PPOL 646 - 01 Interim Visualization Assignment 5

Xiyu. Zhang

TOTAL POINTS

2.67 / 0

QUESTION 1

1 Visualization 1 **0 / 0**

✓ **+ 0 pts** *Really good work!*

+ 0 pts Choice of plot type (bar chart, line chart, etc.) should be appropriate to the type of data.

+ 0 pts Graph should clearly indicate the unit of analysis (U.S. adults?, states? counties? countries?) and each variable's unit of measurement. And perhaps more information about the key measure(s). For example, are the data annual or monthly? A percentage? Per capita? Pounds? Kilograms? US dollars (nominal or constant)?

+ 0 pts Graph should include a title that indicates the content of the visualization.

+ 0 pts Is there a particular substantive insight or finding that you want readers to take away from this visualization? If so, consider adjusting the title and/or adding annotation that helps readers to discover it more quickly. If not, consider posing a question or two that invites the reader to look for potential findings that might not be immediately obvious.

✓ **+ 0 pts** *Titles, subtitles, and annotation should have proper capitalization, spelling, punctuation, and, as appropriate, grammar. If your title/subtitle is effectively a declarative sentence, capitalize it the way you would capitalize a sentence. The title does not, however, require a period at the end. The first occurrence of an acronym should usually be spelled out, unless the acronym is very familiar to the broader audience, e.g., "US" for "United States."*

+ 0 pts Axes and tic marks (if applicable) should be appropriately labeled and have an appropriate range. Axis and tic mark labels should have proper capitalization, spelling, punctuation, and, as appropriate, grammar. Generally, capitalize only the first word of an axis label unless it includes a proper noun. When possible, axis labels should be horizontal. Exclude unnecessary axis names, such as "Year" if the meaning is already clear enough for readers to understand

+ 0 pts Graph should credit the source of the data. Acronyms in the credit should usually be spelled out, unless they are very familiar to the broader audience, e.g., "US" for "United States." The data source credit should be an institution (such as "World Bank" or "US Bureau of the Census" and possibly even the name of a study or a data set produced by that institution. A URL can be unstable.

✓ **+ 0 pts** *Graphs that use more than one value of a mark channel (such as hues, textures, color values, or shapes) should explain the meaning of each value by using labels, a legend, or some other key. The legend should be easy to understand and use correct* *spelling, capitalization, and punctuation.*

✓ **+ 0 pts** *Visual encoding choices (e.g., using hues to represent values of a categorical variable, or position, length, area to represent quantitative variables) could be strengthened.*

+ 0 pts Could improve aesthetics and/or general readability.

+ 0 pts Strengthen the visual hierarchy of the text elements by ensuring that, for example, labels or legends do not compete with the graph title for the reader's attention.

+ 0 pts Could improve visual hierarchy to feature the encoded data more prominently than accessory elements, such as labels, legends, reference lines/rectangles, etc.

+ 0 pts Graph should be free of irrelevant or redundant information.

+ 0 pts The strongest visualizations will provide the reader with insights beyond those obtainable from looking at a table of the numbers used to create the visualization.

+ 0 pts The strongest visualizations will reflect conscious design choices and more exercise of control instead of relying on the software defaults, such as default color choices and the grey background. Graphs should be in color. Lines that encode data should often be thicker than the default from R or ggplot.

+ 0 pts Type sizes of similar graph elements (e.g., value labels, legend labels) should be the same.

+ 0 pts Graph title or annotation should more prominently note the represented population (e.g., USA, China, world, Texas) somewhere. Specify "U.S. States" instead of only "States".

+ 0 pts Graph title or annotation should more prominently indicate the time period represented.

+ 0 pts Stronger graphs will exclude unnecessary axis lines, such as a top axis line that has neither tic marks nor tic labels.

+ 0 pts Data selected for use should match the substantive claim(s) in the annotation.

+ 0 pts A very good application of the code that we learned in class. The strongest work will demonstrate extending the work with more of your own code.

+ 0 pts Reduce the number of alignment points for a cleaner, more professional presentation. See chapter 5 of the Berinato book for details.

+ 0 pts Consider using a sans sarif type face for a more contemporary, professional-looking presentation. (See, for example the data visualizations from the World Bank or Deloitte.)

