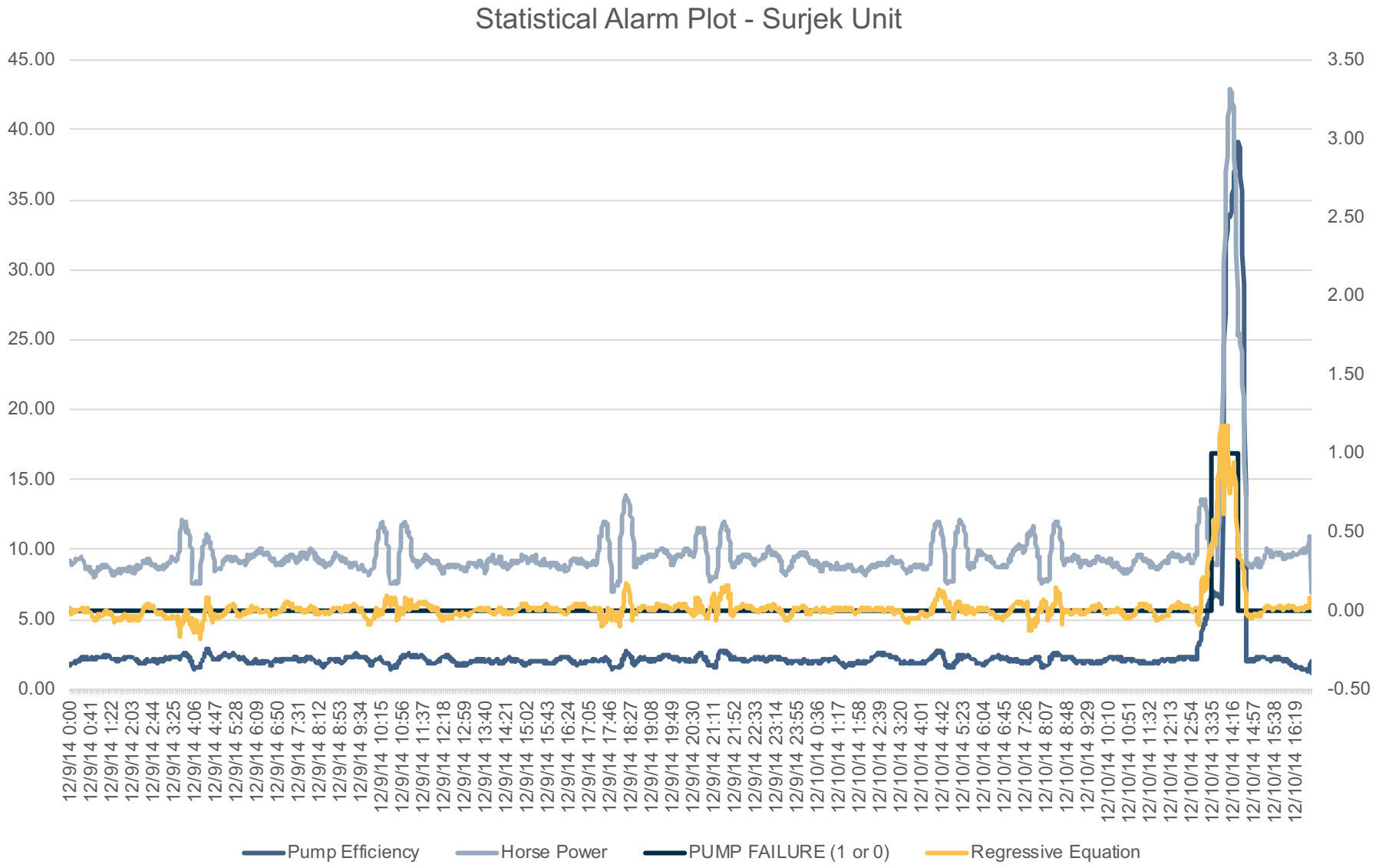
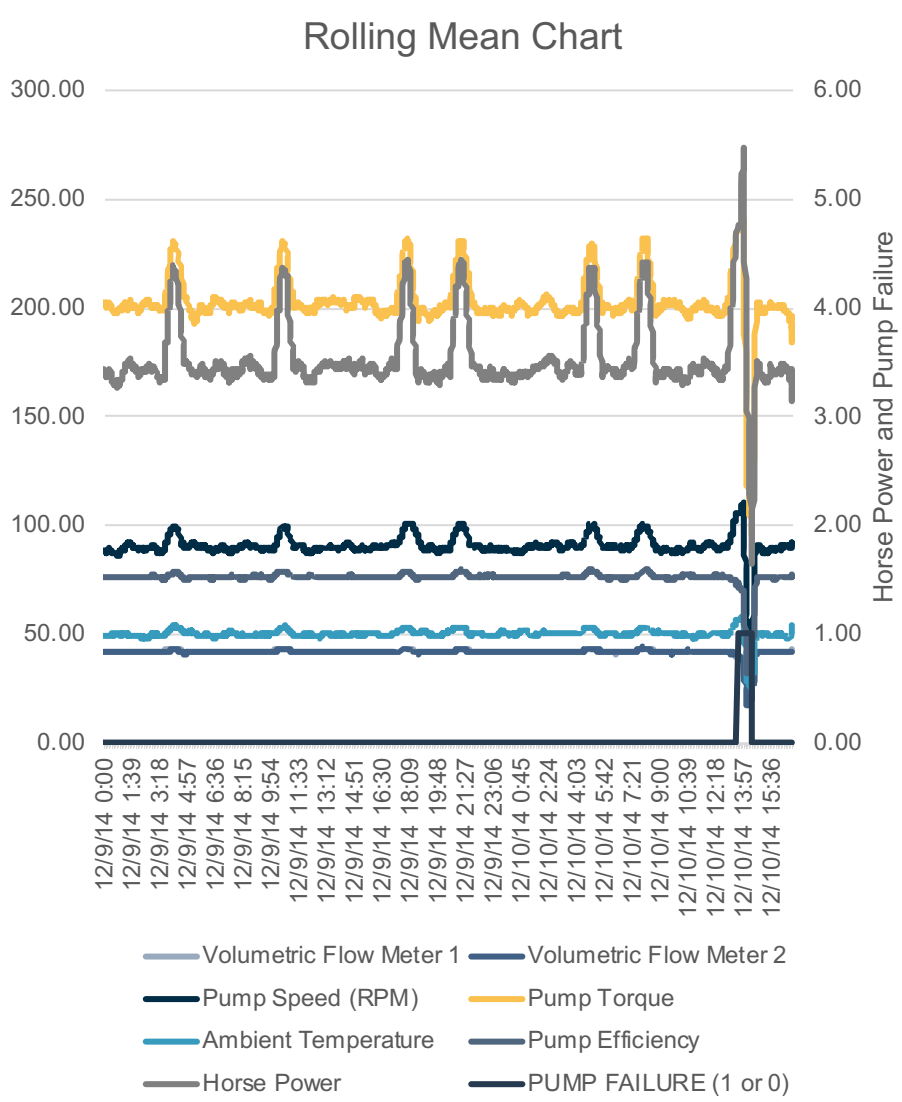
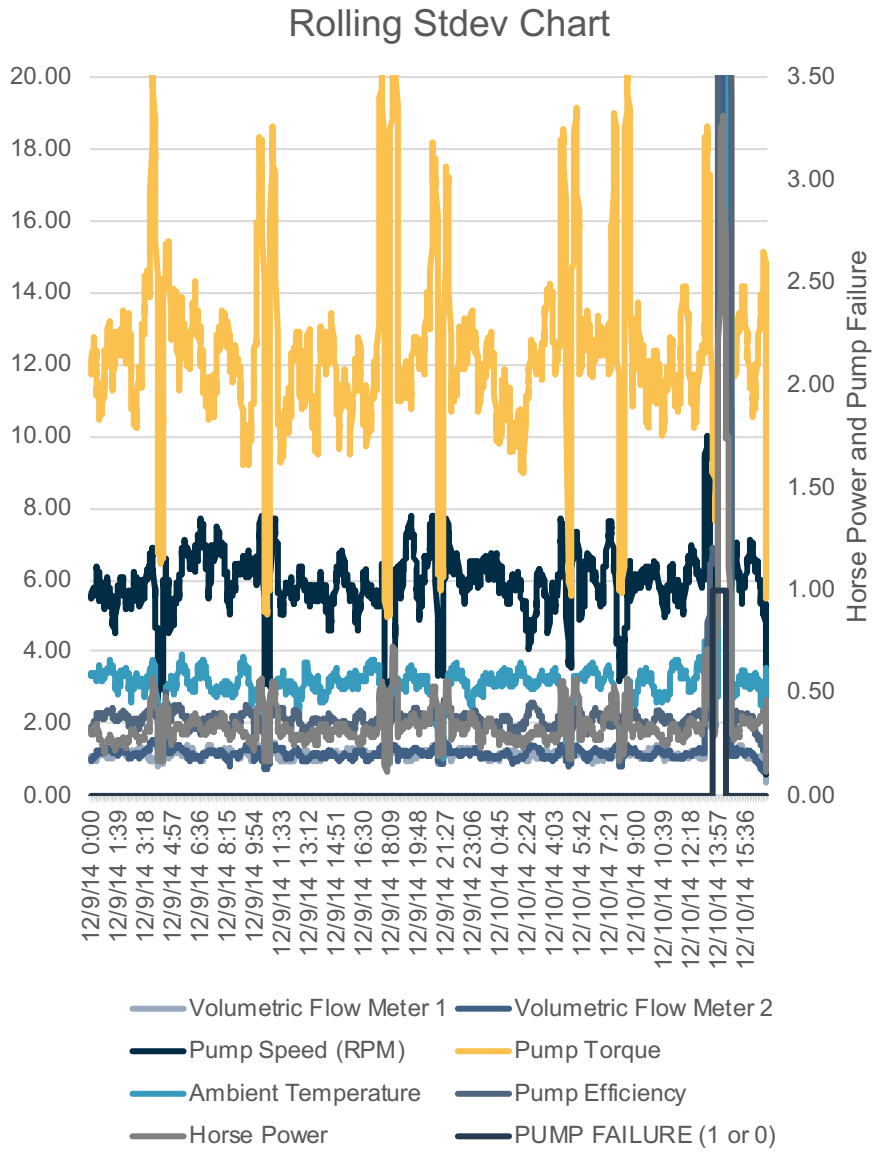
Descriptive and inferential statistical methodologies have proven effective in creating a proactive 'alarm', accurately identifying Pump Failures with Horse Power (HP) and Pump Efficiency (PE) emerging as key variables of interest with deviations of 15 HP and > 3 % PE being our core signal thresholds.

