**Strategic Analysis of BrightTV: Methods and Insights**

**Busile Ndhlovu**

**Table of Contents**

|  |  |
| --- | --- |
| **Topic** | **Page** |
| Objective  Data Overview | 3  3 |
| Tools and Techniques | 3 |
| Understanding the dataset | 3 |
| Trend Analysis | 11 |
| Insights Drawn from Trend Analysis | 14 |
| Factors influencing consumption | 14 |
| Recommendations to increase consumption on low consumption days | 14 |
| Initiative to grow BrightTvs user base | 15 |

**Objective**

Goal: Increase BrightTv’s subscription base within this financial year.

CEO’s request: Insights for CVM(Customer Value Management) team to achieve the objective.

**Data Overview**

* The dataset attached contains information on the user profiles and viewer transactions for BrightTV.
* The dataset was from the months January, February and March from the year 2016.
* Times and dates in the dataset are supplied in UTC.
* Consumption is split per session, i.e. for every session a subscriber has, there will be 1 record.

**Tools and Techniques**

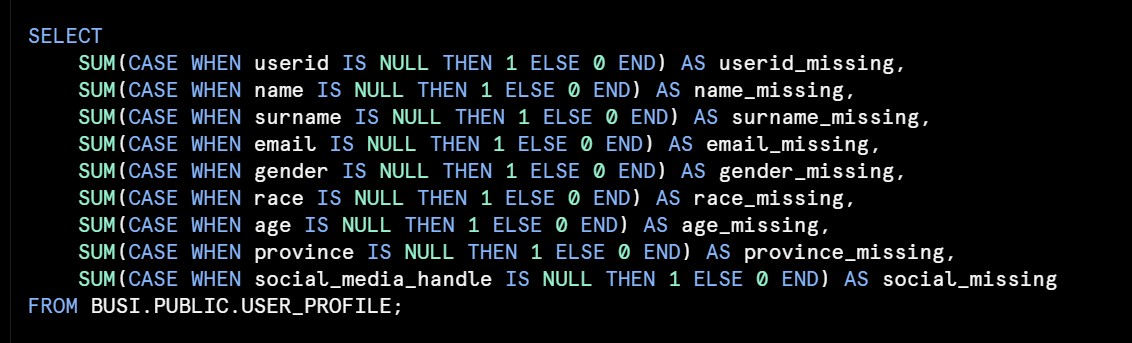
* **Data Analysis Tools**: SQL on Snowflake and MS Excel for statistical analysis and visualization.
* **Visualization Tools**: PowerPoint for visualization.

**Understanding the dataset**

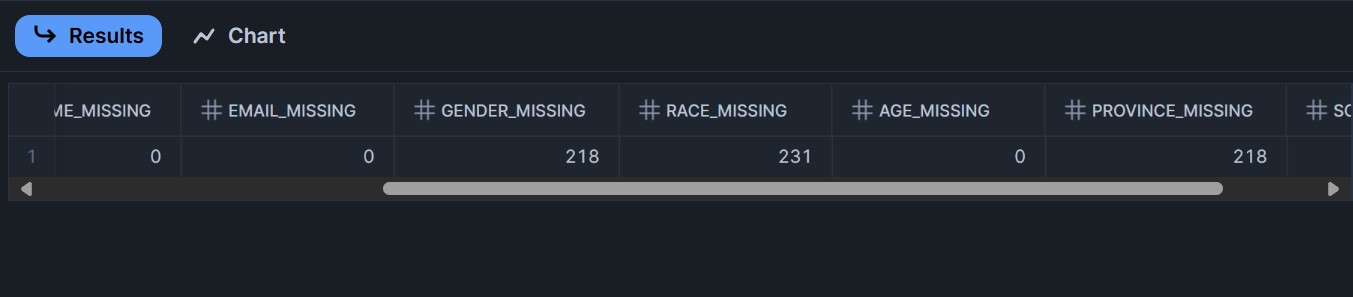
1.Handle Missing Values:

Code used to find missing values in all columns of the user profile table.

Missing values in this case are considered NULL values.



Returned output:



I found that there are three columns with missing values gender, race and province.

Code used to find missing values in all columns of the user Viewership table.

