**DRUG**

**MISUSE**

**CASE STUDY**



**ITC 6000 - DBMS**

**MODULE 6**

**FINAL PROJECT REPORT**

**HUSKY GROUP 3**

**Iti Rohilla**

**Kareena Ramrakhya**

**Monika Gundecha**

**Shachi Dwivedi**

**Ratnesh Mishra**



**INTRODUCTION**

**Project Idea and Interest**: For our project, We are interested in developing a thorough database system that analyzes and maintains data related to substance addiction, focusing in particular on tobacco , alcohol and drug use trends and related public health issues of a particular nation , the data which we chose to process is about substance abuses among different age groups and states , data ranges from 2002 to 2018 both in totals(in thousands of people) and rates(as a percentage of the population are given). We are interested in this issue since it provides an opportunity to examine facts in order to make well-informed decisions and establish policies.

The topic which we have chosen is very crucial and socially relevant , our study aims to serve as an informative and data-driven platform to address the alarming issue of substance abuse, specifically focusing on tobacco alcohol and drug usage .

**Our personal connection** comes from our shared commitment to address the serious issue of substance abuse at such tender age, as well as our aim to help develop data-driven solutions for problems facing public health. We choose this subject because it fits with our shared passion in applying technology to solve problems in the real world and improve society. We want to equip people and organizations with the resources they need to effectively battle drug abuse through our database.

With the help of our database , queries and study organisations can make educated decisions about intervention tactics and policy creation by analyzing age-related data and gaining insights into substance misuse tendencies. Our study will help advance public health and safety.

**BUSINESS ANALYSIS**

**Governmental Institutions**

The data can be used by public health departments to design specialized awareness programs, give funds to facilities that treat addiction, and formulate evidence-based drug policy.

Data insights can help law enforcement authorities pinpoint regions with greater drug misuse rates so they can concentrate their efforts on prevention and control.

NGOs: Non-Governmental Organizations

**Rehabilitation Facilities**:

NGOs that offer addiction treatment might utilize the information to customize their programs to meet the particular requirements of various age groups and substances.

Community Outreach Programs: NGOs that focus on raising public awareness and educating the public can create campaigns that appeal to various populations.

health industry personnel

**Counselors and doctors:**

Healthcare professionals can better adapt their treatment strategies for patients by understanding the demographics of those who seek assistance for substance misuse.

Researchers: Addiction and public health researchers might use the data for research studies and suggestions on public policy.

**Academia:**

Students and Researchers: Academic institutions can be a great resource for anyone researching public health, sociology, psychology, and addiction.

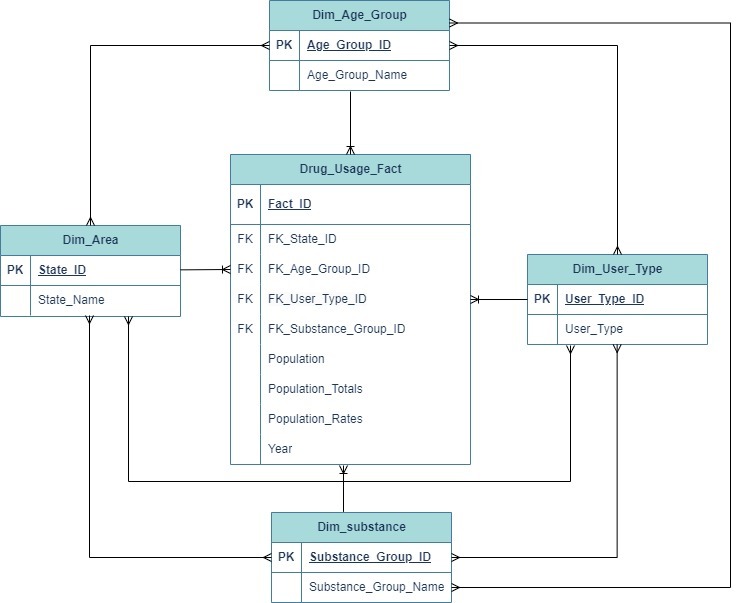
**Community Members:**

By having better access to and knowledge of local drug addiction trends, the general people will be better able to make decisions that will benefit their own and their families' health.

