

# Sample Data Generator

The Sample Data Generator produces realistic, cohesive, and 100% fictional datasets for use in demonstrations and testing.

The SDG prefers statistically realistic patterns (e.g., a student with poor attendance generally tracks to poor grades, students are the appropriate age for their grade level, students who are English learners have home languages that track to their ethnicity, and so forth). The system is configurable, and can produce arbitrarily large datasets.

While the SDG creates data with realistic patterns, it is randomly generated and must not be used in place of real-world data for scenarios such as training for machine learning or other algorithmic approaches.

## Documentation

---

- [What's New](#)
- [Getting Started](#) (*This information is available to Ed-Fi licensees only, and requires you to be logged in.*)
- [User Guide](#)
- [Developers' Guide](#)

