International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 07 | July - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

Smoking and Stroke: Role of Sociodemographic Factors

Anishaa Nopany

Abstract:

"You cannot affirm the power plant and condemn the smokestack, or affirm the smoke and condemn the cough"—Wendell Berry, The Gift of the Good Land, 1981. The 1st law of ecology by Ecologists Barry Commoner states that nature and its web-like workings are more complex and resilient than the evolution of a single organism. The research study provides a deeper understanding of the relationship between smoking tobacco and the occurrence of brain strokes. It uses a secondary dataset 'Brain stroke prediction dataset' provided by Izzet Turkalp Akbasli to search whether the association between smoking and the risk of stroke is causal or is in fact spurious in nature.

Introduction:

"The doctor of the future will give no medicine, but will instruct his patients in the care of the human frame, in diet, and in the cause and prevention of disease." – Thomas Edison

A stroke occurs when the blood supply to the brain is disrupted which prevents the oxygen from reaching the blood vessels of that part of the brain. Due to this oxygen starvation, a brain stroke befalls which could result in either temporary or permanent brain damage, including paralysis, memory, and speech impairment, or it could even lead to death. According to the data from Centres for Disease Control and Prevention, every 40 seconds someone in the United States has a stroke, and out of those, every 3.5 minutes someone dies of that stroke. In the United States, 1 in every 4 individuals has a stroke once in their lifetime which accounts for 25% of the population suffering from this condition. Additionally, strokes affect 15 million people worldwide of which 5 million die, and another 5 million are left permanently crippled. The World Stroke Organization estimated that the treatment, rehabilitation, and indirect costs for stroke account for more than US$700 billion every year. This stroke burden annual estimation is predicted to go over US$1 trillion for the global economy by the year 2030. Another point to note is that these values are highly underestimated as it does not account for the overhead out-of-pocket costs and the income losses that occurred due to the onset of the additional responsibilities to the caregiver. With stroke being a leading cause of death amongst people in the US and having massive opportunity costs as seen from the data values above, it becomes critical for us to evaluate the factors, both personal and social, that might affect this grave ailment in order to establish the methods to address and prevent the shortcomings of this condition.

Literature Review:

Numerous research has examined how smoking impacts the risk of stroke, with a particular focus on the link between tobacco use and brain stroke. Chen et al (2019) in 'Impact of Smoking Status on Stroke Recurrence' stated how there exists a dose-response relationship between smoking quantity and the risk of stroke recurrence wherein persistent smoking raises the likelihood of a stroke recurrence following a first stroke. Shah et al (2010) in 'Smoking and Stroke: The more you smoke the more you stroke' added to the past literature by including the socioeconomic status of individuals as an important influencing factor. Putaala et al in 'Silent brain infarcts and leukoaraiosis in young adults with first-ever ischemic stroke' concludes smoking to be an independent risk factor for silent brain infarcts.

Many papers have worked on this relationship by studying different angles of the impact of smoking on the probability of brain strokes. However, not much work has been done in accessing how confounding factors such as other lifestyle choices, existing biomechanisms, and different individualistic underlying mechanisms affect the relationship between smoking tobacco and the risk of brain stroke along with understanding the adverse impacts of tobacco use on health and aid in developing of effective approaches to reduce the incidence of brain stroke in smoking-affected communities.

Key stakeholders:

Due to the imminent dangers of strokes, data on the quantification of the effects of various factors leading to stroke could be of incredible value to various parties such as-

- Governments around the world to help set out strategic roadmaps and policy interventions for the prevention of stroke by increased taxation and control measures on harmful products such as tobacco and trans-fats, improved public awareness for the various risk factors, development in the technology of medical aids currently available and the provision of universal health coverage for low-cost clinical proceedings.
- Pharmaceutical companies to better understand the causes of strokes and to develop more enhanced measures for its prevention and aid.
- Individuals and their families both those affected by this condition and those who would be willing to make significant changes to their lifestyle to prevent the onset of this stroke condition.

Data:

For this paper, I use the dataset ‘Brain stroke prediction dataset’ by Izzet Turkalp Akbasli, available on Kaggle. This dataset, updated 4 months ago, has been formulated using the information present in a paper by M. S. Pathan et al called ‘Identifying Stroke Indicators Using Rough Sets’, published in the scientific journal published by the Institute of Electrical and Electronics Engineers (IEEE) in 2020.

