 # Sudip Sen

Fruit Characteristics and Yield of brinjal plant

Table 1 demonstrates that treatment has a highly significant effect on fruit weight, diameter, and length indicating that different treatments lead to distinct differences in every variable.
The Duncan test further clarifies these differences, grouping treatments based on their mean fruit weights. Some treatments (e.g., T7) yield significantly higher fruit weights compared to others (e.g., T12). Again, certain treatments produce significantly larger fruit sizes (both diameter and length), with T7 consistently producing the largest fruits in both measures.

Table 1. Fruit Size Attributes and Yield

Treatment	Fruit length (cm)	Fruit diameter (cm)	Fruit weight (gm)
T_1	14.5 ± 0.057e	3.3 ± 0.057f	38.7 ± 0.057e
T_2	14.9 ± 0.057d	3.9 ± 0.057b	42.1 ± 0.057c
T_3	13.4 ± 0.057f	3.5 ± 0.057e	36.2 ± 0.057h
T_4	12.1 ± 0.057h	2.8 ± 0.057h	29.1 ± 0.057j
T_5	15.6 ± 0.057b	3.7 ± 0.057cd	44.5 ± 0.057b
T_6	15.4 ± 0.057c	3.6 ± 0.057de	37.5 ± 0.057g
T_7	16.4 ± 0.057a	4.5 ± 0.057a	47.8 ± 0.057a
T_8	14.8 ± 0.057d	3.7 ± 0.057cd	39.8 ± 0.057d
T_9	15.4 ± 0.057c	3.8 ± 0.057bc	37.9 ± 0.057f
T_{10}	13.5 ± 0.057f	3.1 ± 0.057g	32.5 ± 0.057i
T_{11}	12.7 ± 0.057g	2.8 ± 0.057h	27.8 ± 0.057k
T_{12}	11.8 ± 0.057i	2.1 ± 0.057i	26.1 ± 0.057l

Total 12 treatments are indicated by T1, T2….T12. Different letters in the same column shows extremely significantly difference ($p < 2e\text{-}16$). Data represents mean values of three replicates ± SE (Standard error)

Correlation plot:

Figure 1 displays the correlation coefficients between various variables, ranging from -1 (strong negative correlation) indicated by blue to +1 (strong positive correlation) indicated by red. Fruit weight, diameter, and length show a strongly positive correlation. Fruit weight and "Total chlorophyll" appear to have a strong positive correlation, indicating that as fruit weight increases, total chlorophyll content also tends to increase. To exemplify a negative correlation "Relative water content" and "MDA" (malondialdehyde) might show a negative relationship, indicating that as relative water content decreases, MDA levels increase.

