Question 1: Briefly explain why data visualization is important. Support with an example from the activity?

Answer: Data visualization is important because at the end of the day it is translating machine output into human legible data. The part where it asks - “The summary of data above gives *a lot* of information but is it ***easy*** to remember?” outlines this by following said statement up with easily digestible ***visible*** graphs. Data output can be extensive, and visualizing it makes it legible for people.

Question 2: What do you conclude from the visualization of this data? Support it with atleast 3 findings relevant from given dataset. Also, include a screen shot of the chart that supports your answers. (Attach screenshots to support your answer)

Answer:

Question 3: What does a dataset (df in our case) consists of? What would be the issue in the chart when you replace 'y' with 'x'?
(Click on KernelRestart and run all, wait for a few moments before the output loads, then answer this question)

Answer: Different attributes that influence a response or decision in a given test? It yields an error of Type, it cant be something other than Mapping or Data Frame

Question 4: How many rows and columns the dataset has? What relationships can we deduce from this visual analysis?

Answer: 10 rows, 8 columns. The relationship between attributes, and the numbers they yield given the tests ran?

Question 5: Why do you think we focused on 'y' variable instead of any other variable?

Answer: In this case, it appears they want to find who said yes and what type of individual they are. Which implies that the ‘y’ variable is the dependent variable. In tests such as these, if not generally speaking, it is the Independent variable’s effect on the Dependent variable. X is all the attributes on a Y/N answer. We want to visualize what type of data says yes/no.