**EXECUTIVE SUMMARY:**

This report gives an overview about the performance of the call centre over the last winter. It also details about the activity level of the call centre and gives graphical evidence on the Virtual Hold Technology usage during the last winter. This comprehensive analysis is majorly done on the python platform and GUI has been built for the user to get a visual view of the data. The aggregated data analysis is performed on the given data to understand the given variables better and to identify the relationship between each variable. This helps them in future planning of any event or might be used in designing a protocol to handle the peak hours.

The main reason behind the given data exploration is to comprehend the past peak hour data and prepare strategies for the future to tackle the issues faced in last winter and improve the efficiency level in the company. The Virtual Hold Technology – a experiment conducted by the company seems to have no major impact on the number of calls abandoned and on top of that in most of the months the calls abandoned seems to have been low when the technology was off. Summary statistics for each variable is highlighted in the report to know the granular details of each peak hours.

The call frequency has also been investigated to decide on which part of time the call recorded were the highest so that the user could forecast before hand and apply appropriate resources to handle the peak hours. The Correlation between the each of the variable is displayed and the strong and weak linear relationship has been denoted by the scatterplot.

The conclusion part of the report talks about the major findings found in the call centre data analysis and shows how the user has been provided a permission to add, delete the data in future and thus it makes this system a flexible to use for the future set of data.

**INTRODUCTION:**

The discussion in this report focuses on the call centre management's action regarding how customer enquiry calls were handled over the previous winter using the Virtual Hold Technology (VHT) and how the customer experience might be enhanced going forward. The Data Exploration depends on the 100 sampled data taken from the original data through random number generator.

The report has four sections of data where the first one focuses on the correlation between each given variable, and it is supported with the graphical evidence of its linear relationship. The next section focuses on the data exploration done on the sampled data along with its visual representations and the inference observed based on the analysis. Furthermore, the following section provides evidence of the Virtual Hold Technology impact on the calls abandoned for each month of data.

The final part of the report gives a summary of all the data exploration of the sample data, suggestion to improve the performance level and the appendix which includes the python code from which the GUI and plots have been obtained.

**1.1 VARIABLE CORRELATION:**

Initially, to proceed with the data exploration the given data is sampled to get a random 100 data from the source. And all the analysis has been performed on the sampled data and this gives an idea on the original data performance. The graph below shows the correlation between all the given variables.

Chart

Description automatically generated

Fig 1.1 Correlation Coefficient Plot

The scatterplot above denotes a very strong relationship between the variables – agent, calls offered, calls handled. This indicates that number of agents employed during peak hours has a strong influence on the number of calls handled and received. It is visibly noted from the plot that average handle time has no or weak relationship with any of the given variable. Average speed of answer and calls abandoned have a good linear relationship between each other.

A screenshot of a computer

Description automatically generated with medium confidence

Image 1.1 P- value of Correlation Coefficient

Also from the numerical evidence, it is noticeable that the ASA has a very minute negative relationship with the agents, calls offered, and calls handled. As it is a very slight chance that this negative relationship has an impact on other variables.

Based on this correlation analysis, it can be concluded that the among all the given variables, the agent, calls offered, and calls handled holds the strongest relationship with each other. This also shows the evidence of the dependency among variables. Whereas the ASA and Calls Abandoned holds a decent linear relationship. Thus, out of all given variables, Average handle time is the variable which has very weak relationship with the other variables.

**1.2 DATA EXPLORATION:**

**a. Summary Statistics of Numerical Variable:**

In this data analysis, the minimum, maximum and average values of the agents, calls offered, calls abandoned, calls handled, ASA, Average handle time is found. These aggregations are done based on the month and time of the day value entered by the user.

Table

Description automatically generated

Image 1.2 Aggregated table based on User Input

Based on the above data, it is possible to filter out the data based on the specified month and time of the day to find the aggregated values of all the numerical variable. From this the user can get an idea about the minimum, maximum number of agents working and all the calls attributes. By comparing the data between different time of day and in same month, user can get an understanding of the efficiency of the agents and performance level of calls centre at peak hours. It can also give the average speed of answer which can provide an idea on how quick the agents answer the call during that time of the day. The box plot below shows the minimum, maximum, median, Quartiles for the sample data.

Chart, box and whisker chart

Description automatically generated

Fig 1.2 Box Plot for Sample Data

The plot provides an overall view of the summary statistics of the sample data and thus allows the user to compare get an idea on the sample data statistics. They can then compare the overall data with the aggregated data and conclude on the resource planning of agents. From the plot, it can be seen that in calls offered and calls handled there is a huge difference between minimum and maximum value indicating that the call centre receive and handle huge number of calls over the last winter. One of the interesting facts to notice is that the calls abandoned is very small compared to the other variable and it proves that the call centre functions efficiently. Even though, the maximum number of agents in a sample is near to 50 they receive and handle calls above 230.

