

Are you really driving clean?

Question: *What is the predominant renewable energy source utilized for on-site electricity generation for electric charging stations across the United States of America?*

Introduction: *In the 21st century, electric vehicles (EVs) represent a pivotal aspect of the future of transportation. Beyond the act of driving electric vehicles, it becomes imperative to comprehend the sourcing of electricity for EV stations and assess the environmental cleanliness of such sources. Our focus is directed towards examining how different states across the United States procure their electricity, thereby discerning the implications associated with the advantages that these states may derive.*

To conduct this analysis, we will utilize two distinct datasets. The first dataset encompasses a shapefile containing geometric points utilized for mapping all EV stations across the states. The second dataset, in CSV format, contains information regarding the diverse forms of energy fueling these EV stations.

In reference to the shapefile, we leverage the following columns:

- FUEL_TYPE_: (string) The type of alternative fuel the station provides (CNG/Electric/LPG etc.)
- STATE: (string) The two character code for the U.S. or Canadian state
- COUNTRY: (string) The two character country code of the station's location

Concerning the stations.csv dataset, the relevant columns include:

- COUNTRY: (string) The two character country code of the station's location
- STATE: (string) The two character code for the U.S. or Canadian state
- FUEL_TYPE_CODE: (string) The type of alternative fuel the station provides (CNG/Electric/LPG etc.)
- EV_ON_SITE_RENEWABLE_SOURCE: (string) For electric stations, the type of renewable energy used to generate electricity on-site (SOLAR/WIND/HYDRO etc.)

Approach: *Through the application of geospatial analysis, it becomes apparent how states differ in their procurement of electric vehicle (EV) charging stations and the overall prevalence of such stations. Initial analysis involves examining the spatial distribution of EV stations across various states. Subsequently, a comprehensive analysis delves into the percentage of EV stations within each state that derive their power from environmentally sustainable, green energy sources. To refine focus on clean energy producing minimal emissions, a decision to leave Wastewater and Landfill- related data was left out of the analysis.*

Analysis:

```
# We gather the shapefile to use the point geometry
download.file('https://raw.githubusercontent.com/rfordatascience/tidytuesday/master/data/2022/2022-03-01/Alternative_Fueling_Stations.zip',
             "myshape.zip")
unzip("myshape.zip")
shapefile <- st_read("Alternative_Fueling_Stations.shp")
```

```
## Reading layer `Alternative_Fueling_Stations' from data source
## `D:\MSDS UT Austin\Fall 23\Data Viz\Projects\Alternative_Fueling_Stations.shp'
## using driver `ESRI Shapefile'
## Simple feature collection with 59927 features and 68 fields
## Geometry type: POINT
## Dimension: XY
## Bounding box: xmin: -159.7886 ymin: 0 xmax: 0 ymax: 64.85247
## Geodetic CRS: WGS 84
```

```
# getting the files ready & filter outliers from the shapefile
```

```
ammended_shapefile <- shapefile %>%
  filter(
    COUNTRY == "US" &
    FUEL_TYPE_ == "ELEC" &
    !STATE %in% c("AK", "HI", "ON", "DC") &
    between(LATITUDE, 23, 50) &
    between(LONGITUDE, -125.0, -66.9)
  ) %>%
  select(FUEL_TYPE_, STATE, COUNTRY) %>%
  na.omit()
```

```
# use this built_in package to get the geometry of the states
```

