

Project: Diamond Prices

Step 1: Understanding the Model

$$\text{Price} = -5269 + 8413 \times \text{Carat} + 158.1 \times \text{Cut} + 454 \times \text{Clarity}$$

1. According to the model, if a diamond is 1 carat heavier than another with the same cut, how much more should I expect to pay? Why?

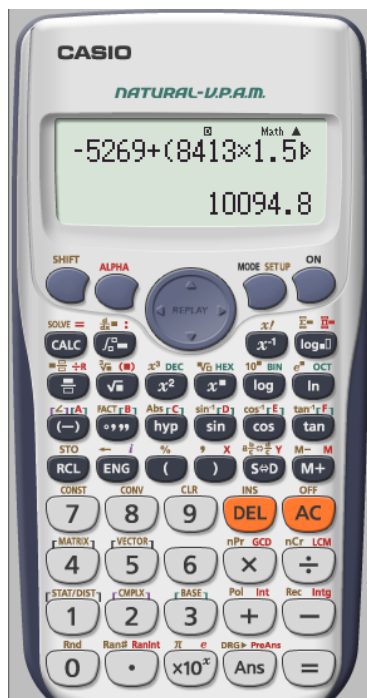
Answer...

We can see that the carat coefficient is 8413. Which means that each one carat is heavier than the others with the same cut. So the price goes up and the increase will be \$8413.

2. If you were interested in a 1.5 carat diamond with a **Very Good** cut (represented by a 3 in the model) and a **VS2** clarity rating (represented by a 5 in the model), how much would the model predict you should pay for it?

Answer...

$$\begin{aligned}\text{Price} &= -5269 + 8413 \times \text{Carat} + 158.1 \times \text{Cut} + 454 \times \text{Clarity} \\ &= -5269 + (8413 \times 1.5) + (158.1 \times 3) + (454 \times 5) \\ &= \$10094.8\end{aligned}$$

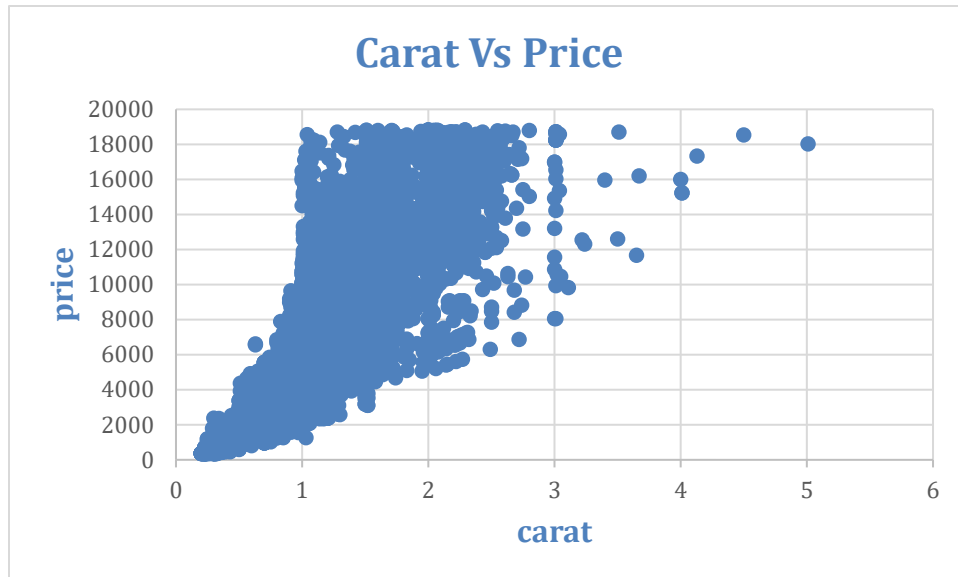


Therefore, for 1.5 carats of diamonds, we must pay \$10094.8.

