**Data Cleaning, Nosie filtering, and Data Analysis**

**How data was recorded?**

Data was recorded by Physics toolbox application. In this application we can record the daily life activities through our cell phone. Multiple functionalities are also user to research in different fields. Our main focus is on Linear Acceleration, which shows the change in variations of dimensions during a certain activities. Our team work on 4 different kinds of activities such as fall, drop, laying and sitting with the main focus only on linear acceleration.

**How many times data was recorded?**

The data was recorded 20 times for each activity by each group member. The data was recorded through android and apple phones. Data was recorded by doing activity and data is of Linear Acceleration and was saved in csv file with three dimensions X, Y, Z and whenever an activity occurs it causes a change in linear acceleration up and down in each dimensions.

**Noise along with data**

While recording the data for each activity there was lot of noise also comes along with the data. With the noise it will be impossible to predict what actually going on each activity. So to remove the noise from the data we use low pass high pass filtering for each activity.

Initially, for each activity append the X axis with other X axis of files and appending Y axis with other Y axis of all files and appending Z axis with other Z axis of all files. The data comes out with the noise. So to remove noise I uses low-pass high-pass filtering. The results after filtering was quite good. Before Filtering it is hard to analyze that what happening in the data but after filtering the data it is easy to analyze and compare the activities.

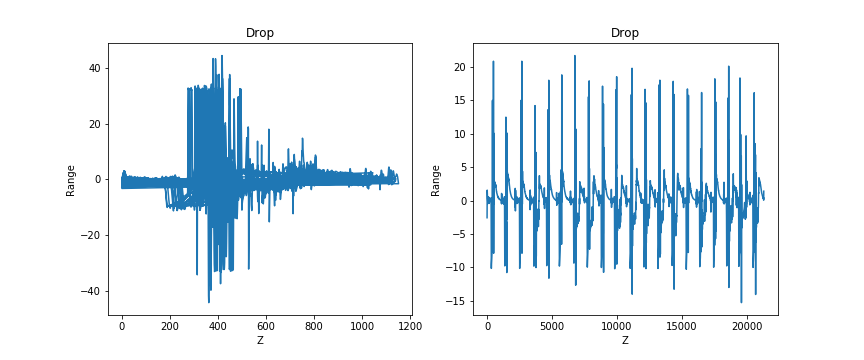
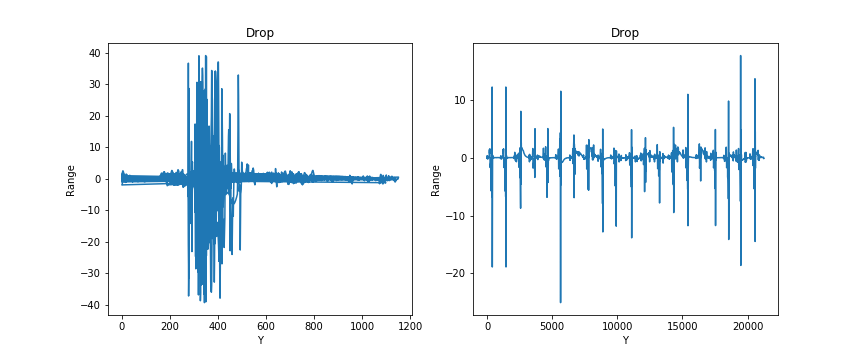
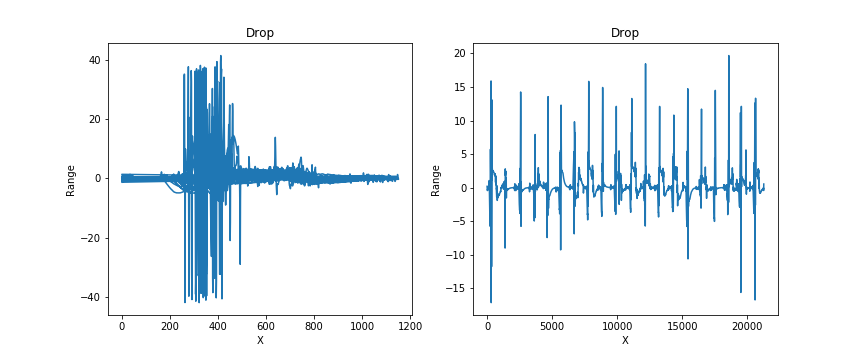
**How different activity data varies?**

Our main goal is to analyze how all these 4 different activities relates to each other. In each activity there is a variation in the data like changing in the values of x axis, y axis and z axis. Some have higher changes in data while some have low changes in data.

