PPOL 646 - 01 Interim Visualization Assignment 4

Xiyu. Zhang

TOTAL POINTS

3.33 / 0

QUESTION 1

1 Visualization 1 0 / 0

✓ **+ 0 pts** *A really nice 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. A legend should be easy to understand and use correct spelling, capitalization, and punctuation. For some*

graphs (such as line graphs), consider directly labeling the elements (such as the lines) instead of using a legend.

+ 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 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 Use ggplot instead of base R graphics for this visualization.

+ 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.

+ 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.

Really good work! I like your font choice, your use of non-default aesthetics, and your use of annotations. One suggestion to consider is breaking up plots by year rather than country because it is hard to read data

for present and past years given the magnitude of predicted increases in vehicle stock for 2025 and 2030. A few additional suggestions are to name the International Energy Agency's dataset you used in your source, and to take a look at your labels' capitalization and grammar. Finally, please remember to state your visualization's intended audience in the introductory paragraph.

QUESTION 2

2 Visualization 2 3.33 / 0

✓ **+ 0 pts** *A really nice 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. A legend should be easy to understand and use correct spelling, capitalization, and punctuation. For some graphs (such as line graphs), consider directly*

labeling the elements (such as the lines) instead of using a legend.

✓ **+ 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 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 Use ggplot instead of base R graphics for this visualization.

+ 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.

+ 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.

+ 3.33 *Point adjustment*

Good job! This scatterplot is a great choice for your data, and I think your annotation for China's plot point is very effective.

Please double-check the scale of the x-axis; it

seems inconsistent. The distance between 0 and 1.2 percent looks larger than the distance between 1.2 and 26.8 percent.

Consider increasing the horizontal dimension of the plot area to allow more separation of points at the low end of the x-axis.

Consider using ggrepel for the point labels to avoid some of the overwriting of the points.

A few suggestions are to reduce the number of colors by giving countries in the same regions the same colors, and to use per capita values for the number of electric vehicles. China's point size is huge, but that makes sense because its population is huge. To better compare across countries, it would be helpful to understand what countries have more electric vehicles accounting for differences in their populations.

1 This should be a legend that shows some sample sizes and associated values.

2 The distance between 0 and 1.2 percent looks larger than the distance between 1.2 and 26.8 percent. The scale should be consistent.

