EV Power - Lab 4 Project Report

Example Solution 1

Part 0: libraries

```
options(repos = c(CRAN = "https://cloud.r-project.org"))
install.packages("tigris")
```

```
The downloaded binary packages are in
    /var/folders/8v/5zpy9md96lq5nf5dzw8jmktc0000gn/T//Rtmpj1YOF5/downloaded_packages
```

```
install.packages("viridis")
```

```
The downloaded binary packages are in
    /var/folders/8v/5zpy9md96lq5nf5dzw8jmktc0000gn/T//Rtmpj1YOF5/downloaded_packages
```

```
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   1.5.2
v ggplot2   4.0.0     v tibble    3.3.0
v lubridate 1.9.4     v tidyr     1.3.1
v purrr     1.1.0
-- 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 becor
```

```
library(ggplot2)
library(readr)
library(sf)
```

```
Linking to GEOS 3.13.0, GDAL 3.8.5, PROJ 9.5.1; 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(viridis)
```

```
Loading required package: viridisLite
```

```
setwd("/Users/vivianlau/Documents/stat133/project4/ev-power-viviaanlau-creator")
renew_2021 <- read_csv("data/renew-use-2021.csv")
```

```
Rows: 260 Columns: 3
-- Column specification --------------------------------------------------------
Delimiter: ","
chr (3): State, Energy_Source, Renewable_Use_2021

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
renew_2022 <- read_csv("data/renew-use-2022.csv")
```

```
Rows: 260 Columns: 3
-- Column specification --------------------------------------------------------
Delimiter: ","
chr (3): State, Energy_Source, Renewable_Use_2022

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
renew_2023 <- read_csv("data/renew-use-2023.csv")
```

```
Rows: 260 Columns: 3
-- Column specification --------------------------------------------------------
Delimiter: ","
chr (3): State, Energy_Source, Renewable_Use_2023

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
avg_price <- read_csv("data/av-energy-price-2021-2023.csv")
```

```
Rows: 54 Columns: 1
-- Column specification --------------------------------------------------------
Delimiter: ","
chr (1): Total energy average price, dollars per million Btu,,,

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
total_2021 <- read_csv("data/total-use-2021.csv")
```

```
Rows: 5 Columns: 53
-- Column specification --------------------------------------------------------
Delimiter: ","
chr  (1): Energy_Source
dbl (52): AK, AL, AR, AZ, CA, CO, CT, DC, DE, FL, GA, HI, IA, ID, IL, IN, KS...

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
total_2022 <- read_csv("data/total-use-2022.csv")
```

```
Rows: 5 Columns: 53
-- Column specification --------------------------------------------------------
Delimiter: ","
chr  (1): Energy_Source
dbl (52): AK, AL, AR, AZ, CA, CO, CT, DC, DE, FL, GA, HI, IA, ID, IL, IN, KS...

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
total_2023 <- read_csv("data/total-use-2023.csv")
```

```
Rows: 5 Columns: 53
-- Column specification --------------------------------------------------------
Delimiter: ","
chr  (1): Energy_Source
dbl (52): AK, AL, AR, AZ, CA, CO, CT, DC, DE, FL, GA, HI, IA, ID, IL, IN, KS...

i Use `spec()` to retrieve the full column specification for this data.
i Specify the column types or set `show_col_types = FALSE` to quiet this message.
```

```
registration <- read_csv("data/ev-registrations-by-state-2023.csv")
```

```
New names:
Rows: 54 Columns: 2
-- Column specification
-------------------------------------------------------- Delimiter: "," chr
(2): electric vehicle registrations_by_state (2023), ...2
i Use `spec()` to retrieve the full column specification for this data. i
Specify the column types or set `show_col_types = FALSE` to quiet this message.
* `` -> `...2`
```

Part 1: Defining Research Question

Chosen Question: Are there higher EV registrations in states with more renewable energy use? Before analyzing, I believe that states that get most energy from renewable sources wll have more EV registrations because they have higher access, making it more widespread.

