**JSGS 803 Assignment 3 – Logistic Regression**

1. **Dataset and resource:** 
   1. Choose one dataset that allows you to perform a logistic regression. For example,

* One of the datasets on Canadian Election Survey <http://www.ces-eec.ca/>;

For example, the 2021 Canadian Election Study at <https://dataverse.harvard.edu/dataset.xhtml?persistentId=doi:10.7910/DVN/XBZHKC>

* You can use any dataset, as long as the dataset has more than 300 observations and allows you to conduct a meaningful logistic regression. You can use the same dataset that you use for your Assignment 2 and 4, but the dependent variable has to be a dummy variable.

You may find a dataset from an online data repository. For example,

UCI collection of datasets: <https://archive.ics.uci.edu/>

* 1. Check this and other STATA logistic regression analysis examples.

<https://stats.idre.ucla.edu/stata/dae/logistic-regression/>

1. **Tasks:**

Your task is to build and interpret a logistic multiple regression model using the dataset.

Please prepare a report of your analysis. Your report should be no more than 2 pages. For your data analysis you should do the following:

1.1 Correctly specify the dependent and major independent variables. Your independent variables should be no more than ten variables. Explain what are the three most important independent variables and why? **[20 points]**

1.2 State the hypothesis with regards to the expected sign and magnitude (if applicable) of the slope coefficient of the three most important independent variables. **[20 points]**

1.3 Run a logistic regression analysis and report Stata output. What is the regression equation? (in the form of *Y=a+bX+e*) **[20 points]**

1.4 Interpret the coefficients and model statistics. **[20 points]**

1.5 If you have access to any data you want, what other variables will you include in your regression equation?Why? **[20 points]**