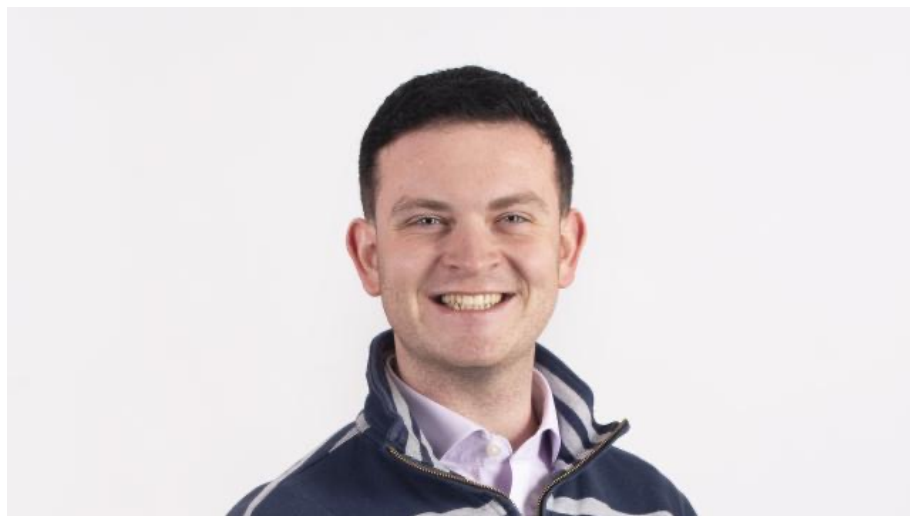


Introduction to Statistics

Welcome!



Chris Oldnall



Ki Tong



Descriptive vs Inferential

Descriptive statistics involve summarizing and presenting the main features of a dataset, such as measures of central tendency (mean, median, mode) and measures of variability (range, standard deviation).

Inferential statistics, aims to make predictions or draw conclusions about a population based on a sample of data, using probability theory and statistical inference techniques like hypothesis testing and confidence intervals.





