



## BUSINESS INTELLIGENCE FOR PLAYSTORE APP ANALYSIS

CIS4008-N-BF1-2020

SCHOOL OF COMPUTING, ENGINEERING AND DIGITAL  
TECHNOLOGY

### **Section -1**



**12 January 2022**

Prepared by

W9533735 , SAICHANDRA MUVVA

## Contents

Acknowledgement .....	3
Data set.....	4
Business Intelligence Queries .....	5
Loading Data.....	6
Creation of fact and dimension Table .....	15
Creating Download Dimension Table .....	17
Creating rating Dimension Table.....	17
Creating size dimension table .....	18
Creating version dimension table.....	18
Dax expression.....	19
M language.....	20
Creating model and relationships.....	21
Business intelligence for playstore app analysis.....	24
Executive summary .....	25
Background.....	25
Key findings.....	25
Recommendations.....	25
Findings based on Analysis and evaluation .....	26
References.....	32

## Acknowledgement

This is to state that my module tutor Annalisa Occhipinti and lab Tutor Mohammad Suleiman helped in Various stages of this project to complete this successfully and I am especially thanks to other Tutors who Helped me in securing a good Data set for completing This Ica project and providing Tips and spending Time to Acquiring a desired success business model and I extend my Thanks to other individuals who helped me through this journey

## Data set

This data set is about Details of applications from the play store it includes 13 columns and 9362 rows of data and other data set is about reviews of applications from play store and it consists of 5 columns both of these data sets are downloaded from <https://www.kaggle.com/lava18/google-play-store-apps>

The Table below gives brief about the Playstore Data set

Column name	Description of Column
App	Name of the app in a play store dataset
Category	Description of category which that belongs to
Rating	Total Rating of app out of 5
Reviews	Total number of users reviewed the app
Size	Storage required for the app in phone
installs	Number of times app is installed in phone
Type	Description of app whether paid or free
price	This column represents paid as numerical values and free as 0
Content Rating	Description of app used by age limits of users
Genre	In extent to Category Further classified
Last updated	Date in which app is last updated
Version	Describes the Features, compatibility of device
Sentiment	Describes whether user is positive or negative response over the app

## Business Intelligence Queries

This selected Data set consists of detail information of apps about their size, downloads, ratings, review and compatibility and few other parameters. by analysing these key parameters we can find the relationships and affecting parameters with the help of powerbi tool

This report senses the following Queries.

- Distribution of application over categories and installs?
- What is the Most Used age group using the application and their categories?
- What are the Most Expensive apps used in play store?
- Most money generated app from play store?
- What is the Most Popular app Used in the play store?
- Which category of apps are mostly used?
- Is there any Relationship between Rating and installs?
- Is there any Relationship between Reviews, Genre, Type And category?
- Key factors Influencing installs, Rating and Price?

## Loading Data

The first step is to Load Raw data for data analysis [1]. Power bi supports various formats of data like csv, .exe, API, space separated format and many more other files

Opening of a powerbi shows a popup window stating load existing data or get data as shown fig below

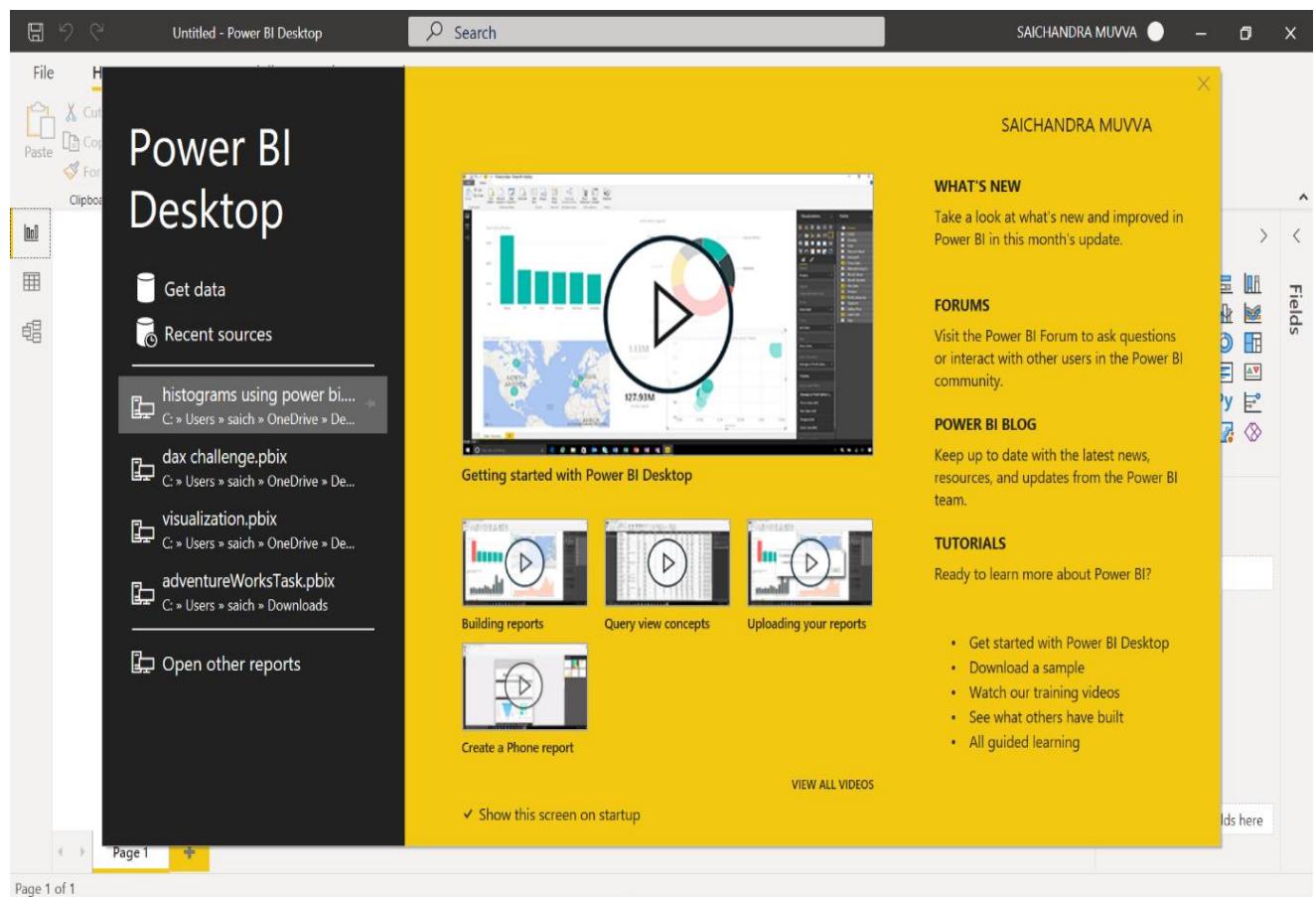


Fig 1- power bi

We must load two data sets of play store data we have to select “get data” and select “csv format” and load two data sets one is play store data and other reviews data file both are in csv format

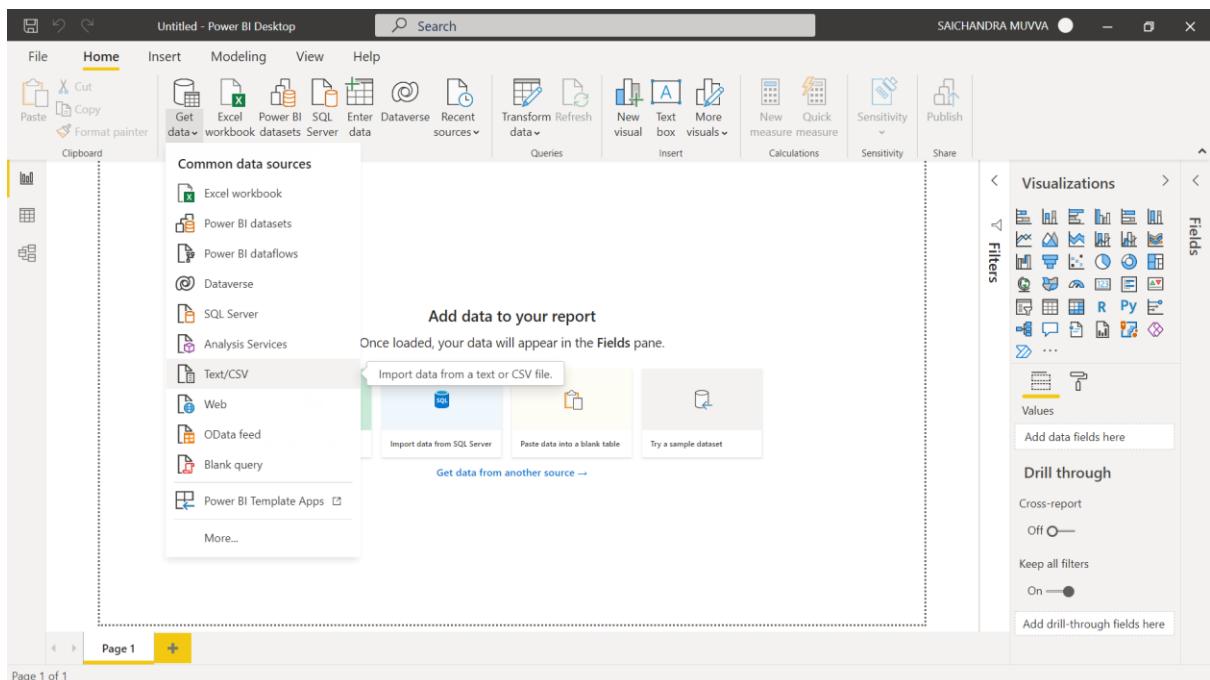


Fig 2- Dialog box

After selecting the get data we will see the options of different formats of data as shown in the above fig after selecting the csv format a popup window is shown below in diagram stating whether “Load Data” or “Transform Data” [2] simply Load the data we will be able to see the structure of data set shown below Fig

