

Assignment_1

Alexis McCartney

2025-09-07

```
# Load my Data
library(readxl)
Cereals <- read_excel("~/Desktop/BA64060/Cereals.xlsx")
```

Quantitative Variable: Calories

I grouped the calories of Cereal brands by variable.

```
summary(Cereals$calories)
```

```
##      Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
##      50.0   100.0   110.0   106.9   110.0   160.0
```

Findings: The Cereal brands with the fewest calories has 50 calories. 25% of cereals have 100 or fewer calories. Half of the cereals are greater or fewer than 110 calories. Average calories across most brands are 106.9 and 75% of cereals have 110 or fewer calories.

Categorical Variable: Manufacturer

I grouped the manufacturer variable

```
table(Cereals$mfr)
```

```
##
##  A  G  K  N  P  Q  R
##  1 22 23  6  9  8  8
```

Findings: The majority of Cereal brands are manufactured by Kellogg & General Mills. Smaller manufacturers like A and N are not as represented in the data, as they appear once.

Variable Transformation

I transformed Cereal names into all lowercase.

```
Cereals$name_lower<- tolower(Cereals$name)
head(Cereals[, c("name", "name_lower")], 10)
```

```
## # A tibble: 10 x 2
##   name                name_lower
##   <chr>              <chr>
## 1 100%_Bran          100%_bran
## 2 100%_Natural_Bran 100%_natural_bran
## 3 All-Bran          all-bran
## 4 All-Bran_with_Extra_Fiber all-bran_with_extra_fiber
## 5 Almond_Delight    almond_delight
## 6 Apple_Cinnamon_Cheerios apple_cinnamon_cheerios
## 7 Apple_Jacks        apple_jacks
## 8 Basic_4            basic_4
## 9 Bran_Chex          bran_chex
## 10 Bran_Flakes        bran_flakes
```

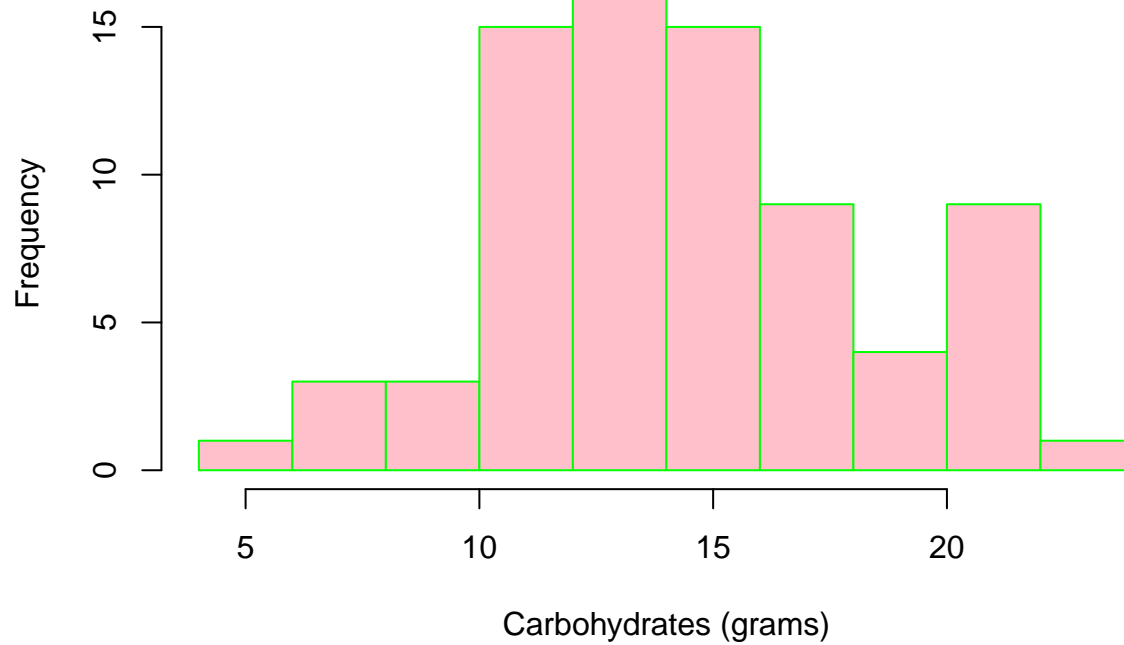
Findings: The original cereal names reveal lower case versions.

