Finding Correlations between Aspects of Automobile Design and Performance

Pattarawadee Saranukul

Explore Dataframe

```
head(mtcars)
```

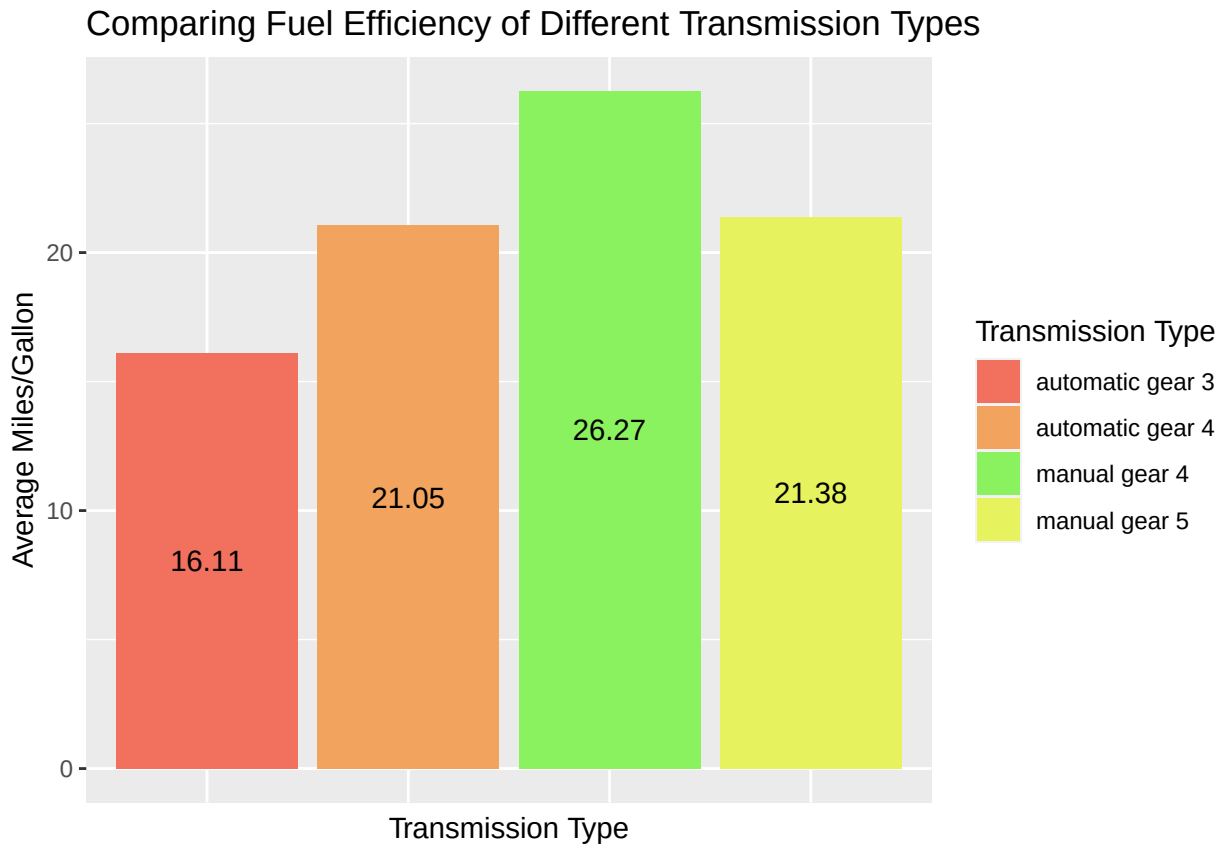
```
##                   mpg cyl disp  hp drat    wt  qsec vs am gear carb
## Mazda RX4         21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
## Datsun 710        22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
## Hornet 4 Drive    21.4   6  258 110 3.08 3.215 19.44  1  0    3    1
## Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02  0  0    3    2
## Valiant           18.1   6  225 105 2.76 3.460 20.22  1  0    3    1
```

Prepare Data

```
# library
library(tidyverse)
library(ggthemes)
library(patchwork)
library(tibble)
library(dplyr)
library(stringr)
library(ggrepel)
library(latexpdf)
library(tinytex)

# Change row names to column
mtcars <- rownames_to_column(mtcars, "model")

# Transform data
mtcars_update <- mtcars %>%
  mutate(am = case_when(
    am == 0 ~ "automatic",
    am == 1 ~ "manual"
  ),
  vs = case_when(
    vs == 0 ~ "V-shaped",
    vs == 1 ~ "straight"
  ))
```

