Applying a Principal Components Analysis to the United States Department of Agriculture Food Composition Database

Maureen Renaud

6/23/2020

INTRODUCTION

In this study, I analyzed a data set from the United States Department of Agriculture Food Composition Database[1]. This database contains data on 8,789 food items and up to 150 food components. This particular study analyzes 2,223 food items and 46 nutrient variables with Principal Component Analysis. After examining the data, I chose to keep the first six Principal Components to adequately represent this data. I found that most prominent nutrients in these components were Water, Energy, Total Lipids, Protein, Magnesium, and Carbohydrates. Foods with higher amounts of water typically had lower levels of other nutrients and foods with higher levels of other nutrients typically had lower levels of water.

DATA

The data set used in this study is from the United States Department of Agriculture Food Composition Database. According to *USDA National Nutrient Database for Standard Reference, Release 28*[1], it was developed in September 2015 and slightly revised in May 2016. The database contains 8,789 food items and up to 150 food components.
Before begining any analyses, the data was pared down to eliminate rows with missing values, duplicate values, and non-nutrient variables. This left a data set with 2,223 food items and 46 nutrient variables.

A complete summary is provided in Figure 1 of the Plots and Tables Section.

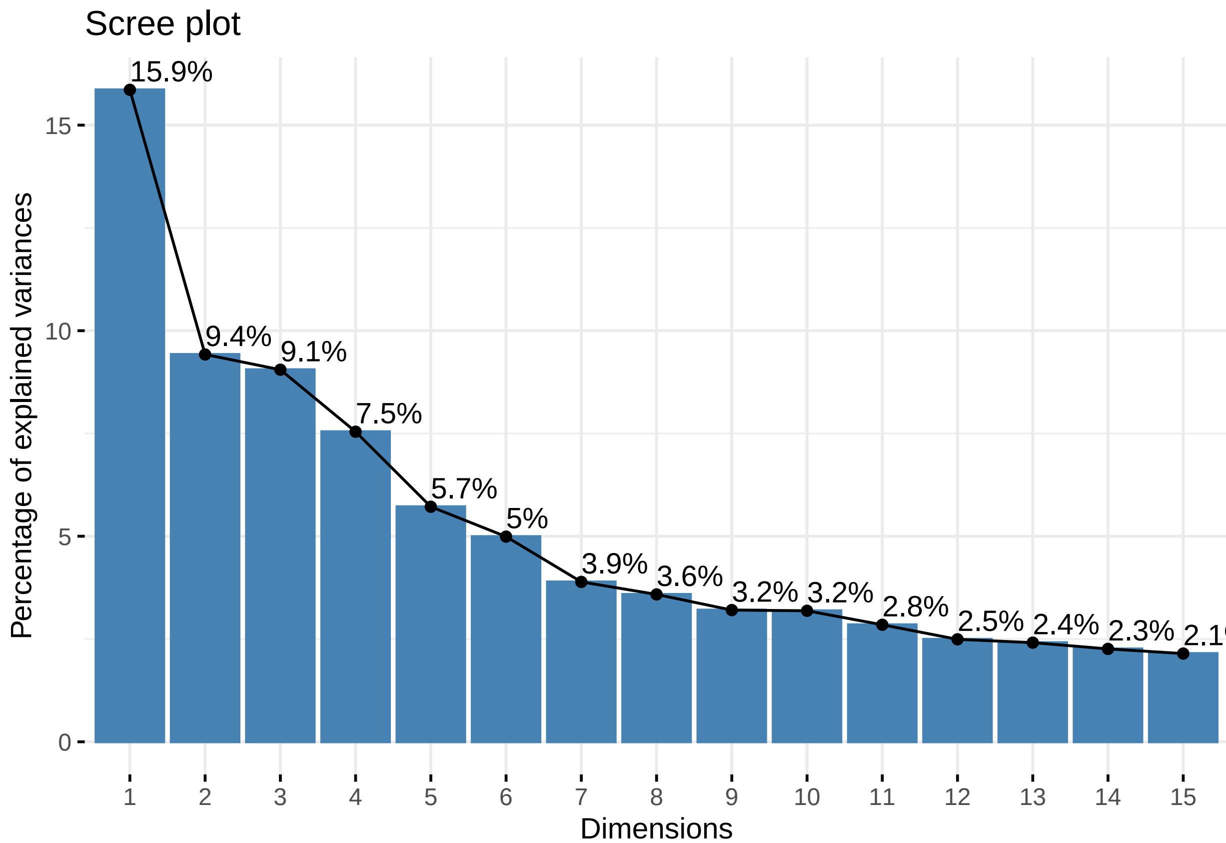
ANALYSES

I first ran a Principal Component Analysis on the cleaned up data set:

```
pr.out <- prcomp(SRp, scale = T)
```

SCREE PLOT AND CHOOSING PRINCIPAL COMPONENTS

After running the PCA, we can examine the scree plot in order to determine how many principal components to focus on.



In this plot you can see that the the first component explains more than 15% of the variance and subsequent components explains significantly less. In order to determine how many of these components we should consider, there are a few schools of thought.

First, we could look for the "elbow" in the plot. There doesn't not appear to be a significant elbow, but a small one does appear after the sixth component. So, we could consider just the first six components.

