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(maps)
```

```

Attaching package: 'maps'

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

    map
```

Part 1: Defining Research Question

Chosen Question: In 2023, which states have the highest share of renewable energy, and what does it say about how clean EV charging might be?

Overview: Electric vehicles don't produce tailpipe emissions, however, that doesn't mean that they're entirely clean. Whether they're actually environmentally beneficially depends on where the electricity used to charge them comes from, and EV charging can still contribute to emission in states that rely on fossil fuels like coal or natural gas. In this project, I used the renewable energy use, total energy use, and EV registrations by state datasets. I used these to calculate the

percentage of each state's electricity that comes from renewable sources versus the percentages with EV adoption in order to see whether states with cleaner energy also tend to have more EVs.

Part 2: Data Preparation and Cleaning

```
ev_raw <- 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_raw <- 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.
```

```
total_raw <- 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.
```

```
ev_clean <- ev_raw |>
    rename(State = 1, EV = 2) |>
    filter(!is.na(State), State != "STATE") |>
    mutate(EV = str_remove_all(EV, "[#,~]"),
        EV = str_remove_all(EV, ","),
        EV = str_remove(EV, " EVs| EV"),
        EV = str_trim(EV),
        EV_Count = as.numeric(EV)) |>
    select(State, EV_Count)
```

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

```
state_lookup <- tibble(State_abbrev=state.abb, State=state.name)

renew23 <- renew_raw |>
    rename(State_abbrev = 1, Energy_Source = 2, Renewable = 3) |>
    filter(State_abbrev %in% state.abb) |>
    mutate(
        Renewable = str_remove(Renewable, " kWh"),
        Renewable = str_remove_all(Renewable, ","),
        Renewable_kWh = as.numeric(Renewable)
  ) |>
    group_by(State_abbrev) |>
    summarise(Renewable_Total_kWh = sum(Renewable_kWh, na.rm=TRUE), .groups =
"drop") |>
    left_join(tibble(State_abbrev = state.abb, State = state.name), by
="State_abbrev") |>
    select(State, Renewable_Total_kWh)
```

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

```
total23 <- total_raw |>
    pivot_longer(cols = -Energy_Source, names_to = "State_abbrev", values_to =
"Total") |>
    filter(State_abbrev %in% state.abb) |>
    mutate(
    Total = str_remove(Total, " kWh"),
    Total = str_remove_all(Total, ","),
```

```
    Total_kWh = as.numeric(Total)) |>
    group_by(State_abbrev) |>
    summarise(Total_Energy_kWh = sum(Total_kWh, na.rm = TRUE), .groups =
"drop") |>
    left_join(tibble(State_abbrev = state.abb, State = state.name), by =
"State_abbrev") |>
    select(State, Total_Energy_kWh)
```

Part 3: Joining / Pivoting Datasets for Analysis

```
energy23 <- renew23 |>
    left_join(total23, by = "State") |>
    mutate(Pct_Renewable_2023 = 100 * Renewable_Total_kWh/Total_Energy_kWh) |>
    left_join(ev_clean, by ="State")

energy23 |>
    select(State, Pct_Renewable_2023, EV_Count) |>
    arrange(desc(Pct_Renewable_2023))
```

```
# A tibble: 32 × 3
   State         Pct_Renewable_2023 EV_Count
   <chr>                      <dbl>    <dbl>
 1 South Dakota                34.8     1675
 2 Iowa                        28.3     9031
 3 Maine                       27.2     7377
 4 Oregon                      26.9    64361
 5 Idaho                       18.3     8501
 6 California                  16.6  1256646
 7 Kansas                      12.9    11271
 8 North Dakota                11.4      959
 9 New Hampshire               11.2     9861
10 Georgia                     11.1    92368
# ℹ 22 more rows
```

```
head(energy23)
```

```
# A tibble: 6 × 5
  State      Renewable_Total_kWh Total_Energy_kWh Pct_Renewable_2023 EV_Count
  <chr>                    <dbl>            <dbl>              <dbl>    <dbl>
1 Alaska                   10088           746979               1.35     2697
2 Alabama                 222189          2265008               9.81    13047
3 California             1065179          6429818              16.6   1256646
4 Colorado                109615          1359507               8.06    90083
5 Delaware                  8040           203487               3.95     8435
6 Georgia                 291462          2627553              11.1     92368
```

```r
summary(energy23$Pct_Renewable_2023)
```

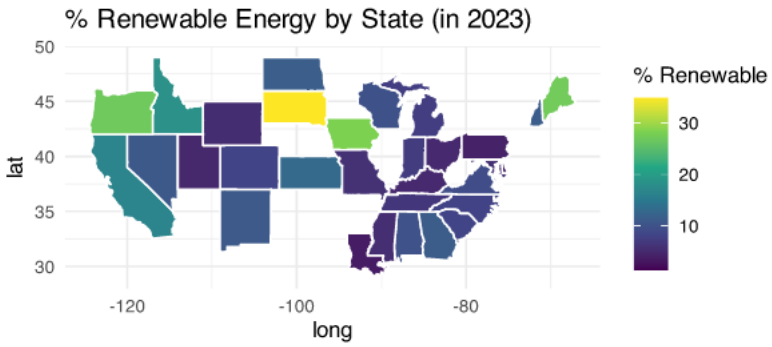
```
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
  1.351   5.363   8.125  10.723  11.267  34.844
```

Part 4: Mapping Visualization

```r
us_map <- map_data("state")

map_df <- energy23 |>
    mutate(state_lower = tolower(State)) |>
    left_join(us_map, by = c("state_lower" = "region"))

ggplot(map_df, aes(long, lat, group = group, fill = Pct_Renewable_2023))+
    geom_polygon(color = "white") +
    coord_fixed(1.3) +
    theme_minimal() +
    scale_fill_viridis_c() +
    labs(title = "% Renewable Energy by State (in 2023)", fill="% Renewable")
```



Analysis

From this project, I found out that renewable energy varied greatly across the U.S. in 2023. The states that had the highest share of renewable electricity were South Dakota, Iowa, and Maine. This might be explained by the greater access to hydro and wind resources, which they can use to generate electricity. The states that had the lowest share of renewable electricity were Alaska, Louisiana, and Delaware.

What was interesting about my findings was that states with a higher percentage of renewable energy didn't necessarily always have the highest number of EVs. One example is South Dakota,

which has the highest percentage of renewable energy but only has around 1,600 EVs. On the other hand, California only has 16.6% of electricity from renewables, but has over 1.2 million EVS.

Overall, EV charging is cleaner in states where renewable energy takes up a large share, however, renewable energy share doesn't determine EV adoption in states.