R Viz

Punchika

2025-05-14

R Visualization

My ggplot2 project

data visualization or charts

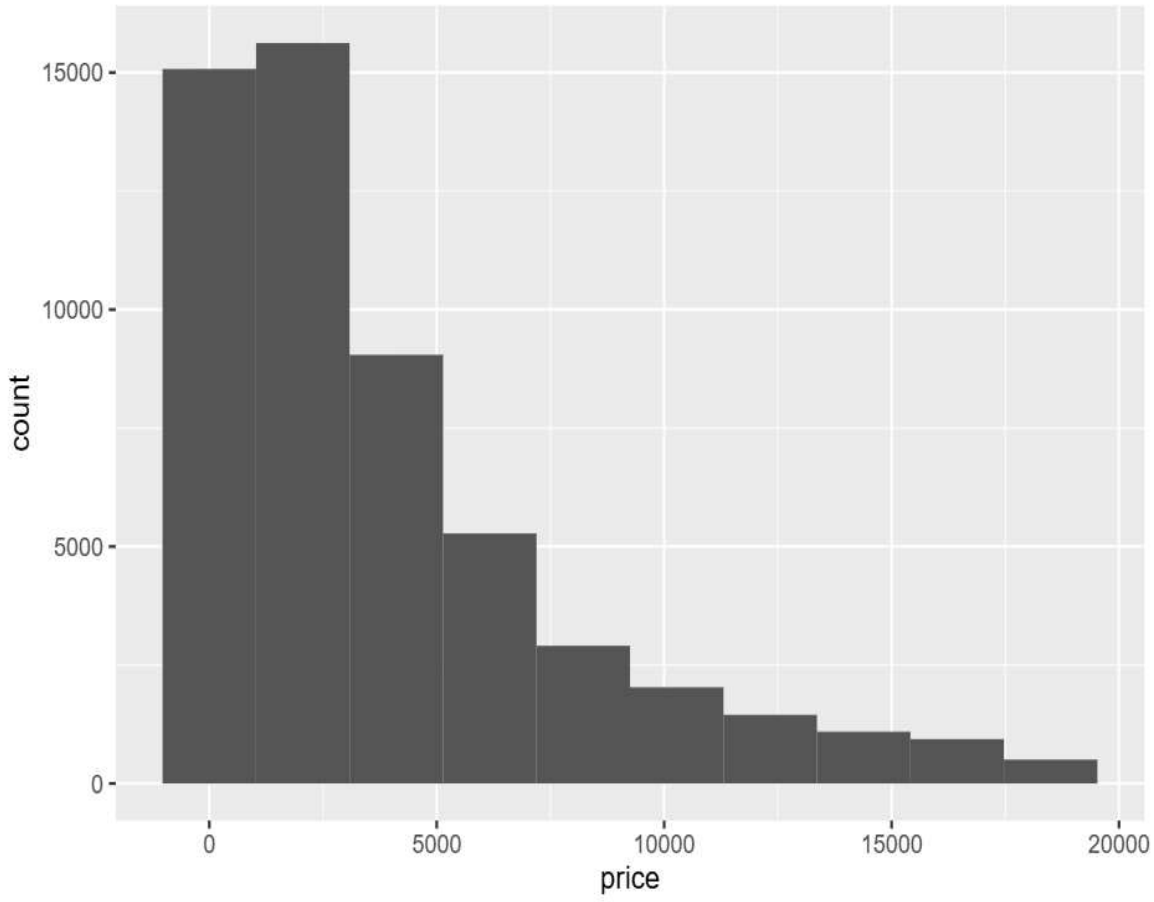
```
library(ggplot2)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'
```

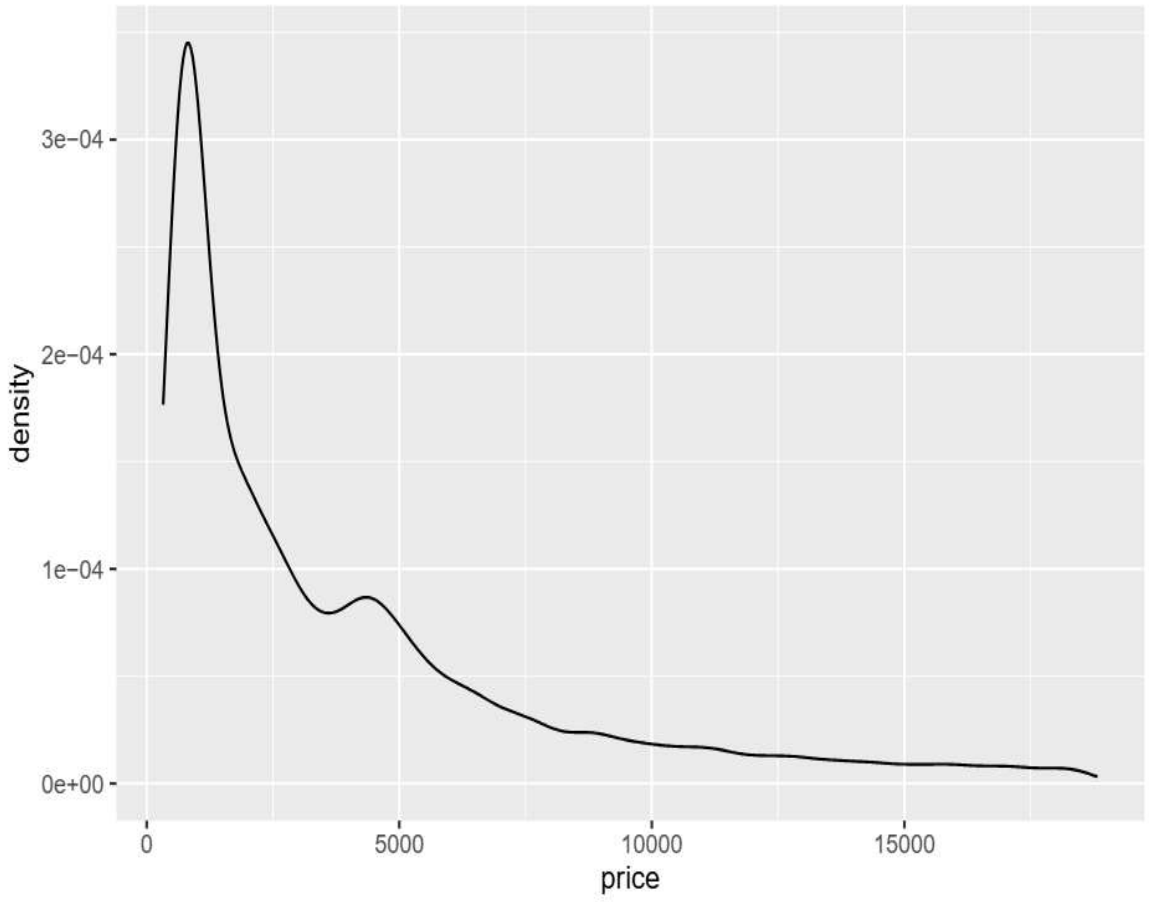
```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
base <- ggplot(data = diamonds,
       mapping = aes(x = price))

base + geom_histogram(bins = 10)
```




```
base + geom_density()
```

