#### **App Insights Unlocked: A Data Analytics Challenge**

***Background:***

*Our company specializes in developing innovative mobile applications and is committed to enhancing user experience and engagement. To stay competitive in the evolving app market, we have acquired a dataset from the* ***Google Play Store*** *containing crucial information about various apps, including ratings, reviews, category, size, price, and installs. By analyzing this dataset, we aim to uncover valuable insights into app performance, user preferences, and market trends. Understanding these factors will enable us to optimize app features, improve customer satisfaction, and drive higher downloads and revenue, ultimately strengthening our position in the mobile application industry.*

***Problem Statement:***

*The main task is to analyze this dataset to uncover patterns and trends that can guide the company's decision-making process. The analysis should focus on understanding what factors contribute to high app ratings, identifying the most popular app categories, and exploring the impact of app size and price on user reviews and installs.*

***Project Scope:***

*The Google Play Store Data Analysis project focuses on examining app performance, user engagement, and market trends using statistical and visualization techniques. The project scope defines the boundaries, data sources, methodologies, and tools used for analysis. The following aspects are included in the project:*

***1.Data Collection & Preprocessing:***

* *Dataset sourced from* ***Kaggle****.*
* *Data cleaning, handling missing values, removing duplicates, and outlier detection.*

***2.Exploratory Data Analysis (EDA):***

* *Descriptive statistics to summarize key features.*
* *Distribution analysis of app categories, ratings, reviews, and installs.*

***3.Statistical Analysis:***

* ***Correlation analysis*** *to examine relationships between ratings, reviews, installs, and other numerical features.*

***4****.****Data Visualization & Reporting:***

* ***Power BI dashboards*** *for interactive data visualization.*
* *Implementation of* ***tooltips and filters*** *for enhanced insights.*
* *Category-wise trends and pricing analysis.*

***5.Insights & Recommendations:***

* *Identifying high-performing app categories.*
* *Analyzing key factors affecting user ratings and downloads.*
* *Providing data-driven recommendations for app developers, marketers, and product managers.*

***Stakeholders:***

***Internal Stakeholders:***

* ***App Developer******–*** *Developers use insights on app performance, user ratings, and category trends to enhance features, optimize app size, and improve user experience.*
* ***Product Managers–*** *They leverage data to shape app development strategies, prioritize updates, and ensure alignment with market demands.*
* ***Marketing Team–*** *Marketing professionals analyze trends in app downloads, user engagement, and pricing strategies to design data-driven campaigns that maximize visibility and retention.*
* ***Senior Management–*** *Decision-makers use this analysis to assess business performance, allocate resources effectively, and define long-term growth strategies.*

***External Stakeholders:***

* ***App Users–*** *End-users benefit from optimized apps that cater to their needs, ensuring a seamless experience based on ratings, reviews, and feature enhancements.*
* ***Advertisers–*** *Brands and ad agencies rely on insights regarding user engagement and app popularity to place targeted advertisements effectively.*
* ***Partners (e.g., other tech companies or platforms)–*** *Third-party service providers, such as payment platforms or cloud service providers, can adjust their strategies based on app growth trends and user preferences.*

### ***Data Requirements:***

*The data provided includes the following columns:*

* *App: Name of the app*
* *Category: Category of the app*
* *Rating: Average user rating of the app*
* *Reviews: Number of user reviews*
* *Size: Size of the app*
* *Installs: Number of installs*
* *Type: Whether the app is free or paid*
* *Price: Price of the app*
* *Content Rating: Age group for which the app is appropriate*
* *Genres: Genres of the app*
* *Last Updated: Date when the app was last updated*
* *Current Ver: Current version of the app*
* *Android Ver: Minimum Android version required*

***Metric Development:***

* *Total Installs: 141bn*
* *Total Apps: 6273*
* *Average Size: 24.73*
* *Average Rating: 4.24*
* *Total Reviews: 4bn*

***Tools & Technologies Used:***

***1.Python - Data Processing & Analysis:***

* ***Data Cleaning & Transformation:*** *Used* ***Pandas and NumPy*** *to handle missing values, filter data, and perform computations.*
* ***Data Visualization:*** *Leveraged* ***Matplotlib and seaborn*** *to create bar charts, scatter plots, and heatmaps for trend analysis.*
* ***Correlation Analysis:*** *Used Pandas and NumPy to identify relationships between variables such as ratings, downloads, and reviews.*

