**Problem Set 4**

This problem set covers material from the two classes after the exam, including material from OIS Sections 5.3 and 6.1; this problem set also draws upon information from OIS Section 5.2. 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. **This problem set is now due on Friday March 5 at 11:59 PM.**

**Part I: Hypothesis Testing Framework (20 pts.)**

1. You are a researcher who would like to examine a topic of interest (you may use your project topic here if you would like). Write a sentence or two describing your topic. Then, write your null hypothesis and an alternative hypothesis for both a “not-equal” and either a “greater-than” or “less-than” scenario (based upon which you think makes more sense here). (12 pts).
2. Having collected data, you would now like to evaluate the evidence. (8 pts.)
   1. After getting your results, what type of evidence would lead you to reject the null? Would you ever accept the null? (4 pts.)
   2. Why might you prefer to use a “not equal” alternative rather than a “greater-than”/ “less than alternative” here? (4 pts.)

**Part II: Confidence Intervals and Hypothesis Testing (70 pts.)**

1. On Monday February 22, [Gallup](https://news.gallup.com/poll/329948/biden-gets-high-marks-covid-response.aspx) released its second reading of President Biden’s approval. This question will focus on that survey. (40 pts.)
   1. In this poll, among likely voters (sample size= 1,021) 56% of adults said that they approved of Biden and 40% said that they disapproved. Find and interpret a 95% confidence interval for Biden’s approval and discuss whether you have evidence to suggest that their percentage differs from 50% and how you know this. (12 pts.)
   2. Now, please find the z-score for Biden’s approval here. (12 pts.)
   3. After finding the z-score, what are two options for finding the p-value associated with this estimate? Using one of these, find and interpret the p-value (95% confidence level). Can we reject the null hypothesis here? (8 pts.)
   4. Using a one-sided hypothesis test associated with the same z-score, do you have evidence to suggest with 95% confidence that Biden’s percentage in the poll is *greater than* 50%? How do you know? (8 pts.)
2. On February 23rd, Marist released a poll of New Yorkers that asked a series of questions about New York politicians and issues facing the Empire State. (30 pts.)
   1. Of the 813 registered voters polled, 41% said they thought that Senator Chuck Schumer was doing an “excellent” or “good” job while 52% said he was doing a “fair” or “poor” job. The rest of the sample was undecided. Using the z-score approach, do you find support for the alternative hypothesis that the percentage who thought Schumer was doing a “fair” or “poor” job is above 50%? Please show your work. (20 pts.)
   2. How small would the margin of error have to be for “excellent”/“good” to not be within the margin of error of “fair”/“poor”? (Hint: this comes from pre-exam material) Is Marist’s sample size big enough for “excellent”/“good” to not be within the margin of error of “fair”/“poor”??How big would they have to make it in order to have a margin of error smaller than this number? Please show your work, but double check using a calculator (or R). For these calculations, you can set *p* equal to 0.5. (10 pts.)

**Part III: P-Values and Confidence Levels (10 pts.)**

1. Is there any reason why we have to have a 95% confidence level when hypothesis testing? Denworth suggests using surprisals as an alternative to p-values. Do you find this approach intuitive? Why or why not? What would a p-value of 0.0625 represent in surprisals? (10 pts.)