**BUSINESS PROBLEM**

Fargo Health Group is a healthcare service provider with 34 clinics across the United States. One of their services is providing disability compensation benefits to thousands of patients each year. The disability examination process involves the patient submitting a request to one of Fargo’s local offices. If a decision cannot be made at the office, the request is forwarded to one of the 34 clinics. On receipt of the request the clinic has 30 days to complete the examination and send the results back to the local office. The clinic can either attempt to complete the exam within the 30 days, or return the request to the local office. If the latter occurs, the local office has the decision to reroute the request to another clinic or send the request to an out-of-network Outpatient Clinic. Due to limited resources such as examining physicians, it is often difficult for the clinics to complete this task under the 30 day time frame. Fargo must pay $200 to the Regional Office of Health Oversight for each day past the deadline. Also, if an exam is sent to an Outpatient Clinic, Fargo must pay an additional $1,250 more than requested for an in-house exam and there is no guarantee that the exam will be completed within the deadline. The request has been made by Fargos’ Quality Assessment Office to have a more data-driven approach to predicting incoming examination volume in order to reduce feeds paid to the Regional Office of Health Oversight and the Outpatient Clinics. Fargo has issued a contract seeking external help to build a predictive analytics product in order to improve the scheduling of physicians. The proposal detailed below is a pilot study conducted on data given by Fargo Health Group and details the approach and forecast methods used to build the predictive model.

**DATA CLEANING PROCEDURE**

The specific task of the pilot study was to model and forecast incoming medical requests for heart related exams at the health center (HC) located in Abbeville, Louisiana. Data was given in various forms: all incoming medical requests to Abbeville from January 2006 to December 2013; incoming medical requests to Lafayette, Baton Rouge, New Orleans, and Violet HCs for May 2007; and all incoming medical exams for all HCs in December 2013. All data was cleaned and filtered to produce the incoming heart exams for the Abbeville HC from January 2006 to December 2013. Two forecasting techniques were then investigated to appropriately model future behavior of incoming exams.

The first step taken by the team was to clean the data of all non-numeric or unrealistic values and outliers. The incoming medical requests to the Abbeville HC data included multiple non-numeric values and some values with ‘999999999’. These values were initially set to missing, or ‘NA’. This was done by changing all values to characters, then to numeric types. The values that were ‘NA’, tested by using the ‘is.na’ function, were then set to ‘NA’. All values that were over 100,000 were also set to ‘NA’, as this seemed like an unrealistic value and also addressed the values that were set to ‘999999999’. 100,000 seemed like an unrealistic value since none of the values in the data set were close to this value, and it was more than 20 times higher than the maximum number of requests in any given month. On plotting the data it was seen that there was an outlier for October 2008 with a value of 3110 exams. The explanation for this data point was that there was a large increase in exams sent to Abbeville from New Orleans due to Hurricane Katrina. As this point does not represent an expected trend or seasonality, and is instead a very special case, this data point was set to ‘NA’ so as to not affect the forecast models.

The data for rerouted exams for the other HCs then needed to be split into month and year as it included data for May 2007 and May, June, and July of 2013 that needed to be included in the Abbeville requests. Due to the way the dates were read into R, this separation of data was done in Excel. The list from each non-Abbeville HC was filtered by date to get the rerouted exams for May 2007, May 2013, June 2013, and July 2013. Separate files were then made for each month and HC that were then read into R for further processing.