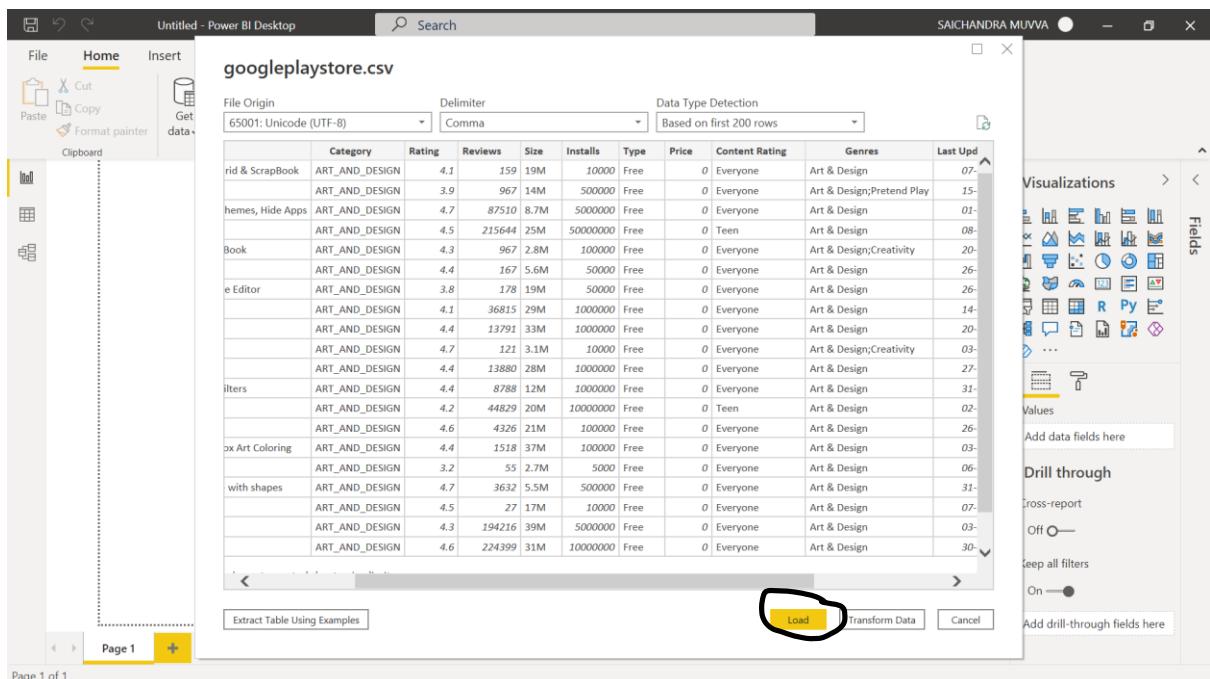


Fig 3 - Loading of play store csv file

Similarly Load the data of reviews in csv format steps as shown below

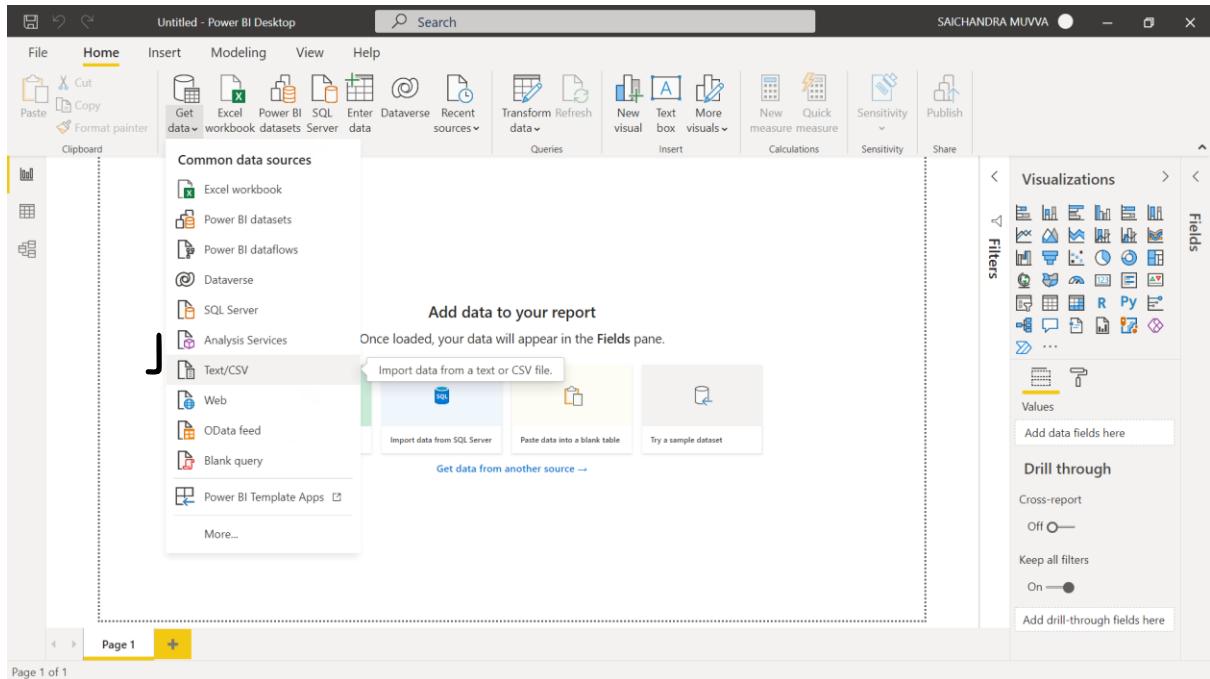


Fig 4: selecting a data menu to load csv file

Reviews File has one sheet, and it is loaded as shown below figure

The screenshot shows the Power BI Desktop interface with the 'Home' tab selected. A data preview window is open, showing a table titled 'googleplaystore\_user\_reviews.csv'. The table has five columns: Column1, Column2, Column3, Column4, and Column5. The first few rows of data are visible. At the bottom of the preview window, there are buttons for 'Extract Table Using Examples.', 'Load' (which is circled in yellow), 'Transform Data', and 'Cancel'. To the right of the preview, there's a 'Fields' pane with a search bar and a list of fields from the loaded CSV file.

Column1	Column2	Column3	Column4	Column5
App	Translated_Review	Sentiment	Sentiment_Polarity	Sentiment_Subjectivity
10 Best Foods for You	I like eat delicious food. That's I'm cooking food myself...	Positive	1.0	0.5333333333333333
10 Best Foods for You	This help eating healthy exercise regular basis	Positive	0.25	0.28846153846153844
10 Best Foods for You	nan	nan	nan	
10 Best Foods for You	Works great especially going grocery store	Positive	0.4	0.875
10 Best Foods for You	Best idea us	Positive	1.0	0.3
10 Best Foods for You	Best way	Positive	1.0	0.3
10 Best Foods for You	Amazing	Positive	0.6000000000000001	0.9
10 Best Foods for You	nan	nan	nan	
10 Best Foods for You	Looking forward app,	Neutral	0.0	0.0
10 Best Foods for You	It helpful site ! It help foods get !	Neutral	0.0	0.0
10 Best Foods for You	good you.	Positive	0.7	0.6000000000000001
10 Best Foods for You	Useful information The amount spelling errors question...	Positive	0.2	0.1
10 Best Foods for You	Thank you! Great app!! Add arthritis, eyes, immunity, ki...	Positive	0.75	0.875
10 Best Foods for You	Greatest ever Completely awesome maintain health.....	Positive	0.9921875	0.8666666666666667
10 Best Foods for You	Good health..... Good health first priority.....	Positive	0.5499999999999999	0.5111111111111112
10 Best Foods for You	nan	nan	nan	
10 Best Foods for You	Health It's important world either life , think? :)	Positive	0.45	1.0
10 Best Foods for You	Mrs sunita bhati i thankful developers,to make kind ap...	Positive	0.6	0.6666666666666666
10 Best Foods for You	Very useful in diabetes age 30. I need control sugar. th...	Positive	0.2950000000000004	0.1

Fig 5: the table must be loaded as shown above

After loading of data sets, we have seen some errors

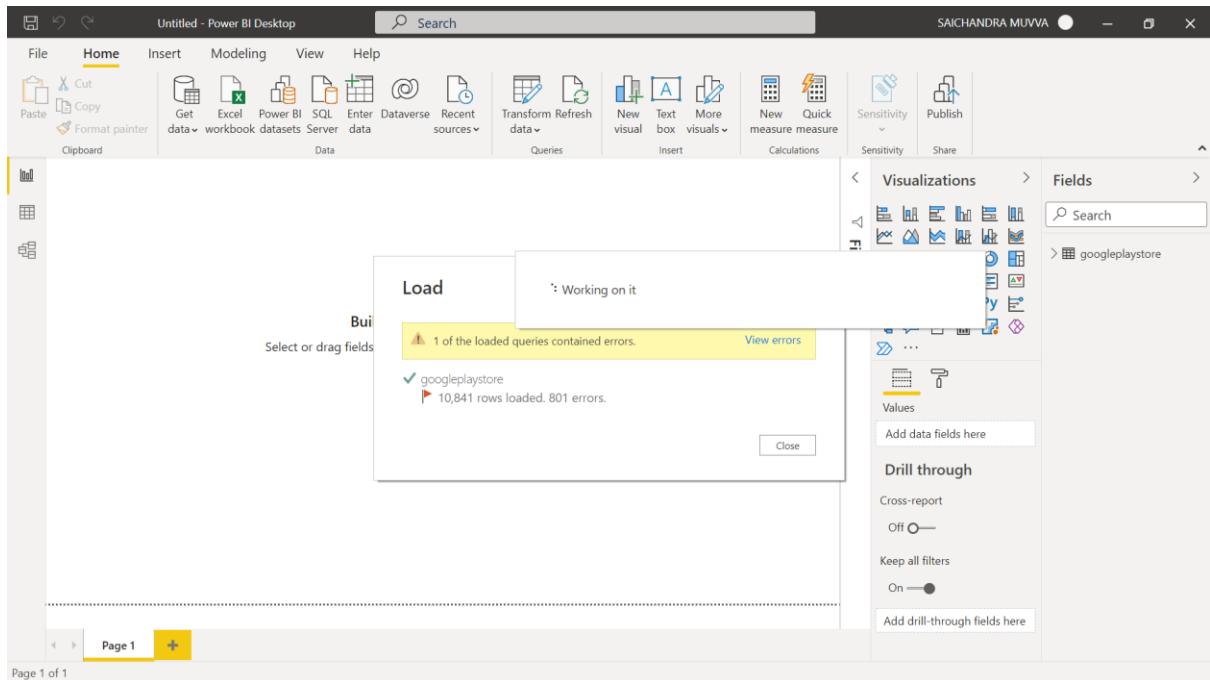


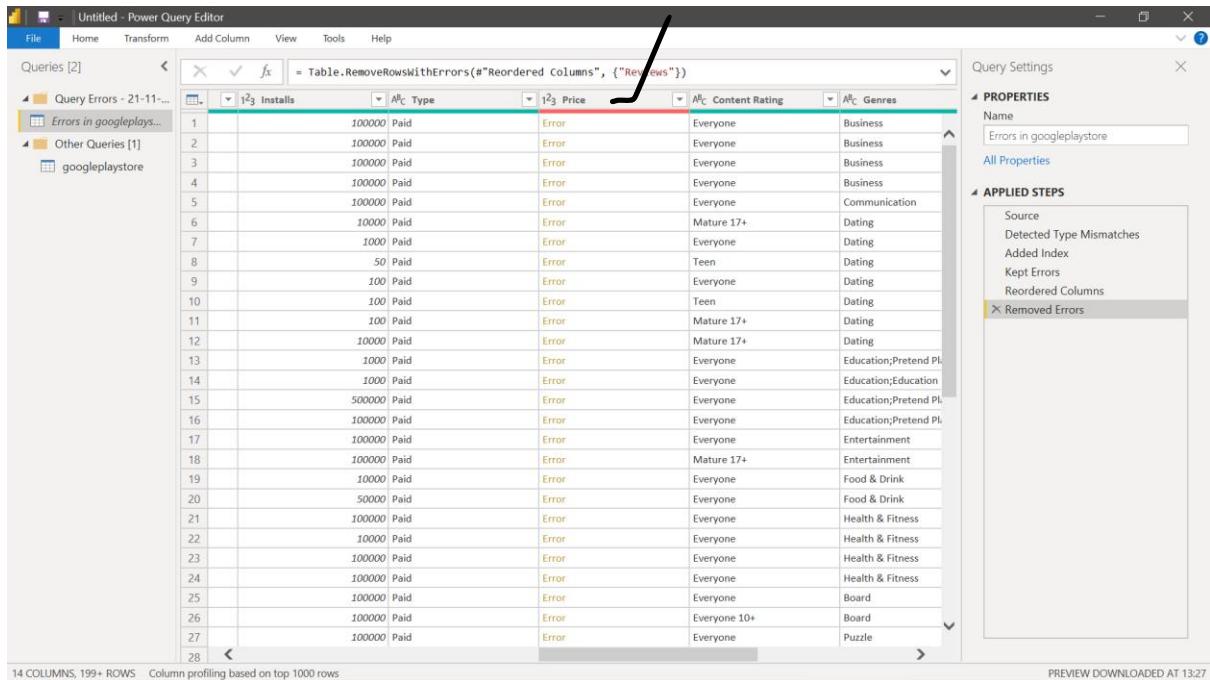
Fig 6: Errors Detected in the power bi

To resolve these errors these are seen in by clicking view errors in above Figure Simultaneously a power window editor is opened showing the error columns

