

TAFADZWA MULAUSI

METHODOLOGY

BRIGHT TV

The viewership table contained two userid columns; therefore, I removed the duplicate column.

```
--delete the userid column since we have two of them
Alter table viewership
Drop column "userid";
```

The viewership and user_profile tables were joined to determine the total number of records resulting from the combined dataset.

```
select
    count (userid)as total_viewers
from user_profile as a
full outer join viewership as b
on a.userid = b."UserID";
```

Results Chart

| | # TOTAL_VIEWERS |
|---|-----------------|
| 1 | 10989 |

This is to identify the number of unique viewers present in the joined dataset.

```
select
    count (distinct userid)as total_viewers
from user_profile as a
full outer join viewership as b
on a.userid = b."UserID";
```

Results Chart

| | # TOTAL_VIEWERS |
|---|-----------------|
| 1 | 5375 |

This analysis aims to identify the channel with the largest viewer base as well as the one with the smallest, based on the joined dataset

```
select
    distinct a.channel2 ,
    count(userid)as total_viewers
from viewership a
join user_profile b
on a."UserID" = b.userid
group by channel2
order by total_viewers desc;
```

| Results | | | Chart |
|---------|----------------------------|---------------|-------|
| | CHANNEL2 | TOTAL_VIEWERS | |
| 1 | Supersport Live Events | 1638 | |
| 2 | ICC Cricket World Cup 2011 | 1465 | |
| 3 | Channel O | 1050 | |
| 4 | Trace TV | 952 | |
| 5 | SuperSport Blitz | 896 | |
| 6 | Africa Magic | 859 | |
| 7 | Cartoon Network | 793 | |
| 8 | Boomerang | 714 | |
| 9 | CNN | 505 | |
| 10 | E! Entertainment | 367 | |
| 11 | SawSee | 251 | |
| 12 | M-Net | 116 | |
| 13 | Vuzu | 111 | |
| 14 | DStv Events 1 | 107 | |
| 15 | Break in transmission | 66 | |
| 16 | kykNET | 45 | |
| 17 | MK | 32 | |

The data reveals the distribution of viewer activity across the days of the week, identifying the days with the highest viewership.

```
select
    dayname(b.recorddate2) as day_of_week ,
    count(userid) as viewers
from user_profile a
join viewership b
on a.userid = b."UserID"
group by day_of_week,
    dayname(b.recorddate2)
order by viewers desc;
```

| Results | | | Chart |
|---------|-------------|---------|-------|
| | DAY_OF_WEEK | VIEWERS | |
| 1 | Sat | 1655 | |
| 2 | Fri | 1642 | |
| 3 | Wed | 1539 | |
| 4 | Thu | 1471 | |
| 5 | Sun | 1398 | |
| 6 | Tue | 1321 | |
| 7 | Mon | 974 | |

This query is used to determine which race watches the most.

```
select
    b.race ,
    count(*)as total_viewers
from viewership a
join user_profile b
on a."UserID" = b.userid
group by race
order by total_viewers desc;
```

| Results Chart | | |
|---------------|--------------|-----------------|
| | A RACE | # TOTAL_VIEWERS |
| 1 | black | 4331 |
| 2 | coloured | 1633 |
| 3 | indian_asian | 1575 |
| 4 | white | 1292 |
| 5 | None | 1057 |
| 6 | other | 102 |
| 7 | null | 10 |

This query is used to determine which age group watches the most. We grouped the ages into categories to make it easier to compare the viewership across different groups.

```
select
  case
    when age = 0 then 'not applicable'
    when age between 1 and 12 then 'kids'
    when age between 13 and 19 then 'teenager'
    when age between 20 and 35 then 'youth'
    when age between 36 and 50 then 'adult'
    when age between 51 and 65 then 'mature adult'
    else 'retired'
  end as age_group ,
  count(*) as total_viewers
from user_profile a
join viewership b
on a.userid = b."UserID"
group by all
order by total_viewers desc;
```

| Results Chart | | |
|---------------|----------------|-----------------|
| | A AGE_GROUP | # TOTAL_VIEWERS |
| 1 | youth | 5733 |
| 2 | adult | 2972 |
| 3 | mature adult | 452 |
| 4 | teenager | 436 |
| 5 | not applicable | 260 |
| 6 | kids | 99 |
| 7 | retired | 48 |

This query is used to find out how many channels were watched by each gender, based on the combined tables

```
select
  b.gender ,
  count(a.channel2)as channels_watched
from viewership a
join user_profile b
on a."UserID" = b.userid
group by b.gender
order by channels_watched desc;
```

| Results | | Chart | | |
|---------|--------|-------|------------------|------|
| | GENDER | | CHANNELS_WATCHED | |
| 1 | male | | | 8761 |
| 2 | female | | | 977 |
| 3 | None | | | 262 |

This query shows how many total viewing records come from each province, helping you see where most of your audience is located.

```
select
    b.province ,
    count(*)as total_viewers
from viewership a
join user_profile b
on a."UserID" = b.userid
group by province
order by total_viewers desc;
```

| Results | | Chart | | |
|---------|---------------|-------|---------------|------|
| | PROVINCE | | TOTAL_VIEWERS | |
| 1 | Gauteng | | | 3654 |
| 2 | Western Cape | | | 1845 |
| 3 | Kwazulu Natal | | | 1001 |
| 4 | Mpumalanga | | | 918 |
| 5 | Limpopo | | | 763 |
| 6 | Eastern Cape | | | 690 |
| 7 | North West | | | 344 |
| 8 | Free State | | | 292 |
| 9 | None | | | 263 |
| 10 | Northern Cape | | | 230 |