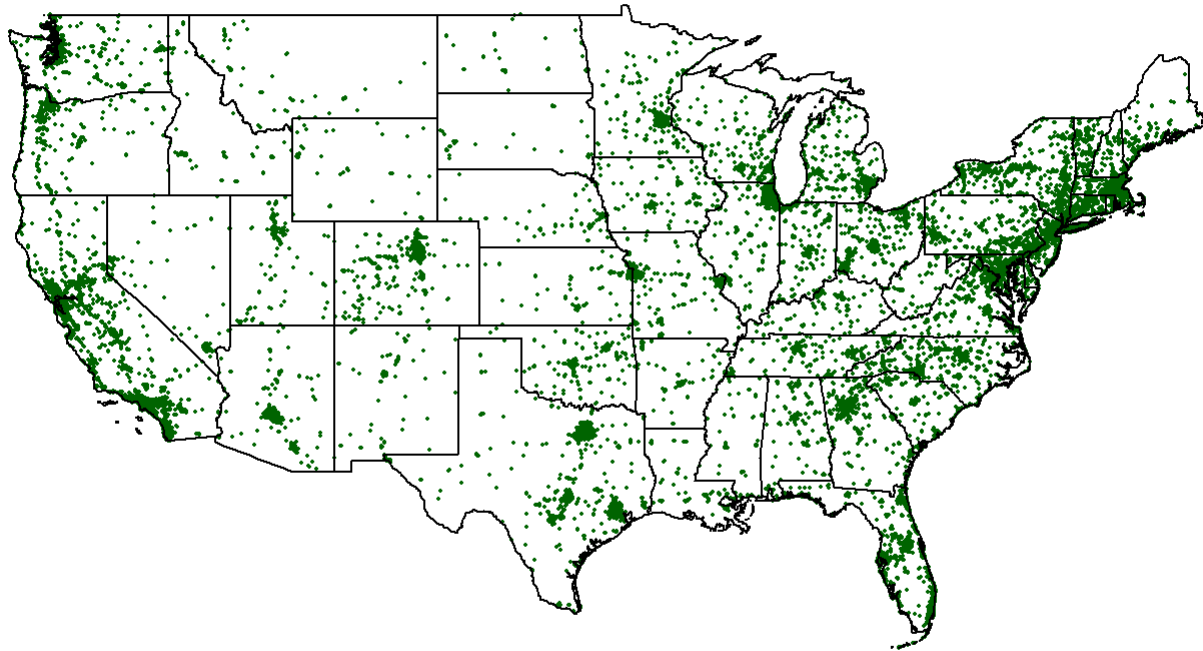
```
sf_us <- ne_states(
  country = "United States of America",
  returnclass = 'sf'
)
# changing names and selecting relevant columns
sf_us <- sf_us %>%
  select(geometry, postal, code_local) %>%
  filter(!code_local %in% c("US02", "US15")) %>%
  rename("STATE" = "postal")
```

```
# join the datasets and plot the geomteries together
```

```
# points_us <- st_join(ammended_shapefile, left = TRUE, sf_us["STATE"])

par(mfrow = c(1, 1), mar = c(0.7, 0.7, 0.7, 0.7))
plot(ammended_shapefile$geometry, pch = 16, cex = 0.25, col = "darkgreen", bg = "white", main =
"Distribution of EV stations in the US")
plot(sf_us$geometry, border="black", col=NA, add=T)
```

Distribution of EV stations in the US



Next we analyze the data further.

```
stations <- readr::read_csv('https://raw.githubusercontent.com/rfordatascience/tidytuesday/master/data/2022/2022-03-01/stations.csv')
```

```
## Warning: One or more parsing issues, call `problems()` on your data frame for details,  
## e.g.:  
##   dat <- vroom(...)  
##   problems(dat)
```

```
## Rows: 59927 Columns: 70  
## — Column specification —————  
## Delimiter: ","  
## chr (43): FUEL_TYPE_CODE, STATION_NAME, STREET_ADDRESS, INTERSECTION_DIRECTI...  
## dbl (15): X, Y, OBJECTID, EV_LEVEL1_EVSE_NUM, EV_LEVEL2_EVSE_NUM, EV_DC_FAST...  
## lgl (12): PLUS4, EV_OTHER_INFO, HYDROGEN_STATUS_LINK, LPG_PRIMARY, E85_BLEND...  
##  
## 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.
```

```

stations_clean <- stations %>%
  select(COUNTRY, STATE, FUEL_TYPE_CODE, EV_ON_SITE_RENEWABLE_SOURCE) %>%
  filter(
    !STATE %in% c("AK", "HI", "ON", "DC", "PR") &
    !EV_ON_SITE_RENEWABLE_SOURCE %in% c("WASTEWATER", "LANDFILL")
  )

# count the ev stations by state
stations_sum <- stations_clean %>%
  mutate(EV_ON_SITE_RENEWABLE_SOURCE = ifelse(is.na(EV_ON_SITE_RENEWABLE_SOURCE), "NONE", EV_ON_
SITE_RENEWABLE_SOURCE)) %>%
  filter(FUEL_TYPE_CODE == "ELEC") %>%
  group_by(STATE) %>%
  summarise(count = n())

stations_sum