THE UNIVERSITY OF EDINBURGH
Centre for Data, Culture & Society

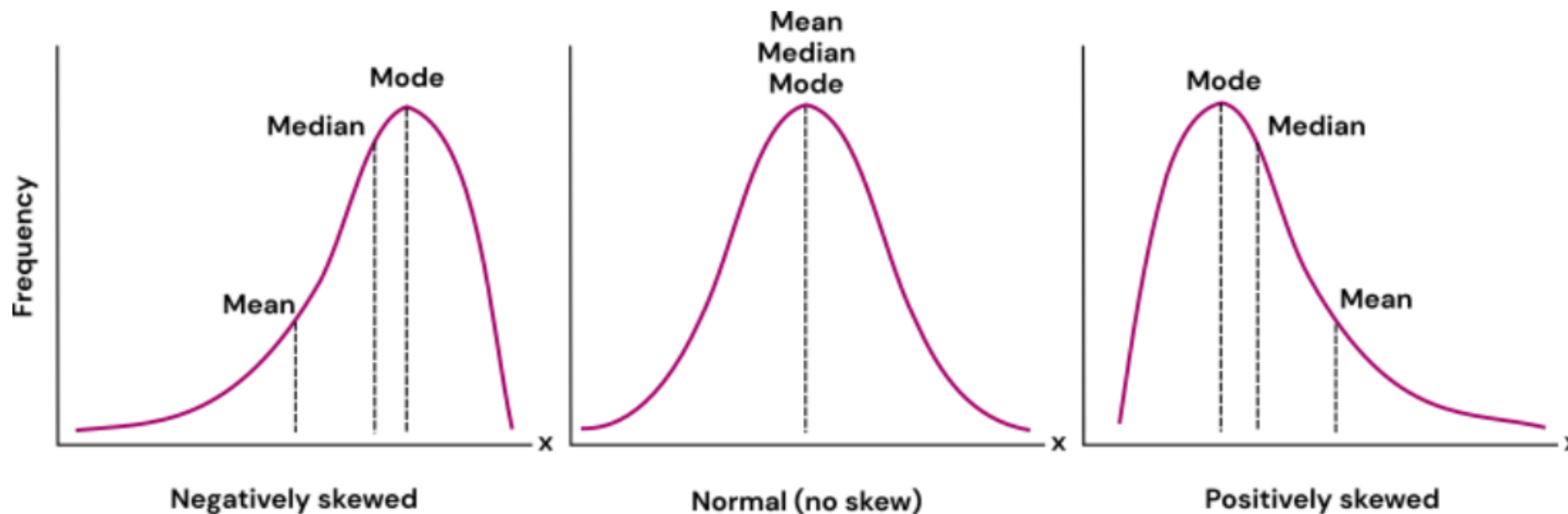


SESSION A: DESCRIPTIVE STATISTICS

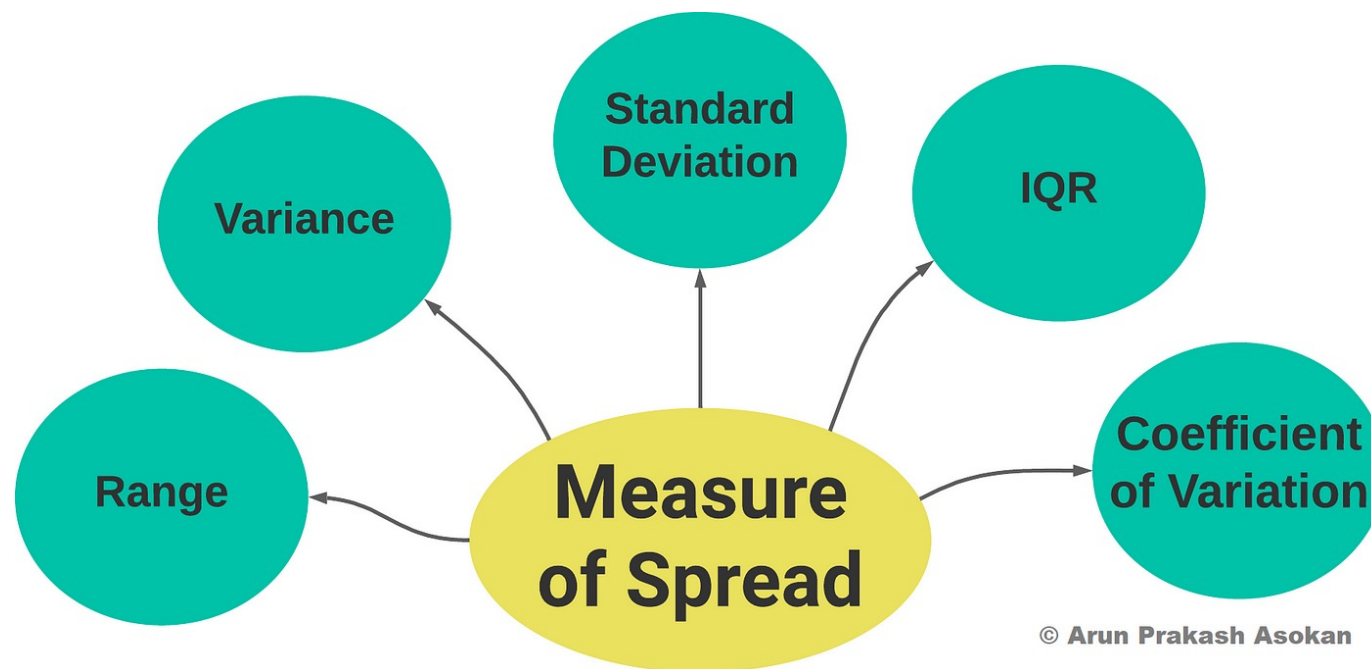


www.cdcs.ed.ac.uk

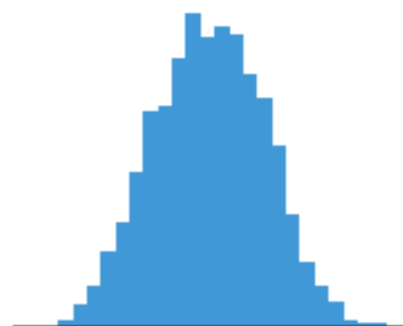
What is a measure of central tendency?