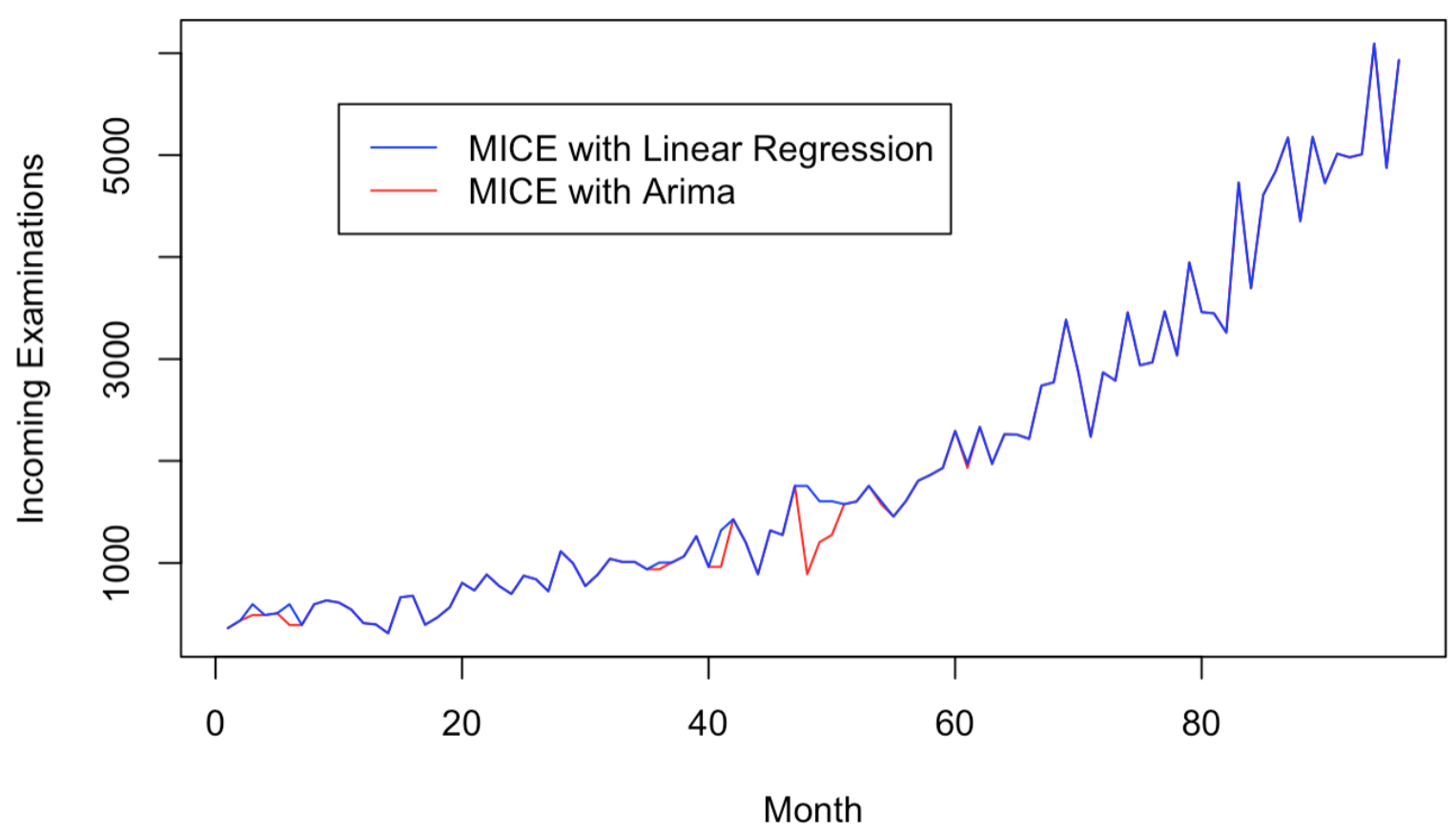
The next step was to determine the correct number of rerouted incoming cardiovascular requests from Abbeville for May 2007, May 2013, June 2013, and July 2013. A number of requests were sent to other HCs, so their requests needed to be filtered out for Abbeville cardiovascular requests for these months. For each HC (Violet, Baton Rouge, New Orleans, and Lafayette) and month the following process was taken. The exams that originated from Abbeville were first filtered out by selecting only those exams where ‘Original Hospital Location’ was ‘Abbeville’. The exams specific to cardiovascular issues then needed to be filtered out. The ‘Abbeville’ exams were then filtered by exams where ‘Examination’ was set to heart related conditions. The conditions were examined in each list and the conditions that were defined as ‘heart issues’ were the following: Heart, Cardiac, Cardiovascular, CAD, coronary Artery Disease (CAD), Heart Palpitations, Ventricular Septal Defect (VSD), VSD, Ischemic Heart Disease, Myocardial Ischemia, Cul Pulmunae, Myocarditis, Premature Ventricular Contraction, Aortic Valve Stenosis, and Arrhythmia. Any condition that involved the heart was considered a ‘heart issue’. The number of exams fitting these conditions were then counted from each HC and added to the existing May 2007 for Abbeville.

The Abbeville cardiovascular exams for December 2013 then needed to be counted. An Excel sheet with routing numbers and SYS IDs for all rerouted exams was given. The Abbeville exams had IDs starting with ‘L839’ and ended with ‘TGU3’ or ‘ROV8’. A list of heart condition codes was also given, each of which would be contained in the SYS ID. The procedure started by filtering out the entire December 2013 list of ID by those containing ‘L839’. This list was then passed to another filter that selected those IDs containing either ‘TGU3’ or ‘ROV8’. The list produced form this function was the medical exams for Abbeville. The next step was to select those exams related to heart conditions, in which the SYS ID contained one of the heart codes given. The Abbeville list was then filtered based on whether or not the ID contained one of the heart codes. This was done by looping through each of the heart codes to see if it was contained in the Abbeville ID. The final list was the Abbeville rerouted heart exams for December 2013.

The next step was to check all Abbeville data for duplicates. The original Abbeville list for January 2006 to December 2013 was ordered by year and then month, and was then visually inspected for repeating month entries. There were no duplicates for this particular list. The Abbeville heart rerouted exams to other HCs for May 2007 were then inspected for duplicate ‘Request IDs’. This field was used to check for duplicates as this number should be a unique identifier and thus should be only used once. There were duplicates found for all lists. There were 51, 41, 48, and 36 duplicates rerouted to the Violet, Lafayette, Baton Rouge, and New Orleans HCs, respectively, during May 2007. There were no duplicates rerouted during 2013. It was deemed impossible to find duplicate values for the December 2013 data as there were no unique identifiers for a rerouted exam, i.e. it was believed possible to have multiple exams with the same Routing SYSID. Once the rerouted values were found for May 2007, and May, June, and July 2013, they were added to the existing values from the Abbeville list.

