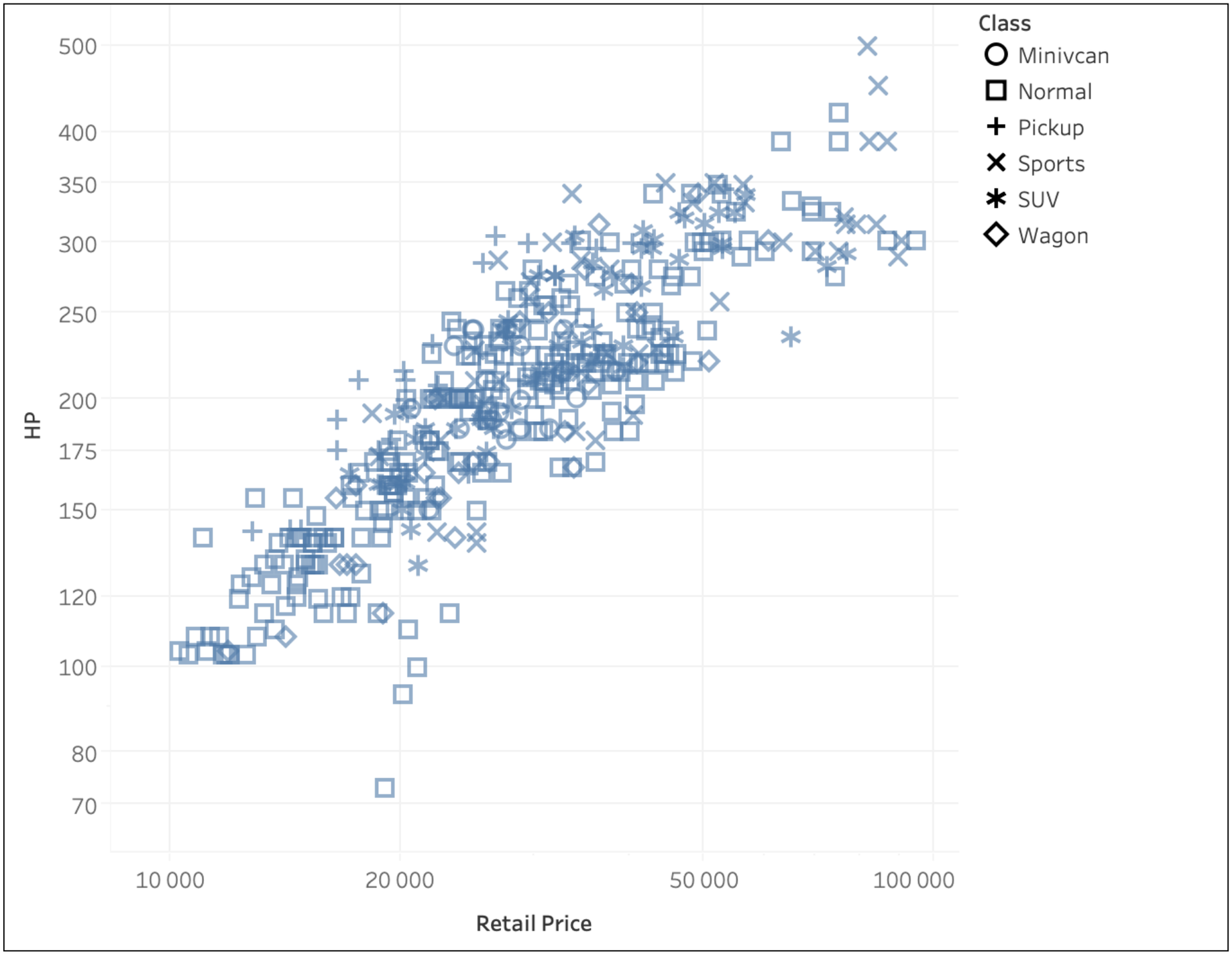
Y - Count of Cars

X - Class (excluding “Normal”)



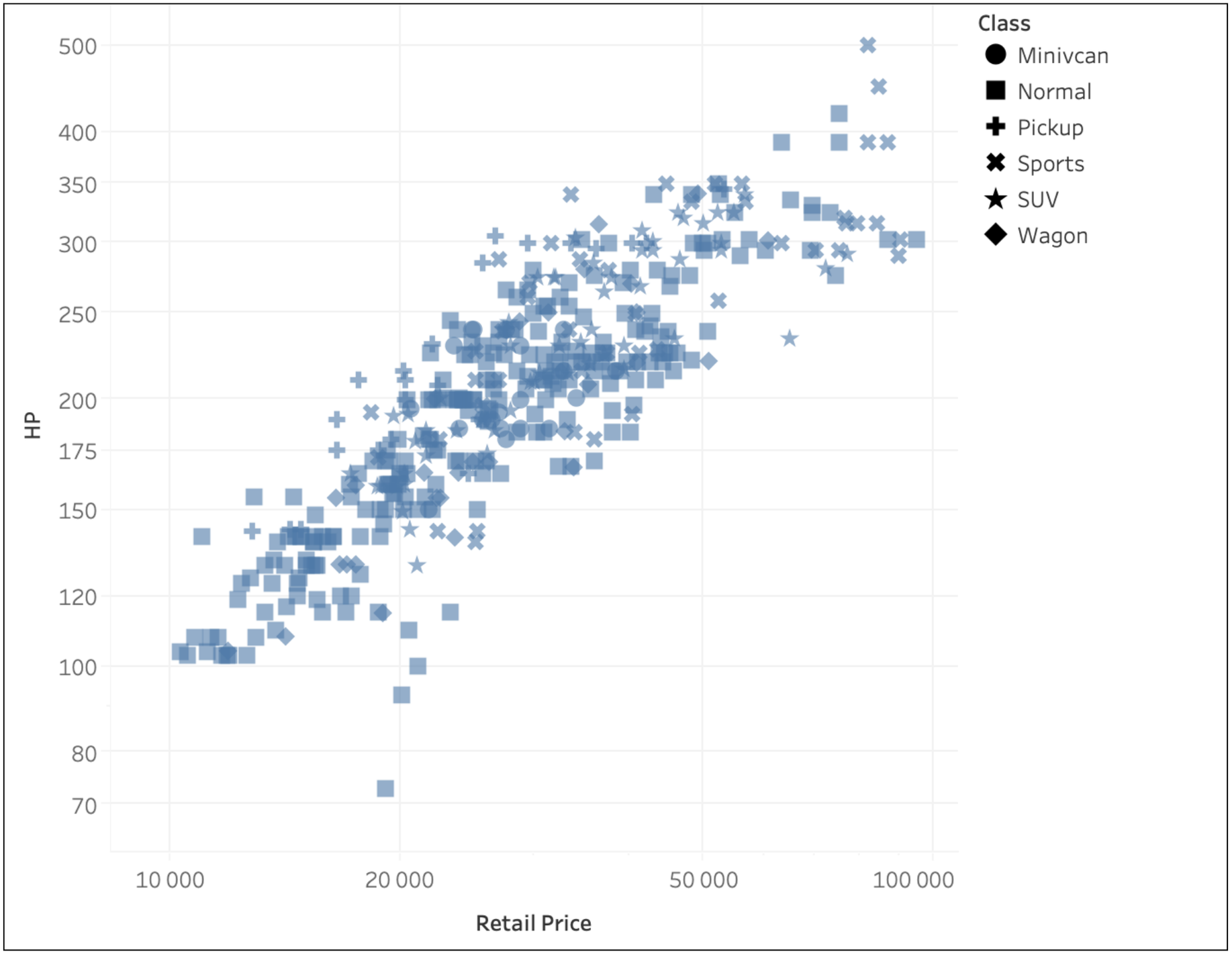
# Eight visual variables: Mark (or shape)




