IB 120/201 - Lab 1

R & Python Basics

Due Date:

University of California, Berkeley *GSI: Ksenia Arzumanova*

In this lab, we will go through how to install and run python and R environments on your computers. We will also write our first program by printing a message to the command prompt in order to verify a successful installation.

Background

R & RStudio

R is a programming language that is well-suited for statistical analysis and working with large datasets. The best way to use R is in an integrated development environment (IDE) called Rstudio, where one could simultaneously view and control the code, command prompt, variables, directory, and figures.

Python

Python is an all-purpose programming language that is suited for a variety of tasks and functionalities (e.g. machine learning to applications with graphical user interfaces). It is known for its intuitive syntax, simplicity, and open-source support. There are many environments and methods

possible to run Python code, and we will first learn how to execute Python code by familiarizing oneself with the command line. A decent interactive development environment for Python is PyTorch, among many others. It provides an iPython interactive command line, meaning that one can execute code with Python syntax through the console, analogous to RStudio.

Jupyter Hub

Nowadays, most data scientists write python code in a notebook, meaning that there are chunks of code executed in a chronological order and output is presented right below the chunk of code. The advantage of notebooks is that debugging becomes much easier and it allows for easier dissection and understanding of one's workflow. Typically, these are called Jupyter notebooks because it was originally developed an organization called Project Jupyter and the code is run on a server that is either instantiated locally or located remotely. For an example, see this link.

Assignment

R & RStudio

1. Open a new window of RStudio on the hub
2. Create a script that writes `print('Hello World')` and save as `first_script.R`.
3. Use the `setwd()` command in Rstudio to navigate to the working directory where the file is located.
4. Run the code either by opening and running with the IDE shortcuts or by typing `source('first_script.R')` and run it. Make sure that the working directory of the command prompt is the same as where the file is located.

Python

1. Open a new text file, write `print('Hello World')`, and save as `first_script.py`.

2. Open a terminal window.

3. Navigate your command prompt to the directory containing the file and write `python3 first_script.py`.

Please submit a screenshot of printing "HelloWorld", or any message of your choosing, in the command prompt in both R and Python.