### D4_casestudy_eda (Score: 2.0 / 2.0)

1. Test cell (Score: 0.5 / 0.5)
2. Test cell (Score: 0.5 / 0.5)
3. Test cell (Score: 0.5 / 0.5)
4. Test cell (Score: 0.5 / 0.5)
5. Written response (Score: 0.0 / 0.0)

# Case Study: Age in American Politics¶

In the last few section workbooks, we've reviewed Python basics, used a little `pandas`, wrangled and cleaned data, and generated a number of basic visualizations. This week, we'll extend this a bit futher, incorporating all that we've learned so far.

Your goal is to, without much structure provided, utilize the data provided to answer the following questions:

1. Does Congress have an age problem? (aka, is Congress much older than the population they represent?)
2. Is this problem exclusive to one of the two major parties?

Think about what your hypotheses are before delving into this analysis!

Code has been provided in Parts I and III - your job is do figure out Part II! And, don't worry if you don't get to everything.

The data you have to start with are available here: congress-terms.csv (https://github.com/fivethirtyeight/data/tree/master/congress-age). They were used in this piece (https://fivethirtyeight.com/features/both-republicans-and-democrats-have-an-age-problem/) at FiveThirtyEight. Note, there is an entry for every member of congress who has served at any point during a particular congress between January 1947 and Februrary 2014. One thing to keep in mind is the fact that elections have occurred since 2014 that are not included in this dataset. Getting up-to-date data will be explored in Part III of this notebook.

# Part I: Setup & Data¶

There's not much for you to do beyond read the instructions and run the code provided here (being sure to understand it). Don't worry though - there's plenty to do in Part II (and III).

In [1]:

```
# import pandas & numpy library
import pandas as pd
import numpy as np

# Import seaborn and apply its plotting styles
import seaborn as sns
sns.set(font_scale=2, style="white")
```