Human Behavior Analytics
Homework 2

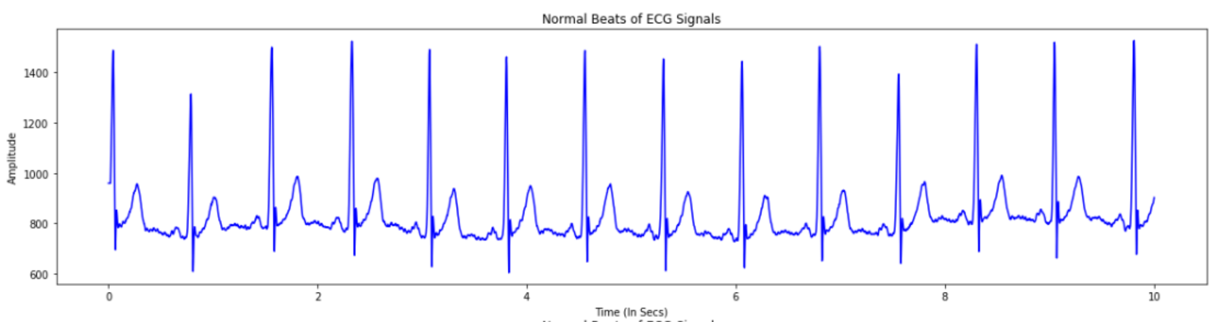
By: Sameer Kumar Behera   **UIN:** 526004296

REMARK – Jupyter Python Codes Files **(HBA_HW_Part1_abc, HBA_HW_Part1_de, HBA_HW_Part2)** attached in the Zip Folder.

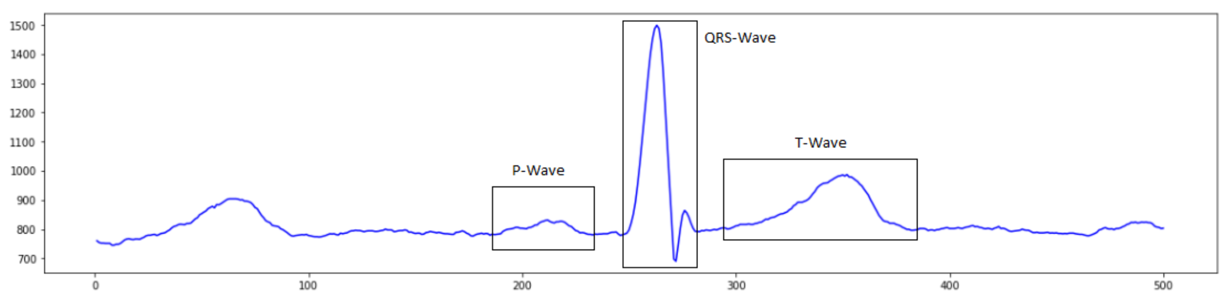
Q1)
ECG Signals with Normal Beats

From the Annotations_NormalAbnormalBeats Folder, we come to know that the Ventricular Beats of Candidate 116 are 109 out of 2410 (Quite Low) and that of Candidate 208 are 991 of 2577 (Quite High).

So, For the Report, I have shown Normal Beats of Candidate-116 and Ventricular Beats of Candidate-208



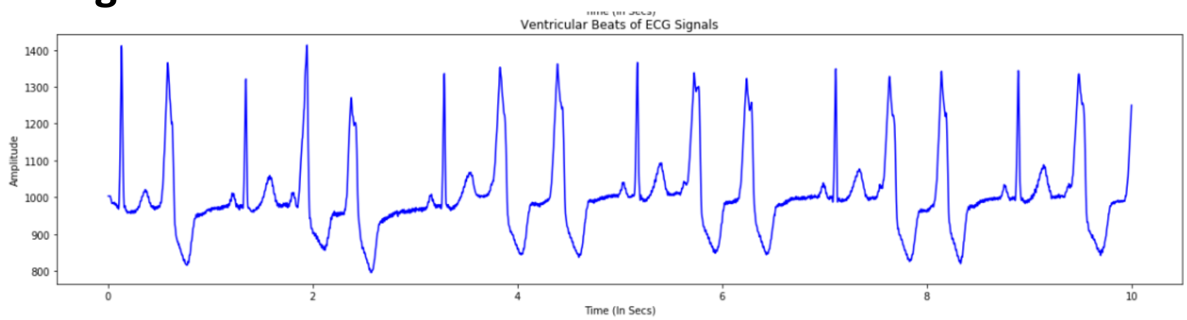
(Displaying ECG Signal of Normal Beat of Candidate – 116)



(L=500, t=np.linspace(1,L,L), plt.plot(t,x116[300:L+300],'b-'), Zoomed In View of 116)

P-Wave of 116 = x_116[487:533]
QRS-Wave of 116 = x_116[547:583]
T-Wave of 116 = x_116[591:682]

ECG Signals with Ventricular Beats

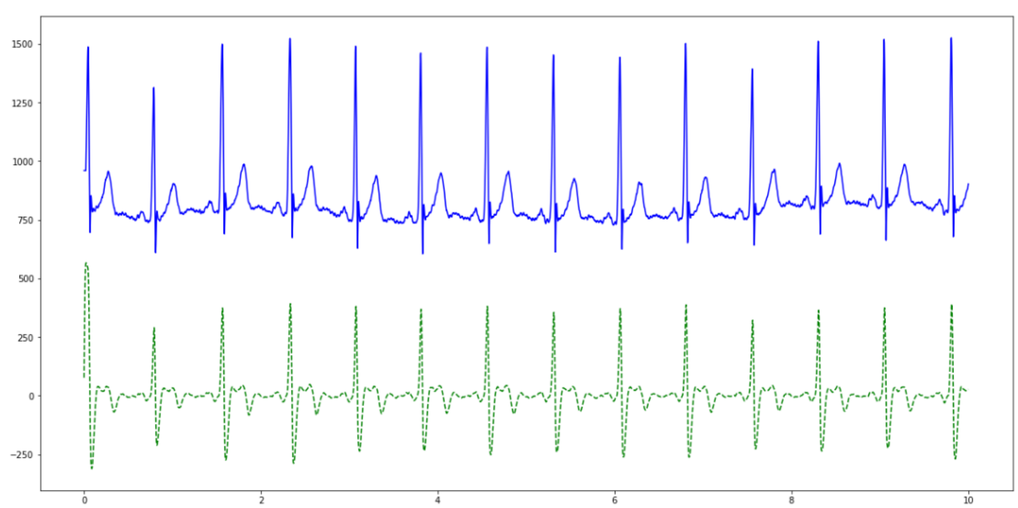


(Displaying ECG Signal of Ventricular Beat of Candidate – 208)

