Variable Analysis: Cut

Katie Dunning (umu5hh)

2025-03-08

Variable Analysis: Cut

Variable Description   According to the client, Blue Nile, diamond cuts refer to the proportion, dimensions, and faceting of a diamond. The cuts of a diamond are descriptive of the quality of the gem's faceting, proportion, and polish, as well as how symmetrical the piece is. Higher grade cuts are more symmetrical and have better light performance. Diamonds in the "good" category are considered to be in the top 25% of diamond cut quality, while "very good" is in the top 15%, and "ideal" is in the top 3%. Astor by Blue Nile is touted to "reflect the most light possible" and are grade/certified by a number of third parties.

Citation: Blue Nile. (n.d.). Diamond cut: Grading scale and buying tips. https://www.bluenile.com/education/diamonds/cut?srsltid=AfmBOop9PWytZgjMIGYvLwuojS7LFcIV_5Pwh_pNHS44fwFoimTfBnbC

### Univariate Analysis

Diamond Cut Counts by Quality
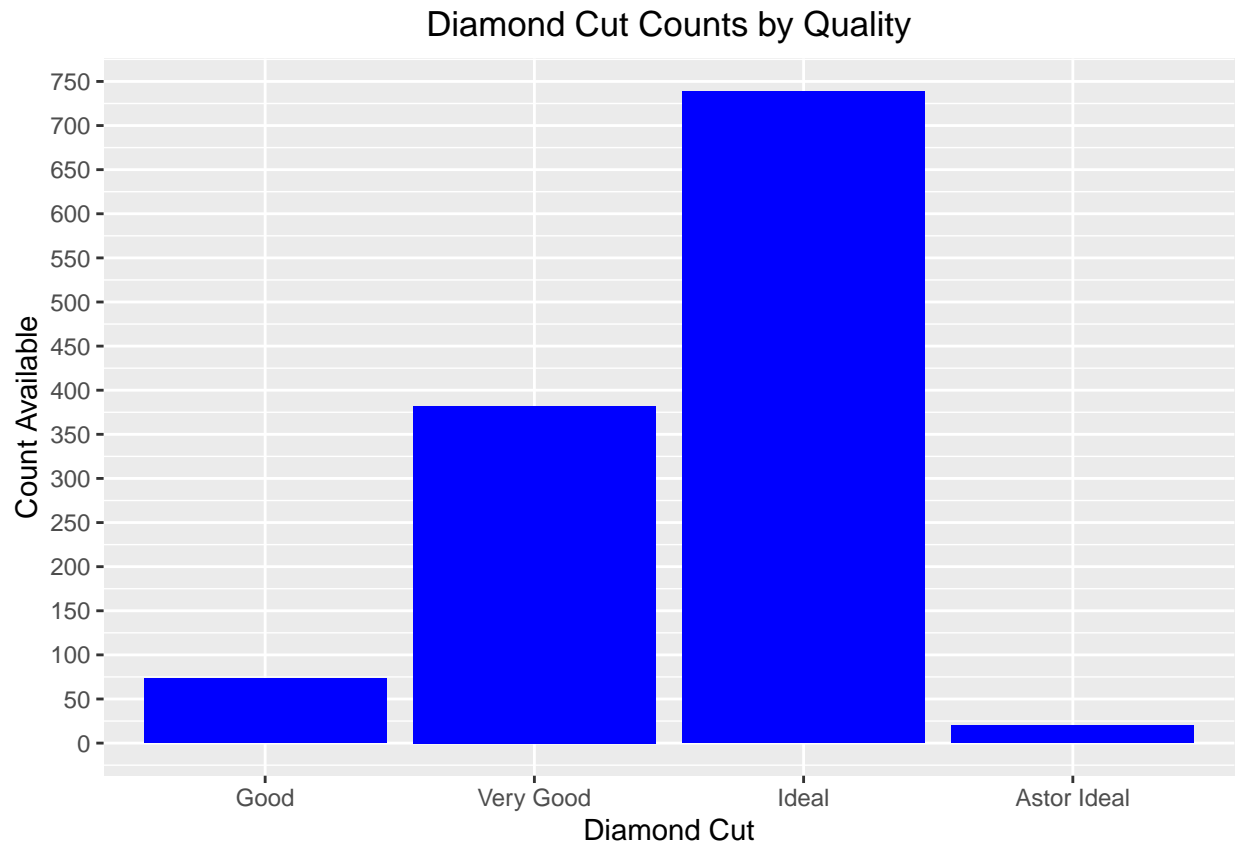
The cut is highly skewed regarding available quantity of each type. There are vastly more ideal cuts (739)

and very good cuts (382) compared to the good cuts (73) and the Astor ideal cuts (20).

