

Department of Statistics

COURSE STATS 330

Assignment 2, 2005

Instructions: Hand in your completed assignment to the Student Resource Centre by 4pm on Thursday 18th August.

The file `houseprices.csv` contains data on 199 houses sold in Auckland in April and May of 2005. The variables in the data set are

SalePrice: Sale price in \$\$,
DaysOnMarket: Number of days between listing and sale,
Bedrooms: Number of bedrooms,
Valuation: Valuation New Zealand valuation in \$\$,
YearValued: Year valuation was done, either 2002 or unknown.

1. Load the data into R and create a data frame `houseprices.df`.
2. Using the 2002 valuation data only, fit a regression model to the data using `SalePrice` as the response. Are there any problems with the model? If so, fix them. Can some variables be eliminated from the model? If so, eliminate them.
3. What is the relationship between the valuation and the sale price? Does it depend on the other variables? Interpret your model to answer this part, and also draw a suitable graph.
4. Suppose you live in a house with a 2002 valuation of \$350,000 and 3 bedrooms. What range of prices should you expect to sell for?
5. Now repeat the analysis for the unknown year of valuation data. Would you expect to get the same result? In what way would you expect the model to change? Has it in fact done so?

Data for Assignment 2 (also in file houseprices.csv)

SalePrice	DaysOnMarket	Bedrooms	Valuation	YearValued
485000	28	3	325000	2002
340000	9	3	250000	Unknown
705000	0	4	410000	Unknown
1100000	44	4	660000	2002
565000	110	3	500000	2002
504000	28	3	320000	Unknown
920000	29	3	590000	2002
565000	57	4	390000	Unknown
240000	11	3	140000	Unknown
4800000	55	3	4300000	2002
207000	52	3	140000	2002
395000	0	2	234000	2002
295000	55	4	215000	2002
670000	29	5	405000	Unknown
1140000	31	4	765000	2002
650000	25	4	430000	Unknown
160500	5	3	132000	2002
310000	5	3	194000	2002
850000	21	4	650000	2002
530000	29	4	370000	2002
405000	37	3	270000	2002
331500	26	3	194000	2002
720000	3	3	450000	2002
835000	26	5	760000	Unknown
305500	7	3	270000	Unknown
660000	162	3	420000	Unknown
398000	31	3	914000	Unknown
268000	4	2	180000	2002
730000	36	3	375000	2002
210000	4	3	139000	2002
680000	38	4	405000	Unknown
284000	65	3	189000	Unknown
570000	30	3	400000	Unknown
702500	28	5	580000	Unknown
320000	54	3	220000	2002
575000	21	3	220000	Unknown
333000	20	4	285000	2002
340000	6	4	222000	Unknown
559000	29	5	395000	2002
2235000	42	4	1750000	Unknown
240000	37	3	168000	Unknown
502000	29	3	340000	Unknown
235000	21	3	137000	2002
145000	10	2	122000	Unknown
556000	21	4	350000	2002
895000	11	4	650000	Unknown
480000	11	3	340000	2002
613000	31	3	465000	Unknown
165000	63	3	45000	2002
483000	30	3	250000	2002
186500	6	3	142000	2002
438000	27	3	265000	Unknown
640000	49	4	385000	2002

