

Staffing Plan Review of a Call Center

Major Research Paper

Yining Liu
“212697607”

Submitted in partial fulfillment of the requirements for the Master of
Science in Management Practice
York University

August 7, 2022

Under the supervision of

Rong Bing Huang

Table of contents

1. Abstract	P3
2. Introduction and objectives.....	P3
3. Literature Review.....	P4-8
4. Methodology and Data.....	P9-11
5. Data Analysis and Results.....	P11-27
6. Discussion	P27-29
7. Limitations.....	P29
8. Theoretical and managerial implications.....	P29-30
9. Future directions.....	P30
10. References.....	P31
11. Appendix.....	P32-39

Abstract

This is a case study using statistical tools to examine the operational stats and investigate the deficiencies from the management perspective. A great amount of research on the service level of call centers uses mathematic modelling to generate an optimal staffing plan; whereas this research focuses on analytical framework and equips call center managers with tools to review their center's performances and discover areas for improvement.

Introduction and Objectives

Most research on improving the performance of call center services follows the presumptions of Erlang models and then uses industry strength maths and optimization algorithms to calculate staffing and schedule multiskilled agents. These complicated models are not practical for daily business practice because they do not include the unique determinants such as the call's arrival process and workforce factors including absenteeism and high turnover. To disregard such critical variables will ultimately cause more issues for staff scheduling. Moreover, the different staffing and service nature will influence the service time, the customer's experience, and their wait time in the queue. This research will analyze two weeks of a call center's statistics and offer insights to improve the status quo. The tools used in this research such as Excel Pivot Table and SPSS are

easy to apply to any call center to check on their operational deficiencies and will consequently make contributions to managerial experiences.

Literature Review

There is a large amount of literature review in the call center studies. Gong et al. (1970) studies the optimal staffing of many-server queues with impatient and repeat-calling customers with a call center modelled as an M/M/s+M queue, which is developed as a behavioural queuing model in which customers come and go based on their satisfaction with waiting time. Optimality is defined as the number of agents that maximize revenues net of staffing costs and we account for the characteristic that revenues are a direct function of staffing.

The research considers customer behaviour patterns that traditional models had not; ultimately proposing how managers might optimally staff with various customer behaviour mechanisms. It offers a solution framework for our research, but we would approach the optimal staffing from the data itself instead of assuming the call demands follow a certain model. The research also gives a thorough explanation on the uncertainties and variabilities in call center business environment: “In the call center settings, customer satisfaction of waiting time usually affects the feedback arrival rate. The service encounter is unlike face-to-face service encounters at other service sites, such as restaurants, hotels, and banks. In a call center with invisible queues, there is no service environment, so that the best means of providing and controlling customers' satisfaction may be providing products and service efficiently and quickly. Therefore, waiting time becomes the main

issue that impacts customer satisfaction, hurts repeat business, and jeopardizes the company's long-term profit in call centers.” (Gong et al., 1970)

There are many factors that cause unstable wait times as they are non-face-to-face communication, and it takes varying periods of time for both customer and agent to find the right track upon their comprehensive ability. For example, seniors who are less literate about technology might need more patient explanations than their youth counterparts who can also refer to the website. Therefore, the root cause of the wait time is variability. This renders wait times unpredictable, both from the perspective of the customer as well as from the perspective of the operation.

Whiting and Donthu (2009) conducted a field experiment in a corporation’s call center and found that the higher the estimation error of callers, the lower the satisfaction. Waiting for information reduces estimation error thus extending more on the importance of the wait time: “First, call centers are a major customer touch point that must be managed and it must be managed well. Most of the customer contact occurs through them and the image of the company from the customer’s eyes can either be demolished or enhanced from interactions with a call center. There are major consequences for companies that do not focus on caller satisfaction. Dissatisfied customers are more likely to spend less, switch to competitors, and spread negative word of mouth communication. Organizations that do not focus heavily on caller satisfaction may want to evaluate how much and how important the business activities are that occur through their call centers. For organizations that conduct critical and/or numerous activities through their call center, the stakes are much higher for increasing caller satisfaction. Second, organizations must focus on estimation error. Estimation error occurs frequently with most consumers thinking that they have waited

significantly longer than they actually have. Decreasing estimation error and overestimation should be an initiative for organizations because of its direct impact on customer satisfaction.”

It is crucial for operational managers to understand what can cause the wait time variables and how best to control the estimation of errors both internally and externally, but this research assumes that is the only factor affecting customer satisfaction. The result is a change in arrival rate as an internal factor. As the variability comes not only from customers, our research will also take into account service time as it is of great importance from the perspective of operations management.

Tanaka et al (2013) offers a well-rounded way to generate the staffing plan by solving linear programming based on the minimal number of agents each hour and meeting the service level (of the abandon rate) as constraints. The researcher considers incidents wherein customers cannot receive the services at the initial call time and explores the extent of the re-calling effect by sending an automatic message to the customer that they need to call back later. The linear programming set up a good example for reaching the optimal level of staffing and showcased the necessary and realistic binged conditions. What makes this research standout is that it generates a plan from the data, considering both the call arrival rate and the service time. This is important as the queues build up when all agents are occupied. Only when the next agent is available, the next customer can be served. The unrealistic expectation is that it will fix the minimum agents available, which can be tricky to maintain on a daily basis, considering absences, technical issues, and employee turnover. Also, this research does not include the cost of hiring agents, simply served for a higher service level. There is no universal rule of what service level is right for a given service operation and it all depends on the importance of the incoming calls for the business.

In a call center, the demand always runs ahead of capacity and the customer cannot be served unless there is at least one available agent. From the management view of matching the supply of agents to the demand of calls, we need to investigate more about their variability and the source. Cachon and Terwiesch (2019) summarized four possible factors for variabilities. The biggest source of variability in service organizations come from the market itself. While some patterns of the customer-arrival process are predictable (e.g., in a hotel there are more guests checking out between 8:00 am and 9:00 am than between 2:00 pm and 3:00 pm), there always remains uncertainty about when the next customer will arrive (Cachon and Terwiesch, 2019).

There are also several customer-related uncertainties to consider. One is dependent on the demographics of the customers, for example senior citizens who typically rise early will most likely be available before 9:00 am whereas for full-time professional, lunch hour may be their best time to call. For national call centers where there are time differences, it is important to include these factors to avoid peak time colliding and to maintain service level. Understanding the demographics of customers can also help predict what to expect at different times of the day.

The second factor is the variability in processing times. Whenever we are dealing with human operators at a resource, it is likely that there will be some variability in their behaviour. Thus, if we would ask a worker at an assembly line to repeat a certain activity 100 times, we would probably find that some of these activities were carried out faster than others (Cachon and Terwiesch, 2019).

The differing levels of skillsets and experiences of each individual agent will certainly cause variabilities as well. Those who are experienced can work more efficiently and the service time is

relatively succinct compared with newcomers. Yet, often, the customer his or herself is involved in many of the tasks constituting the processing time. At a hotel front desk, some guests might require extra time (e.g., the guest requires an explanation for items appearing on his or her bill), while others check out faster (e.g., simply use the credit card that they used for the reservation and only return their room key). Therefore, the management should be aware of the possible directions of certain service types and what is the reasonable average service time they should expect. Some call centers have what's referred to as a Call Etiquette Book to standardize the conversation; however, customers might not follow the predictable script of the conversation as designed by the call center, especially when they have a situation to escalate. Realistically speaking, even though the agent tries to control the service time, it can go longer in difficult cases where customers need to be put on hold for agents to corroborate with different parties.

The third factor to raise is the randomized availability of resources. If resources are subject to random breakdowns, for example, machine failures in manufacturing environments or operator absenteeism in service operations, variability is created (Cachon and Terwiesch, 2019). In the past three years, a lot of companies working remotely using VPN have experienced connectivity issues, for instances where the VPN connection fails or any technical issues in general.

The fourth factor is the random routing in case of multiple flow units in the process. If the path a flow unit takes through the process is random itself, the arrival process at each individual resource is subject to variability. Even if arrival times and processing times are deterministic, this random routing alone is subject to introduce variability (Cachon and Terwiesch, 2019). When calls arrive

and go through the routing systems with computer-generated voice information (Voice Response Unit), the VRU automatically answers, queues, and distributes calls to agents in addition to playing relevant announcements. This system is designed to reduce call waiting times and ensure a seamless customer experience by resolving issues quickly and successfully and optimizing call delivery processes while concurrently improving agent morale, productivity, and efficiency. However, it can also trigger the random routing by user prompts and call attributes, such as the language of the caller, agent availability, company department, call volume, and more.

