Name	Rucha Kulkarni
Class	BE Computer Engineering (Batch F)
UID	2021300067
Exp No.	3

Aim: Design Interactive Dashboards and Storytelling using Tableau / Power BI / R (Shiny) / Python (Streamlit/Flask) / D3.js to be performed on the dataset - Disease spread / Healthcare

● Create interactive dashboard - Write observations from each chart given below

● (Advanced - Word chart, Box and whisker plot, Violin plot, Regression plot (linear and nonlinear), 3D chart, Jitter, Line, Area, Waterfall, Donut, Treemap, Funnel

● Basic - Bar chart, Pie chart, Histogram, Timeline chart, Scatter plot, Bubble plot)

## Description:

Dataset used is Stroke Dataset available at
https://www.kaggle.com/datasets/fedesoriano/stroke-prediction-dataset/data

Following are the columns in the dataset:

1) id: unique identifier

2) gender: "Male", "Female" or "Other"

3) age: age of the patient

4) hypertension: 0 if the patient doesn't have hypertension, 1 if the patient has hypertension

5) heart_disease: 0 if the patient doesn't have any heart diseases, 1 if the patient has a heart disease

6) ever_married: "No" or "Yes"

7) work_type: "children", "Govt_jov", "Never_worked", "Private" or "Self-employed"

8) Residence_type: "Rural" or "Urban"

9) avg_glucose_level: average glucose level in blood

10) bmi: body mass index

11) smoking_status: "formerly smoked", "never smoked", "smokes" or "Unknown"*

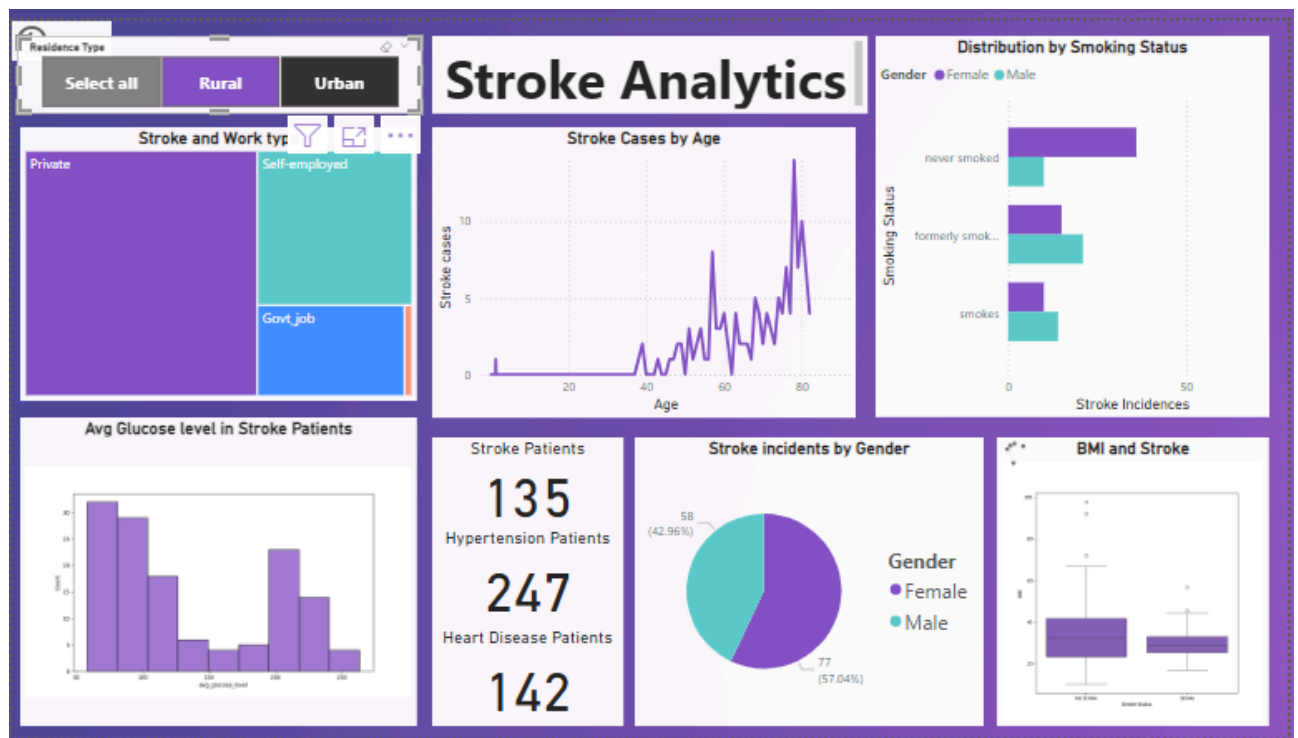12) stroke: 1 if the patient had a stroke or 0 if not

