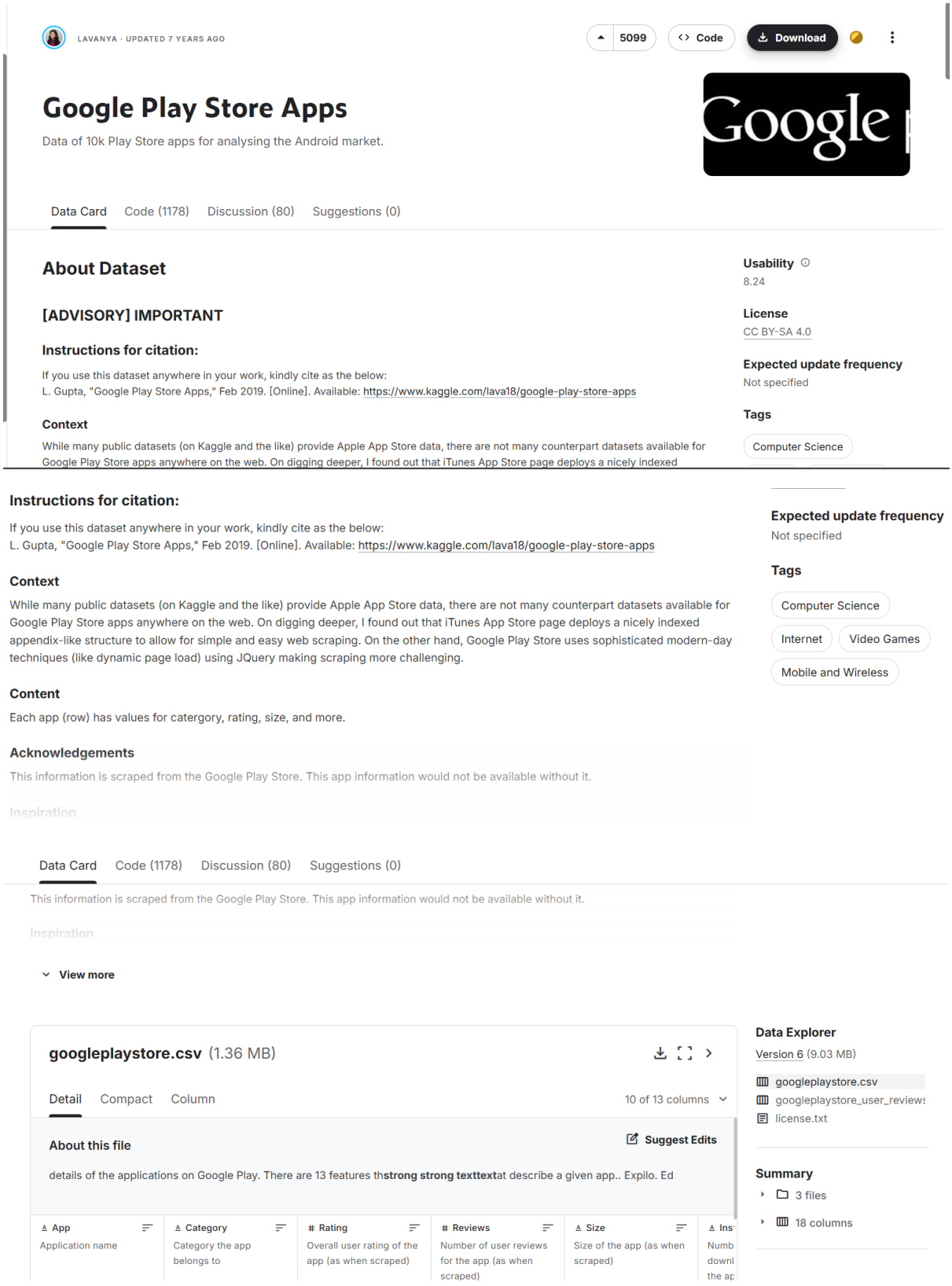
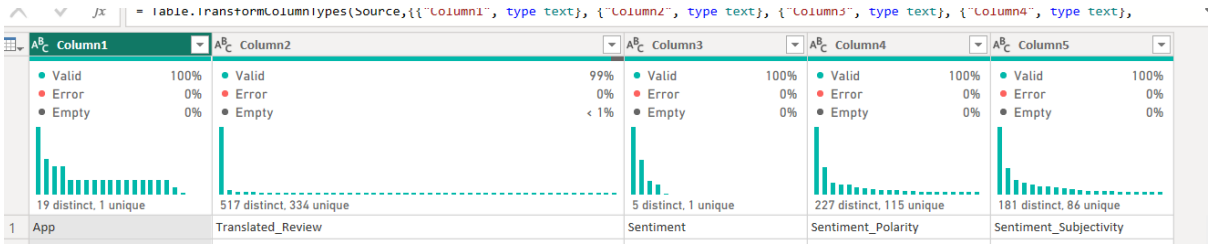


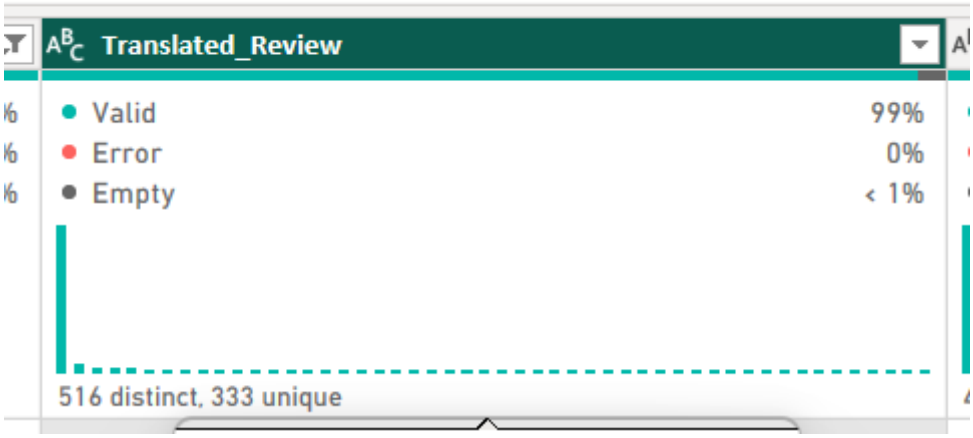
From the screenshots below I have applied column quality, and column distribution from the view tab on my dataset.

