Notebooks:

- data-exploration.ipynb: Test out the climata module and choose a stream to predict.
- data-processing.ipynb: Import data from two sources, combine the data sets, clean up the data, and compute two additional columns.
- watershed-delineation-land-use: Use the pysheds module to delineate the watershed for the selected stream gage, use NLCD data to calculate the percent area for the various land use types.
- zumbro-flow-prediction: Normalize data, build ML model, and visualize results.

Data (included in zip file / GitHub repo):

- final_data_extra_processing.csv: a result file from processing and cleaning data before normalizing it in the zumbro_flow_prediction notebook.
- roch_temp_hourly_cleaned.csv: input data set containing environmental variables used (not including streamflow) that is merged with streamflow and further cleaned to produce final_data_extra_processing.csv.
- sites.csv: a file containing the site codes for the rivers we were initially interested in researching; is only used in data-exploration.ipynb.

Data (included in Google Drive folder):

- digital_elevation_model_30m_15N_2.tif: A tif file of DEM data for all of MN, projected to the UTM zone 15N coordinate system.
- NLCD_2016_masked.tif: Land use data clipped to the extents of the boundary for the Rochester stream gage.
- NLCD_legend.xlsx: Contains a look-up table that matches the number in the NLCD tif file to a descriptor of the land use.