Worksheet 12

Peyton Hall

04/25/2025

Question 1

```
library(readxl)
Supermarket_data <- read_excel("~/Desktop/STAT 301/Week 14/Supermarket data.xlsx")
Supermarket_data
```

```
## # A tibble: 12 x 5
##    supermarket  meat  fish vegetable ownbrand
##    <chr>       <dbl> <dbl>     <dbl>    <dbl>
##  1 Asda         1     2.39      4.14      4.1
##  2 Anderson     1.46  3         3.61      4
##  3 Tesco        8.02  2.59      4         6
##  4 Hyvee        2.1   5.76      1.46      3.4
##  5 Kwiksave     0.43  0.53      3         0
##  6 Asdaton      4.46  0.36      3.72      3
##  7 Winco        2.73  2.4       7         3
##  8 Kwoaski      4.65  1.06      1.54      1
##  9 Waitrose     9.72  5.78     13.1      31.3
## 10 Lund         0.49  2.05      8.32      3.9
## 11 Freedy      12.7   4.86      0.38     16.3
## 12 Safeway     13.8   0.33     11.9      46.9
```

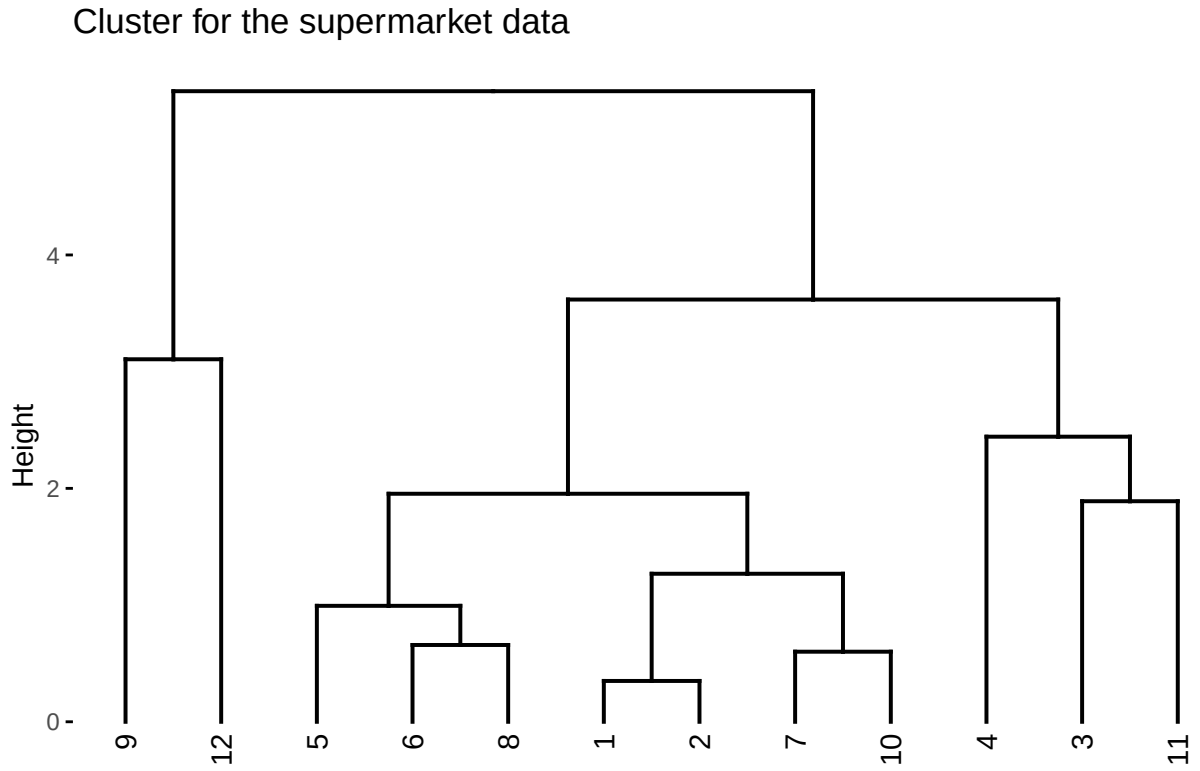
```
library(factoextra)
```

```
## Loading required package: ggplot2
```

```
## Welcome! Want to learn more? See two factoextra-related books at https://goo.gl/ve3WBa
```

```
library(stats)
market1 <- scale(Supermarket_data[,c("meat", "fish", "vegetable", "ownbrand")])
d2 <- dist(market1, method = "euclidean")
model2 <- hclust(d2, method = "complete")
fviz_dend(model2, main = "Cluster for the supermarket data")
```

```
## Warning: The ‘<scale>‘ argument of ‘guides()‘ cannot be ‘FALSE‘. Use "none" instead as
## of ggplot2 3.3.4.
## i The deprecated feature was likely used in the factoextra package.
##   Please report the issue at <https://github.com/kassambara/factoextra/issues>.
## This warning is displayed once every 8 hours.
## Call ‘lifecycle::last_lifecycle_warnings()‘ to see where this warning was
## generated.
```



