R Notebook

PLUM

Hello world

I'm learning R markdown to create document.

```
## [1] "Hello world"
```

```
##                    mpg cyl disp  hp drat    wt  qsec vs am gear carb
## Mazda RX4         21.0   6  160 110 3.90 2.620 16.46  0  1    4    4
## Mazda RX4 Wag     21.0   6  160 110 3.90 2.875 17.02  0  1    4    4
## Datsun 710        22.8   4  108  93 3.85 2.320 18.61  1  1    4    1
## Hornet 4 Drive    21.4   6  258 110 3.08 3.215 19.44  1  0    3    1
## Hornet Sportabout 18.7   8  360 175 3.15 3.440 17.02  0  0    3    2
## Valiant           18.1   6  225 105 2.76 3.460 20.22  1  0    3    1
```

Create New Visualization

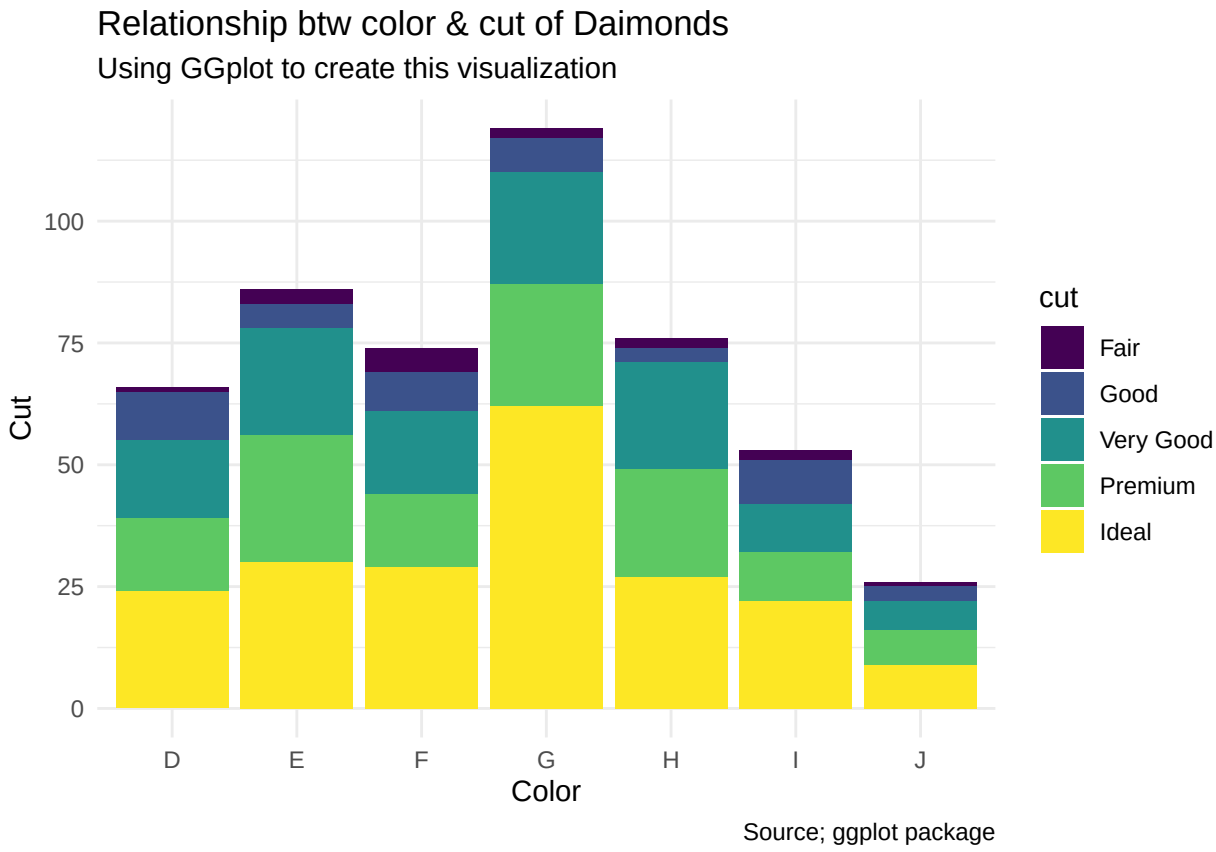
This is a stacked bar chart to show color distribution

```
library(tidyverse)
```