#### ***2.Power BI – Interactive Data Visualization:***

* ***Dynamic Dashboards:*** *Created interactive reports for analyzing app categories, ratings, and pricing strategies.*
* ***DAX (Data Analysis Expressions):*** *Used for advanced calculations and custom measures****.***
* ***Tooltips for Insights:*** *Implemented tooltips to provide additional context when hovering over charts, enhancing data interpretation****.***
* ***Drill-Through & Filters:*** *Enabled users to explore specific app segments based on region, category, and pricing.*

***3.Google Colab:***

* *Colab is the platform used to run Python code for data analysis. It provided a cloud-based platform with built-in support for* ***Pandas, NumPy, and matplotlib****, making data processing and correlation analysis easier. Googlecolab also allowed* ***real-time collaboration****, Google Drive access, and smooth execution without needing a powerful local setup.*

***Methodology:***

***1.Data Collection & Preprocessing:***

* ***Dataset Source:*** *The data was obtained from* ***Kaggle****, a reliable source for structured datasets.*
* *Imported the dataset into* ***Google Colab*** *for analysis.*
* *Cleaned data by handling missing values, duplicates, and inconsistencies****.***

***2. Exploratory Data Analysis (EDA):***

* *Analyzed key metrics like app ratings, downloads, pricing, and category distribution.*
* *Used* ***matplotlib*** *scatter plot to examine correlation between features such as ratings, reviews, and installs.*

***3.Statistical Analysis:***

* *Applied* ***correlation analysis*** *in Python to identify relationships between app features.*
* *Evaluated how factors like app size, price, and reviews impact user ratings and downloads.*

***4.Data Visualization with Power BI:***

* *Created* ***interactive dashboards*** *to display trends and insights.*
* *Used* ***tooltips*** *to provide additional context when hovering over visuals.*
* *Applied* ***filters***  *to explore app performance by category, price, and content rating.*

***5.Insights & Recommendations:***

* *Identified high-performing app categories and pricing strategies.*
* *Recommended optimizing app size and user ratings for better engagement.*

***Technical Process:***

***1.Data Acquisition:***

* ***Source:*** *The dataset was obtained from Kaggle and uploaded to Google Colab.*
* ***Format:*** *CSV file containing app details such as category, installs, ratings, price, and reviews.*

***2. Data cleaning & Preprocessing:***

* ***Loaded*** *the csv dataset by importing*  ***pandas,numpy,matplotlib and seaborn*** *in Python.*
* ***Checked for missing values*** *using isnull() function and handled them by replacing using interpolate() function [Interpolate() is a common technique to estimate missing values based on the surrounding data]**or by imputation.*
* ***Remove duplicates*** *using drop\_duplicates() function to maintain data integrity.*
* ***Converted data types*** *where necessary (e.g., ”Last Updated” column into datetime format, “Reviews” column into integer data type, ”Price” column into float data type and numerical conversion of “Installs” and “Price”).*
* **Standardize Text:** *Ensured consistency in text fields (e.g., standardized Category and Genres).*
* ***Handled outliers*** *in numerical columns (e.g., installs, reviews, and prices).*

***3. Exploratory Data Analysis (EDA):***

* *Used* ***Pandas and NumPy*** *for basic statistical summaries like Describe(), info(),corr(),etc.*
* *Visualized category distributions, app popularity, and pricing trends using* ***Matplotlib & Seaborn****.*
* *Performed* ***correlation analysis*** *to understand relationships between ratings, reviews, and installs.*

***4. Data Visualization in Power BI:***

* *Imported the cleaned dataset into* ***Power BI*** *for dashboard creation.*
* *Created* ***interactive visualizations*** *for trends in app categories, ratings, and pricing.*
* *Implemented* ***tooltips*** *to display additional insights when hovering over charts.*
* *Used* ***filters*** *for deeper data exploration.*

***5.Insights & Reporting:***

* *Extracted meaningful patterns and trends from the analysis.*
* *Summarized insights and recommendations for stakeholders, including app developers, product managers, and marketers.*

