

HW Data Visualization in R

Data Visualization in R

Preparation

```
library(tidyverse)

## -- Attaching packages ----- tidyverse 1.3.2 --
## v ggplot2 3.3.6      v purrr  0.3.4
## v tibble  3.1.8      v dplyr  1.0.10
## v tidyr   1.2.1      v stringr 1.4.1
## v readr   2.1.2      v forcats 0.5.2
## -- Conflicts ----- tidyverse_conflicts() --
## x dplyr::filter() masks stats::filter()
## x dplyr::lag()    masks stats::lag()

library(dplyr)
library(ggthemes)
```

Chart01

Shown relationship between cut and clarity (the more diamonds are clear, the more quality they are)

```
ggplot(diamonds, mapping = aes(clarity, fill = cut)) +
  geom_bar(position = "fill")
```

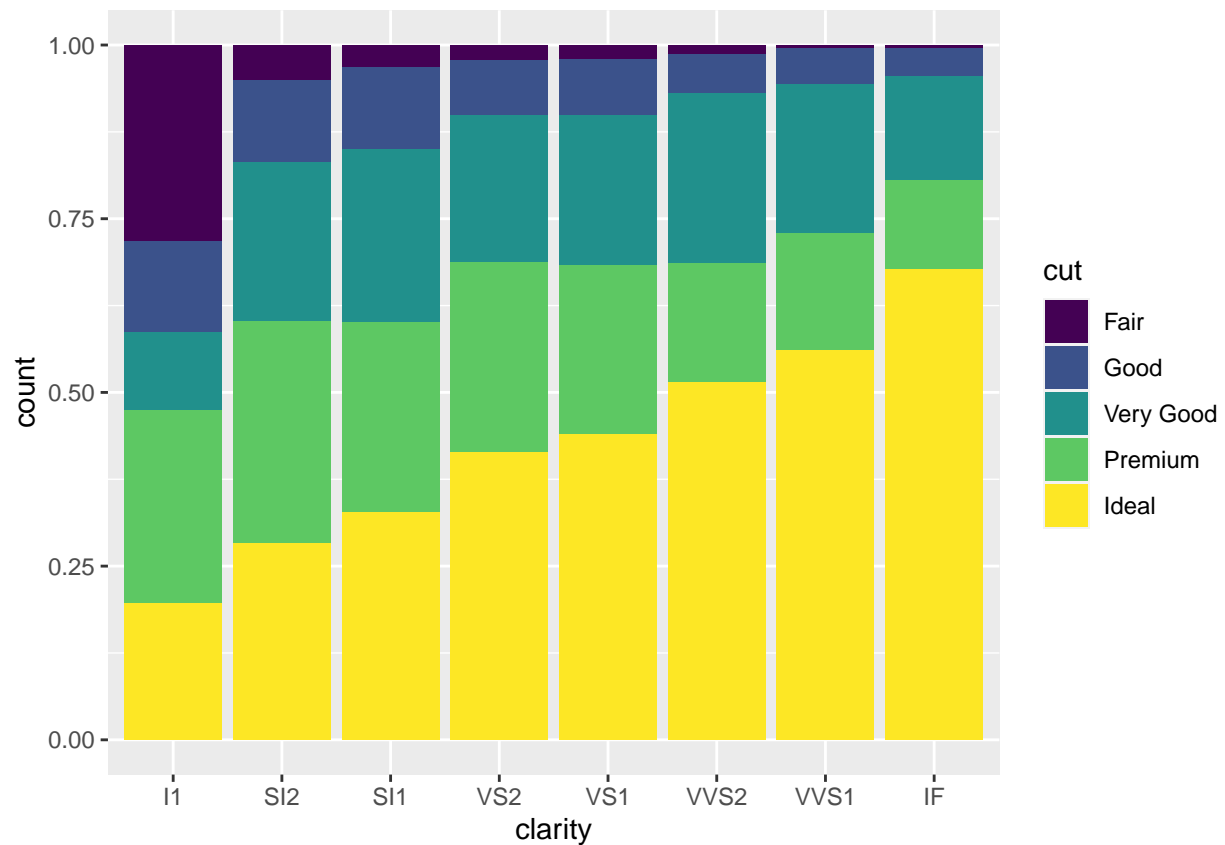


Chart02

Shown relationship betweenvore and total sleep hour of mammals (Omnivores mostly sleep around 10 hours, average insectivore total sleep hour is higher than others)

```
ggplot(data = msleep, aes(x =vore, y = sleep_total)) +
  geom_violin(color = "blue") +
  geom_jitter(color = "salmon") +
  theme_minimal()
```

