Report Structure:

Executive Summary

Analytic Approach and Assumptions

Nature and Structure of Data, Fixes and Resolution

Description of the Data Analysis Approach

Summarization of Key Findings

Ethical Implications of the Forecasting Model and its Implementation

Recommendations

Executive Summary

The goal of this project is to determine whether or not a data-driven analysis of historical levels of cardiovascular examinations provides an accurate enough assessment of future examination level growth, and thus an opportunity to better recognize future staffing needs. A more accurate forecast would provide the Human Resources department better information to plan when additional physicians are needed currently and over a long-term horizon.

Consequently, better service and reductions in cost could be provided to patients with a staffing level that is in tune with growing demand. Time frames for service past the 30-day window have the potential to be shortened with a right-sized staff. This would result in a decrease in fines of $200 per patient per day from the Regional Office of Health Oversight (ROHO). Additionally, any examinations that cannot be performed locally due to staffing limitations require outsourcing to outpatient clinics not in Fargo’s network. This causes Fargo to incur - on average - an extra $1250 per patient outsourced to an organization not bound by the 30-day window. Moreover, the increased costs of clerical overhead to track these inefficiencies cannot be understated.

Analytic Approach and Assumptions

Modeling a time series to the dataset is an appropriate approach to forecasting future levels of cardiovascular examinations. Doing so will take into account any seasonality (if found) and trends for long-term growth. Monthly fluctuations in examination counts can be modeled into the forecast as well. The decision to emphasize patterns of more recent observations as opposed to historical ones can easily be applied automatically in order to make the process more accurate.

Measures of performance of time series models are well developed and have utility across a wide range of endeavors that employ them for decision making. Two separate time series model methodologies will be fit to the data (using ARIMA and an exponential time series using forecast.ets) in order to see which one best suits the data provided. This approach assumes that any missing data in the dataset is noted prior to the analysis so a strategy can be derived for handling it in an appropriate manner.

Nature and structure of data, fixes and resolution:

The collected data for cardiovascular visits spanned a period beginning in January 2006 and concluded in December 2013. It is found in the accompanying DatasetClean spreadsheet on the ‘Abbeville’ tab. A number of imperfections were found in the dataset as seen in Figure 1, ranging from data missing outright, to non-numeric characters and improbable service levels. Details for the issues surrounding each data point at issue are detailed below.

Two methods were employed during the data cleaning process for missing or invalid data. The first involved using data from the additional spreadsheet tabs in order to derive some of the missing information. The second involved the use of the Amelia package for data imputation and simulation. In summary, this package provides replacements for missing data in context of the other points in the dataset. Further background on the usage of the product will be provided later in the report.

