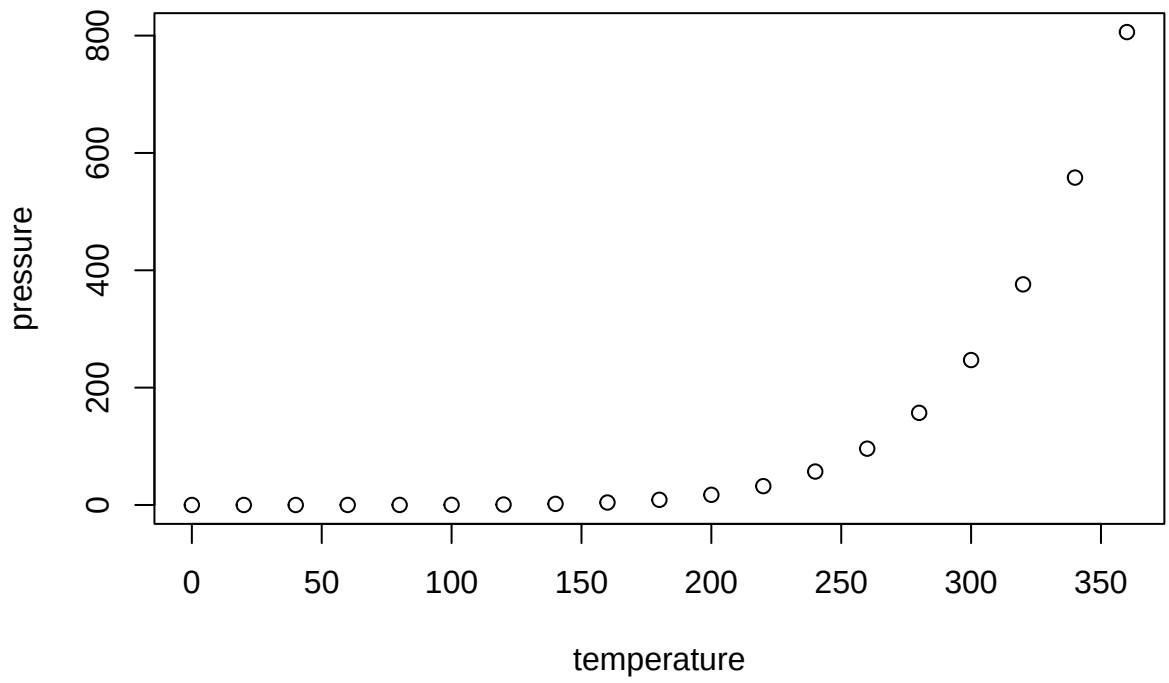
Diamond_Project_2

Ali Valiyev

2024-03-19




```
#Ilk Once "tidyverse" ve "dplyr" paketlerimizi kuruyoruz
clone_diamonds = ggplot2::diamonds
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ------------------------ tidyverse 2.0.0 --
## v dplyr     1.1.4     v readr     2.1.4
## v forcats   1.0.0     v stringr   1.5.0
## v ggplot2   3.4.4     v tibble    3.2.1
## v lubridate 1.9.3     v tidyr     1.3.0
## 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
```

```r
library(dplyr)
```

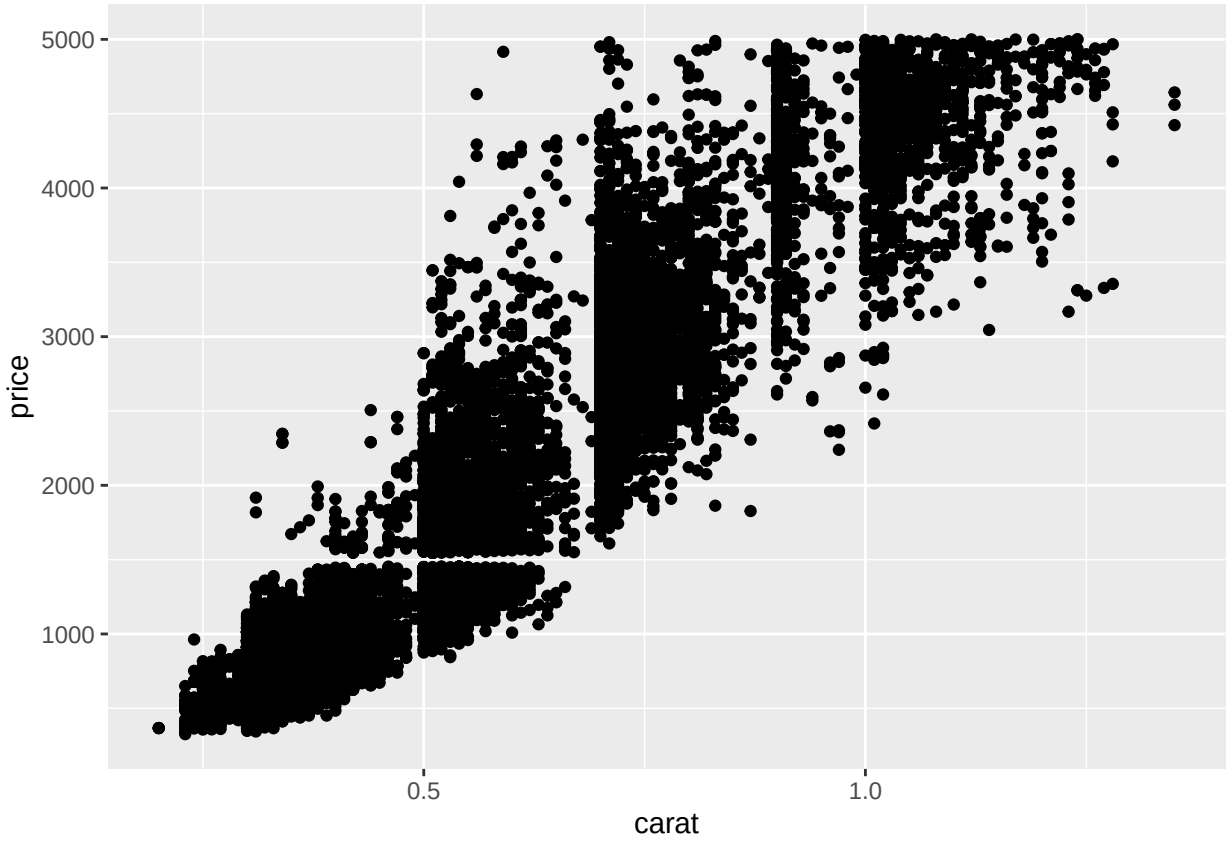
```r
#Verisetinden Bizim kullanabilecegimiz  kismi filter fonksiyonu ile filtreliyoruz.
sub_diamonds = filter(clone_diamonds, price<=5000)
sub_diamonds
```

```
## # A tibble: 39,226 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
##  7  0.24 Very Good I     VVS1     62.3    57   336  3.95  3.98  2.47
##  8  0.26 Very Good H     SI1      61.9    55   337  4.07  4.11  2.53
##  9  0.22 Fair      E     VS2      65.1    61   337  3.87  3.78  2.49
## 10  0.23 Very Good H     VS1      59.4    61   338  4     4.05  2.39
## # i 39,216 more rows
```

```r
