

Project Datasets

Feel free to use one of these datasets for your group project. You can always use a data set not from this list if you prefer, but these have the advantage of being “pre-vetted”, and definitely meet the criteria for the project.

Massachusetts 2020 Traffic Data

- **Dataset:** [Google Sheet](#), [CSV](#)
- **Original Source:** Multiple sources, [detailed information here](#)

Massachusetts 2020 vs. 2019 Traffic Data Comparison

- **Dataset:** [Google Sheet](#), [CSV](#)
- **Original Source:** Multiple sources, [detailed information here](#)

International Powerlifting Results

- **Dataset:** [Google Sheet](#), [CSV](#)
- **Original Source:** [Data and Info here](#)

Avocado Prices by Region

- **Dataset:** [Google Sheet](#), [CSV](#)
- **Original Source:** <https://www.kaggle.com/neuromusic/avocado-prices>

Electric Vehicle Charging Sessions

- **Dataset:** [Google Sheet](#), [CSV](#)
- **Original Source:** [Data here](#), [Info here](#)

Lego Prices

- **Dataset:** [Google Sheet](#), [CSV](#)
- **Original Source:** Modified from [Peterson & Zeigler \(2021\)](#), Updated [info here](#)

Coffee Ratings:

- **Dataset:** [Google Sheet](#), [CSV](#)
- **Original Source:** <https://github.com/jldbc/coffee-quality-database>

Maternal Smoking and Infant Health

- **Dataset:** [Google Sheet](#), [CSV](#)
- **Original Source:** [Info here](#), [Data here](#) (from the link titled “Additional data on the parents in the Child Health and Development Study”)

Movies and the Bechdel Test

- **Dataset:** [Google Sheet](#), [CSV](#)
- **Original Source:** [Data here](#), [Info here](#)

Amazon Books

- **Dataset:** [Google Sheet](#), [CSV](#)
- **Original Source:** <https://dasl.datadescription.com/datafile/amazon-books>

Saratoga, New York Housing Prices

- **Dataset:** [Google Sheet](#), [CSV](#)
- **Original Source:** [Data here](#), units of measurement [here](#)

Mario Kart Auction Prices

- **Dataset:** [Google Sheet](#), [CSV](#)
- **Original Source:** [Data here](#)

Additional Places to Look:

- [fivethirtyeight](#) package
- [moderndive](#) package
- [Kaggle.com](#) (warning: not all data sets are well formatted)
- [UC Irvine Machine Learning Repository](#) (warning: not all data sets are well formatted)