

BRIGHT TV

SELECT * STATEMENT

Used the query below to retrieve all the columns and rows from both the User Profile and Viewership tables

```
1 SELECT
2 *
3 FROM BRIGHT.PUBLIC.USER_PROFILES
4 LIMIT 10;
```

#	USERID	NAME	SURNAME	EMAIL	GENDER	RACE	AGE	PROVINCE	SOCIAL_MEDIA_HANDLE
1	34	Chas	Biggers	Chas 351@yahoo.com	male	white	55	Free State	@Chas 35
2	39	Wilburn	Crotty	Wilburn 303@gmail.com	male	black	54	Eastern Cape	@Wilburn 30
3	63	Emerson	Slane	Emerson 243@gmail.com	male	black	51	Kwazulu Natal	@Emerson 24
4	90	Stacy	Sytma	Stacy 273@gmail.com	male	coloured	49	Kwazulu Natal	@Stacy 27
5	97	Lucas	Woolard	Lucas 233@gmail.com	male	coloured	48	Free State	@Lucas 23
6	104	Randy	Tancredi	Randy 393@gmail.com	male	black	47	Gauteng	@Randy 39
7	107	Trudie	Restivo	Trudie 502@gmail.com	female	black	47	Gauteng	@Trudie 50
8	108	Chas	Luthy	Chas 352@gmail.com	male	coloured	47	Western Cape	@Chas 35
9	120	Jared	Bilyeu	Jared 503@gmail.com	male	coloured	46	Western Cape	@Jared 50
10	132	Carter	Wigfall	Carter 33@gmail.com	male	black	45	Gauteng	@Carter 3

```
8 SELECT
9 *
10 FROM BRIGHT.PUBLIC.VIEWERSHIP
11 LIMIT 10;
```

#	USERID	CHANNEL2	RECORDDATE2	RECORDTIME2	DURATION_2	USER_ID
1	903994	Trace TV	4/1/2016	20:09:00	00:02:00	903994
2	802352	ICC Cricket World Cup 2011	30/3/2016	11:04:00	00:10:06	802352
3	866125	ICC Cricket World Cup 2011	30/3/2016	19:01:00	00:00:42	866125
4	2425604	ICC Cricket World Cup 2011	25/3/2016	17:23:00	00:00:10	2384970
5	793571	SawSee	2/2/2016	19:18:00	00:01:33	793571
6	793571	Boomerang	18/3/2016	18:56:00	00:00:10	793571
7	793571	Channel O	20/3/2016	06:25:00	00:17:15	793571
8	793571	Boomerang	3/2/2016	08:43:00	00:01:13	793571
9	819248	Channel O	31/3/2016	12:01:00	00:04:45	819248
10	820321	Cartoon Network	9/1/2016	10:25:00	00:00:40	820321

Used the **FULL OUTER JOIN** to retrieve all records from both tables.

```
15 -----FULL OUTER JOINS-----
16 SELECT
17 *
18 FROM USER_PROFILES AS U
19 FULL OUTER JOIN VIEWERSHIP AS V
20 ON U.userid = V.userid;
```

