Power BI Project: Google Play Store Apps Analysis

1. Introduction

The Google Play Store is the largest mobile app marketplace globally, hosting millions of apps across various categories. Analyzing this platform offers valuable insights for developers, marketers, and businesses. This project leverages Power BI to perform an in-depth analysis of Google Play Store data, focusing on app ratings, reviews, installs, revenue, and user sentiment.

The objective is to discover trends, identify high-performing categories and apps, understand user feedback, and uncover insights that can guide app development and marketing strategies.

2. Project Objectives

- **To analyze app distribution across categories** – Understand which categories have the most apps and which are most popular based on installs and reviews.
- **To evaluate user engagement** – Measure total installs, reviews, and ratings to determine user interest and satisfaction.
- **To estimate app monetization performance** – Calculate total revenue and identify top-grossing apps and categories.
- **To study user sentiment** – Analyze review sentiment (positive, negative, neutral) and sentiment scores to assess public perception.
- **To correlate financial and user data** – Combine pricing, installs, and sentiment data to uncover what factors lead to higher revenue and better user feedback.
- **To support strategic decisions** – Provide actionable insights for developers and stakeholders to enhance app offerings, marketing, and customer satisfaction.

3. Datasets Used

- **googleplaystore.csv** – App metadata
  - Columns: App, Category, Rating, Reviews, Size, Installs, Type, Price, Content Rating, Genres, Last Updated, etc.

App	Category	Rating	Reviews	Size	Installs	Type	Price	Content Ra	Genres	Last Updat	Current Ve	Android Ver
Photo Edit	ART_AND_	4.1	159	19M	10,000+	Free	0	Everyone	Art & Desi	January 7,	1.0.0	4.0.3 and up
Coloring b	ART_AND_	3.9	967	14M	500,000+	Free	0	Everyone	Art & Desi	January 15	2.0.0	4.0.3 and up
U Launche	ART_AND_	4.7	87510	8.7M	5,000,000+	Free	0	Everyone	Art & Desi	August 1, 2	1.2.4	4.0.3 and up
Sketch - D	ART_AND_	4.5	215644	25M	50,000,000	Free	0	Teen	Art & Desi	June 8, 20	Varies with	4.2 and up
Pixel Draw	ART_AND_	4.3	967	2.8M	100,000+	Free	0	Everyone	Art & Desi	June 20, 20	1.1	4.4 and up
Paper flow	ART_AND_	4.4	167	5.6M	50,000+	Free	0	Everyone	Art & Desi	March 26,	1	2.3 and up
Smoke Eff	ART_AND_	3.8	178	19M	50,000+	Free	0	Everyone	Art & Desi	April 26, 20	1.1	4.0.3 and up
Infinite Pai	ART_AND_	4.1	36815	29M	1,000,000+	Free	0	Everyone	Art & Desi	June 14, 20	6.1.61.1	4.2 and up
Garden Co	ART_AND_	4.4	13791	33M	1,000,000+	Free	0	Everyone	Art & Desi	September	2.9.2	3.0 and up
Kids Paint	ART_AND_	4.7	121	3.1M	10,000+	Free	0	Everyone	Art & Desi	July 3, 201	2.8	4.0.3 and up
Text on Ph	ART_AND_	4.4	13880	28M	1,000,000+	Free	0	Everyone	Art & Desi	October 27	1.0.4	4.1 and up
Name Art	ART_AND_	4.4	8788	12M	1,000,000+	Free	0	Everyone	Art & Desi	July 31, 20	1.0.15	4.0 and up
Tattoo Na	ART_AND_	4.2	44829	20M	10,000,000	Free	0	Teen	Art & Desi	April 2, 20	3.8	4.1 and up
Mandala C	ART_AND_	4.6	4326	21M	100,000+	Free	0	Everyone	Art & Desi	June 26, 20	1.0.4	4.4 and up
3D Color P	ART_AND_	4.4	1518	37M	100,000+	Free	0	Everyone	Art & Desi	August 3, 2	1.2.3	2.3 and up
Learn To D	ART_AND_	3.2	55	2.7M	5,000+	Free	0	Everyone	Art & Desi	June 6, 201	NaN	4.2 and up
Photo Des	ART_AND_	4.7	3632	5.5M	500,000+	Free	0	Everyone	Art & Desi	July 31, 20	3.1	4.1 and up
350 Diy Ro	ART_AND_	4.5	27	17M	10,000+	Free	0	Everyone	Art & Desi	November	1	2.3 and up
FlipaClip -	ART_AND_	4.3	194216	39M	5,000,000+	Free	0	Everyone	Art & Desi	August 3, 2	2.2.5	4.0.3 and up
ibis Paint X	ART_AND_	4.6	224399	31M	10,000,000	Free	0	Everyone	Art & Desi	July 30, 20	5.5.4	4.1 and up
Logo Make	ART_AND_	4	450	14M	100,000+	Free	0	Everyone	Art & Desi	April 20, 20	4	4.1 and up
Boys Photo	ART_AND_	4.1	654	12M	100,000+	Free	0	Everyone	Art & Desi	March 20,	1.1	4.0.3 and up
Superhero	ART_AND_	4.7	7699	4.2M	500,000+	Free	0	Everyone 1	Art & Desi	July 12, 20	2.2.6.2	4.0.3 and up
Mcqueen C	ART_AND_	NaN	61	7.0M	100,000+	Free	0	Everyone	Art & Desi	March 7, 2	1.0.0	4.1 and up
HD Mickey	ART_AND_	4.7	118	23M	50,000+	Free	0	Everyone	Art & Desi	July 7, 201	1.1.3	4.1 and up
Harley Qui	ART_AND_	4.8	192	6.0M	10,000+	Free	0	Everyone	Art & Desi	April 25, 20	1.5	3.0 and up
Colorfit - D	ART_AND_	4.7	20260	25M	500,000+	Free	0	Everyone	Art & Desi	October 11	1.0.8	4.0.3 and up
Animated I	ART_AND_	4.1	203	6.1M	100,000+	Free	0	Everyone	Art & Desi	March 21,	1.03	4.0.3 and up