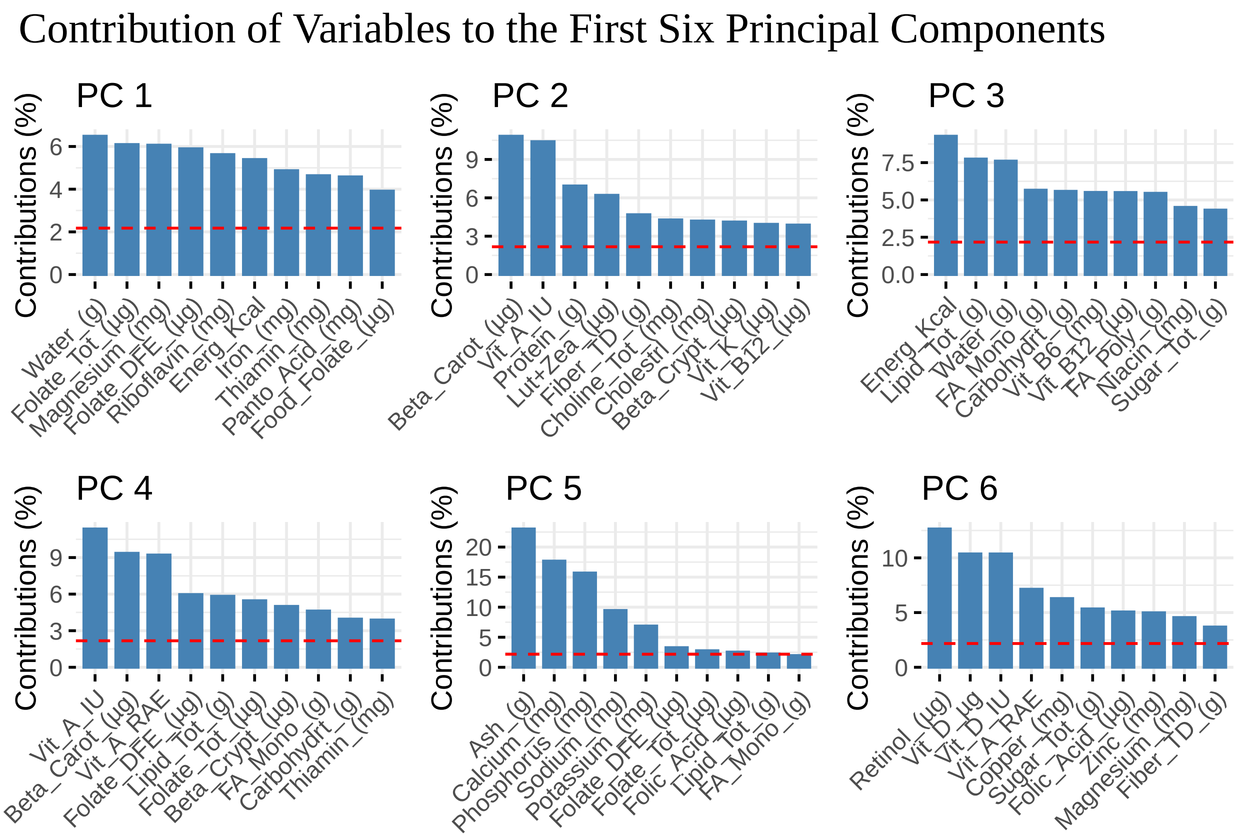
We could consider any component that explains more than it's "fair" share of the variance, fair being defined as more than 1/p of the variance, according to the Tutorial *Principal Components Analysis (PCA)*[2]. In this case, 1/p = 1/46 = .02173913. So, we would consider the first 14 components.

We could also consider the components that explain 65% of the data. In this case, we would need to consider the first ten principal components.

It could actually be counter-productive to evaluate too many principal components due to the overwhelming amount of data, so for this study I will choose the "elbow" method and focus on the first six components.

EXAMINING THE MOST PROMINENT COMPONENTS

Next, I would like to determine the nutrients that are most prominent in the first six components. One way to do this is to look at the most prevalent nutrients in each component.



This is certainly informative, however because the nutrients vary between the components, its hard to determine which ones are most prominent in all six components.

In order to rectify this, we look to the text *An Introduction to Statistical Learning*[3], which notes that the first principal component of a set of features is the normalized linear combination of the features. Normalized means that the sum of the squared loadings is equal to one.

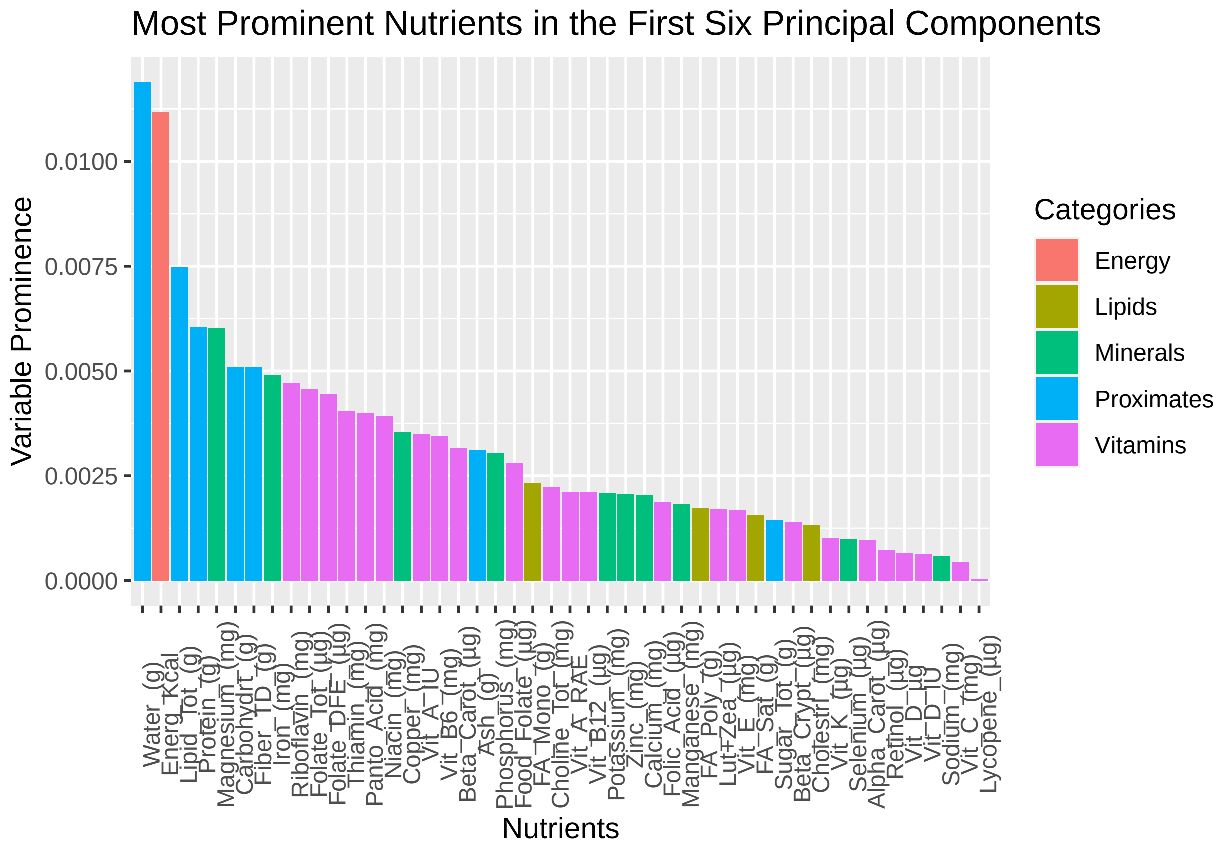
Squaring the loading of each element in a given component gives its relative contribution to that component. If we multiply that contribution by the proportion of variance explained by each component, we will be able to calculate which nutrients have the greatest impact over multiple components.