Diamond Price by Cut
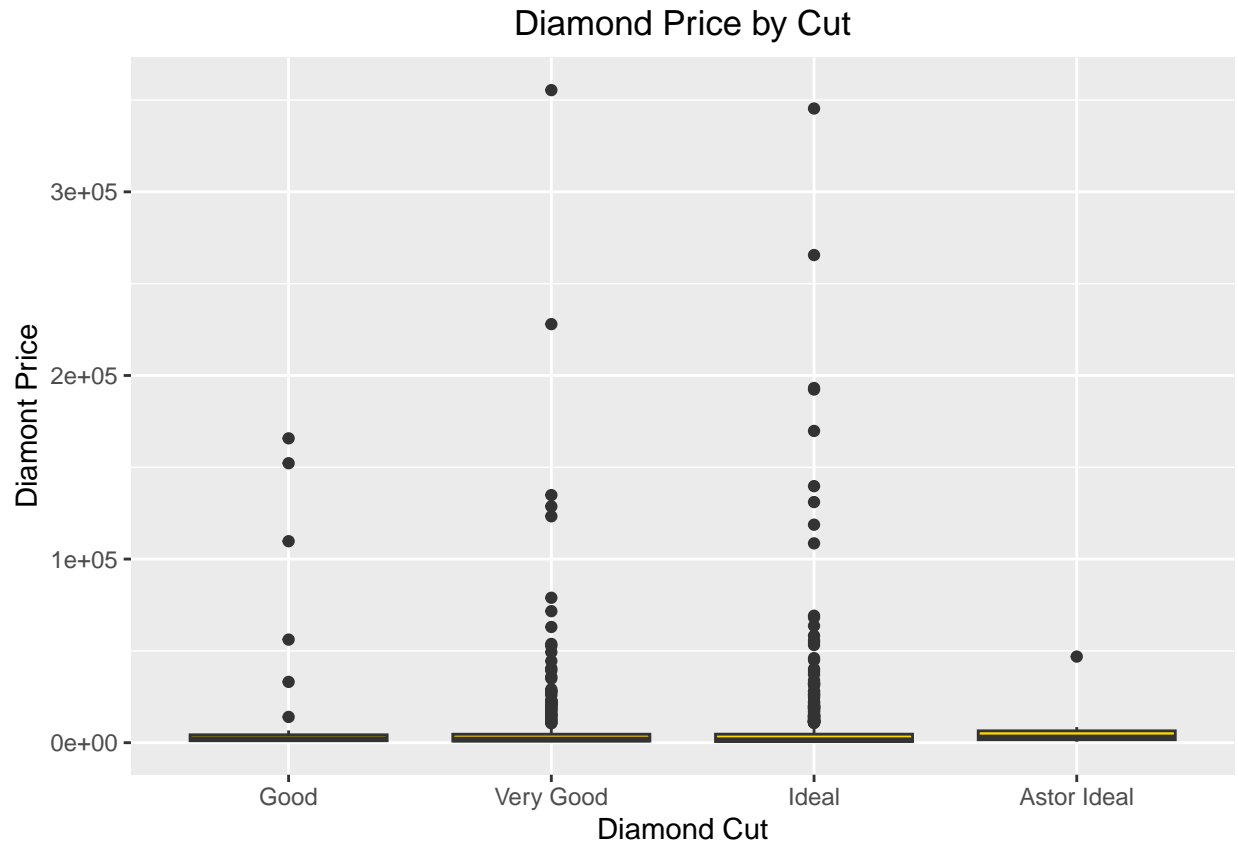
When viewing these original side-by-side boxplots, it is obvious that cut has a great deal of price variation and many outliers. This could suggest the other variables at play (carat, color, and clarity) causing deviations. The most outliers are witnessed in the very good and ideal ranges, however, they are the cuts with the largest quantity. The next graph will represent a closer look at the data, focusing on the boxplot boxes, which account for 50% of the data in each cut category.

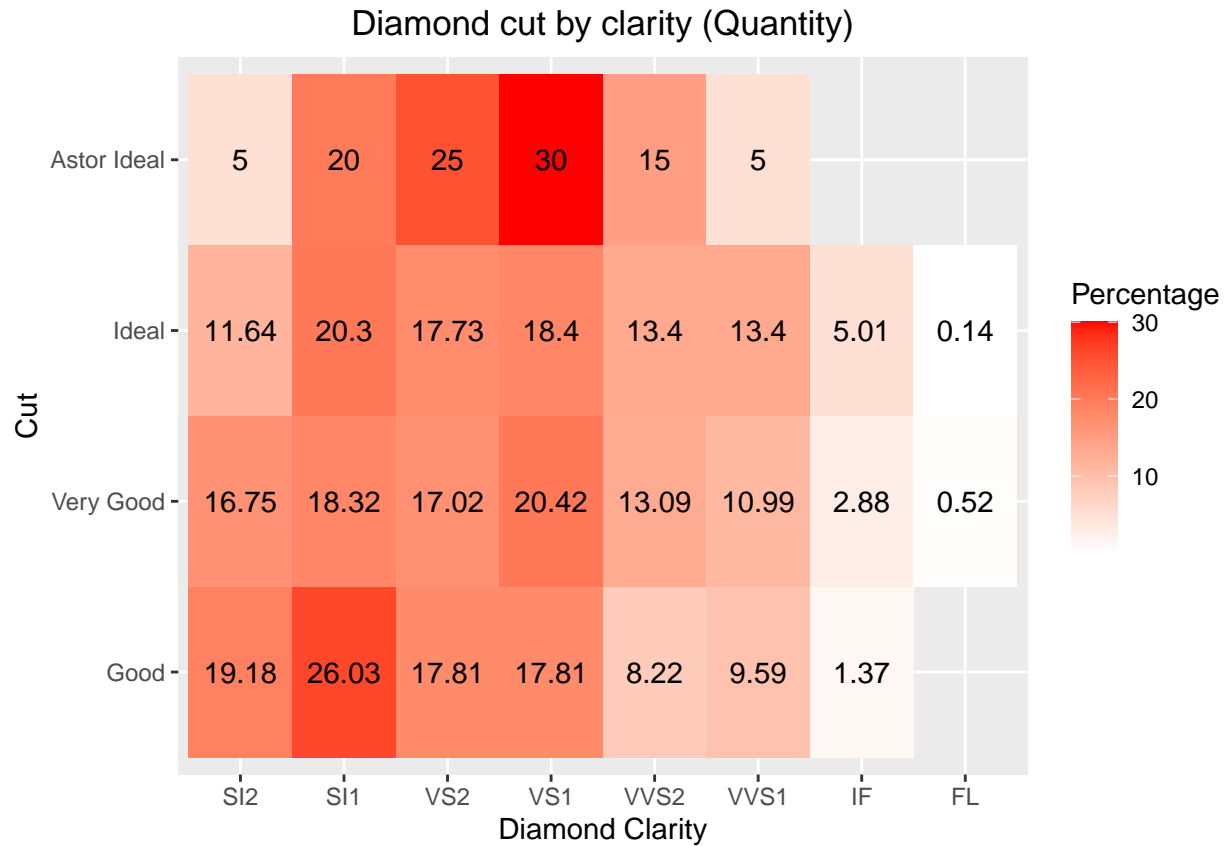| cut	q0	q25	q50	q75	q100	mean
Good	426	1098.00	1903.0	4300.0	165766	9467.288
Very Good	322	901.00	1744.5	4613.0	355403	7757.626
Ideal	354	649.00	1354.0	4643.0	345397	6488.932
Astor Ideal	520	1625.25	2854.0	6452.5	46893	5851.550