#	USERID	NAME	SURNAME	EMAIL	GENDER	RACE	AGE	PROVINCE	SOCIAL_MEDIA_HANDLE	USERID	CHANNEL2	RECORDDATE2	RECORDTIME2	DURATION_2
1	903994	Bernardo	Hier	Bernardo 323@gmail.com	male	coloured	33	Western Cape	@Bernardo 32	903994	Trace TV	4/1/2016	20:09:00	00:02:00
2	802352	None	None	Shelley 34@abcmall.com	None	None	0	None	None	802352	ICC Cricket World Cup 2011	30/3/2016	11:04:00	00:10:06
3	866125	Emerson	Mccollum	Emerson 243@gmail.com	male	indian_asian	30	Gauteng	@Emerson 24	866125	ICC Cricket World Cup 2011	30/3/2016	19:01:00	00:00:42
4	2425604	Tiffani	Pilot	Tiffani 433@gmail.com	female	white	51	Gauteng	@Tiffani 43	2425604	ICC Cricket World Cup 2011	25/3/2016	17:23:00	00:00:10
5	793571	Shelley	Reisinger	Shelley 234@abcmall.com	female	white	9	Eastern Cape	@Shelley 23	793571	SawSee	2/2/2016	19:18:00	00:01:33
6	793571	Shelley	Reisinger	Shelley 234@abcmall.com	female	white	9	Eastern Cape	@Shelley 23	793571	Boomerang	18/3/2016	18:56:00	00:00:10
7	793571	Shelley	Reisinger	Shelley 234@abcmall.com	female	white	9	Eastern Cape	@Shelley 23	793571	Channel O	20/3/2016	06:25:00	00:17:15
8	793571	Shelley	Reisinger	Shelley 234@abcmall.com	female	white	9	Eastern Cape	@Shelley 23	793571	Boomerang	3/2/2016	08:43:00	00:01:13
9	819248	None	None	Bess 381@yahoo.com	None	None	0	None	None	819248	Channel O	31/3/2016	12:01:00	00:04:45
10	820321	Haywood	Singer	Haywood 92@gmail.com	male	indian_asian	26	Kwazulu Natal	@Haywood 9	820321	Cartoon Network	9/1/2016	10:25:00	00:00:40
11	892453	Eli	Caves	Eli 291@yahoo.com	male	indian_asian	34	Gauteng	@Eli 29	892453	Supersport Live 1	24/3/2016	11:09:00	00:00:00

BRIGHTLIGHT DATA ANALYTICS

The **COALESCE** FUNCTION was used to handle NULL values present in the data

```
24 -----NULL VALUES USING COALESCE FUNCTION-----
25 SELECT
26     COALESCE(gender, 'Unknown') AS Gender,
27     COALESCE(race, 'Unknown') AS Race,
28     COALESCE(province, 'Unknown') AS Province,
29     COUNT(*) AS user_count
30 FROM BrightLight_TV
31 GROUP BY Gender, Race, Province
32 ORDER BY user_count DESC;
33
34
35
```

	GENDER	RACE	PROVINCE	# USER_COUNT
1	male	black	Gauteng	1680
2	male	coloured	Western Cape	1002
3	None	None	None	787
4	male	indian_asian	Gauteng	652
5	male	white	Gauteng	601
6	male	black	Mpumalanga	550
7	male	black	Kwazulu Natal	420
8	male	black	Limpopo	329
9	male	indian_asian	Limpopo	315
10	male	indian_asian	Kwazulu Natal	278

A new table was created to after performing the FULL OUTER JOIN to store the combined results from both source tables, including the matching and non-matching records.

This allows for easier access, analysis, and further processing of the complete dataset without rerunning the join repeatedly.

```
37 CREATE OR REPLACE TABLE BrightLight_TV AS(
38     SELECT
39         V.user_id,
40         name,
41         surname,
42         COALESCE(gender, 'Unknown') AS Gender,
43         COALESCE(race, 'Unknown') AS Race,
44         email,
45         age,
46         COALESCE(province, 'Unknown') AS Province,
47         social_media_handle,
48         channel2,
49         recorddate2,
50         recordtime2,
51         duration_2,
52
53         --- Convert hh:mm:ss to total_seconds
54         SPLIT_PART(duration_2, ':',1)::INT * 3600 +
55         SPLIT_PART(duration_2, ':',2)::INT * 60 +
56         SPLIT_PART(duration_2, ':',3)::INT AS Duration_seconds,
57
58         --- Use the formula above for watch_time
59         FLOOR(
60             (SPLIT_PART(duration_2, ':',1)::INT * 3600 +
61              SPLIT_PART(duration_2, ':',2)::INT * 60 +
62              SPLIT_PART(duration_2, ':',3)::INT) /60
63         ) || 'min' AS Watch_time
64     FROM USER_PROFILES AS U
65     FULL OUTER JOIN VIEWERSHIP AS V
66     ON U.userID = V.userID);
67
68
69
```

	status
1	Table BRIGHTLIGHT_TV successfully created.

