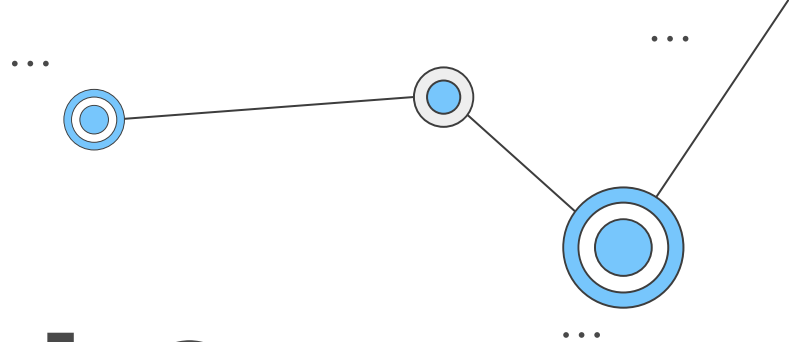




MAISI



CRA Week 8:

Subgroup Work Session

#2 – Finalizing!!! :0

Michigan Data Science Team
Fall 2025

Session 8 Agenda

01

Fun Icebreaker!!

Get to know your projectmates!

...

02

Final Expo Updates

What will the final expo look like?

...

03

Your Next Steps

Applying what you've learned independently of modules.

...

04

Finalize Presentations

Get as close as you can to wrapping up your final deliverable!

...



Quick Icebreaker!!

Share with the people around you :)

If you had to make a time capsule to commemorate the years of your childhood, what items would you put in there?



RECEIPTS

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, **6-7 PM**
 - Where: TBD
 - What: Our final presentations!
- \$200 final prize for the best presentation in MDST!!

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.)



Steps For Project Presentation

Findings: State what you found during your analysis of your research questions. Include model outputs, visualizations, or anything you think will be helpful!

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1 ^a	read	.098	.025	15.199	1	.000	1.103
	science	.066	.027	5.867	1	.015	1.068
	ses			6.690	2	.035	
	ses(1)	.058	.532	.012	1	.913	1.060
	ses(2)	-1.013	.444	5.212	1	.022	.363
	Constant	-9.561	1.662	33.112	1	.000	.000

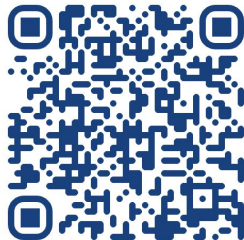
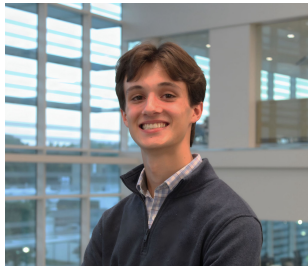
a. Variable(s) entered on step 1: read, science, ses.

Visualizations: Clearly label axes, make sure it is readable, and use colors effectively. Someone should be able to understand your plots easily and quickly.

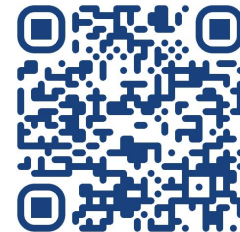


THANK YOU!!

- Thank you all so much for all of your effort this semester! It means a lot to us that you guys have showed up and engaged actively with the contents. We really hope that this project has been informative and exciting for you all, and we hope that you decide to stick with MDST and do another project either with us or our another awesome pair of leads (or even lead a project yourself!!) Feel free to connect with us on LinkedIn :)



...

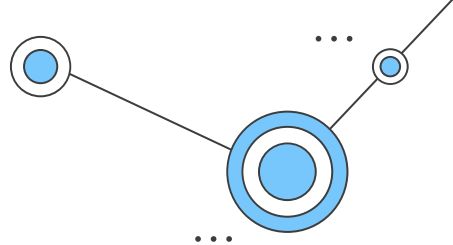




Group Work Time!

