Team name - Girls Squad

Punam Nagrale, Raga Rasagna Paruchuri, Ramya Rao Thella, Smriti Kumari

Module 1 - Playstore Apps analysis

1. Which apps have the highest rating in the given available dataset?

```
%sql
SELECT App AS top_apps, Rating
FROM playstore_apps
WHERE Rating = 5;
```

(1) Spark Jobs

_sqldf: pyspark.sql.dataframe.DataFrame = [top_apps: string, Rating: string]

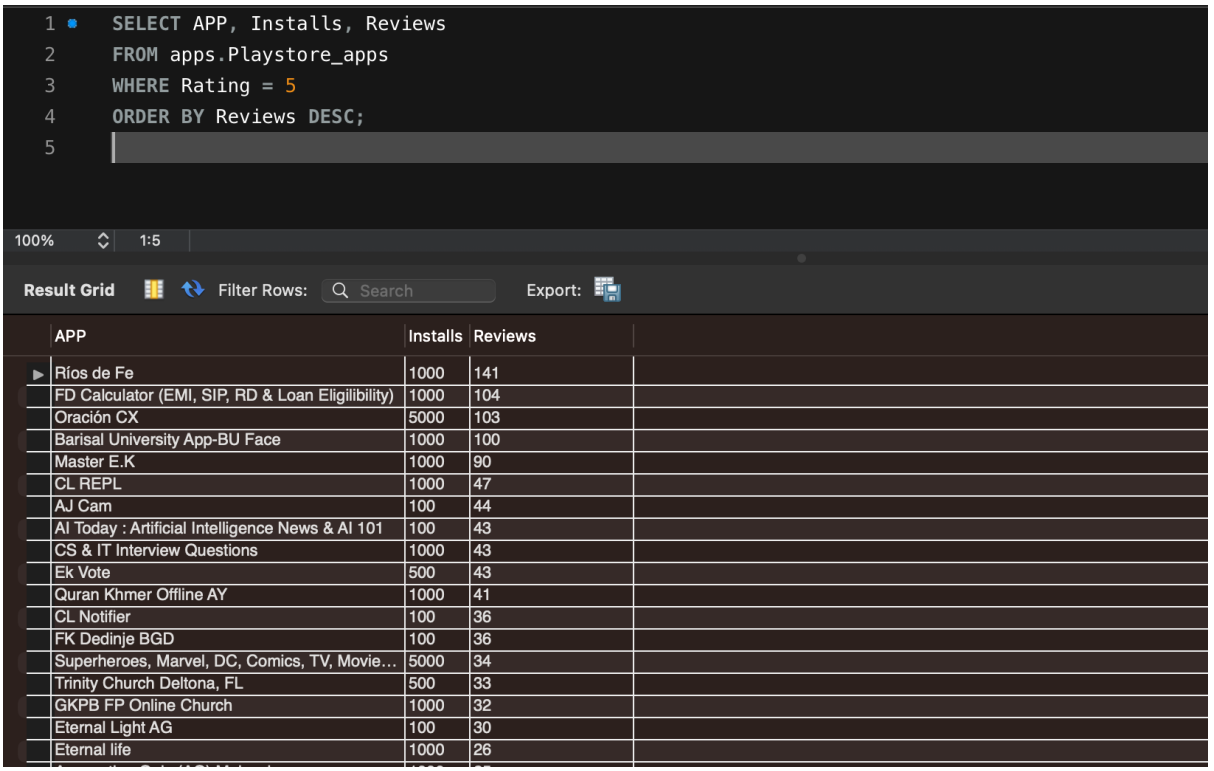
Table +

	top_apps	Rating
1	211:CK	5
2	30WPM Amateur ham radio Koch CW Morse code trainer	5
3	420 BZ Budeze Delivery	5
4	AC DC Power Monitor	5
5	Accounting Quiz (AQ) Malaysia	5
6	ADS-B Driver	5
7	Ag Valley Cooperative	5

Showing all 271 rows. | 1.91 seconds runtime

2. What are the number of installs and reviews for the above apps? Return the apps with the highest reviews to the top.

```
SELECT App, Installs, Reviews
FROM apps.Playstore_apps
WHERE Rating = 5
ORDER BY Reviews DESC;
```



```
1  SELECT APP, Installs, Reviews
2  FROM apps.Playstore_apps
3  WHERE Rating = 5
4  ORDER BY Reviews DESC;
5
```

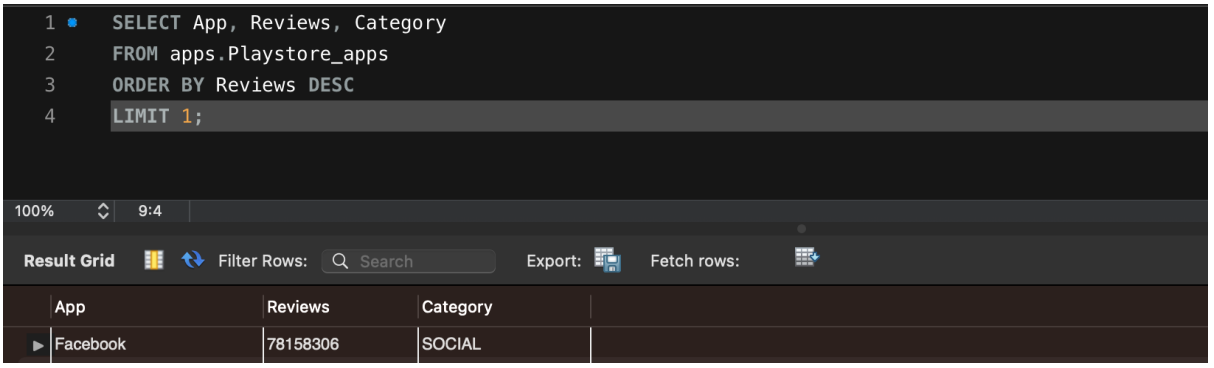
100% 1:5

Result Grid Filter Rows: Search Export:

