Name: Vikram Surya R

Reg No: 19BCE1490

Essentials of Data Analytics

Tasks for Week-8: Hierarchical Clustering

Understand the following operations/functions on 'USArrests' data and perform similar operations on 'iris' dataset based on given instructions.

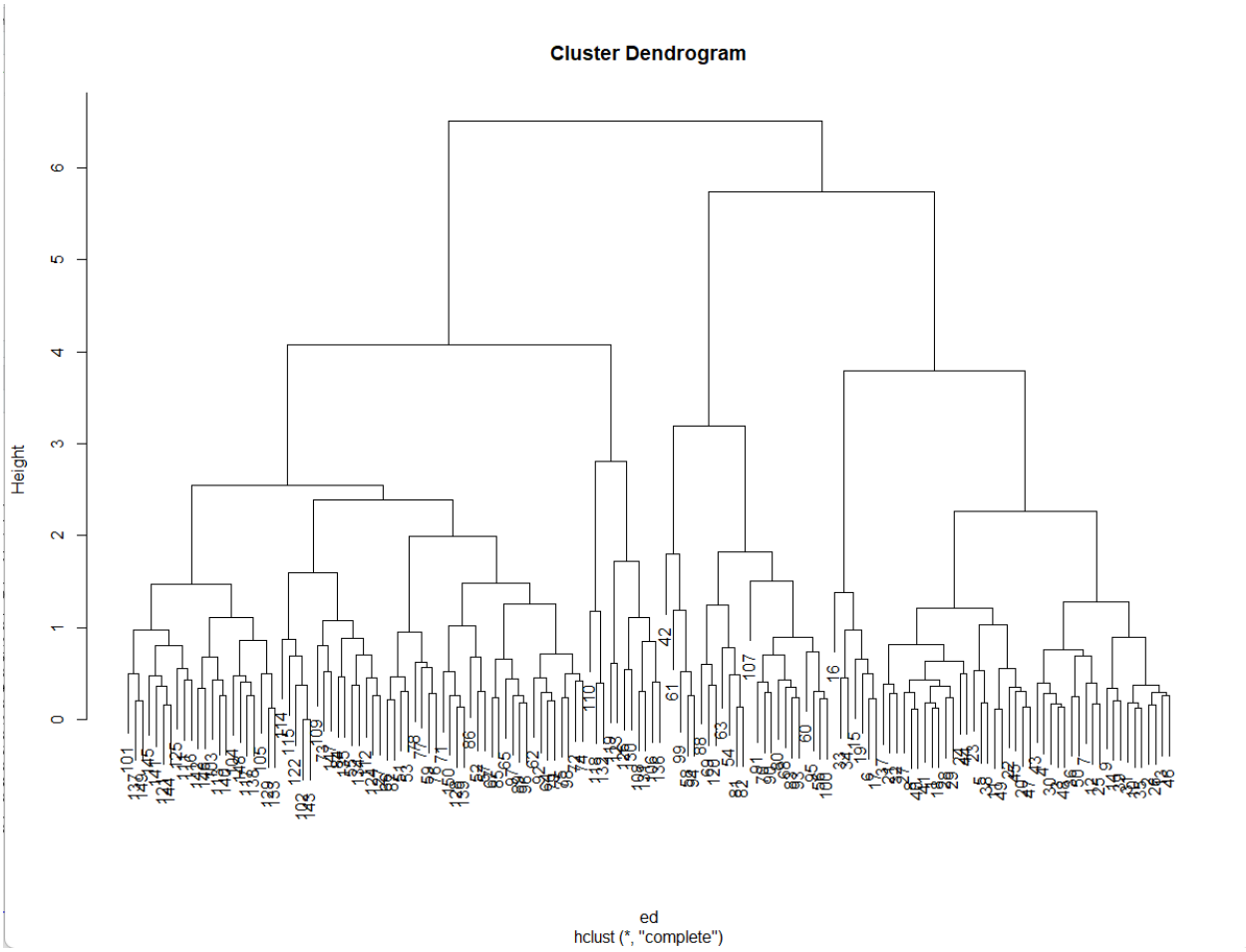
AIM:

To develop a hierarchical clustering model - a dendogram for the given data using R programming and perform operations on the data "iris".

