

Problem Set 2

Colby Pastel

2024-01-29

1. <https://github.com/Colby-P/PSet2.git>
2. See the corrected code chunk below. A total of 8 changes were made.

```
library(tidyverse)

# open my data
gspace = read_csv('greenspace_data_share.csv') # missing quotation marks (1)

# summarize average urban greenspace by region
tab1 = # Renamed "table" to "tab1" to prevent confusion with tidyverse command (2)
  gspace |>
  group_by(Major_Geo_Region) |> # added a pipe (3)
  summarise(
    obs = n(), # added a comma (4)
    avg = mean(annual_avg_2020), # added an underscore (5)
    weighted_avg = mean(annual_weight_avg_2020) # changed a space to an underscore (6)
  )

# output as table
knitr::kable(tab1, digits = 1) # added "knitr::" (7), changed "gspace" to "tab1" (8)
```

Major_Geo_Region	obs	avg	weighted_avg
Africa	154	0.3	0.2
Asia	569	0.3	0.3
Europe	128	0.3	0.3
Latin America and the Caribbean	120	0.3	0.3
Northern America	58	0.3	0.3
Oceania	9	0.3	0.3

3. The greenspace data covers **1,038** urban areas.

```
dim(gspace)
```

```
## [1] 1038 27
```

```
#str(gspace)
#unique(gspace$City)
```

4. The mean average greenspace for urban areas in 2021 was 0.280 and the average greenspace ranged from 0.045 to 0.630. The median and mean average greenspace were identical, 0.280.

```
summary(gspace$annual_avg_2021, digits = 2)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.     NA's  
##    0.045  0.230   0.280   0.280  0.340   0.630         1
```

- 5.a. The number of urban areas which Scored High or above for greenspace in 2015 was **66**.

```
gspace |>  
  filter(  
    indicator_2015 == 'High' |  
    indicator_2015 == 'Very High'  
  ) |>  
  count()
```

```
## # A tibble: 1 x 1  
##       n  
##   <int>  
## 1    66
```

- 5.b. The number of urban areas which Scored Exceptionally Low at any point in the years covered was **240**.

```
gspace |>  
  filter(  
    indicator_2010 == 'Exceptionally Low' |  
    indicator_2015 == 'Exceptionally Low' |  
    indicator_2020 == 'Exceptionally Low' |  
    indicator_2021 == 'Exceptionally Low',  
  ) |>  
  count()
```

```
## # A tibble: 1 x 1  
##       n  
##   <int>  
## 1   240
```

- 5.c. The number of urban areas with arid climates that became greener from 2010 to 2020 was **225**.

```
gspace |>  
  filter(  
    Climate_region == 'Arid',  
    annual_weight_avg_2020 > annual_weight_avg_2010  
  ) |>  
  count()
```

```
## # A tibble: 1 x 1  
##       n  
##   <int>  
## 1   225
```

6. **128** urban areas became less green from 2010 to 2021. These changes were concentrated in Europe and Asia, and urban areas with climates besides arid and polar.

```
gspace |>
  filter(
    annual_avg_2010 > annual_avg_2021
  ) |>
  count()
```

```
## # A tibble: 1 x 1
##       n
##   <int>
## 1   128
```

```
tab2 =
  gspace |>
    filter(
      annual_avg_2010 > annual_avg_2021
    )

tab2 |>
  count(Major_Geo_Region, name = 'Freq')
```

