Data Visualization HW

Songwut Khongrung

2023-07-05

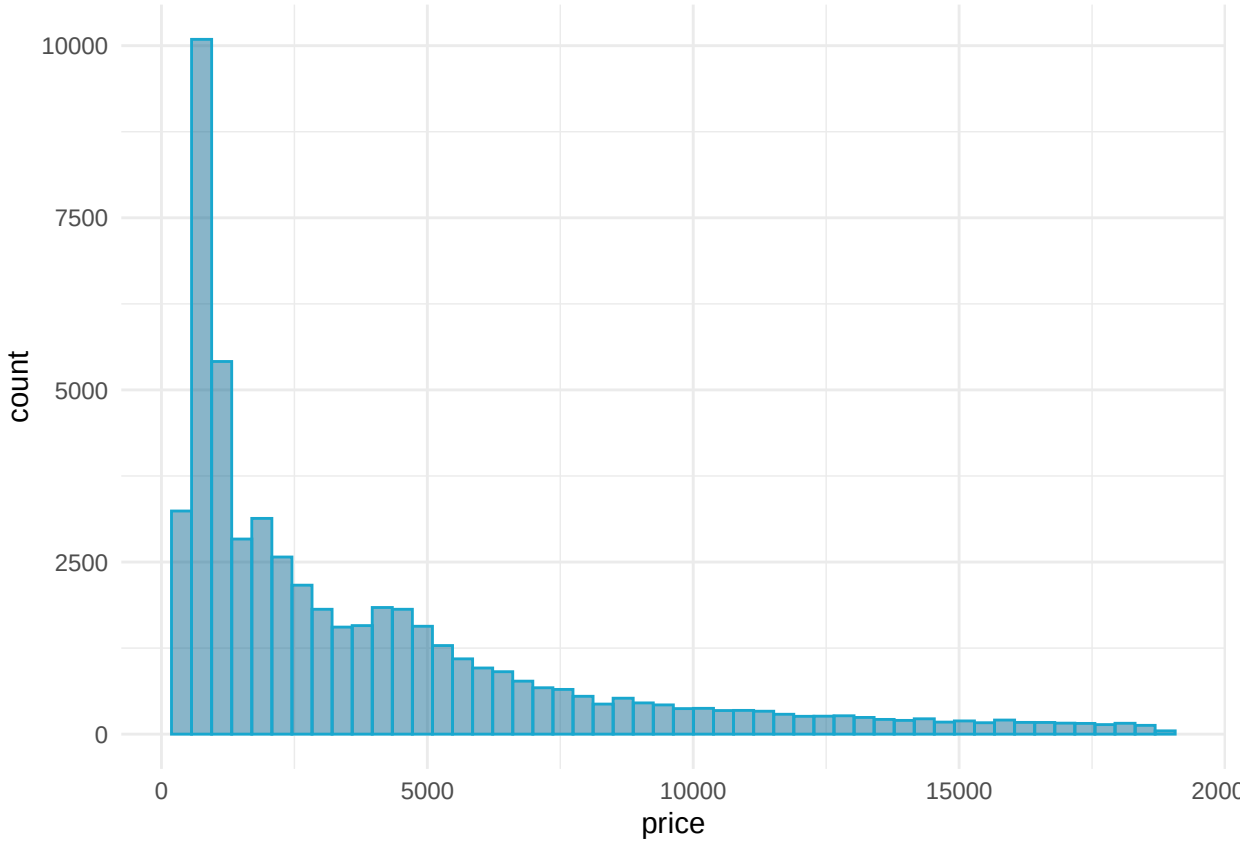
5 Chart in every day used

5 Chart in below i very proud to present .

1. Histogram

It simple chart . we have to use it and easily to understand

```
library(ggplot2)
library(tidyverse)
ggplot(diamonds,
  mapping = aes(x=price))+
  geom_histogram(bins=50,color="#19A7CE",fill="#146C94",alpha=0.5)+
  theme_minimal()
```



