A

Project Report on

**IBM Cloud using Watson Studio**

**Upload Road Accident data to Jupyter Notebook in and create a dashboard.**

Submitted in partial fulfillment of completion of the course

Advanced

Diploma in IT, Networking and Cloud Computing

Submitted by:

**Archana Kumari**

Under Guidance of:

**Arpita Roy (Edunet)**

|  |  |  |
| --- | --- | --- |
| C:\Users\ibmuser\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\2C3D7B61.tmp | C:\Users\ibmuser\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\13EB517.tmp | C:\Users\ibmuser\AppData\Local\Microsoft\Windows\INetCache\Content.MSO\99B363D.tmp |

Year 2023

**PROJECT OVERVIEW**

Abstract

Acknowledgement

Table of Contents

1. Introduction to Problem
2. Proposed Solution
3. Requirements

4.1 Technology Stack

4.2 Hardware

4.3 Software

4.4 Deployment Environment

5. User Requirements

6. Design Documentation

7. Implementation Details

8. Testing

9. Deployment

10. Future Scope

11. Conclusion

Appendix A Screenshot of Project

References

**ABSTRACT**

This project focuses on utilizing IBM Cloud with Watson Studio to upload road accident data into a Jupyter Notebook, subsequently creating a comprehensive dashboard for analysis. Leveraging the powerful tools provided by Watson Studio, the aim is to facilitate efficient data processing, visualization, and interpretation. The resulting dashboard will serve as a valuable resource for gaining insights into road safety and accident trends, enabling informed decision-making and proactive measures for enhancing public safety.

Machine learning models are employed to predict accident severity based on variables such as weather conditions, time of day, and road type. These predictions aid in deploying emergency services effectively and improving response times.

**ACKNOWLEDGEMENT**

We would like to express our sincere gratitude to all those who have supported and contributed to the successful completion of this project. Your assistance and encouragement have been invaluable throughout this journey.

First and foremost, we want to thank my project supervisors, Arpita Roy(Edunet) for their guidance, expertise, and unwavering support. Their insights and feedback have been instrumental in shaping the direction of this project.

We would also like to extend my appreciation to my fellow classmates who provided valuable input, shared resources, and engaged in stimulating discussions that enriched the project. Your collaborative spirit was a driving force behind our achievements.

Furthermore, we want to acknowledge our friends and family for their patience, understanding, and encouragement throughout this endeavour. Your support provided the motivation we needed to see this project through to its completion.

Last but not least, we are grateful to the entire faculty and staff of NSTI Kolkata for providing a conducive learning environment and the necessary resources to undertake this project.

This project has been a rewarding learning experience, and we are thankful for the collective efforts of everyone involved. Your support has been instrumental in making this project a reality.

Archana Kumari,

NSTI KOLKATA

03/10/2023

**INTRODUCTION TO PROBLEM**

The escalating rate of road accidents demands a data-driven approach for effective analysis and mitigation. Leveraging IBM Cloud and Watson Studio, this project addresses the challenge by uploading road accident data into a Jupyter Notebook. The objective is to create a streamlined dashboard, enabling comprehensive analysis and informed decision-making to enhance road safety measures.

1. What is the highest and lowest Road accident during 2010 to 2023

2. What change in accidents per lakh population from 2010 to 2023

3. The correlation between diffrent variables like accident persons killed and road accident per 10k vechicles

4. Trend in the Number of Persons Killed in Road Accidents (2010-2023)

5. The changes in the share of deaths from 2010 to 2023 and states with significant changes

6. states with consistently high or low rates

7. Relationship between the number of persons killed and the population, vehicles, or road length in each state?

8. The rate of persons killed per lakh population for each year

9. Trend in the Number of Persons Killed in Road Accidents (2010-2023) states with consistently high fatality rates

10. The change in injuries per 10,000 km of roads

11. What is the changes in the share Injury\_per\_10,000\_Km\_of\_Roads from 2016 to 2018

**PROPOSED SOLUTION**

In response to the pressing challenges posed by road accidents, our proposed solution involves a comprehensive approach to analyze and interpret road accident data. By harnessing the power of data analytics, this solution aims to uncover patterns, identify key risk factors, and inform evidence-based strategies for enhancing road safety.

