



DataViz 101

The background of the slide is an abstract composition of teal and light blue colors, resembling a watercolor or marbled texture. A solid white horizontal band runs across the middle of the image, serving as a backdrop for the title text.

Data Visualization

Data Visualization

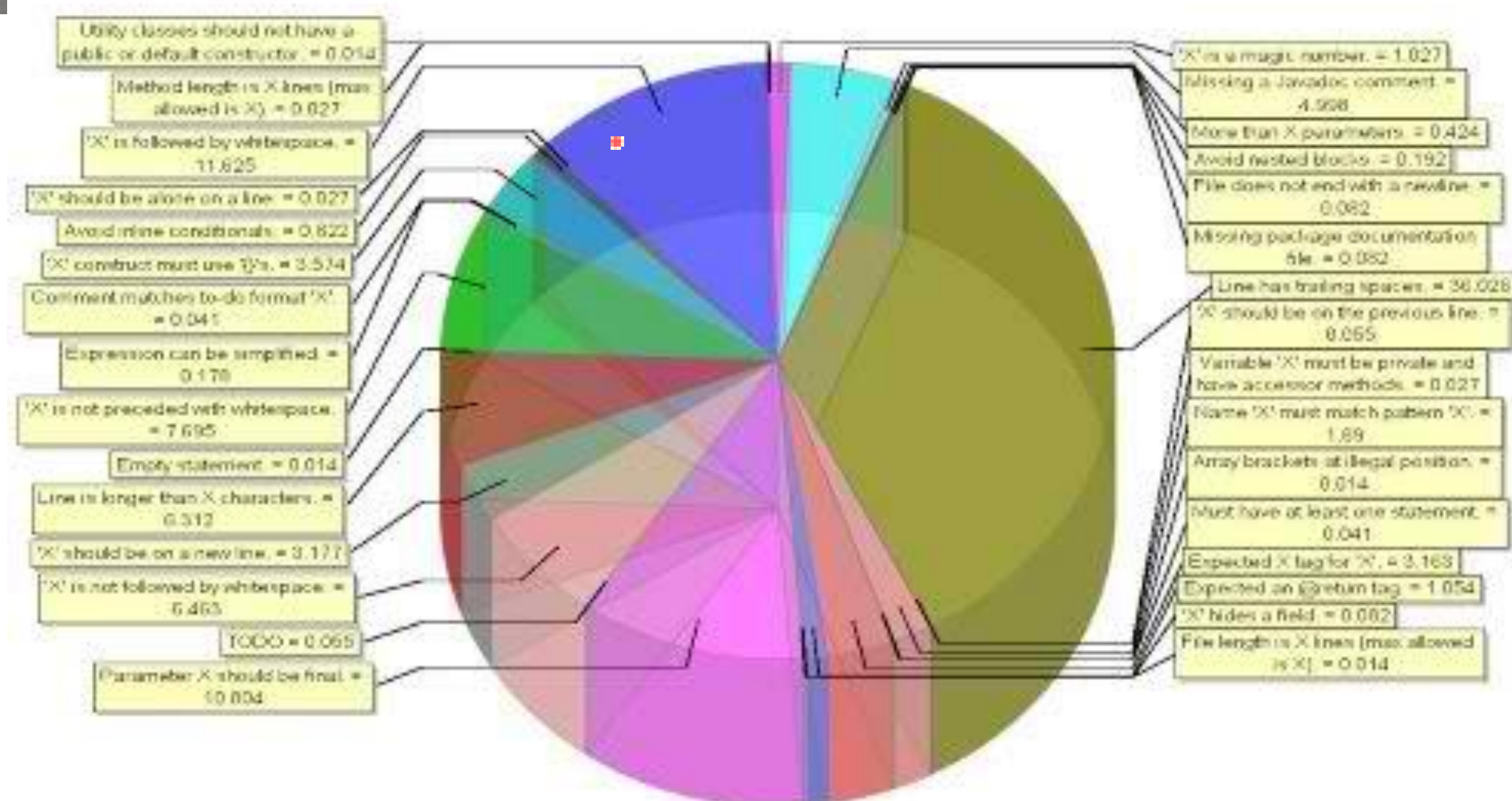
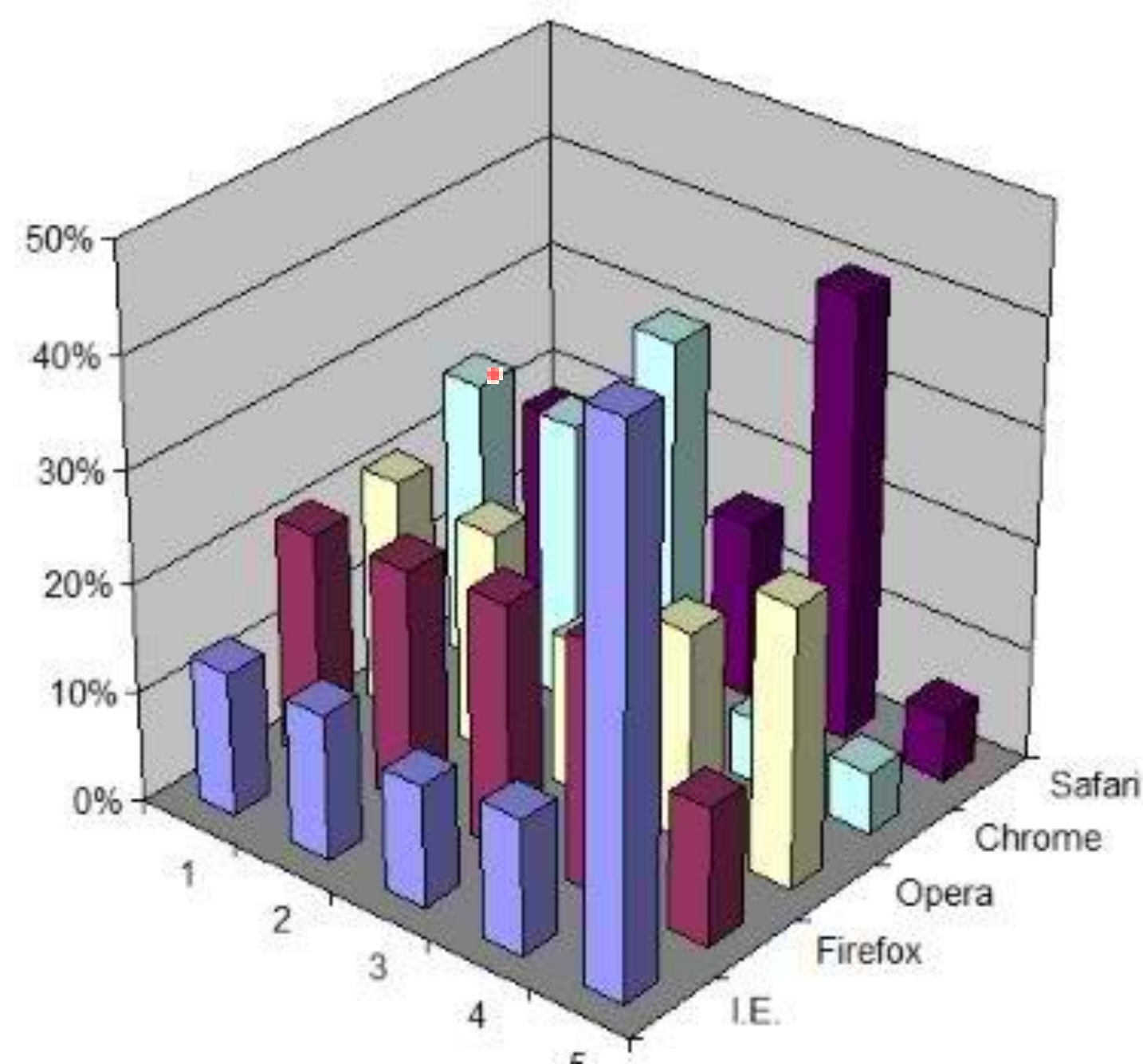
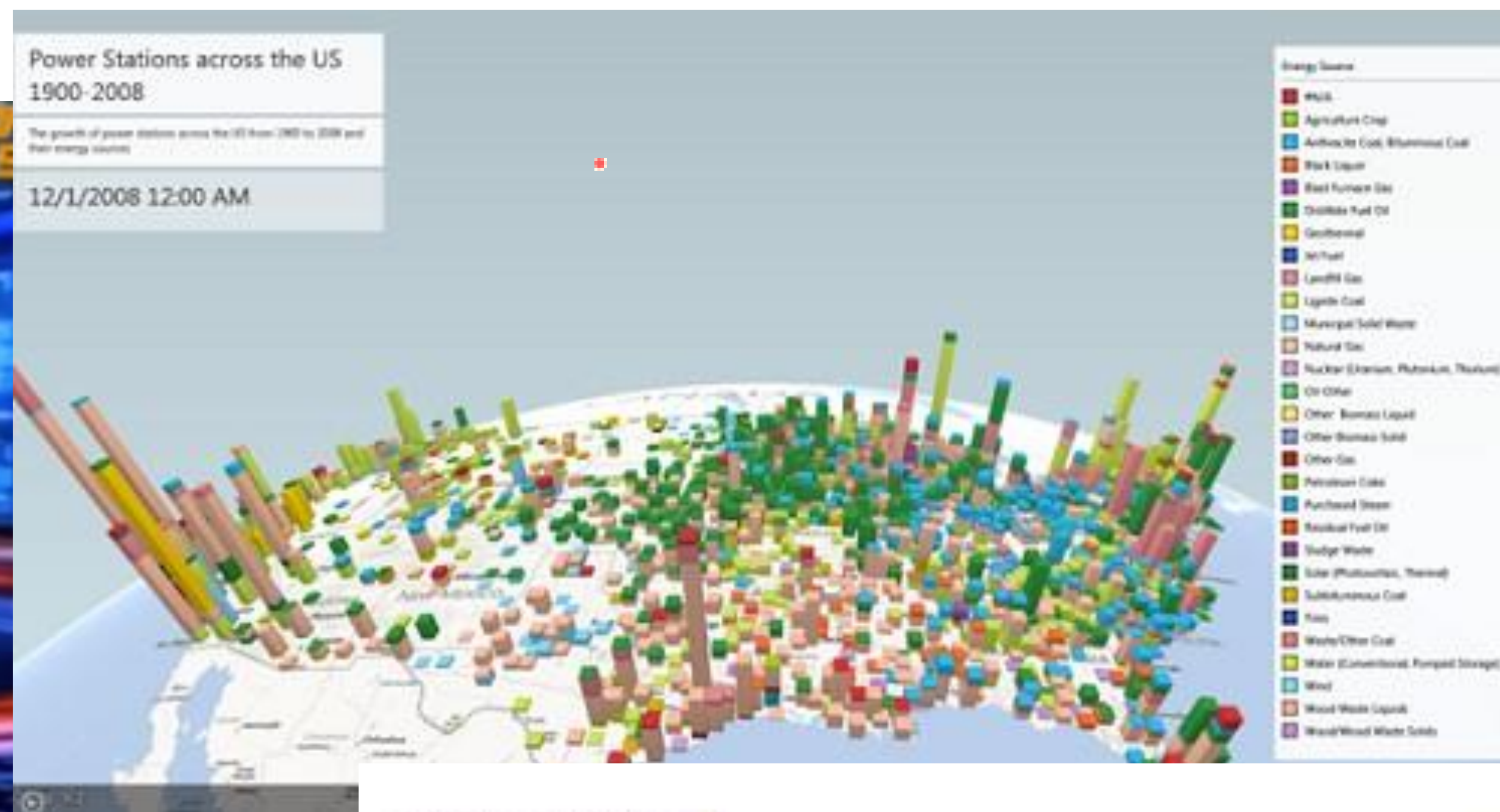
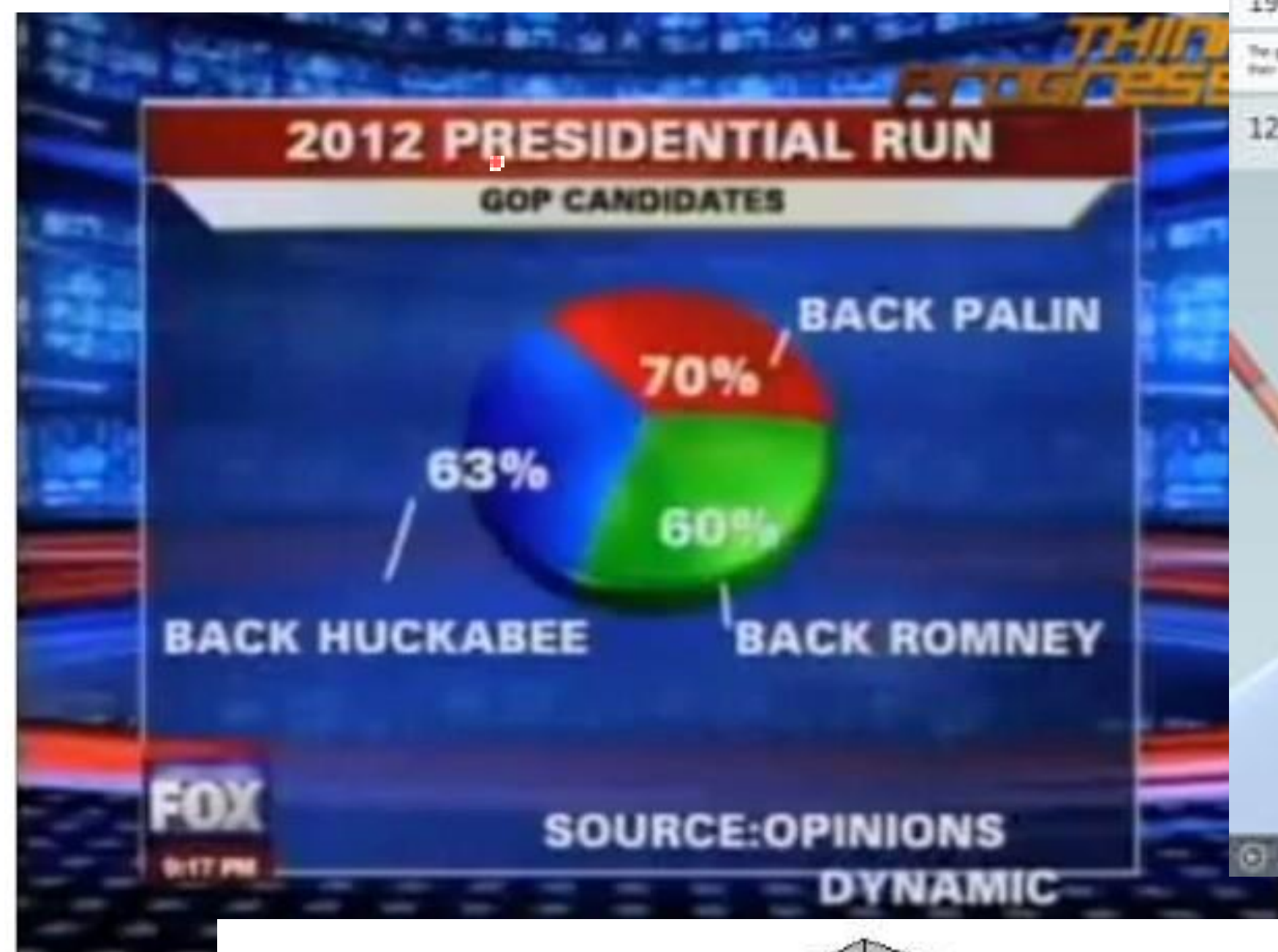
is the representation of data or information in a graph, chart, or other visual format.

It communicates relationships of the data with images.

