Midterm Project

Jeremy Harper
10.12.11
STA4203

A.

DEPVAR Intercept Crime Zoned Industry Chas NOX

HouseVal = 34.2606 -0.14746 +0.048718 +0.094208 +3.42929 -16.8750

Rooms Age Dist Hwy Tax PTeacher Bk Lstat

+4.06319 +0.013773 -1.29310 +0.33113 -0.013617 -0.97817 +.007752692 -0.60901

$R^2 = 0.7607$

B. Yes, there appears to be mild collinearity. The VIF values for tax, hwy, nox, and industry are all above 4, and the condition number is well above 30 (86.84610). Bivariate Pearson correlations reveal that all predictors are correlated (aside from chas), with many explaining above 60% of the variance (see correlation matrix).

Variable	DF	Parameter Estimate	Standard Error	t Value	Pr > |t|	Variance Inflation
Intercept	1	34.26062	6.91368	4.96	<.0001	0
Crime	1	-0.14746	0.06032	-2.44	0.0152	2.10546
Zoned	1	0.04872	0.01923	2.53	0.0119	2.38786
Industry	1	0.09421	0.08678	1.09	0.2787	4.18344
Chas	1	3.42929	1.33571	2.57	0.0109	1.04929
NOX	1	-16.87502	5.20582	-3.24	0.0014	4.42518
Rooms	1	4.06319	0.55655	7.30	<.0001	1.73003
Age	1	0.01377	0.01796	0.77	0.4439	3.13452
Dist	1	-1.29310	0.27096	-4.77	<.0001	3.94670
Hwy	1	0.33113	0.09564	3.46	0.0006	8.32925
Tax	1	-0.01362	0.00547	-2.49	0.0135	10.23342
PTeacher	1	-0.97817	0.18102	-5.40	<.0001	1.83761
Bk	1	0.00775	0.00385	2.01	0.0454	1.39423
Lstat	1	-0.60901	0.07300	-8.34	<.0001	3.00002

Number	Eigenvalue	Index
1	10.14170	1.00000
2	1.61541	2.50561
3	0.95181	3.26423
4	0.65642	3.93064
5	0.20611	7.01464
6	0.15888	7.98947
7	0.09903	10.11997
8	0.07469	11.65241
9	0.03852	16.22515
10	0.02536	19.99625
11	0.01246	28.52531
12	0.01085	30.57910
13	0.00740	37.01260
14	0.00134	86.84610