```

```

## # A tibble: 48 × 2
##   STATE count
##   <chr> <int>
##  1 AL      277
##  2 AR      165
##  3 AZ      891
##  4 CA     14607
##  5 CO      1628
##  6 CT       533
##  7 DE       134
##  8 FL      2668
##  9 GA      1588
## 10 IA       279
## # i 38 more rows

```

```

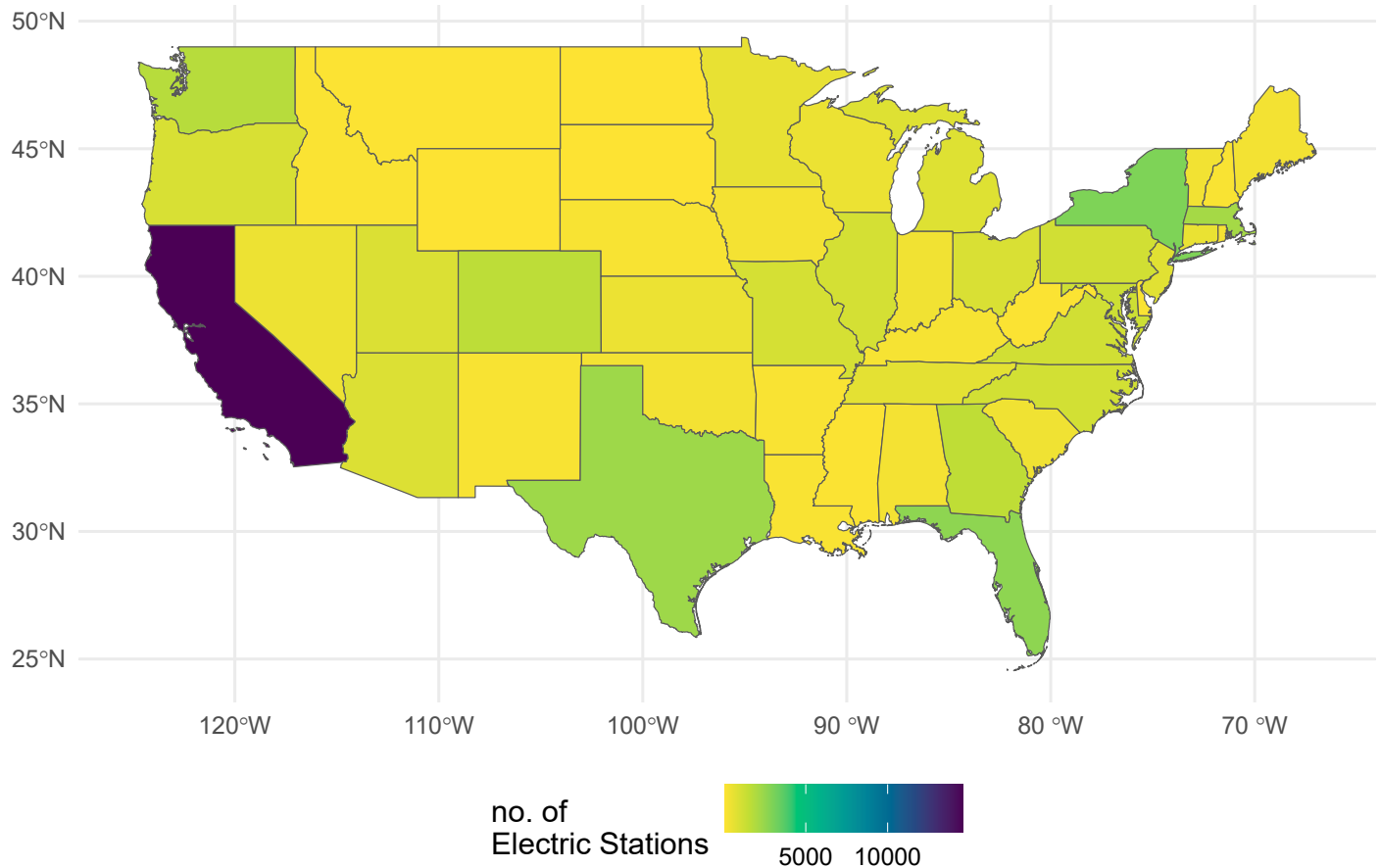
# Left join on STATE to add the geometry for the stations data
joined_data <- inner_join(sf_us, stations_sum, by = 'STATE')

# do some maps with some cool colors, interactive plots
p <- joined_data %>%
  # exclude Alaska (US02), Hawaii (US15)
  ggplot() +
  geom_sf_interactive(
    aes(fill = count, data_id = STATE, tooltip = paste(STATE, ": ", count))
  )+
  scale_fill_continuous_sequential(palette = "Viridis", rev = TRUE) +
  theme_minimal() +
  ggtitle("Spatial distribution of Electric Stations across States") +
  labs(fill = paste("no. of \nElectric Stations")) +
  theme(legend.position = "bottom")

girafe(code = print(p))

