Project 4, Bus. 744, Unsupervised Learning

install.packages(“tidyverse”) library(readr) library(tidyvere) library(cluster) library(datasets) cats <- read.csv(“catsM.csv”,na.strings = "", head = TRUE)

avW <- agnes(cats[,1:2], method =“ward”, stand = TRUE ) h2 <- cutree(as.hclust(avW), k = 2) # Cut into 2 groups avW <- agnes(cats[,1:2], method =“ward”, stand = TRUE ) h3 <- cutree(as.hclust(avW), k = 3) # Cut into 3 groups avW <- agnes(cats[,1:2], method =“ward”, stand = TRUE ) h4 <- cutree(as.hclust(avW), k = 4) # Cut into 4 groups avW <- agnes(cats[,1:2], method =“ward”, stand = TRUE ) h5 <- cutree(as.hclust(avW), k = 5) # Cut into 5 groups avW <- agnes(cats[,1:2], method =“ward”, stand = TRUE ) h6 <- cutree(as.hclust(avW), k = 6) # Cut into 6 groups avW <- agnes(cats[,1:2], method =“ward”, stand = TRUE ) h7 <- cutree(as.hclust(avW), k = 7) # Cut into 7 groups

siH2 <- silhouette (h2, daisy(cats[,1:2])) plot (siH2 , main =“Silhouette Plot Hierarchical”, col =" blue “) siH3 <- silhouette (h3, daisy(cats[,1:2])) plot (siH3 , main =”Silhouette Plot Hierarchical“, col =” blue “) siH4 <- silhouette (h4, daisy(cats[,1:2])) plot (siH4 , main =”Silhouette Plot Hierarchical“, col =” blue “) siH5 <- silhouette (h5, daisy(cats[,1:2])) plot (siH5 , main =”Silhouette Plot Hierarchical“, col =” blue “) siH6 <- silhouette (h6, daisy(cats[,1:2])) plot (siH6 , main =”Silhouette Plot Hierarchical“, col =” blue “) siH7 <- silhouette (h7, daisy(cats[,1:2])) plot (siH7 , main =”Silhouette Plot Hierarchical“, col =” blue ")

I believe that the 4 k cluster was the best solution becasue it divided the results most evenly. library(psych) # Hierarchical hierarchDat <- cbind(cats,group=h2) describeBy(hierarchDat, hierarchDatgroup) hierarchDat <- cbind(cats,group=h4) describeBy(hierarchDat, hierarchDatgroup) hierarchDat <- cbind(cats,group=h6) describeBy(hierarchDat, hierarchDatgroup)

Descriptive statistics by group

group: 1

vars n mean sd median trimmed mad min max range skew

X 1 52 26.50 15.15 26.5 26.50 19.27 1.0 52.0 51.0 0.00

Sex\* 2 52 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 52 2.54 0.25 2.6 2.55 0.30 2.0 2.9 0.9 -0.40

Hwt 4 52 9.72 1.71 9.6 9.66 1.63 6.5 13.5 7.0 0.28

group 5 52 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

kurtosis se

X -1.27 2.10

Sex\* NaN 0.00

Bwt -0.91 0.03

Hwt -0.58 0.24

group NaN 0.00

------------------------------------------------------

group: 2

vars n mean sd median trimmed mad min max range skew

X 1 45 75.00 13.13 75.0 75.00 16.31 53.0 97.0 44.0 0.00

Sex\* 2 45 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 45 3.32 0.26 3.3 3.29 0.30 3.0 3.9 0.9 0.52

Hwt 4 45 13.17 2.05 12.8 13.01 1.78 9.9 20.5 10.6 1.09

group 5 45 2.00 0.00 2.0 2.00 0.00 2.0 2.0 0.0 NaN

kurtosis se

X -1.28 1.96

Sex\* NaN 0.00

Bwt -0.76 0.04

Hwt 1.86 0.31

group NaN 0.00

> hierarchDat <- cbind(cats,group=h3)

> describeBy(hierarchDat, hierarchDat$group)

Descriptive statistics by group

group: 1

vars n mean sd median trimmed mad min max range skew

X 1 31 16.00 9.09 16.0 16.00 11.86 1.0 31.0 30.0 0.00

Sex\* 2 31 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 31 2.37 0.19 2.4 2.39 0.30 2.0 2.6 0.6 -0.42

