Taniya Adhikari

Final Project

DSC 530

**Statistical/Hypothetical Question**

My Statistical Question for the project was Do tax cuts to corporate create jobs for the economy? This concept of trickle-down economics has been widely accepted since 1980’s. I wanted to further analyze this and see if it actually creates jobs or does it have any relationship to the unemployment. If tax cuts creates jobs, then the unemployment rates should go down as well indicating increased jobs. To create dataset for this project, I focused on only economic indicators and consolidated datasets by year to create dataset with multiple variables. For my hypothesis questions, I decided to study two variables closely, Unemployment rate as an economic indicator for jobs, and corporate tax rate. I also decided to compare USA indicators to Japan Indicators, to further see if corporate tax has the same kind of effects on unemployment rate for Japan.

**Outcome of your EDA**

There were few things, I discovered during my EDA process. Economic Indicators usually have uniform distributions. All the variables in my dataset had uniform distributions except Unemployment rates. Both USA and JPN has almost similar distributions for each indicators, except the values were different. Japan historically has lower tax rates compared to USA. Furthermore, Unemployment rates follow a non-linear trendline over the time. This is similar to stock data, so Time series would’ve been a better analysis of this type of data. In summary analysis, the Cohen Effect Size was quite high for almost all variables, which indicates that larger sample size is not a necessity. The difference between the PMFs of variables of the two countries are 0 because they are uniformly distributed. The CDF of Corporate tax rate for US and JPN had almost the straight line, which also indicates uniform distribution. Further analysis of relationship between the two variables unemployment rates and corporate tax for both countries, indicate negative relationship between the two variables with r = -0.32 for US and r = -0.8 for Japan. Spearmen rho = -.31 for US and rho=-.78 for Japan with covariance also negative. Covariance indicate that they both deviates from it’s mean in opposite direction and there is a negative strong relationship for Japan and weak relationship for US.

* H0: There is no mean difference between the unemployment rate between the two countries
* H1: There is a significant difference between the two country’s unemployment rate.

The p-value = 0.0 and I rejected the null hypothesis. Lastly, regression analysis indicates strong relationship but shows non-linearity. Therefore, regression is not the best model for this.

I also performed Hypothesis Testing for the mean difference between Japan’s unemployment rates and US rates. With

**What do you feel was missed during the analysis?**

Japan had lot of missing data for interest rates, there was an analysis I wanted to perform related to monetary policy and fiscal policy for which I needed interest rates data. For Japan, there were no data till 1989. So I ended up skippy that analysis. I also thing Time series would’ve have been a great addition to the project.

**Were there any variables you felt could have helped in the analysis?**

Wage growth rate, but I couldn’t find the data for both USA and JPN.

**Were there any assumptions made you felt were incorrect?**

Well, initially I assumed I will be doing regression analysis because it will have a linear relationship. To my surprise I was dealing with Uniform distribution, so it already violated the normality assumption of regression.

**What challenges did you face, what did you not fully understand?**

The only challenge I had was finding a right dataset. I was not able to find a dataset which will have all economic indicators and also it was quite hard to find a historical data for these indicators.

I still struggle to understand the python coding with regards to statistical analysis. I am more fond of R, I am still learning Python.