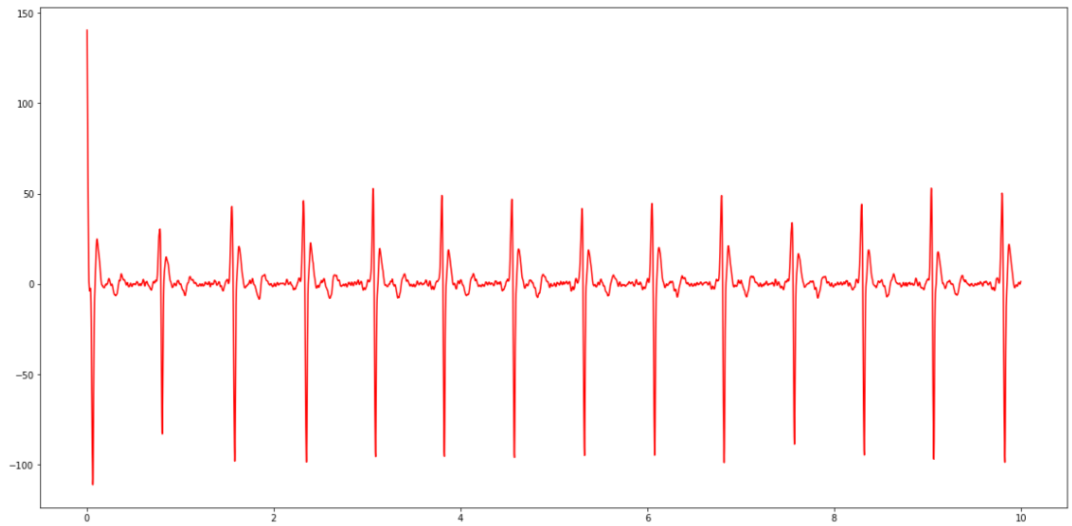
Pan-Tomkins QRS Detection

Parameters Taken:

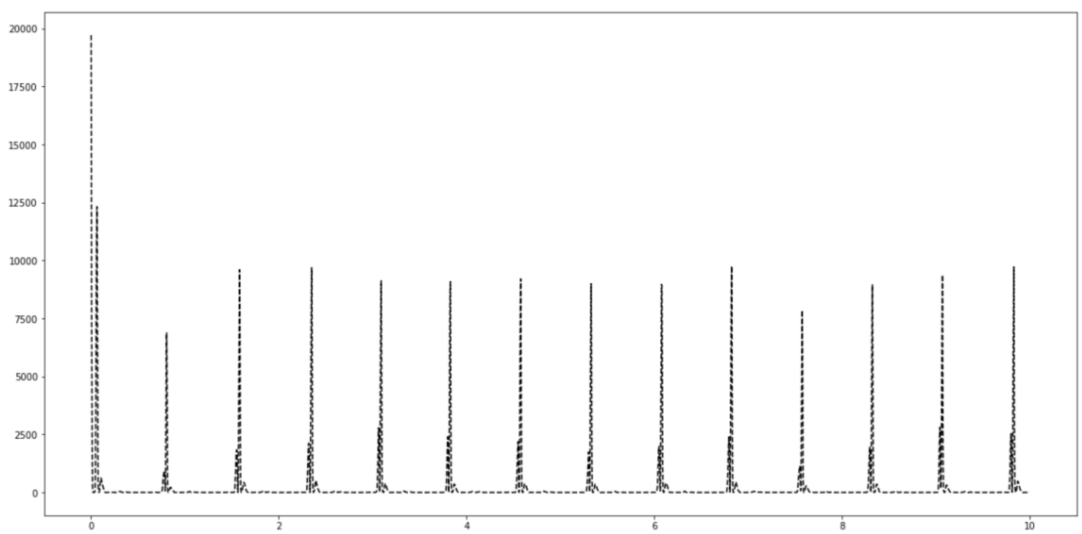
```
# Segment Original Signal
L=3600
x=x116[0:L]
t=np.linspace(1,L,L)/Fs

# Band-Pass Filter
(lowcut,highcut,filter_order)=(5,15,1)
(x_bfilt,b,a)=bandpass_filter(x, lowcut, highcut, Fs, filter_order)
#print(L,len(t),len(x))
plt.figure(figsize=(20,10));plt.plot(np.linspace(1,len(x_bfilt),len(x_bfilt))/Fs,x,'b-',t,x_bfilt,'g--')
```



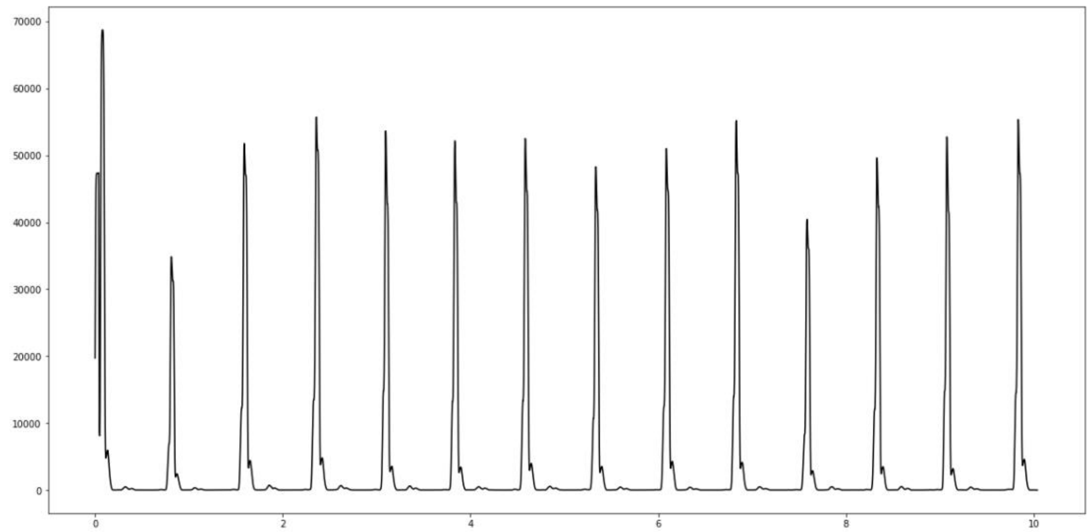
# Differentiate
x_bfilt_diff=np.ediff1d(x_bfilt)
plt.figure(figsize=(20,10));plt.plot(np.linspace(1,len(x_bfilt_diff),len(x_bfilt_diff))/Fs,x_bfilt_diff,'r-')



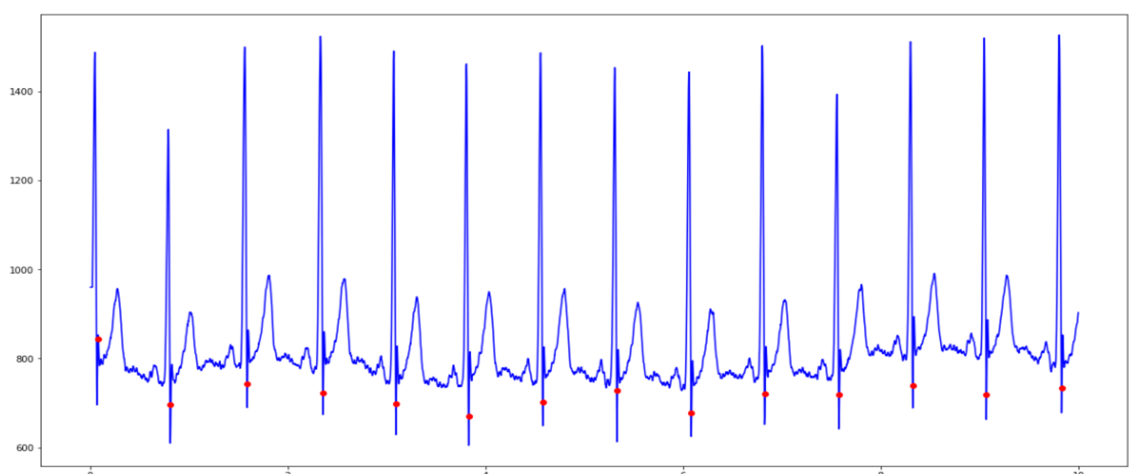
# Square
x_bfilt_diff_sq=x_bfilt_diff**2
plt.figure(figsize=(20,10));plt.plot(np.linspace(1,len(x_bfilt_diff_sq),len(x_bfilt_diff_sq))/Fs,x_bfilt_diff_sq,'k--')



```
# Moving Average Filter
MAfilter_len=15
x_bfilt_diff_sq_smooth = np.convolve(x_bfilt_diff_sq, np.ones(MAfilter_len))
plt.figure(figsize=(20,10));plt.plot(np.linspace(1,len(x_bfilt_diff_sq_smooth),len(x_bfilt_diff_sq_smooth)
)/Fs,x_bfilt_diff_sq_smooth,'k-')
```

