Growth of Renewable Usage Between 2021 and 2023 - Lab 4 Project Report

Example Solution 1

Part 0: libraries

```
library(tidyverse)
```

```
Warning: package 'ggplot2' was built under R version 4.3.3
```

```
Warning: package 'tibble' was built under R version 4.3.3
```

```
Warning: package 'purrr' was built under R version 4.3.3
```

```
-- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
v dplyr     1.1.4     v readr     2.1.4
v forcats   1.0.0     v stringr   1.5.1
v ggplot2   3.5.2     v tibble    3.3.0
v lubridate 1.9.3     v tidyr     1.3.1
v purrr     1.0.4
-- 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(stringr)
library(tidyr)
library(dplyr)
```

```
library(ggplot2)
library(maps)
```

```
Warning: package 'maps' was built under R version 4.3.3
```

```
Attaching package: 'maps'

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

    map
```

```
library(sf)
```

```
Warning: package 'sf' was built under R version 4.3.3

Linking to GEOS 3.13.0, GDAL 3.8.5, PROJ 9.5.1; sf_use_s2() is TRUE
```

Part 1: Defining Research Question

Chosen Question: Which areas in the United States have the highest renewable energy growth and most ethical EV usage from 2021 to 2023?

Part 2: Data Preparation and Cleaning

```
remove_details <- function(x)
  as.numeric(str_replace_all(x, "[^0-9.]", "")
)

first <- read_csv("/Users/abhinavgoel/ev-power-Abh1navgoel459/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.
```

```r
first$State <- toupper(first$State)

summary_one <- first %>%
  mutate(Renewable_Use_2021 = remove_details(Renewable_Use_2021)) %>%
  group_by(State) %>%
  summarise(total = sum(Renewable_Use_2021)) %>%
  mutate(year = 2021)

second <- read_csv("/Users/abhinavgoel/ev-power-Abh1navgoel459/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.
```

```r
second$State <- toupper(second$State)

summary_two <- second %>%
  mutate(Renewable_Use_2022 = remove_details(Renewable_Use_2022)) %>%
  group_by(State) %>%
  summarise(total = sum(Renewable_Use_2022)) %>%
  mutate(year = 2022)

third <- read_csv("/Users/abhinavgoel/ev-power-Abh1navgoel459/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.
```

```r
third$State <- toupper(third$State)

summary_three <- third %>%
```

```
    mutate(Renewable_Use_2023 = remove_details(Renewable_Use_2023)) %>%
    group_by(State) %>%
    summarise(total = sum(Renewable_Use_2023)) %>%
    mutate(year = 2023)

  agg <- bind_rows(summary_one, summary_two, summary_three)

  print(head(agg))
```

```
# A tibble: 6 x 3
  State  total  year
  <chr>  <dbl> <dbl>
1 AK      9598  2021
2 AL    239816  2021
3 AR     89714  2021
4 AZ     99266  2021
5 CA    810020  2021
6 CO    103956  2021
```

Part 3: Joining / Pivoting Datasets for Analysis

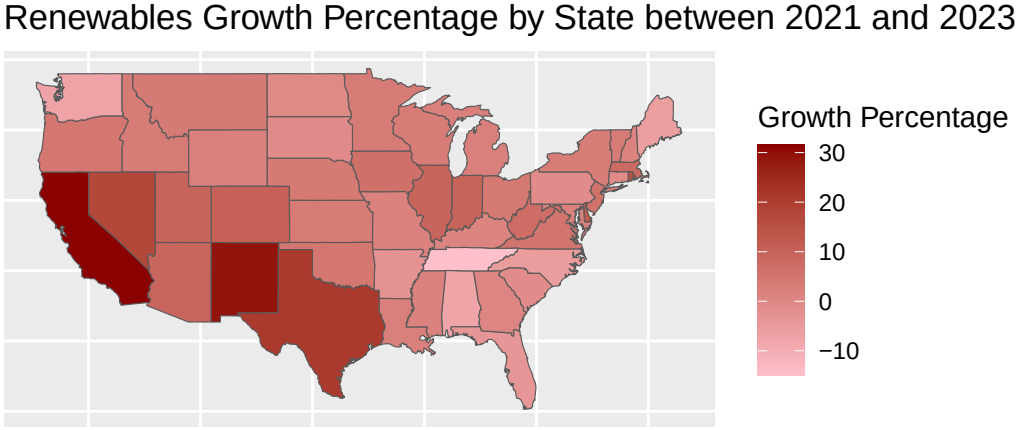
```
  pivoted <- pivot_wider(
    data = agg,
    names_from = year,
    values_from = total,
    names_prefix = "total_"
    )

  pivoted$total_growth<- pivoted$total_2023 - pivoted$total_2021
  pivoted$percent_growth <- ((pivoted$total_2023 - pivoted$total_2021) / pivoted$total_2021)

  print(head(pivoted))
```

```
# A tibble: 6 x 6
  State total_2021 total_2022 total_2023 total_growth percent_growth
  <chr>      <dbl>      <dbl>      <dbl>        <dbl>          <dbl>
1 AK          9598      10410      10088          490           5.11
2 AL        239816     232035     222189       -17627          -7.35
3 AR         89714      90824      87277        -2437          -2.72
```

```
4 AZ          99266      101214      108445          9179           9.25
5 CA         810020      880995     1065179        255159          31.5
6 CO         103956      114918      115062         11106          10.7
```

Part 4: Mapping Visualization

```
ggplot(
  left_join(
    st_as_sf(map("state", fill = TRUE, plot = FALSE)),
    pivoted %>% mutate(State = tolower(state.name[match(State, state.abb)])),
    by = c("ID" = "State")
    )
  ) +
  geom_sf(aes(fill = percent_growth)) +
  scale_fill_gradient(low = "pink",
                      high = "darkred",
                      name = "Growth Percentage") +
  labs(
    title = "Renewables Growth Percentage by State between 2021 and 2023",
  ) +
  theme(
    axis.text = element_blank(),
    axis.ticks = element_blank(),
    plot.title = element_text(),
  )
```



Overview and Analysis

Overview: We have seen EV usage grow dramatically in the past couple of years, fueled by innovation in the EV space from players like Tesla and Rivian. This has led to the slow phasing out of gas, which has been a trend mirroed in all renewables. With the adoption of electric cars, solar panels, and even electric stoves, gas has become less and less of a necessity. From my question, I hope to see how much renewable energy growth has occurred between 2021 and 2023, helping us understand how popular renewables have become and how EVs have thus been able to thrive.

Analysis: From our graph, we that states like Texas, New Mexico, and California have the highest renewable energy usage, which is crucial, because this likely means that these states have some of the highest usages of EVs as well. We can see that more western states have higher renewable usage, whereas more east-coast and central states tend to be lower. Rather, midwestern states, specifically midwestern states tending more towards the east coast, actually had a decrease in renewable energy usage, the most egregious offender being Tennessee, dropping ~10 percent. This is crucial, because with greater implementation of renewables in the west as well, we can see that the west has contributed more to better and more ethical use of EVs, as EVs are not being powered by gas, but rather by the clean and efficient renewable energy that is being produced.