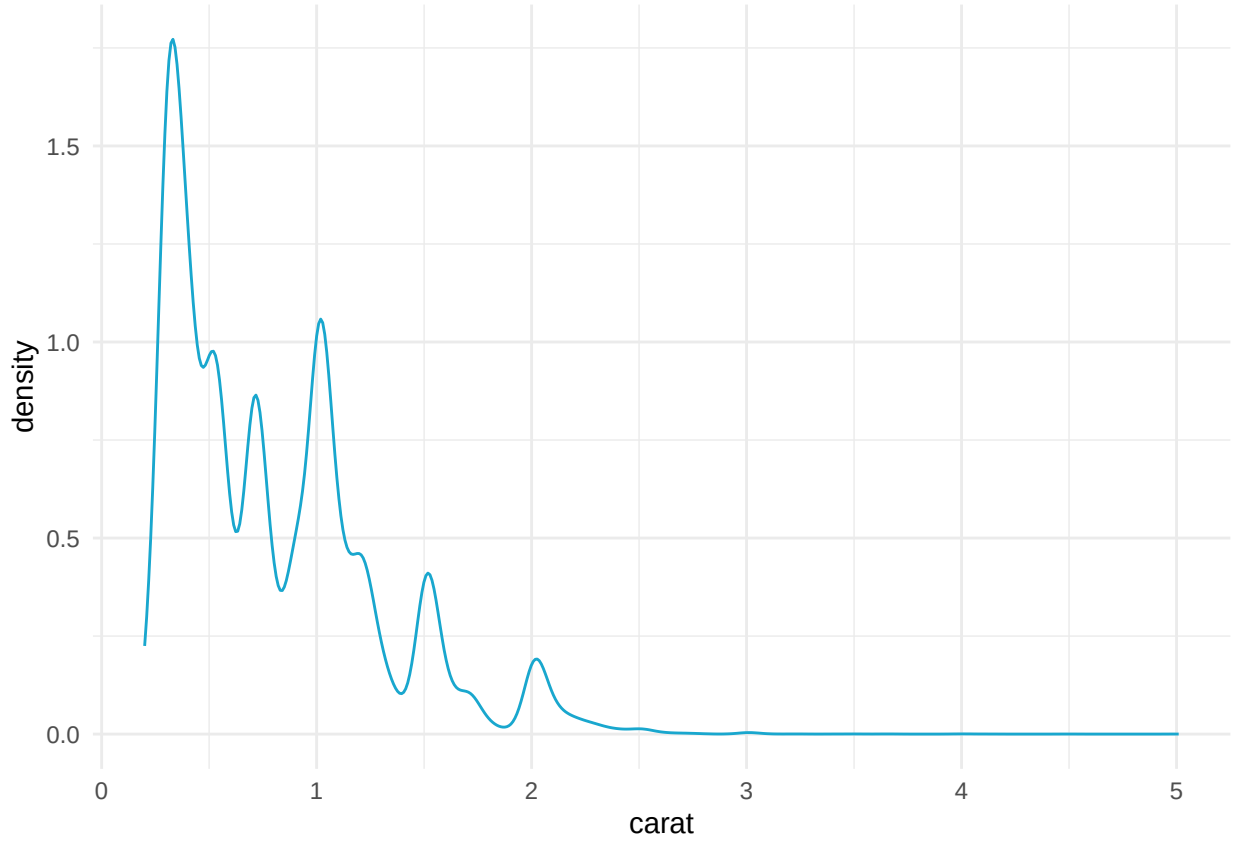
```
#bins
#color
#fill
#alpha
```

This chart show frequency of data. In above chart is Positively Skewed.

2.density

It similar histogram but show line

```
ggplot(diamonds,
 mapping = aes(x=carat,))+
 geom_density(color="#19A7CE")+
 theme_minimal()
```



