**DOCTOR PROJECT VISIT ANALYSIS THAT IS BASED ON THEIR AGA AND OTHER FACTORS**

**Participants names:**

**Team ID :** LTVIP2023TMID09516

**Team size: 5**

**Team Leader : GOLLAPALLI VENKATA NAVEEN**

**Team member : RAHUL MEHTA**

**Team member : PEDAPATI ADITYA**

**Team member : PUDI OMKAR**

**Team member : VANAPALLI MOUNIKA VARALAKSHMI**

**INTRODUCTION** :

The problem statement of my project is doctor visit analysis that is based on their age and other factors how many patients are visited a doctor to cure their diseases.

The data consists of income, gender, age, illness, reduced and some other factors.

So in this project I am analyzing how these factors affects the patients and is they have benefited by visiting the doctor and some other queries.

**AGENDA :**

* The main agenda of this project is to analyze the dataset which contains the health status of different people.so based on this the following data.
* Load the Dataset and display first 15 rows.
* Total number of people based on the count of illness.
* Visualize and analyze maximum and minimum income.
* Correlation between different variables.
* Number of males and females affected by the illness

**OVERVIEW :**

* + Load the data set as “df=pd.read\_csv('DoctorVisit.csv’)”
  + Shape of the data set is defined as “df.shape”.
  + Information of the data set is defined as “df.info()”.
  + The purpose of the project is to identify which age people are mostly affected by illness and which gender people either male or female who are more affecting by the factors which are prescribed above.
  + So in this project I analyzed how the income of the project affected by visiting the hospital.

plt.figure(figsize=(10,10))

plt.scatter(x='income',y='visits',data=df)

plt.xlabel('income’)

plt.ylabel('visits')

* Total number of based on the count of illness it is analyed by using

df['illness'].value\_counts()

* Visualize and analyze maximum and minimum income.

y=list(df.income)

plt.boxplot(y)

plt.show()

* Top 10 Visits based on number of times visited.

top\_10=df.nlargest(10,'visits').set\_index('age’)

sns.barplot(x='visits',y=top\_10.index,data=top\_10)

**END USERS :**

* + The end users of this project are the normal people that they know how their lifestyle is damaged by their health.
  + As we seen in the project that how the income of the patients should be damaged based on their illness and regular visits of doctor.
  + By seeing this project I hope that some of the people may know about the importance of their health and by this way the normal people are benefitted by this project.
  + it is important for healthcare professionals to embrace data science and leverage its power to drive innovation and progress.

**SOLUTION AND ITS VALUE PROPOSITION :**

* + As we seen the project there are some problems regarding their health issues and their illness. So based on the problems which are identified from the project I should suggest some solutions to them.
  + The main solution is every individual should know about the importance of their health and also how the health problems will effect their lifestyle.
  + As observed in the project females are mostly affected by illness so they should take more care.

**CUSTOMIZING THE PROJECT AND MAKE IT YOUR OWN :**

* + I customize my project as one of the innovative creation which was done by me and it was helpful to others.
  + Normally we don’t found these type of data that is which was completely analyzed and identified the problems so based on these results the people’s mind set should change.
  + In this project I use some creative methods like representing the data in bar graphs, visualizing the data with histograms and scatterplots which are some new innovative creations and it may be unique.

**MODELLING :**

**Python Libraries & Frameworks:**

**NumPy Library :**

NumPy is a Python library used for Working with Arrays. It also has functions for working in domain of Linear Algebra,

fourier transform and matrices.

**Pandas Library :**

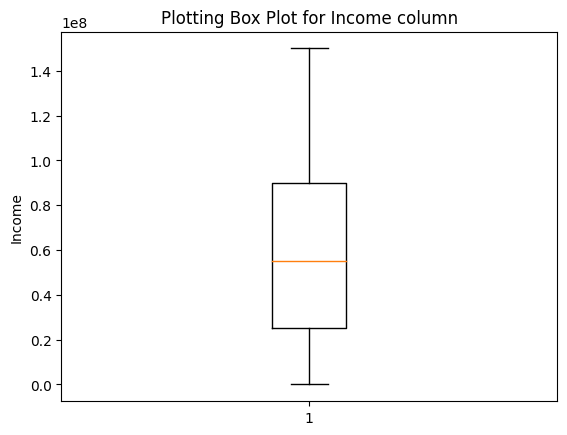
Pandas is a Python library used for Working with data sets. It has functions for Analyzing, Cleaning, Exploring, and

