Data Set Used:

Iris Data Set

Data Set Characteristics:	Multivariate	**Number of Instances:**	150	**Area:**	Life
Attribute Characteristics:	Real	**Number of Attributes:**	4	**Date Donated**	1988-07-01
Associated Tasks:	Classification	**Missing Values?**	No	**Number of Web Hits:**	3661711

Source: https://archive.ics.uci.edu/ml/datasets/Iris

Creator :   R.A. Fisher
Donor   :   Michael Marshall

Data Set Information:

The data set contains 3 classes of 50 instances each, where each class refers to a type of iris plant.

Predicted attribute: Three Class of iris plant.

Attribute Information:

1. Sepal length in cm
2. Sepal width in cm
3. Petal length in cm
4. Petal width in cm
5. Output Classes: - Iris Setosa   Class-1
                        - Iris Versicolour   Class-2
                        - Iris Virginica   Class-3

There are in total 50 data elements in each class, thereby making a total of 150 data rows/elements. For data analysis, only top 10 data rows from each class is used.

% Importing file from my Computer

```
close all;
clear;
clc;
T1 = readtable("C:\Users\salma\OneDrive\Desktop\Matlab_Project\Iris_Data.csv");
```

% Write the imported file in Matlab in .CSV file

```
writetable(T1,'Iris_Data.csv');
```

% Cleaning the data

```
class(T1.(1));
class(T1.(2));
class(T1.(3));
class(T1.(4));
class(T1.(5));

T2 = readtable("Iris_Data.csv","Format","%f%f%f%f%s");

T2.Properties.VariableNames = {'Sepal_Length','Sepal_width','Petal_Length','Petal_width','Class'};

B = T2.Class;
y = strings(size(B));
[y{:}] = B{:};

T2.Class = y;
```

% Saving the corrected file in .CSV file

```
writetable(T2,'Iris_Data.csv')
```

% Extracting Species into parts from the data

```
setosa = T2(T2.Class == "Iris-setosa",:);
versicolor = T2(T2.Class =="Iris-versicolor",:);
virginica = T2(T2.Class =="Iris-virginica",:);
```

*%Data analysis only top 10 species from each classes (3 * 10 = 30 Species)*

```
Iris_setosa = setosa(1:10,:);
Iris_versicolor = versicolor(1:10,:);
Iris_virginica = virginica(1:10,:);


Iris_setosa.Class = ones(10,1);
Iris_versicolor.Class = ones(10,1)*2;
Iris_virginica.Class = ones(10,1)*3;
```

OUTPUT :

Iris_setosa ✖

10x5 table

	1 Sepal_Length	2 Sepal_width	3 Petal_Length	4 Petal_width	5 Class
1	5.1000	3.5000	1.4000	0.2000	.1
2	4.9000	3	1.4000	0.2000	1
3	4.7000	3.2000	1.3000	0.2000	1
4	4.6000	3.1000	1.5000	0.2000	1
5	5	3.6000	1.4000	0.2000	1
6	5.4000	3.9000	1.7000	0.4000	1
7	4.6000	3.4000	1.4000	0.3000	1
8	5	3.4000	1.5000	0.2000	1
9	4.4000	2.9000	1.4000	0.2000	1
10	4.9000	3.1000	1.5000	0.1000	1

Iris_versicolor ✖

10x5 table

	1 Sepal_Length	2 Sepal_width	3 Petal_Length	4 Petal_width	5 Class
1	7	3.2000	4.7000	1.4000	.2
2	6.4000	3.2000	4.5000	1.5000	2
3	6.9000	3.1000	4.9000	1.5000	2
4	5.5000	2.3000	4	1.3000	2
5	6.5000	2.8000	4.6000	1.5000	2
6	5.7000	2.8000	4.5000	1.3000	2
7	6.3000	3.3000	4.7000	1.6000	2
8	4.9000	2.4000	3.3000	1	2
9	6.6000	2.9000	4.6000	1.3000	2
10	5.2000	2.7000	3.9000	1.4000	2

Iris_virginica ✕

10x5 table

	1 Sepal_Length	2 Sepal_width	3 Petal_Length	4 Petal_width	5 Class
1	6.3000	3.3000	6	2.5000	3
2	5.8000	2.7000	5.1000	1.9000	3
3	7.1000	3	5.9000	2.1000	3
4	6.3000	2.9000	5.6000	1.8000	3
5	6.5000	3	5.8000	2.2000	3
6	7.6000	3	6.6000	2.1000	3
7	4.9000	2.5000	4.5000	1.7000	3
8	7.3000	2.9000	6.3000	1.8000	3
9	6.7000	2.5000	5.8000	1.8000	3
10	7.2000	3.6000	6.1000	2.5000	3

