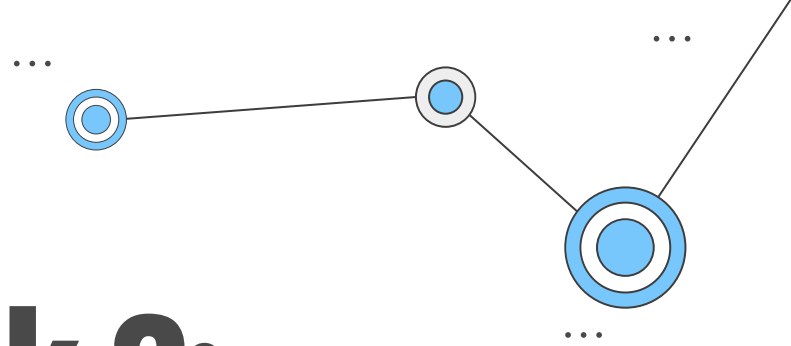


**MAISI**

# **CRA Week 2:**

## **Data Cleaning, EDA, and Intro to COMPAS**



Michigan Data Science Team  
Fall 2025

# Session 2 Agenda

01

## Fun Icebreaker!!

Get to know your projectmates  
(electric boogaloo)

...

03

## Intro to COMPAS

How was COMPAS used, and  
what are its flaws?

...

05

## Practice Time!

Work on the COMPAS dataset  
(with a fun warmup)!

...

02

## History of Risk Assessment

What are risk assessment  
algorithms?

...

04

## AI Safety Mini-Lesson

How does COMPAS relate to the  
use of biased AIs?

...

# Quick Icebreaker!!

What is your go-to karaoke song? (Even if you would never dare to sing it in public)



Share with the people around you!! :)



# What are risk assessment algorithms?

**Risk assessment algorithm:** algorithm used in the criminal justice system to predict the chance of an event: not showing up to trial, reoffending

- Replace human labor in making decisions about risk
- More importantly, replace the emotional, potentially biased nature of human judgement with an unbiased factual decision.
  - **“Hungry judge effect”**
- Be more “objective” in choosing who goes to jail
- Idea that statistics is always better than human judgment



# History of Risk Assessment Algorithms

- One of the earliest-known “risk assessment” algorithms used in courts was called the Burgess Method, developed by UChicago Professor Ernest Burgess.
  - Scored 21 factors using unit-weighted regression (0-1)
- 1990-2000: Computer boom, risk assessment algorithms = common practice
- COMPAS is one of the biggest ones



# What is COMPAS, and what is it used for?

- The COMPAS (Correctional Offender Management Profiling for Alternative Sanctions) algorithm is a risk assessment algorithm that predicts whether a criminal will recidivate in the next 1-3 years
- It's used in different stages of the criminal sentencing process. Broward County - Pretrial Release (bail)
- Created by Northpointe. (Now called Equivant)

equivant

# How Does COMPAS Work?



## Weighted, linear combination of questions

- "A hungry person has a right to steal"
- "Was one of your parents ever sent to jail or prison?"
- No questions directly mention race.
- Additional questions mention job and education.

...



## Proprietary, 'Black-Box' model

Weights on questions/factors are not publicly disclosed

...



## Numerical Scoring System

Violent/Non-Violent  
1-10  
Low/High Risk:  
Low/High Number

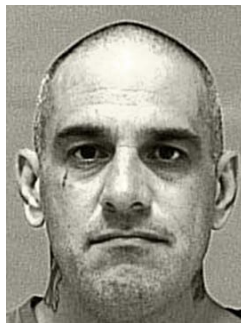
...

# Racial Bias in Algorithmic Risk Assessment

- **Context:** Brisha Borden (18, Black) and Vernon Prater (41, White) committed similar petty crimes.
- **Borden's Incident:** Stole a child's bike and scooter, valued at \$80; had a minor juvenile record.
- **Prater's Incident:** Shoplifted tools worth \$86.35; had a history of armed robbery with previous prison time.
- **Algorithmic Risk Assessment:**
  - Borden was rated as high risk for reoffending. (8)
  - Prater was rated as low risk for reoffending. (3)
- **Outcome:** The algorithm's prediction was incorrect:
  - **Borden:** No new charges after two years.
  - **Prater:** Convicted again, serving an 8-year sentence for burglary.



