



DAB 304 HEALTHCARE ANALYTICS
PROJECT REPORT
“BRAIN STROKE ANALYSIS”

SUBMITTED BY:

VAIBHAV PATEL (0772934)

KUSH PATEL (0792972)

HIREN PATEL (0797548)

AKSHAR PATEL (0793127)

INTRODUCTION:

- Healthcare Analytics is the process of analysing current and historical industry data to predict trends, improve outreach, and even better manage the spread of diseases. The field covers a broad range of businesses and offers insights on both the macro and micro level.
- Healthcare Analytics deals with the health-related data of the patients, their medical history, cost while staying in hospital, data related to hospital management, diagnosis of patients and the record of treatment related to that diagnosis of the diseases.
- A stroke occurs when a blockage or bleed of the blood vessels either interrupts or reduces the supply of blood to the brain. When this happens, the brain does not receive enough oxygen or nutrients, and brain cells start to die.
- Stroke is a cerebrovascular disease. This means that it affects the blood vessels that feed the brain oxygen. If the brain does not receive enough oxygen, damage may start to occur.
- This is a medical emergency. Although many strokes are treatable, some can lead to disability or death.
- Rapid diagnosis is important for reducing brain damage and enabling the doctor to treat the stroke using a suitable method for the type.
- Each type of stroke has a different set of potential causes. Generally, however, stroke is more likely to affect a person if they:
 - have overweight or obesity.
 - are 55 years of age or older.
 - have a personal or family history of stroke.
 - have high blood pressure.
 - have diabetes.
 - have high cholesterol.
 - have heart disease, carotid artery disease, or another vascular disease.
 - are sedentary.
 - consume alcohol excessively.
 - smoke.
 - use illicit drugs.

- Some studies have found that males have a higher risk of death from stroke than females. However, one report suggests that these differences do not take into account adjustments for race, age, the severity of the stroke, and other risk factors.
- The results explain that the risk of stroke mortality often increases due to age and demographic, rather than the biological differences between males and females.

MOTIVATION:

- The main motive behind our project is that when we were looking for various data to select for our project at that time, we were reading different health articles and blogs on Google, from which one of our team members came to know about brain-stroke data.
- Further, when we looked into that article, we were amazed to read that brain-stroke is a kind of disease that is known from so many years but, majority of the people is not aware about the disease itself, its causes and its immediate symptoms. So, we thought we should focus on this topic to bring awareness in the society about this deadly disease if not treated on time.
- This is the motive behind our selection of this topic for our project.

RELATED WORK:

- The similar project was carried out by the other medical students using the same data. In their analysis, they used concepts of Machine Learning Models and other Visualization tools in Python. They made use of just a decision tree model to perform the Exploratory Data Analysis on the data. Further, they only used accuracy metrics to judge the performance of the model. They have not made use of Tableau to do more Exploratory Data Analysis on the data-set to create a visualization dashboard for the audiences. They used concepts of probability to predict the causes of the brain stroke.

- We searched a lot for before selecting this topic to do our project. In this field many students have done different types of their own analysis, so we decided to do our own research analysis than others.
- Their analysis was carried out in python using different features from the data to generate various graphs. We used Tableau to visually present our Exploratory Data Analysis.
- We tried to use different concepts of Machine Learning Models to predict causes of brain stroke based on the various metrics of the Machine Learning Models.
- We tried to do further analysis on this topic which was not done by other students.

METHODS:

- Following methods and tools were used to execute the project:
- **Data Gathering:**
 - There are various ways to collect data for the brain stroke analysis. There are so many repositories in which data related to brain stroke is available.
 - We decided to use brain stroke data-set provided by Kaggle in this project as we felt that the data provided by the Kaggle is best suitable for our project.
 - Our dataset cites the causes of brain-stroke and different types of brain-strokes. There are different reasons due to which brain-stroke occurs such as hypertension, recurring heart disease, the type of work people does, obesity, recurring disease such as diabetes, and smoking. It contains 11 columns and around 5000 rows.
- **Data Cleaning:**
 - Data Cleaning is the most important phase before starting exploratory data analysis in the project. It helps us to find and remove null values, correcting mistakes in the data, locating discrepancies in the data, and to make sure data is fully prepared to do analysis.

- It is crucial to look for the inconsistencies or inaccuracy in the data before proceeding ahead with the project.
- It is important to look for any missing values or duplicates in the data before performing exploratory data analysis to have accurate result. We can use various python functions to fill in the missing values or to remove the missing values. We can also use python functions to remove any duplications in the data-set. It is also very crucial, to check for the outliers and to remove the outliers before using data to get correct and accurate outcomes. To perform Exploratory Data Analysis, we must have clean and accurate data get valid outcomes.

