Homework Batch 11

Kaewkan Kansomboon

2025-04-15

Homework

My First Bar Chart

```
library(tidyverse)
```

```
## -- Attaching core tidyverse packages ---------------------- tidyverse 2.0.0 --
## v dplyr     1.1.4      v readr     2.1.5
## v forcats   1.0.0      v stringr   1.5.1
## v ggplot2   3.5.2      v tibble    3.2.1
## v lubridate 1.9.4      v tidyr     1.3.1
## v purrr     1.0.4
## -- 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)
library(ggplot2)
d <- diamonds
ggplot(d, aes(cut, fill = cut)) +
  geom_bar() +
  scale_fill_manual( values = c(
    "#f0850c",
    "#72c765",
    "#64d1d1",
    "#849bd1",
    "#c87bc9"
  ))
```
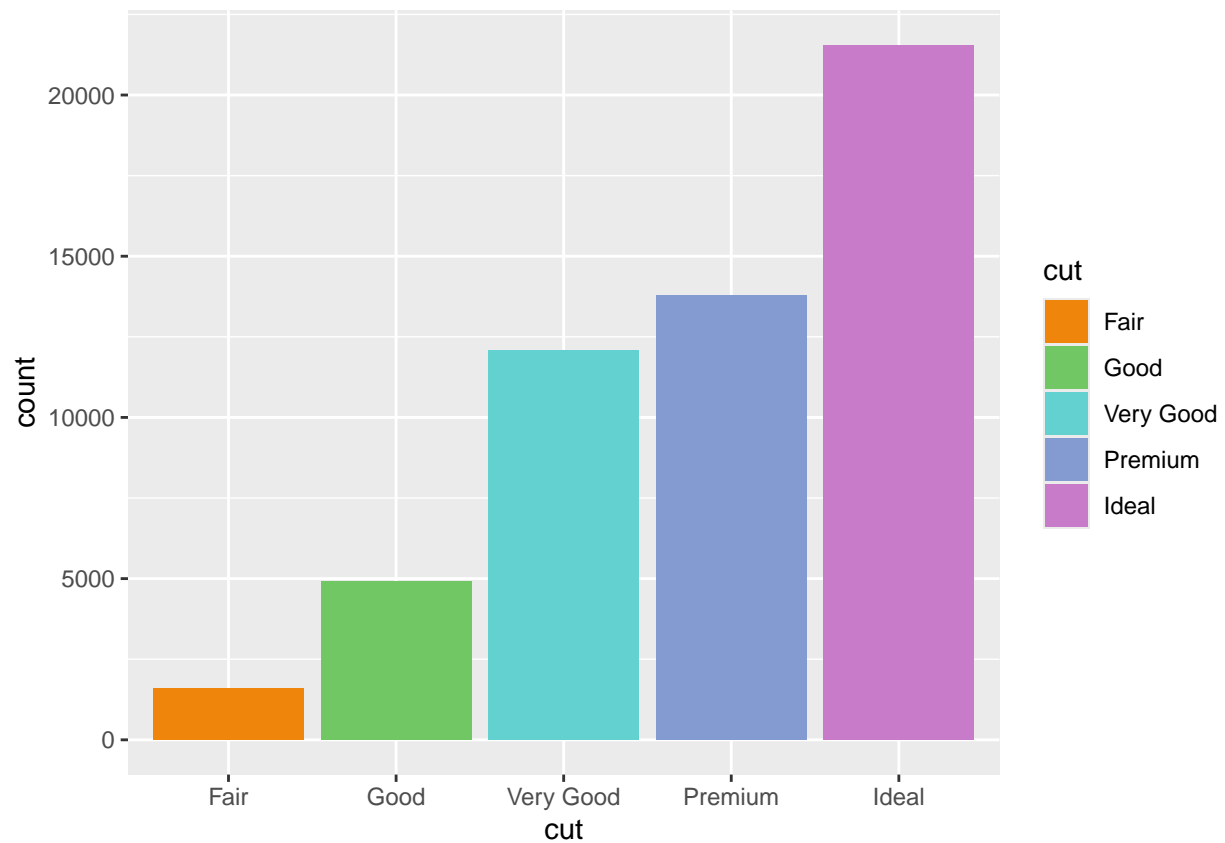
-Ideal cut diamonds are the most popular cut, while Fair is the least common cut in the data.
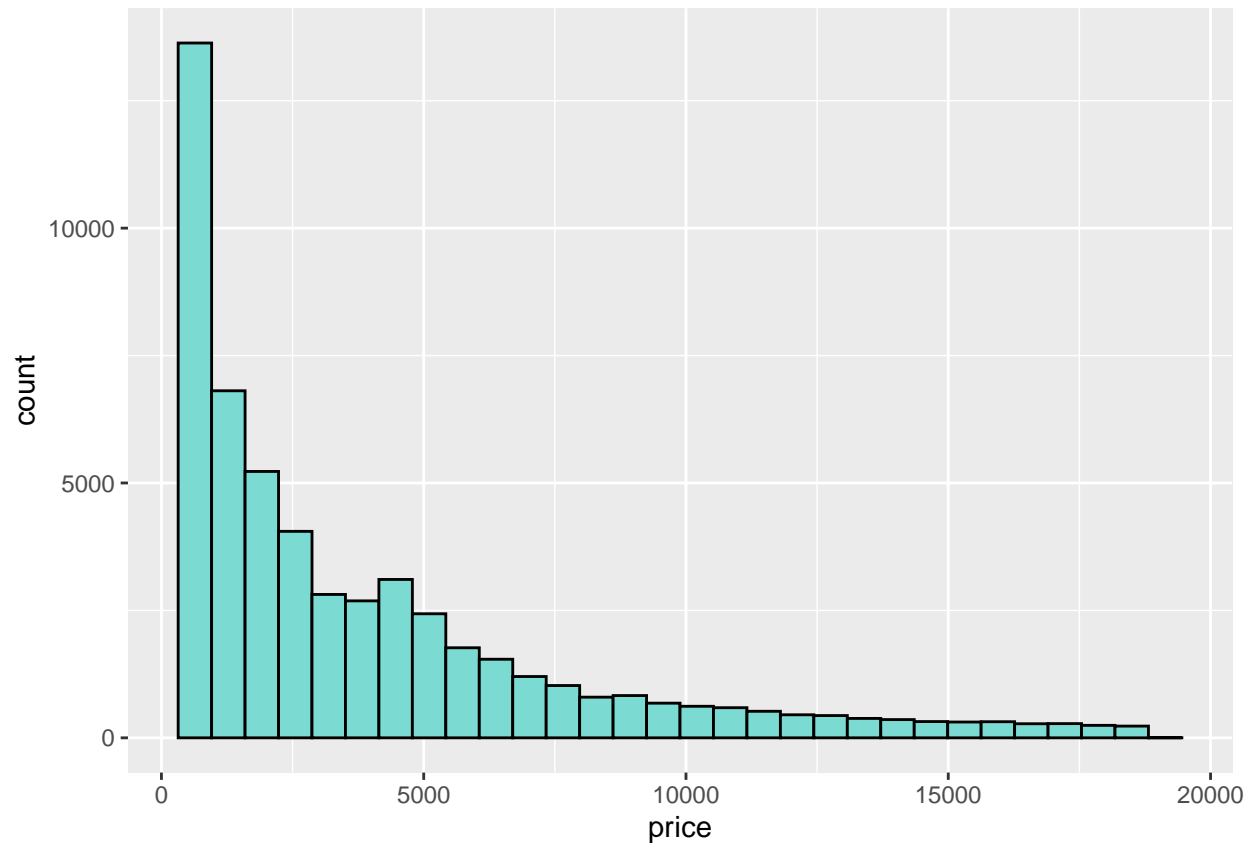
My Second Histogram Chart

```
ggplot(d, aes(price)) +
