THE GEORGE WASHINGTON UNIVERSITY

WASHINGTON, DC

In this LAB, you will learn how to visualize complex data using Tableau software. All the questions in this LAB must be answered using Tableau software. The dataset for this LAB is 'diamond.csv'.

1. Using the Tableau software connect to the 'diamond' dataset and display the mean price versus cut (all cuts) and clarity (all clarities) and color(all colors). Save the table under the tab named "Mean price versus cut-color-clarity". Add an appropriate title to your table.

Average Price versus color-cut-clarity

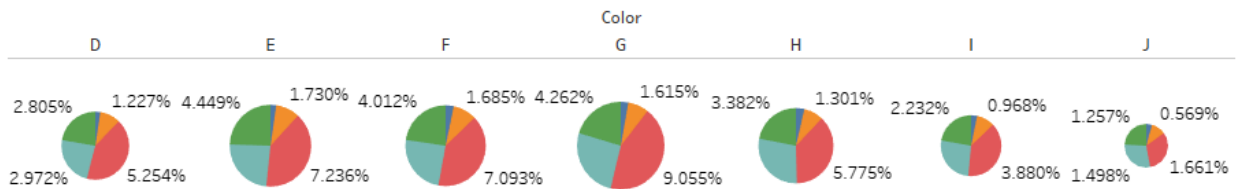| Clarity	Cut	Color						
		D	E	F	G	H	I	J
I1	Fair	7,383	2,095	2,544	3,187	4,213	3,501	5,795
	Good	3,491	4,398	2,570	3,196	3,850	4,175	3,795
	Ideal	3,527	3,559	3,903	4,044	5,415	4,103	9,454
	Premium	3,819	3,199	3,555	4,052	3,904	5,045	4,577
	Very Good	2,623	3,444	4,253	3,195	5,259	6,045	4,478
IF	Fair	1,620		2,344	1,488			
	Good	10,030	1,519	3,133	4,060	5,949	1,749	2,738
	Ideal	6,567	3,259	2,154	2,206	1,983	1,503	2,489
	Premium	9,057	4,525	3,618	3,311	3,385	2,359	7,026
	Very Good	10,298	4,333	4,677	3,525	2,648	4,094	1,074

2. Display the following table that shows the Color, carat(avg), Depth(avg), Price(avg), Table (avg), X(avg), Y(avg), and Z (avg). Makes sure to change the precision to 2 digit decimals. Makes sure to change the price format to $.

Key Perfromance Indicator

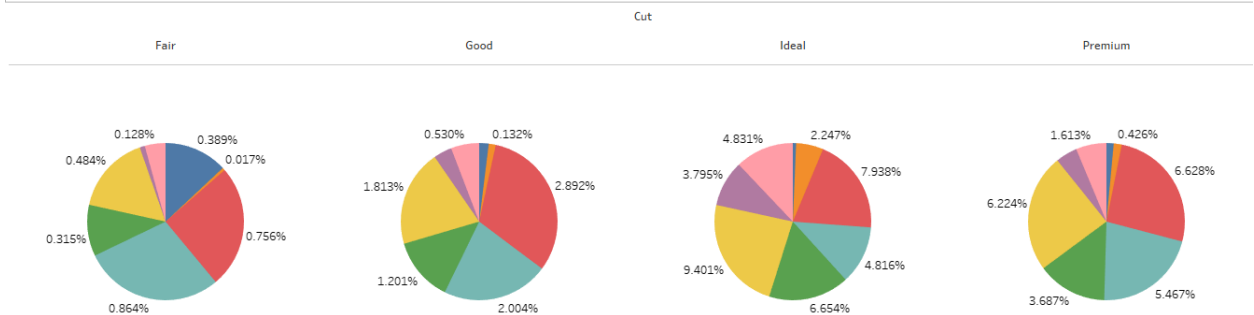| Color	Avg. Carat	Count of diamonds.csv	Avg. Depth	Avg. Price	Table	Avg. X	Avg. Y	Avg. Z
D	0.658	6,775	61.70	$3,169.95	57.40	5.42	5.42	3.34
E	0.658	9,797	61.66	$3,076.75	57.49	5.41	5.42	3.34
F	0.737	9,542	61.69	$3,724.89	57.43	5.61	5.62	3.46
G	0.771	11,292	61.76	$3,999.14	57.29	5.68	5.68	3.51
H	0.912	8,304	61.84	$4,486.67	57.52	5.98	5.98	3.70
I	1.027	5,422	61.85	$5,091.87	57.58	6.22	6.22	3.85
J	1.162	2,808	61.89	$5,323.82	57.81	6.52	6.52	4.03

