#3.

1&2. I have performed linear regression on mpg with various predictors ad also their summary.

Cylinders:

```
> a1<-read.csv("G:/Fall Semester 2017/ISL/Auto-rev.csv",header=TRUE)
> lmod=lm(mpg~cylinders, data=a2)
Error in is.data.frame(data) : object 'a2' not found
> lmod=lm(mpg~cylinders, data=a1)
> summary(lmod)

Call:
lm(formula = mpg ~ cylinders, data = a1)

Residuals:
     Min       1Q   Median       3Q      Max
-14.2413  -3.1832  -0.6332   2.5491  17.9168

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)  42.9155     0.8349   51.40   <2e-16 ***
cylinders    -3.5581     0.1457  -24.43   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.914 on 390 degrees of freedom
Multiple R-squared:  0.6047,    Adjusted R-squared:  0.6037
F-statistic: 596.6 on 1 and 390 DF,  p-value: < 2.2e-16
```

Displacement:

```
> lmod=lm(mpg~displacement, data=a1)
> summary(lmod)

Call:
lm(formula = mpg ~ displacement, data = a1)

Residuals:
     Min       1Q   Median       3Q      Max
-12.9170  -3.0243  -0.5021   2.3512  18.6128

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  35.12064    0.49443   71.03   <2e-16 ***
displacement -0.06005    0.00224  -26.81   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.635 on 390 degrees of freedom
Multiple R-squared:  0.6482,    Adjusted R-squared:  0.6473
F-statistic: 718.7 on 1 and 390 DF,  p-value: < 2.2e-16
```

Weight:

```
> lmod=lm(mpg~weight, data=a1)
> summary(lmod)

Call:
lm(formula = mpg ~ weight, data = a1)

Residuals:
     Min       1Q   Median       3Q      Max
-11.9736  -2.7556  -0.3358   2.1379  16.5194

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 46.216524   0.798673   57.87   <2e-16 ***
weight      -0.007647   0.000258  -29.64   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.333 on 390 degrees of freedom
Multiple R-squared:  0.6926,    Adjusted R-squared:  0.6918
F-statistic: 878.8 on 1 and 390 DF,  p-value: < 2.2e-16
```

Acceleration

```
> lmod=lm(mpg~acceleration, data=a1)
> summary(lmod)

Call:
lm(formula = mpg ~ acceleration, data = a1)

Residuals:
    Min      1Q  Median      3Q     Max
-17.989  -5.616  -1.199   4.801  23.239

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)    4.8332     2.0485   2.359   0.0188 *
acceleration   1.1976     0.1298   9.228   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 7.08 on 390 degrees of freedom
Multiple R-squared:  0.1792,    Adjusted R-squared:  0.1771
F-statistic: 85.15 on 1 and 390 DF,  p-value: < 2.2e-16
```

Year:

```
> lmod=lm(mpg~year, data=a1)
> summary(lmod)

Call:
lm(formula = mpg ~ year, data = a1)

Residuals:
     Min       1Q   Median       3Q      Max
-12.0212  -5.4411  -0.4412   4.9739  18.2088

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) -70.01167    6.64516  -10.54   <2e-16 ***
year          1.23004    0.08736   14.08   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.363 on 390 degrees of freedom
Multiple R-squared:  0.337,     Adjusted R-squared:  0.3353
F-statistic: 198.3 on 1 and 390 DF,  p-value: < 2.2e-16
```

Origin:

```
> lmod=lm(mpg~origin, data=a1)
> summary(lmod)

Call:
lm(formula = mpg ~ origin, data = a1)

Residuals:
     Min       1Q   Median       3Q      Max
-13.2416  -5.2533  -0.7651   3.8967  18.7115

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  14.8120     0.7164   20.68   <2e-16 ***
origin        5.4765     0.4048   13.53   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 6.447 on 390 degrees of freedom
Multiple R-squared:  0.3195,    Adjusted R-squared:  0.3177
F-statistic: 183.1 on 1 and 390 DF,  p-value: < 2.2e-16
```

3. Acceleration has 0.0188 as p-value with least R^2 values has least significant influence on the mpg and then next least significant influences are origin and year.

4.

```
> lmod=lm(mpg~cylinders+displacement+weight, data=al)
> summary(lmod)

Call:
lm(formula = mpg ~ cylinders + displacement + weight, data = al)

Residuals:
     Min       1Q   Median       3Q      Max
-12.5568  -2.8703  -0.3649   2.2708  16.4338

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)  44.3709616  1.4806851  29.967  < 2e-16 ***
cylinders    -0.2677968  0.4130673  -0.648    0.517
displacement -0.0126740  0.0082501  -1.536    0.125
weight       -0.0057079  0.0007139  -7.995  1.5e-14 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 4.297 on 388 degrees of freedom
Multiple R-squared:  0.6993,    Adjusted R-squared:  0.697
F-statistic: 300.8 on 3 and 388 DF,  p-value: < 2.2e-16
```

Here we could see that weight has the least p-value indicating that it's a strong influencing factor but cylinders and displacement has higher P-value. And the R-squared value has come to 0.6993

5.

Here we could see that R^2 value has increased to 0.8215 with all the factors. The total p-value is< 2.2e-16. The predictors like cylinders, horsepower, acceleration got high p-values. But Displacement has a moderate significance on mgp.

```
> lmod=lm(mpg~cylinders+horsepower+displacement+weight+acceleration+year+origin, data=a1)
> summary(lmod)

Call:
lm(formula = mpg ~ cylinders + horsepower + displacement + weight +
    acceleration + year + origin, data = a1)

Residuals:
    Min      1Q  Median      3Q     Max
-9.5903 -2.1565 -0.1169  1.8690 13.0604

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -17.218435   4.644294  -3.707  0.00024 ***
cylinders    -0.493376   0.323282  -1.526  0.12780
horsepower   -0.016951   0.013787  -1.230  0.21963
displacement  0.019896   0.007515   2.647  0.00844 **
weight       -0.006474   0.000652  -9.929  < 2e-16 ***
acceleration  0.080576   0.098845   0.815  0.41548
year          0.750773   0.050973  14.729  < 2e-16 ***
origin        1.426141   0.278136   5.127 4.67e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 3.328 on 384 degrees of freedom
Multiple R-squared:  0.8215,    Adjusted R-squared:  0.8182
F-statistic: 252.4 on 7 and 384 DF,  p-value: < 2.2e-16
```