Patient 1(953):

My data file is Harding.dat. So I am using dlmread Matlab function to read the data.And ploting this data starting from 2 , because the 1st row data is patient ID.

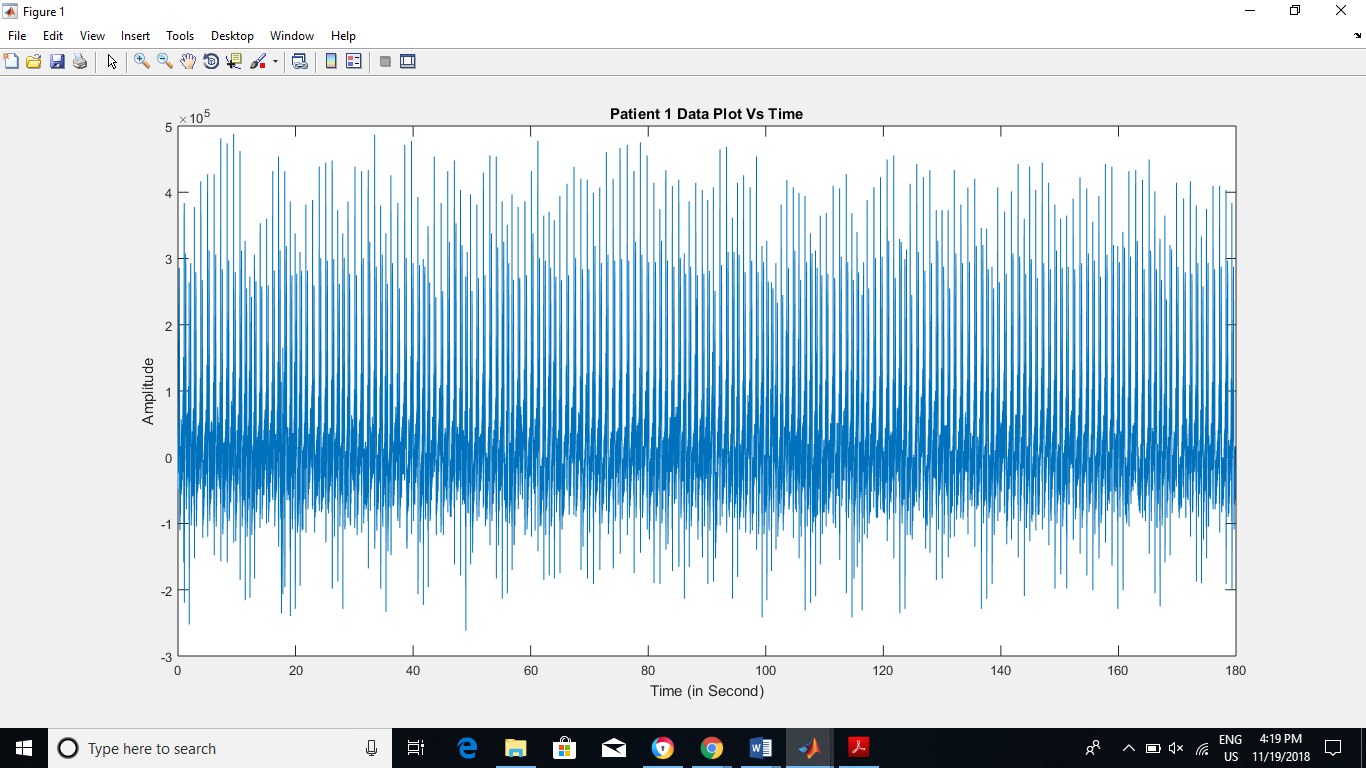


Fig 1: Patient 1 Data Plot

My first approach to evaluation of this data by converting it from time domain into frequency domain. So I am performing FFT for this conversion;

nfft=length(p1);

nfft2=2.^nextpow2(nfft); % because FFT conversion is working with 2^n form, %So finding nearest power of 2 with respect to nfft.

fy=fft(p1,nfft2);% performing fft

fy=fy(1:nfft2/2);

xfft=Fs.\*(0:nfft2/2-1)/nfft2;% Perform 1 side FFT

plot(xfft,abs(fy/max(fy)));% plot the fft.

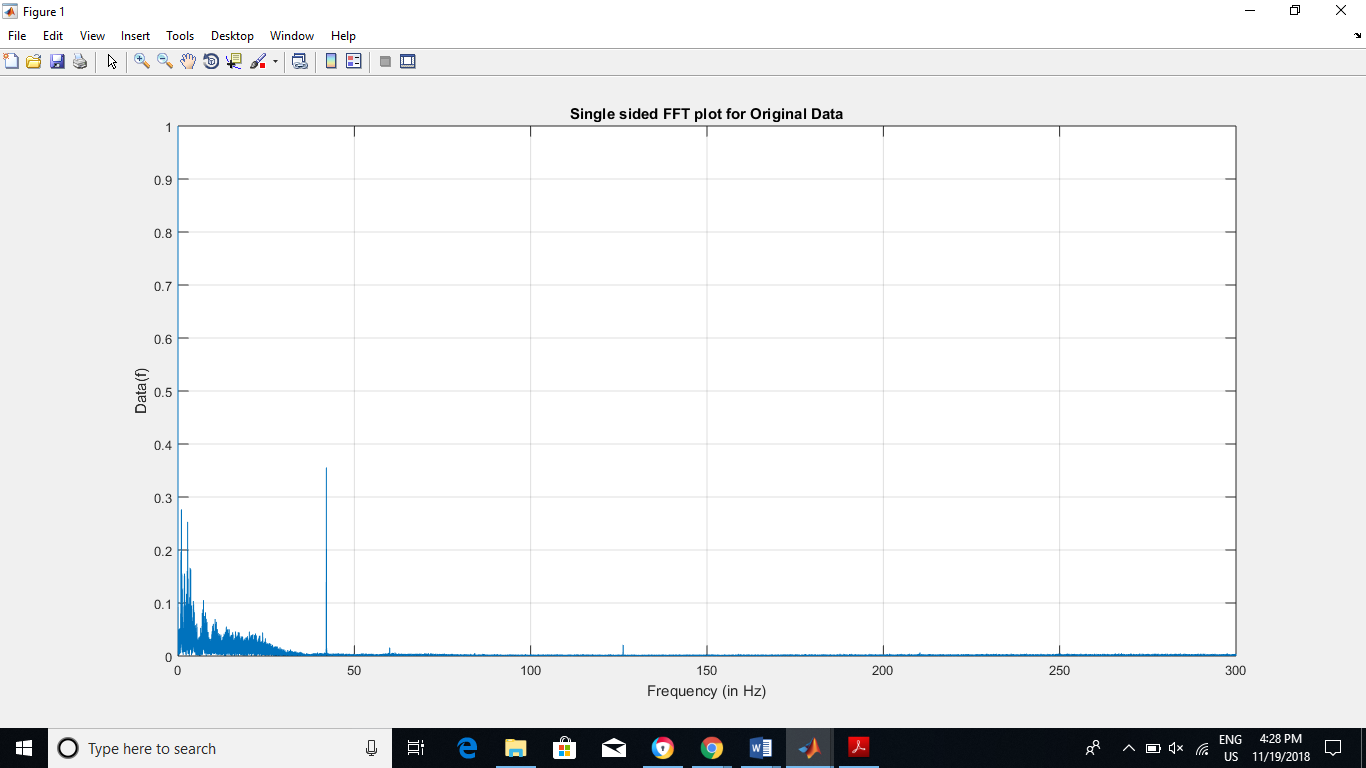


Fig 2: FFT plot for original data

BY converting in the frequency domain, we know that upto 35 Hz frequency is very much important for ECG data. From the readings, 90% spectral energy of the signal focus on 0.25 Hz to 35 Hz. So we cannot use the filter within this range.If we use the filter in this range we lose some of the data of ECG that might be very much important. But we have some baseline drift due to breating and muscle motion is basically between 0.05-2Hz. And powerline interference is at 60Hz.From our FFT plot we can see that near 42Hz we have power line interference.SO we can see in fig 2. It shows that we need to use the Butterworth band stop filter [42,42.1] Hz of order 1. And next we need to remove the high frequency noise. So for this I am going to use Butterworth low pass filter of 43Hz of order 2.



Fig 3: Filtered data plot

After applying the filter, We can easily detect the RR Peak Interval. After filtering the original data, we must analyze this data in Time domain.1st we need to find RR Interval. So for that I am using direct find peak function with giving the window of peak distance which is 0.75s.Here is the MATLAB code for that:

select=y1(1:length(y1));% to select the range of the data

[ygh,x]=findpeaks(select,600,'Minpeakdistance',0.75);% find peak function is finding the highest peak in the window of 0.75s. And this window is going forward after computing that range peak.

