Week 1: Introduction to Statistics

Unit 1: Introduction to Statistics

openSAP Course Overview

Weeks 1 through 6

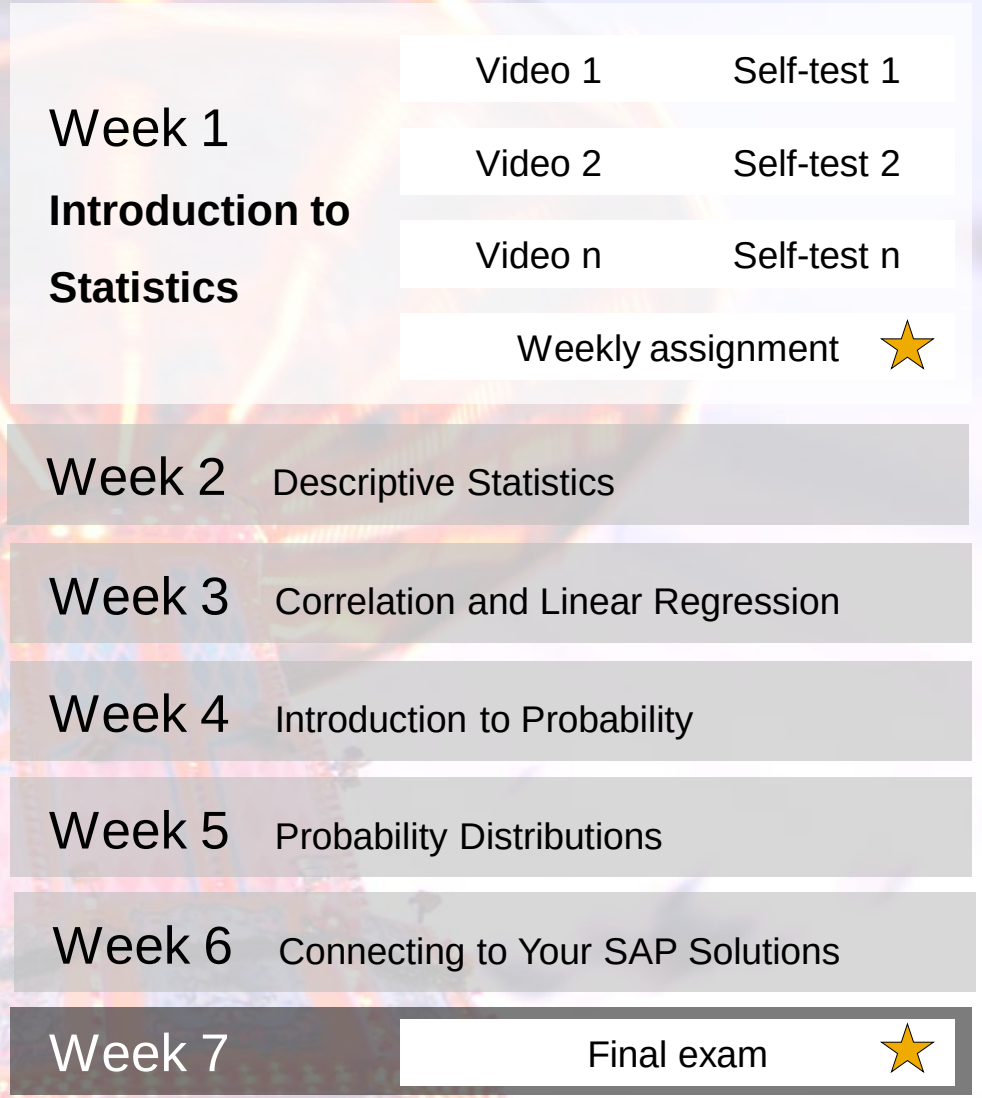
- 4-7 video units and self-tests
- 1 weekly assignment
- Online discussion forum (collaborate, ask questions)
- ~3-4 hours of effort each week

