

GOVT 2304 – Completion of the Research Project

Due Monday, November 30

Read all directions carefully and follow them closely.

I am here to help. I encourage you **strongly** to schedule an appointment with me to guide you through this process. I will repeat this message at the end of this assignment description.

All work, unless otherwise noted, is to be typed, double spaced in 12 point font with exactly 1" margins and left aligned. Follow the rules set forth in the previous exercise.

Part I: Revisions

Revise both the entire hypothesis formulation and literature review stages of your projects to be submitted with the analysis section below into a single, coherent research paper. (Attach with a paper clip the latest revision I have returned to you so I can measure your progress.) I will be grading most stringently on how well you have revised your earlier work based on my comments and guidance. Address every concern I expressed in reviewing your work, paying careful attention to:

1. State your hypotheses clearly and correctly as Le Roy does, with the modifications I have directed. (This includes the null hypotheses.) Describe the data source and collection method in a more detailed manner than simply naming the database with three letters that someone unfamiliar with the data would not recognize.

2. Provide complete univariate analysis of each variable as demonstrated in class. Include detailed operationalization, discussion of the range of values and distribution, and address skewed or missing data concerns if any exist.
3. Make sure your literature review is a relevant “third person” summation of your topic that cites sources correctly; it is meant to be neither a narrative of your search for sources nor a collection of evidence to support your hypothesis. Summarize what research has already been done on similar topics. Make sure your references page meets with APA or MLA standards and include it as the last and separate page(s) of your overall project.
4. The formulation and literature review stages must be separate sections of your paper, although the order is up to you. Do not go back and forth between setting up what you will test and what conclusions others have drawn, but realize the entire paper is one continuous essay and should have inherent flow.

Part IIA: Testing your hypotheses and drawing conclusions (for nominal/ordinal variables)

If you are using interval/ratio variables you will *not* complete this section.

First, test the null hypothesis that there is no relationship between your independent and dependent variables. Get a cross-tabulation of the variables with the independent variables comprising the columns. Use column percentages and create a table that displays how the two variables relate. Tables may be single spaced, but keep any table on a single page. Do not cut and paste from MicroCase; your originally crafted table should look something like the one below (you can use cell borders if you wish):

		Income Level			
		<\$12K	\$12K - \$24K	\$24K<	
Sex	Male	44.3% (43)	53.2% (76)	54.2% (45)	
	Female	55.7% (54)	46.8% (67)	45.8% (38)	
		100%	100%	100%	N = 323

Do not include missing data in the table but mention if a large number of missing data is a concern. Does the table look as if it supports your hypothesis? This may not be obvious with merely a glance; explain whether it is or not. Report any relevant measures of association and the level of significance and explain why you might favor one measure of association over the other. Does the data warrant rejection of the null hypothesis? (Remember, one does not accept the null hypothesis, but may fail to reject it.) What does the analysis reveal about the relationship between these two variables? How strong, if at all, is the relationship?

Next, do two separate cross-tabulations, first using your control variable with your independent variable, then the control with your dependent variable. You need not present a table here but report the relevant statistics for these two steps. Is the control variable significantly associated with either the dependent variable or independent variable? Use this step to explain whether your control variable might be an antecedent or intervening variable in the above relationship, and how this might be important. (If one or both of these relationships is stronger than the relationship between the independent and dependent variables, the control is likely as antecedent or intervening variable in need of mention).

Now do the cross-tabulation that includes the control variable (the full model). Do not present a table for each value of the control variable as does MicroCase®, but craft a much more simple table that places the relevant measure(s) of association and probability as columns and the levels of your control variable as rows. It should resemble the following:

	lambda	Cramer's V	<i>p</i>
Education <12 years	.321	.356	.002
12 years	.381	.401	.030
1-3 years college	.438	.457	.048
college degree+	.476	.503	.002

Once again interpret these result as to what effect, if any, the control variable has on the relationship between your independent and dependent variables and be specific. For which values of the control variable do you reject the null hypothesis that your independent and dependent variables are unrelated? Does the relationship (or lack thereof) between the independent and dependent variables “remain” for each level of the control variable? It might be that the relationship only holds true for the extremes of the spectrum, just the middle values, or only for one value. You might find a linear pattern like the one above (strength of the relationship “increasing” with education, or find that introducing the control variable had no effect whatsoever.