#Biz en iyi 2 kesimin her birinin en uygun optimum pirlantasini bularak 5000 dolari en iyi sekilde dege
sub_diamonds = filter(sub_diamonds, cut=="Ideal")
sub_diamonds
```

```
## # A tibble: 16,566 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  0.23 Ideal J     VS1      62.8    56   340  3.93  3.9   2.46
##  3  0.31 Ideal J     SI2      62.2    54   344  4.35  4.37  2.71
##  4  0.3  Ideal I     SI2      62      54   348  4.31  4.34  2.68
##  5  0.33 Ideal I     SI2      61.8    55   403  4.49  4.51  2.78
##  6  0.33 Ideal I     SI2      61.2    56   403  4.49  4.5   2.75
##  7  0.33 Ideal J     SI1      61.1    56   403  4.49  4.55  2.76
##  8  0.23 Ideal G     VS1      61.9    54   404  3.93  3.95  2.44
##  9  0.32 Ideal I     SI1      60.9    55   404  4.45  4.48  2.72
## 10  0.3  Ideal I     SI2      61      59   405  4.3   4.33  2.63
## # i 16,556 more rows
```

```r
#Geom_point fonksiyonunu kullanarak karat ve fiyat arasindaki dagilimin gorsel halini gorebiliyoruz. Bu
Best_diamonds = ggplot(data=sub_diamonds) +
  geom_point(mapping = aes(x = carat, y = price))
Best_diamonds
```




```
#2500 dolardan az ve karati birden az olan pirlantalari siralayalim.