A black screen with white text

AI-generated content may be incorrect.

Returned Output:

A screenshot of a computer

AI-generated content may be incorrect.

Returns that there are no missing values.

2.Replacing missing values:

Code used to replace missing values in the race,gender and province columns on the user profile table.Every missing value was replaced with the word unknown.

A screenshot of a computer

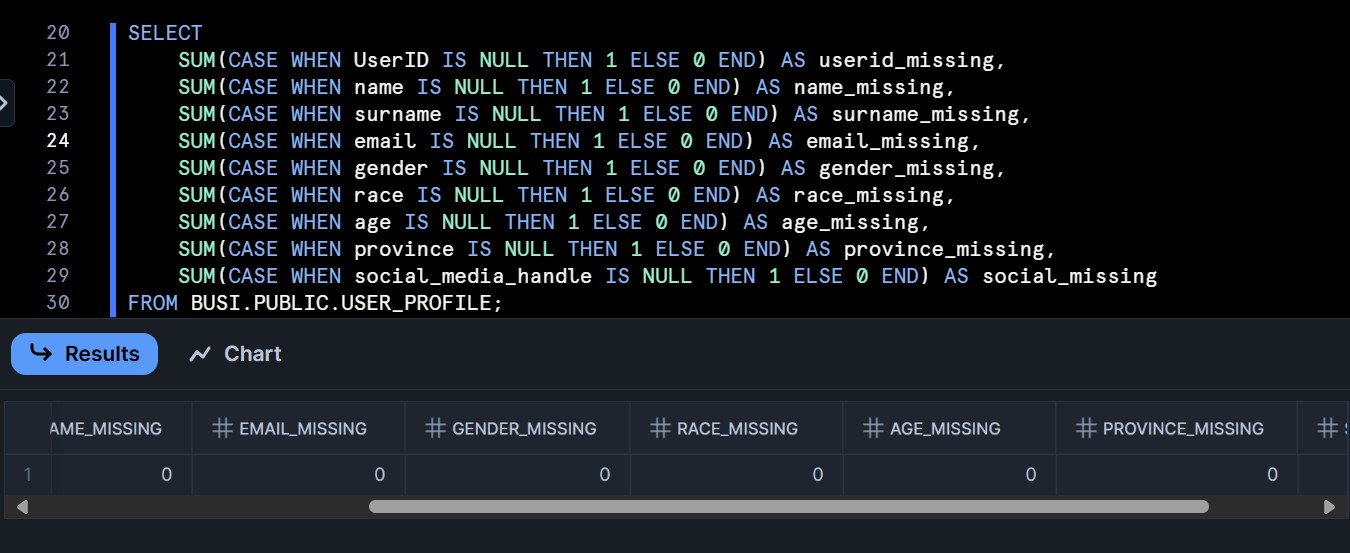
AI-generated content may be incorrect.

A screenshot of a computer

AI-generated content may be incorrect.