APP	Installs	Reviews
Ríos de Fe	1000	141
FD Calculator (EMI, SIP, RD & Loan Eligibility)	1000	104
Oración CX	5000	103
Barisal University App-BU Face	1000	100
Master E.K	1000	90
CL REPL	1000	47
AJ Cam	100	44
AI Today : Artificial Intelligence News & AI 101	100	43
CS & IT Interview Questions	1000	43
Ek Vote	500	43
Quran Khmer Offline AY	1000	41
CL Notifier	100	36
FK Dedinje BGD	100	36
Superheroes, Marvel, DC, Comics, TV, Movie...	5000	34
Trinity Church Deltona, FL	500	33
GKPB FP Online Church	1000	32
Eternal Light AG	100	30
Eternal life	1000	26

3. Which app has the highest number of reviews? Also, mention the number of reviews and category of the app

```
SELECT App, Reviews, Category
FROM apps.Playstore_apps
ORDER BY Reviews DESC
LIMIT 1;
```



```
1  SELECT App, Reviews, Category
2  FROM apps.Playstore_apps
3  ORDER BY Reviews DESC
4  LIMIT 1;
```

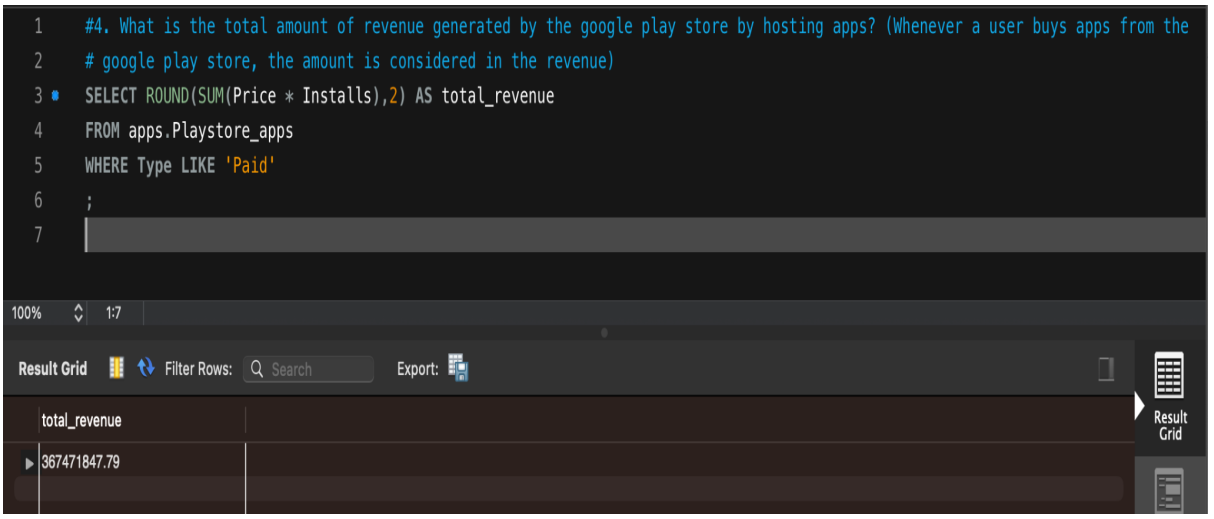
100% 9:4

Result Grid Filter Rows: Search Export: Fetch rows:

App	Reviews	Category
Facebook	78158306	SOCIAL

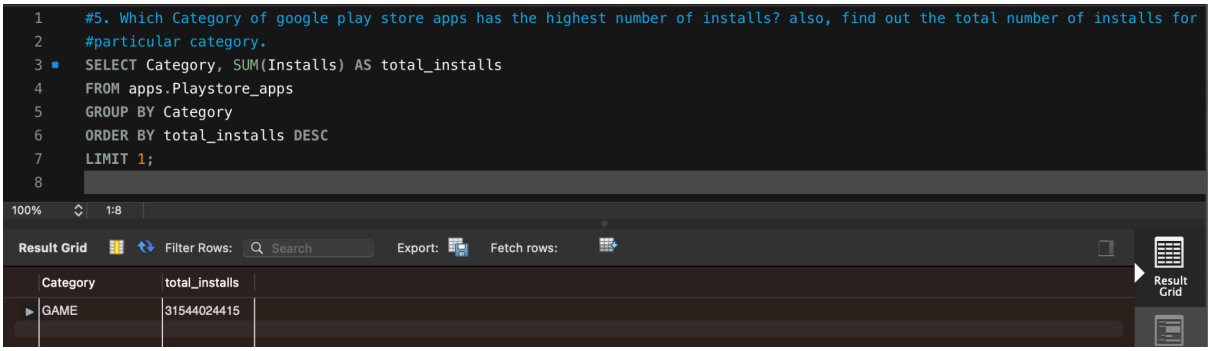
4. What is the total amount of revenue generated by the google play store by hosting apps? (Whenever a user buys apps from the google play store, the amount is considered in the revenue)

```
SELECT ROUND(SUM(Price * Installs),2) AS total_revenue
FROM apps.Playstore_apps
WHERE Type = "Paid";
```