The first data quality mistake was that in the googleplaystore_user_reviews table, the column names were counted as row and the columns were named as numbers. I simply renamed the columns and removed the first row.

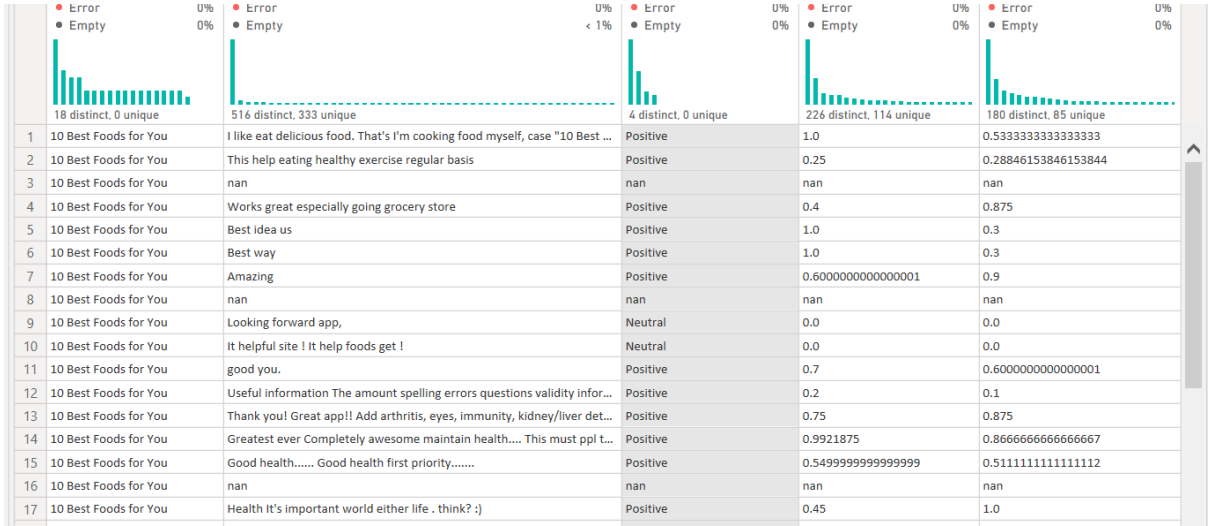


The second data quality issue was on the Translated_Review column that had some missing values that were removed. To fix this I removed I unchecked the (blank) on the filter.

```
SelectRows(#"Renamed Columns", each ([APP] <> "App"))
```

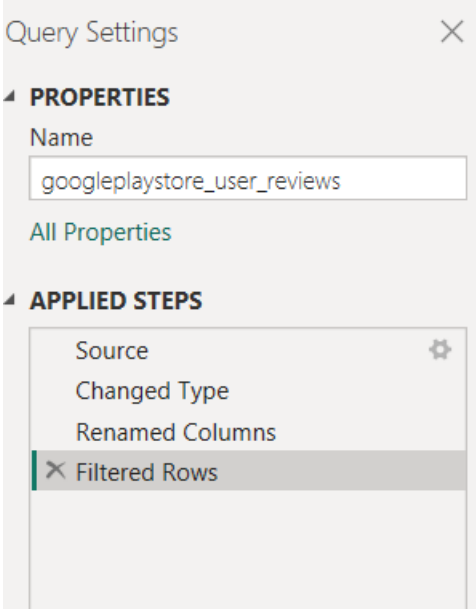


The third data quality issue was that there are rows marked as nan in the same table that had no use. To fix this I also unchecked nan value from one of the columns filter.

