

# Capstone Project Submission

## Instructions:

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

Team Member's Name, Email and Contribution:

**Name – Sandeep K**

Email Id – [hrmsk009@gmail.com](mailto:hrmsk009@gmail.com)

**Contribution :**

**Activities performed to complete end-to-end project individually are as below,**

1. **Import data & important Libraries**
2. **Clean Play Store App Database (App wise data and Sentiment Survey Data)**
3. **Give appropriate treatment to missing values and prepare Data Frames to analyze**
4. **Understand data features and explore important data variables**
5. **Analyze and understand Business Models for App making Business**
6. **Brief overview of Data Variables and their Relationship with each other**
7. **Identify Top 10 Performing : Genres | Categories | Apps**
8. **Analyze important factors responsible for App engagement and success as below,**
  - Impact of Ratings on Installs
  - Content Rating wise Apps on play store (Free App & Paid App)
  - Identify Top 3 Content Rating segments contributing most to the success of an App
  - Top 3 Content rating segments giving best Ratings
  - Prepare Annexures of Top 10 Apps preferred by each Top 3 Content Rating Segment
  - Identify most preferred App Size in general and Content Rating segment wise
  - Identify Top 5 Heavy (big size) App Categories
9. **Correlation Analysis of all Quantitative Variables & their impact as below,**
  - Impact of App Updates on Size of the App
  - Impact of Price on Size of the App
  - App price trend analysis
  - Identify Top 5 most accepted Android & Current Versions
  - Analyze impact of Android & Current Versions on Reviews / Size / Price
  - Investigate Rating / Reviews / Size impact on each other
  - Investigate Price / Rating / Last Update impact on each other
10. **Correlation Analysis of Sentiment data variables & their impact as below,**
  - Investigate Distribution of Sentiment Polarity and Subjectivity
  - Content Rating wise sentiment analysis
  - Analyze Sentiment Relation with Size / Install / rating / last Update
  - Identify Top 10 – Most positive sentiment Genres | Categories | Apps
11. **Understand relation of Sentiment Polarity & Subjectivity**
12. **Understand Overall Translated Reviews and Sentiments through - Word Cloud**
13. **Conclusion**
  - Provided qualitative recommendations to App Making Management Team
  - Technical Insights and specifications to App Development & Quality Team
14. **All Visualizations using libraries like Matplotlib, Seaborn, Plotly, Word Cloud**
15. **Presentation, Summary, Technical Document, Colab Notebbok, Video & Github Commits**

Please paste the GitHub Repo link.

Github Link:-

[https://github.com/Sanddeep-K/Capstone\\_1\\_EDA\\_Play\\_Store\\_App\\_Review.git](https://github.com/Sanddeep-K/Capstone_1_EDA_Play_Store_App_Review.git)

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

### Summary –

ABC Co. Ltd. Is planning to launch a successful app and wanted me to analyze Play Store App data and submit a report with an EDA, which will provide them some actionable insights to build a technically compatible and most accepted App.

Here I analyzed and learnt about Mobile App Business, analyzed the given datasets of Play Store Apps and Reviews dataset. Identified important factors , which affects the success of a Mobile App Business, understood their performance impact and KPIs. Investigated various dependent variables and impact on each other, their trends which may affect business significantly

Also investigated the App market, customer segments, their sentiments, behavior pattern and subjectivity of sentiment and concluded this EDA Report with some important qualitative and actionable insights for App management & Technical Team.

### Components –

1. Introduction
2. Objective
3. Problem statement
4. Data Pipeline
5. Describing Datasets
6. Data Cleaning
7. Data Preparation
8. Data Analysis & Visualization
9. Important Q & A (Annexures)
10. Conclusion

### Problem Statement –

I had received a mandate from ABC Co. Ltd. That, company wants to build a successful Mobile App, which should be accepted by a huge customer base of Play Store and create a positive Brand Image among customers. It should be one of the Successful App on Play Store.

Accordingly, worked out deliverables and approach as below,

#### **Deliverables :**

1. Analyze and identify key factors responsible for app engagement and success.
2. Submit recommendations to select and plan development of a Successful App
3. Suggest Actionable Insights to Developers to build a competitive App in best possible manner

#### **Approach –**

1. Import data & important Libraries, Clean the data, perform Data wrangling and prepare copy datasets (Free App data, paid App data, Merged data etc)
2. Understand data features and explore important data variables
3. Analyze and understand Business Models for App making Business
4. Identify Top 10 Performing : Genres | Categories | Apps
5. Identify Top 3 Content rating categories and their contribution in App Success
6. Identify most preferred factors like Size, Android Version, Categories
7. Analyze important factors responsible for App engagement and success
8. Correlation Analysis of all Quantitative Variables & their impact (Rating, Reviews, Updates, Size, Price, Installs, Android Version, Current Version, Sentiment Polarity, Sentiment, Sentiment Subjectivity)
9. Used Word Cloud to demonstrate, how it can be used to manage sentiments of people
10. Conclusion – Qualitative Recommendations to Management Team and actionable technical insights to App Development Team.

