

## Lab4\_stat123

Koki Itagaki

2023-02-07

#Question 1 #Take a look at the following table. The number of various crops cultivated in #a village is given in the following table.

*#a) Create a data frame named "crops.df" and then calculate the percentage of #crops.*

```
crops.df<-data.frame(Items =  
c("Wheat","Pulses","Jowar","Groundnuts","Vegitables")  
      ,crops = c(140,110,30,60,20))
```

```
crops.df
```

```
##      Items crops  
## 1    Wheat   140  
## 2   Pulses   110  
## 3    Jowar    30  
## 4 Groundnuts   60  
## 5 Vegitables   20
```

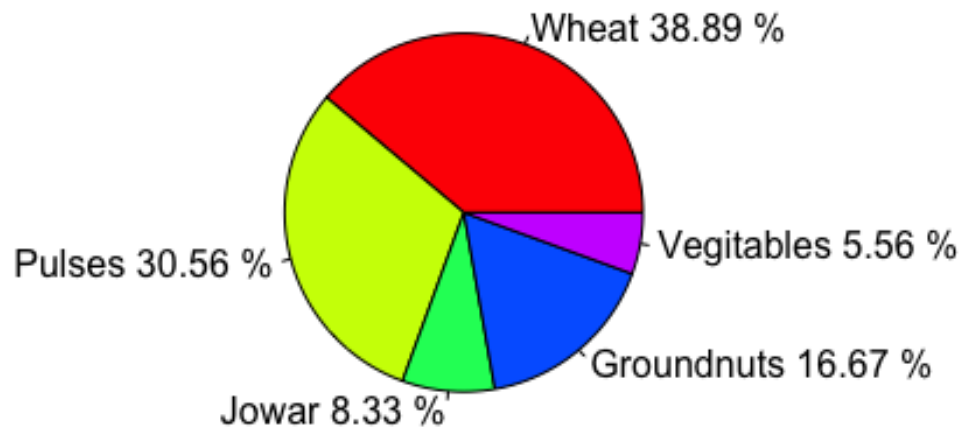
```
percent<- round(crops.df$crops/sum(crops.df$crops)*100,2)  
percent
```

```
## [1] 38.89 30.56  8.33 16.67  5.56
```

*#b) Produce a pie chart showing the items and the percentage of crops with 2 #decimals and change the color of the pie chart. (Using the paste() function) #crops*

```
segment<- paste(crops.df$Items,percent,"%")  
pie(crops.df$crops, labels = segment, col = rainbow(length(crops.df$crops)),  
    main = "The pie chart for crops")
```

## The pie chart for crops



#Question2 Download the data set Superstores.csv and save it #to whatever directory you are using for this course.

```
#a) Read the file and take a look at the dataset to familiarize yourself
#with its contents. Then create a sub-dataset that contains all data and
#columns with restrictions for the Region. The Region can be Central or East.
superstore<- read.csv("/Users/itagakikouki/stat123/Superstores.csv",
                      header = TRUE)
```

```
dim(superstore)
```

```
## [1] 128 21
```

```
head(superstore)
```

```
##   Row.ID      Order.ID Order.Date  Ship.Date      Ship.Mode Customer.ID
## 1      1 CA-2016-152156 11/8/2016 11/11/2016 Second Class    CG-12520
## 2      2 CA-2016-152156 11/8/2016 11/11/2016 Second Class    CG-12520
## 3      3 CA-2016-138688  6/12/2016  6/16/2016 Second Class    DV-13045
## 4      4 US-2015-108966 10/11/2015 10/18/2015 Standard Class    SO-20335
## 5      5 US-2015-108966 10/11/2015 10/18/2015 Standard Class    SO-20335
## 6      6 CA-2014-115812  6/9/2014  6/14/2014 Standard Class    BH-11710
##   Customer.Name Segment      Country      City      State
## 1   Claire Gute  Consumer United States Henderson Kentucky
```

```
## 2 Claire Gute Consumer United States Henderson Kentucky
## 3 Darrin Van Huff Corporate United States Los Angeles California
## 4 Sean O'Donnell Consumer United States Fort Lauderdale Florida
## 5 Sean O'Donnell Consumer United States Fort Lauderdale Florida
## 6 Brosina Hoffman Consumer United States Los Angeles California
## Postal.Code Region Product.ID Category Sub.Category
## 1 42420 South FUR-BO-10001798 Furniture Bookcases
## 2 42420 South FUR-CH-10000454 Furniture Chairs
## 3 90036 West OFF-LA-10000240 Office Supplies Labels
## 4 33311 South FUR-TA-10000577 Furniture Tables
## 5 33311 South OFF-ST-10000760 Office Supplies Storage
## 6 90032 West FUR-FU-10001487 Furniture Furnishings
## Product.Name
```

