The Professor Proposes

ECO404 | April 2nd, 2014

Alejandro Bilbao, Alex Trivanovic, Lisa Guo

Agenda

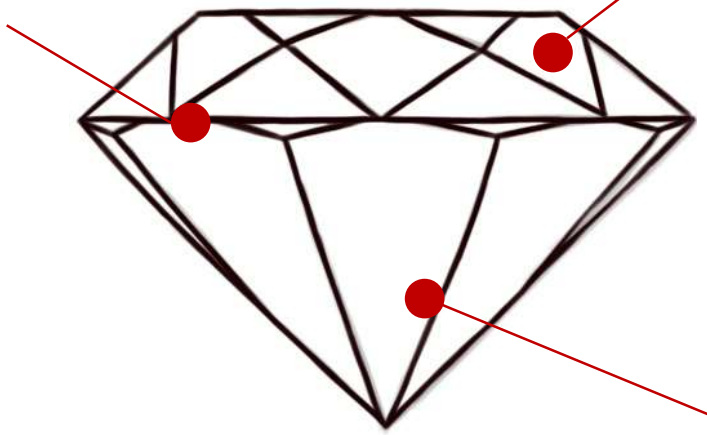
1. Overview and Issue Statement
2. Analysis
3. Recommendation

Problem Statement

What is the fair price of the ring?
Budget: $2,000 - $4,000

Characteristics of Diamonds

- Carat
- Colour
- Clarity
- Cut



- Polish
- Symmetry
- Certification
- Wholesaler

Characteristic Measurements

Carat	1 carat = 0.2 kg
Colour	Colourless | D E F | G H I | J K | L M N | O P Q R S | T U V W X Y Z | Yellow
Clarity	Flawless | FL | IF | VVS 1 – 2 | VS 1 – 2 | SI 1 – 3 | I 1 -3 | Included
Cut	Poor | Fair | Good | Very Good | Excellent | Ideal
Polish	
Symmetry	
Certification	GIA | AGS | EGL | IGI
Wholesaler	1 | 2 | 3

The Engagement Ring

Characteristic	Diamond Rating
Carat Weight	0.9
Cut	Very Good
Color	J (Faint Yellow)
Clarity	S12 (Few inclusions at 10x)
Polish	Good
Symmetry	Very Good
Certification	GIA

Analysis: Overview

Analysis: Comparable Rings

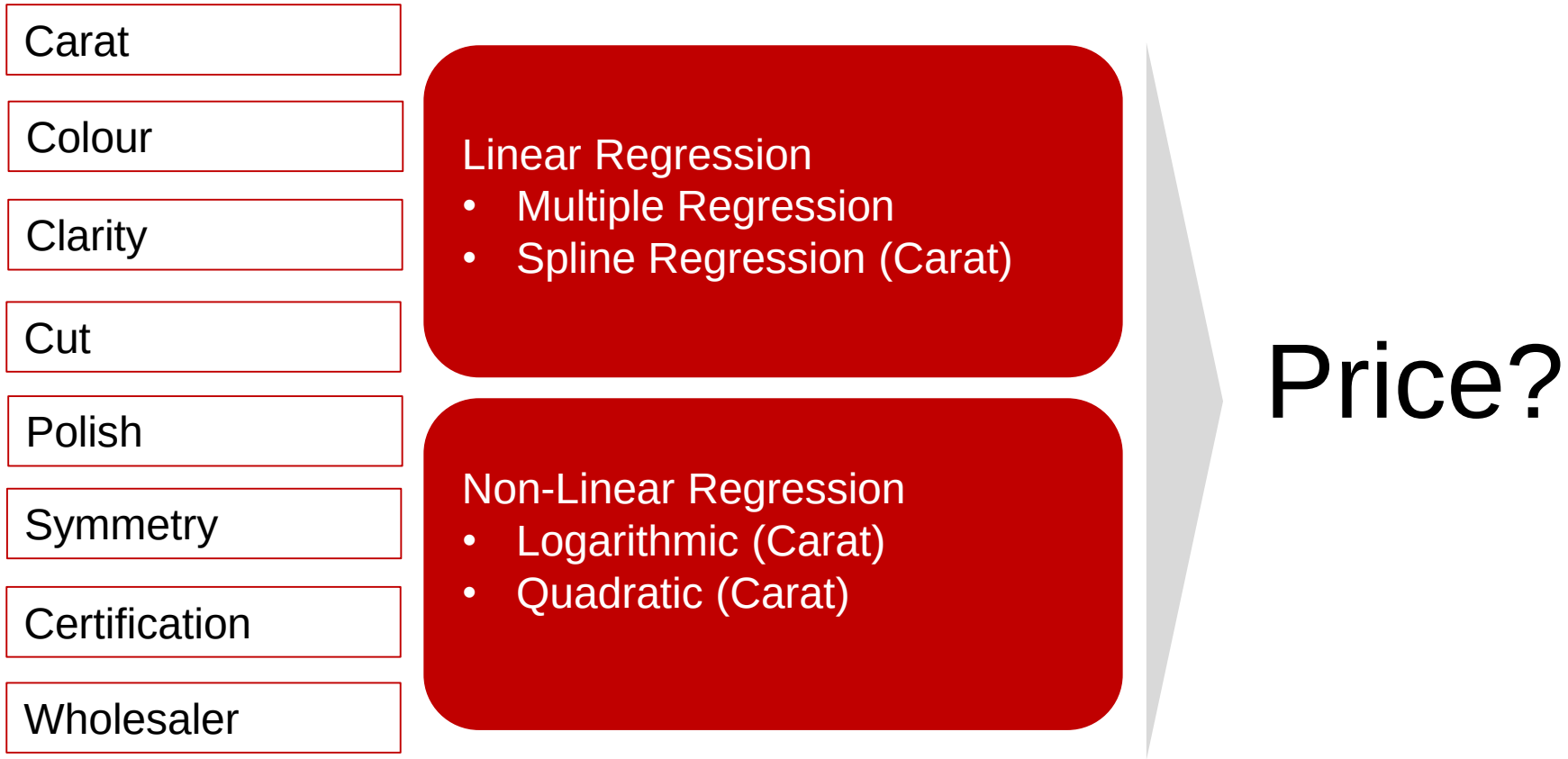
Characteristic	The Diamond	Comparable Set
Carat Weight	0.9	0.88 – 0.92
Cut	Very Good	Good – Excellent
Color	J (Faint Yellow)	G- N (Nearly Colorless – Very Light Yellow)
Clarity	S12 (Few inclusions at 10x)	SI3 – SI1
Polish	Good	Fair – Very Good
Symmetry	Very Good	Good – Excellent
Certification	GIA	GIA