**DATA IMPUTATION METHODS**

Once the above steps were accomplished, the missing values had to be imputed. The data set for Abbeville HC heart related medical requests had 11 missing values. The MICE method was used to impute these values. Linear regression and Arima were both tested with the MICE method to determine which method produced better imputed values. The MICE with linear regression method produced five sets of imputed values. All sets yielded the same coefficients with an intercept of -484.71 and a slope of 51.85 exams/month, so the first set of values was chosen for comparison. The MICE with arima method also produced five sets of imputed values all with the same coefficients with an intercept of 2078.353 and a slope of 171.437 exams/month and an AIC value of 1497.36. Since there was no difference in coefficients or AIC values, the first data set was also chosen for comparison. The two data sets are plotted below in Figure 1.



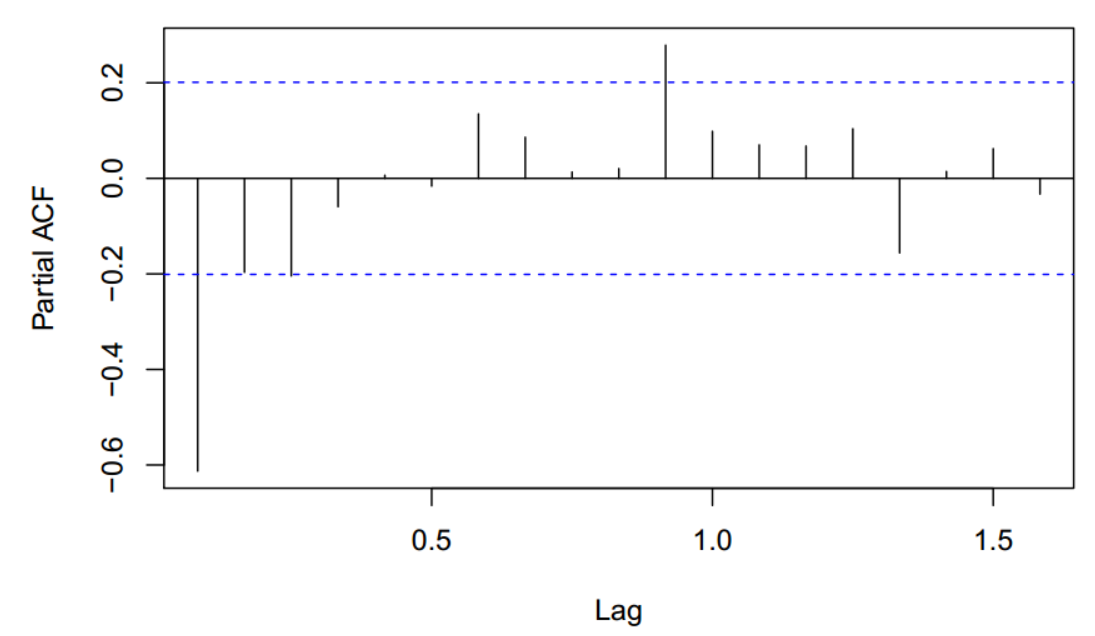
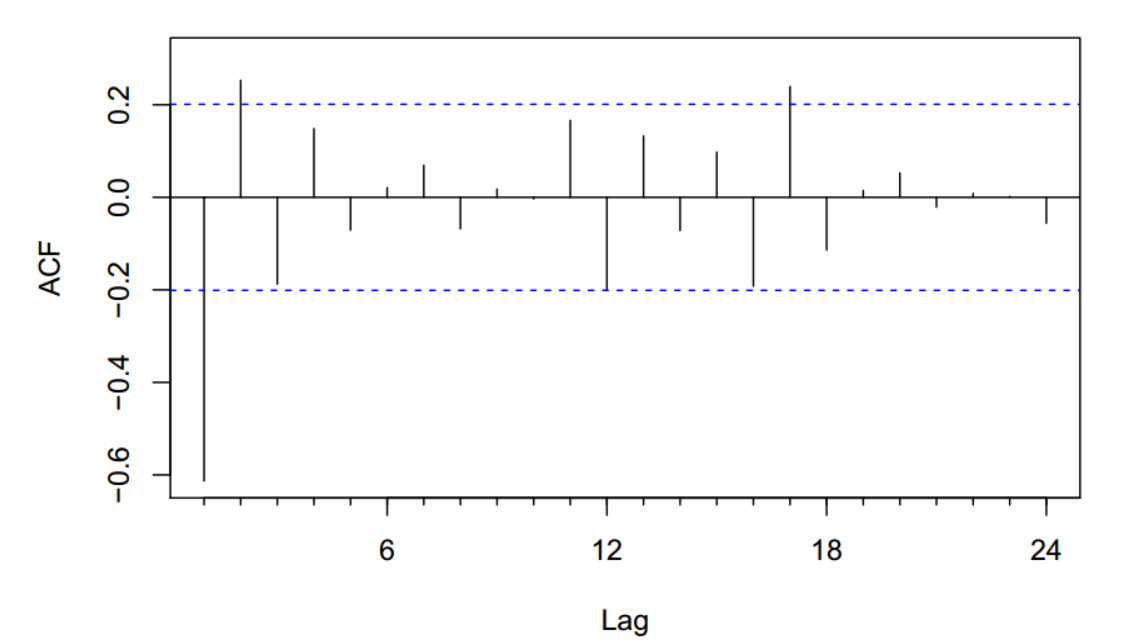
**Figure 1:** Comparison of two data sets using different imputation methods.

