# 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

Cut univariate and Cut with Price

Diamond Cut Counts by Quality
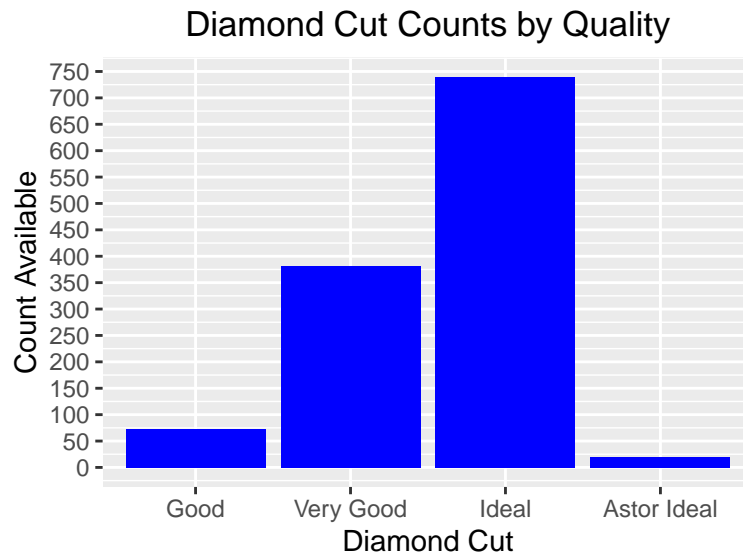
The cut is highly skewed regarding available quantity of each type. There are vastly more ideal cuts (739) and very good cuts (382) compared to the good cuts (73) and the Astor ideal cuts (20).

Table 1: **Cut and Price Quantiles**

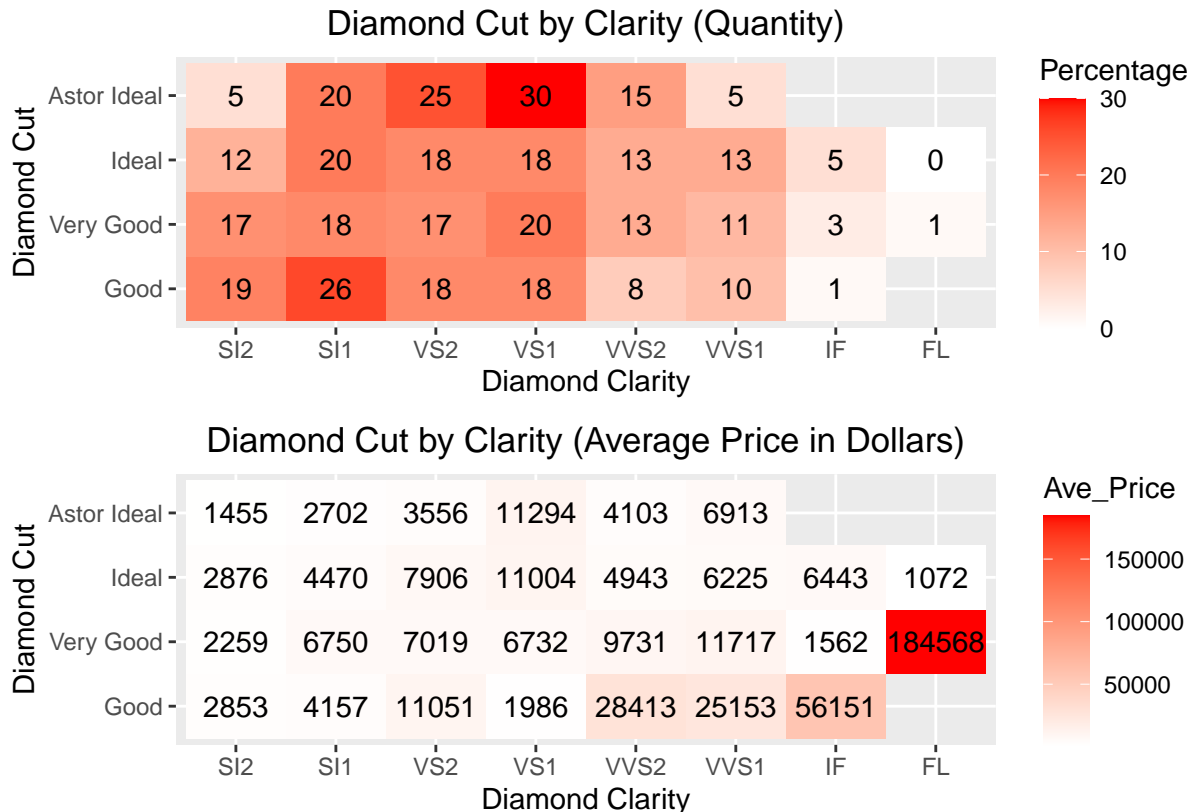| cut	q0	q25	q50	q75	q100	mean
Good	426	1098.00	1903.0	4300.0	165766	9467.288
Very Good	322	901.00	1744.5	4613.0	355403	7757.626
Ideal	354	649.00	1354.0	4643.0	345397	6488.932
Astor Ideal	520	1625.25	2854.0	6452.5	46893	5851.550