Diamond Price by Cut
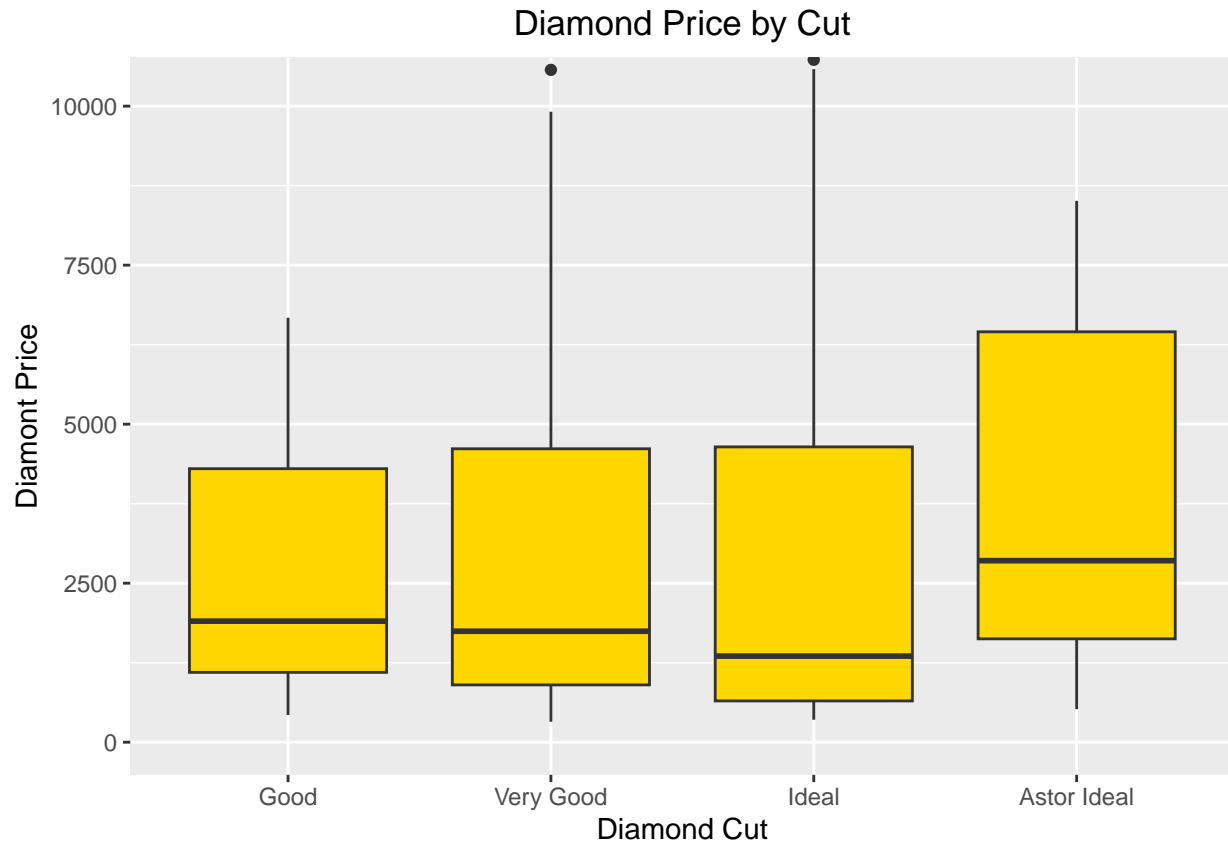
When viewing the boxed ranges, the values are surprisingly close between "good", "very good", and "ideal". "Astor ideal" has a pretty wide range and trends more expensive in the 50% range, but the sample size is very low at only 20 diamonds, and it shares more than half of its range with very good and ideal cuts.
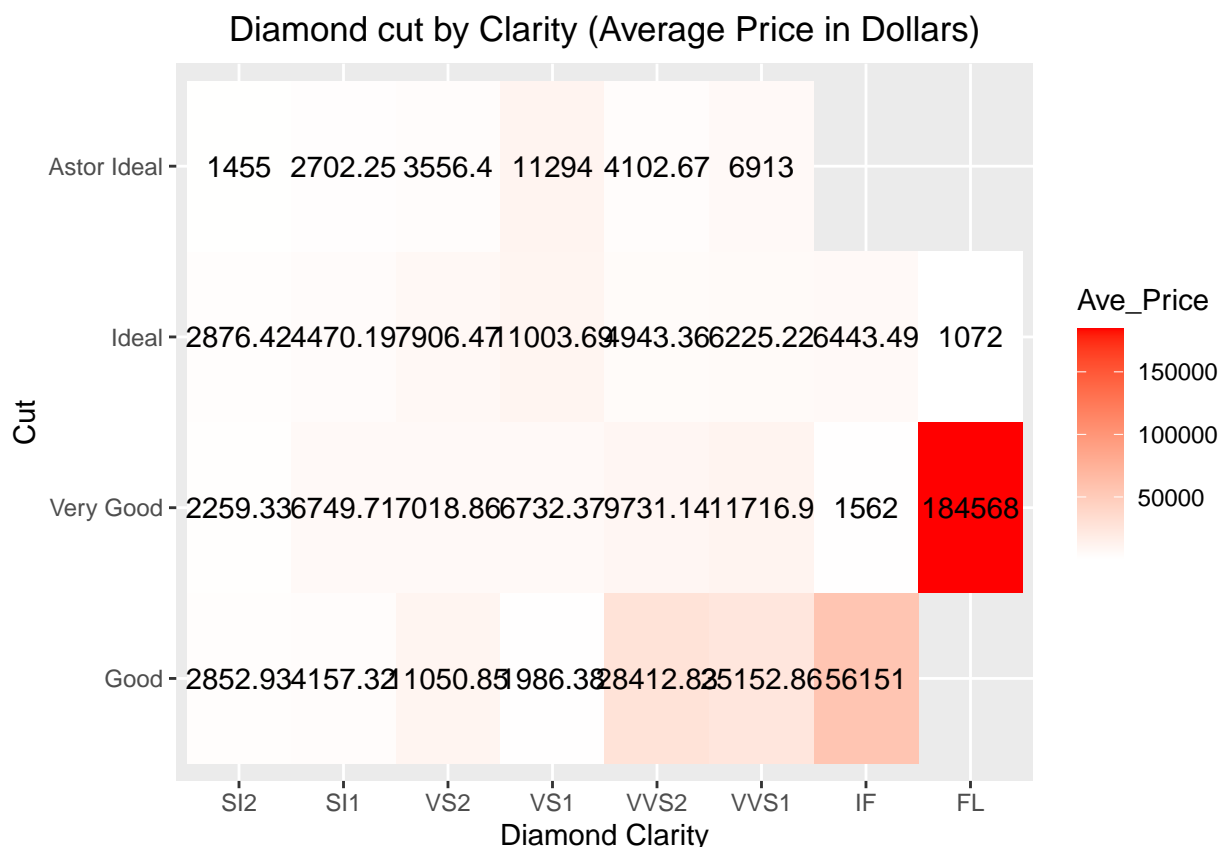
The quantile data reinforces the similarity in the q75(the value in which 75% of the data lies under) for all cuts.

