Improving Staffing and Resource Allocation through Forecasting of Emergency Department Patient Volumes and Admission Rates

Jake Harrison, Parker Galligan, Joshua Jensen, Matthew Millis (PM), Cory Stewart

University of Arkansas College of Engineering || Parkland Health and Hospital System

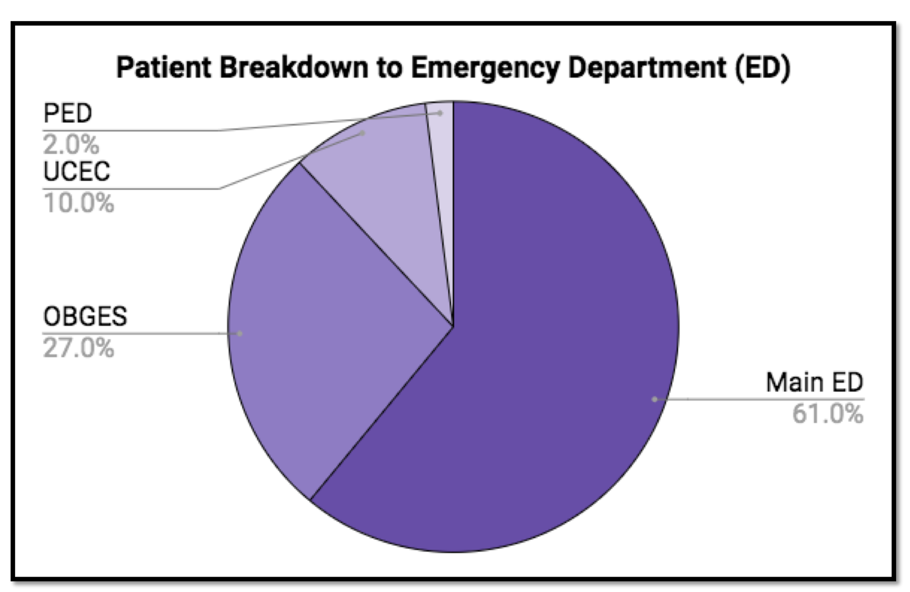
Abstract

Parkland Health & Hospital System's Emergency Department is one of the busiest in the nation with around 250,000 patients visiting annually. Emergency Department patient volumes can swing substantially (±40%) which can affect day-to-day operations. These volumes impact Parkland's Bed Access Management Team, who handle the allocation of beds to admitted patients, and Centralized Staffing Office, who deal with staffing throughout the hospital. Our team sought to provide Parkland with the ability to make informed decisions related to hospital-wide staffing and resource allocation by forecasting patient arrivals and admissions through the Emergency Department using predictive analytics.

We developed a multiple linear regression model with predictors representing Day of Week, Month, and Holidays. We were able to forecast patient arrivals and admissions as far as one year in advance. We have delivered our one-click update forecasts to Parkland including the evaluative metrics Mean Absolute Percent Error and Extreme Error Days. We anticipate these forecasts to have an impact on patient wait times and service levels throughout the hospital.

Industry Partner's Background

Parkland Health & Hospital System (PHHS) first opened its doors in 1894 and is now one of the largest medical systems in the country. Parkland provides a wide variety of services to the Dallas County and surrounding community, which, as of July 2017, has over 2.6 million residents. PHHS as a whole averages more than 1 million patients annually. In 2015, Parkland Hospital opened a brand new, state-of-the-art facility that encompasses 2.8 million square-feet and cost $1.27 billion dollars.



