

Figure

(Left) Percent Cover. (Right) Log of Percent Cover



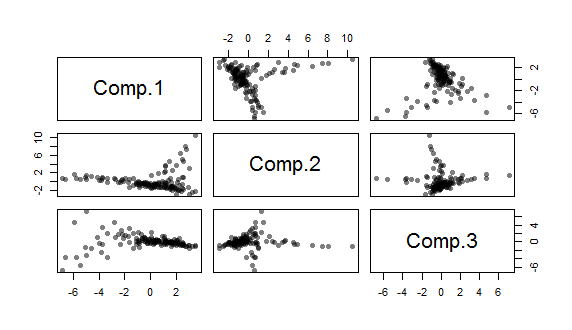
Figure

Data scaled and centered. Variance constant.



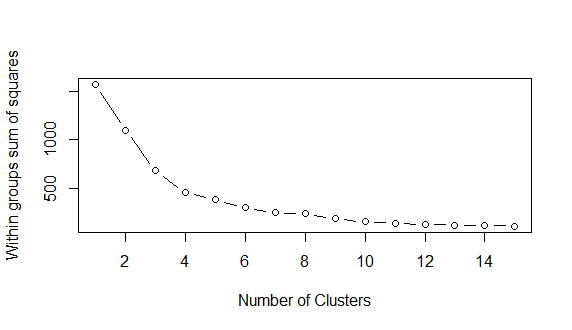
Figure

Visuals of Principal Components using bar and line. Elbow at 3. Does not explain >85% variance(only about 28%, 85% would be 23 principal components).



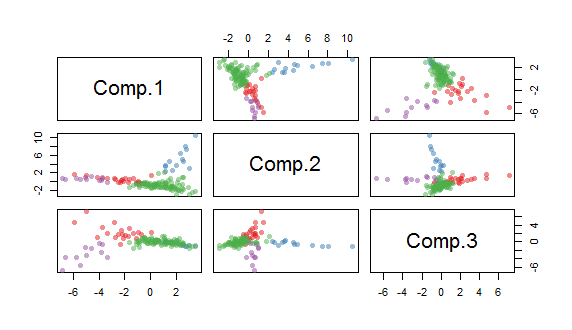
Figure

First 3 principal components. Explains about 28% of variance.



Figure

Determine the number of clusters for K-Means clustering. Elbow at 4 clusters (possibly at 6?).



Figure

K-Means Clustering with 4 clusters. There appear to be four distinct groups. Quite a bit of overlap in the middle. Group 2 has a lot more points than groups 1, 3, & 4

Cluster #: 3 4 1 2

# in Cluster: 11 13 20 92

# First cluster

[1] "B1\_8" "M1\_1" "M1\_2" "M1\_4" "M1\_6" "M2\_10" "M2\_5" "M2\_8"

[9] "M2\_9" "M3\_7" "M4\_5"

# Second Cluster

[1] "1" "2" "10" "12" "14" "15" "17" "18"

[9] "33" "39" "44" "61" "IL4\_8"

# Third Cluster

[1] "11" "67" "68" "B1\_1" "B1\_2" "B1\_3" "B1\_4" "B2\_1"

[9] "B3\_3" "B3\_4" "B3\_5" "B4\_8" "IL1\_3" "IL2\_5" "IL2\_8" "M1\_10"

[17] "M1\_7" "M1\_8" "M2\_6" "M2\_7"

# Fourth Cluster

[1] "16" "19" "20" "21" "23" "24" "32"

[8] "38" "40" "42" "43" "47" "48" "50"

[15] "57" "59" "60" "73" "77" "80" "82"

[22] "90" "B1\_10" "B1\_5" "B1\_6" "B1\_7" "B1\_9" "B2\_10"

[29] "B2\_2" "B2\_4" "B2\_5" "B2\_6" "B2\_7" "B2\_8" "B2\_9"

[36] "B3\_1" "B3\_10" "B3\_2" "B3\_6" "B3\_7" "B3\_8" "B3\_9"

[43] "B4\_1" "B4\_10" "B4\_2" "B4\_3" "B4\_4" "B4\_6" "B4\_7"

[50] "B4\_9" "IL1\_1" "IL1\_10" "IL1\_4" "IL1\_5" "IL1\_8" "IL1\_9"

[57] "IL2\_1" "IL2\_10" "IL2\_2" "IL2\_3" "IL2\_4" "IL2\_7" "IL2\_9"

