D6_casestudy_inference (Score: 2.0 / 2.0)

1. Test cell (Score: 0.5 / 0.5)
2. Test cell (Score: 0.75 / 0.75)
3. Test cell (Score: 0.75 / 0.75)

Case Study : Inference¶

Since this is the third case study we'll be doing, you have some idea of how things work now. The idea here is that we spent in class talking about inference and walking through some examples, and now it's your turn to do an inferential analysis!

Your goal is to, without much structure provided, consider what we've discussed in the course up to this point. Then, utilize the data provided to answer the question: *Do Pulitzers help newspapers keep readers?*, or more specifically *By looking at Pulitzer prizes awarded and changes in readership, can we determine what the effect of prestige is on the viewership at the 50 most popular newspapers between 2004 and 2013?*

The URL to the data for use in answering this question from FiveThirtyEight is https://raw.githubusercontent.com/fivethirtyeight/data/master/pulitzer/pulitzer-circulation-data.csv (https://raw.githubusercontent.com/fivethirtyeight/data/master/pulitzer/pulitzer-circulation-data.csv).

Part I : Setup & Data Wrangling¶

In this section you'll want to:

- import any packages you'll need for your analysis
    - a number are provided for you, but feel free to edit if there are other packages you want to use or changes to the input you want to make (Note: this is allowed in labs; in assignments you're limited to the imports we give you)
- read the dataset in
- understand what's going on in the data
    - how many observations?
    - what variables do you have? what variable types are here
    - which will you need to answer the question?

In [1]:

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

# import matplotlib
import matplotlib as mpl
import matplotlib.pyplot as plt
import matplotlib.style as style
# set plotting size parameter
plt.rcParams['figure.figsize'] = (17, 7)
```