By visual inspection, using this window is appropriate to find the RR peak. For finding the Mean Heart Rate, I am Using the equation Heart Rate=60/mean (RR Interval).

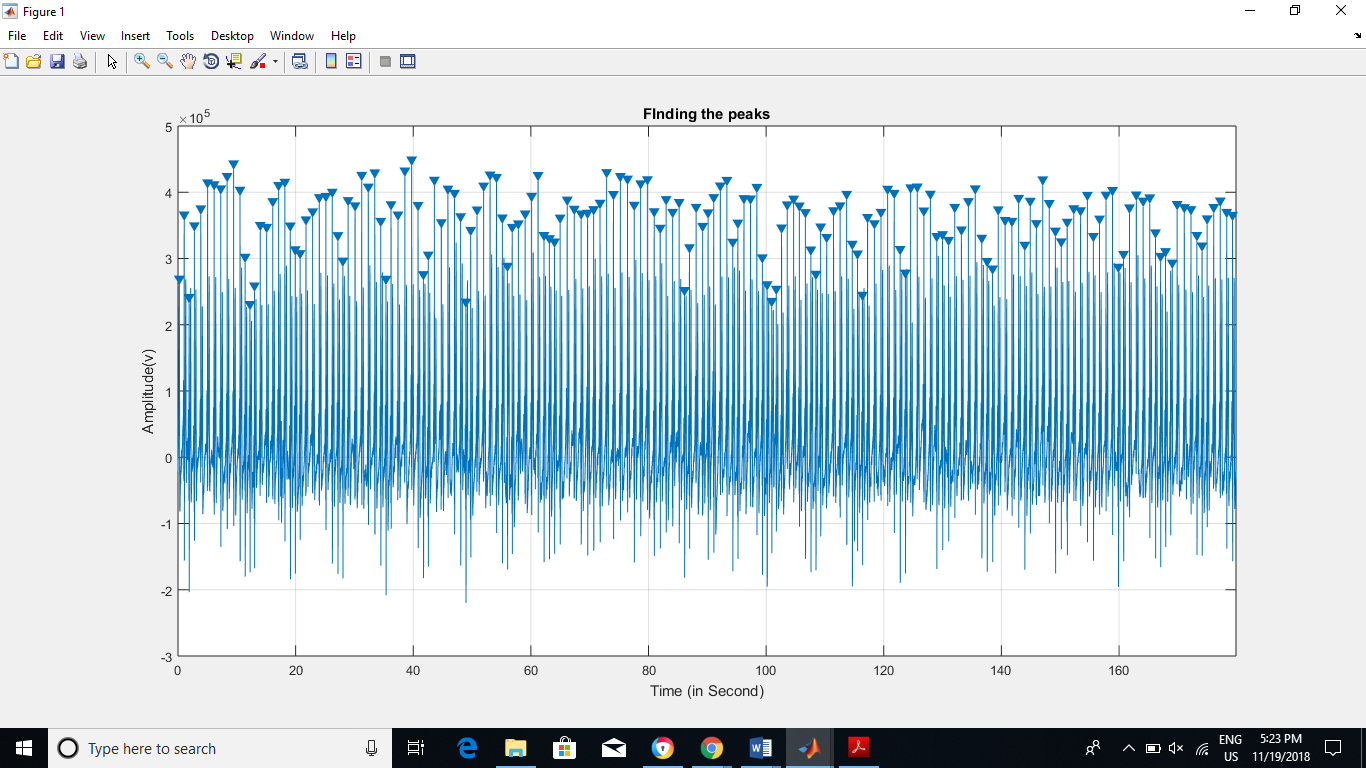


Fig 4: Finding the peak with the window of 0.75s.

And then Calculate the RR interval using this function which is shown below. Where x = vector of peaks of R Curve. Xi is the vector of R-R interval. Here is the MATLAB code for finding RR interval.

i=1:1:size(x)+1;

for i=1:1:size(x)-1

xi(i)=x(i+1)-x(i);% Subtracting the peaks to find the R-R interval in %seconds.

hr(i)=60/xi(i);% to find the heart rate at individual points

End

And this is the code to find the statistics for HRV.

meanNNInterval=sum(xi)/(length(xi)\*1000) % In ms.

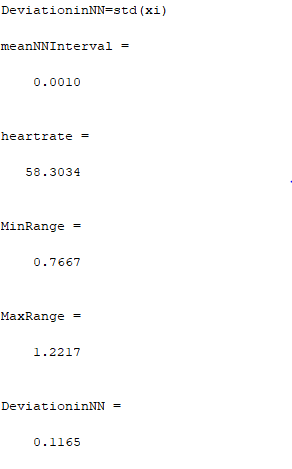
heartrate=60/(meanNNInterval\*1000) in s.

MinRange=min(xi)

MaxRange=max(xi)

DeviationinNN=std(xi)

And the output is:



than we must analyze the frequency domain stuff. For this First I have plotted the RR interval. SDNN is present in seconds. In the code it is in ms.

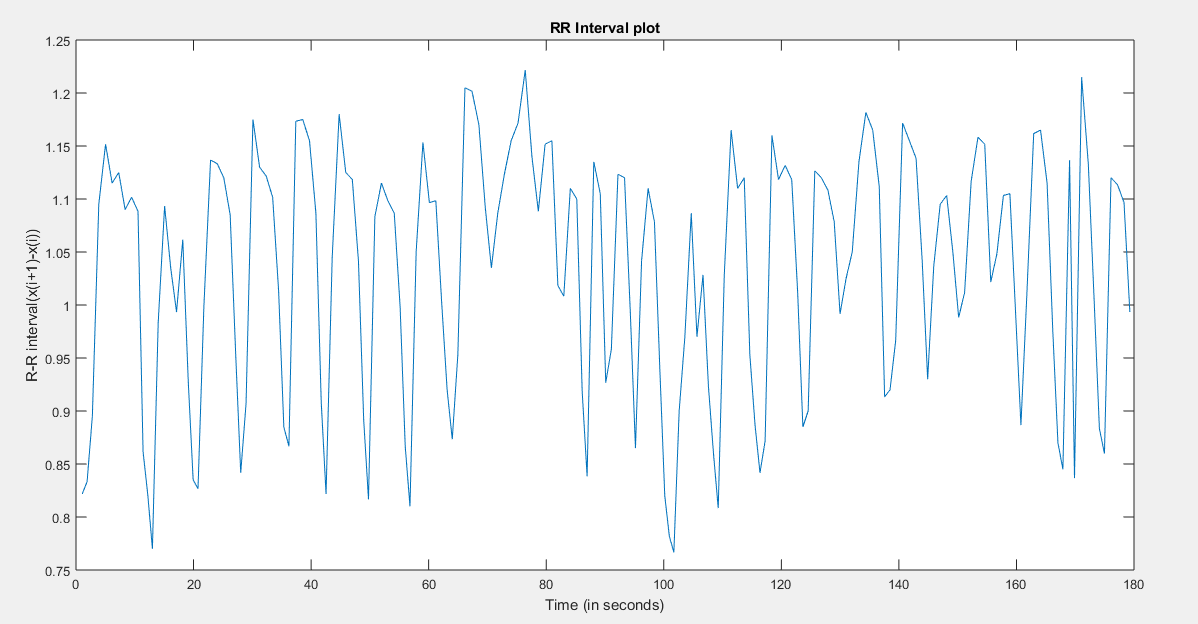


Fig 5: Plot of R-R Interval plot.

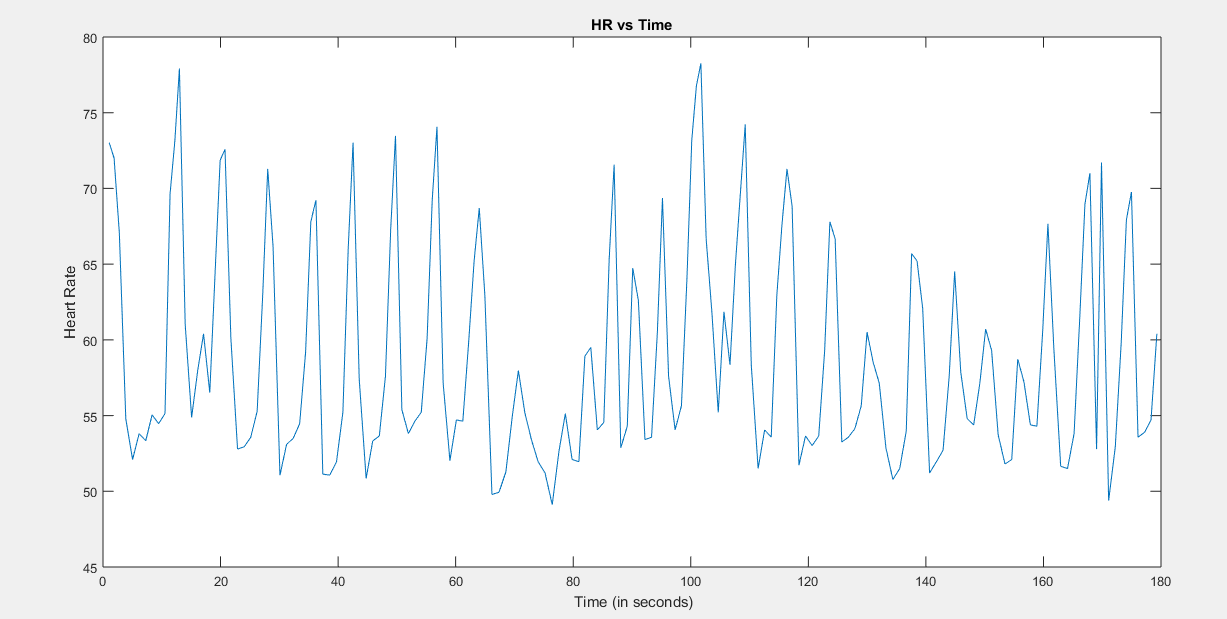


Fig 5a: HR plot vs time

After than I am going to find the change of the RR Interval. By using this technique, we can see that how much data is fluctuated.

