04

Visualization Foundations

Notice

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

- **This material can be freely used for personal or academic purposes without any previous authorization from the author, provided that this notice is kept with.**

- **For commercial purposes the use of any part of this material requires the previous authorisation from the author.**

Bibliography

- **Many examples are extracted and adapted from**
  - **Interactive Data Visualization: Foundations, Techniques, and Applications,**

    Matthew O. Ward, Georges Grinstein, Daniel Keim, 2015
  - **Visualization Analysis & Design,**

    Tamara Munzner, 2015

Our Agenda

Week		Subjects	Event
1	2-6 Mars	Overview	
2	9-13 Mars	Introduction to Data Visualization	**No class**
3	16-20 Mars		**Team registration: Mars 20th**
4	23-27 Mars		
5	30-3 April		
6	6-10 April		**No class - Holliday**
7	13-17 April		
8	20-24 April	Visualization Techniques	**Subject Registration: 25 April**
9	27-1 May		
10	4-8 May		
11	11-15 May		**Paper: May 15th**
12	18-22 May	Advanced Topics	
13	25-29 May		
14	1-5 Jun	Students Support	**TP Implementation: June 5th**
15	8-12 Jun	Oral Sessions	**Oral Sessions**

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Our Agenda

Week		Subjects	Event
1	2-6 Mars	Overview	
2	9-13 Mars	Introduction to Data Visualization	No class
3	16-20 Mars		Team registration: Mars 20th
4	23-27 Mars		
5	30-3 April		
6	6-10 April		No class - Holliday
7	13-17 April		
8	20-24 April	Visualization Techniques	Subject Registration: 25 April
9	27-1 May		
10	4-8 May		
11	11-15 May		Paper: May 15th
12	18-22 May	Advanced Topics	
13	25-29 May		
14	1-5 Jun	Students Support	TP Implementation: June 5th
15	8-12 Jun	Oral Sessions	Oral Sessions

?

FACULDADE DE CIÊNCIAS E TECNOLOGIA UNIVERSIDADE NOVA DE LISBOA

Our Agenda

Week		Subjects	Event
1	2-6 Mars	Overview	
2	9-13 Mars	Introduction to Data Visualization	**No class**
3	16-20 Mars		**Team registration: Mars 20th**
4	23-27 Mars		
5	30-3 April		
6	6-10 April		**No class - Holliday**
7	13-17 April		
8	20-24 April	Visualization Techniques	**Subject Registration: 25 April**
9	27-1 May		
10	4-8 May		
11	11-15 May		**Paper: May 15th**
12	18-22 May	Advanced Topics	
13	25-29 May		
14	1-5 Jun	Students Support	**TP Implementation: June 5th**
15	8-12 Jun	Oral Sessions	**Oral Sessions**

?

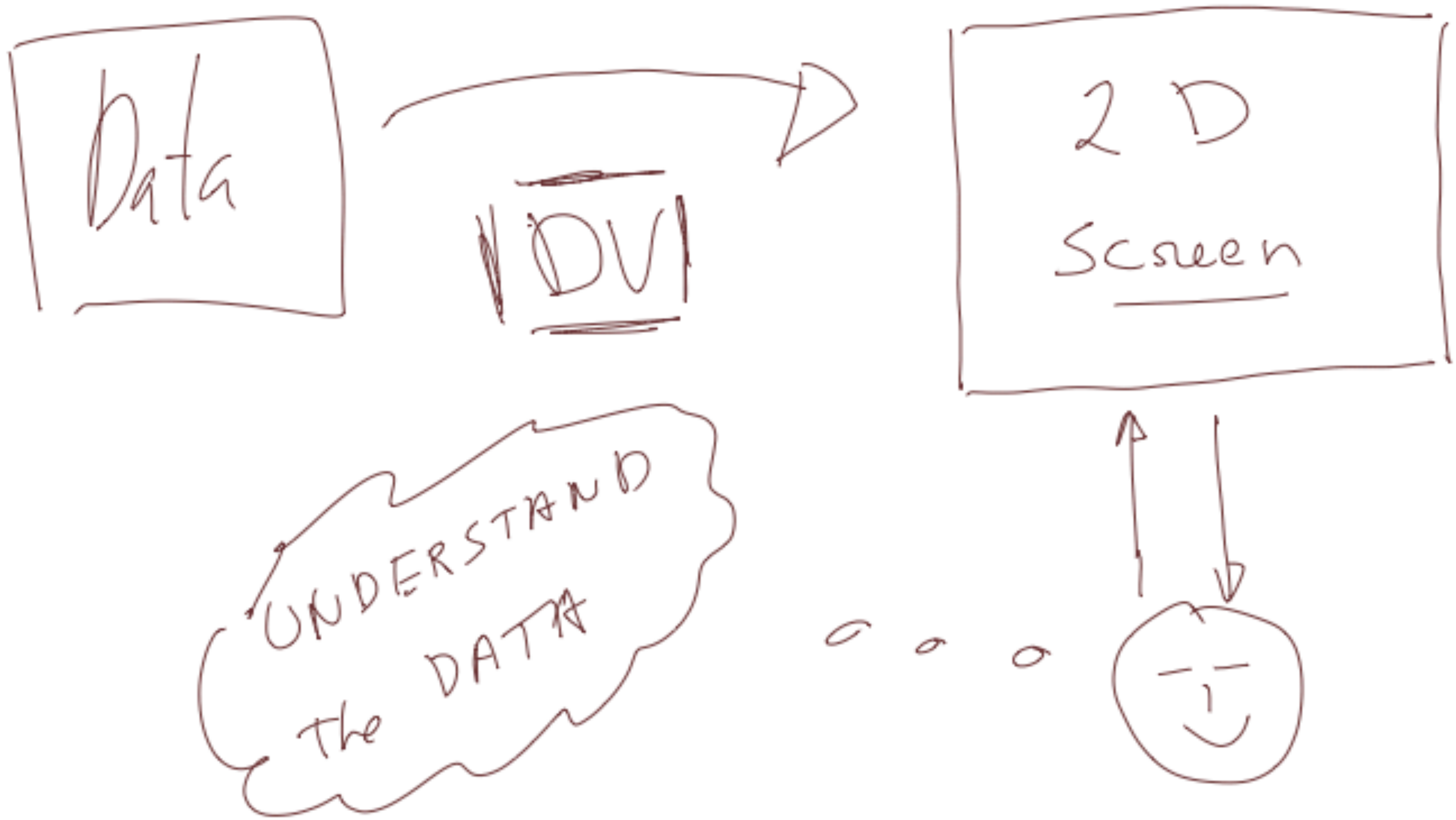
Our Agenda

Week		Subjects	Event
1	2-6 Mars	Overview	
2	9-13 Mars	Introduction to Data Visualization	No class
3	16-20 Mars		Team registration: Mars 20th
4	23-27 Mars		
5	30-3 April		
6	6-10 April		No class - Holliday
7	13-17 April		
8	20-24 April	Visualization Techniques	Subject Registration: 25 April
9	27-1 May		
10	4-8 May		
11	11-15 May		Paper: May 15th
12	18-22 May	Advanced Topics	
13	25-29 May		
14	1-5 Jun	Students Support	TP Implementation: June 5th
15	8-12 Jun	Oral Sessions	Oral Sessions

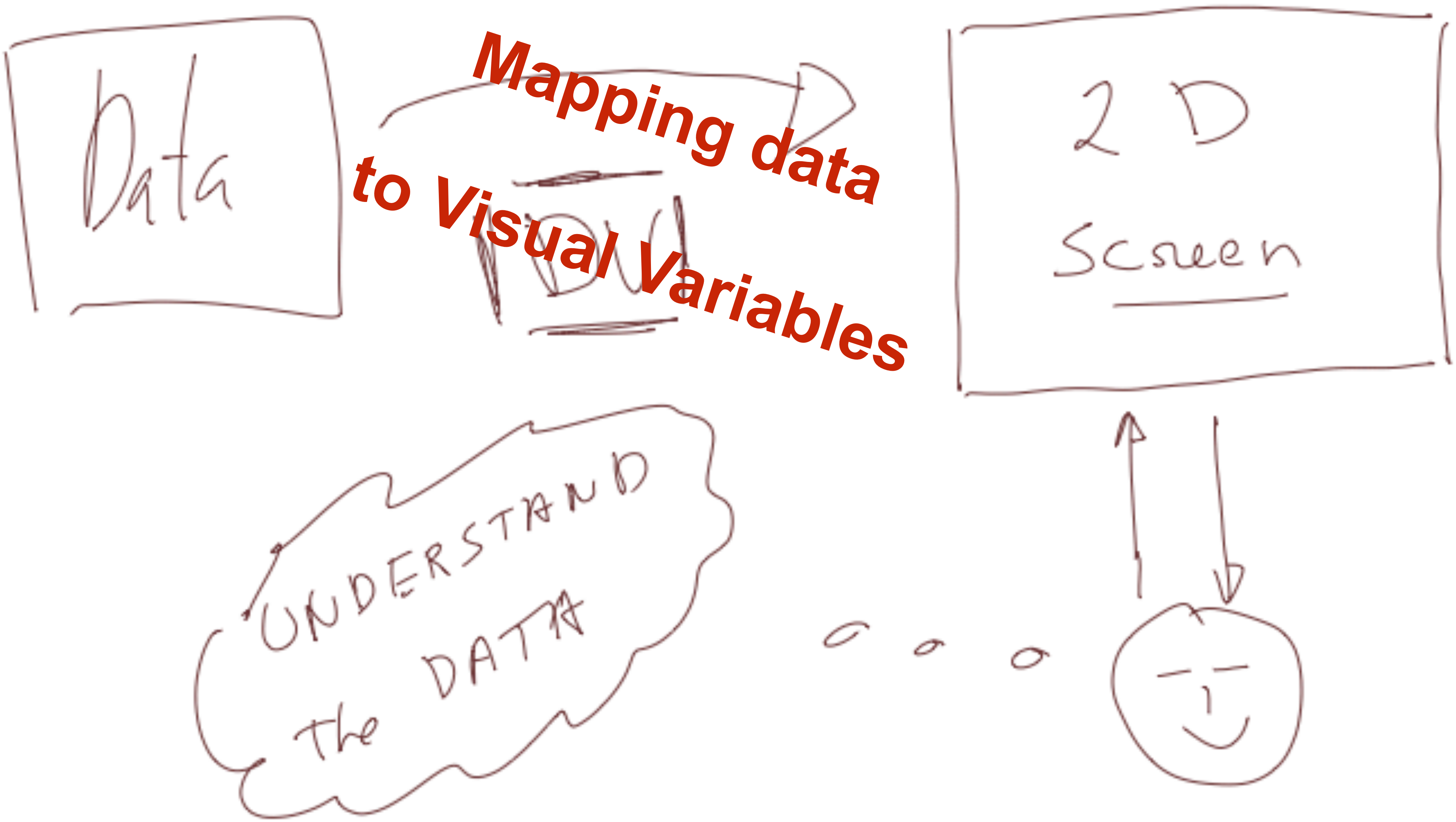
?

Never Forget

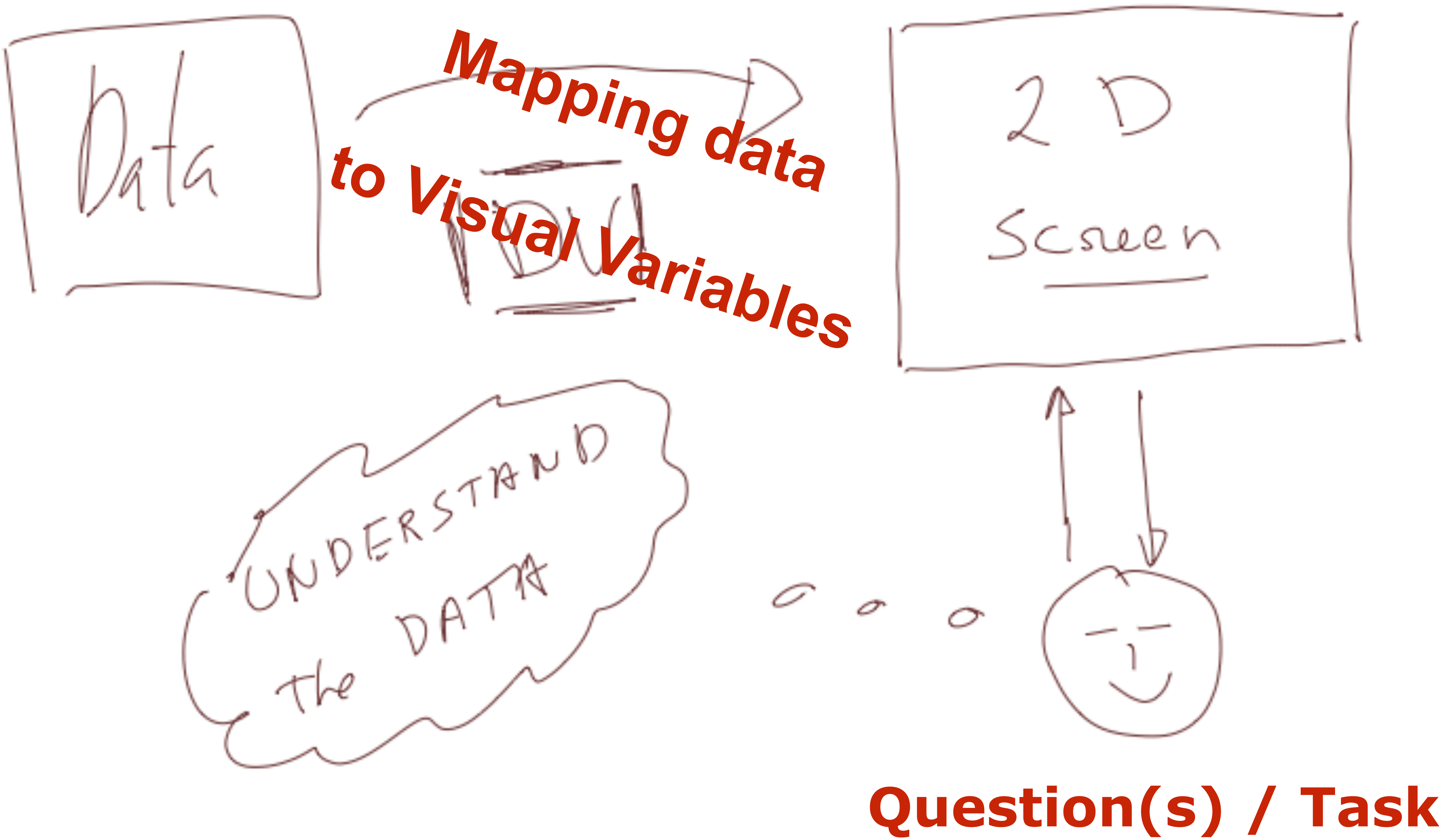
What is the core idea of Interactive Data Visualization?



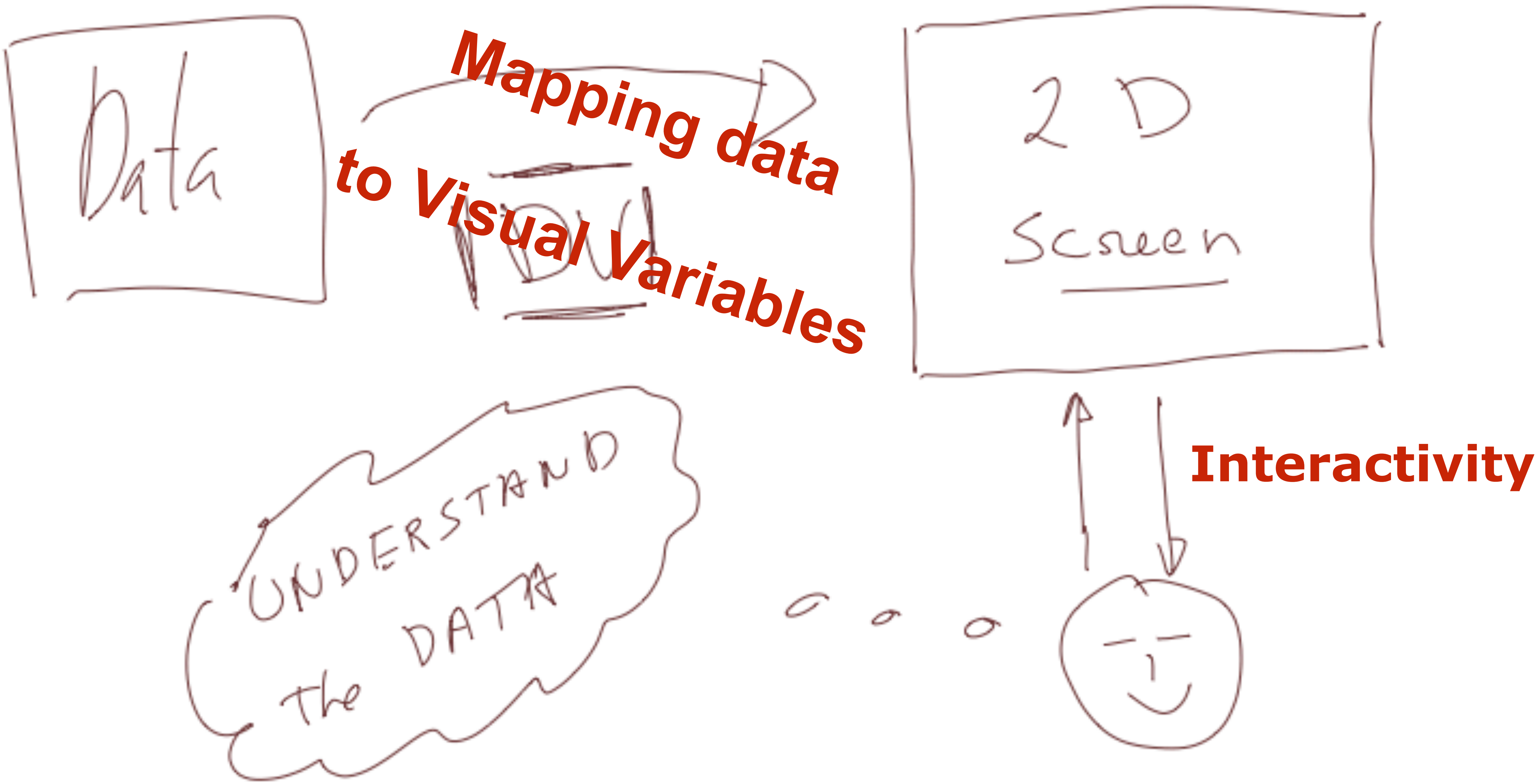
What is the core idea of Interactive Data Visualization?



What is the core idea of Interactive Data Visualization?



What is the core idea of Interactive Data Visualization?



Question(s) / Task

Table of Contents

- **The Visualization Process in Detail**
- **Semiology of Graphical Symbols**
- **The Eight Visual Variables**
- **Historical Perspective**

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

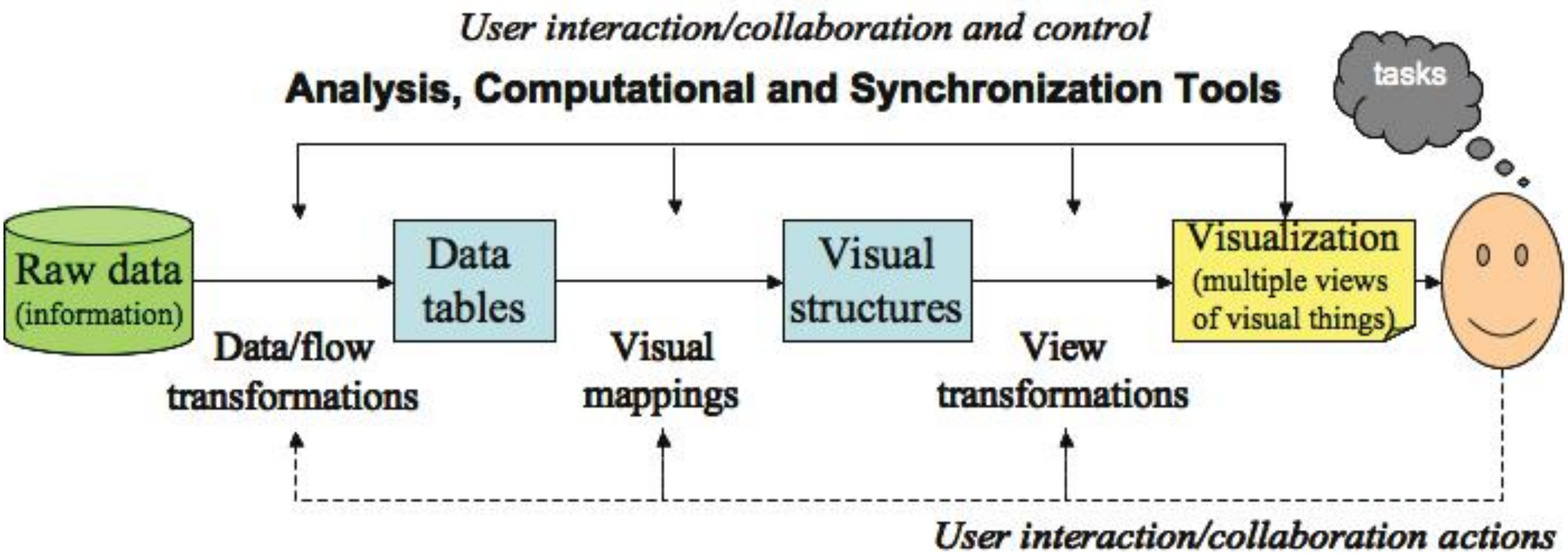
Interaction Operators, Operands and Spaces (screen, object, data, attributes)
Visualization Structure Space (Components of the Data Visualization)
Animating Transformations
Interaction Control
Designing Effective Visualizations
Comparing and Evaluating Visualization Techniques

Visualization Systems

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

The Visualization Process in Detail

The Visualization Process in Detail



FACULDADE DE
CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

The Visualization Process in Detail

- **Data preprocessing and transformation**
  - **Process the raw data into something usable by the visualization system.**
    - The first part is to make sure that the data are mapped to fundamental data types
    - The second step entails dealing with specific application data issues.
- **Mapping for visualizations**
  - **Decide on a specific visual representation.**
    - This requires representation mappings: geometry, color, and sound, for example.
- **Rendering transformations.**
  - **The final stage involves mapping from geometry data to the image**
    - This stage of the pipeline is very dependent on the underlying graphics library.

The Visualization Process in Detail

- **Data preprocessing and transformation**
  - **Process the raw data into something usable by the visualization system.**
    - The first part is to make sure that the data are mapped to fundamental data types
    - The second step entails dealing with specific application data issues.
- **Mapping for visualizations**
  - **Decide on a specific visual representation.**
    - This requires representation mappings: geometry, color, and sound, for example.
- **Rendering transformations.**
  - **The final stage involves mapping from geometry data to the image**
    - This stage of the pipeline is very dependent on the underlying graphics library.

Expressiveness and Effectiveness

Expressiveness and Effectiveness

- **Expressiveness**
  - **An expressive visualization presents all the information, and only the information**

Expressiveness and Effectiveness

- **Expressiveness**
  - **An expressive visualization presents all the information, and only the information**
  - **M_{exp} = The information that we actually display to the user / information we want to present to the user**
  - **$0 \leq M_{exp} \leq 1$.**

Expressiveness and Effectiveness

- **Expressiveness**
  - **An expressive visualization presents all the information, and only the information**
  - **M_{exp} = The information that we actually display to the user / information we want to present to the user**
  - **$0 \leq M_{exp} \leq 1$.**
  - **If $M_{exp} = 1$, we have ideal expressiveness**

Expressiveness and Effectiveness

- **Expressiveness**
  - **An expressive visualization presents all the information, and only the information**
  - **M_{exp} = The information that we actually display to the user / information we want to present to the user**
  - **$0 \le M_{exp} \le 1$.**
  - **If $M_{exp} = 1$, we have ideal expressiveness**
  - **If the information displayed is less than that desired, then $M_{exp} < 1$.**

Expressiveness and Effectiveness

- **Expressiveness**
  - **An expressive visualization presents all the information, and only the information**
  - **M_{exp} = The information that we actually display to the user / information we want to present to the user**
  - **$0 \leq M_{exp} \leq 1$.**
  - **If $M_{exp} = 1$, we have ideal expressiveness**
  - **If the information displayed is less than that desired, then $M_{exp} < 1$.**
  - **If $M_{exp} > 1$, we are presenting too much information.**
    - Expressing additional information is potentially dangerous, because it may not be correct and may interfere with the interpretation of the essential information.

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Expressiveness and **Effectiveness**

- **Effectiveness**
  - **A visualization is effective when it can be interpreted accurately and quickly and when it can be rendered in a cost-effective manner.**

Expressiveness and **Effectiveness**

- **Effectiveness**
  - **A visualization is effective when it can be interpreted accurately and quickly and when it can be rendered in a cost-effective manner.**
  - **Effectiveness thus measures a specific cost of information perception.**

Expressiveness and **Effectiveness**

- **Effectiveness**
  - **A visualization is effective when it can be interpreted accurately and quickly and when it can be rendered in a cost-effective manner.**
  - **Effectiveness thus measures a specific cost of information perception.**
  - $M_{eff} = 1 / (1 + time_{interpret} + \mathrm{time}_{render})$.

Expressiveness and **Effectiveness**

- **Effectiveness**
  - **A visualization is effective when it can be interpreted accurately and quickly and when it can be rendered in a cost-effective manner.**
  - **Effectiveness thus measures a specific cost of information perception.**
  - $M_{eff} = 1 / (1 + time_{interpret} + \text{time}_{render})$.
  - $0 < M_{eff} \leq 1$.

Expressiveness and **Effectiveness**

- **Effectiveness**
  - **A visualization is effective when it can be interpreted accurately and quickly and when it can be rendered in a cost-effective manner.**
  - **Effectiveness thus measures a specific cost of information perception.**
  - $M_{eff} = 1 / (1 + time_{interpret} + time_{render})$.
  - $0 < M_{eff} \leq 1$.
  - **The larger M_{eff} is, the greater the visualization's effectiveness.**

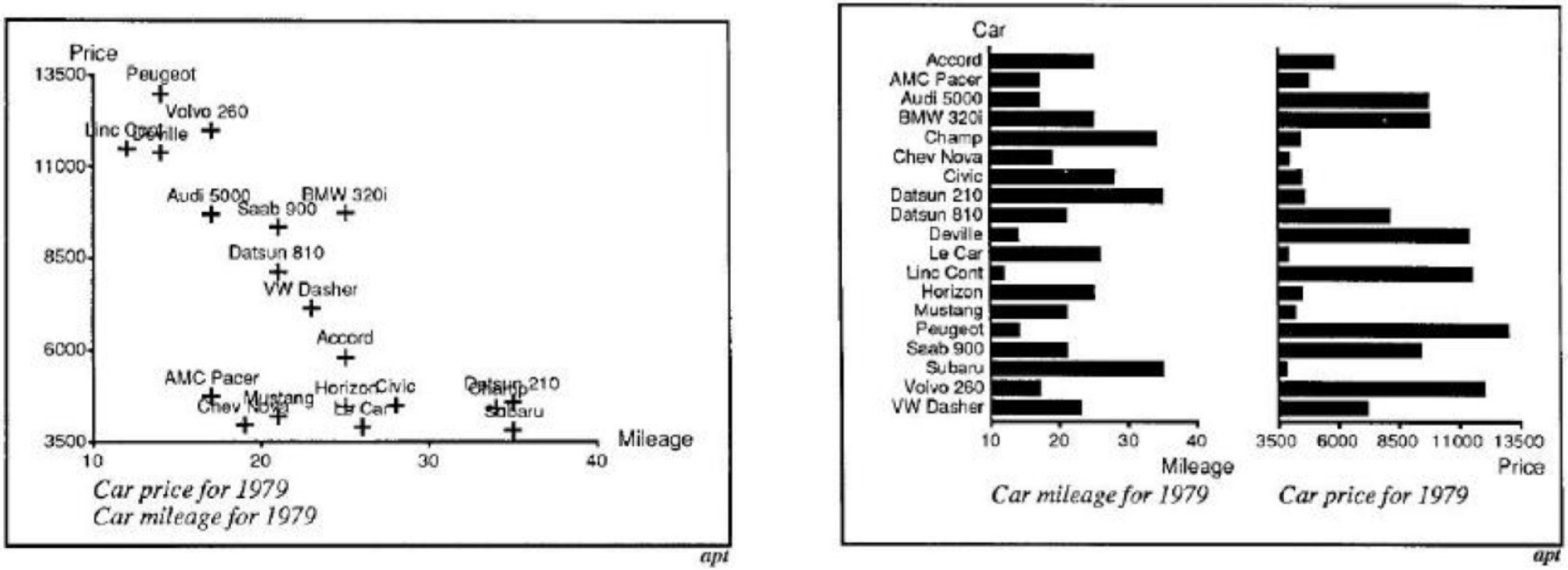
Expressiveness and **Effectiveness**

- **Effectiveness**
  - **A visualization is effective when it can be interpreted accurately and quickly and when it can be rendered in a cost-effective manner.**
  - **Effectiveness thus measures a specific cost of information perception.**
  - $M_{eff} = 1 / (1 + time_{interpret} + time_{render})$.
  - $0 < M_{eff} \leq 1$.
  - **The larger M_{eff} is, the greater the visualization's effectiveness.**
  - **If M_{eff} is small, then either the interpretation time is very large, or the rendering time is large.**

Expressiveness and **Effectiveness**

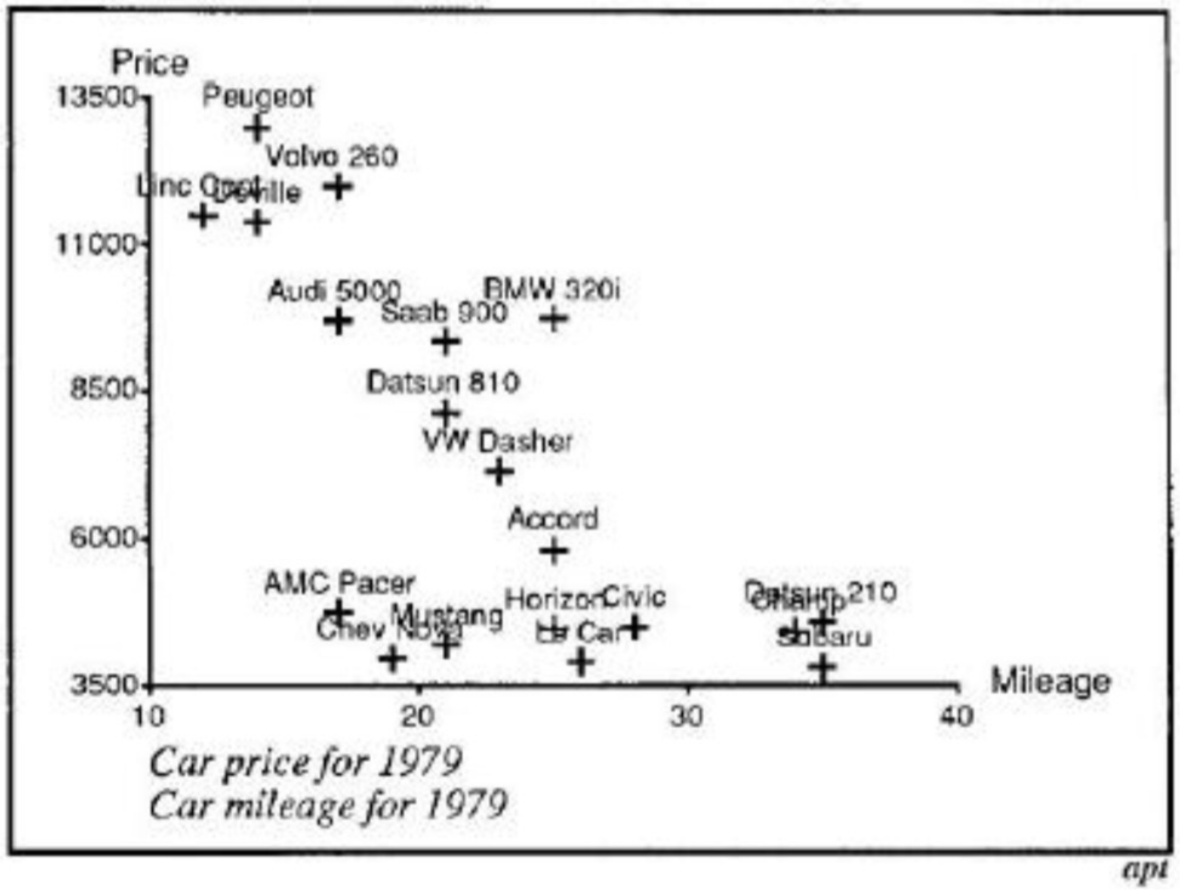
- **Effectiveness**
  - **A visualization is effective when it can be interpreted accurately and quickly and when it can be rendered in a cost-effective manner.**
  - **Effectiveness thus measures a specific cost of information perception.**
  - $M_{eff} = 1 / (1 + time_{interpret} + \text{time}_{render})$.
  - $0 < M_{eff} \leq 1$.
  - **The larger M_{eff} is, the greater the visualization's effectiveness.**
  - **If M_{eff} is small, then either the interpretation time is very large, or the rendering time is large.**
  - **If M_{eff} is large (close to 1), then both the interpretation and the rendering time are very small.**

Expressiveness and Effectiveness

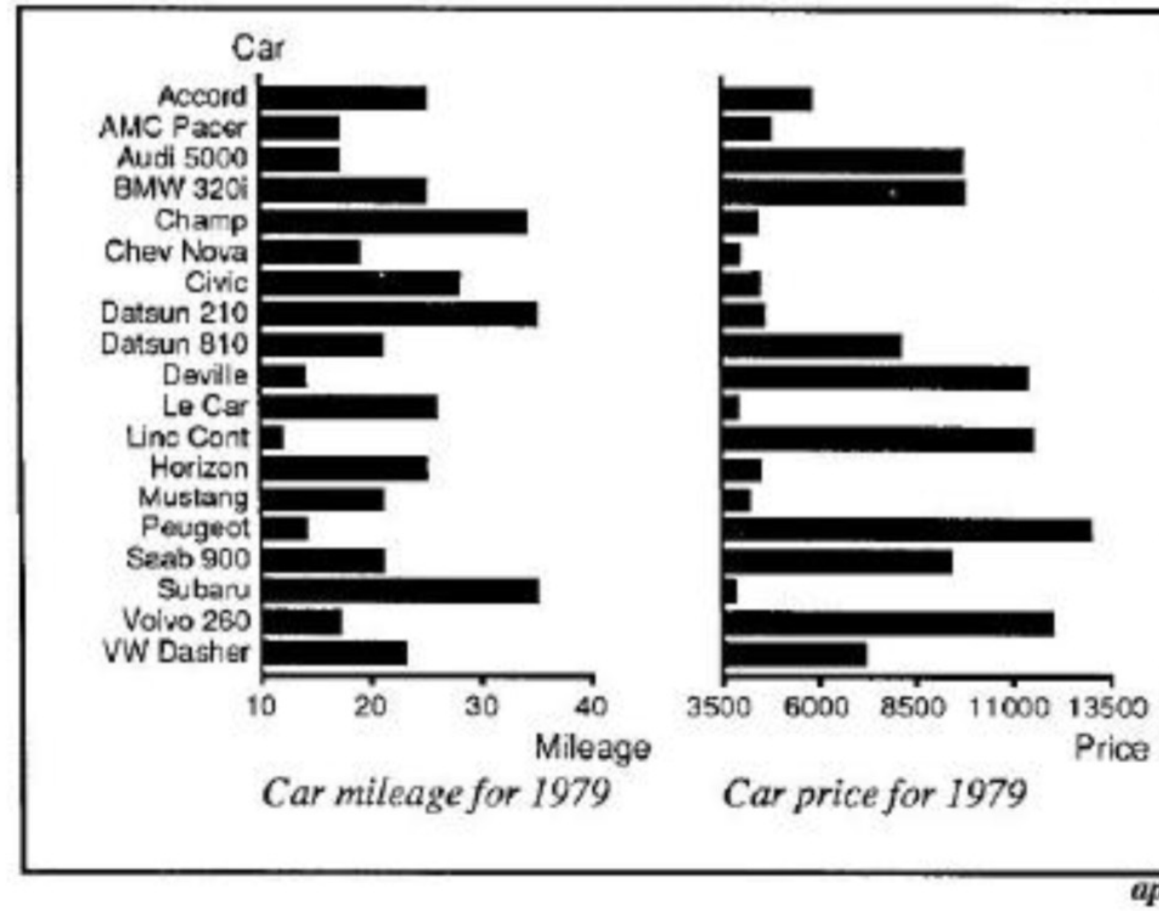


(a) (b)

