**Self-Assessment**
**Weeks 9: Multiple Regression with Both Qualitative and Quantitative Predictors; Multiple Comparisons**

1. What is an adjusted mean? What potential benefit does it offer? How could it potentially be misleading?

> **An adjusted mean is a predicted mean using the regression model that takes into account partial contributions of both qualitative and quantitative predictors (covariates in ANCOVA language). Often researchers will use mean values of quantitative predictors (covariates) when calculating adjusted means although any value of the quantitative predictors could be used to calculate adjusted means.**
>
> **One benefit of the adjusted mean is that it provides an estimated mean for group comparisons in which groups have the same value on quantitative predictors (covariates). So adjusted means provide a way to statistically equate groups on covariates when such groups are not equivalent. This is helpful when determining whether a treatment would be effective if all groups had similar background characteristics. For example, if each of the experimental classes had similar levels of IQ (the covariate), which treatment appears to be most beneficial? Thus, the adjusted mean is a way to equalize groups on covariates that may not be similar across groups.**
>
> **Adjusted means may be misleading because the adjustments made to nonequivalent groups may not be realistic. For example, it may not be reasonable to assume all groups would have similar covariate mean scores, so covariate means used to obtain adjusted means may be unrealistic and provide misleading equivalence among nonequivalent groups. As a concrete example, comparing treatments between regular middle school students and special education middle school students while assuming both groups have mean IQ = 100 when calculating adjusted means could be unrealistic.**
>
> **In experimental research in which subjects are assigned randomly to treatment groups, such adjustments are almost always believed to be acceptable. In correlational and quasi-experimental studies that use intact groups, adjustments are more likely to be misleading so caution should always be used when working with adjusted means.**

2. Below is linked blood pressure data that was presented in previous self-assessments. Perform comparisons of systolic blood pressure among drugs (Ziac, Losartan, Lisinopril 40mg, Lisinopril 12.5mg) while controlling for weight, i.e.,

Systolic Blood Pressure = b0 + bj Drug + bi Weight

Where bj represents several drug coefficients and bi is an unnumbered coefficient for Weight.

Use the Bonferroni adjustment and set familywise alpha to .05. Present results in APA style. Also, present a table showing the predicted systolic blood pressure for someone who weighs 220, 200, and 180 pounds for each of the four treatments.

SPSS
http://www.bwgriffin.com/gsu/courses/edur8132/selfassessments/Week09/BloodPressureDrugs.sav

Excel
http://www.bwgriffin.com/gsu/courses/edur8132/selfassessments/Week09/BloodPressureDrugs.xlsx

**Results**

**Table 1: Descriptive Statistics for Systolic Blood Pressure, Weight, and Drug Treatments**

| Variable | 1 | 2 | 3 | 4 | 5 |
|---|---|---|---|---|---|
| 1. Systolic Blood Pressure | --- | | | | |
| 2. Weight | .04 | --- | | | |
| 3. Lisinopril 40mg | -.22 | .39* | --- | | |
| 4. Losartan | .08 | -.92* | -.33* | --- | |
| 5. Ziac | .07 | .15 | -.29* | -.33* | --- |
| Mean | 153.36 | 220.75 | .22 | .28 | .22 |
| SD | 11.48 | 7.91 | .42 | .45 | .42 |

***Note:*** **Lisinopril 40mg, Losartan, and Ziac are dummy variables (1, 0); n = 76.**
**$^*p<.05$**

**Table 2: Regression of Systolic Blood Pressure on Weight and Drug Treatments**

| Variable | b | se | $\Delta R^2$ | 95%CI | F | t |
|---|---|---|---|---|---|---|
| Drug | | | .197 | | 5.82* | |
| 2 = Lisinopril 40mg | -6.01 | 3.45 | | -12.89, 0.87 | | -1.74 |
| 3 = Losartan | 28.54 | 8.27 | | 12.04, 45.04 | | 3.45* |
| 4 = Ziac | 5.66 | 3.71 | | -1.74, 13.06 | | 1.52 |
| Weight | 1.63 | 0.44 | .152 | 0.75, 2.52 | | 3.68* |
| Intercept | -215.32 | 100.60 | | -415.90, -14.76 | | -2.14* |

***Note:*** **$R^2$ = .20, adj. $R^2$ = .15, $F_{4,71}$ = 4.41*, MSE = 111.58, n = 76. $\Delta R^2$ represents the squared semi-partial correlation or the increment in $R^2$ due to adding the respective variable; Lisinopril 40mg, Losartan, and Ziac are dummy variables (1, 0).**
**$^*p < .05$.**

