Student ID_____

U25 363 - Quantitative Political Methods
Exam Two

Instructions: Answer the following problems in the space provided. The exam is open book, but you are not allowed to consult fellow students. Violations of any of these rules will result in a score of 0 and will be treated as an academic integrity violation. For all questions, use a $\alpha = .05$ threshold for statistical significance and two-sided hypothesis tests (where relevant) unless specifically told otherwise. Make sure to interpret your results in both *statistical* and *substantive* terms when asked to do so. In general, you can round to two significant digits after the decimal. Please be sure to show your work to receive full credit.

Question 1

The following is a regression where the outcome is attitudes towards Hillary Clinton as measured by a standard feeling thermometer ratings (the variable ranges from 0 to 100 where 100 indicates high levels of support for Clinton). Researchers use two explanatory variables in their regression. First, they include a standard 7-point party identification measure that ranges from 1 (Strong Republican) to 7 (Strong Democrat). Second, they include a dummy variable (0 or 1) indicating whether the respondent lives in the South (the 13 states of the former Confederacy). The regression includes N=100 observations.

- Write a paragraph interpreting these results both substantively and statistically. For example, discuss both the statistical significance and size of the effect of each independent variable.
- The researchers claim that the regression indicates that partisanship *causes* changes in attitudes towards Clinton. That is, what assumptions would we need to make to interpret these results as causal? And do these assumptions seem reasonable in this situation? Give your answer in a paragraph.
- Calculate the F-statistic for this regression and interpret. You *do not* need to provide a precise p-value, but you *do* need to provide some indication that you understand what the F-statistic means (i.e., what it tells us about the model).

Outcome variable is thermometer rating of Clinton

Intercept	41.00
	(4.24)
Party ID	6.09
	(2.49)
South	−13.53
	(7.06)
R-squared: 0.24	

Stand. errors are in parentheses.

Question 2

Using the setup in problem 1 (but now using a new sample of N=100), we now fit a regression including an interaction between our two main explanatory variables.

- Explain both statistically and substantively the meaning of each of the parameters in the model.
- Write out the prediction equation for people from the south and another prediction equation for those not from the south. Explain why they are different and provide a substantive interpretation. This means that you need to tell me what this interaction model tells us about the political world.

Outcome variable is thermometer rating of Clinton

Intercept	36.00
	(5.44)
Party ID	8.09
	(2.49)
South	−10.53
	(7.06)
South × Party ID	−3.57
	(0.88)
R-squared: 0.35	

Stand. errors are in parentheses.

Question 3

Imagine that the Trump campaign had selected 30% of respondents in our survey to receive multiple mailers (campaign materials delivered by mail) containing negative information about Clinton. They hire you to tell them whether or not it was effective. The bad news is that the campaign did not choose people at random to send the mailer, but rather targeted them based on demographic characteristics. The good news, however, is we had a panel survey where we measured attitudes towards Clinton *before* and *after* the mailers were delivered.

Below is the result of a regression model where we include the covariates "Received Mailers" (0=Did not receive mailers, 1=Received mailers), "Wave 2" (0=Attitude measured before mailers sent, 1=Attitudes measured after mailers sent), and the interaction of these two variables.

- What was the causal effect of the mailers on attitudes towards clinton?
- What is the key assumption necessary for this causal claim to be valid? Give an example of how it could be violated.

Outcome variable is thermometer rating of Clinton

Intercept	36.00
	(5.44)
Received Mailers	8.09
	(2.49)
Wave 2 (measured after mailers sent)	−10.53
	(7.06)
Received Mailers × Wave 2	−3.57
	(0.88)
R-squared: 0.35	

Stand. errors are in parentheses.

Question 4

Congressional Quarterly tracks the percentage of the time members of Congress vote with their party on partisan roll call votes (i.e., those in which a majority of Repub- licans oppose a majority of Democrats). The measure can take values from 0 to 100. If we take a simple random sample of 11 Democratic incumbents from the group described above and their mean party unity score is 93.87 with a sample standard deviation of 9.50, what is the 90% confidence interval for our estimate of mean party unity among this group as a whole? Please show your work!

