MATLAB LAB ASSIGNMENT -13

~CHARVI JAIN (RA2111047010113)

Q1. Access and explore the weather dataset/patients dataset/Heart failure prediction dataset.

Q2. Preprocess the data and extract features using Live Editor

Tasks: Steps to be implemented for Data preprocessing:

a) Fill Missing Data

b) Fill Outliers

c) Smooth Data

d) Locate Extrema

Write the steps to view the code that a task used to generate the output and visualization.

PATIENTS DATASET

```
load patients
whos
T=table(Age,Gender,Height,Weight,Smoker,...
 'RowNames',LastName)
```

```
% Create plot of T.Height
h3 = plot(T.Height,'DisplayName','Height');
```

```
% Add ylabel, title, and legend
ylabel('Height')
title('Height')
legend
```

```
% Fill missing data
[newTable,missingIndices] = fillmissing(T,"previous");
```

```
% Display results
clf
plot(newTable.Age,"Color",[0 114 189]/255,"LineWidth",1.5,...
 "DisplayName","Cleaned data")
hold on
```

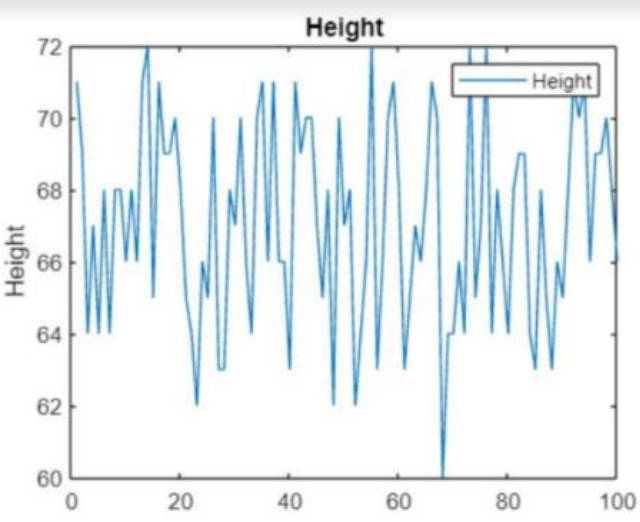
```
% Plot filled missing entries
plot(find(missingIndices(:,1)),newTable.Age(missingIndices(:,1)),".
",...  "MarkerSize",12,"Color",[217 83 25]/255,...
 "DisplayName","Filled missing entries")
title("Number of filled missing entries: " + nnz(missingIndices(:,1)))
```

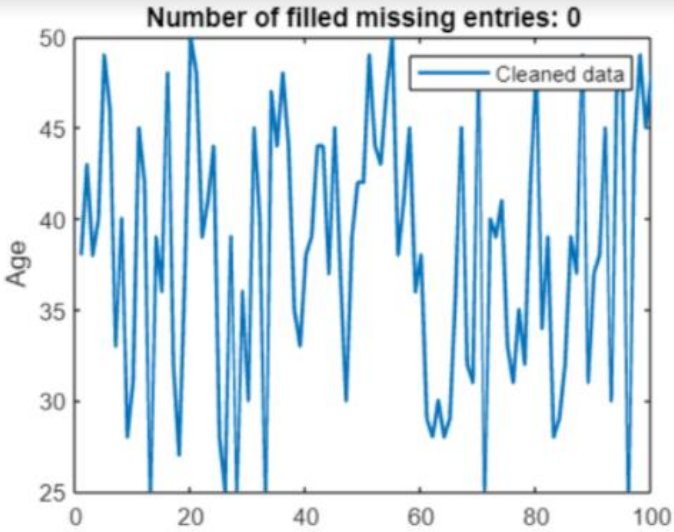
```
hold off
legend
ylabel("Age")
clear missingIndices
```

