Univariate and Bivariate Tests

BUS 230: Business and Economics Research and Communication
Specific goals:
- Be able to distinguish different types of data and prescribe appropriate statistical methods.
- Conduct a number of hypothesis tests using methods appropriate for questions involving only one or two variables.

Learning objectives:
- LO2: Interpret data using statistical analysis.
- LO2.3: Formulate conclusions and recommendations based upon statistical results.
Specific goals:

- Be able to distinguish different types of data and prescribe appropriate statistical methods.
- Conduct a number of hypothesis tests using methods appropriate for questions involving only one or two variables.

Learning objectives:

- LO2: Interpret data using statistical analysis.
- LO2.3: Formulate conclusions and recommendations based upon statistical results.
A **hypothesis** is a claim or statement about a property of a population.

- Example: The population mean for systolic blood pressure is 120.

A **hypothesis test** (or **test of significance**) is a standard procedure for testing a claim about a property of a population.

Recall the example about birth weights with mothers who use drugs.

- Hypothesis: Using drugs during pregnancy leads to an average birth weight of 7 pounds (the same as with mothers who do not use drugs).
A hypothesis is a claim or statement about a property of a population.

Example: The population mean for systolic blood pressure is 120.

A hypothesis test (or test of significance) is a standard procedure for testing a claim about a property of a population.

Recall the example about birth weights with mothers who use drugs.

Hypothesis: Using drugs during pregnancy leads to an average birth weight of 7 pounds (the same as with mothers who do not use drugs).
A hypothesis is a claim or statement about a property of a population.

- Example: The population mean for systolic blood pressure is 120.

A hypothesis test (or test of significance) is a standard procedure for testing a claim about a property of a population.

Recall the example about birth weights with mothers who use drugs.

- Hypothesis: Using drugs during pregnancy leads to an average birth weight of 7 pounds (the same as with mothers who do not use drugs).
A hypothesis is a claim or statement about a property of a population.

- Example: The population mean for systolic blood pressure is 120.

A hypothesis test (or test of significance) is a standard procedure for testing a claim about a property of a population.

Recall the example about birth weights with mothers who use drugs.

- Hypothesis: Using drugs during pregnancy leads to an average birth weight of 7 pounds (the same as with mothers who do not use drugs).
A hypothesis is a claim or statement about a property of a population.

Example: The population mean for systolic blood pressure is 120.

A hypothesis test (or test of significance) is a standard procedure for testing a claim about a property of a population.

Recall the example about birth weights with mothers who use drugs.

Hypothesis: Using drugs during pregnancy leads to an average birth weight of 7 pounds (the same as with mothers who do not use drugs).
The null hypothesis is a statement that the value of a population parameter (such as the population mean) is equal to some claimed value.

- $H_0: \mu = 7$.

The alternative hypothesis is an alternative to the null hypothesis; a statement that says a parameter differs from the value given in the null hypothesis.

Pick only one of the following for your alternative hypothesis. Which one depends on your research question.

- $H_a: \mu < 7$.
- $H_a: \mu > 7$.
- $H_a: \mu \neq 7$.

In hypothesis testing, assume the null hypothesis is true until there is strong statistical evidence to suggest the alternative hypothesis.

Similar to an “innocent until proven guilty” policy.
The **null hypothesis** is a statement that the value of a population parameter (such as the population mean) *is equal to* some claimed value.

- $H_0: \mu = 7$.

The **alternative hypothesis** is an alternative to the null hypothesis; a statement that says a parameter differs from the value given in the null hypothesis.

Pick *only one* of the following for your alternative hypothesis. Which one depends on your research question.

- $H_a: \mu < 7$.
- $H_a: \mu > 7$.
- $H_a: \mu \neq 7$.

In hypothesis testing, assume the null hypothesis is true until there is strong statistical evidence to suggest the alternative hypothesis.

Similar to an “innocent until proven guilty” policy.
The **null hypothesis** is a statement that the value of a population parameter (such as the population mean) is *equal to* some claimed value.

- $H_0: \mu = 7$.

The **alternative hypothesis** is an alternative to the null hypothesis; a statement that says a parameter differs from the value given in the null hypothesis.

Pick *only one* of the following for your alternative hypothesis. Which one depends on your research question.

- $H_a: \mu < 7$.
- $H_a: \mu > 7$.
- $H_a: \mu \neq 7$.

In hypothesis testing, assume the null hypothesis is true until there is strong statistical evidence to suggest the alternative hypothesis.

Similar to an “innocent until proven guilty” policy.
The null hypothesis is a statement that the value of a population parameter (such as the population mean) is equal to some claimed value.

- $H_0: \mu = 7$.

The alternative hypothesis is an alternative to the null hypothesis; a statement that says a parameter differs from the value given in the null hypothesis.

Pick only one of the following for your alternative hypothesis. Which one depends on your research question.

- $H_a: \mu < 7$.
- $H_a: \mu > 7$.
- $H_a: \mu \neq 7$.

In hypothesis testing, assume the null hypothesis is true until there is strong statistical evidence to suggest the alternative hypothesis.

Similar to an “innocent until proven guilty” policy.
Null and Alternative Hypotheses

- The **null hypothesis** is a statement that the value of a population parameter (such as the population mean) is equal to some claimed value.
  - $H_0: \mu = 7$.

- The **alternative hypothesis** is an alternative to the null hypothesis; a statement that says a parameter differs from the value given in the null hypothesis.

- Pick *only one* of the following for your alternative hypothesis. Which one depends on your research question.
  - $H_a: \mu < 7$.
  - $H_a: \mu > 7$.
  - $H_a: \mu \neq 7$.

- In hypothesis testing, assume the null hypothesis is true until there is strong statistical evidence to suggest the alternative hypothesis.

- Similar to an “innocent until proven guilty” policy.
Null and Alternative Hypotheses

- The **null hypothesis** is a statement that the value of a population parameter (such as the population mean) *is equal to* some claimed value.
  - $H_0: \mu = 7$.

- The **alternative hypothesis** is an alternative to the null hypothesis; a statement that says a parameter differs from the value given in the null hypothesis.

- Pick *only one* of the following for your alternative hypothesis. Which one depends on your research question.
  - $H_a: \mu < 7$.
  - $H_a: \mu > 7$.
  - $H_a: \mu \neq 7$.

- In hypothesis testing, assume the null hypothesis is true until there is strong statistical evidence to suggest the alternative hypothesis.

- Similar to an “innocent until proven guilty” policy.
The **null hypothesis** is a statement that the value of a population parameter (such as the population mean) is *equal to* some claimed value.

- $H_0$: $\mu = 7$.

The **alternative hypothesis** is an alternative to the null hypothesis; a statement that says a parameter differs from the value given in the null hypothesis.

Pick *only one* of the following for your alternative hypothesis. Which one depends on your research question.