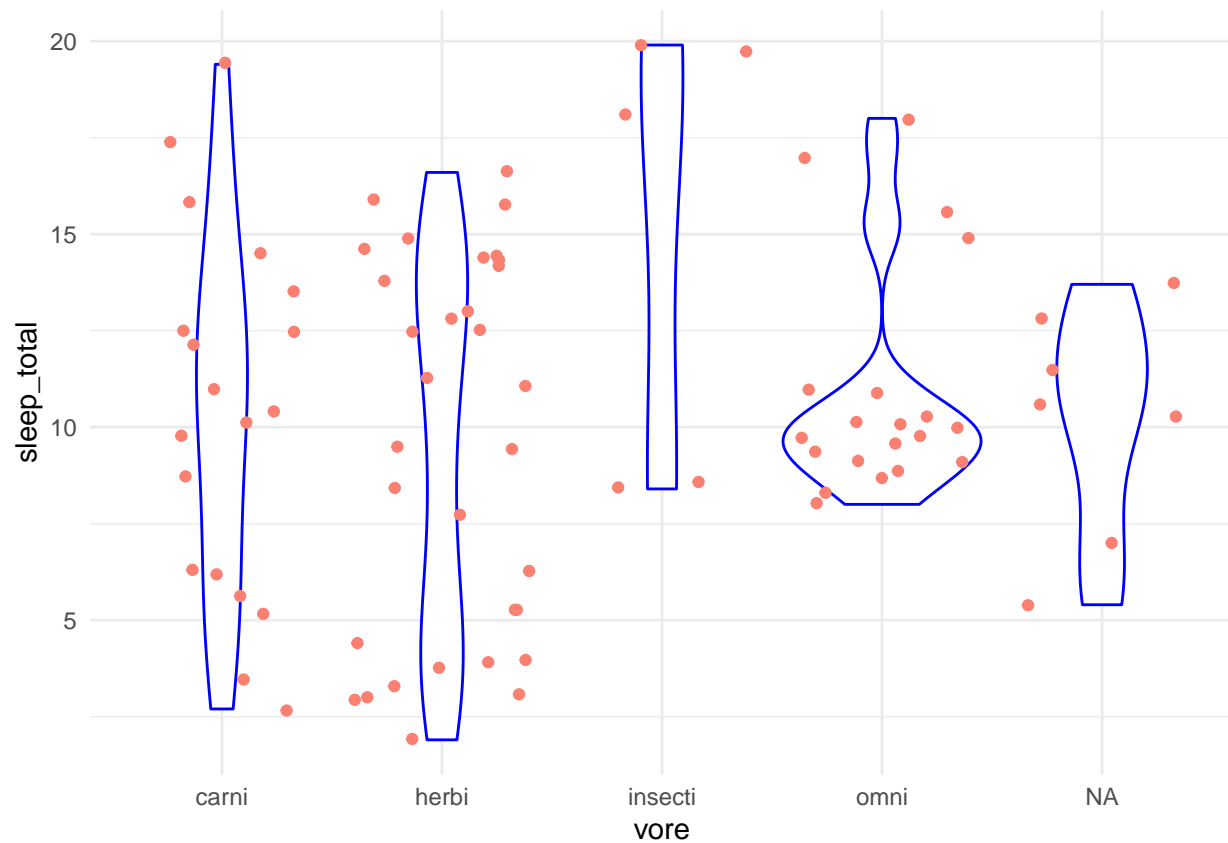


Chart03

Shown relationship between total sleep hour and time spent awake of mammals group by vore (the more time mammals sleep, the less time spent to awake)

```
ggplot(data = msleep, aes(x = sleep_total, y = awake)) +  
  geom_point(size = 1) +  
  facet_wrap(~ vore)
```

