Research Questions

Tofu-FC - Huiwen Wang, Rocky Zhang, Darrick Zhang

Research Question 1

We are interested in investigating which factors end up impacting earnings after graduating from college, such as major choice or gender. We are also interested in seeing if there's any effect from income after graduating and unemployment rate on the choice of college majors. The target population for this research study should be those currently enrolled in, or considering enrolling in, college.

Our motivation for researching this question is to better inform currently enrolled undergraduates or those considering pursuing a higher education. There has always been a sense of uncertainty as to whether a higher education is worth it. As the cost of a higher education increases, from tuition to living costs, more families are concerned about college actually being a positive return on investment. To that extent, some majors have higher median incomes than others, which can be useful information to those trying to not be entrapped in debt post-graduation.

The type of data we might use for such a research question may come from universities themselves or the government, which may publish data about the earnings of their undergraduates after graduating, as well as the number of graduates and composition of graduating class in each major. There may be other data sources online which also suit our needs; the main variables we probably want included are median income after graduating, number of undergraduates enrolled in each major, unemployment rates for each major, gender statistics for each major.

Research Question 2

What environmental factors and physical features of mushrooms can human use to identify toxic/poisonous mushrooms in the wild?

The target population of interest is mushrooms in the wild.

Mushrooms are vital to the general wellness of the ecosystem, decomposing and recycling the nutrients in the soil. Mushrooms also provides a valuable food source full of nutrients for human beings and other important organisms. With this in mind, we are interested to see if there are any environmental factors and physical features of mushrooms that can help curious human avoid poisonous mushrooms that may grow in their yard. Ultimately the study of mushrooms is essential for safety and well-being.

Some data that we need to answer the question could be cap diameter (quantitative), cap shape (qualitative), stem height and width (quantitative), type of stem root and color (qualitative), ring type (qualitative), habitat (qualitative), season (qualitative), and whether the mushroom is edible or poisonous (qualitative).

Data set can be found at: https://github.com/ghattab/secondarydata/blob/main/PrimaryData/primary_data_edited.csv

Includes primary and secondary data, n = 173 and n = 61,069 respectively. Secondary data are hypothetical mushrooms generated for binary classification from the primary data set.

Research Question 3

We want to know what are some potential factors for different kinds of diabetes and to what extent does each factor contributes to or affects different kinds of diabetes.

The target population for this research question includes adult patients diagnosed with different types of diabetes. This population should encompass a diverse range of ages, genders, ethnicities, and socioeconomic backgrounds to ensure comprehensive insights.

Diabetes is a chronic condition that affects millions of people worldwide, leading to significant health complications and impacting quality of life. Understanding how different factors influence diabetes is critical for developing effective interventions and guidelines. This research could provide valuable insights for healthcare providers, patients, and policymakers to improve diabetes care and outcomes. By identifying key factors that positively or negatively affect diabetes, tailored recommendations can be made to help patients manage their condition more effectively, potentially reducing the incidence of complications and improving overall health.

We can have a dataset with a list of patients with different type of diabetes and important metrics about their general health conditions such as age, BMI, etc. as well as some potential factors related to diabetes, such as diets, genetics, etc.