

Data Visualization Bootcamp Homework

Pichayud Udomsud

2023-11-01

Use diamonds dataset (or other datasets) to create 5 charts

install library tidyverse , patchwork

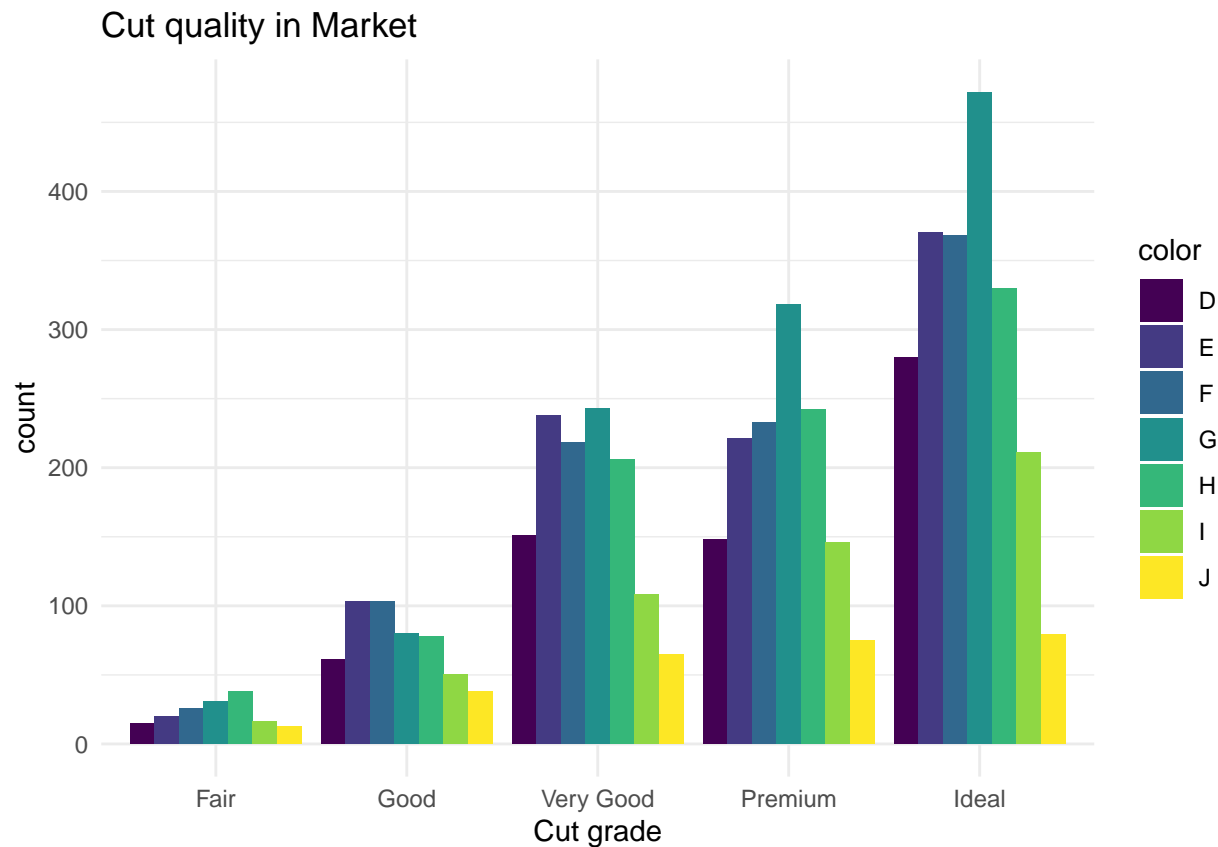
```
library(patchwork)
library(tidyverse)
```

sample bigdata

```
set.seed(20)
sample_diamonds <- sample_frac(diamonds,0.1)
```

1.Quality of cut diamonds in Market

```
sample_diamonds %>%
  ggplot(aes(cut,fill = color)) +
  geom_bar(position = "dodge") +
  labs(title = "Cut quality in Market",
       x = "Cut grade") +
  theme_minimal()
```

