**Group ID: 16**

**Members:** Subrato Chakravorty, Kalvin Goode, Shreyas Rajesh, Jinglong Du

**Proposal: What affects a Ninth Grader’s Science Self Efficacy?**

**Problem:** To investigate the relationship between various student, school, and teacher level variables and students’ science self-efficacy (dependent variable). The project will focus on understanding students' trajectories from the beginning of high school into postsecondary education, the workforce, and beyond.

**Dataset:**

We utilize the data from the [High School Longitudinal Study of 2009 (HSLS:09](https://catalog.data.gov/dataset/high-school-longitudinal-study-of-2009)) to investigate the relationship between various student, school, and teacher level variables and students’ science self-efficacy. HSLS:09 is a nationally representative study that follows a cohort of 9th grade students through their secondary and postsecondary Experiences.

**Proposed Study and Motivation:**

To understand the effect of student, teacher and school level variables on student science self-efficacy, we do analysis in multiple stages. We intend to perform a univariate analysis to provide descriptive statistics of most variables used in the project. We are going to perform bivariate analysis to determine the relationships between various independent variables and the dependent variable (Student’s science self-efficacy).

As the dependent variable in this study – student science self-efficacy – is a continuous variable, three bivariate tests will be utilized: T-tests, ANOVA, and Pearson’s Correlations. We intend to utilize tests such as t-test and anova analysis to explore the relationships between the dependent variable of student science self-efficacy and all categorical independent variables such as student’s sex, school control, science teacher’s sex, science teacher certification, whether the science teacher previously held a science job etc. Pearson’s correlations will be used to investigate the relationship between student science self-efficacy and nine continuous independent variables, including student’s socioeconomic status, student’s science identity, student’s science utility value, student’s science interest, etc. We will also use Ordinary Least Squares (OLS) regression analysis to determine which independent variables have the greatest impact on predicting students’ science self-efficacy.

**Project steps**

|  |  |  |
| --- | --- | --- |
| Step | Estimated completion time | Person(s) in charge |
| 1. Clean Data | 3 days | Kalvin, Subrato |
| 2. Data Analysis | Two weeks | All |
| 3. Data Visualization | One week | Shreyas,Jinglong Du |
| 4. Presentation Slides | 4 days | Kalvin, Shreyas, Jinglong Du |
| 5. Presentation Preparation | 3 days | All |