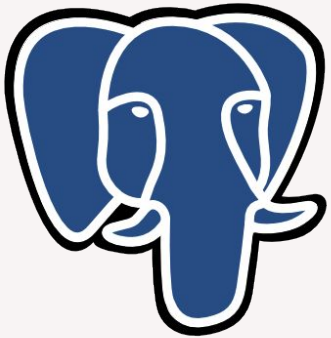
Portfolio SQL DANA Sentiment Analysis

By : Johan Sanjaya

TOOLS

Excel



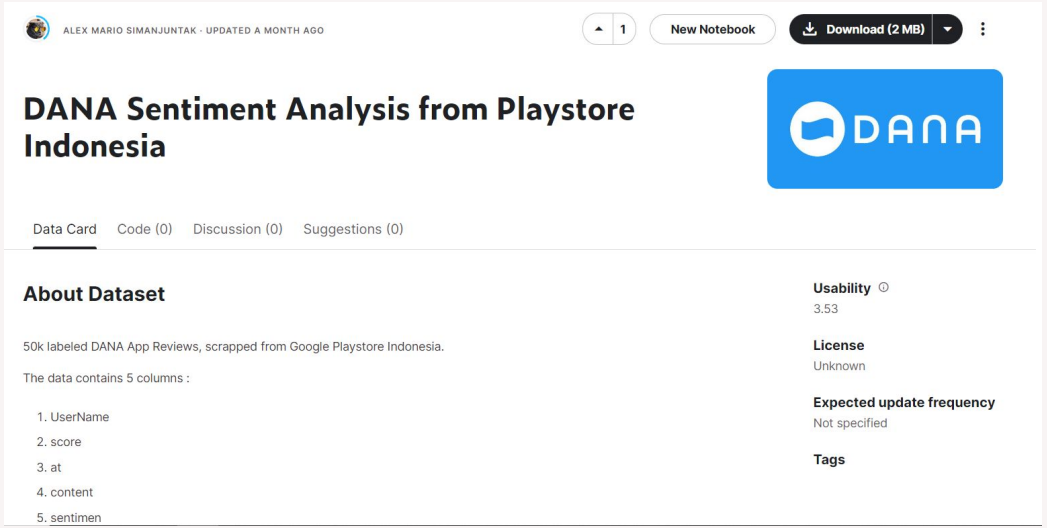
PostgreSQL

GitHub

SOURCE DATA

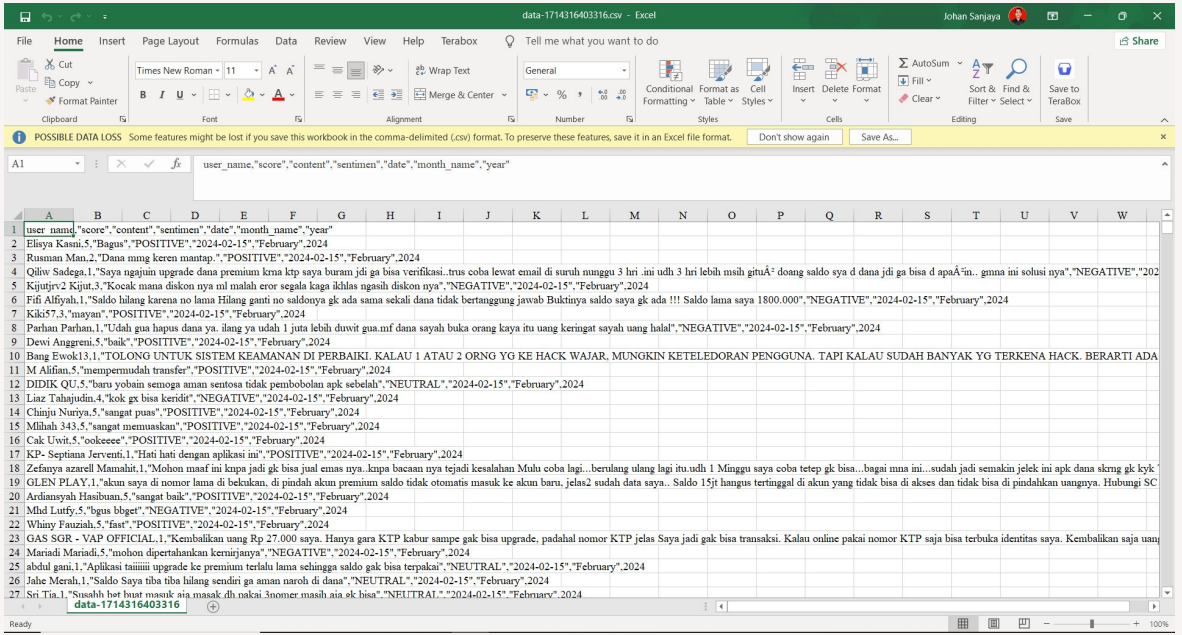
DATA

- The topic to be worked on is “DANA Sentiment Analysis from Playstore Indonesia”
- The dataset is downloaded from web : LINK