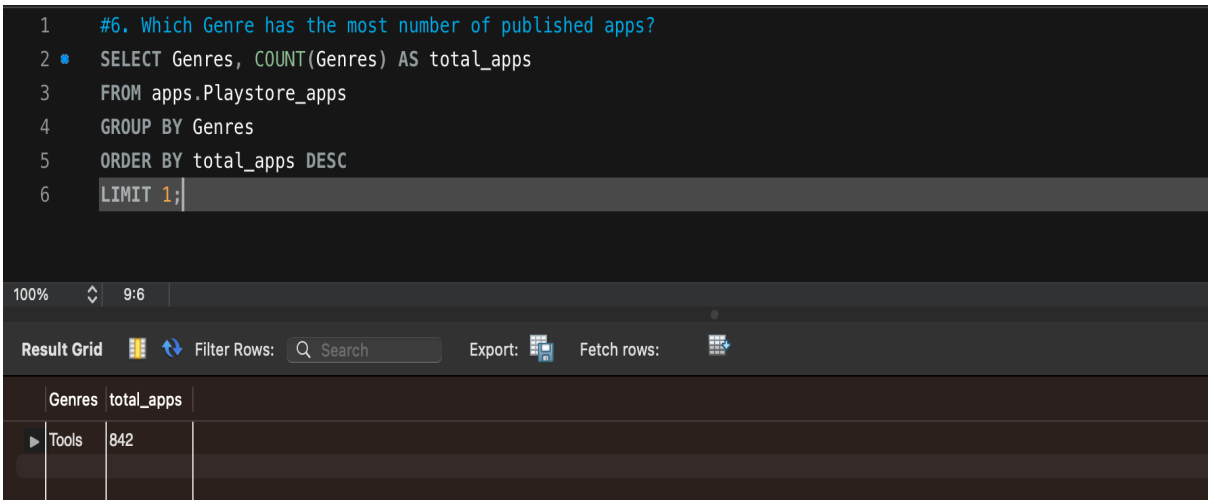
5. Which Category of google play store apps has the highest number of installs? also, find out the total number of installs for that particular category.

```
SELECT Category, SUM(Installs) AS total_installs
FROM apps.Playstore_apps
GROUP BY Category
ORDER BY total_installs DESC
LIMIT 1;
```



6. Which Genre has the most number of published apps?

```
SELECT Genres, COUNT(Genres) AS total_apps
FROM apps.Playstore_apps
GROUP BY Genres
ORDER BY total_apps DESC
LIMIT 1;
```



7. Provide the list of all games ordered in such a way that the game that has the highest number of installs is displayed on the top (to avoid duplicate results use distinct)

```
SELECT DISTINCT App, Installs
FROM playstore_apps
WHERE Category = "GAME"
ORDER BY Installs DESC;
```

```
1  %sql
2  SELECT DISTINCT App, Installs
3  FROM playstore_apps
4  WHERE Category = 'GAME'
5  ORDER BY Installs DESC;
6
```

▸ (2) Spark Jobs

▸ _sqldf: pyspark.sql.dataframe.DataFrame = [App: string, Installs: double]

Table +

	App	Installs
1	Subway Surfers	1000000000
2	Pou	500000000
3	Candy Crush Saga	500000000
4	My Talking Tom	500000000
5	Temple Run 2	500000000
6	Crossy Road	100000000
7	Candy Crush Soda Saga	100000000

Showing all 947 rows. | 1.30 seconds runtime

8. Provide the list of apps that can work on android version 4.0.3 and UP.

a) android version '4.0.3 and UP'

```
SELECT App, `Android Ver`
FROM playstore_apps
WHERE `Android Ver` = "4.0.3 and up";
```

```
%sql
SELECT App, `Android Ver`
FROM playstore_apps
WHERE `Android Ver` = '4.0.3 and up';
```

(1) Spark Jobs

_sqldf: pyspark.sql.dataframe.DataFrame = [App: string, Android Ver: string]

Table +

	App	Android Ver
1	?? Comics - Manga,Novel and Stories	4.0.3 and up
2	?? Football Wallpapers 4K | Full HD Backgrounds ??	4.0.3 and up
3	?? WhatsLov: Smileys of love, stickers and GIF	4.0.3 and up
4	??? - ????, ?????, BJ??	4.0.3 and up
5	??? ??? H	4.0.3 and up
6	????? Astrology - Min Thein Kha BayDin	4.0.3 and up
7	???????????EJ STYLE??????????	4.0.3 and up

Truncated results, showing first 1,000 rows. | 0.52 seconds runtime

b) All android versions including '4.0.3 and UP' and above

```
SELECT App, `Android Ver`
FROM playstore_apps
WHERE `Android Ver`<> "4.0 and up"
AND `Android Ver` LIKE "4%"
OR `Android Ver` LIKE "5%"
OR `Android Ver` LIKE "6%"
OR `Android Ver` LIKE "7%"
OR `Android Ver` LIKE "8%"
ORDER BY `Android Ver`;
```

```
%sql
SELECT App, `Android Ver`
FROM playstore_apps
WHERE `Android Ver` <> '4.0 and up'
AND `Android Ver` LIKE '4%'
OR `Android Ver` LIKE '5%'
OR `Android Ver` LIKE '6%'
OR `Android Ver` LIKE '7%'
OR `Android Ver` LIKE '8%'
ORDER BY `Android Ver`;
```