Estimated Price:

Point Estimate =	$ 2,899.39
Max Estimate =	$ 3,069.40
Min Estimate =	$ 2,638.69

Comparable Rings:	8

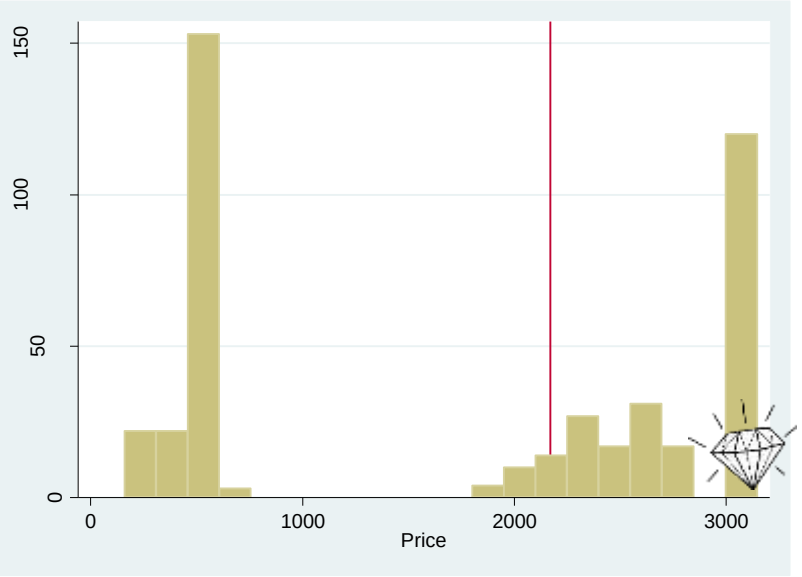
Methodology: Regressions



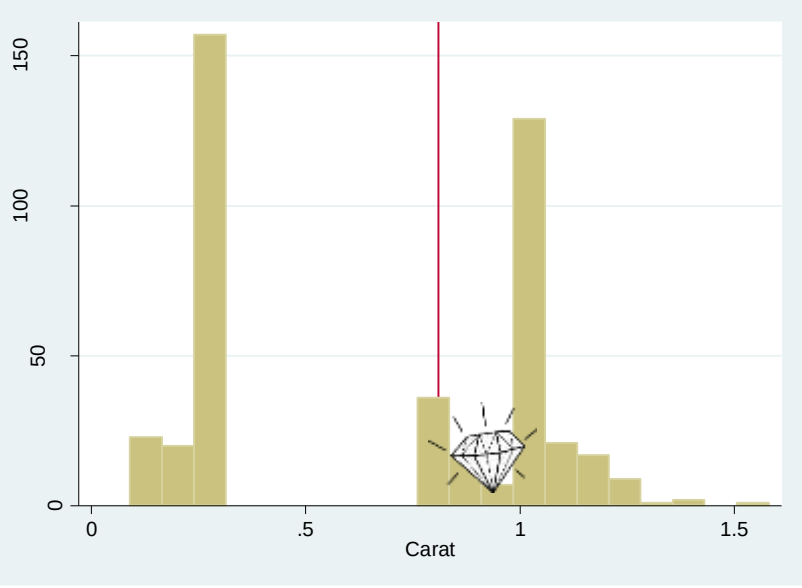
Analysis: Continuous Variables

Number of Observations: 440			
	Average	Range	Std. Dev
Price	$1716.74	$160 - $3145	$1175.689
Carat	0.66925	0.09 – 1.58	0.3798