%Finding MAX,MIN,MEAN,VAR of Attributes for each 3 Species.

```
fprintf('\nClass - Iris-Setosa')

fprintf('\n        Sepal Dim      Petal Dim')
fprintf('\n      Length Width(Cms) Length Width(Cms)')
fprintf('\nMax    : %.2f  %.2f      %.2f  %.2f',max(Iris_setosa.(1)),max(Iris_setosa.(2)),max(Iris_setosa.(3)),max(Iris_setosa.(4)))
fprintf('\nMin    : %.2f  %.2f      %.2f  %.2f',min(Iris_setosa.(1)),min(Iris_setosa.(2)),min(Iris_setosa.(3)),min(Iris_setosa.(4)))
fprintf('\nMean   : %.2f  %.2f      %.2f  %.2f',mean(Iris_setosa.(1)),mean(Iris_setosa.(2)),mean(Iris_setosa.(3)),mean(Iris_setosa.(4)))
fprintf('\nVariance : %.2f  %.2f      %.2f  %.2f',var(Iris_setosa.(1)),var(Iris_setosa.(2)),var(Iris_setosa.(3)),var(Iris_setosa.(4)))


fprintf('\n-----------------------------------------------')

fprintf('\nClass - Iris-Versicolor')

fprintf('\n        Sepal Dim      Petal Dim')
fprintf('\n      Length Width(Cms) Length Width(Cms)')
fprintf('\nMax    : %.2f  %.2f      %.2f  %.2f',max(Iris_versicolor.(1)),max(Iris_versicolor.(2)),max(Iris_versicolor.(3)),max(Iris_versicolor.(4)))
fprintf('\nMin    : %.2f  %.2f      %.2f  %.2f',min(Iris_versicolor.(1)),min(Iris_versicolor.(2)),min(Iris_versicolor.(3)),min(Iris_versicolor.(4)))
fprintf('\nMean   : %.2f  %.2f      %.2f  %.2f',mean(Iris_versicolor.(1)),mean(Iris_versicolor.(2)),mean(Iris_versicolor.(3)),mean(Iris_versicolor.(4)))
fprintf('\nVariance : %.2f  %.2f      %.2f  %.2f',var(Iris_versicolor.(1)),var(Iris_versicolor.(2)),var(Iris_versicolor.(3)),var(Iris_versicolor.(4)))


fprintf('\n-----------------------------------------------')

fprintf('\nClass - Iris-Virginica')

fprintf('\n        Sepal Dim      Petal Dim')
fprintf('\n      Length Width(Cms) Length Width(Cms)')
fprintf('\nMax    : %.2f  %.2f      %.2f  %.2f',max(Iris_virginica.(1)),max(Iris_virginica.(2)),max(Iris_virginica.(3)),max(Iris_virginica.(4)))
fprintf('\nMin    : %.2f  %.2f      %.2f  %.2f',min(Iris_virginica.(1)),min(Iris_virginica.(2)),min(Iris_virginica.(3)),min(Iris_virginica.(4)))
fprintf('\nMean   : %.2f  %.2f      %.2f  %.2f',mean(Iris_virginica.(1)),mean(Iris_virginica.(2)),mean(Iris_virginica.(3)),mean(Iris_virginica.(4)))
fprintf('\nVariance : %.2f  %.2f      %.2f  %.2f\n',var(Iris_virginica.(1)),var(Iris_virginica.(2)),var(Iris_virginica.(3)),var(Iris_virginica.(4)))
```

OUTPUT :

```
Class - Iris-Setosa
                Sepal Dim              Petal Dim
            Length  Width(Cms)     Length   Width(Cms)
Max       :  5.40    3.90           1.70      0.40
Min       :  4.40    2.90           1.30      0.10
Mean      :  4.86    3.31           1.45      0.22
Variance :   0.08    0.09           0.01      0.01
--------------------------------------------------------

Class - Iris-Versicolor
                Sepal Dim              Petal Dim
            Length  Width(Cms)     Length   Width(Cms)
Max       :  7.00    3.30           4.90      1.60
Min       :  4.90    2.30           3.30      1.00
Mean      :  6.10    2.87           4.37      1.38
Variance :   0.53    0.12           0.24      0.03
--------------------------------------------------------

Class - Iris-Virginica
                Sepal Dim              Petal Dim
            Length  Width(Cms)     Length   Width(Cms)
Max       :  7.60    3.60           6.60      2.50
Min       :  4.90    2.50           4.50      1.70
Mean      :  6.57    2.94           5.77      2.04
Variance :   0.65    0.11           0.36      0.08
```

