```
% Exploratory Data Analysis for INCART 2-lead Arrhythmia Dataset

T = readtable('INCART 2-lead 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));
```

```
Dataset Dimensions: 175729 rows, 34 columns
```

```
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
{'I01'}	{'N' }	163	165	0.06961	-0.083281	0.61413	-0.39276	0.047159
{'I01'}	{'N' }	165	166	-0.09703	0.59725	-0.078704	-0.078704	-0.13778
{'I01'}	{'N' }	166	102	0.1094	0.68053	-0.010649	-0.010649	-0.72062
{'I01'}	{'VEB'}	102	231	0.17638	0.25643	-0.1011	-0.70752	-0.1011
{'I01'}	{'N' }	231	165	0.58558	0.60746	-0.083499	-0.083499	-0.16786

```
%% 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));
```

{'I01'}
{'I02'}
{'I03'}
{'I04'}
{'I05'}
{'I06'}
{'I07'}
{'I08'}
{'I09'}
{'I10'}
{'I11'}
{'I12'}
{'I13'}
{'I14'}
{'I15'}
{'I16'}
{'I17'}
{'I18'}
{'I19'}
{'I20'}
{'I21'}
{'I22'}
{'I23'}
{'I24'}
{'I25'}
{'I26'}
{'I27'}
{'I28'}
{'I29'}
{'I30'}
{'I31'}
{'I32'}
{'I33'}
{'I34'}
{'I35'}
{'I36'}
{'I37'}
{'I38'}
{'I39'}
{'I40'}
{'I41'}
{'I42'}
{'I43'}
{'I44'}
{'I45'}
{'I46'}
{'I47'}
{'I48'}
{'I49'}
{'I50'}
{'I51'}
{'I52'}
{'I53'}
{'I54'}
{'I55'}
{'I56'}
{'I57'}
{'I58'}
{'I59'}
{'I60'}
{'I61'}
{'I62'}
{'I63'}
{'I64'}

```
{'I65'}
{'I66'}
{'I67'}
{'I68'}
{'I69'}
{'I70'}
{'I71'}
{'I72'}
{'I73'}
{'I74'}
{'I75'}
```

```
%% Summary Statistics for Numeric Features
numericData = T(:, varfun(@isnumeric, T, 'OutputFormat', 'uniform'));
disp('Mean Values:');
```

```
Mean Values:
```

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

mean_x0_pre_RR	mean_x0_post_RR	mean_x0_pPeak	mean_x0_tPeak	mean_x0_rPeak	mean_x0_sPeak	mean_x0
197.18	197.2	0.039047	0.11772	0.85669	-0.58676	-0.1

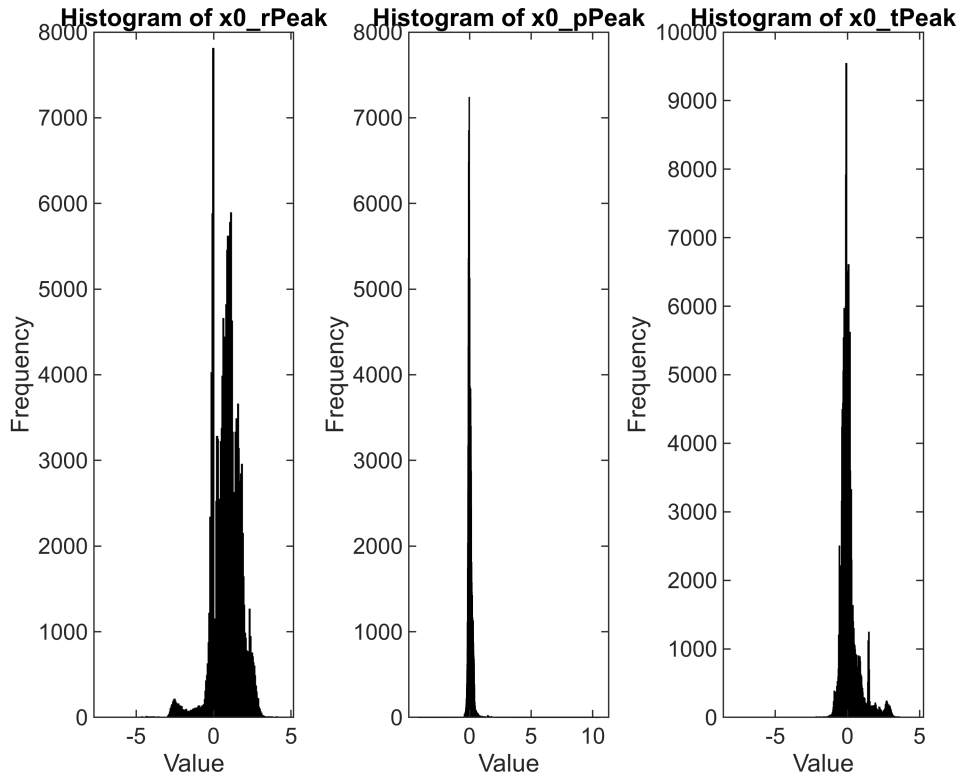
```
disp('Standard Deviation:');
```

