Problem:

(a) Outline the problem, and plot/visualise the data.

(b) Make progress on the problem, by applying the techniques of linear regression to the problem at hand.

(c) Discuss any insights and observations.

Imagine that we're data scientists/ engineers employed by a consulting company. Our consulting firm has been hired by a nonprofit organization whose mission is to advocate for better health outcomes for low-income populations in the United States. We've been asked to examine whether low-income groups are at greater risk for being diagnosed and dying from cancer. If successful, our analysis will help the nonprofit with lobbying and fundraising.

Goals:

1. Demonstrate whether or not cancer incidence and mortality are correlated with socioeconomic status.
2. Provide both quantitative and visual evidence that the nonprofit can take and use to further their mission.

Objectives:

1. Gather, Clean and prepare data
2. Exploratory analysis
3. Statistical model
4. Visualizations