

# A Dive Into the Google Play App Store

Mentorship:  
Dr. Ignacio Becker

Team Outliers:  
Nanda Krishna  
Prayash Panda  
Rajat Gaur  
Shishir Tripathi



# 3 Fundamental Questions!

01

Objective - Collect and analyze the Google Play Store Data to study any relational significance among the apps. Building up a recommendation system based on impactful features.

**What?**

02

To extract insights from the data and to make meaningful recommendations to the users.

**Why?**

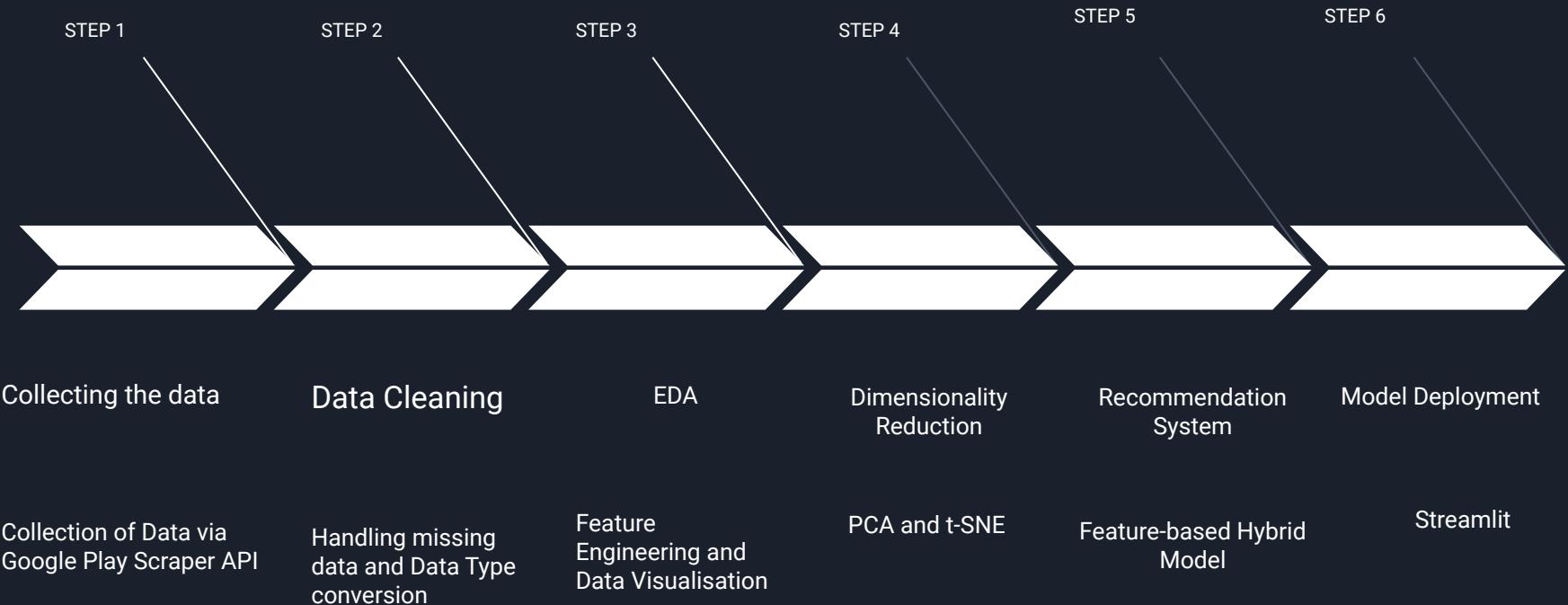
03

Data Collection via Google API and employing EDA techniques to study the relations followed by building up a recommendation system