```

Spatial distribution of Electric Stations across States



```
stations_ev_source <- stations_clean %>%
  mutate(EV_ON_SITE_RENEWABLE_SOURCE = ifelse(is.na(EV_ON_SITE_RENEWABLE_SOURCE), "NONE", EV_ON_SITE_RENEWABLE_SOURCE)) %>%
  filter(FUEL_TYPE_CODE == "ELEC", !EV_ON_SITE_RENEWABLE_SOURCE %in% c("WASTEWATER", "LANDFILL")) %>%
  count(STATE, EV_ON_SITE_RENEWABLE_SOURCE) %>%
  pivot_wider(names_from = "EV_ON_SITE_RENEWABLE_SOURCE", values_from = "n", values_fill = 0)

stations_ev_source <- stations_ev_source %>%
  mutate(Total_Renewable = rowSums(select(., -STATE))) %>%
  mutate(across(NONE:WIND, ~ . / Total_Renewable * 100)) %>%
  rename("NOTHINGU" = "NONE")

stations_ev_source
```

```
## # A tibble: 48 × 6
##   STATE NOTHINGU SOLAR  HYDRO  WIND Total_Renewable
##   <chr>      <dbl> <dbl>  <dbl> <dbl>      <dbl>
## 1 AL          100    0    0      0        277
## 2 AR          99.4 0.606 0      0        165
## 3 AZ          100    0    0      0        891
## 4 CA          99.4 0.602 0      0       14607
## 5 CO          99.3 0.676 0.0614 0        1628
## 6 CT          98.7 1.31  0      0        533
## 7 DE          100    0    0      0        134
## 8 FL          99.9 0.112 0      0       2668
## 9 GA          99.6 0.378 0      0       1588
## 10 IA         98.9 1.08  0      0        279
## # i 38 more rows
```

The following maps represent the distribution of different renewable energy sources across the states.