X - Retail Price;
Y - Horse Power

Marks - Class

# Eight visual variables: Mark (or shape)




X - Retail Price;
Y - Horse Power

Marks - Class

# Eight visual variables

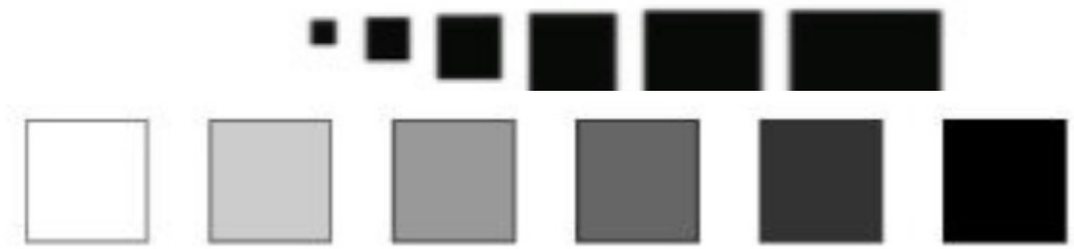
- **The position and marks, are required to define a visualization. Without these two variables there would not be much to see !**

- **The remaining visual variables affect the way individual representations are displayed;**

- **These are the graphical properties of marks other than their shape.**

# Eight visual variables

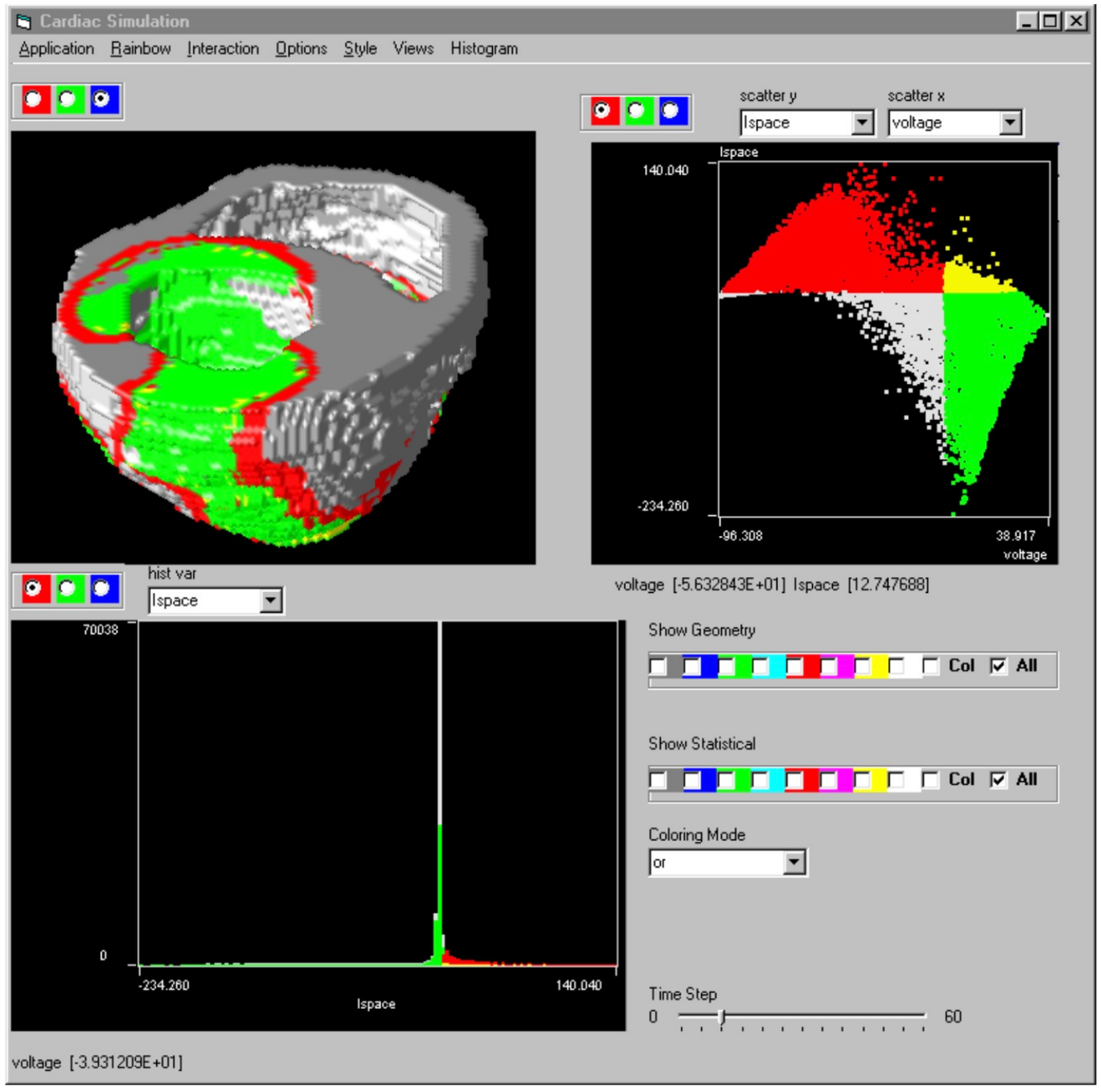
- **eight visual variables:**
  - **position,**
  - **shape,**
  - **size,**
  - **brightness,**
  - **color,**
  - **orientation,**
  - **texture,**
  - **motion**