(1) Spark Jobs

_sqldf: pyspark.sql.dataframe.DataFrame = [App: string, Android Ver: string]

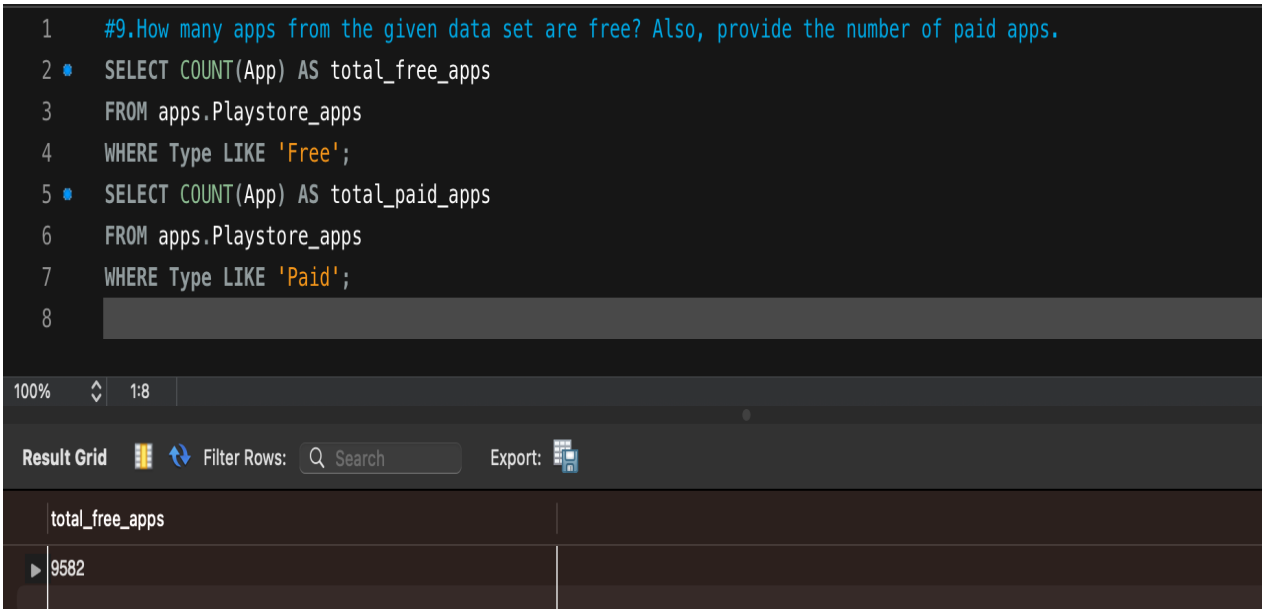
Table +

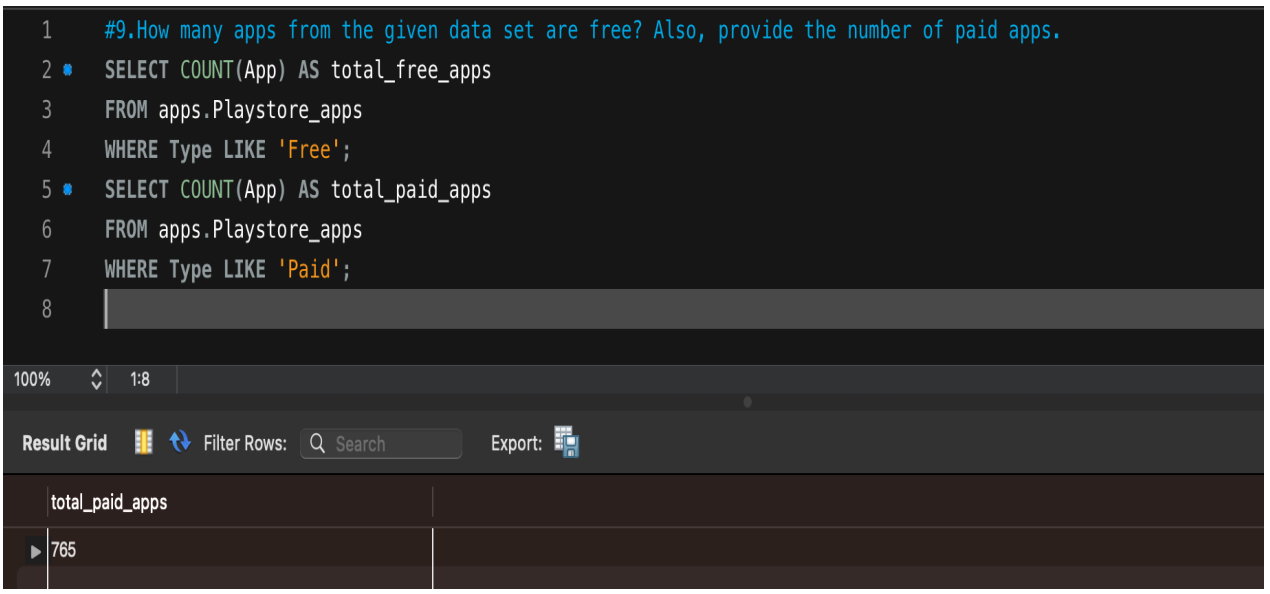
	App	Android Ver
1	Browser 4G	4.0.3 - 7.1.1
2	Web Browser & Explorer	4.0.3 - 7.1.1
3	?? Comics - Manga,Novel and Stories	4.0.3 and up
4	?? Football Wallpapers 4K | Full HD Backgrounds ??	4.0.3 and up
5	?? WhatsLov: Smileys of love, stickers and GIF	4.0.3 and up
6	??? - ????, ?????, BJ??	4.0.3 and up
7	??? ??? H	4.0.3 and up

