Project - MPG Data Visualization

Thundon

2023-12-25

Data Visualization

Load library

```
library(tidyverse) #for manipulate data and visualization(ggplot)
library(ggthemes) # for customize themes
library(patchwork) # for create multiple plot in one chart
library(wesanderson) # for choose color in chart
```

Explore data

```
colnames(mpg)
```

```
##  [1] "manufacturer" "model"        "displ"        "year"         "cyl"
##  [6] "trans"        "drv"          "cty"          "hwy"          "fl"
## [11] "class"
```

```
head(mpg)
```

```
## # A tibble: 6 x 11
##   manufacturer model displ  year   cyl trans      drv     cty   hwy fl    class
##   <chr>        <chr> <dbl> <int> <int> <chr>      <chr> <int> <int> <chr> <chr>
## 1 audi         a4      1.8  1999     4 auto(l5)   f        18    29 p     compa~
## 2 audi         a4      1.8  1999     4 manual(m5) f        21    29 p     compa~
## 3 audi         a4      2    2008     4 manual(m6) f        20    31 p     compa~
## 4 audi         a4      2    2008     4 auto(av)   f        21    30 p     compa~
## 5 audi         a4      2.8  1999     6 auto(l5)   f        16    26 p     compa~
## 6 audi         a4      2.8  1999     6 manual(m5) f        18    26 p     compa~
```