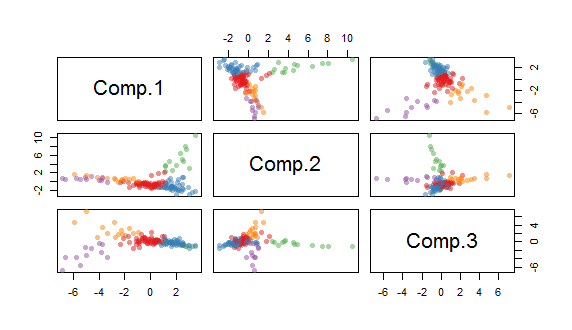
[64] "IL3\_1" "IL3\_10" "IL3\_2" "IL3\_3" "IL3\_4" "IL3\_6" "IL3\_7"

[71] "IL3\_9" "IL4\_10" "IL4\_2" "IL4\_3" "IL4\_4" "IL4\_5" "IL4\_7"

[78] "IL4\_9" "M1\_5" "M2\_1" "M2\_3" "M3\_1" "M3\_2" "M3\_3"

[85] "M3\_5" "M3\_8" "M4\_2" "M4\_3" "M4\_4" "M4\_6" "M4\_7"

[92] "M4\_9"



Figure

K-Means Clustering with 5 clusters. This appears to break up group 2 from above (4 clusters), but doesn’t leave us with a tiny group like in group 2 below (6 clusters). There is still disparity between 3,4,5 and 1,2

Cluster #: 4 3 5 2 1

# in Cluster: 11 13 14 43 55

# First cluster

[1] "B1\_8" "M1\_1" "M1\_2" "M1\_4" "M1\_6" "M2\_10" "M2\_5" "M2\_8"

[9] "M2\_9" "M3\_7" "M4\_5"

# Second Cluster

[1] "1" "2" "10" "12" "14" "15" "17" "18"

[9] "33" "39" "44" "61" "IL4\_8"

# Third Cluster

[1] "11" "B1\_2" "B1\_3" "B3\_4" "B3\_5" "B4\_8" "IL1\_3" "IL2\_5"

[9] "IL2\_8" "M1\_10" "M1\_7" "M1\_8" "M2\_6" "M2\_7"

# Fourth Cluster

[1] "16" "19" "20" "24" "38" "42" "47"

[8] "57" "73" "77" "80" "82" "B1\_10" "B2\_10"

[15] "B2\_7" "B3\_1" "B3\_6" "B3\_7" "B4\_1" "B4\_10" "B4\_3"

[22] "B4\_4" "IL1\_4" "IL1\_5" "IL2\_1" "IL2\_10" "IL2\_7" "IL3\_3"

[29] "IL3\_4" "IL3\_9" "IL4\_4" "IL4\_5" "IL4\_7" "IL4\_9" "M2\_1"

[36] "M2\_3" "M3\_2" "M3\_3" "M3\_8" "M4\_2" "M4\_3" "M4\_6"

[43] "M4\_7"

# Fifth Cluster

[1] "21" "23" "32" "40" "43" "48" "50"

[8] "59" "60" "67" "68" "90" "B1\_1" "B1\_4"

[15] "B1\_5" "B1\_6" "B1\_7" "B1\_9" "B2\_1" "B2\_2" "B2\_4"

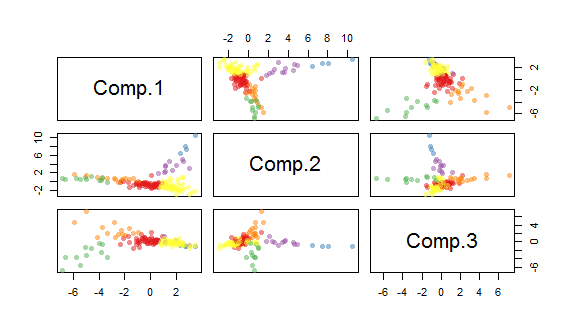
[22] "B2\_5" "B2\_6" "B2\_8" "B2\_9" "B3\_10" "B3\_2" "B3\_3"

[29] "B3\_8" "B3\_9" "B4\_2" "B4\_6" "B4\_7" "B4\_9" "IL1\_1"

[36] "IL1\_10" "IL1\_8" "IL1\_9" "IL2\_2" "IL2\_3" "IL2\_4" "IL2\_9"

[43] "IL3\_1" "IL3\_10" "IL3\_2" "IL3\_6" "IL3\_7" "IL4\_10" "IL4\_2"

[50] "IL4\_3" "M1\_5" "M3\_1" "M3\_5" "M4\_4" "M4\_9"



Figure

K-Means Clustering with 6 clusters. 6 clusters break out fairly well, but groups 1 and 6 have a large number of plots, while group 2 only has 4 plots.