Truncated results, showing first 1,000 rows. | 0.94 seconds runtime

9. How many apps from the given data set are free? Also, provide the number of paid apps.

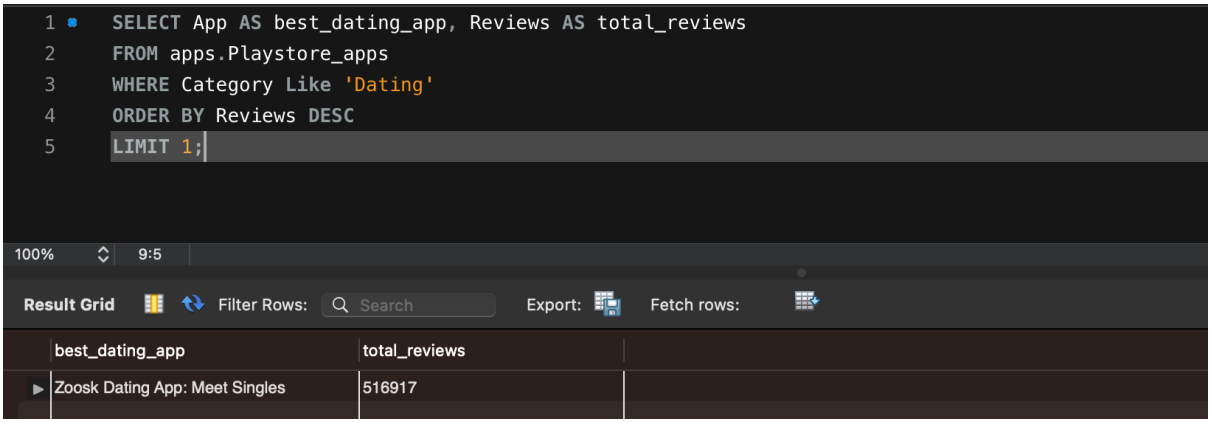
```
SELECT COUNT(App) AS total_free_apps
FROM playstore_apps
WHERE Type = “Free”;

SELECT COUNT(App) AS total_paid_apps
FROM playstore_apps
WHERE Type = “Paid”;
```





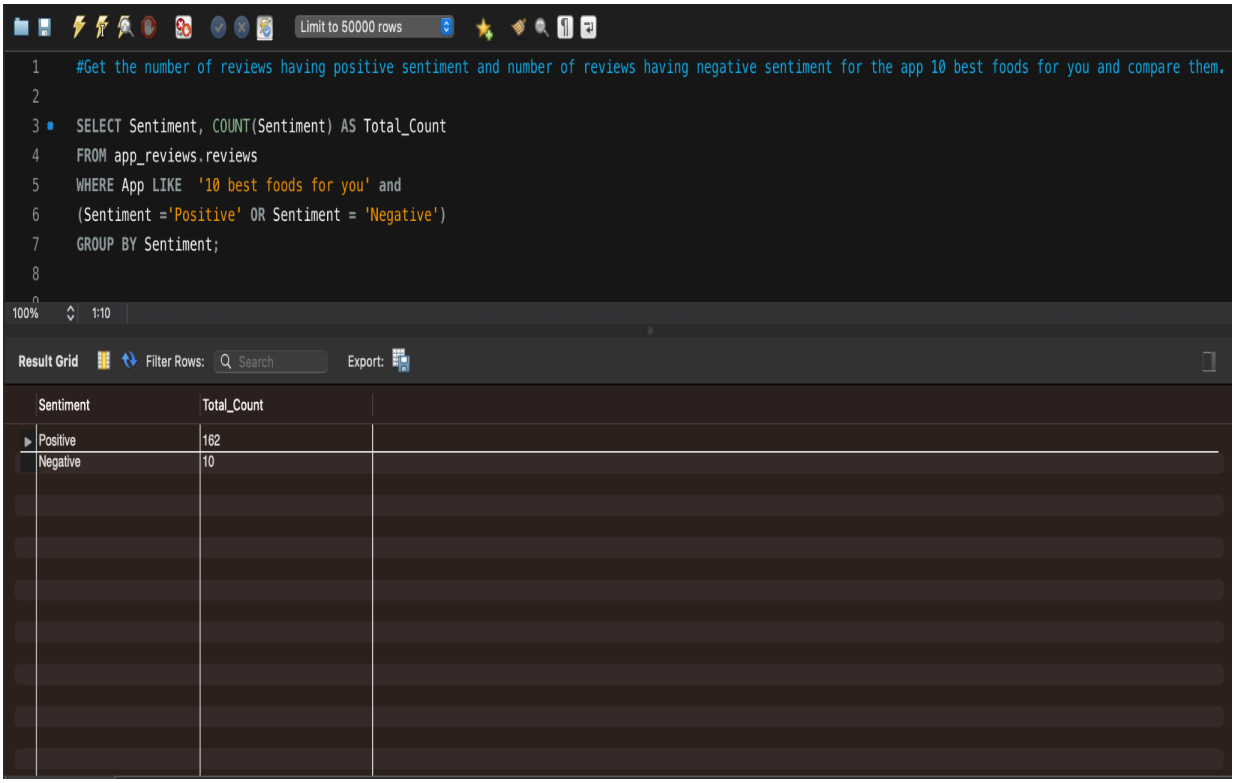
10. Which is the best dating app? (Best dating app is the one having the highest number of Reviews)

```
SELECT App AS best_dating_app, Reviews AS total_reviews
FROM apps.Playstore_apps
WHERE Category = “DATING”
ORDER BY Reviews DESC
LIMIT 1;
```



