Diamonds

Stanislav Fatkhutdinov

University of Maryland Global Campus

Assignment 4

DATA 610 Section 9040

Decision Management Systems

Vahe Heboyan

Diamonds

**Introduction:**

This assignment is a continuation of Assignment 3 with the same dataset that describes the features of 54,000 diamonds. The set consists of several features: price in US dollars; carat weight of the diamond; cut quality; diamond color and clarity; x, y, z dimensions, and the derived from the depth and another derived property – table. The price ranges from $326 to $18,823 per diamond. The carat weight fluctuates between 0.2 and 5.01 carats. The cut quality is divided between Fair, Good, Very Good, Premium, and Ideal. Diamonds are also separated by color, where J is the worst, and D is the best color. The clarity feature is measured by how clear the diamond is; it ranges from I1, SI2, SI1, VS2, VS1, VVS2, VVS1, IF, where 'I1' is the worst clarity and 'IF' is the best possible clarity for a diamond. Other features of the diamonds in the dataset are x, y, and z dimensions representing length, width, and depth. Depth and Table are calculated properties of the diamonds describing the total depth percentage and the width of the top of a diamond relative to the widest point, respectively (Magrawal, 2017). The following formula calculates depth:

Figure 1 below represents the snapshot of the data described above. I've decided to hide-from-user several columns in the dataset. The number column is redundant since Cognos Analytics automatically creates Row Id; plus, the number column is not relevant to the diamonds' features. I also hid the x, y, z column since their derived property of depth is the one I'm going to use for my analysis. For this assignment, I will use the Diamond dataset to create a dashboard and a storyboard using Cognos Analytics to convey insights gleamed from the diamond data.