The CORR Procedure

Pearson Correlation Coefficients, N = 253
Prob > |r| under H0: Rho=0

	Crime	Zoned	Industry	Chas	NOX	Rooms	Age	Dist	Hwy
Crime	1.00000	-0.22165 0.0004	0.46091 <.0001	-0.05675 0.3687	0.47336 <.0001	-0.20911 0.0008	0.38974 <.0001	-0.41981 <.0001	0.69025 <.0001
Zoned	-0.22165 0.0004	1.00000	-0.53974 <.0001	0.01926 0.7605	-0.50724 <.0001	0.32412 <.0001	-0.56026 <.0001	0.66202 <.0001	-0.30102 <.0001
Industry	0.46091 <.0001	-0.53974 <.0001	1.00000	0.03094 0.6243	0.77092 <.0001	-0.37035 <.0001	0.64076 <.0001	-0.70029 <.0001	0.60156 <.0001
Chas	-0.05675 0.3687	0.01926 0.7605	0.03094 0.6243	1.00000	0.03801 0.5472	0.08176 0.1949	0.06362 0.3135	-0.05745 0.3628	-0.02143 0.7344
NOX	0.47336 <.0001	-0.50724 <.0001	0.77092 <.0001	0.03801 0.5472	1.00000	-0.27805 <.0001	0.72778 <.0001	-0.76670 <.0001	0.61455 <.0001
Rooms	-0.20911 0.0008	0.32412 <.0001	-0.37035 <.0001	0.08176 0.1949	-0.27805 <.0001	1.00000	-0.23174 0.0002	0.19502 0.0018	-0.18371 0.0034
Age	0.38974 <.0001	-0.56026 <.0001	0.64076 <.0001	0.06362 0.3135	0.72778 <.0001	-0.23174 0.0002	1.00000	-0.74613 <.0001	0.45267 <.0001
Dist	-0.41981 <.0001	0.66202 <.0001	-0.70029 <.0001	-0.05745 0.3628	-0.76670 <.0001	0.19502 0.0018	-0.74613 <.0001	1.00000	-0.49076 <.0001
Hwy	0.69025 <.0001	-0.30102 <.0001	0.60156 <.0001	-0.02143 0.7344	0.61455 <.0001	-0.18371 0.0034	0.45267 <.0001	-0.49076 <.0001	1.00000
Tax	0.64972 <.0001	-0.32126 <.0001	0.73424 <.0001	-0.04743 0.4526	0.67858 <.0001	-0.25908 <.0001	0.51491 <.0001	-0.54813 <.0001	0.91726 <.0001
PTeacher	0.31725 <.0001	-0.41241 <.0001	0.38771 <.0001	-0.08744 0.1656	0.20111 0.0013	-0.36279 <.0001	0.27854 <.0001	-0.23610 0.0002	0.46408 <.0001
Bk	-0.41599 <.0001	0.16763 0.0075	-0.36674 <.0001	0.01515 0.8105	-0.39837 <.0001	0.08034 0.2028	-0.28409 <.0001	0.29288 <.0001	-0.44700 <.0001
Lstat	0.50187 <.0001	-0.42418 <.0001	0.64214 <.0001	-0.01305 0.8364	0.62383 <.0001	-0.54474 <.0001	0.63861 <.0001	-0.52543 <.0001	0.52147 <.0001

Pearson Correlation Coefficients, N = 253
Prob > |r| under H0: Rho=0

	Tax	PTeacher	Bk	Lstat
Crime	0.64972 <.0001	0.31725 <.0001	-0.41599 <.0001	0.50187 <.0001
Zoned	-0.32126 <.0001	-0.41241 <.0001	0.16763 0.0075	-0.42418 <.0001
Industry	0.73424 <.0001	0.38771 <.0001	-0.36674 <.0001	0.64214 <.0001
Chas	-0.04743 0.4526	-0.08744 0.1656	0.01515 0.8105	-0.01305 0.8364
NOX	0.67858 <.0001	0.20111 0.0013	-0.39837 <.0001	0.62383 <.0001
Rooms	-0.25908 <.0001	-0.36279 <.0001	0.08034 0.2028	-0.54474 <.0001
Age	0.51491 <.0001	0.27854 <.0001	-0.28409 <.0001	0.63861 <.0001
Dist	-0.54813 <.0001	-0.23610 0.0002	0.29288 <.0001	-0.52543 <.0001
Hwy	0.91726 <.0001	0.46408 <.0001	-0.44700 <.0001	0.52147 <.0001
Tax	1.00000	0.46602 <.0001	-0.44609 <.0001	0.58599 <.0001
PTeacher	0.46602 <.0001	1.00000	-0.14653 0.0197	0.39179 <.0001
Bk	-0.44609 <.0001	-0.14653 0.0197	1.00000	-0.39993 <.0001
Lstat	0.58599 <.0001	0.39179 <.0001	-0.39993 <.0001	1.00000

C. Models from housing_1.

#	Predictors	R^2
1	Lstat	0.5489
2	Rooms Lstat	0.6625
3	Rooms Pteacher Lstat	0.7023
4	Rooms Dist Pteacher Lstat	0.7152
5	NOX Rooms Dist Pteacher Lstat	0.7273
6	Chas Nox Rooms Dist Pteacher Lstat	0.7377
7	Chas Nox Rooms Dist Pteacher Bk Lstat	0.7421
8	Zoned Chas NOX Rooms Dist PTeacher Bk Lstat	0.7462
9	Crime Chas NOX Rooms Dist Hwy PTeacher Bk Lstat	0.7502
10	Crime Zoned Chas NOX Rooms Dist Hwy Tax PTeacher Lstat	0.7548
11	Crime Zoned Chas NOX Rooms Dist Hwy Tax PTeacher Bk Lstat	0.7590