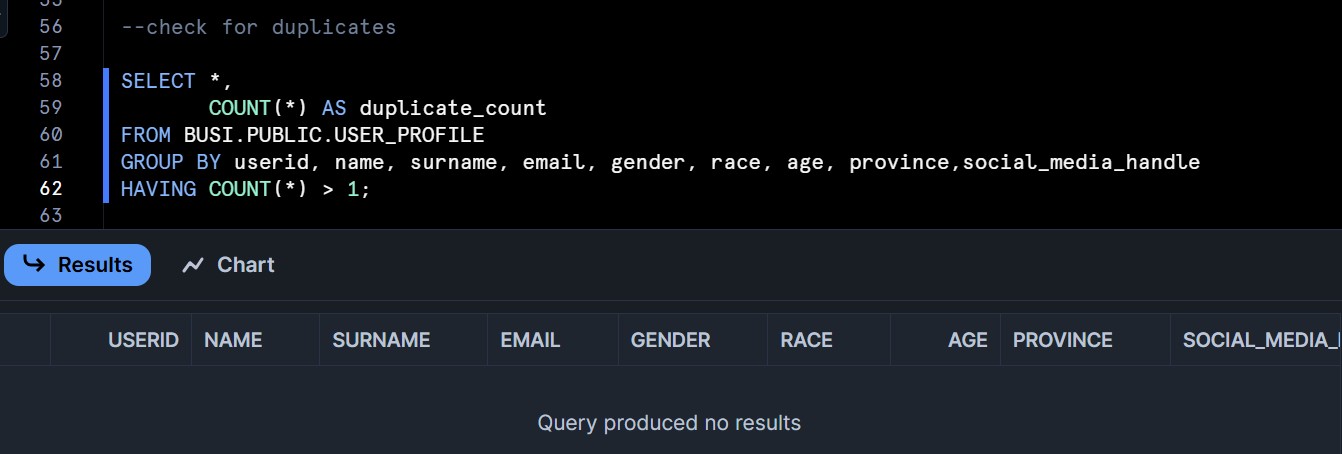


Returned output:



Displays that there are 0 missing values in the race, gender and province tables.

3.Checking for duplicates in the user profile table



Found no duplicates.

Checking for duplicates in the viewership table.

A screenshot of a computer

AI-generated content may be incorrect.

Found that there are duplicates in all columns.

4.Deleting Duplicates

A screenshot of a computer

AI-generated content may be incorrect.

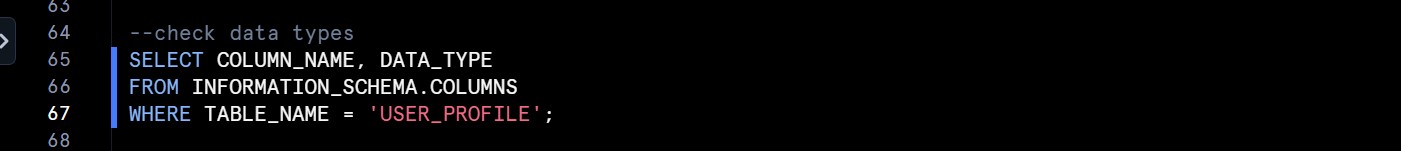
Checking if duplicates are still present:

A screenshot of a computer

AI-generated content may be incorrect.

No duplicates found

5.Check the data types for table user profile

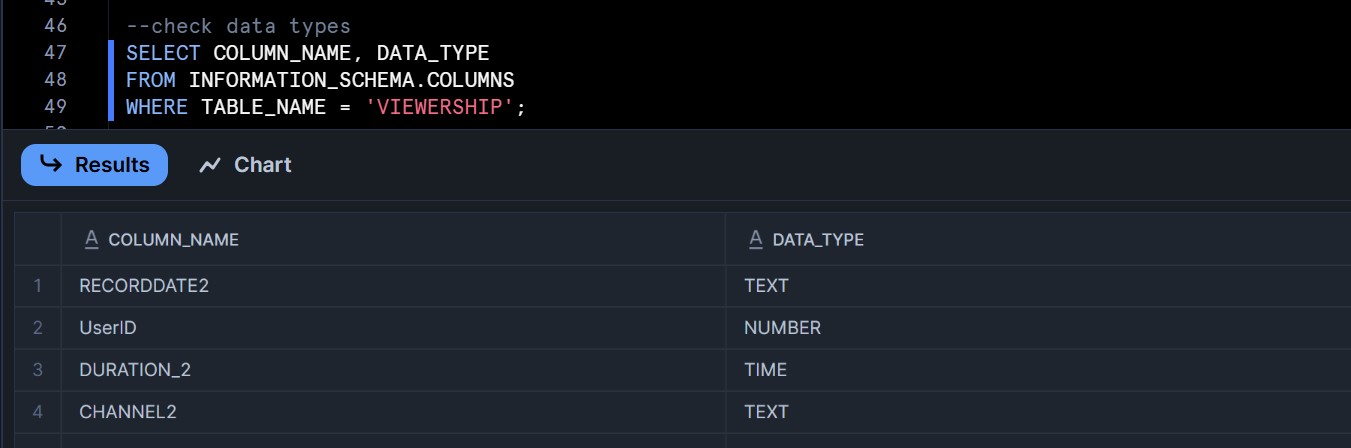


Returned Output:

A screenshot of a computer

AI-generated content may be incorrect.

