Threshold Analysis:

Objective: The objective of the analysis is to identify a robust threshold for metrics to be used in real-time performance monitoring

Context: The payment success rate is one of the most important performance metrics which governs the payment journey and its related customer experience. It is defined as the proportion of successful transactions processed to the total number of transactions initiated by the customer. In real-time scenarios, it is important to monitor this to identify volatility and take immediate steps to address this which can either be a tech-related solve or a bank-related issue.

Methodology: There are two aspects to keep in mind while solving for this:

1. The thresholds should indicate a significant variation in the metric and not just random volatility which otherwise would warrant unneeded attention and resources
2. The thresholds should be defined in a robust way, normalizing most of the external/internal variables

I used a data-driven approach to tackle this. For this, the last 6 months' data around Success Rate was considered at an hourly level. The Mean and Standard Deviation (SD) of Success Rate for each hour of the day across the 6 month period was calculated to understand the central tendency and spread of the data in each hour. The distribution of the metric fairly followed a normal distribution. Keeping in mind the laws of normal distribution i.e. 16% of the data lies below (mean -1SD) value and also to have a strict governing rule on Success Rate, we chose mean-1SD as the bottom threshold for Success Rate. This solves for the two aspects we discussed above i.e. the threshold value is such that it varies significantly from the mean value for that hour by 1SD which commands attention and also this is robust in the sense that each hour will have different values of thresholds so as to minimize and localize the variations in the external and internal factors in that hour.
One more aspect to consider here is the ROI. A decrease in Success Rate eventually means a decrease in the value obtained from customers in terms of sales transactions and the monetary aspect. Different values of bottom thresholds would give us an idea of the value lost from customers if the metric goes below the thresholds, thus providing a definite monetary impact to trade-off with the cost involved in solving for this.
The final implementation involved a bottom threshold for Success Rate for each hour of the day based on the historical variation in the metric

Results/Impact: Once this got implemented, it provided an automated way by which the thresholds could change based on the rolling period of time so as to take into account the variations in customer sentiment and process/tech initiatives and also helped take better decisions based on the monetary value lost to the cost trade-off

Sentiment Analysis:

Objective: The objective here was to find out the sentiment of the Customers writing Emails to Customer Care so as to prioritize the Email response

Context: The contact center of Flipkart receives around 100K Emails per week expressing issues with the pre or post-purchase scenarios. The Contact works in a First In First Out reactive mode and hence does not involve any sort of prioritization with respect to customer anxiety. This leads to a non-optimized way of handling customer issues leading to the most anxious customers vetting out their anger by escalating it to social media. To avoid escalation and to optimally solve the customer issues, it mandates to know the anxiety of the customer.

Methodology: The approach was to find out the anxiety of a customer based on the email text classification which defines the sentiment of the email.
While text classification was the first choice for this type of an unstructured data problem, we pinned down on the knowledge based technique of sentiment analysis specifically for this problem, which works on the premise of identifying the polarity of a statement whether ‘positive’, ‘negative’ or ‘neutral’ based on the unambiguous affect words like ‘sad’, ‘happy’ etc. These words based on their individual polarity and also based on combination of affect words define an overall text to be positive, negative or neutral.
While knowledge based technique of sentiment analysis a robust way for classifying text and identifying the inherent sentiment, there are two aspects to be dealt with while using this in a real world application like the one we have

1. The first one is to define what accuracy means for an unstructured text classification and how to measure it
2. The 2nd part of the problem is to design a framework to improve the accuracy

To solve for the accuracy definition part, the algorithm was run on a significant amount of data, which classified the words into their polarity and a metric was defined called the Sentiment Score which was basically the number of positive words subtracted by the number of negative words. A sample of emails from the above were manually checked and the same score was generated. The scores from the model and the manual process were compared to analyze the accuracy of the model in predicting the sentiment. This process was robust as we were comparing the customer sentiment as understood by a human agent. The first version of the model generated only about 50-55% accuracy which didn’t seem fit.

The 2nd part of the problem of improving the accuracy first required us to identify what was the actual reason why the model was not able to match human performance. When looked deep into the working of the algorithm, it became apparently clear that the standard dictionary of unambiguous affect words provided in the statistical package in R was not robust enough for this context. To achieve the creation of a robust dictionary, the new words encountered during the manual calibration were tagged to their polarity and added to the existing dictionary. This process was iterated with multiple samples of emails so as to achieve a robust dictionary.