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.

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.

Cut and Clarity

Diamond Cut by Clarity (Quantity)

Diamond Cut	SI2	SI1	VS2	VS1	VVS2	VVS1	IF	FL
Astor Ideal	5	20	25	30	15	5		
Ideal	12	20	18	18	13	13	5	0
Very Good	17	18	17	20	13	11	3	1
Good	19	26	18	18	8	10	1	

Percentage: 30 — 20 — 10 — 0

Diamond Clarity

Diamond Cut by Clarity (Average Price in Dollars)

Diamond Cut	SI2	SI1	VS2	VS1	VVS2	VVS1	IF	FL
Astor Ideal	1455	2702	3556	11294	4103	6913		
Ideal	2876	4470	7906	11004	4943	6225	6443	1072
Very Good	2259	6750	7019	6732	9731	11717	1562	184568
Good	2853	4157	11051	1986	28413	25153	56151	

Ave_Price: 150000 — 100000 — 50000

Diamond Clarity

A majority of the diamond stock for all cuts are in the lower quality half of the clarity categories 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.

As clarity increases in the "good" cut category, price typically increases per clarity category (SI2 and SI1 grouped, VVS2 and VVS1 grouped, etc.). 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)

Diamond Cut	D	E	F	G	H	I	J
Astor Ideal	15	10	35	10	5	15	10
Ideal	17	14	18	15	13	15	7
Very Good	15	17	19	19	11	11	8
Good	26	11	18	12	14	11	8

Diamond Color — Percentage

Diamond Cut by Color (Average Price in Dollars)