***Recommended Analysis:***

**1. *What is the average rating of apps in the dataset?***

* ***Understanding the Analysis:****The goal is to determine the average rating of apps to assess overall user satisfaction and app quality. A higher rating indicates better user experience, while lower ratings suggest areas for improvement.*

#### ***Process of analysis:***

#### ***1. Data Cleaning:*** *Remove missing or invalid ratings (outside 1-5 range).*

#### ***2.Computation:*** *Calculate the mean of the "Rating" column.*

#### **Impact of the Solution**

#### **User Insights:** Identifies overall app quality and satisfaction.

#### **Market Trends:** Helps developers and businesses improve app offerings.

#### **Investment Decisions:** Guides stakeholders in evaluating app success.

### **2. What is the distribution of app sizes?**

#### **Understanding the Analysis:** This analysis examines **how app sizes vary**, helping identify trends in storage needs, user preferences, and app optimization.

* **Process of Analysis:**
* **Clean Data:** *Convert app sizes (e.g., "10M", "200k", "1.5k") into MB.*
* **Analyze & Visualize:** *Compute stats (average, min, max) and create a* ***histogram*** *to see size distribution.*

#### **Impact of the Solution:**

* **Users:** *Know which apps take more storage.*
* **Developers:** *Optimize app sizes for better performance.*
* **Market Trends:** *Identify whether smaller or larger apps dominate.*

**3.How many free vs paid apps are there?**

#### ***Understanding the Analysis:*** *This analysis compares the number of free and paid apps, helping understand pricing trends and user preferences.*

#### ***Process of Analysis:***

#### **Data Cleaning:** *Check and filter the "Type" column (Free/Paid).*

#### **Counting Apps:** *Count the number of free and paid apps***.**

#### **Visualization:** *Use a bar chart for comparison.*

#### ***Impact of the Solution:***

#### **Users:** *Helps understand how many apps are free.*

#### **Developers:** *Identifies market trends for pricing strategies.*

#### **Business Insights:** *Assists in monetization and investment decisions*

**4. What is the most common content rating for apps?**

#### ***Understanding the Analysis:*** *This analysis identifies the most frequent content rating (e.g., Everyone, Teen, Mature) to understand which age groups apps are designed for.*

#### **Process of Analysis:**

* ***Data Cleaning:*** *Check the "Content Rating" column.*
* ***Counting Ratings:*** *Find the most frequent rating using value counts.*
* **Visualization:** *Use a bar chart to show distribution.*

#### ***Impact of the Solution:***

* ***Users:*** *Helps find apps suitable for their age group.*
* ***Developers:*** *Guides app classification for target audiences.*
* ***Market Trends:*** *Shows which content category dominates the Play Store*

***5.*What are the top 5 most installed apps?**

#### ***Understanding the Analysis:****This analysis identifies the top 5 apps with the highest installs, helping understand user preferences and app popularity.*

#### ***Process of Analysis:***

#### **Data Cleaning:** *Convert "Installs" column to numeric (remove symbols like +).*

#### **Sorting:** *Sort apps by the highest number of installs.*

#### **Top 5 Extraction:** *Select the top 5 apps.*

#### ***Impact of the Solution:***

* ***Users:*** *Discover the most popular apps.*
* ***Developers:*** *Analyze trends and competition in top apps.*
* ***Business Insights:*** *Helps investors and businesses target high-demand app categories.*

### **6.How many apps have a rating of 4.0 and above?**

#### ***Understanding the Analysis:****This analysis finds the* ***number of apps with a rating of 4.0 or higher****, helping assess how many apps are well-rated by users.*

#### ***Process of Analysis:***

* ***Data Cleaning:*** *Remove missing or invalid ratings.*
* ***Filtering:*** *Count apps where* ***Rating ≥ 4.0****.*
* ***Impact of the Solution:***
* ***Users:*** *Find quality apps easily.*
* ***Developers:*** *Understand the percentage of highly-rated apps.*
* ***Business Insights:*** *Helps in analyzing user satisfaction and competition.*

### **7. What is the average number of reviews for free vs paid apps?**

#### ***Understanding the Analysis:****This analysis compares the* ***average number of reviews for free and paid apps****, helping understand user engagement levels based on pricing models.*

