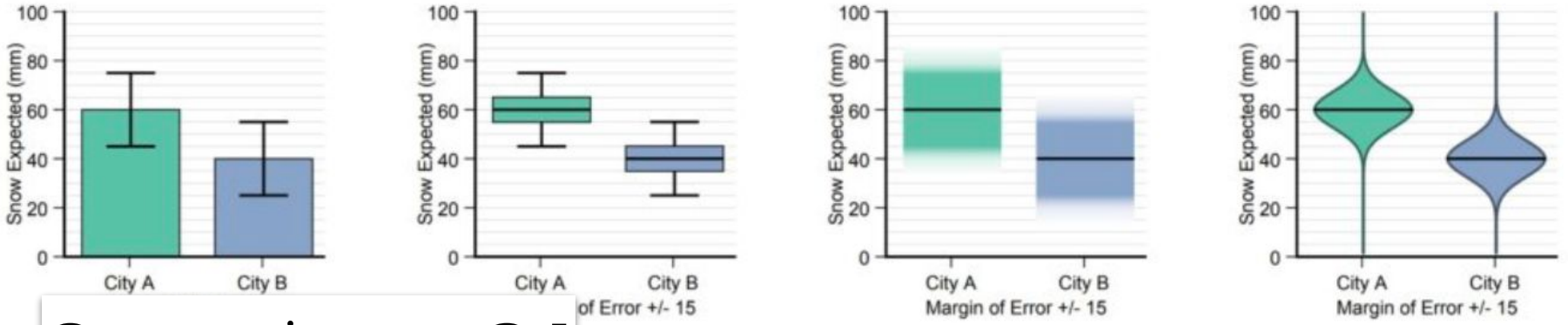
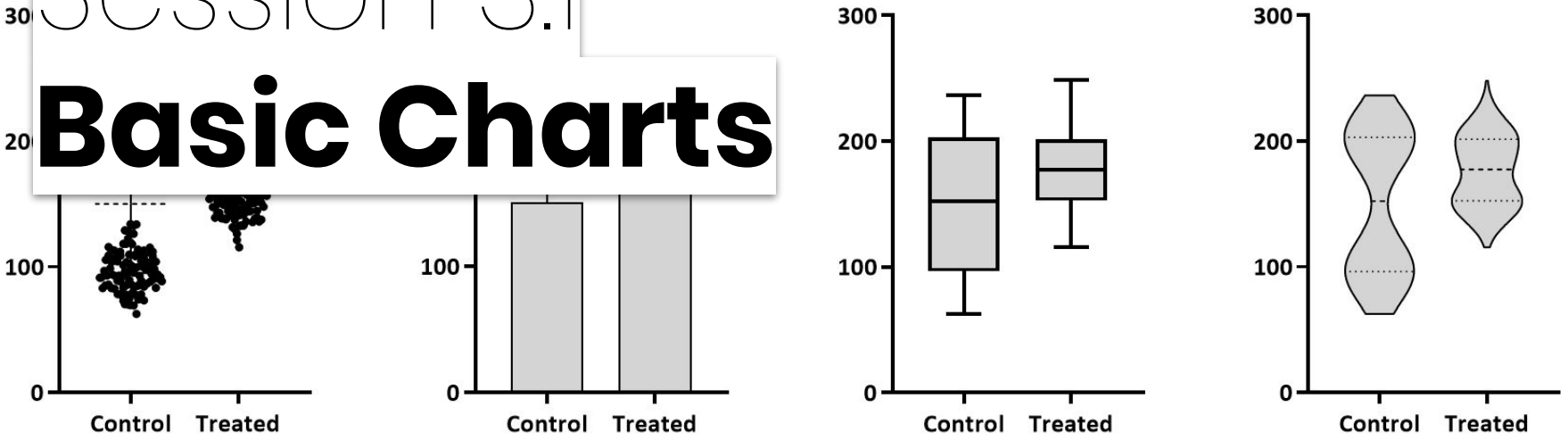


Session 3.1

Basic Charts



Online Course
Data Visualization for Professionals

THE UNIVERSITY of EDINBURGH

Benjamin Bach

June 2020

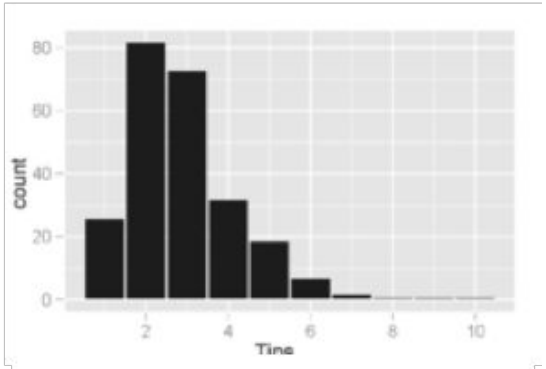
http://benjbach.me

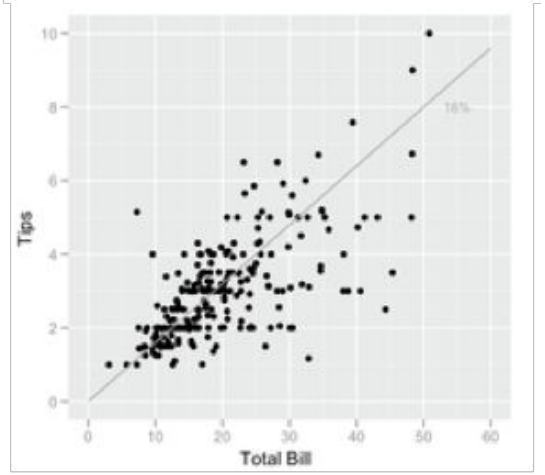
https://datavis-online.github.io

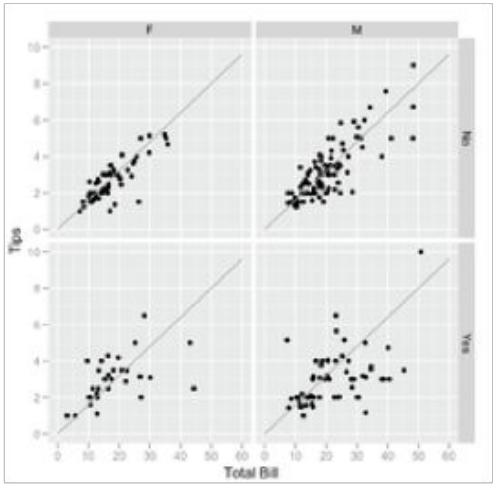
-- Not for external use --

John W. Tukey

EXPLORATORY DATA ANALYSIS

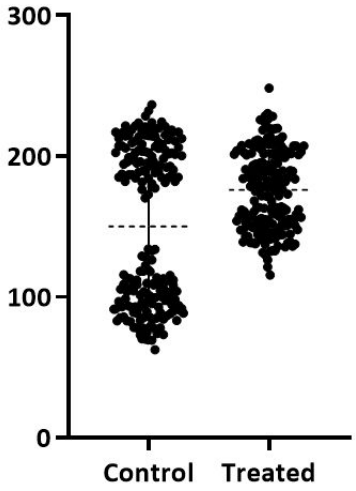




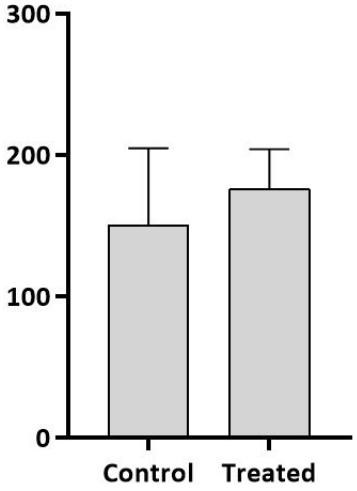


Basic Charts

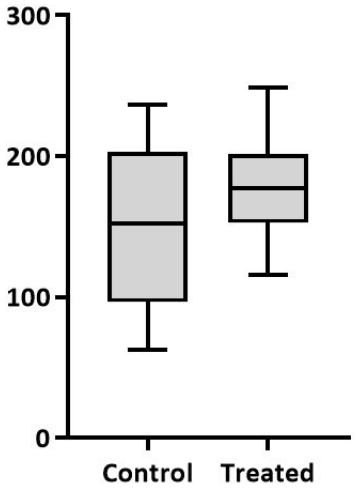
?



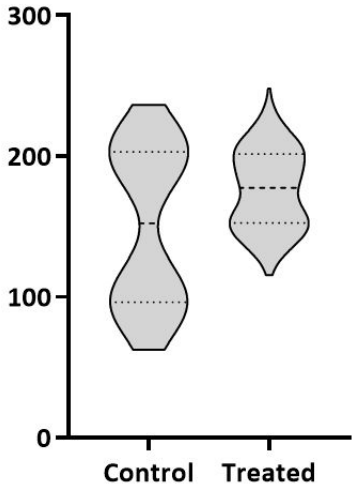
?



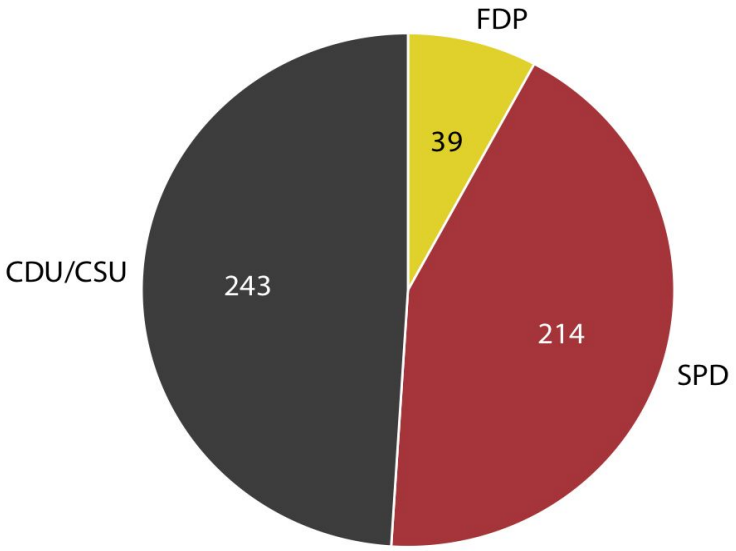
?

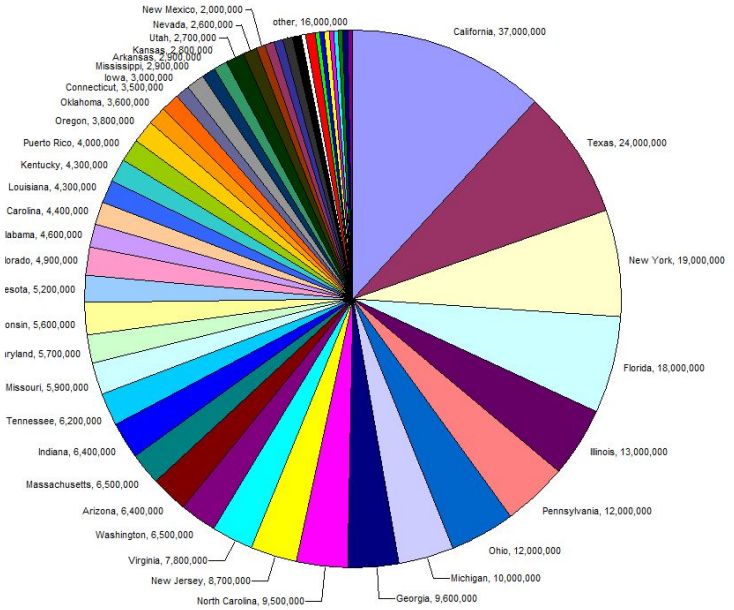


?



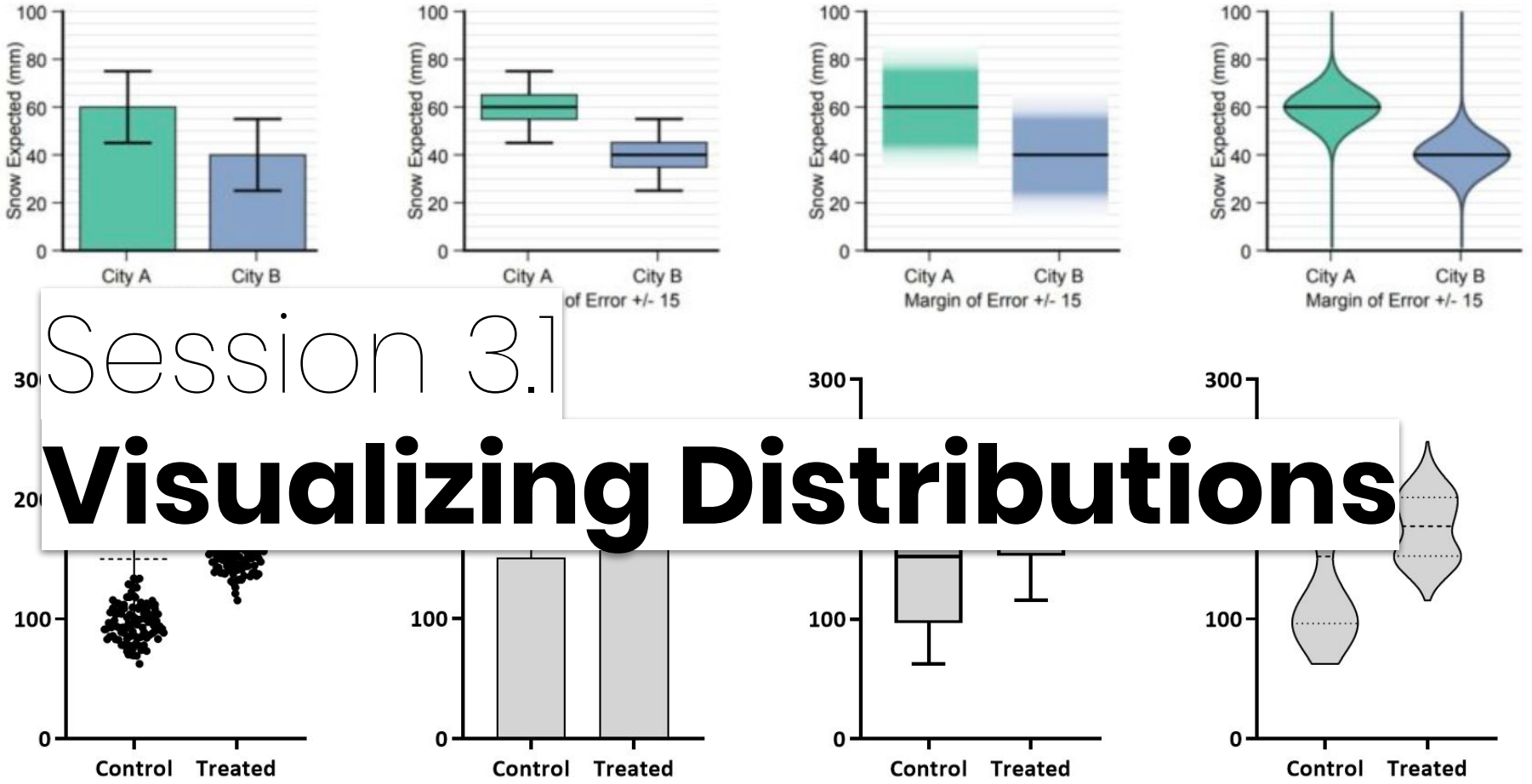
Basic Charts





Outline

- Distributions & Histograms
- Mean, Mode, Average
- Chart types
- Uncertainty
- Pie Charts