In this research, we will take these factors into the analysis of the data and decide if they play a role in the waiting time for this specific call center.

Methodology

The research is a case study on a call center with two weeks of data and analysis. By analyzing the data within the waiting time problem context, the research aims to find out the specific factors for the waiting time variability in this chosen call center. This will include the variability of waiting time in the queue as well as the variability of the processing time. The analysis consisted of two parts: first, we used SPSS to do descriptive statistics, correlation, and regression. From the results, we process the second part analysis from a micro perspective in terms of the performance of agents and the shifts of coverage.

Data

The research used secondary data, recorded by a graduate student at an anonymous bank in Israel from January 1st, 1999, to December 31st, 1999, from the Internet (Guedj et al., 2000). During weekdays (Sunday to Thursday), the call center is staffed from 7:00 am to midnight. During weekends (Friday-Saturday), it closes at 2:00 pm on Friday and reopens at around 8:00 pm on Saturday. The automated service (VRU) operates seven days a week, 24 hours a day.

The data is free to use for research and interpretation, and there is a PDF document explaining the variables, including:

- 1) vru+line - Each call is assigned a VRU number and a line number
- 2) Call_id Each entering call is assigned a caller id. Although they are different, the IDs are not necessarily consecutive due to being assigned to different VRUs
- 3) Customer_id This is the identification number of the caller, which identifies the customer uniquely; the ID is zero if the caller is not identified by the system.
- 4) Priority - There are two types of customers: 0 and 1 indicate unidentified customers or regular customers (to be elaborated on below) ; 2 indicates priority customers Customers are served in the order of their "Time in Queue". • Priority customers are allocated at the outset of their call 1.5 minutes of waiting-time (in order to advance their position in the queue.)
- 5) Type - 2 digits There are six different types of services: PS - regular activity (coded 'PS' for 'Peilut Shotefet'); PE - regular activity in English (coded 'PE' for 'Peilut English'); IN -

internet consulting (coded 'IN' for 'Internet'); NE -stock exchange activity (coded 'NE' for 'Niarot Erech'); NW - potential customer getting information; TT – customers who left a message asking the bank to return their call but, while the system returned their call, the calling-agent became busy hence the customers were put on hold in the queue.

- 6) Date-month-day
- 7) vru_entry - Time that the phone call enters the call center. More specifically, each calling customer must first be identified, which is done by providing the VRU with the customer-id. Hence this is the time the call enters the VRU.
- 8) vru_exit - 6 digits Time of exit from the VRU: either to the queue, or directly to receive service, or to leave the system (abandonment).
- 9) vru_time - Time (in seconds) spent in the VRU.
- 10) q_start - Time of joining the queue (being put on “hold”). This entry is 00:00:00, for customers who have not reached the queue (abandoned from the VRU).
- 11) q_exit - Time (in seconds) of exiting the queue: either to receive service or due to abandonment.
- 12) q_time - Time spent in the queue (calculated by $q_exit - q_start$)
- 13) Outcome - There are three possible outcomes for each phone call: - AGENT - service - HANG - hung up - PHANTOM - a virtual call to be ignored (unclear to us – fortunately, there are only few of these.)
- 14) ser_start - Time of beginning of service by an agent.
- 15) ser_exit - Time of end of service by agent.
- 16) ser_time - Service duration in seconds (calculated by $ser_exit - ser_start$)

17) Server - text Name of the agent who served the call. This field is NO_SERVER if no service was provided.

The sample is from March 1st, 1999, to March 14th, 1999. Since the research aims to recommend the analysis tools and suggest the management actions on a biweekly basis, there is no need to choose two weeks in particular. The sample is chosen by random sampling. We did data cleaning before analysis by filtering out the wrong records. There are several negative results in 9) vru_time - Time (in seconds) spent in the VRU deleted. As well, we define 7) vru_entry as the arrival time for calls, as it reflects the true time of point when customers demand service. Since it is mentioned that the call center is staffed at different times on different dates, we added a week-day variable manually from Monday to Sunday.

Data analysis and results

The first part of the analysis would display the general information of this data set and offer a start point to check in to the service level. It starts with descriptive statistics. From Table 1 below, we can make a summary of the 859 observations. Most calls are from priority 1, the average vru_time is 10s, q_time is 88s and ser_time is 129s, which means that the average customer who is not on top priority, in general, has to wait for 98s to get a service done in 129s. However, once it is averaged out, we might be able to estimate the total calls of the day, but it does not show us the change of the demand during the different time periods, which is the seasonality. It also does not show the individual customer experience at their own moment of service and if we want to

accurately capture the variability from queue time. We need to look closer into the distribution of the call's arrival.

Statistics	N	Mean	Std. Deviator	Minimum	P25	Median	P75	Maximum
priority	859	0.98	0.897	0	1	1	2	2
date	859	990306.98	4.14	990301	990303	990307	990310	990314
vru_entry	859	14:05:33	4:31:10	0:00:17	10:19:57	13:45:42	17:19:10	23:48:36
vru_exit	859	14:05:55	4:31:05	0:00:33	10:20:00	13:46:57	17:19:25	23:48:46
vru_time	859	10.13	26.276	0	5	6	10	504
q_start	859	10:39:50	7:01:59	0:00:00	7:28:16	11:30:23	15:57:13	23:44:27
q_exit	859	10:41:19	7:02:42	0:00:00	7:27:12	11:30:48	15:58:25	23:44:57
q_time	859	88.61	118.86	0	3	43	133	938
ser_start	859	10:19:28	07:30:56	0:00:00	0:00:00	10:53:48	16:22:55	23:48:45
ser_exit	859	10:21:38	07:32:01	0:00:00	0:00:00	10:57:40	16:23:26	23:48:45
ser_time2	859	129.83	233	0	0	62	154	23:51:56

Table 1 Descriptive Stats for Numerical Variables

In the call center, the demand always runs ahead of capacity, the customer cannot be served unless there is at least one agent available. From the management view of matching the supply of agents to the demand of calls, we need to investigate more about the variability and the source of it.

In the variables of this data set, from the literature review, we can make assumptions that the variability may come from the factors as follows and also assign the variables into these categories.

The inflow of the flow units: weekday, vru_entry, and q_time. The call center has an operational time schedule for different weekdays and the vru_entry time reflects the arrival time of the calls

while the `q_time` reflects the time period until the customer is being served by the next available customer.

Processing time (`ser_time`): This is a straightforward category.

Random availability of resources: (Outcome - AGENT - service - HANG - hung up - PHANTOM)

In the category of random routing in case of multiple flow units in the process, there should be a priority; the priority 2 customer will be pulled from the queue, which causes priorities 1 and 0 to wait even longer. If the system successfully recognizes the customer ID, it will shorten the verification time of the customer. Consequently, the service type may cause the rerouting issue. For example, as we are not sure if every agent is bilingual, if it is a “PE” call that enters the “PS” line, it needs to be transferred. At least `vru_time` can be also a factor; each person goes through system recognition at a different time. Some of them know their ID, some do not and may need more time to resort to paperwork.

Table 2 shows the correlations analysis for both Spearman and Pearson done from SPSS. Most variables are correlated. What is worthy of attention are the insignificant results. The highlighted results are insignificant. From the table, we can observe that there is not a significant correlation between service time and vru_time, neither between service time and q_time, nor service time and date. However, vru_time has a significant correlation with q_time.