Data Visualization

*is the secret art
of turning data into understandable visuals.*



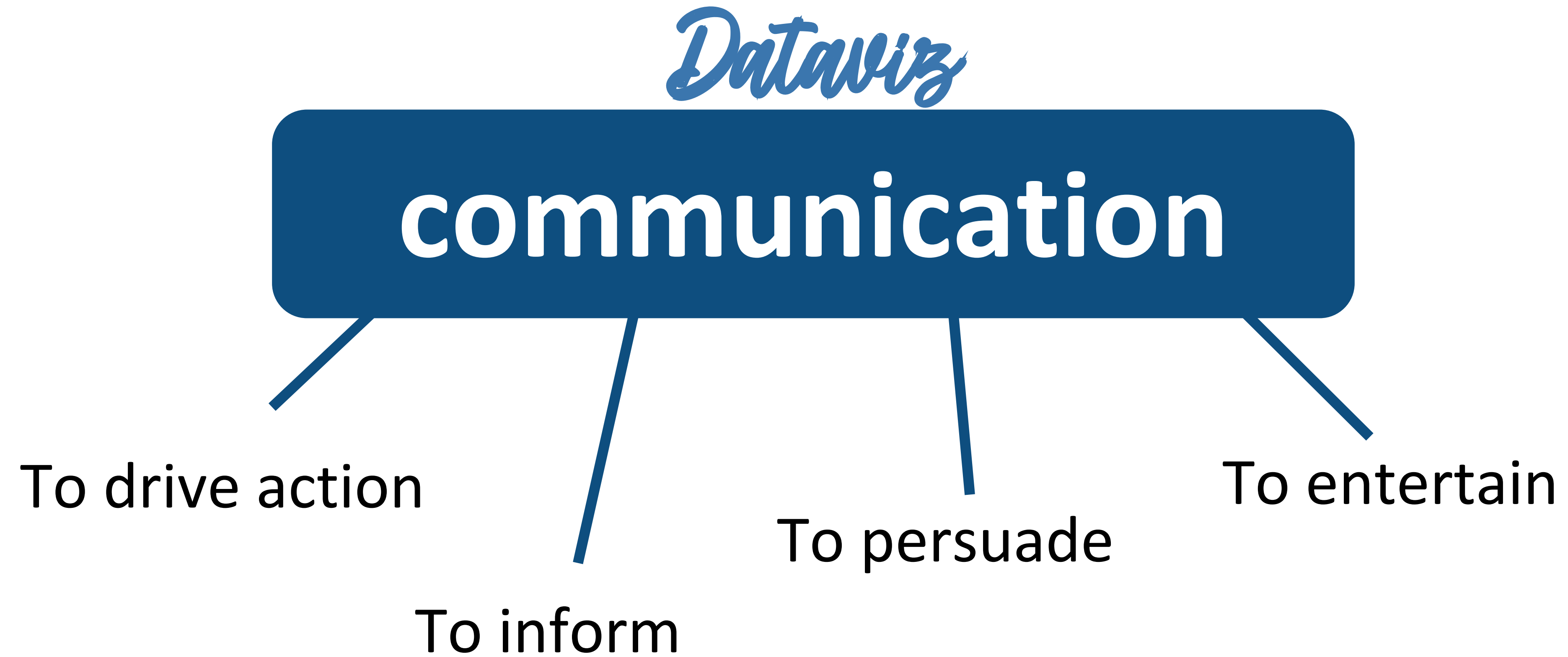


Let's have a warm up

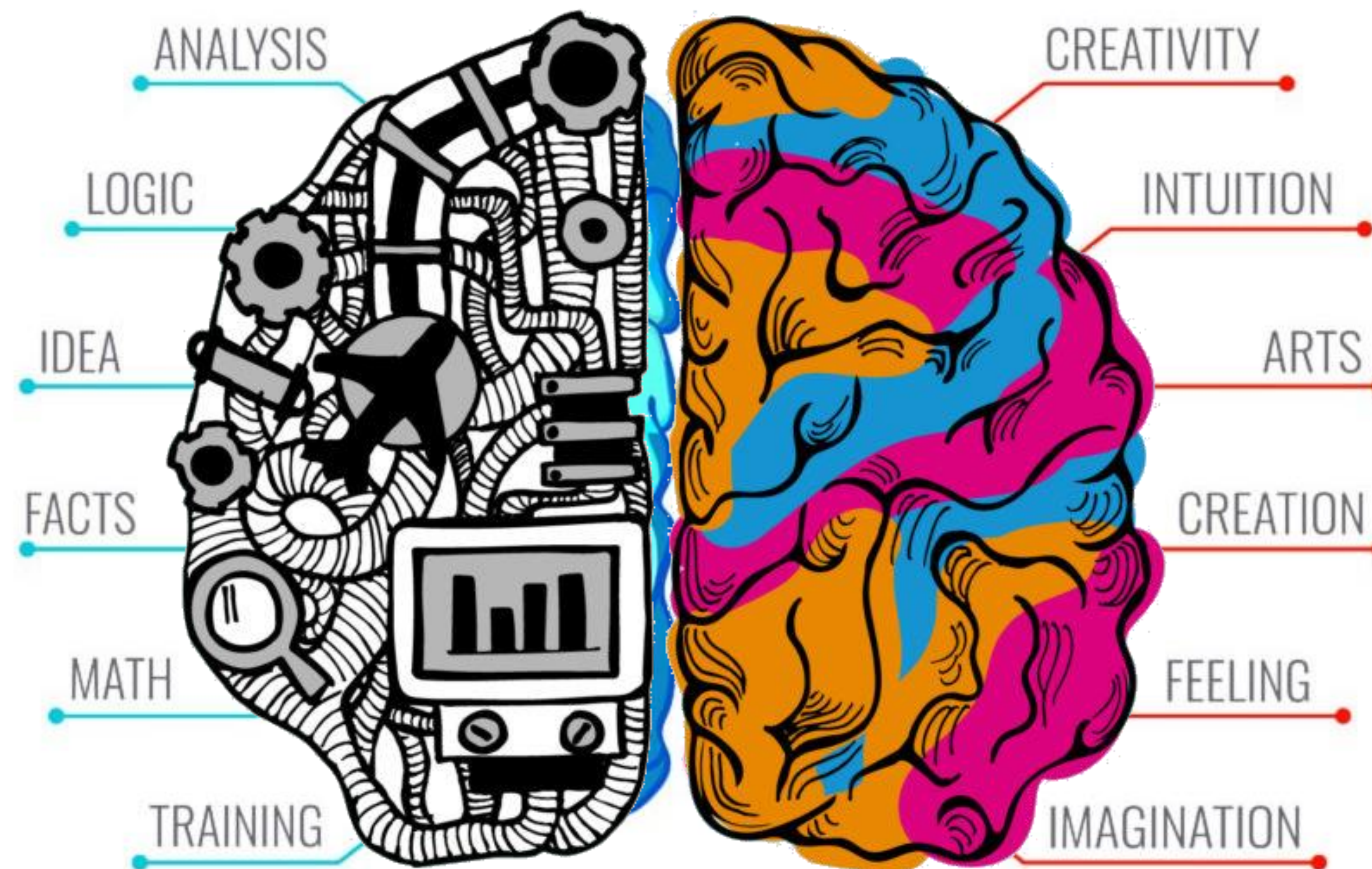
What's a good
Data Visualization?



Data Visualization
is communication



Dataviz touches both side



Dataviz Questions

1. Which data is important to show?

2. What do I want to emphasize in the data?

3. What options do I have for displaying this data?

4. Which option is most effective in communicating the data?

Dataviz Questions

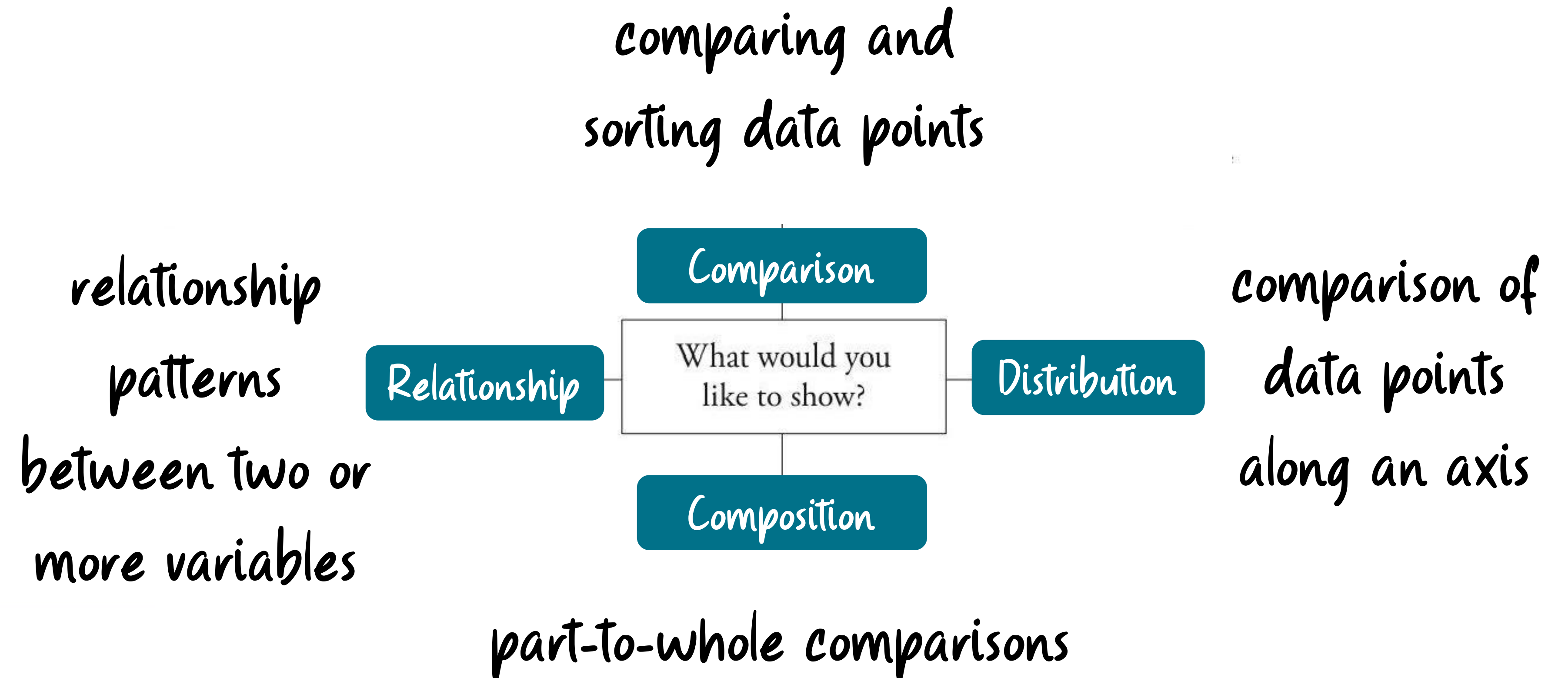
1. Which data is important to show?

2. What do I want to emphasize in the data?

3. What options do I have for displaying this data?

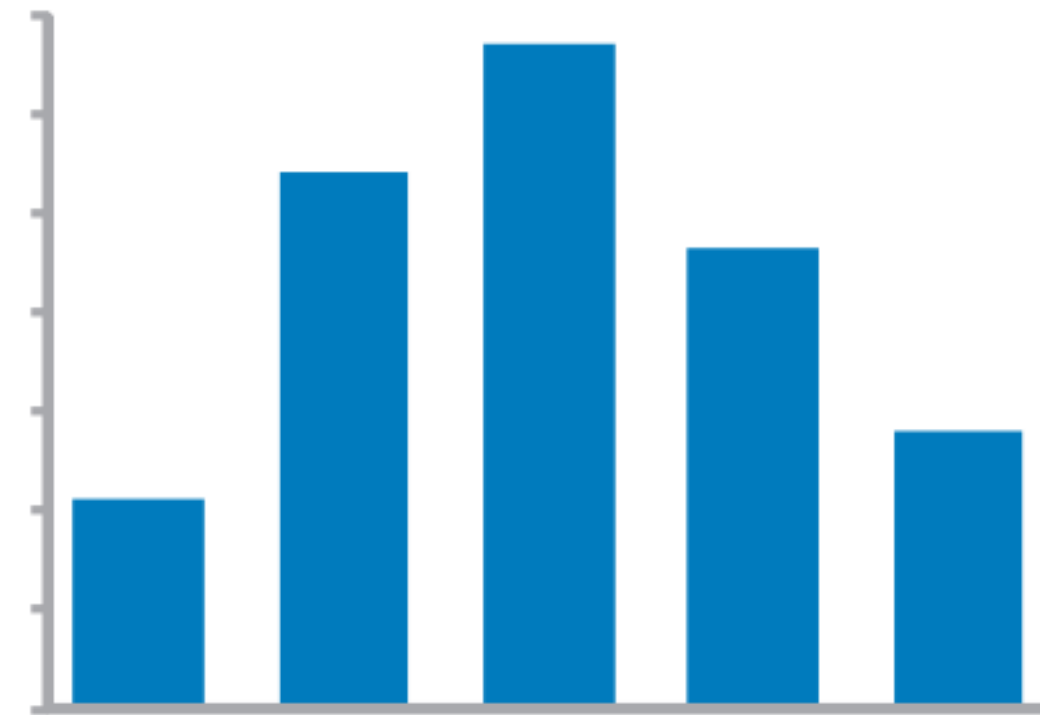
4. Which option is most effective in communicating the data?

Chart Suggestions



Time series

values display how something changed over time



Bar Graph (vertical)



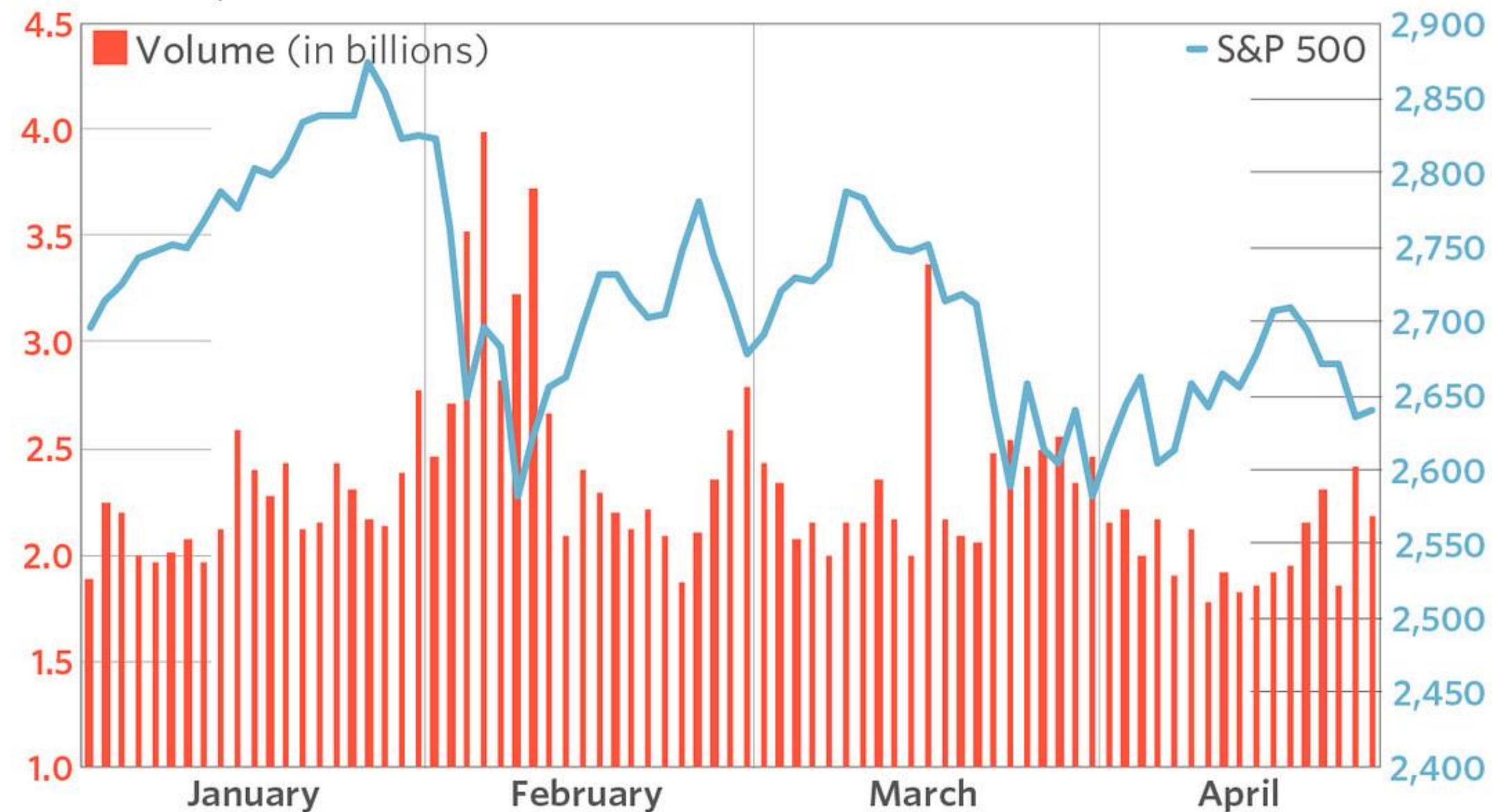
Line Graph

Time series

values display how something changed over time

Does volume follow trend?