- $H_a$: $\mu < 7$.
- $H_a$: $\mu > 7$.
- $H_a$: $\mu \neq 7$.

In hypothesis testing, assume the null hypothesis is true until there is strong statistical evidence to suggest the alternative hypothesis.

Similar to an “innocent until proven guilty” policy.
The **null hypothesis** is a statement that the value of a population parameter (such as the population mean) *is equal to* some claimed value.

- \( H_0: \mu = 7. \)

The **alternative hypothesis** is an alternative to the null hypothesis; a statement that says a parameter differs from the value given in the null hypothesis.

Pick *only one* of the following for your alternative hypothesis. Which one depends on your research question.

- \( H_a: \mu < 7. \)
- \( H_a: \mu > 7. \)
- \( H_a: \mu \neq 7. \)

In hypothesis testing, assume the null hypothesis is true until there is strong statistical evidence to suggest the alternative hypothesis.

Similar to an “innocent until proven guilty” policy.
The null hypothesis is a statement that the value of a population parameter (such as the population mean) is equal to some claimed value.

- $H_0: \mu = 7$.

The alternative hypothesis is an alternative to the null hypothesis; a statement that says a parameter differs from the value given in the null hypothesis.

Pick only one of the following for your alternative hypothesis. Which one depends on your research question.

- $H_a: \mu < 7$.
- $H_a: \mu > 7$.
- $H_a: \mu \neq 7$.

In hypothesis testing, assume the null hypothesis is true until there is strong statistical evidence to suggest the alternative hypothesis.

Similar to an “innocent until proven guilty” policy.
(Many) hypothesis tests are all the same:

$$z \text{ or } t = \frac{\text{sample statistic} - \text{null hypothesis value}}{\text{standard deviation of the sampling distribution}}$$

Example: hypothesis testing about $$\mu$$:
- Sample statistic = $$\bar{x}$$.
- Standard deviation of the sampling distribution of $$\bar{x}$$:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$
(Many) hypothesis tests are all the same:

\[ z \text{ or } t = \frac{\text{sample statistic} - \text{null hypothesis value}}{\text{standard deviation of the sampling distribution}} \]

Example: hypothesis testing about \( \mu \):
- Sample statistic = \( \bar{x} \).
- Standard deviation of the sampling distribution of \( \bar{x} \):

\[ \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \]
(Many) hypothesis tests are all the same:

$$z \text{ or } t = \frac{\text{sample statistic} - \text{null hypothesis value}}{\text{standard deviation of the sampling distribution}}$$

Example: hypothesis testing about $$\mu$$:

- Sample statistic = $$\bar{x}$$.
- Standard deviation of the sampling distribution of $$\bar{x}$$:

$$\sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}}$$
(Many) hypothesis tests are all the same:

\[ z \text{ or } t = \frac{\text{sample statistic} - \text{null hypothesis value}}{\text{standard deviation of the sampling distribution}} \]

Example: hypothesis testing about \( \mu \):

- Sample statistic = \( \bar{x} \).
- Standard deviation of the sampling distribution of \( \bar{x} \):

\[ \sigma_{\bar{x}} = \frac{\sigma}{\sqrt{n}} \]
The P-value of the test statistic, is the area of the sampling distribution from the sample result in the direction of the alternative hypothesis.

Interpretation: If the null hypothesis is correct, then the p-value is the probability of obtaining a sample that yielded your statistic, or a statistic that provides even stronger evidence of the null hypothesis.

The p-value is therefore a measure of statistical significance.

- If p-values are very small, there is strong statistical evidence in favor of the alternative hypothesis.
- If p-values are large, there is insignificant statistical evidence. When large, you fail to reject the null hypothesis.

Best practice is writing research: report the p-value. Different readers may have different opinions about how small a p-value should be before saying your results are statistically significant.
P-values

- The P-value of the test statistic, is the area of the sampling distribution from the sample result in the direction of the alternative hypothesis.
- Interpretation: If the null hypothesis is correct, then the p-value is the probability of obtaining a sample that yielded your statistic, or a statistic that provides even stronger evidence of the null hypothesis.
- The p-value is therefore a measure of statistical significance.
  - If p-values are very small, there is strong statistical evidence in favor of the alternative hypothesis.
  - If p-values are large, there is insignificant statistical evidence. When large, you fail to reject the null hypothesis.
- Best practice is writing research: report the p-value. Different readers may have different opinions about how small a p-value should be before saying your results are statistically significant.
The P-value of the test statistic, is the area of the sampling distribution from the sample result in the direction of the alternative hypothesis.

Interpretation: If the null hypothesis is correct, than the p-value is the probability of obtaining a sample that yielded your statistic, or a statistic that provides even stronger evidence of the null hypothesis.

The p-value is therefore a measure of statistical significance.

- If p-values are very small, there is strong statistical evidence in favor of the alternative hypothesis.
- If p-values are large, there is insignificant statistical evidence. When large, you fail to reject the null hypothesis.

Best practice is writing research: report the p-value. Different readers may have different opinions about how small a p-value should be before saying your results are statistically significant.
The P-value of the test statistic, is the area of the sampling distribution from the sample result in the direction of the alternative hypothesis.

Interpretation: If the null hypothesis is correct, than the p-value is the probability of obtaining a sample that yielded your statistic, or a statistic that provides even stronger evidence of the null hypothesis.

The p-value is therefore a measure of statistical significance.

- If p-values are very small, there is strong statistical evidence in favor of the alternative hypothesis.
- If p-values are large, there is insignificant statistical evidence. When large, you fail to reject the null hypothesis.

Best practice is writing research: report the p-value. Different readers may have different opinions about how small a p-value should be before saying your results are statistically significant.
The P-value of the test statistic, is the area of the sampling distribution from the sample result in the direction of the alternative hypothesis.

Interpretation: If the null hypothesis is correct, then the p-value is the probability of obtaining a sample that yielded your statistic, or a statistic that provides even stronger evidence of the null hypothesis.

The p-value is therefore a measure of statistical significance.

- If p-values are very small, there is strong statistical evidence in favor of the alternative hypothesis.
- If p-values are large, there is insignificant statistical evidence. When large, you fail to reject the null hypothesis.

Best practice is writing research: report the p-value. Different readers may have different opinions about how small a p-value should be before saying your results are statistically significant.
The P-value of the test statistic, is the area of the sampling distribution from the sample result in the direction of the alternative hypothesis.

Interpretation: If the null hypothesis is correct, than the p-value is the probability of obtaining a sample that yielded your statistic, or a statistic that provides even stronger evidence of the null hypothesis.

The p-value is therefore a measure of statistical significance.

If p-values are very small, there is strong statistical evidence in favor of the alternative hypothesis.

If p-values are large, there is insignificant statistical evidence. When large, you fail to reject the null hypothesis.