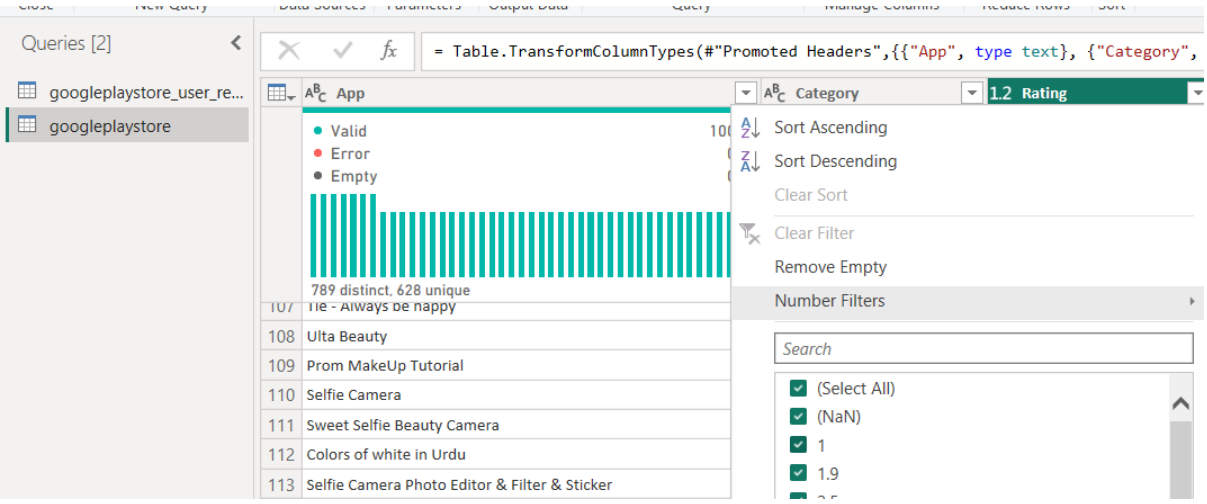


	Error 0% Empty 0%	Error 0% Empty < 1%	Error 0% Empty 0%	Error 0% Empty 0%	Error 0% Empty 0%
	18 distinct, 0 unique	516 distinct, 333 unique	4 distinct, 0 unique	226 distinct, 114 unique	180 distinct, 85 unique
1	10 Best Foods for You	I like eat delicious food. That's I'm cooking food myself, case "10 Best ...	Positive	1.0	0.5333333333333333
2	10 Best Foods for You	This help eating healthy exercise regular basis	Positive	0.25	0.28846153846153844
3	10 Best Foods for You	nan	nan	nan	nan
4	10 Best Foods for You	Works great especially going grocery store	Positive	0.4	0.875
5	10 Best Foods for You	Best idea us	Positive	1.0	0.3
6	10 Best Foods for You	Best way	Positive	1.0	0.3
7	10 Best Foods for You	Amazing	Positive	0.6000000000000001	0.9
8	10 Best Foods for You	nan	nan	nan	nan
9	10 Best Foods for You	Looking forward app,	Neutral	0.0	0.0
10	10 Best Foods for You	It helpful site ! It help foods get !	Neutral	0.0	0.0
11	10 Best Foods for You	good you.	Positive	0.7	0.6000000000000001
12	10 Best Foods for You	Useful information The amount spelling errors questions validity infor...	Positive	0.2	0.1
13	10 Best Foods for You	Thank you! Great app!! Add arthritis, eyes, immunity, kidney/liver det...	Positive	0.75	0.875
14	10 Best Foods for You	Greatest ever Completely awesome maintain health.... This must ppl t...	Positive	0.9921875	0.8666666666666667
15	10 Best Foods for You	Good health..... Good health first priority.....	Positive	0.5499999999999999	0.5111111111111112
16	10 Best Foods for You	nan	nan	nan	nan
17	10 Best Foods for You	Health It's important world either life . think? :)	Positive	0.45	1.0

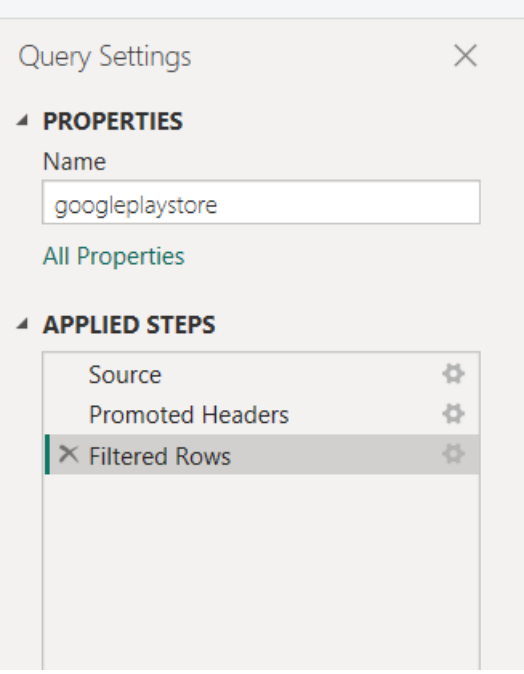
This is the applied steps for the issues above



Another data quality issue is on the table googleplaystore column Rating, where there are nan values as well. To fix this I did the same by unselecting (NaN).



Here are the applied steps for this tables issue

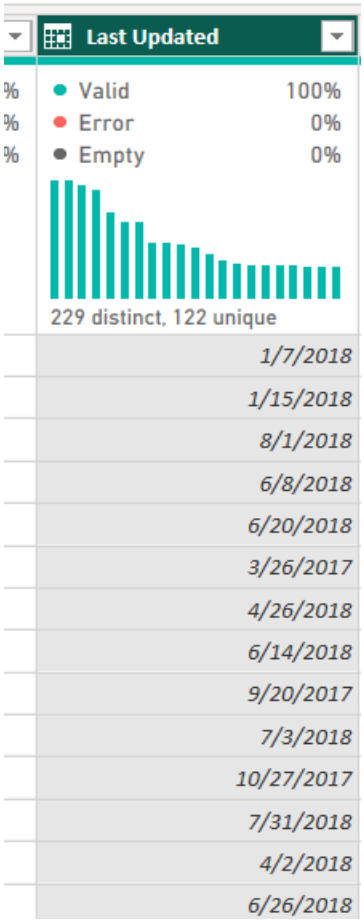


For q2

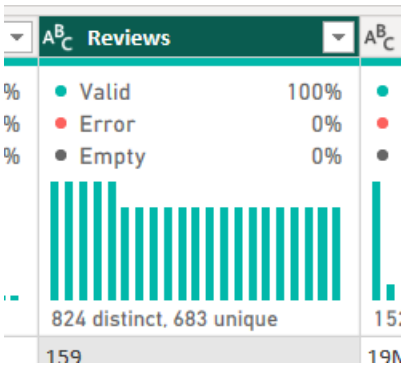
The date type was text and I changed it to date formart

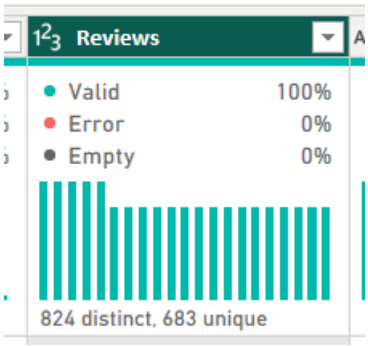
Last Updated	
0%	Valid 100%
0%	Error 0%
0%	Empty 0%
	229 distinct, 122 unique
	January 7, 2018
	January 15, 2018
	August 1, 2018
	June 8, 2018
	June 20, 2018
	March 26, 2017
	April 26, 2018
	June 14, 2018
	September 20, 2017
	July 3, 2018
	October 27, 2017
	July 31, 2018
	April 2, 2018
	June 26, 2018
	August 3, 2018
	June 6, 2018
	July 31, 2018
	November 7, 2017
	August 3, 2018
	July 30, 2018
	April 20, 2018
	March 20, 2018
	July 12, 2018

And here is after.

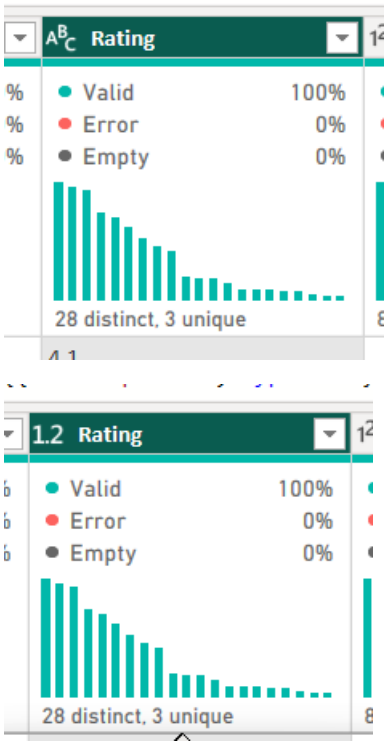


The column reviews was set as text as well and I changed it to a whole number, as the number of reviews made cannot be words. Here are the before and after.

