Major statistics
Econometrics
Dr/ Sara Osama
Alaa Abdallah Mohamed

The aim of this project is to use R statistical Package to answer questions about the impact of indicator variables on the price of the car, which is the dependent variable. In this project, a random sample was chosen from the Cars dataset to answer the seven questions. The conclusions from this study can assist car dealerships in maximizing their sales and car hobbyists to further understand how chosen car qualities affect the selling prices in the used car market.

Model 1:

$Y=\beta_0 + \beta_1 D_{1i} + \beta_2 D_{2i} + \beta_3 D_{3i} + \beta_4 D_{4i} + \beta_5 D_{5i} + u_i$

$Y = 7.0392 - 1.9813 D_{1i} + 0.5425 D_{2i} - 1.2392 D_{3i} - 1.2103 D_{4i} + 1.6326 D_{5i} + u_i$

$H_0: \beta_1 = \beta_2 = \beta_3 = \beta_4 = \beta_5 = 0$

H_1: at least one of them not equal zero

```
Call:
lm(formula = price ~ Make, data = sample_s)

Coefficients:
  (Intercept)   MakeCadillac  MakeChevrolet    MakePontiac       MakeSAAB     MakeSaturn
       7.0392        -1.9813         1.0891         0.5425        -1.2392         1.6326

> summary(model1)

Call:
lm(formula = price ~ Make, data = sample_s)

Residuals:
     Min       1Q   Median       3Q      Max
-3.11030 -0.40680  0.02424  0.45341  2.55057

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept)     7.0392     0.1726  40.790  < 2e-16 ***
MakeCadillac   -1.9813     0.2303  -8.602 4.76e-16 ***
MakeChevrolet   1.0891     0.1907   5.711 2.75e-08 ***
MakePontiac     0.5425     0.2040   2.660  0.00824 **
MakeSAAB       -1.2392     0.2136  -5.802 1.70e-08 ***
MakeSaturn      1.6326     0.2589   6.307 1.04e-09 ***
---
```

Beta 1: the expected price in (Cadillac) is less than the expected price of (Buick) by 1.9813

Beta 2: the expected price in (Chevrolet) is more than the expected price of (Buick) by 1.0891

Beta 3: the expected price in (Pontiac) is more than the expected price of (Buick) by 0.5425

Beta 4: the expected price in (SAAB) is less than the expected price of (Buick) by 1.2392

Beta5: the expected price in (Saturn) is more than the expected price of (Buick) by 1.4917

All estimators are significant here with assuming alpha equal 0.05 that is the mean of the price of the car differs according to the make of the car for all the included car makers.

Model2:

$Y=\beta_0+\beta_1 D_{1i}+ \beta_2 D_{2i}+\beta_3 D_{3i}+ \beta_4 D_{4i}+ \beta_5 D_{5i}+ \beta_6 D_{6i}+ \beta_7 D_{7i}+ \beta_8 D_{8i}+ \beta_9 D_{9i}+ \beta_{10} D_{10i}+ \beta_{11} D_{11i}+u_i$

$Y=5.2418 -1.1092 D_{1i}+0.2522\ D_{2i}+0.3398\beta_3 D_{3i}-2.3198 D_{4i}+ 0.2953 D_{5i}+ 0.6024 D_{6i}+ 1.3823 D_{7i}+ 0.8109 D_{8i}+ 0.2232 D_{9i}+ 2.5041 D_{10i}+ 0.9866 D_{11i}$

$H_0: \beta_1= \beta_2= \beta_3= \beta_4= \beta_5= \beta_6= \beta_7= \beta_8= \beta_9= \beta_{10}= \beta_{11}$=zero

