Using Breast Cancer Data Set -II

Load Dataset

```
table = readtable("dataR2.csv");
disp(table(1:5, :));
```

```
    Age     BMI      Glucose    Insulin     HOMA      Leptin    Adiponectin    Resistin    MCP_1     Classification
    ___    ______    _______    _______    _______    ______    ___________    ________    ______    ______________

    48       23.5      70        2.707     0.46741    8.8071      9.7024        7.9958     417.11          1       
    83      20.69      92        3.115      0.7069    8.8438      5.4293         4.064     468.79          1       
    82     23.125      91        4.498      1.0097    17.939      22.432        9.2772      554.7          1       
    68     21.368      77        3.226     0.61272    9.8827      7.1696        12.766     928.22          1       
    86     21.111      92        3.549     0.80539    6.6994      4.8192        10.576     773.92          1       
```

Getting numeric values from the table

```
tablenum = table2array(table(:, 1:end-1)); %classification is excluded since it is
categorical value

%displaying first five records of numeric values
disp(tablenum(1:5, :));
```

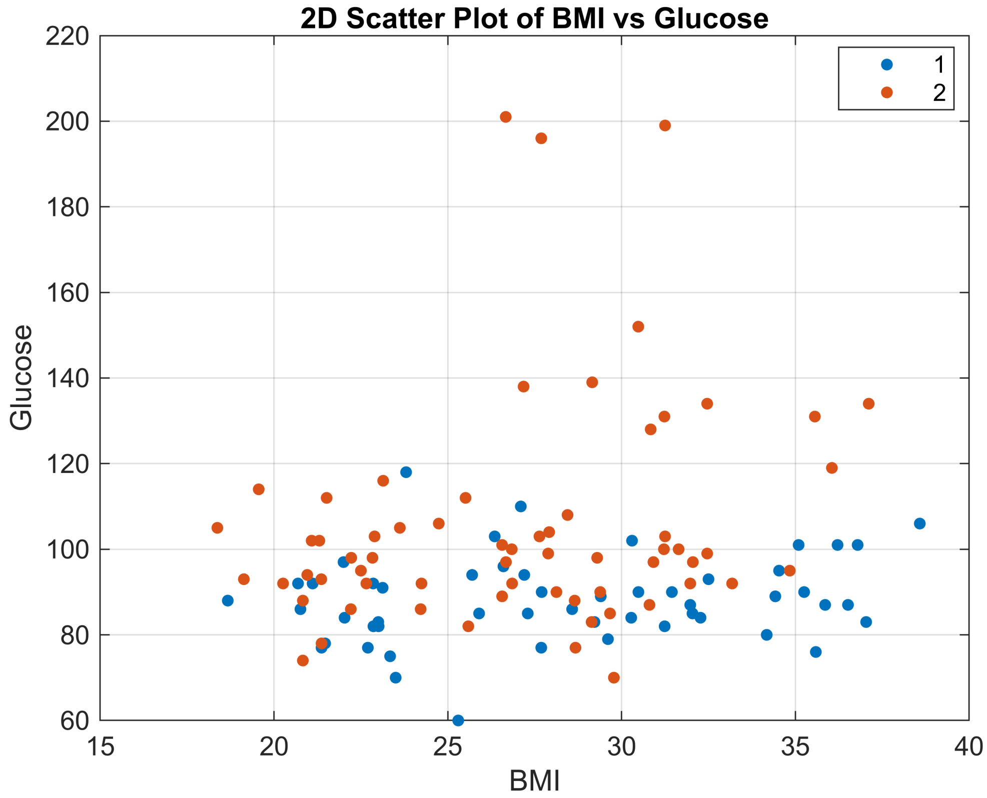
```
   48.0000   23.5000   70.0000    2.7070    0.4674    8.8071    9.7024    7.9958  417.1140
   83.0000   20.6905   92.0000    3.1150    0.7069    8.8438    5.4293    4.0640  468.7860
   82.0000   23.1247   91.0000    4.4980    1.0097   17.9393   22.4320    9.2772  554.6970
   68.0000   21.3675   77.0000    3.2260    0.6127    9.8827    7.1696   12.7660  928.2200
   86.0000   21.1111   92.0000    3.5490    0.8054    6.6994    4.8192   10.5763  773.9200
```

```

% Find minimum and maximum values for each attribute
tab_min = min(tablenum);
tab_max = max(tablenum);

% Compute the range for each attribute
temp = tab_max - tab_min;

% Initialize the normalized feature matrix
n_tab = zeros(size(tablenum));

% Perform min-max normalization for each attribute
for i = 1 : size(tablenum, 2)
    n_tab(:, i) = 10*((tablenum(:, i) - tab_min(i)) ./ temp(i));
end
```

