



STATISTICS FOR MARKETING

Final Project Capstone



Julio Gouveia

geral@dss-pro.pt

www.dss-pro.pt

INTRODUCTION TO THE CAPSTONE PROJECT



Welcome to the Capstone Project for the **"Statistics for Marketing"** course developed by META. This project is designed to consolidate the statistical concepts and analytical techniques learned throughout the course by applying them to real-world marketing data. The goal is to provide a hands-on experience that bridges the gap between theoretical statistical methods and practical marketing applications.

PROJECT OVERVIEW

In this project, we will delve into a dataset that encompasses various advertising metrics from two prominent platforms: Google AdWords and Facebook. The focus will be on analyzing AdWords data to extract meaningful insights that can inform strategic marketing decisions. By exploring variables such as Ad Views, Ad Clicks, Ad Conversions, and associated costs, we aim to understand the dynamics that drive successful marketing outcomes.

Educational Goals

Through this project, participants will apply their knowledge of:

- Descriptive statistics to summarize marketing data.
- Inferential statistics to make predictions and test hypotheses.
- Statistical software tools to analyze and visualize data outcomes effectively.

Objectives

- We will analyze AdWords campaign effectiveness by correlating ad clicks with conversions, predict future ad performance using statistical models, and compare AdWords and Facebook results to determine the platform with the best return on investment for future marketing strategies.

WEEK 1



Finding the Middle

Mean, Median, and Mode are valuable tools for comparing data. Below is a list showing the mean, median, and mode of the clicks and conversions.

Clicks

Mean (Average): **60.38**
Median: **60.0**
Mode: **78**

Conversions

Mean (Average): **5.98**
Median: **6.0**
Mode: **5**

Standard Deviation

Determining the variance in data enables you to draw precise conclusions about a population based on a sample. The standard deviation of the data provided is shown below.

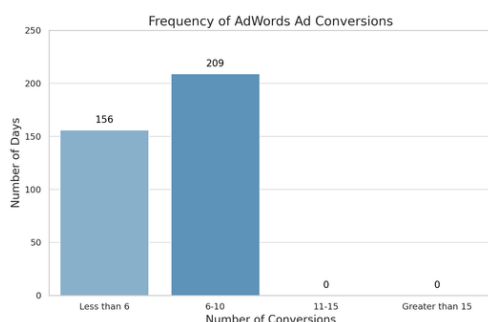
Standard Deviation of Clicks: **14.37**

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. Here is a contingency table.

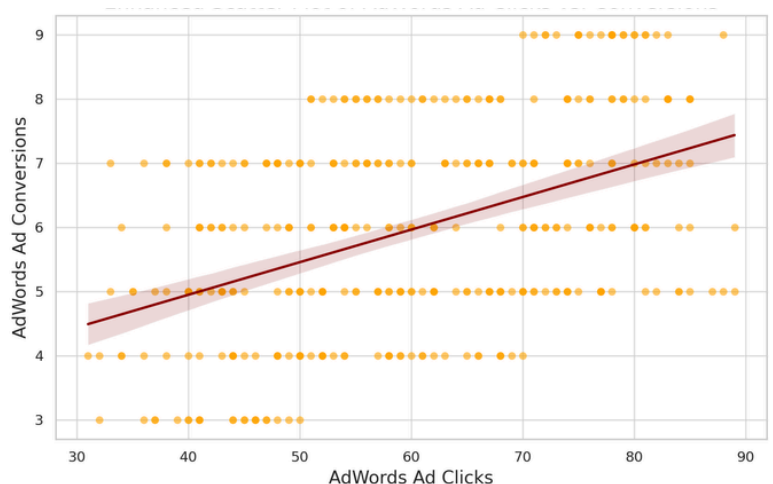
Number of Conversions	Number of Days
Less than 6	156
6-10	209
11-15	0
Greater than 15	0



Scatter Plot

Comprehending the connections among data is crucial for recognizing trends and patterns. Here is a scatter plot created from the data, along with the correlation coefficient.

