# 02

Xiyu Zhang

2023-04-06

```r
# 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 <-
```

```r
  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')
```