sub_diamonds = filter(sub_diamonds, price<2500 & carat<=1)
sub_diamonds
```

```
## # A tibble: 12,854 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  0.23 Ideal J     VS1      62.8    56   340  3.93  3.9   2.46
##  3  0.31 Ideal J     SI2      62.2    54   344  4.35  4.37  2.71
##  4  0.3  Ideal I     SI2      62      54   348  4.31  4.34  2.68
##  5  0.33 Ideal I     SI2      61.8    55   403  4.49  4.51  2.78
##  6  0.33 Ideal I     SI2      61.2    56   403  4.49  4.5   2.75
##  7  0.33 Ideal J     SI1      61.1    56   403  4.49  4.55  2.76
##  8  0.23 Ideal G     VS1      61.9    54   404  3.93  3.95  2.44
##  9  0.32 Ideal I     SI1      60.9    55   404  4.45  4.48  2.72
## 10  0.3  Ideal I     SI2      61      59   405  4.3   4.33  2.63
## # i 12,844 more rows
```

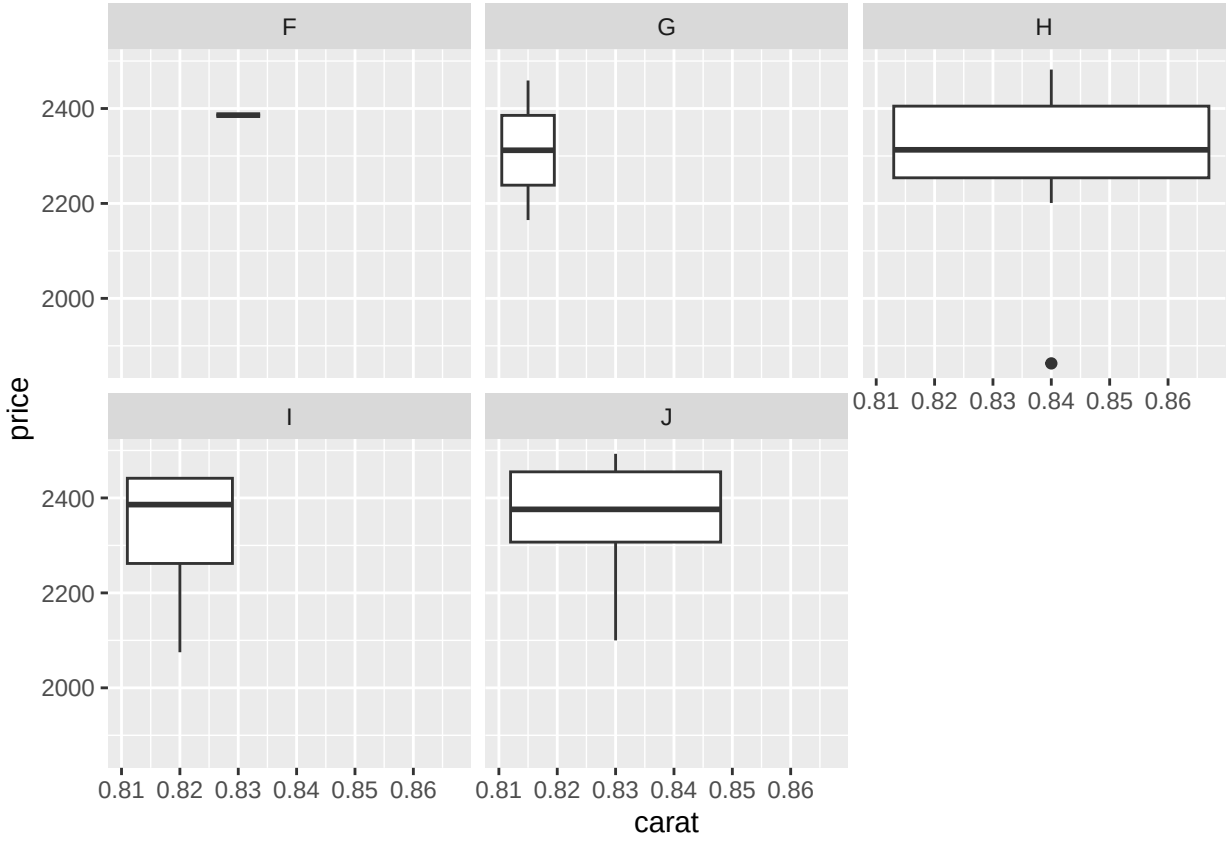
```
#Urunler arasinda clarity acisindan en kotu diyebelcegimiz pirlantalari eliyoruz.