The below table is based on this method. Additonally, I added a new column indicating whether each nutrient was a Proximate, Vitamin, Mineral, or Lipid to help with the visualizations later in this study. The USDA nutritional reference[1] does not include Energy in any of these categories, so I have left it as a category of its own.

```
##                        value Categories
## Water_(g)       1.189419e-02 Proximates
## Energ_Kcal      1.117020e-02     Energy
## Protein_(g)     6.055474e-03 Proximates
## Lipid_Tot_(g)   7.487195e-03 Proximates
## Ash_(g)         3.101866e-03 Proximates
## Carbohydrt_(g)  5.086410e-03 Proximates
## Fiber_TD_(g)    5.083987e-03 Proximates
## Sugar_Tot_(g)   1.443864e-03 Proximates
## Calcium_(mg)    2.048153e-03   Minerals
## Iron_(mg)       4.911252e-03   Minerals
## Magnesium_(mg)  6.027040e-03   Minerals
```

```
## Phosphorus_(mg)  3.045058e-03   Minerals
## Potassium_(mg)   2.083900e-03   Minerals
## Sodium_(mg)      5.733231e-04   Minerals
## Zinc_(mg)        2.052686e-03   Minerals
## Copper_(mg)      3.537498e-03   Minerals
## Manganese_(mg)   1.826133e-03   Minerals
## Selenium_(µg)    1.001107e-03   Minerals
## Vit_C_(mg)       4.440036e-04   Vitamins
## Thiamin_(mg)     4.049980e-03   Vitamins
## Riboflavin_(mg)  4.699868e-03   Vitamins
## Niacin_(mg)      3.913026e-03   Vitamins
## Panto_Acid_(mg)  4.000273e-03   Vitamins
## Vit_B6_(mg)      3.440858e-03   Vitamins
## Folate_Tot_(µg)  4.557056e-03   Vitamins
## Folic_Acid_(µg)  1.883179e-03   Vitamins
## Food_Folate_(µg) 2.811786e-03   Vitamins
## Folate_DFE_(µg)  4.443458e-03   Vitamins
## Choline_Tot_(mg) 2.234566e-03   Vitamins
## Vit_B12_(µg)     2.102527e-03   Vitamins
## Vit_A_IU         3.484647e-03   Vitamins
## Vit_A_RAE        2.109882e-03   Vitamins
## Retinol_(µg)     7.188890e-04   Vitamins
## Alpha_Carot_(µg) 9.607933e-04   Vitamins
## Beta_Carot_(µg)  3.159296e-03   Vitamins
## Beta_Crypt_(µg)  1.390050e-03   Vitamins
## Lycopene_(µg)    4.562848e-05   Vitamins
## Lut+Zea_(µg)     1.705266e-03   Vitamins
## Vit_E_(mg)       1.677682e-03   Vitamins
## Vit_D_µg         6.543447e-04   Vitamins
## Vit_D_IU         6.291424e-04   Vitamins
## Vit_K_(µg)       1.019093e-03   Vitamins
## FA_Sat_(g)       1.565459e-03     Lipids
## FA_Mono_(g)      2.331181e-03     Lipids
## FA_Poly_(g)      1.726460e-03     Lipids
## Cholestrl_(mg)   1.333976e-03     Lipids
```

We can now reorder and plot these elements to get an idea of which ones have the greatest contribution to the first six principal components.