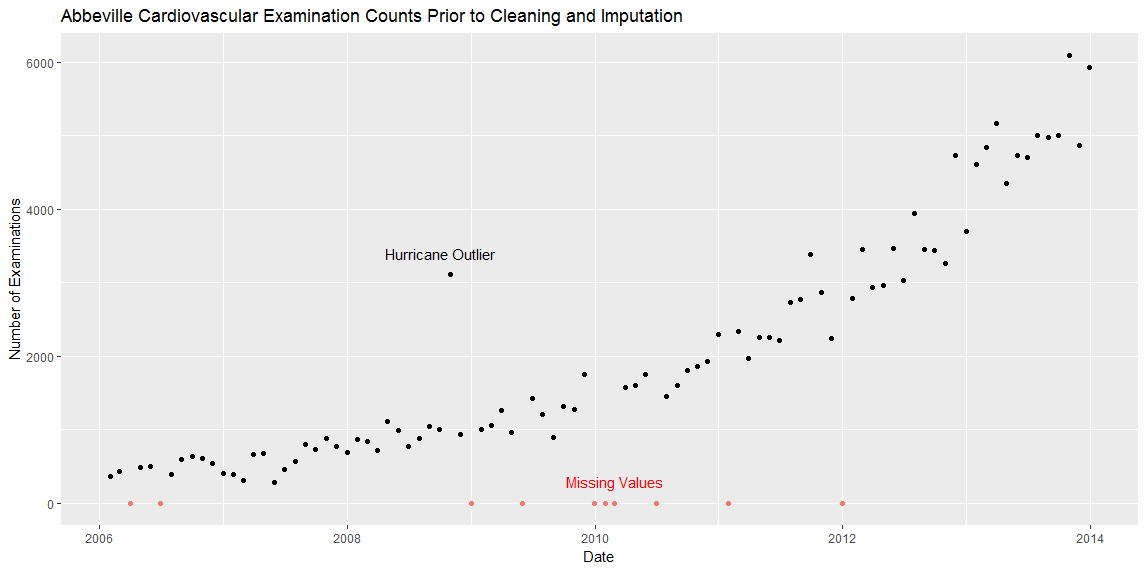


Figure 1 - Dataset prior to cleaning

The at-issue data points were broken down into the following categories:

1. Missing or invalid data, no caveats:

March and June 2006, December 2008, May 2009, June 2010, January 2011 and December 2011. This data will be imputed through results of Amelia simulations.

1. Missing data that can be cross-referenced through Excel:

The May 2007 data has an aberration; the Abbeville location was closed for repairs on the second week of the month. Checking the tabs for the Violet, New Orleans, Lafayette, and Baton Rouge locations for any May 2007 visits where the referring location was Abbeville, 173 cardiovascular-related exams were visually identified. Consequently, May 2007 should be increased from 107 to 280 examinations.

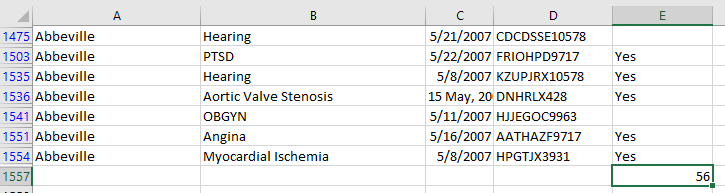


Figure 2 - May 2007 aberration resolution

1. Outlier data:

October 2008 was artificially inflated due to the New Orleans facility being closed due to the effects of a hurricane. This value will be removed during processing in order to allow Amelia to impute it in the context of the rest of the time series.

1. Missing, with caveats:

Details for December 2009 through February 2010 examination counts are missing, but it is known that the sum of the three is 5129. Simulations of various Amelia imputations will be performed to see which simulation has those three values close to 5129. This simulation is in a separate R file, and can show the set.seed() value to be used to get Amelia to reproduce that scenario in the main R file.

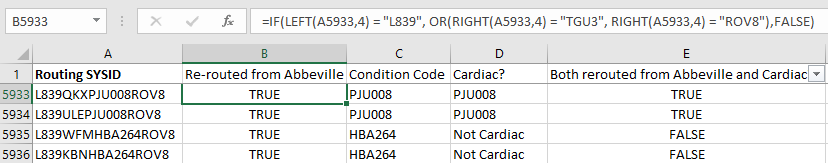
source('~/DS700 Final Project - Cleaning and Forecasting Seed evaluation.R')

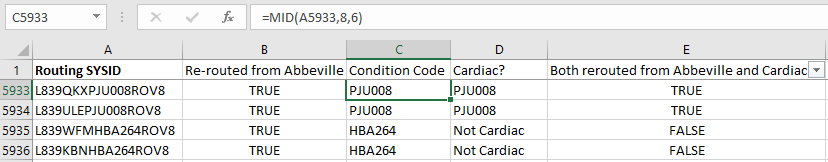
[1] "Seed to set total visits of months 48/49/50 to near 5100 is: 13 total: 5255"

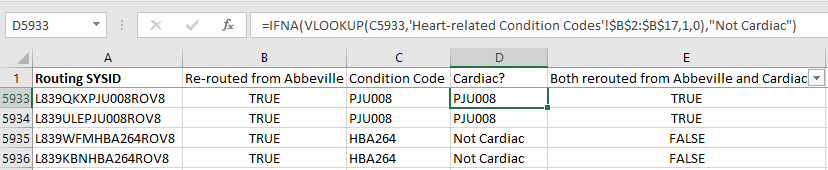
December 2013 examination visit data for Abbeville could be reconstructed from data in the spreadsheet. The conditions tested for were whether the routing ID started with “L839” and ended with either “TGU3” or “ROV8”, which identifies it as an examination being re-routed from Abbeville, as seen below in Figure 2.