Step 2: Visualize the Data

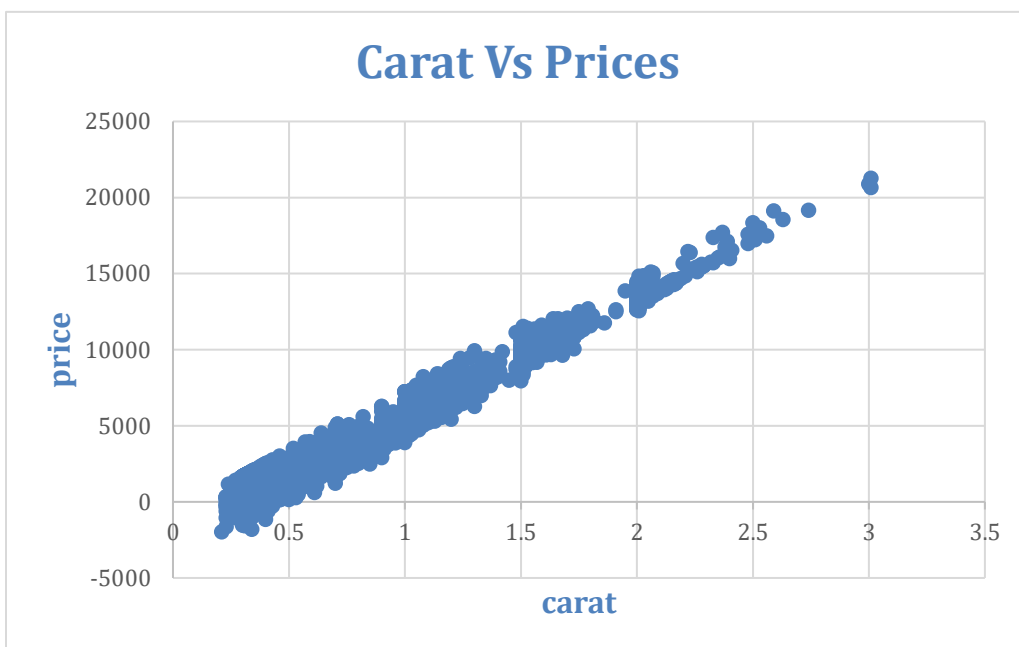
Make sure to plot and include the visualizations in this report. For example, you can create graphs in Excel and copy and paste the graphs into this Word document.

1. Plot 1 - Plot the data for the [diamonds in the database](#), with carat on the x-axis and price on the y-axis.



carat	price
0.51	1749
2.25	7069
0.7	2757
0.47	1243
0.3	789
0.33	728
2.01	18398
0.51	2203
1.7	15100
0.53	1857
0.39	834
1.5	7708
1	6272
1.29	5676
2.01	16776
1.13	7404
0.7	1702
0.38	606
1.17	5423
1.51	8033
0.4	1279
0.41	863
0.51	1893
1	3584
1.09	10196
0.39	1082

2. Plot 2 - Plot the data for the diamonds for which [you are predicting prices](#) with carat on the x-axis and predicted price on the y-axis.



carat	predicting prices
1.22	6989.26
1.01	5814.33
0.71	3448.53
1.01	4926.63
0.27	517.01
0.52	1554.16
1.01	5222.53
0.59	1847.17
1.01	4906.33
2.03	13507.89
1.35	8990.95
0.74	3109.12
0.9	4297.1
0.3	-158.9
1.01	5814.33
1.02	4536.46
2.05	13972.05
0.54	1426.52
0.72	3848.86
2	13097.4
1.57	9479.81
0.89	4212.97
0.33	863.69
0.3	1361.2
1.79	12534.57
1.11	5767.93

3. What strikes you about this comparison? After seeing this plot, do you feel confident in the model's ability to predict prices?