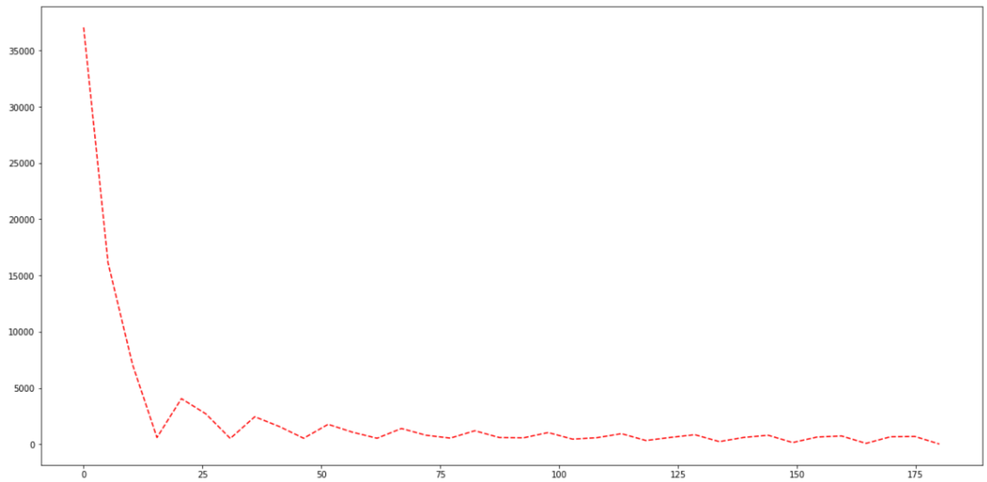


```
# Find Peaks
(findpeaks_limit,findpeaks_spacing)=(1,100)
peaks_indices = findpeaks(x_bfilt_diff_sq_smooth,findpeaks_spacing,findpeaks_limit)
print(peaks_indices)
(precision,recall) = calc_pre_rec(peaks_indices)
print("Precision: "+str(precision))
print("Recall: "+str(recall))
plt.figure(figsize=(20,10));plt.plot(np.linspace(1,len(x),len(x))/Fs,x,'b-',peaks_indices/Fs,x[peaks_indices],
'ro')
```

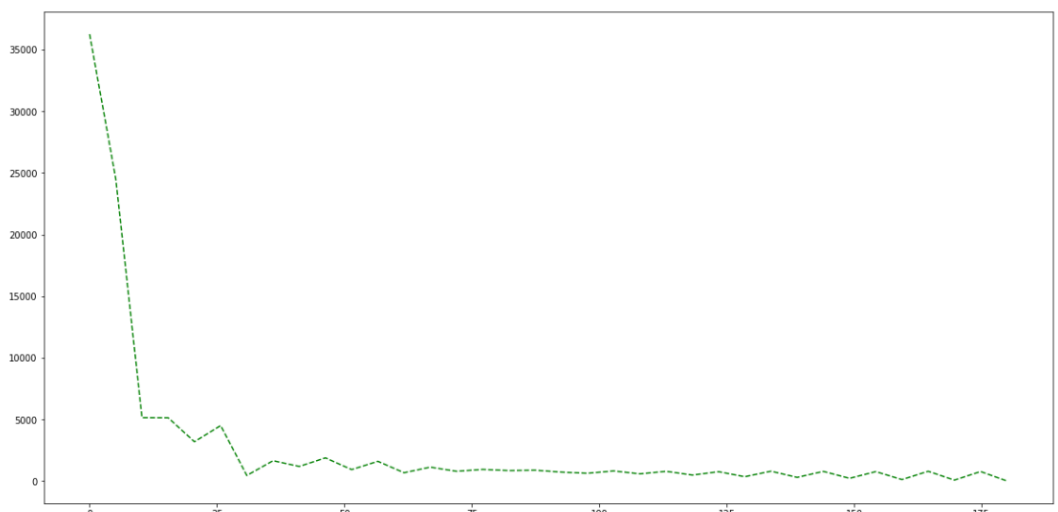


Precision for 3600 Samples of Candidate 116 = 0.9285714285714286
Recall for 3600 Samples of Candidate 116 = 1

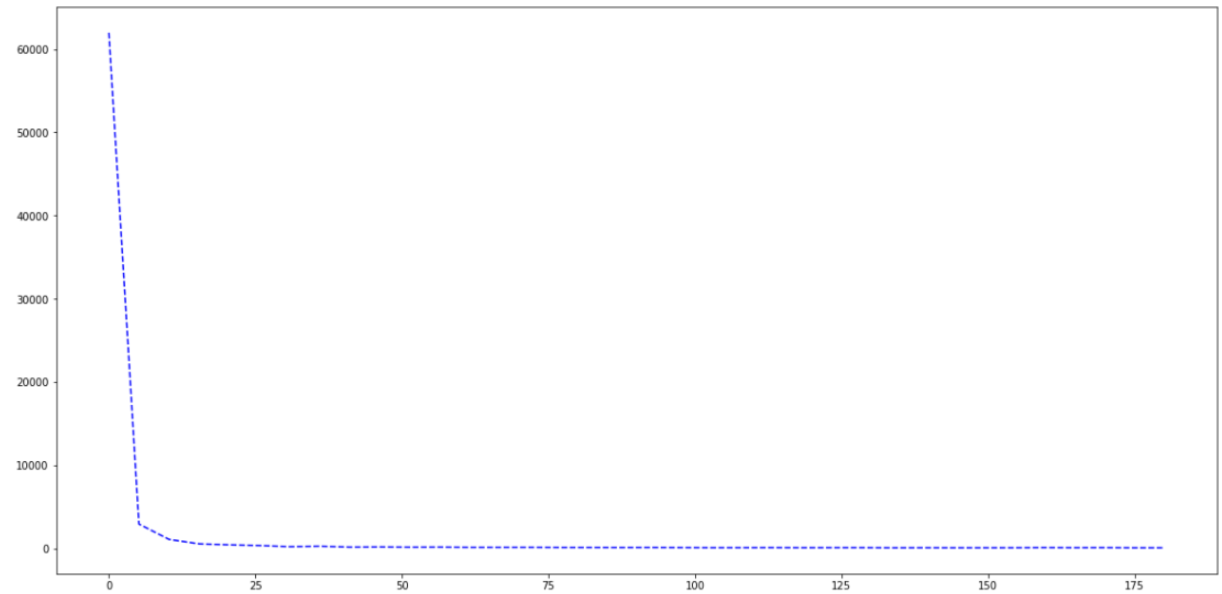
```
# P-Wave Visualization
x116_P=x116[487:533];
N=70
fft_x116_P=np.fft.rfft(x116_P,N)
f = np.linspace(0,(Fs/2),N/2+1)
plt.figure(figsize=(20,10));plt.plot(f,np.abs(fft_x116_P),'r--');plt.show()
```




```
# QRS-Wave Visualization
x116_QRS=x116[547:583];
N=70
fft_x116_QRS=np.fft.rfft(x116_QRS,N)
f = np.linspace(0,(Fs/2),N/2+1)
plt.figure(figsize=(20,10));plt.plot(f,np.abs(fft_x116_QRS),'g--');plt.show()
```