Bivariate Analysis

Cut and Clarity

Diamond cut by clarity (Quantity)

Cut	SI2	SI1	VS2	VS1	VVS2	VVS1	IF	FL
Astor Ideal	5	20	25	30	15	5		
Ideal	11.64	20.3	17.73	18.4	13.4	13.4	5.01	0.14
Very Good	16.75	18.32	17.02	20.42	13.09	10.99	2.88	0.52
Good	19.18	26.03	17.81	17.81	8.22	9.59	1.37	

Diamond Clarity

Percentage: 30, 20, 10

A majority of the diamond stock for all cuts are in the lower quality half of the clarities carried (SI2-VS1). "Very good" and "Ideal" diamonds, which are the cuts of the highest quantities, have more than a quarter of their stock in the higher quality half of the clarities available (VVS2-FL). Additionally, "Very good" and "Ideal" cuts are the only cuts with "Flawless" diamonds.

Diamond cut by Clarity (Average Price in Dollars)

Cut	SI2	SI1	VS2	VS1	VVS2	VVS1	IF	FL
Astor Ideal	1455	2702.25	3556.4	11294	4102.67	6913		
Ideal	2876.42	4470.19	7906.47	11003.69	4943.36	6225.22	6443.49	1072
Very Good	2259.33	6749.71	7018.86	6732.37	9731.14	11716.9	1562	184568
Good	2852.93	4157.32	11050.85	1986.38	28412.82	35152.86	56151	

(Heatmap legend: Ave_Price ranging from 0 to over 150000, with markers at 50000, 100000, 150000)
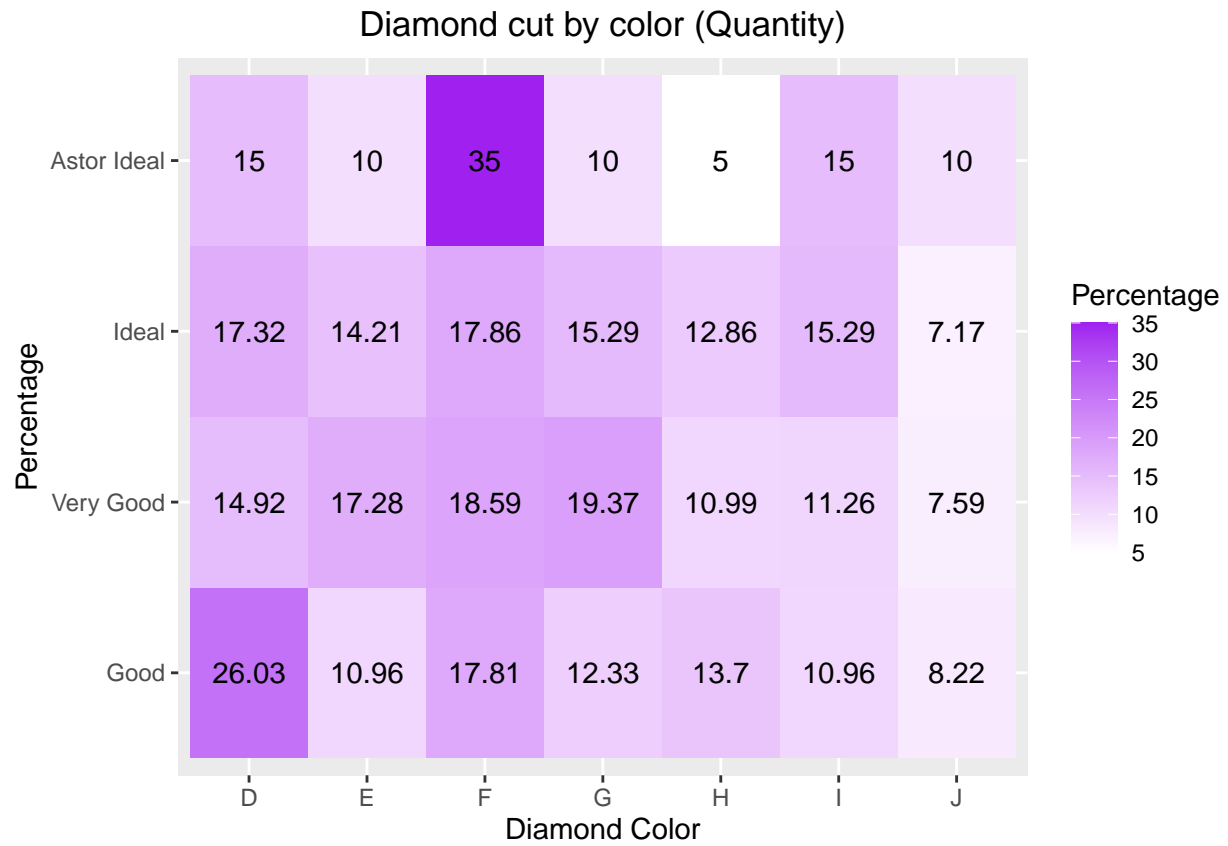
As clarity increases in the "good" cut category, price typically increases per clarity category (SI2 and SI1 combined, VVS2 and VVS1 combined). The average price nearly doubles for most instances of increasing clarity category. An anomaly in VS1 can be described by a majority of the diamonds in this category being of low carat size (under 1 carat) and of comparatively poor color (G,H,I, and J). There is only one "good" cut with "IF" clarity. This diamond is 3.35 carats and worth $56151.