Utilizing the robust capabilities of Watson Studio on IBM Cloud, we seamlessly upload road accident data into a Jupyter Notebook. This integrated platform empowers us to leverage advanced analytics, enabling the creation of a dynamic dashboard for comprehensive analysis. The solution harnesses the power of data visualization and interpretation, providing actionable insights to inform proactive measures and improve road safety.

**REQUIREMENTS**

**TECHNOLOGY STACK**

* : Data Analysis and Machine Learning:

Building a robust technology stack for road accident data analysis involves selecting tools and technologies that can handle data collection, preprocessing, analysis, and visualization efficiently. Here's a suggested technology stack for road accident data analysis:

Python is a versatile language with rich libraries for data analysis and machine learning. Pandas and NumPy help with data manipulation, and Scikit-Learn provides tools for predictive modeling.

* Programming Languages: Python with Pandas, NumPy, and Scikit-Learn

**HARDWARE**

* Server: A dedicated server or cloud-based virtual machine (VM) to host jupyter notebook
* CPU and RAM
* Storage: Adequate storage space for the csv file, application code, and uploaded files.
* Network Infrastructure: Reliable internet connectivity, firewalls, and security measures to protect the system.

**SOFTWARE**

* Operating System
* Python with Pandas, NumPy, and Scikit-Learn, along with any necessary packages and dependencies. Use virtual environments to manage Python dependencies.
* Development Tools: jupyter notebook
* Version Control Tools: GitHub or GitLab
* Web Server
* IBM Cloud

