

# Module 7: Data Wrangling with Pandas

## CPE311 Computational Thinking with Python

Submitted by: LastName, FirstName

Performed on: MM/DD/YYYY

Submitted on: MM/DD/YYYY

Submitted to: Engr. Roman M. Richard

---

## 7.1 Supplementary Activity

Using the datasets provided, perform the following exercises:

### Exercise 1

We want to look at data for the Facebook, Apple, Amazon, Netflix, and Google (FAANG) stocks, but we were given each as a separate CSV file. Combine them into a single file and store the dataframe of the FAANG data as `faang` for the rest of the exercises:

1. Read each file in.
2. Add a column to each dataframe, called `ticker`, indicating the ticker symbol it is for (Apple's is `AAPL`, for example). This is how you look up a stock. Each file's name is also the ticker symbol, so be sure to capitalize it.
3. Append them together into a single dataframe.
4. Save the result in a CSV file called `faang.csv`.

In [ ]:

### Exercise 2

- With `faang`, use type conversion to change the date column into a datetime and the volume column into integers. Then, sort by date and ticker.
- Find the seven rows with the highest value for volume.
- Right now, the data is somewhere between long and wide format. Use `melt()` to make it completely long format. Hint: date and ticker are our ID variables (they uniquely identify each row). We need to melt the rest so that we don't have separate columns for open, high, low, close, and volume.

In [ ]:

## Exercise 3

- Using web scraping, search for the list of the hospitals, their address and contact information. Save the list in a new csv file, hospitals.csv.
- Using the generated hospitals.csv, convert the csv file into pandas dataframe. Prepare the data using the necessary preprocessing techniques.

In [ ]:

---

## 7.2 Conclusion:

Write your conclusion here.