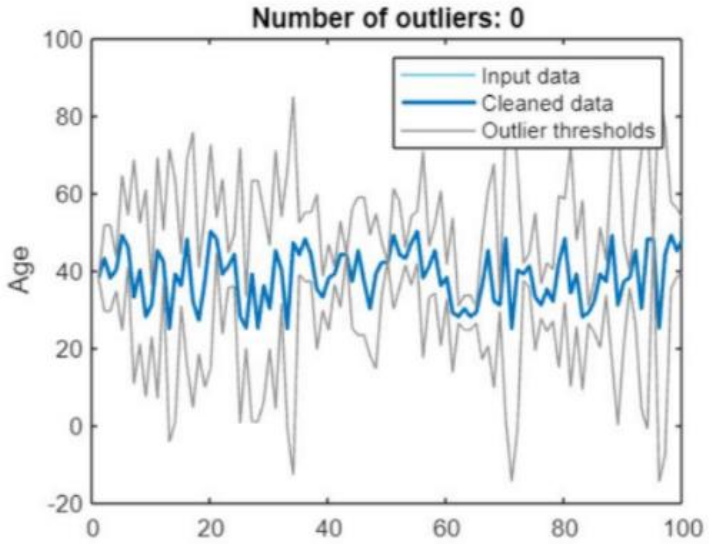
```
% Fill outliers
```

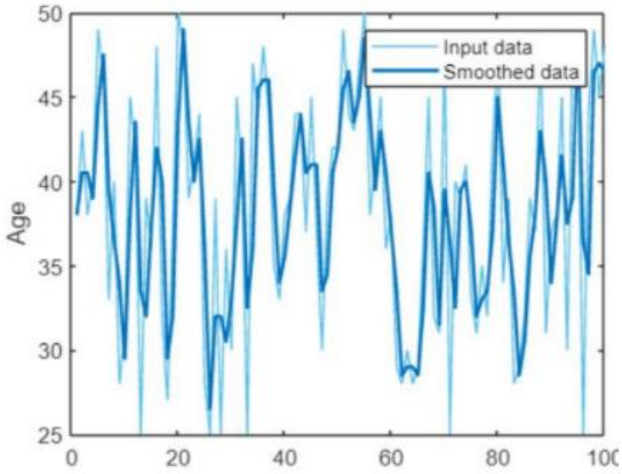
```
[newTable2,outlierIndices,thresholdLow,thresholdHigh] =
filloutliers(newTable,...  "linear","movmedian",2,"DataVariables",["Age","He
ight","Weight"]);
% Display results
clf
plot(newTable.Age,"Color",[77 190 238]/255,"DisplayName","Input
data") hold on
plot(newTable2.Age,"Color",[0 114
189]/255,"LineWidth",1.5,...  "DisplayName","Cleaned
data")


% Plot outliers
plot(find(outlierIndices(:,1)),newTable.Age(outlierIndices(:,1)),"x
",...  "Color",[64 64 64]/255,"DisplayName","Outliers")
title("Number of outliers: " + nnz(outlierIndices(:,1)))


% Plot filled outliers
plot(find(outlierIndices(:,1)),newTable2.Age(outlierIndices(:,1)),".",.
..  "MarkerSize",12,"Color",[217 83 25]/255,"DisplayName","Filled
outliers")


% Plot outlier thresholds
plot([(1:numel(newTable.Age))'; missing;
(1:numel(newTable.Age))'],...  [thresholdHigh.Age(:); missing;
thresholdLow.Age(:)],...  "Color",[145 145
145]/255,"DisplayName","Outlier thresholds")


hold off
legend
ylabel("Age")
clear outlierIndices thresholdLow thresholdHigh

% Smooth input data
newTable3 =
smoothdata(newTable2,"movmean","SmoothingFactor",0.25,...  "Data
Variables",["Age","Height","Weight","Smoker"]);


% Display results
clf
plot(newTable2.Age,"Color",[77 190 238]/255,"DisplayName","Input
data") hold on
plot(newTable3.Age,"Color",[0 114
189]/255,"LineWidth",1.5,...  "DisplayName","Smoothed
data")
hold off
legend
ylabel("Age")

% Find local maxima
maxIndices3 = islocalmax(newTable3,...
 "DataVariables",["Age","Height","Weight","Smoker"]);
% Display results
clf
plot(newTable3.Age,"Color",[77 190 238]/255,"DisplayName","Input
data") hold on
```