Session 3.1

Visualizing Distributions

Online Course
Data Visualization for Professionals

THE UNIVERSITY of EDINBURGH

Benjamin Bach

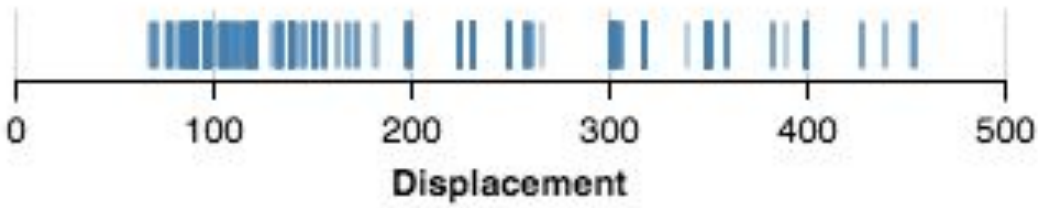
June 2020

http://benjbach.me

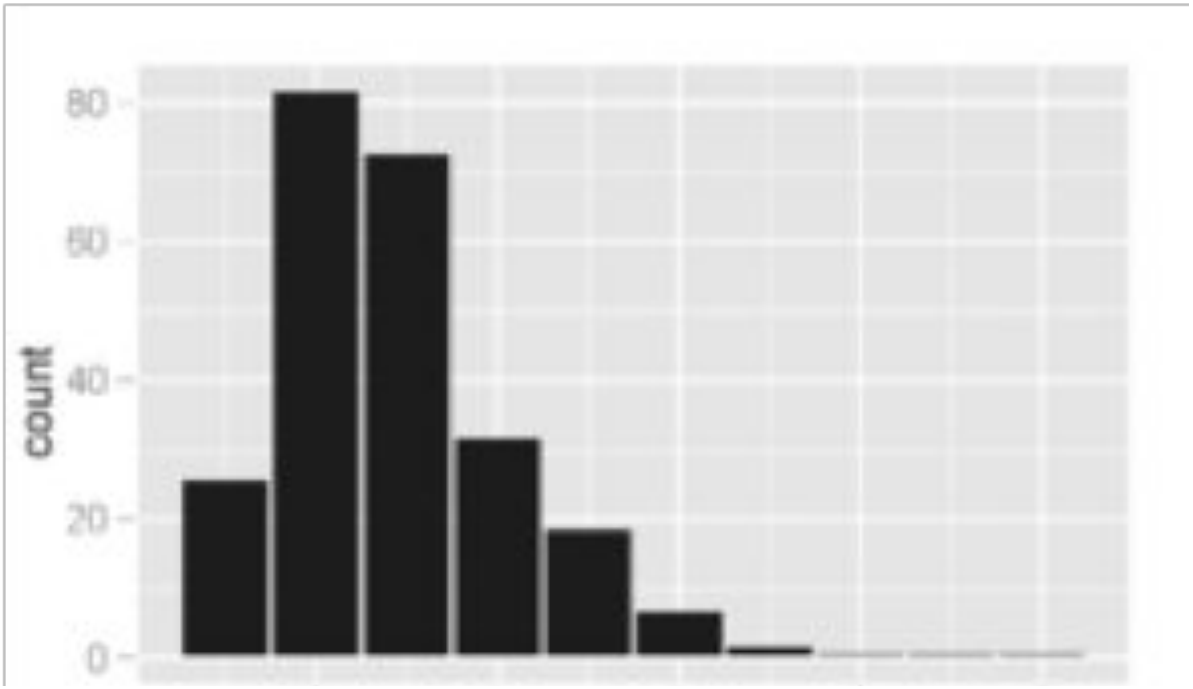
https://datavis-online.github.io

-- Not for external use --

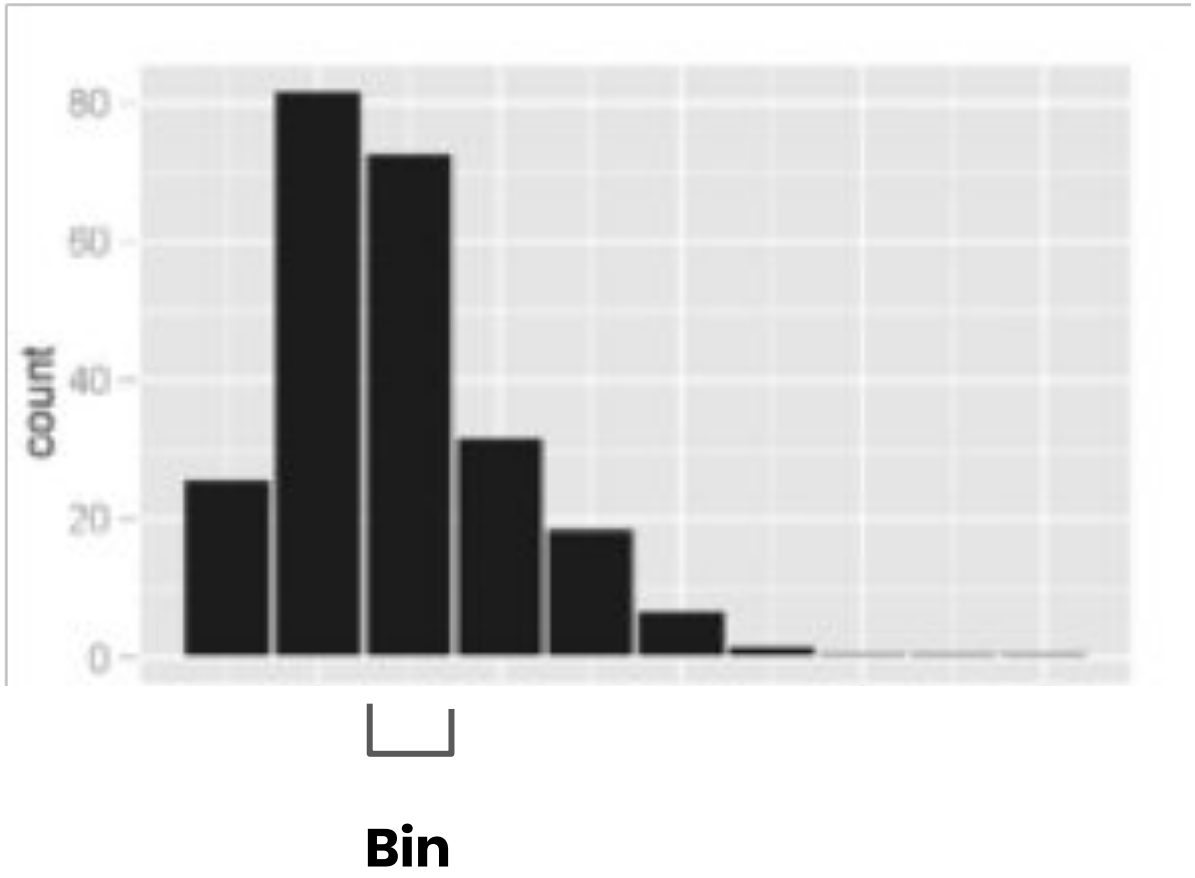
Distributions



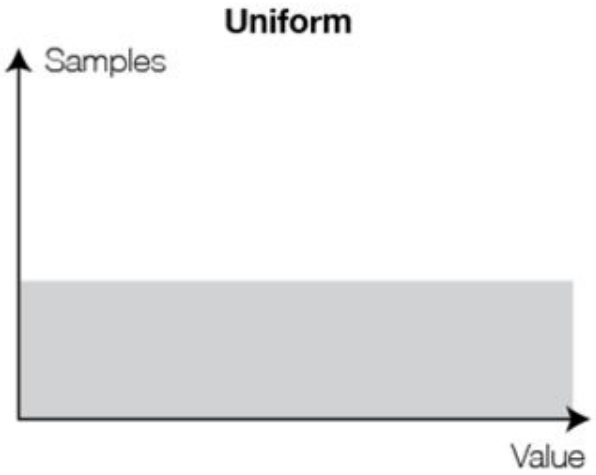
Distributions: Binning



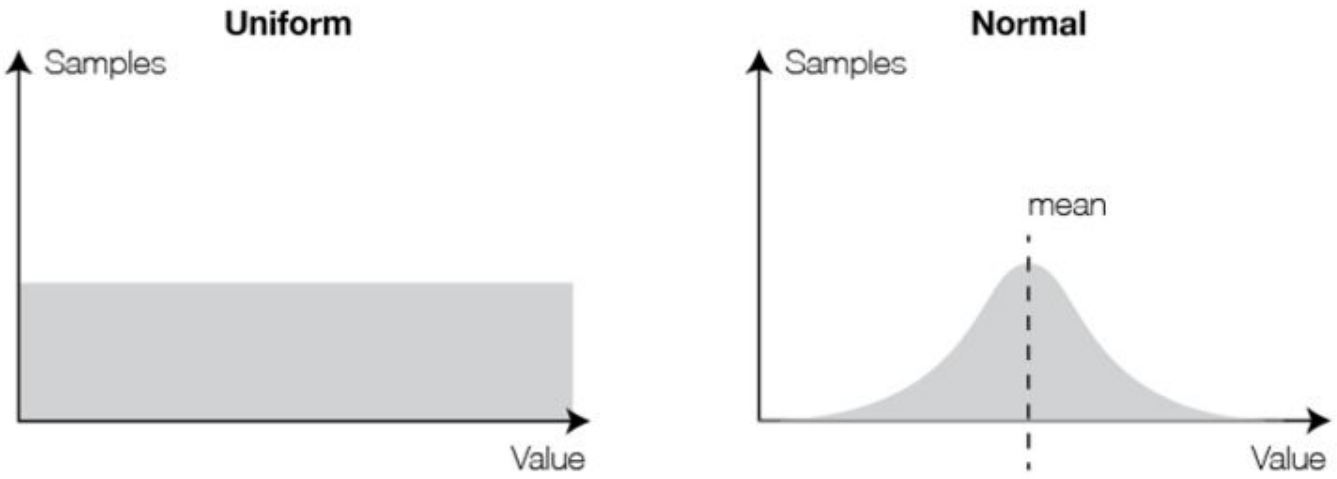
Distributions: Binning



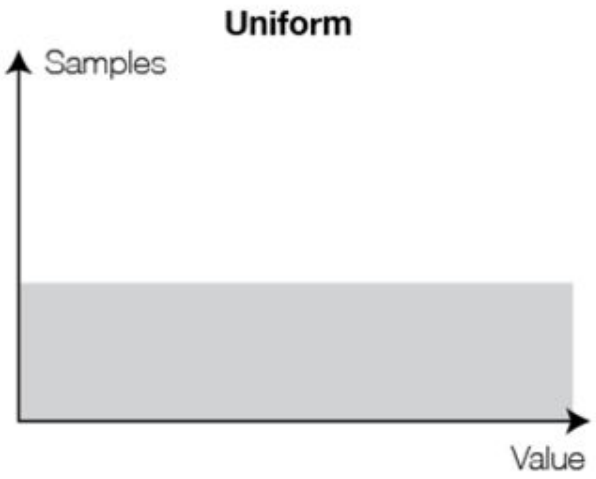
Distributions

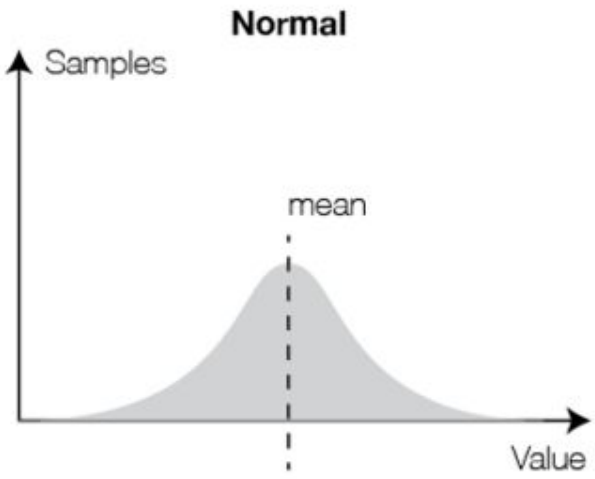


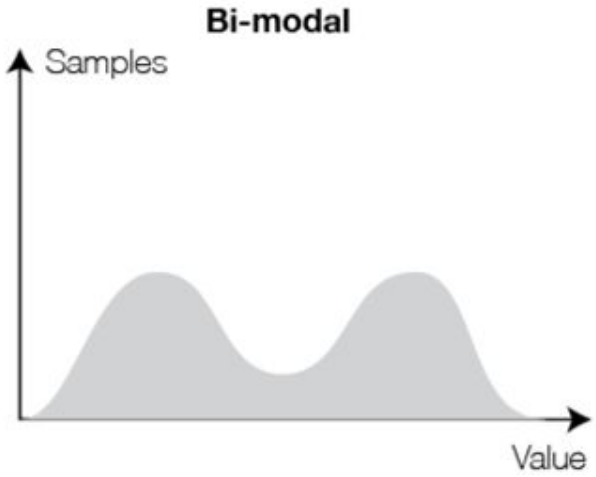
Distributions



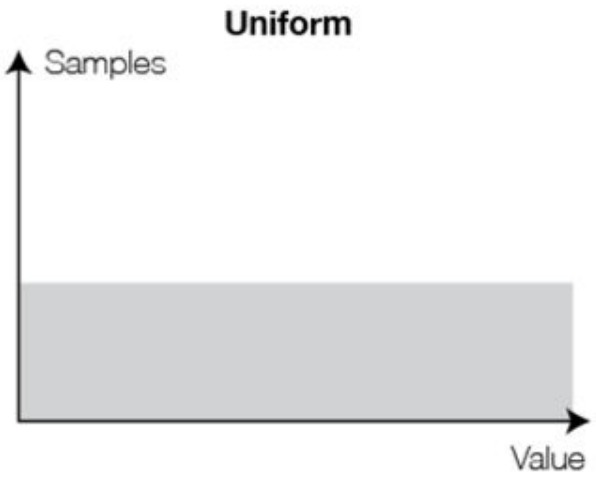
Distributions

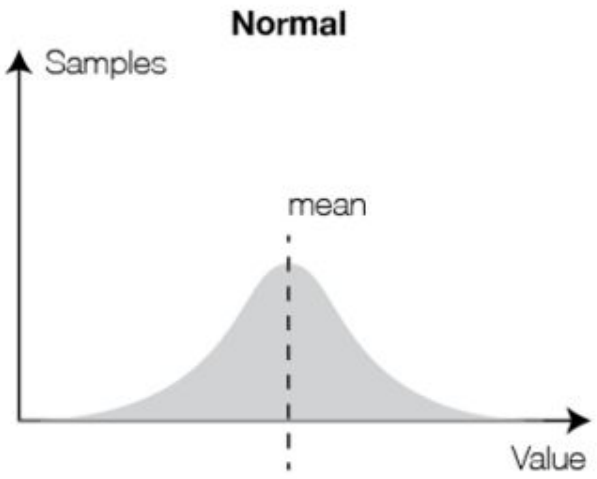


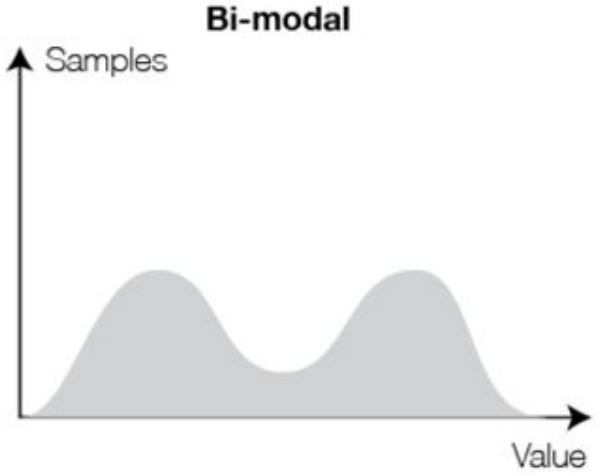