H_1 : at least one doesn't equal zero

```
Coefficients:
     (Intercept)      MakeCadillac     MakeChevrolet       MakePontiac          MakeSAAB       MakeSaturn         TypeCoupe     TypeHatchback
          5.2418           -1.1092            0.2522            0.3398           -2.3198            0.2953            0.6024            1.3823
       TypeSedan         TypeWagon       Cylinderlow  Cylindermoderate
          0.8109            0.2232            2.5041            0.9866

> summary(model2)

Call:
lm(formula = price ~ Make + Type + Cylinder, data = sample_s)

Residuals:
     Min       1Q   Median       3Q      Max
-1.08364 -0.36313 -0.02299  0.32808  1.86991

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)       5.2418     0.2007  26.122  < 2e-16 ***
MakeCadillac     -1.1092     0.1627  -6.818 5.42e-11 ***
MakeChevrolet     0.2522     0.1245   2.027 0.043604 *
MakePontiac       0.3398     0.1252   2.715 0.007023 **
MakeSAAB         -2.3198     0.1545 -15.015  < 2e-16 ***
MakeSaturn        0.2953     0.1713   1.724 0.085834 .
TypeCoupe         0.6024     0.1576   3.822 0.000162 ***
TypeHatchback     1.3823     0.1842   7.504 7.73e-13 ***
TypeSedan         0.8109     0.1407   5.763 2.12e-08 ***
TypeWagon         0.2232     0.1568   1.423 0.155779
Cylinderlow       2.5041     0.1288  19.436  < 2e-16 ***
Cylindermoderate  0.9866     0.1205   8.190 8.60e-15 ***
```

Beta 1: the expected price of (Cadillac) is less than the expected price of (buick) by 1.1092, controlling cylinder and type constant.

Beta 2: the expected price of (Chevrolet) is more than the expected price of (buick) by 0.2522, holding cylinder and type constant.

Beta 3: the expected price of (Pontiac) is more than the expected price of (buick) by 0.3398, holding cylinder and type constant.

Beta 4: the expected price of (saap) is less than the expected price of (buick) by 2.3198, holding cylinder and type constant.

Beta 5: no difference significant between expected price in (Saturn) and (Buick), holding cylinder and type constant.

Beta 6: the expected price of (coupe) is more than expected price of Convertible type by 0.6024, holding make and cylinder constant.

Beta 7: the expected price of hatchback is more than expected price of Convertible by 1.3823, holding make and cylinder constant.

Beta 8: the expected price of sedan is more than expected price of Convertible by 0.8109, holding make and cylinder constant.

Beta 9: no difference significant between expected price in wegon and convertible, holding cylinder and make constant.

Beta 10: the expected price of cylinder-low is more than expected price of cylinder-high by 2.5041, holding make and type constant.

Beta 11: the expected price of cylinder-moderate is more than expected price of cylinder-high by 0.9866, holding make and type constant.

which means that the alternate hypothesis H1 is accepted, meaning that the mean of the price of the car differs according to the make of the car, for any car type, and any Cylinder size

Model 3:

Y=β0 + β1D1i+ β2D2i+ β3D3i+ β4D4i+ β5D5i+ β6D6i+ β7D7i+ β8D1iD6i + β10D3iD6i+ β13D1iD7i+ β14D2iD7i +u

Y= 6.00305 -1.04430D1i-0.76014D2i+ 0.39607D3i-2.09110D4i+ 0.82335D5i+ 1.88808D6i+ 1.03619D7i+ 1.8024D1iD6i + 0.03152D3iD6i-0.58267D1iD7i+ 1.08987D2iD7i