Dashboard:

(Initial/Default):
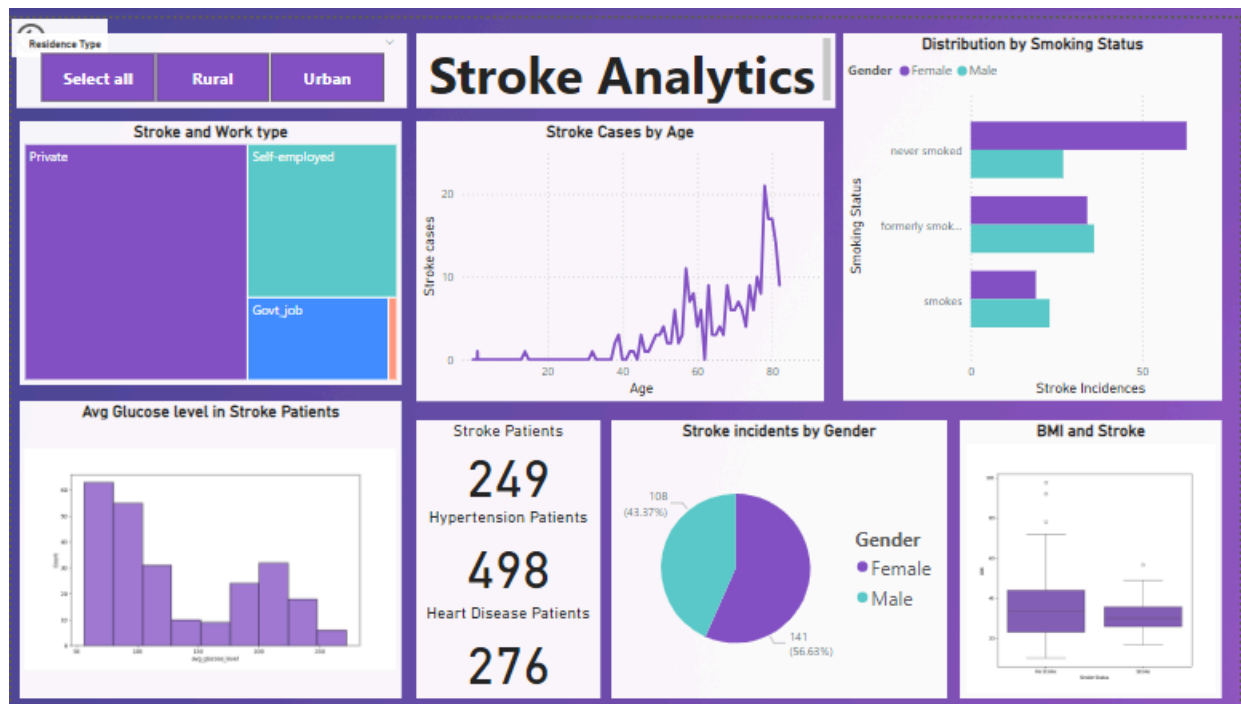
Upon Clicking Rural:
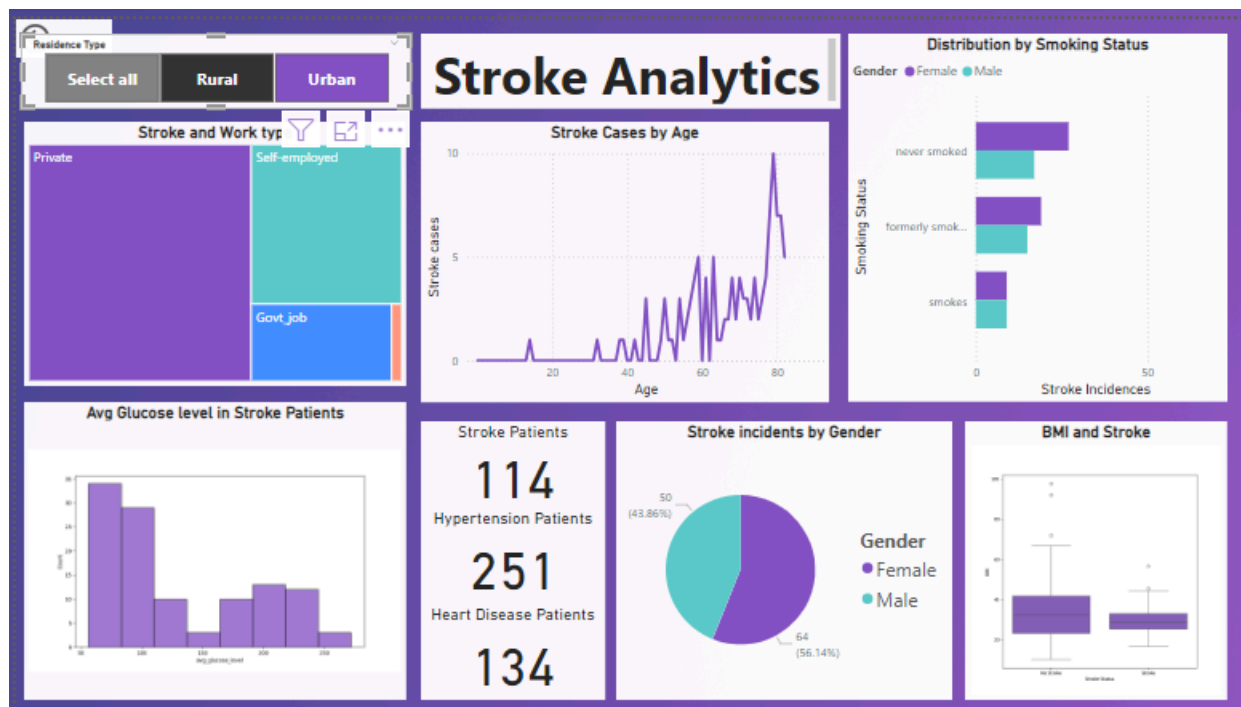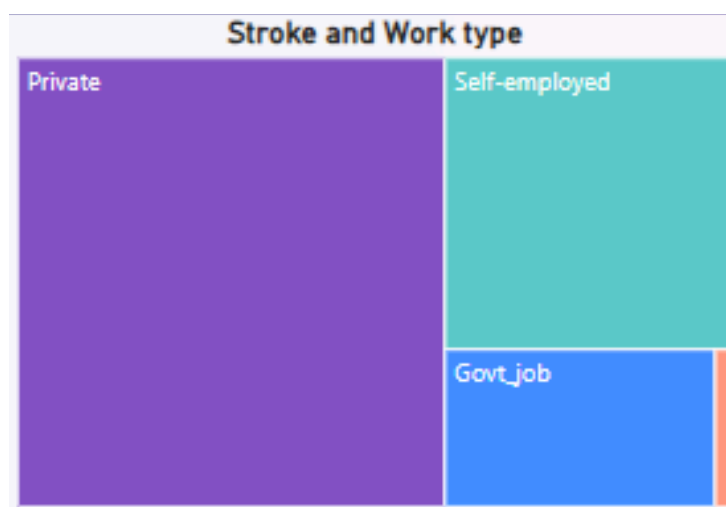
Upon Clicking Urban:

Observation from dashboard:

The dashboard provides insights into various aspects related to stroke patients. There is a filter at top allowing users to select rural or urban residence types of the patients. In rural areas, there are a comparatively lesser number of people who got strokes than in the urban areas.
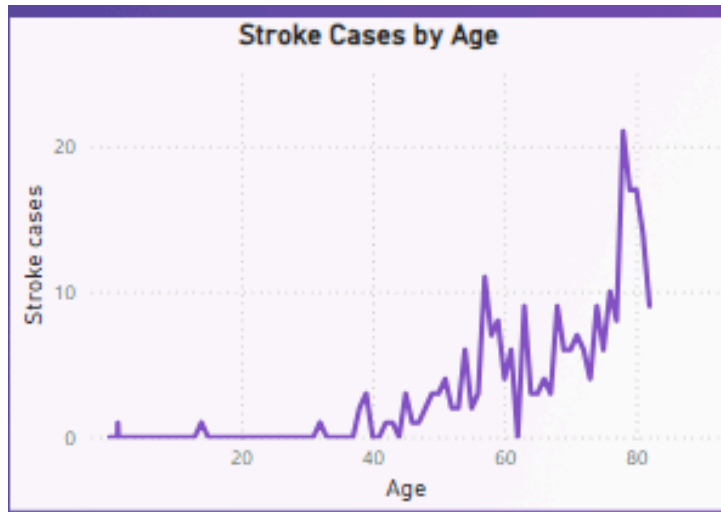
Graphs:

TreeMap:

This TreeMap shows the distribution of stroke cases across different work types.

Most stroke cases are in Private and Self-employed work types, fewer in Govt. jobs.

This may indicate higher stress or lifestyle risks in non-government jobs. Better access to healthcare in government jobs might contribute to fewer cases.
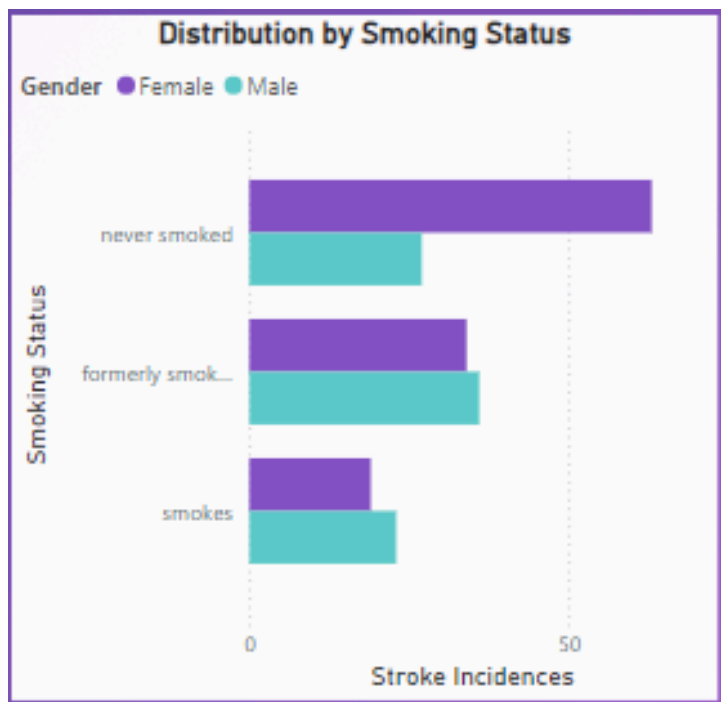
Line Chart:

The line chart displays the number of stroke cases by age, with an upward trend in the number of cases as age increases.

Stroke cases increase sharply with age, especially after 50, peaking around 70-80 years.

So Age is a major risk factor for stroke incidences.

Bar Chart:

This chart illustrates the distribution of stroke incidences across different smoking statuses: never smoked, formerly smoked, and smokes, with differentiation by gender (Female and Male).Most stroke cases are among never smoked individuals, followed by formerly smoked.

Histogram:
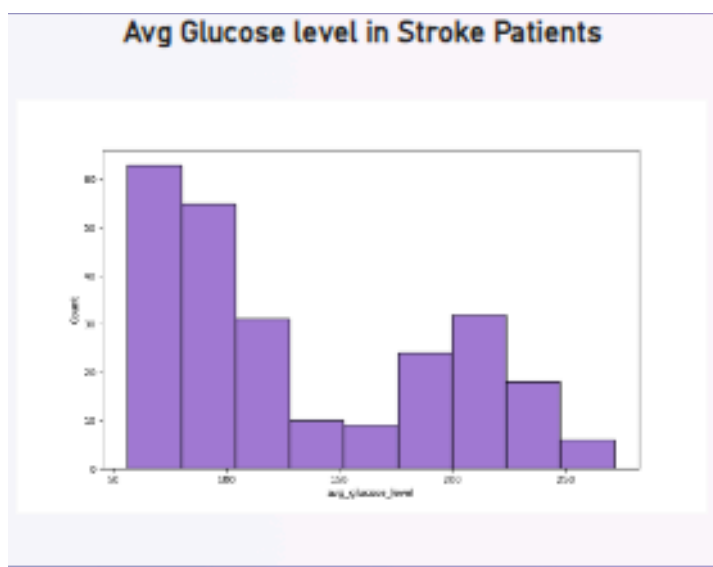
The histogram depicts the distribution of average glucose levels among stroke patients.The distribution is skewed towards higher glucose levels, suggesting that stroke patients have generally higher glucose levels.