EXPLORATORY DATA ANALYSIS:

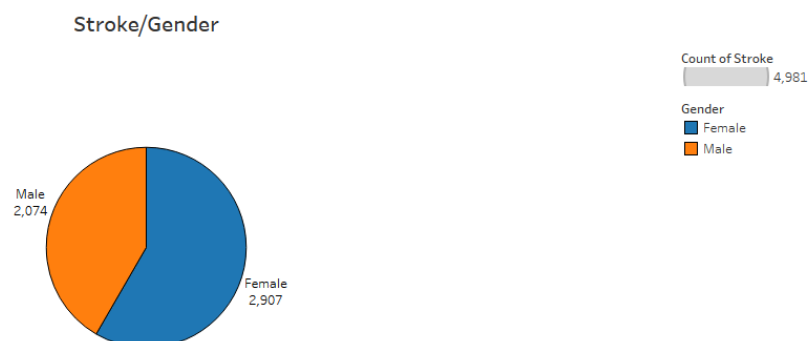
- This is very crucial stage in any project. Exploratory data analysis is used to explore the data and to use that data to predict an outcome.
- There are different ways to explore any data using python or Tableau.
- We can use python to explore data and to find any outliers or unnecessary data that can harm the accuracy of the data.
- We can use Tableau to visually explore the data and to present data in visual form.
- We did Exploratory Data Analysis and Descriptive Analysis on the causes of brain-stroke based on our data with the help of various Machine learning concepts and we are also going to make various predictions on the causes of brain-stroke and also one can avoid such situation by following healthy life. Further, to support our predictions we also built an interactive dashboard using the concepts of Tableau.
- Exploratory Data Analysis helps us to find any unusual data in the data-set and it helps to tackle such situations.

DATA VISUALIZATIONS:

- We used different types of charts in Tableau to present our data in visual form.
- We used Tableau to create charts and to create dashboard containing all the charts to provide better understand about the data.
- We created different charts in Tableau to explore our data and to present those outcomes in visual form.
- We predicted causes of brain-stroke based on visuals created in Tableau.
- We also perform Exploratory Data Analysis on our data using various charts and visuals of Tableau.
- Data Visualization is one of the most important tool to visualize our outcomes in visuals in order to convey our motive to the audience in easy manner.

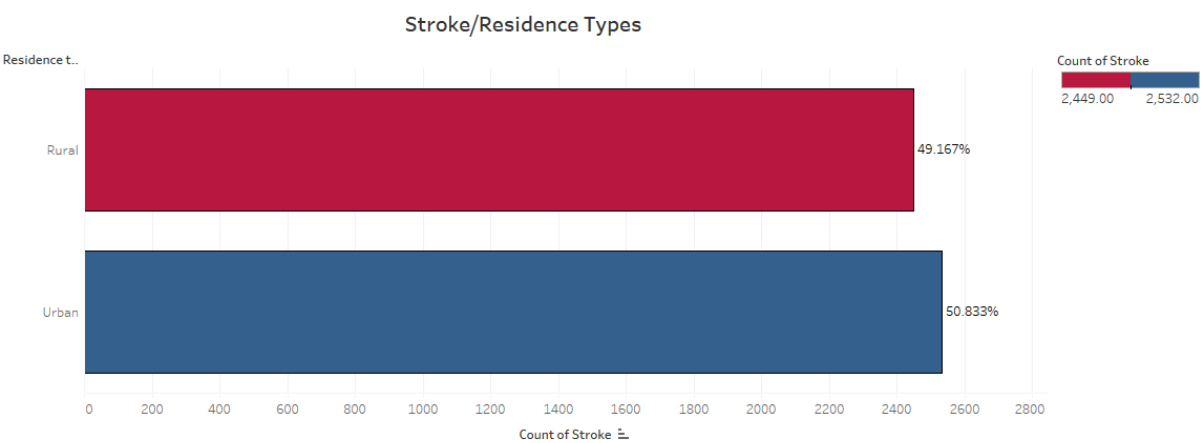
RESULTS:

Gender affected by Brain-Stroke



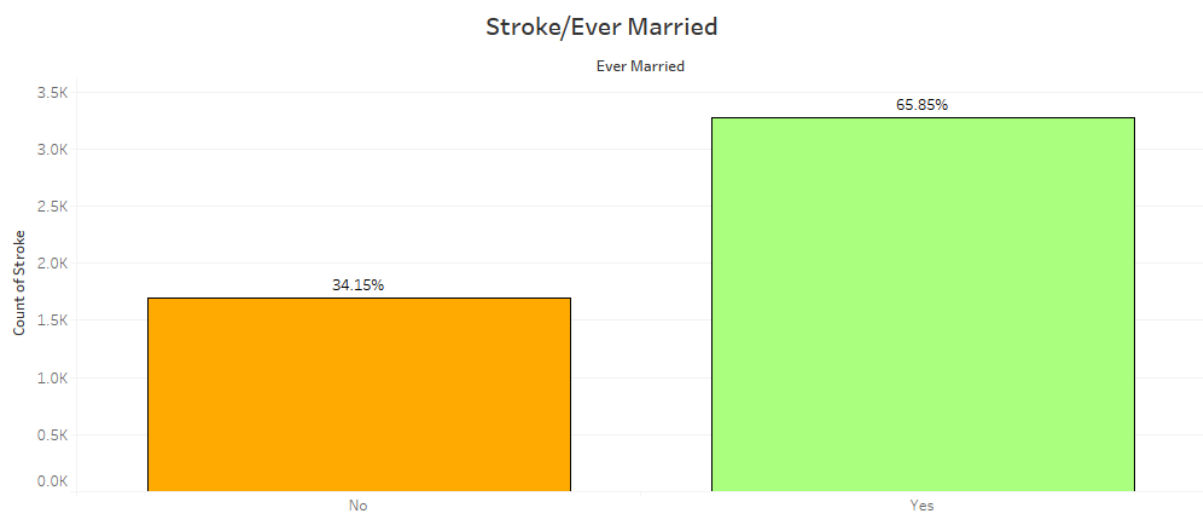
Gender and count of Stroke. Color shows details about Gender. Size shows count of Stroke. The marks are labeled by Gender and count of Stroke.

Residence got affected by Brain-Stroke



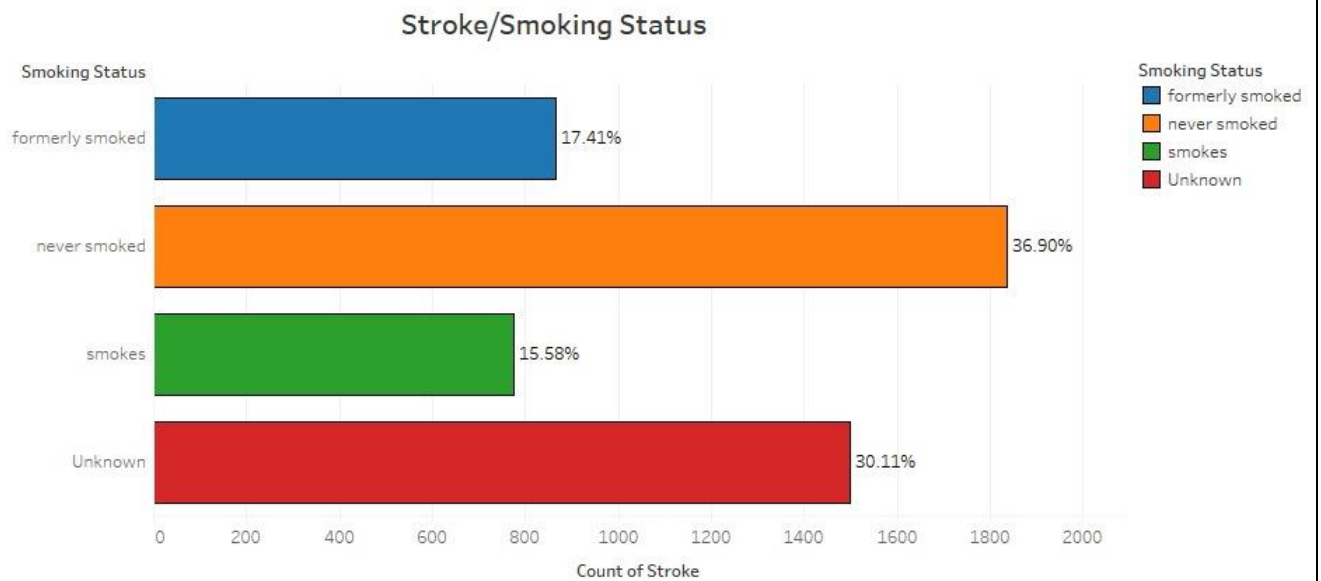
Count of Stroke for each Residence type. Color shows count of Stroke. The marks are labeled by % of Total Count of Stroke.

Stroke by Martial Status



Count of Stroke for each Ever Married. Color shows count of Stroke. The marks are labeled by % of Total Count of Stroke.

