PS3 Solutions

Jingle Fu, Chaitanya Venkateswaran

1 R Code

```
#1.1
data(mtcars)


#1.2
ncol(mtcars)
```

[1] 11

```
nrow(mtcars)
```

[1] 32

```
#1.3
library(dplyr)
selected_data <- select(mtcars, mpg, cyl, hp)
mtcars %>% select('mpg', 'cyl', 'hp')
```

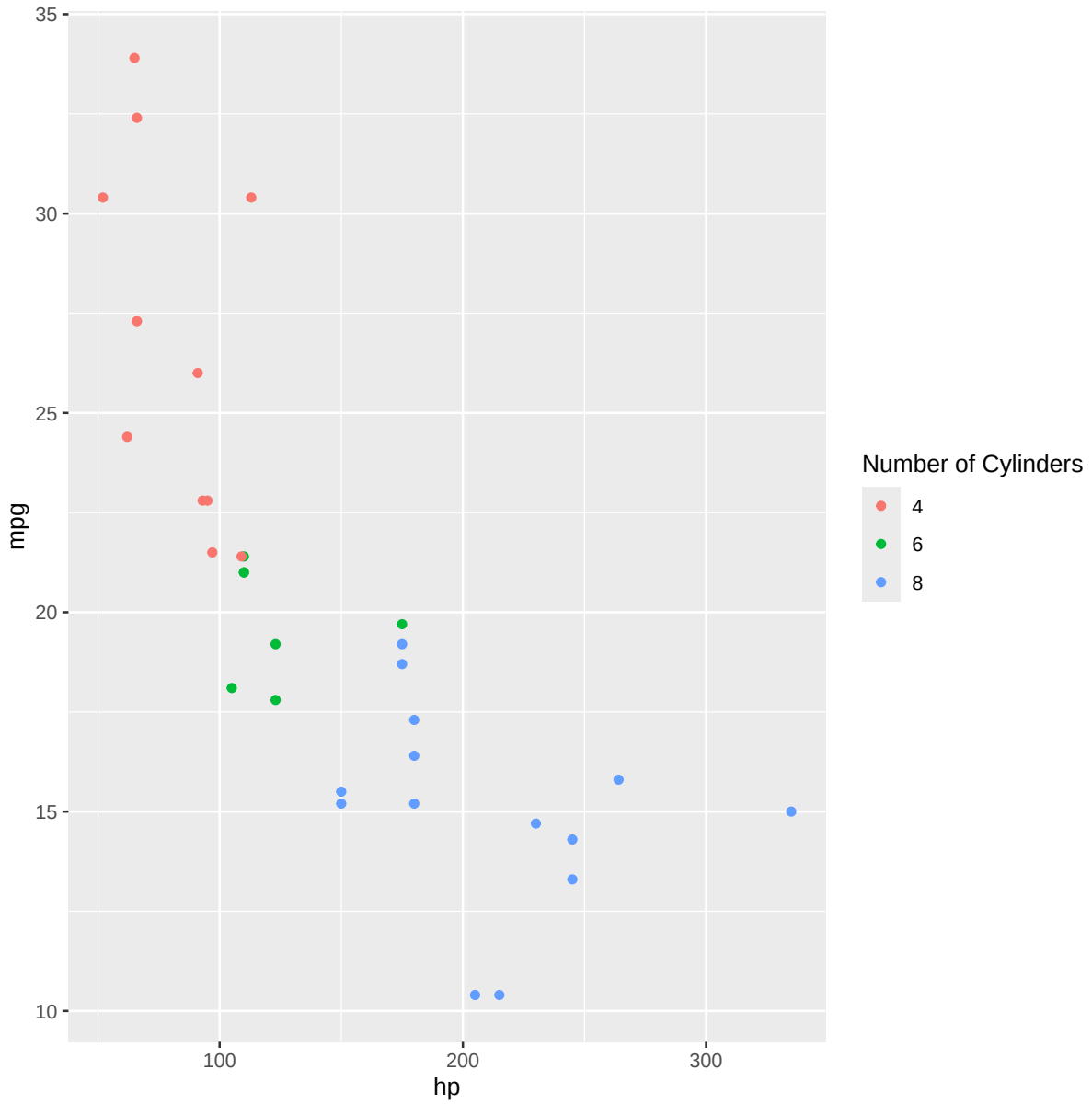
mpg cyl hp Mazda RX4 21.0 6 110 Mazda RX4 Wag 21.0 6 110 Datsun 710 22.8 4 93 Hornet 4 Drive 21.4 6 110 Hornet Sportabout 18.7 8 175 Valiant 18.1 6 105 Duster 360 14.3 8 245 Merc 240D 24.4 4 62 Merc 230 22.8 4 95 Merc 280 19.2 6 123 Merc 280C 17.8 6 123 Merc 450SE 16.4 8 180 Merc 450SL 17.3 8 180 Merc 450SLC 15.2 8 180 Cadillac Fleetwood 10.4 8 205 Lincoln Continental 10.4 8 215 Chrysler Imperial 14.7 8 230 Fiat 128 32.4 4 66 Honda Civic 30.4 4 52 Toyota Corolla 33.9 4 65 Toyota Corona 21.5 4 97 Dodge Challenger 15.5 8 150 AMC Javelin 15.2 8 150 Camaro Z28 13.3 8 245 Pontiac Firebird 19.2 8 175 Fiat X1-9 27.3 4 66 Porsche 914-2 26.0 4 91 Lotus Europa 30.4 4 113

