**Output:-**

|  |
| --- |
| > summary(Zoo)  animal.name hair feathers eggs milk airborne aquatic  Length:101 Min. :0.0000 Min. :0.000 Min. :0.0000 Min. :0.0000 Min. :0.0000 Min. :0.0000  Class :character 1st Qu.:0.0000 1st Qu.:0.000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000  Mode :character Median :0.0000 Median :0.000 Median :1.0000 Median :0.0000 Median :0.0000 Median :0.0000  Mean :0.4257 Mean :0.198 Mean :0.5842 Mean :0.4059 Mean :0.2376 Mean :0.3564  3rd Qu.:1.0000 3rd Qu.:0.000 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:0.0000 3rd Qu.:1.0000  Max. :1.0000 Max. :1.000 Max. :1.0000 Max. :1.0000 Max. :1.0000 Max. :1.0000  predator toothed backbone breathes venomous fins legs  Min. :0.0000 Min. :0.000 Min. :0.0000 Min. :0.0000 Min. :0.00000 Min. :0.0000 Min. :0.000  1st Qu.:0.0000 1st Qu.:0.000 1st Qu.:1.0000 1st Qu.:1.0000 1st Qu.:0.00000 1st Qu.:0.0000 1st Qu.:2.000  Median :1.0000 Median :1.000 Median :1.0000 Median :1.0000 Median :0.00000 Median :0.0000 Median :4.000  Mean :0.5545 Mean :0.604 Mean :0.8218 Mean :0.7921 Mean :0.07921 Mean :0.1683 Mean :2.842  3rd Qu.:1.0000 3rd Qu.:1.000 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:0.00000 3rd Qu.:0.0000 3rd Qu.:4.000  Max. :1.0000 Max. :1.000 Max. :1.0000 Max. :1.0000 Max. :1.00000 Max. :1.0000 Max. :8.000  tail domestic catsize type  Min. :0.0000 Min. :0.0000 Min. :0.0000 Min. :1.000  1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:1.000  Median :1.0000 Median :0.0000 Median :0.0000 Median :2.000  Mean :0.7426 Mean :0.1287 Mean :0.4356 Mean :2.832  3rd Qu.:1.0000 3rd Qu.:0.0000 3rd Qu.:1.0000 3rd Qu.:4.000  Max. :1.0000 Max. :1.0000 Max. :1.0000 Max. :7.000 |
|  |

> table(Zoo$type)

1 2 3 4 5 6 7

41 20 5 13 4 8 10

> str(Zoo)

'data.frame': 101 obs. of 18 variables:

$ animal.name: chr "aardvark" "antelope" "bass" "bear" ...

$ hair : int 1 1 0 1 1 1 1 0 0 1 ...

$ feathers : int 0 0 0 0 0 0 0 0 0 0 ...

$ eggs : int 0 0 1 0 0 0 0 1 1 0 ...

$ milk : int 1 1 0 1 1 1 1 0 0 1 ...

$ airborne : int 0 0 0 0 0 0 0 0 0 0 ...

$ aquatic : int 0 0 1 0 0 0 0 1 1 0 ...

$ predator : int 1 0 1 1 1 0 0 0 1 0 ...

$ toothed : int 1 1 1 1 1 1 1 1 1 1 ...

$ backbone : int 1 1 1 1 1 1 1 1 1 1 ...

$ breathes : int 1 1 0 1 1 1 1 0 0 1 ...

$ venomous : int 0 0 0 0 0 0 0 0 0 0 ...

$ fins : int 0 0 1 0 0 0 0 1 1 0 ...

$ legs : int 4 4 0 4 4 4 4 0 0 4 ...

$ tail : int 0 1 1 0 1 1 1 1 1 0 ...

$ domestic : int 0 0 0 0 0 0 1 1 0 1 ...

$ catsize : int 1 1 0 1 1 1 1 0 0 0 ...

$ type : int 1 1 4 1 1 1 1 4 4 1 ...