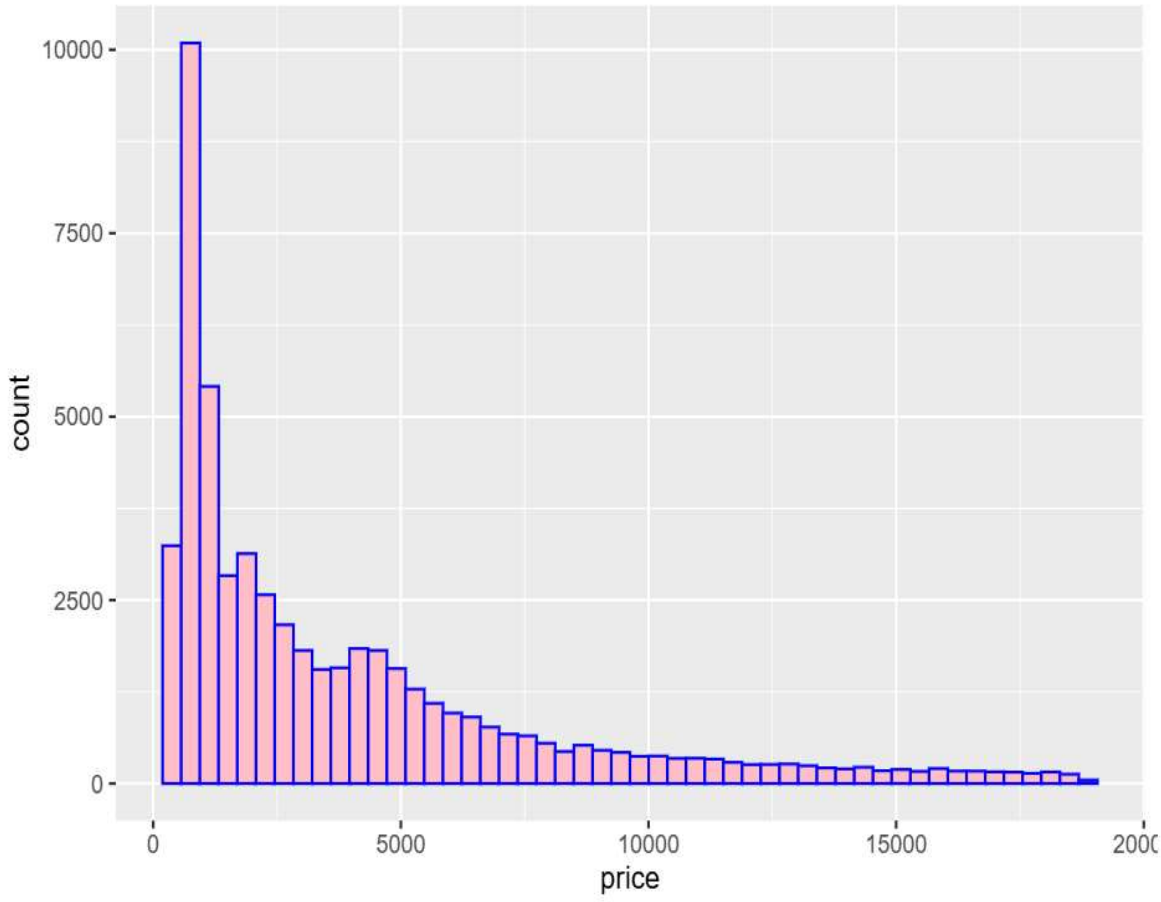


Mapping (MAp data from dataframe into chart)

```
base <- ggplot(data = diamonds,
       mapping = aes(x = price))
```

