# 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(sf)
```

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

```
library(rnaturalearth)
library(lubridate)
library(stringr)
```

### Part 1: Defining Research Question

Chosen Question: How has the share of renewable energy changed from 2021-2023 across states?

### Part 2: Data Preparation and Cleaning

```
# Open each data file and rename them:
# Adding a new year column because eventually when I combine all these
different years into one dataset, I am able to recognize in each row which
year the observation belongs to.
renew_2021 <- read_csv("~/stat133/ev-power-leyna-liu/data/renew-use-2021.csv")
```

```
|>
    mutate(year = 2021)
```

```
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.
```

```
renew_2022 <- read_csv("~/stat133/ev-power-leyna-liu/data/renew-use-2022.csv")
|>
    mutate(year = 2022)
```

```
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.
```

```
renew_2023 <- read_csv("~/stat133/ev-power-leyna-liu/data/renew-use-2023.csv")
|>
    mutate(year = 2023)
```

```
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_2021 <- read_csv("~/stat133/ev-power-leyna-liu/data/total-use-2021.csv")
|>
    mutate(year = 2021)
```

```
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.
```

```
total_2022 <- read_csv("~/stat133/ev-power-leyna-liu/data/total-use-2022.csv")
|>
    mutate(year = 2022)
```

```
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.
```

```
total_2023 <- read_csv("~/stat133/ev-power-leyna-liu/data/total-use-2023.csv")
|>
    mutate(year = 2023)
```

```
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.
```

```
# Now, I am going to Standardize the column names, or basically just make sure
all the column names are the same/consistent throughout all datasets
# So, I am going to make all column names lower case through all datasets
# names(dataframe) -> outputs the names of all columns
# tolower(names(dataframe)) -> makes all the column names lowercase
# names(dataframe) <- tolower(names(dataframe)): then assigning the lower case
column names back to the original column names to save it back into the
dataframe
names(renew_2021) <- tolower(names(renew_2021))
names(renew_2022) <- tolower(names(renew_2022))
names(renew_2023) <- tolower(names(renew_2023))
names(total_2021) <- tolower(names(total_2021))
names(total_2022) <- tolower(names(total_2022))
names(total_2023) <- tolower(names(total_2023))
```

```
# Now that the column names are in lowercase, I will make sure all state names
are in lower case too (in the State column), to make sure state names are
consistent across all files
renew_2021 <- renew_2021 |>
    mutate(state = tolower(state))
renew_2022 <- renew_2022 |>
    mutate(state = tolower(state))
renew_2023 <- renew_2023 |>
    mutate(state = tolower(state))

# Since, in the total_2021, total_2022, total_2023 datasets the state names
are across the columns I can't use the tolower() function across all columns
that would be way too tedious, so I will deal with them in Part 3, where I can
use pivots.
```

```
# I noticed how in the total_2021, total_2022, total_2023 datasets, the
Energy_Source column's cells have different spellings for each energy source
(ex. Coal vs. ), so to make it all the same I will use strings.

# Here, the str_replace() follows this format:
# str_replace(dataset or column you want to change, what you want to change in
that column or dataset, and what you want to change that word into)
# In: "Natural Gas.*" -> the . is a wildcard meaning that it will output any
character followed by gas, and the * means any amount, so together it means
any amount of characters after gas
total_2021 <- total_2021 |>
```

```
    mutate(
        energy_source = str_replace(energy_source, "Coal", "coal"),
        energy_source = str_replace(energy_source, "Natural Gas.*", "natural
gas"),
        energy_source = str_replace(energy_source, "Petroleum.*",
"petroleum"),
        energy_source = str_replace(energy_source, "nuclear", "nuclear"),
        energy_source = str_replace(energy_source, "total_renewable_energy",
"total_renewable_energy")
    )
# Now, I will repeat this process with total_2022 and total_2023
total_2022 <- total_2022 |>
    mutate(
        energy_source = str_replace(energy_source, "coal.*", "coal"),
        energy_source = str_replace(energy_source, "Natural-Gas", "natural
gas"),
        energy_source = str_replace(energy_source, "petroleum.*",
"petroleum"),
        energy_source = str_replace(energy_source, "Nuclear.*", "nuclear"),
        energy_source = str_replace(energy_source, "total_renewable.*",
"total_renewable_energy")
    )