Row Number	App	Category	Rating	Reviews	Size	Type
1			4.7		100000	Paid
2			4.8		100000	Paid
3			4.7		100000	Paid
4			4.8		100000	Paid
5			4		100000	Paid
6			4.2		10000	Paid
7			2.6		50	Paid
8			NaN		100	Paid
9			NaN		100	Paid
10			NaN		100000	Paid
11			3.5		100000	Paid
12			4.2		10000	Paid
13			4.9		1000	Paid
14			4.7		500000	Paid
15			4.7		100000	Paid
16			4.7		100000	Paid
17			4.6		100000	Paid
18			4.6		100000	Paid
19			4.6		100000	Paid
20			4.1		50000	Paid
21			4.4	2107 Varies with device	100000	Paid
22			4.6	1026 29M	100000	Paid
23			4.8	19543 Varies with device	100000	Paid
24			4.4	2107 Varies with device	100000	Paid
25			4.4	18621 63M	100000	Paid
26			4.6	19922 35M	100000	Paid
27			4.9	21119 48M	100000	Paid
28						

Fig 7: query editor displaying the errors

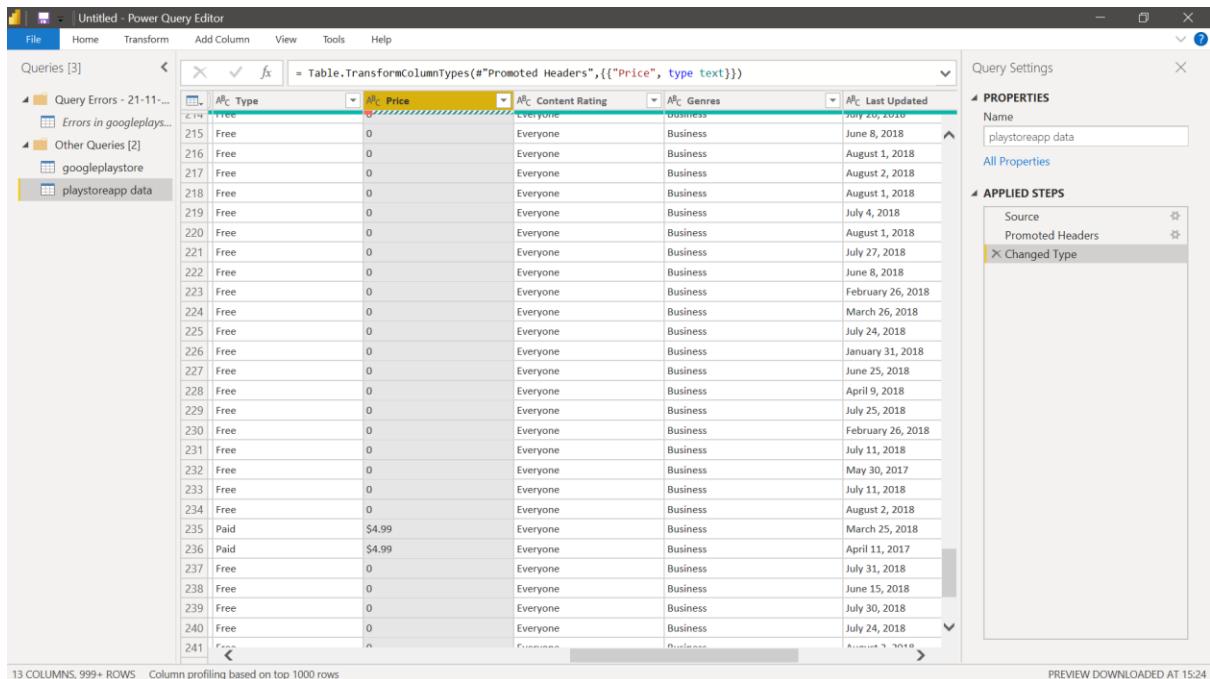
These errors can be removed by right click on the column and “remove errors” so errors are removed from column of reviews as shown in fig above



The screenshot shows the Power Query Editor interface with a table titled "Errors in googleplaystore". The table has columns: Type, Price, Content Rating, and Genres. A red arrow points to the "Price" column, which contains numerous "Error" values. The "Properties" pane on the right shows the query name is "Errors in googleplaystore" and the "Applied Steps" section highlights the "Reordered Columns" step.

Fig 8: price column showing all errors in column

To figure this out we can't delete whole column its because of “\$” symbol so before we need to remove those symbols open the query editor and change the format from 123 to ABC as shown below Fig



The screenshot shows the Power Query Editor interface with a table titled "playstoreapp data". The table has columns: Type, Price, Content Rating, Genres, and Last Updated. The "Price" column now contains text values like "Free", "0", and "\$4.99". The "Properties" pane on the right shows the query name is "playstoreapp data" and the "Applied Steps" section highlights the "Promoted Headers" and "Changed Type" steps.

Fig 9: price column is changed from numeric to text

Then replace the “\$” with empty as shown in the below figure

The screenshot shows the Power Query Editor interface. A context menu is open over the 'Price' column header. The menu includes options like Copy, Remove, Remove Other Columns, Duplicate Column, Add Column From Examples..., Replace Values..., Replace Errors..., Split Column, Group By..., Fill, Unpivot Columns, Unpivot Other Columns, Unpivot Only Selected Columns, Rename..., Move, Drill Down, and Add as New Query. The 'Replace Values...' option is highlighted.

Fig 10: showing Replace values by right clicking column

The screenshot shows the Power Query Editor with the 'Replace Values' dialog box open. The dialog has fields for 'Value To Find' containing '\$' and 'Replace With' containing an empty string. There is also an 'Advanced options' section. The background shows the same data grid as Fig 10, with the 'Price' column now containing empty values where '\$' was present.

Fig 11: showing replace values from \$ to null

The screenshot shows the Power Query Editor interface with the 'playstoreapp data' query selected. The 'Price' column, which previously contained 'Free' values, now contains '0' values. The 'APPLIED STEPS' pane on the right shows the 'Replaced Value' step, indicating that the 'Price' column was changed from 'Free' to '0'. The 'Properties' pane shows the query name as 'playstoreapp data'.

Fig 12: Price column is now updated without errors as shown in above

The screenshot shows the Power Query Editor interface with the 'playstoreapp data' query selected. The 'Price' column now displays decimal values such as '0' and '4.99'. A specific value '4.99' is circled in black. The 'APPLIED STEPS' pane on the right shows the 'Changed Type1' step, indicating that the data type of the 'Price' column was changed. The 'Properties' pane shows the query name as 'playstoreapp data'.

Fig 13: price column is updated to decimal values

Similarly, for install column change the format by deleting “+” in values and “,” as shown in the below diagrams so that column of install is cleansed

The screenshot shows the Power Query Editor interface with the following details:

- File** tab selected.
- Queries [3]** pane on the left lists "Query Errors - 21-11...", "Errors in googleplay...", "Other Queries [2]", "googleplaystore", and "playstoreapp data".
- Applied Steps** pane on the right shows steps: "Source", "Promoted Headers", "Changed Type", "Replaced Value", "Changed Type1", and "Filtered Rows".
- Table View**: The main area displays a table with columns: Size, Installs, Type, Price, and Content Rating. The "Installs" column contains values like "1,000,000+", "500,000", etc., with many entries ending in a plus sign ("+").
- Properties** pane on the right shows "Name: playstoreapp data" and "All Properties".
- Preview**: At the bottom right, it says "PREVIEW DOWNLOADED AT 15:27".

Fig 14: column of install column

By use of “Replace” command we replace “+” with empty as shown in the below figure

The screenshot shows the Power Query Editor interface after applying the "Replace" command:

- File** tab selected.
- Queries [3]** pane on the left lists "Query Errors - 21-11...", "Errors in googleplay...", "Other Queries [2]", "googleplaystore", and "playstoreapp data".
- Applied Steps** pane on the right shows steps: "Source", "Promoted Headers", "Changed Type", "Replaced Value", "Changed Type1", and "Filtered Rows".
- Table View**: The main area displays the same table as Fig 14, but the "Installs" column now contains values like "10,000", "500,000", etc., with no plus signs.
- Properties** pane on the right shows "Name: playstoreapp data" and "All Properties".
- Preview**: At the bottom right, it says "PREVIEW DOWNLOADED AT 15:28".

Fig 15: updated after using replace command for “+”

Untitled - Power Query Editor

File Home Transform Add Column View Tools Help

Queries [3]

Query Errors - 21-11... Errors in googleplaystore Other Queries [2] googleplaystore playstoreapp data

	Type	1.2 Price	Content Rating	Genres	Last Up
6	50000 Free		Everyone	Art & Design	March 2
7	50000 Free		Everyone	Art & Design	April 26
8	1000000 Free		Everyone	Art & Design	June 14
9	1000000 Free		Everyone	Art & Design	September
10	10000 Free		Everyone	Art & Design;Creativity	July 3, 2
11	1000000 Free		Everyone	Art & Design	October
12	1000000 Free		Everyone	Art & Design	July 31,
13	10000000 Free		Teen	Art & Design	April 2,
14	100000 Free		Everyone	Art & Design	June 26
15	100000 Free		Everyone	Art & Design	August
16	50000 Free		Everyone	Art & Design	June 6,
17	500000 Free		Everyone	Art & Design	July 31,
18	10000 Free		Everyone	Art & Design	November
19	5000000 Free		Everyone	Art & Design	August
20	100000000 Free		Everyone	Art & Design	July 30,
21	1000000 Free		Everyone	Art & Design	April 20
22	1000000 Free		Everyone	Art & Design	March 2
23	500000 Free		Everyone 10+	Art & Design	July 12,
24	50000 Free		Everyone	Art & Design	July 3, 2
25	10000 Free		Everyone	Art & Design	April 25
26	5000000 Free		Everyone	Art & Design;Creativity	October
27	1000000 Free		Everyone	Art & Design	March 2
28	10000 Free		Everyone	Art & Design	July 12,
29	1000000 Free		Everyone	Art & Design	August
30	1000000 Free		Everyone	Art & Design	July 12,
31	50000 Free		Everyone	Art & Design	May 31,
32	1000000 Free		Everyone	Art & Design	July 19,
33					

13 COLUMNS, 999+ ROWS Column profiling based on top 1000 rows PREVIEW DOWNLOADED AT 15:29

Fig 16: final column after using command for “,”

f Untitled - Power Query Editor

File Home Transform Add Column View Tools Help

Queries [2]

Query Errors - 27-12... Errors in googleplaystore Other Queries [1] googleplaystore