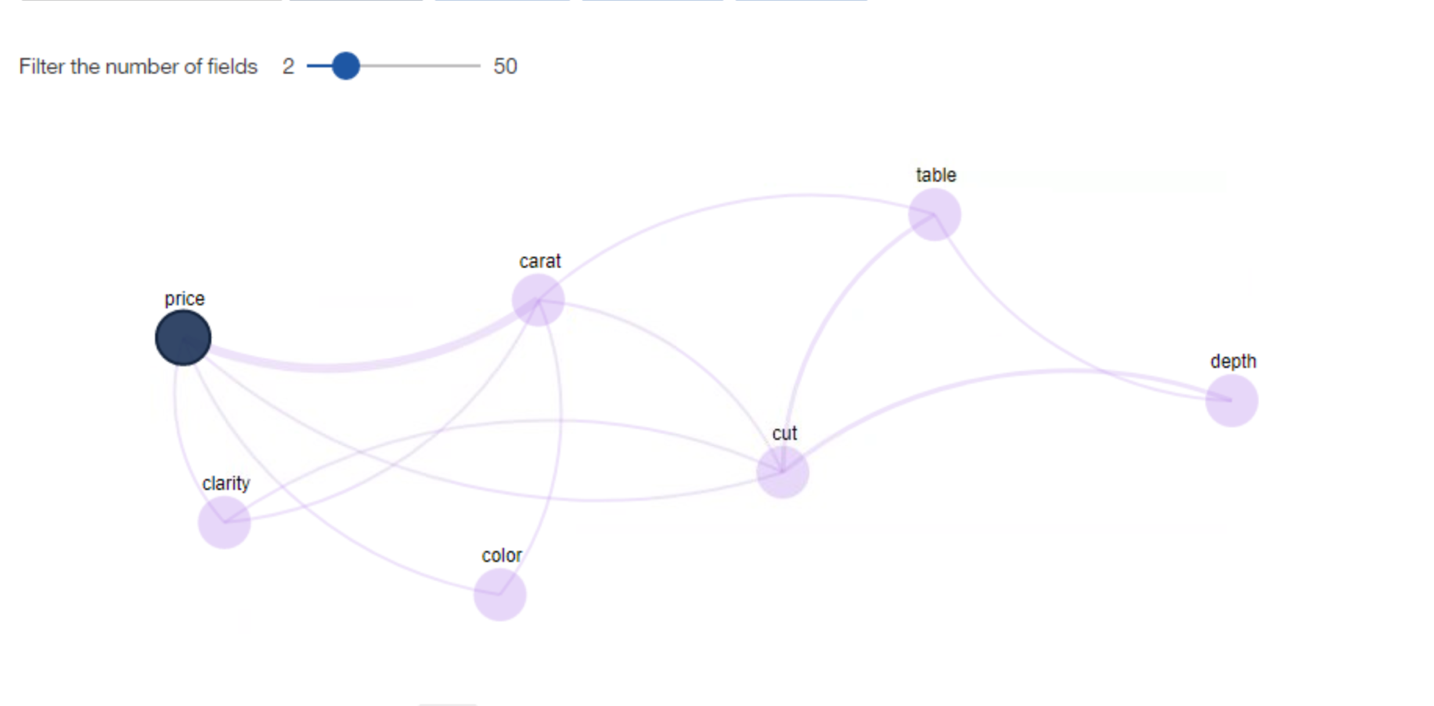


Figure 1 Partial Data Snapshot

**Dashboard Assembly:**

Using my findings from assignment 3, I've decided to create a dashboard that can quickly determine the price and carat of diamonds based on their color, cut, and clarity. I also used diamond carat weight and cost to visualize the comparative data. Figures 2 and 3 below reflect an average price and carat weight of all the diamonds in the dataset.

Icon

Description automatically generated

Figure 2 Average Price

Icon

Description automatically generated

Figure 3 Average Weight

I needed to create a meaningful dependency between the diamond dataset features to quickly and accurately predict the two numbers reflected in Figures 2 and 3. Using my previous findings, I decided to promptly use the diamonds' cut to display the most expensive types of diamonds. However, my initial display was a bit untidy, so I've decided to group the amounts by colors of diamonds. The results in Figure 4 present that 'J' and 'I' colors of the 'Premium' cut are the most expensive diamonds.

Chart, bubble chart

Description automatically generated

Figure 4 Cut by Price and Color

Now that I've figured out the diamonds' price breakdown, I wanted to focus on the weight and clarity. I tried to present which diamonds are the heaviest based on their transparency. Figure 5 demonstrates that 'I1' clarity diamonds are the heaviest among the provided dataset, followed closely by the 'SI2' clarity diamonds.

Chart, bubble chart

Description automatically generated

Figure 5 Weight by Clarity

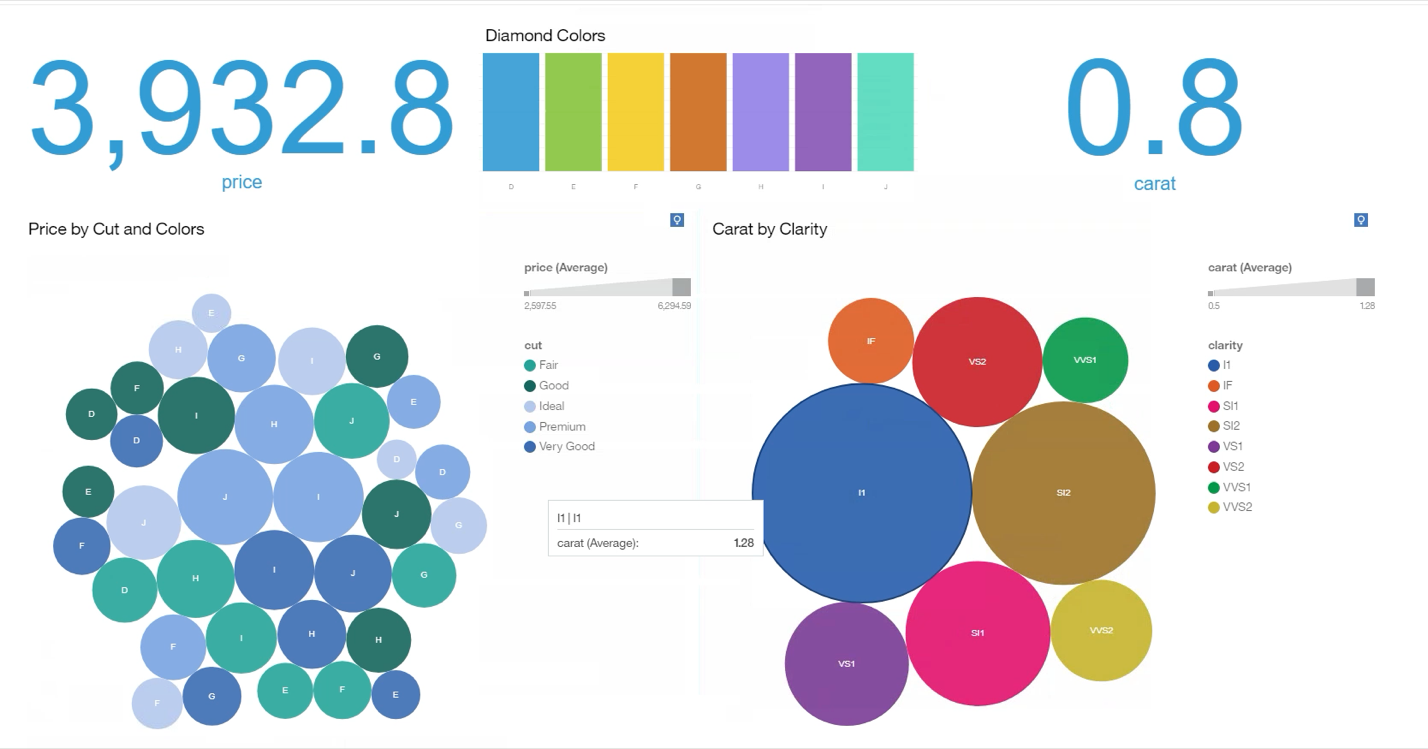
After creating the visualizations mentioned above, I wanted to use the Cognos Analytics capabilities of quick sorting (IBM Cognos Analytics). I've decided to sort my data by colors fast using the dashboard features. Figure 6 demonstrates the quick clickable assortment of diamond colors that the user can use to gain insight into the diamond data.