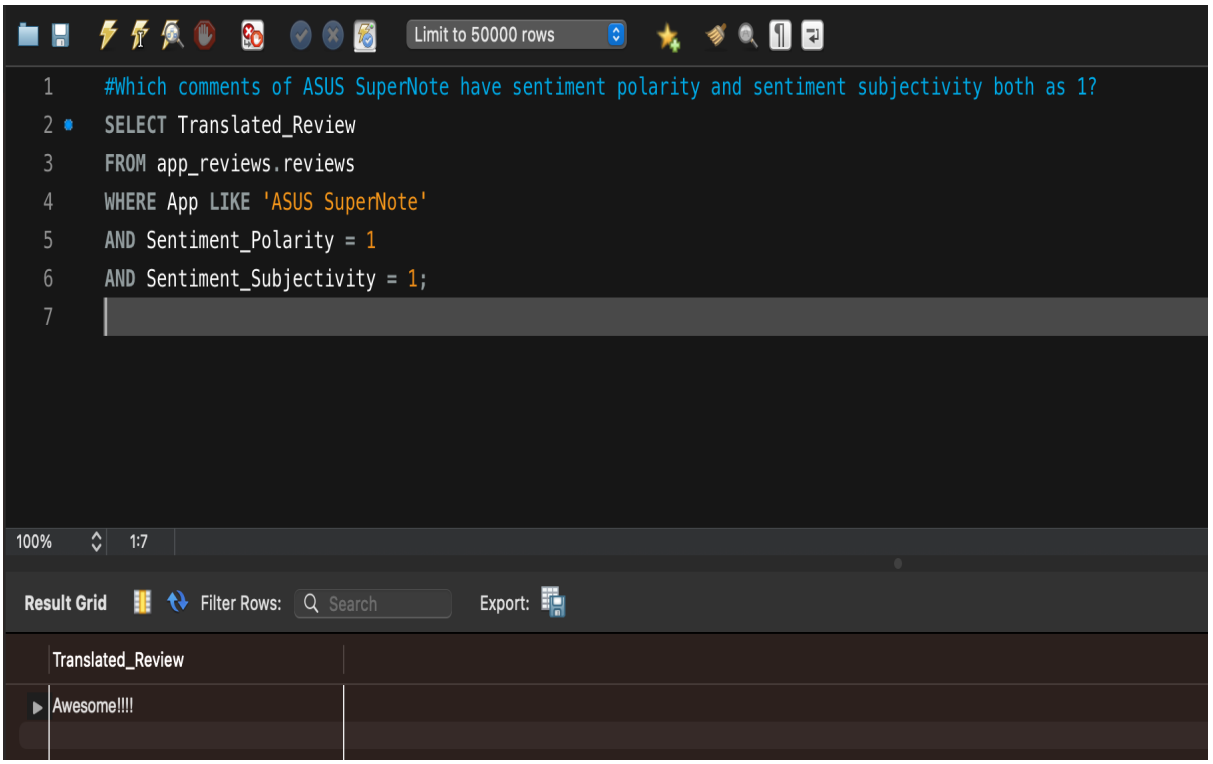
11. Get the number of reviews having positive sentiment and number of reviews having negative sentiment for the app **10 best foods for you** and compare them.

```
SELECT Sentiment, COUNT(Sentiment) AS Total_Count
FROM reviews
WHERE App LIKE “10 best foods for you”
AND (Sentiment = “Positive” OR  Sentiment = “Negative”)
GROUP BY Sentiment;
```



