create TABLE AppleStore\_description\_combined AS

SELECT \* FROM appleStore\_description1

union ALL

SELECT \* FROM appleStore\_description2

UNION ALL

SELECT \* FROM appleStore\_description3

UNION all

SELECT \* FROM appleStore\_description4

--Check the number of unique apps in tables in the AppleStore.

SELECT count(DISTINCT id) AS UniqueAppIDs

FROM AppleStore

SELECT count(DISTINCT id) AS UniqueAppIDs

FROM AppleStore\_description\_combined

--check for missing values

SELECT COUNT(\*) as MissingValues

FROM AppleStore

where track\_name IS null or user\_rating is NULL or prime\_genre is NULL

SELECT COUNT(\*) as MissingValues

FROM AppleStore\_description\_combined

where app\_desc IS null

--Find number of apps per genre.

SELECT prime\_genre, count(\*) as NumApps

from AppleStore

group by prime\_genre

order by NumApps DESC

--Get an overview of app ratings

SELECT

min(user\_rating) as MinRating,

max(user\_rating) as MaxRating,

avg(user\_rating) as AvgRating

FROM AppleStore

--Determine whether paid apps have higher ratings than free apps in the AppleStore.

SELECT CASE

when price > 0 then 'Paid'

else 'Free'

end as App\_Type,

avg(user\_rating) as Avg\_Rating

FROM AppleStore

GROUP by App\_Type

--Determine if apps with more languages have higher ratings?

SELECT CASE

when lang\_num < 10 then '<10 languages'

when lang\_num BETWEEN 10 and 30 THEN '10-30 languages'

else '>30 languages'

end as Language\_Bucket,

avg(user\_rating) as Avg\_Rating

from AppleStore

group BY Language\_Bucket

order by Avg\_Rating desc

-- Which genres have the lowest ratings?

SELECT prime\_genre,

avg(user\_rating) as Avg\_Rating

from AppleStore

group by prime\_genre

order by Avg\_Rating aSC

limit 10

--Is there any correlation between app rating and length of the description in AppleStore?

SELECT CASE

when length(b.app\_desc) < 500 then 'Short'

when length(b.app\_desc) between 500 and 1000 then 'Medium'

ELSE 'Long'

end as descrition\_length\_bucket,

avg(a.user\_rating) as Average\_Rating

from AppleStore as A

Join AppleStore\_description\_combined as B

ON A.id=B.id

group by descrition\_length\_bucket

order by Average\_Rating desc

-- What are the top rated apps for each genre?

SELECT

prime\_genre,

track\_name,

user\_rating

from (

SELECT

prime\_genre,

track\_name,

user\_rating,

rank() over(partition by prime\_genre order by user\_rating desc, rating\_count\_tot desc) as rank

FROM AppleStore

) as A

WHERE a.rank= 1