total_2023 <- total_2023 |>
    mutate(
        energy_source = str_replace(energy_source, "coal.*", "coal"),
        energy_source = str_replace(energy_source, "Natural.*", "natural
gas"),
        energy_source = str_replace(energy_source, "petroleum.*",
"petroleum"),
        energy_source = str_replace(energy_source, "nuclear.*", "nuclear"),
        energy_source = str_replace(energy_source, "total renewable.*",
"total_renewable_energy")
    )
```

## Part 3: Joining / Pivoting Datasets for Analysis

```
# Now, I will combine all the years into a single dataset. (The new year
column I created in Part 2 with each year listed in each row will come in
handy now to be able to distinguish what year the observations are from.)
total_years_joined <- total_2021 |>
    full_join(total_2022, by = "energy_source") |>
    full_join(total_2023, by = "energy_source")

# Since this is using the original total_2021, total_2022, total_2023
datasets, I need to pivot it, since the columns are each states.
total_years_pivoted <- pivot_longer(
    total_years_joined,
```

```
    cols = -c("energy_source"),
    names_to = "states.year",
    values_to = "count"
)
view(total_years_pivoted)
# cols = -c("energy_source") -> because I want all the columns except
energy_source and it would be too tedious to write each of them out, so I just
"minus" energy_source
# Since the original datasets total_2021, total_2022, total_2023 all have the
same column names, when the columns of the same state names from the other
datasets are joined onto total_2021, R automatically gives the state names
different suffixes added to it.
# Therefore, ak.x = year 2021, ak.y = year 2022, ak = year 2023
```

```
# Now, I will also join renew datasets into one single renew dataset:
renew_joined <- renew_2021 |>
    full_join(renew_2022, by = c("state", "energy_source")) |>
    full_join(renew_2023, by = c("state", "energy_source"))
view(renew_joined)

# Here, it is crucial that I have joined by both "state" and "energy_source"
because both columns are matching keys in this case that are the same in all
renew datasets. If I just match by "state" then it will take every row from
2021 for a given state and try to match it with every row from 2022 with the
same state. (for ex. if there is 3 rows of ak in 2021 and 3 rows of ak in 2022
-> 3*3 = 9 rows after full_join)
# But if I match with both state and energy source it will look for a match of
the same state AND energy source for 2021 to 2022. (for ex. ak Biomass in 2021
matches with ak biomass in 2022)
```

The method I chose is to join the total-use datasets into a single total_years_joined datasets that combines all the years. I had to pivot it to make sure it was more legible. Then I did the same thing to the renew-use datasets. I combined all the years of renew-use into a single dataset. The reason I did this is so that later is it easier to make the map visualization, and I know I am trying to explore the question of renewable usage OVER TIME, so it would be to my advantage to group all the years together in a centralized dataset.

```
# Now, I will create a new variable, which is percentage/ratio of renewable
energy source out of total energy by year
# First, I will pivot each year of total dataset
total_2021_pivoted <- pivot_longer(total_2021, cols = -c(energy_source, year),
names_to = "state", values_to = "count")

total_2022_pivoted <- pivot_longer(total_2022, cols = -c(energy_source, year),
names_to = "state", values_to = "count")
```

```r
total_2023_pivoted <- pivot_longer(total_2023, cols = -c(energy_source, year),
names_to = "state", values_to = "count")

# Then, I will create the total energy for each year
total_energy_2021 <- total_2021_pivoted |>
    group_by(state) |>
    summarize(total_energy = sum(count))

total_energy_2022 <- total_2022_pivoted |>
    group_by(state) |>
    summarize(total_energy = sum(count))

total_energy_2023 <- total_2023_pivoted |>
    group_by(state) |>
    summarize(total_energy = sum(count))

