

# Data visualization

Phanuwat

2022-10-12

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

## 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 3.3.6      v purrr  0.3.5
## v tibble  3.1.8      v dplyr  1.0.10
## v tidyr   1.2.1      v stringr 1.4.1
## 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(patchwork)
library(dplyr)
library(ggplot2)

tibble(diamonds)

## # A tibble: 53,940 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
## # ... with 53,930 more rows
```

## Relation Weight and Price

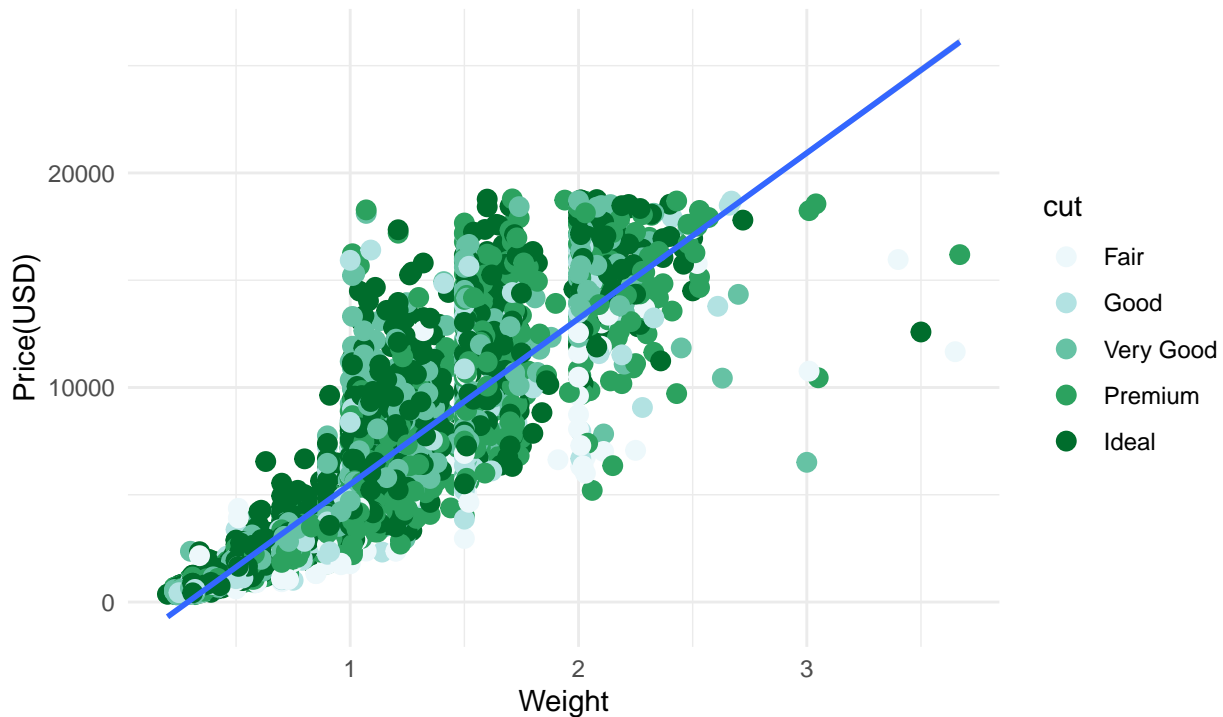
```
set.seed(16)
ggplot(sample_n(diamonds, 10000), aes(carat, price)) +
  geom_point(size=3,
             aes(color = cut)) +
```

```
geom_smooth(method = "lm") +
theme_minimal() +
scale_color_brewer(type = "seq", palette = "BuGn") +
labs(title = "Relation Weight and Price of African diamonds",
      x = "Weight",
      y = "Price(USD)",
      subtitle = "Using ggplot to create this visualization",
      caption = "Source : ggplot package")
```

## `geom\_smooth()` using formula 'y ~ x'

## Relation Weight and Price of African diamonds

Using ggplot to create this visualization



Source : ggplot package

#

Relationship Weight and Volume