%Plotting classes vs Sepal length, Sepal Width, Petal Length, Petal Width for each Species

```matlab
figure('Name','Classes vs Species Attributes','NumberTitle','off');

subplot(2,2,1);
plot(Iris_setosa.Class,Iris_setosa.(1),'LineStyle','none','Marker','*');
hold on
plot(Iris_versicolor.Class,Iris_versicolor.(1),'LineStyle','none','Marker','*');
hold on
plot(Iris_virginica.Class,Iris_virginica.(1),'LineStyle','none','Marker','*');
xlim([0,4]);
hold off
title("The length of Sepal of 3 Classes");
xlabel("Classes--->");
ylabel("Sepal Length(Cms)--->");

subplot(2,2,2);
plot(Iris_setosa.Class,Iris_setosa.(2),'LineStyle','none','Marker','*');
hold on
plot(Iris_versicolor.Class,Iris_versicolor.(2),'LineStyle','none','Marker','*');
hold on
plot(Iris_virginica.Class,Iris_virginica.(2),'LineStyle','none','Marker','*');
xlim([0,4]);
hold off
title("The Width of Sepal of 3 Classes");
xlabel("Classes--->");
ylabel("Sepal Width(Cms)--->");

subplot(2,2,3);
plot(Iris_setosa.Class,Iris_setosa.(3),'LineStyle','none','Marker','*');
hold on
plot(Iris_versicolor.Class,Iris_versicolor.(3),'LineStyle','none','Marker','*');
hold on
plot(Iris_virginica.Class,Iris_virginica.(3),'LineStyle','none','Marker','*');
xlim([0,4]);
ylim([0,7]);
hold off
title("The Length of Petal of 3 Classes");
xlabel("Classes--->");
ylabel("Petal Length(Cms)--->");

subplot(2,2,4);
plot(Iris_setosa.Class,Iris_setosa.(4),'LineStyle','none','Marker','*');
hold on
plot(Iris_versicolor.Class,Iris_versicolor.(4),'LineStyle','none','Marker','*');
hold on
plot(Iris_virginica.Class,Iris_virginica.(4),'LineStyle','none','Marker','*');
xlim([0,4]);
ylim([0,3]);
hold off
title("The Width of Petal of 3 Classes");
xlabel("Classes--->");
ylabel("Petal Width(Cms)--->");
```
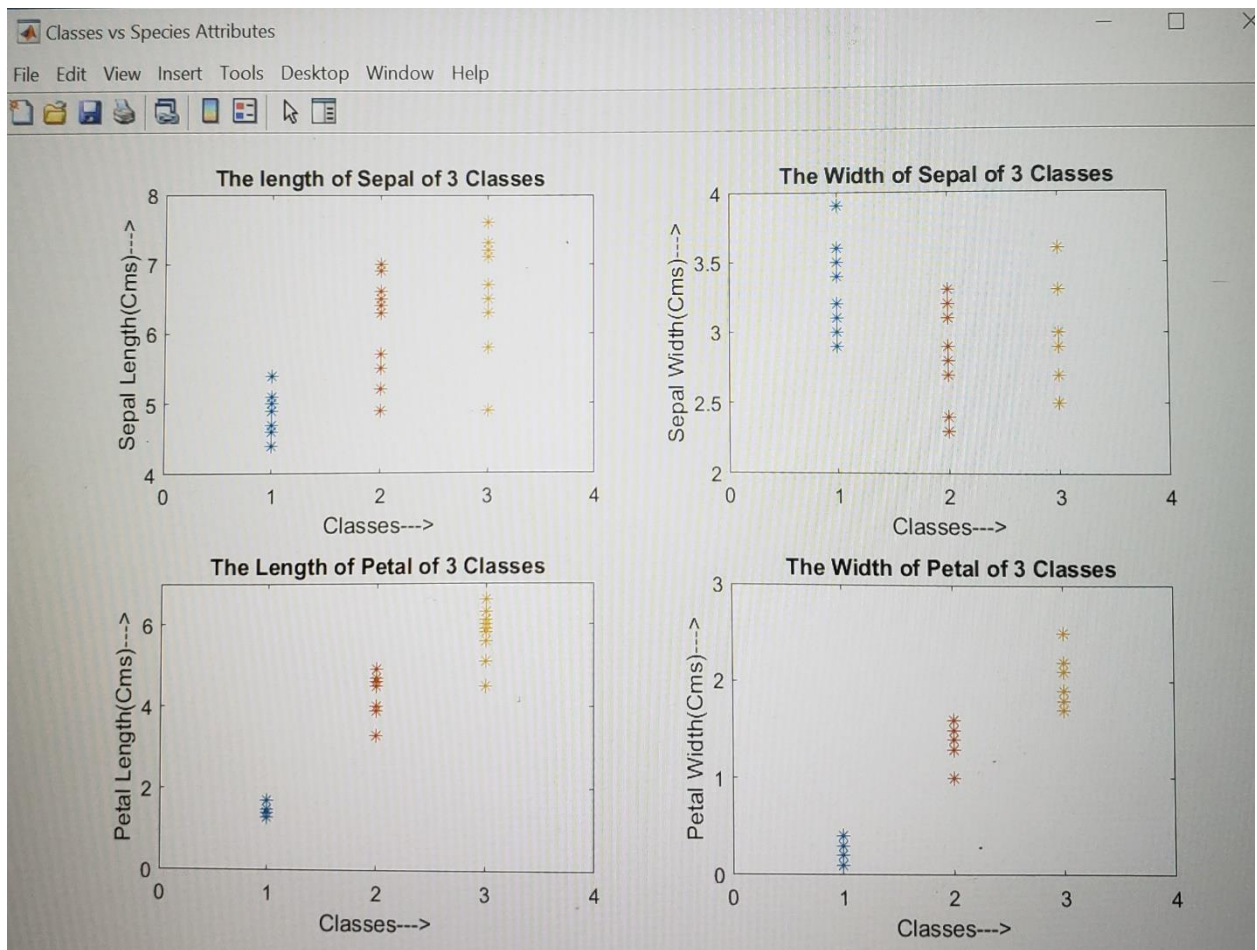
grid off;

