



Week 1: Introduction to Statistics

## Unit 1: Introduction to Statistics

# 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

#### Week 1

#### Introduction to Statistics

Video 1

Self-test 1

Video 2

Self-test 2

Video n

Self-test n

Weekly assignment



#### Week 2

Descriptive Statistics

#### Week 3

Correlation and Linear Regression

#### Week 4

Introduction to Probability

#### Week 5

Probability Distributions

#### Week 6

Connecting to Your SAP Solutions

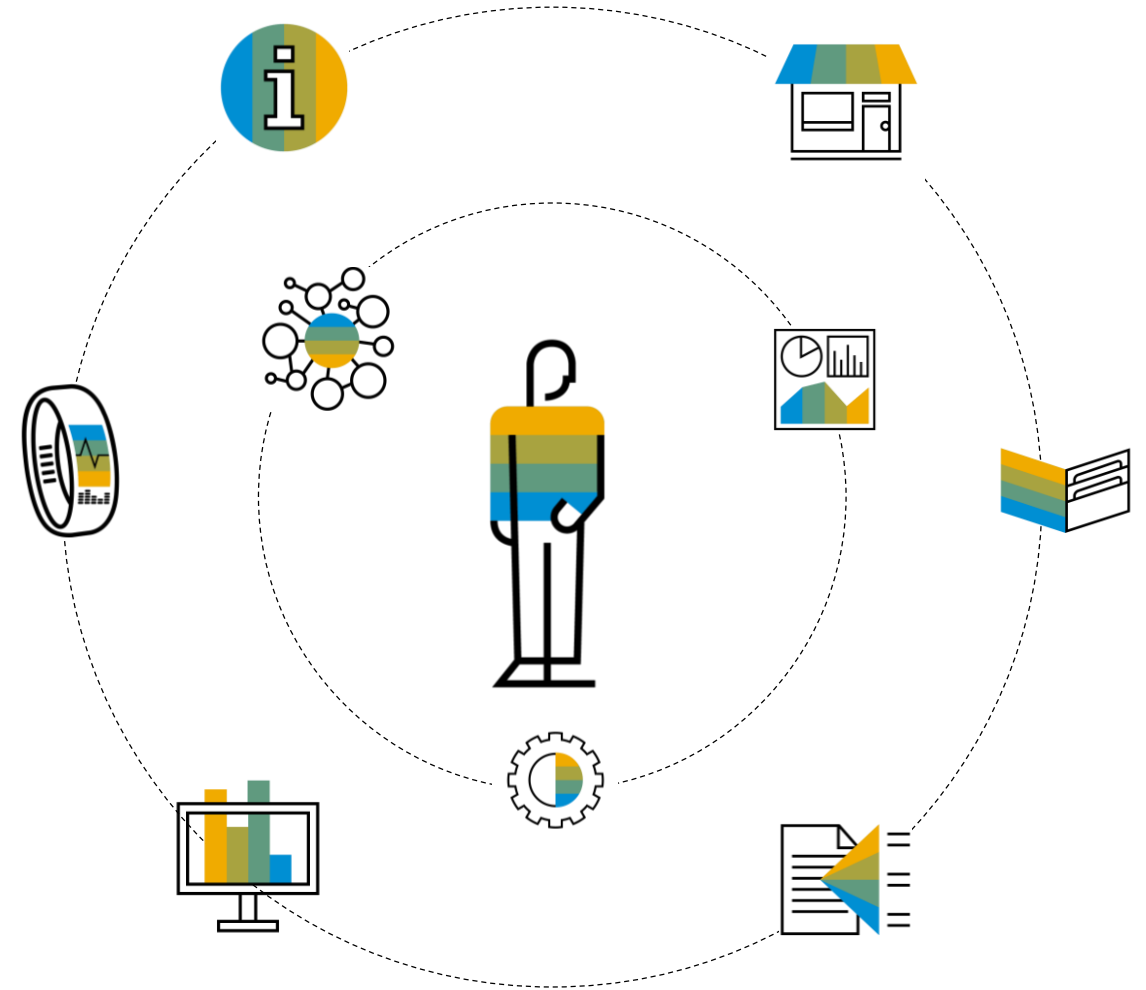
#### Week 7

Final exam



## Numbers are everywhere!

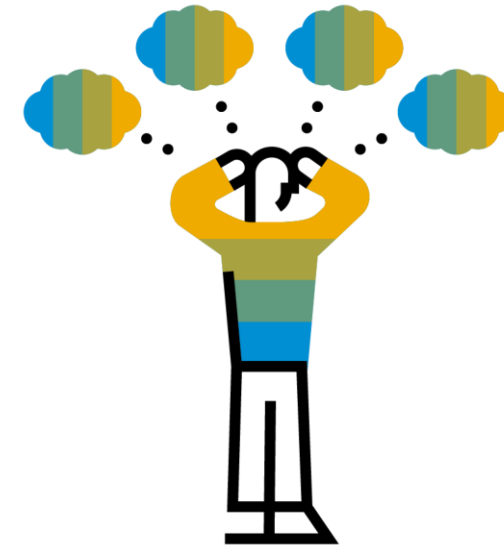