**Table 3: Comparisons of Adjusted Mean Differences in Systolic Blood Pressure among Drug Treatments**

| Contrast | Estimated Mean Difference | Standard Error of Difference | 95% Bonferroni Corrected CI of Mean Difference |
|---|---|---|---|
| Lis. 40mg vs. Lis. 12.5mg | -6.01 | 3.45 | -15.37, 3.35 |
| Losartan vs. Lis. 12.5mg | 28.54* | 8.27 | 6.09, 51.00 |
| Ziac vs. Lis. 12.5mg | 5.66 | 3.71 | -4.42, 15.73 |
| Losartan vs. Lis. 40mg | 34.55* | 8.48 | 11.54, 57.56 |
| Ziac vs. Lis. 40mg | 11.67* | 3.93 | 1.00, 22.33 |
| Ziac vs. Losartan | -22.88* | 7.12 | -42.20, -3.57 |

**$^*p < .05$, where p-values are adjusted using the Bonferroni method.**

**Regression results show that both weight and drug treatment are associated with systolic blood pressure. The greater the weight, the higher blood pressure. The table of comparisons reveal that Losartan resulted in statistically higher blood pressure than the other three drug treatments tested even after controlling for weight. Additionally, the remaining three drug treatments appear to result in similar mean levels of systolic blood pressure after controlling for weight since there are no statistically significant differences noted in Table 3 except for the Ziac vs. Lisinopril (40mg) comparison which shows that systolic blood pressure is higher when taking Ziac than when taking Lisinopril (40mg).**

**Predicted Systolic Blood Pressure**

**The predicted mean systolic blood pressure for each drug treatment and for weights of 180, 200, and 220 are tabled below.**

**Table 4: Predicted Systolic Blood Pressure for Drug Treatments and Different Weights**

| | Drug Treatment | | | |
|---|---|---|---|---|
| **Weight** | **Lisinopril 12.5** | **Lisinopril 40** | **Losartan** | **Ziac** |
| **180** | **78.93** | **72.92** | **107.47** | **84.59** |
| **200** | **111.63** | **105.62** | **140.17** | **117.28** |
| **220** | **144.32** | **138.31** | **172.86** | **149.98** |

