Using SQL workbench

1. SELECT DISTINCT prime\_genre

FROM IHAMSFT.app;

1. SELECT prime\_genre, sum(rating\_count\_tot) as Total\_ratings

FROM IHAMSFT.app

GROUP BY prime\_genre

ORDER BY Total\_ratings DESC

LIMIT 1;

1. SELECT DISTINCT prime\_genre, count(\*) as no\_apps

FROM IHAMSFT.app

GROUP BY prime\_genre

ORDER BY no\_apps DESC  
LIMIT 1;

1. SELECT DISTINCT prime\_genre, count(\*) as no\_apps

FROM IHAMSFT.app

GROUP BY prime\_genre

ORDER BY no\_apps ASC

LIMIT 1;

1. SELECT track\_name, rating\_count\_tot as Total\_ratings

FROM IHAMSFT.app

ORDER BY Total\_ratings DESC

LIMIT 10;

1. SELECT track\_name, user\_rating\_ver

FROM IHAMSFT.app

ORDER BY user\_rating\_ver DESC

LIMIT 10;

1. Social networking, music and games seem to be high on the list.
2. There are a lot more than 10 apps with a 5 start rating. Not completely clear how it decides which apps are shown in the top 10.
3. There’s not a great deal of crossover between the two. Of the top 10 most rated apps, only the bible app has a 5 star rating.
4. Find a way to combine the scores into a single expression. This could be multiplication or use powers. In the case of multiplication, you may want to weight the scores first somehow. It may take some experimentation to find the correct balance.
5. Firstly, I noticed price doesn’t sort properly so my first query was:  
   SELECT track\_name, rating\_count\_tot, price, user\_rating\_ver

FROM IHAMSFT.app

ORDER BY length(price) DESC, price DESC;  
  
I also sorted in the opposite direction.

From this, I could see that free apps receive considerably more reviews than paid apps do. There wasn’t an obvious consistent increase in downloads as price got cheaper but, when I copied the data and plotted a scatter plot in excel, it could be seen that cheaper apps tended to get more reviews. It would be interesting to see how much of a difference simply having to pay for an app makes, even a very nominal fee.

It may also be that some of the more expensive apps are expensive because they have a narrow target audience.

Reviews scores do not appear to be as affected. This could be that the paid apps, while better, have higher expectations to meet or may simply reflect. I don’t think I could currently prove this theory one way or another.

I did attempt to find ways to group the prices to run some stats however it was getting complicated with the aforementioned issues sorting price.