

Capstone Project Submission

Instructions:

- i) Please fill in all the required information.
- ii) Avoid grammatical errors.

Team Member's Name, Email and Contribution:

- Contributor Role:

PARUL SRIVASTAVA(shivastava98parul@gmail.com)

1. Data Analyzing
 - Performing Data Wrangling.
 - Data Gathering, processing and transforming.
2. Data Cleaning
 - Finding missing values.
 - Removing null values.
 - Removing replicated data.
3. Drawing Inferences from the data.
4. Data Visualization
 - Implementing bar plot
 - Implementing pie plot
 - Implementing dis plot
 - Implementing joint plot
5. Conclusion

Please paste the GitHub Repo link.

Github Link:- <https://github.com/ParulSrivastava98/PlayStore-App-Data-Analysis>

Google drive Link:-

<https://drive.google.com/drive/folders/11sfFUpZufh2hh2sqKWV/QH691HrxuHxC0?usp=sharing>

Please write a short summary of your Capstone project and its components. Describe the problem statement, your approaches and your conclusions. (200-400 words)

Google play store is widely used by users to download the required applications onto their android devices. Here we are provided with two data sets one is of play store data and user review data to draw inferences from it and find out how developers can work on enhancing it more to capture market.

Firstly, in my capstone project I started with data analyzing by using data wrangling operations to know more about our data. Further I performed various operations on raw data to gather process and transform it and draw insights from it.

Secondly, I did data cleaning by removing any null, nan or replicated values and after that performed mean operations on column to fill any missing values like for content rating, current version, android version etc.

After that I was able to draw basic observations like apps with highest rating, apps with maximum reviews all apps with five-star rating, count of free and paid apps, average overall rating and various other.

Followed by this I performed data visualization to see which category was more preferred by play store users that has helped in play store in engaging more users and acquiring market compared to other applications. Also, I analyzed and visualized various categories of play store using pie chart followed by this I further visualized relationship between installs, size and rating columns using scatter plot and pair plot. And got the inferences that how size of applications affects the rating and installs as smaller size is preferred more for easy download. And how installs will be more for 4- or 5-star rating applications.

By doing this the developers can draw conclusions as what should be an average size of application and where they can focus more to improve play store application further more for better engagement.