# Then, I will create the renewable energy per state per year
# renewable_energy_count = count -> I renamed the count column to
renewable_energy_count specifically because later when I do the calculations
it will be easier and it would be confusing with two count columns (1 total
count, 1 renewable count)
renewable_2021 <- total_2021_pivoted |>
    filter(energy_source == "total_renewable_energy") |>
    select(state, renewable_energy_count = count)

renewable_2022 <- total_2022_pivoted |>
    filter(energy_source == "total_renewable_energy") |>
    select(state, renewable_energy_count = count)

renewable_2023 <- total_2023_pivoted |>
    filter(energy_source == "total_renewable_energy") |>
    select(state, renewable_energy_count = count)

# Finally, we can calculate the percentage
# Since we are looking for percentage of renewable energy out of total energy
I multipled 100 to total_energy
percentage_2021 <- total_energy_2021 |>
    left_join(renewable_2021, by = "state") |>
    mutate(ratio_percent = renewable_energy_count / total_energy * 100)

percentage_2022 <- total_energy_2022 |>
    left_join(renewable_2022, by = "state") |>
    mutate(ratio_percent = renewable_energy_count / total_energy * 100)

percentage_2023 <- total_energy_2023 |>
    left_join(renewable_2023, by = "state") |>
    mutate(ratio_percent = renewable_energy_count / total_energy * 100)
```

```
view(percentage_2021)
view(percentage_2022)
view(percentage_2023)
```

Suprisingly, it seems like in 2021 South Dakota had the highest percentage of renewable energy usage at 34.87%. In 2022, again it is South Dakota with a percentage of 34.98%. In 2023, again South Dakota with 34.84%.

## Part 4: Mapping Visualization

```
# This creates the data that goes into geom_sf() later:
united_states_map <- ne_states(country = "United States of America",
returnclass = "sf")
united_states_map
```

```
Simple feature collection with 51 features and 121 fields
Geometry type: MULTIPOLYGON
Dimension:     XY
Bounding box:  xmin: -179.1435 ymin: 18.90612 xmax: 179.7809 ymax: 71.4125
Geodetic CRS:  WGS 84
First 10 features:
                        featurecla scalerank adm1_code diss_me iso_3166_2
1236 Admin-1 states provinces lakes         2  USA-3519    3519      US-WA
1238 Admin-1 states provinces lakes         2  USA-3518    3518      US-ID
1239 Admin-1 states provinces lakes         2  USA-3515    3515      US-MT
1242 Admin-1 states provinces lakes         2  USA-3516    3516      US-ND
1244 Admin-1 states provinces lakes         2  USA-3514    3514      US-MN
1246 Admin-1 states provinces lakes         2  USA-3562    3562      US-MI
1247 Admin-1 states provinces lakes         2  USA-3550    3550      US-OH
1248 Admin-1 states provinces lakes         2  USA-3560    3560      US-PA
1249 Admin-1 states provinces lakes         2  USA-3559    3559      US-NY
1251 Admin-1 states provinces lakes         2  USA-3540    3540      US-VT
                                          wikipedia iso_a2 adm0_sr
1236 http://en.wikipedia.org/wiki/Washington_(state)     US       6
1238              http://en.wikipedia.org/wiki/Idaho     US       1
1239            http://en.wikipedia.org/wiki/Montana     US       1
1242       http://en.wikipedia.org/wiki/North_Dakota     US       1
1244          http://en.wikipedia.org/wiki/Minnesota     US       1
1246            http://en.wikipedia.org/wiki/Michigan     US       1
1247                http://en.wikipedia.org/wiki/Ohio     US       1
1248       http://en.wikipedia.org/wiki/Pennsylvania     US       1
1249            http://en.wikipedia.org/wiki/New_York     US       3
1251             http://en.wikipedia.org/wiki/Vermont     US       1
             name                        name_alt name_local  type type_en
1236   Washington                      WA|Wash.       <NA> State   State
1238        Idaho                      ID|Idaho       <NA> State   State
1239      Montana                      MT|Mont.       <NA> State   State
```

