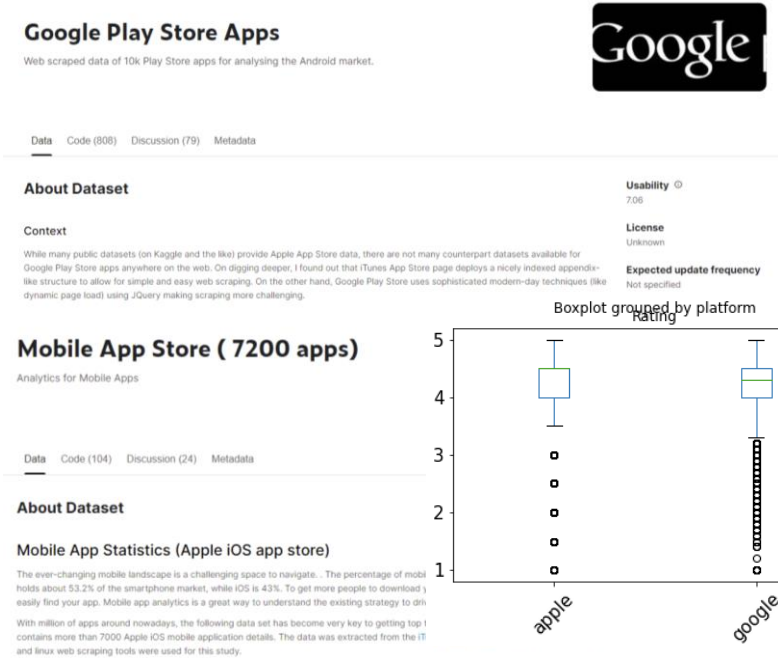
Problem to Solve

Did Apple Store apps receive better reviews than Google Play apps?

Approach

01 – Data Sourcing & Cleaning

- Download the mobile app statistics datasets for Google and Apple Stores from Kaggle
- Examine the datasets to fix errors and remove outliers
  - Remove the records in the datasets where the app price has an error
  - Remove the records in the datasets where the app review count is 0
  - Convert the review counts in the datasets to numerical numbers



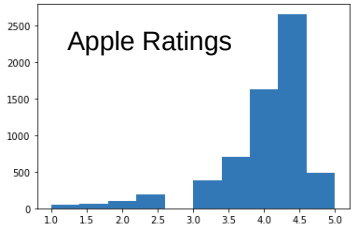
Approach

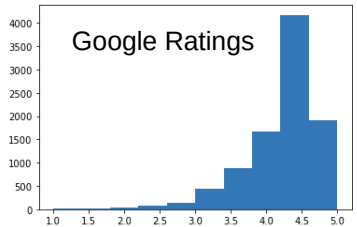
02 – Modeling (Hypothesis Test: Significance Level of 0.05)

- Hypothesis test formulation
  - H_0: 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).
  - H_a: the observed difference in the mean ratings of Apple and Google users is not due to chance (and is actually due to platform)
- Get the distribution of the data
  - Check whether the ratings for Apple and Google apps are in normal distribution
- Create 10,000 permutations
  - Calculate the mean ratings for Google and Apple apps and the difference between these for each one
  - Take the average of all of these differences
  - The p-value of our observed data is the proportion of the data given the Null that's at least as extreme as that observed data
  - If p-value <=0.05, we reject the null hypothesis; otherwise, we accept the null hypothesis.

Results

- Both the ratings of Apple and Google apps are not in normal distribution





- Observed mean difference in the ratings between Apple and Google apps is 0.14

platform	count	mean	std	min	25%	50%	75%	max
apple	6268.0	4.049697	0.726943	1.0	4.0	4.5	4.5	5.0
google	9366.0	4.191757	0.515219	1.0	4.0	4.3	4.5	5.0

- Calculate the number of the differences which are at least as extreme as the observed mean difference
  - There are **zero** for both positive and negative extremes compared to the observed mean difference from the permutation test
- Therefore, the p-value of our observed data is 0 and we reject the null hypothesis

Conclusion

- Our findings suggest that platform does impact on ratings
- Specifically, we suggest to integrate **only Google Play** into their operating system interface