It can be seen that the biggest difference between the two methods occurs near month 50, where the MICE with Arima method predicts a much lower value than the MICE with linear regression method. This is most likely due to Arima being a better method with seasonality. Since there is some seasonality in the data, the MICE with Arima method imputed values were used for further analysis.

**FORECAST MODELS**

Once all of the missing values were imputed and the final data set was produced, two forecasting models were investigated. The two forecasting models used were built using Arima and Holt-Winters. These techniques were used as there was a seemingly multiplicative trend in the data, thus making these methods more acceptable than using a weighted average or simple exponential smoothing. There is also some seasonality in the data which can be accounted for in both of these models.

The Arima model was built first. As this model cannot work with data with trends, the ndiff function was used to remove the trend. The Dickey-Fuller test showed that one difference was enough to make the data stationary. The ACF and PACF were then calculated with this data to determine the p and q values for the Arima model. Both graphs are shown below in Figure 2.

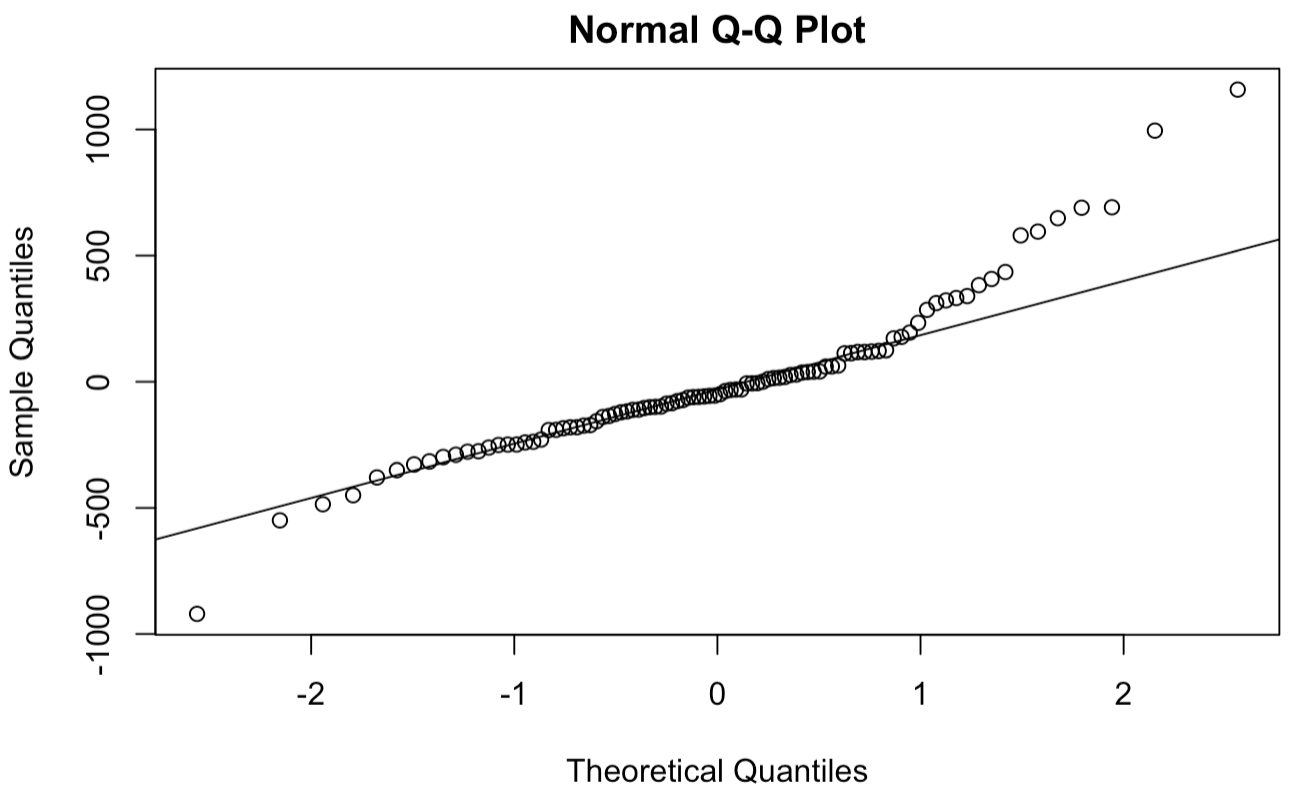
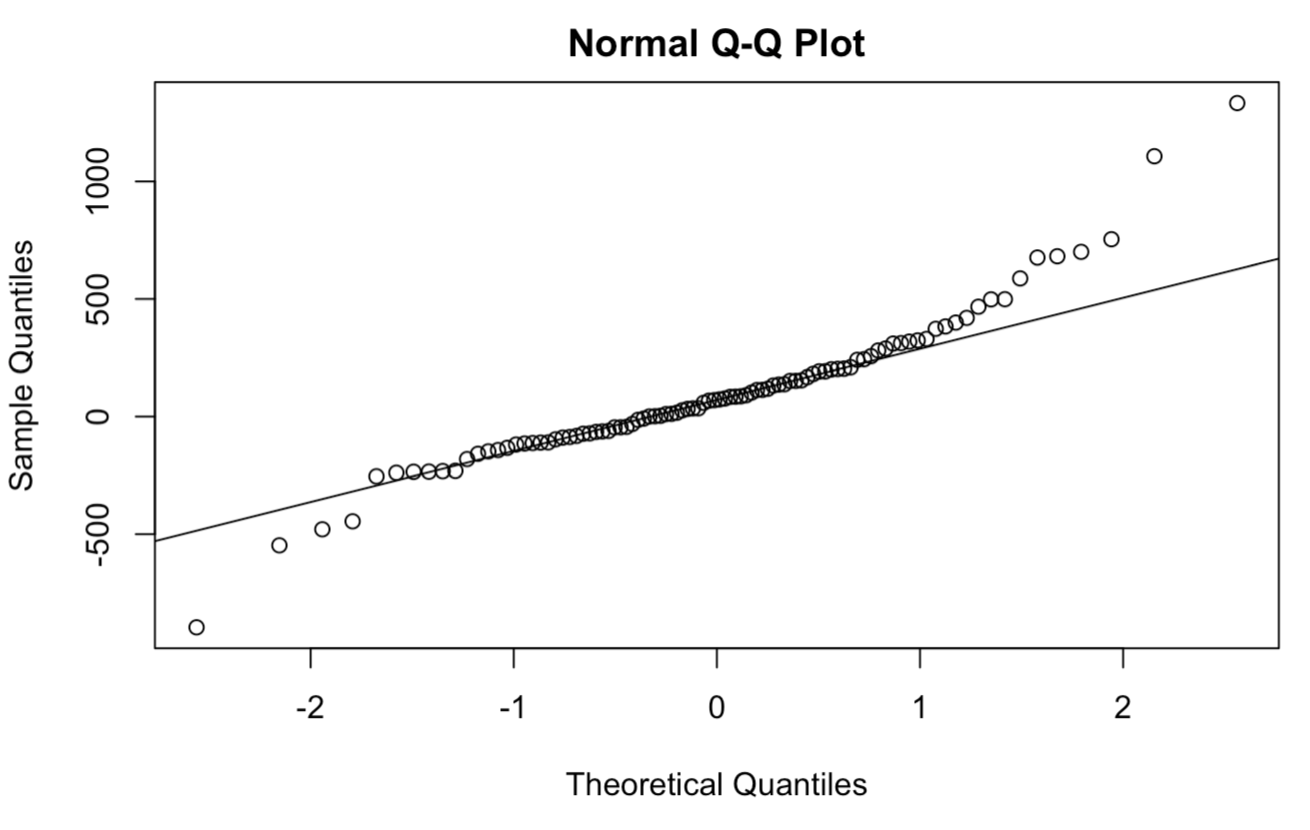