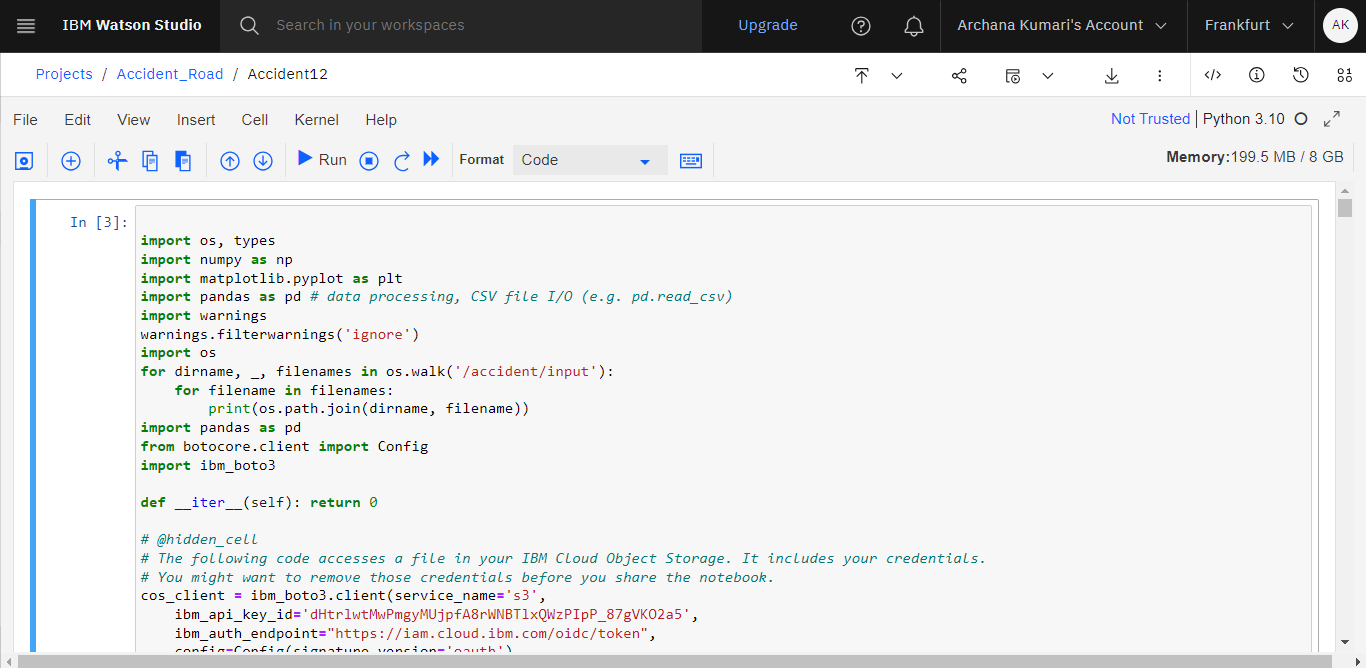
**DEPLOYMENT ENVIRONMENT**

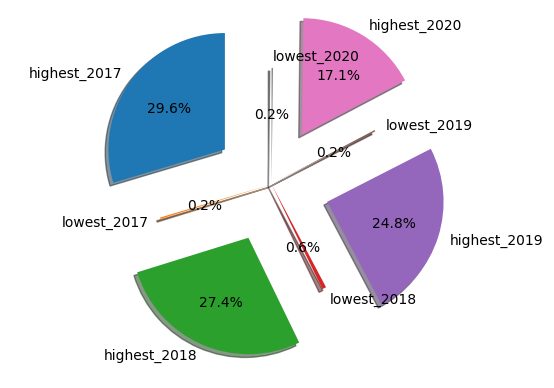
* The size and complexity of the project
* The security requirements of the project
* The budget for the project
* The technical expertise of the team

**USER REQUIREMENTS**

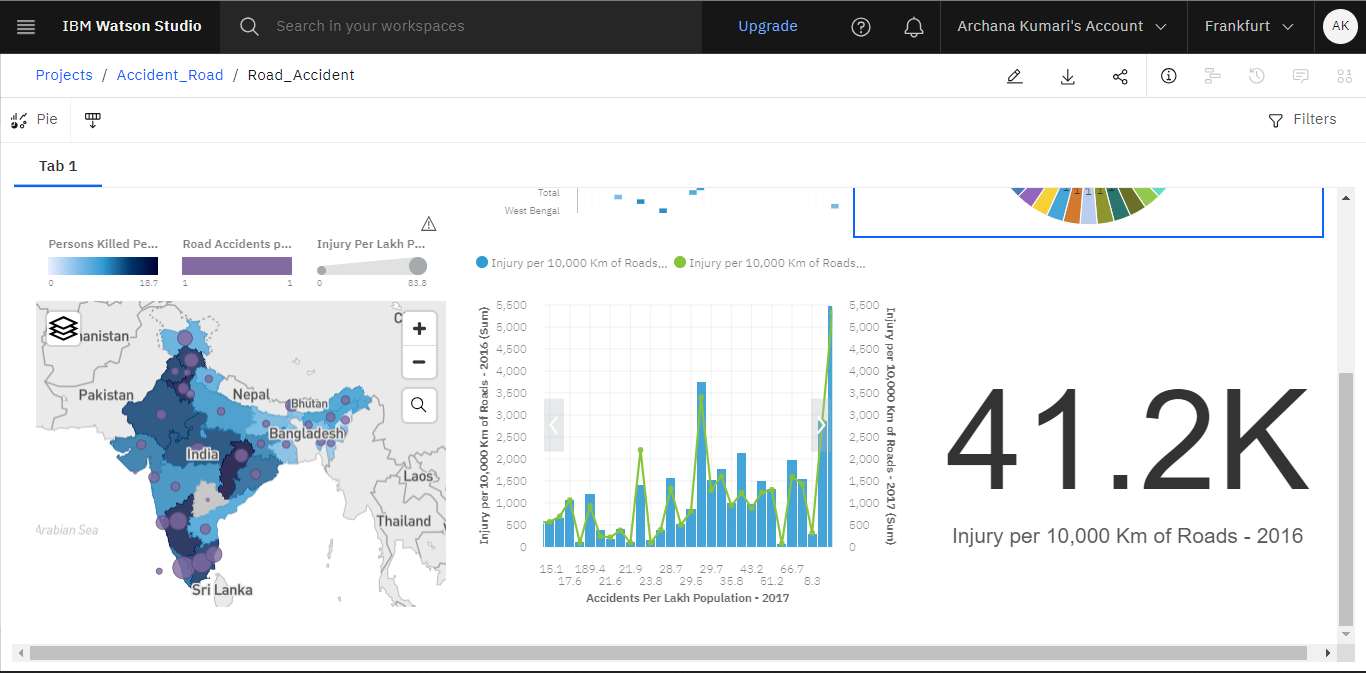
* Mobile Accessibility or PCs
* Internet Access
* Supported Browser
* Agreement to Terms