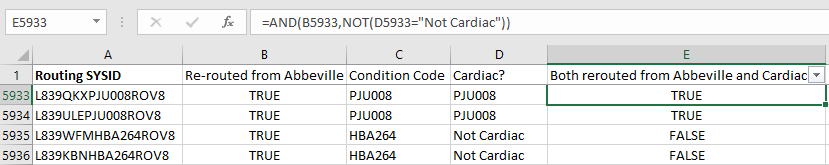
Another test was made to see if the segment of the routing ID that contains conditions code matches against entries in the ‘Heart-Related Condition Code’ tab. The rows where column E is true are the number of cardiac examinations referred from Abbeville that month, totaling 5933.

Figure 3 - Excel methods for finding Dec 2013 data









For missing values, the Amelia package was the sole method of imputation of the missing variables. It uses a bootstrap+EM algorithm to survey the dataset and provide ranges of permissible values that are consistent with the rest of the dataset. It also provides output such that ranges of data points within + 2 standard deviations can be plotted and visually identified for reasonableness, as seen in Figure 4:

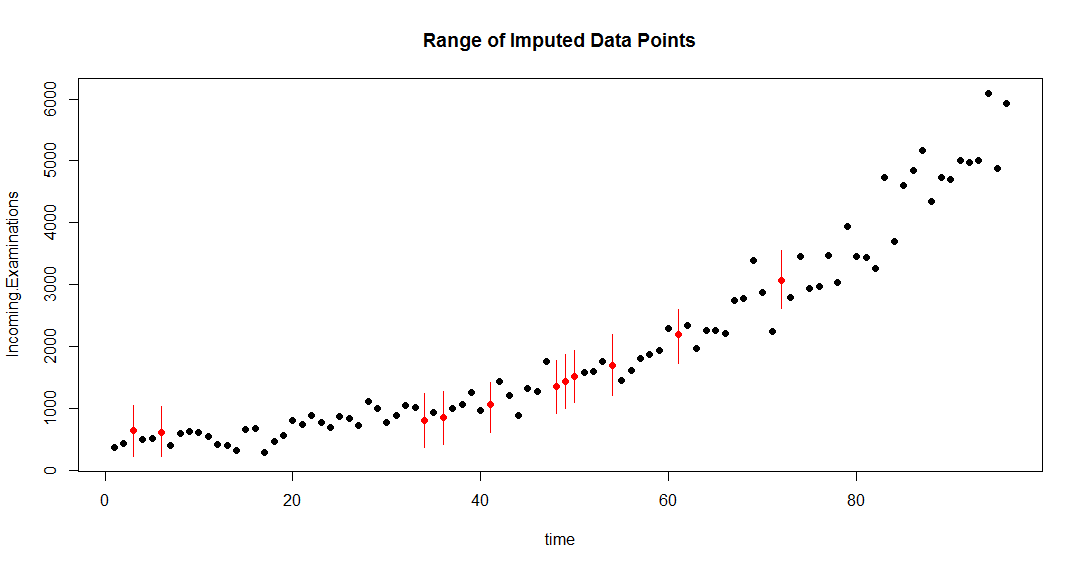


Figure 4- Cleaned dataset with missing data imputed

Default settings for Amelia allow for five different imputation possibilities. The value was instead set to one, so simulations could more easily be performed (next paragraph). There are provisions for parameters to control the style of curve that can fit the time series. In this instance, polytime and splinetime values were set to 2, allowing for non-linear fits for replacement values. This gives the model freedom to fit other than straight-line effects to potential replacements:

a.out <- amelia(abbeville2, ts="index", cs="groupID", polytime=2, splinetime=2, m=1, p2s=0)

To replace the missing points, the process keyed off of the three missing months (December 2009 through February 2010) where the total number of examinations was known (5129) but the individual monthly details were not. Simulations were performed until such time as the sum of the imputed values approached that amount. Once that iteration was found, the values found on that run were the basis for replacing not only the missing months but also the rest of the missing items.