#### ***Process of Analysis:***

* ***Data Cleaning:*** *Convert "Reviews" column to numeric and check "Type" (Free/Paid).*
* ***Grouping & Aggregation:*** *Compute the average reviews by finding mean for each type.*

#### ***Impact of the Solution:***

* ***Users:*** *Understand which apps get more engagement.*
* ***Developers:*** *Analyze user behavior based on pricing strategy.*
* ***Business Insights:*** *Helps optimize app monetization strategies.*

***8.*What is the average app size for each category?**

#### ***Understanding the Analysis:****This analysis finds the* ***average app size per category****, helping understand storage trends across different app types.*

#### ***Process of Analysis:***

* ***Data Cleaning:*** *Convert "Size" to MB (remove "M", "K", "G") and handle missing values.*
* ***Grouping & Aggregation:*** *Compute the average size for each category.*

#### ***Impact of the Solution:***

* **Users:** *Helps manage storage needs based on app categories.*
* **Developers:** *Optimizes app size for better performance.*
* **Business Insights:** *Identifies trends in storage requirements across different app types.*

### **9.How many apps were last updated in 2018?**

#### ***Understanding the Analysis:****This analysis finds how many* ***apps were last updated in 2018****, helping understand how many apps are outdated or actively maintained.*

#### ***Process of Analysis:***

* ***Data Cleaning:*** *Convert the "Last Updated" column to datetime format.*
* ***Filtering:*** *Count apps last updated in* ***2018****.*

#### ***Impact of the Solution:***

* ***Users:*** *Helps identify how many apps might be outdated.*
* ***Developers:*** *Understand how frequently apps are updated.*
* ***Business Insights:*** *Assists in analyzing app maintenance trends.*

### **10.What is the correlation between the number of installs and the app rating?**

#### ***Understanding the Analysis*:***This analysis examines the* ***relationship between the number of installs and app ratings*** *to determine if popular apps tend to have higher ratings.*

#### ***Process of Analysis:***

* ***Data Cleaning:*** *Convert "Installs" to numeric (remove +, ,) and drop missing ratings.*
* ***Correlation Calculation:*** *Use* ***Pearson correlation [corr()]*** *to measure the relationship .*

#### ***Impact of the Solution:***

* ***Users:*** *Understand if highly installed apps are also highly rated.*
* ***Developers:*** *Optimize strategies to improve both downloads and ratings.*
* ***Business Insights:*** *Helps in predicting app success based on installs and user feedback.*

### **11.Which app categories have the highest average rating?**

#### ***Understanding the Analysis:****This analysis finds* ***which app categories have the highest average ratings****, helping identify the most well-received types of apps.*

#### ***Process of Analysis:***

* ***Data Cleaning:*** *Remove missing ratings.*
* ***Grouping & Aggregation:*** *Calculate the* ***average rating per category*** *and sort in descending order.*

#### ***Impact of the Solution:***

* ***Users:*** *Discover high-quality app categories.*
* ***Developers:*** *Focus on well-rated categories for app development.*
* ***Business Insights:*** *Identify sectors with strong user satisfaction and engagement.*

***12.*How does the price of an app affect its average rating?**

#### ***Understanding the Analysis:****This analysis examines whether paid apps receive higher ratings than free apps, helping understand if pricing influences user satisfaction.*

#### ***Process of Analysis:***

* ***Data Cleaning:*** *Convert "Price" to numeric (remove $).*
* ***Grouping & Aggregation:*** *Calculate the average rating for paid apps by filtering the “Type” column as paid type.*

#### ***Impact of the Solution:***

* ***Users:*** *Helps decide if paid apps provide better quality.*
* ***Developers:*** *Understand if pricing influences user satisfaction.*
* ***Business Insights:*** *Helps in pricing strategies for better engagement*

***13.*What is the distribution of app ratings across different content ratings?**

#### ***Understanding the Analysis:****This analysis visualizes how* ***app ratings vary for different content ratings*** *(e.g., Everyone, Teen, Mature) using a box plot to show the spread and distribution.*

#### ***Process of Analysis:***

#### ***Data Cleaning:*** *Remove missing ratings.*

#### ***Visualization:*** *Use a* ***box plot*** *to display the rating distribution across content ratings.*

#### ***Impact of the Solution:***

* ***Users:*** *Helps find stable and well-rated content categories.*
* ***Developers:*** *Understand rating variations for different audiences.*
* ***Business Insights:*** *Guides app classification for better market positioning.*