H0: β9 = β10= β13= β14 =zero

H1: at least one of them not equal zero

```
Call:
lm(formula = price ~ Make * Cylinder, data = sample_s)

Coefficients:
                 (Intercept)                   MakeCadillac                   MakeChevrolet                    MakePontiac                        MakeSAAB
                     6.00305                      -1.04430                      -0.76014                       0.39607                      -2.09110
                  MakeSaturn                    Cylinderlow                Cylindermoderate          MakeCadillac:Cylinderlow  MakeChevrolet:Cylinderlow
                     0.82335                       1.88808                       1.03619                             NA                       1.80243
      MakePontiac:Cylinderlow            MakeSAAB:Cylinderlow            MakeSaturn:Cylinderlow  MakeCadillac:Cylindermoderate  MakeChevrolet:Cylindermoderate
                     0.03152                             NA                            NA                      -0.58267                       1.08987
 MakePontiac:Cylindermoderate  MakeSAAB:Cylindermoderate  MakeSaturn:Cylindermoderate
                           NA                            NA                            NA
```

```
Coefficients: (6 not defined because of singularities)
                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                      6.00305    0.21911  27.397  < 2e-16 ***
MakeCadillac                    -1.04430    0.24207  -4.314 2.20e-05 ***
MakeChevrolet                   -0.76014    0.30351  -2.504 0.012815 *
MakePontiac                      0.39607    0.13640   2.904 0.003973 **
MakeSAAB                        -2.09110    0.54295  -3.851 0.000145 ***
MakeSaturn                       0.82335    0.52462   1.569 0.117653
Cylinderlow                      1.88808    0.56214   3.359 0.000888 ***
Cylindermoderate                 1.03619    0.19345   5.356 1.74e-07 ***
MakeCadillac:Cylinderlow              NA         NA      NA       NA
MakeChevrolet:Cylinderlow        1.80243    0.60358   2.986 0.003067 **
MakePontiac:Cylinderlow          0.03152    0.54699   0.058 0.954081
MakeSAAB:Cylinderlow                  NA         NA      NA       NA
MakeSaturn:Cylinderlow                NA         NA      NA       NA
MakeCadillac:Cylindermoderate   -0.58267    0.29295  -1.989 0.047645 *
MakeChevrolet:Cylindermoderate   1.08987    0.29588   3.683 0.000275 ***
MakePontiac:Cylindermoderate          NA         NA      NA       NA
MakeSAAB:Cylindermoderate             NA         NA      NA       NA
MakeSaturn:Cylindermoderate           NA         NA      NA       NA
---
```

Beta 8 ,14: they are significant but Beta10,13 is insignificant so, it has been found that: There is an interaction between cylinder low and Chevrolet. No interaction between cylinder low

and pontiac cars. There is no interaction between cylinder moderate and Cadillacs. There is an interaction between cylinder moderate and Chevrolet cars

Model 4: $Y= \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \beta_3 D_{1i} + \beta_4 D_{2i} + \beta_5 D_{3i} + \beta_6 D_{4i} + \beta_7 D_{5i} + u_i$

$Y= 5.624 + 0.2257 x_{1i} + 0.00002391 x_{2i} + -1.927 D_{1i} + 1.250 D_{2i} + 0.6631 D_{3i} - 1.132 D_{4i} + 1.785 D_{5i}$

H0: β1= β2= β3= β4= β5=0

H1: at least one of them not equal zero

```
Call:
lm(formula = price ~ Doors + Mileage + Make, data = sample_s)

Coefficients:
  (Intercept)          Doors        Mileage   MakeCadillac  MakeChevrolet    MakePontiac       MakeSAAB     MakeSaturn
    5.046e+00      3.756e-01      2.325e-05     -1.949e+00      1.337e+00      6.086e-01     -9.829e-01      1.740e+00

> summary(model4)

Call:
lm(formula = price ~ Doors + Mileage + Make, data = sample_s)

Residuals:
    Min      1Q  Median      3Q     Max
-2.6997 -0.3974  0.0355  0.3630  2.1719

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)    5.046e+00  3.035e-01  16.628  < 2e-16
Doors          3.756e-01  6.025e-02   6.234 1.58e-09
Mileage        2.324e-05  5.759e-06   4.037 6.93e-05
MakeCadillac  -1.949e+00  1.999e-01  -9.751  < 2e-16
MakeChevrolet  1.337e+00  1.651e-01   8.096 1.55e-14
MakePontiac    6.086e-01  1.743e-01   3.491 0.000555
MakeSAAB      -9.829e-01  1.886e-01  -5.211 3.56e-07
MakeSaturn     1.740e+00  2.198e-01   7.913 5.26e-14
```

Beta 1: when doors increase by one unit, the expected price will increase by 0.2257.