```
x <- diamonds$x
y <- diamonds$y
z <- diamonds$z
diamonds$volume <- x * y * z

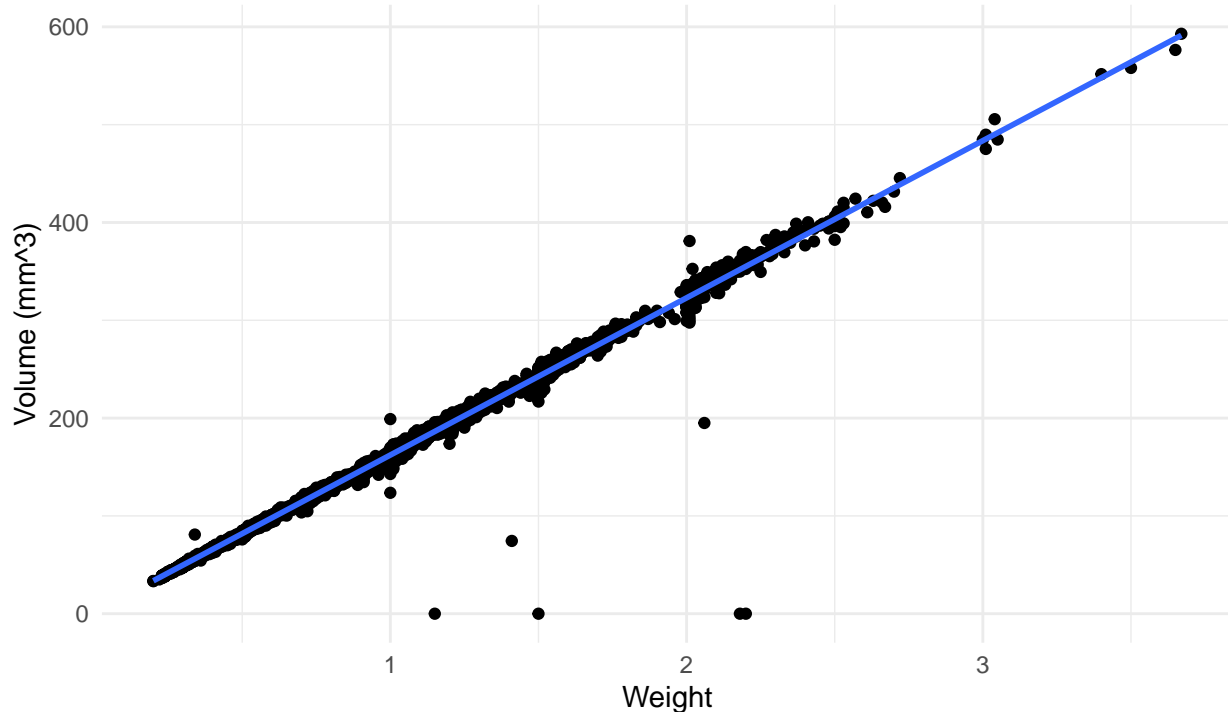
set.seed(16)
ggplot(sample_n(diamonds, 10000),
        aes(carat, volume)) +
  geom_point() +
  geom_smooth(method = "lm") +
  theme_minimal() +
  labs(title = "Relationship Weight and Volume of African diamonds",
        x = "Weight",
        y = "Volume (mm^3)",
        subtitle = "Using ggplot to create this visualization",
```

```
caption = "Source : ggplot package")
```

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

## Relationship Weight and Volume of African diamonds

Using ggplot to create this visualization



Source : ggplot package

```
# Relationship Weight and Volume Group by Clarity
```

```
diamonds %>%
  select(carat, clarity, volume)
```

```
## # A tibble: 53,940 x 3
##   carat clarity volume
##   <dbl> <ord>    <dbl>
## 1  0.23 SI2      38.2
## 2  0.21 SI1      34.5
## 3  0.23 VS1      38.1
## 4  0.29 VS2      46.7
## 5  0.31 SI2      51.9
## 6  0.24 VVS2      38.7
## 7  0.24 VVS1      38.8
## 8  0.26 SI1      42.3
## 9  0.22 VS2      36.4
## 10 0.23 VS1      38.7
## # ... with 53,930 more rows
```

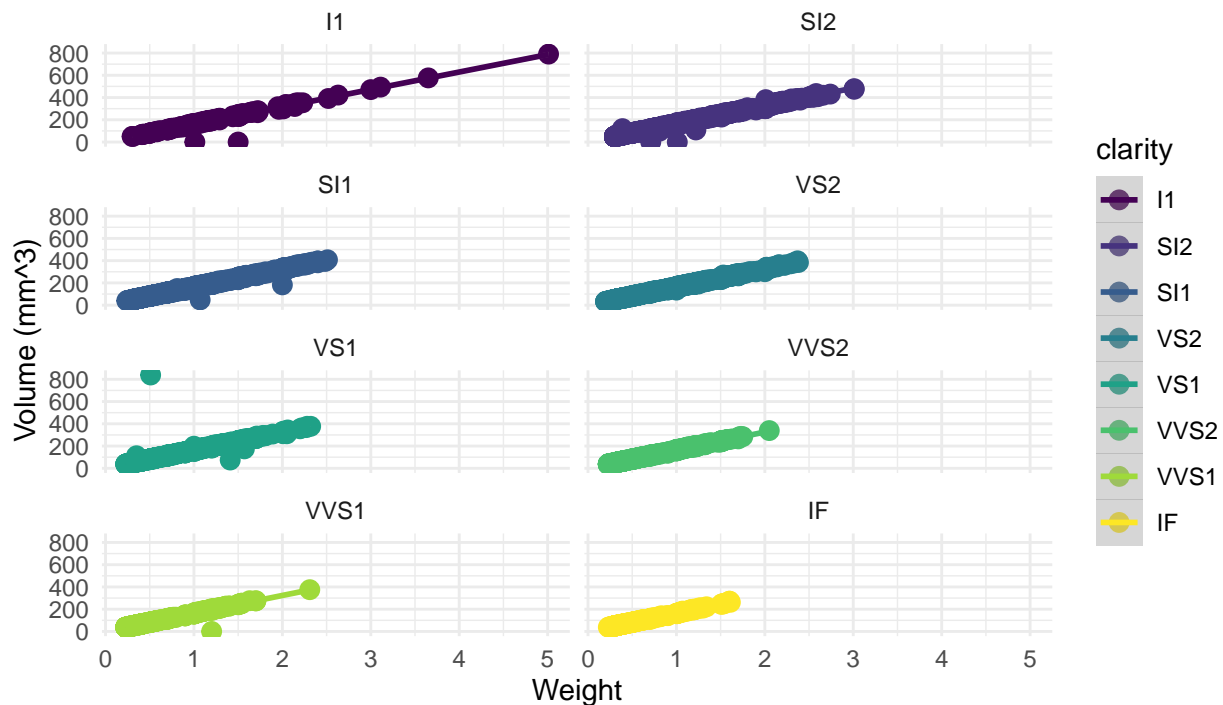
```
ggplot(sample_n(diamonds, 10000),
  aes(carat, volume, color=clarity)) +
  geom_point(size = 3) +
  geom_smooth(method = "lm") +
  facet_wrap(~clarity, ncol=2) +
```

