Configuring the SDG

The SDG relies on several different inputs to produce its output data. Each of these files must be edited before running the SDG in order to produce a valid output data set.

Contents

CSV Input files

There are two types of CSV input files: Name files and Interchange data files. These files are required to operate and should be located in the DataFiles directory.

Name Files

Name files may be edited as you see fit, though they contain a large set of default data. These files include first and last names for people based on ethnicity and gender. There's also a set of street names (in *StreetNames.csv*) which are used when generating addresses. Name files are stored in the root of the DataFiles directory and named as follows:

- FirstName-{Ethnicity}-{Gender}.csv
- Surname-{Ethnicity}.csv
- StreetNames.csv

Interchange Data Files

Interchange data files contain data from several Ed-Fi interchanges that are required to bootstrap an output data set. These files are located in subfolders of the DataFiles directory based on Interchange. This includes the following interchanges:

- AssessmentMetadata
- Descriptors
- EducationOrganization
- EducationOrgCalendar
- MasterSchedule
- Standards

It's highly recommended to use the example files for the NorthRidge data set as a starting point for creating a new data set. Every one of these CSV files must be customized for your specific output data set.

XML Configuration

Parameters for shaping the school district for which you'd like to generate are all found within the XML configuration file. Here you'll outline what your school district and output data looks like by configuring things like:

- Periods for which you want to generate data
- District info
    - Schools in the district
- School info
    - Discipline
    - Attendance
    - Grade levels
    - Staff population demographics
        - Race
        - Gender
- Grade level info
    - Student count
    - Student profile
    - Assessment participation
    - Graduation plans
- Student population profiles
    - Race
    - Gender
    - Status (e.g., Homeless, Immigrant, Economically Disadvantaged)

The SampleConfig.xml file is a good starting point for understanding the XML configuration. The file contains comments that explain the elements used to configure a student population. NorthRidge.xml also provides a more complete example after you understand the basics of the configuration format.

Seed Input File

The seed is used as an *input* to the SDG only when the tool is run in *Standard* mode and the *-seedFilePath* command line argument has been provided. This file can be hand-crafted, though it may be easier to run the SDG in seed output mode to produce this file and then use it as an input in subsequent runs of the SDG.

See Seed Data documentation for further details.