Hwt 4 31 9.06 1.57 9.1 8.96 1.78 6.5 12.7 6.2 0.57

group 5 31 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

kurtosis se

X -1.32 1.63

Sex\* NaN 0.00

Bwt -1.12 0.03

Hwt -0.13 0.28

group NaN 0.00

------------------------------------------------------

group: 2

vars n mean sd median trimmed mad min max range skew

X 1 21 42.00 6.20 42.0 42.00 7.41 32.0 52.0 20.0 0.00

Sex\* 2 21 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 21 2.78 0.08 2.8 2.78 0.15 2.7 2.9 0.2 0.33

Hwt 4 21 10.70 1.45 10.4 10.64 1.48 8.0 13.5 5.5 0.26

group 5 21 2.00 0.00 2.0 2.00 0.00 2.0 2.0 0.0 NaN

kurtosis se

X -1.37 1.35

Sex\* NaN 0.00

Bwt -1.49 0.02

Hwt -0.87 0.32

group NaN 0.00

------------------------------------------------------

group: 3

vars n mean sd median trimmed mad min max range skew

X 1 45 75.00 13.13 75.0 75.00 16.31 53.0 97.0 44.0 0.00

Sex\* 2 45 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 45 3.32 0.26 3.3 3.29 0.30 3.0 3.9 0.9 0.52

Hwt 4 45 13.17 2.05 12.8 13.01 1.78 9.9 20.5 10.6 1.09

group 5 45 3.00 0.00 3.0 3.00 0.00 3.0 3.0 0.0 NaN

kurtosis se

X -1.28 1.96

Sex\* NaN 0.00

Bwt -0.76 0.04

Hwt 1.86 0.31

group NaN 0.00

> hierarchDat <- cbind(cats,group=h4)

> describeBy(hierarchDat, hierarchDat$group)

Descriptive statistics by group

group: 1

vars n mean sd median trimmed mad min max range skew

X 1 31 16.00 9.09 16.0 16.00 11.86 1.0 31.0 30.0 0.00

Sex\* 2 31 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 31 2.37 0.19 2.4 2.39 0.30 2.0 2.6 0.6 -0.42

Hwt 4 31 9.06 1.57 9.1 8.96 1.78 6.5 12.7 6.2 0.57

group 5 31 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

kurtosis se

X -1.32 1.63

Sex\* NaN 0.00

Bwt -1.12 0.03

Hwt -0.13 0.28

group NaN 0.00

------------------------------------------------------

group: 2

vars n mean sd median trimmed mad min max range skew

X 1 21 42.00 6.20 42.0 42.00 7.41 32.0 52.0 20.0 0.00

Sex\* 2 21 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 21 2.78 0.08 2.8 2.78 0.15 2.7 2.9 0.2 0.33

Hwt 4 21 10.70 1.45 10.4 10.64 1.48 8.0 13.5 5.5 0.26

group 5 21 2.00 0.00 2.0 2.00 0.00 2.0 2.0 0.0 NaN

kurtosis se

X -1.37 1.35

Sex\* NaN 0.00

Bwt -1.49 0.02

Hwt -0.87 0.32

group NaN 0.00

------------------------------------------------------

group: 3

vars n mean sd median trimmed mad min max range skew

X 1 24 64.50 7.07 64.50 64.50 8.90 53.0 76.0 23.0 0.00

Sex\* 2 24 1.00 0.00 1.00 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 24 3.11 0.11 3.10 3.10 0.15 3.0 3.3 0.3 0.37

Hwt 4 24 12.24 1.23 12.25 12.28 1.11 9.9 14.3 4.4 -0.24

group 5 24 3.00 0.00 3.00 3.00 0.00 3.0 3.0 0.0 NaN

kurtosis se

X -1.35 1.44

Sex\* NaN 0.00

Bwt -1.29 0.02

Hwt -0.84 0.25

group NaN 0.00

------------------------------------------------------

group: 4

vars n mean sd median trimmed mad min max range skew

X 1 21 87.00 6.20 87.0 87.00 7.41 77.0 97.0 20.0 0.00

Sex\* 2 21 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 21 3.55 0.18 3.5 3.54 0.15 3.3 3.9 0.6 0.55