3. Plot the pie chart for number of sales in percentage per cut and color. Save the code for this part under the tab named pie chart in % per color and cut. Add an appropriate title to your graph.

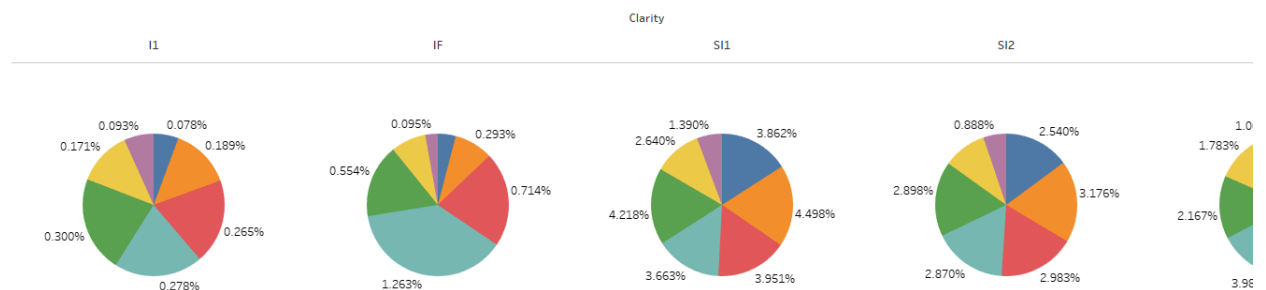Pie Chart Sales Count per color and cut

Color

| D | E | F | G | H | I | J |

2.805%   1.227% 4.449%   1.730% 4.012%   1.685% 4.262%   1.615% 3.382%   1.301% 2.232%   0.968% 1.257%   0.569%

2.972%   5.254%   7.236%   7.093%   9.055%   5.775%   3.880% 1.498%   1.661%

4. Plot the pie chart for number of sales in percentage per cut and clarity. Save the code for this part under the tab named pie chart in % per cut and clarity. Add an appropriate title to your graph.

Pie Chart percentage sales count per cut and clairty

Cut

| Fair | Good | Ideal | Premium |

Fair
0.128% 0.389%
0.484% 0.017%
0.315% 0.756%
1.201%
0.864%

Good
0.530% 0.132%
1.813% 2.892%
2.004%

Ideal
4.831% 2.247%
3.795% 7.938%
9.401% 4.816%
6.654%

Premium
1.613% 0.426%
6.628%
6.224%
3.687% 5.467%

5. Plot the pie chart for number of sales in percentage per color and clarity. Save the code for this part under the tab named pie chart in % per color and clarity. Add an appropriate title to your graph.