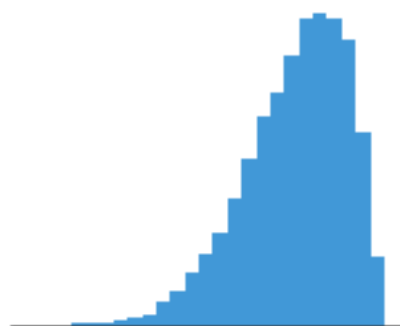
What is a measure of spread?



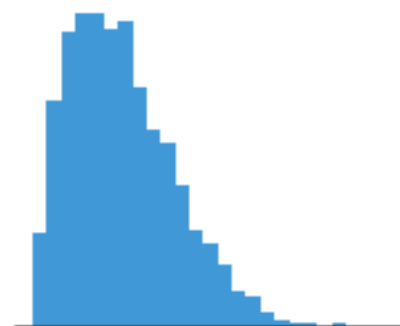
What is a distribution?



symmetric, unimodal



skew left



skew right



uniform



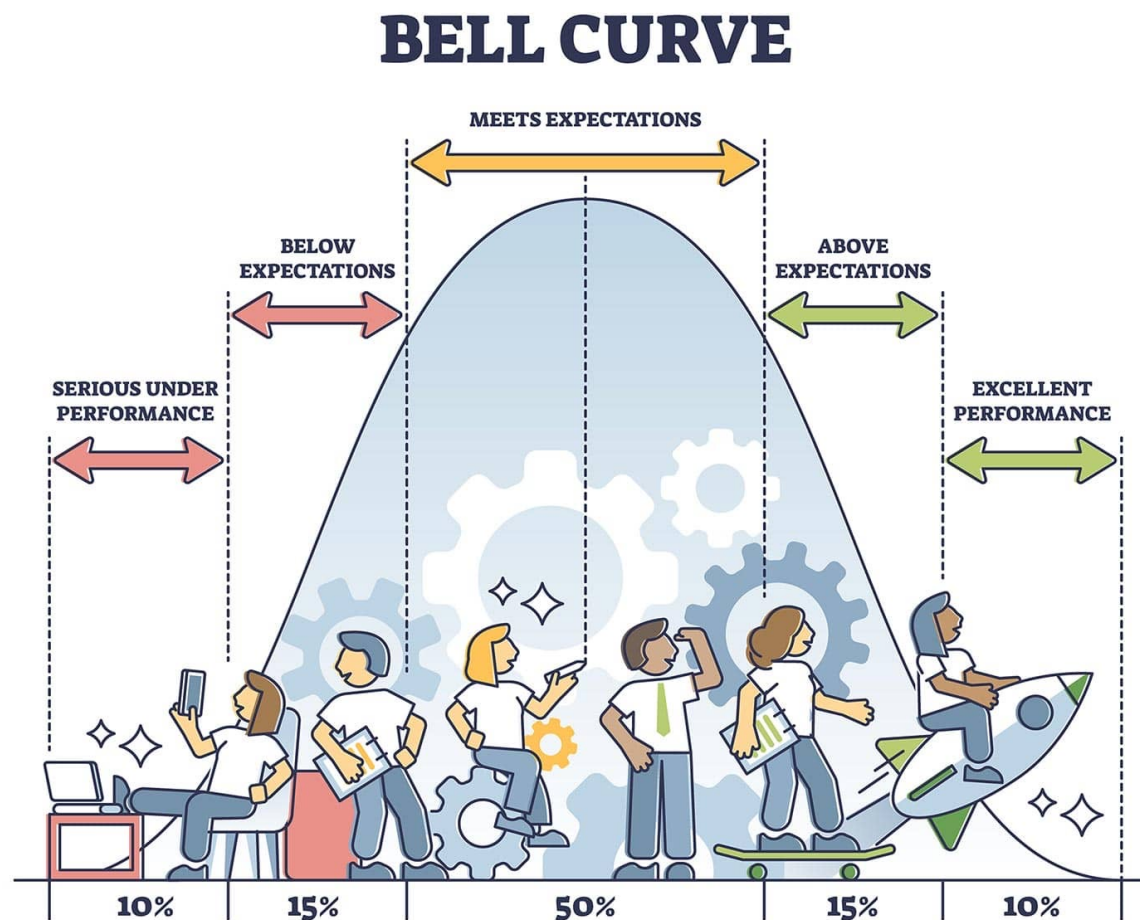
bimodal



multimodal



The normal distribution





Let's do some programming

Notebook A: What is a summary statistic?





