FORECASTING AND OPTIMIZING CALL CENTER STAFFING

Hanzhang Fu | Jessie Le | Jieying Wang | Parika Gupta

Faculty: Dr. Hande Benson

Table of Contents

Executive Summary

A team of Drexel LeBow graduate students investigated the Interactive Voice Response Calls dataset, call center activity dataset and the case studies in the energy and electric industry to answer the following questions.

- How many staff members are needed at the call center to maintain an acceptable level of service at minimum cost? (80% of calls handled within 30 seconds)
- How can the model adjust to changes happening in real-time business conditions?

Detailed data analysis showed that:

- There are extremely high number of calls in the week 1 of March 2018 because of Snowstorms.
- There were a greater number of calls on July 22nd, 2019, as compared to the whole month because of heavy rainfall and strong winds.
- After removing the outliers, the distribution of the calls become stationary.
- The distribution of the calls for the queues is different, but they follow the same pattern within there category over the period of 3 years.
- Different time-series forecasting models will be used to predict the call volumes and staffing.

Analysis from Mock-up Solution:

- Predicted values from Simple Exponential Smoothing are weighted sum of past observations and does not account for seasonality and trend.
- ARIMA model is only good for predicting short-term forecasting, like a weekly forecast.
- Outlier did not fit the pattern, so need to build new ARIMA model on the outliers.
- ARIMA is good for seasonality trend and missing values.
- Erlang A is more accurate than Erlang C as it accounts for the abandonment rate.

Introduction to Business Challenge

PECO is an electric and natural gas utility subsidiary of Exelon Corporation who based on Philadelphia and employs about 2,600 employees in the region. Every year, the company serves 1.6 million electric customers in Southeastern Pennsylvania and over 500,000 natural gas customers in Southeastern Pennsylvania (excluding the city of Philadelphia). PECO operates and maintains a network with 550 electric substations, 21,000 miles of distribution and transmission lines, 29 natural gas gate stations and 6,600 miles of underground gas mains.

Call center plays an important role in providing customer support and assistant at PECO. The center currently supports customers via calls, emails, faxes and in person. In 2018, the call center received 6.7 million of calls and managed to answer 88.0% in 30 seconds with capacity of 190 agents. Customer inquiries are usually divided into four different lines: Commercial, Emergency, Residential with billing and administrative questions and Residential with transfer questions.

As human resource costs account for 60%-70% of operating expenses in most call centers, PECO is searching for better solutions to forecast number of staff needed monthly using both analytical approaches and simulation models. The Drexel Lebow team investigated the call volume, skill performance and shrink data sets provided by PECO from Jan 2017 to Sep 2019 to answer the following questions:

- How many staff members are needed at the call center to maintain an acceptable level of service at minimum cost? (80% of calls handled within 30 seconds)
- How can the model adjust to changes happening in real-time business conditions?

Using personal knowledge and insights provided by PECO, the team developed a number of potential hypothesis to be investigated as follow:

- The volume of calls varies depending on specific weather conditions/unexpected events
- Call volumes are associated with seasonality and trend
- Staff productivity differs according to different skills in different seasons
- Number of required staff depends on events other than call volumes such as staff's expertise and company's hiring policies.

Literature/Industry Review

We reviewed several academic papers that related to call center analysis and modeling, service operation management, and call volume forecasting. We found that there are three important characteristics of call arrival process. First characteristic is time-variability. Some of the researchers mentioned that call arrival rates are very temporally over the day. There is significant dependency between arrival counts on successive day and there is a strong correlation in the successive period. For example, peak hour arrival rate can be significantly higher than the level of the average daily arrival rate (Brown et al., 2005). Second characteristic is inter-day correlation. There is significant dependency between arrival counts on successive days. Third is intra-day correlation. Successive periods within the same day exhibit strong correlations.

For the forecasting methods in academic papers, ARIMA model is one of the most widely used model. Many result used ARMA model in the early forecasting studies and some of them used transform function to help predict outliners; add exogenous variables for tacking the calendar effect (Aldor-Noiman et al., 2009). One of the examples is the FedEx case. In the case Weidong Xu used a combination of Exponential Smoothing, ARIMA, Linear Regression and Time Series Decomposition to develop the forecasts model (Xu, 2000). Since many empirical studies found several characteristics of the calls arrival process which we concluded in the last paragraph, the arrival process of calls follows Poisson distributed. The important characteristics enabled researchers to use Bayesian technique to forecast call volume. In these years, machine learning become more and more popular. For example, Setzer et al. employed an artificial neural network to forecast the emergency medical service demand volumes of specific areas during different time of the day (Setzler et al., 2009).

Queueing theory is to predict queue lengths and waiting times and Agner Krarup Erlang is the pioneer of queueing theory. The basic idea of Erlang is “First come, first serve” and the simplest and the most popular model is Erlang C model. The assumption of Erlang C model is calls arrival process is Poisson distributed and the calls are served by a defined number of agents which follows an exponential distribution. Also, Erlang C model ignores busy signals, customer impatience, and services that span multiple visits (Gans et al., 2003). However, Erlang C model is not easy to obtain insights from its answer and it can be inaccurate since some situations violate the assumptions (Gans et al., 2003). Erlang A model and Square-Root Safety Staffing are two improvement methods. Erlang A model is an extension of Erlang C model and it accommodate abandonment, and in the assumption of Erlang A model, customer patience time is exponentially distributed. Square-Root Safety Staffing, also known as Quality and Efficiency Driven (QED) regime, it is an asymptotically optimal of both the calls arrival rate and the number of agents. QED regime requires a balance between service quality.

Case Study

The case study is available as Reference 7

Scope of study

The purpose of the research paper is to evaluate univariate time series methods for forecasting intraday arrivals for lead times from one half-hour ahead to two weeks ahead. First, the research discusses characteristic of each method and then compares the performance of these forecasting methods on the data set. Finally, the recommendation is drawn from comparing performance of these models.

Data of the case

The data was collected from five series of half-hourly arrivals at call centers operated by a retail bank in UK for the 36-week period from January 2004 to 10 September 2004. The motivation for us to choose this case study is because of the similar patterns between our data and the data used in the paper.

- The data shows no obvious trend but very clear seasonality
- The data has a repeating intra-week cycle when the call volume generally peaks on Monday and is clearly much lower on Sundays
- The intraday cycle from data in the research paper is quite similar to PECO data: there is a peak around 11 am and then followed by a second peak around 2 pm (in PECO data the second peak is at 3 pm)

Forecasting Methods

The research considers these 5 methods to forecast the call volume: Seasonal ARMA modeling, periodic AR modeling; moving average modeling; average smoothing; an extension of Holt-Winters exponential smoothing for the case of two seasonal cycles; robust exponential smoothing based on exponentially weighted least absolute deviations regression; and dynamic harmonic regression, which is a form of unobserved component state space modeling.

Research Findings

The results indicate strong potential for the use of seasonal ARMA modeling and the extension of Holt-Winters for predicting up to about two to three days ahead and that, for longer lead times, a simplistic historical average is difficult to beat. The research also finds a similar ranking of methods for call center data from an Israeli bank which make their finding even more convincing.

DREXEL UNIVERSITY LeBow College of Business

Data Review

IVR and Call data

The IVR and call dataset contains the total number of IVR calls received (IVR Calls), calls offered (Offered), calls handled (Handled), calls answered under 30 seconds (AnsInSvcl) and service level covering the time period from 1st January 2017 to 15th September 2019. There are total 988 observations and 6 attributes. There is an unusual observation on 15th June 2019, i.e. all the columns have value '0' for this observation. All the variables have numeric datatype.

Skills Performance Data

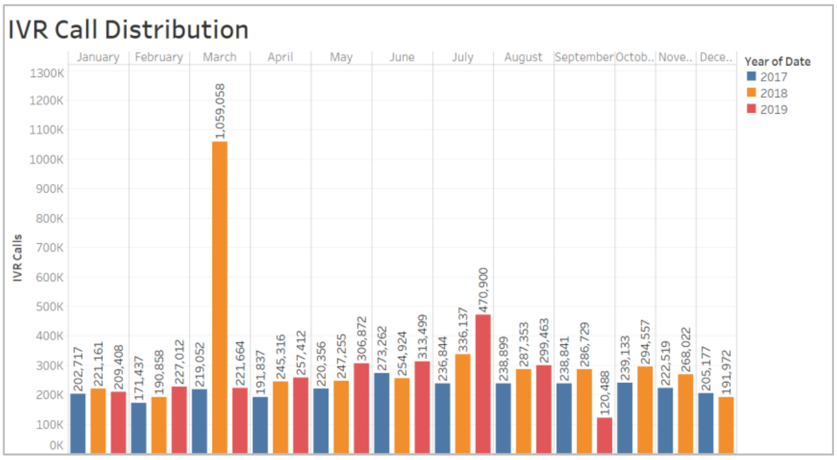
The skill performance dataset contains calls covering time period from 1st January 2017 to 15th September 2019 for every 30-minute interval. The dataset mentions calls offered, handled, abandoned, answered within target time (AnsInSvcl), skill of the agent based on training, Average length of time it takes for an agent to handle a customer inquiry (AHT), Average length of time an agent speaks with a customer (Talk), Average length of time an agent has customers on hold (Hold), Average length of time an agent does after call work (Wrap), percentage of calls answered within target (Service level), Average speed of answer and percentage of calls abandoned (Abandoned Rate). The dataset also states what type of calls were received (Queue) such as emergency calls/transfer calls/business commercial calls/residential calls. Emergency calls are offered 24 hours a day, 7 days a week. There are total 185556 observations and 16 attributes. Variables except Skills and Queue have numeric datatype, whereas the former variables have character datatype.

Shrink Data

The shrink data states the number of hours that were taken off by the agents daily with respect to Activity Category and Activity Code. The time period for this dataset ranges from 1st January 2017 to 14th September 2019. There are total 21534 observations and 4 attributes. Variables except Activity Category and Activity Code have numeric datatype, whereas the former variables have character datatype.

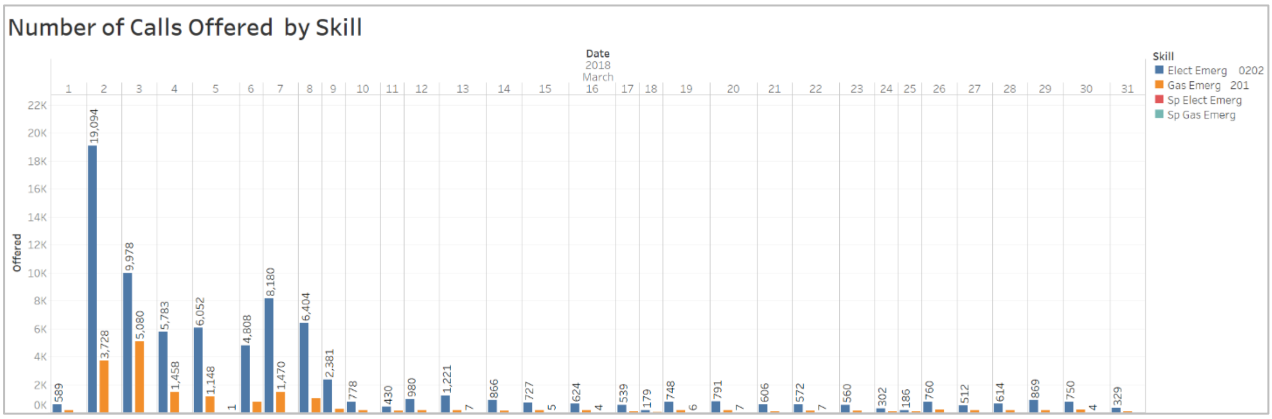
After the analysis we found that the total number of calls for all categories in IVR and Call Data does not match when the number of daily calls with 30-minute intervals are added from the Skill Performance Data. This is because the IVR and Calls data does not include transfer calls whereas Skills performance data contains observation for transfer calls. We decided to do our analysis and forecasting based on the Skill Performance Data as it will be more accurate because we have to forecast the call volume for different queues.

Data Analysis



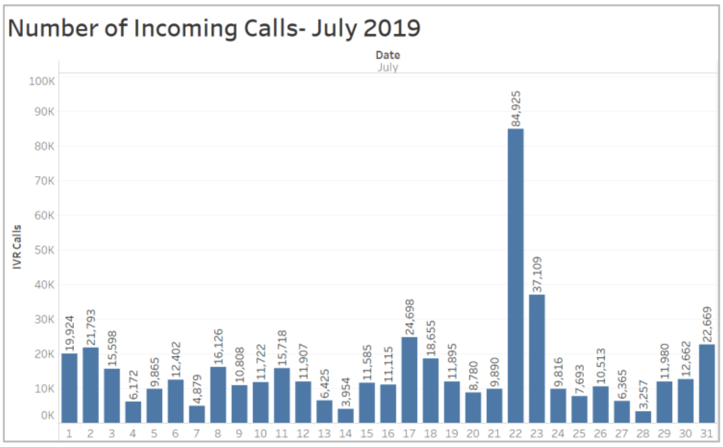
Graph 1: IVR Call Distribution

From the graph, the data is quite stationary except March 2018 and July 2019 have more calls as compared to other months. We wanted to see if there was any seasonal effect in the data, therefore analyzed the unusual months by themselves.