Results from the New table created

```

63 -----
64 | SELECT * FROM BrightLight_TV;
65 -----
66

```

	# USER_ID	NAME	SURNAME	GENDER	RACE	EMAIL	AGE	PROVINCE	SOCIAL_MEDIA_HANDLE	CHANNEL2	RECORDDATE2
1	903994	Bernardo	Hier	male	coloured	Bernardo 323@mail.com	33	Western Cape	@Bernardo 32	Trace TV	4/1/2016
2	802352	None	None	None	None	Shelley 34@abcmail.com	0	None	None	ICC Cricket Work	30/3/2016
3	866125	Emerson	Mccollum	male	indian_asian	Emerson 243@mail.com	30	Gauteng	@Emerson 24	ICC Cricket Work	30/3/2016

Ran this query to verify if the date is formatted correctly and displays as expected for consistency, and accuracy reporting requirements

```

63 -----
64 | SELECT DISTINCT recorddate2 FROM BrightLight_TV
65 | LIMIT 10;
66 -----
67

```

	RECORDDATE2
1	4/1/2016
2	25/3/2016
3	29/3/2016
4	20/3/2016
5	3/2/2016
6	31/3/2016
7	28/2/2016
8	26/2/2016
9	10/2/2016
10	18/3/2016

Altered the new table to store a formatted version of an existing date for display and reporting

```

68 -----
69 | ALTER TABLE BrightLight_TV ADD COLUMN formatted_recorddate2 STRING;
70 -----

```

	status
1	Statement executed successfully.

Changed the data type from DD/MM/YYYY to YYYY/MM/DD to standardize the date format for international compatibility, improve sorting and comparison accuracy and align with ISO 8601 standards

```

76 -----
77 UPDATE BrightLight_TV
78 SET formatted_recorddate2 = TO_CHAR(TO_DATE(recorddate2, 'DD/MM/YYYY'), 'YYYY/MM/DD');
79 -----
80

```

Results		Chart
# number of rows updated	10989	# number of multi-joined rows updated
1		0

Results after running the query to compare the dates.

```

/8 -----
79 SELECT recorddate2, formatted_recorddate2
80 FROM BrightLight_TV
81 LIMIT 10;
82 -----
83
84

```

	RECORDDATE2	FORMATTED_RECORDDATE2
1	4/1/2016	2016/01/04
2	30/3/2016	2016/03/30
3	30/3/2016	2016/03/30
4	25/3/2016	2016/03/25
5	2/2/2016	2016/02/02
6	18/3/2016	2016/03/18
7	20/3/2016	2016/03/20
8	3/2/2016	2016/02/03
9	31/3/2016	2016/03/31
10	9/1/2016	2016/01/09

Used the query below to delete the old column

```

84 -----
85 ALTER TABLE BrightLight_TV DROP COLUMN recorddate2;
86 -----

```

Results		Chart
status		
1	Statement executed successfully.	

Ran the below Query to rename the new column name

```
90 ALTER TABLE BrightLight_TV RENAME COLUMN formatted_recorddate2 TO recorddate2;
```

Results Chart

status

1 Statement executed successfully.

Results after all the altering done

```
93 SELECT * FROM BRIGHTLIGHT_TV;
```

#	USER_ID	NAME	SURNAME	GENDER	RACE	EMAIL	AGE	PROVINCE	SOCIAL_MEDIA_HANDLE	CHANNEL2	RECORDTIME2	DURATION_2	RECORDDATE2
1	903994	Bernardo	Hier	male	coloured	Bernardo 323@mail.com	33	Western Cape	@Bernardo 32	Trace TV	20:09:00	00:02:00	2016/01/04
2	802352	None	None	None	None	Shelley 34@abcmail.com	0	None	None	ICC Cricket Work	11:04:00	00:10:06	2016/03/30
3	866125	Emerson	Mccollum	male	indian_asian	Emerson 243@mail.com	30	Gauteng	@Emerson 24	ICC Cricket Work	19:01:00	00:00:42	2016/03/30
4	2384970	Tiffani	Pilot	female	white	Tiffani 433@mail.com	51	Gauteng	@Tiffani 43	ICC Cricket Work	17:23:00	00:00:10	2016/03/25
5	793571	Shelley	Reisinger	female	white	Shelley 234@abcmail.com	9	Eastern Cape	@Shelley 23	SawSee	19:18:00	00:01:33	2016/02/02
6	793571	Shelley	Reisinger	female	white	Shelley 234@abcmail.com	9	Eastern Cape	@Shelley 23	Boomerang	18:56:00	00:00:10	2016/03/18
7	793571	Shelley	Reisinger	female	white	Shelley 234@abcmail.com	9	Eastern Cape	@Shelley 23	Channel O	06:25:00	00:17:15	2016/03/20
8	793571	Shelley	Reisinger	female	white	Shelley 234@abcmail.com	9	Eastern Cape	@Shelley 23	Boomerang	08:43:00	00:01:13	2016/02/03
9	819248	None	None	None	None	Bess 381@yahoo.com	0	None	None	Channel O	12:01:00	00:04:45	2016/03/31
10	820321	Haywood	Singer	male	indian_asian	Haywood 92@gmail.com	26	Kwazulu Natal	@Haywood 9	Cartoon Network	10:25:00	00:00:40	2016/01/09
11	892453	Eli	Caves	male	indian_asian	Eli 291@yahoo.com	34	Gauteng	@Eli 29	Supersport Live I	11:09:00	00:00:06	2016/03/24
12	892453	Eli	Caves	male	indian_asian	Eli 291@yahoo.com	34	Gauteng	@Eli 29	ICC Cricket Work	17:33:00	05:01:12	2016/03/26
13	822560	Romeo	Mastrangelo	male	None	Romeo 375@mailxyz.com	38	Mpumalanga	@Romeo 37	Channel O	10:09:00	00:00:00	2016/02/28
14	787656	Garry	Sytma	male	black	Garry 284@abcmail.com	26	Gauteng	@Garry 28	Channel O	10:40:00	00:11:15	2016/02/15