- **googleplaystore_user_reviews.csv** – User review data
  - Columns: App, Translated Review, Sentiment, Sentiment Polarity, Sentiment Subjectivity

App	Translated	Sentiment	Sentiment	Sentiment_Subjectivity
10 Best Fo	I like eat d	Positive	1	0.533333
10 Best Fo	This help e	Positive	0.25	0.288462
10 Best Fo	nan	nan	nan	nan
10 Best Fo	Works grea	Positive	0.4	0.875
10 Best Fo	Best idea u	Positive	1	0.3
10 Best Fo	Best way	Positive	1	0.3
10 Best Fo	Amazing	Positive	0.6	0.9
10 Best Fo	nan	nan	nan	nan
10 Best Fo	Looking fo	Neutral	0	0
10 Best Fo	It helpful s	Neutral	0	0
10 Best Fo	good you.	Positive	0.7	0.6
10 Best Fo	Useful info	Positive	0.2	0.1
10 Best Fo	Thank you	Positive	0.75	0.875
10 Best Fo	Greatest e	Positive	0.992188	0.866667
10 Best Fo	Good heal	Positive	0.55	0.511111
10 Best Fo	nan	nan	nan	nan
10 Best Fo	Health It's	Positive	0.45	1
10 Best Fo	Mrs sunita	Positive	0.6	0.666667
10 Best Fo	Very Usefu	Positive	0.295	0.1
10 Best Fo	One greate	Positive	1	1
10 Best Fo	good nice	Positive	0.65	0.8
10 Best Fo	Healthy Re	Positive	0.35	0.35
10 Best Fo	God health	Neutral	0	0
10 Best Fo	HEALTH SH	Positive	0.78125	0.5
10 Best Fo	An excelle	Positive	0.65	0.5
10 Best Fo	I found lot	Neutral	0	0
10 Best Fo	Because I	Positive	0.4	1
10 Best Fo	Healthy Ea	Positive	0.5	0.5

4. Attributes in Google Play Store Data & User Reviews Data

The dataset attributes should be structured in this schema format, where each attribute is paired with its datatype for accurate data representation and handling:

Columns	Datatype
App	Varchar
Category	Varchar
Rating	Decimal
Reviews	Integer
Size	Varchar
Installs	Integer
Type	Varchar
Price	Integer
Content Rating	Varchar
Genres	Varchar
Last Updated	Date
Current version	Varchar
Android Version	Varchar

1. **App**- Application name   -   Varchar
2. **Translated Review**- User review   -   Text
3. **Sentiment**- Positive/Negative/Neutral   -   Text
4. **Sentiment Polarity**- Sentiment polarity score   -   Float
5. **Sentiment Subjectivity**- Sentiment subjectivity score - Float

5. Data Cleaning & Transformation (Power Query)

- **Price Column Cleaning**:
    - Removed currency symbols like $ in Excel.
    - Converted cleaned Price column to Fixed Decimal Number type.
    - Ensured all prices are in USD for consistency in revenue calculations.
- **Sentiment Columns**:
    - Converted Sentiment Polarity and Sentiment Subjectivity to Decimal Number.
    - Removed missing or erroneous sentiment values.

6. Data Modeling

- Established a Many-to-many relationship between:
    - googleplaystore_user_reviews[App] → googleplaystore[App]
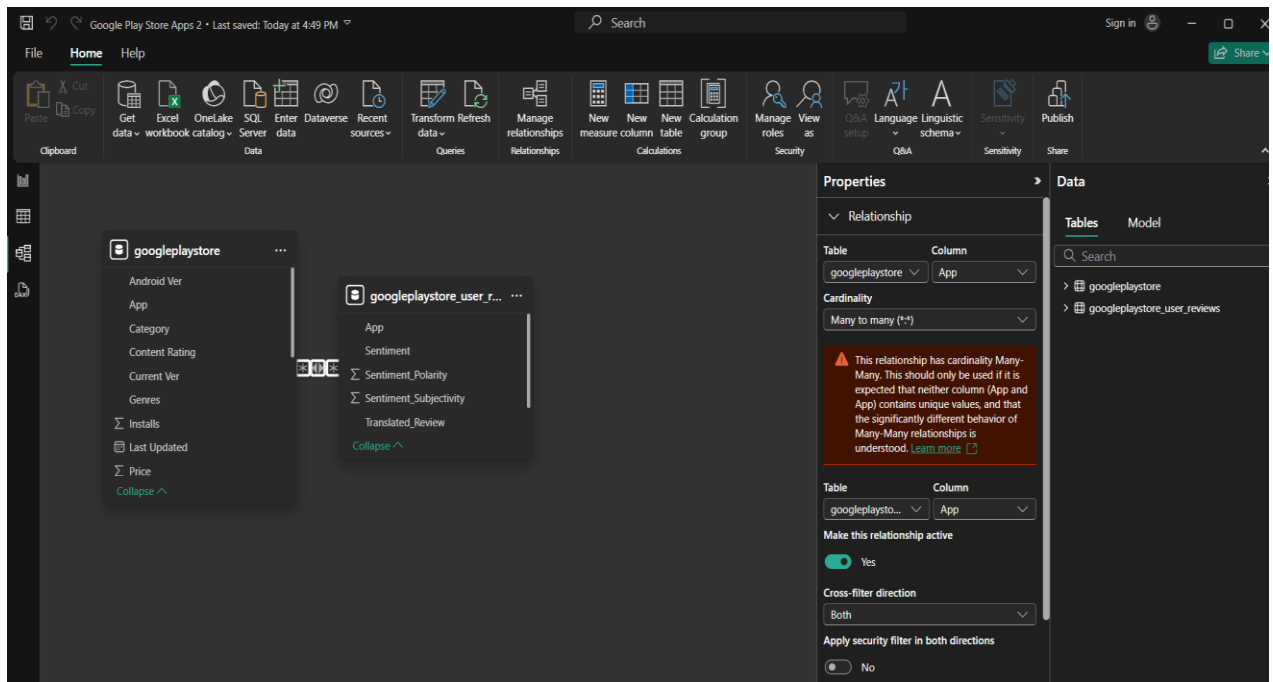
Result: This joins app metadata with user reviews for combined analysis**.**