Week 7

- Final exam

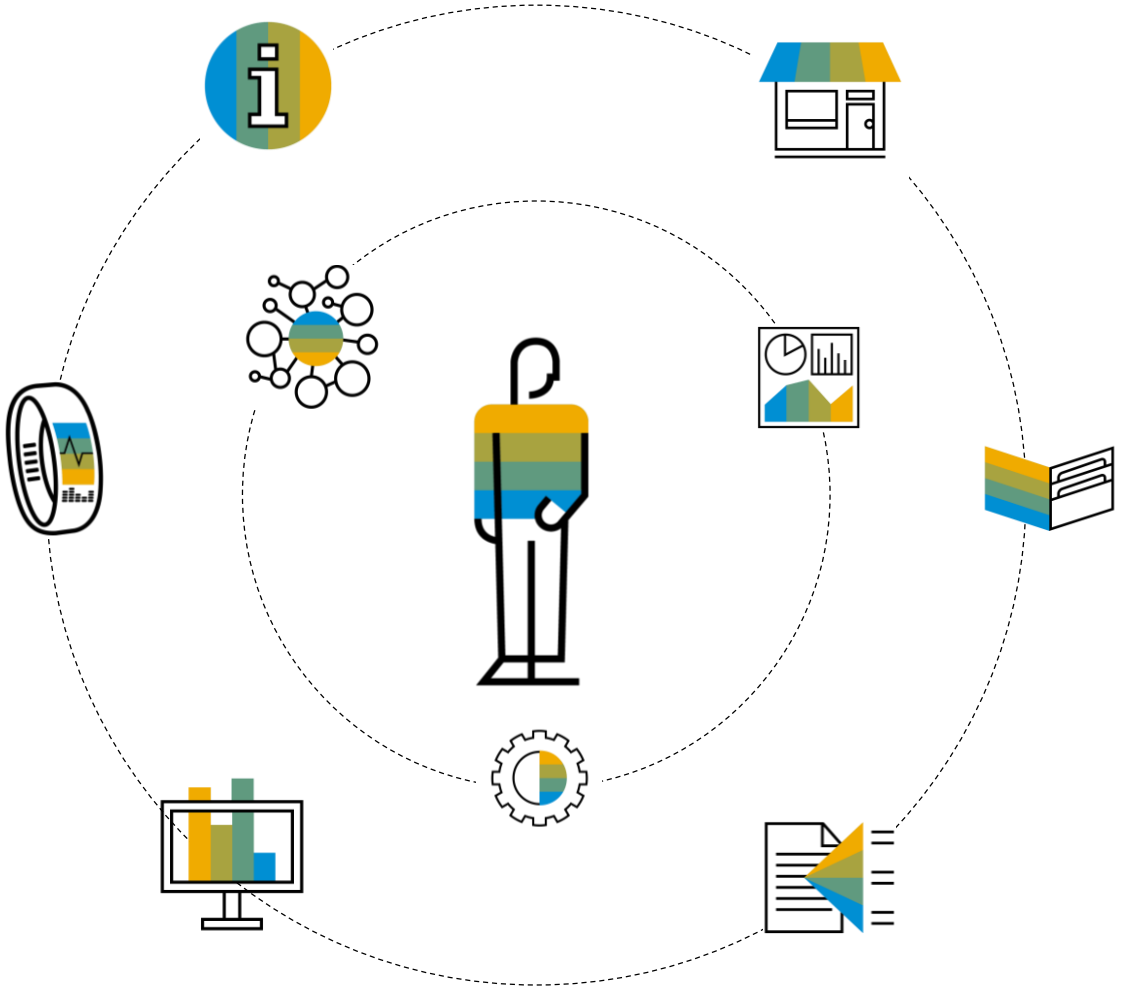
Record of achievement

- Collect at least 50% of the total points available in all online tests during the course



