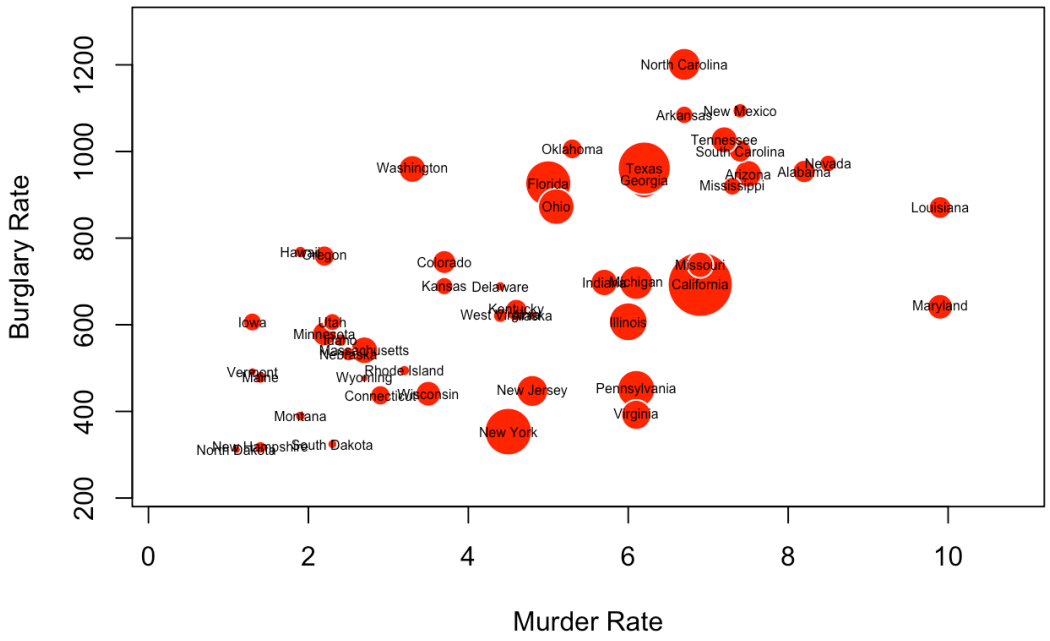
1. Replication of Yau's visualization



2. My brief statement of the purpose of my original visualization

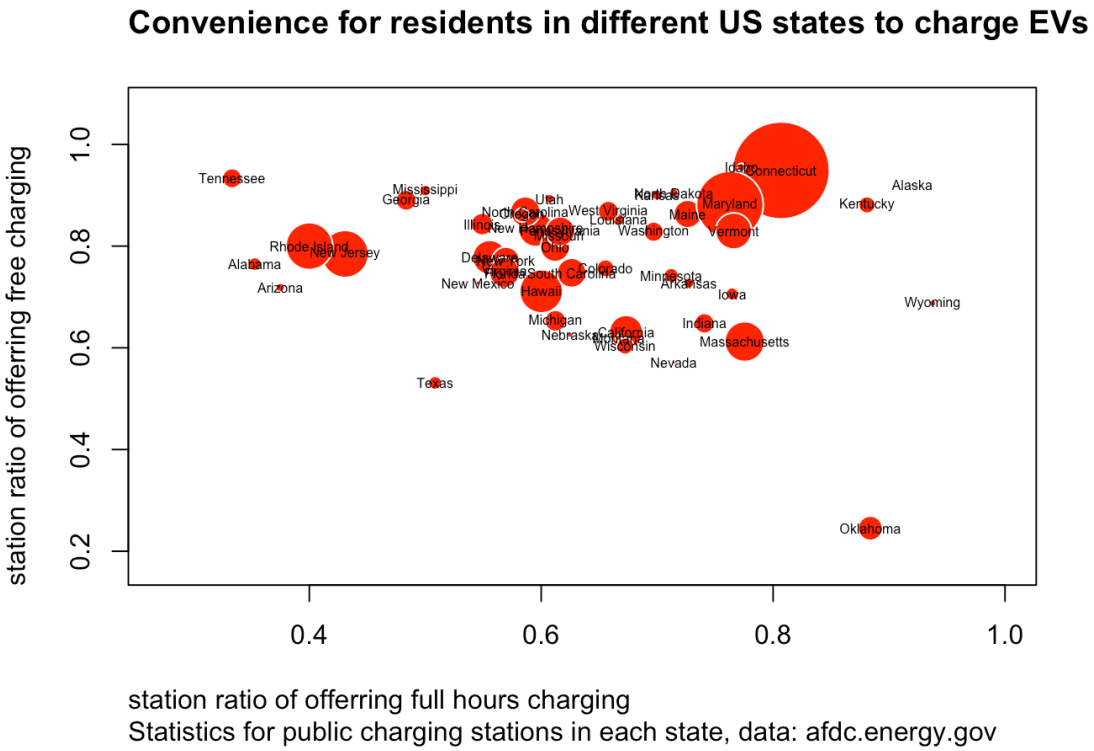
I tried to construct a bubble graph like this assignment's replication practice.

Here, the two dimensions in the x-axis and y-axis show each state's electronic vehicle public charging station's free charging rate and full hours charging rate, respectively. The third dimension shows the electronic vehicle charging station density, which that means on an average area of land in a state, how many publicly accessible electronic vehicles charging stations residents can find.

Overall, the chart tries to present the convenience for residents in different states to charge their vehicles in daily life. As one state approaches the upper and right frontiers of this chart with a large bubble, that indicates charging one's electronic vehicle in this state will be pretty convenient.