Pie Chart Sales count percentage per color and clarity

Clarity

| I1 | IF | SI1 | SI2 |

I1
0.093% 0.078%
0.171% 0.189%
0.300% 0.265%
0.278%

IF
0.095% 0.293%
0.554% 0.714%
1.263%

SI1
1.390% 3.862%
2.640%
4.218% 4.498%
3.663% 3.951%

SI2
0.888% 2.540%
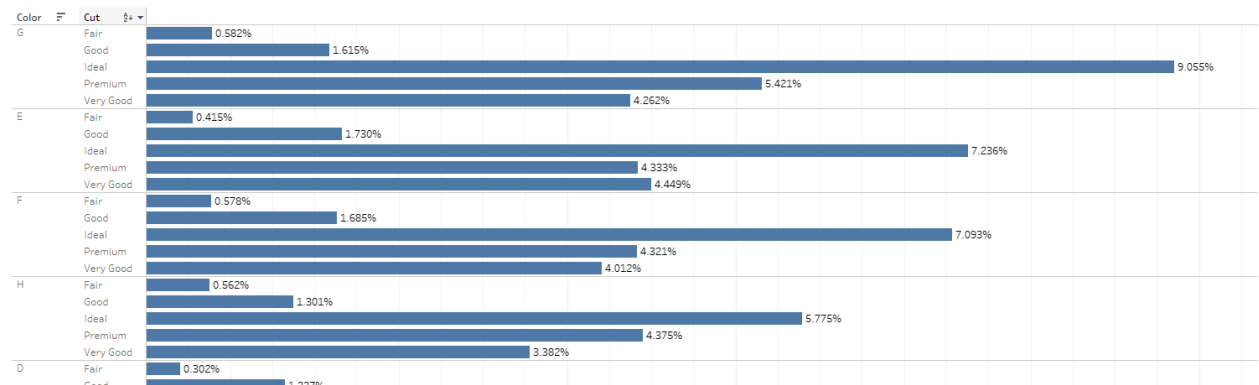2.898% 3.176%
2.870% 2.983%

1.0
1.783%
2.167%
3.98

6. Plot the horizontal bar plot (sorted descending order) of the number of diamond sales per color. Display the number of sales in %. Save the code for this plot under the tab 'Sales count-Barplot per color'. Add an appropriate title to your graph.

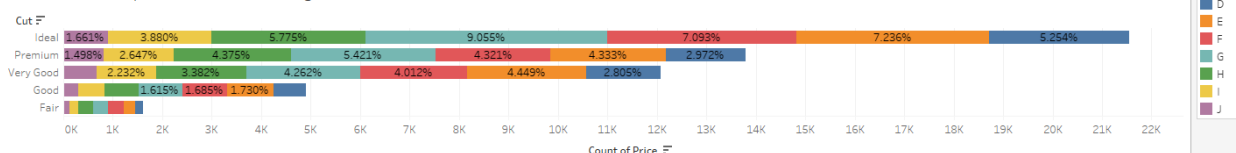Number of Sales per color and Percentage
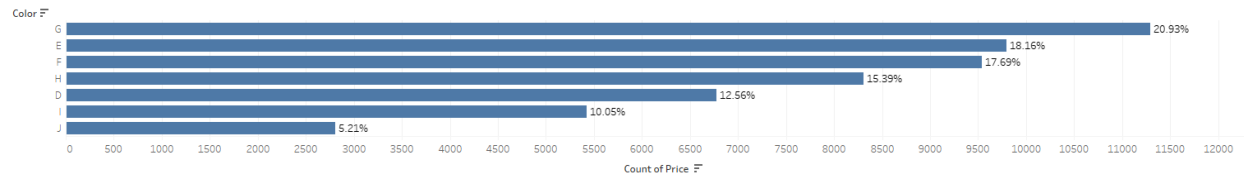
7. Plot the horizontal bar plot of the number of diamond sales per color and cut. Display the number of sales in %. Save the code for this plot under the tab 'Sales count-Barplot per color and cut'. Add an appropriate title to your graph.

Sales count -Barplot per color and cut

8. Plot the horizontal bar plot of the number of diamond sales per cut and color (legend) . Display the number of sales in %. Use the color as the filter. Save the code for this plot under the tab 'Sales count-Barplot per color and cut'. Add an appropriate title to your graph.