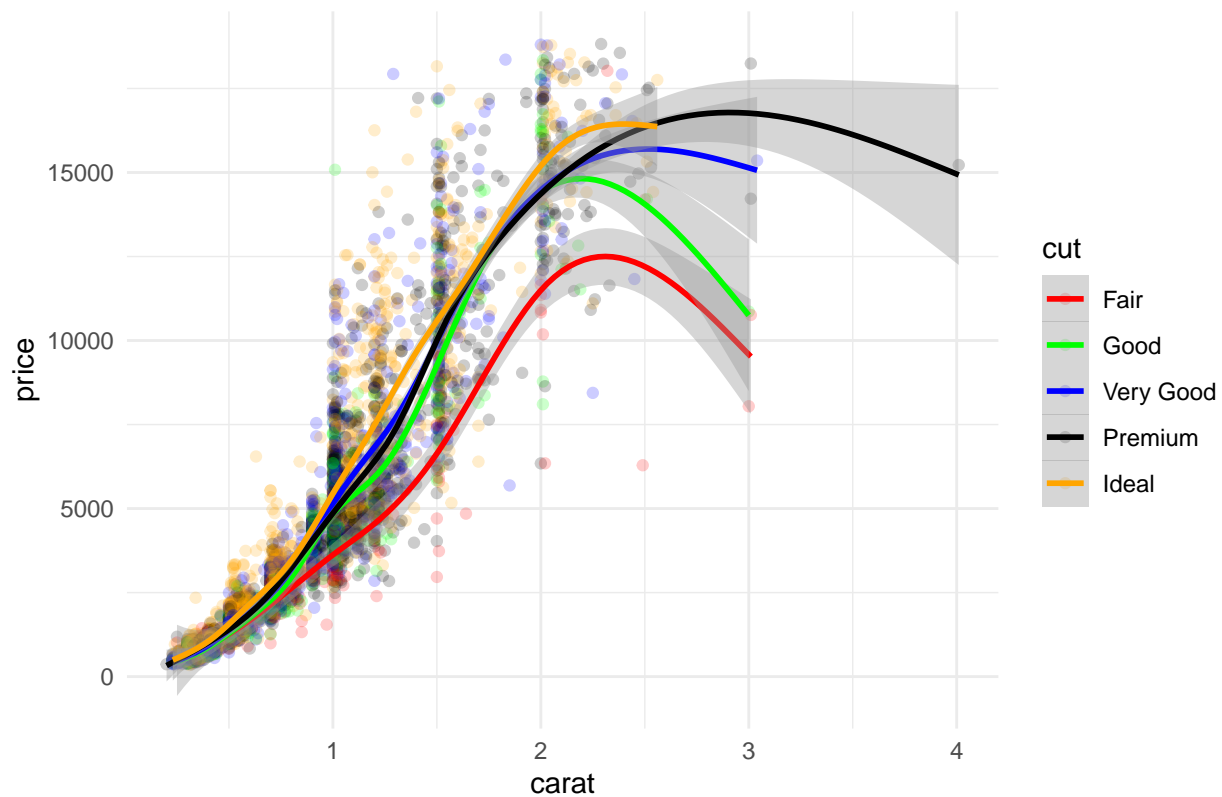


It can be noted that cut can be controlled because it is a process that is done by humans. Therefore, most diamonds on the market have ideal cut quality.

2. Relationship between carat/price groupby cut

```
sample_diamonds %>%
  ggplot(aes(carat, price, col = cut)) +
  geom_point(alpha = 0.2) +
  geom_smooth() +
  scale_color_manual(values = c("red", "green", "blue", "black", "orange")) +
  labs(title = "Relationship between Carat and Price by cut",) +
  theme_minimal()
```

Relationship between Carat and Price by cut



This graph supports the first point that the market gives higher prices to diamonds with higher cut grades, in the same carat size.

summary The graph supports the claim that the market values the quality of cut, and is willing to pay more for diamonds with higher cut grades.

