Phase 1 analysis for a Manufacturing Process Control

ISEN 614 Advance Quality Control

Authors: Ayan Patel (UIN - 727004956)

Dhruvil Shah (UIN - 128001129)

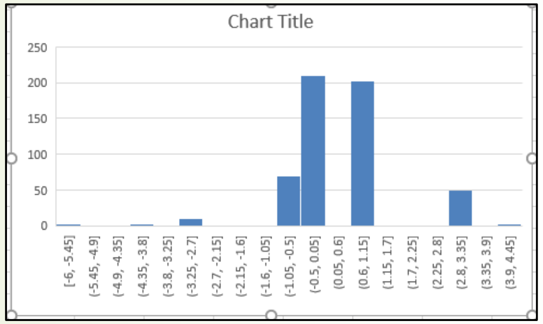
Krisil Patel (UIN - 527003189)

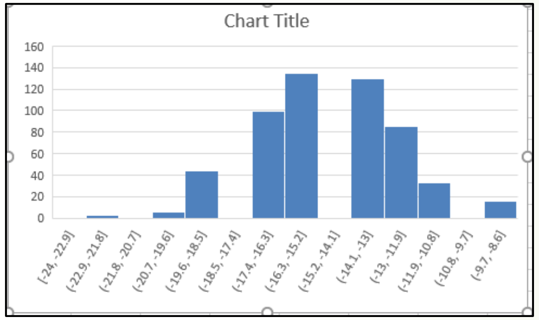
Presentation Outline

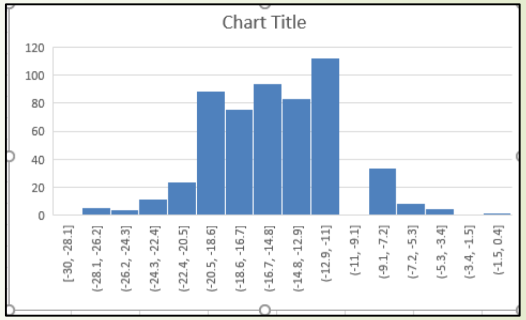
1. Preliminary Analysis
    1. Type of Process, Type of Data and Trend Analysis
2. Principal Component Analysis
    1. Conversion of original data to principal components
    2. Pareto and Scree plot to find components having maximum variance.
3. Control Charting
    1. Use of T^2 and M-CUSUM. Iterations to remove out of control data points and generate an in control chart.
    2. Find the in control Parameters i.e. μ_0 and Σ_0 .
4. Key Learnings

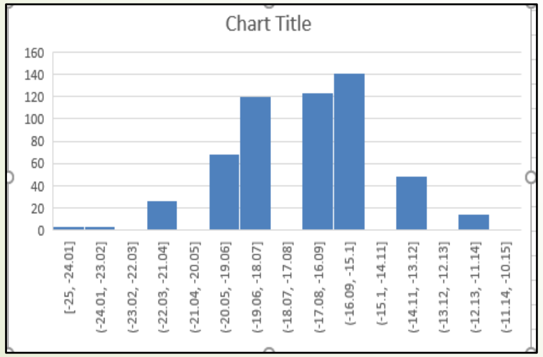
Preliminary Analysis.

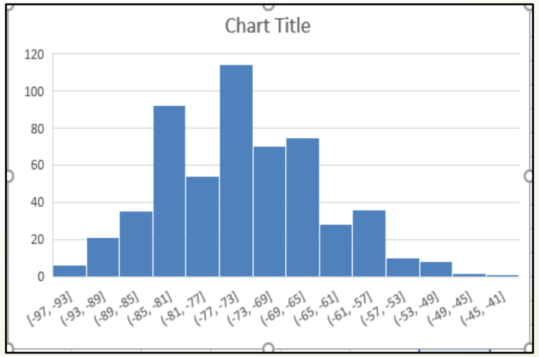
1. We calculated the mean of each of the variables and plotted a graph from it.
2. We concluded that the process followed a unimodal trend having only one highest value.
3. We chose covariance matrix instead of correlation matrix in order to maintain original relationship between the variables of the process.
4. From the graph we assumed that the data follows normal distribution.