Number of Sales per cut and Percentage- Bar Plot

9. Plot the horizontal bar plot of the number of diamond sales per color (legend) and clarity and color. Display the number of sales in %. Use the color as the filter. Save the code for this plot under the tab 'Sales count-Barplot per color, cut and clarity'. Add an appropriate title to your graph.
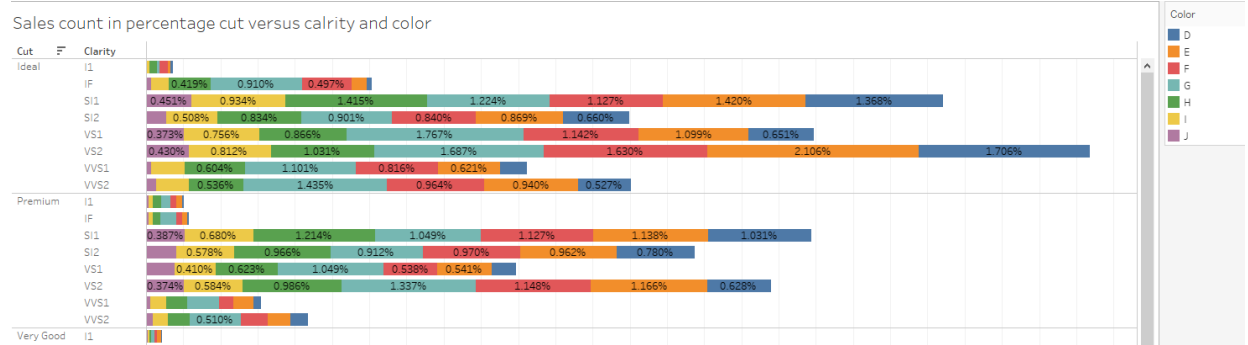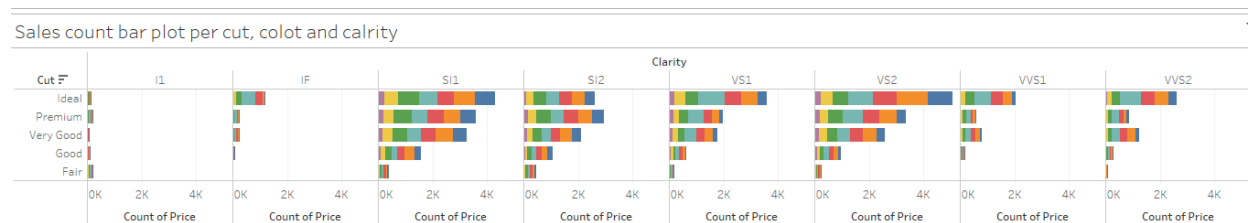
Sales count in percentage cut versus calrity and color

Color: D, E, F, G, H, I, J

| Cut	Clarity	Percentages
Ideal	I1	
	IF	0.419% 0.910% 0.497%
	SI1	0.451% 0.934% 1.415% 1.224% 1.127% 1.420% 1.368%
	SI2	0.508% 0.834% 0.901% 0.840% 0.869% 0.660%
	VS1	0.373% 0.756% 0.866% 1.767% 1.142% 1.099% 0.651%
	VS2	0.430% 0.812% 1.031% 1.687% 1.630% 2.106% 1.706%
	VVS1	0.604% 1.101% 0.816% 0.621%
	VVS2	0.536% 1.435% 0.964% 0.940% 0.527%
Premium	I1	
	IF	
	SI1	0.387% 0.680% 1.214% 1.049% 1.127% 1.138% 1.031%
	SI2	0.578% 0.966% 0.912% 0.970% 0.962% 0.780%
	VS1	0.410% 0.623% 1.049% 0.538% 0.541%
	VS2	0.374% 0.584% 0.986% 1.337% 1.148% 1.166% 0.628%
	VVS1	
	VVS2	0.510%
Very Good	I1	

10. Plot the horizontal bar plot of the number of diamond sales per cut and color (legend) and clarity (column). Display the number of sales in %. Use the color as the filter. Save the code for this plot under the tab 'Sales count-Barplot per color and cut and clarity'. Add an appropriate title to your graph.

Sales count bar plot per cut, colot and calrity

Clarity: I1, IF, SI1, SI2, VS1, VS2, VVS1, VVS2

Cut: Ideal, Premium, Very Good, Good, Fair

(x-axis for each: 0K 2K 4K — Count of Price)

