# Data Collection Report

Jianing Yang

## Get Data

**Where to get data?**

I chose to use <https://api.census.gov/> as the source of data, because this is the official government’s website, which constitutes an authoritative, reliable and up-to-date source.

**What variables to use?**

I selected 12 variables based 4 main categories: race, income, gender and education. Variables in these categories are likely to be helpful for our high school graduation intervention project because I think there could be a high correlation between this family background information and the success of a student in high school.

|  |  |  |
| --- | --- | --- |
| Race | Total | B02001\_001E |
| White alone | B02001\_002E |
| Black or African American alone | B02001\_003E |
| Asian alone | B02001\_005E |
| Some other race alone | B02001\_007E |
| Income | Median household income in the past 12 months (in 2018 inflation-adjusted dollars) | B25119\_001E |
| Gender | Male | B01001\_002E |
| Female | B01001\_026E |
| Education | Total: Male | B15001\_002E |
| Male, 18 to 24 years, Bachelor's degree | B15001\_009E |
| Total: Female | B15001\_043E |
| Female, 18 to 24 years, Bachelor's degree | B15001\_050E |