HW1-108020033

Che-Wei, Chang

2023-02-17

1.What is the 5th element in the original list of ages?

```
my_data <- read.delim("customers.txt")# import data from customers.txt in my_data
age <- my_data[,1]                    # the array age will get the data in my_data
age[5]
```

```
## [1] 45
```

We put the data into a vector which called age and use age[5] to show the result

2.What is the fifth lowest age?

```
sorted_age <- sort(age, decreasing = F) # sort the data in ascending order and
sorted_age[5]                           # put the result into sorted_age
```

```
## [1] 19
```

We use a vector to store the sorted age in ascending order and sorted_age[5] will show the fifth lowest age

3.Extract the five lowest ages together

```
extract_sorted_age <- sorted_age[1:5]   # put the 1st ~ 5th element of sorted_age
extract_sorted_age                      # into extract_sorted_age
```

```
## [1] 18 19 19 19 19
```

We use vector called extract_sorted_age to store the five lowest ages together and show the result

4.Get the five highest ages by first sorting them in decreasing order first.

```
sorted_age <- sort(age, decreasing = TRUE)# sort the data in decreasing order and
extract_sorted_age <- sorted_age[1:5]     # put into sorted_age and store into the
extract_sorted_age                        # extract_sorted_age
```

```
## [1] 85 83 82 82 81
```

We use a vector to store the sorted age in decreasing order and use extract_sorted_age to store the result

5.What is the average (mean) age?

```
mean(age)                                  # use mean() to get the average age
```

```
## [1] 46.80702
```

We use mean() to get the average age

6.What is the standard deviation of ages?

```
sd(age)                        # use sd() to get the standard deviation of age
```

```
## [1] 16.3698
```