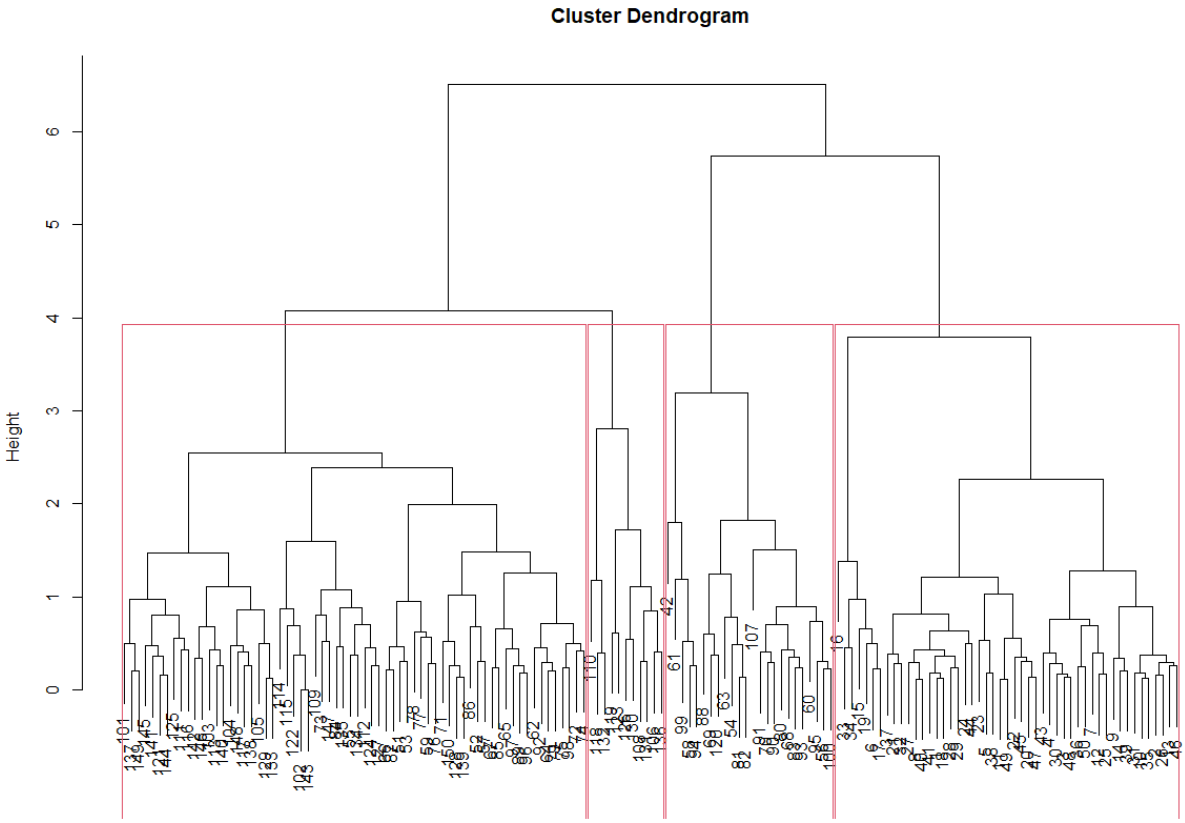
Algorithm:

- Set the working directory to the dataset where we store by using setwd() and read data.
- By using scale function. We scale the data and store it in another variable.
- Find the Euclidean distance using the dist() function and method="euclidean".
- Find the hierarchical clustering using hclust and method = "complete.
- For plotting the dendrogram we use plot function.
- Grouping the dataset into number of cluster using cutree function and plotting them.
- Use rec.hclust to find the cluster.

RESULTS:




```
> cluster
  1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17  18  19  20  21  22  23  24  25  26  27  28  29  30
  1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1   1
 31  32  33  34  35  36  37  38  39  40  41  42  43  44  45  46  47  48  49  50  51  52  53  54  55  56  57  58  59  60
  1   1   1   1   1   1   1   1   1   1   1   2   1   1   1   1   1   1   1   1   3   3   3   2   3   2   3   2   3   2
 61  62  63  64  65  66  67  68  69  70  71  72  73  74  75  76  77  78  79  80  81  82  83  84  85  86  87  88  89  90
  2   3   2   3   3   3   3   2   2   2   3   3   3   3   3   3   3   3   3   2   2   2   2   3   3   3   3   2   3   2
 91  92  93  94  95  96  97  98  99 100 101 102 103 104 105 106 107 108 109 110 111 112 113 114 115 116 117 118 119 120
  2   3   2   2   2   3   3   3   2   2   3   3   3   3   3   4   2   4   3   4   3   3   3   3   3   3   3   4   4   2
121 122 123 124 125 126 127 128 129 130 131 132 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150
  3   3   4   3   3   4   3   3   3   4   4   4   3   3   3   4   3   3   3   3   3   3   3   3   3   3   3   3   3   3
> rect.hclust(hercluster,k=4)
```



CODE:

```
rm(list=ls())
setwd("C:/Users/VIKRAM SURYA/Desktop/EDA_LAB")
data <- read.csv("iris.csv",row.names=1)
View(data)
df<-scale(data)
ed<-dist(df,method="euclidean")
hercluster<-hclust(ed,method='complete')
```

```
plot(hercluster)
cluster<-cutree(hercluster,k=4)
cluster
rect.hclust(hercluster,k=4)
```