Expressiveness and Effectiveness



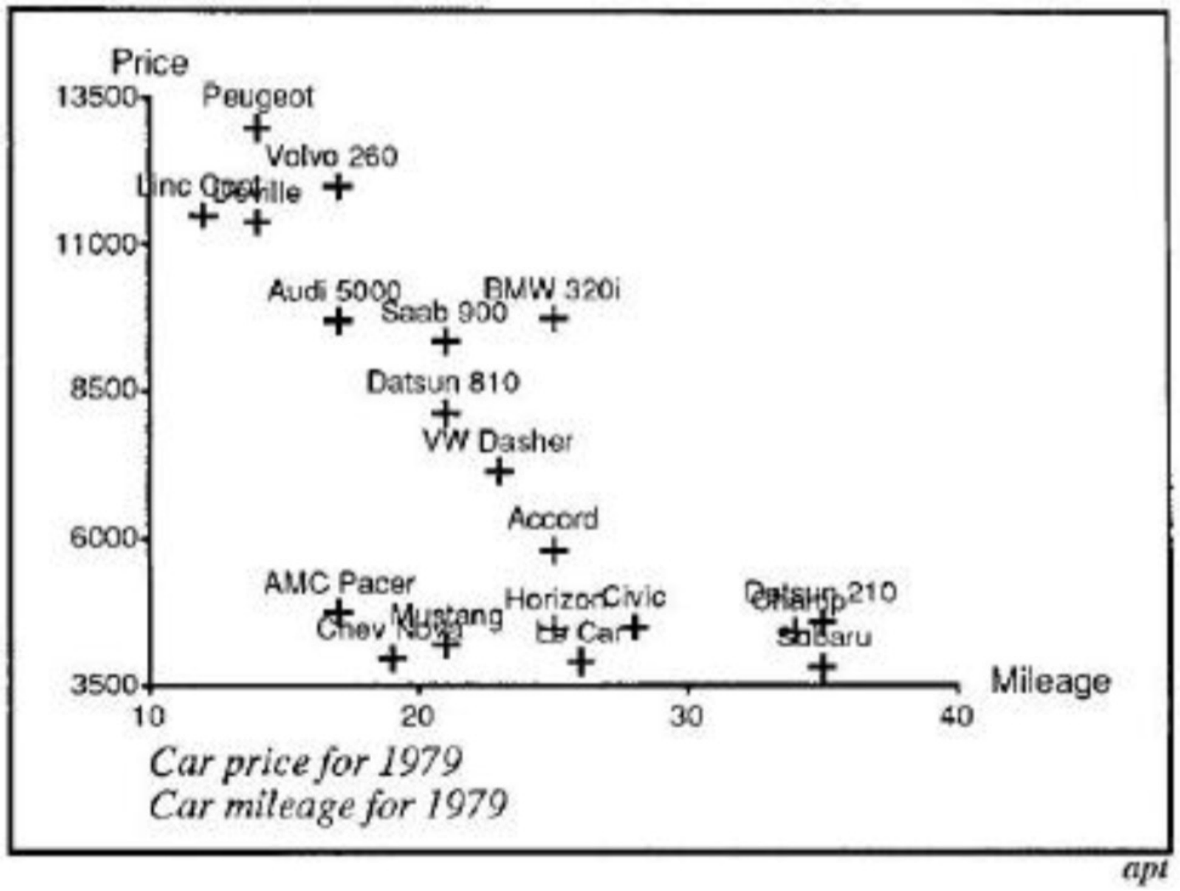
(a)



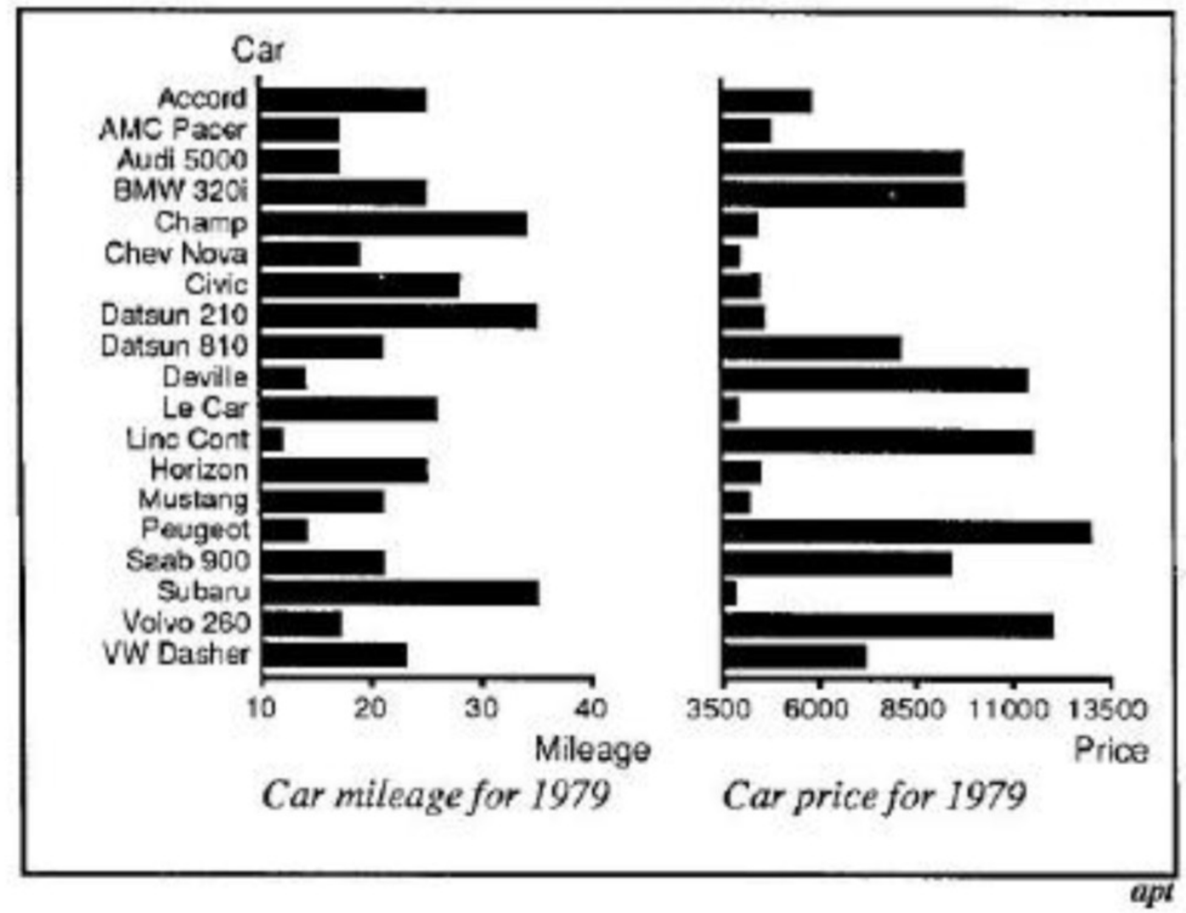
(b)

$$\boxed{M_{exp}(\text{a}) \approx M_{exp}(\text{b})}$$

Expressiveness and Effectiveness



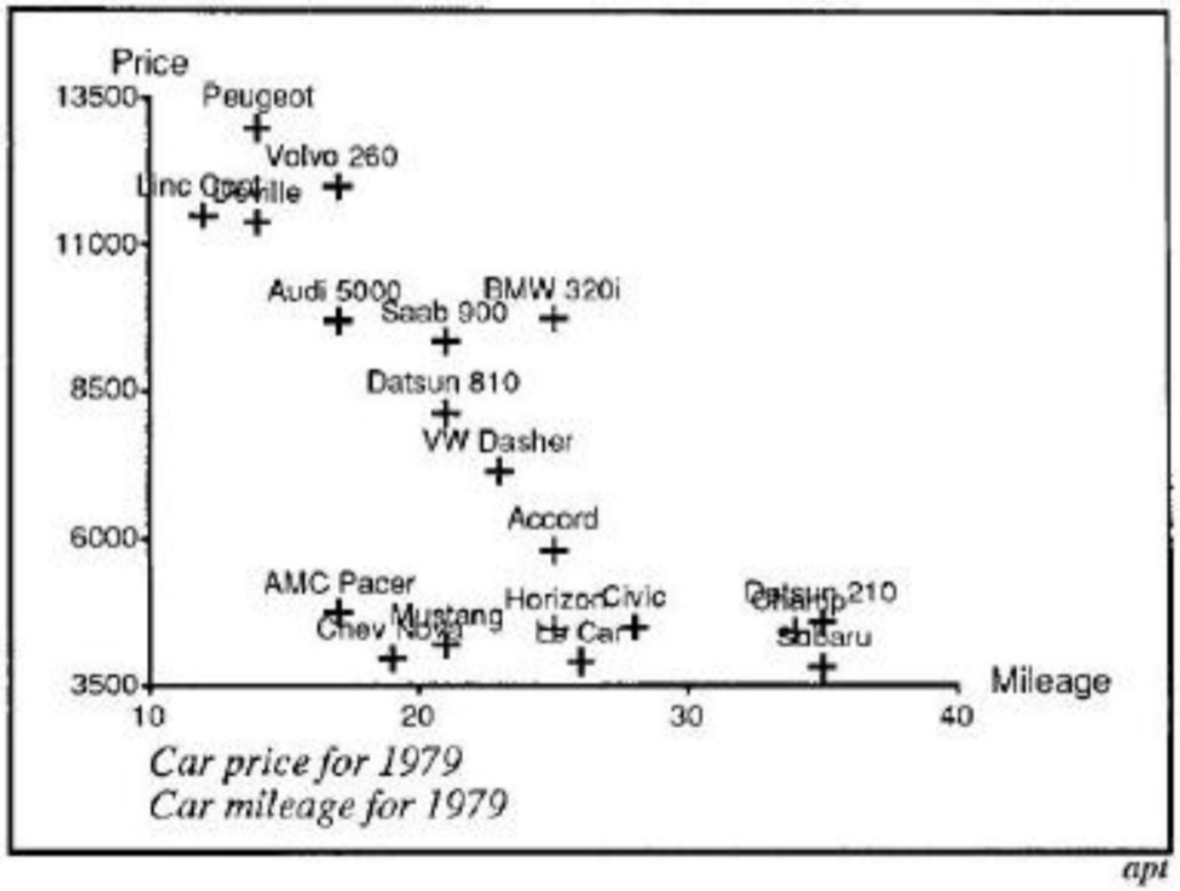
(a)



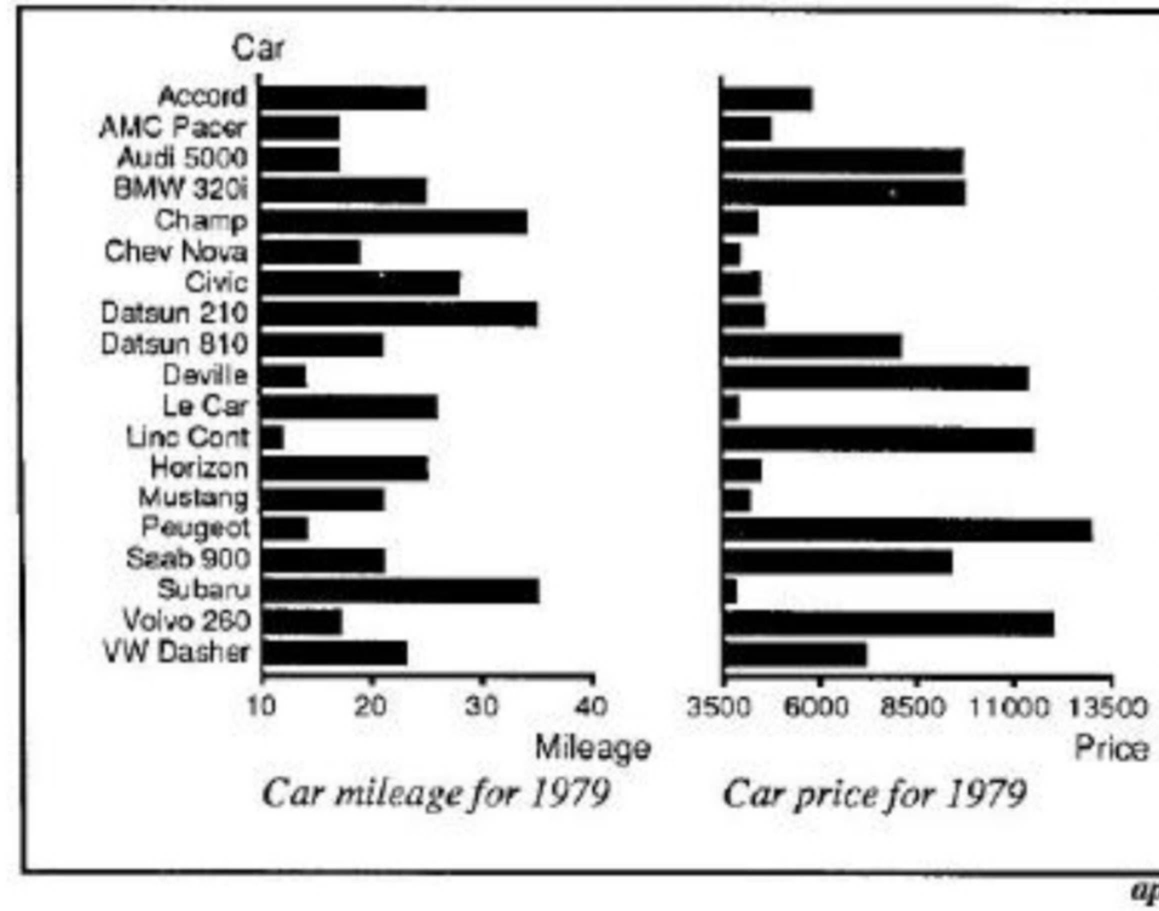
(b)

The information in (b) can be interpreted more accurately or more quickly than that in (a) for some questions. For example, which car has the best mileage?

Expressiveness and Effectiveness



(a)

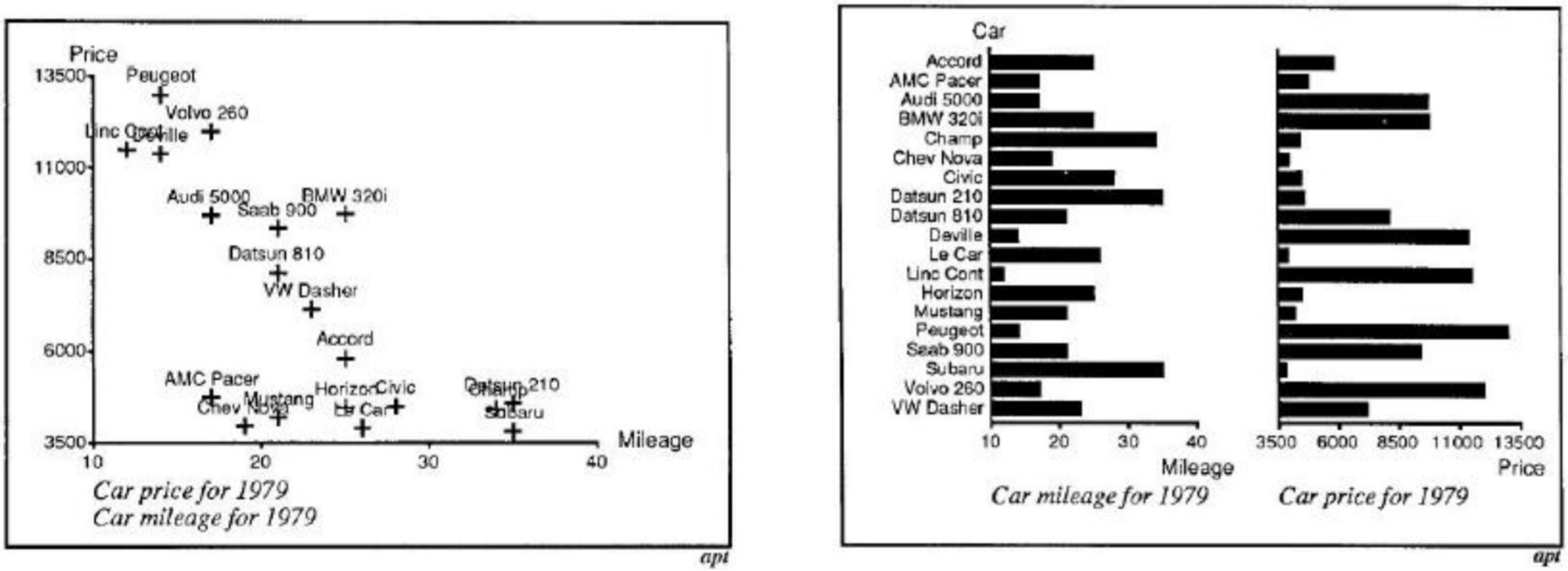


(b)

The information in (b) can be interpreted more accurately or more quickly than that in (a) for some questions. For example, which car has the best mileage?

However, if we ask which car has the best mileage under $11,000?

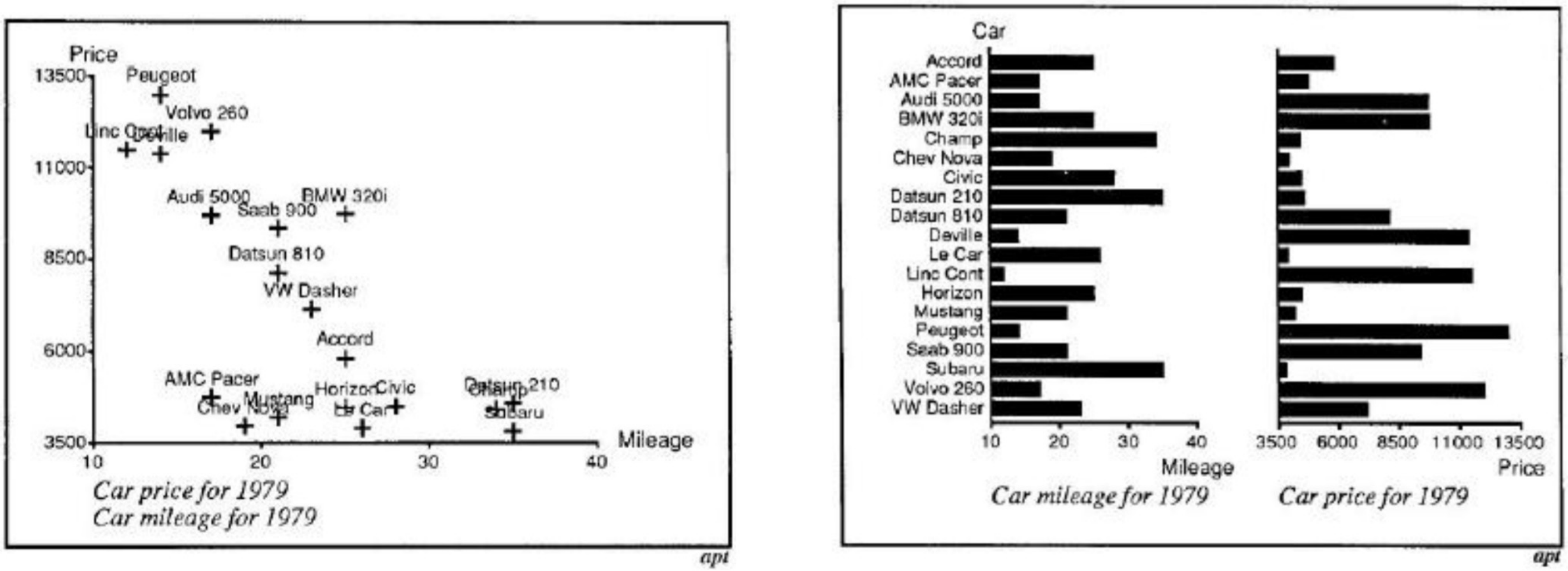
Expressiveness and Effectiveness



(a) (b)

Figure 4.3. (a) Scatterplot using plus as symbol provides good query-answering capabilities, but is slower for simple one-variable queries. (b) Bar charts clearly display cost and mileage, but don't provide as much flexibility in answering some other queries.

Expressiveness and Effectiveness



(a) (b)

Figure 4.3. (a) Scatterplot using plus as symbol provides good query-answering capabilities, but is slower for simple one-variable queries. (b) Bar charts clearly display cost and mileage, but don't provide as much flexibility in answering some other queries.

Semiology of Graphical Symbols

Semiology of Graphical Symbols

Semiology of Graphical Symbols

- **The science of graphical symbols and marks is called semiology.**

Semiology of Graphical Symbols

- **The science of graphical symbols and marks is called semiology.**

- **Every possible construction in the Euclidean plane is a graphical representation made up of graphical symbols (diagrams, networks, maps, plots, and other common visualizations).**

Semiology of Graphical Symbols

- **The science of graphical symbols and marks is called semiology.**

- **Every possible construction in the Euclidean plane is a graphical representation made up of graphical symbols (diagrams, networks, maps, plots, and other common visualizations).**

- **Semiology uses the qualities of the plane and objects on the plane to produce similarity features, ordering features, and proportionality features of the data that are visible for human consumption.**

Semiology of Graphical Symbols

- **Discovery of relations or patterns occurs through a mapping between any relationship of the graphic symbols and the data that these symbols represent.**

Semiology of Graphical Symbols

- **Discovery of relations or patterns occurs through a mapping between any relationship of the graphic symbols and the data that these symbols represent.**
  - **any pattern on the screen must imply a pattern in the data.**

Semiology of Graphical Symbols

- **Discovery of relations or patterns occurs through a mapping between any relationship of the graphic symbols and the data that these symbols represent.**
  - **any pattern on the screen must imply a pattern in the data.**
    - **If it does not, then it is an artifact of the selected representation (and is disturbing).**

Semiology of Graphical Symbols

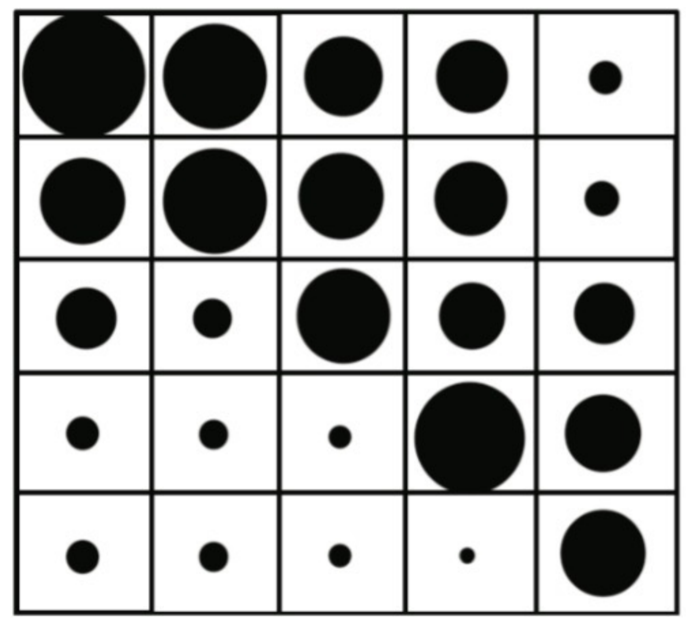
- Discovery of relations or patterns occurs through a mapping between any relationship of the graphic symbols and the data that these symbols represent.
  - any pattern on the screen must imply a pattern in the data.
    - If it does not, then it is an artifact of the selected representation (and is disturbing).
  - Similarly, any perceived pattern variation in the graphic or symbol cognitively implies such a similar variation in the data.

Semiology of Graphical Symbols

- Discovery of relations or patterns occurs through a mapping between any relationship of the graphic symbols and the data that these symbols represent.
  - any pattern on the screen must imply a pattern in the data.
    - If it does not, then it is an artifact of the selected representation (and is disturbing).
  - Similarly, any perceived pattern variation in the graphic or symbol cognitively implies such a similar variation in the data.
  - Any perceived order in graphic symbols is directly correlated with a perceived corresponding order between the data, and vice versa

Features of Graphics

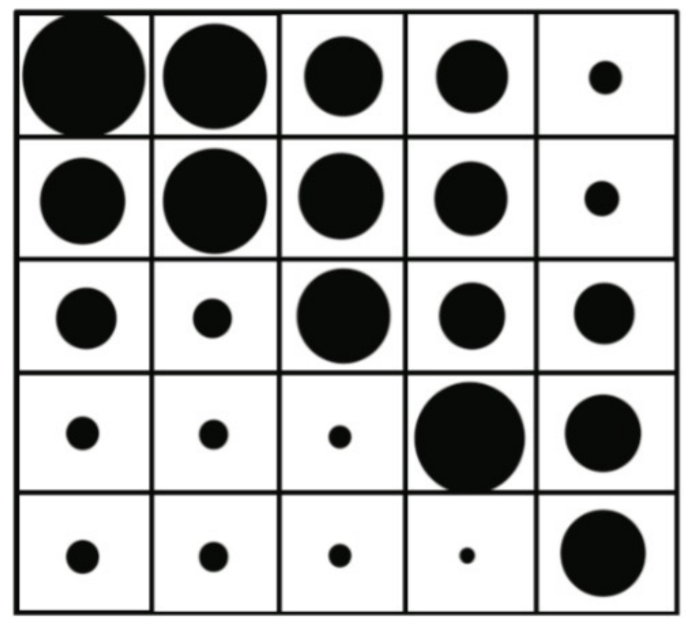
- **Graphics have three (or more) dimensions.**



Matrix representation of a set of relationships between nodes in a graph. The size represents the strength of the relationship.

Features of Graphics

- **Graphics have three (or more) dimensions.**



Matrix representation of a set of relationships between nodes in a graph. The size represents the strength of the relationship.

- **Every point of the graphic can be interpreted as a relation between a position in x and a position in y. The points vary in size, providing a third dimension or variable to interpret.**

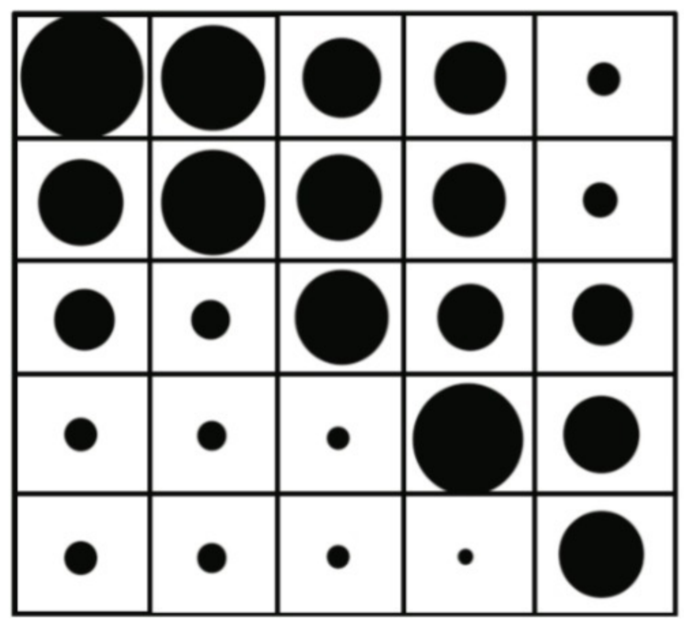
Rules of Graphics

Rules of Graphics

- **The aim of a graphic is to discover groups or orders in x, and groups or orders in y, that are formed on z-values;**

Rules of Graphics

- **The aim of a graphic is to discover groups or orders in x, and groups or orders in y, that are formed on z-values;**



Matrix representation of a set of relationships between nodes in a graph. The size represents the strength of the relationship.

Rules of Graphics

- **The aim of a graphic is to discover groups or orders in x, and groups or orders in y, that are formed on z-values;**
- **(x, y, z)-construction enables in all cases the discovery of these groups;**

Rules of Graphics

- **The aim of a graphic is to discover groups or orders in x, and groups or orders in y, that are formed on z-values;**
- **(x, y, z)-construction enables in all cases the discovery of these groups;**
- **Within the (x,y,z)-construction, permutations and classifications solve the problem of the upper level of information;**

Rules of Graphics

- The aim of a graphic is to discover groups or orders in x, and groups or orders in y, that are formed on z-values;
- (x, y, z)-construction enables in all cases the discovery of these groups;
- Within the (x,y,z)-construction, permutations and classifications solve the problem of the upper level of information;
- Every graphic with more than three factors that differs from the (x, y, z)-construction destroys the unity of the graphic and the upper level of information;

Rules of Graphics

- The aim of a graphic is to discover groups or orders in x, and groups or orders in y, that are formed on z-values;
- (x, y, z)-construction enables in all cases the discovery of these groups;
- Within the (x,y,z)-construction, permutations and classifications solve the problem of the upper level of information;
- Every graphic with more than three factors that differs from the (x, y, z)-construction destroys the unity of the graphic and the upper level of information;
- Pictures must be read and understood by the human.

The Eight Visual Variables

Spatial arrangement of marks

Spatial arrangement of marks

- **For the most part, all graphic primitives will be termed marks.**

Spatial arrangement of marks

- **For the most part, all graphic primitives will be termed marks.**
- **One way to encode data for display is to map different data values to different marks and their attributes.**

Spatial arrangement of marks

- **For the most part, all graphic primitives will be termed marks.**
- **One way to encode data for display is to map different data values to different marks and their attributes.**
- **However, marks by themselves do not define informative displays, since all the marks would simply obscure all previously drawn marks; it is only through the spatial arrangement of marks that informative displays are created.**

Spatial arrangement of marks

- For the most part, all graphic primitives will be termed marks.
- One way to encode data for display is to map different data values to different marks and their attributes.
- However, marks by themselves do not define informative displays, since all the marks would simply obscure all previously drawn marks; it is only through the spatial arrangement of marks that informative displays are created.
- Once the layout and types of marks are specified, then additional graphical properties can be applied to each mark.

Spatial arrangement of marks

- For the most part, all graphic primitives will be termed **marks**.
- One way to encode data for display is to **map different data values to different marks and their attributes**.
- However, marks by themselves do not define informative displays, since all the marks would simply obscure all previously drawn marks; **it is only through the spatial arrangement of marks that informative displays are created**.
- Once the layout and types of marks are specified, then additional graphical properties can be applied to each mark.
  - Marks can vary in **size**, can be displayed using different **colors**, and can be mapped to different **orientations**, all of which can be **driven by data to convey information.**

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

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Eight visual variables

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

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

Eight visual variables: Position

- **The first and most important visual variable is that of position, the placement of representative graphics within some display space, be it one-, two-, or three-dimensional.**

Eight visual variables: Position

- The first and **most important visual variable is that of position**, the placement of representative graphics within some display space, be it one-, two-, or three-dimensional.
- 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.

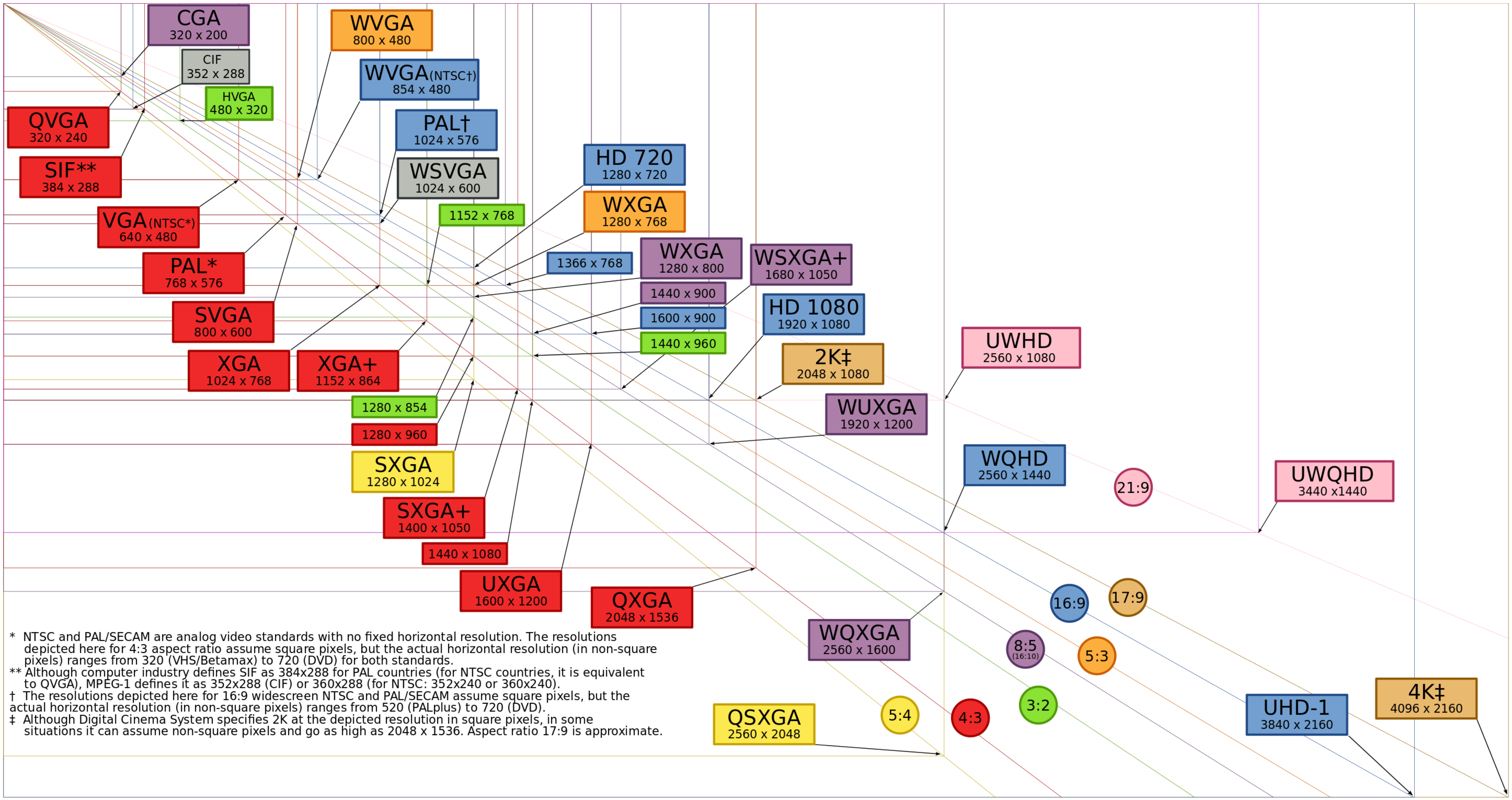
Eight visual variables: Position

- The first and **most important visual variable is that of position**, the placement of representative graphics within some display space, be it one-, two-, or three-dimensional.
- 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**

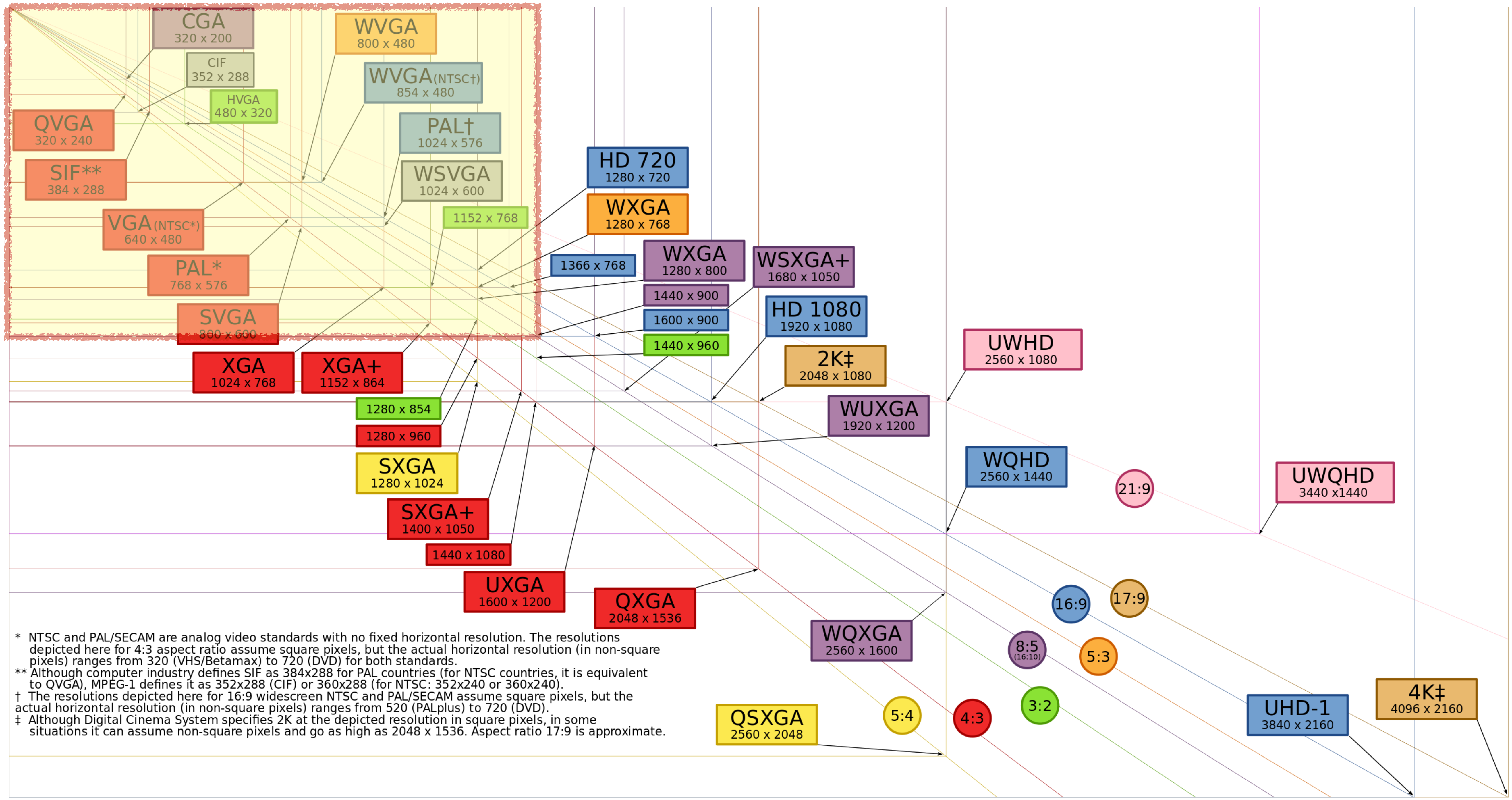
Eight visual variables: Position

- The first and most important visual variable is that of position, the placement of representative graphics within some display space, be it one-, two-, or three-dimensional.
- 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: Screen resolution