We use sd() to get standard deviation of ages

7.Make a new variable called age_diff, with the difference between each age and the mean age

```
age_diff <- age - mean(age) # let the element of age minus the average of age
age_diff                    # show the result
```

```
##   [1]   2.1929825  22.1929825  -5.8070175  26.1929825  -1.8070175  24.1929825
##   [7]   3.1929825  -3.8070175  23.1929825 -14.8070175   0.1929825  30.1929825
##  [13]  17.1929825   3.1929825   3.1929825  -1.8070175   2.1929825   0.1929825
##  [19]  15.1929825   3.1929825   0.1929825  25.1929825   0.1929825  16.1929825
##  [25] -25.8070175   2.1929825   3.1929825   1.1929825 -11.8070175  30.1929825
##  [31]   1.1929825   1.1929825   3.1929825   0.1929825 -17.8070175  -4.8070175
##  [37]  -4.8070175  38.1929825  -1.8070175   2.1929825  -1.8070175  -3.8070175
##  [43]   2.1929825  21.1929825  -4.8070175   1.1929825  25.1929825  32.1929825
##  [49]   1.1929825   3.1929825   0.1929825  -1.8070175 -16.8070175  29.1929825
##  [55] -15.8070175   2.1929825  27.1929825  25.1929825   1.1929825   2.1929825
##  [61]  26.1929825   3.1929825   0.1929825   0.1929825  36.1929825  25.1929825
##  [67]  28.1929825   3.1929825   3.1929825   2.1929825   1.1929825  -1.8070175
##  [73]   2.1929825   2.1929825   2.1929825  25.1929825   3.1929825  28.1929825
##  [79]  27.1929825  25.1929825  27.1929825  29.1929825   2.1929825   3.1929825
##  [85]  29.1929825 -10.8070175  -1.8070175 -11.8070175 -22.8070175  -1.8070175
##  [91]   3.1929825  -4.8070175 -24.8070175  13.1929825  12.1929825  -1.8070175
##  [97]   4.1929825  -0.8070175   0.1929825 -12.8070175  16.1929825  24.1929825
## [103]  -9.8070175 -25.8070175  -3.8070175 -14.8070175   0.1929825 -11.8070175
## [109]  23.1929825 -20.8070175  16.1929825   7.1929825  -1.8070175   0.1929825
## [115] -20.8070175 -11.8070175 -24.8070175 -15.8070175  23.1929825   4.1929825
## [121]  -9.8070175  -5.8070175   6.1929825 -12.8070175  -1.8070175 -12.8070175
## [127]  -3.8070175   3.1929825 -17.8070175   2.1929825  -0.8070175  -2.8070175
## [133] -20.8070175   2.1929825   1.1929825 -20.8070175 -12.8070175 -21.8070175
## [139]  -8.8070175 -21.8070175  31.1929825  -1.8070175 -15.8070175   0.1929825
## [145]  10.1929825 -18.8070175  28.1929825   2.1929825 -20.8070175   2.1929825
## [151] -12.8070175 -21.8070175   2.1929825 -12.8070175 -27.8070175 -14.8070175
```

```
## [157]   5.1929825  26.1929825  -7.8070175 -15.8070175   1.1929825  35.1929825
## [163] -13.8070175 -16.8070175  -9.8070175 -13.8070175   0.1929825 -17.8070175
## [169]   0.1929825  -9.8070175 -17.8070175  -6.8070175  15.1929825   1.1929825
## [175] -10.8070175  -5.8070175  10.1929825  10.1929825 -12.8070175 -21.8070175
## [181]  31.1929825 -23.8070175 -14.8070175  -5.8070175 -26.8070175  26.1929825
## [187]   2.1929825   3.1929825  -0.8070175   3.1929825 -19.8070175  -1.8070175
## [193] -17.8070175   9.1929825  28.1929825   6.1929825   0.1929825  -7.8070175
## [199]  31.1929825  -3.8070175  -1.8070175   5.1929825   1.1929825 -10.8070175
## [205]  31.1929825   0.1929825 -23.8070175 -12.8070175   2.1929825 -21.8070175
## [211]  -0.8070175  -6.8070175   3.1929825  -9.8070175   4.1929825 -11.8070175
## [217]  -1.8070175   2.1929825 -25.8070175  -9.8070175  -4.8070175  10.1929825
## [223]   2.1929825  -6.8070175   0.1929825   2.1929825   5.1929825  -4.8070175
## [229]   1.1929825 -18.8070175 -13.8070175   2.1929825   6.1929825 -25.8070175
## [235]  -8.8070175 -26.8070175 -14.8070175  30.1929825  -1.8070175   2.1929825
## [241] -25.8070175   1.1929825   3.1929825  15.1929825  -7.8070175  -1.8070175
## [247]  -2.8070175  -0.8070175  28.1929825  -4.8070175  -0.8070175   3.1929825
## [253]  23.1929825  -9.8070175  -8.8070175  -0.8070175 -14.8070175  -2.8070175
## [259]   3.1929825 -16.8070175  -7.8070175  -8.8070175 -19.8070175 -27.8070175
## [265]   2.1929825 -27.8070175  -6.8070175  18.1929825 -19.8070175   3.1929825
## [271]   1.1929825  -8.8070175  -2.8070175  23.1929825  17.1929825  25.1929825
## [277]   2.1929825 -14.8070175   2.1929825   2.1929825  26.1929825 -16.8070175
## [283] -16.8070175   8.1929825  -4.8070175  11.1929825  32.1929825 -18.8070175
## [289] -13.8070175 -20.8070175 -10.8070175 -15.8070175  26.1929825 -15.8070175
## [295]  24.1929825  21.1929825   3.1929825 -21.8070175  34.1929825 -22.8070175
## [301]   3.1929825 -25.8070175 -23.8070175   4.1929825  15.1929825  33.1929825
## [307]  19.1929825 -17.8070175 -16.8070175  -6.8070175 -27.8070175  24.1929825
## [313]  32.1929825  30.1929825 -14.8070175  -6.8070175   2.1929825 -20.8070175
## [319]   2.1929825  20.1929825   9.1929825 -22.8070175   0.1929825 -18.8070175
## [325]  11.1929825  -1.8070175 -27.8070175  25.1929825 -12.8070175  -0.8070175
## [331] -27.8070175 -13.8070175  33.1929825 -16.8070175  26.1929825 -26.8070175
## [337] -27.8070175  -6.8070175  29.1929825   1.1929825  -8.8070175  29.1929825
## [343]  29.1929825  -9.8070175 -11.8070175 -20.8070175 -21.8070175  20.1929825
## [349] -15.8070175 -19.8070175 -10.8070175 -25.8070175 -18.8070175  -7.8070175
## [355]   2.1929825  -1.8070175  13.1929825   1.1929825  -1.8070175   0.1929825
## [361] -19.8070175  32.1929825  -1.8070175   4.1929825 -23.8070175  27.1929825
## [367] -15.8070175 -26.8070175   3.1929825 -16.8070175  35.1929825  23.1929825
## [373]  -3.8070175 -26.8070175   3.1929825   1.1929825 -28.8070175  -1.8070175
## [379]  15.1929825  -5.8070175  24.1929825 -27.8070175  26.1929825 -20.8070175
## [385]  28.1929825  -5.8070175  -0.8070175   2.1929825   2.1929825 -23.8070175
## [391]  27.1929825   6.1929825 -23.8070175   4.1929825  24.1929825   3.1929825
## [397]   3.1929825  20.1929825  27.1929825
```

8.What is the average "difference between each age and the mean age"?