+ 0 pts Consider reordering the categorical data values in some way other than alphabetical. You can use the fct_reorder() function from the forcats package to do this. See chapter 15 of "R for Data Science" and let me know if you have any questions about it.

+ 0 pts Keep Mackinlay's "expressiveness principle" in mind. The most important variable(s), given your substance claim, should be represented by the marks/channels that readers can most quickly and accurately decode/interpret.

+ 0 pts Work based on sample data, such as survey data, should indicate whether the figures were calculated using sample weights.

This work looks nice. But the assignment instructions, titled "Two More Visualizations in R," states in the first paragraph: "Please create both of the graphs using R and refine each graph in Adobe Illustrator."

The small multiples style maps effectively communicate your point about the decrease in the EV to charger ratios over time. Consider using a continuous color scale for Evs per charger rather than green to red, cropping the bottom maps to remove the little country bits in the lower righthand corners, and defining EVs as an acronym for "electric vehicle."

1 This kind of graph variation could confuse readers. Please use the map idiom consistently for all states.

QUESTION 2

2 Visualization 2 **2.67 / 0**

+ **0 pts** A terrific original visualization in R!

+ **0 pts** Choice of plot type (bar chart, line chart, etc.) should be appropriate to the type of data.

✓ **+ 0 pts** *Graph should clearly indicate the unit of analysis (U.S. adults?, states? counties? countries?) and each variable's unit of measurement. And perhaps more information about the key measure(s). For example, are the data annual or monthly? A percentage? Per capita? Pounds? Kilograms? US dollars (nominal or constant)?*

+ **0 pts** Graph should include a title that indicates the content of the visualization.

+ **0 pts** Is there a particular substantive insight or finding that you want readers to take away from this visualization? If so, consider adjusting the title and/or adding annotation that helps readers to discover it more quickly. If not, consider posing a question or two that invites the reader to look for potential findings that might not be immediately obvious.

+ **0 pts** Titles, subtitles, and annotation should have proper capitalization, spelling, punctuation, and, as appropriate, grammar. If your title/subtitle is effectively a declarative sentence, capitalize it the way you would capitalize a sentence. The title does not, however, require a period at the end. The first occurrence of an acronym should usually be spelled out, unless the acronym is very familiar to the broader audience, e.g., "US" for "United States."

+ **0 pts** Axes and tic marks (if applicable) should be appropriately labeled and have an appropriate range. Axis and tic mark labels should have proper capitalization, spelling, punctuation, and, as appropriate, grammar. Generally, capitalize only the first word of an axis label unless it includes a proper noun. When possible, axis labels should be horizontal. Exclude unnecessary axis names, such as "Year" if the meaning is already clear enough for readers to understand

+ **0 pts** Graph should credit the source of the data. Acronyms in the credit should usually be spelled out, unless they are very familiar to the broader audience, e.g., "US" for "United States." The data source credit should be an institution (such as "World Bank" or "US Bureau of the

Census" and possibly even the name of a study or a data set produced by that institution. A URL can be unstable.

+ 0 pts Graphs that use more than one value of a mark channel (such as hues, textures, color values, or shapes) should explain the meaning of each value by using labels, a legend, or some other key. The legend should be easy to understand and use correct spelling, capitalization, and punctuation.

+ 0 pts Visual encoding choices (e.g.,using different hues to represent values of a categorical variable, or position, length, area to represent quantitative variables) could be strengthened.

✓ **+ 0 pts** *Could improve aesthetics and/or general readibility.*

+ 0 pts Strengthen the visual hierarchy of the text elements by ensuring that, for example, labels or legends do not compete with the graph title for the reader's attention.

+ 0 pts Could improve visual hierarchy to feature the encoded data more prominently than accessory elements, such as labels, legends, reference lines/rectangles, etc.

+ 0 pts Graph should be free of irrelevant or redundant information.

+ 0 pts The strongest visualizations will provide the reader with insights beyond those obtainable from looking at a table of the numbers used to create the visualization.

+ 0 pts The strongest visualizations will reflect conscious design choices instead of relying on the software defaults. For example, lines that encode data should often be thicker than the default from R. Graphs should be in color.

+ 0 pts Type sizes of similar graph elements (e.g., value labels, legend labels) should be the same.

+ 0 pts Graph title or annotation should more prominently note the represented population (e.g., USA, China, world, Texas) somewhere. Specify "U.S. States" instead of only "States".