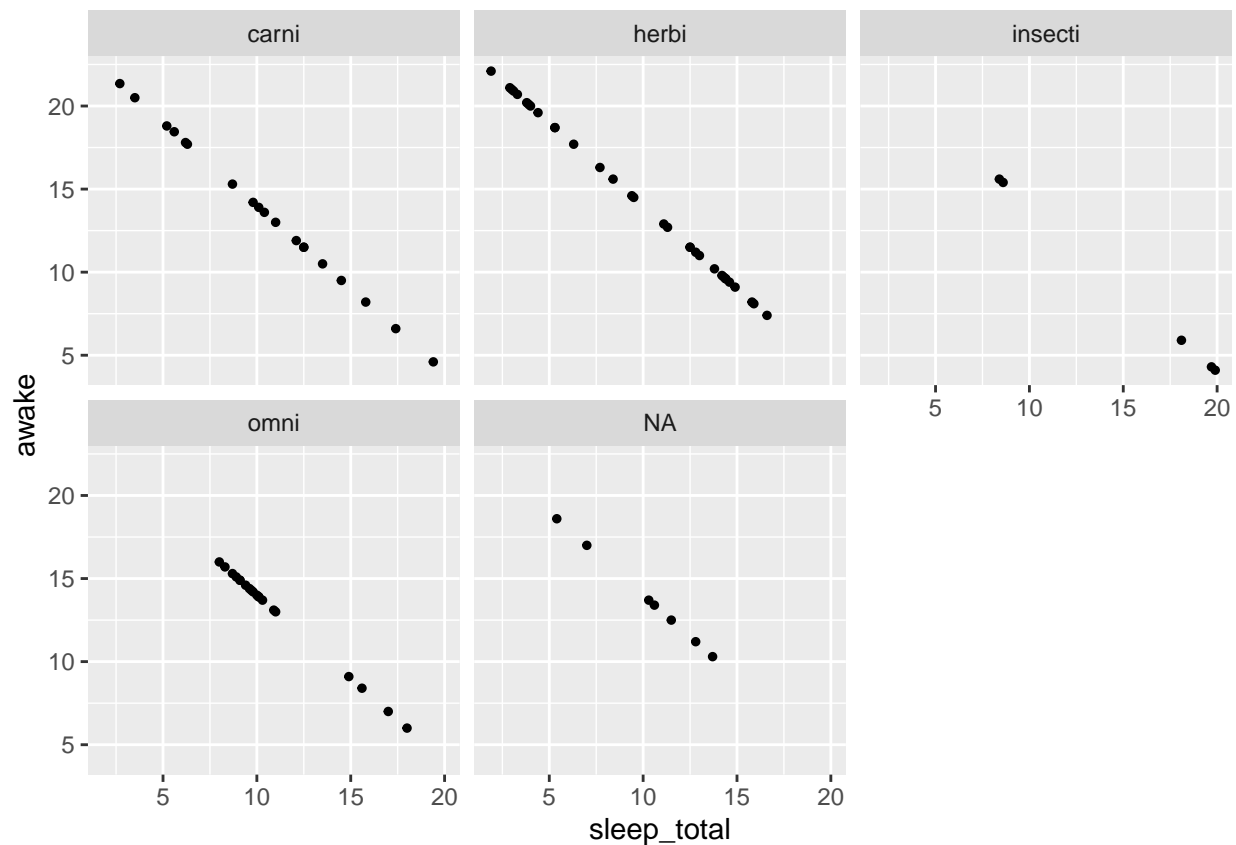


Chart04

Shown relationship between total sleep hour and body weight of mammals (there is not clearly relationship between total sleep hour and body weight of mammals)

```
## filter out outliers
msleep_no_out <- msleep %>%
  filter(bodywt < 250)

## create chart
ggplot(data = msleep_no_out, aes(x = bodywt, y = sleep_total)) +
  geom_line(alpha = 0.4) +
  geom_point(size = 1) +
  theme_minimal()
```

