Problem Set 2

Jacob Álvarez

```r
knitr::opts_chunk$set(
  echo = TRUE,
  eval = FALSE
)
```

Answer the questions below to the best of your ability. Write clearly, and format your tables and visuals appropriately. You must use `R Markdown` to compose and compile your work. For full credit, `echo` all code chunks, and include your `setup` chunk. Submit your work in hard copy at the beginning of class.

You need the Global Greenspace Indicator Data for this assignment. Review the `README.txt` file for information about the data, variables, etc.

1. Show me that you're all set on GitHub. Create a public repository named `PSet2`. Clone it, include all your project files (including your `.Rmd` and `.pdf` files) for the work below, and commit/push your work to your repository. Include the link to your repo as your answer to this question.

https://github.com/ja2905a/PSet2/tree/main

2. The script below doesn't work. Type the corrected code chunk into your problem set. Annotate any line you correct to note your fix (i.e. `# unbalanced parentheses`). *There are more than five errors.*

```r
knitr::opts_chunk$set(
  echo = TRUE,
  eval = FALSE
)

#uncorrected

library(tidyverse)

# open my data
```

```r
  gspace = read_csv(greenspace_data_share.csv)

# summarize average urban greenspace by region
  table =
    gspace |>
    group_by(Major_Geo_Region)
    summarise(
      obs = n()
      avg = mean(annual_avg2020),
      weighted avg = mean(annual_weight_avg_2020)
    )

# output as table
  kable(gspace, digits = 1)

#corrected

library(tidyverse)

# open my data
gspace = read_csv('greenspace_data_share.csv') #added quotes around file name

# summarize average urban greenspace by region
table =
  gspace |>
  group_by(Major_Geo_Region) |> #added pipe operator
summarise(
  obs = n(), #added comma
  avg = mean(annual_avg_2020), #added underscore in annual_avg_2020
  weighted_avg = mean(annual_weight_avg_2020) #added underscore in weighted_avg
)

# output as table
knitr::kable(table) #performed function on table instead of gspace
#pulled kable function from unloaded knitr package
```

3. How many urban areas does the greenspace data cover?

4. In a couple of sentences and with reference to a well-formatted tabulation, describe the greenspace classification scores for urban areas in 2021.

5. Report the number of urban areas that satisfy the conditions below. Either write your code inline or echo the code that generated the answer.

   a. Scored `High` or above for greenspace in 2015.

    b. Scored `Exceptionally Low` at any point in the years covered.

    c. Urban areas in arid climate that became greener (as measured by annual weighted average) from 2010 to 2020.

6. How many urban areas became less green (measured by annual average) from 2010 to 2021? Were these changes concentrated in a particular geographic or climate region? Explain (with evidence, of course).

7. Present a histogram showing the change in greenspace (annual average) from 2010 to 2021. Note that you will need to create a new variable equal to this difference.

8. Present a scatter plot of population weighted greenspace in 2021 over the greenspace in 2010.

BONUS OPPORTUNITY: Use color-coding to differentiate urban areas that added versus lost greenspace in that time. Then include a 45-degree line to further highlight the change.