Hwt 4 21 14.23 2.30 14.4 14.05 2.22 11.0 20.5 9.5 0.73

group 5 21 4.00 0.00 4.0 4.00 0.00 4.0 4.0 0.0 NaN

kurtosis se

X -1.37 1.35

Sex\* NaN 0.00

Bwt -0.85 0.04

Hwt 0.37 0.50

group NaN 0.00

> hierarchDat <- cbind(cats,group=h5)

> describeBy(hierarchDat, hierarchDat$group)

Descriptive statistics by group

group: 1

vars n mean sd median trimmed mad min max range skew

X 1 13 7.00 3.89 7.0 7.00 4.45 1.0 13.0 12.0 0.00

Sex\* 2 13 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 13 2.18 0.11 2.2 2.18 0.00 2.0 2.4 0.4 -0.10

Hwt 4 13 8.48 1.40 8.5 8.45 1.63 6.5 10.7 4.2 -0.01

group 5 13 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

kurtosis se

X -1.48 1.08

Sex\* NaN 0.00

Bwt -0.22 0.03

Hwt -1.58 0.39

group NaN 0.00

------------------------------------------------------

group: 2

vars n mean sd median trimmed mad min max range skew

X 1 18 22.50 5.34 22.5 22.50 6.67 14.0 31.0 17.0 0.00

Sex\* 2 18 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 18 2.51 0.08 2.5 2.51 0.15 2.4 2.6 0.2 -0.16

Hwt 4 18 9.49 1.57 9.2 9.40 1.63 7.7 12.7 5.0 0.83

group 5 18 2.00 0.00 2.0 2.00 0.00 2.0 2.0 0.0 NaN

kurtosis se

X -1.40 1.26

Sex\* NaN 0.00

Bwt -1.35 0.02

Hwt -0.58 0.37

group NaN 0.00

------------------------------------------------------

group: 3

vars n mean sd median trimmed mad min max range skew

X 1 21 42.00 6.20 42.0 42.00 7.41 32.0 52.0 20.0 0.00

Sex\* 2 21 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 21 2.78 0.08 2.8 2.78 0.15 2.7 2.9 0.2 0.33

Hwt 4 21 10.70 1.45 10.4 10.64 1.48 8.0 13.5 5.5 0.26

group 5 21 3.00 0.00 3.0 3.00 0.00 3.0 3.0 0.0 NaN

kurtosis se

X -1.37 1.35

Sex\* NaN 0.00

Bwt -1.49 0.02

Hwt -0.87 0.32

group NaN 0.00

------------------------------------------------------

group: 4

vars n mean sd median trimmed mad min max range skew

X 1 24 64.50 7.07 64.50 64.50 8.90 53.0 76.0 23.0 0.00

Sex\* 2 24 1.00 0.00 1.00 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 24 3.11 0.11 3.10 3.10 0.15 3.0 3.3 0.3 0.37

Hwt 4 24 12.24 1.23 12.25 12.28 1.11 9.9 14.3 4.4 -0.24

group 5 24 4.00 0.00 4.00 4.00 0.00 4.0 4.0 0.0 NaN

kurtosis se

X -1.35 1.44

Sex\* NaN 0.00

Bwt -1.29 0.02

Hwt -0.84 0.25

group NaN 0.00

------------------------------------------------------

group: 5

vars n mean sd median trimmed mad min max range skew

X 1 21 87.00 6.20 87.0 87.00 7.41 77.0 97.0 20.0 0.00

Sex\* 2 21 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 21 3.55 0.18 3.5 3.54 0.15 3.3 3.9 0.6 0.55

Hwt 4 21 14.23 2.30 14.4 14.05 2.22 11.0 20.5 9.5 0.73

group 5 21 5.00 0.00 5.0 5.00 0.00 5.0 5.0 0.0 NaN

kurtosis se

X -1.37 1.35

Sex\* NaN 0.00

Bwt -0.85 0.04

Hwt 0.37 0.50

group NaN 0.00

Descriptive statistics by group

group: 1

vars n mean sd median trimmed mad min max range skew

X 1 13 7.00 3.89 7.0 7.00 4.45 1.0 13.0 12.0 0.00

Sex\* 2 13 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 13 2.18 0.11 2.2 2.18 0.00 2.0 2.4 0.4 -0.10