After the creation of the dictionary, the model was again run on a new sample of data and the accuracy of the data is again calibrated with the new dictionary and this time an accuracy of 80% was achieved which was good for implementation.
The model process was automated in real time with each of the emails received by Flipkart going through this model calibration and an output was generated which was basically the Sentiment Score i.e. number of positive words subtracted by the number of negative words which gives both the polarity as well as the severity of the email. The emails were then stack ranked based on the Sentiment Score and were routed to agents in that order for resolution.

Results/Impact: The prioritization of the customer emails based on the Sentiment of the customer caused the customer escalations to come down by 30% and the Customer Satisfaction Score to improve by about 300 bps

Sustainability Analysis:

Objective: The objective here was to identify what is the Incidents per Unit (IPU) sustainability affected by and what might potentially drive it in the future

Context: The Contact Centre at Flipkart receives around 1Mn Contacts per Week from Customers either through Calls, Emails or Chats. These are spread across the journey of the Customer right from Pre-Purchase to Delivery to Delivery to Returns to Refunds and also across various product categories like Mobiles, Lifestyle, Large Appliances, etc. A lot of initiatives have been carried out in the recent past to tackle these contacts from a process and policy level so as to provide a better resolution to the customer. While the past has seen improvements in many of the pinpoints of the customer journey and Categories, it is imperative to find, what holds for the future in terms of potential areas to tackle.

Methodology: The Correlation of Overall Flipkart IPU was established with each of the Business Unit wise IPU as well as Journey node wise IPU at a WoW level historically. The correlations from a historical perspective were established and the important Business Units and Journey nodes were identified which significantly decreased IPU. Along with this the recency effect of correlation also gave us the saturated nodes in which further decrease would cause insignificant changes and the potential areas which have been slumber for quite some time and carry a lot of momentum and which become the major areas to decrease IPU in the next phase

Insights/Impact: One of the clear call-outs from this analysis was that Returns has shown the greatest correlation to IPU historically and is saturated to quite an extent and now Pre-Delivery would drive the next phase of IPU drop followed by Pre-Purchase as well. Major initiatives were planned out the potential areas while maintaining status - quo in the existing ones and this drove down the Incidents per unit by almost 200 bps

UPI A/B Testing:

Objective: The objective here was to design the A/B framework for a new payment method to be launched and also analyze the performance of the new Payment instrument post-A/B

Context: UPI was planned to be launched as a new payment instrument on the Flipkart Payments platform in two ways - Intent and Collect, the former having a platform which opens up UPI on the Flipkart interface itself while the latter redirects the customer to a 3rd party UPI app for authentication.

Methodology: One of the major considerations while launching a new payment instrument is to have a seamless customer payment journey experience for the customer enabling an order to get placed. This experience can be quantified using a metric called Payments to Order (P2O) which basically defines the proportion of Orders placed to the Payments Initiated by the customer indicating success in the Payments process.

We need to measure the performance of UPI - Intent and Collect flows through an A/B experiment compared to other instruments so as to compare the success metrics like P2O for the new instrument to the existing ones as well as between the Intent and Collect flows.

We started off with the design of the A/B experiment by considering the baseline of the P2O metric as the baseline conversion metric and an expected improvement in P2O as 100 bps, with the statistical power at 80% and statistical significance level at 5% which gives minimum sample size to consider for both Test and Control set. A simple random sampling was setup for the customers visiting the Flipkart App/Site with the Test group receiving the UPI option of payment while the Control group doesn't.

Once the required sample size was achieved the experiment was stopped and we proceeded to analyze the post A/B performance. Apart from measuring the major performance metric, P2O, the journey of the customer is also analyzed with respect to drop-offs at each stage so as to pinpoint the journey phases/processes which caused the maximum proportion of drop-offs.

Results/Impact: The UPI instrument as a potential future payment method was established based on the A/B experiment. Apart from this the loopholes in the payment journey experience was analyzed for both the Intent and Collect flows for future enhancements

Salary:

CTC: 11.15 lakhs

Base Pay: 9.7 lakhs

Variable: 10% of Base Pay = 97,000

Benefits (Insurance+Gratuity) = 48,000
20 ESOPs worth 80K

Expected:
40% on Base Pay = 140% of 9.7 = 140% * 9.7 = 13.58 lakhs
35% on Base Pay = 135% of 9.7 = 135% * 9.7 = 13.09 lakhs

NOTE: ROI (**Return on Investment**) measures the gain or loss generated on an investment relative to the amount of money invested. ROI is usually expressed as a percentage and is typically used for personal financial decisions, to compare a company's profitability or to compare the efficiency of different investments.

ROI = (Net Profit / Cost of Investment) x 100