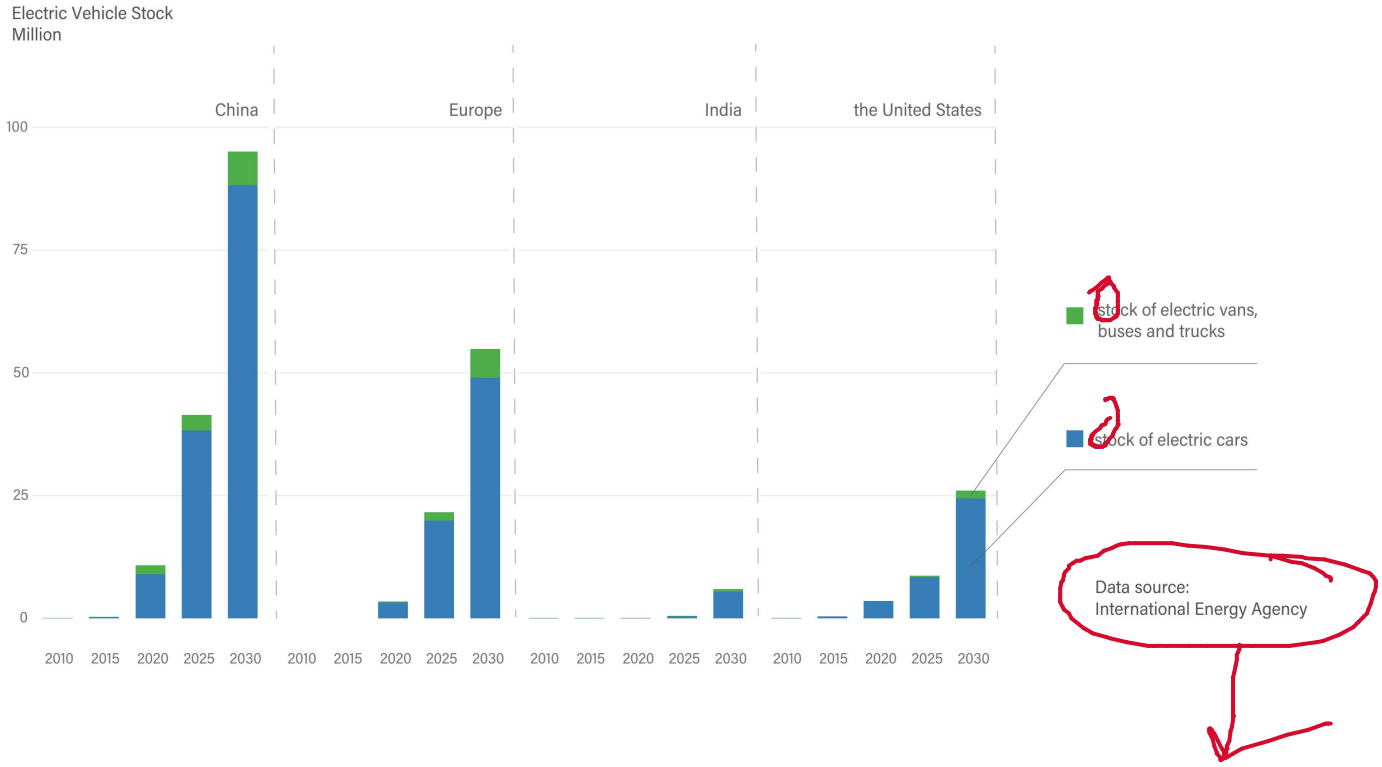
Statement of Purpose

In this diagram, as the title suggests, I want to convey the message that electric cars are dominant in the world among all types of electric vehicles. I plan to include this graph in my thesis. This will give my thesis readers a strong explanation as to why I chose electric cars and no other types of electric vehicles as the subject of my analysis.

By visualizing the International Energy Agency (IEA) dataset, using the agency's records of historical data and forecasts of the number of electric vehicles over the next decade under the IEA Stated Policies Scenario (STEPS), I show the market compositions of the world's four major electric vehicle markets: China, Europe, India, and the United States, between 2010 and 2030. For clarity, I have combined the categories other than cars: vans, buses, and trucks, because their combined numbers (in green) are far smaller than the share of cars alone.

Cars dominate vehicle electrification in the past, present, and foreseeable future

Electric vans, buses and trucks altogether account for little proportion of electric vehicles stock in key markets worldwide, 2010 – 2030



01

Xiyu Zhang

2023-04-06

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

library(tidyverse)
library(ggplot2)

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

EV_2010_to_2021 <-
  read_csv('data_own/IEA-EV-data.csv')

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

EV_cars_quantity <-
  EV_2010_to_2021 %>%
  filter(region %in%
           c('China', 'Europe', 'USA', 'India')) %>%
  filter(year %in%
           seq(2010, 2030, 5)) %>%
  filter(category != 'Projection-APS') %>%
  filter(parameter %in%
           c('EV stock')) %>%
  mutate(mode_new =
           if_else(mode == 'Cars',
                   'Cars',
                   'Vans, buses & trucks')) %>%
  group_by(region, year, mode_new) %>%
  summarise(value_new =
              sum(value))

  # Set the categorical variables as factors with certain orders

EV_cars_quantity$mode_new <-
  factor(EV_cars_quantity$mode_new,
         levels =
           c('Vans, buses & trucks', 'Cars'))

EV_cars_quantity$region <-
  factor(EV_cars_quantity$region,
         levels =

           # in alphabetical order

           c('China', 'Europe', 'India', 'USA'))
```

```
# Add percent numbers for reference

EV_cars_percent <-
  EV_cars_quantity %>%
  pivot_wider(names_from = mode_new,
              values_from = value_new) %>%
  mutate(percentage =
           formattable::percent(
             (`Vans, buses & trucks` /
                 (`Vans, buses & trucks` + `Cars`)))) %>%
  select(region, year, percentage)