There are 11 variables and their observation points in 4983 columns. The variables used in the data set to measure stroke are-

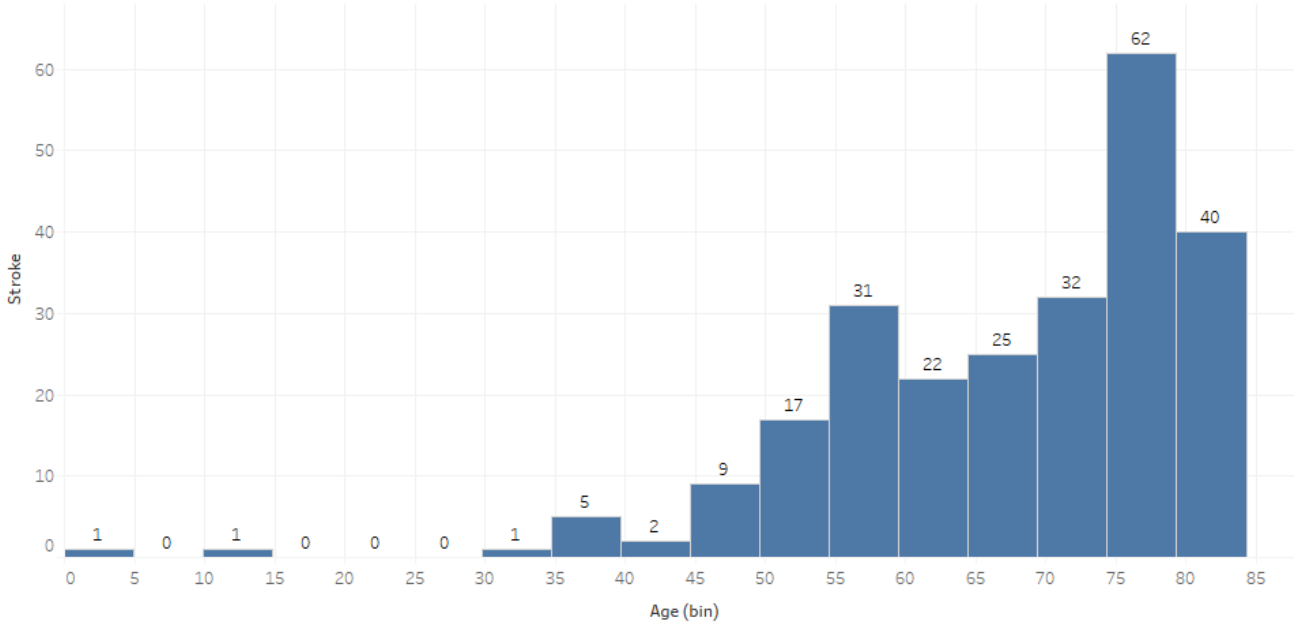
- Gender- “Male” or “Female”
- Age- The age of the patient
- Hypertension- 0 if the patient does not have hypertension, 1 if the patient has hypertension
- Heart Disease- 0 if the patient does not have any heart diseases, 1 if the patient has a heart disease
- Marriage- “Yes” if married, otherwise “No”
- Work type- "Children", "Govt job", "Never worked", "Private" or "Self-employed"
- Residence type- "Rural" or "Urban"
- Average glucose level- Average glucose level in the blood
- BMI- The body mass index of the patient
- Smoking status- "Formerly smoked", "Never smoked", "Smokes" or "Unknown"
- Stroke- 1 if the patient had a stroke and 0 if not

All these variables are discrete variables out of which-

- Gender, Hypertension, Heart disease, Marriage, Residence type, and Stroke are binary variables taking only two values.
- Age, Average glucose level, and BMI take numerical values.
- Work type is a categorical variable that cannot be quantified and Smoking status is an ordinal variable that can be ranked based on preferences.

Before any analysis of the data set, I would first like to establish, according to my data values, the effects on age and the average glucose level, BMI, heart disease, and hypertension according to the residence and work type of the individual patients.

