EV Power - Lab 4 Project Report

Example Solution 1

Part 0: libraries

```
library(tidyverse)
```

```
── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0
──
✔ dplyr     1.1.4     ✔ readr     2.1.5
✔ forcats   1.0.1     ✔ stringr   1.5.2
✔ ggplot2   4.0.0     ✔ tibble    3.3.0
✔ lubridate 1.9.4     ✔ tidyr     1.3.1
✔ purrr     1.1.0
── Conflicts ────────────────────────────────────────── tidyverse_conflicts()
──
✖ dplyr::filter() masks stats::filter()
✖ dplyr::lag()    masks stats::lag()
ℹ Use the conflicted package (<http://conflicted.r-lib.org/>) to force all
conflicts to become errors
```

```
library(readr)
library(sf)
```

```
Linking to GEOS 3.13.1, GDAL 3.11.0, PROJ 9.6.0; sf_use_s2() is TRUE
```

```
library(tigris)
```

```
To enable caching of data, set `options(tigris_use_cache = TRUE)`
in your R script or .Rprofile.
```

```
library(dplyr)
library(ggplot2)

renew_2021 <- read.csv("data/renew-use-2021.csv")
renew_2022 <- read.csv("data/renew-use-2022.csv")
renew_2023 <- read.csv("data/renew-use-2023.csv")
```

```r
total_2021 <- read.csv("data/total-use-2021.csv")
total_2022 <- read.csv("data/total-use-2022.csv")
total_2023 <- read.csv("data/total-use-2023.csv")

av_energy_2021_2023 <- read.csv("data/av-energy-price-2021-2023.csv")
ev_registration_2021_2023 <- read.csv("data/ev-registrations-by-
state-2023.csv")
```

Part 1: Defining Research Question

Chosen Question: What is the share of electricity that comes from clean sources by state, and how has that proportion changed over time?

