Data Visualization with R

```
install.packages(c("tidyverse","patchwork","lubridate","ggrepel"))
```

```
## Installing packages into '/cloud/lib/x86_64-pc-linux-gnu-library/4.2'
## (as 'lib' is unspecified)
```

```
library(tidyverse)
```

```
## -- Attaching packages --------------------------------------- tidyverse 1.3.2 --
```

```
## v ggplot2 3.4.0      v purrr   1.0.0
## v tibble  3.1.8      v dplyr   1.0.10
## v tidyr   1.2.1      v stringr 1.5.0
## v readr   2.1.3      v forcats 0.5.2
## -- Conflicts ------------------------------------------ tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()
```

```
library(ggplot2)
library(dplyr)
library(RColorBrewer)
library(patchwork)
library(ggrepel)
```

```
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
```

Chart 1 : Number of Diamonds by cut (One Discrete Variable)

```
ggplot(diamonds, aes(cut)) +
  geom_bar(fill=c("#f2ebe9",
                  "#8c7462",
                  "#8e6248",
                  "#3f2a1d",
                  "#24150e")) +
  geom_text(stat = 'count',aes(label = after_stat(count)), vjust=-0.5) +
  labs(title = "No.of Diamond by cut") +
  theme_minimal() +
  theme(panel.grid.minor = element_blank(),
        axis.title.y = element_blank(),
        axis.text.y = element_blank(),
        axis.title.x = element_blank())
```