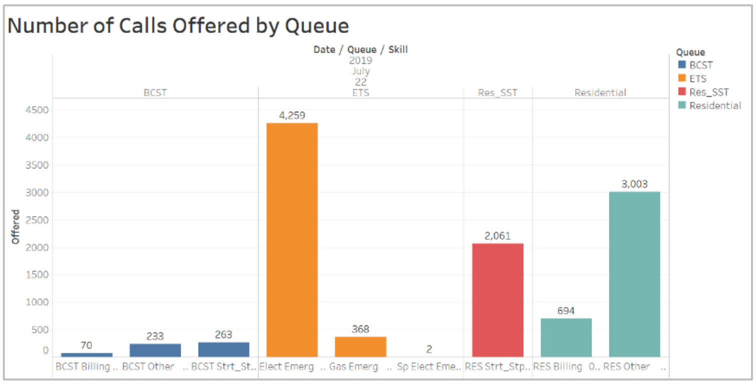


Graph 2: Number of Calls Offered by Skill in March 2018

In March 2018, week 1 which is March 2nd to March 8th has high number of emergency calls offered as compared to other weeks. After looking into this unusual week, we found that in week 1 had snowstorms and there were power outages.

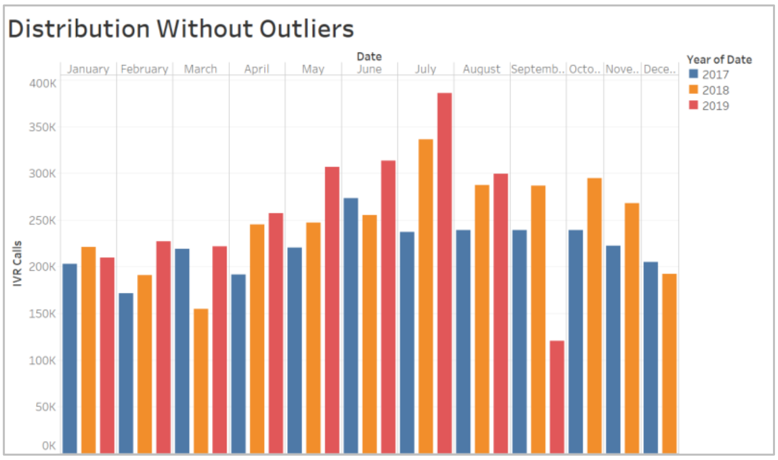


Graph 3: Number of IVR Calls for July 2019



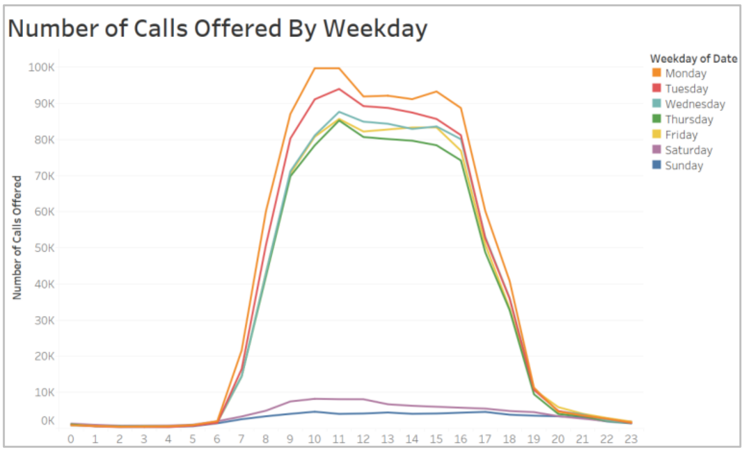
Graph 4: Number of Calls Offered by Queue

After analyzing the number of calls for the month of July 2019 (Graph 3), it was seen that July 22nd received extremely high calls. Looking further into it by queue (Graph 4), it was found that a lot of emergency calls were received. When explored, the team found that July 22nd had bad weather of heavy rains and strong winds, which lead to extremely high IVR calls.



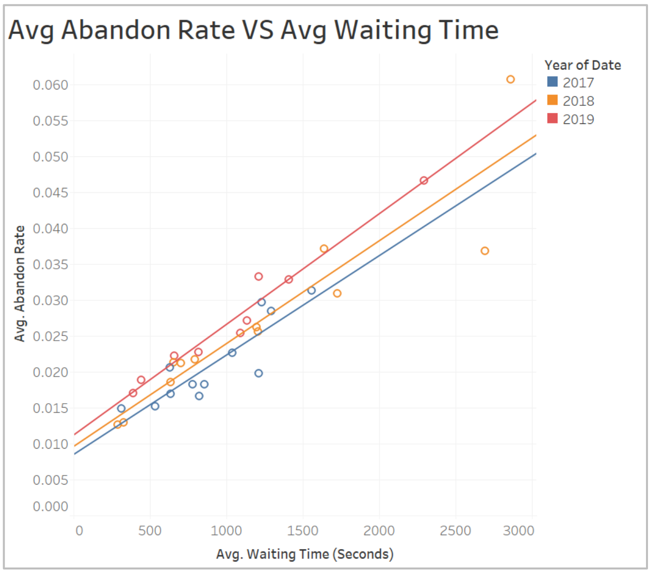
Graph 5: Distribution Without Outliers

We do not want the outliers (July 22nd and Week 1 of March 2018) to disrupt the forecasting numbers, so we decided to remove the outliers and after removing the outliers, the data looks stationary all over the years.



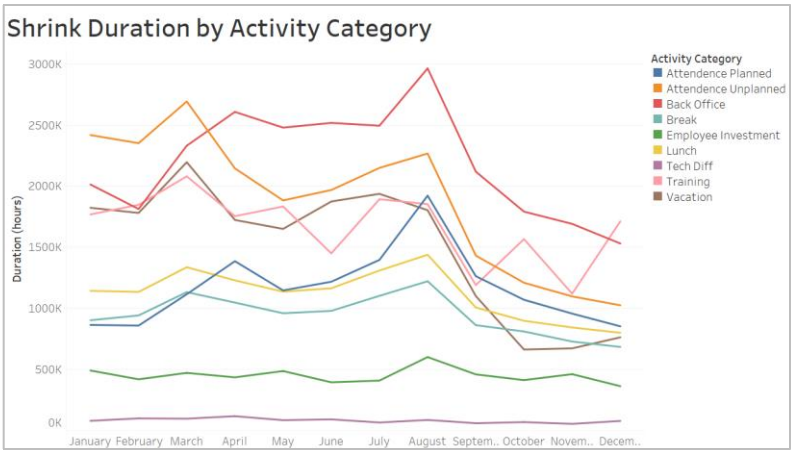
Graph 6: Number of Calls Offered By Weekday

From the graph it is visible that Monday to Friday, the call distribution follows the same pattern and, on the weekends, i.e. Saturday and Sunday they follow same calls offered distribution pattern. The number of calls offered are highest on the Monday and then decrease till Thursday, but on Friday the calls offered increase again.



Graph 7: Average Abandon Rate vs. Average Waiting Time

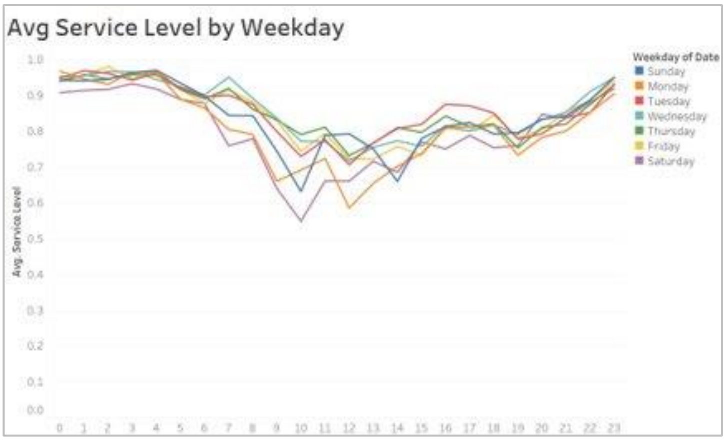
Next, we analyzed abandon rate with waiting time. We found that there is a positive correlation in both variables. As the average waiting time increase, customers tend to abandon the calls more.



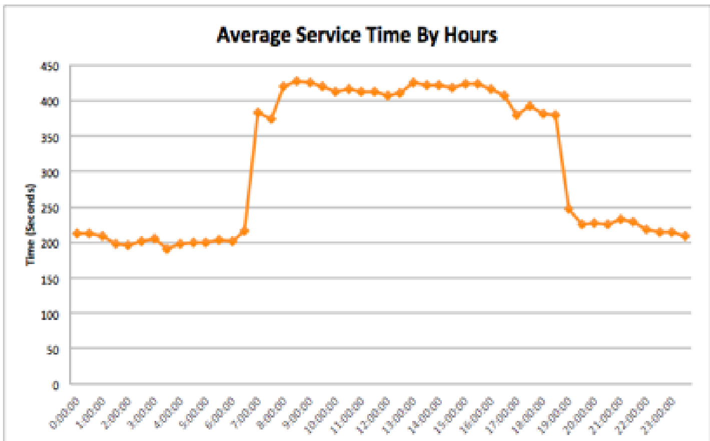
Graph 8: Shrink Duration by Activity Category

Activity Categories of Attendance Unplanned, attendance planned, back office, break, lunch, vacation and training follow the same pattern of shrink duration over the months. Employment investment and tech diff are the activity categories where agents take less time off.

Service Time and Service Level Data Analysis



Graph 9: Average Service Level by Weekday

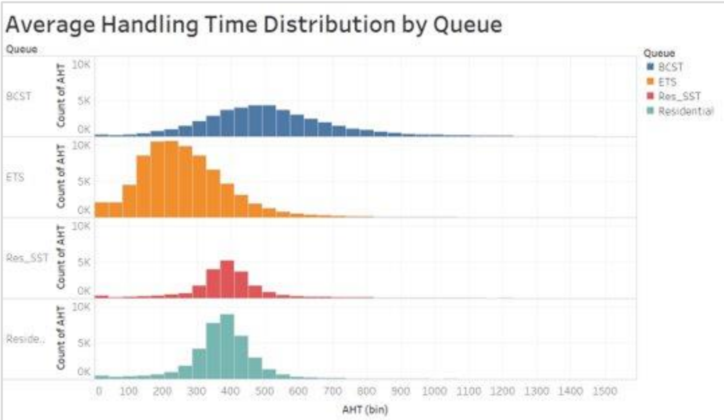


Graph 10: Average Service Time by Hour

Service quality is a complex and important topic that is closely related to the understanding of agents and customer's behavior, and we briefly review the service level during 24 hours in different days of the week and average service time during the day.

As can be seen from Graph 9, different weekdays demonstrate different daily pattern of average service level. The average service level reaches the bottom point at 10 am on Saturday, followed by 12 am on Monday. The peak time for service level to drop is from 10 to 11 in the morning and from 2 to 3 in the afternoon.

As can be seen from the Graph 10 the average service time patterns resemble the step function with the mean service time around 200 seconds during night-time and the mean service time at almost double duration (400 seconds) from the morning to the night.



Graph 11: Average Handling Time Distribution By Queue

In the Graph 11, we can see the distribution of average handling time by queue. Every queue has long tail, which means that every queue has many outliners. BCST has the highest mean among four queues, whereas ETS has the smallest means. BCST queue often have longer average handling time but the range of the time also big. For emergency call, people may prefer to talk fast, and agents have higher efficiency on working the cases. The distribution of Residential and transfer are similar, which means that average handling time of these two queues are close to each other, but transfer queue has larger standard deviation that the call length of transfer queue has wider range than residential.

Forecasting Models

Call Volumes Forecasting

To forecast the call volumes, we divided the skill performance dataset into training and testing data. To test the model, we are going to use 2017 and 2018 data as training data to predict 2019 data (testing data). The process is to defined patterns during day-of-week, week-of-month and month-of-year. Then, call volume will be forecasted based on different queues. The training data will be fitted to different models, and then models will be evaluated using Root mean squared error (RMSE) and error rate. Finally, the models with be modified and optimized.

Call Volume Forecasting Overview

Table 1 (Call volume forecast by PECO) and Table 2 (Call volume forecast by Group 4) compare the error rate for the 2019 forecasted call volume by months, starting from January till August. We also calculated the total average error rate and also average error rate by different queues. Group 4 was able to decrease the error rate for Residential calls by 7.58%, Business Commercial calls by 5.79% and Emergency calls by 6.31%. The average total error rate was also decreased by 0.19%.

