Google Play Store Report and Analysis

Does price affect downloads for the ENTERTAINMENT category?

February 2021 - Catharina van Veen

# Introduction

DigitTech Advertising has acquired a new mobile gaming company client. This client wants to broaden its targeted market in the mobile gaming app industry. But before they can settle on the types of games, they need to know what kind of games are being played and by whom.

The question we are tasked with here is: “Does price affect the number of downloads for the ENTERTAINMENT category?” In this report I will be discussing the data, methods used, analysis performed, and the overall results.

The null hypothesis I will be testing in this report is: The price has no impact on the number of Installs.

# Data

For this report I used the data and information about Google Play Store contained in the provided file, googleplaystoremaster.csv. The data consists of 10,841 rows and 13 columns all related to Google Play Store statistics and metrics. There are 10,841 apps with the following information per app contained within the data: Name, Category, Rating, Number of Reviews, Size, Number of Installs, Type, Price, Content Rating, Genres, Last Updated, Current Version, Android Version.

The data has 34 unique Category values: 'ART\_AND\_DESIGN', ‘AUTO\_AND\_VEHICLES', 'BEAUTY', 'BOOKS\_AND\_REFERENCE', 'BUSINESS', 'COMICS', 'COMMUNICATION', 'DATING', 'EDUCATION', 'ENTERTAINMENT', 'EVENTS', 'FINANCE', 'FOOD\_AND\_DRINK', 'HEALTH\_AND\_FITNESS', 'HOUSE\_AND\_HOME', 'LIBRARIES\_AND\_DEMO', 'LIFESTYLE', 'GAME', 'FAMILY', 'MEDICAL', 'SOCIAL', 'SHOPPING', 'PHOTOGRAPHY', 'SPORTS', 'TRAVEL\_AND\_LOCAL', 'TOOLS', 'PERSONALIZATION', 'PRODUCTIVITY', 'PARENTING', 'WEATHER', 'VIDEO\_PLAYERS', 'NEWS\_AND\_MAGAZINES', 'MAPS\_AND\_NAVIGATION', '1.9'.

Closer inspection showed that the record with 1.9 as Category is corrupt. Since no category is given, I ignored this record and filtered for only the records with ENTERTAINMENT as category. This leaves us with 149 rows.

For the purpose of this report, I needed only the columns: Price, Number of Installs.

The filtered data has 21 unique Installs values: '1,000,000,000+', '100,000,000+', '50,000,000+', '10,000,000+', '5,000,000+', '1,000,000+', '500,000+', '100,000+', '50,000+', and '10,000+' stored as text values. I added a column with these values as numeric values. I kept the original column so that I could generate categorical frequency table as well.

The data has 3 unique Price values: '0', '$4.99', and '$2.99' stored as text values. I converted these to numeric values.

# Method

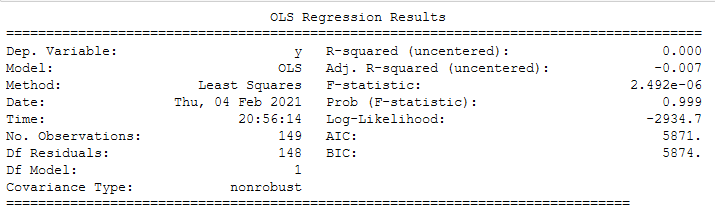
I performed a Simple Linear Regression Analysis.

To get better insight in the data I analyzed the frequency tables of number of Installs for each of the 3 price categories.

# Analysis

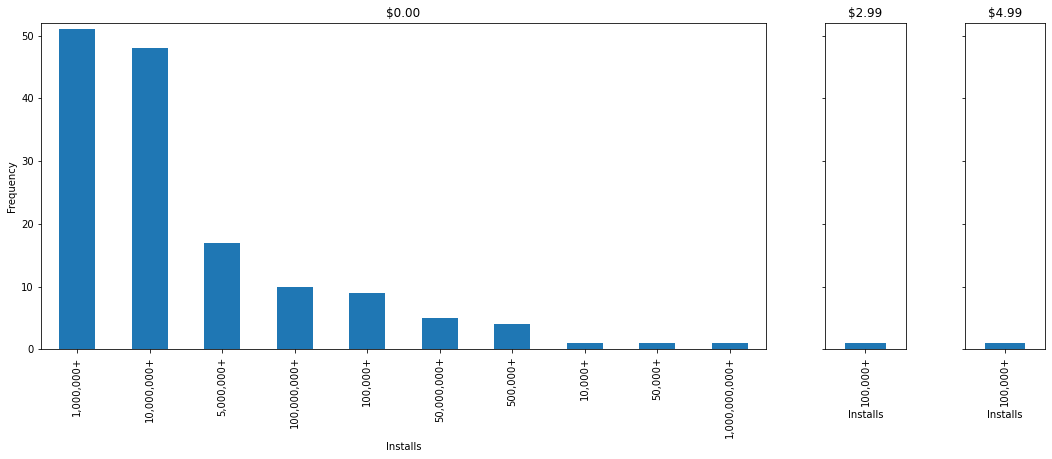
## Linear Regression Analysis

The results of the Simple Linear Regression Analysis are as follows:



Since the p-value is close to 1 the regression is not statistically significant.

## Frequency Tables



From the frequency table it is clear that we do not have enough data for the aps with prices other than $0.00. Each price other than $0.00 has data for only one app. The number of Installs for each of these other prices falls well within the range of the $0.00 priced apps.

# Conclusion

Based on the analysis I do not reject the Null hypothesis: The price has no impact on the number of Installs.

In answer to the question asked, “Does price affect the number of downloads for the ENTERTAINMENT category?”, I conclude that there is insufficient data for the prices other than $0.00 to determine if there is a correlation between Price and number of Installs.