

# Module 7 Homework: Manipulating Colorado River data

*Natalie Nelson, PhD*

## 1. Overview

The Colorado River provides water to one in eight Americans and supports one-seventh of the nation's crops (*Unplugging the Colorado River*, NYT), making it one of the most important water bodies in the US. The Colorado River's flows are regulated by two major dams: the Hoover Dam (construction completed in 1936) and Glen Canyon Dam (construction completed in 1963). Lake Mead and Lake Powell are the reservoirs of the Hoover and Glen Canyon Dams, respectively. The Colorado River crosses the US-Mexico border and discharges into the Pacific Ocean. Lees Ferry is located in the reach of the Colorado River just downstream of the Glen Canyon Dam.

**Data.** Data analyzed in this session are from the Department of the Interior (US Geological Survey and Bureau of Reclamation) and include flow measurements in cubic feet per second (cfs) from three locations along the Colorado River: Lake Powell, Lees Ferry, and the US-Mexico border (Fig. 1). Observations were collected over different periods for each of the three sites, with the earliest dating back to October 1921. Additionally, depth measurements in feet are provided for Lake Mead. The data files are as follows:

- Flow (cfs) at the US-Mexico border: `border_flow.csv`
  - Depth (ft) of Lake Mead: `lake_mead_depth.csv`
  - Flow (cfs) upstream of Lake Powell: `lake_powell_flow.csv`
  - Flow (cfs) at Lees Ferry, downstream of the Glen Canyon Dam: `lees_ferry_flow.csv`
-

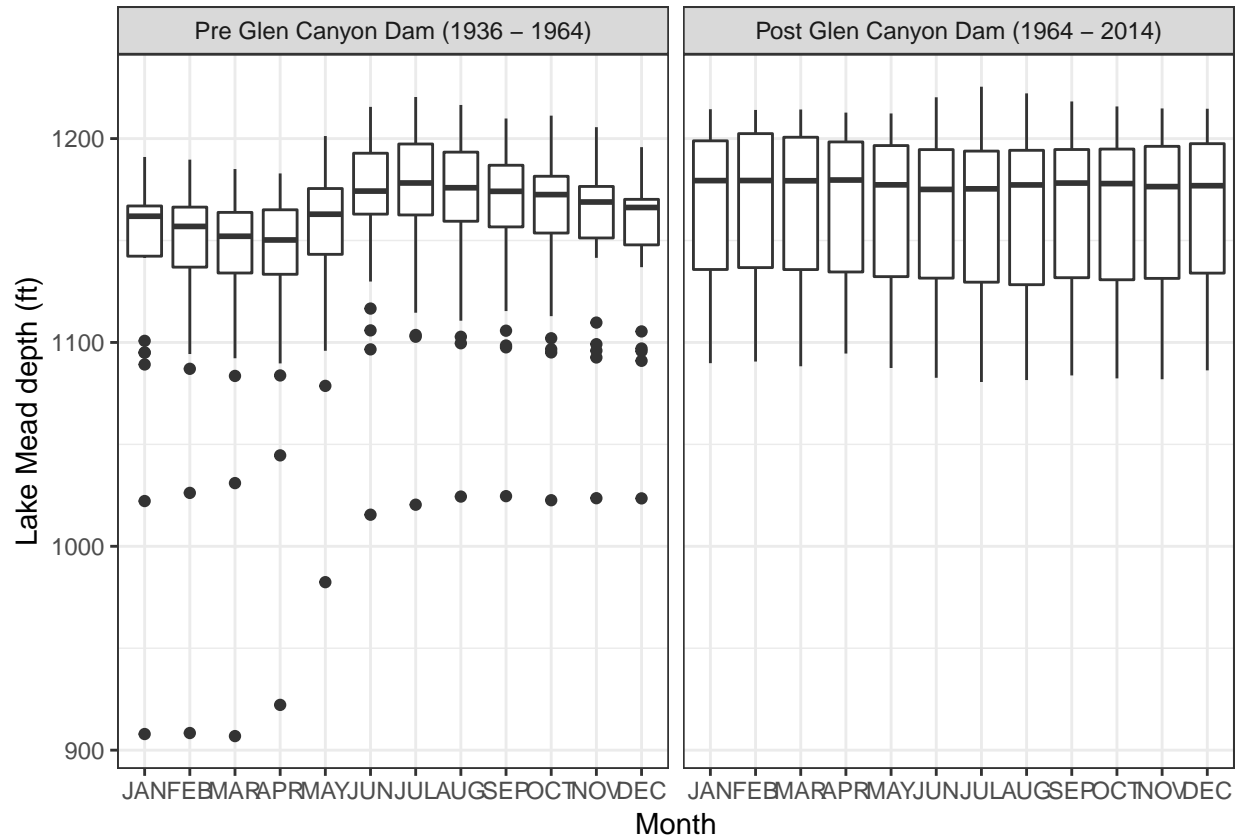


Fig. 1. Major dams and reservoirs of the Lower Colorado River. Map adapted from the New York Times.