# (Data) Visualization today

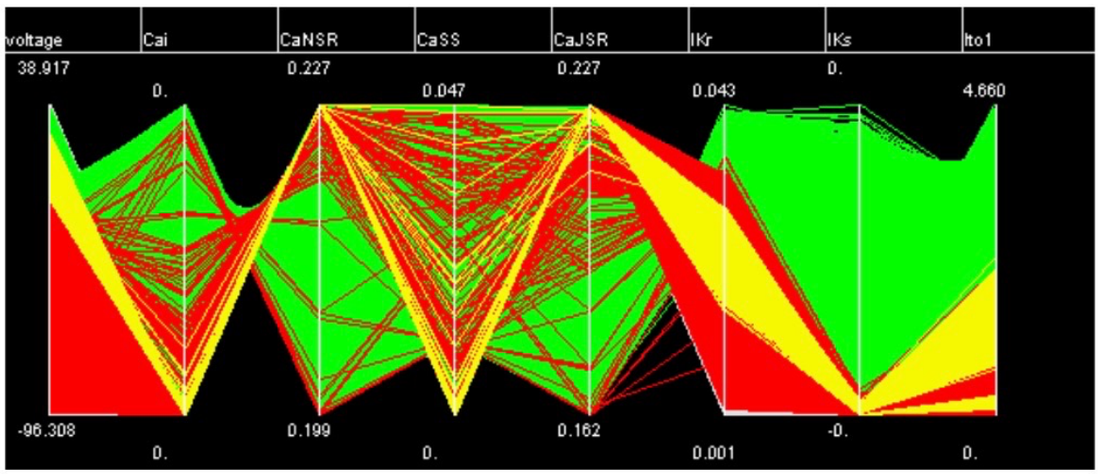
# Visualization today

- **Single visualization versus Multiple visualization**




Heart 3D Model
Additional parameters

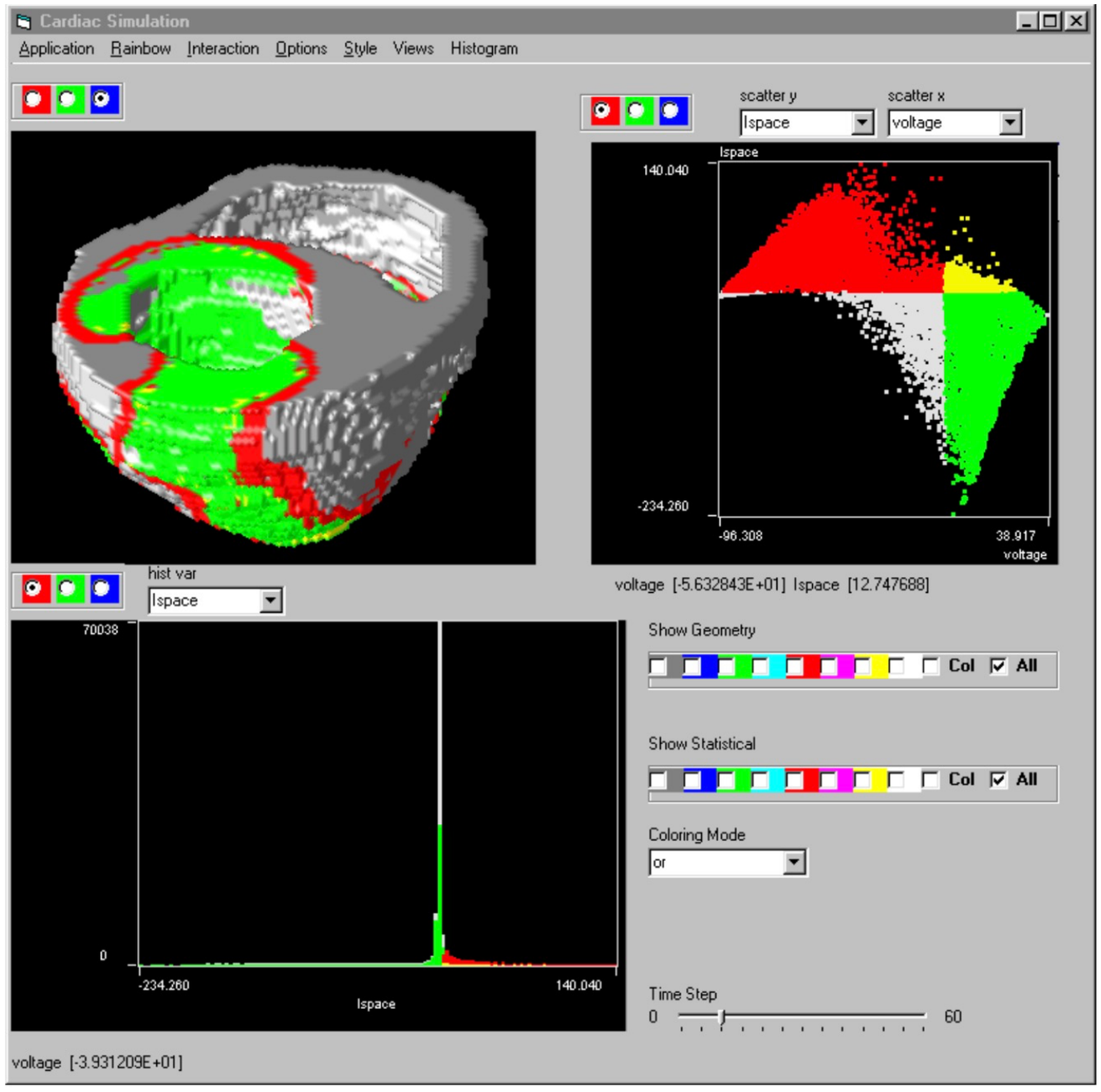
**Linked** parallel coordinates presentation




D. L. Gresh, B. E. Rogowitz, R. L. Winslow, D. F. Scollan, and C. K. Yung. “WEAVE: A System for Visually Linking 3D and Statistical Visualizations, Applied to Cardiac Simulation and Measurement Data.”

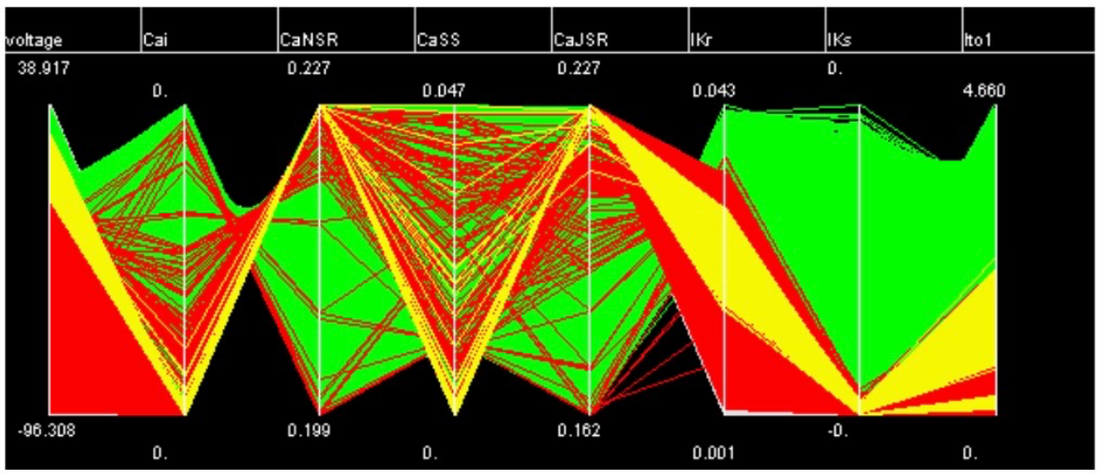
# Visualization today

- **Static versus Interactive**




In an **interactive** visualization the user can query the display and thus interact with the application display directly rather than menus