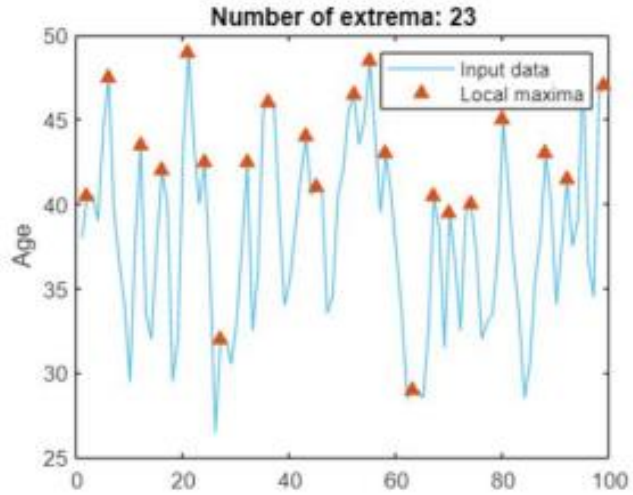
```
% Plot local maxima
plot(find(maxIndices3(:,1)),newTable3.Age(maxIndices3(:,1)),"^
",...  "Color",[217 83 25]/255,"MarkerFaceColor",[217 83
25]/255,...  "DisplayName","Local maxima")
title("Number of extrema: " + nnz(maxIndices3(:,1)))
hold off
legend
ylabel("Age")
```







