

# Module: Displaying Data

## Introduction to Statistics in Kinesiology

Furtado JR., O

Cal State Northridge

updated: 2022-02-22



```
xaringanExtra::use_webcam()
```





It is not knowledge, but the act of learning, not possession but the act of getting there, which grants the greatest enjoyment.

— Gauss, C.

# Overview



Topics to be covered:

# Data set



Download the data set from the link below. You will need the data set to complete the **jamovi** exercises.

Data set: <https://bit.ly/iskdata>

Note: for more information about the data set above, click visit link below:

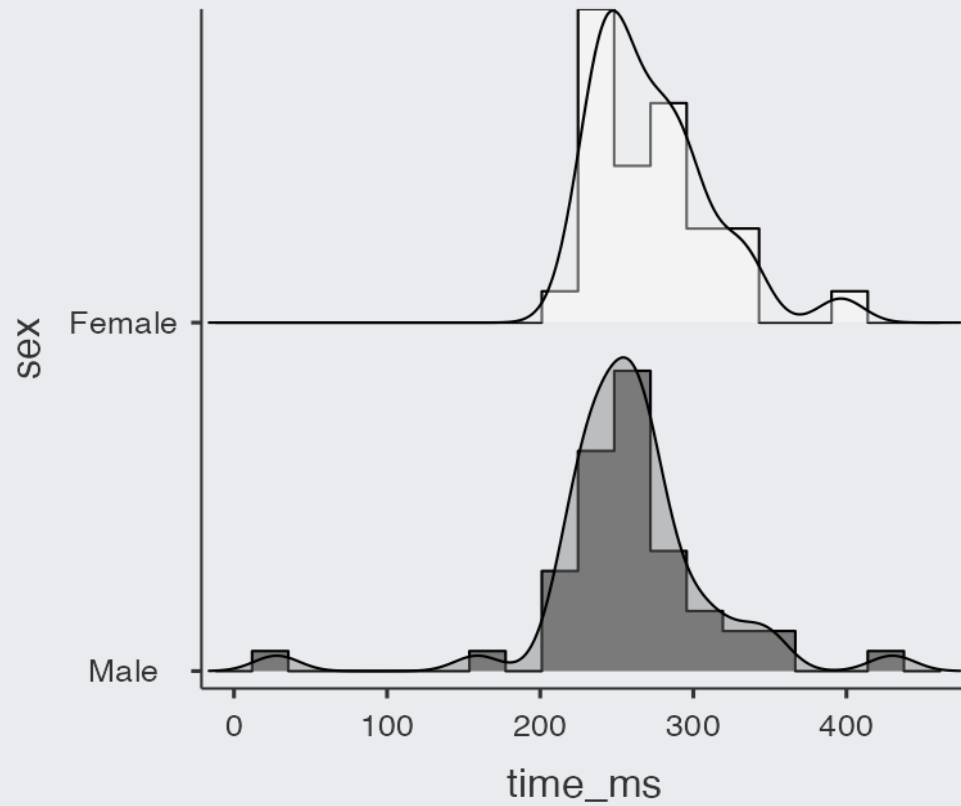


# Histograms

- simplest and most useful ways of visualizing data
  - overall impression of the data set
  - use when data are interval or ratio
- possible values are divided up into bins and then count the number of observations that fall within each bin