#	Predictors	R^2
12	Crime Zoned Industry Chas NOX Rooms Dist Hwy Tax PTeacher Bk Lstat	0.7601
13	Crime Zoned Industry Chas NOX Rooms Age Dist Hwy Tax PTeacher Bk Lstat	0.7607

D. Models from housing_1.

#	RSS_1	RSS_2
1	9445.72608	10070.26
2	7066.62094	8524.19
3	6234.23001	7593.27
4	5963.21575	7394.85
5	5711.46560	6884.79
6	5492.64295	6838.22
7	5401.31192	6666.37
8	5314.37287	6549.29
9	5231.07524	6555.37
10	5134.92375	6456.40
11	5047.56326	6274.89
12	5023.75256	6326.70
13	5011.41952	6370.57

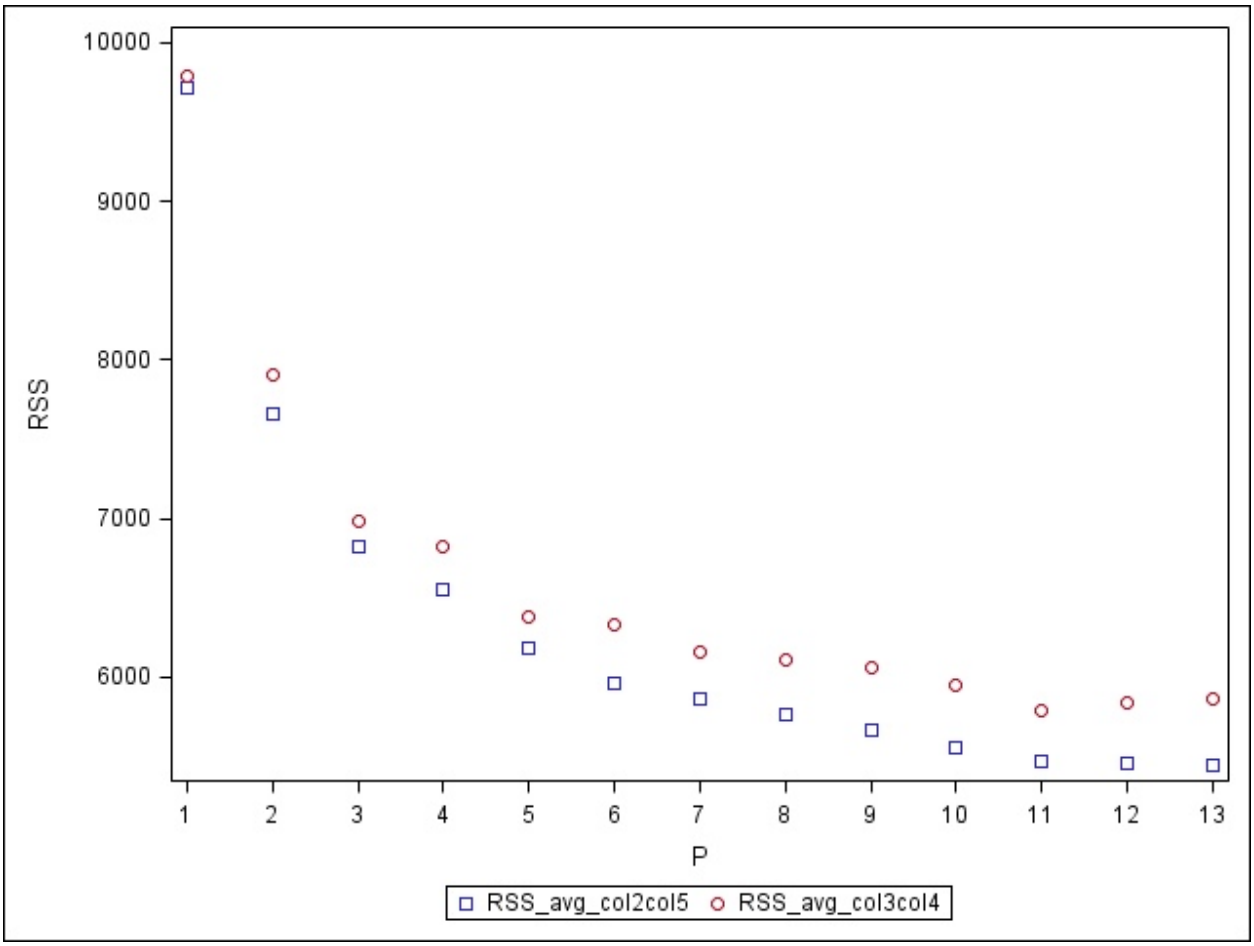
E. Models from housing_2.