Ford Pantera L 15.8 8 264 Ferrari Dino 19.7 6 175 Maserati Bora 15.0 8 335 Volvo 142E 21.4 4 109

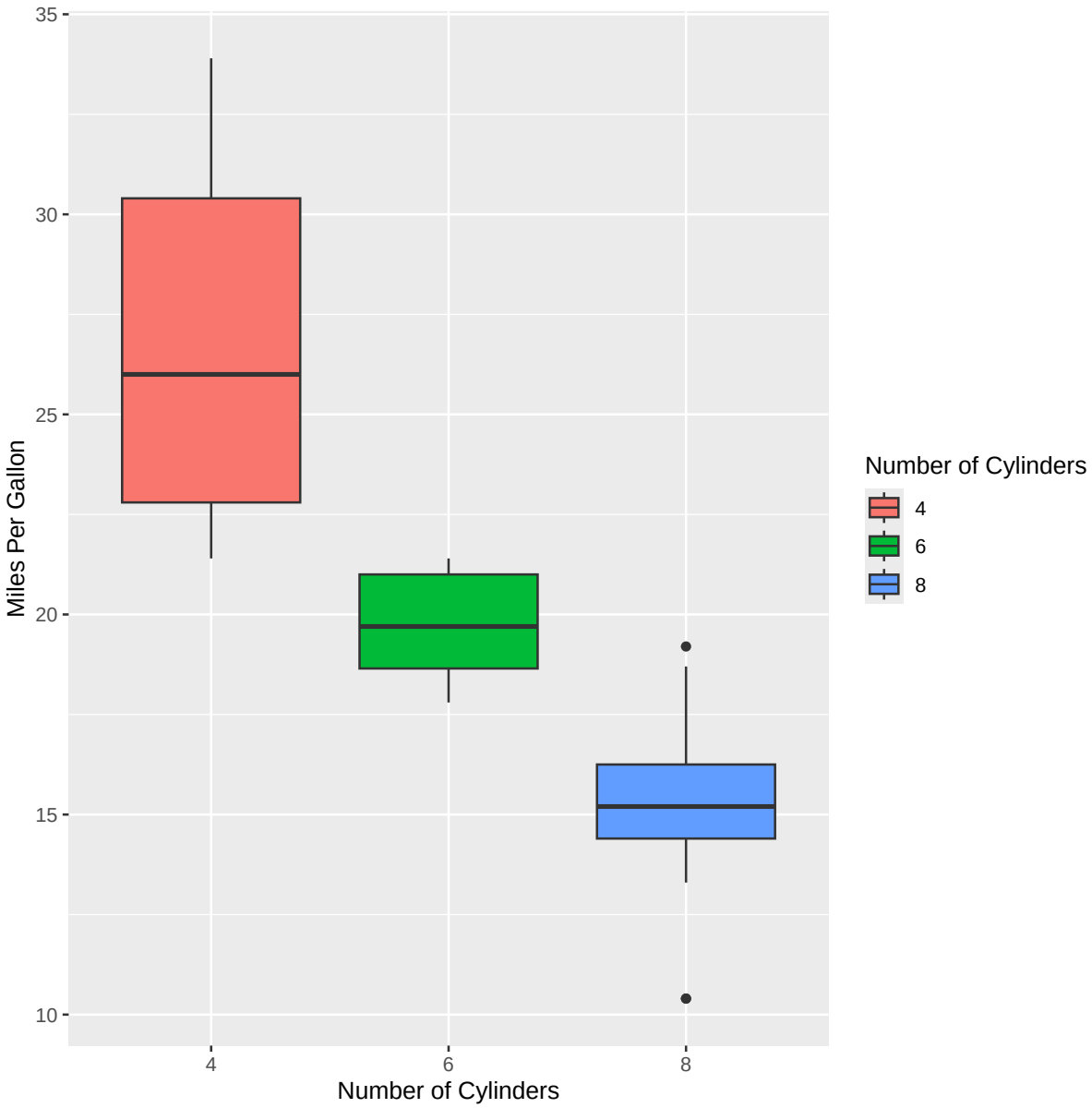
```
#1.4
data1_4 <- filter(mtcars, cyl == 6, hp > 100)
mtcars %>% filter(cyl == 6, hp > 100)
```