```
1242 North Dakota                         ND|N.D.       <NA> State   State
1244    Minnesota                        MN|Minn.       <NA> State   State
1246     Michigan                        MI|Mich.       <NA> State   State
1247         Ohio                         OH|Ohio       <NA> State   State
1248 Pennsylvania Commonwealth of Pennsylvania|PA       <NA> State   State
1249     New York                         NY|N.Y.       <NA> State   State
1251      Vermont                              VT       <NA> State   State
     code_local code_hasc note hasc_maybe    region region_cod provnum_ne
1236       US53     US.WA <NA>       <NA>      West       <NA>          0
1238       US16     US.ID <NA>       <NA>      West       <NA>          0
1239       US30     US.MT <NA>       <NA>      West       <NA>          0
1242       US38     US.ND <NA>       <NA>   Midwest       <NA>          0
1244       US27     US.MN <NA>       <NA>   Midwest       <NA>          0
1246       US26     US.MI <NA>       <NA>   Midwest       <NA>          0
1247       US39     US.OH <NA>       <NA>   Midwest       <NA>          0
1248       US42     US.PA <NA>       <NA> Northeast       <NA>          0
1249       US36     US.NY <NA>       <NA> Northeast       <NA>          0
1251       US50     US.VT <NA>       <NA> Northeast       <NA>          0
     gadm_level check_me datarank abbrev postal area_sqkm sameascity labelrank
1236          1       20        1  Wash.     WA         0        -99         0
1238          1       20        1  Idaho     ID         0        -99         0
1239          1       20        1  Mont.     MT         0        -99         0
1242          1       20        1   N.D.     ND         0        -99         0
1244          1       20        1  Minn.     MN         0        -99         0
1246          1       20        1  Mich.     MI         0        -99         0
1247          1       20        1   Ohio     OH         0        -99         0
1248          1       20        1    Pa.     PA         0        -99         0
1249          1       20        1   N.Y.     NY         0        -99         0
1251          1       20        1    Vt.     VT         0        -99         0
     name_len mapcolor9 mapcolor13 fips fips_alt  woe_id
1236       10         1          1 US53     <NA> 2347606
1238        5         1          1 US16     <NA> 2347571
1239        7         1          1 US30     <NA> 2347585
1242       12         1          1 US38     <NA> 2347593
1244        9         1          1 US27     <NA> 2347582
1246        8         1          1 US26     <NA> 2347581
1247        4         1          1 US39     <NA> 2347594
1248       12         1          1 US42     <NA> 2347597
1249        8         1          1 US36     <NA> 2347591
1251        7         1          1 US50     <NA> 2347604
                          woe_label     woe_name latitude longitude sov_a3
1236   Washington, US, United States   Washington  47.4865 -120.3610    US1
1238        Idaho, US, United States        Idaho  43.7825 -114.1330    US1
1239      Montana, US, United States      Montana  46.9965 -110.0440    US1
1242 North Dakota, US, United States North Dakota  47.4675 -100.3020    US1
1244    Minnesota, US, United States    Minnesota  46.0592  -93.3640    US1
1246     Michigan, US, United States     Michigan  43.4343  -84.9479    US1
1247         Ohio, US, United States         Ohio  40.0924  -82.6719    US1
```