**DESIGN DOCUMENTATION**









Creating a comprehensive road accident data analysis in a Jupyter Notebook involves several steps, from data loading and preprocessing to visualization and insights generation. In this example, I'll provide a simplified demonstration using Python and popular data science libraries. For a real-world scenario, you would need to adapt and expand this code based on the specifics of your dataset and analysis goals.

**Step 1: Setup and Dependencies**

# Install required packages

pip install pandas matplotlib seaborn scikit-learn geopandas

# Import necessary libraries

import numpy as np

import pandas as pd # data processing, CSV file I/O (e.g. pd.read\_csv)

import warnings

warnings.filterwarnings('ignore')

import os

for dirname, \_, filenames in os.walk('/accident/input'):

for filename in filenames:

print(os.path.join(dirname, filename))

**Step 2: Data Loading**

Assuming you have a dataset in CSV format, load it into a Pandas DataFrame.

rac = pd.read\_csv("Road\_Accident.csv")

rac

**Step 3: Data Preprocessing**

is a command used to display the first 5 rows of a DataFrame named.

rac.head(5)

**step 4: how many rows and columns of the DataFrame**

rac.shape

**step 5: used to understand the total number of elements in the DataFrame**

rac.size

**step 6: show all Csv file datas**

rac.describe().round(2)

**step 7:** **used to identify and count the missing values (NaN or null values) in each column**

rac.isna().sum()

**step 8 : #statement is used to remove rows containing missing values**

* rac.dropna(inplace=True
* rac.isnull().any()
* rac.info()
* rac.columns

rac.columns=[i.replace(' ','\_')for i in rac.columns]

