INTERNSHIP REPORT

INTRODUCTION:

During my internship, I had the opportunity to work on data analytics project, focusing on visualizing Google Play Store data. The main objective was to analyze app installations, user engagement and trends through various data visualization techniques. This experience enhanced by technical skills and provided insights into real world data analysis.

BACKGROUND:

The internship was centered around data analytics and visualization, requiring professionals in Python, Pandas and Plotly. The primary dataset contained information about various apps, including installs, reviews, ratings and other attributes. The goal was to extract meaningful insights and present them through interactive charts and graphs.

LEARNING OBJECTIVES:

- Develop a deeper understanding of data visualization techniques.
- Enhanced proficiency in Python and data manipulation libraries.
- Gain experience in using Plotly to create interactive visualization.
- Improve problem solving skills by tackling real world data challenges.
- Learn to filter and present data effectively for business insights.

ACTIVITIES AND TASKS:

Throughout the internship I worked on multiple data visualization tasks with included:

1. **World Cloud Generation**: Created a word cloud to visualize the most frequently used words in app names and descriptions.
2. **Dual-Axis Chart**: Developed a dual-axis chart to compare installs and reviews overtime, highlighting key trends.
3. **Choropleth Map**: design the Choropleth map to represent global installs by category, applying specific filters and conditions.
4. **Bubble Chart**: Created a bubble chart to compare installs and reviews across categories, emphasizing key data points.
5. **Time-Series Chart**: Plotted a time-series chart to showcase trends in app installs highlighting significant growth periods.

SKILLS AND COMPETENCIES:

- **Technical Skills**: Proficiency in Python, Pandas and Plotly.
- **Data Analysis**: Cleaning, filtering, and processing large datasets to extract meaningful insights.
- **Visualization Techniques**: Creating engaging and informative charge maps and graphs.
- **Problem-Solving**: Entering data related challenges and implementing solutions effectively.
- **Time Management**: Completing multiple tasks within tight deadlines.

FEEDBACK AND EVIDENCE:

The internship provided valuable feedback on data visualization techniques and best practices. Through reviews and discussions with mentors, I refined my approach to presenting data efficiently. The completed visualizations were well-received, demonstrating clarity and effectiveness in conveying insights.

CHALLENGES AND SOLUTIONS:

1. **Missing Data in the Dataset**: Some required columns were absent, such as country codes for the Choropleth map.
    - *Solution*: Created an empty map while ensuring the data filters were correctly applied.
2. **Annotations annotation errors in Time-Series Chart**: incorrect positioning of labels caused visualization issues.
    - *Solution*: Adjusted annotation placement to comply with Pilates requirement.

OUTCOMES AND IMPACT:

The internship allowed me to successfully complete all assigned tasks demonstrating my ability to handle data visualization challenges. The skills acquired during this experience will be instrumental in future data analytics roles. Additionally, the project outcomes provided valuable insights into app trends, which could help stakeholders make informed decisions.

CONCLUSION:

This internship has been an enriching experience, significantly enhancing my technical and analytical skills. Working with real-world data and implementing advanced visualization techniques has strengthened my ability to derive insights from complex datasets. The challenges faced and the solutions implemented have contributed to my growth as a data analyst, preparing me for the future opportunities in the field.