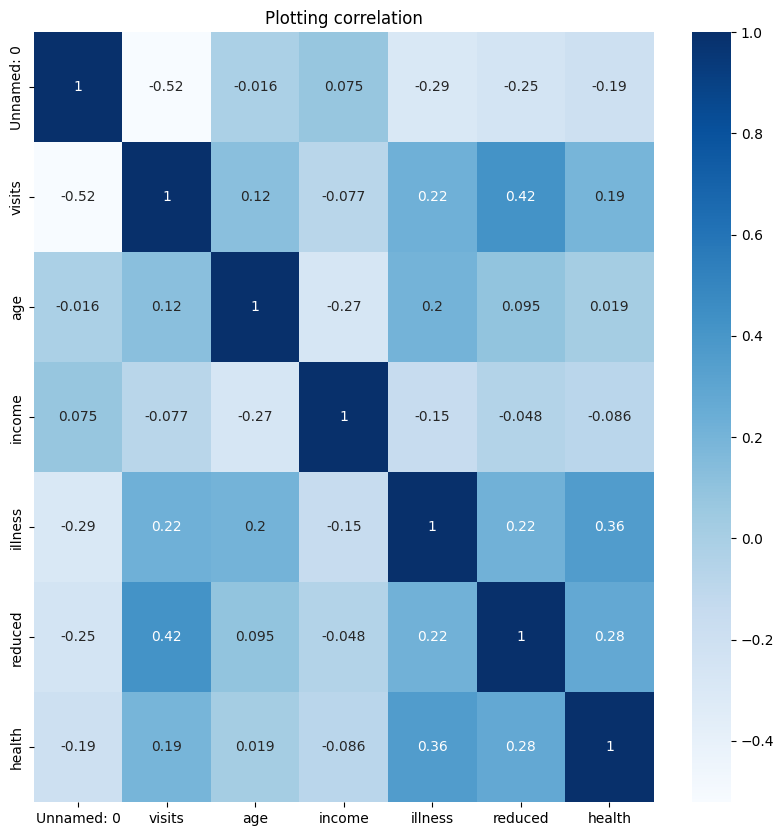
manipulating data.

**Matplotlib Library :**

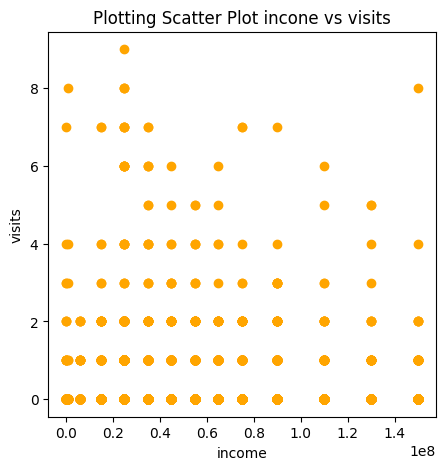
Matplotlib Library is a High Level Graph Plotting library in python that serves as a Visualization utility.