The "very good" category also increases as the category of clarity increases, though SI1 through VS1 carry a fairly similar average price. The "IF" clarity diamonds that were also of the "very good" cut were all under 1 carat in weight and nearly two-thirds of the diamonds had a color of G or below.
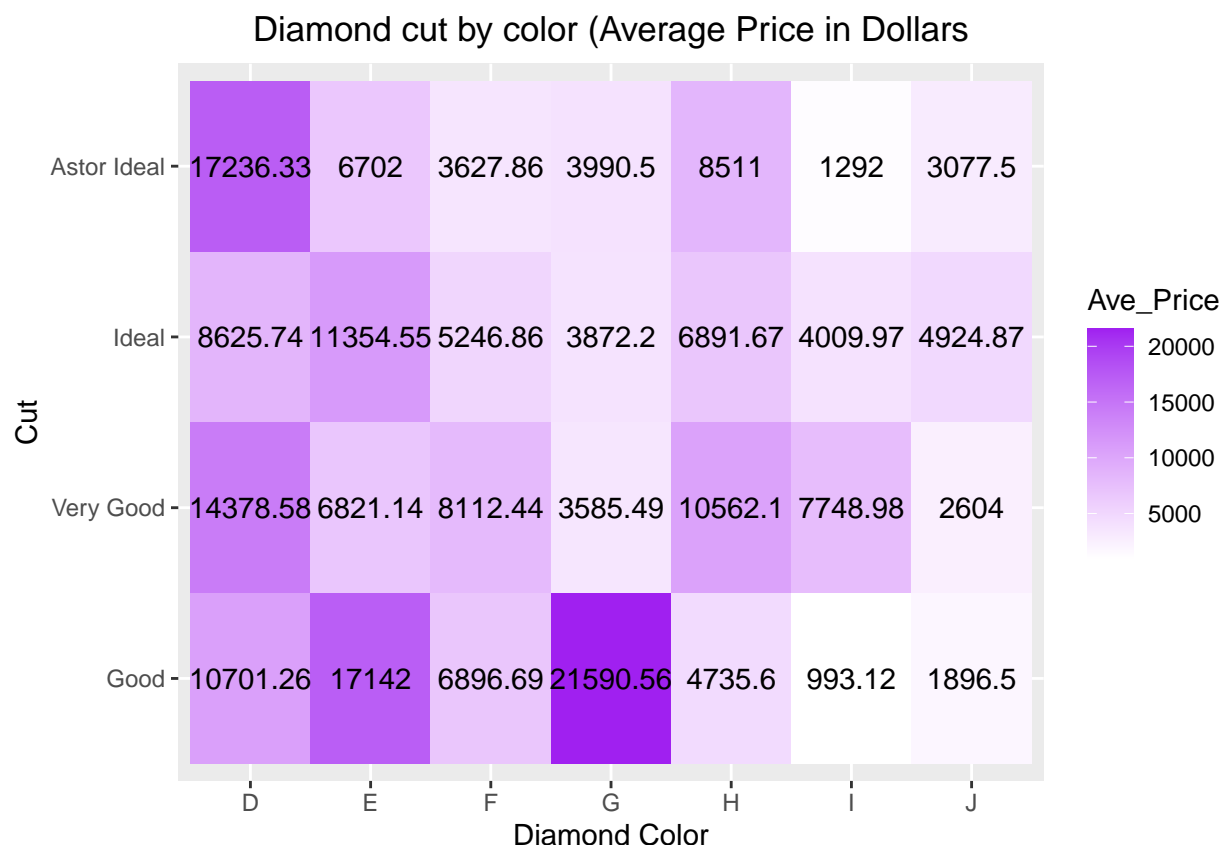
The "Ideal" category of cut increases from the SI category into the VS category, then drops in price for the VVS and If-FL categories. There are nearly 100 diamonds in each of the VVS category sections. A majority of these diamonds (three-quarters or more for each category section) were less than 1 carat in weight and more than half of the diamonds in each VVS section were color "G" or below.

The "Astor Ideal" category increases in average price as clarity increases, similarly to the other categories, with some oscillation between VS1 and VVS1 categories. Since there are only 20 "Astor Ideal" cut diamonds, the small sample size may contribute to some slightly unexpected fluctuations at the higher clarities. For example, the price for VSS1 represents the only diamond with those specifications. Additionally, VS1 has a comparatively high average due to a 2 carat diamond priced at $46,893.

Cut and Color

Diamond cut by color (Quantity)

Percentage	D	E	F	G	H	I	J
Astor Ideal	15	10	35	10	5	15	10
Ideal	17.32	14.21	17.86	15.29	12.86	15.29	7.17
Very Good	14.92	17.28	18.59	19.37	10.99	11.26	7.59
Good	26.03	10.96	17.81	12.33	13.7	10.96	8.22

Diamond Color

Percentage: 35, 30, 25, 20, 15, 10, 5

The distribution of diamond color per cut appears to be a bit more widely distributed than for clarity per cut. Inversely to clarity and cut, approximately half of all of the diamonds for each cut category reside in the most premium colors (D,E,and F). The other half are distributed between the G,H,I, and J colors.

Diamond cut by color (Average Price in Dollars

Cut	D	E	F	G	H	I	J
Astor Ideal	17236.33	6702	3627.86	3990.5	8511	1292	3077.5
Ideal	8625.74	11354.55	5246.86	3872.2	6891.67	4009.97	4924.87
Very Good	14378.58	6821.14	8112.44	3585.49	10562.1	7748.98	2604
Good	10701.26	17142	6896.69	21590.56	4735.6	993.12	1896.5