Cluster #: 2 3 4 5 6 1

# in Cluster: 4 11 11 14 46 50

# First cluster

[1] "14" "15" "18" "33"

# Second Cluster

[1] "B1\_8" "M1\_1" "M1\_2" "M1\_4" "M1\_6" "M2\_10" "M2\_5" "M2\_8"

[9] "M2\_9" "M3\_7" "M4\_5"

# Third Cluster

[1] "1" "2" "10" "12" "17" "39" "44" "50"

[9] "61" "IL1\_9" "IL4\_8"

# Fourth Cluster

[1] "11" "B1\_2" "B1\_3" "B3\_4" "B3\_5" "B4\_8" "IL1\_3" "IL2\_5"

[9] "IL2\_8" "M1\_10" "M1\_7" "M1\_8" "M2\_6" "M2\_7"

# Fifth Cluster

[1] "16" "19" "20" "21" "24" "38" "42"

[8] "47" "57" "73" "77" "80" "82" "B1\_10"

[15] "B2\_10" "B2\_7" "B3\_1" "B3\_6" "B3\_7" "B4\_1" "B4\_10"

[22] "B4\_3" "B4\_4" "B4\_9" "IL1\_4" "IL1\_5" "IL2\_1" "IL2\_10"

[29] "IL2\_7" "IL3\_3" "IL3\_4" "IL3\_9" "IL4\_3" "IL4\_4" "IL4\_5"

[36] "IL4\_7" "IL4\_9" "M2\_1" "M2\_3" "M3\_2" "M3\_3" "M3\_8"

[43] "M4\_2" "M4\_3" "M4\_6" "M4\_7"

# Sixth Cluster

[1] "23" "32" "40" "43" "48" "59" "60"

[8] "67" "68" "90" "B1\_1" "B1\_4" "B1\_5" "B1\_6"