218000	38	3	162000	2002
800000	52	5	565000	Unknown
800000	46	3	638000	Unknown
406000	2	3	260000	2002
630000	7	7	465000	2002
730000	47	3	425000	Unknown
495000	24	4	400000	Unknown
500000	18	2	350000	2002
415000	45	3	290000	2002
660000	28	3	465000	2002
578000	45	4	345000	2002
151000	47	3	108000	2002
310050	23	2	210000	Unknown
465000	29	2	295000	2002
530000	17	4	400000	2002
379000	24	3	184000	2002
472000	29	3	300000	Unknown
485000	14	3	300000	2002
640000	27	3	405000	2002
304000	3	3	161000	2002
905000	57	4	715000	2002
290000	18	3	215000	2002
510000	39	3	300000	2002
630000	9	4	355000	Unknown
467500	75	4	345000	2002
603000	5	4	440000	Unknown
480000	76	3	335000	2002
474500	37	5	325000	Unknown
490000	13	3	250000	Unknown
420000	8	3	250000	2002
570000	24	4	400000	2002
675000	94	5	500000	2002
700000	37	4	410000	Unknown
331000	17	2	207000	2002
540000	88	4	465000	Unknown
470000	274	5	330000	Unknown
320000	13	3	240000	Unknown
338000	14	3	166000	Unknown
349000	17	2	173000	Unknown
552000	81	3	395000	2002
572000	29	3	310000	2002
510000	18	2	335000	2002
280000	8	3	195000	2002
334000	55	3	215000	2002
310000	73	3	255000	2002
185000	9	3	109000	2002
860000	13	3	625000	2002
285000	30	4	225000	2002
520000	115	4	315000	2002
172000	36	3	139000	2002
275000	26	3	163000	2002
1488800	144	5	1025000	2002
225000	86	3	200000	Unknown
450000	44	4	315000	2002
750000	29	3	435000	2002
288000	26	2	176000	Unknown
230000	50	3	148000	2002

535000	20	3	305000	Unknown
720000	19	3	350000	Unknown
383000	20	3	189000	2002
192000	13	3	145000	Unknown
416000	1	3	280000	2002
705000	33	3	460000	2002
550000	47	4	410000	Unknown
595000	47	4	410000	2002
171500	74	3	144000	2002
390000	30	5	278000	2002
690000	38	6	400000	2002
536200	21	3	345000	Unknown
175000	5	4	150000	2002
314000	22	3	220000	2002
275000	4	4	187000	Unknown
320000	17	3	215000	2002
705000	19	3	415000	Unknown
475000	13	4	380000	2002
374000	42	2	125000	Unknown
308000	6	4	200000	2002
550000	10	2	330000	Unknown
310000	29	5	229000	2002
580000	32	4	490000	Unknown
452000	24	4	345000	2002
601000	0	3	375000	Unknown
280000	16	3	220000	2002
152000	38	3	117000	2002
210000	33	2	150000	Unknown
150000	0	3	117000	2002
377500	19	3	305000	2002
300000	9	2	175000	2002
320250	49	3	242000	2002
490000	185	4	405000	2002
194000	79	5	252000	Unknown
640000	21	3	345000	Unknown
470000	8	3	320000	2002
320000	52	2	196000	2002
1322000	32	4	800000	Unknown
835000	37	3	475000	Unknown
2150000	18	6	1075000	2002
490000	89	3	260000	Unknown
1195000	19	4	780000	2002
520000	57	5	410000	2002
1500000	71	4	805000	Unknown
210000	32	3	161000	2002
323000	17	3	230000	Unknown
325000	143	3	240000	2002
525000	49	3	290000	Unknown
435000	34	2	270000	Unknown
480000	32	3	370000	Unknown
420000	16	3	260000	2002
345000	17	4	225000	Unknown
700000	34	4	440000	2002
450000	70	4	365000	Unknown
562000	47	4	410000	Unknown
145000	73	3	127000	2002
290000	25	2	172000	2002

382500	6	3	215000	2002
810750	10	3	445000	2002
433000	27	3	250000	2002
520000	177	3	370000	Unknown
340000	47	4	235000	Unknown
625000	29	3	395000	2002
240000	10	3	142000	Unknown
375000	41	4	198000	2002
555000	16	3	380000	Unknown
660000	157	5	380000	Unknown
420000	44	4	285000	Unknown
256000	26	2	187000	Unknown
435000	48	3	285000	2002
890000	33	3	595000	Unknown
285000	60	3	265000	Unknown
1050000	40	3	705000	Unknown
355000	62	5	252000	Unknown
335000	18	3	235000	2002
666000	32	4	305000	2002
440000	35	3	315000	Unknown
470000	31	3	320000	Unknown
339000	17	3	210000	Unknown
960000	19	5	620000	Unknown
342500	30	3	230000	2002
475000	43	4	430000	2002
462000	19	4	340000	2002
420000	15	4	290000	2002
620000	48	5	390000	Unknown
445000	46	2	290000	2002
150000	0	3	115000	2002
510000	68	3	400000	Unknown
850000	29	3	525000	Unknown