

## % Exploratory Data Analysis for MIT-BIH Arrhythmia Database

```
T = readtable ("MIT-BIH Arrhythmia Database.csv");
```

Warning: Column headers from the file were modified to make them valid MATLAB identifiers before creating variable names for the table. The original column headers are saved in the VariableDescriptions property. Set 'VariableNamingRule' to 'preserve' to use the original column headers as table variable names.

```
T = rmmissing(T);
if any(strcmp('0_rPeak', T.Properties.VariableNames))
    T = T(T.('0_rPeak') > -3000 & T.('0_rPeak') < 3000, :);
end

%(z-score normalization)
numericVars = varfun(@isnumeric, T, 'OutputFormat', 'uniform');
T_norm = T;
T_norm(:, numericVars) = varfun(@(x) (x - mean(x)) / std(x), T(:, numericVars));

disp('First 5 Rows:');
```

First 5 Rows:

```
disp(head(T, 5));
```

record	type	x0_pre_RR	x0_post_RR	x0_pPeak	x0_tPeak	x0_rPeak	x0_sPeak	x0_qPeak	x0_oPeak
101	{'N'}	76	313	0.074347	-0.16055	1.0364	-0.28566	-0.026824	-0.000000
101	{'N'}	313	315	-0.052079	-0.26478	0.8866	-0.3663	-0.05971	-0.000000
101	{'N'}	315	321	-0.062151	-0.29698	0.99186	-0.41031	-0.065686	-0.000000
101	{'N'}	321	336	-0.063322	-0.28139	1.0349	-0.40388	-0.07175	-0.000000
101	{'N'}	336	344	-0.062915	1.0469	1.0464	1.0464	-0.074639	-0.000000

```
%% Unique Types and Records
disp('Unique Beat Types:');
```

Unique Beat Types:

```
disp(unique(T.type));
```

```
{'F' }
{'N' }
{'Q' }
{'SVEB'}
{'VEB' }
```

```
disp('Unique Records:');
```

Unique Records:

```
disp(unique(T.record));
```

```
100
101
103
105
```

106  
108  
109  
111  
112  
113  
114  
115  
116  
117  
118  
119  
121  
122  
123  
124  
200  
201  
202  
203  
205  
207  
208  
209  
210  
212  
213  
214  
215  
219  
220  
221  
222  
223  
228  
230  
231  
232  
233  
234

```
% Summary Statistics for Numeric Features
```

```
numericData = T(:, varfun(@isnumeric, T, 'OutputFormat', 'uniform'));  
disp('Mean Values:');
```

Mean Values:

```
disp(varfun(@mean, numericData));
```

<u>mean_record</u>	<u>mean_x0_pre_RR</u>	<u>mean_x0_post_RR</u>	<u>mean_x0_pPeak</u>	<u>mean_x0_tPeak</u>	<u>mean_x0_rPeak</u>	<u>mean_x0_s</u>
173.73	279.4	278.59	0.049772	0.37226	1.0545	-0.134

```
disp('Standard Deviation:');
```

Standard Deviation:

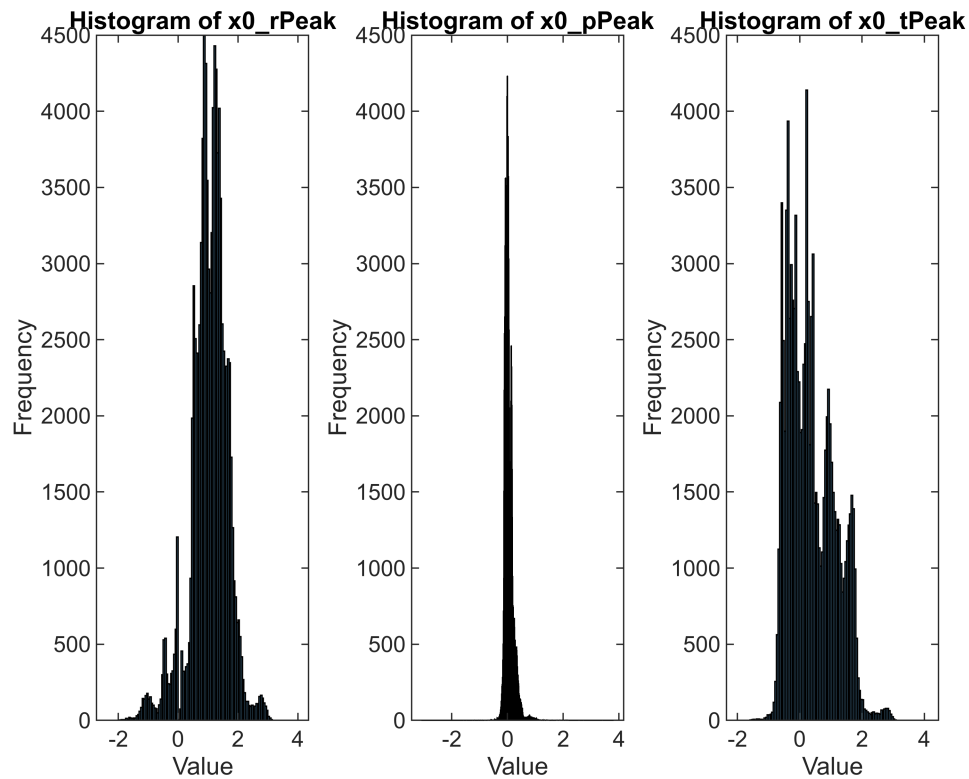
```
disp(varfun(@std, numericData));
```