# Combine the data together

EV_cars_all <-
  EV_cars_quantity %>%
  full_join(EV_cars_percent,
            by = join_by(region, year))

# Data visualization ----------------------------------------------------

p1 <-
  EV_cars_all %>%
  ggplot(mapping =
           aes(x = year,
               y = value_new,
               fill = mode_new)) +

  # only add some of the grid lines

  geom_hline(yintercept = 25000000,
             color = 'gray93') +
  geom_hline(yintercept = 50000000,
             color = 'gray93') +
  geom_hline(yintercept = 75000000,
             color = 'gray93') +
  geom_hline(yintercept = 100000000,
             color = 'gray93') +
  geom_bar(stat = 'identity',
           width = 3.2) +
  facet_wrap(~ region,
             nrow = 1) +
  geom_text(aes(label = percentage),
            size = 2.5,
            check_overlap = T) +
  scale_y_continuous(
    breaks = seq(0, 125000000, 25000000),
    labels = c('0', '25', '50', '75', '100', ''),
    limits = c(0, 125000000)) +
```

```
scale_fill_manual(values =
                    c('#4daf4a', '#377eb8')) +
labs(title = paste('Cars dominate vehicle electrification in the past,',
                   'present, and foreseeable future'),
     subtitle = paste('Electric vans, buses and trucks account for a little',
                      'proportion of electric vehicles in key markets worldwide'),
     caption = 'Data source: International Energy Agency',
     y = 'Million') +
theme_minimal() +
theme(axis.title.x = element_blank(),
      panel.grid = element_blank(),
      legend.position = 'none')
```

1 Visualization 1 0 / 0

✓ **+ 0 pts** *A really nice 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. A legend should be easy to understand and use correct spelling, capitalization, and punctuation. For some graphs (such as line graphs), consider directly labeling the elements (such as the lines) instead of using a legend.*

+ 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 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 Use ggplot instead of base R graphics for this visualization.

+ 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.

+ 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.

Really good work! I like your font choice, your use of non-default aesthetics, and your use of annotations. One suggestion to consider is breaking up plots by year rather than country because it is hard to read data for present and past years given the magnitude of predicted increases in vehicle stock for 2025 and 2030. A few additional suggestions are to name the International Energy Agency's dataset you used in your source, and to take a look at your labels' capitalization and grammar. Finally, please remember to state your visualization's intended audience in the introductory paragraph.

Statement of Purpose

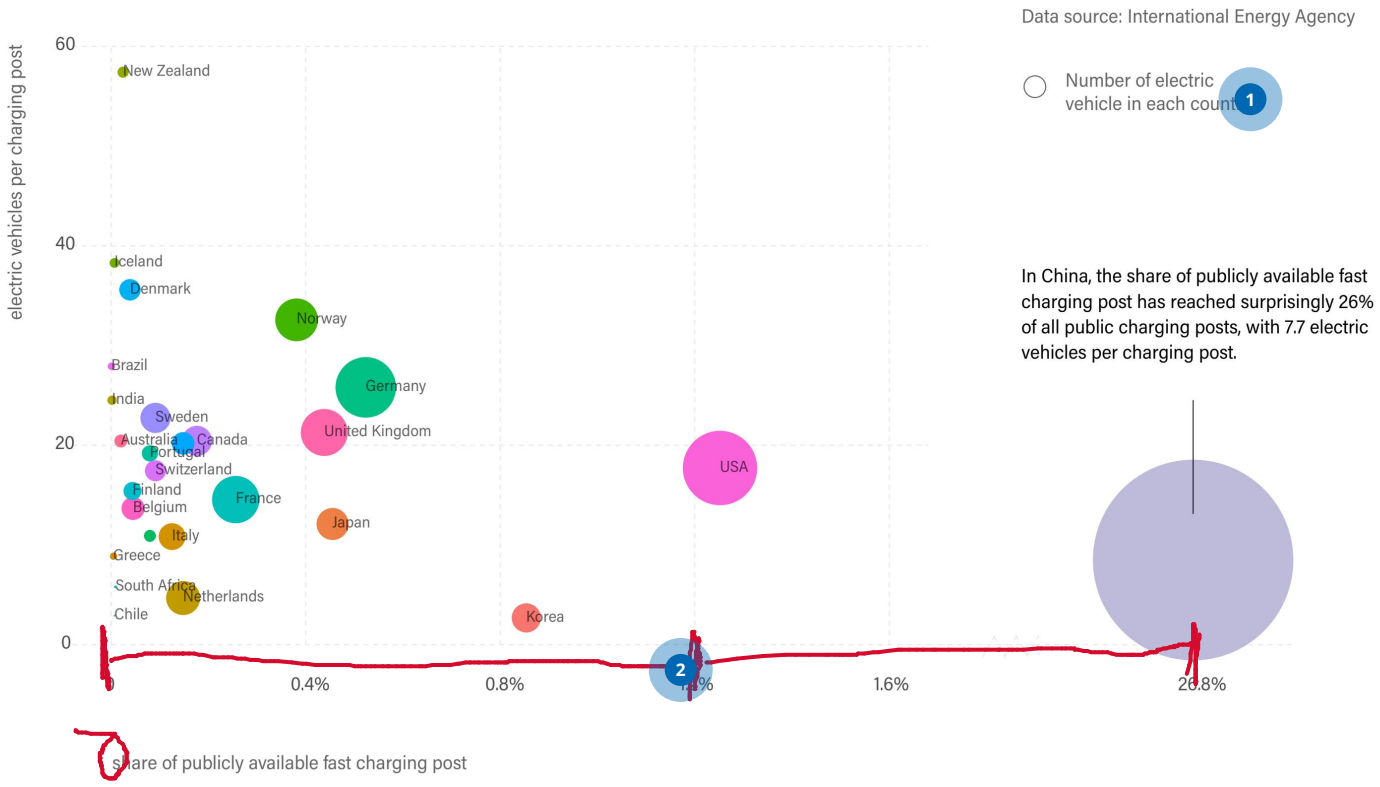
The deployment of publicly available chargers for electric vehicles typically lags behind the growth of electric vehicles (IEA, 2022). That is, in today's electrification of transportation, people sometimes buy electric vehicles (EVs) first, and then find out that publicly available charging facilities are insufficient. The deployment of EV charging facilities is the policy issue I want to examine in my paper.

The purpose of this graph is to show the level of EV charging deployment in different countries. The horizontal axis depicts the share of publicly available fast-charging posts, and the publicly available charging posts consist of both "slow" and "fast" charging posts. Fast charging posts can reduce charging time per vehicle and serve more vehicle owners in a given time frame, i.e., provide more charging capacity. The larger this number is, the fewer public charging posts are available for each EV, and the closer to 0, the more charging posts are available for each vehicle. Finally, the size of the bubbles shows the current number of registered electric vehicles in a country and represents the demand for electric vehicle charging in that country.