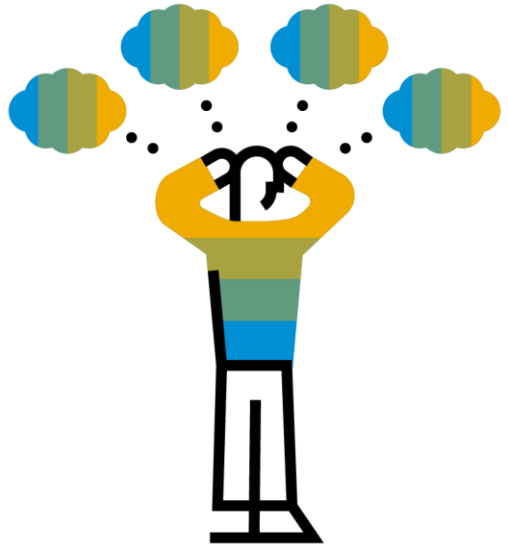
Numbers are everywhere!

- During every hour of every day, we make decisions and judgments based on data. For example, a house purchase …
  - Location
  - Size of town
  - Proximity to services, shops, and sea
  - Crime rates
  - Property prices
  - Size and number of rooms
  - Condition of house



How do I know?

- Crime has gone down/up
- By stopping eating meat you will improve your health and save the environment
- Climate change cannot be true because we had a cold winter!
- Our government has invested 14% more in social services



How do we know? How do we test the claims?!

What is a statistical problem?

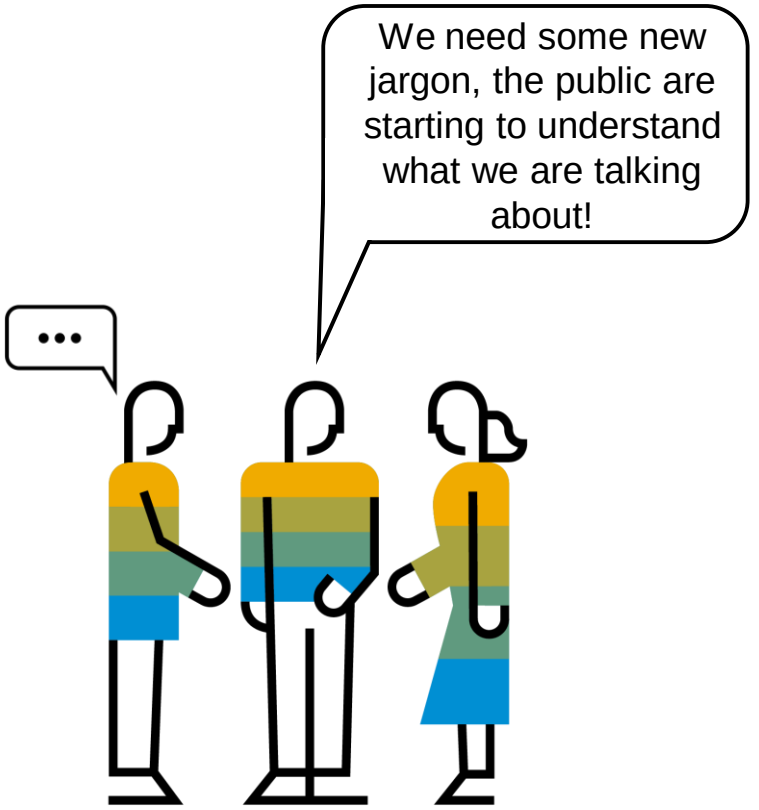
- How much sleep does the average person get?
- Is there a difference by age/gender/ethnicity etc?
- What lifestyle characteristics influence sleep quantity and quality?
- How do I test whether a sleep intervention is effective or not?

What is a statistical problem in business?

- How much profit are we making?
- How many product defects are we discovering? What's the trend?
- Has the process change led to a significant increase or decrease in employee satisfaction?
- What is our customer churn rate?