Check the data types for table viewership



6.Dropping extra column in viewership table

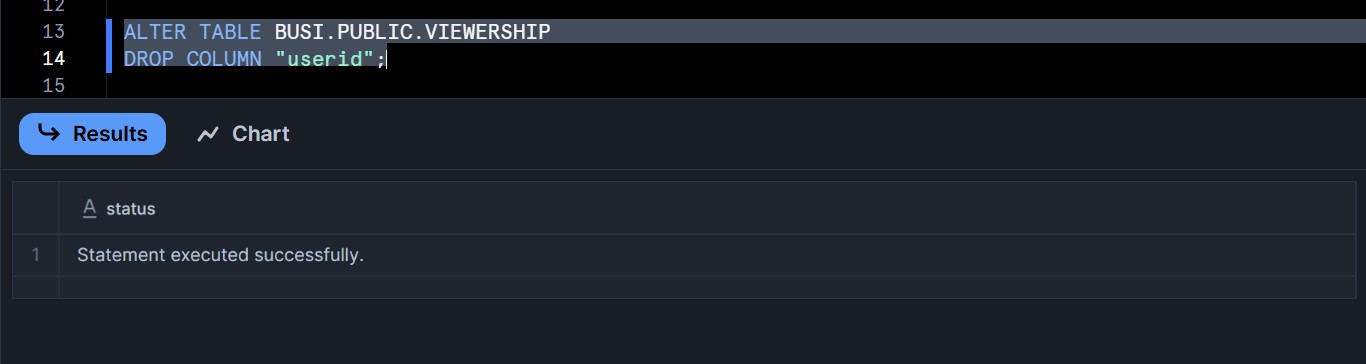
When analysing the data I found that the viewership table has a repeated column called userid



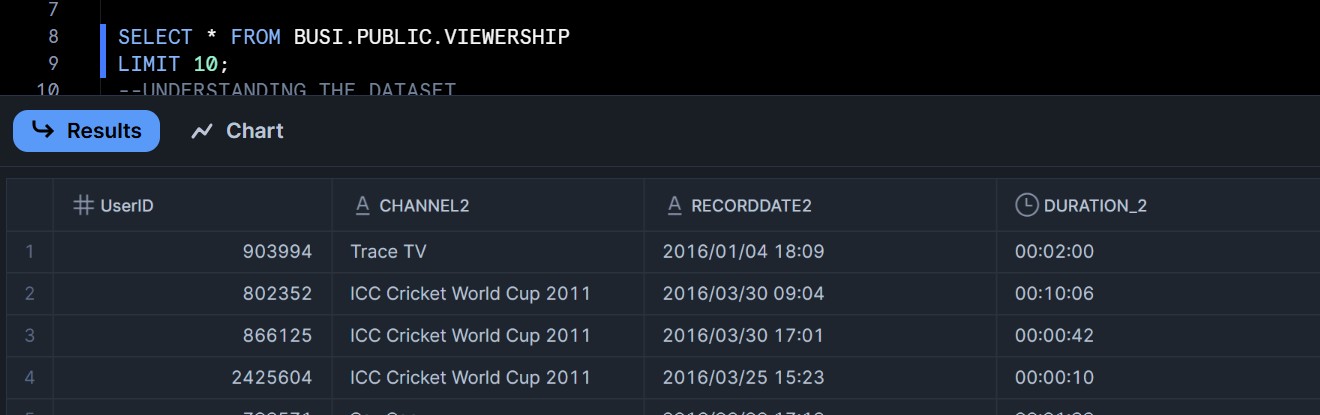
Code used to drop the extra table:

The extra table is being dropped to avoid confusion as there already is a column named userid.

Please note that the reason userid is in parenthesis is due to the auto fix that occurred while uploading the data into snowflake, the system automatically renamed the column with parentheses.



Returned results:



The extra userid column is removed from the table.

7.Joining the two tables using Inner Join

A computer screen with text on it

AI-generated content may be incorrect.