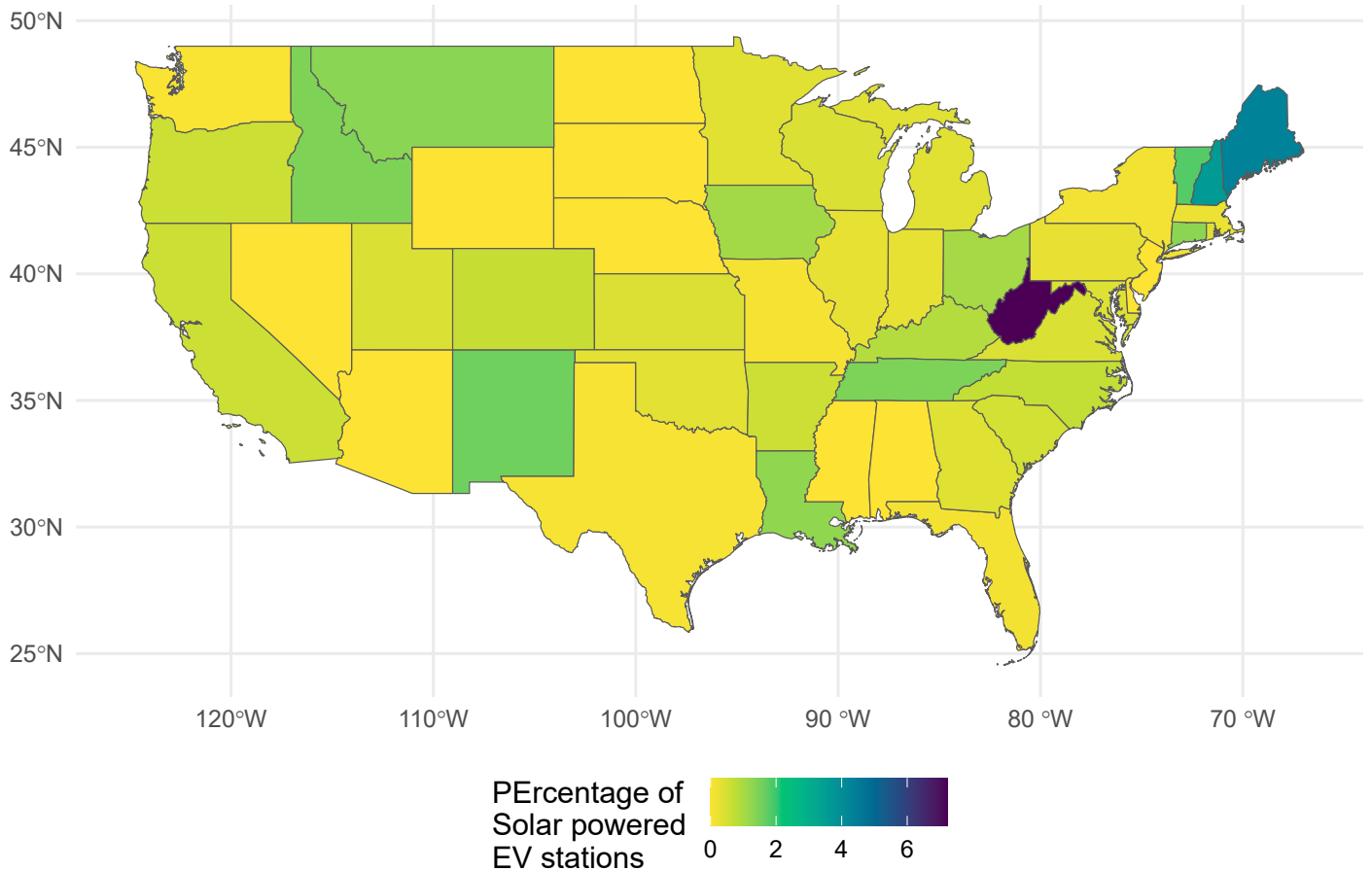
```
# List of renewable energy sources to create heatmaps for
new_join <- inner_join(sf_us, stations_ev_source, by = 'STATE')

# SOLAR
solar <- new_join %>%

ggplot() +
geom_sf_interactive(
  aes(fill = SOLAR, data_id = STATE, tooltip = paste(STATE, ": ", SOLAR))
)+
scale_fill_continuous_sequential(palette = "Viridis", rev = TRUE) +
theme_minimal() +
ggtitle("7% West Virginia EV stations are solar powered") +
labs(fill = paste("PErcentage of \nSolar powered \nEV stations")) +
theme(legend.position = "bottom")

girafe(code = print(solar))
```

