

# Preliminary Analysis of Surface Water Toxicity Dataset

Created by: Travis Ludlum

Date: 04/19/2016

# General Information

- Data set used:
  - Surface Water Toxicity 2003-2014
- Language used:
  - Python 2.7

# Original Goals

- Analyze change in water quality over time at each station based on sample results
- Compare water quality in different regions of CA based on common organisms and analytes measured
- Identify any inconsistencies in reported data as well as bad data entries

# Changes to Be Made:

- Use Python's CSV package to read in CSV file instead of splitting by commas
- Remove duplicate data points and bad data points from analysis
- Compare changing water quality over time for each station/region
- Identify which stations are in which regions

# Capabilities

- Creates an object for each entry in the dataset
- Converts dates and times into numerical values for the purpose of sorting the data chronologically
- Calculates averages for each organism and its various analytes, on a by-project and by-station-name basis.
- Outputs results by project to a CSV results file.