Avg. Age	43.4
Stroke	248.0



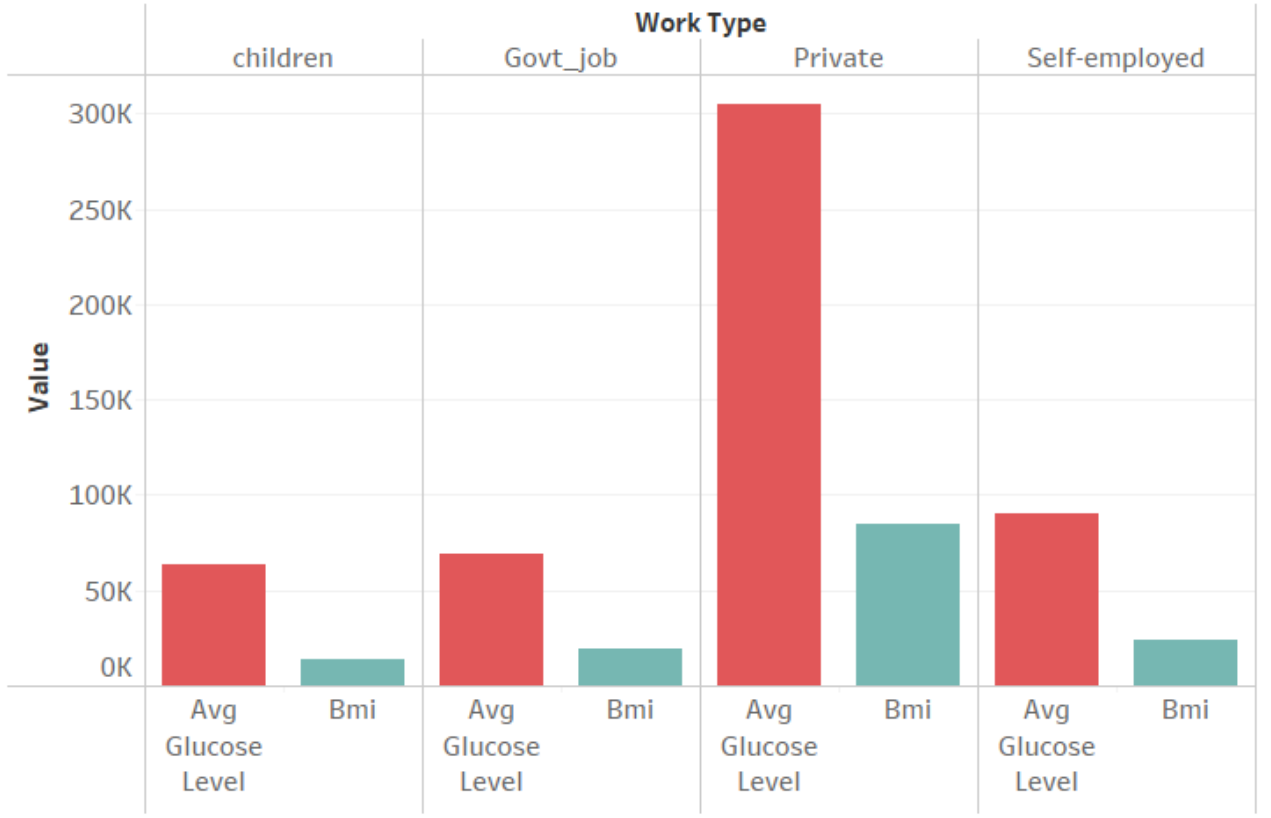
- Out of our extensive data set, the average age of patients having a stroke is 43.4 and the second highest number of strokes occurred has been for the age group 55-60. This is surprising as according to the World Stroke Organisation; the incidence of stroke increases with age and 60% of the strokes occur to those aged between 65 to 70 years of age with 16% strokes occurring to those under 55. As