	Error Rate for 2019 Forecasted Call Volume (PECO)				
	Total	Residential	Transfer	BCST	ETS
Jan	11.46%	20.73%	1.91%	11.71%	0.07%
Feb	5.58%	17.98%	7.72%	15.62%	28.46%
March	4.27%	10.74%	3.00%	7.49%	14.83%
April	1.27%	6.11%	0.63%	10.02%	29.87%
May	6.58%	24.48%	2.45%	5.31%	24.95%
June	8.47%	20.27%	4.54%	11.07%	14.52%
July	0.06%	12.90%	3.23%	14.71%	29.45%
August	0.47%	0.33%	1.21%	2.19%	2.46%
Average	4.77%	14.19%	3.08%	9.77%	18.08%

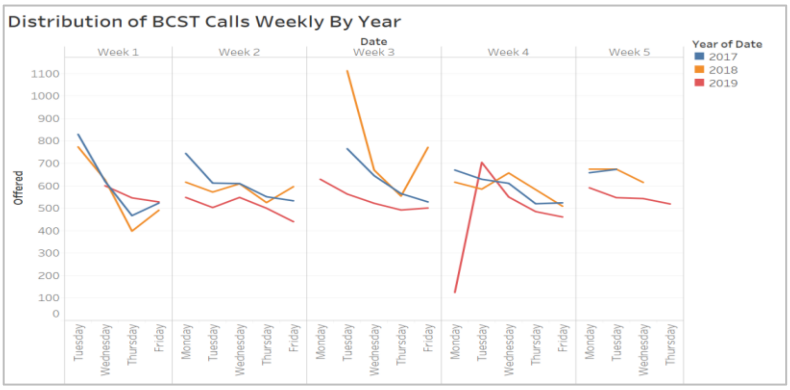
Table 1: 2019 Call volume forecast by PECO

	Error Rate for 2019 Forecasted Call Volume (Group 4)				
	Total	Residential	Transfer	BCST	ETS
Jan	5.53%	13.58%	2.10%	9.07%	6.99%
Feb	0.70%	8.55%	2.17%	9.56%	21.99%
March	2.40%	0.51%	6.78%	2.37%	7.03%
April	3.53%	2.57%	4.70%	0.96%	5.81%
May	6.24%	6.92%	5.98%	2.04%	6.84%
June	13.19%	10.20%	23.65%	4.81%	8.33%
July	4.18%	5.21%	0.64%	2.74%	27.38%
August	3.68%	5.35%	11.58%	0.30%	9.83%
Average	4.93%	6.61%	7.20%	3.98%	11.77%

Table 2: 2019 Call volume forecast by Group 4

BCST Calls

To forecast the call volume of commercial line, we first analyzed the distribution of the calls offered by year. Number of calls offered through Commercial line (BCST) (Graph 12) follow the same pattern over the year except some outliers. Number of calls offered are highest on Monday which decreases till Wednesday, and then increases as the weekend approaches.



Graph 12: Distribution of BCST Calls Weekly

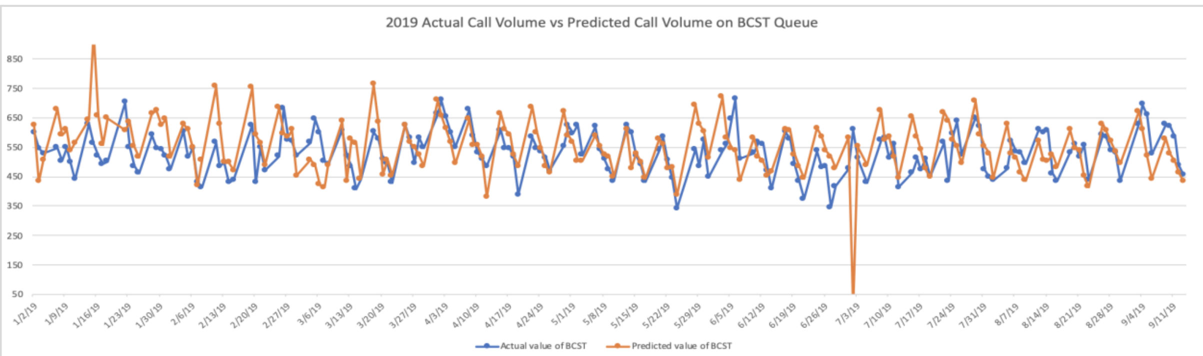
Further analyzing the distribution for BCST Calls for 2017 (Graph 13), 2018 (Graph 14) and 2019 (Graph 15), distribution for all the three years is stationary, except 2019 which has some outliers. Since different days of a week follow the same pattern and there are "weekly" cycles in the Graph 11, Graph 12 and Graph 13. We will use average smoothing forecasting model to predict days of the week in the year.

Mean Moving Average

Mean Moving Average time-series forecasting method used for data. In this method the predicted values are average of past observations. It does not need to choose a smoothing factor. This method gives the fact that "what has happened before will be done again".

We built up the mean moving average model by calculating the days of the week. For instance, the average of the Monday on the first week in 2017 and 2018 was used to predict the value of Monday in the first week of 2019. Prior to forecasting the call volume, outliers such as weekends and federal holidays were removed. When there is a holiday on the weekday, we just put zero as the predicted value. If the week has holiday such as July 4th, we manually inputted the holiday data into a day before.

Graph 13 is the actual value versus predicted value and the model performance is in Table 3. The week on week and day on day method provides a good result, the error rate is 3.98% and the standard deviation is 267. Even though the RMSE is larger than PECO's forecasting, we believe that our model still can have a great performance on the call volume forecasting not only in short term but also in a long term. The model is in the appendix.



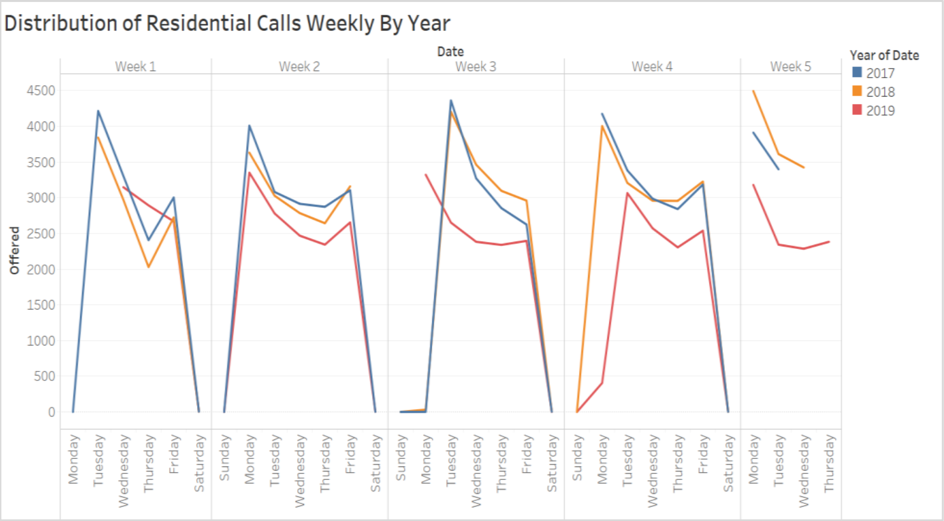
Graph 13: Actual values and predicted values of BCST Calls 2019

BCST QUEUE			
	Actual	PECO Forecast	Group 4 Forecast
RMSE	-	75	94
Average % difference	-	9.77%	3.98%
Mean	369	402	380
S.D	252	274	267

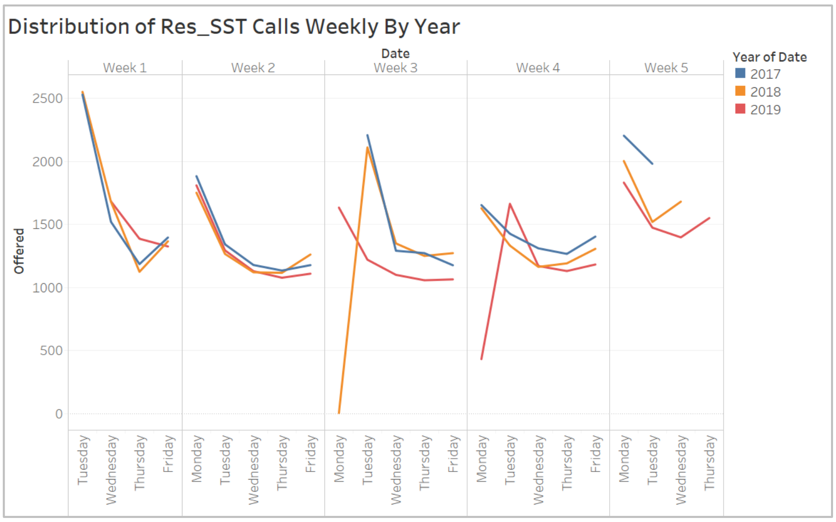
Table 3: BCST Calls Forecasting model result

Residential and Transfer Calls

To forecast the call volume for Residential and Transfer line, first we analyzed the distribution over the three years to see patterns. From graph 14 and graph 15, calls offered follow the same pattern over all the three years. There are slight ups and downs in the distribution, which is because of the outliers and the possibility that PECO acquires new customers every year.



Graph 14: Distribution of Residential Calls Weekly



Graph 15: Distribution of Transfer Calls Weekly

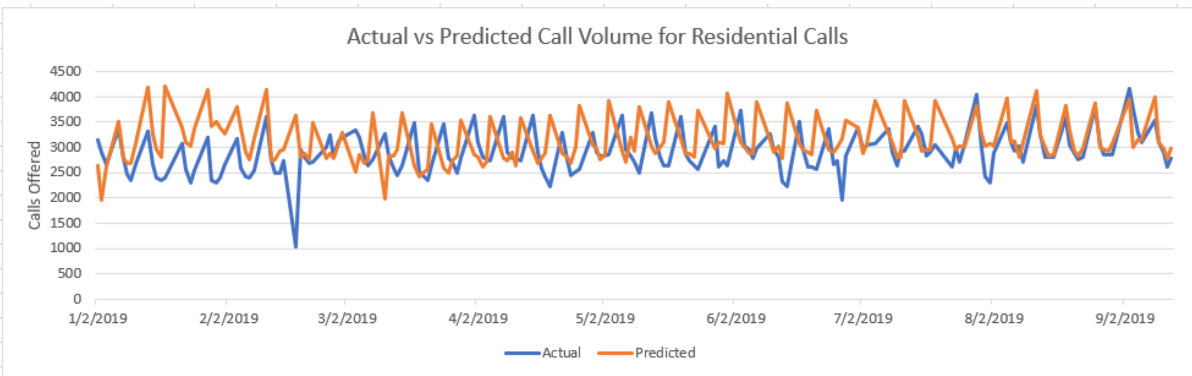
Further analyzing the distribution for Residential calls and Transfer for 2017, 2018 and 2019 (in Appendix), distribution for all the three years is stationary. In 2017, there is a slight decrease in the trend of the calls offered, and same for 2018, there is slight decrease in the trend and for 2019 there is increase in the trend of the calls offered. The analysis for the Residential calls is done yearly, but over the three years, the calls follow the same pattern weekly. For instance, week 20 for 2017, 2018 and 2019 have the same pattern for the calls offered.

Holt Winters Forecasting Model

From the distribution of the calls offered in 2017, 2018 and 2019 for both residential and transfer calls, we analyzed that there are trends that are changing over the three-year period. The calls also show seasonality variations. Considering the factors that are visible, we chose Triple Exponential method to forecast the call volumes for 2019. Triple Exponential method is also known as the Holt-Winters Forecasting model. This model accounts for level, trend and seasonality factors. There are smoothing parameters for each of the factor: Alpha-smoothing parameter for LEVEL, Beta- smoothing parameter for TREND and Gamma- smoothing parameter for SEASONALITY.

The values for these smoothing parameters should be selected in way that minimizes the RMSE. The forecasting model for residential calls, focuses on the weekly distribution and forecasting model for transfer calls focuses on the weekday distribution over the three-year period. Calls offered in 2017 were used as the initial values to calculate the seasonality. Starting from Jan 1, 2018 the calls were forecasted till September 15, 2019. To forecast the call volume, following method was used for residential calls. The seasonality value for the weekday of week 1 of 2017, values of level and trend for last weekday of week 52 of 2017 was used to forecast the call volume for a weekday in week 1 of 2018. In other words, the day of the week to be forecasted uses previous years' same weeks' seasonality and the values of trend and level from the previous day from the day to be forecasted. For transfer calls, same method was followed, except the calls were forecasted for individual weekdays.

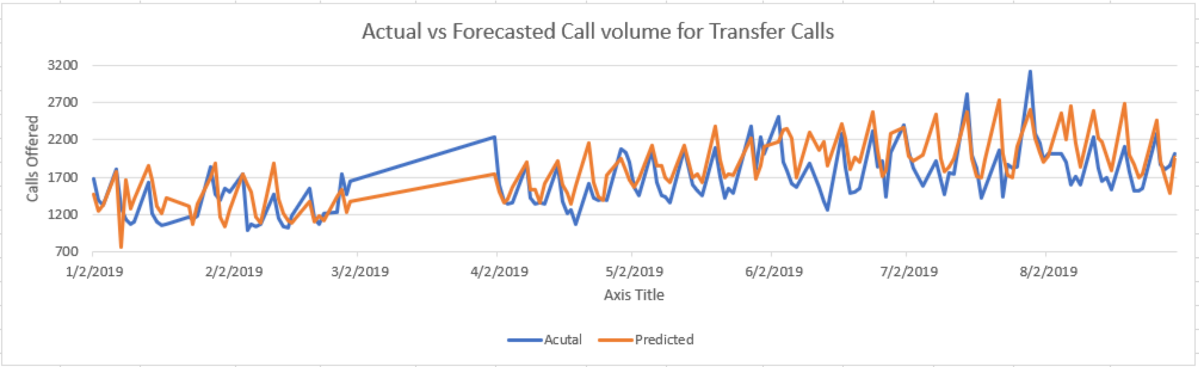
After the call volume was forecasted, the RMSE was calculated for 2018 and 2019. The RMSE was minimized using the solver and the values for alpha, beta and gamma were chosen by solver. While forecasting the values for 2019, outliers such as Saturdays, Sundays, federal holidays such as Martin Luther King's Day, Independence Day, Thanksgiving days, were excluded from the model. The model for forecasted call volume for both Residential and Transfer calls is in appendix.