. And by plotting this graph we can see the non-uniformity in the plot. So in figure 6, we can see that after every detection of peak, here if t1=0.98s and t2=1.02s. So this showing change is T=t2-t1. And it is 0.02. that is plotting in fig 6.

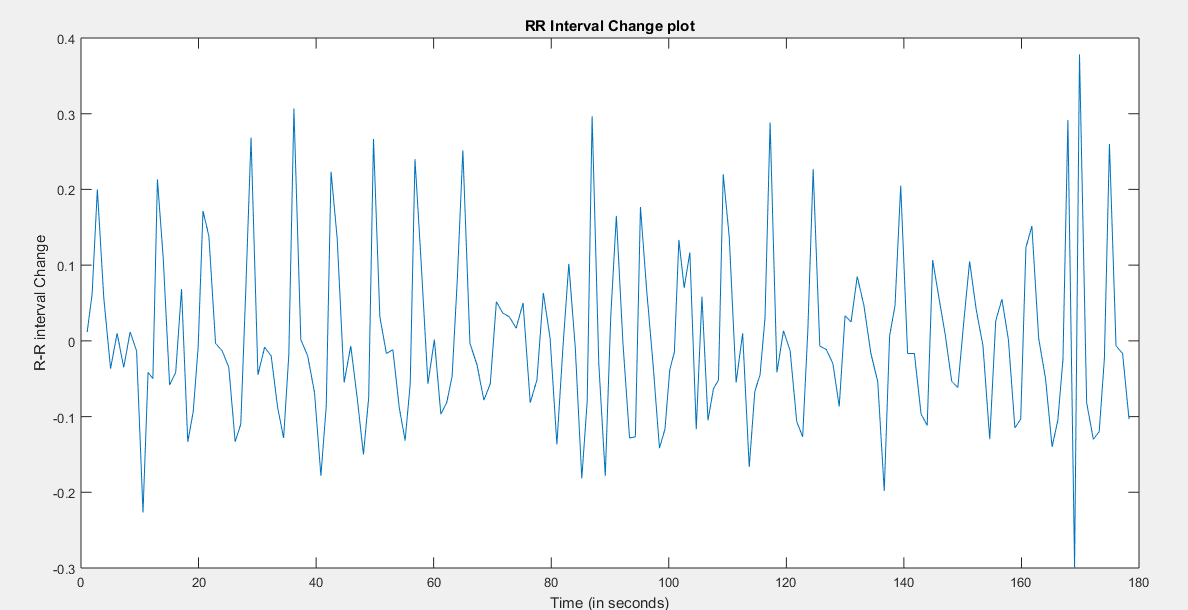


Fig 6: Change of R-R interval plot:

This non-uniformity leads us to resample this data, that will evenly distributed. This means that this data will follow some pattern. For this, I am Directly using resample MATLAB function to resample the data. The MATLAB function is showing below.

Xcs=resample(xi,rsf,1);% doing the resampling with p/q ratio. Where P is my resampling frequency.

And in Fig 7, I am plotting without resampling and with resampling data. Here I am resampling my data at 20Hz, Which is above the Nyquist frequency. Here we can sample at any rate above the Nyquist plot. But by using the20Hz, I am getting good resolution of the signal.

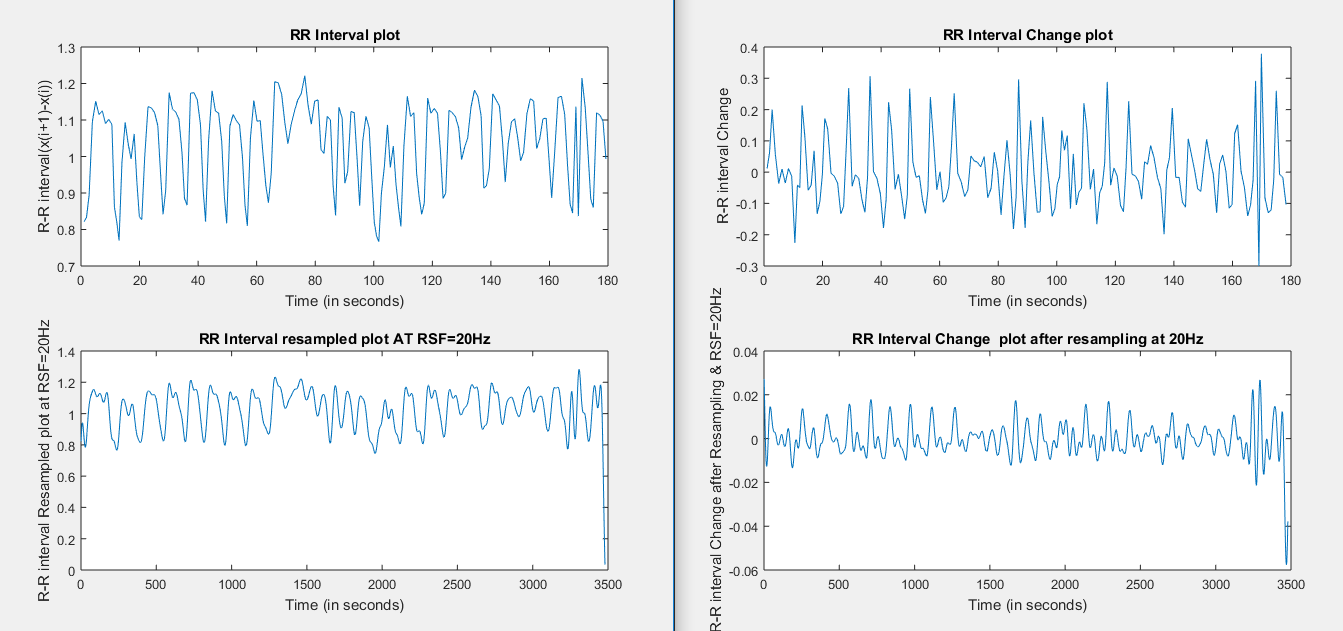


Fig 7: Resampled and Before resampled plot of RR Interval with the change in RR Interval.

In the fig 7, you can see in right top figure, the change is 0.2 seconds between 2 peaks. But after resampling it should be 0.02s (Right bottom). SO if we are comparing this 2 plot, the right bottom plot is uniformly resampled. and in left bottom plot you can see it is uniformly sampled.

After this step, we must plot PSD of this data. So for this, My approach to use FFT to find the periodogram. And for this my MATLAB code is:

N = length(Xcs);% Xcs is my Sampled data vector

xdft = fft(Xcs);

xdft = xdft(1:N/2+1);

psdx = (1/(Fs\*N)) \* abs(xdft).^2;% Finding the power from FFT.

psdx(2:end-1) = 2\*psdx(2:end-1);

freq = 0:Fs/length(Xcs):Fs/2;% Nyquist Range

In fig 8, You can see that periodogram has some lower frequency components that we have to focus. And curve is smooth because I am not removing any the data from bottom left plot in fig 7.

After getting this curve, I am going to find out the coherence ratio, that is Peak Power/(Total Power-Peak Power). For finding the power, I am Using the trapz Function within in the range of ULF, LF, HF. And the results are showing below.