Best practice is writing research: report the p-value. Different readers may have different opinions about how small a p-value should be before saying your results are statistically significant.
Types of Data

- **Nominal data**: consists of categories that cannot be ordered in a meaningful way.
- **Ordinal data**: order is meaningful, but not the distances between data values.
  - Excellent, Very good, Good, Poor, Very poor.
- **Interval data**: order is meaningful, and distances are meaningful. However, there is no natural zero.
  - Examples: temperature, time.
- **Ratio data**: order, differences, and zero are all meaningful.
  - Examples: weight, prices, speed.
Types of Data

- Nominal data: consists of categories that cannot be ordered in a meaningful way.
- Ordinal data: order is meaningful, but not the distances between data values.
  - Excellent, Very good, Good, Poor, Very poor.
- Interval data: order is meaningful, and distances are meaningful. However, there is no natural zero.
  - Examples: temperature, time.
- Ratio data: order, differences, and zero are all meaningful.
  - Examples: weight, prices, speed.
Types of Data

- **Nominal data**: consists of categories that cannot be ordered in a meaningful way.
- **Ordinal data**: order is meaningful, but not the distances between data values.
  - Excellent, Very good, Good, Poor, Very poor.
- **Interval data**: order is meaningful, and distances are meaningful. However, there is no natural zero.
  - Examples: temperature, time.
- **Ratio data**: order, differences, and zero are all meaningful.
  - Examples: weight, prices, speed.
Types of Data

- Nominal data: consists of categories that cannot be ordered in a meaningful way.
- Ordinal data: order is meaningful, but not the distances between data values.
  - Excellent, Very good, Good, Poor, Very poor.
- Interval data: order is meaningful, and distances are meaningful. However, there is no natural zero.
  - Examples: temperature, time.
- Ratio data: order, differences, and zero are all meaningful.
  - Examples: weight, prices, speed.
Nominal data: consists of categories that cannot be ordered in a meaningful way.

Ordinal data: order is meaningful, but not the distances between data values.
  - Excellent, Very good, Good, Poor, Very poor.

Interval data: order is meaningful, and distances are meaningful. However, there is no natural zero.
  - Examples: temperature, time.

Ratio data: order, differences, and zero are all meaningful.
  - Examples: weight, prices, speed.
Types of Data

- **Nominal data**: consists of categories that cannot be ordered in a meaningful way.
- **Ordinal data**: order is meaningful, but not the distances between data values.
  - Excellent, Very good, Good, Poor, Very poor.
- **Interval data**: order is meaningful, and distances are meaningful. However, there is no natural zero.
  - Examples: temperature, time.
- **Ratio data**: order, differences, and zero are all meaningful.
  - Examples: weight, prices, speed.
Types of Data

- **Nominal data**: consists of categories that cannot be ordered in a meaningful way.
- **Ordinal data**: order is meaningful, but not the distances between data values.
  - Excellent, Very good, Good, Poor, Very poor.
- **Interval data**: order is meaningful, and distances are meaningful. However, there is no natural zero.
  - Examples: temperature, time.
- **Ratio data**: order, differences, and zero are all meaningful.
  - Examples: weight, prices, speed.
Different types of data require different statistical methods.

Why? With interval data and below, operations like addition, subtraction, multiplication, and division are meaningless!

Parametric statistics:
- Typically take advantage of central limit theorem (imposes requirements on probability distributions)
- Appropriate only for interval and ratio data.
- More powerful than nonparametric methods.

Nonparametric statistics:
- Do not require assumptions concerning the probability distribution for the population.
- There are many methods appropriate for ordinal data, some methods appropriate for nominal data.
- Computations typically make use of data's ranks instead of actual data.
Different types of data require different statistical methods.

Why? With interval data and below, operations like addition, subtraction, multiplication, and division are *meaningless!*

- **Parametric statistics:**
  - Typically take advantage of central limit theorem (imposes requirements on probability distributions)
  - Appropriate only for interval and ratio data.
  - More powerful than nonparametric methods.

- **Nonparametric statistics:**
  - Do not require assumptions concerning the probability distribution for the population.
  - There are many methods appropriate for ordinal data, some methods appropriate for nominal data.
  - Computations typically make use of data’s *ranks* instead of actual data.
Different types of data require different statistical methods.

Why? With interval data and below, operations like addition, subtraction, multiplication, and division are meaningless!

Parametric statistics:
- Typically take advantage of central limit theorem (imposes requirements on probability distributions)
- Appropriate only for interval and ratio data.
- More powerful than nonparametric methods.

Nonparametric statistics:
- Do not require assumptions concerning the probability distribution for the population.
- There are many methods appropriate for ordinal data, some methods appropriate for nominal data.
- Computations typically make use of data’s ranks instead of actual data.
Different types of data require different statistical methods.

Why? With interval data and below, operations like addition, subtraction, multiplication, and division are *meaningless!*

Parametric statistics:
- Typically take advantage of central limit theorem (imposes requirements on probability distributions)
- Appropriate only for interval and ratio data.
- More *powerful* than nonparametric methods.

Nonparametric statistics:
- Do not require assumptions concerning the probability distribution for the population.
- There are many methods appropriate for ordinal data, some methods appropriate for nominal data.
- Computations typically make use of data’s *ranks* instead of actual data.
Types of Tests

- Different types of data require different statistical methods.
- Why? With interval data and below, operations like addition, subtraction, multiplication, and division are meaningless!
- Parametric statistics:
  - Typically take advantage of central limit theorem (imposes requirements on probability distributions)
  - Appropriate only for interval and ratio data.
  - More powerful than nonparametric methods.
- Nonparametric statistics:
  - Do not require assumptions concerning the probability distribution for the population.
  - There are many methods appropriate for ordinal data, some methods appropriate for nominal data.
  - Computations typically make use of data’s ranks instead of actual data.
Different types of data require different statistical methods.

Why? With interval data and below, operations like addition, subtraction, multiplication, and division are meaningful!

Parametric statistics:
- Typically take advantage of central limit theorem (imposes requirements on probability distributions)
- Appropriate only for interval and ratio data.
- More powerful than nonparametric methods.

Nonparametric statistics:
- Do not require assumptions concerning the probability distribution for the population.
- There are many methods appropriate for ordinal data, some methods appropriate for nominal data.
- Computations typically make use of data’s ranks instead of actual data.
Different types of data require different statistical methods.

Why? With interval data and below, operations like addition, subtraction, multiplication, and division are *meaningless*!

**Parametric statistics:**
- Typically take advantage of central limit theorem (imposes requirements on probability distributions)
- Appropriate only for interval and ratio data.
- More **powerful** than nonparametric methods.

**Nonparametric statistics:**
- Do not require assumptions concerning the probability distribution for the population.
- There are many methods appropriate for ordinal data, some methods appropriate for nominal data.
- Computations typically make use of data’s *ranks* instead of actual data.
Different types of data require different statistical methods.

Why? With interval data and below, operations like addition, subtraction, multiplication, and division are *meaningless*!

