Google Play Store Analysis Design Document

Overview

The goal of this project is to analyse trends and patterns in the Google Play Store dataset to address critical business questions regarding app popularity, pricing strategies, and user satisfaction. By leveraging data-driven insights, this project aims to provide actionable recommendations for app developers and businesses.

Objective

To uncover insights that help:
1. Identify app categories driving the highest user engagement and installs.
2. Examine the impact of pricing strategies on app downloads.
3. Highlight factors contributing to higher ratings and better user satisfaction.

Scope

The analysis focuses on:
- App categories, ratings, pricing, install counts, and user reviews.
- Relationships between numerical variables to identify correlations.
- Trends influencing app performance and user engagement.

Methodology

1. Data Wrangling and Cleaning

- Remove irrelevant or incomplete records (e.g., missing values for key columns like `Category`, `Rating`, `Price`, or `Installs`).
- Standardize data formats (e.g., converting prices and installs into numeric fields).
- Identify and exclude outliers to ensure accurate analysis.

2. Exploratory Data Analysis (EDA)

Visualize and summarize data to identify patterns and trends:
- Examine app rating distributions to understand user feedback.

- Calculate total installs across app categories to determine popularity.
- Review pricing trends for insights into user preferences for free vs. paid apps.

3. Correlation Analysis

Investigate relationships between numerical variables (e.g., `Rating`, `Price`, `Installs`) using correlation matrices and visualizations. This helps understand the impact of pricing and ratings on app engagement.

4. Advanced Visualization

Create meaningful visuals to address specific business questions:
- Bar charts to highlight popular app categories by installs.
- Scatter plots to explore the relationship between price and installs.
- Heat maps to show correlations between metrics like ratings, installs, and prices.

5. Insights and Recommendations

Analyse findings and present actionable recommendations:
- Target high-install app categories for development.
- Optimize pricing strategies to balance user acquisition and revenue generation.
- Focus on improving user experience in lower-rated categories for better engagement.

Deliverables

1. **Key Findings:** Insights derived from the data, backed by visuals and statistics.
2. **Recommendations:** Actionable steps for stakeholders to improve app performance.
3. **Documentation:** A comprehensive report summarizing the analysis and outcomes.
4. **Visuals:** Clear and impactful charts and graphs to present insights.

Technical Requirements

- **Libraries/Tools:** Python (Pandas, Seaborn, Matplotlib), Jupyter Notebook.
- **Dataset:** Google Play Store dataset.
- **Output:** PDF report or hosted GitHub repository with project files and documentation.

Benefits

This analysis offers businesses a roadmap for optimizing app development and marketing strategies. Developers can use insights to focus on high-demand categories and pricing models while enhancing user satisfaction.