THE UNIVERSITY OF EDINBURGH
Centre for Data, Culture & Society

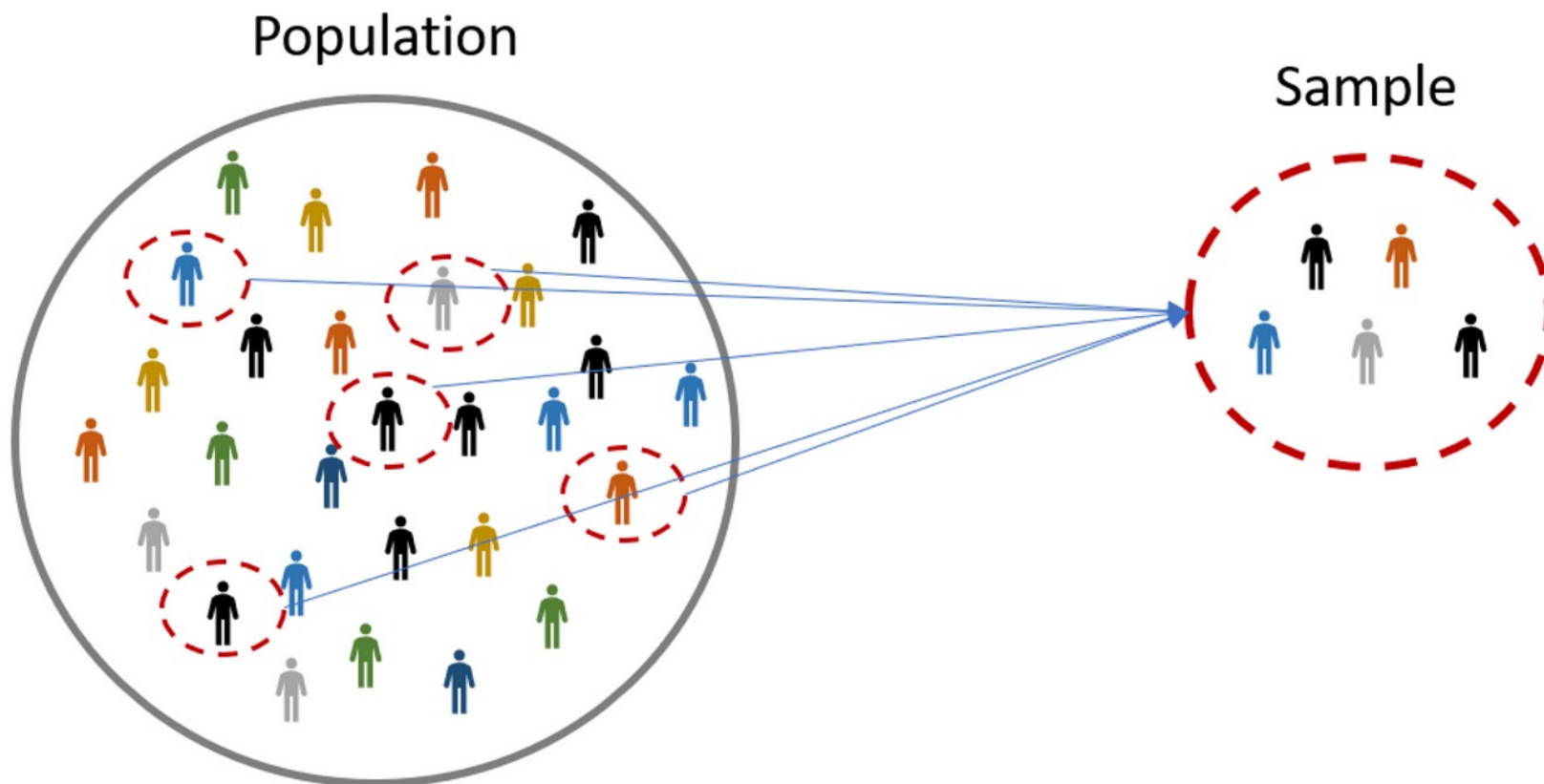


DATA
CULTURE
SOCIETY


SESSION B: INFERENTIAL STATISTICS



www.cdcs.ed.ac.uk




What is a hypothesis?