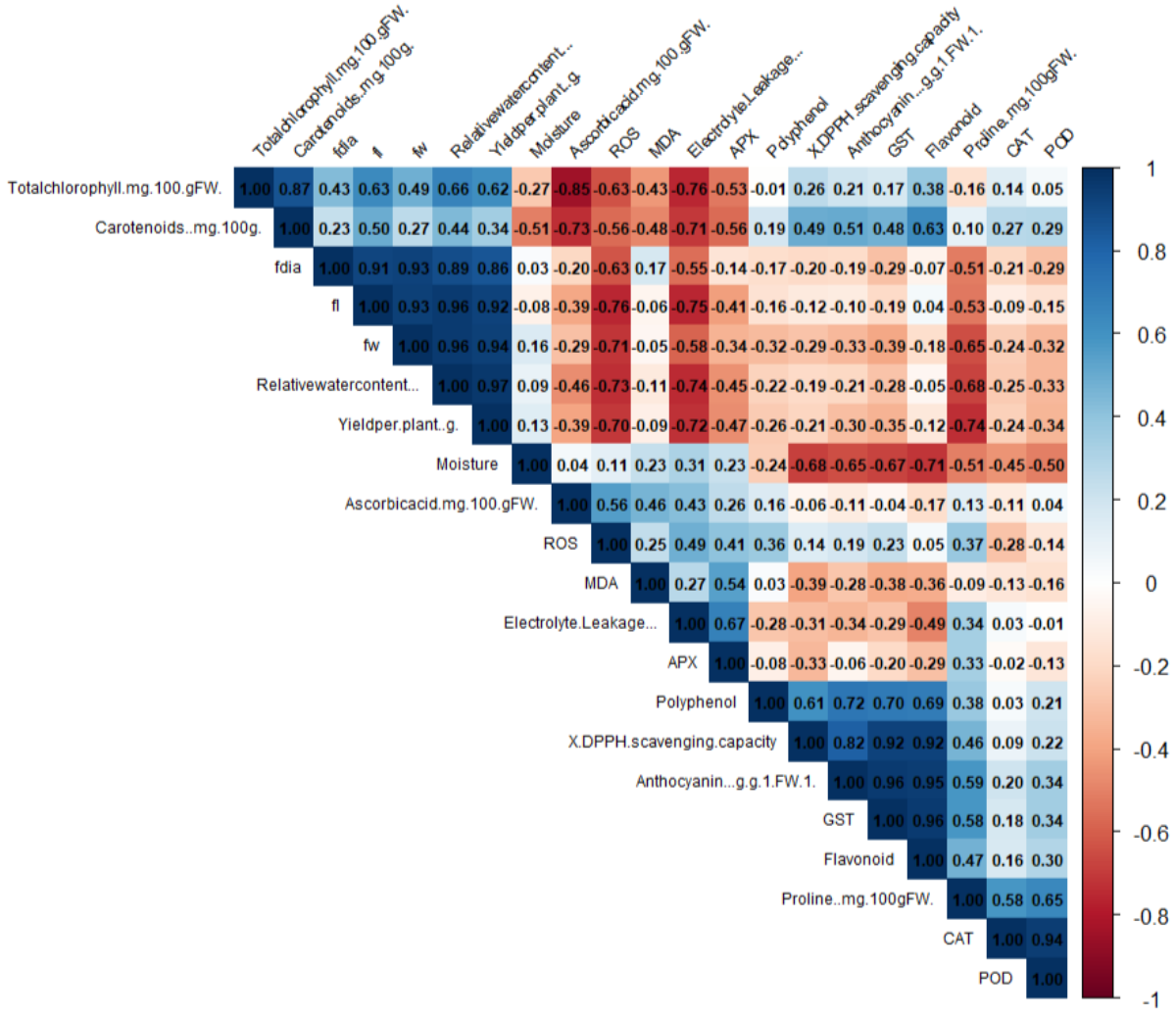

Figure 1: Correlation Plot

Principle component analysis:

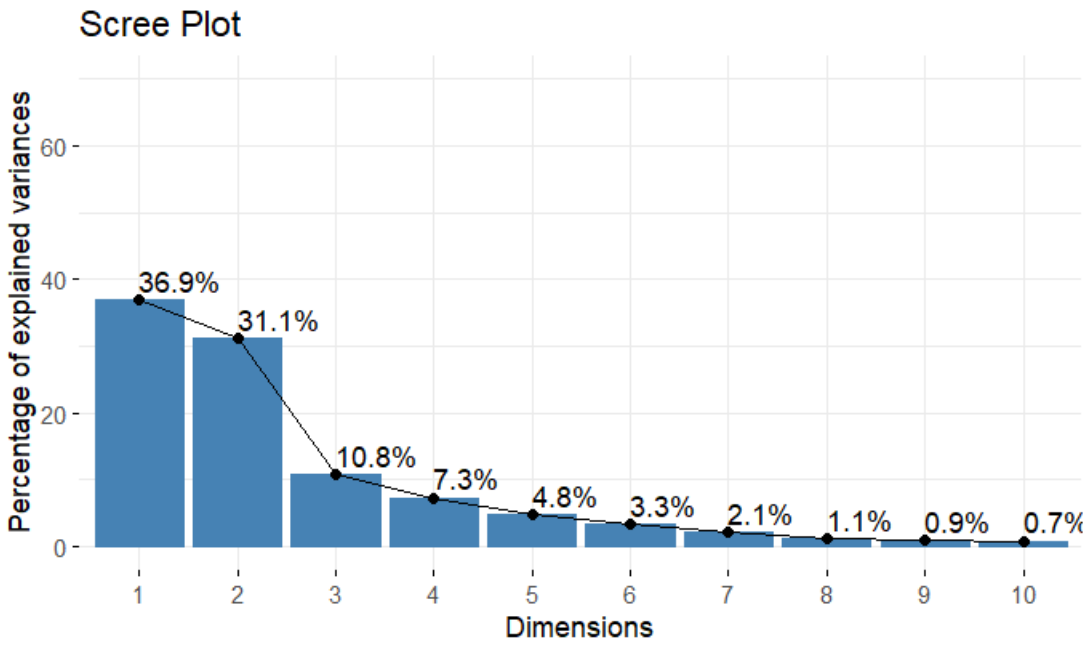

Figure 2: Scree plot

The scree plot (Figure 2) shows that the first three principal components explain 67.99% of total variations. Therefore, working with three principal components may loss of 32.01% of the information.