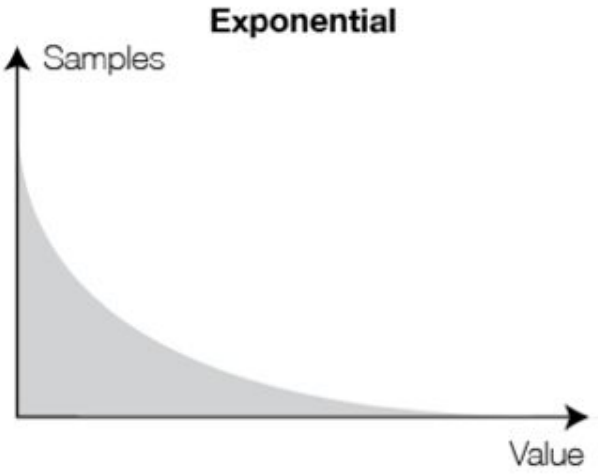


Distributions

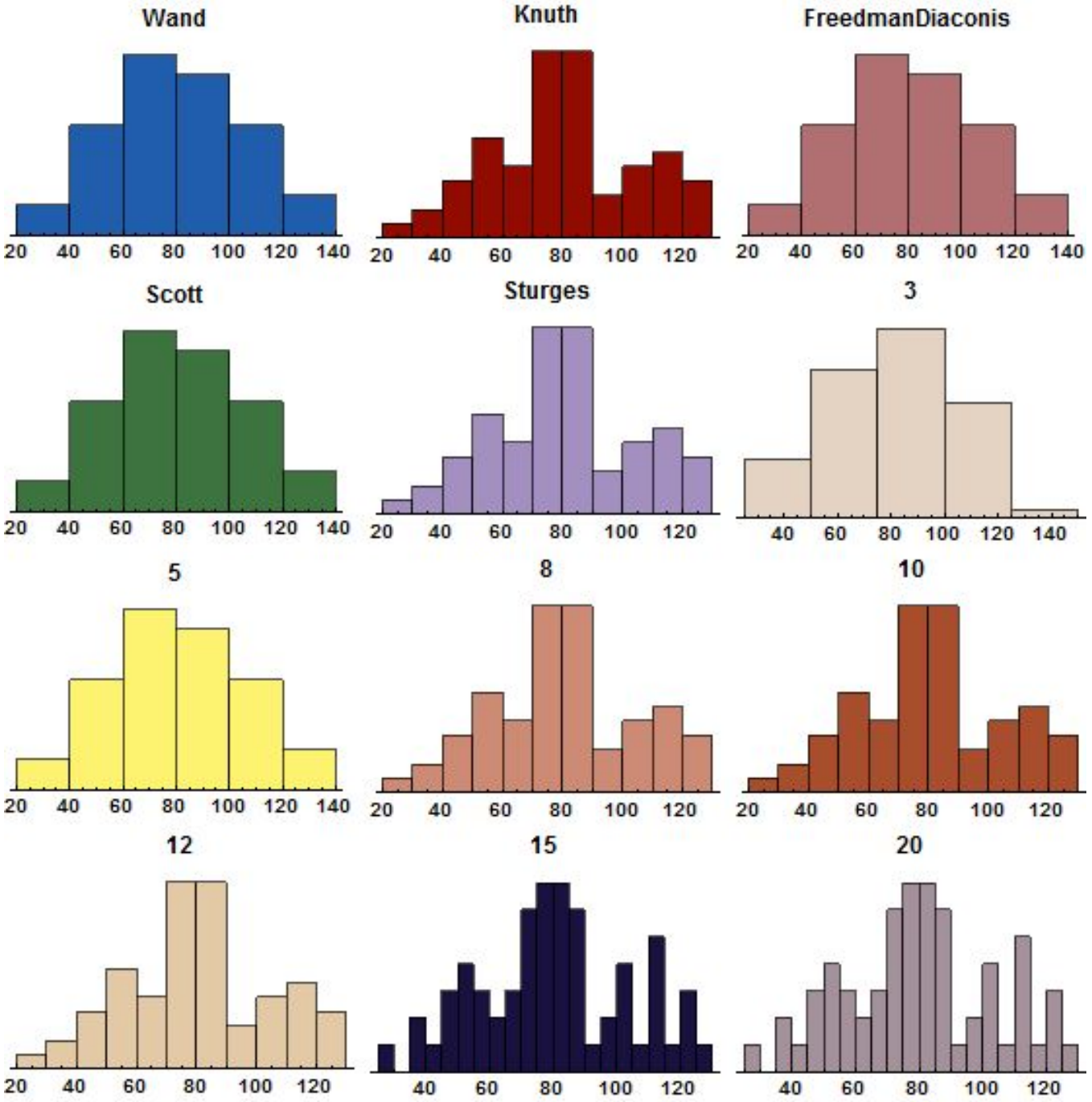




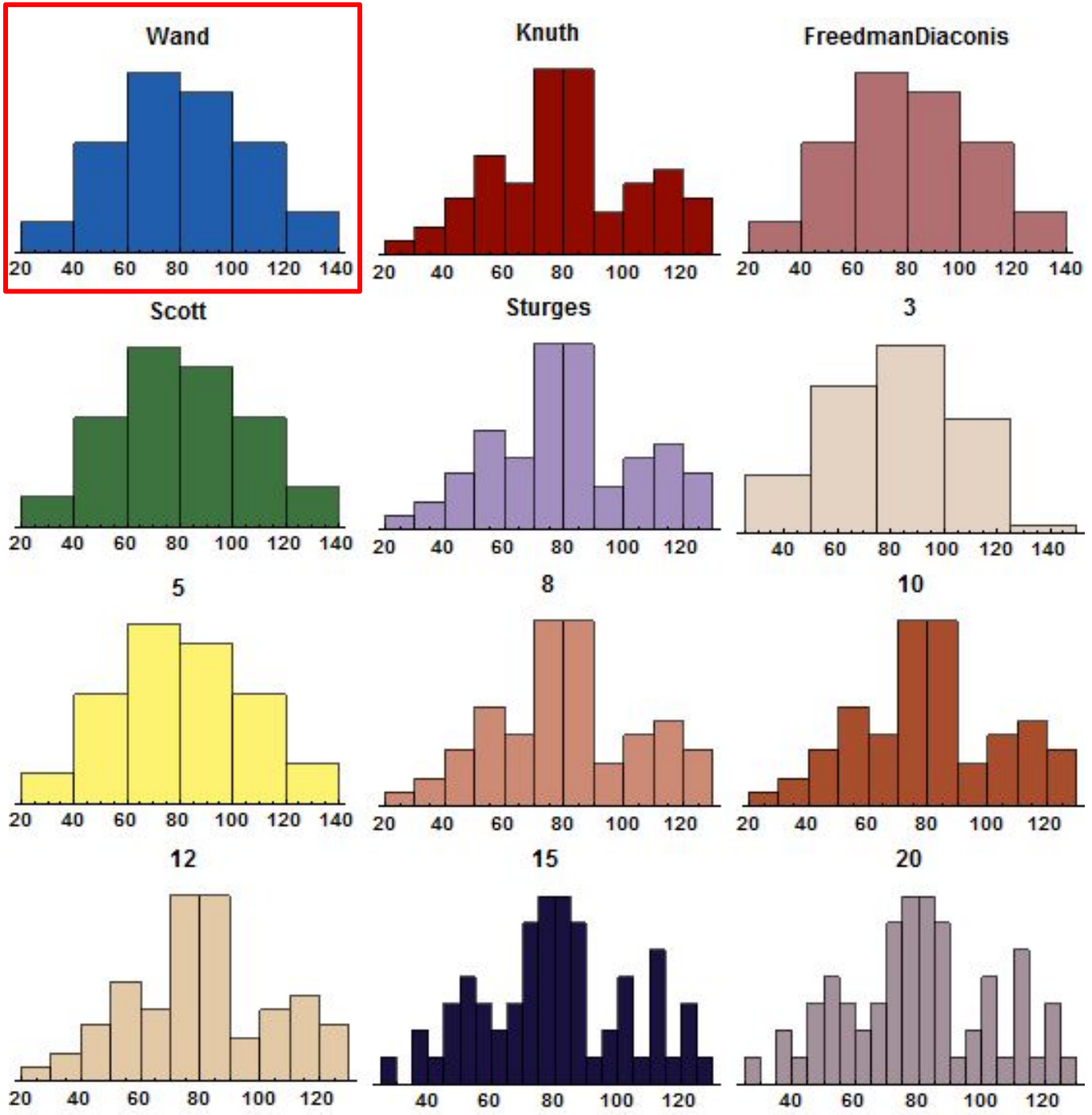




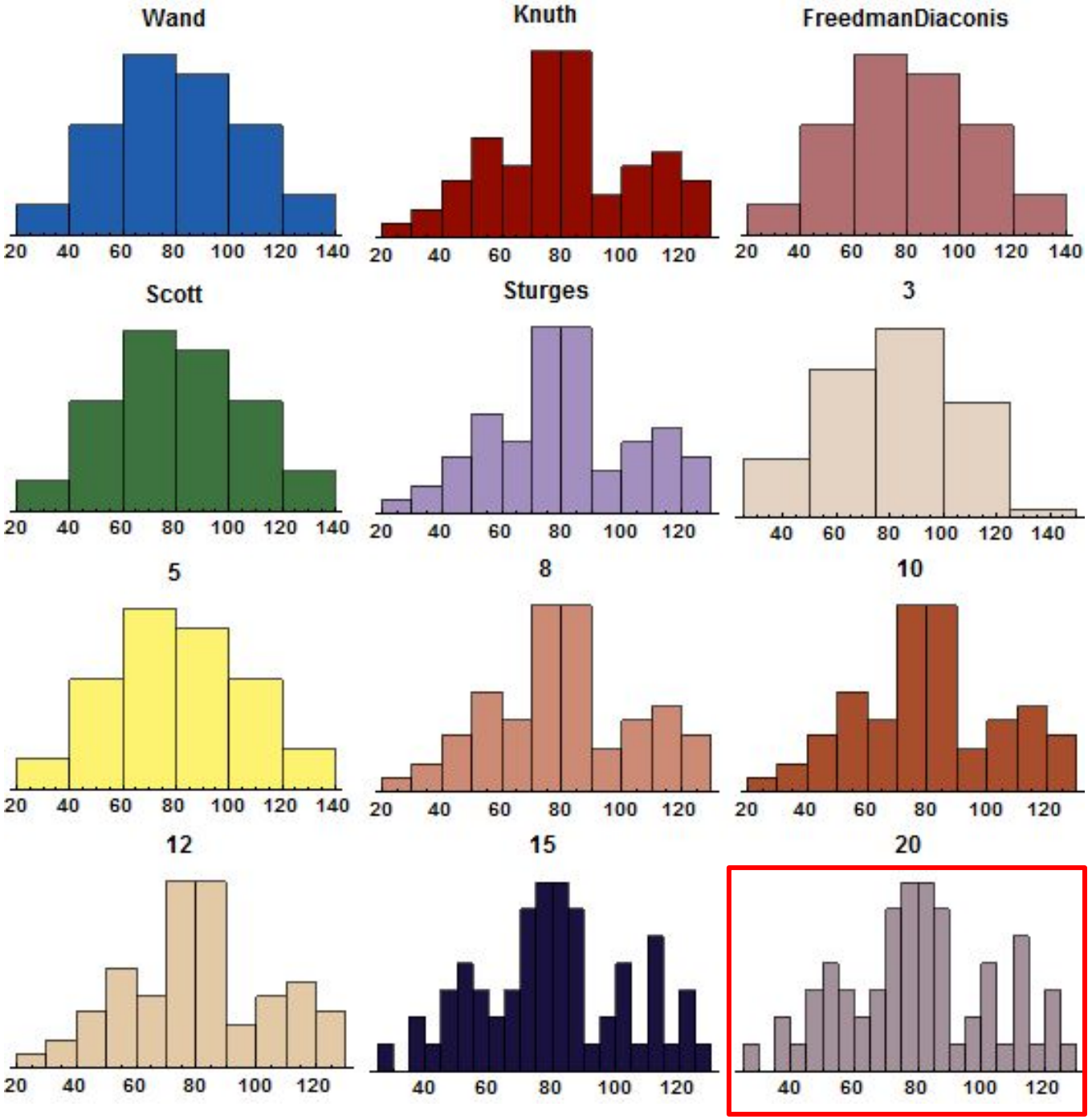
Histogram Binning



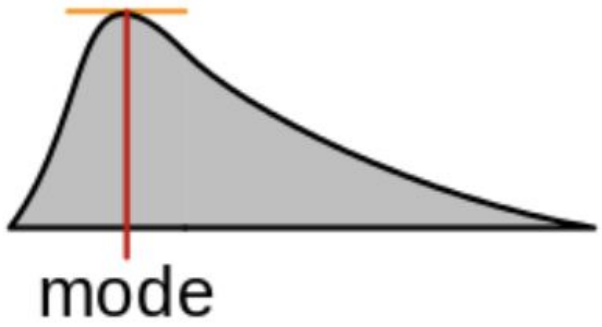
Histogram Binning



Histogram Binning

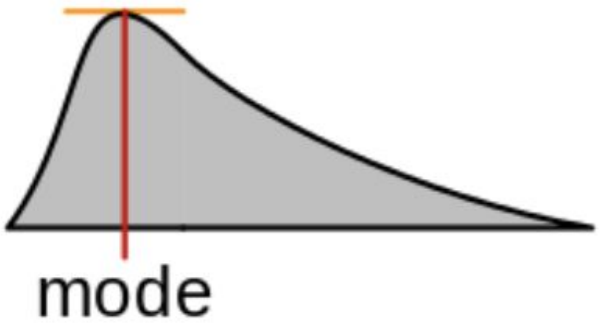


Descriptive Statistics

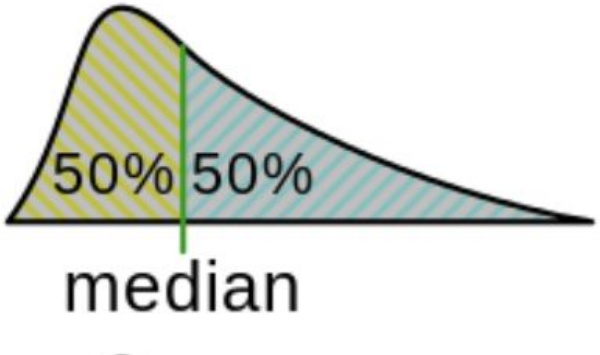


Mode: Most frequent value in data

Descriptive Statistics

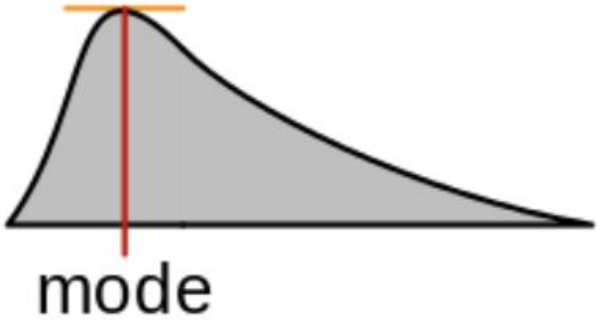


Mode: Most frequent value in data

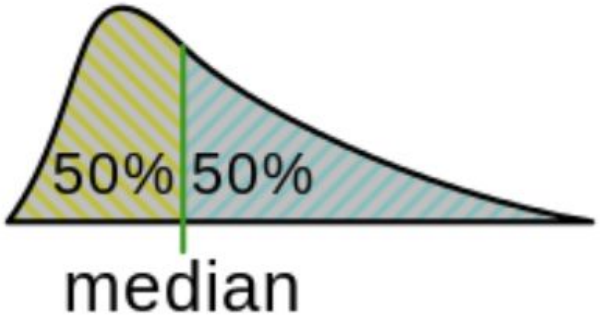


Median: splits samples in two equally sized sets (50:50)

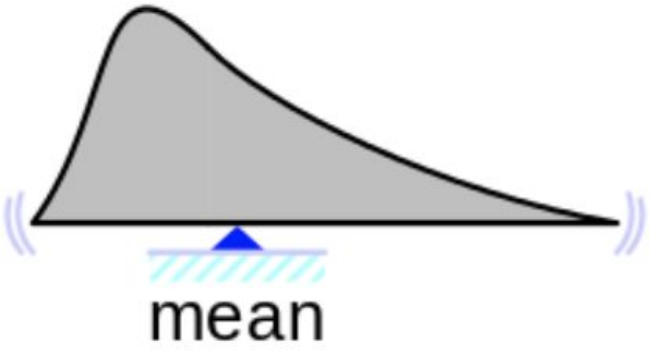
Descriptive Statistics



Mode: Most frequent value in data

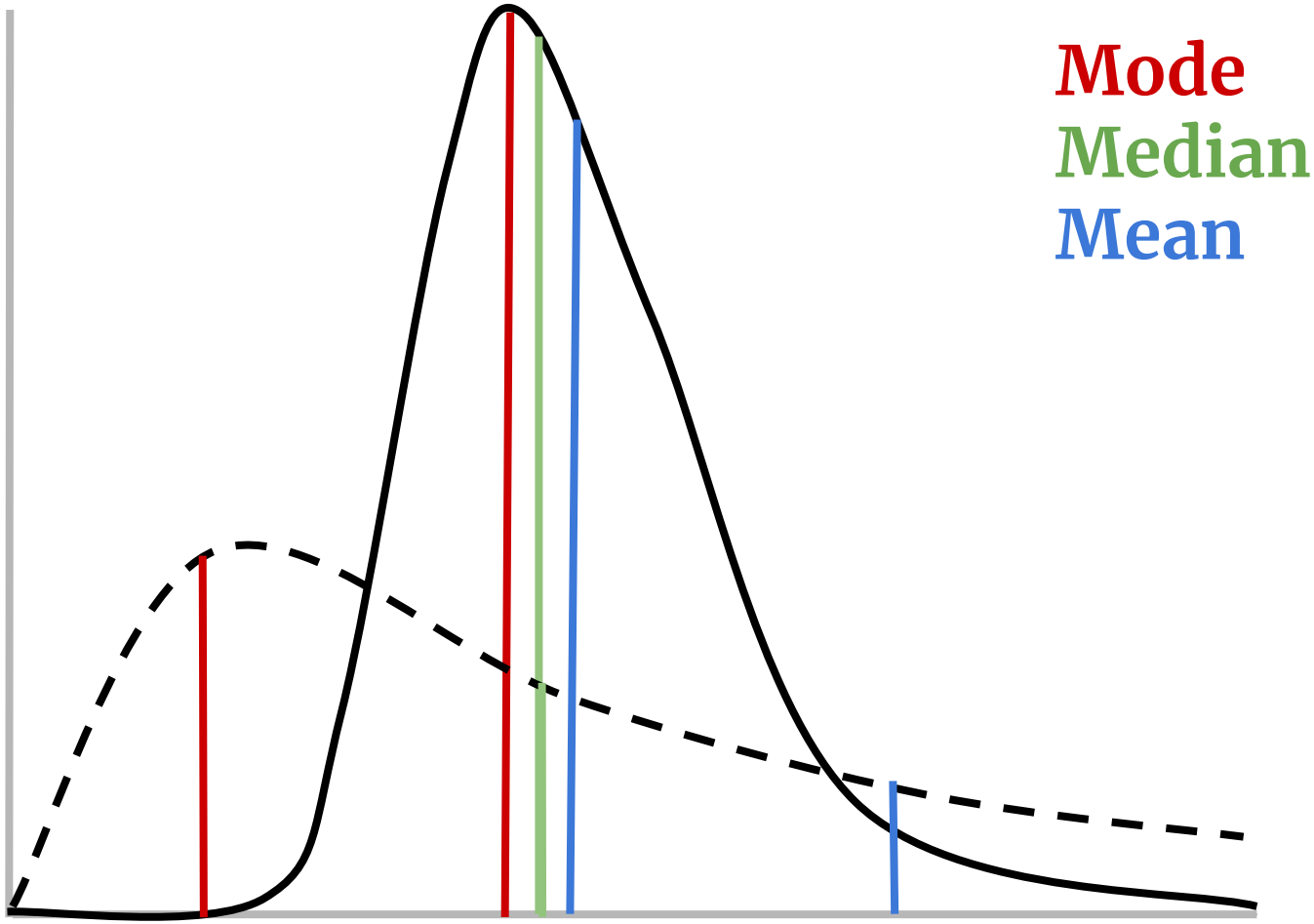


Median: splits samples in two equally sized sets (50:50)

