Homework Data Viz Batch 10

Carereen

2024-08-21

load library

```
library(tidyverse)
```

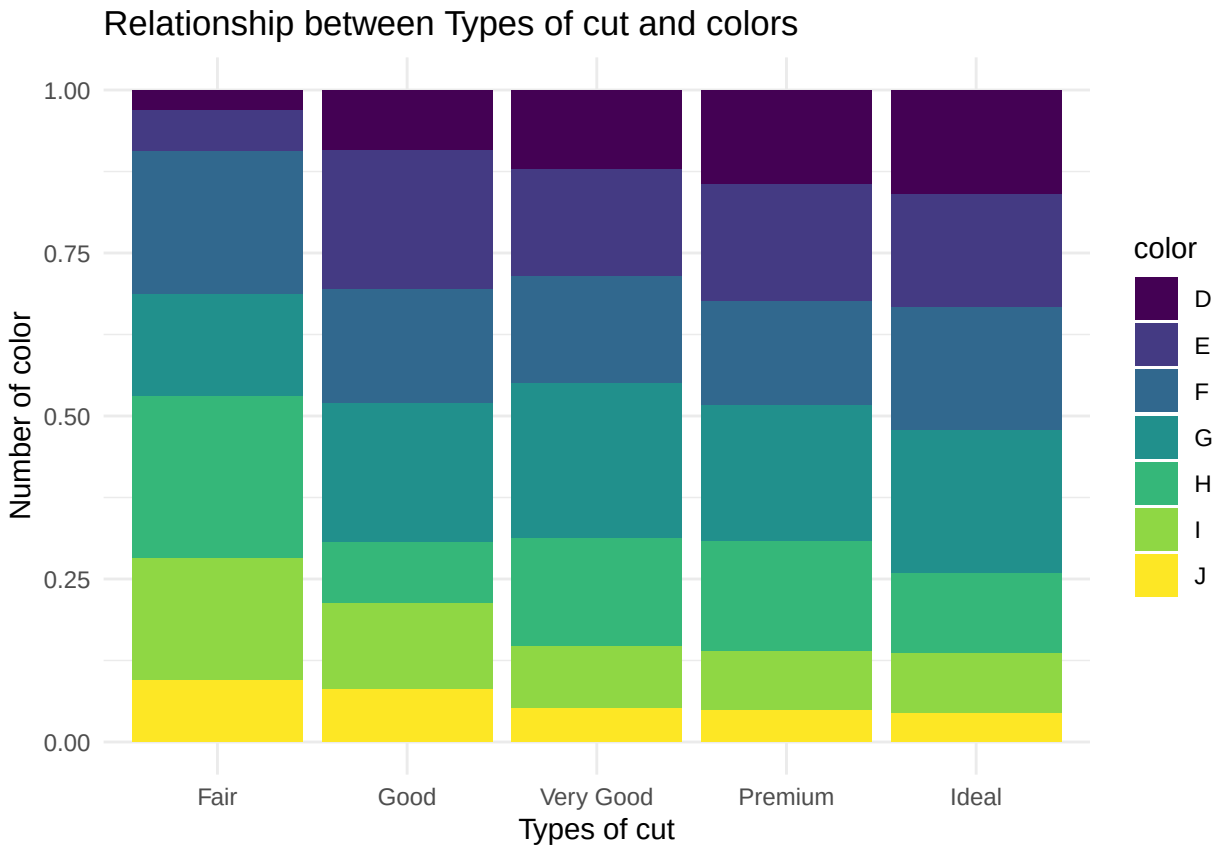
```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.5
## v forcats   1.0.0     v stringr   1.5.1
## v ggplot2   3.5.1     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.1
## v purrr     1.0.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
## i Use the conflicted package (<http://conflicted.r-lib.org/>) to force all conflicts to become errors
```

prepare data

```
small_df <- diamonds %>%
  filter(carat < 2) %>%
  sample_n(1000)
```