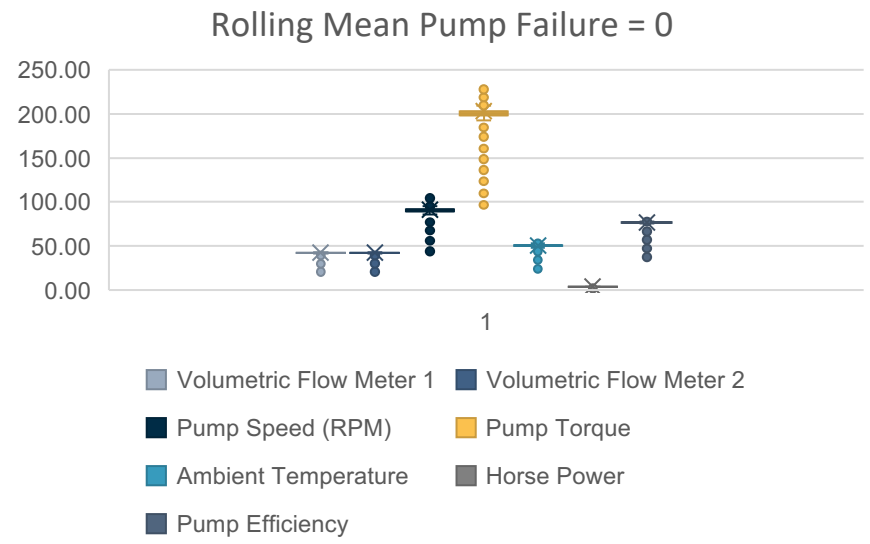


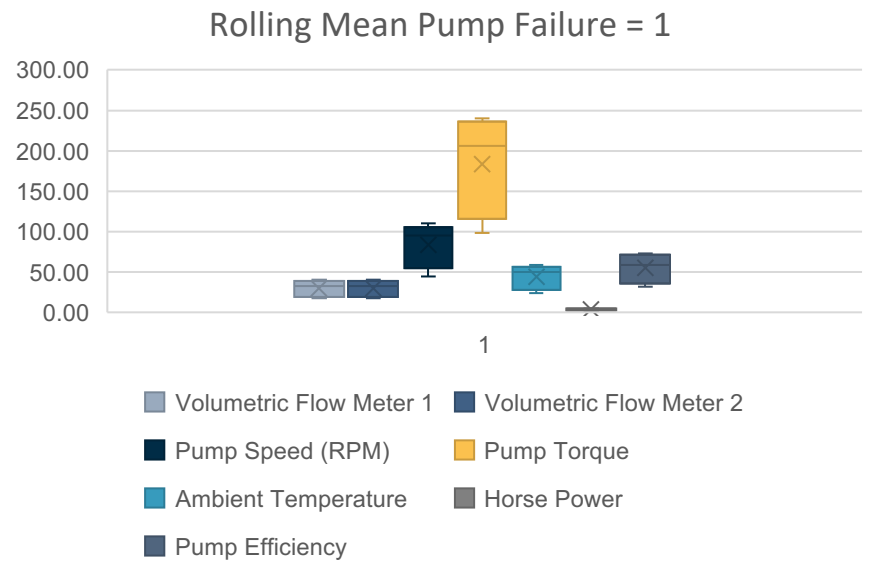
Descriptive Analysis has enabled us to clearly identify particular signature abnormalities showing clear signature changes in both Rolling Standard Deviation and Rolling Mean Datasets when observed over the respective failure period of interest.

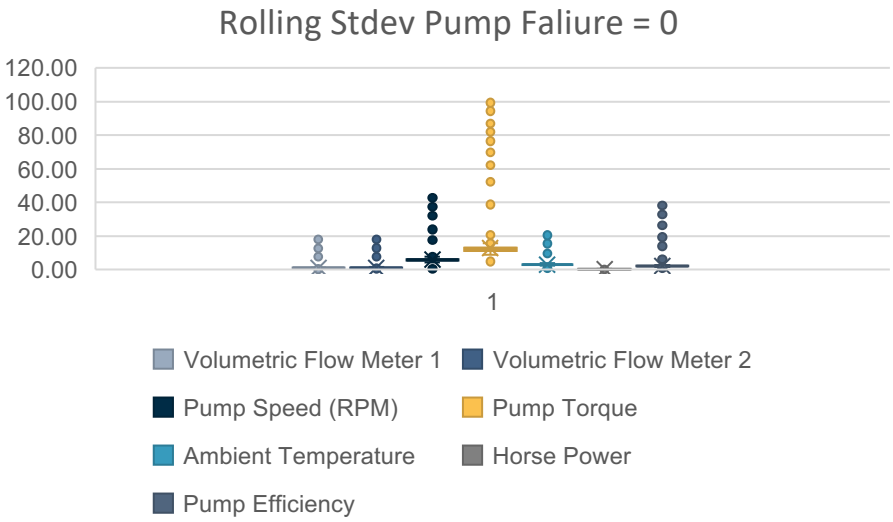


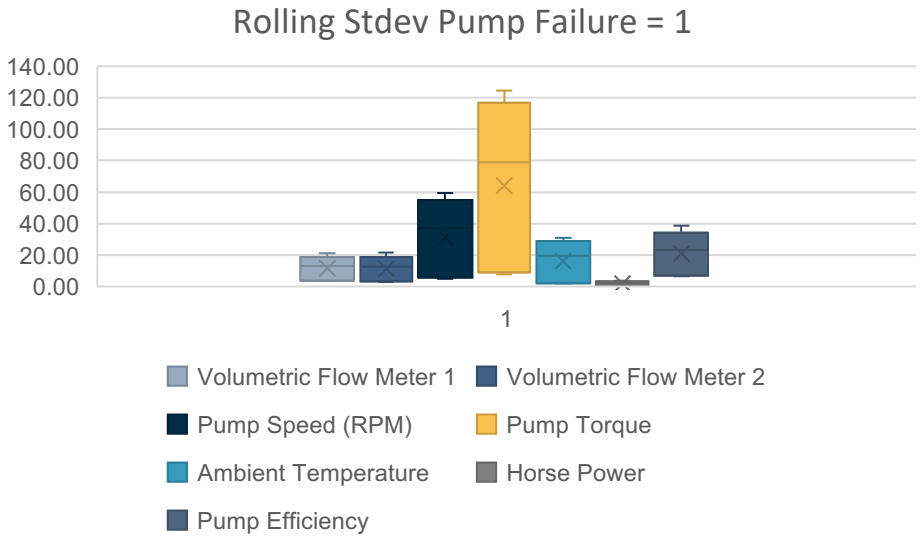