OUTPUT :

%Extracting data from the table for Correlation Coefficient

corr_Setosa= [Iris_setosa.(1),Iris_setosa.(2),Iris_setosa.(3),Iris_setosa.(4)];
corr_versicolor= [Iris_versicolor.(1),Iris_versicolor.(2),Iris_versicolor.(3),Iris_versicolor.(4)];
corr_virginica = [Iris_virginica.(1),Iris_virginica.(2),Iris_virginica.(3),Iris_virginica.(4)];

%Displaying Correlation Matrix for Each classes i.e Species

corr = [corr_Setosa,corr_versicolor,corr_virginica];
Cm = corrcoef(corr);

[rows, cols] = size(corrcoef(corr));

fprintf('\n\n\n')

fprintf(' Correlation Matrix of Iris-Setosa vs Iris-Versicolor vs Iris-Virginica ')
fprintf('\n\n   Iris-Setosa   Iris-Versicolor   Iris-Virginica')
fprintf('\n   Sepal   Septal   Petal   Petal   Sepal   Sepal   Petal   Petal   Sepal   Sepal
Petal   Petal ')

```matlab
fprintf('\n        length   width   length   width   length   width   length   width   length   width   length   width \n')

for i = 1 : rows
    if i == 1
        fprintf('Sepal len ')
    end
        if i == 2
        fprintf('Sepal wid ')
        end
        if i == 3
        fprintf('petal len ')
        end
        if i == 4
        fprintf('petal wid ')

        end
        if i == 5
        fprintf('Sepal len ')
        end
        if i == 6
        fprintf('Sepal wid ')
        end
        if i == 7
        fprintf('petal len ')
        end
        if i == 8
        fprintf('Petal wid ')
        end
        if i == 9
        fprintf('Sepal len ')
        end
        if i == 10
        fprintf('Sepal wid ')
        end
        if i == 11
        fprintf('petal len ')
        end
        if i == 12
        fprintf('Petal wid ')
        end

    for j=1:cols
        fprintf('%10.3f',Cm(i,j))
    end
    fprintf('\n')
end
```
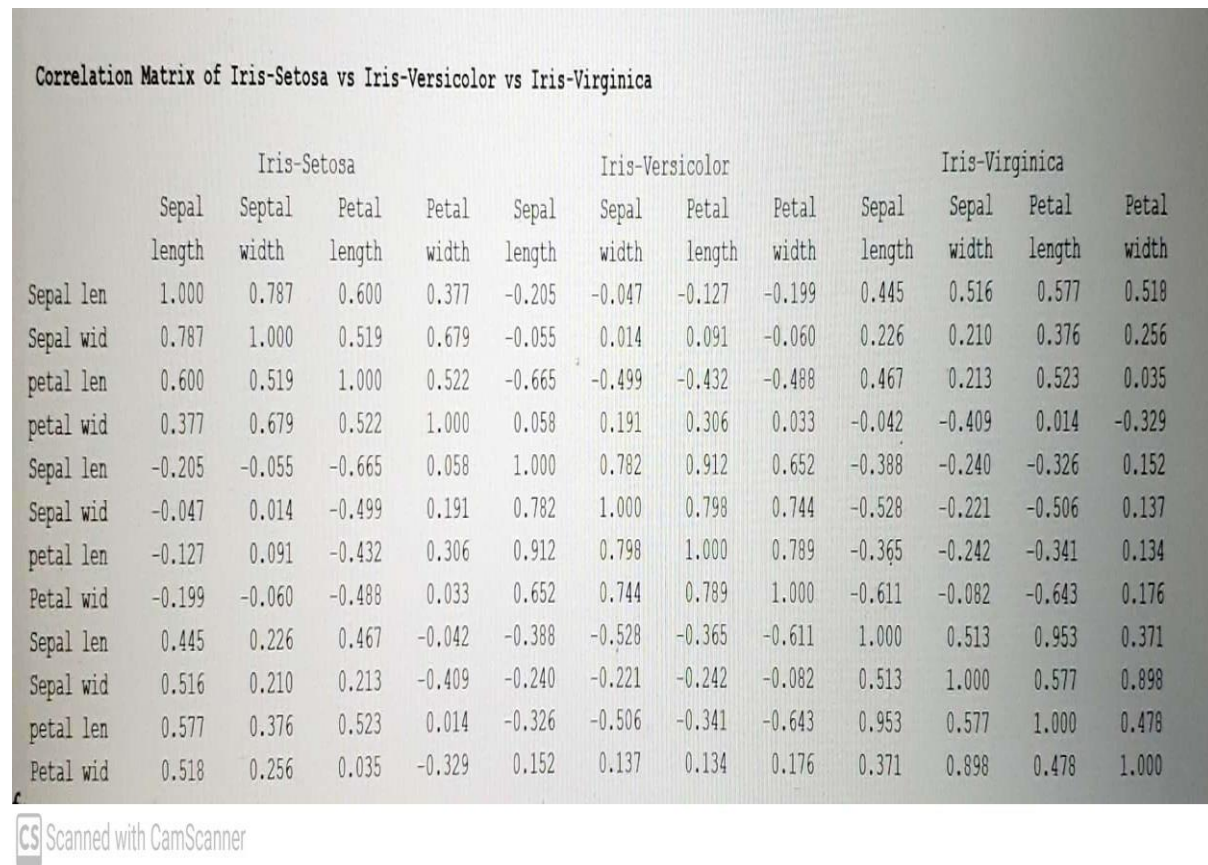
OUTPUT :