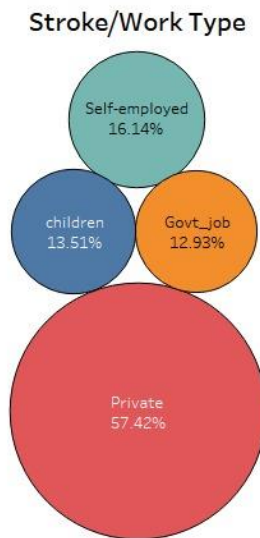
Stroke by Smoking habits



Brain Stroke due to Hypertension

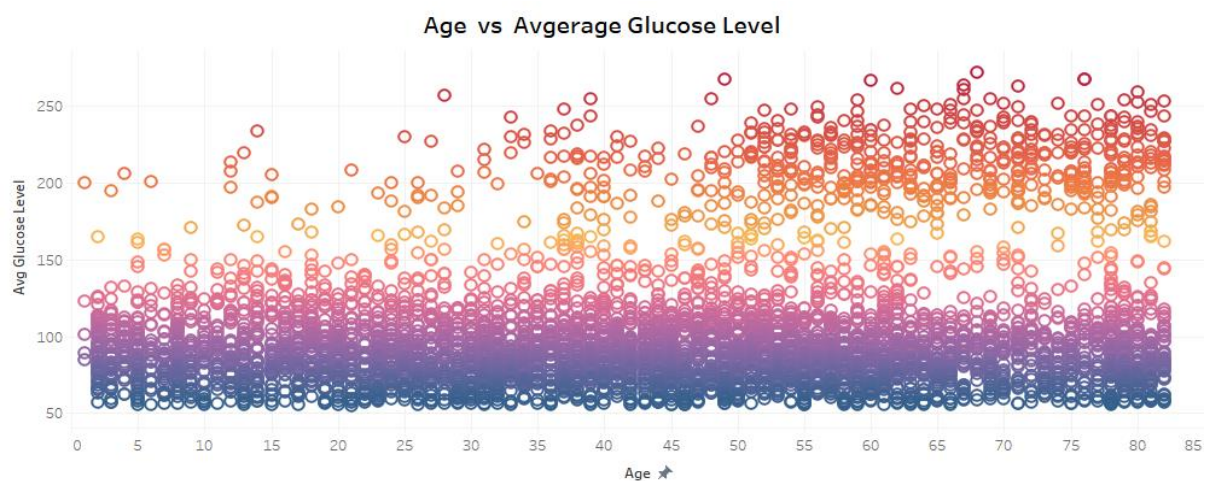


Types of work causes Brain Stroke



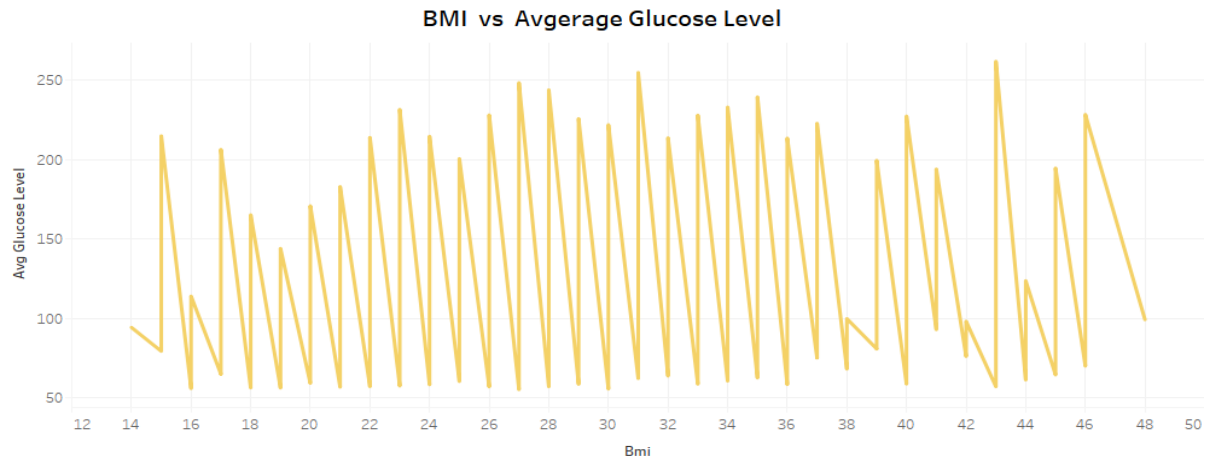
Work Type. Color shows details about Work Type. Size shows count of Stroke. The marks are labeled by Work Type.