1. Comparing Fuel Efficiency of Different Transmission Types

```r
# Calculate average fuel efficiency by transmission type
base1 <- mtcars_update %>%
  group_by(am, gear) %>%
  summarise(avg_mpg = round(mean(mpg), 2))

# join two columns
base1 <- base1 %>%
  mutate(transmission_type = str_c(am," gear ", gear))

# plot
ggplot(base1, aes(transmission_type, avg_mpg, fill = transmission_type)) +
  geom_col() +
  theme(axis.text.x=element_blank(),
        axis.ticks.x=element_blank()) +
  scale_fill_manual(values=c("#f2705e", "#f2a35e", "#8af25e", "#e6f25e")) +
  labs(fill = "Transmission Type") +
  geom_text(aes(label = avg_mpg),
            position = position_stack(vjust = 0.5)) +
  ggtitle("Comparing Fuel Efficiency of Different Transmission Types") +
  xlab("Transmission Type") + ylab("Average Miles/Gallon")
```

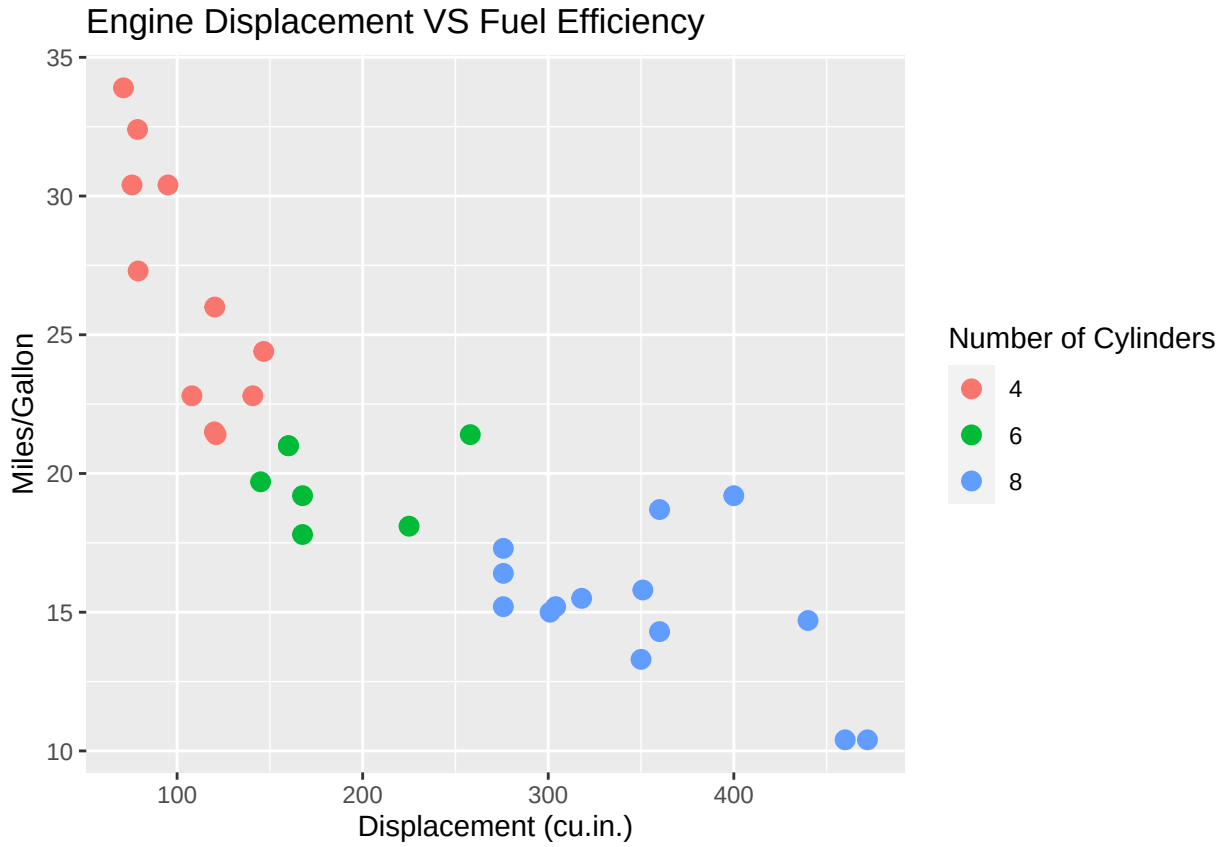


The graph shows that both transmission type and number of gears affect fuel efficiency with manual gear 4 being the most fuel-efficient, followed by manual gear 5 and automatic gear 4, and the least being automatic

gear 3.