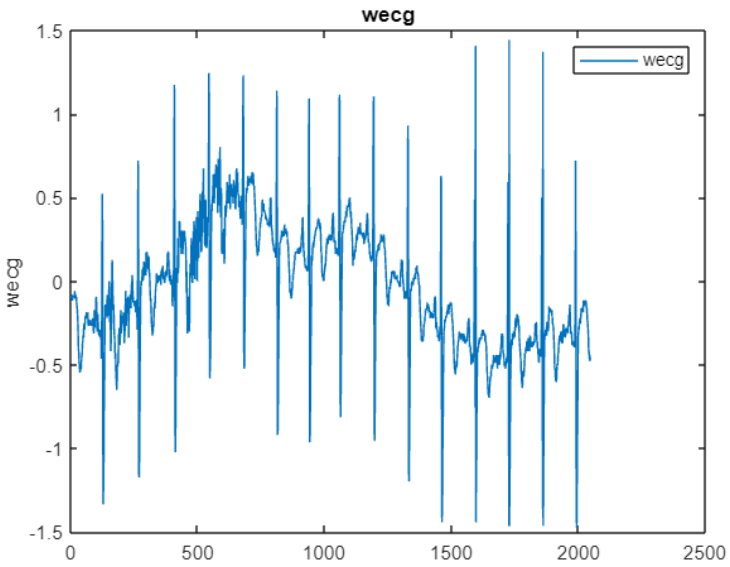


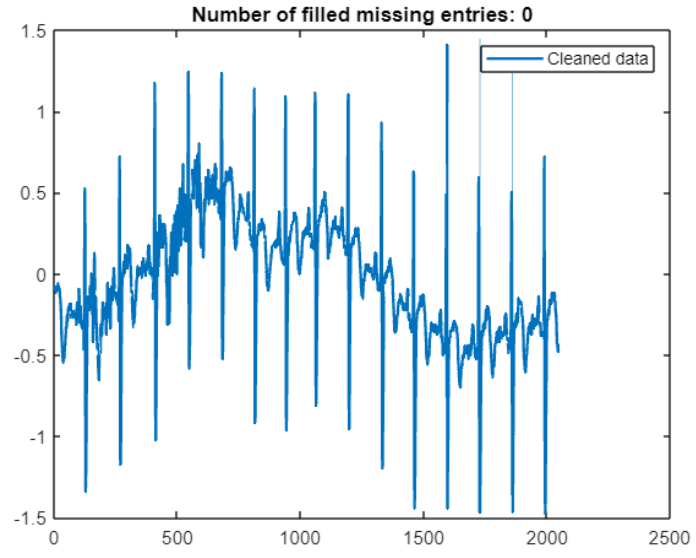


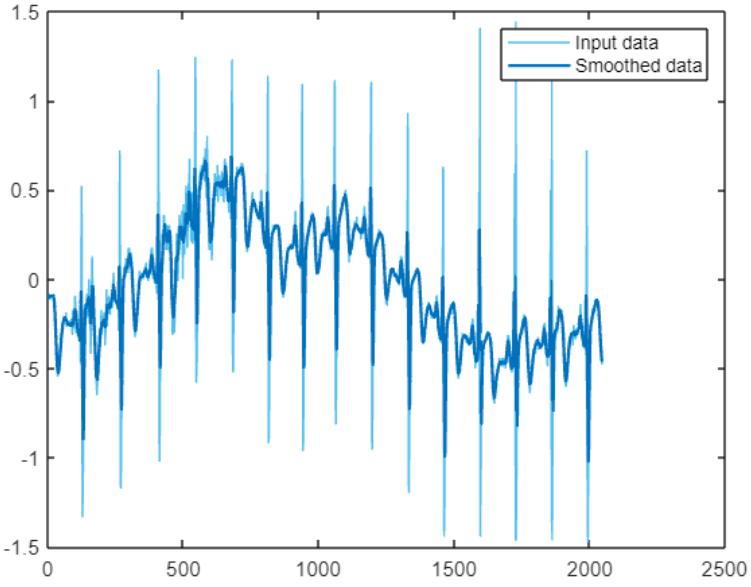
HEART SOUND - ECG DATASET

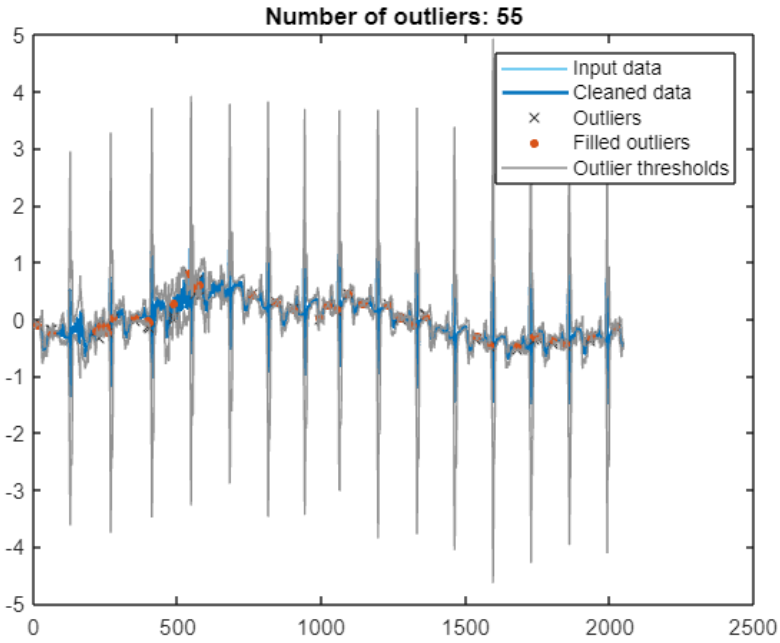
```
load wecg;
[a,d] = haart(wecg);

% Create plot of wecg
h = plot(wecg,'DisplayName','wecg');
% Add ylabel, title, and legend
ylabel('wecg')
title('wecg')
legend

% Fill missing data
[cleanedData,missingIndices] = fillmissing(wecg,"linear");


% Display results
clf
plot(cleanedData,"Color",[0 114 189]/255,"LineWidth",1.5,...
    "DisplayName","Cleaned data")
hold on


% Plot filled missing entries
plot(find(missingIndices),cleanedData(missingIndices),".","MarkerSize",12,...
    "Color",[217 83 25]/255,"DisplayName","Filled missing entries")
title("Number of filled missing entries: " + nnz(missingIndices))


hold off
legend
clear missingIndices

% Fill outliers
[cleanedData2,outlierIndices,thresholdLow,thresholdHigh] = ...
    filloutliers(cleanedData,"linear","movmedian",8);


% Display results
clf
plot(cleanedData,"Color",[77 190 238]/255,"DisplayName","Input data")
hold on
plot(cleanedData2,"Color",[0 114 189]/255,"LineWidth",1.5,...
    "DisplayName","Cleaned data")


% Plot outliers
plot(find(outlierIndices),cleanedData(outlierIndices),"x",...
    "Color",[64 64 64]/255,"DisplayName","Outliers")
title("Number of outliers: " + nnz(outlierIndices))


% Plot filled outliers
plot(find(outlierIndices),cleanedData2(outlierIndices),".","MarkerSize",12,...
    "Color",[217 83 25]/255,"DisplayName","Filled outliers")


% Plot outlier thresholds
plot([(1:numel(cleanedData))'; missing; (1:numel(cleanedData))'],...
```