Beta 2: when mileage increase by one unit, the expected price will increase by 0.00002391.

Beta 3: the expected price of (Cadillac) is less than expected price of (buick) by 1.927, holding mileage and doors constant.

Beta 4: the expected price of (Chevrolet) is more than expected price of makebuick by 1.25, holding mileage and doors constant.

Beta 5: the expected price of (Pontiac) is more than expected price of (buick) by 0.6631, holding mileage and doors constant.

Beta 6: the expected price of (saap) is less than expected price of (buick) by 1.132, holding mileage and doors constant.

Beta 7: the expected price of (Saturn) is more than expected price of (buick) by 1.785, holding mileage and doors constant.

All estimator is significant which mean that on average, the make of the car affects the price of the car holding the number of miles driven and the number of doors constant. And according to the model, by increasing the miles driven the price of the car will increase, and by increasing the number of doors the price of the car increases

Model 5: $Y= \beta_0 + \beta_1 D_{1i} + \beta_2 D_{2i} + \beta_3 D_{3i} + \beta_4 D_{4i} + \beta_5 x_i + \beta_6 D_{1i} x_i + \beta_7 D_{1i} x_i + \beta_8 D_{1i} x_i + \beta_9 D_{1i} x_i + u_i$

$Y= 4.548 + 0.00002677 x_i + 2.973 D_{1i} + 3.567 D_{2i} + 1.980 D_{3i} + 1.886 D_{4i} - 0.00001075 D_{1i} x_i + 0.000004273 D_{1i} x_i + 0.000006381 D_{1i} x_i - 0.00001433 D_{1i} x_i$

H0: β6= β7= β8= β9=zero

H1: at least one of them not equal zero

```
Coefficients:
          (Intercept)               Mileage             TypeCoupe         TypeHatchback
            4.548e+00              2.677e-05              2.973e+00              3.567e+00
             TypeSedan             TypeWagon      Mileage:TypeCoupe  Mileage:TypeHatchback
            1.980e+00              1.886e+00             -1.075e-05              4.273e-06
    Mileage:TypeSedan      Mileage:TypeWagon
            6.381e-06             -1.433e-05

> summary(model5)

Call:
lm(formula = price ~ Mileage * Type, data = sample_s)

Residuals:
    Min      1Q  Median      3Q     Max
-2.6470 -0.9501  0.1925  0.9086  2.9977

Coefficients:
                      Estimate Std. Error t value Pr(>|t|)
(Intercept)           4.548e+00  7.985e-01   5.697    3e-08 ***
Mileage               2.677e-05  3.356e-05   0.798 0.425800
TypeCoupe             2.973e+00  8.909e-01   3.337 0.000957 ***
TypeHatchback         3.567e+00  1.042e+00   3.423 0.000709 ***
TypeSedan             1.980e+00  8.327e-01   2.378 0.018059 *
TypeWagon             1.886e+00  1.000e+00   1.885 0.060420 .
Mileage:TypeCoupe    -1.075e-05  3.817e-05  -0.282 0.778396
Mileage:TypeHatchback 4.273e-06  4.727e-05   0.090 0.928041
Mileage:TypeSedan     6.381e-06  3.525e-05   0.181 0.856496
Mileage:TypeWagon    -1.433e-05  4.286e-05  -0.334 0.738334
---
```

β6,β7, β8 and β9 are insignificant which mean that no interaction between Mileage and Type

Beta2 is significant so the expected price of (coupe) is more than expected price of Convertible type by 2.973, holding Mileage constant , Beta 3 is significant so the expected price of (Hatchback) is more than expected price of Convertible type by 3.567, holding Mileage constant , beta 4 is significant so the expected price of (Sedan) is more than expected price of Convertible type by 3.567, holding Mileage constant and Beta5 is significant so the expected price of (Wagon) is more than expected price of Convertible type by 3.567, holding Mileage constant