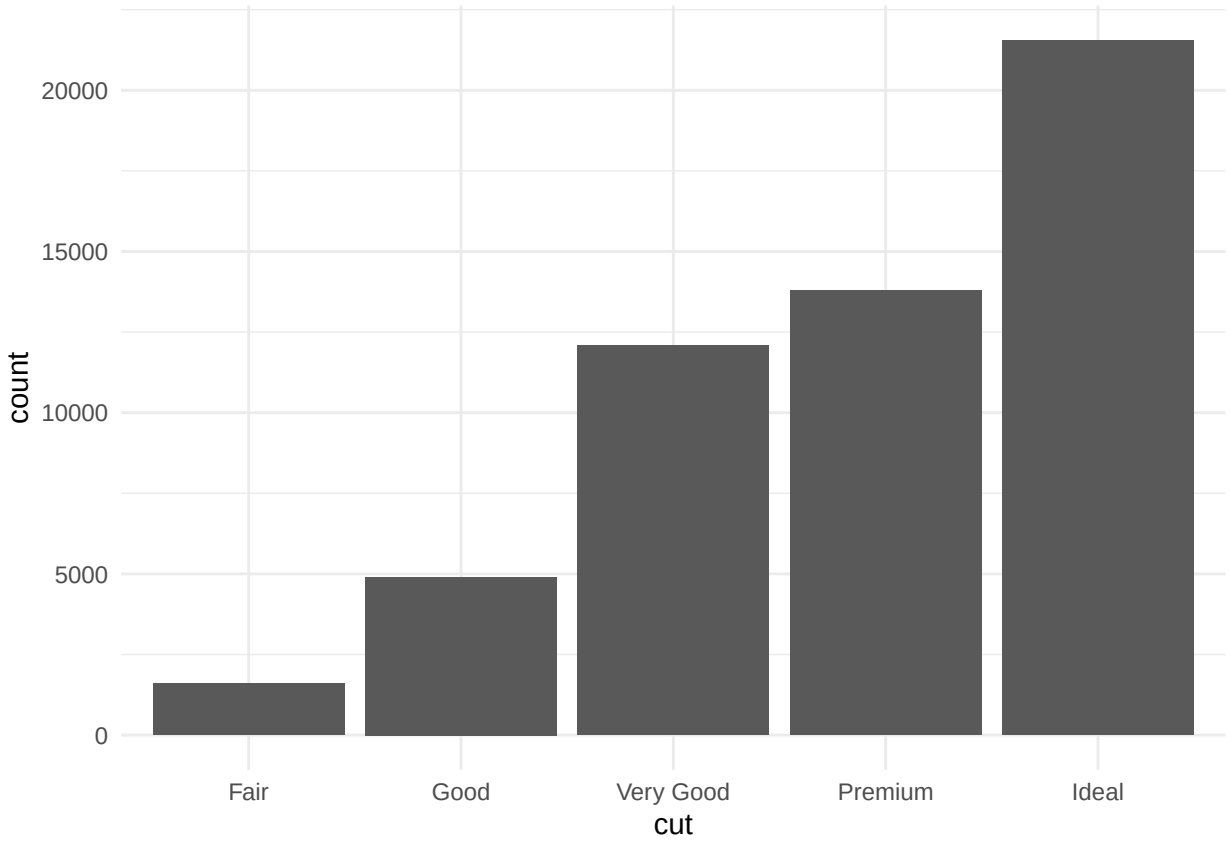
This result seem histogram . most frequency have between 0-1 data make it Positively Skewed.

If we use data not cleansing data. make it error result

3.Bar

This bar chart very popular .I think everyone know this chart

```
ggplot(diamonds ,
  mapping = aes(x=cut))+
  geom_bar()+
  theme_minimal()
```