Graph 16: Actual vs Predicted Call Volume for Residential Calls

Residential			
	Actual	PECO Forecast	Group 4 Forecast
RMSE	-	462	388
Average % difference	-	14.19%	6.61%
Mean	2010	2271	2105
S.D	1359	1552	1455

Table 4: Residential Calls Forecasting model result

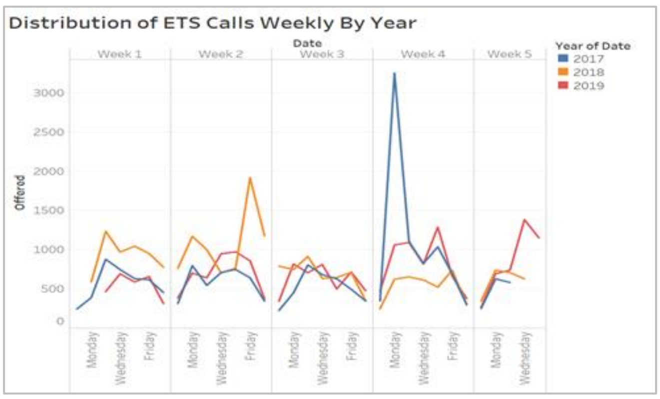


Graph 17: Actual vs Predicted Call Volume for Transfer Calls

TRANSFER QUEUE			
	Actual	PECO Forecast	Group 4 Forecast
RMSE	-	203	321
Average % difference	-	3.08%	7.20%
Mean	1134	1156	1251
S.D	791	829	913

Table 5: Transfer Calls Forecasting model result

Emergency Calls



Graph 18: Distribution of ETS Calls Weekly By Year

As shown in Graph 18, Emergency calls doesn't have strong pattern. So, we decided use different models for emergency and non-emergency calls. There are three things shown in the chart. First, Monday or Tuesday are likely to have more calls, but it is not as clear as non-emergency. We can only guess what happened. Emergency is for 7/24, so they don't have to wait until Monday to call. This might because people leave their home during the weekends and found emergency when they go back. Weekday might be a factor. Secondly, we found that the call volume goes up a little bit as time goes by. And third, after we check the sudden acceleration, it is mostly because of the weather: snow or storm. So, we would like to collect more weather data to forecast the ETS calls use Linear Regression. ETS including emergency for gas and electric. Usually we use gas for heating and electric for cooling so they might have different mode.

We collected those data for 2018 from "weather underground" website. The variables include date, weekday, Maximum, Minimum, Average Temperature, Wind Speed, Pressure, Humidity, Dew points and Precipitation. Combining with real life, the temperature or pressure change might be related -- people would start using the facility and find an emergency when the weather change. We calculated the daily change of weather as well. The wind speed or other factor might not cause a problem until threshold. For example, wind speed higher than 24mph is strong wind and would have whistling heard in telegraph wires. We put some 0-1 variable into consider.

For the data processing, we removed 3/2/2018-3/8/2018 data. Those 7 days have high call volume because an unexpected storm, we consider it as outliner and remove it to avid outliner. Then we build the liner regression model and use lasso to the feature selection. (More detail can be found in appendix)

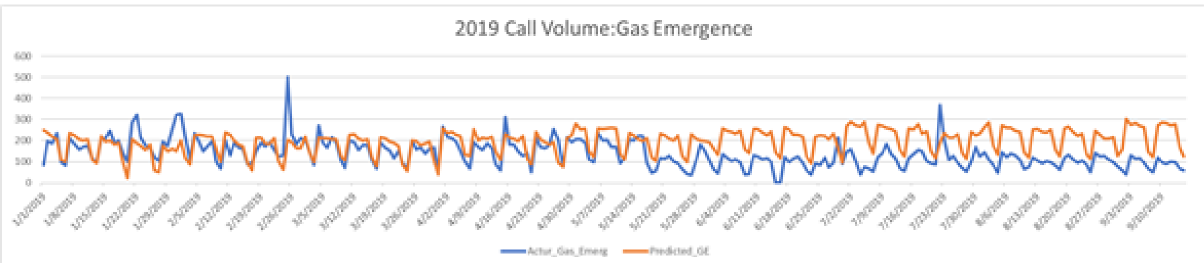
	(Intercept)	sun	sat	Avg WindSpeed	mon	MaxTemp	Date	tue	month	Avg Humidity	MinWindSpeed _change	MinTemp	MW
Estimate	-18.4581	-281.341	-251.195	25.4627	104.3285	4.0617	-3.4606	76.7772	-7.9757	2.557	-9.7945	1.9698	296.0528

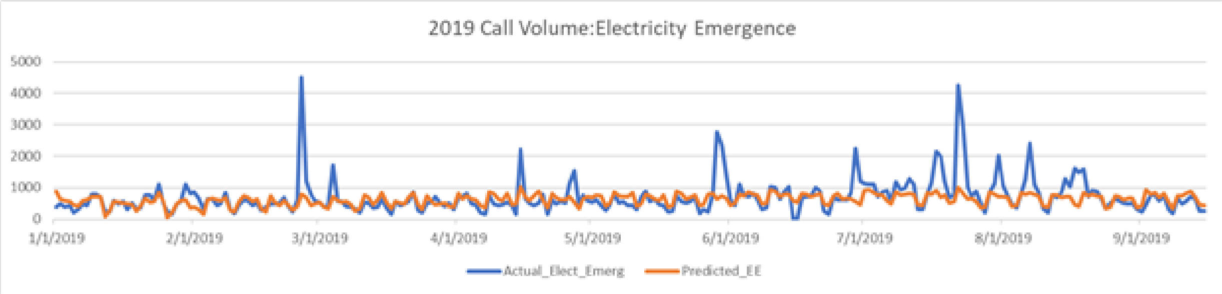
Table 6: Variables and Coefficient for Electricity Emergency

Gas	(Intercept)	Sun	Sat	Min DewPoint	Date	month	Avg DewPoint	Mon	AvgHumidity _change	Max Pressure	Tue	MaxWind Speed	MaxTemp_ change	AvgPressure _change	Min Pressure	Thur	MW
Estimate	-2165.00	-120.50	-92.22	1.35	-1.81	1.35	-0.08	15.15	-0.26	46.34	8.94	1.27	-0.47	-15.47	31.48	-7.53	38.56

Table 7: Variables and Coefficient for Gas Emergency

To forecast the emergency call volume, we will need Date, Week, and weather forecast data. Weather forecasting data includes: Average Wind Speed, Maximum wind speed, Minimum Wind Speed, Maximum Temperature, Minimum Temperature, Average Dew Point, Minimum Dew Point, Average Pressure, Maximum Pressure, Minimum Pressure, Average Humidity.





Graph 19: Call Volume predicted VS actual

This is the result compare between actual and predicted, the 2019 gas emergence have a slightly downswing and the model failed to predict. The high call volume has some error but overall the result is better than PECO's.

EMERGENCY QUEUE			
	Actual	**PECO Forecast**	**Group 4 Forecast**
RMSE	-	562	522
Average % difference	-	18.08%	11.77%
Mean	841	702	807
S.D	581	285	232

Table 8: Emergency Calls Forecasting model result

Service Level and Service Time Forecasting

Service Time

Average Handling Time (AHT) also has trend within three years. There was a slight increase in AHT at the end of 2017, followed by an increase in AHT in January 2018. From this observation, we conclude that AHT was also affected by time factors. By visualizing AHT by year and queues, we can see that there are many outliers in AHT data. Most outliers are from BSCT Queue, resulting in large variation in AHT distribution for this queue. The percentage that each queue accounted for also varies from time to time which is also critical for us to know when building forecasting models.

We use AHT data from 2017 and 2018 as training data and 2019 as testing data. As we learn from our literature review and data analysis results, AHT differs during different days of week and different months of year. Therefore, we create new variable Weekday indicating specific date of the week. We build two linear regression models which include different number of variables to predict AHT and compare results between them. In the first linear regression model we use all variables that can help to explain AHT such as: Queue, Year, Month, Weekday and Time. Moreover, we all have assumption that different weekday and queue as well as different month of year will have specific effect on AHT, we also include these factors in our first predicting model. Model 2 is the generalized regression version of model 1, which does not include the inter-day effect as well as effect between queue and month. Results show that Model 1 can estimate the relationship between AHT and other variables better than Model 2. However, this model has over-fitting problem which can perform well on 2017 and 2018 data but not on 2019 data. In contract, model 1 can perform better when predicting AHT om 2019 data set. The table

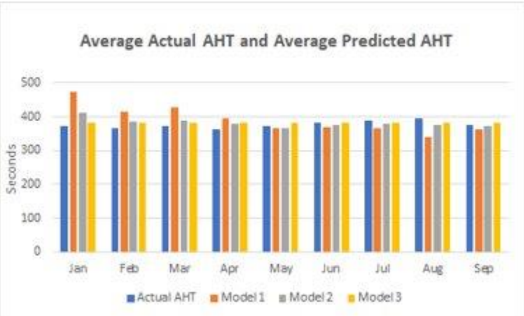
9 shows our forecasting values for different queues and actual value of AHT 2019. As we can see from the table, predicted AHT values from Model 2 are closer to actual AHT than Model 1.

Unit: Seconds

Month	BCST			ETS			Res_SST			Residential		
	Actual	Model 1	Model 2	Actual	Model 1	Model 2	Actual	Model 1	Model 2	Actual AHT	Model 1	Model 2
Jan	512	629	567	282	360	298	392	524	462	383	483	420
Feb	511	573	545	270	303	275	395	469	441	385	428	398
Mar	540	583	547	261	312	276	393	476	442	383	437	400
Apr	519	547	530	255	281	261	376	441	425	369	402	384
May	525	521	518	272	252	249	379	416	414	369	375	372
Jun	550	525	532	287	256	262	386	418	427	363	378	385
Jul	555	525	534	303	256	265	379	394	426	375	380	387
Aug	553	497	530	312	227	261	389	369	422	370	350	384
Sep	529	522	533	299	251	262	387	412	427	357	374	387
Grand Total	**533**	**548**	**537**	**282**	**280**	**268**	**386**	**436**	**432**	**374**	**404**	**392**

Table 9: Comparison between Forecasted AHT value and Actual AHT value per Queue

The second method we use to predict AHT is simple moving average. By visualizing AHT over time, we can see the correlation for AHT between months. Therefore, we choose to build another model using time -series technique to predict AHT and compare with linear regression models. The technique is quite simple: the value of AHT for a particular month will be the average AHT of the two previous months (The full three models can be found in Appendix). Graph 19 shows the comparison on average AHT between our models and actual values. It can be seen from the Graph 19 that model 2 and model 3 can predict values closer to actual values compared to model 1. Table 10 shows the calculation of error rate for each model and PECO's forecast. Model 3 which uses simple moving average gives us the least error rate. However, our error rate is still higher than PECO's forecast which only use the same forecasted value for every month. We suggest using PECO method or our moving average model (Model 3) for long-term forecasting. However, if PECO need to forecast AHT for each individual queue in short-term period, our group suggests using Linear Regression Model (Model 2) because this model can provide different prediction on AHT for different queues per days.



Graph 20: Average Actual vs Predicted AHT

Model 1 Error	Model 2 Error	Model 3 Error	Average Prediction Error	Peco Prediction Error
27%	10%	4%	6%	2%
13%	5%	1%	3%	0%
14%	5%	0%	2%	2%
9%	4%	2%	0%	0%
1%	2%	1%	5%	1%
3%	1%	1%	4%	0%
6%	3%	0%	7%	2%
14%	5%	3%	7%	5%
4%	1%	2%	4%	0%
10.1%	**4.0%**	**1.6%**	**4.2%**	**1.2%**

Table 10: Error rate between Group 4 and PECO

Service Level

We want to estimate the relationship between service level, abandon rate and number of agents by using linear regression method. As can be interpreted from the model, Abandon Rate, Number of Agents and Number of Offered Call are important factors to predict service level. While Number of Agents is positively correlated with Service Level, Abandon Rate and Number of Offered Call are negatively correlated with Service Level, which means the higher the call volume and abandon rate, the lower service level. We also build another build another model including only Queues, Time, Number of Agent and Abandon Rate to estimate the relationship and use coefficient from these models to predict for service level. We only train and test model on the same data (2019) because we can't get staffing data for 2017 and 2018. The table 11 below shows our forecasted values and actual service level per queue. The model can predict service level for ETS queue better than other queues (Full model can be found in Appendix)

	BCST		ETS		Res_SST		Residential			
Month	**Actual**	**Forecast**	**Actual**	**Forecast**	**Actual**	**Forecast**	**Actual**	**Forecast**	**Total Actual**	**Total Forecast**
Jan	65%	51%	93%	91%	93%	84%	89%	76%	85%	78%
Feb	71%	52%	91%	90%	91%	84%	86%	76%	85%	78%
Mar	55%	51%	92%	91%	89%	83%	78%	74%	80%	76%
Apr	53%	50%	92%	91%	82%	83%	74%	74%	77%	76%
May	49%	50%	92%	90%	78%	82%	69%	73%	74%	75%
Jun	37%	46%	90%	91%	70%	82%	63%	72%	69%	74%
Jul	30%	44%	86%	89%	68%	78%	56%	69%	64%	72%
Aug	35%	46%	88%	91%	78%	78%	69%	74%	69%	75%
Sep	41%	46%	93%	92%	85%	81%	77%	75%	77%	77%
Grand Total	**49%**	**49%**	**91%**	**91%**	**81%**	**81%**	**74%**	**74%**	**76%**	**76%**