### **14.Which genres have the most apps with over 1 million installs?**

#### ***Understanding the Analysis:****This analysis identifies which* ***app genres have the highest number of apps with over 1 million installs****, helping understand popular categories.*

#### ***Process of Analysis:***

* ***Data Cleaning:*** *Convert "Installs" to numeric (remove +,,).*
* ***Filtering:*** *Select apps with* ***Installs > 1,000,000****.*
* ***Grouping & Counting:*** *Count apps and total Installs per genre and sorting in descending order.*

#### ***Impact of the Solution:***

* ***Users:*** *Helps identify trending app genres.*
* ***Developers:*** *Understand high-demand categories for new app development.*
* ***Business Insights:*** *Supports investment and marketing decisions based on genre popularity.*

### **15.How frequently do apps get updated? Calculate the average time between updates.**

#### ***Understanding the Analysis:****This analysis determines* ***how often apps are updated*** *by calculating the average time between updates. It helps understand developer engagement and app maintenance trends.*

#### **Process of Analysis:**

#### ***Data Cleaning:***

#### *The* ***"Last Updated"*** *column contains dates in string format.*

#### *Convert it to a* ***datetime format*** *for accurate date-based calculations.*

#### *Handle missing or incorrect values by removing or correcting them.*

#### ***Sorting the Data:***

* *Sort the dataset by* ***App Name*** *and* ***Last Updated Date*** *to ensure updates are in the correct order.*
* *This allows us to track the sequence of updates for each app.*
* ***Calculating Update Frequency:***
* U*se the .diff() function to find the time gap between consecutive updates* ***for each app****.*
* *This provides a list of time differences between updates in days.*
* ***Computing the Average Time Between Updates:***
* *Take the* ***mean*** *of all update time gaps to find the* ***average update frequency*** *across apps.*

***Impact of the Solution:***

* ***Users:*** *Helps identify frequently updated apps for better security and features.*
* ***Developers:*** *Understand update trends to stay competitive.*
* ***Business Insights:*** *Helps businesses track app maintenance trends and assess developer commitment*

### **16.What is the impact of app size on the number of installs?**

#### ***Understanding the Analysis:****This analysis visually explores* ***the relationship between app size and the number of installs*** *using a scatter plot to see if larger or smaller apps are preferred by users.*

#### ***Process of Analysis:***

* ***Data Cleaning:*** *Convert "Size" to MB and "Installs" to numeric values.*
* ***Visualization:*** *Plot a* ***scatter plot*** *where each point represents an app, showing its size vs. number of installs.*

#### ***Impact of the Solution:***

#### ***Users:*** *Helps in deciding whether to install larger or smaller apps.*

#### ***Developers:*** *Understand how app size affects user downloads.*

#### ***Business Insights:*** *Provides insights into optimizing app size for better reach.*

### **17.Which apps have the highest number of reviews, and what are their ratings?**

#### ***Understanding the Analysis:****This analysis helps identify the* ***apps with the highest number of user reviews*** *and their corresponding ratings. It provides insights into which apps are the most discussed and whether high engagement correlates with high or low ratings.*

#### ***Process of Analysis:***

* *Convert the "Reviews" column to a numeric data type.*
* *Sort the dataset in* ***descending order*** *based on the number of reviews.*
* *Extract the* ***top 5 apps*** *with the highest number of reviews, along with their ratings.*
* *Show the top-reviewed apps in a* ***tabular format*** *to highlight their popularity and ratings.*

### **18.How does the content rating distribution differ between free and paid apps?**

#### ***Understanding the Analysis:***This analysis examines how **content ratings** (e.g., Everyone, Teen, Mature) differ between **free and paid apps**, helping understand if paid apps are designed for a specific audience.