7. Data Analysis using DAX Measures

created custom measures using DAX (Data Analysis Expressions) to calculate KPIs:

Measures Name:

```
Avg No.Of Apps = AVERAGE(googleplaystore[App])
```

```
Avg Rating = AVERAGE(googleplaystore[Rating])
```

```
Avg Review = AVERAGE(googleplaystore[Reviews])
```

```
No.Of Apps = DISTINCTCOUNT(googleplaystore[App])
```

```
Total App Cost = SUM(googleplaystore[Price])
```

```
Total Installs = SUM(googleplaystore[Installs])
```

```
Total Reviews = SUM(googleplaystore[Reviews])
```

```
Total Revenue = SUMX(googleplaystore,[Price]*[Installs])
```

```
Sentiment Count = COUNT(googleplaystore_user_reviews[Sentiment])
```

```
Total Categories = DISTINCTCOUNT(googleplaystore[Category])
```

8. Visualizations & Dashboards

8.1 Dashboard 1: Category Analysis
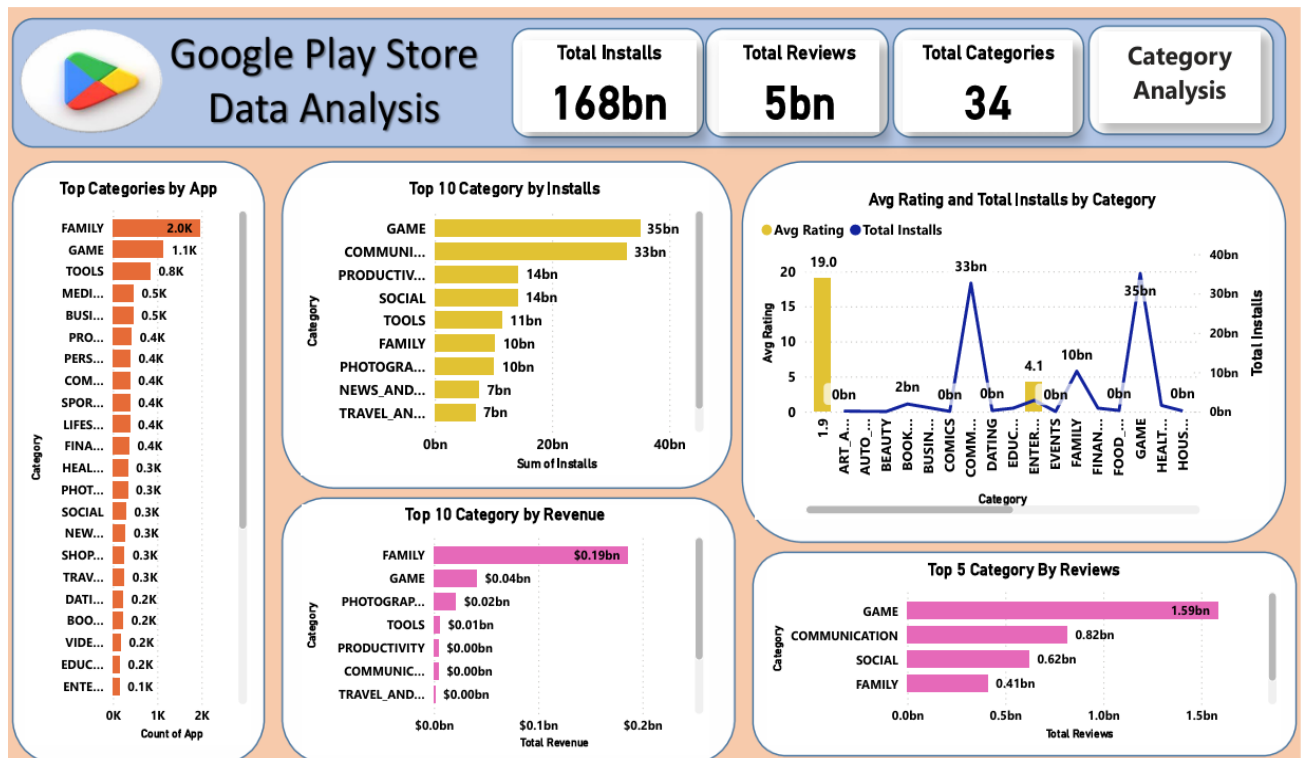
Observation:

- **Top Categories by App Count**:
  - Shows that FAMILY, GAME, and TOOLS are the most populated app categories.
- **Top Categories by Installs**:
  - GAME (35bn), COMMUNICATION (33bn) are the most downloaded.
- **Average Rating VS Installs by Category**:
  - Combines quality (rating) and popularity (installs).
  - For instance, ART_AND_DESIGN Category has a high avg rating (~19) but low installs.
- **Top Categories by Revenue**:
  - FAMILY has the highest estimated revenue ($0.19bn).
- **Top Categories by Reviews**:
 - GAME, COMMUNICATION, SOCIAL are the most discussed categories.

Insights:

- GAME is the most popular in terms of installs, reviews, and engagement, but FAMILY apps dominate in number and revenue.
- COMMUNICATION apps are highly installed and reviewed, but not monetized as strongly.

- Some categories (like ART_AND_DESIGN) show outlier data (19.0 avg rating is unusually high; needs verification).
- Revenue generation is not directly proportional to installs or app count—FAMILY is a strong monetization performer despite fewer installs than GAME.

Strategic Recommendations: Focusing on GAME and FAMILY categories may offer the best opportunities depending on goals (volume vs revenue). COMMUNICATION also holds strong user engagement but weaker monetization.

