Data Flow and Architecture Diagram

Start
Understand the ask and business goals

→

Identify data source:
2 x Excel file (BrightTV data)

→

Plan analysis and coding

→

Perform data conversion:
Convert Excel to CSV format for compatibility with Snowflake ingestion

→

Setup Snowflake:
- Create a Database on Snowflake

→

Setup Snowflake:
- Create a schema within the database

Setup Snowflake (data storage):
- Create a folder and worksheet

→

Complete data ingestion:
- Upload CSV to Snowflake

→

Perform data transformation (ETL):
- Use SQL scripts to transform and query

→

Perform data exploration and analysis:
- Query transformed data to understand the data
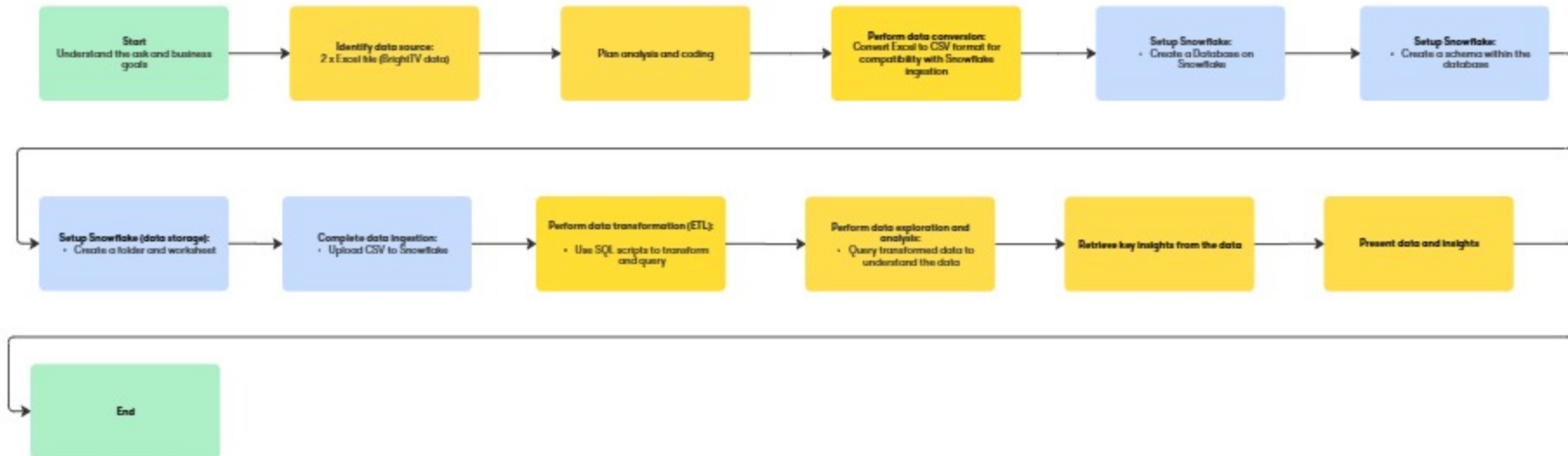
→

Retrieve key insights from the data

→

Present data and insights

End

CEO Objective:
Grow the company's subscription base for this financial year.

Project Objective:
1. Provide insights on user and usage trends of BrightTV.
2. What type of factors influence consumption?
3. What content would you recommend to increase consumption on the days with low consumption?
4. What type of initiatives would you recommend to further grow BrightTV's User base.

Key Insights to deliver:
1. Gender/ race of those with highest viewership. What time of day/day is the highest viewership?
Which channels are most watched?
In which province do they watch the most?
2. Is it time of day, gender, province?
3. See which channel has highest consumption and recommend the content for days with low consumption.
4. Based on insights and facts

Data Planning

Data Types

Date/Time:
- RecordDate2
- Duration 2

Aggregate total viewing duration per day/week/month

Create time buckets using CASE statement in SQL.

Create time buckets to identify morning, afternoon and evening

Extract the day name and day of the month

Extract the day of the month, name and month_id

Numerical:
- UserID
- Age

Create age buckets and analyze trends per age group

Unique Identifiers:
- UserID
- Email
- Social media handle

User UserID to join datasets

Check is users with social media are more active?

Categorical:
- Name
- Surname
- Gender
- Race
- Province
- Channel2

Join Viewership and User Profiles to analyse usage by age, gender, race, province.
- See which user segments are most/least active.

Aggregate viewing by Channel2.
- Identify most/least popular channels.

Which channels/content types drive longer engagement?
- Recommend similar content to boost usage.

Which demographic groups have low registration or engagement?

Quantitative data:
- Duration 2

Find days with lowest total viewing..

On high-usage days, which channels/content are most watched?

Identify peak and low usage periods

Recommend the top-performing content from high-usage days to be promoted or scheduled on low-usage days.

Analyze viewing by day of week and time of day.
- Identify when users are most/least active.