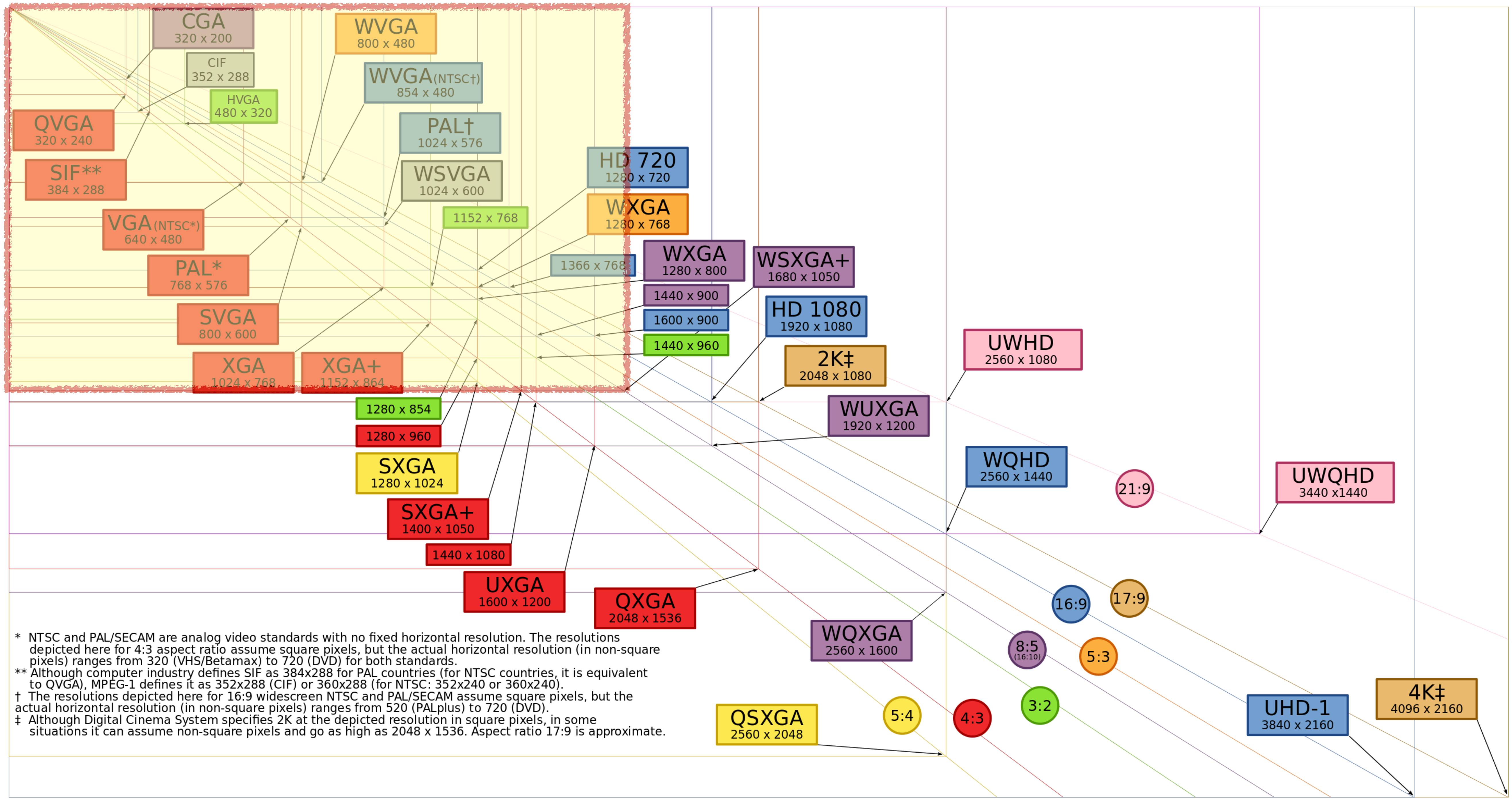


Eight visual variables: Screen resolution

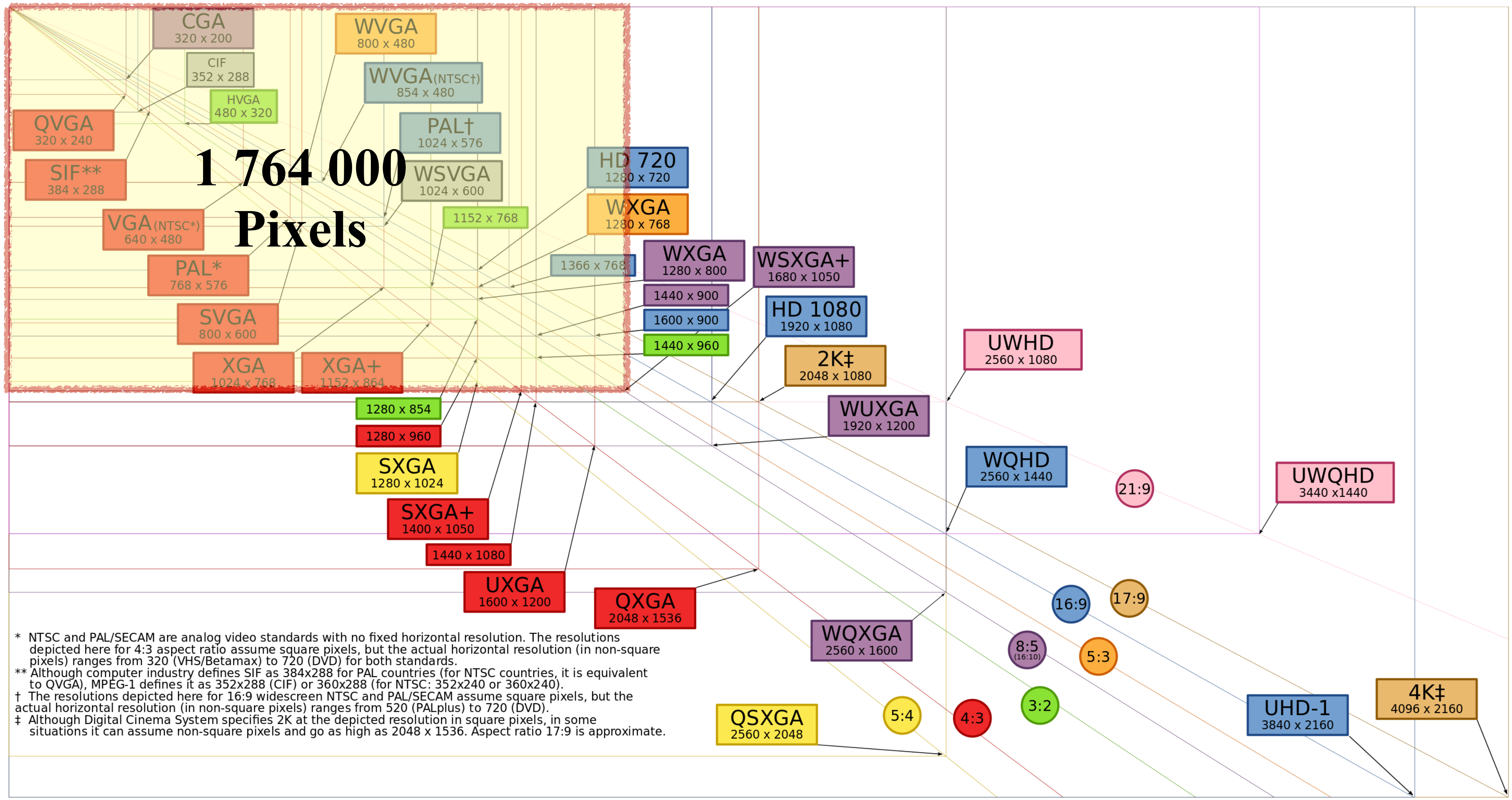


* NTSC and PAL/SECAM are analog video standards with no fixed horizontal resolution. The resolutions depicted here for 4:3 aspect ratio assume square pixels, but the actual horizontal resolution (in non-square pixels) ranges from 320 (VHS/Betamax) to 720 (DVD) for both standards.
** Although computer industry defines SIF as 384x288 for PAL countries (for NTSC countries, it is equivalent to QVGA), MPEG-1 defines it as 352x288 (CIF) or 360x288 (for NTSC: 352x240 or 360x240).
† The resolutions depicted here for 16:9 widescreen NTSC and PAL/SECAM assume square pixels, but the actual horizontal resolution (in non-square pixels) ranges from 520 (PALplus) to 720 (DVD).
‡ Although Digital Cinema System specifies 2K at the depicted resolution in square pixels, in some situations it can assume non-square pixels and go as high as 2048 x 1536. Aspect ratio 17:9 is approximate.

Eight visual variables: Screen resolution

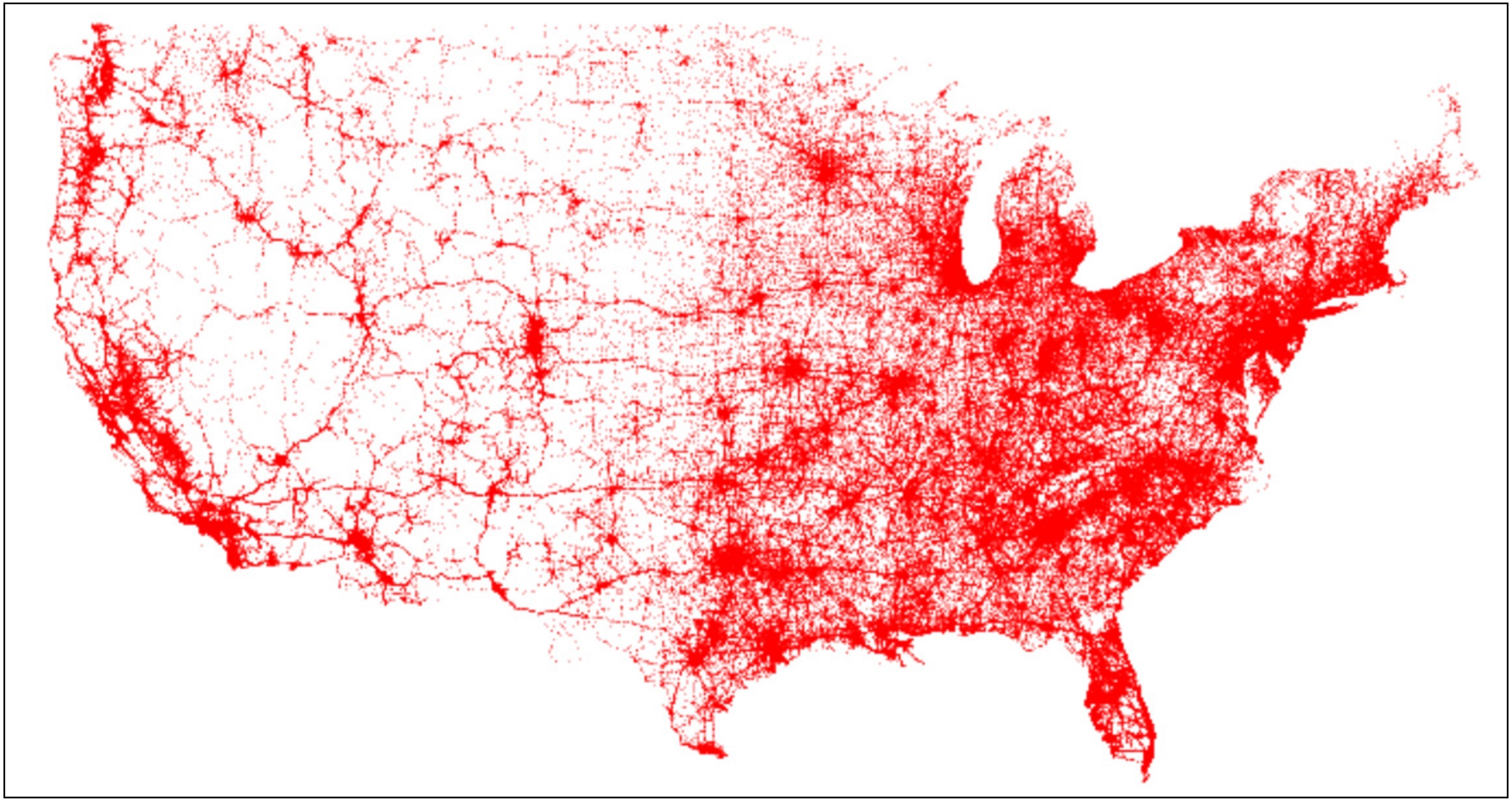


Eight visual variables: Screen resolution



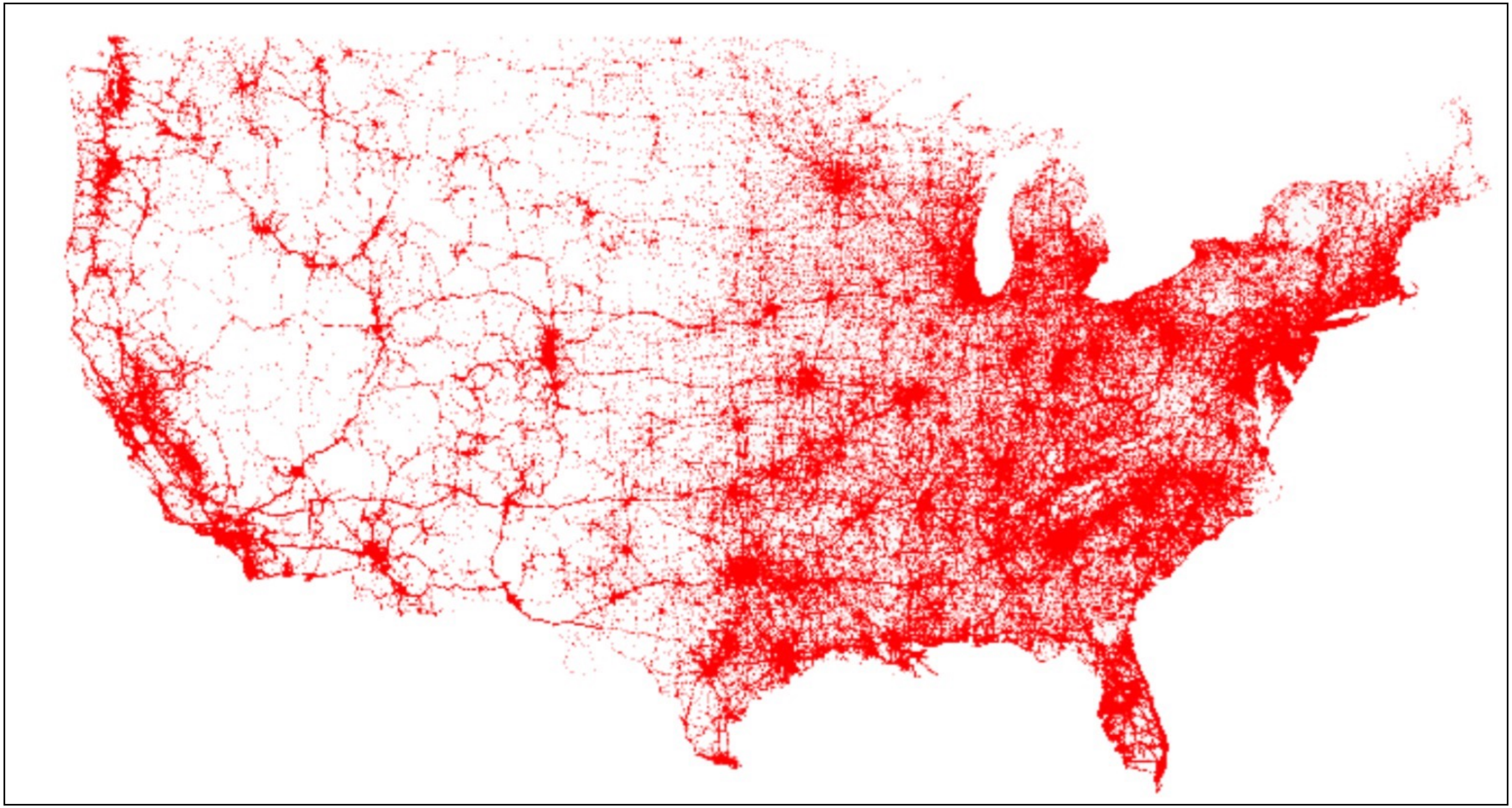
Eight visual variables: Screen resolution

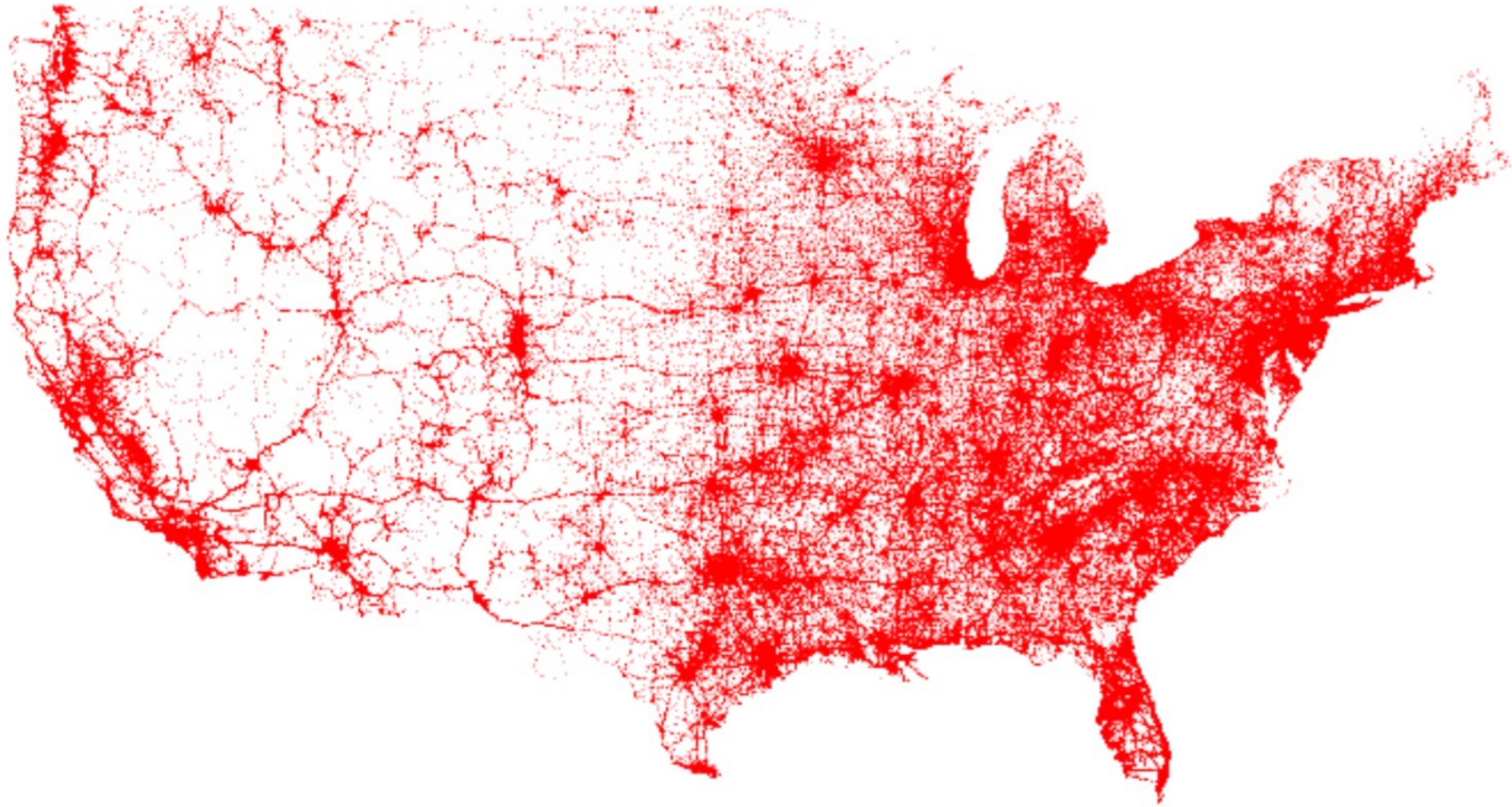
- **450.710 geo-referenced accidents between 2001 and 2013 in US**

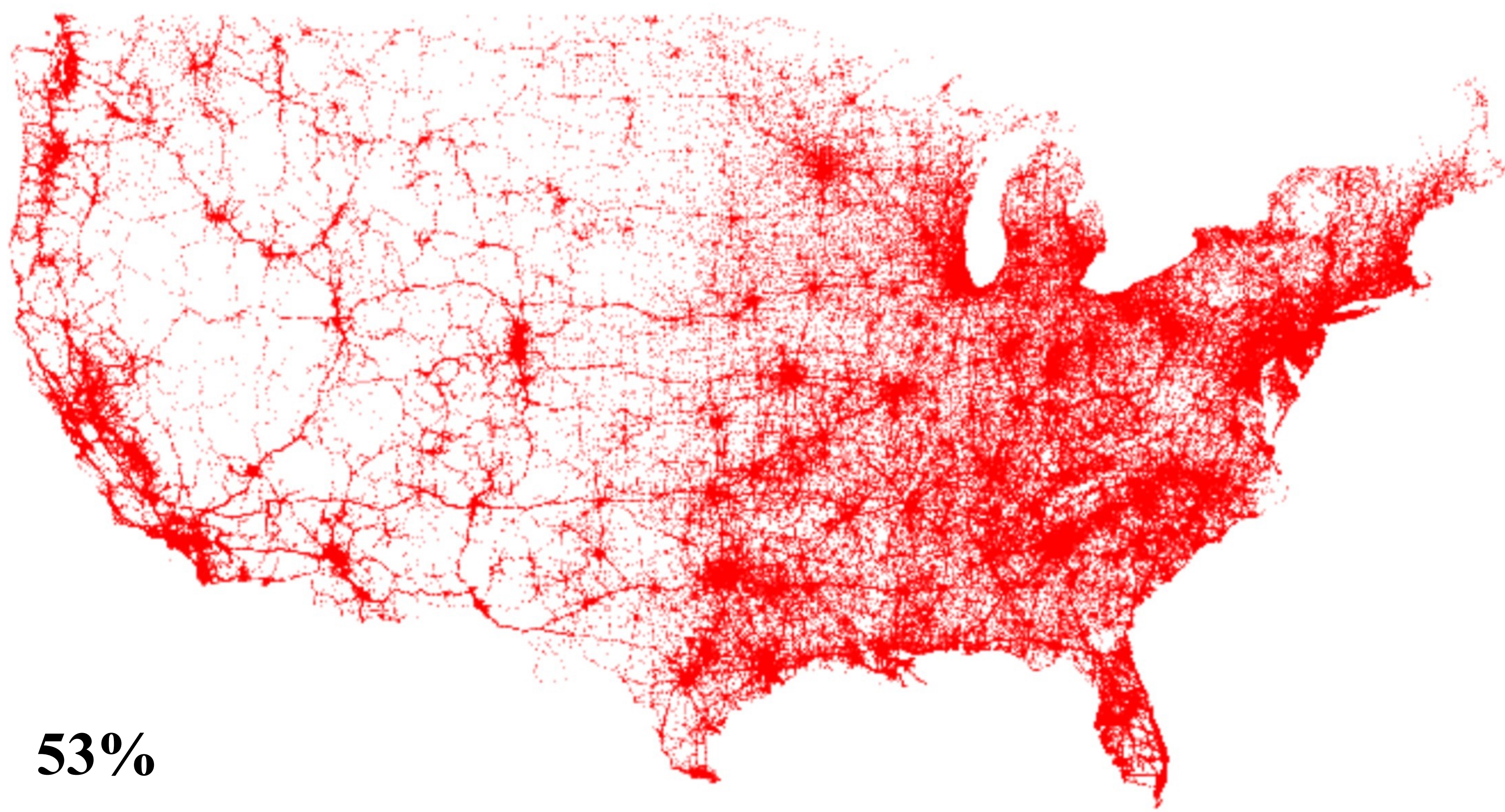


Eight visual variables: Screen resolution

- **Preprocessed data: 53% of items from original data set**

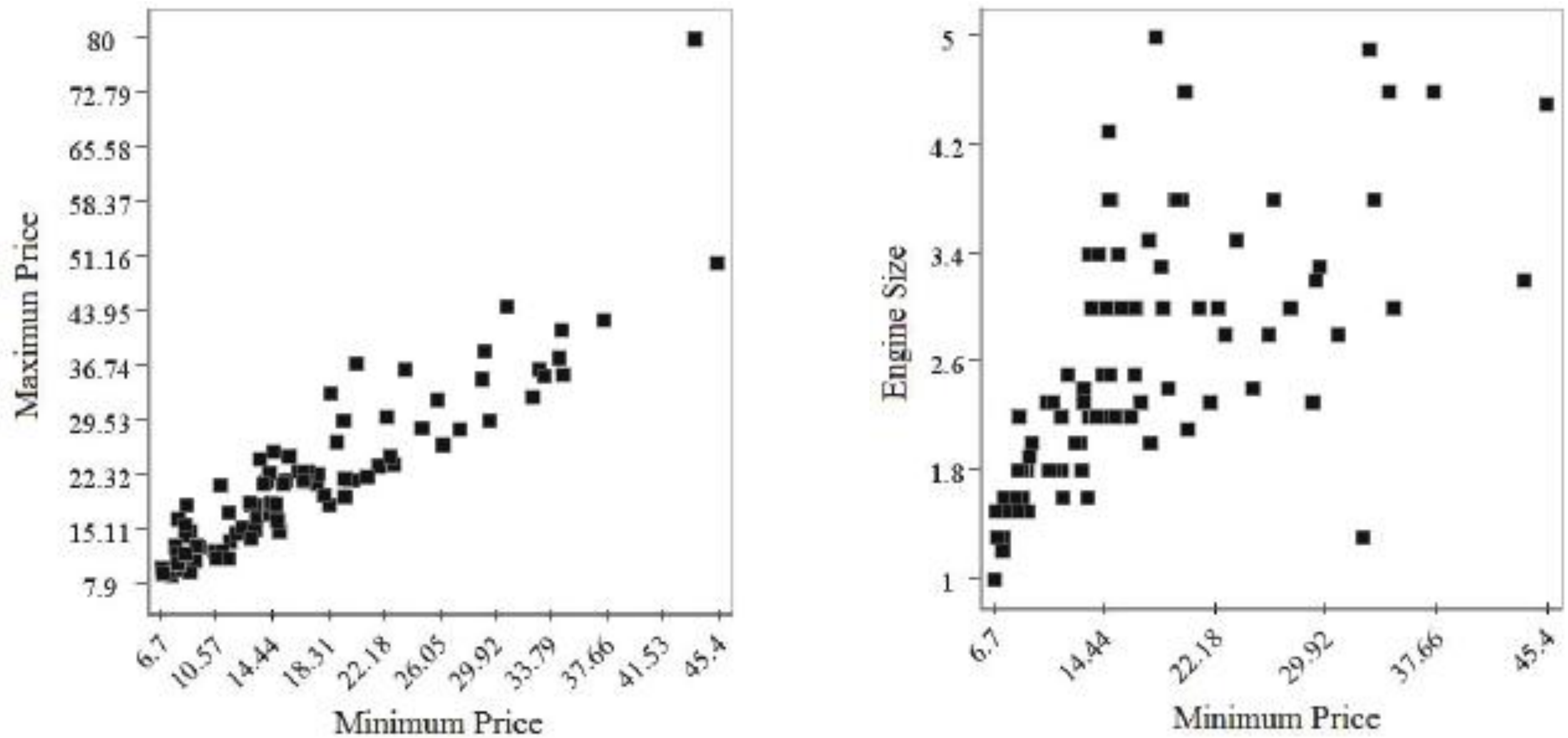






53%

Eight visual variables: Position - Scales

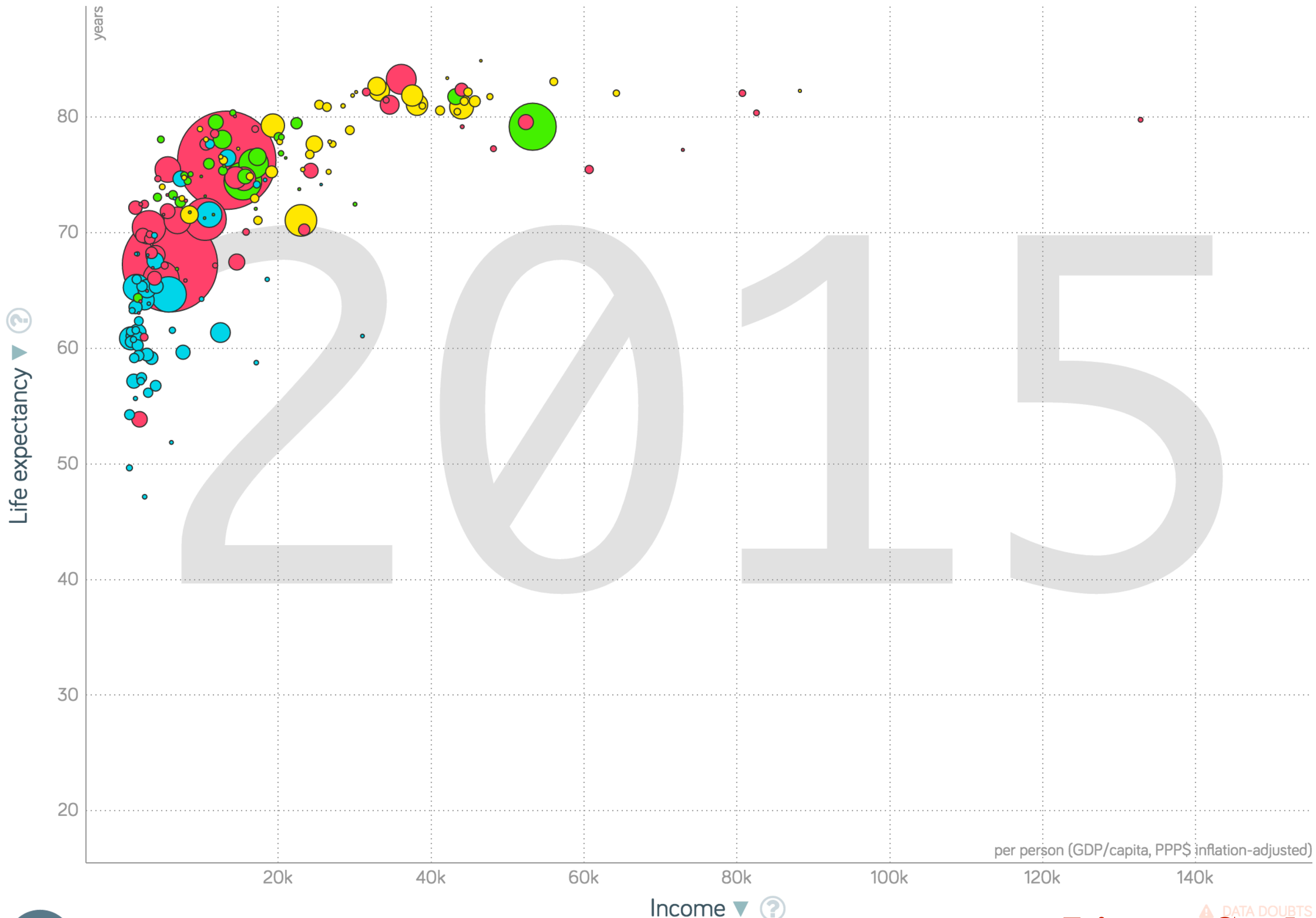


(a) (b)

Example visualizations: (a) using position to convey information. Displayed here is the minimum price versus the maximum price for cars with a 1993 model year. The spread of points appears to indicate a linear relationship between minimum and maximum price; (b) another visualization using a different set of variables. This figure compares minimum price with engine size for the 1993 cars data set. Unlike (a), there does not appear to be a strong relationship between these two variables.

