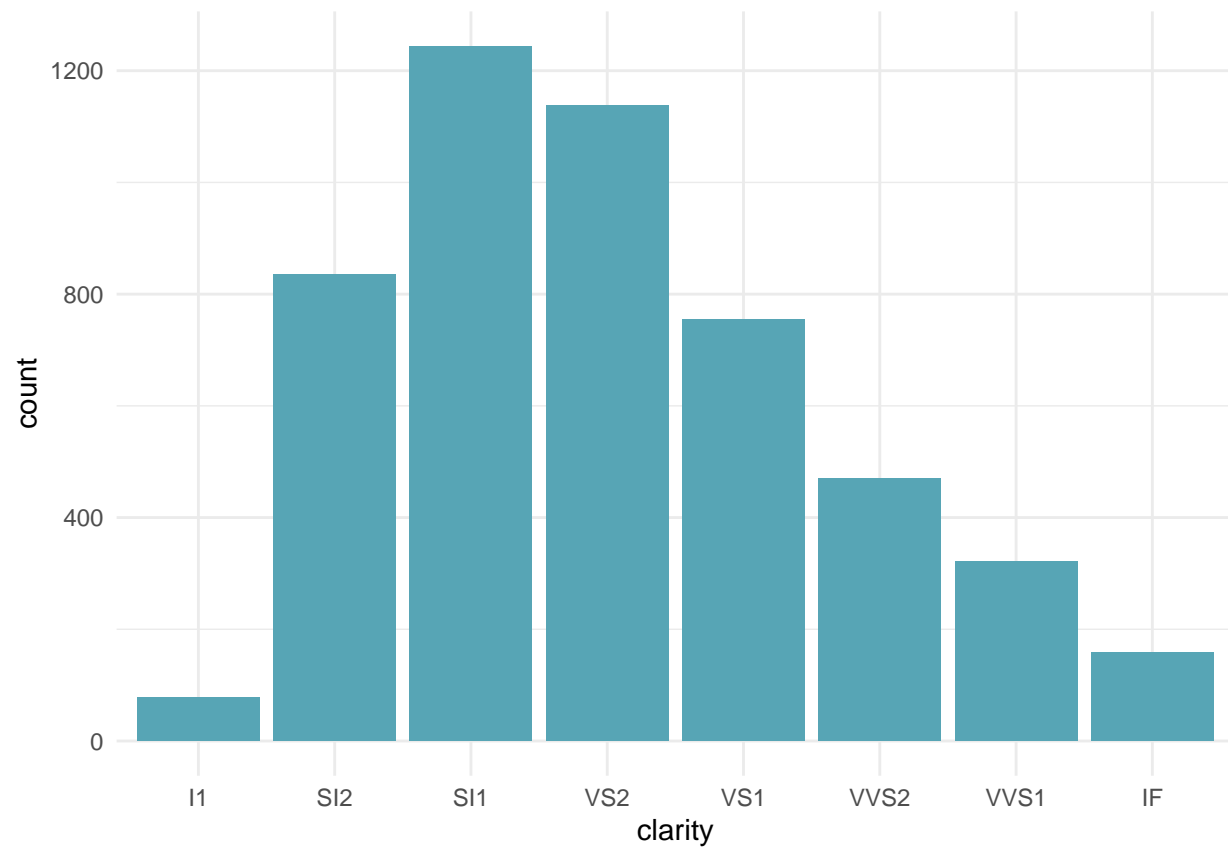# Data Visualization in R - Diamonds

Nappaporn Ellermann

Library

```
library(ggplot2)
library(dplyr)
library(tidyverse)
library(patchwork)
library(lubridate)
library(pander)
```
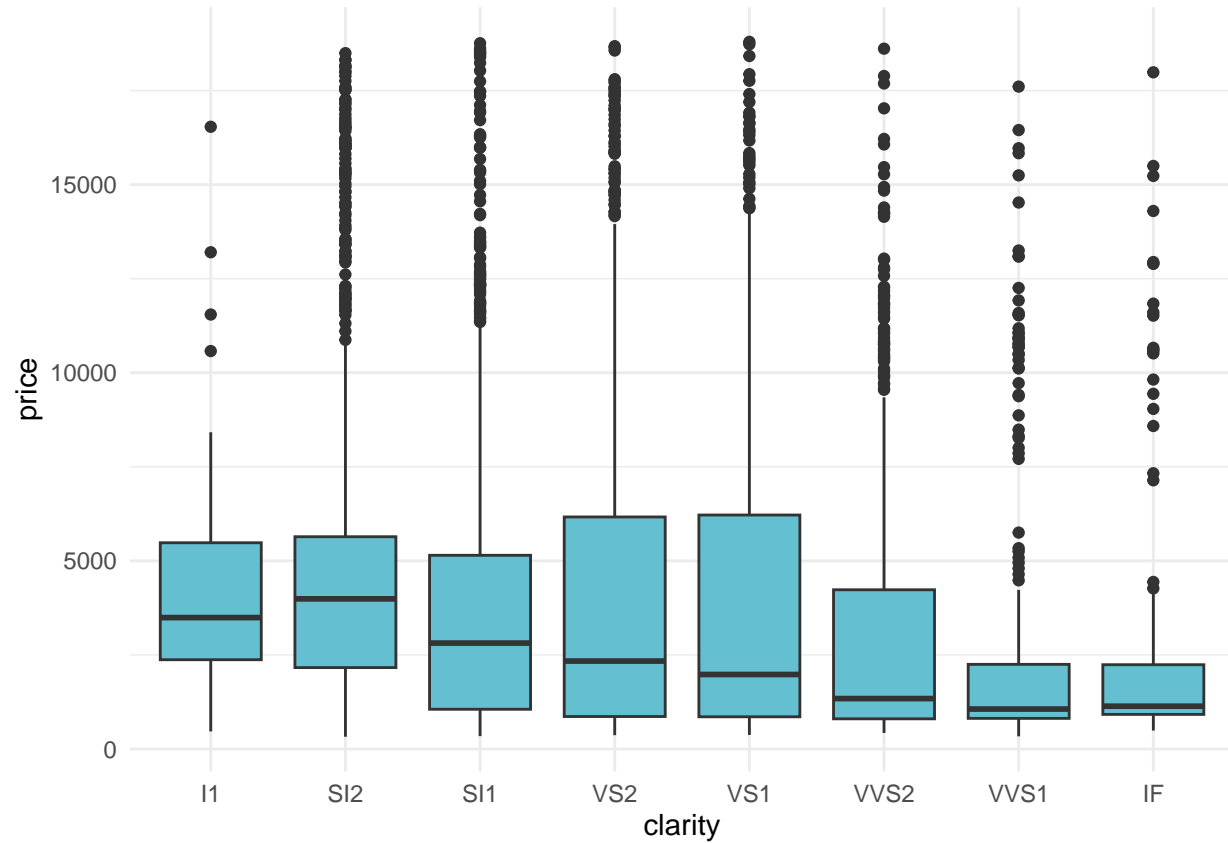
1. Distribution of Clarity

```
ggplot(sample_n(diamonds, 5000), aes(clarity))+
  geom_bar(fill = "#57a5b5")+
  theme_minimal()
```
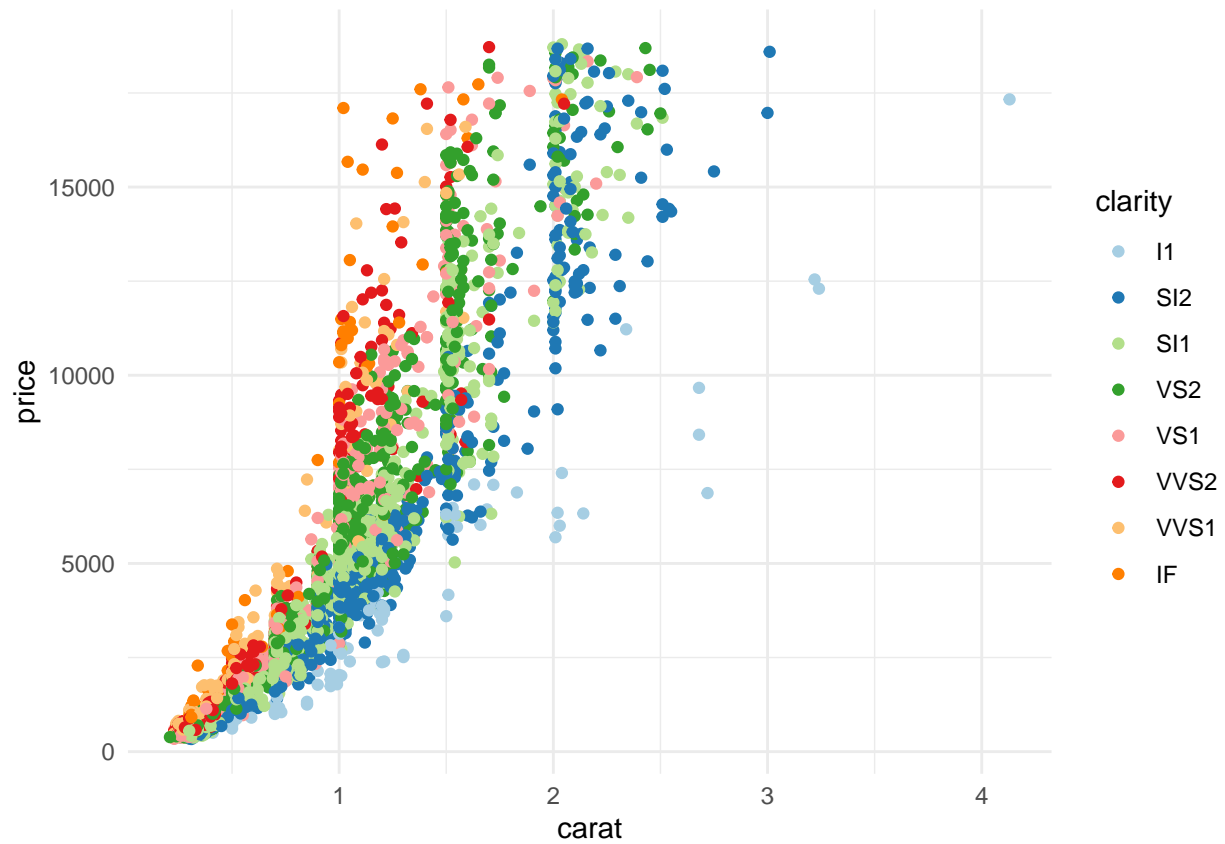