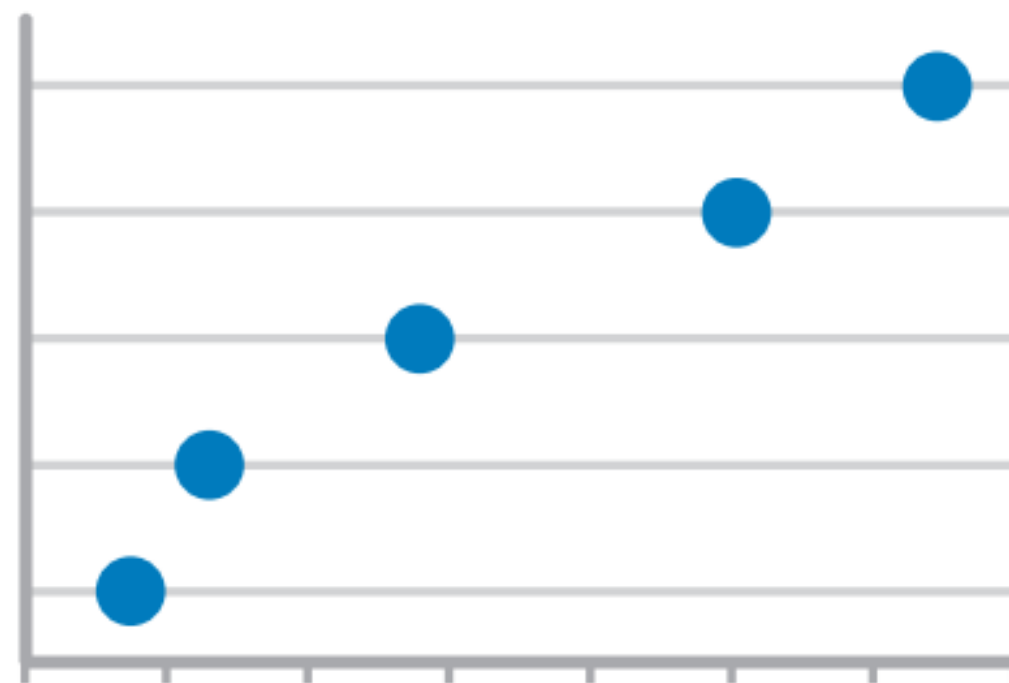
S&P 500's price and volume so far in 2018



Source: FactSet

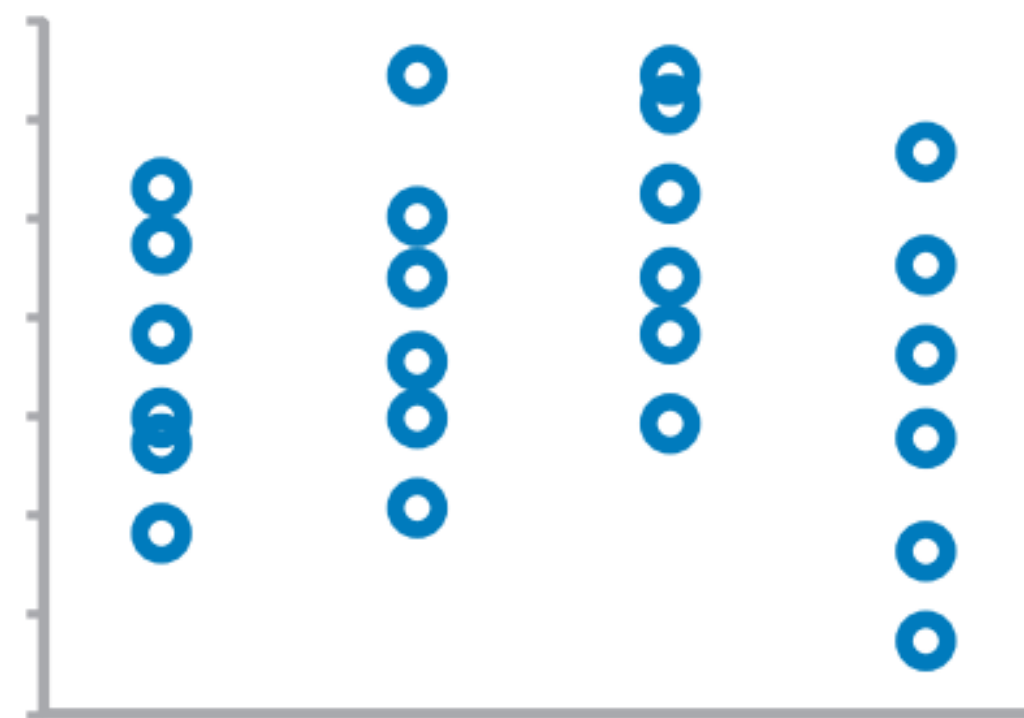
Time series

values display how something changed over time



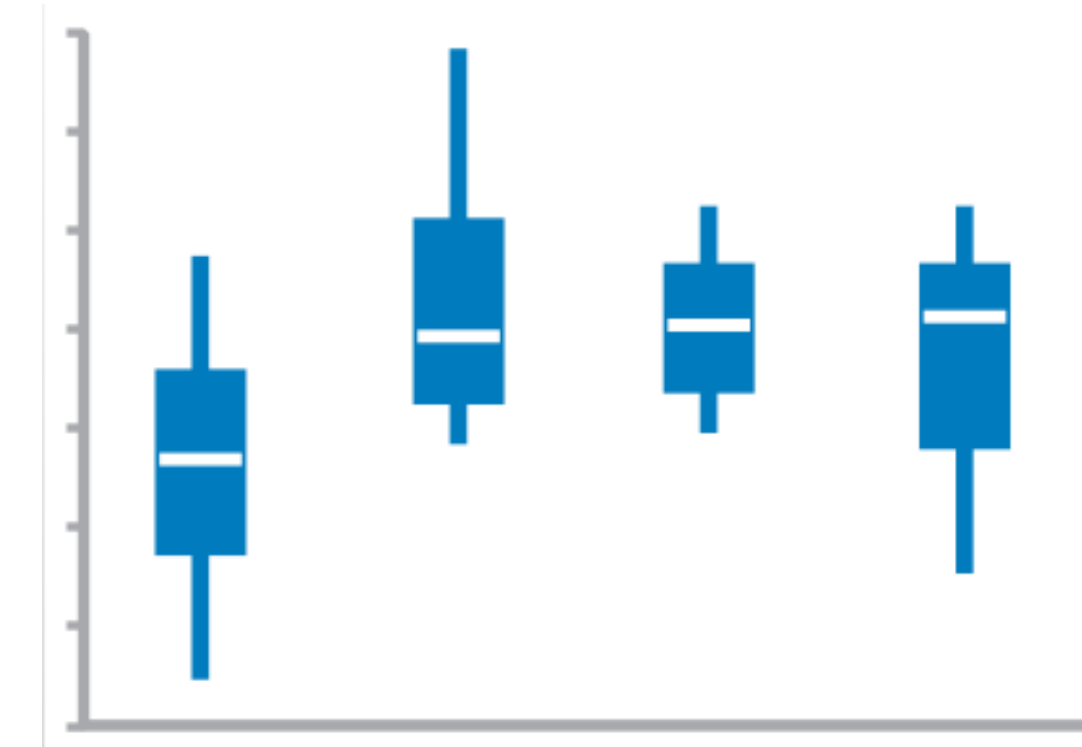
Dot Plot (vertical)

When you do not have a value for every interval of time



Strip Plot (multiple)

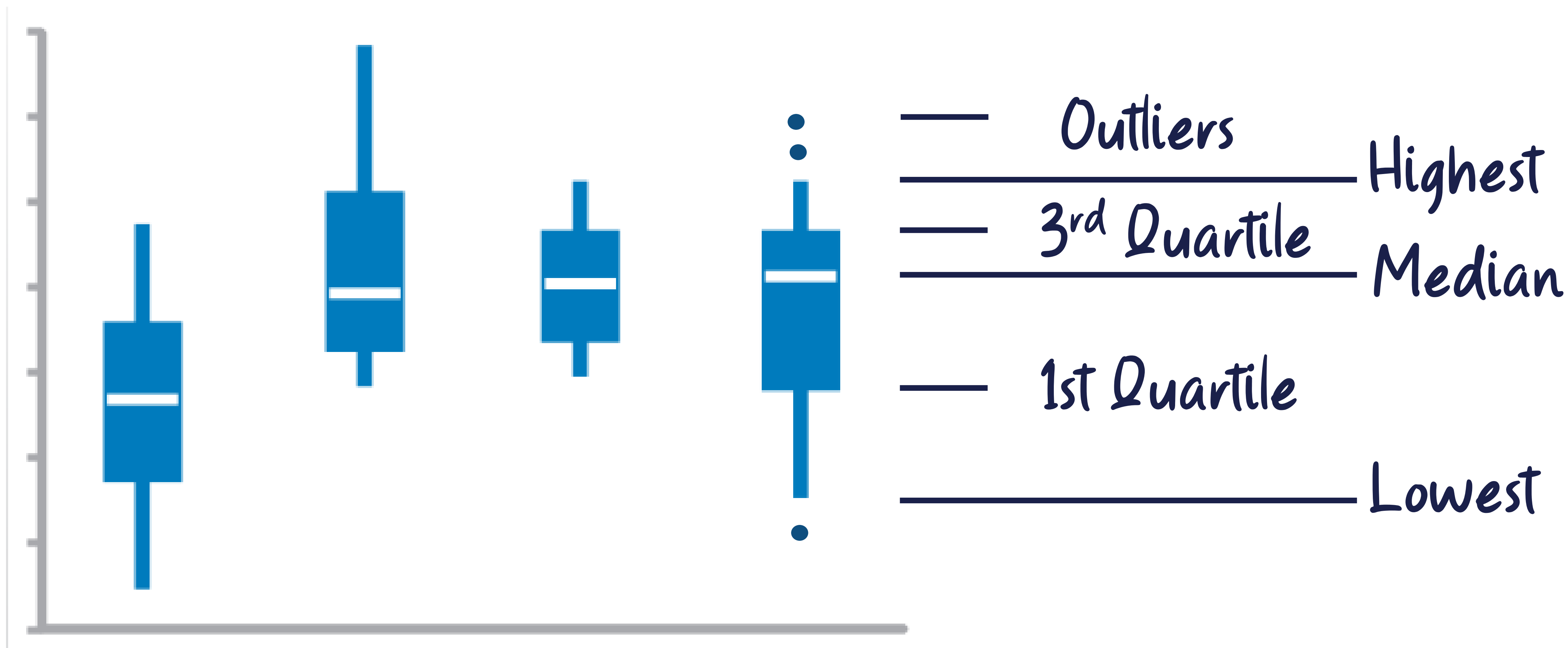
Only when also featuring distributions



Box Plot (vertical)

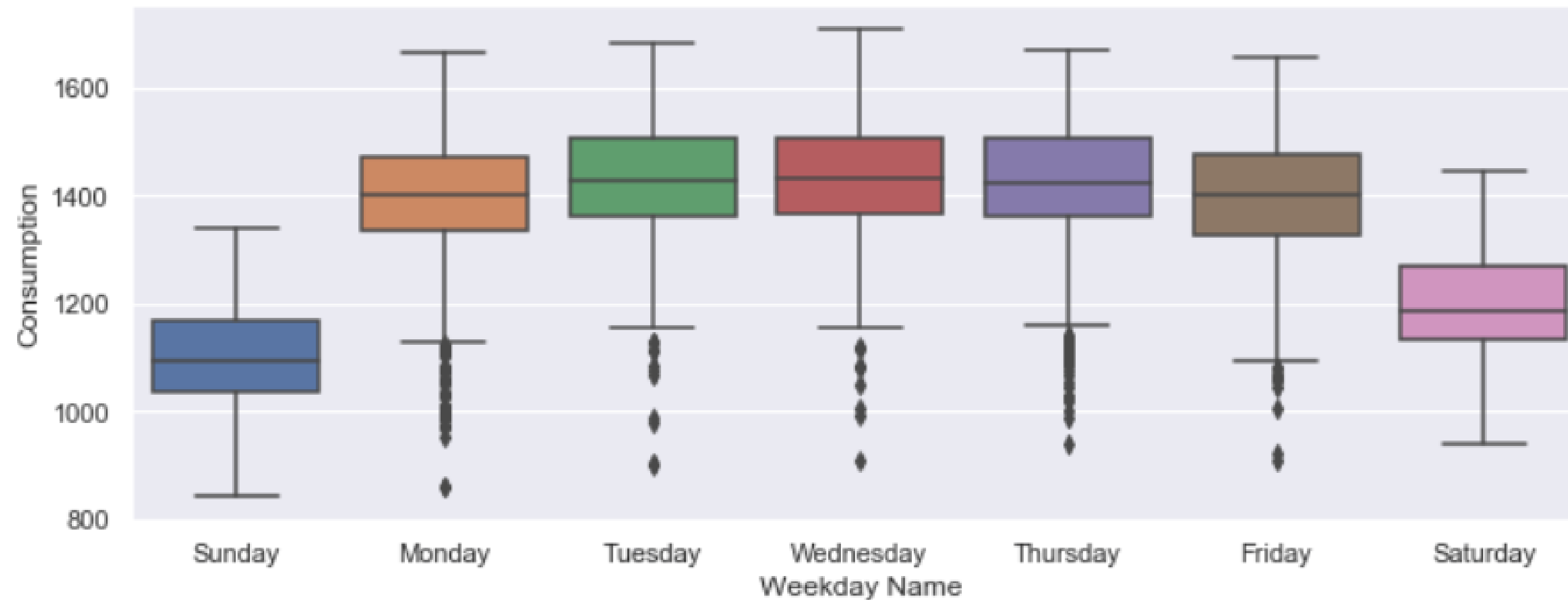
Time series

BOX PLOT (VERTICAL)



Time series

Sample BOX PLOT (VERTICAL)