8.2   Dashboard 2: Revenue Analysis
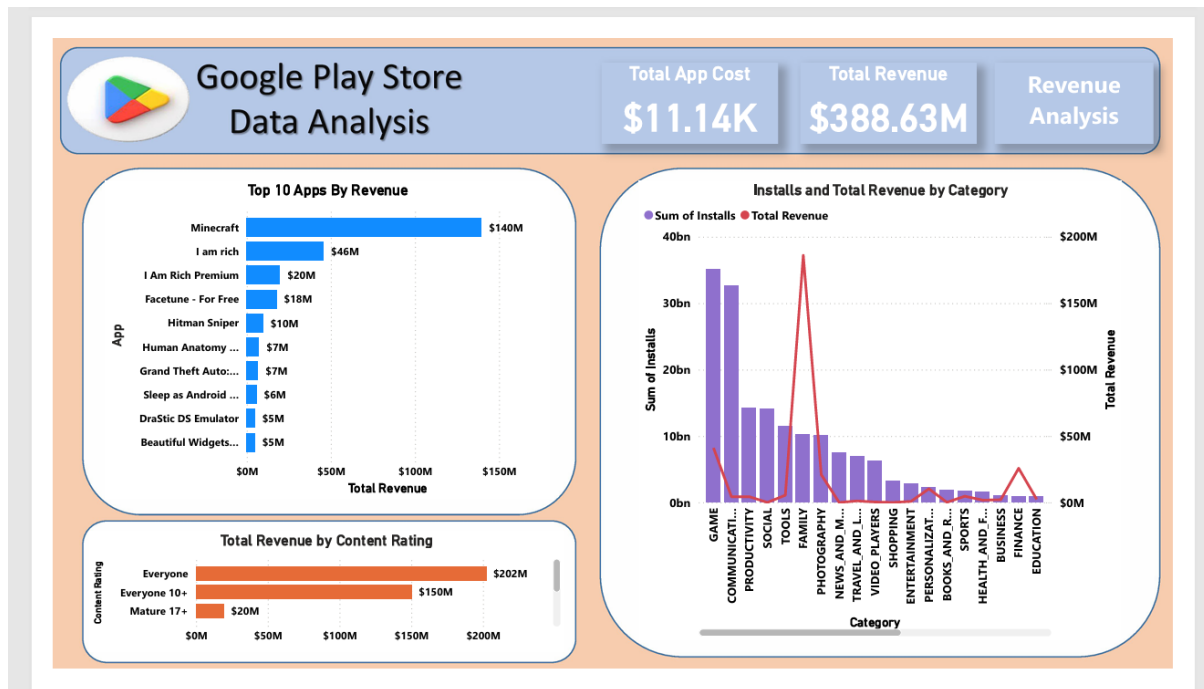
Observation:

- **Top 10 Apps by Revenue**:
  - Minecraft alone generated $140M, followed by I Am Rich ($46M).
- **Revenue vs Installs by Category**:
  - GAME category leads in installs (~35bn installs) but its revenue curve is not the highest.
  - FAMILY category shows a high revenue spike (~$190M) despite fewer installs than GAME and COMMUNICATION.
- **Revenue by Content Rating**:
  - Everyone rating dominates with $202 million revenue.
  - Everyone 10+ and Mature 17+ follow with ~$150 million and ~$20 million respectively.

Insights:

- Focus on developing "Everyone" content-rated apps to maximize reach and revenue potential.
- Consider targeting FAMILY and GAME categories:

o   FAMILY offers higher revenue per install.
o   GAME offers volume and engagement.
*   High installs in COMMUNICATION and SOCIAL may help user growth but are harder to monetize directly.