**Figure 2:** ACF and PACF of time series data.

There is one large autocorrelation at lag one. Both the ACF and PCF trail to zero overall as the lags increase. Due to this and the fact that the time series was differenced once, an Arima(1,1,1) model was first tried. In order to find the best model, the auto.arima function was also used for comparison. This function produced an Arima(1,1,1) with drift model. The accuracy measurements are shown below in Table 1. It can be seen that the auto.arima function did indeed produce a better model in that it produced a lower error value for almost all of the measurements, including the most sited measurement, RMSE, and the AIC value. It should be noted here as well that both models yielded Q-Q plots with data that mostly fell along the line, as shown in Figure 3. Due to this, the decision was made to use the Arima(1,1,1) with Drift model for comparison with a Holt-Winters model.

**Table 1:** Comparison of Arima models.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **ME** | **RMSE** | **MAE** | **MPE** | **MAPE** | **MASE** | **AIC** |
| Arima(1,1,1) | 98.862 | 325.405 | 224.025 | 2.765 | 14.098 | 0.814 | 1377 |
| Arima(1,1,1) with Drift | 0.212 | 305.190 | 212.920 | -7.064 | 15.357 | 0.320 | 1366 |



**Figure 3:** Q-Q plots for Arima models; Arima(1,1,1) on left and Arima(1,1,1) with Drift on right.

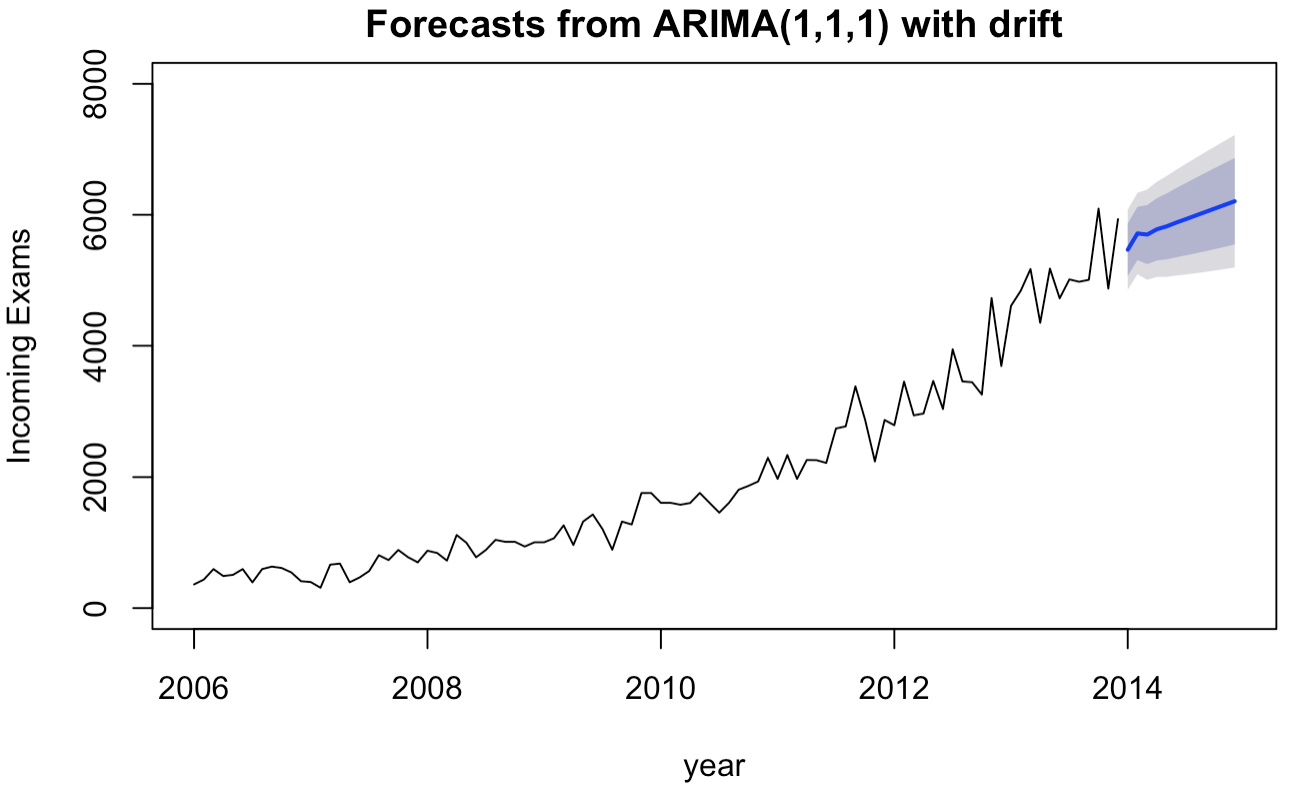
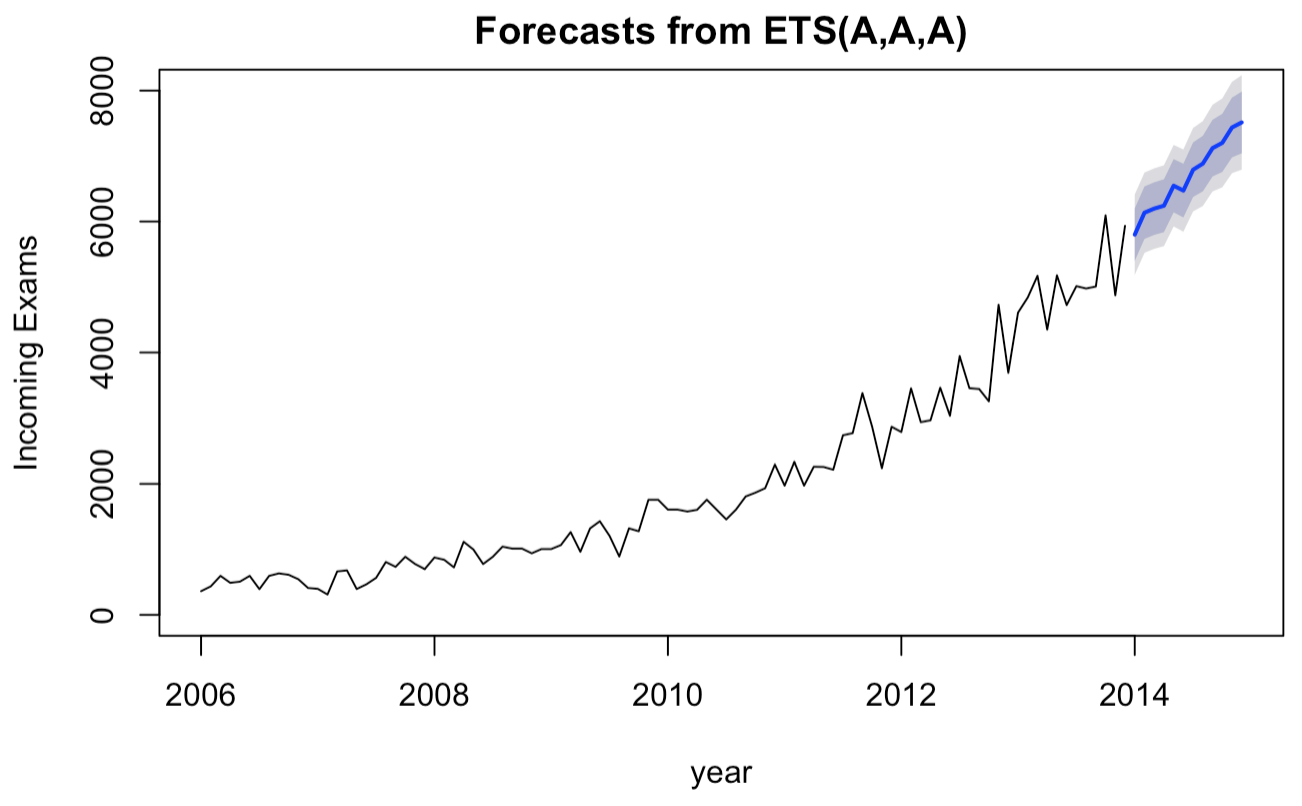
A number of Holt-Winters models were also built and compared. The ets function was used to build these models. The first model built used an untransformed data set with setting ‘AAA’, as there is a trend and some seasonality in the data. The ‘AAA’ implies that there is an additive level, slope, and seasonal parameter fit in the model. The second model was built using the log transformation of the time series data and a setting of ‘AAA’, as there was a slight multiplicative trend in the data. The ets function also features an auto selection where it will select the best model if none is supplied, this action used on the non-transformed data set resulted in a (M,A,N) model; which is a multiplicative level, additive trend, and no seasonal parameter. The accuracy measurements for all three models are below in Table 2. **It should be noted that the values for the log transformed data should not be compared with the other values as these are not the same units.**