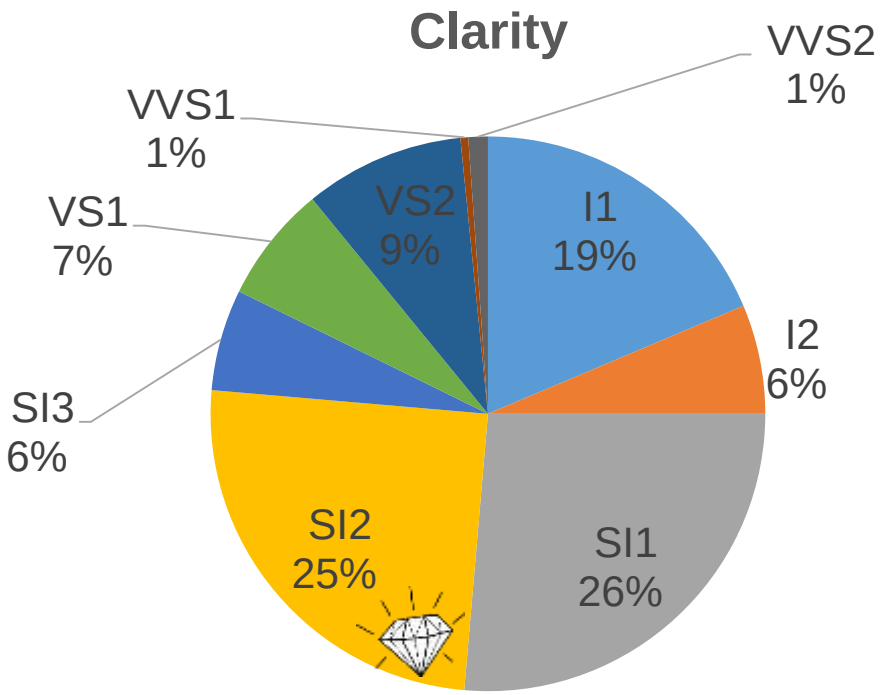
Price Distribution

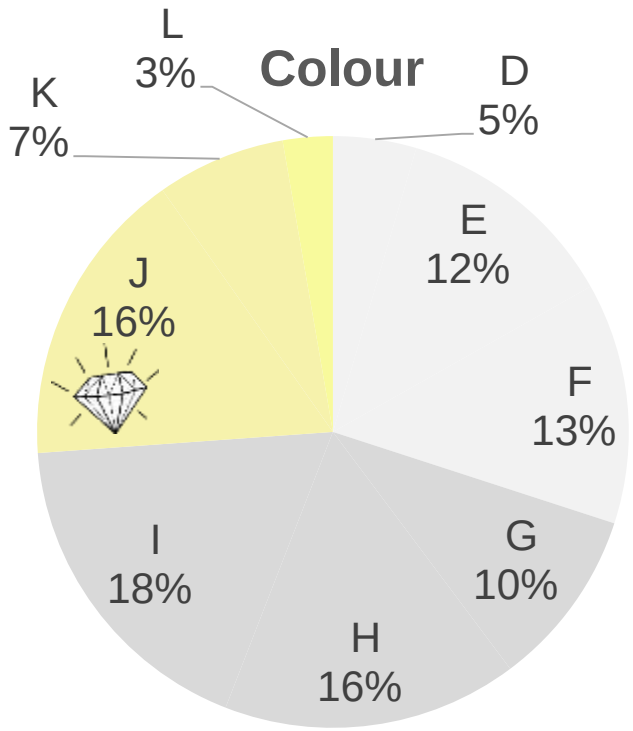


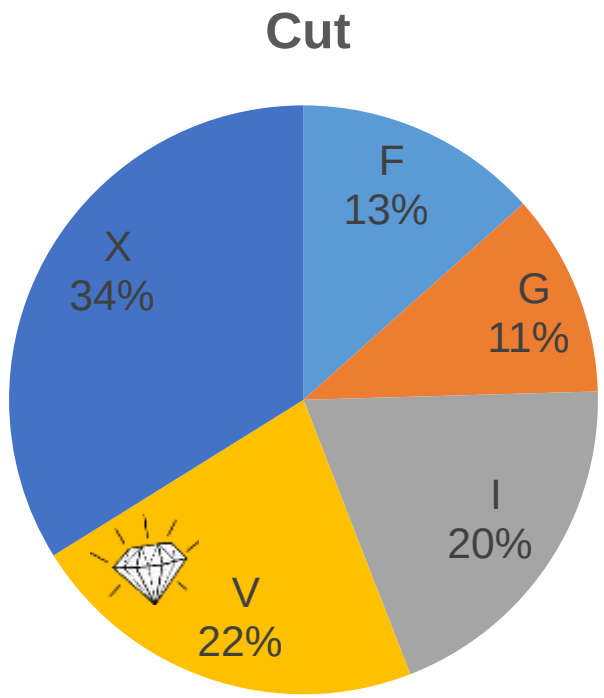
Carat Distribution



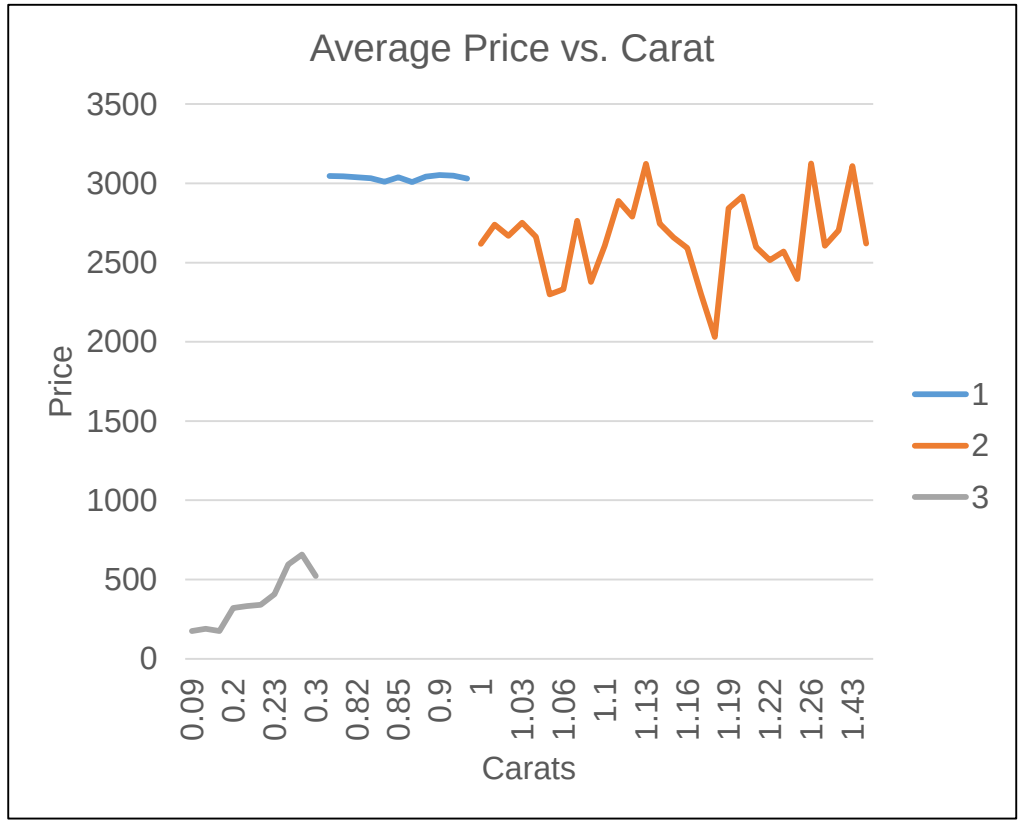
Analysis: Discrete Variables







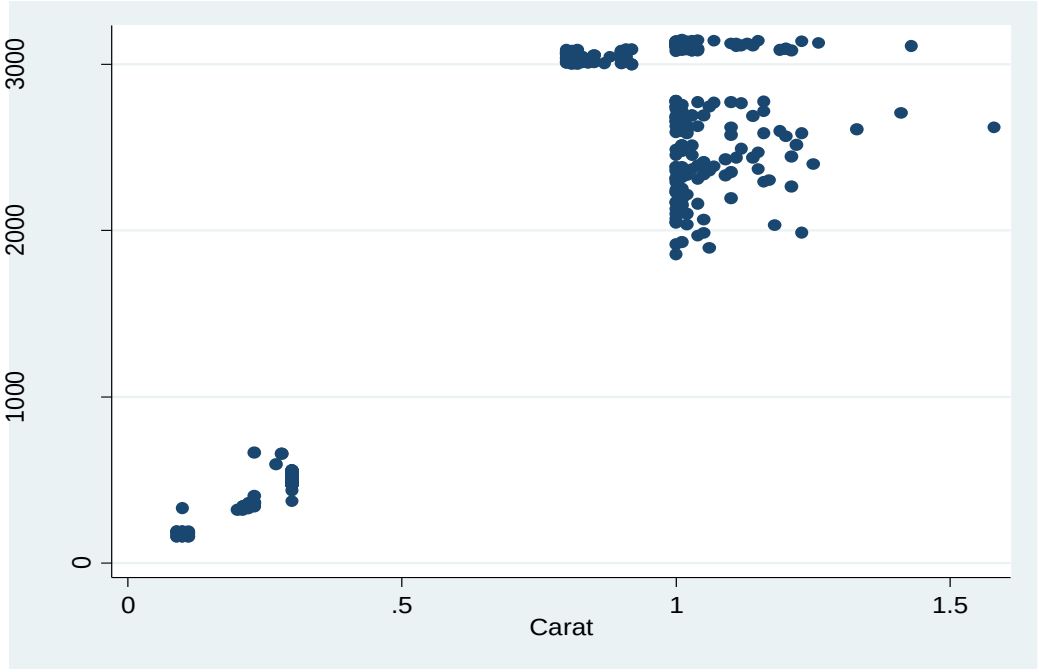
Analysis: Wholesaler



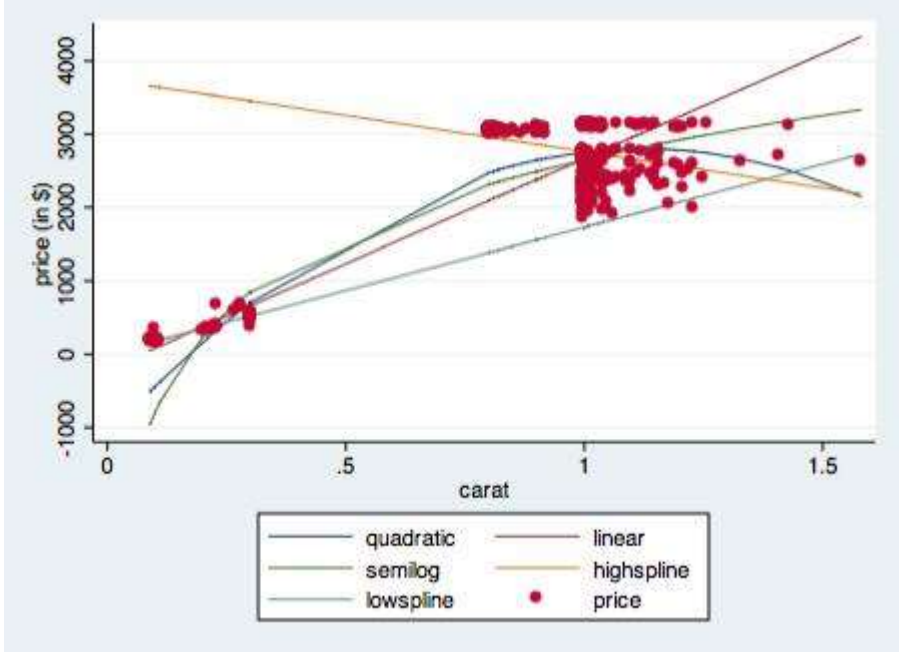
Wholesaler	3	2	1
Number	200	180	60
Average Carat Size	0.27	1.06	0.84
Average Price	468	2662	3043
Range of Color	D - K (Colorless - Faint yellow)	D - L (Colorless - Very light yellow)	D - J (Colorless - Faint yellow)
Range of Cut	Fair - Ideal	Fair - Ideal	Fair - Ideal
Range of Clarity	I1 - VVS2	I1 - VS2	SI1 - VS2
Range of Polish	Good - Excellent	Fair - Excellent	Good - Ideal
Range of Symmetry	Good - Excellent	Fair - Excellent	Good - Ideal
Certification	GIA, IGI	EGL, DOW, GIA	AGS, GIA

The Price and Carat Relationship

Scatterplot of Price vs. Carat



Fitted Lines of Price vs. Carat



Creating the Best Model

Find the best price and carat relationship

Find the optimal division of discrete variables

Optimal Model for Predicting Price

Linear | Spline | Quadratic | Logarithmic

Colour | Cut | Clarity | Polish

Symmetry | Certification | Wholesaler

Use entire scale? Dummy for a portion of the scale?

Which regression gives the best results?