1. bar chart

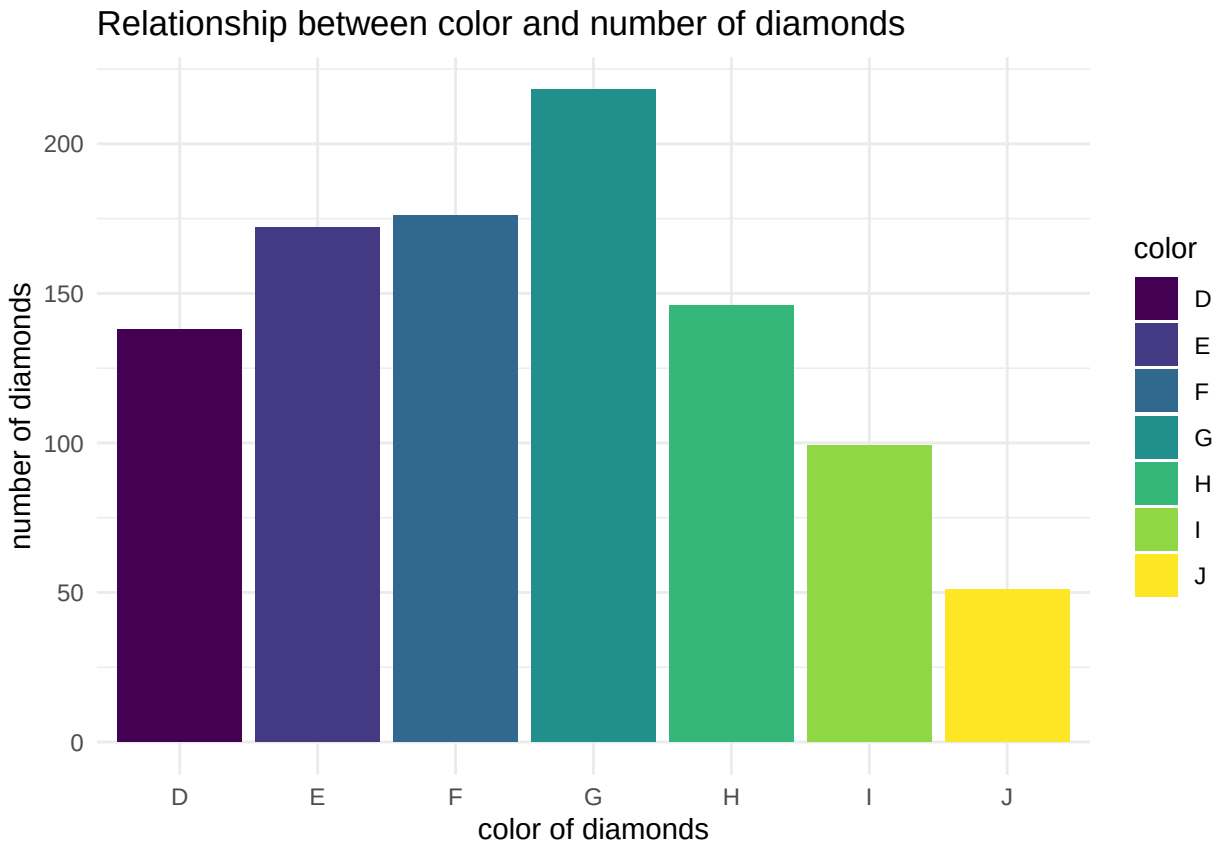
```
ggplot(small_df, aes(cut, fill=color)) +
  geom_bar(position="fill") +
  labs(title = "Relationship between Types of cut and colors",
       x = "Types of cut",
       y = "Number of color") +
  theme_minimal()
```



This is relationship between Types of cut and colors chart, the bar chart shows the types of cut then the area of color is large, there will be a lot of diamonds.