<u>std_record</u>	<u>std_x0_pre_RR</u>	<u>std_x0_post_RR</u>	<u>std_x0_pPeak</u>	<u>std_x0_tPeak</u>	<u>std_x0_rPeak</u>	<u>std_x0_sPeak</u>
51.673	81.915	81.125	0.16879	0.73988	0.61751	1.0276

```

%% Histograms of Selected Features
selectedFeatures = {'x0_rPeak', 'x0_pPeak', 'x0_tPeak'};
figure;
for i = 1:length(selectedFeatures)
    subplot(1, length(selectedFeatures), i);
    histogram(T.(selectedFeatures{i}));
    title(['Histogram of ', strcmp(selectedFeatures{i}, '_'), '\_']);
    xlabel('Value'); ylabel('Frequency');
end

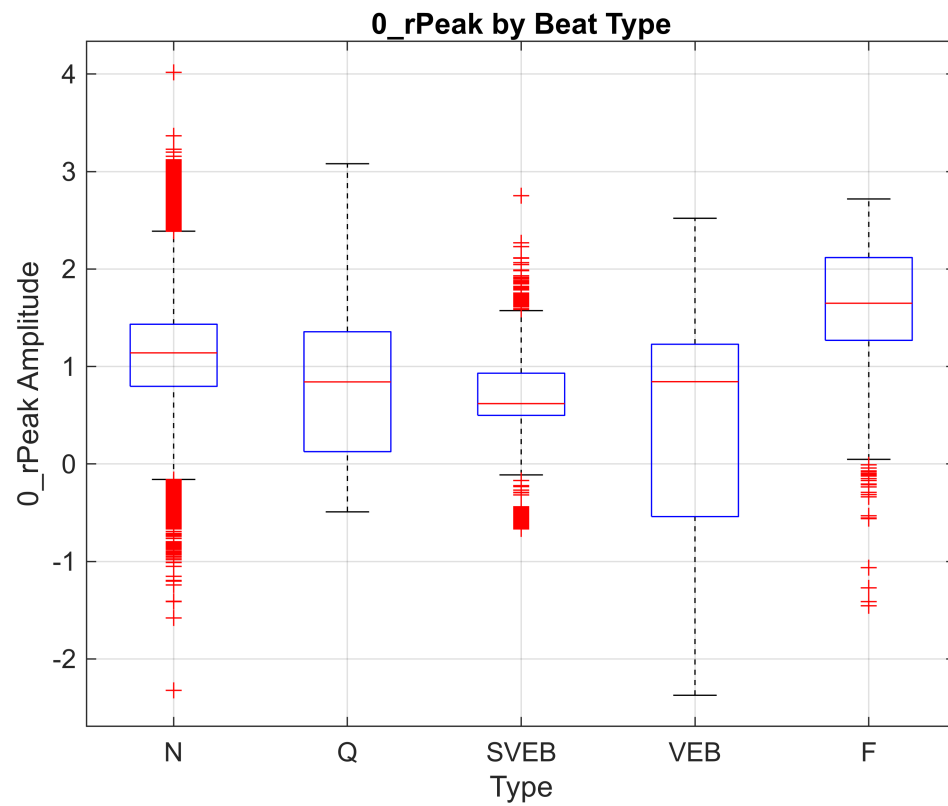
```



```

%% Boxplot of 0_rPeak by Beat Type
figure;
boxplot(T.('x0_rPeak'), T.type);
title('0\_rPeak by Beat Type');
xlabel('Type'); ylabel('0\_rPeak Amplitude');
grid on;

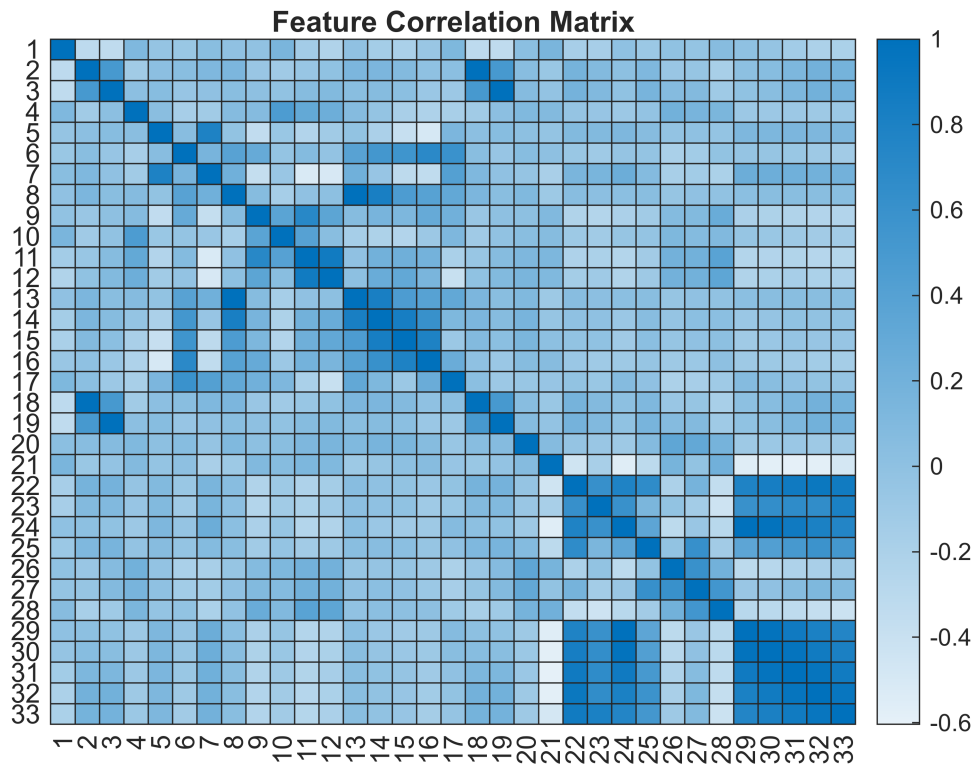
```



```

%% Correlation Matrix
numericMatrix = table2array(numericData);
corrMatrix = corr(numericMatrix, 'Rows', 'complete');
figure;
heatmap(corrMatrix);
title('Feature Correlation Matrix');

```



```
% Beat Type Distribution
typeCounts = groupcounts(T.type);
figure;
bar(typeCounts);
xticks(1:length(typeCounts));
xticklabels(unique(T.type));
title('Class Distribution (Beat Types)');
xlabel('Beat Type'); ylabel('Count');
grid on;
```