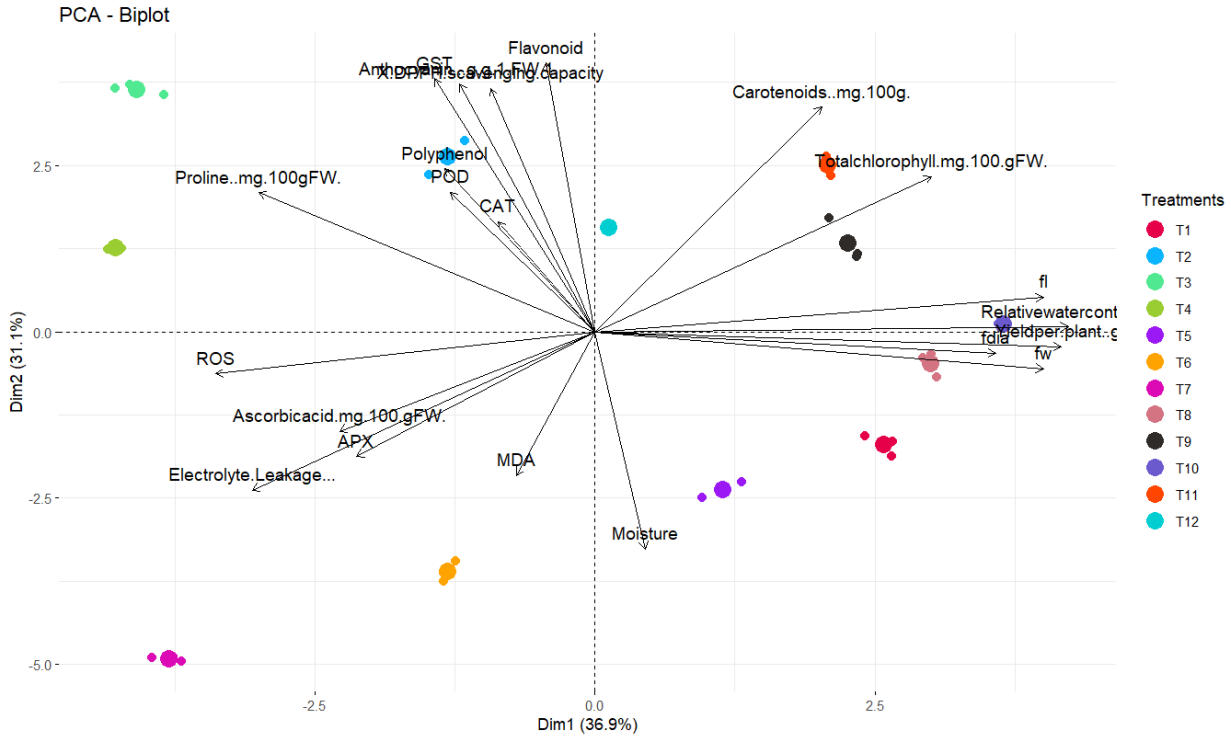

Figure 3: PCA Biplot

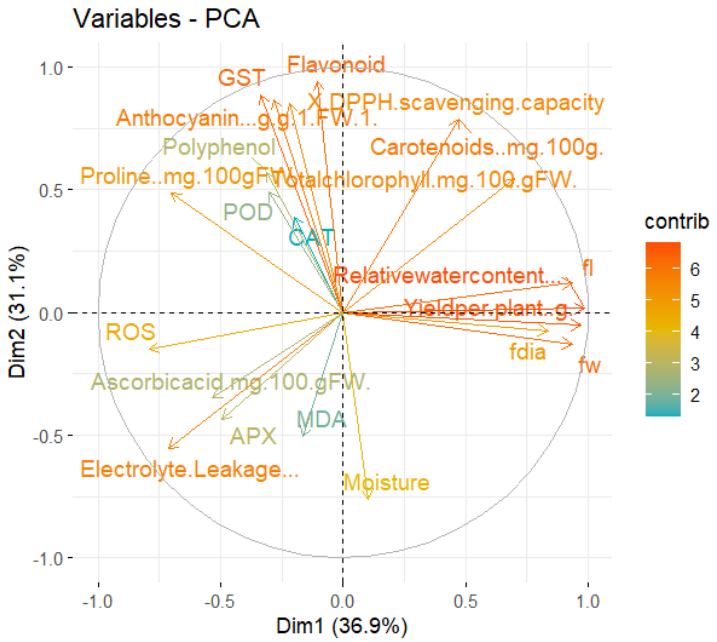

Figure 4: Variable contribution in PCA 1 and PCA 2

In Figure 3, we can see the PCA biplot indicating the variables and treatments. The arrow toward the treatment indicates the higher value of the respective treatments. For example, Treatment 10 shows a higher yield. Treatments 9 and 11 show higher chlorophyll values and less MDA, APX, and Ascorbic acid.