2. Engine Displacement VS Fuel Efficiency

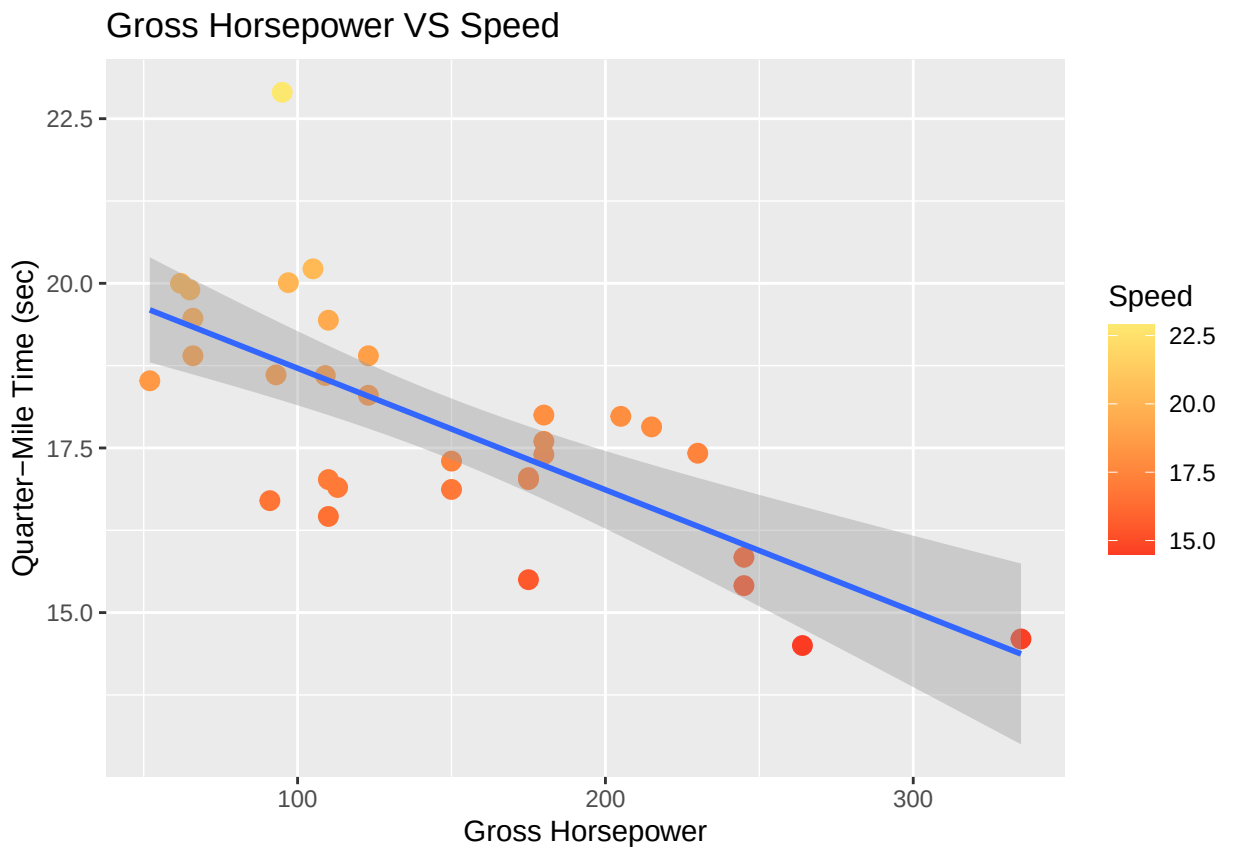
```
# as.factor() to convert continuous value into discreet value
ggplot(mtcars, aes(disp, mpg, color = as.factor(cyl))) +
  geom_point(size = 3) +
  labs(color = "Number of Cylinders") +
  ggtitle("Engine Displacement VS Fuel Efficiency") +
  xlab("Displacement (cu.in.)") + ylab("Miles/Gallon")
```



It could be seen that the higher an engine's displacement is, the more fuel it can consume. The number of cylinders also directly vary with engine's displacement.

