**** **Bansilal Ramnath Agarwal Charitable Trust’s**

**Vishwakarma Institute of Information Technology, Pune-48**

**(An Autonomous Institute affiliated to Savitribai Phule Pune University)**

**Department of Computer Science and Engineering (Artificial Intelligence)**

**LAB SUBMISSION**

**Data Science and Machine Learning**

**CAUA22201**

*Submitted by:*

*Vedant Santosh Deshpande,*

*PRN: 22210198,*

*Roll Number: 281015*

*Second Year*

*Semester II Academic Year 2023-24*

**Assignment: 5**

Aim: To visualize the data using R/Python by plotting the graphs for assignment no. 1 and 2. Consider a

suitable data set.

a) Use Scatter plot, bar plot, Box plot and Histogram

OR

b) Perform the data visualization operations using Tableau for the given dataset.

Theory:

To visualize the car crashes dataset using R or Python, we first need to load the dataset into our environment. Once loaded, we can use various plotting libraries such as matplotlib or seaborn in Python to create graphical representations of the data.

In Python, we can use matplotlib and seaborn libraries to create similar plots. With matplotlib, we can create scatter plots, histograms, and other types of plots to visualize the data. Seaborn provides high-level functions to create more sophisticated plots with less code, such as pair plots to visualize relationships between multiple variables simultaneously.

Overall, visualizing the car crashes dataset allows us to gain insights into patterns, trends, and relationships within the data, which can help in understanding the factors contributing to crashes and informing decision-making processes aimed at improving road safety. Additionally, visualizations provide a more intuitive way to communicate findings and results to stakeholders and decision-makers.

Here is the dataset that we have used:

“<https://drive.google.com/file/d/10z6R86pi-GGU4CWzQNJGqOAoDsxOW8db/view>”

Matplotlib.pyplot:

Matplotlib.pyplot is a state-based interface to matplotlib. It provides an implicit, MATLAB-like, way of plotting. It also opens figures on your screen, and acts as the figure GUI manager. Pyplot is mainly intended for interactive plots and simple cases of programmatic plot generation. It is imported as:

For Python Environment: pip install matplotlib

For Anaconda Environment: conda install matplotlib

For Google Colab: import matplotlib.pyplot as plt

Seaborn:

Seaborn is an amazing visualization library for statistical graphics plotting in Python. It provides beautiful default styles and colour palettes to make statistical plots more attractive. It is built on top matplotlib library and is also closely integrated with the data structures from pandas. Seaborn aims to make visualization the central part of exploring and understanding data. It provides dataset-oriented APIs so that we can switch between different visual representations for the same variables for a better understanding of the dataset.

For Python Environment: pip install seaborn

For Anaconda Environment: conda install seaborn

For Google Colab: import seaborn as sns

Results:

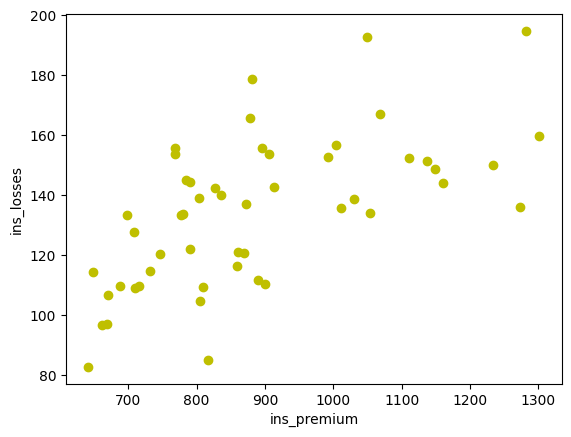
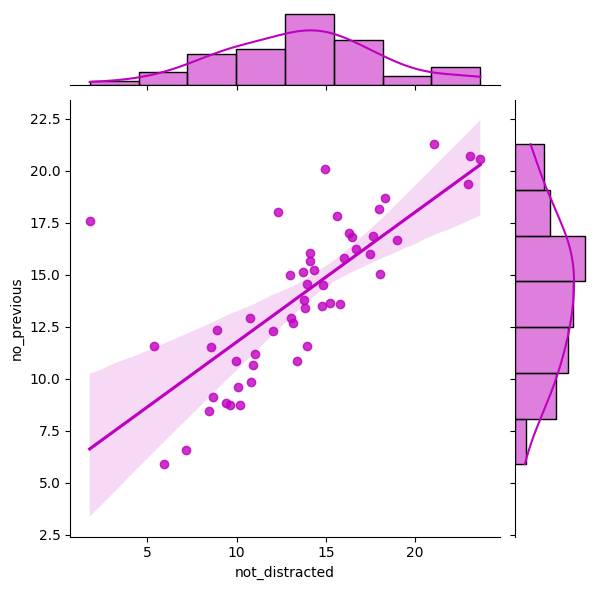
 