**Journalism and the media:**

Journalists: Media experts can use the information to produce enlightening stories, articles, and reports that will promote awareness and change.

**ER DIAGRAM**



**TABLE RELATIONSHIPS**

1. **Dim\_Fact has foreign keys to Dim\_Area, Dim\_Age\_Group, Dim\_User\_Type, Dim\_Substance.**

This indicates:

**Many to many relationship exists among all the Dimension tables like:**

Dim\_Area & Dim\_Age\_Group

Dim\_User\_Type & Dim\_Substance

Dim\_Area & Dim\_Substance

Dim\_User\_Type & Dim\_Area

Dim\_Age\_Group & Dim\_User\_Type

(b) One to many relationship between:

Drug\_Usage\_Fact & Dim\_Area

Drug\_Usage\_Fact & Dim\_Age\_Group

Drug\_Usage\_Fact & Dim\_User\_Type

Drug\_Usage\_Fact & Dim\_Substance

**DATABASE IMPLEMENTATION**

**Inquiry/Overview/Basic Insight of our dataset:**

**1. Finding the total number of drug usage records for each year.**

--Finding the total number of drug usage for each year:

**SELECT Year, sum(Population\_Totals) AS Total\_Usage FROM Drug\_Usage\_Fact GROUP BY Year;**

A screenshot of a table

Description automatically generated

**2. Breakdown of the number of users for each substance group.**

-- Breakdown of the number of users for each substance group:

**SELECT ds.Substance\_Group\_Name, sum(Population\_Totals) AS Total\_UsersFROM Drug\_Usage\_Fact fJOIN Dim\_Substance ds ON f.FK\_Substance\_Group\_ID = ds.Substance\_Group\_IDGROUP BY ds.Substance\_Group\_Name;**

**A screenshot of a computer

Description automatically generated**

**3. Determining the year with the highest drug usage rate.**

-- Determining the year with the highest drug usage rate:

**SELECT Year, SUM(Population\_Totals) AS Total\_UsageFROM Drug\_Usage\_FactGROUP BY YearORDER BY Total\_Usage DESCLIMIT 1;**

****

**4. Identifying the substance group with the highest usage.**

-- Identifing the substance group with the highest usage:

**SELECT ds.Substance\_Group\_Name, SUM(Population\_Totals) AS Total\_UsageFROM Drug\_Usage\_Fact fJOIN Dim\_Substance ds ON f.FK\_Substance\_Group\_ID = ds.Substance\_Group\_IDGROUP BY ds.Substance\_Group\_NameORDER BY Total\_Usage DESCLIMIT 1;**

****

**5. Finding the age group with the highest drug usage.**

-- Finding the age group with the highest drug usage:

**SELECT da.Age\_Group\_Name, SUM(Population\_Totals) AS Total\_UsageFROM Drug\_Usage\_Fact fJOIN Dim\_Age\_Group da ON f.FK\_Age\_Group\_ID = da.Age\_Group\_IDGROUP BY da.Age\_Group\_NameORDER BY Total\_Usage DESCLIMIT 1;**

**SELECT ds.Substance\_Group\_Name, du.User\_Type, Year, SUM(Population\_Totals) AS Total\_Users**

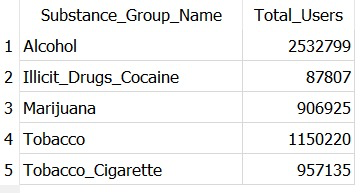
**FROM Drug\_Usage\_Fact f**

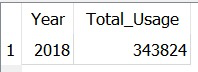
**JOIN Dim\_Substance ds ON f.FK\_Substance\_Group\_ID = ds.Substance\_Group\_ID**