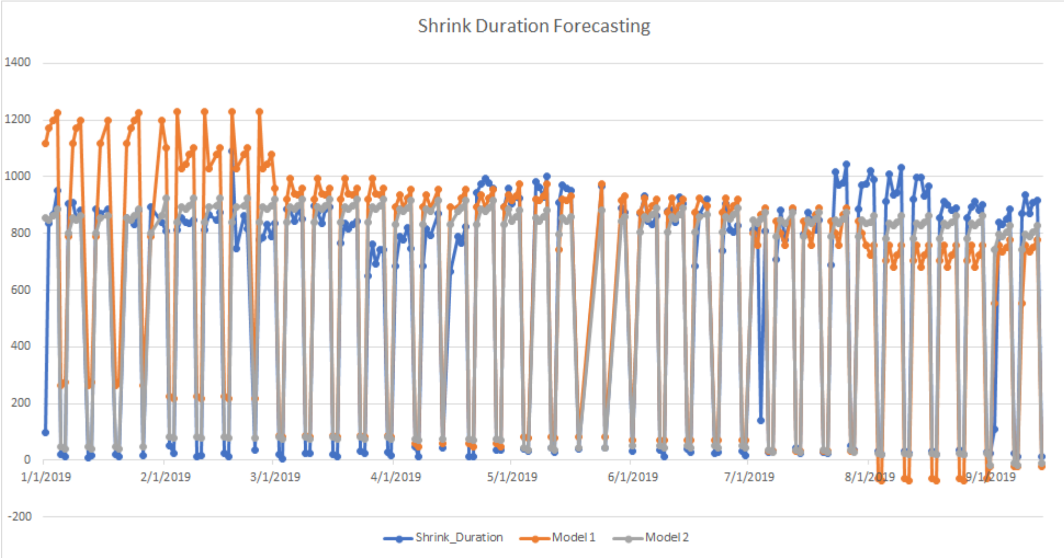
Table 11: Comparison between Forecast and Actual Service Level Per Queue

Shrink Percentage Forecasting

We use Multiple Linear Regression to predict for Shrink Duration. Model 1 includes the interaction between month for different years and weekday for different months. Model 2 is a simplified version of Model 1 and includes only Year, Month and weekday to predict shrink duration. These are listed variables for these models:

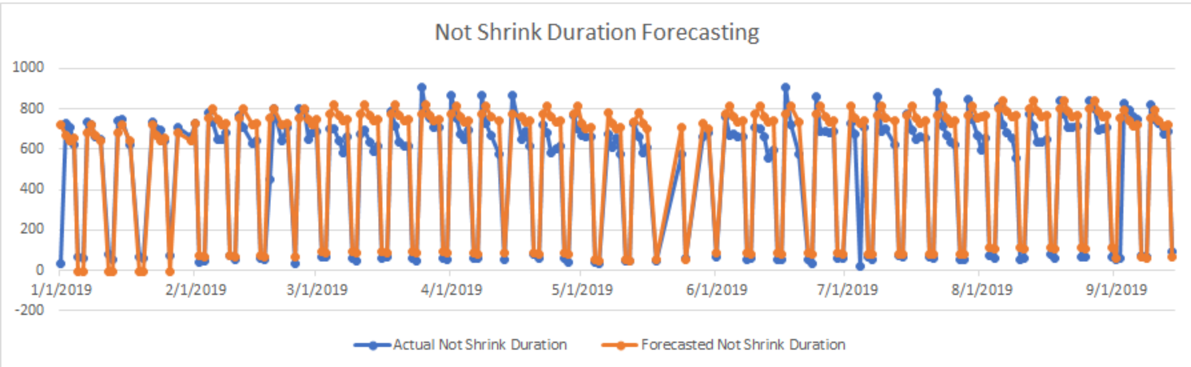
- Model 1 Variables: Year + Month + Weekday + Year*Month + Month*Weekday
- Model 2 Variables: Year + Month + Weekday

Model 2 performs better than Model 1 on both training and testing data. Both models can explain high proportion of variance in shrink duration using only 3 exploratory factors. Therefore, our group suggests using Model 2 to predict Shrink Duration because it is easy to implement on another environment such as Excel and the result is also good.



Graph 21: Shrink Duration Forecasting

For not shrink duration, we also use Multiple regression with only 3 variables: Year, Month and Weekday for forecasting. It can be seen from the graph, the model performs well and predicted values are close to actual values.



Graph 22: Non-Shrink Duration Forecasting

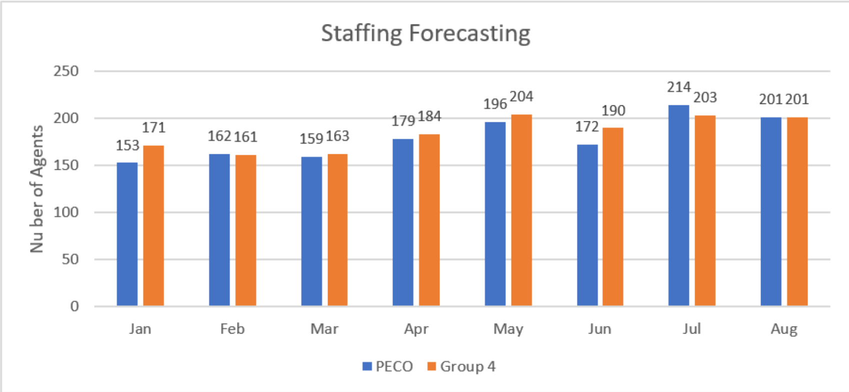
Month	Actual Percentage	Predicted Percentage	Error	Average Percentage	Error
Jan	55.1%	56.3%	1.0%	55.8%	0.7%
Feb	54.3%	54.2%	2.0%	54.1%	0.1%
Mar	53.8%	53.3%	2.7%	53.4%	0.4%
Apr	54.0%	53.2%	1.6%	53.2%	0.8%
May	59.0%	53.9%	2.2%	53.4%	5.6%
Jun	54.6%	52.0%	0.3%	52.3%	2.3%
Jul	54.8%	51.7%	0.1%	51.5%	3.2%
Aug	56.2%	50.2%	2.9%	50.2%	6.0%
Sep	53.1%	50.5%	0.7%	50.3%	2.8%
Average Error Rate			1.49%		2.44%

Table 12: Comparison between Forecast and Actual Shrink Percentage Level Per Month

Staffing Forecasting

To calculate staffing number, we are using two methods to compare between PECO forecast and our group's forecast:

Method 1: We will use PECO formula, AHT and Shrink to calculate number of FTE. However, our group will use our predicted call offered to compare between our forecasting model and PECO's forecasting model.



Graph 23: Staffing Forecasting

January, April and June have the largest difference in our group's calculation and PECO due to the biggest difference between our forecasted call volume and PECO.

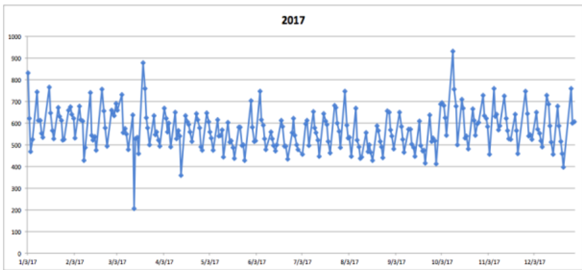
Method 2: We will change the formula using: new abandon rate, our predicted call offered, AHT and Shrink. For new calculation formula, our assumption is that each agent works 40h/week and 22 days a month. We also recalculate the abandon rate using new coefficient from regression models with service level and number of agents. Our new abandon rate is 3.3% to get service level above 80%. Therefore, we will use this abandon rate to calculate number of agents needed to maintain the required service level. The table 13 shows our forecasting for number of agents required to meet the service level of 80%. The last row explains the difference between our required staff from forecasting models and the current number of agents that PECO is having right now.

	Jan	Feb	Mar	Apr	May	Jun	Jul	Aug
Call Offered	133,040	112,483	124,285	138,103	144,653	141,625	154,601	153,043
Abandon Rate	3.3%	3.3%	3.3%	3.3%	3.3%	3.3%	3.3%	3.3%
Call Handle	128,650	108,771	120,184	133,545	139,879	136,951	149,499	147,993
AHT	382	381	382	381	381	381	381	381
Call Load (h)	13651	11512	12753	14134	14804	14494	15822	15663
Occupancy	81%	81%	81%	81%	81%	81%	81%	81%
Adjusted Call Volume	16853	14212	15744	17449	18276	17894	19533	19337
Raw number of Agent	96	81	89	99	104	102	111	110
Shrink	56.3%	54.2%	53.3%	53.9%	52.0%	51.7%	50.2%	50.5%
FTE	170	149	168	184	200	197	221	218
Prod	152	159	157	155	153	149	158	157
OT	4	4	3	7	3	3	3	3
Events		0	0	0	0	0	-3	0
Prod+OT	156	163	160	162	156	149	161	160
Net Staff	**-14**	**14**	**-8**	**-22**	**-44**	**-47**	**-60**	**-58**

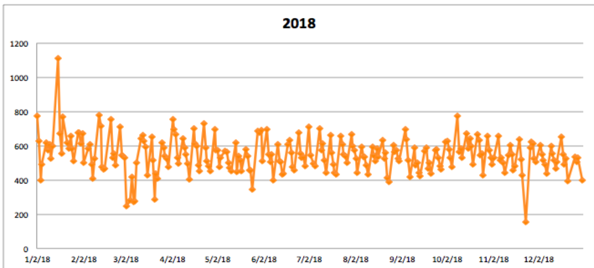
Table 13: Comparison between Forecast and Actual Shrink Percentage Level Per Month

