## Compiling the Bright TV case study

Two table were joined on full outer join so that we can have the relevant information on the viewers watching. Data set had to be cleaned as it also contained user profile information but some of those viewers have not watched any channel. Decided to exclude the users name, surname, email and addresses as that wasn't going to provident any significant value to the insights.

Syntax for Full Outer Join to join the two tables (viewership and the user profile)

```
---JOINING USERPROFILE TABLE AND VIEWERS TABLE

SELECT A.USERID,

GENDER,

RACE,

AGE,

PROVINCE,

B.CHANNEL,

"RecordDate2 (UTC)",

DAY,

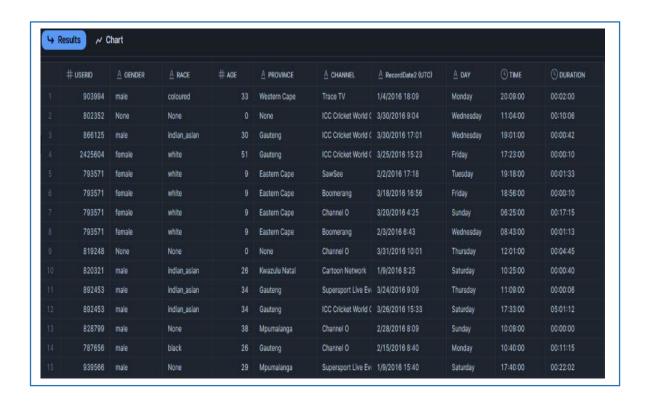
TIME,

DURATION,

FROM USER_PROFILE AS A

FULL OUTER JOIN VIEWERS AS B

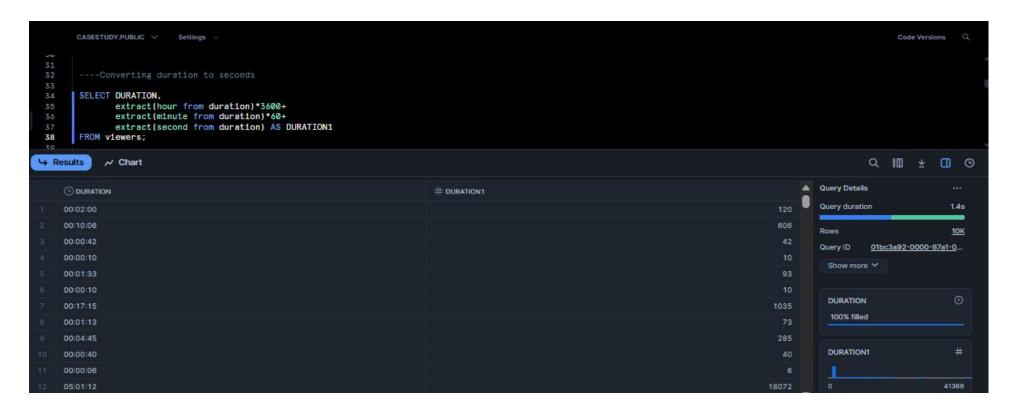
ON A.USERID = B.USERID;
```



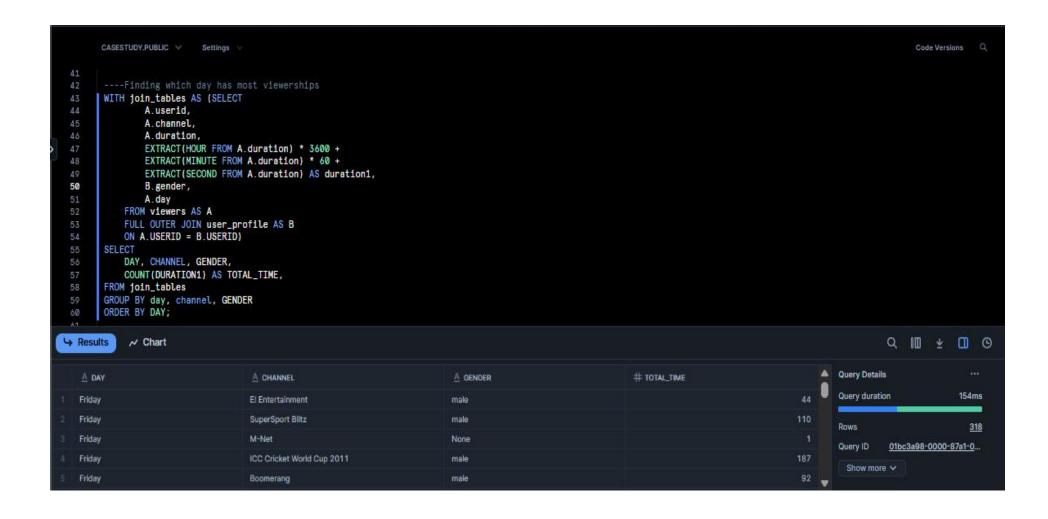
To determine the total viewers subscribed to the Bright TV Channels