Some key statistical terms

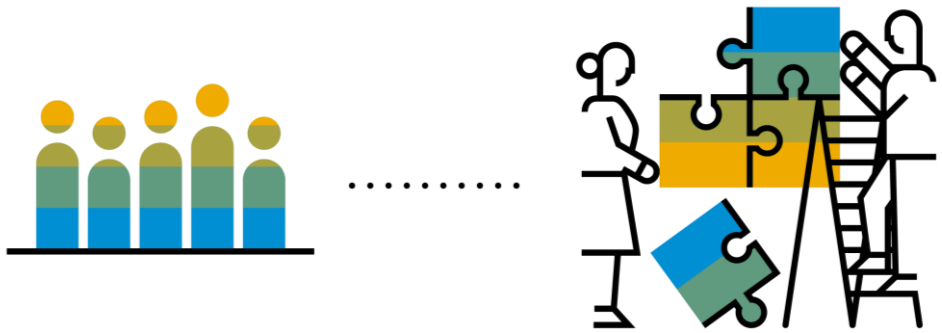
- Population versus sample
- Randomness
- Descriptive statistics
- Distributions
- Inference
- Probability
- Correlation



Population versus sample

“**Population**" data sets and "**sample**" data sets

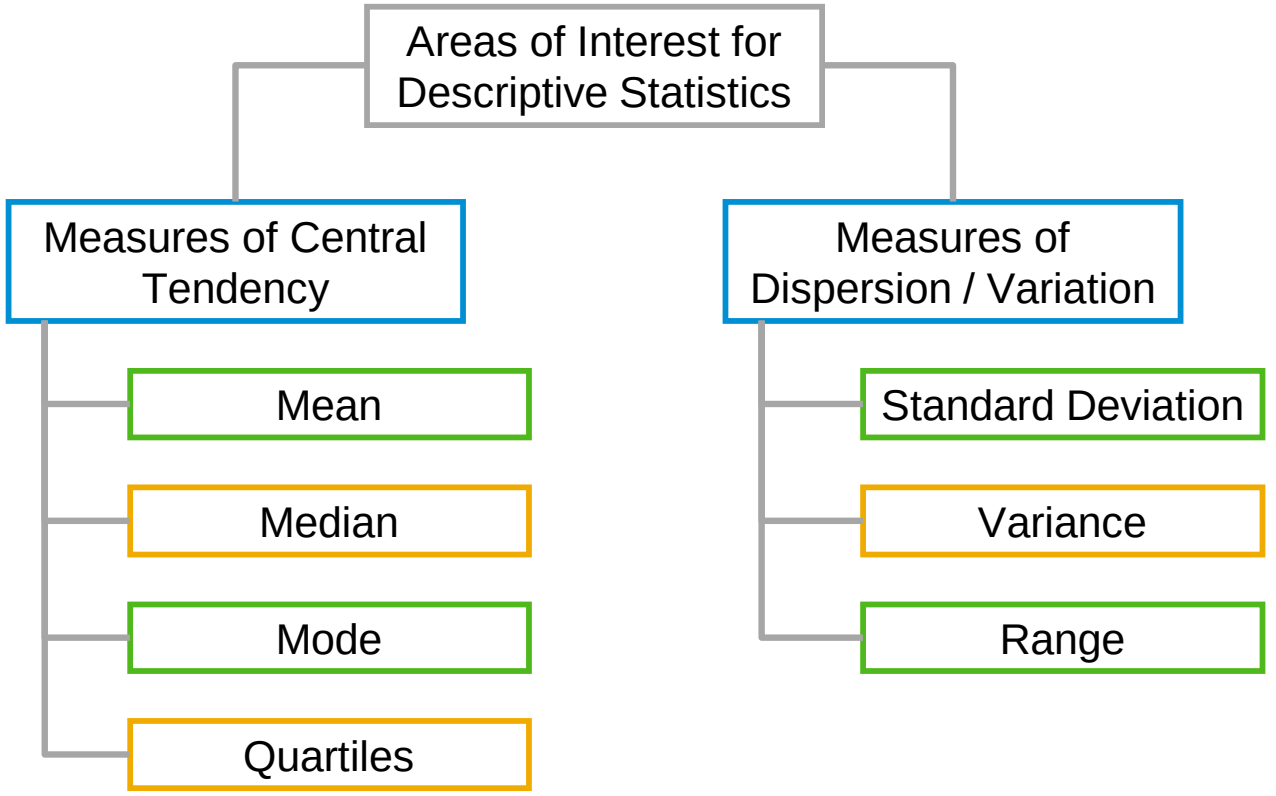
- A **population** data set contains all members of a specified group (the entire list of possible data values).
- A **sample** data set contains a part, or a subset, of a **population**. The size of a **sample** is always less than the size of the **population** from which it is taken.