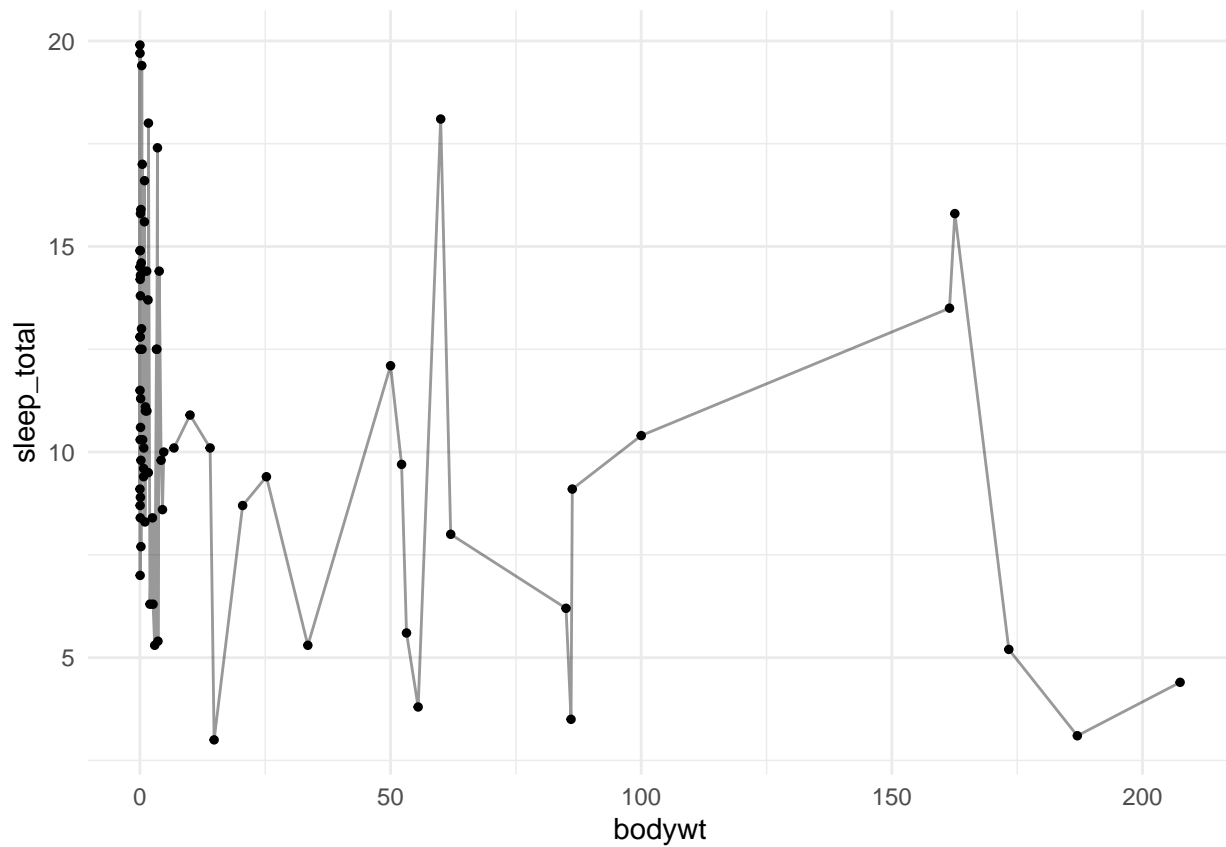


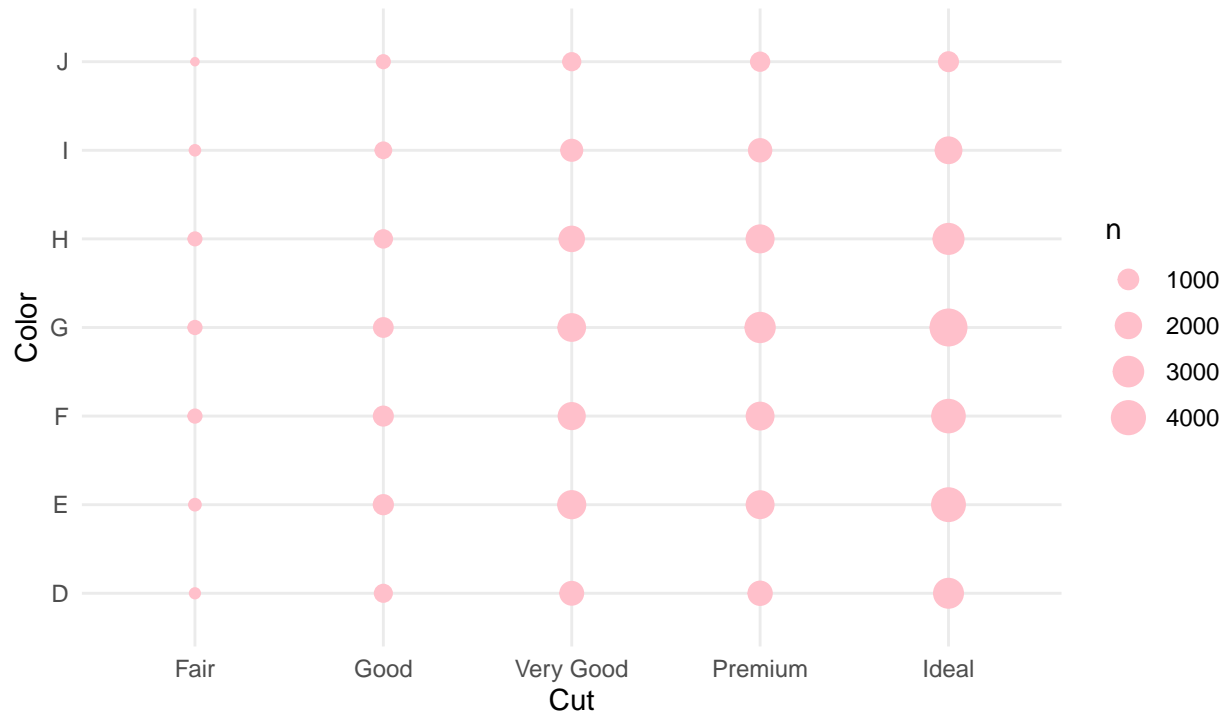
Chart05

Shown distribution of color in each cut (The most color in cut Ideal is color G and the least color in cut Ideal is color J)

```
ggplot(data = diamonds, aes(cut, color)) +
  geom_count(color = "pink") +
  theme_minimal() +
  labs(
    title = "Distribution of color in each cut of diamonds",
    x = "Cut",
    y = "Color",
    subtitle = "Using ggplot to create this visualization",
    caption = "Source: ggplot package"
  )
```

Distribution of color in each cut of diamonds

Using ggplot to create this visualization



Source: ggplot package