

Open Source Libraries for Privacy Newcomers

If you're familiar with their languages, following open source libraries are a great way to start producing some deidentified data to submit to the CRC or test out the sdnist report generator. Feel free to try more than one to compare them. These libraries have user documentation and engaged creators -- just reach out to the creators if you have any issues or questions with them!

R:

- **R Synthpop Library**
- **Non-Parametric Bayesian Multiple Imputation for Categorical Data (NPBayesImputeCat Library)**

Python (with Pandas):

- **SmartNoise-Synth**
- **The Synthetic Data Vault**
- **Tumult Analytics Open-Source Library**
- **Synthetic Minority Oversampling Technique (SMOTE)**

Rust:

- **OpenDP Library**

Reference Links:

1. R Synthpop Library

- Location: <https://www.synthpop.org.uk/get-started.html>
- Installation: <https://www.synthpop.org.uk/get-started.html#installation>
- Documentation: <https://www.synthpop.org.uk/resources.html>

2. NPBayesImputeCat Library

- Location: <https://CRAN.R-project.org/package=NPBayesImputeCat>
- Installation: https://cran.r-project.org/src/contrib/NPBayesImputeCat_0.5.tar.gz
- Documentation: <https://cran.r-project.org/web/packages/NPBayesImputeCat/NPBayesImputeCat.pdf>

3. SmartNoise-Synth

- Location: <https://github.com/opendp/smartnoise-sdk/tree/main/synth>
- Installation: <https://github.com/opendp/smartnoise-sdk/tree/main/synth#installation>
- Documentation: <https://docs.smartnoise.org/synth/index.html>

4. The Synthetic Data Vault

- Location: <https://sdv.dev/SDV/>
- Installation: https://sdv.dev/SDV/getting_started/install.html
- Documentation: https://sdv.dev/SDV/user_guides/index.html

5. Tumult Analytics Open-Source Library

- Location: <https://docs.tmlt.dev/analytics/latest/index.html>
- Installation: <https://docs.tmlt.dev/analytics/latest/installation.html>
- Documentation: <https://docs.tmlt.dev/analytics/latest/tutorials/index.html>

6. SMOTE

- Location: https://imbalanced-learn.org/stable/references/generated/imblearn.over_sampling.SMOTE.html
- Installation: <https://imbalanced-learn.org/stable/install.html>
- Documentation: https://imbalanced-learn.org/stable/user_guide.html

7. Open DP Library

- Location: <https://github.com/opensdp/opensdp>
- Installation: <https://github.com/opensdp/opensdp#installation>
- Documentation: <https://docs.opensdp.org>