Comparing 2 means

t-test

Mario E. Bermonti-Pérez, MA, PhD

```
************
Welcome to faux. For support and examples visit:
https://debruine.github.io/faux/
- Get and set global package options with: faux_options()
************
```

```
-- Attaching packages --------------------------------------- tidyverse 1.3.2 --
v ggplot2 3.3.6     v purrr   0.3.4
v tibble  3.1.8     v dplyr   1.0.9
v tidyr   1.2.0     v stringr 1.4.1
v readr   2.1.2     v forcats 0.5.2
-- Conflicts ------------------------------------------ tidyverse_conflicts() --
x purrr::%||%()   masks faux::%||%()
x dplyr::filter() masks stats::filter()
x dplyr::lag()    masks stats::lag()
```

Objectives

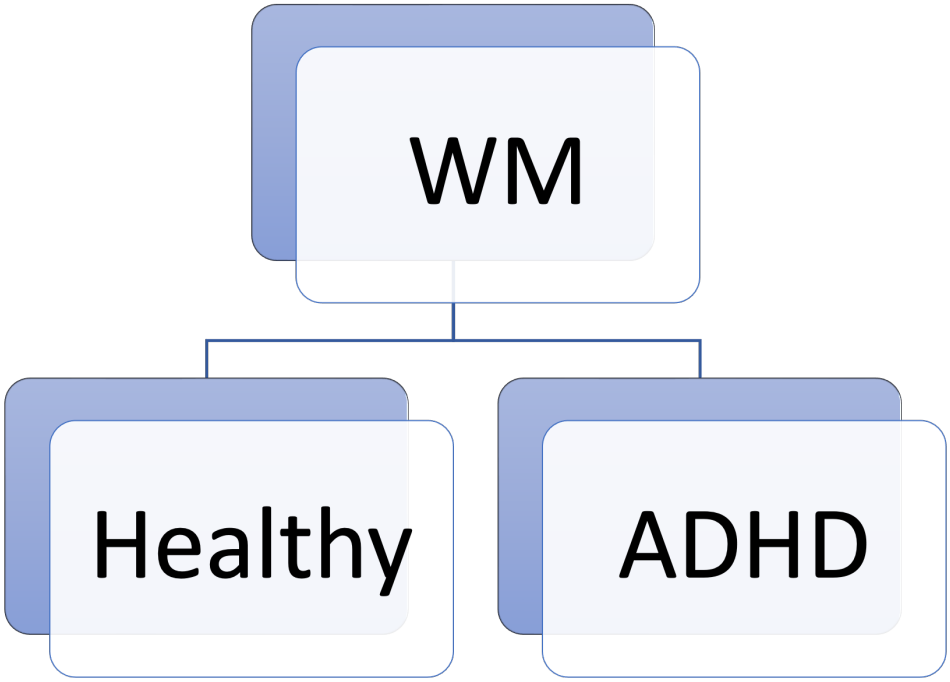
- Understand the usefulness of the independent t-test
- Understand the basic concepts of the independent t-test
- Conduct the independent t-test
- Interpret the results of the independent t-test

t-test

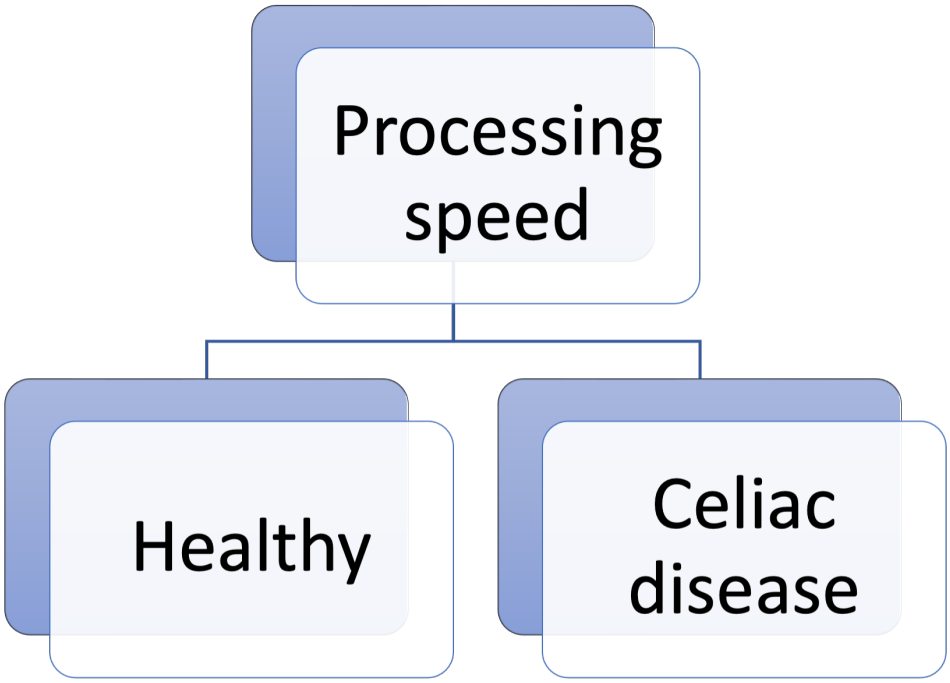
What is a t-test?