#### ***Process of Analysis:***

* ***Categorize Apps:*** *Classify apps as* ***Free*** *or* ***Paid*** *using the "Type" column.*
* ***Generate Crosstab Table:*** *Use pd.crosstab() to count the number of apps per content rating for each type.*
* ***Display Table:*** *Show the data in a structured format.*

#### ***Impact of the Solution:***

* ***Users:*** *Helps in choosing apps based on content suitability.*
* ***Developers:*** *Assists in selecting the right content rating for app monetization.*
* ***Business Insights:*** *Helps in marketing strategies for different audience segments*

### **19.What are the top 5 categories with the most installs?**

#### ***Understanding the Analysis:****This analysis identifies the* ***top 5 app categories*** *with the highest number of installs, helping understand which types of apps are most popular among users.*

#### ***Process of Analysis:***

* ***Convert Installs to Numeric:*** *Remove symbols (+, ,) and convert to numbers.*
* ***Group by Category:*** *Sum the total installs for each app category.*
* ***Sort and Select Top 5:*** *Rank categories by total installs and display the top 5.*

#### ***Impact of the Solution:***

* ***Users:*** *Understand which app categories are most used.*
* ***Developers:*** *Focus on high-demand categories for app development.*
* ***Business Insights:*** *Identify market trends for investment and growth strategies.*

### **20.What are the top 10 apps with the highest ratings, and how do their number of reviews and installs compare?**

#### ***Understanding the Analysis:****This analysis identifies the* ***top 10 apps with the highest ratings*** *and compares their* ***number of reviews*** *and* ***installs*** *to see if high ratings correlate with popularity.*

#### ***Process of Analysis:***

* ***Filter Data:*** *Select relevant columns: "App," "Rating," "Reviews," and "Installs."*
* ***Convert Data Types:*** *Ensure "Rating" is numeric and clean "Installs" data.*
* ***Sort & Select:*** *Rank apps by* ***highest rating*** *and pick the* ***top 10****.*
* ***Compare Metrics:*** *Display ratings, reviews, and installs together for insights.*

#### ***Impact of the Solution:***

* ***Users:*** *Find the best-rated apps with high user engagement.*
* ***Developers:*** *Identify trends among top-rated apps.*
* ***Business Insights:*** *Understand how ratings impact popularity and installs.*

### **21. Analyze the trend of app updates over time. Are there any noticeable patterns or seasonal trends?**

#### ***Understanding the Analysis:****This analysis examines* ***how frequently apps are updated over time****, identifying any* ***patterns or seasonal trends*** *in app maintenance and improvements.*

#### ***Process of Analysis:***

* ***Convert "Last Updated" to Date Format*** *for accurate time-based analysis.*
* ***Grouping & Aggregation:****Group the data by* ***year*** *and* ***month*** *to observe long-term and seasonal trends and count the number of app updates in each time period.*
* ***Trend Visualization:***
* *Use a* ***line chart*** *to display the number of updates over time.*
* *Look for* ***spikes or dips*** *that indicate seasonal patterns (e.g., more updates before holidays).*
* ***Pattern Analysis:***
* *Identify trends like* ***increased updates in recent years*** *(showing active maintenance).*
* *Check for* ***seasonal peaks*** *(e.g., updates before major shopping seasons or app store policy changes).*

#### ***Impact of the Solution:***

* ***Users:*** *Understand how actively apps are maintained.*
* ***Developers:*** *Identify the best time for updates based on trends.*
* ***Business Insights:*** *Spot seasonal update patterns for marketing and engagement strategies.*

### **22.How does the average rating of apps change with the number of installs? Create a binned analysis.**

#### ***Understanding the Analysis:****This analysis examines how an app’s average rating changes as its number of installs increases by grouping installs into bins.*

#### ***Process of Analysis:***

* ***Clean & Convert Data:*** *Convert "Installs" and "Rating" columns to numeric format.*
* ***Binning the Data:***
* *Define* ***install bins*** *(e.g., <1K, 1K-10K, 10K-100K, etc.) to group apps based on their popularity.*
* *Categorize each app into one of these bins using the pandas* ***cut function****.*

### ***Calculating the Average Rating per Bin:***

### ***Group the data*** *by install bins.*

### ***Calculate the average rating*** *within each bin to see how it changes across different install levels.*

### ***Data Visualization:***

* ***Bar Chart:*** *Plot the average rating for each install range to identify trends.*
* ***Insights from Visualization:***
* *Do apps with higher installs generally have higher or lower ratings?*
* *Are there significant differences in ratings across install levels?*