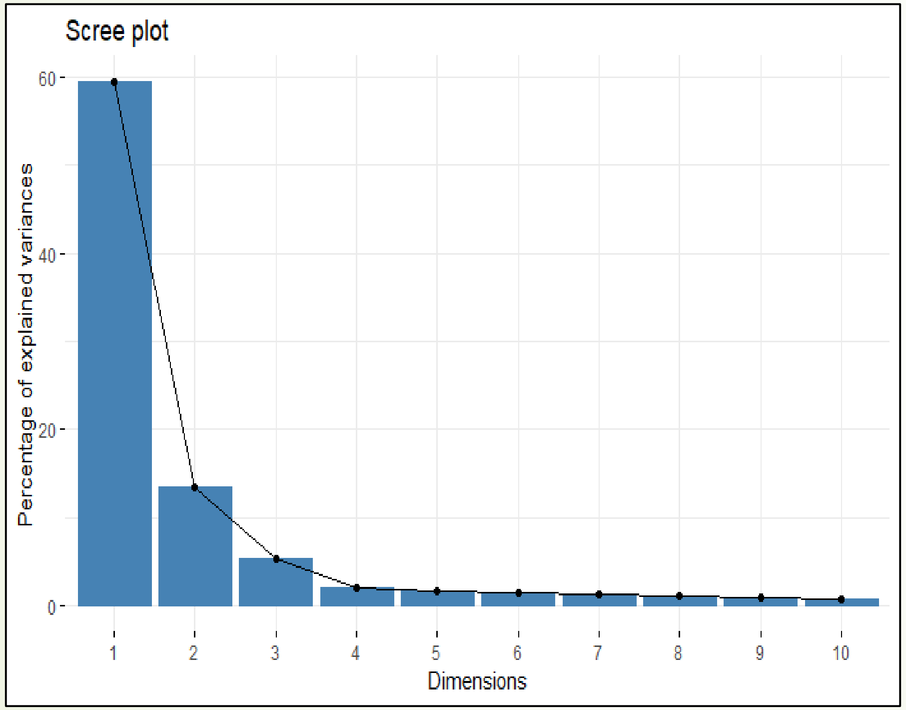




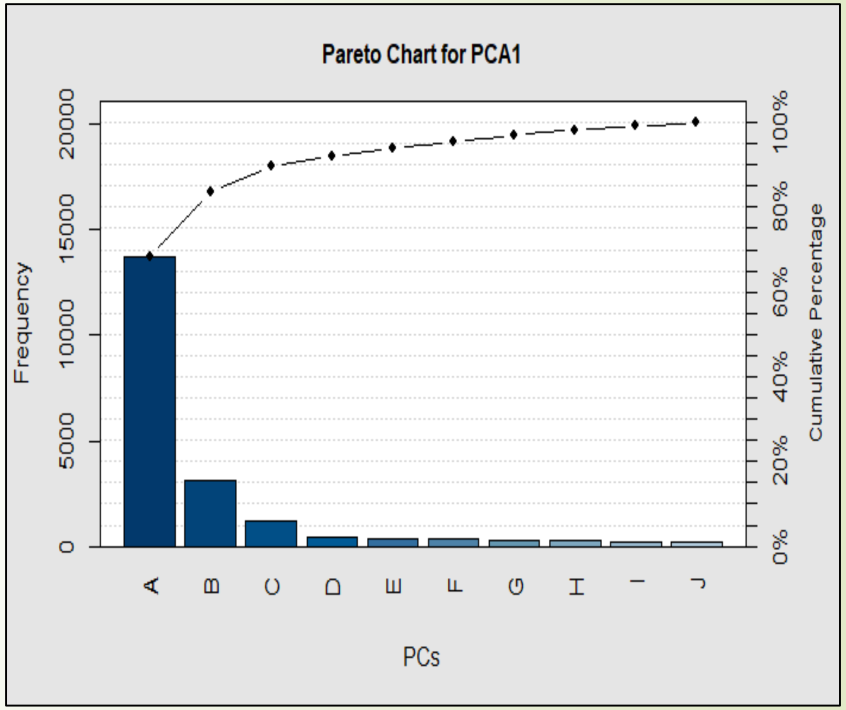
These Graphs shows that the data really follows a normal distribution and if we overlap them we can get a single normal distribution curve.