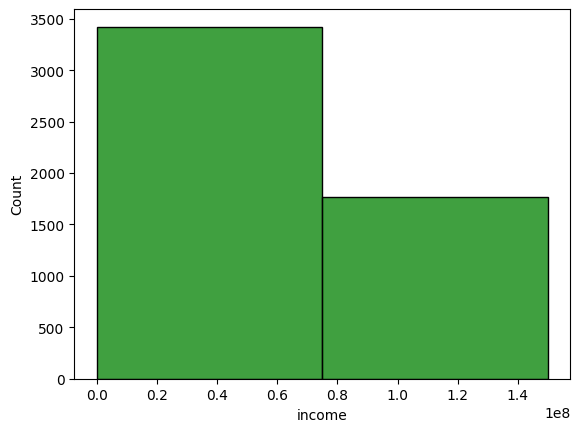
**RESULT :**

****

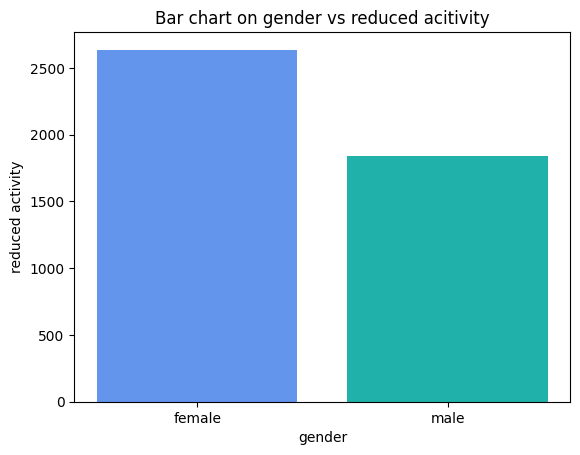
****

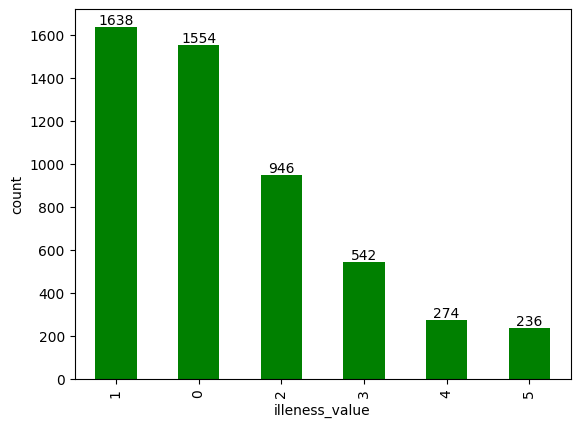
**RESULT :**

****

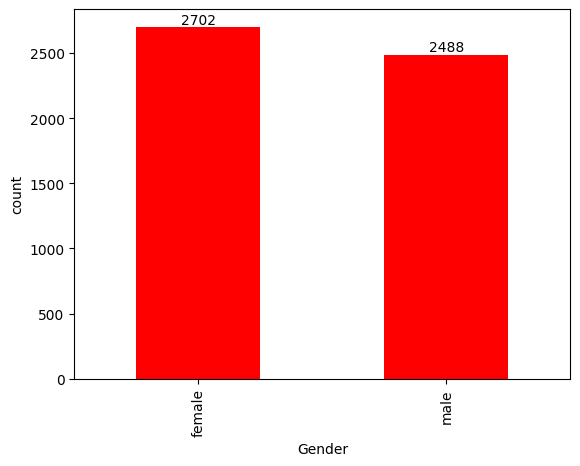
****

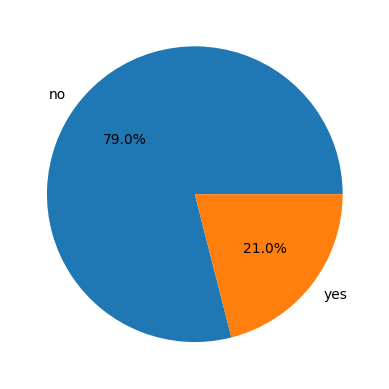
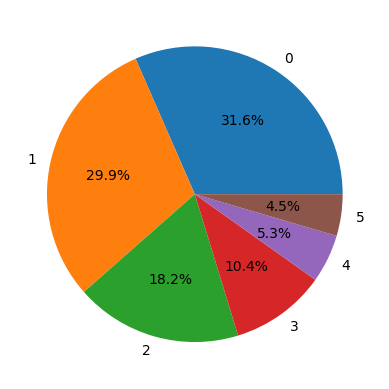
**RESULT :**

****

****

**RESULT :**

****

**** 

**LINKS:**

* Google colab link
* <https://colab.research.google.com/drive/1F4SqdTk-GXMIrxB92Q6gpJRDGz_1YsxL?usp=sharing>

**CONCLUSION:**

In Conclusion, we have learned that Data Science is a vital field that can provide valuable insights and inform important decisions. By understanding the fundamentals of data collection, preparation, analysis, and visualization, we can unlock the full potential of Data and use it to Drive Innovation and progress.