### ***Interpretation & Insights:***

* *Identify if* ***more popular apps*** *tend to have* ***higher or lower*** *ratings.*
* *Understand how install count influences* ***user perception and satisfaction****.*
* *Use insights to* ***optimize app development and marketing strategies.***
* ***Impact of the Solution:***
* ***Users:*** *Understand if highly installed apps tend to have better ratings.*
* ***Developers:*** *See how install count affects user satisfaction.*
* ***Business Insights:*** *Identify rating trends to optimize app strategies*

### **23.What is the relationship between app genre and user ratings? Are certain genres consistently rated higher or lower?**

#### ***Understanding the Analysis:****This analysis examines how* ***user ratings vary across different app genres****, identifying genres that are consistently rated higher or lower.*

#### ***Process of Analysis:***

#### ***Extract & Clean Data:*** *Select "Category" (Genre) and "Rating" columns, converting ratings to numeric.*

#### ***Grouping & Aggregation:***

* ***Calculate the mean rating*** *for each genre to see the* ***average user sentiment****.*
* ***Calculate the median rating*** *to reduce the effect of extreme values and provide a* ***more robust central tendency****.*

### ***Interpretation & Insights:***

* ***Genres with high mean & median ratings*** *indicate consistently well-rated apps.*
* ***Large differences between mean & median*** *suggest skewed ratings due to* ***outliers or a few very low/high-rated apps****.*
* *Helps* ***developers focus on highly-rated genres*** *and* ***businesses optimize app investments****.*

#### ***Impact of the Solution:***

#### ***Users:*** *Identify genres with higher-rated apps for better choices.*

#### ***Developers:*** *Focus on genres with* ***better user reception****.*

#### ***Business Insights:*** *Target high-rated genres for investment and marketing strategies.*

#### ***Key Findings and Insights:***

***1.Category & Popularity Trends:***

* ***Top 3 Categories by Installations:***
* *“Games” (30-40% of total installs)*
* *“Communication” (15-20%)*
* *“Productivity” (10-15%)*
* ***Underperforming Categories:***
* *“Beauty”,"events", “Parenting” (<2% of total installs).*
* *Games & Social Media apps dominate , accounting for over 50% of total installs.*

***2. Ratings & Reviews Impact:***

* *Apps with the prices under $50 with the ratings ranges between 3.7 to 4.7, demonstrate higher user satisfaction and preferences by user.*
* *Apps with ratings above 4.5 stars receive 35% more installs on average than lower-rated apps.*
* *Apps with fewer than 100 reviews have a 60% lower chance of appearing in top charts.*
* *Correlation between 'Installs' and 'Ratings' is approximately 0.5, indicating a weak positive relationship. This suggests that while there is a slight tendency for apps with more installs to have higher ratings, the relationship is not strong enough to imply a significant.*

***3.Pricing Strategy: Free vs. Paid Apps:***

* *free apps dominate with 92% of total apps on the Play Store.*
* *Paid apps account for only 8% but generate more revenue per download.*
* *Paid apps with prices below $5 perform 3x better in downloads than apps priced above $5.*

***4. Content Rating & User Preferences:***

* *Apps rated “Everyone” receive 60% of total installs compared to apps rated “Teen” or “Mature.*
* *Mature content apps receive 30% fewer installs but have higher user engagement per session.*

***5. App Size & Performance:***

* *Apps under 50MB have 25% higher download rates than larger apps.*
* *Apps exceeding 100MB tend to have 40% lower retention rates, as users uninstall due to storage concerns.*

# ***Actionable Recommendations:***

1. ***Target Popular Categories:*** *Invest in high-demand categories like Games & Productivity for better user transaction.*
2. ***Optimize App Ratings:*** *Apps should maintain a 4.5+ rating to boost downloads by 35%.*
3. ***Strategic Pricing:*** *Consider premium models or pricing below $5 for better conversions.*
4. ***Manage App Size:*** *Keep apps under 50MB to maximize installs.*
5. ***Use the Right Content Rating:*** *“Everyone” rated apps get the most downloads, but Mature content apps see higher engagement per session.*

***Dataset:***

[*https://www.kaggle.com/datasets/lava18/google-play-store-apps?select=googleplaystore.csv*](https://www.kaggle.com/datasets/lava18/google-play-store-apps?select=googleplaystore.csv)

***Project Owner:***

***Name: Diya Surana***

***Date: 28,February 2025***