Both the tools for GWR and those for OLS require that data to be processed conform to a particular variable naming format due to the level of automation which was necessary for early versions of this project as well as in possible future expansions of it, where ethnic breaks in the data are used to estimate the role of pollution in producing interethnic achievement gaps.

**Grades should be formatted following an sy\_d format, where:**

S is r for reading or m for math (other subjects must be added to the code as needed)

R is 14 for 2014 data or 15 for 2015 data (other years must be added to the code as needed)

\_ is an underscore

D is a three or four letter shortening of the demographic the data was collected from.

|  |  |
| --- | --- |
| White | whi |
| Black | bla |
| Native American | nati |
| Asian | asi |
| Latinx | lati |
| Pacific Islander | paci |
| Multiracial/ethnic | mult |
| Overall Population | over |

**Predictors** instead follow a dvy format, where

D is a four or five letter shortening of the name of the demographic the data was collected from.

|  |  |
| --- | --- |
| White | White |
| Black | Black |
| Native American | Nativ |
| Asian | Asian |
| Pacific Islander | Pacif |
| Multiracial/ethnic | Multi |
| Overall Population | Over |

V is an abbreviation of the proper name of the variable in question. Currently variables supported by any provided version of either tool are listed below

|  |  |
| --- | --- |
| Estimated average home block group pollution | Pol |
| Estimated average family income | Inc |
| Estimated average home block group population | Pop |
| Estimated average parental education level | Sch |
| Estimated average percent of parents working in business | Bus |
| Estimated average percent of parents working in the sciences | Sci |
| Estimated average percent of parents working in educational, legal, or artistic fields | Edl |
| Estimated average percent of parents working in healthcare | Hea |
| Estimated average percent of parents working in the service sector | Ser |
| Estimated average percent of parents working in sales or office related positions | Sal |
| Estimated average percent of parents working in the resource extraction, manufacturing, or installation sectors | Ext |
| Estimated average percent of parents working in the logistics or shipping sectors | Log |
| Disadvantage Index A: Title I and Free/Reduced Lunch ranked equally to ethnic variables | DisA |
| Disadvantage Index B: Title I and Free/Reduced Lunch ranked higher than ethnic variables | DisB |
| Disadvantage Index C: Title I and Free/Reduced Lunch ranked lower than ethnic variables | DisC |

Y is 14 for 2014 data or 15 for 2015 data (other years must be added to the code as needed)

**Formatting rules for disadvantage index inputs are explained internally to the tool itself.**