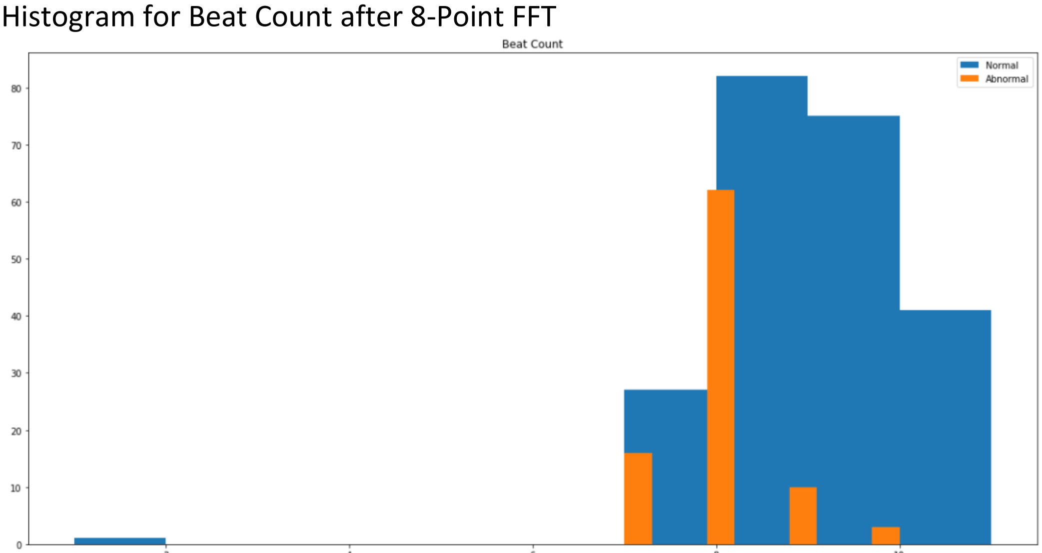
```
# T-Wave Visualization
x116_T=x116[591:682];
N=70
fft_x116_T=np.fft.rfft(x116_T,N)
f = np.linspace(0,(Fs/2),N/2+1)
plt.figure(figsize=(20,10));plt.plot(f,np.abs(fft_x116_T),'b--');plt.show()
```



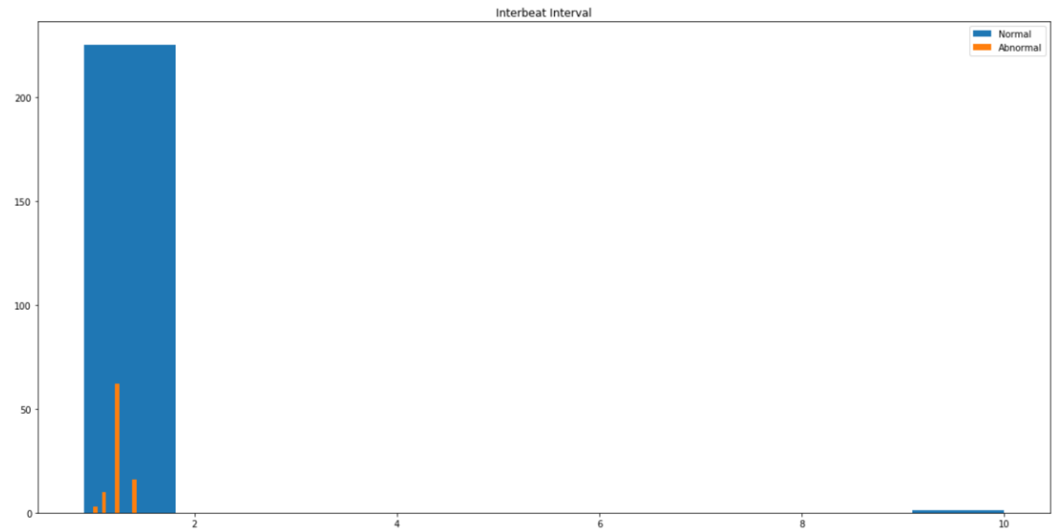
ECG Feature Extraction

ECG Feature Extraction was performed for Candidate 208 which was having higher ratio of Ventricular Beats to Normal Beats in comparison to the other candidates.

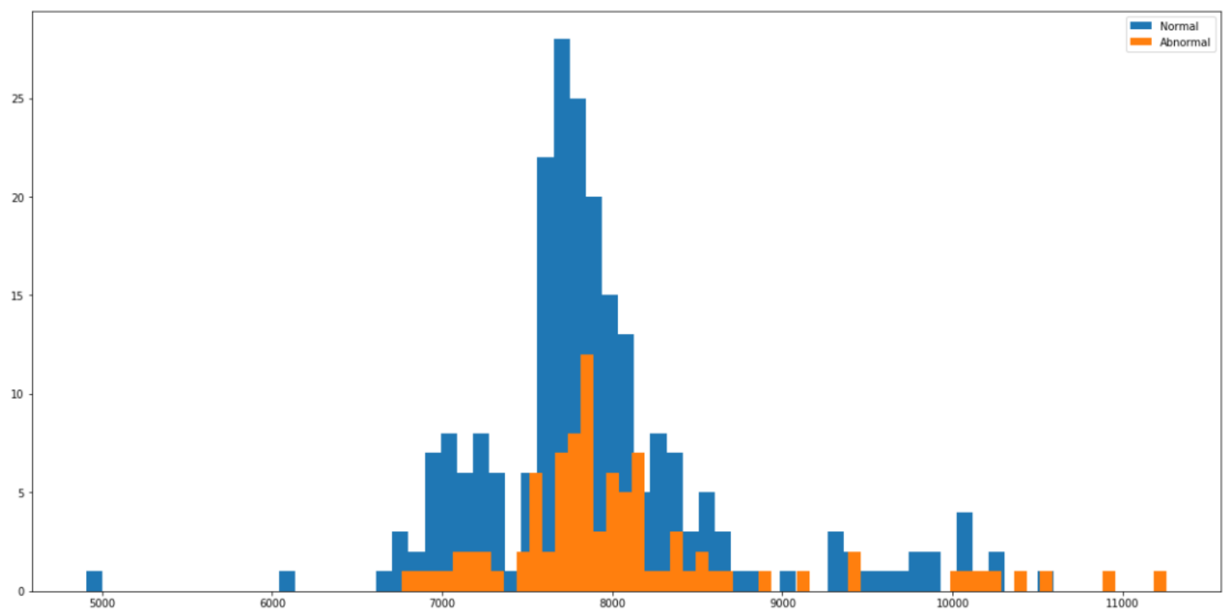
Histogram for Beat Count after 8-Point FFT