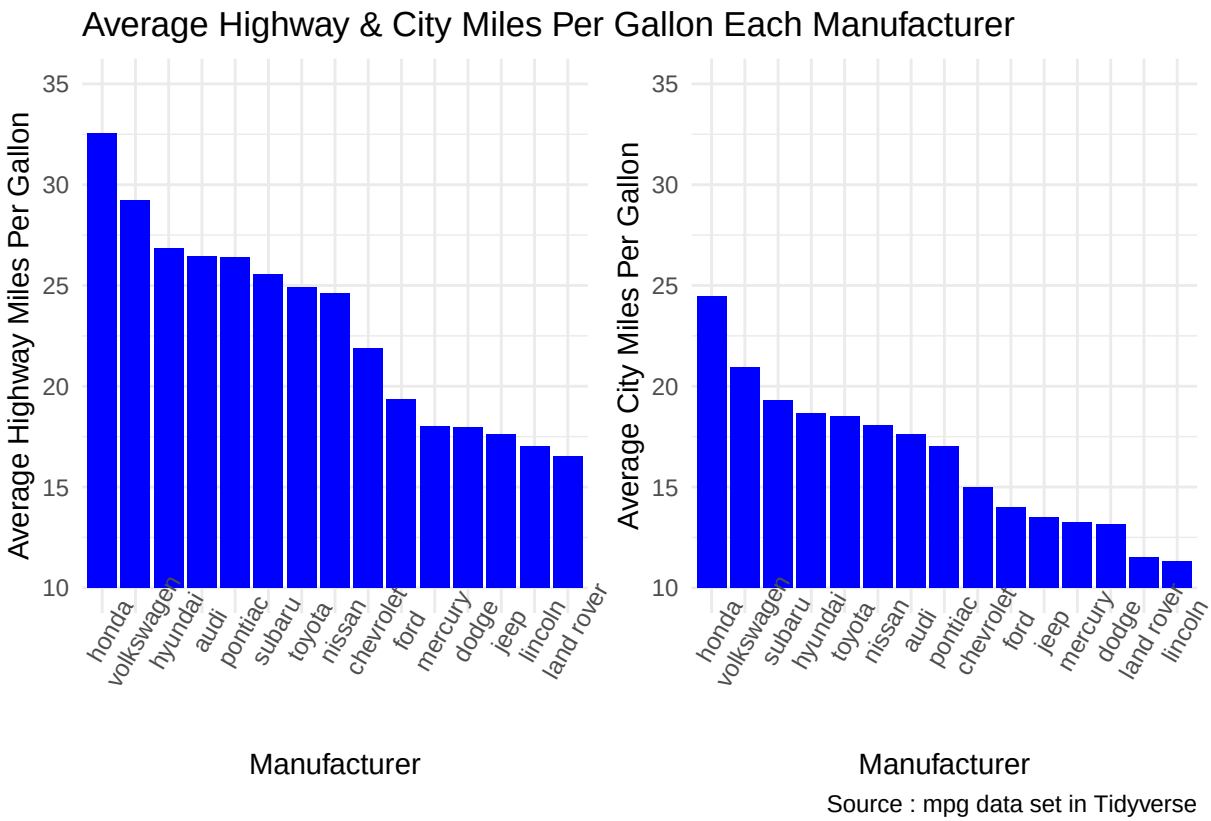
Data Description

Vairable	Description
manufacturer	manufacturer name
model	model name
displ	engine displacement, in litres
year	year of manufacture
cyl	number of cylinders
trans	type of transmission
drv	the type of drivetrain, where f = front-wheel drive, 4 = 4wd
cty	city miles per gallon
hwy	highway per gallon

Vairable	Description
fl	fuel type
class	type of car

Show Average cty and hwy in each manufacturer

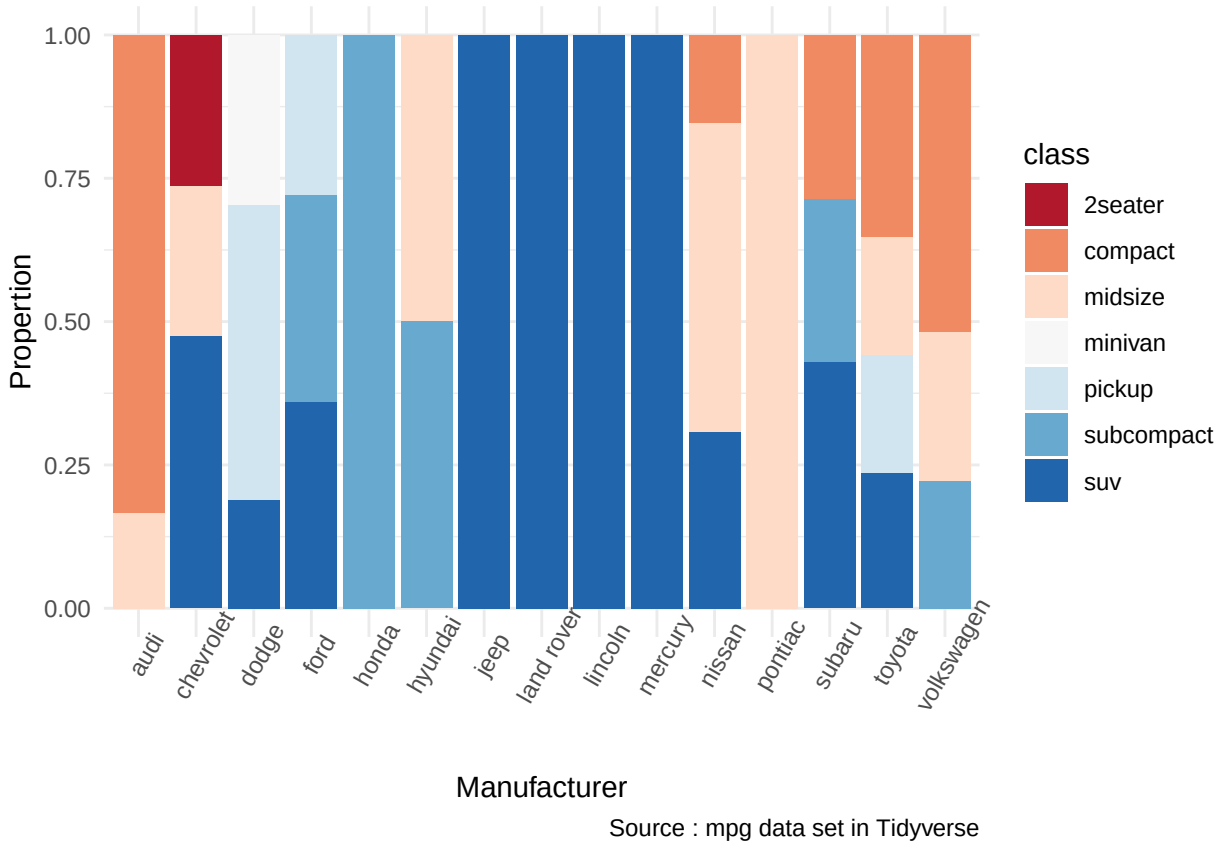
```
avg_manu<- mpg %>%
  group_by(manufacturer) %>%
  summarise( avg_hwy = mean(hwy),
             avg_cty = mean(cty)) %>%
  arrange(desc(avg_hwy))

bar1 <- ggplot(avg_manu, aes(x = reorder(manufacturer, -avg_hwy), y =avg_hwy )) +
  labs(x = "Manufacturer",
       y = "Average Highway Miles Per Gallon",
       title = "Average Highway & City Miles Per Gallon Each Manufacturer") +
  geom_col( fill = "blue") +
  scale_y_continuous(limit = c(10,35),
                     oob = scales::squish) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 60))

bar2 <- ggplot(avg_manu, aes(x = reorder(manufacturer, -avg_cty), y =avg_cty )) +
  labs(x = "Manufacturer",
       y = "Average City Miles Per Gallon",
       caption = "Source : mpg data set in Tidyverse") +
  geom_col( fill = "blue") +
  scale_y_continuous(limit = c(10,35),
                     oob = scales::squish) +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 60))

bar1 + bar2
```