Comparison of Models

	Linear	Spline	Quadratic	Logarithmic
Constant	303.33	-144.86	467.10	1816.59
Carat	1510.94	1812.96	Carat: 1071.99	lnCarat: 285.66
			Carat-Sq: 232.150	
Discrete Variables				
Full Scale Significant				
Colour		Clarity		"Wholesaler"
(-)		(+)		(-) for 3, (+) for 2
Partial Scale Significant				
Certificate			Cut	
GIA/AGS vs. EGL/DOW			Poor to Very Good vs. Excellent	
No Significance				
Polish			"Wholesaler"	

Comparison of Models

	Linear	Spline	Quadratic	Logarithmic
R^2	0.9818	0.7057	0.9820	0.9784
MSE	161.75	206.25	161.52	176.75
Advantages	• Simple	• Captures carat “gap”	• Best fit • Captures non-linearity	• Captures non-linearity • All variables significant
Disadvantages	• Does not capture non-linearity • Wholesaler is insignificant	• Worst fit • Insignificant constant • Does not capture non-linearity • Reduces sample size	• Carat-squared coefficient insignificant • Wholesaler is insignificant	• Lower R^2 than linear

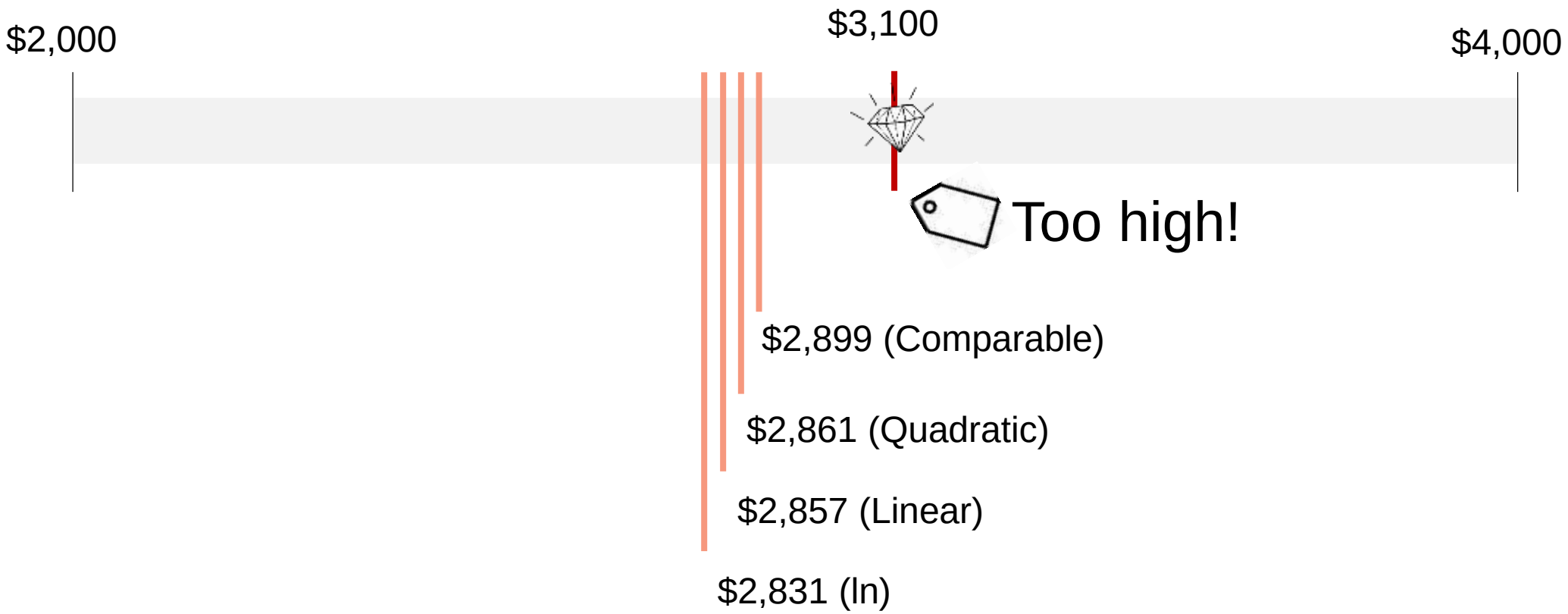
Example Model: Logarithmic

Variable	Coefficient
lnCarat	285.6625
Scale For Colour	
2	-102.9674
3	-242.1503
4	-554.7601
Scale For Clarity	
3	426.3774
4	602.8649
5	867.3755
6	930.583
7	952.4909
8	1003.027
9	950.4456
10	1054.666
Wholesaler	
2	208.7881
3	-2193.272
New Scale For Symmetry	
3	188.6193
4	192.0782
5	182.9909
Dummies and Constant	
Dummy cut	46.78337
dummy_bestlab	239.6189
_Constant	1816.598