My electronic vehicle charging station data comes from afdc.energy.gov. The state area data comes from an R built-in data set named ‘state’.

3. My original visualization



station ratio of offerring full hours charging
Statistics for public charging stations in each state, data: afdc.energy.gov

4. My R code for my original visualization

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

library(sf)
library(tidyverse)

# Data preparation ---------------------------------------------------------

# Read in the initial dataset

electricity_station_initial <-

   # read in the dataset

   st_read('data_own/alt_fuel_stations.geojson') %>%

   # convert an sf object into a pure tibble

   as_tibble()

# Filter for the wanted data

electricity_station <-
   electricity_station_initial %>%

   # filter for the wanted types

   filter(

      # only include public electricity stations but not private ones

      access_code == 'public',

      # only include those are currently available but not planned nor
      # temporarily unavailable

      status_code == 'E',

      # only include those in the US

      country == 'US',

      # only include the charging stations open to the public
```

```
      restricted_access == FALSE) %>%

    # select the wanted traits of those electricity charging stations

    select(
      c(access_days_time, id, open_date, owner_type_code, state,
      ev_pricing, ev_renewable_source, facility_type))

# Create the wanted variables

# To create a dataset, as for every state, including the free pricing rate,
# 24-hour pricing rate, and the density of the charging stations in different
# states open to the public

# The intention of this data visualization is to visualize the convenience for
# people to charge their private electronic vehicles

# develop wanted variables

elec_station_by_state <-
    electricity_station %>%

    # select the wanted features

    select(id, state, access_days_time, ev_pricing) %>%

    # exclude missing values

    filter(!is.na(access_days_time),
                !is.na(ev_pricing)) %>%

    # construct two Boolean values describing whether a station operates for 24
    # hours or not, and whether this station offers free charging, respectively

    transmute(
      id,
      state,
      x =
        if_else(
          str_detect(access_days_time, '24'),
          TRUE,
          FALSE),
      y =
```

```
            if_else(
               str_detect(ev_pricing, 'Free'),
               TRUE,
               FALSE)) %>%

    # renames these two Boolean values to full_hours and free_charging

    rename(full_hours = x,
                free_charging = y,
                state.abb = state)

# Gather states information in R build-in data sets

data(state)

# select wanted features and construct a tibble

state_features <-

    # including state abbreviation, state area, and state name

    tibble(state.abb, state.area, state.name)

# I don't know how to do these concisely so I hard-code to build the wanted
# variables

# 1. construct a variable naming state.amount to describe the full amount of these
# electricity charging station in each US state

temp <-
    elec_station_by_state %>%
    group_by(state.abb) %>%
    summarise(state.amount = n())

# 2. join this variable with the state features tibble construct above

temp_2 <-
    elec_station_by_state %>%
    left_join(temp) %>%
    left_join(state_features)

# 3. Respectively, calculate free_charging_rate and full_hour_rate for each
# state, referring to among all of the electric vehicle charging station, the
# ratio of free charging stations and the ratio of charging stations operating
```

```
# 24 hours everyday

temp_3 <-
    temp_2 %>%
    group_by(state.abb, free_charging, state.amount) %>%
    summarise(free_charging_amount = n(),

                    # ignore 'state.amount', only group by the first two variables

                    .groups = 'drop_last') %>%

    # calculating the charging stations that offers free charging

    filter(free_charging == TRUE) %>%

    # calculate the wanted variable by divide the free charging station amount by
    # the whole charging station amount in each state

    mutate(free_charging_rate = free_charging_amount / state.amount) %>%

    # select useful variables for future data visualization

    select(state.abb, free_charging_rate)

# temp_4 is basically the same as the previous one, but for full-hours rate

temp_4 <-
    temp_2 %>%
    group_by(state.abb, full_hours, state.amount) %>%
    summarise(full_hours_amount = n(),
                    .groups = 'drop_last') %>%
    filter(full_hours == TRUE) %>%
    mutate(full_hours_rate = full_hours_amount / state.amount) %>%
    select(state.abb, full_hours_rate, state.amount)

# 4. Finally, combine the constructed variables together in one tibble

temp_fin <-
    left_join(state_features,
                    temp_3) %>%
    left_join(temp_4) %>%

    # calculate the electronic vehicle charging station density in each state
    # by divide the amount in each state by the stata area
```

```
    mutate(station_density = state.amount / state.area)

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

# I tried to construct a bubble graph just like the replication practice in this
# assignment.
# Here, the two dimension in x-axis and y-axis shows the free charging rate and
# the full hours charging rate in each state, respectively; the third dimension
# shows the electronic vehicle charging station density in each state.
# Overall, the chart tries to present the convenience for residents in different
# states to charge their vehicles in daily life. As one state approaching the
# upper and right frontiers of this chart with a large bubble, that indicates
# charging one's electronic vehicle in this state will be pretty convenient.

symbols(temp_fin$full_hours_rate,
              temp_fin$free_charging_rate,
              circles =
                 sqrt(temp_fin$station_density / pi),
              inches = 0.3,
              fg = 'white',
              bg = 'red',
              main = paste('Convenience for residents in different US states to',
                           'charge EVs'),
              sub = paste('Statistics for public charging stations in each',
                          'state, data: afdc.energy.gov'),
              xlab = 'station ratio of offerring full hours charging',
              ylab = 'station ratio of offerring free charging',
              adj = 0) +
   text(temp_fin$full_hours_rate,
            temp_fin$free_charging_rate,
            temp_fin$state.name,
            cex = 0.5)
```