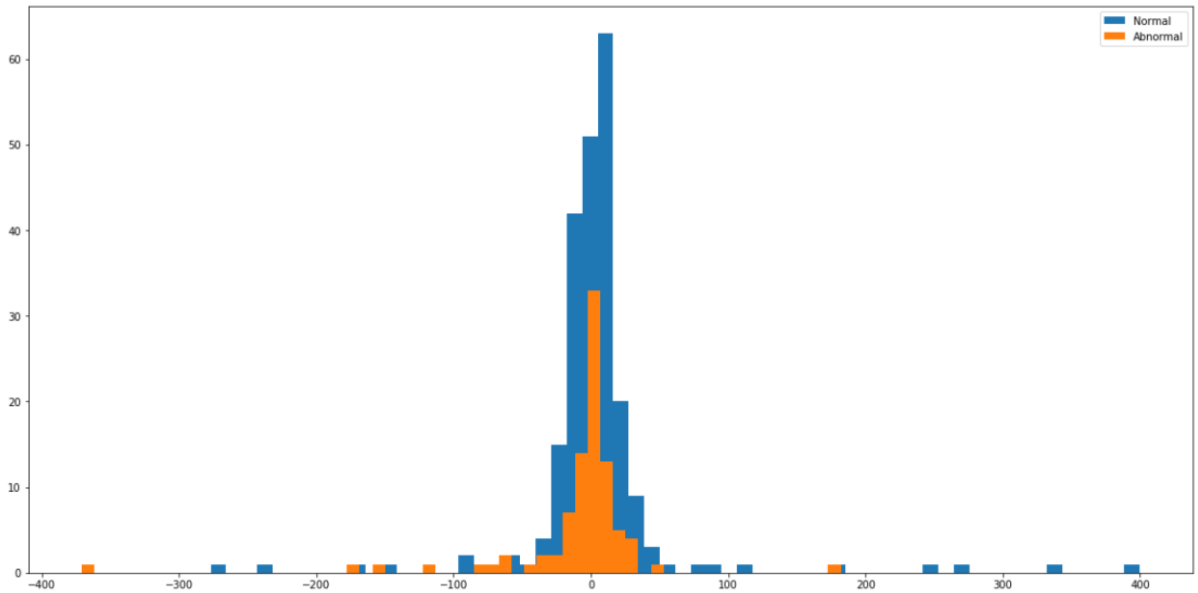
Histogram for Beat Interval after 8-Point FFT



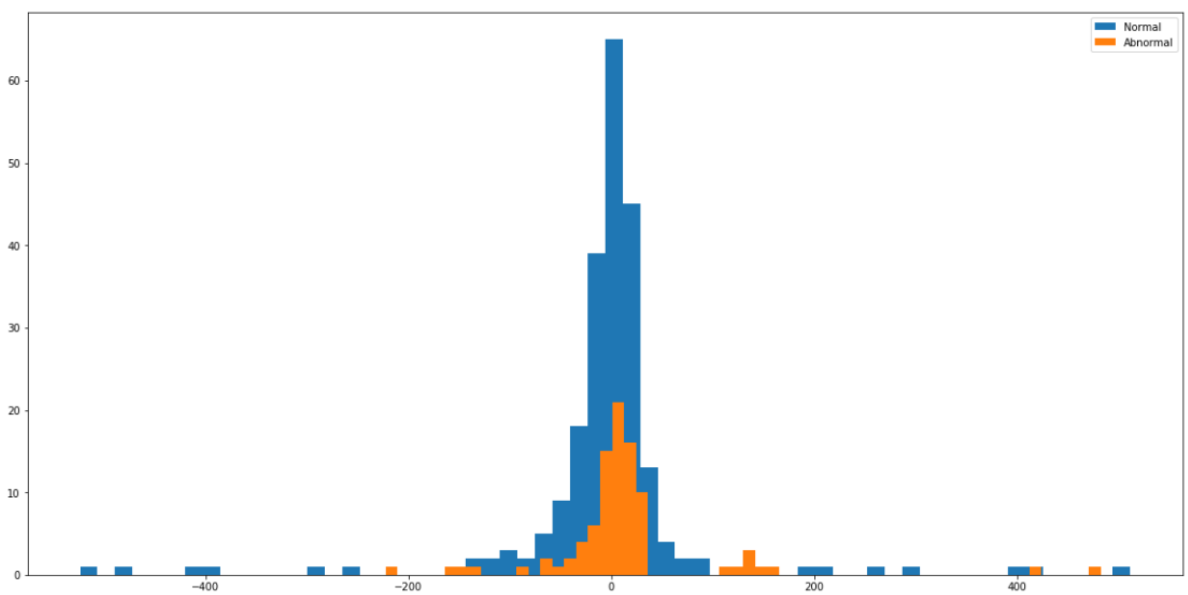
Histogram for Fourier Coefficient-1 after 8-Point FFT



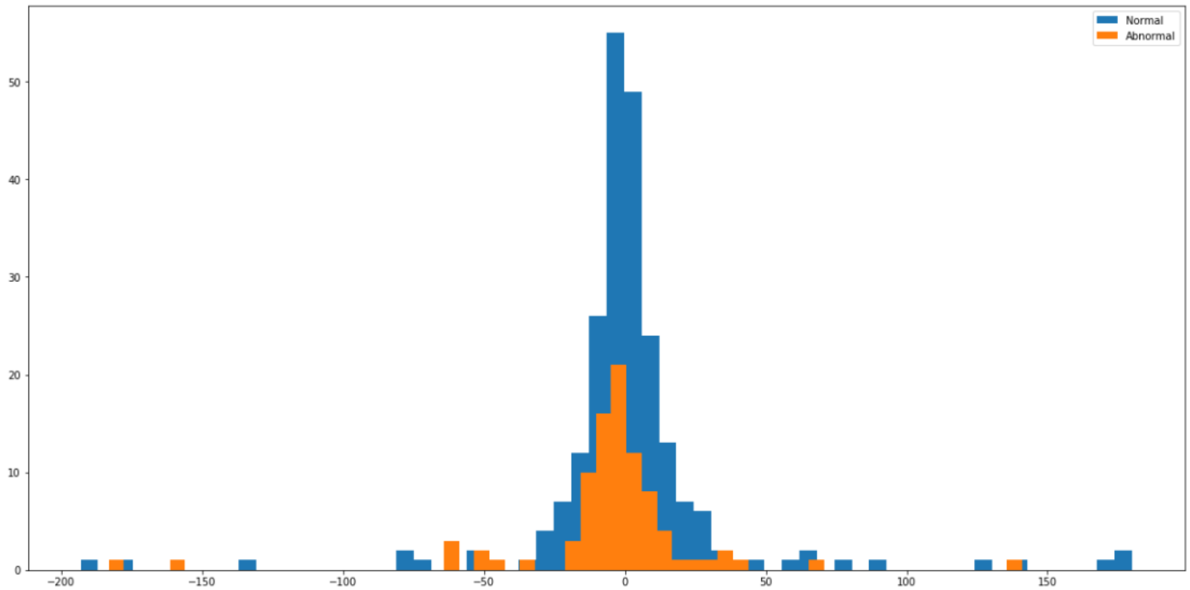
Histogram for Fourier Coefficient-2 after 8-Point FFT