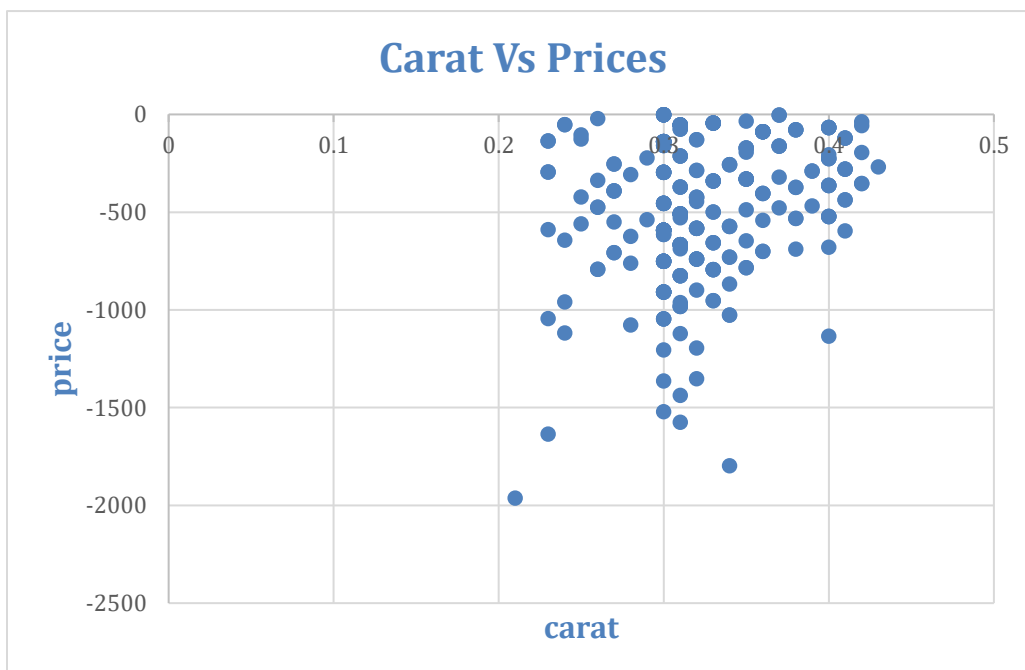
Caught my attention:

Old Diamond Price:

- The graph between carats and prices shows that (0 to 1) is in a linear trend, then (1 to 3) it begins to spread and has weak wear.
- There are some factors independent of the expected prices that affect the price of diamonds (for example: color, clarity, etc.).

New Diamond Price:

- To be clear about the price of diamonds, we noticed that in the new graph, a strong positive correlation occurred between carats and prices, so I had a (linear regression). However, we noted some examples where the price turned negative. This incorrect approach indicated a non-linear regression and a solution to this problem (we expect diamond prices).



carat	predicting prices
0.3	-158.9
0.35	-330.05
0.33	-44.31
0.25	-125.55
0.3	-0.8
0.33	-794.21
0.33	-656.41
0.33	-498.31
0.3	-750.7
0.3	-592.6
0.37	-161.79
0.3	-750.7
0.38	-373.56
0.31	-54.47
0.31	-54.47
0.37	-319.89
0.28	-306.86
0.31	-508.47
0.4	-363.4
0.29	-222.73
0.35	-784.05
0.27	-390.99
0.38	-531.66
0.3	-138.6
0.39	-289.43
0.32	952.31

Step 3: Make a Recommendation

1. What price do you recommend the jewelry company to bid? Please explain how you arrived at that number.

Answer...

❖ \$8,213,465

- ❖ The sum of the price prediction values equals (11733522.76), then multiplied the total by 70% and gave me the value of the price recommended by the jewelry company for bidding (\$ 8,213,465).

	B	C	D	E	F	G	H	I	J	K	L	M
	carat	cut	cut_ord	color	clarity	clarity_ord	carat	predicting prices				
1	1.22	Premium	4	G	SI1	3	1.22	6989.26				
2	1.01	Good	2	G	VS2	5	1.01	5814.33	11733522.76			
3	0.71	Very Good	3	I	VS2	5	0.71	3448.53				
4	1.01	Ideal	5	D	SI2	2	1.01	4926.63	8213465.932			
5	0.27	Ideal	5	H	VVS2	6	0.27	517.01				
6	0.52	Premium	4	G	VS1	4	0.52	1554.16				
7	1.01	Premium	4	F	SI1	3	1.01	5222.53				
8	0.59	Ideal	5	D	SI1	3	0.59	1847.17				
9	1.01	Good	2	E	SI1	3	1.01	4906.33				
10	2.03	Ideal	5	F	SI2	2	2.03	13507.89				
11	1.35	Premium	4	H	VS2	5	1.35	8990.95				
12	0.74	Ideal	5	G	SI1	3	0.74	3109.12				
13	0.9	Premium	4	D	SI1	3	0.9	4297.1				
14	0.3	Good	2	G	VS2	5	0.3	-158.9				
15	1.01	Good	2	F	VS2	5	1.01	5814.33				
16	1.02	Good	2	H	SI2	2	1.02	4536.46				
17	2.05	Premium	4	G	SI1	3	2.05	13972.05				
18	0.54	Ideal	5	I	SI1	3	0.54	1426.52				
19	0.72	Ideal	5	G	VS2	5	0.72	3848.86				
20	2	Premium	4	J	SI2	2	2	13097.4				
21	1.57	Good	2	G	SI2	5	1.57	6470.81				