_parametric* statistics:
- Typically take advantage of central limit theorem (imposes requirements on probability distributions)
- Appropriate only for interval and ratio data.
- More *powerful* than nonparametric methods.

_nonparametric* statistics:
- Do not require assumptions concerning the probability distribution for the population.
- There are many methods appropriate for ordinal data, some methods appropriate for nominal data.
- Computations typically make use of data’s *ranks* instead of actual data.
Different types of data require different statistical methods.

Why? With interval data and below, operations like addition, subtraction, multiplication, and division are *meaningless*!

**Parametric statistics:**
- Typically take advantage of central limit theorem (imposes requirements on probability distributions)
- Appropriate only for interval and ratio data.
- More *powerful* than nonparametric methods.

**Nonparametric statistics:**
- Do not require assumptions concerning the probability distribution for the population.
- There are many methods appropriate for ordinal data, some methods appropriate for nominal data.
- Computations typically make use of data’s *ranks* instead of actual data.
Different types of data require different statistical methods.

Why? With interval data and below, operations like addition, subtraction, multiplication, and division are meaningless!

Parametric statistics:
- Typically take advantage of central limit theorem (imposes requirements on probability distributions)
- Appropriate only for interval and ratio data.
- More powerful than nonparametric methods.

Nonparametric statistics:
- Do not require assumptions concerning the probability distribution for the population.
- There are many methods appropriate for ordinal data, some methods appropriate for nominal data.
- Computations typically make use of data’s ranks instead of actual data.
Single Mean T-Test

- Test whether the population mean is equal or different to some value.
- Uses the sample mean as its statistic.
- T-test is used instead of Z-test for reasons you may have learned in a statistics class. Interpretation is the same.
- Parametric test that depends on results from Central Limit Theorem.
- Hypotheses
  - Null: The population mean is equal to some specified value.
  - Alternative: The population mean is [greater/less/different] than the value in the null.
Single Mean T-Test

- Test whether the population mean is equal or different to some value.
- Uses the sample mean as its statistic.
- T-test is used instead of Z-test for reasons you may have learned in a statistics class. Interpretation is the same.
- Parametric test that depends on results from Central Limit Theorem.
- Hypotheses
  - Null: The population mean is equal to some specified value.
  - Alternative: The population mean is [greater/less/different] than the value in the null.
Single Mean T-Test

- Test whether the population mean is equal or different to some value.
- Uses the sample mean as its statistic.
- T-test is used instead of Z-test for reasons you may have learned in a statistics class. Interpretation is the same.
- Parametric test that depends on results from Central Limit Theorem.
- Hypotheses
  - Null: The population mean is equal to some specified value.
  - Alternative: The population mean is [greater/less/different] than the value in the null.
Single Mean T-Test

- Test whether the population mean is equal or different to some value.
- Uses the sample mean as its statistic.
- T-test is used instead of Z-test for reasons you may have learned in a statistics class. Interpretation is the same.
- Parametric test that depends on results from Central Limit Theorem.

Hypotheses
- Null: The population mean is equal to some specified value.
- Alternative: The population mean is [greater/less/different] than the value in the null.
Single Mean T-Test

- Test whether the population mean is equal or different to some value.
- Uses the sample mean its statistic.
- T-test is used instead of Z-test for reasons you may have learned in a statistics class. Interpretation is the same.
- Parametric test that depends on results from Central Limit Theorem.
- Hypotheses
  - Null: The population mean is equal to some specified value.
  - Alternative: The population mean is [greater/less/different] than the value in the null.
Single Mean T-Test

- Test whether the population mean is equal or different to some value.
- Uses the sample mean as its statistic.
- T-test is used instead of Z-test for reasons you may have learned in a statistics class. Interpretation is the same.
- Parametric test that depends on results from Central Limit Theorem.
- Hypotheses
  - Null: The population mean is equal to some specified value.
  - Alternative: The population mean is [greater/less/different] than the value in the null.
Single Mean T-Test

- Test whether the population mean is equal or different to some value.
- Uses the sample mean as its statistic.
- T-test is used instead of Z-test for reasons you may have learned in a statistics class. Interpretation is the same.
- Parametric test that depends on results from Central Limit Theorem.
- Hypotheses
  - Null: The population mean is equal to some specified value.
  - Alternative: The population mean is [greater/less/different] than the value in the null.
Suppose we have a dataset of individual schools and the average pay for teachers in each school, and the level of spending as a ratio of the number of students (spending per pupil).

- Show some descriptive statistics for teacher pay and expenditure per pupil.
- Is there statistical evidence that teachers make less than $50,000 per year?
- Is there statistical evidence that expenditure per pupil is more than $7,500?
Example Questions

- Suppose we have a dataset of individual schools and the average pay for teachers in each school, and the level of spending as a ratio of the number of students (spending per pupil).
  - Show some descriptive statistics for teacher pay and expenditure per pupil.
  - Is there statistical evidence that teachers make less than $50,000 per year?
  - Is there statistical evidence that expenditure per pupil is more than $7,500?
Suppose we have a dataset of individual schools and the average pay for teachers in each school, and the level of spending as a ratio of the number of students (spending per pupil).

- Show some descriptive statistics for teacher pay and expenditure per pupil.
- Is there statistical evidence that teachers make less than $50,000 per year?
- Is there statistical evidence that expenditure per pupil is more than $7,500?
Suppose we have a dataset of individual schools and the average pay for teachers in each school, and the level of spending as a ratio of the number of students (spending per pupil).

- Show some descriptive statistics for teacher pay and expenditure per pupil.
- Is there statistical evidence that teachers make less than $50,000 per year?
- Is there statistical evidence that expenditure per pupil is more than $7,500?
Example: Public School Spending


- Download dataset eduspending.sav.

- Conduct the following exercises:
  - Show some descriptive statistics for teacher pay and expenditure per pupil.
  - Is there statistical evidence that teachers make less than $25,000 per year?
  - Is there statistical evidence that expenditure per pupil is more than $3,500?
Example: Public School Spending

- Download dataset eduspending.sav.
- Conduct the following exercises:
  - Show some descriptive statistics for teacher pay and expenditure per pupil.
  - Is there statistical evidence that teachers make less than $25,000 per year?
  - Is there statistical evidence that expenditure per pupil is more than $3,500?
Example: Public School Spending

- Download dataset eduspending.sav.
- Conduct the following exercises:
  - Show some descriptive statistics for teacher pay and expenditure per pupil.
  - Is there statistical evidence that teachers make less than $25,000 per year?
  - Is there statistical evidence that expenditure per pupil is more than $3,500?
Example: Public School Spending

- Download dataset eduspending.sav.
- Conduct the following exercises:
  - Show some descriptive statistics for teacher pay and expenditure per pupil.
  - Is there statistical evidence that teachers make less than $25,000 per year?
  - Is there statistical evidence that expenditure per pupil is more than $3,500?
Example: Public School Spending

- Download dataset eduspending.sav.
- Conduct the following exercises:
  - Show some descriptive statistics for teacher pay and expenditure per pupil.
  - Is there statistical evidence that teachers make less than $25,000 per year?
  - Is there statistical evidence that expenditure per pupil is more than $3,500?
Example: Public School Spending

- Download dataset eduspending.sav.
- Conduct the following exercises:
  - Show some descriptive statistics for teacher pay and expenditure per pupil.
  - Is there statistical evidence that teachers make less than $25,000 per year?
  - Is there statistical evidence that expenditure per pupil is more than $3,500?
Proportion: Percentage of times some characteristic occurs.