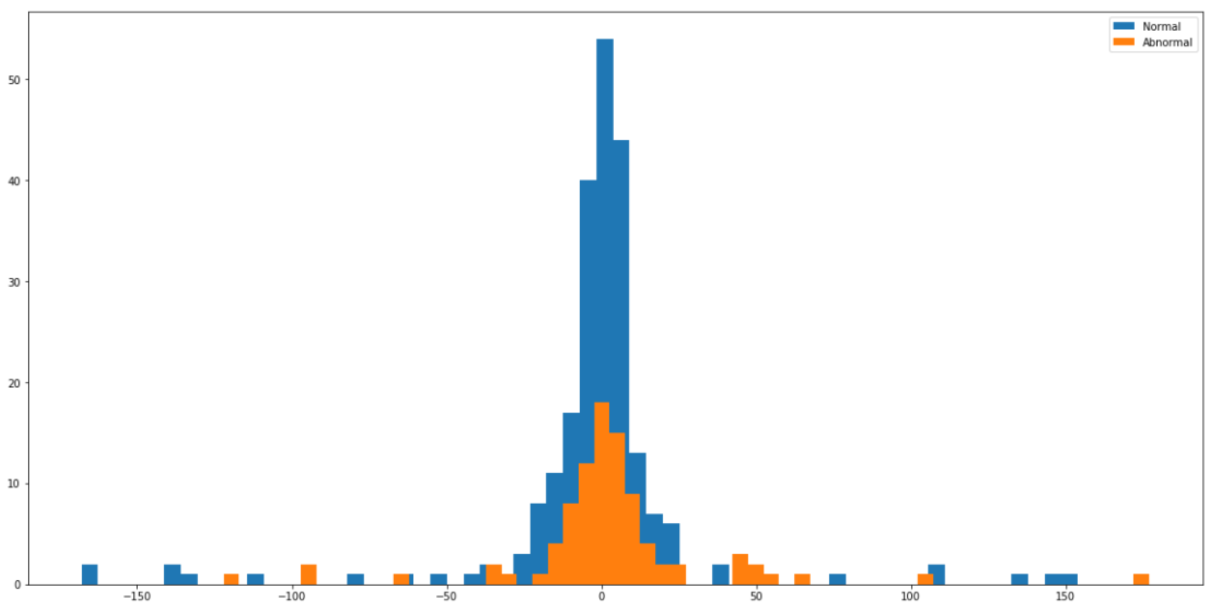
Histogram for Fourier Coefficient-3 after 8-Point FFT



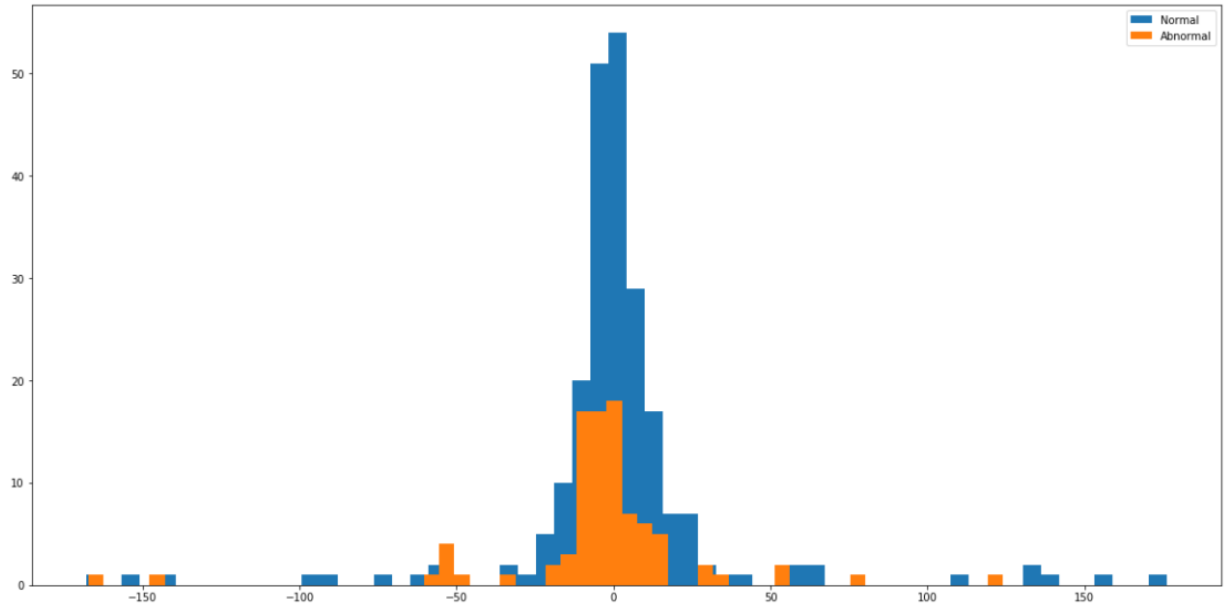
Histogram for Fourier Coefficient-4 after 8-Point FFT



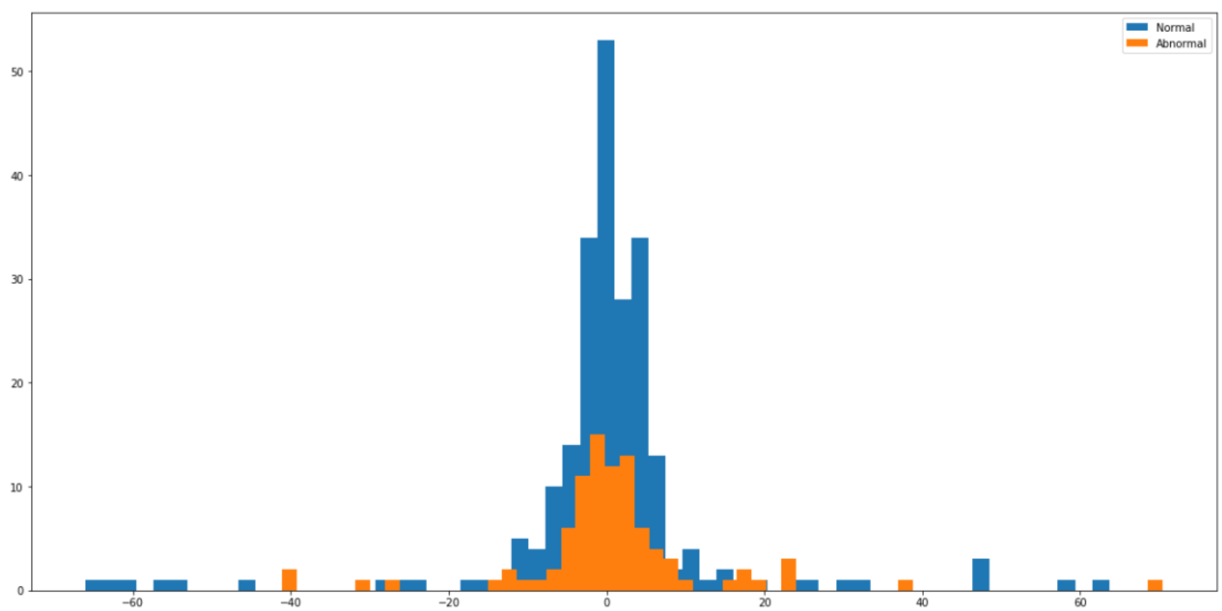
Histogram for Fourier Coefficient-5 after 8-Point FFT



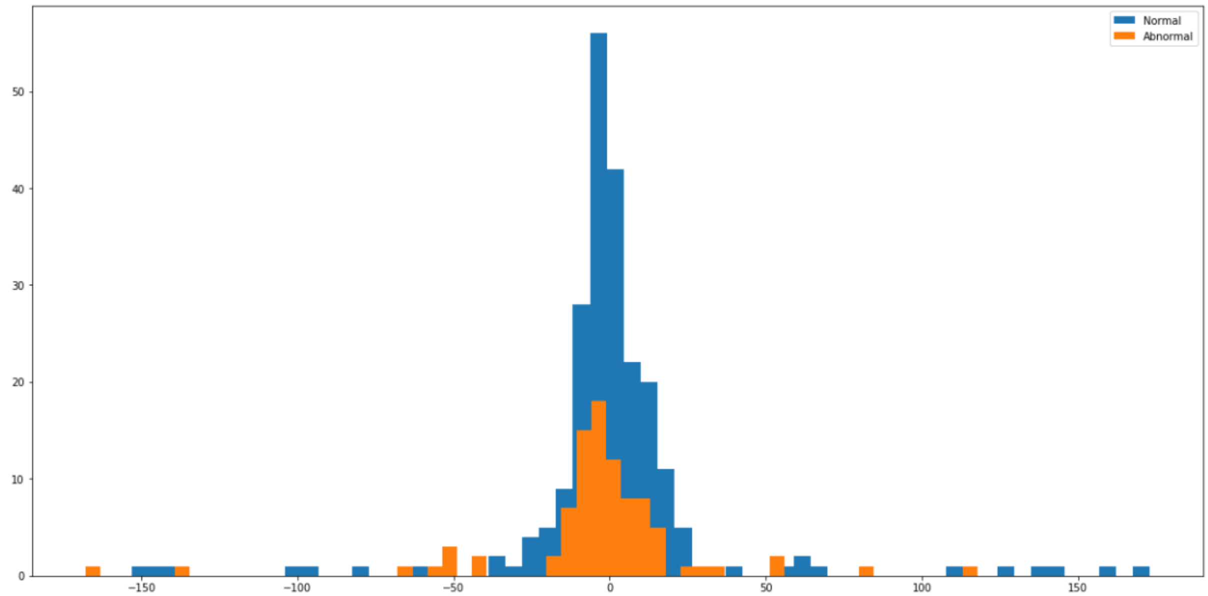
Histogram for Fourier Coefficient-6 after 8-Point FFT