Diamond Color

Ave_Price

20000
15000
10000
5000

The average value of diamonds with a "good" quality cut decreases as the color decreases. D,E, and F are considered the most premium colors, so it is not strange that "E" is a bit higher than "D." Additionally, "E" has an ~4 carat diamond worth $109,744 in its set of 8 diamonds. Good cuts with G color have 9 diamonds associated with the set. There is a large, 6 carat outlier worth $165,766. This inflates the price dramatically.
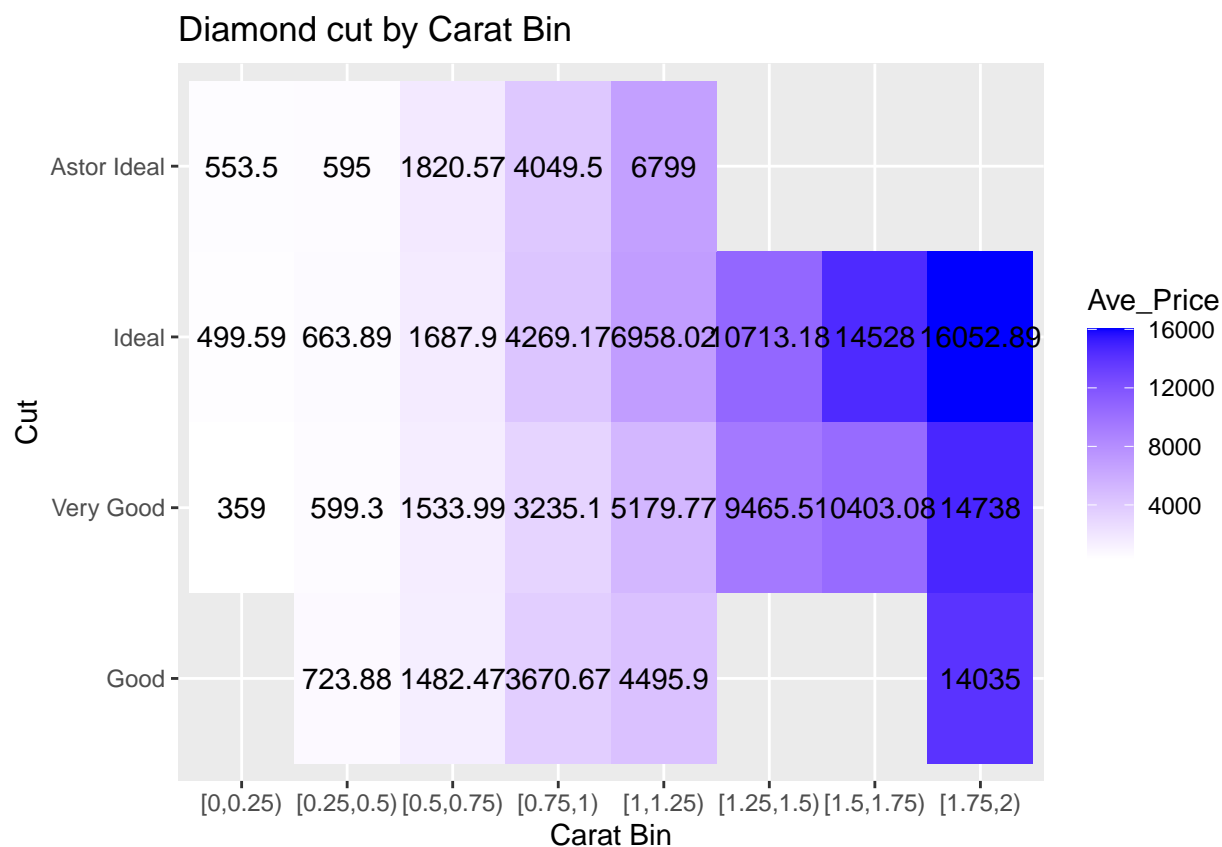
The "Very Good" section is not very price predictable by color. The other sections above follow this same trend. While with clarity, one or two combinations of color for each cut could be analyzed to mostly explain deviations, it would be nearly every combination of color within a cut that would have to be analyzed to explain why price does not neatly descend with decreasing color quality.

Cut and Carat

Diamond cut by Carat
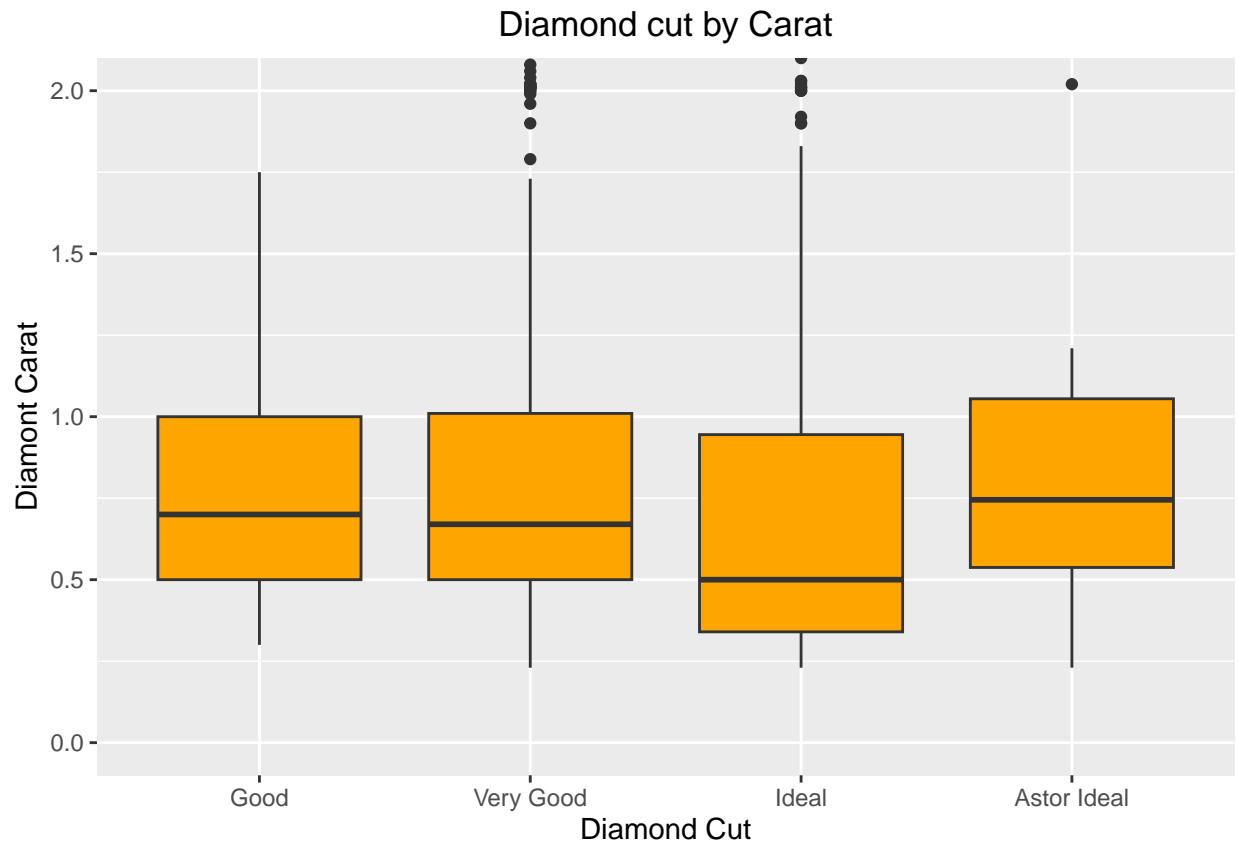
The graph of diamond carat against diamond cut had a great deal of outliers, which presented similarly to the diamond price by cut graph in the univariate section. Therefore, this assessment utilizes the truncated graph, containing all values that are not considered outliers.