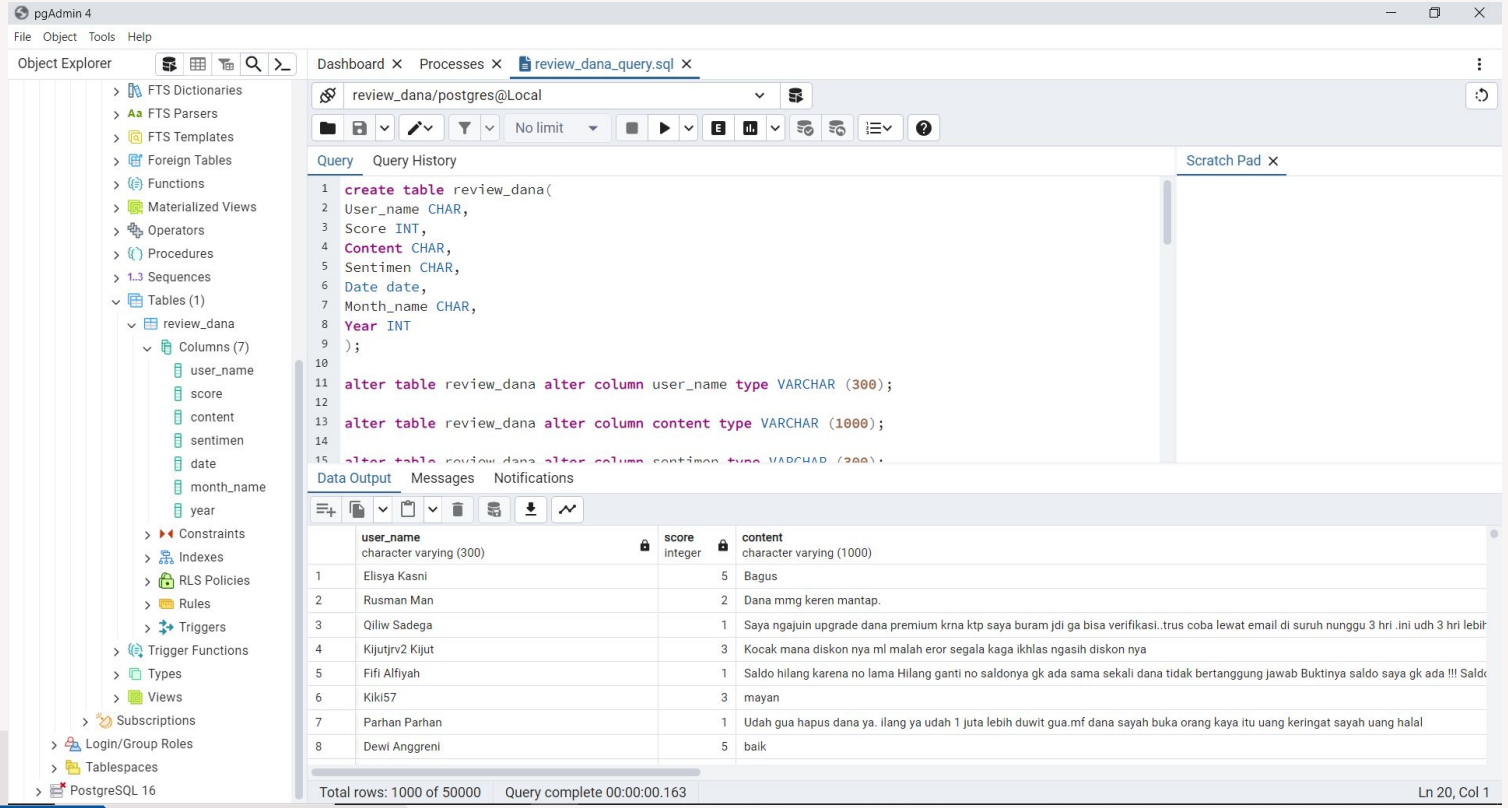
DATA

- The data is presented in csv format
- Data imported to GUI PostgreSQL pgAdmin
- Data can be accessed in here : LINK