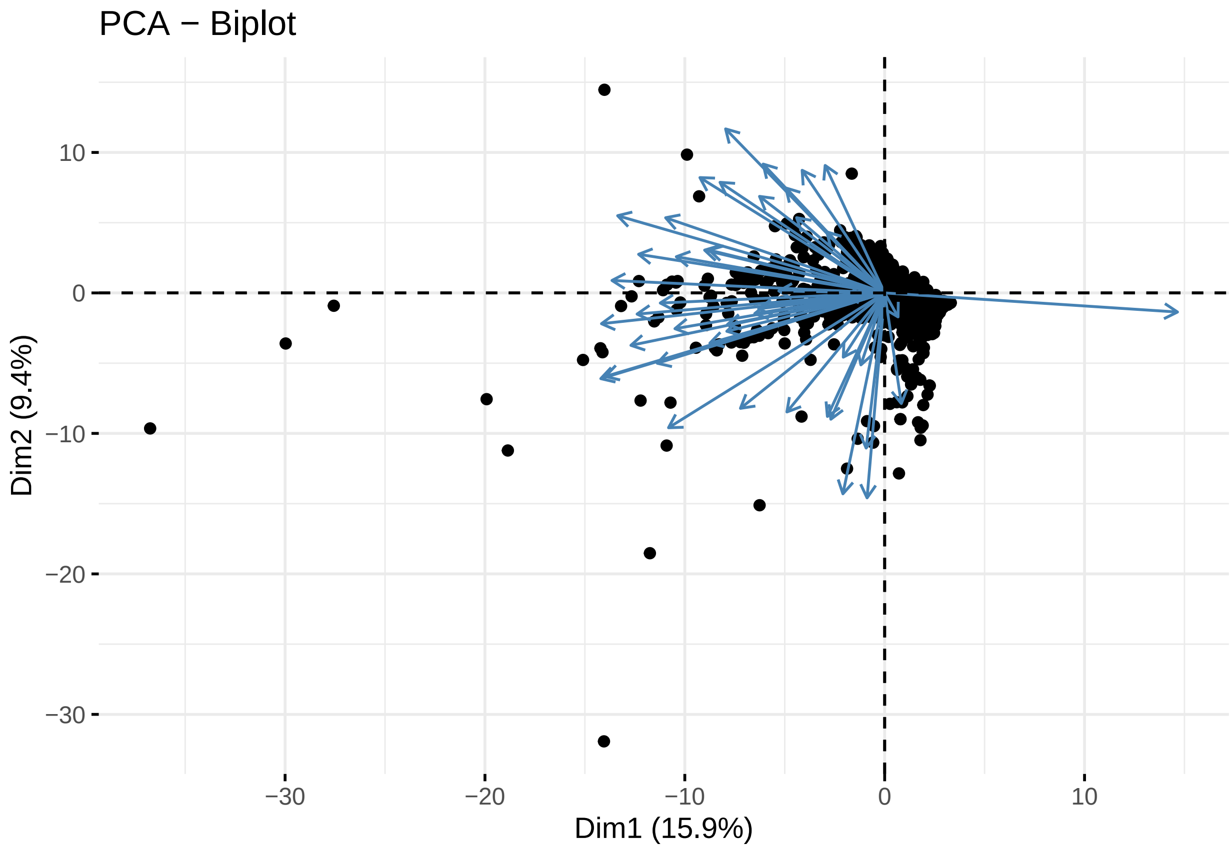


The plot is also color coded based on whether the nutrient is a proximate, vitamin, mineral, or lipid.

It's interesting to note that the proximates make up five of the seven largest contributers to the top six principal components. According to the nutritional reference[1], proximates are the macronutrients that make up our food. So, it is logical that they are going to be among the most prominent nutrients in the various food items. I also find it interesting to see Magnesium in the top six.

BIPLOT

Next we move to a biplot of the individuals and variables in the first two principal components. Because there are so many individuals and variables in this study, the biplot is rather cluttered:

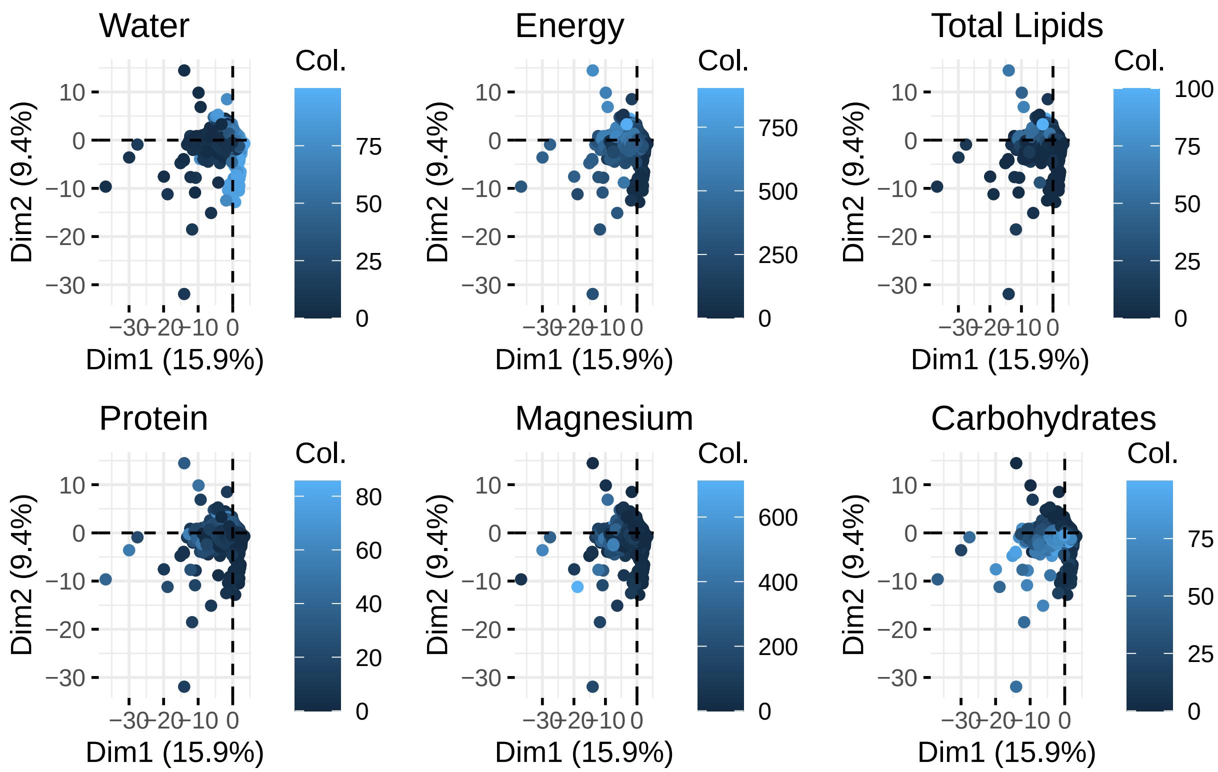


To combat this situation, I have broken the plots down into individuals plots and variables plots.

INDIVIDUALS PLOT

Now that I have identified the most prominent variables, we will examine plots of the indiviudals (foods) in the first and second components. The color gradient represents the contribution of the specified nutrient to those indiviudals. Lighter blue colors are indicative of a greater amount of the specified nutrient.