Correlations	1	2	3	4	5	6	7	8	9	10	11
priority	1	.091**	0.022	0.021	-.142**	.407**	.407**	.163**	.094**	.095**	.137**
date	.090**	1	.088*	.088*	0.001	.207**	.208**	.237**	-0.033	-0.033	date
vru_entry	0.033	.083*	1	1.000**	-0.004	.403**	.402**	-0.014	.506**	.505**	vru_entry
vru_exit	0.033	.083*	1.000**	1	-0.003	.402**	.402**	-0.014	.506**	.505**	vru_exit
vru_time	-.497**	-.106**	-0.013	-0.012	1	-.205**	-.205**	-.126**	-.127**	-.127**	-.072*
q_start	.361**	.188**	.549**	.549**	-.421**	1	1.000**	.355**	.085*	.084*	q_start
q_exit	.361**	.189**	.549**	.548**	-.421**	1.000**	1	.359**	.085*	.084*	q_exit
q_time	.316**	.228**	0.009	0.009	-.436**	.555**	.558**	1	.071*	.072*	.084*
ser_start	.101**	-0.02	.572**	.572**	-0.027	.207**	.208**	0.019	1	1.000**	.278**
ser_exit	.102**	-0.019	.572**	.572**	-0.027	.207**	.208**	0.019	1.000**	1	.286**
ser_time	.151**	-0.049	0.033	0.033	-0.061	-0.039	-0.038	0.044	.618**	.622**	1
** Correlation is significant at the 0.01 level (2-tailed).											
* Correlation is significant at the 0.05 level (2-tailed).											

The bottom left half of the table contains Pearson's correlation coefficient, whereas the upper right half of the table shows Spearman's correlation coefficient.

Table 2 Correlation Analysis

With this analysis, we can think of vru_time and q_time with their variability as a whole and processing time with its variability as a whole. Next, we need to examine the distribution of call's arrival.

In general, a simple analysis determines whether a process is stationary. First, sort all arrival times so that they are increasing in time (label them as $AT_1 \dots AT_n$), second plot ($x = AT_i; y = i$) as illustrated by Figure 1. Test for Stational Arrival is a full analysis on one day's data on March 3rd.

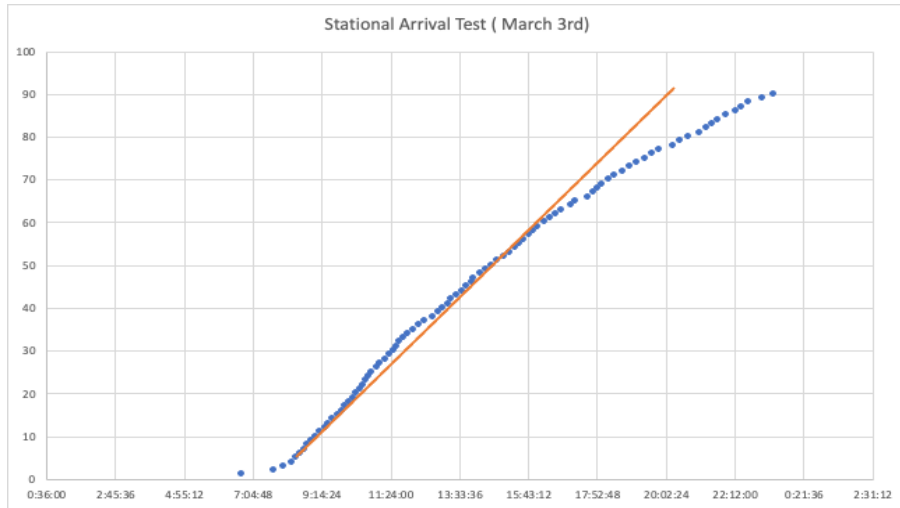


Figure 1 Stational Arrival Test (March 3rd)

The dotted straight line indicates the estimated arrival; therefore, it is observed strong seasonalities with several deviations between the straight line and the actual data. Specifically, it is observed that fewer calls come in compared to the average arrival rate before 8:00 am and after 4:00 pm. Besides, the arrival rate becomes higher and between 10:00 am to 12:00 pm then the rate falls gradually again between 12:00 pm to 2:00 pm. Thus, the analysis indicates that the arrival process for March 3rd is not stationary.

To proceed, it needs to divide up the day (the week, the month) into smaller time intervals and have a separate arrival rate for each interval. For example, if we verify the calls' arrival between 2:30 p.m. to 4:30 pm with 15-minute intervals, we come relatively close to a stationary arrival stream as shown in Figure 2. The stationary behaviour of the interarrivals within a 15-minute interval is illustrated in Figure 2.

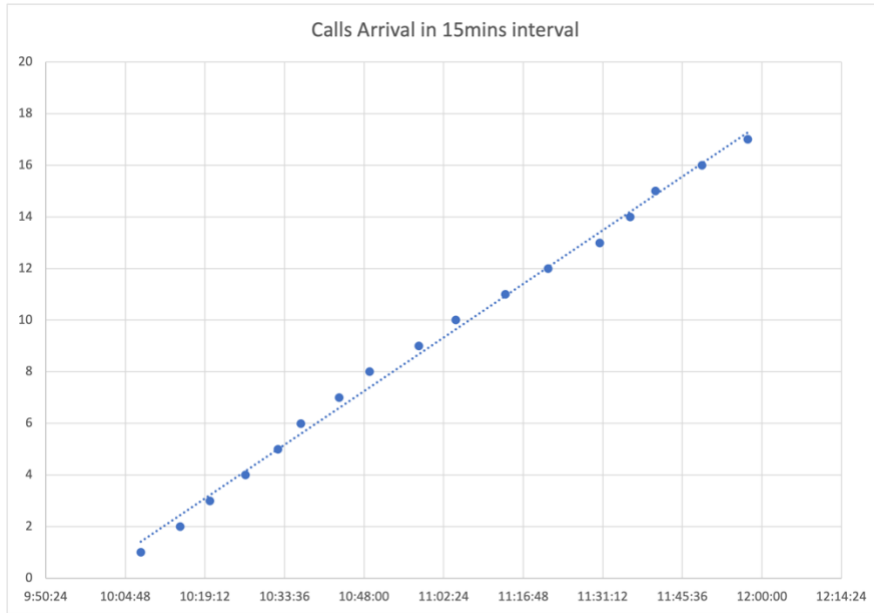


Figure 2 Test for Stational Arrival (15 mins).

As we break down the timeline to small interval, we can tell that the calls follow stational arrival. Next steps are now needed to further examine if the interarrival times falls under Poisson distribution.

First, we computed the interarrival times from IA_1 to IA_{88} . Then we sorted the interarrival times in increasing order; let a_i denote the i th smallest interarrival time (a_1 is the smallest interarrival time; a_{88} is the largest). Plot pairs $(x = a_i, y = i/88)$. The resulting graph is called an empirical distribution function as Figure 3 below.

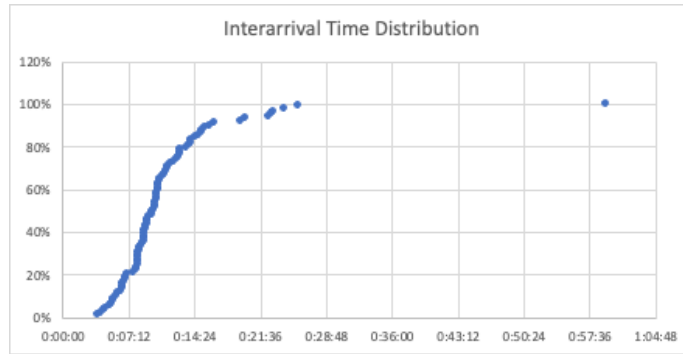


Figure 3. Interarrival Time Distribution

The graph shape is rather ambiguous compared with exponential distribution. Further, we did Chi-Square Test on Excel to check on the distribution.

<i>Descriptive</i>	
Mean	00:11:15
Standard Error	00:00:44
Median	00:09:56
Mode	00:08:05
Standard Deviation	00:06:52
Sample Variance	00:00:02
Kurtosis	06:29:18
Skewness	07:22:05
Range	00:55:34
Minimum	00:03:52
Maximum	00:59:26
Sum	16:41:43
Count	88

<i>Bin</i>	<i>Frequency</i>
0:06:59	17
0:12:59	53
0:18:59	11
0:24:59	6
0:30:59	1
More	0

Bin Item	Bin Range		Observed	Expected	Chi-Square Cal
1	00:00:00	00:06:59	17	41.68	14.61
2	00:07:00	00:12:59	53	61.12	1.08

3	00:13:00	00:18:59	11	72.22	51.90
4	00:19:00	00:24:59	7	82.22	68.82
5	00: 25:00	00:30:59			
				Chi-Square Statistic	136.41
				Chi-Square @5%	5.991464547

Figure 4. Chi-Square Test

Bin 4 and Bin 5 were combined as one since there were only one item in Bin 5. The result shows that the Chi-Square Stat was too high, therefore we must reject the H_0 , the distribution is not exponential distribution. Following our earlier definition of the coefficient of variation, we can measure the variability of an arrival (demand) process as

$$CV_a = \frac{\text{Standard Deviation of interarrival time}}{\text{Average Interarrival Time}} = \frac{00:06:52}{00:11:15} \approx 0.61$$

Next, we need to measure the variability on processing time. Since processing time = $ser_exit - ser_start$, we get all results and look on the shape of the distributions.