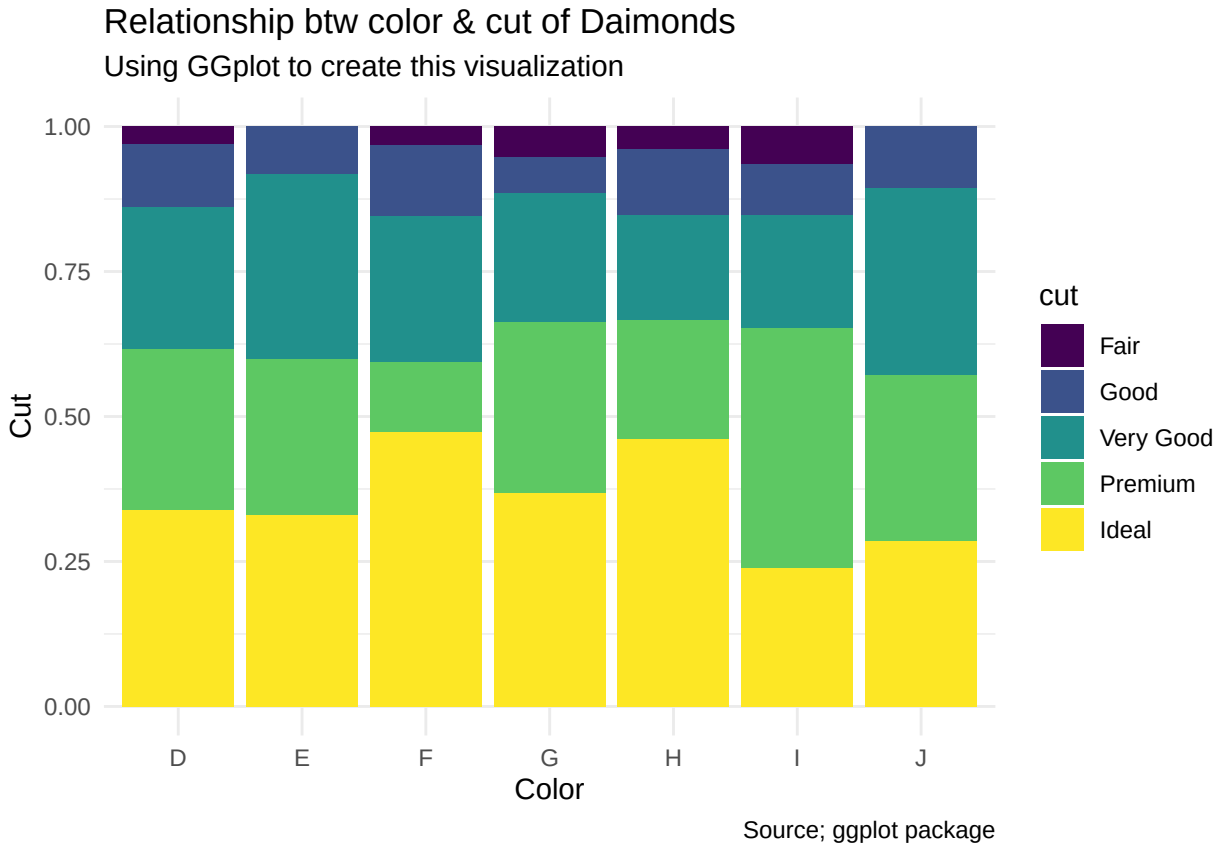
```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
## v ggplot2 3.3.6      v purrr   0.3.4
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.4.1
## v readr   2.1.2      v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(ggplot2)
library(dplyr)

ggplot(sample_n(diamonds, 500), aes(color, fill = cut))+ geom_bar() +
  theme_minimal()+
  labs(
    title = "Relationship btw color & cut of Daimonds",
    x = "Color",
    y = "Cut",
    subtitle = "Using GGplot to create this visualization",
    caption = "Source; ggplot package")
```




```
ggplot(sample_n(diamonds, 500), aes(color, fill = cut))+ geom_bar(position = "fill")+ theme_minimal()+
  labs(
    title = "Relationship btw color & cut of Daimonds",
    x = "Color",
    y = "Cut",
    subtitle = "Using GGplot to create this visualization",
    caption = "Source; ggplot package")
```