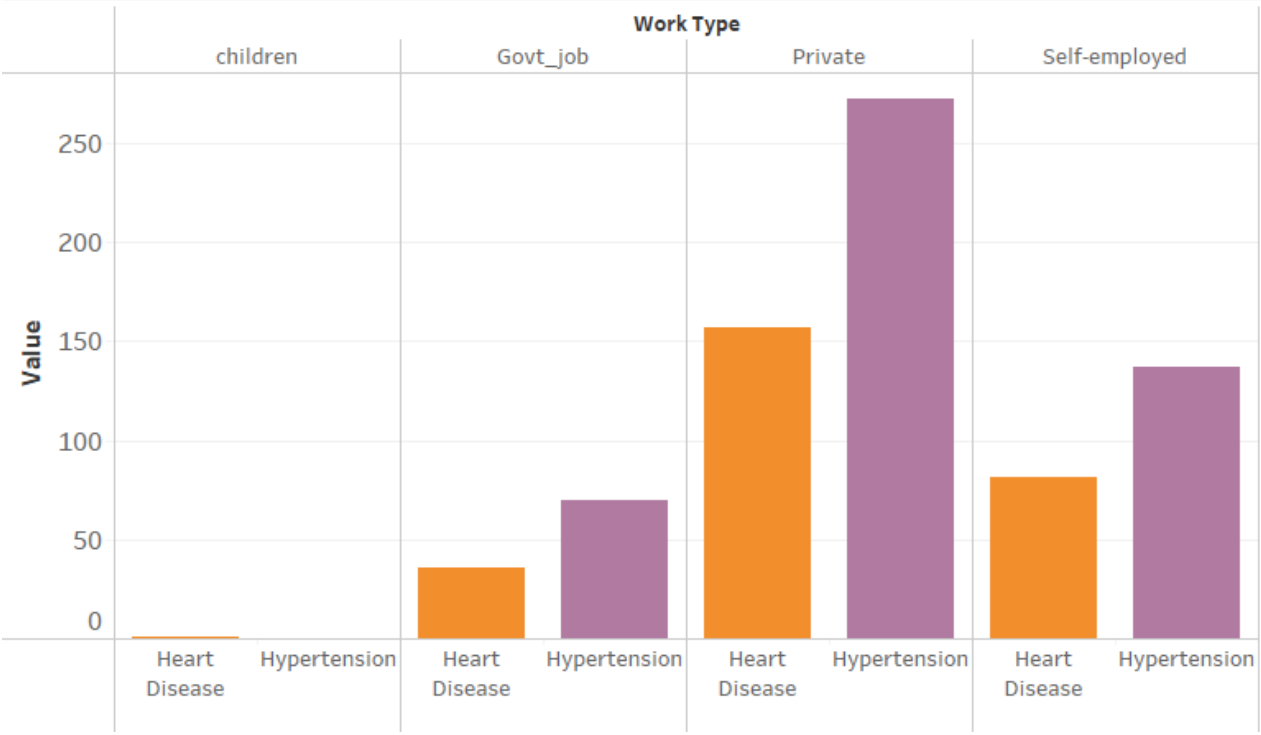
our mean age of stroke occurrence stands at 43.4, it implies that the prevalence of stroke has been at an increasing trend and has started greatly affecting people, regardless of their age groups. This could be due to the changes in our lifestyles which has depreciated our body strength and made us more susceptible to ailments.

- We see that the residence type- whether the individuals are residents of rural or urban areas plays no significant role on the average glucose level, BMI, heart disease or hypertension. We see a trend of these variables being slightly higher for the patients residing in urban areas which could occur due to differences in the lifestyles, but these variances in values are not large enough for us to make causal interpretations.

Residence type	Avg Glucose Level	Bmi	Heart Disease	Hypertension
Urban	268,400	72,380	141	240
Rural	259,305	69,569	134	239

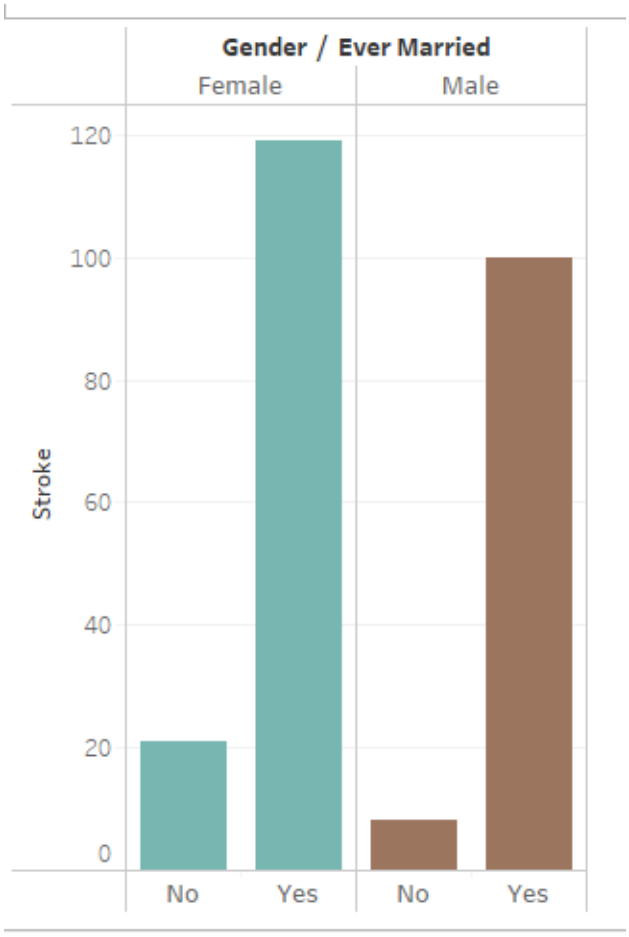
- For classification under work type- Children, Government job, Private and Self-employed, we see that the average glucose, as well as the body max index, is highest for private jobs, followed by those patients who are self-employed. This I believe could occur due to various factors. According to National Centre for Biotechnical Information, Medical News Today, and various other verified sources, higher levels of stress and linked to obesity which is the increase in average body mass index. Stress is also a cause of hypertension as mentioned by the Cleveland Clinic, Healthline publications, etc. The differences could also occur due to environmental factors like worse air quality in the urban sectors as compared to the rural areas.