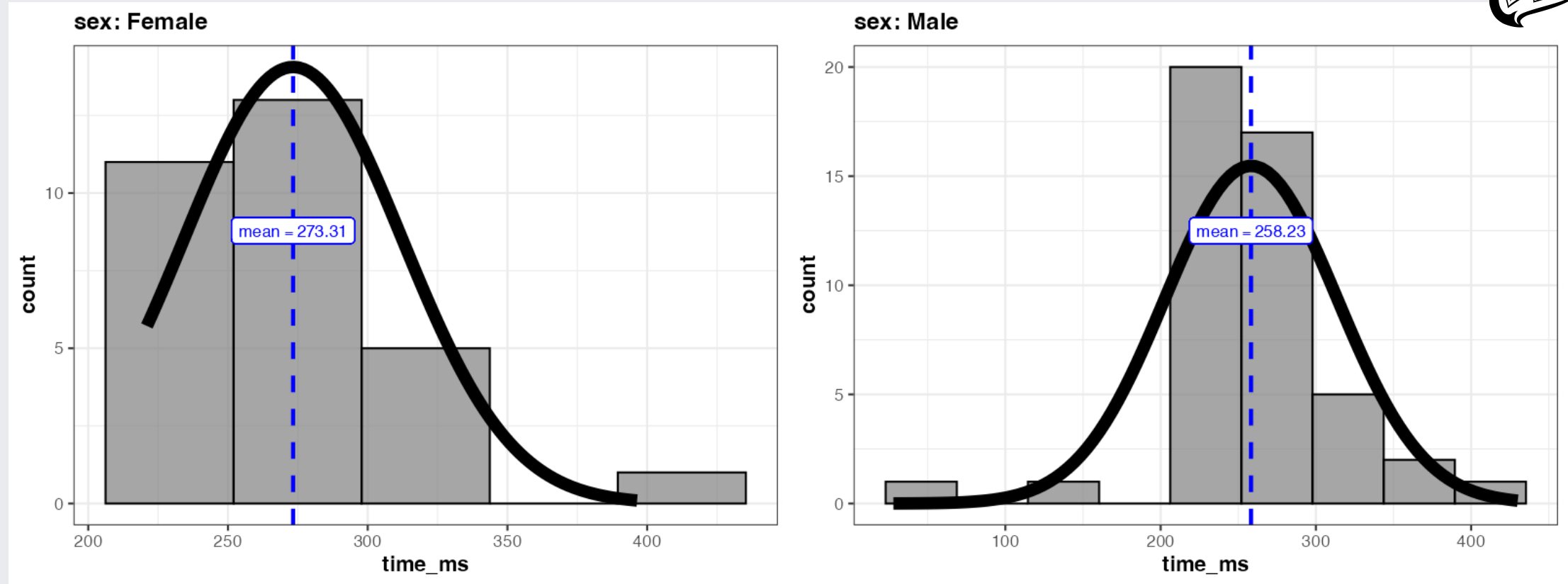


# Histograms created using the core package in jamovi





# Histograms created using the jjstatsplot module in jamovi





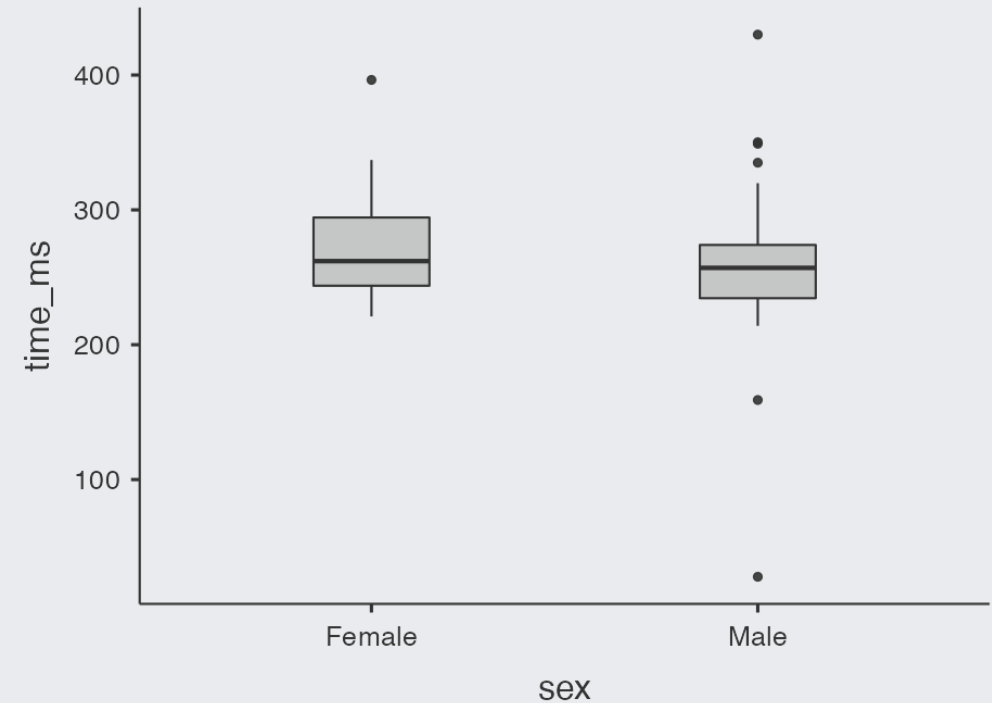
# Histogram - jamovi practice

- Open the `iskdata` with jamovi
  - **Exploration** --> **Descriptives** and click the **Histogram** check box
  - Check **Density** for a density curve - allows for a smoother distribution
- Observations that falls outside this range is plotted as a circle or dot instead of being covered by the whiskers, and is commonly referred to as an outlier.