2D Scatter Plot

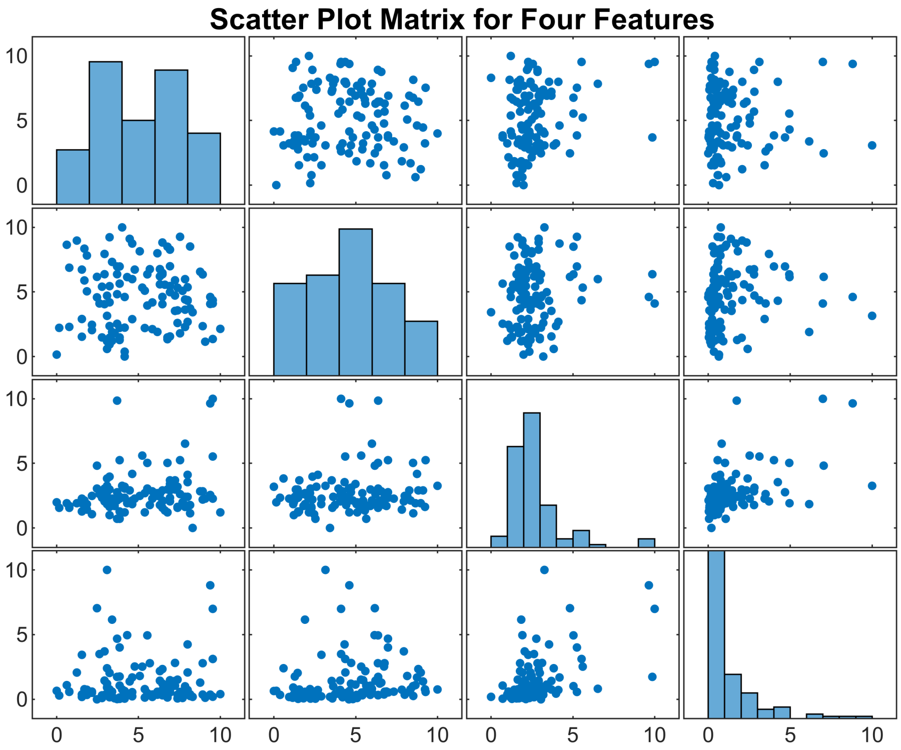
```
figure;
gscatter(table.BMI, table.Glucose, table.Classification);
xlabel('BMI');
ylabel('Glucose');
title('2D Scatter Plot of BMI vs Glucose');
```

```
grid on;
```



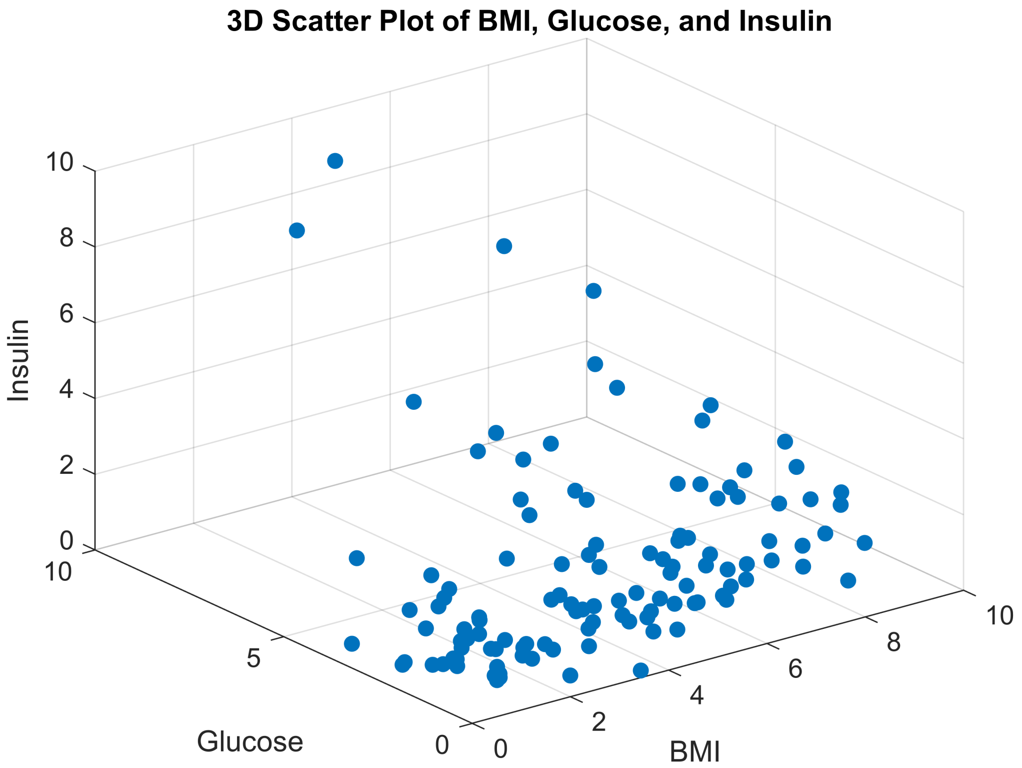
Matrix of Scatter Plots (at least four features)

```
selected_features = n_tab(:, 1:4);  % First four features

% Generate a matrix of scatter plots
figure;
plotmatrix(selected_features);
title('Scatter Plot Matrix for Four Features');
```