Appendix 1

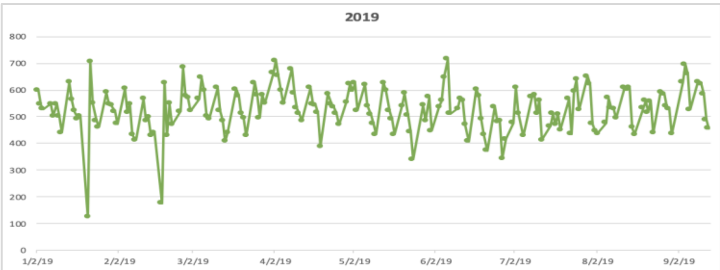
BCST CALLS DISTRIBUTION BY YEAR



Graph 21: Distribution of BCST Calls 2017

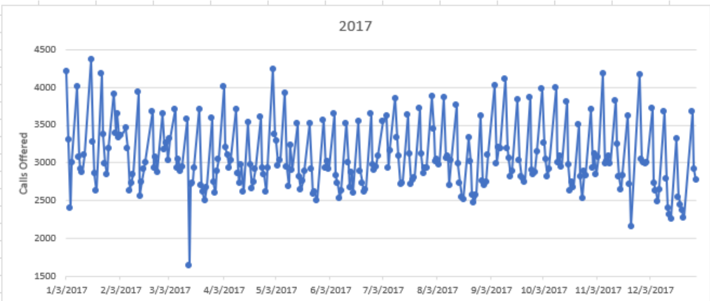


Graph 22: Distribution of BCST Calls 2018

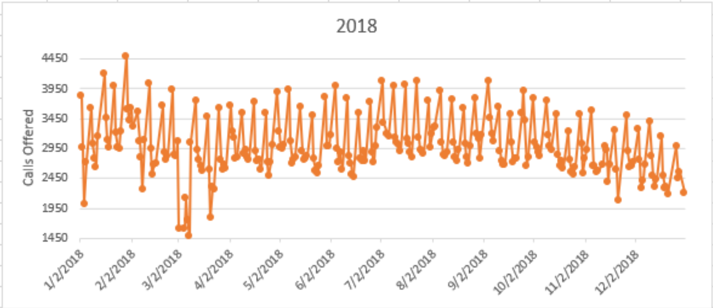


Graph 23: Distribution of BCST Calls 2019

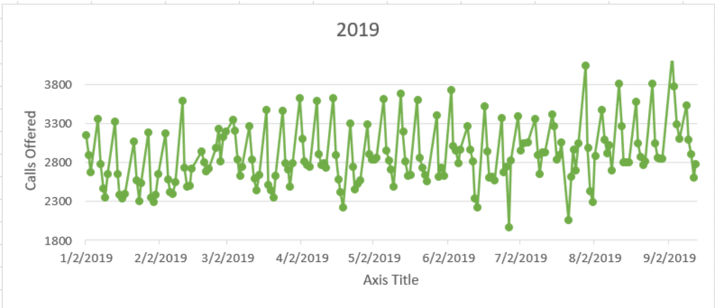
RESIDENTIAL CALLS DISTRIBUTION BY YEAR



Graph 24: Distribution for Residential Calls 2017

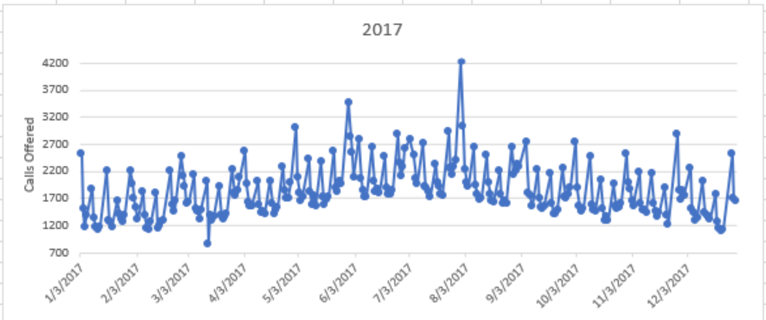


Graph 25: Distribution for Residential Calls 2018

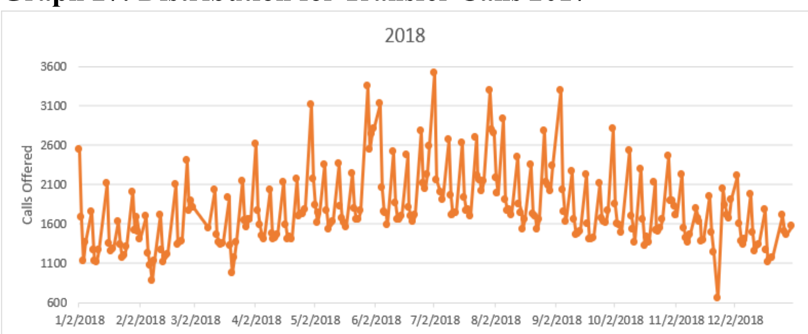


Graph 246 Distribution for Residential Calls 2019

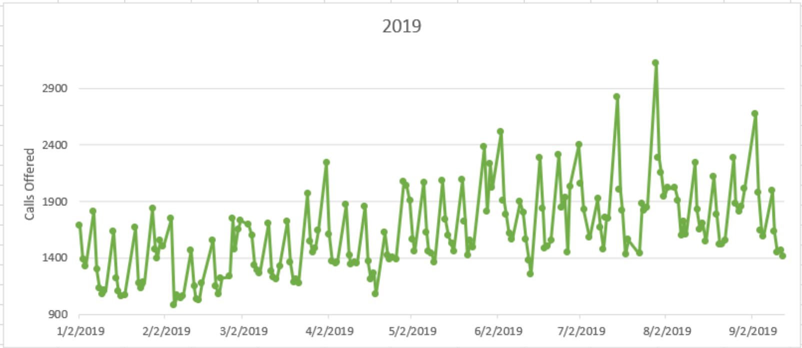
<u>**TRANSFER CALLS DISTRIBUTION BY YEAR**</u>



Graph 27: Distribution for Transfer Calls 2017



Graph 28: Distribution for Transfer Calls 2018



Graph 29: Distribution for Transfer Calls 2019

Mean Moving Average Estimate For BCST Data

Week	Weekday	DOW	Date	BCST Offered	Prediction
1	2	12	1/1/19	0	0
1	3	13	1/2/19	600	623.5
1	4	14	1/3/19	546	432.5
1	5	15	1/4/19	528	507
1	6	16	1/5/19	0	0
1	7	17	1/6/19	0	0
2	1	21	1/7/19	548	680
2	2	22	1/8/19	503	592
2	3	23	1/9/19	548	609.5
2	4	24	1/10/19	500	538
2	5	25	1/11/19	440	564.5
2	6	26	1/12/19	0	0
2	7	27	1/13/19	0	0
3	1	31	1/14/19	629	0
3	2	32	1/15/19	563	938.5
3	3	33	1/16/19	522	657.5
3	4	34	1/17/19	492	559.5
3	5	35	1/18/19	501	649.5
3	6	36	1/19/19	0	0
3	7	37	1/20/19	0	0
4	1	41	1/21/19	125	643
4	2	42	1/22/19	704	607
4	3	43	1/23/19	550	634
4	4	44	1/24/19	485	551.5
4	5	45	1/25/19	461	516.5
4	6	46	1/26/19	0	0
4	7	47	1/27/19	0	0
5	1	51	1/28/19	591	666
5	2	52	1/29/19	547	673.5
5	3	53	1/30/19	543	626.5
5	4	54	1/31/19	519	646
5	5	55	2/1/19	474	516.5
5	6	56	2/2/19	0	0

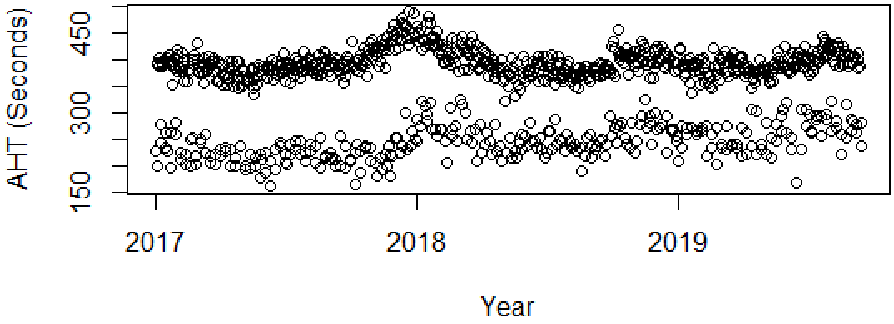
Liner Regression For Gas_Emergency

```
Call:
lm(formula = Gas_Emerg ~ Sun + Sat + MinDewPoint + Date + month +
    AvgDewPoint + Mon + AvgHumidity_change + MaxPressure + Tue +
    MaxWindSpeed + MaxTemp_change + AvgPressure_change + MinPressure +
    Thur, data = dgas)

Residuals:
    Min      1Q  Median      3Q     Max
-162.03  -40.09  -11.84   24.04  461.75

Coefficients:
                    Estimate Std. Error t value Pr(>|t|)
(Intercept)       -2001.8859   701.1526  -2.855  0.00456 **
Sun                -119.0163    12.5837  -9.458  < 2e-16 ***
Sat                 -90.5808    12.3248  -7.349 1.48e-12 ***
MinDewPoint           1.4952     1.0670   1.401  0.16205
Date                 -1.8376     0.4351  -4.224 3.08e-05 ***
month                 1.3786     1.5176   0.908  0.36431
AvgDewPoint          -0.2393     1.0698  -0.224  0.82312
Mon                  16.1243    12.3473   1.306  0.19246
AvgHumidity_change   -0.2625     0.3587  -0.732  0.46476
MaxPressure          55.1139    33.2180   1.659  0.09800 .
Tue                   8.7322    12.4333   0.702  0.48296
MaxWindSpeed          2.3408     0.9117   2.568  0.01067 *
MaxTemp_change       -0.5090     0.7004  -0.727  0.46788
AvgPressure_change  -17.3797    24.7961  -0.701  0.48384
MinPressure          16.7754    31.1494   0.539  0.59055
Thur                 -8.6163    12.4966  -0.689  0.49098
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 71.66 on 342 degrees of freedom
Multiple R-squared:  0.4144,	Adjusted R-squared:  0.3887
F-statistic: 16.13 on 15 and 342 DF,  p-value: < 2.2e-16
```

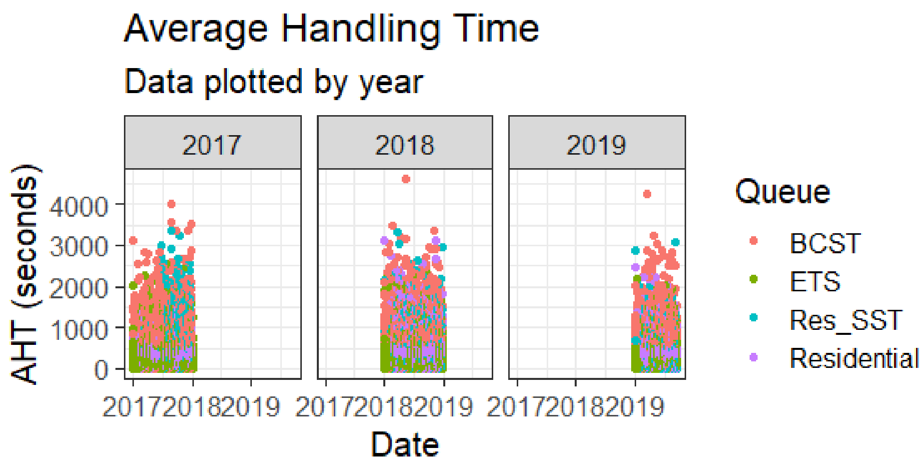
Liner Regression For Electric_Emergency

```
Call:
lm(formula = Elect_Emerg ~ sun + sat + AvgWindSpeed + mon + MaxTemp +
    Date + tue + month + AvgHumidity + MinWindSpeed_change +
    MinTemp + +MW, data = dgas)

Residuals:
    Min      1Q  Median      3Q     Max
-574.48 -131.39  -27.37  105.41 1397.65

Coefficients:
                     Estimate Std. Error t value Pr(>|t|)
(Intercept)          -18.4581   103.9749  -0.178  0.85920
sun                 -281.3409    36.7846  -7.648 2.05e-13 ***
sat                 -251.1950    36.5916  -6.865 3.10e-11 ***
AvgWindSpeed          25.4627     4.5124   5.643 3.49e-08 ***
mon                  104.3285    36.4527   2.862  0.00447 **
MaxTemp                4.0617     2.3138   1.755  0.08007 .
Date                  -3.4606     1.3645  -2.536  0.01165 *
tue                   76.7772    36.4464   2.107  0.03588 *
month                 -7.9757     3.7333  -2.136  0.03335 *
AvgHumidity            2.5570     0.8471   3.018  0.00273 **
MinWindSpeed_change   -9.7945     3.5354  -2.770  0.00590 **
MinTemp                1.9698     2.5334   0.778  0.43736
MW                   296.0528    50.9887   5.806 1.45e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 224.9 on 345 degrees of freedom
Multiple R-squared:  0.4858,	Adjusted R-squared:  0.4679
F-statistic: 27.16 on 12 and 345 DF,  p-value: < 2.2e-16
```

