



DATA WAREHOUSE AND MINING PROJECT

GOOGLE PLAY STORE APP REVIEW ANALYSIS

Presented by:-

- Shreyashi Shankar Sinha - AP21110010273
- S. Tejaswini - AP21110010299
- Mayank Shukla - AP21110010157
- Koduri C S N V Satyanarayana - AP21110011199



Why analyse play store ?



Mobile App Market
is set to grow 20%
by 2023



Android Apps
comprise 90% of the
Mobile App Market



What makes an App
popular? Can we predict
how popular it's going to
be?



What are some
interesting patterns in
user behavior related
to app usage &
feedback?



Introduction

Play store is an Android Market serves as the official app store for certified devices running on the Android Operating system. Developed and Operated by Google, launched on 6th March, 2012.

Approximately 3.48 million apps are in the Play store. Play store apps have their own features such as Ratings, Reviews, Size and more.

From the problem statement given, we should analyze the given database and should come up with the key factors that increased the number of users, long term usage etc., the objective of this project is to deliver insights to understand customer demands better and thus help developers to popularize the product.



Problem Statement

- ❑ Two datasets are provided, one with **basic information (play store Data.csv)**
- ❑ We must examine and evaluate the data in datasets in order to identify the important characteristics that influence app engagement and success.



Dataset Preparation

- **Loading the data sets:** Play store app dataset.
- **Import Libraries:** NumPy, Pandas, Seaborn and Matplotlib
- **Data cleaning:** Null values, Finding and removing Outliers, Removing duplicate data.
- **Data Imputation:** Filling the missing categorical values with mode and numerical values with median. Conversion of price, installs, reviews into numerical values.
- **Exploratory Data Analysis:** Analyzing the data sets to summarize their main characteristics using statistical graphics and data visualizations method.



Attributes in Google Play store Data

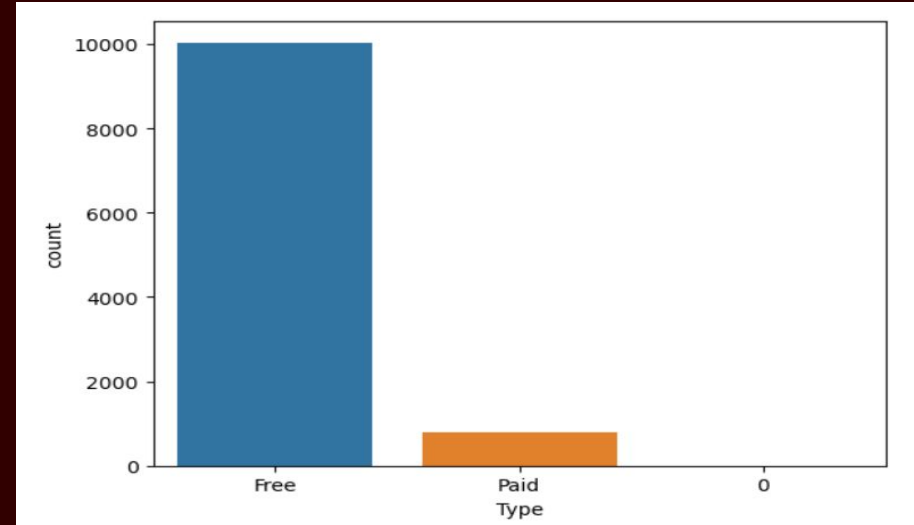
- 1.App :** This column Contains the name of the app for each observation.
- 2.Category :** This column Contains Category to which the app belongs.
- 3.Rating :** This column contains the average rating for the app.
- 4.Reviews :** This column contains the number of reviews that the app has received on the play store.
- 5.Size :** This column contains the amount of memory the app occupies on the device.
- 6.Installs :** This column contains the number of times that the app has been downloaded and installed from the play store.
- 7.Type :** This column contains the information whether the app is free or paid.
- 8.Price:** If the app is a paid app, this column contains the data about its price.
- 9.Content Rating:** This column contains the maturity rating of the app i.e. the age group of the audience for which it is suitable.
- 10.Genres:** This column contains the data about to which genre the app belongs. Genres can be considered as a further division of the group of Category.
- 11.Last Updated:** Contains the date on which the latest update of the app was released.
- 12.Current Version:** Contains information on the current version of the app available on the play store.
- 13.Android Version:** Contains information about the android versions on which the app is supported.



