

STAT 154, Spring 2021

Midterm - data analysis part

By putting your name and student ID below, you agree to abide by the University Code of Conduct.

In particular, you are allowed to use all class materials for this exam, as well as material on the web.

You are not allowed to specifically search for answers to the problems online or in printed material, get or request help from anyone, post questions on online forums and similar entities, reproduce data analysis you found on the web. Answers/code in this problem set need to be written by you and you only.

You are not allowed to post this assignment online or reproduce it electronically and further disseminate it without consent of instructor.

Name:

Student ID:

Problem 1 [30 points]

Analyze the concrete mixture data at <http://archive.ics.uci.edu/ml/datasets/Concrete+Slump+Test>

The details of the dataset are there, explaining what is the response (please use 28-day Compressive Strength (Mpa)) and what are the predictors.

Provide a small report (at most 3 pages) describing the techniques you've used (and why), what sort of results you got and a small conclusion in plain English about what you've learned about the corresponding problem. It is preferable to use a few (e.g 2-3) methods well rather than run 10 different things not so well. Add your code as an appendix. (The code does not count towards the 3-page quota, so don't worry if your code is more than 3 pages.)

You are free to choose the format: either notebook, or any other methods that allows you to mix text and data analysis. .pdf is preferred, but an html file is fine, too.