```
theme_minimal() +
labs(title = "Relationship Weight and Volume of African diamonds",
      x = "Weight",
      y = "Volume (mm^3)",
      subtitle = "Using ggplot to create this visualization",
      caption = "Source : ggplot package")
```

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

## Relationship Weight and Volume of African diamonds

Using ggplot to create this visualization



Source : ggplot package

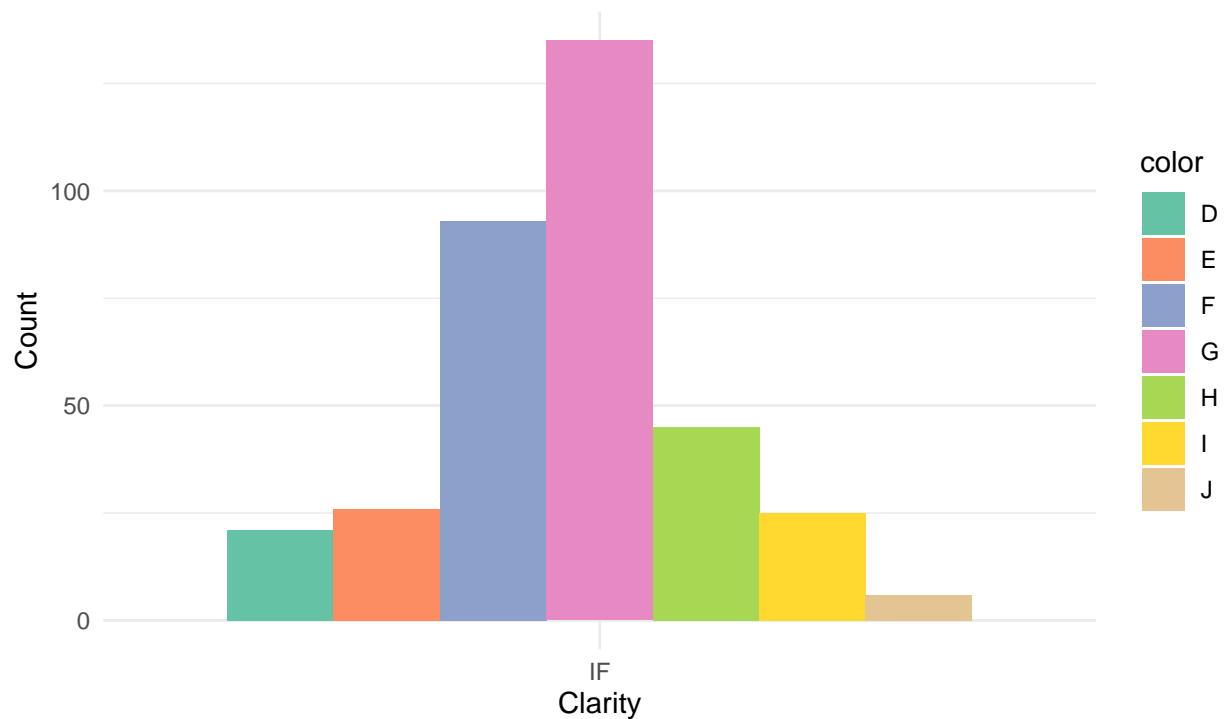
#

Diamonds color of the best clarity of African diamonds

