**Project 2- ETL Proposal**

Amy Brunet, PeiChi Hockaday, and Trang Bui

**Motivations**

Our motivations for this project are an extension of our motivations from project 1, which was to find a relationship between student test scores and various factors affecting student academic performance. In project 1, we looked at the relationship between funding and student achievement, and we found there was not a significant relationship between the two variables. We concluded that more research would need to be done in order to determine which factors truly influence and increase student achievement. Because of those findings, we have decided to continue our research into the factors that influence student achievement, and in this project we will look at the relationship, if any, between poverty rates, median household income, and student achievement in each of the 50 states in the United States over three separate academic years. We will seek to find a correlation between poverty rates and median household income to determine whether that has an impact on student achievement, or perhaps whether it is a predicting factor of how a student will perform academically.

**Datasets**

We found three new datasets (as csv files) from the United States Census Bureau website on the income and poverty rates in each of the 50 states over various years. In addition, we will be using one of the data sets from our previous project that contains data on the average math and reading scores for 4th and 8th grade students in each of the 50 states over the same years. Location of where we found each of the data sets is linked below:

US Income and Poverty Rates, 2013

<https://www.census.gov/data/datasets/2013/demo/saipe/2013-state-and-county.html>

US Income and Poverty Rates, 2015

<https://www.census.gov/data/datasets/2015/demo/saipe/2015-state-and-county.html>

US Income and Poverty Rates, 2017

<https://www.census.gov/data/datasets/2017/demo/saipe/2017-state-and-county.html>

US Average Math and Reading Scores, 2013-2017

<https://www.nationsreportcard.gov/>

**Database**

For our project, we plan to transform and clean the data through python and pandas. Once we have transformed the data, we will then load the data into a PostgreSQL database.