11. Plot the vertical bar plot of the diamond price versus the number of sales in a descending order. Save the code for this plot under the tab 'Price versus number of sales. Add an appropriate title to your graph.
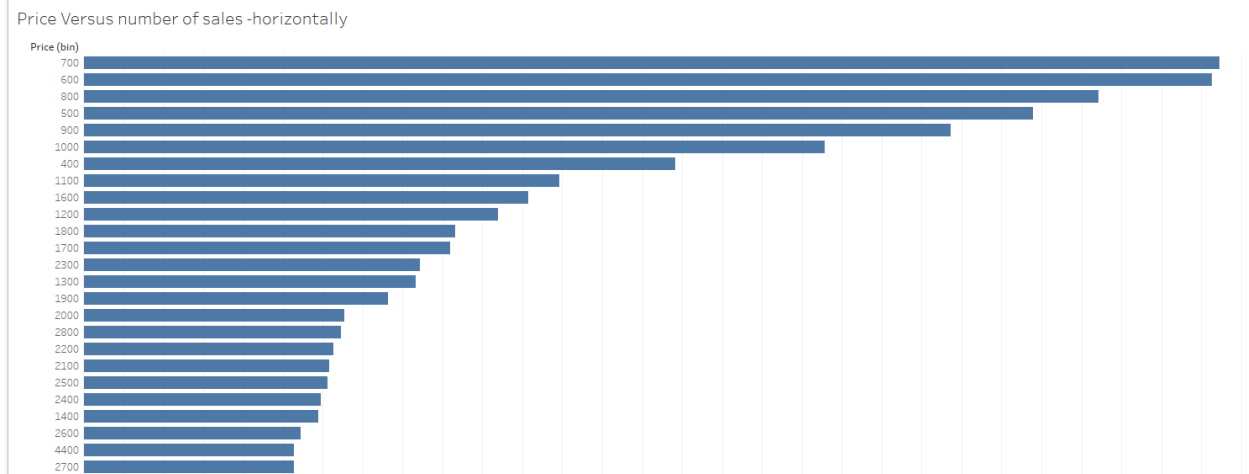
Price Versus number of sales
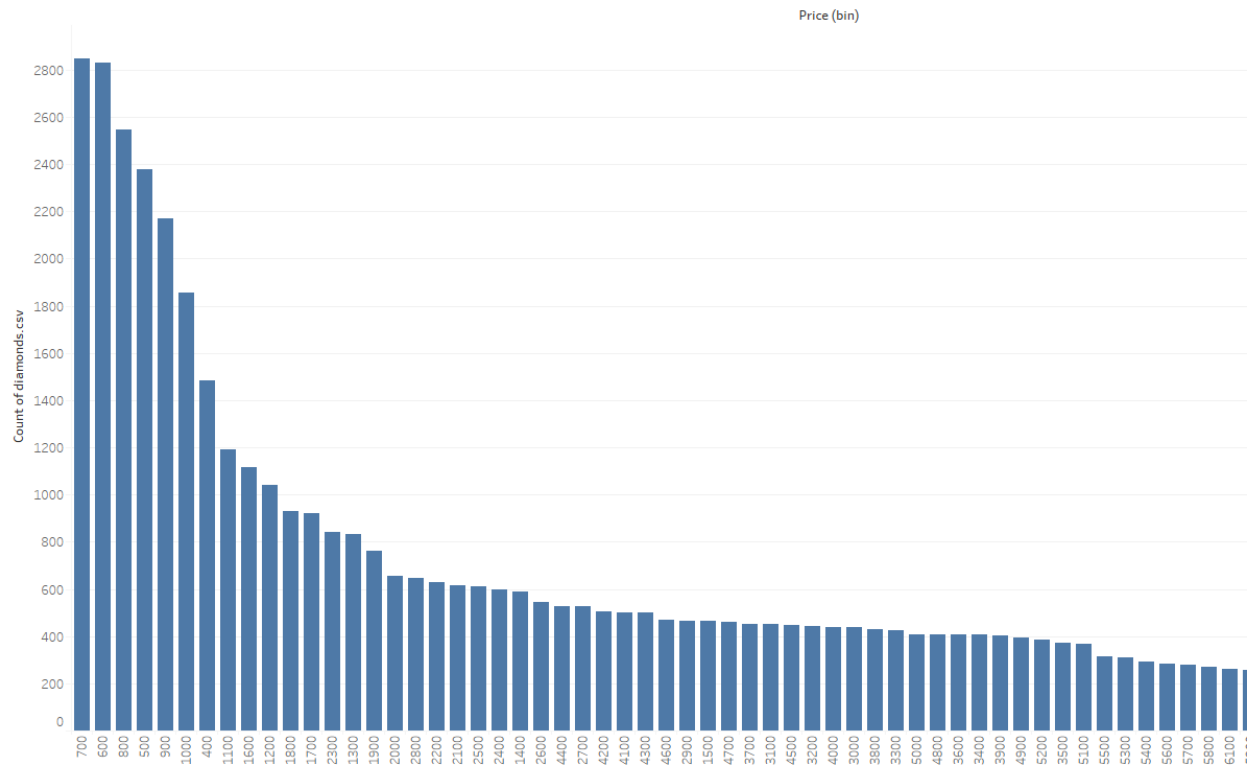
12. Plot the horizontal bar plot of the diamond price versus the number of sales in a descending order. Save the code for this plot under the tab 'Price versus number of sales. Add an appropriate title to your graph.