Model6:

$Y= \beta_0 + \beta_1 D_{1i} + \beta_2 D_{2i} + \beta_3 D_{3i} + \beta_4 D_{1i}D_{2i} + \beta_5 D_{1i}D_{3i} + \beta_6 D_{2i}D_{3i} + \beta_7 D_{1i}D_{3i}D_{2i} + u_i$

$Y= 8.92448 - 1.73224 D_{1i} + 0.09822 D_{2i} - 0.32378 D_{3i} - 0.96060 D_{1i}D_{2i}$

$- 0.21313 D_{1i}D_{3i} - 0.36172 D_{2i}D_{3i} + 1.12639 D_{1i}D_{3i}D_{2i}$

H0: β1 = β2 = β3= Zero

H1:at least one doesn't equal zero

```
Coefficients:
        (Intercept)              Cruise              Leather                Sound       Cruise:Leather         Cruise:Sound       Leather:Sound
            8.92448            -1.73224             0.09822            -0.32378            -0.96060             0.21313            -0.36172
Cruise:Leather:Sound
            1.12639

> summary(model6)

Call:
lm(formula = price ~ Cruise * Leather * Sound)

Residuals:
    Min      1Q  Median      3Q     Max
-3.2245 -0.8258 -0.1011  0.8091  4.0841

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)          8.92448    0.33784  26.416  < 2e-16 ***
Cruise              -1.73224    0.43845  -3.951 9.77e-05 ***
Leather              0.09822    0.78022   0.126    0.900
Sound               -0.32378    0.54737  -0.592    0.555
Cruise:Leather      -0.96060    0.84548  -1.136    0.257
Cruise:Sound         0.21313    0.64638   0.330    0.742
Leather:Sound       -0.36172    0.90769  -0.399    0.691
Cruise:Leather:Sound  1.12639    0.99135   1.136    0.257
```

with assume alpha equal 0.05 which mean Beta 2, 3 are insig and beta 1 is sig and no leather and no sound difference between cruise

Secondly,

$H_0: \beta_4 = \beta_5 = \beta_6 = \beta_7 = 0$

H_1: at least one of them not equal zero

Assume alpha equal 0.05. beta 4, beta 5 and beta 6 and Beta7 are insignificant so, there are no interaction between leather, cruise and sound

Model 7

Y= 6.904+0.00001591xi+ β2(xi-15000) Di

H0: β2 =zero

H1: β2 not equal zero

```
Call:
lm(formula = price ~ (Mileage < 15000) * Mileage + (Mileage >=
    15000) * Mileage)

Coefficients:
                 (Intercept)           Mileage < 15000TRUE                      Mileage
                   6.904e+00                    1.545e-01                    1.593e-05
       Mileage >= 15000TRUE   Mileage < 15000TRUE:Mileage  Mileage:Mileage >= 15000TRUE
                          NA                   -1.566e-05                           NA

> summary(model7)

Call:
lm(formula = price ~ (Mileage < 15000) * Mileage + (Mileage >=
    15000) * Mileage)

Residuals:
    Min      1Q  Median      3Q     Max
-3.2577 -1.2134  0.1614  1.0646  3.1886

Coefficients: (2 not defined because of singularities)
                             Estimate Std. Error t value Pr(>|t|)
(Intercept)                 6.904e+00  4.146e-01  16.653   <2e-16 ***
Mileage < 15000TRUE         1.545e-01  5.670e-01   0.272    0.785
Mileage                     1.593e-05  1.696e-05   0.939    0.348
Mileage >= 15000TRUE               NA         NA      NA       NA
Mileage < 15000TRUE:Mileage -1.566e-05  4.202e-05  -0.373    0.710
Mileage:Mileage >= 15000TRUE        NA         NA      NA       NA
```

Assuming that alpha equals 0.05 and that beta 2 is significant, the regression model of price and number of miles changed after 15 thousand miles.