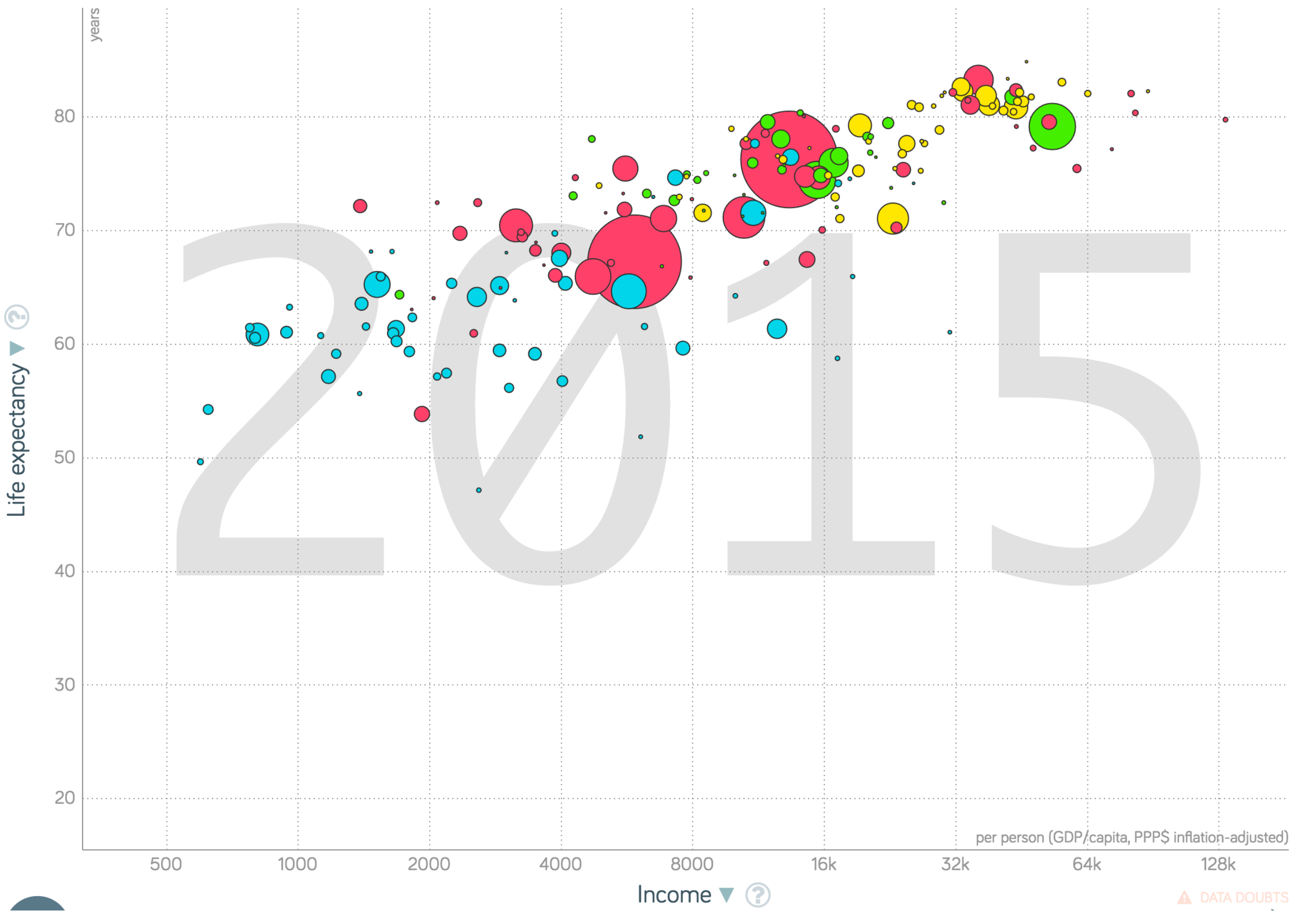
Eight visual variables: Position - Scales

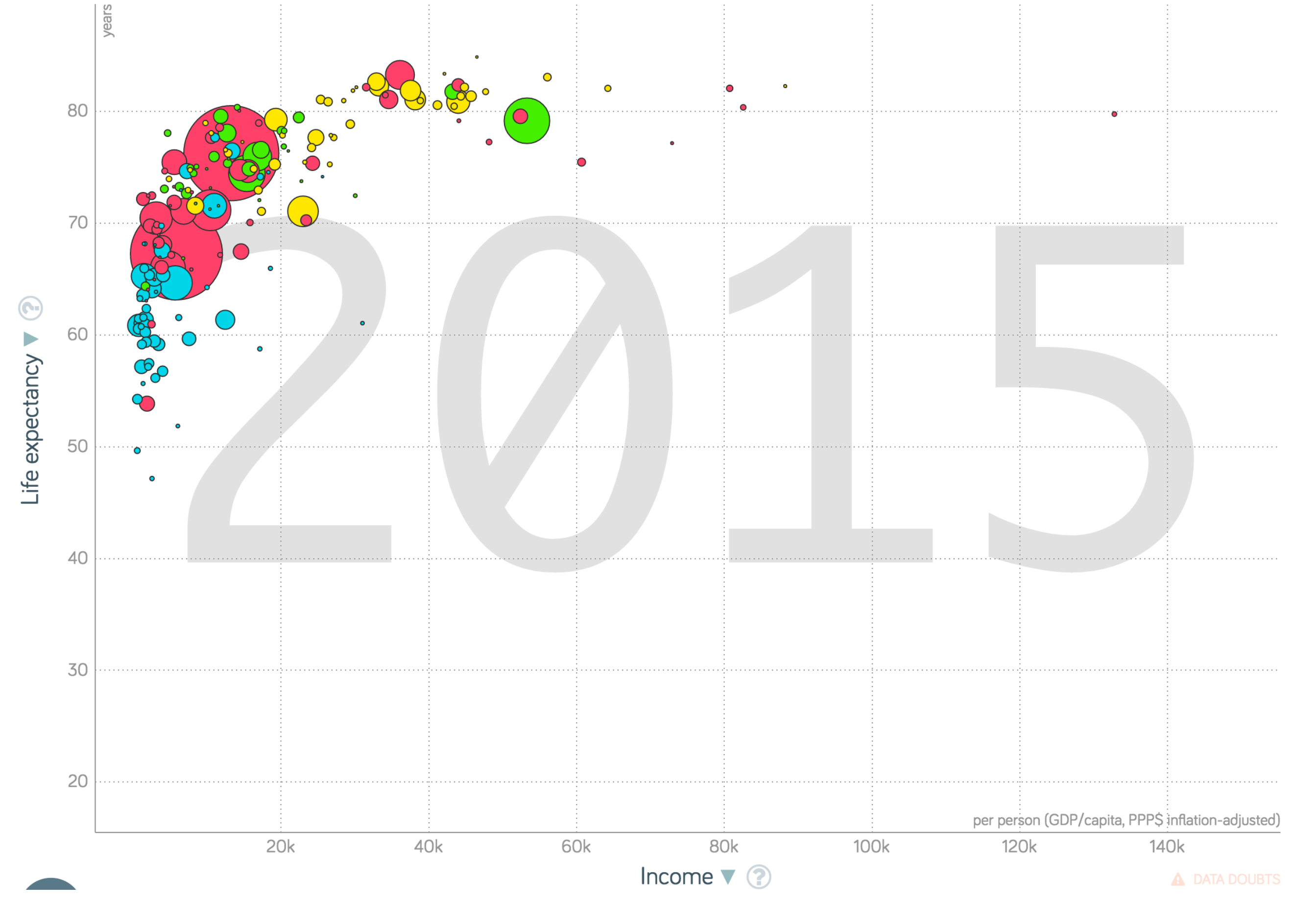
Linear Scale



Linear Scale

Eight visual variables: Position - Scales





Log Scale **Linear Scale**

Eight visual variables: Mark (or shape)

Eight visual variables: Mark (or shape)

- **The second visual variable is the mark or shape: points, lines, areas, volumes, and their compositions.**

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.

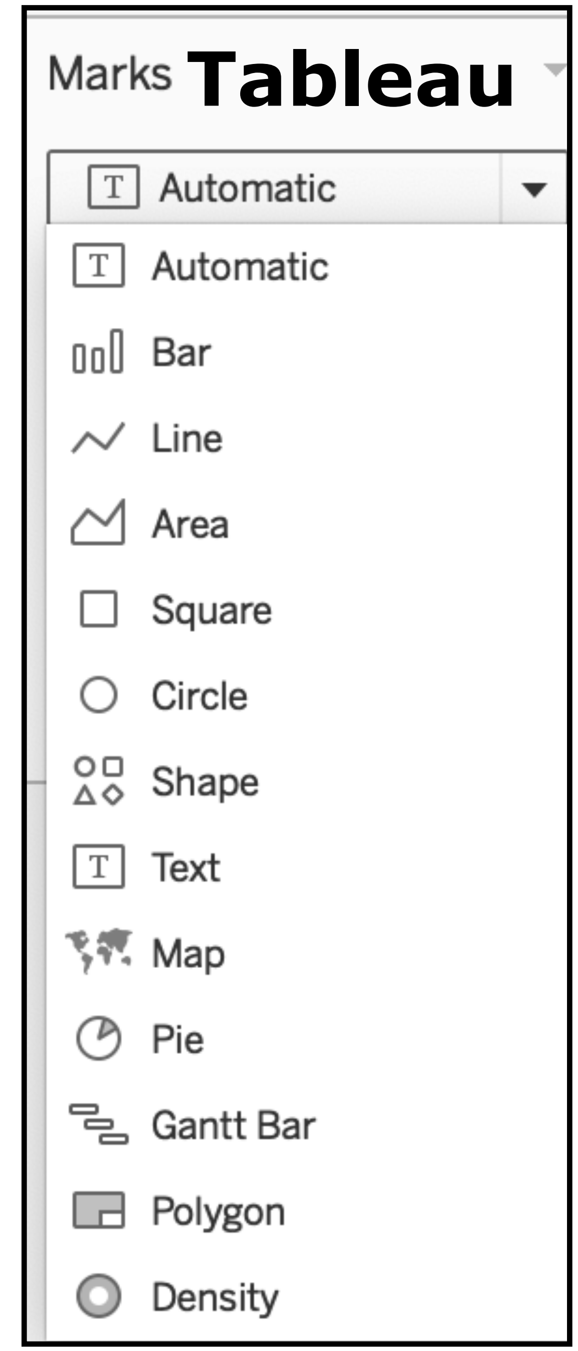
  - When using marks, it is important to consider how well one mark can be differentiated from other marks

Eight visual variables: Mark (or shape)

- **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

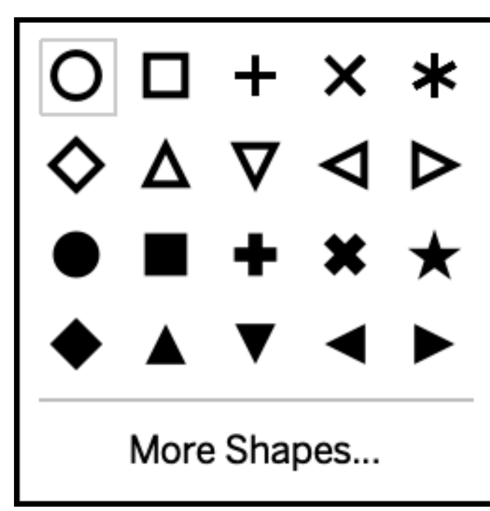


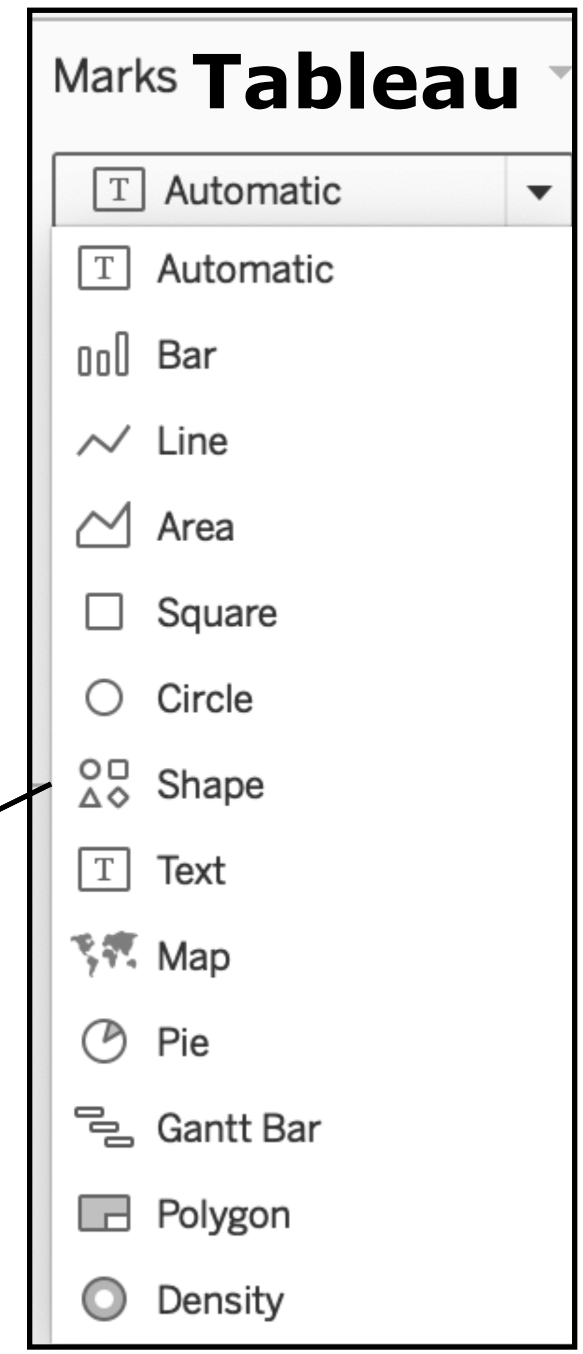
Eight visual variables: Mark (or shape)

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

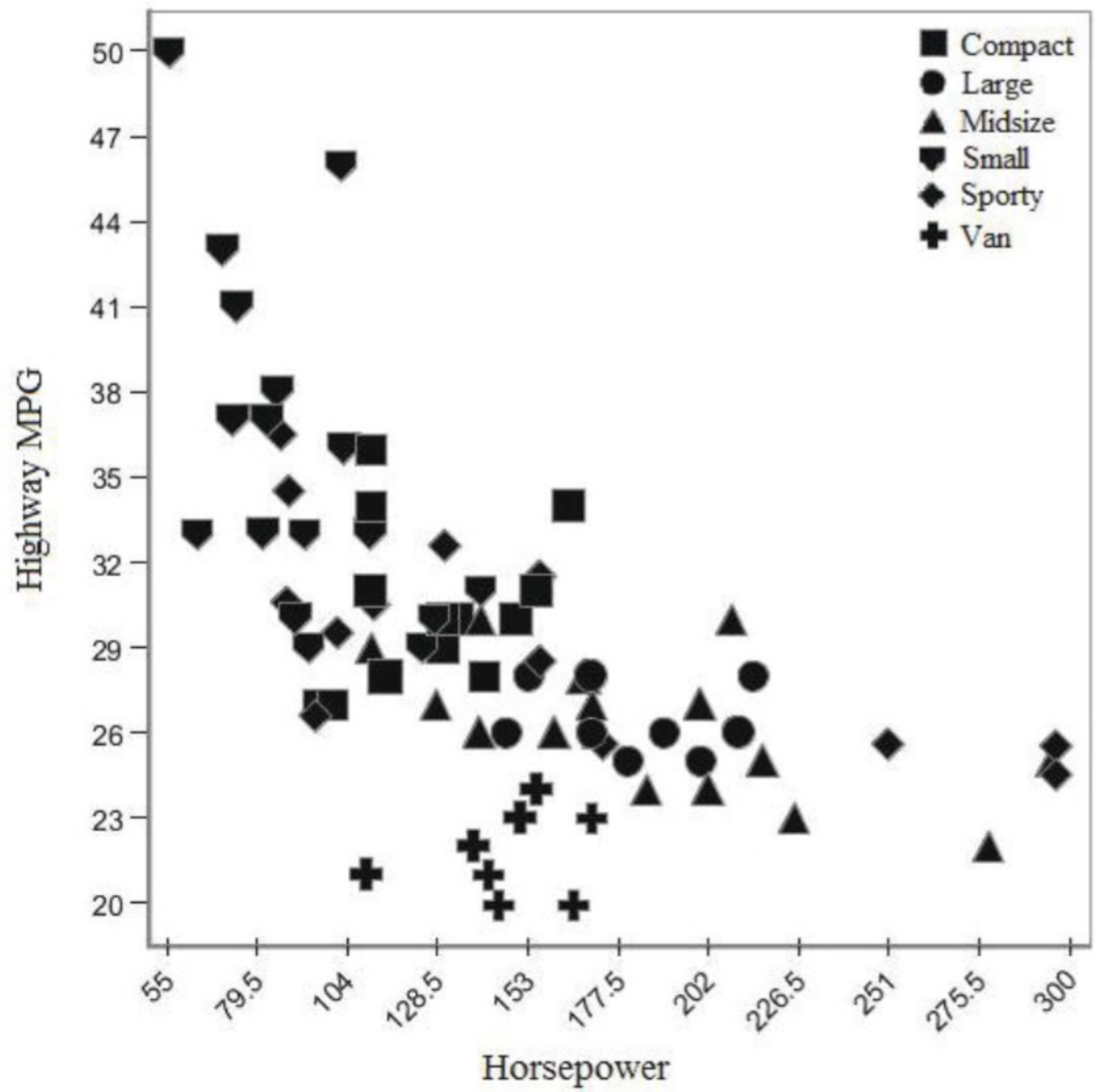
Several examples of different marks or glyphs that can be used.

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





Eight visual variables: Mark (or shape)



This visualization uses shapes to distinguish between different car types in a plot comparing highway MPG and horsepower. Clusters are clearly visible, as well as some outliers.

10 Min Interval

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: Size

Example sizes to encode data.

Eight visual variables: Size

Example sizes to encode data.

- **Size easily maps to interval and continuous data variables, because that property supports gradual increments over some range.**

Eight visual variables: Size

Example sizes to encode data.

- **Size easily maps to interval and continuous data variables, because that property supports gradual increments over some range.**
- **It is more difficult to distinguish between marks of near similar size, and thus size can only support categories with very small cardinality.**

Eight visual variables: Size

Example sizes to encode data.

- **Size easily maps to interval and continuous data variables, because that property supports gradual increments over some range.**
- **It is more difficult to distinguish between marks of near similar size, and thus size can only support categories with very small cardinality.**
- **A confounding problem with using size is the type of mark.**
  - **For points, lines, and curves the use of size works well**
  - **when marks are represented with graphics that contain sufficient area, the quantitative aspects of size fall, and the differences between marks becomes more qualitative.**

Eight visual variables: Size

Example sizes to encode data.

- **Size easily maps to interval and continuous data variables, because that property supports gradual increments over some range.**
- **It is more difficult to distinguish between marks of near similar size, and thus size can only support categories with very small cardinality.**
- **A confounding problem with using size is the type of mark.**
  - **For points, lines, and curves the use of size works well**
  - **when marks are represented with graphics that contain sufficient area, the quantitative aspects of size fall, and the differences between marks becomes more qualitative.**

Stevens' Law

Eight visual variables: Size

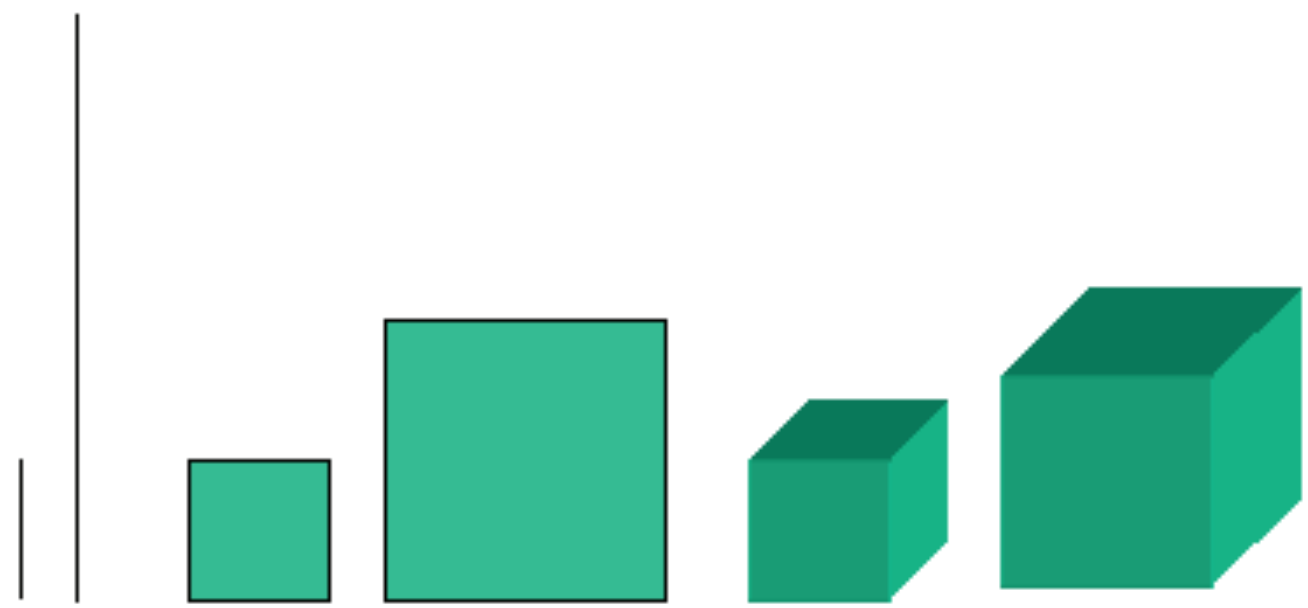


Illustration of Stevens' Law. The size ratio for each pair is 1:4. This magnitude is readily apparent in the lines, but it is easily underestimated in the squares and cubes.

Eight visual variables: Size

Stevens' Law

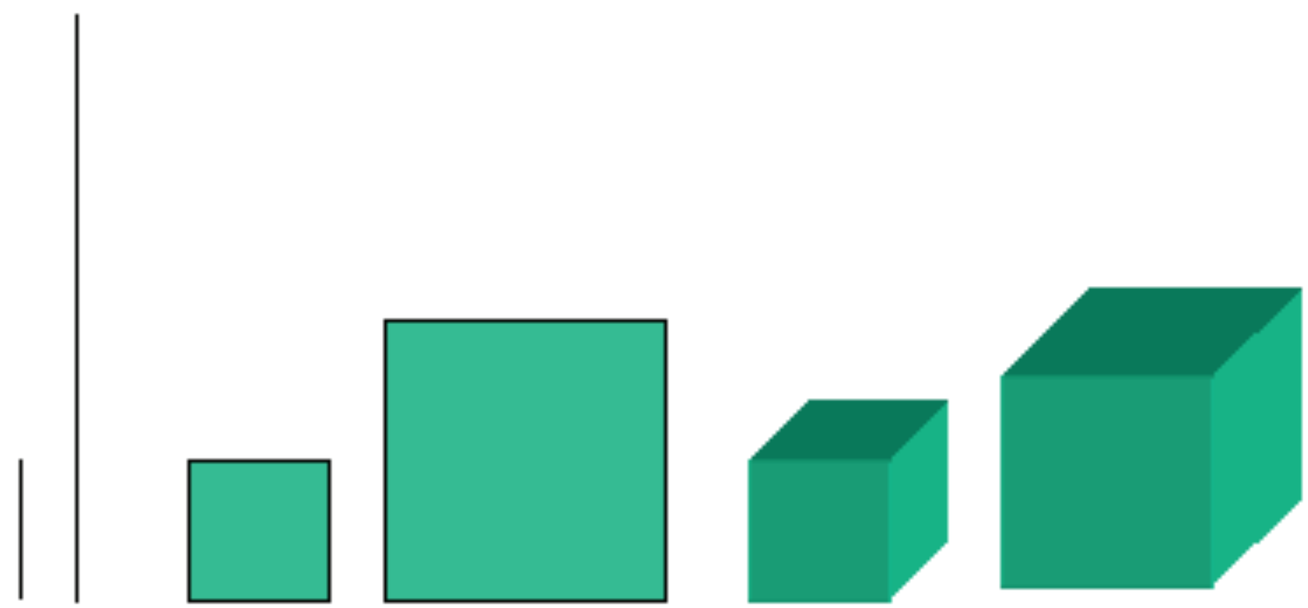
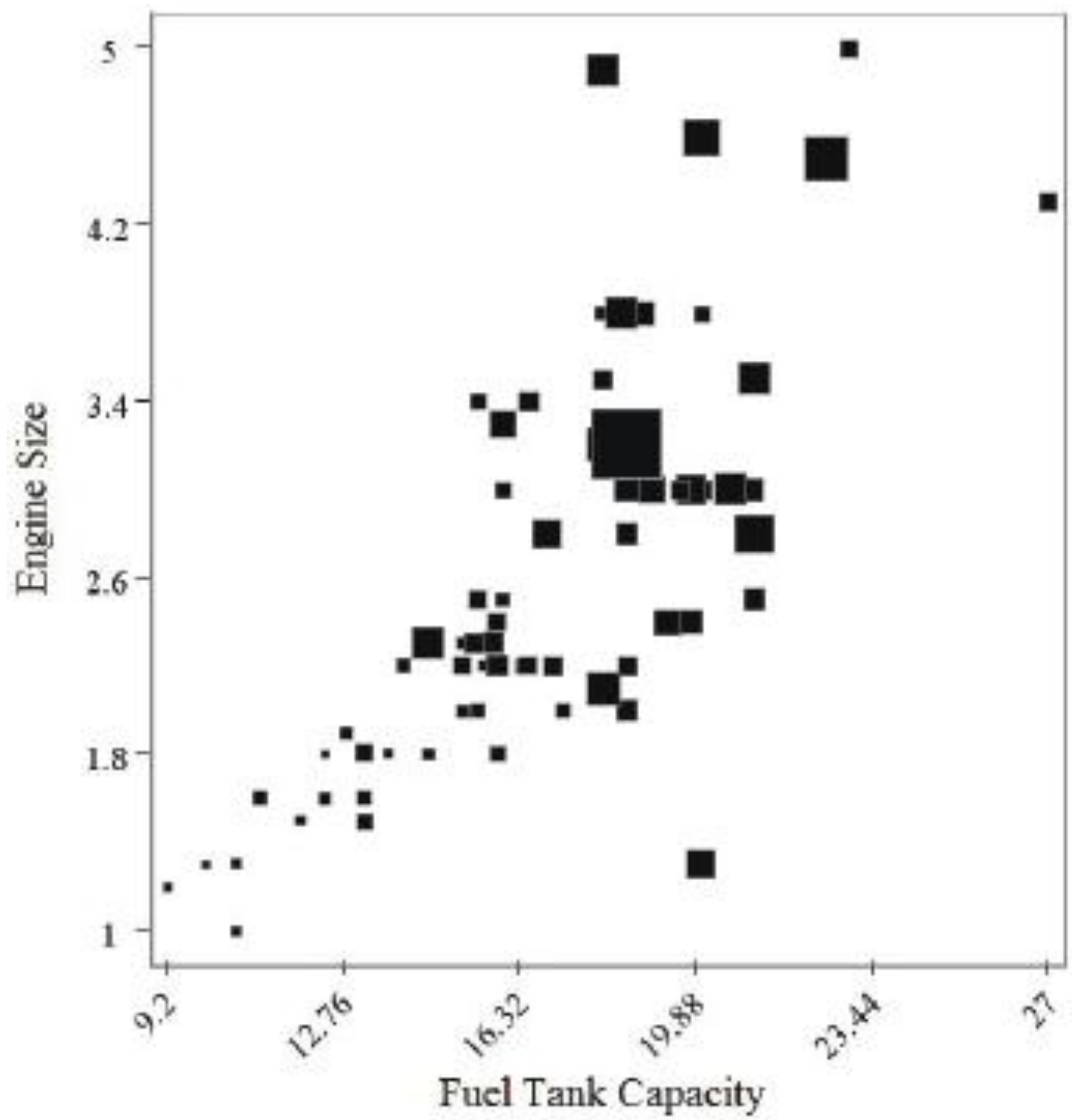


Illustration of Stevens' Law. The size ratio for each pair is 1:4. This magnitude is readily apparent in the lines, but it is easily underestimated in the squares and cubes.

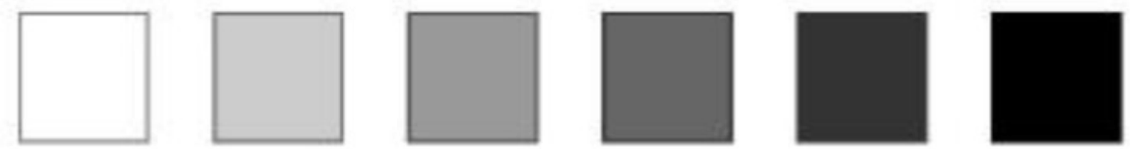
Eight visual variables: Size



This is a visualization of the 1993 car models data set, showing engine size versus fuel tank capacity. Size is mapped to maximum price charged.

Eight visual variables: Brightness (ou luminance)

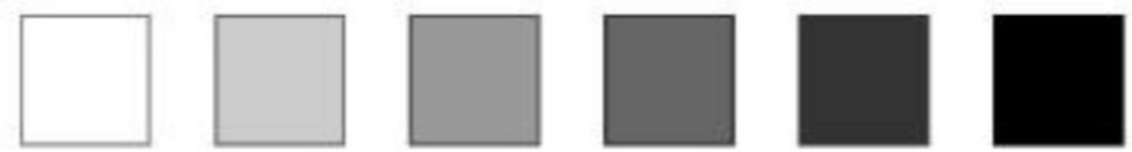
- **Brightness is the second visual variable used to modify marks to encode additional data variables.**



Brightness scale for mapping values to the display.

Eight visual variables: Brightness (ou luminance)

- **Brightness is the second visual variable used to modify marks to encode additional data variables.**



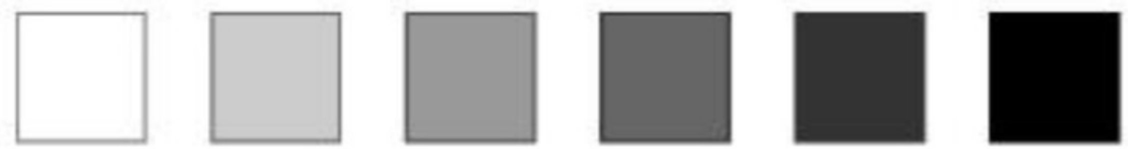
Brightness scale for mapping values to the display.

- **While it is possible to use the complete numerical range of brightness values, human perception cannot distinguish between all pairs of brightness values.**

FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Eight visual variables: Brightness (ou luminance)

- **Brightness is the second visual variable used to modify marks to encode additional data variables.**

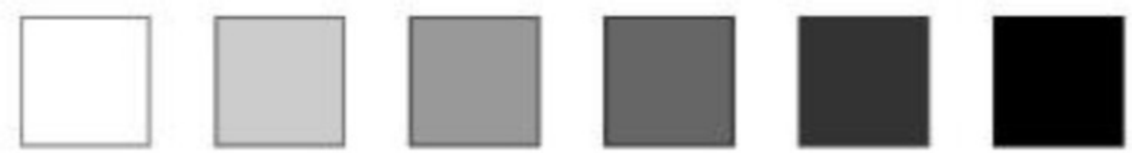


Brightness scale for mapping values to the display.

- **While it is possible to use the complete numerical range of brightness values, human perception cannot distinguish between all pairs of brightness values.**
- **Brightness can be used to provide relative difference for large interval and continuous data variables,**

Eight visual variables: Brightness (ou luminance)

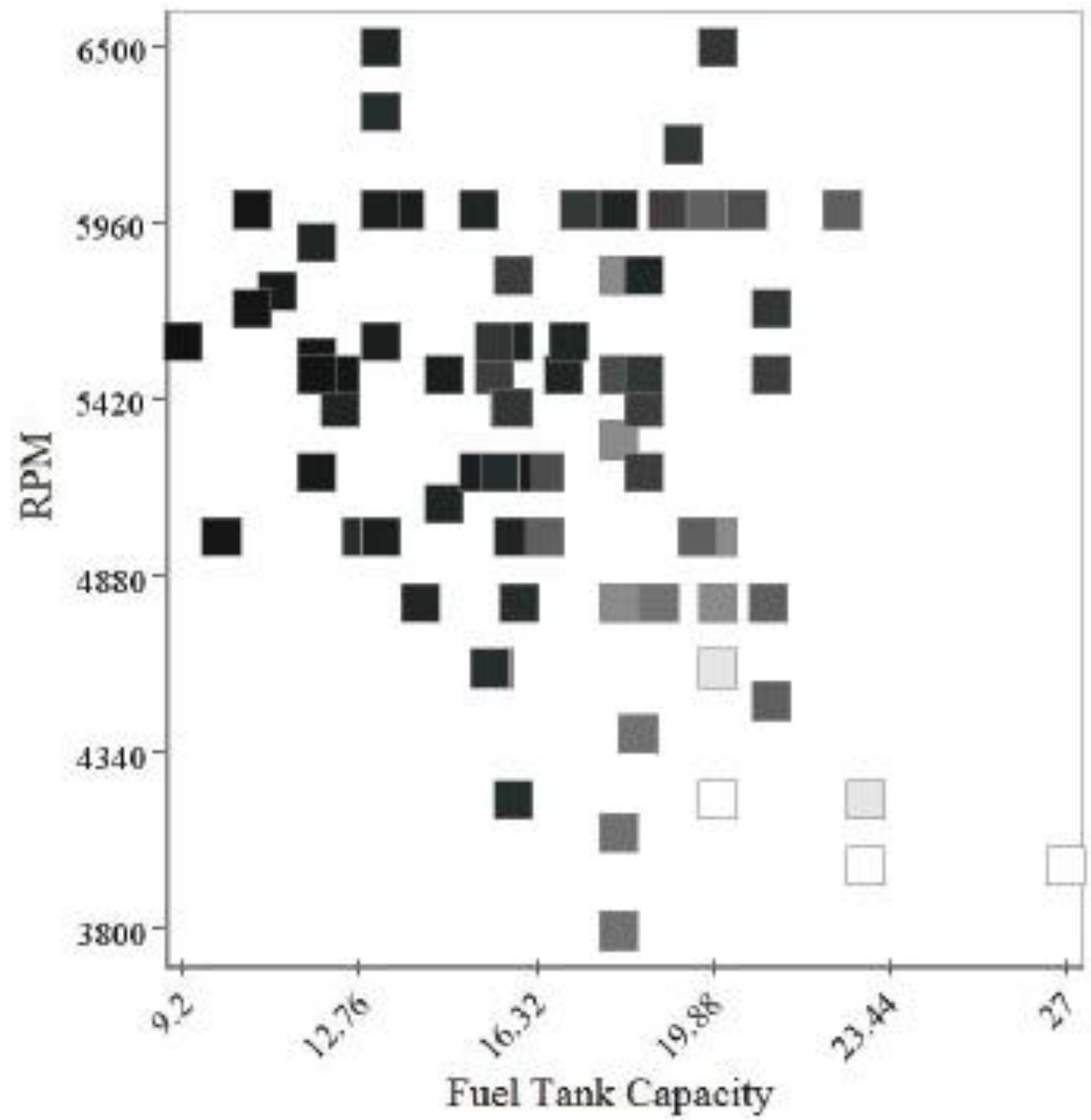
- **Brightness is the second visual variable used to modify marks to encode additional data variables.**



Brightness scale for mapping values to the display.

- **While it is possible to use the complete numerical range of brightness values, human perception cannot distinguish between all pairs of brightness values.**
- **Brightness can be used to provide relative difference for large interval and continuous data variables,**
- **or for mark distinction for marks drawn using a reduced sampled brightness scale.**

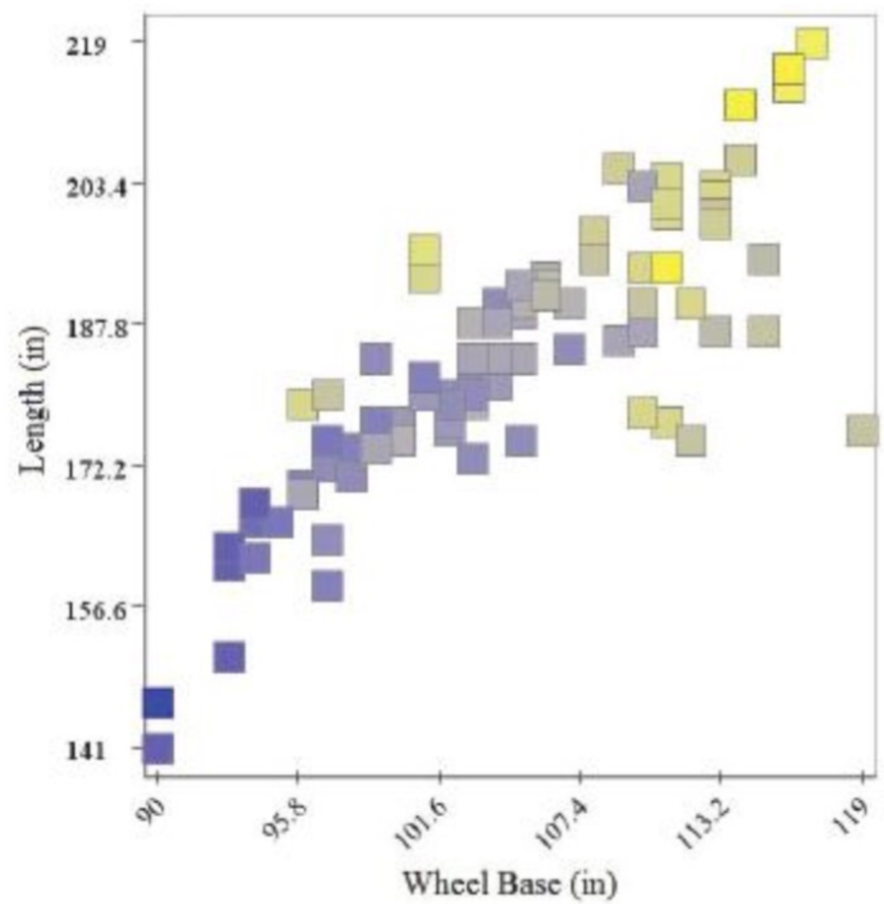
Eight visual variables: Brightness (ou luminance)



Another visualization of the 1993 car models data set, this time illustrating the use of brightness to convey car width (the darker the points, the wider the vehicle).