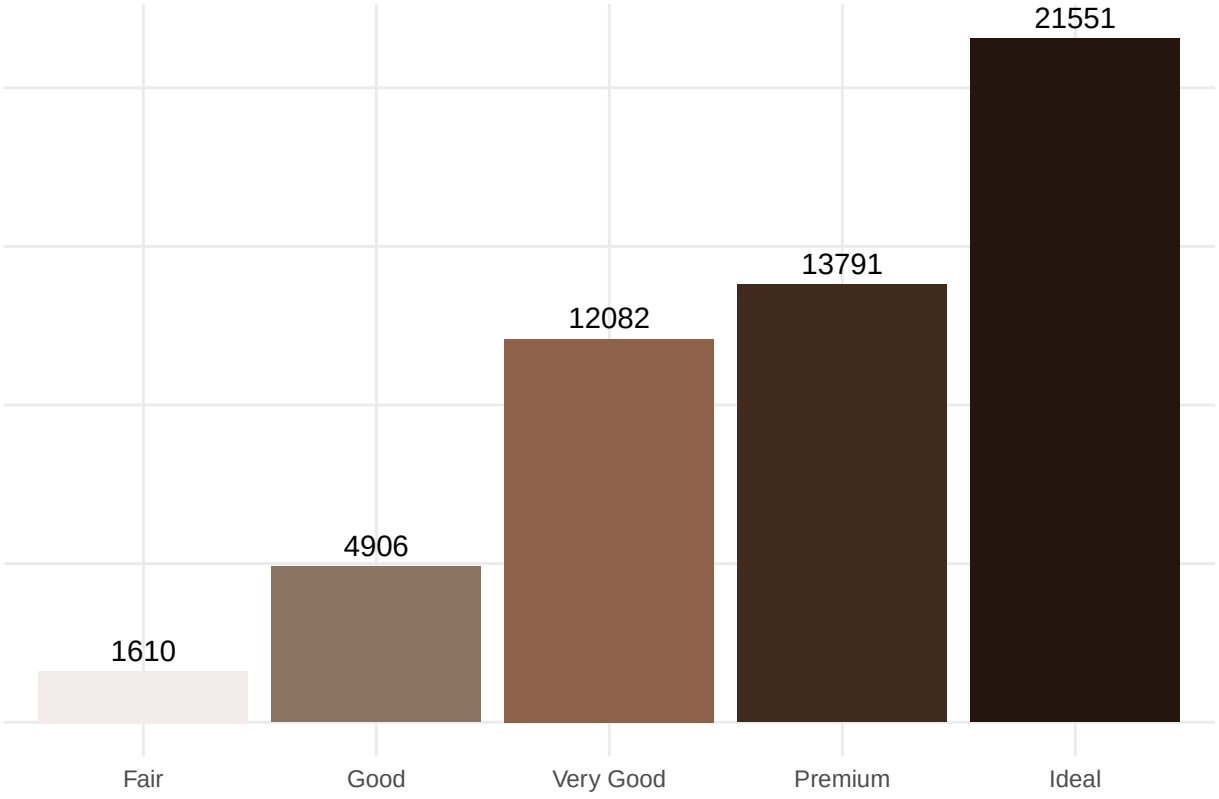


Chart 2 : Quantity of Diamonds by Price (One Continuous Variable)

```
ggplot(diamonds, aes(price)) +
  geom_density(color="#A2D3F0",linewidth=2) +
  theme_classic() +
  labs(title = "Quantity of Diamonds by price(USD)",
       y="Diamonds (Qty)",
       x="Price (USD)") +
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element_blank())
```

