

## Introduction

A large energy provider offers a call center to assist customers. The management of the call center is interested in learning about the activity and performance levels that the call center has seen over the previous winter and what they may change to make the following winter better. They have gathered data for the "peak" hours of the morning, afternoon, and evening because they are aware that activity levels vary significantly throughout the day. They want to have a better understanding of their workloads and how they connect to the experience that customers have when they call the call center. If it is possible, they are especially interested in lowering the number of callers who hang up. They have been experimenting with virtual hold technologies during the past Winter (VHT). When VHT is active, callers who cannot be handled right away are given the option of being rung back without losing their place in line. In this manuscript, the analysis of caller record is done to check the waiting the caller spent to talk to agent, the calls abandoned in waiting of agent, average waiting time agent spend in handling single caller or call offered in that day of the morning and evening.

## Dataset

The dataset contains the following columns data for 504 'peak' hours of operation from last Winter:

- **Month** – refers to three 2-month long periods of time;
- **VHT** – specifies whether Virtual Hold Technology was on or off;
- **ToD** – specifies which peak hour the data is from;
- **Agents** – an indication of the number of agents available to handle calls during the hour (not very accurate);
- **CallsOffered** – number of callers calling during the hour;
- **CallsAbandoned** – number of calls arriving during the hour which rang off before speaking to an agent;
- **CallsHandled** – number of calls arriving during the hour which spoke to an agent;
- **ASA** – average speed of answer (minutes), i.e. average time between the caller first ringing the call centre and speaking to an agents;
- **Avehandletime** – the average time that calls require from an agent, including 'wrap-up' time.

The shape of dataset is (504,9) which mean it contain the 504 instances of 9 columns. As the part of requirement, the sample of 100 random rows is taken from whole dataset to perform analysis

(Univariant and Bivariant) on it. The code 100 sample extraction and head of the dataset is shown in Figure 1 where in figure second cell is for resetting the index values.

df1 = df.sample(100)									
df1 = df1.reset_index().drop(['index'], axis=1)									
df1.head()									
	Month	VHT	ToD	Agents	CallsOffered	CallsAbandoned	CallsHandled	ASA	Avehandletime
0	Feb-Mar	Off	evening	9	57	1	56	0.10	4.9
1	Feb-Mar	Off	afternoon	3	23	0	23	0.00	4.0
2	Dec-Jan	Off	evening	8	58	0	58	0.05	4.6
3	Oct-Nov	On	evening	11	74	1	73	0.38	4.5
4	Oct-Nov	Off	evening	11	67	1	66	0.17	5.0

Figure 1 Code of extracting sample and head of Dataset is shown

## Description of dataset

The statistical description (std, mean, max, min, count) of numerical columns dataset is shown in figure 2. By detail review of the statistical description it can be seen that there are presence of outlier in few columns such as ASA, Call abandon and Calloffered.

	Agents	CallsOffered	CallsAbandoned	CallsHandled	ASA	Avehandletime
count	100.000000	100.000000	100.000000	100.000000	100.000000	100.000000
mean	20.530000	129.480000	4.180000	125.300000	0.925000	5.028000
std	11.435771	70.197050	9.400602	66.846047	1.987360	0.596637
min	3.000000	23.000000	0.000000	23.000000	0.000000	2.800000
25%	10.000000	58.750000	0.000000	58.000000	0.000000	4.800000
50%	21.000000	141.000000	0.000000	141.000000	0.100000	5.000000
75%	29.250000	186.250000	2.250000	183.500000	0.747500	5.400000
max	49.000000	297.000000	49.000000	279.000000	10.280000	6.300000

Figure 2 Description of Dataset

## Analysis

### Univariant Analysis

Uni means "one," thus your data only has one variable. Unlike regression, it doesn't deal with causes or relationships; rather, its main goal is to depict; Data is taken, summarized, and patterns in the data are found. Figure 3 shows the VHT (Virtual Hold Technology) status in 100 sample datasets i.e., out of 100 55 rows are those in which VHT is off and while remaining 45 have status of VHT ON. If we look at figure 4, it shows the call time status most of call are made in evening and afternoon time while morning call are lower in number than other two but keep in mind these are from 100 sample dataset not whole dataset.