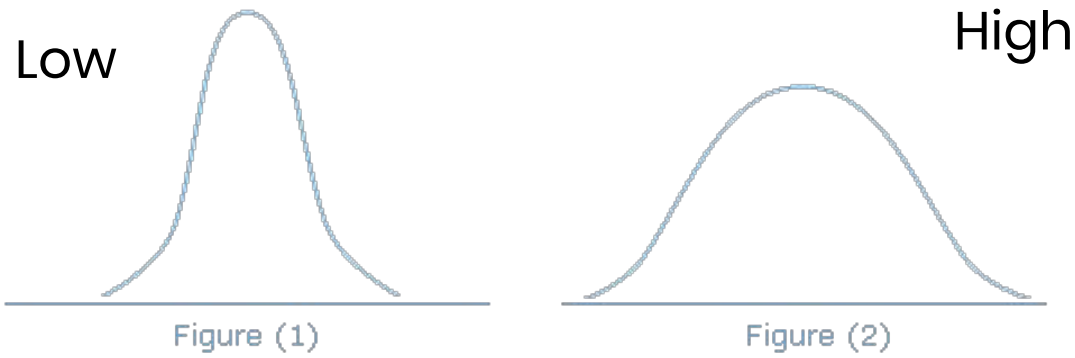


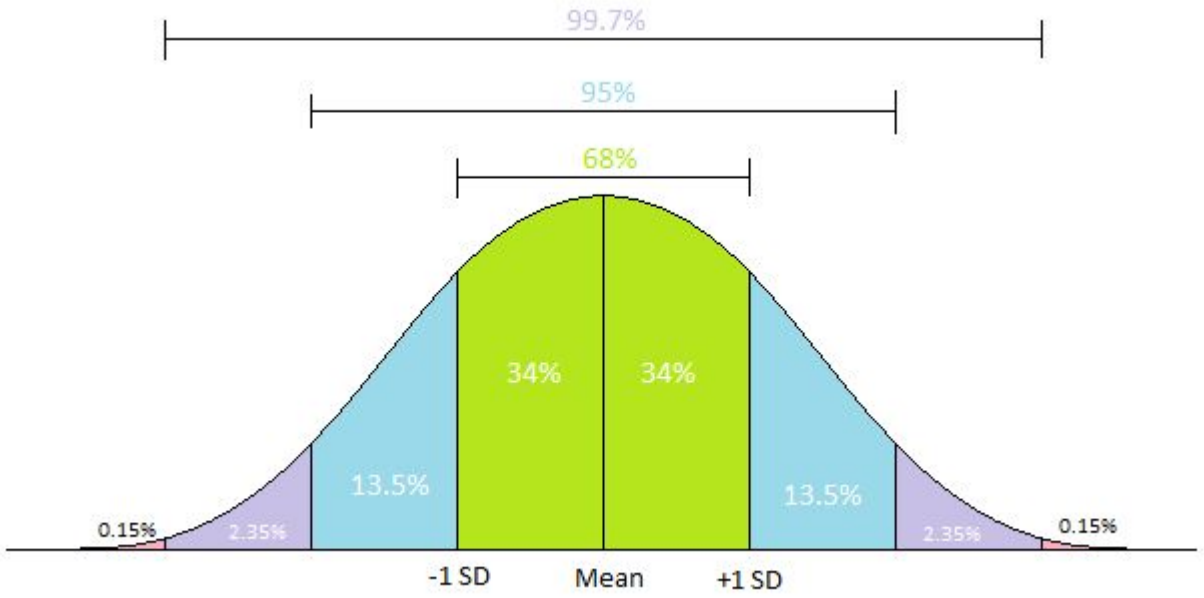
Mean: average value

Descriptive Statistics

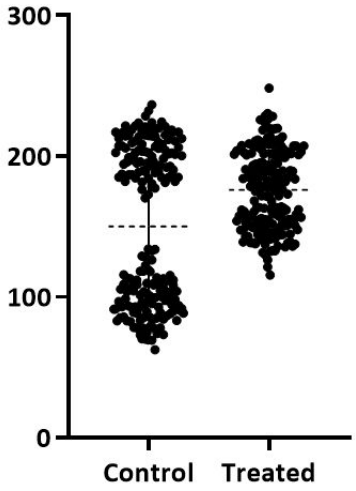


Standard deviation

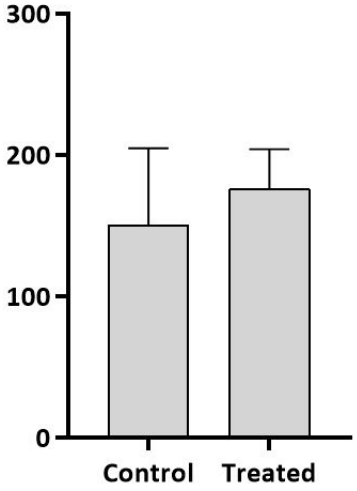




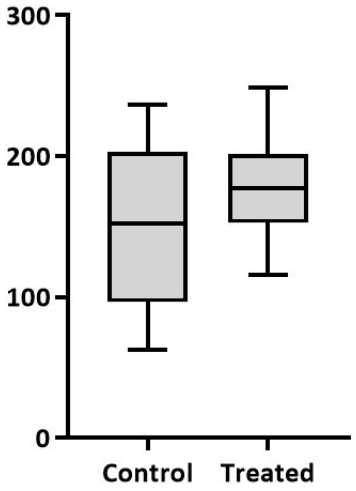
Plotting Distributions



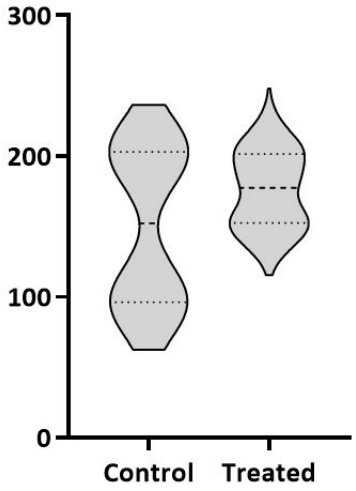
Dot plot/
Beeswarm plot



Bar charts +
error bars

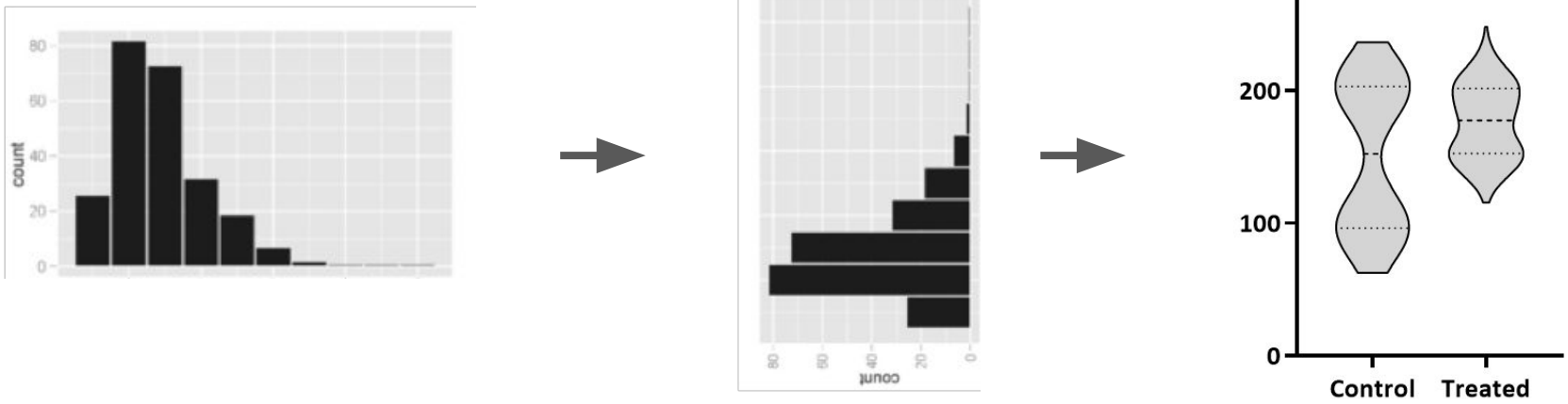


Box plots

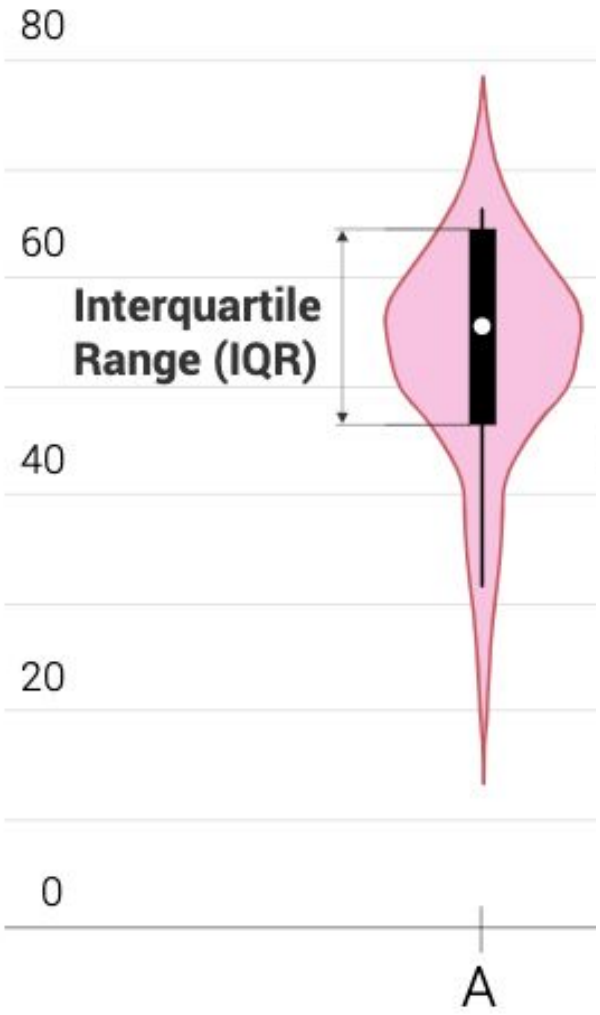


Violin plots

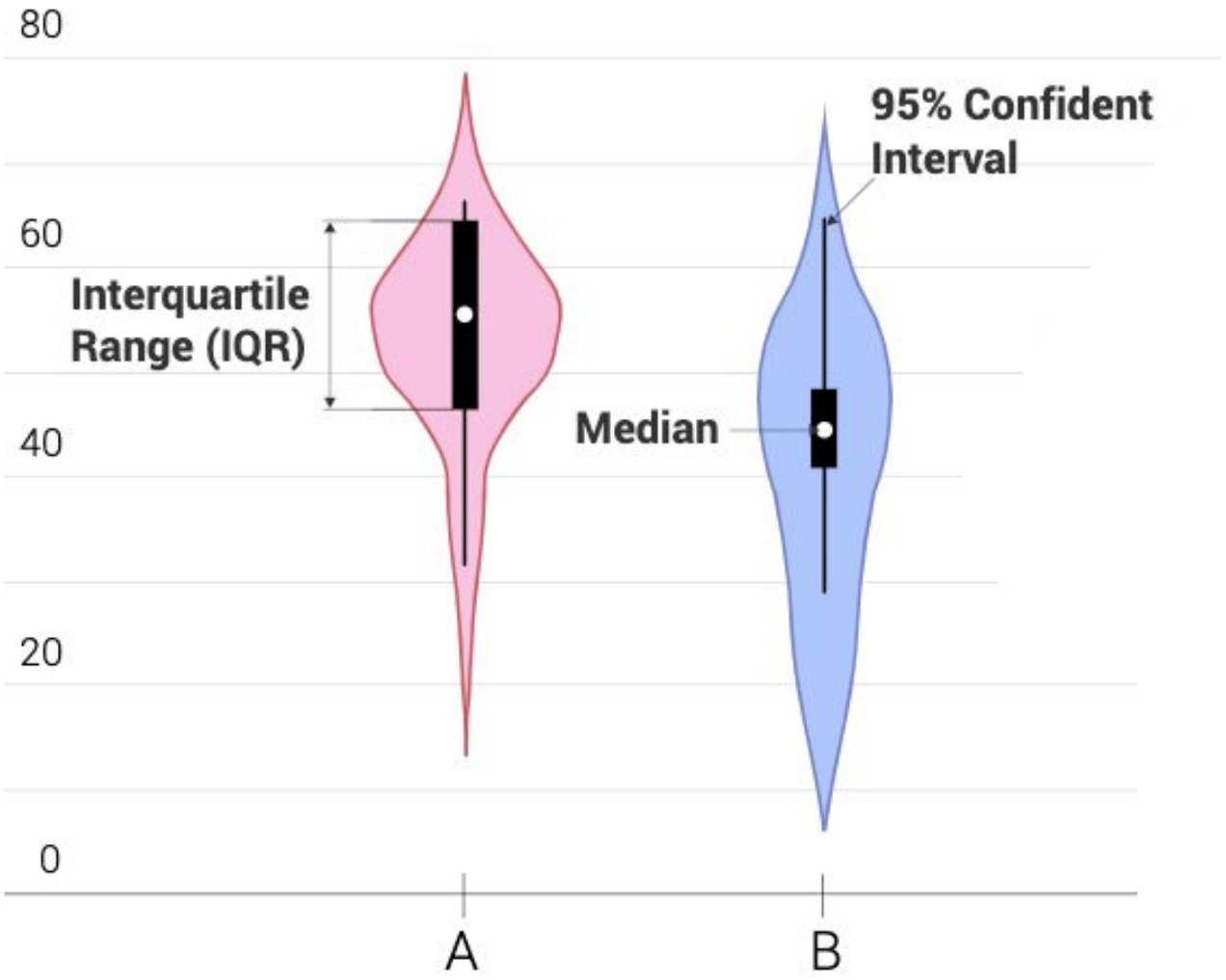