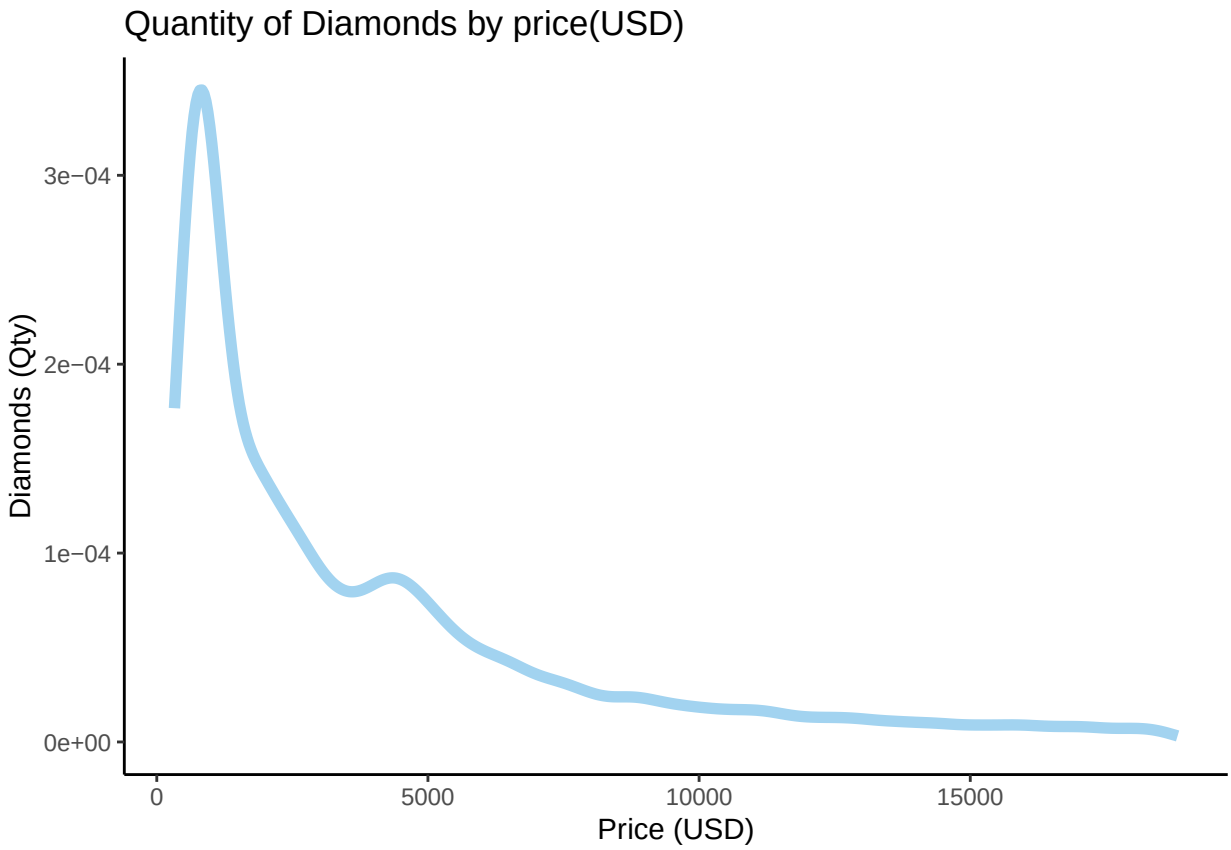


Chart 3 : Relationship between Color & Carat (Two Variables)

```
set.seed(36)
small_diamonds <- sample_n(diamonds, 2000)
ggplot(small_diamonds, aes(color,carat, fill=color)) +
  geom_boxplot() +
  scale_fill_brewer(palette="Blues", direction= -1) +
  theme_light() +
  labs(title = "Relationship between Color & Carat") +