Basics

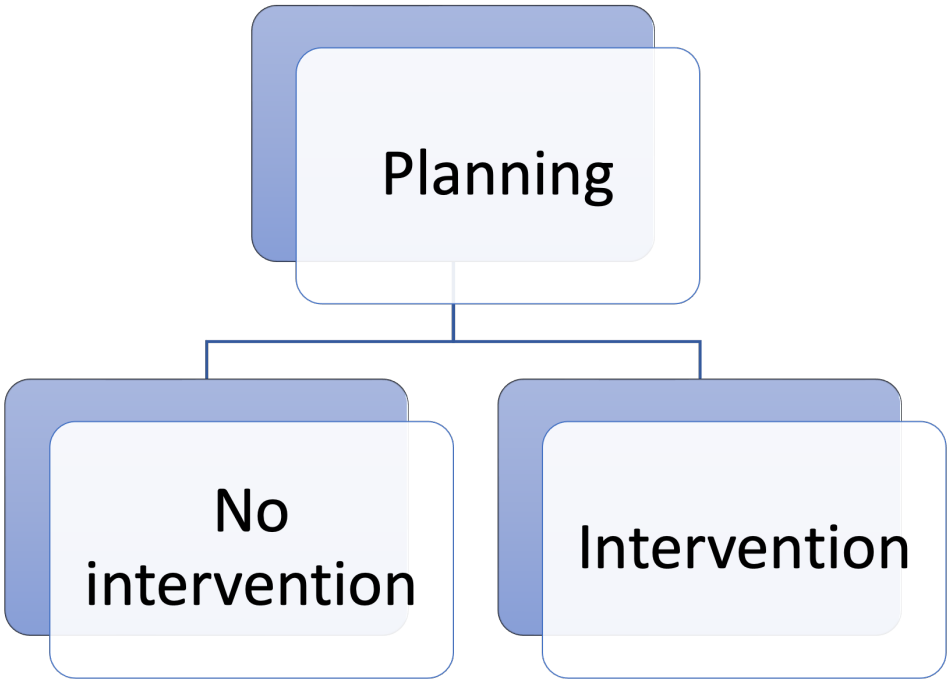
Usefulness



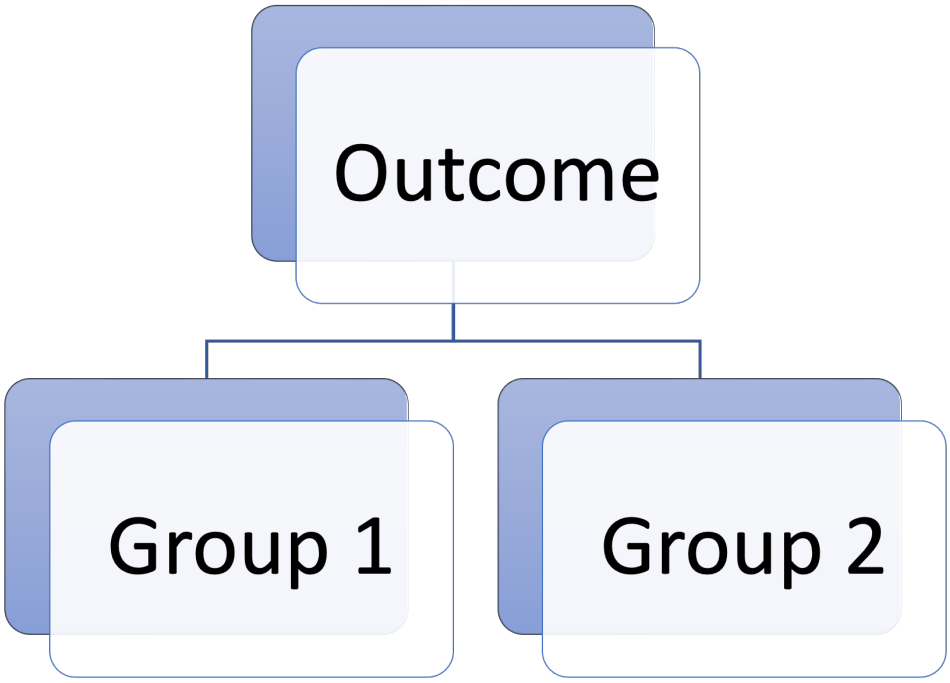
Usefulness



Usefulness