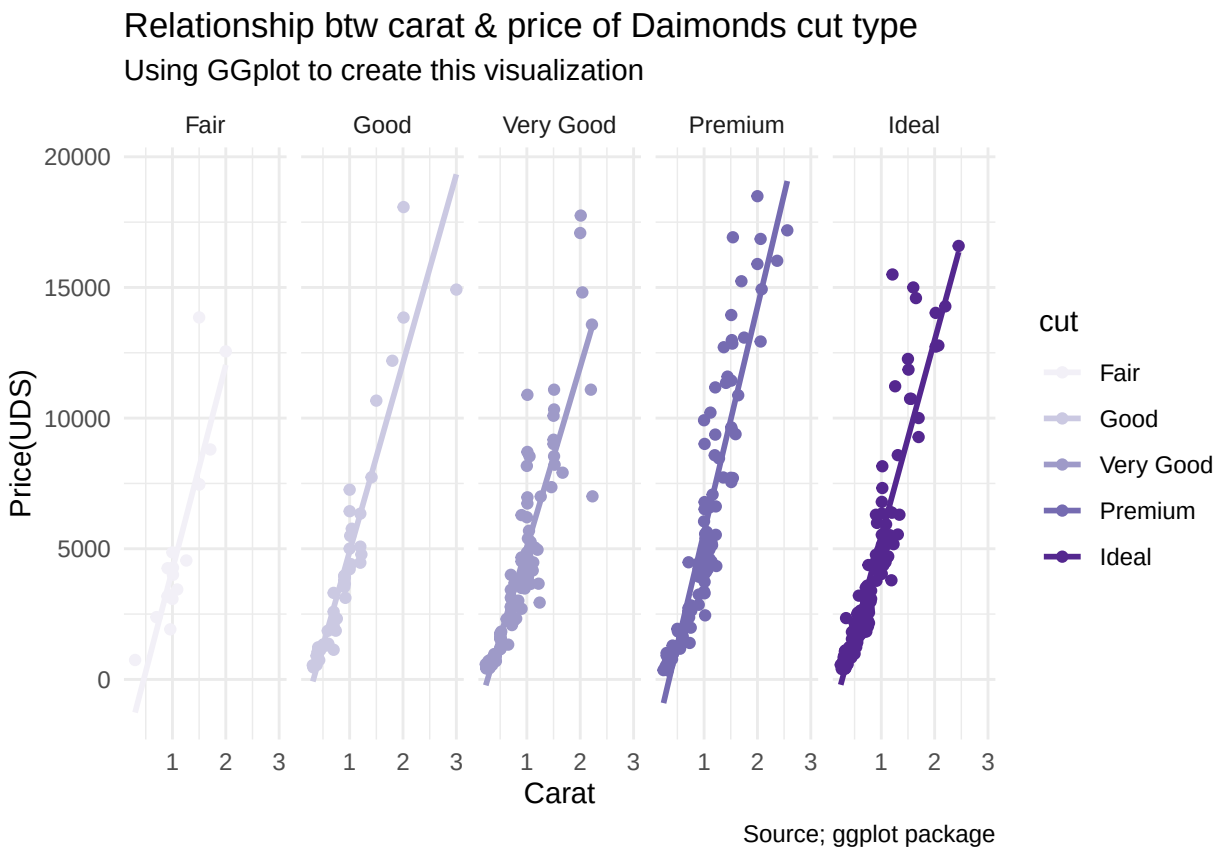
Homework

```
head(diamonds)
```

```
## # A tibble: 6 x 10
##   carat cut       color clarity depth table price     x     y     z
##   <dbl> <ord>     <ord> <ord>   <dbl> <dbl> <int> <dbl> <dbl> <dbl>
## 1  0.23 Ideal     E     SI2      61.5    55   326  3.95  3.98  2.43
## 2  0.21 Premium   E     SI1      59.8    61   326  3.89  3.84  2.31
## 3  0.23 Good      E     VS1      56.9    65   327  4.05  4.07  2.31
## 4  0.29 Premium   I     VS2      62.4    58   334  4.2   4.23  2.63
## 5  0.31 Good      J     SI2      63.3    58   335  4.34  4.35  2.75
## 6  0.24 Very Good J     VVS2     62.8    57   336  3.94  3.96  2.48
```

```
set.seed(42)
p <- ggplot(sample_n(diamonds, 500), aes(carat, price, color = cut))
p + geom_point()+ geom_smooth(method ="lm", se = F) +
  labs(
    title = "Relationship btw carat & price of Daimonds cut type",
    x = "Carat",
    y = "Price(UDS)",
    subtitle = "Using GGplot to create this visualization",
    caption = "Source; ggplot package"
  )+theme_minimal()+ scale_color_brewer(palette = "Purples")+facet_wrap(~cut, ncol =5)
```

