2.

a) Population: the spending and consumption patterns of coffee shop's customers

Sample: Consumption patterns of customers from population that are investigated.

Individual: Consumption pattern of one customer.

b) Response Variables: Total consumption expand (Quantitative), Average Times of consumption per month (Quantitative); Regular customer or not (Categorical), Drinks buyer/Snacks buyer/Both (Categorical)

Explannatory covariates: Temperature (Quantitative), Selling Price (Quantitative); With competitor/Without competitor (Categorical), With promotes/Without promotes (Categorical)

c) We can get the selling and income data from dealer. As for the competitors, it is practical to collect data by searching around.

3. Using R

b. The data generated is like follows:

> grades

[1] "C" "C" "B" "B" "C" "C" "C" "A" "C" "B" "F" "F" "C" "D" "B" "B" "B"

[18] "B" "C" "A" "C" "C" "A" "C" "B" "A" "C" "B" "B" "C"

> ages

[1] 19.0 18.8 23.7 20.0 24.8 27.2 19.3 24.3 22.0 28.8 24.3 22.3 12.7

[14] 16.7 21.6 23.0 18.1 23.7 24.1 23.1 23.8 18.4 17.8 24.5 20.5 25.4

[27] 21.9 13.7 18.0 21.1

c. length(grades) will get the numbers of grades, which is 30.

sort(grades) will place the grades by increasing order.

[1] "A" "A" "A" "A" "B" "B" "B" "B" "B" "B" "B" "B" "B" "B" "C" "C" "C"

[18] "C" "C" "C" "C" "C" "C" "C" "C" "C" "C" "D" "F" "F"

table(grades) will give us a table like this:

grades

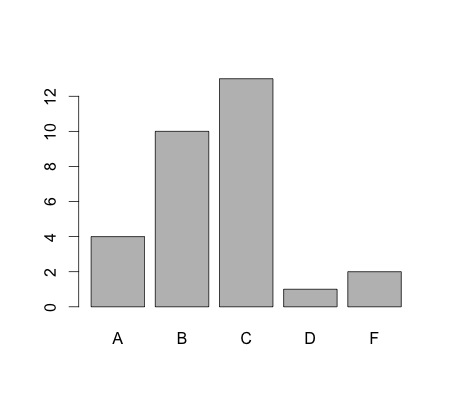
A B C D F

4 10 13 1 2

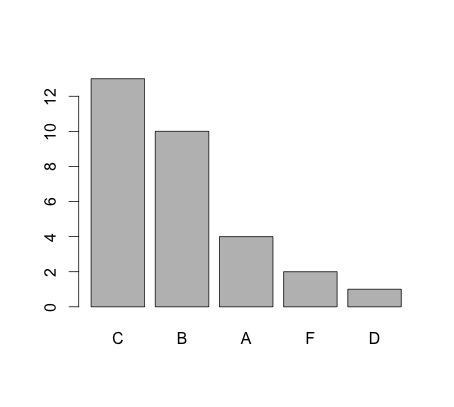
table(grades)/length(grades)

The numbers of different grades will dived by the total number of grades.

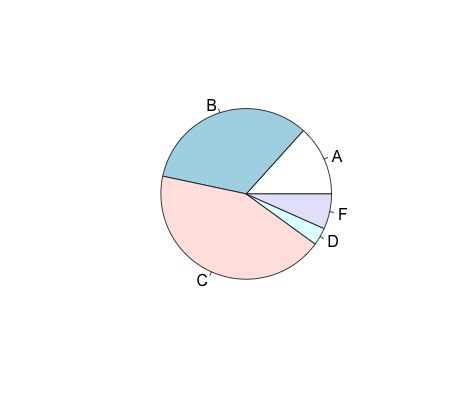
barplot(table(grades)) will generate a plot like follows based on the table we gived.



barplot(sort(table(grades), decreasing = TRUE)) will sort the table by the order that we gave and draw the bar diagram.



pie(table(grades)) will give us the pie diagram based on the table we gave.



d. sort(ages) will sort the data set of ages by increasing order.

Stem(ages) will give us the stemplot of ages as follows.

The decimal point is at the |

12 | 77

14 |

16 | 78

18 | 014803

20 | 05169

22 | 0301778

24 | 133584

26 | 2

28 | 8

stem(ages, scale = 2) will give the actual stemplot like this:

The decimal point is at the |

12 | 7

13 | 7

14 |

15 |

16 | 7

17 | 8

18 | 0148

19 | 03

20 | 05

21 | 169

22 | 03

23 | 01778

24 | 13358

25 | 4

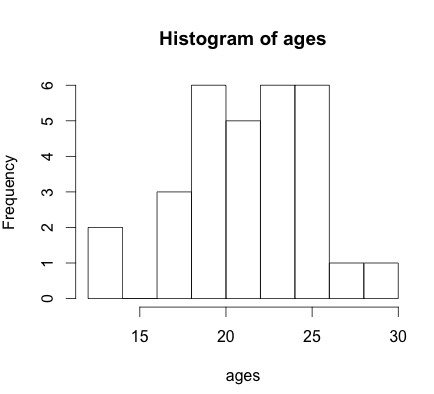
26 |

27 | 2

28 | 8

The scale variable will scale the stemplot. When the scale is 1, it is the actual stemplot without any redundancy data.

