

1: R Intro and Data Visualization Challenge

Water Data Analytics | Kateri Salk

Spring 2022

Lesson Objectives

1. Explore components of effective data visualization
2. Apply R coding skills to create effective visualizations
3. Review and/or develop skills in R coding and syntax

Opening Discussion

What makes an effective data visualization? What are the essential components/characteristics, and what are optional but nice-to-have components/characteristics?

Exploring web resources for visualization

Here are a couple of my favorites. Browse through these to get some useful ideas for R code that will improve the functionality and graphic design of your data visualization:

From Data to Viz

Data Visualization Example from Our Coding Club

A note on color palettes

In general, the ggplot base color palettes are not the most effective option for data visualization. Compiled below is a list of color palettes available in R that may serve your purposes better. Note that when working with multiple colors, you must choose between sequential, diverging, and qualitative color palettes. Under what circumstances might each be useful?

RColorBrewer (package)

- <http://colorbrewer2.org>
- <https://moderndata.plot.ly/create-colorful-graphs-in-r-with-rcolorbrewer-and-plotly/>

viridis and viridisLite (packages)

- <https://cran.r-project.org/web/packages/viridis/vignettes/intro-to-viridis.html>
- https://ggplot2.tidyverse.org/reference/scale__viridis.html

colorRamp (function; comes with base R as part of the grDevices package)

- <https://bookdown.org/rdpeng/exdata/plotting-and-color-in-r.html#colorramp>

MetBrewer: Palettes inspired by works at the Metropolitan Museum of Art in New York (package)

- <https://github.com/BlakeRMills/MetBrewer>

Session Set Up

```
getwd()

## [1] "/Users/katerisalk/Box Sync/Courses/Water Data Analytics"

# install.packages("tidyverse")
# install.packages("dataRetrieval")
# install.packages("zoo")
# install.packages("ggrepel")

library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.1 --

## v ggplot2 3.3.5      v purrr  0.3.4
## v tibble  3.1.6      v dplyr  1.0.7
## v tidyr   1.1.4      v stringr 1.4.0
## v readr   2.1.1      v forcats 0.5.1

## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()     masks stats::lag()

library(dataRetrieval)
library(zoo)

##
## Attaching package: 'zoo'

## The following objects are masked from 'package:base':
##
##      as.Date, as.Date.numeric
```

Data Visualization Challenge

We will be exploring discharge data for the Eno River, aided by the dataRetrieval package. A basic ggplot is provided for you here. Your assignment this week will be to edit the dataset and update the graph with some data visualization best practices.

```
# Import data
EnoDischarge <- readNWISdv(siteNumbers = "02096500",
                           parameterCd = "00060", # discharge (ft3/s)
                           startDate = "2012-01-01",
                           endDate = "2021-12-31")

# Look at the data frame in your Environment tab.

# Renaming columns (one method of multiple)
names(EnoDischarge)[4:5] <- c("Discharge_cfs", "Approval.Code")

# dataRetrieval also includes attribute information
attr(EnoDischarge, "variableInfo")

##      variableCode      variableName      variableDescription
## 1      00060 Streamflow, ft3/s Discharge, cubic feet per second
##      valueType unit options noDataValue
## 1 Derived Value ft3/s      Mean          NA
```

```
# note: imperial, not metric
attr(EnoDischarge, "siteInfo")
```

```
##              station_nm  site_no agency_cd timeZoneOffset
## 1 HAW RIVER AT HAW RIVER, NC 02096500      USGS      -05:00
##   timeZoneAbbreviation dec_lat_va dec_lon_va      srs siteTypeCd   hucCd
## 1                     EST   36.08722  -79.36611 EPSG:4326      ST 03030002
##   stateCd countyCd network
## 1      37    37001    NWIS
```

```
# Build a ggplot
```

```
EnoPlot <-
  ggplot(EnoDischarge, aes(x = Date, y = Discharge_cfs)) +
    geom_line() +
    xlab("Year")
EnoPlot
```