<i>Descriptive Statistics</i>	
Mean	135
Standard Error	25
Median	84
Mode	4
Standard Deviation	206
Sample Variance	42308
Kurtosis	34
Skewness	5
Range	1550
Minimum	2
Maximum	1552
Sum	9300
Count	69

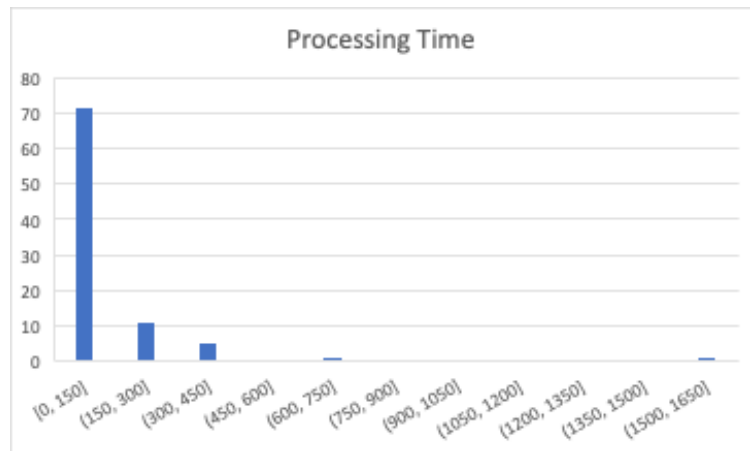


Figure 5. Processing Time Histogram

$$CV_p = \frac{\text{Standard Deviation of processing time}}{\text{Average processing Time}} = \frac{206}{135} \approx 1.51$$

With CV_a and CV_p known, we can proceed with the calculation of waiting time in the queue.

For example, from 14:00:00, there are three calls arriving in 21 mins, average interarrival time $a=7\text{min}=420\text{s}$. Since there are only three agents available, Sharon, Gili, Yifat, therefore $m=3$, from previous calculations we also know processing time $p=206$, then we can get the results:

$$\text{Utilization} = \frac{\text{Flow rate}}{\text{capacity}} = \frac{1/\text{Interarrival time}}{\text{Number of resources}/\text{Processing time}} = \frac{1/a}{m/p} = \frac{p}{a*m} = \frac{206\text{s}}{420\text{s}*3} \approx 0.16$$

$$\text{Time in queue} = \left(\frac{\text{Processing Time}}{m} \right) \times \left(\frac{\text{Utilization}^{\sqrt{2m+1}-1}}{1-\text{Utilization}} \right) \times \left(\frac{CVa^2 + CVp^2}{2} \right)$$

$$= \left(\frac{206}{3} \right) \times \left(\frac{0.16^{\sqrt{2 \cdot 3 + 1} - 1}}{1 - 0.16} \right) \times \left(\frac{0.61^2 + 1.51^2}{2} \right) \approx 5.3s$$

$$\text{Flow time} = \text{Waiting time in queue} + \text{Processing time} = 206 + 5.3 = 211.3s$$

From the above result, we can ascertain that for this call center, the time customer spent in the queue is short compared with the processing time. However, if we check each inbound call's waiting time in 14 days as the graph 4 below, we can identify two groups: about 20% customers did not have to wait and received immediate service and the remaining 80% customers experienced a waiting time strongly resembles an exponential distribution. As we do not have the detailed context for this call center, we are not able to know if this is their ideal service level. However, we can look more into their abandon rate and what factors have a significant effect on the outcome of calls. To investigate these, we used multinomial logistic regression from SPSS.

As you can see, the "ser_time" for both results "agent" and "hang" is statistically significant ($p = 0.025$ and $p = 0.004$, respectively; the "Sig." column). This would suggest that there is a statistical association between service time and the outcome. However,

there is a statistical significance value for "Hang" but not "Agent" for the q_time variable. This would suggest that there is a statistical association between q_time and abandoned calls rather than connected calls. Besides, the variables "priority=1", "type=NW", and "week day=Tuesday" have a significant effect on connected calls.

Outcome		Parameter Estimates							95% Confidence Interval for Exp(B)	
		B	Std. Error	Wald	df	Sig.	Exp(B)	Lower Bound	Upper Bound	
		AGENT	Intercept	16.274	2444.785	0	1	0.995		
	vru_time	0.787	1.013	0.603	1	0.437	2.196	0.302	15.987	
	q_time	-0.006	0.005	1.5	1	0.221	0.994	0.984	1.004	
	ser_time	0.112	0.05	5.011	1	0.025	1.119	1.014	1.234	
	[priority=0]	17.648	2100.757	0	1	0.993	46174148	0	.b	
	[priority=1]	-2.194	1.14	3.705	1	0.054	0.111	0.012	1.041	
	[priority=2]	0c	.	.	0	
	[type=NE]	-1.042	1.446	0.519	1	0.471	0.353	0.021	6.002	
	[type=NW]	-3.976	0.514	59.824	1	<.001	0.019	0.007	0.051	
	[type=PE]	16.251	7165.084	0	1	0.998	11425011	0	.b	
	[type=PS]	0c	.	.	0	
	[Week Day=Friday]	-1.298	0.739	3.086	1	0.079	0.273	0.064	1.162	
	[Week Day=Monday]	-18.5	2444.779	0	1	0.994	9.24E-09	0	.b	
	[Week Day=Saturday]	-0.079	1.135	0.005	1	0.944	0.924	0.1	8.545	
	[Week Day=Sunday]	-17.414	2444.779	0	1	0.994	2.74E-08	0	.b	
	[Week Day=Thursday]	-17.559	2444.779	0	1	0.994	2.37E-08	0	.b	
	[Week Day=Tuesday]	-1.559	0.582	7.184	1	0.007	0.21	0.067	0.658	
	[Week Day=Wednesday]	0c	.	.	0	
HANG	Intercept	17.688	2444.785	0	1	0.994				
	vru_time	0.805	1.013	0.631	1	0.427	2.236	0.307	16.276	
	q_time	-0.01	0.005	4.206	1	0.04	0.99	0.981	1	
	ser_time	-0.123	0.06	4.237	1	0.04	0.885	0.787	0.994	
	[priority=0]	16.446	2100.757	0	1	0.994	13886710	0	.b	
	[priority=1]	-1.05	1.07	0.962	1	0.327	0.35	0.043	2.852	
	[priority=2]	0c	.	.	0	
	[type=NE]	-1.517	1.415	1.148	1	0.284	0.219	0.014	3.516	
	[type=NW]	-2.39	0	.	1	.	0.092	0.092	0.092	
	[type=PE]	16.287	7165.084	0	1	0.998	11844589	0	.b	
	[type=PS]	0c	.	.	0	
	[Week Day=Friday]	-0.348	0	.	1	.	0.706	0.706	0.706	
	[Week Day=Monday]	-17.94	2444.779	0	1	0.994	1.62E-08	0	.b	
	[Week Day=Saturday]	1.564	0	.	1	.	4.778	4.778	4.778	
	[Week Day=Sunday]	-17.164	2444.779	0	1	0.994	3.52E-08	0	.b	
	[Week Day=Thursday]	-16.586	2444.779	0	1	0.995	6.26E-08	0	.b	
	[Week Day=Tuesday]	-0.701	0	.	1	.	0.496	0.496	0.496	
	[Week Day=Wednesday]	0c	.	.	0	

a The reference category is: PHANTOM.

b Floating point overflow occurred while computing this statistic. Its value is therefore set to system missing.

c This parameter is set to zero because it is redundant.