Part 2: Data Preparation and Cleaning

```r
new_names <- c(
    "coal_use",
    "natural_gas",
    "petroleum",
    "nuclear_energy",
    "total_renewable"
)

set_names <- function(df) {
  df$Energy_Source <- new_names
  return(df)
}

total_2022_clean <- set_names(total_2022)
total_2023_clean <- set_names(total_2023)

summary_generator <- function(df) {
  df_long <- df |>
    pivot_longer(
      cols = -Energy_Source,
      names_to = "State",
      values_to = "Usage"
    )
  nonrenewable_summary <- df_long |>
    filter(Energy_Source != "total_renewable") |>
    group_by(State) |>
    summarise(Usage = sum(Usage, na.rm = TRUE), .groups = 'drop') |>
    mutate(Energy_Source = "total_nonrenewable")
  df_with_totals <- bind_rows(df_long, nonrenewable_summary)
  total_energy_summary <- df_with_totals |>
    filter(Energy_Source %in% c("total_renewable", "total_nonrenewable")) |>
    group_by(State) |>
    summarise(Total_Energy = sum(Usage, na.rm = TRUE), .groups = 'drop')
```

```
  proportion_summary <- df_with_totals |>
    filter(Energy_Source %in% c("total_renewable", "total_nonrenewable")) |>
    inner_join(total_energy_summary, by = "State") |>
    mutate(
      Usage = Usage / Total_Energy,
      Energy_Source = paste0("prop_", Energy_Source)
    ) |>
    select(-Total_Energy)

  return(
    bind_rows(df_with_totals, proportion_summary) %>%
      pivot_wider(
        names_from = State,
        values_from = Usage
      )
  )
}

total_2022_final <- summary_generator(total_2022_clean)
total_2023_final <- summary_generator(total_2023_clean)
total_2022_final
```

```
# A tibble: 8 × 53
  Energy_Source       AK      AL      AR      AZ      CA      CO      CT
DC
  <chr>            <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
<dbl>
1 coal_use       1.86e+4 2.98e+5 2.12e+5 1.54e+5 3.00e+4 2.33e+5 0
0
2 natural_gas    4.38e+5 7.87e+5 3.98e+5 4.68e+5 2.13e+6 5.25e+5 3.07e+5
3.02e+4
3 petroleum      2.63e+5 5.78e+5 3.28e+5 5.95e+5 3.02e+6 5.38e+5 3.03e+5
1.80e+4
4 nuclear_energy 0       4.42e+5 1.50e+5 3.34e+5 1.84e+5 0       1.72e+5
0
5 total_renewab… 1.04e+4 2.32e+5 9.08e+4 1.01e+5 8.81e+5 1.15e+5 4.91e+4
2.62e+3
6 total_nonrene… 7.20e+5 2.11e+6 1.09e+6 1.55e+6 5.36e+6 1.30e+6 7.82e+5
4.82e+4
7 prop_total_re… 1.43e-2 9.93e-2 7.71e-2 6.13e-2 1.41e-1 8.14e-2 5.91e-2
5.16e-2
8 prop_total_no… 9.86e-1 9.01e-1 9.23e-1 9.39e-1 8.59e-1 9.19e-1 9.41e-1
9.48e-1
# ℹ 44 more variables: DE <dbl>, FL <dbl>, GA <dbl>, HI <dbl>, IA <dbl>,
#   ID <dbl>, IL <dbl>, IN <dbl>, KS <dbl>, KY <dbl>, LA <dbl>, MA <dbl>,
#   MD <dbl>, ME <dbl>, MI <dbl>, MN <dbl>, MO <dbl>, MS <dbl>, MT <dbl>,
```

```
#   NC <dbl>, ND <dbl>, NE <dbl>, NH <dbl>, NJ <dbl>, NM <dbl>, NV <dbl>,
#   NY <dbl>, OH <dbl>, OK <dbl>, OR <dbl>, PA <dbl>, RI <dbl>, SC <dbl>,
#   SD <dbl>, TN <dbl>, TX <dbl>, UT <dbl>, VA <dbl>, VT <dbl>, WA <dbl>,
#   WI <dbl>, WV <dbl>, WY <dbl>, US <dbl>
```

```
total_2023_final
```

```
# A tibble: 8 × 53
  Energy_Source       AK      AL      AR      AZ      CA      CO      CT
DC
  <chr>            <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>   <dbl>
<dbl>
1 coal_use       1.84e+4 2.25e+5 1.80e+5 1.38e+5 2.87e+4 2.05e+5 0
0
2 natural_gas    4.48e+5 7.76e+5 4.00e+5 5.37e+5 2.15e+6 5.25e+5 3.05e+5
2.62e+4
3 petroleum      2.70e+5 5.66e+5 3.27e+5 6.00e+5 3.00e+6 5.14e+5 2.93e+5
1.73e+4
4 nuclear_energy 0       4.76e+5 1.56e+5 3.29e+5 1.85e+5 0       1.43e+5
0
5 total_renewab… 1.01e+4 2.22e+5 8.73e+4 1.08e+5 1.07e+6 1.15e+5 4.90e+4
2.79e+3
6 total_nonrene… 7.37e+5 2.04e+6 1.06e+6 1.60e+6 5.36e+6 1.24e+6 7.41e+5
4.35e+4
7 prop_total_re… 1.35e-2 9.81e-2 7.58e-2 6.33e-2 1.66e-1 8.46e-2 6.20e-2
6.03e-2
8 prop_total_no… 9.86e-1 9.02e-1 9.24e-1 9.37e-1 8.34e-1 9.15e-1 9.38e-1
9.40e-1
# ℹ 44 more variables: DE <dbl>, FL <dbl>, GA <dbl>, HI <dbl>, IA <dbl>,
#   ID <dbl>, IL <dbl>, IN <dbl>, KS <dbl>, KY <dbl>, LA <dbl>, MA <dbl>,
#   MD <dbl>, ME <dbl>, MI <dbl>, MN <dbl>, MO <dbl>, MS <dbl>, MT <dbl>,
#   NC <dbl>, ND <dbl>, NE <dbl>, NH <dbl>, NJ <dbl>, NM <dbl>, NV <dbl>,
#   NY <dbl>, OH <dbl>, OK <dbl>, OR <dbl>, PA <dbl>, RI <dbl>, SC <dbl>,
#   SD <dbl>, TN <dbl>, TX <dbl>, UT <dbl>, VA <dbl>, VT <dbl>, WA <dbl>,
#   WI <dbl>, WV <dbl>, WY <dbl>, US <dbl>
```

Here I just wanted to make some useful summary statistics to work with that would potentially give me insight on what I could do next. I specifically chose to work with the data from 2022 and 2023 as I felt their more recent information would provide a better insight than the 2021 data and so I saw that one as less important.