Correlation Matrix of Iris-Setosa vs Iris-Versicolor vs Iris-Virginica

	Iris-Setosa				Iris-Versicolor				Iris-Virginica			
	Sepal length	Septal width	Petal length	Petal width	Sepal length	Sepal width	Petal length	Petal width	Sepal length	Sepal width	Petal length	Petal width
Sepal len	1.000	0.787	0.600	0.377	-0.205	-0.047	-0.127	-0.199	0.445	0.516	0.577	0.518
Sepal wid	0.787	1.000	0.519	0.679	-0.055	0.014	0.091	-0.060	0.226	0.210	0.376	0.256
petal len	0.600	0.519	1.000	0.522	-0.665	-0.499	-0.432	-0.488	0.467	0.213	0.523	0.035
petal wid	0.377	0.679	0.522	1.000	0.058	0.191	0.306	0.033	-0.042	-0.409	0.014	-0.329
Sepal len	-0.205	-0.055	-0.665	0.058	1.000	0.782	0.912	0.652	-0.388	-0.240	-0.326	0.152
Sepal wid	-0.047	0.014	-0.499	0.191	0.782	1.000	0.798	0.744	-0.528	-0.221	-0.506	0.137
petal len	-0.127	0.091	-0.432	0.306	0.912	0.798	1.000	0.789	-0.365	-0.242	-0.341	0.134
Petal wid	-0.199	-0.060	-0.488	0.033	0.652	0.744	0.789	1.000	-0.611	-0.082	-0.643	0.176
Sepal len	0.445	0.226	0.467	-0.042	-0.388	-0.528	-0.365	-0.611	1.000	0.513	0.953	0.371
Sepal wid	0.516	0.210	0.213	-0.409	-0.240	-0.221	-0.242	-0.082	0.513	1.000	0.577	0.898
petal len	0.577	0.376	0.523	0.014	-0.326	-0.506	-0.341	-0.643	0.953	0.577	1.000	0.478
Petal wid	0.518	0.256	0.035	-0.329	0.152	0.137	0.134	0.176	0.371	0.898	0.478	1.000

Observation:

From the correlation matrix, it is observed that the output classes are highly correlated with length parameter (Septal and Petal) for all 3 Classes. The width Parameter (Septal and Petal) are little low correlated than length parameter but not too bad. This shows that all the classes are directly related with septal and petal length and width.