Quantitative Variable Plot

Histogram of carbohydrates (carbo)

```
hist(Cereals$carbo,
main = "Carbohydrates in Cereals",
xlab = "Carbohydrates (grams)",
col = "pink",
border = "green")
```

Carbohydrates in Cereals



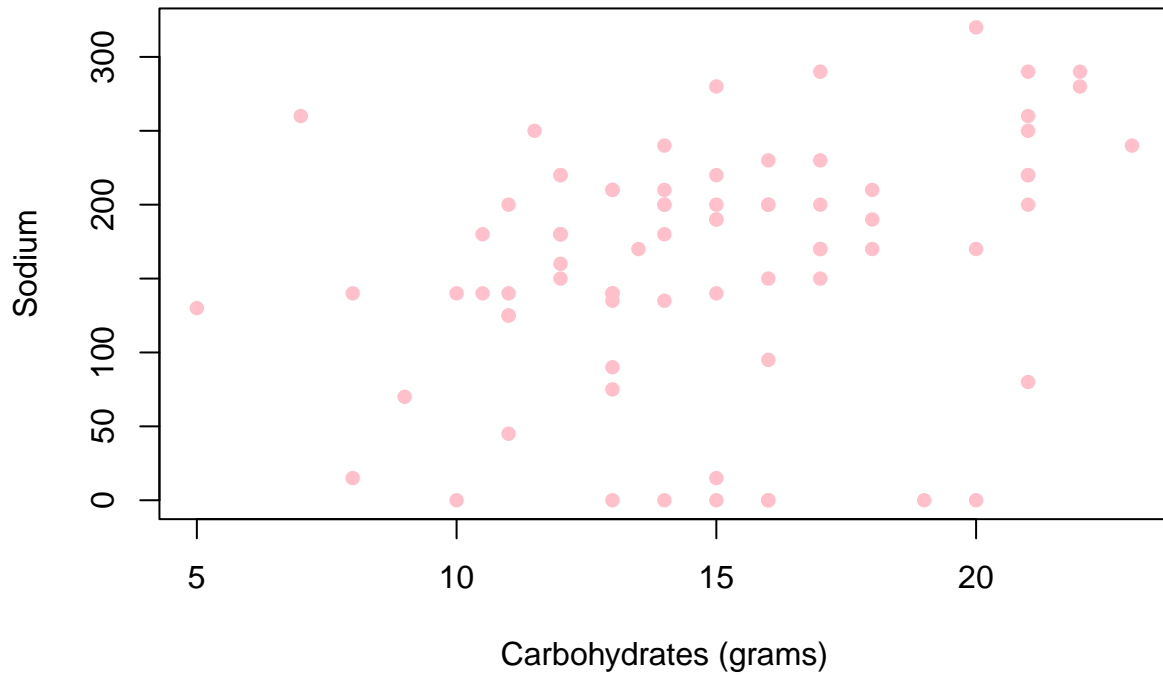
Findings: Most Cereals are in the 10-15g carbohydrate range.

Quantitative Variable Scatterplot

Scatterplot of **Carbohydrates** versus **Sodium**

```
plot(Cereals$carbo, Cereals$sodium,  
main = "Sodium vs Carbohydrates in Cereals",  
xlab = "Carbohydrates (grams)",  
ylab = "Sodium",  
col = "pink",  
pch = 16)
```

Sodium vs Carbohydrates in Cereals



Findings: Each point represents a Cereal. The findings show if cereals with higher carbohydrates concentration also has more sodium.