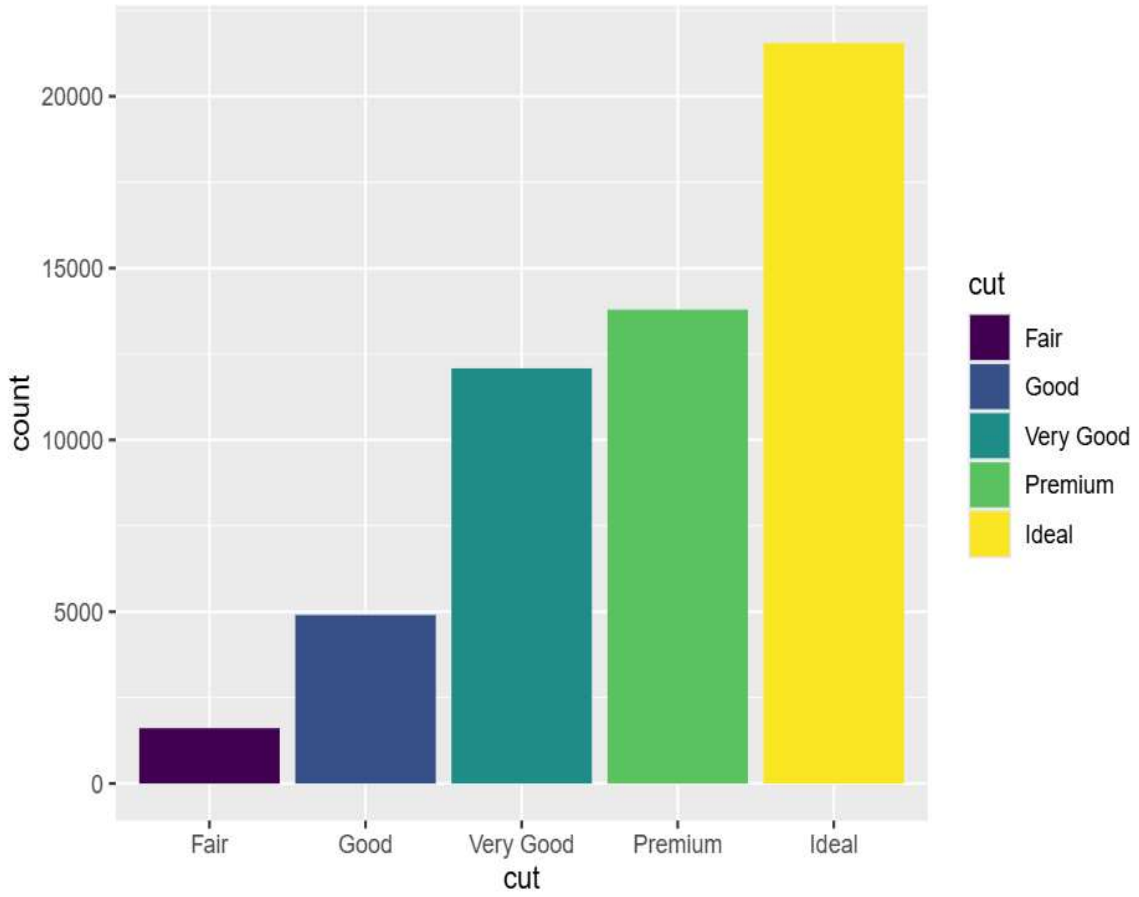
.### Setting (Set format)

```
base + geom_histogram(bins = 50,
                      fill = "pink",
                      color = "blue")
```



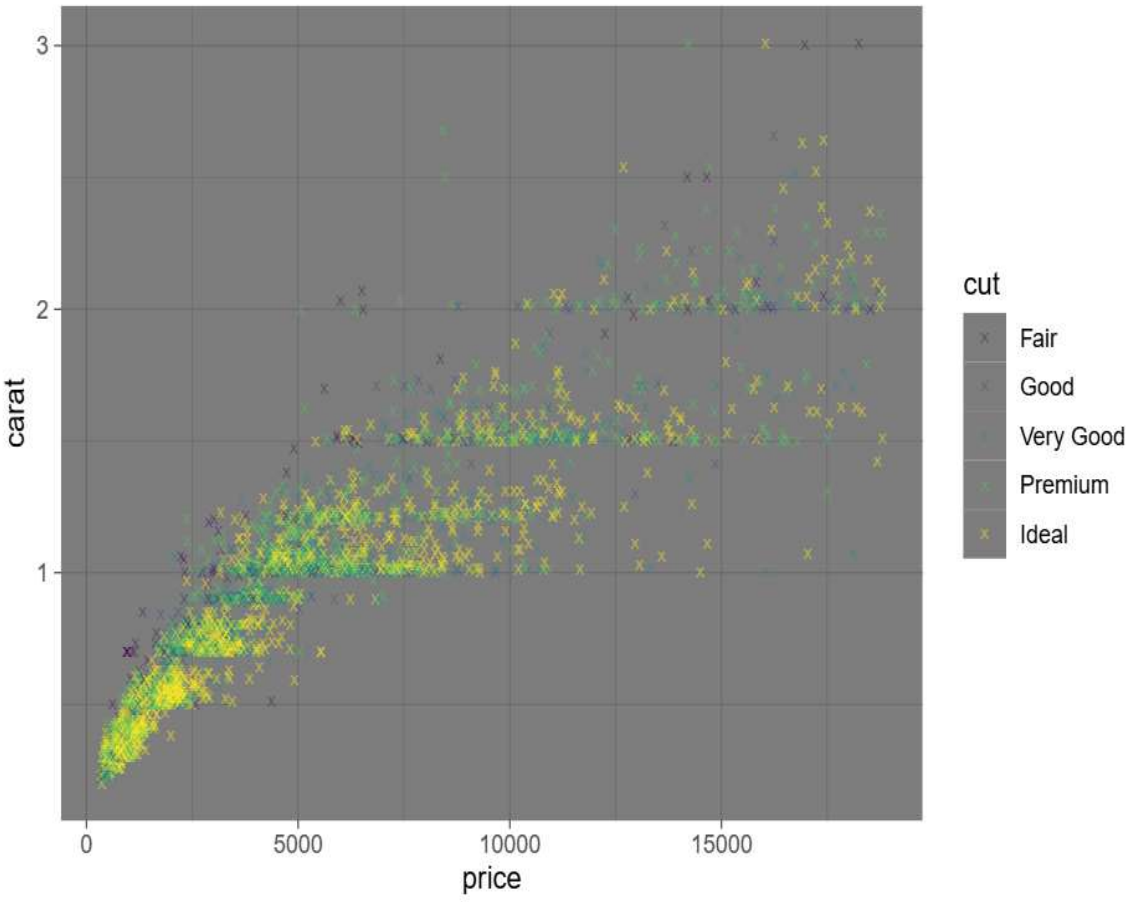
one variable, non-number factor

```
base2 <- ggplot(data = diamonds,
                        mapping = aes(x = cut))

base2 +
  geom_bar(mapping = aes(fill = cut))
```