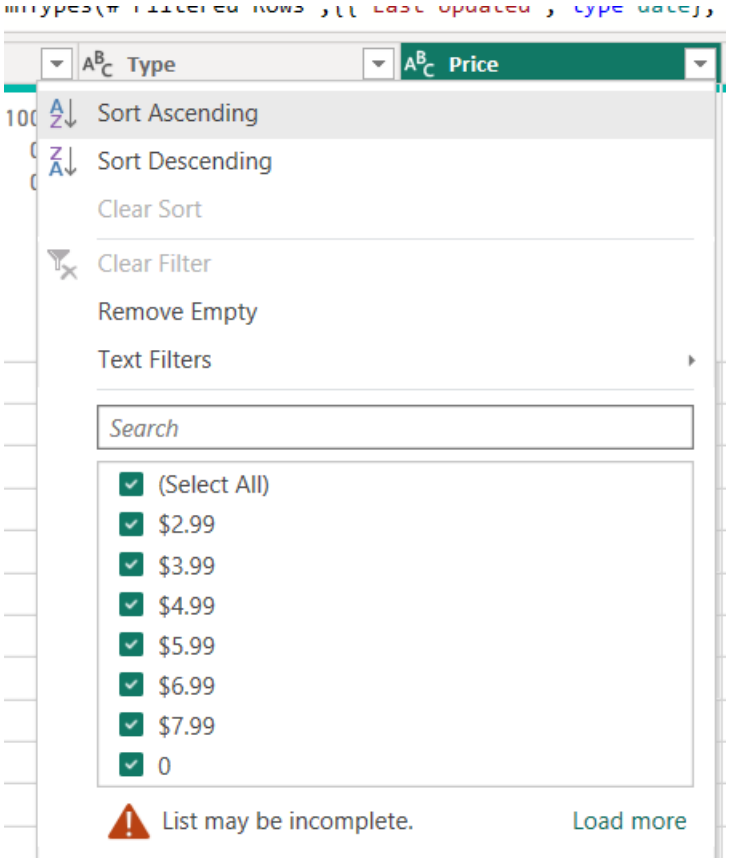


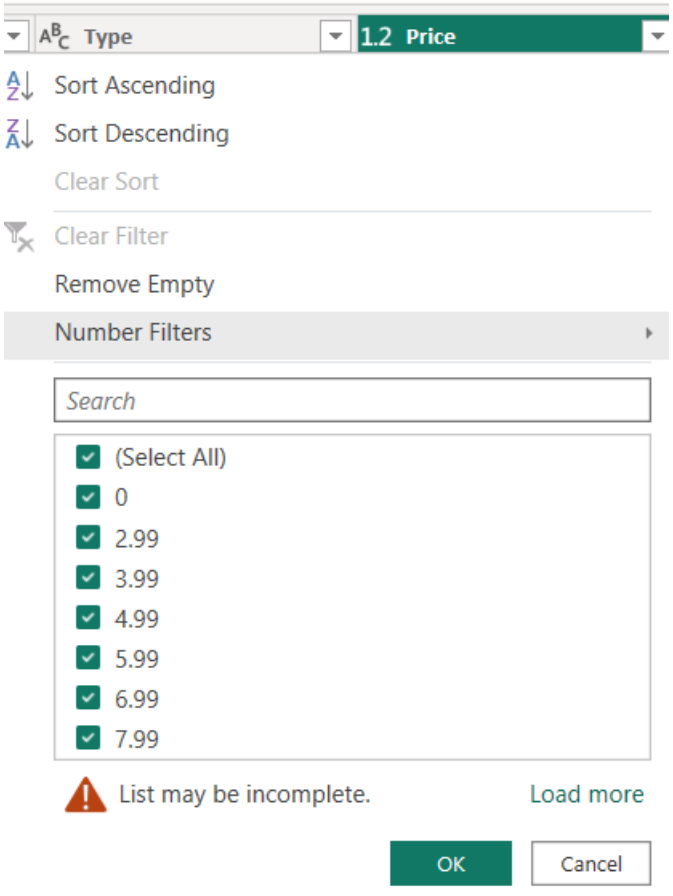


The values on rating column were also set to text as there were some values that were labled as NaN, I changed this to decimal number as it fits rating more. Here are the before and after.

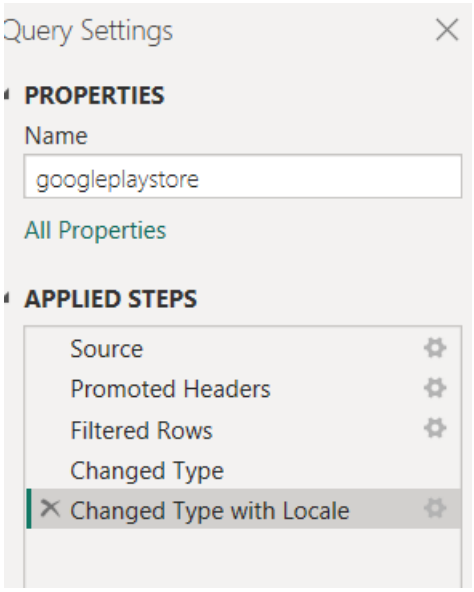


The values on the price column have the currency sign $ making them have text as their type. To fix this I used locale using the English(United States) to parse the $ sign. Here are the before and after.