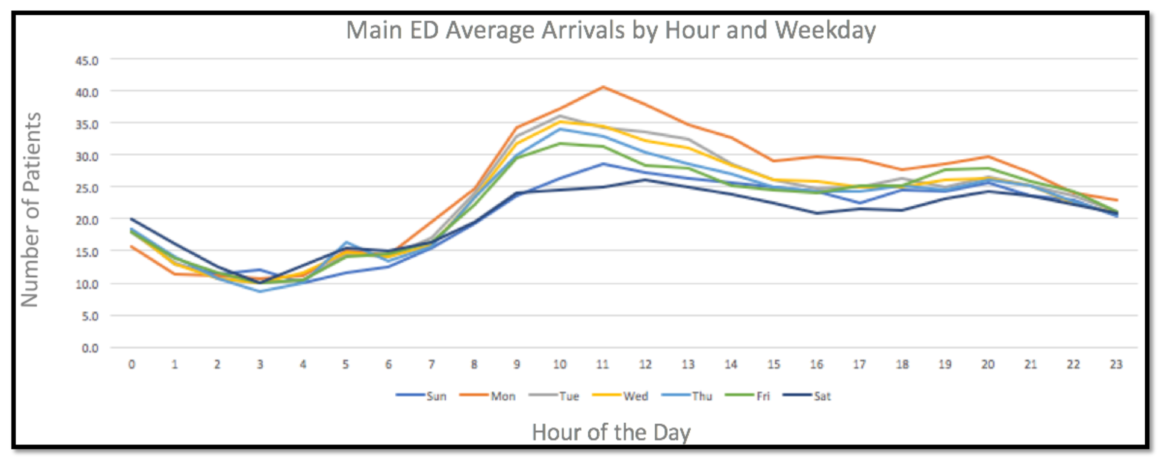
Literature Review

In order to assist Parkland in preparing for and treating complex patient volumes, our team decided to provide forecasted arrivals and admissions. Predictive analytics in healthcare is a widely researched field, and a paper written by Fogliatto et al provided a framework for how to approach these problems. These researchers investigated multiple external factors, primarily associated with weather, and provide forecasts of 1, 2, 7, 14, 21, and 30 days out using four different methodologies.

In evaluating the forecasts used by Fogliatto et al, mean absolute percentage error (MAPE) was the most common metric for measuring error. However, the team felt an additional metric could be helpful to evaluate different methodologies since we are looking to provide PHHS with an extremely accurate prediction. Therefore, extreme error days (EED), or occurrences of predictions beyond a specific threshold was added to our model.

Data Analysis

Parkland has provided three fiscal years of data beginning in October of 2015, the first full fiscal year "New Parkland" was open, and ending in September of 2018. The data contains numerous descriptors of the patient, arrival and discharge information, and their treatment needs.



Parkland was under the impression that patient volume behaviors were influenced by external factors such as weather (temperature and precipitation), seasonality, the lunar calendar, local events, and/or holidays. Our team analyzed which factors were relevant to our data using Minitab correlation analysis. In an effort to provide an preliminary forecasting method, the team created a forecast of arrivals through simple seasonal exponential smoothing.

Source Of Variation	P-Value
Weather	0.05
Month	0.00
Day of Week	0.00
Holidays	0.00
Lunar Calendar	0.70
Mavs' Game	0.77
R-Squared Value	**89.88%**

When we began our data analysis on the relevant external factors, we created the table shown below by taking the average number of arrivals to the Emergency Department by both Day of Week and Month.

Average Number of Arrivals by Day and Month	Sunday	Monday	Tuesday	Wednesday	Thursday	Friday	Saturday
January	424.6	705.4	645.8	619.3	608.1	581.8	503.9
February	431.8	729.9	640.5	648.3	607.5	612.1	510.9
March	426.3	718.1	651.6	628.2	617.9	610.4	502.6
April	415.1	731.0	661.9	628.5	637.0	612.8	497.3
May	419.6	702.0	654.0	638.0	627.5	616.0	504.3
June	430.9	744.6	682.6	656.4	626.4	630.9	514.3
July	437.1	724.6	660.5	674.1	646.8	636.6	525.7
August	443.5	724.3	675.6	639.9	618.0	622.1	514.0
September	450.1	727.9	715.8	669.1	632.1	632.6	546.9
October	407.6	698.7	645.5	614.1	583.7	573.4	472.7
November	388.8	676.2	598.4	588.9	516.8	572.9	475.6
December	379.3	652.4	609.1	587.4	539.8	540.8	450.7

Low — High

The main issue with this initial model was the number of EEDs, which we defined as any single prediction that is off by 10% in either direction. This initial model yielded 64 EEDs out of the 365 days tested. Holidays resulted in 11 of those EEDs, including six of the eight most extreme errors of the year. The model was consistently underpredicting holidays because fewer people come into the ED on these days.

In order to improve our predictions, we created a multiple linear regression model with the following categorical predictors: Day of Week, Month, and Holiday. Once this analysis was completed we were given a regression equation that could be used to make our predictions. The resulting coefficients we received can be seen in the figure below.

$$y = a + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n$$

Alpha Value	Month	Beta Value	Day of Week	Beta Value
609.3	January	12.7	Sunday	-179.3
	February	16.9	Monday	118.4
	March	7.2	Tuesday	52.5
	April	0.0	Wednesday	28.6
	May	-2.9	Thursday	5.7
	June	14.0	Friday	0.0
	July	21.6	Saturday	-98.2
	August	4.3		
	September	25.3		
	October	2.9		
	November	-16.9		
	December	-22.5		

When predicting holidays, we found that a more manual approach would be required. While holidays consistently saw fewer patients than normal, this decrease was never by a constant amount or a constant proportion. To tackle this problem, we divided holidays into two separate groups:

Holidays

Fixed-Date Holidays

Examples: *Christmas*, *Independence Day*

Method: *Use regression equation, but "downgrade" predictions*

Fixed-Week Day Holidays

Examples: *Thanksgiving*, *Labor Day*

Method: *Take the average of all past observations of that holiday*

Due to the regular values for number of admissions not being normally distributed with a constant variance, predicting the number of hospital admissions required a minor adjustment to our model.