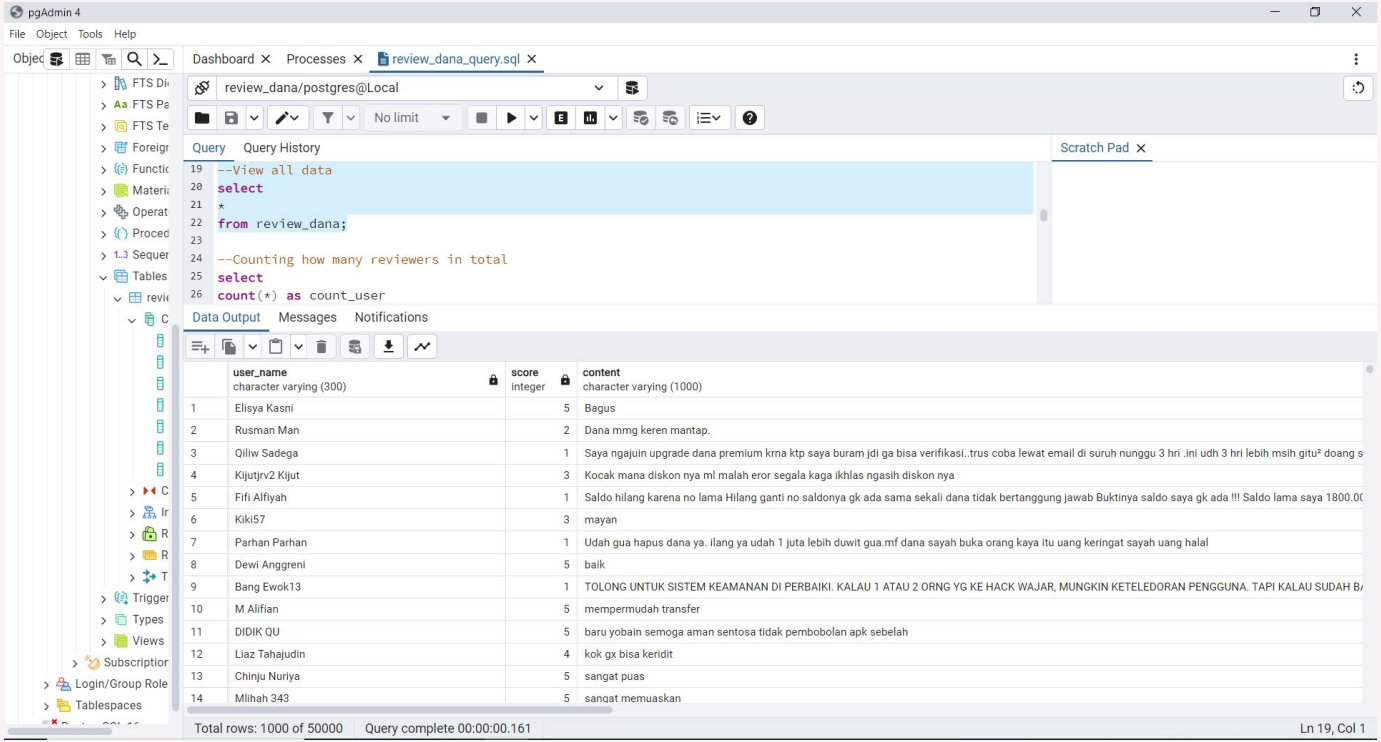
LINK GITHUB

PGADMIN DISPLAY



QUERIES

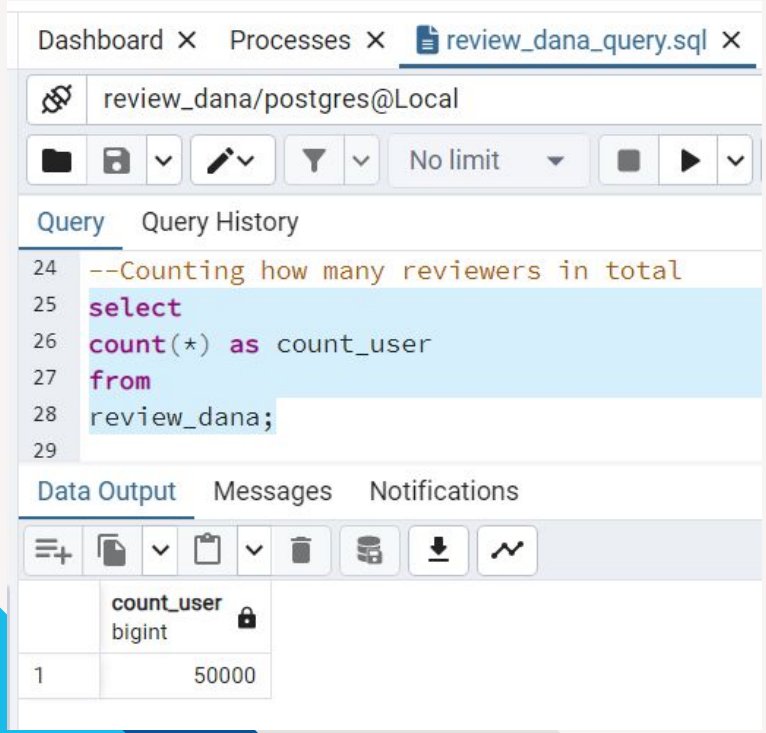
1. Viewing all data
   a. This query is to view all data.
   b. The data contains 50000 rows



