Week 1: Introduction to Statistics

# Unit 5: Different Kinds of Analytic Approaches

# Qualitative vs. quantitative data

- The type of analytical approach you take depends on the type of data you have collected and the question you are answering.
- There are two types of data: **qualitative** and **quantitative**.

## Qualitative Data

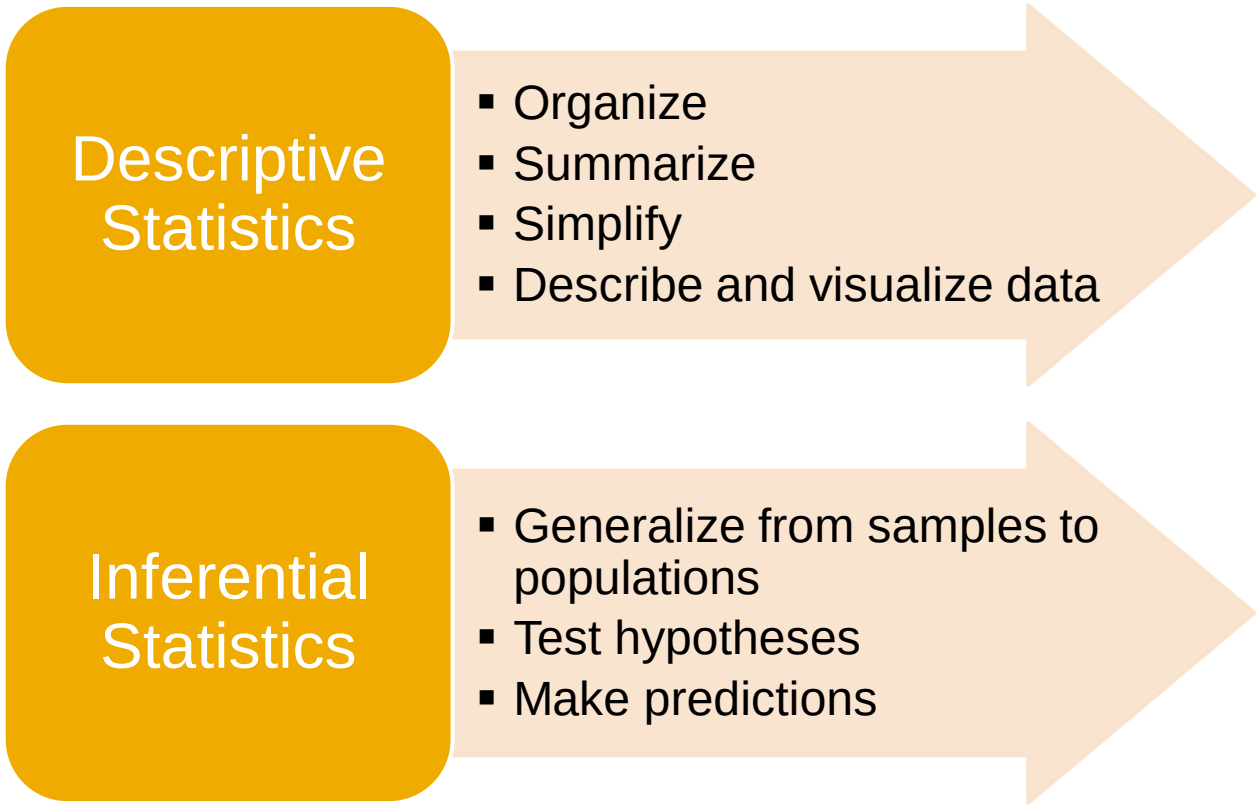
- Qualitative → Quality
- Deals with descriptions
- Data can be observed but not measured
- Colors, textures, smells, tastes, appearance, etc.