**b. Call Frequencies Based on Month:**

From this exploration, the user can get an idea of the call frequency received during the day based on the particular month. This gives an insight on expected number of calls to be expected in the future and based on that value the agents resourcing can be adjusted. This observation might help the call centre to increase their performance as well as reduce their budget on agent resourcing. The below plot shows the frequency of the calls during the time of day when user selects a particular month.

Chart, bar chart

Description automatically generated

Fig 1.3 The frequency of calls received during Dec-Jan Month

It can be noticed from the graph that the frequency of calls received during the evening is very low compared to the afternoon and morning. The call frequency on morning and afternoon don’t differ much in the numbers. From the GUI, on selecting all the months, it can be inferred that among the given three period of months, the number of calls received during the afternoon are higher in the Dec-Jan and Oct-Nov periods. And the Feb-Mar period receives the highest calls in the morning. From this, it can be inferred that during the end of year the callers mostly call during the afternoon. It is clearly visible that in all the time period, evenings receive the least number of calls.

**c. Agent Efficiency Based on ASA:**

On analysing the relationship between the ASA and time of day with respect to the given month gives an insight on how quick the agents answer the call when the call is received.

Chart

Description automatically generated

Fig 1.4 Time of Day vs ASA

The figure 1.4 shows how the ASA has varied along the time of the day. On visually analysing the graph, it is noticed that the in the given months the average speed of answer was minimum during the afternoon and higher during the evenings. This combined with the analysis of the call frequency gives an idea that when the call frequency is less the agents take lot of time to answer the calls but when the call frequency is high the agents answer the calls quickly. From this it can be understood that the agents are efficient when working in the peak hours of the day. The Morning period of the day is moderate even compared to the other time of day.

It is clearly evident that in the month of Dec-Jan the average speed of answer follow the downward trend along the time of the day. But for the other given months fluctuates over the time of day.

**1.3 VHT EFFICIENCY:**

Chart, bar chart

Description automatically generated

Fig 1.5 VHT Efficiency for the month Feb - Mar

The Call Centre has experimented with the Virtual Hold Technology in order to reduce the calls abandoned during last winter. The evidence to show the VHT efficiency is done by aggregating the sample data based on month. The graphical evidence of the VHT efficiency for a particular month is provided in the graph above in Fig 1.5. The stacked bar chart also shows the abandoned calls at every peak hour when both VHT is ‘ON’ & ‘OFF’. It is clearly evident from the graph that number of abandoned calls for Feb – Mar when VHT is ‘OFF’ far less compared to when the VHT is ‘ON’. On analysing the data for the other month as well in the GUI, the VHT ‘OFF’ seems to give the best result comparing to the ‘ON’ data. Thus, according to the given sample data, ‘OFF’ seems to be more efficient

Table

Description automatically generated

Image 1.3 Mean of call abandoned for the Month of Feb – Mar

The Numerical evidence for the VHT efficiency is provided in the table above in the Image 1.3. The values denote the mean of number of abandoned calls. From the image 1.3, it is noted that the during the morning hour of Feb – Mar the average calls abandoned is 0 even when the VHT is ‘OFF’. Comparing the averages of both the VHT functionality, it is clearly visible that the mean value when the VHT is ‘ON’ is greater in all aspects of the day.

Thus, from the graphical and numerical evidence it can be concluded that the Virtual Hold Technology does not achieve its objective of reducing the number of abandoned calls. Hence, I recommend that the Virtual Hold Technology not be used in the future as it will hinder the performance of the call centre and the call centre objective.

**CONCLUSION:**

From the initial analysis, the strength of the relationship between the given variables was identified by the Pearson’s coefficient method. Although, the relationship can be evidently seen through the scatterplot graph it is analysed deeper by segregating each variable and combing with other variables. The Numerical and Graphical evidence of this analysis is detailed in each sub section of the data exploration.

From the Data Exploration, it can clearly seen that Virtual Hold Technology experiment conducted by the call centre has failed miserably because the number of calls abandoned seem to be higher on implementation of VHT. Thus, the company must avoid the usage of VHT in future to improve its performance as well as reduce the number of abandoned calls. On other perspective, call frequency and efficiency of the agency based in average speed of answer provides a deeper knowledge on the call centre activities. From these analysis, the call centre can take initiative steps in the future to perform at its best and also provide a good customer experience and also decrease number of agents when the need for it is lower and increase when required. These explorations give us a descriptive statistic of the given data and helps in the decision planning for the call centre.

On a conclusion note, the call centre is performing efficiently during the peak hours as well as from the data analysis evidence it is proven that agent efficiency is also higher when required. Few changes in the resource planning and turning off the virtual hold technology might help call centre in achieving its objective.