Histogram for Fourier Coefficient-7 after 8-Point FFT



Histogram for Fourier Coefficient-8 after 8-Point FFT



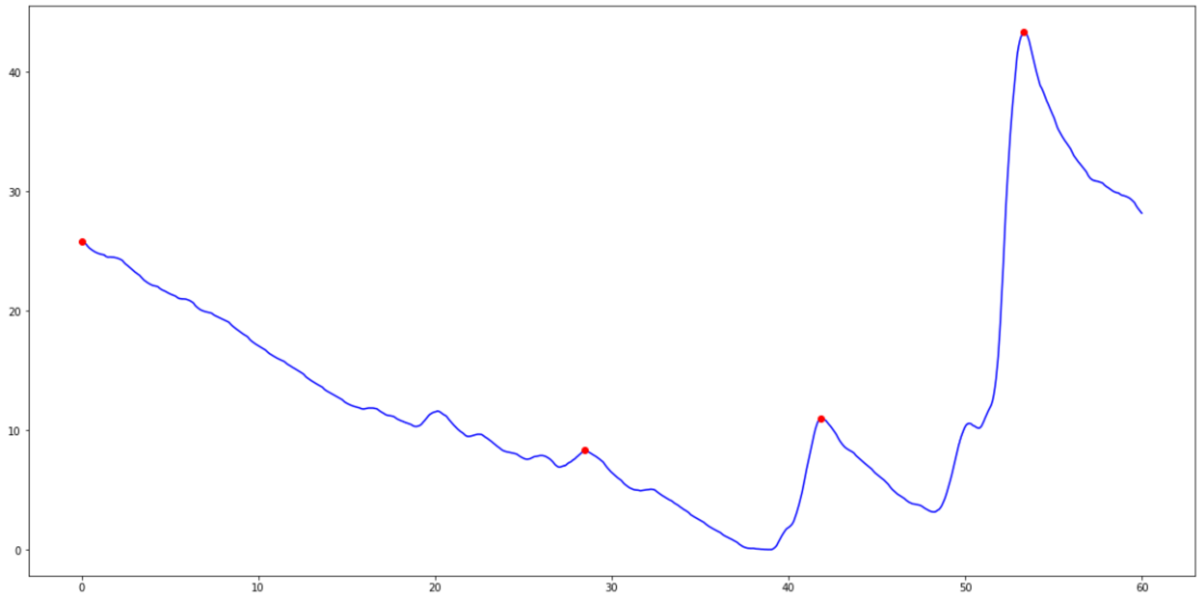
Bonus Ttest:

```
Ttest for Beat Count: Ttest_indResult(statistic=4.661003172536448, pvalue=4.652167257838253e-06)
Ttest for Beat Interval: Ttest_indResult(statistic=-0.6752256180967557, pvalue=0.5000277689559829)
```

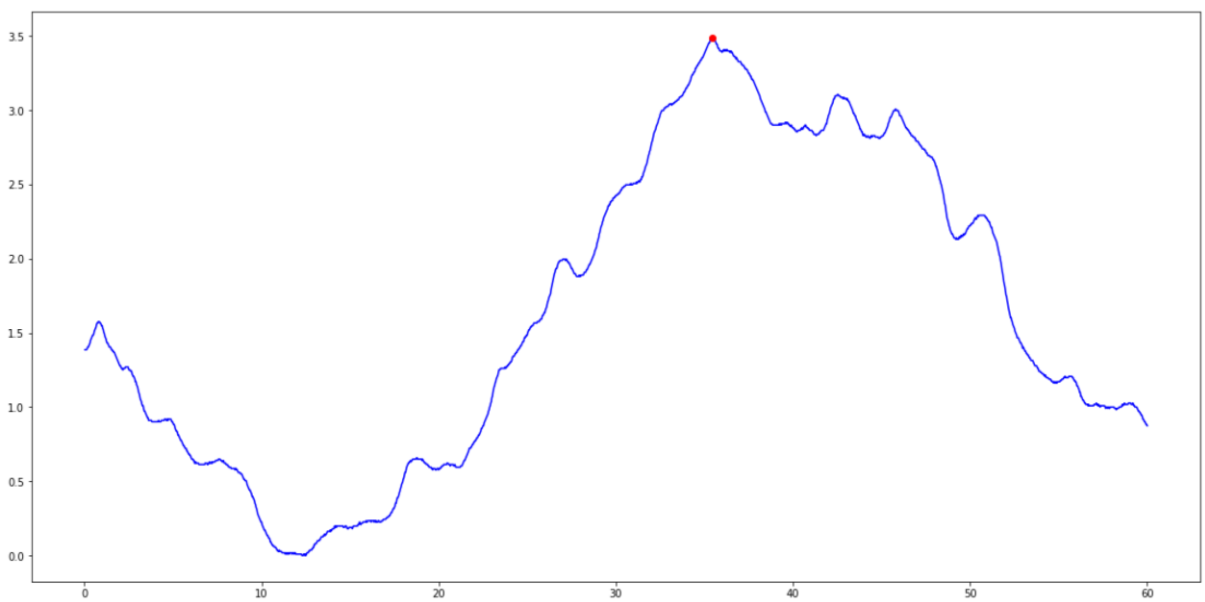
```
Ttest for Fourier Coefficient 1: Ttest_indResult(statistic=array([-2.06422438]), pvalue=array([0.03981491]))
Ttest for Fourier Coefficient 2: Ttest_indResult(statistic=array([1.85467942]), pvalue=array([0.06457634]))
Ttest for Fourier Coefficient 3: Ttest_indResult(statistic=array([-1.46617743]), pvalue=array([0.14359749]))
Ttest for Fourier Coefficient 4: Ttest_indResult(statistic=array([1.56632311]), pvalue=array([0.11827699]))
Ttest for Fourier Coefficient 5: Ttest_indResult(statistic=array([-0.81017369]), pvalue=array([0.41845198]))
Ttest for Fourier Coefficient 6: Ttest_indResult(statistic=array([1.31950913]), pvalue=array([0.1879572]))
Ttest for Fourier Coefficient 7: Ttest_indResult(statistic=array([-0.57401882]), pvalue=array([0.5663647]))
Ttest for Fourier Coefficient 8: Ttest_indResult(statistic=array([1.20305719]), pvalue=array([0.22985762]))
```

The Low Probability Values of the Ttest Results tell us that there is very less chance of Similarity between the Normal and Abnormal Values for the different features extracted.