- During every hour of every day, we make decisions and judgments based on data. For example, a house purchase ...
  - Location
  - Size of town
  - Proximity to services, shops, and sea
  - Crime rates
  - Property prices
  - Size and number of rooms
  - Condition of house



### How do I know?

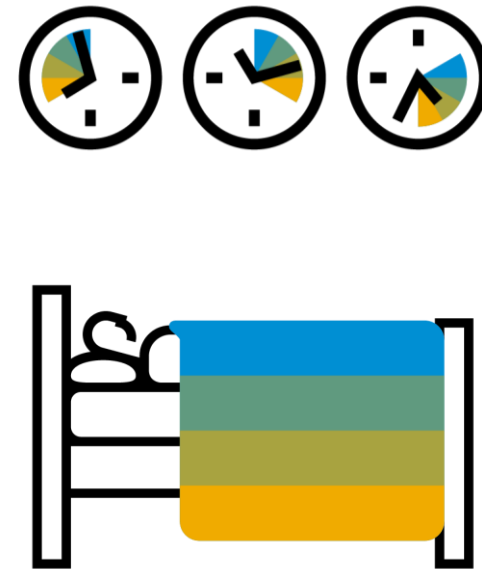
- Crime has gone down/up
- By stopping eating meat you will improve your health and save the environment
- Climate change cannot be true because we had a cold winter!
- Our government has invested 14% more in social services

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?



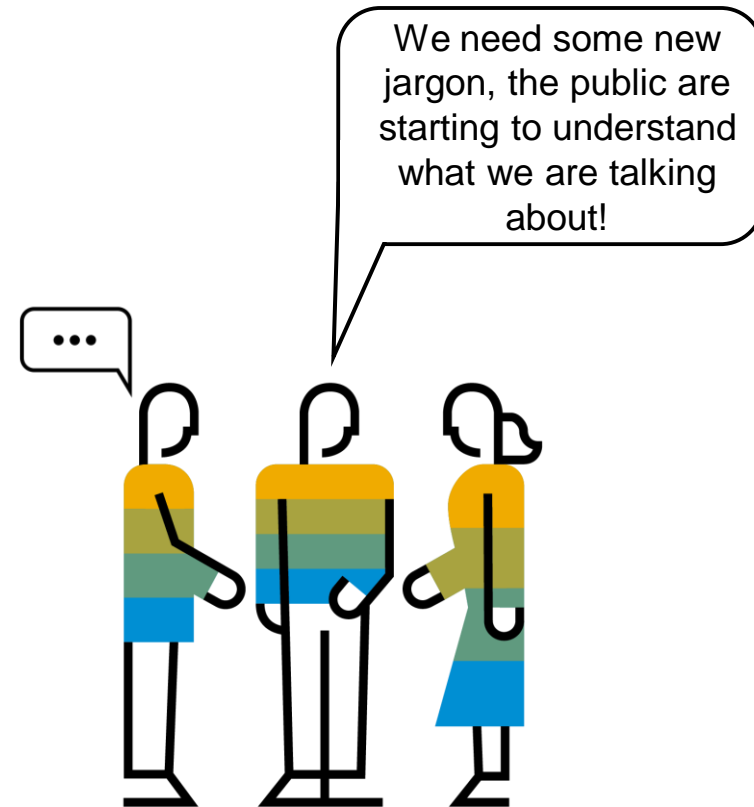
## What is a statistical problem in business?