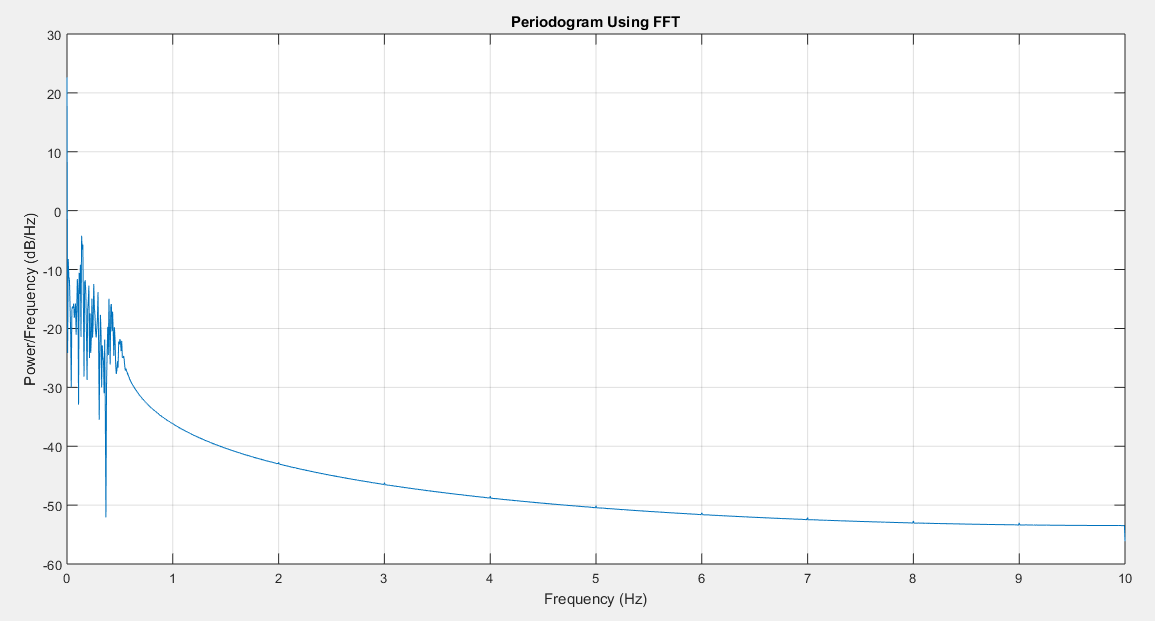
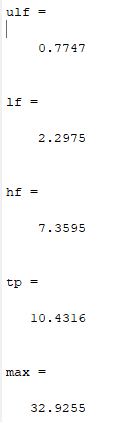
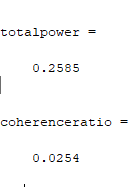


Fig 8: PSD of resampled data.

Points:

1. BY analyse the FFT, we can say that ECG data is bound in [1-35]Hz. After that it is noise. So we removed all the noise by using the filter to smooth the data.
2. Task 5 answer: mean HR is 58.3. which is quite low. And if we analyse the HR plot , we can see that there is one pattern where RR interval is increase and decreasing, and this pattern is follows most of the time. That indicates that patient has some breathing problem or other, that is why patient’s HR is increase and decrease in small time duration and this pattern is going on.
3. So HR is not stable.

Patient2(954):

First Plotting the Data Which is given:

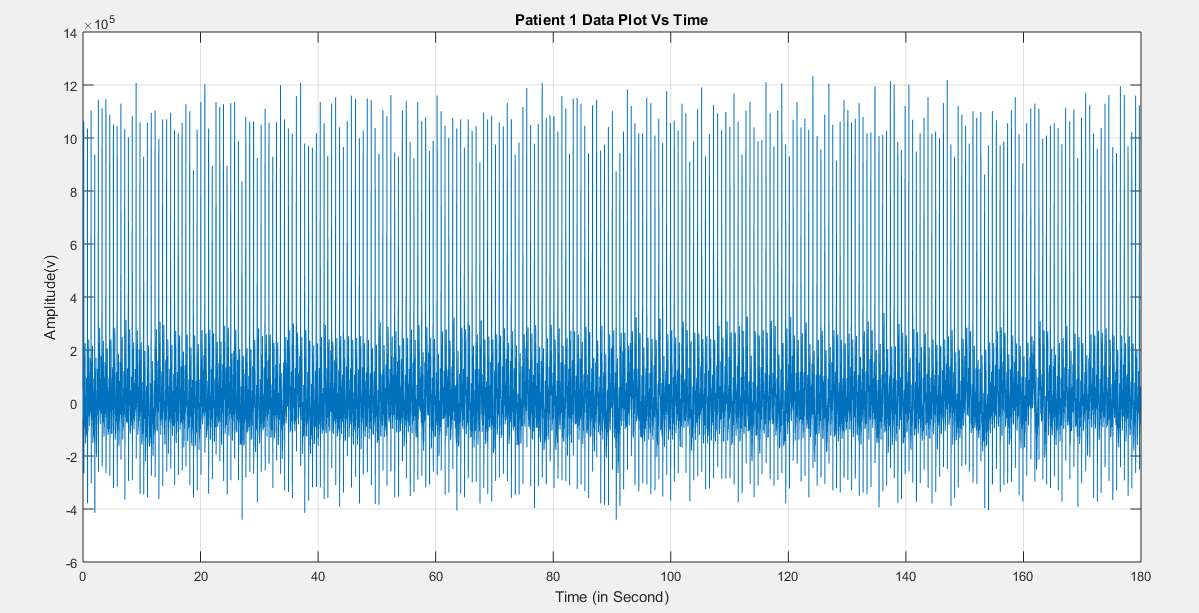


Fig 8: Plot of original data:

Using FFT function from the MATLAB to convert from time domain to frequency domain. Shown in fig 9.

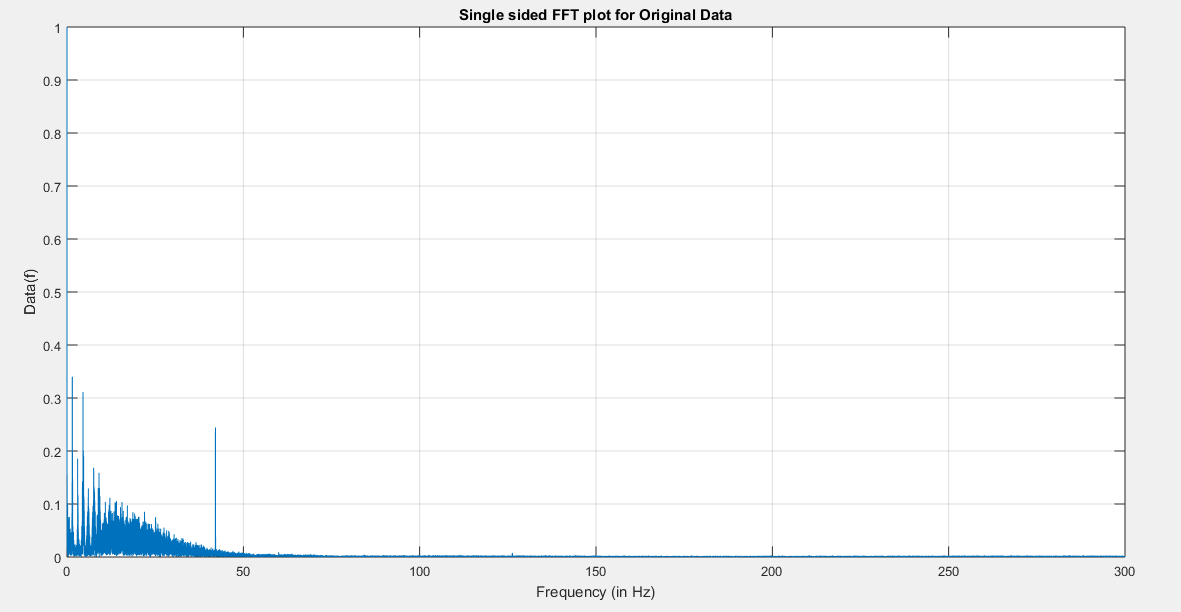


Fig 9: FFT of patient 2 data:

Form this plot we can see that the power interference noise is near 42Hz. So I am going to apply 2 filter.1) Band stop filter [42,42.15] Hz to remove power interference. 2) low pass filter of order 2 at 43Hz.Fig 10 shows these 3 different plots. And after applying this filter, we can easily detect the RR interval.