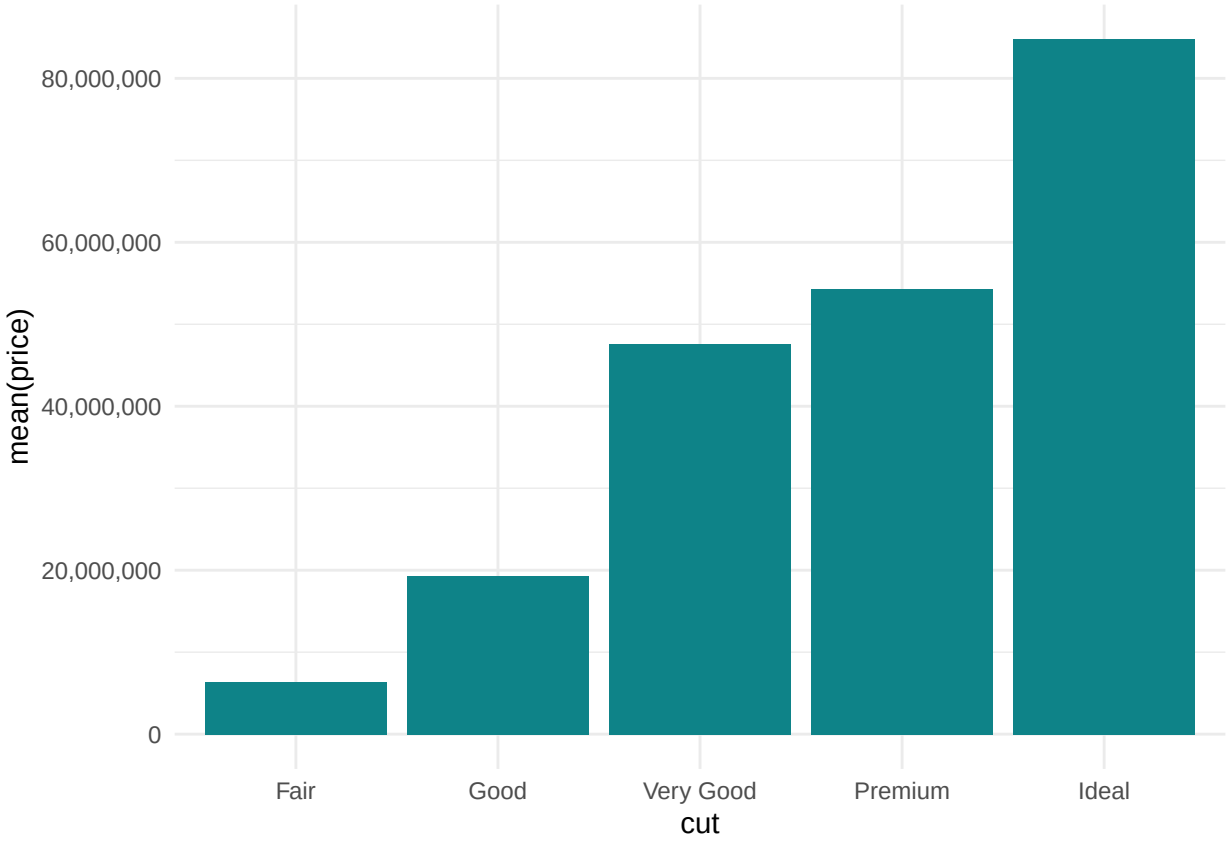


quality of diamonds in data set we have quality ideal the most .

quality ideal many times more than quality fair

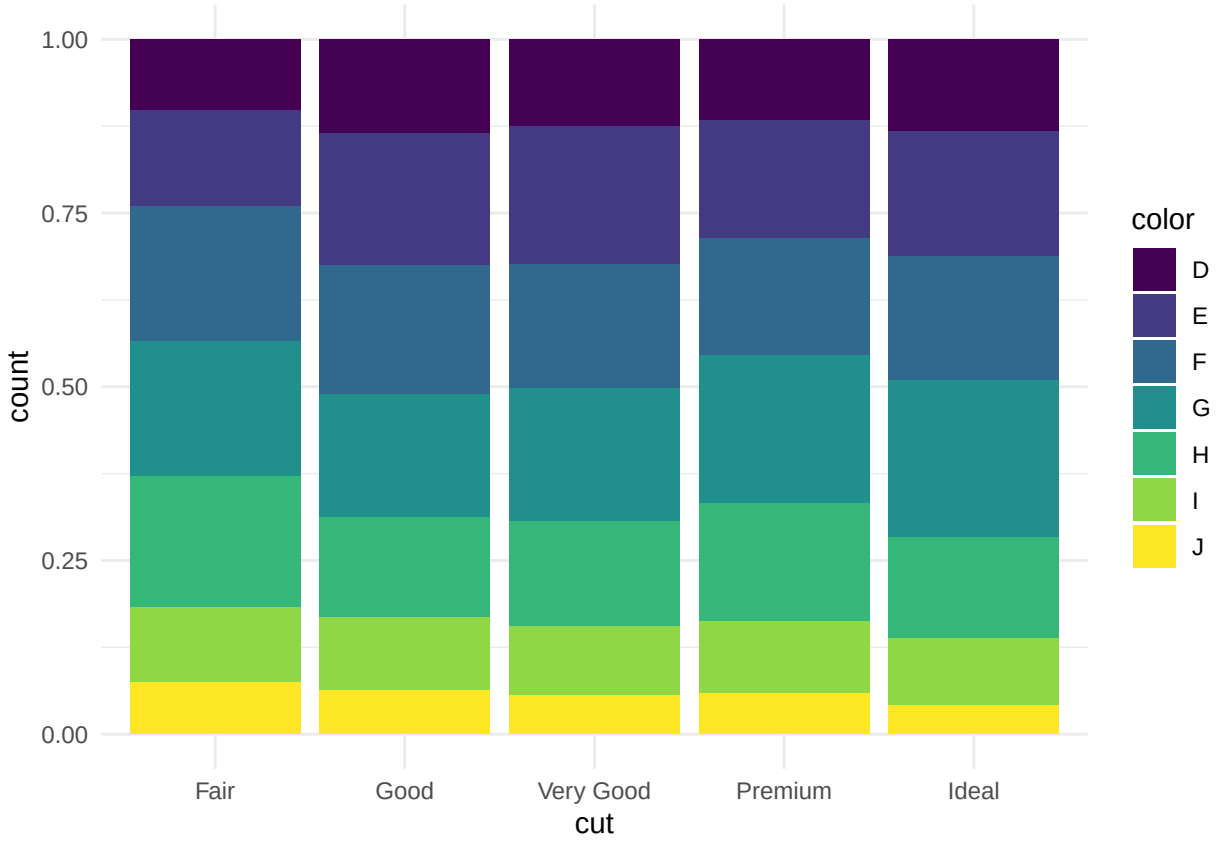
I need to show other chart but it seem above chart

```
ggplot( diamonds,
 mapping = aes(x=cut,y=mean(price)))+
 geom_col(fill="#0E8388")+
 theme_minimal()+
 scale_y_continuous(labels = scales::comma)# Format y-axis labels with commas
```