Hypothesis Testing

[hī-'pä-thə-səs 'te-stiŋ]

An act in statistics whereby an analyst tests an assumption regarding a population parameter.

 Investopedia

How do I formulate a hypothesis?

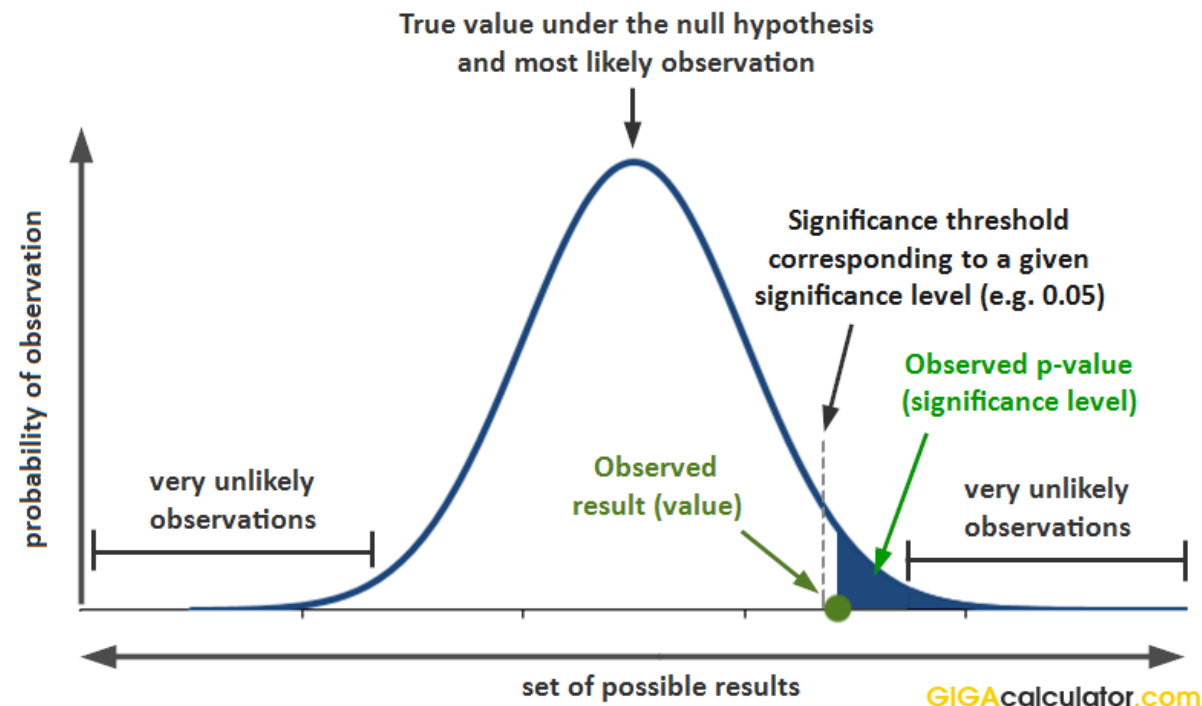
H0: The current status quo is this.

H1: I believe that it is this instead.

