WeilinWang_A06_CraftingReports.Rmd

Weilin wang

Fall 2024

Objectives:

1. More practice with R code chunk options
2. Gain proficiency with figures, tables (w/`Kable`) table of contents, etc.
3. Debugging knitting issues

Directions

1. Rename this file `<FirstLast>_A06_CraftingReports.Rmd` (replacing `<FirstLast>` with your first and last name).
2. Change "Student Name" on line 3 (above) with your name.
3. Work through the tasks, **creating code and output** that fulfill each instruction.
4. Be sure your code is tidy; use line breaks to ensure your code fits in the knitted output.
5. Be sure to **answer the questions** in this assignment document.
6. When you have completed the assignment, **Knit** the text and code into a single PDF file.
7. **Be sure that you also commit and push your final Rmd document to your GitHub account**.

Task 1 - Basic Markdown

Using markdown, create a 2-column table beneath the `Table: EPA Air Quality` line below that summarizes the metadata of the EPA Air Quality data.

- The first column should have the header `Item`, below which are three rows listing these metadata attribute item names: `Filename`, `Date`, and `Source`.

- The second column should have the header `Value` and include the metadata values: `EPAair_O3_PM25_NC1819_Processed`, `2018-2019`, and `EPA Air Quality System (AQS)`.

- The first column should be aligned to the right and the second to the left.

Table 1: EPA Air Quality

Item	Value
Filename	EPAair_O3_PM25_NC1819_Processed.csv
Date	2018-2019
Source	EPA Air Quality System (AQS)

Task 2 - Import packages and data, suppressing messages

Set the following R code chunk so that it runs when knit, but no messages, errors, or any output is shown. The code itself, however, should be displayed.

```r
#Import libraries
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ----------------------- tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.1
## v purrr     1.0.2
## -- Conflicts ---------------------------------------- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become error
```

```r
library(lubridate)
library(here)
```

```
## here() starts at /home/guest/EDE_Fall2024
```

```r
library(knitr)

#Import EPA data (from the processed_KEY folder) & fix dates
epa_data <- read.csv(
  here("Data","Processed","EPAair_O3_PM25_NC1819_Processed.csv"),
  stringsAsFactors = TRUE) %>%
  mutate(Date = ymd(Date))
```

Task 3: Creating tables

Set the following R code chunk to display two tables, using knitr's `kable()` function, one listing the mean PM2.5 concentrations for each county, and the other the same except for Ozone.

- The titles should be "Mean Particulates (2.5mm)" and "Mean Ozone", respectively.
- The column names should be "County" and "µg/m3" for both tables. (See tip below.)
- Finally, round the concentration values in the 2nd column to two decimal places.

Customize the chunk options such that the code is run but is not displayed in the knitted document. The output, however, should be displayed.

   TIPS:

- Use `"$\\mu g/m^3$"` as a column name to generate a nicely formatted string via markdown/MathJax notation

- If your output table spans across two pages, try inserting a page break in the markdown just before your code chunk.

```r
#Compute mean PM2.5 concentration for each county, sorted by value
summary_PM25 <- epa_data %>%
  drop_na(AQI_PM2.5) %>%
  group_by(COUNTY) %>%
  summarize(mean_PM2.5 = mean(AQI_PM2.5)) %>%
  arrange(mean_PM2.5)

#Compute mean ozone concentration for each county, sorted by value
summary_Ozone <- epa_data %>%
  drop_na(AQI_Ozone) %>%
  group_by(COUNTY) %>%
  summarize(
    mean_Ozone = mean(AQI_Ozone)) %>%
  arrange(mean_Ozone)

#Use the `kable` function to produce the 2 tables in markdown format
# Display the PM2.5 table
kable(summary_PM25, col.names = c("County", "$\\mu g/m^3$"), caption = "Mean Particulates (2.5mm)")
```

Table 2: Mean Particulates (2.5mm)

County	$\mu g/m^3$
Haywood	13.98400
New Hanover	15.60681
Avery	18.27941
Edgecombe	26.06503
Pitt	27.37166
Guilford	29.14163
Swain	30.62780
Johnston	33.02695
Durham	33.53770
Mecklenburg	33.63038
Forsyth	35.09282
Wake	37.45423

```r
# Display the Ozone table
kable(summary_Ozone, col.names = c("County", "$\\mu g/m^3$"), caption = "Mean Ozone")
```

Table 3: Mean Ozone

County	$\mu g/m^3$
Swain	35.58367
Avery	38.39308
Wake	38.61345
New Hanover	39.11688
Edgecombe	39.22154
Johnston	40.33849

County	$\mu g/m^3$
Mecklenburg	40.45746
Durham	40.69882
Pitt	41.64147
Forsyth	44.02352
Haywood	44.75049
Guilford	45.86681