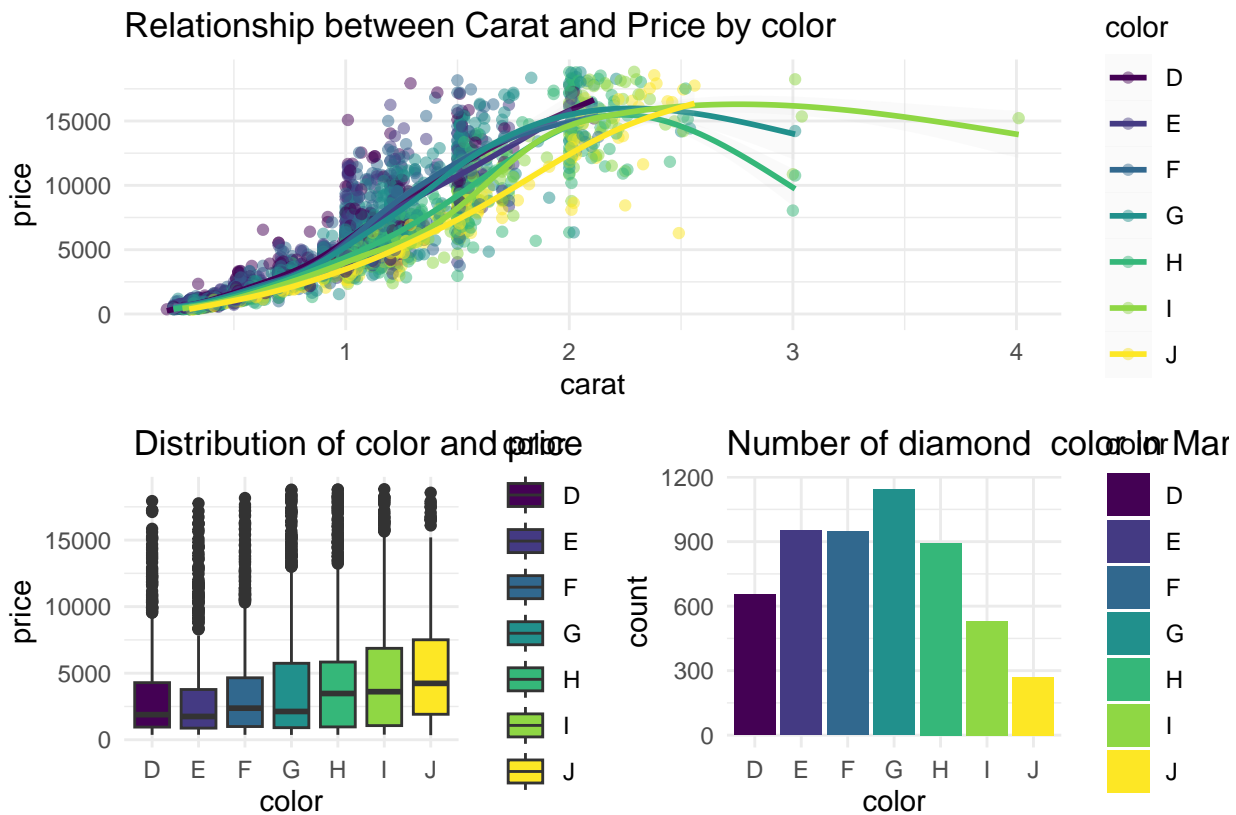
3. Relationship between carat/price groupby color

```
a1 <- sample_diamonds %>%
  ggplot(aes(carat,price,col = color)) +
  geom_point(alpha = 0.5) +
  geom_smooth(fill = "lightgray",alpha =0.1) +
  labs(title = "Relationship between Carat and Price by color") +
  theme_minimal()

a2 <- sample_diamonds %>%
  ggplot(aes(color,price,fill = color)) +
  geom_boxplot() +
  labs(title = "Distribution of color and price") +
  theme_minimal()

a3 <- sample_diamonds %>%
  ggplot(aes(color,fill = color)) +
  geom_bar() +
  labs(title = "Number of diamond color in Market") +
  theme_minimal()
```

```
a1 / (a2+a3)
```