two variables, number x number

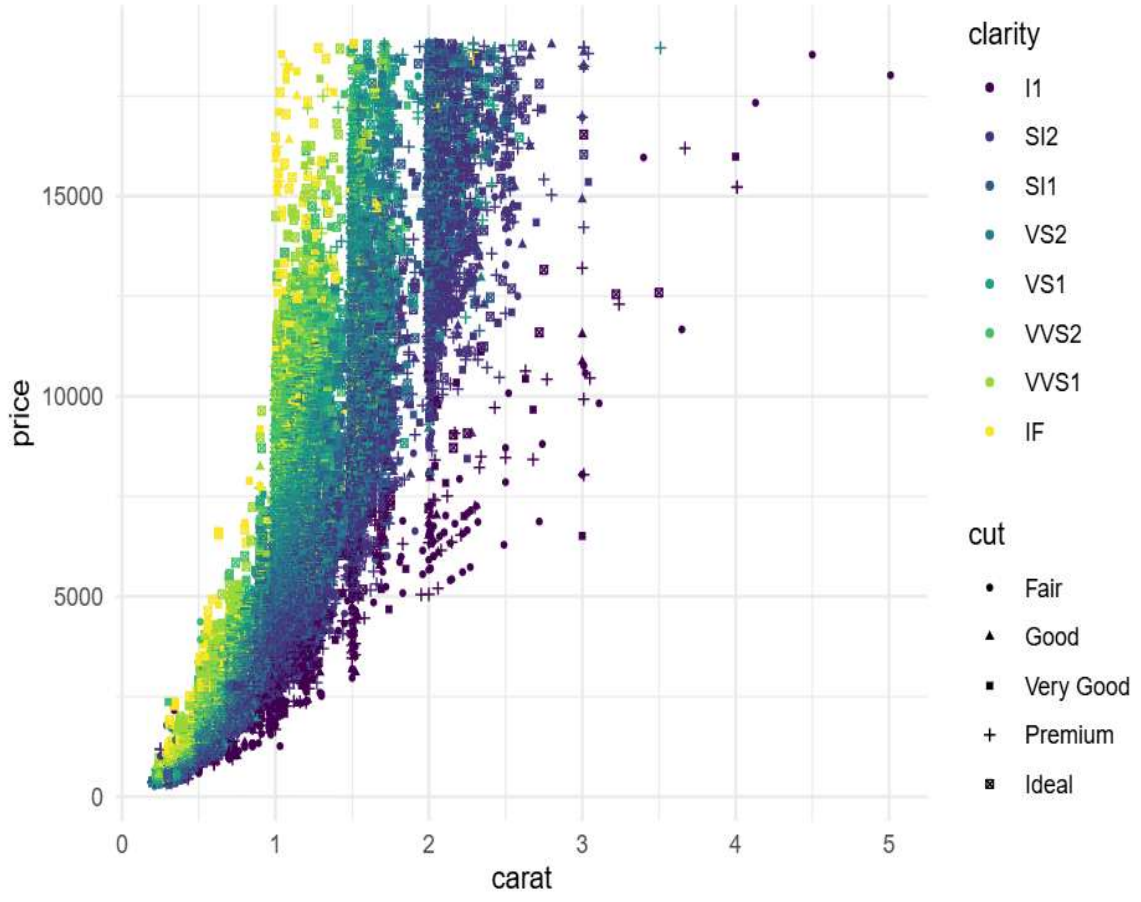
```
small_dia <- diamonds %>%
  sample_n(5000)

ggplot(data = small_dia,
            mapping = aes(x = price, y = carat)) +
    geom_point(alpha = 0.5,shape = "X",
                        mapping = aes(color = cut)) +
    theme_dark()
```



add more variables

```
ggplot(diamonds,
       aes(carat, price)) +
  geom_point(aes(color = clarity,
                 shape = cut),
                     size = 1) +
  theme_minimal()
```

