Google Play or Apple Store?

Which one to implement in our OS?

Problem Statement

It is no secret that the mobile and desktop **apps world** has skyrocketed user engagement with electronic devices in the past decade. Thus, choosing the most popular store for apps is a paramount for an OS successful integration.

Steps for Decision Making

1.- SOURCING DATA FROM KAGGLE IN CSV FORMAT TO GET INSIGHTS FROM APPLE API AND GOOGLE WEBPAGE REGARDING APPLE APP STORE AND GOOGLE PLAY STORE .

1.1 Source and Load the data

1.2 Pick useful features

1.3 Subset the data frame

2.- CLEANING, TRANSFORMING AND VISUALIZING DATA FOR GETTING THE CORRECT DATATYPES TO WORK WITH.

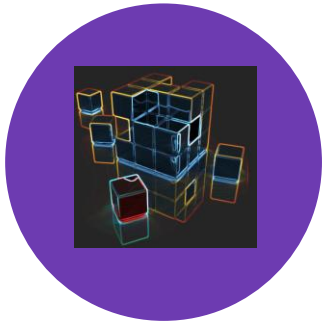
2.1 Check all data types in data frames

1.2 Adding features

1.3 Changing columns names for merging

1.4 Join datasets and eliminate NaN's

1.5 Filter data and summarize data analytically and visually



3.- MODELING THE DATA FOR HYPOTHESIS FORMULATION AND STATISTICAL ANALYSIS.

3.1 Hypothesis formulation

3.2 Getting the distribution of the data

3.3 Permutation testing

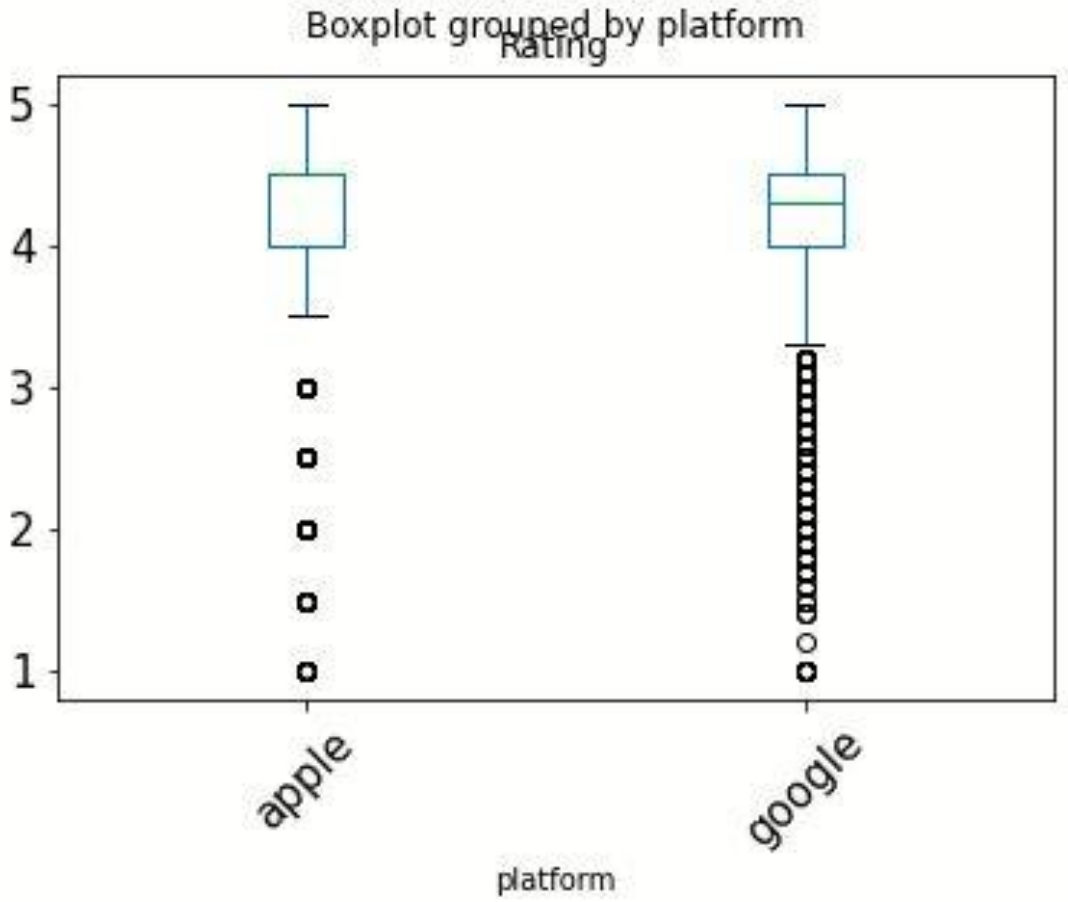
Graphs for Decision Making

We gather the data and combine the two Apple and Google datasets to **group the data by on the ratings by platform**. Being 1 the lowest rating and 5 the maximum rating.

Interesting! Our means of **4.049697** and 4.191757 don't **seem** all that different! Perhaps there is no significant difference between Google and Apple reviews??

Observed Difference:

(4.191757 - 4.049697) = **0.14206.** Actual difference that we observed between the mean rating for apps from Google, and the mean rating for apps from Apple.

