**JENNIE MAZE LIMITED**

***Enhancing Call Center Performance Using Predictive Analytics***

Team 8

Katherine Broderick, Antonio Jimenez, Nate Gentry,

Anna Nolan, John Power

**Introduction: Background and Business Problem**

Jennie Maze Limited is a student loan lender that has seen its revenue and market share plummet over the past seven years. The company who was dominating the private loan-lending industry in 2007 with approximately 37% market share ended up at the bottom of the four leading lenders in 2013. The leadership reviewed several potential causes for this decline and has eliminated Jennie Maze’s product offerings as the cause. As such, they are pursuing a strategy to improve customer service, based on the assumption that above average wait times and abandoned calls are responsible for their dropping market share. As part of this strategy, Jennie Maze is seeking to better understand the demand for their call centers to better staff and better meet their customers’ needs. Our goal is to develop a prediction engine that can forecast demand for Jennie Maze’s Toledo, Ohio call center, based on the data set provided by Jennie Maze, which includes data on Answered and Unanswered calls from January 2002 to August 2013.

We focused our analysis on time series models, which use time as the primary predictor of the response variable which, in this case, is call volume. In the following pages, we will discuss the techniques we employed to explore and clean the data and prepare it for analysis; the modeling approaches we used to perform our analysis; and our recommendations for a predictive model for Jennie Maze.

**Data Structure**

The data set provided by Jennie Maze Limited regarding their Toledo, Ohio call center was delivered in two forms: individual call-level data on all calls received by Jennie Maze in August of 2013 and month-level data on answered and unanswered calls from January 2002 to July 2013. Given our objective--to forecast call volume--we looked at this data as a time series, with time operating as the independent variable and call volume (answered + unanswered calls) as the dependent variable.

As our exploratory analysis (Exhibit 1) demonstrates, the data has a level, trend, and seasonality. Based on an autoregression analysis, we confirmed, by looking at the autocorrelation, that Total Calls operates on an annual cycle (see Exhibit 2). Similar annual trends were seen in Answered Calls and Unanswered Calls, though the seasonality was much stronger in Answered Calls.

As the exploratory analysis demonstrated, Total Calls, Answered Calls, and Unanswered Calls all follow an upward trend over time. The exact nature of that trend will be discussed in further detail in our analysis.

**Missing Data**

In all, there were 10 values missing, through omission or entry error, representing 3.57% of the data. Within these 10, six were for Unanswered Calls, representing 4.29% of Unanswered Calls. Four of the missing values were for Answered Calls, representing 2.86% of Answered Calls. The client also explained that the data for September 2003 was missing completely due to Toledo’s center being closed for renovation. Because we only had data for 140 periods and given that three of the months for which data was missing were September, which we considered an important month as it is likely the month many students are beginning school and tuition payments are coming due, we were inclined to impute the data for the missing values, rather than disregarding the missing values. We considered interpolation but felt that we did not have enough data for each month, particularly given the missing values, to reliably interpolate. Before deciding to impute, however, we met with the client to discuss some missing values and discovered that most of the missing values were missing due to human error that seemed anomalous, rather than a systematic issue or an issue related to the nature of the data. Because the missing data did not seem related to the nature of the data, i.e. there was no conceptual relationship between the fact that certain data points were missing and the nature of the time series call center data, we decided it was safe to impute the missing data points.

To perform the imputation, we considered Answered Calls and Unanswered Calls separately, rather than trying to impute values for Total Calls, the sum of Answered and Unanswered. Our concern about imputing a sum for Total Calls was that it would obfuscate information inherent in the primary data sets of Answered and Unanswered calls, leaving us with a less reliable imputed value and, moreover, it would prohibit us from using true values in the case when only one value--Answered or Unanswered--was missing for a given month.