- 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

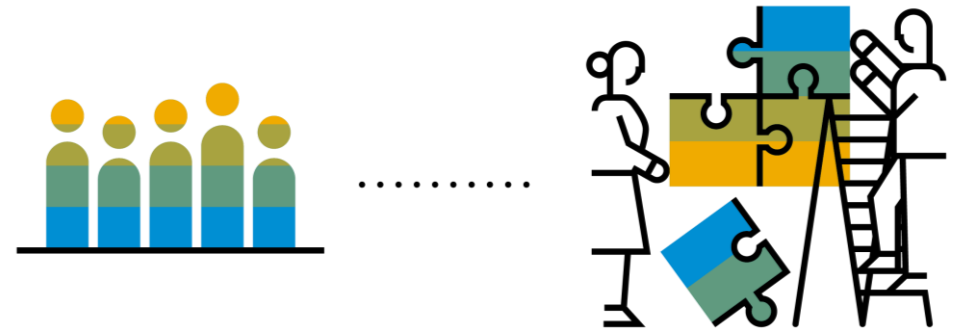
- Population versus sample
- Randomness
- Descriptive statistics
- Distributions
- Inference
- Probability
- Correlation



## 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.



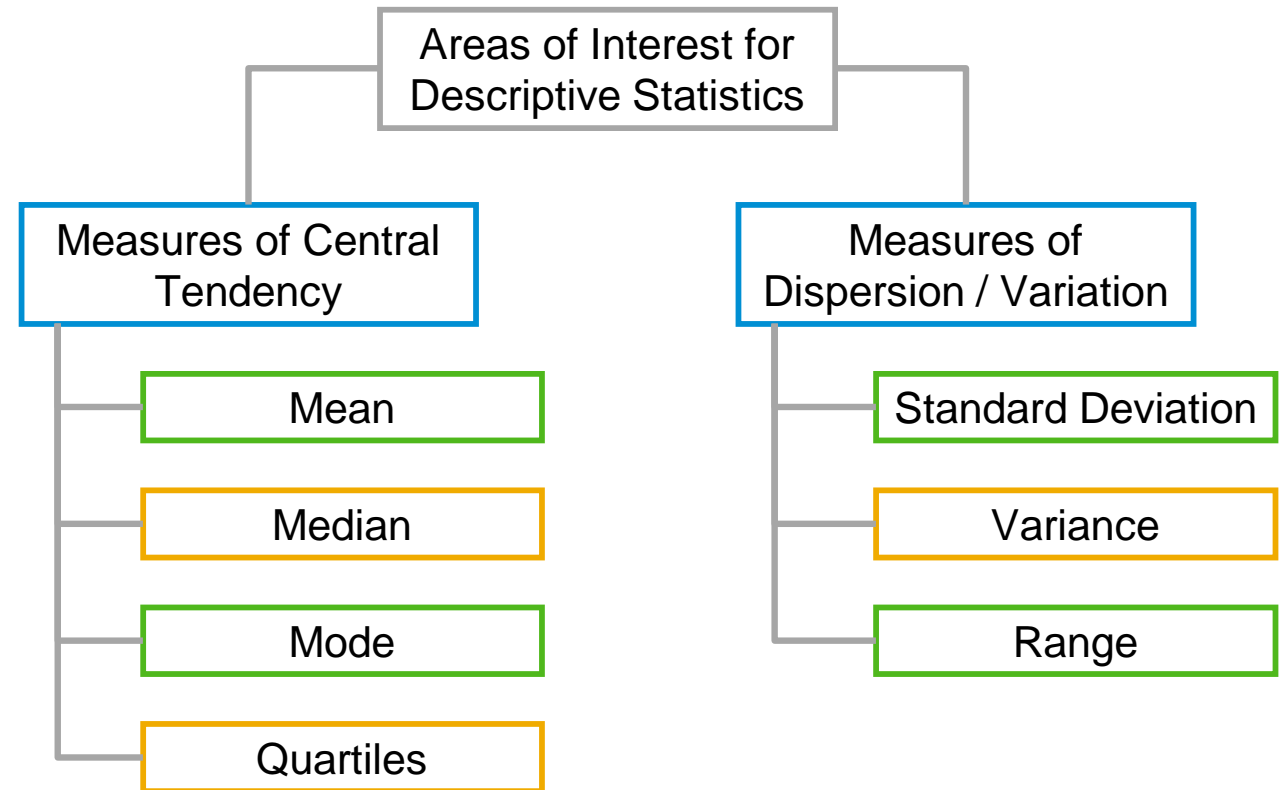
<https://mathbitsnotebook.com/Algebra1/StatisticsData/STPopSample.html>