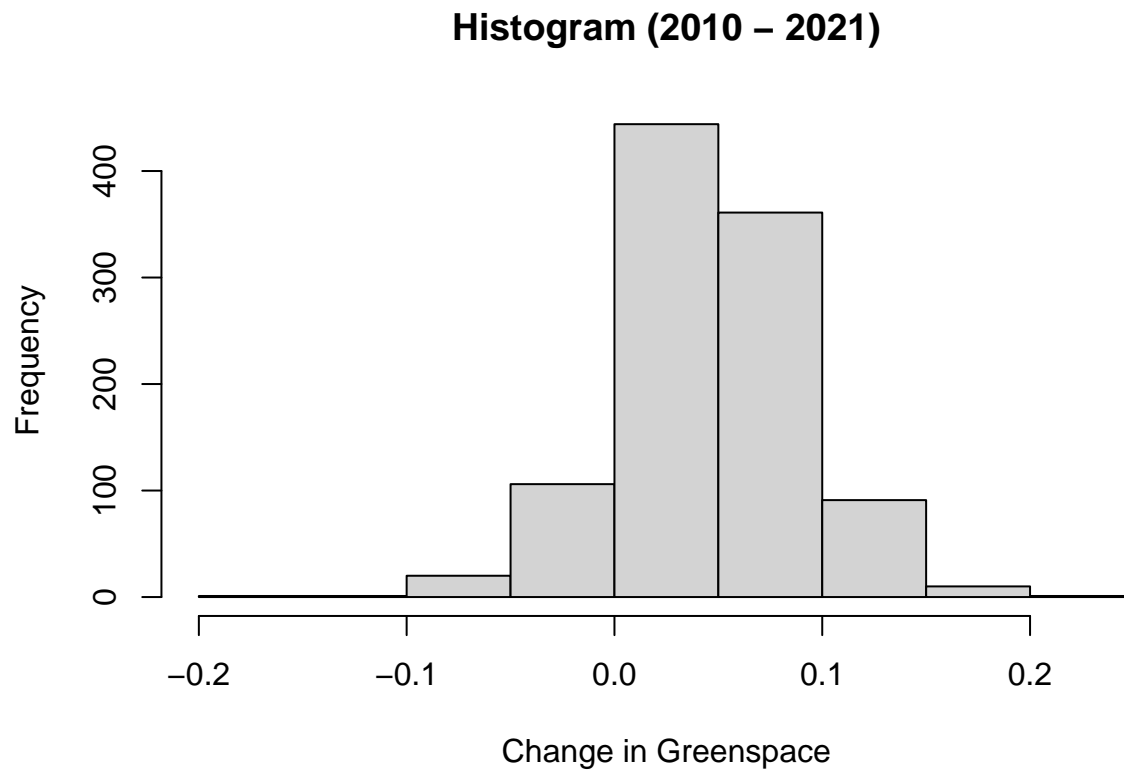
```
## # A tibble: 5 x 2
##   Major_Geo_Region      Freq
##   <chr>             <int>
## 1 Africa              25
## 2 Asia                 35
## 3 Europe               47
## 4 Latin America and the Caribbean 12
## 5 Northern America     9
```

```
tab2 |>
  count(Climate_region, name = 'Freq')
```

```
## # A tibble: 5 x 2
##   Climate_region      Freq
##   <chr>             <int>
## 1 Arid                12
## 2 Continental         36
## 3 Polar                 1
## 4 Temperate           45
## 5 Tropical            34
```

7. See histogram below.

```
tab3 =  
  gspace |>  
    summarize(  
      gspace_2010 = annual_avg_2010,  
      gspace_2021 = annual_avg_2021,  
      Diff = annual_avg_2021 - annual_avg_2010  
    )  
  
hist(tab3$Diff,  
     main = 'Histogram (2010 - 2021)',  
     xlab = 'Change in Greenspace')
```



8. See scatter plot below.

```
plot(gspace_2021 ~ gspace_2010, tab3,  
     main = 'Population Weighted Greenspace',  
     xlab = 'Avg 2010',  
     ylab = 'Avg 2021')
```



Bonus Opportunity: See scatter plot below.

```
plot(gspace_2021 ~ gspace_2010, tab3,  
     col = ifelse(Diff > 0, 'darkgreen', 'black'),  
     main = 'Population Weighted Greenspace',  
     xlab = 'Avg 2010', ylab = 'Avg 2021')  
abline(0,1, col = 'red')
```

