

**Program:**

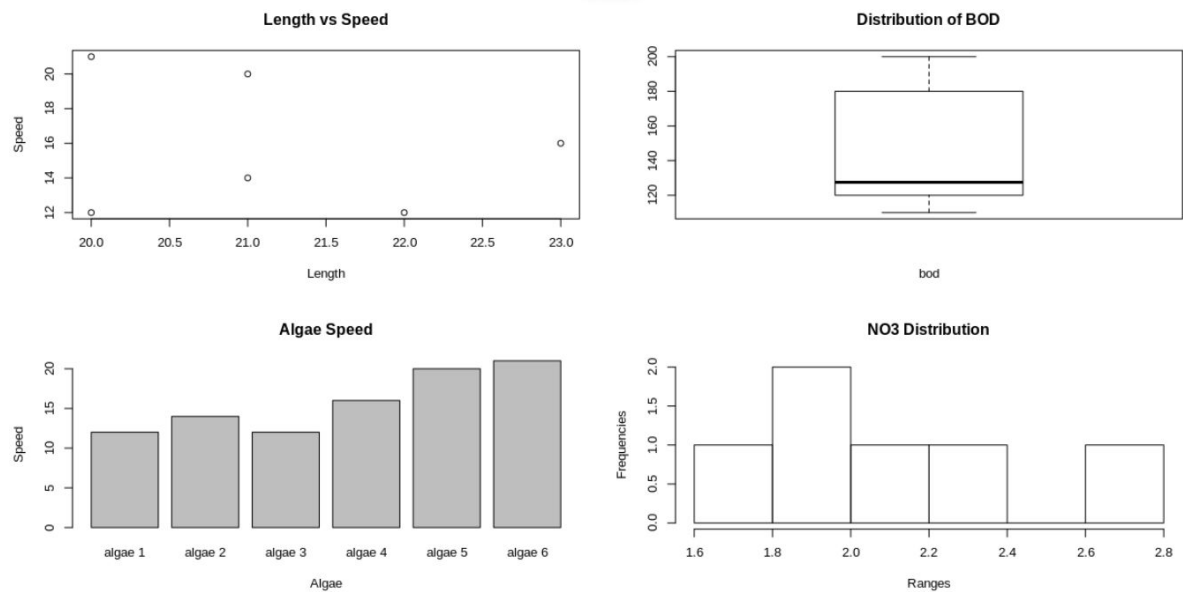
```
# question 5a
length = c(20,21,22,23,21,20)
speed = c(12,14,12,16,20,21)
algae = c(40,45,45,80,75,65)
no3 = c(2.25,2.15,1.75,1.95,1.95,2.75)
bod = c(200,180,135,120,110,120)
regions = par(mfrow=c(2,2))
plot(x=length, y=speed, xlab="Length", ylab="Speed", main="Length vs Speed")
boxplot(bod, xlab="bod", main="Distribution of BOD")
barplot(speed, names.arg=c("algae 1", "algae 2", "algae 3", "algae 4", "algae 5",
"algae 6"), xlab="Algae", ylab="Speed", main="Algae Speed")
hist(no3, xlab="Ranges", ylab="Frequencies", main="NO3 Distribution")

# question 5b
age = c("youth", "youth", "middleaged", "senior", "senior", "senior", "middleaged",
"youth", "youth", "senior", "youth", "middleaged", "middleaged", "senior")
income = c("high", "high", "high", "medium", "low", "low", "low", "medium", "low",
"medium", "medium", "medium", "high", "medium")
student = c("no", "no", "no", "no", "yes", "yes", "yes", "no", "yes", "yes", "yes", "no",
"yes", "no")
credit_ranking = c("fair", "excellent", "fair", "fair", "fair", "excellent", "excellent", "fair",
"fair", "fair", "excellent", "excellent", "fair", "excellent")
class_buys = c("no", "no", "yes", "yes", "yes", "no", "yes", "no", "yes", "yes", "yes",
"yes", "yes", "no")

train_df = data.frame(age, income, student, credit_ranking, class_buys)
library("e1071")
model = naiveBayes(class_buys~., data=train_df)
model

test_df = data.frame(age=c("senior"), income=c("high"), student=c("no"),
credit_ranking=c("fair"))
predict(model, test_df)
```

## Outputs:



Output 5a

```
no      yes
0.3571429 0.6428571
```

Conditional probabilities:

```
age
Y    middleaged  senior  youth
no   0.0000000  0.4000000  0.6000000
yes  0.4444444  0.3333333  0.2222222
```

```
income
Y    high  low  medium
no   0.4000000  0.2000000  0.4000000
yes  0.2222222  0.3333333  0.4444444
```

```
student
Y    no  yes
no   0.8000000  0.2000000
yes  0.3333333  0.6666667
```

```
credit_ranking
Y    excellent  fair
no   0.6000000  0.4000000
yes  0.3333333  0.6666667
```

```
> test_df = data.frame(age=c("senior"), income=c("high"), student=c("no"), credit_ranking=c("fair"))
> predict(model, test_df)
[1] no
Levels: no yes
```

Output 5b