Eight visual variables: Color

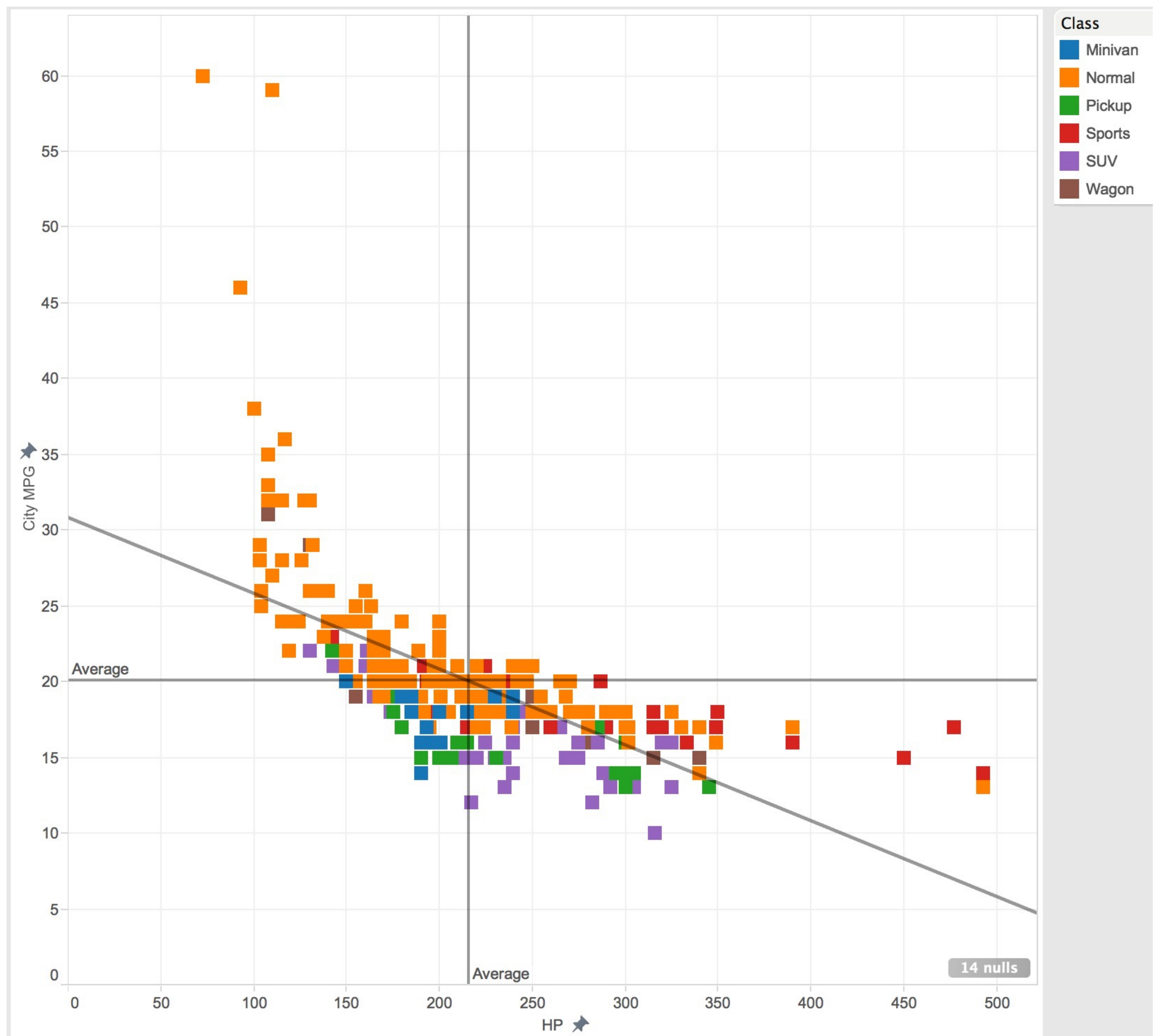
- **Color maps are useful for handling both interval and continuous data variables, since a color map is generally defined as a continuous range of hue and saturation values**



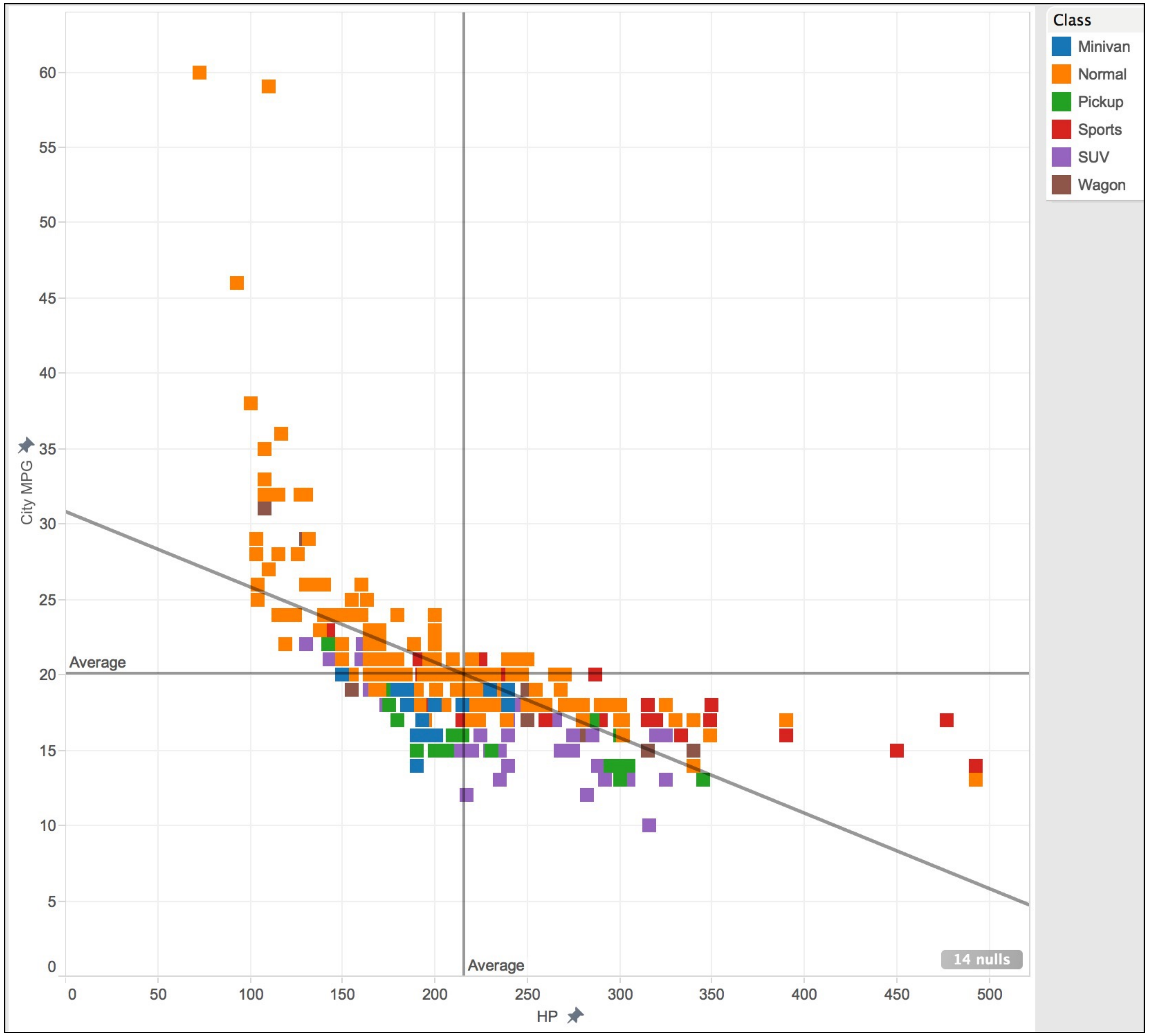
A visualization of the 1993 car models, showing the use of color to display the car's length. Here length is also associated with the y-axis and is plotted against wheelbase. In this figure, blue indicates a shorter length, while yellow indicates a longer length.

Eight visual variables: Color

- **When working with categorical or interval data with very low cardinality, it is generally acceptable to manually select colors for individual data values, which are selected to optimize the distinction between data types**

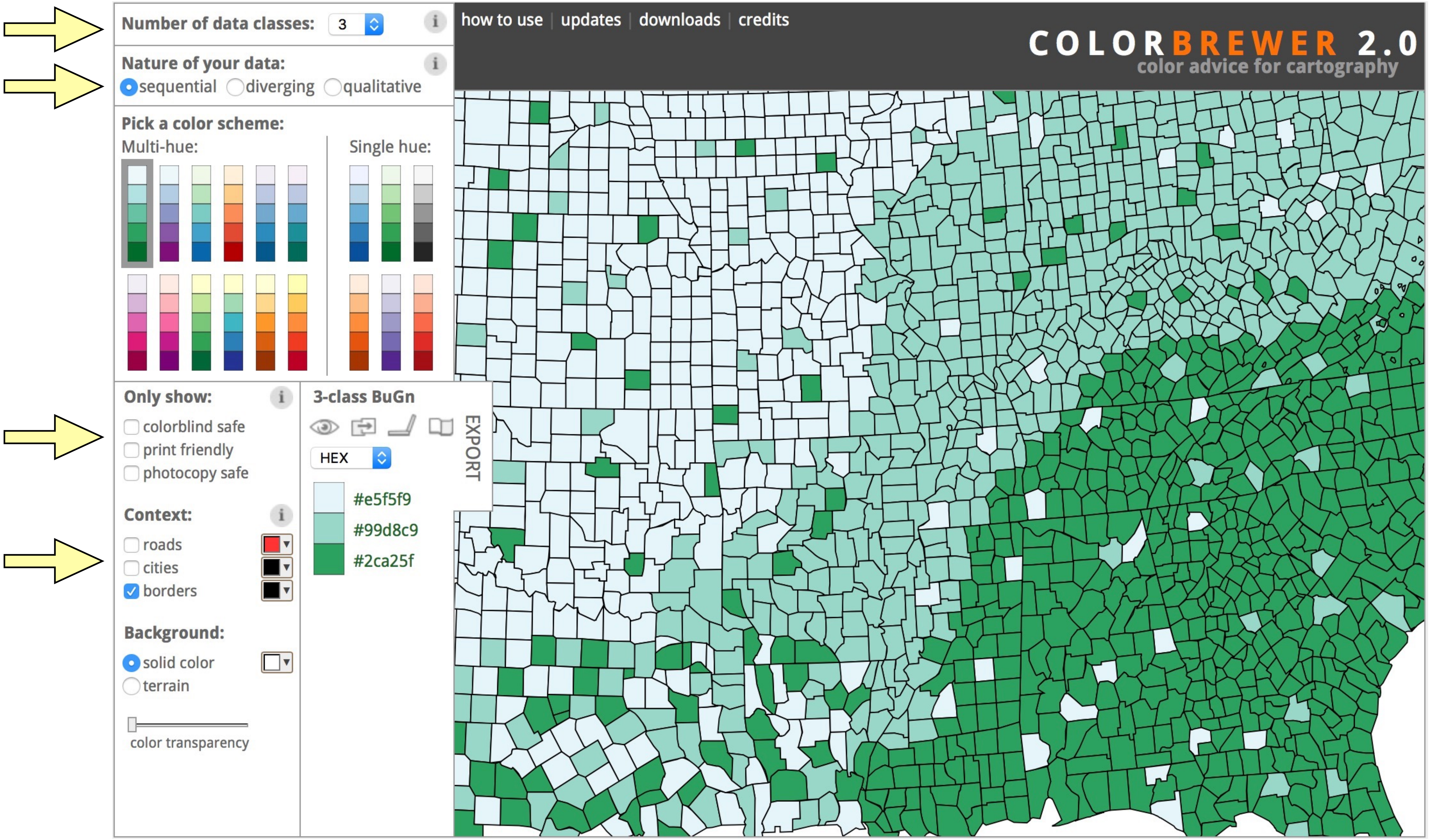


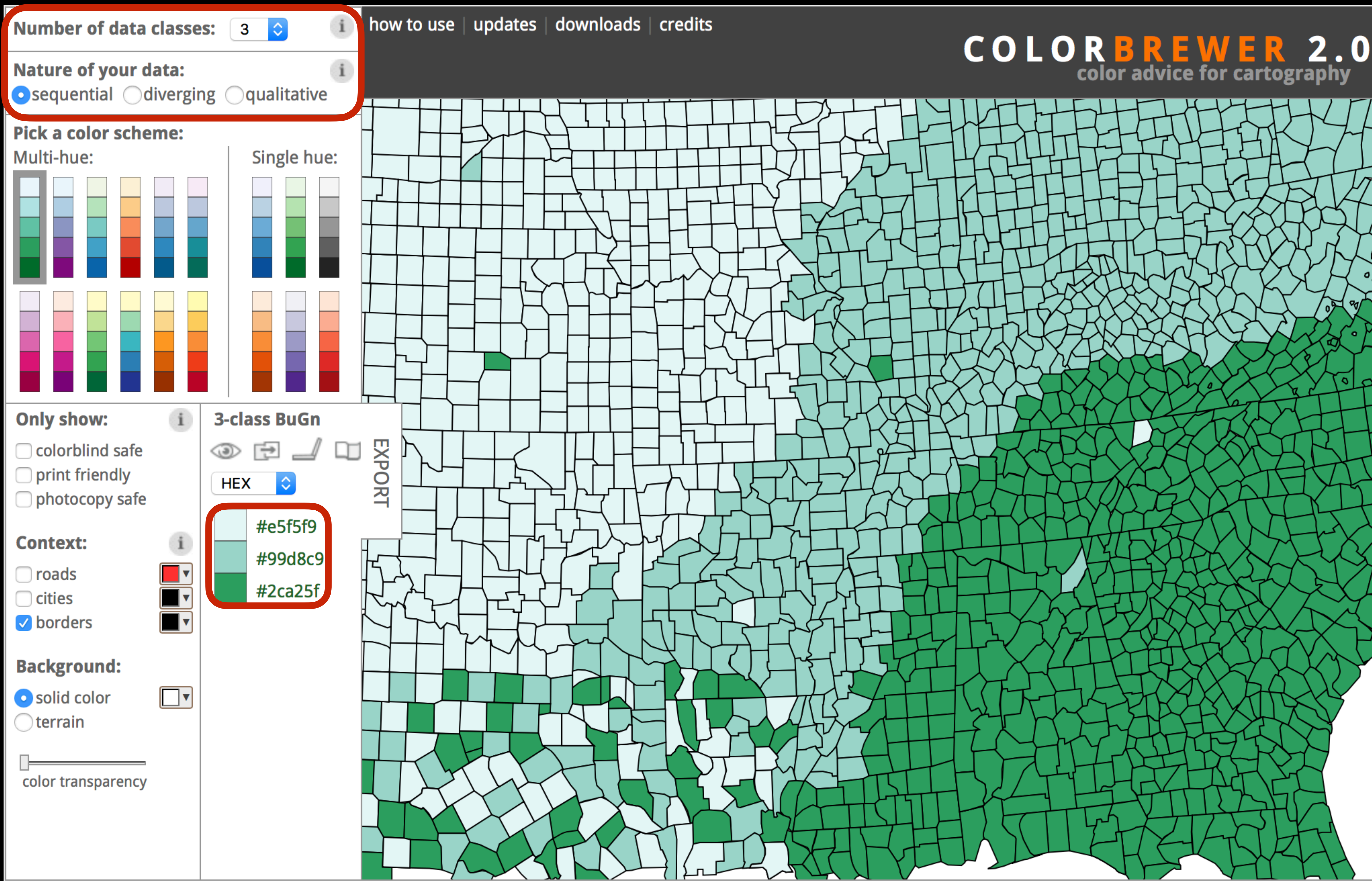
Eight visual variables: Color



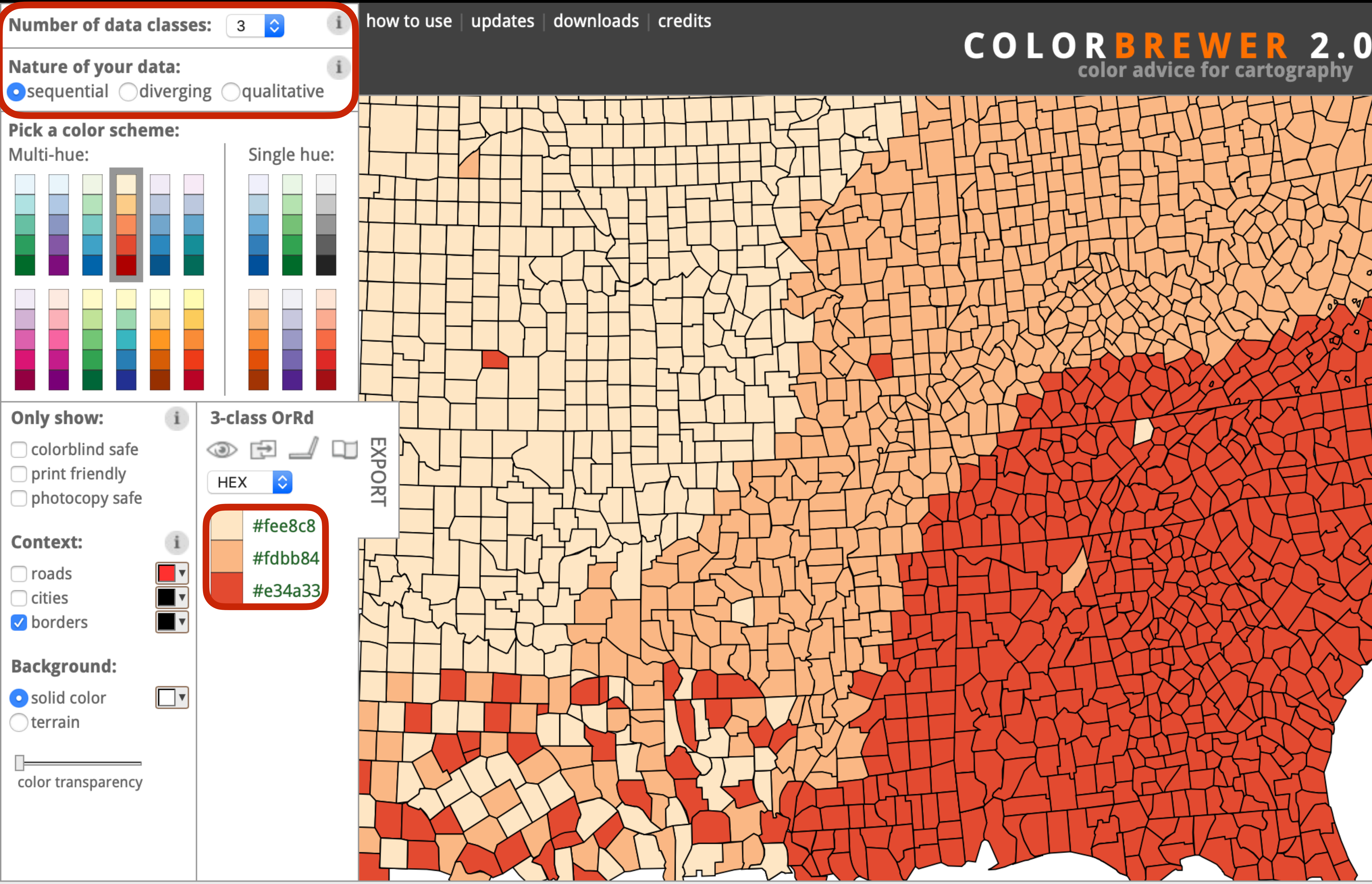
Eight visual variables: Color

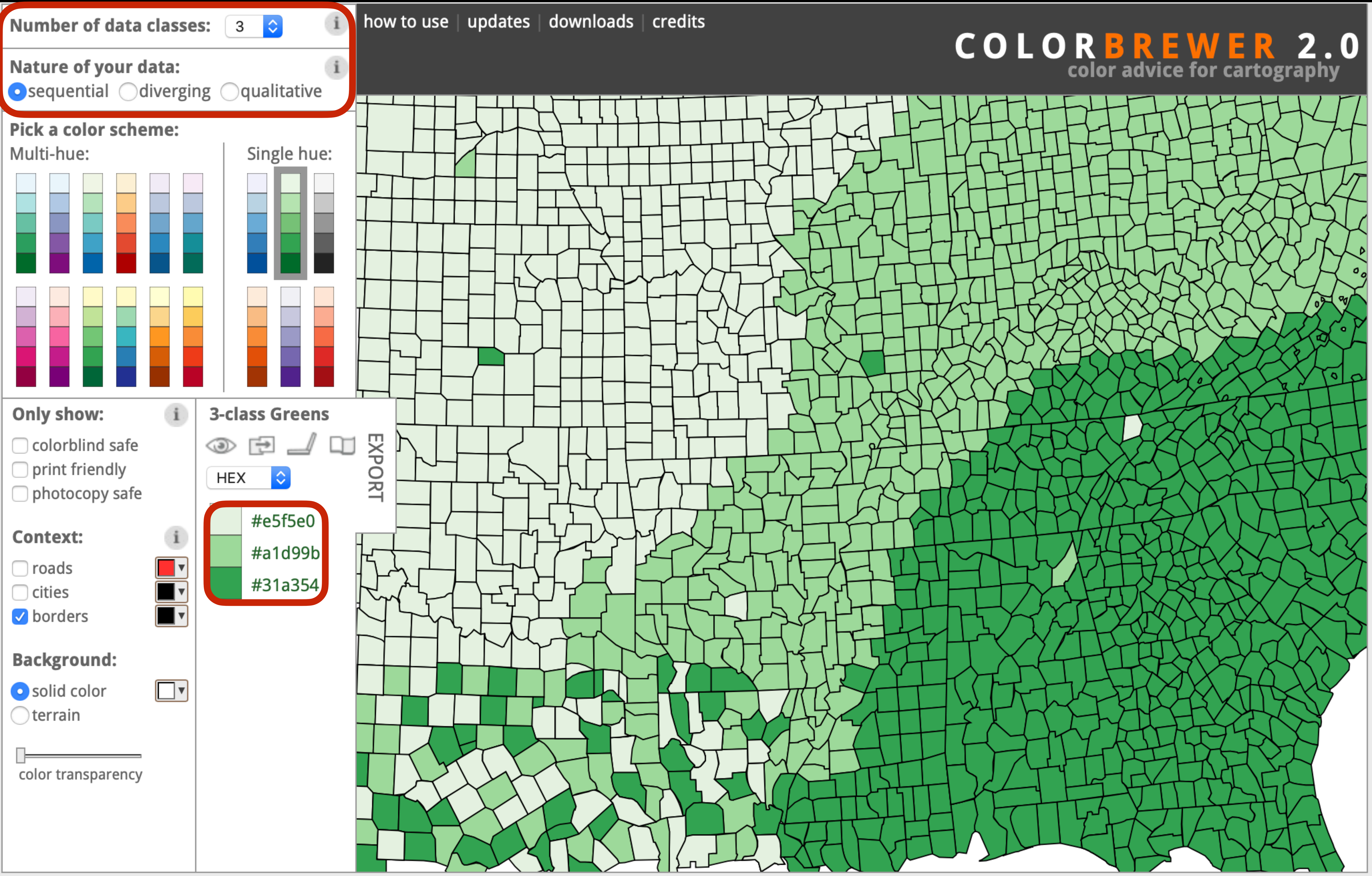
- **Check and try with: www.colorbrewer2.org**

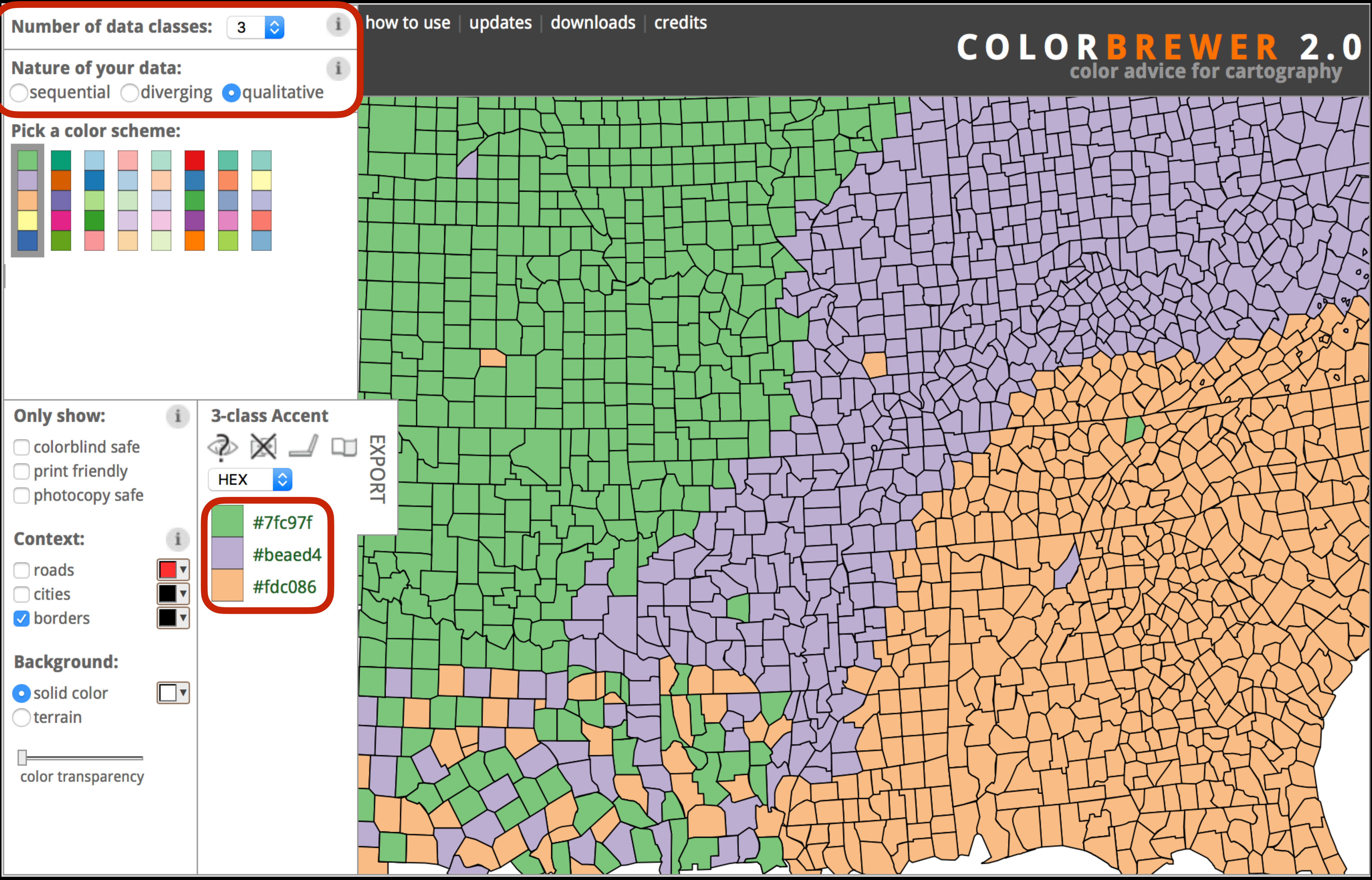


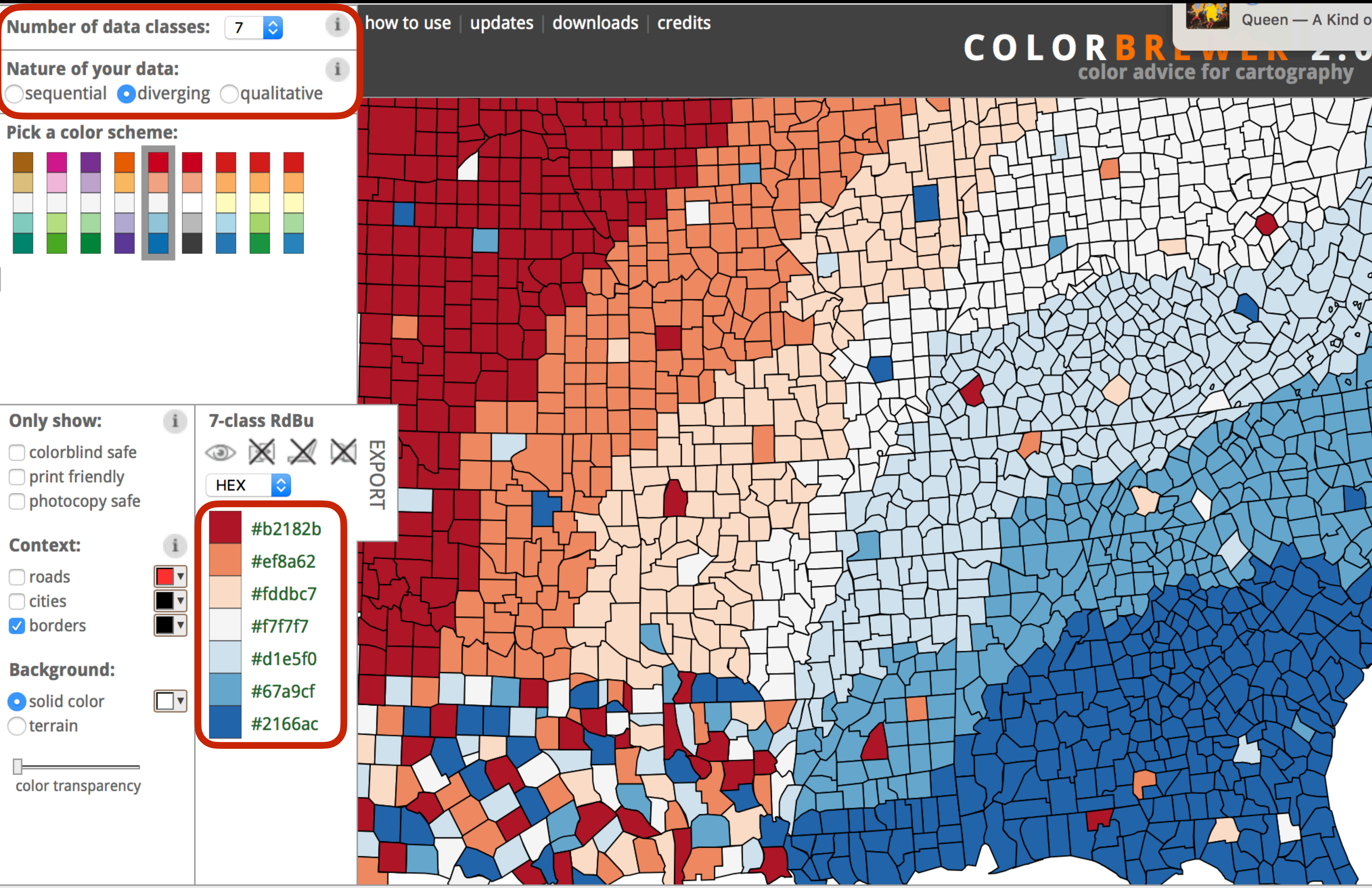


Number of data classes:
3
Nature of your data:
sequential
diverging
qualitative
how to use
updates
downloads
credits
COLORBREWER 2.0
color advice for cartography
Pick a color scheme:
Multi-hue:
Single hue:
Only show:
colorblind safe
print friendly
photocopy safe
3-class BuGn
HEX
EXPORT
#e5f5f9
#99d8c9
#2ca25f
Context:
roads
cities
borders
Background:
solid color
terrain
color transparency


Number of data classes:
3
Nature of your data:
sequential
diverging
qualitative
Pick a color scheme:
Multi-hue:
Single hue:
Only show:
colorblind safe
print friendly
photocopy safe
Context:
roads
cities
borders
Background:
solid color
terrain
color transparency
3-class OrRd
HEX
EXPORT
#fee8c8
#fdbb84
#e34a33
how to use
updates
downloads
credits
COLORBREWER 2.0
color advice for cartography


Number of data classes: 3
Nature of your data:
sequential
diverging
qualitative
Pick a color scheme:
Multi-hue:
Single hue:
Only show:
colorblind safe
print friendly
photocopy safe
Context:
roads
cities
borders
Background:
solid color
terrain
color transparency
3-class Greens
HEX
EXPORT
#e5f5e0
#a1d99b
#31a354
how to use
updates
downloads
credits
COLORBREWER 2.0
color advice for cartography


Number of data classes: 3
Nature of your data:
sequential
diverging
qualitative
Pick a color scheme:
how to use | updates | downloads | credits
COLORBREWER 2.0
color advice for cartography
Only show:
colorblind safe
print friendly
photocopy safe
Context:
roads
cities
borders
Background:
solid color
terrain
color transparency
3-class Accent
HEX
EXPORT
#7fc97f
#beaed4
#fdc086


Number of data classes: 7
Nature of your data:
sequential
diverging
qualitative
Pick a color scheme:
how to use | updates | downloads | credits
Queen — A Kind of
COLORBREWER 2.0
color advice for cartography
Only show:
colorblind safe
print friendly
photocopy safe
Context:
roads
cities
borders
Background:
solid color
terrain
color transparency
7-class RdBu
HEX
EXPORT
#b2182b
#ef8a62
#fddbc7
#f7f7f7
#d1e5f0
#67a9cf
#2166ac

Eight visual variables: Orientation

- **Orientation is a principal graphic component behind iconographic stick figure displays, and is tied directly to preattentive vision.**

Example orientations of a representation graphic, where the lowest value maps to the mark pointing upward and increasing values rotate the mark in a clockwise rotation.

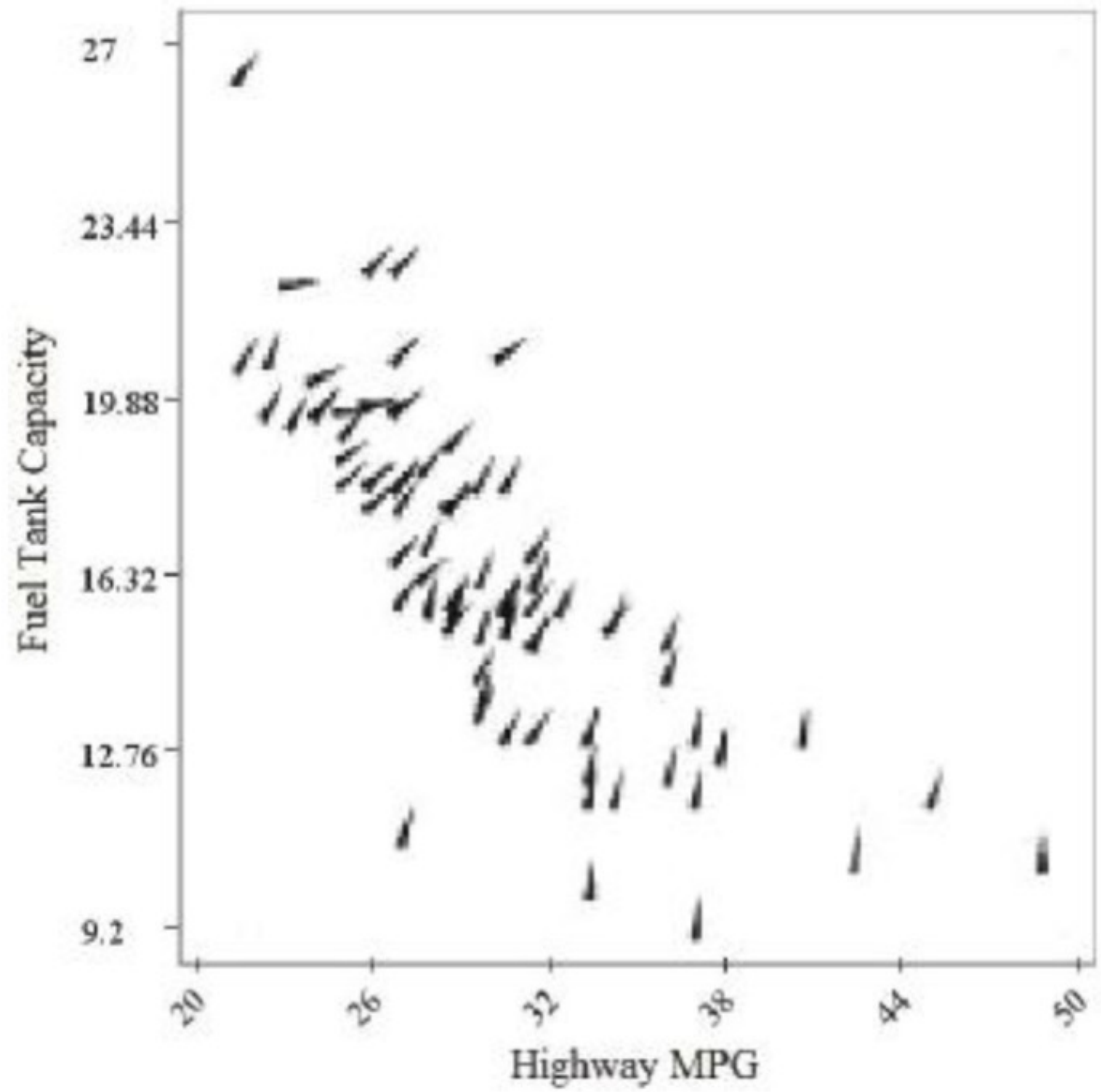
Eight visual variables: Orientation

- **Orientation is a principal graphic component behind iconographic stick figure displays, and is tied directly to preattentive vision.**

Example orientations of a representation graphic, where the lowest value maps to the mark pointing upward and increasing values rotate the mark in a clockwise rotation.