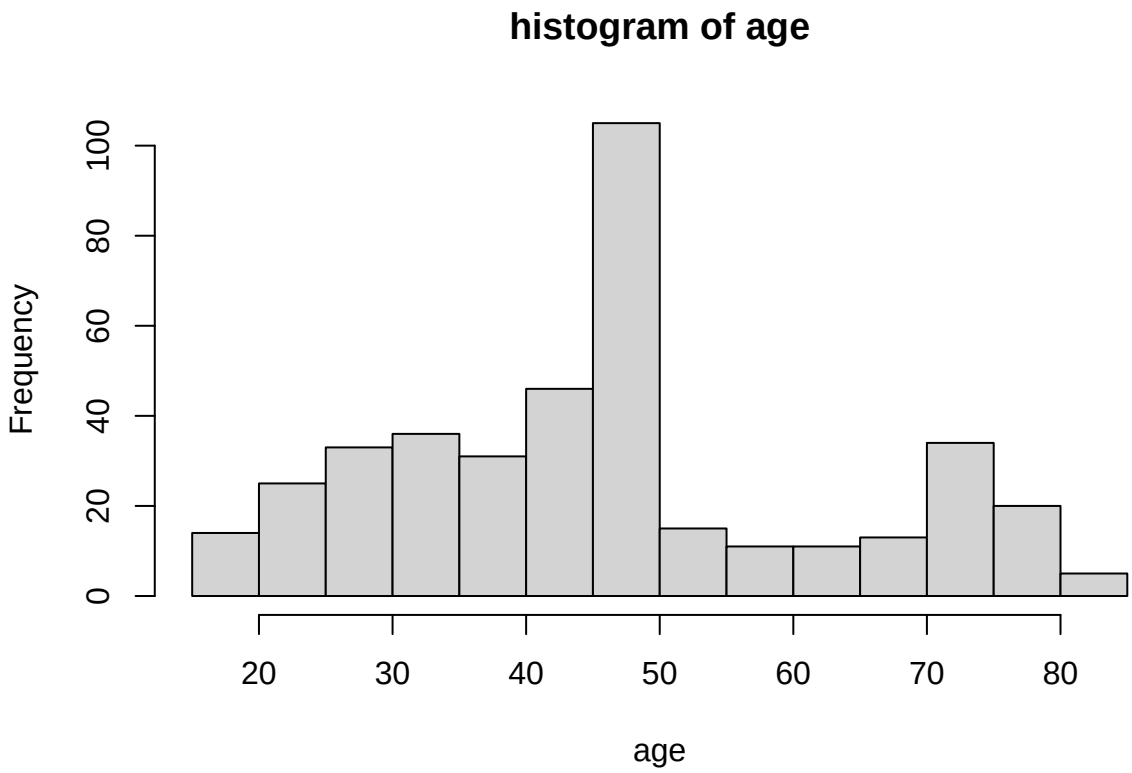
```
mean(age_diff)
```

```
## [1] -1.623275e-15
```

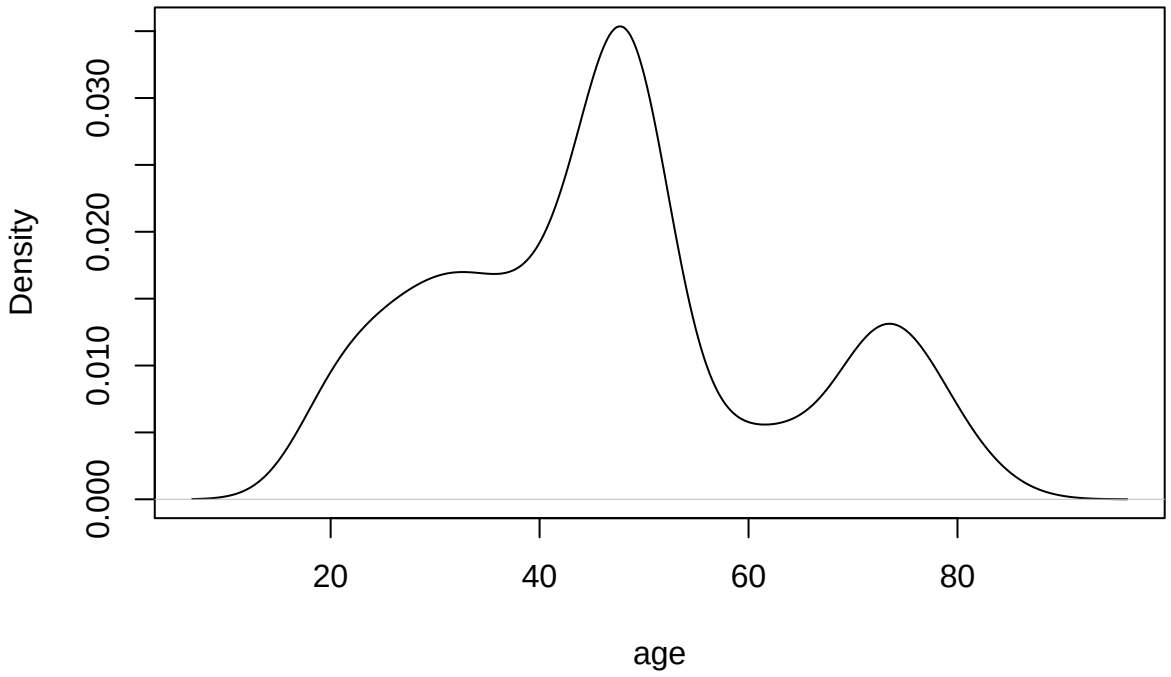
Since computer uses exponent notation to represent the floating number, the answer is very close but not equal to 0.

