Abstract:

There are more than 3.04 million apps found on Google Play Store. The information we will use in this project is scraped from Google Play

The goal of this project is to deliver insights to understand customer demands and helping developers to popularize the product. The dataset is chosen from Kaggle. It is of 10k Play Store apps for analyzing the Android market.

This information is scraped from the Google Play Store. I did the analysis to answer some of questions

- **Highest apps category?**
- **Which Apps type is the highest?**
- **The highest installed app categories is it free or paid?**
- **What is Highest Content Rating?**

DATA:

This information is scraped from the Google Play Store. Its contain 10841 rows × 13 columns consists of :

App: Application name

Category: Category the app belongs to

Rating: Overall user rating of the app (as when scraped)

Reviews: Number of user reviews for the app (as when scraped)

Size: Size of the app (as when scraped)

Installs: Number of user downloads/installs for the app (as when scraped)

Type : Paid or Free

Price: Price of the app (as when scraped)

Content Rating: Age group the app is targeted at - Children / Mature 21+ / Adult

Genres : An app can belong to multiple genres (apart from its main category). For eg, a musical family game will belong to

Last update: Date of the last update

Current Ver : Current Version of the app Android Ver Which Android Ver its compatible with the app

Algorithms:

Data Analyzing:

I analyzed the data using two methods:

Visually: by looking to the dataset and try to understand the data and it's pattern

Programmatically: by using pandas methods like .info() , describe() , head() , tail()

Data Cleaning:

- Check if there are any missing values
- Check Duplicate
- Checking for nulls.
- Drop unnecessary column
- Strip column names

Tools:

- **Technologies**: Python, Jupyter notebook

- **Libraries**: Numpy, Pandas, Matplotlib, Seaborn

Communication:

some of screenshots of my presentation