How to change the hh:mm:ss format to seconds to calculate the total viewership time



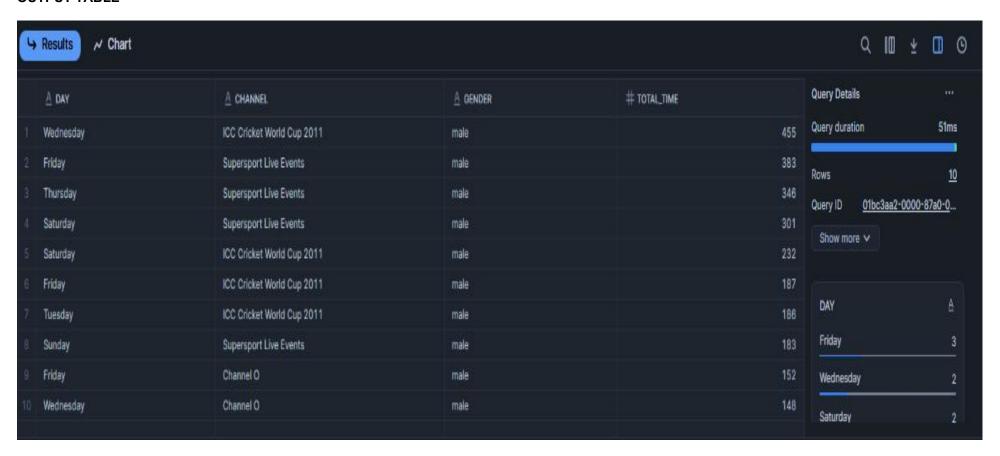
To establish which weekday has most viewers (split by gender), what channels they watching and how much time they spending on each channel.



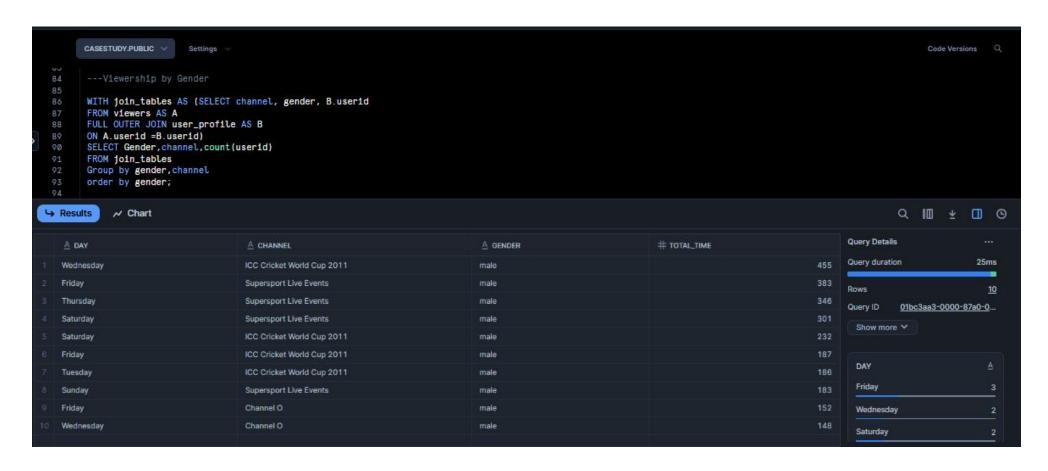
Establishing weekday with most viewers limit it to top ten viewers who sent most time viewing

```
CASESTUDY.PUBLIC V
                         Settings
      ----Finding which day has most viewerships, LIMIT TO TEN
      WITH join_tables AS (SELECT
              A.userid,
              A. channel,
              A. duration,
              EXTRACT(HOUR FROM A.duration) * 3600 +
              EXTRACT(MINUTE FROM A.duration) * 60 +
              EXTRACT(SECOND FROM A.duration) AS duration1,
              B. gender,
              A. day
          FROM viewers AS A
          FULL OUTER JOIN user_profile AS B
          ON A.USERID = B.USERID)
      SELECT
          DAY, CHANNEL, GENDER,
          COUNT(DURATION1) AS TOTAL_TIME,
      FROM join_tables
      GROUP BY day, channel, GENDER
      ORDER BY TOTAL_TIME DESC
80
      LIMIT 10;
81
```