sub_diamonds = filter(sub_diamonds, carat>0.8 & clarity!="I1")
sub_diamonds
```

```
## # A tibble: 26 x 10
##    carat cut   color clarity depth table price     x     y     z
```

```
##    <dbl> <ord> <ord> <ord>   <dbl> <dbl> <int> <dbl> <dbl> <dbl>
##  1  0.83 Ideal H     SI2      61.8    57  1863  6.08  5.99  3.73
##  2  0.82 Ideal I     SI2      61.1    57  2075  6.05  6.09  3.71
##  3  0.81 Ideal J     SI2      62.2    57  2100  5.95  5.92  3.69
##  4  0.82 Ideal G     SI2      61      55  2165  6.04  5.99  3.67
##  5  0.83 Ideal H     SI2      60.7    57  2201  6.08  6.05  3.68
##  6  0.83 Ideal I     VS2      62.9    56  2201  5.98  5.94  3.75
##  7  0.83 Ideal J     VS2      61.5    57  2241  6.08  6.03  3.72
##  8  0.87 Ideal H     SI2      63      57  2307  6.07  6.02  3.81
##  9  0.81 Ideal J     SI1      62      56  2307  6     6.03  3.73
## 10  0.81 Ideal H     SI2      63      57  2313  5.94  5.9   3.73
## # i 16 more rows
```

```
#Elimizdeki son 26 urun icersinde karat ve fiyatinin renk acisindan nasil degerlendirebilecegimizi anla
Best_diamonds = ggplot(data=sub_diamonds) +
  geom_boxplot(mapping = aes(x = carat, y = price)) +
  facet_wrap(~color)
Best_diamonds
```

```
## Warning: Continuous x aesthetic
## i did you forget `aes(group = ...)`?
```



#Gorselden de anlayacagimiz uzere Fiyat-Performans olarak F - rengine sahip pirlanda en uygun secim olacaktir

```r
#F rengine sahip pirlantayi secelim.
sub_diamonds = filter(sub_diamonds, color=="F")
sub_diamonds
```

```
## # A tibble: 1 x 10
##   carat cut   color clarity depth table price     x     y     z
##   <dbl> <ord> <ord> <ord>   <dbl> <dbl> <int> <dbl> <dbl> <dbl>
## 1  0.83 Ideal F     SI2      59.5    57  2386  6.13   6.1  3.64
```

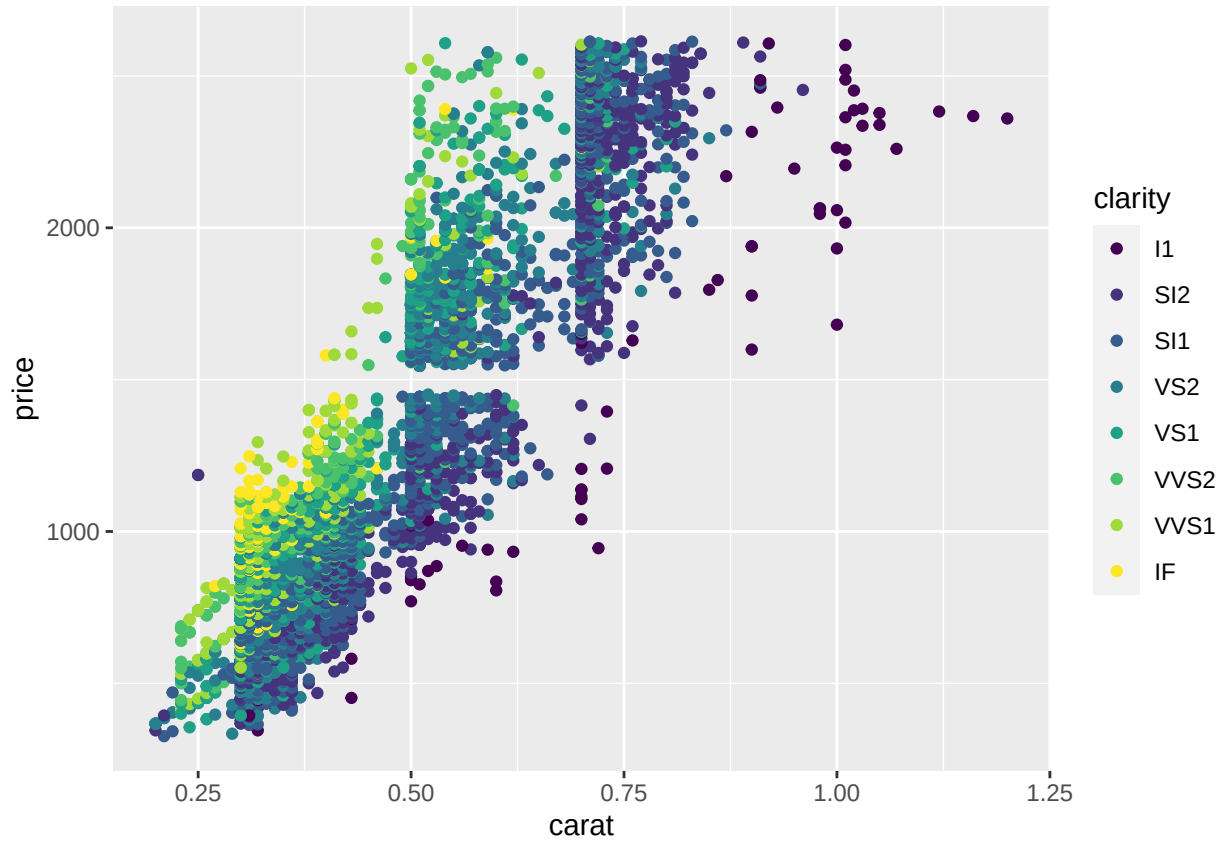
```r
#Paramizin kalan kismiyla alabilecegimiz pirlantalari filtreliyelim.