Part IIB: Testing your hypotheses and drawing conclusions (for interval/ratio variables)

If you are using nominal/ordinal variables you will *not* complete this section.

First, visually test the null hypothesis that there is no relationship between your independent and dependent variables. Present a scatterplot of the variables with the independent variable along the x -axis, the dependent variable along the y -axis to convey the visual relationship between the variables (yes, you may copy-and-paste here from the MicroCase® output). Be sure to include the regression line that helps to make sense of the relationship. You should comment on the observed algebraic relationship between the independent and dependent variables at this point.

Next, obtain and present a correlation matrix among all three variables. By including all three in a correlation matrix, you can learn Pearson's r (the correlation coefficient) for each of the IV-DV, CV-IV, and CV-DV bivariate (only two at a time) relationships. You need not present a table here but report the relevant statistics for these three bivariate relationships (r and p). Is the control variable significantly associated with either the dependent variable or independent variable? Use this step to explain whether your control variable might be an antecedent or intervening variable in the above relationship, and how this might be important. (If one or both of these relationships is stronger than the relationship between the independent and dependent variables, the control is likely as antecedent or intervening variable in need of mention).

Complete the regression analysis with only your independent and dependent variables. When click on the ANOVA summaries, you should be able to find the estimates (they will be the same values revealed when requesting the regression line on the scatterplot). Record the estimates computed to complete the table below. Now, run a regression model that includes your control variable as if it were an additional

independent variable. Once you have run the two regression models, you will need to present a simple table to summarize the results of both the models with and without the control variables that should look something like the one below:

Independent Variable	Estimate	Estimate
Median income	.213** (.057)	.136* (.096)
Percent rural	—	.213 (.057)
Constant (Y-intercept)	.864	.559
Standard error	.073	.122
<i>n</i>	425	424
Multiple R^2	.029	.076

Note: Dependent variable is the House voter turnout. Cell entries are regression coefficients with standard errors in parentheses. * $p < .05$, ** $p < .01$, two-tailed t test.

Tables may be single spaced but any table must be on a single page. Notice, the model on the left does not present parameter estimates for the control variable because it was not included in the original regression. Regarding the relationship solely between your independent and dependent variable, do the results support your hypothesis? Does the data warrant rejection of the null hypothesis? (Remember, one does not accept the null hypothesis, but may fail to reject it.) What does the analysis reveal about the relationship between these two variables?

Interpret the results from the second regression as to what effect the inclusion of the control variable has on the relationship between your independent and dependent variables. Does the relationship (or lack thereof) between the independent and dependent variables “remain” when introducing the control variable? Does the

introduction of the control variable strengthen or weaken that relationship? Is the control variable itself a significant explanatory factor of the dependent variable? Are both the independent and control variables simultaneously and significantly related to the dependent variable?

Part III: Discussion

Finally, make substantive (politically relevant) conclusions about this relationship in light of your completed analysis. Does your analysis contradict, refine or extend the base of existing knowledge on this subject or your own assumptions? Remember, finding no support for your original hypotheses or discovering contradictory evidence is not failure; quite the contrary. The point is to have posited, explored, measured and explained.

Make sure the paper is typed, double spaced with 1" margins. You may use single spacing only in the presentation of tables, but make certain that any tables you present do not appear cluttered. Place one vertically oriented staple in the upper-left corner of the paper and page numbers in the upper right; do not use plastic covers or any other superfluous packaging techniques. *There should be absolutely no contractions in your paper.* The project is due at the beginning of class on Thursday, April 23. Late papers will not be accepted absent extraordinary circumstances, in which case a significant grading penalty will likely be assessed. I expect you to cite sources appropriately and take all necessary steps to avoid even the appearance of plagiarism. Papers including plagiarized material, misrepresented sources or falsified results will receive a score of zero and any student submitting one will receive a failing grade for the course

Final instruction

So that I can check your work, include a separate attachment (affixed to your Research Project with a paper clip). The first page of this stapled attachment will indicate which database you used and list the names and numbers of each of the three variables you employed; be specific regarding which variables are which (label as “IV,” “DV” and “CV”). The remainder of the attachment will consist of the printout of every step of analysis you perform using the MicroCase® software. This should be the only place in your work that you reveal the internal names and numbers MicroCase® uses.

I am here to help. I encourage you **strongly** to schedule an appointment with me to guide you through this process.

After completing your work, reread the assignment to make sure you have followed every direction correctly and performed each task fully. I am not kidding.