In Figure 4, those variables that are blue colored (MDA, CAT, etc.), contribute less in PC1 and PC2. Again, those with warmer colors contribute more in PC1 and PC2. Yield per plant, Relative water content, fw, fl, fdia, etc for Dim 1 and Flavonoid, GST, Anthocyanin, etc. for Dim 2.

Dendrogram:

In Figure 5, samples in the same cluster are more similar to each other than to samples in different clusters. For example, samples 36 and 12 in the red cluster are closely related, as are samples 37 and 29 in the cyan cluster. In this hierarchical clustering, we have clustered the total 36 samples in 4 broad groups based on their similarities in respect of their values in variables.

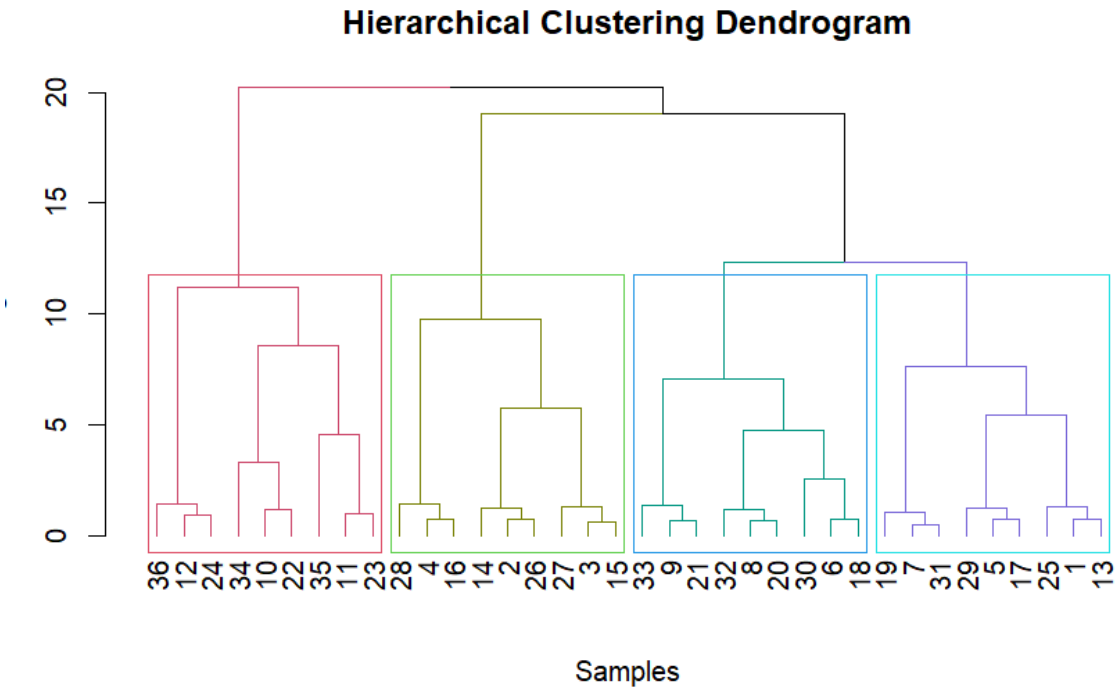

Figure 5: Dendrogram

Scripts:

```
setwd("F:/R for Agricultural Research/JHassan_R_Training_24/")
getwd()
datass <- read.csv(file.choose(), header = TRUE)# Raw data for assignment
str(datass)
library(dplyr)
library(agricolae)
library(ggplot2)
library(ggpubr)

# Select columns by names
data <- datass %>% select(Replication,Treatment,fw, fdia,fl)

str(data)
data$Replication=as.factor(data$Replication)
data$Treatment=as.factor(data$Treatment)
attach(data)
str(data)

# ANOVA RCBD
# FW= Fruit weight
FWanova <- aov(fw ~ Replication + Treatment)
summary(FWanova)

duncan.FW = duncan.test(y = fw,
                        trt = Treatment,
                        DFerror = FWanova$df.residual,
                        MSerror = deviance(FWanova)/FWanova$df.residual,
                        group = T,
```

```
                console = T)


FW_SE = data %>%
 group_by(Treatment) %>%
 summarise(GR.mean = mean(fw),
        std = sd(fw),
        SE = sd(fw)/sqrt(n()))

FW_SE
print(FW_SE)

###### ANOVA ########
# Fdia= Fruit diameter
Fdanova <- aov(fdia ~ Replication + Treatment)
summary(Fdanova)

duncan.Fd = duncan.test(y = fdia,
                trt = Treatment,
                DFerror = Fdanova$df.residual,
                MSerror = deviance(Fdanova)/Fdanova$df.residual,
                group = T,
                console = T)


Fd_SE = data %>%
 group_by(Treatment) %>%
 summarise(GR.mean = mean(fdia),
        std = sd(fdia),
```

```
        SE = sd(fdia)/sqrt(n()))

Fd_SE
print(Fd_SE)



#####
###### ANOVA ########
# Fl= Fruit length
Flanova <- aov(fl ~ Replication + Treatment)
summary(Flanova)

duncan.Fl = duncan.test(y = fl,
                        trt = Treatment,
                        DFerror = Flanova$df.residual,
                        MSerror = deviance(Flanova)/Flanova$df.residual,
                        group = T,
                        console = T)



Fl_SE = data %>%
  group_by(Treatment) %>%
  summarise(GR.mean = mean(fl),
        std = sd(fl),
        SE = sd(fl)/sqrt(n()))

Fl_SE
print(Fl_SE)
```

```
##### Correlation Plot ######


# Select only the dependent variables (excluding Replication and Treatment)
dependent_vars <- datass[, c("Totalchlorophyll.mg.100.gFW.", "Carotenoids..mg.100g.",
