Student ID: 112077423

```
library(ggplot2)
```

```
# read file
customers_df <- read.delim('customers.txt')
# print first 6 elements in dataframe
head(customers_df)
```

```
##   age
## 1  49
## 2  69
## 3  41
## 4  73
## 5  45
## 6  71
```

Question 1

Find the 5th element in the original list.

```
fifth <- customers_df[5,]
print(paste('The 5th element is', fifth, sep=' '))
```

```
## [1] "The 5th element is 45"
```

Question 2

Find the fifth lowest age.

```
# sort in ascending order
tmp <- customers_df[order(customers_df$age, decreasing=FALSE), ]
fifth_lowest <- tmp[5]
print(paste('The 5th lowest element is', fifth_lowest, sep=' '))
```

```
## [1] "The 5th lowest element is 19"
```

Question 3

Extract the five lowest ages together.

```
five_lowest <- head(tmp, 5)
five_lowest
```

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

Student ID: 112077423

Question 4

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

```
# sort in descending order
tmp2 <- customers_df[order(customers_df$age, decreasing=TRUE), ]
five_highest <- head(tmp2, 5)
five_highest
```

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

Question 5

Find the average (mean) age.

```
mean_age <- mean(customers_df$age)
print(paste('The average is', mean_age, sep=' '))
```

```
## [1] "The average is 46.8070175438597"
```

Question 6

Find the standard deviation of ages.

```
std <- sd(customers_df$age)
print(paste('The standard deviation is', std, sep=' '))
```

```
## [1] "The standard deviation is 16.3697954137071"
```

Question 7

Create a new variable with the difference between each age and the mean age.

```
age_diff <- customers_df$age - mean_age
```

