```
# User Values Survey Data Analysis
# Charlotte Zoe Smith - 20 April 2019

#Import required libaries
library(reshape2)
library(plyr)
library(Hmisc)
library(dplyr)
library(data.table)
library(cluster)
library(factoextra) # clustering algorithms & visualization


# Import the CSV data file output for the survey. Change the
file path to point to the file locally on your own personal
computer
data =
read.csv("/Users/Char/Documents/user_values_responses.csv",
header = TRUE, na.strings=c("","NA"), stringsAsFactors = FALSE)

# View the csv as a table
View(data)

summary(data)
### Rename the header of the consent form column to consent
names(data)[2] <- "consent"

##Rename the column headers for frequency of use questions
names(data)[3] <- "freq.visit"
names(data)[4] <- "freq.scenery"
names(data)[5] <- "freq.swimming"
names(data)[6] <- "freq.fishing"
names(data)[7] <- "freq.limu"
names(data)[8] <- "freq.surfing"
names(data)[9] <- "freq.cleanup"
names(data)[10] <- "freq.paddling"
names(data)[11] <- "freq.hiking"
names(data)[12] <- "freq.culture"
names(data)[13] <- "freq.bbq"
names(data)[14] <- "freq.comfish"
names(data)[15] <- "freq.work"

#Rename column headers for the conservation values
names(data)[16] <- "bio.conserv"
names(data)[17] <- "livelihoods"
names(data)[18] <- "economy"
names(data)[19] <- "tour.beauty"
```

```
names(data)[20] <- "prop.values"
names(data)[21] <- "transgen.trad"
names(data)[22] <- "water.qual"
names(data)[23] <- "corals"
names(data)[24] <- "birds"
names(data)[25] <- "limu"
names(data)[26] <- "fish"
names(data)[27] <- "end.spec"

# Rename column headers for the sense of place values
names(data)[28] <- "plc.wildlife"
names(data)[29] <- "plc.network"
names(data)[30] <- "plc.job"
names(data)[31] <- "plc.landscape"
names(data)[32] <- "plc.lifestyle.values"
names(data)[33] <- "plc.part.of.me"
names(data)[34] <- "plc.best.activity"
names(data)[35] <- "plc.identity"
names(data)[36] <- "plc.satisfaction"
names(data)[37] <- "plc.attachement"
names(data)[38] <- "plc.who.i.am"
names(data)[39] <- "plc.time.invest"
names(data)[40] <- "plc.lifestyle"
names(data)[41] <- "plc.friends"
names(data)[42] <- "plc.phys.landscape"
names(data)[43] <- "plc.outdoor.act"
names(data)[44] <- "plc.land.connection"
names(data)[45] <- "plc.people.connection"

#Rename the column headers for the demographic information
names(data)[46]<-"gender"
names(data)[47]<-"age"
names(data)[48]<-"ethnicity"
names(data)[49]<-"tour.fish.ind"
names(data)[50]<-"civic.group"
names(data)[51]<-"env.group"
names(data)[52]<-"funding"
names(data)[53]<-"vol.awareness"
names(data)[54]<-"education"
names(data)[55]<-"family.income"
names(data)[56]<-"residency.locat"
names(data)[57]<-"residency"
names(data)[58]<-"months.in.hi"
names(data)[59]<-"comfish.policy"

# Subset the survey responses to ONLY include those that signed
the consent form
```

```
only.consent <- data[!is.na(data$consent),]

#view filtered dataset in new window
View(only.consent)
#print a summary of this dataset to console
summary(only.consent)

### Change likert responses to numeric
res <- only.consent %>%
  mutate_at(.vars =vars(3), funs(recode(.,"Never" = 0, "2-3
times a year" = 1,
                                        "1-2 times a month"=
2,
                                        "1-2 times a week"=
3,
                                        "Almost every day"=
4,
                                        "I live here" = 5,
.default = NaN)))

res1 <- res %>%
  mutate_at(.vars =vars(4:15), funs(recode(.,"Never" = 1,
                                    "Rarely (Once a month or
less)"= 2,
                                    "Occasionally (2-3 times a
month)"= 3,
                                    "Frequently (2-3 times a
week)"= 4,
                                   "Always (1-2 times a day)" =