https://mathbitsnotebook.com/Algebra1/StatisticsData/STPopSample.html

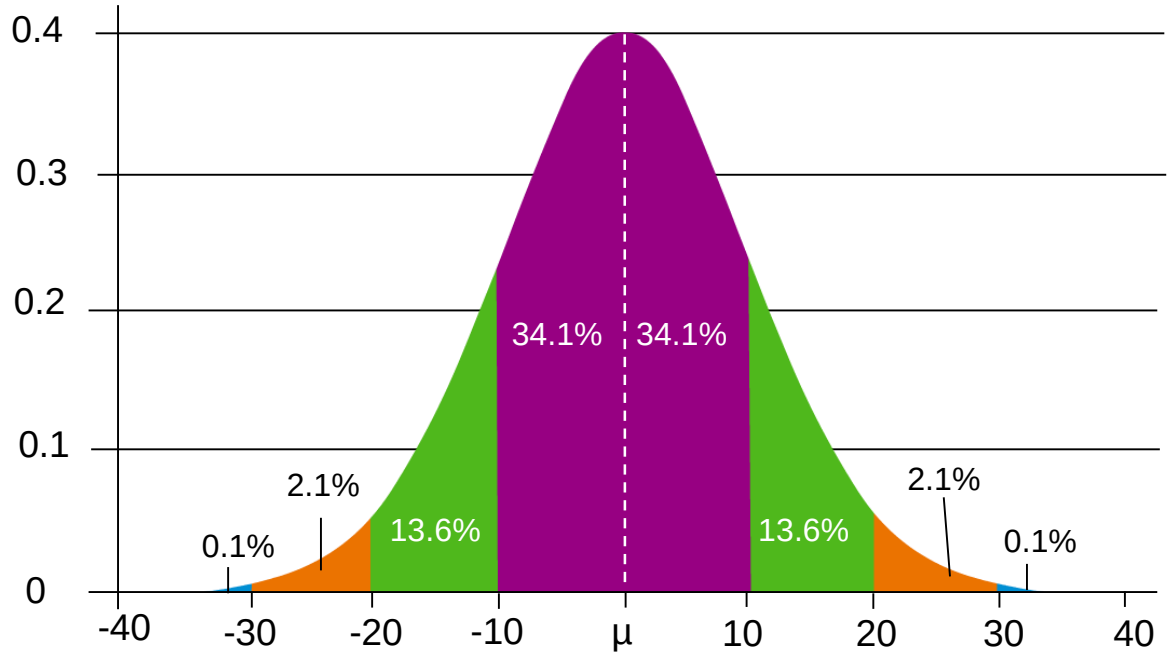
Descriptive statistics

- **Descriptive statistics** attempt to summarize a large body of data so that you can highlight key information.
- This is mainly through measures of central tendency and measures of dispersion.