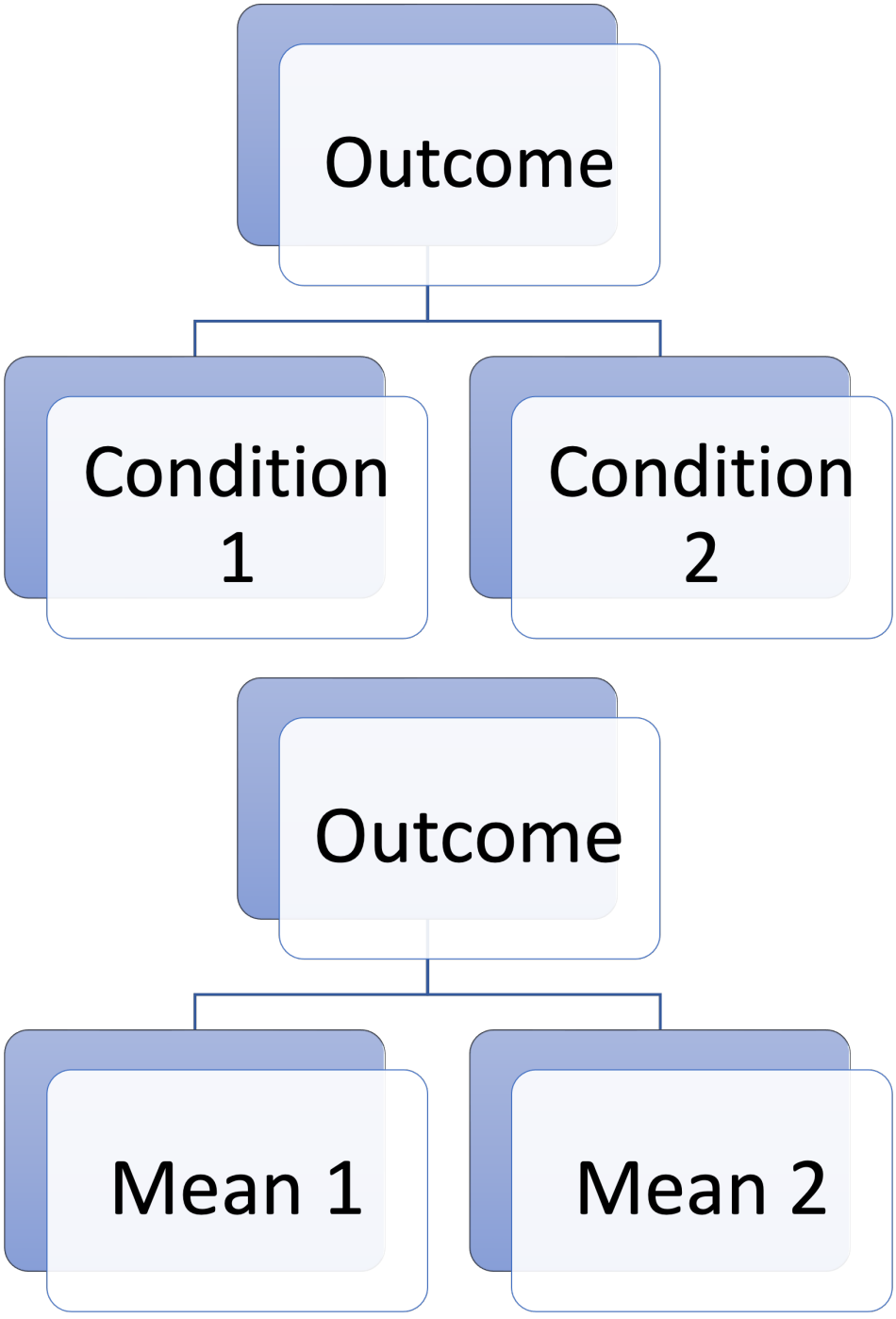


General form



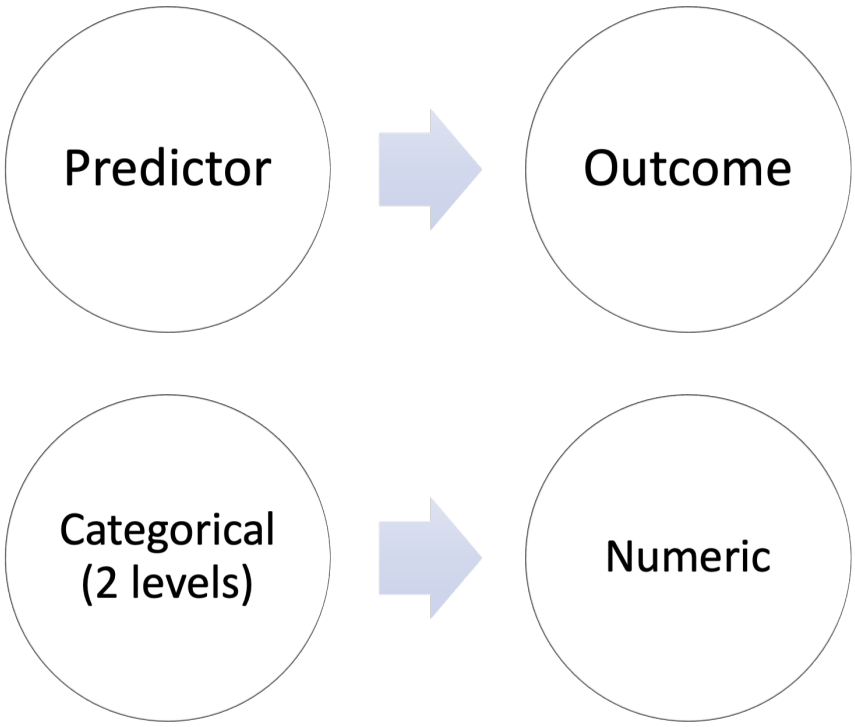
Compare 2 groups

Even more general

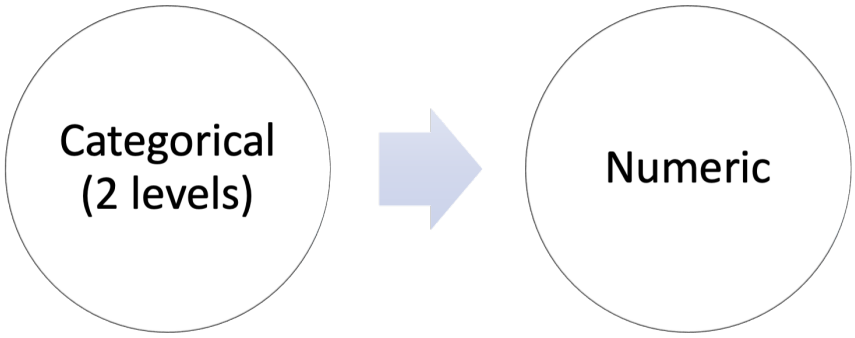