	Genres	Last Updated	Current Ver	Android Ver	Index
1	Copy Entire Table	25-03-2018	1.5.2	4.0 and up	1
2	Use First Row as Headers	11-04-2017	3.4.6	3.0 and up	2
3	Add Custom Column...	25-03-2018	1.5.2	4.0 and up	3
4	Add Column From Examples...	11-04-2017	3.4.6	3.0 and up	4
5	Invoke Custom Function...	05-07-2018	7.5.3.20547	4.1 and up	5
6	Add Conditional Column...	19-06-2018	2.1.139	4.1 and up	6
7	Add Index Column	From 0	1.6	4.0 and up	7
8	Choose Columns...	From 1	1.6	4.0 and up	8
9	Keep Top Rows...	Custom...	1.0.1	4.0 and up	9
10	Keep Bottom Rows...	02-08-2018	1.2	4.0.3 and up	10
11	Keep Range of Rows...	21-01-2017	1.0.1	4.0 and up	11
12	Keep Duplicates	19-06-2018	2.6.139	4.1 and up	12
13	Keep Errors	24-07-2017	1	4.0.3 and up	13
14	Remove Top Rows...	23-07-2017	1.3	4.1 and up	14
15	Remove Bottom Rows...	06-07-2018	1.5-play	4.4 and up	15
16	Remove Alternate Rows...	12-06-2018	1.1.1-play	4.4 and up	16
17	Remove Duplicates	30-06-2018	Varies with device	Varies with device	17
18	Remove Errors	03-08-2018	4.426	4.1 and up	18
19	Merge Queries...	28-06-2018	Varies with device	Varies with device	19
20	Append Queries...	03-06-2018	1.4.4	4.0 and up	20
21	From 0 to N Rows...	22-12-2015	Varies with device	Varies with device	21
22	Ith & Fitness	15-05-2018	1.0.6	4.3 and up	22
23	Ith & Fitness	19-01-2018	Varies with device	Varies with device	23
24	Ith & Fitness	22-12-2015	Varies with device	Varies with device	24
25	rd	04-07-2018	2.1.2	4.4 and up	25
26					

14 COLUMNS, 800 ROWS Column profiling based on top 1000 rows PREVIEW DOWNLOADED AT 18:35

fig 17: adding a index coloumn to data set

Table: playstore android (9,364 rows) Column: Index (9,364 distinct values)

Fig 18: Final model after the data set cleaning and removing errors

## Creation of fact and dimension Table

After loading the data and cleaning the errors in the data we are creating dimension tables [4] in further so for creation we need to open the power query editor window for this click on the transform data shown on the Fig 19 so that a new window is popped on the desktop

Table: playstore android (9,364 rows) Column: Index (9,364 distinct values)

Fig 19 – Transform Data to open power query editor

Power query editor is opened as shown in the figure and we need to create a duplicate tables and name them as following dimension tables for creation of duplicate table right click on the “play store android” as shown in the below and then select the “Duplicate” similar to original query as shown in the fig20

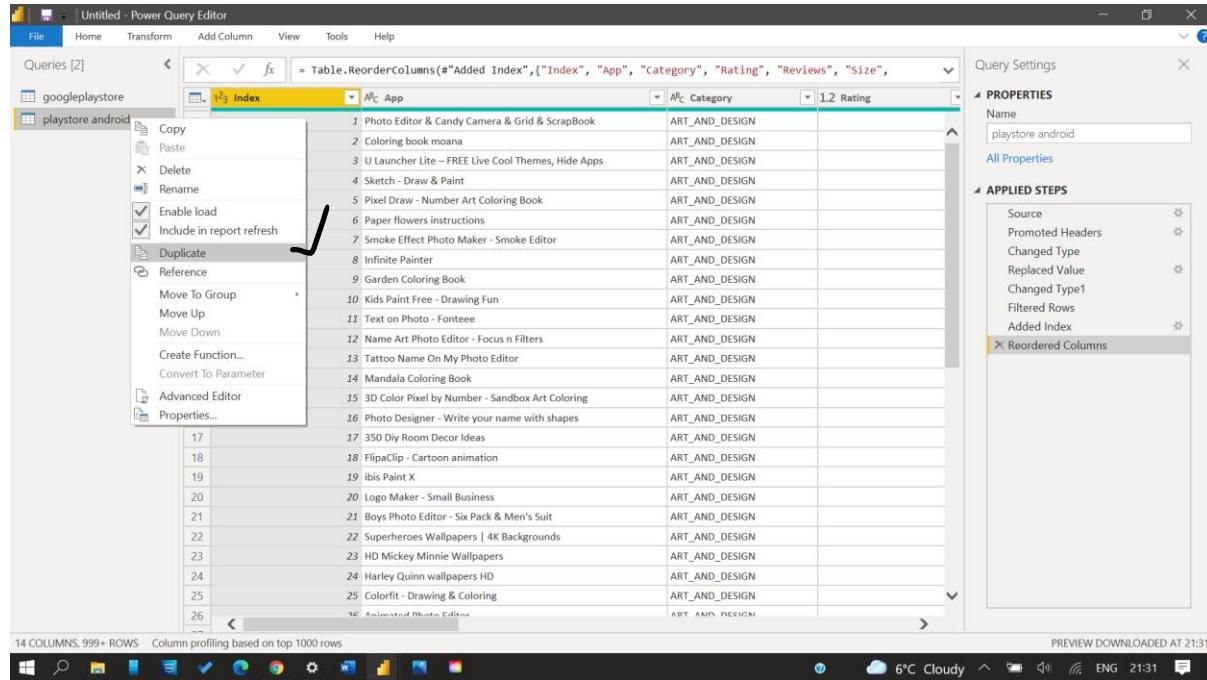


Fig 20 – creation Duplicate in power query editor

Select the column of needed by selecting following column and right click on the selected column and then “Remove other columns” so that another column is deleted

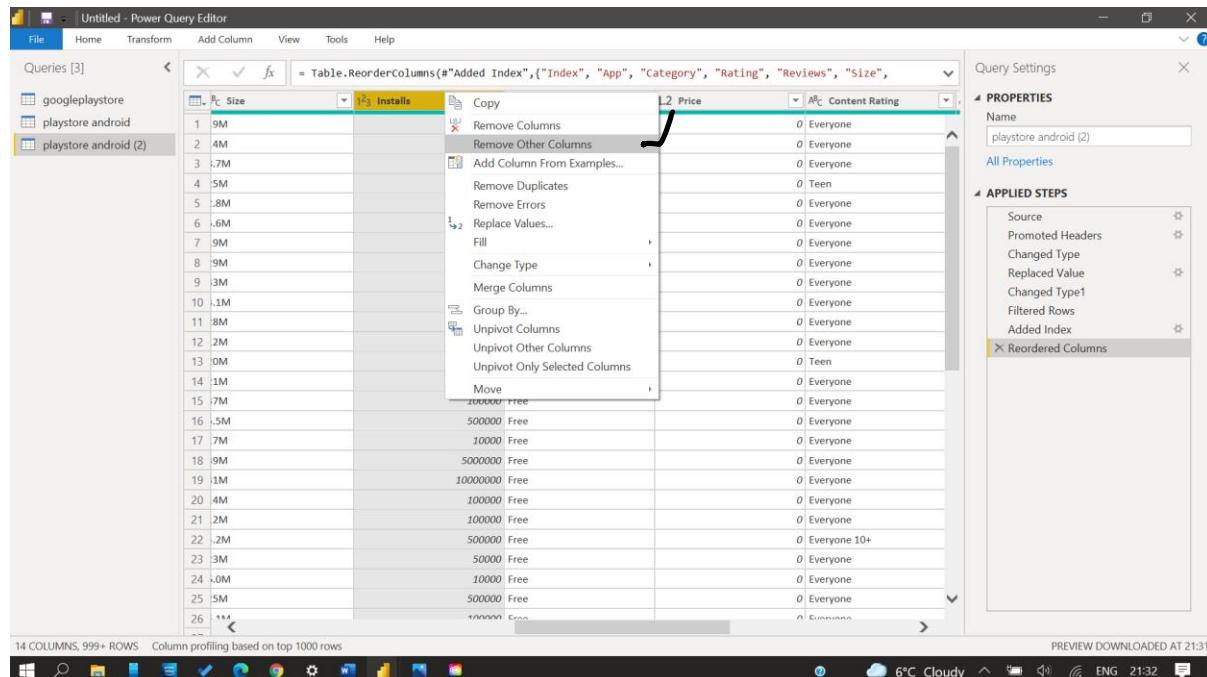


Fig 21- column removal in the duplicate query

## Creating Download Dimension Table

After the removal of unwanted column in the query the final table looks below in fig22 and rename it as the “download dimension table”

Index	App	Category	downloads
1	Photo Editor & Candy Camera & Grid & ScrapBook	ART_AND_DESIGN	10000
2	Coloring book moana	ART_AND_DESIGN	50000
3	U Launcher Lite – FREE Live Cool Themes, Hide Apps	ART_AND_DESIGN	500000
4	Sketch - Draw & Paint	ART_AND_DESIGN	5000000
5	Pixel Draw - Number Art Coloring Book	ART_AND_DESIGN	100000
6	Paper Flowers Instructions	ART_AND_DESIGN	50000
7	Smoke Effect Photo Maker - Smoke Editor	ART_AND_DESIGN	50000
8	Infinite Painter	ART_AND_DESIGN	1000000
9	Garden Coloring Book	ART_AND_DESIGN	1000000
10	Kids Paint Free - Drawing Fun	ART_AND_DESIGN	10000
11	Text on Photo - Fontee	ART_AND_DESIGN	1000000
12	Name Art Photo Editor - Focus n Filters	ART_AND_DESIGN	1000000
13	Tattoo Name On My Photo Editor	ART_AND_DESIGN	10000000
14	Mandala Coloring Book	ART_AND_DESIGN	100000
15	3D Color Pixel by Number - Sandbox Art Coloring	ART_AND_DESIGN	100000
16	Photo Designer - Write your name with shapes	ART_AND_DESIGN	500000
17	350 DIY Room Decor Ideas	ART_AND_DESIGN	10000
18	FlipClip - Cartoon animation	ART_AND_DESIGN	5000000
19	ibis Paint X	ART_AND_DESIGN	10000000
20	Logo Maker - Small Business	ART_AND_DESIGN	100000
21	Boys Photo Editor - Six Pack & Men's Suit	ART_AND_DESIGN	100000
22	Superheroes Wallpapers   4K Backgrounds	ART_AND_DESIGN	50000
23	HD Mickey Minnie Wallpapers	ART_AND_DESIGN	50000
24	Harley Quinn Wallpapers HD	ART_AND_DESIGN	10000
25	Colorfit - Drawing & Coloring	ART_AND_DESIGN	500000
26	Animated Photo Editor	ART_AND_DESIGN	1000000

Fig 22 – Download Dimension Table

## Creating rating Dimension Table

Similarly create a duplicate query[3] from the “play store android” query and select the required column shown in the fig23 and remove other columns and rename the query as “rating dimension table” and table looks same as shown below

Index	App	Rating
1	Photo Editor & Candy Camera & Grid & ScrapBook	4.1
2	Coloring book moana	3.9
3	U Launcher Lite – FREE Live Cool Themes, Hide Apps	4.7
4	Sketch - Draw & Paint	4.5
5	Pixel Draw - Number Art Coloring Book	4.3
6	Paper Flowers Instructions	4.4
7	Smoke Effect Photo Maker - Smoke Editor	3.8
8	Infinite Painter	4.1
9	Garden Coloring Book	4.4
10	Kids Paint Free - Drawing Fun	4.7
11	Text on Photo - Fontee	4.4
12	Name Art Photo Editor - Focus n Filters	4.4
13	Tattoo Name On My Photo Editor	4.2
14	Mandala Coloring Book	4.6
15	3D Color Pixel by Number - Sandbox Art Coloring	4.4
16	Photo Designer - Write your name with shapes	4.7
17	350 DIY Room Decor Ideas	4.5
18	FlipClip - Cartoon animation	4.3
19	ibis Paint X	4.6
20	Logo Maker - Small Business	4
21	Boys Photo Editor - Six Pack & Men's Suit	4.1
22	Superheroes Wallpapers   4K Backgrounds	4.7
23	HD Mickey Minnie Wallpapers	4.7
24	Harley Quinn Wallpapers HD	4.8
25	Colorfit - Drawing & Coloring	4.7
26	Animated Photo Editor	4.1

Fig 23 -Rating dimension Table

## Creating size dimension table

Similar to Rating dimension Table above create a duplicate and rename it as “size dimension table” and keep the columns that are needed as shown the fig24 below

The screenshot shows the Power Query Editor interface. The main area displays a table with two columns: 'Index' and 'App'. The 'Index' column contains values from 1 to 26. The 'App' column lists various mobile applications. To the right, the 'Query Settings' pane shows the 'Name' as 'size dimension table'. The 'APPLIED STEPS' pane details the process: 'Reordered Columns' (Index, App) and 'Removed Other Columns' (Rating, Current Ver, Android Ver).

