Vehicles Registered

2023-01-17

Vehicles Registered in California

The California DMV releases data on the number of registered vehicles by model year, make, and fuel type by Zip Code. This allows for calculating what percent of vehicles registered are electric vehicles or by brand for every model year.

Clearly, the number of electric vehicles is growing in California. Most of these electric vehicles are from Tesla. The data can be found here: https://data.ca.gov/dataset/vehicle-fuel-type-count-by-zip-code

```r
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 4.2.3
```

```r
library(dplyr)
```

```
## Warning: package 'dplyr' was built under R version 4.2.3
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```r
library(tidyr)
```

```
## Warning: package 'tidyr' was built under R version 4.2.3
```

```r
vehicles = read.csv("Data/vehicle-fuel-type-count-by-zip-code-2022.csv")

tab = vehicles %>%
  group_by(Zip.Code, Model.Year) %>%
  mutate(percent =  100 * Vehicles/sum(Vehicles)) %>%
  ungroup %>%
  group_by(Zip.Code, Model.Year, Make) %>%
  summarize(Vehicles = sum(Vehicles), Percent = sum(percent))
```

```
## 'summarise()' has grouped output by 'Zip.Code', 'Model.Year'. You can override
## using the '.groups' argument.

zi = tab %>%
  filter(Model.Year != "<2009" &  Model.Year != "2023") %>% ungroup() %>% select(-Zip.Code)
other_zi = zi %>%
  mutate(Make = recode(Make, "LEXUS" = "TOYOTA", "SCION" = "TOYOTA", "ACURA" = "HONDA", "INFINITI" = "N
                       "KIA" = "HYUNDAI", "AUDI" = "VOLKSWAGEN", "PORSCHE" = "VOLKSWAGEN")) %>%
  group_by(Model.Year, Make) %>%
  summarize(Vehicles = sum(Vehicles), Percent = sum(Percent))


## 'summarise()' has grouped output by 'Model.Year'. You can override using the
## '.groups' argument.

comb = other_zi %>%
  filter(!(Make %in% c("TOYOTA", "HONDA", "TESLA", "NISSAN", "SUBARU", "HYUNDAI", "FORD", "VOLKSWAGEN")
  group_by(Model.Year) %>%
  summarize(Make = "OTHER", Vehicles = sum(Vehicles), Percent = sum(Percent))
comb = rbind(other_zi, comb)
result = comb %>%
  filter(Make %in% c("TOYOTA", "HONDA", "TESLA", "NISSAN", "LEXUS", "SUBARU", "HYUNDAI", "FORD", "VOLKS
  mutate(Make = recode(Make, "TOYOTA" = "Toyota+Lexus", "HONDA" = "Honda+Acura", "NISSAN" = "Nissan+Inf
                       "VOLKSWAGEN" = "Volkswagen Group", "HYUNDAI" = "Hyundai+Kia",
                       "LEXUS" = "Lexus", "SUBARU" = "Subaru", "FORD" = "Ford", "OTHER" = "Other Brands
  ggplot(aes(x = Model.Year, y = Vehicles, group = Make, color = Make)) +
    geom_point() + geom_line() + theme_bw() +
    xlab("Model Year") +
    guides(x = guide_axis(angle = 45)) +