- **The best marks for using orientation are those with a natural single axis; the graphic exhibits symmetry about a major axis.**

Eight visual variables: Orientation

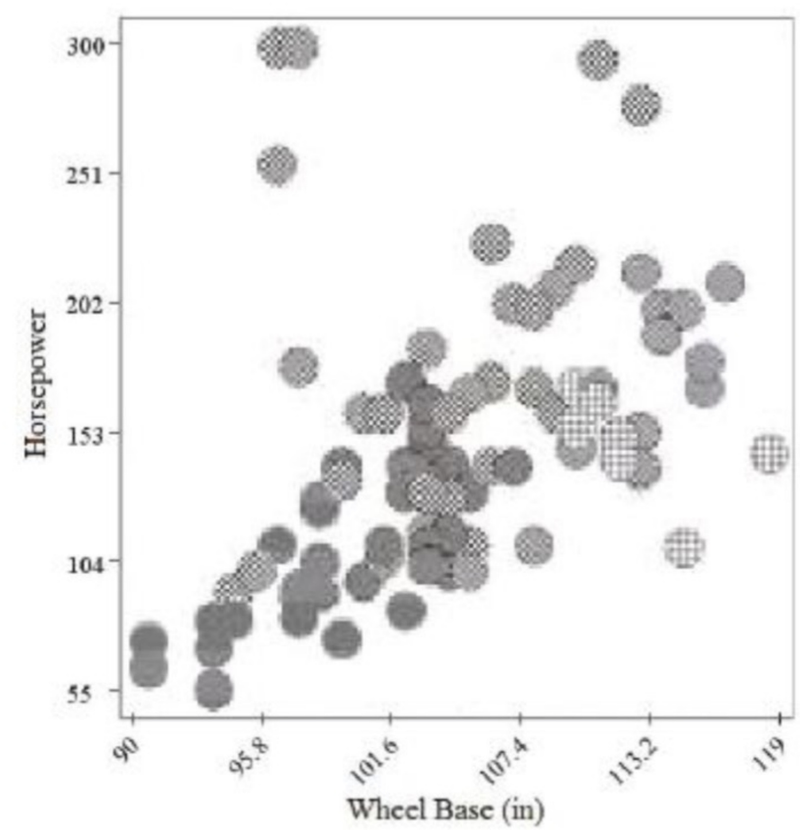


Sample visualization of the 1993 car models data set depicting using highway miles-per-gallon versus fuel tank capacity (position) with the additional data variable, midrange price, used to adjust mark orientation.

FACULDADE DE CIÊNCIAS E TECNOLOGIA UNIVERSIDADE NOVA DE LISBOA

Eight visual variables: Texture

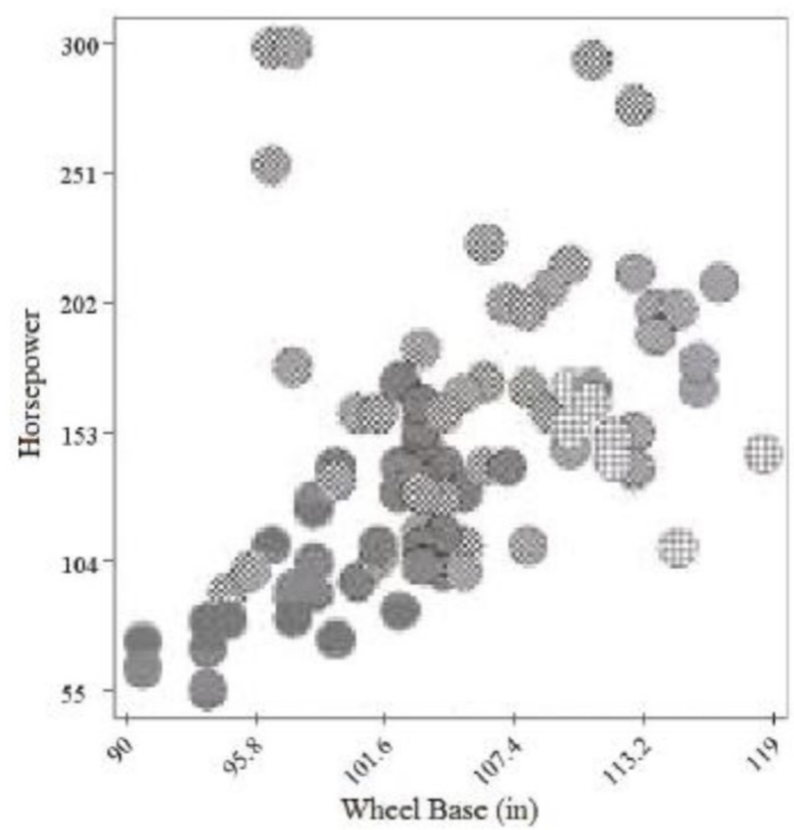
- **Texture can be considered as a combination of many of the other visual variables, including marks (texture elements), color (associated with each pixel in a texture region), and orientation (conveyed by changes in the local color).**



Example visualization using texture to provide additional information about the 1993 car models data set, showing the relationship between wheelbase versus horsepower (position) as related to car types, depicted by different textures.

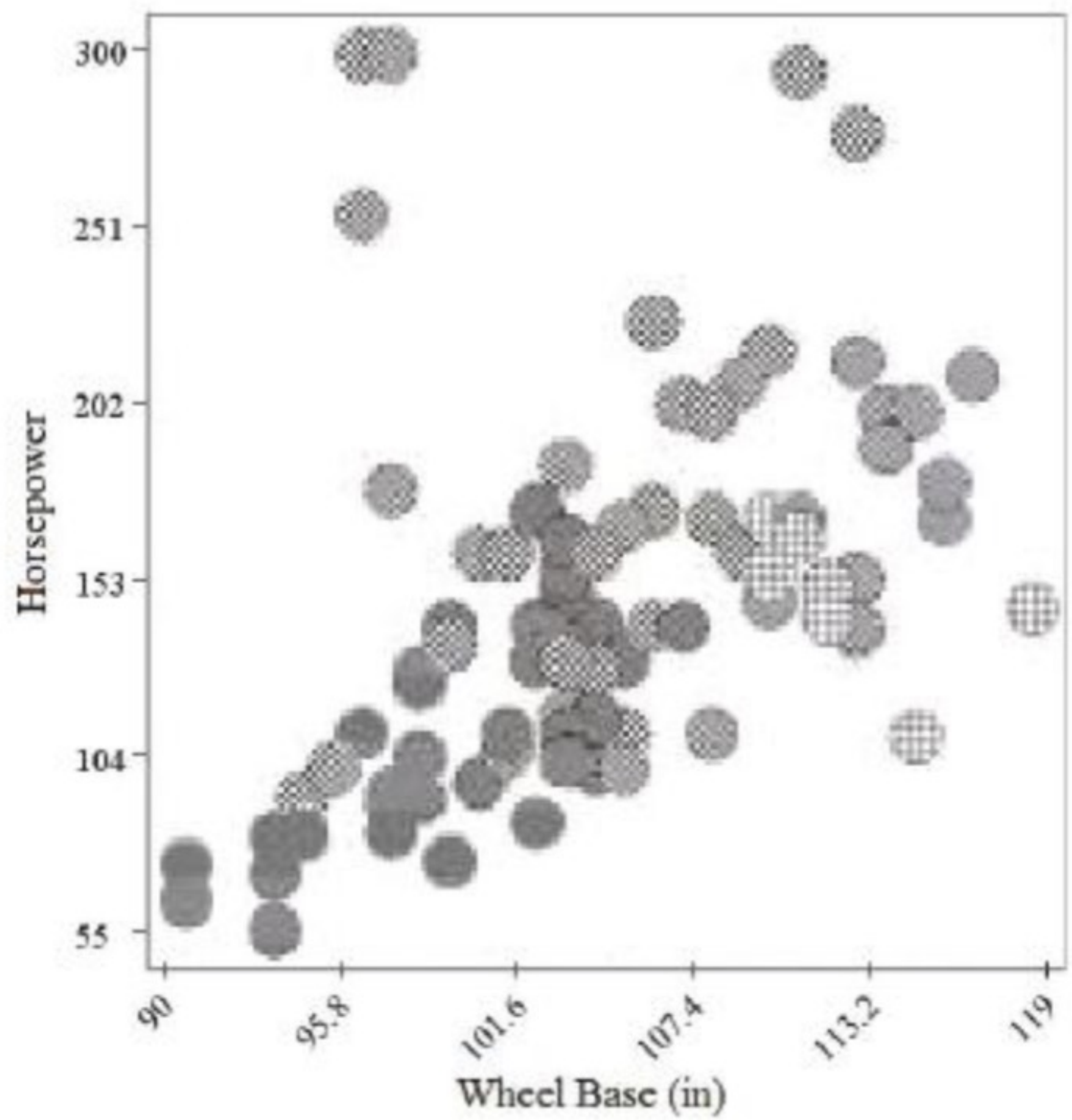
Eight visual variables: Texture

- **Texture can be considered as a combination of many of the other visual variables, including marks (texture elements), color (associated with each pixel in a texture region), and orientation (conveyed by changes in the local color).**
- **Texture is most commonly associated with a polygon, region, or surface.**



Example visualization using texture to provide additional information about the 1993 car models data set, showing the relationship between wheelbase versus horsepower (position) as related to car types, depicted by different textures.

Eight visual variables: Texture



Example visualization using texture to provide additional information about the 1993 car models data set, showing the relationship between wheelbase versus horsepower (position) as related to car types, depicted by different textures.

Eight visual variables: Motion

- **Motion can be associated with any of the other visual variables, since the way a variable changes over time can convey more information.**

Eight visual variables: Motion

- **Motion can be associated with any of the other visual variables, since the way a variable changes over time can convey more information.**

- **One common use of motion is in varying the speed at which a change is occurring (such as position change or flashing, which can be seen as changing the opacity).**

Eight visual variables: Motion

- **Motion can be associated with any of the other visual variables, since the way a variable changes over time can convey more information.**
- **One common use of motion is in varying the speed at which a change is occurring (such as position change or flashing, which can be seen as changing the opacity).**
- **The other aspect of motion is in the direction for position, this can be up, down, left, right, diagonal, or basically any slope, while for other variables it can be larger/ smaller, brighter/dimmer, steeper/shallower angles, and so on.**

Effects of Visual Variables

Effects of Visual Variables

- **Selective visual variables:**
  - **After coding with such variables, different data values are spontaneously divided by the human into distinguished groups (e.g., for visualizing nominal values).**

Effects of Visual Variables

- **Selective visual variables:**
  - **After coding with such variables, different data values are spontaneously divided by the human into distinguished groups (e.g., for visualizing nominal values).**
- **Associative visual variables:**
  - **All factors have same visibility (e.g., for visualizing nominal values).**

Effects of Visual Variables

- **Selective visual variables:**
  - **After coding with such variables, different data values are spontaneously divided by the human into distinguished groups (e.g., for visualizing nominal values).**
- **Associative visual variables:**
  - **All factors have same visibility (e.g., for visualizing nominal values).**
- **Ordinal visual variables:**
  - **After coding with such variables, different data values are spontaneously ordered by the human into distinguished groups (e.g., for visualizing ordinal and quantitative data).**

Effects of Visual Variables

- **Check the slides by Sheelagh Carpendale, University of Calgary**
  - **https://pages.cpsc.ucalgary.ca/~saul/hci_topics/pdf_files/visual-variables.pdf**
- **For each graphic attribute evaluates its use for each visual variable:**

Effects of Visual Variables

- **Check the slides by Sheelagh Carpendale, University of Calgary**
  - **https://pages.cpsc.ucalgary.ca/~saul/hci_topics/pdf_files/visual-variables.pdf**
- **For each graphic attribute evaluates its use for each visual variable:**
  - **selective (is a change enough to allow us to select it from a group?)**

Effects of Visual Variables

- **Check the slides by Sheelagh Carpendale, University of Calgary**
  - **https://pages.cpsc.ucalgary.ca/~saul/hci_topics/pdf_files/visual-variables.pdf**
- **For each graphic attribute evaluates its use for each visual variable:**
  - **selective (is a change enough to allow us to select it from a group?)**
  - **associative (is a change enough to allow us to perceive them as a group?)**

Effects of Visual Variables

- **Check the slides by Sheelagh Carpendale, University of Calgary**
  - **https://pages.cpsc.ucalgary.ca/~saul/hci_topics/pdf_files/visual-variables.pdf**
- **For each graphic attribute evaluates its use for each visual variable:**
  - **selective (is a change enough to allow us to select it from a group?)**
  - **associative (is a change enough to allow us to perceive them as a group?)**
  - **quantitative (is there a numerical reading obtainable from changes in this variable?)**

Effects of Visual Variables

- **Check the slides by Sheelagh Carpendale, University of Calgary**
  - **https://pages.cpsc.ucalgary.ca/~saul/hci_topics/pdf_files/visual-variables.pdf**
- **For each graphic attribute evaluates its use for each visual variable:**
  - **selective (is a change enough to allow us to select it from a group?)**
  - **associative (is a change enough to allow us to perceive them as a group?)**
  - **quantitative (is there a numerical reading obtainable from changes in this variable?)**
  - **order (are changes in this variable perceived as ordered?)**

Effects of Visual Variables

- **Check the slides by Sheelagh Carpendale, University of Calgary**
  - **https://pages.cpsc.ucalgary.ca/~saul/hci_topics/pdf_files/visual-variables.pdf**
- **For each graphic attribute evaluates its use for each visual variable:**
  - **selective (is a change enough to allow us to select it from a group?)**
  - **associative (is a change enough to allow us to perceive them as a group?)**
  - **quantitative (is there a numerical reading obtainable from changes in this variable?)**
  - **order (are changes in this variable perceived as ordered?)**
  - **length (across how many changes in this variable are distinctions perceptible?)**

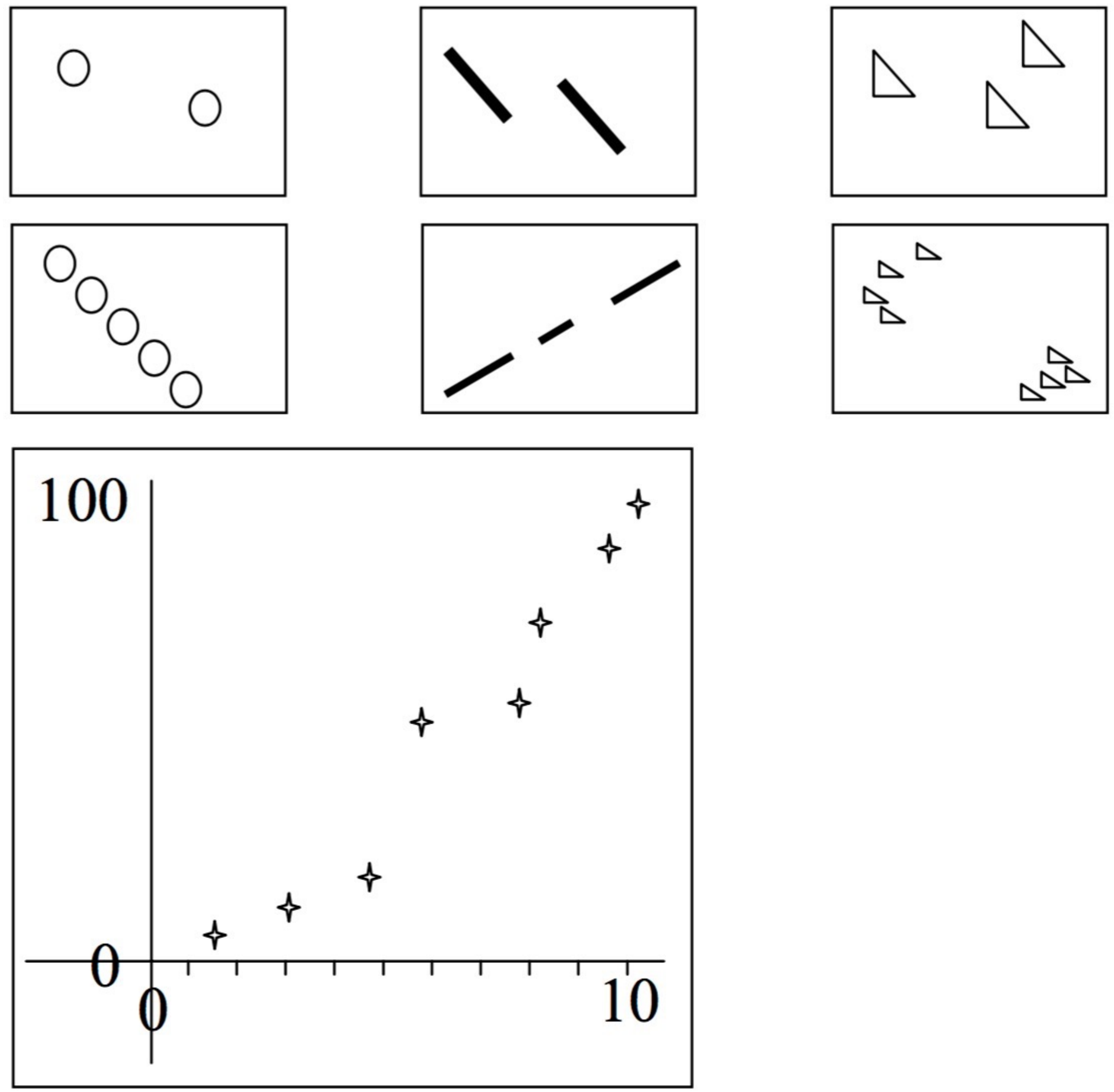
Effects of Visual Variables

- **Check the slides by Sheelagh Carpendale, University of Calgary**
  - **https://pages.cpsc.ucalgary.ca/~saul/hci_topics/pdf_files/visual-variables.pdf**
- **For each graphic attribute evaluates its use for each visual variable:**
  - **selective (is a change enough to allow us to select it from a group?)**
  - **associative (is a change enough to allow us to perceive them as a group?)**
  - **quantitative (is there a numerical reading obtainable from changes in this variable?)**
  - **order (are changes in this variable perceived as ordered?)**
  - **length (across how many changes in this variable are distinctions perceptible?)**

Effects of Visual Variables (by Sheelagh Carpendale)

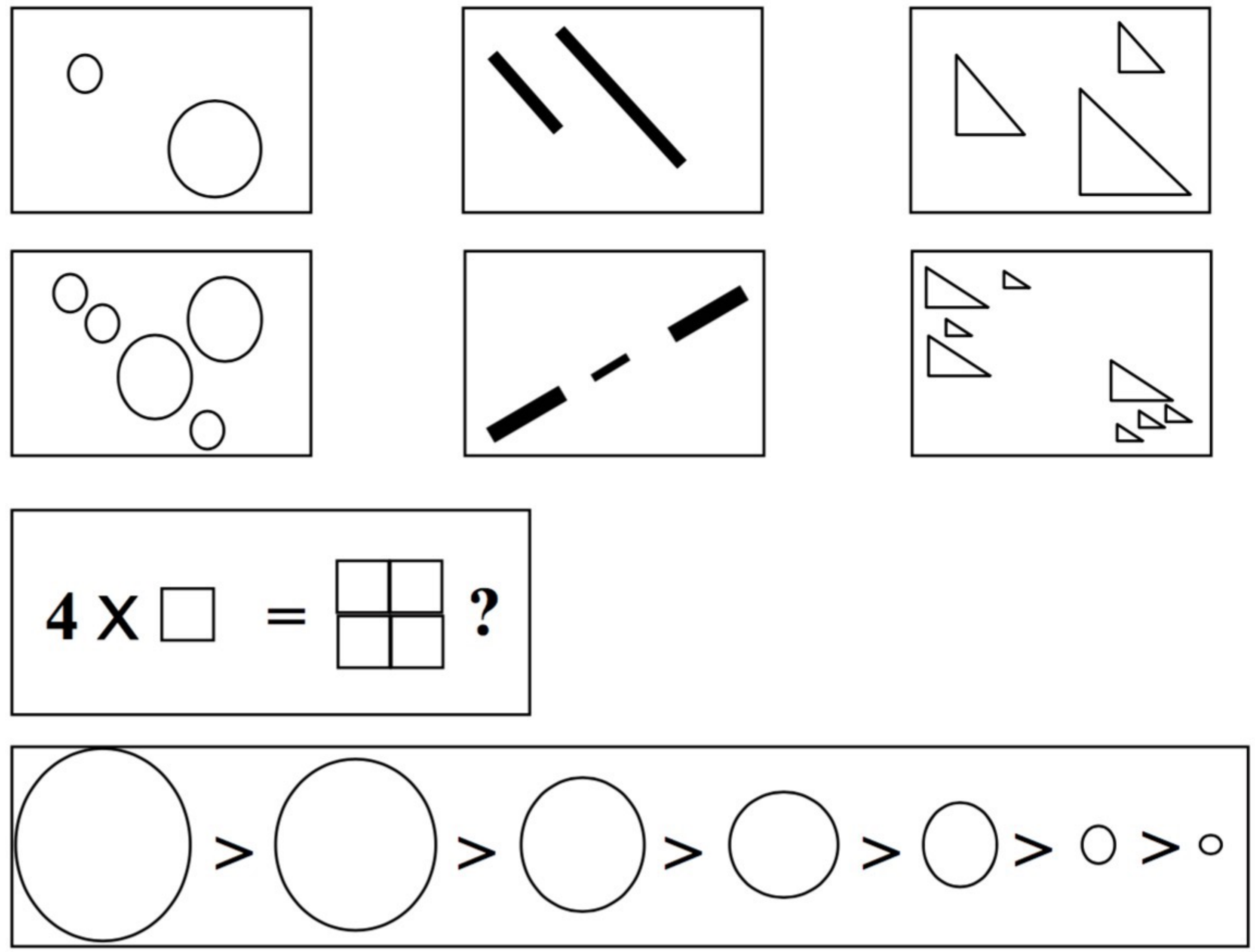
Position

- ✓ selective
- ✓ associative
- ✓ quantitative
- ✓ order
- ✓ length



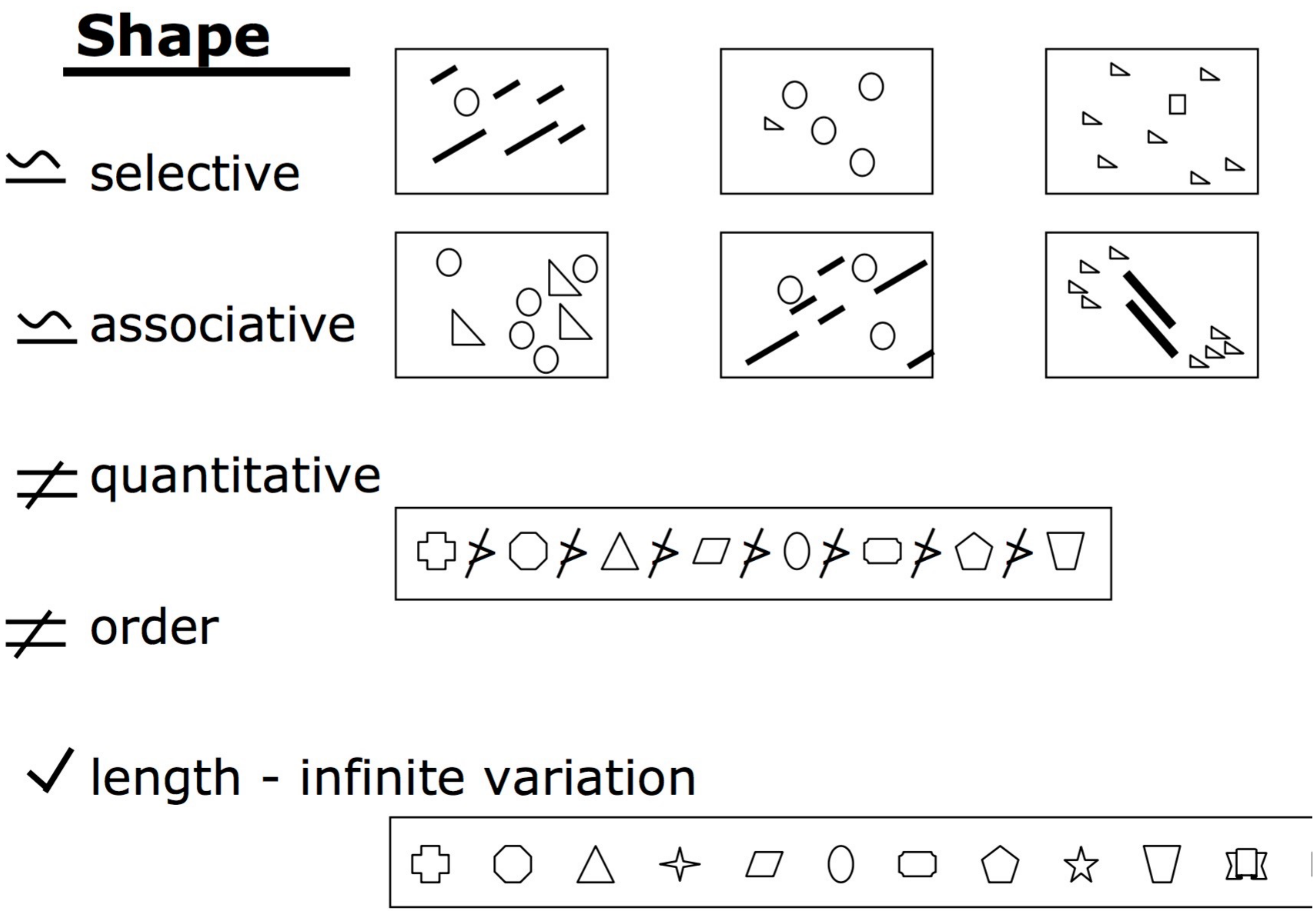
Effects of Visual Variables (by Sheelagh Carpendale)

Size



- ✓ selective
- ✓ associative
- ≏ quantitative
- ✓ order
- ✓ length
  - theoretically infinite but practically limited
  - association and selection ~ 5 and distinction ~ 20

Effects of Visual Variables (by Sheelagh Carpendale)

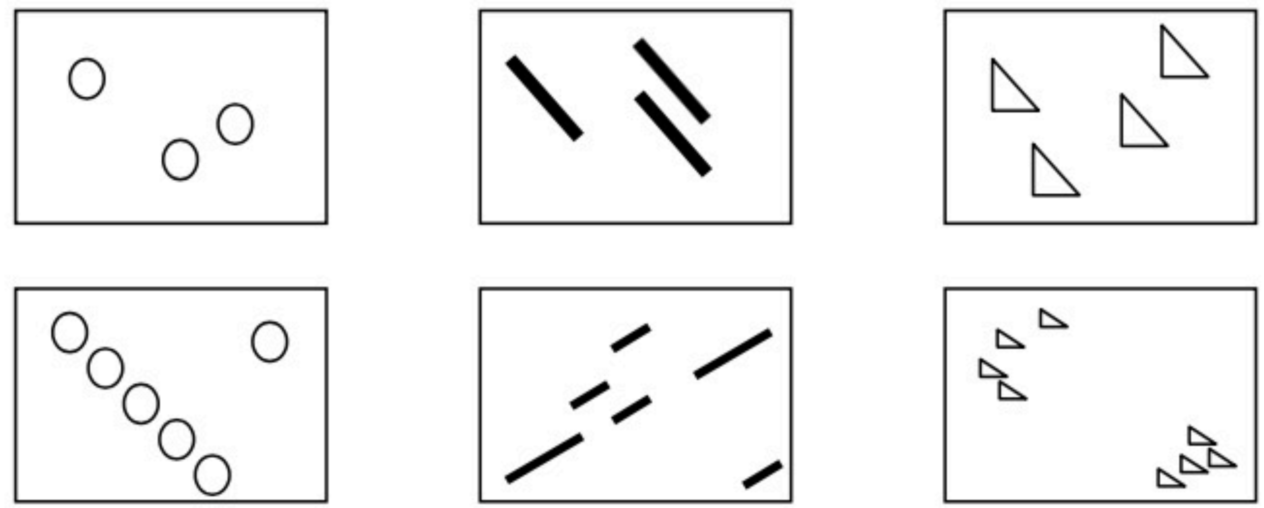


Effects of Visual Variables (by Sheelagh Carpendale)

Value

(Brightness)

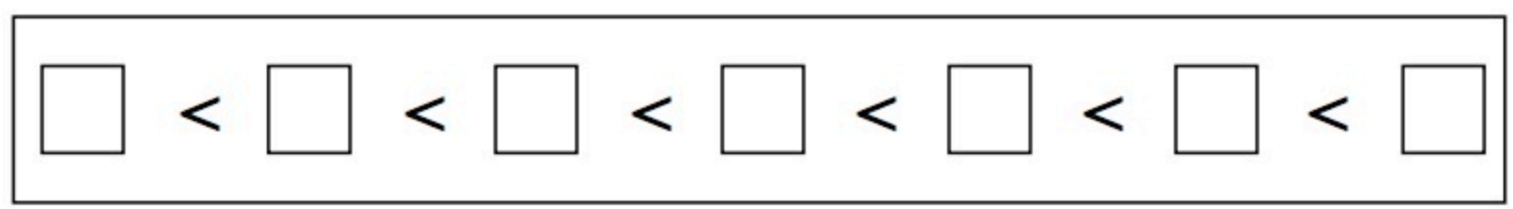
✓ selective



✓ associative

≠ quantitative

✓ order



✓ length

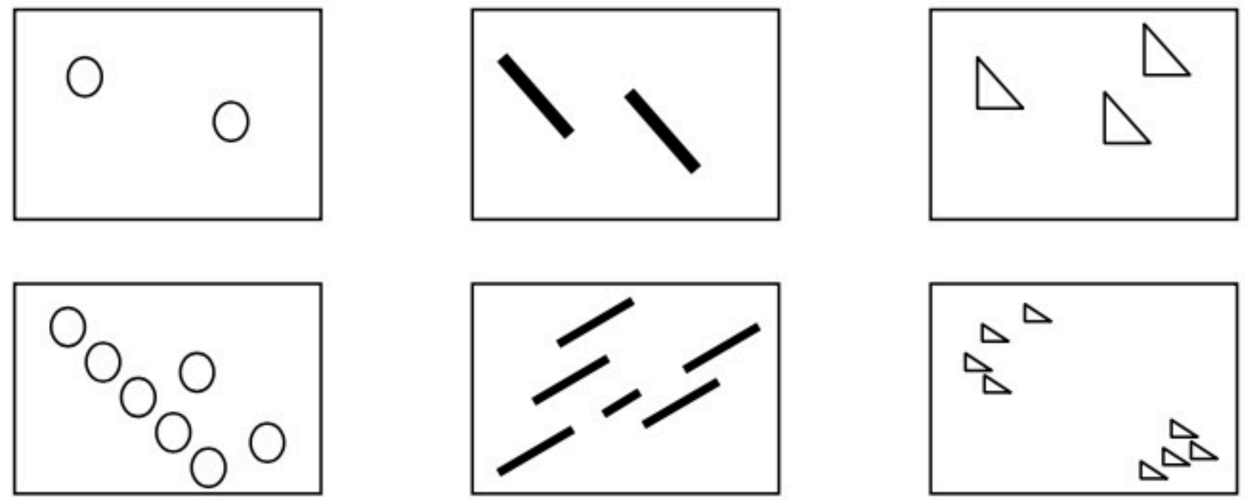
- theoretically infinite but practically limited
- association and selection ~ < 7 and distinction ~ 10

Effects of Visual Variables (by Sheelagh Carpendale)

Color

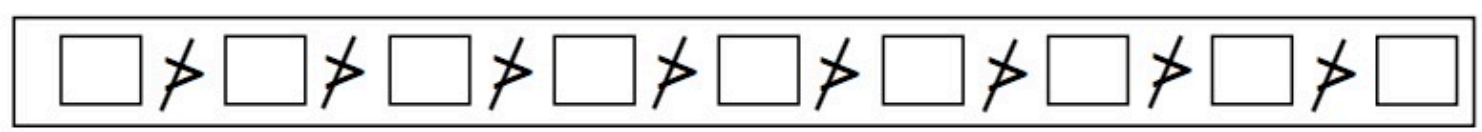
✓ selective

✓ associative



≠ quantitative

≠ order



✓ length

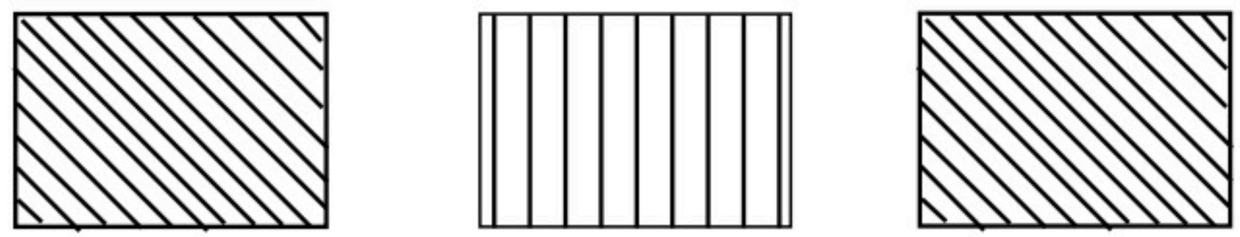
- theoretically infinite but practically limited
- association and selection ~ < 7 and distinction ~ 20

Effects of Visual Variables (by Sheelagh Carpendale)

Orientation

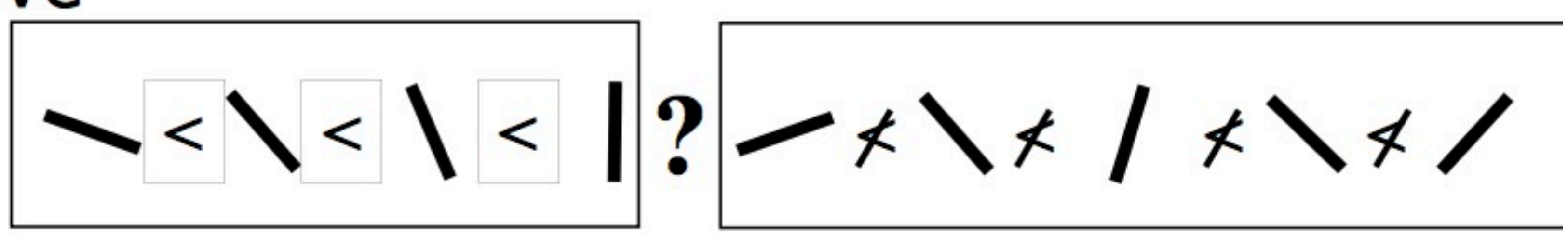
✓selective

✓ associative



≠quantitative

≠order



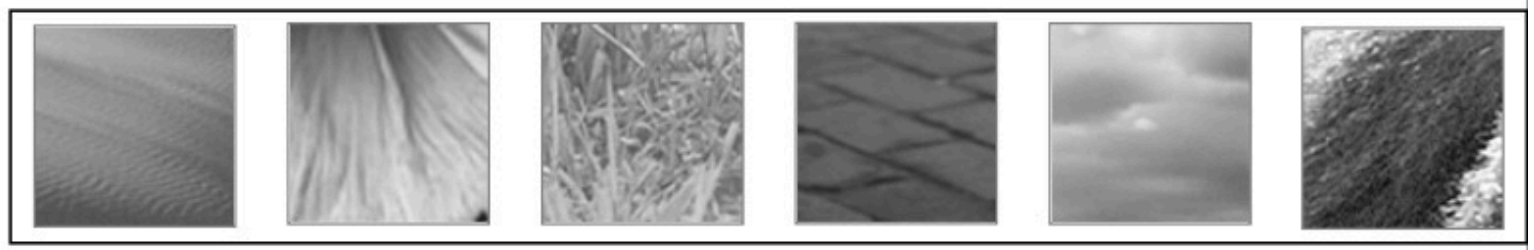
✓ length

- ~5 in 2D; ? in 3D

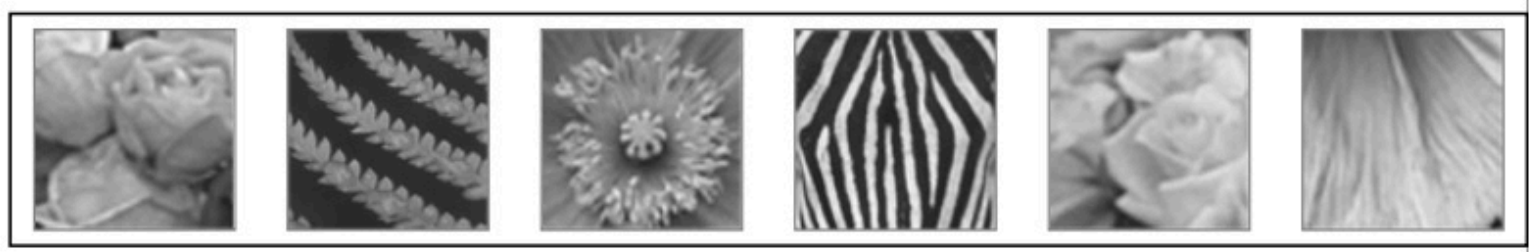
Effects of Visual Variables (by Sheelagh Carpendale)

Texture

✓ selective



✓ associative



≠ quantitative

≠ order

✓ length

- theoretically infinite

Effects of Visual Variables (by Sheelagh Carpendale)

Motion

✓ selective
- motion is one of our most powerful attention grabbers

✓ associative
- moving in unison groups objects effectively

≠ quantitative
- subjective perception

≠ order

? length
- distinguishable types of motion?