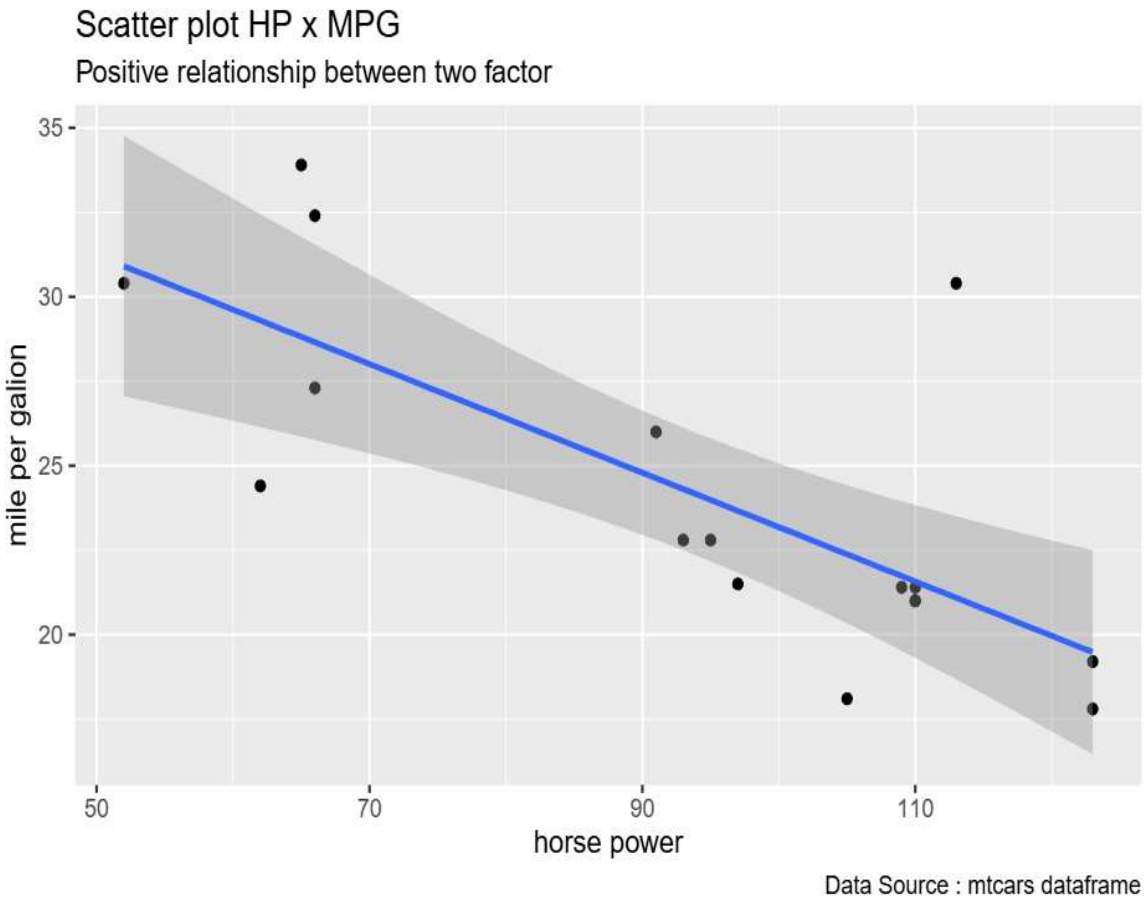
```
## Warning: Using shapes for an ordinal variable is not advised
```



Multiple chart in one + add title

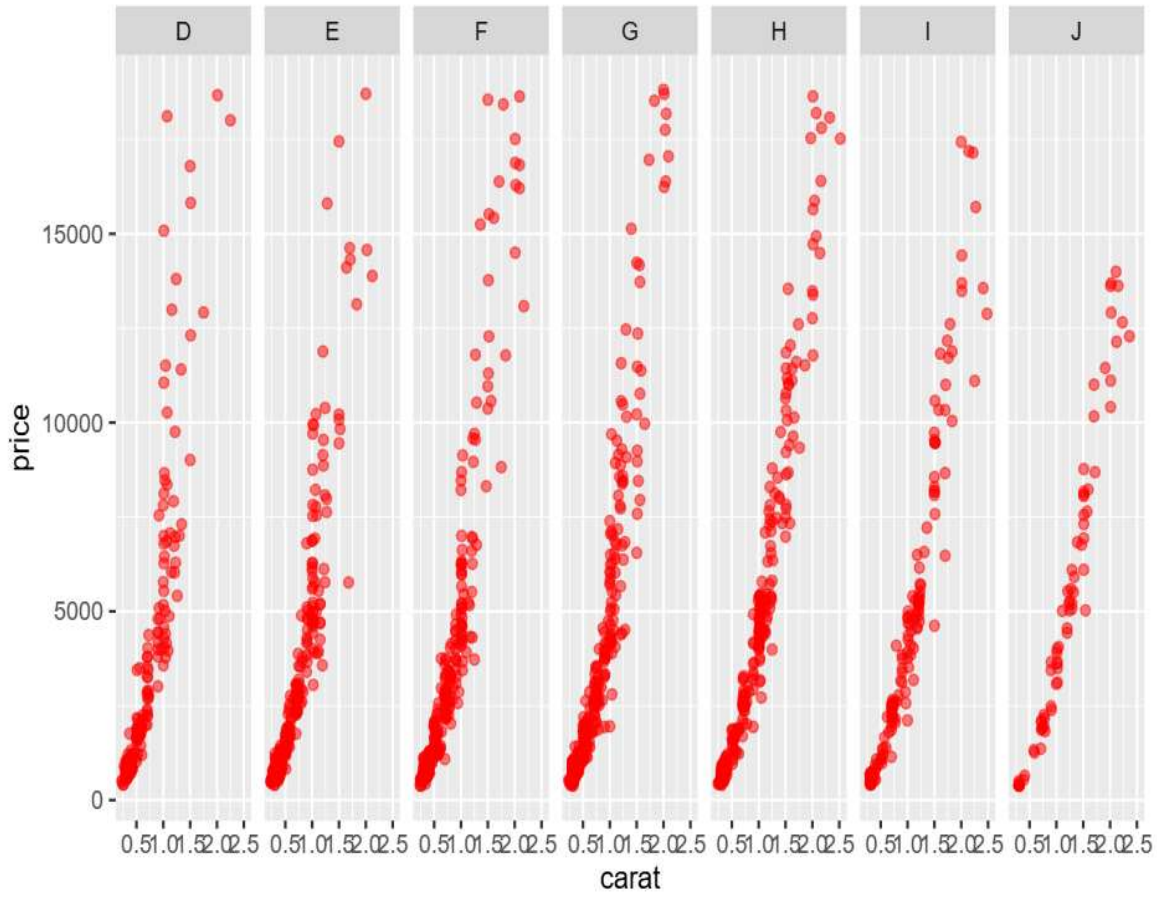
```
ggplot(mtcars %>%
         filter(hp < 130, mpg > 16),
       aes(hp, mpg)) +
  geom_point() +                          ## chart 1
  geom_smooth(method = "lm") +            ## chart 2
  labs(title = "Scatter plot HP x MPG",
       subtitle = "Positive relationship between two factor",
       caption = "Data Source : mtcars dataframe",
       x = "horse power",
       y = "mile per galion")
```