Fig 24 – Size dimension table

## Creating version dimension table

Version dimension table is created from the duplicate of play store query and then removed unwanted columns and final table looks as shown in the fig25

The screenshot shows the Power Query Editor interface. The main area displays a table with four columns: 'Index', 'App', 'Current Ver', and 'Android Ver'. The 'Index' column contains values from 1 to 26. The 'App' column lists various mobile applications. The 'Current Ver' and 'Android Ver' columns provide specific version information for each app. To the right, the 'Query Settings' pane shows the 'Name' as 'version dimension Table'. The 'APPLIED STEPS' pane details the process: 'Reordered Columns' (Index, App, Current Ver, Android Ver) and 'Removed Other Columns' (Rating, Price, Date).

Fig25 – version dimension table

## Dax expression

we use “data analytics expression” as Dax to create a new column in the price dimension table for this click on the new column as shown in the fig26

The screenshot shows the Power BI Desktop interface with the 'Table tools' tab selected. A tooltip in the center of the screen says "Write a DAX expression that creates a new column in the selected table and calculates values for each row." The Fields pane on the right lists columns such as App, Index, Sum Installs, Price, Type, Rating, and others. The main table view shows a list of apps with columns for App, Type, Price, and Installs.

Fig 26 – creating new column

A new column is generated and rename this column as “revenue through paid apps” and use the formula as shown in the below fig27 and then press tick mark shown in the left a new column is formed with the help of Dax

The screenshot shows the Power BI Desktop interface with the 'Column tools' tab selected. A new column named 'revenue through paid apps' is being created with the formula '= 'price dimension Table'[Installs]\*'price dimension Table'[Price]'. The Fields pane on the right shows the new column along with other columns like App, Index, Sum Installs, Price, Type, Rating, and others. The main table view shows a list of apps with columns for App, Type, Price, and the newly created 'revenue through paid apps' column.

Fig27a- column with help of Dax

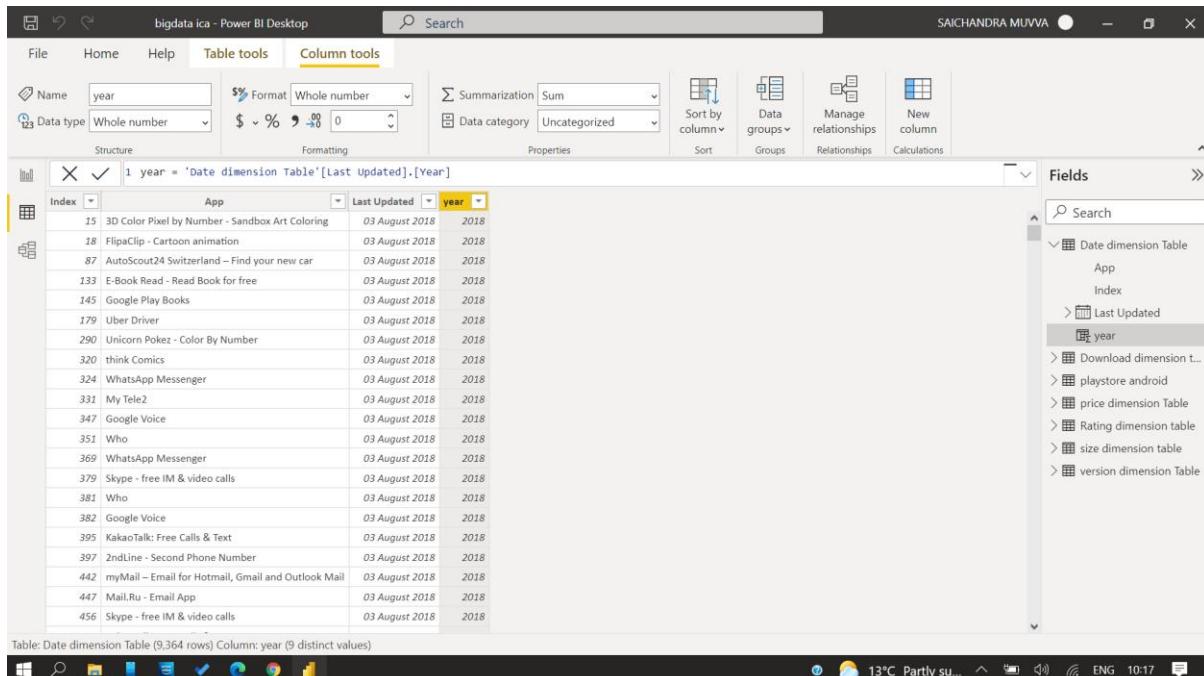


Fig 27 b – year column using Dax

In date dimension table for attaining the year column we used Dax pression above in fig 27b and rename the column as year

M language

In size dimension table we seen size of megabyte as “M” and kilobyte as “k” the actual notation is “MB”, “KB” so change it with the help of M language [5] For this open power query editor and select “view” and then select “Advanced editor” as shown in the fig28



Fig29 – Advanced editor

A pop-up window is opened as shown in the fig29 m lang code is written there and hence M is changed as “MB” and K as “KB” as shown in the below fig30

The screenshot shows the Power Query Editor interface with the following details:

- File Bar:** Home, Transform, Add Column, View, Tools, Help.
- Queries List:** playstore android, Download dimension ta..., Rating dimension table, size dimension table (selected), price dimension Table, Date dimension Table, version dimension Table.
- Table View:** A table titled "size dimension table" with columns "Index" and "Size". The table contains 26 rows of data, such as "Photo Editor & Candy Camera & Grid & ScrapBook" (19MB) and "Garden Coloring Book" (33MB).
- Query Settings Panel:**
  - PROPERTIES:** Name: size dimension table.
  - APPLIED STEPS:** A list of steps including "Source", "Promoted Headers", "Changed Type", "Replaced Value", etc., with "Sorted Rows" highlighted.
- Bottom Status:** PREVIEW DOWNLOADED ON TUESDAY, system icons for weather, battery, and time.

Fig30 – size table after M Lang

### Creating model and relationships

After creating the tables, we need to attain the relationship between the tables to play store android to all other dimension tables so click the below one shown in the fig 31 so the models are loaded with dimension tables and click on the manage relationships as shown on the below

The screenshot shows the Power BI Desktop interface with the following details:

- File Bar:** File, Home, Help.
- Get Data:** Get data (dropdown), Refresh.
- Manage Relationships:** A button circled in yellow.
- Model Data View:** A grid showing relationships between tables:
  - playstore android (parent table)
  - Download dimension table
  - Date dimension Table
  - price dimension Table
  - Rating dimension table (parent table)
  - size dimension table (parent table)
  - version dimension Ta... (parent table)
- Bottom Buttons:** All tables (highlighted in yellow), +, Publish.
- Bottom Status:** 5°C Cloudy, system icons for battery, signal, and time.

Fig31 – model data view

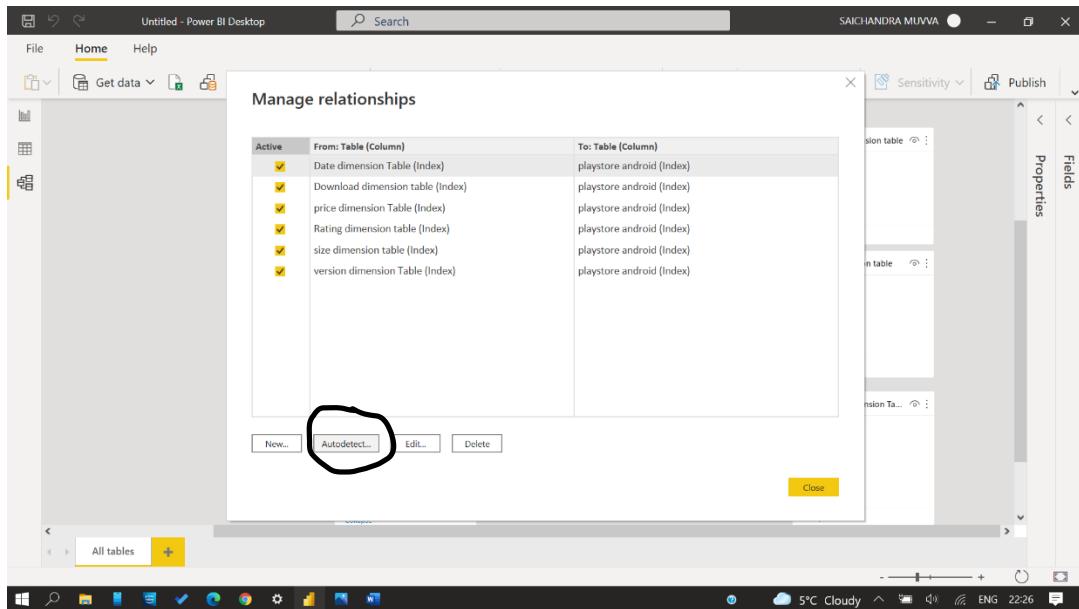


Fig32 – Manage relationships between models

In fig 32 we can see the auto detect after selecting the option we relationships are formed