+ 0 pts Graph title or annotation should more prominently indicate the time period represented.

+ 0 pts Stronger graphs will exclude unnecessary axis lines, such as a top axis line that has neither tic marks nor tic labels.

+ 0 pts Data selected for use should match the substantive claim(s) in the annotation.

+ 0 pts A very good application of the code that we learned in class. The strongest work will demonstrate extending the work with more of your own code.

✓ **+ 0 pts** *Reduce the number of alignment points for a cleaner, more professional presentation. See chapter 5 of the Berinato book for details.*

+ 0 pts Consider using a sans sarif type face for a more contemporary, professional-looking presentation. (See, for example the data visualizations from the World Bank or Deloitte.)

+ 0 pts Consider reordering the categorical variable values in some way other than alphabetical. You can use the fct_reorder() function from the forcats package to do this. See chapter 15 of "R for Data Science" and let me know if you have any questions about it.

+ 0 pts Keep Mackinlay's "expressiveness principle" in mind. The most important variable(s), given your substance claim, should be represented by the marks/channels that readers can most quickly and accurately decode/interpret.

+ 0 pts Work based on sample data, such as survey data, should indicate whether the figures were calculated using sample weights.

+ 0 pts Please submit the pages in the order specified in the assignment prompt. When the pages are out of order, it increases the amount of time required to review the assignment. As the prompt indicates, assignments submitted out of order forfeit 0.2 points.

+ 2.67 *Point adjustment*

Great work! I particularly like your color choices and that you reordered states from least to most adequate deployment of EV charging points. A few suggestions to consider are centering the state labels under the bars, reversing the order of years so the plot can be read from top to bottom, and aligning the colors for the ratios better with the recommended ratio ways to align ratios better with the recommended ratio (e.g., you could do the first color group as 0-10 to easily show what states meet the recommended ratio of 10).

01

Statement of Purpose

Typically, the deployment of electric vehicle charging infrastructure in a country lags behind the purchase of electric vehicles[1]. The same holds in the United States. This set of choropleth maps shows that as time went by, the electric vehicle (EV) to charger ratios were decreasing during 2018 – 2021. That indicates more charger points were built as time went by compared to the purchase of electric vehicles in the US states.

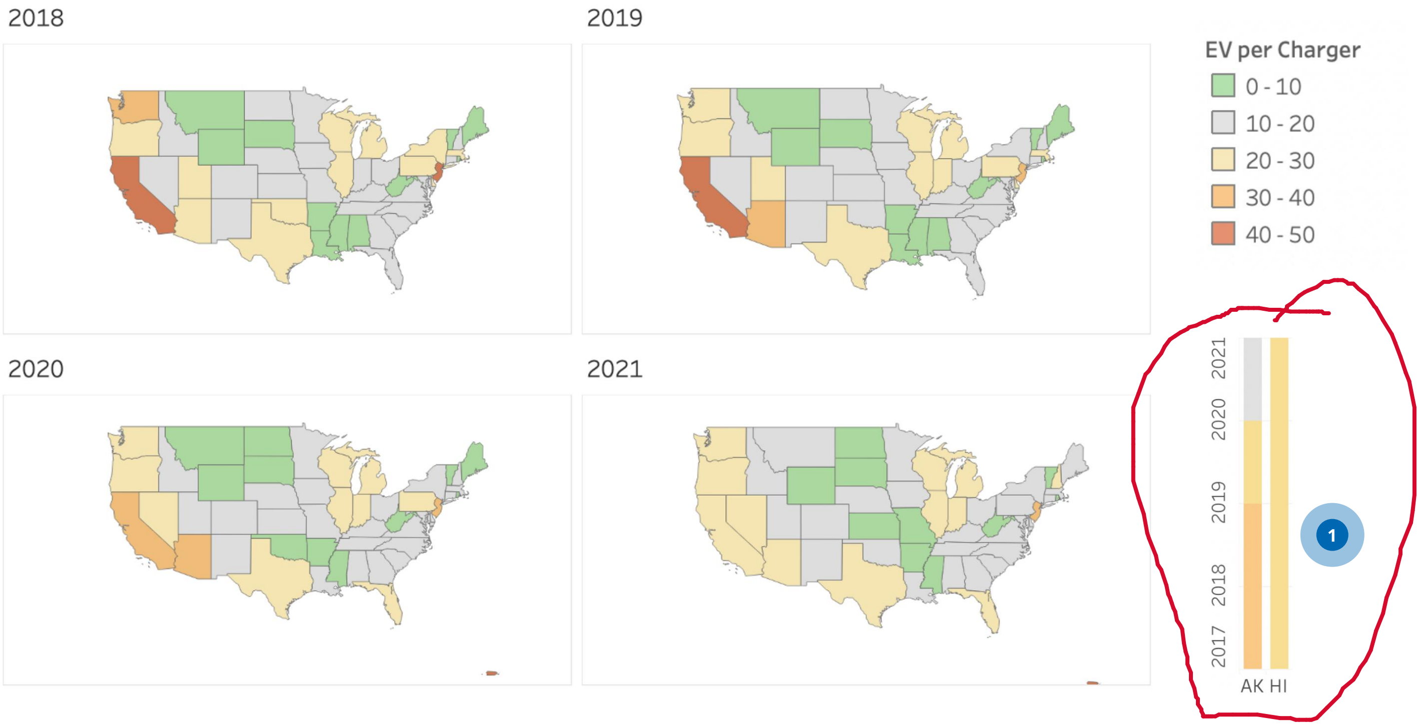
However, the deployment of EV charging infrastructure is still insufficient according to the most recent data. In 2014, the Alternative Fuel Infrastructure Directive (AFID) recommended that countries reach 10 electric light-duty vehicles (LDVs) per public charger by 2020. This set of maps marks the US states with warm tones (yellow, orange, and dark orange) while the EV-to-charger ratios are no less than 20, which is a threshold even higher than the recommendation, to indicate the US states that are still urgent for EV charging infrastructures. On the contrary, the US states marked as green or gray indicate their sufficient or mediocre deployment level of EV charging infrastructures compared to the numbers of registered EVs in the states.