#	Predictors	R^2
1	Lstat	0.5411
2	Rooms Lstat	0.6209
3	Rooms Pteacher Lstat	0.6594
4	Rooms Pteacher Bk Lstat	0.6717
5	NOX Rooms Dist PTeacher Lstat	0.6941
6	NOX Rooms Dist PTeacher Bk Lstat	0.7044
7	Chas NOX Rooms Dist PTeacher Bk Lstat	0.7095
8	Crime NOX Rooms Dist Hwy PTeacher Bk Lstat	0.7141
9	Crime NOX Rooms Dist Hwy Tax PTeacher Bk Lstat	0.7200
10	Crime Zoned NOX Rooms Dist Hwy Tax PTeacher Bk Lstat	0.7257
11	Crime Zoned Chas NOX Rooms Dist Hwy Tax PTeacher Bk Lstat	0.7298
12	Crime Zoned Industry Chas NOX Rooms Dist Hwy Tax PTeacher Bk Lstat	0.7301
13	Crime Zoned Industry Chas NOX Rooms Age Dist Hwy Tax PTeacher Bk Lstat	0.7303

	Model from housing_1		Model from housing_2	
#	RSS_1	RSS_2	RSS_1	RSS_2
1	9445.72608	10070.26	9512.41	9990.30274
2	7066.62094	8524.19	7280.18	8253.33019
3	6234.23001	7593.27	6376.40	7414.91196
4	5963.21575	7394.85	6262.20	7148.08551
5	5711.46560	6884.79	5885.24	6660.70179
6	5492.64295	6838.22	5813.59	6436.63499
7	5401.31192	6666.37	5649.39	6325.74425
8	5314.37287	6549.29	5674.18	6224.29757

	Model from	housing_1	Model from	housing_2
#	RSS_1	RSS_2	RSS_1	RSS_2
9	5231.07524	6555.37	5573.40	6094.99556
10	5134.92375	6456.40	5437.34	5970.87469
11	5047.56326	6274.89	5306.05	5882.44873
12	5023.75256	6326.70	5337.76	5875.83593
13	5011.41952	6370.57	5351.02	5870.88399

F. RSS_avg_col2col5 = (column2 + column5)/2
RSS_avg_col3col4 = (column3 + column4)/2



G.

model from	housing_a	model from	housing_b
RSS_a	RSS_b	RSS_a	RSS_b
7618.93703	6135.39	3047.49194	12931.13