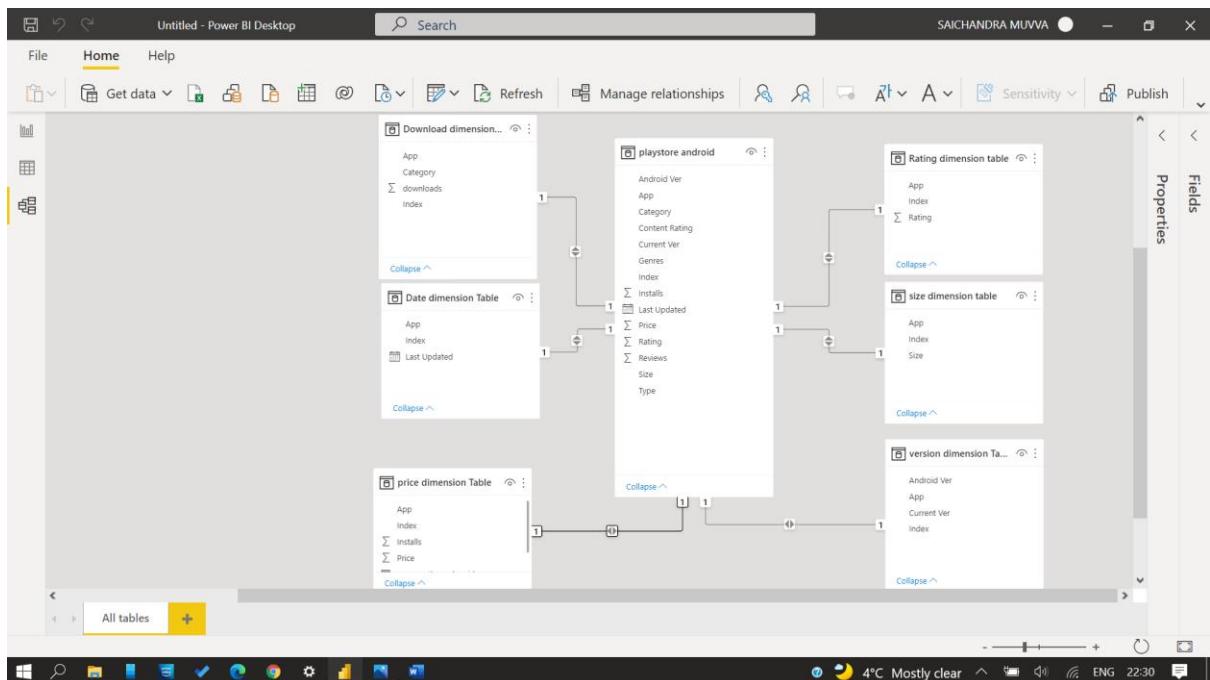


Fig33 – relationships are formed from one to another model

After selecting the auto detect models are in one-to-one relationship [6]and configured of both the side as shown in the below fig34

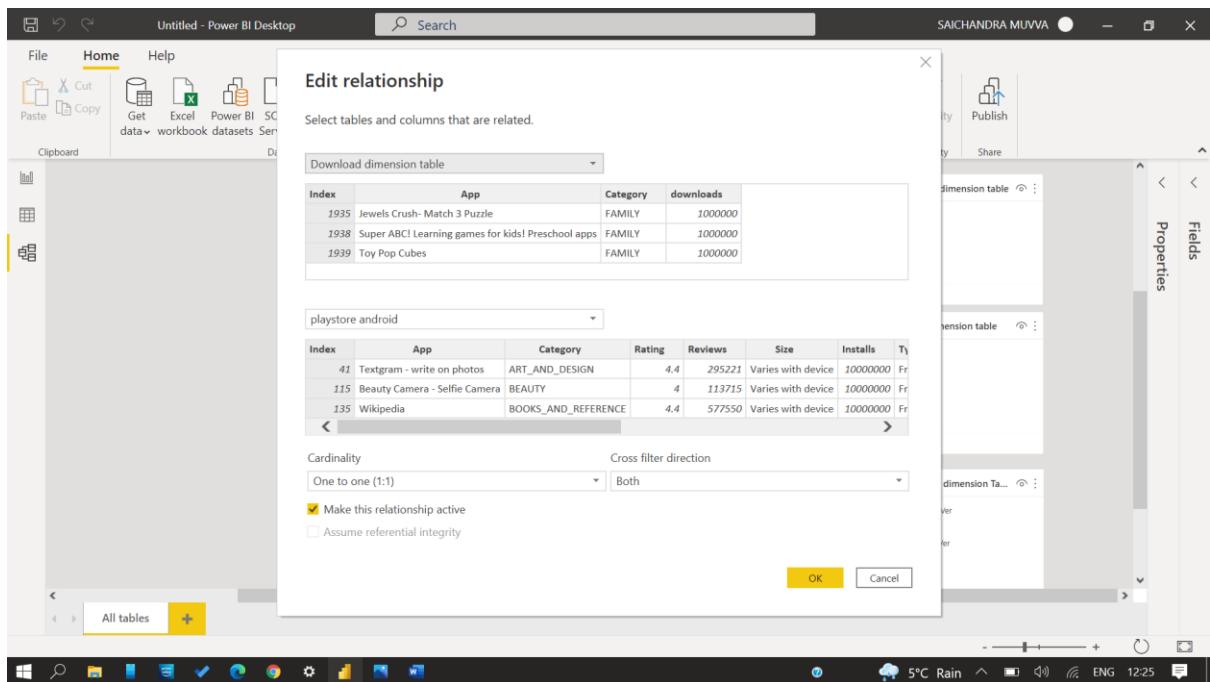


Fig34 – Relationship status between models



## BUSINESS INTELLIGENCE FOR PLAYSTORE APP ANALYSIS

CIS4008-N-BF1-2020

SCHOOL OF COMPUTING, ENGINEERING AND DIGITAL TECHNOLOGY

### **Section -2**

### **BUSINESS REPORT**



**12 January 2022**

Prepared by

W9533735, SAICHANDRA MUVVA

## EXECUTIVE SUMMARY

### Background

This data is from play store from 2013 to 2018 it compromises of parameters affecting the app popularity and revenue and this data set is taken from the Kaggle with the help of power bi we have analysed the key parameters effecting the app and we have some insights of this data so that it could be helpful in building the better

### Key findings

- Paid App of Minecraft that generated the highest revenue from play store
- From play store \$344M is paid to play store paid apps
- Android version influences the most downloads and least influence is type
- With the increase from 2013 to 2018 we can forecast lift in increase of downloads
- Subway surfers is the most downloaded app with highest rating
- Communication genre has the greatest number of downloads
- People show less interest in education apps
- Content rating with everyone shows more interested and mature 17+ has less popularity
- Rating with everyone has good feedback and adult apps have poor feedback

### Recommendations

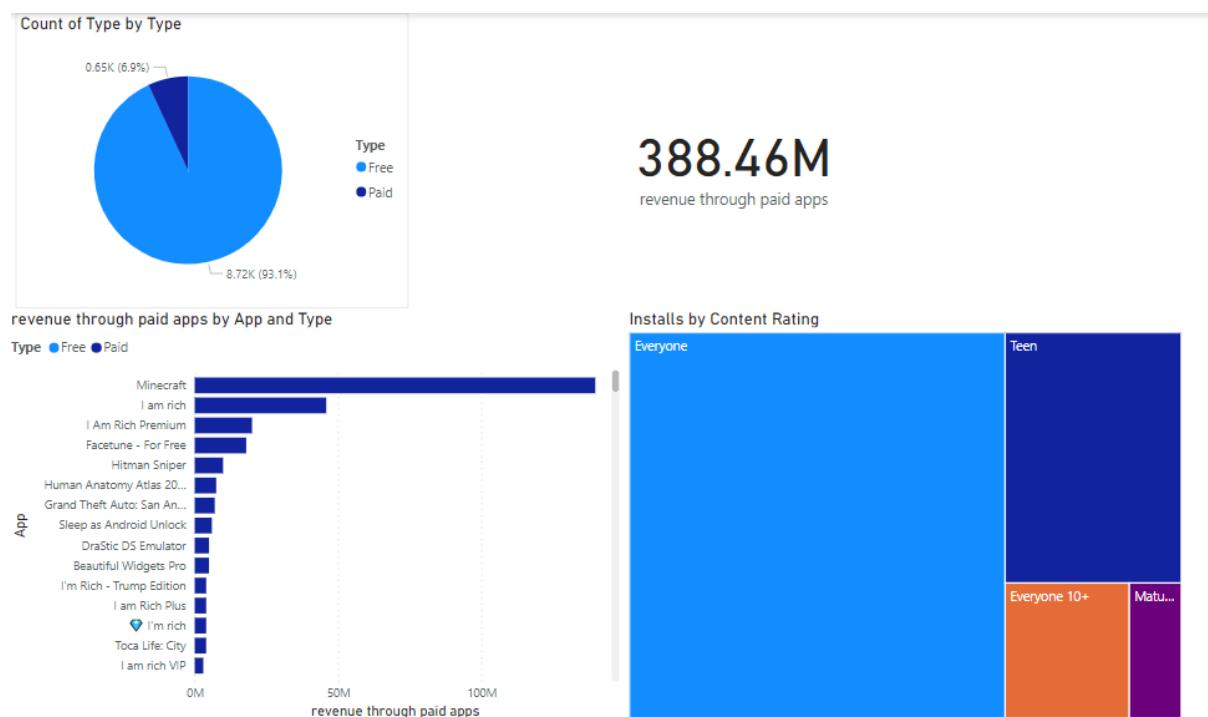
- Apps of education are built with less standards so increase the quality of apps
- Android version of varies with device has most key influence factor so try to develop the apps in the greatest number of compatible phones
- Money is mostly generated from gaming apps in play store
- We can clearly see the forecast of increment of downloads may be doubled by 2030
- Mature 17+ apps are built with better standard because they have less rating
- People shows adult apps have most disliked try to remove these apps

This report addresses the following questions

- Distribute of app over categories and installs?
- What is the Most Usage group using the application and their category?
- What is the Costliest app used in play store?
- What is the Most Popular app Used in the play store?
- Which category of apps are mostly used?
- Is there any interlink between Rating and installs?
- Is there any Relationship between Reviews, Genre, Type And category?
- Key factors Influencing installs, Rating and Price?

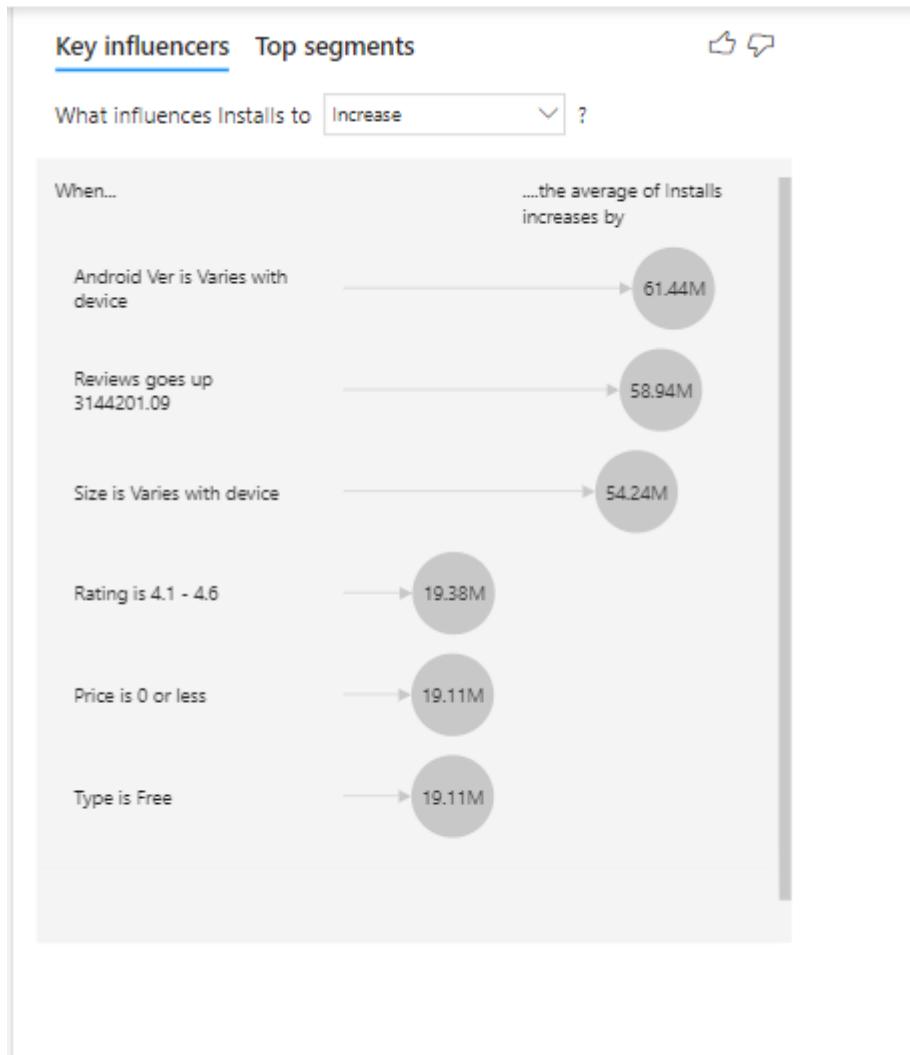
### Findings based on Analysis and evaluation