2. Boxplots of Price by Clarity

```
ggplot(sample_n(diamonds, 5000), aes(clarity, price))+ #color
  geom_boxplot(fill = "#64bfd1")+
  theme_minimal()
```
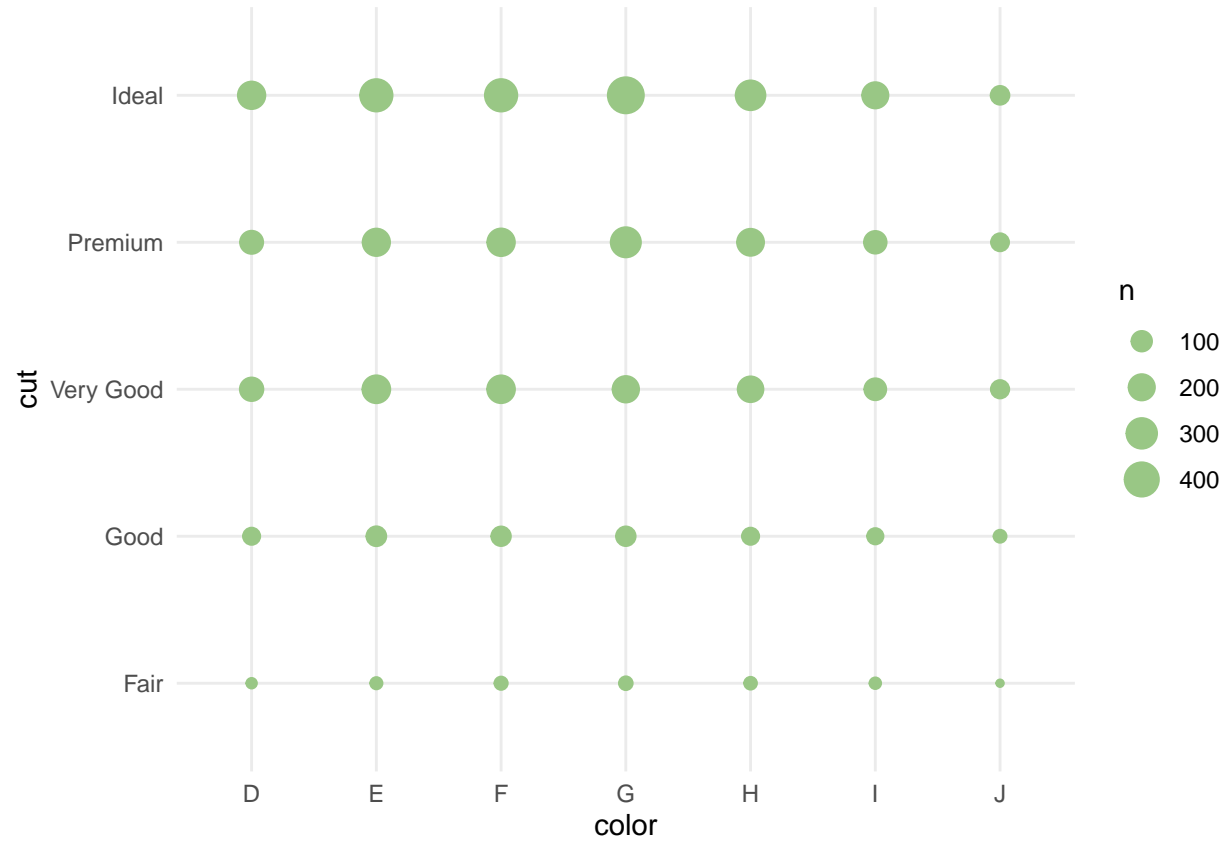
3. Relationship between Price and Carat

```
ggplot(sample_n(diamonds, 5000), aes(carat, price, color=clarity)) +
  geom_point()+
  theme_minimal()+
  scale_color_brewer(type = "qua", palette = 3)
```

4. Counts Cut by Color

```
ggplot(sample_n(diamonds, 5000), aes(color, cut))+ #color, cut #carat, clarity
  geom_count(color="#99c785")+
  theme_minimal()
```

5. Relationship between Price and Carat by Cut

```
ggplot(sample_n(diamonds, 5000), aes(carat, price)) +
  geom_point() +
  geom_smooth(method=lm) +
  facet_wrap(~cut, ncol = 3)+
  theme_light()
```

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