# Course 3 Capstone

## Data Collection

# Finding the Middle

Mean, Median, and Mode help you compare data. Below, list the mean, median, and mode of the clicks in the provided data.

Mean: 60.38

Median: 60.00

Mode: 78.00

# Finding the Middle

Mean, Median, and Mode help you compare data. Below, list the mean, median, and mode of the conversions in the provided data.

Mean: 5.98

Median: 6.00

Mode: 5.00

# Standard Deviation

Determining variance in data helps you [why this is helpful]. Below, enter the standard deviation of the provided data.

Standard Deviation of Clicks: 14.37

Standard Deviation of Conversions: 1.63

# Frequency and Contingency Tables

Understanding how often something happens is important to understanding trends and patterns in your data. Create and insert a contingency table generated from your data.

Placeholder, replace this image

- CMD+Click or Right Click and select "Replace Image"
- Then, select the visualization, graph, etc. that you want to include

# Scatter Plot

Understanding the ==relationships between data is important to understanding trends and patterns==. Create and insert a scatter plot generated from your data. Then, include the input the correlation coefficient as well.

Scatter Plot of your data:

Correlation coefficient: 0.45

AdWords Ad Conversions vs. AdWords Ad Clicks
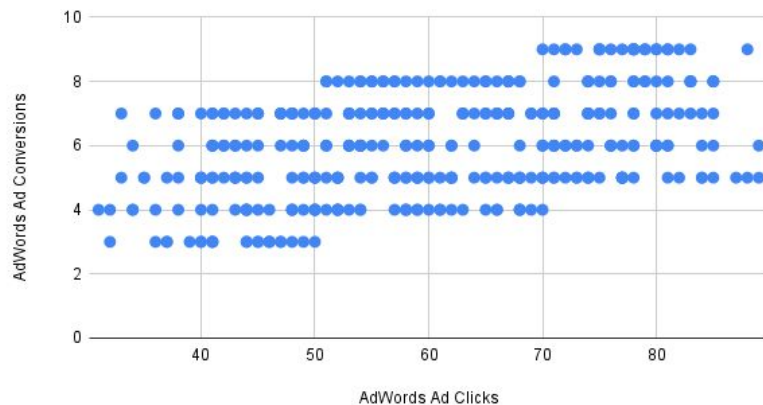
# End of Section 1

# Sample Type

It's important to understand the sample you're using in your analysis. Fill in the information below about the sample you have received:
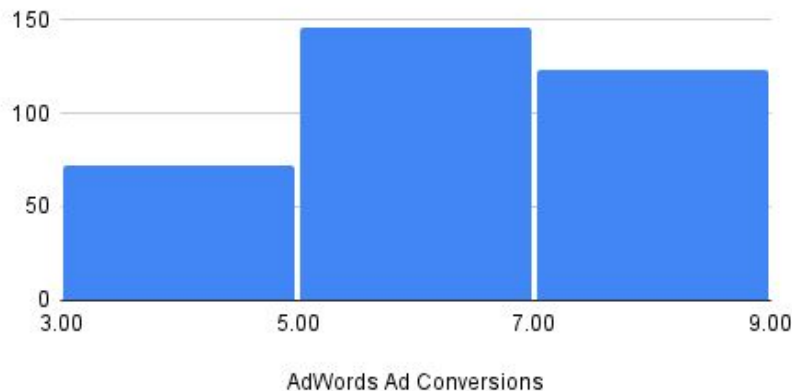
Histogram of your clicks data:

Histogram of conversions data:



Histogram of AdWords Ad Clicks



Histogram of AdWords Ad Conversions

# Sample Type

Does the clicks data have a normal distribution? Yes

Does the conversions data have a normal distribution? Yes

# Variable Types

Determining the types of variables your working with is an important skill. Below, list the variables from your data that are:

Quantitative:

Continuous: AdWords Click-Through Rate (Clicks / View), AdWords Conversion Rate (Conversions / Click), AdWords Cost per Click (Ad Cost / Clicks)

Discrete: AdWords Ad Views, AdWords Ad Clicks, AdWords Ad Conversions, Cost per AdWords Ad

Qualitative:

Ordinal: AdWords Ad Campaign

# End of Section 2

# Question and Hypothesis

The question you hope to answer and your hypothesized answer are necessary to complete an analysis. Answer the following questions

What is your hypothesis based off the evaluation question? Our number of conversions will be greater if we advertise on Facebook rather than AdWords

# Question and Hypothesis

What is your independent variable? Conversions

What is your dependent variable? Campaign

# Running a Test

With your question and hypothesis ready, run the test on the two sets of data. Fill in the information below.