Violin plots



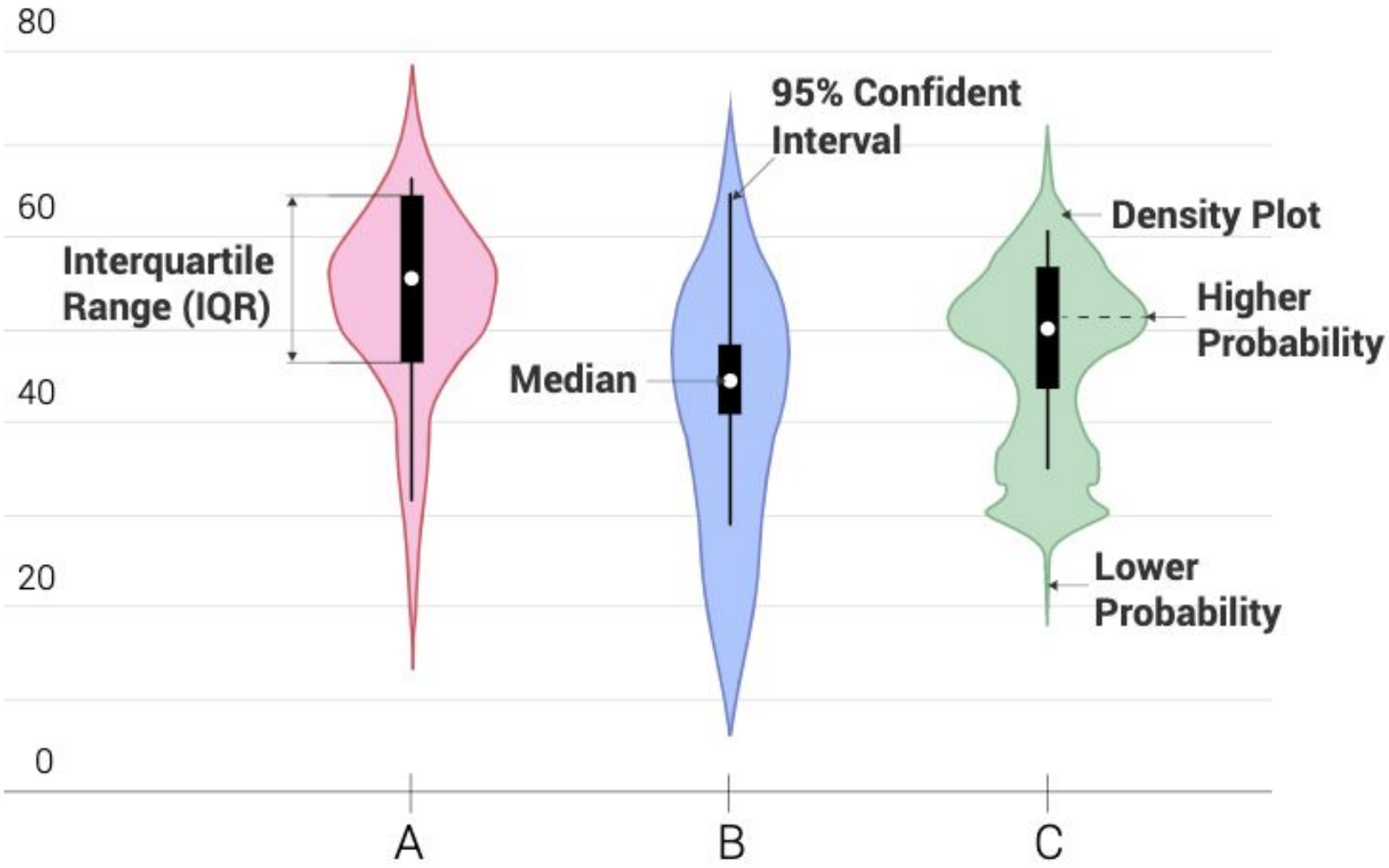
Violin plots: measures



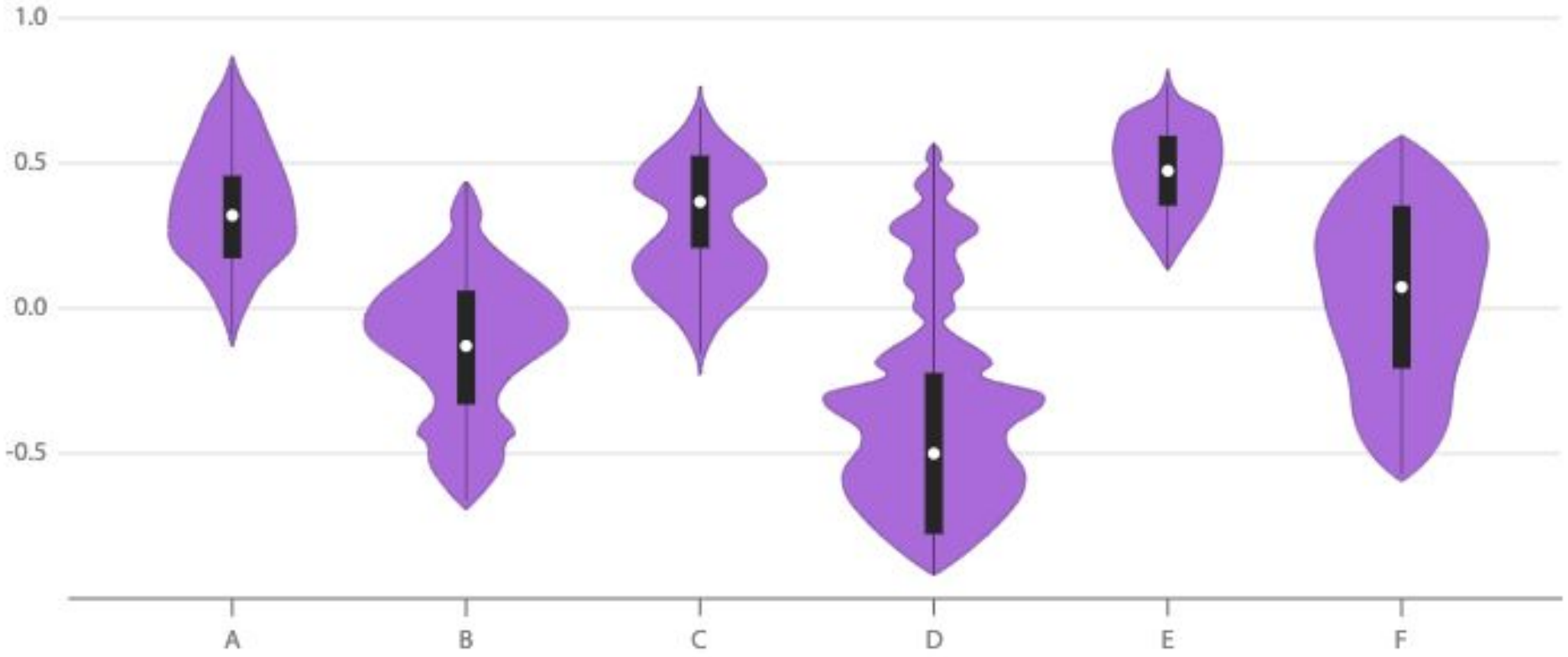
Violin plots: measures



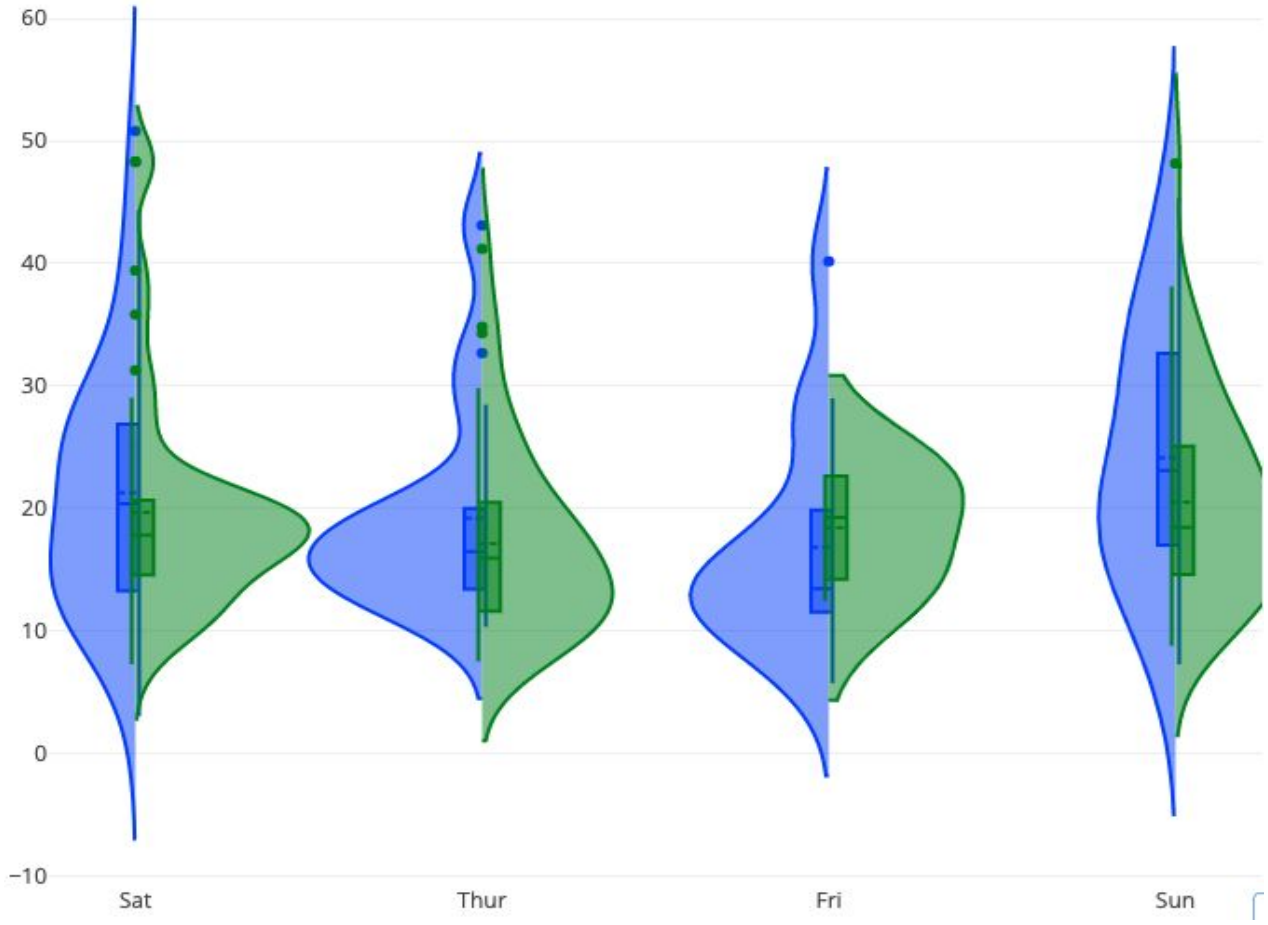
Violin plots: measures



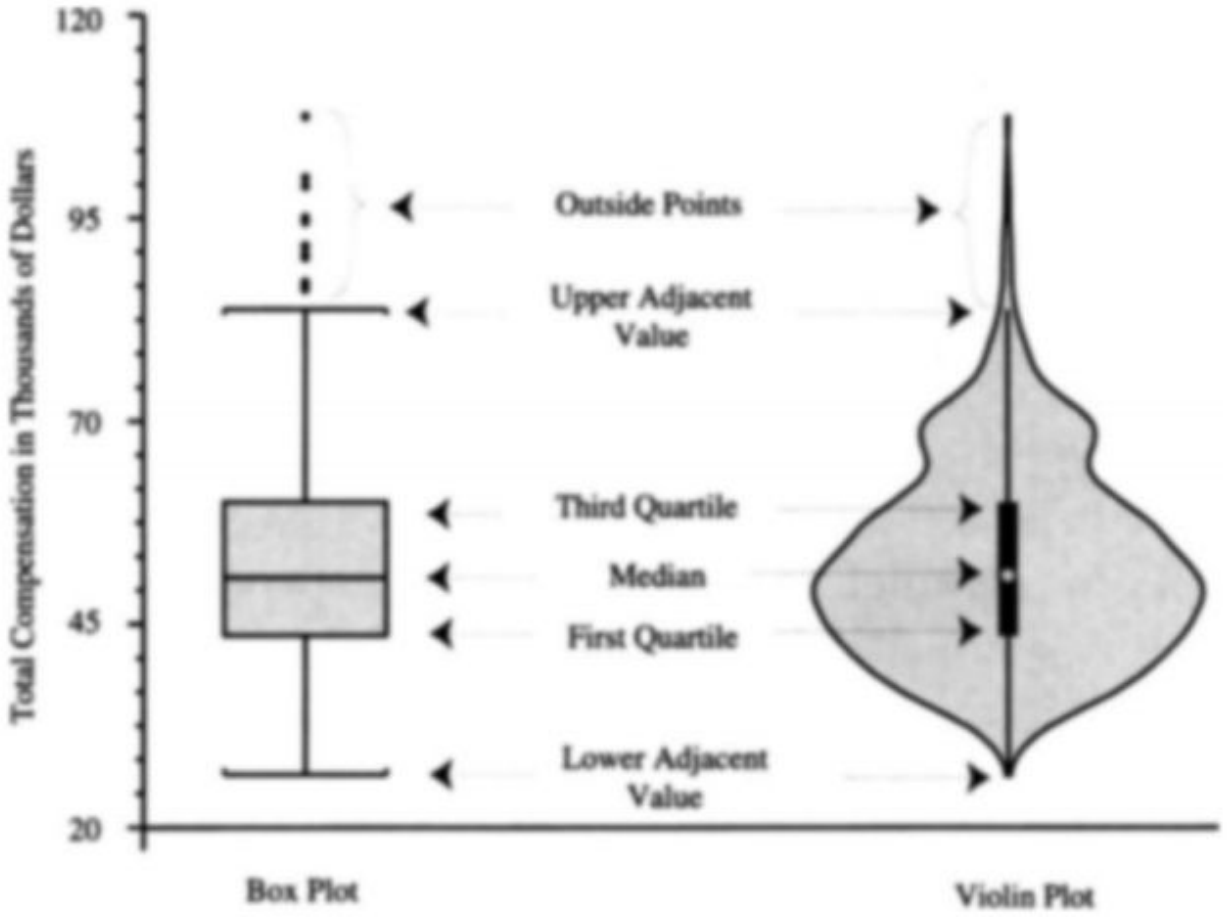
Violin plots: example shapes



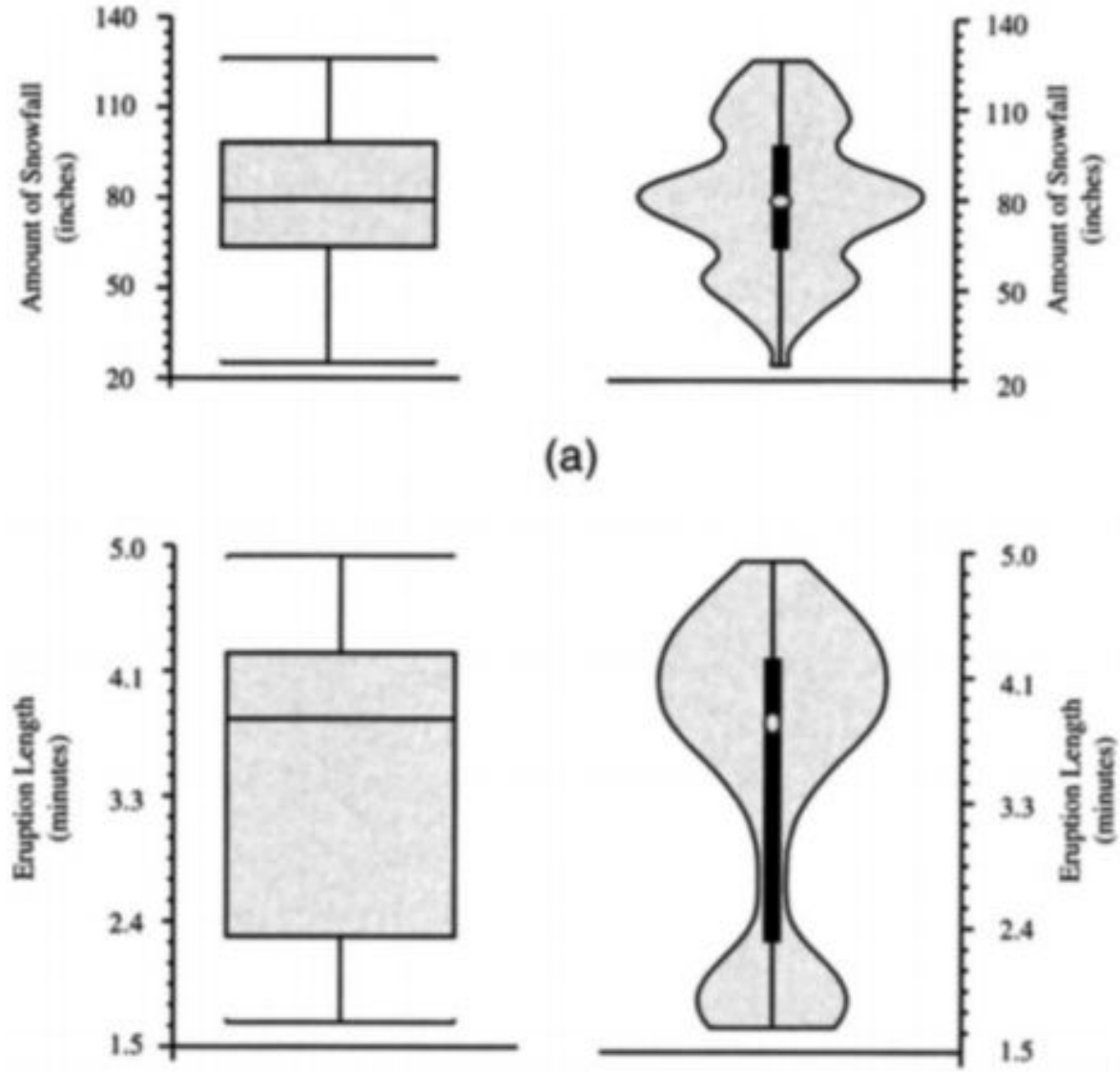
Violin plots for comparison



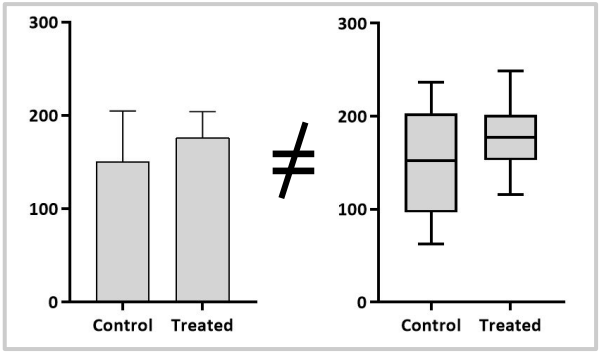
Box plots



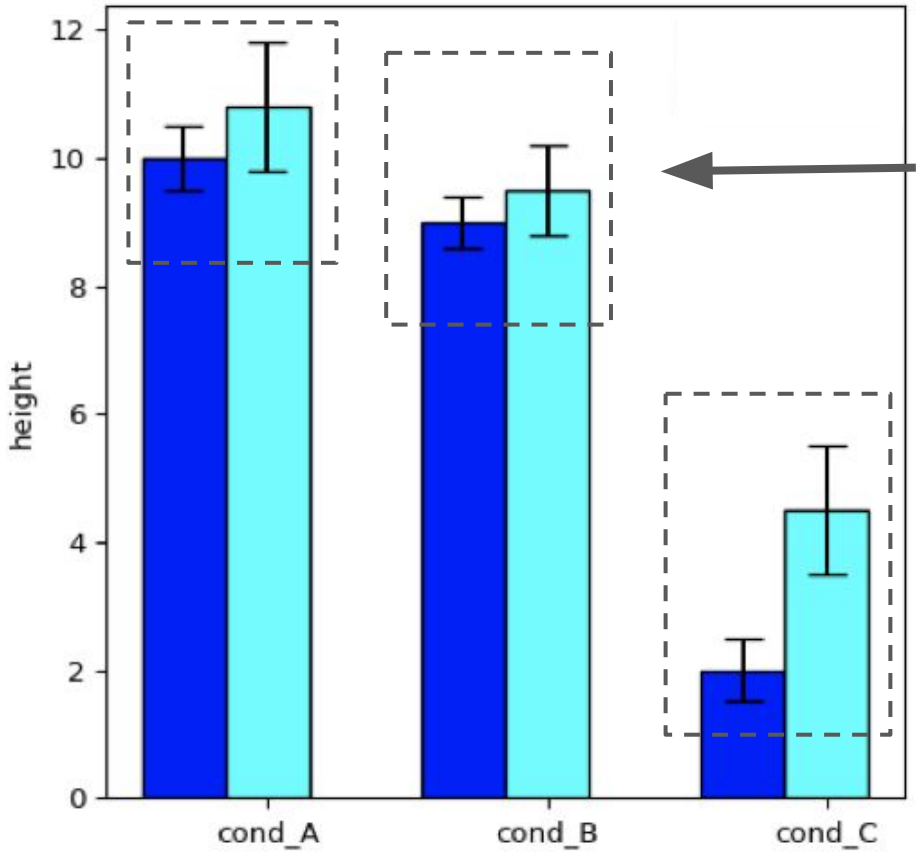
http://www.stat.cmu.edu/~rnugent/PCMI2016/papers/ViolinPlots.pdf