|  |
| --- |
| Zoo$type <- factor(Zoo$type, levels = c('1', '2','3','4','5', '6','7'), labels=c('category1','category2', 'category3','category4', 'category5', 'category6', 'category7'))  > str(Zoo)  'data.frame': 101 obs. of 18 variables:  $ animal.name: chr "aardvark" "antelope" "bass" "bear" ...  $ hair : int 1 1 0 1 1 1 1 0 0 1 ...  $ feathers : int 0 0 0 0 0 0 0 0 0 0 ...  $ eggs : int 0 0 1 0 0 0 0 1 1 0 ...  $ milk : int 1 1 0 1 1 1 1 0 0 1 ...  $ airborne : int 0 0 0 0 0 0 0 0 0 0 ...  $ aquatic : int 0 0 1 0 0 0 0 1 1 0 ...  $ predator : int 1 0 1 1 1 0 0 0 1 0 ...  $ toothed : int 1 1 1 1 1 1 1 1 1 1 ...  $ backbone : int 1 1 1 1 1 1 1 1 1 1 ...  $ breathes : int 1 1 0 1 1 1 1 0 0 1 ...  $ venomous : int 0 0 0 0 0 0 0 0 0 0 ...  $ fins : int 0 0 1 0 0 0 0 1 1 0 ...  $ legs : int 4 4 0 4 4 4 4 0 0 4 ...  $ tail : int 0 1 1 0 1 1 1 1 1 0 ...  $ domestic : int 0 0 0 0 0 0 1 1 0 1 ...  $ catsize : int 1 1 0 1 1 1 1 0 0 0 ...  $ type : Factor w/ 7 levels "category1","category2",..: 1 1 4 1 1 1 1 4 4 1 ... |
|  |
| |  |  |  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | --- | --- | --- | | |  | | --- | | > norm(c(1,2,3,4,5))  [1] -1.0 -0.8 -0.6 -0.4 -0.2  > summary(Zoo)  hair feathers eggs milk airborne aquatic predator  Min. :0.0000 Min. :0.000 Min. :0.0000 Min. :0.0000 Min. :0.0000 Min. :0.0000 Min. :0.0000  1st Qu.:0.0000 1st Qu.:0.000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000 1st Qu.:0.0000  Median :0.0000 Median :0.000 Median :1.0000 Median :0.0000 Median :0.0000 Median :0.0000 Median :1.0000  Mean :0.4257 Mean :0.198 Mean :0.5842 Mean :0.4059 Mean :0.2376 Mean :0.3564 Mean :0.5545  3rd Qu.:1.0000 3rd Qu.:0.000 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:0.0000 3rd Qu.:1.0000 3rd Qu.:1.0000  Max. :1.0000 Max. :1.000 Max. :1.0000 Max. :1.0000 Max. :1.0000 Max. :1.0000 Max. :1.0000    toothed backbone breathes venomous fins legs tail  Min. :0.000 Min. :0.0000 Min. :0.0000 Min. :0.00000 Min. :0.0000 Min. :0.000 Min. :0.0000  1st Qu.:0.000 1st Qu.:1.0000 1st Qu.:1.0000 1st Qu.:0.00000 1st Qu.:0.0000 1st Qu.:2.000 1st Qu.:0.0000  Median :1.000 Median :1.0000 Median :1.0000 Median :0.00000 Median :0.0000 Median :4.000 Median :1.0000  Mean :0.604 Mean :0.8218 Mean :0.7921 Mean :0.07921 Mean :0.1683 Mean :2.842 Mean :0.7426  3rd Qu.:1.000 3rd Qu.:1.0000 3rd Qu.:1.0000 3rd Qu.:0.00000 3rd Qu.:0.0000 3rd Qu.:4.000 3rd Qu.:1.0000  Max. :1.000 Max. :1.0000 Max. :1.0000 Max. :1.00000 Max. :1.0000 Max. :8.000 Max. :1.0000    domestic catsize type  Min. :0.0000 Min. :0.0000 category1:41  1st Qu.:0.0000 1st Qu.:0.0000 category2:20  Median :0.0000 Median :0.0000 category3: 5  Mean :0.1287 Mean :0.4356 category4:13  3rd Qu.:0.0000 3rd Qu.:1.0000 category5: 4  Max. :1.0000 Max. :1.0000 category6: 8  category7:10 | |  | | |  |  |  |  |  | | --- | --- | --- | --- | --- | | #building a model  > KnnModel <- knn(Train[1:16], Test[1:16], Train$type, k=3)  > mean(KnnModel==Test$type)  [1] 0.9642857  CrossTable(KnnModel,Test$type)    Cell Contents  |-------------------------|  | N |  | Chi-square contribution |  | N / Row Total |  | N / Col Total |  | N / Table Total |  |-------------------------|    Total Observations in Table: 28    | Test$type  KnnModel | category1 | category2 | category3 | category4 | category5 | category6 | category7 | Row Total |  -------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|  category1 | 12 | 0 | 0 | 0 | 0 | 0 | 0 | 12 |  | 9.143 | 2.571 | 0.429 | 1.286 | 0.429 | 0.857 | 1.286 | |  | 1.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.429 |  | 1.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | |  | 0.429 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | |  -------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|  category2 | 0 | 6 | 0 | 0 | 0 | 0 | 0 | 6 |  | 2.571 | 17.286 | 0.214 | 0.643 | 0.214 | 0.429 | 0.643 | |  | 0.000 | 1.