

## Assignment 2: Data Exploration

In this assignment we will work with a (fake) dataset encompassing people's age, number of steps taken and income. The main purpose of this assignment relates to data wrangling / data cleaning - but we will also follow through on some basic analysis and predictions with this data.

### Tasks / Learning Goals

Hands-on experience with:

- Data Wrangling: Connecting together data from
- Data Visualization: Looking at data to check for problems
- Data Cleaning: Removing 'bad' data, dealing with outliers, dealing with missing data
- Basic Analysis & Prediction: Looking at how different variables relate to each other

### Due Date

11:59 pm on Sunday, February 4th, submitted on TritonED.

### Submitting Assignments

You will submit a Jupyter notebook file (.ipynb) to TritonED. Make sure that the file you submit has the following filename (filled in with unique course ID number):

'A2\_#####.ipynb'

### Grading Rubric

This assignment is worth 12% of your grade (12 points).

There are 8 parts to this assignment, with the following point values:

Part 1: Data Wrangling	2.5 points
Part 2: Data Cleaning	1.5 points
Part 3: Data Visualization	1 point
Part 4: Data Pre-Processing	1.5 points
Part 5: Outliers	1 point
Part 6: Basic Analysis	1.5 points
Part 7: Predictions	2 points
Part 8: Revisiting Outliers	1 point