% Graphing Frequency for all 3 classes Attributes

```
figure('Name','Iris-Setosa (Sepal and Petal(cms)) Frequency','NumberTitle','off');
subplot(2,2,1);
histogram(Iris_setosa.(1),7);
ylim([0,4]);
title("Sepal Length(Iris-Setosa)");
xlabel("Iris-Setosa Sepal Length(cms)");
ylabel("Frequency");

subplot(2,2,2);
histogram(Iris_setosa.(2),7);
ylim([0,4]);
title(" Sepal Width(Iris-Setosa)");
xlabel("Iris-Setosa Sepal Width(cms)");
ylabel("Frequency");

subplot(2,2,3);
```
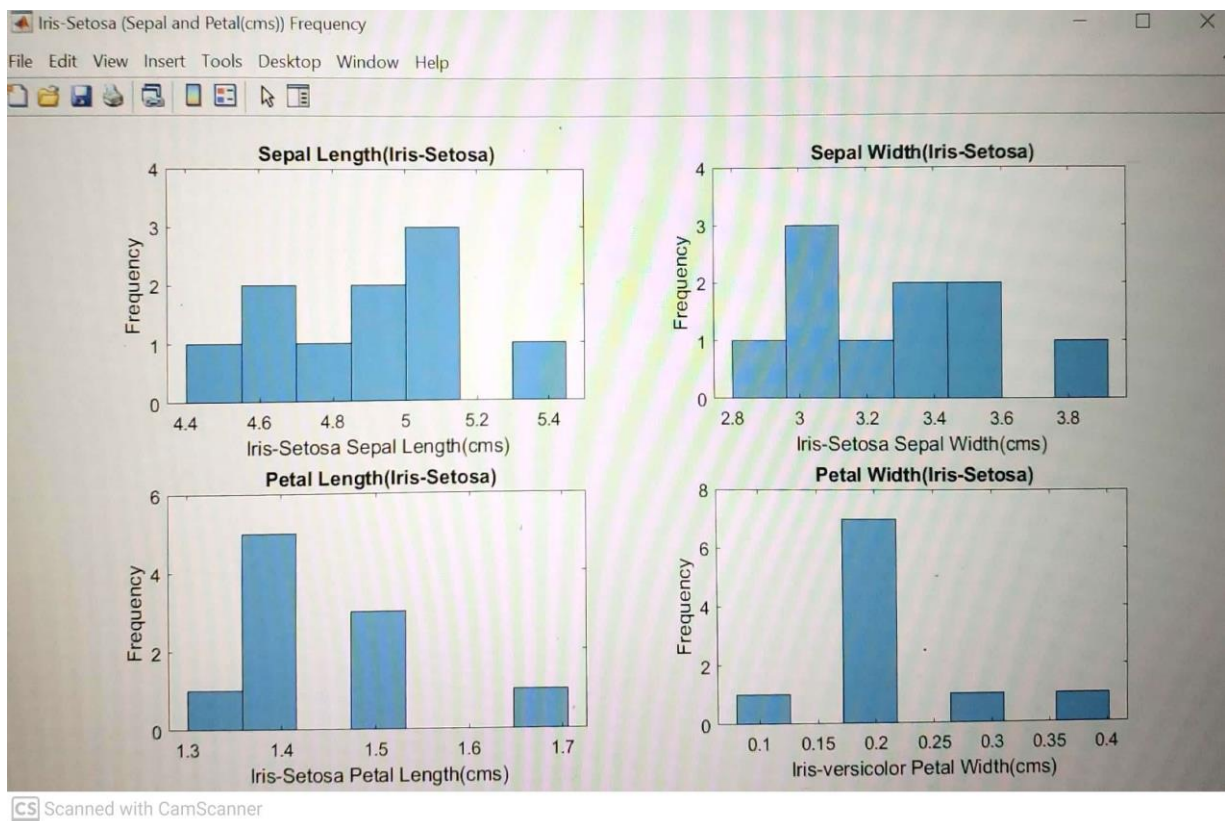
```
histogram(Iris_setosa.(3),7);
ylim([0,6]);
title(" Petal Length(Iris-Setosa)");
xlabel("Iris-Setosa Petal Length(cms)");
ylabel("Frequency");

subplot(2,2,4);
histogram(Iris_setosa.(4),7);
title(" Petal Width(Iris-Setosa)");
xlabel("Iris-versicolor Petal Width(cms)");
ylabel("Frequency");
grid off;
```

