

First Part Of Assignment:

Console:



```
> install.packages("dplyr")
trying URL 'https://cran.rstudio.com/bin/macosx/big-sur-arm64/contrib/4.4/dplyr_1.1.4.tgz'
Content type 'application/x-gzip' length 1599250 bytes (1.5 MB)
=====
downloaded 1.5 MB

The downloaded binary packages are in
  /var/folders/kd/2twdkrq94q1bq50x8twq4c700000gn/T//RtmpzuIjpB/downloaded_packages
> install.packages("ape")
also installing the dependency 'Rcpp'

trying URL 'https://cran.rstudio.com/bin/macosx/big-sur-arm64/contrib/4.4/Rcpp_1.0.14.tgz'
Content type 'application/x-gzip' length 3354635 bytes (3.2 MB)
=====
downloaded 3.2 MB

trying URL 'https://cran.rstudio.com/bin/macosx/big-sur-arm64/contrib/4.4/ape_5.8-1.tgz'
Content type 'application/x-gzip' length 3365942 bytes (3.2 MB)
=====
downloaded 3.2 MB

The downloaded binary packages are in
  /var/folders/kd/2twdkrq94q1bq50x8twq4c700000gn/T//RtmpzuIjpB/downloaded_packages
> install.packages("ggplot2")
trying URL 'https://cran.rstudio.com/bin/macosx/big-sur-arm64/contrib/4.4/ggplot2_3.5.2.tgz'
Content type 'application/x-gzip' length 4969589 bytes (4.7 MB)
=====
downloaded 4.7 MB

The downloaded binary packages are in
  /var/folders/kd/2twdkrq94q1bq50x8twq4c700000gn/T//RtmpzuIjpB/downloaded_packages
> library(dplyr)

Attaching package: 'dplyr'

The following objects are masked from 'package:stats':

  filter, lag

The following objects are masked from 'package:base':
```

Source

Console Terminal x Background Jobs x

R 4.4.2 · ~/

```
> library(ape)
```

Attaching package: 'ape'

The following object is masked from 'package:dplyr':

where

```
> library(ggplot2)
```

```
> data <- read.csv("/mnt/data/TomatoFirst (2).csv")
```

Error in file(file, "rt") : cannot open the connection

In addition: Warning message:

In file(file, "rt") :

cannot open file '/mnt/data/TomatoFirst (2).csv': No such file or directory

```
> data <- read.csv(file.choose())
```

```
> View(data)
```

```
> tomato_data <- data %>%
```

```
+ select(Tomato, Price, Sweet, Color, Texture, Overall)
```

```
> View(tomato_data)
```

```
> rownames(tomato_data) <- tomato_data$Tomato
```

```
> head(tomato_data)
```

| | Tomato | | | |
|--------------------|---------|--------------------|-------|--|
| Simpson SM | | Simpson SM | | |
| Tuttorosso (blue) | | Tuttorosso (blue) | | |
| Tuttorosso (green) | | Tuttorosso (green) | | |
| La Fede SM DOP | | La Fede SM DOP | | |
| Cento SM DOP | | Cento SM DOP | | |
| Cento Organic | | Cento Organic | | |
| | Price | Sweet | Color | |
| Simpson SM | 3.99 | 2.8 | 3.7 | |
| Tuttorosso (blue) | 2.99 | 3.3 | 3.4 | |
| Tuttorosso (green) | 0.99 | 2.8 | 3.3 | |
| La Fede SM DOP | 3.99 | 2.6 | 3.0 | |
| Cento SM DOP | 5.49 | 3.3 | 2.9 | |
| Cento Organic | 4.99 | 3.2 | 2.9 | |
| | Texture | Overall | | |
| Simpson SM | 3.4 | 3.4 | | |
| Tuttorosso (blue) | 3.0 | 2.9 | | |
| Tuttorosso (green) | 2.8 | 2.9 | | |
| La Fede SM DOP | 2.3 | 2.8 | | |
| Cento SM DOP | 2.8 | 3.1 | | |
| Cento Organic | 3.1 | 2.9 | | |

```
> distances <- dist(tomato_data)
```

Warning message:

In dist(tomato_data) : NAs introduced by coercion

Source

Console Terminal x Background Jobs x

R 4.4.2 · ~/

```
cannot open file '/mnt/data/TomatoFirst (2).csv': No such file or directory
```

```
> data <- read.csv(file.choose())
> View(data)
> tomato_data <- data %>%
+   select(Tomato, Price, Sweet, Color, Texture, Overall)
> View(tomato_data)
> rownames(tomato_data) <- tomato_data$Tomato
> head(tomato_data)
```

| | Tomato |
|--------------------|--------------------|
| Simpson SM | Simpson SM |
| Tuttorosso (blue) | Tuttorosso (blue) |
| Tuttorosso (green) | Tuttorosso (green) |
| La Fede SM DOP | La Fede SM DOP |
| Cento SM DOP | Cento SM DOP |
| Cento Organic | Cento Organic |

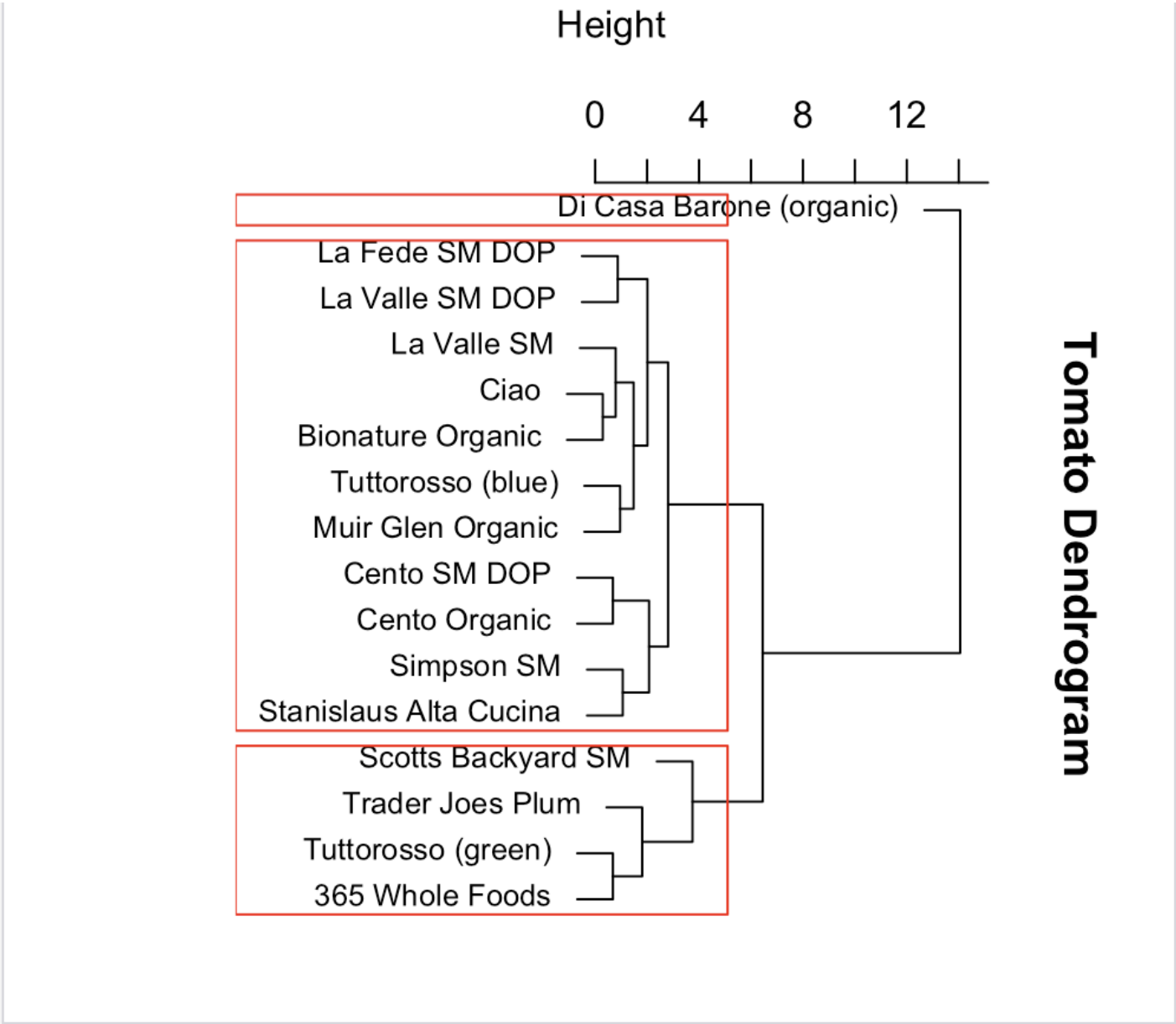
| | Price | Sweet | Color |
|--------------------|-------|-------|-------|
| Simpson SM | 3.99 | 2.8 | 3.7 |
| Tuttorosso (blue) | 2.99 | 3.3 | 3.4 |
| Tuttorosso (green) | 0.99 | 2.8 | 3.3 |
| La Fede SM DOP | 3.99 | 2.6 | 3.0 |
| Cento SM DOP | 5.49 | 3.3 | 2.9 |
| Cento Organic | 4.99 | 3.2 | 2.9 |