```
1248 Pennsylvania, US, United States Pennsylvania  40.8601  -77.6094    US1
1249     New York, US, United States     New York  43.1988  -75.3242    US1
1251      Vermont, US, United States      Vermont  44.0886  -72.7317    US1
     adm0_a3 adm0_label                   admin                 geonunit
gu_a3
1236     USA          2 United States of America United States of America
USA
1238     USA          2 United States of America United States of America
USA
1239     USA          2 United States of America United States of America
USA
1242     USA          2 United States of America United States of America
USA
1244     USA          2 United States of America United States of America
USA
1246     USA          2 United States of America United States of America
USA
1247     USA          2 United States of America United States of America
USA
1248     USA          2 United States of America United States of America
USA
1249     USA          2 United States of America United States of America
USA
1251     USA          2 United States of America United States of America
USA
       gn_id      gn_name gns_id gns_name gn_level gn_region gn_a1_code
1236 5815135   Washington     -1     <NA>        1      <NA>      US.WA
1238 5596512        Idaho     -1     <NA>        1      <NA>      US.ID
1239 5667009      Montana     -1     <NA>        1      <NA>      US.MT
1242 5690763 North Dakota     -1     <NA>        1      <NA>      US.ND
1244 5037779    Minnesota     -1     <NA>        1      <NA>      US.MN
1246 5001836     Michigan     -1     <NA>        1      <NA>      US.MI
1247 5165418         Ohio     -1     <NA>        1      <NA>      US.OH
1248 6254927 Pennsylvania     -1     <NA>        1      <NA>      US.PA
1249 5128638     New York     -1     <NA>        1      <NA>      US.NY
1251 5242283      Vermont     -1     <NA>        1      <NA>      US.VT
             region_sub sub_code gns_level gns_lang gns_adm1 gns_region
1236            Pacific     <NA>        -1     <NA>     <NA>       <NA>
1238           Mountain     <NA>        -1     <NA>     <NA>       <NA>
1239           Mountain     <NA>        -1     <NA>     <NA>       <NA>
1242 West North Central     <NA>        -1     <NA>     <NA>       <NA>
1244 West North Central     <NA>        -1     <NA>     <NA>       <NA>
1246 East North Central     <NA>        -1     <NA>     <NA>       <NA>
1247 East North Central     <NA>        -1     <NA>     <NA>       <NA>
1248    Middle Atlantic     <NA>        -1     <NA>     <NA>       <NA>
1249    Middle Atlantic     <NA>        -1     <NA>     <NA>       <NA>
1251        New England     <NA>        -1     <NA>     <NA>       <NA>
     min_label max_label min_zoom wikidataid         name_ar      name_bn
```