From this chart show Honda and Volksawagen is Top 2 manufacturer that produce cars have great fuel-efficient when measure by average highway&city miles per gallon

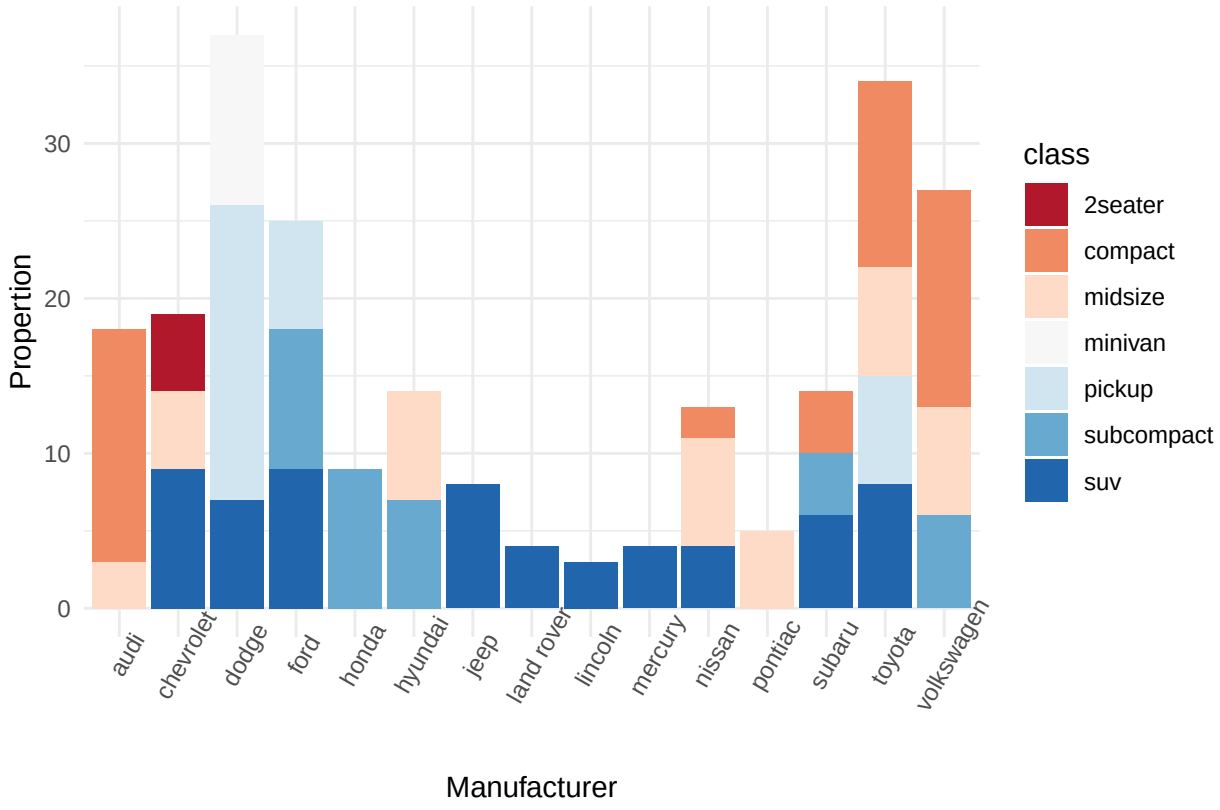
Show propertion of type car in Each Manufacturer

```
prop1 <- ggplot(mpg,aes(manufacturer, fill = class)) +
  geom_bar(position = "fill") +
  scale_fill_brewer(palette = "RdBu") +
   labs(x = "Manufacturer",
        y = "Propertion",
        caption = "Source : mpg data set in Tidyverse")  +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 60))

prop2 <- ggplot(mpg,aes(manufacturer, fill = class)) +
  geom_bar() +
  scale_fill_brewer(palette = "RdBu") +
   labs(x = "Manufacturer",
        y = "Propertion",
        caption = "Source : mpg data set in Tidyverse")  +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 60))


prop1
```