Table 4. Multinomial Logistics Regression

With these findings, we will look at these variables in a micro way with the help of Excel Pivot Table.

Type & Outcome	PHANTOM	HANG	AGENT	Grand Total	Abandon rate
NE	1	18	57	76	24%
NW		83	93	176	47%
PE		2	5	7	29%
PS	4	138	459	600	23%
Grand Total	5	241	613	859	28%

Type	Average of ser_time	StdDev of ser_time	Coefficient of Variation
NE	279	540	1.93
NW	56	131	2.35
PE	188	197	1.05
PS	132	177	1.34
Grand Total	129.83	233.15	1.80

From the table above, it is found that the average abandonment rate is 28% during these two weeks, while the “PE” has the highest abandonment rate 47%, which means the half of English calls were not picked up, even though there is only a minimal amount of them. Moreover, the “NW” stock activity has the highest average service time 279s with the CV of 1.93. It is understandable that “PS” has the highest CV due to the fact that the new customers can be unpredictable, but “NE” might require more quality control to reduce the variabilities.

Priority	AGENT	HANG	PHANTOM	Grand Total	Abandonment Rate
0	236	118		354	33%
1	116	49	3	168	29%
2	261	74	2	337	22%
Grand Total	613	241	5	859	28%

Row Labels	NE	NW	PE	PS	Grand Total
0	12	176	2	164	354
1	17			151	168
2	47		5	285	337
Grand Total	76	176	7	600	859

From these two tables we are able to know that the “NW” new customer type takes half of the priority 0 calls, which also has the highest abandonment rate compared with priority 1 and priority 2. It makes sense that the call center prioritizes the current customers and naturally the higher abandonment can be observed among potential new customers. These calls could also mean new business and further research needs to be conducted to see how effective the calls help to gain new customers. If it is an effective tool to expand the business, the management should take actions to improve the service level with potential customers.

Next, we will investigate the days of the week and staffing. Based on what has been explained by the document, the call center closes at 2:00 pm on Friday and reopens at around 8:00 pm on Saturday.

Weekdays	Count of vru_entry
Sunday	178.0
Monday	118.0
Tuesday	114.0
Wednesday	181.0
Thursday	199.0
Friday	48.0
Saturday	21.0
Grand Total	859.0

If we see the distribution based on the weekdays during the past, Saturday and Friday receive a very small portion of calls while Sunday is really congested with the highest call volumes. The call center should revise the current schedule to make the service more evenly distributed on weekdays. It is important for the call center to consider the cost of running four hours on Saturday night.

Based on what has been explained by the recorder, the call center constitutes of: eight agent positions, one shift-supervisor position, five agent positions for internet services (in an adjacent room), but we do not have any more information regarding the shifts schedule or the agents' skillset. The only possible way to make assumptions is to filter out all types of calls the agent pick up and check his/her skillset.

Agent Name	NE	NW	PE	PS	Grand Total
SHARON	3	20		44	67
STEREN		12		31	43
MIKI		11	2	31	44
SHLOMO		9		22	31
MORIAH	4	8		35	47
AVNI	1	8		16	25
KAZAV	7	6		32	45
TOVA	7	5		32	44
IDIT	3	4		23	30
YIFAT		3	1	39	43
PINHAS		3			3
BASCH	2	1		21	24

ELI	14			4	18
ZOHARI	7			23	30
GILI	4			19	23
AVIDAN	1			6	7
YITZ			2	27	29
NAAMA				7	7
DARMON				7	7
GELBER				6	6
ANAT	4			29	33

Table 5. Calls summary in type by Agents

The table shows the stats that the types of calls the agents served. Once we further organize it, we can get tables below to assign an agent to four categories: Multi Skill Agent (Unilingual), Two-Skill Agent, Single Skill Agent, and Bilingual Agent.

Multi Skill Agent (Unilingual)					
Agent	NE	NW	PE	PS	Grand Total
SHARON		3	20		44
MORIAH		4	8		35
AVNI		1	8		16
KAZAV		7	6		32
TOVA		7	5		32
IDIT		3	4		23
BASCH		2	1		21

Two-Skill Agent					
Agent	NE	NW	PE	PS	Grand Total
STEREN		12		31	43
SHLOMO		9		22	31
ELI	14			4	18
ZOHARI	7			23	30
GILI	4			19	23
AVIDAN	1			6	7
ANAT	4			29	33

Single Skill Agent					
Agent	NE	NW	PE	PS	Grand Total
PINHAS		3			3
NAAMA				7	7
DARMON				7	7
GELBER				6	6

Bilingual Agent					
Agent	NE	NW	PE	PS	Grand Total
MIKI		11	2	31	44
YIFAT		3	1	39	43
YITZ			2	27	29

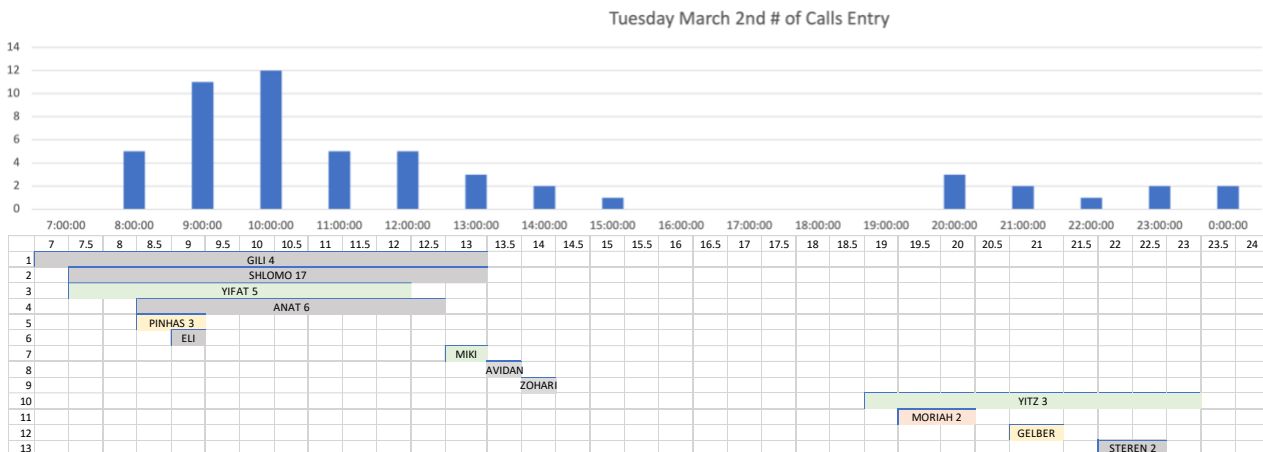
We recognize that no single agent in this call center can cover all types of calls, maximum three skills. For the single skilled agent with small number of calls, it can be assumed that it is a supervisor onsite to handle the difficult customers or escalations in general.

To further address Tuesday's concerns, we also need to know which staff were on duty and their specific shifts that day. We took the starting time of the first call, "min of vru_entry" and the ending time of the last call 'max of sev_exit" on Tuesday, March 2nd, 1999, and made the table as below. There is no other way to restore the employee's shift time and we can only assume that the

closest clock or half time will the starting or ending point of the shift. For example, Gili might started from 7:00 am and finished 1:00 pm.

Agent	Min of ser_start	Max of ser_exit
GILI	07:10:51 AM	12:58:59 PM
SHLOMO	07:39:47 AM	12:41:34 PM
YIFAT	07:58:42 AM	11:39:52 AM
ANAT	08:19:54 AM	12:09:52 PM
PINHAS	08:39:48 AM	08:56:38 AM
ELI	09:16:12 AM	09:20:20 AM
MIKI	01:12:46 PM	01:13:27 PM
AVIDAN	01:29:41 PM	01:31:07 PM
ZOHARI	02:00:44 PM	02:03:53 PM
YITZ	07:23:54 PM	11:04:34 PM
MORIAH	07:45:48 PM	08:06:05 PM
GELBER	08:53:47 PM	08:54:50 PM
STEREN	10:09:13 PM	10:40:37 PM

Based on this table, we can restore the shifts on Tuesday and then compare it with the calls distribution on Tuesday.



There is no multi-skilled agent scheduled in the busy morning. Yifat is the only agent who can help NW calls in the morning.