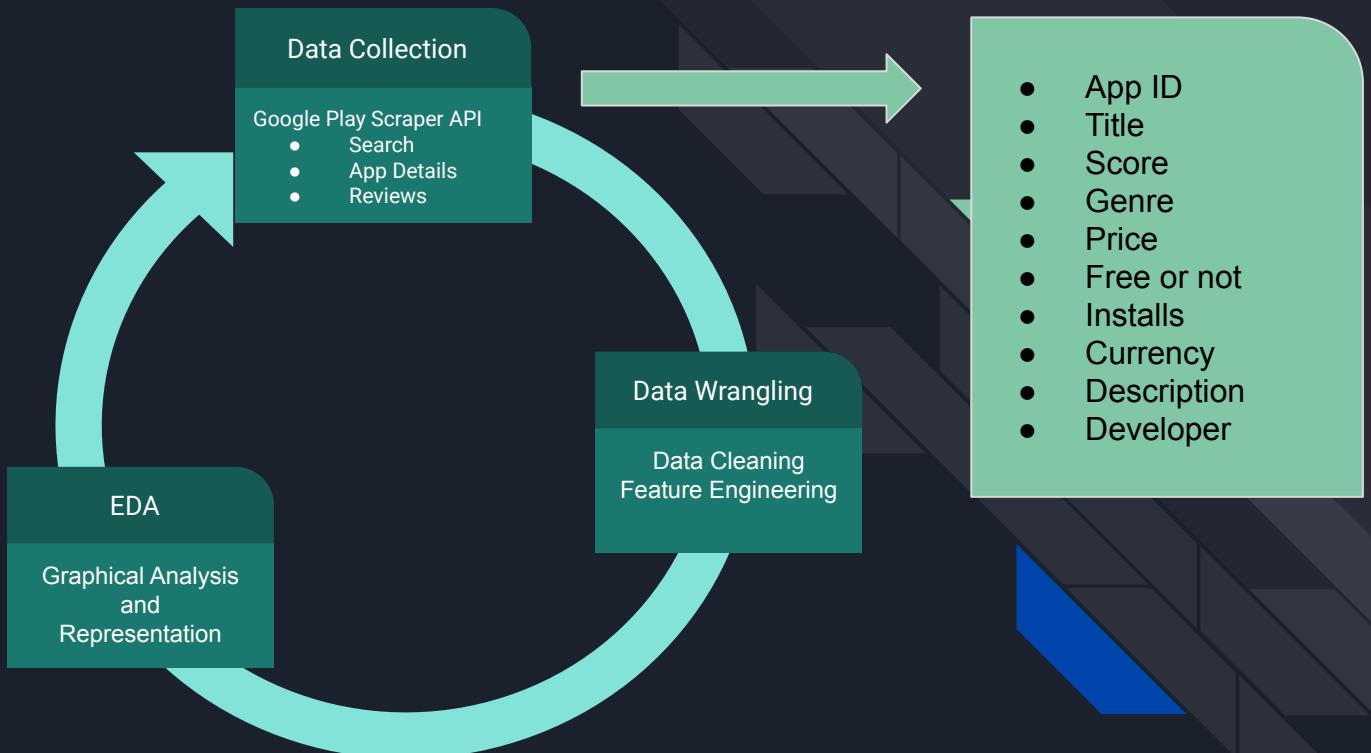
**How?**



# Methodology

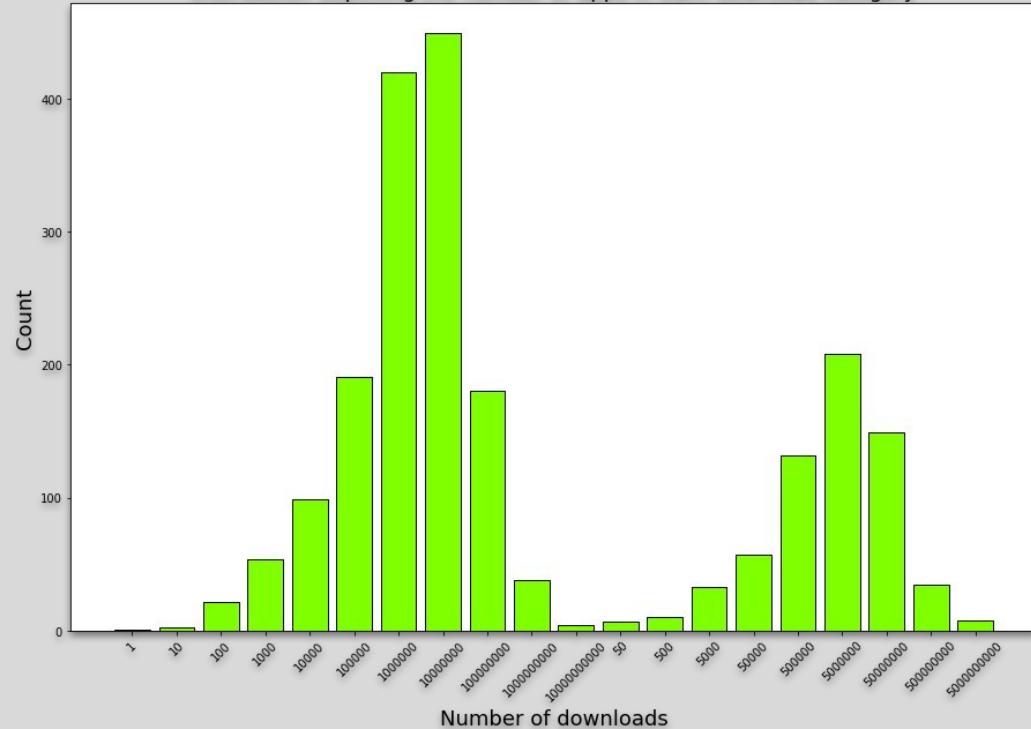


# And, the How Part ?

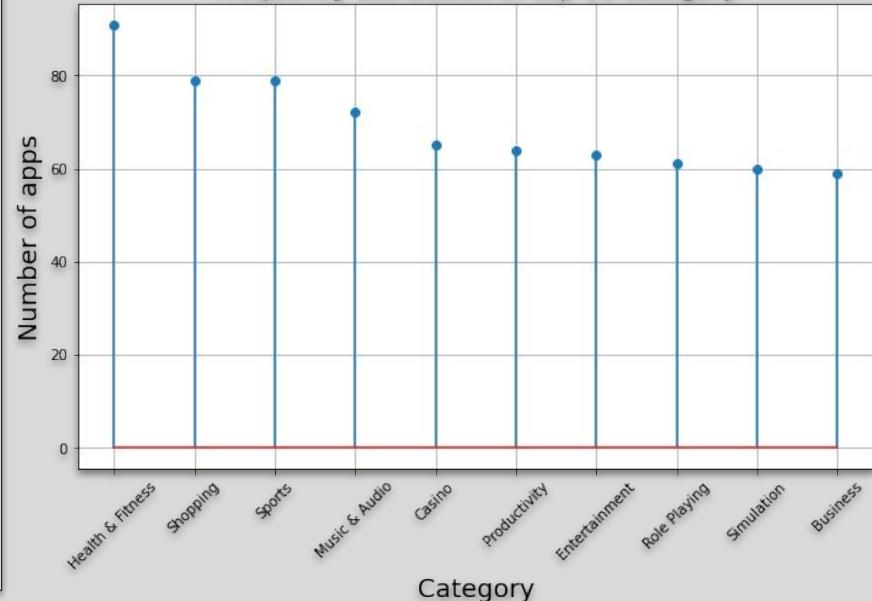