China (26.8%) is far ahead of all other countries (less than 2%) on the horizontal axis, while also being close to zero on the vertical axis. At the same time, China has the largest bubble, indicating that it has the most electric vehicles. All this reflects the booming development of electric vehicles in China.

China is leading in EV charging infrastructure depolyment, far more than the next 2: the US and Korea

The deployment of EV charging posts in 2021



02

Xiyu Zhang

2023-04-06

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

library(tidyverse)
library(ggplot2)
library(colorspace)

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

EV_2010_to_2021 <-
  read_csv('data_own/IEA-EV-data.csv')

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

# Create a initial data set

EV_initial <-
  EV_2010_to_2021 %>%
  filter(!region %in%

           # exclude the regional data, in order to focus on countries

           c('Europe', 'Rest of the world', 'Other Europe', 'World')) %>%
  filter(parameter %in% c('EV stock', 'EV charging points'),
         year == 2021,
         category == 'Historical')

# Number of electric vehicles per charging point

EV_per_charging <-
  EV_initial %>%
  group_by(region, parameter) %>%
  summarise(value_new =
              sum(value)) %>%
  pivot_wider(
    names_from = parameter,
    values_from = value_new) %>%
  mutate(EV_per_charger =
           `EV stock` / `EV charging points`) %>%
  select(region, EV_per_charger)

# The share of fast chargers in the total number of chargers

Fast_charger_share <-
```

```
  EV_initial %>%
  filter(parameter == 'EV charging points') %>%
  select(region, powertrain, value) %>%
  pivot_wider(
    names_from = powertrain,
    values_from = value) %>%
  mutate(fast_percentage =
           `Publicly available fast` /
           sum(`Publicly available slow`, `Publicly available fast`)) %>%
  mutate(fast_percentage =
           formattable::percent(fast_percentage)) %>%
  select(region, fast_percentage)

# The total number of electric vehicles

EV_number <-
  EV_initial %>%
  filter(parameter == 'EV stock') %>%
  group_by(region) %>%
  summarise(EV_total = sum(value))

# Combine the created variables into one data frame

EV_analysis <-
  EV_per_charging %>%
  full_join(Fast_charger_share) %>%
  full_join(EV_number) %>%

  # exclude the missing values

  filter(!is.na(EV_per_charger)) %>%

  # to make visualization neater, arrange the order

  arrange(EV_number) %>%

  # we found that the fast charger percentage for China is an outlier, since
  # China's percentage of fast chargers are way higher than other countries
  # Exclude China for now

  filter(region != 'China')

# Data visualization ----------------------------------------------------

# generate a color palette

palette <-
  rainbow_hcl(27)

# create a bubble chart
```

```
p2 <-
  EV_analysis %>%
  ggplot() +
  geom_point(
    aes(x = EV_per_charger,
        y = fast_percentage,
        size = EV_total,
        color = palette)) +
  scale_size(range = c(.1, 30),
             name = 'Number of Electric Vehicles') +
  coord_flip() +
  geom_text(
    aes(
      x = EV_per_charger,
      y = fast_percentage,
      label = region),
    hjust = 'left',
    size = 3,
    check_overlap = T) +
  scale_x_continuous(
    breaks = seq(0, 60, 20)) +
  scale_y_continuous(
    breaks = seq(0, 0.016, 0.004),
    labels = c('0', '0.4%', '0.8%', '1.2%', '1.6%'),
    limits = c(0, 0.016)) +
  labs(title = paste('China is leading in EV charging infrastructure depolyment,',
                     'follows by the US and Korea'),
       subtitle = 'The deployment of EV charging posts in 2021',
       caption = 'Data source: International Energy Agency',
       x = 'electric vehicles per charging post',
       y = 'share of publicly available fast charging post') +
  theme_minimal() +
  theme(panel.grid.minor = element_blank(),
        panel.grid.major = element_line(linetype = 'dashed'),
        legend.position = 'none')
```