Note: There is no R code for data visualization since this set of choropleth maps is created by Tableau. The R code here only includes the data wrangling process.

[1] IEA. (2022). Global EV Outlook 2022. https://www.iea.org/reports/global-ev-outlook-2022

U.S. Commits to Deploy Electric Vehicle Charging Infrastructure

the EV to Charger Ratio is decreasing in U.S. states, 2018 - 2021



According to Alternative Fuel Infrastructure Directive (AFID), countries were recommended to reach 10 electric light-duty vehicles (LDVs) per public charger by 2020. The US states did not meet the standard and EV-to-charger ratios larger than 20 are marked with warm tone (yellow, orange, and dark orange). Although the EV-to-charger ratios were decreasing, there were still plenty of states that were insufficient of electric vehicle charging infrastructures.

Data Source: Alternative Fuels Data Center: Vehicle Registration Counts by State, Alternative Fueling Station Counts by State.

```r
# Load packages ----------------------------------------------------------

library(tidyverse)
library(readxl)
library(writexl)

# Load data --------------------------------------------------------------

# Stations

stations <-
  read_csv('assignment5/alt_fuel_stations.csv')

# EV registration

LDV_2021 <-
  read_excel('assignment5/2021_LDV_registration.xlsx')

LDV_2020 <-
  read_excel('assignment5/2020_LDV_registration.xlsx')

LDV_2019 <-
  read_excel('assignment5/2019_LDV_registration.xlsx')

LDV_2018 <-
  read_excel('assignment5/2018_LDV_registration.xlsx')

# State name and code

state <-
  read_csv('assignment5/state.csv') %>%
  select(-'abbrev') %>%
  rename('State' = 'state')

# Data for 2021 ---------------------------------------------------------

# Electric vehicle supply equipment (EVSE, charging points) in 2021 by state

charger_2021 <-
  stations %>%

  # the open date is ealier than 2022, so this is the existence of chargers in
  # 2021

  filter(year(`Open Date`) < 2022) %>%

  # 'ON' is in Canada

  filter(State != 'ON') %>%
  select(`Station Name`, `EV Level1 EVSE Num`, `EV Level2 EVSE Num`,
         `EV DC Fast Count`, State) %>%
```

```r
  # 'NA' indicates that there is no this type of EVSE in the station, so there
  # is 0 this type of EVSE in the station

  replace(is.na(.), 0) %>%

  # the total number of charging points in a charging station is the sum of
  # 3 types of EVSEs

  mutate(EVSE = `EV Level1 EVSE Num` + `EV Level2 EVSE Num` +
           `EV DC Fast Count`) %>%

  # calculate the total number of EVSE in each state in 2021

  group_by(State) %>%
  summarise(EVSE_state = sum(EVSE))

# the Plug-in Electric Vehicles in 2021 by state

PEV_2021 <-
  LDV_2021 %>%

  # exclude the total amount

  filter(State != 'United States') %>%

  # only calculate the types of EV that needed chargers

  mutate(PEV = `Electric (EV)` + `Plug-In Hybrid Electric (PHEV)`) %>%
  select(State, PEV)

# Merge the needed data for 2021

merge_2021 <-
  full_join(state, PEV_2021) %>%
  select(code, PEV) %>%
  rename(State = code) %>%
  full_join(charger_2021) %>%
  mutate(EV_per_charger = PEV / EVSE_state) %>%
  select(State, EV_per_charger) %>%
  mutate(year = 2021)

# Data for 2020 ------------------------------------------------------------

# Electric vehicle supply equipment (EVSE, charging points) in 2020 by state

charger_2020 <-
  stations %>%
  filter(year(`Open Date`) < 2021) %>%