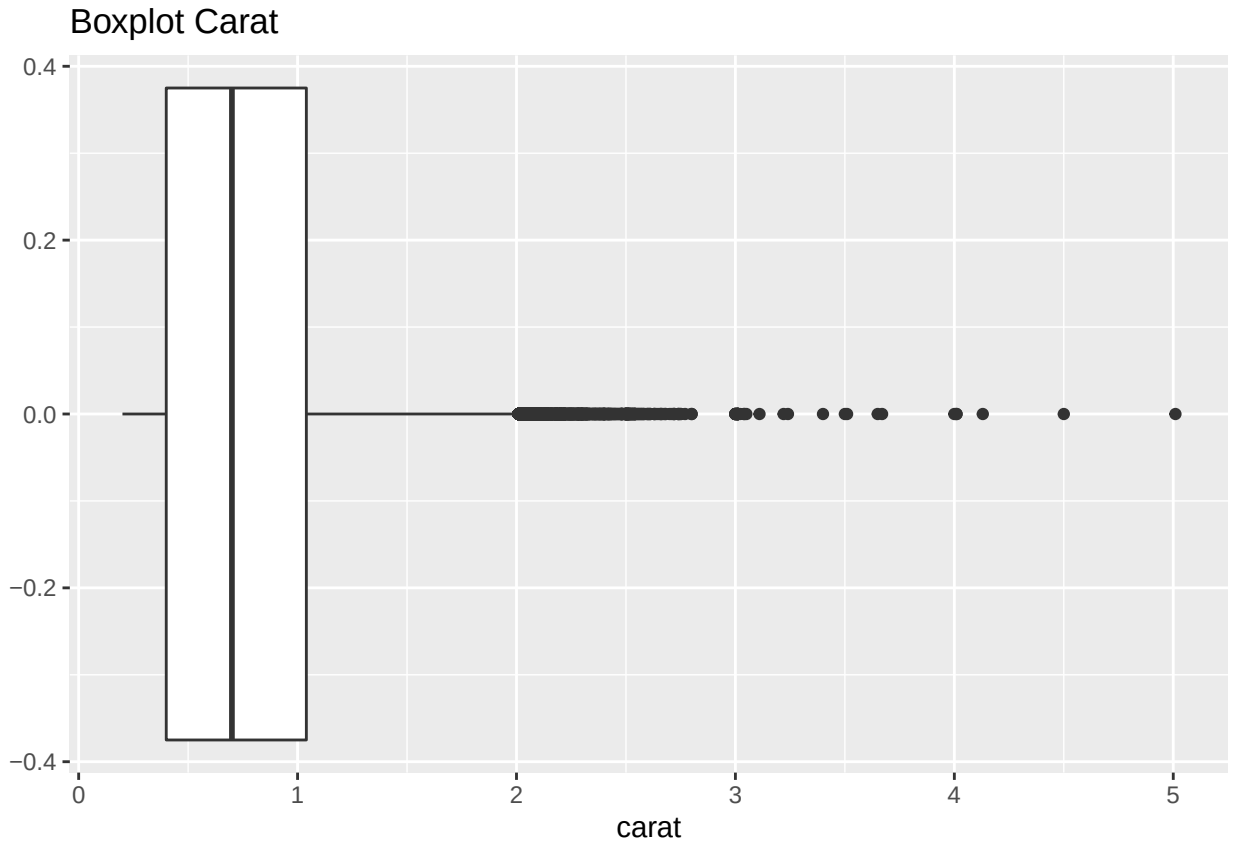
```
## `geom_smooth()` using formula 'y ~ x'
```



