Tidyverse CREATE Assignment: Groceries Data

DH Kim

2020-10-26

This Assignment uses `readr`, `dplyr',stringr, andggplot2‘` to do the following data work.

```
# libraries in Tidyverse
library(readr)
library(dplyr)
library(stringr)
library(ggplot2)
```

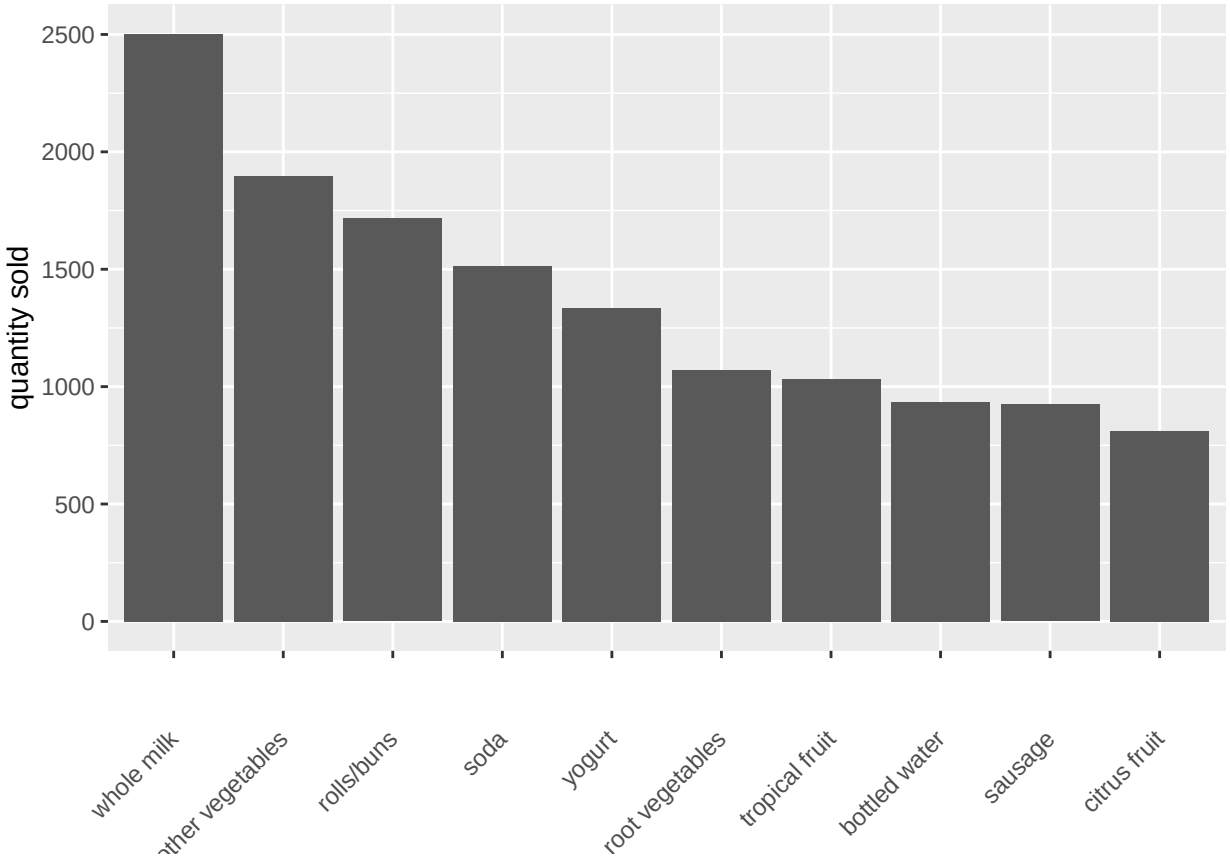
```
groceries <-
  read.csv(url("https://raw.githubusercontent.com/HwanKim2/data_repo/main/Groceries_dataset.csv"))
glimpse(groceries)
```

```
## Rows: 38,765
## Columns: 3
## $ Member_number   <int> 1808, 2552, 2300, 1187, 3037, 4941, 4501, 3803, 276...
## $ Date            <chr> "21-07-2015", "05-01-2015", "19-09-2015", "12-12-20...
## $ itemDescription <chr> "tropical fruit", "whole milk", "pip fruit", "other...
```

```
item_count <- groceries %>%
  dplyr::count(itemDescription) %>%
  arrange(desc(n))
item_count[1:10,]
```

```
##     itemDescription    n
## 1        whole milk 2502
## 2  other vegetables 1898
## 3        rolls/buns 1716
## 4              soda 1514
## 5            yogurt 1334
## 6   root vegetables 1071
## 7    tropical fruit 1032
## 8     bottled water  933
## 9           sausage  924
## 10     citrus fruit  812
```

```
ggplot(item_count[1:10,],
       aes(x=reorder(itemDescription, -n), y = n)) +
geom_bar(stat="identity") +
labs(x = "", y = "quantity sold") +
theme(axis.text.x = element_text(angle = 45, vjust = 0.5, hjust=1))
```


quantity sold
2500
2000
1500
1000
500
0
whole milk
other vegetables
rolls/buns
soda
yogurt
root vegetables
tropical fruit
bottled water
sausage
citrus fruit