Used the CASE Statement to perform conditional logic that allows different output for Time Slots, Age Group and Month to enable categorization within a Query.

The results from this Query are used as the primary dataset for creating visualisations, ensuring the data is pre-processed, categorized or formatted appropriately for accurate and meaningful graphical representation

```

109 SELECT
110 *
111 CASE
112 WHEN recordtime2 BETWEEN '00:00:01' AND '04:00:00' THEN '12AM-4AM'
113 WHEN recordtime2 BETWEEN '04:00:01' AND '08:00:00' THEN '4AM-8AM'
114 WHEN recordtime2 BETWEEN '12:00:01' AND '16:00:00' THEN '12PM-4PM'
115 WHEN recordtime2 BETWEEN '16:00:01' AND '20:00:00' THEN '4PM-8PM'
116 ELSE '8PM-12AM'
117 END AS Time_Slot,
118 CASE
119 WHEN age BETWEEN 0 AND 1 THEN 'Infancy'
120 WHEN age BETWEEN 2 AND 4 THEN 'Toddler'
121 WHEN age BETWEEN 5 AND 12 THEN 'Child'
122 WHEN age BETWEEN 13 AND 19 THEN 'Adolescence'
123 WHEN age BETWEEN 20 AND 39 THEN 'Young Adult'
124 WHEN age BETWEEN 40 AND 59 THEN 'Middle Adult'
125 ELSE 'Senior Citizen'
126 END AS Age_group,
127 CASE
128 WHEN TRY_TO_DATE(formatted_recorddate2, 'YYYY/MM/DD') BETWEEN '2016-01-01' AND '2016-01-31' THEN 'JAN'
129 WHEN TRY_TO_DATE(formatted_recorddate2, 'YYYY/MM/DD') BETWEEN '2016-02-01' AND '2016-02-29' THEN 'FEB'
130 WHEN TRY_TO_DATE(formatted_recorddate2, 'YYYY/MM/DD') BETWEEN '2016-03-01' AND '2016-03-31' THEN 'MAR'
131 ELSE 'APR'
132 END AS Record_month
133 FROM BrightLight_TV;
134

```

#	USER_ID	NAME	SURNAME	GENDER	RACE	EMAIL	AGE	PROVINCE	SOCIAL_MEDIA_HANDLE	CHANNEL2	RECORDTIME2	DURATION_2	DURATION_SECONDS	WATCH_TIME	FORM
1	903994	Bernardo	Hier	male	coloured	Bernardo 323@gmail.com	33	Western Cape	@Bernardo 32	Trace TV	20:09:00	00:02:00	120	2min	2016/01
2	802352	None	None	None	None	Shelley 34@abcmail.com	0	None	None	ICC Cricket Work	11:04:00	00:10:06	606	10min	2016/03
3	866125	Emerson	Mccollum	male	Indian,Asian	Emerson 243@gmail.com	30	Gauteng	@Emerson 24	ICC Cricket Work	19:01:00	00:00:42	42	0min	2016/03