SQL QUERY USED FOR ANALYSIS(on Snowflake)

https://app.snowflake.com/ljatxpp/rs17262/wwbB2V05E1B#query

select \* from tbl\_yelp\_business limit 10;

select \* from tbl\_yelp\_reviews limit 10;

--1.find the number of businesses in each category

with cte as (

select business\_id,trim(A.value) as category

from tbl\_yelp\_business,

lateral split\_to\_table(categories,',') A

)

select category,count(\*)

from cte

group by 1

select trim(A.value) as category,count(\*) as no\_of\_businesses

from tbl\_yelp\_business,

lateral split\_to\_table(categories,',') A

group by category

order by no\_of\_businesses DESC

--2.Find the top 10 users who have reviewed the most businesses in the restaurant category

select a.user\_id,count(distinct a.business\_id) as cnt

from tbl\_yelp\_reviews a

inner join tbl\_yelp\_business b

on a.business\_id=b.business\_id

where b.categories ilike '%restaurant%'

group by user\_id

order by cnt DESC

limit 10

--3.Find the most popular categories of businesses(based on the number of reviews)

with category as (select trim(A.value) as category,business\_id

from tbl\_yelp\_business,

lateral split\_to\_table(categories,',') A

)

select category,count(user\_id) as ttl\_reviews

from category as a

join tbl\_yelp\_reviews as b

on a.business\_id=b.business\_id

group by category

order by ttl\_reviews DESC

--4. Find the top 3 most recent reviews for each business

select business\_id,name,review\_date,review\_text,review\_star from (

select a.business\_id,b.name,a.review\_date,a.review\_text,a.review\_star,

row\_number() over (partition by a.business\_id order by review\_date DESC) AS rnk

from tbl\_yelp\_reviews a

join tbl\_yelp\_business b

on a.business\_id=b.business\_id

)

where rnk<=3

order by name

--5 Find the month with the highest number of reviews

select month(review\_date) as month, count(user\_id) as cnt

from tbl\_yelp\_reviews

group by month

order by cnt DESC

--6. Find the percentage of 5-star reviews for each business

select a.business\_id,name,

count(\*) as ttl\_count,

sum(case when review\_star=5 then 1 else 0 end) as Five\_star

,Five\_star\*100/ttl\_count as perc

from tbl\_yelp\_reviews a

join tbl\_yelp\_business b

on a.business\_id=b.business\_id

group by 1,2

--7 Find the top 5 most reviewed businesses in each city

with cte as (select city,a.business\_id,name,count(\*) as reviews,

row\_number() over (partition by city order by count(\*) DESC) AS rnk

from tbl\_yelp\_reviews as a

join tbl\_yelp\_business as b

on a.business\_id=b.business\_id

group by 1,2,3

)

select city,business\_id,name,reviews,rnk

from cte

where rnk<=5

order by city

--8 Find the average rating of businesses that have atleast 100 reviews

select a.business\_id,name,count(\*) as ttl\_review\_cnt,avg(review\_star) as avg\_rating

from tbl\_yelp\_reviews as a

join tbl\_yelp\_business as b

on a.business\_id=b.business\_id

group by 1,2

having ttl\_review\_cnt>=100

select \* from tbl\_yelp\_reviews limit 10;

--9 List the top 10 users who have written the most review, along with the business they reviewed

WITH CTE AS (select

user\_id,count(\*) as ttl\_reviews,

row\_number() over(order by count(\*) DESC) AS RNK

from tbl\_yelp\_reviews

group by 1

order by rnk

limit 10

)

SELECT user\_id,business\_id from tbl\_yelp\_reviews

where user\_id in(select user\_id from cte)

order by user\_id

--10 Find the top 10 businesses with highest positive sentiment reviews.

select a.business\_id,name,count(\*) as total\_reviews

from tbl\_yelp\_reviews as a

join tbl\_yelp\_business as b

on a.business\_id=b.business\_id

where sentiments='Positive'

group by 1,2

order by 3 desc

limit 10