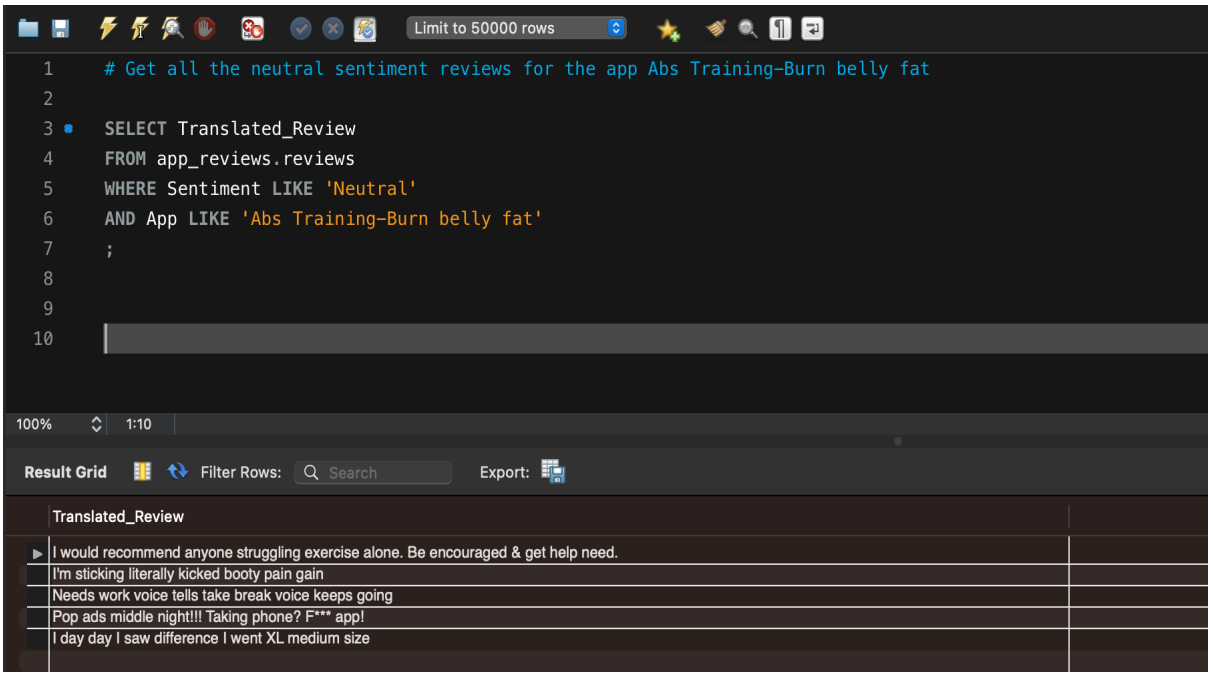
12. Which comments of **ASUS SuperNote** have sentiment polarity and sentiment subjectivity both as 1?

```
SELECT Translated_Review
FROM reviews
WHERE App LIKE 'ASUS SuperNote'
AND Sentiment_Polarity = 1
AND Sentiment_Subjectivity = 1;
```



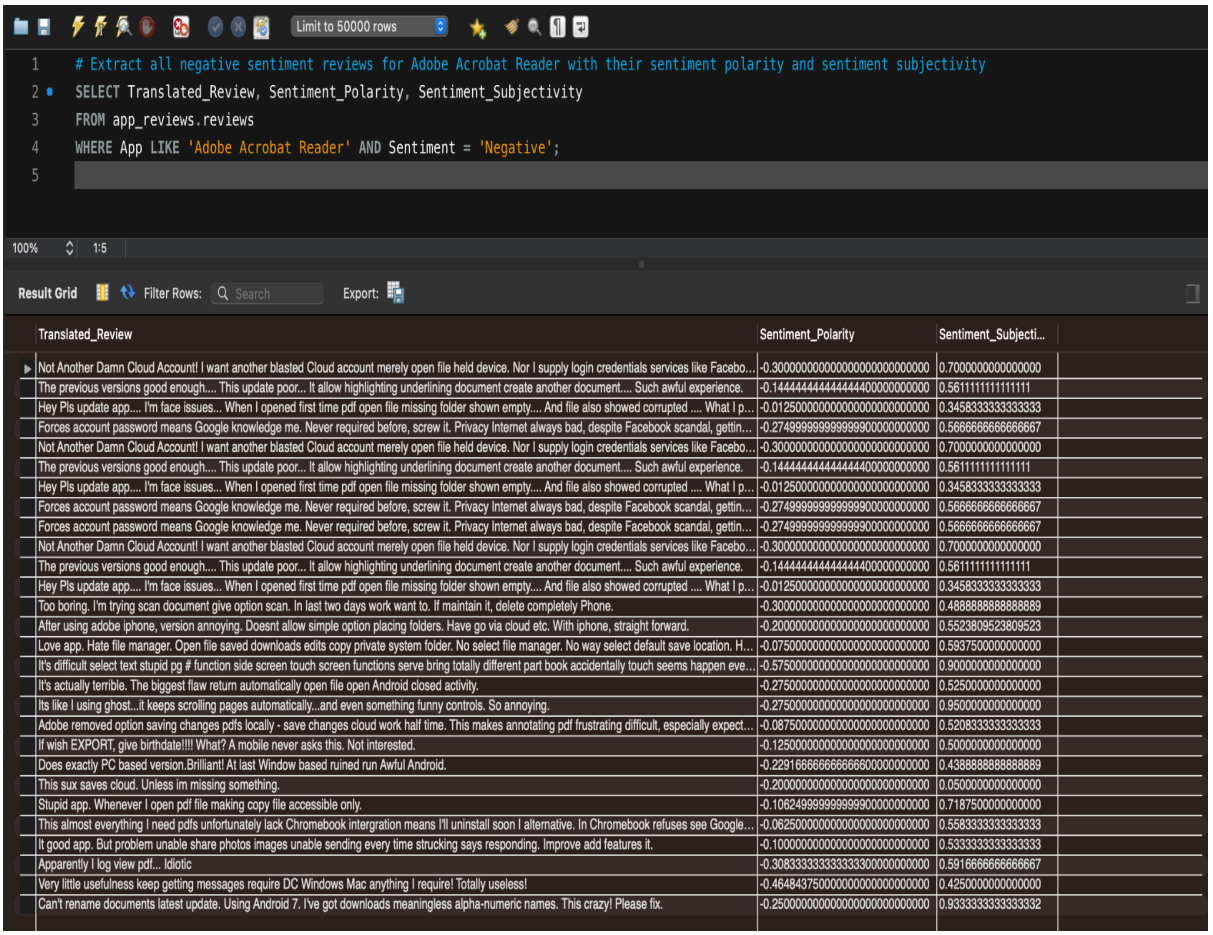
13. Get all the neutral sentiment reviews for the app **Abs Training-Burn belly fat**

```
SELECT Translated_Review
FROM reviews
WHERE Sentiment LIKE “Neutral”
AND App LIKE “Abs Training-Burn belly fat”;
```