For a majority of the cuts ("Good","Very good", and "Astor Ideal"), 50% of the diamonds fall within the range of 0.5 carats to 1 carat, or just slightly over. For the "Ideal" cut range, this is slightly lowered at a range of 0.3 carats to just under 1 carat.

Diamond cut by Carat Bin

Cut	[0,0.25)	[0.25,0.5)	[0.5,0.75)	[0.75,1)	[1,1.25)	[1.25,1.5)	[1.5,1.75)	[1.75,2)
Astor Ideal	553.5	595	1820.57	4049.5	6799			
Ideal	499.59	663.89	1687.9	4269.17	6958.02	10713.18	14528	16052.89
Very Good	359	599.3	1533.99	3235.1	5179.77	9465.5	10403.08	14738
Good		723.88	1482.47	3670.67	4495.9			14035

Ave_Price
16000
12000
8000
4000

The relationship with cut and carat is quite apparent with the heat map. As the carat bin size increases within all cut categories, the price increases. This price increase is gradual and sequential. However, when you look through the columns, analyzing the cuts per each bin, the pricing is usually somewhat similar, with "Ideal" and "Astor Ideal" typically being worth a bit more than "good" and "very good." This bin range of up to 2 carats was selected because this represents approximately 93% of the diamond data. Some evidence against cut being the most important of the 4 Cs is that in almost all cases, moving horizontally to the next 0.25 carat wide bin size raises the average price more than moving vertically to the most expensive premium cut in the column. That evidence is explored more in the graphics below.

Diamond Price by Cut
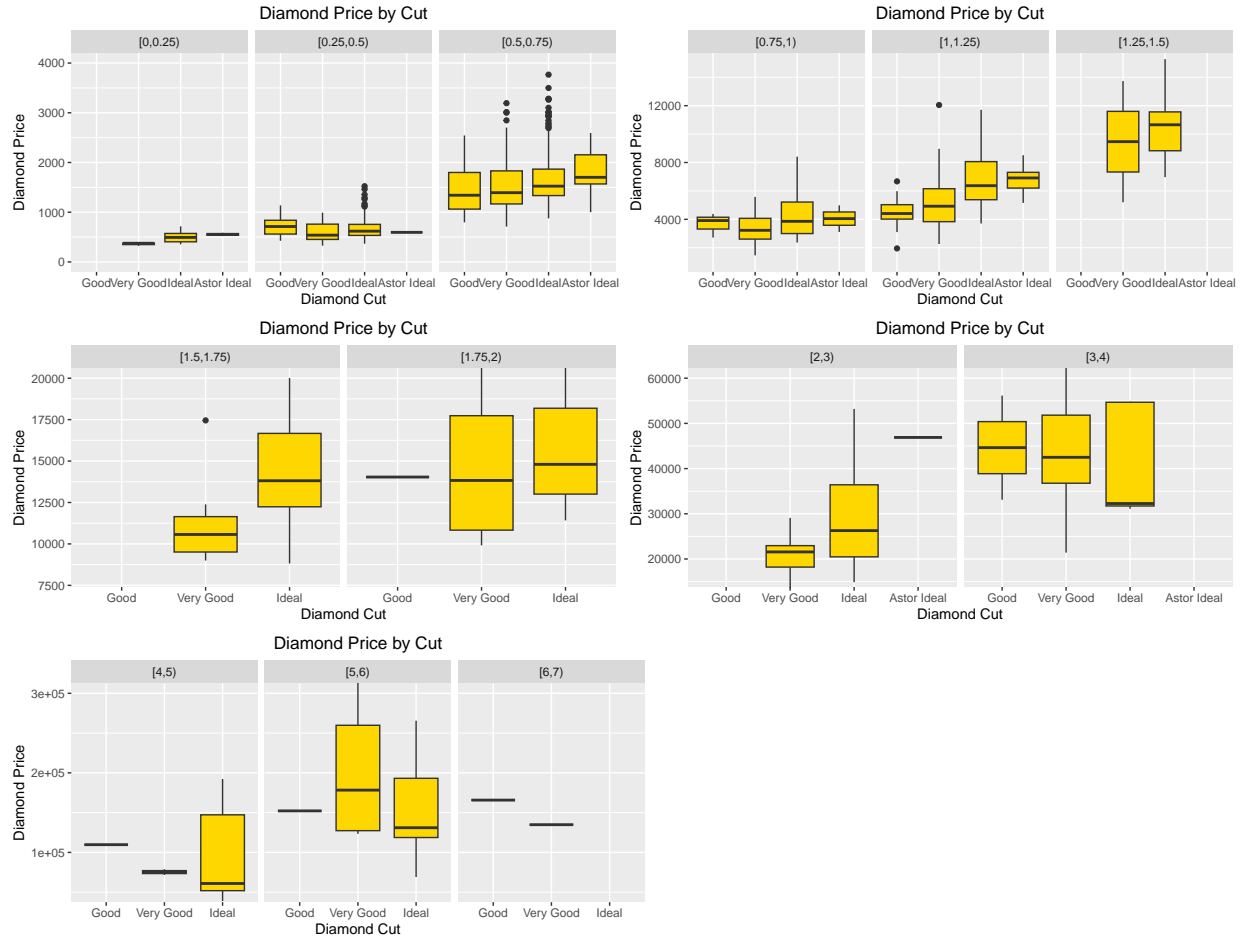
At the 0-0.25 carat level, the change between the q75 price between the lowest value cut (very good), and the highest value cut (Astor Ideal), is (~562(Astor Ideal)-375(Very good)) for a total of $187.

At the 0.25-0.5 carat range, we're omitting Astor ideal for lack of product. The change between the q75 price of the lowest value cut in this range (Ideal) and the highest value cut in this range (Good), is ($_{812(Good)}$-750(Ideal)) for a total of $62.

At the 0.5-0.75 carat range, the change between the q75 price of the lowest value cut in this range (Good) and the highest value cut in this range (Astor Ideal), is (~2200(Astor Ideal)~1750(Good)) for a total of $450.

At the 0.75-1 carat range, the change for the q75 price between the lowest value cut (Very Good) and the highest value cut (Ideal) is ($_{5320(Ideal)}$-4000(Very Good)) for a total of $1320.

