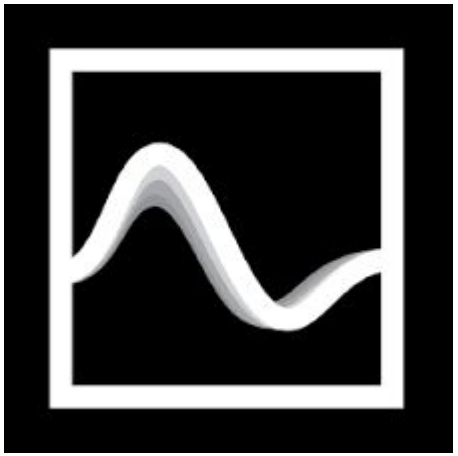
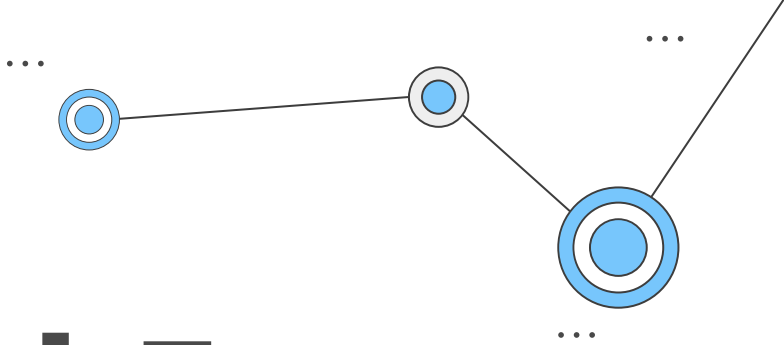


MAISI



CRA Week 7:

Subgroup Work Session #1 – Brainstorming!

Michigan Data Science Team

Fall 2025

Session 7 Agenda

01

Fun Icebreaker!!

Get to know your projectmates!

02

Final Expo Overview

What is our final deliverable going to look like?

03

Your Next Steps

Applying what you've learned independently of modules.

04

Subgroup Formation

Group up with others who have the same research ideas.

05

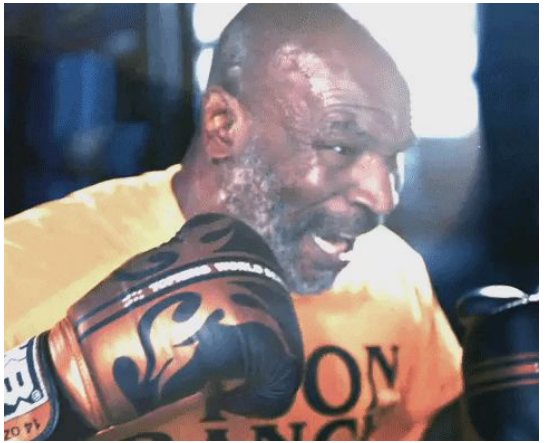
Brainstorming Time!

Get the ball rolling on your final presentation for the project.

Quick Icebreaker!!

Share with the people around you :)

If you were a boxer/professional athlete, what would your theme song be?



Final Expo Presentation Logistics

- Each semester, MDST hosts a Project Expo for all our teams and their members to show off what they have learned over the past two months!
  - This event hosted in two different sessions and is typically attended by other MDST members, DS enthusiasts on campus, professors, and sometimes even our corporate sponsors!
- General Logistical Notes
  - When: Friday, November 21st (likely around 6:00 PM)
  - Where: CCCB
  - What: Our final presentations!

Steps for Project Presentation

Hypothesis: Define the hypothesis or question you are aiming to answer with the COMPAS data. Feel free to use past workbooks as inspiration!

Research Questions: Clearly state some research questions that will guide your analysis. To create a strong research question include the *who, what, when, where, why, and how*

Tools and Techniques: Mention the data analysis tools your group will be using to solve the research questions. (pandas, confusion matrices, logistic regression, etc.)

pandas

NumPy

Sample Ideas

- T1: How does COMPAS score middle-aged defendants when compared to (otherwise similar) older/younger defendants?
- T2: Is there a disparity in how COMPAS scores women vs. men that is not reflected in actual recidivism rates?
- T3: Does COMPAS give White defendants with prior convictions high scores at the same rate as Black defendants w/priors?
- T4: To what extent do the data suggest that the COMPAS algorithm has an issue with racial profiling?