  geom_histogram(bin = 10, fill ="#7bdbd3", color ="black")
```

```
## Warning in geom_histogram(bin = 10, fill = "#7bdbd3", color = "black"):
## Ignoring unknown parameters: `bin`
```
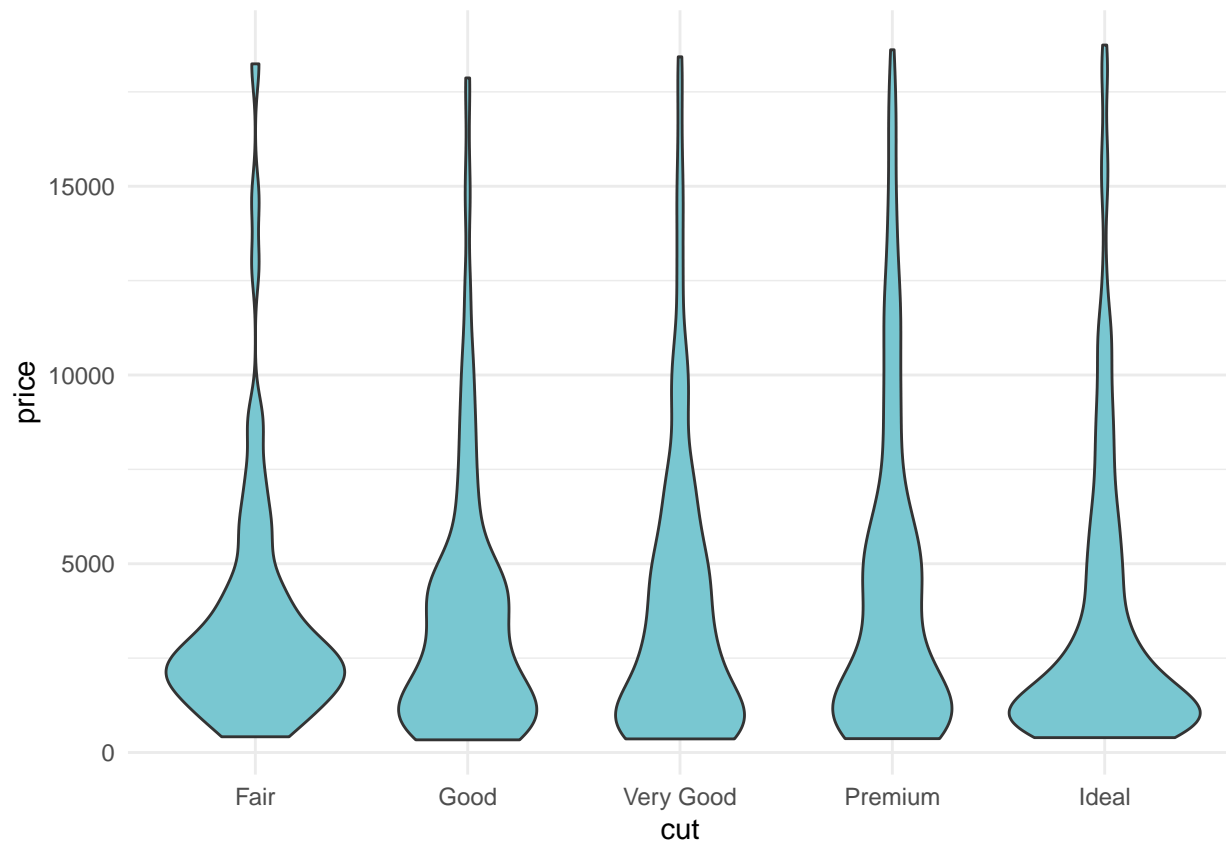
```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```

This chart shows that diamond prices are skewed right, with the majority in the low price range (0-5,000 USD) and significantly fewer in the high price range.

My Third violin chart