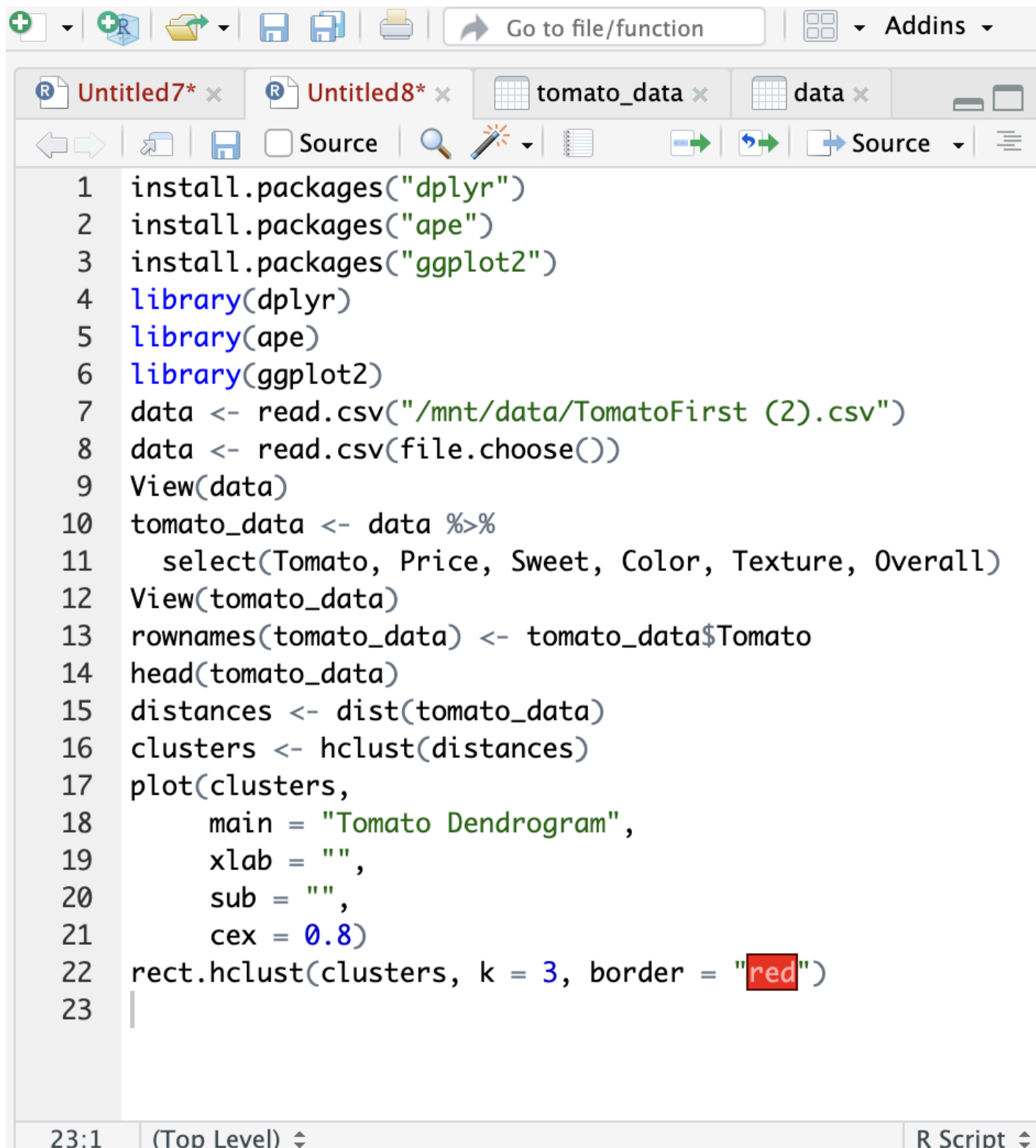
| | Texture | Overall |
|--------------------|---------|---------|
| Simpson SM | 3.4 | 3.4 |
| Tuttorosso (blue) | 3.0 | 2.9 |
| Tuttorosso (green) | 2.8 | 2.9 |
| La Fede SM DOP | 2.3 | 2.8 |
| Cento SM DOP | 2.8 | 3.1 |
| Cento Organic | 3.1 | 2.9 |

```
> distances <- dist(tomato_data)
Warning message:
In dist(tomato_data) : NAs introduced by coercion
```

```
> clusters <- hclust(distances)
> plot(clusters,
+       main = "Tomato Dendrogram",
+       xlab = "",
+       sub = "",
+       cex = 0.8)
> rect.hclust(clusters, k = 3, border = "red")
> plot(as.phylo(clusters),
+       type = "fan",
+       tip.color = c("red", "blue", "green")[cutree(clusters, 3)],
+       main = "Fan Dendrogram of Tomatoes")
>
```