An example of the first six months (with the imputed values from the missing March 2006 and June 2006 data in red) in the imputation results are shown below.

head(a.out$imputations$imp1$Incoming.Examinations)

[1] 362.0000 436.0000 703.0702 490.0000 508.0000 588.1838

Once the missing values were replaced, this is a view of the dataset to be forecasted against:

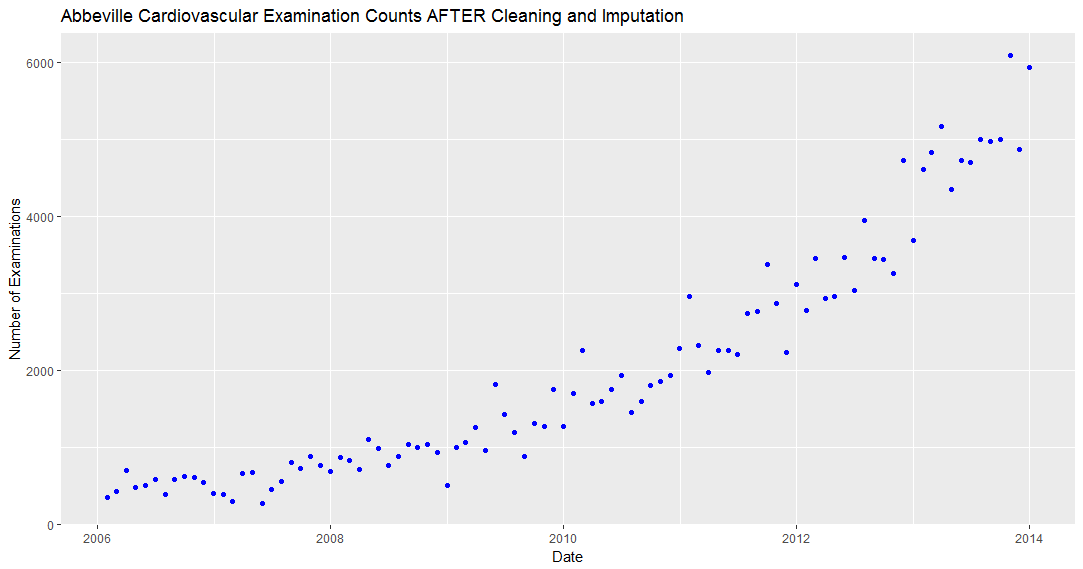


Figure 5 – Plot of cleaned data

Description of the Data Analysis Approach:

The first approach attempted was ARIMA. It suggested a solution that accounts for random variation, trend, and seasonal variation, but no auto-regression (as observed with the three resulting factors – 0,1,1). It has been differenced one time, and future results predicted on not the previous values but on a single one-period moving average.