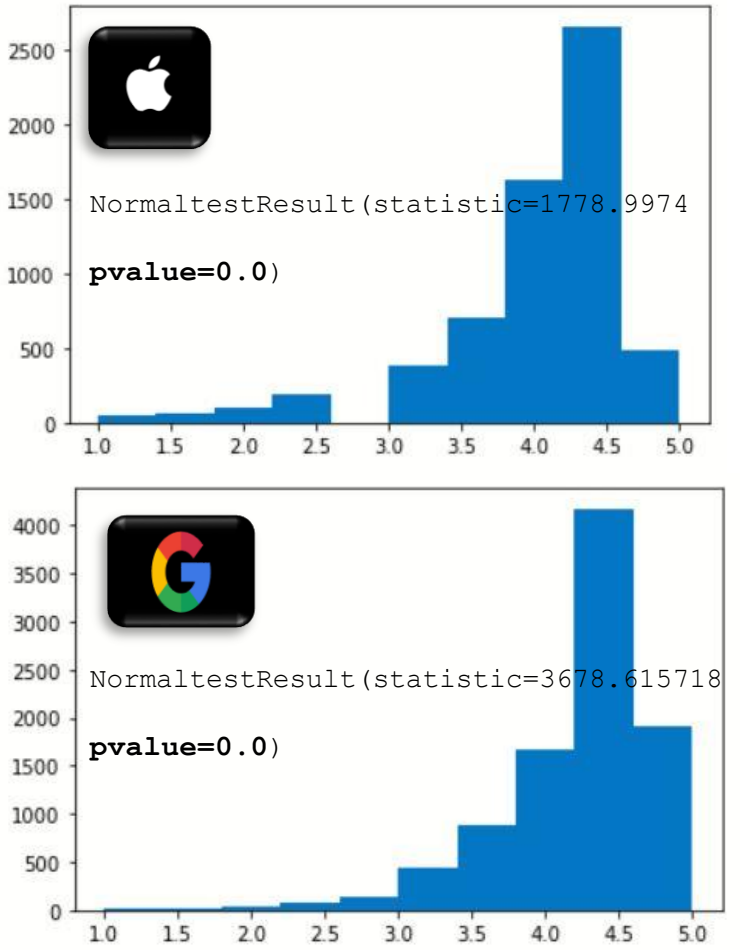


Hypothesis Formulation and Data Distribution

Null Hypothesis:
The observed difference in the mean rating of Apple Store and Google Play apps is due to chance (and thus not due to the platform).

Alternate Hypothesis:
The observed difference in the average ratings of apple and google users is not due to chance (and is due to platform).

We also picked a **significance level** of 0.05 for statistically testing both hypotheses and accepting one of them.



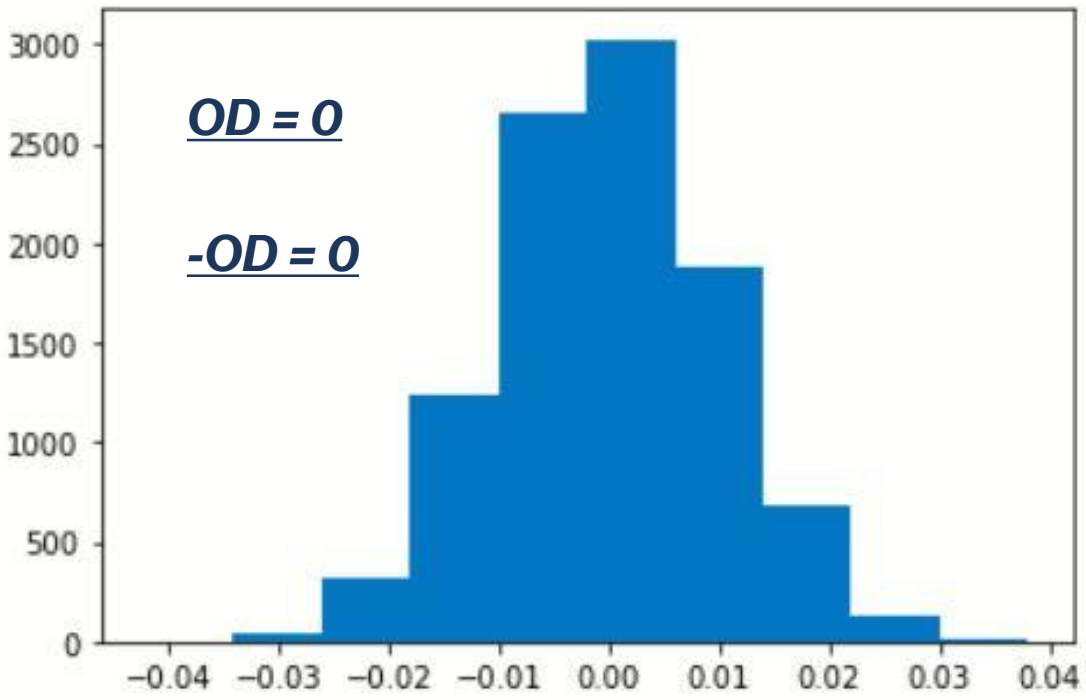
Statistical Testing

Based on the distribution of the Apple and Google sets and the **P value of 0** indicating that the data is not normally distributed, we decided to compute a **Permutation Test** to estimate the statistical significancy of the results where we computed the rating mean differences 10,000 times so we could:

(1) Get the average value

(2) Count how many of the differences in our difference list are at least as extreme as our ***observed difference (OD, -OD).***

Histogram of the Difference List



CONCLUSIONS

Zero **differences** are at least as extreme as our **observed difference**!

The **p-value** of our observed data is **0**.

It doesn't matter which significance level we pick; our observed data is statistically significant, and we **reject the Null Hypothesis**.

We conclude that **platform does impact on ratings**. Specifically, we should advise to integrate **only Google Play** into their operating system interface.