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element_blank())
```

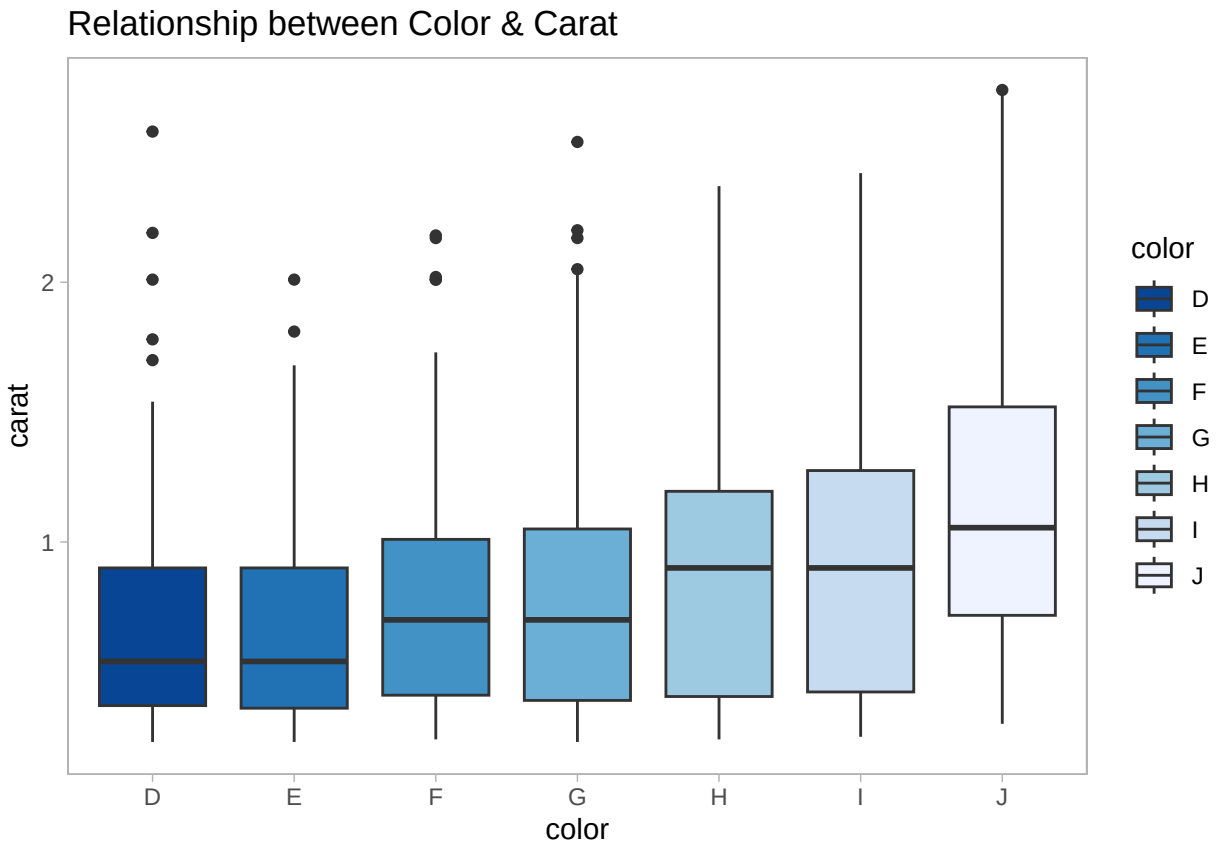


Chart 4 : Relationship betweenCarat and Price (Two Variables)

```
set.seed(36)
small_diamonds <- sample_n(diamonds, 2000)
ggplot(small_diamonds, aes(carat, price, color= cut)) +
  geom_point() +
  facet_wrap(~cut, ncol=5) +
  theme_linedraw() +
  labs(title = "Relationship between Carat and Price(USD)")
```