(If you want to use any of these as your topic, please type “T1”, “T2”, “T3”, or “T4” in the Google Sheets column asking for your choice. We encourage you to be creative!)

Getting into Subgroups (2–4 People)

- If you have already formed a group by filling out the Google Sheet, get together and start working on your presentation following the steps outlined in the previous slide.

- If you aren't quite sure of what you want to do or who you want to work with, talk with the people around you and find others who have a similar research interest to you!
  - Everyone needs to be in a group! (no solo presentations)

- **When you have your group, come talk to us!!**

Group Work Time!

Let's break into our teams and start researching topics to present on!

Hands-On Data Science!! :O

Next Steps:

1. Get in your groups and create a shared Google Drive for files (notebooks, slides, Docs, etc.)

2. **Fill out the Google Sheet!!** Link

3. Create a blank notebook and research a topic of your choice!

4. Ask us if you need any inspiration or help!

Pandas Cheat Sheet Seaborn Cheat Sheet

Reminders

- Don’t share colab notebooks with teammates if you are working at the same time

- Where to put csv and data files
  - Google Drive
    - Need to include:

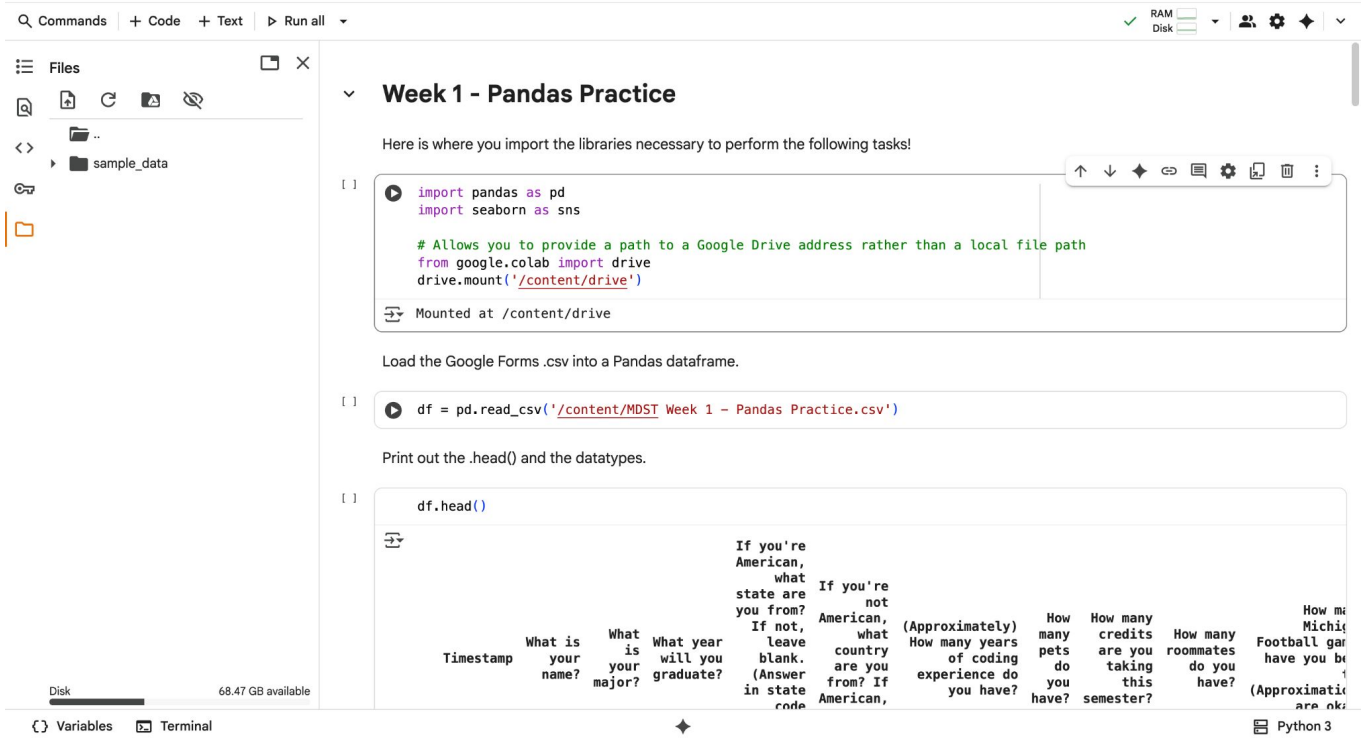
```
from google.colab import drive
drive.mount('/content/drive')
```

```
pd.read_csv('/content/drive/MyDrive/[FILE NAME]')
```