CODE

```
PROC IMPORT OUT= WORK.housing_1
        DATAFILE= "Z:\Documents\University\Fall_2011\STA4203\midterm
\housing_1.csv" 877
        DBMS=CSV REPLACE;
    GETNAMES=YES;
    DATAROW=2;
RUN;

PROC IMPORT OUT= WORK.housing_2
        DATAFILE= "Z:\Documents\University\Fall_2011\STA4203\midterm
\housing_2.csv"
        DBMS=CSV REPLACE;
    GETNAMES=YES;
    DATAROW=2;
RUN;

/* a */

proc reg data = housing_1 outest = problema;
            model houseval = crime zoned industry chas NOX rooms age dist hwy tax pteacher bk lstat;
run; quit;

proc print data = problema;
run;

/* b */

libname mylib 'Z:Documents\University\Fall_2011\STA4203\midterm\';

ods pdf file='b_corr.pdf' style=journal;

proc corr data = housing_1;
            var crime zoned industry chas NOX rooms age dist hwy tax pteacher bk lstat;
run;

ods pdf close;

proc reg data = housing_1;
            model houseval = crime zoned industry chas NOX rooms age dist hwy tax pteacher bk lstat / collin vif;
run; quit;

/* c */

proc reg data = housing_1;
            model houseval = crime zoned industry chas NOX rooms age dist hwy tax pteacher bk lstat / selection=rsquare;
run; quit;

/* d */

            /* 1 */

proc reg data = housing_1 outest=model1;
            model houseval = lstat;
run; quit;

proc score data=housing_2 score=model1 out=score1 residual type=parms;
```

```
	var lstat houseval;
run;

data oneresid;
	set score1;
	squareresid = (model1)**2;
run;

proc print data = oneresid;
	var model1 squareresid;
	sum model1 squareresid;
run;

	/* 2 */

proc reg data = housing_1 outest=model2;
	model houseval = Rooms Lstat;
run; quit;

proc score data=housing_2 score=model2 out=score2 residual type=parms;
	var rooms lstat houseval;
run;

data tworesid;
	set score2;
	squareresid = (model1)**2;
run;

proc print data = tworesid;
	var model1 squareresid;
	sum model1 squareresid;
run;

	/* 3 */

proc reg data = housing_1  outest=model3;
	model houseval = Rooms Pteacher Lstat;
run; quit;

proc score data=housing_2 score=model3 out=score3 residual type=parms;
	var rooms pteacher lstat houseval;
run;

data threeresid;
	set score3;
	squareresid = (model1)**2;
run;

proc print data = threeresid;
	var model1 squareresid;
	sum model1 squareresid;
run;

	/* 4 */

proc reg data = housing_1 outest = model4;
	model houseval = Rooms Dist Pteacher Lstat;
run; quit;

proc score data=housing_2 score=model4 out=score4 residual type=parms;
	var rooms dist pteacher lstat houseval;
run;

data fourresid;
	set score4;
	squareresid = (model1)**2;
run;

proc print data = fourresid;
	var model1 squareresid;
	sum model1 squareresid;
run;
```

```
		/* 5 */

proc reg data = housing_1 outest = model5;
		model houseval = NOX Rooms Dist Pteacher Lstat;
run; quit;

proc score data=housing_2 score=model5 out=score5 residual type=parms;
		var NOX Rooms Dist Pteacher Lstat houseval;
run;

data fiveresid;
		set score5;
		squareresid = (model1)**2;
run;

proc print data = fiveresid;
		var model1 squareresid;
		sum model1 squareresid;
run;

		/* 6 */

proc reg data = housing_1 outest=model6;
		model houseval = Chas Nox Rooms Dist Pteacher Lstat;
run; quit;

proc score data=housing_2 score=model6 out=score6 residual type=parms;
		var Chas Nox Rooms Dist Pteacher Lstat houseval;
run;

data sixresid;
		set score6;
		squareresid = (model1)**2;
run;

proc print data = sixresid;
		var model1 squareresid;
		sum model1 squareresid;