# Introduction to Statistics

## Descriptive statistics

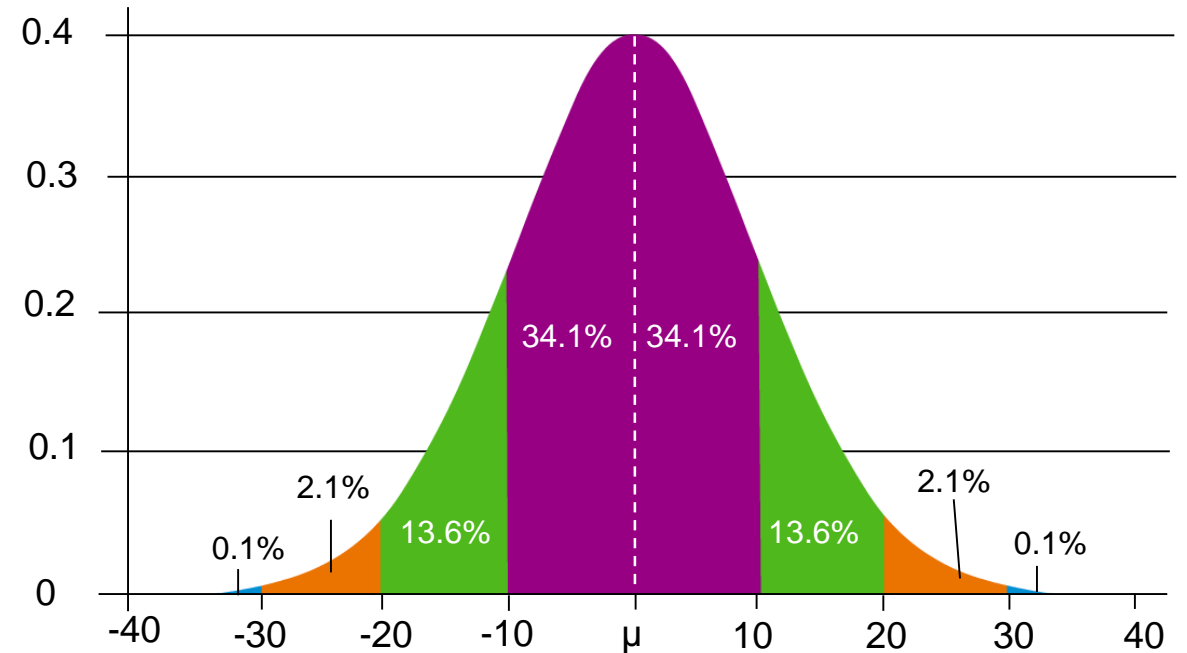
- **Descriptive statistics** attempt to summarize a large body of data so that you can highlight key information.
- This is mainly through measures of central tendency and measures of dispersion.



# Introduction to Statistics

## Distributions

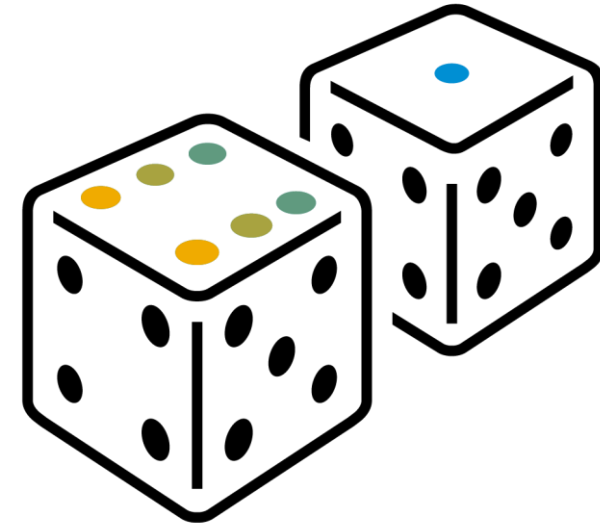
- A *frequency* distribution provides a way of viewing all the values of a sample in a table view or a histogram.
- A *probability* distribution is a mathematical function that describes the probability of getting any particular result, such as the outcome when you roll 2 dice.