Discussions

The call center has a relatively short queue time, but the abandonment rate is high from March 1st, 1999 to March 14th, 1999. The variabilities mainly come from the processing time due to the business schedule and the shortage of the agents and limited skillset. The recommendations are as follows:

Use the morning time and evening time of the call center: From calls distributions on weekdays (in Appendix), we can see on Monday, Wednesday, Thursday and Sunday, there are two peaks of volume. But before 9:00 am and after 6:00 pm, the calls entered tremendously decreased. Before 9:00 am, there are two agents, while during the morning peak time, they only schedule two more agents. It is unreasonable and unproductive to create this idle time in the morning. If they cannot hire more agents to help with the peak time, they should offer incentive to shift calls to the time before 8:00 am and after 7:00 pm, especially for priority 0 and 1.

Tuesday has three hour shut-down time, and it is unknown why. It could be system maintenance or depart meeting. The bank can choose to keep partial team available and take turns to join the meeting. Or if they necessarily need the entire team for participance, the meeting should be moved to Friday afternoon to extend the business hour longer. At the same time, Saturdays should be closed entirely due to the low volume of calls.

The type of “PE” is not everyday demand and should be changed to “appointment” system. For this call center, based on how little number of English calls, being bilingual is not the prioritized qualification for agents. In general, the bilingual agents are paid in a higher rate and the management should use the language talent in a focused way by setting a “English call day” once or twice week. On the other hand, there is no need to schedule the bilingual rep on other days. This aims to regulate the demand and have better control of the follow.

Moreover, NE and PS calls mainly come in between 9:00 am to 5:00 pm, but NW is distributed quite evenly among different time of the day and different day in a week. The scheduler should be aware of these when planning the shifts. There is obvious shortage in their shift scheduling, where the certain call types enter the system, but the agent does not have that skill.

Limitations

The biggest limitation of this case study is how little information we have on the data itself. The research was conducted with a lot of deductive method to restore the real working environment, and it can cause a lot of inaccuracy. For examples, under the call’s coverage topic, we only made assumptions of the employee’s skills based on the data. Yet, just based on data, there is still a lot of information missing, such as agents’ break, agent’s salary, agent’s rating model based on the call quality, and whether this call center has outbound calls or not. If we have more information about the work type content and expectations on agents, we could give more concrete advice on agent’s performance. However, in real business practice, this missing information should be highly feasible for management to obtain to make comprehensive decision on improving the call center’s efficiency.

Another limitation comes from data itself. It is outdated and the daily amount of calls are relatively small compared with the recent research average 200 calls per day in past decade. For recent call centers with a large number of inbound calls, the tools we used in this case study should be effective to diagnose the bottle neck in the system.

Theoretical and managerial implications

The research displayed a framework of stats analysis for call center managers. It reviewed the stats in both macro and micro perspectives, which is a good model for call centers to do bi-weekly review on the service performances. All the tools and analysis used in this research are actionable, practical, and easy to apply.

Future Directions

In this dataset, the sample we took unfortunately did not take into any “TT”– customers who left a message asking the bank to return their call but while the system returned their call, the calling-agent became busy hence the customers were put on hold in the queue. If we continuously do the extended research, we can investigate the effect of “TT”.

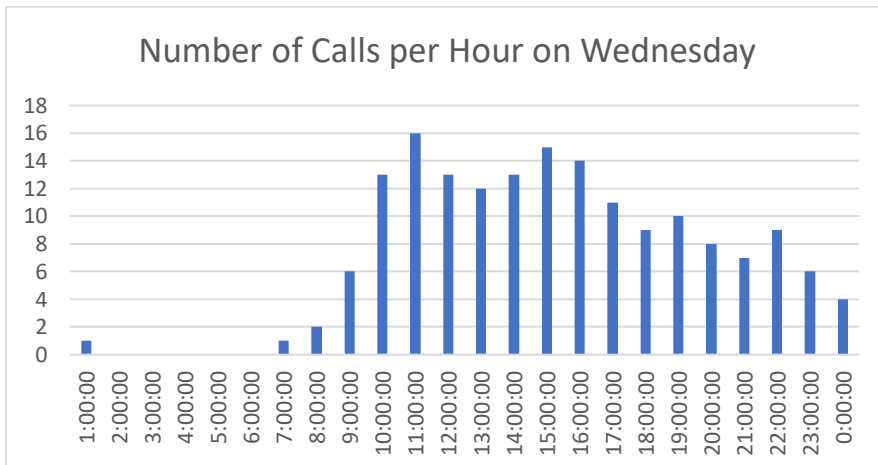
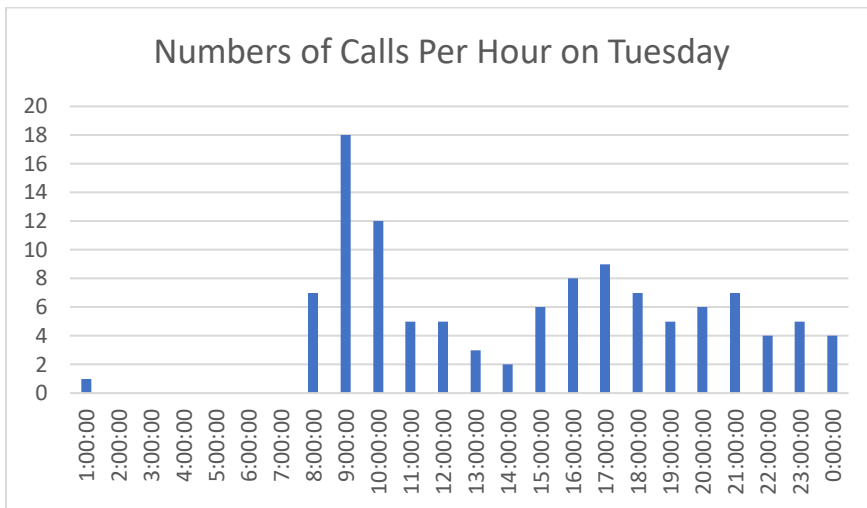
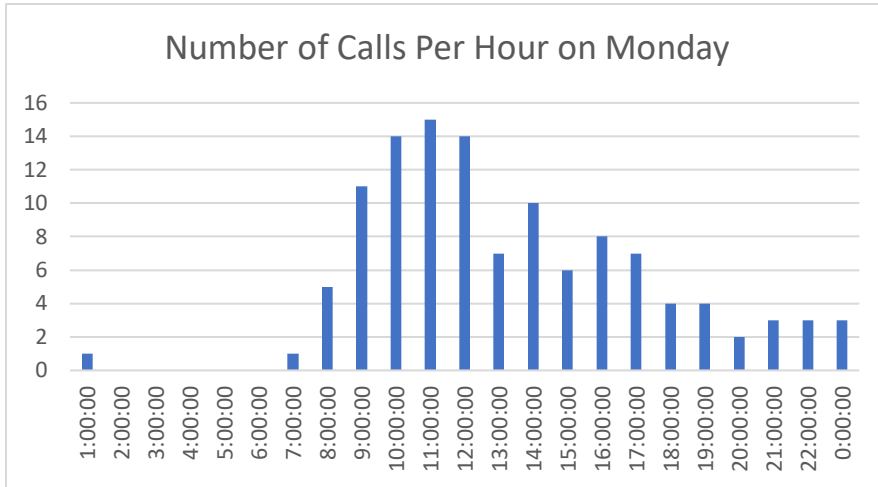
The priority in this dataset is very much close to the idea of “appointment.” The priority customers were pulled to serve from system, while the regular and unidentified were required to wait longer. Judging from the data, the priority 2 customer’s average queue time is also above average, and the service time is longest with the biggest coefficient of variation. These types of calls need more