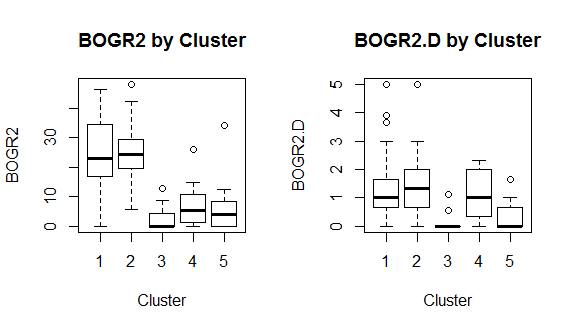
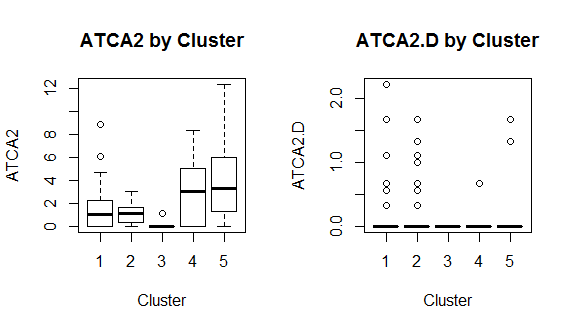
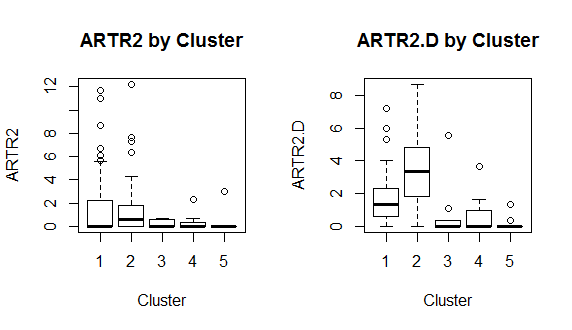
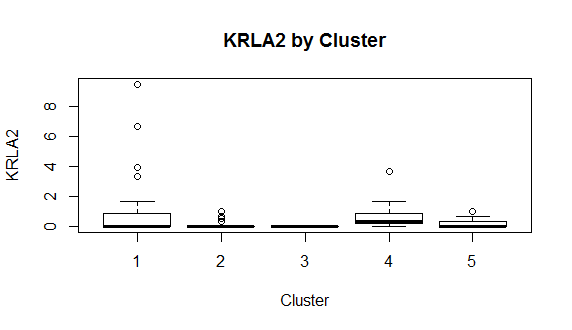
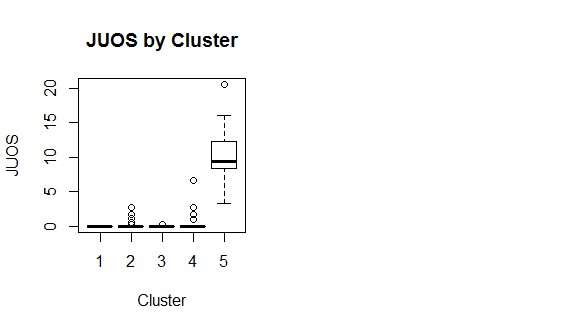
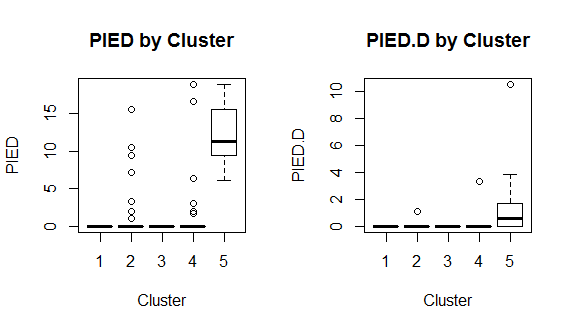
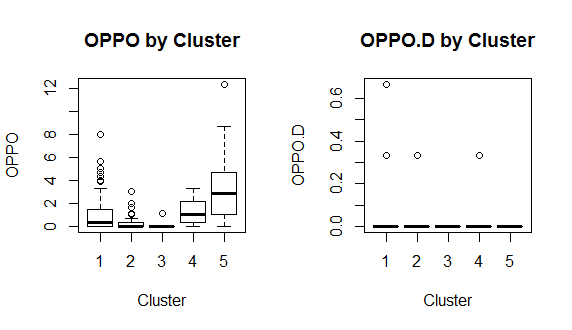
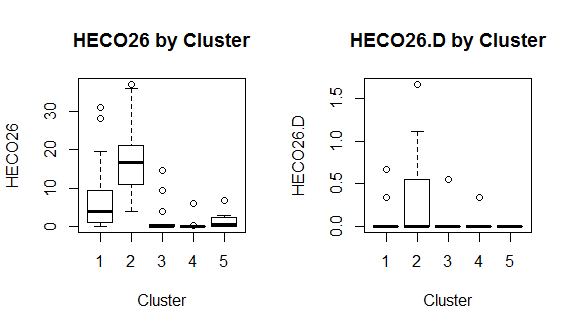
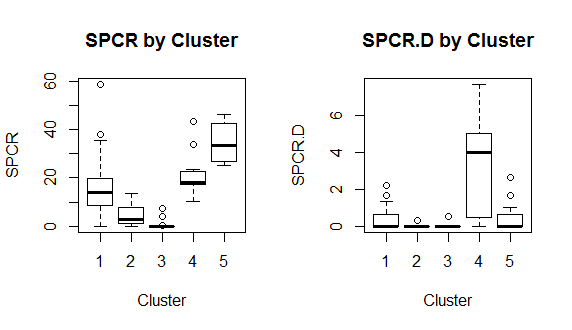
[15] "B1\_7" "B1\_9" "B2\_1" "B2\_2" "B2\_4" "B2\_5" "B2\_6"

[22] "B2\_8" "B2\_9" "B3\_10" "B3\_2" "B3\_3" "B3\_8" "B3\_9"

[29] "B4\_2" "B4\_6" "B4\_7" "IL1\_1" "IL1\_10" "IL1\_8" "IL2\_2"

[36] "IL2\_3" "IL2\_4" "IL2\_9" "IL3\_1" "IL3\_10" "IL3\_2" "IL3\_6"

[43] "IL3\_7" "IL4\_10" "IL4\_2" "M1\_5" "M3\_1" "M3\_5" "M4\_4"

[50] "M4\_9" ****

# PCA before K-Means

# LPI Data

data <- read.csv("F:/LPI/Output/USGSLPIPercentCover.csv",header=TRUE, row.names=1)