The Results

Model	Predicted Price	Lower 95% CI	Upper 95% CI	Lower 95% PI	Upper 95% PI
Logarithmic (Wholesaler)	$2,831.08	$2,754.77	$2,907.38	$2,474.99	$3,187.16
Quadratic (Wholesaler)	$2,861.20	$2,793.53	$2,928.86	$2,536.10	$3,186.29
Linear (Wholesaler)	$2,857.23	$2,789.67	$2,924.80	$2,531.68	$3,182.79
Logarithmic (No Wholesaler)	$2,384.86	$2,282.44	$2,487.27	$1,691.66	$3,078.05
Quadratic (No Wholesaler)	$2,615.68	$2,552.07	$2,679.29	$2,190.49	$3,040.87
Linear (No Wholesaler)	$2,584.29	$2,515.65	$2,652.94	$2,122.67	$3,045.91

You’re paying too much!



Recommendation

Characteristic	Original Ring	Ring A	Ring B
Price	$3,100	$3,006	$3,064
Carat Weight	0.9	0.9	0.9
Cut	Very Good	Very Good	Good
Color	J (Faint Yellow)	J (Faint Yellow)	J (Faint Yellow)
Clarity	S12 (Few inclusions at 10x)	VS2 (Very Slightly Included)	VS2 (Very Slightly Included)
Polish	Good	Very Good	Excellent
Symmetry	Very Good	Very Good	Excellent
Certification	GIA	GIA	AGS

Thank-you,
Questions?

Appendix

- Link to spreadsheet – for comparability

Appendix: Comparable Diamonds

Comparable Diamonds														
Carat	Dummy_Colour	Colour	Scale for Clarity	Clarity	Scale for Cut	Cut	Dummy_Best LAB	Certification	Scale for Polish	Polish	Scale for Symmetry	Symmetry	Price	Wholesaler
0.91	3	J	6	SI1	4	V	1	GIA	4	V	4	V	3023	1
0.9	2	I	6	SI1	4	V	1	GIA	5	X	4	V	3028	1
0.91	2	H	5	SI2	3	G	1	GIA	4	V	3	G	3038	1
0.9	2	H	5	SI2	4	V	1	GIA	4	V	3	G	3043	1
0.9	2	I	5	SI2	3	G	1	GIA	4	V	4	V	3069	1
0.9	2	I	5	SI2	4	V	1	GIA	4	V	4	V	3069	1
0.9	2	I	6	SI1	3	G	1	GIA	4	V	3	G	3081	1
0.92	3	J	6	SI1	4	V	1	GIA	4	V	4	V	3091	1

Recommended Diamonds														
Carat	Dummy_Colour	Colour	Scale for Clarity	Clarity	Scale for Cut	Cut	Dummy_Best LAB	Certification	Scale for Polish	Polish	Scale for Symmetry	Symmetry	Price	Wholesaler
0.9	**3**	J	7	VS2	4	V	1	GIA	4	V	4	V	3006	1
0.9	**3**	J	7	VS2	3	G	2	AGS	5	X	5	X	3064	1

Appendix: Results

- Carat Graphs

Appendix: Final Quadratic Regression

Final Quadratic Regression						
Source	**SS**	**df**	**MS**	**Number of obs**	=	441
				F(20, 420)	=	1142.07
Model	597723839	20	29886191.9	**Prob > F**	=	0
Residual	10990755.2	420	26168.4649	**R-squared**	=	0.9819
				Adj R-squared	=	0.9811
Total	608714594	440	1383442.26	**Root MSE**	=	161.77
price	**Coef.**	**Std. Err.**	**t**	**P>|t|**	**[95% Conf.**	**Interval]**
carat	1071.998	319.2041	3.36	0.001	444.5618	1699.435
caratsqr	232.1507	153.1811	1.52	0.13	-68.94634	533.2477
scaleforcolour						
2	-124.6927	22.64757	-5.51	0	-169.2094	-80.17596
3	-304.5214	30.06927	-10.13	0	-363.6264	-245.4164
4	-661.7041	55.57112	-11.91	0	-770.9362	-552.4719
scaleforclarity						
3	558.1136	41.78554	13.36	0	475.9788	640.2485
4	776.4684	55.5983	13.97	0	667.1829	885.754
5	1058.521	51.4329	20.58	0	957.4235	1159.619
6	1136.534	57.30882	19.83	0	1023.886	1249.182
7	1158.617	62.15503	18.64	0	1036.443	1280.791
8	1223.329	67.64517	18.08	0	1090.363	1356.294
9	1156.878	91.44662	12.65	0	977.1281	1336.628
10	1295.136	131.6794	9.84	0	1036.303	1553.968
wholesaler						
2	63.46869	52.78531	1.2	0.23	-40.28761	167.225
3	-1760.534	103.3932	-17.03	0	-1963.767	-1557.302
newscaleforsymmetry						
3	191.4221	39.62702	4.83	0	113.5301	269.3141
4	222.7717	41.24056	5.4	0	141.7081	303.8353
5	219.6325	46.28347	4.75	0	128.6564	310.6086
dummycut	40.14575	16.59549	2.42	0.016	7.525181	72.76632
dummy_bestlab	264.4796	29.81245	8.87	0	205.8794	323.0798
_cons	467.1027	172.7448	2.7	0.007	127.5507	806.6547

