

- We will most likely be looking at period 3 data which drops the number of participants once again

Elevator pitch:

Our project is about finding the relationship and correlation between smoking habits and coronary heart disease. We will be using the Framingham Heart Study data (1956 and 1968) as our data set. Our analysis will be done across the study population but also across the following subgroups:

- Age - continuous variable
- Sex - binary
- Attained level of education - categorical (1 to 4)
- Total cholesterol - continuous
- HDLC - continuous
- LDLC - continuous

Age and sex will be to demonstrate trends across demographics.

Total Chol, HDLC, and LDLC will demonstrate if there's correlation between smoking habits and CHD when considering cholesterol levels. It will also show the different trends between, totchol, high and low density lipoproteins.

Analyzing the relationship between CHD and smoking habits across the different levels of education.

- Person X2 correlation test