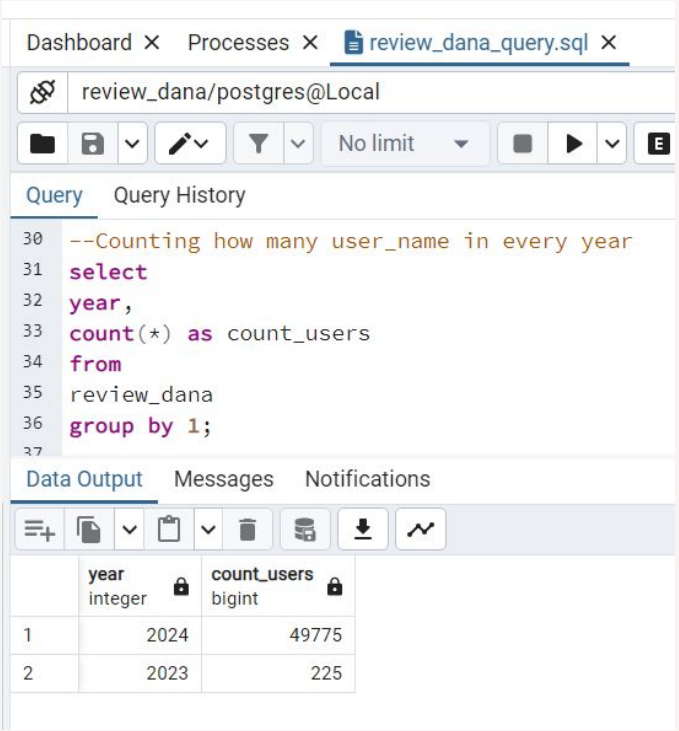
	user_name character varying (300)	score integer	content character varying (1000)
1	Elisya Kasni	5	Bagus
2	Rusman Man	2	Dana mmg keren mantap.
3	Qiliw Sadega	1	Saya ngajuin upgrade dana premium krna ktp saya buram jdi ga bisa verifikasi..trus coba lewat email di suruh nunggu 3 hri .ini udh 3 hri lebih msih gitu² doang s
4	Kijutjrv2 Kijut	3	Kocak mana diskon nya ml malah eror segala kaga ikhlas ngasih diskon nya
5	Fifi Alfiyah	1	Saldo hilang karena no lama Hilang ganti no saldonya gk ada sama sekali dana tidak bertanggung jawab Buktinya saldo saya gk ada !!! Saldo lama saya 1800.0(
6	Kiki57	3	mayan
7	Parhan Parhan	1	Udah gua hapus dana ya. ilang ya udah 1 juta lebih duwit gua.mf dana sayah buka orang kaya itu uang keringat sayah uang halal
8	Dewi Anggreni	5	baik
9	Bang Ewok13	1	TOLONG UNTUK SISTEM KEAMANAN DI PERBAIKI. KALAU 1 ATAU 2 ORNG YG KE HACK WAJAR, MUNGKIN KETELEDORAN PENGGUNA. TAPI KALAU SUDAH BA
10	M Alifian	5	mempermudah transfer
11	DIDIK QU	5	baru yobain semoga aman sentosa tidak pembobolan apk sebelah
12	Liaz Tahajudin	4	kok gx bisa keridit
13	Chinju Nuriya	5	sangat puas
14	Mlihah 343	5	sangat memuaskan

QUERIES

2. Counting how many reviewers in total



3. Counting how many user_name in every year



QUERIES

4. Finding for score percentage in 2024

Dashboard × Processes × review_dana_query.sql ×

review_dana/postgres@Local

No limit

Query Query History

```
38 --Finding for score percentage in 2024
39 select
40 score,
41 count(*) as count_users,
42 round(100*count(*)/sum(count(*)) over(),2) || '%' as percentage
43 from
