Grouping Things

JuJu

## Grab the Data

Data from an in-class survey was inputted into Excel and then exported to CSV. The file is imported from the CSV file into R.

likes = read.csv("namesandlikes2016.csv")  
 head(likes)

## Fname Gender Sport1 Sport2 Sport3 Politics AngelsDodgers  
## 1 Chris Male Baseball Football Basketball No Dodgers  
## 2 Cameron Male Golf Basketball Football No Neither  
## 3 Michael Male Swimming Water Polo Hockey No Dodgers  
## 4 Samantha Female Water Polo Swimming Basketball Yes Dodgers  
## 5 Melissa Female Soccer Football Baseball Yes Dodgers  
## 6 Monica Female Soccer Swimming Football Yes Dodgers  
## EatSleep CatsDogs SandWater BurgerHotDog  
## 1 Sleep Dogs Both Burger  
## 2 Sleep Neither Sand Burger  
## 3 Eat Dogs Water Burger  
## 4 Eat Dogs Water Burger  
## 5 Eat Dogs Sand Hot Dog  
## 6 Sleep Neither Water Burger

## Chernoff Faces

People do a good job of classifying things by using visual cues. Herman Chernoff realized that we can use a graphical face with features tied to variables to classify things.

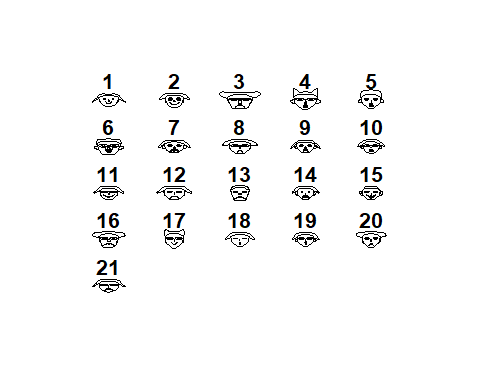
p\_load(aplpack)  
  
 ### Convert factor variables to numeric for faces  
 temp <- likes[,-1] ### Make a copy of likes without the Fname (1st) column  
   
 as.numeric.factor <- function(x) { ### Create a function that acts on a vector  
 x = as.factor(x) ### Convert the string values to factor  
 seq\_along(levels(x))[x] ### Change the factor values (character) into their levels (numeric)  
 }  
   
 likes ### Take a look at what's in likes

## Fname Gender Sport1 Sport2 Sport3 Politics AngelsDodgers  
## 1 Chris Male Baseball Football Basketball No Dodgers  
## 2 Cameron Male Golf Basketball Football No Neither  
## 3 Michael Male Swimming Water Polo Hockey No Dodgers  
## 4 Samantha Female Water Polo Swimming Basketball Yes Dodgers  
## 5 Melissa Female Soccer Football Baseball Yes Dodgers  
## 6 Monica Female Soccer Swimming Football Yes Dodgers  
## 7 Topher Male Basketball Football Boxing Yes Neither  
## 8 Derrick Male Golf Soccer Basketball No Dodgers  
## 9 Blake Male Baseball Basketball Boxing Yes Dodgers  
## 10 Cole Male Football Baseball Basketball Yes Dodgers  
## 11 Will Male Football Basketball Basketball No Neither  
## 12 Alfonso Male Football Soccer Basketball No Neither  
## 13 Charly Female Lacrosse Basketball Tennis No Neither  
## 14 Samuel Male Football Baseball Basketball Yes Neither  
## 15 Katryna Female Football Basketball Baseball Yes Neither  
## 16 Carlos Male Soccer Football Track No Neither  
## 17 Sarah Female Tennis Basketball Baseball No Dodgers  
## 18 Mitchell Male Football Basketball Baseball No Angels  
## 19 Jonathan Male Football Baseball Basketball Yes Angels  
## 20 Christian Male Soccer Basketball Boxing No Dodgers  
## 21 David Male Basketball Football Baseball No Neither  
## EatSleep CatsDogs SandWater BurgerHotDog  
## 1 Sleep Dogs Both Burger  
## 2 Sleep Neither Sand Burger  
## 3 Eat Dogs Water Burger  
## 4 Eat Dogs Water Burger  
## 5 Eat Dogs Sand Hot Dog  
## 6 Sleep Neither Water Burger  
## 7 Eat Dogs Sand Burger  
## 8 Eat Dogs Water Burger  
## 9 Eat Dogs Sand Burger  
## 10 Eat Dogs Water Burger  
## 11 Sleep Dogs Water Burger  
## 12 Eat Dogs Sand Burger  
## 13 Eat Dogs Sand Burger  
## 14 Eat Both Both Burger  
## 15 Sleep Dogs Sand Burger  
## 16 Eat Dogs Water Burger  
## 17 Sleep Dogs Water Burger  
## 18 Eat Both Sand Burger  
## 19 Eat Dogs Water Burger  
## 20 Eat Dogs Sand Burger  
## 21 Eat Dogs Water Burger

temp <- apply(temp,2,as.numeric.factor) ### Apply the function as.numeric.factor to the columns (2) of temp  
 temp ### Look at temp to see if the numeric coding has worked