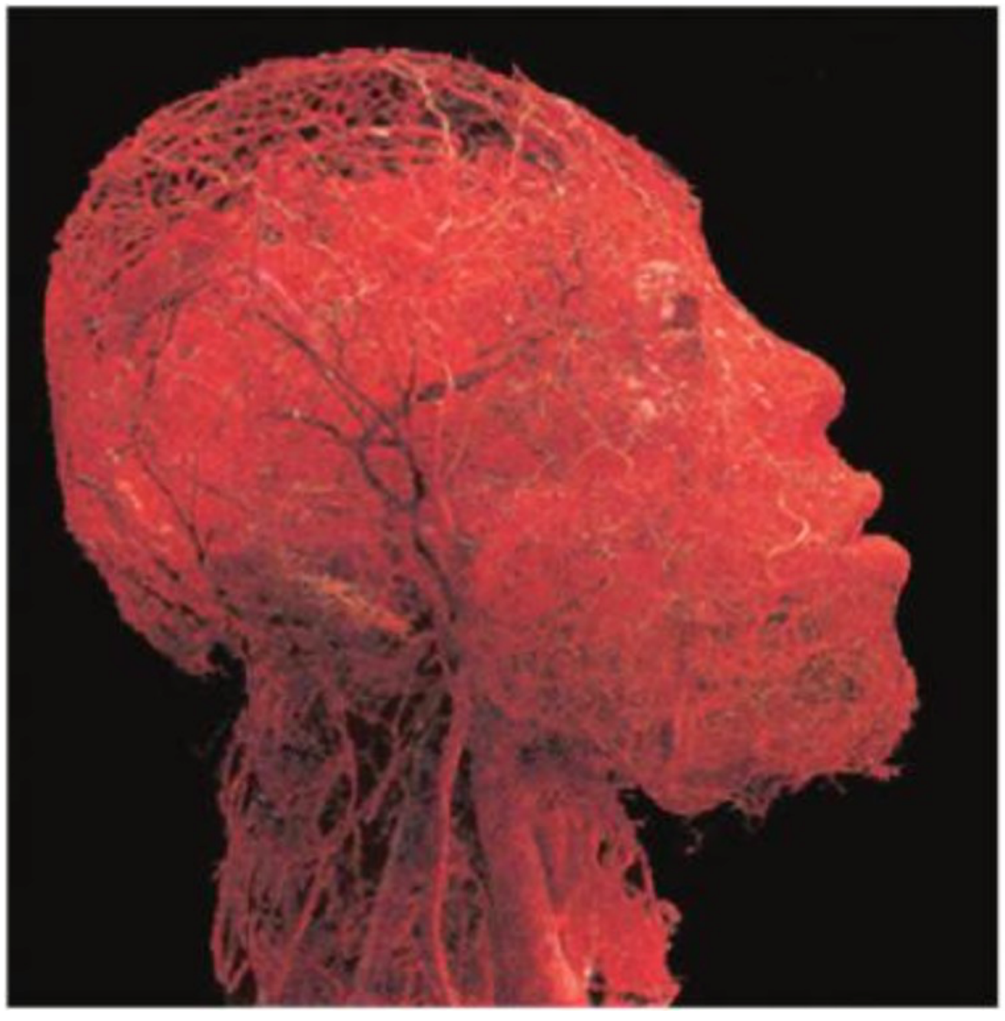
**Linked** parallel coordinates presentation



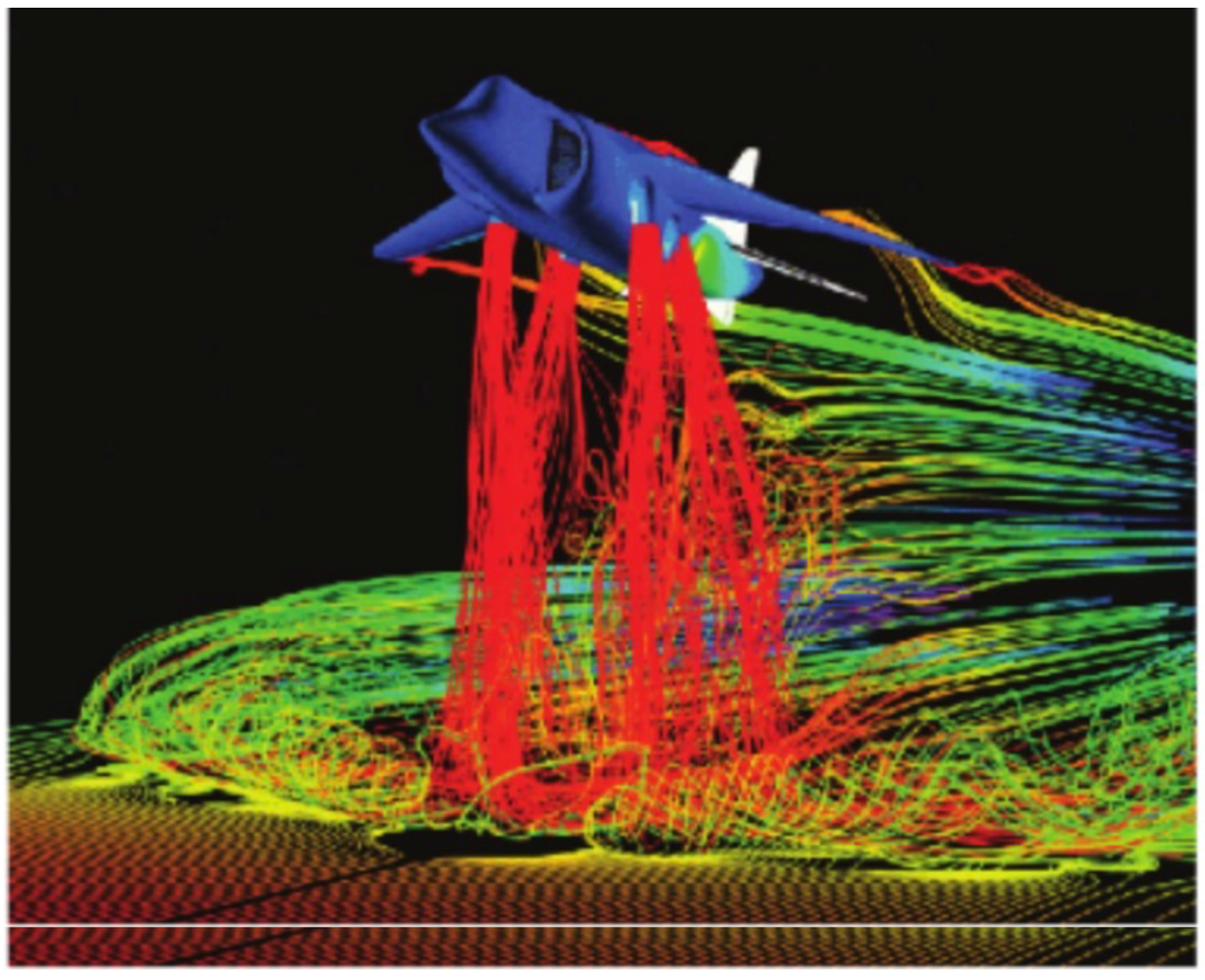

D. L. Gresh, B. E. Rogowitz, R. L. Winslow, D. F. Scollan, and C. K. Yung. “WEAVE: A System for Visually Linking 3D and Statistical Visualizations, Applied to Cardiac Simulation and Measurement Data.”