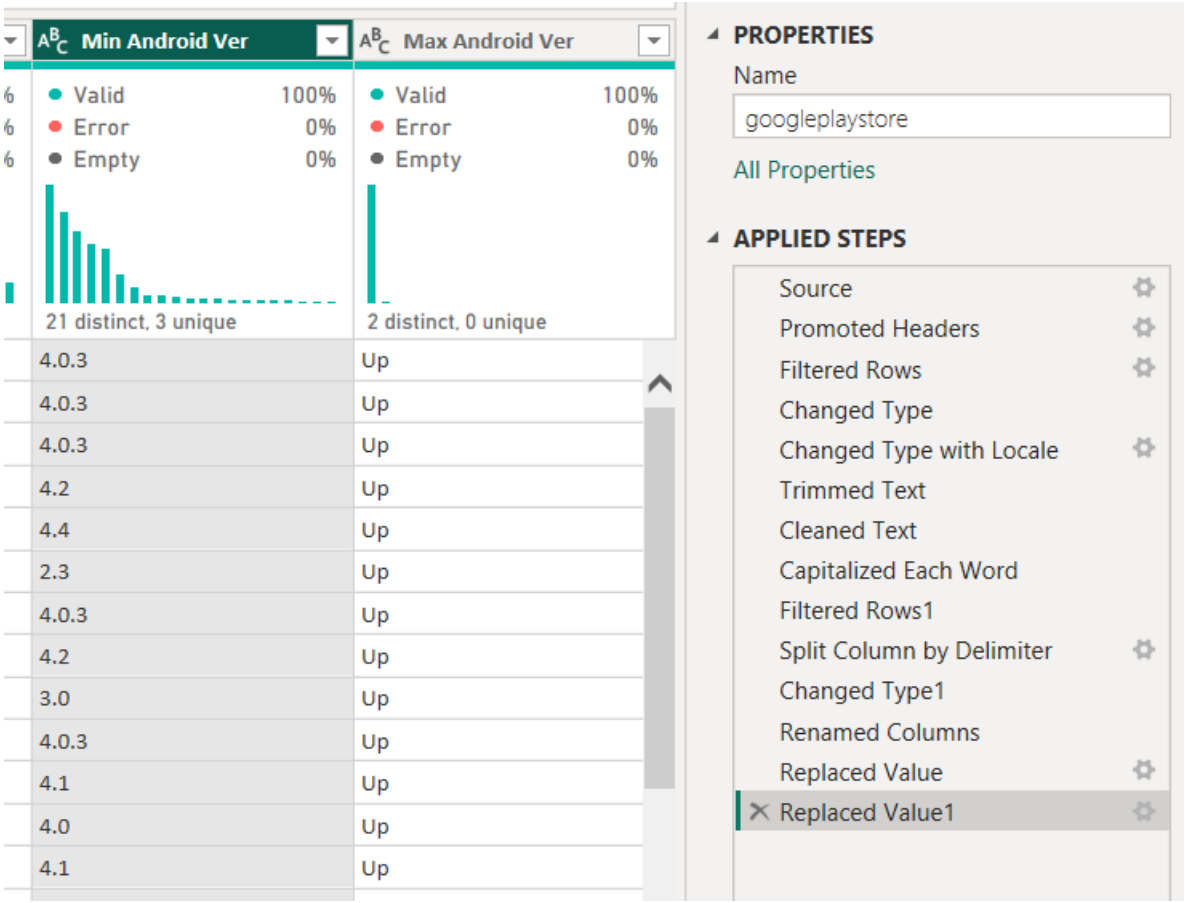

Type
Price
Sort Ascending
Sort Descending
Clear Sort
Clear Filter
Remove Empty
Text Filters
Search
(Select All)
$2.99
$3.99
$4.99
$5.99
$6.99
$7.99
0
List may be incomplete.
Load more


Type
1.2 Price
Sort Ascending
Sort Descending
Clear Sort
Clear Filter
Remove Empty
Number Filters
Search
(Select All)
0
2.99
3.99
4.99
5.99
6.99
7.99
List may be incomplete.
Load more
OK
Cancel

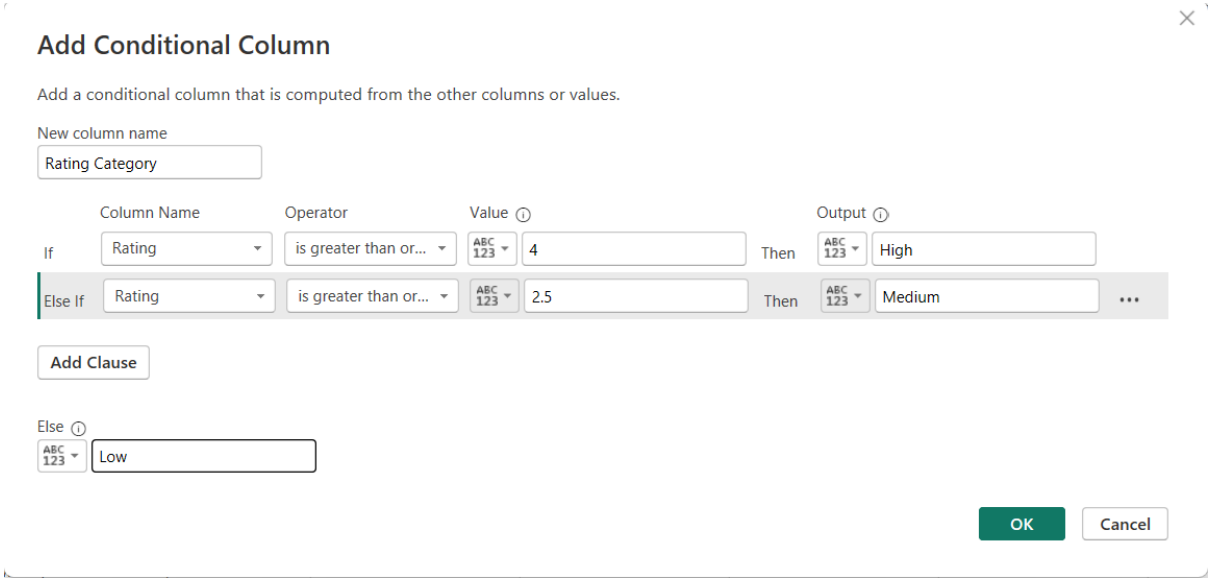
Here are the applied steps for question 2