mpg cyl disp hp drat wt qsec vs am gear carb Mazda RX4 21.0 6 160.0 110 3.90 2.620 16.46 0 1 4 4 Mazda RX4 Wag 21.0 6 160.0 110 3.90 2.875 17.02 0 1 4 4 Hornet 4 Drive 21.4 6 258.0 110 3.08 3.215 19.44 1 0 3 1 Valiant 18.1 6 225.0 105 2.76 3.460 20.22 1 0 3 1 Merc 280 19.2 6 167.6 123 3.92 3.440 18.30 1 0 4 4 Merc 280C 17.8 6 167.6 123 3.92 3.440 18.90 1 0 4 4 Ferrari Dino 19.7 6 145.0 175 3.62 2.770 15.50 0 1 5 6

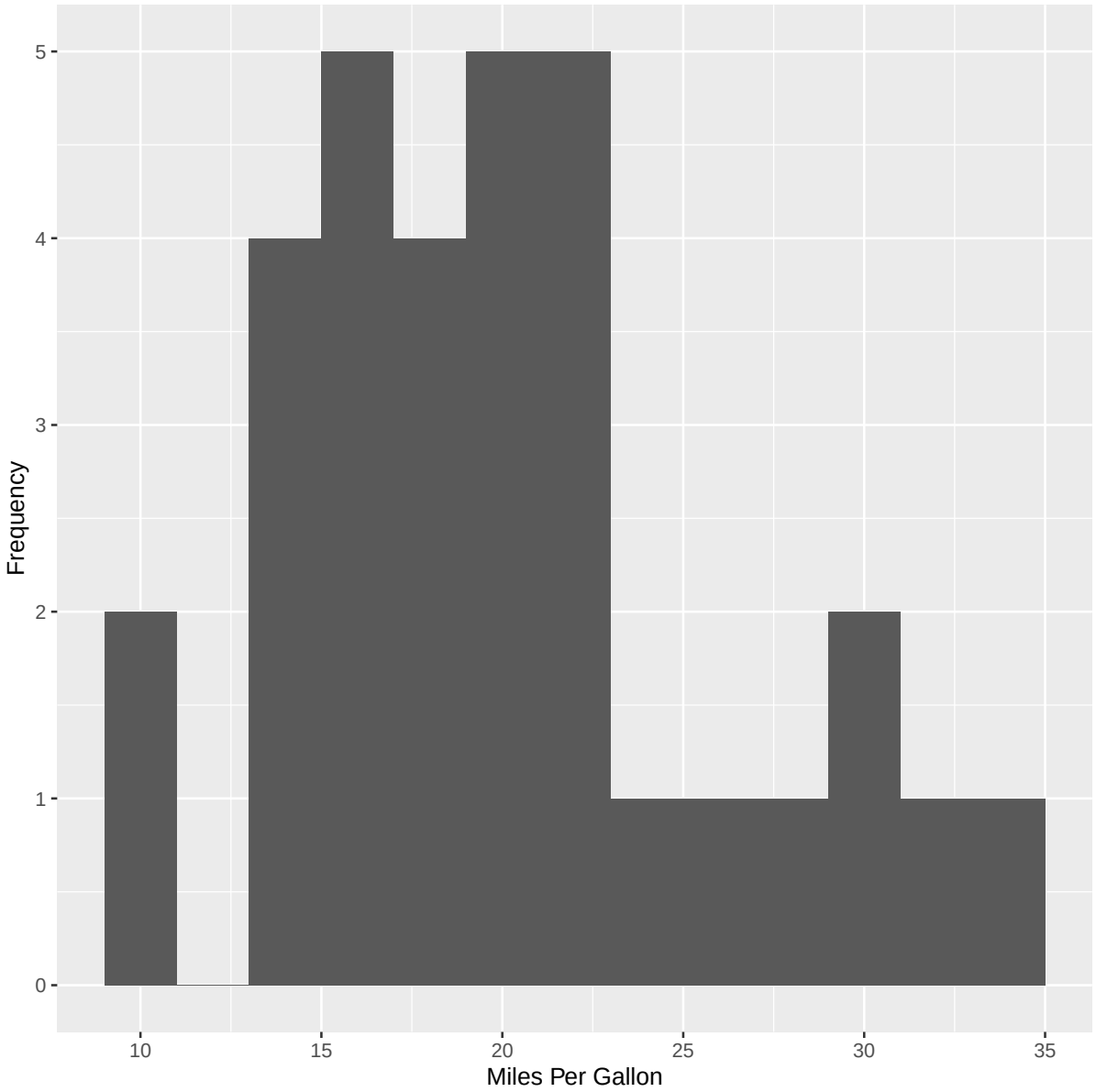
```
#1.5
mtcars$hp_per_cyl <- mtcars$hp / mtcars$cyl


#1.6
average_mpg <- mtcars %>%
  group_by(cyl) %>%
  summarize(avg_mpg = mean(mpg))


#2.1
library(ggplot2)
ggplot(mtcars, aes(x = hp, y = mpg, color = as.factor(cyl))) +
  geom_point() +
  labs(color = "Number of Cylinders")
```




```
#2.2
ggplot(mtcars, aes(x = as.factor(cyl), y = mpg, fill = as.factor(cyl))) +
  geom_boxplot() +
  scale_fill_discrete(name = "Number of Cylinders") +
  labs(x = "Number of Cylinders", y = "Miles Per Gallon")
```




```
#2.3
ggplot(mtcars, aes(x = mpg)) +
  geom_histogram(binwidth = 2) +
  xlab("Miles Per Gallon") +
  ylab("Frequency")
```




```r
#3.1
mean_mpg <- mean(mtcars$mpg)
mean_hp <- mean(mtcars$hp)

#3.2
var_mpg <- var(mtcars$mpg)
var_hp <- var(mtcars$hp)

#3.3
cov_mpg_hp <- cov(mtcars$mpg, mtcars$hp)
```

```r
#3.4
cor_mpg_hp <- cor(mtcars$mpg, mtcars$hp)

#4.1
car_names <- data.frame(
  car_model = rownames(mtcars),
  origin = c(rep('USA', 10), rep('Europe', 10), rep('Japan', 12))
)

mtcars$car_model <- rownames(mtcars)
merged_data <- merge(mtcars, car_names, by = "car_model")

#4.2
library(tidyr)
long_format <- pivot_longer(mtcars,
                            cols = -car_model,
                            names_to = "variable",
                            values_to = "value")

#4.3
short_format <- pivot_wider(long_format,
                            names_from = variable,
                            values_from = value,
                            id_cols = car_model)

#5.1
library(stargazer)
model <- lm(mpg ~ hp + wt, data = mtcars)

#5.2
summary(model)
```