**JOIN Dim\_User\_Type du ON f.FK\_User\_Type\_ID = du.User\_Type\_ID**

**GROUP BY ds.Substance\_Group\_Name, du.User\_Type, Year**

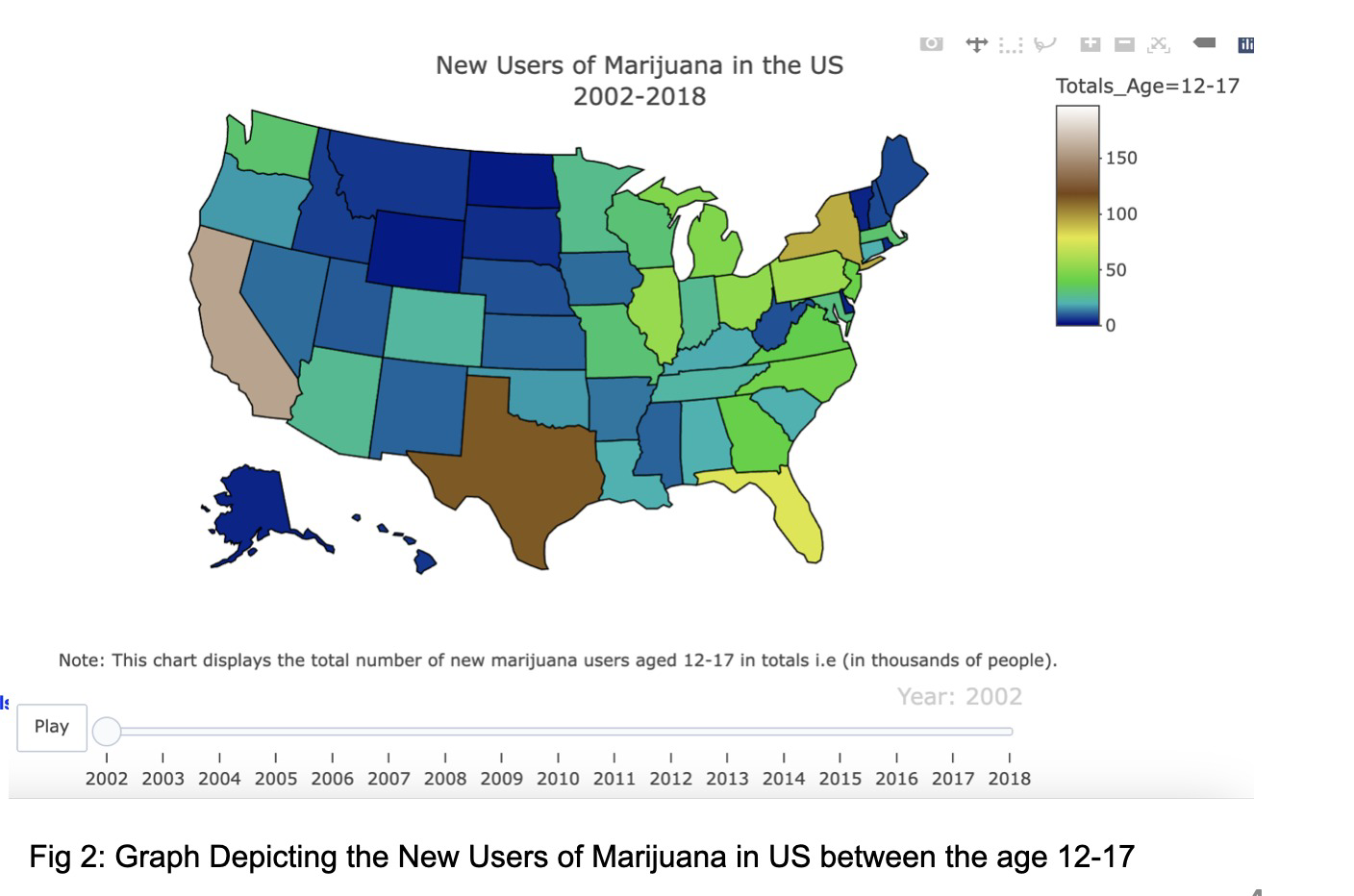
**ORDER BY ds.Substance\_Group\_Name, du.User\_Type, Year;**