  - Colab Files
    - See next slides

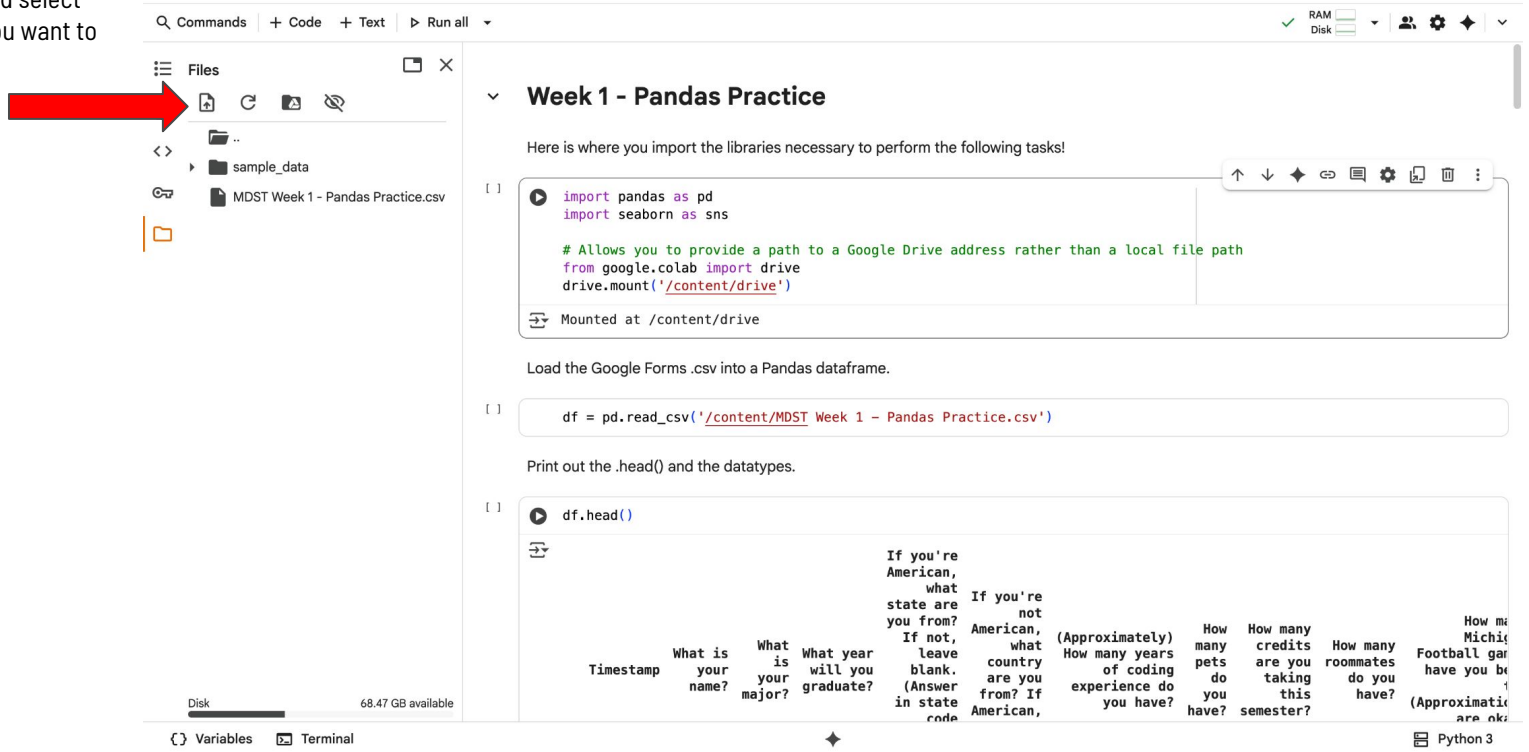
Reminders

Click on the folder in the sidebar



Reminders

Click the upload button and select the file you want to upload



Reminders

Click the three dots and copy the path. Put this in your read function