Elizabeth Evans

MSIS 510, Autumn 2020 Purple

Assignment 1

1. Categorical: cut, color, clarity; Numerical: price, carat, x(length), y(width), z(depth), depth, table

Chart, scatter chart

Description automatically generated

1. The clustering along vertical lines indicates that 1) many of the observations over 1 carat are of similar weight and, more importantly, 2) there are additional factors besides carat that play into the price of the diamonds. As per the plots below, it would appear that clarity is the strongest determining factor for a price at a particular carat, followed by color than cut.

Chart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generated

Chart, scatter chart

Description automatically generated

1. Bar chart of average price per cut:Chart, bar chart

   Description automatically generated
2. Bar chart of count of observations per cut:Chart, bar chart

   Description automatically generated
3. Bar chart of average carat per cut:Chart, bar chart

   Description automatically generated
4. Based on the bar charts, there seems to be less observations in the lower quality cuts, particularly in “Fair”, than in “Ideal” with a higher average carat which would skew the average price based on cut. Additionally, as per the above scatter plot, clarity and color appear to be stronger predictors of price than cut.