# Google PlayStore App launch Study

## Complete project Requirement Specification

A leading software development and distribution firm ,that helps individuals /Firms in designing and launching comercially viable mobile apps has conducted a survey of the user download behaviour on the apps across all categories on the google play store. They also have pursued the user review for the contents that the users have downloaded and used over a period of time. The databases for both the survey results have been provided for further study. The Company aims to take crucial decisions before their future launches on playstore based on the results provided by the research analysts. The company is looking for at least the following insights from the data analysis,

1) What is the percentage download in each category on the playstore.
2) How many apps have managed to get the following number of downloads
   a) Between 10,000 and 50,000
   b) Between 50,000 and 150000
   c) Between 150000 and 500000
   d) Between 500000 and 5000000
   e) More than 5000000
3) Which category of apps have managed to get the most,least and an average of 2,50,000 downloads atleast.
4) Which category of apps have managed to get the highest maximum average ratings from the users.Display the result using suaitable visualization tool(s) and also update the data into the database.
5) What is the number of installs for the following app sizes.
   a) Size between 10 and 20 mb
   b) Size between 20 and 30 mb
   c) More than 30 mb
6) For the years 2016,2017,2018 what are the category of apps that have got the most and  the least downloads

7) All those apps , whose android version is not an issue and can work with varying devices ,what is the percentage increase or decrease in the downloads.
8) Amongst sports, entertainment,social media,news,events,travel and games,which is the category of app that is most likely to be downloaded in the coming years, kindly make a prediction and back it with suitable findings.Also update the number of downloads that these categories have received into a database .(Hint:create a new database using WAMP server)

9) All those apps who have managed to get over 1,00,000 downloads, have they managed to get an average rating of 4.1 and above? An we conclude something in co-relation to the number of downloads and the ratings received.

10) Across all the years ,which month has seen the maximum downloads fr each of the category. What is the ratio of downloads for the app that qualifies as teen versus mature17+

11) Which quarter of which year has generated the highest number of install for each app used in the study?

12) Which of all the apps given have managed to generate the most positive and negative sentiments.Also figure out the app which has generated approximately the same ratio for positive and negative sentiments.

13) Study and find out the relation between the Sentiment-polarity and sentiment-subjectivity of all the apps. What is the sentiment subjectivity for a sentiment polarity of 0.4.

14) Generate an interface where the client can see the reviews categorized as positive.negative and neutral ,once they have selected the app from a list of apps available for the study.

15) Is it advisable to launch an app like '10 Best foods for you'? Do the users like these apps?

16) Which month(s) of the year , is the best indicator to the avarage downloads that an app will generate over the entire year?

17) Does the size of the App influence the number of installs  that it gets ? if,yes the trend is positive or negative with the increase in the app size.

18) Provide an interface to add new data to both the datasets provided.The data needs to be added to the excel sheets.

## NOTE : Add at least two Unique additional feature to your design /Analysis that can help the client with a better insight into something that they might have not even thought of.

# Important Instructions

1 ) The Complete  design has to be GUI based.

2) The application has to be Simple to use and  powerful at the same time.

3) The colour combination and layout designing has to be soothing,

4) It should cover all the requirements mentioned in the requirement specifications.

5) it should generate non ambigous results.

6) any additional feature and insght is highly appreciated.

7) Exhaustive use of data visualization tools is expected.

8) If any assumptions are made or formulae are used ,they have to be specified explicitly.

9) if any regression methods are used, their accuracy has to be calculated using some metrics.

10) Validation are absolutely necessary at all possible places in the design.