Returned output:

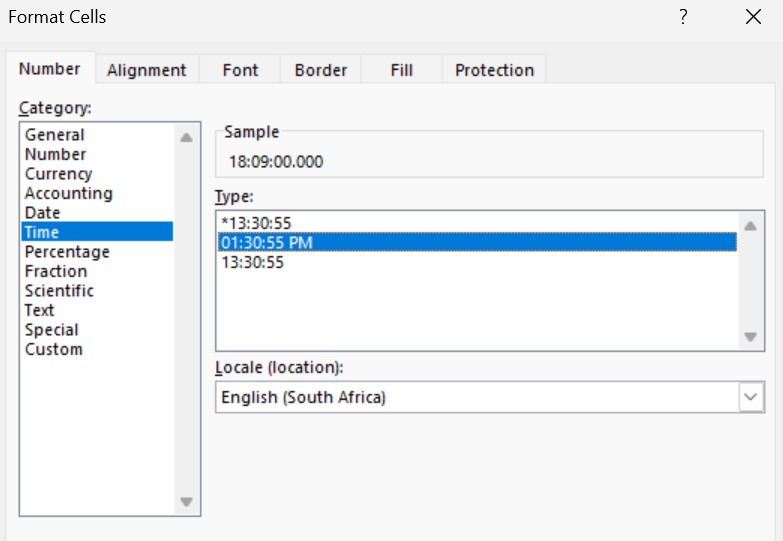
A screenshot of a computer

AI-generated content may be incorrect.

Joined the two tables User Profile and Viewership using the inner join function to filter and combine data from the two tables based on matching value userId. Separated the recorddate2 column which was a timestamp into two separate columns Date and Time to Identify specific hours when consumption is highest and Compare across dates by finding trends over weeks or months. Used the separated Date column to create another column to display days of the week to find peak consumption days.

8.Converting time to SA time

Times and dates in the dataset were supplied in UTC and needed to be changed. To convert the time I extracted the data from snowflake after the inner join of the two tables. And changed it on MS Excel.



9.Age bracket

A computer screen with text and numbers

AI-generated content may be incorrect.

Returned output:

A screenshot of a computer

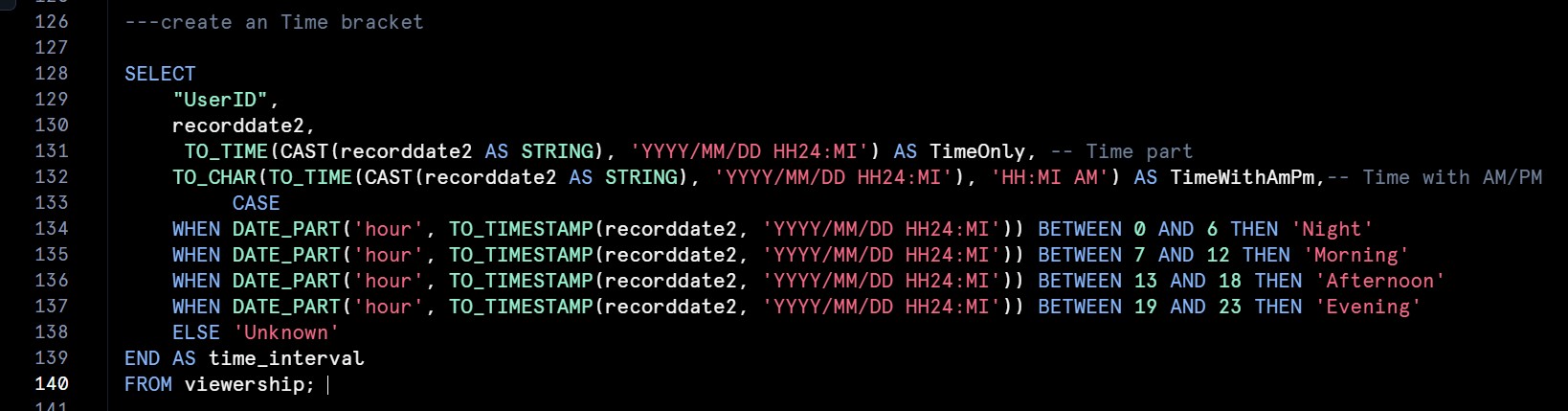
AI-generated content may be incorrect.