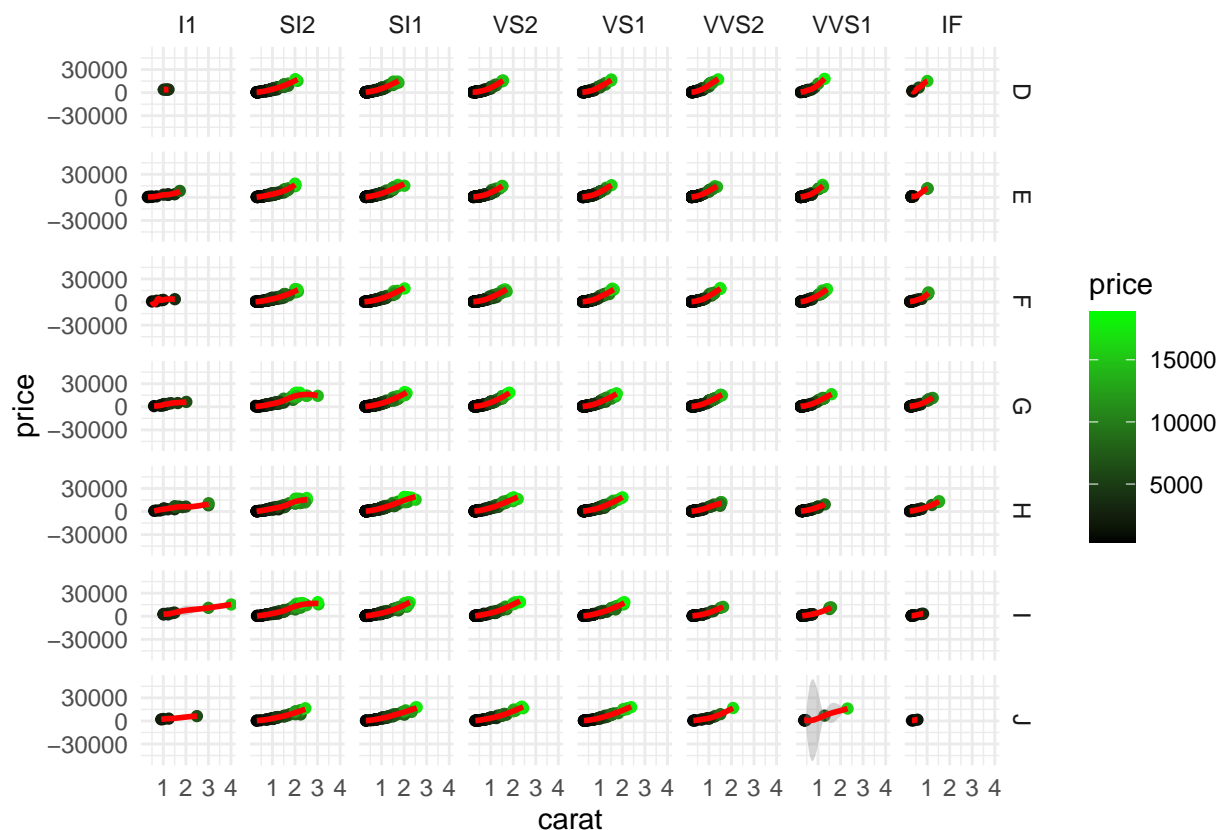


The results of the comparison are similar. The boxplot and barplot of color do not match the scatter plot, suggesting that the color of diamonds is naturally occurring and the supply is limited. This results in a relationship with demand and supply in the market.

summary In addition to cut, carat and color are also important factors that affect the price of diamonds.

4. Try to compare color and clarity between carat and price

```
a4 <- sample_diamonds %>%
  ggplot(aes(carat, price, col = price)) +
  geom_point() +
  geom_smooth(col = "red") +
  facet_grid(color ~ clarity) +
  theme_minimal() +
  scale_color_gradient(low = "black", high = "green")
a4
```



We can compare clarity and color to carat and price to see if they have the same effect on price.

It can be concluded that clarity and color still directly affect the price of diamonds.