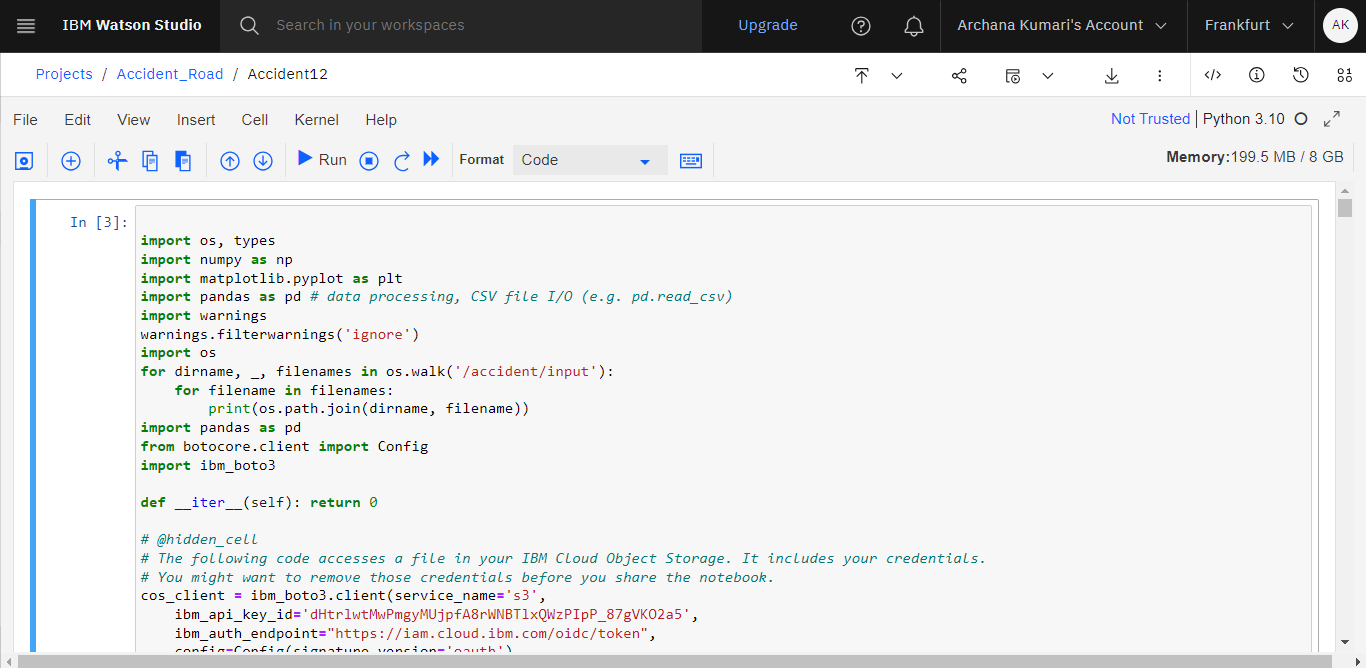
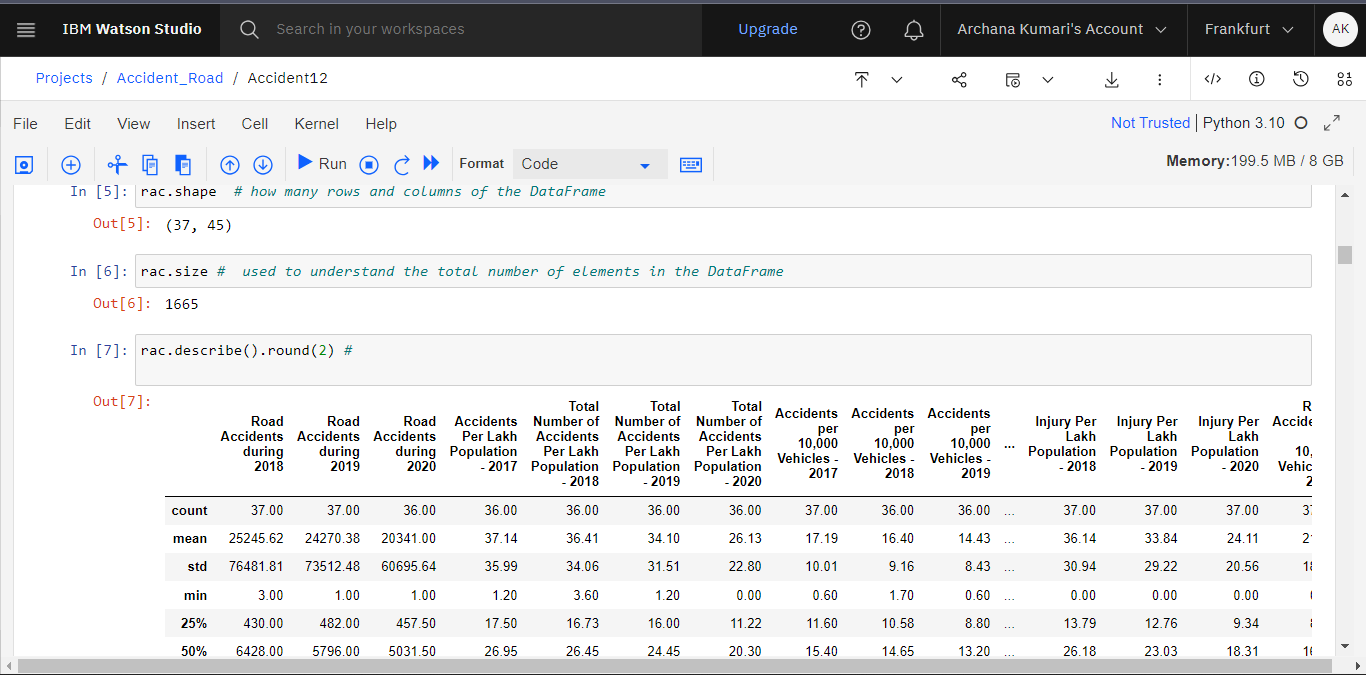
* rac.columns
* states = sorted\_rac['State']