```
Standard Deviation:
```

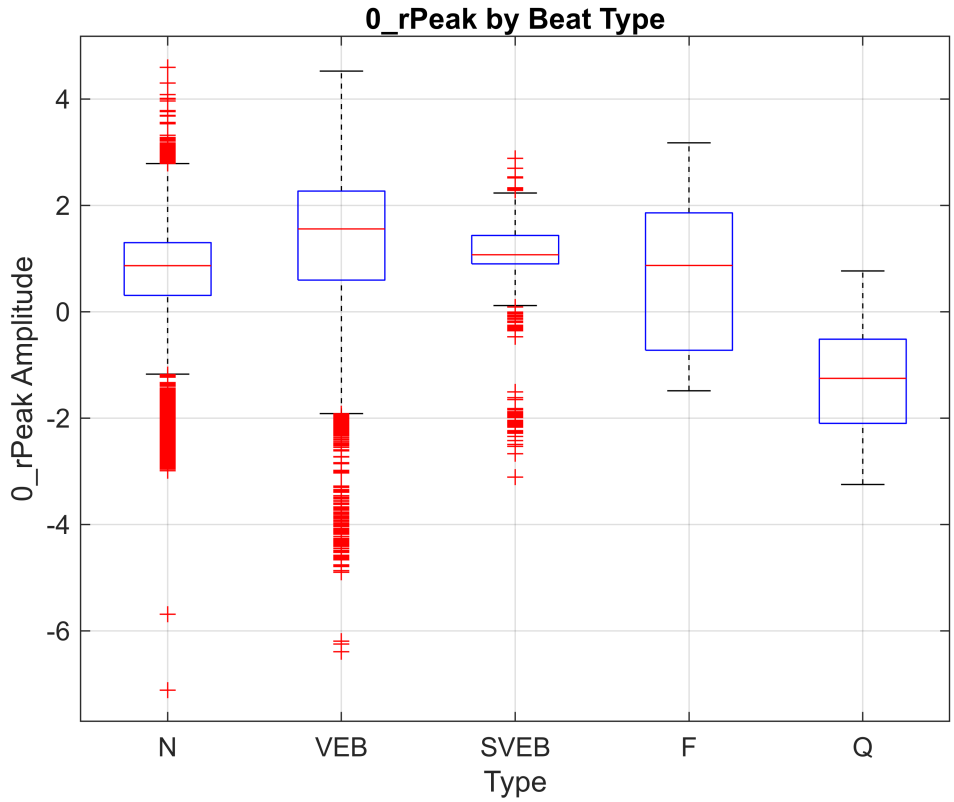
```
disp(varfun(@std, numericData));
```

std_x0_pre_RR	std_x0_post_RR	std_x0_pPeak	std_x0_tPeak	std_x0_rPeak	std_x0_sPeak	std_x0_qPeak
61.704	61.722	0.19216	0.63696	0.86691	0.79955	0.3688

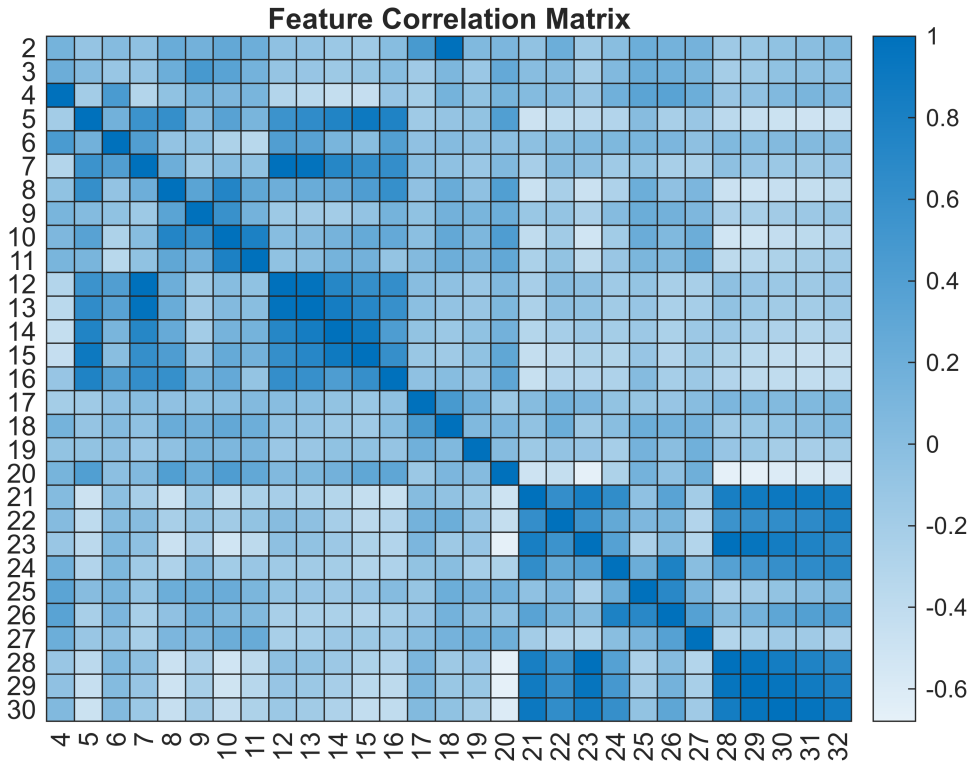
```
%% Histograms of Selected Features
selectedFeatures = {'x0_rPeak', 'x0_pPeak', 'x0_tPeak'};
figure;
for i = 1:length(selectedFeatures)
    subplot(1, length(selectedFeatures), i);
    histogram(T.(selectedFeatures{i}));
    title(['Histogram of ', strrep(selectedFeatures{i}, '_', '\_')]);
    xlabel('Value');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



