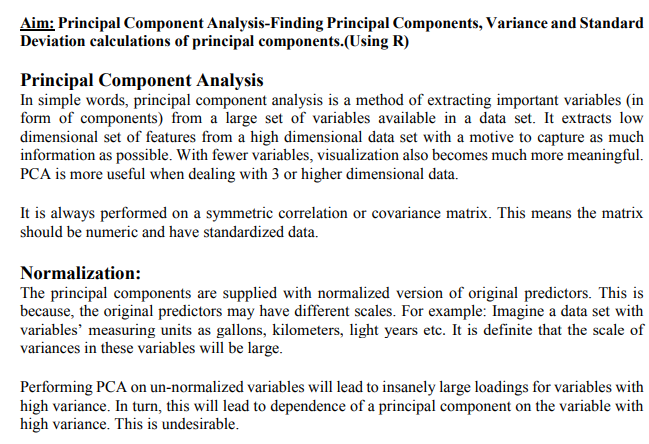
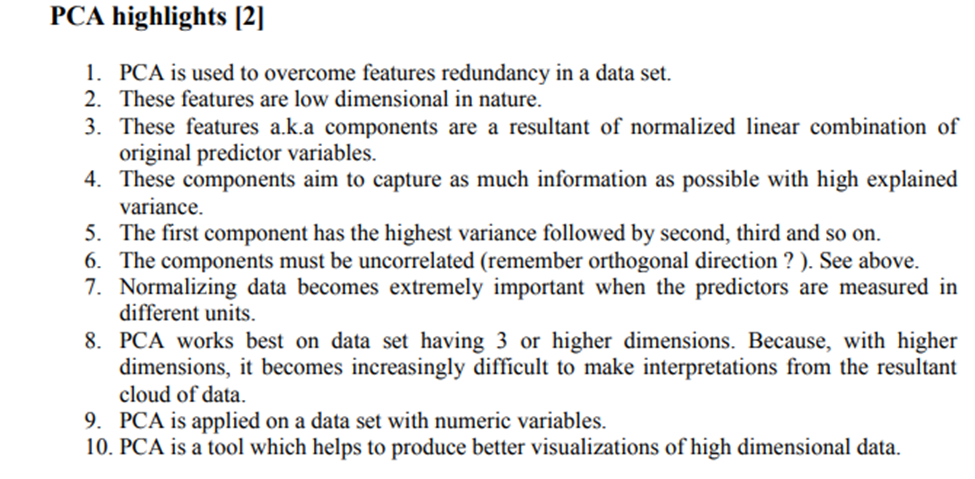
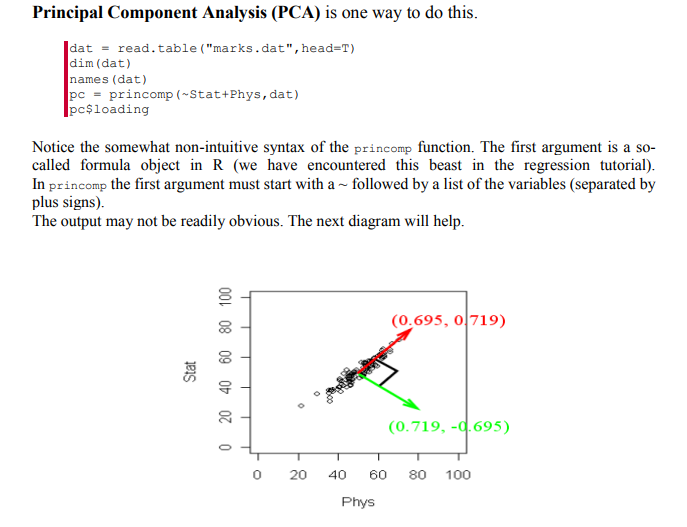
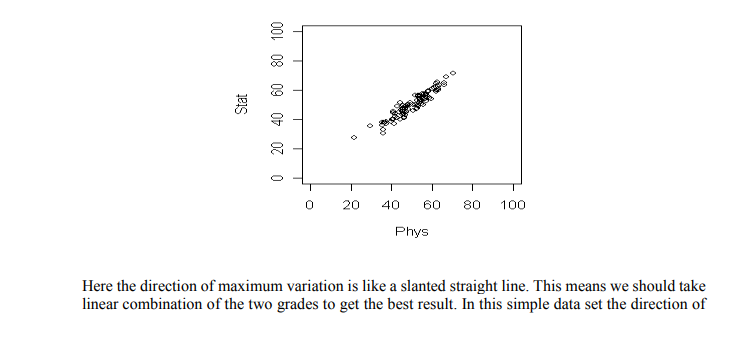
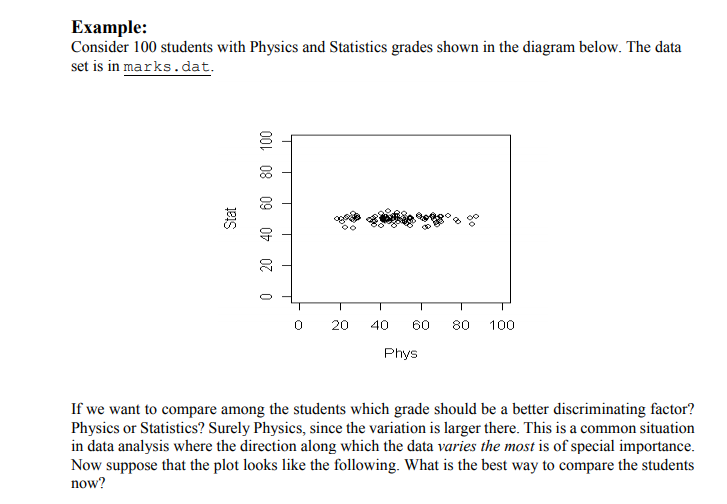
**Name:** Aditya Somani **Roll No:** BE1851061 **PRN:** 71901204L