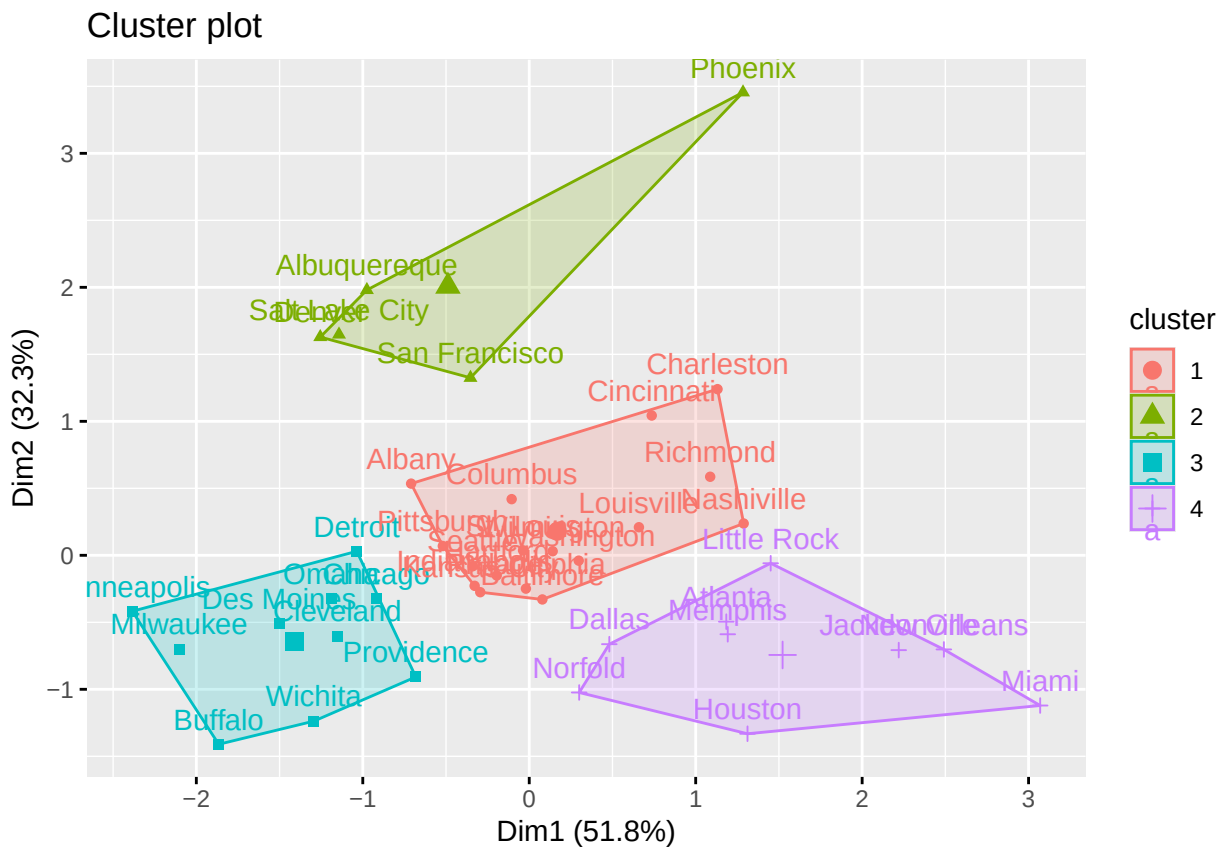
Question 2

```
library(readxl)
Weather_data <- read_excel("~/Desktop/STAT 301/Week 14/Weather data.xlsx")
Weather_data
```

```
## # A tibble: 41 x 4
##    City          Temp  Wind Precipitation
##    <chr>        <dbl> <dbl>         <dbl>
##  1 Albany        47.6   8.8          33.4
##  2 Albuquereque  56.8   8.9           7.77
##  3 Atlanta       61.5   9.1          48.3
##  4 Baltimore     55     9.6          41.3
##  5 Buffalo       47.1  12.4          36.1
##  6 Charleston    55.2   6.5          40.8
##  7 Chicago       50.6  10.4          34.4
##  8 Cincinnati    54     7.1          39.0
##  9 Cleveland     49.7  10.9          35.0
## 10 Columbus      51.5   8.6          37.0
## # i 31 more rows
```

```
x1 <- scale(Weather_data[,c (2,3,4)])

model4 <- kmeans(x1, centers = 4, nstart = 41)
rownames(x1) <- Weather_data$City
fviz_cluster(model4, data = x1)
```


Cluster plot
Phoenix
Albuquerque
Salt Lake City
Denver
San Francisco
Charleston
Cincinnati
Richmond
Albany
Columbus
Louisville
Nashville
Detroit
Little Rock
Chicago
Des Moines
Cleveland
Milwaukee
Providence
Wichita
Buffalo
Dallas
Memphis
Atlanta
Norfolk
Houston
Miami
Dim1 (51.8%)
Dim2 (32.3%)
cluster
1
2
3
4