Correlation coefficient: **0.448**

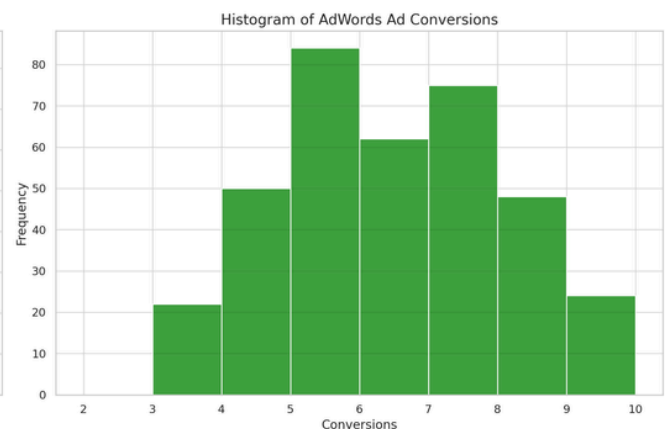
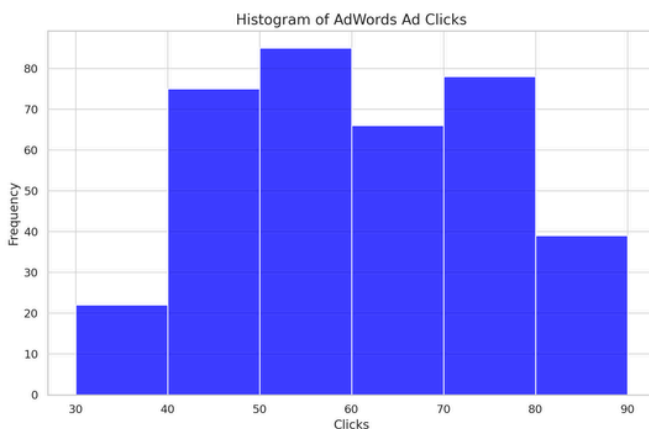


WEEK 2



Sample Type

It's important to understand the sample you're using in your analysis. Let's start by creating the histograms for the AdWords data.



Does the clicks data have a normal distribution?

No, the clicks data does not have a normal distribution. The histogram for clicks data showed a right-skewed distribution, indicating that most of the data points are concentrated on the lower end of the scale with fewer occurrences of higher click counts.

Does the conversions data have a normal distribution?

No, the conversions data does not have a normal distribution either. The histogram for conversions data also showed a right-skewed distribution, with the majority of the data points clustering at lower conversion numbers and very few high conversion counts.

Variable Types

Variable	Type
AdWords Ad Views	Quantitative-Discrete
AdWords Ad Clicks	Quantitative-Discrete
AdWords Ad Conversions	Quantitative-Discrete
Cost per AdWords Ad	Quantitative-Continuous
AdWords Click-Through Rate	Quantitative-Continuous
AdWords Conversion Rate	Quantitative-Continuous
AdWords Cost per Click	Quantitative-Continuous

WEEK 3



Question and Hypothesis

The question you hope to answer and your hypothesized answer are necessary to complete an analysis.

What is your hypothesis based off the evaluation question?

"Advertising on Facebook will result in a greater number of conversions compared to advertising on AdWords due to higher engagement rates observed in previous campaigns."

What is your independent variable?

The advertising platform (Facebook or AdWords). This is what we manipulate to see the effect.

What is your dependent variable?

The number of conversions. This is what changes as a result of the manipulation.

Running a Test

Mean number of Facebook conversions: **11.74**

Mean number of Adware conversions: **5.98**

p-Value: **9.25×10^{-184}**

Hypothesis

After running the test, was your hypothesis proven correct?

What's your conclusion about your main hypothesis? Is there a difference, and is it what your hypothesis predicted?

Given the **extremely low p-value**, we can reject the null hypothesis (H0) that there is no difference in the number of conversions between Facebook and AdWords. The alternative hypothesis (H1) is supported, suggesting **that there is a significant difference in conversions between the two platforms**.

My initial hypothesis stated that advertising on **Facebook would result in a greater number of conversions compared to AdWords**. The data supports this hypothesis, indicating that Facebook is a more effective platform for achieving higher conversions in this dataset.