*   Small number of superstar apps dominate revenue → Competitive market for top revenue positions.
*   Some categories (e.g., PHOTOGRAPHY, TOOLS) could be under-monetized opportunities or need diversified revenue models.

Strategic Recommendations:

*   Prioritize FAMILY or highly monetizable niche apps (e.g., Minecraft, premium content).
*   If aiming for installs/userbase:
    o   Focus on GAME, COMMUNICATION but complement with advertising or freemium models.

8.3   Dashboard 3: Sentiment Analysis
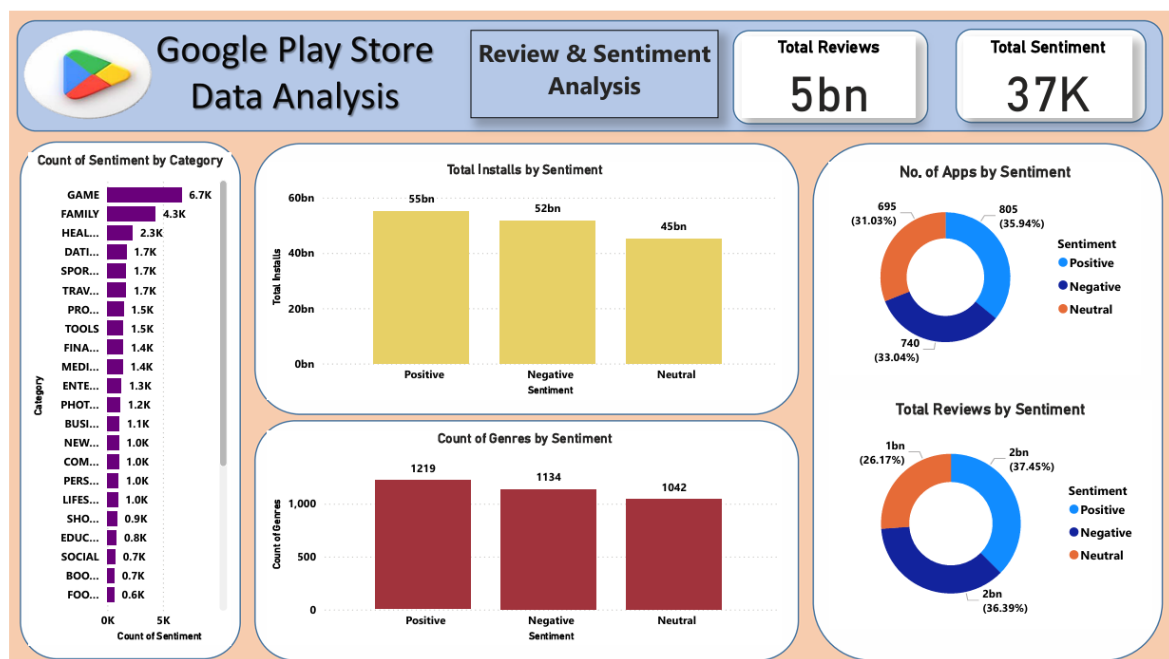
Observations:

*   **Total Reviews & Sentiments:**
    o   Total Reviews = 5 billion
    o   Total Sentiment count = 37K sentiments analyzed
*   **Sentiment Distribution by Number of Apps:**
    o   Positive sentiment → 805 apps (35.94%)
    o   Negative sentiment → 740 apps (33.04%)
    o   Neutral sentiment → 695 apps (31.03%)
*   **Total Reviews by Sentiment:**