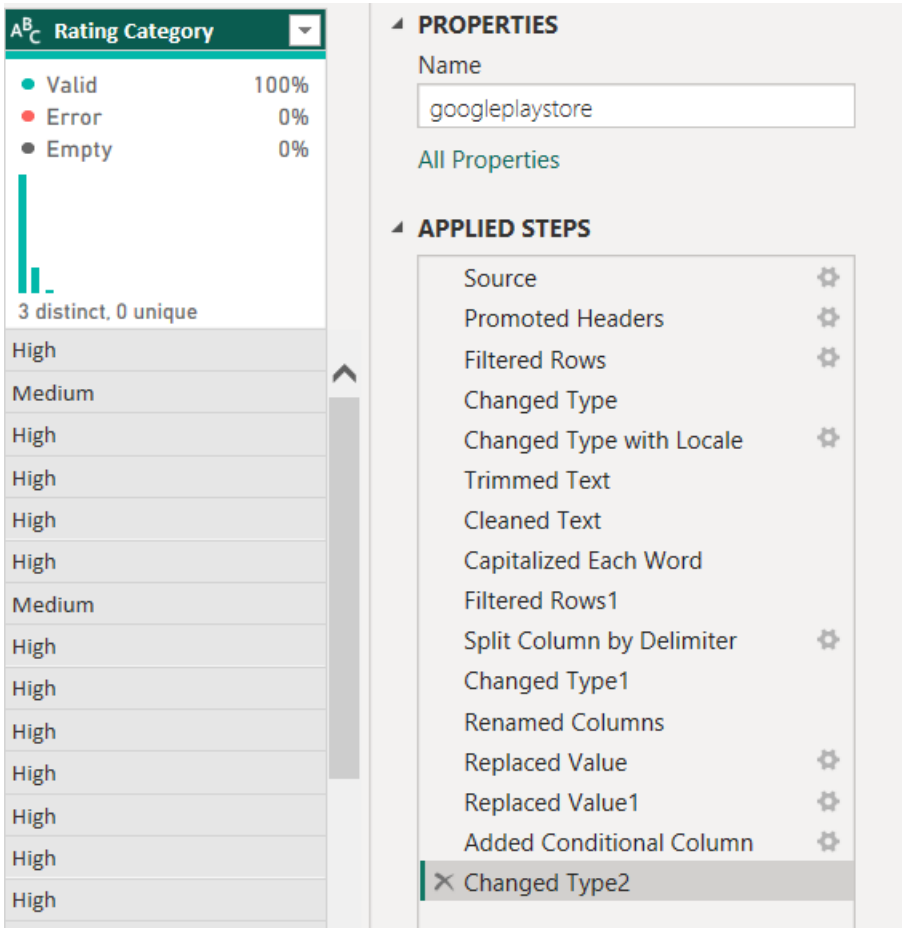


For q3 I trimmed, cleaning, capitalized, and split the android ver column. I also renamed the new columns and also filtered out and replaced values that were obvious in this column's case.