2. bar chart

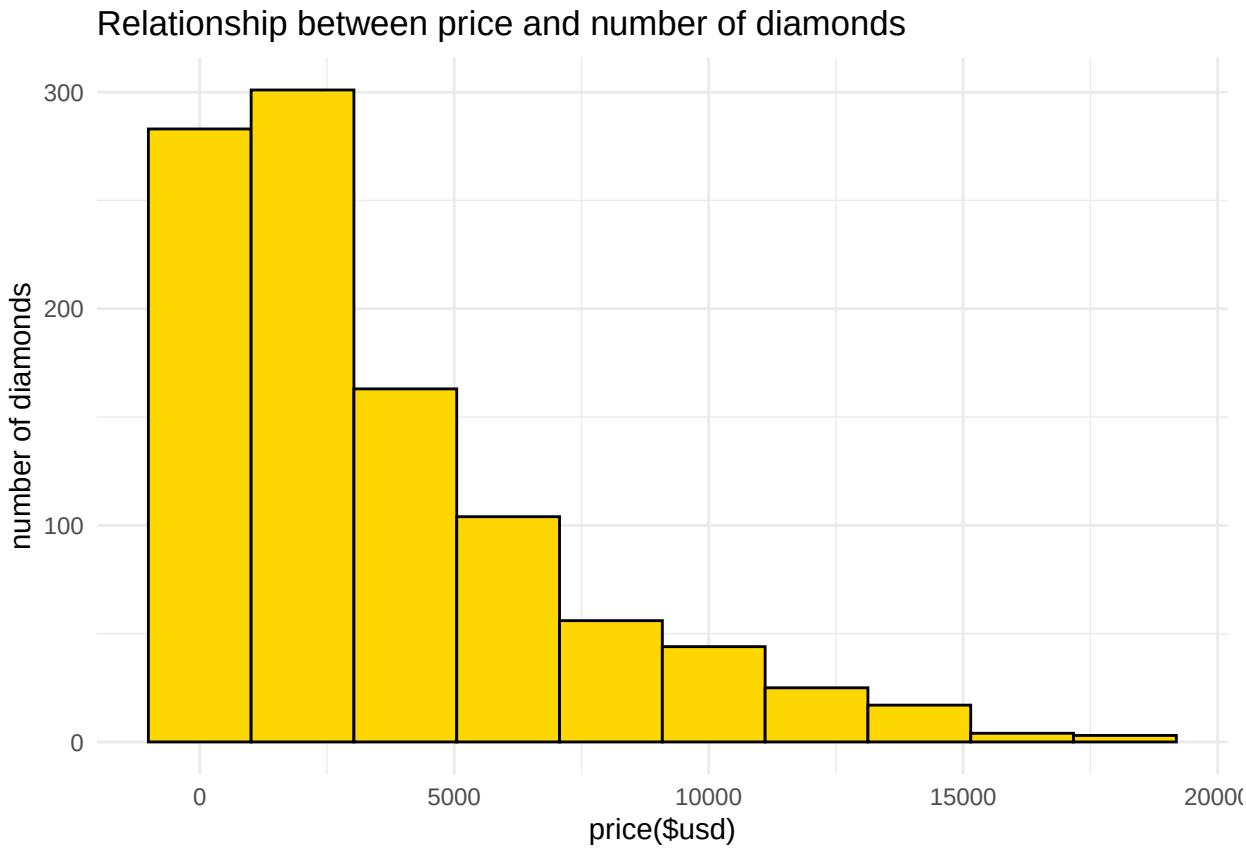
```
ggplot(small_df, aes(color, fill=color)) +
  geom_bar() +
  labs(title = "Relationship between color and number of diamonds",
       x = "color of diamonds",
       y = "number of diamonds") +
  theme_minimal()
```



This is Relationship between color and number of diamonds chart, the bar chart shows number of diamonds by color of diamonds.