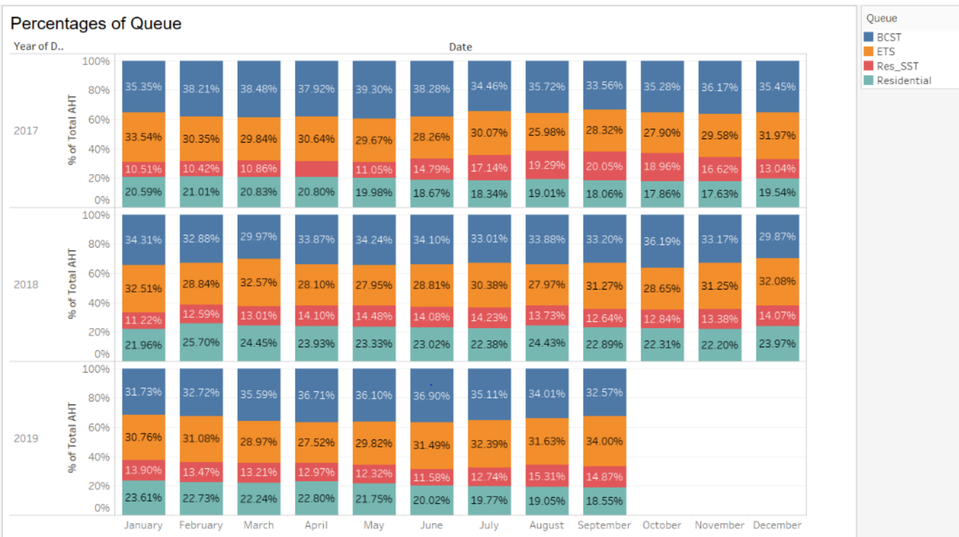
Average Handling Time Over Year



Outlier In Average Handling Time By Year



Percentages Of Queue By Month



AHT Forecasting Models

Model 1:

```
> fit_aht6 <- lm(AHT ~ Year + Month + weekday + time_24 + Queue + weekday*Queue + Month*Year, data = atrain)
> summary(fit_aht6)

Call:
lm(formula = AHT ~ Year + Month + weekday + time_24 + Queue + 
    weekday * Queue + Month * Year, data = atrain)

Residuals:
   Min     1Q Median     3Q    Max 
-575.5  -90.6  -19.3   57.7 4104.9 

Coefficients: (2 not defined because of singularities)
                                   Estimate Std. Error t value Pr(>|t|)    
(Intercept)                      -8.044e+05  6.714e+04 -11.981  < 2e-16 ***
Year                              1.997e+01  3.685e+00   5.418 6.03e-08 ***
MonthAugust                       7.238e+04  1.016e+04   7.127 1.03e-12 ***
MonthDecember                     1.153e+05  1.048e+04  11.001  < 2e-16 ***
MonthFebruary                    -1.136e+04  1.071e+04  -1.061 0.288798    
MonthJanuary                     -5.741e+04  1.054e+04  -5.445 5.19e-08 ***
MonthJuly                         3.820e+04  1.032e+04   3.702 0.000214 ***
MonthJune                         3.608e+04  1.034e+04   3.489 0.000485 ***
MonthMarch                       -2.241e+04  1.032e+04  -2.173 0.029810 *  
MonthMay                          2.337e+04  1.035e+04   2.259 0.023887 *  
MonthNovember                     5.660e+04  1.036e+04   5.462 4.71e-08 ***
MonthOctober                      2.350e+04  1.030e+04   2.281 0.022571 *  
MonthSeptember                    4.234e+04  1.045e+04   4.051 5.11e-05 ***
weekdayMonday                     1.204e+00  3.378e+00   0.356 0.721482    
weekdaySaturday                  -1.212e+02  8.303e+00 -14.600  < 2e-16 ***
weekdaySunday                    -2.021e+02  2.039e+01  -9.912  < 2e-16 ***
weekdayThursday                  -1.950e-02  3.309e+00  -0.006 0.995299    
weekdayTuesday                   -5.998e+00  3.303e+00  -1.816 0.069365 .  
weekdayWednesday                 -2.401e+00  3.296e+00  -0.728 0.466338    
time_24                          -3.461e-04  3.020e-05 -11.459  < 2e-16 ***
QueueETS                         -2.661e+02  3.132e+00 -84.961  < 2e-16 ***
QueueRes_SST                     -1.155e+02  4.053e+00 -28.501  < 2e-16 ***
QueueResidential                 -1.521e+02  3.463e+00 -43.923  < 2e-16 ***
weekdayMonday:QueueETS            2.790e-01  4.480e+00   0.062 0.950344    
weekdaySaturday:QueueETS          1.000e+02  8.825e+00  11.334  < 2e-16 ***
weekdaySunday:QueueETS            1.669e+02  2.062e+01   8.095 5.79e-16 ***
weekdayThursday:QueueETS          4.383e+00  4.427e+00   0.990 0.322087    
weekdayTuesday:QueueETS           1.141e+01  4.419e+00   2.583 0.009803 ** 
weekdayWednesday:QueueETS         6.702e+00  4.415e+00   1.518 0.129032    
weekdayMonday:QueueRes_SST        2.019e+01  5.871e+00   3.439 0.000584 ***
weekdaySaturday:QueueRes_SST     -3.885e+01  5.398e+01  -0.720 0.471749    
weekdaySunday:QueueRes_SST       -1.880e+02  1.373e+02  -1.370 0.170740    
weekdayThursday:QueueRes_SST      9.990e+00  5.725e+00   1.745 0.080950 .  
weekdayTuesday:QueueRes_SST       2.326e+01  5.690e+00   4.088 4.35e-05 ***
weekdayWednesday:QueueRes_SST     2.592e+00  5.703e+00   0.455 0.649412    
weekdayMonday:QueueResidential    1.729e+01  5.019e+00   3.445 0.000572 ***

weekdayMonday:QueueResidential    1.729e+01  5.019e+00   3.445 0.000572 ***
weekdaySaturday:QueueResidential         NA         NA      NA       NA    
weekdaySunday:QueueResidential           NA         NA      NA       NA    
weekdayThursday:QueueResidential  4.134e+00  4.898e+00   0.844 0.398700    
weekdayTuesday:QueueResidential   1.683e+01  4.888e+00   3.444 0.000574 ***
weekdayWednesday:QueueResidential 7.393e+00  4.877e+00   1.516 0.129517    
Year:MonthAugust                 -3.588e+01  5.034e+00  -7.127 1.03e-12 ***
Year:MonthDecember               -5.715e+01  5.197e+00 -10.997  < 2e-16 ***
Year:MonthFebruary                5.639e+00  5.309e+00   1.062 0.288188    
Year:MonthJanuary                 2.847e+01  5.226e+00   5.449 5.09e-08 ***
Year:MonthJuly                   -1.893e+01  5.114e+00  -3.702 0.000214 ***
Year:MonthJune                   -1.788e+01  5.126e+00  -3.489 0.000486 ***
Year:MonthMarch                   1.112e+01  5.113e+00   2.174 0.029688 *  
Year:MonthMay                    -1.159e+01  5.128e+00  -2.260 0.023815 *  
Year:MonthNovember               -2.804e+01  5.136e+00  -5.460 4.78e-08 ***
Year:MonthOctober                -1.164e+01  5.106e+00  -2.279 0.022688 *  
Year:MonthSeptember              -2.099e+01  5.181e+00  -4.050 5.11e-05 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 191.9 on 136458 degrees of freedom
Multiple R-squared:  0.2514,	Adjusted R-squared:  0.2511 
F-statistic: 935.3 on 49 and 136458 DF,  p-value: < 2.2e-16
```

Model 2:

```
Call:
lm(formula = AHT ~ Year + Month + weekday + time_24 + Queue,
    data = atrain)

Residuals:
   Min      1Q Median      3Q     Max
-561.5   -91.2  -18.6    58.6 4104.6

Coefficients:
                   Estimate Std. Error  t value Pr(>|t|)
(Intercept)      -7.750e+05  6.690e+04  -11.584  < 2e-16 ***
Year              6.526e+00  1.043e+00    6.258 3.91e-10 ***
MonthAugust      -2.029e-01  2.520e+00   -0.081  0.93581
MonthDecember     3.988e+01  2.600e+00   15.336  < 2e-16 ***
MonthFebruary     1.413e+01  2.656e+00    5.321 1.03e-07 ***
MonthJanuary      3.604e+01  2.615e+00   13.782  < 2e-16 ***
MonthJuly         2.783e+00  2.559e+00    1.088  0.27669
MonthJune         1.614e+00  2.566e+00    0.629  0.52927
MonthMarch        1.622e+01  2.559e+00    6.339 2.32e-10 ***
MonthMay         -1.224e+01  2.566e+00   -4.770 1.84e-06 ***
MonthNovember     2.680e+01  2.569e+00   10.430  < 2e-16 ***
MonthOctober      2.006e+01  2.556e+00    7.850 4.18e-15 ***
MonthSeptember    2.818e+00  2.594e+00    1.086  0.27737
weekdayMonday     7.769e+00  1.776e+00    4.374 1.22e-05 ***
weekdaySaturday  -2.922e+01  2.513e+00  -11.629  < 2e-16 ***
weekdaySunday    -3.837e+01  2.623e+00  -14.626  < 2e-16 ***
weekdayThursday   3.956e+00  1.747e+00    2.264  0.02356 *
weekdayTuesday    5.314e+00  1.744e+00    3.046  0.00232 **
weekdayWednesday  2.197e+00  1.743e+00    1.261  0.20746
time_24          -3.451e-04  3.028e-05  -11.399  < 2e-16 ***
QueueETS         -2.597e+02  1.402e+00 -185.196  < 2e-16 ***
QueueRes_SST     -1.043e+02  1.824e+00  -57.160  < 2e-16 ***
QueueResidential -1.457e+02  1.555e+00  -93.722  < 2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 192.4 on 136485 degrees of freedom
Multiple R-squared:  0.2475,    Adjusted R-squared:  0.2474
F-statistic:  2041 on 22 and 136485 DF,  p-value: < 2.2e-16
```

Model 3: Moving Average Model

Period	Year	Month	Actual	Forecasted AHT	Difference	Error Rate
1	2017	1	392			
2		2	386			
3		3	378	389	11	3%
4		4	372	382	10	3%
5		5	369	375	6	2%
6		6	372	371	-2	0%
7		7	374	371	-4	1%
8		8	383	373	-10	3%
9		9	384	379	-6	1%
10		10	393	384	-10	2%
11		11	405	389	-17	4%
12		12	445	399	-46	10%
13	2018	1	443	425	-18	4%
14		2	407	444	37	9%
15		3	369	425	56	15%
16		4	385	388	3	1%
17		5	369	377	8	2%
18		6	383	377	-6	2%
18		7	382	376	-6	2%
20		8	373	383	10	3%
21		9	378	378	-1	0%
22		10	399	376	-24	6%
23		11	384	389	5	1%
24		12	380	392	12	3%
25	2019	1	369	382	13	4%
26		2	376	381	5	1%
27		3	381	382	1	0%
28		4	375	381	6	2%
29		5	377	381	4	1%
30		6	376	381	5	1%
31		7	382	381	-1	0%
32		8	394	381	-13	3%
33		9	375	381	6	2%

Drexel University LeBow College of Business

Service Level Forecasting Model

Simplified Model:

```
Call:
lm(formula = Service_Level ~ Abandon_Rate + Agents + Queue +
    time_24, data = newda1)

Residuals:
     Min       1Q   Median       3Q      Max
-0.97046 -0.11507  0.04747  0.14692  0.92231

Coefficients:
                    Estimate Std. Error  t value Pr(>|t|)
(Intercept)       -1.111e+03  1.432e+02   -7.756 8.91e-15 ***
Abandon_Rate      -1.678e+00  1.371e-02 -122.362  < 2e-16 ***
Agents             5.313e-04  3.951e-05   13.449  < 2e-16 ***
QueueETS           4.112e-01  3.155e-03  130.346  < 2e-16 ***
QueueRes_SST       2.916e-01  3.844e-03   75.844  < 2e-16 ***
QueueResidential   2.289e-01  3.303e-03   69.302  < 2e-16 ***
time_24           -5.031e-07  6.483e-08   -7.760 8.66e-15 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 0.2466 on 49040 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.4379,    Adjusted R-squared:  0.4379
F-statistic:  6369 on 6 and 49040 DF,  p-value: < 2.2e-16
```

Full model:

```
Call:
lm(formula = Service_Level ~ ., data = newda1)

Residuals:
     Min       1Q   Median       3Q      Max
-0.94660 -0.06406  0.01954  0.07891  2.71102

Coefficients: (6 not defined because of singularities)
                          Estimate Std. Error  t value Pr(>|t|)
(Intercept)             -1.166e+03  1.046e+02  -11.144  < 2e-16 ***
Date                    -1.708e-04  9.420e-05   -1.813  0.06988 .
time_24                 -5.294e-07  4.736e-08  -11.179  < 2e-16 ***
QueueETS                 3.573e-01  1.779e-02   20.087  < 2e-16 ***
QueueRes_SST             3.636e-01  5.708e-03   63.704  < 2e-16 ***
QueueResidential         3.594e-01  6.656e-03   54.001  < 2e-16 ***
SkillBCST Other    0212  3.867e-02  4.110e-03    9.409  < 2e-16 ***
SkillBCST Strt_Stp Xfr 02  3.781e-02  4.088e-03    9.248  < 2e-16 ***
SkillElect Emerg     0202 -3.858e-02  1.766e-02   -2.184  0.02897 *
SkillGas Emerg   201    -2.449e-02  1.762e-02   -1.390  0.16464
SkillRES Billing   0230 -1.634e-01  6.631e-03  -24.648  < 2e-16 ***
SkillRES Other     0211 -6.335e-02  7.633e-03   -8.299  < 2e-16 ***
SkillRES Strt_Stp Xfr 022 -8.759e-02  6.259e-03  -13.995  < 2e-16 ***
SkillSp Elect Emerg      8.190e-03  1.922e-02    0.426  0.67001
SkillSp Gas Emerg               NA         NA       NA       NA
SkillSpn RES Billing    -1.185e-02  1.290e-02   -0.919  0.35819
SkillSpn RES Other              NA         NA       NA       NA
SkillSpn RES Strt_Stp Xfr       NA         NA       NA       NA
Offered                  6.733e-03  2.751e-04   24.469  < 2e-16 ***
Handled                 -1.903e-02  2.838e-04  -67.047  < 2e-16 ***
Abandon                         NA         NA       NA       NA
AnsInSvcl                1.388e-02  1.109e-04  125.171  < 2e-16 ***
AnsMinutes               1.675e-05  5.226e-07   32.054  < 2e-16 ***
AHT                     -1.746e-04  1.943e-05   -8.982  < 2e-16 ***
Talk                     1.776e-04  2.089e-05    8.505  < 2e-16 ***
Hold                     1.572e-04  2.717e-05    5.786 7.26e-09 ***
Wrap                            NA         NA       NA       NA
ASA                     -8.616e-04  8.061e-06 -106.884  < 2e-16 ***
Abandon_Rate            -9.557e-01  1.243e-02  -76.867  < 2e-16 ***
MonthAugust             -1.816e-02  1.203e-02   -1.510  0.13098
MonthFebruary            1.317e-02  6.511e-03    2.022  0.04314 *
MonthJanuary             9.907e-03  8.959e-03    1.106  0.26883
MonthJuly               -2.866e-02  9.257e-03   -3.096  0.00196 **
MonthJune               -1.760e-02  6.706e-03   -2.625  0.00867 **
MonthMarch               1.783e-03  4.356e-03    0.409  0.68222
MonthMay                -1.269e-03  4.392e-03   -0.289  0.77262
MonthSeptember           2.141e-02  1.438e-02    1.489  0.13654
Year                            NA         NA       NA       NA
Agents                   4.949e-04  3.652e-05   13.549  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1


Residual standard error: 0.1787 on 49014 degrees of freedom
  (1 observation deleted due to missingness)
Multiple R-squared:  0.7051,    Adjusted R-squared:  0.7049
F-statistic:  3662 on 32 and 49014 DF,  p-value: < 2.2e-16
```