Mean number of Facebook conversions: 11.74

Mean number of Adware conversions: 5.98

p-Value: 0

# Hypothesis

After running the test, was your hypothesis proven correct? Yes

Do your findings support a null or an alternative hypothesis? Alternative

My data tells me that indeed, there is a difference between population A and population B, population A is greater.

# End of Section 3

# Determining a Model

Based off what you know so far, you'll need to determine if your data meets the assumptions for a chosen model. Including:

Simple Linear Regression is the most appropriate choice as it directly addresses the question and allows for the estimation of the expected number of Facebook ad conversions given a specific number of ad clicks.

# Modeling

Finally, include a visualization of your complete model.



Expected Conversions by Number of Clicks
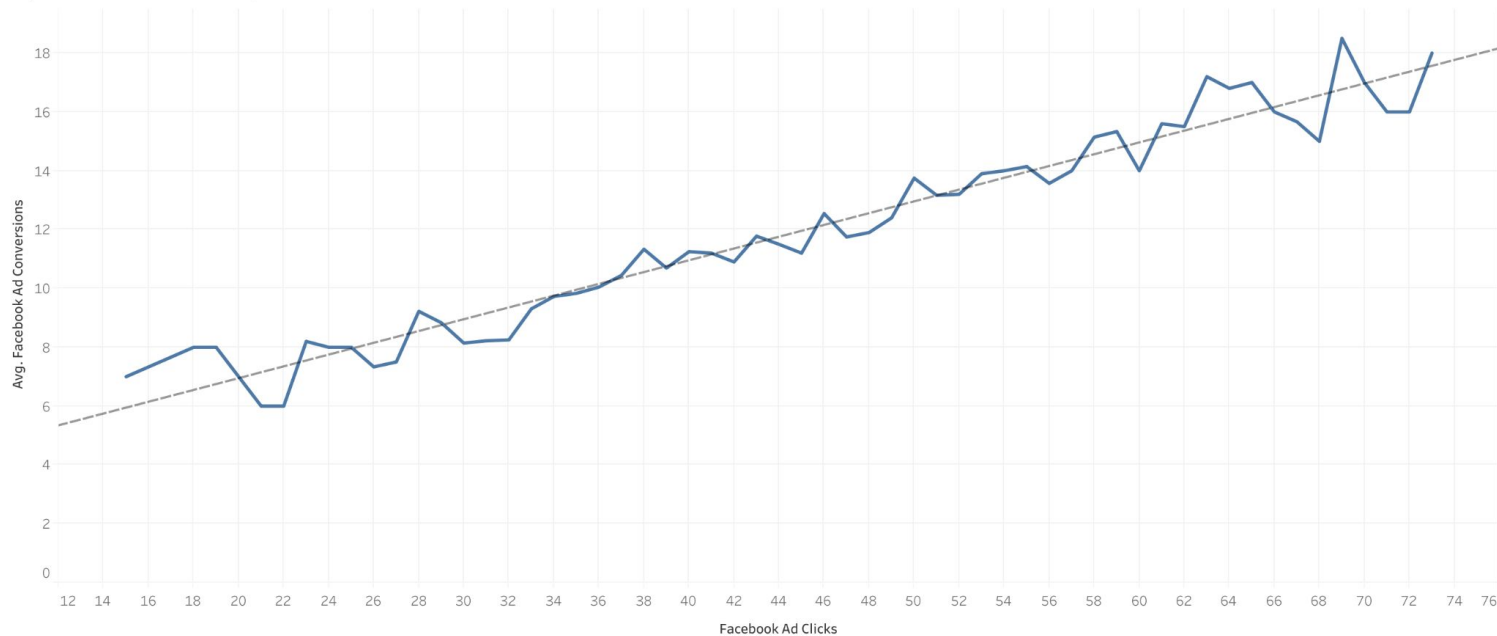
# End of Section 4

# Final Insights

Now, knowing what you do about the results of your test, what are the final insights that you would share with your client? What did you learn and what would you recommend? Is there anything you would do differently next time?

Based on the analysis, it is recommended to use Simple Linear Regression to predict the number of Facebook ad conversions based on ad clicks, while considering the standard deviation as a measure of variability in the data; for future improvements, exploring alternative models and seeking domain expertise are advised.