5, .default = NaN)))


res2 <- res1 %>%
  mutate_at(.vars =vars(16:45), funs(recode(.,"Not Applicable" =
0,
                                        "Strongly Disagree" =
1,
                                        "Disagree" = 2,
                                        "Neutral" = 3,
                                        "Agree" = 4,
                                        "Strongly Agree" = 5
, .default = NaN)))

summary(res2)

View(res2)
```

```
##### Create new columns that are the average of the different
values groups
library(data.table)

# Create a new column plc.biophysical to contain the mean of the
biophysical sense of place values
res2$plc.biophysical <- rowMeans(subset(res2, select =
c(plc.wildlife,plc.landscape,plc.outdoor.act,plc.phys.landscape)
), na.rm = TRUE)

# Create a new column plc.psychological to contain the mean of
the psychological sense of place values
res2$plc.psychological <- rowMeans(subset(res2, select =
c(plc.lifestyle.values,plc.part.of.me,plc.best.activity,plc.iden
tity,plc.satisfaction,plc.attachement,plc.who.i.am,plc.lifestyle
)), na.rm = TRUE)

# Create a new column plc.economical to contain the mean of the
economical sense of place values
res2$plc.economical <- rowMeans(subset(res2, select =
c(plc.job,plc.time.invest)), na.rm = TRUE)

# Create a new column plc.sociocultural to contain the mean of
the sociocultural sense of place values
res2$plc.sociocultural <- rowMeans(subset(res2, select =
c(plc.friends,plc.people.connection,plc.land.connection,plc.netw
ork)), na.rm = TRUE)

# Create a new column val.environmental to contain environmental
values statements
res2$val.environment <- rowMeans(subset(res2, select =
c(bio.conserv,water.qual,corals,birds,fish,end.spec)), na.rm =
TRUE)

# Create a new column val.econ to contain the mean of the
economic values statements
res2$val.econ <- rowMeans(subset(res2, select =
c(livelihoods,economy,tour.beauty,prop.values)), na.rm = TRUE)

# Create a new column val.culture to contain the mean of the
culture values statements
res2$val.culture <- rowMeans(subset(res2, select =
c(transgen.trad,limu)), na.rm = TRUE)

# Create a new column freq.uses to contain the mean of the
frequency of use values
```

```
res2$freq.use<- rowMeans(subset(res2, select =
c(freq.scenery,freq.swimming,freq.fishing,freq.limu,freq.surfing
,freq.cleanup,freq.paddling,freq.hiking,freq.culture,freq.bbq,
freq.comfish,freq.work)), na.rm = TRUE)

View(res2)


## Scale data
scaled.dat <- scale(res2[60:67])
summary(scaled.dat)

View(scaled.dat)


###### k-means clustering for env values

#Determine the optimal number of centroids
fviz_nbclust(scaled.dat, kmeans, method = "wss")

results <- kmeans(na.omit(scaled.dat), centers = 3, nstart = 25)
# this helps in omitting NA
results

# visualize the clustering
fviz_cluster(results, data = scaled.dat)

#Add a column with the environmental cluster information to the
original data frame
res2$cluster <- as.factor(results$cluster)

View(res2)

#Create a table of the clusters and the means of each cluster
per each variable
cluster.table<-res2 %>%
  mutate(cluster = res2$cluster) %>%
  group_by(cluster) %>%
  summarise_all("mean")

View(cluster.table)


#### Calculate the demographic characteristics of each cluster
group

demog<-res2[46:59]
```

```
# Add in the column with the cluster data
demog$cluster<-res2$cluster

View(demog)

# Add an ID field
demog$ID <- seq.int(nrow(demog))

demog[, .N, by = .(ID, gender)] %>% dcast(ID ~ gender)

#make separate tables per each cluster
cluster1 <-subset(demog,cluster==1)
cluster2 <-subset(demog,cluster==2)
cluster3 <-subset(demog,cluster==3)
cluster4 <-subset(demog,cluster==4)

View(cluster1)
View(cluster2)
View(cluster3)
View(cluster4)


####END####
```