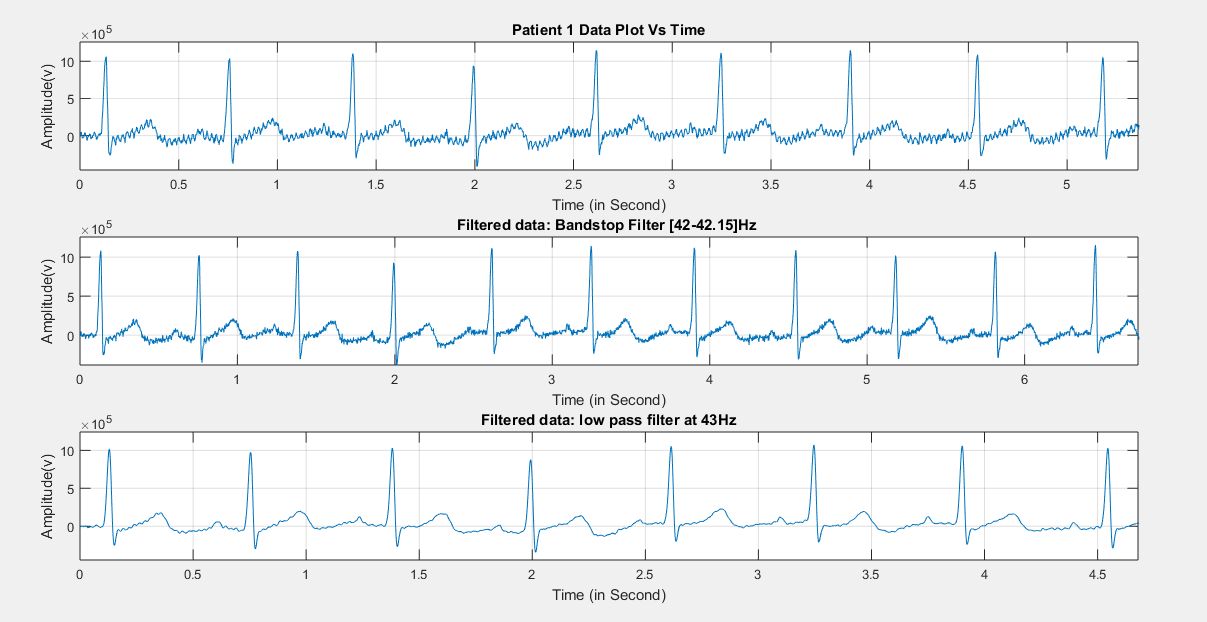


Fig 10: After Applying filter, the data is plotted.

As the MATLAB code which is shown above, I am finding the heart rate. And the heart rate data is shown in fig 11.

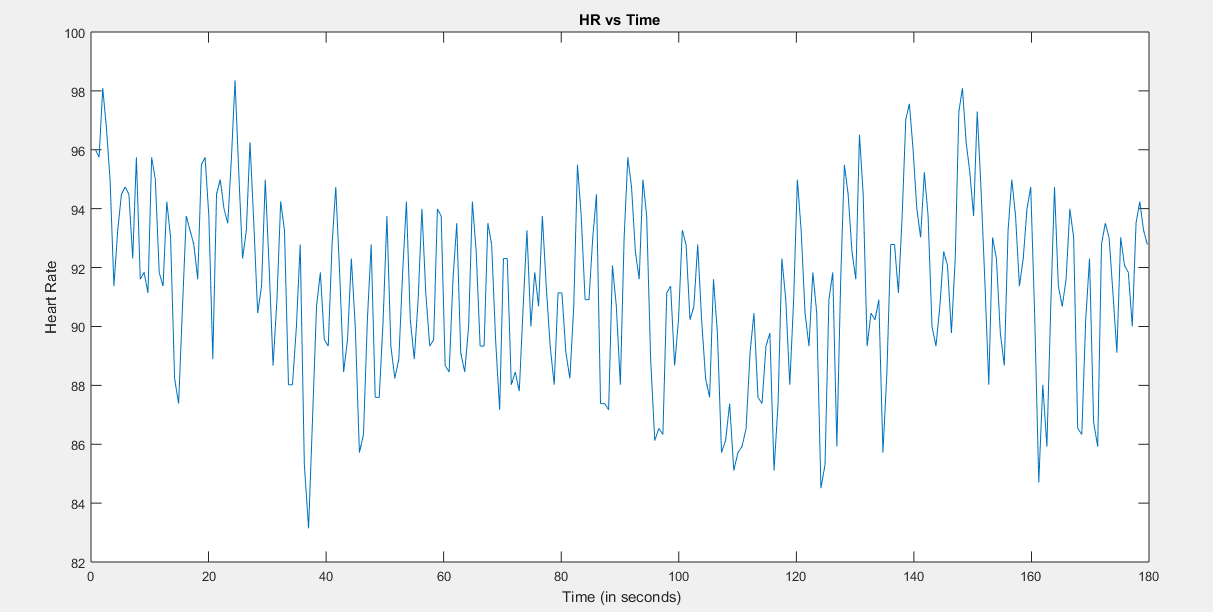


Fig 11: Heart rate data vs Time plot

For time domain analysis:

For finding the RR peaks, I am using the window of 0.5s of peak distance. This 0.5 second window can cover all RR peaks. This averaging window is basically depending on mean RR peaks. So according to this I am using this function. Fig 12 shows that RR peak curve.

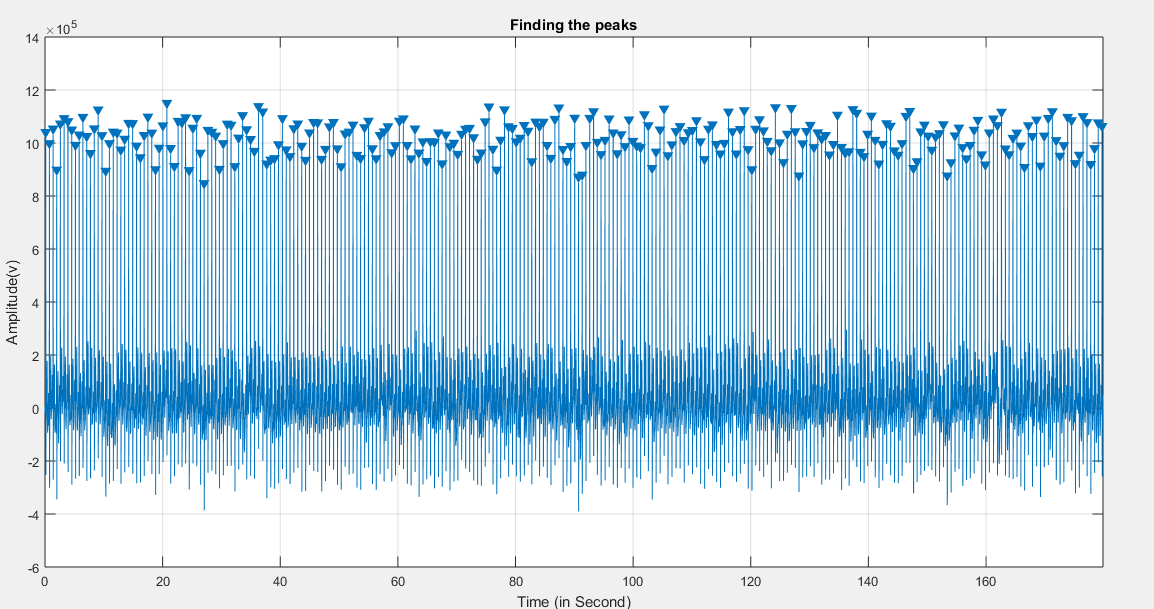
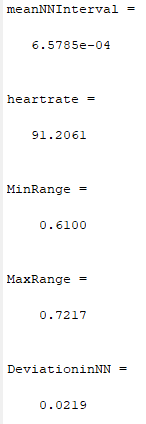


Fig 12: RR Peak plot of patient 2

After this these are the statistical results of HRV. And mean RR Interval is in ms(milli second).



For frequency domain analysis,

We are plotting beat techogram. That is shown in Fig 13.

After that I am plotting the RR Interval Plot.