For imputing Answered calls, we used three regression models. We relied heavily on testing residuals of multiple models to achieve the best regression. Two models used a logarithmic transformation on the response variable. The first model incorporated quarterly seasonality using a categorical variable for a quarter and included an interaction term of a number of periods squared, in addition to a number of periods. The second model used only the number of periods and quarter. The third model, which did not have a transformed response variable, used a number of periods, a number of periods squared, and a quarter. These models achieved high r-squared values, between .91 and .98, which was very concerning to us at first, but as these models were used to impute data rather than to forecast and because the residual diagnostics were satisfactory--normally distributed and lacking structure--we moved forward with these models (see Exhibit 3.a-3.c). To impute the final value for each missing Answered Calls, we calculated a simple average of the values generated by the three models.

We followed a similar technique to impute the values for Unanswered Calls. Rather than a logarithmic transformation, however, we looked at the reciprocal of Unanswered Calls, which seemed to remedy a light heteroskedastic trend in the data. In our first model, we used periods and a number of periods squared. In the second model, our response variable was 1/Unanswered Calls squared and our independent variables were a number of periods, a number of periods squared, and a quarter. The third model used 1/Log[Unanswered Calls] as the response variable and used a quarter, a number of periods, and a number of periods squared as the independent variables. These models again produced high r-squared values (.91, .97, and .96, respectively) but, as these were not predictive models, we were comfortable with the clear sign of overfitting. Additionally, the residual diagnostics were satisfactory. As with Answered Calls, we then calculated the mean of the three modeled values to generate the imputed value (for residuals, see Exhibits 3.d-3.f).

**August 2013**

August 2013 was provided to us as a coded call-level data, rather than an aggregate monthly data for Answered and Unanswered calls. Based on the information given to us by the client, we were able to calculate Answered and Unanswered call volumes. Answered and unanswered calls were determined by filtering out codes with OHIO in it. Rerouted calls were computed by filtering out the codes with $OH$ in them. While the Unanswered call volume appeared low, it was only about 1.5 standard deviations away from the mean, so the null hypothesis was rejected and calculated data points were not treated as outliers. Total call volume came up to 52,018 calls for August 2013.

**Outliers**

Based on our conversation with the client, the June 2009 Unanswered Call volume was higher than average due to a high number of calls forwarded from Maryland. Since the client believed that this event is unlikely to repeat, we considered this data point as an outlier. To avoid any adverse impact on our model, we used an imputed value for this month. Finally, to ensure there were no other outliers, we calculated z-scores for our Answered and Unanswered Calls and affirmed that no data points were more than three standard deviations away from the mean. With the addition of this imputed value, the total percentage of imputed values was 3.929% of the data.

**Testing for a Random Walk**

Before building any models, we decided to make sure that our dataset can be used to predict the future call volume. In other words, we wanted to ensure that the past month-to-month movements in call volume did not follow an unpredictable path. The solution to this dilemma was testing for a Random Walk. This hypothesis was previously applied to the stock market movements and proved that no active manager can predict the future prices as the past market movements are random, thus, unpredictable. Hence, we ran an autoregressive model to ensure our data series does not exhibit a Random Walk. An AR model showed the slope coefficient of 0.951 with an error of .025. Thus, the slope coefficient was about 2 standard errors away from 1. We were able to reject the null hypothesis with 97.5% confidence level. We determined it was sufficient enough to use the data for the forecasting and started building the models.

**Method 1: Regression Analysis**

We started by running a regression on the log of Total Calls. We used the log of the calls to account for an exponential trend line. For predictors, we used Number of Periods as our timeline as well as a month number to account for annual seasonality. We also used the last 12 months of data as a validation data (see Exhibit 4). The resulting graph showed fairly good predictability with r-squared of validation data of over 86%. As expected, RASE is lower for training data than for the validation data (see Exhibit 5).

Residuals were not normally distributed with three apparent outliers. Two of these outliers were imputed and another one was from August 2013 decoded data. We performed z-score analysis and realized that none of the values were greater than three standard deviations from the mean and, thus, we did not exclude them from the dataset. From the distribution analysis (see Exhibit 6), we can see that, on average, our model overpredicts (the mean for the residuals for the validation data is -1,456) with over half of the errors in the [-2000,-1100] range.

