Call Center Regression Data Analysis

Michael Marcaccio

November 14 2022

Abstract

With over 15 million people employed, a compound annual growth rate (CAGR) of 5.6 percent between 2020 and 2027, and 28,000 in the United States alone, call centers play a pivotal role in a businesses success.In this paper, call center volume was forecasted and models were created to predict the number of agents needed to meet critical attributes such as waittime, calltime, and holdtime. Forecasting gives businesses the ability to make informed business decision and develop data-driven strategies. In the literature it is very common to see call center volume being predicted using different techniques, but there is very limited studies on attempting to correlate the number of agents needed. Through Regression techniques, there are 8 models proposed to model the number of agents needed based on waittime, calltime, goaltime, and the amount of calls handled.

Introduction

With over 15 million people employed, a compound annual growth rate (CAGR) of 5.6 percent between 2020 and 2027, and 28,000 in the United States alone, call centers play a pivotal role in a businesses success. Call Centers are a key part of customer service, that will save a company time, money, and unneccessary stress. In the banking industry, calls can range from inquires, transfers, payments, reporting, to processing. This means members can be calling about their account balance, credit card bills, loan applications, or unauthorized transactions. It is crucial that a bank is prepared for spikes in calls and have agents knowledgeable in all aspects of the bank.

Many studies have been done working with call center volumes such as Modeling and Forecasting Call Center Arrivals (Ibrahim et al., 2016). Here an Autoregressive moving average (ARIMA) model is used, which is depicted in the paper as:

$$\Phi(B)(x_i - \mu) = \theta_q(B)\varepsilon_t \tag{1}$$

Ibrahim uses multiple ARIMA models to depict seasonality with a combonination of exponential smoothing. Holt-Winters smoothing is also used, with three equations of:

$$M_t = \alpha_0(X_t - S_{t-s}) + (1 - \alpha_0)(M_{t-1} + B_{t-1}), \tag{2}$$

$$B_t = \alpha_1(M_t - M_{t-1}) + (1 - \alpha_1)B_{t-1}, \tag{3}$$

$$S_t = \alpha_2(X_t - M_t) + (1 - \alpha_2)S_{t-s}, \tag{4}$$

where B_t is the slope component, M_t is the level component, and S_t is the seasonal component, and s is the period of seasonality. Arima models are a great model to create as "they allow the representation of a wide array of potentially useful predictor functions in models which contain relatively few parameters" (Newbold, 1983). With the combination of Holt-Winters smoothing, Ibrahim creates a great model for forecasting call volume, however the paper never goes into detail about how many agents there should be.

(Evensen et al., 1999) goes into detail about effectve service delivery stating the expecations of customers are of "function of the customer's own experiences...and when judging their own service quality, financial instiutions need to evalute themselves on objective measures which span across industries".When taking this into consideration, I focused my models on using a calls per agent approach, so agents are not overworked and can treat each member with respect and honor to represent the business in good light. (Avramidis and L'Ecuyer, 2005) confirms this as it is described as the "call-to-agent" problem, and that an efficiency-driven call center is important.

The contributions of the paper creates models on the basis to maximize quality of the service in relation to the number of agents. Based on this, models were created to minimize:

1. WaitTime
2. CallTime
3. Reduce GoalTime not being Met
4. Reduce Calls not being Handled

The rest of the paper will go through the dataset and breaking it up to present a clear view on the historical data. I have also created a forecast on the projected call volume throughout a week through a 15 minute interval.

Data

The data comes from an undisclosed bank where I wrote two SQL queries to obtain. This comes directly from the banks records dating back from Febuary 2020 to Septemeber 2022, and I was lucky enough to do a data analysis on real-life data. This dataset consists of 466,565 observations of 31 variables. Some significant variables consist of Calldate, Caller_ID, Caller_ID, QueuedYN, AnsweredYN, AbandonedYN, QueuedName, InteractionOutcome, AnsweredGoalMet, StartTime, EndTime, WaitTime, Holdtime, AgentTalkDuration, AgentHandleTime, WrapUpTime, DayoftheYear, WeekoftheYear, and NbrofAgents.

It is important to note that AgentTalkDuration refers to the actual amount of time the agent the agent was talking while AbandonedYN refers to if the caller ended the call before speaking to a representative. The QueuedYN is if the member was put in a queue and had to wait to speak to an agent. The queue is an automated service or an IVR (Interactive Voice Response). The latest generation of speech-recognition technology allows IVRs to interpret complex user commands, so customers may be able to "self-serve", i.e., complete the service interaction at the IVR. (Avramidis and L'Ecuyer, 2005). Holdtime is how long the member waits when an agent places on them hold, while waittime is how long they have to wait before someone answers there call. WrapUpTime refers to how long the member takes to hang up the phone after the agent stops talk. AgentHandleTime is the total of Holdtime, WrapUpTime, and AgentHandleTime.

Methods

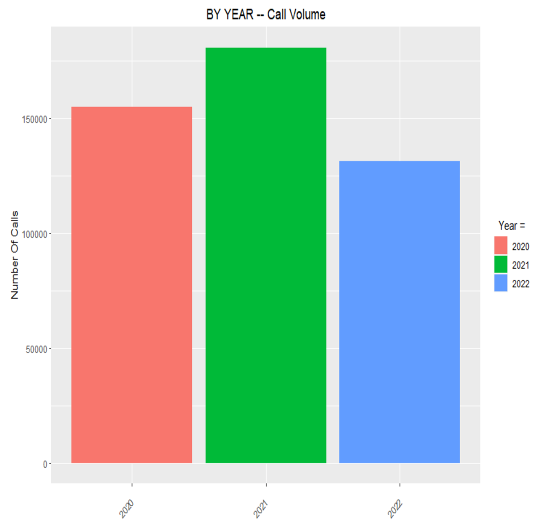

Figure 1: This is my first figure.

Simulation

Discussion

References

Avramidis, A. N. and P. L'Ecuyer (2005). Modeling and simulation of call centers. In *Proceedings of the Winter Simulation Conference, 2005.*, pp. 9–pp. IEEE.

Evensen, A., F. X. Frei, and P. T. Harker (1999). *Effective call center management: evidence from financial services*. Citeseer.

Ibrahim, R., H. Ye, P. L'Ecuyer, and H. Shen (2016). Modeling and forecasting call center arrivals: A literature survey and a case study. *International Journal of Forecasting 32*(3), 865–874.

Newbold, P. (1983). Arima model building and the time series analysis approach to forecasting. *Journal of forecasting 2*(1), 23–35.