  filter(State != 'ON') %>%
  select(`Station Name`, `EV Level1 EVSE Num`, `EV Level2 EVSE Num`,
         `EV DC Fast Count`, State) %>%
  replace(is.na(.), 0) %>%
```

```r
  mutate(EVSE = `EV Level1 EVSE Num` + `EV Level2 EVSE Num` +
           `EV DC Fast Count`) %>%
  group_by(State) %>%
  summarise(EVSE_state = sum(EVSE))

# the Plug-in Electric Vehicles in 2020 by state

PEV_2020 <-
  LDV_2020 %>%
  filter(State != 'United States') %>%
  mutate(PEV = `Electric (EV)` + `Plug-In Hybrid Electric (PHEV)`) %>%
  select(State, PEV)

# Merge the needed data for 2020

merge_2020 <-
  full_join(state, PEV_2020) %>%
  select(code, PEV) %>%
  rename(State = code) %>%
  full_join(charger_2020) %>%
  mutate(EV_per_charger = PEV / EVSE_state)%>%
  select(State, EV_per_charger) %>%
  mutate(year = 2020)

# Data for 2019 ------------------------------------------------------------

# Electric vehicle supply equipment (EVSE, charging points) in 2019 by state

charger_2019 <-
  stations %>%
  filter(year(`Open Date`) < 2020) %>%
  filter(State != 'ON') %>%
  select(`Station Name`, `EV Level1 EVSE Num`, `EV Level2 EVSE Num`,
         `EV DC Fast Count`, State) %>%
  replace(is.na(.), 0) %>%
  mutate(EVSE = `EV Level1 EVSE Num` + `EV Level2 EVSE Num` +
           `EV DC Fast Count`) %>%
  group_by(State) %>%
  summarise(EVSE_state = sum(EVSE))

# the Plug-in Electric Vehicles in 2019 by state

PEV_2019 <-
  LDV_2019 %>%
  filter(State != 'United States') %>%
  mutate(PEV = `Electric (EV)` + `Plug-In Hybrid Electric (PHEV)`) %>%
  select(State, PEV)

# Merge the needed data for 2019

merge_2019 <-
  full_join(state, PEV_2019) %>%
```

```
  select(code, PEV) %>%
  rename(State = code) %>%
  full_join(charger_2019) %>%
  mutate(EV_per_charger = PEV / EVSE_state)%>%
  select(State, EV_per_charger) %>%
  mutate(year = 2019)

# Data for 2018 ---------------------------------------------------------------

# Electric vehicle supply equipment (EVSE, charging points) in 2018 by state

charger_2018 <-
  stations %>%
  filter(year(`Open Date`) < 2019) %>%
  filter(State != 'ON') %>%
  select(`Station Name`, `EV Level1 EVSE Num`, `EV Level2 EVSE Num`,
         `EV DC Fast Count`, State) %>%
  replace(is.na(.), 0) %>%
  mutate(EVSE = `EV Level1 EVSE Num` + `EV Level2 EVSE Num` +
           `EV DC Fast Count`) %>%
  group_by(State) %>%
  summarise(EVSE_state = sum(EVSE))

# the Plug-in Electric Vehicles in 2018 by state

PEV_2018 <-
  LDV_2018 %>%
  filter(State != 'United States') %>%
  mutate(PEV = `Electric (EV)` + `Plug-In Hybrid Electric (PHEV)`) %>%
  select(State, PEV)

# Merge the needed data for 2018

merge_2018 <-
  full_join(state, PEV_2018) %>%
  select(code, PEV) %>%
  rename(State = code) %>%
  full_join(charger_2018) %>%
  mutate(EV_per_charger = PEV / EVSE_state)%>%
  select(State, EV_per_charger) %>%
  mutate(year = 2018)

# Merge the data in different years ---------------------------------------

All_years <-
  rbind(merge_2018,
        merge_2019) %>%
  rbind(merge_2020) %>%
  rbind(merge_2021) %>%
  mutate(EV_per_Charger =
           case_when(EV_per_charger <= 10 ~ '0 - 10',
                     EV_per_charger <= 20 ~ '10 - 20',
```

```
                   EV_per_charger <= 30 ~ '20 - 30',
                   EV_per_charger <= 40 ~ '30 - 40',
                   .default = '40 - 50'))

# Export an excel file for Tableau

write_xlsx(All_years, 'assignment5/merge_all.xlsx')
```

