**Problem Set 5**

This problem set covers material from class on September 30/October 1 and October 5/6, including material from OIS Sections 6.2, 7.1, and 7.2. It makes use similar data to that we discussed in class. Partial credit may be given for answers that are correct in part, but not in full. As a reminder, you are allowed to work with others, but your answer should be your own. What that means is that on problems that require you to write interpretations or responses, you should make sure that you and those you work with do not provide answers that are constructed similarly. This problem set is due at 11:59 pm on October 7, 2020.

**Part I: Rasmussen vs. Rasmussen 2 (40 pts.)**

In addition to the approval rating numbers from over the summer that we used in class, both Rasmussens also more recently came out with numbers for the presidential race.

These poll results are summarized below:

* Rasmussen Reports: Joe Biden 48% Donald Trump 47%, 3000 likely voters.
* RMG Research (Scott Rasmussen’s firm): Joe Biden 51% Donald Trump 45%, 752 likely voters.

1. For this question, we are first going to focus on **Joe Biden’s support** in these two polls.
   1. Please calculate and interpret the 95% confidence interval for the difference in **Biden’s** support in the two polls. Please show your work. Does there appear to be a significant difference between the two polls? How do you know? (For this problem, you can assume that the necessary conditions have been met and you can set aside that polls have a slightly different sample.) (10 points.)
   2. Now please calculate the z-score for the difference in **Biden’s** support in these polls and then find and interpret the associated p-value. Please show your work for the z-score and explain how you got the p-value. (15 points.)
   3. Finally, please calculate the z-score for the difference in **Trump’s** support in these polls and then find and interpret the associated p-value. Please show your work for the z-score and explain how you got the p-value. (15 points.)

**Part II: A “Tea” Test (30 pts.)**

Following the 2008 Election, a conservative movement known as the Tea Party rose up in opposition to President Obama. While their influence has since [waned](https://apnews.com/64b634a91a2d4933b8bca4c95baa1309), the caucus had significant influence following the 2010 Election. In this section, you will be working with data from the 2012 [American National Election Study](https://electionstudies.org/).

1. In the 2012 ANES, respondents were asked after the election to rate the Tea Party on a scale from 0 to 100, where 0 indicates that they have very cold feelings towards the Tea Party and 100 indicates they have very warm feelings towards the group. The mean rating for the Tea Party after the election was 41.63, the sample standard deviation was 27.39, and 5,310 respondents answered that question. (25 points)
   1. Please find and interpret a 95% confidence interval for the mean feeling towards the Tea Party in the 2016 ANES. Please explain how you found your t-critical value. Does this t-distribution closely approximate a normal distribution? Why? (10 pts.)
   2. Let’s say that your null hypothesis was that the mean rating of the Tea Party was 50. Find a t-score and p-value at the 95% confidence level to test the alternative hypothesis that the mean rating for the Tea Party is different from 50%. Please show your work, interpret your p-value, and explain how you found the p-value. (10 pts.)
   3. The 2012 ANES also asked about support of the Tea Party on a seven-point scale from 1 to 7, where 1 was most supportive and 7 was least supportive. The mean score was 4.49, with a standard deviation of 1.98. 5,506 respondents answered this question. Your null hypothesis is that the mean rating of the Tea Party was 4 (at the middle of the scale). Please test the alternative hypothesis that the mean rating for the Tea Party is *greater* than 4. Please show your work, interpret your p-value, and explain how you found the p-value. (15 pts.)
2. The World Apple and Pear Foundation has collected data on the number of pears (and apples) produced for each country from 2003 to 2013. The United States and Argentina tend to be among the five largest producers of pears. (25 pts.)
   1. The mean difference in the number of pears produced (i.e., US pears produced minus pears produced in Argentina) is 98,316.64 and the standard deviation of the difference is 78,084.04 for the 11 years included in the data. Please conduct and interpret a “peared” t-test at the 95% confidence level to examine whether there is a difference between the number of pears produced in the United States and Argentina. Please discuss how you found your p-value. (15 pts.)
   2. The below histogram shows the distribution of the difference between the number of pears produced in the United States and the number of pears produced in Argentina. Given the number of years included in the study, should this histogram give us pause in conducting this test and interpreting our results? Why or why not? (10 pts.)