# plot variance of columns

mar <- par()$mar

par(mar=mar+c(0,5,0,0))

par(mfrow=c(1,2))

barplot(sapply(data, var), horiz=T, las=1, cex.names=0.5)

barplot(sapply(data, var), horiz=T, las=1, cex.names=0.5, log='x')

par(mar=mar)

# Scale

data2 <- data.frame(scale(data))

# Verify variance is uniform

plot(sapply(data2, var))

# Proceed with principal components

pc <- princomp(data2)

plot(pc)

plot(pc, type='l')

summary(pc) # 3 components is 'elbow' but does not explain >85% variance(would need 23 components)(only explains about 7%)

# First few principal components

comp <- data.frame(pc$scores[,1:3])

# Plot

plot(comp, pch=16, col=rgb(0,0,0,0.5))

library(rgl)

# Multi 3D plot

plot3d(comp$Comp.1, comp$Comp.2, comp$Comp.3)

### K-Means ###

# Determine number of clusters

wss <- (nrow(comp)-1)\*sum(apply(comp,2,var))

for (i in 2:15) wss[i] <- sum(kmeans(comp,

centers=i)$withinss)

plot(1:15, wss, type="b", xlab="Number of Clusters",

ylab="Within groups sum of squares")

# From scree plot elbow occurs at k = 4 (or 6?)

# Apply k-means with k=4 (then try 6)

k <- kmeans(comp, 5, nstart=25, iter.max=1000)

library(RColorBrewer)

library(scales)

palette(alpha(brewer.pal(9,'Set1'), 0.5))

plot(comp, col=k$clust, pch=16)

# 3D plot

plot3d(comp$Comp.1, comp$Comp.2, comp$Comp.3, col=k$clust)

# Cluster sizes

sort(table(k$clust))

clust <- names(sort(table(k$clust)))

# First cluster

row.names(data[k$clust==clust[1],])

# Second Cluster

row.names(data[k$clust==clust[2],])

# Third Cluster

row.names(data[k$clust==clust[3],])

# Fourth Cluster

row.names(data[k$clust==clust[4],])

# Fifth Cluster

row.names(data[k$clust==clust[5],])

# Sixth Cluster

row.names(data[k$clust==clust[6],])

# Compare accommodation by cluster in boxplot

boxplot(data$KRLA2 ~ k$cluster,

xlab='Cluster', ylab='KRLA2',

main='KRLA2 by Cluster')

par(mfrow=c(1,2))

boxplot(data$ARTR2 ~ k$cluster,

xlab='Cluster', ylab='ARTR2',

main='ARTR2 by Cluster')

boxplot(data$ARTR2.D ~ k$cluster,

xlab='Cluster', ylab='ARTR2.D',

main='ARTR2.D by Cluster')

boxplot(data$ATCA2 ~ k$cluster,

xlab='Cluster', ylab='ATCA2',

main='ATCA2 by Cluster')

boxplot(data$ATCA2.D ~ k$cluster,

xlab='Cluster', ylab='ATCA2.D',

main='ATCA2.D by Cluster')

boxplot(data$BOGR2 ~ k$cluster,

xlab='Cluster', ylab='BOGR2',

main='BOGR2 by Cluster')

boxplot(data$BOGR2.D ~ k$cluster,

xlab='Cluster', ylab='BOGR2.D',

main='BOGR2.D by Cluster')

boxplot(data$SPCR ~ k$cluster,

xlab='Cluster', ylab='SPCR',

main='SPCR by Cluster')

boxplot(data$SPCR.D ~ k$cluster,

xlab='Cluster', ylab='SPCR.D',

main='SPCR.D by Cluster')

boxplot(data$HECO26 ~ k$cluster,

xlab='Cluster', ylab='HECO26',

main='HECO26 by Cluster')

boxplot(data$HECO26.D ~ k$cluster,

xlab='Cluster', ylab='HECO26.D',

main='HECO26.D by Cluster')

boxplot(data$OPPO ~ k$cluster,

xlab='Cluster', ylab='OPPO',

main='OPPO by Cluster')

boxplot(data$OPPO.D ~ k$cluster,

xlab='Cluster', ylab='OPPO.D',

main='OPPO.D by Cluster')

boxplot(data$PIED ~ k$cluster,

xlab='Cluster', ylab='PIED',

main='PIED by Cluster')

boxplot(data$PIED.D ~ k$cluster,

xlab='Cluster', ylab='PIED.D',

main='PIED.D by Cluster')

boxplot(data$JUOS ~ k$cluster,

xlab='Cluster', ylab='JUOS',

main='JUOS by Cluster')