

Greg Finkelberg: HW05

R packages

```
library(MASS)
library(qgraph)
```

Loading and preparing data

```
load("C:/Users/Greg/Downloads/tobacco_clr.Rdata")

microbes_data <- tobacco_clr$data[, 1:10]

new_col_names <- c("OTU1", "OTU2", "OTU3", "OTU4", "OTU5", "OTU6", "OTU7", "OTU8", "OTU9", "OTU10")

colnames(microbes_data) <- new_col_names
```

I changed the column names in the dataset so they would be easier to read when I constructed the network later on.

PCA analysis and covariance matrix

```
pca_result <- princomp(microbes_data, cor = TRUE)

biplot(pca_result, scale = 0)
```

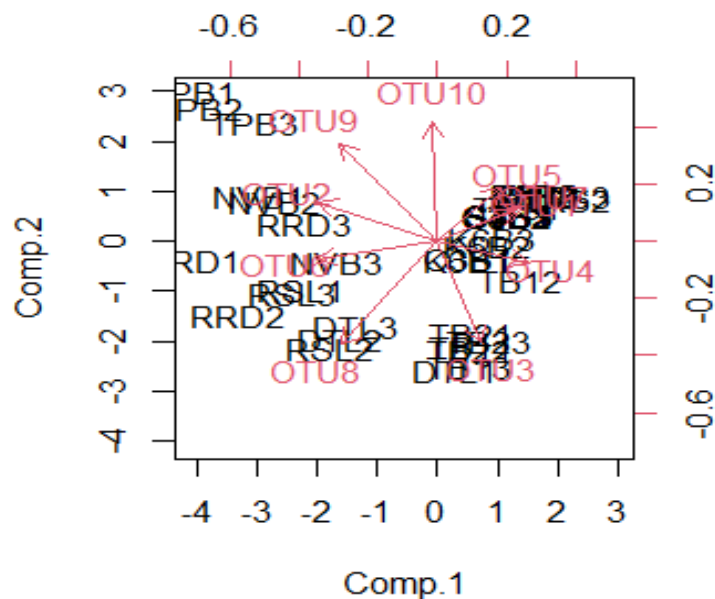


Figure 1.

The biplot in Figure 1 shows where each variable falls in components 1 and 2 of the PCA analysis. Variables that fall on the positive end of an axis explain a larger portion of the variation in that component.

```
cov_matrix <- cov(microbes_data) * (pca_result$sdev ^2)

heatmap(cov_matrix, col = colorRampPalette(c("blue", "white", "red"))(100), m
ain = "Covariance Matrix Heatmap", xlab = "Microbes", ylab = "Microbes")
```

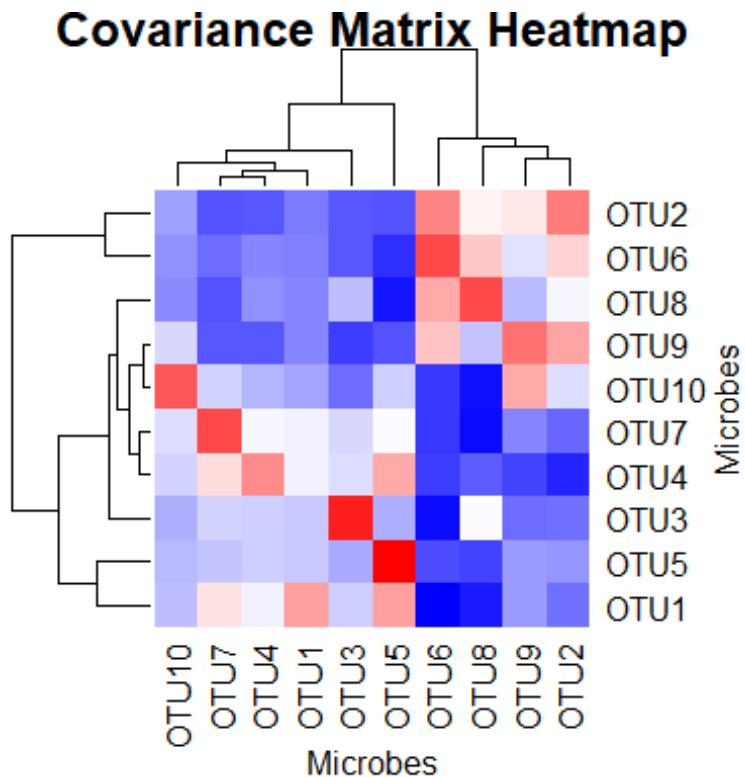


Figure 2.

The heatmap in Figure 2 shows the strength of covariance relationships between variables. Darker squares indicate stronger relationships. Blue squares indicate negative relationships, while red ones represent positive relationships.

```
ggm_result <- qgraph(cov_matrix, layout = "spring")
```