Part 2: Data Preparation and Cleaning

```
renew_2021 <- renew_2021 |> rename(state = State, renewable_use = Renewable_Use_2021)
renew_2022 <- renew_2022 |> rename(state = State, renewable_use = Renewable_Use_2022)
renew_2023 <- renew_2023 |> rename(state = State, renewable_use = Renewable_Use_2023)

renew_2021$year <- 2021
renew_2022$year <- 2022
renew_2023$year <- 2023
```

```r
renew_all <- bind_rows(renew_2021, renew_2022, renew_2023)
renew_all$state <- str_to_title(renew_all$state)

total21_longer <- total_2021 |>
  pivot_longer(
    cols = -Energy_Source,
    names_to = "state",
    values_to = "energy_use"
  ) |>
  group_by(state) |>
  summarize(total_energy_use = sum(energy_use, na.rm = TRUE)) |>
  mutate(year = 2021)

total22_longer <- total_2022 |>
  pivot_longer(
    cols = -Energy_Source,
    names_to = "state",
    values_to = "energy_use"
  ) |>
  group_by(state) |>
  summarize(total_energy_use = sum(energy_use, na.rm = TRUE)) |>
  mutate(year = 2022)

total23_longer <- total_2023 |>
  pivot_longer(
    cols = -Energy_Source,
    names_to = "state",
    values_to = "energy_use"
  ) |>
  group_by(state) |>
  summarize(total_energy_use = sum(energy_use, na.rm = TRUE)) |>
  mutate(year = 2023)

total_all <- bind_rows(total21_longer, total22_longer, total23_longer)
total_all$state <- str_to_title(total_all$state)

ev <- registration |>
  rename(
    state = `electric vehicle registrations_by_state (2023)`,
    ev_registrations = `...2`
  )
```

```r
ev$state <- str_to_title(ev$state)


head(renew_all)
```

```
# A tibble: 6 x 4
  state Energy_Source renewable_use  year
  <chr> <chr>         <chr>         <dbl>
1 Ak    Biomass        3153          2021
2 Ak    Geothermal    186 MMBtu      2021
3 Ak    Hydropower    5763 about     2021
4 Ak    Solar Energy  ~45            2021
5 Ak    Wind Energy   451 USD        2021
6 Al    Biomass       198543 est.    2021
```

```r
head(total_all)
```

```
# A tibble: 6 x 3
  state total_energy_use  year
  <chr>            <dbl> <dbl>
1 Ak              684975  2021
2 Al             2352656  2021
3 Ar             1136025  2021
4 Az             1681257  2021
5 Ca             6142252  2021
6 Co             1364155  2021
```

```r
head(ev)
```

```
# A tibble: 6 x 2
  state    ev_registrations
  <chr>    <chr>
1 <NA>     <NA>
2 State    Count-EVs
3 Alabama  #13047
4 Alaska   ~2697
5 Arizona  89798
6 Arkansas 7108 EVs
```

Part 3: Joining / Pivoting Datasets for Analysis

```
state_find <- tibble(
  state = str_to_title(state.abb),
  full_state = state.name
)

energy_combined <- renew_all |>
  left_join(total_all, by = c("state", "year")) |>
  mutate(
    renewable_use = as.numeric(gsub("[^0-9.]", "", renewable_use)),
    total_energy_use = as.numeric(gsub("[^0-9.]", "", total_energy_use)),
    pct_renew = renewable_use / total_energy_use * 100
  ) |>
  left_join(state_find, by = "state") |>
  mutate(state = full_state) |>
  select(-full_state)

ev_energy_2023 <- energy_combined |>
  filter(year == 2023) |>
  left_join(ev, by = "state") |>
  mutate(
    ev_registrations = as.numeric(gsub("[^0-9.]", "", ev_registrations)),
    ev_per_energy = ev_registrations / total_energy_use
  )

head(energy_combined)
```

```
# A tibble: 6 x 6
  state   Energy_Source renewable_use  year total_energy_use pct_renew
  <chr>   <chr>                 <dbl> <dbl>            <dbl>     <dbl>
1 Alaska  Biomass                3153  2021           684975   0.460
2 Alaska  Geothermal              186  2021           684975   0.0272
3 Alaska  Hydropower             5763  2021           684975   0.841
4 Alaska  Solar Energy             45  2021           684975   0.00657
5 Alaska  Wind Energy             451  2021           684975   0.0658
6 Alabama Biomass              198543  2021          2352656   8.44
```

```
head(ev_energy_2023)
```

```
# A tibble: 6 x 8
```

```
  state   Energy_Source renewable_use  year total_energy_use pct_renew
  <chr>   <chr>                 <dbl> <dbl>            <dbl>     <dbl>
1 Alaska  Biomass                3404  2023           746979   0.456
2 Alaska  Geothermal              186  2023           746979   0.0249
3 Alaska  Hydropower             6051  2023           746979   0.810
4 Alaska  Solar Energy             67  2023           746979   0.00897
5 Alaska  Wind Energy             380  2023           746979   0.0509
6 Alabama Biomass              189040  2023          2265008   8.35
# i 2 more variables: ev_registrations <dbl>, ev_per_energy <dbl>
```

```
# Check
sum(!is.na(ev_energy_2023$ev_registrations))
```

```
[1] 250
```

```
# Summary statistics
summary(ev_energy_2023$pct_renew)
```

```
    Min.  1st Qu.   Median     Mean  3rd Qu.     Max.
 0.00000  0.06978  0.45138  2.05981  2.92614 21.43650
```

```
summary(ev_energy_2023$ev_registrations)
```

```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.    NA's
    959    8150   25833   70948   72139 1256646      10
```

```
cor(ev_energy_2023$pct_renew, ev_energy_2023$ev_registrations, use = "complete.obs")
```

```
[1] 0.02980727
```