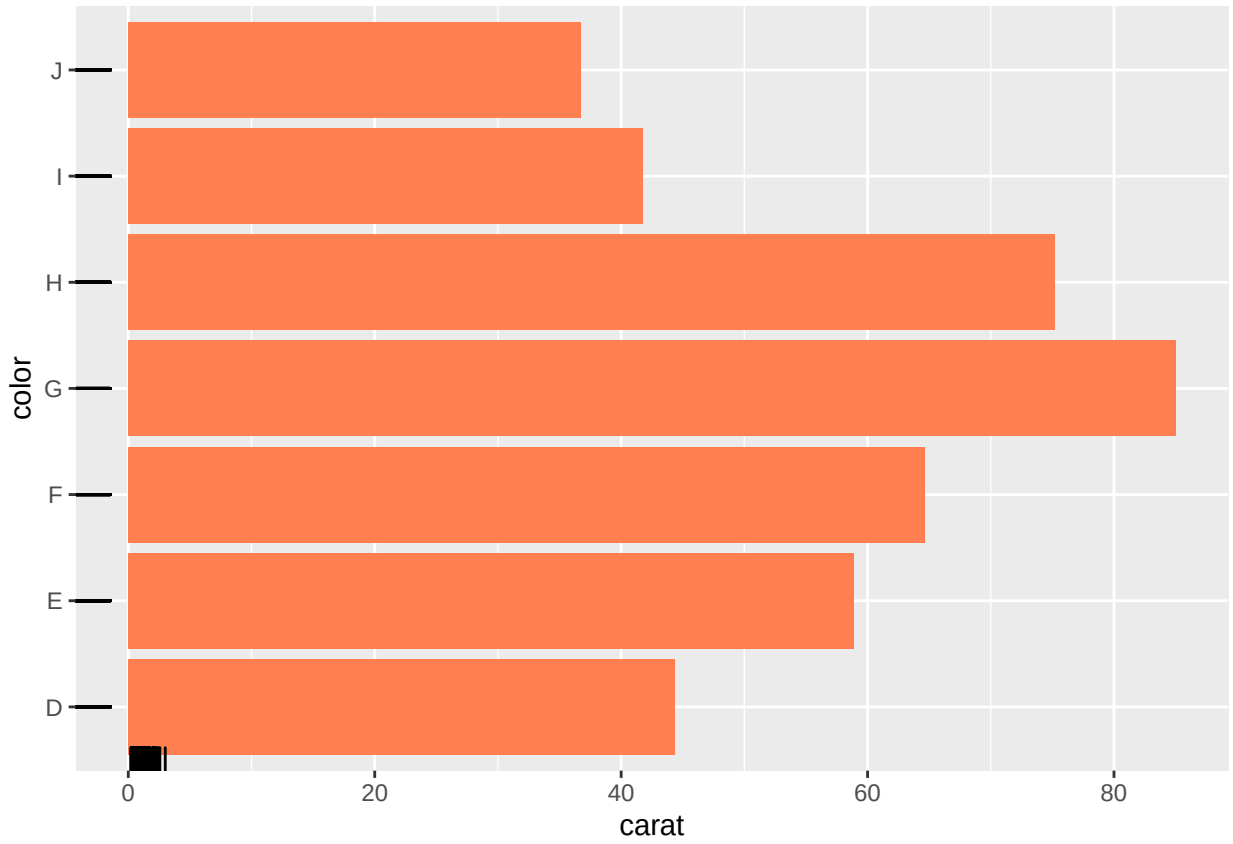

```
diamonds %>%
  summarise(
    min_carat = min(carat),
    max_carat = max(carat),
    q1 = quantile(carat, 0.25),
    q2 = quantile(carat, 0.50),
    q3 = quantile(carat, 0.75)
  )
```

```
## # A tibble: 1 x 5
##   min_carat max_carat    q1    q2    q3
##       <dbl>     <dbl> <dbl> <dbl> <dbl>
## 1       0.2      5.01   0.4   0.7  1.04
```