**Stata Commands and Results**

```
. regress meanhighsystolic i.drug_num weight
2 = Lisinopril 40
3 = Losartan
4 = Ziac
. corr meanhighsystolic weight lisinopril40 losartan ziac, means
. pwcorr meanhighsystolic weight lisinopril40 losartan ziac,  sig
. testparm i.drug_num
. di 5.83 * (1-.1989) / 71*3    " = Delta R-squared "
. testparm weight
. di 13.51 * (1-.1989) / 71*1    " = Delta R-squared "
. margins drug_num, mcompare(bonferroni) pwcompare level(95)
. margins drug_num, at( weight=(180)) asbalanced
. margins drug_num, at( weight=(200)) asbalanced
. margins drug_num, at( weight=(220)) asbalanced

. regress meanhighsystolic i.drug_num weight

      Source |       SS           df       MS      Number of obs   =        76
-------------+----------------------------------   F(4, 71)        =      4.41
       Model |  1967.38559         4  491.846397   Prob > F        =    0.0031
    Residual |  7922.02231        71  111.577779   R-squared       =    0.1989
-------------+----------------------------------   Adj R-squared   =    0.1538
       Total |  9889.40789        75  131.858772   Root MSE        =    10.563

------------------------------------------------------------------------------
meanhighsy~c |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    drug_num |
          2  |  -6.009033   3.449123    -1.74   0.086    -12.88639    .8683232
          3  |   28.54137   8.273454     3.45   0.001     12.04457    45.03817
          4  |   5.656523   3.711093     1.52   0.132    -1.743186    13.05623
             |
      weight |   1.634732   .4448129     3.68   0.000     .7478002    2.521664
       _cons |  -215.3194   100.5965    -2.14   0.036    -415.9031    -14.7357
------------------------------------------------------------------------------
```

```
. corr meanhighsystolic weight lisinopril40 losartan ziac, means
(obs=76)

    Variable |         Mean    Std. Dev.          Min          Max
-------------+----------------------------------------------------
meanhighsy~c |     153.3553     11.48298        127.5        180.5
      weight |       220.75     7.911384          205          236
lisinopril40 |     .2236842     .4194817            0            1
    losartan |     .2763158     .4501462            0            1
        ziac |     .2236842     .4194817            0            1


             | meanhi~c   weight lisin~40 losartan     ziac
-------------+---------------------------------------------
meanhighsy~c |   1.0000
      weight |   0.0395   1.0000
lisinopril40 |  -0.2146   0.3867   1.0000
    losartan |   0.0826  -0.9238  -0.3317   1.0000
        ziac |   0.0719   0.1537  -0.2881  -0.3317   1.0000

. pwcorr meanhighsystolic weight lisinopril40 losartan ziac,  sig

             | meanhi~c   weight lisin~40 losartan     ziac
-------------+---------------------------------------------
meanhighsy~c |   1.0000
             |
             |
      weight |   0.0395   1.0000
             |   0.7347
             |
lisinopril40 |  -0.2146   0.3867   1.0000
             |   0.0626   0.0006
             |
    losartan |   0.0826  -0.9238  -0.3317   1.0000
             |   0.4779   0.0000   0.0034
             |
        ziac |   0.0719   0.1537  -0.2881  -0.3317   1.0000
             |   0.5373   0.1850   0.0116   0.0034
             |

. testparm i.drug_num
 ( 1)  2.drug_num = 0
 ( 2)  3.drug_num = 0
 ( 3)  4.drug_num = 0

       F(  3,    71) =    5.83
            Prob > F =    0.0013

. di 5.83 * (1-.1989) / 71*3    " = Delta R-squared "
.19734139 = Delta R-squared

. testparm weight
 ( 1)  weight = 0

       F(  1,    71) =   13.51
            Prob > F =    0.0005

. di 13.51 * (1-.1989) / 71*1    " = Delta R-squared "
.15243466 = Delta R-squared
```

```
. margins drug_num, mcompare(bonferroni) pwcompare level(95)

Pairwise comparisons of predictive margins
Model VCE    : OLS
Expression   : Linear prediction, predict()

---------------------------
             |    Number of
             |  Comparisons
-------------+-------------
    drug_num |            6
---------------------------

--------------------------------------------------------------
             |            Delta-method        Bonferroni
             |   Contrast   Std. Err.     [95% Conf. Interval]
-------------+------------------------------------------------
    drug_num |
     2 vs 1  |  -6.009033   3.449123     -15.37061    3.352543
     3 vs 1  |   28.54137   8.273454      6.085642     50.9971
     4 vs 1  |   5.656523   3.711093     -4.416089    15.72913
     3 vs 2  |    34.5504   8.477104      11.54193    57.55888
     4 vs 2  |   11.66556   3.928091       1.00397    22.32714
     4 vs 3  |  -22.88485   7.117365     -42.20273   -3.566963
--------------------------------------------------------------

. margins drug_num, at( weight=(180)) asbalanced
Adjusted predictions                             Number of obs     =         76
Model VCE    : OLS

Expression   : Linear prediction, predict()
at           : drug_num                        (asbalanced)
               weight          =         180

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    drug_num |
          1  |   78.93236   20.63292     3.83   0.000     37.79149    120.0732
          2  |   72.92332   20.80291     3.51   0.001     31.44351    114.4031
          3  |   107.4737    13.1039     8.20   0.000     81.34529    133.6022
          4  |   84.58888   19.29777     4.38   0.000     46.11023    123.0675
------------------------------------------------------------------------------
```

```
. margins drug_num, at( weight=(200)) asbalanced
Adjusted predictions                              Number of obs     =         76
Model VCE    : OLS

Expression   : Linear prediction, predict()
at           : drug_num                      (asbalanced)
               weight          =         200

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    drug_num |
          1  |    111.627   11.83416     9.43   0.000     88.03036    135.2236
          2  |    105.618   12.02438     8.78   0.000     81.64202    129.5939
          3  |   140.1684   4.619498    30.34   0.000     130.9573    149.3794
          4  |   117.2835   10.54659    11.12   0.000     96.25422    138.3128
------------------------------------------------------------------------------

. margins drug_num, at( weight=(220)) asbalanced

Adjusted predictions                              Number of obs     =         76
Model VCE    : OLS

Expression   : Linear prediction, predict()
at           : drug_num                      (asbalanced)
               weight          =         220

------------------------------------------------------------------------------
             |            Delta-method
             |     Margin   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
    drug_num |
          1  |   144.3216   3.558659    40.56   0.000     137.2259    151.4174
          2  |   138.3126   3.833733    36.08   0.000     130.6684    145.9568
          3  |    172.863   5.408706    31.96   0.000     162.0783    183.6477
          4  |   149.9782    2.88862    51.92   0.000     144.2184    155.7379
------------------------------------------------------------------------------
```

3. Below is a data file containing the following variables for cars taken between 1970 and 1982:

| | |
|---|---|
| mpg: | miles per gallon |
| engine: | engine displacement in cubic inches |
| horse: | horsepower |
| weight: | vehicle weight in pounds |
| accel: | time to accelerate from 0 to 60 mph in seconds |
| year: | model year (70 = 1970, to 82 = 1982) |
| origin: | country of origin (1=American, 2=Europe, 3=Japan) |
| cylinder: | number of cylinders |

SPSS Data: http://www.bwgriffin.com/gsu/courses/edur8132/selfassessments/Week04/cars_missing_deleted.sav
(Note: There are underscore marks between words in the SPSS data file name.)
Other Data Format: If you prefer a data file format other than SPSS, let me know.