To analyze the data I used graphs and plotting’s for each dimension and for each activity. The graphs shows the data before filtering the data after filtering. The data before filtering does not even give an idea about what is actually going on. After appending all the x, y, z axis and plotting them we should get some peaks where we can say that the activity has occur. But the filtering helps a lot in analyzing data as through filtering we can analyze that how many number of times activity has occur and how the range varies for each activity.

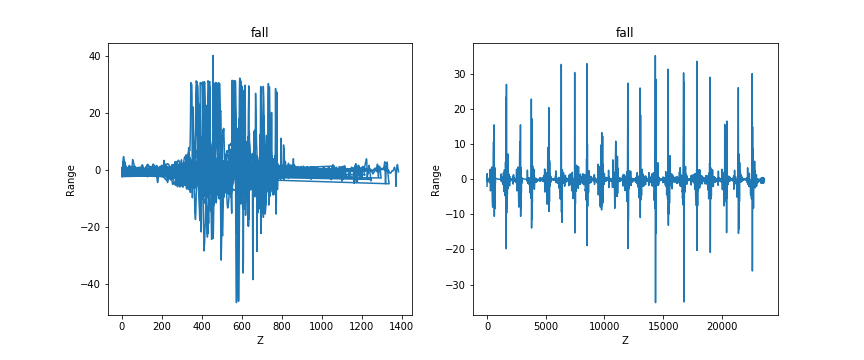
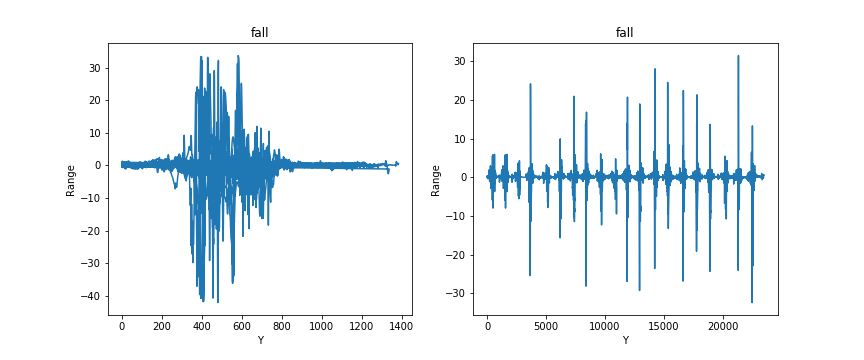
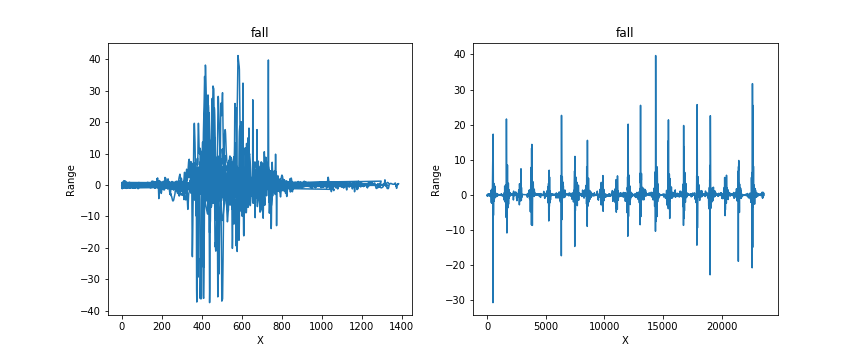
Activity 1- Dropping the phone

From the filtered graph, we can actually notice that there are 20 peaks in which that the phone was drop 20 times. There are also values close to each other in between peaks as that is the data when either the activity has been done or has not started. Change in linear acceleration cause X varies from (-40,40), Y varies from (-40,40) and Z varies from (-40,40).