Example: percentage of consumers of soda who prefer Pepsi over Coke.

Sample proportion = \frac{\text{Number of items that has characteristic}}{\text{sample size}}

Example questions:
- Are more than 50% of potential voters most likely to vote for Barack Obama in the next presidential election?
- Suppose typical brand-loyalty turn-over in the mobile phone industry is 10%. Is there statistical evidence that AT&T has brand-loyalty turnover more than 10%?
Proportion: Percentage of times some characteristic occurs.

Example: percentage of consumers of soda who prefer Pepsi over Coke.

Sample proportion = \[ \frac{\text{Number of items that has characteristic}}{\text{sample size}} \]

Example questions:

- Are more than 50% of potential voters most likely to vote for Barack Obama in the next presidential election?
- Suppose typical brand-loyalty turn-over in the mobile phone industry is 10%. Is there statistical evidence that AT&T has brand-loyalty turnover more than 10%?
Single Proportion T-Test

- **Proportion**: Percentage of times some characteristic occurs.
- Example: percentage of consumers of soda who prefer Pepsi over Coke.

\[
\text{Sample proportion} = \frac{\text{Number of items that has characteristic}}{\text{sample size}}
\]

- Example questions:
  - Are more than 50% of potential voters most likely to vote for Barack Obama in the next presidential election?
  - Suppose typical brand-loyalty turn-over in the mobile phone industry is 10%. Is there statistical evidence that AT&T has brand-loyalty turnover more than 10%?
Proportion: Percentage of times some characteristic occurs.
Example: percentage of consumers of soda who prefer Pepsi over Coke.

Sample proportion = \( \frac{\text{Number of items that has characteristic}}{\text{sample size}} \)

Example questions:
- Are more than 50% of potential voters most likely to vote for Barack Obama in the next presidential election?
- Suppose typical brand-loyalty turnover in the mobile phone industry is 10%. Is there statistical evidence that AT&T has brand-loyalty turnover more than 10%?
Proportion: Percentage of times some characteristic occurs.

Example: percentage of consumers of soda who prefer Pepsi over Coke.

Sample proportion = \[ \frac{\text{Number of items that has characteristic}}{\text{sample size}} \]

Example questions:
- Are more than 50% of potential voters most likely to vote for Barack Obama in the next presidential election?
- Suppose typical brand-loyalty turn-over in the mobile phone industry is 10%. Is there statistical evidence that AT&T has brand-loyalty turnover more than 10%?
**Proportion:** Percentage of times some characteristic occurs.

Example: percentage of consumers of soda who prefer Pepsi over Coke.

\[
\text{Sample proportion} = \frac{\text{Number of items that has characteristic}}{\text{sample size}}
\]

Example questions:

- Are more than 50% of potential voters most likely to vote for Barack Obama in the next presidential election?
- Suppose typical brand-loyalty turnover in the mobile phone industry is 10%. Is there statistical evidence that AT&T has brand-loyalty turnover more than 10%?
Data from Montana residents in 1992 concerning their outlook for the economy.

All data is ordinal or nominal:
- AGE = 1 under 35, 2 35-54, 3 55 and over
- SEX = 0 male, 1 female
- INC = yearly income: 1 under $20K, 2 20-35$K, 3 over $35K
- POL = 1 Democrat, 2 Independent, 3 Republican
- AREA = 1 Western, 2 Northeastern, 3 Southeastern Montana
- FIN = Financial status 1 worse, 2 same, 3 better than a year ago
- STAT = 0, State economic outlook better, 1 not better than a year ago

Do the majority of Montana residents feel their financial status is the same or better than one year ago?

Do the majority of Montana residents have a more positive economic outlook than one year ago?
Example: Economic Outlook

- Data from Montana residents in 1992 concerning their outlook for the economy.
- All data is ordinal or nominal:
  - AGE = 1 under 35, 2 35-54, 3 55 and over
  - SEX = 0 male, 1 female
  - INC = yearly income: 1 under $20K, 2 20-35$K, 3 over $35K
  - POL = 1 Democrat, 2 Independent, 3 Republican
  - AREA = 1 Western, 2 Northeastern, 3 Southeastern Montana
  - FIN = Financial status 1 worse, 2 same, 3 better than a year ago
  - STAT = 0, State economic outlook better, 1 not better than a year ago

- Do the majority of Montana residents feel their financial status is the same or better than one year ago?
- Do the majority of Montana residents have a more positive economic outlook than one year ago?
Example: Economic Outlook

- Data from Montana residents in 1992 concerning their outlook for the economy.
- All data is ordinal or nominal:
  - AGE = 1 under 35, 2 35-54, 3 55 and over
  - SEX = 0 male, 1 female
  - INC = yearly income: 1 under $20K, 2 20-35$K, 3 over $35K
  - POL = 1 Democrat, 2 Independent, 3 Republican
  - AREA = 1 Western, 2 Northeastern, 3 Southeastern Montana
  - FIN = Financial status 1 worse, 2 same, 3 better than a year ago
  - STAT = 0, State economic outlook better, 1 not better than a year ago
- Do the majority of Montana residents feel their financial status is the same or better than one year ago?
- Do the majority of Montana residents have a more positive economic outlook than one year ago?
Example: Economic Outlook

- Data from Montana residents in 1992 concerning their outlook for the economy.
- All data is ordinal or nominal:
  - **AGE**: 1 under 35, 2 35-54, 3 55 and over
  - **SEX**: 0 male, 1 female
  - **INC**: yearly income: 1 under $20K, 2 20-35$K, 3 over $35K
  - **POL**: 1 Democrat, 2 Independent, 3 Republican
  - **AREA**: 1 Western, 2 Northeastern, 3 Southeastern Montana
  - **FIN**: Financial status 1 worse, 2 same, 3 better than a year ago
  - **STAT**: 0, State economic outlook better, 1 not better than a year ago
- Do the majority of Montana residents feel their financial status is the same or better than one year ago?
- Do the majority of Montana residents have a more positive economic outlook than one year ago?
Single Median Nonparametric Test

**Why?**

- Ordinal data: cannot compute sample means (they are meaningless), only median is meaningful.
- Small sample size and you are not sure the population is not normal.
- Sign test: can use tests for proportions for testing the median.