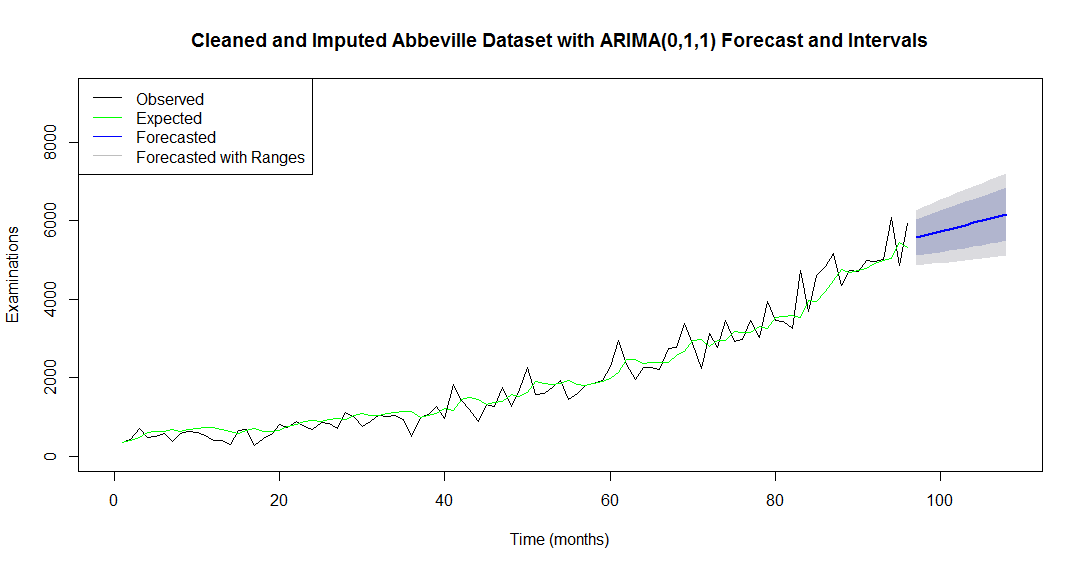


Figure 6 - ARIMA Forecast

> print(abbevilleArima)

ARIMA(0,1,1) with drift

Coefficients:

ma1 drift

-0.6725 54.2265

s.e. 0.0638 12.1710

sigma^2 estimated as 128124: log likelihood=-692.73

AIC=1391.45 AICc=1391.71 BIC=1399.11

In other words, the forecast was based on a single-term moving average of the differences plotted below:

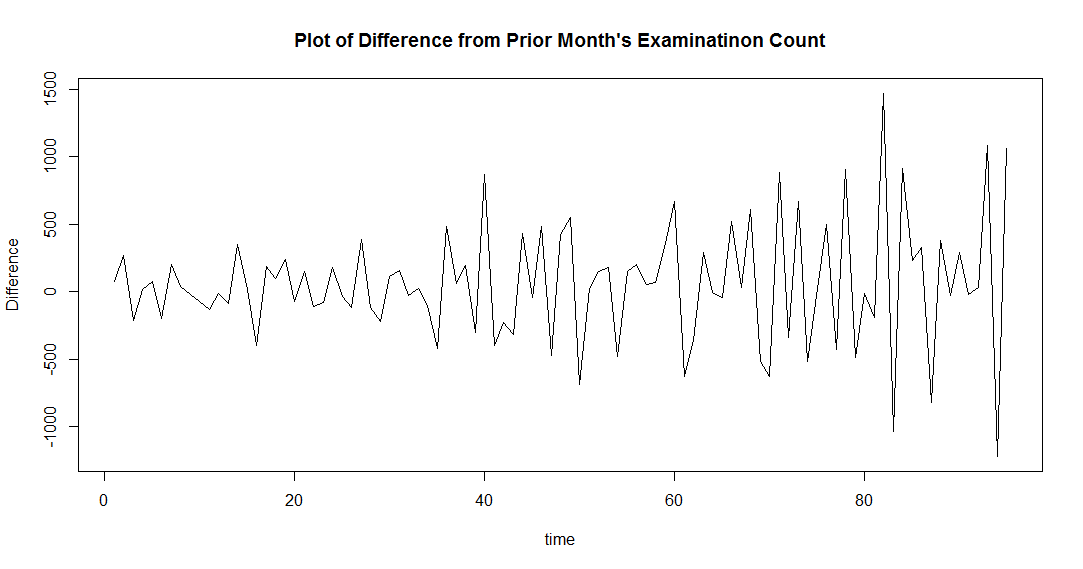


Figure 7 - Plot of Differences in Examination Count

The second approach attempted was an exponential time series (ETS) with smoothing factors. It as well suggested a solution that accounts for random variation and trend. No seasonal variation was noted as there was no gamma factor present.