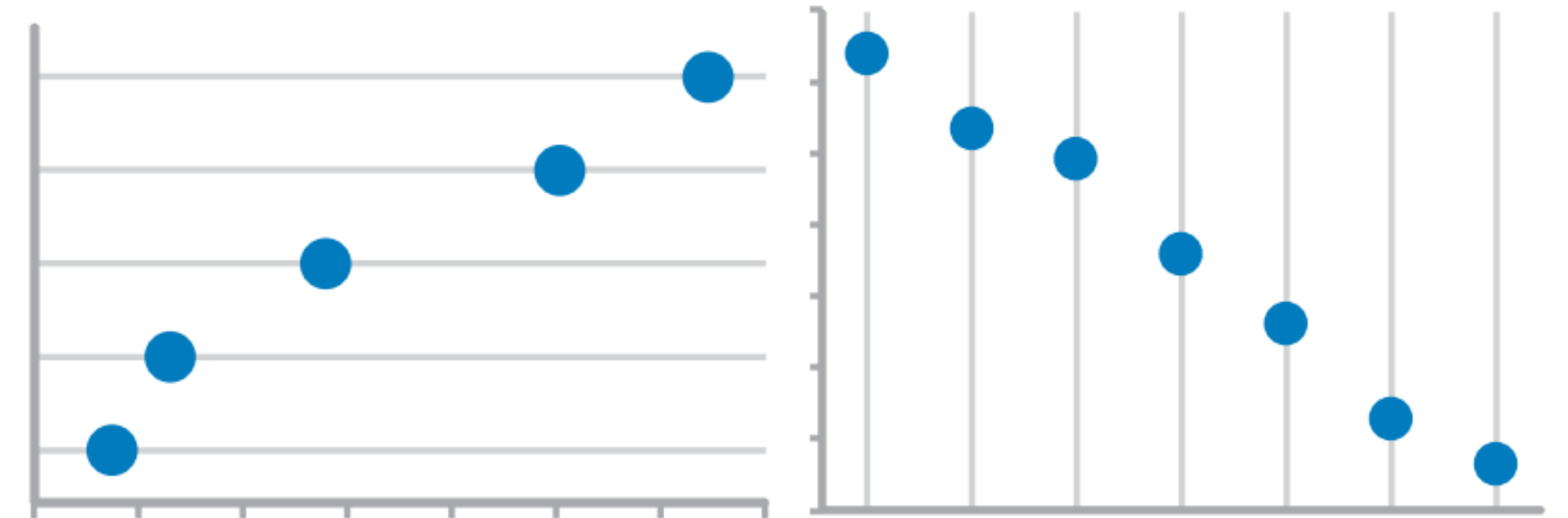
Electricity consumption per day

Ranking

values are ordered by size (descending or ascending)



Bar Graphs











Dot Plots

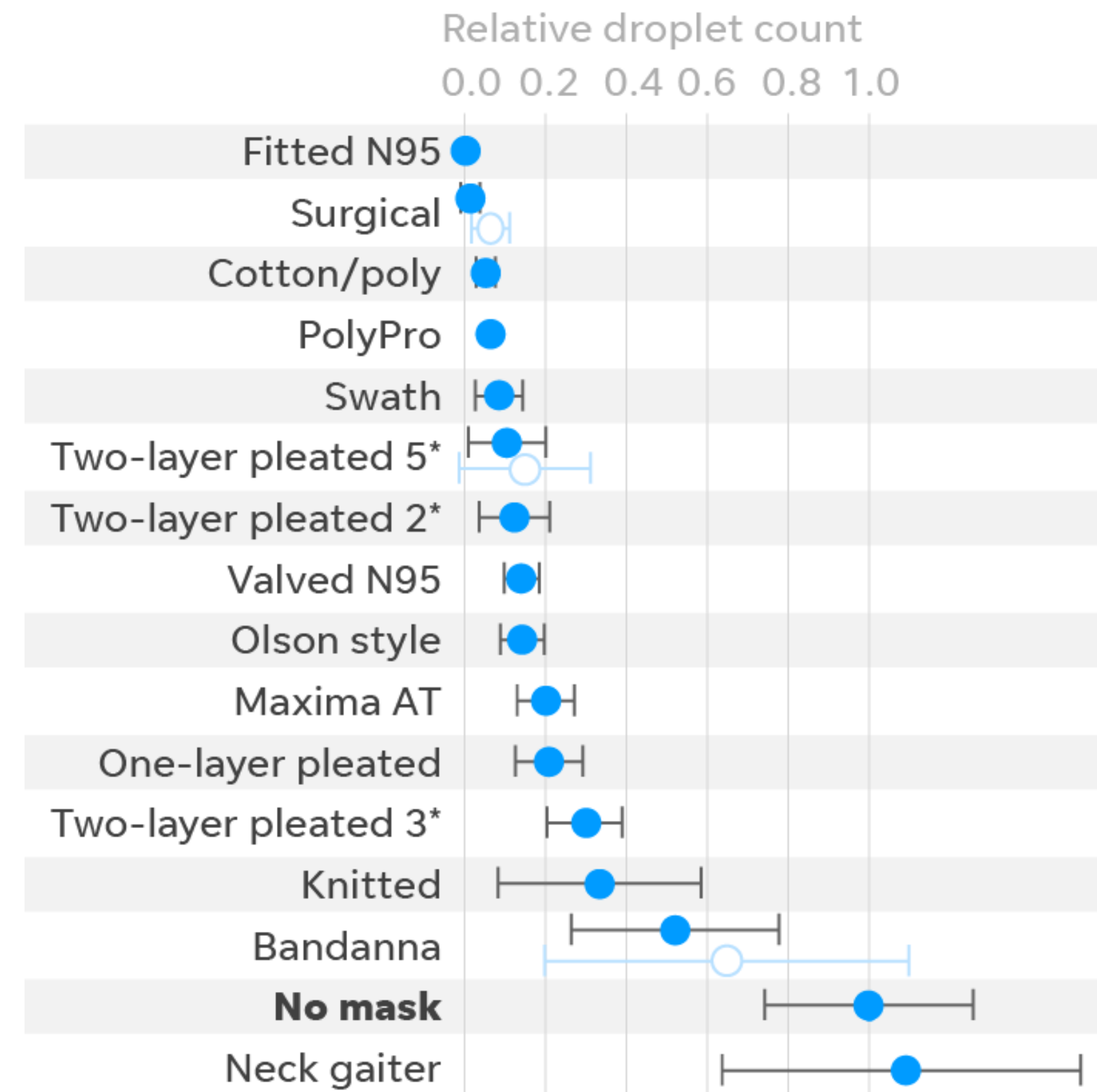
Quantitative scale must begin at zero

Ranking

values are ordered by size (descending or ascending)

Best and worst face coverings

TYPE OF COVERING	EFFICIENCY AT FILTERING LARGE DROPLETS	EFFICIENCY AT FILTERING AEROSOLS	WHERE IT CAN BE WORN
 N95 mask	99.9%	95%	Healthcare settings
 Surgical mask	98.5%	89.5%	Healthcare settings
 Hybrid mask	96%	94%	Public, indoor, and/or crowded settings
 Two-layer cotton mask	99.5%	82%	Public, indoor, and/or crowded settings
 Tea towel or dishcloth	98%	72.5%	Outdoor areas
 100% cotton t-shirt	97%	51%	Outdoor areas
 Natural silk	56%	54%	Outdoor areas
 Scarf or bandana	44%	49%	As a last resort

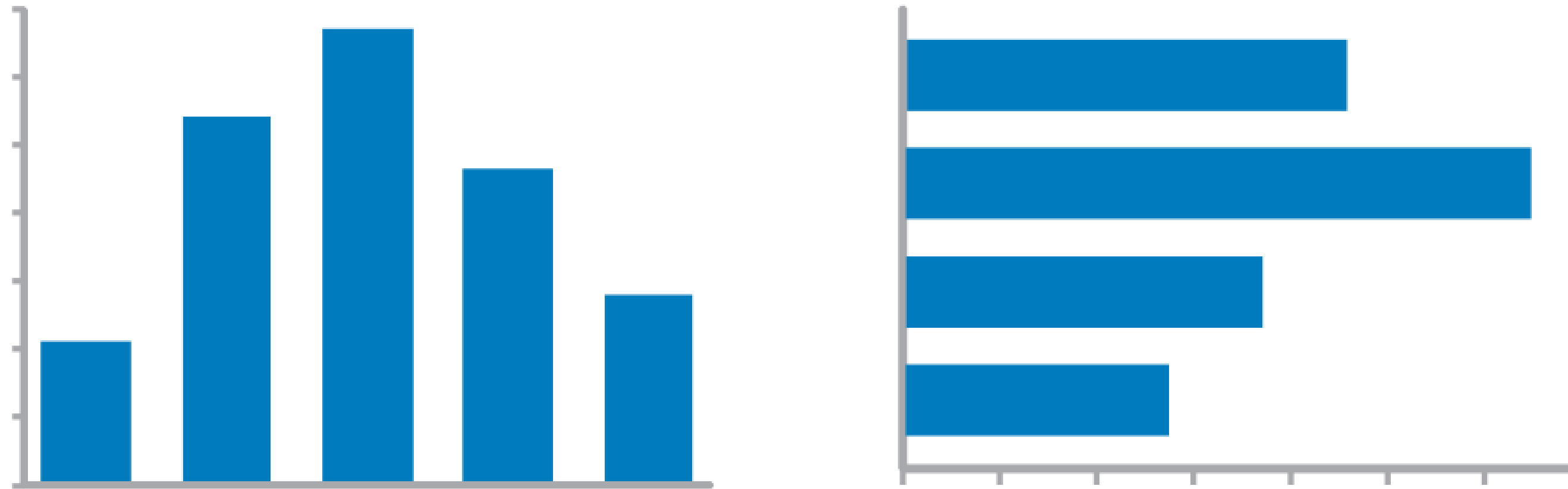


* Three different two-layer pleated masks were tested

NOTE Results are normalized to the control trial (no mask)