" `

Task 4: Plots

Below this paragraph, but above the horizonal line (`---`), create two separate code chunks that create boxplots of the distribution of Ozone levels by month using, one for only records collected in 2018 and one for records in 2019. Customize the chunk options such that the final figures are displayed but not the code used to generate the figures. In addition, align the plots on the left side of the page and set the figure heights so both plots fit on the same page with minimal space remaining. Lastly, add a `fig.cap` chunk option to add a caption (title) to your plot that will display underneath the figure.
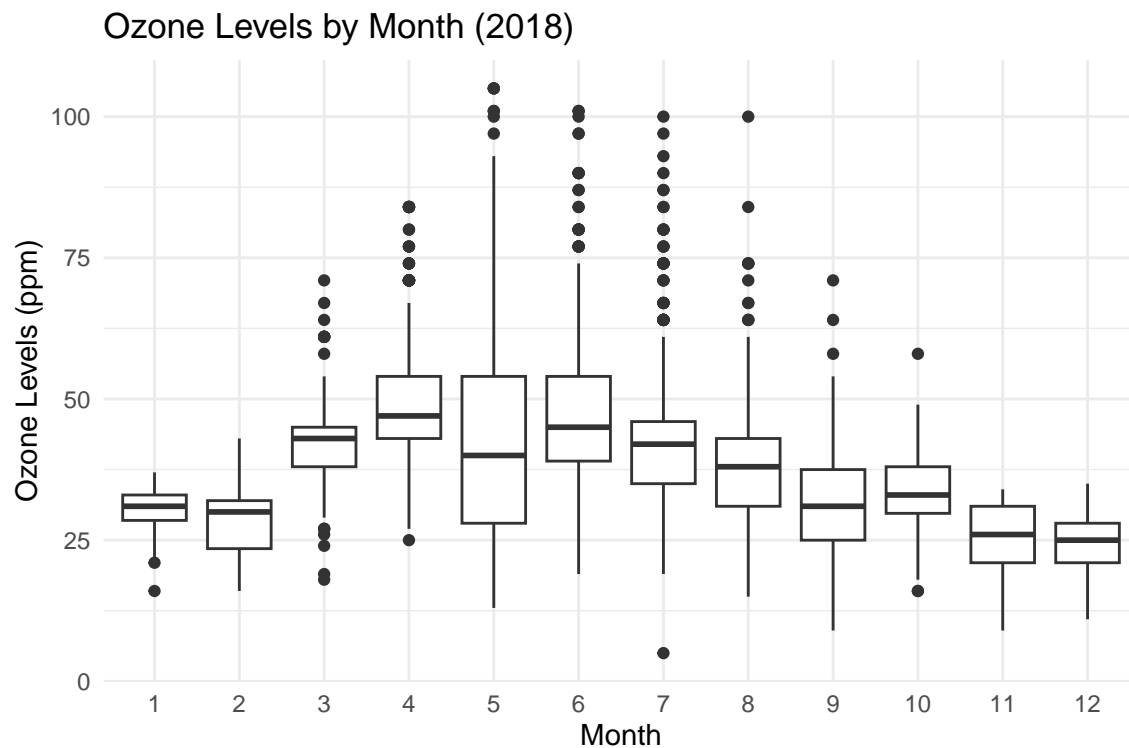

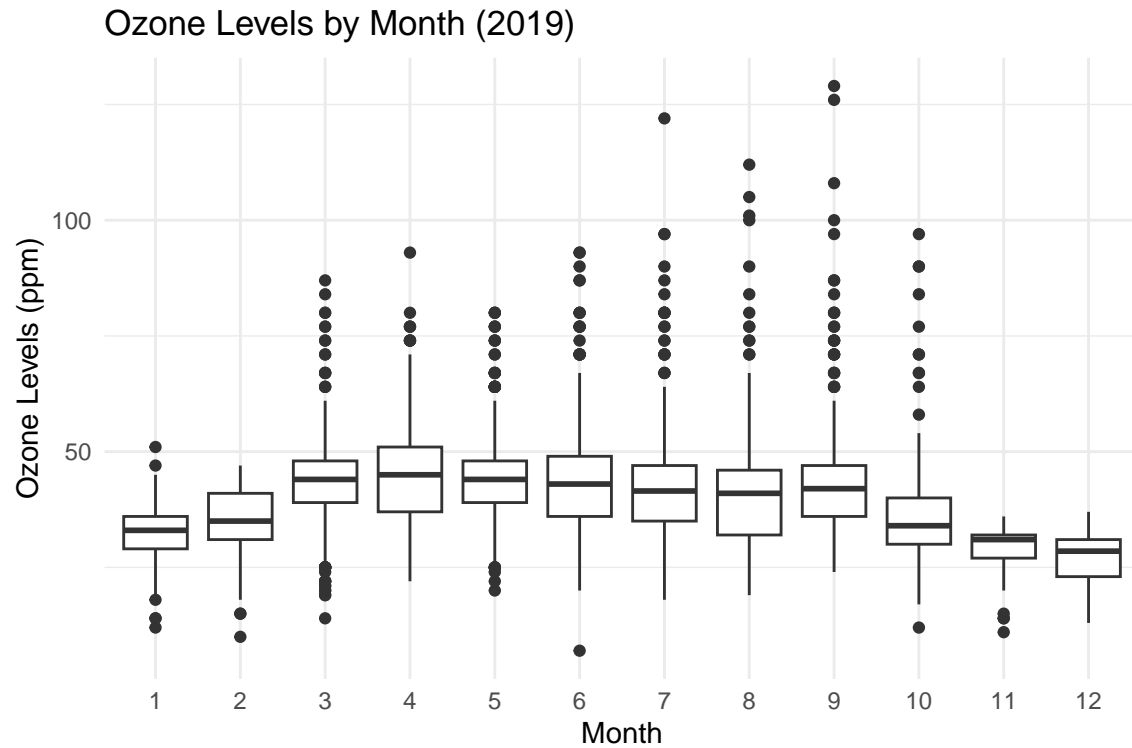Figure 1: Ozone levels by month in 2018

Figure 2: Ozone levels by month in 2019

Task 5: Knit and submit.

Add a table of contents and list of figures to your document and knit to a PDF. Submit your PDF to Sakai, but also be sure to commit and push your Rmd file used to create this knit document to GitHub. In the section below, add a link to your GitHub repository.

Git Repository