Work Type
children
Govt_job
Private
Self-employed
300K
250K
200K
150K
100K
50K
0K
Value
Avg Glucose Level
Bmi

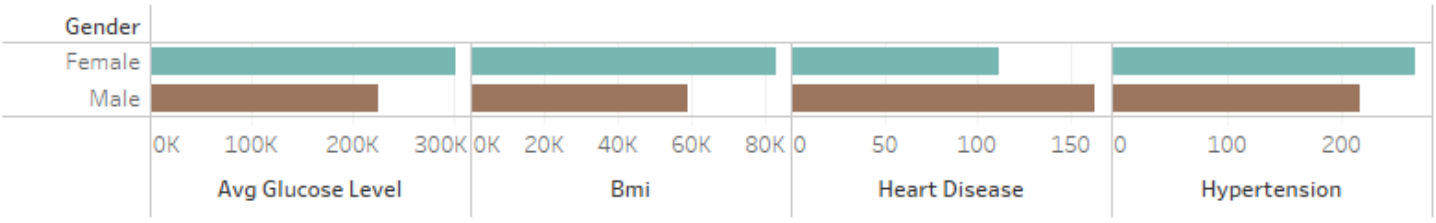

Work Type
children
Govt_job
Private
Self-employed
250
200
150
100
50
0
Value
Heart Disease
Hypertension

Analysis:

Now, coming to the analysis of the determinants of stroke, we see that gender and marital status play a key element role. Females are much more likely to be a candidate for stroke compared to males. Marriage further increases this likeness for both genders. Thus, a married female is much more probable (around 20%) of suffering from a stroke sometime in her life as compared to married males and married females are approximately 1100% points more likely to suffer from a stroke when compared with unmarried males.



The estimated reason for such an occurrence is as follows-



We notice that apart from the variable for heart diseases, the female counterparts suffer from higher levels of average glucose level, body mass index, and hypertension which we know are determinants of a stroke