3. histrogram chart

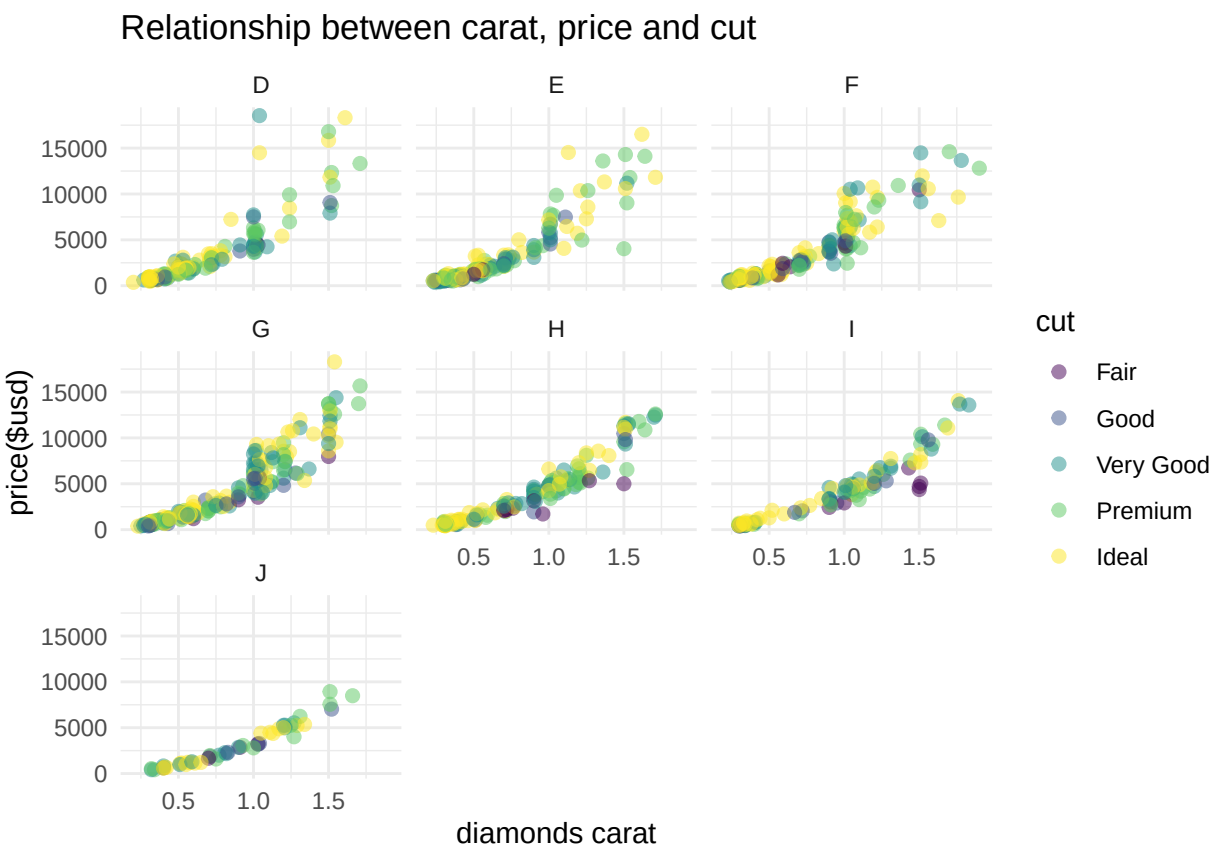
```
ggplot(small_df, aes(price)) +
  geom_histogram(bins=10, fill="gold", col="black") +
  labs(title = "Relationship between price and number of diamonds",
       x = "price($usd)",
       y = "number of diamonds") +
  theme_minimal()
```



This is Relationship between price and number of diamonds chart, the point chart shows in USA the most diamonds price less than 5000 usd.