run;

		/* 7 */

proc reg data = housing_1 outest=model7;
		model houseval = Chas Nox Rooms Dist Pteacher Bk Lstat;
run; quit;

proc score data=housing_2 score=model7 out=score7 residual type=parms;
		var Chas Nox Rooms Dist Pteacher Bk Lstat houseval;
run;

data sevenresid;
		set score7;
		squareresid = (model1)**2;
run;

proc print data = sevenresid;
		var model1 squareresid;
		sum model1 squareresid;
run;

		/* 8 */

proc reg data = housing_1 outest=model8;
		model houseval = Zoned Chas NOX Rooms Dist PTeacher Bk Lstat;
run; quit;

proc score data=housing_2 score=model8 out=score8 residual type=parms;
		var Zoned Chas NOX Rooms Dist PTeacher Bk Lstat houseval;
run;

data eightresid;
```

```
	set score8;
	squareresid = (model1)**2;
run;

proc print data = eightresid;
	var model1 squareresid;
	sum model1 squareresid;
run;

	/* 9 */

proc reg data = housing_1 outest = model9;
	model houseval = Crime Chas NOX Rooms Dist Hwy PTeacher Bk Lstat;
run; quit;

proc score data=housing_2 score=model9 out=score9 residual type=parms;
	var Crime Chas NOX Rooms Dist Hwy PTeacher Bk Lstat houseval;
run;

data nineresid;
	set score9;
	squareresid = (model1)**2;
run;

proc print data = nineresid;
	var model1 squareresid;
	sum model1 squareresid;
run;

	/* 10 */

proc reg data = housing_1 outest = model10;
	model houseval = Crime Zoned Chas NOX Rooms Dist Hwy Tax PTeacher Lstat;
run; quit;

proc score data=housing_2 score=model10 out=score10 residual type=parms;
	var Crime Zoned Chas NOX Rooms Dist Hwy Tax PTeacher Lstat houseval;
run;

data tenresid;
	set score10;
	squareresid = (model1)**2;
run;

proc print data = tenresid;
	var model1 squareresid;
	sum model1 squareresid;
run;

	/* 11 */

proc reg data = housing_1 outest = model11;
	model houseval = Crime Zoned Chas NOX Rooms Dist Hwy Tax PTeacher Bk Lstat;
run; quit;

proc score data=housing_2 score=model11 out=score11 residual type=parms;
	var Crime Zoned Chas NOX Rooms Dist Hwy Tax PTeacher Bk Lstat houseval;
run;

data elevenresid;
	set score11;
	squareresid = (model1)**2;
run;

proc print data = elevenresid;
	var model1 squareresid;
	sum model1 squareresid;
run;

	/* 12 */

proc reg data = housing_1 outest = model12;
```

```
	model houseval = Crime Zoned Industry Chas NOX Rooms Dist Hwy Tax PTeacher Bk Lstat;
run; quit;

proc score data=housing_2 score=model12 out=score12 residual type=parms;
	var Crime Zoned Industry Chas NOX Rooms Dist Hwy Tax PTeacher Bk Lstat houseval;
run;

data twelveresid;
	set score12;
	squareresid = (model1)**2;
run;

proc print data = twelveresid;
	var model1 squareresid;
	sum model1 squareresid;
run;

	/* 13 */

proc reg data = housing_1 outest = model13;
	model houseval = Crime Zoned Industry Chas NOX Rooms Age Dist Hwy Tax PTeacher Bk Lstat;
run; quit;

proc score data=housing_2 score=model13 out=score13 residual type=parms;
	var Crime Zoned Industry Chas NOX Rooms Age Dist Hwy Tax PTeacher Bk Lstat houseval;
run;

data thirteenresid;
	set score13;
	squareresid = (model1)**2;
run;

proc print data = thirteenresid;
	var model1 squareresid;
	sum model1 squareresid;
run;

	/* e */

proc reg data = housing_2;
	model houseval = crime zoned industry chas NOX rooms age dist hwy tax pteacher bk lstat / selection=rsquare;
run; quit;


	/* 1 */

proc reg data = housing_2 outest = model1;
	model houseval = lstat;
run; quit;

proc score data=housing_1 score=model1 out=score1 residual type=parms;