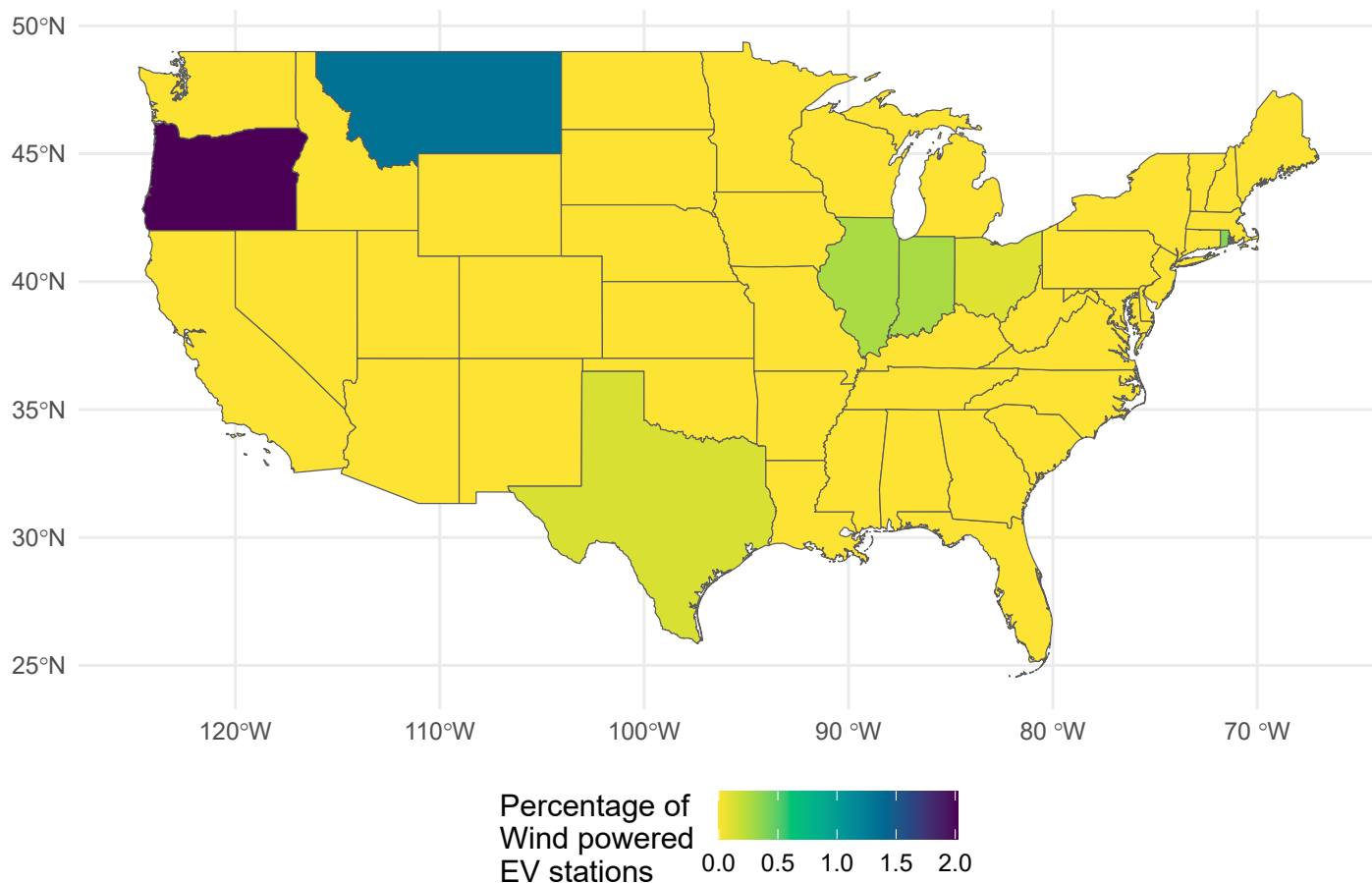
7% West Virginia EV stations are solar powered



```
# WIND
wind <- new_join %>%
  ggplot() +
  geom_sf_interactive(
    aes(fill = WIND, data_id = STATE, tooltip = paste(STATE, ": ", WIND))
  ) +
  scale_fill_continuous_sequential(palette = "Viridis", rev = TRUE) +
  theme_minimal() +
  ggtitle("2% of Oregon EV Stations are wind powered") +
  labs(fill = paste("Percentage of \nWind powered \nEV stations")) +
  theme(legend.position = "bottom")

girafe(code = print(wind))
```