At the 1-1.25 carat range, the change for the q75 price between the lowest value cut (Good) and the highest value cut (Ideal) is ($_{8000(Ideal)}$-5000(Good)) for a total of $3000.

At the 1.25-1.5 carat range, the change for the q75 price between the lowest value cut and the highest value cut (between Ideal and Very good), is negligible.

At the 1.5-1.75 carat range, the change for the q75 price between the lowest value cut (Very Good) and the highest value cut (Ideal) is ($_{16600(Ideal)}$-11600(Very Good)) for a total of $5000.

At the 1.75-2 carat range, the change for the q75 price between the lowest value cut (Very Good) and the highest value cut (Ideal) is ($_{18125(Ideal)}$-17800(Very Good)) for a total of $325.

Diamond Carat Bin
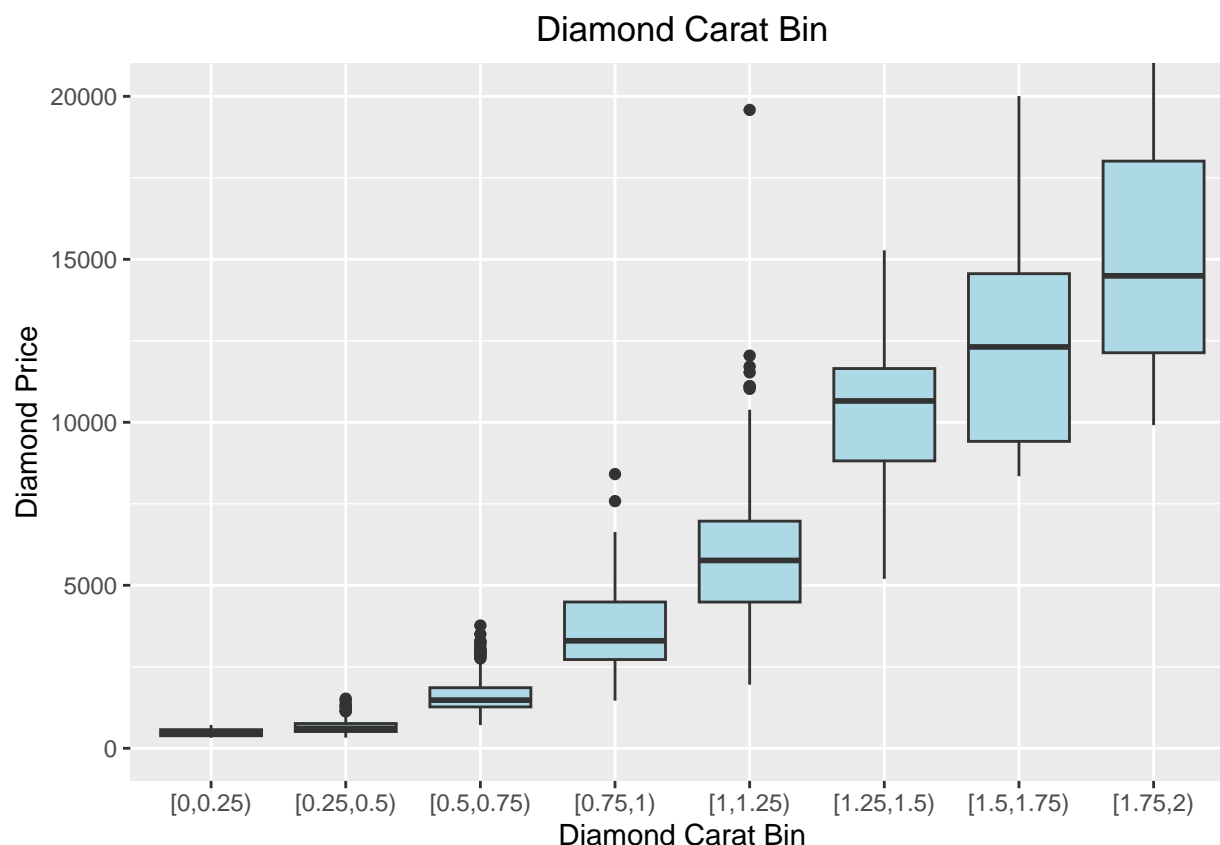
The difference between the q75 for 0-0.25 carat and q75 for 0.25-0.5 carat is approximately $150. The difference between q75 for 0.25-0.5 carat and 0.5-0.75 carat ranges is ~$1,125. The difference between q75 of 0.5-0.75 carat and q75 of 0.75-1 carat is ~$2,625. The difference between the q75 of the 0.75-1 carat and 1-1.25 carat range is ~$2,500. The difference between the q75 of the 1-1.25 carat range and 1.25-1.5 carat range is ~$4,750. The difference between the q75 of the 1.25-1.5 range and the 1.75-2 carat range is ~$6,250. This carat range includes ~93% of the data. Prices were compared by analyzing the internal change in price based on cut compared to the external change in price based on carat. This was done by zooming in graphs where required and analyzing the difference in price using the q75 mark, which constitutes the value under which 75% of the data falls and is depicted by the top horizontal line of each boxplot.

For the internal variation, the "most expensive" q75 mark for cut was read, typically "Ideal," which had subtracted from it the "least expensive" q75 mark for cut, typically "very good" or "good." This value represents the price increase between the q75 marks of the "most" and "least" expensive cut in each carat grouping. For carat grouping, the q75 mark of the next grouping is subtracted from the q75 marking of the former grouping. This value represents the increase in q75 value when raising the carat into the next 0.25 carat wide grouping.

In a majority of cases, increasing the carat size into the next 0.25 carat wide grouping raises the price more than increasing the quality within the 0.25 carat wide grouping. An additional note is that on rare occasion, the most expensive cut q75 in a range will not be the highest quality cut. For example, at the 0.25-0.5 carat range, the good diamond q75 was the most expensive. When increasing the quality from "very good" to ideal" in the 1.5-1.75 carat range, the q75 price increases by $5,000, which is the highest internal increase attributed to changing cut. However, when increasing the carat range from 1.5-1.75 to the next 0.25 carat wide grouping of 1.75-2, the q75 value increases by approximately $6,250. Based on the data in these graphs, carat may be more important than cut for price. This is additional evidence against the claim that cut is the most important variable in price.