Fig. Scatter Plot Fig. Jointplot

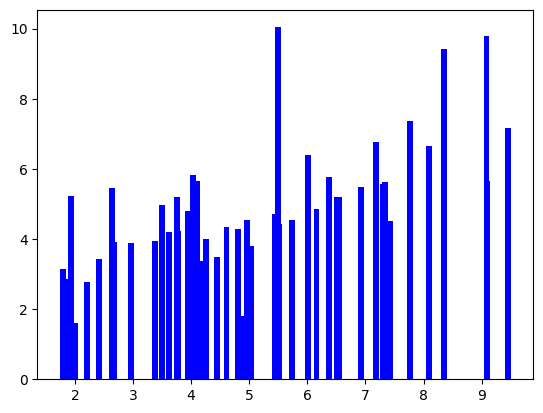
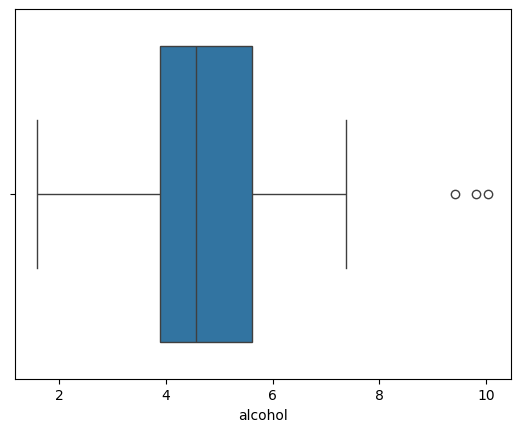
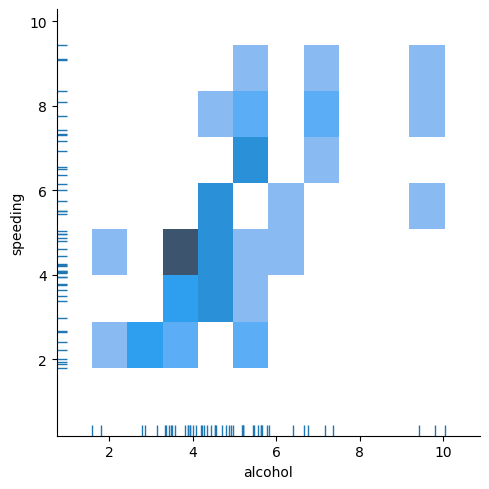
 

Fig. Bar Chart Fig. Boxplot

 Fig. Displot

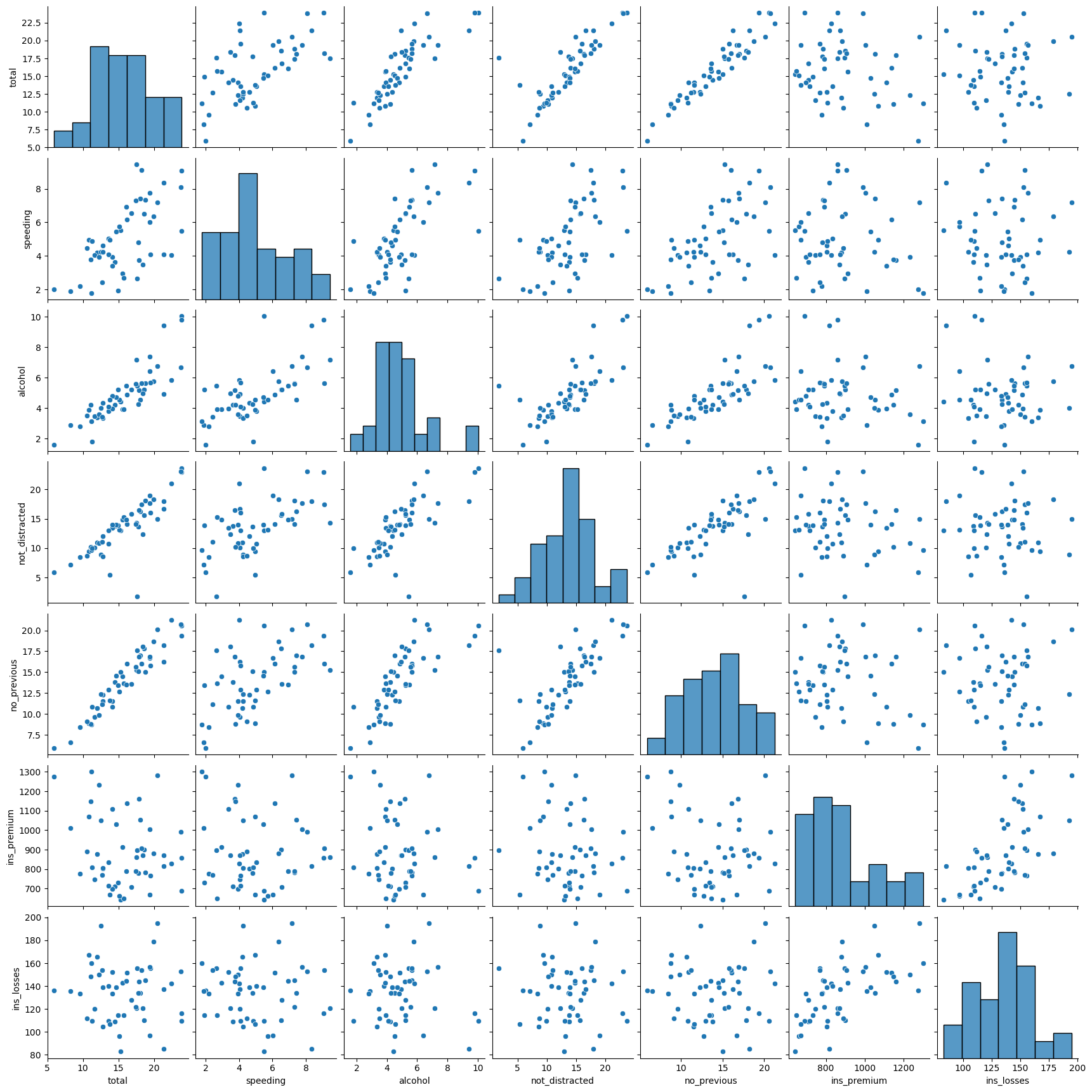


Fig. Pairplot

Conclusion:

In this assignment, we were able to visualise various charts and plots using most common libraries in Python- Matplotlib.pyplot and Seaborn.