Risk:  
**HIGH**



Risk:  
**LOW**



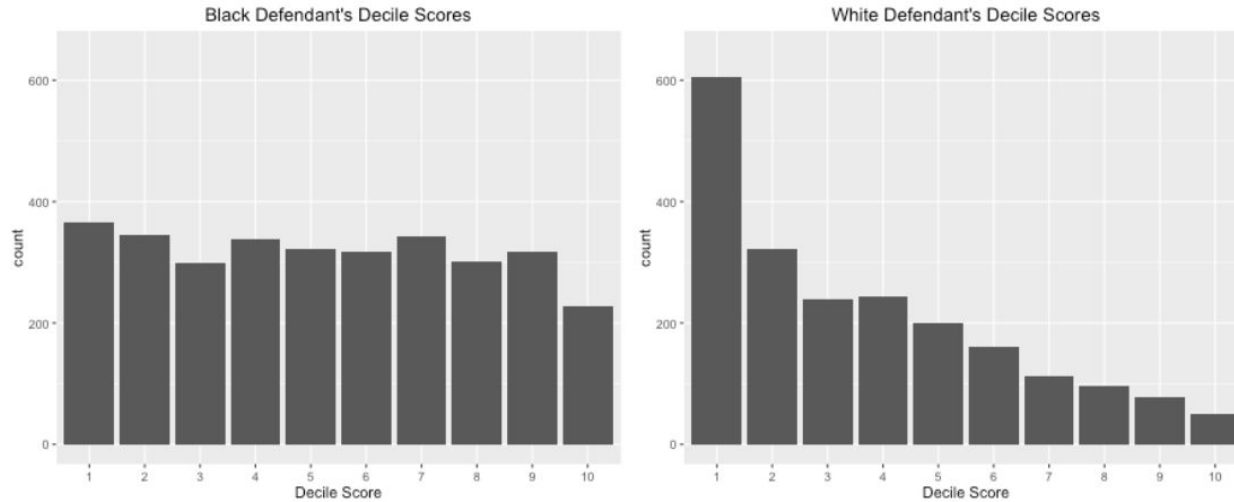
# Who are ProPublica, and why did they choose to investigate?

- Independent, nonprofit newsroom founded in 2007-2008.
- **Mission:** Expose abuses of power through investigative journalism. (Watchdog role in bureaucracy)
- COMPAS lacked independent studies on its **accuracy** and **fairness**
- Concerns about **bias** injected into the judicial process due to risk scores.
- Ensure fairness in the criminal justice system by scrutinizing widely used tools like COMPAS.



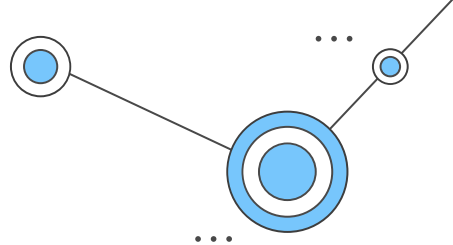
# What Did ProPublica Find?

Upon doing exploratory data analysis on a dataset of COMPAS risk scores...

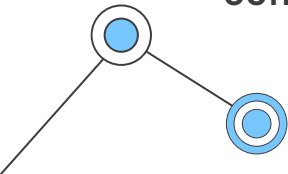
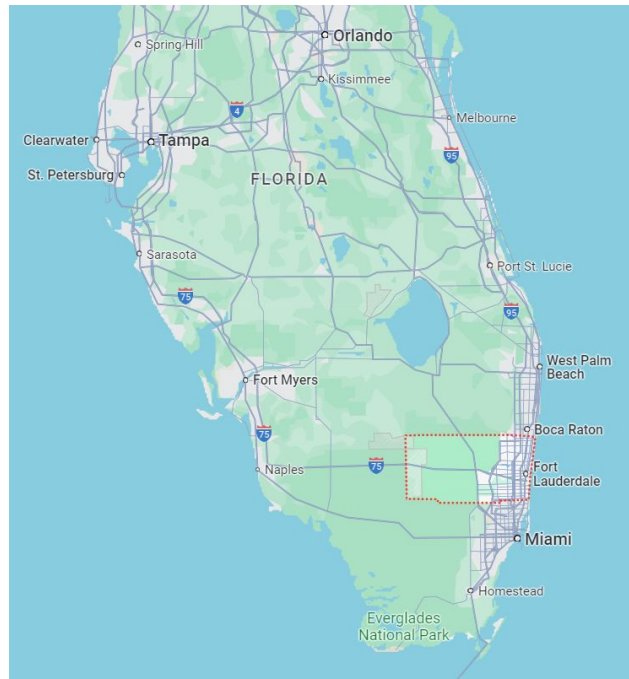