Part-to-whole

values represent parts (ratios) of a whole

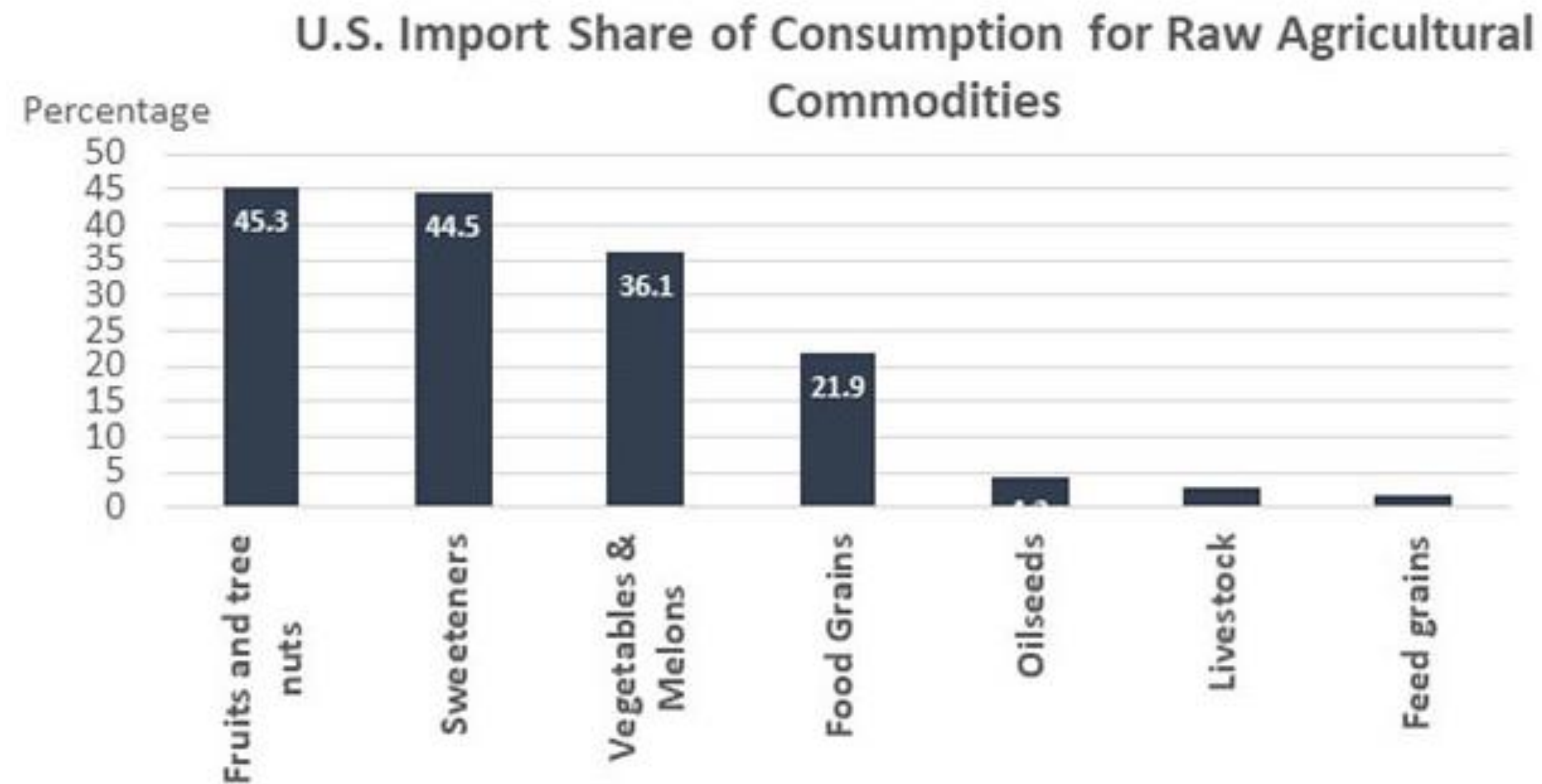


Bar Graphs

Quantitative scale must begin at zero

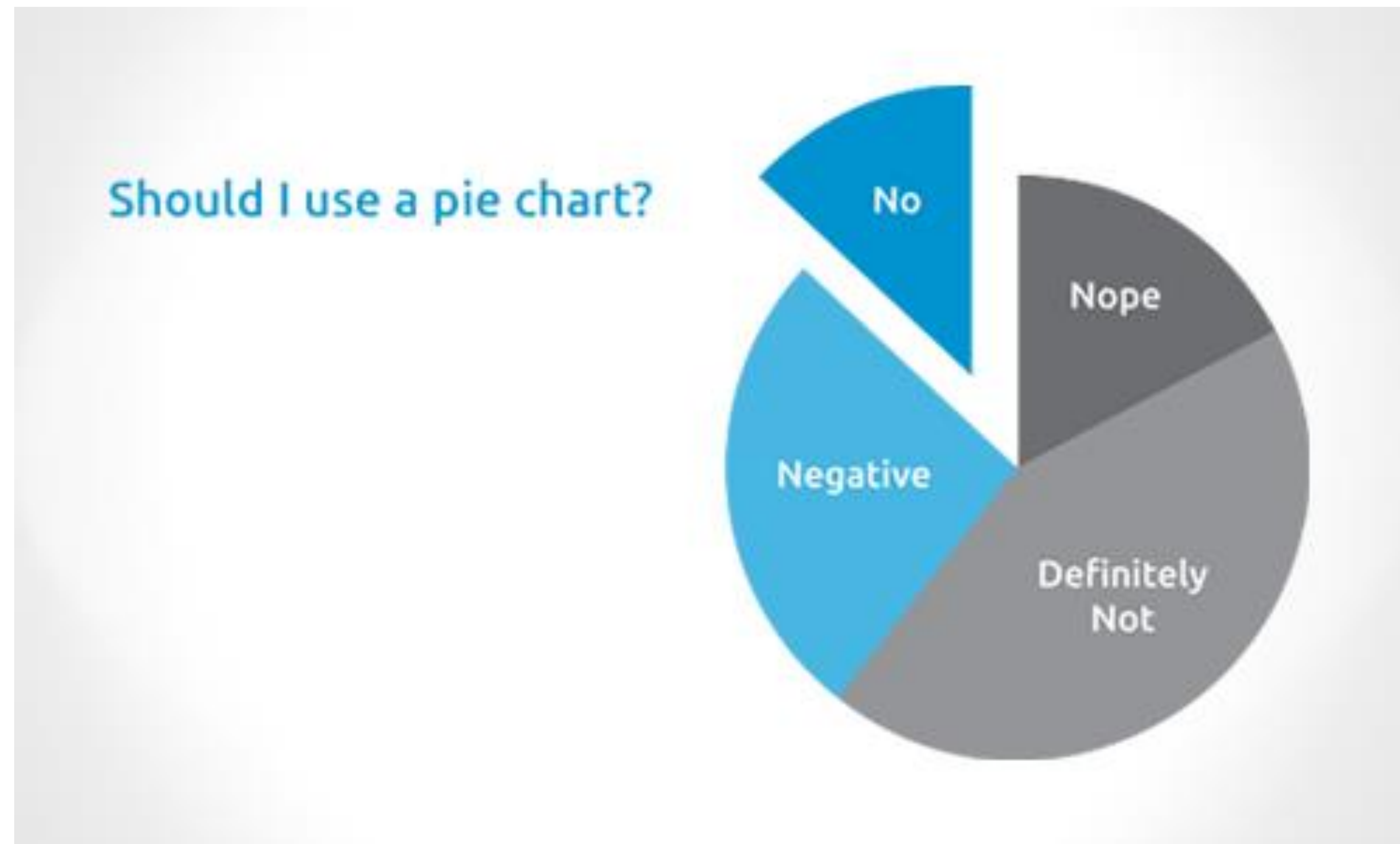
Part-to-whole

values represent parts (ratios) of a whole



Part-to-whole

values represent parts (ratios) of a whole

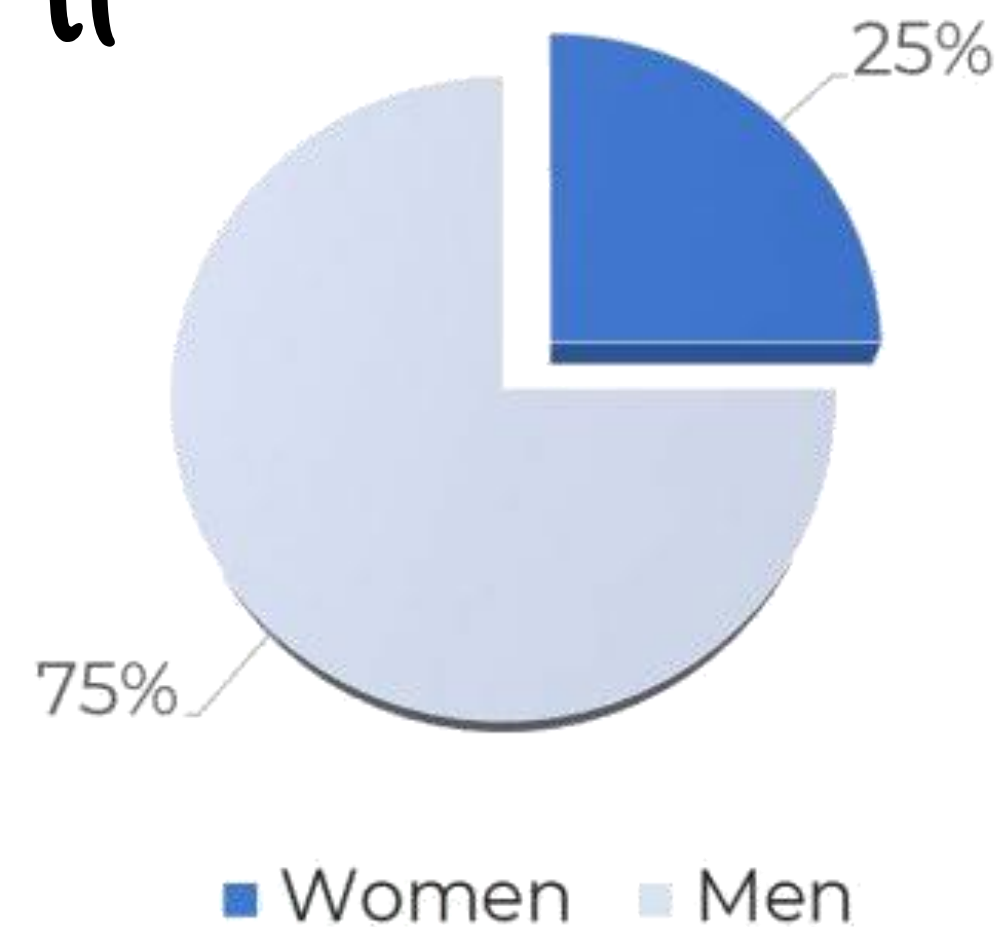


Part-to-whole

values represent parts (ratios) of a whole

Use Pie Charts Only if...