3. Gross Horsepower VS Speed

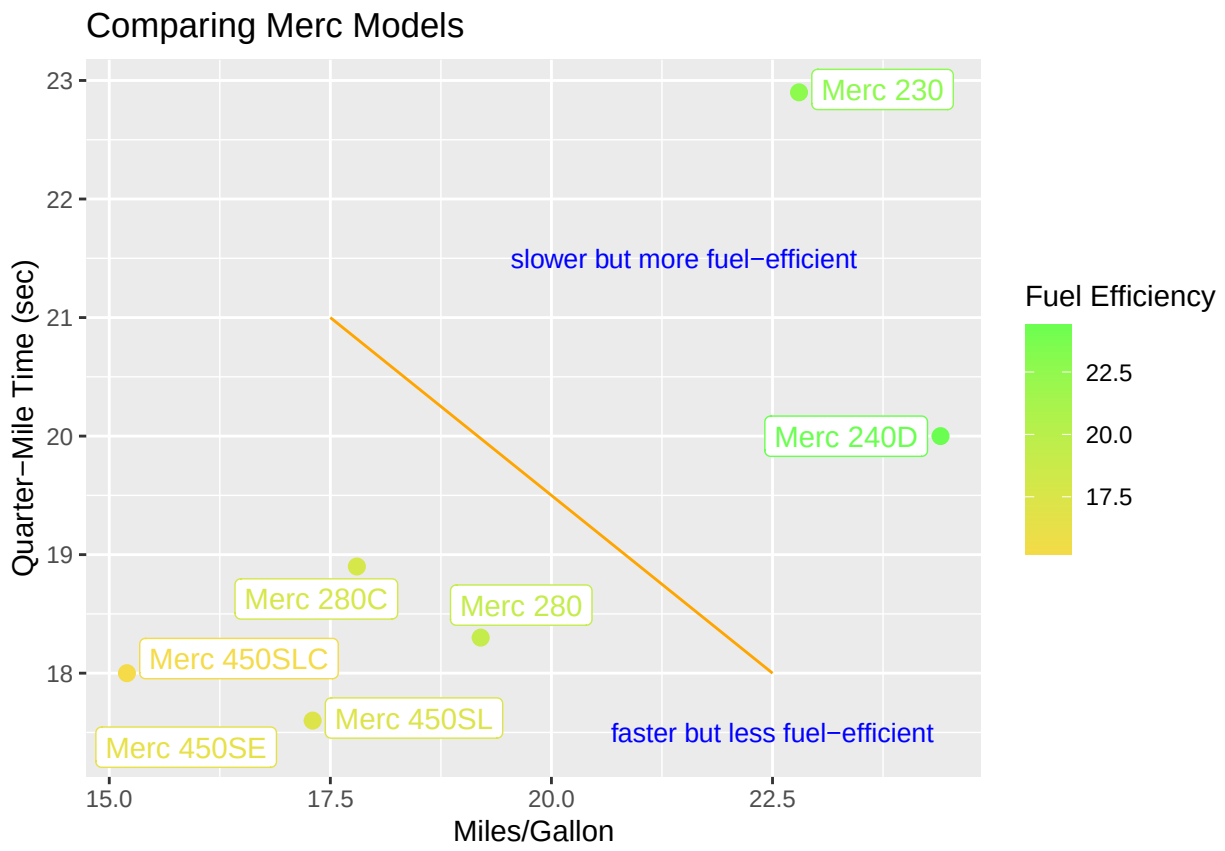
```
ggplot(mtcars, aes(hp, qsec, color = qsec)) +
  geom_point(size = 3) +
  scale_color_gradient(low = "#fb3c23", high = "#fde86f") +
  labs(color = "Speed") +
  geom_smooth(method=lm, se=TRUE) +
  ggtitle("Gross Horsepower VS Speed ") +
  xlab("Gross Horsepower") + ylab("Quarter-Mile Time (sec)")
```



Similarly to the previous graph, gross horsepower also directly varies with car's speed. In other words, the more gross horse power, the faster the car.