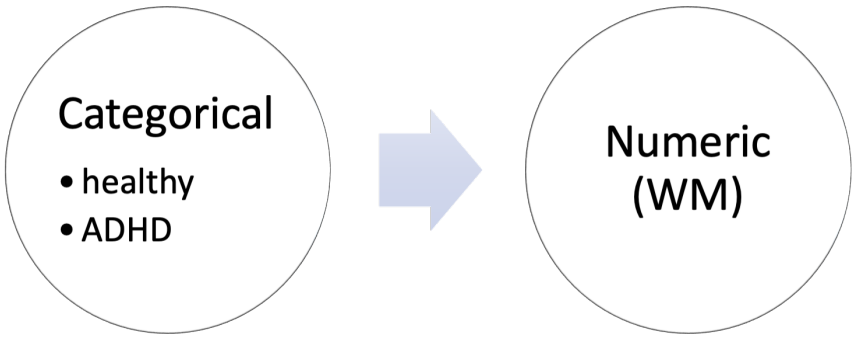
More general, powerful

GLM form

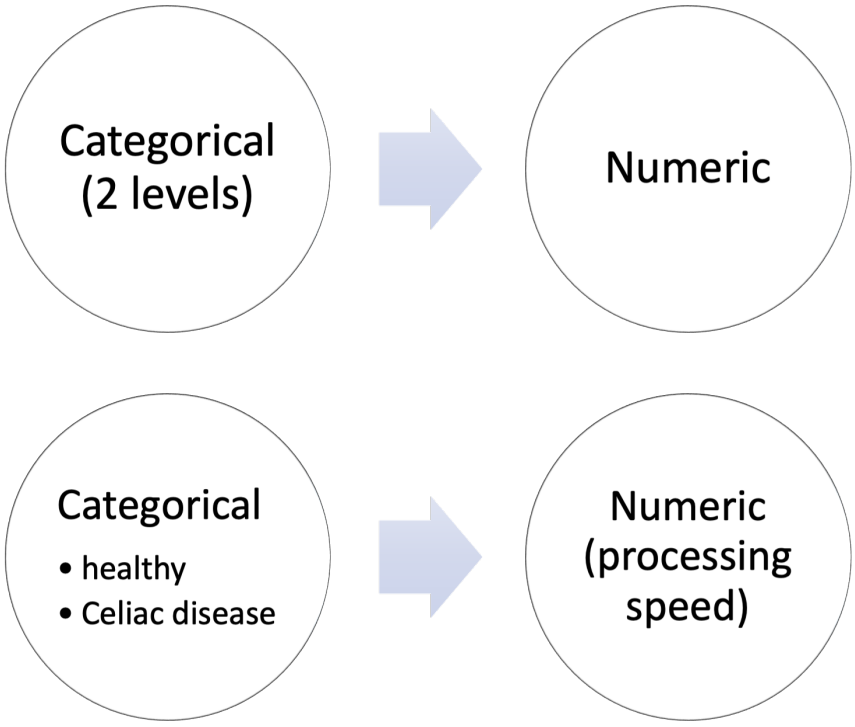


GLM form - examples