Pie Chart:
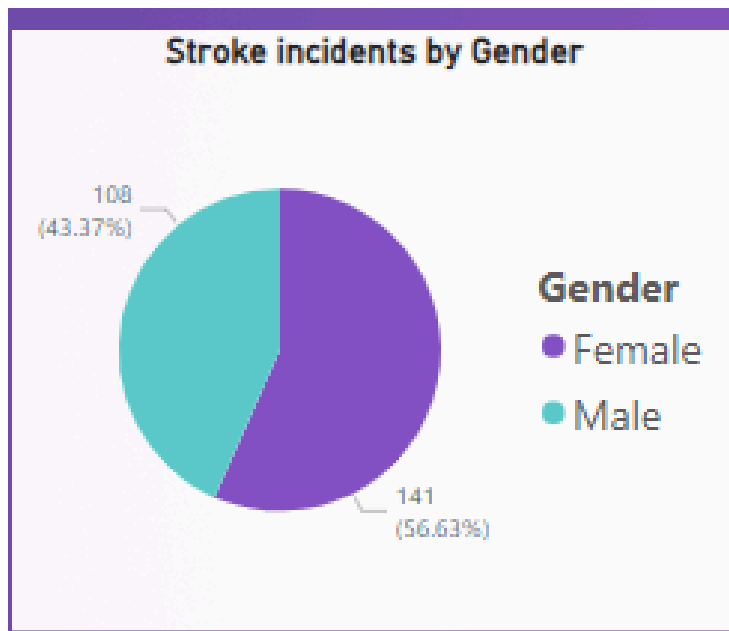
The pie chart shows the proportion of stroke incidence by gender, with females having a higher percentage of incidents compared to males.

Box Plot:
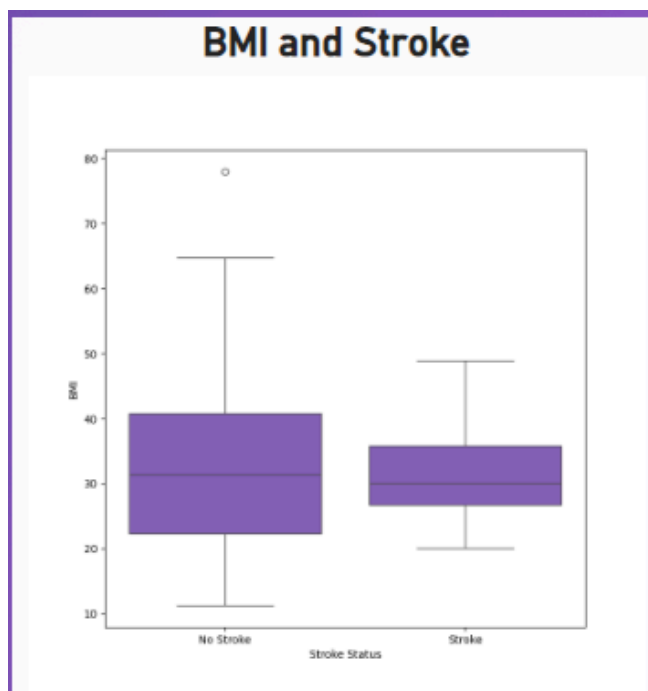
The box plot shows the distribution of BMI values for two groups: individuals with "No Stroke and those with "Stroke.

The median BMI is similar for both No Stroke and Stroke groups.

The No Stroke group shows slightly more variability in BMI.

There is a BMI outlier in the No Stroke group.

Conclusion:

From this experiment,we learnt to create a dashboard using Power BI. It made it clear how factors like age, gender, and BMI affect stroke risk and helped spot patterns and trends easily. This showed how useful visual tools can be for making complex data easier to understand and use in healthcare decisions.