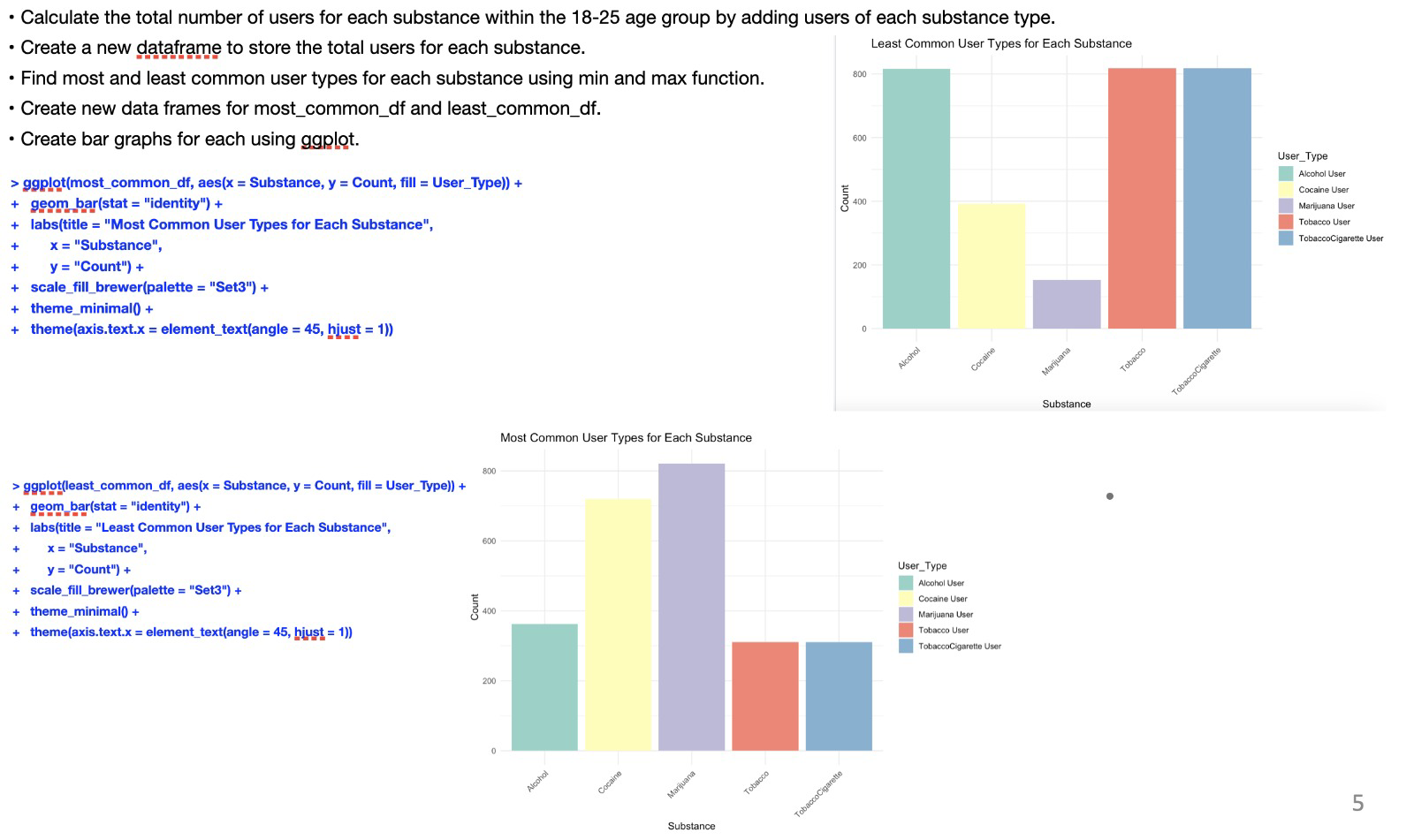
****

****

**A close up of a computer screen

Description automatically generated**

****

****

**Create/Insert Queries :**

1. CREATE TABLE "Dim\_Age\_Group" (

"Age\_Group\_ID" INTEGER,

"Age\_Group\_Name" TEXT,

PRIMARY KEY("Age\_Group\_ID")

)