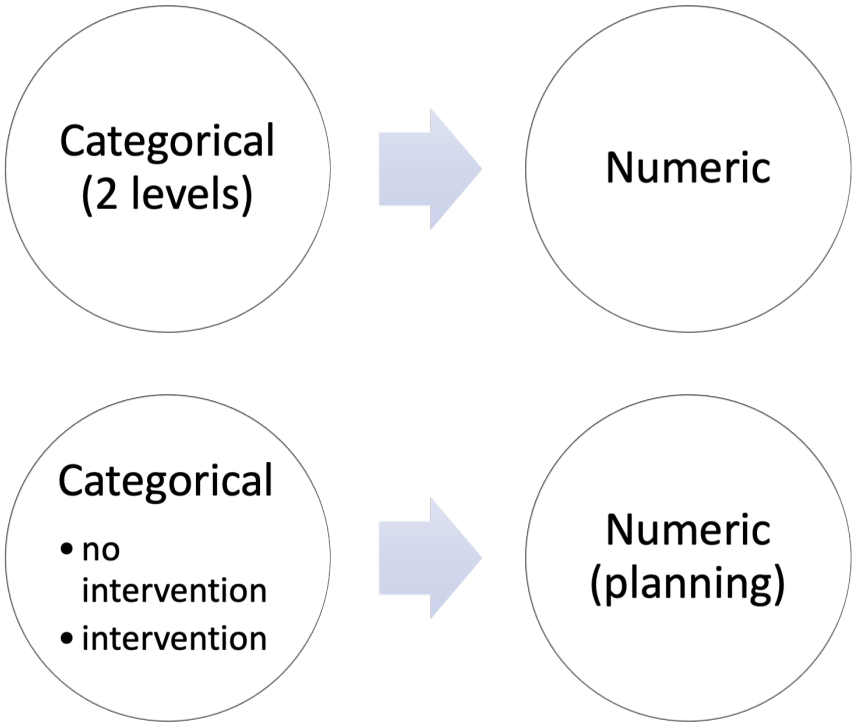




GLM form - examples



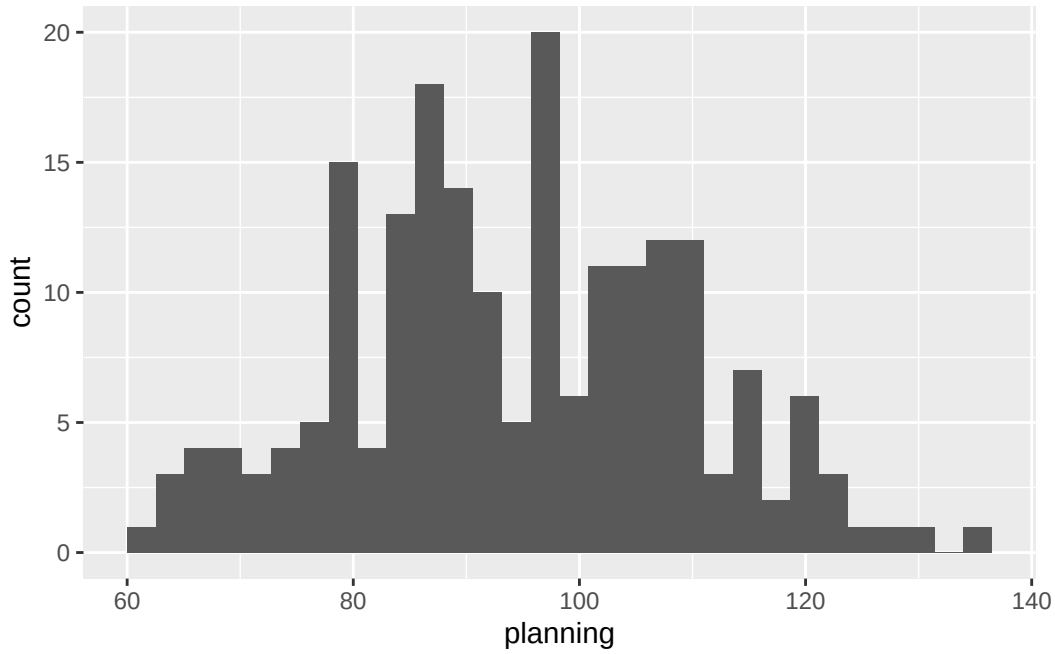
GLM form - examples



