Individual Project Proposals

The first question I want to talk about is employment. Nowadays, employment has become an issue that everyone discusses. I want to research this topic because it is highly related to our lives, which can visualize the employment situation directly. In addition, the government and experts can carry out specific improvements for employment. Specifically, the variable I am interested in is Q7 of 2023_cross: Are you currently..., which asks for the employment situation in 2022. To visualize, I will use barplots to show the employment situation in 2022 because the findings are categorical variables with two or more categories, and we can see which category is most common or least common, or how many times more common is one category than another directly. To analyze the question, I want to use bootstrapping. The reason for it is that it can show bootstrap sampling distribution as a histogram and create a range of plausible values for the population parameter. I assume the shape of the distribution will be bimodal, which will concentrate on the categories employed full-time or part-time (including self-employed) and retired. I think the result of the assumption will be true, and the sample in the employed full-time or part-time (including self-employed) will be greater than in the retired. This result means a large proportion of people are working so that the overall employment situation is good.

According to surveys, aging is becoming more and more serious, so the second research question I want to focus on is the aging situation in total population. The variable I find is Q1 of 2023_cross: What is your age (in years) in 2023? To show the graph of the distribution, I prefer to use a histogram because it can show the skewness, number of modals and how concentrated the values are. Therefore, we can know whether the overall trend of the age distribution is skew to the old. To analyze this question, I will also use bootstrapping because I think I can bootstrap several times in different samples, show the sampling distribution, and compare the possibility of the sample in different ranranges. ssume that the bootstrapping sample distribution will be right skewness because I think aging indeed has become more and more series. I think the result of the assumption will be true because if we follow the aging standard of 60 years old, there will be a large proportion of the sample staying in that range. Back to the research question, the aging situation indeed is more serious, and the government may suggest policies like family planning and delayed retirement to mitigate the situation.

The last research question is about weight and age. Some people say the weight increases as we get older. I think it's an interesting statement, and I want to know if it's a rumor or fact. I will use variable Q1 in 2023_cross: What is your age (in years), and Q186 in 2023_cross: What is your weight in 2023? To visualize, I will use the scatter plot because a scatter plot is useful to show the relationship between two continuous variables. We can use the x-axis to represent the age and the y-axis to represent the weight. In addition, a positive slope means they are positively correlated, and random scatters mean there is no strong relationship. To analyze, I will use simple linear regression because simple linear regression can show the relationship between two variables. I will calculate the correlation first to check whether age and weight are linear. If it is linear, I will calculate the true modal and the estimate modal to check their relationship. I assume the relationship will be positively correlated. And the result of it, I think, will be true. The reason for it is that as people grow up, the laziness will increase and the digestion ability will also degenerate. In the sample, the distribution of the scatter plot also shows a positive linear relationship of age and weight. For the research question, if the result is the same as we assume, the statement of weight increase as we get older will be a fact. People need to control their weight as they grow up to decrease the incidence of some diseases, like running and fitness.

Last but not least, the group members I want to cooperate in the future are Hongzhe Liu and Xuyang Tuli.