price average follow quality . High quality high price .low quality low price

```
ggplot(diamonds,
 mapping = aes(x=cut,fill=color))+
 geom_bar(position = "fill")+ #position have  fill  stack indensity   dodge
 theme_minimal()
```



geom_bar in R can use function position for different output.

this chart show volume percentage. we can see volume for quality by color

4.Box plot and violin plot

Box charts and violin charts.

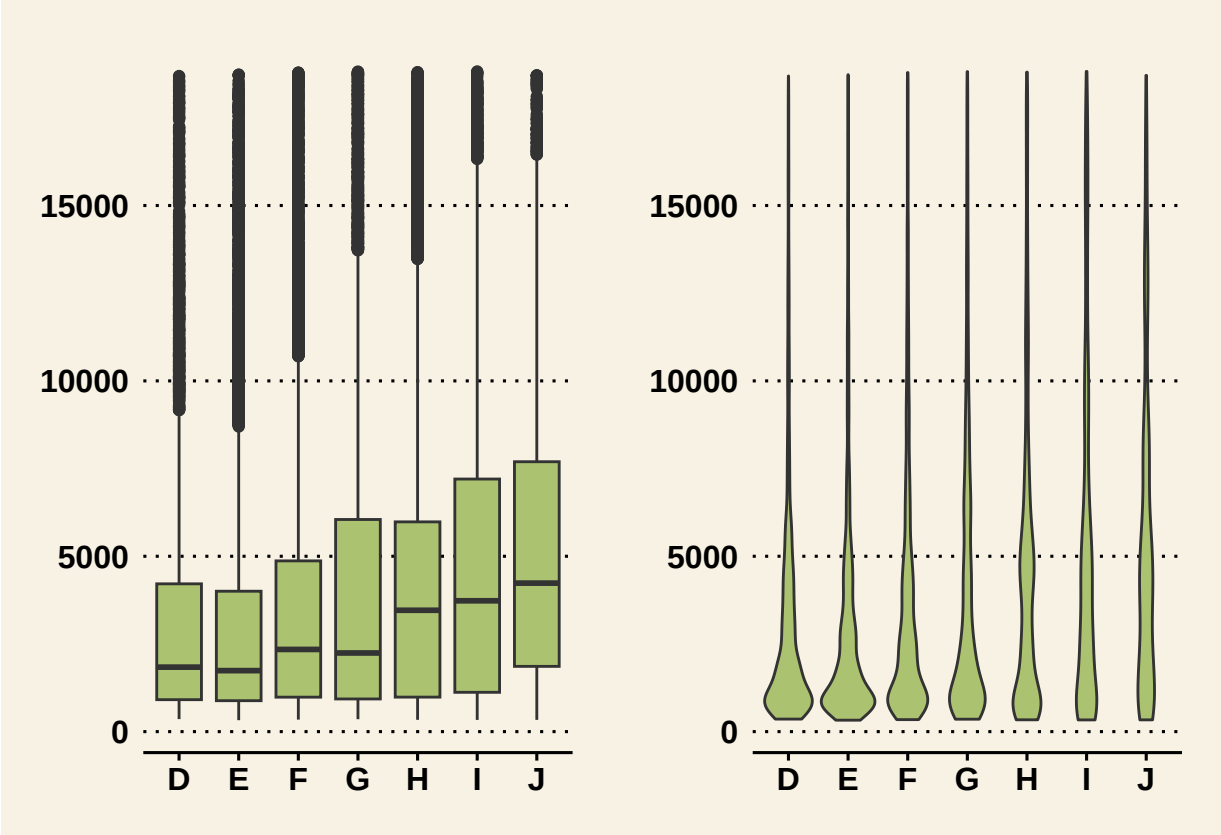
very used in data job .