```
ggplot(d |> sample_n(2000)
       , aes(cut, price)) +
  geom_violin(fill = "#79c7d1") +
  theme_minimal()
```

This graph shows the price behavior of diamonds in different cut grades, with Premium and Ideal having the widest price spread and covering very high prices, while Fair has the lowest price spread and the narrowest spread.

My Fourth point chart

```r
very_good <- diamonds |>
  filter(cut == "Very Good") |>
  sample_n(500)

good_d <- diamonds |>
  filter(cut == "Good") |>
  sample_n(500)

ggplot()+
  geom_point(data = very_good,
             mapping = aes(price, carat),
             color = "#55e084")+
  geom_point(data = good_d,
             mapping = aes(price, carat),
             color = "#f28763", alpha = 0.5)+
  theme_minimal()
```
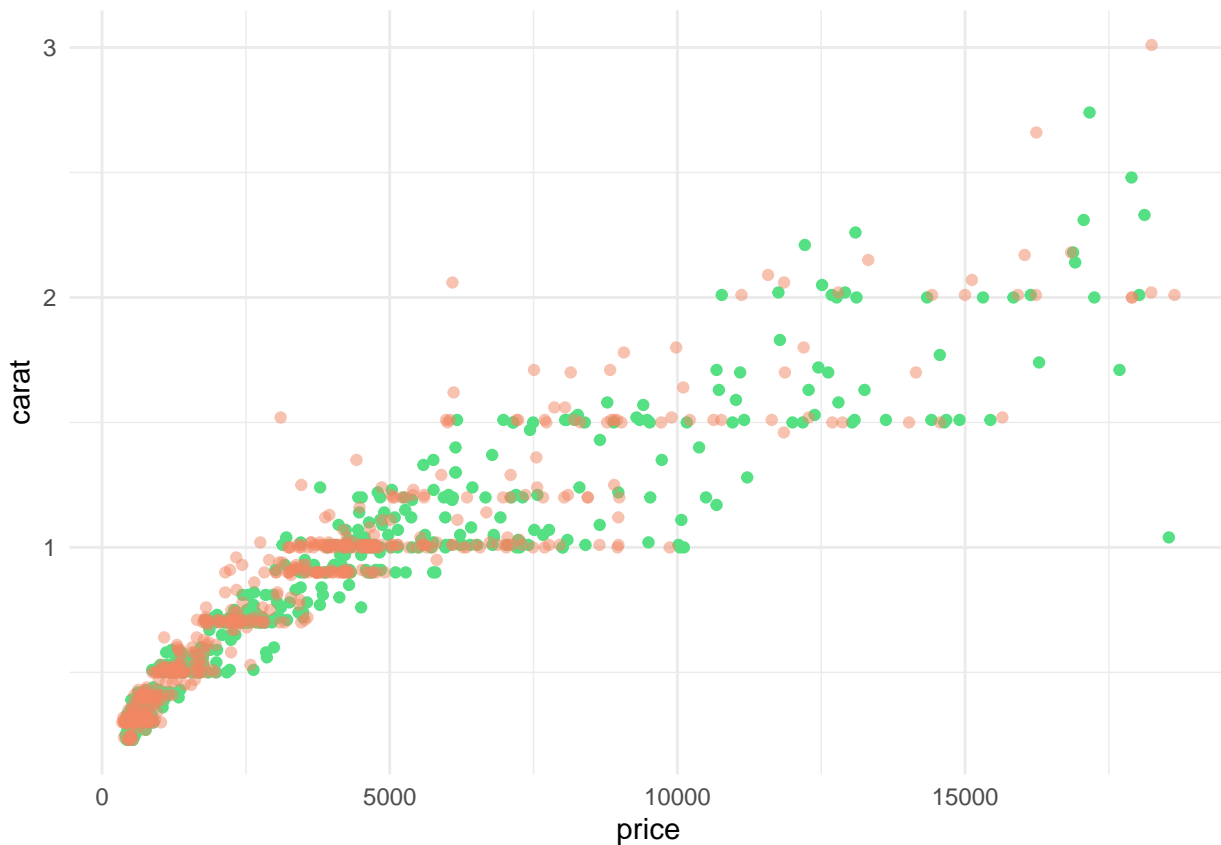
-This chart clearly shows the relationship between diamond weight and price, using colors to separate the data. Carat weight and price increase in the same direction with higher volatility as the Carat value increases.

This my Five Scater Chart