Below visual shows the more number of apps in the play store are free and 6.9% apps in the play store are paid apps and 93.1% apps are free apps and total revenue generated in the play store is 388.46 million usd and most revenue is generated by Minecraft app and it's a gaming app it is the highest money generated app from play store app followed by the I am rich app



Some of the factors influencing the conditions of downloads from the play store app are seen below fig are version, reviews, size, rating, price, type these influences the downloads increase if the android version is compatible with device increases downloads increases and reviews increases then downloads of app increases similarly rating is from 4.1 to

4.6 then downloads increases and price with free similar incline in downloads and free type of apps have higher downloads than paid apps



The below fig shows the apps with higher price “I am rich” is the app having higher price in the whole play store app and apps with lower price rather than zero is 0.99 usd

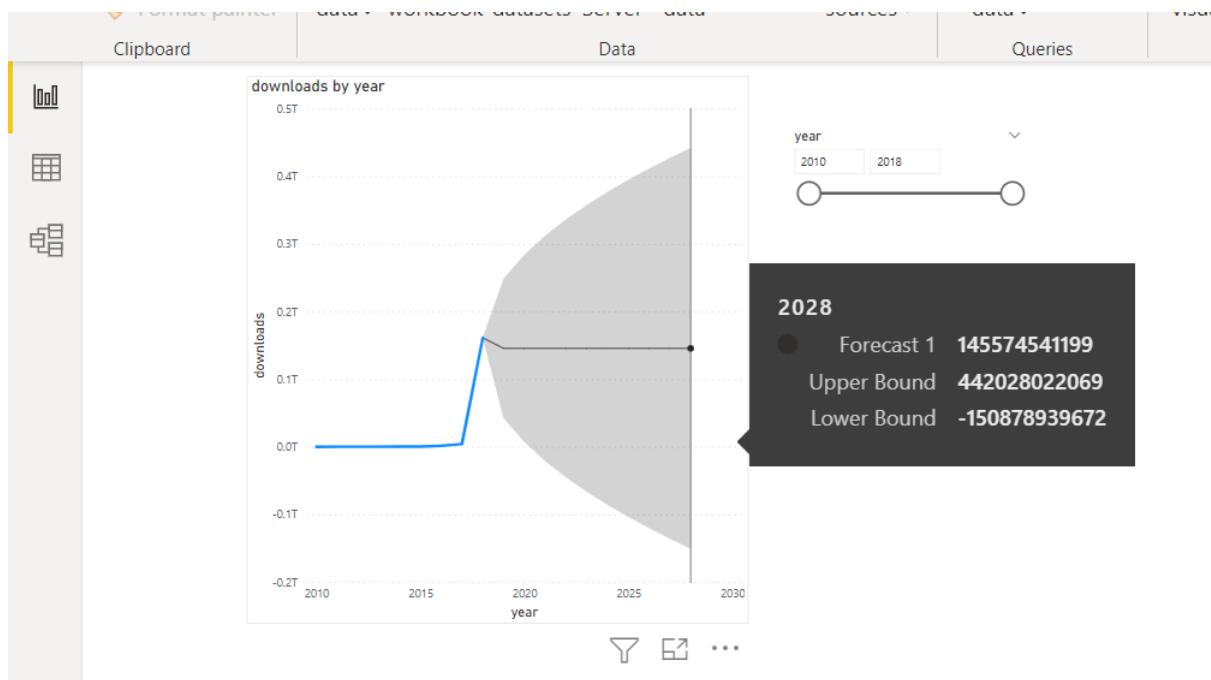
The image shows two separate screenshots of mobile app store interfaces. The left screenshot displays a list of apps with their prices, including 'I am rich' at \$1,189.97 and 'I'm Rich - Trump Edition' at \$400.00. The right screenshot shows a similar list of apps, with the total price for all items being \$8,999.06.

App	Price
I am rich	1,189.97
I'm Rich - Trump Edition	400.00
I am rich (Most expensive app)	399.99
I am Rich Plus	399.99
I Am Rich Premium	399.99
I Am Rich Pro	399.99
I AM RICH PRO PLUS	399.99
I am Rich!	399.99
I am rich(premium)	399.99
most expensive app (H)	399.99
⌚ I'm rich	399.99
I am extremely Rich	379.99
I am rich VIP	299.99
Vargo Anesthesia Mega App	159.98
Human Anatomy Atlas 2018: Complete 3D Human Body	74.97
A Manual of Acupuncture	67.98
EMT PASS	59.98
PTA Content Master	59.98
Muscle Premium - Human Anatomy, Kinesiology, Bones	49.98
LTC AS Legal	39.99
Hospitalist Handbook	39.98
I am Rich Person	37.99
2017 EMRA Antibiotic Guide	33.98
AP Art History Flashcards	29.99
<b>Total</b>	<b>8,999.06</b>

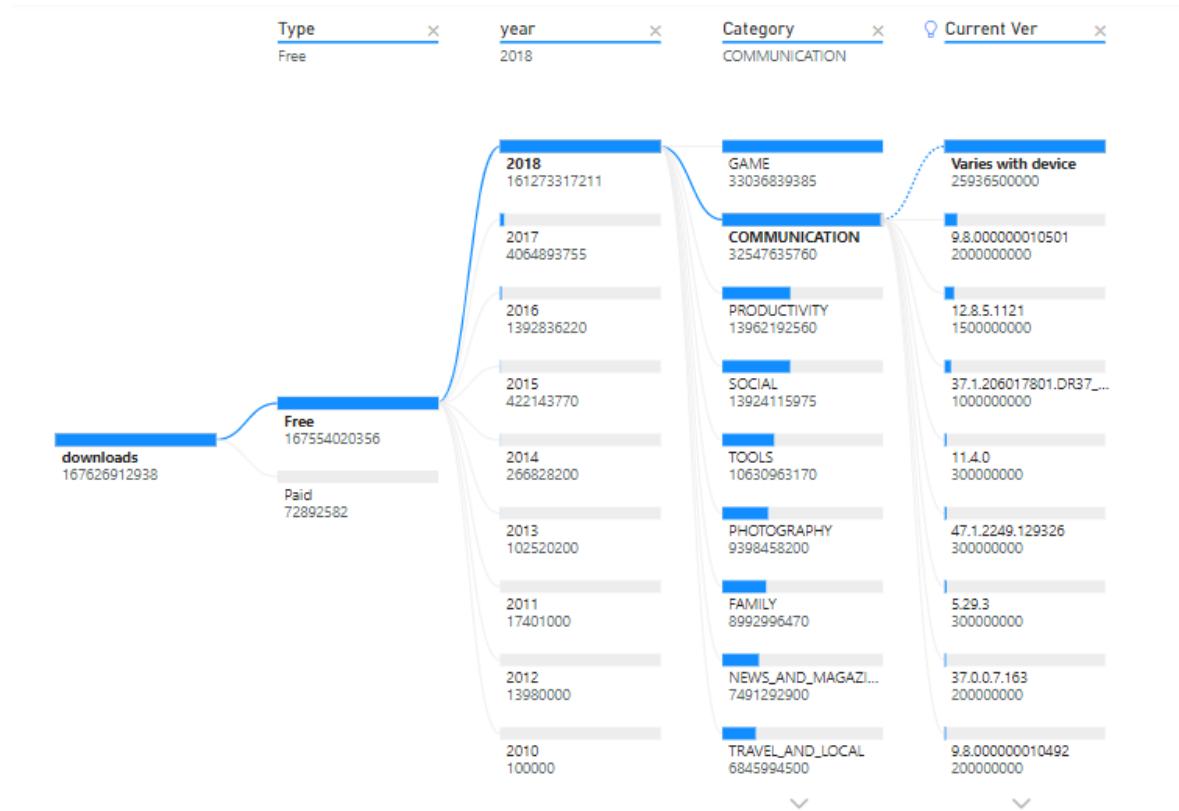
  

App	Price
[Sub/EMUI] P Pro - EMUI 8.1/8.0/5.X Theme	0.99
10,000 Quotes DB (Premium)	0.99
211CK	0.99
A41 WatchFace for Android Wear Smart Watch	0.99
AG Subway Simulator Mobile	0.99
AJ Blue Icon Pack	0.99
AJ Gray Dark Icon Pack	0.99
AJ Gray Icon Pack	0.99
All-in-One Mahjong 3	0.99
An Elite Warrior Ex	0.99
Android P Style Icon Pack	0.99
Antenna Tool Premium	0.99
B-52 Spirits of Glory Deluxe	0.99
Basket Manager 2016 Pro	0.99
Basket Manager 2017 Pro	0.99
BringGo AU & NZ	0.99
CA Speakers	0.99
CB Frequencies	0.99
CE Map - Interactive Conan Exiles Map	0.99
Chinese Chess 3D Online (Xiangqi, 象棋, co tuong)	0.99
Chronolink DX	0.99
Cl 174 Gray Icon Pack	0.99
Circle Colors Pack-FN Theme	0.99
Color Changer Pro [root]	0.99
<b>Total</b>	<b>8,999.06</b>

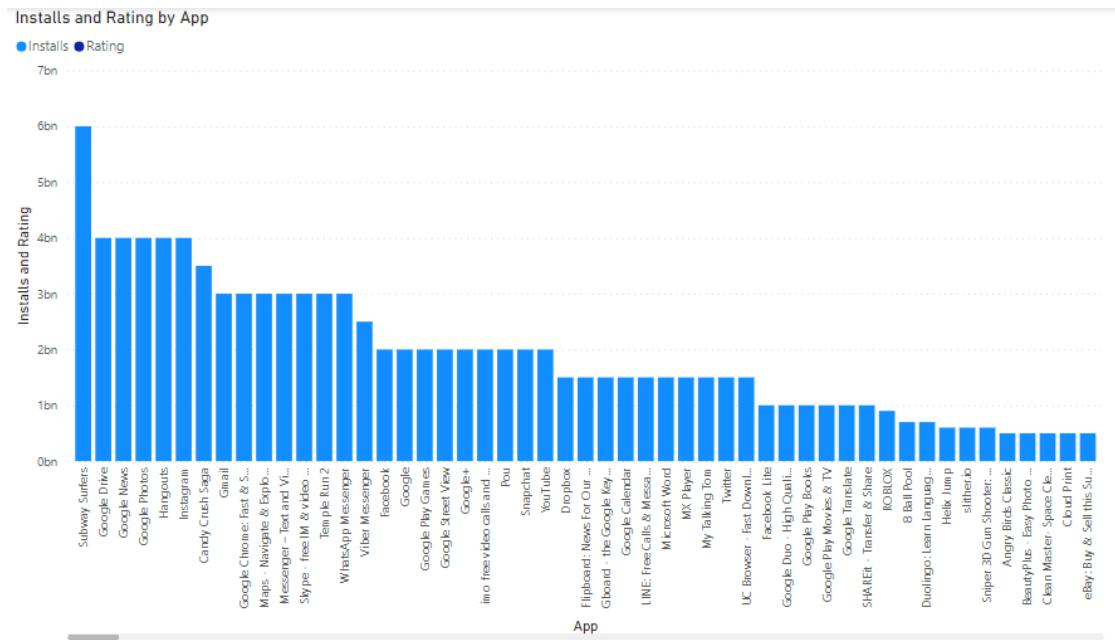
The below visual is about to forecast the downloads counts from 2020 to 2028 and gradual increase of download count in the next 10 years and the data model is trained from the 2010 to 2018 and forecast model is from 2018 to 2028