Source : mpg data set in Tidyverse

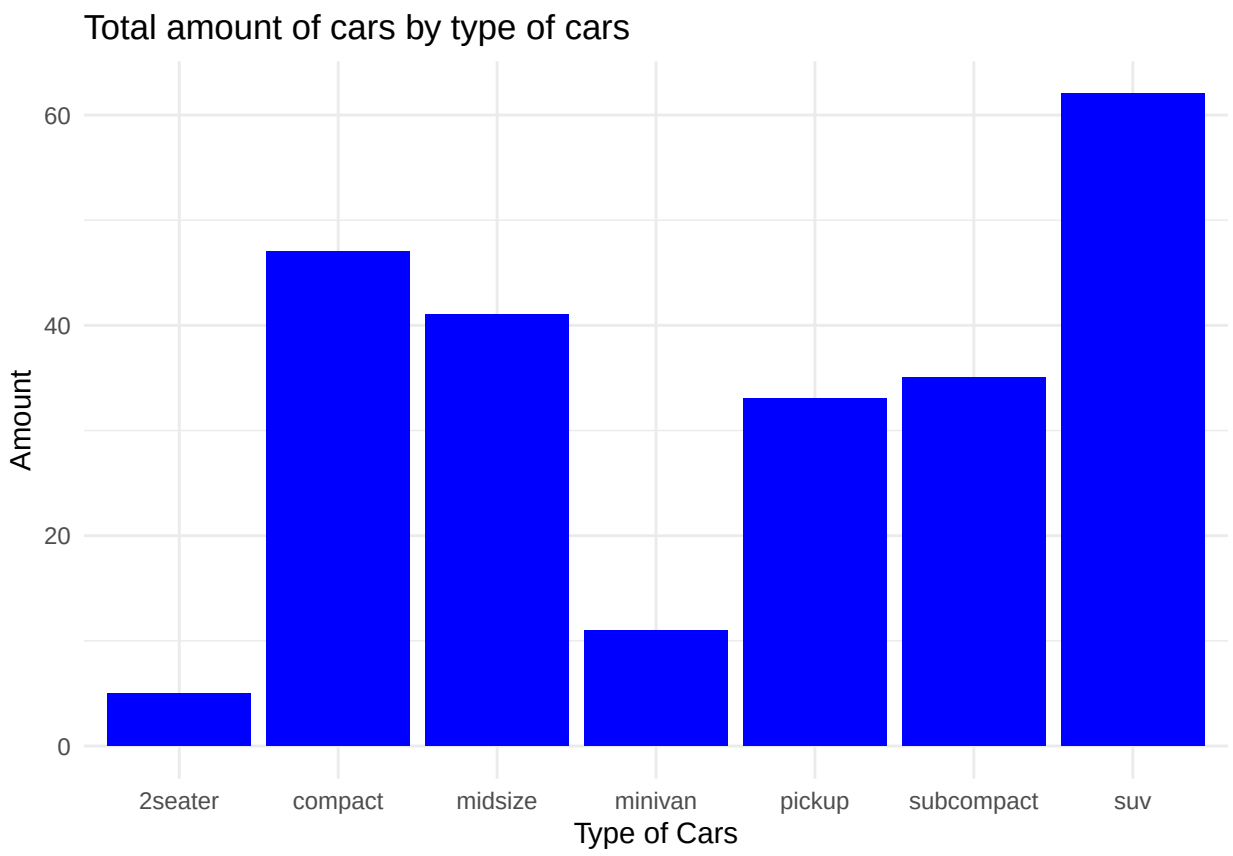
```
prop2
```



Source : mpg data set in Tidyverse

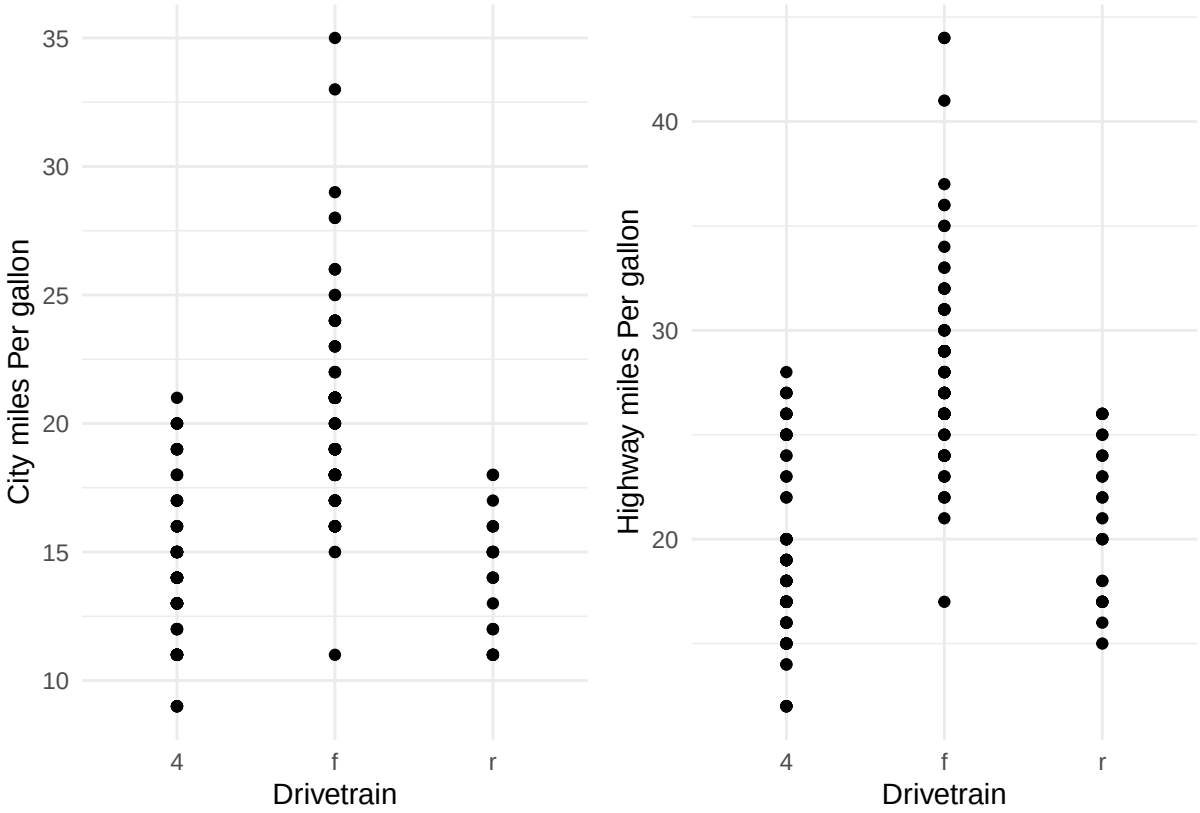
How many amount of car in Each Class

```
ggplot(mpg, aes(class)) +
  geom_bar(fill = "blue") +
  labs( x = "Type of Cars",
        y = "Amount",
        title = "Total amount of cars by type of cars") +
  theme_minimal()
```