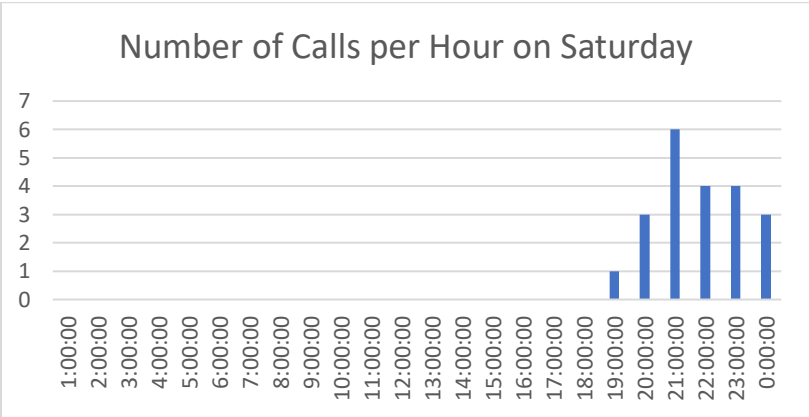
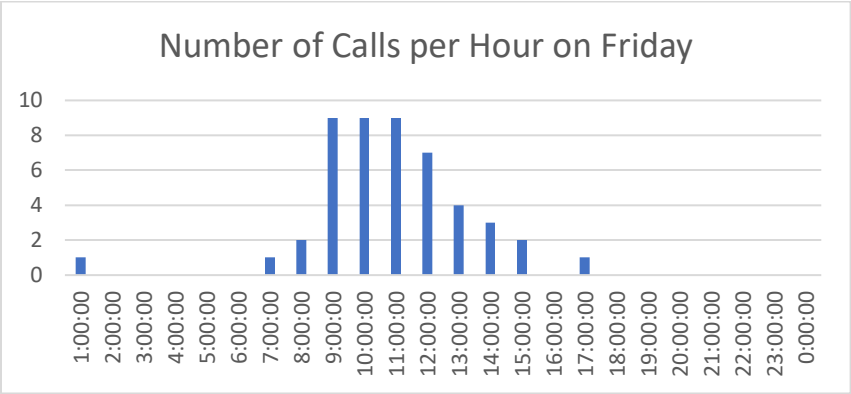
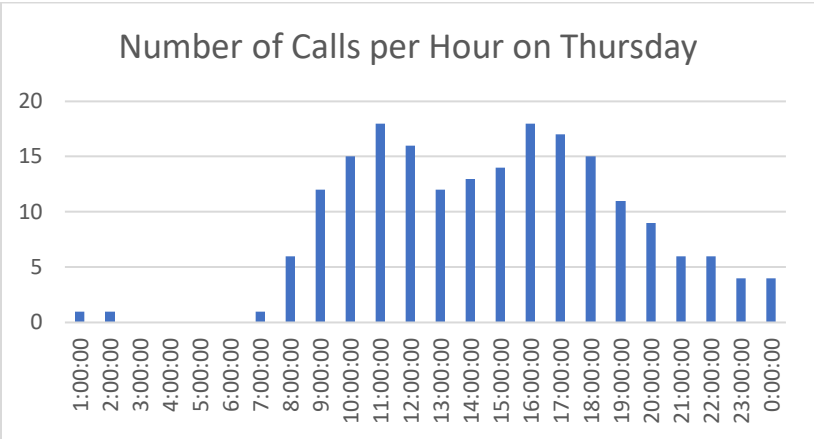
content to diagnose why it is like this. The call center potentially needs to restructure priority calls and the call routine to reduce the service time and variability.

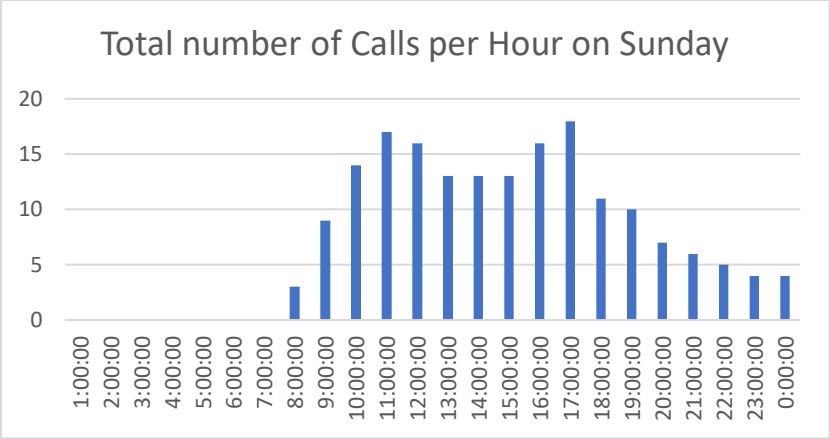
References

- Gong, J., Yu, M., Tang, J., & Li, M.-hon. (1970, January 1). [PDF] *staffing to maximize profit for call centers with impatient and repeat-calling customers: Semantic scholar*. undefined. Retrieved July 27, 2022, from <https://www.semanticscholar.org/paper/Staffing-to-Maximize-Profit-for-Call-Centers-with-Gong-Yu/922cf4bf94c9a684b9094f44baea83f50f15dfcc>
- Whiting, A., & Donthu, N. (2009, July 31). *Closing the gap between perceived and actual waiting times in a call center: Results from a field study*. *Journal of Services Marketing*. Retrieved July 27, 2022, from <https://www.emerald.com/insight/content/doi/10.1108/08876040910973396/full/html>
- Tanaka, M., Luo, H., Inada, S., & Nakatsuka, A. (2013, March). *A Case Study on Optimal Allocation of Call Center Staff —With Un-used Phones and Repeated Calls Situation Considered*. *J-Stage*. Retrieved July 27, 2022, from https://www.jstage.jst.go.jp/article/iscm/7/1/7_25/_pdf/-char/en
- Cachon Gérard, & Terwiesch, C. (2019). *Variability and Its Impact on Process Performance: Waiting Time Problems*. In *Matching supply with demand: An introduction to operations management* (3e ed., pp. 144–177). essay, McGraw-Hill Education.
- Guedj, I., & Mandelbaum, A. (2000, February 9). *"Anonymous Bank" Call-Center Data Documentation by Ilan Guedj and Avi Mandelbaum February 9, 2000*. *Call center data*. Retrieved August 4, 2022, from <https://iew.technion.ac.il/serveng/callcenterdata/index.html>

Appendix







"Anonymous Bank" Call-Center Data

Documentation by Ilan Guedj and Avi Mandelbaum February 9, 2000

This document describes telephone data, recorded over 12 month (from 1/01/99 till 31/12/99), at the telephone call-center of "Anonymous Bank" in Israel.

The data was organized by Ilan Guedj, who was at the time a graduate student of the Faculty of Industrial Engineering and Management at the Technion, Haifa.

- The data is free for use. If used, please notify Avi Mandelbaum, at avim@ie.technion.ac.il and do acknowledge the data source in any of your work.
- A sample of the data, as an EXCEL worksheet, is included on the last page of the document.
- All files are zipped; after unzipping, they become an ASCII file, namely Plain Text (*.txt).
- The Internet site <http://ie.technion.ac.il/~serveng> contains a lot of material that is based on analysis of the present data.

General Description

The call center of "Anonymous Bank" provides several different services:

- Information on and transactions of checking and saving, to bank-customers
- Computer generated voice information (through VRU = Voice Response Unit)
- Information for prospective customers
- Support for the customers of "Anonymous Bank" web-site (internet customers)

The call center constitutes of:

- 8 agent positions
- 1 shift-supervisor position
- 5 agent positions for internet services (in an adjacent room)

During weekdays (Sunday to Thursday), the call center is staffed from 7:00am to midnight. During

weekends (Friday-Saturday), it closes at 14:00 on Friday and reopens at around 20:00 on Saturday. The automated service (VRU) operates 7 days a week, 24 hours a day.

Data Structure

The data archives all the calls handled by the call center, over the period of 12 months from January 1999 till December 1999.

The data consists of 12 files, a file per month. Each file consists of records (lines), a record per phone call (between 20,000 to 30,000 calls per month). Each record has 17 fields, which will now be described in details.

1) vru+line - 6 digits

Each entering phone-call is first routed through a VRU: There are 6 VRUs labeled AA01 to AA06. Each VRU has several lines labeled 1-16. There are a total of 65 lines. Each call is assigned a VRU number and a line number.

2) Call_id - 5 digits

Each entering call is assigned a call id. Although they are different, the id's are not necessarily consecutive due to being assigned to different VRUs.

3) Customer_id - 0 to 12 digits

This is the identification number of the caller, which identifies the customer uniquely; the ID is zero if the caller is not identified by the system (as is the case for prospective customers, for example).

4) Priority - 1 digit

The priority is taken from an off-line file.