13. Plot the diamond price with sum aggregation versus the color with the cut as the legend. Add the price (sum) as the label to each bar.
    a. Which diamond color has the highest price (sum)?
    b. Which diamond color has the highest price (sum)?
    c. Which diamond color and cut has the highest price (sum)?
    d. Which diamond color and cut has the least price (sum)?

Diomand Price versus color and cut legend
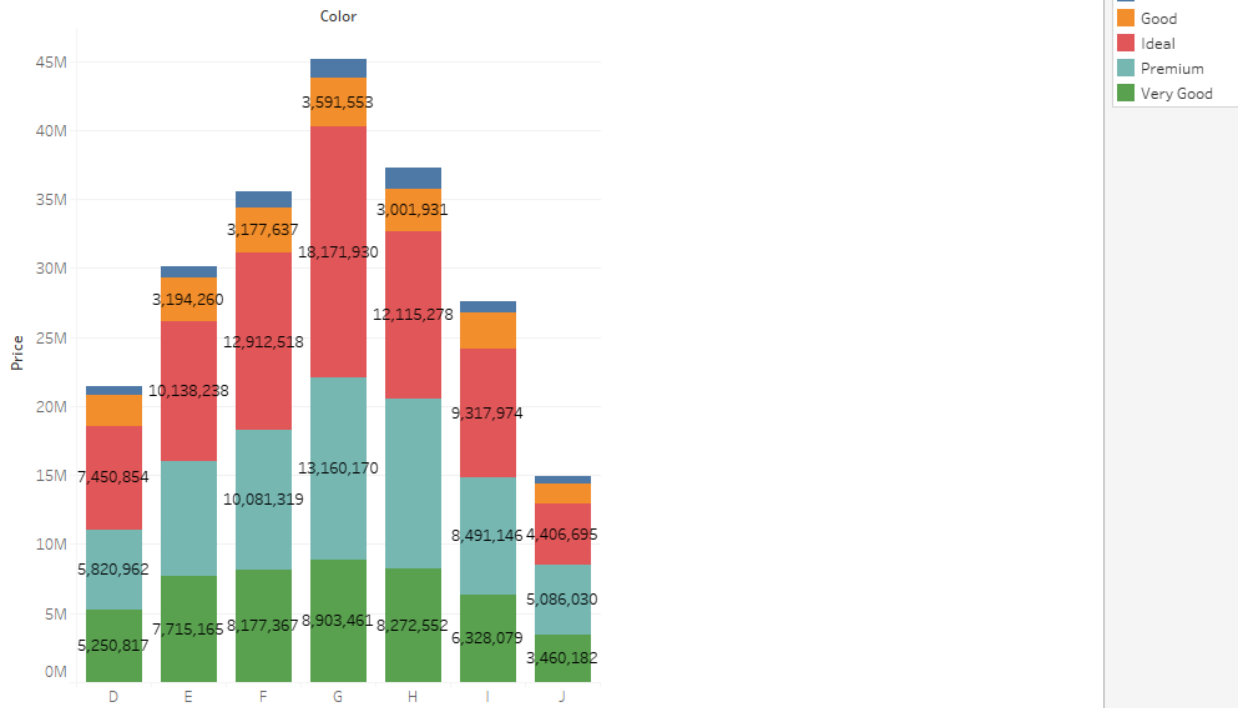
14. Plot the heat map graph of the average sales price of color versus clarity. Save the code for this plot under the tab 'Heat Map of average price color versus clarity'. Add an appropriate title to your graph.