	var Lstat houseval;
run;

data oneresid;
	set score1;
	squareresid = (model1)**2;
run;

proc print data = oneresid;
	var model1 squareresid;
	sum model1 squareresid;
run;


	/* 2 */

proc reg data = housing_2 outest = model2;
	model houseval = Rooms Lstat;
run; quit;

proc score data=housing_1 score=model2 out=score2 residual type=parms;
```

```
	var Rooms Lstat houseval;
run;

data tworesid;
	set score2;
	squareresid = (model1)**2;
run;

proc print data = tworesid;
	var model1 squareresid;
	sum model1 squareresid;
run;

	/* 3 */

proc reg data = housing_2 outest = model3;
	model houseval = Rooms Pteacher Lstat;
run; quit;

proc score data=housing_1 score=model3 out=score3 residual type=parms;
	var Rooms Pteacher Lstat houseval;
run;

data threeresid;
	set score3;
	squareresid = (model1)**2;
run;

proc print data = threeresid;
	var model1 squareresid;
	sum model1 squareresid;
run;

	/* 4 */

proc reg data = housing_2 outest = model4;
	model houseval = Rooms Pteacher Bk Lstat;
run; quit;

proc score data=housing_1 score=model4 out=score4 residual type=parms;
	var Rooms Pteacher Bk Lstat houseval;
run;

data fourresid;
	set score4;
	squareresid = (model1)**2;
run;

proc print data = fourresid;
	var model1 squareresid;
	sum model1 squareresid;
run;

	/* 5 */

proc reg data = housing_2 outest = model5;
	model houseval = NOX Rooms Dist PTeacher Lstat;
run; quit;

proc score data=housing_1 score=model5 out=score5 residual type=parms;
	var NOX Rooms Dist PTeacher Lstat houseval;
run;

data fiveresid;
	set score5;
	squareresid = (model1)**2;
run;

proc print data = fiveresid;
	var model1 squareresid;
	sum model1 squareresid;
run;
```

```
		/* 6 */

proc reg data = housing_2 outest = model6;
		model houseval = NOX Rooms Dist PTeacher Bk Lstat;
run; quit;

proc score data=housing_1 score=model6 out=score6 residual type=parms;
		var NOX Rooms Dist PTeacher Bk Lstat houseval;
run;

data sixresid;
		set score6;
		squareresid = (model1)**2;
run;

proc print data = sixresid;
		var model1 squareresid;
		sum model1 squareresid;
run;

		/* 7 */

proc reg data = housing_2 outest = model7;
		model houseval = Chas NOX Rooms Dist PTeacher Bk Lstat;
run; quit;

proc score data=housing_1 score=model7 out=score7 residual type=parms;
		var Chas NOX Rooms Dist PTeacher Bk Lstat houseval;
run;

data sevenresid;
		set score7;
		squareresid = (model1)**2;
run;

proc print data = sevenresid;
		var model1 squareresid;
		sum model1 squareresid;
run;

		/* 8 */

proc reg data = housing_2 outest = model8;
		model houseval = Crime NOX Rooms Dist Hwy PTeacher Bk Lstat;
run; quit;

proc score data=housing_1 score=model8 out=score8 residual type=parms;
		var Crime NOX Rooms Dist Hwy PTeacher Bk Lstat houseval;
run;

data eightresid;
		set score8;
		squareresid = (model1)**2;
run;

proc print data = eightresid;
		var model1 squareresid;
		sum model1 squareresid;
run;

		/* 9 */

proc reg data = housing_2 outest = model9;
		model houseval = Crime NOX Rooms Dist Hwy Tax PTeacher Bk Lstat;
run; quit;

proc score data=housing_1 score=model9 out=score9 residual type=parms;
		var Crime NOX Rooms Dist Hwy Tax PTeacher Bk Lstat houseval;
run;
```

```
data nineresid;
	set score9;
	squareresid = (model1)**2;
run;

proc print data = nineresid;
	var model1 squareresid;
	sum model1 squareresid;
run;

	/* 10 */

proc reg data = housing_2 outest = model10;
