📊 Google Play Store Data Analysis Project

🔹 Overview

This project focuses on analyzing the Google Play Store dataset to understand app trends, user behavior, and performance across different categories. Using Python and popular data analysis libraries, we performed data cleaning, exploratory data analysis (EDA), and data visualization to extract meaningful insights.

🔹 Tools & Technologies

- **Python**
- **Pandas** – Data manipulation
- **NumPy** – Numerical operations
- **Matplotlib & Seaborn** – Data visualization
- **Jupyter Notebook**

🔹 Key Tasks Performed

1. **Data Loading**
    - Imported the Google Play Store dataset using Pandas.
    - Inspected the structure and basic statistics of the data.
2. **Data Cleaning**
    - Handled missing values and removed inconsistencies.
    - Converted columns like `Reviews`, `Installs`, and `Price` into proper numeric formats.
    - Removed duplicate and invalid entries.
3. **Exploratory Data Analysis (EDA)**
    - Analyzed app distribution across categories.
    - Studied the relationship between ratings, reviews, and installs.

- Compared free vs paid apps.

4. **Data Visualization**
    - Created bar charts, count plots, and distribution plots.
    - Visualized trends in installs, ratings, and categories.

🔹 Insights & Results

- Most apps on the Play Store are **Free apps**.
- **Game** and **Family** categories have the highest number of apps.
- Apps with higher installs generally have **better ratings**.
- Paid apps are fewer but often show higher quality indicators.
- There is a strong relationship between user reviews and number of installs.

🔹 Conclusion

This project demonstrates how data analysis can be used to extract valuable insights from real-world datasets. It highlights trends in app categories, user preferences, and performance metrics, which can help developers and businesses make data-driven decisions.

🔹 Author

Md Sakib Azam
Data Analyst | *Python* | *SQL* | *Power BI* | *Machine Learning*