R has returned two principal components. (Two because we have two variables). These are a unit vector at right angles to each other. You may think of PCA as choosing a new coordinate system for the data, the principal components being the unit vectors along the axes. The first principal component gives the direction of the maximum spread of the data. The second gives the direction of maximum spread perpendicular to the first direction. These two directions are packed inside the matrix pc$loadings. Each column gives a direction. The direction of maximum spread (the first principal component) is in the first column, the next principal component in the second and so on.

Conclusion:

Thus we have found principal components, calculated the mean and variance of principal components.

**CODE:**

**> mydata<-read.csv("pca.csv")**

**> attach(mydata)**

**> names(mydata)**

[1] "calory" "breakfast" "lunch" "dinner" "exercise"

[6] "body\_shape"

**> X <- cbind(calory, breakfast, lunch, dinner, exercise, body\_shape)**

**> X <- mydata[1:10,1:5]**

**> summary(X)**

calory breakfast lunch dinner exercise

Min. :1200 Min. :0.00 Min. :0.00 Min. :0.00 Min. :0.00

1st Qu.:1700 1st Qu.:0.25 1st Qu.:0.25 1st Qu.:0.25 1st Qu.:0.00

Median :2900 Median :1.00 Median :1.50 Median :1.00 Median :1.00

Mean :2780 Mean :1.00 Mean :1.20 Mean :0.90 Mean :1.00

3rd Qu.:3375 3rd Qu.:1.75 3rd Qu.:2.00 3rd Qu.:1.00 3rd Qu.:1.75

Max. :5000 Max. :2.00 Max. :2.00 Max. :2.00 Max. :3.00

**> cor(X)**

calory breakfast lunch dinner exercise

calory 1.0000000 0.7954175 0.6910415 0.8043890 -0.8386183

breakfast 0.7954175 1.0000000 0.4442617 0.7377111 -0.6454972

lunch 0.6910415 0.4442617 1.0000000 0.3605104 -0.6882472

dinner 0.8043890 0.7377111 0.3605104 1.0000000 -0.5714286

exercise -0.8386183 -0.6454972 -0.6882472 -0.5714286 1.0000000

**> pcal<-princomp(X, scores=TRUE, cor=TRUE)**

**> summary(pcal)**

Importance of components:

Comp.1 Comp.2 Comp.3 Comp.4 Comp.5

Standard deviation 1.9128933 0.8688388 0.51717050 0.49982709 0.26204168

Proportion of Variance 0.7318322 0.1509762 0.05349307 0.04996542 0.01373317

Cumulative Proportion 0.7318322 0.8828083 0.93630141 0.98626683 1.00000000

**> loadings(pcal)**

Loadings:

Comp.1 Comp.2 Comp.3 Comp.4 Comp.5

calory 0.508 0.102 0.116 0.847

breakfast 0.447 0.369 -0.500 -0.631 -0.128

lunch 0.384 -0.708 0.382 -0.403 -0.207

dinner 0.428 0.532 0.590 0.222 -0.369

exercise -0.460 0.282 0.496 -0.613 0.295

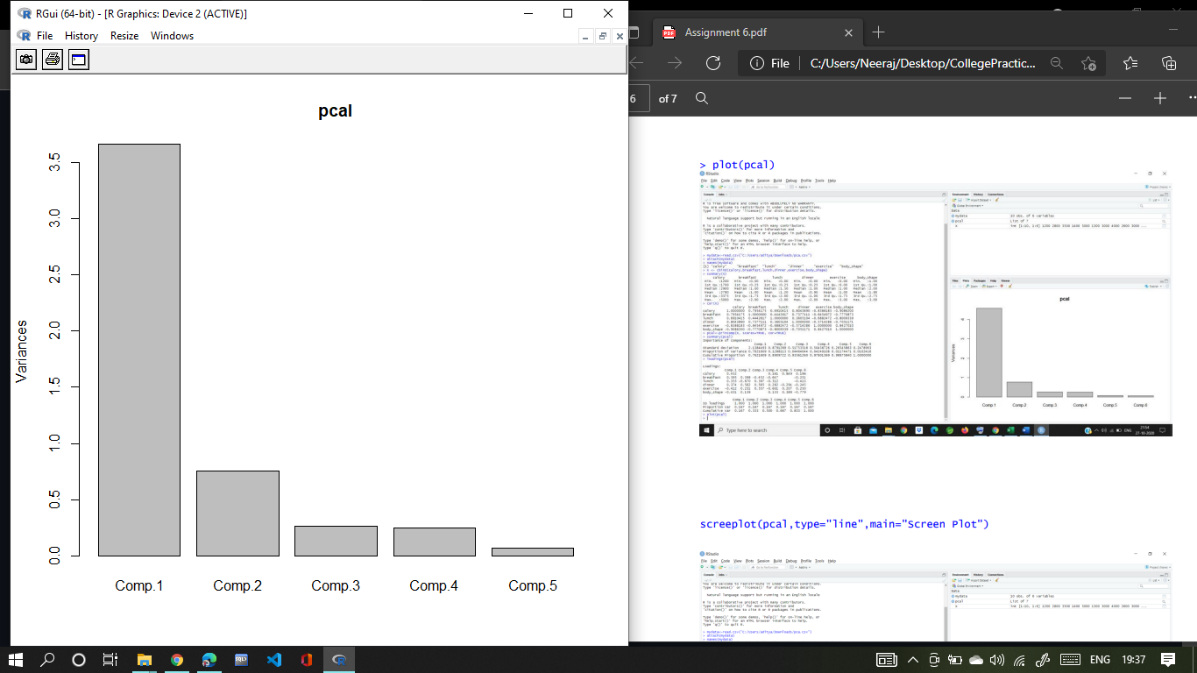
Comp.1 Comp.2 Comp.3 Comp.4 Comp.5

SS loadings 1.0 1.0 1.0 1.0 1.0

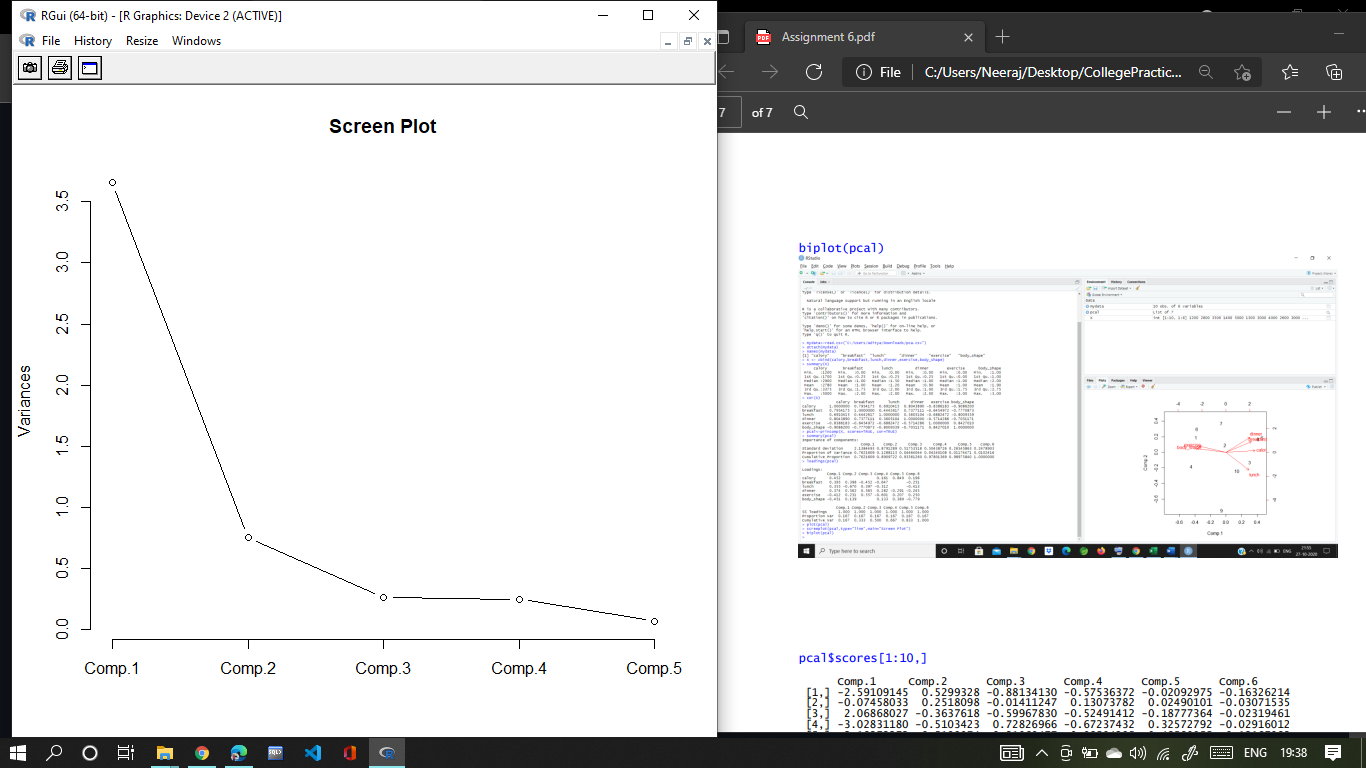
Proportion Var 0.2 0.2 0.2 0.2 0.2

Cumulative Var 0.2 0.4 0.6 0.8 1.0

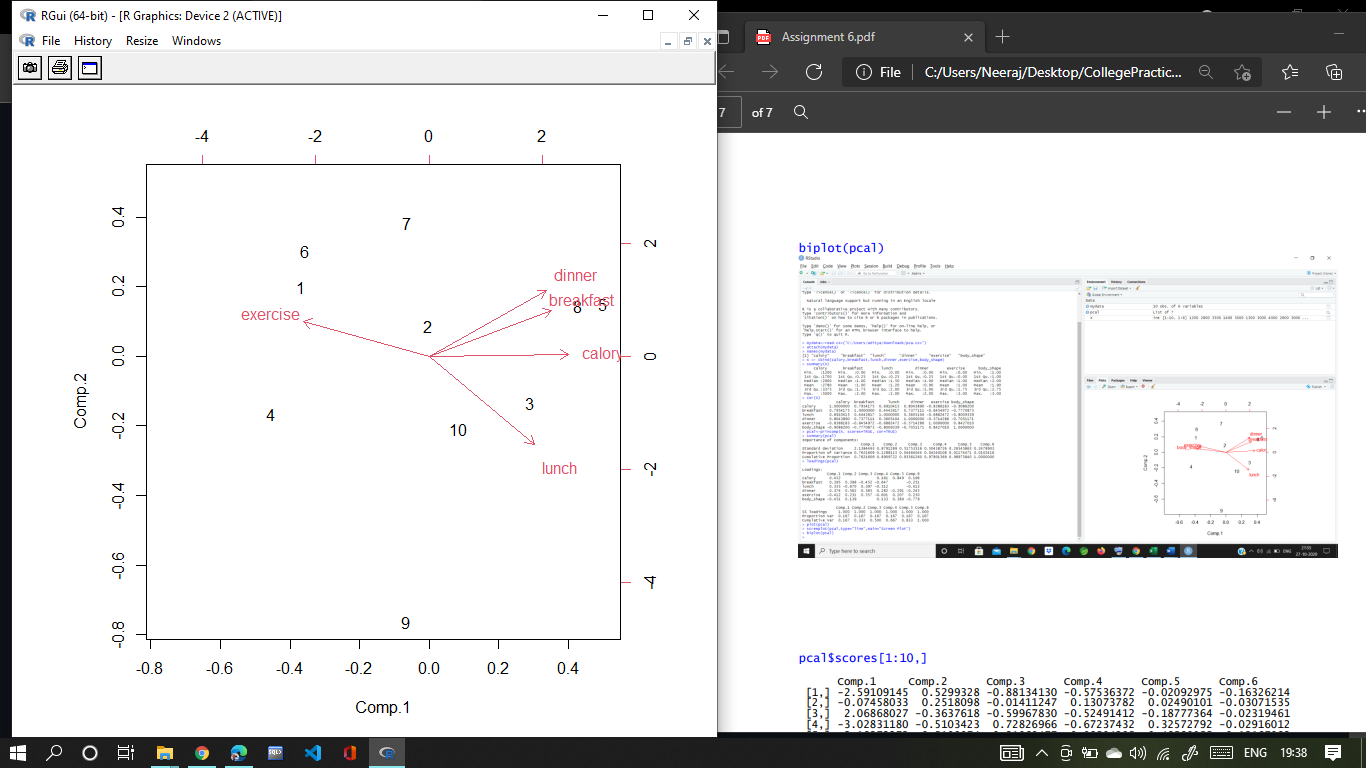
**> plot(pcal)**

****

**> screeplot(pcal, type="line", main="Screen Plot")**

****

**> biplot(pcal)**

****

**> pcal$scores[1:10,]**

Comp.1 Comp.2 Comp.3 Comp.4 Comp.5

1 -2.2260094 0.5493651 -0.926748204 -0.5017618788 -0.091168759

2 -0.0181432 0.2387190 -0.001736852 0.1262129315 0.009165519

3 1.7629661 -0.3686868 -0.642992171 -0.4677611479 -0.180484240

4 -2.7354241 -0.4544983 0.670516979 -0.7428230654 0.276439242

5 3.0271154 0.4130506 0.330717058 -0.0008315232 0.379090547

6 -2.1470258 0.8344911 0.569382701 0.6405158598 -0.381107428

7 -0.3714247 1.0538059 -0.422756474 0.6084633748 0.391373597

8 2.5923988 0.3985085 0.243047350 -0.1004204621 -0.345812836

9 -0.3934781 -2.0942040 -0.273781041 0.7546078475 0.025585563

10 0.5090250 -0.5705512 0.454350653 -0.3162019362 -0.083081206