USF UNIVERSITY OF SOUTH FLORIDA

Statistical and Econometric Methods

Assignment #1
(Continuous Data - Regression Analysis)

You are given 151 observations of a travel survey collected in State College, Pennsylvania. All of the households in the sample are making the morning commute to work. They are all departing from the same origin (a large residential complex in the suburbs) and going to work in the Central Business District. They have the choice of three alternate routes; 1) a four-lane arterial (speed limit = 35mph, 2 lanes each direction), 2) a two-lane rural road (speed limit = 35mph, 1 lane each direction) and 3) a limited access four-lane freeway (speed limit = 55mph, 2 lanes each direction).

Your task is to estimate a model of individual average travel speed to work using standard regression techniques. Your solution to this problem should include:

1. The results of your best model specification.

2. A discussion of the logical process that led you to the selection of your final specification. (e.g. Discuss the theory behind the inclusion of your selected variables). Include t-statistics and justify the sign of your variables.

Variables available for your specification are: (file *trt.out*)

Variable	Explanation
x1	Actual in-vehicle travel time in minutes
x2	Route chosen: 1 - arterial, 2 - rural road, 3 - freeway
x3	Traffic flow rate at time of departure in vehicles per hour
x4	Number of traffic signals on the selected route
x5	Distance along the selected route in tenths of miles
x6	Seat belts: 1 - if wear, 0 - if not
x7	Number of passengers in car
x8	Driver age in years: 1 - 18 to 23, 2 - 24 to 29, 3 - 30 to 39, 4 - 40 to 49, 5 - 50 and above
x9	Gender: 1 - male, 0 - female
x10	Marital status: 1 - single, 0 - married
x11	Number of children
x12	Annual income: 1 - less than 20000, 2 - 20000 to 29999, 3 - 30000 to 39999, 4 - 40000 to 49999, 5 - more than 50000
x13	Model year of car (e.g. 86 = 1986)
x14	Origin of car: 1 - domestic, 0 - foreign

```
|-> sample;1-151$
|-> read;nvar=14;nobs=151;file=U:\00Work-Purdue\new_laptop\CE697N-
disk\trt.out.txt$
Reading data file as space delimited format.
|-> create;speed=(x5/10)/(x1/60)$
|-> dstat;rhs=speed$

--------+---------------------------------------------------------------------
        |                  Standard                                      Missing
Variable|         Mean    Deviation      Minimum      Maximum    Cases  Values
--------+---------------------------------------------------------------------
   SPEED|     26.37265     8.545642     11.36842        62.25      151       0
--------+---------------------------------------------------------------------

Descriptive Statistics for   1 variables
DSTAT results are matrix LASTDSTA in current project.
|-> create;if(x2=3) frwy=1$
|-> create;if(x2=1) art=1$
|-> create;cage=86-x13$
|-> regress;lhs=speed;rhs=one,frwy,art,cage,x6,x3$

------------------------------------------------------------------------------
Ordinary     least squares regression .............
LHS=SPEED    Mean                 =       26.37265
             Standard deviation   =        8.54564
----------   No. of observations  =            151  DegFreedom   Mean square
Regression   Sum of Squares       =        2850.88            5     570.17535
Residual     Sum of Squares       =        8103.32          145      55.88498
Total        Sum of Squares       =        10954.2          150      73.02800
----------   Standard error of e  =        7.47563  Root MSE          7.32560
Fit          R-squared            =         .26025  R-bar squared      .23475
Model test   F[  5,   145]        =       10.20266  Prob F > F*        .00000
--------+---------------------------------------------------------------------
        |                 Standard            Prob.      95% Confidence
   SPEED|  Coefficient       Error       t    |t|>T*         Interval
--------+---------------------------------------------------------------------
Constant|    27.5020***     2.59796    10.59  .0000     22.4101   32.5939
    FRWY|    11.9336***     2.35348     5.07  .0000      7.3209   16.5464
     ART|    3.40147**      1.71104     1.99  .0487      .04789   6.75505
    CAGE|    -.24239         .15614    -1.55  .1228     -.54842    .06364
      X6|    2.05650        1.35486     1.52  .1312     -.59899   4.71198
      X3|    -.00714*         .00429    -1.66  .0984     -.01555    .00127
--------+---------------------------------------------------------------------
***, **, * ==>  Significance at 1%, 5%, 10% level.
Model was estimated on Aug 31, 2017 at 10:21:45 AM
------------------------------------------------------------------------------

|-> create;if(x2=3)frwytl=x4$
|-> create;if(x2=1)arttl=x4$
|-> dstat;rhs=frwytl,arttl$


--------+---------------------------------------------------------------------
