#11 (Part 1)

(a) Fuelcon = 783.8 – 4.1* GASTAX – 0.007* INCOME

```
Call:
lm(formula = FUELCON ~ GASTAX + INCOME, data = FUELCON4)

Residuals:
     Min       1Q   Median       3Q      Max 
-171.762  -34.014    3.668   31.704  200.403 

Coefficients:
              Estimate Std. Error t value Pr(>|t|)    
(Intercept) 783.759824  67.762851  11.566 1.75e-15 ***
GASTAX       -4.132910   1.933209  -2.138 0.037649 *  
INCOME       -0.007316   0.001857  -3.939 0.000264 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 62.12 on 48 degrees of freedom
Multiple R-squared:  0.2931,	Adjusted R-squared:  0.2637 
F-statistic: 9.952 on 2 and 48 DF,  p-value: 0.0002422
```

(b) The assumed per capita fuel consumption in gallons in 50 states and the District of Columbia is $783.8, as 1 unit of the tax of gas increasing, 4.13 units of the consumption would decreasing; As 1 unit of income of a person increasing, 0.007 unit of the consumption would decreasing.

(c) The hypothesis is:

H0: Beta_1 = Beta_2 = 0

Ha: At least one Beta is not zero

```
> qf(0.95,df1 = 4,df2 = 46)
[1] 2.574035
```

F (a; K, n-K-1) = F (0.05; 4, 46) = 2.57

F – statistic = 9.952 > 2.57, so we reject H0. We conclude that no coefficients (Beta_1, Beta_2) is equal to zero. So, Gastax and Income both important in explaining the variation in Fuelcon.

#11 Part 2

(a)

```
lm(formula = FUELCON ~ GASTAX + INCOME + DRIVERS + HWYMILES,
    data = FUELCON4)

Residuals:
     Min       1Q   Median       3Q      Max
-161.696  -33.229    0.745   35.469  161.584

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  9.161e+02  7.327e+01  12.503  < 2e-16 ***
GASTAX      -3.690e+00  1.772e+00  -2.082  0.04290 *
INCOME      -5.492e-03  1.767e-03  -3.108  0.00322 **
DRIVERS     -2.182e+02  6.213e+01  -3.512  0.00101 **
HWYMILES    -7.756e-04  1.006e-03  -0.771  0.44444
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 56.28 on 46 degrees of freedom
Multiple R-squared:  0.444,	Adjusted R-squared:  0.3956
F-statistic: 9.183 on 4 and 46 DF,  p-value: 1.537e-05
```

For DRIVERS,

H0: Beta_D = 0

Ha: Beta_D<> 0

Since 0.001 < 0.05, the null hypotheses is rejected at 5% level of significance. Therefore, we could conclude that DRIVERS is linearly related to FUELCON and significant.

For HWYMILES,

H0: Beta_H = 0

Ha: Beta_H <> 0

Since 0.444 > 0.05, the null hypotheses is fail to reject at 5% level of significance. Therefore, we could conclude that HWYMILES is not linearly related to FUELCON and not significant.

(b)

If a person want to reduce the cost on fuel during a year, he need not to care the how many number of miles of federally funded highways (HWYMILES) of his car runs, what the person need to care is to use private motor vehicles and earn more salary.

(c)

35	Ohio	458.31	0.74	16807	22.00	28619

```
> predict(Fuelcon_3, data.frame(DRIVERS = 0.74, GASTAX = 22, INCOME = 28619), interval = "confidence",level = 0.95)
       fit      lwr      upr
1 507.6525 485.4509 529.8542
```

Ohio is a state, so we use confidence.

We are 95% sure that the per capita fuel consumption of the state of Ohio for next year with 22 tax per gallon of gasoline in cents (GASTAX), 0.74 ratio of licensed drivers to private and commercial motor vehicles registered (DRIVERS) and 28619 of the average household income in dollars (INCOME) will be between $485.45 and $529.85.

#17

(a)

```
Call:
lm(formula = WINS ~ BA + ERA + HR, data = MLB4)

Residuals:
    Min      1Q  Median      3Q     Max 
-9.1555 -2.5054  0.4665  2.1392  9.1389 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) -21.88065   28.92622  -0.756   0.4562    
BA          606.30599  100.76891   6.017 2.36e-06 ***
ERA         -16.89736    1.75825  -9.610 4.82e-10 ***
HR            0.09759    0.03572   2.732   0.0112 *  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 5.002 on 26 degrees of freedom
Multiple R-squared:  0.8969,	Adjusted R-squared:  0.885 
F-statistic: 75.42 on 3 and 26 DF,  p-value: 5.895e-13
```

For HR,

H0: Beta_HR = 0

Ha: Beta_HR <> 0

Since 0.0112 < 0.05, the null hypotheses is rejected at 5% level of significance. Therefore, we could conclude that HR is linearly related to FUELCON and significant.

(b)

Based on the output of p-value, we know all the variables are important. We also can know that the positive coefficients are BA and HR, the negative coefficient is ERA. Thus, the teams with higher number of home runs and average batting and lower earned run average tend to be more successful.

(c)

5	Cleveland Indians	74	192	0.249	4.91

```
> predict(Mlb_3, data.frame(BA = 0.249, ERA = 4.91, HR = 192), interval = "predict",level = 0.95)
       fit      lwr      upr
1 64.86106 53.75506 75.96707
> 
```

There is only one team, so we use predict.

We are 95% sure that the individual team - Cleveland Indians - with 192 HR, 0.249 BA, and 4.91 ERA would win 54 to 76 times next year.