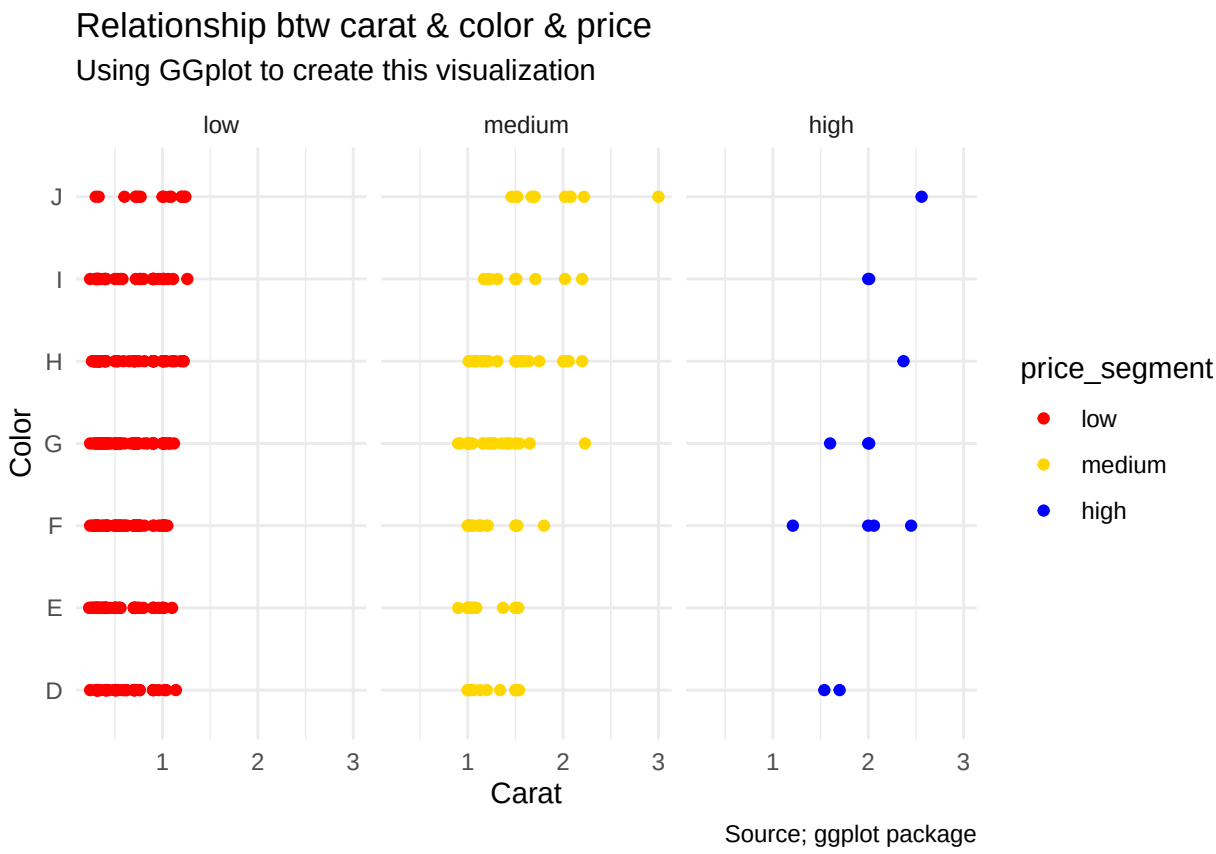
```
ggplot(diamonds, aes(carat))+
  geom_boxplot()+ labs(
    title = "Boxplot Carat")
```




```
set.seed(42)
ggplot(sample_n(diamonds, 500), aes(carat,color))+geom_col(fill="coral") + geom_rug()
```




```
set.seed(42)
sample_n(diamonds, 500) %>%
  select(carat, color, price) %>%
  mutate(price_segment = case_when(
    price < 5000 ~ "low",
    price < 15000 ~ "medium",
    TRUE ~ "high"
  )) %>%
  mutate(price_segment = factor(
    price_segment,
    labels = c("low", "medium", "high"),
    levels = c("low", "medium", "high"),
    ordered = "TRUE")) %>%
  ggplot(aes(carat, color, color = price_segment)) +
  geom_point() +
  scale_color_manual(
    values = c("red", "gold", "blue"))+
  labs(
    title = "Relationship btw carat & color & price",
    x = "Carat",
    y = "Color",
    subtitle = "Using GGplot to create this visualization",
    caption = "Source; ggplot package"
  )+theme_minimal()+facet_wrap(~price_segment, ncol =5)
```


Relationship btw carat & color & price
Using GGplot to create this visualization
low
medium
high
J
I
H
G
F
E
D
Color
1
2
3
Carat
price_segment
low
medium
high
Source; ggplot package