This document outlines Data Model, SQL queries necessary to load data into BigQuery tables, handle regular data updates, and create fact tables for mortality rates and net worth by age group.

**Data Model**

The data model comprises five main tables designed to store and analyse data related to causes of death, household income and wealth, and user information.

A screenshot of a computer

Description automatically generated

1. **dim\_safe\_user**
   * user\_id (PK): Unique identifier for each user.
   * user\_region: Geographic region of the user.
   * user\_dob: Date of birth of the user.
   * age\_group: Age group classification of the user.
2. **dim\_household\_income\_and\_wealth**
   * item\_type (PK): Category of the item (e.g., financial assets, Non-financial assets).
   * item (PK): Specific item within the category.
   * age\_15\_24 to age\_75\_and\_over: Values for different age groups.
3. **dim\_cause\_of\_death\_2015\_2017**
   * cause\_of\_death (PK): Specific cause of death.
   * number\_males: Number of males who died from this cause.
   * number\_females: Number of females who died from this cause.
   * number\_persons: Total number of persons who died from this cause.
   * rate\_males, rate\_females, rate\_person: Mortality rates per 100,000 for males, females, and total persons**.(assumed counts)**
   * age\_group: Age group classification for the cause of death data.
4. **fact\_age\_group\_mortality\_rates**
   * age\_group (PK): Age group classification.
   * rate\_person\_2015 to rate\_person\_2017: Mortality rates per 100,000 persons for each year.
   * rate\_males\_2015 to rate\_males\_2017: Mortality rates per 100,000 males for each year.
   * rate\_females\_2015 to rate\_females\_2017: Mortality rates per 100,000 females for each year.
   * cause\_of\_death (FK): Cause of death reference.
5. **fact\_age\_group\_across\_net\_worth**
   * age\_group\_across\_net\_worth\_id (PK): Unique identifier for each record.
   * user\_id (FK): Reference to the user ID.
   * age\_group (FK): Age group classification.
   * net\_worth: Combined net worth of the age group.

Relationships

* dim\_safe\_user links to fact\_age\_group\_across\_net\_worth through the user\_id and age\_group fields.
* dim\_household\_income\_and\_wealth contributes data to fact\_age\_group\_across\_net\_worth for calculating net worth by age group.
* dim\_cause\_of\_death\_2015\_2017 data is summarized in fact\_age\_group\_mortality\_rates which provides mortality rates by age group.
* fact\_age\_group\_mortality\_rates links to fact\_age\_group\_across\_net\_worth through the age\_group field.

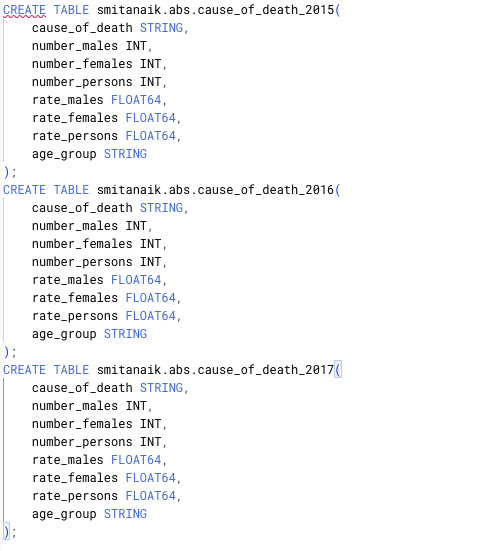
Top of Form

Bottom of Form

**SQL**

The following tables are created to store the data:

1. Cause of Death Tables

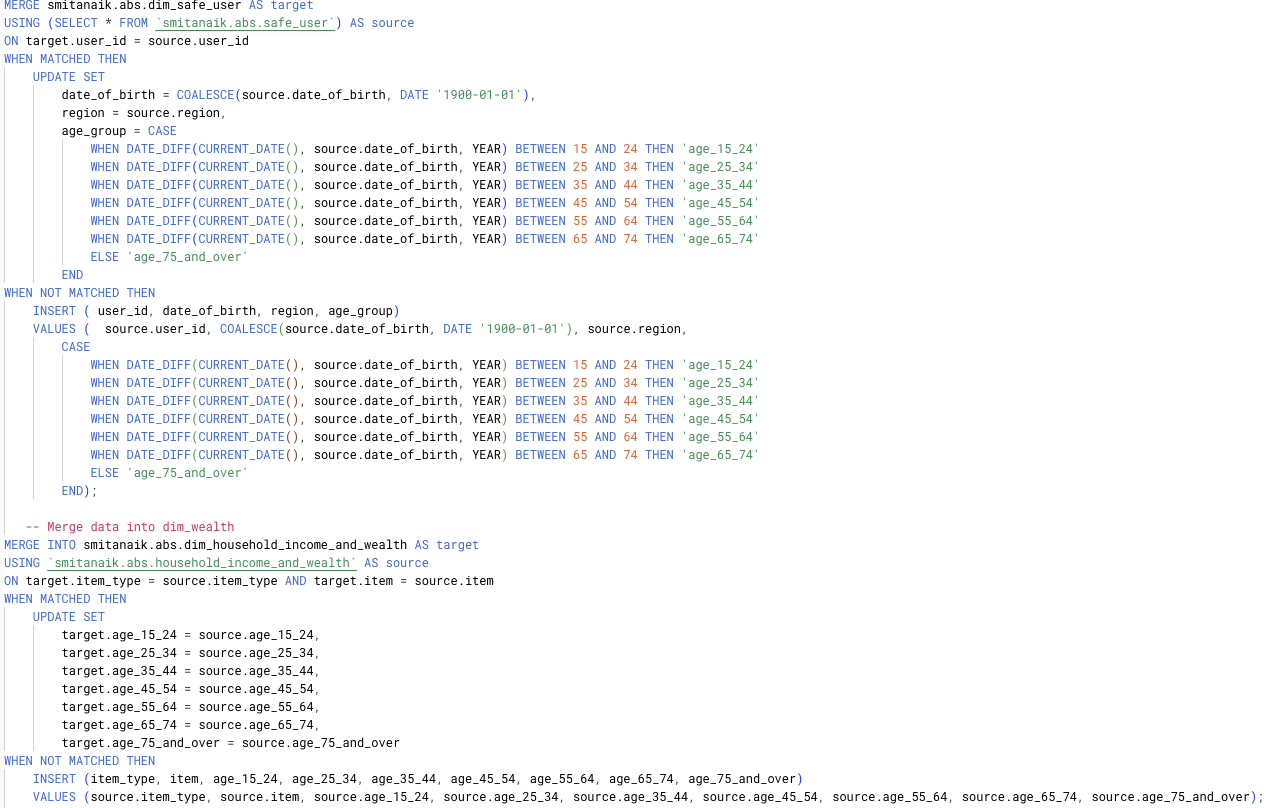


1. Houseold\_income\_and\_wealth and safe\_userTable

A screenshot of a computer

Description automatically generated

The MERGE statement ensures the tables are updated with new data and existing records are modified if necessary.



Fact Table for Mortality Rates by Age Group

This query aggregates the mortality rates by age group across the years 2015, 2016, and 2017.

A screenshot of a computer program

Description automatically generated

A screenshot of a computer program

Description automatically generated

Fact Table for net worth by Age Group

This query aggregates the net worth = (all assets – liabilities) for the age group for the users.  
  