**Is this evidence of bias?**

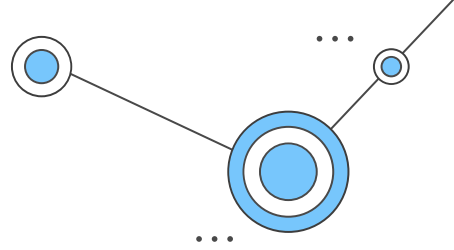
# What Did ProPublica Find?



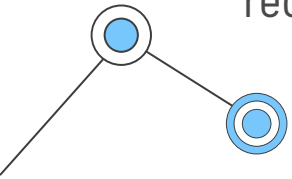
- These graphs alone are **not sufficient evidence** of bias
- Another way to assess the bias of the model: cross-reference recidivism risk scores with REAL recidivism
  - What we're doing!!
- Look at criminal histories of Broward County, FL residents and compare them with COMPAS scores



# What Did ProPublica Find?



- **False positives:** Black defendants who *did not recidivate* in two years were nearly twice as likely to be misclassified as higher risk compared to white defendants (45% vs 23%)
- **False negatives:** white defendants who reoffended in two years were nearly twice as likely to be misclassified as lower risk compared to black defendants (48% vs 28%)
- **Gender bias:** a high risk woman has a lower chance of recidivating than a male counterpart
- Even disregarding bias across races, COMPAS correctly predicted an offender's recidivism only **61 percent** of the time!



# What is Fairness, and How Does It Relate to AI Safety?

- Fairness is complicated because no algorithm can be perpetually accurate, especially if its outputs are associated with inputs with different traits.
- 3 different definitions of Fairness:
  - **Statistical Parity:** an algorithm makes positive decisions at an equal rate across all groups
  - **Equalized odds:** False positive & False negative rate are equal between all groups
  - **Calibration:** Prediction matches the real outcome at the same rate across groups

Which of these fairness definition does **COMPAS** fail?

Answer: **ALL OF THEM!!**

# Critical Issues

## Inaccuracy

- Only 20% of those predicted to commit violent crimes did so.
- When all crimes were considered, only 61% of those predicted to reoffend did so within two years.

## Racial Disparities

- Black defendants are nearly twice as likely to be falsely labeled as high risk compared to White defendants.
- White defendants are more often labeled as low risk.

# Our Partnership with MAISI

- Student org on campus focusing on AI ethics and current events (Will is a board member!)
- Bi-weekly meetings in East Quad B804
- Weekly reading groups on Alignment and Governance
- Feel free to stop in and say hi!!



M A I S I

...

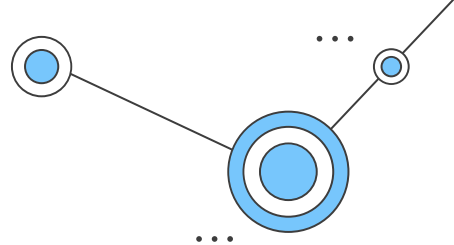


# Practice Time!

Time to apply our skills to the Titanic  
and COMPAS datasets!



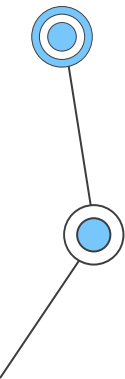
# Reminders



- 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.