clone_diamonds = ggplot2::diamonds
sub_diamonds = filter(clone_diamonds, price<=2614)
sub_diamonds
```

```
## # A tibble: 28,234 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
##  7  0.24 Very Good I     VVS1     62.3    57   336  3.95  3.98  2.47
##  8  0.26 Very Good H     SI1      61.9    55   337  4.07  4.11  2.53
##  9  0.22 Fair      E     VS2      65.1    61   337  3.87  3.78  2.49
## 10  0.23 Very Good H     VS1      59.4    61   338  4     4.05  2.39
## # i 28,224 more rows
```

```r
#Pirlantalar icinde "Premium" kesime sahip olan pirlantalari ayiralim
sub_diamonds = filter(sub_diamonds, cut=="Premium")
sub_diamonds
```

```
## # A tibble: 6,241 x 10
##    carat cut     color clarity depth table price     x     y     z
##    <dbl> <ord>   <ord> <ord>   <dbl> <dbl> <int> <dbl> <dbl> <dbl>
##  1  0.21 Premium E     SI1      59.8    61   326  3.89  3.84  2.31
##  2  0.29 Premium I     VS2      62.4    58   334  4.2   4.23  2.63
##  3  0.22 Premium F     SI1      60.4    61   342  3.88  3.84  2.33
##  4  0.2  Premium E     SI2      60.2    62   345  3.79  3.75  2.27
##  5  0.32 Premium E     I1       60.9    58   345  4.38  4.42  2.68
##  6  0.24 Premium I     VS1      62.5    57   355  3.97  3.94  2.47
##  7  0.29 Premium F     SI1      62.4    58   403  4.24  4.26  2.65
##  8  0.22 Premium E     VS2      61.6    58   404  3.93  3.89  2.41
##  9  0.22 Premium D     VS2      59.3    62   404  3.91  3.88  2.31
## 10  0.3  Premium J     SI2      59.3    61   405  4.43  4.38  2.61
## # i 6,231 more rows
```

```r
#Geom_point fonksiyonunu kullanarak karat ve fiyat arasindaki dagilimin clarity-e nazaran gorsel halini
Best_diamonds = ggplot(data=sub_diamonds) +
  geom_point(mapping = aes(x = carat, y = price,color=clarity))
Best_diamonds
```




```
#2000 dolardan cok ve karati 0.8-den cok olan pirlantalari siralayalim.