1. You have Binary Data
2. You can format it

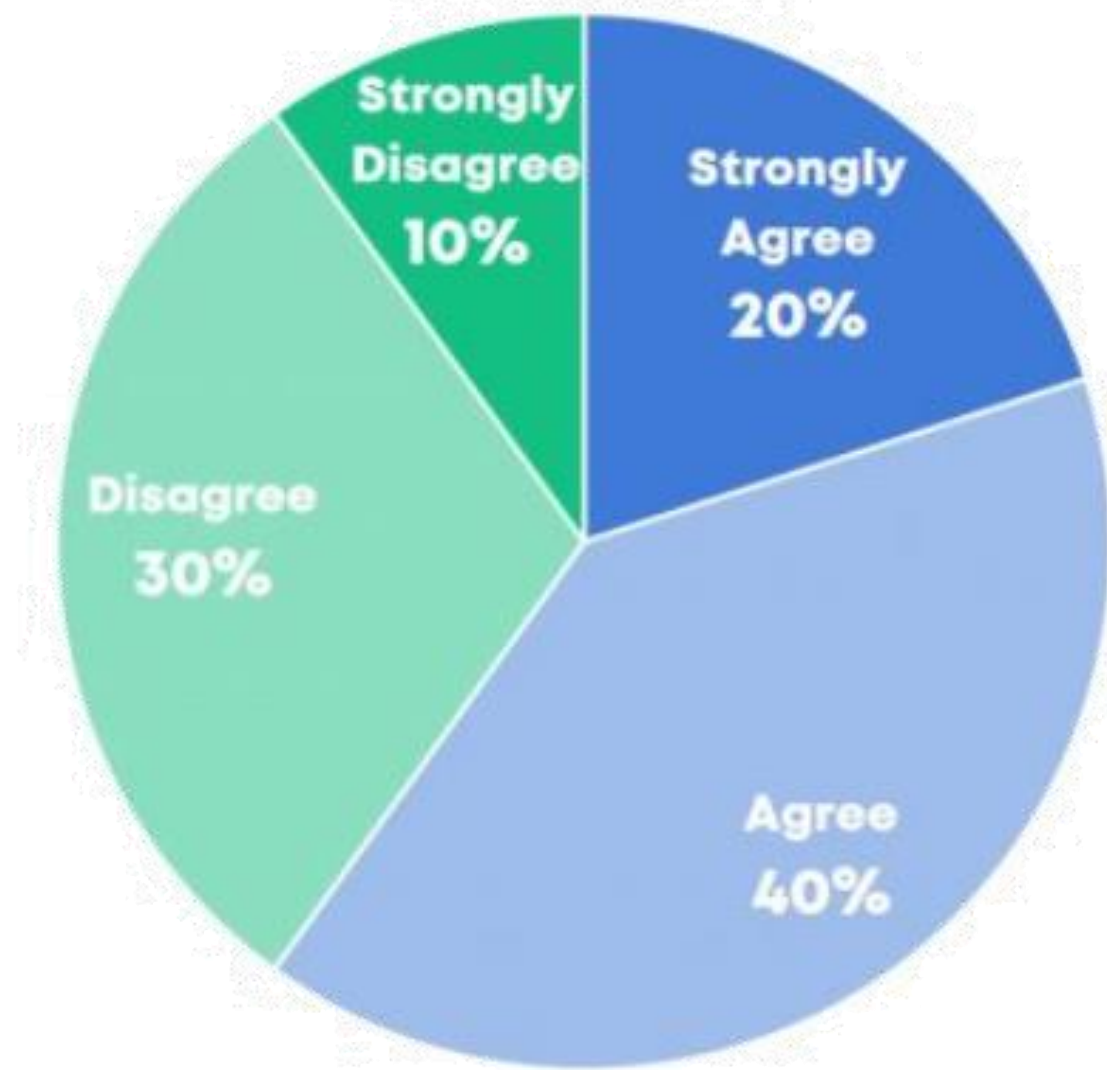


Part-to-whole

values represent parts (ratios) of a whole

When not to use Pie Charts

1. You have Ordinal Data

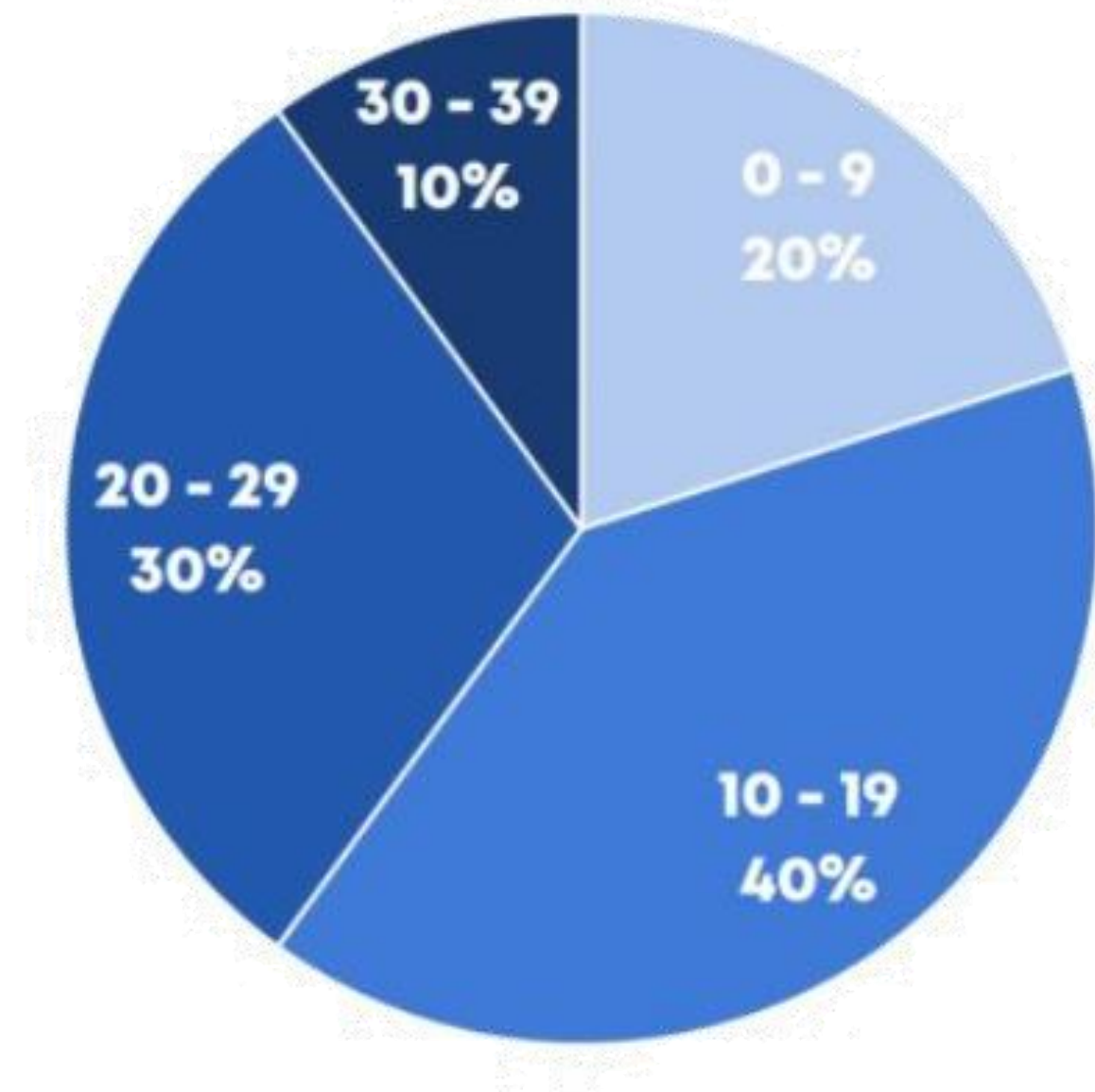


Part-to-whole

values represent parts (ratios) of a whole

When not to use Pie Charts

1. You have Ordinal Data

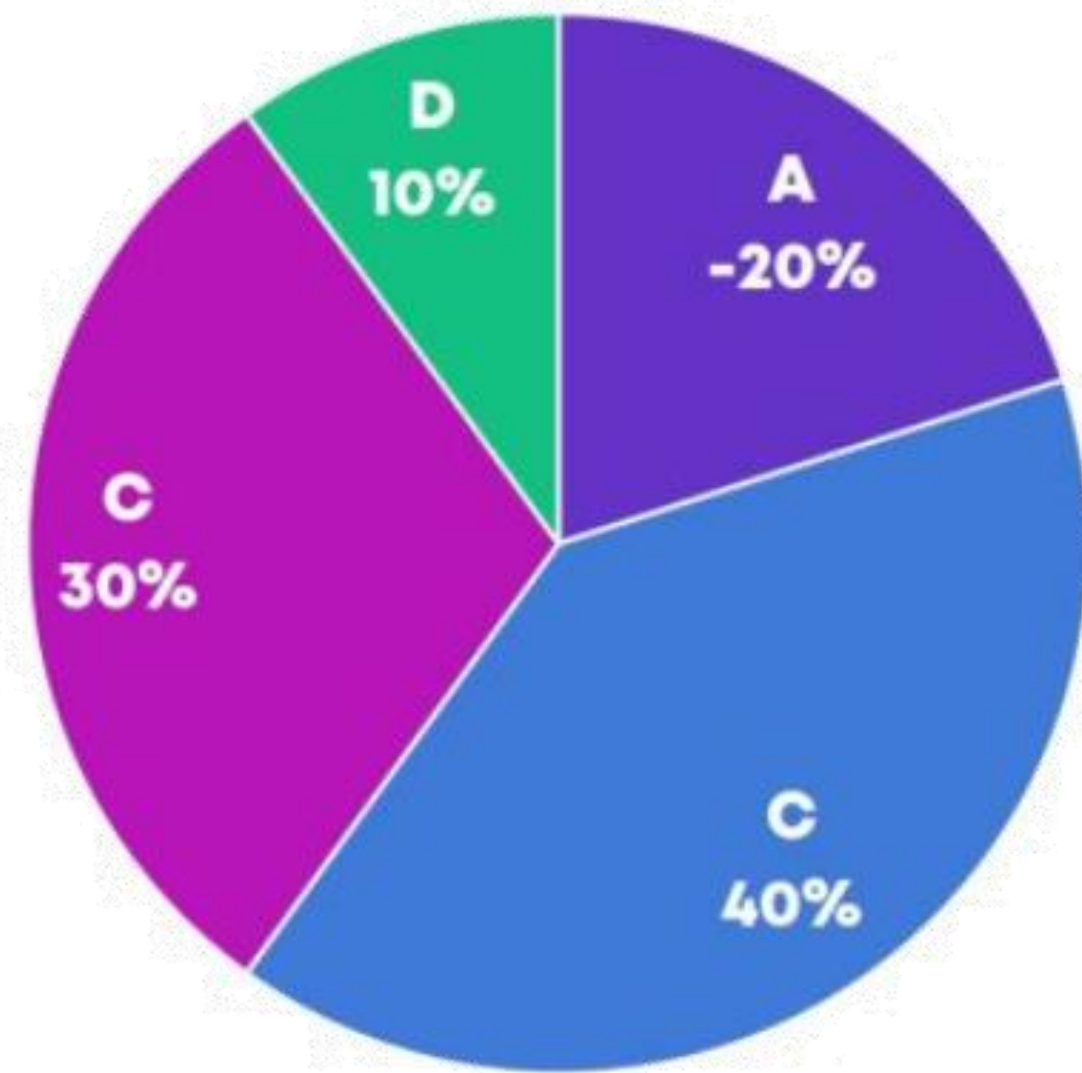


Part-to-whole

values represent parts (ratios) of a whole

When not to use Pie Charts

2. You have a Negative Number

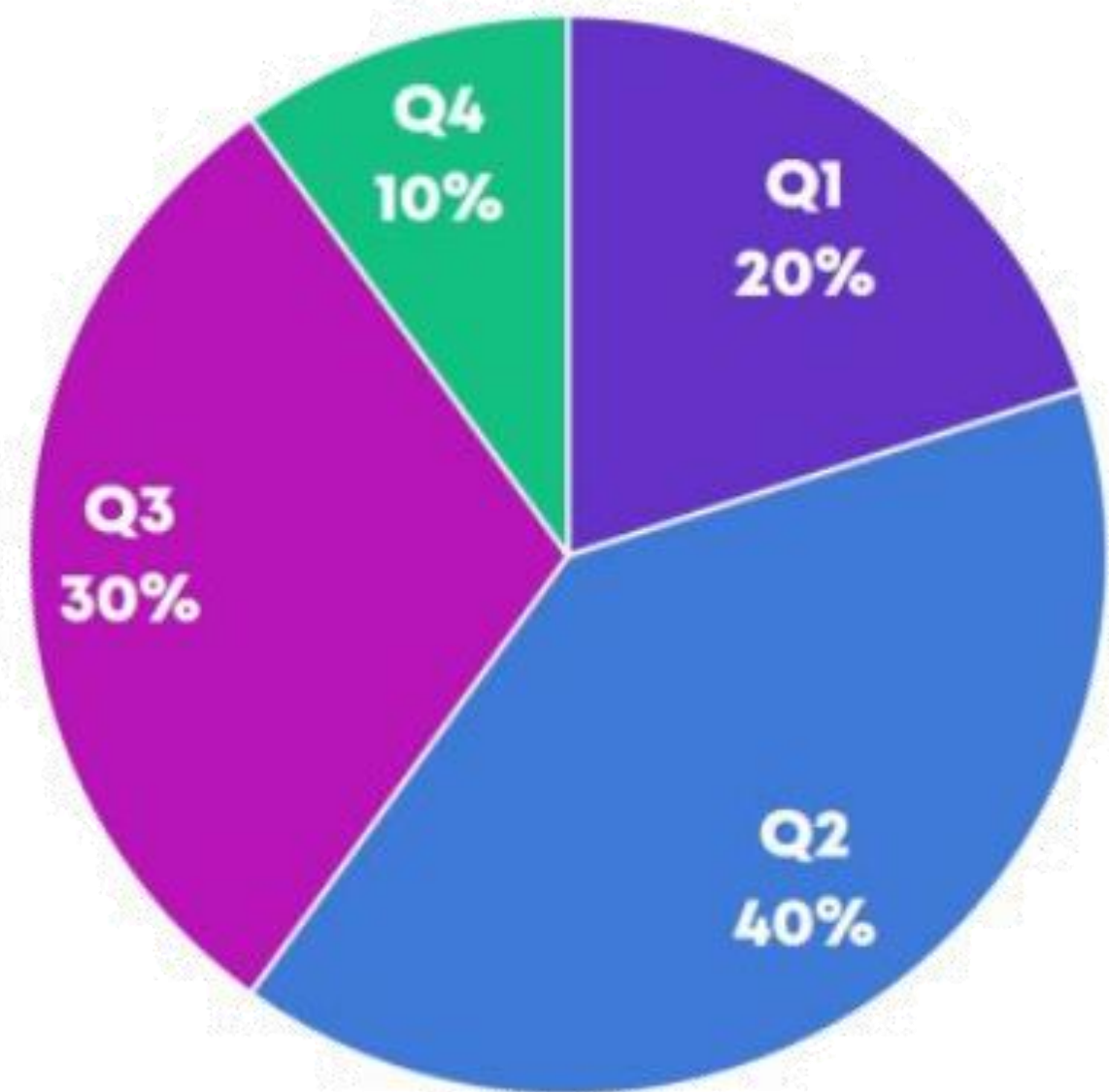


Part-to-whole

values represent parts (ratios) of a whole

When not to use Pie Charts

3. You can form a Pattern



Deviation

difference between two sets of values



Bar Graphs

Quantitative scale must be at zero

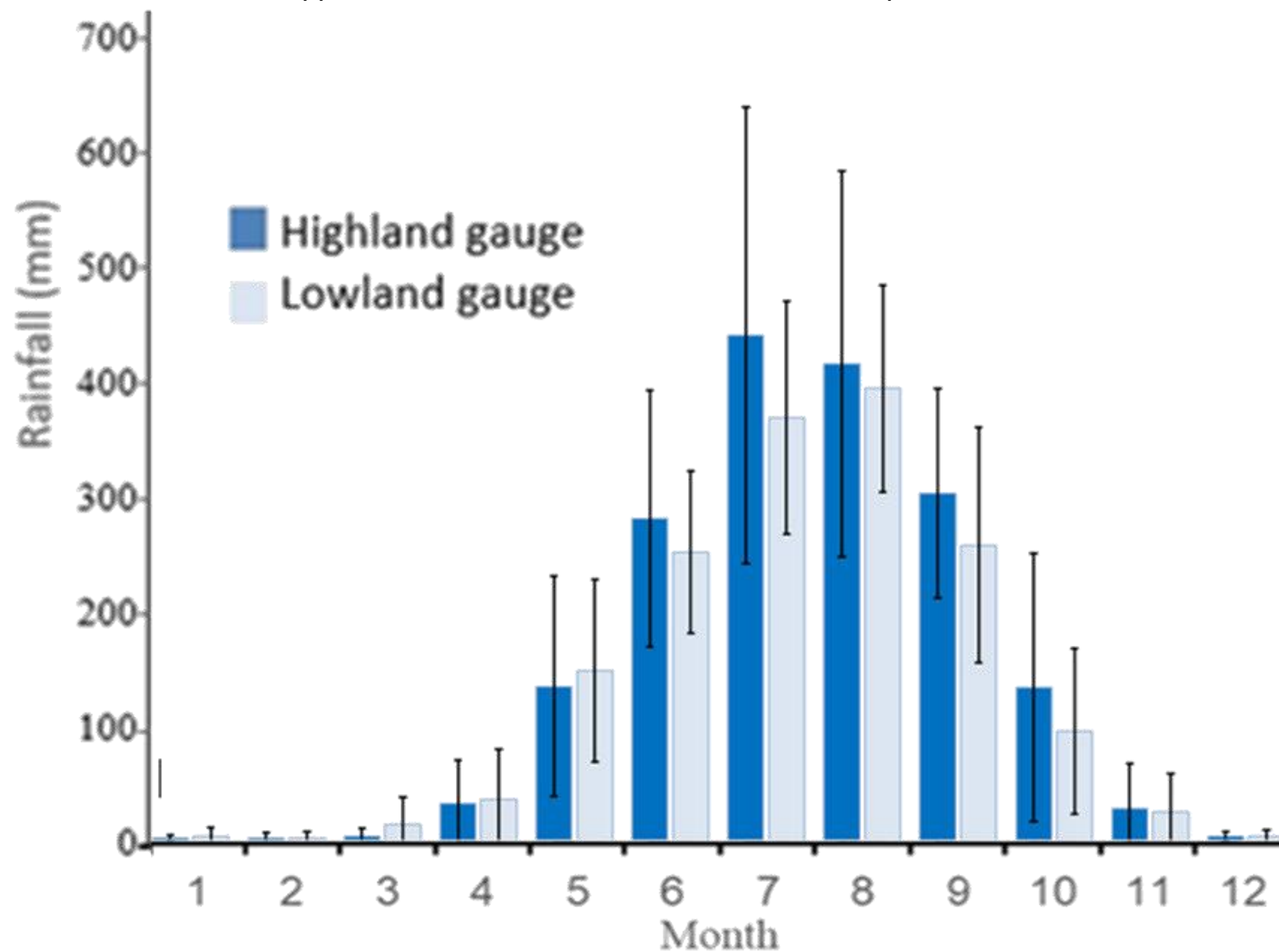


Line Graph

Only when also featuring time series or
single distribution

Deviation

difference between two sets of values



Distribution

count of values per interval along quantitative scale



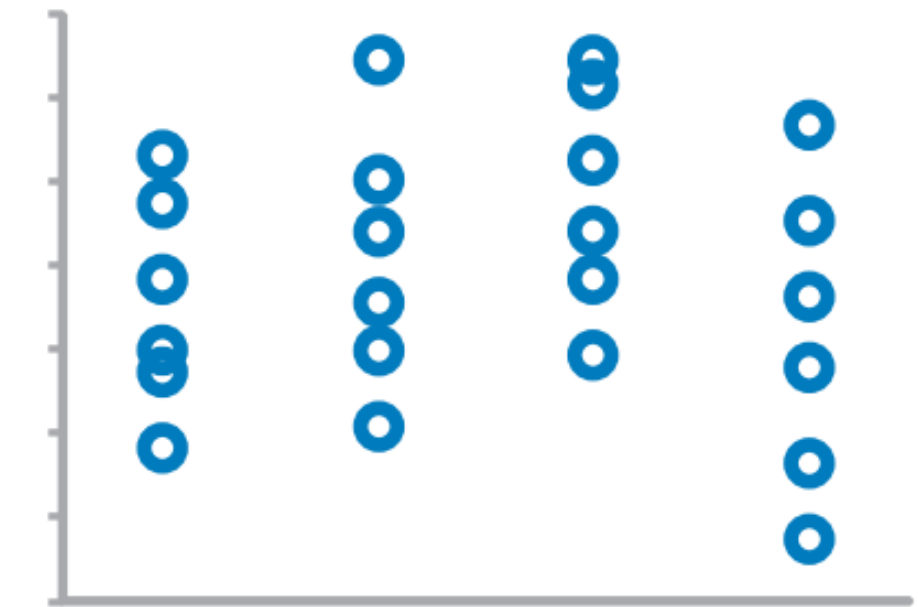
Bar Graphs

Quantitative Scale, must begin at zero



Strip Plot (single)

When you want to see each value



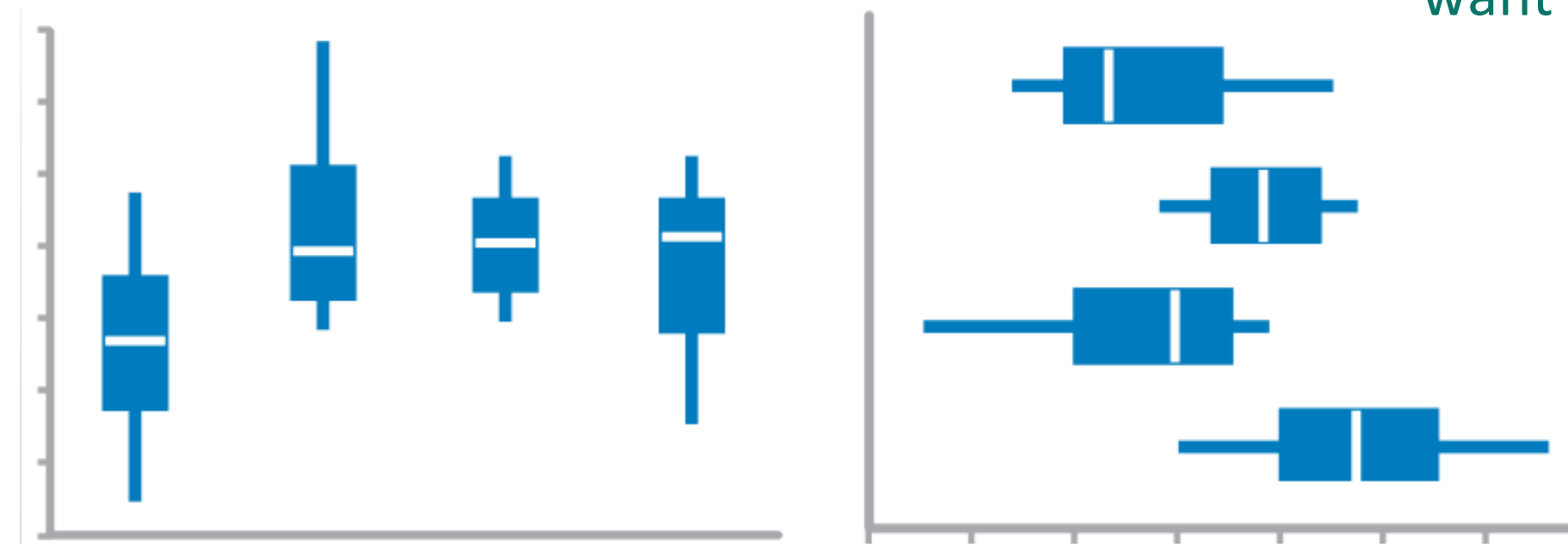
Strip Plot (multiple)