states

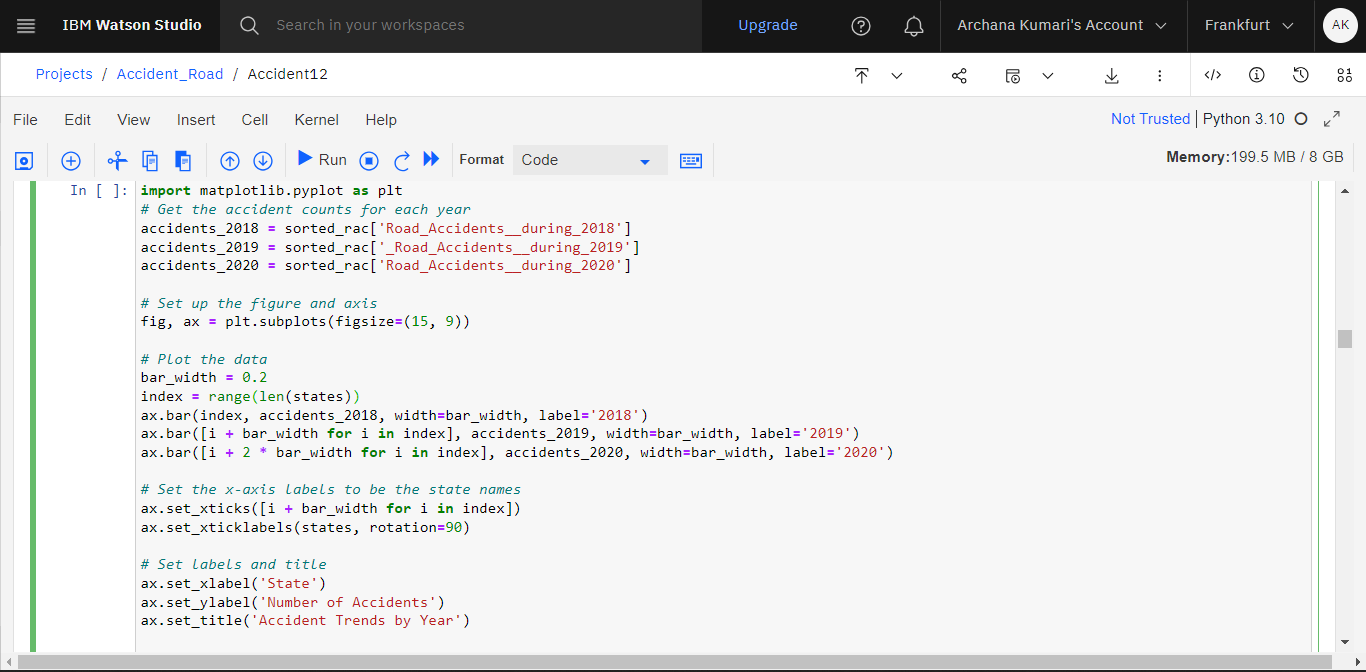
**TESTING**

**SCREENSHOT OF PROJECT**

**This is my overall how to create project image**

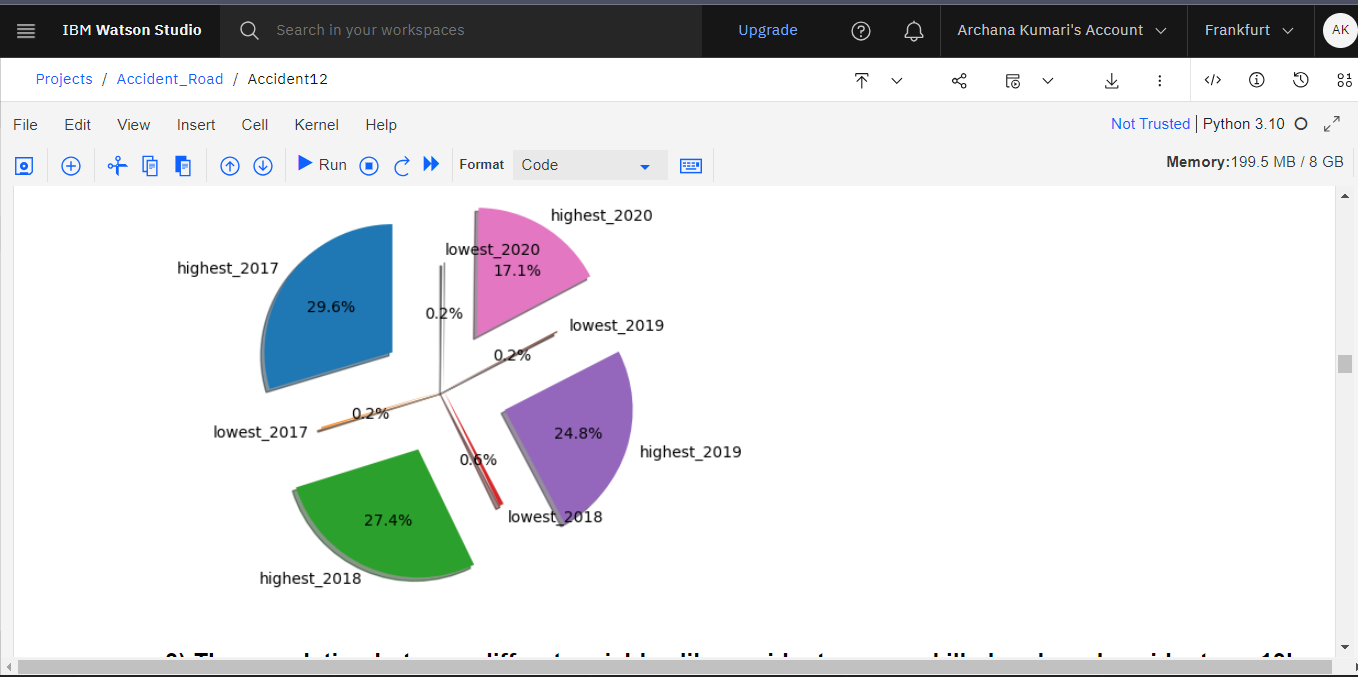


**Jupyter notebook some code**

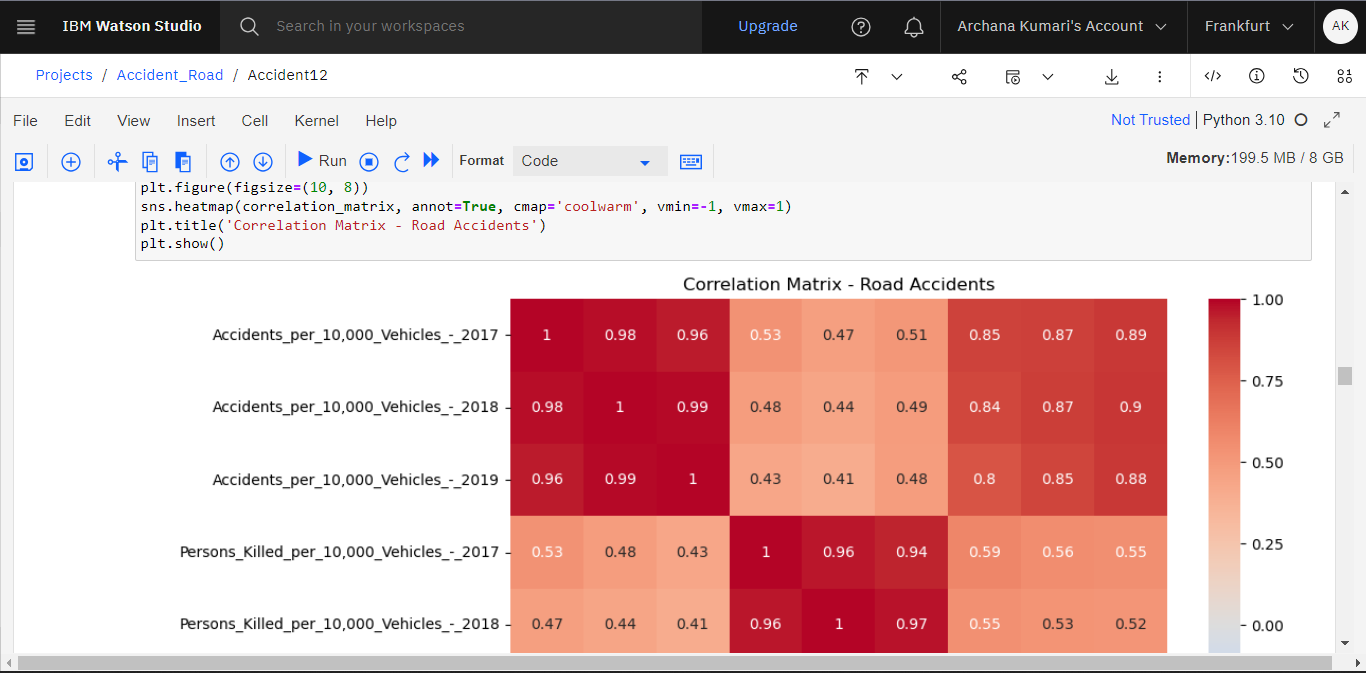
****

**Some Graph about road accident data analysis**

**In jupyter notebook**

****

**This is road accident data analysis details output**

****

**FUTURE SCOPE**

Here are some of the potential future features that could be added to an data analysis and Machine learning

Road accident data analysis is an evolving field that continues to leverage technological advancements and data-driven insights to enhance road safety. The future scope of road accident data analysis includes:

* Integration with other systems: The system could be integrated with other systems, such as a human resources system or a payroll system. This would allow for seamless data exchange and make it easier to manage library