3D Scatter Plot

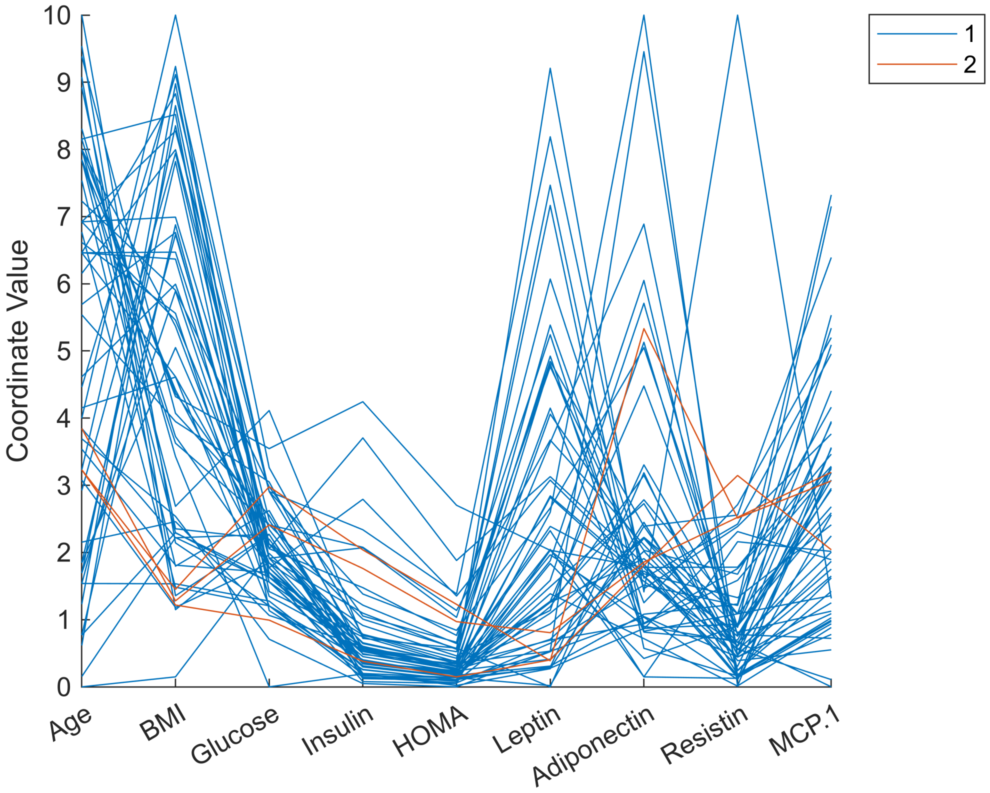
```
figure;
scatter3(n_tab(:, 2), n_tab(:, 3), n_tab(:, 4), 'filled');  % BMI, Glucose, and
Insulin
xlabel('BMI');
ylabel('Glucose');
zlabel('Insulin');
title('3D Scatter Plot of BMI, Glucose, and Insulin');
grid on;
```