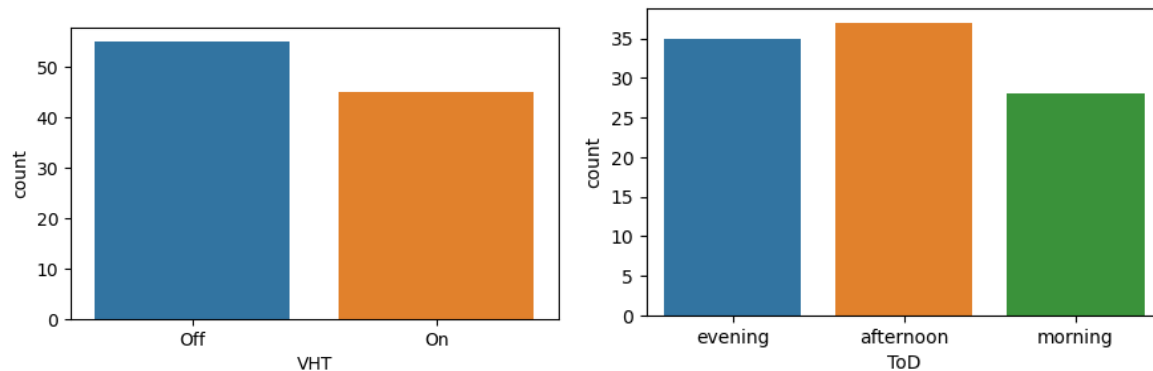


Figure 3 Frequency of call in morning, afternoon and evening in 100 sample dataset

Figure 4 VHT ON and OFF frequency in 100 sample dataset

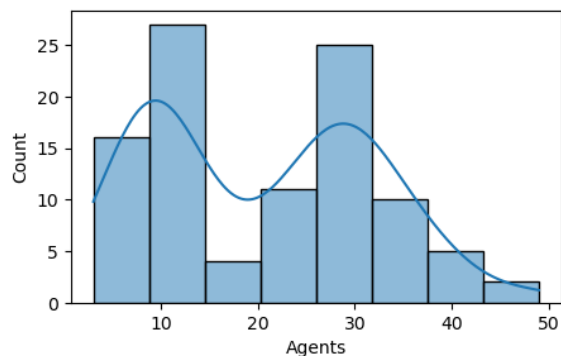


Figure 5 Agents Count

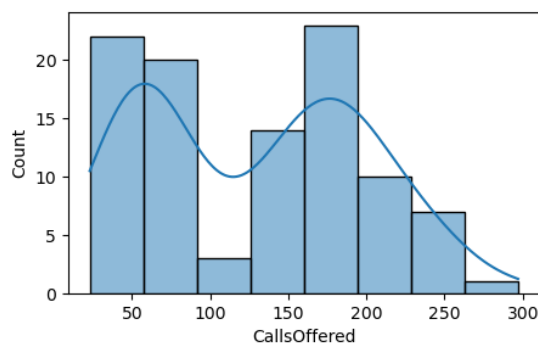


Figure 6 CallsOffered

Figure 5 shows the agents count that is available that time to accommodate caller to resolve their issues. It can be seen that mostly agent have number in 10-15 and 25-35 range while remaining are lower in number. Figure 6 shows the calls offered in each time slot of 100 collected sample. The highest call record at each session are 300 calls which occur 2 times only and mostly 150 to 200 or 30 to 100 calls happens. These two ranges are occurring most in the dataset.

## Bivariant Analysis

When two variables are being studied, you have bivariate data. For instance, if you are researching a group of college students to determine their typical SAT score and age, you need to uncover two parts of the puzzle (SAT score and age). The first bivariant analysis is done between VHT and ToD columns which is shown in figure 7. The analysis shows that most of the times in evening the VHT is off and in the morning the VHT is ON and OFF status is in equal amount. If we look in afternoon then most of the time the VHT is ON. Figure 8 shows the number agents when VHT status is ON and OFF. It can be seen in figure 8 that when VHT is ON then the number of agent operating is higher than the OFF status.