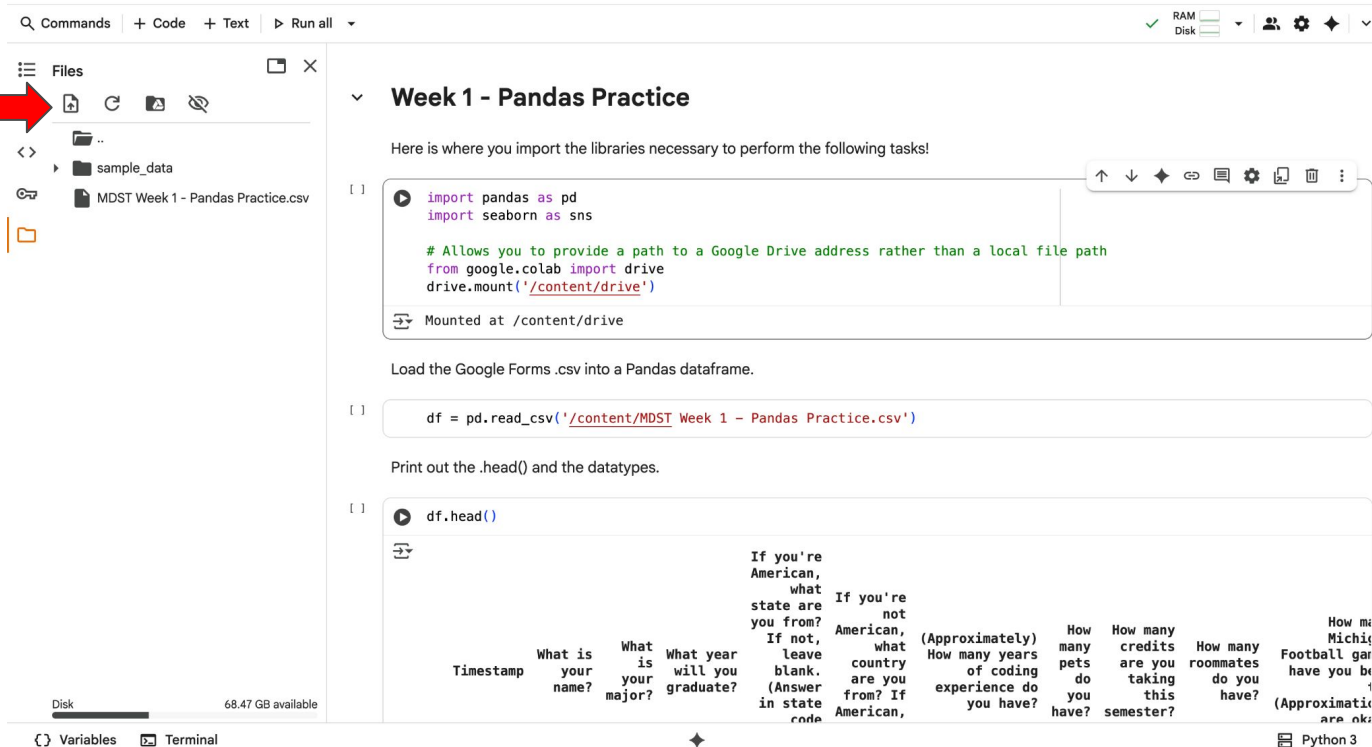

```
[ ] df.head()
```

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 file explorer. 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 from the text 'Click the three dots and copy the path. Put this in your read function' to this option. The main code editor shows 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')

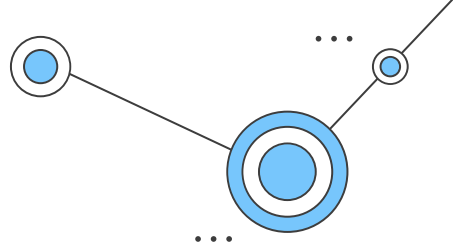
Print out the .head() and the datatypes.

[ ] df.head()
```

The output of `df.head()` is displayed as a table with 10 columns and 1 row of data. The columns are: 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?, and How many Michigan Football games have you been to? (Approximately are ok).

At the bottom of the interface, there are tabs for 'Variables', 'Terminal', and 'Python 3'.

# Hands-On Data Science!! :0



## Next Steps:

1. Find the F25 CRA repo in the [MDST GitHub](#) and Download all the files in the Week 2 directory
  - a. You can just Google "https://github.com/MichiganDataScienceTeam"
2. Split into teams of 2-3 (new ones or same as last week)
  - a. If you're working with new people, introduce yourselves!!
3. Work on the exercises in the notebooks!
  - a. Start off with the Titanic dataset and see how far you can get. We'll let you know when you should move on to the COMPAS module.

[Pandas Cheat Sheet](#)

[Seaborn Cheat Sheet](#)

