

# Problem Set 3

Vismante Dringelyte 17341481

Due: November 19, 2022

## Instructions

- Please show your work! You may lose points by simply writing in the answer. If the problem requires you to execute commands in **R**, please include the code you used to get your answers. Please also include the `.R` file that contains your code. If you are not sure if work needs to be shown for a particular problem, please ask.
- Your homework should be submitted electronically on GitHub.
- This problem set is due before 23:59 on Sunday November 19, 2023. No late assignments will be accepted.

In this problem set, you will run several regressions and create an add variable plot (see the lecture slides) in **R** using the `incumbents_subset.csv` dataset. Include all of your code.

## Question 1

We are interested in knowing how the difference in campaign spending between incumbent and challenger affects the incumbent's vote share.

1. Run a regression where the outcome variable is `voteshare` and the explanatory variable is `difflog`.

Table 1:

	<i>Dependent variable:</i>
	voteshare
difflog	0.042*** (0.001)
Constant	0.579*** (0.002)
Observations	3,193
R <sup>2</sup>	0.367
Adjusted R <sup>2</sup>	0.367
Residual Std. Error	0.079 (df = 3191)
F Statistic	1,852.791*** (df = 1; 3191)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

2. Make a scatterplot of the two variables and add the regression line.

3. Save the residuals of the model in a separate object.

4. Write the prediction equation.

## Question 2

We are interested in knowing how the difference between incumbent and challenger's spending and the vote share of the presidential candidate of the incumbent's party are related.

1. Run a regression where the outcome variable is `presvote` and the explanatory variable is `difflog`.

Table 2:

<i>Dependent variable:</i>	
	presvote
difflog	0.024*** (0.001)
Constant	0.508*** (0.003)
Observations	3,193
R <sup>2</sup>	0.088
Adjusted R <sup>2</sup>	0.088
Residual Std. Error	0.110 (df = 3191)
F Statistic	307.715*** (df = 1; 3191)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

2. Make a scatterplot of the two variables and add the regression line.

3. Save the residuals of the model in a separate object.

4. Write the prediction equation.

### Question 3

We are interested in knowing how the vote share of the presidential candidate of the incumbent's party is associated with the incumbent's electoral success.

1. Run a regression where the outcome variable is **voteshare** and the explanatory variable is **presvote**.

Table 3:

<i>Dependent variable:</i>	
	voteshare
presvote	0.388*** (0.013)
Constant	0.441*** (0.008)
Observations	3,193
R <sup>2</sup>	0.206
Adjusted R <sup>2</sup>	0.206
Residual Std. Error	0.088 (df = 3191)
F Statistic	826.950*** (df = 1; 3191)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

2. Make a scatterplot of the two variables and add the regression line.

3. Write the prediction equation.

## Question 4

The residuals from part (a) tell us how much of the variation in **voteshare** is *not* explained by the difference in spending between incumbent and challenger. The residuals in part (b) tell us how much of the variation in **presvote** is *not* explained by the difference in spending between incumbent and challenger in the district.

1. Run a regression where the outcome variable is the residuals from Question 1 and the explanatory variable is the residuals from Question 2.

Table 4:

	<i>Dependent variable:</i>
	residuals1
residuals2	0.257*** (0.012)
Constant	-0.000 (0.001)
Observations	3,193
R <sup>2</sup>	0.130
Adjusted R <sup>2</sup>	0.130
Residual Std. Error	0.073 (df = 3191)
F Statistic	476.975*** (df = 1; 3191)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

2. Make a scatterplot of the two residuals and add the regression line.

3. Write the prediction equation.

## Question 5

What if the incumbent's vote share is affected by both the president's popularity and the difference in spending between incumbent and challenger?

1. Run a regression where the outcome variable is the incumbent's `voteshare` and the explanatory variables are `difflog` and `presvote`.

Table 5:

	<i>Dependent variable:</i>
	voteshare
difflog	0.036*** (0.001)
presvote	0.257*** (0.012)
Constant	0.449*** (0.006)
Observations	3,193
R <sup>2</sup>	0.450
Adjusted R <sup>2</sup>	0.449
Residual Std. Error	0.073 (df = 3190)
F Statistic	1,302.947*** (df = 2; 3190)
<i>Note:</i>	*p<0.1; **p<0.05; ***p<0.01

2. Write the prediction equation.

3. What is it in this output that is identical to the output in Question 4? Why do you think this is the case?