# Analysis of Downloads and Categories

Distribution depicting the number of apps in each download category

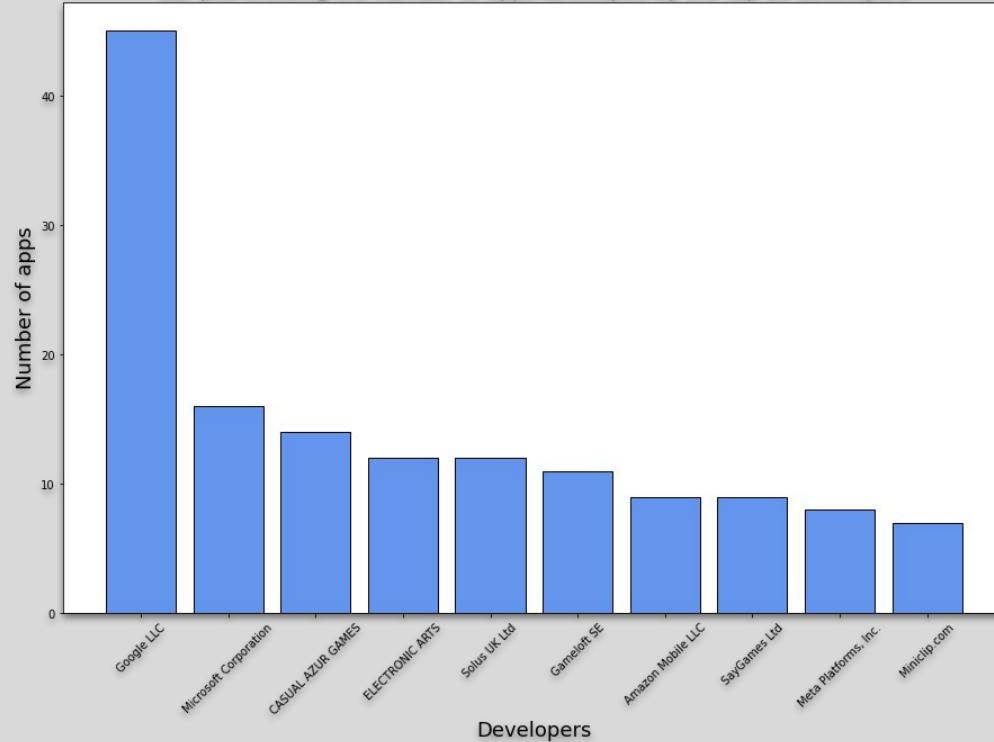