When comparing multiple distributions AND you want to see each value



Line Graph

To feature overall shape of distribution

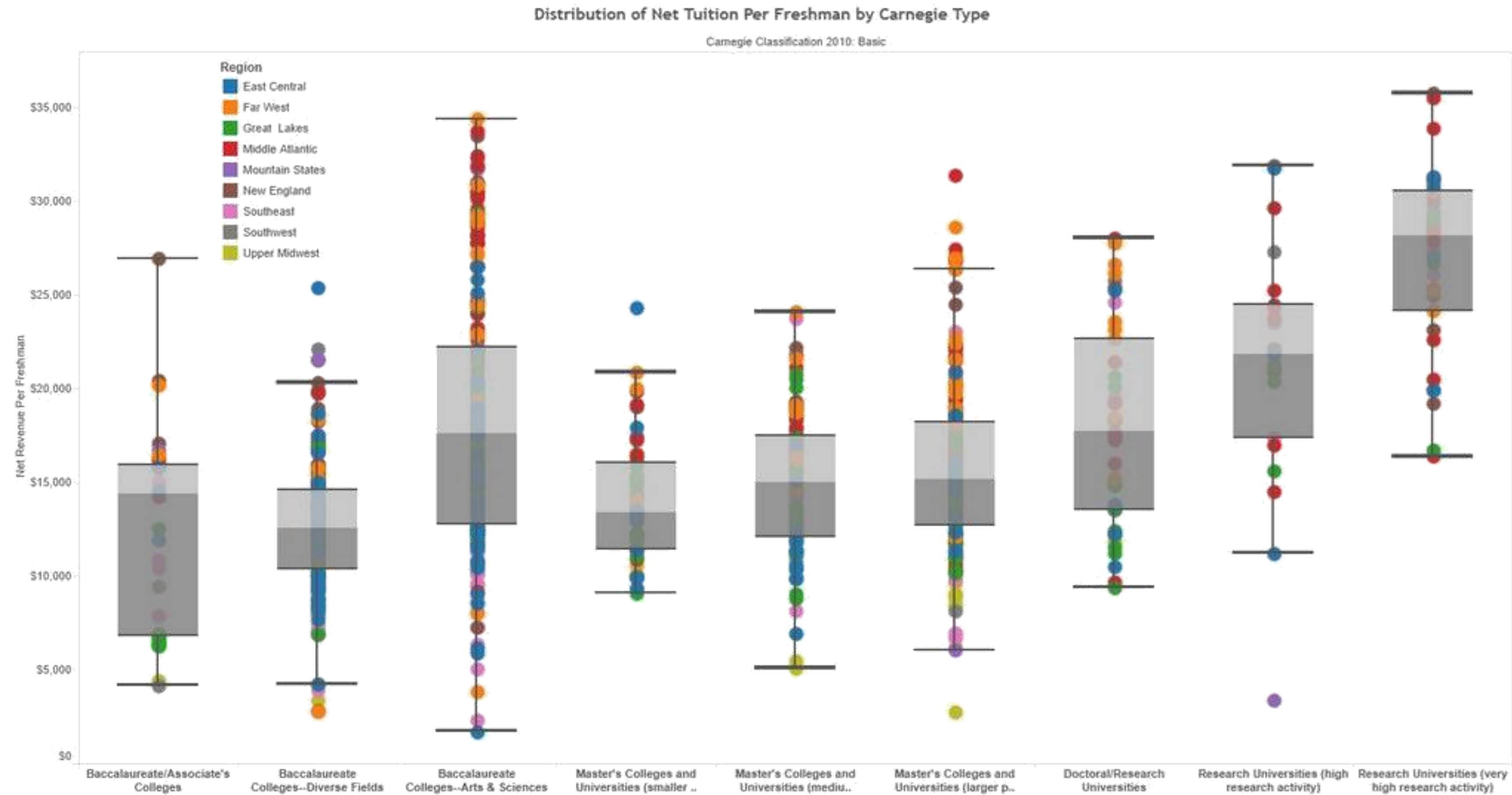


Box Plots

When Comparing Multiple Distributions

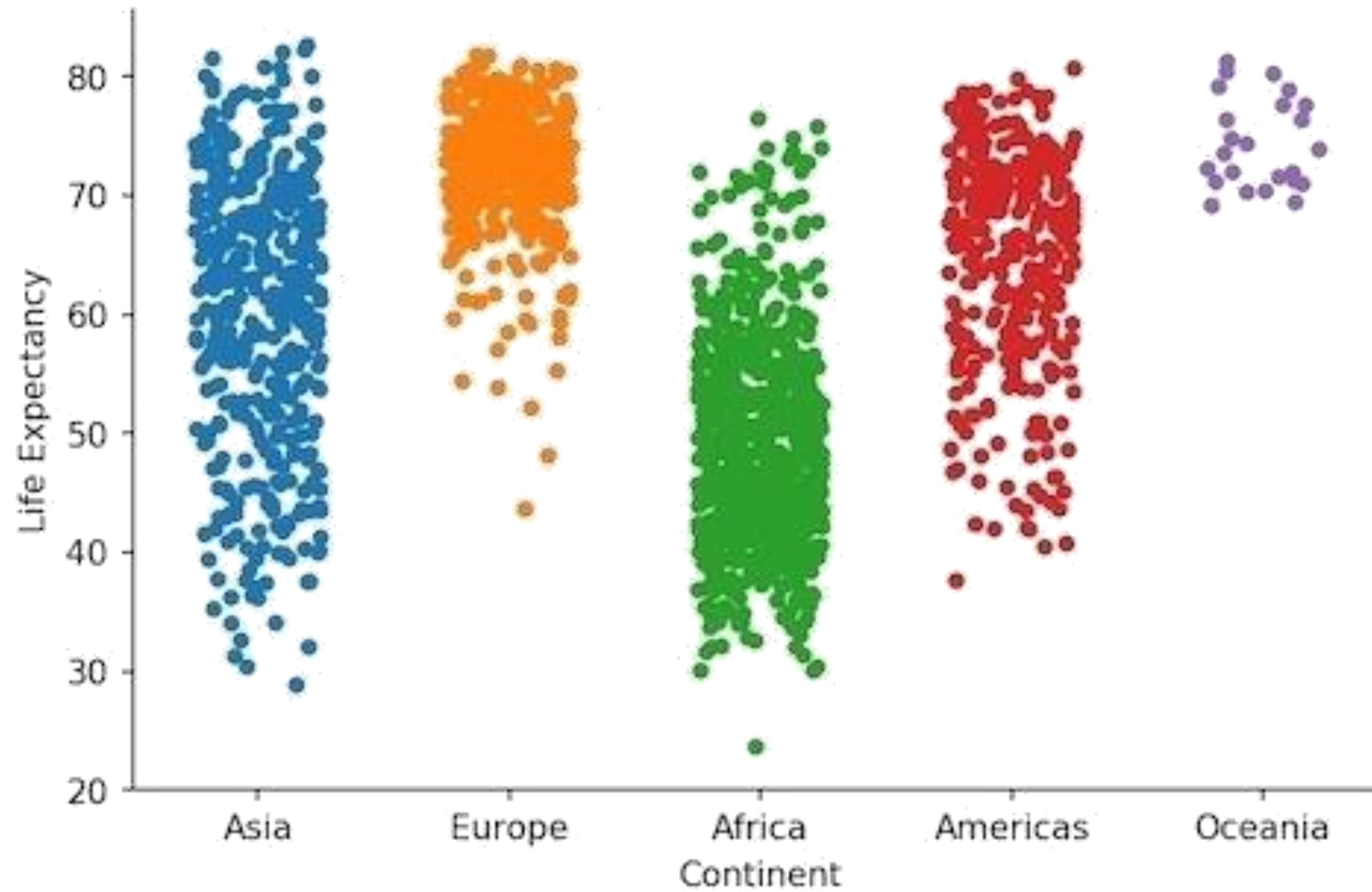
Distribution

count of values per interval along quantitative scale



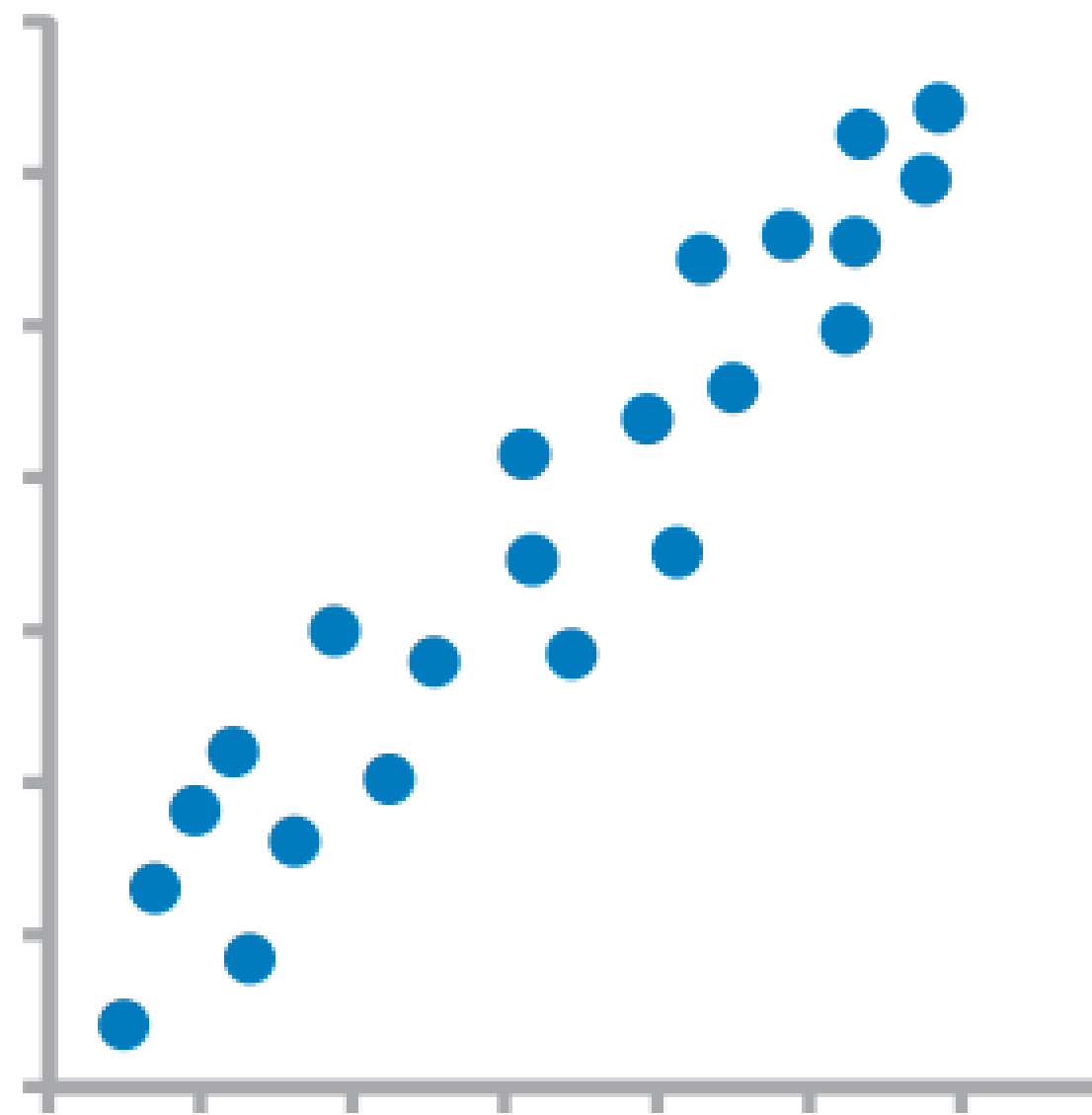
Distribution

count of values per interval along quantitative scale



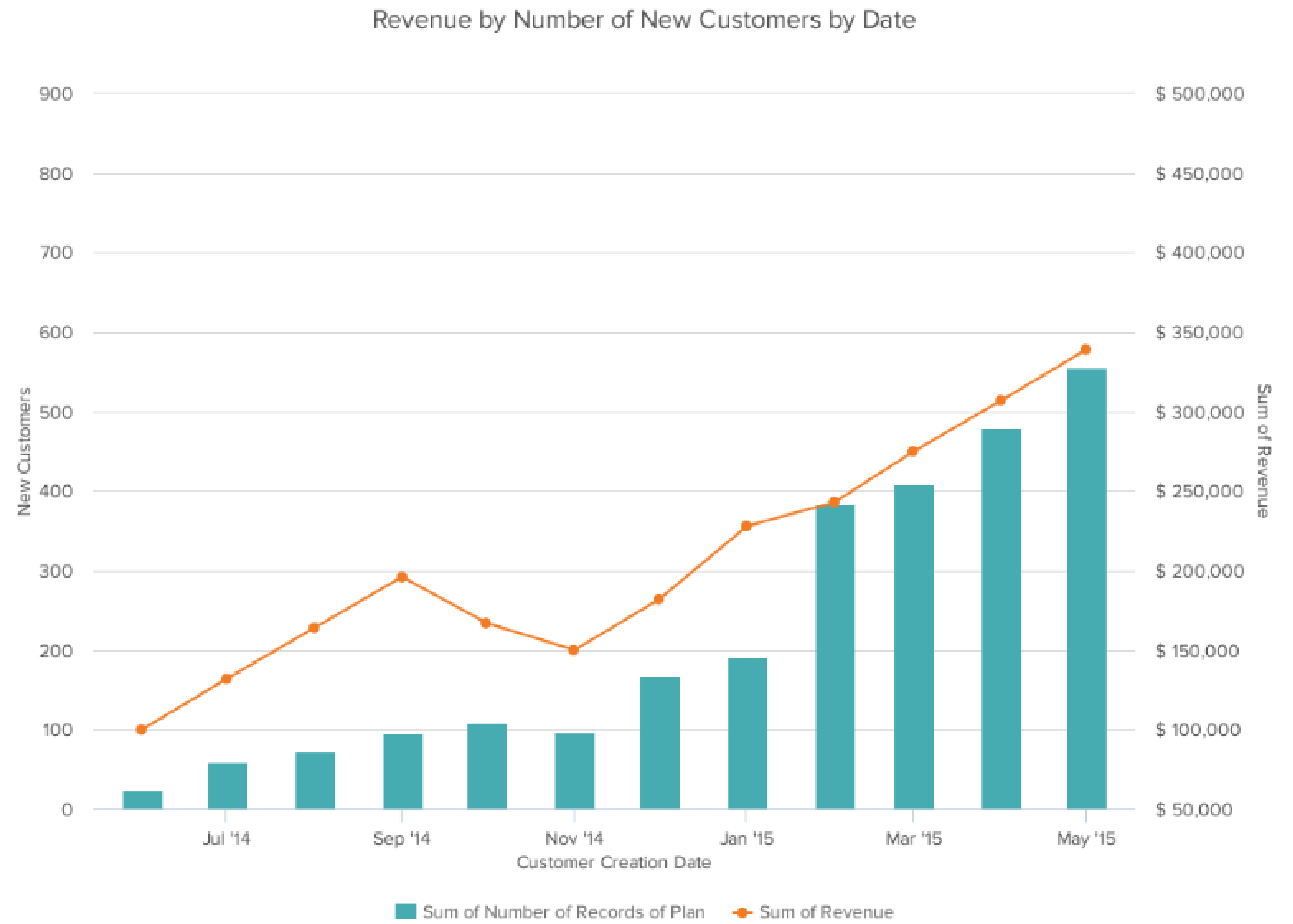
Correlation

Comparison of two paired sets of values to determine if there is a relationship between them



Scatter Plot

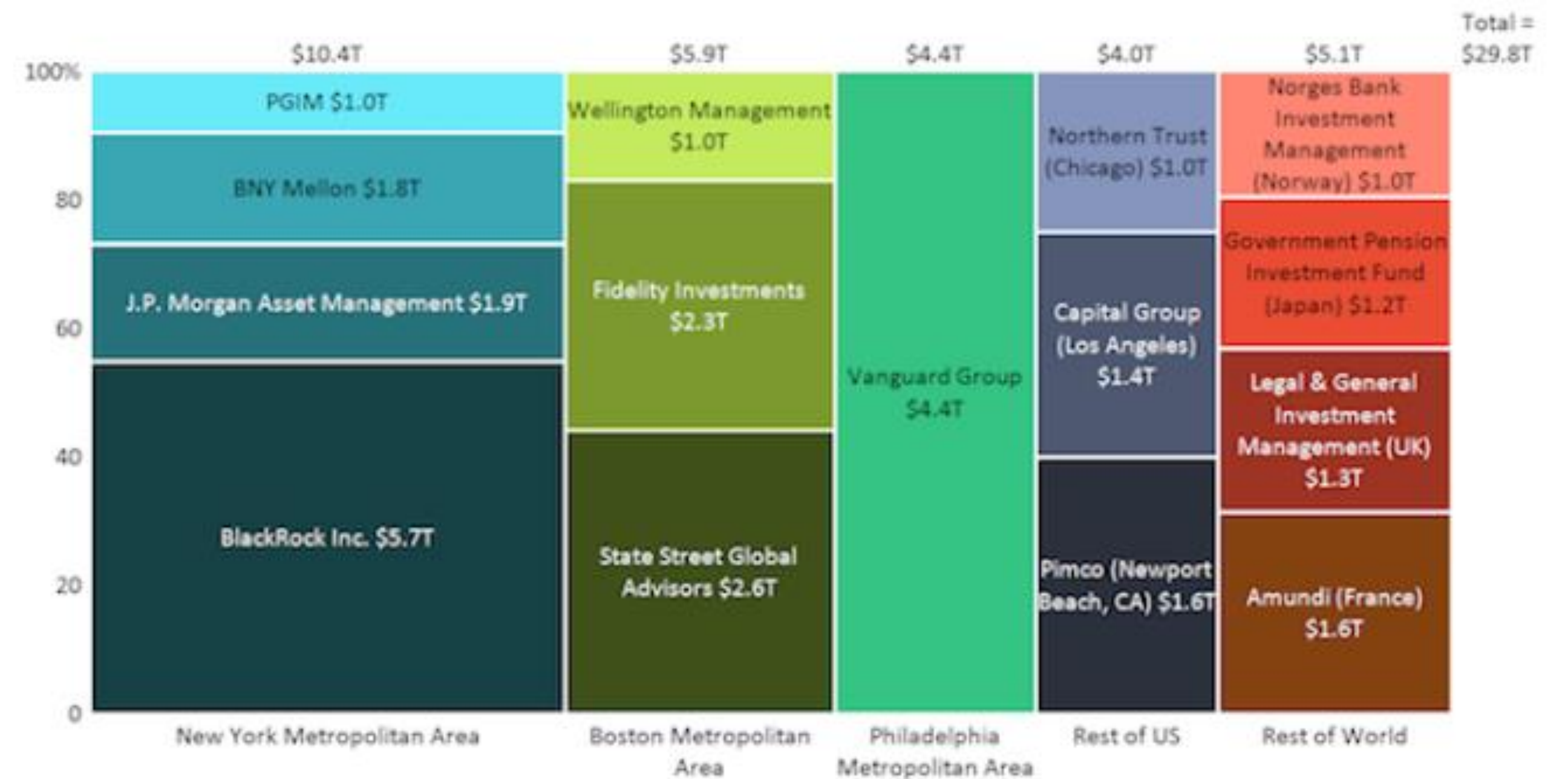
Examples



Examples

World's Largest Asset Managers

Most of the world's largest asset managers are grouped in the Northeast US. Eight of the 14 firms that manage \$1T or more are in the NY, Boston or Philadelphia areas.

