MLDS 422 – Fall 2023
Project 2
Due Monday, 11/13/23 at 11:59pm

Exercise 1: Web Scraping + Coding Best Practices

From the MLDS website (https://www.mccormick.northwestern.edu/machine-learning-data-science/people/alumni/), use Python to scrape every alumni's **name**, **graduation year**, **current role** and **current company** from graduation years 2013 – 2022. There are multiple approaches to this problem – as long as you create a DataFrame with the four **bolded** columns, you will receive full credit.

Answer the following questions:

- What are the top 5 companies that alumni are working at?
- What are two other interesting insights you can gather from the data?

Out of 50 points, you will receive:

- +35 points for correctly creating the DataFrame and finding the top 5 companies
- +5 points for using loops and/or functions so you're not repeating code
- +5 points for using an approach that can apply to unseen cohorts (you aren't manually fixing each small problem, but rather have reuseable code)
- +5 points for coming up with two original insights

This roughly follows the coding philosophy – make it work, make it right, make it beautiful, make it fast.

Exercise 2: API

Write a Python program that gathers, summarizes and plots weather data.

1. Go to the World Weather API site (https://www.worldweatheronline.com/weather-api/) and get a free 30-day trial. Remember your API key.
2. Get one month of historical temperature data for 20 distinct locations.
3. Create a summary table that summarizes the min, mean and max temperature values during that month for each location.
4. Plot the mean temperature for each location, compare the temperature data and share your insights.

NOTE: Please start this assignment early. You are limited to 500 API calls per day.