Part 3: Joining / Pivoting Datasets for Analysis

```
long_2022 <- total_2022_final |>
  pivot_longer(cols = -Energy_Source, names_to = "State", values_to =
"Usage_2022")
```

```
long_2023 <- total_2023_final |>
  pivot_longer(cols = -Energy_Source, names_to = "State", values_to =
"Usage_2023")
comparison_data <- inner_join(
  long_2022,
  long_2023,
  by = c("Energy_Source", "State")
) |>
  mutate(Change_2022_2023 = Usage_2023 - Usage_2022) |>
  select(Energy_Source, State, Usage_2022, Usage_2023, Change_2022_2023)
comparison_data
```

```
# A tibble: 416 × 5
   Energy_Source State Usage_2022 Usage_2023 Change_2022_2023
   <chr>         <chr>      <dbl>      <dbl>            <dbl>
 1 coal_use      AK         18615      18414             -201
 2 coal_use      AL        297654     224926           -72728
 3 coal_use      AR        211724     180262           -31462
 4 coal_use      AZ        154007     137885           -16122
 5 coal_use      CA         30049      28746            -1303
 6 coal_use      CO        233256     204826           -28430
 7 coal_use      CT             0          0                0
 8 coal_use      DC             0          0                0
 9 coal_use      DE          1846        338            -1508
10 coal_use      FL        171953     129387           -42566
# ℹ 406 more rows
```

The way the data was set up was pretty difficult to work with so I made sure to make the dataset longer rather than wider so I could more easily use the information gathered in the mapping step.

Part 4: Mapping Visualization

```
states_sf <- states(cb = TRUE) |>
  shift_geometry()
```

```
Retrieving data for the year 2024
```

```


|
  |                                                                      |
0%

|
  |=                                                                     |
```

```
1%

|
  |=                                                                |
2%

|
  |==                                                               |
2%

|
  |==                                                               |
3%

|
  |===                                                              |
4%

|
  |===                                                              |
5%

|
  |====                                                             |
5%

|
  |====                                                             |
6%

|
  |=====                                                            |
6%

|
  |=====                                                            |
7%

|
  |=====                                                            |
8%

|
  |======                                                           |
8%

|
  |======                                                           |
```

```
9%

|
  |=======                                                           |
10%

|
  |=======                                                           |
11%

|
  |========                                                          |
11%

|
  |========                                                          |
12%

|
  |=========                                                         |
13%

|
  |==========                                                        |
14%

|
  |==========                                                        |
15%

|
  |===========                                                       |
15%

|
  |===========                                                       |
16%

|
  |============                                                      |
17%

|
  |============                                                      |
18%

|
  |=============                                                     |
```

```
18%

|
  |=============                                                   |
19%

|
  |==============                                                  |
20%

|
  |==============                                                  |
21%

|
  |===============                                                 |
21%

|
  |===============                                                 |
22%

|
  |================                                                |
22%

|
  |================                                                |
23%

|
  |=================                                               |
24%

|
  |=================                                               |
25%

|
  |==================                                              |
25%

|
  |==================                                              |
26%

|
  |===================                                             |
```

```
27%

|
  |==================                                                    |
28%

|
  |==================                                                    |
28%

|
  |===================                                                   |
29%

|
  |====================                                                  |
30%

|
  |=====================                                                 |
31%

|
  |=====================                                                 |
32%

|
  |======================                                                |
32%

|
  |======================                                                |
33%

|
  |=======================                                               |
34%

|
  |=======================                                               |
35%

|
  |========================                                              |
35%

|
  |========================                                              |
```

```
36%

|
  |=========================                                              |
37%

|
  |=========================                                              |
38%

|
  |=========================                                              |
38%

|
  |=========================                                              |
39%

|
  |=========================                                              |
39%

|
  |=========================                                              |
40%

|
  |=========================                                              |
41%

|
  |=========================                                              |
42%

|
  |=========================                                              |
42%

|
  |=========================                                              |
43%

|
  |=========================                                              |
44%

|
  |=========================                                              |
```

```
45%

|
  |================================                                 |
45%

|
  |================================                                 |
46%

|
  |================================                                 |
46%

|
  |================================                                 |
47%

|
  |================================                                 |
48%

|
  |================================                                 |
49%

|
  |================================                                 |
49%

|
  |================================                                 |
50%

|
  |================================                                 |
51%

|
  |================================                                 |
52%

|
  |================================                                 |
52%

|
  |================================                                 |
```

```
53%

|
  |=====================================                                |
54%

|
  |=====================================                                |
54%

|
  |=====================================                                |
55%

|
  |=====================================                                |
55%

|
  |=====================================                                |
56%

|
  |=====================================                                |
57%

|
  |=====================================                                |
58%

|
  |=====================================                                |
58%

|
  |=====================================                                |
59%

|
  |=====================================                                |
60%

|
  |=====================================                                |
61%

|
  |=====================================                                |
```

```
61%

|
  |=========================================                            |
62%

|
  |=========================================                            |
63%

|
  |=========================================                            |
64%

|
  |=========================================                            |
65%

|
  |=========================================                            |
65%

|
  |=========================================                            |
66%

|
  |=========================================                            |
67%

|
  |=========================================                            |
68%

|
  |=========================================                            |
68%

|
  |=========================================                            |
69%

|
  |=========================================                            |
70%

|
  |=========================================                            |
```

```
71%

|
  |================================================                 |
71%

|
  |================================================                 |
72%

|
  |================================================                 |
72%

|
  |================================================                 |
73%

|
  |================================================                 |
74%

|
  |================================================                 |
75%

|
  |================================================                 |
75%

|
  |================================================                 |
76%

|
  |================================================                 |
77%

|
  |================================================                 |
78%

|
  |================================================                 |
78%

|
  |================================================                 |
```