## Quantitative Data

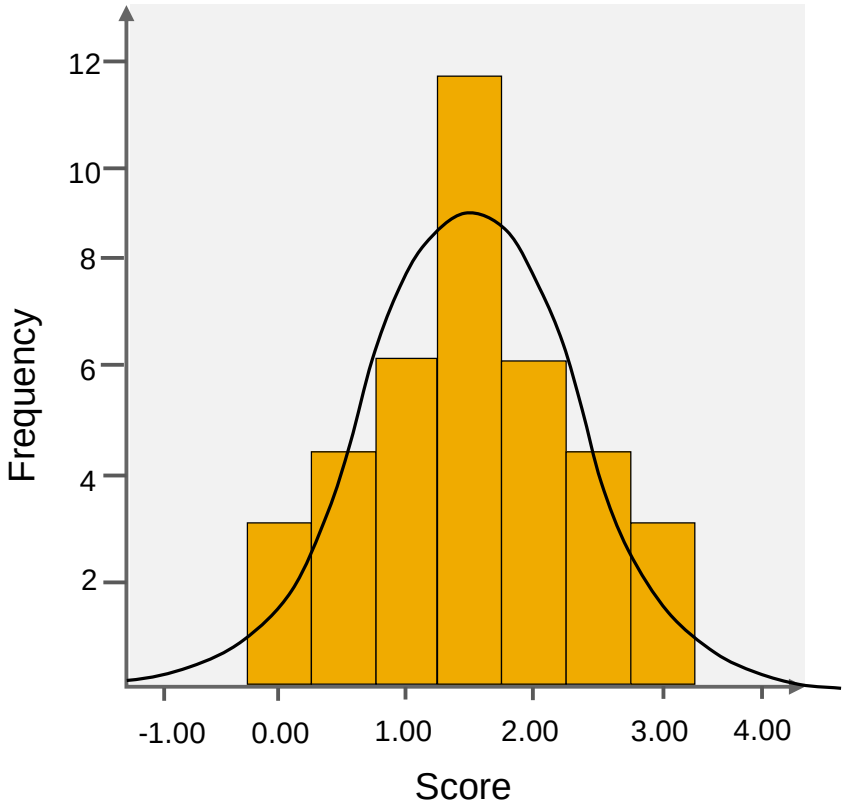
- Quantitative → Quantity
- Data that can be measured
- Length, height, area, volume, weight, speed, time, temperature, cost, etc.

# Descriptive vs. inferential analysis

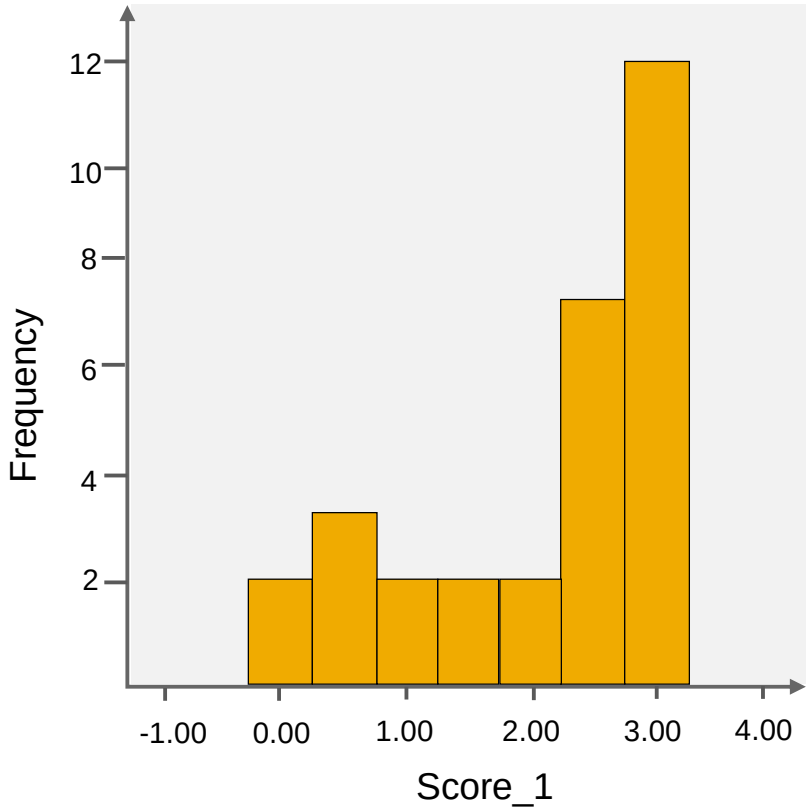
- There are two common types of analysis that are referred to as “descriptive” and “inferential”.



# Normal vs. non-normal




A **normal distribution** looks like the bell curve




A **non-normal distribution**

https://cyfar.org/inferential-analysis for more information

# Common parametric statistical tests

| Correlation: Analyze the association between variables | |
|---|---|
| Pearson correlation | Tests for the strength of the association between two continous variables |
| Spearman correlation | Tests for the strength of the association between two ordinal variables (does not assume that data is normally distributed) |
| Chi-square | Tests for the strength of the association between two categorical variables |

| Comparison of means: Analyze the difference between the means of variables | |
|---|---|
| One sample T-test | Compares the mean of a sample to a pre-specified value and tests for a deviation from that value |
| Paired T-test | Tests for the difference between two related variables |
| Independent T-test | Tests for the difference between two independent variables |
| Analysis of Variance (ANOVA) | Tests for the difference between group means in a sample after any other variance in the outcome variable is accounted for |

| Regression: Analyze how change in one variable predicts change in another variable | |
|---|---|
| Simple regression | Tests how change in the predictor variable predicts the level of change in the outcome variable |
| Multiple regression | Tests how change in the combination of two or more predictor variables predicts the level of change in the outcome variable |