```
1236       3.5       7.5        2      Q1223          واشنطن      ওয়াশিংটন
1238       3.5       7.5        2      Q1221           أيداهو        আইডাহো
1239       3.5       7.5        2      Q1212          مونتانا         মন্টানা
1242       3.5       7.5        2      Q1207 داكوتا الشمالية  নর্থ ডাকোটা
1244       3.5       7.5        2      Q1527         مينيسوتا      মিনেসোটা
1246       3.5       7.5        2      Q1166         ميشيغان       মিশিগান
1247       3.5       7.5        2      Q1397          أوهايو         ওহাইও
1248       3.5       7.5        2      Q1400      بنسيلفانيا পেনসিলভেনিয়া
1249       3.5       7.5        2      Q1384         نيويورك      নিউ ইয়র্ক
1251       3.5       7.5        2     Q16551         فيرمونت        ভার্মন্ট
          name_de      name_en          name_es          name_fr
name_el
1236   Washington   Washington       Washington       Washington
Ουάσινγκτον
1238        Idaho        Idaho            Idaho            Idaho
Άινταχο
1239      Montana      Montana          Montana          Montana
Μοντάνα
1242 North Dakota North Dakota Dakota del Norte   Dakota du Nord Βόρεια
Ντακότα
1244    Minnesota    Minnesota        Minnesota        Minnesota
Μινεσότα
1246     Michigan     Michigan         Míchigan         Michigan
Μίσιγκαν
1247         Ohio         Ohio             Ohio             Ohio
Οχάιο
1248 Pennsylvania Pennsylvania      Pensilvania     Pennsylvanie
Πενσιλβάνια
1249     New York     New York       Nueva York État de New York      Νέα
Υόρκη
1251      Vermont      Vermont          Vermont          Vermont
Βερμόντ
          name_hi      name_hu      name_id         name_it          name_ja
1236 वॉशिंगटन राज्य   Washington   Washington      Washington     ワシントン州
1238       आयडाहो        Idaho        Idaho           Idaho       アイダホ州
1239      मोन्टाना      Montana      Montana         Montana       モンタナ州
1242   उत्तर डेकोटा Észak-Dakota Dakota Utara Dakota del Nord   ノースダコタ州
1244      मिनेसोटा    Minnesota    Minnesota       Minnesota       ミネソタ州
1246       मिशिगन     Michigan     Michigan        Michigan       ミシガン州
1247        ओहायो         Ohio         Ohio            Ohio       オハイオ州
1248   पेन्सिलवेनिया Pennsylvania Pennsylvania    Pennsylvania ペンシルベニア州
1249      न्यूयॉर्क     New York     New York        New York   ニューヨーク州
1251       वर्मांट      Vermont      Vermont         Vermont     バーモント州
          name_ko      name_nl         name_pl         name_pt         name_ru
1236       워싱턴   Washington      Waszyngton      Washington       Вашингтон
1238     아이다호        Idaho           Idaho           Idaho          Айдахо
1239       몬태나      Montana         Montana         Montana         Монтана
1242   노스다코타 Noord-Dakota Dakota Północna Dakota do Norte Северная Дакота
```

| | | | | | | |
|---|---|---|---|---|---|---|
| 1244 | 머네소타 | Minnesota | Minnesota | Minnesota | Миннесота | |
| 1246 | 머시갼 | Michigan | Michigan | Michigan | Мичиган | |
| 1247 | 오하이오 | Ohio | Ohio | Ohio | Огайо | |
| 1248 | 펜싴베이니아 | Pennsylvania | Pensylwania | Pensilvânia | Пенсильвания | |
| 1249 | 뉴욬 | New York | Nowy Jork | Nova Iorque | Нью-Йорк | |
| 1251 | 버몬트 | Vermont | Vermont | Vermont | Вермонт | |

| | name_sv | name_tr | name_vi | name_zh | ne_id |
|---|---|---|---|---|---|
| 1236 | Washington | Vaşington | Washington | 华盛顿州 | 1159309547 |
| 1238 | Idaho | Idaho | Idaho | 爱达荷州 | 1159315339 |
| 1239 | Montana | Montana | Montana | 蒙大拿州 | 1159315333 |
| 1242 | North Dakota | Kuzey Dakota | Bắć Dakota | 北达科他州 | 1159315337 |
| 1244 | Minnesota | Minnesota | Minnesota | 明尼苏达州 | 1159315297 |
| 1246 | Michigan | Michigan | Michigan | 密歇根州 | 1159314665 |
| 1247 | Ohio | Ohio | Ohio | 俄亥俄州 | 1159315315 |
| 1248 | Pennsylvania | Pensilvanya | Pennsylvania | 宾夕法尼亚州 | 1159315331 |
| 1249 | New York | New York | New York | 纽约州 | 1159312155 |
| 1251 | Vermont | Vermont | Vermont | 佛蒙特州 | 1159315305 |

| | name_he | name_uk | name_ur | name_fa | name_zht |
|---|---|---|---|---|---|
| 1236 | וושינגטון | Вашингтон | ریاست واشنگٹن ایالت واشینگتن | | 華盛頓州 |
| 1238 | איידהו | Айдахо | آیداهو | ایڈاہو | 愛達荷州 |
| 1239 | מונטנה | Монтана | ایالت مونتانا | مونٹانا | 蒙大拿州 |
| 1242 | דקוטה הצפונית | Північна Дакота | داکوتای شمالی | شمالی ڈکوٹا | 北達科他州 |
| 1244 | מינסוטה | Міннесота | مینهسوتا | مینیسوٹا | 明尼蘇達州 |
| 1246 | מישיגן | Мічиган | میشیگان | مشی گن | 密歇根州 |
| 1247 | אוהיו | Огайо | اوھایو | اوہائیو | 俄亥俄州 |
| 1248 | פנסילבניה | Пенсильванія | پنسلوانیا | پنسلوانیا | 賓夕法尼亞州 |
| 1249 | ניו יורק | штат Нью-Йорк | نیویورک | نیویارک | 紐約州 |
| 1251 | ורמונט | Вермонт | ورمونت | ورمونٹ | 佛蒙特州 |

| | FCLASS_ISO | FCLASS_US | FCLASS_FR | FCLASS_RU | FCLASS_ES | FCLASS_CN | FCLASS_TW |
|---|---|---|---|---|---|---|---|
| 1236 | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> |
| 1238 | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> |
| 1239 | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> |
| 1242 | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> |
| 1244 | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> |
| 1246 | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> |
| 1247 | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> |
| 1248 | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> |
| 1249 | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> |
| 1251 | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> |

| | FCLASS_IN | FCLASS_NP | FCLASS_PK | FCLASS_DE | FCLASS_GB | FCLASS_BR | FCLASS_IL |
|---|---|---|---|---|---|---|---|
| 1236 | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> |
| 1238 | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> |
| 1239 | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> |
| 1242 | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> |
| 1244 | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> |
| 1246 | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> |
| 1247 | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> |
| 1248 | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> | <NA> |