**Comparison to the Naive Forecast**

We further performed a comparison to the Naive Forecast by taking a lag-1 of the actual total calls. The resulting graph showed little difference among the actual values, regression values, and the naive values (see Exhibit 7).

However, residuals of the validation data for the naive prediction are not normally distributed. Though the model barely overpredicts on average (the residuals mean is ~ -100), the distribution range of the half of the errors is alarmingly large [-2700,2500] (see Exhibit 8). The non-normal distribution of residuals is also noticeable in the chart comparison between the regression model residuals and the naive prediction residuals (see Exhibit 9). The naive forecast doesn’t take the trend or seasonality into account, which is not suitable for the given dataset with such distinct patterns and trends.

**Method 2: Holt-Winter’s Method**

In the regression models, we performed the analysis on the Log(TotalCalls) to account for trend. For this time series analysis, we used the actual TotalCalls data, because the Holt-Winter exponential smoothing method takes both trend and seasonality into account.

Holt-Winter’s method works with both seasonality and trend, because it is an adaptive smoothing method, which can account for trend and seasonality to change over time. The traditional Holt-Winter’s smoothing method is a multiplicative method, not directly available in JMP. JMP’s version of Holt-Winter is an additive method, which means the seasons are accounted for by using a constant change, as opposed to the multiplicative method, which uses a percentage change.

Exhibit 11 illustrates graphical representation of our forecast using Holt-Winter’s smoothing. The forecast appears to follow a similar pattern as the historical call data, capturing both the trend and the seasonality, but we cannot assume that we have an accurate prediction, simply because the forecasts appear to make sense. To develop a better sense of how the model is actually performing, we can analyze other factors, like the r-squared value. Exhibit 12 shows that with this model we have an r-squared value of 0.9913. This means our model can explain over 99% of the variation from the mean. Often times, a high r-squared value suggests an accurate model, but a high r-squared can also indicate a model is over-fitting the data. Given this extremely high r-squared and the relatively small data set used to create the model we are concerned with the possibility of over-fitting. This model might work well at using the data to explain the historical call level data, but it might not accurately predict future demand.

Analyzing the residuals is another approach to determine the quality of this model. In a good model, we would expect a scatter plot of the residuals to look totally random, with no visible patterns, and the residuals’ distances from the mean should follow a Gaussian distribution. Exhibit 13 shows a scatterplot of the residuals, which appears random, and has no visible outliers. The quantile plot (exhibit 14) for the most part, appears to be linear, which indicates that the residuals follow a fairly normal distribution.

Finally, we can test our model using the most recent twelve months of data as validation data. Comparing the residuals to the actual values from those 12 months, we calculate an r-squared of 0.829117 for the validation data (exhibit 15). Although much lower than the r-squared for the training data, this is still a high R Squared value. As with the training set, we cannot assume a high R Squared equals a good model, especially because this validation set has only 12 data points.

**Method 3: Simple Exponential Smoothing**

To improve monthly forecasting of call volume, we explored using a simple exponential smoothing model. Since simple exponential smoothing can only provide benefit to data sets that have no trend or seasonality, we used the residual of our regression forecast against the actual historical call volume as the data to be smoothed. For the model we used a standard fixed .2 smoothing weight. The output of this simple exponential smoothing model is shown in Exhibit 16.

For exploratory purposes we also ran the exponential smoothing model with a Zero to One “best-fit” parameter to see if there would be marked improvement in model accuracy. The result of that model is shown in Exhibit 17. Since the “best-fit” weight only marginally improves the fit of the model, we decided to stick with the 0.2 weight in order to avoid overfitting and to utilize a greater portion of the data at a higher weight.

With a smoothing weight chosen and a simple exponential now built we can use it to project the residual for our next month of call data, September 2013. The model predicts that our regression will over-project call volume by 1,276-2,276 calls based on how the regression model has performed against actual call volume over the course of the data set. By plugging September 2013 into the regression model we find that it forecasts call volume to be 51,443 - 52,421 calls. We can then subtract the range of calls that the smoothing model projects as overestimates of the regression’s predicted volume to get our adjusted forecast for the month of September to be in a range of 48,981 to 51,314 calls.

Since simple exponential smoothing models require the most recent month of actual data to forecast a future month, we can only use it to improve one month of the regression’s forecast at a time. As such it will not be useful for long term demand projection but can help Jennie Maze more accurately plan staffing at the beginning of each month.

**Recommendations**

Our recommendation is to use the simple exponential smoothing model to enhance the performance of the regression. This model improved the r-squared from 0.869 produced by the regression to 0.954. Moreover, the simple exponential smoothing model also demonstrated superior root average square error and average absolute error, measures of the model’s ability to predict (for full comparison of models, see Exhibit 18). This model not only outperformed the Holt-Winters and Regression models, but it also outperformed the naive benchmark (see Exhibit 18 and Exhibit 19). We used a one-month lag to develop the naive benchmark, i.e. using the previous month’s actual calls as the prediction for the next month. As both Exhibit 18 and Exhibit 19show, neither the naive benchmark nor the Holt-Winters or unadjusted regression models outperform the simple exponential smoothing model on fit, measured by r-squared, or predictive performance, measured by RASE and AAE.

In Exhibits 19.a - 19.c, we have included the full forecast for the next 12 months where September 2013 is predicted with the regression and adjusted with simple exponential smoothing and the rest the data is predicted just using the regression. The ranges are based on the combination of the 95% confidence intervals around the forecasted errors, produced with the exponential smoothing and the 95% confidence intervals around the forecasted values for call volume. By combining the full range of confidence intervals, we can ensure we are fully accounting for the uncertainty in our models.

**Next Steps**

While we do recommend the simple exponential smoothing model as the best prediction engine for forecasting call volume at the Toledo, Ohio call center, we have several other recommendations for ways we can help Jennie Maze regain market share. Our first recommendation is to engage with us to perform call volume analysis at all of the Jennie Maze call centers so that we can develop a more holistic understanding of how demand functions. By analyzing the total demand for Jennie Maze, we can ensure that the trends and seasonality we saw in Toledo were not idiosyncratic and develop a better picture and more in-depth insights of the timing and trends of calls to Jennie Maze. We can also produce prediction engines for the other call centers, which, given the weak KPIs for most of the other call centers, we feel would be a valuable first step.

Our next recommendation is to perform a deeper analysis on call volume using more granular data. While predicting monthly call volume is certainly valuable for understanding macro trends, we feel that we can produce much more actionable forecasts with more detailed data. We would like to look at day-level call data, so we can predict the days in a month when demand is likely to be highest. For example, if the spike in demand for a month occurs over just a couple of days, it would be wasteful to boost staffing for the entire month. Ideally, we would like to use data that includes both date and time of day for each call so we can more accurately predict when call volume peaks, so we can help Jennie Maze to optimize its staffing shifts to best meet customer needs and preferences. Using these insights, we can help develop cost-benefit analysis to illuminate how Jennie Maze should staff its call centers for optimal performance, providing excellent value to the company.

Our final recommendation is to expand the scope of our analysis beyond call center performance. Jennie Maze’s goal is to halt and reverse its drop in market share. It is certainly possible that poor call center performance is the partly to blame for Jennie Maze’s weakening performance against competitors, but we would like to use rigorous market analysis and data analytics to perform a root cause analysis to identify the true cause of the loss of market share. As Exhibit 20shows, the percentage of missed calls has remained relatively constant despite a 35.9% increase in call volume from the first full calendar year (January 2002 - December 2002) to the last full calendar year of data (January 2012 - December 2012). More specifically, the median percentage missed for the first calendar year is 28.5% versus 32.1% for the last full calendar year, just a 3.5% increase in the percent of missed calls despite the 35.9% increase in total volume between those periods. Based on these observations and in conjunction with the 25% observed drop in market share over the same period, we are concerned that poor call center performance may not completely explain the downward trend in market share. As a firm, we strongly believe in analytics and, as such, we strongly suggest Jennie Maze contract with us for further analysis on its industry and itself in order to determine the true cause for the market share drop and to develop a data-driven strategy to reverse the trend.