**OBJECTIVE :-**

Finding the BEST INFERENCE MODEL / Pattern OF KW consumption for every 5 minutes of a 90 days data.

**METHODOLOGY :-**

*Outlier detection*

X= average

S=standard deviation

Lower bound= X - 1\*S

Upper bound=X + 1\*S

**IMPORTANT POINTS:-**

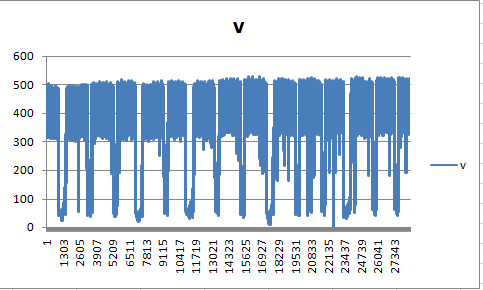
1. I am taking KW CONSUMPTION as an independent variable.

2. First I have observed the behavior of KW consumption for all 90 days data with outliers and without outliers.

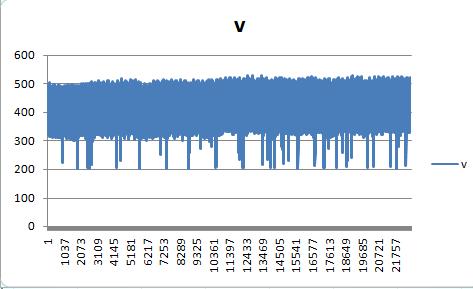
3. Then I have grouped the 90 days according to months and then observed the behavior of KW consumption(independent variable).

**GRAPHICAL REPRESENTATION:-**

*KW Consumption for 90 days data(actual data)*



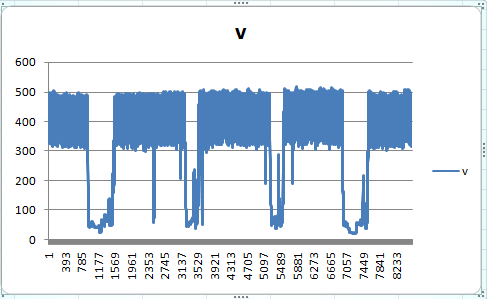
*KW Consumption for 90 days data.(without outliers)*



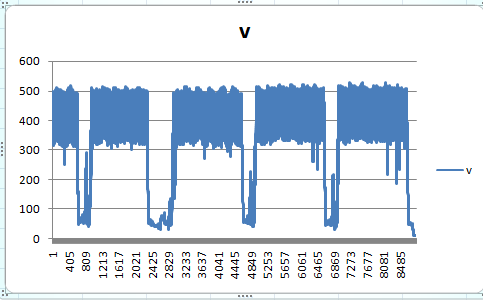
After using the standard deviation method for detecting outliers we are getting this graph.

We can say that from the point 300 to 500 the pattern of the graph is slightly smooth but there are still many outliers existing and If we try to remove all these outliers then there will be a loss of lots of data points from the actual data set.

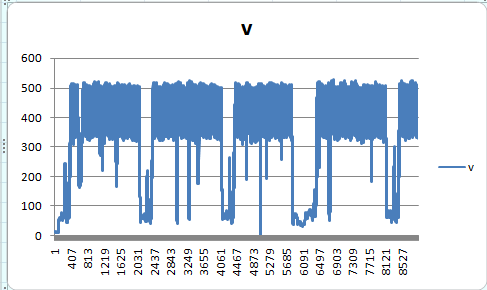
KW Consumption for month 11 of 2017(actual data)



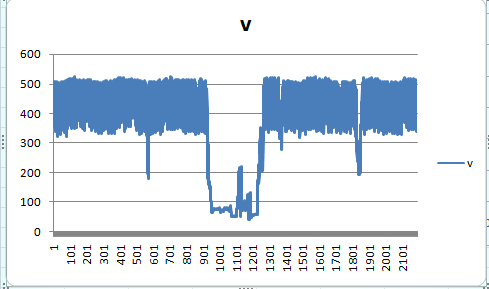
*KW Consumption for month 12 of 2017.(actual data)*



*KW Consumption for month 01 of 2018.(actual data)*



*KW Consumption for month 02 (1st 8 days) of 2018.(actual data)*



If we observe all the above graphs which are grouped according to months, we can see and conclude that KW consumption output pattern for each month is quite same and for each month the outliers are also lying between the same range of 100 to 300.