Hwt 4 13 8.48 1.40 8.5 8.45 1.63 6.5 10.7 4.2 -0.01

group 5 13 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

kurtosis se

X -1.48 1.08

Sex\* NaN 0.00

Bwt -0.22 0.03

Hwt -1.58 0.39

group NaN 0.00

------------------------------------------------------

group: 2

vars n mean sd median trimmed mad min max range skew

X 1 18 22.50 5.34 22.5 22.50 6.67 14.0 31.0 17.0 0.00

Sex\* 2 18 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 18 2.51 0.08 2.5 2.51 0.15 2.4 2.6 0.2 -0.16

Hwt 4 18 9.49 1.57 9.2 9.40 1.63 7.7 12.7 5.0 0.83

group 5 18 2.00 0.00 2.0 2.00 0.00 2.0 2.0 0.0 NaN

kurtosis se

X -1.40 1.26

Sex\* NaN 0.00

Bwt -1.35 0.02

Hwt -0.58 0.37

group NaN 0.00

------------------------------------------------------

group: 3

vars n mean sd median trimmed mad min max range skew

X 1 21 42.00 6.20 42.0 42.00 7.41 32.0 52.0 20.0 0.00

Sex\* 2 21 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 21 2.78 0.08 2.8 2.78 0.15 2.7 2.9 0.2 0.33

Hwt 4 21 10.70 1.45 10.4 10.64 1.48 8.0 13.5 5.5 0.26

group 5 21 3.00 0.00 3.0 3.00 0.00 3.0 3.0 0.0 NaN

kurtosis se

X -1.37 1.35

Sex\* NaN 0.00

Bwt -1.49 0.02

Hwt -0.87 0.32

group NaN 0.00

------------------------------------------------------

group: 4

vars n mean sd median trimmed mad min max range skew

X 1 10 57.50 3.03 57.5 57.50 3.71 53.0 62.0 9.0 0.00

Sex\* 2 10 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 10 3.01 0.03 3.0 3.00 0.00 3.0 3.1 0.1 2.28

Hwt 4 10 11.69 1.40 11.9 11.65 2.00 9.9 13.8 3.9 0.04

group 5 10 4.00 0.00 4.0 4.00 0.00 4.0 4.0 0.0 NaN

kurtosis se

X -1.56 0.96

Sex\* NaN 0.00

Bwt 3.57 0.01

Hwt -1.70 0.44

group NaN 0.00

------------------------------------------------------

group: 5

vars n mean sd median trimmed mad min max range skew

X 1 14 69.50 4.18 69.5 69.50 5.19 63.0 76.0 13.0 0.00

Sex\* 2 14 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 14 3.19 0.08 3.2 3.18 0.15 3.1 3.3 0.2 0.21

Hwt 4 14 12.64 0.96 12.4 12.59 1.04 11.5 14.3 2.8 0.37

group 5 14 5.00 0.00 5.0 5.00 0.00 5.0 5.0 0.0 NaN

kurtosis se

X -1.46 1.12

Sex\* NaN 0.00

Bwt -1.41 0.02

Hwt -1.39 0.26

group NaN 0.00

------------------------------------------------------

group: 6

vars n mean sd median trimmed mad min max range skew

X 1 21 87.00 6.20 87.0 87.00 7.41 77.0 97.0 20.0 0.00

Sex\* 2 21 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 21 3.55 0.18 3.5 3.54 0.15 3.3 3.9 0.6 0.55

Hwt 4 21 14.23 2.30 14.4 14.05 2.22 11.0 20.5 9.5 0.73

group 5 21 6.00 0.00 6.0 6.00 0.00 6.0 6.0 0.0 NaN

kurtosis se

X -1.37 1.35

Sex\* NaN 0.00

Bwt -0.85 0.04

Hwt 0.37 0.50

group NaN 0.00

> hierarchDat <- cbind(cats,group=h7)

> describeBy(hierarchDat, hierarchDat$group)

Descriptive statistics by group

group: 1

vars n mean sd median trimmed mad min max range skew

X 1 13 7.00 3.89 7.0 7.00 4.45 1.0 13.0 12.0 0.00

Sex\* 2 13 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 13 2.18 0.11 2.2 2.18 0.00 2.0 2.4 0.4 -0.10