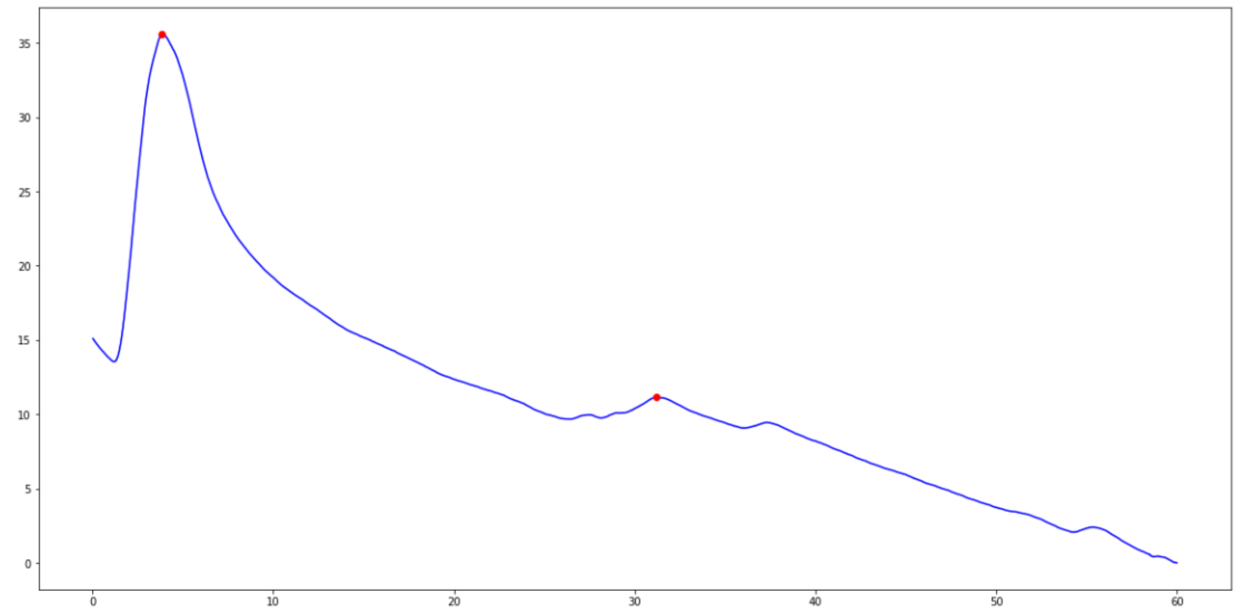
Q2)
EDA Signal Visualization
EDA Signals of Participants 1,2,3,4 with Trial = 20 were plotted.



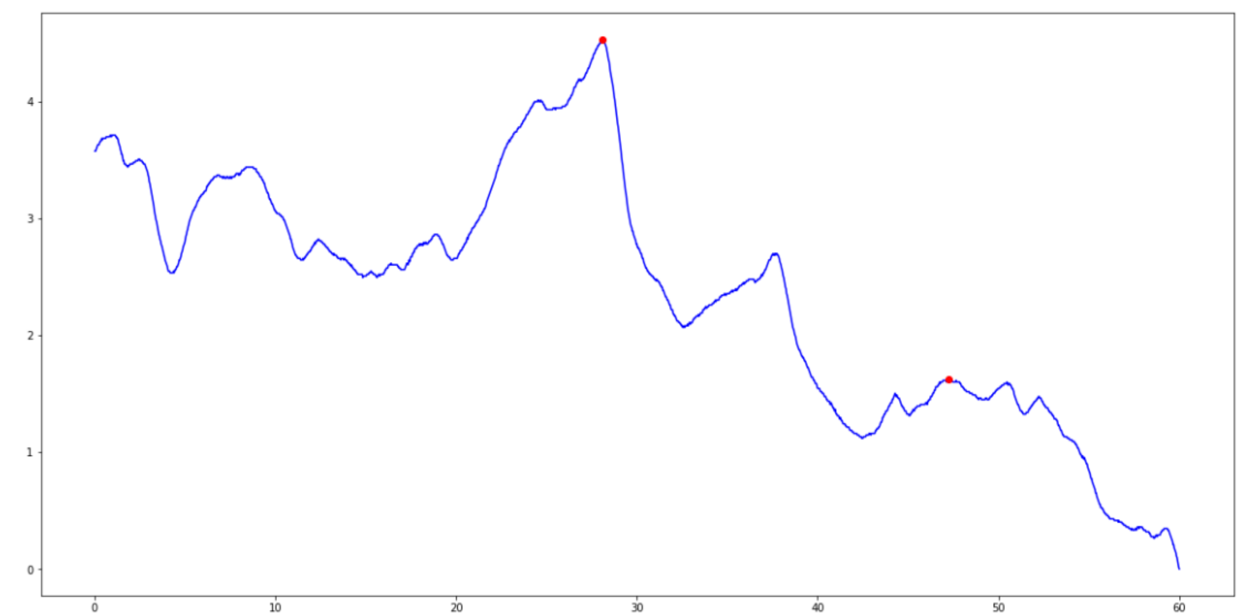
Displaying EDA Signal of Participant 1 – Trial 20
Red Dots are the Marked SCRs



Displaying EDA Signal of Participant 2 – Trial 20
Red Dots are the Marked SCRs



Displaying EDA Signal of Participant 3 – Trial 20
Red Dots are the Marked SCRs



Displaying EDA Signal of Participant 4 – Trial 20
Red Dots are the Marked SCRs

I observed varying levels of Valence and Arousal in all the 4 plots. There is a high amount of variability in the data.

Feature Extraction

For Participant – 1:
SCL = Mean EDA = 13.569166960416668
SCR Frequency = 19
SCR Amplitude = 14.315322315789475

For Participant – 2:
SCL = Mean EDA = 1.624496295833333
SCR Frequency = 43
SCR Amplitude = 15.310941395348838

For Participant – 3:
SCL = Mean EDA = 11.3343698390625
SCR Frequency = 21
SCR Amplitude = 12.541862857142855

For Participant – 4:
SCL = Mean EDA = 2.36254698125
SCR Frequency = 47
SCR Amplitude = 12.493362765957448

Co-Relation

Created Dataframe:

	Id	Trial	Average EDA	SCR Frequency	SCR Amplitude	Valence	Arousal
0	1	20	13.569167	19	14.315322	7.35	6.95
1	2	20	1.624496	43	15.310941	8.01	7.10
2	3	20	11.334370	21	12.541863	7.91	2.97
3	4	20	2.362547	47	12.493363	1.77	2.06

Correlation Values:

	Id	Trial	Average EDA	SCR Frequency	SCR Amplitude	Valence	Arousal
Id	1.000000	NaN	-0.504865	0.550161	-0.766833	-0.722897	-0.922513
Trial	NaN	NaN	NaN	NaN	NaN	NaN	NaN
Average EDA	-0.504865	NaN	1.000000	-0.982611	-0.157720	0.466393	0.135970
SCR Frequency	0.550161	NaN	-0.982611	1.000000	0.073216	-0.612670	-0.205172
SCR Amplitude	-0.766833	NaN	-0.157720	0.073216	1.000000	0.559370	0.954951
Valence	-0.722897	NaN	0.466393	-0.612670	0.559370	1.000000	0.659384
Arousal	-0.922513	NaN	0.135970	-0.205172	0.954951	0.659384	1.000000

As we can see in the highlighted portions, there are some significant correlations between the Features and Self-Reported Ratings. We can improve on the correlations by using Linear, Logistic or Polynomial Regressions, PCA, i.e. Principal Component Analysis and other Machine Learning Techniques.