Effects of Visual Variables (by Sheelagh Carpendale)

- **Check the slides by Sheelagh Carpendale, University of Calgary**
  - **https://pages.cpsc.ucalgary.ca/~saul/hci_topics/pdf_files/visual-variables.pdf**
- **For each graphic attribute evaluates its use for each visual variable:**
  - **selective (is a change enough to allow us to select it from a group?)**
  - **associative (is a change enough to allow us to perceive them as a group?)**
  - **quantitative (is there a numerical reading obtainable from changes in this variable?)**
  - **order (are changes in this variable perceived as ordered?)**
  - **length (across how many changes in this variable are distinctions perceptible?)**

10 Min Interval

Marks and Channels by Tamara Munzner

Channel Rankings

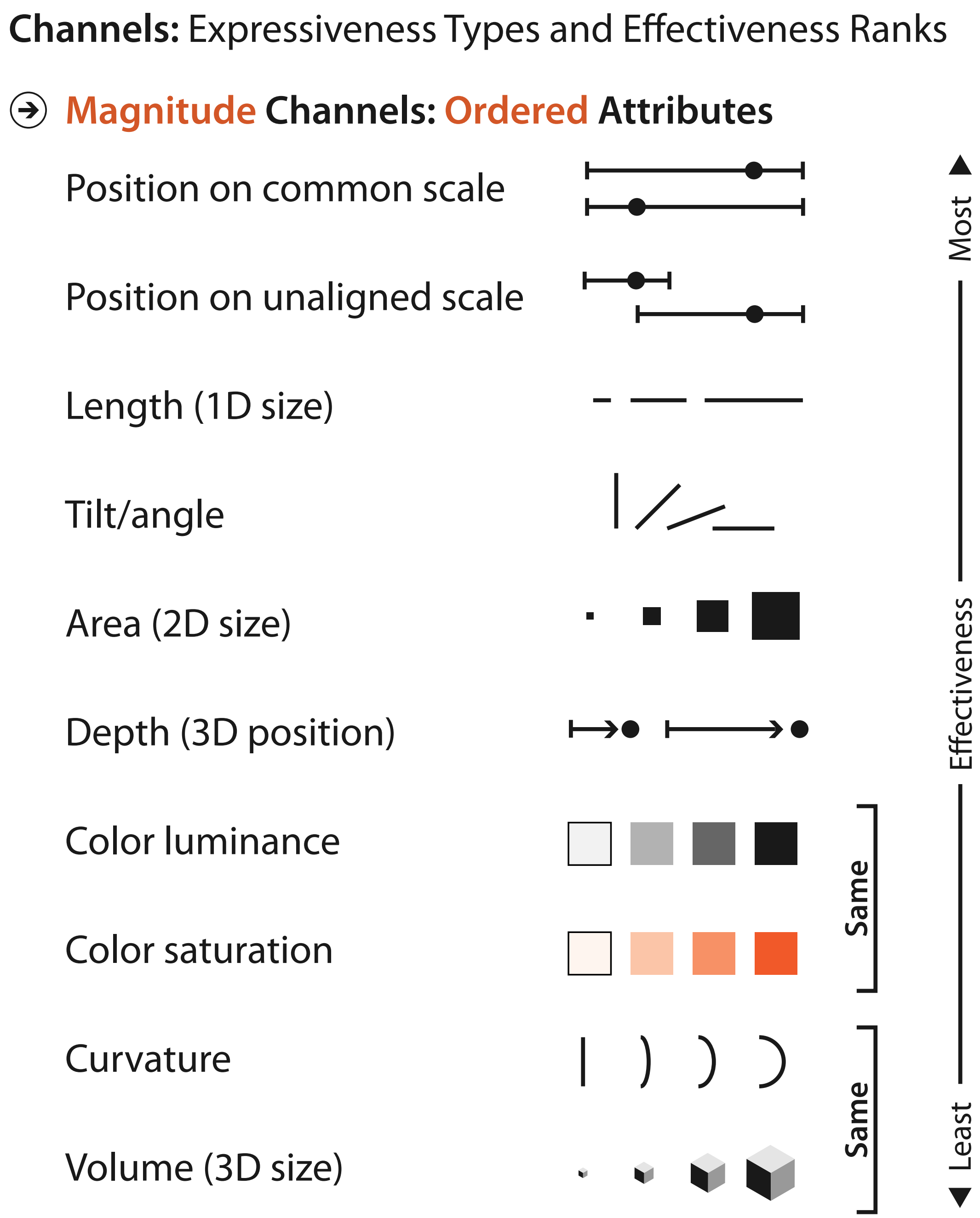


Figure 5.6. Channels ranked by effectiveness according to data and channel type. Ordered data should be shown with the magnitude channels, and categorical data with the identity channels.

Channel Rankings

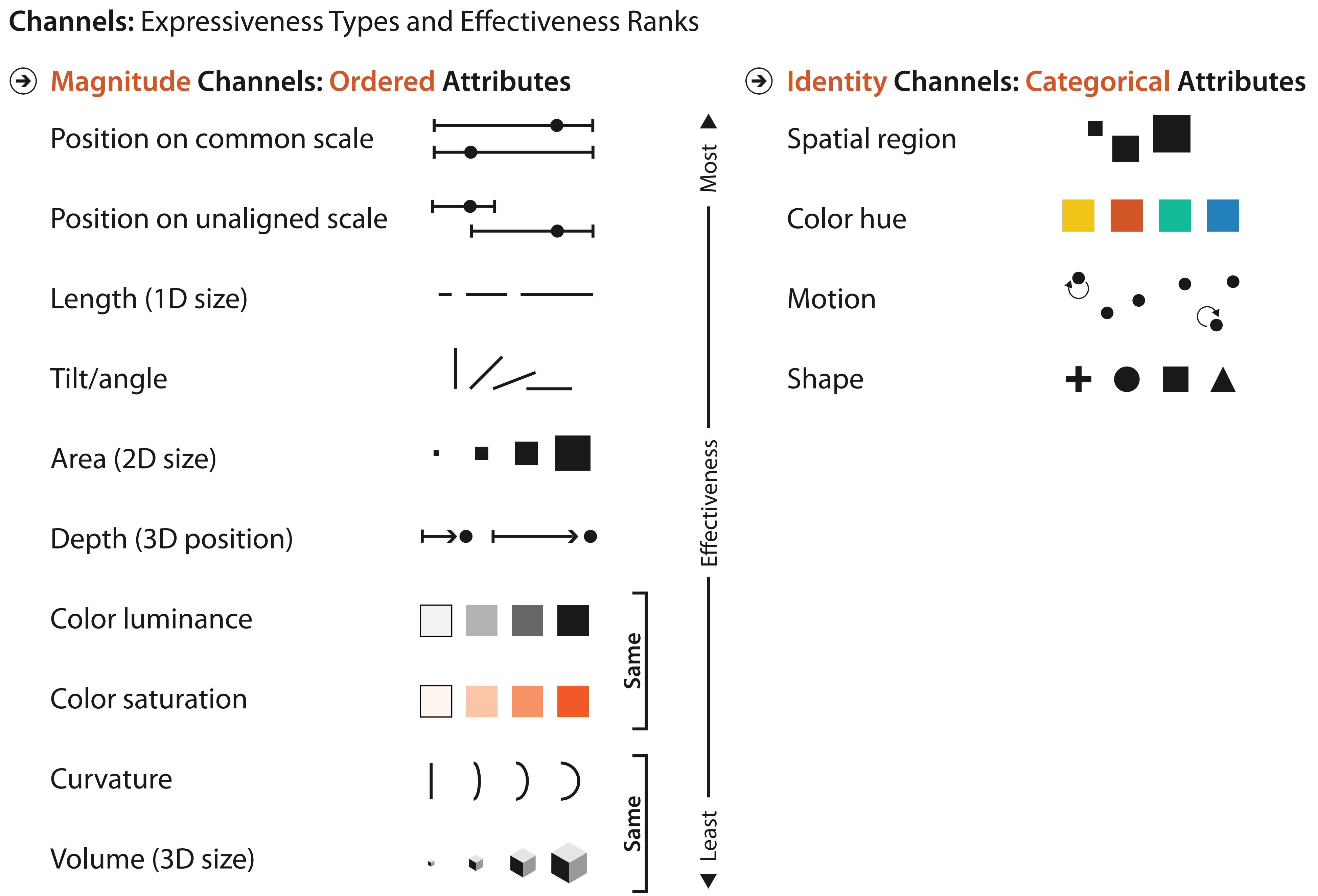


Figure 5.6. Channels ranked by effectiveness according to data and channel type. Ordered data should be shown with the magnitude channels, and categorical data with the identity channels.

Accuracy

Steven's Psychophysical Power Law: S= I^{N}

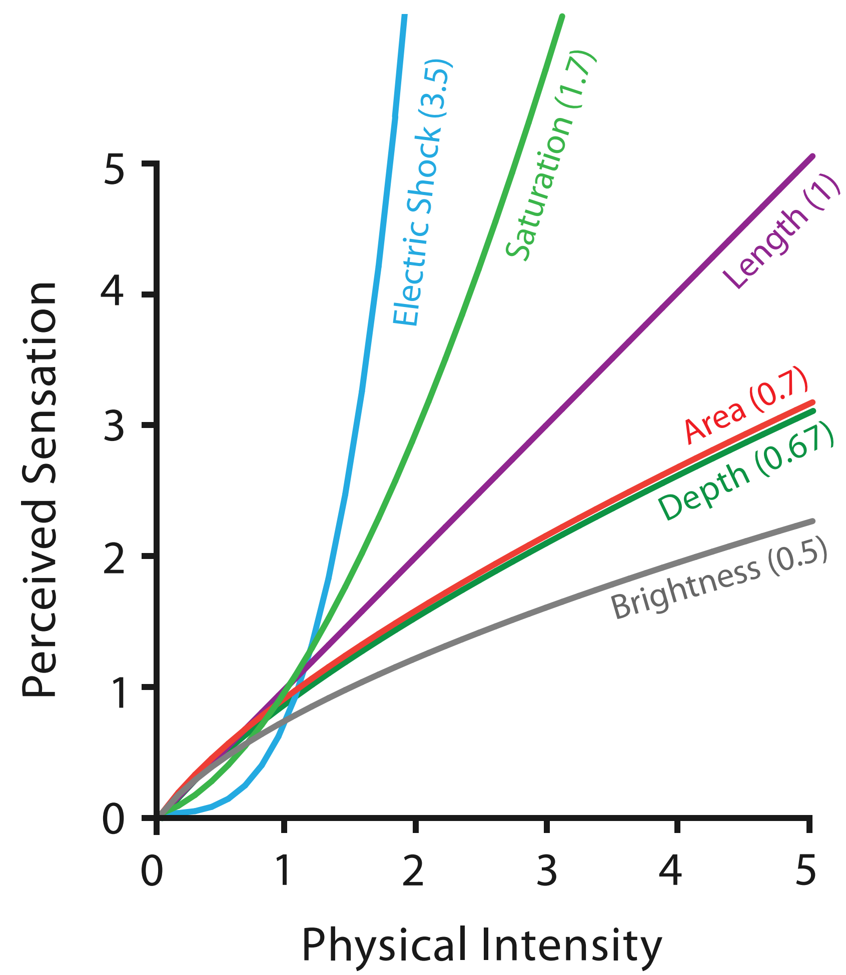
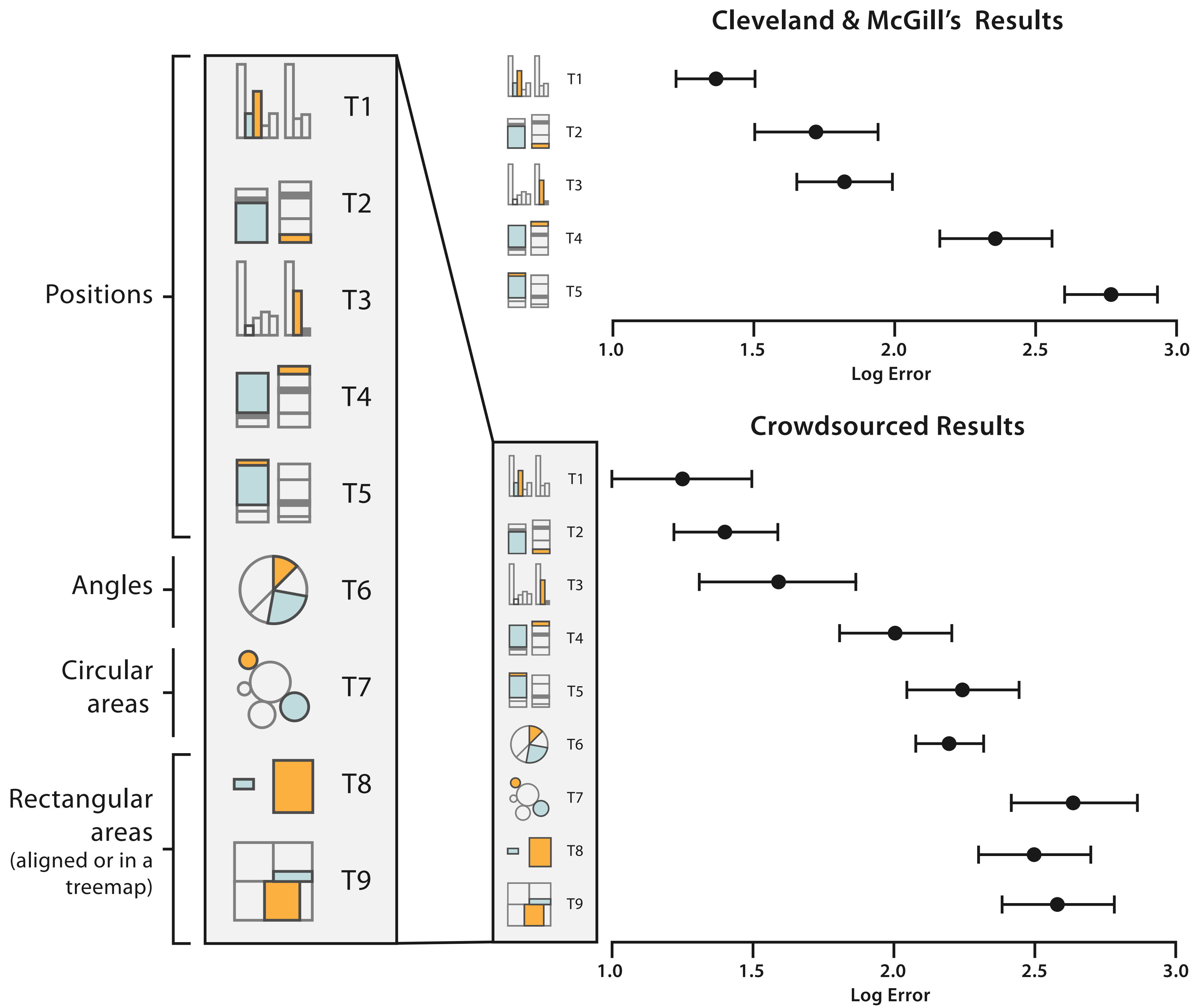
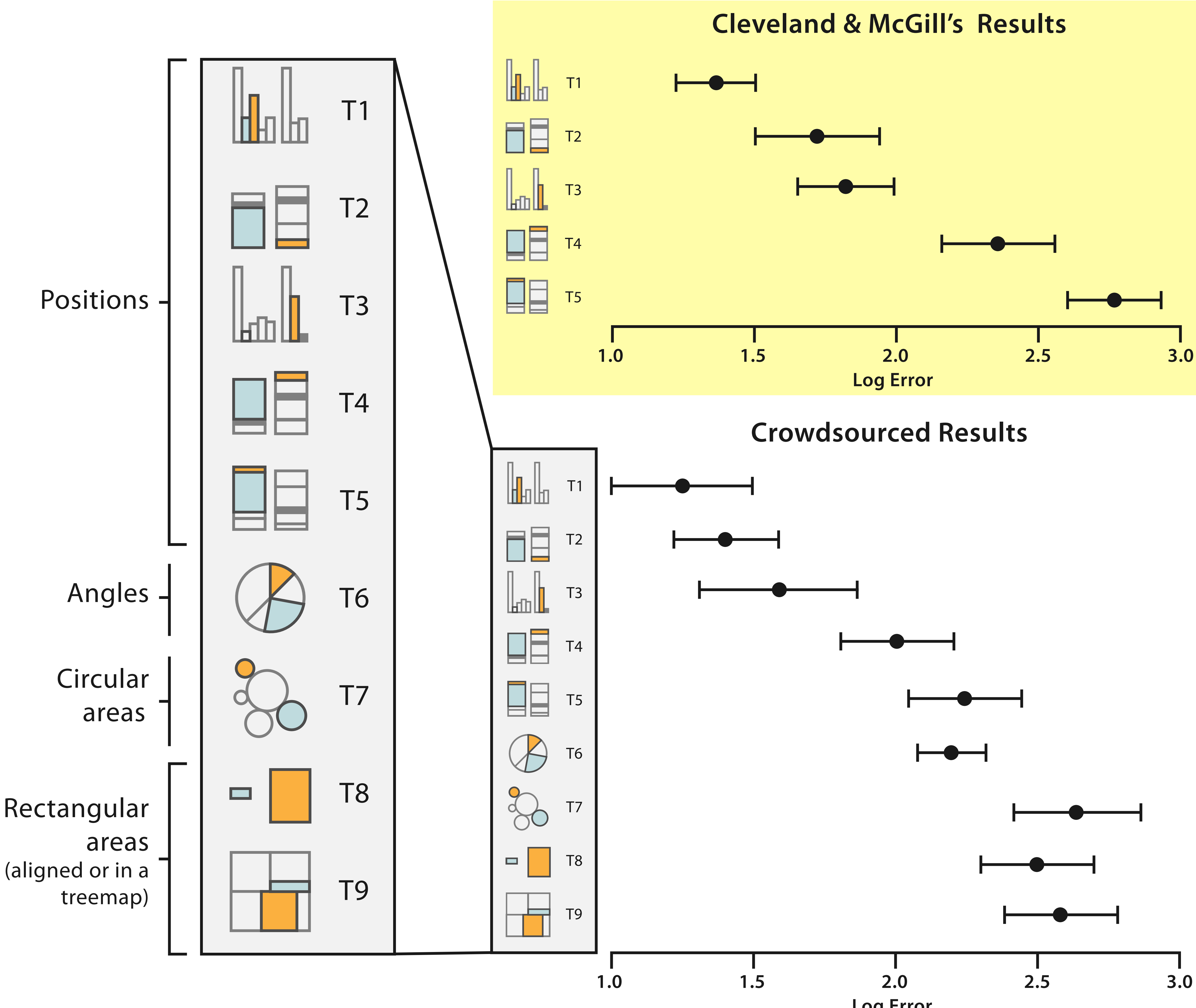


Figure 5.7. Stevens showed that the apparent magnitude of all sensory channels follows a power law $S = I^n$, where some sensations are perceptually magnified compared with their objective intensity (when $n > 1$) and some compressed (when $n < 1$). Length perception is completely accurate, whereas area is compressed and saturation is magnified. Data from Stevens [Stevens 75, p. 15].

Error rates (Cleveland and McGill [Cleveland and McGill 84a]. After [Heer and Bostock])





Cleveland & McGill's Results
T1
T2
T3
T4
T5
1.0
1.5
2.0
2.5
3.0
Log Error
Crowdsourced Results
T1
T2
T3
T4
T5
T6
T7
T8
T9
1.0
1.5
2.0
2.5
3.0
Log Error
Positions
Angles
Circular areas
Rectangular areas (aligned or in a treemap)
T1
T2
T3
T4
T5
T6
T7
T8
T9

Discriminability

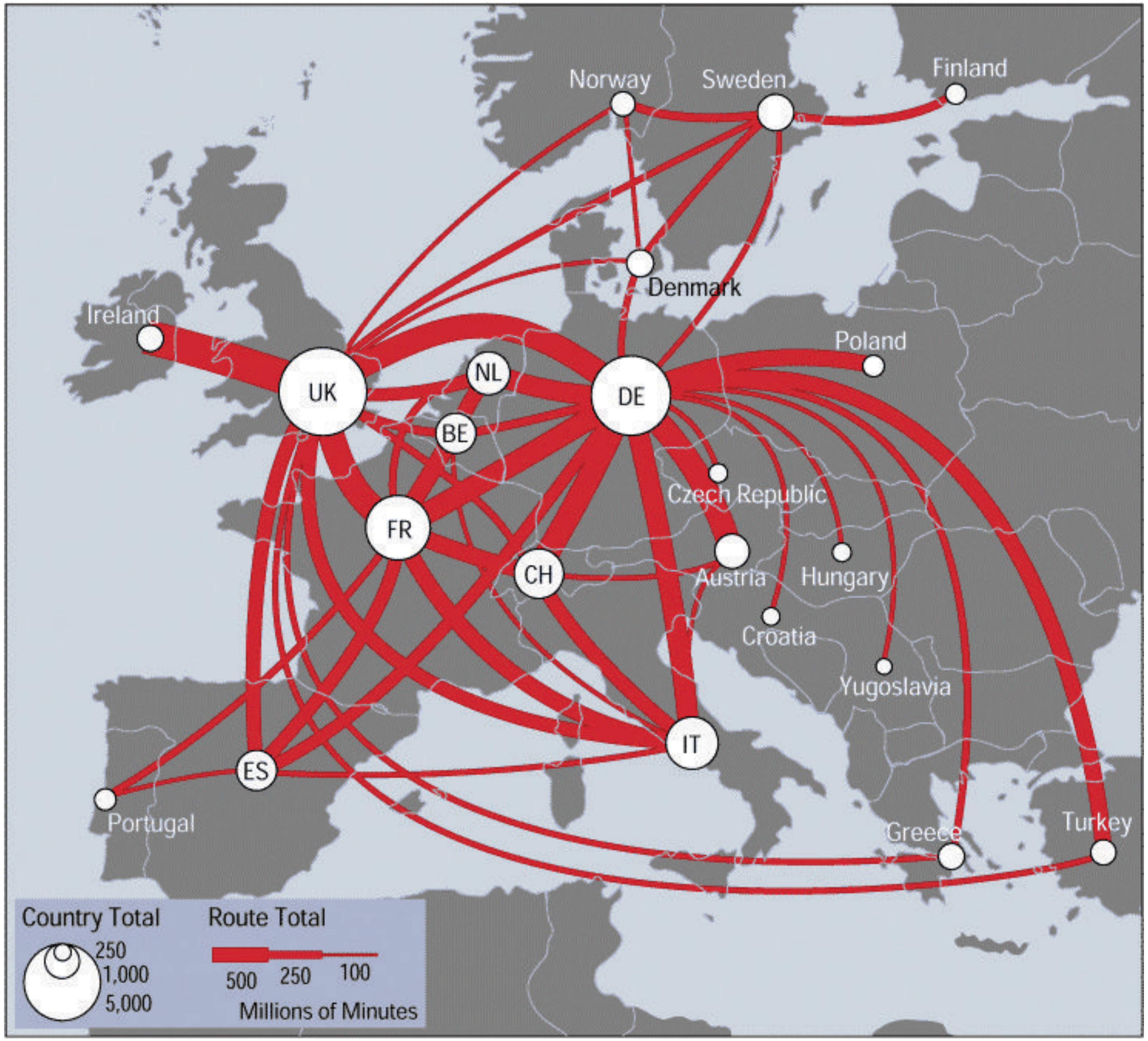


Figure 5.9. Linewidth has a limited number of discriminable bins.

Separability

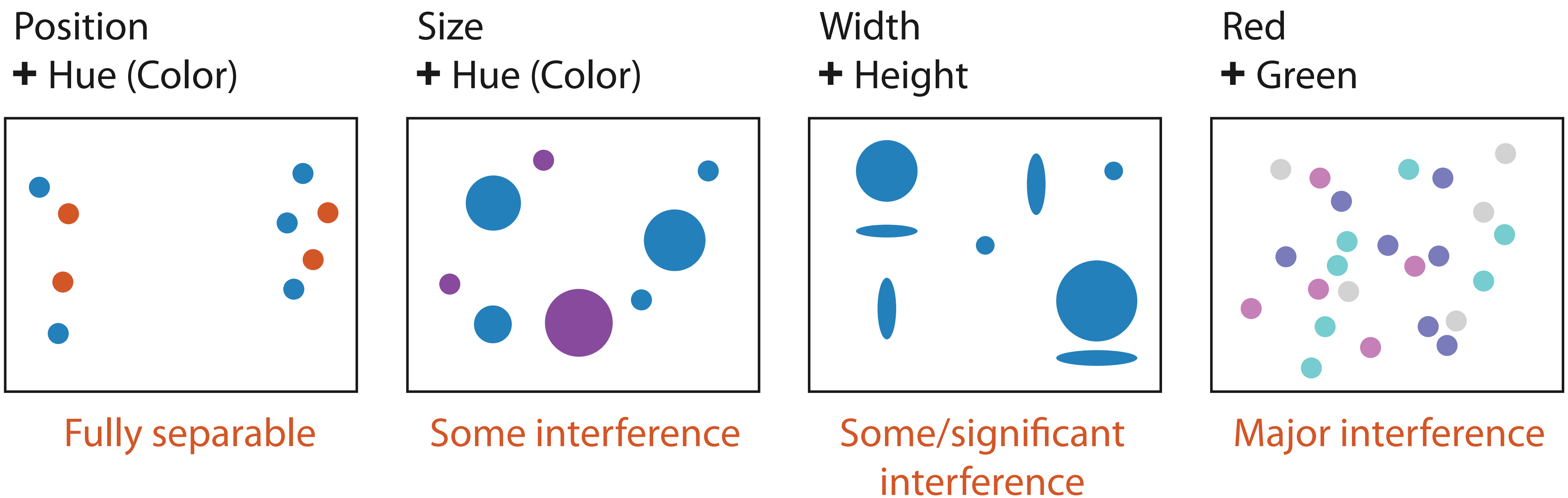


Figure 5.10. Pairs of visual channels fall along a continuum from fully separable to intrinsically integral. Color and location are separable channels well suited to encode different data attributes for two different groupings that can be selectively attended to. However, size interacts with hue, which is harder to perceive for small objects. The horizontal size and and vertical size channels are automatically fused into an integrated perception of area, yielding three groups. Attempts to code separate information along the red and green axes of the RGB color space fail, because we simply perceive four different hues. After [Ware 13, Figure 5.23].

Historical Perspective

Historical Perspective

- Bertin (1967) Semiology of Graphics
- Mackinlay (1986) APT
- Bergeron and Grinstein (1989) Visualization Reference Model
- Wehrend and Lewis (1990)
- Robertson (1990) Natural Scene Paradigm
- Roth (1991) Visage and SAGE
- Casner (1991) BOZ
- Beshers and Feiner (1992) AutoVisual

Historical Perspective

- Senay and Ignatius (1994) VISTA
- Hibbard (1994) Lattice Model
- Golovchinsky (1995) AVE
- Card, Mackinlay, and Shneiderman (1999) Spatial Substrate
- Kamps (1999) EAVE
- Wilkinson (1999) Grammar of Graphics
- Hoffman (2000) Table Visualizations

Historical Perspective

- **In 1967, Jacques Bertin, possibly the most important figure in visualization theory, published his *Sémiologie Graphique.***

Marks	Points, lines, and areas
Positional	Two planar dimensions
Retinal	Size, value, texture, color, orientation, and shape

Bertin's graphical vocabulary.

Historical Perspective

- Mackinlay (1986) introduced a design for an automated graphical presentation designer of relational information, named APT (A Presentation Tool)
- Mackinlay went on to describe graphical languages, defining graphical presentations as sentences of these languages. Two graphic design criteria: expressiveness criterion, the effectiveness criterion,
- The important aspect of Mackinlay’s work pertains to his composition algebra, a collection of primitive graphic languages and composition operators that can form complex presentations.

Historical Perspective

Marks	Points, lines, and areas
Positional	1D, 2D, and 3D
Temporal	Animation
Retinal	Color, shape, size, saturation, texture, and orientation

Mackinlay's graphical vocabulary, extended from Bertin.

Encoding Technique	**Primitive Graphical Language**
Retinal-list	Color, shape, size, saturation, texture, orientation
Single-position	Horizontal axis, vertical axis
Apposed-position	Line chart, bar chart, plot chart
Map	Road map, topographic map
Connection	Tree, acyclic graph, network
Misc. (angle, contain, ...)	Pie chart, Venn diagram, ...

Mackinlay's basis set of primitive graphical languages.

Keller and Keller (1994) Taxonomy of Visualization Goals

- **Task list**
  - **identify:** establish characteristics by which an object is recognizable
  - **locate:** ascertain the position (absolute or relative);
  - **distinguish:** recognize as distinct or different (identification is not needed);
  - **categorize:** place into divisions or classes;
  - **cluster:** group similar objects
  - **rank:** assign an order or position relative to other objects
  - **compare:** notice similarities and differences;
  - **associate:** link or join in a relationship that may or may not be of the same type;
  - **correlate:** establish a direct connection, such as causal or reciprocal.

Further Reading and Summary

Q&A

Further Reading

- **Pag 139 - 180 from Interactive Data Visualization: Foundations, Techniques, and Applications, Matthew O. Ward, Georges Grinstein, Daniel Keim, 2015**

- **Pag 42 - 64 from Visualization Analysis & Design, Tamara Munzner**

- **Check the slides by Sheelagh Carpendale, University of Calgary**
  - https://pages.cpsc.ucalgary.ca/~saul/hci_topics/pdf_files/visual-variables.pdf

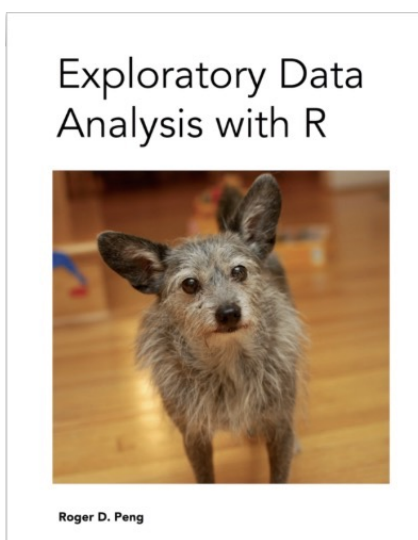
What you should know

- **The Visualization Process**
- **Expressiveness and Effectiveness**
- **The fundamental ideas of Semiology of Graphical Symbols**
  - data -> (x, y, z*)
- **The eight visual variables(VV)**
  - position, shape - Why they are the most important !
  - the others VVs
- **Effects of Visual Variables**
  - selective, associative, quantitative, order
- **Tasks list(s)**
  - Why it is important to consider a task; Why it is important to consider a taxonomy

One more thing !

General Rules for Exploratory Data Analysis

General Rules for Exploratory Data Analysis



Exploratory Data Analysis with R

Roger D. Peng

Principles of Analytic Graphics

- **Principle 1: Show comparisons**
  - **Evidence for a hypothesis is always relatives another competing hypothesis**
  - **Always ask “Compared to What?"**

Principles of Analytic Graphics

Principle 1: Show comparisons

- **Evidence for a hypothesis is always relatives another competing hypothesis**
- **Always ask “Compared to What?"**

Testing whether an **air cleaner installed in a child’s home improves their asthma-related symptoms**.

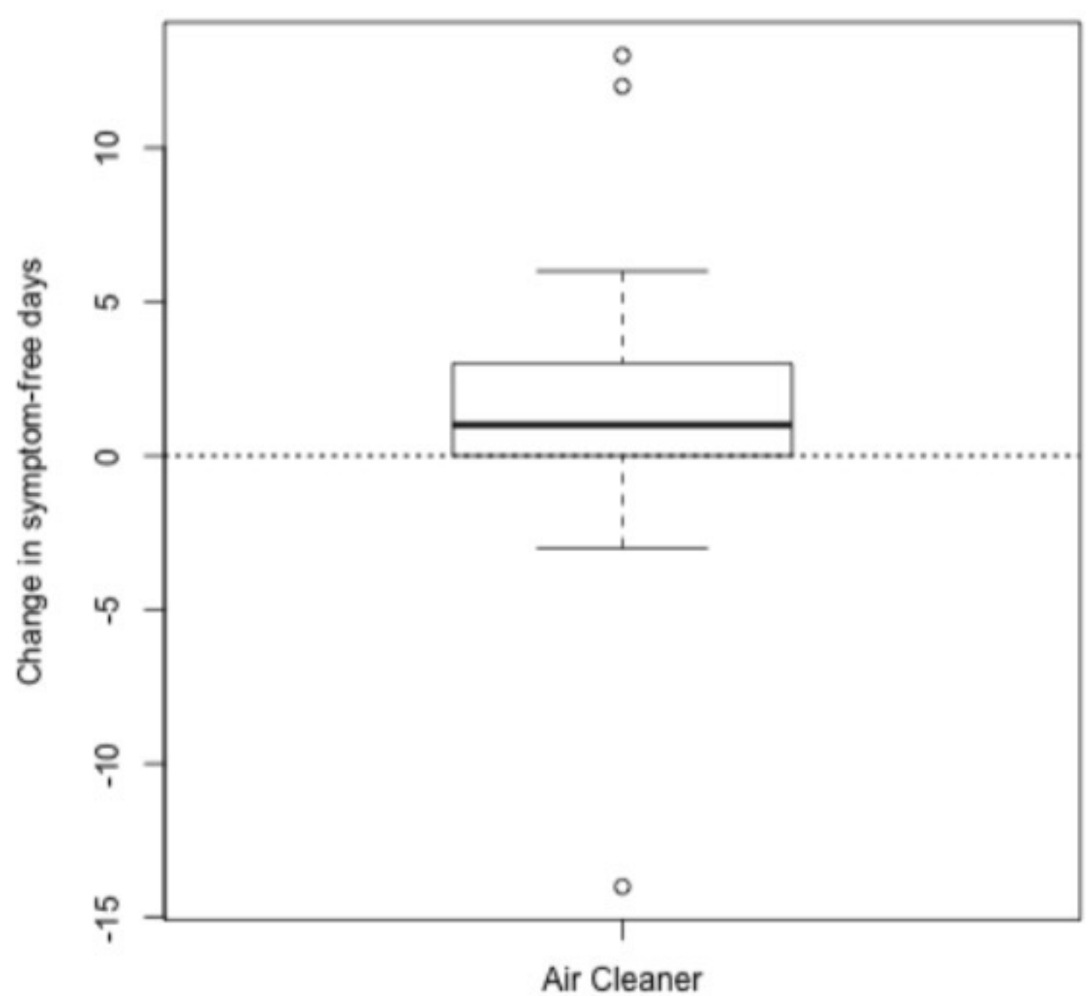
This study was conducted at the Johns Hopkins University School of Medicine and was conducted in homes where a smoker was living for at least 4 days a week.

Each child was assessed at baseline and then 6-months later at a second visit. The aim was to improve a child’s symptom-free days over the 6-month period. In this case, a higher number is better, indicating that they had *more* symptom-free days.

Reference: Butz AM, *et al.*, *JAMA Pediatrics*, 2011.