Amount of Snowfall (inches)
140
110
80
50
20
(a)
Eruption Length (minutes)
5.0
4.1
3.3
2.4
1.5
(b)

Bar plots + Error bars



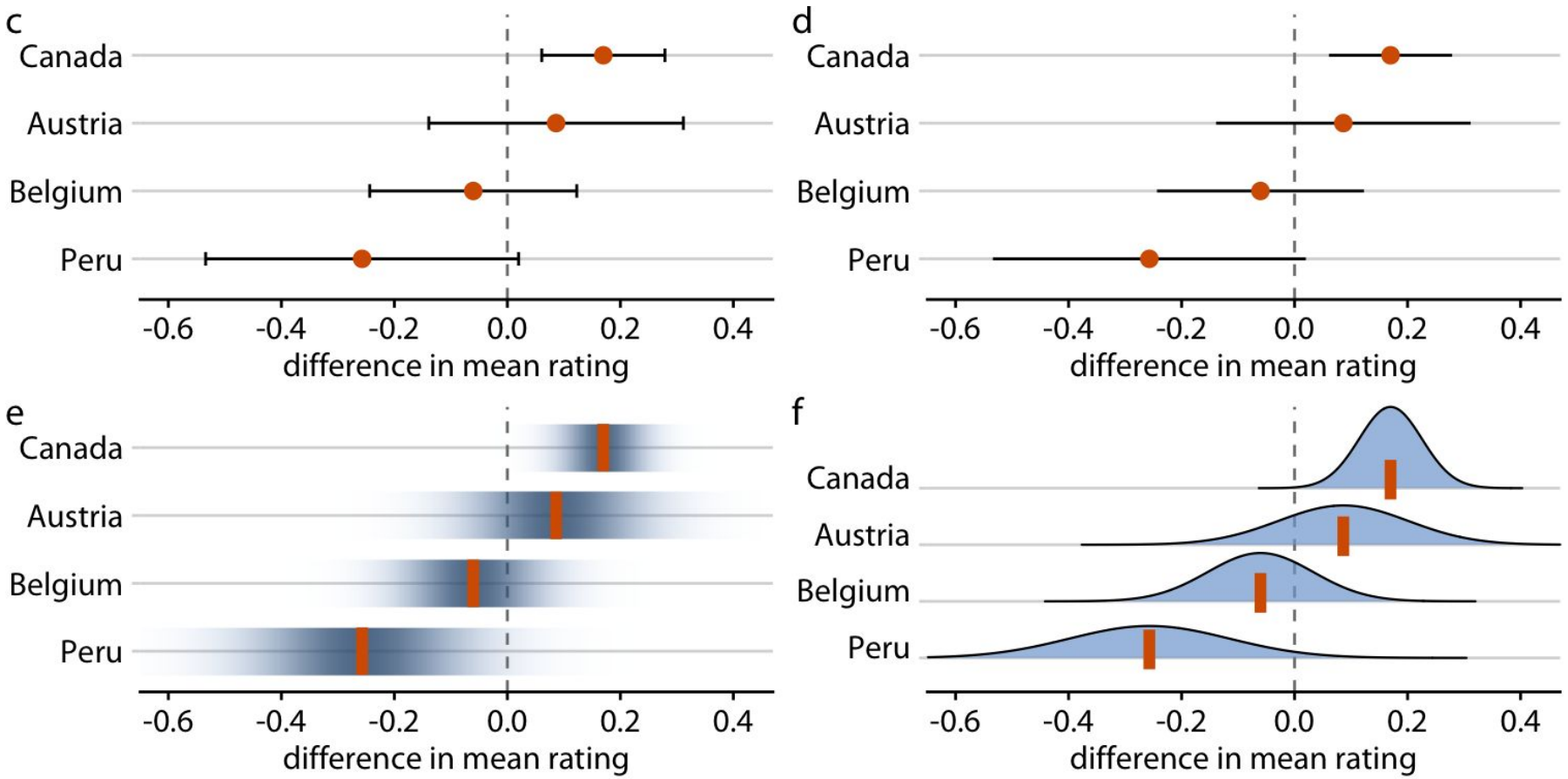
?



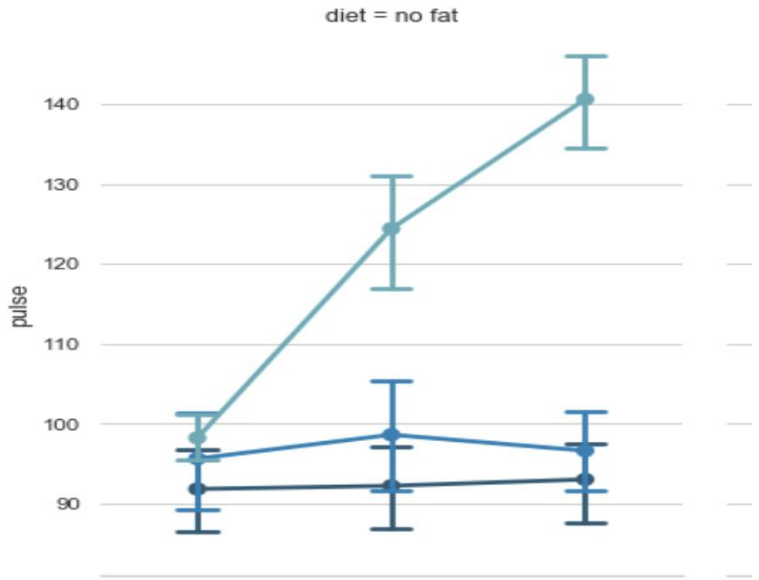
No significant difference

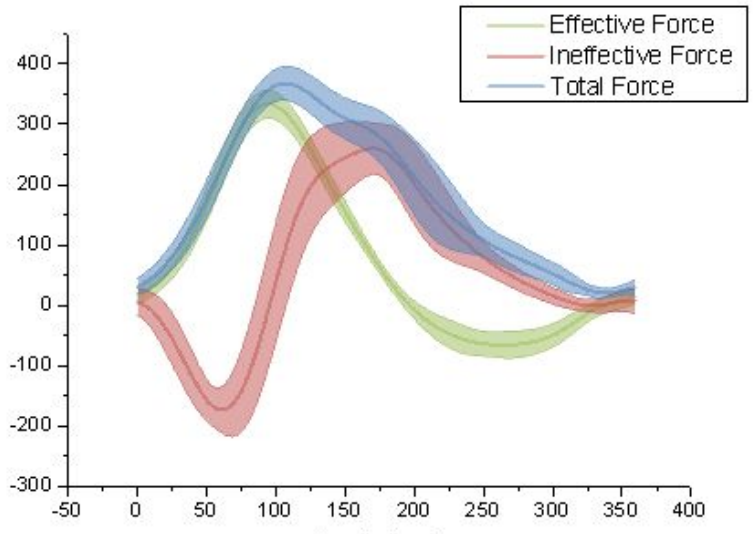
Significance difference

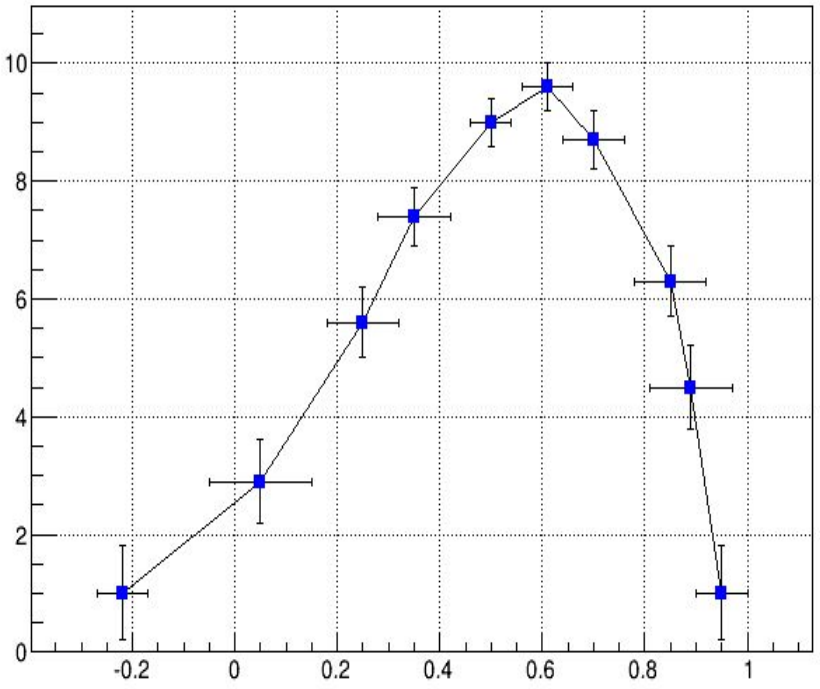
Uncertainty



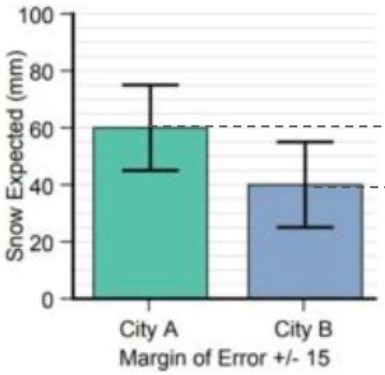
Uncertainty





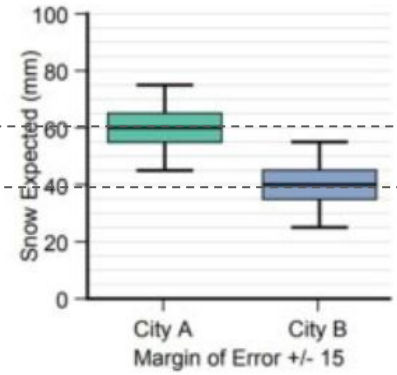


Comparing plots

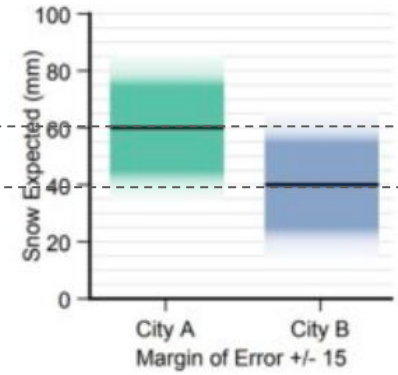


Bar plot

!

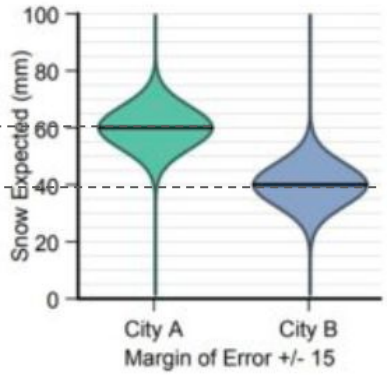


Box plot

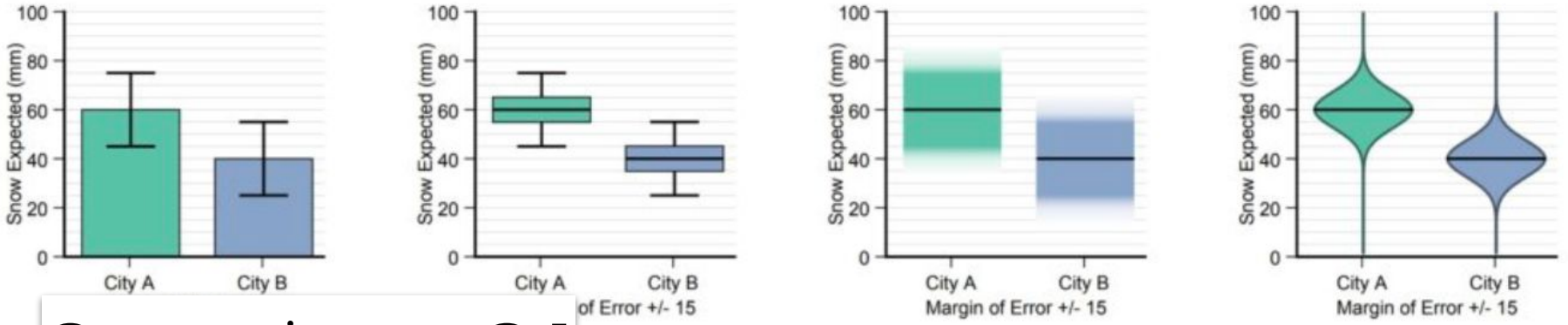


Gradient plot

!