# Introduction to Statistics

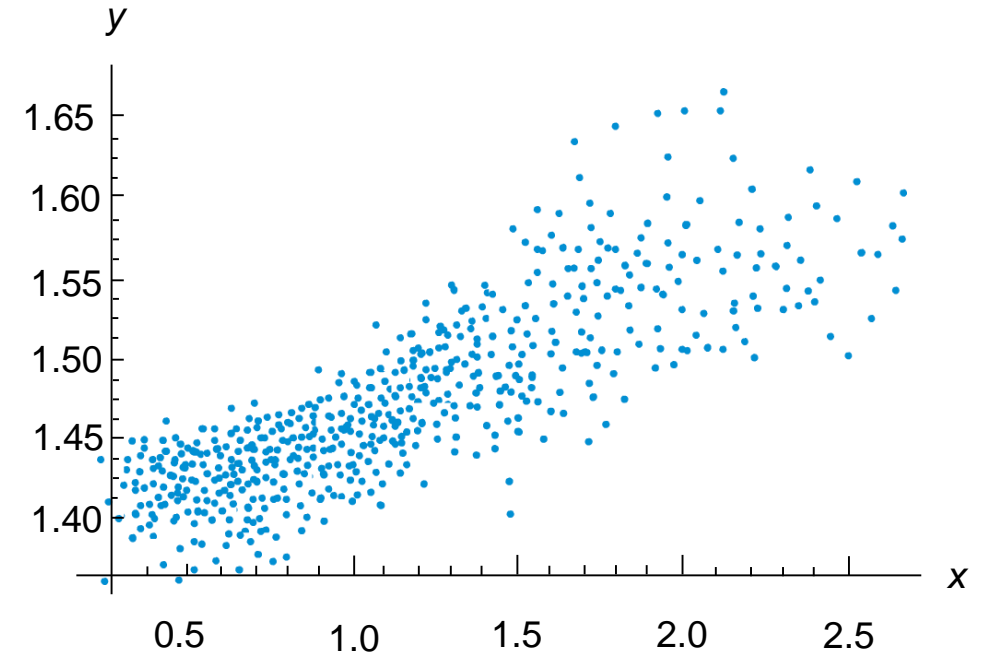
## Probability

- 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](http://www.sap.com/contactsap)

© 2019 SAP SE or an SAP affiliate company. All rights reserved.

No part of this publication may be reproduced or transmitted in any form or for any purpose without the express permission of SAP SE or an SAP affiliate company.

The information contained herein may be changed without prior notice. Some software products marketed by SAP SE and its distributors contain proprietary software components of other software vendors. National product specifications may vary.

These materials are provided by SAP SE or an SAP affiliate company for informational purposes only, without representation or warranty of any kind, and SAP or its affiliated companies shall not be liable for errors or omissions with respect to the materials. The only warranties for SAP or SAP affiliate company products and services are those that are set forth in the express warranty statements accompanying such products and services, if any. Nothing herein should be construed as constituting an additional warranty.

In particular, SAP SE or its affiliated companies have no obligation to pursue any course of business outlined in this document or any related presentation, or to develop or release any functionality mentioned therein. This document, or any related presentation, and SAP SE's or its affiliated companies' strategy and possible future developments, products, and/or platforms, directions, and functionality are all subject to change and may be changed by SAP SE or its affiliated companies at any time for any reason without notice. The information in this document is not a commitment, promise, or legal obligation to deliver any material, code, or functionality. All forward-looking statements are subject to various risks and uncertainties that could cause actual results to differ materially from expectations. Readers are cautioned not to place undue reliance on these forward-looking statements, and they should not be relied upon in making purchasing decisions.

SAP and other SAP products and services mentioned herein as well as their respective logos are trademarks or registered trademarks of SAP SE (or an SAP affiliate company) in Germany and other countries. All other product and service names mentioned are the trademarks of their respective companies.

See [www.sap.com/copyright](http://www.sap.com/copyright) for additional trademark information and notices.