* Data analytics: The system could be used to collect and analyse data about students. This data could be used to improve decision-making and identify areas for improvement.
* AI and machine learning: AI and machine learning could be used to improve the system's features and functionality. For example, AI could be used to identify patterns in library data or to automate tasks.

Advanced Predictive Analytics: Implement machine learning models to predict potential accident hotspots based on historical data, improving preemptive safety measures.

Real-time Monitoring: Develop a system for real-time accident monitoring using IoT devices and sensors, enhancing emergency response and reducing response times.

Integration with Traffic Management Systems: Collaborate with city authorities to integrate accident data with traffic management systems, enabling dynamic traffic rerouting for safer and more efficient transportation.

**CONCLUSION**

Enable data-driven decision-making, enhance road safety, and contribute to smarter urban planning through intuitive accident analysis dashboards on IBM Cloud's Watson Studio.

Road accident data analysis is pivotal for informed decision-making in road safety. Through data-driven insights, proactive measures can be implemented to prevent accidents. Continuous technological advancements, holistic approaches, and global collaboration contribute to creating safer roads and communities. The field's dynamic nature emphasizes the need for ongoing improvement and adaptation to address evolving challenges.

**REFERENCES**

[1] https://www.kaggle.com/

[2] https://www.youtube.com/watch?v=sGQfiyXOvF0