Glucose level per Age factor which affects brain stroke



Age vs. Avg Glucose Level. Color shows details about Avg Glucose Level. The view is filtered on Age, which ranges from 1 to 82. Avg Glucose Level is high among old people.

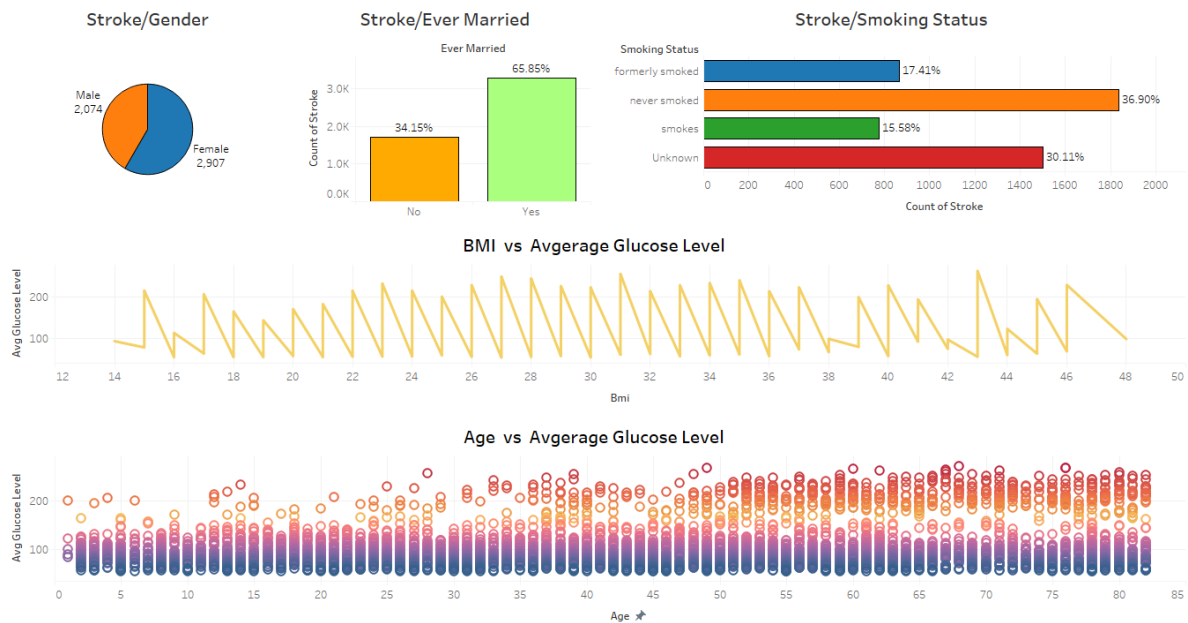
Glucose level per Body Mass Index which affects brain stroke



Bmi vs. Avg Glucose Level. Color shows count of Bmi. The data is filtered on Avg Glucose Level and Bmi. The Avg Glucose Level filter keeps 3,895 of 3,895 members. The Bmi filter has multiple members selected. Bmi > 40 have low avg glucose.

DASHBOARD

BRAIN STROKE ANALYSIS DASHBOARD



CONCLUSIONS:

- The best way to prevent a stroke is to address the underlying causes. People can achieve this by making lifestyle changes such as:
 - eating a healthful diet.
 - maintaining a moderate weight.
 - exercising regularly.
 - not smoking tobacco.
 - avoiding alcohol, or only drinking moderately.
- Eating a nutritious diet including plenty of:
 - Fruits.
 - Vegetables.
 - Whole grains.
 - Nuts.
 - Seeds.
 - Legumes.
- Be sure to limit the amount of red and processed meat in the diet, as well cholesterol and saturated fats. Also, moderate salt intake to support healthy blood pressure levels.
- This, are the important measures one can take and maintain in their life to live a healthy life.

CONTRIBUTIONS:

- We have 4 people in our group namely, Vaibhav, Kush, Hiren and Akshar.
- We distributed all work evenly between 4 of us. In which, Kush and Hiren were responsible for the finding various articles, studying them after, that we all make contributions to decide the topic and related dataset. Moreover, after deciding everything Vaibhav and Akshar made the project proposal.
- After that we all were in every step of the project phase's right from the start to the end.
- We all equally contributed in making this project and also in making this report and power-point presentations.

REFERENCES:

<https://www.kaggle.com/datasets/jillanisofttech/brain-stroke-dataset>

<http://www.google.com>

<https://www.medicalnewstoday.com/articles/7624>

<http://www.keras.io>

<http://www.tableau.com>