Individiual Plots

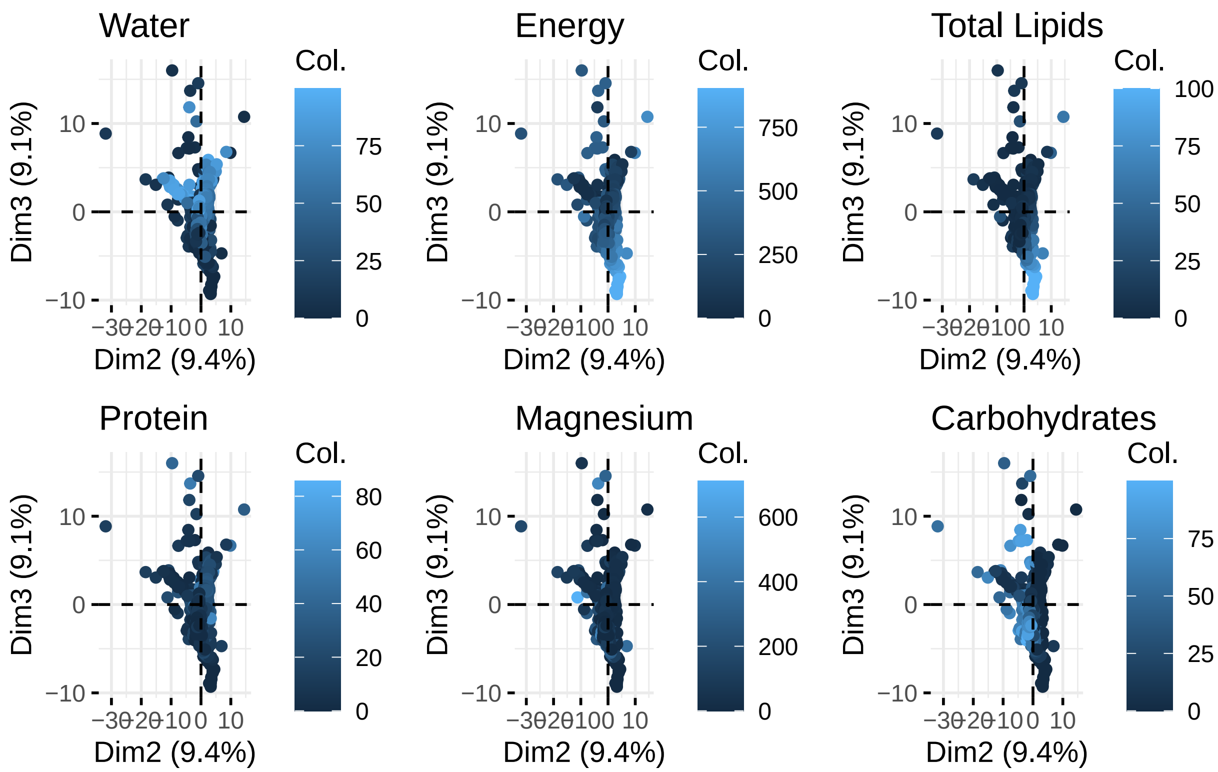


There is clear clustering between foods that contain significant amounts of water and those that do not. There is moderate evidence of clustering between foods that contain significant amounts of energy, lipids, and carbohydrates and those that do not.

Another intersting observation comparing all six plots is that foods with higher amounts of water tend to have less of all other categories (energy, total lipids, protein, magnesium, and carbohydrates). This is particularly noticeable when you compare the Water plot and the Carbohydrate plot.

Just to check if there are any major differences, we will also take a look at the individuals plot for principal components two and three:

Individiual Plots

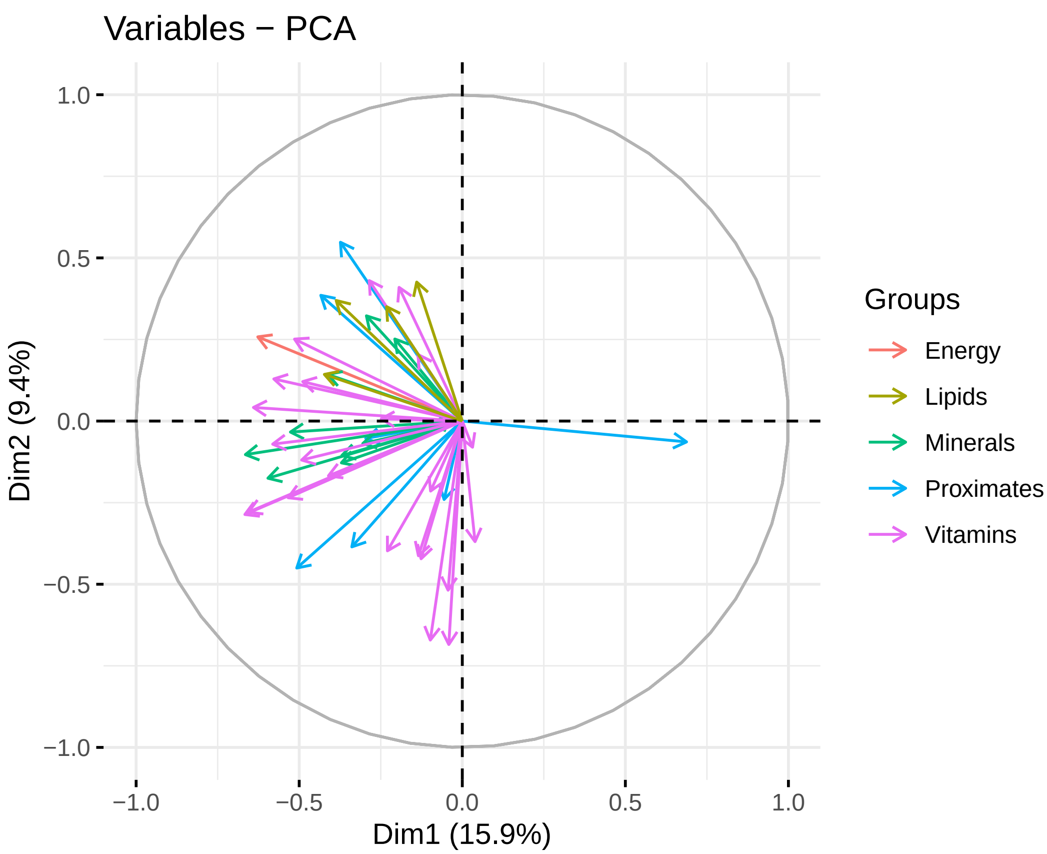


The pattern evident between the first two principal components is also evident between the second and third component. There is a clear clustering between foods that contain significant amounts of water and those that do not and moderate evidence of clustering between foods that contain significant amounts of energy, lipids, and carbohydrates and those that do not.

VARIABLES PLOT

Now that I have identified the most prominent variables, we will examine a plot of the variables (nutrients) in the first and second components. Rather than consider all 46 nutrients, I used the previously developed groups (Proximates, Vitamins, Minerals, Lipids, and Energy) to compare nutrients.

```
fviz_pca_var(pr.out, labels = FALSE, habillage = new.contrib.cat$Categories)
```



It is clear that there is a fair amount of correlation between all of the nutrients but water. It is the lone vector opposing the rest of the plot. This reinforces what we observed in the individuals plots, where foods with higher amounts of water had less of all other nutrients.

It appears that the vitamins are slightly clustered in the third quadrant, but for the most part the main nutrient groups seem to intermingle.

PLOTS AND TABLES

Below is a summary of our initial data set:

```
##     Water_(g)        Energ_Kcal      Protein_(g)     Lipid_Tot_(g)
##  Min.    : 0.00   Min.    :   0.0   Min.    : 0.00   Min.    :  0.00
##  1st Qu.:41.00   1st Qu.: 83.0   1st Qu.: 2.21   1st Qu.:  0.74
##  Median :64.96   Median :182.0   Median :10.08   Median :  5.34
##  Mean   :57.57   Mean   :210.9   Mean   :12.85   Mean   : 10.15
##  3rd Qu.:79.00   3rd Qu.:292.0   3rd Qu.:22.46   3rd Qu.: 13.37
##  Max.   :99.90   Max.   :902.0   Max.   :85.60   Max.   :100.00
##      Ash_(g)          Carbohydrt_(g)    Fiber_TD_(g)      Sugar_Tot_(g)
##  Min.    : 0.000   Min.    : 0.00   Min.    : 0.000   Min.    : 0.000
##  1st Qu.: 0.840   1st Qu.: 0.00   1st Qu.: 0.000   1st Qu.: 0.000
##  Median : 1.130   Median : 5.50   Median : 0.000   Median : 1.070
##  Mean   : 1.868   Mean   :17.55   Mean   : 1.763   Mean   : 7.182
##  3rd Qu.: 1.900   3rd Qu.:22.70   3rd Qu.: 2.100   3rd Qu.: 6.370
##  Max.   :99.800   Max.   :99.98   Max.   :53.200   Max.   :99.800
##   Calcium_(mg)        Iron_(mg)        Magnesium_(mg)    Phosphorus_(mg)
```

```
##  Min.   :   0.00   Min.   : 0.000   Min.   :  0.00   Min.   :   0.0
##  1st Qu.:  11.00   1st Qu.: 0.445   1st Qu.: 12.00   1st Qu.:  50.0
##  Median :  19.00   Median : 1.080   Median : 21.00   Median : 152.0
##  Mean   :  76.16   Mean   : 1.890   Mean   : 33.58   Mean   : 177.0
##  3rd Qu.:  56.00   3rd Qu.: 2.340   3rd Qu.: 28.00   3rd Qu.: 221.5
##  Max.   :7364.00   Max.   :89.800   Max.   :711.00   Max.   :9918.0
##  Potassium_(mg)     Sodium_(mg)      Zinc_(mg)       Copper_(mg)
##  Min.   :    0.0   Min.   :    0   Min.   : 0.000   Min.   :0.0000
##  1st Qu.:  135.0   1st Qu.:   41   1st Qu.: 0.350   1st Qu.:0.0490
##  Median :  241.0   Median :   76   Median : 1.000   Median :0.0810
##  Mean   :  298.6   Mean   :  339   Mean   : 2.185   Mean   :0.1716
##  3rd Qu.:  339.0   3rd Qu.:  349   3rd Qu.: 3.100   3rd Qu.:0.1370
##  Max.   :16500.0   Max.   :38758   Max.   :90.950   Max.   :6.1000
##  Manganese_(mg)     Selenium_(µg)       Vit_C_(mg)        Thiamin_(mg)
##  Min.   :  0.0000   Min.   :   0.00   Min.   :   0.000   Min.   : 0.0000
##  1st Qu.:  0.0130   1st Qu.:   1.30   1st Qu.:   0.000   1st Qu.: 0.0360
##  Median :  0.0550   Median :  13.00   Median :   0.000   Median : 0.0710
##  Mean   :  0.4587   Mean   :  17.08   Mean   :   6.356   Mean   : 0.1753
##  3rd Qu.:  0.2800   3rd Qu.:  26.85   3rd Qu.:   1.800   3rd Qu.: 0.1725
##  Max.   :125.0000   Max.   :1917.00   Max.   :1900.000   Max.   :10.9900
##  Riboflavin_(mg)  Niacin_(mg)     Panto_Acid_(mg)    Vit_B6_(mg)
##  Min.   :0.000   Min.   : 0.000   Min.   : 0.0000   Min.   :0.0000
##  1st Qu.:0.057   1st Qu.: 0.438   1st Qu.: 0.1970   1st Qu.:0.0500
##  Median :0.175   Median : 2.486   Median : 0.4330   Median :0.1520
##  Mean   :0.212   Mean   : 3.317   Mean   : 0.5864   Mean   :0.2565
##  3rd Qu.:0.265   3rd Qu.: 5.279   3rd Qu.: 0.7320   3rd Qu.:0.4035
##  Max.   :4.000   Max.   :45.249   Max.   :22.6240   Max.   :5.8990
##  Folate_Tot_(µg)   Folic_Acid_(µg)    Food_Folate_(µg)  Folate_DFE_(µg)
##  Min.   :   0.00   Min.   :   0.000   Min.   :   0.00   Min.   :   0.0
##  1st Qu.:   3.00   1st Qu.:   0.000   1st Qu.:   3.00   1st Qu.:   3.0
##  Median :   9.00   Median :   0.000   Median :   8.00   Median :   9.0
##  Mean   :  31.08   Mean   :   9.052   Mean   :  22.01   Mean   :  37.4
##  3rd Qu.:  28.00   3rd Qu.:   0.000   3rd Qu.:  20.00   3rd Qu.:  31.0
##  Max.   :2340.00   Max.   :1286.000   Max.   :2340.00   Max.   :2340.0
##  Choline_Tot_(mg)   Vit_B12_(µg)       Vit_A_IU         Vit_A_RAE
##  Min.   :   0.00   Min.   : 0.000   Min.   :    0.0   Min.   :   0.00
##  1st Qu.:  10.60   1st Qu.: 0.000   1st Qu.:    1.0   1st Qu.:   0.00
##  Median :  29.10   Median : 0.200   Median :   16.0   Median :   3.00
##  Mean   :  46.63   Mean   : 0.961   Mean   :  499.5   Mean   :  48.21
##  3rd Qu.:  70.45   3rd Qu.: 1.390   3rd Qu.:  142.5   3rd Qu.:  22.00
##  Max.   :2403.30   Max.   :20.090   Max.   :49254.0   Max.   :4220.00
##   Retinol_(µg)     Alpha_Carot_(µg) Beta_Carot_(µg)   Beta_Crypt_(µg)
##  Min.   :   0.00   Min.   :   0.0   Min.   :    0.0   Min.   :   0.00
##  1st Qu.:   0.00   1st Qu.:   0.0   1st Qu.:    0.0   1st Qu.:   0.00
##  Median :   0.00   Median :   0.0   Median :    0.0   Median :   0.00
##  Mean   :  27.89   Mean   :  26.8   Mean   :  225.2   Mean   :  10.92
##  3rd Qu.:   8.00   3rd Qu.:   0.0   3rd Qu.:   12.0   3rd Qu.:   0.00
##  Max.   :4220.00   Max.   :4342.0   Max.   :26162.0   Max.   :6186.00
##  Lycopene_(µg)      Lut+Zea_(µg)       Vit_E_(mg)        Vit_D_µg
##  Min.   :    0.0   Min.   :    0.0   Min.   : 0.000   Min.   : 0.0000
##  1st Qu.:    0.0   1st Qu.:    0.0   1st Qu.: 0.110   1st Qu.: 0.0000
##  Median :    0.0   Median :    0.0   Median : 0.250   Median : 0.0000
##  Mean   :  162.4   Mean   :  173.9   Mean   : 0.999   Mean   : 0.5171
##  3rd Qu.:    0.0   3rd Qu.:   26.5   3rd Qu.: 0.630   3rd Qu.: 0.2000
```

```
##  Max.   :45902.0   Max.   :19697.0   Max.   :41.080   Max.   :31.9000
##     Vit_D_IU         Vit_K_(µg)        FA_Sat_(g)      FA_Mono_(g)
##  Min.   :   0.00   Min.   :   0.00   Min.   : 0.000   Min.   : 0.0000
##  1st Qu.:   0.00   1st Qu.:   0.10   1st Qu.: 0.136   1st Qu.: 0.1215
##  Median :   0.00   Median :   1.50   Median : 1.617   Median : 1.8820
##  Mean   :  20.68   Mean   :  13.27   Mean   : 3.396   Mean   : 3.9440
##  3rd Qu.:   7.00   3rd Qu.:   4.25   3rd Qu.: 4.249   3rd Qu.: 5.2865
##  Max.   :1276.00   Max.   :1714.50   Max.   :86.503   Max.   :83.6890
##   FA_Poly_(g)     Cholestrl_(mg)
##  Min.   : 0.000   Min.   :   0.0
##  1st Qu.: 0.164   1st Qu.:   0.0
##  Median : 0.559   Median :   8.0
##  Mean   : 1.910   Mean   :  42.2
##  3rd Qu.: 1.619   3rd Qu.:  73.0
##  Max.   :57.740   Max.   :3100.0
```

Figure 1

CONCLUSION

It can be overwhelming when encountering a database such as the United States Department of Agriculture Food Composition Database with data on 8,789 food items and up to 150 food components. Principal Component Analysis enabled us to make unique observations that we might not be able to make otherwise.

In order to narrow down the data, I settled on making observations based on the first six principal components. The most prominet nutrients in these first six components were Water, Energy, Total Lipids, Protein, Magnesium, and Carbohydrates. Taking a closer look at these nutrients, it was evident that foods with higher levels of water tended to have lower levels of other nutrients.

SOURCES

1) US Department of Agriculture, Agricultural Research Service, Nutrient Data Laboratory. USDA National Nutrient Database for Standard Reference, Release 28 (Slightly revised). Version Current: May 2016. Internet: http://www.ars.usda.gov/ba/bhnrc/ndl

2) Holland, S. (2019). *Principal Components Analysis (PCA)*. Athens, GA.

3) James, G et al. (2017). *An Introduction To Statistical Learning*. New York, NY.

APPENDIX

R Code

```
#Provided code to clean up the data

SR = read.table("ABBREV.txt", header=F, row.names=1, sep="^", quote="~")
SR = na.omit(SR) # remove rows with missing values
SR = SR[row.names(SR) != "13352",] # remove "duplicate" entry
row.names(SR) = SR[,1] # set more meaningful row names
SR = SR[,-1]
names(SR) = c("Water_(g)", "Energ_Kcal", "Protein_(g)", "Lipid_Tot_(g)", "Ash_(g)",
              "Carbohydrt_(g)", "Fiber_TD_(g)", "Sugar_Tot_(g)", "Calcium_(mg)",
              "Iron_(mg)", "Magnesium_(mg)", "Phosphorus_(mg)", "Potassium_(mg)",
```

```
              "Sodium_(mg)", "Zinc_(mg)", "Copper_(mg)", "Manganese_(mg)",
              "Selenium_(µg)", "Vit_C_(mg)", "Thiamin_(mg)", "Riboflavin_(mg)",
              "Niacin_(mg)", "Panto_Acid_(mg)", "Vit_B6_(mg)", "Folate_Tot_(µg)",
              "Folic_Acid_(µg)", "Food_Folate_(µg)", "Folate_DFE_(µg)",
              "Choline_Tot_(mg)", "Vit_B12_(µg)", "Vit_A_IU", "Vit_A_RAE",
              "Retinol_(µg)", "Alpha_Carot_(µg)", "Beta_Carot_(µg)", "Beta_Crypt_(µg)",
              "Lycopene_(µg)", "Lut+Zea_(µg)", "Vit_E_(mg)", "Vit_D_µg", "Vit_D_IU",
              "Vit_K_(µg)", "FA_Sat_(g)", "FA_Mono_(g)", "FA_Poly_(g)", "Cholestrl_(mg)",
              "GmWt_1", "GmWt_Desc1", "GmWt_2", "GmWt_Desc2", "Refuse_Pct")
SRp = SR[,c(1:46)] # restrict to just the nutrient variables

# Running the PCA

pr.out <- prcomp(SRp, scale = T)

#Scree Plot

library(factoextra)
fviz_eig(pr.out, ncp = 15, choice = c("variance", "eigenvalue"), addlabels = T)

# Looking at the most prevalent nutrients in each component.

library(cowplot)

One <- fviz_contrib(pr.out, choice = "var", axes = 1, top = 10, title ="PC 1")

Two <- fviz_contrib(pr.out, choice = "var", axes = 2, top = 10, title ="PC 2")

Three <- fviz_contrib(pr.out, choice = "var", axes = 3, top = 10, title ="PC 3")

Four <- fviz_contrib(pr.out, choice = "var", axes = 4, top = 10, title ="PC 4")

Five <- fviz_contrib(pr.out, choice = "var", axes = 5, top = 10, title ="PC 5")

Six <- fviz_contrib(pr.out, choice = "var", axes = 6, top = 10, title ="PC 6")

Individual.plot <- plot_grid(One, Two, Three, Four, Five, Six, align = 'h')

title <- ggdraw() +
  draw_label(
    "Contribution of Variables to the First Six Principal Components",
    x = 0,
    hjust = 0, , fontfamily = "serif", size = 16
  ) +
  theme(
    plot.margin = margin(0, 0, 0, 7)
  )
plot_grid(
  title, Individual.plot,
  ncol = 1,
  rel_heights = c(0.1, 1)
)
```

```r
# Determing the most prevalent variables in the first six PCs
library(reshape2)

pr.var =pr.out$sdev^2
pve=pr.var/sum(pr.var)
pve.mat = matrix(rep(pve, each = 6), nrow=length(pve))

contribution = apply(pr.out$rotation[,c(1:6)]^2 * pve.mat, 1, sum)
new.contribution <- melt(contribution)

# Creating the new categories

Categories <- c(rep.int("Proximates", 1), rep.int("Energy", 1), rep.int("Proximates", 6),
                rep.int("Minerals", 10), rep.int("Vitamins", 24), rep.int("Lipids",4))
new.contrib.cat <- data.frame(cbind(new.contribution, Categories))
new.contrib.cat

# Plotting the most prominent variables

ggplot(data=new.contrib.cat, aes(fill=Categories)) +
   geom_col(aes(x=reorder(rownames(new.contrib.cat), -value), y=value)) +
   theme(axis.text.x = element_text(angle = 90)) +
   labs(x="Nutrients", y="Variable Prominence") +
   ggtitle("Most Prominent Nutrients in the First Six Principal Components")

# Creating Biplot

fviz_pca_biplot(pr.out, labels = FALSE)


#Creating Individuals Plots

library(cowplot)

Water <- fviz_pca_ind(pr.out, label = "none", col.ind = SRp$`Water_(g)`,
                      title = "Water")

Energy <- fviz_pca_ind(pr.out, label = "none", col.ind = SRp$Energ_Kcal,
                       title = "Energy")

Lipid <- fviz_pca_ind(pr.out, label = "none", col.ind = SRp$'Lipid_Tot_(g)',
                      title = "Total Lipids")

Protein <- fviz_pca_ind(pr.out, label = "none", col.ind = SRp$'Protein_(g)',
                        title = "Protein")

Magnesium <- fviz_pca_ind(pr.out, label = "none", col.ind = SRp$'Magnesium_(mg)',
                          title = "Magnesium")

Carbohydrates <- fviz_pca_ind(pr.out, label = "none", col.ind = SRp$'Carbohydrt_(g)',
                              title = "Carbohydrates")

Individual.plot <- plot_grid(Water, Energy, Lipid, Protein, Magnesium, Carbohydrates,
```

```r
                             align = 'h')

title <- ggdraw() +
  draw_label(
    "Individiual Plots- First and Second Component",
    x = 0,
    hjust = 0, , fontfamily = "serif", size = 16
  ) +
  theme(
    plot.margin = margin(0, 0, 0, 7)
  )
plot_grid(
  title, Individual.plot,
  ncol = 1,
  rel_heights = c(0.1, 1)
)

#Creating Individuals Plots Looking at the Second and Third Component

library(cowplot)

Water <- fviz_pca_ind(pr.out, axes = c(2, 3), label = "none",
                      col.ind = SRp$`Water_(g)`, title = "Water")

Energy <- fviz_pca_ind(pr.out, axes = c(2, 3), label = "none",
                       col.ind = SRp$Energ_Kcal, title = "Energy")

Lipid <- fviz_pca_ind(pr.out, axes = c(2, 3), label = "none",
                      col.ind = SRp$'Lipid_Tot_(g)', title = "Total Lipids")

Protein <- fviz_pca_ind(pr.out, axes = c(2, 3), label = "none",
                        col.ind = SRp$'Protein_(g)', title = "Protein")

Magnesium <- fviz_pca_ind(pr.out, axes = c(2, 3), label = "none",
                          col.ind = SRp$'Magnesium_(mg)', title = "Magnesium")

Carbohydrates <- fviz_pca_ind(pr.out, axes = c(2, 3), label = "none",
                              col.ind = SRp$'Carbohydrt_(g)', title = "Carbohydrates")

Individual.plot <- plot_grid(Water, Energy, Lipid, Protein, Magnesium, Carbohydrates,
                             align = 'h')

title <- ggdraw() +
  draw_label(
    "Individiual Plots- Second and Third Component",
    x = 0,
    hjust = 0, , fontfamily = "serif", size = 16
  ) +
  theme(
    plot.margin = margin(0, 0, 0, 7)
  )
plot_grid(
  title, Individual.plot,
```

```
  ncol = 1,
  rel_heights = c(0.1, 1)
)

#Creating Variables Plot

fviz_pca_var(pr.out, labels = FALSE, habillage = new.contrib.cat$Categories)

#Examining the Summary of the Initial Data Set

summary(SRp)
```