Tomato Dendrogram:

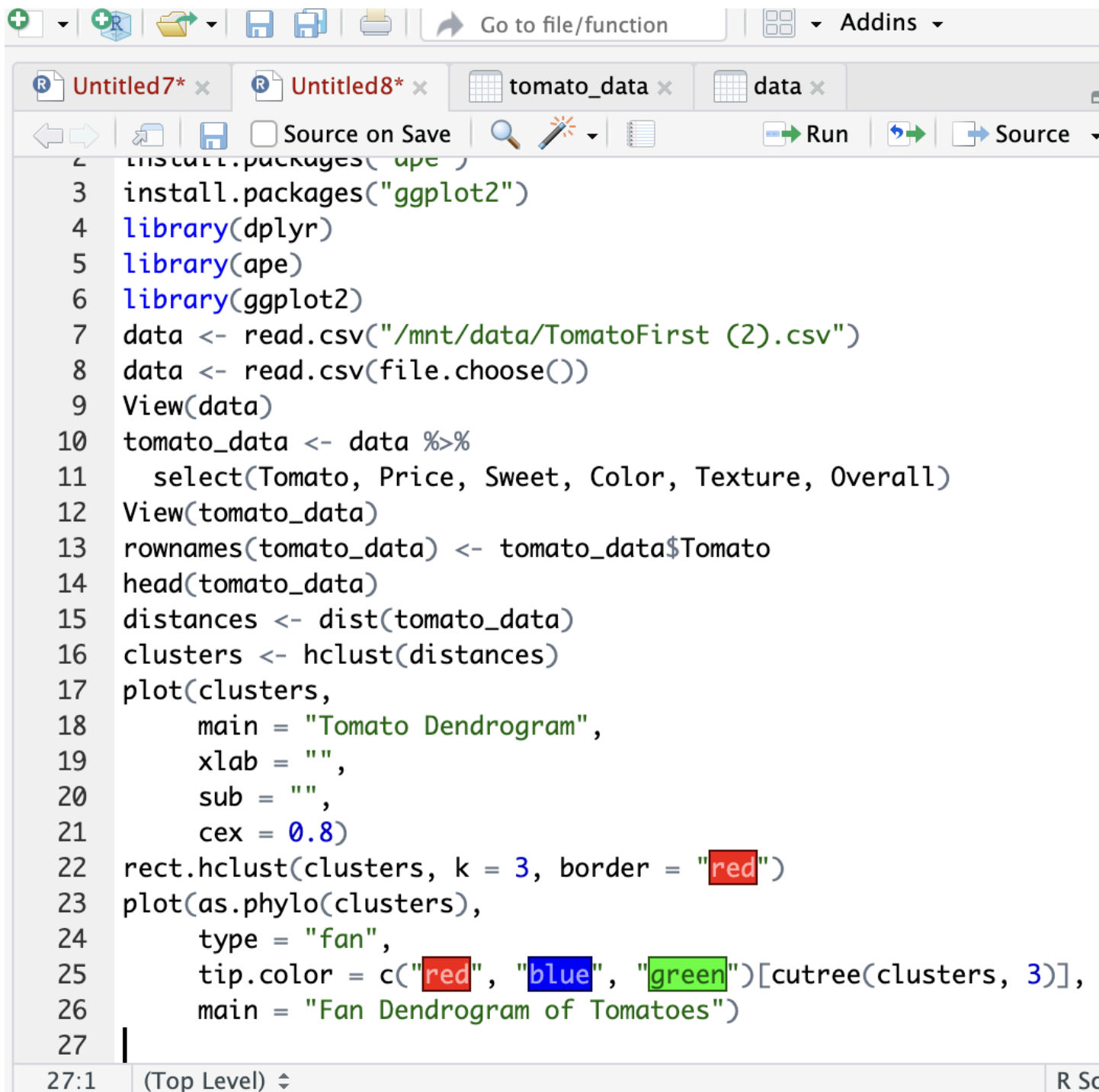




The image shows a screenshot of the RStudio application. At the top, there is a toolbar with icons for file operations (new, open, save, print) and a search bar labeled 'Go to file/function'. Below the toolbar, the 'Source' pane is active, displaying an R script. The script is numbered 1 through 23. The script performs the following actions: 1. Install 'dplyr' package. 2. Install 'ape' package. 3. Install 'ggplot2' package. 4. Load 'dplyr' library. 5. Load 'ape' library. 6. Load 'ggplot2' library. 7. Read a CSV file from '/mnt/data/TomatoFirst (2).csv' into a variable named 'data'. 8. Prompt the user to choose a file to read into 'data'. 9. View the 'data' object. 10. Create a new object 'tomato_data' by selecting columns 'Tomato', 'Price', 'Sweet', 'Color', 'Texture', and 'Overall' from 'data'. 11. View the 'tomato_data' object. 12. Set the row names of 'tomato_data' to the values in the 'Tomato' column. 13. Display the first few rows of 'tomato_data'. 14. Calculate pairwise distances between rows of 'tomato_data'. 15. Perform hierarchical clustering on the distances. 16. Plot the hierarchical clustering result. 17. Set the main title of the plot to 'Tomato Dendrogram'. 18. Set the x-axis label to an empty string. 19. Set the subtitle to an empty string. 20. Set the font size (cex) to 0.8. 21. Perform rectangular hierarchical clustering with k=3 clusters and red borders. 22. The script ends at line 23. The status bar at the bottom shows '23:1' and '(Top Level)'. The file type is 'R Script'.

```
1 install.packages("dplyr")
2 install.packages("ape")
3 install.packages("ggplot2")
4 library(dplyr)
5 library(ape)
6 library(ggplot2)
7 data <- read.csv("/mnt/data/TomatoFirst (2).csv")
8 data <- read.csv(file.choose())
9 View(data)
10 tomato_data <- data %>%