Heat Map of average price color versus clar

| Color	I1	IF	SI1	SI2	VS1	VS2	VVS1	VVS2
			Clarity					
D	3,863	8,307	2,976	3,931	3,030	2,587	2,948	3,351
E	3,488	3,669	3,162	4,174	2,856	2,751	2,220	2,500
F	3,342	2,751	3,714	4,473	3,797	3,757	2,804	3,476
G	3,546	2,558	3,775	5,022	4,131	4,416	2,867	3,845
H	4,453	2,288	5,032	6,100	3,781	4,722	1,846	2,649
I	4,302	1,995	5,355	7,003	4,633	5,691	2,035	2,968
J	5,254	3,364	5,186	6,521	4,884	5,311	4,034	5,142

15. Plot the heat map graph of the average sales price of color versus clarity and cut. Save the code for this plot under the tab 'Heat Map of average price color versus clarity and cut'. Add an appropriate title to your graph.

Heat Map of average price color versus clarity and cut

Color	Cut	I1	IF	SI1	SI2	VS1	VS2	VVS1	VVS2
D	Fair	7,383	1,620	4,273	4,355	2,921	4,513	4,473	3,607
	Good	3,491	10,030	3,021	3,595	3,557	3,588	2,586	2,346
	Ideal	3,527	6,567	2,490	3,142	2,576	2,112	2,706	3,619
	Premium	3,819	9,057	3,236	4,351	4,178	2,919	3,771	3,888
	Very Good	2,623	10,298	3,235	4,425	2,955	3,145	2,988	2,615
E	Fair	2,095		3,901	4,172	3,308	3,042	4,115	3,119
	Good	4,398	1,519	3,162	3,785	3,713	3,772	1,906	3,390
	Ideal	3,559	3,259	2,884	3,891	2,176	2,163	2,206	2,556
	Premium	3,199	4,525	3,363	4,490	3,722	3,070	2,700	2,941
	Very Good	3,444	4,333	3,228	4,279	3,089	3,329	1,997	2,042
F	Fair	2,544	2,344	3,785	4,520	4,103	3,400	4,680	4,018
	Good	2,570	3,133	3,261	4,427	2,788	3,791	2,190	3,192
	Ideal	3,903	2,154	3,710	4,336	3,504	3,317	2,611	3,324
	Premium	3,555	3,618	4,040	4,747	4,758	4,221	3,969	4,099
	Very Good	4,253	4,677	3,574	4,250	3,881	3,996	2,827	3,462
G	Fair	3,187	1,488	3,579	5,665	3,498	5,384	2,216	3,099
	Good	3,196	4,060	4,129	4,776	4,302	4,141	2,705	3,310
	Ideal	4,044	2,206	3,441	4,612	4,117	4,310	2,909	3,796
	Premium	4,052	3,311	4,303	5,617	4,436	4,556	2,934	4,324
	Very Good	3,195	3,525	3,482	4,699	3,770	4,427	2,719	3,712
H	Fair	4,213		5,196	6,022	4,605	5,111	4,115	3,482
	Good	3,850	5,949	4,179	5,530	3,819	4,433	1,720	2,428
	Ideal	5,415	1,983	4,770	5,589	3,613	4,039	1,916	2,591
	Premium	3,904	3,385	5,708	6,719	3,949	5,554	1,454	2,651
	Very Good	5,259	2,648	4,934	6,112	3,750	4,620	2,042	2,768
I	Fair	3,501		4,575	6,658	4,500	3,856	4,194	2,995
	Good	4,175	1,749	4,743	6,933	4,597	5,957	2,651	2,758
	Ideal	4,103	1,503	5,179	7,192	3,944	4,663	2,034	2,859
	Premium	5,045	2,359	6,092	7,148	5,339	7,156	1,831	3,191
	Very Good	6,045	4,094	5,195	6,622	5,277	5,755	2,056	3,060
J	Fair	5,795		4,554	5,132	5,906	4,068	1,691	2,998
	Good	3,795	2,738	4,628	5,306	3,663	4,803	4,633	4,371
	Ideal	9,454	2,489	5,116	6,555	4,734	4,867	2,000	4,122
	Premium	4,577	7,026	5,727	7,550	5,817	6,176	7,244	6,423
	Very Good	4,478	1,074	5,027	5,993	4,340	5,326	3,176	5,960

16. Graph the scatter plot of between x and y dimension and show the trend line. Take a screen shot of the 'describe trend model' and interpret the result. Save the code for this plot under the tab 'scatter plot between x and y'. Add an appropriate title to your graph.