Principal Component Analysis.

1. R Code to find Eigen values and Eigen Vectors.
2. Using these Eigen values and Eigen vectors, we made Scree plot and Pareto plot.
3. Using Scree and Pareto Plot we reached to the result of using 4 PCs for the control chart



For scree Plot the elbow comes at 4 and so we selected PC=4



The result of scree plot was validated using Pareto analysis and found that I=4 contributed more than 90% of variance.

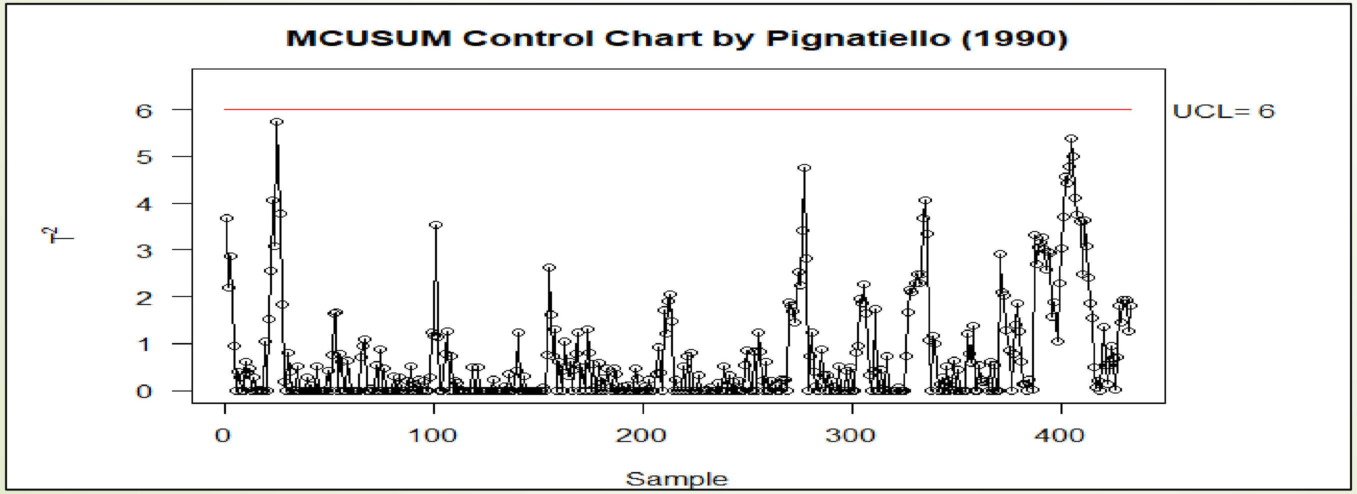
Control Charting

M- Cusum

1. We plot the Cusum chart.
2. We removed all the outliers/ Out of control samples.
3. We assumed mean shift of 3 and our desired ARL is 200.
4. For UCL we used Interpolation from p=2,3,10 to find value of p=4. We got UCL as 6.00 and ARL was 201.

Phase-I Analysis	Number of Out-of-Control Samples Observed
After the 1st iteration	91
After the 2nd iteration	23
After the 3rd iteration	4
After the 4th iteration	1
After the 5th iteration	0

The table shows the number of iterations and the number of out of control points. We had 6 iterations after which we go all the data points in control. We used R programming to do those iterations.



The data points seem in control for the M_CUSUM chart however we can not say that all the data points are in control based on only CUSUM chart. We need T^2 Chart too for determining the in control data points.