# Visualization today

- **Abstraction versus “real images”**



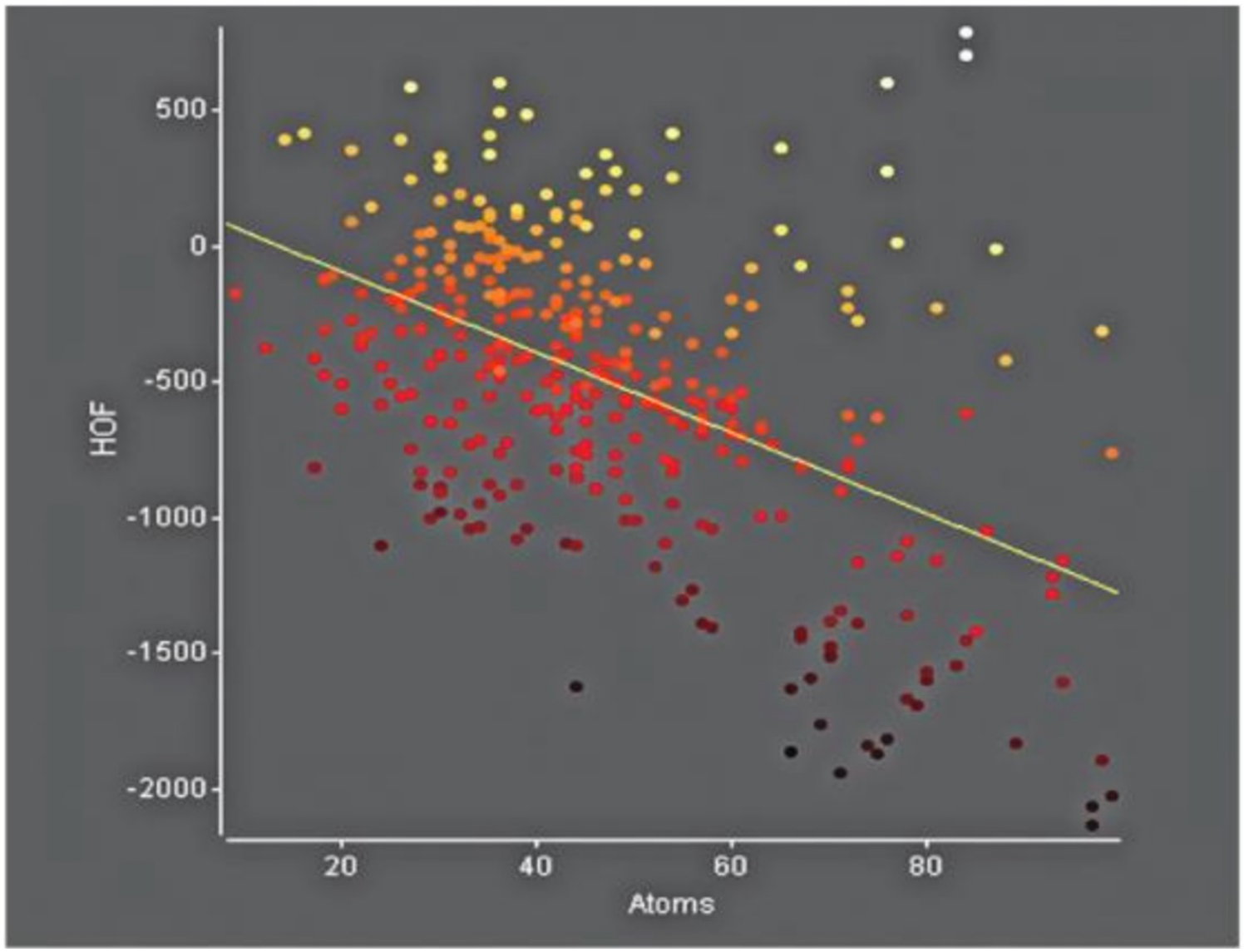
Blood vessel configuration of the head and Brain (http://www.bodyworlds.com/)



Simulation visualization of the air generated by a Harriet Jet (http://quest.nasa.gov/aero/background/tools/)

# Visualization today

- **Abstraction versus "real images"**




Mechanism of action for yeast (image generated by UMass Lowell UVP Software)

*x*-coordinate: number of atoms;
*y*-coordinate: heat information;

$y = mx + b$; $m$ = -12.5 and $b$ = 50

Color of each point: Gibs energy

Visualization provides **visual representation of objects** that may include data, algorithms, results of computations, process and many other components of the application

The ability to provide rich descriptions of data is one of the strengths of visualization

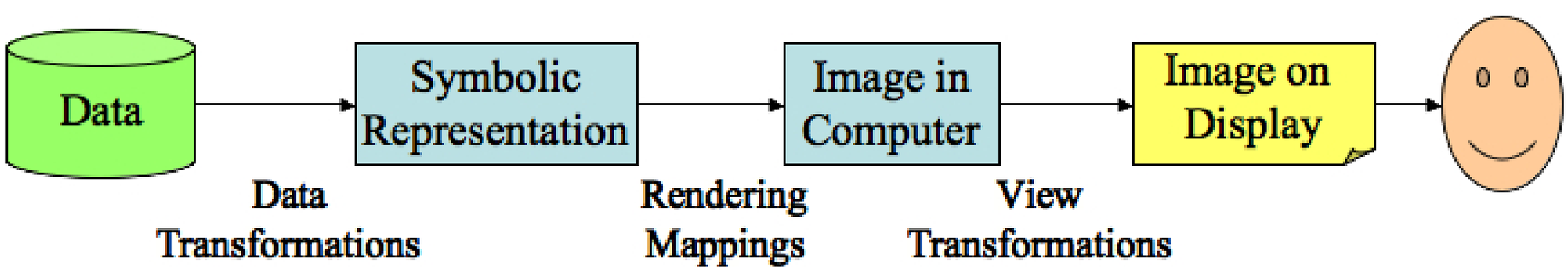
# (Data) Visualization and other fields

# Visualization and other fields

- **Visualization emerges as a sub-field of Computer Graphics, and is now a new field that encompasses aspects from *human-computer interaction*, *perceptual psychology*, *databases*, *statistics*, *data mining*, and *computer graphics*, and others.**
- ***Computer graphics* focus on graphical objects and the organization (and implementation) of graphical primitives.**
- **Visualization is the application of graphics to display data by mapping data to graphical primitives and rendering the display.**
- **In Computer Graphics the visual realism is often one major concern. In Visualization the focus is on finding an effective communication of information.**

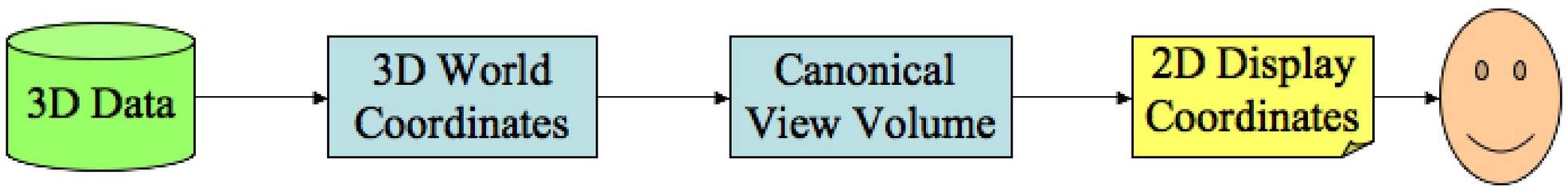
# Visualization Process

- **What is involved in the Visualization process?**
  - **Type of data available for display**
  - **Type of the information the Viewer hopes to extract from (exploration; confirm hypotheses) or convey with the display (present results)**



