

# EV Power - Lab 4 Project Report

## Example Solution 1

### Part 0: libraries

```
knitr::opts_knit$set(root.dir = "E:/Users/W1421/OneDrive/ /HW")
library(tidyverse)
```

```
-- Attaching core tidyverse packages ----- tidyverse 2.0.0 --
v dplyr      1.1.4      v readr      2.1.5
v forcats    1.0.1      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 become
```

```
library(janitor)
```

Attaching package: 'janitor'

The following objects are masked from 'package:stats':

chisq.test, fisher.test

```
library(sf)
```

Linking to GEOS 3.13.1, GDAL 3.11.0, PROJ 9.6.0; sf\_use\_s2() is TRUE

```
library(ggplot2)
library(leaflet)
library(viridis)
```

Loading required package: viridisLite

## Part 1: Defining Research Question

Chosen Question: Do states with higher shares of renewable energy also have higher numbers of EV registrations?

## Part 2: Data Preparation and Cleaning

```
renew21 <- read_csv("data/renew-use-2021.csv") |> clean_names()
```

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.

```
renew22 <- read_csv("data/renew-use-2022.csv") |> clean_names()
```

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.

```
renew23 <- read_csv("data/renew-use-2023.csv") |> clean_names()
```

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.

```
total21 <- read_csv("data/total-use-2021.csv") |> clean_names()
```

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.

```
total22 <- read_csv("data/total-use-2022.csv") |> clean_names()
```

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.

```
total23 <- read_csv("data/total-use-2023.csv") |> clean_names()
```

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.

```
ev2023 <- read_csv("data/ev-registrations-by-state-2023.csv") |> clean_names()
```

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`

```
renew21 <- renew21 |> mutate(year = 2021)
renew22 <- renew22 |> mutate(year = 2022)
renew23 <- renew23 |> mutate(year = 2023)

renew_all <- bind_rows(renew21, renew22, renew23) |>
rename(renewable_use = renewable_use_2021) |>
select(state, year, renewable_use)

total21_long <- total21 |>
pivot_longer(cols = -energy_source,
names_to = "state",
values_to = "total_energy_use") |>
mutate(year = 2021)
total22_long <- total22 |>
pivot_longer(cols = -energy_source,
names_to = "state",
values_to = "total_energy_use") |>
mutate(year = 2022)
total23_long <- total23 |>
pivot_longer(cols = -energy_source,
names_to = "state",
values_to = "total_energy_use") |>
mutate(year = 2023)
```

## Part 3: Joining / Pivoting Datasets for Analysis

```
total_all <- bind_rows(total21_long, total22_long, total23_long) |>
group_by(state, year) |>
summarise(total_energy_use = sum(as.numeric(total_energy_use), na.rm = TRUE)) |>
ungroup()
```

``summarise()`` has grouped output by 'state'. You can override using the ``.groups`` argument.

```
energy_joined <- left_join(renew_all, total_all, by = c("state", "year")) |>
mutate(across(c(renewable_use, total_energy_use), as.numeric)) |>
mutate(renew_share = renewable_use / total_energy_use)
```

Warning: There was 1 warning in ``mutate()``.  
i In argument: ``across(c(renewable_use, total_energy_use), as.numeric)``.  
Caused by warning:  
! NAs introduced by coercion

```
ev2023 <- ev2023 |>
set_names(c("state", "ev_registrations")) |>
mutate(state = str_to_title(state),
ev_registrations = as.numeric(ev_registrations))
```

Warning: There was 1 warning in ``mutate()``.  
i In argument: ``ev_registrations = as.numeric(ev_registrations)``.  
Caused by warning:  
! NAs introduced by coercion

```
ev_energy <- energy_joined |>
filter(year == 2023) |>
left_join(ev2023, by = "state") |>
mutate(ev_per_energy = ev_registrations / total_energy_use)
```

## Part 4: Mapping Visualization

```
library(maps)
```

Attaching package: 'maps'

The following object is masked from 'package:viridis':

unemp

The following object is masked from 'package:purrr':

map

```
library(sf)
```

```
us_states <- st_as_sf(map("state", plot = FALSE, fill = TRUE)) |>  
  mutate(state = str_to_title(ID))
```

```
map_data_2023 <- left_join(us_states, ev_energy, by = "state")
```

```
ggplot(map_data_2023, aes(fill = renew_share)) +  
  geom_sf(color = "white", size = 0.2) +  
  scale_fill_viridis_c(option = "C", na.value = "grey90") +  
  labs(  
    title = "Renewable Energy Share by State (2023)",  
    fill = "Renewable\nShare"  
  ) +  
  theme_minimal()
```

Renewable Energy Share by State (2023)