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



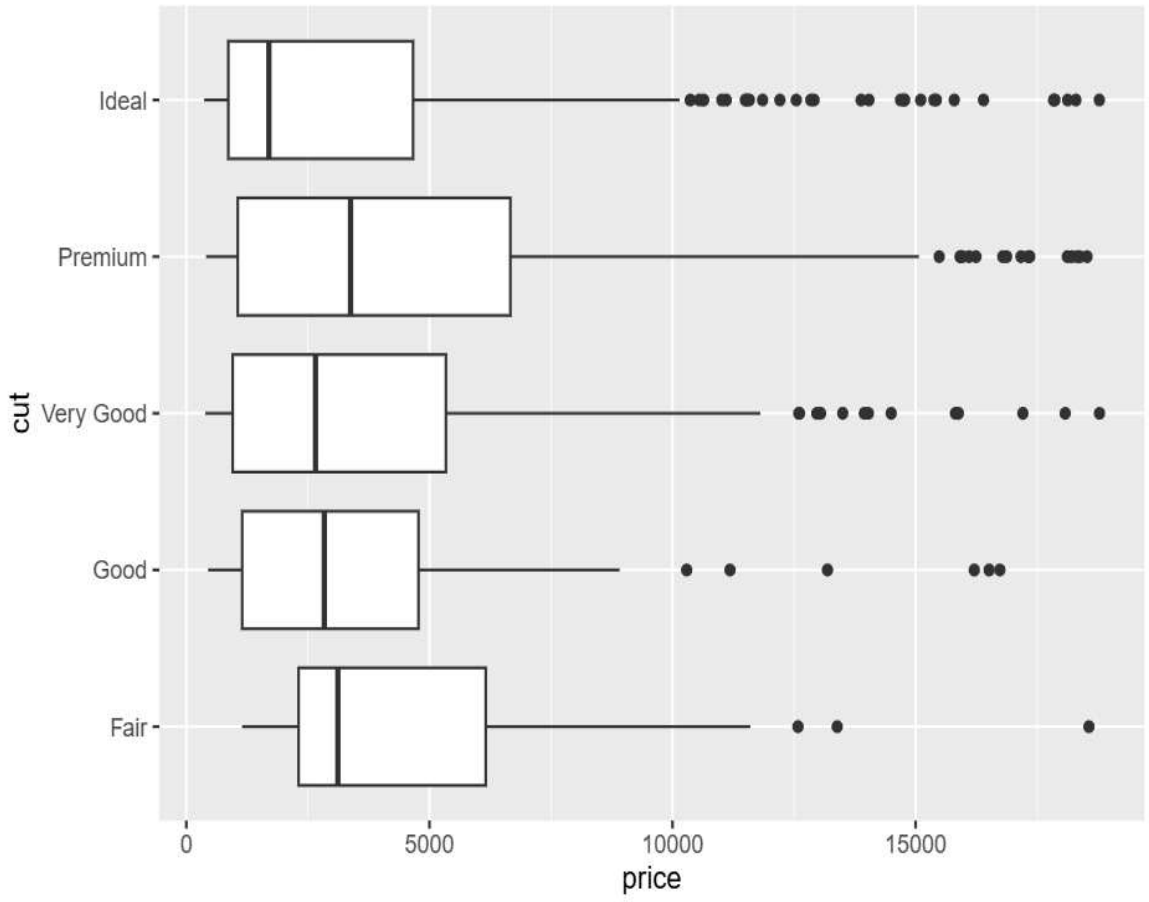
faceting breaks down big chart

```
ggplot(diamonds %>% sample_n(1500),
       aes(x = carat, y = price)) +
  geom_point(color = "red", alpha = 0.5) +
  facet_wrap(~color, ncol = 7)
```



