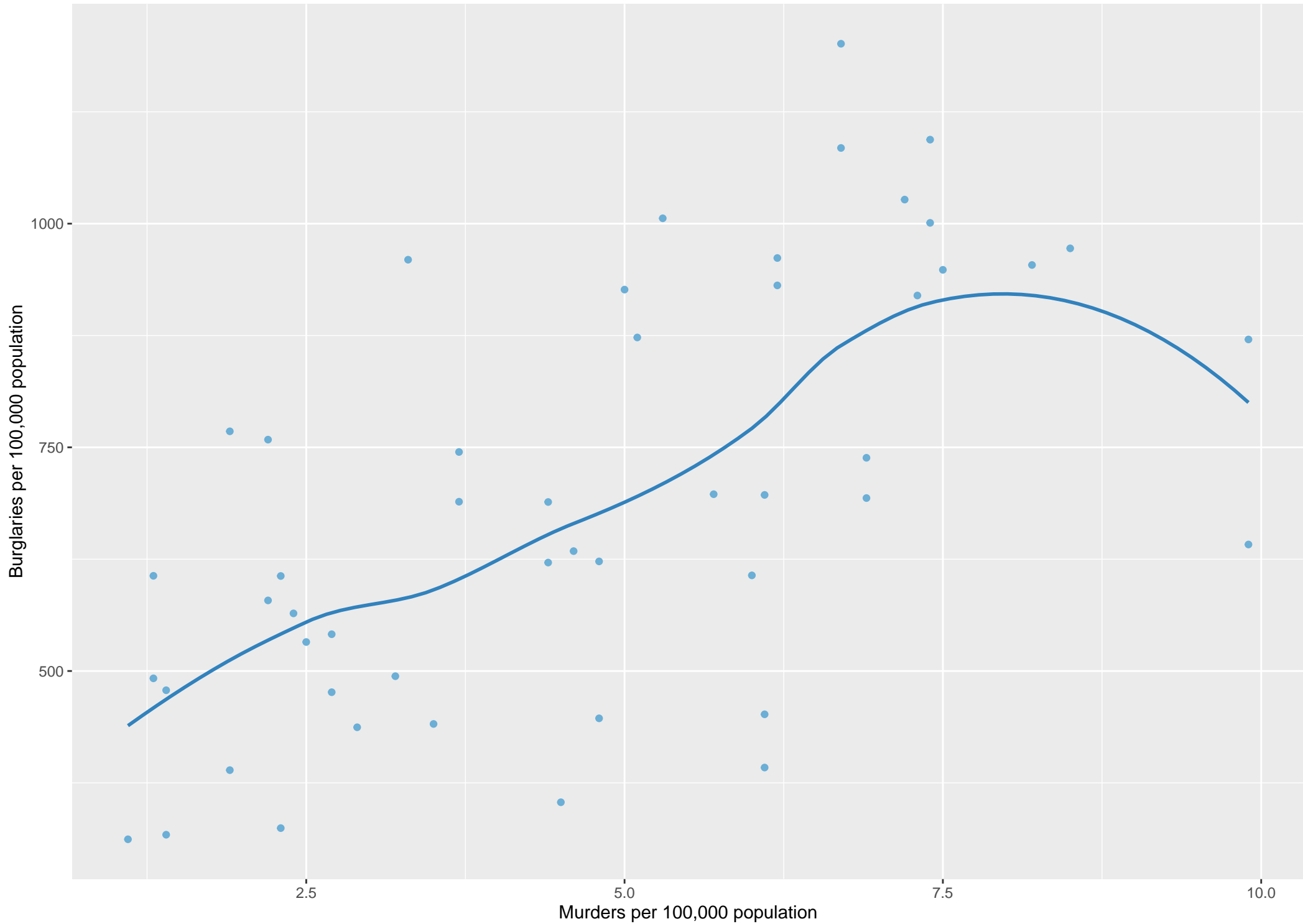# MURDERS VERSUS BURGLARIES IN THE UNITED STATES

