Homework 2 - Exploratory Data Analysis and Visualization in Python

Download the **Exploratory Data Analysis.zip** file and unzip it in a folder to which you can navigate in Anaconda. I would recommend your Documents or Desktop directory if you're using Windows or Mac. I've included links on unzipping a file to a specific folder for Windows and Mac OS/X; ask in Piazza if you're having problems unpacking the assignment.

This assignment is structured as a Python Jupyter notebook. The notebook contains a series of four problems. Make sure that you read the Homework 2 Guide.pdf file for help on this assignment.

The problems correspond to:

1. Basic statistical calculations using NumPy vector
2. Basic statistical calculations using Pandas dataframes and series
3. Plotting data using matplotlib
4. Plotting data using Pandas

Part 1 assumes you've read chapter 16 of the MIPS book or the OpenIntro stats chapter. The other parts assume you've gone through the video lectures for the NumPy, matplotlib and Pandas tools.

Each problem in the assignment is followed a series of **assert** statements that allow you to test your solution. An **assert** statement evaluates a condition or expression; if that expression or condition evaluates to **False**, an error is generated. For example, executing

```
assert 1==2
```

would generate an **AssertError** when evaluated. When the homework is graded, these assertions (and others that are hidden from you) will be used to grade the assignment.

Some of the problems will be graded by hand. One question (1d) requires you to show an algebraic derivation; you should express this using the LaTeX equation forms covered in the earlier tutorial.

You should upload your solution as a single file, **Lastname_Exploratory Data Analysis.ipynb**, that contains the modifications for your solutions. You should only modify locations asking for your name and those identified as "**# your code here**"