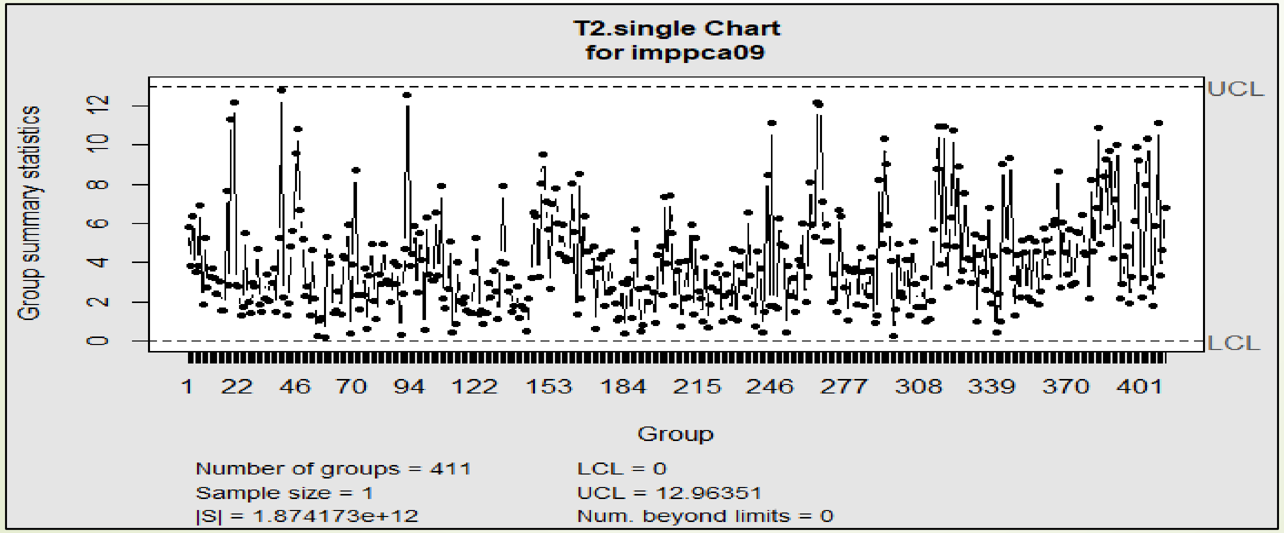
Control Charting

Hotelling T^2 Control Chart

1. As every variable from 0 to 209 have unique observational value, n=1
2. The number of PCA used, p=4.
3. The sigma limit used is similar to Shewhart control chart with 3 Sigma. Control limit Therefore $\alpha = 0.0027$ and using these parameters and Chi Square table we calculated UCL of ~13.
4. We then plotted the in control M-CUSUM chart data points in T^2 chart.

Phase-I Analysis	Number of Out-of-Control Samples Observed
After the 1st iteration	11
After the 2nd iteration	4
After the 3rd iteration	5
After the 4th iteration	2
After the 5th iteration	0

The table shows the number of iterations and the number of out of control points. We had 5 iterations after which we go all the data points in control. We used R programming to do those iterations.



After the iterations the chart seems in control for T^2 and we already removed outliers for M-Cusum chart. However we run M-CUSUM chart once again so that we can be sure of the result.