2 Visualization 2 3.33 / 0

✓ **+ 0 pts** *A really nice 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. A legend should be easy to understand and use correct spelling, capitalization, and punctuation. For some graphs (such as line graphs), consider directly labeling the elements (such as the lines) instead of using a legend.*

✓ **+ 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 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 Use ggplot instead of base R graphics for this visualization.

+ 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.

+ 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.

+ 3.33 *Point adjustment*

Good job! This scatterplot is a great choice for your data, and I think your annotation for China's plot point is very effective.

Please double-check the scale of the x-axis; it seems inconsistent. The distance between 0 and 1.2 percent looks larger than the distance between 1.2 and 26.8 percent.

Consider increasing the horizontal dimension of the plot area to allow more separation of points at the low end of the x-axis.

Consider using ggrepel for the point labels to avoid some of the overwriting of the points.

A few suggestions are to reduce the number of colors by giving countries in the same regions the same colors, and to use per capita values for the number of electric vehicles. China's point size is huge, but that makes sense because its population is huge. To better compare across countries, it would be helpful to understand what countries have more electric vehicles accounting for differences in their populations.

1 This should be a legend that shows some sample sizes and associated values.

2 The distance between 0 and 1.2 percent looks larger than the distance between 1.2 and 26.8 percent. The scale should be consistent.