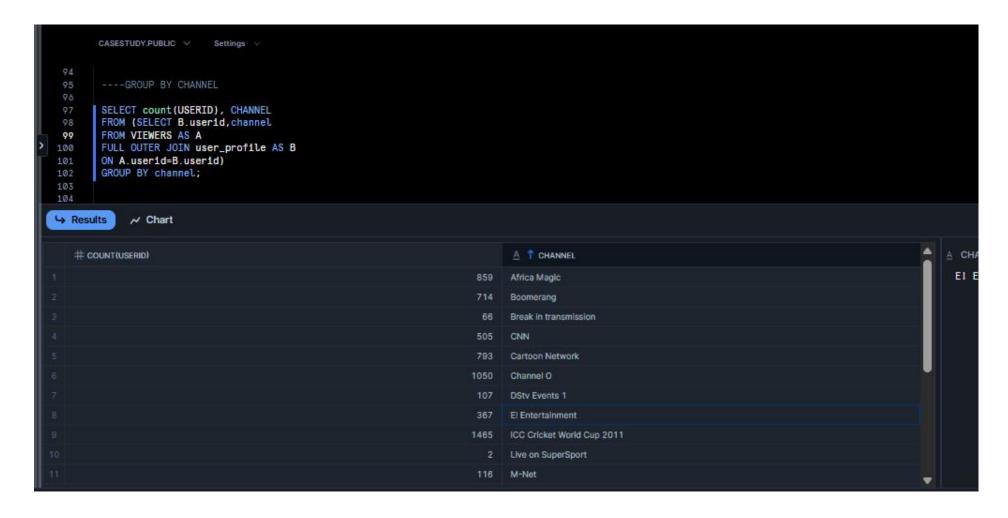
## **OUTPUT TABLE**



Determining viewership by gender and which gender spend most of they time watching



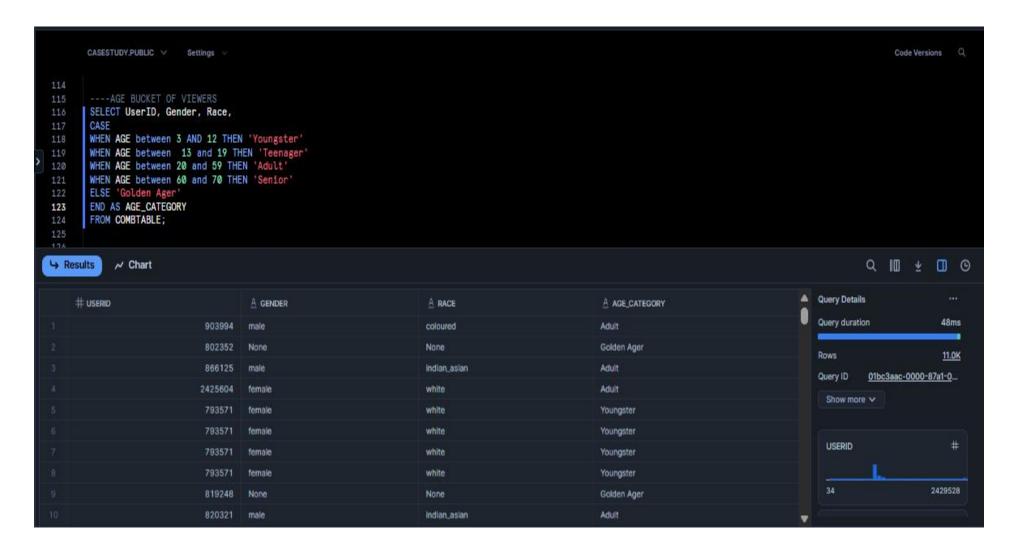
Group the information by channel to see which channel are watched by the viewers



How many Null values are in this data set



## Group by the viewer per age bucket



## Total viewers per province