Further segmentation of the data via binary means (Pump Failure = 0 or 1) illustrated through box charts, show a clear signature difference between that of normal behaviour and that of Failure with Pump Torque, Pump Speed, and Pump Efficiency showing the 3 largest variances.

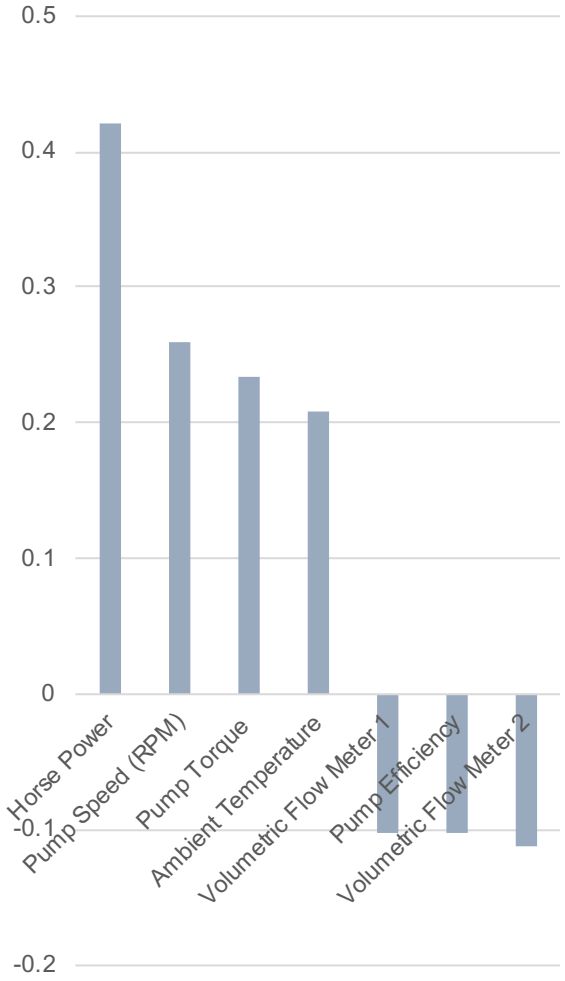


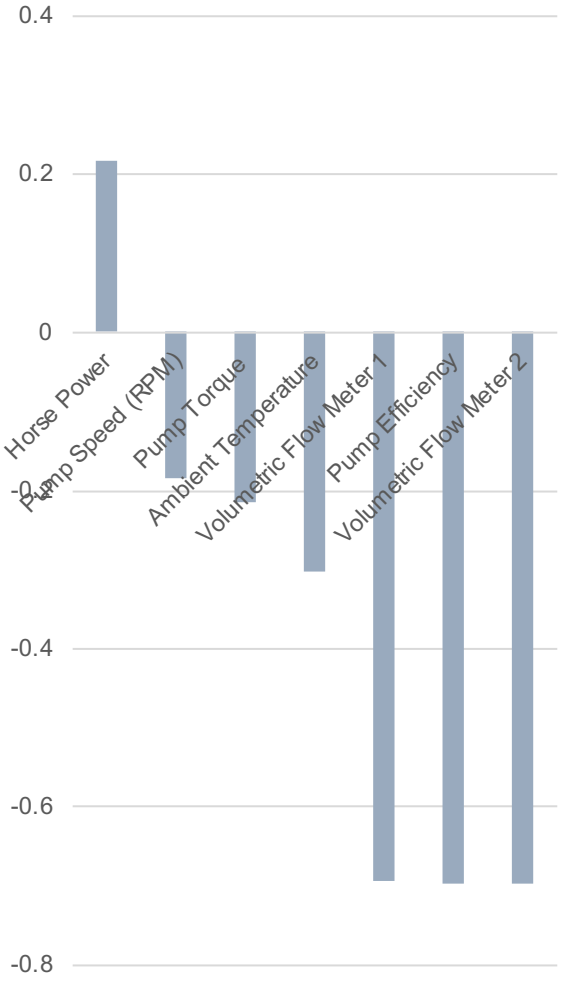


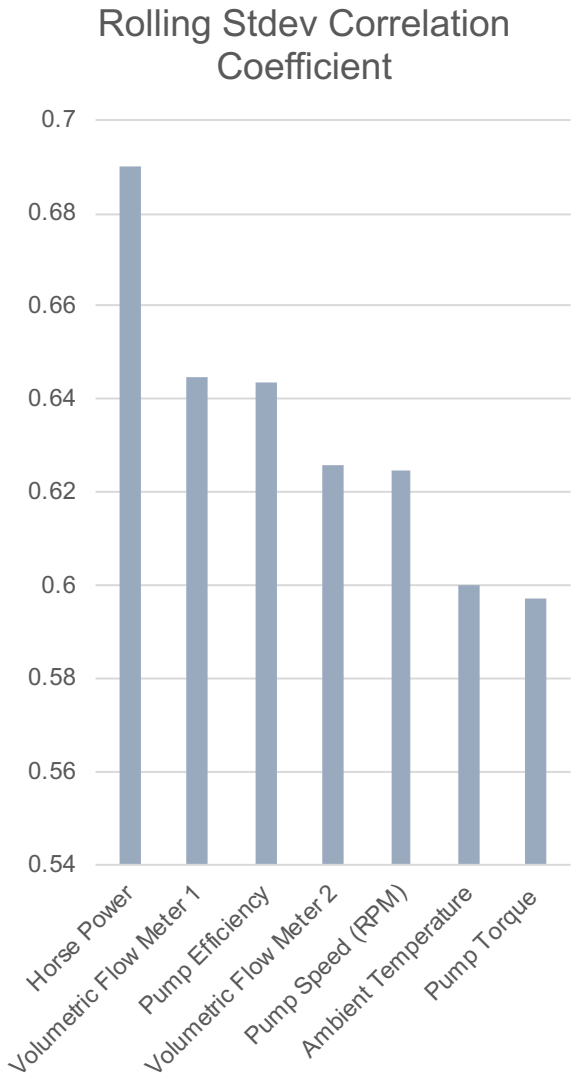




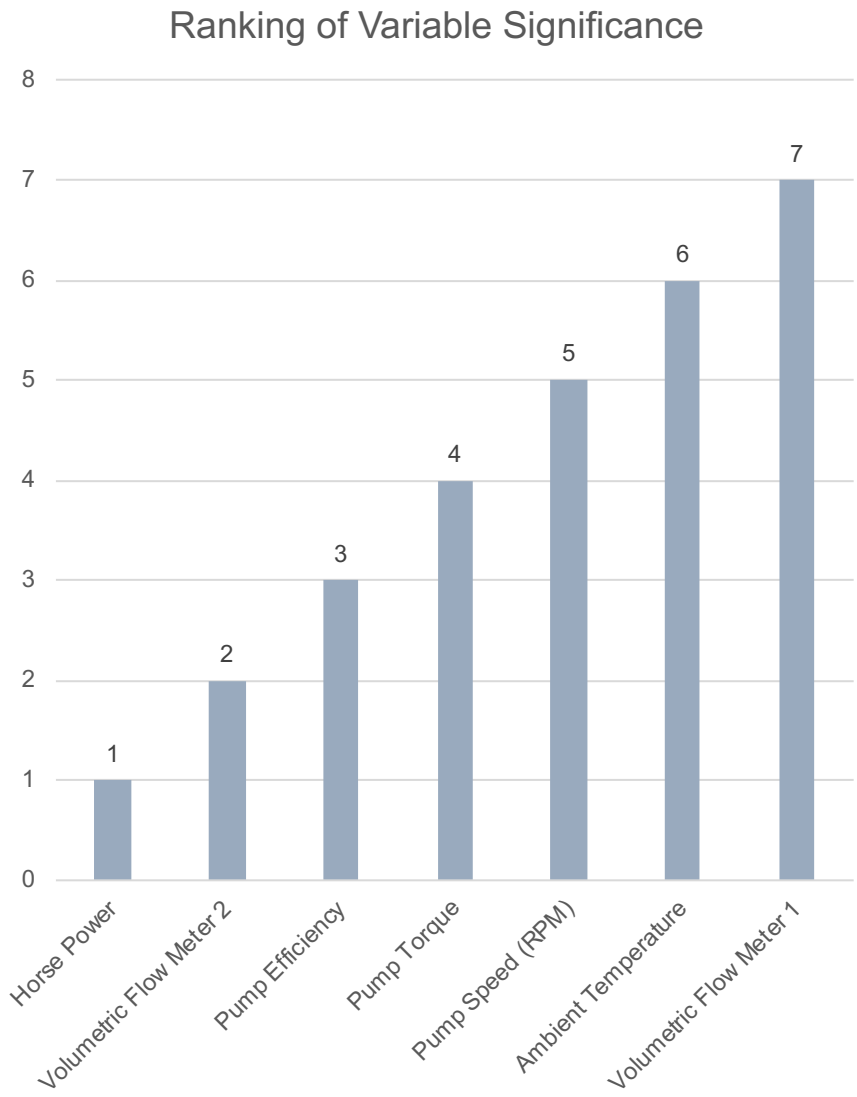
Correlation analyses across datasets yield particularly interesting insights with Pump Efficiency and Volumetric Flow negatively correlated with Pump Failure in the Rolling Mean Dataset, whilst Horse Power and Pump Efficiency show a subsequently strong positive correlation in the Rolling Standard Deviation Dataset.







Lastly, analysis of the statistical significance of variables contributing towards Pump Failure reveal that with a R Squared of 0.779, a linear model is a good fit for the data with both Rolling Mean and Rolling Standard Deviation datasets contributing key information to understand Pump Failure mechanics.



Variable	P-Value Significance
Horse Power	3.5524E-219
Volumetric Flow Meter 2	2.033E-174
Pump Efficiency	2.592E-152
Pump Torque	1.0551E-102
Pump Speed (RPM)	1.68299E-19
Ambient Temperature	4.20516E-09
Volumetric Flow Meter 1	4.01554E-07