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.214 |  | 0.000 | 1.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | |  | 0.000 | 0.214 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | |  -------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|  category3 | 0 | 0 | 1 | 0 | 0 | 0 | 0 | 1 |  | 0.429 | 0.214 | 26.036 | 0.107 | 0.036 | 0.071 | 0.107 | |  | 0.000 | 0.000 | 1.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.036 |  | 0.000 | 0.000 | 1.000 | 0.000 | 0.000 | 0.000 | 0.000 | |  | 0.000 | 0.000 | 0.036 | 0.000 | 0.000 | 0.000 | 0.000 | |  -------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|  category4 | 0 | 0 | 0 | 3 | 0 | 0 | 0 | 3 |  | 1.286 | 0.643 | 0.107 | 22.321 | 0.107 | 0.214 | 0.321 | |  | 0.000 | 0.000 | 0.000 | 1.000 | 0.000 | 0.000 | 0.000 | 0.107 |  | 0.000 | 0.000 | 0.000 | 1.000 | 0.000 | 0.000 | 0.000 | |  | 0.000 | 0.000 | 0.000 | 0.107 | 0.000 | 0.000 | 0.000 | |  -------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|  category5 | 0 | 0 | 0 | 0 | 1 | 0 | 0 | 1 |  | 0.429 | 0.214 | 0.036 | 0.107 | 26.036 | 0.071 | 0.107 | |  | 0.000 | 0.000 | 0.000 | 0.000 | 1.000 | 0.000 | 0.000 | 0.036 |  | 0.000 | 0.000 | 0.000 | 0.000 | 1.000 | 0.000 | 0.000 | |  | 0.000 | 0.000 | 0.000 | 0.000 | 0.036 | 0.000 | 0.000 | |  -------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|  category6 | 0 | 0 | 0 | 0 | 0 | 2 | 1 | 3 |  | 1.286 | 0.643 | 0.107 | 0.321 | 0.107 | 14.881 | 1.433 | |  | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.667 | 0.333 | 0.107 |  | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 1.000 | 0.333 | |  | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.071 | 0.036 | |  -------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|  category7 | 0 | 0 | 0 | 0 | 0 | 0 | 2 | 2 |  | 0.857 | 0.429 | 0.071 | 0.214 | 0.071 | 0.143 | 14.881 | |  | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 1.000 | 0.071 |  | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.667 | |  | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.000 | 0.071 | |  -------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|  Column Total | 12 | 6 | 1 | 3 | 1 | 2 | 3 | 28 |  | 0.429 | 0.214 | 0.036 | 0.107 | 0.036 | 0.071 | 0.107 | |  -------------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|-----------|   |  | | --- | | confusionMatrix(KnnModel,Test$type)  Confusion Matrix and Statistics  Reference  Prediction category1 category2 category3 category4 category5 category6 category7  category1 12 0 0 0 0 0 0  category2 0 6 0 0 0 0 0  category3 0 0 1 0 0 0 0  category4 0 0 0 3 0 0 0  category5 0 0 0 0 1 0 0  category6 0 0 0 0 0 2 1  category7 0 0 0 0 0 0 2  Overall Statistics    Accuracy : 0.9643  95% CI : (0.8165, 0.9991)  No Information Rate : 0.4286  P-Value [Acc > NIR] : 1.906e-09    Kappa : 0.9518    Mcnemar's Test P-Value : NA  Statistics by Class:  Class: category1 Class: category2 Class: category3 Class: category4 Class: category5  Sensitivity 1.0000 1.0000 1.00000 1.0000 1.00000  Specificity 1.0000 1.0000 1.00000 1.0000 1.00000  Pos Pred Value 1.0000 1.0000 1.00000 1.0000 1.00000  Neg Pred Value 1.0000 1.0000 1.00000 1.0000 1.00000  Prevalence 0.4286 0.2143 0.03571 0.1071 0.03571  Detection Rate 0.4286 0.2143 0.03571 0.1071 0.03571  Detection Prevalence 0.4286 0.2143 0.03571 0.1071 0.03571  Balanced Accuracy 1.0000 1.0000 1.00000 1.0000 1.00000  Class: category6 Class: category7  Sensitivity 1.00000 0.66667  Specificity 0.96154 1.00000  Pos Pred Value 0.66667 1.00000  Neg Pred Value 1.00000 0.96154  Prevalence 0.07143 0.10714  Detection Rate 0.07143 0.07143  Detection Prevalence 0.10714 0.07143  Balanced Accuracy 0.98077 0.83333 | |  | | |  | | --- | | > > KNNFInal\_model <- knn(Train[1:16],Test[1:16],Train$type,k=5)  > mean(KNNFInal\_model==Test$type)  [1] 0.8571429 | | | | | |