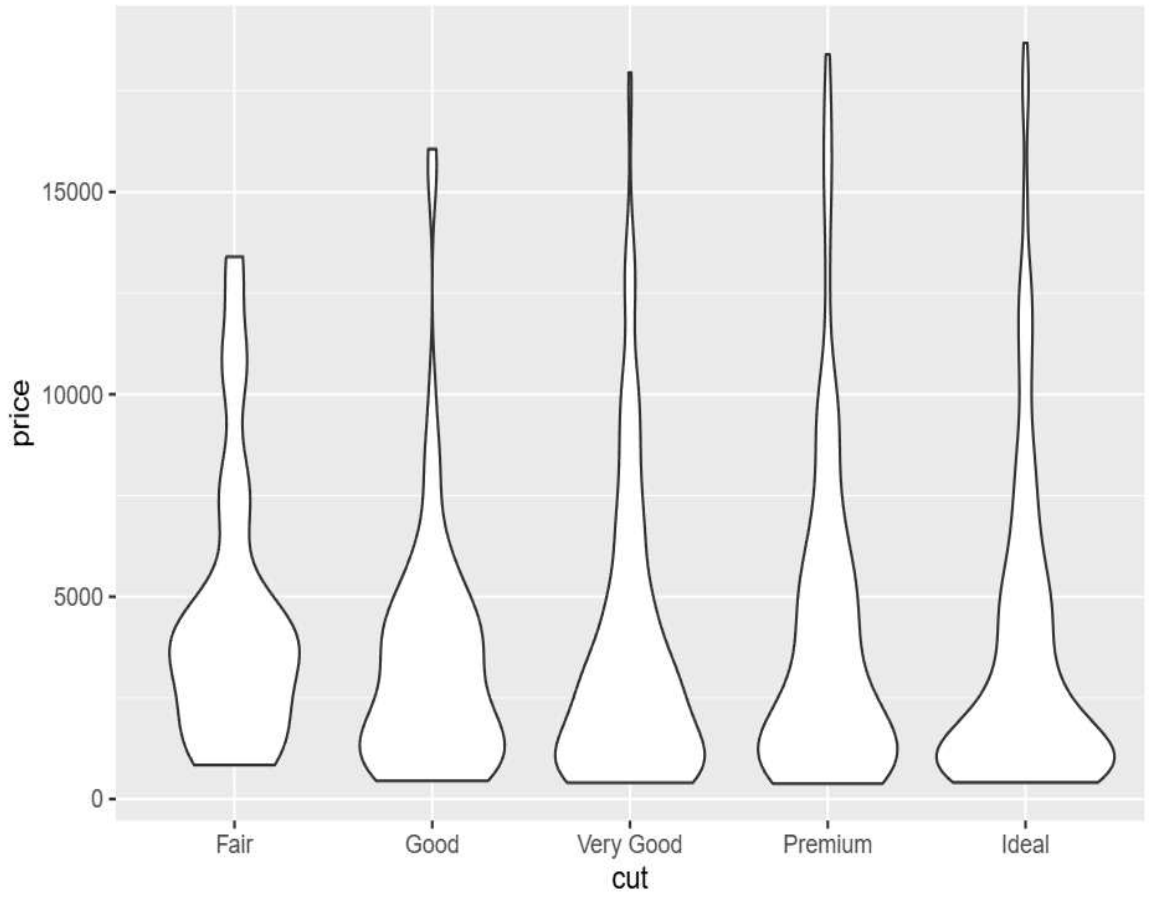
Box plot

```
ggplot(diamonds %>% sample_n(1000),
       aes(y = cut, x = price)) +
  geom_boxplot()
```



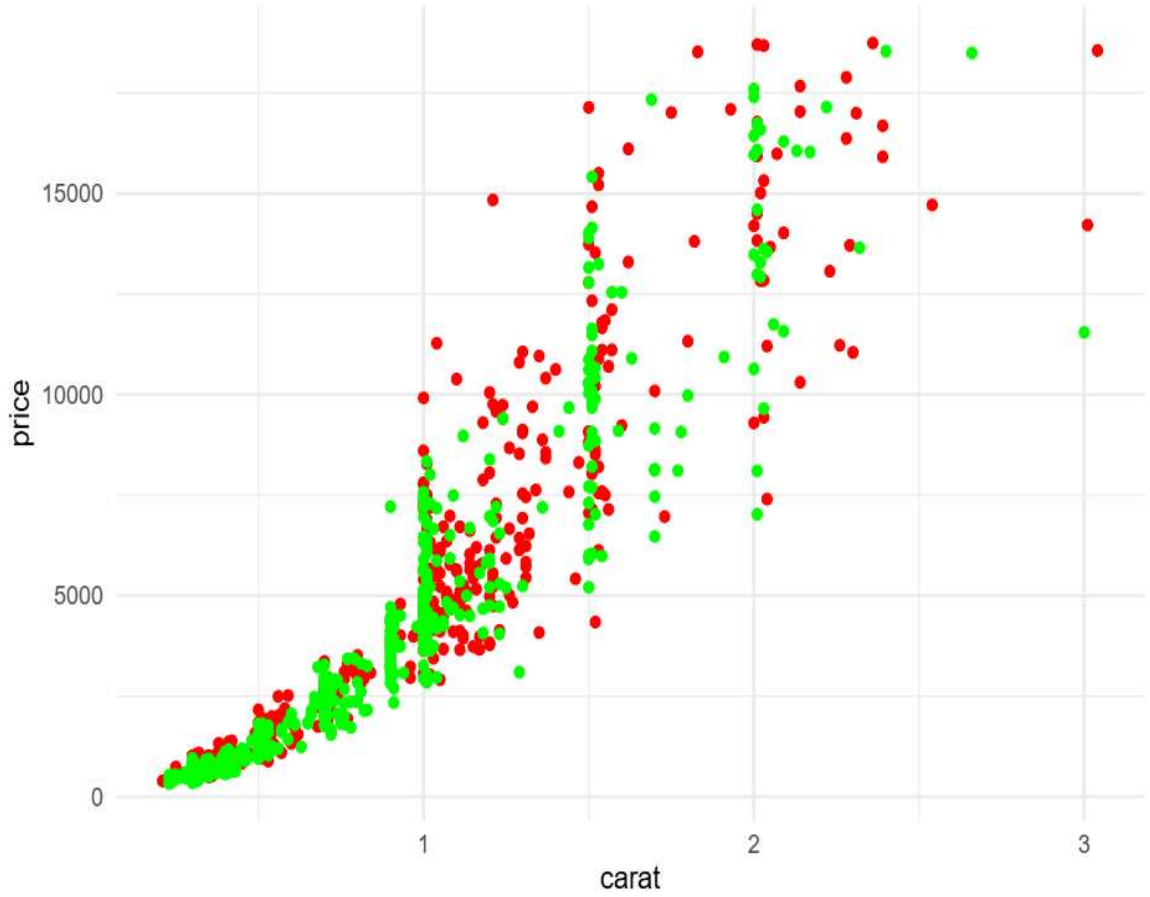
Violin plot

```
ggplot(diamonds %>% sample_n(1000),
       aes(x = cut, y = price)) +
  geom_violin()
```