	model houseval = Crime Zoned NOX Rooms Dist Hwy Tax PTeacher Bk Lstat;
run; quit;

proc score data=housing_1 score=model10 out=score10 residual type=parms;
	var Crime Zoned NOX Rooms Dist Hwy Tax PTeacher Bk Lstat houseval;
run;

data tenresid;
	set score10;
	squareresid = (model1)**2;
run;

proc print data = tenresid;
	var model1 squareresid;
	sum model1 squareresid;
run;

	/* 11 */

proc reg data = housing_2 outest = model11;
	model houseval = Crime Zoned Chas NOX Rooms Dist Hwy Tax PTeacher Bk Lstat;
run; quit;

proc score data=housing_1 score=model11 out=score11 residual type=parms;
	var Crime Zoned Chas NOX Rooms Dist Hwy Tax PTeacher Bk Lstat houseval;
run;

data elevenresid;
	set score11;
	squareresid = (model1)**2;
run;

proc print data = elevenresid;
	var model1 squareresid;
	sum model1 squareresid;
run;

	/* 12 */

proc reg data = housing_2 outest = model12;
	model houseval = Crime Zoned Industry Chas NOX Rooms Dist Hwy Tax PTeacher Bk Lstat;
run; quit;

proc score data=housing_1 score=model12 out=score12 residual type=parms;
	var Crime Zoned Industry Chas NOX Rooms Dist Hwy Tax PTeacher Bk Lstat houseval;
run;

data twelveresid;
	set score12;
	squareresid = (model1)**2;
run;

proc print data = twelveresid;
	var model1 squareresid;
	sum model1 squareresid;
run;

	/* 13 */
```

```
proc reg data = housing_2 outest= model13;
		model houseval = Crime Zoned Industry Chas NOX Rooms Age Dist Hwy Tax PTeacher Bk Lstat;
run; quit;

proc score data=housing_1 score=model13 out=score13 residual type=parms;
		var Crime Zoned Industry Chas NOX Rooms Age Dist Hwy Tax PTeacher Bk Lstat houseval;
run;

data thirteenresid;
		set score13;
		squareresid = (model1)**2;
run;

proc print data = thirteenresid;
		var model1 squareresid;
		sum model1 squareresid;
run;

/* f */

PROC IMPORT OUT= WORK.RSS
	DATAFILE= "Z:\Documents\University\Fall_2011\STA4203\midterm
\RSS.csv"
	DBMS=CSV REPLACE;
    GETNAMES=YES;
    DATAROW=2;
RUN;

data f;
		set RSS;
		RSS_avg_col2col5 = mean(One_RSS_1, Two_RSS_2);
		RSS_avg_col3col4 = mean(One_RSS_2, Two_RSS_1);
run;

ods graphics on;
proc sgplot data = f;
		scatter y = RSS_avg_col2col5 x = P / markerattrs=(symbol=square);
		scatter y = RSS_avg_col3col4 x = P / markerattrs=(symbol=circle);
		xaxis integer values = (1 to 13 by 1);
		yaxis label = "RSS";
run;


/* g */

PROC IMPORT OUT= WORK.housing_a
	DATAFILE= "Z:\Documents\University\Fall_2011\STA4203\midterm
\housing_a.csv"
	DBMS=CSV REPLACE;
    GETNAMES=YES;
    DATAROW=2;
RUN;

PROC IMPORT OUT= WORK.housing_b
	DATAFILE= "Z:\Documents\University\Fall_2011\STA4203\midterm
\housing_b.csv"
	DBMS=CSV REPLACE;
    GETNAMES=YES;
    DATAROW=2;
RUN;

proc reg data = housing_a outest = modela;
		model houseval = lstat pteacher dist rooms nox;
run; quit;

proc score data=housing_b score=modela out=scoreb residual type=parms;
		var lstat pteacher dist rooms nox houseval;
run;

data bresid;
		set scoreb;
		squareresid = (model1)**2;
```

```
run;

proc print data = bresid;
	var model1 squareresid;
	sum model1 squareresid;
run;


proc reg data = housing_b outest = modelb;
	model houseval = lstat pteacher dist rooms nox;
run; quit;

proc score data=housing_a score=modelb out=scorea residual type=parms;
	var lstat pteacher dist rooms nox houseval;
run;

data aresid;
	set scorea;
	squareresid = (model1)**2;
run;

proc print data = aresid;
	var model1 squareresid;
	sum model1 squareresid;
run;
```