Capstone Project

**Introduction**

1. Background and Problem

In today’s job market, a college degree is essential for many high-paying jobs. There are many articles and studies that show the relationship between salary and educational attainment. For example, according to an article published by Northeastern University, the average salary for those with high school diploma is $38,792, compared to $64,896 for those with bachelor’s degree[[1]](#footnote-1). However, there are many high school students who have been struggling to keep their grades high enough to enter colleges.

Based on an assumption that grades in high school are positively related to college entrance, this project will examine relationships between academic scores (e.g., ACT) and college enrollment rates in high schools in Chicago. This project will also try to identify nearby universities who could possibly provide educational services (e.g., mentoring or tutoring) to students in high schools who have the lowest academic scores in Chicago.

1. Interest

The results of this project may be utilized by high school teachers, parents, non-profit organizations, universities, and other community members to develop plans to provide educational services to high school students who have been struggling to keep up their classes.

**Data**

1. Data Sources

This project will use data that is provided by Chicago Public Schools, and is available from the website below:

https://data.cityofchicago.org/Education/Chicago-Public-Schools-Progress-Report-Cards-2011-/9xs2-f89t?cm\_mmc=Email\_Newsletter-\_-Developer\_Ed%2BTech-\_-WW\_WW-\_-SkillsNetwork-Courses-IBMDeveloperSkillsNetwork-DB0201EN-SkillsNetwork-20127838&cm\_mmca1=000026UJ&cm\_mmca2=10006555&cm\_mmca3=M12345678&cvosrc=email.Newsletter.M12345678&cvo\_campaign=000026UJ

Due to limited availability of the most current datasets, this study will use the data that was published in “Progress Report Cards (2011-2012).” Variables used for this project are:

* 9th Grade EXPLORE (2010): average score of EXPLORE test
* 10th Grade PLAN (2010): average score of PLAN test
* 11th Grade Average ACT (2011)
* College Eligibility %
* Graduation Rate %
* College Enrollment Rate %

Scores of EXPLORE and PLAN tests in 2011 were not available in this dataset. Therefore, the ones in 2010 will be used. There were multiple scores that may be specific to the 12th grade students, but no detailed descriptions were available for those scores, so academic scores of the 12th grade students will not included.

Since this project will also try to identify locations of universities in Chicago, geospatial data will be obtained from Foursquare.

1. Data Cleaning

There were 93 high schools in this dataset. Among these high schools, there were 18 schools who did not have the complete data for this project. Using Nupmy, the schools with incomplete data will be dropped.

1. <https://www.northeastern.edu/bachelors-completion/news/average-salary-by-education-level/> [↑](#footnote-ref-1)