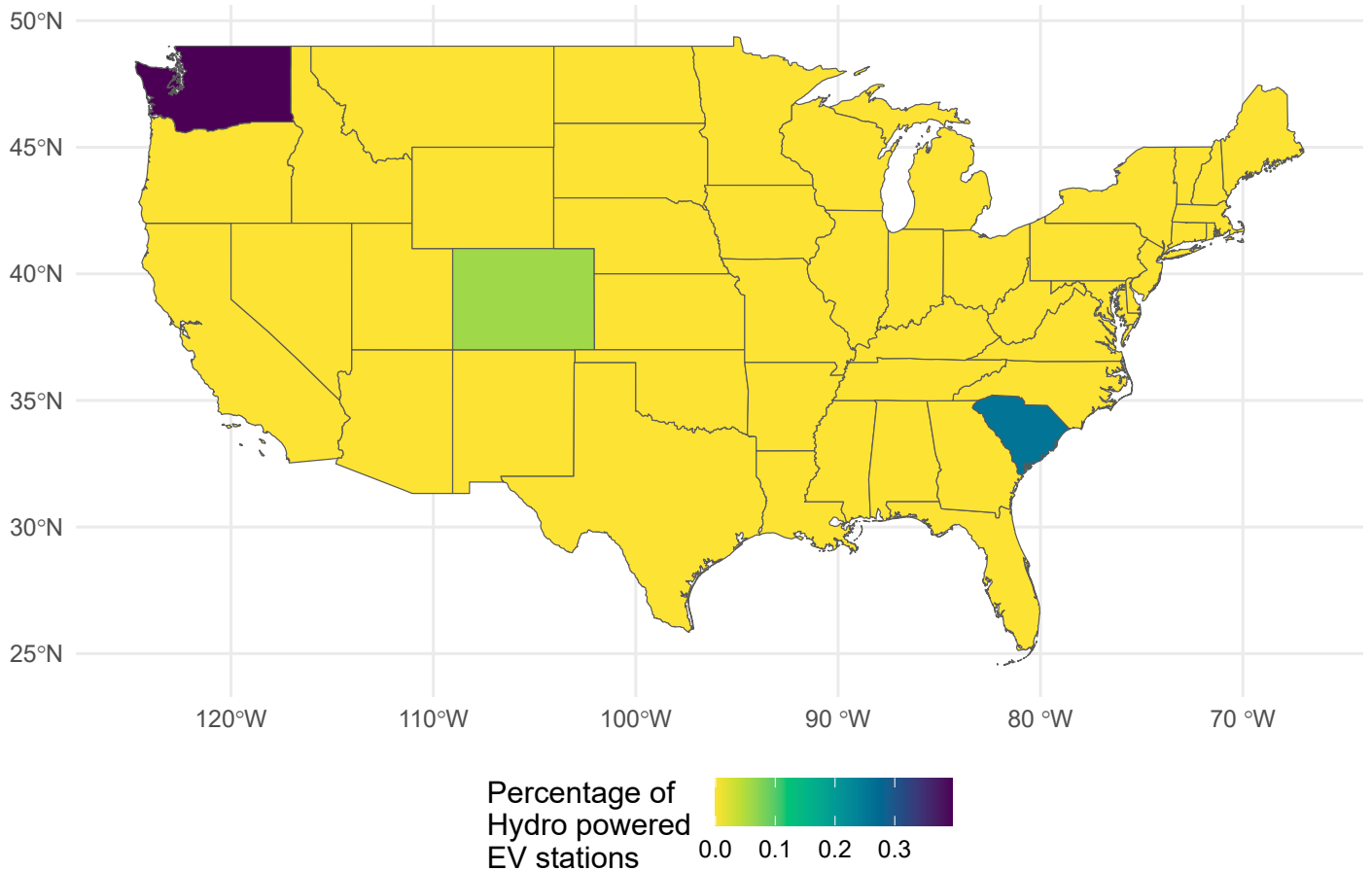
2% of Oregon EV Stations are wind powered



```
# HYDRO
hydro <- new_join %>%
ggplot() +
geom_sf_interactive(
  aes(fill = HYDRO, data_id = STATE, tooltip = paste(STATE, ": ", HYDRO))
)+
scale_fill_continuous_sequential(palette = "Viridis", rev = TRUE) +
theme_minimal() +
labs(fill = paste("Percentage of \nHydro powered \nEV stations")) +
theme(legend.position = "bottom") +
ggtitle("0.39% Washington EV stations are Hydro powered")

girafe(code = print(hydro))
```

