Instructor: McHale, Michael, Research Hydrologist, USGS

Code transcribed by: Charles Stoll

R-class Date: 02/09/2018

Course: Intro to mapping

Introduction to several R packages used to import and visualize data. Libraries include: maps; readxl; dplyr; dataRetrieval; maptools; rgdal; sp. Additional topics covered included using the unique and filter functions, as well as, the %in% command. Data and supplementary materials supplied by course instructor (M. McHale, Research Hydrologist, USGS). Reference R script for specific code and additional commentary.

This lesson introduces users to the Base Map package and some of its functionality. The lesson gives an overview of some basic functions; e.g. map generation, accessing map attributes, table query/manipulation, and map output as a JPEG. It also introduces the user to the dataRetrieval tool, which enables users to access State and Federal water quality data repositories; NWIS was used for this lesson plan.

**Script Workflow**

Initial steps in the lesson script load required packages and setup the working directory; required packages may need to be installed prior to running. Of note, this script and future lessons will use a specific file structuring; refer script commentary for details.

After setup, the script introduces basic map generation using the supplementary excel sheet “QWsitesformap.xlsx.” After reading the excel table into a data frame, headers are displayed for the user to see; this is a useful function, and an alternative method is described in script commentary. The exercise then introduces the ***map*** command.

The first map generated is of the contiguous United States. Currently, this package does not include Alaska, Hawaii, Puerto Rico, or any other territories; there are alternative methods that allow the user to call on external base maps, shapefiles, or other spatial datasets. A second map is generated of the State of NY, demonstrating the ability to limit map output to specific regions. The function to reset the plot window is written into the script after the display of both maps to illustrate that when new maps are generated, it will overwrite the existing map in the plot window. To avoid losing a map, it needs to be saved prior to generating a new one; this is discussed later.

To show how the user can add point features, a single point is added to the map of NY by specifying coordinates. A whole point layer is then created using the site locations previously loaded into a data frame. During the creation of a layer it is noted that the type, size, and color of symbols can be specified.

In addition to adding point features to the map, the lesson demonstrates how to add a title, and subtitle to the map. Of note, and further described in the script commentary, is that if a title is generated when there is already an existing map title, the new title will overlay the existing title. Unlike the ***map*** command that overwrites a previous map, generating multiple titles will place both in the same map; you must regenerate the map to remove the old title.

To avoid map loss and/or issues with overlaying titles, you can quickly save a map as a jpeg file. The script demonstrates the steps necessary to output a JPEG of the map. There are several input options to JPEG creation that control image production that is was recommended the user play around with; e.g. it is not recommended to us a resolution >300 due to file size. Of particular note is the use of the ***dev.off()*** command *(the same to reset the plot window)*; this instructs R to close the function and finalize the JPEG.

The introduction to Base Map ends with demonstrating the command function to call on the help documentation associated with the map package. This details all the functions and source materials included with the map package. This can be accessed using the ***?map*** command. There is also a short demonstration of user ability to access attributes with the map object; e.g. county names.

The last section included with this lesson script demonstrates the use of the dataRetrieval tool. As described in the script commentary, this enables users to access repositories of data, like NWIS. This portal provides access to data from State and Federal agencies, including the USGS & EPA among others.

The dataRetrieval tool demonstration uses a different supplementary excel table; QWsites.xlsx. This table has the station id associated with many sampling locations throughout the country. The script uses station id to create a variable used by the dataRetrieval tool. Station id, along with the parameter code for Dissolved Organic Carbon (DOC) water chemistry (ug/l) and a specific date range, are used to query and retrieve the records from the NWIS portal.

The lesson ends by working through creating a layer from the retrieved data where DOC values measure >20, and by illustrating how to create a map using two distinct layers of data. The layer containing all DOC values is symbolized using one symbol definition, and the layer containing DOC values >20 are symbolized by a different symbol definition. Finally, both point features are added to the plot display over a map of the contiguous United States.