

EV Power - Lab 4 Project Report

Example Solution 1

Part 0: libraries

```
setwd("c:/Users/light/downloads/ev-power-oscarcheng-1/data")  
  
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(dplyr)  
library(readr)  
library(stringr)  
# Read all datasets  
renew_2021 <- read_csv("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.
```

```
renew_2022 <- read_csv("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.
```

```
renew_2023 <- read_csv("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.
```

```
energy_price <- read_csv("av-energy-price-2021-2023.csv")
```

```
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.
```

```
total_energy_2021 <- read_csv("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.
```

```
ev_registrations <- read_csv("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`
```

Part 1: Defining Research Question

Chosen Question: How much of each state's total electricity comes from renewable sources between 2021–2023?

Part 2: Data Preparation and Cleaning

```
# Clean column names manually  
names(renew_2021) <- tolower(names(renew_2021))  
names(renew_2021) <- str_replace_all(names(renew_2021), " ", "_")  
names(renew_2021) <- str_replace_all(names(renew_2021), "%", "percent")  
  
names(renew_2022) <- tolower(names(renew_2022))  
names(renew_2022) <- str_replace_all(names(renew_2022), " ", "_")  
names(renew_2022) <- str_replace_all(names(renew_2022), "%", "percent")  
  
names(renew_2023) <- tolower(names(renew_2023))  
names(renew_2023) <- str_replace_all(names(renew_2023), " ", "_")  
names(renew_2023) <- str_replace_all(names(renew_2023), "%", "percent")
```

Part 3: Joining / Pivoting Datasets for Analysis

```
library(tidyverse)  
setwd("c:/Users/light/downloads/ev-power-oscarcheng-1/data")  
# Read & clean column names  
# Read and clean each dataset  
renew_2021 <- read_csv("renew-use-2021.csv") |>  
  rename(state = State, energy_source = Energy_Source, renewable_use =
```

```
Renewable_Use_2021) |>  
  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("renew-use-2022.csv") |>  
  rename(state = State, energy_source = Energy_Source, renewable_use =  
Renewable_Use_2022) |>  
  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("renew-use-2023.csv") |>  
  rename(state = State, energy_source = Energy_Source, renewable_use =  
Renewable_Use_2023) |>  
  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.

```
renew_all <- renew_2021 |>
  full_join(renew_2022, by = names(renew_2021)) |>
  full_join(renew_2023, by = names(renew_2021))

head(renew_all)
```

```
# A tibble: 6 × 4
  state energy_source renewable_use year
<chr> <chr>          <chr>    <dbl>
1 AK    Biomass         ≈3153    2021
2 AK    Geothermal       186 MMBtu 2021
3 AK    Hydropower       5763 about 2021
4 AK    Solar Energy     ~45      2021
5 AK    Wind Energy      451 USD   2021
6 AL    Biomass          198543 est. 2021
```

Part 4: Mapping Visualization