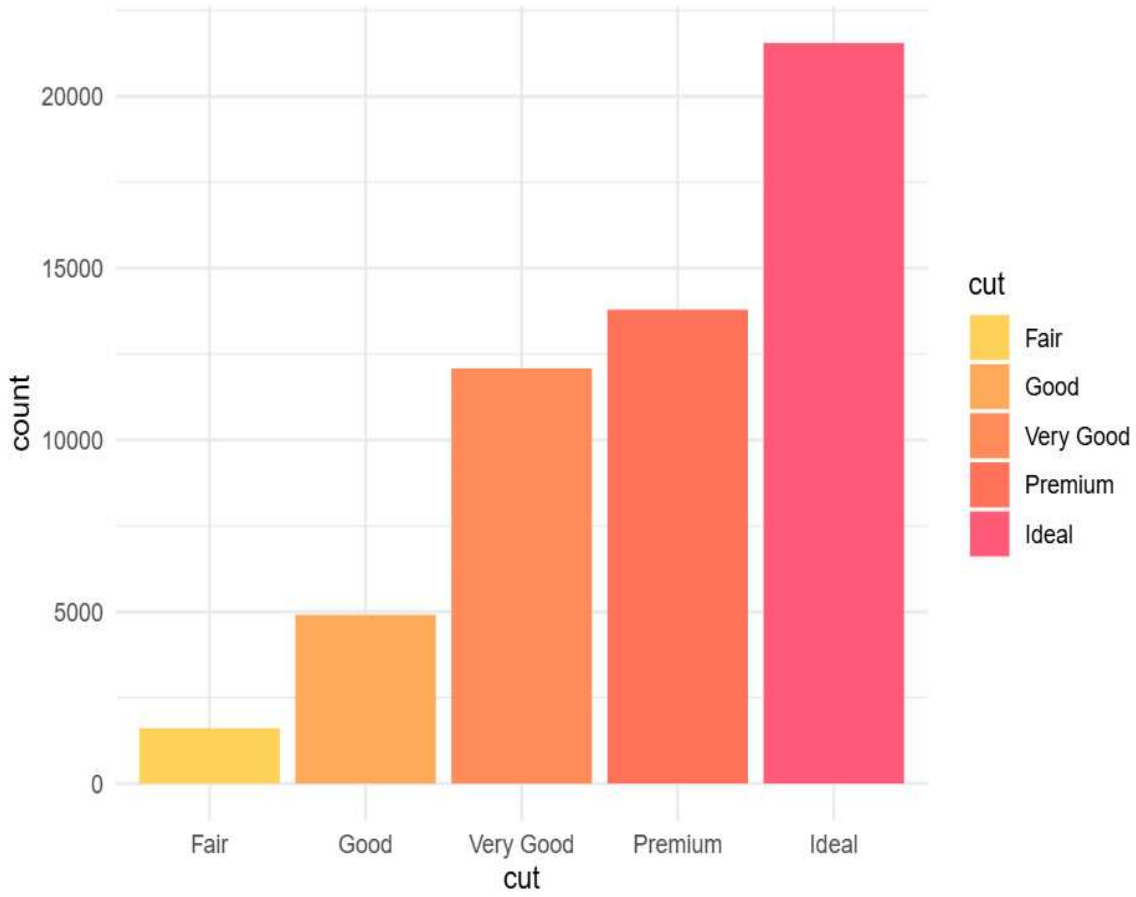
Multiple data set

```
## Set1
premium_di <- diamonds %>%
  filter(cut == "Premium") %>%
  sample_n(500)
## Set2
good_di <- diamonds %>%
  filter(cut == "Good") %>%
  sample_n(500)

ggplot() +
  geom_point(data = premium_di,
             mapping = aes(carat,price),
             color = "red") +
  geom_point(data = good_di,
             mapping = aes(carat,price),
             color = "green") +
  theme_minimal()
```



Manual change color

```
diamonds %>%
  ggplot(aes(cut , fill = cut)) +
  geom_bar() +
  theme_minimal() +
  scale_fill_manual(values = c(
    "#ffd45e",      ## hex color
    "#ffaf5e",
    "#ff8e5e",
    "#ff735e",
    "#ff5e79"
  ))
```


20000
15000
10000
5000
0
count
Fair
Good
Very Good
Premium
Ideal
cut
cut
Fair
Good
Very Good
Premium
Ideal