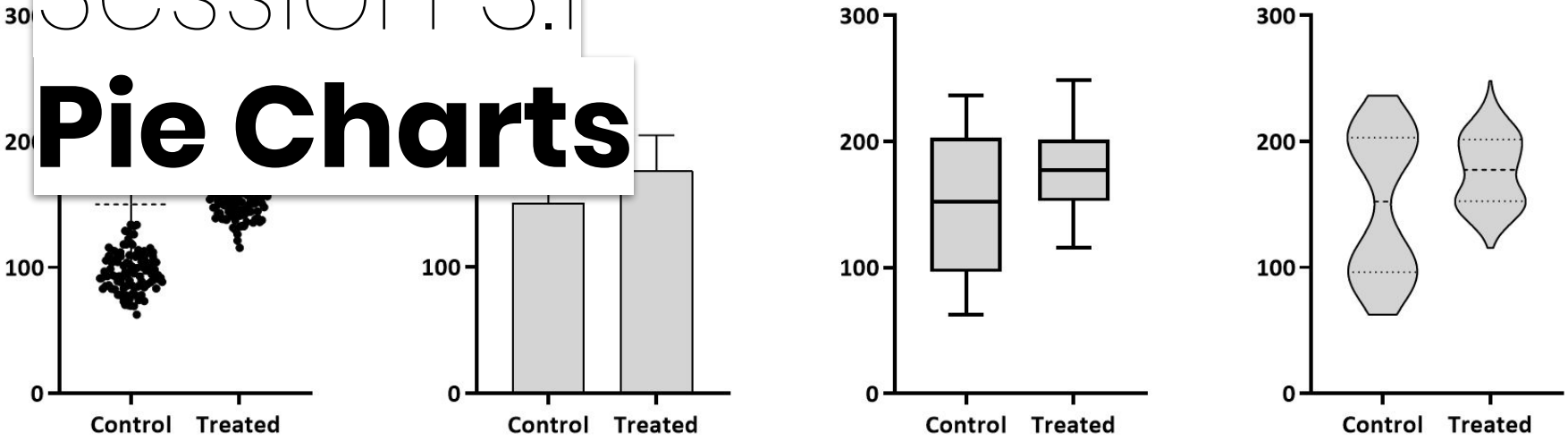


Violin plot



Session 3.1

Pie Charts



Online Course
Data Visualization for Professionals

THE UNIVERSITY of EDINBURGH

Benjamin Bach

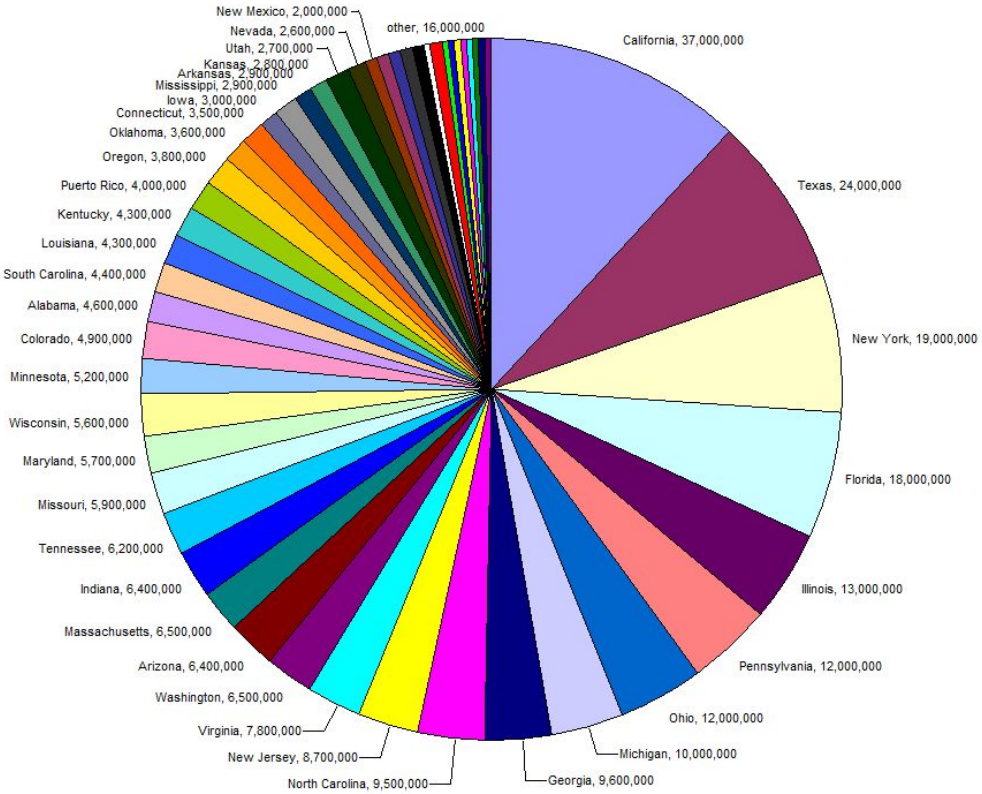
June 2020

http://benjbach.me

https://datavis-online.github.io

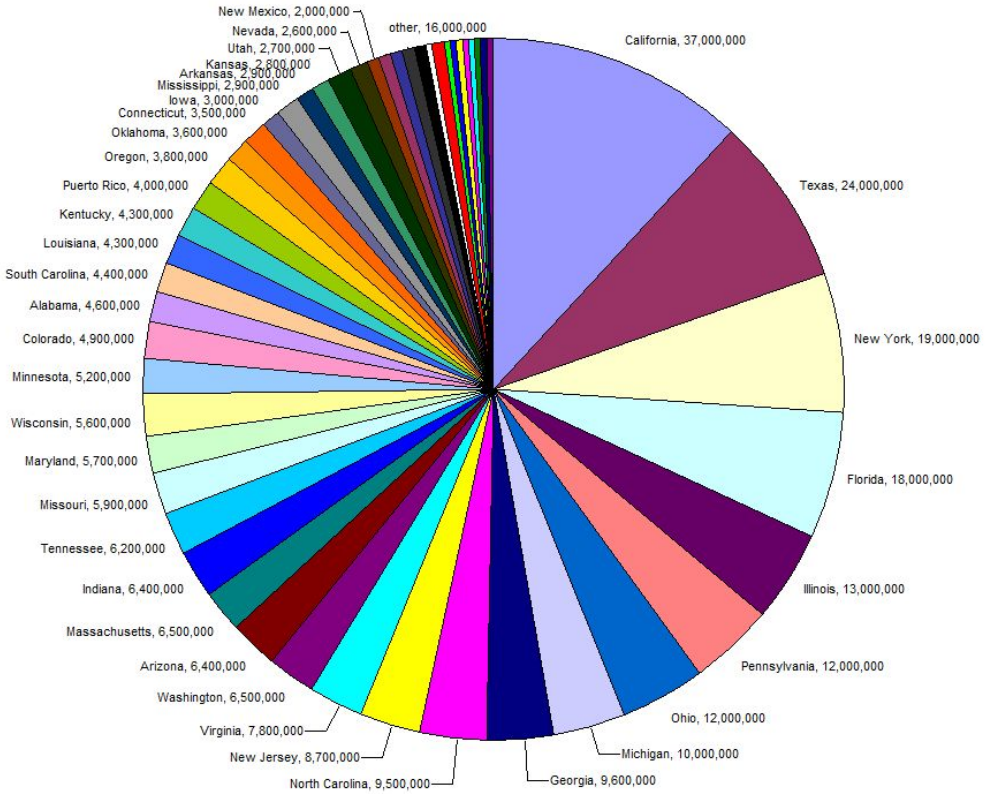
-- Not for external use --

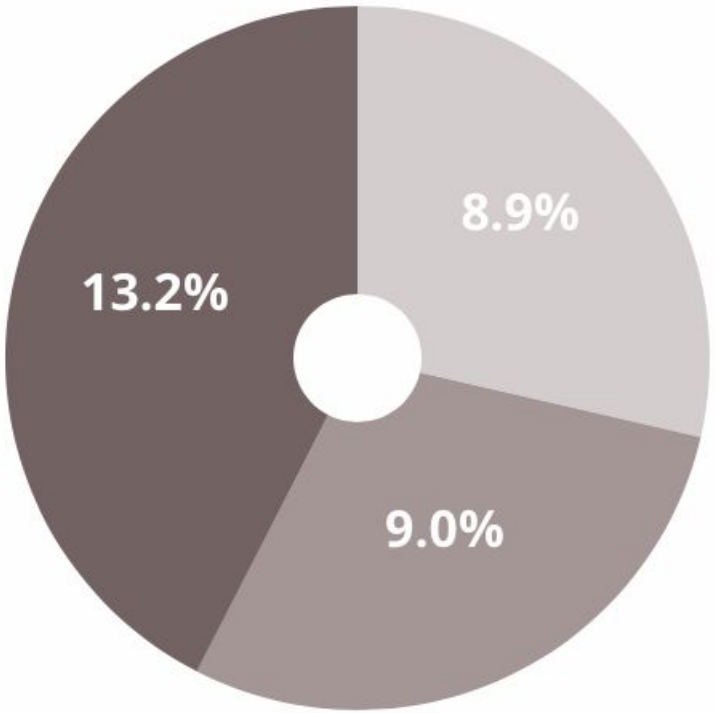
Bad Pie Charts



- Too many values
- Differences hard to understand
- Distribution?
- (Confusing colors)

Bad Pie Charts

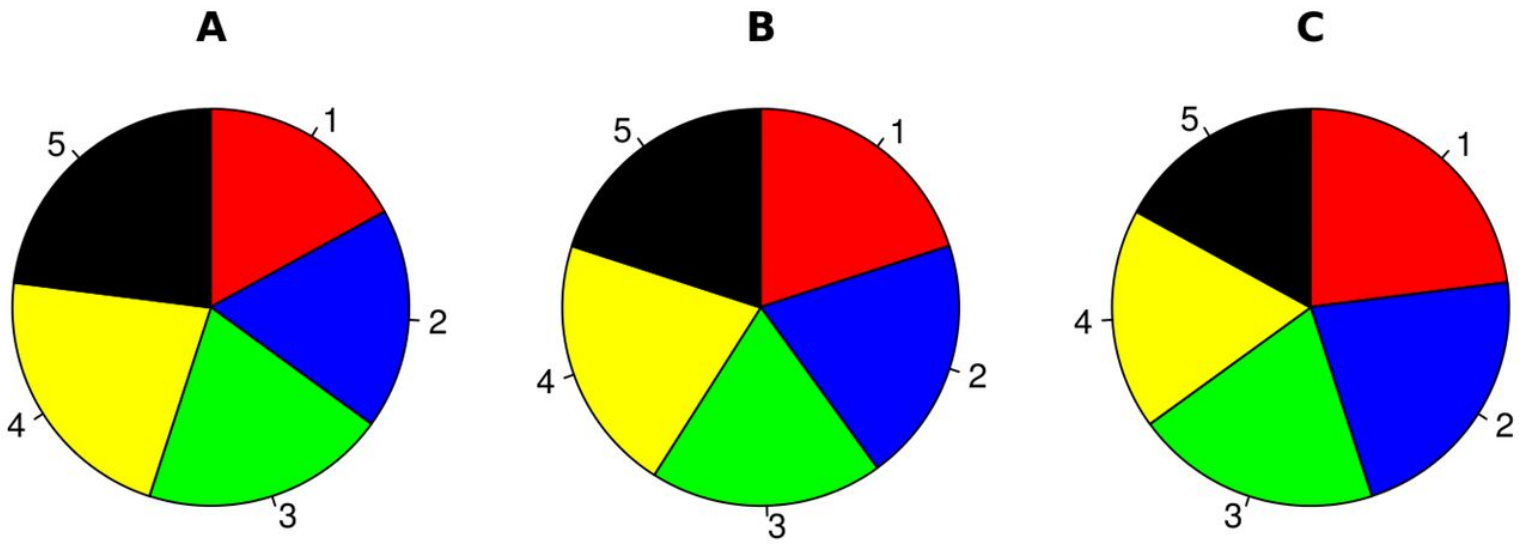




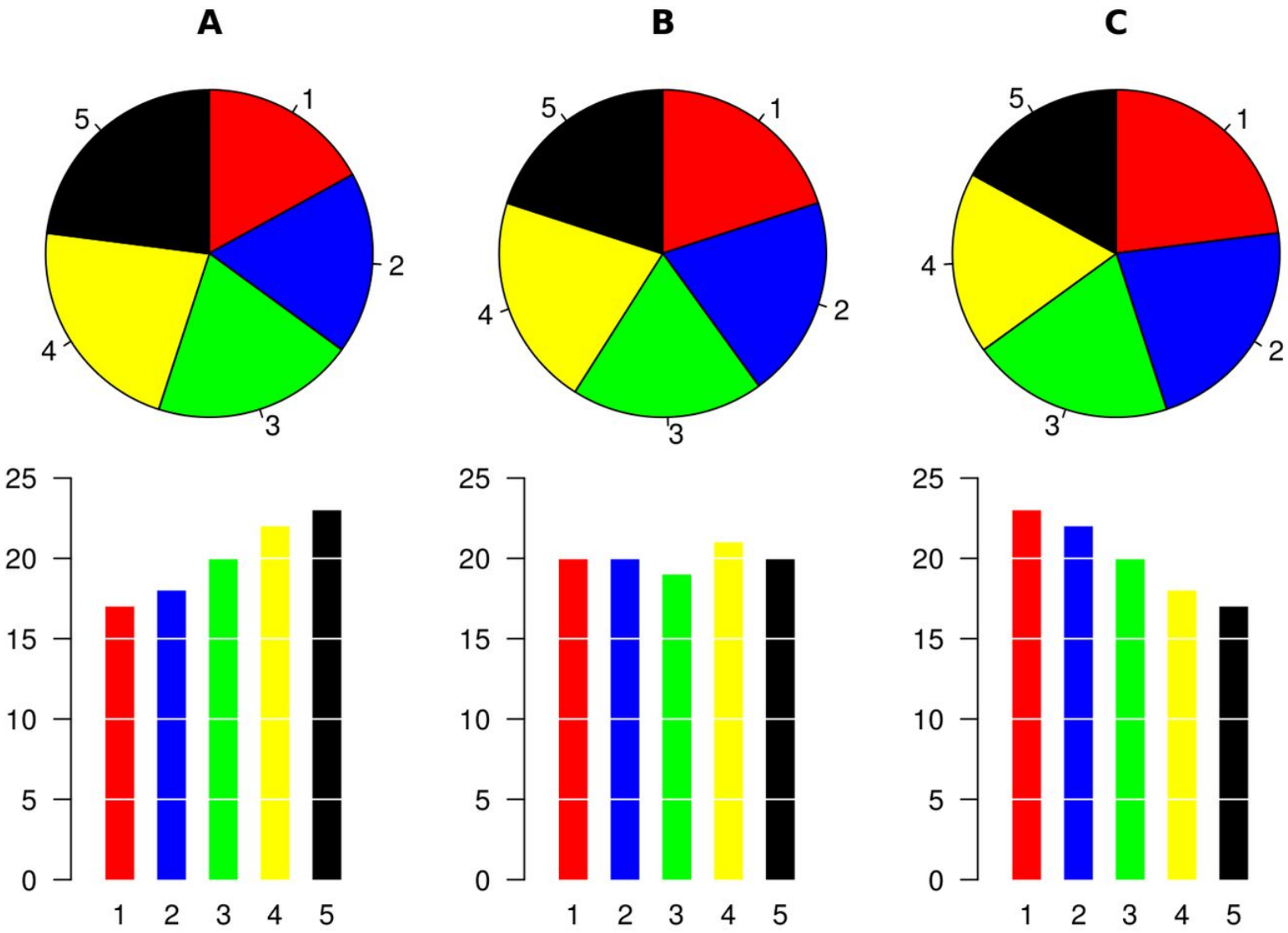
- Too many values
- Differences hard to understand
- Distribution?
- (Confusing colors)

- Not adding up to 100%

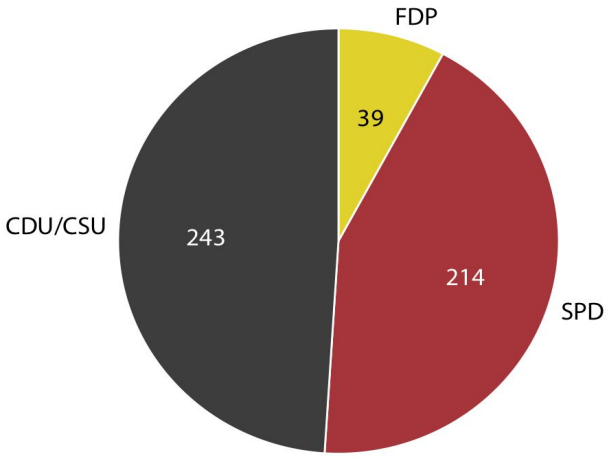
Pie Charts



Pie Charts

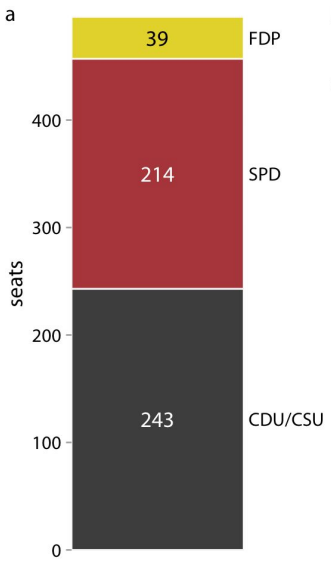


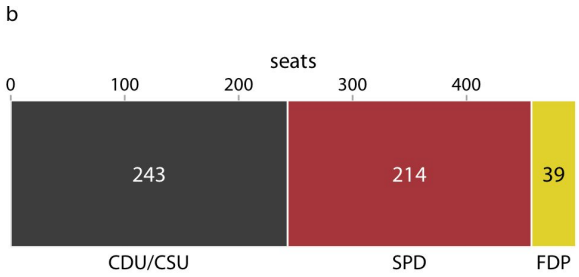
Alternatives?



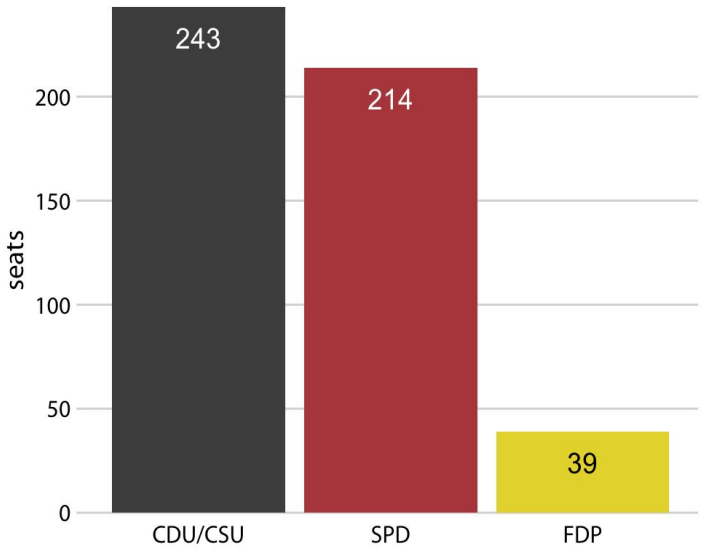
Pie chart

Stacked bars





Bar chart



Wilke, Claus O. *Fundamentals of data visualization: a primer on making informative and compelling figures*. O'Reilly Media, 2019.

Pie Chart Alternatives

Table 10.1: Pros and cons of common approaches to visualizing proportions: pie charts, stacked bars, and side-by-side bars.