**Table 2:** Comparison of Holt-Winters models.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **ME** | **RMSE** | **MAE** | **MPE** | **MAPE** | **MASE** | **AIC** |
| A,A,A | 64.051 | 285.277 | 203.691 | 1.199 | 13.793 | 0.306 | 1558 |
| log transformed; A,A,A\* | -0.003 | 0.162 | 0.121 | -0.114 | 1.733 | 0.341 | 113 |
| M,A,N | 81.139 | 307.250 | 206.342 | 0.034 | 12.800 | 0.310 | 1450 |
| \*Accuracy values have not been transformed back into original units and should not be compared with other model values. | | | | | | |
|

It can be seen from the table that the ‘A,A,A’ model produced the best accuracy values for most of the measurements, including the most sited measurement, RMSE. The ‘M,A,N’ model does produce a lower AIC value, however it does not possess a seasonal component. The original data does show a small seasonal component, and therefore the ‘AAA’ model was chosen for comparison with the Arima model.

A period of 12 months was used as the forecast period as it seemed an appropriate percentage of the total eight years of data supplied. This forecast value was also chosen as it was long enough to predict any seasonality in the forecasted data. The forecasted values with 85 and 95 % confidence intervals are plotted below for each model in Figure 4.

**Figure 4:** Forecast values with confidence intervals for Arima and Holt-Winters models.

The figures above show that the Holt-Winters model predicts a steeper increase in incoming exams for 2014, following the most recent trend in the data. The Arima model predicts a slight increase in incoming exams, however it does not follow the somewhat exponential trend from 2010 to 2013. Neither model shows much seasonality, the Holt-Winters showing only a small amount, and the Arima model showing none. It can also been seen that the Holt-Winters model has much smaller confidence intervals. Table 3 shows the accuracy measurements for each model again for easier comparison.

**Table 3:** Accuracy measurements for Arima and Holt-Winters models.

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **ME** | **RMSE** | **MAE** | **MPE** | **MAPE** | **MASE** | **AIC** |
| Arima(1,1,1) with Drift | 0.212 | 305.190 | 212.920 | -7.064 | 15.357 | 0.320 | 1366 |
| Holt-Winters A,A,A | 64.051 | 285.277 | 203.691 | 1.199 | 13.793 | 0.306 | 1558 |

On taking a closer look at the two models, the Holt-Winters model has more accuracy measurements with values below the Arima model. The Holt-Winters model also has a lower RMSE value, though the Arima model has a lower AIC value. Due to having more low accuracy measurements and the overall trend, the Holt-Winters model was chosen as the best model for predicting future incoming exam volume. The final forecasted values from this model are shown below in Table 4.