Activity 2- Walking and Falling

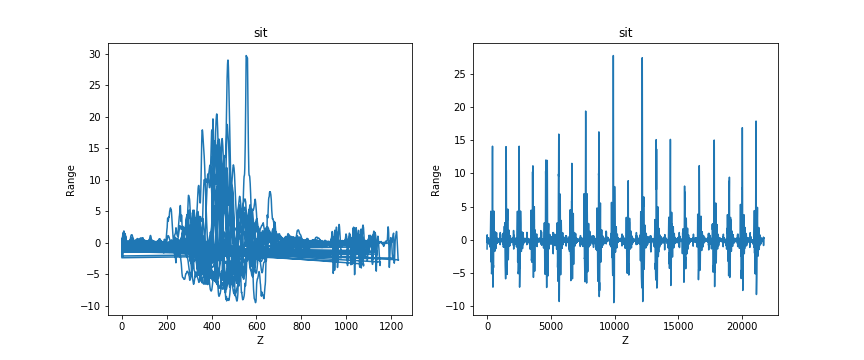
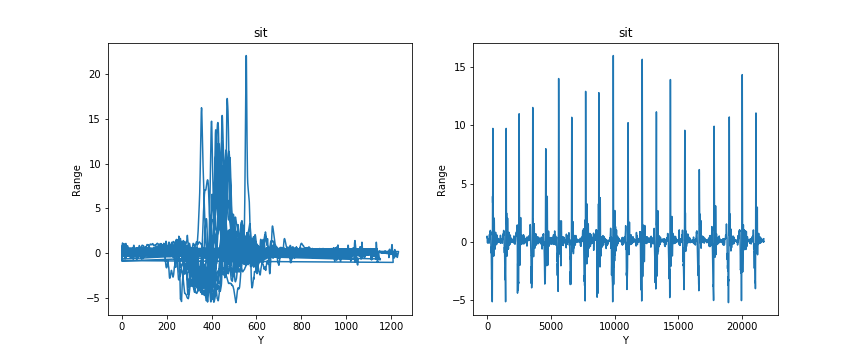
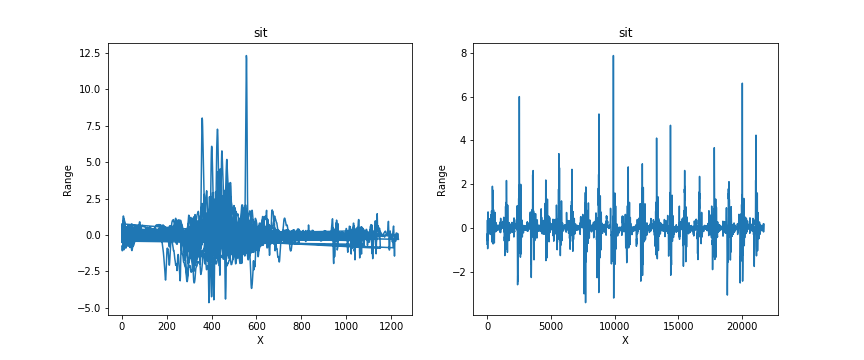
Second Case consider a falling of person during the walk. The recorded data filtered and uses low pass high pass filtering to remove noise.



Filtered shows the peaks where the actuall fall occur. The data after filter is quite understandable and the peaks shows when the fall occur. The change in range of x, y,z during the fall activity is high . The range in x axis changes from (-30 to 40),in Y dimension it changes from (-30 to 30) and in Z dimension it changes from (-40 to 40). So there is a big change in linear acceleration when the fall acitivity performed.