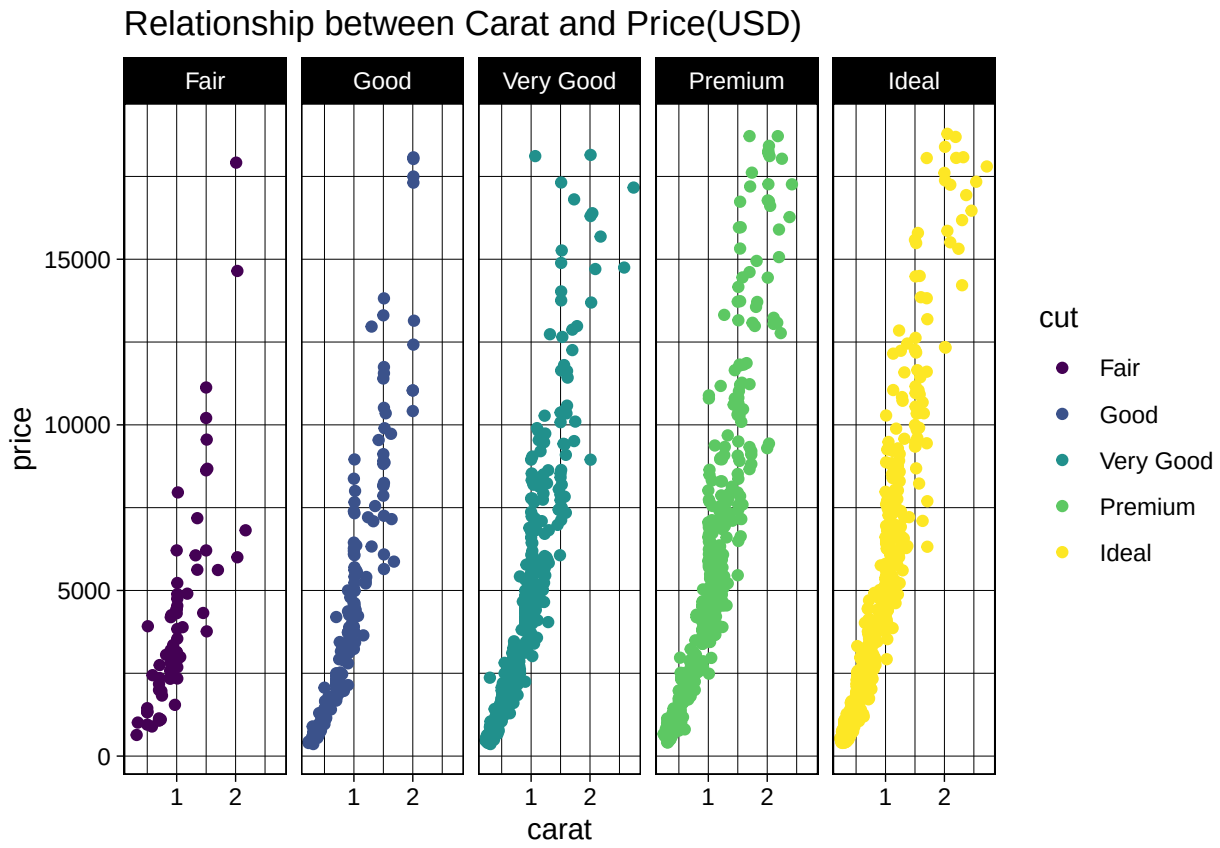
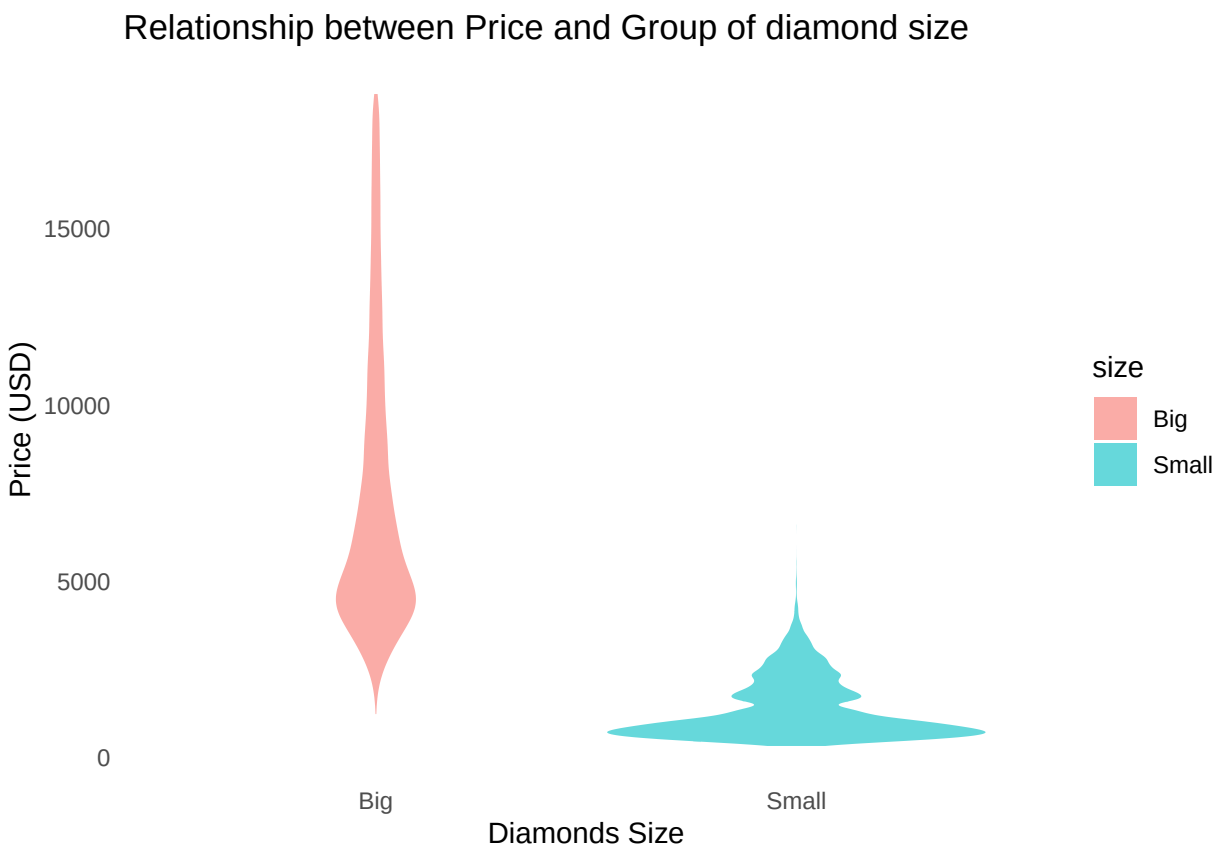


Chart 5 : Relationship between Price and Group of diamond size (Two Variables)

```
new_diamonds <- diamonds %>%
  mutate(size = factor(if_else(carat > mean(diamonds$carat),
                               "Big","Small")))

ggplot(new_diamonds, aes(size,price,fill=size))+
  geom_violin(alpha =0.6 , linetype="blank") +
  theme_minimal() +
  labs(title = "Relationship between Price and Group of diamond size",
       x = "Diamonds Size",
       y = "Price (USD)") +
  theme(panel.grid.major = element_blank(),
        panel.grid.minor = element_blank())
```


Relationship between Price and Group of diamond size
15000
10000
5000
0
Price (USD)
Big
Small
Diamonds Size
size
Big
Small