sub_diamonds = filter(sub_diamonds, price>2000& carat>0.75)
sub_diamonds
```

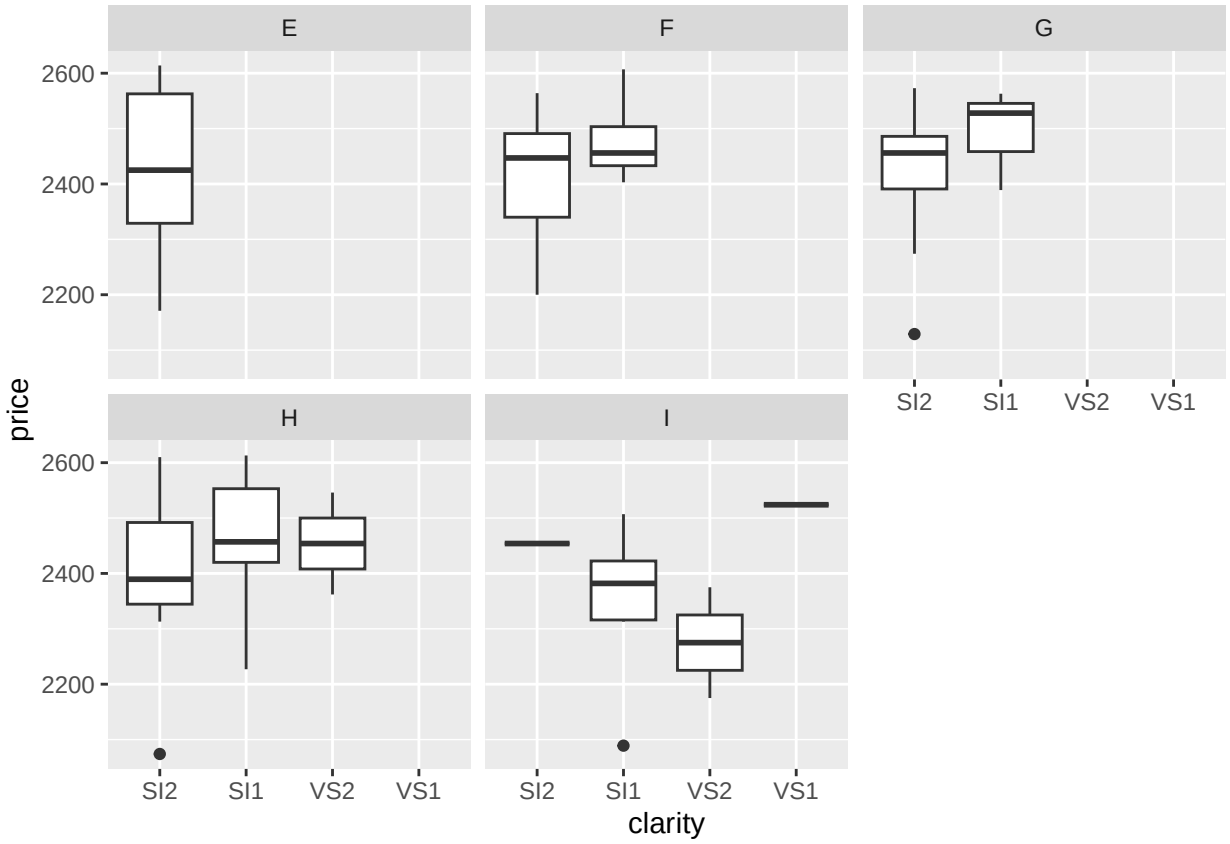
```
## # A tibble: 150 x 10
##    carat cut     color clarity depth table price     x     y     z
##    <dbl> <ord>   <ord> <ord>   <dbl> <dbl> <int> <dbl> <dbl> <dbl>
##  1  0.77 Premium J     VS1      62.2    61  2005  5.85  5.82  3.63
##  2  0.77 Premium J     SI1      61.8    59  2010  5.86  5.82  3.61
##  3  0.8  Premium J     SI1      58.4    59  2016  6.12  6.04  3.55
##  4  1.01 Premium I     I1       59.2    57  2017  6.58  6.44  3.85
##  5  0.8  Premium J     SI2      60.2    59  2022  6.07  5.95  3.62
##  6  0.83 Premium J     SI1      59.3    59  2022  6.17  6.1   3.64
##  7  0.78 Premium J     SI2      60.6    58  2022  5.94  5.91  3.59
##  8  0.8  Premium J     VS2      62.7    61  2045  5.89  5.84  3.68
##  9  0.98 Premium H     I1       59.6    60  2046  6.61  6.48  3.89
## 10  0.76 Premium J     VS2      59.9    59  2052  5.94  5.92  3.55
## # i 140 more rows
```

```
#Urunler arasinda clarity ve color acisindan en kotu diyebelcegimiz pirlantalari eliyoruz.
sub_diamonds = filter(sub_diamonds, color != "J" & clarity != "I1")
sub_diamonds
```

```
## # A tibble: 95 x 10
##    carat cut     color clarity depth table price     x     y     z
```

```
##    <dbl> <ord>   <ord> <ord>   <dbl> <dbl> <int> <dbl> <dbl> <dbl>
##  1  0.78 Premium H     SI2      59.7    59  2074  6.03  5.97  3.58
##  2  0.77 Premium I     SI1      60.4    58  2089  5.96  5.99  3.61
##  3  0.77 Premium G     SI2      61.3    58  2129  5.94  5.88  3.62
##  4  0.77 Premium E     SI2      62.5    60  2171  5.84  