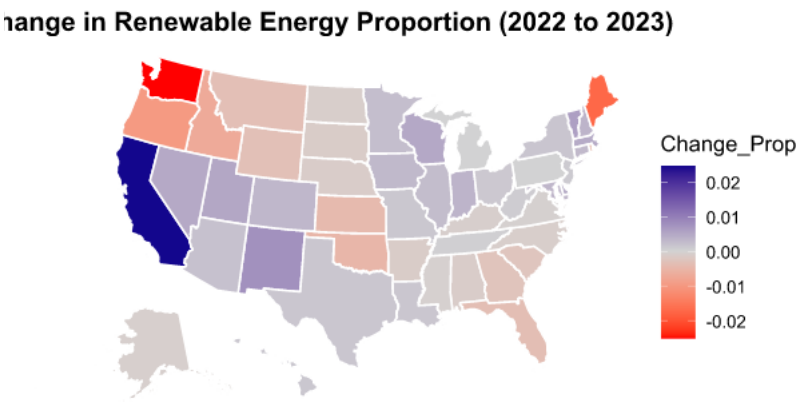
```
79%

|
  |=======================================================               |
80%

|
  |========================================================              |
81%

|
  |========================================================              |
82%

|
  |========================================================              |
82%

|
  |========================================================              |
83%

|
  |========================================================              |
84%

|
  |========================================================              |
85%

|
  |========================================================              |
85%

|
  |========================================================              |
86%

|
  |========================================================              |
87%

|
  |========================================================              |
88%

|
  |========================================================              |
```

```
88%

|
  |============================================================        |
89%

|
  |============================================================       |
89%

|
  |============================================================       |
90%

|
  |============================================================      |
91%

|
  |============================================================      |
92%

|
  |============================================================     |
92%

|
  |============================================================     |
93%

|
  |============================================================    |
94%

|
  |============================================================    |
95%

|
  |============================================================   |
95%

|
  |============================================================   |
96%

|
  |============================================================  |
```

```
97%

|
  |======================================================================== |
98%

|
  |======================================================================== |
99%

|
  |========================================================================|
99%

|
  |========================================================================|
100%
```

```
map_change_data <- comparison_data |>
  filter(Energy_Source == "prop_total_renewable") |>
  select(State, Change_Prop = Change_2022_2023)

map_data_final_change <- states_sf |>
  filter(!(NAME %in% c("Puerto Rico", "United States Virgin Islands", "Guam",
"Commonwealth of the Northern Mariana Islands", "American Samoa"))) |>
  left_join(map_change_data, by = c("STUSPS" = "State"))

change_map <- ggplot(data = map_data_final_change) +
  geom_sf(
    aes(fill = Change_Prop),
    color = "white",
    linewidth = 0.5
  ) +
  scale_fill_gradient2( #See if you can make it so darker means more intense
    low = "red",
    mid = "lightgrey",
    high = "darkblue",
    midpoint = 0
  ) +

  labs(
    title ="Change in Renewable Energy Proportion (2022 to 2023)"
  ) +
  theme_void() +
  theme(
    legend.position = "right",
    plot.title = element_text(hjust = 0.5, face = "bold")
```

```
  )

change_map
```


hange in Renewable Energy Proportion (2022 to 2023)
Change_Prop
0.02
0.01
0.00
-0.01
-0.02