Sales

```
## 1 Bush Somerset Collection Bookcase
261.9600
## 2 Hon Deluxe Fabric Upholstered Stacking Chairs, Rounded Back
731.9400
## 3 Self-Adhesive Address Labels for Typewriters by Universal
14.6200
## 4 Bretford CR4500 Series Slim Rectangular Table
957.5775
## 5 Eldon Fold 'N Roll Cart System
22.3680
## 6 Eldon Expressions Wood and Plastic Desk Accessories, Cherry Wood
48.8600
```

```
## Quantity Discount Profit
## 1 2 0.00 41.9136
## 2 3 0.00 219.5820
## 3 2 0.00 6.8714
## 4 5 0.45 -383.0310
## 5 2 0.20 2.5164
## 6 4 0.00 14.1694
```

```
sample<-superstore[superstore$Region == c("Central","East"),]
```

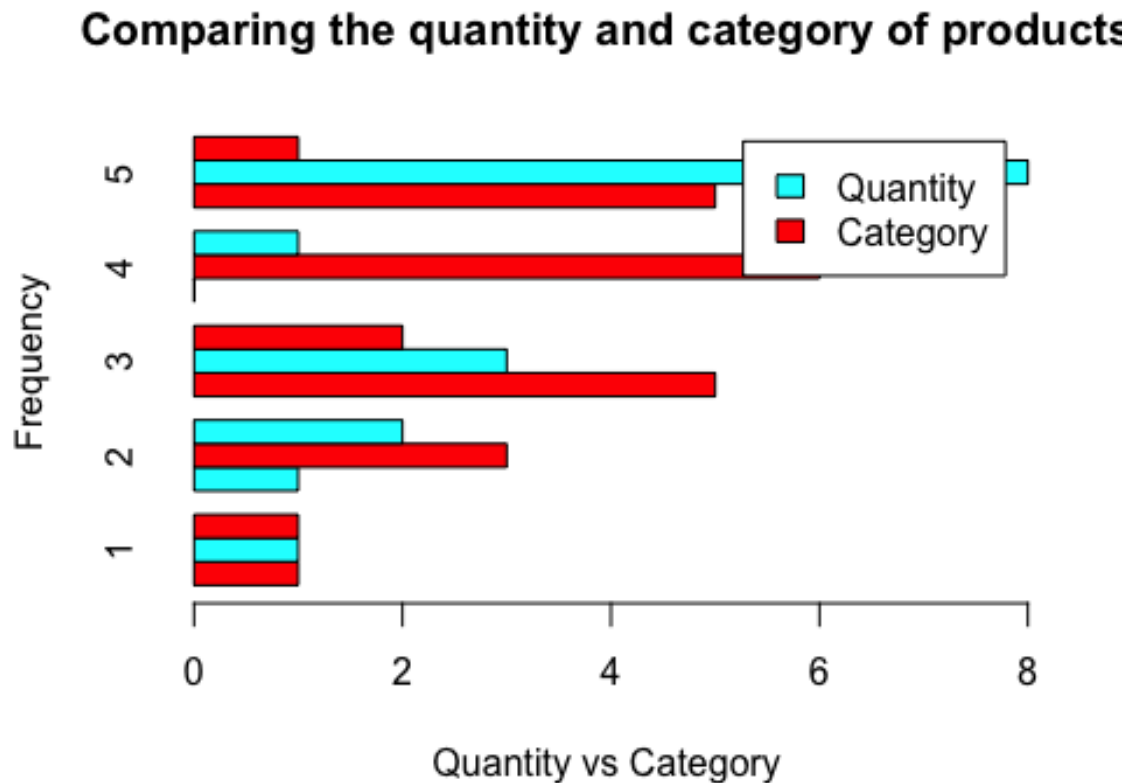
*#b) Produce a horizontal bar plot to compare the quantity and category.  
 #You should set the bars beside each other. Set the colour of the bars to be  
 #the rainbow as you saw in Chapter 10's Lecture and consider a title bar  
 #as "Comparing the quantity and category of products," the x-axis  
 #as "Quantity vs Category," and the y-axis as "Frequency."*

```
value <- table(sample$Category,sample$Quantity)
value
```

```
##
## 1 2 3 4 5
## Furniture 1 1 5 0 5
```

```
## Office Supplies 1 3 3 6 8
## Technology      1 2 2 1 1

barplot(value, main = "Comparing the quantity and category of products", xlab =
"Quantity vs Category", ylab = "Frequency", col = rainbow(2),
legend= c("Category", "Quantity"), beside = TRUE, horiz = TRUE)
```



#Question3. Use the Superstores.csv. It contains the profits in the last column.

```
#a) Produce a line plot with the x-axis as "Index," the y-axis as "Profit,"
#and the title as "Profits among superstores." The x-axis value should be
#between 0 and 100
plot(superstore$Profit, type = "l", xlab = "Index", ylab = "Profit",
      xlim = c(0,100), col = "darkblue", main = "Profits among superstores.")
```

### Profits among superstores.