# Traditional Boxplots

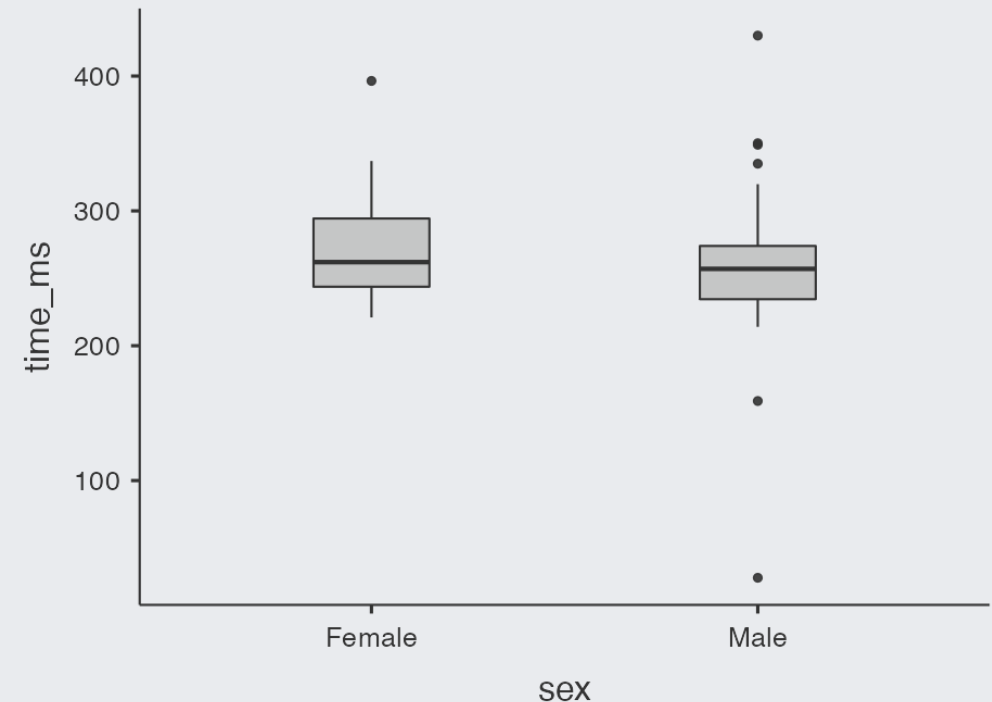
- suited for `interval` or `ratio` scale data
- `median`: line in the middle of the box
- box upper and lower bounds: 75th and 25th percentiles, respectively





# Traditional Boxplots, cont

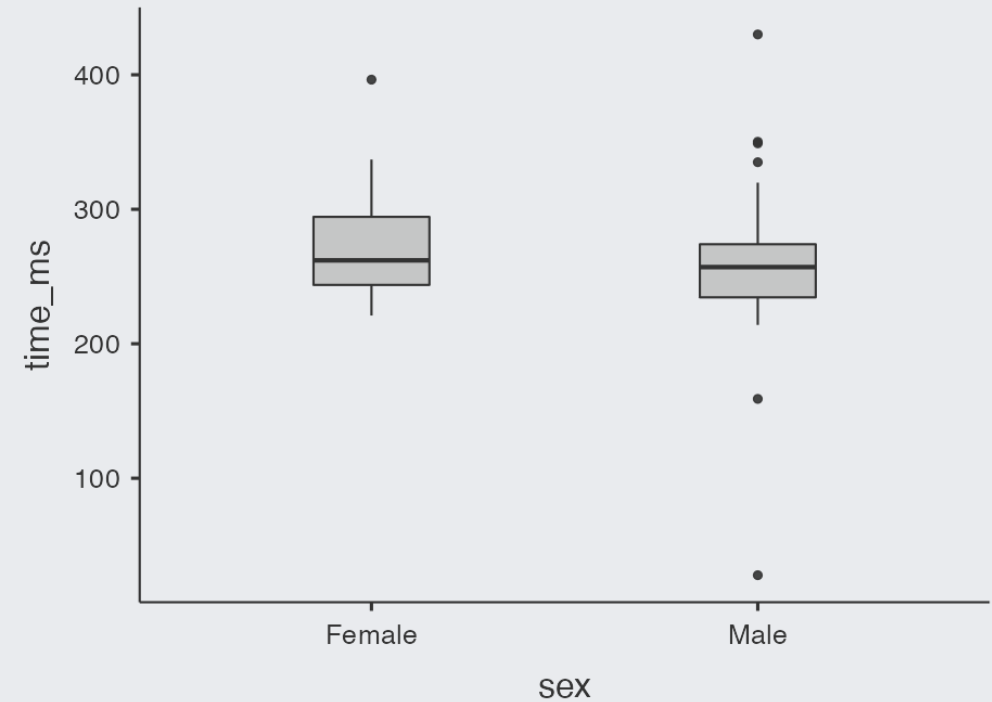
- two ends of the vertical lines: most extreme data points but do not exceed a certain bound
- calculated as 25th percentile -  $(1.5 \text{ IQR})$  for the lower boundary, and 75th percentile +  $(1.5 \text{ IQR})$  for the upper boundary
- observations lying outside of these "fences" can be considered outliers