Created an age bracket to group individuals into meaningful, manageable categories for analysis and decision-making. Specifically, here's how I created an age bracket with differences like 12 years between categories:

|  |  |
| --- | --- |
| Age bracket | Category |
| 0-12 | Child |
| 13-17 | Teenager |
| 18-24 | Young Adult |
| 25-39 | Adult |
| 40- 50 | Middle Aged Adult |
| 50+ | Older Adult |

Grouping people this way helps capture behaviours and needs unique to each stage.

10.Age bracket



Returned output:

A screenshot of a computer

AI-generated content may be incorrect.

Created a time interval to group individuals into meaningful, manageable categories for analysis and decision-making. Specifically, here's how I created an time interval:

|  |  |
| --- | --- |
| Age bracket | Category |
| 00:00 – 06:59 | Night |
| 07:00 – 12:59 | Morning |
| 13:00 – 18:59 | Afternoon |
| 19:00 – 23:59 | Evening |
|  |  |

**Trend Analysis**

User Demographics:

1.Consumption By Age

Downloaded document created from the age bracket join on snowflake for consumption by age into MS Excel to create a graph. Sorted the graph in descending order by count of age bracket to display the age group with the highest to lowest consumption.

A screenshot of a graph

AI-generated content may be incorrect.

2.Consumption by race

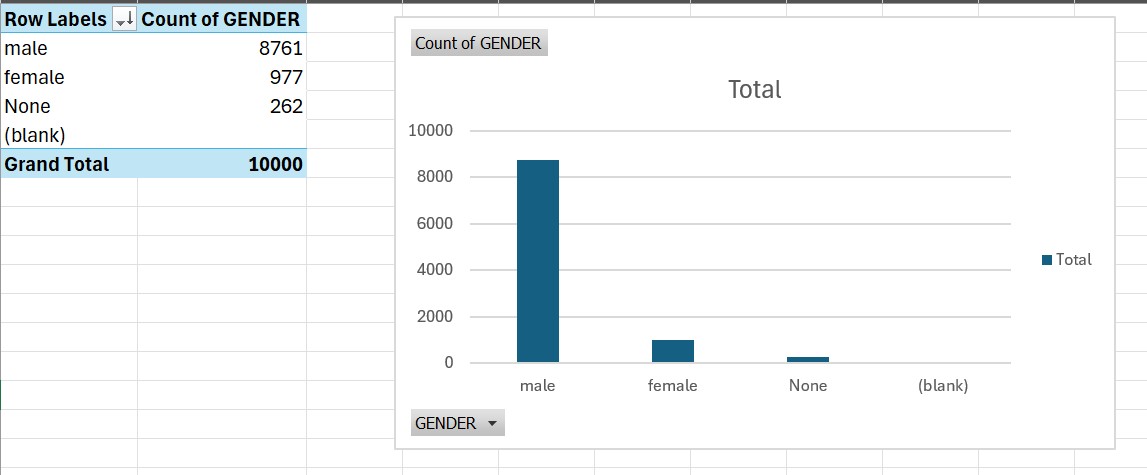
Downloaded document created from the join on snowflake for consumption by race into MS Excel to create a graph. Made use of the pie chart to effectively display all the races and percentage of consumption.

A pie chart with numbers and a number

AI-generated content may be incorrect.

3.Consumption By Gender

Downloaded document created from the join on snowflake for consumption by race into MS Excel to create a graph. Sorted the graph in descending order by count of gender to display the gender with the highest to lowest consumption.



4.Consumption By location.

Downloaded document created from the join on snowflake for consumption by race into excel to create a graph. Sorted the graph in descending order by count of location to display the location with the highest to lowest consumption.

A screenshot of a graph

AI-generated content may be incorrect.

User Trends:

1.Peak Consumption days

Downloaded document created from the inner join on snowflake for consumption by race into excel to create a graph. Sorted the graph in descending order by count of days of the week to display the days of the week with the highest to lowest consumption.

A screenshot of a computer

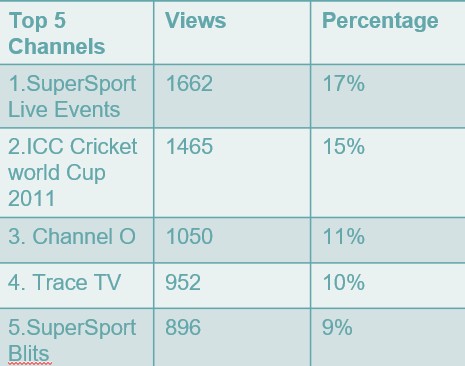
AI-generated content may be incorrect.