4. point chart

```
ggplot(small_df, aes(carat, price, col=cut)) +
  geom_point(size=2, alpha=0.5) +
  facet_wrap(~color, ncol=3) +
  labs(title = "Relationship between carat, price and cut",
       x = "diamonds carat",
       y = "price($usd)") +
  theme_minimal()
```

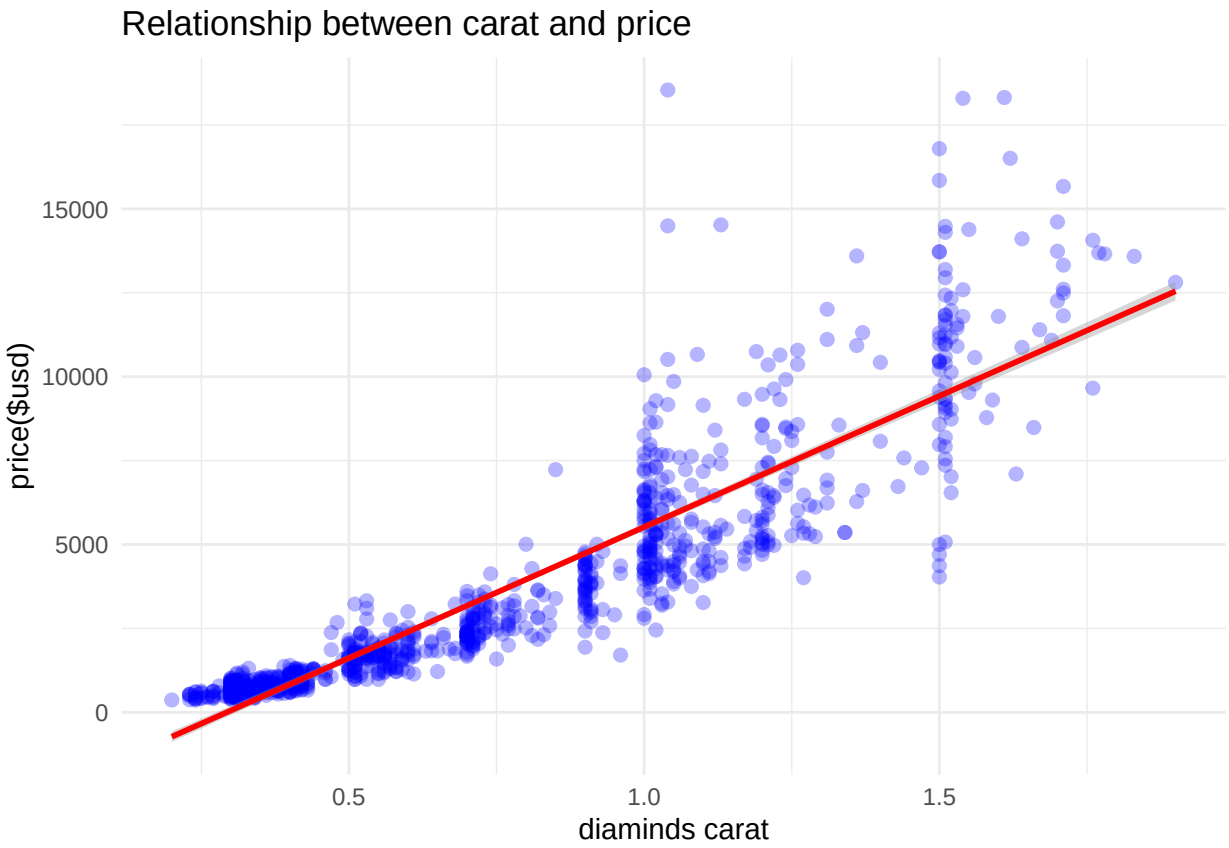


This is relationship between carat, price and cut chart, If the carat of a diamond increases, the price will increase.the diamonds of cut have 7 types is D, E, F, G, H, I, J and types of cut is Fair, Premium, Ideal.

5. point chart

```
ggplot(small_df, aes(carat, price)) +
  geom_point(col="blue",
             size=2,
             alpha=0.3) +
  geom_smooth(method = "lm", col="red") +
  labs(title = "Relationship between carat and price",
       x = "diaminds carat",
       y = "price($usd)") +
  theme_minimal()
```

```
## `geom_smooth()` using formula = 'y ~ x'
```



This is relationship between carat and price chart, If the carat of a diamond increases, the price will increase.