## using filters in jamovi to inspect outliers



- focus on one particular case = below 100ms -filter data to inspect value

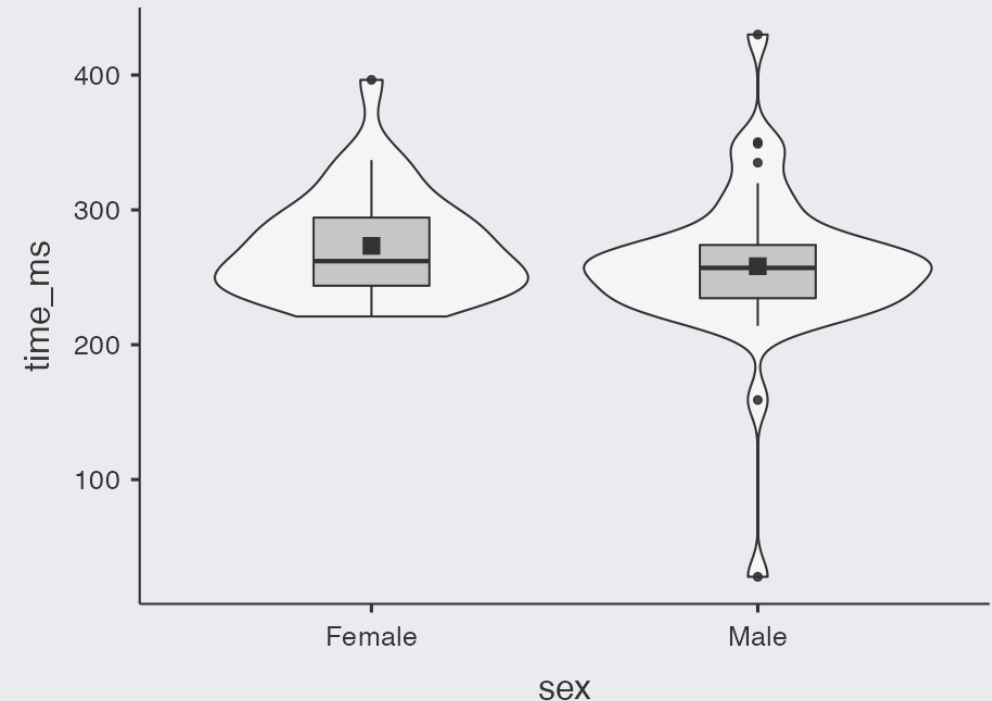




# Violin Boxplots

similar to traditional boxplots, but

- shows the kernel probability density (KPD) of the data at different values
- median (horizontal bar inside the box); square dot (mean)



KPD estimation - fundamental data smoothing problem where inferences about the population are made, based on a finite data sample.

# Boxplots jamovi example



term

year

vision-aid

bmi

Variables

time\_ms

Split by

sex

Descriptives Variables across columns

Frequency tables

Statistics

Plots

Histograms

Histogram

Density

Q-Q Plots

Q-Q

Box Plots

Box plot

Violin

Data

Jittered

Mean

Bar Plots

Bar plot

- move variables (1,2)
- expand Plots(3)
- Check desired options (4)
  - for a traditional boxplot: uncheck Violin



# Bar graphs

- Outcome (DV) continuous variable goes on **y-axis**
- The IV discrete variable goes on the **x-axis**





# Contingency tables

A contingency table is a tabular representation of categorical data . A contingency table usually shows frequencies for particular combinations of values of two discrete random variables  $X$  and  $Y$ .

Each cell in the table represents a mutually exclusive combination of  $X$ - $Y$  values.

## References



- The jamovi project (2022). jamovi. (Version 2.3) [Computer Software]. Retrieved from <https://www.jamovi.org>.
- R Core Team (2021). R: A Language and environment for statistical computing. (Version 4.1) [Computer software]. Retrieved from <https://cran.r-project.org>. (R packages retrieved from MRAN snapshot 2022-01-01).
- Wickham, H., Chang, W., Henry, L., Pedersen, T. L., Takahashi, K., Wilke, C., Woo, K., & - RStudio (2018). ggplot2: Create Elegant Data Visualisations Using the Grammar of Graphics. [R package]. Retrieved from <https://CRAN.R-project.org/package=ggplot2>.
- Patil, I. (2018). ggstatsplot: 'ggplot2' Based Plots with Statistical Details. [R package]. Retrieved from <https://CRAN.R-project.org/package=ggstatsplot>.