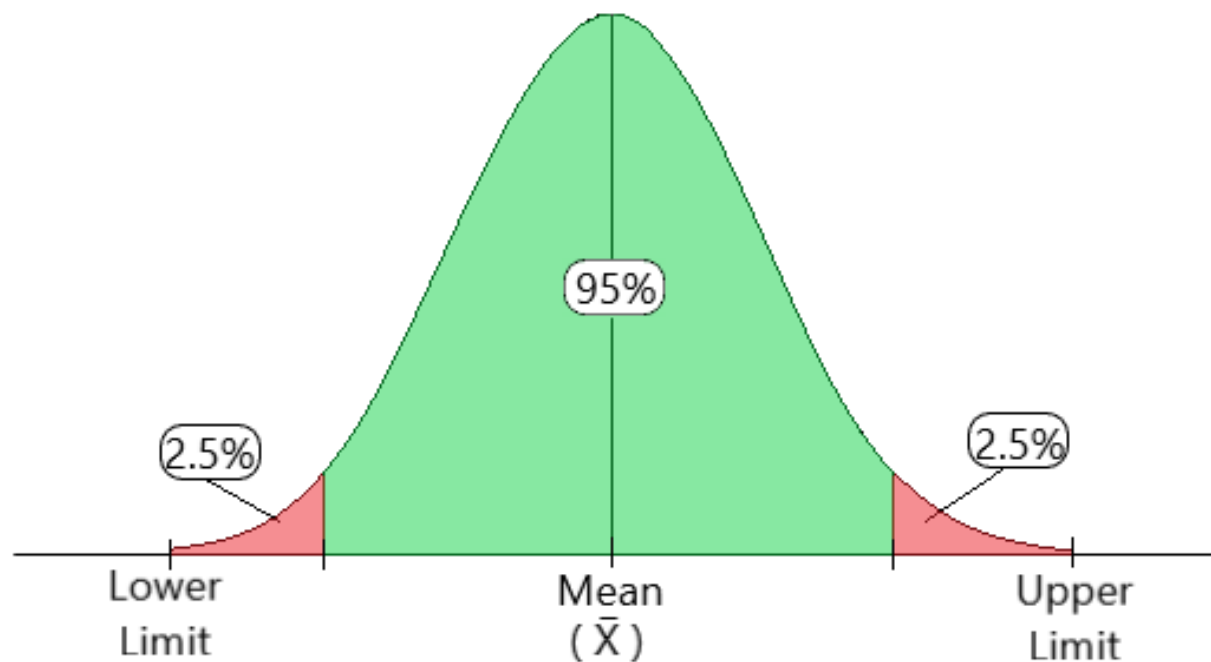


What is a p-value?

P-values and statistical significance explained



What is a confidence interval?



Testing a hypothesis.

1. A null and alternative hypothesis,
2. Some assumptions,
3. A test statistic,
4. Some significance threshold (typically 0.95)



