

BF550: Fall 2021

Project 2 (due by 11:59 pm on Monday, December 13)

Scope and Goals

The goal of the project is to practice data analysis in Python. To that end, each student should choose a paper and reproduce some of its figures. It is important to make sure that you have access to the data. The data could have been deposited in a public repository or available from the authors upon request. Papers, which are part of the [Human Microbiome Project](#) could be a suitable choice. Unless you have interest in processing raw data, it is often better to start with processed data. Processing data could however be the main aspect of the project especially if it is done using your own code. Because the amount of work involved in reproducing different results varies greatly, you should consult the instructor ideally during office hours and discussion sessions about what you would like to do. As a rough guide, reproducing one figure that involves p-values and one figure that involves PCA or clustering is about the right amount of work when starting with processed microbiome data. Of course the nature of the plots could be very different especially in projects based on other types of data.

Write-up

The grading of this assignment will be based on your report. The report should include up to one page of introduction describing the motivation for the proposed study and the type of data used. Please do not spend too much time on this part as it contributes little to your score. The following section should describe how the data was analyzed and provide a description of how you structured your code. This methods section should be limited to one page. The last two pages should describe the results, show the figures, and make any conclusions that you reached while working on the project. The total length including figures and references is capped at four pages. The font size should be 11pt or greater; Arial and Times New Roman or similar fonts are recommended. The margins should be between 0.5 and 1 inch. The text should be single spaced.

Submission

Please submit it on the blackboard by 11:59 pm on Monday, December 13. Each student should submit their own report.