```
library(tidyverse)
library(sf)
# Original scatterplot -------
# Read in the initial dataset
# motor gasoline price for each state, from EIA, in 2021
gasoline_21 <-
 read_csv('data_own/fuel_mg.csv') %>%
 filter(MSN == 'MGACD') %>%
 rename(motor_gasoline = '2021')
# all sector electricity price for each state, from EIA, in 2021
electricity_21 <-
 read_csv('data_own/fuel_es.csv') %>%
 filter(MSN == 'ESTCD') %>%
 rename(electricity = '2021')
# join the two forms into one
price 21 <-
 left_join(gasoline_21,
           electricity 21,
           by = 'State') %>%
 select(State, motor gasoline, electricity)
# create an elementary scatterplot
scatter temp <-
 price 21 %>%
 ggplot(
   aes(x = motor_gasoline,
       y = electricity)) +
 geom point(colour = 'blue')
# It seems there exists an outlier, so descending the electricity price and
# remove the outlier, and create the scatterplot again
scatter_fin <-
 price 21 %>%
 arrange(desc(electricity)) %>%
 filter(State != 'HI') %>%
 ggplot(
   aes(x = motor gasoline,
       y = electricity,
       label = State)) +
  geom_point(colour = '#3182bd',
```

```
alpha = 0.7) +
geom_smooth(
 method = loess,
 colour = '#6baed6',
 se = F) +
geom_text(hjust = 0,
          vjust = -1,
          size = 3,
          colour = '#6baed6') +
labs(title = paste('The cost of driving gasoline and electric cars in the',
                   'United States'),
     subtitle = paste('Motor gasoline price and electricity price in each',
                      'state of US in 2021'),
     caption = 'Data: eia.gov',
     x = 'Motor gasoline price in transportation sector (USD per MMBtu)',
     y = 'Electricity price in all sector (USD per MMBtu)') +
theme_bw() +
theme(
 panel.border = element_blank(),
 panel.grid = element_blank(),
 axis.line = element_line(colour = '#6baed6'),
 axis.text.x = element_text(colour = '#3182bd'),
 axis.text.y = element_text(colour = '#3182bd'),
  axis.ticks = element_line(colour = '#6baed6'))
```