Appendix: Final Linear Regression

Final Linear Regression						
Source	**SS**	**df**	**MS**	**Number of obs**	=	440
				F(19, 420)	=	1198.52
Model	595816375	19	31358756.6	**Prob > F**	=	0
Residual	10989146.4	420	26164.6343	**R-squared**	=	0.9819
				Adj R-squared	=	0.9811
Total	606805521	439	1382244.92	**Root MSE**	=	161.75
price.	**Coef**	**Std. Err.**	**t**	**P>|t|**	**[95% Conf.**	**Interval]**
carat	1510.937	137.2017	11.01	0	1241.25	1780.625
scaleforcolour						
2	-132.0719	22.27041	-5.93	0	-175.8473	-88.29659
3	-314.9913	30.1648	-10.44	0	-374.2841	-255.6985
4	-662.1205	55.57713	-11.91	0	-771.3645	-552.8765
scaleforclarity						
3	538.7522	39.42788	13.66	0	461.2517	616.2528
4	750.2744	52.12226	14.39	0	647.8214	852.7274
5	1039.661	49.18782	21.14	0	942.9764	1136.346
6	1123.757	55.9773	20.08	0	1013.726	1233.788
7	1148.799	61.28565	18.74	0	1028.334	1269.264
8	1217.562	67.16188	18.13	0	1085.547	1349.577
9	1141.67	90.46201	12.62	0	963.8552	1319.485
10	1283.091	130.9818	9.8	0	1025.63	1540.553
wholesaler						
2	56.39354	51.97516	1.09	0.279	-45.77031	158.5574
3	-1657.753	81.22857	-20.41	0	-1817.418	-1498.087
newscaleforsymmetry						
3	189.9703	39.60967	4.8	0	112.1124	267.8281
4	210.6689	40.59178	5.19	0	130.8805	290.4573
5	207.0838	45.48911	4.55	0	117.6691	296.4985
dummycut	44.19004	16.43469	2.69	0.007	11.88554	76.49454
dummy_bestlab	247.2777	26.94654	9.18	0	194.3108	300.2446
_cons	303.33	141.4398	2.14	0.033	25.31188	581.3481

Appendix: Final Logarithmic Regression

Final Logarithmic Regression						
Source	**SS**	**df**	**MS**	**Number of obs**	=	441
				F(19, 421)	=	998.15
Model	595495210	19	31341853.2	**Prob > F**	=	0
Residual	13219384	421	31399.962	**R-squared**	=	0.9783
				Adj R-squared	=	0.9773
Total	608714594	440	1383442.26	**Root MSE**	=	177.2
price	**Coef.**	**Std. Err.**	**t**	**P>|t|**	**[95% Conf.**	**Interval]**
lncarat	285.6625	50.72263	5.63	0	185.9614	385.3637
scaleforcolour						
2	-102.9674	24.40753	-4.22	0	-150.9432	-54.99155
3	-242.1503	31.89968	-7.59	0	-304.8527	-179.4478
4	-554.7601	59.5737	-9.31	0	-671.8591	-437.6611
scaleforclarity						
3	426.3774	42.67159	9.99	0	342.5015	510.2533
4	602.8649	56.97219	10.58	0	490.8795	714.8503
5	867.3755	51.47011	16.85	0	766.2051	968.5459
6	930.583	57.97386	16.05	0	816.6287	1044.537
7	952.4909	63.65608	14.96	0	827.3676	1077.614
8	1003.027	69.42115	14.45	0	866.572	1139.482
9	950.4456	97.47485	9.75	0	758.8476	1142.044
10	1054.666	141.445	7.46	0	776.6392	1332.692
wholesaler						
2	208.7881	53.14136	3.93	0	104.3326	313.2435
3	-2193.272	62.87429	-34.88	0	-2316.859	-2069.685
newscaleforsymmetry						
3	188.6193	43.39228	4.35	0	103.3268	273.9118
4	192.0782	44.75892	4.29	0	104.0994	280.057
5	182.9909	50.25239	3.64	0	84.21407	281.7678
dummycut	46.78337	18.10801	2.58	0.01	11.19	82.37674
dummy_bestlab	239.6189	31.98712	7.49	0	176.7446	302.4933
_cons	1816.598	85.39026	21.27	0	1648.753	1984.442