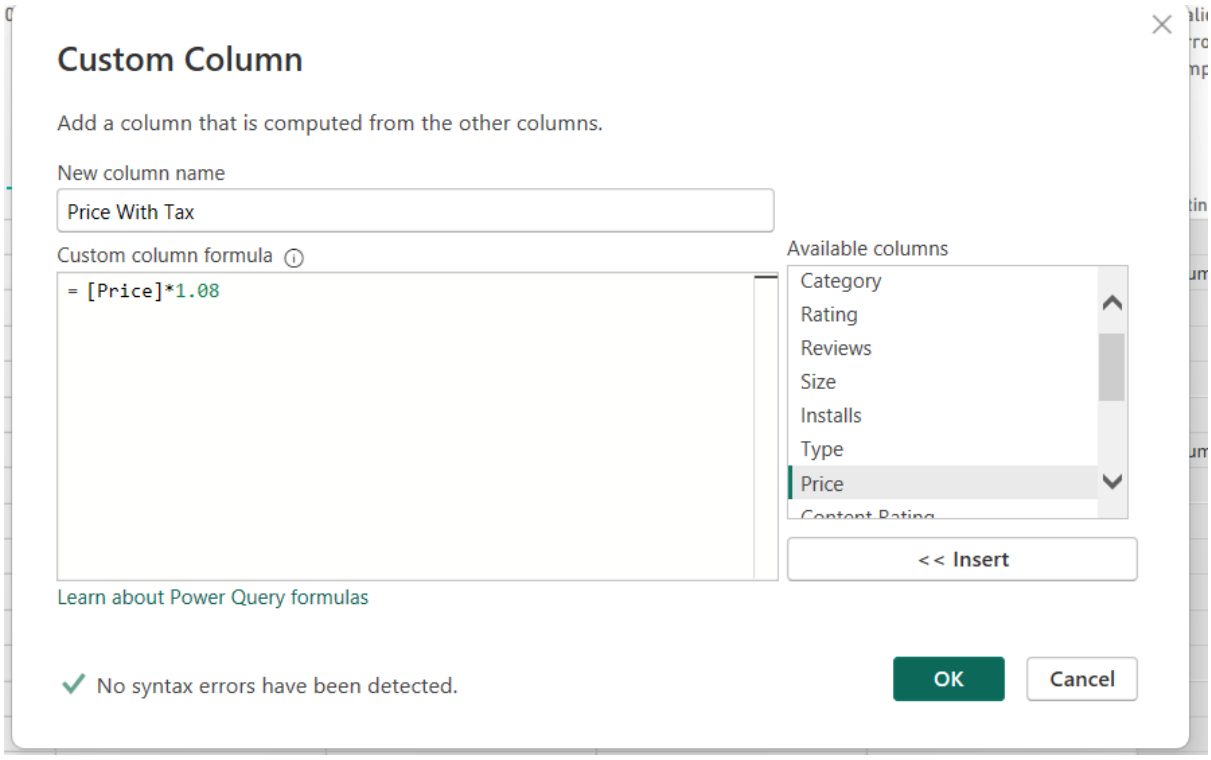


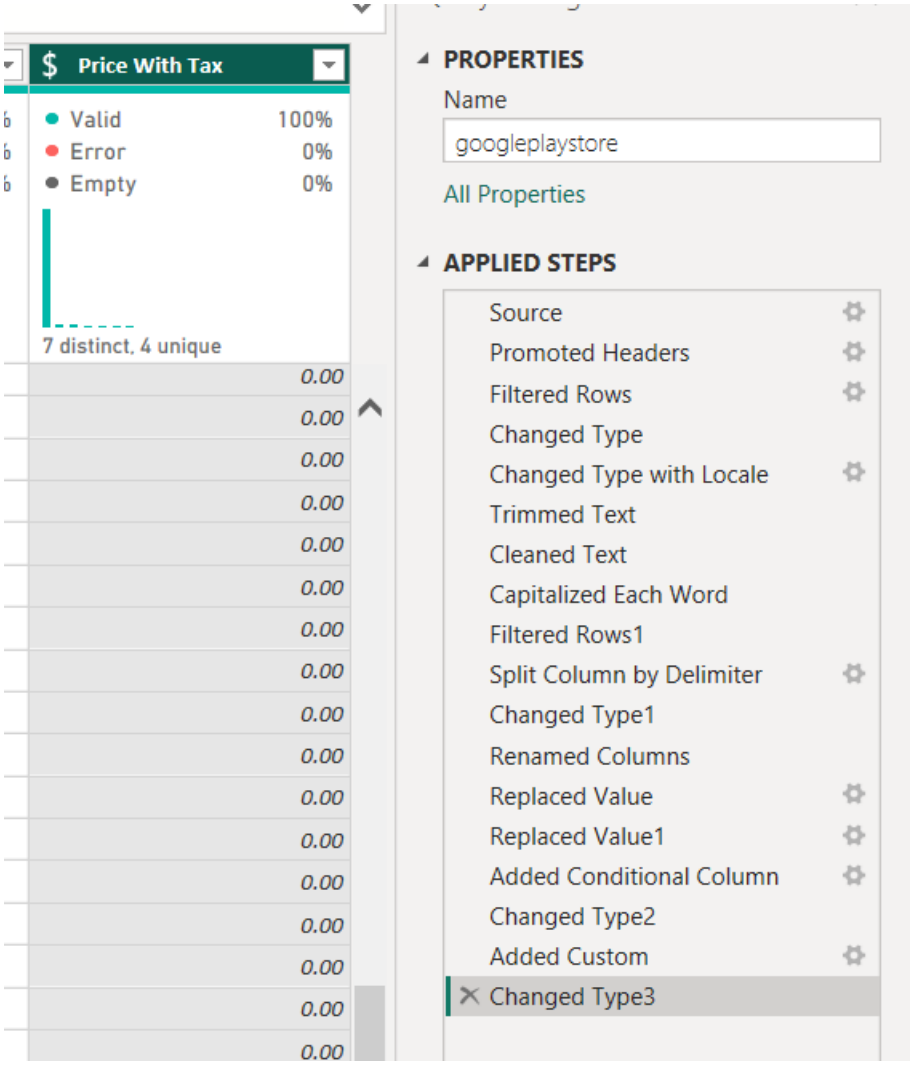
For q4 I created a column called Ratings category and here are the conditions and results.





For the custom column here is the m formula and results.





For q5 I went with group by. I duplicated the table googleplaystore and renamed it category stats and grouped it by the below conditions

Group By

Specify the columns to group by and one or more outputs.

Basic Advanced

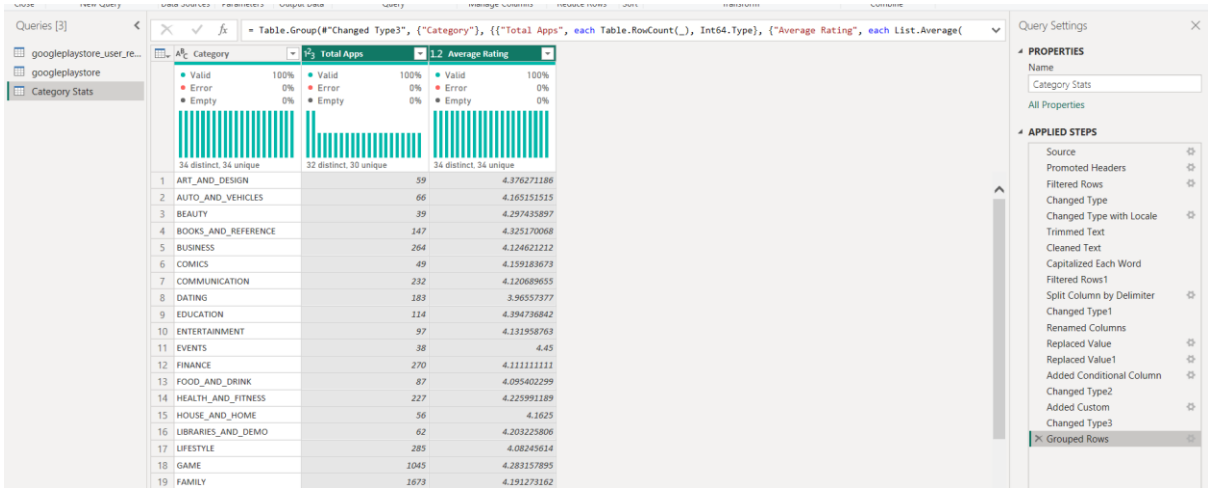
Category

Add grouping

New column name	Operation	Column
Total Apps	Count Rows	
Average Rating	Average	Rating