For a null hypothesized population median...
- Count how many observations are above the median.
- Test whether that proportion is greater, less than, or not equal to 0.5.
- For small sample sizes, use binomial distribution instead of normal distribution.
Why?

- Ordinal data: cannot compute sample means (they are meaningless), only median is meaningful.
- Small sample size and you are not sure the population is not normal.

Sign test: can use tests for proportions for testing the median.

- For a null hypothesized population median...
- Count how many observations are above the median.
- Test whether that proportion is greater, less than, or not equal to 0.5.
- For small sample sizes, use binomial distribution instead of normal distribution.
Single Median Nonparametric Test

Why?
- Ordinal data: cannot compute sample means (they are meaningless), only median is meaningful.
- Small sample size and you are not sure the population is not normal.
- Sign test: can use tests for proportions for testing the median.

For a null hypothesized population median...
- Count how many observations are above the median.
- Test whether that proportion is greater, less than, or not equal to 0.5.
- For small sample sizes, use binomial distribution instead of normal distribution.
Why?

- Ordinal data: cannot compute sample means (they are meaningless), only median is meaningful.
- Small sample size and you are not sure the population is not normal.

Sign test: can use tests for proportions for testing the median.

- For a null hypothesized population median...
- Count how many observations are above the median.
- Test whether that proportion is greater, less than, or not equal to 0.5.
- For small sample sizes, use binomial distribution instead of normal distribution.
Single Median Nonparametric Test

- **Why?**
  - Ordinal data: cannot compute sample means (they are meaningless), only median is meaningful.
  - Small sample size and you are not sure the population is not normal.
  - Sign test: can use tests for proportions for testing the median.

- **For a null hypothesized population median...**
  - Count how many observations are above the median.
  - Test whether that proportion is greater, less than, or not equal to 0.5.
  - For small sample sizes, use binomial distribution instead of normal distribution.
Single Median Nonparametric Test

Why?

- Ordinal data: cannot compute sample means (they are meaningless), only median is meaningful.
- Small sample size and you are not sure the population is not normal.

Sign test: can use tests for proportions for testing the median.

- For a null hypothesized population median...
- Count how many observations are above the median.
- Test whether that proportion is greater, less than, or not equal to 0.5.
- For small sample sizes, use binomial distribution instead of normal distribution.
Why?
- Ordinal data: cannot compute sample means (they are meaningless), only median is meaningful.
- Small sample size and you are not sure the population is not normal.

Sign test: can use tests for proportions for testing the median.
- For a null hypothesized population median...
- Count how many observations are above the median.
- Test whether that proportion is greater, less than, or not equal to 0.5.
- For small sample sizes, use binomial distribution instead of normal distribution.
Why?

- Ordinal data: cannot compute sample means (they are meaningless), only median is meaningful.
- Small sample size and you are not sure the population is not normal.
- Sign test: can use tests for proportions for testing the median.

For a null hypothesized population median...
- Count how many observations are above the median.
- Test whether that proportion is greater, less than, or not equal to 0.5.
- For small sample sizes, use binomial distribution instead of normal distribution.
Instructor evaluations have an ordinal scale: Excellent, Very Good, Good, Poor, Very Poor.

- Is there statistical evidence that the median rating for a professor is below 'Very Good'?

Suppose you have a frequency determination question on your survey. Is this an ordinal scale? Is the median an appropriate measure of center?
Instructor evaluations have an ordinal scale: Excellent, Very Good, Good, Poor, Very Poor.

Is there statistical evidence that the median rating for a professor is below 'Very Good'? 

Suppose you have a frequency determination question on your survey. Is this an ordinal scale? Is the median an appropriate measure of center?
Instructor evaluations have an ordinal scale: Excellent, Very Good, Good, Poor, Very Poor.

Is there statistical evidence that the median rating for a professor is below 'Very Good'?

Suppose you have a frequency determination question on your survey. Is this an ordinal scale? Is the median an appropriate measure of center?
Example: Attitudes Grade School Kids

- Dataset: 438 students in grades 4 through 6 were sampled from three school districts in Michigan. Students ranked from 1 (most important) to 5 (least important) how important grades, sports, being good looking, and having lots of money were to each of them.

- Open dataset gradeschools.sav. Choose second worksheet, titled Data.

- Answer some of these questions:
  - Is the median importance for grades is greater than 3?
  - Is the median importance for money less than 3?
Example: Attitudes Grade School Kids

- Dataset: 438 students in grades 4 through 6 were sampled from three school districts in Michigan. Students ranked from 1 (most important) to 5 (least important) how important grades, sports, being good looking, and having lots of money were to each of them.

- Open dataset gradeschools.sav. Choose second worksheet, titled Data.

- Answer some of these questions:
  - Is the median importance for grades is greater than 3?
  - Is the median importance for money less than 3?
Example: Attitudes Grade School Kids

Dataset: 438 students in grades 4 through 6 were sampled from three school districts in Michigan. Students ranked from 1 (most important) to 5 (least important) how important grades, sports, being good looking, and having lots of money were to each of them.

Open dataset gradeschools.sav. Choose second worksheet, titled Data.

Answer some of these questions:
- Is the median importance for grades is greater than 3?
- Is the median importance for money less than 3?
Example: Attitudes Grade School Kids

- Dataset: 438 students in grades 4 through 6 were sampled from three school districts in Michigan. Students ranked from 1 (most important) to 5 (least important) how important grades, sports, being good looking, and having lots of money were to each of them.

- Open dataset gradeschools.sav. Choose second worksheet, titled Data.

- Answer some of these questions:
  - Is the median importance for grades is greater than 3?
  - Is the median importance for money less than 3?
Dataset: 438 students in grades 4 through 6 were sampled from three school districts in Michigan. Students ranked from 1 (most important) to 5 (least important) how important grades, sports, being good looking, and having lots of money were to each of them.

Open dataset gradeschools.sav. Choose second worksheet, titled Data.

Answer some of these questions:
- Is the median importance for grades is greater than 3?
- Is the median importance for money less than 3?
Suppose you want to know whether the mean from one population is larger than the mean for another.

Independent samples means you have different individuals in your two sample groups.

Examples:
- Compare sales volume for stores that advertise versus those that do not.
- Compare production volume for employees that have completed some type of training versus those who have not.

Statistic: Difference in the sample means ($\bar{x}_1 - \bar{x}_2$).

Hypotheses:
- Null hypothesis: the difference between the two means is zero.
- Alternative hypothesis: the difference is [above/below/not equal] to zero.
Suppose you want to know whether the mean from one population is larger than the mean for another. Independent samples means you have different individuals in your two sample groups.

Examples:

- Compare sales volume for stores that advertise versus those that do not.
- Compare production volume for employees that have completed some type of training versus those who have not.

Statistic: Difference in the sample means ($\bar{x}_1 - \bar{x}_2$).