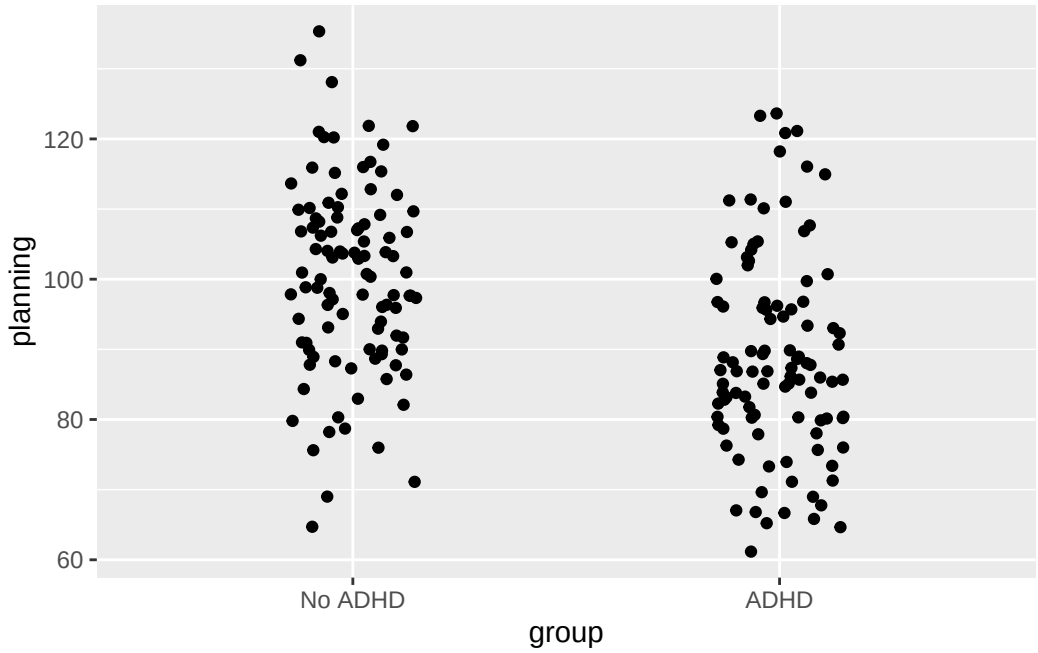
General process (conceptual)

Get data

`stat_bin()` using `bins = 30`. Pick better value with `binwidth`.