44 review_dana
45 where year=2024
46 group by 1
47 order by 1 desc;
48
```

Data Output Messages Notifications

	score integer	count_users bigint	percentage text
1	5	28921	58.10%
2	4	4026	8.09%
3	3	2564	5.15%
4	2	2041	4.10%
5	1	12223	24.56%

5. Finding how many user and percentage in every sentimen and every year

Dashboard × Processes × review_dana_query.sql ×

review_dana/postgres@Local

No limit

Query Query History

```
49 --Finding how many user and percentage in every sentimen and every year
50 select
51 year,
52 sentimen,
53 count(*) as count_users,
54 round(100*count(*)/sum(count(*)) over(),2) || '%' as percentage
55 from
56 review_dana
57 group by 1,2
58 order by 1,3 asc;
```

Data Output Messages Notifications

	year integer	sentimen character varying (300)	count_users bigint	percentage text
1	2023	NEUTRAL	32	0.06%
2	2023	NEGATIVE	75	0.15%
3	2023	POSITIVE	118	0.24%
4	2024	NEUTRAL	6340	12.68%
5	2024	NEGATIVE	16998	34.00%
6	2024	POSITIVE	26437	52.87%

QUERIES THAT IMPLEMENTED

1. **SELECT** : retrieve data from table
2. **COUNT** : aggregate function to get number of rows with specific condition
3. **ROUND** : round numeric value
4. **OVER** :execute aggregate inside aggregate queries
5. **ORDER BY** :sort data ascending/descending
6. **GROUP BY** : aggregate same values into summary rows

CONCLUSION

1. Data can be analyzed using query SQL, the data contains 50.000 rows
2. There are 225 reviewers in 2023, 49.775 reviewers in 2024
3. There are more than half of the total reviewers give 5 score rating (58.1%) in 2024, and overall give positive sentiment (52.87%)

THANK YOU