For this problem we wish to know whether MPG differs among car origins and number of cylinders while controlling for the weight of the car. The regression model for this study follows:

Predicted MPG = b0 + origin of car + number of cylinders + car weight

Origin of car is categorical. Number of cylinders may appear to be ratio, but since observed categories of this variable are limited, it is best to treat this variable as categorical. Note the following number of cylinders reported:

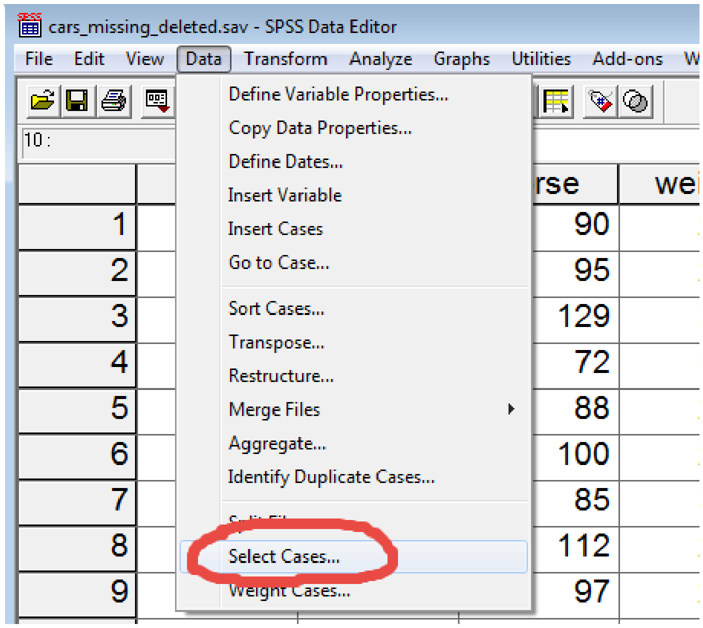
**Number of Cylinders**

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|---|---|---|---|---|---|
| Valid | 3 Cylinders | 4 | 1.0 | 1.0 | 1.0 |
| | 4 Cylinders | 199 | 50.9 | 50.9 | 51.9 |
| | 5 Cylinders | 3 | .8 | .8 | 52.7 |
| | 6 Cylinders | 83 | 21.2 | 21.2 | 73.9 |
| | 8 Cylinders | 102 | 26.1 | 26.1 | 100.0 |
| | Total | 391 | 100.0 | 100.0 | |

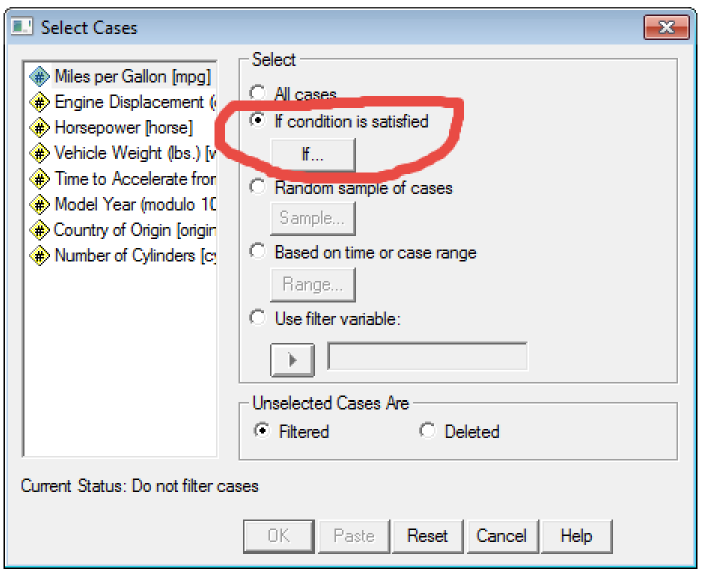
As the frequency display above shows, the number of cylinders include 3, 4, 5, 6, and 8. However, only 4 cars had 3 cylinders and only 3 cars had 5 cylinders. Given the small sample sizes for these categories, it is best to remove these cases from the regression analysis. There are several ways to accomplish this. Four approaches are (a) manually delete these cases after sorting all cases on number of cylinders, (b) telling SPSS to treat these 7 cases as missing values so they will not be included in any analysis (use Recode into Same Variable and set 3 Cylinders and 5 Cylinders as system missing), (c) defining 3 and 5 Cylinders as missing values in the variable missing values, or (d) using the Select Cases command to filter these cases from all analyses. Other possibilities also exist.

Of these four, option (d) works well and does not require deletion of any cases. This option is explained below.