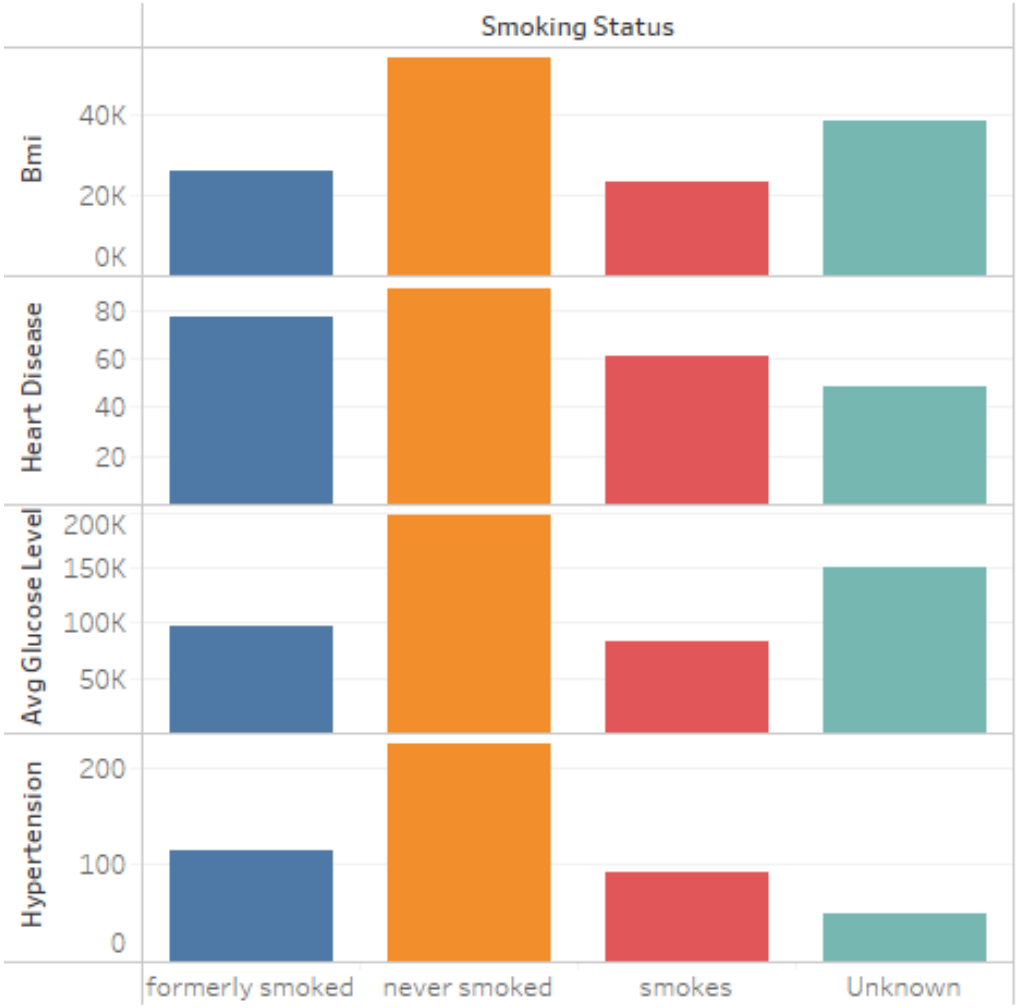
as mentioned in various previous literature by the American Diabetes Association, Mayo Clinic, etc. Ochsner Lafayette General even mentions that "Each unit increase in body mass index (BMI) increases the risk of stroke by 5 percent. With a baseline BMI of 20, this means that a person with a BMI of 30, for example, increases their risk of stroke by 50 percent".

Thus, this gives us enough evidence to approximate female genders, due to their greater incidence of higher average glucose levels, BMI, and hypertension, are more likely to suffer from a stroke at least once in their lifetime than compared to males and this statistic increases with marriage due to greater responsibilities of the family household and levels of stress.

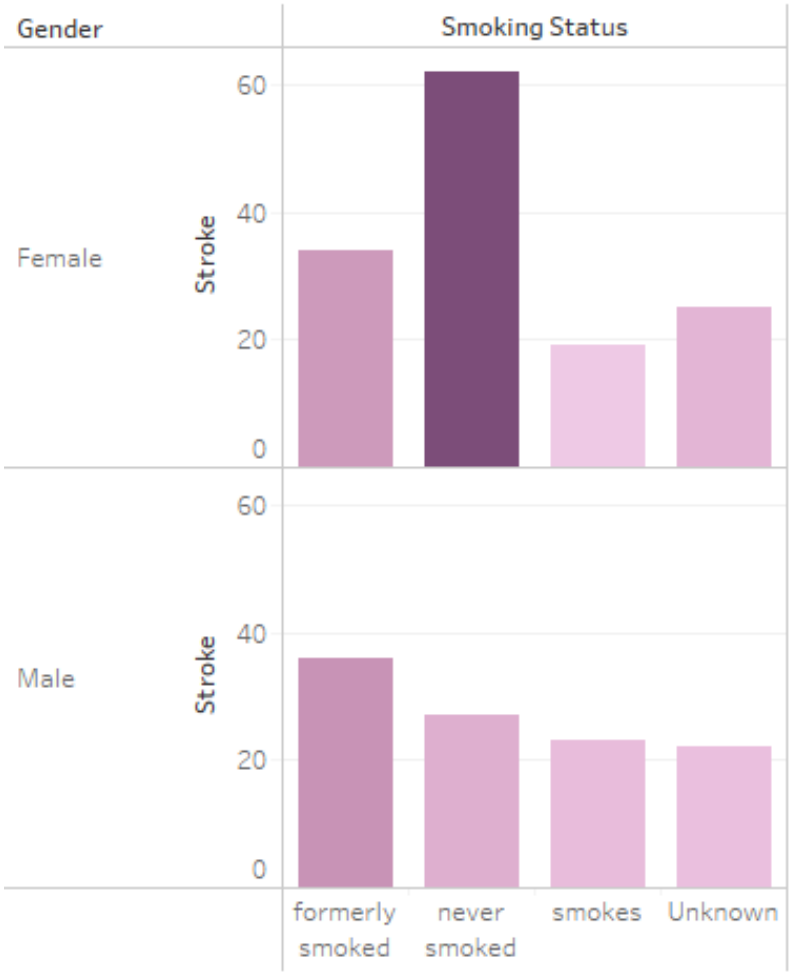
Now, we will analyze the effects of smoking on the variables- average glucose level, BMI, hypertension, and heart diseases and its effect on our outcome of interest strokes.