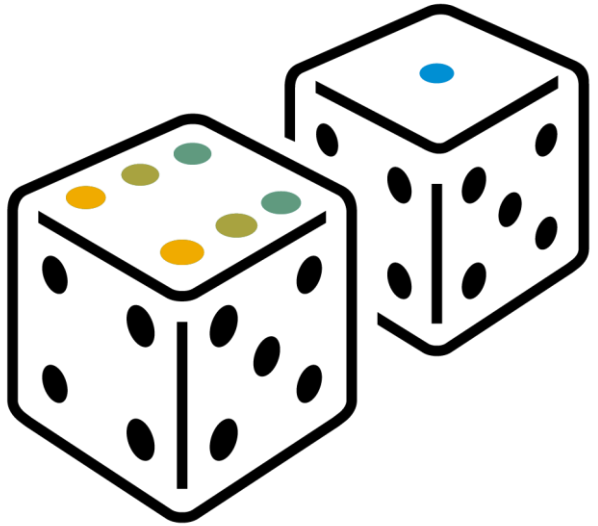
Distributions

- A *frequency* distribution provides a way of viewing all the values of a sample in a table view or a histogram.
- A *probability* distribution is a mathematical function that describes the probability of getting any particular result, such as the outcome when you roll 2 dice.



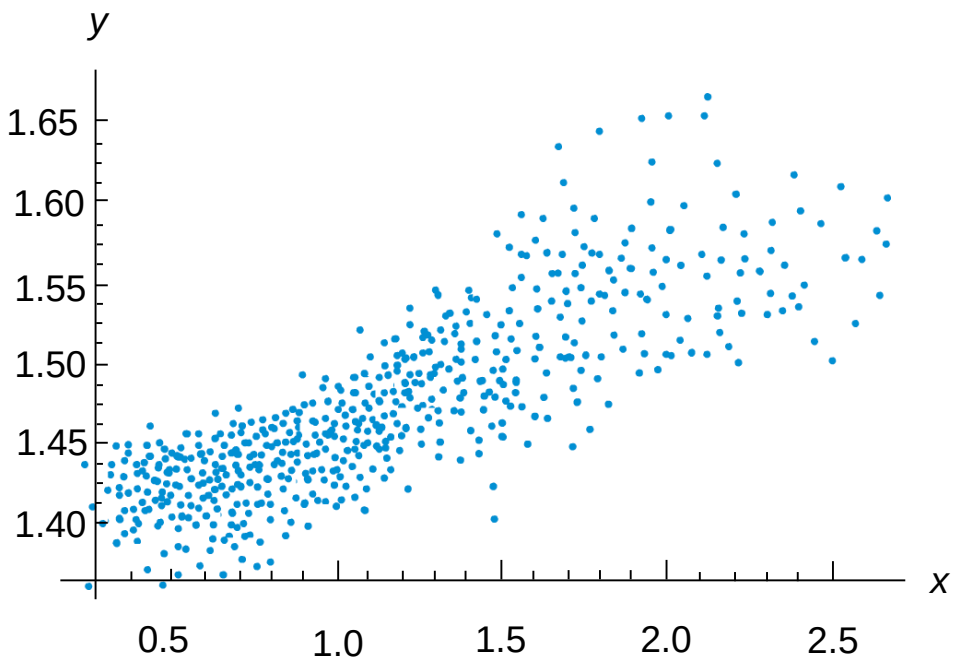
Probability

- Probability is a statistical measure of the likelihood of an event happening.
- Probability is measured between 0 and 1, so that at 0, an event definitely will not happen and at 1 it certainly will.
- Most probabilities fall between the two figures.



Correlation

- In statistics, dependence or association is any statistical relationship, whether causal or not, between two random variables.
- This statistical relationship is described as **correlation**.



Summary

- Statistics is everywhere, and everyday we are expected to make different kinds of statistical estimations.
- As part of the above, we are bombarded with different kinds of statistical claims from parties who want our vote, our time, or our money.
- We will learn how to evaluate those claims.
- There are some key statistical terms, which we need to understand to be able to develop our skills in understanding statistics.
- The first terms we need to understand are population vs. sample, randomness, descriptive statistics, distributions, probability, and correlation.

Thank you.

Contact information:

open@sap.com

Follow all of SAP

www.sap.com/contactsap

THE BEST RUN SAP®