Hwt 4 13 8.48 1.40 8.5 8.45 1.63 6.5 10.7 4.2 -0.01

group 5 13 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

kurtosis se

X -1.48 1.08

Sex\* NaN 0.00

Bwt -0.22 0.03

Hwt -1.58 0.39

group NaN 0.00

------------------------------------------------------

group: 2

vars n mean sd median trimmed mad min max range skew

X 1 18 22.50 5.34 22.5 22.50 6.67 14.0 31.0 17.0 0.00

Sex\* 2 18 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 18 2.51 0.08 2.5 2.51 0.15 2.4 2.6 0.2 -0.16

Hwt 4 18 9.49 1.57 9.2 9.40 1.63 7.7 12.7 5.0 0.83

group 5 18 2.00 0.00 2.0 2.00 0.00 2.0 2.0 0.0 NaN

kurtosis se

X -1.40 1.26

Sex\* NaN 0.00

Bwt -1.35 0.02

Hwt -0.58 0.37

group NaN 0.00

------------------------------------------------------

group: 3

vars n mean sd median trimmed mad min max range skew

X 1 9 36.00 2.74 36.0 36.00 2.97 32.0 40.0 8.0 0.00

Sex\* 2 9 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 9 2.70 0.00 2.7 2.70 0.00 2.7 2.7 0.0 NaN

Hwt 4 9 10.22 1.44 9.8 10.22 1.19 8.0 12.5 4.5 0.18

group 5 9 3.00 0.00 3.0 3.00 0.00 3.0 3.0 0.0 NaN

kurtosis se

X -1.60 0.91

Sex\* NaN 0.00

Bwt NaN 0.00

Hwt -1.32 0.48

group NaN 0.00

------------------------------------------------------

group: 4

vars n mean sd median trimmed mad min max range skew

X 1 12 46.50 3.61 46.50 46.50 4.45 41.0 52.0 11.0 0.00

Sex\* 2 12 1.00 0.00 1.00 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 12 2.84 0.05 2.80 2.84 0.00 2.8 2.9 0.1 0.30

Hwt 4 12 11.06 1.42 10.95 11.01 1.33 9.1 13.5 4.4 0.36

group 5 12 4.00 0.00 4.00 4.00 0.00 4.0 4.0 0.0 NaN

kurtosis se

X -1.50 1.04

Sex\* NaN 0.00

Bwt -2.06 0.01

Hwt -1.22 0.41

group NaN 0.00

------------------------------------------------------

group: 5

vars n mean sd median trimmed mad min max range skew

X 1 10 57.50 3.03 57.5 57.50 3.71 53.0 62.0 9.0 0.00

Sex\* 2 10 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 10 3.01 0.03 3.0 3.00 0.00 3.0 3.1 0.1 2.28

Hwt 4 10 11.69 1.40 11.9 11.65 2.00 9.9 13.8 3.9 0.04

group 5 10 5.00 0.00 5.0 5.00 0.00 5.0 5.0 0.0 NaN

kurtosis se

X -1.56 0.96

Sex\* NaN 0.00

Bwt 3.57 0.01

Hwt -1.70 0.44

group NaN 0.00

------------------------------------------------------

group: 6

vars n mean sd median trimmed mad min max range skew

X 1 14 69.50 4.18 69.5 69.50 5.19 63.0 76.0 13.0 0.00

Sex\* 2 14 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 14 3.19 0.08 3.2 3.18 0.15 3.1 3.3 0.2 0.21

Hwt 4 14 12.64 0.96 12.4 12.59 1.04 11.5 14.3 2.8 0.37

group 5 14 6.00 0.00 6.0 6.00 0.00 6.0 6.0 0.0 NaN

kurtosis se

X -1.46 1.12

Sex\* NaN 0.00

Bwt -1.41 0.02

Hwt -1.39 0.26

group NaN 0.00

------------------------------------------------------

group: 7

vars n mean sd median trimmed mad min max range skew

X 1 21 87.00 6.20 87.0 87.00 7.41 77.0 97.0 20.0 0.00

Sex\* 2 21 1.00 0.00 1.0 1.00 0.00 1.0 1.0 0.0 NaN

Bwt 3 21 3.55 0.18 3.5 3.54 0.15 3.3 3.9 0.6 0.55

Hwt 4 21 14.23 2.30 14.4 14.05 2.22 11.0 20.5 9.5 0.73

group 5 21 7.00 0.00 7.0 7.00 0.00 7.0 7.0 0.0 NaN

kurtosis se

X -1.37 1.35

Sex\* NaN 0.00

Bwt -0.85 0.04

Hwt 0.37 0.50

group NaN 0.00