The below fig is about the decomposition tree of downloads on what basis app downloads are done and below figure illustrates the division of downloads done on their basis. it gives the clear view on the downloads and the greatest number of apps are downloaded in the 2018 a gradual increase from 2010-18 and gaming apps are mostly downloaded followed by the communication similarly most apps are in the version of compatible with the device



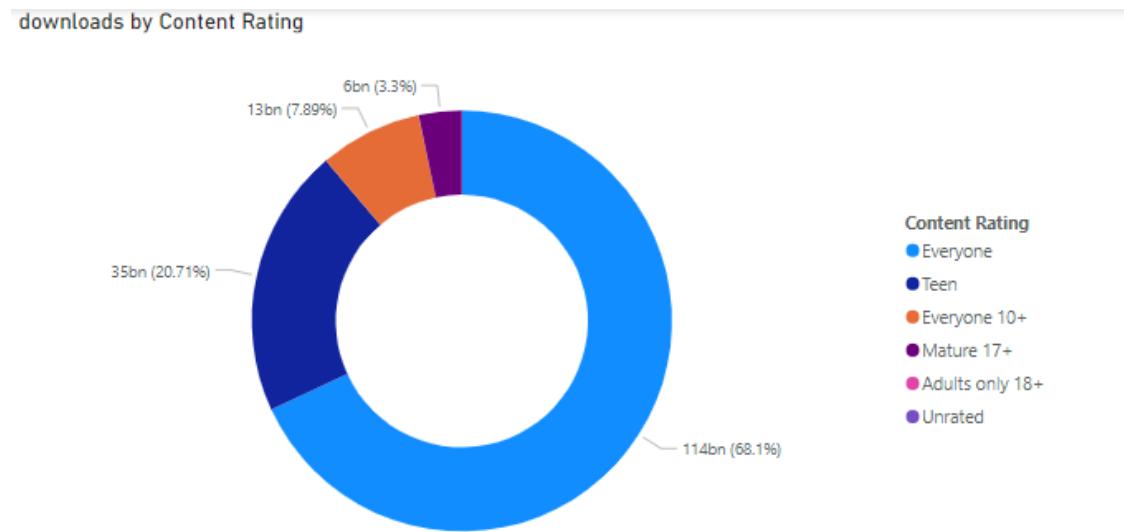
Most installs done with the better rating in the play store is subway surfers app followed by google drive this shows people are favour over this game and as most liked app maximum number of users



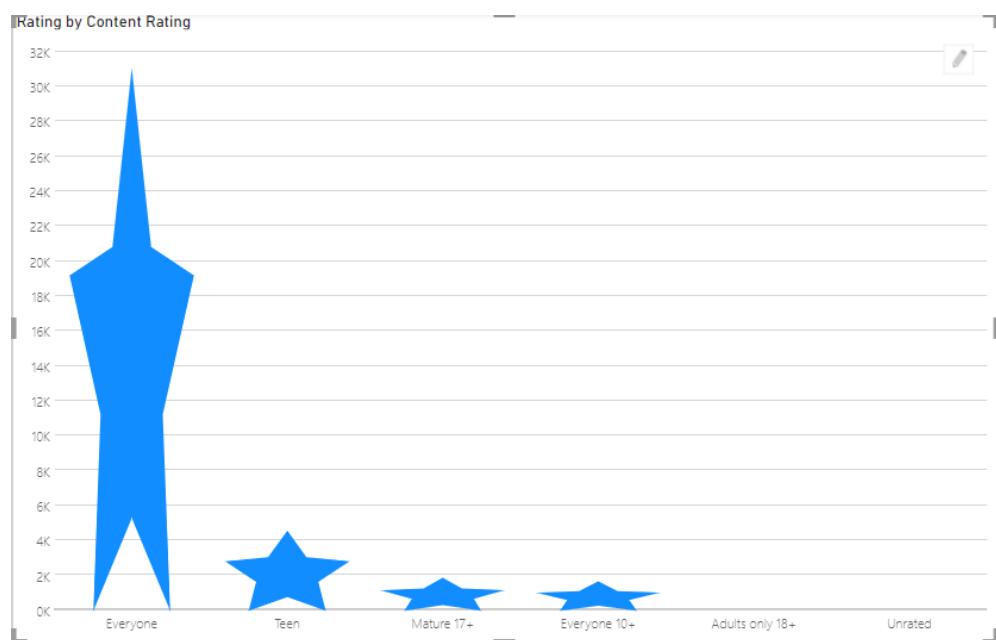
Communication genre shown in the below diagram gives the clear evidence that most number of downloads are done in this criteria followed by productivity and social genre and pretend play genre is the less downloaded genre apps and followed by education genre app by this evidence we have a clear sight of people shows less interest in the education apps and pretend play and quality in the education and pretend play genre must be increased

Genres	downloads	Genres	downloads
Communication	32647241530	Board;Pretend Play	100
Productivity	14176070180	Books & Reference;Education	11000
Social	14069841475	Comics;Creativity	50000
Tools	11440234500	Adventure;Brain Games	100000
Arcade	10727077727	Card;Brain Games	100000
Photography	10088243130	Health & Fitness;Education	100000
Casual	9662778970	Lifestyle;Education	100000
Action	9342037470	Puzzle;Education	100000
News & Magazines	7496210650	Travel & Local;Action & Adventure	100000
Travel & Local	6868759300	Communication;Creativity	500000
Video Players & Editors	6206897200	Music & Audio;Music & Video	500000
Entertainment	5064124820	Strategy;Education	500000
Shopping	3247831540	Simulation;Education	600500
Sports	2867131465	Arcade;Pretend Play	1000000
Personalization	2325352930	Art & Design;Pretend Play	1000000
Racing	2000246020	Entertainment;Education	1000000
Strategy	1974137500	Health & Fitness;Action & Adventure	1000000
Books & Reference	1921291655	Parenting;Brain Games	1000000
Puzzle	1668421070	Racing;Pretend Play	1000000
Health & Fitness	1583056220	Strategy;Creativity	1000000
Business	1001768120	Strategy;Action & Adventure	1100000
Adventure;Action & Adventure	906150000	Puzzle;Creativity	1500000
Finance	876612400	Casual;Education	2010000
Simulation	814469810	Simulation;Pretend Play	2110000
Education;Education	778013560	Art & Design;Creativity	2210000
Maps & Navigation	724267560	Parenting;Education	3000000
Lifestyle	527561120	Education;Action & Adventure	4020000
Education	500767210	Entertainment;Pretend Play	6000000
Casual;Action & Adventure	500001000	Board;Action & Adventure	6005000
Weather	426096500	Parenting;Music & Video	6710000
Adventure	406695300	Entertainment;Action & Adventure	7000000
Total	167626912938	Total	167626912938

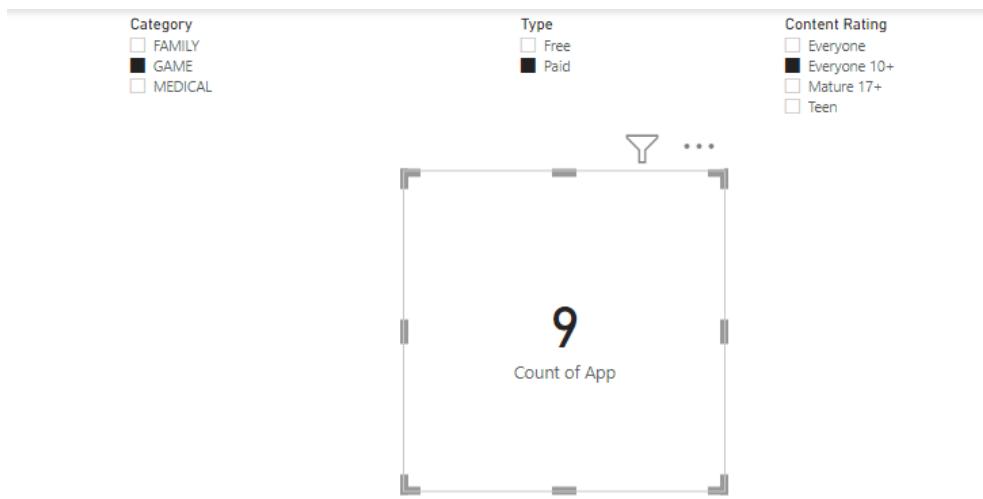
App are built with content rating this check has been done by the play store app and there are 6 categories in that they are everyone, teen, everyone 10+, mature 17+, adults and unrated below visual shows most of the apps in the play store are built and downloaded for everyone category and then teen, followed by 10+ and unrated majority of 68.1% of apps are downloaded in everyone category and 3.3% apps are unrated by comparing with this adult apps are negligible in the below diagram



We used infographic designer for the below diagram and this graph is about the content rating genre with the ratings on app in the play store it compromises of 6 kinds of data and but majority of users are rated on 4 content rating genre apps they are everyone, teen , everyone 10+, everyone 17+ people shows no interest on remaining genre apps and content rating with everyone has better rating apps followed by teen, everyone 10+,17+



## Use of slicers



Slicers used in the visualization we can slice the data into categories and the above visualization gives a overview that apps are categorized into type, category, content rating by these filters we can receive the number of apps in the playstore with these qualities slicers can be used as filters to the required data from the table

## References

- [1] Richterich, Annika. "Introduction." *The Big Data Agenda: Data Ethics and Critical Data Studies*, vol. 6, University of Westminster Press, 2018, pp. 1–14,  
<https://doi.org/10.2307/j.ctv5vddsw>.
- [2] J. Patel, "An Effective and Scalable Data Modeling for Enterprise Big Data Platform," 2019 IEEE International Conference on Big Data (Big Data), 2019, pp. 2691-2697, doi: 10.1109/BigData47090.2019.9005614.
- [3] Bhadani, A., Jothimani, D. (2016), Big data: Challenges, opportunities and realities, In Singh, M.K., & Kumar, D.G. (Eds.), Effective Big Data Management and Opportunities for Implementation (pp. 1-24), Pennsylvania, USA, IGI Global
- [4] Duncan, O., & Kinsman, M. (2018, October 17). Power Query M reference [Web site]. Retrieved from <https://docs.microsoft.com/en-us/powerquery-m/power-query-m-reference>
- [5] Kline, K. (2014, August-September). PowerBI: The self-service BI app that can change your life. *Database Trends & Applications*
- [6] Simon, P. (2014). *The visual organization data visualization, big data, and the quest for better decisions (Wiley and SAS business series)*. Hoboken, NJ: John Wiley & Sons.