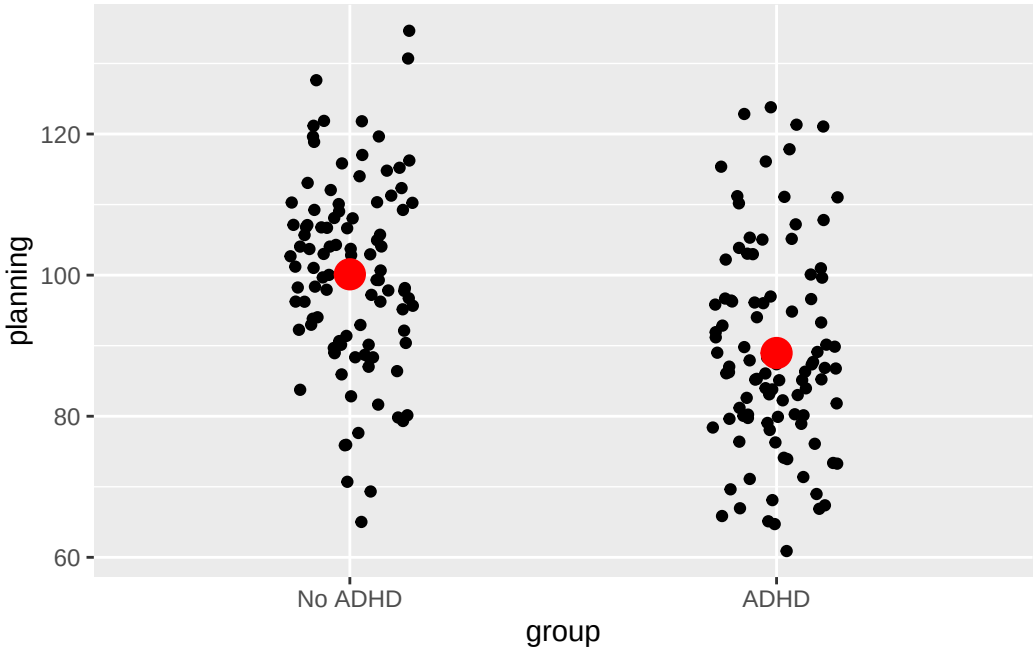
Group data



what is next?

Estimate means

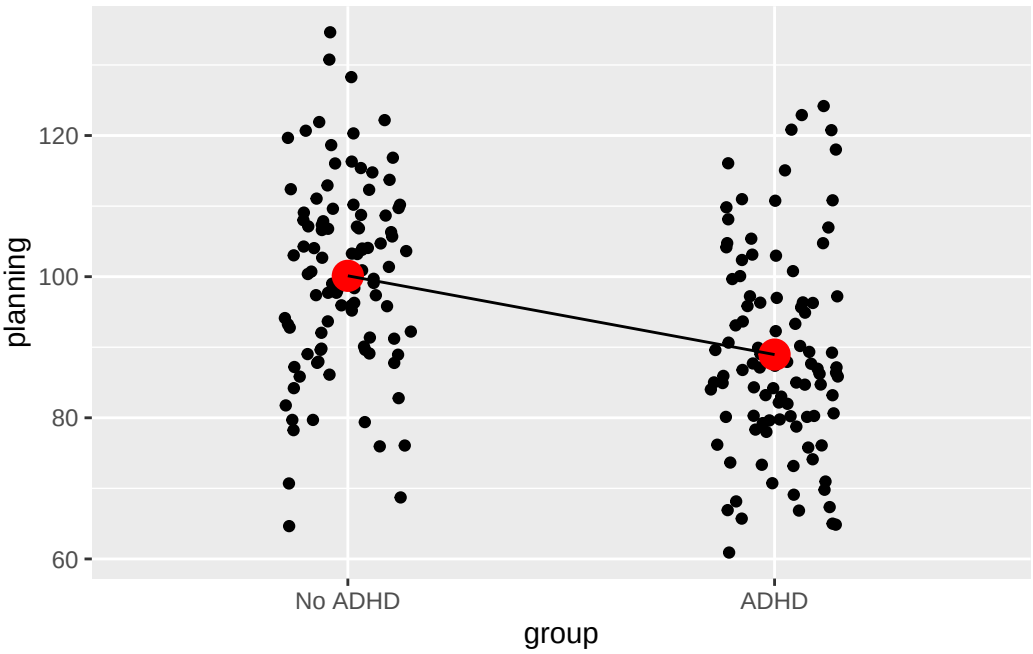
```
Warning: `fun.y` is deprecated. Use `fun` instead.
```



what is next?

Estimate effect (difference)

```
Warning: `fun.y` is deprecated. Use `fun` instead.
```



That is all folks!

Conducting and interpreting

Hands on exercises

Get ready!