## Frequency distribution of top 10 category

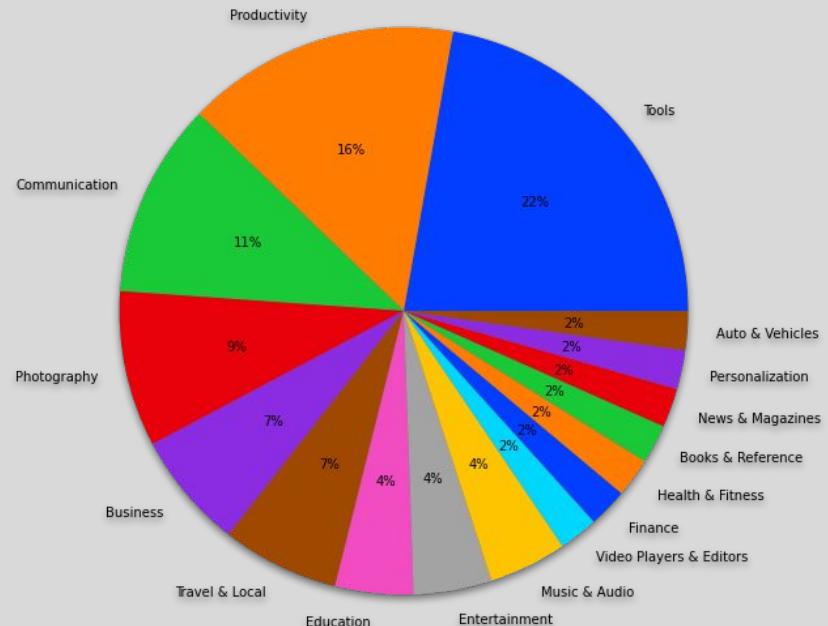


# Exploratory Analysis of Top Developers