**Table 4:** Forecasted incoming heart issue exams for Abbeville HC in 2014.

|  |  |
| --- | --- |
| **Month** | **Forecasted Exams** |
| January 2014 | 5799 |
| February 2014 | 6134 |
| March 2014 | 6197 |
| April 2014 | 6241 |
| May 2014 | 6546 |
| June 2014 | 6472 |
| July 2014 | 6791 |
| August 2014 | 6884 |
| September 2014 | 7121 |
| October 2014 | 7201 |
| November 2014 | 7436 |
| December 2014 | 7513 |

Each of the models compared above make certain assumptions about the data and have certain limitations. The Holt-Winters model, since it is a form of an exponential smoothing model, should not be used for long term forecasts. Therefore, it is recommended that this type of model not be used to forecast periods longer than 12 months with the same amount of given data (8 years). Also, since the form of model used for the Holt-Winter model included only parameters for a multiplicative level and additive trend with no seasonal component, it will not predict much seasonality in the forecasted values. The other model compared, using ARIMA, assumes that the data is stationary, or that the moving average does not change with time. Since the data in this pilot study was not stationary, it had to be ‘differenced’ before making the ARIMA model.

**RECOMMENDATIONS**

**Data Entry**

Throughout the data cleaning process, a few common types of incorrect data entries were noticed. All incorrect entries were made in the field where ‘incoming examinations’ should have been. The two most common incorrect entries were ‘\*’ or ‘999999999’. Out of the 96 months that data was given for, 11 of them were missing incoming exam values. Ensuring that values entered for incoming exams are correct, numerical values would increase the accuracy of the model and prevent having to impute values. Perhaps writing a program for the data entry software such that a numerical value or ‘NA’ would need to be recorded before submitting an entry would help with this. It was also noticed that multiple duplicate entries existed for rerouted exams during May 2017. As no duplicates were rerouted for other months, this was perhaps a user error, however making sure that there is only one unique ‘Request ID’ in the system would ensure that no duplicates exist for rerouted exams.

**Data Analysis**

As it is unlikely that the number of incoming exams will continue to increase indefinitely, better models could be produced by correlating number of exams to other occurrences or population attributes. Another, more proactive, approach would be to make a predictive model to determine if a patient requires disability benefits for the local offices to use, therefore reducing the amount of exams requested to clinics along with examining physicians needed for all clinics. A model such as this would take time to make, and many attributes of past data would need to be analyzed to ensure that the model would properly diagnose a patient as ‘disabled’ or ‘not disabled’.

**Resource Staffing**

The final model used for this pilot study predicts that incoming exams will increase at a rate similar to that in 2013. It is also noted that while there are some periods throughout the years 2006-2013 that experience lower volumes than others, the forecast for 2014 increases fairly consistently throughout the year and staffing of examining physicians at the Abbeville HC should be done with this in mind.

**ETHICAL IMPLICATIONS**

This pilot study involved using data collected from multiple health centers at Fargo Health Group containing information on disability exams requested by patients. The specific data used for the forecast models was number of exams, location of health center, and type of exam requested. This was the only patient data used for the forecast model. Although there was minimal patient data used, the paragraphs below address some concerns about this and future data analysis work.

The data used for this analysis originated from data collected from a patient. This data was originally intended to be used to determine if the patient had a disability. Some of this data was used to forecast future volumes of incoming requests in order to save Fargo Health Group money as well as to preserve Fargo’s good reputation. During this pilot study there was no evidence of patient acknowledgement or consent of their information being used for this purpose. Since the data is being used for a reason different than originally intended, which is probably unknown to the patient, the patient should be informed prior to submitting their request of the intended and possible uses of their information.

Incoming exam data was given for approximately 8 years, from which a forecast was built for the next year. The 8 years of prior data was deemed reasonable for this forecast as it gave enough insight into trends and seasonality for incoming exams. It is believed that having more data prior to 2006 would have been of little use since the trends prior to and after 2010 are different; pre-2010 increasing slightly over time and post-2010 increasing at a much faster rate. Although the data given was sufficient for the model, the 2014 forecast is given with caution, as it is unlikely that the number of incoming exams will increase at this rate indefinitely.

Although it appears that the patients were not aware that some of their data would be used in a forecast model, the actions recommended in this pilot study would certainly prove beneficial to not only Fargo Health Group corporate and employees, but also the patients as they will be able to complete their exams in a timely manner and receive their benefits on time. This model should help Fargo properly staff their health centers such that they have more physicians available during high demand times and less available during slow periods. Although this may suggest to Fargo to reduce staff if a steady decrease in volume is predicted, the reverse is true for increasing volume which would create new jobs. Assuming that Fargo would like to keep the number of staffed employees constant, this model could also be used to rearrange or relocate employees where they are most needed, thus reducing the need to reduce total head count.

Fargo Health Group is the owner of all data and forecast models. It is at their discretion to make the recommended actions in order to reduce fees to outside organizations and ensure that requests are processed on time. As Fargo is ultimately responsible for processing these requests from patients, it is their responsibility to ensure they are processed on time. Due to the time in which outside requests are processed being outside of Fargo’s control, it is in their best interest to be prepared to process as many requests as possible in-house.

Fargo Health Group outsourced the analysis of this pilot study and as such the outsourced party is somewhat responsible for any mistakes or unintended consequences of the model. It is the responsibility of the outsourced party to convey the assumptions of the model and the accuracy of the forecasted values. However, since Fargo is the organization taking the action, they are the main party responsible for any unintended consequences of the suggested actions.