EV Power - Lab 4 Project Report

Part 0: libraries

```
library(tidyverse)
```

```
── Attaching core tidyverse packages ──────────────────────── tidyverse 2.0.0
──
✔ dplyr     1.1.4     ✔ readr     2.1.5
✔ forcats   1.0.0     ✔ 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(dplyr)
library(usmap)
library(janitor)
```

```

Attaching package: 'janitor'

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

    chisq.test, fisher.test
```

Part 1: Defining Research Question

Chosen Question: Are EV registrations per unit of total energy use higher in states that rely more on renewable energy?

Part 2: Data Preparation and Cleaning

```
#load my data in, and got the clean_names() function to help me clean the
data
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.
```

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

```
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.
```

```
ev <- 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`
```

Part 3: Joining / Pivoting Datasets for Analysis

```
total_long <- total23 |>
  pivot_longer(cols = -energy_source, names_to = "state_abbr", values_to =
"total_use") |>
mutate(year = 2023) |>
mutate(
    state_abbr = toupper(state_abbr),
    year = 2023,
    state = state.name[match(state_abbr, state.abb)]
  )

renew_long <- renew23|>
    pivot_longer(cols = -energy_source, names_to = "state_abbr", values_to =
"renew_use")|>
mutate(year = 2023)

data_joined <- left_join(
  renew_long,
  total_long,
  by = c("state_abbr", "year", "energy_source")
)
```

Part 4: Mapping Visualization

```
bargraph <- data_joined |>
  mutate(
    renew_use = as.numeric(renew_use),
    total_use = as.numeric(total_use),
```

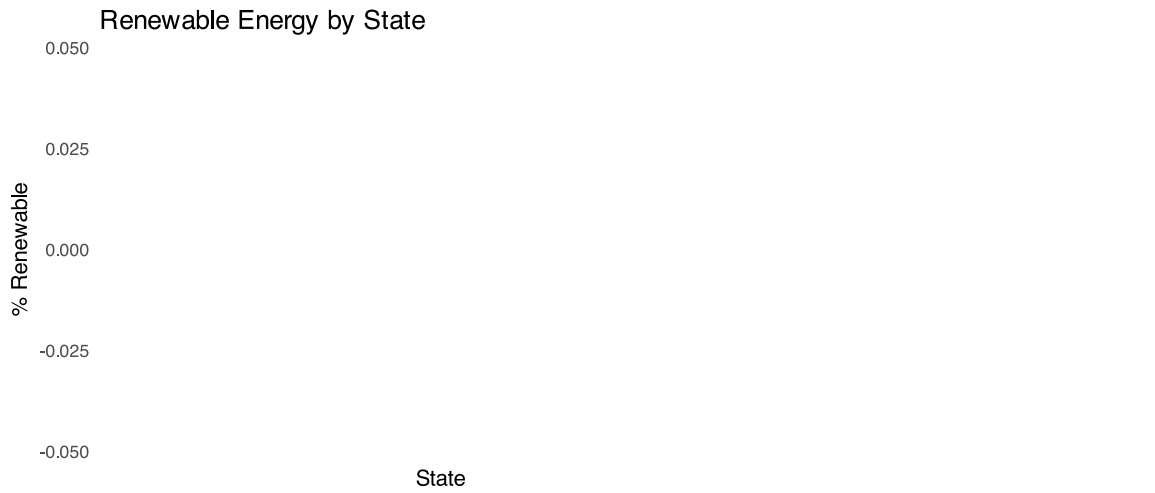
```
    renew_share = (renew_use / total_use) * 100
  )
```

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

```
ggplot(bargraph, aes(x = state_abbr, y = renew_share)) +
  geom_col(fill = "hotpink") +
  labs(
    title = "Renewable Energy by State",
    x = "State",
    y = "% Renewable"
  ) +
  theme_minimal()
```

```
Warning: Position guide is perpendicular to the intended axis.
ℹ Did you mean to specify a different guide `position`?
```

```
Warning: Removed 520 rows containing missing values or values outside the
scale range
(`geom_col()`).
```



I struggled to get my graph to work but from the data it seems like some states have very little renewable energy while others are making great progess. Thus some states, renewable cars are not being run off of reneable energy and instead my be doing more hard than good. But that is not true across all situations as some statees, do have robost networks of renewable energy.