Show correlation between DRV and cty&hwy

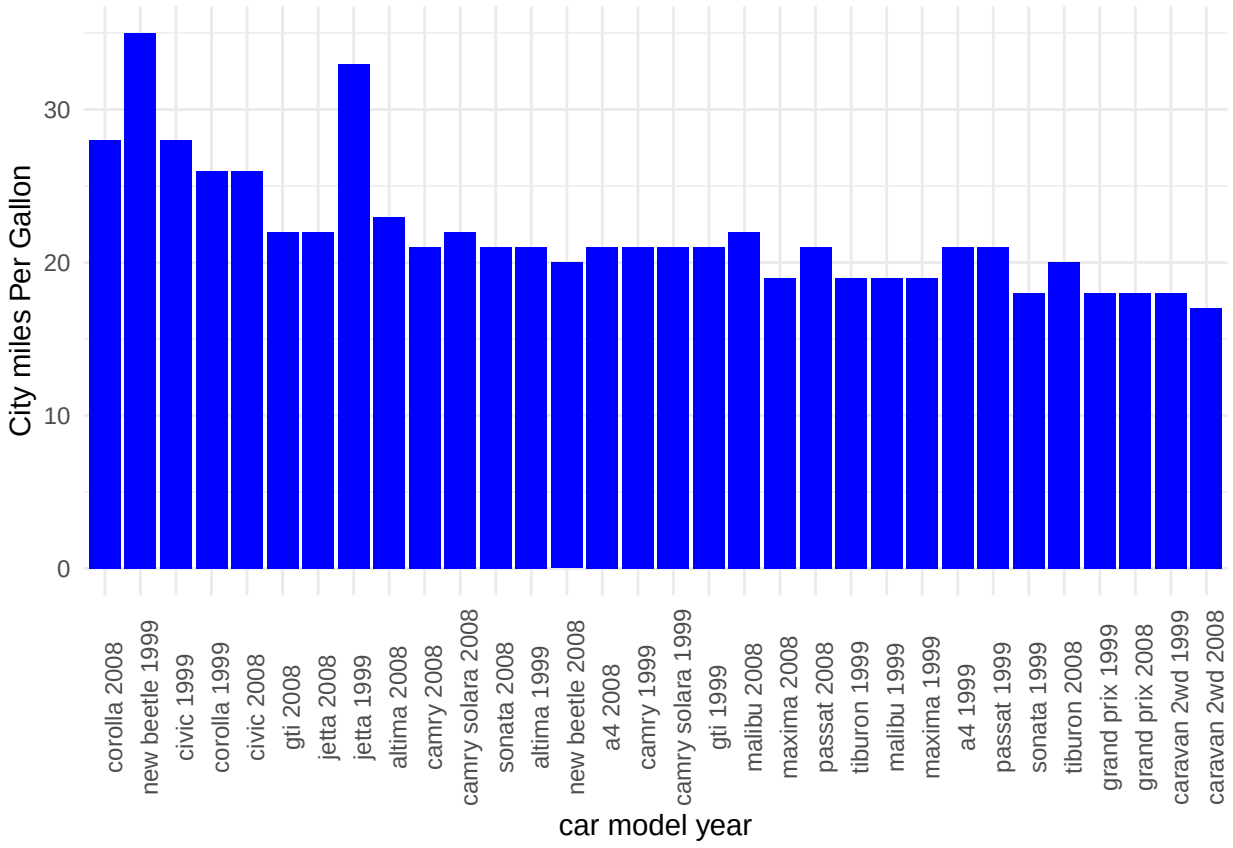
```
cty_drv <- ggplot(mpg, aes(drv, cty)) +
  geom_point() +
  labs( x = "Drivetrain",
        y = "City miles Per gallon") +
  theme_minimal()

hwy_drv <- ggplot(mpg, aes(drv, hwy)) +
  geom_point() +
  labs( x = "Drivetrain",
        y = "Highway miles Per gallon") +
  theme_minimal()

cty_drv + hwy_drv
```