BIG STATS POINT

REJECT the *NULL* if
 $p < 0.05$ or NOTHING



A difference in means = Paired T-Testing

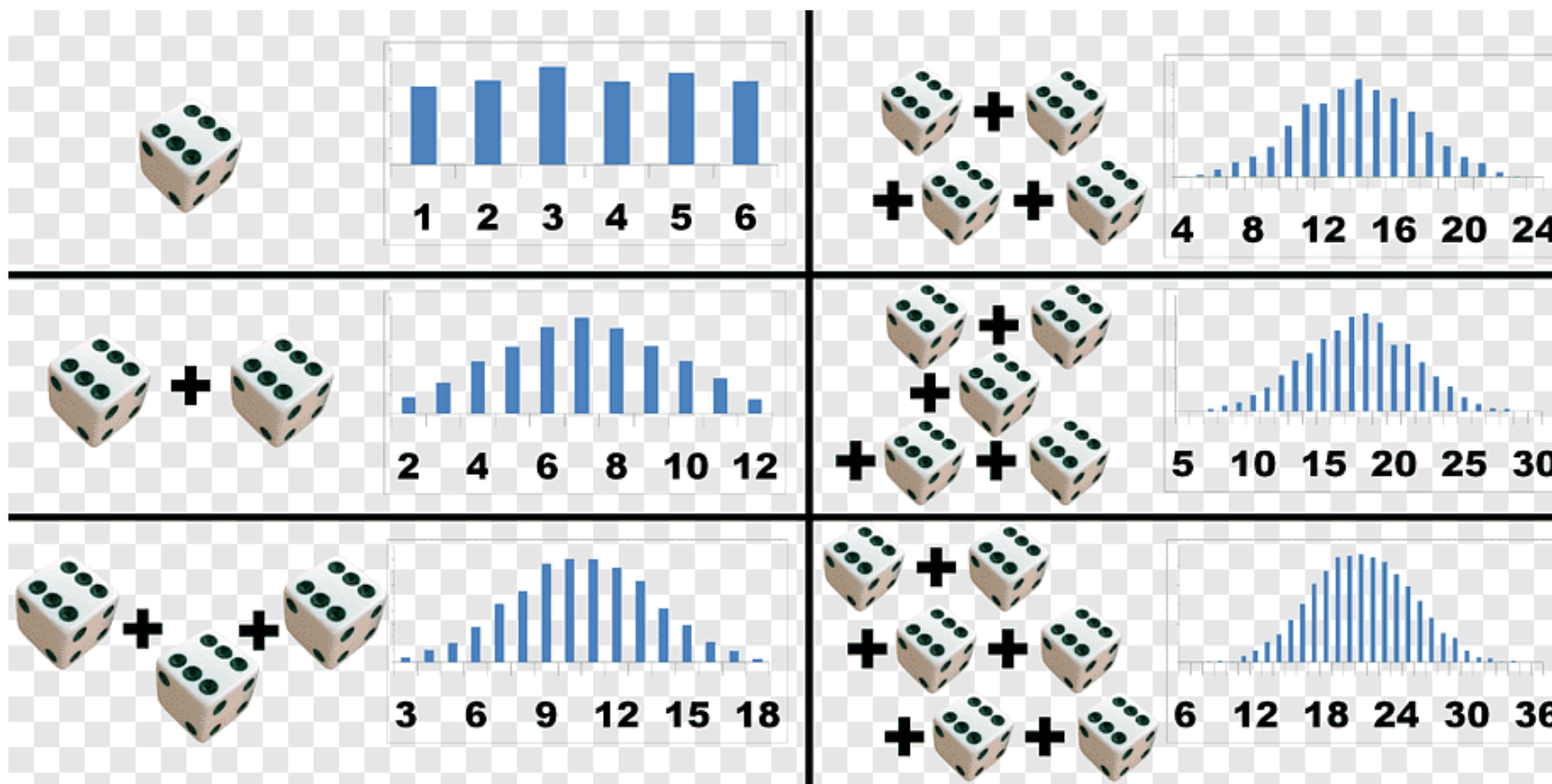
H0: My two groups have the same mean.

H1: My two groups of data have a different mean.

Assume: The data are independent and the means of the data are normally distributed



Central Limit Theorem



Central Limit Theorem

When I have a sample size over 30, I can assume that the mean of the data is normally distributed.



Let's do some programming

Notebook B: How do I make inferential statistics?



Feedback for us...

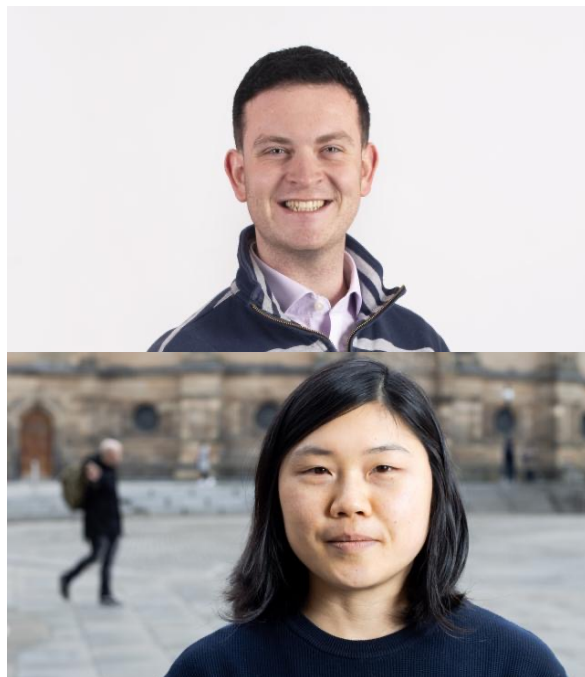
- We hope you've enjoyed the course as much as we did.
- It is really useful for us to hear your feedback

<https://forms.office.com/r/YYNrqvNr8>

Should be really quick and only take 5 mins (maximum!)



Thank you from us!



- **After the festive break we will be back with a range of course about using statistics. Come to the Digital Method of the Month in Jan/Feb to hear what is on offer.**