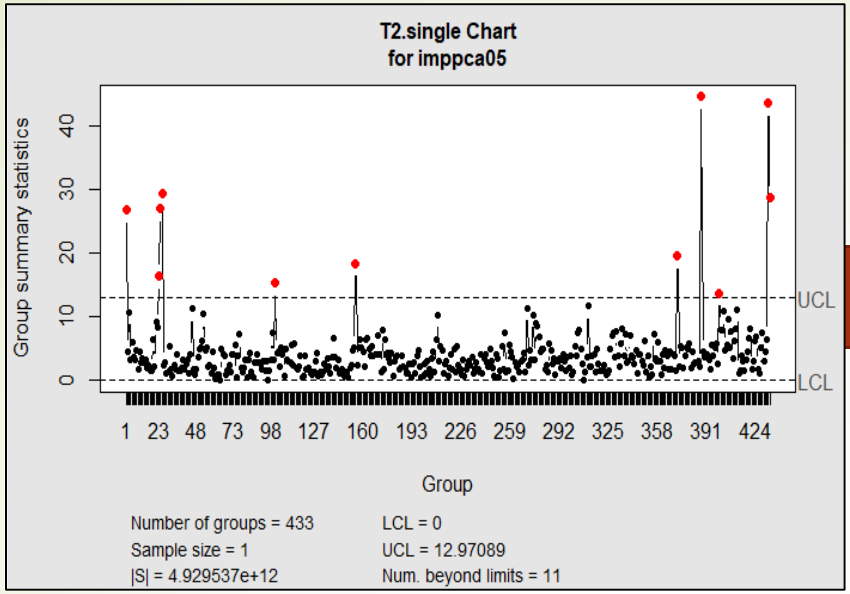
Control Charting

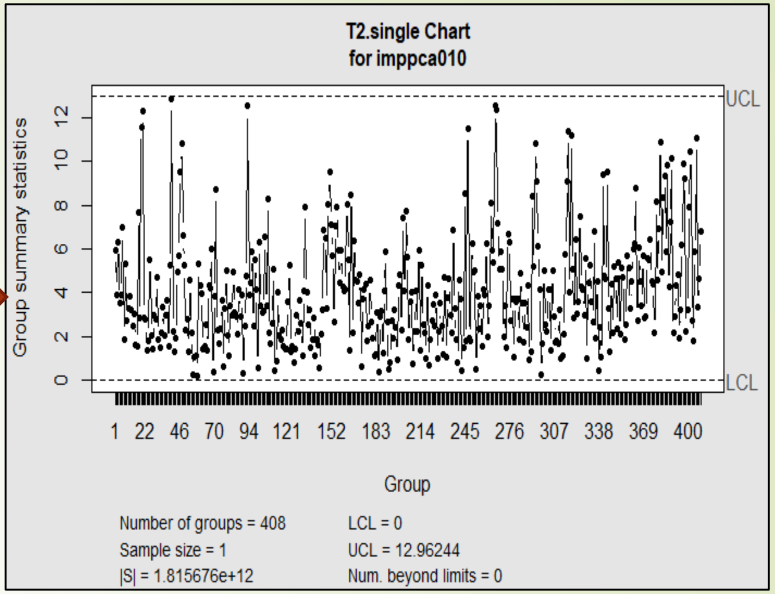
1. We run M-CUSUM chart again and ensure that data points are in control. When we run the M-CUSUM chart we get 3 data points out of control.
2. So we do phase one again and after 1 iteration, there were Zero out of control data points.
3. With these data we ran T2 chart again and there were no out of control data points. This concluded that 408 out of 552 samples were in control and didn't had any spike change or mean shift.
4. This data points can be used as control chart for future observations.

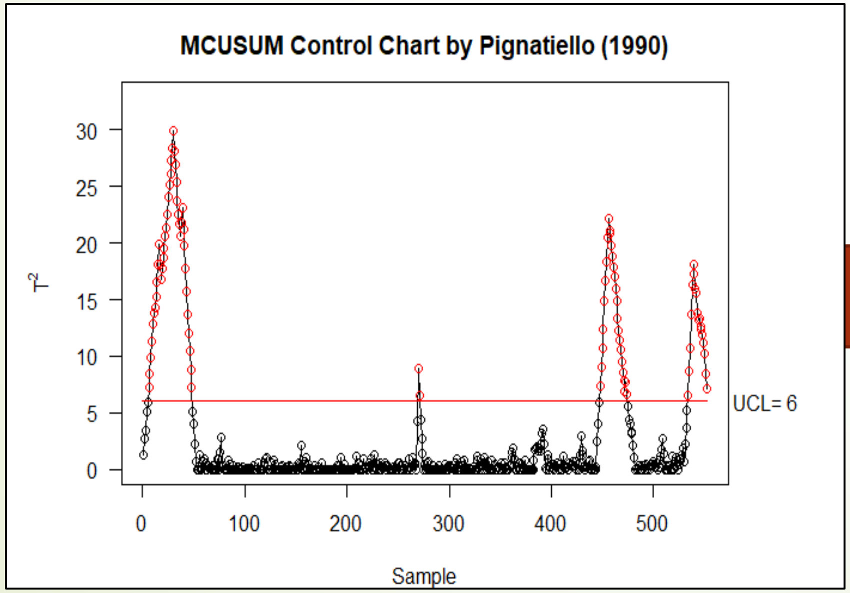
The final data points we obtained are 408 out of 552 samples. These can be used to plot a chart with α = 0.0027 for Hotelling T2 chart and sustained mean shift of 3 for M-CUSUM chart.