- o Positive Sentiment → 2 billion reviews (37.45%)
    - o Negative Sentiment → 2 billion reviews (36.39%)
    - o Neutral Sentiment → 1 billion reviews (26.17%)
- **Total Installs by Sentiment:**
    - o Positive sentiment → 55 billion installs
    - o Negative sentiment → 52 billion installs
    - o Neutral sentiment → 45 billion installs
- **Count of Sentiment by Category**
    - o GAME → highest sentiment mentions (6.7K)
    - o FAMILY → 4.3K sentiments
    - o Followed by HEALTH (2.3K), DATING, SPORTS, TRAVEL (1.7K each).
- **Count of Genres by Sentiment:**
    - o Positive → 1219 genres
    - o Negative → 1134 genres
   - o Neutral → 1042 genres

Insights:

- **Positive sentiment is slightly higher** in installs, reviews, and apps, but not dominant → competition remains on app quality & user expectations.
- **Game category leads in user sentiments** → highest engagement, likely more passionate or opinionated userbase.
- **Similar proportions of positive/negative sentiment indicate polarized audiences** → developers must focus on addressing user complaints even in popular apps.
- **Higher installs don't guarantee positive sentiment** → apps with billions of downloads still receive high negative reviews.
- **Balanced sentiment across genres** → opportunity in niche categories to improve perception or reviews.

Strategy Recommendtaions:

- Focus on improving sentiment in high-engagement categories like Games and Family → even small improvements in reviews can influence perception.
- Invest in user experience, bug fixes, and support in negatively skewed apps, especially those with large install bases.
- Encourage more positive reviews via incentives or feedback prompts → leverage the slight positive tilt in sentiment for marketing.

