**Social Economic Factors and Health Outcomes**

Amber Greer and Ashlyn Mcconaha

University of Arizona

APCV 361: Data Analysis and Visualization

Professor Li Xu

October 29, 2023

**Social Economic Factors and Health Outcomes**

**Introduction and Project Objectives**

For our project, we decided on the research question, “How do socioeconomic factors such as income, education, and employment affect health outcomes such as premature death rates and overall health quality in counties?” We arrived at this after reading the article, “Formulation of Research Question – Stepwise Approach” by Ratan, Anand, and Ratan. We believe it details the problem statement, describes the issue, and guides data collection and analysis. It also meets most of the criteria of the FINERMAPS recommendations by the authors, because it is feasible, interesting, ethical, relevant, manageable, appropriate, and systematic.

 In order to hone in on what was feasible, we had to review the data that was available in the Excel document “2023 County Health Rankings National Data” and try to determine an output that would be interesting. We then had to decide on the social economic factors we wanted to consider. We determined that income, education, and employment status were a good start in our research. By evaluating these variables in conjunction with quality-of-life variables such as poor or fair health, poor physical health days, poor mental health days, and low birthweight and/or length of life variables such as premature death (years of potential life lost before age 75) it would paint a robust picture of any correlations in those areas. We plan to try out a few methods, as the social economic factors and health outcomes that we have chosen will give us a variety of different results. We will do this across 3 different geographies to also view any variances there as well.

**Project Design and Development**

As we continue to explore the data over the next week, we will determine if we need to handle any missing values or if we need to implement any imputation strategies (mean, median, or mode). We will also check for any outliers using box plots. Since we want to do some comparative analysis, we may need to do some normalization or standardization. After the data is “clean,” we will look to uncover trends, patterns, and relationships in the data using descriptive statistics such as mean, median, mode, range, variance and visualize them using histograms, box plots, and bar charts, depending on which seems most appropriate.

For correlation and regression analysis, neither of us feel as strong in this area, so we have not decided which methods we want to use, but right now are planning to look at the last few assignments and see what code we can reuse to try and represent the correlations we are looking for and the strength of the correlations while also ensuring we aren’t inferring causation or inflicting any bias. Our immediate focus is on understanding the data we have thoroughly and getting it in a “ready” state and we will build on that over the coming weeks.

**Project Timeline**

Having a detailed timeline is critical to our project, it will guide us moving forward in our research and development as well as helping us meet deadlines. We plan to complete the project in six weeks following the submission of this proposal. In order to meet this deadline, we will split our tasks into a realistic timeline to be achievable weekly and designate individual roles that will be responsible for their section of the project. Being a team of two, we have decided to split the project in half and work together on sections we find the most difficult. We communicate regularly to ensure that we are staying on task with our weekly objectives, and we have established a calendar to keep up with due dates. On a weekly basis, we plan to meet once, complete our project log and add a few details to the project notebook just to ensure that we keep all our research on track with no conflicting data. The weeks following the submission of this proposal we plan to focus on the research of our topic in greater detail, build upon our research question, establish code for the method chosen for our data and ultimately complete our final project.

A list of project tasks

Description automatically generated with medium confidence

References

“Explore Health Rankings: 2023 Measures.” County Health Rankings & Roadmaps, www.countyhealthrankings.org/explore-health-rankings/county-health-rankings-measures. Accessed 26 Oct. 2023.

“Explore Health Rankings: County Health Rankings Model.” County Health Rankings & Roadmaps, www.countyhealthrankings.org/explore-health-rankings/county-health-rankings-model. Accessed 26 Oct. 2023.

“Explore Health Rankings: Rankings Data & Documentation.” County Health Rankings & Roadmaps, www.countyhealthrankings.org/explore-health-rankings/rankings-data-documentation. Accessed 22 Oct. 2023.

Ratan SK, Anand T, Ratan J. “Formulation of Research Question - Stepwise Approach.” J Indian Assoc Pediatr Surg. 2019 Jan-Mar; 24(1):15-20. Accessed 24 Oct. 2023