States with higher murder rates tend to have higher burglary rates
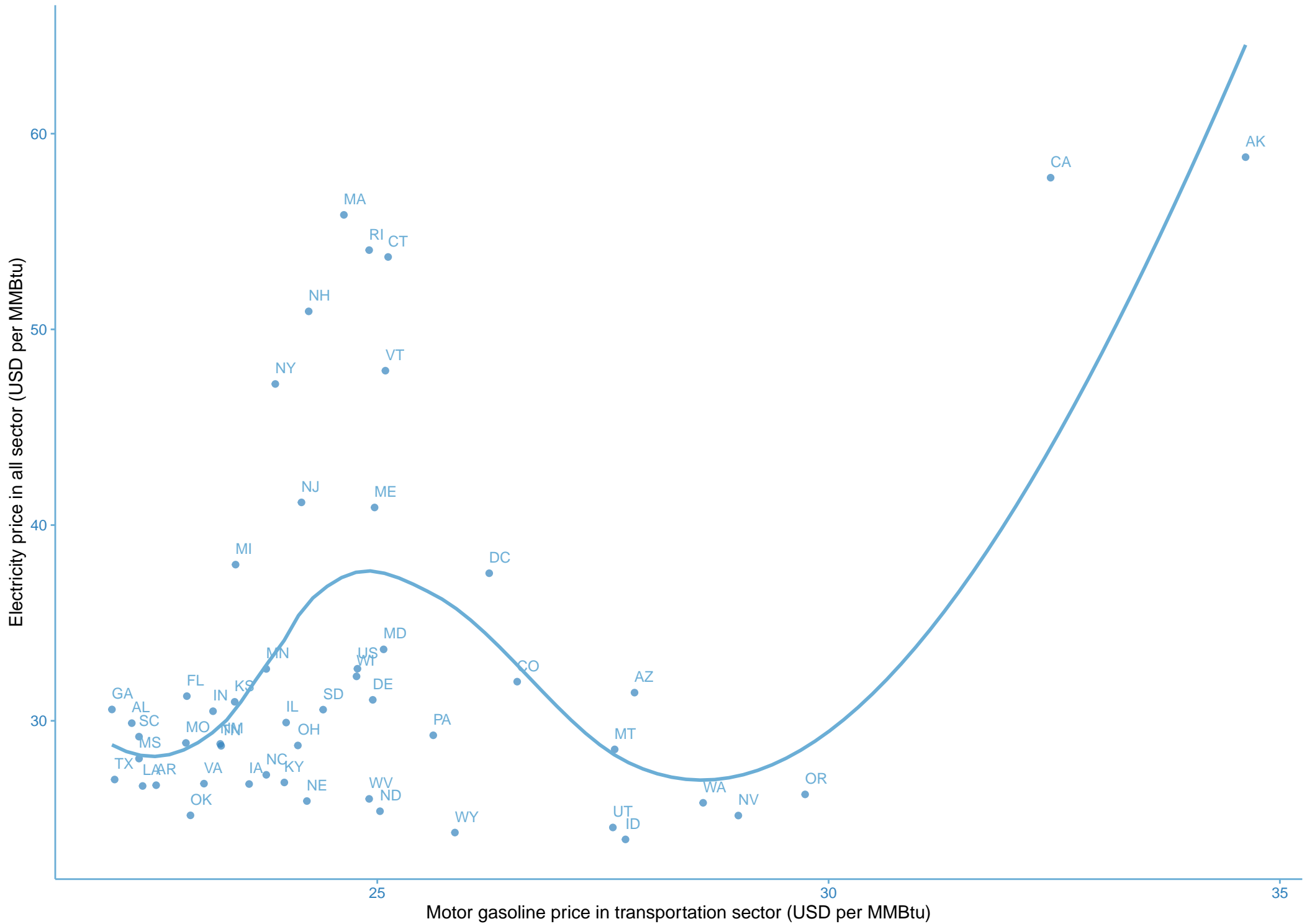

Burglaries per 100,000 population

Murders per 100,000 population

The purpose of my original visualization

First, to practice generating a scatterplot with a similar format to Yau's figure, I am trying to develop my own ability to create this kind of figure.

Second, I created this figure with consideration of my thesis topic. For my thesis, I focused on the factors that influence people's consumption decisions of electric vehicles across different states in the US. Given that the cost of fuel is a key independent variable that affects the overall cost of driving, I chose to represent each state's motor gasoline price and electricity price on the x-axis and the y-axis, respectively, as a measurement of the cost of each fuel. Then, I tried to use each dot to represent a state, and use their locations to position each state's fuel cost for driving both gasoline vehicles and electric vehicles. By doing so, I tried to highlight the potential influence of fuel prices on individuals' consumption decisions across different states.

The cost of driving gasoline and electric cars in the United States

Motor gasoline price and electricity price in each state of US in 2021

Electricity price in all sector (USD per MMBtu)

Motor gasoline price in transportation sector (USD per MMBtu)

Data: eia.gov

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

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'))
```