Step 1: Open the Select Case window




Step 2: Choose the select If option




Step 3: Define the filter so SPSS can determine which cases NOT to select.

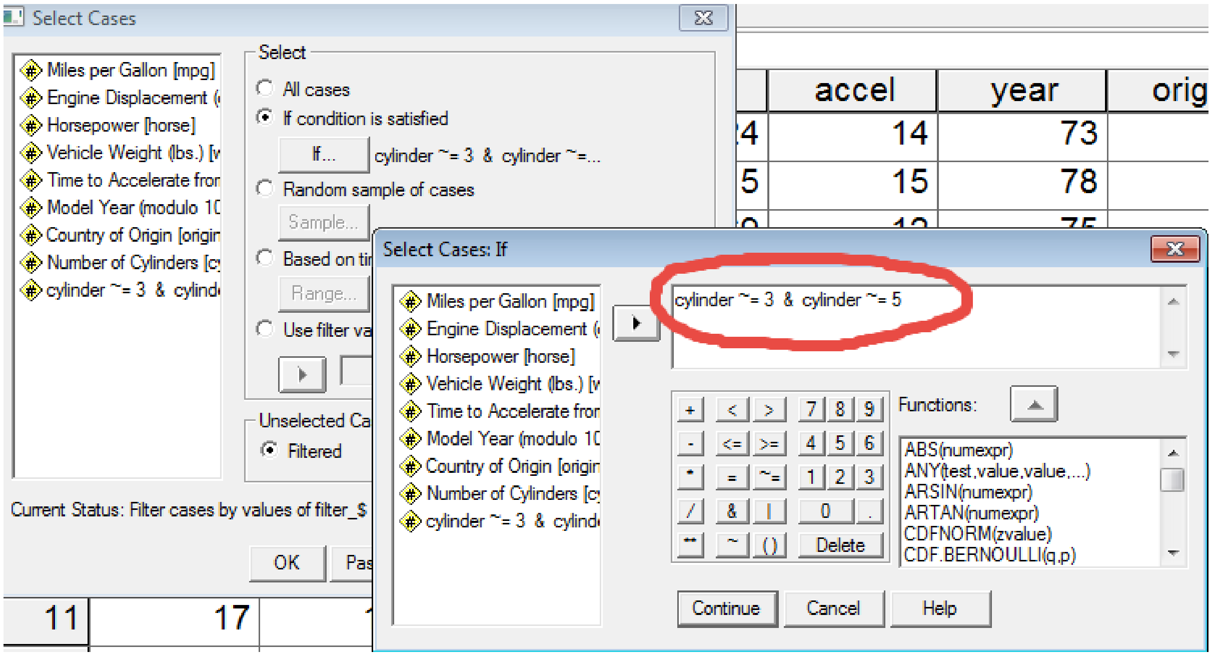
We do not want cylinders of 3 or 5, so in the Select Cases IF box, write

cylinders ~= 3

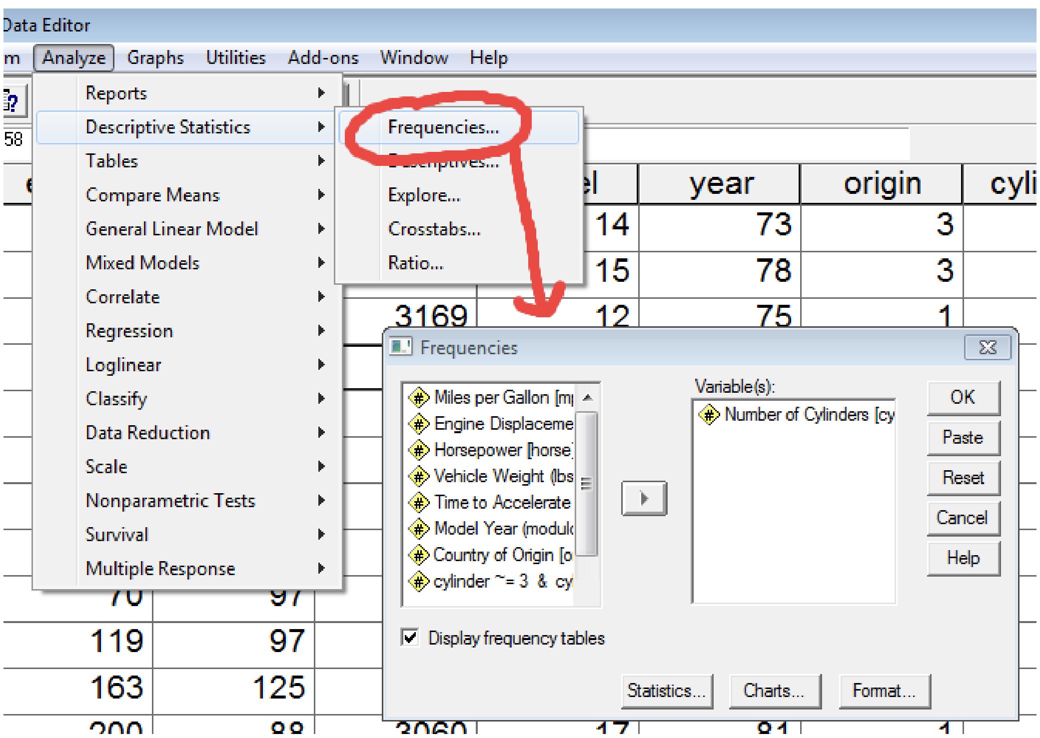
The symbol ~= means “not equal”; this tells SPSS not to select any cases in which cylinders are 3. Also, write