8.4 Dashboard 4: App & Content Analysis
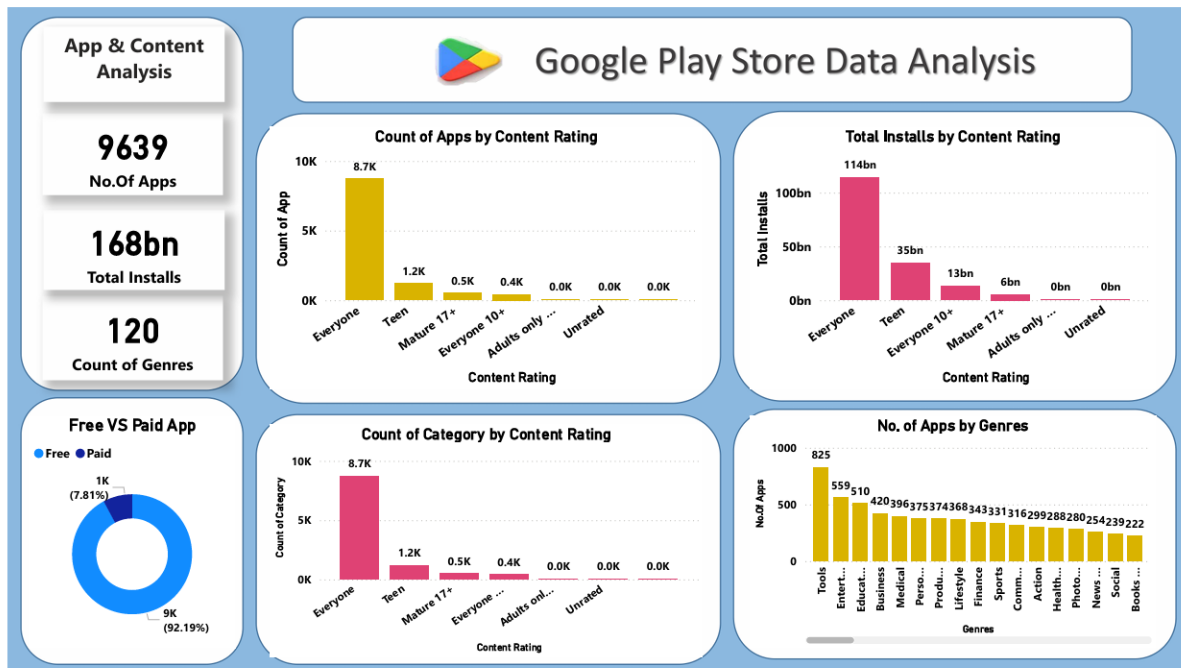
Observations:

- **Total apps:** 9639 apps
  Total installs (downloads): 168 billion
  Total genres (categories): 120 types of apps
- **Count of Apps by Content Rating**
  - Most apps are rated "Everyone" →anyone can use
  - Few apps are for "Teen" (1.2K apps), "Mature 17+" (500 apps), and "Everyone 10+" (400 apps)
  - Very few or no apps marked Adults only or Unrated
- **Total Installs by Content Rating**
  - Apps rated "Everyone" are installed the most: 114 billion installs!
  - Then comes Teen → 35 billion installs
  - Then Everyone 10+ → 13 billion installs
  - Other ratings have much fewer installs
- **Free vs Paid Apps**
  - 92% apps are FREE (9000 apps)
  - Only 8% apps are PAID (1000 apps)
- **Count of Category by Content Rating**
  - Most app categories are for "Everyone" content rating.
- **No. of Apps by Genres**
  - Most popular categories:
    - Tools (825 apps)
    - Entertainment (559 apps)
    - Education (510 apps)
    - Business (420 apps)
    - Medical (396 apps)

- ▪ Fewer apps in Social, News, Books, etc.

Insights:

- Most apps are free and made for everyone
- Most downloads are for apps safe for all ages
- Tools, Entertainment, and Education are top categories by number of apps
- Paid apps are very few

Strategy Recommendations:

- **Focus on "Everyone" content**: It's the largest market (highest number of apps and installs)
- **Grow in "Teen" & "Everyone 10+" categories**: Less competition, but high installs (good opportunity)
- **Invest in Tools, Entertainment, and Education genres**: These have the most apps and user demand.
- **Attract users with free apps & monetize through ads or in-app purchases**: Since most users prefer free apps
- **Explore niche genres for innovation (Social, News, Books)**: Lower competition, opportunity for unique apps

9. Conclusion

- **The Google Play Store is dominated by free apps aimed at all audiences:**
  - o 92% of apps are free**;** users strongly prefer free apps over paid ones.
  - o **"Everyone"** content rating accounts for most apps and installs (114 billion downloads: 68% of total installs).
  - o Developers should focus on free, family-friendly apps to reach the largest user base.
- **Top categories differ by popularity, installs, revenue, and reviews.**
  - o Games dominate installs (35 billion), reviews, and engagement.
  - o FAMILY leads in app count and revenue ($190M).
  - o Tools, Entertainment, and Education are top by number of apps → highly competitive but popular.
  - o Developers must balance between high-volume competitive categories (Games, Tools, Education) and niche opportunities (News, Social, Books).
- **Revenue is not directly proportional to installs or app count.**
  - o Despite more installs, Games generate less revenue per install than FAMILY.
  - o A few superstar apps dominate revenue (e.g., Minecraft $140M).
  - o FAMILY and niche premium apps can yield better monetization per user; volume doesn't guarantee revenue.
- **User sentiment is balanced: neither overwhelmingly positive nor negative.**
  - o Reviews are ~37% positive, 36% negative, 26% neutral.
  - o Games have the most sentiment mentions: high user engagement but also polarized opinions.

- o Developers should invest in improving reviews, bug fixes, user support: sentiment matters for reputation and installs.
- **Teen and Everyone 10+ categories are underserved but have high demand.**
  - o Fewer apps (1.2K Teen, 400 Everyone 10+) but still large install numbers: opportunity for growth.
  - o Developers can target these content ratings for lower competition with strong demand.