1. INSERT INTO Dim\_Age\_Group(Age\_Group\_ID,Age\_Group\_Name) VALUES (1001,'12\_17');
2. INSERT INTO Dim\_Age\_Group(Age\_Group\_ID,Age\_Group\_Name) VALUES (1002,'18\_25');
3. INSERT INTO Dim\_Age\_Group(Age\_Group\_ID,Age\_Group\_Name) VALUES (1003,'26+');
4. CREATE TABLE "Dim\_Substance" (
5. "Substance\_Group\_ID" INTEGER,
6. "Substance\_Group\_Name" TEXT,
7. PRIMARY KEY("Substance\_Group\_ID")
8. )
9. INSERT INTO Dim\_Substance(Substance\_Group\_ID,Substance\_Group\_Name) VALUES (2501,'Alcohol');
10. INSERT INTO Dim\_Substance(Substance\_Group\_ID,Substance\_Group\_Name) VALUES (2502,'Tobacco\_Cigarette');
11. INSERT INTO Dim\_Substance(Substance\_Group\_ID,Substance\_Group\_Name) VALUES (2503,'Illicit\_Drugs\_Cocaine');
12. INSERT INTO Dim\_Substance(Substance\_Group\_ID,Substance\_Group\_Name) VALUES (2504,'Marijuana');
13. INSERT INTO Dim\_Substance(Substance\_Group\_ID,Substance\_Group\_Name) VALUES (2505,'Tobacco');
14. CREATE TABLE "Dim\_User\_Type" (
15. "User\_Type\_ID" INTEGER,
16. "User\_Type" TEXT,
17. PRIMARY KEY("User\_Type\_ID")
18. )
19. INSERT INTO Dim\_User\_Type(User\_Type\_ID,User\_Type) VALUES (3001,'Used Past Year');
20. INSERT INTO Dim\_User\_Type(User\_Type\_ID,User\_Type) VALUES (3002,'Used Past Month');
21. INSERT INTO Dim\_User\_Type(User\_Type\_ID,User\_Type) VALUES (3003,'New Users');
22. ﻿CREATE TABLE "Dim\_Area" (
23. "State\_ID" INTEGER,
24. "State\_Name" TEXT,
25. PRIMARY KEY("State\_ID")
26. )
27. INSERT INTO Dim\_Area(State\_ID,State\_Name) VALUES (2100,'Alabama');
28. INSERT INTO Dim\_Area(State\_ID,State\_Name) VALUES (2101,'Alaska');
29. INSERT INTO Dim\_Area(State\_ID,State\_Name) VALUES (2102,'Arizona');
30. INSERT INTO Dim\_Area(State\_ID,State\_Name) VALUES (2103,'Arkansas');
31. INSERT INTO Dim\_Area(State\_ID,State\_Name) VALUES (2104,'California');
32. INSERT INTO Dim\_Area(State\_ID,State\_Name) VALUES (2105,'Colorado');
33. INSERT INTO Dim\_Area(State\_ID,State\_Name) VALUES (2106,'Connecticut');
34. INSERT INTO Dim\_Area(State\_ID,State\_Name) VALUES (2107,'Delaware');
35. INSERT INTO Dim\_Area(State\_ID,State\_Name) VALUES (2108,'District of Columbia');
36. CREATE TABLE Drug\_Usage\_Fact(
37. Fact\_ID INTEGER NOT NULL PRIMARY KEY
38. ,State\_ID INTEGER NOT NULL
39. ,Year INTEGER NOT NULL
40. ,Population INTEGER NOT NULL
41. ,FK\_Substance\_Group\_ID INTEGER NOT NULL
42. ,FK\_User\_Type\_ID INTEGER NOT NULL
43. ,Population\_Totals INTEGER NOT NULL
44. ,Population\_Rates NUMERIC(7,3) NOT NULL
45. ,FK\_Age\_Group\_ID INTEGER NOT NULL
46. );
47. 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);
48. 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 (40002,2101,2002,69400,2504,3003,4,77.736,1001);
49. 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 (40003,2102,2002,485521,2504,3003,25,64.179,1001);
50. **USING CONCATENATE FUNCTION IN ADVANCED EXCEL WE CAN CREATE REST OF THE INSERT STATEMENTS QUICKLY BY STRETCHING THE FUNCTION FOR ALL THE ROWS**

**SECURITY & PRIVACY CONCERNS**

Given that our dataset is not user-specific, we can confidently state that there are no specific security and privacy concerns, including user privacy, authentication, data encryption, and access control.

**PROJECT ARCHITECTURE REQUIREMENTS**

**Sensitive Information:** The dataset contains information related to substance use, including alcohol, tobacco, cocaine, and marijuana. This data can be considered sensitive, especially when broken down by age groups and states. It may reveal patterns of substance use that can be potentially sensitive or stigmatized.