    ylab("Number of Registered Vehicles") +
    ggtitle("Number of Registered Vehicles by Brand and Model Year for California")


tab2 = vehicles %>%
  group_by(Model.Year) %>%
  mutate(percent =  100 * Vehicles/sum(Vehicles)) %>%
  ungroup %>%
  group_by(Model.Year, Make) %>%
  summarize(Vehicles = sum(Vehicles), Percent = sum(percent)) %>%
  filter(Model.Year != "<2009" & Model.Year != "2023")


## 'summarise()' has grouped output by 'Model.Year'. You can override using the
## '.groups' argument.

other_zi = tab2 %>%
  mutate(Make = recode(Make, "LEXUS" = "TOYOTA", "SCION" = "TOYOTA", "ACURA" = "HONDA", "INFINITI" = "N
                       "KIA" = "HYUNDAI", "AUDI" = "VOLKSWAGEN", "PORSCHE" = "VOLKSWAGEN")) %>%
  group_by(Model.Year, Make) %>%
  summarize(Vehicles = sum(Vehicles), Percent = sum(Percent))


## 'summarise()' has grouped output by 'Model.Year'. You can override using the
## '.groups' argument.
```

```r
comb = other_zi %>%
  filter(!(Make %in% c("TOYOTA", "HONDA", "TESLA", "NISSAN", "SUBARU", "HYUNDAI", "FORD", "VOLKSWAGEN"))
  group_by(Model.Year) %>%
  summarize(Make = "OTHER", Vehicles = sum(Vehicles), Percent = sum(Percent))
comb = rbind(other_zi, comb)
result2 = comb %>%
  filter(Make %in% c("TOYOTA", "HONDA", "TESLA", "NISSAN", "LEXUS", "SUBARU", "HYUNDAI", "FORD", "VOLKSW
  mutate(Make = recode(Make, "TOYOTA" = "Toyota+Lexus", "HONDA" = "Honda+Acura", "NISSAN" = "Nissan+Infi
                       "VOLKSWAGEN" = "Volkswagen Group", "HYUNDAI" = "Hyundai+Kia",
                       "LEXUS" = "Lexus", "SUBARU" = "Subaru", "FORD" = "Ford", "OTHER" = "Other Brands"
  ggplot(aes(x = Model.Year, y = Percent, group = Make, color = Make)) +
    geom_point() + geom_line() + theme_bw() +
    xlab("Model Year") +
    guides(x = guide_axis(angle = 45)) +
    ylab("Percent of Registered Vehicles") +
    ggtitle("Percent of Registered Vehicles by Brand and Model Year for California")
```

```r
tab2 = vehicles %>%
  group_by(Model.Year) %>%
  mutate(percent =  100 * Vehicles/sum(Vehicles)) %>%
  ungroup %>%
  group_by(Model.Year, Fuel) %>%
  summarize(Vehicles = sum(Vehicles), Percent = sum(percent)) %>%
  filter(Model.Year != "<2009" & Model.Year != "2023") %>%
  ggplot(aes(x = Model.Year, y = Percent, color = Fuel, group = Fuel)) +
    geom_point() + geom_line() + theme_bw() +
    xlab("Model Year") +
    guides(x = guide_axis(angle = 45)) +
    ylab("Percent of Registered Vehicles") +
    ggtitle("Percent of Vehicles by Fuel Type for California")
```

```
## `summarise()` has grouped output by 'Model.Year'. You can override using the
## `.groups` argument.
```
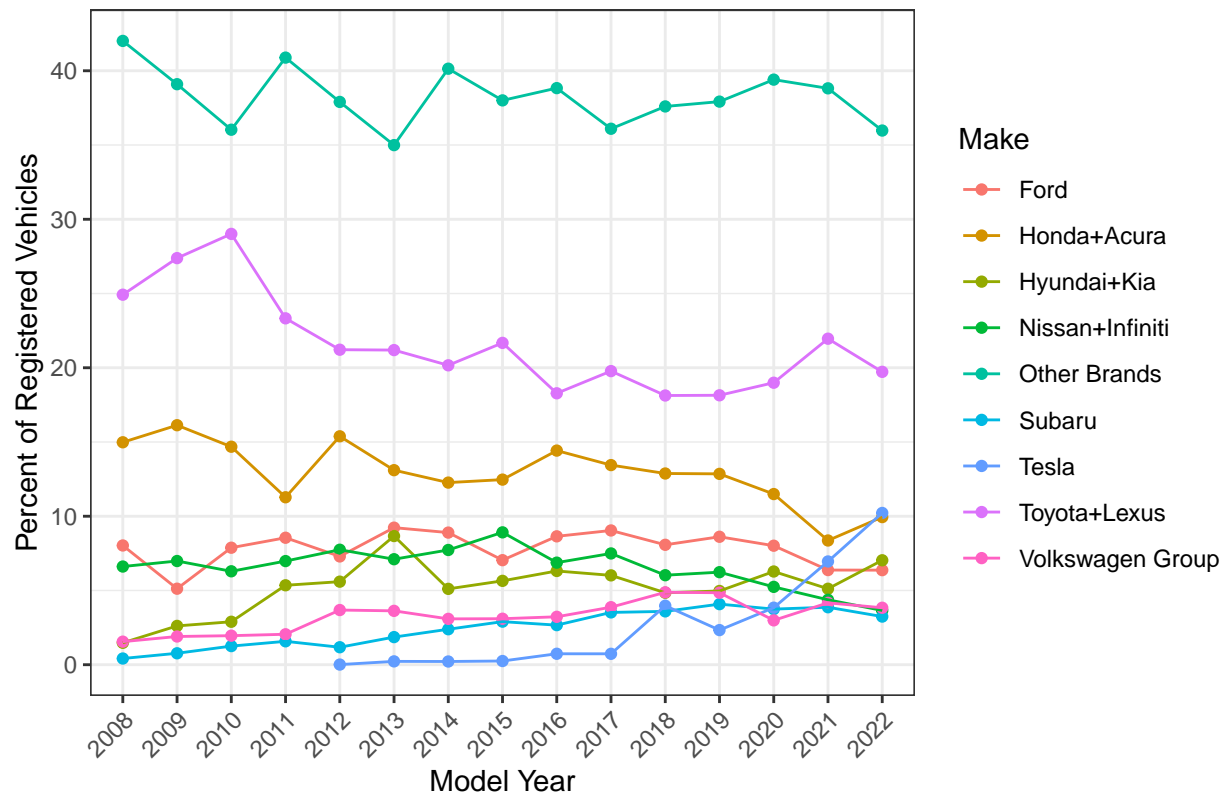
```r
result
```

