INTERNSHIP REPORT

PROJECT NAME: Build Real-Time Google Play Store Data Analytics using Python

🏛️ ORGANIZATION DETAILS:

- Internship program by : NullClass EdTech Private Limited
- Internship Duration: 27th May 2025 – 27th June 2025 (1 Month)
- Position: Data Analyst Intern

📌 OBJECTIVE:

To design and implement a real-time, interactive data analytics dashboard using Python that analyzes and visualizes trends from the Google Play Store app and user review datasets.

🗂️ DATASETS USED:

- Googleplaystore.csv: Contains app metadata such as ratings, size, installs, last update, etc.
- Googleplaystore_user_reviews.csv: Contains user sentiment data including sentiment polarity and subjectivity.

🎯 TASKS COMPLETED:

✅ **Task 1: Sentiment Distribution Visualization**

- Created a stacked bar chart segmented by app rating groups (1-2, 3-4, 4-5 stars).
- Filtered for top 5 categories and apps with more than 1,000 reviews.
- Visualized positive, neutral, and negative sentiment distribution.

✅ **Task 2: Install-Based Category Analysis (Time-Gated: 3 PM – 5 PM IST)**

- Built a grouped bar chart comparing:
  - Average Rating
  - Total Reviews
  - Average App Size
- Filtered categories:
  - Rating ≤ 4.0
  - Size ≤ 10MB
  - Last Updated in January
- Conditional display based on current IST time.

✅ **Task 3: Bubble Chart of App Size vs Rating (Time-Gated: 5 PM – 7 PM IST)**

- Plotted a bubble chart:
  X-axis: App Size (MB)
  Y-axis: Rating
  Bubble size: Installs
- Applied filters:
  Rating > 3.5
  Reviews > 500
  Sentiment Subjectivity > 0.5
  Installs > 50,000
  App name does not contain the letter "S"
  Specific categories only
- Translated categories:
  Beauty → सुंदरता (Hindi)
  Business → வணிகம் (Tamil)
  Dating → Verabredung (German)
  Highlighted the Game category in pink.

🛠️ TECH STACK USED:

- Language: Python
- IDE: Visual Studio/Jupyter notebook
- Libraries: pandas, numpy, plotly, ipywidgets, datetime, pytz

📊 OUTPUT:

- A fully functional HTML dashboard that dynamically updates visualizations based on the time of day (IST).
- Used plotly.to_html() and time filtering logic for real-time responsiveness.
- Dashboard launch via a button using ipywidgets.

🌟 LEARNING OUTCOMES:

- Hands-on experience with data preprocessing, feature engineering, and time-aware visualization logic.
- Learned to work with real-world datasets and build interactive visual analytics using Plotly.
- Understood real-time dashboard deployment techniques and user-centric design logic.

🙋 CONTACT:

Monisha M.
✉️ **Email: monishamani2908@gmail.com**
🌐 **GitHub: Monisha3029**