Principles of Analytic Graphics

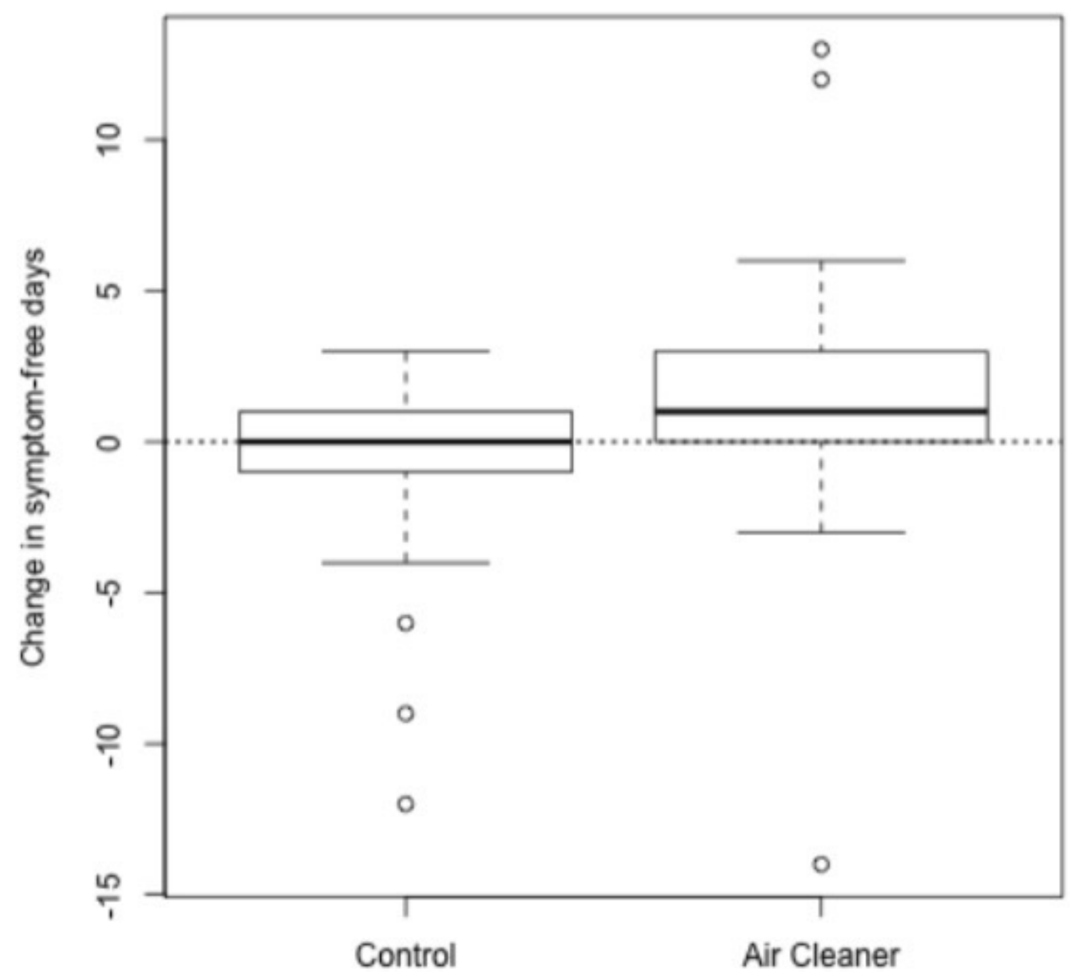
- **Principle 1: Show comparisons**
  - **Evidence for a hypothesis is always relatives another competing hypothesis**
  - **Always ask “Compared to What?"**



Reference: Butz AM, *et al.*, *JAMA Pediatrics*, 2011.

Principles of Analytic Graphics

- **Principle 1: Show comparisons**
  - **Evidence for a hypothesis is always relatives another competing hypothesis**
  - **Always ask “Compared to What?"**



Reference: Butz AM, *et al.*, *JAMA Pediatrics*, 2011.

Principles of Analytic Graphics

- **Principle 1: Show comparisons**
  - **Evidence for a hypothesis is always relatives another competing hypothesis**
  - **Always ask “Compared to What?”**
- **Principle 2: Show causality, mechanism, explanation, systematic structure**
  - **What is your causal framework for thinking about a question?**

Principles of Analytic Graphics

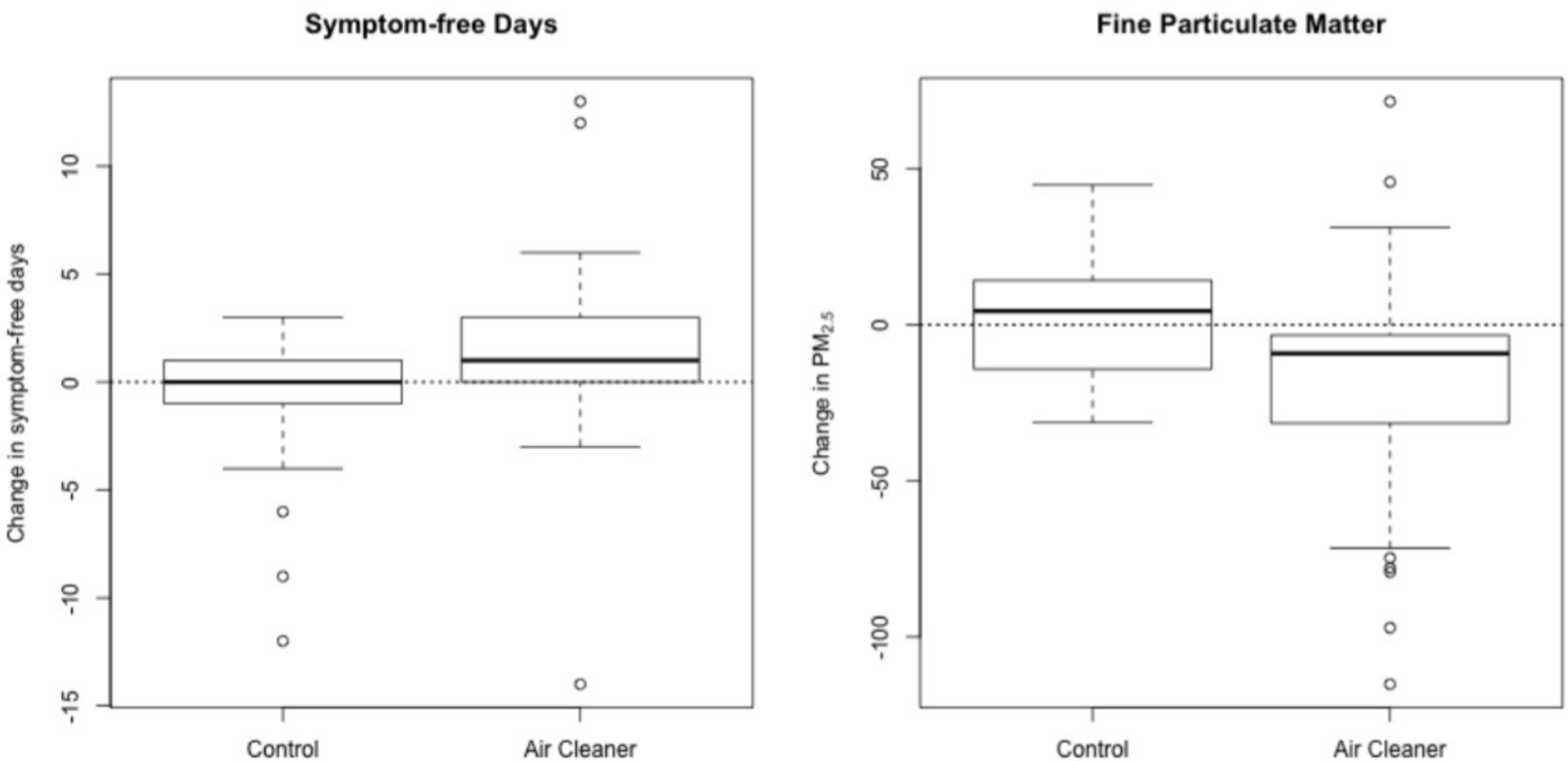
Air quality standards in the U.S. concerns the long-term average level of fine particle pollution, also referred to as PM2.5

The standard says that the "annual mean, averaged over 3 years" cannot exceed 12 micrograms per cubic meter.

FACULDADE DE CIÊNCIAS E TECNOLOGIA UNIVERSIDADE NOVA DE LISBOA

Principles of Analytic Graphics

- **Principle 2: Show causality, mechanism, explanation, systematic structure**
  - **What is your causal framework for thinking about a question?**



Reference: Butz AM, *et al.*, *JAMA Pediatrics*, 2011.

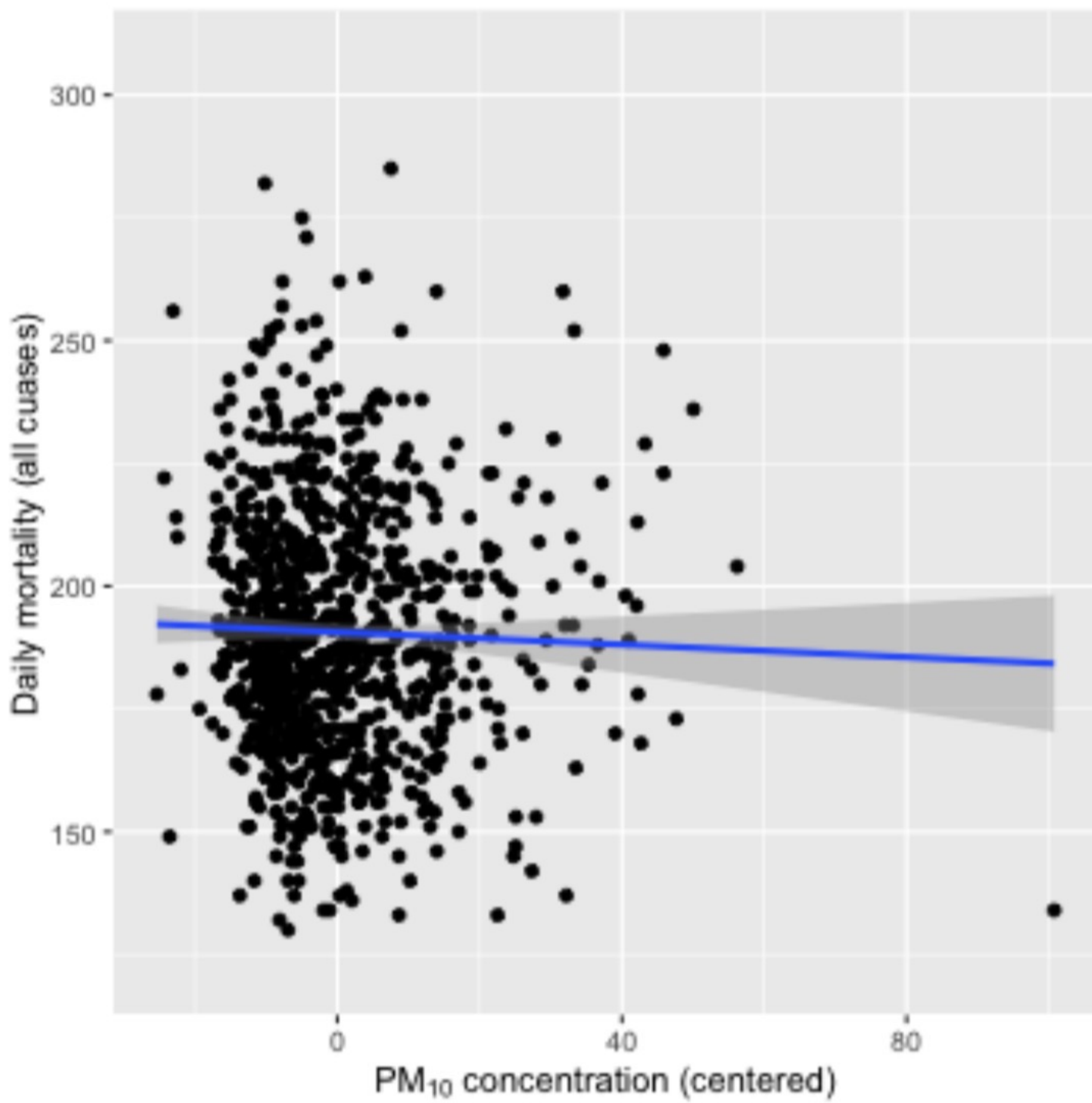
FACULDADE DE CIÊNCIAS E TECNOLOGIA
UNIVERSIDADE NOVA DE LISBOA

Principles of Analytic Graphics

- **Principle 1: Show comparisons**
- **Principle 2: Show causality, mechanism, explanation, systematic structure**
- **Principle 3: Show multivariate data**
  - **Multivariate = more than 2 variables**
  - **The real world is multivariate**
  - **Need to “escape flatland"**

Principles of Analytic Graphics

- **Principle 3: Show multivariate data**



PM10 and mortality in New York City

it **seems** that there is a **slight negative relationship between the two variables**.

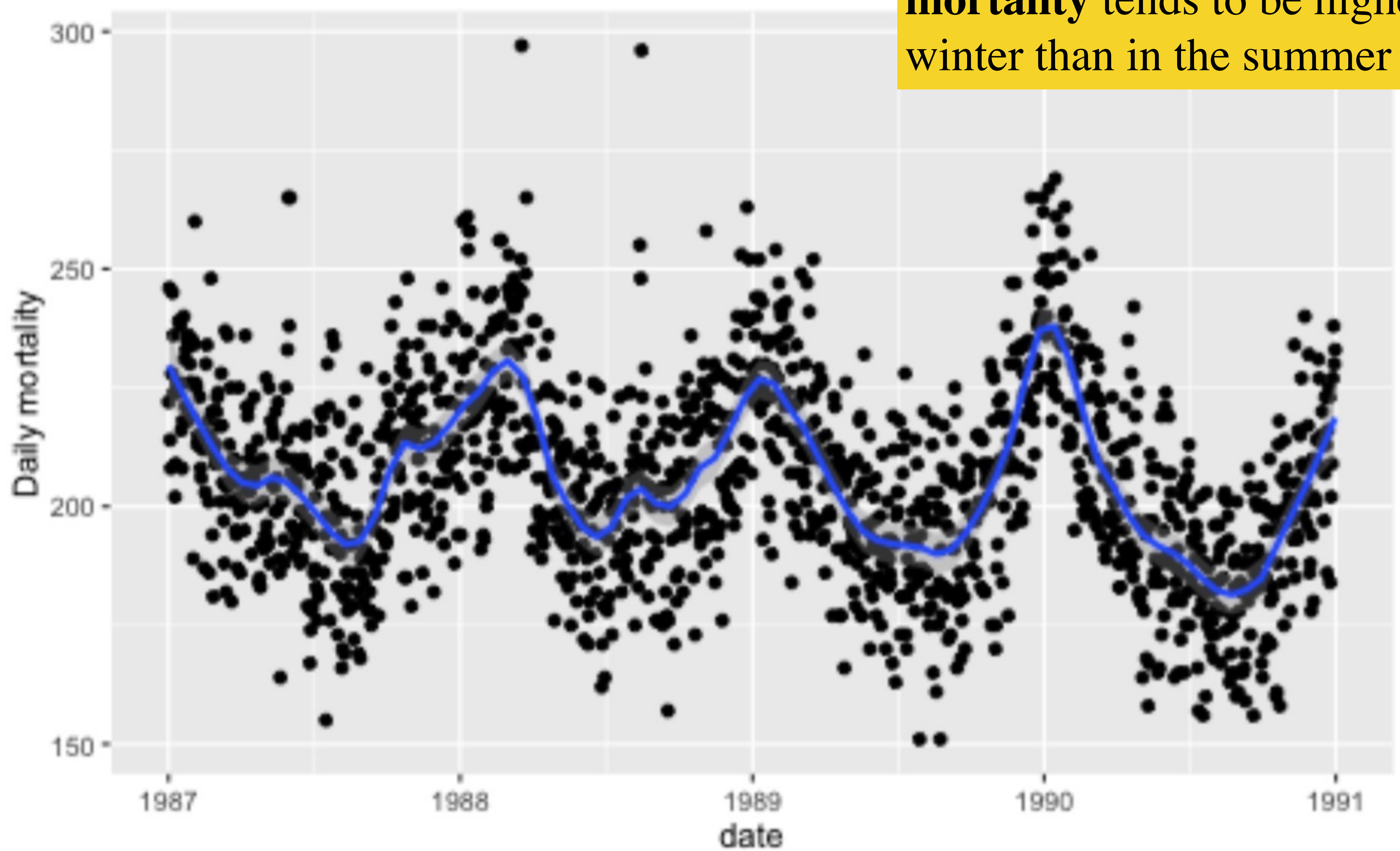
That is, higher daily average levels of PM10 appear to be associated with lower levels of mortality (fewer deaths per day).

Reference: Butz AM, *et al.*, *JAMA Pediatrics*, 2011.

Principles of Analytic Graphics

- **Principle 3: Show multivariate data**

mortality tends to be higher in the winter than in the summer



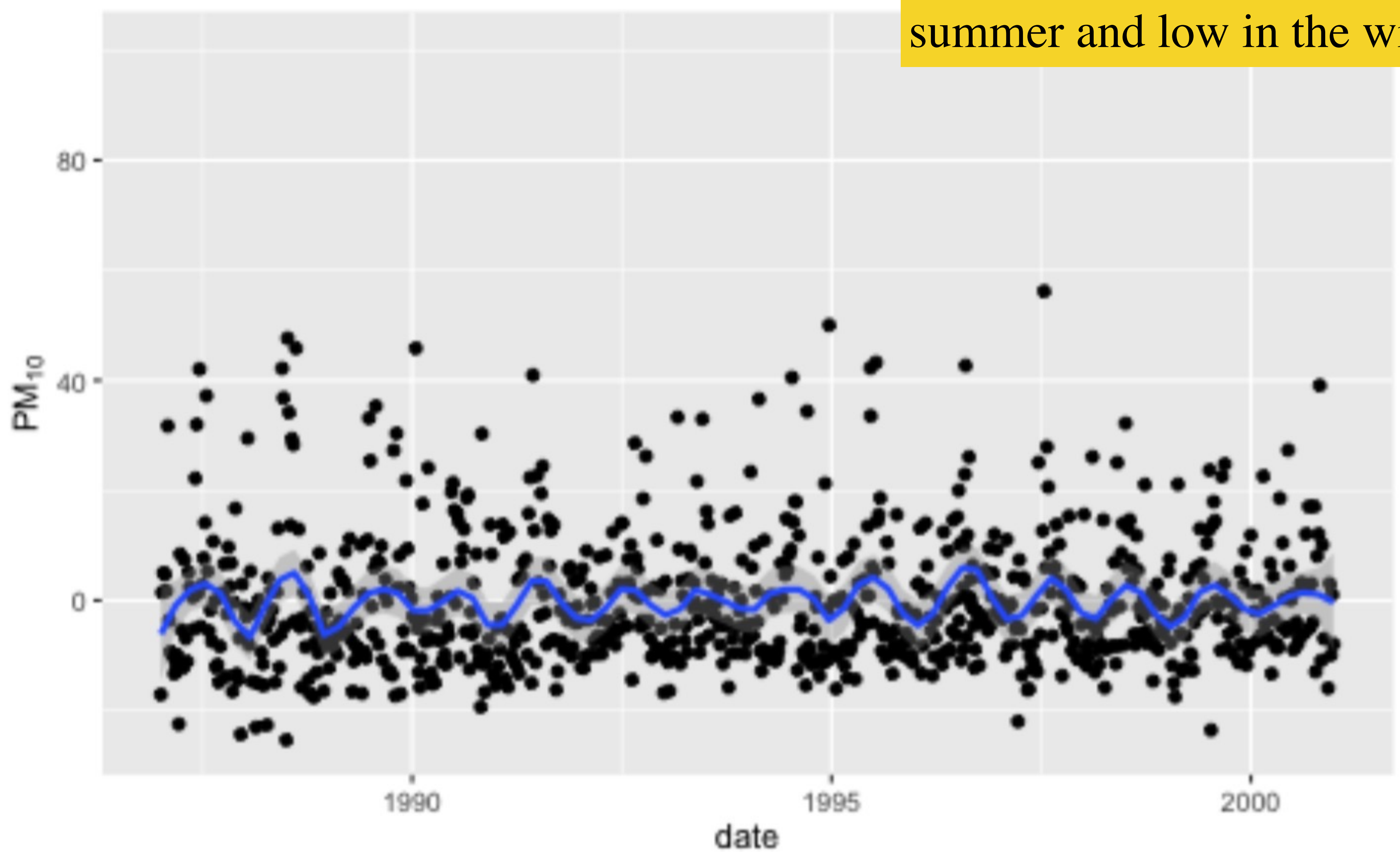
Daily mortality in New York City

Reference: Butz AM, *et al.*, *JAMA Pediatrics*, 2011.

Principles of Analytic Graphics

- **Principle 3: Show multivariate data**

PM10 levels tend to be high in the summer and low in the winter.



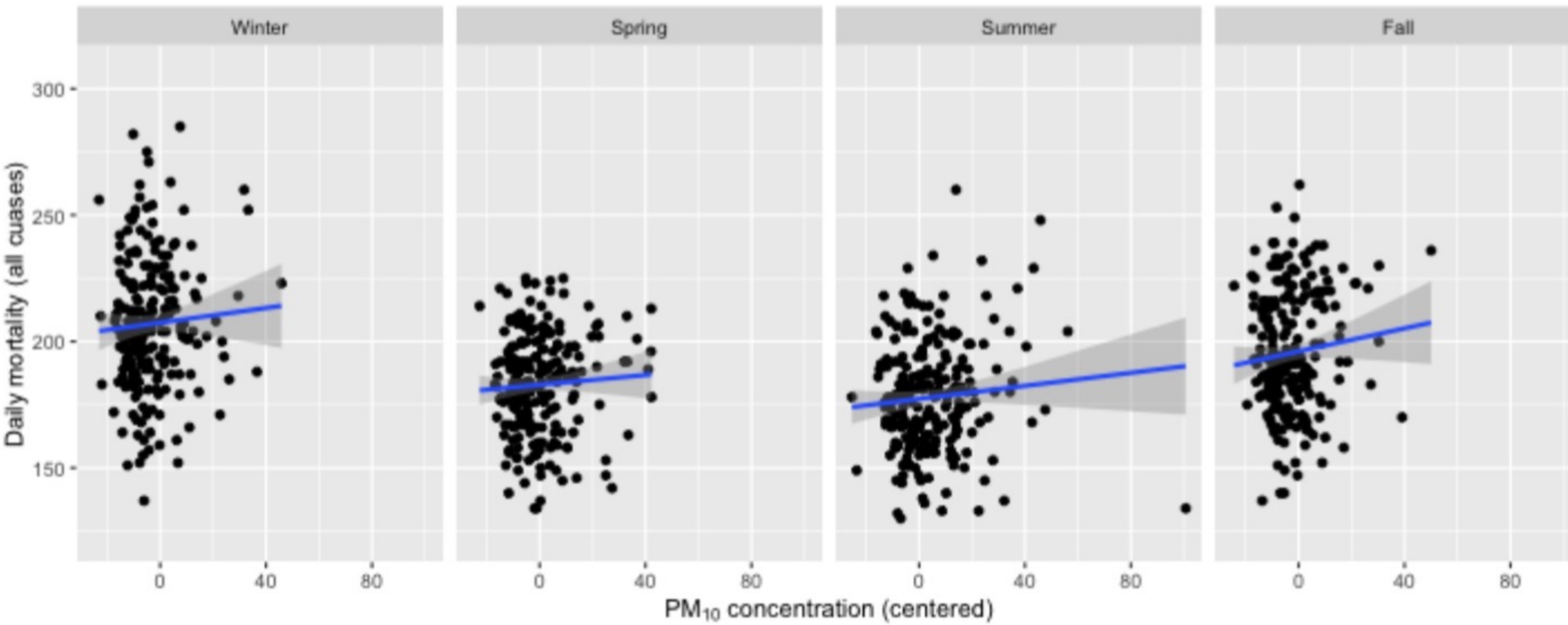
Daily PM10 in New York City

Reference: Butz AM, *et al.*, *JAMA Pediatrics*, 2011.

Principles of Analytic Graphics

- **Principle 3: Show multivariate data**

There is a **slight positive** relationship between the two variables in each season



PM10 and mortality in New York City by season

Reference: Butz AM, *et al.*, *JAMA Pediatrics*, 2011.

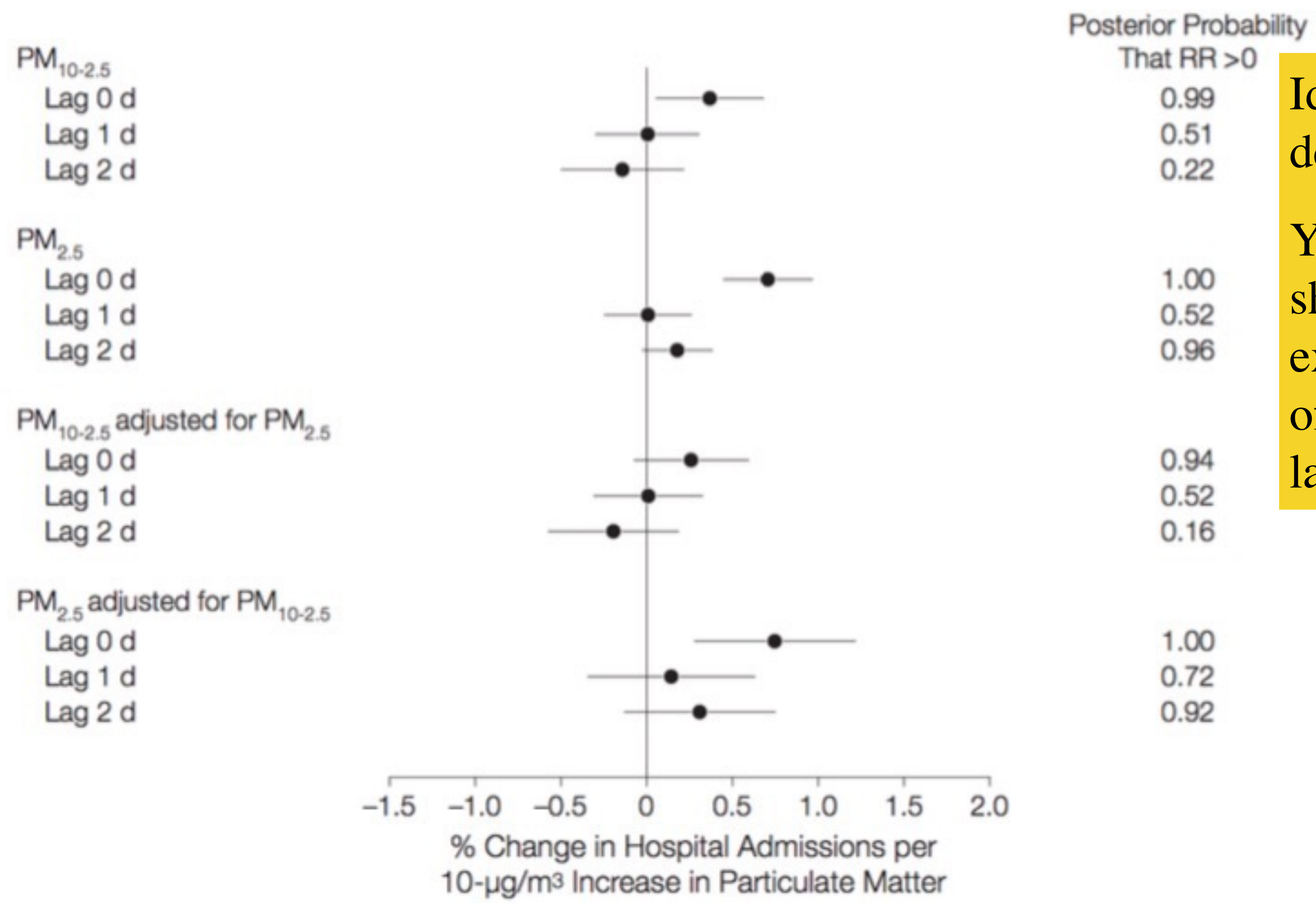
Principles of Analytic Graphics

- Principle 1: Show comparisons
- Principle 2: Show causality, mechanism, explanation, systematic structure
- Principle 3: Show multivariate data
- Principle 4: Integration of evidence
  - Completely integrate words, numbers, images, diagrams
  - Data graphics should make use of many modes of data presentation
  - Don't let the tool drive the analysis

Principles of Analytic Graphics

- **Principle 4: Integration of evidence**

Figure 2. Percentage Change in Emergency Hospital Admissions Rate for Cardiovascular Diseases per a 10-μg/m³ Increase in Particulate Matter



Estimates are on average across 108 counties. $PM_{2.5}$ indicates particulate matter is 2.5 μm or less in aerodynamic diameter; PM_{10}, particulate matter is 10 μm or less in aerodynamic diameter; $PM_{10\text{-}2.5}$, particulate matter is greater than 2.5 μm and 10 μm or less in aerodynamic diameter; RR, relative risk. Error bars indicate 95% posterior intervals.

Ideally, a plot would have all of the necessary descriptions attached to it.

You might think that this level of documentation should be reserved for “final” plots as opposed to exploratory ones, but it’s good to get in the habit of documenting your evidence sooner rather than later.

Reference: Butz AM, *et al.*, *JAMA Pediatrics*, 2011.

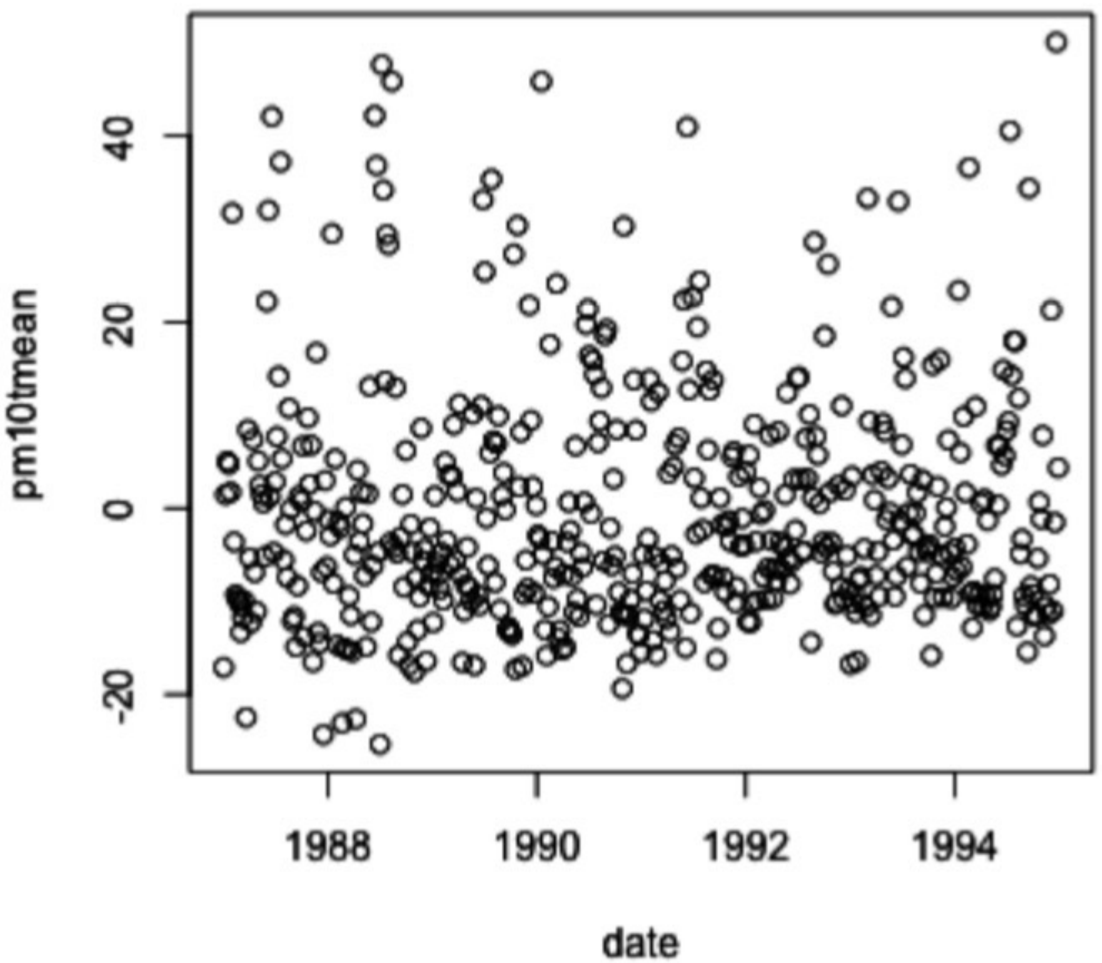
Principles of Analytic Graphics

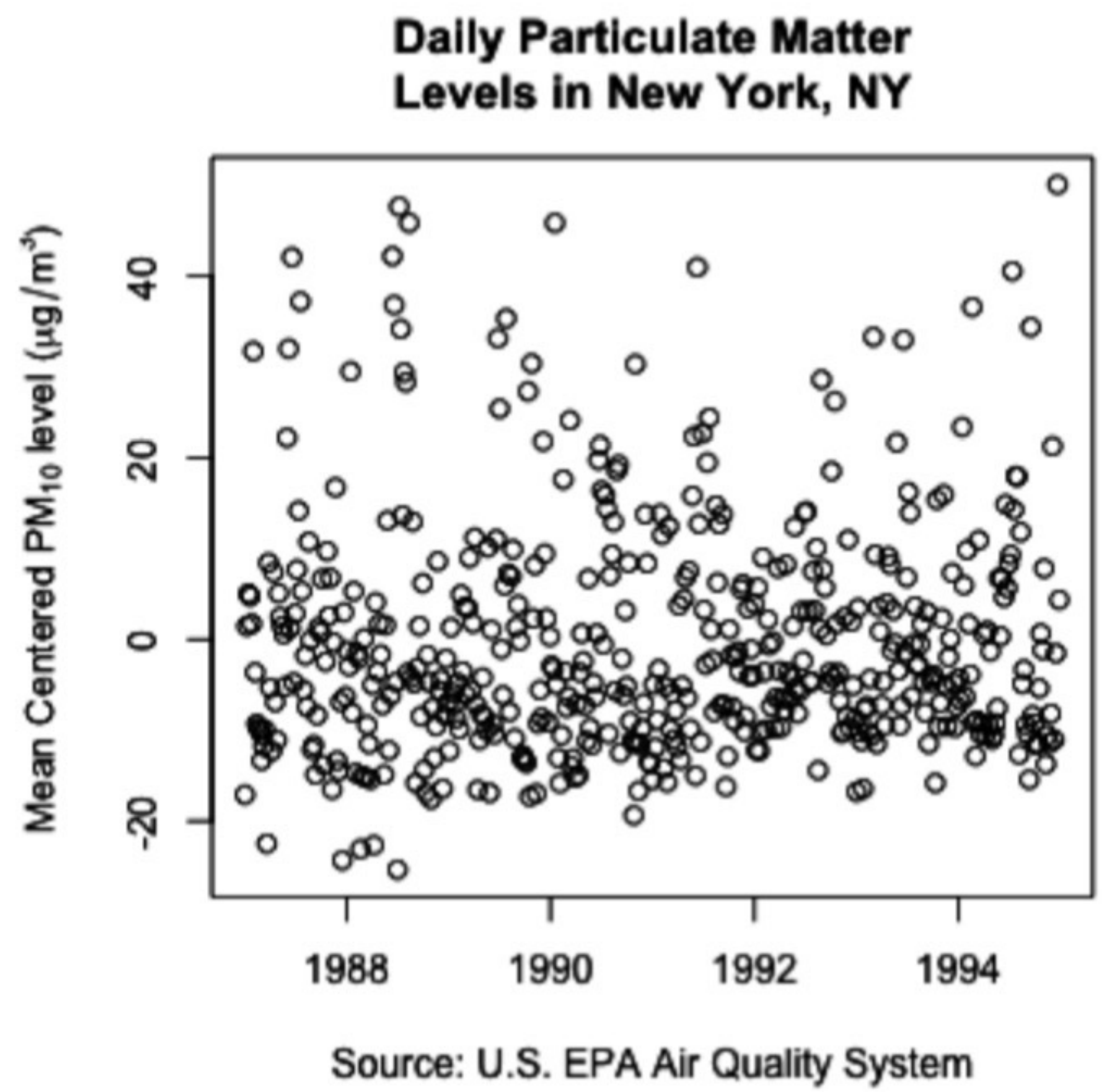
- **Principle 1: Show comparisons**
- **Principle 2: Show causality, mechanism, explanation, systematic structure**
- **Principle 3: Show multivariate data**
- **Principle 4: Integration of evidence**
- **Principle 5: Describe and document the evidence with appropriate labels, scales, sources**

Principles of Analytic Graphics

- **Principle 5: Describe and document the evidence with appropriate labels, scales, sources**

Defaults





Reference: Butz AM, *et al.*, *JAMA Pediatrics*, 2011.

Principles of Analytic Graphics

- Principle 1: Show comparisons
- Principle 2: Show causality, mechanism, explanation, systematic structure
- Principle 3: Show multivariate data
- Principle 4: Integration of evidence
- Principle 5: Describe and document the evidence with appropriate labels, scales, sources
- Principle 6: Content is King
  - Analytical presentations ultimately stand or fall depending on the quality, relevance, and integrity of their content.

Principles of Analytic Graphics

- **Principle 1: Show comparisons**
- **Principle 2: Show causality, mechanism, explanation, systematic structure**
- **Principle 3: Show multivariate data**
- **Principle 4: Integration of evidence**
- **Principle 5: Describe and document the evidence with appropriate labels, scales, sources**
- **Principle 6: Content is King**

Edward Tufte (2006). *Beautiful Evidence*, Graphics Press LLC.

www.edwardtufte.com

Reference: Butz AM, *et al.*, *JAMA Pediatrics*, 2011.