```
d |>
  filter(depth < 60 , price < 400) |>
  ggplot(aes(depth, price)) +
  geom_point() +
  ## regression  | lm = linea model
  geom_smooth(method ="loess",
              se = F, ## standard eror
              color = "red"
  ) +
  geom_rug() +
  theme_minimal() +
  ## add label to chart
  labs(title = "Scater lot depth x price",
       subtitle = "relationship between depth x price",
       caption = "Data Source : diamonds dataframe",
       x = "Carat",
       y = "Price"
  )
```

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

Scater lot depth x price
relationship between depth x price

Data Source : diamonds dataframe
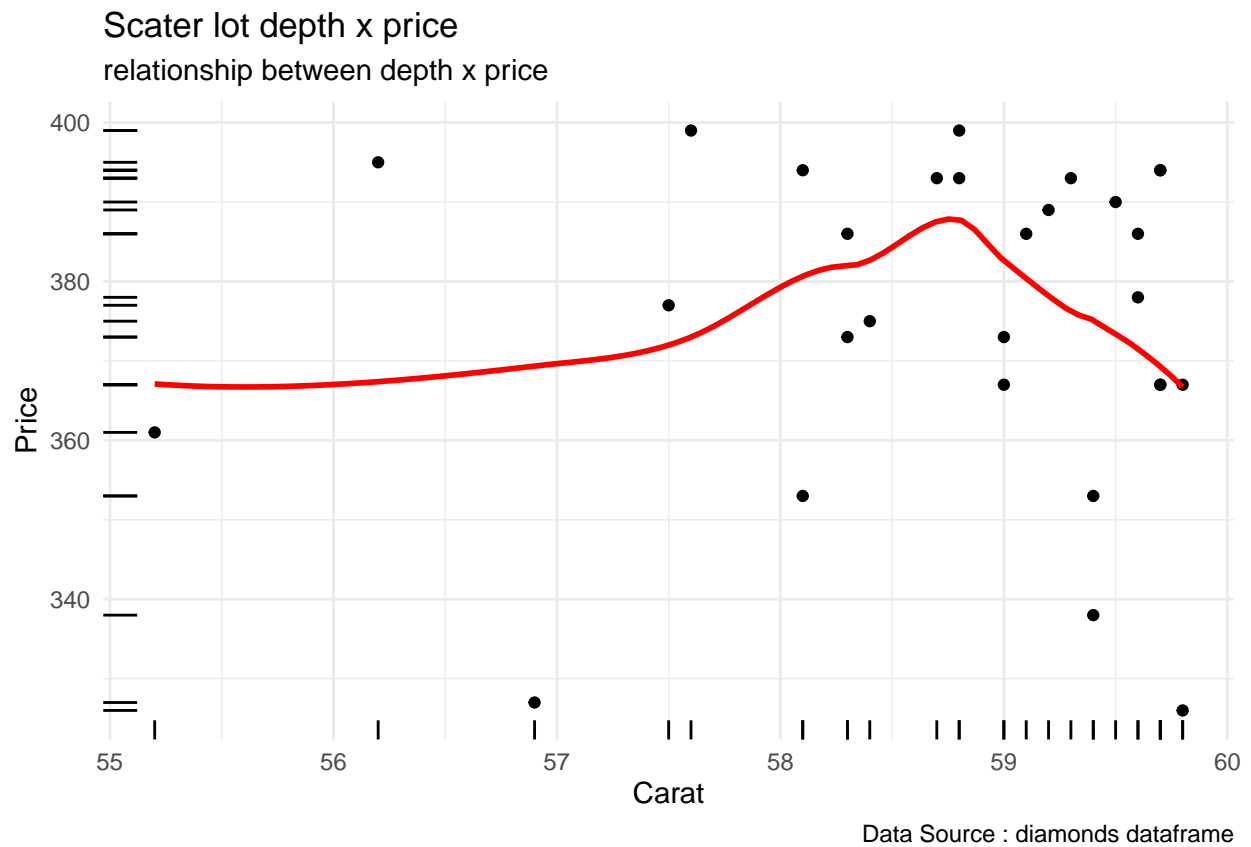
This graph show relationship of depth with price. The price initially increases slowly until reaching its peak (~ Carat 58-59) before decreasing as the Carat gets higher.