EV Power - Lab 4 Project Report

Example Solution 1

Part 0: libraries

```
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   1.5.1
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 beco
```

```
library(stringr)
library(ggplot2)
```

Part 1: Defining Research Question

Chosen Question: How has renewable energy use changed across U.S. states between 2021 and 2023?

Part 2: Data Preparation and Cleaning

```
renew21 <- 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.
```

```
renew22 <- 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.
```

```
renew23 <- 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.
```

```
total21 <- 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.
```

```
total22 <- 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.
```

```
total23 <- 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.
```

```
ev23 <- 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`
```

```
renew_all <- bind_rows(
renew21 %>% mutate(year = 2021),
renew22 %>% mutate(year = 2022),
renew23 %>% mutate(year = 2023)
)
```

```
total_all <- bind_rows(
total21 %>% mutate(year = 2021),
total22 %>% mutate(year = 2022),
total23 %>% mutate(year = 2023)
)

names(ev23)
```

```
[1] "electric vehicle registrations_by_state (2023)"
[2] "...2"
```

Part 3: Joining / Pivoting Datasets for Analysis

```
summary_data <- renew_all %>%
  group_by(year) %>%
  summarise(avg_use = mean(renew_all, na.rm = TRUE))
```

```
Warning: There were 3 warnings in `summarise()`.
The first warning was:
i In argument: `avg_use = mean(renew_all, na.rm = TRUE)`.
i In group 1: `year = 2021`.
Caused by warning in `mean.default()`:
! argument is not numeric or logical: returning NA
i Run `dplyr::last_dplyr_warnings()` to see the 2 remaining warnings.
```

Part 4: Mapping Visualization

```
ggplot(summary_data, aes(x = year, y = avg_use)) +
  geom_line(color = "steelblue", size = 1.2) +
  geom_point(size = 3, color = "darkblue") +
  labs(
    title = "Average Renewable Energy Use (2021-2023)",
    x = "Year",
    y = "Average Renewable Energy Use"
  ) +
  theme_minimal()
```

```
Warning: Using `size` aesthetic for lines was deprecated in ggplot2 3.4.0.
i Please use `linewidth` instead.
```

```
Warning: Removed 3 rows containing missing values or values outside the scale range
(`geom_line()`).
```

```
Warning: Removed 3 rows containing missing values or values outside the scale range
(`geom_point()`).
```