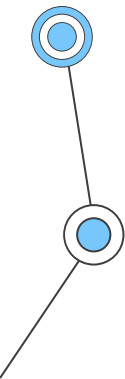
Time to start wrapping up our presentations to
prepare for the Final Expo on Friday!!

Hands-On Data Science!! :0



Next Steps:

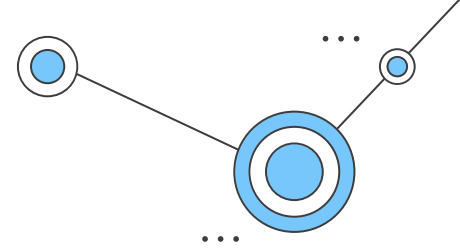
1. Get in your groups and get back to doing the good work you started last week!
2. **Fill out the Google Sheet!!** [Link](#) (we need to know how many groups we have to get enough tables at expo night)
3. Ask us if you need any inspiration or help, and show us when your slides are done!



[Pandas Cheat Sheet](#)

[Seaborn Cheat Sheet](#)

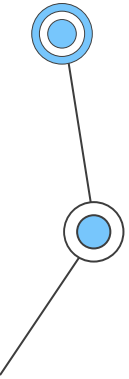
Reminders



- 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



Q Commands + Code + Text ▶ Run all

Files

- ..
- sample_data

Week 1 - Pandas Practice

Here is where you import the libraries necessary to perform the following tasks!

```
[ ] ▶ import pandas as pd
import seaborn as sns

# Allows you to provide a path to a Google Drive address rather than a local file path
from google.colab import drive
drive.mount('/content/drive')
```

Mounted at /content/drive

Load the Google Forms .csv into a Pandas dataframe.

```
[ ] ▶ df = pd.read_csv('/content/MDST Week 1 - Pandas Practice.csv')
```

Print out the .head() and the datatypes.