9.Visualize the raw data as we did in class: (a) histogram, (b) density plot, (c) box-plot+stripchart

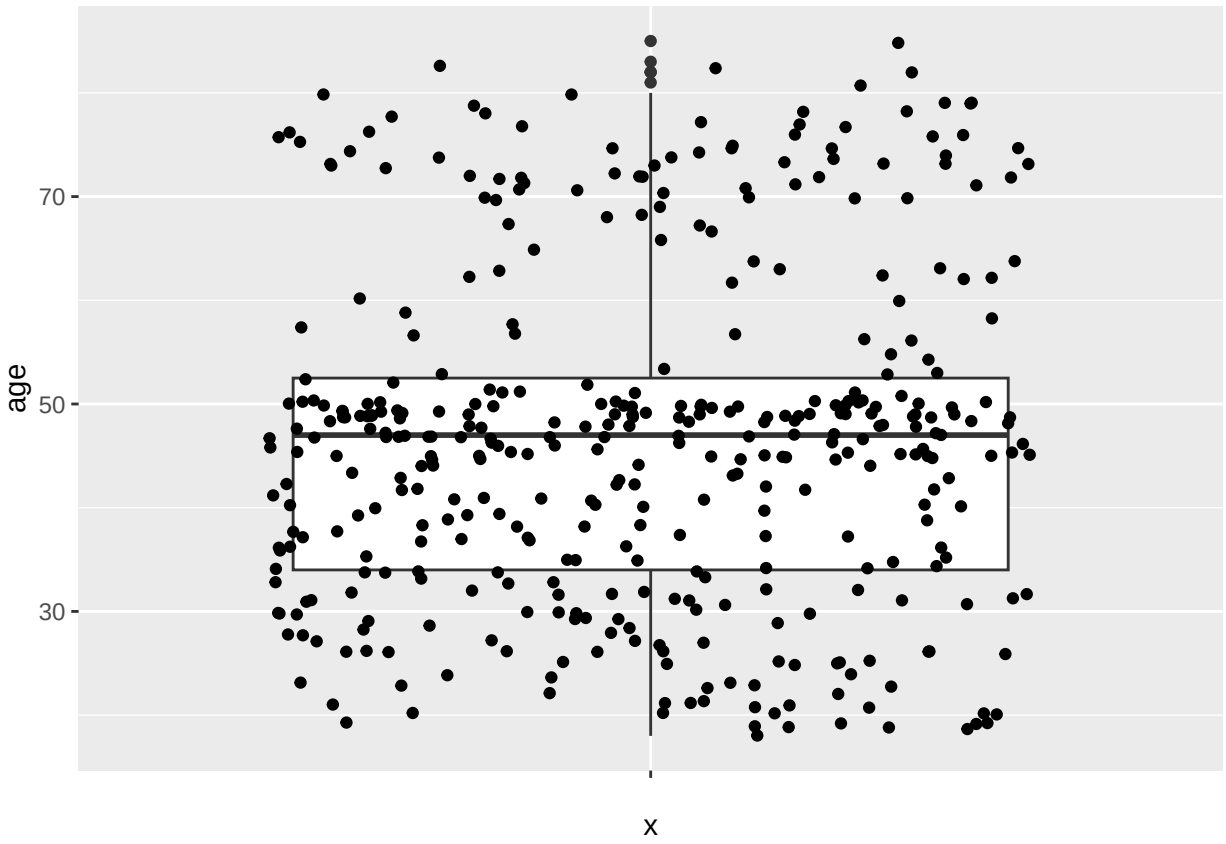
```
# (a) Plot a histogram of age
hist(age, main = "histogram of age", xlab = "age")  # draw the histogram and show the result
```




```
# (b) density plot
plot(density(age), main = "Density plot of age", xlab = "age")  # draw the density plot
```




```
# (c) boxplot + stripchart
library(ggplot2)                                    # use ggplot2 to draw boxplot + stripchart
ggplot(my_data, aes(x = "", y = age)) +
  geom_boxplot()+
  geom_jitter()
```


70
50
30
age
x