Hypotheses:

- Null hypothesis: the difference between the two means is zero.
- Alternative hypothesis: the difference is [above/below/not equal] to zero.
Suppose you want to know whether the mean from one population is larger than the mean for another. Independent samples means you have different individuals in your two sample groups.

Examples:

- Compare sales volume for stores that advertise versus those that do not.
- Compare production volume for employees that have completed some type of training versus those who have not.

Statistic: Difference in the sample means ($\bar{x}_1 - \bar{x}_2$).

Hypotheses:

- Null hypothesis: the difference between the two means is zero.
- Alternative hypothesis: the difference is [above/below/not equal] to zero.
Suppose you want to know whether the mean from one population is larger than the mean for another. Independent samples means you have different individuals in your two sample groups.

**Examples:**

- Compare sales volume for stores that advertise versus those that do not.
- Compare production volume for employees that have completed some type of training versus those who have not.

**Statistic:** Difference in the sample means ($\bar{x}_1 - \bar{x}_2$).

**Hypotheses:**

- Null hypothesis: the difference between the two means is zero.
- Alternative hypothesis: the difference is [above/below/not equal] to zero.
Suppose you want to know whether the mean from one population is larger than the mean for another. Independent samples means you have different individuals in your two sample groups.

Examples:
- Compare sales volume for stores that advertise versus those that do not.
- Compare production volume for employees that have completed some type of training versus those who have not.

Statistic: Difference in the sample means ($\bar{x}_1 - \bar{x}_2$).

Hypotheses:
- Null hypothesis: the difference between the two means is zero.
- Alternative hypothesis: the difference is [above/below/not equal] to zero.
Suppose you want to know whether the mean from one population is larger than the mean for another. Independent samples means you have different individuals in your two sample groups.

Examples:
- Compare sales volume for stores that advertise versus those that do not.
- Compare production volume for employees that have completed some type of training versus those who have not.

Statistic: Difference in the sample means ($\bar{x}_1 - \bar{x}_2$).

Hypotheses:
- Null hypothesis: the difference between the two means is zero.
- Alternative hypothesis: the difference is [above/below/not equal] to zero.
Suppose you want to know whether the mean from one population is larger than the mean for another.

Independent samples means you have different individuals in your two sample groups.

Examples:
- Compare sales volume for stores that advertise versus those that do not.
- Compare production volume for employees that have completed some type of training versus those who have not.

Statistic: Difference in the sample means ($\bar{x}_1 - \bar{x}_2$).

Hypotheses:
- Null hypothesis: the difference between the two means is zero.
- Alternative hypothesis: the difference is [above/below/not equal] to zero.
Suppose you want to know whether the mean from one population is larger than the mean for another.

Independent samples means you have different individuals in your two sample groups.

Examples:
- Compare sales volume for stores that advertise versus those that do not.
- Compare production volume for employees that have completed some type of training versus those who have not.

Statistic: Difference in the sample means ($\bar{x}_1 - \bar{x}_2$).

Hypotheses:
- Null hypothesis: the difference between the two means is zero.
- Alternative hypothesis: the difference is [above/below/not equal] to zero.
Suppose you want to know whether the mean from one population is larger than the mean for another.

Independent samples means you have different individuals in your two sample groups.

Examples:
- Compare sales volume for stores that advertise versus those that do not.
- Compare production volume for employees that have completed some type of training versus those who have not.

Statistic: Difference in the sample means ($\bar{x}_1 - \bar{x}_2$).

Hypotheses:
- Null hypothesis: the difference between the two means is zero.
- Alternative hypothesis: the difference is [above/below/not equal] to zero.
Example

- Test the following hypotheses:
  - Does spending per pupil differ in the North (region 1) and the South (region 2)?
  - Does teacher salary differ in the North and the West (region 3)?
- Do you see any weaknesses in our statistical analysis?
Example


- Test the following hypotheses:
  - Does spending per pupil differ in the North (region 1) and the South (region 2)?
  - Does teacher salary differ in the North and the West (region 3)?
  - Do you see any weaknesses in our statistical analysis?
Example

- Test the following hypotheses:
  - Does spending per pupil differ in the North (region 1) and the South (region 2)?
  - Does teacher salary differ in the North and the West (region 3)?
- Do you see any weaknesses in our statistical analysis?
Example

- Test the following hypotheses:
  - Does spending per pupil differ in the North (region 1) and the South (region 2)?
  - Does teacher salary differ in the North and the West (region 3)?
- Do you see any weaknesses in our statistical analysis?
Example

- Test the following hypotheses:
  - Does spending per pupil differ in the North (region 1) and the South (region 2)?
  - Does teacher salary differ in the North and the West (region 3)?
- Do you see any weaknesses in our statistical analysis?
Mann-Whitney U test: nonparametric test to determine difference in medians.

Can you suggest some examples?

Assumptions:

- Samples are independent of one another.
- The underlying distributions have the same shape (i.e. only the location of the distribution is different).
- It has been argued that violating the second assumption does not severely change the sampling distribution of the Mann-Whitney U test.

Null hypothesis: medians for the two populations are the same.

Alternative hypotheses: medians for the two populations are different.
Mann-Whitney U test: nonparametric test to determine difference in medians.

Can you suggest some examples?

Assumptions:
- Samples are independent of one another.
- The underlying distributions have the same shape (i.e., only the location of the distribution is different).
- It has been argued that violating the second assumption does not severely change the sampling distribution of the Mann-Whitney U test.

Null hypothesis: medians for the two populations are the same.

Alternative hypotheses: medians for the two populations are different.
**Nonparametric Tests for Differences in Medians**

- **Mann-Whitney U test**: nonparametric test to determine difference in *medians*.

- **Can you suggest some examples?**

- **Assumptions:**
  - Samples are independent of one another.
  - The underlying distributions have the same shape (i.e. only the location of the distribution is different).
  - It has been argued that violating the second assumption does not severely change the sampling distribution of the Mann-Whitney U test.

- **Null hypothesis**: medians for the two populations are the same.

- **Alternative hypotheses**: medians for the two populations are different.
Mann-Whitney U test: nonparametric test to determine difference in *medians*.

Can you suggest some examples?

**Assumptions:**

- Samples are independent of one another.
- The underlying distributions have the same shape (i.e. only the location of the distribution is different).
- It has been argued that violating the second assumption does not severely change the sampling distribution of the Mann-Whitney U test.

**Null hypothesis:** medians for the two populations are the same.

**Alternative hypotheses:** medians for the two populations are different.
Nonparametric Tests for Differences in Medians

- Mann-Whitney U test: nonparametric test to determine difference in medians.
- Can you suggest some examples?
- Assumptions:
  - Samples are independent of one another.
  - The underlying distributions have the same shape (i.e. only the location of the distribution is different).
  - It has been argued that violating the second assumption does not severely change the sampling distribution of the Mann-Whitney U test.
- Null hypothesis: medians for the two populations are the same.
- Alternative hypotheses: medians for the two populations are different.
Nonparametric Tests for Differences in Medians

- Mann-Whitney U test: nonparametric test to determine difference in *medians*.