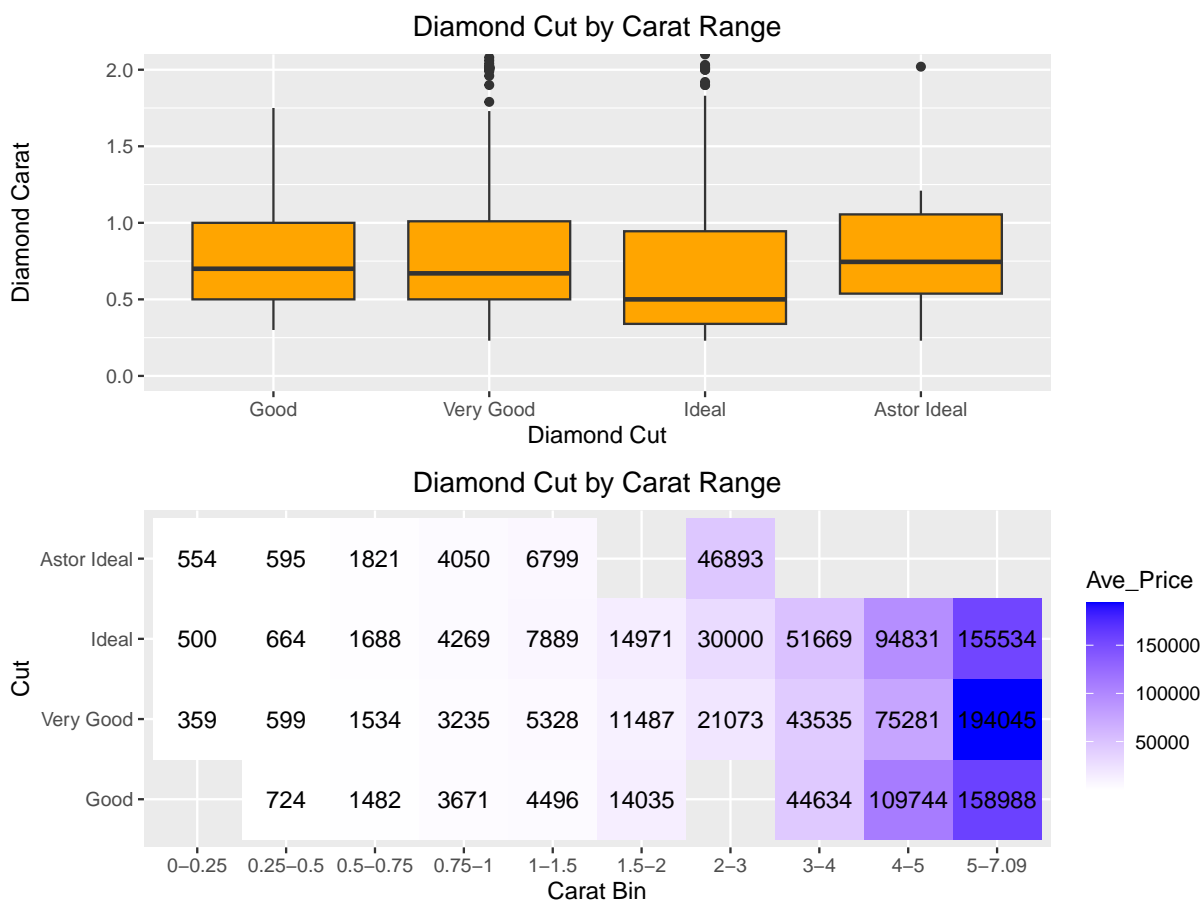| Diamond Cut	D	E	F	G	H	I	J
Astor Ideal	17236	6702	3628	3990	8511	1292	3078
Ideal	8626	11355	5247	3872	6892	4010	4925
Very Good	14379	6821	8112	3585	10562	7749	2604
Good	10701	17142	6897	21591	4736	993	1896

Diamond Color — Ave_Price

The distribution of diamond color per cut appears to be a bit more widely varied 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.

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, most increases across cuts are consistent, with one or two deviations per cut being explainable by carat, or a combination of carat and color, it would be nearly every category 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 Range

Diamond Cut by Carat Range

Cut	0–0.25	0.25–0.5	0.5–0.75	0.75–1	1–1.5	1.5–2	2–3	3–4	4–5	5–7.09
Astor Ideal	554	595	1821	4050	6799		46893			
Ideal	500	664	1688	4269	7889	14971	30000	51669	94831	155534
Very Good	359	599	1534	3235	5328	11487	21073	43535	75281	194045
Good		724	1482	3671	4496	14035		44634	109744	158988

Ave_Price: 150000 / 100000 / 50000

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 previous 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.