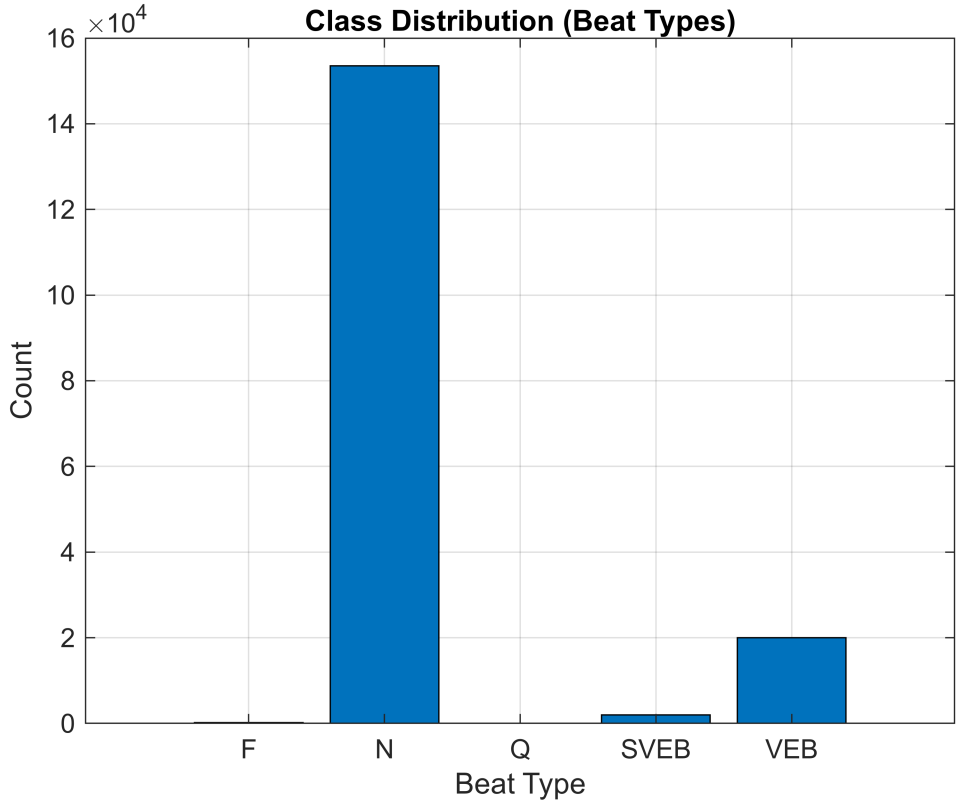

```
%% Boxplot of 0_rPeak by Beat Type
figure;
boxplot(T.('x0_rPeak'), T.type);
title('0\_rPeak by Beat Type');
xlabel('Type'); ylabel('0\_rPeak Amplitude');
grid on;
```




```
%% 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;
```


Class Distribution (Beat Types)
×10^4
16
14
12
10
8
6
4
2
0
Count
F
N
Q
SVEB
VEB
Beat Type