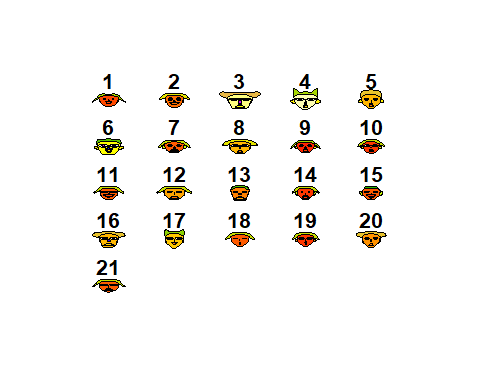
## Gender Sport1 Sport2 Sport3 Politics AngelsDodgers EatSleep CatsDogs  
## [1,] 2 1 3 2 1 2 2 2  
## [2,] 2 4 2 4 1 3 2 3  
## [3,] 2 7 6 5 1 2 1 2  
## [4,] 1 9 5 2 2 2 1 2  
## [5,] 1 6 3 1 2 2 1 2  
## [6,] 1 6 5 4 2 2 2 3  
## [7,] 2 2 3 3 2 3 1 2  
## [8,] 2 4 4 2 1 2 1 2  
## [9,] 2 1 2 3 2 2 1 2  
## [10,] 2 3 1 2 2 2 1 2  
## [11,] 2 3 2 2 1 3 2 2  
## [12,] 2 3 4 2 1 3 1 2  
## [13,] 1 5 2 6 1 3 1 2  
## [14,] 2 3 1 2 2 3 1 1  
## [15,] 1 3 2 1 2 3 2 2  
## [16,] 2 6 3 7 1 3 1 2  
## [17,] 1 8 2 1 1 2 2 2  
## [18,] 2 3 2 1 1 1 1 1  
## [19,] 2 3 1 2 2 1 1 2  
## [20,] 2 6 2 3 1 2 1 2  
## [21,] 2 2 3 1 1 3 1 2  
## SandWater BurgerHotDog  
## [1,] 1 1  
## [2,] 2 1  
## [3,] 3 1  
## [4,] 3 1  
## [5,] 2 2  
## [6,] 3 1  
## [7,] 2 1  
## [8,] 3 1  
## [9,] 2 1  
## [10,] 3 1  
## [11,] 3 1  
## [12,] 2 1  
## [13,] 2 1  
## [14,] 1 1  
## [15,] 2 1  
## [16,] 3 1  
## [17,] 3 1  
## [18,] 2 1  
## [19,] 3 1  
## [20,] 2 1  
## [21,] 3 1

### Plot faces  
 faces(temp[,c(2:10,1)],face.type=0) ### Note that the Gender variable has been moved to



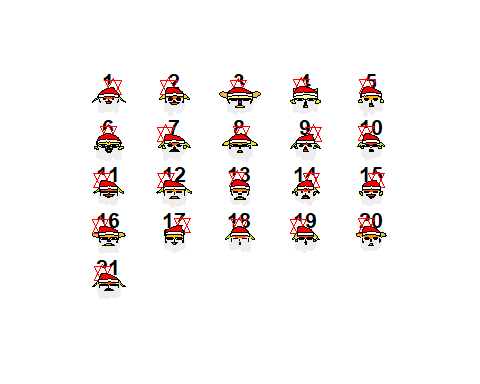
## effect of variables:  
## modified item Var   
## "height of face " "Sport1"   
## "width of face " "Sport2"   
## "structure of face" "Sport3"   
## "height of mouth " "Politics"   
## "width of mouth " "AngelsDodgers"  
## "smiling " "EatSleep"   
## "height of eyes " "CatsDogs"   
## "width of eyes " "SandWater"   
## "height of hair " "BurgerHotDog"   
## "width of hair " "Gender"   
## "style of hair " "Sport1"   
## "height of nose " "Sport2"   
## "width of nose " "Sport3"   
## "width of ear " "Politics"   
## "height of ear " "AngelsDodgers"

faces(temp[,c(2:10,1)],face.type=1) ### last position as it is a characteristic and not a like



## effect of variables:  
## modified item Var   
## "height of face " "Sport1"   
## "width of face " "Sport2"   
## "structure of face" "Sport3"   
## "height of mouth " "Politics"   
## "width of mouth " "AngelsDodgers"  
## "smiling " "EatSleep"   
## "height of eyes " "CatsDogs"   
## "width of eyes " "SandWater"   
## "height of hair " "BurgerHotDog"   
## "width of hair " "Gender"   
## "style of hair " "Sport1"   
## "height of nose " "Sport2"   
## "width of nose " "Sport3"   
## "width of ear " "Politics"   
## "height of ear " "AngelsDodgers"

faces(temp[,c(2:10,1)],face.type=2)



## effect of variables:  
## modified item Var   
## "height of face " "Sport1"   
## "width of face " "Sport2"   
## "structure of face" "Sport3"   
## "height of mouth " "Politics"   
## "width of mouth " "AngelsDodgers"  
## "smiling " "EatSleep"   
## "height of eyes " "CatsDogs"   
## "width of eyes " "SandWater"   
## "height of hair " "BurgerHotDog"   
## "width of hair " "Gender"   
## "style of hair " "Sport1"   
## "height of nose " "Sport2"   
## "width of nose " "Sport3"   
## "width of ear " "Politics"   
## "height of ear " "AngelsDodgers"