cylinders ~= 5

so SPSS knows not to select cases when cylinders are 5. To combine these two, we use the ampersand symbol, &, which means select all cases which are not 3 and 5 cylinders. See image below.




Once these cases are defined, click Continue then OK to process this command. Next, check that the appropriate cases were selected by running the Frequency command for cylinders as shown below.




And the results should look like this:

**Frequencies**

**Statistics**

Number of Cylinders

| | | |
|---|---|---|
| N | Valid | 384 |
| | Missing | 0 |

**Number of Cylinders**

| | | Frequency | Percent | Valid Percent | Cumulative Percent |
|---|---|---|---|---|---|
| Valid | 4 Cylinders | 199 | 51.8 | 51.8 | 51.8 |
| | 6 Cylinders | 83 | 21.6 | 21.6 | 73.4 |
| | 8 Cylinders | 102 | 26.6 | 26.6 | 100.0 |
| | Total | 384 | 100.0 | 100.0 | |

Note that no cases of cylinders equal to 3 or 5 were selected.

Present an APA styled regression analysis with DV = MPG, IV = origin, IV = Cylinders (4, 6, and 8 only), and IV = vehicle weight. Set alpha = .01. You will have to create the dummy variables for origins and cylinders. Also present Scheffé confidence intervals comparisons among origins and among cylinders.

In addition to APA styled results, present literal interpretations for each regression coefficient.

**Results**

**Table 1: Descriptive Statistics for MPG, Origin, and Number of Cylinders**

| Variable | 1 | 2 | 3 | 4 | 5 | 6 |
|---|---|---|---|---|---|---|
| 1. MPG | --- | | | | | |
| 2. Origin 2 (Europe) | .24* | --- | | | | |
| 3. Origin 3 (Japan) | .47* | -.22* | --- | | | |
| 4. Cylinder 6 | -.24* | -.17* | -.16* | --- | | |
| 5. Cylinder 8 | -.65* | -.27* | -.30* | -.32* | --- | |
| 6. Weight | -.84* | -.31* | -.44* | .14* | .80* | --- |
| Mean | 23.48 | 0.17 | 0.20 | 0.22 | 0.27 | 2978.07 |
| SD | 7.81 | 0.38 | 0.40 | 0.41 | 0.44 | 850.69 |

***Note:*** **Origin 2 and 3 are dummy variables (1, 0) and Cylinder 6 and 8 are dummy variables (1, 0); n = 384.**
***p<.01**

Table 2: Regression of MPG on Origins and Cylinders

| Variable | b | se | $\Delta R^2$ | 99%CI | F | t |
|---|---|---|---|---|---|---|
| Origin | | | .010 | | 7.42* | |
| 2 = Europe | -.214 | .662 | | -1.93, 1.50 | | -0.32 |
| 3 = Japan | 2.144 | .650 | | 0.46, 3.83 | | 3.30* |
| Number of Cylinders | | | .017 | | 12.41* | |
| 6 | -3.719 | .758 | | -5.68, -1.76 | | -4.91* |
| 8 | -3.415 | 1.142 | | -6.37, -0.46 | | -2.99* |
| Weight8 | -0.005 | .0006 | .067 | -0.007, -0.004 | | -9.84* |
| Intercept | 41.604 | 1.461 | | 37.82, 45.39 | | 28.47* |

*Note:* $R^2$ = .74, adj. $R^2$ = .73, $F_{5,378}$ = 212.97*, MSE = 16.20, n = 384. $\Delta R^2$ represents the squared semi-partial correlation or the increment in $R^2$ due to adding the respective variable.
*p < .01.

Table 3: Comparisons of Adjusted MPG among Vehicle Origins

| Contrast | Estimated Mean Difference | Standard Error of Difference | 99% Scheffé Corrected CI of Mean Difference |
|---|---|---|---|
| Europe vs USA | -0.21 | 0.66 | -2.24, 1.81 |
| Japan vs USA | 2.14* | 0.65 | 0.16, 4.13 |
| Japan vs Europe | 2.36* | 0.69 | 0.24, 4.47 |

*p < .01, where p-values are adjusted using the Scheffé method.