02

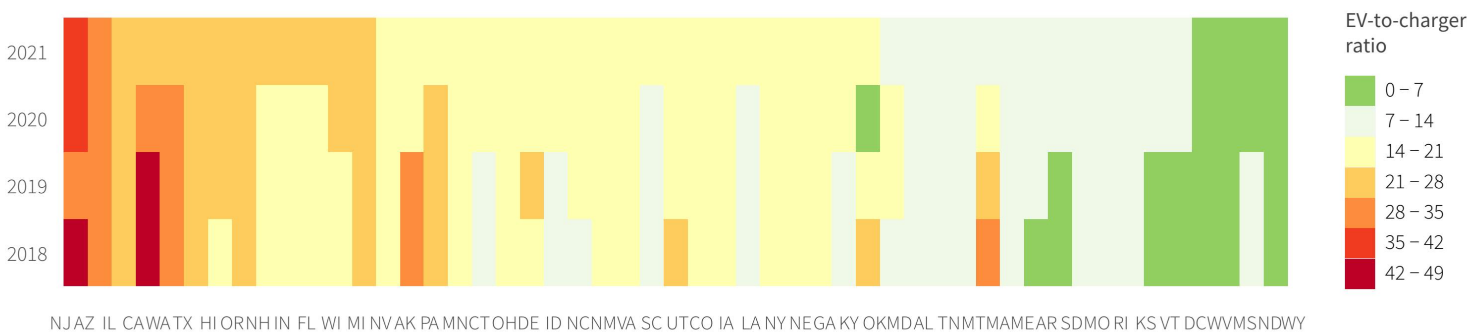
Statement of Purpose

This data set is almost identical to the one generated in the previous section, demonstrating the change in electric vehicle (EV) to charger ratios across U.S. states from 2018 to 2021, with only a few adjustments to fit the graph type better.

However, while a set of choropleth maps are mainly indicating to readers the general trend of the decrease of EV-to-charger ratio across U.S. states, the heat map here focuses on the spectrum of different levels of EV charging infrastructure deployment across U.S. states. On the most left there is New Jersey, with the highest EV-to-charger ratio in 2021, while on the most right there is Wyoming with the lowest ratio. From this aspect, readers could see that New Jersey, Arizona, Illinois, and California are among the states that need EV charging infrastructure the most, while Wyoming, North Dakota, and Mississippi are the states that need the infrastructure the least.

Electric Vehicle Charging Point Deployment by State in the U.S., 2018 – 2021

From left to right: the state with the least (NJ) to most (WY) adequate deployment of EV charging points



From left to right are the states with the least (NJ) to most (WY) adequate deployment of EV charging points.
From 2018 to 2021, every state in the U.S. was committed to deploying electric vehicle charging infrastructure.
The EV-to-charger ratio decreases in most states.
However, most states still do not meet the Alternative Fuel Infrastructure Directive's (AFID) recommendation to reduce the EV-to-charger ratio to 10 by 2020 (white and green tiles in the chart).

