**Exploring USGS streamflow data with dataRetrieval**

**Author:** David Manning, Department of Biology, University of Nebraska at Omaha. Email: davidmanning@unomaha.edu

**Course Information**

Department: Biology/Environmental Science/Water Resources

Level: **Upper Undergraduate**

Course type: **Lab**

Students: **Majors**

Number of Students: **20**

**Focus:** This Swirl lesson walks through how to retrieve, view, and visually analyze streamflow and water quality data from the National Water Information System (US Geological Survey).

**Overview:** In this Swirl lesson, we explore publicly available data collected by USGS to explore patterns of hydrology in Nebraska streams, using the R package 'dataRetrieval' to access the data, and then visualize the data in R for graphical analysis.

**Learning objectives:** By the end of the Swirl lesson, students should be able to:

1) Access data from USGS using the dataRetrieval package.

2) Explore trends in stream discharge from any site where these data are available (e.g., Elkhorn River, Nebraska) using graphs.

**Lesson sequence:**

1. Load Swirl course, and 'dataRetrieval' in R.
2. Examine capabilities of dataRetrieval functions, starting with the information that's provided for a given USGS site using the function siteINFO().
3. Examine at what data are available at given sites using the function whatNWISdata().
4. Access some data. The lesson starts with daily means of discharge from the Elkhorn River in Nebraska using the function readNWISdv(), but any USGS site of local interest could be used.
5. Plot the discharge data for the Elkhorn River in a given water year (2018-2019, for example) and describe patterns.
6. Explore the range and the dates when minimum and maximum flow occurred for a given water year and site.

**Pre-lesson activities:** Students should have a basic understanding of how streamflow can be measured through time, and how it should change in response to precipitation events. A complementary Swirl lesson or other pre-lesson module that introduces R basics (vectors, data frames), and ggplot2 would be beneficial.

**Implementation notes:**

* This is a lesson that could occur during a lab, lecture or discussion that introduces streamflow and effects of human activities in watersheds on river flow regimes (e.g., urbanization, dams, climate change). This would help with defining some of the metrics that are used, like discharge (volume of water/time).
* A brief pre-lesson module or introduction to basic R coding, and making plots with ggplot2 would be beneficial.
* Toggling between the syntax of the two packages that are used, dataRetrieval vs ggplot2, might cause some confusion.
* Be aware that some of the functions in dataRetrieval have idiosyncratic arguments, even among functions unique to the package. See <https://cran.r-project.org/web/packages/dataRetrieval/vignettes/dataRetrieval.html> for helpful package vignettes.
* In my experience a key piece for helping students understand these lessons is highlighting that streamflow data are often collected in real-time just down the street across the USA. To this end, I have also used this lesson to prompt students to look for and access data from stream sites that are near their hometown (instead of the example sites given in the lesson).

**R Package citation:**

De Cicco, L.A., Hirsch, R.M., Lorenz, D., Watkins, W.D., 2018, dataRetrieval: R packages for discovering and retrieving water data available from Federal hydrologic web services, doi:10.5066/P9X4L3GE