Discussion Section #0 : Introduction

Mitchell Valdes

Introduction

- **files needed** = None
- For the first week, we're covering some basics stuff and doing a little introduction.
- You can get the slides and Python notebooks fo all sections on, Canvas or the Github repository

About Me

- Originally from **La Habana, Cuba**
- **Education:**
  - B.A. in Mathematics from the University of Habana
  - M.A. in Economics from CIDE (Mexico)
  - 4th year Ph.D. student in Economics at UW-Madison
- **Research Interests:**
  - Macroeconomics
  - Labor Economics
  - Computational Economics

How to get in touch

- **Email:** valdsbobes@wisc.edu
  - Expect a response within 24 hours (usually much faster).
  - If you don't get a response within 24 hours, send me another email.

- **Office Hours:**

  - **Thursday** 3:00 - 4:00 pm
  - My office is 7308 Social Science

Discussion Structure

- Material should be covered in ~**30 minutes** with the remaining time for questions.
  - **Questions** are encouraged throughout the presentation.
- Attendance is **not** mandatory, but **highly** recommended, since we will be covering material that is not in the lectures.

Difficulty

- The first part of the course is can be both very simple and very difficult, depending on your background.

Difficulty

- If you have never programmed before, it will be difficult, but you will learn a lot.
- It is like learning a new language.
- actually, it is learning a new language.
- actually, it is learning a new language that is very different from any other language you have learned.
- … and a different way of thinking.
- after all you are learning how to talk to a computer.

Difficulty

- If you have programmed before, it will be easy, but you will still learn a lot (I hope).
- Challenge yourself to learn new things.
- How can I do this, but better?
- Different?
- Faster?
- More elegantly?

Difficulty

- **The most important thing is to not get discouraged.**
  - Also, study groups are encouraged.
  - Don't fall behind, it's very hard to catch up.
  - **Go to office hours!.**

. . .

- Programing is like a superpower, once you learn it, you will be able to do things that you never thought possible.

Finding good data

- What makes a "good" dataset? Why do we lean towards certain sources over others?
- Use this link from the course page as a starting point this semester when trying to find data.

Usefull exercise

Pick two datasets and compare them (even if they're not about the same topic). Some questions worth considering:

- Is the data easily digestible?
- Is the formatting easy to "process"?
  - Somethimes the hardest part of data work is data cleaning.

Suggest some new sources!

Feel free to discuss some less-known sources of data that you may be familiar with. Some examples:

- Sports-Reference (advanced features paid) for individual and team stats on numerous American sports
- UCI's Machine Learning Repository for machine learning datasets
- Open data portals for cities, states, and countries, such as data.gov, data.cityofchicago.org, and data.gov.uk
- This *awseome* repository of public datasets on GitHub.

Running Python

- Two ways to run Python:
  - Locally on your computer.
    - *If* you are intent on running Python on your own laptop, make sure you use Python 3.9.7, preferably with Anaconda 3
    - If interested reach out to me and I can help you set it up.
  - WinStat/SSCC, is the recommended way.

WinStat/SSCC:

- WinStat is a server (a remote computer) that you can access from any computer with an internet connection.
- First got to this link and log in with your NetID and password.
- Intructions how to install WinStat can be found here.

Runing Python on WinStat

- There are two ways to run Python on WinStat:
  - **Python scripts**
    - Basically a text file with Python code in it.
    - Python interprets the code in the file and runs it.
    - The file extension is `.py`.
  - **Jupyter Notebooks**, which is what we will be using in this course.
    - Has an interface that is similar to a web browser.
    - Can combine text, code, and output in one file.
    - The file extension is `.ipynb`.

Runing Juptyer Notebooks on WinStat

- To run **Jupyter Notebooks** on I'll prefer to use **VS Code**.
  - VS Code is a text editor that has a lot of nice features for programming.
  - It is preinstalled on WinStat. If you want to install it on your own computer, you can download it here.
- Alternatively there are two browser based tools that you can use to run Jupyter Notebooks:
  - **Jupyter Lab** (recommended)
  - **Jupyter Notebook** (not recommended)

Next I'll do a quick demo on how to run Python on WinStat using VS Code and Jupyter Lab.

Plot Twist

. . .

- This presentation is actually a Jupyter Notebook.

. . .

```
x = 223
print(f"{x} to the power of 8 is equal to {x**8}")
```

```
223 to the power of 8 is equal to 61155976398913804 81
```