                             "Proline..mg.100gFW.", "Electrolyte.Leakage...",
                             "Relativewatercontent...", "fw", "fdia", "fl",
                             "Yieldper.plant..g.", "Ascorbicacid.mg.100.gFW.",
                             "MDA", "ROS", "CAT", "APX", "POD", "GST",
                             "Moisture", "Anthocyanin...g.g.1.FW.1.",
                             "X.DPPH.scavenging.capacity", "Polyphenol", "Flavonoid")]

# Create correlation matrix
cor_matrix <- cor(dependent_vars, method = "pearson")

library(corrplot)

# Create correlation plot
corrplot(cor_matrix, method = "color", type = "upper",
         order = "hclust", tl.col = "black", tl.srt = 45,
         addCoef.col = "black", number.cex = 0.6,
         tl.cex = 0.6) # Reduced text size for better fit


####### PCA #############
# Load required libraries
library(ggplot2)
```

```
library(factoextra)
library(dendextend)
library(FactoMineR)

# Data preparation
numeric_data <- datass[, -c(1,2)]  # Remove Replication and Treatment columns
scaled_data <- scale(numeric_data) # Scaling data
# ============= PCA VISUALIZATION =============
# Perform PCA
pca_result <- PCA(numeric_data, scale.unit = TRUE, graph = FALSE)

# 1. Scree plot
fviz_eig(pca_result,
         addlabels = TRUE,
         ylim = c(0, 70),
         main = "Scree Plot",
         barcolor = "steelblue",
         barfill = "steelblue")

# 2. PCA Biplot
col_unique<-as.character(datass$Treatment)
fviz_pca_biplot( pca_result, geom.ind="point",
                 pointsize = 2.5, col.var="black",
                 col.ind=col_unique,pointshape=19,  )+

  scale_color_manual(name="Treatments", labels= unique(col_unique),
                     values= c("#e60049", "#0bb4ff", "#50e991", "#9ACD32",
                               "#9b19f5", "#ffa300", "#dc0ab4", "#D47382",
                               "#2e2b28", "#6A5ACD", "#FF4500", "#00CED1"))
```

```
# COntribution of variables to DIm 1 and Dim 2
fviz_contrib(pca_result, choice = "var", axes = 1, top = 21)
fviz_contrib(pca_result, choice = "var", axes = 2, top = 21)

fviz_pca_var(pca_result,
             col.var="contrib",gradient.cols=c("#00AFBB","#E7B800","#FC4E07"),
             repel=TRUE)
pca_result$eig



# ============= DENDROGRAM VISUALIZATION =============
# Calculate distance matrix
dist_matrix <- dist(scaled_data, method = "euclidean")

# Hierarchical clustering
hc <- hclust(dist_matrix, method = "ward.D2")

# Create and color dendrogram
dend <- as.dendrogram(hc)
dend_colored <- color_branches(dend, k = 4)  # Change k for different number of clusters

# Plot dendrogram
par(mar = c(5,3,3,3))
plot(dend_colored,
    main = "Hierarchical Clustering Dendrogram",
    ylab = "Height",
```

```
    sub = "",
    xlab = "Samples")

# Add rectangles to show clusters
rect.hclust(hc, k = 4, border = 2:5)  # Change k as needed
```