Journal Report 2
9/8/19-9/14/19
Sarah Luthra
Computer Systems Research Lab
Period 2, White

Daily Log

Monday September 9

I found two more papers that I wanted to read about international immigration policies, especially those related to the United States. I read one of the papers.

Tuesday September 10

I read the second paper. Once I was satisfied I had a pretty good understanding about the policy aspect of the project, I downloaded R and RStudio so I could begin writing my regression script. I watched a couple videos on how to write scripts in R and how to use RStudio, and read some example websites on projects that people had used multi linear regressions for.

Thursday September 12

I started writing my own script for an example scenario I had found online. I ran and debugged this script, which took a while since there were many errors I had to look up to continue to understand how to use R and RStudio. I think I have a pretty good grasp of it now, though, and I hope to begin my actual script next week.

Timeline

Date	Goal	Met
Aug 26	Learn how Sys Lab works, find good data sets	Yes, but I didn't download any data sets or do any calculations with them
Sept 2	Download and process data sets, start researching and creating lists of factors for emigration to put through a multi linear regression. Downloaded R packages to write the multi linear regression	Yes, I made lists for 2 countries. I have not finished the regression
Sept 9	Finalize factor list and finish a script in R to run a multi-linear regression to determine p values	Yes, but now I need to do this same thing with my actual immigration factors
Sept 16	Run a regression on my factors and get p values for each factor. Run this by Mr. White to see if this makes sense. Start researching the best way to make my predictive analytics model	
Sept 23	Choose appropriate weightings based on the p values, and start to research how to make my predictive analytics model	

Reflection

I'm glad I have finished my background research so that I can get to the coding part of my project. The titles of the two papers I read are "International Migration Policies in Latin America" by S Torrado and "U. S. Immigration Policy and Latin America: In Search of the "Special Relationship" by RA Pastor. I also think I have a good grasp on how to write scripts in R, so I can accurately run my regression to get the weights I need to run my model. My initial attempt at writing a script took a lot of debugging since I was still trying to learn the syntax and ins and outs of R. Here is a piece of the code that I ran:

```
## predicting the test set result by different modeling output
y_pred1<- predict(regressor1, newdata= testSet)
y_pred2<- predict(regressor2, newdata= testSet)
y_pred3<- predict(regressor3, newdata= testSet)
y_pred4<- predict(regressor4, newdata= testSet)
testSet$Predict1 <- y_pred1
testSet$Predict2 <- y_pred2
testSet$Predict3 <- y_pred3
testSet$Predict4 <- y_pred4
## exporting testset data with predictions in different models
write.csv(testSet, "Profit comparision of different model.csv")
getwd()
```