Table 4: Comparisons of Adjusted MPG among Number of Cylinders

| Contrast | Estimated Mean Difference | Standard Error of Difference | 99% Scheffé Corrected CI of Mean Difference |
|---|---|---|---|
| 6 vs 4 | -3.72* | 0.76 | -6.03, -1.40 |
| 8 vs 4 | -3.42 | 1.14 | -6.90, 0.07 |
| 8 vs 6 | 0.30 | 0.78 | -2.09, 2.70 |

*p < .01, where p-values are adjusted using the Scheffé method.

Results show that there are statistical differences in MPG by vehicles' origin and number of cylinders, and that vehicle weight is negatively associated with MPG. The greater the vehicle weight, the lower will be MPG. For origins, cars from Japan appear to have about a 2 MPG advantage over cars from Europe and the USA once vehicle weight and number of cylinders are taken into account, and there seems to be little to no difference in MPG between cars from Europe and the USA. For number of cylinders, cars with 4 cylinders appear to obtain about 3 MPGs more than cars with 6 and 8 cylinders once vehicle origin and weight is controlled, but only the comparison between 4 and 6 cylinders shows a statistically significant difference at the .01 level of significance. There appears to be no difference in MPG between cans with 6 and 8 cylinders after controlling for weight and origins.

## Literal Interpretations

b0 = 41.60: The predicted MPG for a car made in USA, with 4 cylinders, and with a weight of 0.

b1 = -.214: European cars expected to obtain .214 MPG less than American cars controlling for weight and cylinders.

**b2 = 2.14: Japanese cars expected to obtain 2.14 MPG more than American cars controlling for weight and cylinders.**

**b3 = -3.72: Cars with 6 cylinders expected to obtain 3.72 MPG less than cars with 4 cylinders controlling for weight and origins.**

**b4 = -3.41: Cars with 8 cylinders expected to obtain 3.41 MPG less than cars with 4 cylinders controlling for weight and origins.**

**b5 = -0.005: For every 1 additional pound added to weight of a car, the MPG is expected to decline by .005 controlling for vehicle origin and cylinders.**

**STATA Commands and Results**

```
. regress mpg i.origin  i.cylinder weight if cylinder==4 | cylinder==6 | cylinder==8, level(99)
. tabulate origin , gen(origin)
. tabulate cylinder , gen( cylinder )
. corr mpg origin2 origin3 cylinder4 cylinder5  weight if cylinder==4 | cylinder==6 | cylinder==8,
means
. pwcorr mpg origin2 origin3 cylinder4 cylinder5  weight if cylinder==4 | cylinder==6 |
cylinder==8, sig
. testparm i.origin
. di 7.42 * (1-.7380) / 378*2 " = Delta R-squared"
. testparm i.cylinder
.  di 12.41 * (1-.7380) / 378*2 " = Delta R-squared"
. testparm weight
.  di 96.76 * (1-.7380) / 378*1 " = Delta R-squared"
. margins origin , mcompare(scheffe) pwcompare level(99)
. margins cylinder , mcompare(scheffe) pwcompare level(99)

. regress mpg i.origin  i.cylinder weight if cylinder==4 | cylinder==6 | cylinder==8, level(99)

      Source |       SS           df       MS      Number of obs   =       384
-------------+----------------------------------   F(5, 378)       =    212.97
       Model |  17251.2188         5  3450.24377   Prob > F        =    0.0000
    Residual |  6123.76114       378  16.2004263   R-squared       =    0.7380
-------------+----------------------------------   Adj R-squared   =    0.7346
       Total |    23374.98       383  61.0312793   Root MSE        =     4.025

------------------------------------------------------------------------------
         mpg |      Coef.   Std. Err.      t    P>|t|     [99% Conf. Interval]
-------------+----------------------------------------------------------------
      origin |
          2  |  -.2140693   .6625454    -0.32   0.747    -1.929332    1.501193
          3  |    2.14367   .6502408     3.30   0.001     .4602624    3.827077
             |
    cylinder |
          6  |  -3.719466   .7581589    -4.91   0.000    -5.682262    -1.75667
          8  |  -3.415028   1.142405    -2.99   0.003    -6.372598    -.457459
             |
      weight |  -.0056387   .0005732    -9.84   0.000    -.0071227   -.0041547
       _cons |   41.60402   1.461099    28.47   0.000     37.82138    45.38665
------------------------------------------------------------------------------
```

```
. tabulate origin , gen(origin)
 Country of |
     Origin |      Freq.     Percent        Cum.
------------+-----------------------------------
          1 |        244       62.40       62.40
          2 |         68       17.39       79.80
          3 |         79       20.20      100.00
------------+-----------------------------------
      Total |        391      100.00

. tabulate cylinder , gen( cylinder )
  Number of |
  Cylinders |      Freq.     Percent        Cum.
------------+-----------------------------------
          3 |          4        1.02        1.02
          4 |        199       50.90       51.92
          5 |          3        0.77       52.69
          6 |         83       21.23       73.91
          8 |        102       26.09      100.00
------------+-----------------------------------
      Total |        391      100.00

. corr mpg origin2 origin3 cylinder4 cylinder5  weight if cylinder==4 | cylinder==6 | cylinder==8,
means
(obs=384)

    Variable |         Mean    Std. Dev.          Min          Max
-------------+----------------------------------------------------
         mpg |     23.48307     7.812252           10         46.6
     origin2 |     .1692708     .3754802            0            1
     origin3 |     .1953125     .3969583            0            1
   cylinder4 |     .2161458     .4121519            0            1
   cylinder5 |      .265625     .4422416            0            1
      weight |     2978.065      850.694         1613         5140


             |      mpg  origin2  origin3 cylind~4 cylind~5   weight
-------------+------------------------------------------------------
         mpg |   1.0000
     origin2 |   0.2390   1.0000
     origin3 |   0.4733  -0.2224   1.0000
   cylinder4 |  -0.2362  -0.1696  -0.1630   1.0000
   cylinder5 |  -0.6522  -0.2715  -0.2963  -0.3158   1.0000
      weight |  -0.8412  -0.3058  -0.4444   0.1385   0.8003   1.0000
```