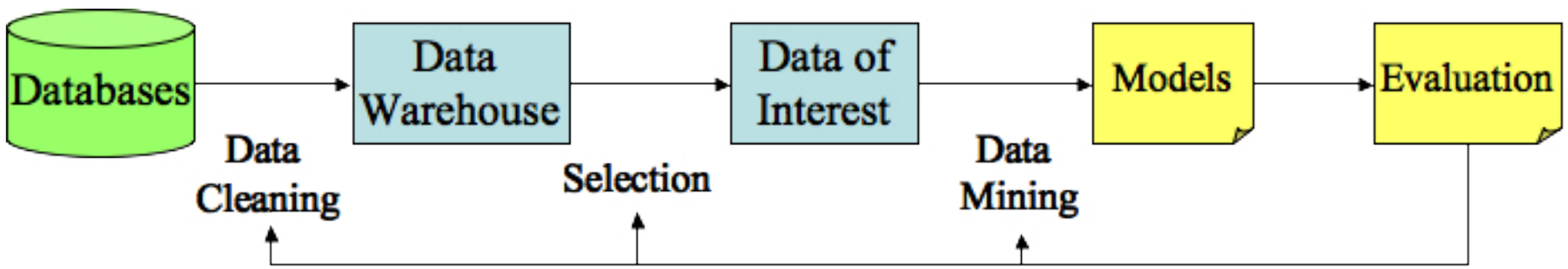
# Visualization Process: **computer graphics** pipeline

- **For computer graphics the stages are:**
  - **Modeling: 3D model**
  - **Viewing: virtual camera**
  - **Clipping: bounds of the desired image**
  - **Hidden surface removal & Projection: mapping to a 2D system**
  - **Rendering: color, illumination, etc.**



# Visualization Process: the **knowledge discovery** pipeline

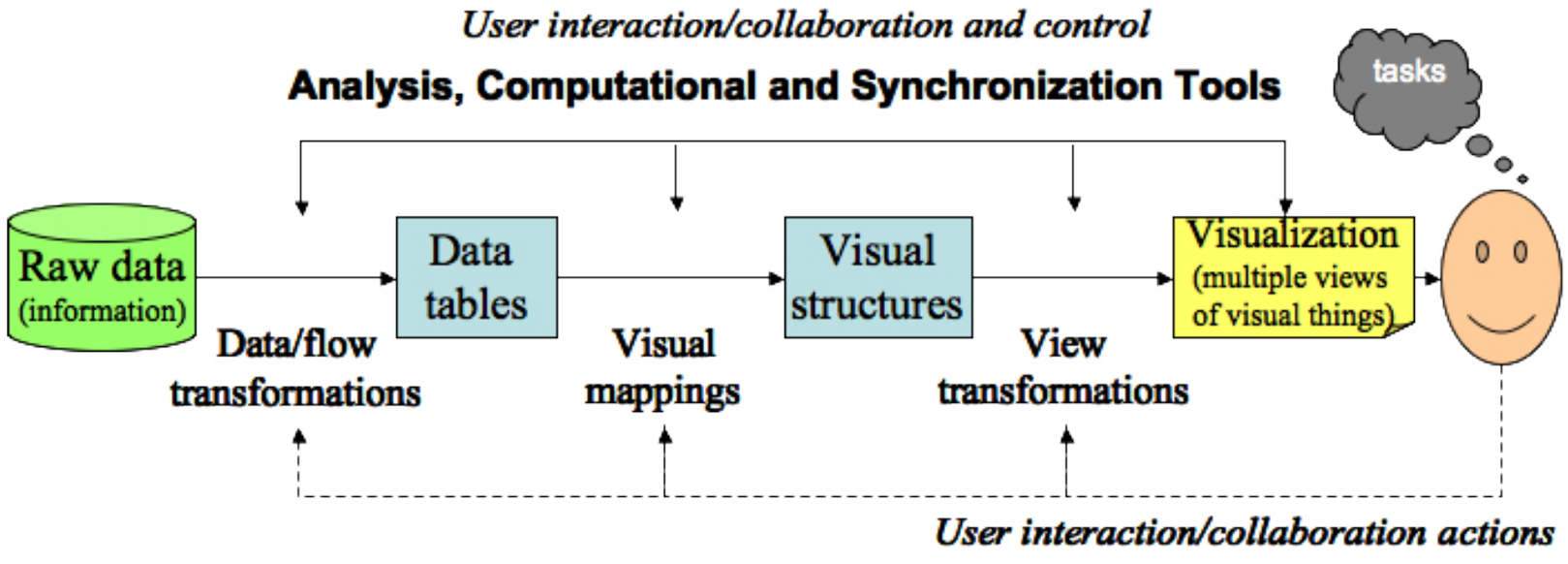
- **For knowledge discovery the stages are:**
  - **Data**
  - **Data integration, cleaning, warehousing and selection**
  - **Data mining**
  - **Pattern evaluation**
- **Rendering or visualization:**




(*) Interactive visualization can be used at every step of KD pipeline

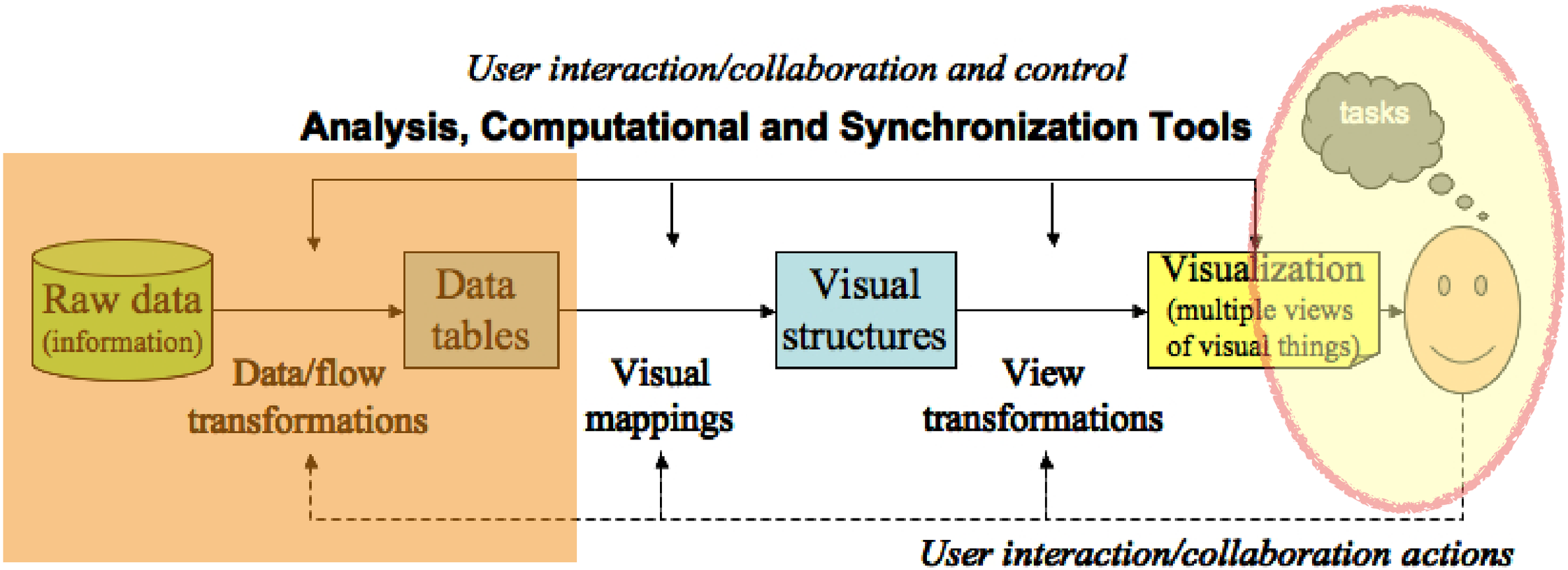
# Visualization Process: **Visualization pipeline**