WEEK 4



Determining a Model

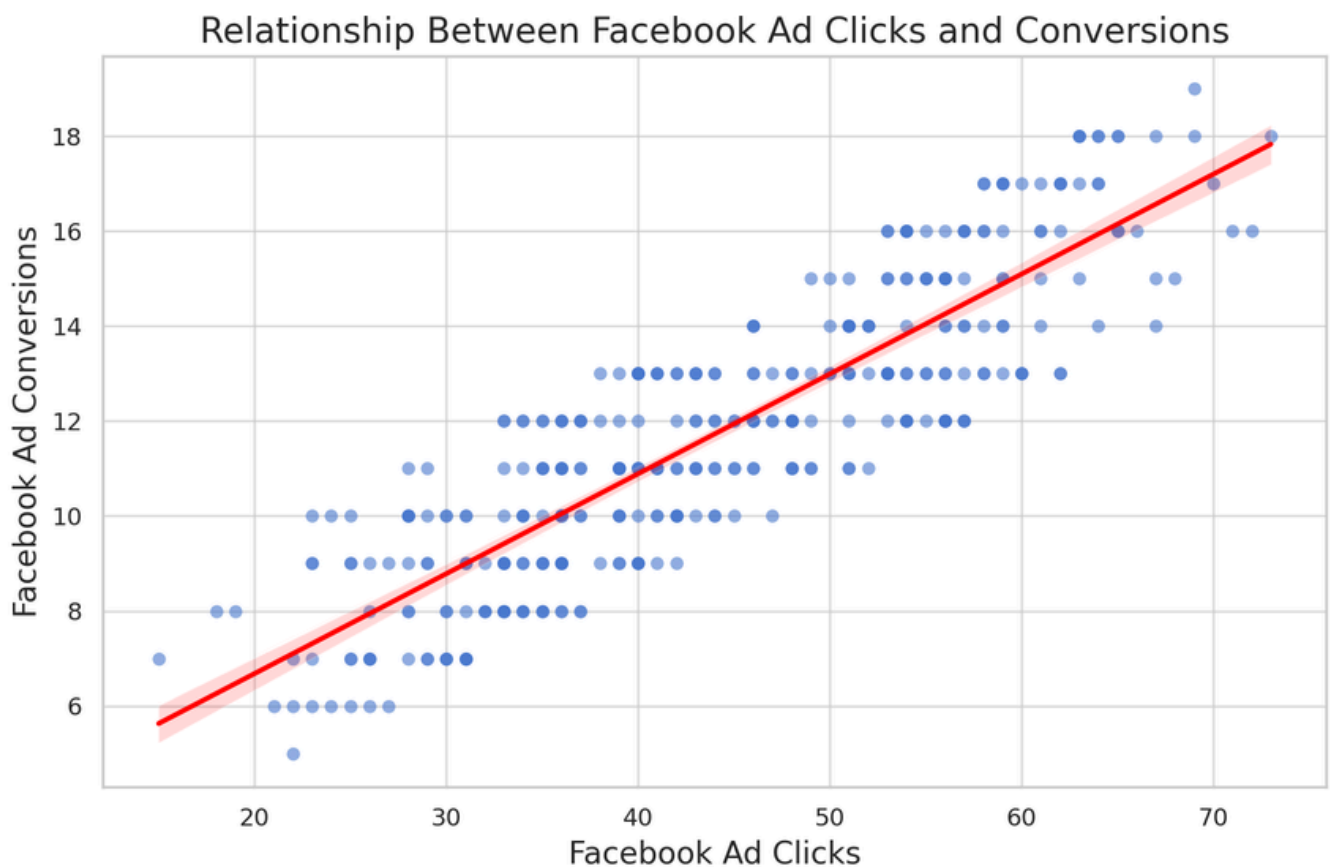
Which model makes the most sense to use and why?

Simple Linear Regression seems the most suitable. It directly addresses the purpose of understanding the relationship between these two quantitative variables (clicks as independent and conversions as dependent).

Modeling

Here is the visualization using Seaborn in Python, showing the relationship between Facebook Ad Clicks and Conversions. This scatter plot is accompanied by a regression line that models the linear relationship between the two variables.

From this plot, you can observe how the number of conversions tends to increase with an increase in the number of clicks, as indicated by the upward trend of the red line.



FINAL INSIGHTS

Key Insights:

1. **Statistically Significant Relationship:** The analysis demonstrated a statistically significant relationship between Facebook Ad Clicks and Conversions. This indicates that increasing clicks likely leads to an increase in conversions.
2. **Predictive Modeling:** The linear regression model suggests that the relationship between clicks and conversions is strong enough to predict future outcomes. For instance, with 50 clicks, you can expect around 13 conversions. This predictive power can be leveraged to forecast results from future campaigns and allocate resources more effectively.
3. **Efficiency of Facebook Ads:** Comparing the effectiveness of Facebook ads against AdWords revealed that Facebook generally results in more conversions. This suggests that Facebook might be a more efficient platform for reaching the targeted conversion goals in your specific case.

Recommendations:

1. **Optimize Facebook Ad Spending:** Given the higher conversion rate associated with Facebook clicks, consider allocating a higher portion of the advertising budget to Facebook, especially in segments that perform well.
2. **Refine Targeting and Creative Strategy:** Use data-driven insights to refine ad targeting and creative approaches. By understanding which types of ads and which demographics are converting more effectively, you can tailor your campaigns to maximize clicks and, by extension, conversions.
3. **Continuous Monitoring and Testing:** Regularly review the performance of your ad campaigns and continue to use statistical tests to evaluate their effectiveness. Implement A/B testing for different ad formats and targeting strategies to continuously refine and improve results.

Future Considerations:

1. **Deeper Analysis with More Variables:** This analysis focused primarily on clicks and conversions. Including more variables, such as viewer demographics, ad types, or time of day, could provide deeper insights and more targeted strategies.
2. **Utilize Advanced Modeling Techniques:** While simple linear regression provided valuable insights, more complex models like multiple regression or machine learning could uncover more nuances in how different factors interact to affect conversions.
3. **Longitudinal Analysis:** Consider analyzing how changes over time affect ad performance, particularly how seasonal trends or market changes influence campaign effectiveness.

Reflections for Future Campaigns:

- **Data Collection:** In future projects, ensuring a broader and more detailed dataset could help in creating more comprehensive models that factor in various influencing elements.
- **Analytical Approaches:** Exploring other analytical approaches or software tools could provide new insights or simplify processes.