**Setting goals is an important step in any data analysis project, as it helps to focus your analysis and ensure that you are using your time and resources effectively.**

**Goal of the analysis:** Identify risk factors for stroke: Another goal could be to identify risk factors for stroke in the dataset. For example, you could use logistic regression analysis to model the relationship between different variables and stroke and identify the variables that are most strongly associated with stroke.

**Descriptive questions:**

1. What are the key characteristics of stroke data, such as demographic information, medical history and lifestyle factors associated with stroke risk?
2. What are the risk factors for stroke?
3. What is the average BMI among the sample?
4. What is the percentage of male and female among sample?
5. What is the frequency of gender across smoking status?
6. What is the max average glucose level?
7. What is the average BMI and glucose level among individuals who have had a stroke?
8. What is the percentage of individuals with hypertension and heart disease in the dataset with stroke?
9. How many individuals in the dataset have had a stroke?
10. What is the distribution of age, gender, BMI, and other demographic and health factors among individuals who have had a stroke?

**Exploratory question:**

1. Is there is a pattern between age and BMI?
2. Is there a relationship between BMI and average glucose level.
3. Is there a relationship between smoking status and gender.
4. Are there any patterns or relationships in the quantitative data that can provide insights into the risk factors or early indicators of stroke?
5. Can we identify potential Relationship between certain variables and stroke occurrence?

**Predictive question:**

1. Predict stroke based on the variables included in the data:
2. Can we identify the most influential risk factors for stroke prediction?

**Inferential question:**

1. What is the accuracy of logistic regression algorithms in predicting stroke risk?
2. Is there is a difference of the average BMI between gender groups?
3. Is there a difference in BMI with different work type.
4. Can we identify potential associations between certain variables and stroke occurrence?

**Causal question:**

1. The existence between variables on the data and stroke risk?