Some of these tests will be described in more detail later in this course.

https://towardsdatascience.com/statistical-tests-when-to-use-which-704557554740

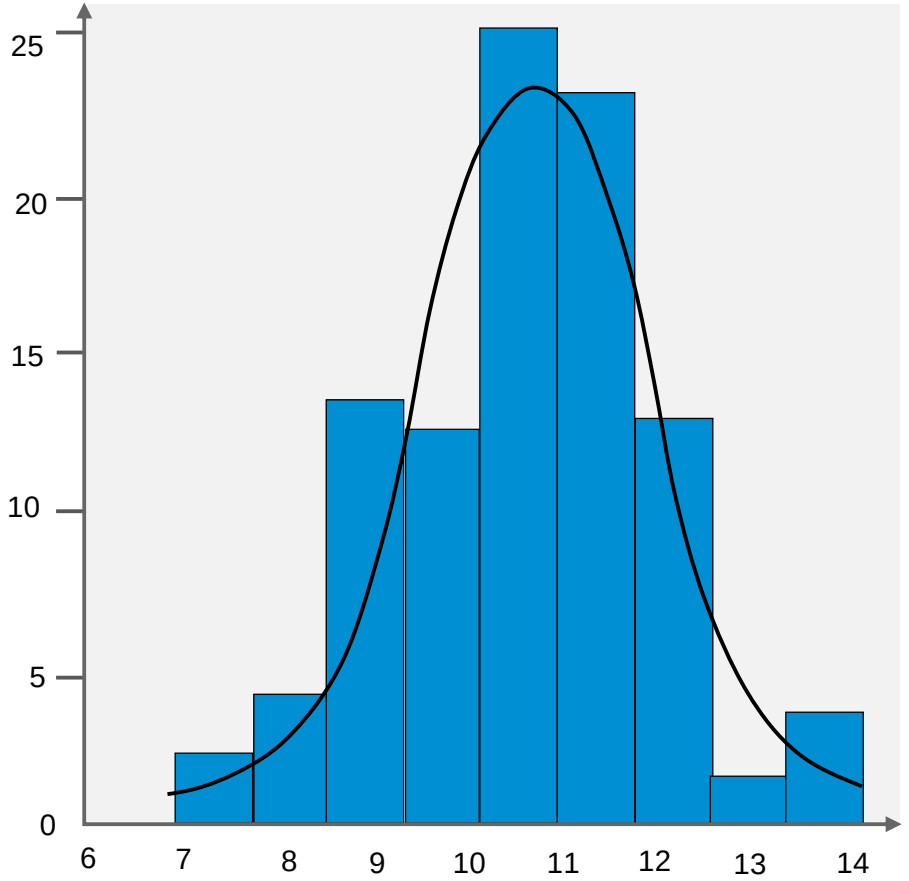
# Common nonparametric statistical tests

| Nonparametric: used when the data does not meet the assumptions required for parametric tests | |
|---|---|
| Sign test | Tests if two related variables are different – ignores the magnitude of change, only takes into account direction. The sign is an alternative to one sample T-test or a paired T-test. |
| Wilcoxon rank-sum test | Tests for the difference between two independent variables – takes into account magnitude and direction of difference |
| Wilcoxon sign-rank test | Tests for the difference between two related variables – takes into account the magnitude and direction of difference |

For more information, see
http://sphweb.bumc.bu.edu/otlt/mph-modules/bs/bs704_nonparametric/BS704_Nonparametric_print.html

# Choosing parametric or nonparametric statistical tests

- It can sometimes be difficult to assess whether a continuous outcome follows a normal distribution and whether a parametric or nonparametric test is appropriate.
- The most practical approach to assessing normality involves analyzing the distribution of the outcome in the sample using a histogram.




https://en.wikipedia.org/wiki/Goodness_of_fit for more information

# Summary

- **Descriptive analysis** informs you about the basic qualities of the data.
- **Inferential analysis** uses statistical tests to analyze whether a pattern in the data is due to chance or due to the intervention that is observed, and what the strength of that relationship is.
- In this course, you'll learn about some of these descriptive and inferential statistical techniques, and how these techniques can be misused.

# Thank you.

**Contact information:**

**open@sap.com**

Follow all of SAP

**www.sap.com/contactsap**



THE BEST RUN SAP®