Chart, bar chart

Description automatically generated

Figure 6 Diamond Color Picker

The completed dashboard, presented in Figure 7, shows the combination of the abovementioned visualizations to give the user a quick insight into the diamond dataset. The average price of a diamond is $3,932.8, and the average weight is 0.8 carats. Once the user interacts with the dashboard, they can quickly price and weigh the diamonds based on their color, cut, and clarity.

Chart, bubble chart

Description automatically generated

Media 1

Figure 7 Completed Dashboard

Media 1 demonstrates the simple use features of the dashboard to drill down the diamond data's insights.

**Story Assembly:**

For my story, I decided to focus on the contributing factors of the price of diamonds. Figure 8 demonstrates the layout of my account and five essential factors affecting the price.

Graphical user interface

Description automatically generated

Figure 8 Storyboard Layout

I've focused on cut, clarity, color, and weight as my main features affecting price. The premium cut diamonds are the most expensive in the dataset. Clarity of diamonds, as I found out, plays a less significant role. However, SI2 clarity diamonds were the most costly diamonds despite being lighter than some other diamonds, as seen in Figure 9.

Chart, scatter chart

Description automatically generated

Figure 9 Clarity of Diamonds

On average, the J color diamonds were the most expensive, closely followed by the I color diamonds. The weight or the diamonds' carat weight is the most important contributing factor to the price of diamonds. The story concludes with the derived knowledge of the contributing factors to the price by the weight of diamonds in carats and their cut. Media 2 demonstrates the story unfolding.

Media 2 CA Story

**Organizational Impact:**

I found a storyboard a less useful tool to convey insights due to some of its limitations. My organization would probably benefit the most from the interactive dashboard. The ability to quickly sort and organize data with a real-time visualization tool like the CA dashboard will be a great addition to our projects. The storyboard seems to be less useful for a day to day operations but probably can be used during contract bidding to present the highlights of what my company can offer.

**Conclusion:**

The dashboard and storyboard tools by Cognos Analytics can be a great addition to my work environment. Using these tools, I was able to get the relevant insights from the diamond dataset. I determined that the price can be predicted by the weight and cut of the diamonds and that color and clarity of those diamonds can play a role when appraising them.

# References

IBM Cognos Analytics. (n.d.). *Explorations User Guide.* Retrieved September 2020, from www.ibm.com: https://www.ibm.com/support/knowledgecenter/SSEP7J\_11.1.0/com.ibm.swg.ba.cognos.ca\_explorations.doc/ca\_explorations.pdf?view=kc

Magrawal, S. (2017, May 24). *Diamonds.* Retrieved from www.kaggle.com: https://www.kaggle.com/shivam2503/diamonds/metadata