17.

scatter plot x and y with line
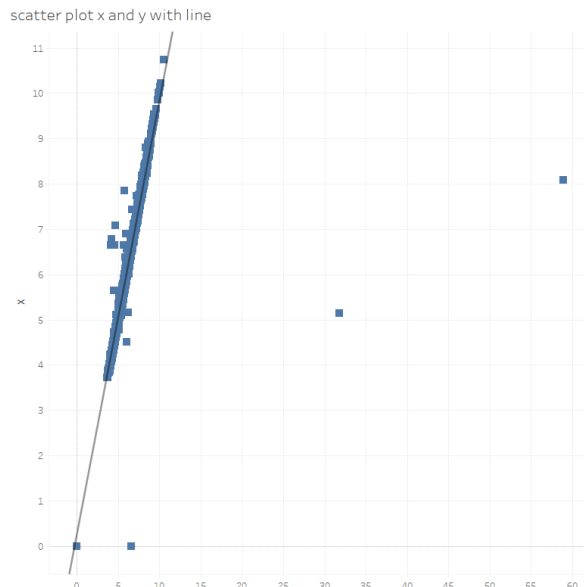
18. In this question, you need to find out that price has mostly correlated with which of the following variables : x, y, z, depth, and carat. Display the result in one graph that shows the scatter plot between price→ and (x, y , z depth and carat). Write down the equation of the line for each model. Which of the x, y , z depth, carat variable has the highest correlation with price? Save the code for this plot under the tab 'scatter plot between price and others'. Add an appropriate title to your graph.

19. Create an interactive dashboard as shown below. The interactive means that all charts in the dashboard are connected.

Pie Chart percentage sales count per cut and clairty

Cut

Fair	Good	Ideal	Premium	Very Good

0.017%
2.892%
7.938%
6.628%
6.007%

0.315%   0.756%
9.401%   4.816%   4.803%
1.201%   3.687%   3.291%   3.893%
0.864%   2.004%   6.654%   5.467%

Cut
- Fair
- Good
- Ideal
- Premium
- Very Good

% of Total Count of diamo..
- 5.206%
- 10.000%
- 15.000%
- 20.934%

Number of Sales per cut and Percentage- Bar Plot

Cut

Ideal	3.880%	5.775%	9.055%	7.093%	7.236%	5.254%
Premium	2.647%	4.375%	5.421%	4.321%	4.333%	2.972%
Very Good	3.382%	4.262%	4.012%	4.449%	2.805%	
Good						

OK 2K 4K 6K 8K 10K 12K 14K 16K 18K 20K 22K

Count of Price

Clarity
- I1
- IF
- SI1
- SI2
- VS1
- VS2
- VVS1
- VVS2

Pie Chart Sales Count per color and cut

Color

D	E	F	G	H	I	J

2.805%   1.227%   1.730%   1.685%   1.615%   1.301% 2.232%   0.968% 1.257%   0.569%

2.972%   5.254%   7.236%   7.093%   9.055%   5.775%   3.880%   1.661%

Color
- D
- E
- F
- G
- H
- I
- J

Pie Chart Sales count percentage per color and clarity

Clarity

I1	IF	SI1	SI2	VS1	VS2	VVS1	VVS2

0.093%   0.293%   3.862%   2.540%   3.146%   1.025%
0.189%   0.714%   2.898%   3.176%   2.375%   1.216%   1.837%
0.300%   0.265%   4.218%   4.498%   2.167%   4.579%
0.278% 1.263%   3.663%   3.951%   2.983%   2.529%   1.361%   1.808%
4.080%

Upload the **formal report (as a single pdf**) plus **.twb file** through BB. Make sure to include all answers and the corresponding graphs into your report.