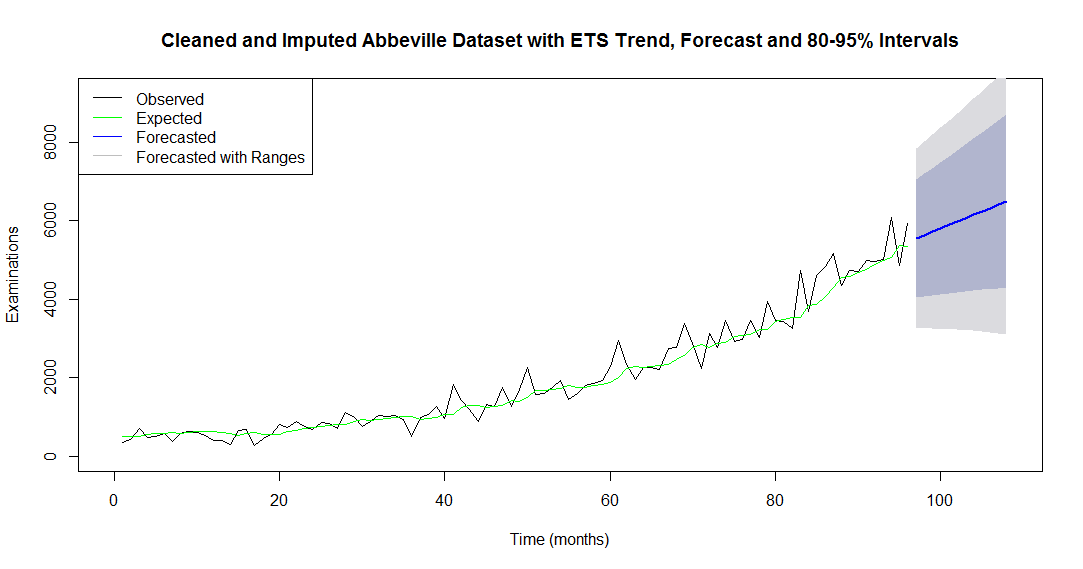


Figure 8 - ETS Forecast

print(fitETS)

ETS(M,A,N)

Call:

ets(y = abbevilleTS, model = "ZZZ")

Smoothing parameters:

alpha = 0.2009

beta = 0.0072

Initial states:

l = 495.7252

b = 21.7955

sigma: 0.2111

AIC AICc BIC

1551.914 1552.581 1564.736

The factors used to score the model comparison include:

Mean Absolute Deviation (MAD), which is the average absolute difference between estimated and actual data. Across all data points considered,

Mean Absolute Percentage Error (MAPE), which is the average percentage difference between estimated and actual data.

Mean Squared Error (MSE), which is the squared value of the difference between estimated and actual data. Root Mean Square Error (RMSE) is simply the square root of MSE.

The Akaike Information Criterion (AIC) is a common scoring process for model comparison. Its utility, to quote its Wikipedia page, is to “deal with the trade-off between the goodness of fit of the model and the complexity of the model”. A lower score indicates a better combination of fit and complexity. Further information can be found in the accompanying spreadsheet on the ‘Residual Analysis’ tab.

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | MAD | MAPE | MSE / RMSE | | AIC |
| ARIMA | 236.5 | 17.75% | 110454 | 332.3 | 1370.6 |
| ETS | 241.7 | 16.20% | 124393 | 352.7 | 1526.3 |

*Figure 8 -Measurements of model accuracy*

We are presented with results emphasizing different strengths: The ARIMA model provided the best trade-off between goodness of fit and complexity (AIC) as well as a lower magnitude of errors while attempting to fit the forecasted trend line (MSE). It also resulted in a slightly lower average absolute deviation. The ETS model realized a 2% lower overall percentage error in its results.

Referencing the plots in figures 6 and 8 respectively, the shaded confidence intervals denoting the 80 and 95 percent ranges for forecasted values are much tighter in the ARIMA model.

The ARIMA model suffered from significant normality departures in its residual Q-Q plot at the top end of the model. This is the section which utilizes the most recent set of cardiovascular exam data. If that model were chosen and used as-is, residual errors (actual – forecasted) will increase dramatically at later stages of the model. Worse so, in data that is to be forecasted. Future analyses - if undertaken - will likely see the number of cardiovascular visits needing to be log-transformed or additional differencing applied to the data points prior to the generation of a forecast.