Part 4: Mapping Visualization

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

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

```
|
|                                                                      |   0%
|
|                                                                      |   1%
|
|=                                                                     |   1%
|
|=                                                                     |   2%
|
|==                                                                    |   2%
|
|==                                                                    |   3%
|
|===                                                                   |   4%
|
|===                                                                   |   5%
|
|====                                                                  |   5%
|
|====                                                                  |   6%
|
|=====                                                                 |   7%
|
|=====                                                                 |   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%
|
|===================================                              |  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%
```

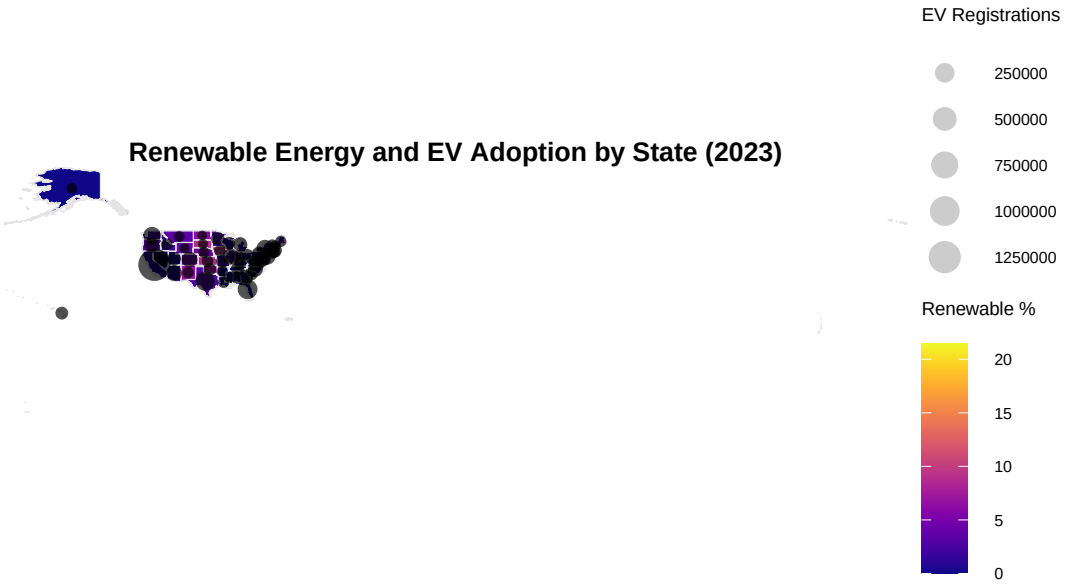
```r
map_df <- left_join(states_sf, ev_energy_2023, by = c("NAME" = "state"))

centers <- st_centroid(map_df)
```

```
Warning: st_centroid assumes attributes are constant over geometries
```

```r
ggplot() +
  geom_sf(data = map_df, aes(fill = pct_renew), color = "grey90", size = 0.2) +
  geom_sf(data = centers, aes(size = ev_registrations), color = "black", alpha = 0.2) +
  scale_fill_viridis(option = "plasma", name = "Renewable %") +
  scale_size(range = c(1, 5), name = "EV Registrations") +
  labs(title = "Renewable Energy and EV Adoption by State (2023)") +
  coord_sf(expand = FALSE) +
  theme_void(base_size = 20) +
  theme(
    plot.title = element_text(size = 10, face = "bold", hjust = 0.5),
    legend.title = element_text(size = 7),
    legend.text = element_text(size = 6),
    legend.position = "right"
  )
```

```
Warning: Removed 6 rows containing missing values or values outside the scale range
(`geom_sf()`).
```




```
##Part 5:
#Based off the map, the first thing I noticed is that more states on the coast have higher E
```