From this chart, Front-Wheel Drivetrain is most fuel-efficient in two measurement

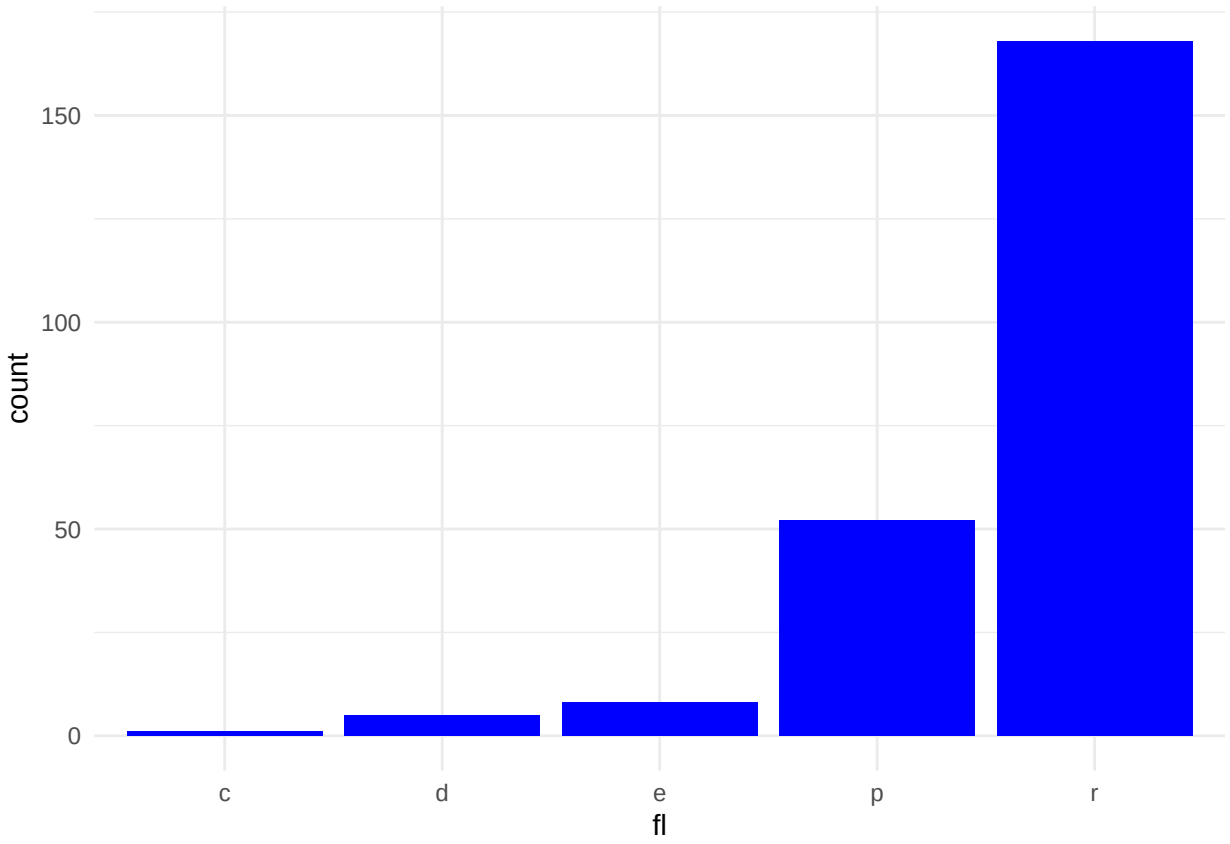
Find model in Front-Wheel Drive train is most Fuel-efficient

```
drv_f_cty <- mpg %>%
  filter(drv == "f") %>%
  mutate(model = paste(model, year)) %>%
  select(model, cty, hwy, manufacturer) %>%
  arrange(desc(cty))



ggplot(drv_f_cty, aes(reorder(model, -cty), cty)) +
  geom_col(position = "dodge",
           fill = "blue") +
  labs( x = "car model year",
        y = "City miles Per Gallon") +
  theme_minimal() +
  theme(axis.text.x = element_text(angle = 90))
```