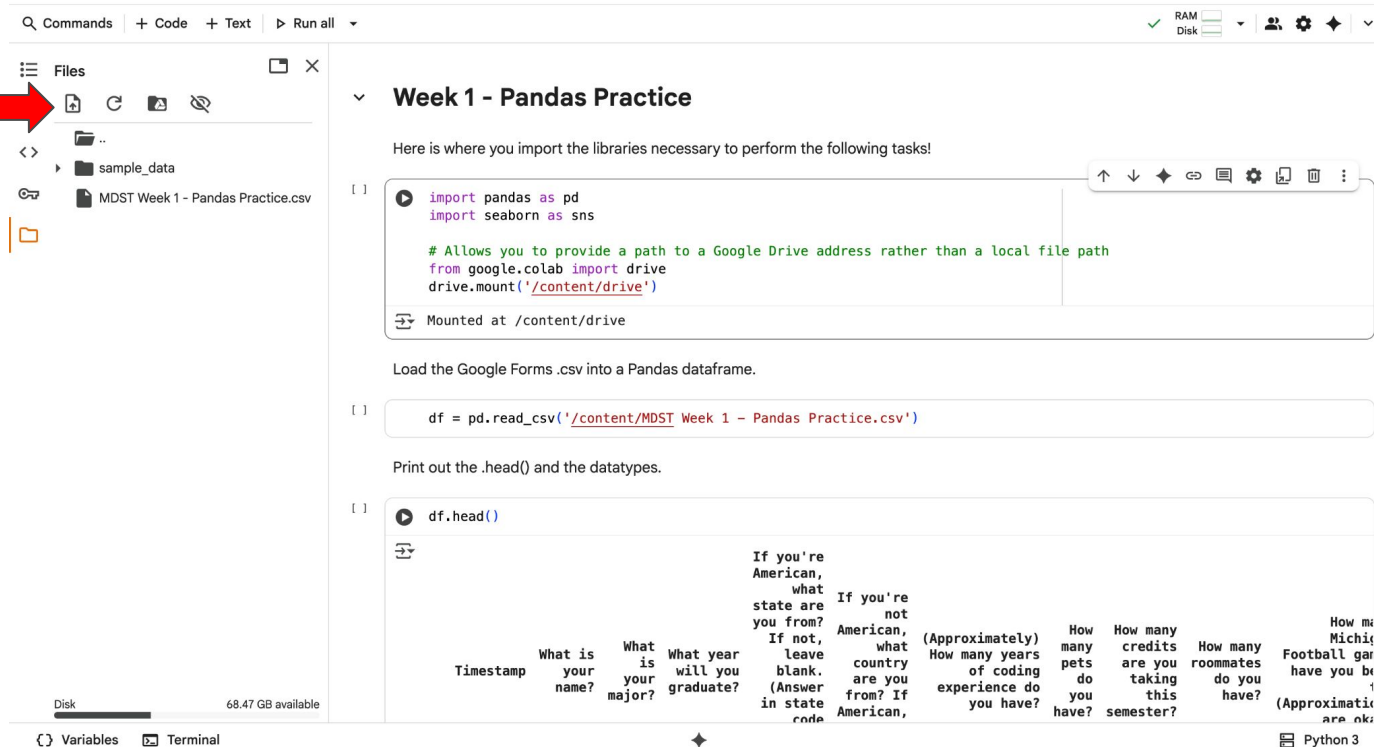

```
[ ] df.head()
```

Disk 68.47 GB available

Variables Terminal Python 3

Reminders

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



The screenshot shows the Google Colab interface. On the left, the 'Files' pane shows a folder named 'sample_data' and a file named 'MDST Week 1 - Pandas Practice.csv'. A red arrow points to the upload button (a square with a plus sign) in the 'Files' pane. The main area displays a Jupyter Notebook titled 'Week 1 - Pandas Practice'. The notebook contains the following code:

```
import pandas as pd
import seaborn as sns

# Allows you to provide a path to a Google Drive address rather than a local file path
from google.colab import drive
drive.mount('/content/drive')
```

Below the code, it says 'Mounted at /content/drive'. The next cell contains the following code:

```
df = pd.read_csv('/content/MDST Week 1 - Pandas Practice.csv')
```

Below the code, it says 'Load the Google Forms .csv into a Pandas dataframe.' The next cell contains the following code:

```
df.head()
```

Below the code, it shows the output of the `df.head()` command, which is a preview of the data from the CSV file. The output is a table with columns: Timestamp, What is your name?, What is your major?, What year will you graduate?, If you're American, what state are you from?, If you're not American, what country are you from?, (Approximately) How many years of coding experience do you have?, How many pets do you have?, How many credits are you taking this semester?, How many roommates do you have?, and How many Michigan Football games have you been to? (Approximately).

Reminders

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



The screenshot shows a Google Colab environment. On the left, the 'Files' pane displays a directory structure with 'sample_data' and 'MDST'. A context menu is open for the 'MDST' folder, with 'Copy path' highlighted. A red arrow points to this option. The main code editor area is titled 'Week 1 - Pandas Practice' and contains the following code:

```
[ ] import pandas as pd
import seaborn as sns

# Allows you to provide a path to a Google Drive address rather than a local file path
from google.colab import drive
drive.mount('/content/drive')

Mounted at /content/drive

Load the Google Forms .csv into a Pandas dataframe.

[ ] df = pd.read_csv('/content/MDST Week 1 - Pandas Practice.csv')

Print out the .head() and the datatypes.

[ ] df.head()
```

The output of the code shows the first few rows of the dataset:

Timestamp	What is your name?	What is your major?	What year will you graduate?	If not, leave blank. (Answer in state code)	If you're American, what state are you from?	If you're not American, what country are you from?	(Approximately) How many years of coding experience do you have?	How many pets do you have?	How many credits are you taking this semester?	How many roommates do you have?	How many Football games have you been to?	How many Michigan games have you been to?

The bottom of the interface shows 'Variables' and 'Terminal' tabs, and a 'Python 3' runtime indicator.