Question 8

Find the average of age_diff.

```
diff_mean <- mean(age_diff)
diff_mean
```

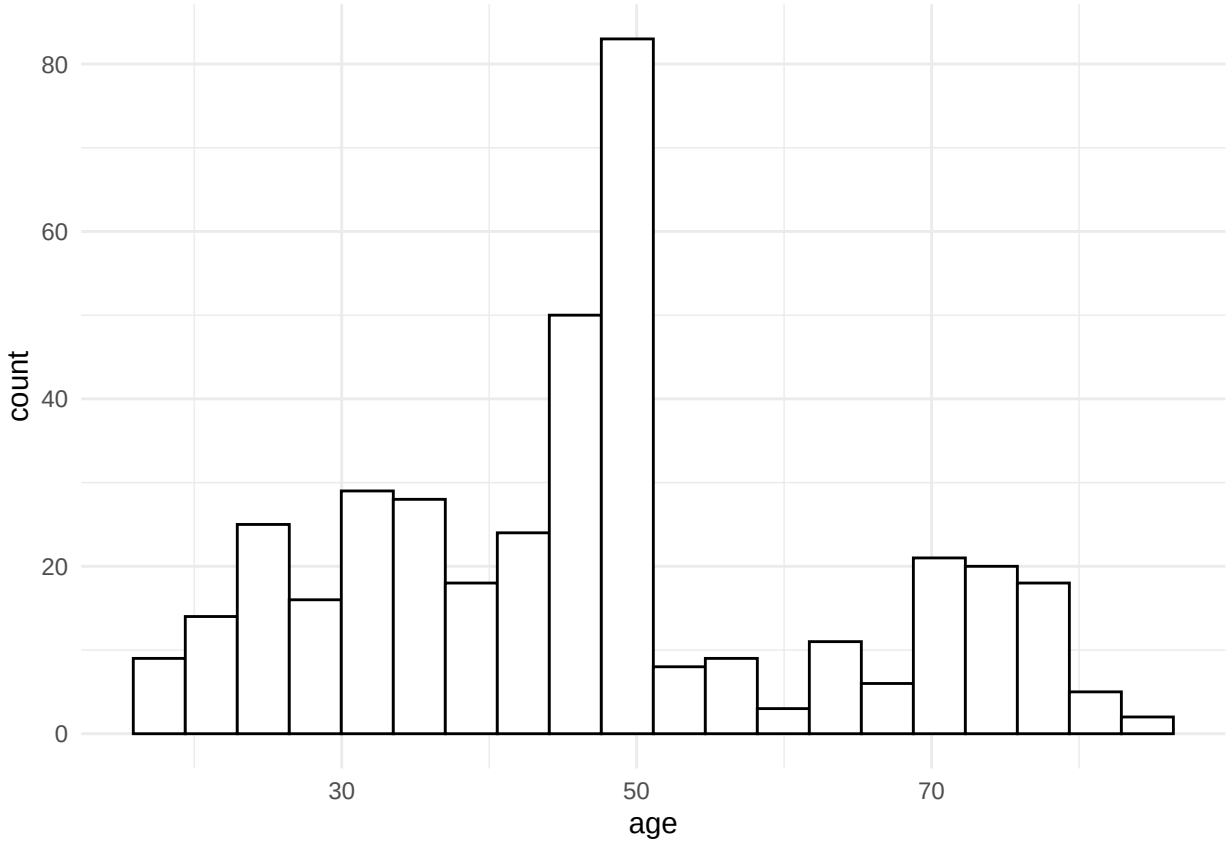
```
## [1] -1.623275e-15
```

Student ID: 112077423

Question 9a

Visualize the raw data in form of histogram.

```
ggplot(data=customers_df, aes(x=age)) +
  geom_histogram(color="black", fill="white", bins=20) +
  theme_minimal()
```

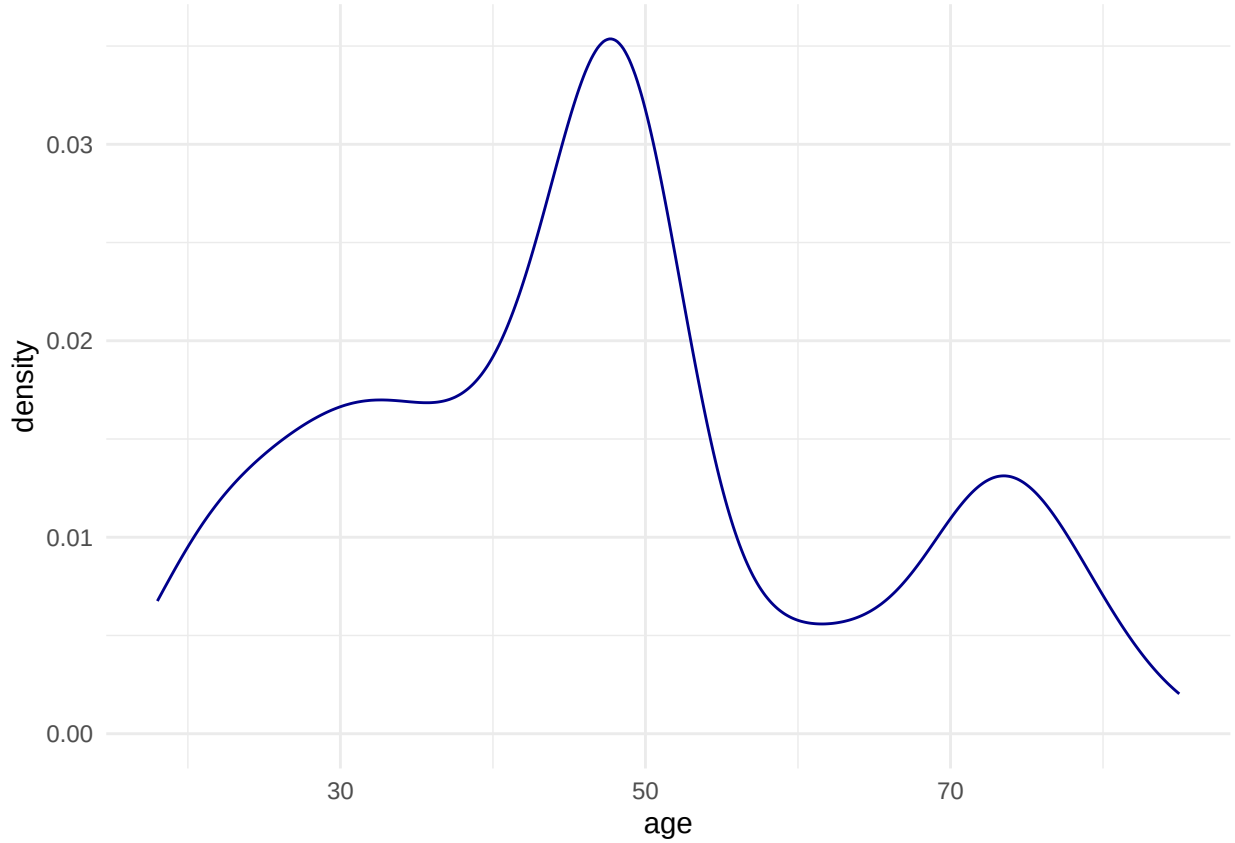


Student ID: 112077423

Question 9b

Visualize the raw data in form of density plot.

```
ggplot(data=customers_df, aes(x=age)) +
  geom_density(color="darkblue") +
  theme_minimal()
```



Student ID: 112077423

Question 9c

Visualize the raw data in form of boxplot+stripchart.

```
ggplot(data=customers_df, aes(x='', y = age)) +
  geom_boxplot() +
  geom_jitter(position=position_jitter(0.2), color='darkblue', shape=17) +
  theme_minimal() +
  labs(x="")
```