Final project

STA 138 | A. Farris | Fall 2023

Due Thursday, December 14^{th} at 5:00pm in pdf format; submit your work via Gradescope.

Under 'Files/final-project' on Canvas, find Byssinosis.csv .

In 1973, a large cotton textile company in North Carolina participated in a study to investigate the prevalence of byssinosis, a form of pneumoconiosis to which workers exposed to cotton dust are subject. Data was collected on 5,419 workers, including:

- Type of work place [1 (most dusty), 2 (less dusty), 3 (least dusty)]
- Employment, years [< 10, 10–19, 20–]
- Smoking [Smoker, or not in last 5 years]
- Sex [Male, Female]
- Race [White, Other]
- Byssinosis [Yes, No]

Your task is to investigate relationships between this disease on the one hand and smoking status, sex, race, length of employment, smoking, and dustiness of workplace on the other. To do so, select, fit, and then interpret an appropriate statistical model. Can you conclude in particular that workplace dustiness contributes to the chance of byssinosis? You will be graded on succinctly using valid statistical tools to uncover meaningful associations in the data and convincingly communicating them.

To complete this project, you may work if you would like with a group of size up to five. Each group needs to submit only one copy of their report, but name each group member in the assignment, as well as on gradescope when you upload it[1].

[1] *This is important, please do make sure everyone's name is entered into gradescope when your work is submitted!*