**Personally Identifiable Information (PII):** While the dataset does not appear to contain direct PII like names or addresses, it does include information on populations and rates related to specific states. In some cases, it might be possible to indirectly identify individuals or regions based on the data.

**Privacy Risks:** The detailed breakdown of substance use by age group and state can pose privacy risks. The identification of trends or patterns within specific age groups or regions could potentially lead to the identification of individuals or communities, raising privacy concerns.

**Compliance:** Data related to substance use and disorders might be subject to legal and ethical compliance requirements, such as health data protection regulations. Ensuring that the data is handled in compliance with these regulations is essential.

In summary, the dataset contains potentially sensitive information related to substance use and disorders, with possible privacy implications. It's important to handle and store this data with appropriate security measures, ensuring compliance with relevant regulations to protect the privacy of individuals and communities. Additionally, access to this data should be restricted to authorized personnel to prevent any unauthorized use or disclosure.

**WHAT WE’VE LEARNED**

With a focus on **tobacco, alcohol, and drug use trends**, this project has provided us with invaluable insights into the pressing problem of substance addiction. With an emphasis on a particular country, our database system was created to maintain and analyze data about substance abuse, looking at information from 2002 to 2018. This is what we now know:

**Recognizing the Importance of Data on Substance Abuse**

We now recognize how crucial it is to gather and examine data regarding drug abuse. Addiction to substances is a serious public health concern that affects people individually, in families, and in communities. We have identified this data's potential to inform evidence-based policies and interventions by looking through and preserving it.

**The Potency of Solutions Driven by Data**

Our project demonstrated the effectiveness of data-driven approaches in resolving challenging societal issues. We have developed a platform with our database that enables different stakeholders, including NGOs, government agencies, medical professionals, and researchers, to make well-informed decisions in the fight against substance abuse. This confirms our conviction that technology has the power to bring about constructive change.

**Considerations for Security and Privacy**

We now understand how crucial it is to manage sensitive data appropriately. Even in the absence of direct PII, the substance abuse dataset can provide insightful information with potential privacy implications. Therefore, it is crucial to put strong security measures in place and make sure that all applicable laws are followed in order to safeguard people's privacy as well as that of their communities.

**The Linkages Among Stakeholders**

We have seen from our project how access to well-structured data can benefit a wide range of stakeholders, including government agencies, non-governmental organizations, medical professionals, academic institutions, and the general public. It has brought attention to the necessity of cooperation and information exchange in order to successfully handle complicated public health issues.

**PROJECT WRAP UP AND FUTURE CONSIDERATIONS**

It's critical to evaluate this project's impact and possible future directions as we draw to a close:

**Project Effects**

The battle against substance abuse could be greatly impacted by our database system. It offers a useful resource for public awareness campaigns, policy development, and data analysis. Evidence-based policies and interventions may result from its adoption by NGOs, academics, medical professionals, and governmental institutions.

**Constant Updates of Data**

Constant data updates are necessary to guarantee the database’s accuracy and relevance. Substance abuse-related public health trends are always changing, so it’s important to keep data up to date so that stakeholders can modify their plans in response to new information.

**User Assistance and Education**

Future project considerations should include user support and training. Offering training materials and assistance can improve the database's usability and efficacy as more businesses and individuals use it.

**Integration of Data**

Integrating information from various sources, including treatment results, financial effects, and law enforcement data, may offer a more complete picture of the drug abuse issue. Future iterations of the project might include this expansion.

To sum up, our project has the potential to make a substantial impact on the fight against substance abuse. We now understand the worth of data, the significance of security and privacy, and the necessity of stakeholder collaboration. There are prospects for growth, ongoing development, and a more extensive influence on public health in the future.

**REFERENCES**

*1.Creating a SQL INSERT statement*. (n.d.). <https://docs.devart.com/querybuilder-for-sql-server/building-queries/how-to-create-sql-insert-statement.html>

2. Tahaghoghi, S. M. (n.d.). *Learning MySQL*. O’Reilly Online Learning. <https://www.oreilly.com/library/view/learning-mysql/0596008643/ch04s03.html>

3. *Articles | DrugBank Help Center*. (n.d.). DrugBank. https://dev.drugbank.com/guides/sql