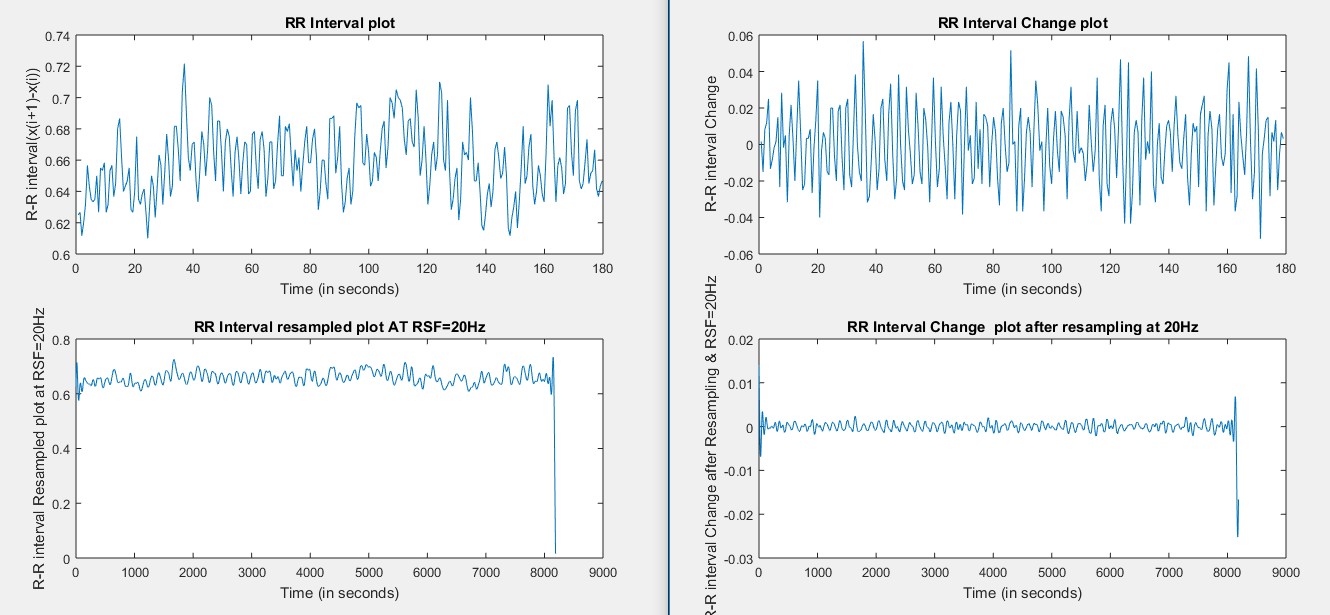


Fig 13: RR Interval Plot & RR Interval Change plot for without resampled data & With resampled data.

In fig 13, In this left top graph we can see the non-uniformity. But after resampling it as 20Hz (), you can see that most of the points are bound between [-0.005,0.005](right bottom). And that seems to be uniform as compare to left top.

After resampling the RR Interval data, I have plotted the graph of PSD. Which is shown in fig 14.

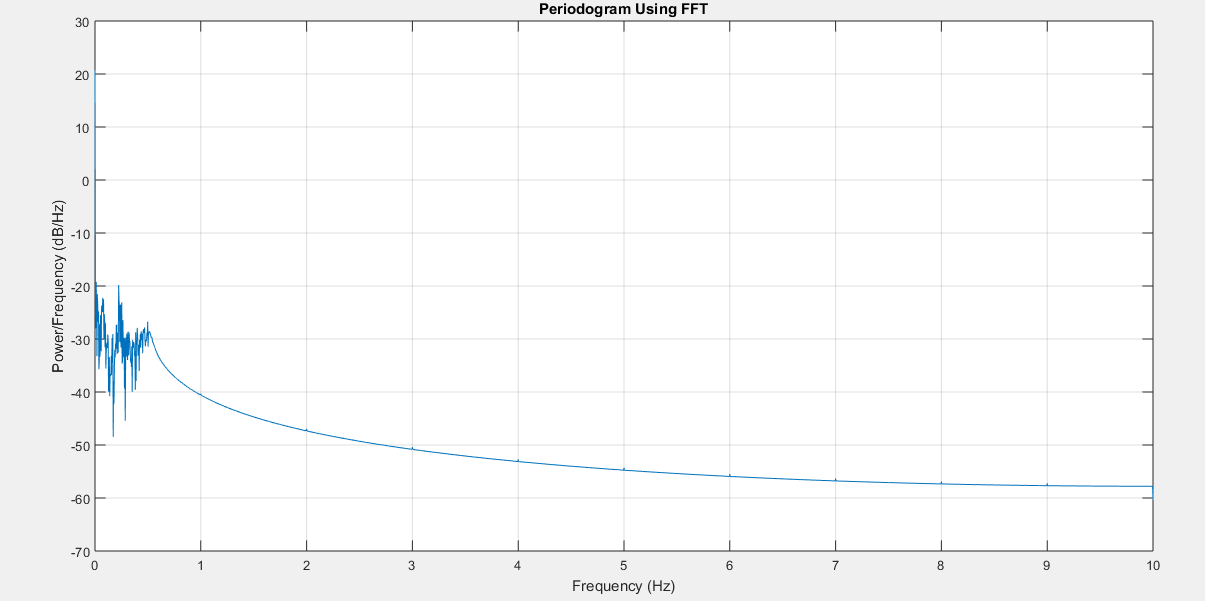
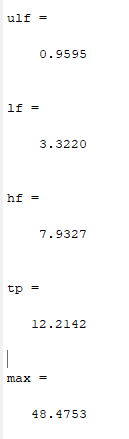
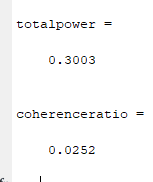


Fig 14: PSD of resampled data at 20Hz.

Here I am doing resampling at 20Hz again because, It is above the Nyquist bound and gave good resolution of the signal in the range of (0.15-0.4) Hz.

And the results of task 7 are:

Points:

1. In this data , after applying FFT, we can see that most of the ECG data lies the frequency range of [1-40]Hz.
2. Task 5 answer: The mean heart rate is 91.20 per minute. Which is quite high. So I can say that patient is contineoulsly doing some activity.and if we can see the HR plot, it is not stable.After 80 seconds, we can see so much change is RR time, So data is flucuating so much. We can see some stable HR in the initial stage of 0-30 s and 40-80s.but after 30 secnonds the RR time going little high.

Patient 3(955):

First the original ECG Data and the FFT of that time series ECG data is plotted.

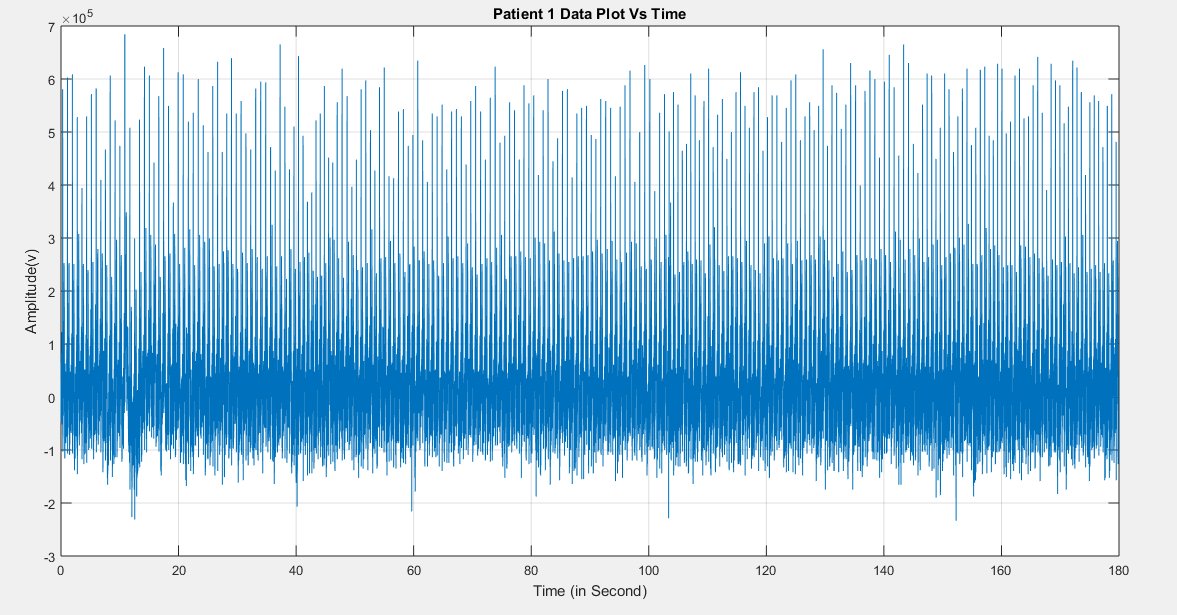
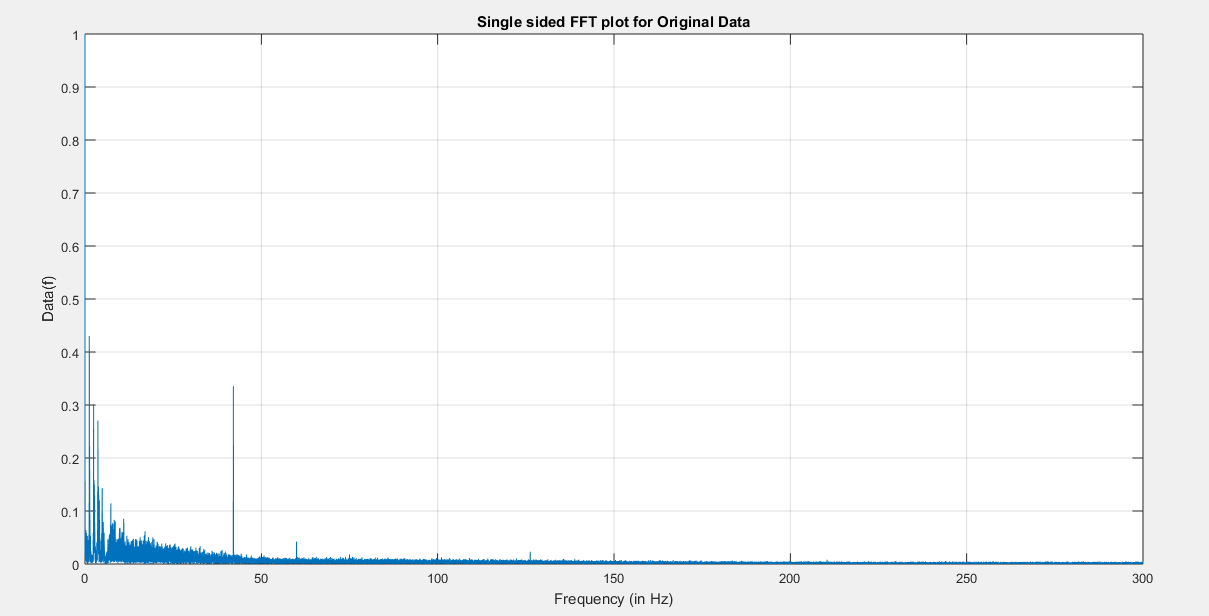