Note: cylinder4 = 6 cylinders and cylinder5 = 8 cylinders.

```
. pwcorr mpg origin2 origin3 cylinder4 cylinder5  weight if cylinder==4 | cylinder==6 |
cylinder==8, sig

             |      mpg  origin2  origin3 cylind~4 cylind~5   weight
-------------+------------------------------------------------------
         mpg |   1.0000 
             |
             |
     origin2 |   0.2390   1.0000 
             |   0.0000
             |
     origin3 |   0.4733  -0.2224   1.0000 
             |   0.0000   0.0000
             |
   cylinder4 |  -0.2362  -0.1696  -0.1630   1.0000 
             |   0.0000   0.0009   0.0014
             |
   cylinder5 |  -0.6522  -0.2715  -0.2963  -0.3158   1.0000 
             |   0.0000   0.0000   0.0000   0.0000
             |
      weight |  -0.8412  -0.3058  -0.4444   0.1385   0.8003   1.0000 
             |   0.0000   0.0000   0.0000   0.0066   0.0000

. testparm i.origin
 ( 1)  2.origin = 0
 ( 2)  3.origin = 0
       F(  2,   378) =    7.42
            Prob > F =    0.0007
. di 7.42 * (1-.7380) / 378*2 " = Delta R-squared"
.01028593 = Delta R-squared

. testparm i.cylinder
 ( 1)  6.cylinder = 0
 ( 2)  8.cylinder = 0
       F(  2,   378) =   12.41
            Prob > F =    0.0000

.  di 12.41 * (1-.7380) / 378*2 " = Delta R-squared"
.01720328 = Delta R-squared

. testparm weight
 ( 1)  weight = 0
       F(  1,   378) =   96.76
            Prob > F =    0.0000

.  di 96.76 * (1-.7380) / 378*1 " = Delta R-squared"
.06706646 = Delta R-squared

. margins origin , mcompare(scheffe) pwcompare level(99)
Pairwise comparisons of predictive margins
Model VCE    : OLS
Expression   : Linear prediction, predict()
------------------------------------------------------------
             |            Delta-method          Scheffe
             |   Contrast   Std. Err.     [99% Conf. Interval]
-------------+----------------------------------------------
      origin |
     2 vs 1  |  -.2140693   .6625454     -2.237109     1.80897
     3 vs 1  |    2.14367   .6502408      .1582014    4.129138
     3 vs 2  |   2.357739   .6922867      .2438863    4.471592
------------------------------------------------------------
```

```
. margins cylinder , mcompare(scheffe) pwcompare level(99)
Pairwise comparisons of predictive margins
Model VCE    : OLS
Expression   : Linear prediction, predict()
--------------------------------------------------------------
             |            Delta-method         Scheffe
             |   Contrast   Std. Err.     [99% Conf. Interval]
-------------+------------------------------------------------
    cylinder |
     6 vs 4  |  -3.719466   .7581589     -6.034455   -1.404477
     8 vs 4  |  -3.415028   1.142405     -6.903287    .0732304
     8 vs 6  |   .3044378   .7841679     -2.089968    2.698844
--------------------------------------------------------------

.
```