        |                  Standard                                      Missing
Variable|         Mean    Deviation      Minimum      Maximum    Cases  Values
--------+---------------------------------------------------------------------
  FRWYTL|      .523179     1.652612          0.0          7.0      151       0
   ARTTL|     3.225166     6.202335          0.0         23.0      151       0
--------+---------------------------------------------------------------------

Descriptive Statistics for   2 variables
DSTAT results are matrix LASTDSTA in current project.
```

```
|-> regress;lhs=speed;rhs=one,frwy,art,cage,x6,x3,x5,frwytl,arttl$

------------------------------------------------------------------------------
Ordinary     least squares regression .....
LHS=SPEED    Mean                 =       26.37265
             Standard deviation   =        8.54564
----------   No. of observations  =            151  DegFreedom   Mean square
Regression   Sum of Squares       =        3944.25           8     493.03079
Residual     Sum of Squares       =        7009.95         142      49.36587
Total        Sum of Squares       =        10954.2         150      73.02800
----------   Standard error of e  =        7.02608  Root MSE         6.81348
Fit          R-squared            =         .36007  R-bar squared     .32401
Model test   F[  8,   142]        =        9.98728  Prob F > F*       .00000
--------+---------------------------------------------------------------------
        |                 Standard            Prob.      95% Confidence
   SPEED|  Coefficient       Error       t    |t|>T*         Interval
--------+---------------------------------------------------------------------
Constant|    16.1633***     3.99287     4.05  .0001      8.3374   23.9891
    FRWY|    16.9494**      7.63159     2.22  .0279      1.9917   31.9070
     ART|    1.98863        8.66231      .23  .8188   -14.98918  18.96644
    CAGE|    -.23282         .14913    -1.56  .1207     -.52511    .05946
      X6|    1.99289        1.27522     1.56  .1203     -.50648   4.49227
      X3|    -.00820**       .00407    -2.02  .0457     -.01617   -.00023
      X5|     .26615***      .07440     3.58  .0005      .12032    .41198
  FRWYTL|   -2.30106*       1.24576    -1.85  .0668    -4.74271    .14059
   ARTTL|     .02508         .58845      .04  .9661    -1.12826   1.17842
--------+---------------------------------------------------------------------
***, **, * ==>  Significance at 1%, 5%, 10% level.
Model was estimated on Aug 31, 2017 at 10:21:45 AM
------------------------------------------------------------------------------

|-> create;if(x2=2)rurtl=x4$
|-> create;if(x8<3&x9=1)youngm=1$
|-> regress;lhs=speed;rhs=one,frwy,art,cage,x6,x3,x5,frwytl,rurtl,youngm$

------------------------------------------------------------------------------
Ordinary     least squares regression .....
LHS=SPEED    Mean                 =       26.37265
             Standard deviation   =        8.54564
----------   No. of observations  =            151  DegFreedom   Mean square
Regression   Sum of Squares       =        4227.02           9     469.66850
Residual     Sum of Squares       =        6727.18         141      47.71052
Total        Sum of Squares       =        10954.2         150      73.02800
----------   Standard error of e  =        6.90728  Root MSE         6.67464
Fit          R-squared            =         .38588  R-bar squared     .34668
Model test   F[  9,   141]        =        9.84413  Prob F > F*       .00000
--------+---------------------------------------------------------------------
        |                 Standard            Prob.      95% Confidence
   SPEED|  Coefficient       Error       t    |t|>T*         Interval
--------+---------------------------------------------------------------------
Constant|    22.9006***     5.11966     4.47  .0000     12.8662   32.9349
    FRWY|    -.00463       10.37452      .00  .9996   -20.33832  20.32907
     ART|   -9.05273*       5.27328    -1.72  .0882   -19.38816   1.28270
    CAGE|    -.21880         .14657    -1.49  .1377     -.50607    .06848
      X6|    1.54567        1.27031     1.22  .2257     -.94410   4.03543
      X3|    -.00946**       .00401    -2.36  .0197     -.01732   -.00160
      X5|     .37580***      .08444     4.45  .0000      .21031    .54129
  FRWYTL|   -1.82157        1.23742    -1.47  .1432    -4.24686    .60373
   RURTL|   -1.46879**       .64959    -2.26  .0253    -2.74197   -.19561
  YOUNGM|    1.30948        1.28648     1.02  .3105    -1.21197   3.83094
--------+---------------------------------------------------------------------
***, **, * ==>  Significance at 1%, 5%, 10% level.
Model was estimated on Aug 31, 2017 at 10:21:45 AM
------------------------------------------------------------------------------
```

```
|-> create;if(x9=1)maleage=x8$
|-> histogram;rhs=maleage$
```