OUTPUT :


```
figure('Name','Iris-Versicolor (Sepal and Petal(cms)) Frequency','NumberTitle','off');
subplot(2,2,1);
histogram(Iris_versicolor.(1),7);
ylim([0,4]);
title("Sepal Length(Iris-versicolor)");
xlabel("Iris-versicolor Sepal Length(cms)");
ylabel("Frequency");

subplot(2,2,2);
histogram(Iris_versicolor.(2),7);
ylim([0,4]);
title(" Sepal Width(Iris-versicolor)");
xlabel("Iris-versicolor Sepal Width(cms)");
ylabel("Frequency");

subplot(2,2,3);
histogram(Iris_versicolor.(3),7);
```
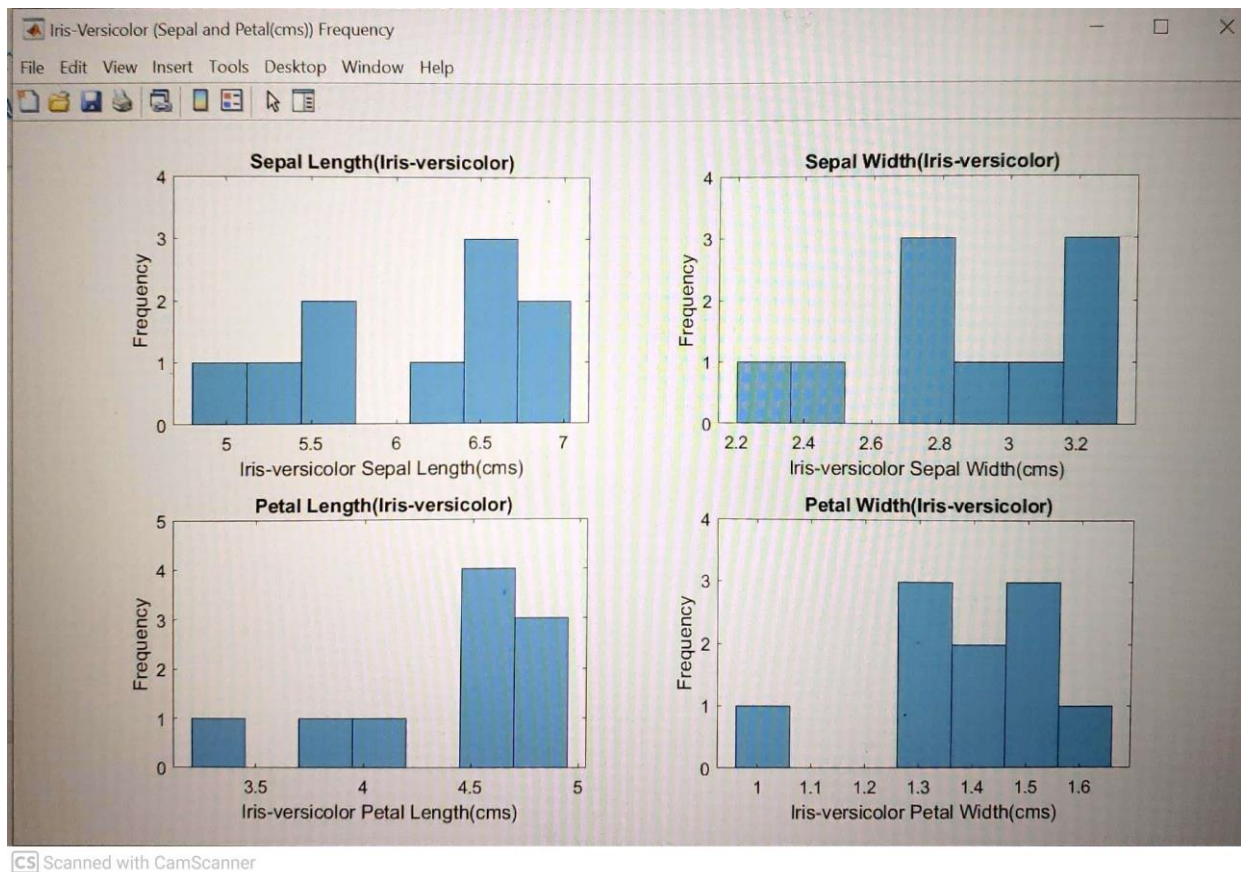
```
ylim([0,5]);
title(" Petal Length(Iris-versicolor)");
xlabel("Iris-versicolor Petal Length(cms)");
ylabel("Frequency");

subplot(2,2,4);
histogram(Iris_versicolor.(4),7);
title(" Petal Width(Iris-versicolor)");
ylim([0,4]);
xlabel("Iris-versicolor Petal Width(cms)");
ylabel("Frequency");
grid off;
```