2.Top 5 Most watched Channels

Downloaded document created from the inner join on snowflake for consumption by race into MS Excel to create a graph. Sorted the graph in descending order by count of channel to display the channel with the highest to lowest consumption. From this bar graph I was able to create a table with the top 5 most watched channels with number of views and percentage of total viewership.

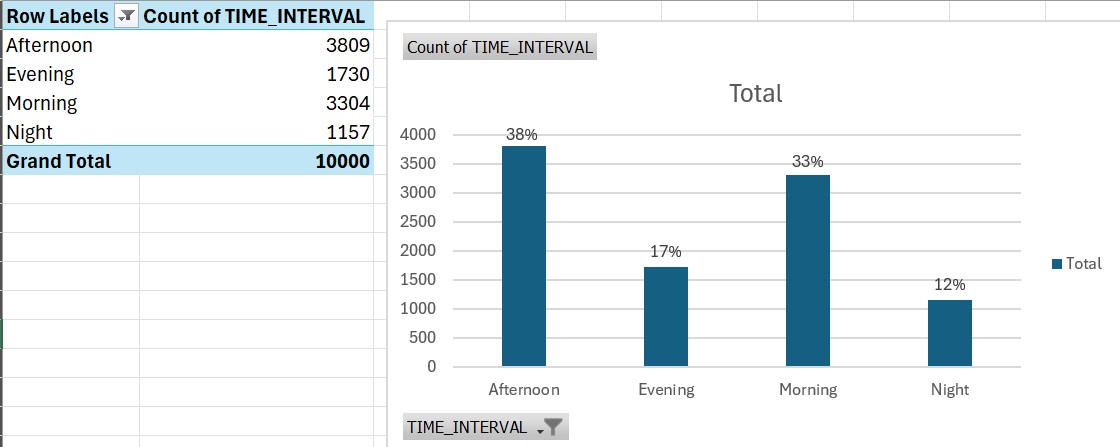
A screenshot of a graph

AI-generated content may be incorrect.



3.Peak consumption days

Downloaded document created from the time bracket on snowflake for consumption by timeonly into MS Excel to create a graph. From this bar graph I was able to create a table with the top 5 peak consumption times.



**Insights drawn from Trend Analysis**

From the users demographics I can conclude that BrightTvs main consumers are black men that range from the ages of 18-39 who reside in Gauteng.

From the users trends I can conclude that BrightTvs main consumers days are Saturday and Friday with our top 5 channels falling under music and sports channels.

**Factors Influencing Consumption**

* Sports events draw massive viewership, particularly live events.
* Music channels attract consistent engagement.
* Prime time evenings, weekends(Friday and Saturday) and public holidays are high-consumption periods.
* Group activities such as family viewings or sports gatherings likely amplify engagement.

**Recommendations to increase consumption on low consumption days**

* Run promotional campaigns targeting Gauteng residents using ads highlighting popular content like sports and music.
* Partner with local influencers, particularly those in sports and entertainment industries.
* Offer “refer-a-friend” discounts or free trial subscription(1 month free)
* Partner with sports leagues(e.g. cricket or football) to offer live-streaming of events exclusively.
* Collaborate with artists featured on music channels (e.g. Channel O, Trace TV) for exclusive content.
* Organize viewing events for big sports matches or concerts in local venues. Collaborate with TikTok influencers for visibility and engagement with their audiences.
* Expand BrightTvs reach by leveraging partnership with mobile carries or bundling services with internet providers.

**Initiative to grow BrightTvs user base**

* Introduce weekday- specific programming ( Monday Sports Recap or Midweek movie night)
* For Sunday: Cultural or inspirational programming(e.g. African Storytelling, religious content)
* Allow viewers to vote for favourite shows or curate playlist, particularly on low-consumption days.
* Replay iconic matches or moments from SuperSport Live Event to engage sports fans even during off-peak days.