5.8   3.64
##  5  0.76 Premium I     VS2      62.8    61  2175  5.78  5.74  3.62
##  6  0.78 Premium F     SI2      62.8    56  2200  5.9   5.86  3.69
##  7  0.76 Premium H     SI1      59.8    57  2227  5.93  5.91  3.54
##  8  0.77 Premium F     SI2      61.6    58  2247  5.9   5.88  3.63
##  9  0.81 Premium G     SI2      59      57  2274  6.14  6.1   3.61
## 10  0.77 Premium I     VS2      61.6    58  2275  5.92  5.96  3.66
## # i 85 more rows
```

```
#Elimizdeki son 95 urun icersinde clarity ve fiyatinin renk acisindan nasil degerlendirebilecegimizi an
Best_diamonds = ggplot(data=sub_diamonds) +
  geom_boxplot(mapping = aes(x = clarity, y = price)) +
  facet_wrap(~color)
Best_diamonds
```



#Goruldugu uzere "F" rengine sahip "SI1" Clarity tipli pirlanta Fiyat-performans olarak "Premium" kesim olan en iyi pirlanta diyebiliriz

```
#aradigimiz pirlantayi secelim
sub_diamonds = filter(sub_diamonds, color == "F" & clarity == "SI1")
sub_diamonds
```

```
## # A tibble: 4 x 10
```

```
##   carat cut     color clarity depth table price     x     y     z
##   <dbl> <ord>   <ord> <ord>   <dbl> <dbl> <int> <dbl> <dbl> <dbl>
## 1  0.81 Premium F     SI1      59.4    60  2403  6.11  6.07  3.62
## 2  0.76 Premium F     SI1      62.5    59  2443  5.79  5.77  3.61
## 3  0.78 Premium F     SI1      62.3    57  2469  5.86  5.85  3.65
## 4  0.76 Premium F     SI1      60.3    58  2607  5.91  5.96  3.58
```

#bu 4 Pirlanta icinde 0.81 karatlik pirlanta satinalabilecegimiz en iyi pirlanta diyebiliriz Note that the `echo = FALSE` parameter was added to the code chunk to prevent printing of the R code that generated the plot.