## 2. Assignment

For the assignment, you will need to create code to generate the two figures shown below. **Pay close attention to detail, and ensure your code recreates all elements of the figures.** Be sure to include a code header, code sections, and annotations throughout. You will use functions from `dplyr`, `tidyr`, `ggplot2`, `forcats`, and `lubridate`. You should only need to use functions we've reviewed in the class; you are not expected to find functions online that have not been reviewed in course materials. When in doubt, refer to your cheat sheets! Hints are provided below each figure.

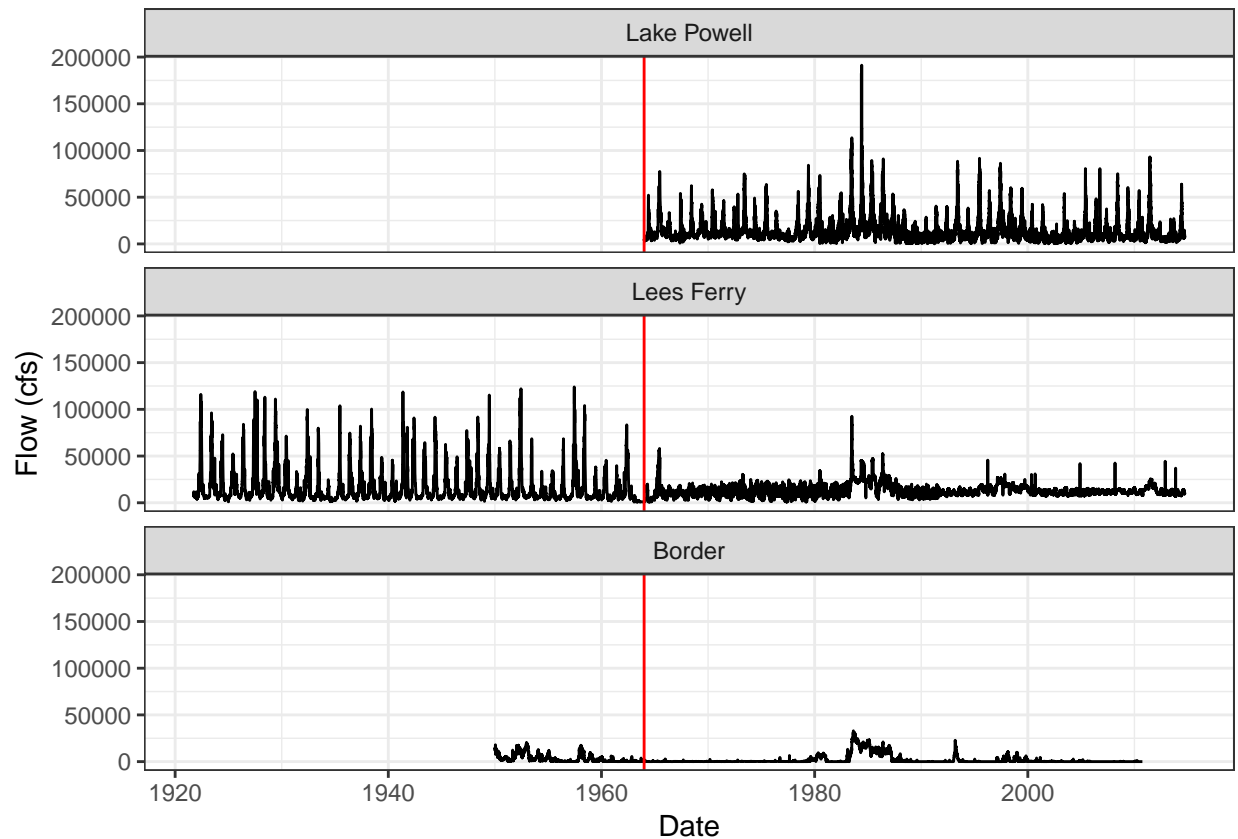
### 2.1. Water depths in Lake Mead



#### Hints:

- You will need to create a new column with the “Pre Glen Canyon Dam” and “Post Glen Canyon Dam” labels. If you’re unsure how to do this, refer back to the AQI exercise from Module 5. You can assume the dam construction was complete by 1964-01-01.
- The theme used in the figures is `theme_bw()`

## 2.2. Flows along the Colorado River



#### Hints:

- When you join your data, you should not exclude any rows. All rows from the datasets should remain in the data frame you create the plot from.
- You can only join two data frames at a time. For example, if you have data frames A, B, and C, you cannot use `*_join(A, B, C)`. Instead, you should chain the join functions together with pipes:

```
A %>%
  *_join(B) %>%
  *_join(C)
```

- The vertical line in the plot corresponds to the date in which the Glen Canyon Dam construction was completed (1964-01-01). In the geom you use to add the vertical line, you will need to specify the `xintercept`. Since the x-axis is created from a vector with data type `date`, the `xintercept` will also have to be a date. To do this, you can use one of the `lubridate` parse functions, e.g. `xintercept = choose-a-parse-function("1964-01-01")`.