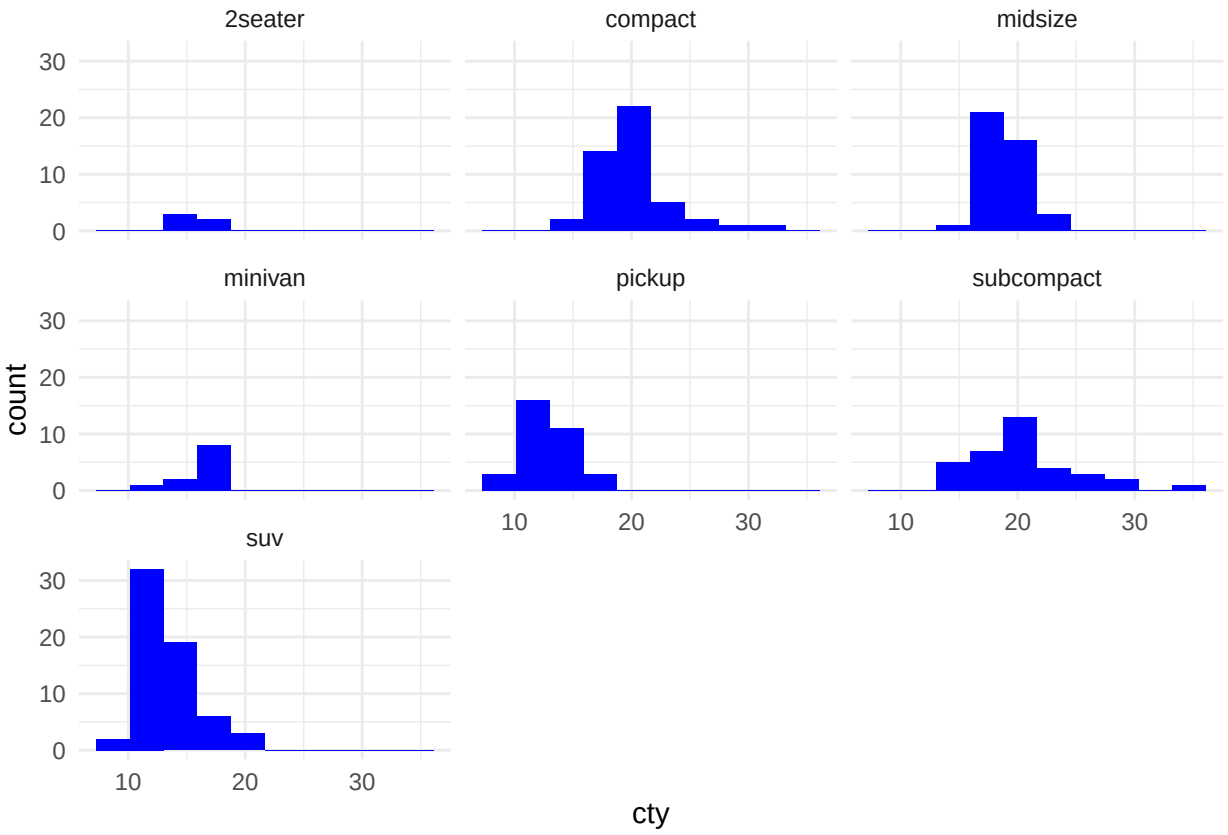
What is the popular fuel type

```
ggplot(mpg, aes(fl)) +
  geom_bar( fill = "blue") +
  theme_minimal()
```



Histogram of City miles Pergallon for each car class

```
ggplot(mpg, aes(cty)) +
  geom_histogram(bins = 10, fill = "blue") +
  facet_wrap(~class) +
  theme_minimal()
```


2seater
compact
midsize
minivan
pickup
subcompact
suv
count
cty
0
10
20
30