Question 5

A group of researchers are examining attitudes about the Affordable Care Act. They asked the following question to 781 respondents, "From what you've heard or read, do you approve or disapprove of the health care law that was enacted last year?"

Unfortunately, the researchers have had computer trouble, and they have only been able to retrieve the information presented in the following table. It contains partial information for each cell, including some observed counts, some expected frequencies (in parentheses), and some column and row totals.

	Response		
	Yes	**No**	**Total**
Democrats	221 (205.58)	225	446
Republicans	(154.42)	(180.58)	335
Total	360		781

- Use the information listed to complete the table. Be sure to calculate both the observed and expected frequencies for each cell.
- Calculate the cell component for the χ^2 statistic for the lower-right cell of the table (i.e., Republicans who responded No).
- The χ^2 statistic for this table is (approximately) 5.02. Specify and conduct a hypothesis test using this number.
- In one sentence, what does this table tell us about the relationship between party and attitudes on the Affordable Care Act?

Question 6

You are interested in studying the effect of political knowledge on partisan identification. Party ID is measured on a 7-point scale where 1 indicates a "Strong Republican" and a 7 indicates a "Strong Democrat." The randomly assigned treatment group for your study completed a short class on basic *civics* and the control group completed a short course on *art* appreciation. The data you get back is as follows:

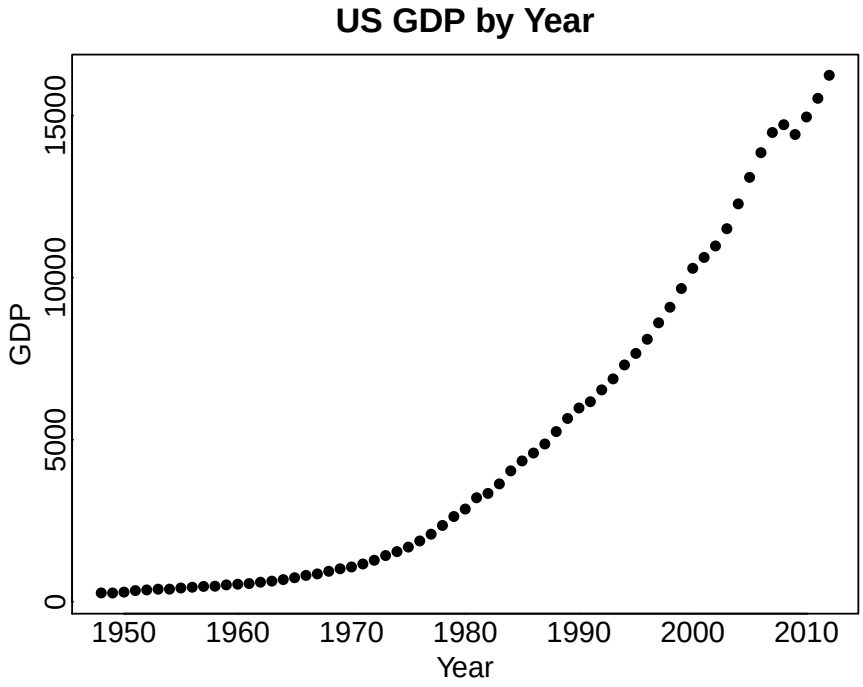
	Civics Class	Art Class
Mean	3.8	3.5
SD	2.4	2.2
N	288	242

- Provide a point and interval estimate (use $\alpha = 0.95$) for the difference in Party ID for the treatment and control groups.
- Test the theory that the civics class changed Party ID.
- Is it OK to treat this estimate as causal? Why or why not?

Question 7

Answer the questions below. It is possible to answer each with only two sentences and receive full credit.

- Explain how a sample distribution, a sampling distribution, and a population distribution are different *and how they are related*.
- Suppose we were interested in studying GDP in the United States. Here is a plot of this variable (the y-axis is in millions of dollars):



- What concerns might we have about using spending as a dependent variable in regression? How could we address these concerns?

Question 8

Define five (5) of the following six (6) terms. Each definition is worth two (2) points. If you provide a definition for all terms, all will be graded and the **highest** score will be dropped.

P-value

Outliers

Counterfactual

Instrument

Autocorrelation

Standard error