Add aggregation

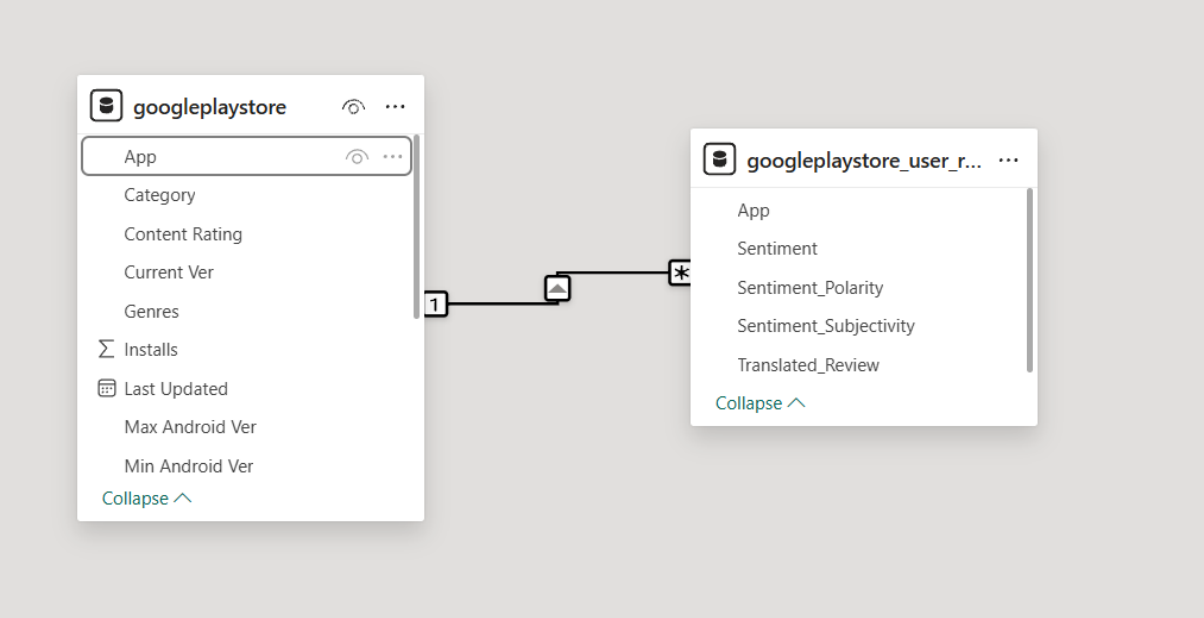
OK Cancel



Q6

googleplaystore_user_reviews is the fact table because it contains the events, reviews submitted by users and changes frequently. While googleplaystore is the dimension table it contains the descriptive attributes about the apps and serves as a lookup table with unique app names.

Q7



Tables: googleplaystore to googleplaystore_user_reviews

Cardinality: 1-to-Many

Cross filter direction: Single

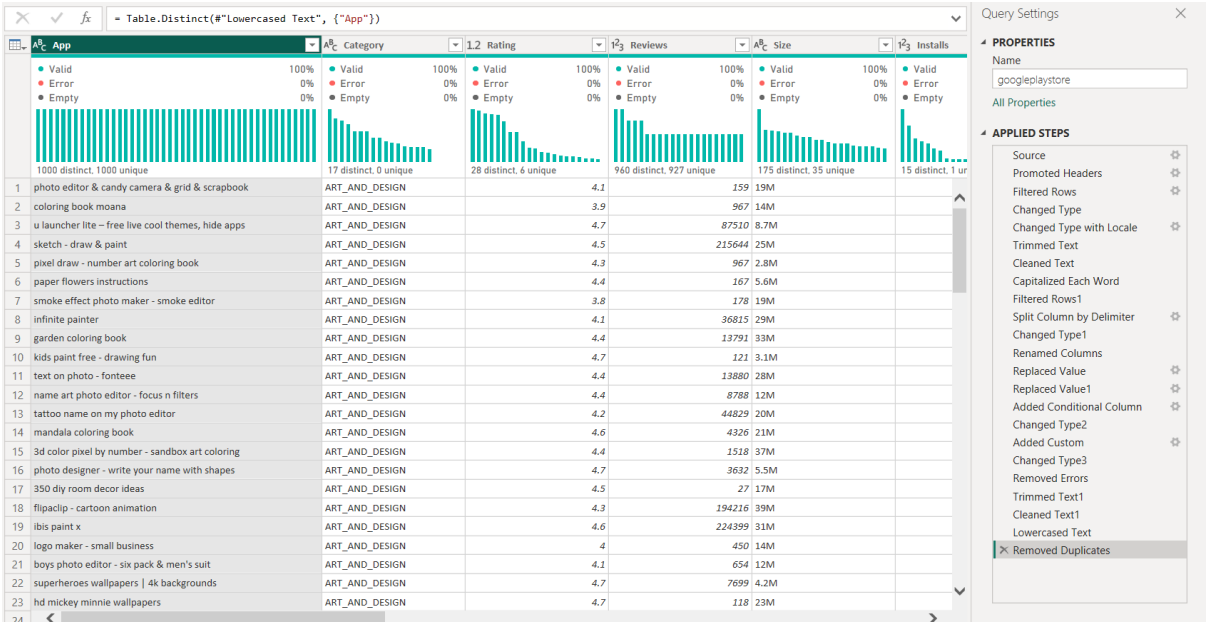
Active/inactive: Active

Why correct: This is correct because one single app in the dimension table can have many individual user reviews in the fact table.

A modeling mistake that would break totals is setting a many-to-many relationship instead of a 1-to-many relationship. If both tables had duplicate app names, the relationship would become many-to-many, which causes filter ambiguity and results in incorrect, inflated totals when visualizing the data.

Q8

I verified the dimension key (App) was unique by checking the model relationship. Power BI initially flagged a many-to-many cardinality, indicating duplicates existed. I checked for duplicates and resolved them in Power Query by Trimming the text, converting it to lowercase, and applying the 'Remove Duplicates' function to ensure strictly unique keys for a valid 1-to-many relationship.



Q9

Yes. It should exist.

1. It allows you to use Power BI's built-in Time Intelligence DAX functions like Year-to-Date or Month-over-Month calculations.

2. It acts as a central filter, allowing you to slice multiple fact tables using one continuous, standardized calendar without missing dates.

It should contain: Date, Year, and Month Name

Bonus points

Table	Column	Meaning	Type	Key/Attribute
googleplaystore	App	Unique name of the application	Text	Primary Key
googleplaystore	Price	Cost of the app (after locale conversion)	Currency	Attribute
googleplaystore	Rating Category	Custom conditional grouping (High/Medium/Low)	Text	Attribute
googleplaystore	Min Android Version	Minimum OS required (from split column step)	Text	Attribute
googleplaystore	Last Updated	The date the application was last updated	Date	Attribute
googleplaystore_user_reviews	App	Name of the app being reviewed	Text	Foreign Key
googleplaystore_user_reviews	Translated_Review	The user's written feedback in English	Text	Attribute
googleplaystore_user_reviews	Sentiment	The emotional tone of the review	Text	Attribute

Table	Column	Meaning	Type	Key/Attribute
		(Positive/Negative)		