



ITC 6000

Database Management Systems

Final Project Presentation

"DRUG MISUSE CASE STUDY"

Group Name – HUSKY 3

Team Members: Iti Rohilla (Leader)

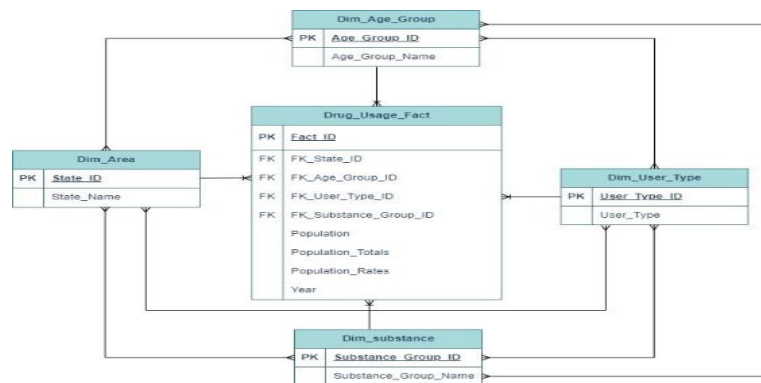
Kareena Ramrakhya

Monika Gundecha

Shachi Dwivedi

Ratnesh Mishra

Entity Relationship Diagram for the Project:



SQL COMMANDS FOR DATA IMPORT:

The data used for the project was imported from csv files obtained from the data source onto the DBBrowser using following commands

Create Table:

<pre> CREATE TABLE "Dim_Age_Group" ("Age_Group_ID" INTEGER, "Age_Group_Name" TEXT, PRIMARY KEY("Age_Group_ID")) INSERT INTO Dim_Age_Group(Age_Group_ID, Age_Group_Name) VALUES (1001, '12_17'); </pre>	<pre> CREATE TABLE "Dim_User_Type" ("User_Type_ID" INTEGER, "User_Type" TEXT, PRIMARY KEY("User_Type_ID")) INSERT INTO Dim_User_Type(User_Type_ID, User_Type) VALUES (3001, 'Used Past Year'); </pre>
<pre> CREATE TABLE "Dim_Substance" ("Substance_Group_ID" INTEGER, "Substance_Group_Name" TEXT, PRIMARY KEY("Substance_Group_ID")) INSERT INTO Dim_Substance(Substance_Group_ID, Substance_Group_Name) VALUES (2501, 'Alcohol'); </pre>	<pre> CREATE TABLE "Dim_Area" ("State_ID" INTEGER, "State_Name" TEXT, PRIMARY KEY("State_ID")) INSERT INTO Dim_Area(State_ID, State_Name) VALUES (2100, 'Alabama'); </pre>
<pre> CREATE TABLE Drug_Usage_Fact(Fact_ID INTEGER NOT NULL PRIMARY KEY ,State_ID INTEGER NOT NULL ,Year INTEGER NOT NULL ,Population INTEGER NOT NULL ,FK_Substance_Group_ID INTEGER NOT NULL ,FK_User_Type_ID INTEGER NOT NULL ,Population_Totals INTEGER NOT NULL ,Population_Rates NUMERIC(7,3) NOT NULL ,FK_Age_Group_ID INTEGER NOT NULL </pre>	<pre> INSERT INTO Drug_Usage_Fact(Fact_ID, State_ID, Year, Population, FK_Substance_Group_ID, FK_User_Type_ID, Population_Totals, Population_Rates, FK_Age_Group_ID) VALUES (40001, 2100, 2002, 380805, 2504, 3003, 20, 59.732, 1001); </pre>

SQL STATEMENTS –

We ran following SQL queries to answer few interesting questions about data:

<p>1. Among different age groups, which state has the highest number of new users for various substance usage over the years?</p> <p>Query: SELECT dag.Age_Group_Name, da.State_Name, SUM(Population_Totals) AS Total_New_Users FROM Drug_Usage_Fact f JOIN Dim_Age_Group dag ON f.FK_Age_Group_ID = dag.Age_Group_ID JOIN Dim_Area da ON f.FK_State_ID = da.State_ID JOIN Dim_Substance ds ON f.FK_Substance_Group_ID = ds.Substance_Group_ID WHERE f.FK_User_Type_ID = 3003 GROUP BY dag.Age_Group_Name, da.State_Name ORDER BY dag.Age_Group_Name, Total_New_Users DESC;</p>	<p>2. Breakdown of the number of users for each substance group.</p> <p>Query: SELECT ds.Substance_Group_Name, sum(Population_Totals) AS Total_Users FROM Drug_Usage_Fact f JOIN Dim_Substance ds ON f.FK_Substance_Group_ID = ds.Substance_Group_ID GROUP BY ds.Substance_Group_Name;</p>
<p>3. Determining the year with the highest drug usage rate.</p> <p>Query: SELECT Year, SUM(Population_Totals) AS Total_Usage FROM Drug_Usage_Fact GROUP BY Year ORDER BY Total_Usage DESC LIMIT 1;</p>	<p>4. Identifying the substance group with the highest usage.</p> <p>Query: Identifying the substance group with the highest usage: SELECT ds.Substance_Group_Name, SUM(Population_Totals) AS Total_Usage FROM Drug_Usage_Fact f JOIN Dim_Substance ds ON f.FK_Substance_Group_ID = ds.Substance_Group_ID GROUP BY ds.Substance_Group_Name ORDER BY Total_Usage DESC LIMIT 1;</p>
<p>5. Finding the age group with the highest drug usage.</p> <p>Query: - SELECT da.Age_Group_Name, SUM(Population_Totals) AS Total_Usage FROM Drug_Usage_Fact f JOIN Dim_Age_Group da ON f.FK_Age_Group_ID = da.Age_Group_ID GROUP BY da.Age_Group_Name ORDER BY Total_Usage DESC LIMIT 1;</p>	

Data Architecture and Storage Requirements:

1) Architecture: Azure Data Studio is the client-side, while Azure SQL Database is the server-side of your solution, following the common client-server model.

2) Cloud Hosting: Both Azure Data Studio and Azure SQL Database are hosted in the cloud, providing benefits such as scalability, high availability, and easy management without the need for physical hardware or on-premises infrastructure.

3) Storage Requirements: The well-structured and optimized database design efficiently manages the fact table with 20,808 rows and four dimension tables with a total of 62 rows, resulting in a total database size of 3.4 MB. Azure automatically manages and scales the storage as needed, making it a cost-effective solution for your storage needs.