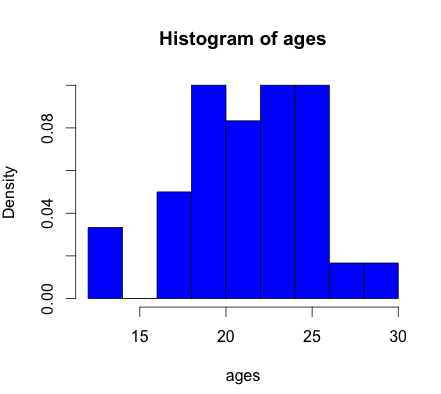
hist(ages) will give us the histogram of ages like this:



hist(ages, breaks = 2)



hist(ages, freq = FALSE, col = “blue”)

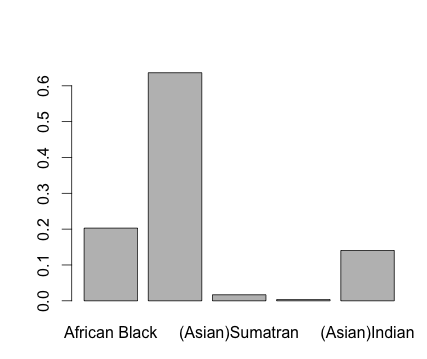


The freq argument can control the function to draw the histogram as density of frequency. Col controls the color of the bars. Breaks decide the width of the graph.

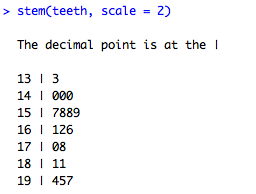
4. Korean geneology

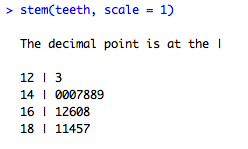
(a)

(b)



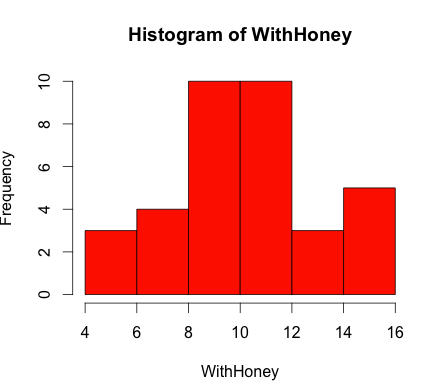
6. Stem and leaf of ancient teeth

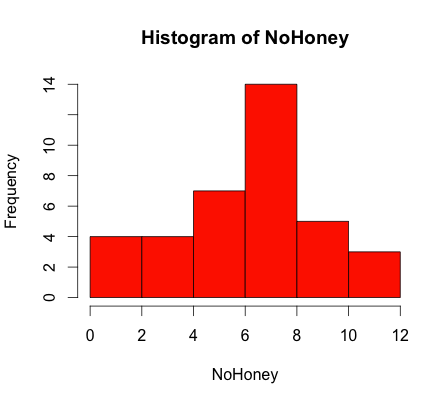


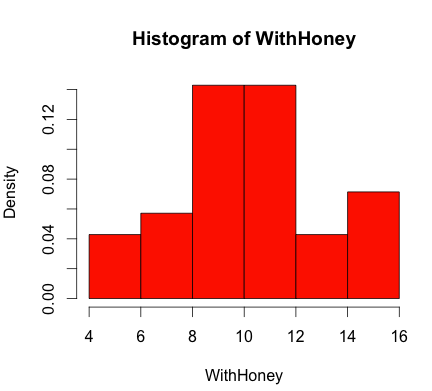


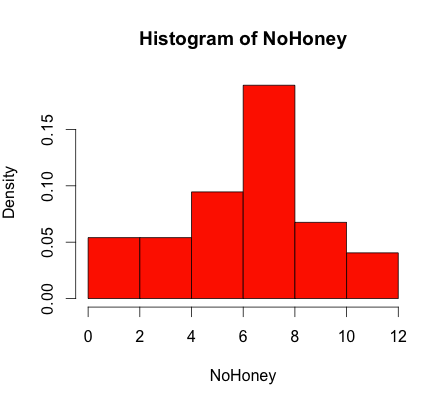
At about 14mm depth has the most frequency.

7. Honey Histograms

(a) 



(b) 



(c)