Natural Log Transformation of Admissions

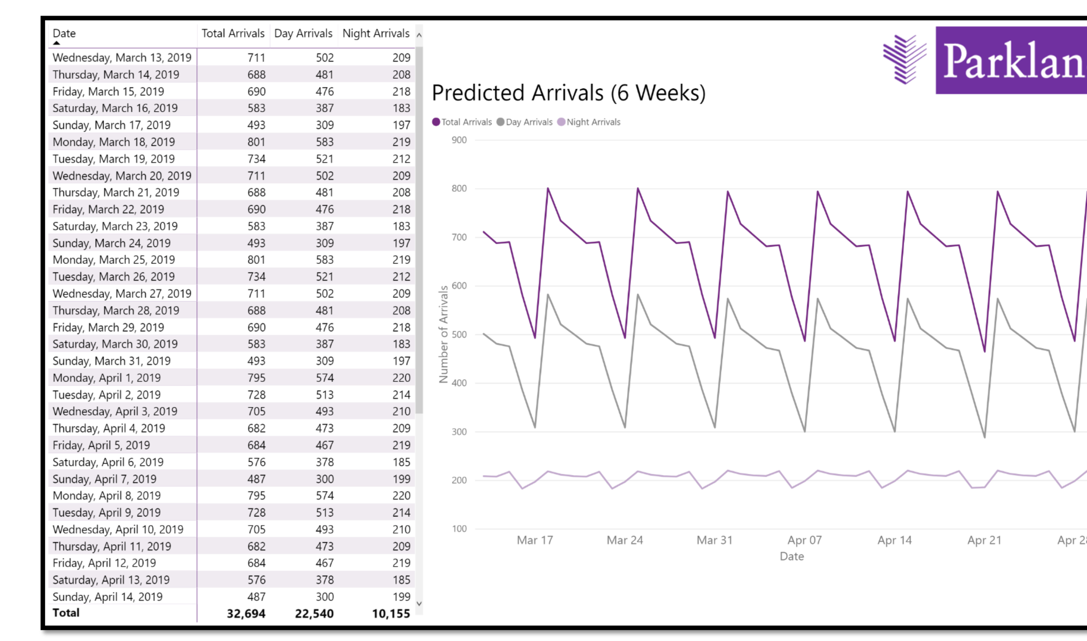
+

$$y = a + \beta_1 X_1 + \beta_2 X_2 + \cdots + \beta_n X_n$$

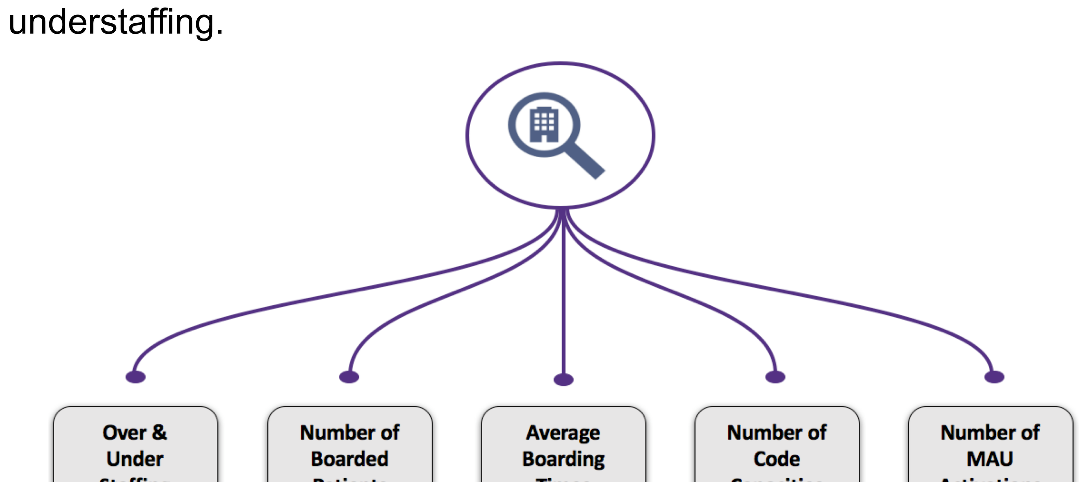
Conclusion

Forecasted ARRIVALS	Forecasted ADMISSIONS
MAPE = 4.02%	MAPE = 8.33%
EED = 22	EED = 82

The team concluded our deliverable will be a combination of extracting data from Parkland's Databases using Access and a Power BI application using VBA capable of performing the statistical tests we had been performing in Minitab.



Our team's goal was for our predictions to improve Parkland's operations by decreasing the number of Code Capacities and activations of the MAU due to the high costs associated with them as well as improve the scheduling process to avoid over and understaffing.



The team sought to measure the impact of our work in part by comparing Parkland's current operations with our forecast, as well as a theoretical best scheduling methodology where we will be using daily averages as a baseline.

Parkland's Current Operations	VS	Our Team's Forecasts	VS	Theoretical Best
MAPE = 13.71%		MAPE = 4.02%		MAPE = 3.86%