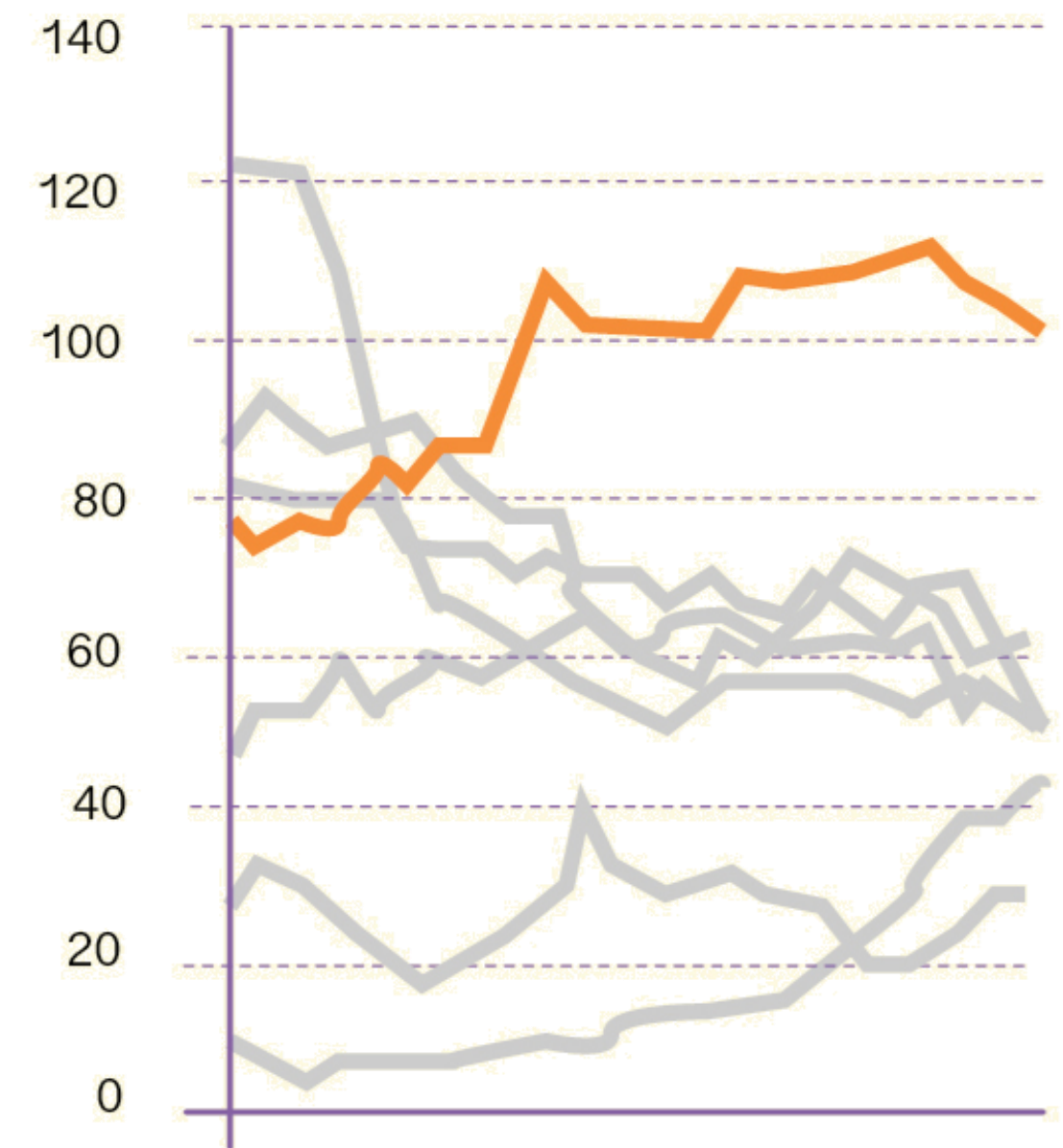
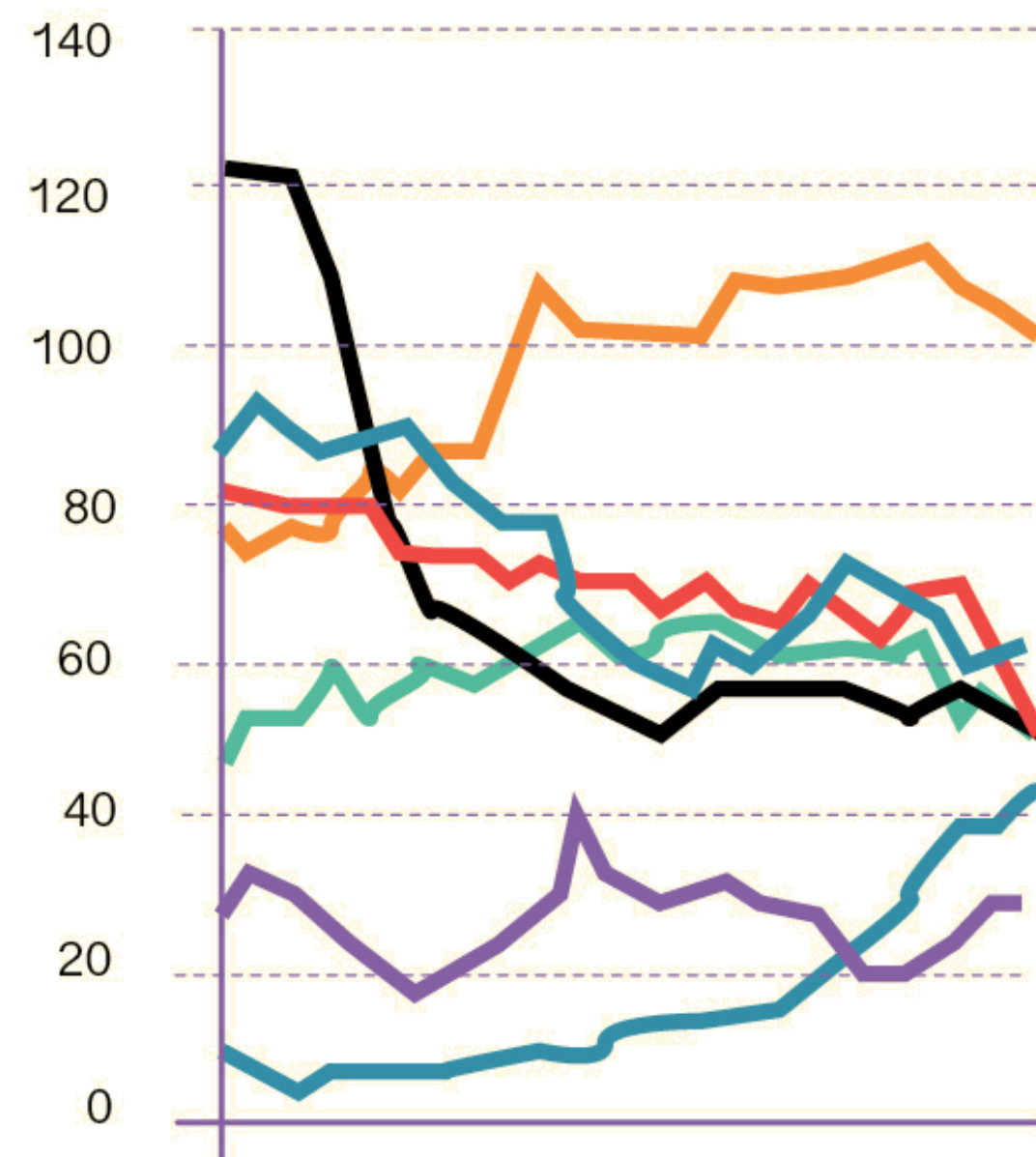


Examples



Best Practices

Emphasize most important data



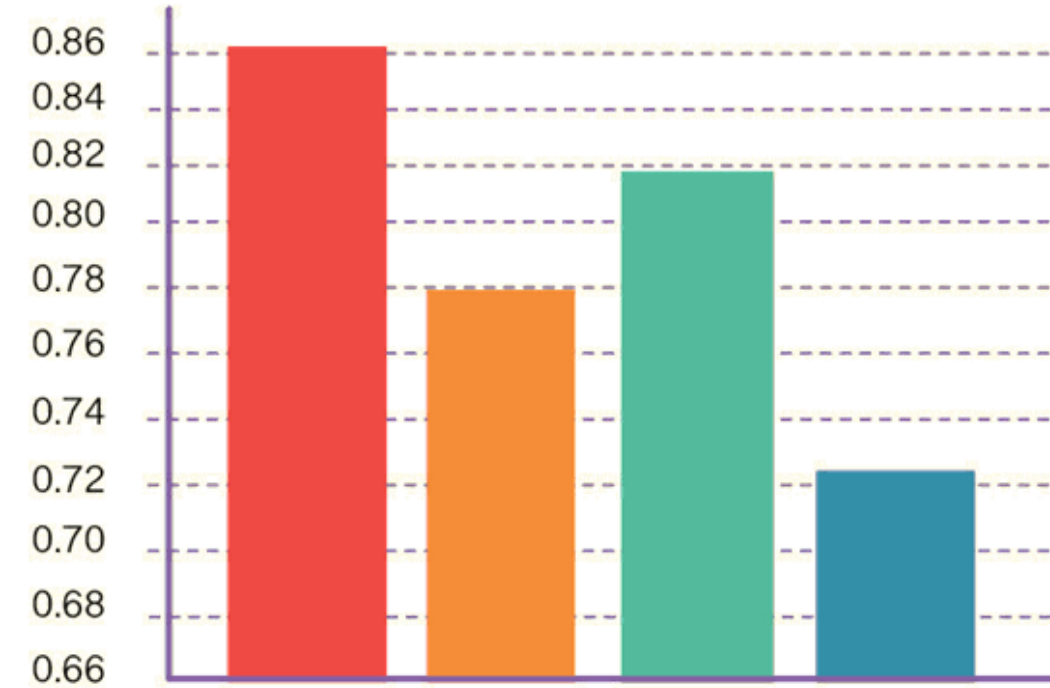
Best Practices

Emphasize

most important data

Orient

graphs for legibility



Best Practices

Emphasize

most important data

Orient

graphs for legibility

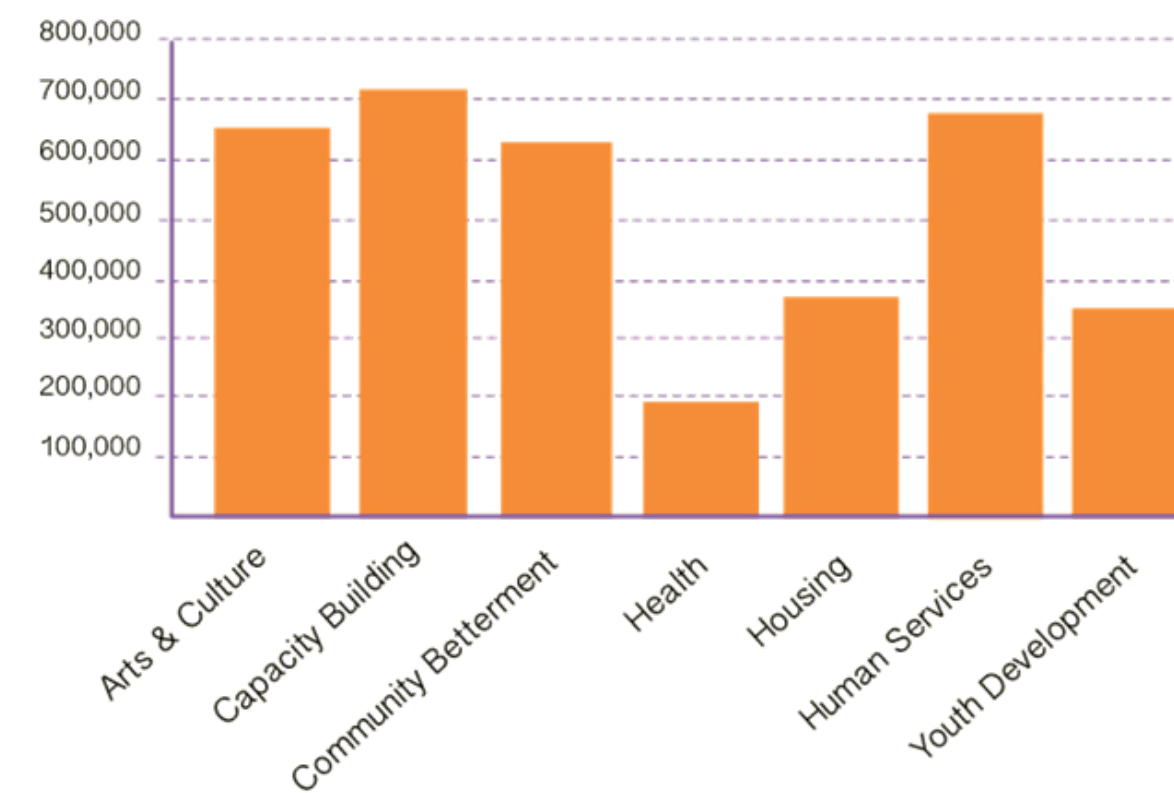
Organize

graph/table



Investment by area of impact

2006-Present



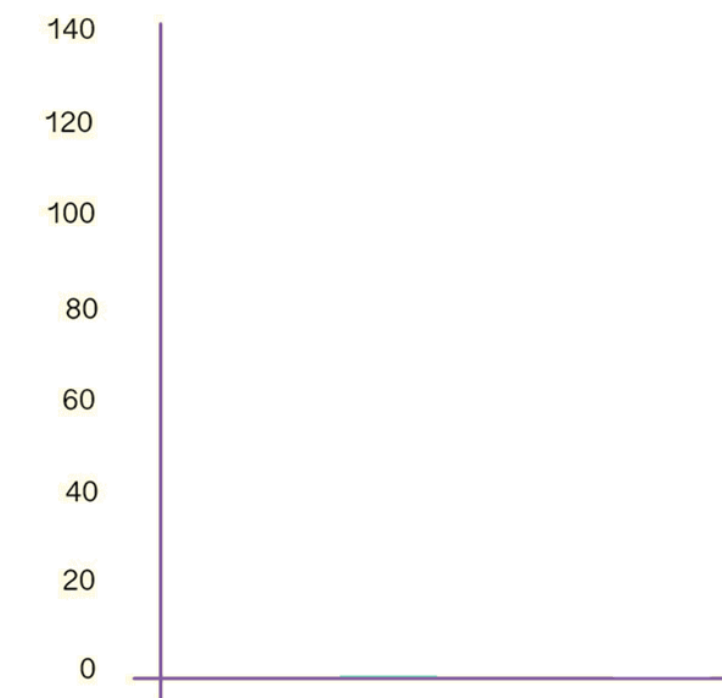
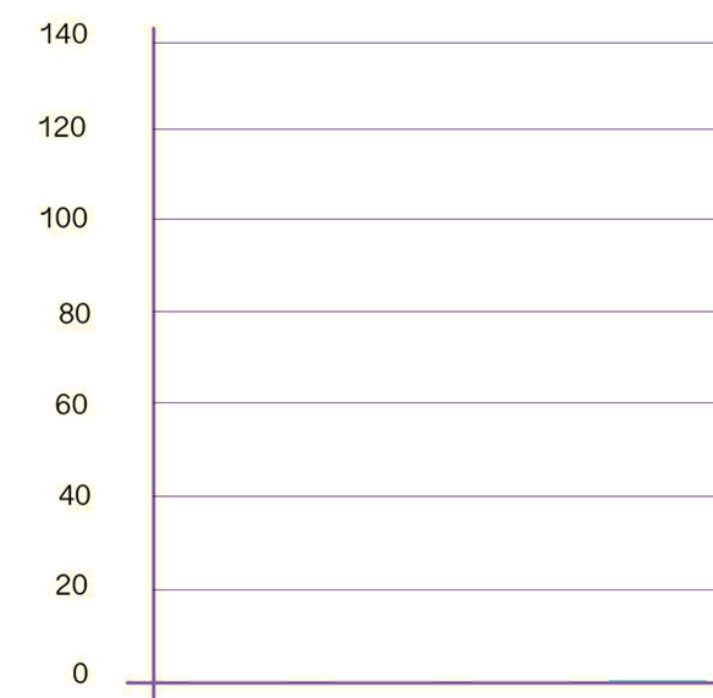
Investment by area of impact

2006-Present / Dollars in '000s



Best Practices

Emphasize	most important data
Orient	graphs for legibility
Organize	graph/table
Avoid	overloading graphs



“The greatest value of a picture is when it forces us to notice what we never expected to see.”

- John Tukey, 1977





DataViz 101