```
1249      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>
1251      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>
     FCLASS_PS FCLASS_SA FCLASS_EG FCLASS_MA FCLASS_PT FCLASS_AR FCLASS_JP
1236      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>
1238      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>
1239      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>
1242      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>
1244      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>
1246      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>
1247      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>
1248      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>
1249      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>
1251      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>
     FCLASS_KO FCLASS_VN FCLASS_TR FCLASS_ID FCLASS_PL FCLASS_GR FCLASS_IT
1236      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>
1238      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>
1239      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>
1242      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>
1244      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>
1246      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>
1247      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>
1248      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>
1249      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>
1251      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>      <NA>
     FCLASS_NL FCLASS_SE FCLASS_BD FCLASS_UA FCLASS_TLC
1236      <NA>      <NA>      <NA>      <NA>       <NA>
1238      <NA>      <NA>      <NA>      <NA>       <NA>
1239      <NA>      <NA>      <NA>      <NA>       <NA>
1242      <NA>      <NA>      <NA>      <NA>       <NA>
1244      <NA>      <NA>      <NA>      <NA>       <NA>
1246      <NA>      <NA>      <NA>      <NA>       <NA>
1247      <NA>      <NA>      <NA>      <NA>       <NA>
1248      <NA>      <NA>      <NA>      <NA>       <NA>
1249      <NA>      <NA>      <NA>      <NA>       <NA>
1251      <NA>      <NA>      <NA>      <NA>       <NA>
                          geometry
1236 MULTIPOLYGON (((-122.753 48...
1238 MULTIPOLYGON (((-117.0382 4...
1239 MULTIPOLYGON (((-116.0482 4...
1242 MULTIPOLYGON (((-104.0476 4...
1244 MULTIPOLYGON (((-97.22609 4...
1246 MULTIPOLYGON (((-84.4913 46...
1247 MULTIPOLYGON (((-80.52023 4...
1248 MULTIPOLYGON (((-79.76301 4...
1249 MULTIPOLYGON (((-79.06523 4...
1251 MULTIPOLYGON (((-73.35134 4...
```

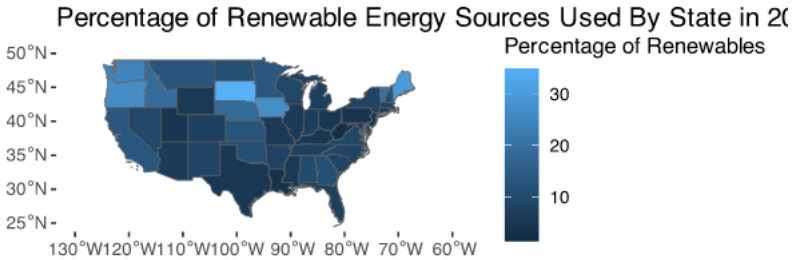
```r
# after doing view(united_states_map) you can see the state names under postal column AND the state names are in UPPER CASE
# therefore, we have to alter our percentage_2021 to have state names in upper case too
# then join the new percentage_2021 data with upper case states to the map with key "state" which corresponds to woe_name in the map
percentage_2021 <- percentage_2021 |>
    mutate(state = toupper(state))
map_joined <- united_states_map |>
    left_join(percentage_2021, by = c("postal" = "state"))