Call: lm(formula = mpg hp + wt, data = mtcars)

Residuals: Min 1Q Median 3Q Max -3.941 -1.600 -0.182 1.050 5.854

Coefficients: Estimate Std. Error t value Pr(>|t|) (Intercept) 37.22727 1.59879 23.285 < 2e-16 *** hp -0.03177 0.00903 -3.519 0.00145 ** wt -3.87783 0.63273 -6.129 1.12e-06 *** — Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 2.593 on 29 degrees of freedom Multiple R-squared: 0.8268, Adjusted R-squared: 0.8148 F-statistic: 69.21 on 2 and 29 DF, p-value: 9.109e-12

```
stargazer(model, type = "latex",
          out = "PS3-5_2.tex",
          title = "Regression Results",
          single.row = TRUE,
          header = FALSE, no.space = TRUE)
```

Table 1: Regression Results

	Dependent variable:
	mpg
hp	-0.032^{***} (0.009)
wt	-3.878^{***} (0.633)
Constant	37.227^{***} (1.599)
Observations	32
R^2	0.827
Adjusted R^2	0.815
Residual Std. Error	2.593 (df = 29)
F Statistic	69.211^{***} (df = 2; 29)
Note:	*p<0.1; **p<0.05; ***p<0.01

```
#5.3
predict_mpg <- predict(model, newdata = data.frame(hp = 150, wt = 3.0))
stargazer(predict_mpg, type = "latex")
```

```
source("E:/IHEID/2024Fall/Mathematics and Statistics for Economists (EI071)/PS3.R")
```

Table 2:

1
20.828

Table 3: Regression Results

	Dependent variable:
	mpg
hp	-0.032^{***} (0.009)
wt	-3.878^{***} (0.633)
Constant	37.227^{***} (1.599)
Observations	32
R^2	0.827
Adjusted R^2	0.815
Residual Std. Error	2.593 (df = 29)
F Statistic	69.211^{***} (df = 2; 29)
Note:	*p<0.1; **p<0.05; ***p<0.01