Bar plot showing the number of apps developed by the top 10 developers

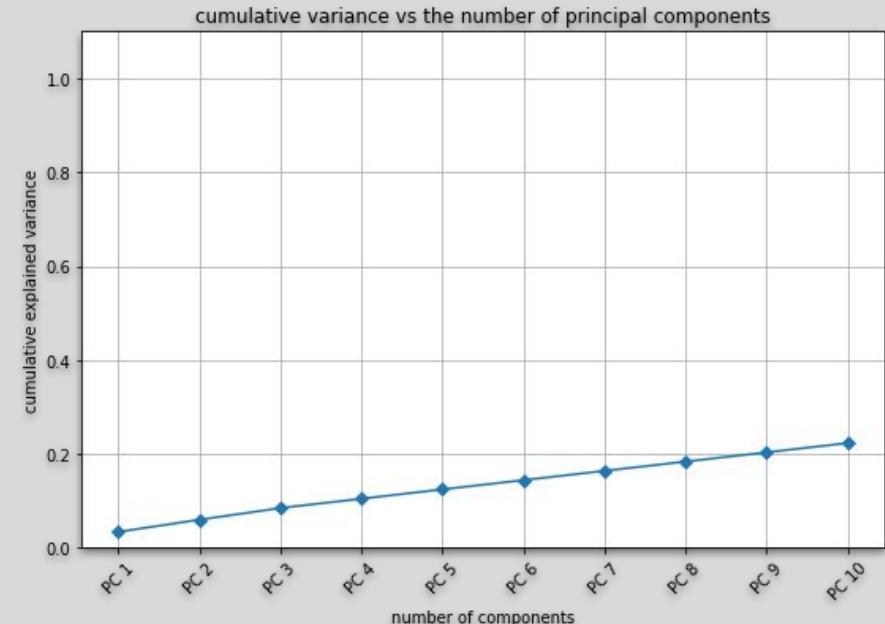
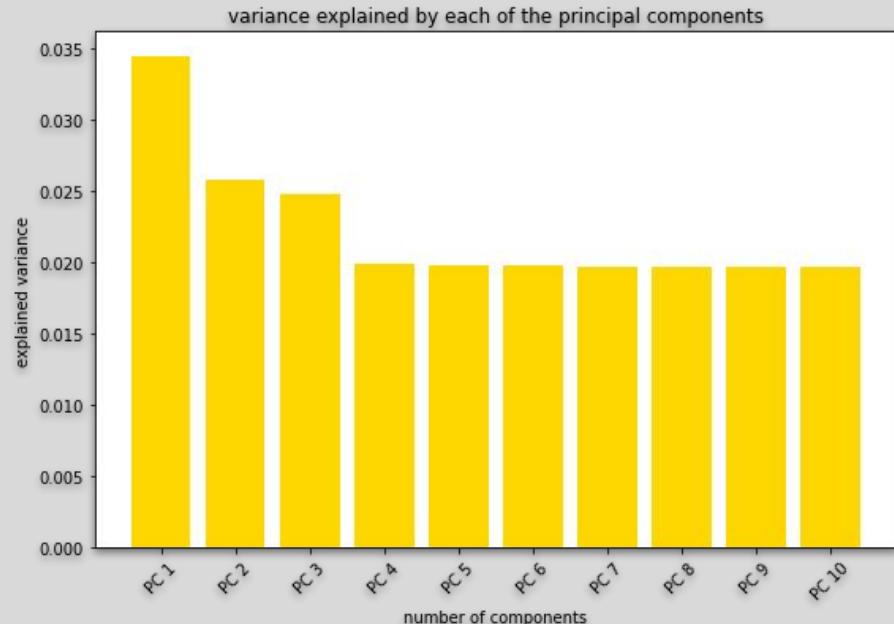