Data Source:
Alternative Fuels Data Center:
Vehicle Registration Counts by State
Alternative Fueling Station Counts by State

```r
# Load packages ------------------------------------------------------------

library(tidyverse)
library(ggplot2)
library(readxl)

# Load data ----------------------------------------------------------------

# Stations

stations <-
  read_csv('assignment5/alt_fuel_stations.csv') %>%

  # ON is a Canada state, exclude non-US states.

  filter(State != 'ON')

# EV registration

LDV_2021 <-
  read_excel('assignment5/2021_LDV_registration.xlsx')

LDV_2020 <-
  read_excel('assignment5/2020_LDV_registration.xlsx')

LDV_2019 <-
  read_excel('assignment5/2019_LDV_registration.xlsx')

LDV_2018 <-
  read_excel('assignment5/2018_LDV_registration.xlsx')

# State name and code

state <-
  read_csv('assignment5/state.csv') %>%
  select(-'abbrev') %>%
  rename('State' = 'state')

# Data wrangling -----------------------------------------------------------

# 1. Repeat the process to generate the 'All_years' data frame as same as
# the first graph

# Electric vehicle supply equipment (EVSE, charging points) in 2021 by state

charger_2021 <-
  stations %>%
  filter(year(`Open Date`) < 2022) %>%
  filter(State != 'ON') %>%
  select(`Station Name`, `EV Level1 EVSE Num`, `EV Level2 EVSE Num`,
         `EV DC Fast Count`, State) %>%
  replace(is.na(.), 0) %>%
```

```
  mutate(EVSE = `EV Level1 EVSE Num` + `EV Level2 EVSE Num` +
           `EV DC Fast Count`) %>%
  group_by(State) %>%
  summarise(EVSE_state = sum(EVSE))

# the Plug-in Electric Vehicles in 2021 by state

PEV_2021 <-
  LDV_2021 %>%
  filter(State != 'United States') %>%
  mutate(PEV = `Electric (EV)` + `Plug-In Hybrid Electric (PHEV)`) %>%
  select(State, PEV)

# Merge the needed data for 2021

merge_2021 <-
  full_join(state, PEV_2021) %>%
  select(code, PEV) %>%
  rename(State = code) %>%
  full_join(charger_2021) %>%
  mutate(EV_per_charger = PEV / EVSE_state) %>%
  select(State, EV_per_charger) %>%
  mutate(year = 2021)

# Electric vehicle supply equipment (EVSE, charging points) in 2020 by state

charger_2020 <-
  stations %>%
  filter(year(`Open Date`) < 2021) %>%
  filter(State != 'ON') %>%
  select(`Station Name`, `EV Level1 EVSE Num`, `EV Level2 EVSE Num`,
         `EV DC Fast Count`, State) %>%
  replace(is.na(.), 0) %>%
  mutate(EVSE = `EV Level1 EVSE Num` + `EV Level2 EVSE Num` +
           `EV DC Fast Count`) %>%
  group_by(State) %>%
  summarise(EVSE_state = sum(EVSE))

# the Plug-in Electric Vehicles in 2020 by state

PEV_2020 <-
  LDV_2020 %>%
  filter(State != 'United States') %>%
  mutate(PEV = `Electric (EV)` + `Plug-In Hybrid Electric (PHEV)`) %>%
  select(State, PEV)

# Merge the needed data for 2020

merge_2020 <-
  full_join(state, PEV_2020) %>%
  select(code, PEV) %>%
  rename(State = code) %>%
```

```
  full_join(charger_2020) %>%
  mutate(EV_per_charger = PEV / EVSE_state)%>%
  select(State, EV_per_charger) %>%
  mutate(year = 2020)

# Electric vehicle supply equipment (EVSE, charging points) in 2019 by state

charger_2019 <-
  stations %>%
  filter(year(`Open Date`) < 2020) %>%
  filter(State != 'ON') %>%
  select(`Station Name`, `EV Level1 EVSE Num`, `EV Level2 EVSE Num`,
         `EV DC Fast Count`, State) %>%
  replace(is.na(.), 0) %>%
  mutate(EVSE = `EV Level1 EVSE Num` + `EV Level2 EVSE Num` +
           `EV DC Fast Count`) %>%
  group_by(State) %>%
  summarise(EVSE_state = sum(EVSE))

# the Plug-in Electric Vehicles in 2019 by state

PEV_2019 <-