- **For visualization the stages are:**
  - **Modeling: the data to be visualized**
  - **Data Selection: similar to clipping**
  - **Data to visual mappings: the heart of the visualization is mapping data values to graphical entities or their attributes; may involve scaling, shifting, filtering, interpolating, or subsampling.**
  - **Scene parameter setting: (ex: color mapping)**
  - **Rendering or generation of the visualization**



# Visualization Process: visualization pipeline

- **What is involved in the Visualization process?**
  - **Type of data available for display**
  - **Type of the information the Viewer hopes to extract from (exploration; confirm hypotheses) or convey with the display (present results)**



# Further Reading and Summary

**Q&A**

# Further Reading

- **Recommend Readings**
  - Interactive Data Visualization: Foundations, Techniques, and Applications, Matthew O. Ward et all, 2015, pages 1 - 38.
- **Supplemental readings:**
  - Cholera map's John Snow:
    - https://en.wikipedia.org/wiki/1854_Broad_Street_cholera_outbreak
  - Napoleon
    - https://en.wikipedia.org/wiki/Charles_Joseph_Minard
  - William Playfair:
    - https://en.wikipedia.org/wiki/William_Playfair
  - Florence Nightingale:
    - https://pt.wikipedia.org/wiki/Florence_Nightingale
  - Periodic table:
    - https://en.wikipedia.org/wiki/Periodic_table

**Check - vis25timeline**

# What you should know

- **What is Visualization.**
  - grocking the data => take decisions
- **Data Visualization can be extremely powerful**
  - Uncover new patterns; confirm hypothesis;
- **Why Visualization is important.**
  - Stats not enough; communication needs; exploratory needs
- **Key aspects of today Visualizations.**
  - Interactions; visual abstractions; multiple (linked) visualizations.
- **The general steps of a Visualization Process**
  - **Raw data -> data -> viz structures -> images -> perception + feedback**
- **The role of Perception.**
  - The role and the importance of the user.

# Further Reading and Summary

**Q&A**

# Course Organization and Overview

# Syllabus

## Introduction to Data Visualization

What Is Visualization?
Relationship between Visualization and Other Fields.
The Visualization Process.
Data Foundations.
Human Perception and Information Processing.
Semiology of Graphical Symbols.
The Visual Variables.

## Visualization Techniques

Visualization Techniques for Spatial Data
Visualization Techniques for Geospatial Data
Visualization Techniques for Time-Oriented Data
Visualization Techniques for Multivariate Data
Visualization Techniques for Trees, Graphs, and Networks
Text and Document Visualization

## Interaction Concepts and Techniques

Interaction Operators, Operands and Spaces (screen, object, data, attributes)
Visualization Structure Space (Components of the Data Visualization)
Animating Transformations
Interaction Control
Designing Effective Visualizations
Comparing and Evaluating Visualization Techniques

## Visualization Systems

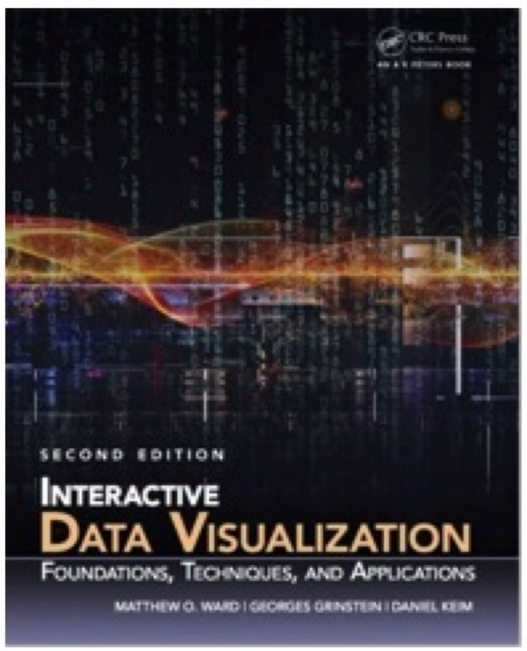
Systems Based on Data Type
Systems Based on Analysis Type
Text Analysis and Visualization
Modern Integrated Visualization Systems
Toolkits

## Research Directions in Visualization

# Bibliography

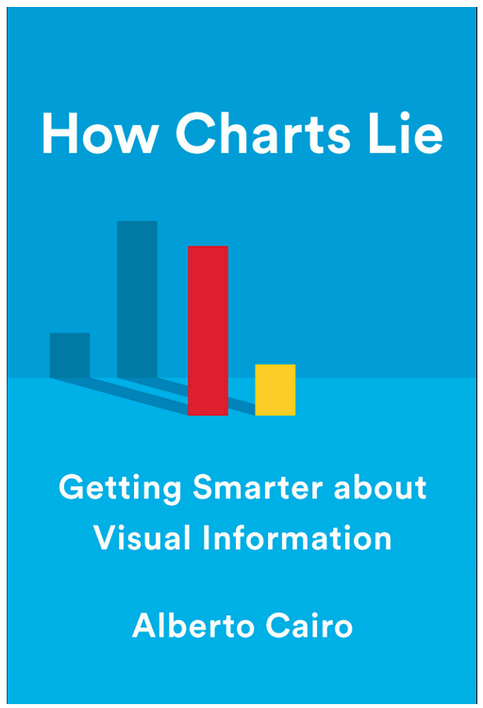


**Visualization Analysis & Design**
Tamara Munzner
2015,
ISBN: 9781466508910
ISBN (e-Book): 9781498707763

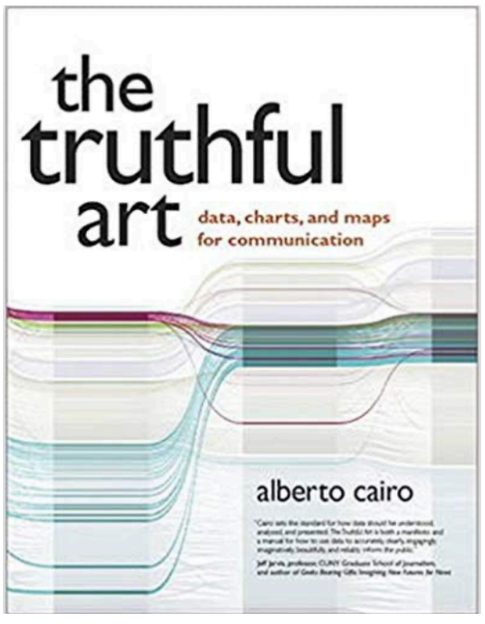



**Interactive Data Visualization: Foundations, Techniques, and Applications**
Matthew O. Ward, Georges Grinstein, Daniel Keim
2015, 2nd Edition
ISBN: 9781482257373
ISBN (e-Book): 9781482257397