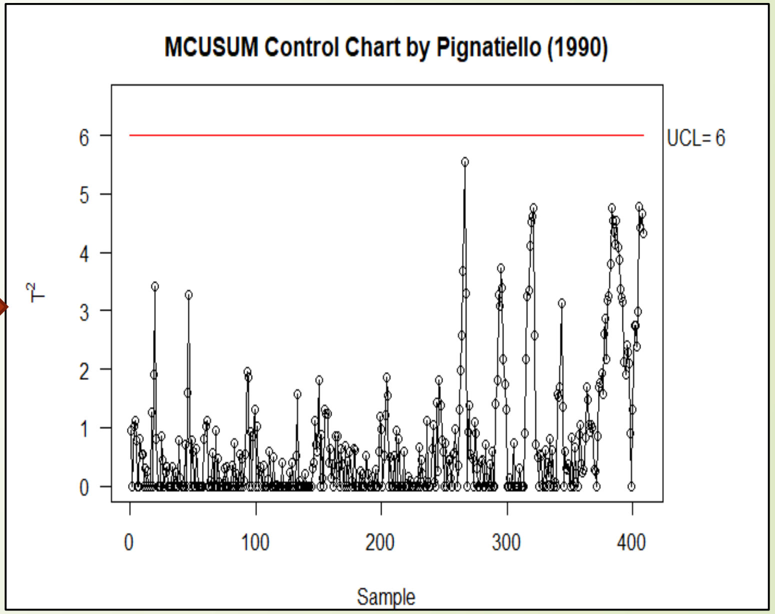
Using M-Cusum chart we removed 91, 23, 4 and 1 i.e. 119 samples. So the remaining samples are 433 samples from which we removed 11,4,5 and 2 i.e. 22 samples using Hotelling T^2 chart. The remaining samples were 411. From these 411 samples we again removed 3 samples using T^2 hoteling resulting in 408 samples having no mean shift or spike changes.

These 408 samples are used in Monte Carlo simulation to create long loops of simulated data to calculate Average Run Length (ARL) values which specify when the control chart would give false alarm and when would it fail to detect. The preference if small ARL_1 and Large ARL_0
i.e. less false alarms and quick detection of out of control data point.


T2.single Chart
for imppca05
Group summary statistics
Group
UCL
LCL
Number of groups = 433
Sample size = 1
|S| = 4.929537e+12
LCL = 0
UCL = 12.97089
Num. beyond limits = 11


T2.single Chart
for imppca010
Group summary statistics
Group
UCL
LCL
Number of groups = 408
Sample size = 1
|S| = 1.815676e+12
LCL = 0
UCL = 12.96244
Num. beyond limits = 0


MCUSUM Control Chart by Pignatiello (1990)
T2
Sample
UCL= 6


MCUSUM Control Chart by Pignatiello (1990)
T2
Sample
UCL= 6

Key learnings

1. Using multiple Univariate control chart for large dimensional data can create a lot of errors but they can help in understanding the nature of variables.
2. Using only T^2 control chart or M-Cusum control chart can be problematic. As seen in this project when M-Cusum was in control there were still some of the out of control data points present that were out of control detected by T^2.
3. The order of using the chart is also important. It matters if we use T^2 first or M-Cusum in terms of number of iterations required. We should use M-Cusum before T^2 can reduce the number of required iterations.
4. We should check for the out of control points for other method even if one method is in in control as observed when M-Cusum was in control but T^2 still had out of control data points.
5. We used interpolation because PCS were less than 10 but if they were more we could not have solved the project because we can not extrapolate.
6. The use of MDL, Scree Plot and Pareto can drastically reduce the number of PCA's.
7. A lot of small variables can create noise in the system and that should be kept in mind while designing a control chart.