  LDV_2019 %>%
  filter(State != 'United States') %>%
  mutate(PEV = `Electric (EV)` + `Plug-In Hybrid Electric (PHEV)`) %>%
  select(State, PEV)

# Merge the needed data for 2019

merge_2019 <-
  full_join(state, PEV_2019) %>%
  select(code, PEV) %>%
  rename(State = code) %>%
  full_join(charger_2019) %>%
  mutate(EV_per_charger = PEV / EVSE_state)%>%
  select(State, EV_per_charger) %>%
  mutate(year = 2019)

# Electric vehicle supply equipment (EVSE, charging points) in 2018 by state

charger_2018 <-
  stations %>%
  filter(year(`Open Date`) < 2019) %>%
  filter(State != 'ON') %>%
  select(`Station Name`, `EV Level1 EVSE Num`, `EV Level2 EVSE Num`,
         `EV DC Fast Count`, State) %>%
  replace(is.na(.), 0) %>%
  mutate(EVSE = `EV Level1 EVSE Num` + `EV Level2 EVSE Num` +
           `EV DC Fast Count`) %>%
  group_by(State) %>%
  summarise(EVSE_state = sum(EVSE))
```

```
# the Plug-in Electric Vehicles in 2018 by state

PEV_2018 <-
  LDV_2018 %>%
  filter(State != 'United States') %>%
  mutate(PEV = `Electric (EV)` + `Plug-In Hybrid Electric (PHEV)`) %>%
  select(State, PEV)

# Merge the needed data for 2018

merge_2018 <-
  full_join(state, PEV_2018) %>%
  select(code, PEV) %>%
  rename(State = code) %>%
  full_join(charger_2018) %>%
  mutate(EV_per_charger = PEV / EVSE_state)%>%
  select(State, EV_per_charger) %>%
  mutate(year = 2018)

# Merge the data of different years

All_years <-
  rbind(merge_2018,
        merge_2019) %>%
  rbind(merge_2020) %>%
  rbind(merge_2021)

# 2. Minor adjustments

All_years_heat <-
  All_years %>%
  filter(!is.na(EV_per_charger))

# factorize the state vector

level <-
  All_years_heat %>%
  filter(year == 2021) %>%
  arrange(desc(EV_per_charger)) %>%
  pull(State)

All_years_heat$State <-
  factor(All_years_heat$State, levels = level)

# Reset the discrete intervals that suit heat map better

All_years_heat <-
  All_years_heat %>%
  mutate(EV_per_Charger =
         case_when(EV_per_charger <= 7 ~ '0 - 7',
                   EV_per_charger <= 14 ~ '7 - 14',
                   EV_per_charger <= 21 ~ '14 - 21',
```

```r
                    EV_per_charger <= 28 ~ '21 - 28',
                    EV_per_charger <= 35 ~ '28 - 35',
                    EV_per_charger <= 42 ~ '35 - 42',
                    .default = '42 - 49'))

All_years_heat$EV_per_Charger <-
  factor(All_years_heat$EV_per_Charger,
         levels = c('0 - 7', '7 - 14', '14 - 21', '21 - 28', '28 - 35',
                    '35 - 42', '42 - 49'))

# Heat map ------------------------------------------------------------------

heat_map <-
  All_years_heat %>%
  ggplot(mapping =
           aes(x = State,
               y = year)) +
  geom_tile(
    aes(fill = EV_per_Charger)) +

  # the mid point is EV-to-charger ratio equals to 10, which is the recommended
  # level of EV charger deployment by the Alternative Fuel Infrastructure
  # Directive (AFID).

  scale_fill_manual(values =
                      c('#91cf60', '#f0f9e8', '#ffffb2',
                        '#fecc5c', '#fd8d3c', '#f03b20', '#bd0026')) +
  labs(title = paste('Electric Vehicle Charging Point Deployment by State in',
                     'the U.S., 2018 - 2021'),
       subtitle = paste('From left to right: the state with the least (NJ) to',
                        'most (WY) adequate deployment of EV charging points'),
       caption = 'Data Source: Alternative Fuels Data Center') +
  theme(axis.title = element_blank(),
        axis.ticks = element_blank(),
        panel.background = element_blank())
```