

Sta 108 - Project 1

The CDI Data

- *Background.* The CDI data is described on pages 1349 and 1350 of the book.
- *Requirements:*
 1. You are encouraged to work in small groups with 2-3 people in each group. Each group will submit one report, with the names of the group members on the front page.
 2. The report should begin with a brief introduction (about half a page), followed by the parts listed below.
 3. The due date of the project is **Wednesday, April 28, 2021** by midnight on canvas.

- *Data:* Available at the course website, or on CD attached to the book.

- The project has five parts:

Part I: Fitting regression models. This part consists of Projects 1.43 and 1.44 of the book. Note that there are three parts (a, b, c) in each project.

Part II: Measuring linear associations. Project 2.62 in the book.

Part III. Inference about regression parameters. Project 2.63 in the book. Also carry out the analysis of variance (ANOVA) for each regression model and state the results of the F-tests. What do you conclude in each case?

Part IV: Regression diagnostics. Project 3.25 in the book.

Part V: Discussion. Interpret your results using nontechnical words, that is, those that could be understood by someone who has not taken a statistics course before. Are there any thing else that you would like to comment on? Any suggestions on how to improve the linear regression models?

Note: Please provide your computer codes, as well as screenshots (limit to one page for each Part) of how the codes are run, as an attachment to your project report.