From the many advertisements, flyers, and even the display on cigarette own packets we have encountered, the message stating 'Smoking is injurious to health' is bolded and even highlighted before its own other taglines and thus we tend to believe this statement without any hesitation. Nearly every organ in the body sustains damage from smoking, which further increases the risk of disease and also lowers smokers' overall health. In the United States, smoking-related deaths account for more than 480,000 deaths annually, or roughly one out of every five, according to the Centres for Disease Control and Prevention. Smokers are more likely to develop cardiovascular disorders, which affect the heart and blood arteries. Smoking harms blood arteries and may cause them to thicken and become more constrictive. This causes your heart to beat more quickly, raising your blood pressure and promoting the development of clots that cause strokes.

Therefore, it should be clear at least from these figures that smoking increases the risk of heart disease and high blood pressure or hypertension. However, we observe a pattern in the data that all the variables—BMI, heart disease, hypertension, and average blood glucose level—are consistently highest for non-smokers, followed by patients who have smoked in the past, but not for the patients who are present smokers. This discrepancy in the results is striking, since it's vital to note that the data points on this dataset are fairly extensive, wherein data has been gathered over many years and these variables are consistently greatest for the non-smoker patients across our four variables.



When the difference is measured between the sexes, we notice this discrepancy clearly. In the case of men, the differences in the data values are not particularly significant wherein the male patients who were previous smokers are at the highest chances of strokes, but in the case of women, we observe that the patients who have never smoked have the highest risk of stroke than come female patients who have smoked in the past. I'm underlining this topic strongly since women who have smoked in the past have a doubled risk of stroke compared to women who have never smoked thus this value difference may not have just arisen out of coincidence, and it may have some unobservable factors leading to such a shocking result. The figures under the observations for men make sense both intuitively and scientifically as smoking harms the body not instantly but it has a recurring deterioration process that occurs over time. Therefore, it makes logical sense for the stroke likelihood values to be greatest for former smokers as the bodies of the patients who used to smoke before are now suffering from the consequences of the effects of smoking.



Conclusion:

In this manner, we demonstrate that smoking doesn't causally affect the frequency of strokes. Rather, it has a backhanded spurious correlation with strokes, wherein smoking affects and damages different body functions and capabilities which clubbed with ecological and social elements, lead to the decay of the circulatory system of the human body, prompting strokes. As noted by the World Stroke Organisation, "Smoking tobacco increases your risk of having a stroke. Someone who smokes 20 cigarettes a day is six times more likely to have a stroke compared to a non-smoker". Here as well, the Organisation is talking in terms of probability and likelihood as researching past literature, I have found that smoking does not have a direct causal impact on the incidence of smoking. Essentially, the elements like a higher average glucose level, high BMI, rate of hypertension, and the presence of heart diseases, make the body susceptible to brain strokes however once more, don't causally affect the occurrence of strokes. For instance, one can't claim that having a higher body mass index prompts strokes, it might act as a contributor to strokes but doesn't lead to strokes. The factors are all interrelated as having a higher body mass index leads to higher blood pressure, which in turn may put pressure on the blood vessels, thickening them and leading to heart diseases. These heart diseases may disable the smooth workings of the cardiovascular system, increasing the likelihood of having a brain stroke.