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." 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
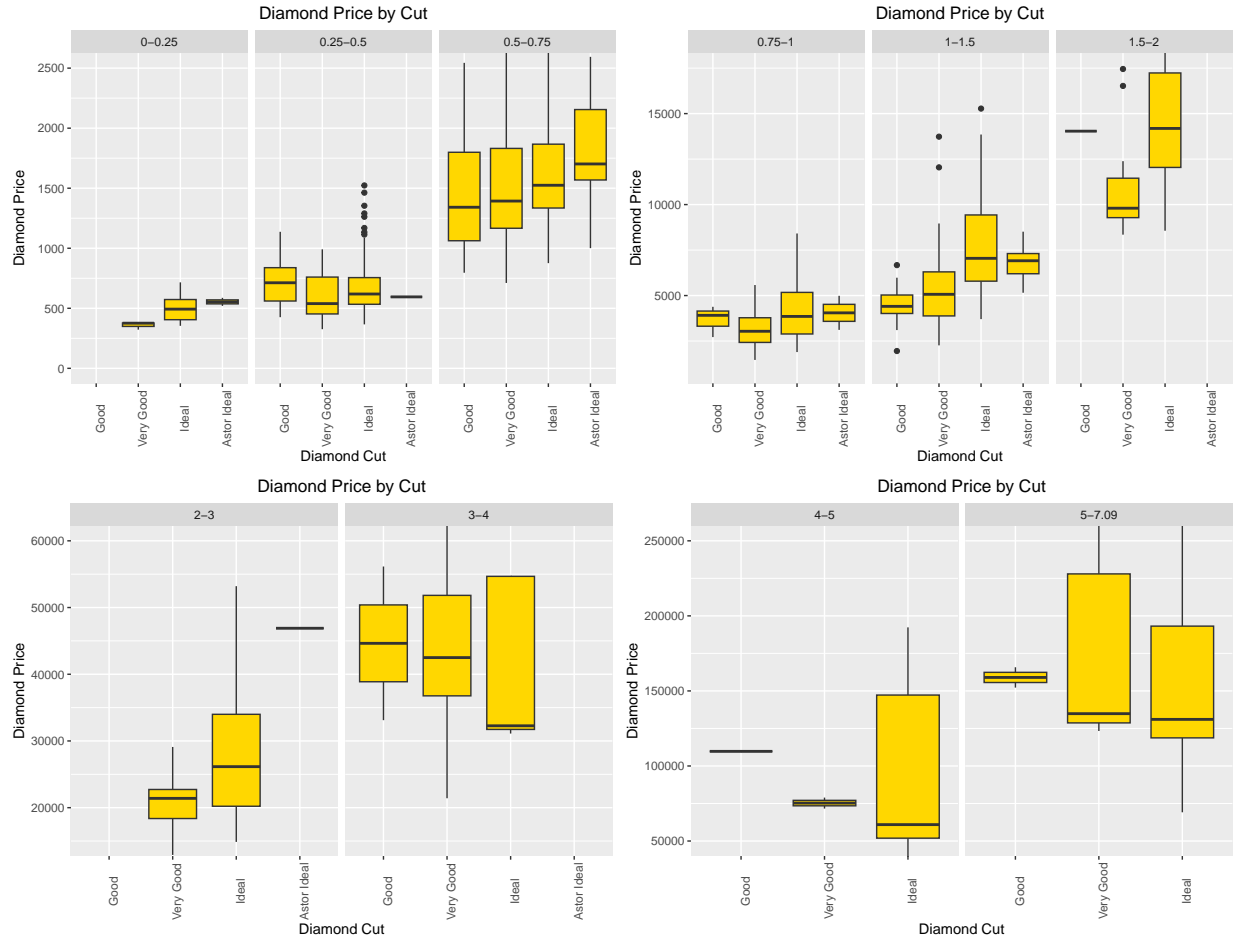
The q75 mark was selected for comparison, as this represents the value in which 75% of the data in each category falls under. The table depicts the cuts in each bin with the highest q75 price and the lowest q75 price. The lowest q75 mark is subtracted from the highest q75 mark and this constitutes the difference in q75 price between the "cheapest" cut in the range and the "most expensive" cut in the range. This represents the inner difference, the difference between the q75 value of cuts within the same carat range. This can be compared to the outer difference, the difference in q75 price when a bin increases to the next bin size. Due to both the decrease in data at upper carat ranges, and due to the way Blue Nile structures their diamond sizing guide, the bins change to increasing by half a carat size between 1 and 2 carats, then 1 carat in size after the 2 carat mark. This is metric was used for both inner and outer change graphs/tables.

Table 2: **Minimum and Maximum 75th Quantile Analysis per Carat Bin (Inner change)**

cut	carat_bin	Minimum_q75	Maximum_q75	q75_range
Very Good	0-0.25	377.50	NA	195.50
Ideal	0-0.25	NA	573.00	195.50
Astor Ideal	0.25-0.5	595.00	NA	243.25
Good	0.25-0.5	NA	838.25	243.25
Good	0.5-0.75	1799.25	NA	355.75
Astor Ideal	0.5-0.75	NA	2155.00	355.75
Very Good	0.75-1	3780.50	NA	1393.00
Ideal	0.75-1	NA	5173.50	1393.00
Good	1-1.5	5030.00	NA	4400.00
Ideal	1-1.5	NA	9430.00	4400.00
Very Good	1.5-2	11450.00	NA	5784.50
Ideal	1.5-2	NA	17234.50	5784.50
Very Good	2-3	22728.00	NA	24165.00
Astor Ideal	2-3	NA	46893.00	24165.00
Good	3-4	50392.25	NA	4281.75
Ideal	3-4	NA	54674.00	4281.75
Very Good	4-5	77107.50	NA	70136.50
Ideal	4-5	NA	147244.00	70136.50
Good	5-7.09	162377.00	NA	65583.00
Very Good	5-7.09	NA	227960.00	65583.00

