**Capstone Project Submission**

**Instructions:**

i) Please fill in all the required information.

ii) Avoid grammatical errors.

|  |
| --- |
| **Team Member’s Name, Email and Contribution:**  **Afridi Firoj Pathan**  **Email –** [**pathanafridi212@gmail.com**](mailto:pathanafridi212@gmail.com)  • Data visualization.  • Approach towards plain.  • Data sorting.  • Bar plot and Heat map.  • Data analysis.  • Frame work of project.  • Debug all Errors  • Pi-plot and Histogram plot  • Sample PPT  .• Technical documentation.  • PPT presentation.  • Project summery template. |
| **Problem definition:**  The Play Store apps data has enormous potential to drive app-making businesses to success. Actionable insights can be drawn for developers to work on and capture the Android market. Each app (row) has values for category, rating, size, and more. Another dataset contains customer reviews of the android apps. Objective of the project to Explore and analyze the data to discover key factors responsible for app engagement and success.  **EDA on given Data set:**  **There are two dataset:**   1. **Play Store Data**(App, Category, Rating, Review, Size, Install, Type, current rating ,genres , Last update, Current Var ,Android Var) 2. **User Review Data**(App, Sentiment ,Sentiment Polarity, Sentiment Subjectivity) Digging into data we understand that   • There are 13 columns of properties with 10841 rows of data.  • Column 'Reviews', 'Size', 'Installs' are in the type of 'object'  • Values of column 'Size' are strings representing size in 'M' as Megabytes, 'k' as kilobytes and also 'Varies with devices'  • Values of column 'Installs' are strings representing install amount with symbols such as ',' and '+'  **Conclusion**   * Maximum number of sentiment subjectivity lies between 0.4 to 0.7 From this we can conclude that maximum number of users give reviews to the applications, according to their experience. * Install and rating highly positive correlated data * Most of the rating is in between 4.0 to 4.5 * In merge tow data frame We can determine from above that each column are Positively Correlated. * Majority of the genres have high positive sentiments, when compared to neutral and negative ones. * Size of maximum number of applications present in the dataset * Size is impact the number of installations |
| **Please paste the GitHub Repo link.** |
| Github Link:- <https://github.com/afridipathan/Play-Store-App-Review-Analysis.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)** |
|  |