

MODULE: a premade library of functions that you can use in your program

Interesting modules

- **Numpy / pandas / matplotlib:** data analysis
- **Tensorflow:** machine learning
- **Pytest:** testing your code
- **Howdoi:** asking questions about anything
- **Emoji:** adding emoji's to your code
- **Wikipedia:** import all of wikipedia for use in your program
- **Disassemble python:** shows you your code under the hood

PANDAS REVIEW: introductory concepts from class

Setting it up

pip install pandas
OR
conda install pandas
OR
[advanced install instructions](#)

Handy commands

- **import pandas as pd:** import the module
- **.sample(number):** preview your data where number = the desired amount of rows you want to see
- **.describe():** basic stats on your data
- **dropna():** drop any columns having null data
- **fillna(new_value):** bar graph
- **.loc[row_number]:** to select a specific row

PANDAS PRACTICE: in class

1. Download the csv located in this repository with the title "practice.csv"
2. Drop null values from your data set or replace them with empty strings
3. Write a program in the desired text editor/notebook thof your choice:
 - a. Sample a few rows of the dataset to explore what's there
 - b. Store the column "**common_name**" in its own variable
 - c. Pick a tree type and count how many trees there are of that type in Pittsburgh
 - d. What is the mean height and width of trees in Pittsburgh?
 - e. What neighborhood has the most trees?

PANDAS PRACTICE: on your own

1. Pandas allows you to concatenate two data sets in their entirety with **pd.concat**. This will horizontally stack each dataset on top of each other so similar datasets are required.
2. You can also merge two datasets and only combine elements of each that you want with **pd.merge**. This is similar to the concept [a SQL join](#).

PANDAS RESOURCES

- Pandas cheat sheet: https://pandas.pydata.org/Pandas_Cheat_Sheet.pdf
- Pandas users guide: https://pandas.pydata.org/docs/user_guide/index.html
- Mode Pandas tutorial: <https://mode.com/python-tutorial/pandas-dataframe/>