PURDUE
UNIVERSITY

School of Civil Engineering

CE 614 – 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=D:\old_drive_d\new_laptop\CE697N-disk\trt.out....
--> create;speed=(x5/10)/(x1/60)$
--> dstat;rhs=speed$
```

Descriptive Statistics
All results based on nonmissing observations.

Variable	Mean	Std.Dev.	Minimum	Maximum	Cases
SPEED	.263726534D+02	.854564197D+01	.113684211D+02	.622500000D+02	151

```
--> 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 Weighting variable = none
Dep. var. = SPEED Mean= 26.37265340 , S.D.= 8.545641969
Model size: Observations = 151, Parameters = 6, Deg.Fr.= 145
Residuals: Sum of squares= .8103322769D+04, Std.Dev.= 7.47563
Fit: R-squared= .260254, Adjusted R-squared = .23475
Model test: F[5, 145] = 10.20, Prob value = .00000
Diagnostic: Log-L = -514.9573, Restricted(b=0) Log-L = -537.7167
LogAmemiyaPrCrt.= 4.062, Akaike Info. Crt.= 6.900
Autocorrel: Durbin-Watson Statistic = 1.93247, Rho = .03376

Variable	Coefficient	Standard Error	t-ratio	P[|T|>t]	Mean of X
Constant	27.50201489	2.5979595	10.586	.0000	
FRWY	11.93363775	2.3534831	5.071	.0000	.99337748E-01
ART	3.401468353	1.7110414	1.988	.0487	.21854305
CAGE	-.2423910189	.15614209	-1.552	.1228	4.0927152
X6	2.056499369	1.3548641	1.518	.1312	.70860927
X3	-.7138526638E-02	.42918384E-02	-1.663	.0984	493.57616

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

Descriptive Statistics
All results based on nonmissing observations.

Variable	Mean	Std.Dev.	Minimum	Maximum	Cases
FRWYTL	.523178808D+00	.165261182D+01	.000000000D+00	.700000000D+01	151
ARTTL	.322516556D+01	.620233524D+01	.000000000D+00	.230000000D+02	151

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

Ordinary least squares regression Weighting variable = none
Dep. var. = SPEED Mean= 26.37265340 , S.D.= 8.545641969
Model size: Observations = 151, Parameters = 9, Deg.Fr.= 142
Residuals: Sum of squares= .7009953198D+04, Std.Dev.= 7.02608
Fit: R-squared= .360067, Adjusted R-squared = .32401
Model test: F[8, 142] = 9.99, Prob value = .00000
Diagnostic: Log-L = -504.0141, Restricted(b=0) Log-L = -537.7167
LogAmemiyaPrCrt.= 3.957, Akaike Info. Crt.= 6.795
Autocorrel: Durbin-Watson Statistic = 1.91592, Rho = .04204

Variable	Coefficient	Standard Error	t-ratio	P[|T|>t]	Mean of X
Constant	16.16325632	3.9928669	4.048	.0001	
FRWY	16.94939297	7.6315911	2.221	.0279	.99337748E-01
ART	1.988634670	8.6623071	.230	.8188	.21854305
CAGE	-.2328219473	.14912800	-1.561	.1207	4.0927152
X6	1.992892731	1.2752154	1.563	.1203	.70860927
X3	-.8198678320E-02	.40677382E-02	-2.016	.0457	493.57616
X5	.2661454594	.74404251E-01	3.577	.0005	48.099338
FRWYTL	-2.301061742	1.2457614	-1.847	.0668	.52317881
ARTTL	.2508103259E-01	.58844954	.043	.9661	3.2251656

```
--> 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 Weighting variable = none
Dep. var. = SPEED Mean= 26.37265340 , S.D.= 8.545641969
Model size: Observations = 151, Parameters = 10, Deg.Fr.= 141
Residuals: Sum of squares= .6727183008D+04, Std.Dev.= 6.90728
Fit: R-squared= .385881, Adjusted R-squared = .34668
Model test: F[9, 141] = 9.84, Prob value = .00000
Diagnostic: Log-L = -500.9054, Restricted(b=0) Log-L = -537.7167
LogAmemiyaPrCrt.= 3.929, Akaike Info. Crt.= 6.767
Autocorrel: Durbin-Watson Statistic = 1.91023, Rho = .04489

Variable	Coefficient	Standard Error	t-ratio	P[|T|>t]	Mean of X
Constant	22.90059405	5.1196606	4.473	.0000	
FRWY	-.4626342414E-02	10.374523	.000	.9996	.99337748E-01
ART	-9.052728265	5.2732767	-1.717	.0882	.21854305
CAGE	-.2187957231	.14657167	-1.493	.1377	4.0927152
X6	1.545666270	1.2703105	1.217	.2257	.70860927
X3	-.9457421614E-02	.40102172E-02	-2.358	.0197	493.57616
X5	.3757958689	.84435209E-01	4.451	.0000	48.099338
FRWYTL	-1.821566472	1.2374165	-1.472	.1432	.52317881
RURTL	-1.468789815	.64959125	-2.261	.0253	5.1523179
YOUNGM	1.309483785	1.2864788	1.018	.3105	.27814570

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