Download data from teams

Study

Effect of ADHD on planning ability

Variables

- ADHD
  - healthy
  - ADHD
- Plannig: Standardized test

Standardized tests: mean? SD?

Hypothesis

- H0: No effect
- H1: Effect
- NHST - null = 0
- H0: Hardheaded

Info needed

- Effect size
- Statistical significance

What are these?

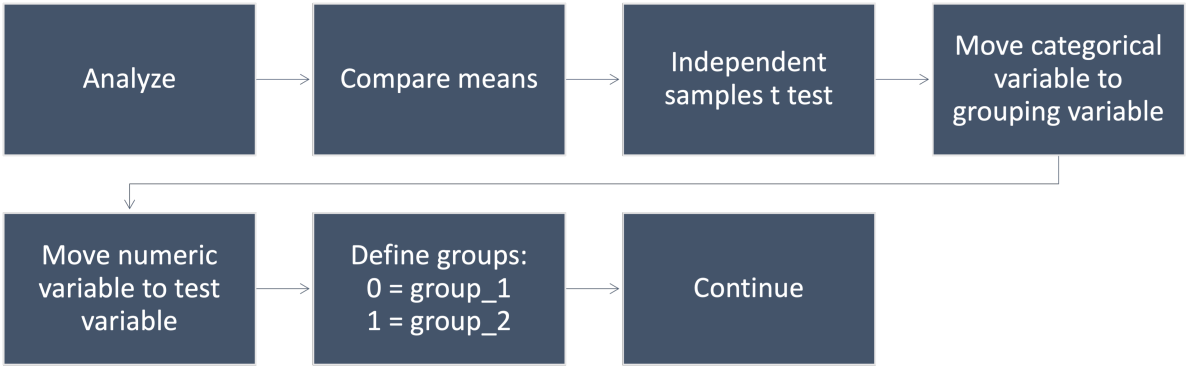
Viz

- Bar plot
- Box plot
- Which provides more info?
- Why use both? More info = better

Descriptive stats

- groupby
- detailed: mean, median, SD, Range

Conducting an independent t-test



Interpretation

Effect size

Unstandardized

- Mean difference
- 95% CI

Diff in raw scores?

Standardized

- Hedges g
- 95% CI

Interpretation

```
+ small: .2 - 0.49
+ medium: 0.5 - 0.79
+ large: 0.8+
```

- Diff in terms of SD?
- Why Hedges?

Statistical significance

Where to look

- Equal variances not assumed
- Significance, two sided

Interpretation

- p-value < 0.05 = Statistical significant
- p-value > 0.05 = not statistical significant

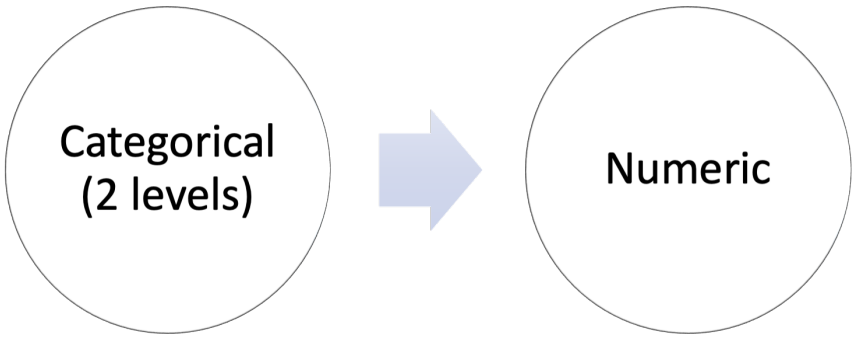
Assumptions

Assumptions

- Types of variables (GLM)
- Normal distribution
- Homegeinity of variance (Welch's test)
- No outliers

What are assumptions?

Variable types



Effect of ADHD on planning

- ADHD
- Planning

Do we meet them?

Normal distribution

- Histograms
- QQ plots
- QQ plots: quantiles
- group by condition
- normal: line
- devs: kurtosis, extremely values (at ends)
- S shape: skewness (negative = high scores, positive = low scores)

Outliers

- Boxplot

Do we meet them?

Closing

Recap

- What analysis?
- What is its use?
- Predictor type?
- Type of outcome?
- Pieces of info needed?
- Assumptions?

- t test
- compare means
- categorical

- numeric
- ES, stats sig
- var types, normal distribution, homo variances, outliers

Questions?

- What was easiest?
- What was hardest?