# Now I will repeat this for every year
percentage_2022 <- percentage_2022 |>
    mutate(state = toupper(state))
map_joined2 <- united_states_map |>
    left_join(percentage_2022, by = c("postal" = "state"))

percentage_2023 <- percentage_2023 |>
    mutate(state = toupper(state))
map_joined3 <- united_states_map |>
    left_join(percentage_2023, by = c("postal" = "state"))
```

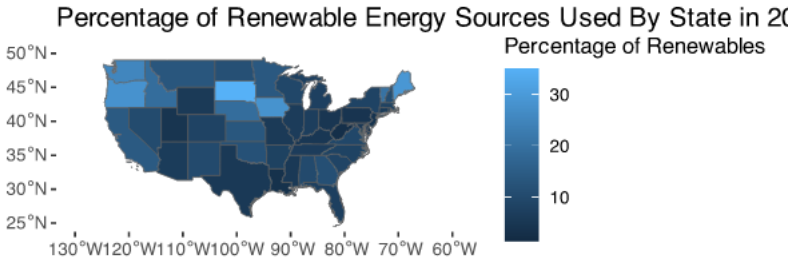
```r
# now finally to create the map for 2021:
map2021 <- ggplot() + geom_sf(data = united_states_map) + geom_sf(data = map_joined, aes(fill = ratio_percent)) + coord_sf(xlim = c(-130, -60), ylim = c(25, 50)) + labs(title = "Percentage of Renewable Energy Sources Used By State in 2021", fill = "Percentage of Renewables") + theme(panel.background = element_blank())

map2021
```



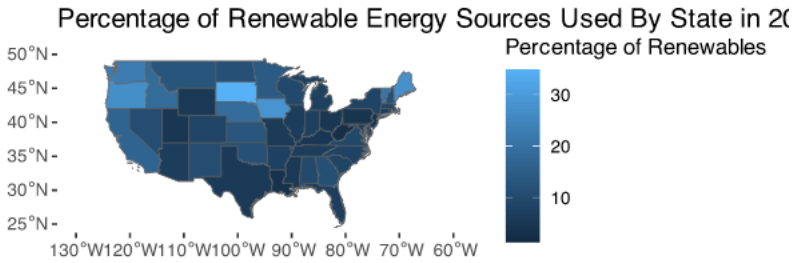
```
# map for 2022:
map2022 <- ggplot() + geom_sf(data = united_states_map) + geom_sf(data =
map_joined2, aes(fill = ratio_percent)) + coord_sf(xlim = c(-130, -60), ylim =
c(25, 50)) + labs(title = "Percentage of Renewable Energy Sources Used By
State in 2022", fill = "Percentage of Renewables") + theme(panel.background =
element_blank())

map2022
```




```
# map for 2023:
map2023 <- ggplot() + geom_sf(data = united_states_map) + geom_sf(data =
map_joined3, aes(fill = ratio_percent)) + coord_sf(xlim = c(-130, -60), ylim =
c(25, 50)) + labs(title = "Percentage of Renewable Energy Sources Used By
State in 2023", fill = "Percentage of Renewables") + theme(panel.background =
element_blank())

map2023
```




### Part 5: Final Deliverable

## Analysis

The patterns I notice is that over time there seems to be a slight increase in Renewable Energy usage throughout the United States. However, the increase is so minimal it is almost non-existent. South Dakota does seem to have the highest percentage of Renewable Energy sources used. I expected California to be higher in Renewable Energy Usage, but it seems that California seems to have average percentages of Renewables over the years 2021 to 2023. My maps help to answer the main research question of: How has the share of renewable energy changed from 2021-2023 across states? This is because through the three seperate maps by years 2021, 2022, and 2023 one can see distinctly the differences in percentages of Renewables used across the states. From the map, it is evident that there is not drastic enough changes to be able to see the color gradients change that much.