Fig 15: Patient 3 Real ECG data vs time plot

Fig 16: FFT of patient data 3:

After analysing FFt in Fig 16, I am applying 2 filters 1) TO stop power interference at [42.06,42.1] Hz with band stop filter and 2) low pass butterworth filter at 43Hz of order 2.

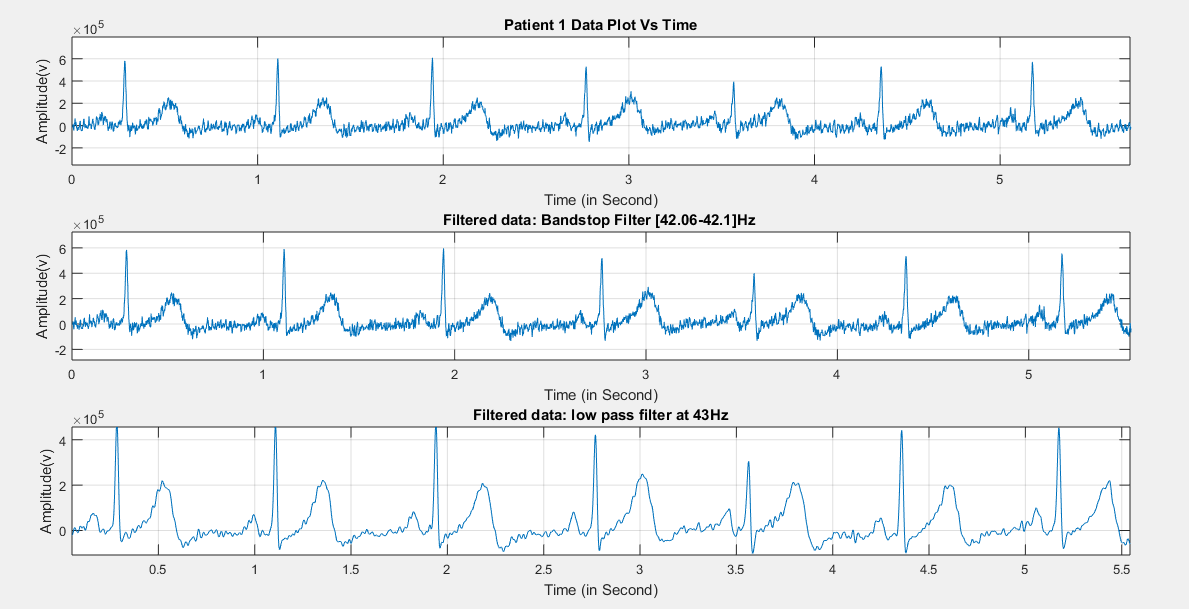


Fig 17: After aplying filter ECG data vs time plot of patient 3.

Then I found the RR peaks with the window of 0.7s window, because that can cover all my peaks.

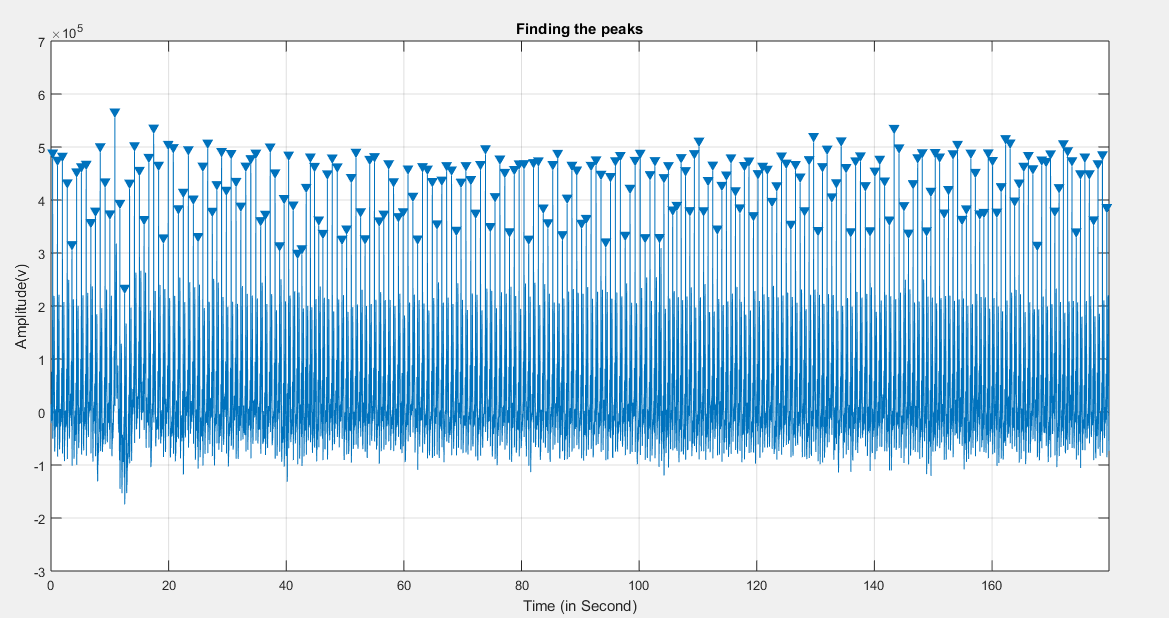
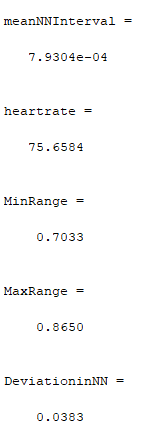


Fig 18: RR peak detection vs time plot.

For time domain statistical analysis results are shown below.



And fig 19 shows the plot of HR vs time.

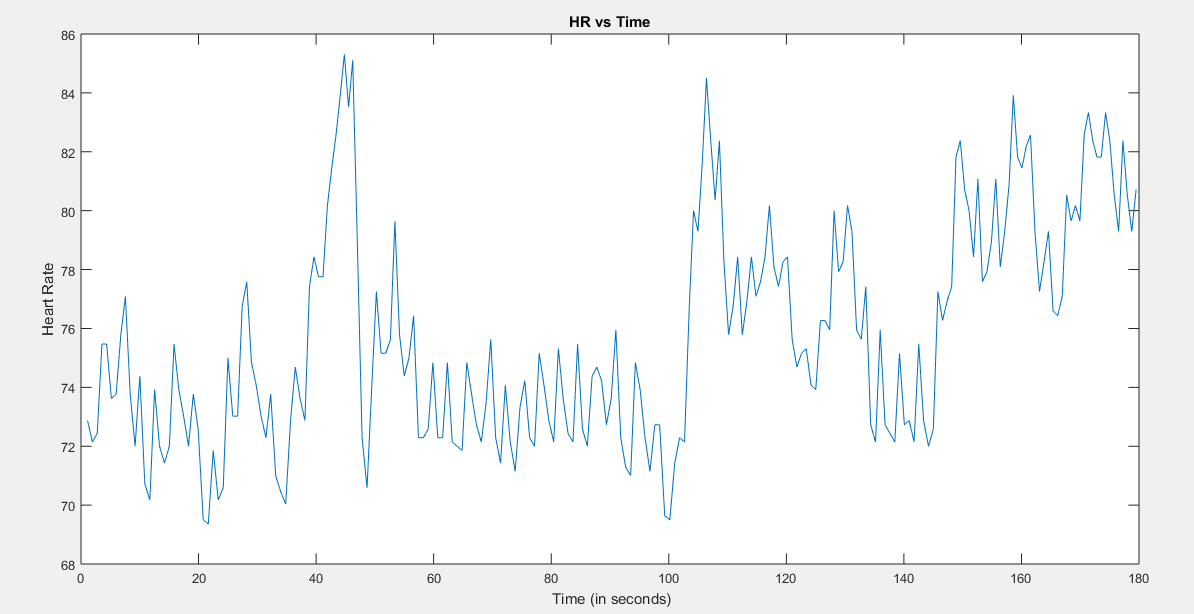


Fig 19: HR plot vs time

For frequency domain analysis I have plotted the graph of RR interval with respect to time. We can see in the fig 20, left upper plot is unsampled data and it is not uniformly distributed. So I resampled it at 20Hz, TO get the high resolution signal between required band [0.14-0.4]Hz.