There are two types of customers: (high-)priority and regular:

- 0 and 1 indicate unidentified customers or regular customers (to be elaborated on below)
- 2 indicates priority customers

- Customers are served in the order of their "Time in Queue".
- Priority customers are allocated at the outset of their call 1.5 minutes of waiting-time (in order to advance their position in the queue.) They are also exempt from paying a NIS 7 monthly fee, which regular customers must pay.
- Customers have not been told about the existence of priorities.
- Until August 1996, all the customers had the same priority - 0. Priorities 1 and 2 were introduced in August 1st, 1996. There still are 0 priority customers, but they are treated as Priority 1. (As we understand it, priority 0 corresponds to those customers that were assigned priority 0 before August 1st and whose priority has not been upgraded.)
- Due to a system bug, customer I.D. was not recorded for those who did not wait in queue, hence, their priority is 0.

5) Type - 2 digits

There are 6 different types of services:

- PS - regular activity (coded 'PS' for 'Peilut Shotefet')
- PE - regular activity in English (coded 'PE' for 'Peilut English')
- IN - internet consulting (coded 'IN' for 'Internet')
- NE -stock exchange activity (coded 'NE' for 'Niarot Erech')
- NW - potential customer getting information
- TT – customers who left a message asking the bank to return their call but, while the system returned their call, the calling-agent became busy hence the customers were put on hold in the queue.

6) Date - 6 digits

year-month-day

7) vru_entry - 6 digits

Time that the phone-call enters the call-center. More specifically, each calling customer must

first be identified, which is done by providing the VRU with the customer-id. Hence this is the time the call enters the VRU.

8) vru_exit - 6 digits

Time of exit from the VRU: either to the queue, or directly to receive service, or to leave the system (abandonment).

9) vru_time - 1 to 3 digits

Time (in seconds) spent in the VRU (calculated by exit_time – entry_time) .

10) q_start - 6 digits

Time of joining the queue (being put on “hold”). This entry is 00:00:00, for customers who have not reached the queue (abandoned from the VRU).

11) q_exit - 6 digits

Time (in seconds) of exiting the queue: either to receive service or due to abandonment.

12) q_time - 1 to 3 digits

Time spent in queue (calculated by q_exit – q_start)

13) Outcome - 4,5 or 7 digits

There are 3 possible outcomes for each phone call:

- AGENT - service

- HANG - hung up

- PHANTOM - a virtual call to be ignored (unclear to us – fortunately, there are only few of these.)

14) ser_start - 6 digits

Time of beginning of service by agent.

15) ser_exit - 6 digits

Time of end of service by agent.

16) ser_time - 1 to 3 digits

Service duration in seconds (calculated by ser_exit – ser_start)

17) Server - text

Name of the agent who served the call. This field is NO_SERVER, if no service was provided.

Sample examples

vru+line	call_id	customer	priority	type	date	Week Da	vru_entry	vru_exit	vru_t	q_start	q_exit	q_time	outcom	ser_start	ser_exit	ser_t	server
AA0101	36227	0	0	PS	990301	Monday	0:14:46	0:15:01	15	0:00:00	0:00:00	0	HANG	0:00:00	0:00:00	0	NO_SERVER
AA0101	36228	0	0	PS	990301	Monday	7:19:13	7:19:23	10	0:00:00	0:00:00	0	AGENT	7:19:22	7:19:29	7	NO_SERVER
AA0101	36229	0	0	NW	990301	Monday	7:40:24	7:40:37	13	0:00:00	0:00:00	0	AGENT	7:40:36	7:41:01	25	MIKI
AA0101	36230	0	0	PS	990301	Monday	7:54:44	7:54:54	10	0:00:00	0:00:00	0	AGENT	7:54:53	7:57:41	168	TOVA
AA0101	36231	0	0	PS	990301	Monday	8:09:10	8:09:19	9	0:00:00	0:00:00	0	AGENT	8:09:18	8:09:23	5	NO_SERVER
AA0101	36232	0	0	NW	990301	Monday	8:22:31	8:22:45	14	0:00:00	0:00:00	0	AGENT	8:22:44	8:23:15	31	MIKI
AA0101	36233	0	0	NW	990301	Monday	8:30:00	8:30:09	9	8:30:09	8:30:13	4	AGENT	8:30:12	8:31:00	48	MIKI
AA0101	36234	0	0	NW	990301	Monday	8:35:34	8:35:43	9	8:35:43	8:36:25	42	AGENT	8:36:24	8:38:22	118	MIKI
AA0101	36235	0	0	PS	990301	Monday	8:42:31	8:42:42	11	0:00:00	0:00:00	0	AGENT	8:42:40	8:45:54	194	MIKI
AA0101	36236	0	0	NW	990301	Monday	8:48:28	8:48:37	9	8:48:37	8:50:14	97	HANG	0:00:00	0:00:00	0	NO_SERVER
AA0101	36237	0	0	NW	990301	Monday	9:01:06	9:01:21	15	0:00:00	0:00:00	0	AGENT	9:01:20	9:01:52	32	STEREN
AA0101	36238	0	0	PS	990301	Monday	9:10:19	9:10:31	12	0:00:00	0:00:00	0	AGENT	9:10:30	9:13:36	186	STEREN
AA0101	36239	2E+07	2	PS	990301	Monday	9:17:16	9:17:22	6	9:17:22	9:17:34	12	AGENT	9:17:33	9:20:08	155	NAAMA
AA0101	36240	0	0	NW	990301	Monday	9:22:32	9:22:41	9	9:22:41	9:23:53	72	AGENT	9:23:53	9:27:04	191	MIKI
AA0101	36241	6E+07	2	PS	990301	Monday	9:32:29	9:32:34	5	9:32:34	9:33:29	55	AGENT	9:33:28	9:33:51	23	BASCH
AA0101	36242	6E+07	1	PS	990301	Monday	9:40:45	9:40:50	5	9:40:50	9:41:01	11	AGENT	9:41:00	9:43:05	125	MIKI
AA0101	36243	6E+07	1	PS	990301	Monday	9:49:00	9:49:06	6	9:49:06	9:50:40	94	AGENT	9:50:40	9:51:22	42	YITZ
AA0101	36244	0	0	PS	990301	Monday	9:58:36	9:58:47	11	0:00:00	0:00:00	0	AGENT	9:58:46	9:59:51	65	KAZAV
AA0101	36245	0	0	PS	990301	Monday	10:06:44	10:06:57	13	0:00:00	0:00:00	0	AGENT	10:06:55	10:10:08	193	TOVA
AA0101	36246	7E+07	2	PS	990301	Monday	10:14:35	10:14:41	6	10:14:41	10:14:58	17	AGENT	10:14:56	10:15:47	51	BASCH
AA0101	36247	8E+07	2	PS	990301	Monday	10:19:54	10:20:00	6	10:20:00	10:20:12	12	AGENT	10:20:10	10:24:03	233	KAZAV
AA0101	36248	3E+07	1	PS	990301	Monday	10:24:02	10:24:08	6	10:24:08	10:24:26	18	HANG	0:00:00	0:00:00	0	NO_SERVER
AA0101	36249	3E+07	1	PS	990301	Monday	10:29:18	10:29:23	5	10:29:23	10:30:12	49	AGENT	10:30:12	10:31:14	62	YIFAT
AA0101	36250	6E+07	2	PS	990301	Monday	10:37:23	10:37:28	5	10:37:28	10:38:16	48	AGENT	10:38:15	10:39:09	54	KAZAV
AA0101	36251	0	0	PS	990301	Monday	10:48:15	10:48:25	10	0:00:00	0:00:00	0	AGENT	10:48:23	10:50:53	150	KAZAV
AA0101	36252	0	0	PS	990301	Monday	10:56:56	10:57:07	11	0:00:00	0:00:00	0	AGENT	10:57:05	10:57:40	35	NAAMA
AA0101	36253	0	0	PS	990301	Monday	11:02:31	11:02:40	9	0:00:00	0:00:00	0	AGENT	11:02:39	11:04:22	103	NAAMA
AA0101	36254	0	0	NE	990301	Monday	11:12:09	11:12:22	13	0:00:00	0:00:00	0	AGENT	11:12:21	11:12:24	3	GIU