

# DS2001 - CS Practicum

Spring 2023

## Practicum 2 - Visualizing Data and Conditionals

Practicum is DUE at the end of your scheduled practicum.

We'll be working in pairs this week! With your partner, work on the programming assignment below. You may not finish every problem every week, but we expect your best effort.

We will let you know when there are 15 minutes left in class – at that point, it is time to wrap up your code and move on to the self-reflection portion of the assignment! Read the self-reflection questions below, write your answers, and save them in a PDF.

Submit your code and self-reflection to gradescope. You only need one submission per group.

### Grading Policy

You will receive full credit if you submit your self-reflection (with thoughtful answers) and any code by the deadline. All questions in the self-reflection must be answered for full credit.

### Feedback

We will provide feedback on your code if requested in your self-reflection. Otherwise, we will not look at your code. Our feedback will mirror the expectations of your DS200 homeworks, and we will use the same criteria for grading your projects in DS2001. So, make sure you ask for feedback when helpful, and read our notes!

## Today's Goals

- Visualize data from a file
- Use conditionals for computations and plotting

## Working in Pairs

Each pair will work at one computer, and everyone contributes. Here's how we split up the work:

- Navigator: Dictates the code to be written. Explains the *why* as we go. Checks for syntax errors.
- Driver: Writes the code. Listens closely to the navigator. Asks questions whenever there is a lack of clarity.
- Both driver and navigator are responsible for contributing to the work, and for making space for the other person to make contributions.

# Programming Assignment

## 1. Gather data

Download the datafile punxsutawney.txt (this data is from the source: <https://www.kaggle.com/datasets/groundhogclub/groundhog-day>). Save it in the same directory as the Python program you will write.

The data file contains data relating to Punxsutawney Phil. Legend has it if Punxsutawney Phil (a groundhog) comes out of his hole and sees his shadow on groundhog's day, then we will have 6 more weeks of winter in the US.

In punxsutawney.txt the data is as follows:

```
Year
Did Punxsutawney see his shadow?
Average US temperature in February
```

Under main, start your program by reading in the punxsutawney.txt data and saving each line in a variable.

## 2. Computations

How accurate is Punxsutawney Phil?

The average US temperature in spring is 52 degrees F. For each year in the data, use conditionals to determine if the average February temperature was greater than or equal to 52 degrees F. Save this information as variables.

Let's say Punxsutawney Phil's prediction is accurate if (a) he does not see his shadow and average US temperature in February is above or equal to 52 degrees F, or (b) he sees his shadow and average US temperature in February is less than 52 degrees. For each year in the data, use conditionals to determine if Punxsutawney Phil's prediction was accurate. Save this information as variables.

## 3. Communicate to the user

Illustrate Punxsutawney Phil's accuracy with a visualization. Create a scatterplot that shows year on the x-axis and average US temperature in February on the y-axis. Be sure to add axis labels!

Color each point according to whether Punxsutawney Phil's prediction was accurate. Be sure to add a legend to explain what your coloring represents!

## Self-Reflection

Create a PDF with answers to the following questions (each should be answered collaboratively with your partner unless instructed otherwise):

- Does this work reflect the best effort of both partners?
- In what way(s) did each partner (1) make contributions to the assignment, and (2) give the other person space to make contributions? (Each partner should respond to this question separately.)

Which problem(s) would you like feedback on and why?

## Submitting Your Work

Only one person per group needs to submit on gradescope, but that person needs to tag their partner to make sure everyone gets credit. Look back at labs 0 and 1 if you want step by step instructions.

## Keep Going (if you have time)!

- Add a line to your scatterplot to mark 52 degrees F, our threshold for spring weather.
- Decide if Punxutawney Phil is better than random guessing. For each year in the dataset, generate a random value. Check if that value corresponds to spring temperatures or not, and add it to the scatterplot. Are the random values more, the same, or less accurate than Punxutawney Phil?
- Can you come up with another interesting analysis question for this data? Brainstorm and then try to write code to answer your question.