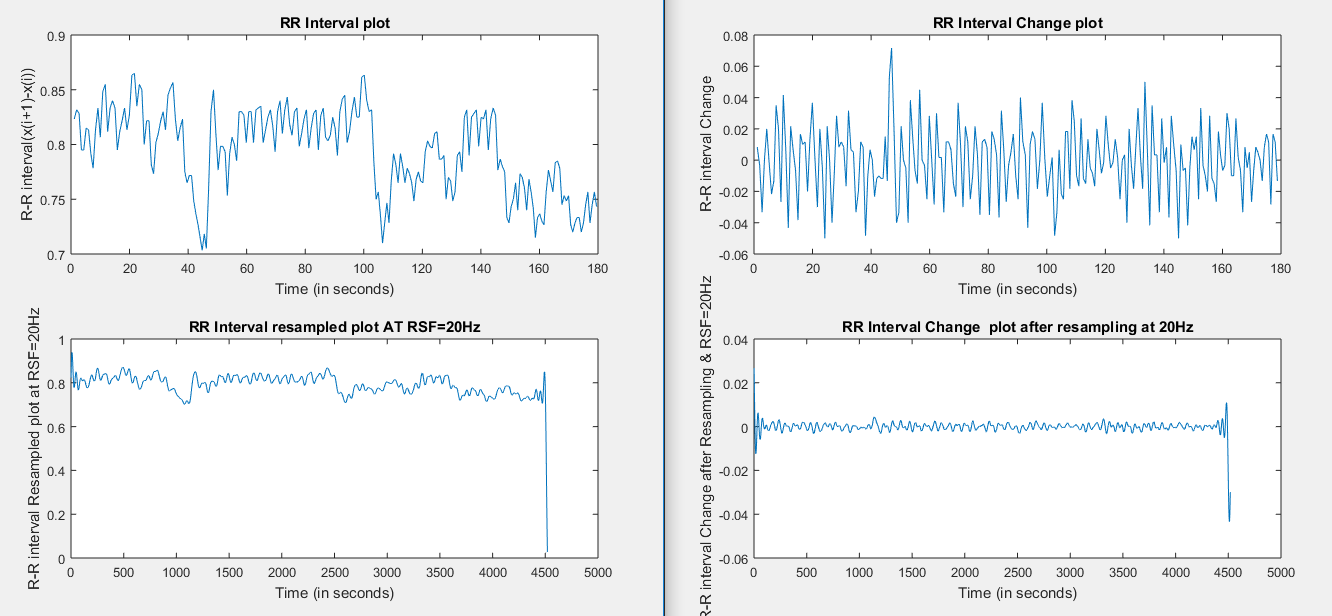


Fig 20: RR interval plot without sampled and resampled data.

Here in right both plot, we can see that after the resampling the data is uniformly distributed(right bottom), because that plot shows the change of time between 2 RR peak interval time.

And at last I plotted the PSD of resampled signal to find the power.

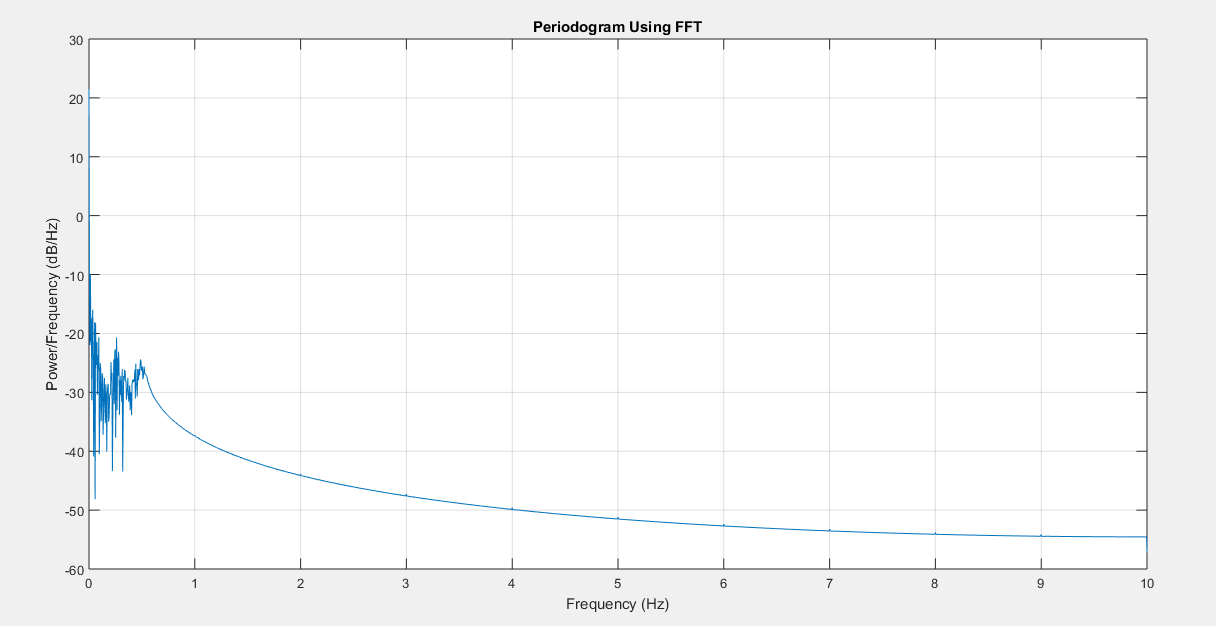
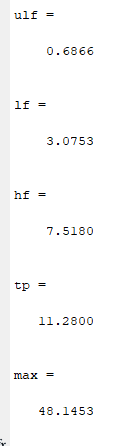
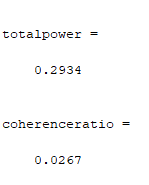


Fig 21: PSD curve of resampled RR data at 20Hz.

After that, results of the power are shown below with respective bands.

Points:

1. Here by observing the FFT, I can do the analyses in which frequency boundary the ECG data lies. And that boundary is [1-40Hz].
2. Ans task 4. After seeing the HR plot, At 40 second, the RR distance is lower, SO at that time some condition is changed of the patient. That might be motion. And after that 100 to 135seconds, RR time is decrease and HR is increase, that after RR time increase from 135 s to 150s. and till the end of the graph, RR time is decrease. SO, where RR peak is decrease at that time patient is doing some activity that is why heart rate is increasing. So, I can see some pattern in the patient activity. So it is not stable at the baseline. And the average heart rate is 75.65 per minute.
3. For resampling the RR data, I am taking the RR Time vector, and find the change between the vector, that can show us how much the time difference between 2 RR time. That can helpful to resampled data. And after resampling this difference, our difference in the RR difference is going low, because we are adding some points between them. And at 20Hz we can see uniform data and change is quite low.

Discussion:

1. Here we have 3 min of data instead of 5 min data. In the statistic the more the data you have the more accurate results you should get with less error. So, if we talk about the standard deviation of NN or RR interval, for 5 min data we can get more depth result. So, after 5 min data, SDNN is going to be decrease (in general).
2. In my data, patient 1 has SDNN is more than 100ms but in patient 2 and 3, it is below 50ms . So, patient 2 and 3 are highly depressed. In my data, LF/HF is represented by variable a. and by evaluating them, it says that none of the patient is healthy. Because HF has higher power in all patient case. In heart rate coherence, we can see that HF is dominant as compare to LF.
3. The basic reason to convert time domain signal into frequency domain signal is to see that in which frequency range you have data. And in which frequency range you have noise. So according to that we can analyses the signal
4. In patient 1, HR is 58 bpm. And we can see that, this patient has variability in his RR distance. It follows the pattern of suddenly increase and decrease in small window and this window follow till end. In patient data 2, Hr. is 92bpm. And this patient data shows that after the sometime interval (Big interval than patient 1) RR distance is decrease and increase. So this patient has different variability than patient 1. But in patient 1 has the RR time interval, that is changing in small range that is why SDNN is Low. But in patient 2 it is larger because after 40s and 80s HRV is changing.in patient 3, the HR is increasing after the 100s. by evaluating the data, we can say that this all patient has different variability, and patient 1 has less variability in his/her HR.