14. Extract all negative sentiment reviews for **Adobe Acrobat Reader** with their sentiment polarity and sentiment subjectivity

```
SELECT Translated_Review, Sentiment_Polarity, Sentiment_Subjectivity
FROM reviews
WHERE App LIKE "Adobe Acrobat Reader"
AND Sentiment = "Negative";
```



```
Limit to 50000 rows

1  # Extract all negative sentiment reviews for Adobe Acrobat Reader with their sentiment polarity and sentiment subjectivity
2  SELECT Translated_Review, Sentiment_Polarity, Sentiment_Subjectivity
3  FROM app_reviews.reviews
4  WHERE App LIKE 'Adobe Acrobat Reader' AND Sentiment = 'Negative';
5
```

100% 1:5

Result Grid | Filter Rows: Search | Export:

Translated_Review	Sentiment_Polarity	Sentiment_Subjecti...
Not Another Damn Cloud Account! I want another blasted Cloud account merely open file held device. Nor I supply login credentials services like Facebo...	-0.300000000000000000000000000	0.7000000000000000
The previous versions good enough.... This update poor... It allow highlighting underlining document create another document.... Such awful experience.	-0.144444444444444400000000000	0.5611111111111111
Hey Pls update app.... I'm face issues... When I opened first time pdf open file missing folder shown empty.... And file also showed corrupted .... What I p...	-0.012500000000000000000000000	0.3458333333333333
Forces account password means Google knowledge me. Never required before, screw it. Privacy Internet always bad, despite Facebook scandal, gettin...	-0.274999999999999900000000000	0.5666666666666667
Not Another Damn Cloud Account! I want another blasted Cloud account merely open file held device. Nor I supply login credentials services like Facebo...	-0.300000000000000000000000000	0.7000000000000000
The previous versions good enough.... This update poor... It allow highlighting underlining document create another document.... Such awful experience.	-0.144444444444444400000000000	0.5611111111111111
Hey Pls update app.... I'm face issues... When I opened first time pdf open file missing folder shown empty.... And file also showed corrupted .... What I p...	-0.012500000000000000000000000	0.3458333333333333
Forces account password means Google knowledge me. Never required before, screw it. Privacy Internet always bad, despite Facebook scandal, gettin...	-0.274999999999999900000000000	0.5666666666666667
Forces account password means Google knowledge me. Never required before, screw it. Privacy Internet always bad, despite Facebook scandal, gettin...	-0.274999999999999900000000000	0.5666666666666667
Not Another Damn Cloud Account! I want another blasted Cloud account merely open file held device. Nor I supply login credentials services like Facebo...	-0.300000000000000000000000000	0.7000000000000000
The previous versions good enough.... This update poor... It allow highlighting underlining document create another document.... Such awful experience.	-0.144444444444444400000000000	0.5611111111111111
Hey Pls update app.... I'm face issues... When I opened first time pdf open file missing folder shown empty.... And file also showed corrupted .... What I p...	-0.012500000000000000000000000	0.3458333333333333
Too boring. I'm trying scan document give option scan. In last two days work want to. If maintain it, delete completely Phone.	-0.300000000000000000000000000	0.4888888888888889
After using adobe iphone, version annoying. Doesnt allow simple option placing folders. Have go via cloud etc. With iphone, straight forward.	-0.200000000000000000000000000	0.5523809523809523
Love app. Hate file manager. Open file saved downloads edits copy private system folder. No select file manager. No way select default save location. H...	-0.075000000000000000000000000	0.5937500000000000
It's difficult select text stupid pg # function side screen touch screen functions serve bring totally different part book accidentally touch seems happen eve...	-0.575000000000000000000000000	0.9000000000000000
It's actually terrible. The biggest flaw return automatically open file open Android closed activity.	-0.275000000000000000000000000	0.5250000000000000
Its like I using ghost...it keeps scrolling pages automatically...and even something funny controls. So annoying.	-0.275000000000000000000000000	0.9500000000000000
Adobe removed option saving changes pdfs locally - save changes cloud work half time. This makes annotating pdf frustrating difficult, especially expect...	-0.087500000000000000000000000	0.5208333333333333
If wish EXPORT, give birthdate!!!! What? A mobile never asks this. Not interested.	-0.125000000000000000000000000	0.5000000000000000
Does exactly PC based version.Brilliant! At last Window based ruined run Awful Android.	-0.229166666666666600000000000	0.4388888888888889
This sux saves cloud. Unless im missing something.	-0.200000000000000000000000000	0.0500000000000000
Stupid app. Whenever I open pdf file making copy file accessible only.	-0.106249999999999900000000000	0.7187500000000000
This almost everything I need pdfs unfortunately lack Chromebook intergration means I'll uninstall soon I alternative. In Chromebook refuses see Google...	-0.062500000000000000000000000	0.5583333333333333
It good app. But problem unable share photos images unable sending every time strucking says responding. Improve add features it.	-0.100000000000000000000000000	0.5333333333333333
Apparently I log view pdf... Idiotic	-0.308333333333333300000000000	0.5916666666666667
Very little usefulness keep getting messages require DC Windows Mac anything I require! Totally useless!	-0.464843750000000000000000000	0.4250000000000000
Can't rename documents latest update. Using Android 7. I've got downloads meaningless alpha-numeric names. This crazy! Please fix.	-0.250000000000000000000000000	0.9333333333333332