	Pie chart	Stacked bars	Side-by-side bars
Clearly visualizes the data as proportions of a whole	✔	✔	✖

Wilke, Claus O. *Fundamentals of data visualization: a primer on making informative and compelling figures*. O'Reilly Media, 2019.

Pie Chart Alternatives

Table 10.1: Pros and cons of common approaches to visualizing proportions: pie charts, stacked bars, and side-by-side bars.

	Pie chart	Stacked bars	Side-by-side bars
Clearly visualizes the data as proportions of a whole	✔	✔	✖
Allows easy visual comparison of the relative proportions	✖	✖	✔

Wilke, Claus O. *Fundamentals of data visualization: a primer on making informative and compelling figures*. O'Reilly Media, 2019.

Pie Chart Alternatives

Table 10.1: Pros and cons of common approaches to visualizing proportions: pie charts, stacked bars, and side-by-side bars.

	Pie chart	Stacked bars	Side-by-side bars
Clearly visualizes the data as proportions of a whole	✔	✔	✖
Allows easy visual comparison of the relative proportions	✖	✖	✔
Visually emphasizes simple fractions, such as 1/2, 1/3, 1/4	✔	✖	✖

Wilke, Claus O. *Fundamentals of data visualization: a primer on making informative and compelling figures*. O'Reilly Media, 2019.

Pie Chart Alternatives

Table 10.1: Pros and cons of common approaches to visualizing proportions: pie charts, stacked bars, and side-by-side bars.

	Pie chart	Stacked bars	Side-by-side bars
Clearly visualizes the data as proportions of a whole	✔	✔	✖
Allows easy visual comparison of the relative proportions	✖	✖	✔
Visually emphasizes simple fractions, such as 1/2, 1/3, 1/4	✔	✖	✖
Looks visually appealing even for very small datasets	✔	✖	✔

Wilke, Claus O. *Fundamentals of data visualization: a primer on making informative and compelling figures*. O'Reilly Media, 2019.

Pie Chart Alternatives

Table 10.1: Pros and cons of common approaches to visualizing proportions: pie charts, stacked bars, and side-by-side bars.

	Pie chart	Stacked bars	Side-by-side bars
Clearly visualizes the data as proportions of a whole	✔	✔	✖
Allows easy visual comparison of the relative proportions	✖	✖	✔
Visually emphasizes simple fractions, such as 1/2, 1/3, 1/4	✔	✖	✖
Looks visually appealing even for very small datasets	✔	✖	✔
Works well when the whole is broken into many pieces	✖	✖	✔

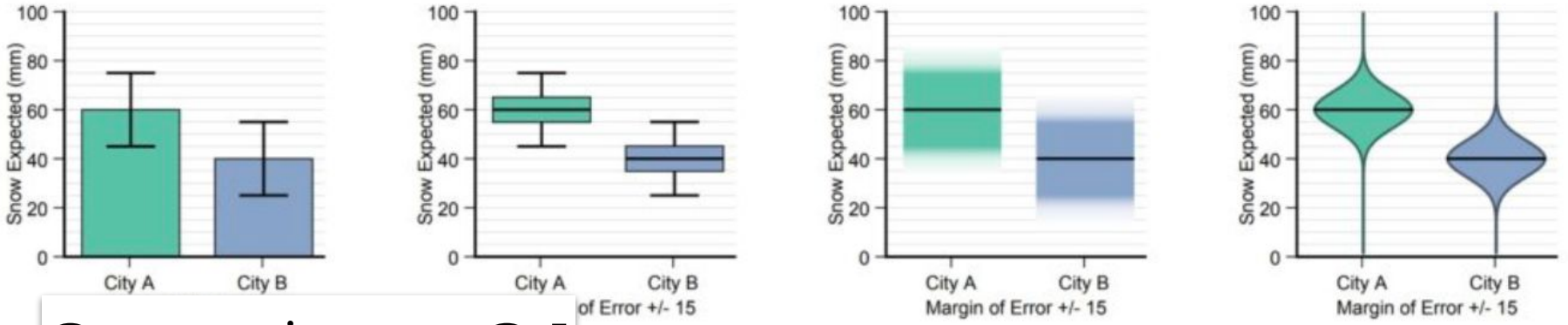
Wilke, Claus O. *Fundamentals of data visualization: a primer on making informative and compelling figures*. O'Reilly Media, 2019.

Pie Chart Alternatives

Table 10.1: Pros and cons of common approaches to visualizing proportions: pie charts, stacked bars, and side-by-side bars.

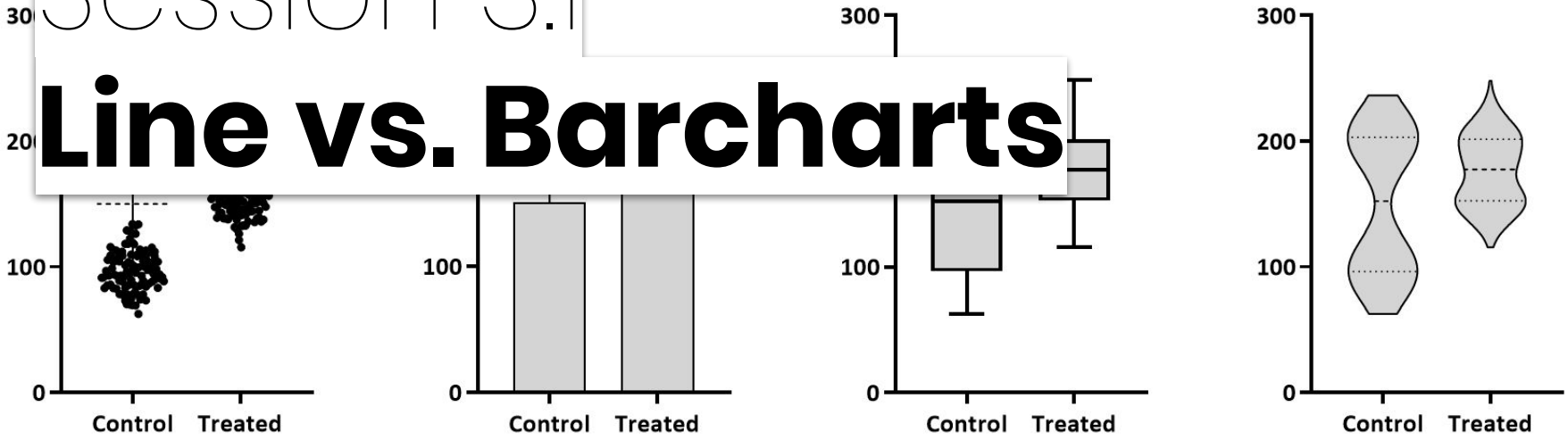
	Pie chart	Stacked bars	Side-by-side bars
Clearly visualizes the data as proportions of a whole	✔	✔	✖
Allows easy visual comparison of the relative proportions	✖	✖	✔
Visually emphasizes simple fractions, such as 1/2, 1/3, 1/4	✔	✖	✖
Looks visually appealing even for very small datasets	✔	✖	✔
Works well when the whole is broken into many pieces	✖	✖	✔
Works well for the visualization of many sets of proportions or time series of proportions	✖	✔	✖

Wilke, Claus O. *Fundamentals of data visualization: a primer on making informative and compelling figures*. O'Reilly Media, 2019.



Session 3.1

Line vs. Barcharts



Online Course
Data Visualization for Professionals

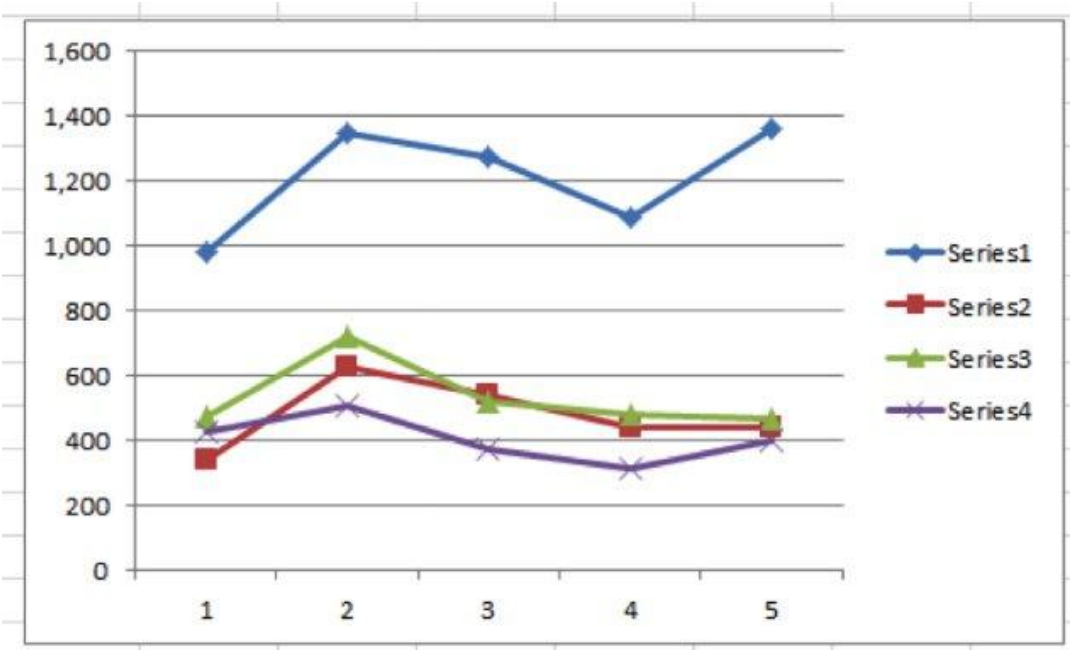
THE UNIVERSITY of EDINBURGH

Benjamin Bach
June 2020
http://benjbach.me
https://datavis-online.github.io

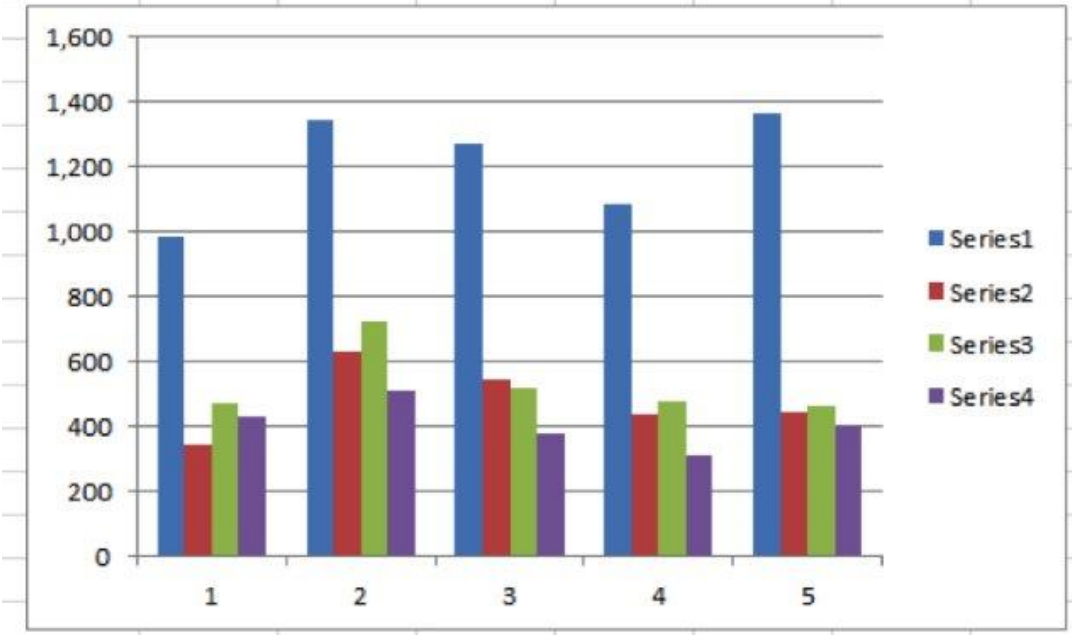
-- Not for external use --

Line Charts vs. Bar Charts?

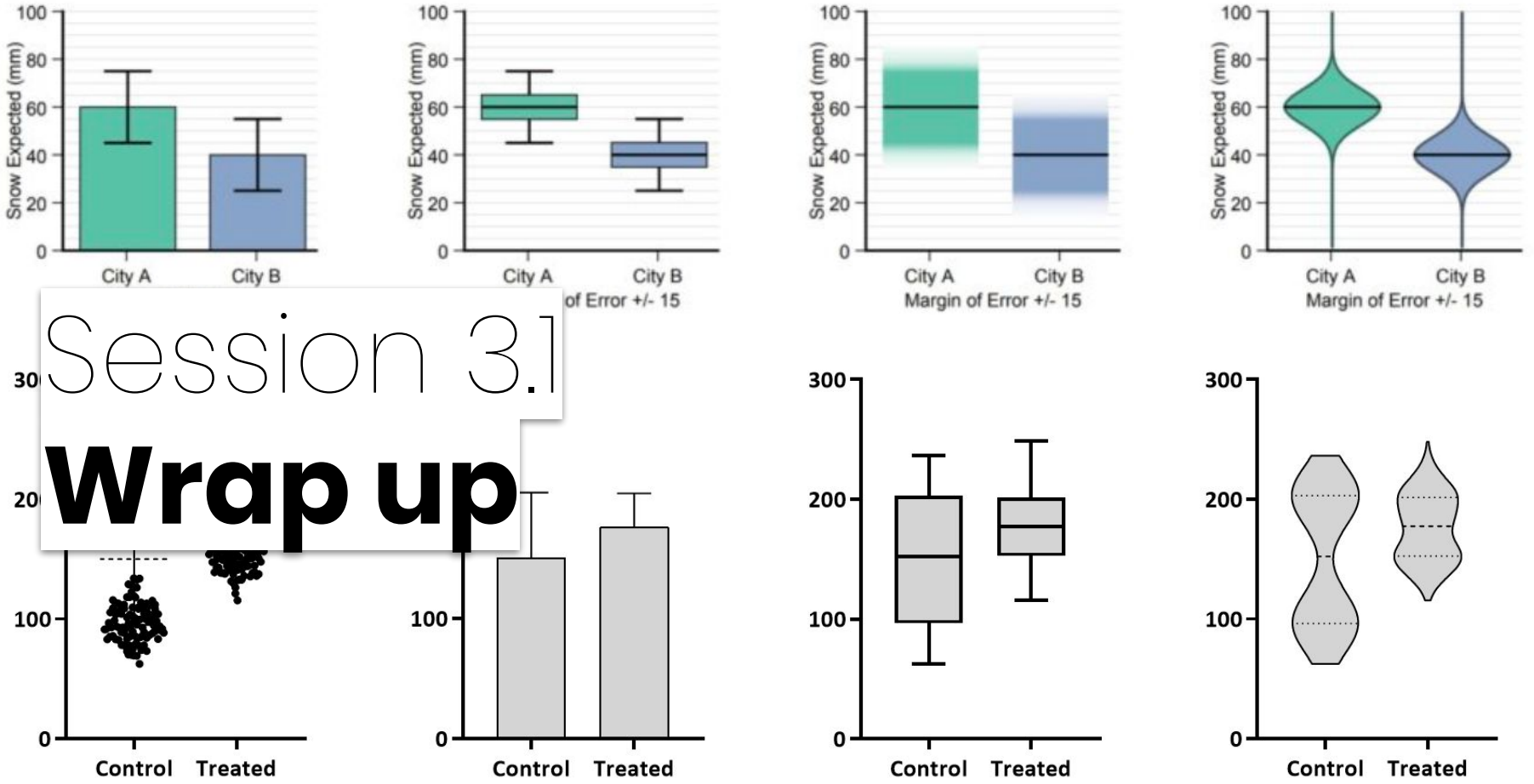
Continuous values
e.g., time series



Discrete values
e.g., countries



https://online-behavior.com/analytics/data-visualization



Session 3.1

Wrap up

Online Course
Data Visualization for Professionals

THE UNIVERSITY of EDINBURGH

Benjamin Bach

June 2020

http://benjbach.me

https://datavis-online.github.io

-- Not for external use --

Wrap up

- Choose appropriate bin-sizes
- violin plots for compare distributions
- Label error bars
- Don't confound box plots and error bars
- Avoid pie charts
- Line charts for continuous data

Further Readings

- Wilke, Claus O. **Fundamentals of data visualization: a primer on making informative and compelling figures**. O'Reilly Media, 2019.
  - https://serialmentor.com/dataviz/visualizing-proportions.htm
  - https://serialmentor.com/dataviz/histograms-density-plots.htmll
  - https://serialmentor.com/dataviz/boxplots-violins.html
- Hullman, Jessica, Paul Resnick, and Eytan Adar. "**Hypothetical outcome plots outperform error bars and violin plots for inferences about reliability of variable ordering**." *PloS one* 10.11 (2015).
- Skau, Drew, and Robert Kosara. "**Arcs, angles, or areas: Individual data encodings in pie and donut charts**." *Computer Graphics Forum*. Vol. 35. No. 3. 2016.
- Cairo, Alberto. **The truthful art: data, charts, and maps for communication**. New Riders, 2016.
  - Chapter 6: Exploring Data with Simple Charts
  - Chapter 7: Visualizing Distributions