11   select(Tomato, Price, Sweet, Color, Texture, Overall)
12 View(tomato_data)
13 rownames(tomato_data) <- tomato_data$Tomato
14 head(tomato_data)
15 distances <- dist(tomato_data)
16 clusters <- hclust(distances)
17 plot(clusters,
18       main = "Tomato Dendrogram",
19       xlab = "",
20       sub = "",
21       cex = 0.8)
22 rect.hclust(clusters, k = 3, border = "red")
23
```

EXTRA CREDIT:



The screenshot shows the RStudio IDE with the following elements:

- Toolbar:** Includes icons for adding files, saving, printing, and navigating. The text "Go to file/function" is visible.
- Session Information:** Shows "Addins" and a dropdown menu.
- File Explorer:** Displays open files: "Untitled7*", "Untitled8*", "tomato_data", and "data".
- Source Editor:** Contains R code for installing packages, reading a CSV file, and creating dendrograms. The code is as follows:

```
2 install.packages("ape")
3 install.packages("ggplot2")
4 library(dplyr)
5 library(ape)
6 library(ggplot2)
7 data <- read.csv("/mnt/data/TomatoFirst (2).csv")
8 data <- read.csv(file.choose())
9 View(data)
10 tomato_data <- data %>%
11   select(Tomato, Price, Sweet, Color, Texture, Overall)
12 View(tomato_data)
13 rownames(tomato_data) <- tomato_data$Tomato
14 head(tomato_data)
15 distances <- dist(tomato_data)
16 clusters <- hclust(distances)
17 plot(clusters,
18       main = "Tomato Dendrogram",
19       xlab = "",
20       sub = "",
21       cex = 0.8)
22 rect.hclust(clusters, k = 3, border = "red")
23 plot(as.phylo(clusters),
24       type = "fan",
25       tip.color = c("red", "blue", "green")[cutree(clusters, 3)],
26       main = "Fan Dendrogram of Tomatoes")
27
```
- Console:** Shows the current line number and column: "27:1 (Top Level)".
- Status Bar:** Displays "R Sc" (likely R Session).

Fan Dendrogram of Tomatoes:

