paper\_mat<-indata[ ,1:4]  
paper\_mat

## BL EM SF BS  
## 1 21.312 7.039 5.326 0.932  
## 2 21.206 6.979 5.237 0.871  
## 3 20.709 6.779 5.060 0.742  
## 4 19.542 6.601 4.479 0.513  
## 5 20.449 6.795 4.912 0.577  
## 6 20.841 6.919 5.108 0.784  
## 7 19.060 6.447 4.246 0.358  
## 8 18.597 6.261 4.032 0.215  
## 9 19.346 6.572 4.358 0.432  
## 10 18.720 6.455 4.072 0.372  
## 11 18.587 6.295 4.068 0.239  
## 12 19.813 6.775 4.604 0.637  
## 13 19.989 6.737 4.686 0.779  
## 14 19.116 6.512 4.299 0.588  
## 15 18.769 6.335 4.089 0.470  
## 16 18.708 6.271 3.978 0.457  
## 17 19.323 6.550 4.404 0.588  
## 18 17.433 5.948 3.486 0.104  
## 19 19.195 6.213 4.300 0.405  
## 20 19.436 6.387 4.404 0.519  
## 21 20.136 6.725 4.723 0.652  
## 22 16.740 6.168 3.201 0.104  
## 23 18.589 6.531 3.989 0.336  
## 24 19.422 6.615 4.382 0.432  
## 25 24.420 7.874 6.999 1.730  
## 26 25.288 8.034 7.406 1.873  
## 27 26.119 8.222 7.771 1.946  
## 28 23.113 7.288 6.329 1.513  
## 29 25.209 7.955 7.296 1.792  
## 30 25.444 8.045 7.477 1.847  
## 31 23.699 7.593 6.609 1.482  
## 32 24.303 7.775 6.861 1.583  
## 33 24.793 8.123 7.202 1.703  
## 34 23.438 7.650 6.457 1.477  
## 35 24.197 7.794 6.833 1.583  
## 36 24.741 7.996 7.152 1.728  
## 37 24.170 7.766 6.846 1.615  
## 38 24.174 7.877 6.826 1.692  
## 39 25.052 8.287 7.332 1.773  
## 40 23.846 7.639 6.615 1.560  
## 41 24.822 8.041 7.129 1.721  
## 42 25.200 8.270 7.356 1.785  
## 43 23.695 7.460 6.567 1.543  
## 44 24.941 7.929 7.286 1.703  
## 45 25.007 8.081 7.287 1.787  
## 46 21.183 7.156 5.388 0.924  
## 47 21.875 7.336 5.762 1.068  
## 48 22.095 7.447 5.790 1.182  
## 49 25.166 7.913 7.211 1.813  
## 50 24.560 7.854 7.020 1.701  
## 51 22.007 8.259 7.322 1.169  
## 52 21.115 7.913 6.557 0.928  
## 53 26.194 8.454 7.816 2.145  
## 54 25.674 8.208 7.534 2.046  
## 55 25.930 8.100 7.669 2.037  
## 56 21.390 7.475 5.294 0.875  
## 57 18.441 6.652 3.946 0.140  
## 58 16.441 6.315 2.997 -0.400  
## 59 16.294 6.572 3.017 -0.478  
## 60 20.289 7.719 4.866 0.239  
## 61 17.163 7.086 3.396 -0.236  
## 62 20.289 7.437 4.859 0.470

# Calculate the mean vector  
mean\_vec = colMeans(paper\_mat)  
mean\_vec

## BL EM SF BS   
## 21.7228 7.2662 5.6375 1.0188

# Calculate the variance vector  
var\_vec = apply(paper\_mat, 2, var)  
var\_vec

## BL EM SF BS   
## 8.30287 0.51336 2.14005 0.48027

# Another way of getting the covariance matrix  
covMat = cov(paper\_mat)  
covMat

## BL EM SF BS  
## BL 8.3029 1.88664 4.14732 1.97206  
## EM 1.8866 0.51336 0.98759 0.43431  
## SF 4.1473 0.98759 2.14005 0.98797  
## BS 1.9721 0.43431 0.98797 0.48027

# Calculate the correlation matrix  
corMat = cor(paper\_mat)  
corMat

## BL EM SF BS  
## BL 1.00000 0.91383 0.98388 0.98756  
## EM 0.91383 1.00000 0.94222 0.87467  
## SF 0.98388 0.94222 1.00000 0.97451  
## BS 0.98756 0.87467 0.97451 1.00000

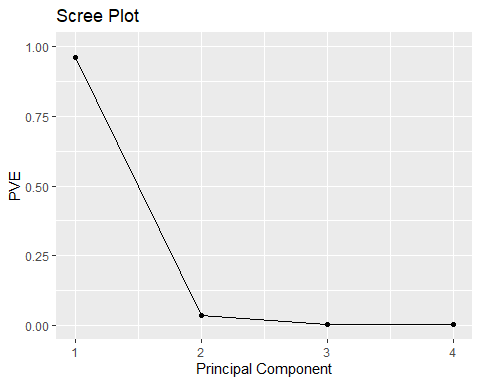
# Scaled data  
scaled\_indata = scale(paper\_mat)  
eigenRes = eigen(corMat)  
eigenRes

## eigen() decomposition  
## $values  
## [1] 3.8395058 0.1403040 0.0126039 0.0075863  
##   
## $vectors  
## [,1] [,2] [,3] [,4]  
## [1,] -0.50617 -0.261102 0.565177 0.59682  
## [2,] -0.48549 0.819048 0.193505 -0.23667  
## [3,] -0.50807 -0.020209 -0.800196 0.31803  
## [4,] -0.49996 -0.510468 0.053073 -0.69760

PVE <- eigenRes$values / sum(eigenRes$values)  
# Percent variance explained  
PVE

## [1] 0.9598764 0.0350760 0.0031510 0.0018966

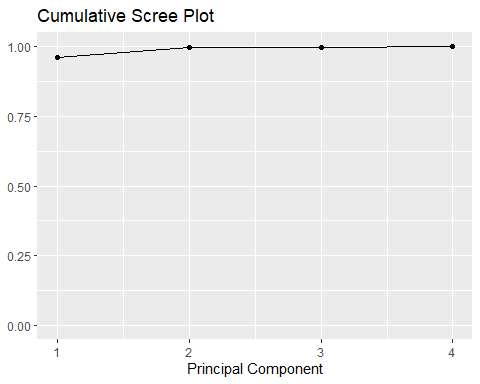
# PVE (aka scree) plot  
PVEplot <- qplot(c(1:4), PVE) +   
 geom\_line() +   
 xlab("Principal Component") +   
 ylab("PVE") +  
 ggtitle("Scree Plot") +  
 ylim(0, 1)  
PVEplot



#Cumulative percent variance explained  
cumsum(PVE)

## [1] 0.95988 0.99495 0.99810 1.00000

# Cumulative PVE plot  
cumPVE <- qplot(c(1:4), cumsum(PVE)) +   
 geom\_line() +   
 xlab("Principal Component") +   
 ylab(NULL) +   
 ggtitle("Cumulative Scree Plot") +  
 ylim(0,1)  
  
cumPVE



evecs = -eigenRes$vectors[,1:2]  
colnames(evecs) = c("PC1", "PC2")  
row.names(evecs) = colnames(scaled\_indata)  
evecs

## PC1 PC2  
## BL 0.50617 0.261102  
## EM 0.48549 -0.819048  
## SF 0.50807 0.020209  
## BS 0.49996 0.510468

PC1 <- as.matrix(scaled\_indata)%\*% evecs[,1]  
PC2 <- as.matrix(scaled\_indata)%\*% evecs[,2]  
PC1

## [,1]  
## [1,] -0.39690  
## [2,] -0.53109  
## [3,] -0.90845  
## [4,] -1.60105  
## [5,] -1.11372  
## [6,] -0.74343  
## [7,] -1.98282  
## [8,] -2.36767  
## [9,] -1.75559  
## [10,] -2.08745  
## [11,] -2.31657  
## [12,] -1.30268  
## [13,] -1.16659  
## [14,] -1.74460  
## [15,] -2.08355  
## [16,] -2.18556  
## [17,] -1.64602  
## [18,] -3.05393  
## [19,] -2.06500  
## [20,] -1.78640  
## [21,] -1.22767  
## [22,] -3.12558  
## [23,] -2.11376  
## [24,] -1.70477  
## [25,] 1.87159  
## [26,] 2.37700  
## [27,] 2.82979  
## [28,] 0.85569  
## [29,] 2.21295  
## [30,] 2.41776  
## [31,] 1.24017  
## [32,] 1.62998  
## [33,] 2.15686  
## [34,] 1.17655  
## [35,] 1.61451  
## [36,] 2.06234  
## [37,] 1.61839  
## [38,] 1.74291  
## [39,] 2.40913  
## [40,] 1.35552  
## [41,] 2.09402  
## [42,] 2.44060  
## [43,] 1.17877  
## [44,] 2.08058  
## [45,] 2.25611  
## [46,] -0.32452  
## [47,] 0.15278  
## [48,] 0.35861  
## [49,] 2.16257  
## [50,] 1.86900  
## [51,] 1.41605  
## [52,] 0.58536  
## [53,] 3.15936  
## [54,] 2.73197  
## [55,] 2.74415  
## [56,] -0.14000  
## [57,] -2.21410  
## [58,] -3.51294  
## [59,] -3.41395  
## [60,] -0.77554  
## [61,] -2.60680  
## [62,] -0.80241

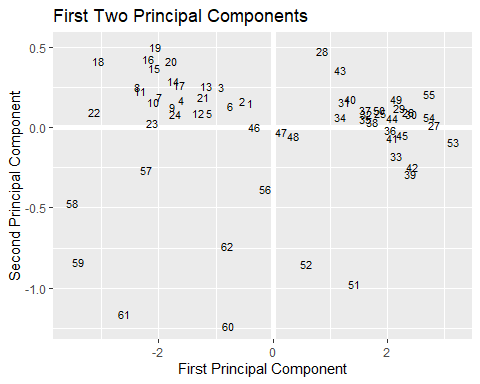
PC <- data.frame(Week = row.names(indata), PC1, PC2)  
head(PC)

## Week PC1 PC2  
## 1 1 -0.39690 0.154256  
## 2 2 -0.53109 0.167078  
## 3 3 -0.90845 0.253205  
## 4 4 -1.60105 0.174232  
## 5 5 -1.11372 0.087773  
## 6 6 -0.74343 0.136726

{r components description} #The number of components to effectively summarize the variability is one.ie..PC1 #As both in PVE and cumPVE,the bent or elbow is observed at 2nd component.

{PCA analysis} #PCA has been applied and found useful in very many disciplines. #The first two PCs account for 95.99% and 3.51%, respectively, of the total variation in the datasets #All of the loadings in the first PC is a weighted average of all variables

ggplot(PC, aes(PC1, PC2)) +   
 modelr::geom\_ref\_line(h = 0) +  
 modelr::geom\_ref\_line(v = 0) +  
 geom\_text(aes(label = Week), size = 3) +  
 xlab("First Principal Component") +   
 ylab("Second Principal Component") +   
 ggtitle("First Two Principal Components")



summary(evecs)

## PC1 PC2   
## Min. :0.485 Min. :-0.81905   
## 1st Qu.:0.496 1st Qu.:-0.18961   
## Median :0.503 Median : 0.14066   
## Mean :0.500 Mean :-0.00682   
## 3rd Qu.:0.507 3rd Qu.: 0.32344   
## Max. :0.508 Max. : 0.51047

{r strength index} The proportion of variance explained and the scree plot(PVEplot and CVE) depict that first principal component PC1 has equal variable load and strong corelation between variable and first principal component,hence it high strength index .

r{outliers} #As per graph , 60 and 61 seems to be little distant but not potential outliers.