Activity3- Sitting

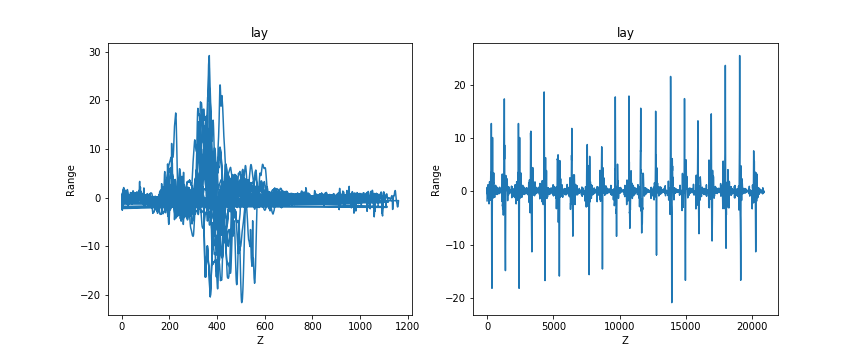
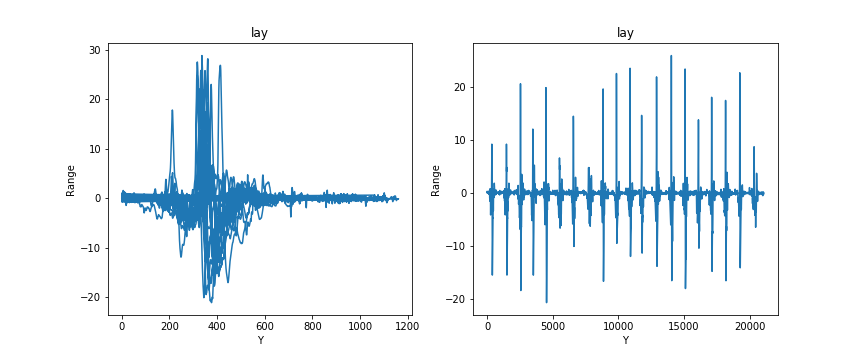
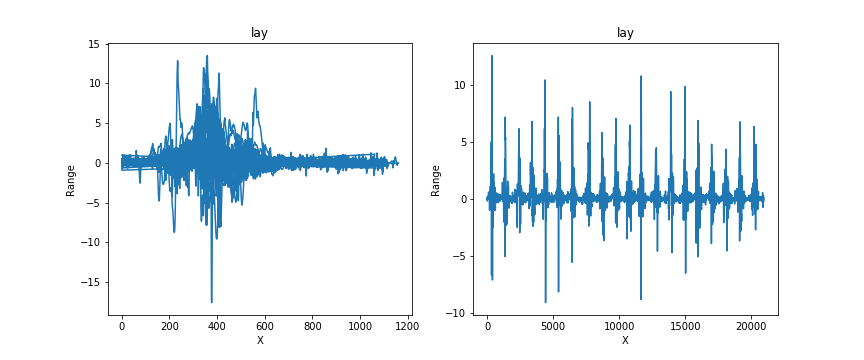
In this activity data was recorded by sitting on the cough and use low pass high filtering to remove noise from data.



The change in variations in the linear acceleration is the lowest as X lies between (-5 to 12), y lies between (-5, 20) and z lies between (-10, 30) .Comparing the sit with other activity it can be concluded that the sit data does not have a lot of variation in the linear acceleration. The peaks in each axis are very closely related to each other and are uniform as compared to the other activities.

Activity 4- Laying down

Data recorded by laying down on couch and is filtered by low-pass high-pass to remove noise.



The laying activity somehow closely relates to the sit activity. But the variation in linear acceleration of the laying activity is higher than the sit activity. The range in X dimension varies from (-15 to 15), in Y dimension varies from (-20 to 30), and in Z dimension varies from (-20 to 30).

**DATA ANALYSIS**

DROP

The change in linear acceleration for the drop and fall activity is very high in all dimension as compared to the others. When a phone drops the change in linear acceleration is very high. Focusing on how in each dimension data is varying. For the drop the change in linear acceleration cause X dimension varies from (-40, 40), Y dimension varies from (-40, 40) and Z dimension varies from (-40, 40). This is the highest change in data we have found as compared to other activities. The data variation shows how rapidly data changes when an activity occur. While dropping the phone the linear acceleration goes highest affected and change in each dimension is maximum.

FALL

The change in linear acceleration for the fall is second highest also closely relates to Drop activity. When a fall activity occur the falls cause the linear acceleration change rapidly in each dimension. The range in x axis changes from (-30 to 40),in Y dimension it changes from (-30 to 30) and in Z dimension it changes from (-40 to 40). So there is a big change in linear acceleration when the fall acitivity performed. But as we compared with the drop activity the variation is somehow less or equal.

LAYING

The change in linear acceleration for the laying activity is less than the fall and drop activity but higher than the sit activity. When the laying activity occur the variation in change in data is not much. . The range in X dimension varies from (-15 to 15), in Y dimension varies from (-20 to 30), and in Z dimension varies from (-20 to 30). As comparing with fall and drop the change in linear acceleration lies between (-20 to 30) which is small.

SIT

The change in linear acceleration for the sit activity is very small as compared to other activities. The change in variations in the linear acceleration is the lowest as X lies between (-5 to 12), in Y dimension it lies between (-5, 20) and in Z dimension it lies between (-10, 30). In general for each axis most of the data lies in between (-10 to 20) which is actually very small. Even from the graph we can analyze that there is no much difference in the peaks as well as most of the data resides in a very small range.

From the above scenarios, It can be easily conclude that the dropping of phone activity has the highest change in linear acceleration, falling has the second highest change in linear acceleration, laying down activity has the third highest and sitting activity is at last with very small change in linear acceleration.