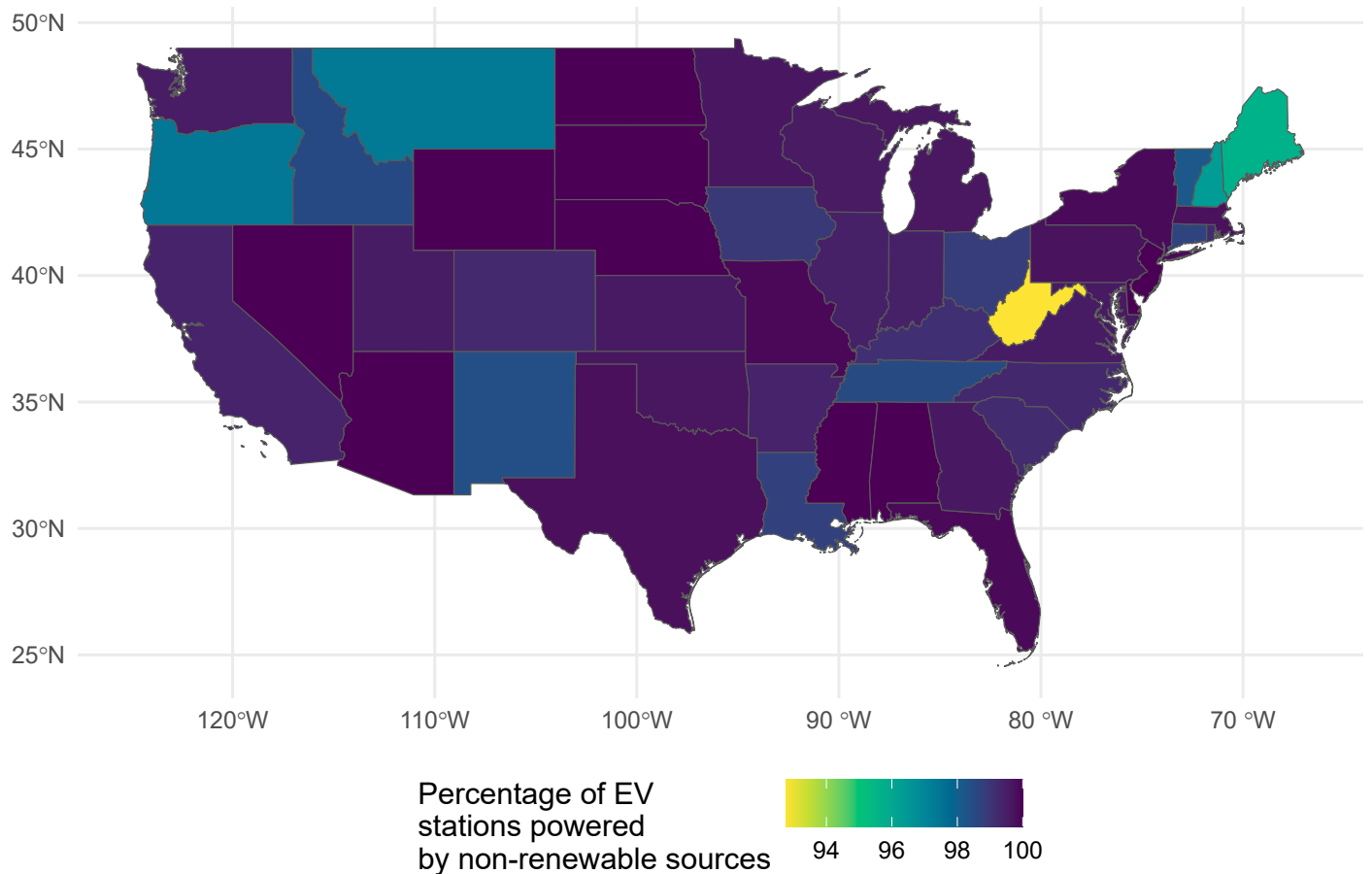

0.39% Washington EV stations are Hydro powered



```
# NON-RENEWABLE
non_green <- new_join %>%
ggplot() +
geom_sf_interactive(
  aes(fill = NOTHINGU, data_id = STATE, tooltip = paste(STATE, ": ", NOTHINGU))
)+
scale_fill_continuous_sequential(palette = "Viridis", rev = TRUE) +
theme_minimal() +
labs(fill = paste("Percentage of EV \nstations powered \nby", "non-renewable sources")) +
ggtitle("Most states heavily rely on non-renewable sources") +
theme(legend.position = "bottom")
#title("Percentage of "+ HYDRO + " energy produced across the states")

girafe(code = print(non_green))
```

Most states heavily rely on non-renewable sources



Discussion: *In the examination of electric vehicle (EV) stations across various states, it was observed that 7% of stations in West Virginia, 2% in Oregon, and 0.39% in Washington are powered by solar, wind, and hydro energy, respectively. Notably, despite California having the highest number of EV stations, a comparatively limited proportion is sourced from renewable energy.*

Amidst the electric vehicle surge, the haunting question lingers: Did you unknowingly fuel your EV with non-renewable energy? Are you really driving clean? And we have miles to go before we green. We must urgently propel the relentless journey towards green technology.