## Conclusion –

1. App makers need to finalise a Business Model first. Either Free App model or Paid App Model
2. In Free App category – Most successful categories are **Communication | Travel | Productivity | Social | Family**. Annexure – 1 of our report will help you with all these details
3. In Paid Type of App below are the top Paid App categories **Gaming Apps | Communication | Lifestyle | Sports | Family**. Annexure – 2 of our report will help you with all these details
4. Top 3 Customer segments to focus on for **high success rate – Everyone, Everyone 10+ & Teen** for Free Apps
5. Top 3 Customer segments to focus on for **high success rate – Mature 17+, Everyone and Everyone 10+** for paid Apps
6. Mature 17+ category **alone contributing @ 50%** to the success of paid apps
7. In Free app category people are preferring an app size of **12 – 30 MB** and in paid App category people are preferring an avg. app size between **23 to 45 MB**
8. **Varies with device** app versions are also most accepted versions.
9. App needs to get **updated frequently** with **latest technology**
10. **Periodical updates** affects user experience, reviews and rating of app.
11. In terms of technology people are preferring Android Version apps having version series 4.0 and above. These updated versions are mostly appreciated and liked by people.
12. Certain categories like **Sport / family / Game includes video streaming and storage** hence these apps are most heavier apps in terms of size on the play store
13. As we investigated correlation of various important factors which affects performance or success of apps, we found below observations
  - **Size and Ratings in Free App category are negatively** correlated with each other.
  - **Installs and Reviews are strongly** correlated with each other. Reviews are mostly positive.
  - These positive **reviews are pushing installations** of app
  - After an installation **threshold of 0.5 Billion** rating rate slows down, Need to review customer sentiment at this stage of 0.5 B installs of app.
  - In paid apps also we saw a **positive correlation** of Installs and reviews.
14. Hence for success of an app, reviews are very important, they are driving success of the app. You need to develop a mechanism in app which will frequently ask for reviews and **gain maximum reviews**.
15. In paid apps **Size and Price** are adversely affecting apps success.
16. Make sure to have App **pricing between Rs 10 -20/-** Maximum people are tend to pay this amount on an average.
17. Sentiment **Subjectivity is high within Mature 17+ , Teen and Everyone 10+** categories.  
So while planning of these category apps need to take this feedback redressal mechanism

## Conclusion – Technical Recommendations

### To App Development Team

Technical Recommendations	Details
<b>Technology</b>	Android Version - Series 4.0 and above or Android version - Varies with Device versions
<b>Size</b>	Free App - between 12 - 30 MB   Paid App - 23 - 45 MB   Heavy App - Max. 60 MB for paid Apps and 40 MB for Free Apps
<b>Review Mechanism</b>	A pop-up mechanism for review after every long activity on App Periodical review mechanism for people who gives frequent review Negative feedback tracing and tracking of redressal - Mechanism / Feature
<b>User Interface</b>	Free App - Can be of moderate technology   Paid App - Should be HD technology for better user experience
<b>App Updates</b>	App updation should be done periodically with latest technology
<b>App Experience</b>	Our Annexures consist of some top play store apps, study their technical specification to get an edge over other competitors
<b>Annexures</b>	Annexure 1 - Top 10 Genres   Categories   Apps (Free Apps)
	Annexure 2 - Top 10 Genres   Categories   Apps (Paid Apps)
	Annexure 3 - Top 10 Apps (Everyone Segment - Free apps)
	Annexure 4 - Top 10 Apps (Everyone 10+ Segment - Free apps)
	Annexure 5 - Top 10 Apps (Teen Segment - Free apps)
	Annexure 6 - Top 10 Apps (Mature 17+ Segment - Paid apps)
	Annexure 7 - Top 10 Apps (Everyone 10+ Segment - Paid apps)
	Annexure 8 - Top 10 Apps (Everyone Segment - Paid apps)
	Annexure 9 - Top 10 Apps (Most Positive Sentiments - Free apps)
	Annexure 10 - Top 10 Apps (Most Positive Sentiments - Paid apps)