Paid apps v/s Free apps

When comparing the both plots, people are showing more interest on free apps like art & design. And The number of free apps are more than 10000

When it's coming, commercial people are preferring apps like Business, communication and Personalization and the number of paid apps are less than free apps.

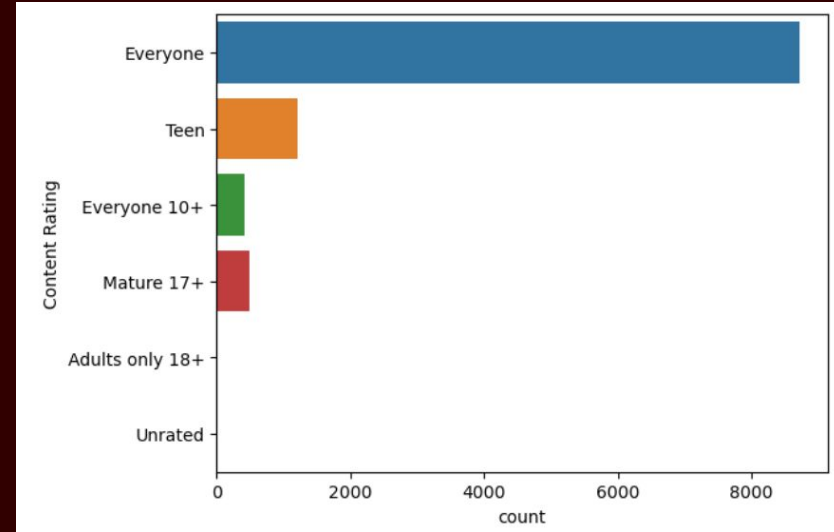




Content Rating

By this analysis we can know that who can use the app most there are so many apps we taken the count of the app (Everyone, Teen, Everyone 10+, Mature 17+, Adults only 18+, Unrated).

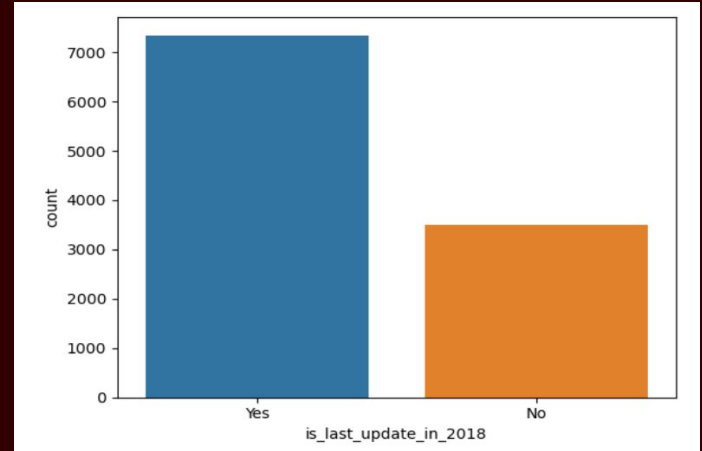
Here we can know how many apps are there for the for the teen, everyone, etc...





Last updated year

We can check the updated year.
Suppose if the year is 2018 it will
check if the last update is in 2018 it
display Yes or else it will display No.



4	Pixel Draw - Number Art Coloring Book	ART_AND_DESIGN	4.3	967.0	2.8	100,000+	Free	0	Everyone	Art & Design,Creativity	June 20, 2018	4.4 and up	Yes
...
36	Sya9a Maroc - FR	FAMILY	4.5	38.0	53.0	5,000+	Free	0	Everyone	Education	July 25, 2017	4.1 and up	No



Conclusions

After analyzing the dataset we have got answers to some of the serious & interesting facts which any of the android users would love to know.

- a) Top 5 Highest rating Apps in google play store.
- b) Free and Paid Apps.
- c) Content rating.
- d) checking last updated year.