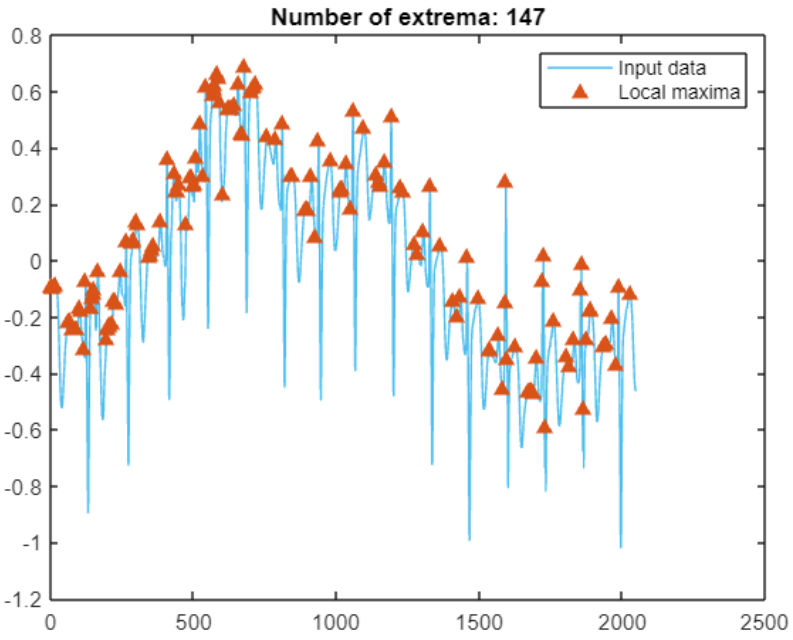
```
    [thresholdHigh(:); missing; thresholdLow(:)],"Color",[145 145 145]/255,...
    "DisplayName","Outlier thresholds")


hold off
legend
clear outlierIndices thresholdLow thresholdHigh

% Smooth input data
smoothedData = smoothdata(cleanedData2,"movmean","SmoothingFactor",0.25);


% Display results
clf
plot(cleanedData2,"Color",[77 190 238]/255,"DisplayName","Input data")
hold on
plot(smoothedData,"Color",[0 114 189]/255,"LineWidth",1.5,...
    "DisplayName","Smoothed data")
hold off
legend

% Find local maxima
maxIndices = islocalmax(smoothedData);


% Display results
clf
plot(smoothedData,"Color",[77 190 238]/255,"DisplayName","Input data")
hold on


% Plot local maxima
plot(find(maxIndices),smoothedData(maxIndices),"^","Color",[217 83 25]/255,...
    "MarkerFaceColor",[217 83 25]/255,"DisplayName","Local maxima")
title("Number of extrema: " + nnz(maxIndices))
hold off
legenD
```






Input data
Smoothed data
1.5
1
0.5
0
-0.5
-1
-1.5
0
500
1000
1500
2000
2500


Number of outliers: 55
Input data
Cleaned data
Outliers
Filled outliers
Outlier thresholds
5
4
3
2
1
0
-1
-2
-3
-4
-5
0
500
1000
1500
2000
2500


Number of extrema: 147
Input data
Local maxima
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
-1.2
0
500
1000
1500
2000
2500