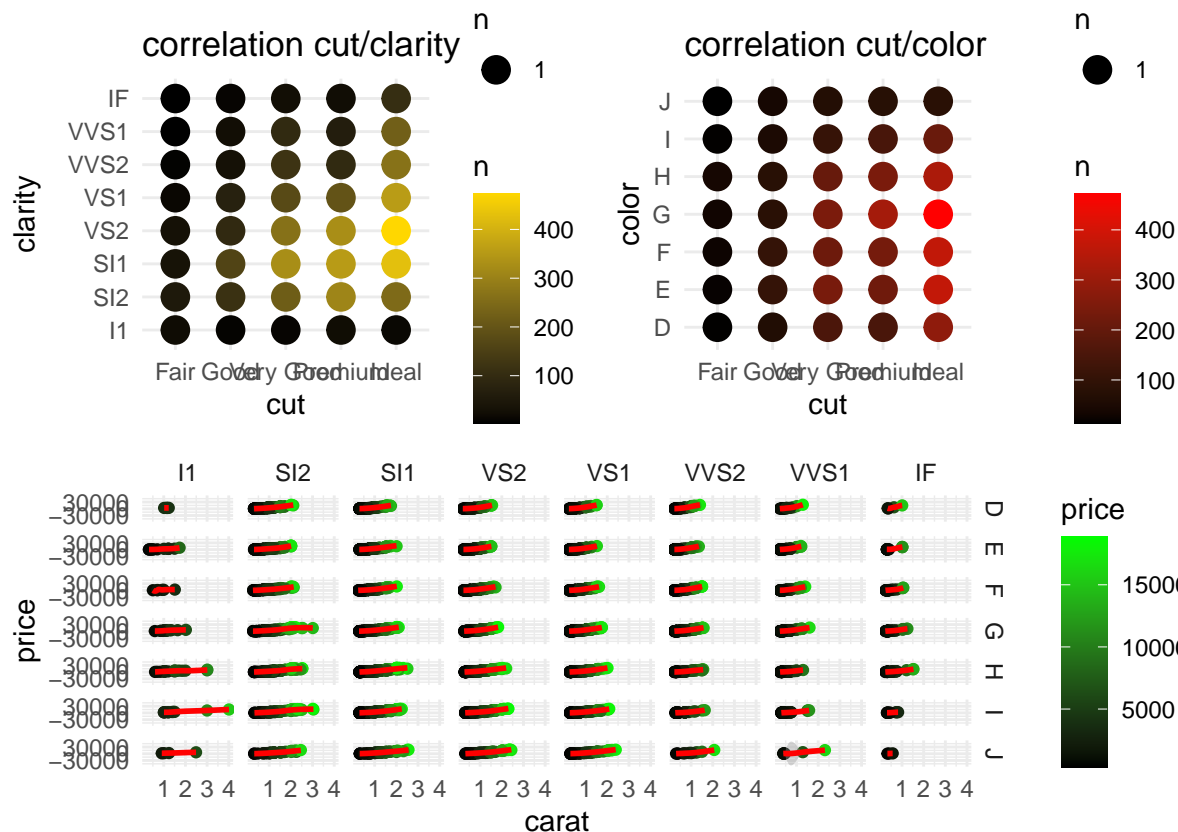
5. What if we want a reasonable and fair price for a diamond?

1. We need to combine all the comparison tables to analyze them, first based on cut.
2. We choose from the largest number in the market in the best spec, making the price reasonable.

```
a5 <- sample_diamonds %>%
  count(cut,clarity) %>%
  ggplot(aes(cut,clarity, color = n)) +
  geom_count() +
  labs(title = "correlation cut/clarity") +
  theme_minimal() +
  scale_color_gradient(low = "black",high = "gold")

a6 <- sample_diamonds %>%
  count(cut,color) %>%
  ggplot(aes(cut,color, color = n)) +
  geom_count() +
  labs(title = "correlation cut/color") +
  theme_minimal() +
  scale_color_gradient(low = "black",high = "red")
#use a4

(a5 + a6) / a4
```



After plotting a graph of the relationship between cut, clarity, and color, and comparing it to price/carat below, we can conclude that if we are to buy a diamond at a reasonable price, we will choose

Based on the graph of the relationship between cut, clarity, color, and price/carat, we can conclude that the best way to get a reasonable price for a diamond is to choose an ideal cut, a clarity range of SI1 to VS1, and a color range of E to H.

It is also important to remember that the price of a diamond is subjective. What one person considers to be a reasonable price, another person may not.