- Can you suggest some examples?

- Assumptions:
  - Samples are independent of one another.
  - The underlying distributions have the same shape (i.e. only the location of the distribution is different).
  - It has been argued that violating the second assumption does not severely change the sampling distribution of the Mann-Whitney U test.

- Null hypothesis: medians for the two populations are the same.

- Alternative hypotheses: medians for the two populations are different.
Nonparametric Tests for Differences in Medians

- Mann-Whitney U test: nonparametric test to determine difference in *medians*.
- Can you suggest some examples?
- Assumptions:
  - Samples are independent of one another.
  - The underlying distributions have the same shape (i.e. only the location of the distribution is different).
  - It has been argued that violating the second assumption does not severely change the sampling distribution of the Mann-Whitney U test.
- Null hypothesis: medians for the two populations are the same.
- Alternative hypotheses: medians for the two populations are different.
Nonparametric Tests for Differences in Medians

- Mann-Whitney U test: nonparametric test to determine difference in *medians*.
- Can you suggest some examples?
- Assumptions:
  - Samples are independent of one another.
  - The underlying distributions have the same shape (i.e. only the location of the distribution is different).
  - It has been argued that violating the second assumption does not severely change the sampling distribution of the Mann-Whitney U test.
- Null hypothesis: medians for the two populations are the same.
- Alternative hypotheses: medians for the two populations are different.
Use a **paired sampled test** if instead the two samples have the same individuals before and after some treatment.

**Examples:**
- The Biggest Loser: Compare the weight of people on the show before the season begins and one year after the show concludes.
- Training session: Are workers more productive 6 months after they attended some training session versus before the training session.

Really simple: for each individual subtract the before treatment measure from the after treatment measure (or vice-versa).

Treat your new series as a single series.

Conduct one-sample tests.
Use a **paired sampled test** if instead the two samples have the same individuals before and after some treatment.

**Examples:**

- The Biggest Loser: Compare the weight of people on the show before the season begins and one year after the show concludes.
- Training session: Are workers more productive 6 months after they attended some training session versus before the training session.
- Really simple: for each individual subtract the before treatment measure from the after treatment measure (or vice-versa).
- Treat your new series as a single series.
- Conduct one-sample tests.
Use a **paired sampled test** if instead the two samples have the same individuals before and after some treatment.

Examples:

- The Biggest Loser: Compare the weight of people on the show before the season begins and one year after the show concludes.
- Training session: Are workers more productive 6 months after they attended some training session versus before the training session.

Really simple: for each individual subtract the before treatment measure from the after treatment measure (or vice-versa).

Treat your new series as a single series.

Conduct one-sample tests.
Use a **paired sampled test** if instead the two samples have the same individuals before and after some treatment.

**Examples:**
- The Biggest Loser: Compare the weight of people on the show before the season begins and one year after the show concludes.
- Training session: Are workers more productive 6 months after they attended some training session versus before the training session.

Really simple: for each individual subtract the before treatment measure from the after treatment measure (or vice-versa).

Treat your new series as a single series.

Conduct one-sample tests.
**Dependent Samples - Paired Samples**

- **Use a paired sampled test** if instead the two samples have the same individuals before and after some treatment.
- **Examples:**
  - The Biggest Loser: Compare the weight of people on the show before the season begins and one year after the show concludes.
  - Training session: Are workers more productive 6 months after they attended some training session versus before the training session.
- **Really simple:** for each individual subtract the before treatment measure from the after treatment measure (or vice-versa).
- **Treat your new series as a single series.**
- **Conduct one-sample tests.**
Use a **paired sampled test** if instead the two samples have the same individuals before and after some treatment.

**Examples:**
- **The Biggest Loser:** Compare the weight of people on the show before the season begins and one year after the show concludes.
- **Training session:** Are workers more productive 6 months after they attended some training session versus before the training session.

**Really simple:** for each individual subtract the before treatment measure from the after treatment measure (or vice-versa).

**Treat your new series as a single series.**

**Conduct one-sample tests.**
Use a **paired sampled test** if instead the two samples have the same individuals before and after some treatment.

**Examples:**
- The Biggest Loser: Compare the weight of people on the show before the season begins and one year after the show concludes.
- Training session: Are workers more productive 6 months after they attended some training session versus before the training session.

**Really simple:** for each individual subtract the before treatment measure from the after treatment measure (or vice-versa).

**Treat your new series as a single series.**

**Conduct one-sample tests.**
Ideas to keep in mind:

- What is a sampling distribution? What does it imply about p-values and statistical significance?
- When it is appropriate to use parametric versus non-parametric methods.
- Most univariate and bivariate questions have a parametric and non-parametric approach.

Next class: closed-book, closed note quiz (so study!); and an in-class exercise practicing this stuff in SPSS.

Next: Statistics to understand the relationship between two variables.
Ideas to keep in mind:

- What is a sampling distribution? What does it imply about p-values and statistical significance?
- When it is appropriate to use parametric versus non-parametric methods.
- Most univariate and bivariate questions have a parametric and non-parametric approach.

Next class: closed-book, closed note quiz (so study!); and an in-class exercise practicing this stuff in SPSS.

Next: Statistics to understand the relationship between two variables.
Ideas to keep in mind:

- What is a sampling distribution? What does it imply about p-values and statistical significance?
- When it is appropriate to use parametric versus non-parametric methods.
- Most univariate and bivariate questions have a parametric and non-parametric approach.

Next class: closed-book, closed note quiz (so study!); and an in-class exercise practicing this stuff in SPSS.

Next: Statistics to understand the relationship between two variables.
Ideas to keep in mind:

- What is a sampling distribution? What does it imply about p-values and statistical significance?
- When it is appropriate to use parametric versus non-parametric methods.
- Most univariate and bivariate questions have a parametric and non-parametric approach.

Next class: closed-book, closed note quiz (so study!); and an in-class exercise practicing this stuff in SPSS.

Next: Statistics to understand the relationship between two variables.
Ideas to keep in mind:

- What is a sampling distribution? What does it imply about p-values and statistical significance?
- When it is appropriate to use parametric versus non-parametric methods.
- Most univariate and bivariate questions have a parametric and non-parametric approach.

Next class: closed-book, closed note quiz (so study!); and an in-class exercise practicing this stuff in SPSS.

Next: Statistics to understand the relationship between two variables.
• Ideas to keep in mind:
  • What is a sampling distribution? What does it imply about p-values and statistical significance?
  • When it is appropriate to use parametric versus non-parametric methods.
  • Most univariate and bivariate questions have a parametric and non-parametric approach.

• Next class: closed-book, closed note quiz (so study!); and an in-class exercise practicing this stuff in SPSS.

• Next: Statistics to understand the relationship between two variables.