We should to know and use it

```
library(ggthemes)
library(patchwork)
P1<- ggplot(diamonds,
       aes(x=color,y=price)) +
  geom_boxplot(fill="#ABC270")+
  theme_wsj()

P2<- ggplot(diamonds,
            aes(x=color,y=price)) +
  geom_violin(fill="#ABC270")+
  theme_wsj()

P1+P2 # from library(patchwork)        +                    /
```



boxplot we know about outlier.

we can see outlier but box plot can't see frequency.

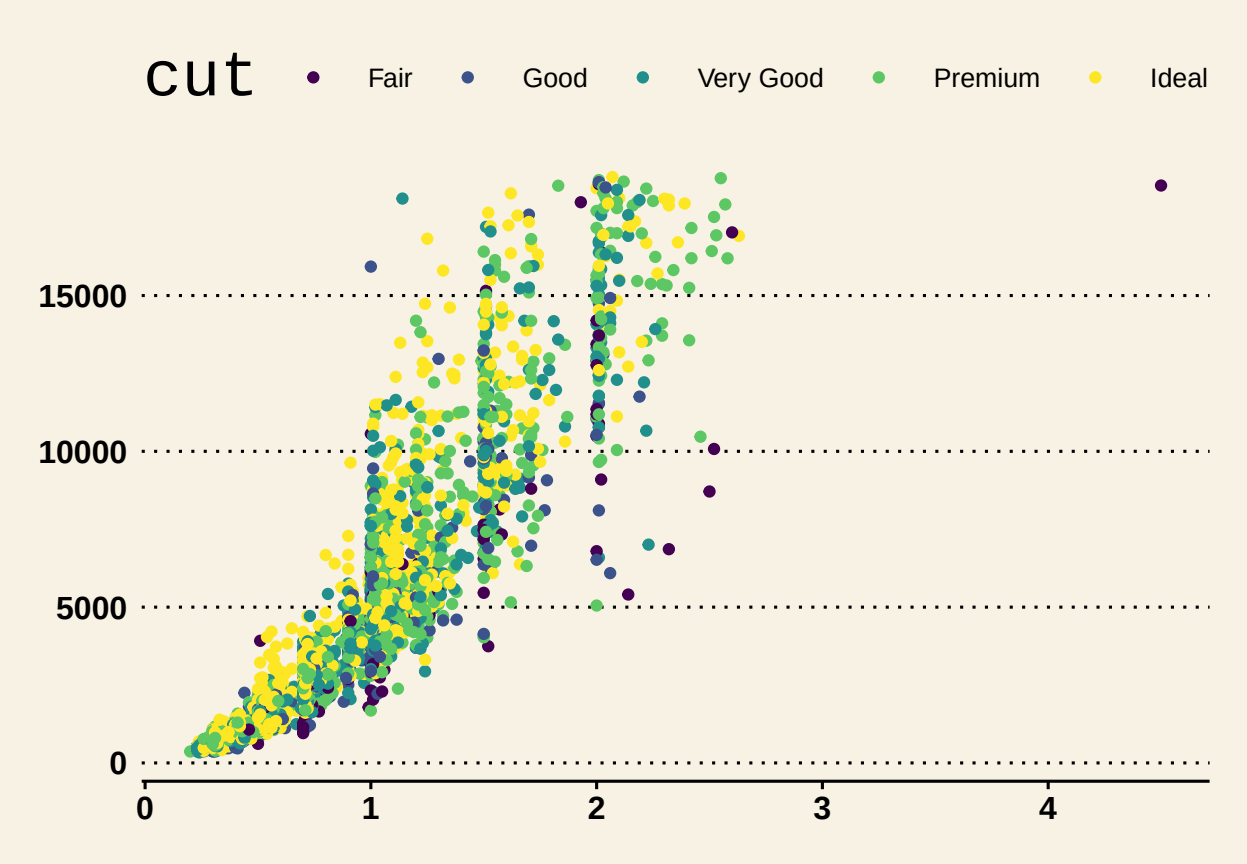
chart above so many outlier can cut outlier for quality data.

violin can see frequency but not show outlier.