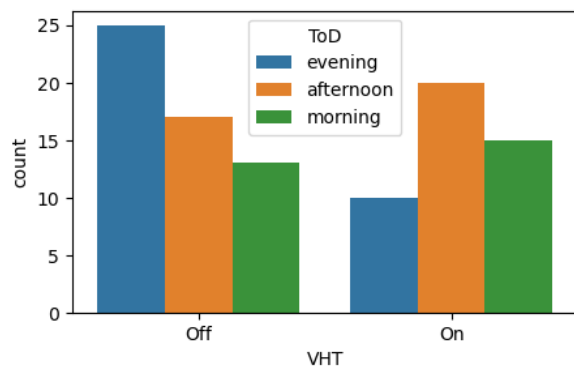


Figure 7 VHT status with respect to ToD

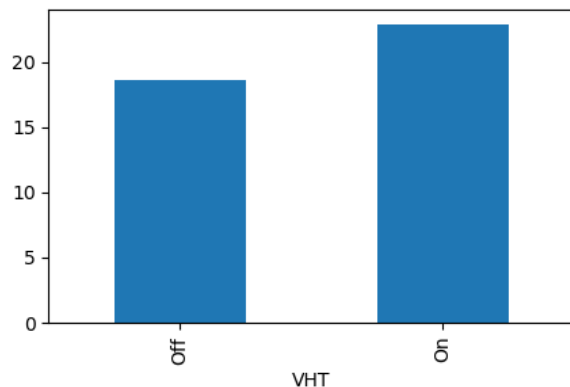


Figure 8 Agents Count When VHT is ON and OFF

If we look at call offered during VHT ON then figure 9 clearly shows that there is more call offered during VHT is ON which mean that VHT technology is allocating more calls then traditional method while the number is slightly above according to 100 sample dataset. While figure 10 shows that ASA (Average speed of answer) increase by using VHT technology which means speed of answering to caller is increased as by using the VHT technology and it increase to double by the case of not using the VHT.

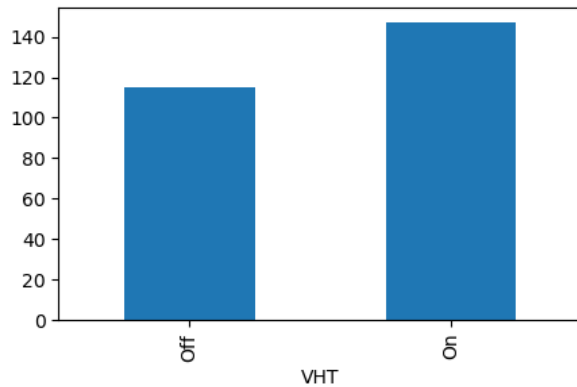


Figure 9 Frequency of Call Offered when VHT ON or OFF

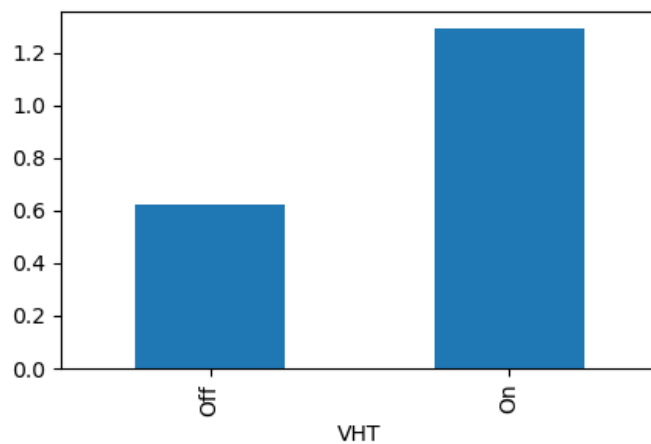


Figure 10 ASA during VHT ON and OFF

Lastly, VHT status is measured across call abandon column. It can be seen in figure 11 that almost double call is abandoned when VHT is ON then when it is OFF which means most of the call are abandon or dropped when VHT is ON.

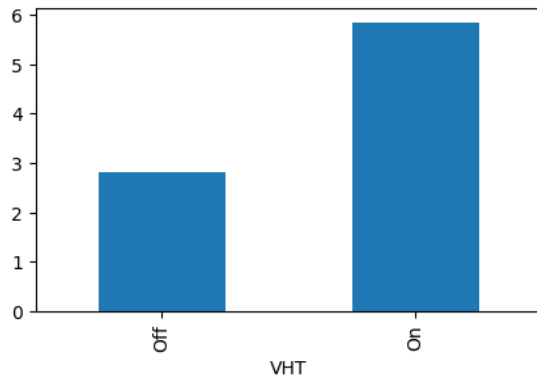


Figure 11 Call Abandon when VHT ON and OFF

## Impact of VHT

The VHT has following impacts on call center

- Most of the caller abandon their call even when VHT is ON, so there is no benefit of using it
- Average handling remains same for both when VHT is ON or OFF
- Average speed of answering is doubled using VHT
- Number of calls offered to agents is increased when VHT is ON
- Number of Agents working also increase when VHT is ON
- In evening, when VHT is OFF most of the call from caller are happen in this time while in morning and afternoon most calls were happened when VHT is ON