Number of Registered Vehicles by Brand and Model Year for California
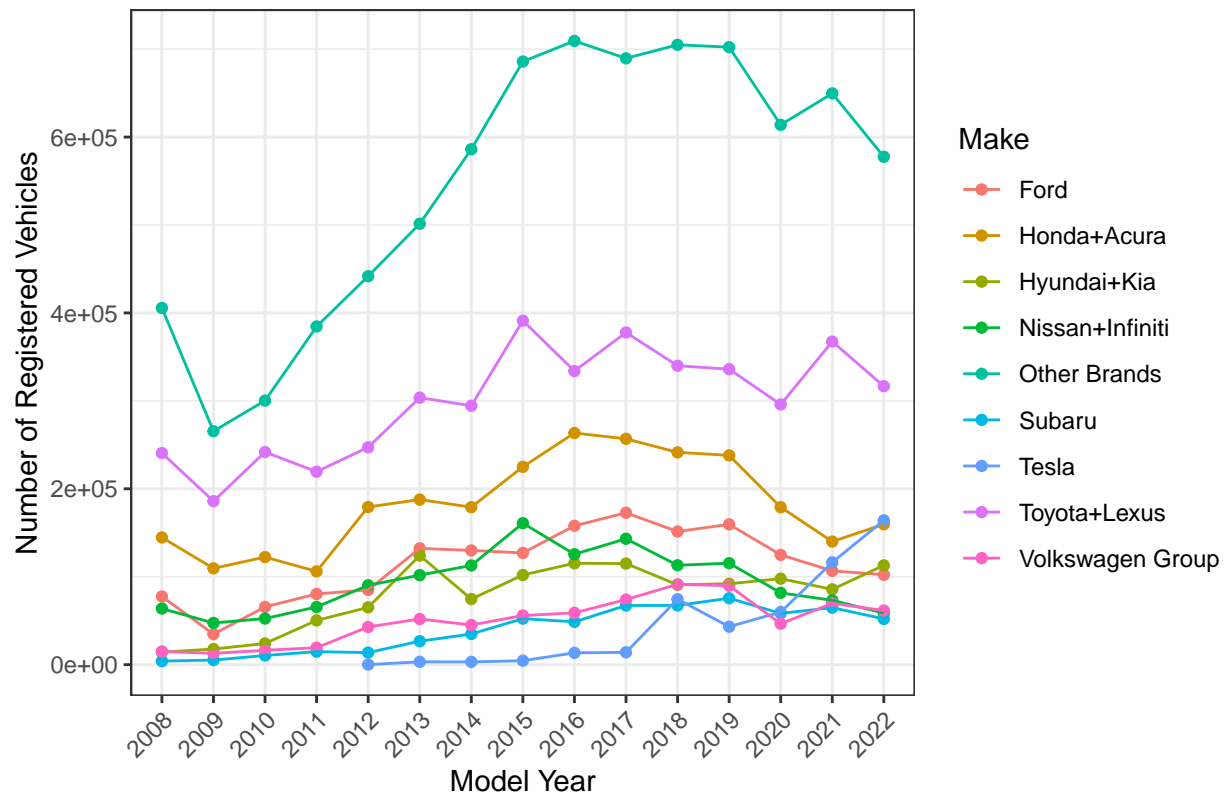
result2

Percent of Registered Vehicles by Brand and Model Year for California

tab2

Percent of Vehicles by Fuel Type for California