4. Comparing Merc Models

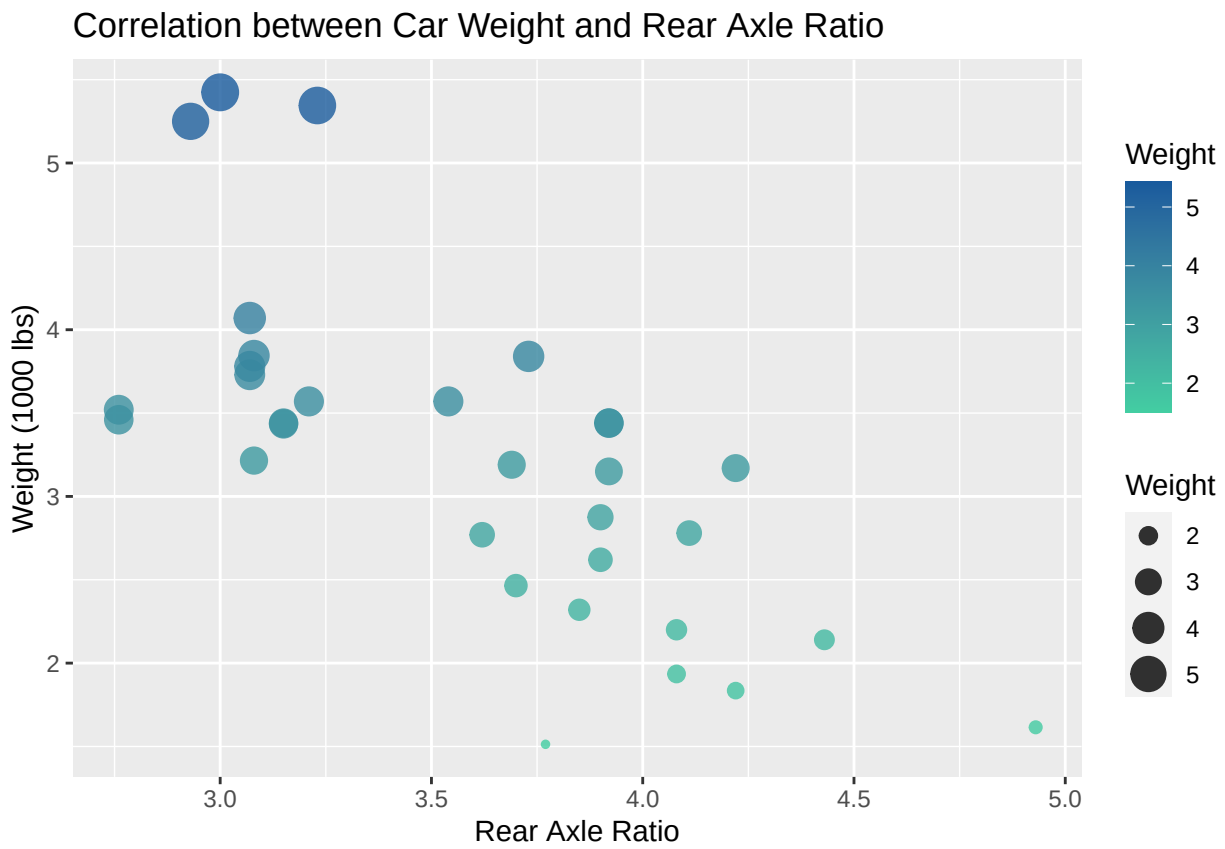
```
# subset data
merc_mtcars <- mtcars_update %>% filter(grepl("^Merc", model))

# plot
ggplot(merc_mtcars, aes(mpg, qsec, color = mpg)) +
  geom_point(size = 2.5) +
  scale_color_gradient(low = "#f4db47", high = "#6aff50") +
  labs(size = "Weight (1000 lbs)", color = "Fuel Efficiency") +
  ggtitle("Comparing Merc Models ") +
  xlab("Miles/Gallon") + ylab("Quarter-Mile Time (sec)") +
  annotate("segment", x = 17.5, xend = 22.5, y = 21, yend = 18,
           colour = "orange") +
  geom_label_repel(aes(label = model)) +
  annotate("text", label = "slower but more fuel-efficient",
    x = 21.5, y = 21.5, size = 3.5, colour = "blue") +
  annotate("text", label = "faster but less fuel-efficient",
           x = 22.5, y = 17.5, size = 3.5, colour = "blue")
```



The graph is divided into two segments. The left-handed segment shows 5 models, Merc 280, Merc 280C, Merc 450SL, Merc 450SLC, Merc 450SE, that are *faster but less fuel-efficient* while the right-handed segment shows 2 models, Merc 230 and Merc 240 D, that are slower but more fuel-efficient.

5. Correlation between Car Weight and Rear Axle Ratio

```
ggplot(mtcars_update, aes(drat, wt, size = wt, color = wt)) +
  geom_point(alpha = 0.8) +
  scale_color_gradient(low = "#43cea2", high = "#185a9d") +
  labs(size = "Weight", color = "Weight") +
  ggtitle("Correlation between Car Weight and Rear Axle Ratio") +
  xlab("Rear Axle Ratio") + ylab("Weight (1000 lbs)")
```



Rear axle ratio represents the number of revolutions the driveshaft must make to spin the axle one full turn. From the graph it seems like the vehicle weight doesn't affect rear axle ratio as cars with varied weight got similar ratio.