OUTPUT :

```
figure('Name','Iris-Virginica (Sepal and Petal(cms)) Frequency','NumberTitle','off');
subplot(2,2,1);
histogram(Iris_virginica.(1),7);
ylim([0,3]);
title("Sepal Length(Iris-virginica)");
xlabel("Iris-virginica Sepal Length(cms)");
ylabel("Frequency");

subplot(2,2,2);
histogram(Iris_virginica.(2),7);
ylim([0,4]);
title(" Sepal Width(Iris-virginica)");
xlabel("Iris-virginica Sepal Width(cms)");
ylabel("Frequency");

subplot(2,2,3);
```
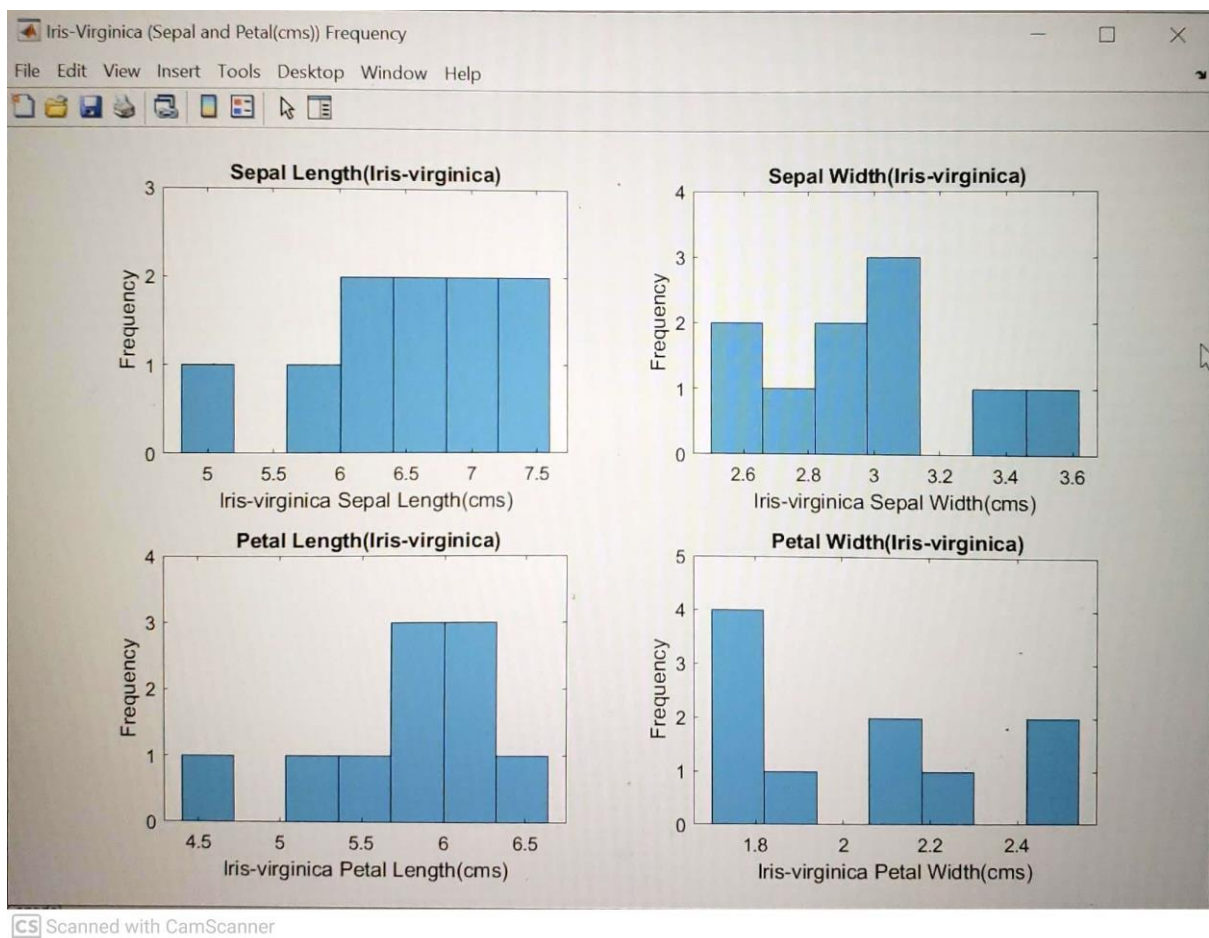
```
histogram(Iris_virginica.(3),7);
ylim([0,4]);
title(" Petal Length(Iris-virginica)");
xlabel("Iris-virginica Petal Length(cms)");
ylabel("Frequency");

subplot(2,2,4);
histogram(Iris_virginica.(4),7);
ylim([0,5]);
title(" Petal Width(Iris-virginica)");
xlabel("Iris-virginica Petal Width(cms)");
ylabel("Frequency");
grid off;
```

OUTPUT :

Explanation:

- The sepal length and width and Petal length and width are the features that are used to predict the plant classes.
- The mean, max, min and variance were computed using the following formulas:

$$Mean\ (\mu) = \frac{1}{N}\sum_{i=1}^{N} X_i$$

$$Variance\ (Var) = \frac{1}{N}\sum_{i=1}^{N}(X_i - \mu)^2$$

Where Xi is the length in cms.

- A correlation matrix is also computed, and histogram is used to understand frequency 0f each classes.
- Plotted classes vs all species length and width (Sepal and Petal) i.e., classes to examine 3 species length and width (sepal and petal).