Parallel Coordinates

only displaying 55 samples to get a clear view

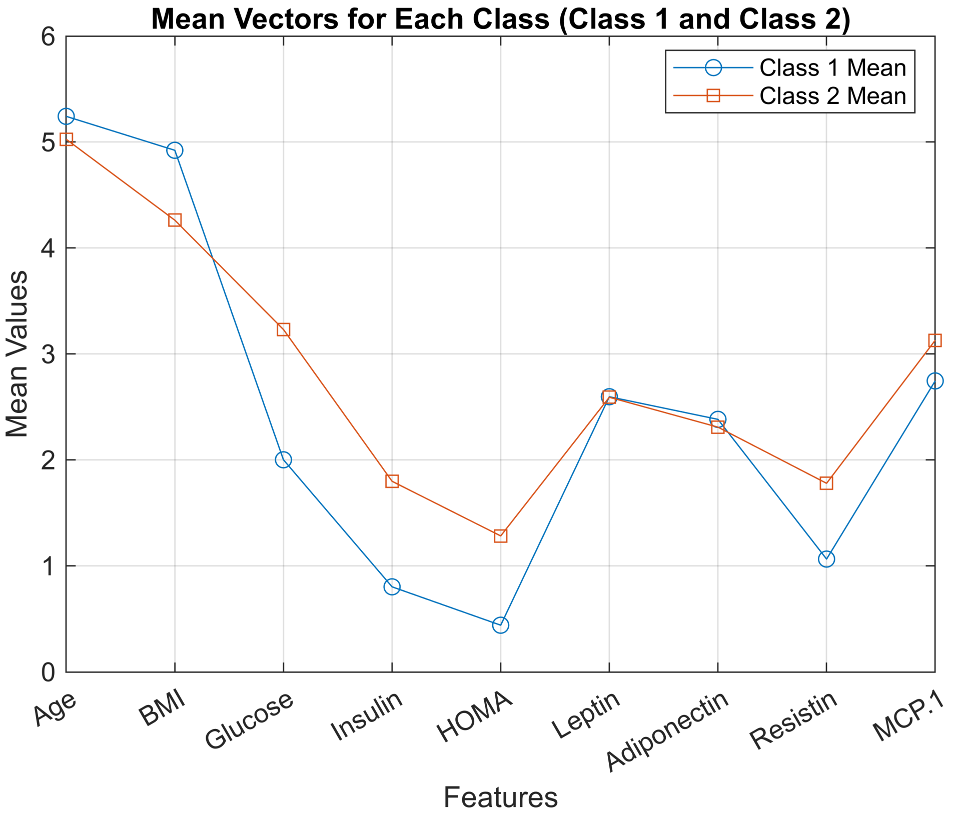
```
p_tab = n_tab(1:55, :);
species = table.Classification(1:55,:);
labels = {'Age', 'BMI', 'Glucose', 'Insulin', 'HOMA', 'Leptin', 'Adiponectin',
'Resistin', 'MCP.1'};
parallelcoords(p_tab, 'Group', species, 'Labels', labels);
```



Mean Vectors

```
% Separate data by class using 'table.Classification' for labels
class1 = n_tab(table.Classification == 1, :);  % Class 1 data
class2 = n_tab(table.Classification == 2, :);  % Class 2 data

% Finding mean vectors for each class
mclass1 = mean(class1, 1);  % Mean vector for class 1
mclass2 = mean(class2, 1);  % Mean vector for class 2

% Plot the mean vectors for each class
figure;
plot(mclass1, '-o', 'DisplayName', 'Class 1 Mean'); hold on;
plot(mclass2, '-s', 'DisplayName', 'Class 2 Mean'); hold off;

% Define your custom labels for the x-axis
labels = {'Age', 'BMI', 'Glucose', 'Insulin', 'HOMA', 'Leptin', 'Adiponectin',
'Resistin', 'MCP.1'};

% Set x-axis ticks and labels
xticks(1:length(labels));  % Set the number of ticks to match the number of features
xticklabels(labels);  % Set the custom feature names as x-axis labels

% Add labels and title
xlabel('Features');
```

```
ylabel('Mean Values');
title('Mean Vectors for Each Class (Class 1 and Class 2)');
legend('Location', 'best');
grid on;
```


Mean Vectors for Each Class (Class 1 and Class 2)
Class 1 Mean
Class 2 Mean
Mean Values
Age
BMI
Glucose
Insulin
HOMA
Leptin
Adiponectin
Resistin
MCP.1
Features