Distribution of each category in which Google LLC launched an app



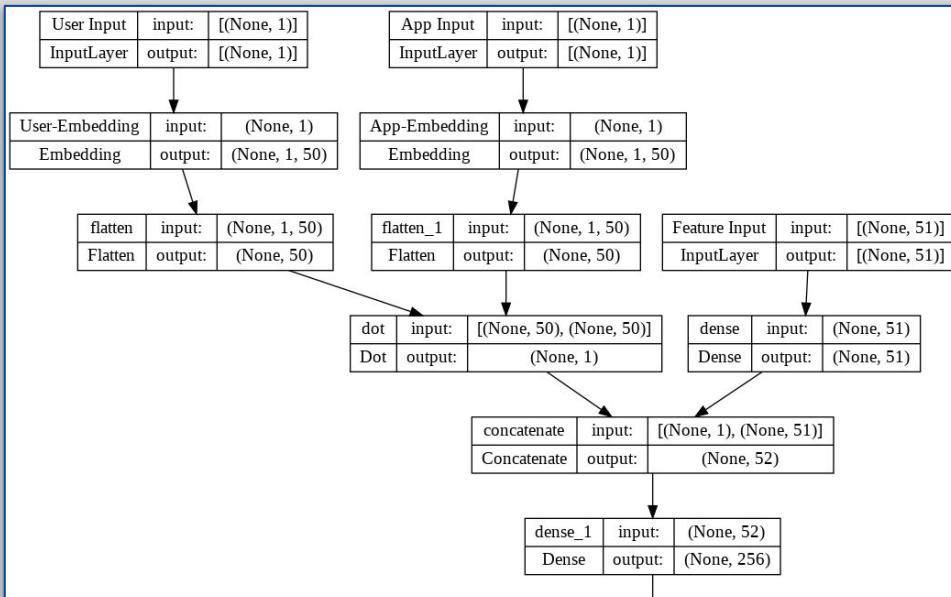
# Principal Component Analysis

Considering first 10 components explaining maximum variance



# Recommendation System

## Feature-Based Hybrid Model



## Model Deployment

## Generating Recommendations

Recommendation For Mikeala Russell	
Apps	Recommendation Score
Black Desert Mobile	75.17 %
LiveBoard: Online Whiteboard	69.93 %
Cashman Casino Las Vegas Slots	68.85 %
BladeBound: RPG Adventure Game	58.95 %
Japan Official Travel App	53.63 %

## A dive into the Google Playstore!!

Get your next recommendation by answering few questions.

What are your favorite genres?

Action Games [ ] Comedy [ ]



Previously played app

Free [ ] Paid [ ]

What do you want the minimum number of installs to be?

- 100K
- 1M
- 10M
- 1B

Select the app from the following

Call of Duty [ ] Clash Royale [ ] Farm Heroes [ ] Candy Crush [ ]



Search for apps [ ]

Click on the button above to get your recommendations.



# Conclusion & Findings

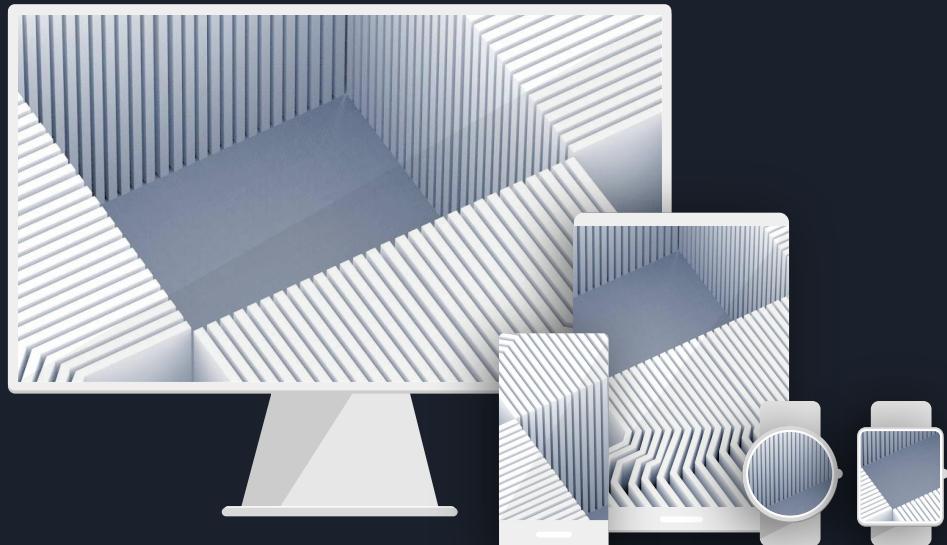
- Exploratory analysis for top developers show that tech giants such as Google, Microsoft and EA have a significant no. of apps compared to others.
- Another interesting observation being a majority of Fitness and Shopping apps, with Sports and Music apps coming just behind them.
- As observed, not much variance has been explained by the first 10 principal components. Only 20% of the variance has been captured by them.
- Incorporating the apps data, we built a feature-based hybrid model to recommend apps to a user based on his preferences selected. You can try this out by accessing the deployed model on Streamlit.



# Future Scope of Work

- We can extend this project to a big data study by involving more number of apps in our dataset and more features as well.
- To extend this study, we can take into account the reviews suggested by users and use Natural Language Processing tools to build a Sentiment Analysis model for the reviews gathered; and further we can include them as well in the analysis and user recommendation system.
- The reviews also have features such as Thumbs up count, No. of replies and reply content, which could also be used to analyse the importance of each review on its own. They also contain some demographic information about the user which can add context.

Thank you!



Any Questions?