5.point

most very popular to use

```
ggplot(sample_frac(diamonds,0.08),
       aes(x=carat,y=price,col=cut))+
      geom_point()+
  theme_wsj()
```

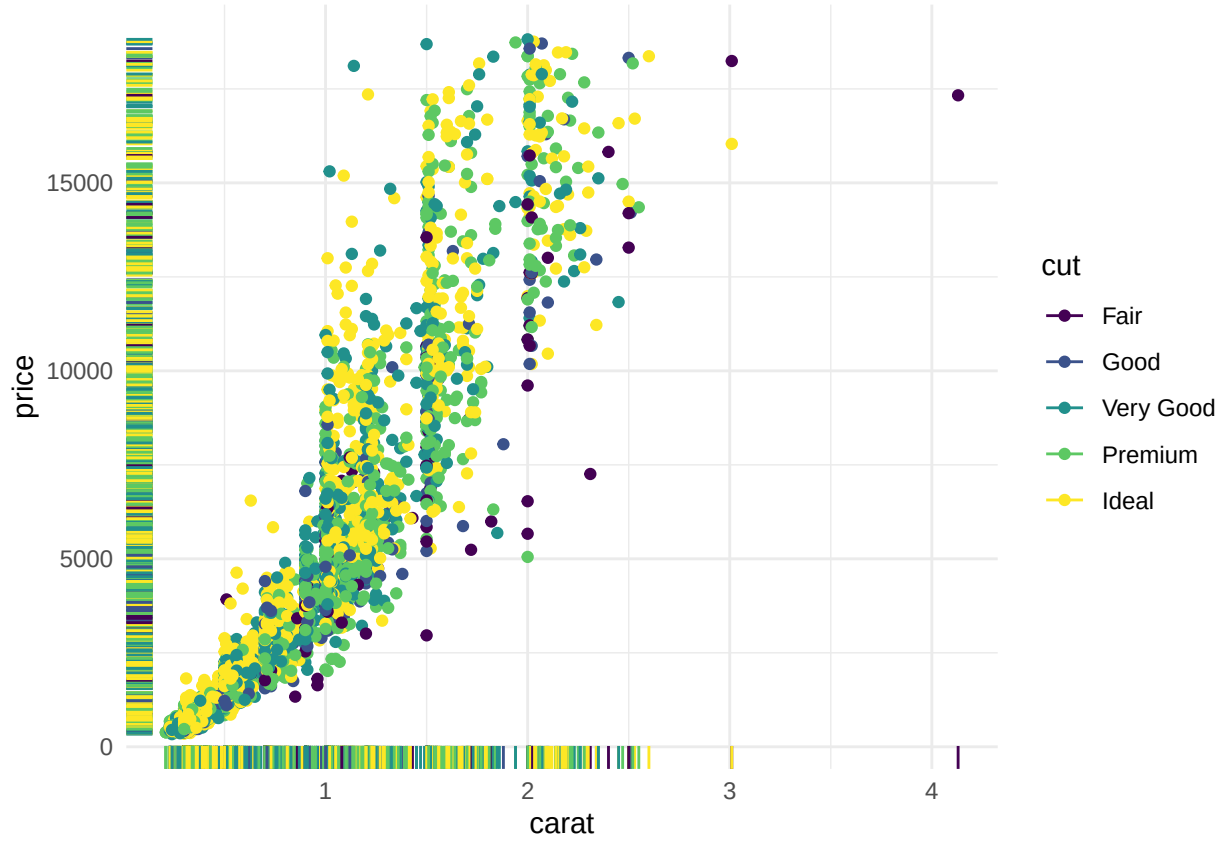


this chart show frequency between two columns.

In data low price and low weight most frequency .

but we can see quality by color

```
ggplot(sample_frac(diamonds,0.08),
       aes(x=carat,y=price,col=cut))+
      geom_point()+
      geom_rug()+
  theme_minimal()
```



can put more chart in chart . in chart above we see frequency very clery