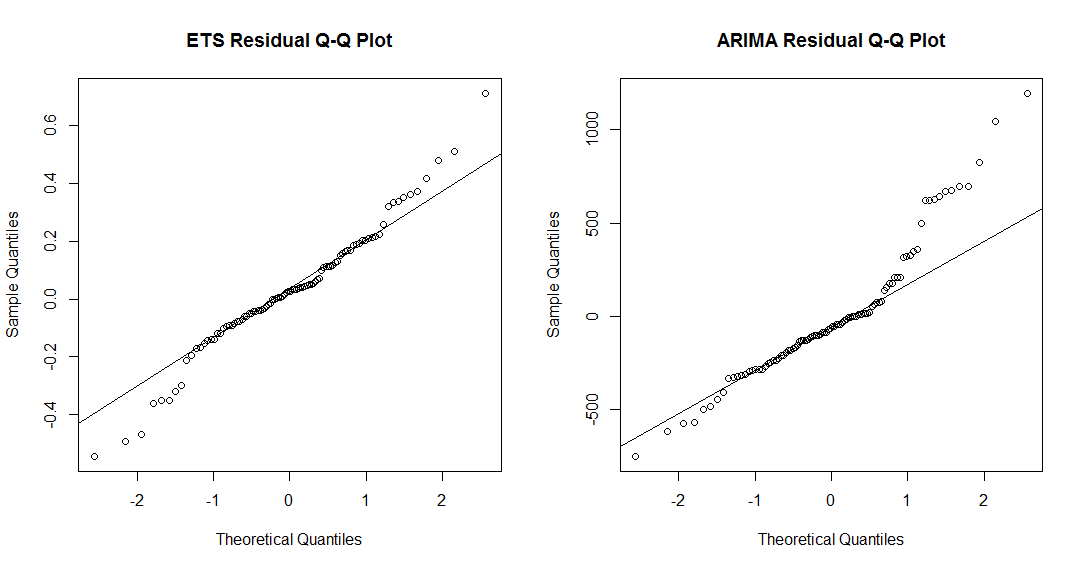


Figure 9 - Normality check of residuals between ETS and ARIMA(0,1,1) models

To further confirm the fit of the models, and to see if the normality displacement is significant, a Box Test was performed to test whether the autocorrelations are all zero. The results were non-significant in both cases, so the residuals cannot be assumed to be different from zero.

Box-Ljung test

data: abbevilleArima$residuals

X-squared = 1.0569, df = 1, p-value = 0.3039

Box-Ljung test

data: fitETS$residuals

X-squared = 1.0475, df = 1, p-value = 0.3061

As a result, the Arima model will be chosen due to the better results for all overall scoring factors. The forecast (as found on the ‘Cleaned w-Forecast Results’ tab in the spreadsheet) for the next twelve months using this model is:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| ARIMA | | | | | |
| Jan 2014 | Feb 2014 | Mar 2014 | Apr 2014 | May 2014 | Jun 2014 |
| 5570 | 5624 | 5678 | 5733 | 5787 | 5841 |
| Jul 2014 | Aug 2014 | Sep 2014 | Oct 2014 | Nov 2014 | Dec 2014 |
| 5895 | 5950 | 6004 | 6058 | 6112 | 6166 |

*Figure 10: 12-month forecast window for examinations using the ARIMA(0,1,1) model*

Summarization of key findings:

1. The issues with quality of forecast data were successfully reconciled by the use of the Amelia imputation package.
2. An ARIMA (Auto-regressive Integrated Moving Average) forecast can be applied to provide a mean percentage error of roughly 18% in the number of cardiovascular cases from month-to-month.
3. The same time series can also provide 80% and 95% confidence values for future forecasts so that worst-case ranges are known.