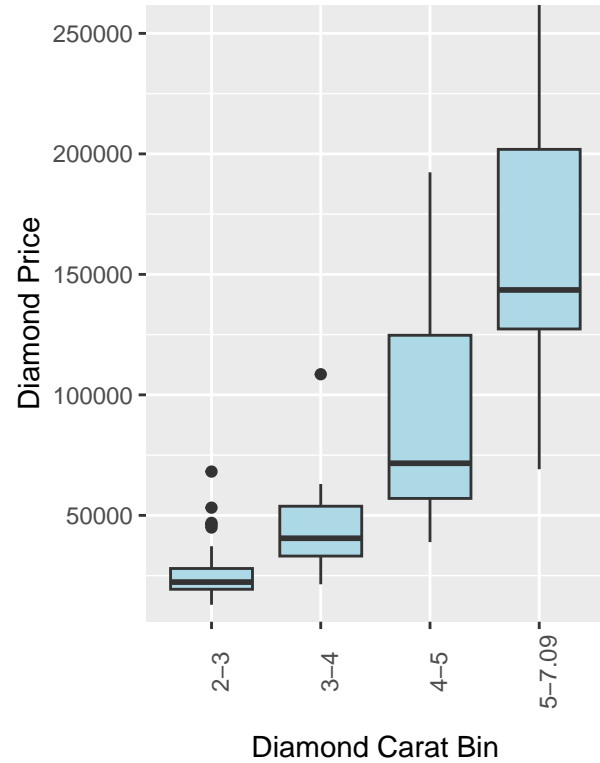Price change between Carat Bins
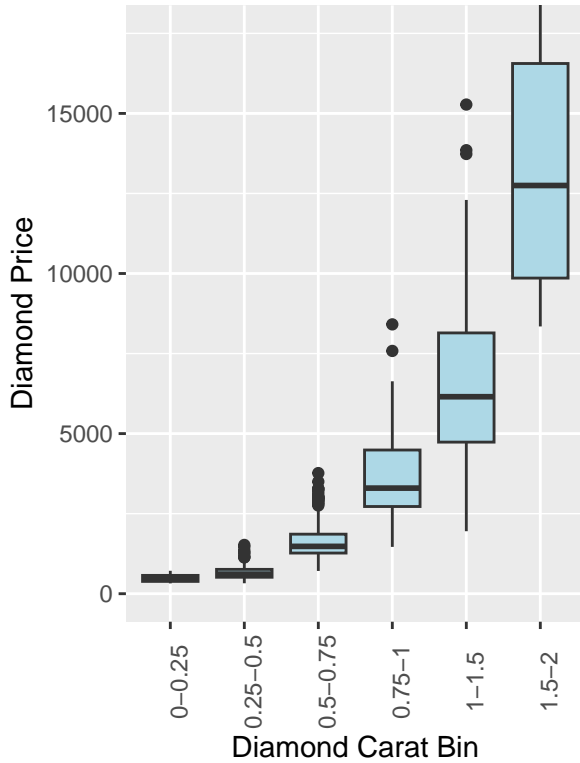
Price change between Carat Bins

Table 3: **Minimum and Maximum 75th Quantile Analysis per Carat Bin (Outer change)**

carat_bin	q75	q75_Change
0-0.25	571.50	NA
0.25-0.5	760.00	188.50
0.5-0.75	1859.25	1099.25
0.75-1	4488.00	2628.75
1-1.5	8144.75	3656.75
1.5-2	16558.75	8414.00
2-3	27963.00	11404.25
3-4	53794.00	25831.00
4-5	124735.50	70941.50
5-7.09	201864.50	77129.00

When comparing the two tables, based on difference in q75 values, the price always increases more by increasing the carat size (outer change), regardless of cut, compared to increasing the quality of the cut within a carat bin (inner change). An exception is the final line of the first table. This is because this is the largest carat size offered, so there are no further values to compare it to in the second table.