```
set.seed(16)
ggplot(sample_n(diamonds, 10000) %>%
  filter(clarity == "IF"),
  aes(clarity, fill=color)) +
geom_bar(position = "dodge") +
scale_fill_brewer(type = "qual", palette = "Set2") +
theme_minimal() +
labs(title = "Count diamonds color is the best clarity of African diamonds",
      x = "Clarity",
      y = "Count",
      subtitle = "Using ggplot to create this visualization",
      caption = "Source : ggplot package")
```

## Count diamonds color is the best clarity of African diamonds

Using ggplot to create this visualization



Source : ggplot package

#

Relationship table and depth of African diamonds

```
set.seed(16)
ggplot(sample_n(diamonds, 10000),
  aes(table, depth, fill=clarity)) +
  geom_violin() +
  facet_wrap(~clarity, ncol = 4) +
  scale_fill_brewer(type = "div", palette = "RdYlGn") +
  theme_minimal() +
  labs(title = "Relationship table and depth of African diamonds",
    subtitle = "Using ggplot to create this visualization",
    caption = "Source : ggplot package")
```

## Relationship table and depth of African diamonds

Using ggplot to create this visualization

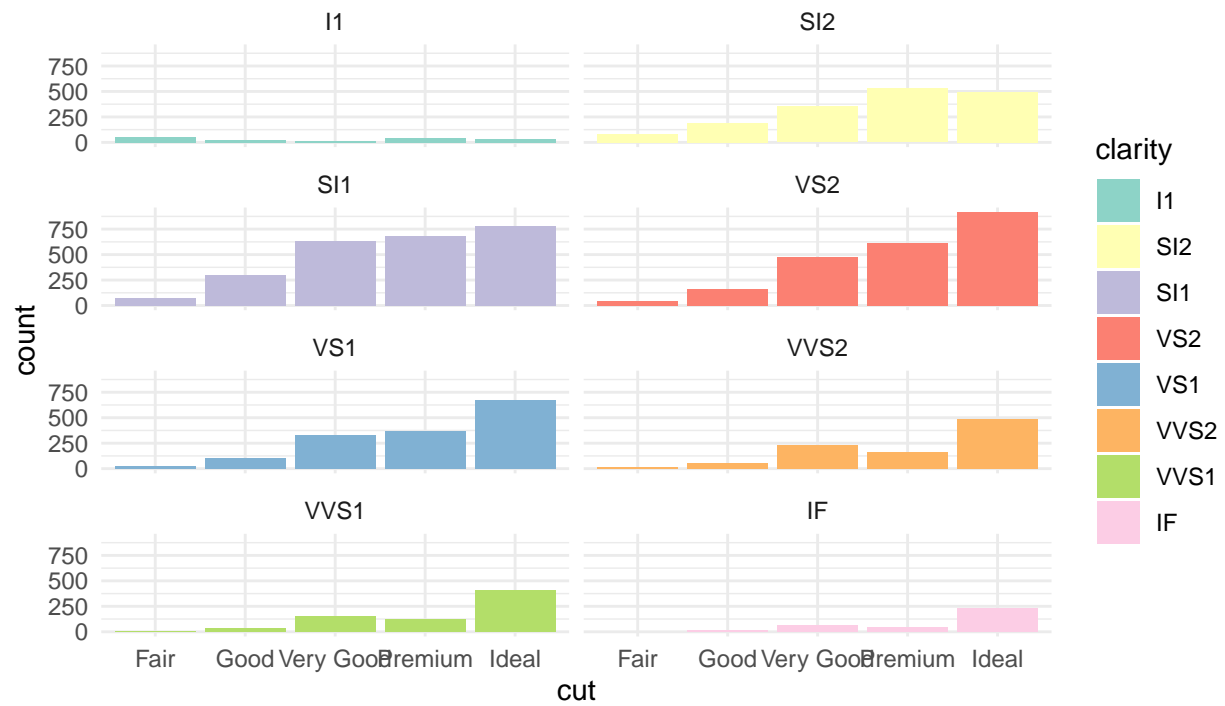


Source : ggplot package

```
set.seed(16)
ggplot(sample_n(diamonds, 10000),
  aes(cut, fill=clarity)) +
  geom_bar() +
  facet_wrap(~clarity, ncol = 2) +
  scale_fill_brewer(type = "qual", palette = "Set3") +
  theme_minimal() +
  labs(title = "Cut diamonds of African diamonds",
    subtitle = "Using ggplot to create this visualization",
    caption = "Source : ggplot package")
```

## Cut diamonds of African diamonds

Using ggplot to create this visualization



Source : ggplot package