Ethical Implications of the Forecasting Model and its Implementation:

Context: A patient submits a disability request and if valid, is processed by the local office (LO) or is forwarded to a regional health center (HC). The patient has a direct one-on-one with a Customer Service representative to maximize Fargo’s value to the customer. It does not seem any sort of stretch to simultaneously use that contact as a data point to be used in determining how to make services more efficient.

Consent: Since the patient submitted the request for services, there was ample opportunity for Fargo to offer up information that would constitute informed consent. If the forecasting model is executed, it would behoove them to modify their consent form to encompass provision of data to third parties – if – the decision is made to outsource the analytics. Should they decide to do that in-house, the disclaimer should be more along the lines of consent to use the information provided in order to offer better service to all patients.

Reasonability: The collection (or imputation) of data over 96 months provides sufficient resources to generate reasonable forecasts of cardiac examination counts. The narrow focus, while sufficient in terms of assisting with the determination of appropriate staffing levels of cardiac-related physician services, is less valuable when attempting to predict the staffing needs for other, more generalized positions. Direct care positions such as nurses and a wide assortment of back office positions such as casework adjudicators, management, and accounting positions are not as well-served by the limited scope of the model.

Fairness: As long as HIPAA guidelines are followed in respect to data privacy and appropriateness of collection, Fargo is providing for a predominantly fair process. However, should this be successful in making services as streamlined as possible, that would place other services that Fargo provides (OBGYN and emergency care, for example) as being ‘allowed’ to be less efficient. A successful program for forecasting cardiovascular examinations should be applied to other units as well. Additionally, no mention of ethnic or socio-economic mix in Abbeville was noted in the article; dimensions of local ethnic mix as well as the mix of services provided by ethnicity should be considered in order to ensure that gains in efficiency are not realized solely by a narrow set of individuals.

Ownership: Fargo would own the dataset, analysis and insights, but would have an obligation to the greater medical community to share any discoveries with other medical service providers that could better improve health-care services to the population at large. There is a moral obligation to act upon the results of the forecasting model: A more efficient system allows either more patients to be cared for, or allows the same level of patients to be seen more quickly, incrementally improving patients’ quality of life.

Accountability: Data quality would have to be monitored much more closely in order to reduce the need to impute months’ worth of missing data. Once the process is in place, a forecasted-to-actual assessment could be conducted quarterly to focus attention on the effectivity of the model and whether any revisions need to be made to it. With better data collection, an additional analysis could be done to determine if specific days of the week have consistent spikes that could influence efficient staffing levels.

Transparency: Visibility of forecasted-to-actual results and the decisions made (hiring or lack thereof) should be made available to affected parties. Moreover, the methodology should be just as easily accessible to anyone affected.

Recommendations:

1. Institution of a pilot program to implement actions based on the results of these findings in the cardiovascular unit:
   1. Evaluate the need for changes in hiring rates for cardiovascular physicians given the rate of growth, assuming that the desire to keep the number of cases serviced per physician remains the same.
   2. HR to develop a plan to meet physician hiring rates to satisfy the targets presented in the 12-month forecast.
2. Institution of a committee to meet – at maximum – on a quarterly basis with consideration of the following agenda as the pilot expands:
   1. Evaluation of the continued effectivity of the model with the goal of determining how to best improve its accuracy. For example: What steps need to be taken to improve data entry quality so that less data imputation will be necessary.
   2. Establishment of timeframes to expand the process to include other direct-care services that the hospital provides.
   3. Establishment of timeframes to expand the process to other Fargo locations and a reporting of the basis on which those decisions are made.
   4. Evaluation of whether the model reduces the number of fines paid out as a result of increased service times. Is the overall days-to-service decreasing? A reported projection on yearly cost savings to Fargo should be the deliverable here.
   5. Evaluation of whether the model reduces the number of times patients were sent to out-of-network clinics. This could be delivered with the financial analysis in part d.
   6. Evaluation of whether the model - when executed – influences service times/patient as well as patient sentiment.
   7. Establishment of a process to make the findings available to other institutions and affected parties once out of the pilot stage.
   8. Establishment of a process to ensure that no bias is introduced by the actions of the committee. Access to services should remain equitable for all.