DREXEL UNIVERSITY
LeBow
College of Business

Shrink Forecasting

Model 1:

```
> fit7 <- lm(Shrink_Duration ~ Year + Month + weekday + Year*Month + Month*weekday, data = strain)
> summary(fit7)

Call:
lm(formula = Shrink_Duration ~ Year + Month + weekday + Year *
    Month + Month * weekday, data = strain)

Residuals:
    Min      1Q  Median      3Q     Max
-657.86  -29.17    2.11   33.34  683.48

Coefficients:
                                  Estimate Std. Error t value Pr(>|t|)
(Intercept)                     -4.188e+04  6.739e+04  -0.621 0.534517
Year                             2.122e+01  3.340e+01   0.635 0.525550
MonthAugust                      1.637e+05  9.453e+04   1.731 0.083896 .
MonthDecember                    1.316e+05  9.551e+04   1.378 0.168758
MonthFebruary                   -2.258e+05  9.741e+04  -2.318 0.020754 *
MonthJanuary                    -2.990e+05  9.661e+04  -3.095 0.002060 **
MonthJuly                        3.460e+04  9.453e+04   0.366 0.714428
MonthJune                       -1.891e+04  9.765e+04  -0.194 0.846548
MonthMarch                      -3.222e+04  9.453e+04  -0.341 0.733307
MonthMay                        -3.224e+04  9.530e+04  -0.338 0.735240
MonthNovember                    1.866e+05  9.587e+04   1.946 0.052102 .
MonthOctober                     2.284e+05  9.453e+04   2.416 0.015970 *
MonthSeptember                   1.023e+05  9.530e+04   1.073 0.283467
weekdayMonday                   -6.098e+01  6.277e+01  -0.972 0.331653
weekdaySaturday                 -8.945e+02  6.277e+01 -14.251  < 2e-16 ***
weekdaySunday                   -9.078e+02  6.125e+01 -14.821  < 2e-16 ***
weekdayThursday                 -3.200e+01  6.456e+01  -0.496 0.620316
weekdayTuesday                  -1.812e+01  6.456e+01  -0.281 0.779003
weekdayWednesday                -6.638e+01  6.456e+01  -1.028 0.304309
Year:MonthAugust                -8.115e+01  4.685e+01  -1.732 0.083751 .
Year:MonthDecember              -6.523e+01  4.734e+01  -1.378 0.168752
Year:MonthFebruary               1.119e+02  4.828e+01   2.318 0.020766 *
Year:MonthJanuary                1.482e+02  4.789e+01   3.095 0.002057 **
Year:MonthJuly                  -1.717e+01  4.685e+01  -0.367 0.714102
Year:MonthJune                   9.346e+00  4.840e+01   0.193 0.846955
Year:MonthMarch                  1.596e+01  4.685e+01   0.341 0.733478
Year:MonthMay                    1.598e+01  4.724e+01   0.338 0.735291
Year:MonthNovember              -9.261e+01  4.752e+01  -1.949 0.051741 .
Year:MonthOctober               -1.133e+02  4.685e+01  -2.417 0.015926 *
Year:MonthSeptember             -5.076e+01  4.724e+01  -1.075 0.282992
MonthAugust:weekdayMonday        7.793e+00  8.877e+01   0.088 0.930071
MonthDecember:weekdayMonday     -1.795e+02  8.896e+01  -2.018 0.044025 *
MonthFebruary:weekdayMonday      1.882e+02  9.005e+01   2.090 0.036987 *
MonthJanuary:weekdayMonday      -3.742e+02  8.877e+01  -4.216 2.86e-05 ***
MonthJuly:weekdayMonday          1.638e+01  8.770e+01   0.187 0.851863
MonthJune:weekdayMonday          1.479e+01  9.041e+01   0.164 0.870143
MonthMarch:weekdayMonday         2.393e+01  8.770e+01   0.273 0.785015
MonthMay:weekdayMonday          -1.689e+02  8.877e+01  -1.903 0.057495 .
MonthNovember:weekdayMonday      1.890e+02  8.877e+01   2.129 0.033638 *
MonthOctober:weekdayMonday       1.504e+02  8.770e+01   1.714 0.086946 .
```

```
MonthSeptember:weekdaySaturday   9.711e+01  8.639e+01   1.124 0.261401
MonthAugust:weekdaySunday        7.931e+01  8.770e+01   0.904 0.366178
MonthDecember:weekdaySunday      1.225e+01  8.672e+01   0.141 0.887694
MonthFebruary:weekdaySunday      2.675e+01  8.899e+01   0.301 0.763828
MonthJanuary:weekdaySunday      -4.012e+01  8.899e+01  -0.451 0.652231
MonthJuly:weekdaySunday          5.135e+01  8.662e+01   0.593 0.553511
MonthJune:weekdaySunday          5.912e+01  8.936e+01   0.662 0.508470
MonthMarch:weekdaySunday         3.445e+01  8.662e+01   0.398 0.690972
MonthMay:weekdaySunday           1.225e+01  8.899e+01   0.138 0.890559
MonthNovember:weekdaySunday      2.930e+02  8.770e+01   3.341 0.000885 ***
MonthOctober:weekdaySunday       9.273e+01  8.770e+01   1.057 0.290743
MonthSeptember:weekdaySunday     1.110e+02  8.643e+01   1.285 0.199396
MonthAugust:weekdayThursday     -3.041e+00  8.770e+01  -0.035 0.972350
MonthDecember:weekdayThursday    1.488e+01  9.140e+01   0.163 0.870756
MonthFebruary:weekdayThursday    1.125e+01  9.130e+01   0.123 0.901977
MonthJanuary:weekdayThursday     7.375e+00  9.130e+01   0.081 0.935647
MonthJuly:weekdayThursday        9.500e+00  9.130e+01   0.104 0.917165
MonthJune:weekdayThursday        1.156e+01  8.873e+01   0.130 0.896426
MonthMarch:weekdayThursday       9.000e+00  8.662e+01   0.104 0.917279
MonthMay:weekdayThursday        -1.062e+01  9.130e+01  -0.116 0.907397
MonthNovember:weekdayThursday    3.904e+01  8.770e+01   0.445 0.656330
MonthOctober:weekdayThursday     1.075e+01  9.130e+01   0.118 0.906313
MonthSeptember:weekdayThursday   4.822e+00  9.005e+01   0.054 0.957312
MonthAugust:weekdayTuesday       1.902e+01  8.881e+01   0.214 0.830476
MonthDecember:weekdayTuesday    -1.777e+02  9.140e+01  -1.945 0.052260 .
MonthFebruary:weekdayTuesday    -5.400e+01  9.130e+01  -0.591 0.554446
MonthJanuary:weekdayTuesday     -8.909e+01  9.005e+01  -0.989 0.322849
MonthJuly:weekdayTuesday        -7.099e+01  9.005e+01  -0.788 0.430784
MonthJune:weekdayTuesday         2.378e+01  9.167e+01   0.259 0.795360
MonthMarch:weekdayTuesday        5.245e+01  8.899e+01   0.589 0.555821
MonthMay:weekdayTuesday         -3.335e+01  8.899e+01  -0.375 0.707972
MonthNovember:weekdayTuesday     1.880e+02  9.005e+01   2.088 0.037203 *
MonthOctober:weekdayTuesday     -1.980e+01  8.899e+01  -0.222 0.824004
MonthSeptember:weekdayTuesday    6.968e-01  9.005e+01   0.008 0.993828
MonthAugust:weekdayWednesday    -9.866e+00  8.770e+01  -0.112 0.910465
MonthDecember:weekdayWednesday   4.745e+01  9.294e+01   0.511 0.609862
MonthFebruary:weekdayWednesday   9.885e+00  9.295e+01   0.106 0.915340
MonthJanuary:weekdayWednesday    1.400e+01  9.130e+01   0.153 0.878184
MonthJuly:weekdayWednesday      -6.338e+01  9.130e+01  -0.694 0.487872
MonthJune:weekdayWednesday       1.707e+01  9.005e+01   0.190 0.849746
MonthMarch:weekdayWednesday      4.808e+01  8.770e+01   0.548 0.583693
MonthMay:weekdayWednesday        8.344e+00  9.005e+01   0.093 0.926200
MonthNovember:weekdayWednesday   2.253e+02  9.023e+01   2.497 0.012771 *
MonthOctober:weekdayWednesday    2.559e+01  9.005e+01   0.284 0.776397
MonthSeptember:weekdayWednesday  2.445e+01  9.005e+01   0.271 0.786102
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 129.1 on 618 degrees of freedom
Multiple R-squared:  0.9067,    Adjusted R-squared:  0.8923
F-statistic: 63.21 on 95 and 618 DF,  p-value: < 2.2e-16
```

Model 2:

```
> fit6 <- lm(Shrink_Duration ~ Year + Month + weekday, data = strain)
> summary(fit6)

Call:
lm(formula = Shrink_Duration ~ Year + Month + weekday, data = strain)

Residuals:
    Min      1Q  Median      3Q     Max 
-740.30  -31.61    7.77   50.03  870.79 

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)    
(Intercept)      -13611.158  21544.557  -0.632 0.527746    
Year                  7.195     10.679   0.674 0.500700    
MonthAugust         -51.198     25.835  -1.982 0.047907 *  
MonthDecember       -45.378     26.033  -1.743 0.081760 .  
MonthFebruary         7.482     26.619   0.281 0.778726    
MonthJanuary        -30.065     26.380  -1.140 0.254815    
MonthJuly           -40.802     25.821  -1.580 0.114519    
MonthJune           -26.270     26.635  -0.986 0.324320    
MonthMarch            6.446     25.833   0.250 0.803039    
MonthMay            -32.778     26.041  -1.259 0.208569    
MonthNovember      -115.095     26.152  -4.401 1.25e-05 ***
MonthOctober        -40.826     25.827  -1.581 0.114392    
MonthSeptember      -87.414     26.034  -3.358 0.000829 ***
weekdayMonday       -83.689     19.935  -4.198 3.04e-05 ***
weekdaySaturday    -838.455     19.974 -41.977  < 2e-16 ***
weekdaySunday      -844.599     19.934 -42.369  < 2e-16 ***
weekdayThursday     -24.418     19.918  -1.226 0.220647    
weekdayTuesday      -29.238     19.981  -1.463 0.143838    
weekdayWednesday    -37.320     20.123  -1.855 0.064072 .  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 142.6 on 695 degrees of freedom
Multiple R-squared:  0.8721,    Adjusted R-squared:  0.8687
F-statistic: 263.2 on 18 and 695 DF,  p-value: < 2.2e-16
```

Not-Shrink Forecasting Model:

```
> fitn1 <- lm(Not_Shrink_Duration ~ Year + Month + weekday, data = strain)
> summary(fitn1)

Call:
lm(formula = Not_Shrink_Duration ~ Year + Month + weekday, data = strain)

Residuals:
    Min      1Q  Median      3Q     Max 
-748.37  -20.48    8.48   52.64  292.95 

Coefficients:
                   Estimate Std. Error t value Pr(>|t|)    
(Intercept)      -1.144e+04  2.057e+04  -0.556 0.578372    
Year              6.031e+00  1.019e+01   0.592 0.554259    
MonthAugust       2.632e+01  2.466e+01   1.067 0.286315    
MonthDecember    -6.413e+01  2.485e+01  -2.581 0.010068 *  
MonthFebruary    -1.478e+01  2.541e+01  -0.582 0.560990    
MonthJanuary     -8.923e+01  2.518e+01  -3.543 0.000421 ***
MonthJuly        -2.918e+00  2.465e+01  -0.118 0.905804    
MonthJune         4.241e-01  2.543e+01   0.017 0.986696    
MonthMarch        5.015e+00  2.466e+01   0.203 0.838918    
MonthMay         -3.404e+01  2.486e+01  -1.370 0.171270    
MonthNovember    -3.725e+01  2.496e+01  -1.492 0.136163    
MonthOctober      9.533e-01  2.465e+01   0.039 0.969166    
MonthSeptember   -1.899e+01  2.485e+01  -0.764 0.444965    
weekdayMonday     2.937e+01  1.903e+01   1.543 0.123200    
weekdaySaturday  -6.567e+02  1.907e+01 -34.441  < 2e-16 ***
weekdaySunday    -6.607e+02  1.903e+01 -34.719  < 2e-16 ***
weekdayThursday  -7.955e+00  1.901e+01  -0.418 0.675811    
weekdayTuesday    7.207e+01  1.907e+01   3.778 0.000171 ***
weekdayWednesday  1.858e+01  1.921e+01   0.967 0.333640    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 136.1 on 695 degrees of freedom
Multiple R-squared:  0.8427,	Adjusted R-squared:  0.8386 
F-statistic: 206.8 on 18 and 695 DF,  p-value: < 2.2e-16
```

DREXEL UNIVERSITY LeBow College of Business

Reference

1. Tomas R. Robins, D.J. Medeiros, Terry P. and Harrison. Evaluating the Erlang C and Erlang A Models for Call Center Modeling. Working paper.

2. Noah Gans, Ger Koole, Avishai Mandelbaum (2003).Telephone Call Centers: Tutorial, Review, and Research Prospects. Manufacturing & Service Operations Management 5(2):79-141.

3. Sivan Aldor-Noiman, Paul D. Feigin and Avishai Mandelbaum, (2009) Workload Forecasting for a call center: methodology and a case study. The annals of Applied Statistics.

4. Hubert Setzler, Cem Saydam, Sungjune Park (2009). EMS call volume predictions: A comparative study. Computers & Operations Research.

5. Xu, W. (2000). Long range planning for call centers at Fedex. The Journal of Business Forecasting Methods and Systems.

6. Lawrence Brown, Noah Gans, Avishai Mandelbaum, Anat Sakow, Haipeng Shen, Sergey Zeltyn and Linda Zhao(2005). Statistical Analysis of a Telephone Call Center: A Queueing-Science Perspective. Journal of the American Statistical Association

7. James W. Taylor, A Comparison of Univariate Time Series Methods for Forecasting Intraday Arrivals at a Call Center, Management Science, 2008, Vol. 54, pp. 253-265.