# Bibliography


**How Charts Lie: Getting Smarter about Visual Information**
Alberto Cairo
2019




**The Truthful Art: Data, Charts, and Maps for Communication (Voices That Matter)**
Alberto Cairo
2016

# Bibliography

https://www.tableau.com/learn

# Weekly routine

- **Lectures - 1 x 2 h**

- **The lab sessions - 1 x 2 h**
  - **Demoing and Training**
  - **Project support**

- **The recommended readings**

- **The recommended actions**

- **Meetings for student support if required**

# Evaluation rules

- **Two mid-term written individual tests (25% each)**
- **One project (for team of two students), with several phases:**
  - **Specification**
  - **Paper**
  - **Code/implementation**
    - **(*) an oral discussion will be required to validate the project components**
- **Course approval requires the following minimal grades:**
  - **(mean (Test1; Test2) >= 10)**
  - **Project >= 10**
- **Final exam may replace mean (Test1; Test2) if project is approved.**
- **Tests and Exam presencial if possible; Without consultation.**

# Summaries

## [T01]: Course overview

06 Mar 2020, 04:10 PM Filed in: Lectures

**What we mean by "Interactive Data Visualization"?** What is Visualization? Why Visualization is important? Early Visualizations; Visualization today; Visualization and other fields. Visualization Process; The role of Perception.

**Course Organization and Overview**: Syllabus; Bibliography; Evaluation rules; important dates, etc..

**Recommended Readings: (i)** Interactive Data Visualization: Foundations, Techniques, and Applications, Matthew O. Ward et all, 2010, pages 1 - 33.

**Recommended Activities: (ii)** Visit the various sections of this site; **(iii)** instal Tableau software o your computer. Follow the link http://www.tableau.com/academic/students.

**To Know:**

- What is Visualization.
- The main "applications" of Visualization.
- Why Visualization is important.
- Key aspects of today Visualizations.
- Some important landmarks of early visualizations. For each one why is a landmark.
- The relation between Visualization and computer graphics. The differences between them.
- The relation of Visualization with other fields.
- The general steps of a Visualization Process
- The role of Perception.
- The role and the importance of the user.

# Important Dates

| Week | Subjects | Event |
|---|---|---|
| 1 | Overview | |
| 2 | Introduction to Data Visualization | **DATES WILL BE PUBLISHED** |
| 3 | | |
| 4 | | |
| 5 | | |
| 6 | Visualization Techniques | |
| 7 | | |
| 8 | | |
| 9 | | |
| 10 | | |
| 11 | Advanced Topics: Evaluation; Research directions | |
| 12 | | |
| 13 | | |
| 14 | Students Support | |
| 15 | Oral Sessions | |

**Working in progress**

# Further Reading and Summary

# Further Reading

- **Recommend Readings**
  - Interactive Data Visualization: Foundations, Techniques, and Applications, Matthew O. Ward et all, 2015, pages 1 - 38.
  - Visualization Analysis & Design, from Tamara Munzner, Chapter 1
- **Supplemental readings:**
  - Cholera map's John Snow:
    - https://en.wikipedia.org/wiki/1854_Broad_Street_cholera_outbreak
  - Napoleon
    - https://en.wikipedia.org/wiki/Charles_Joseph_Minard
  - William Playfair: (https://en.wikipedia.org/wiki/William_Playfair)
  - Florence Nightingale: (https://pt.wikipedia.org/wiki/Florence_Nightingale)
  - Periodic table: (https://en.wikipedia.org/wiki/Periodic_table)

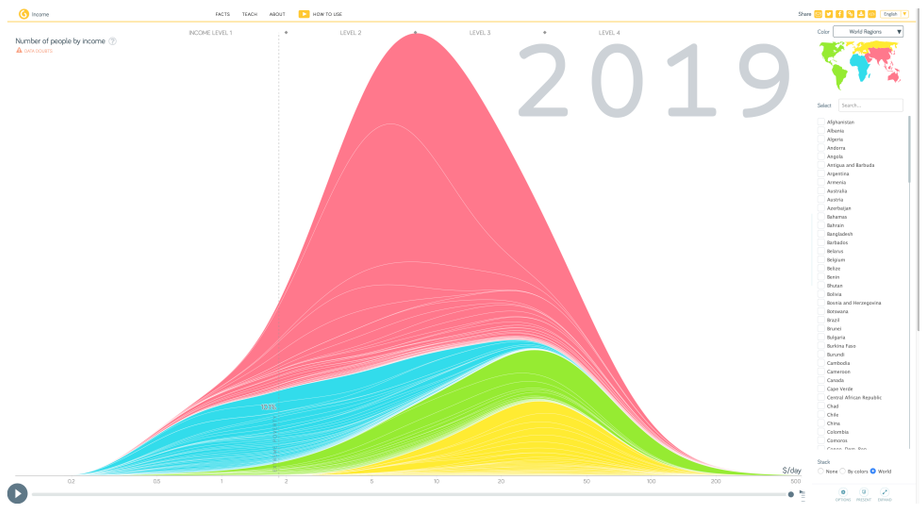
# What you should know

- **What is Visualization.**
  - Understanding the data => take decisions
- **Data Visualization can be extremely powerful**
  - Uncover new patterns; confirm hypothesis;
- **Why Visualization is important.**
  - Stats not enough; communication needs; exploratory needs
- **Key aspects of today Visualizations.**
  - Interactions; visual abstractions; multiple (linked) visualizations.
- **The general steps of a Visualization Process**
  - **Raw data -> data -> viz structures -> images -> perception + feedback**
- **The role of Perception.**
  - The role and the importance of the user.

# Recommended Activities

- **See the video of Hans Rosling, “Debunking third-world myths”**
  - http://www.gapminder.org/videos/ted-talks/hans-rosling-ted-2006-debunking-myths-about-the-third-world/
  - https://www.youtube.com/watch?v=RUwS1uAdUcI


- **Play with Gapminder**
  - https://www.gapminder.org/tools/
    - Bubble Charts
    - Income



# Recommended Actions

- **Check the Summaries and follow its recommendations**
- **Install Tableau software**
  - https://www.tableau.com/products/desktop/download
  - **Get the academic license**
    - http://www.tableau.com/academic/students