

Forecasting Applied to a Service Industry Call Center

by

Harmeet Singh
h_singh17@hotmail.com
Health Products Research, Inc.
Somerset, NJ

Matthew E. Elam
Matthew_Elam@tamu-commerce.edu
Dept. of Industrial Engineering and Technology
Texas A&M University-Commerce

Ben D. Cranor
Ben_Cranor@tamu-commerce.edu
Dept. of Industrial Engineering and Technology
Texas A&M University-Commerce

Gary P. Moynihan
gmoynihan@eng.ua.edu
Industrial Engineering Program
University of Alabama-Tuscaloosa

Robert G. Batson
rbatson@eng.ua.edu
Industrial Engineering Program
University of Alabama-Tuscaloosa

Abstract: This paper applies forecasting to a service industry call center. The information forecasted is transaction volume for the purpose of appropriately staffing call centers. First, candidate forecasting methods are selected based on a time series plot of the call center transaction data. The candidate forecasting methods are then applied to the call center transaction data with the help of Minitab and Excel. The methodology to do this is presented in detail so that it can be followed by other service industries to forecast their call center transaction volume when their time series plots follow a similar pattern. The results are compared with each other using the mean square error. This paper concludes with a discussion of different approaches to forecast call center transaction volume.

I. Introduction

A forecast is a statement about the future. People make and use forecasts all the time, both in their jobs and in everyday life. In everyday life, they forecast answers and then make decisions based on their forecasts. Typical questions they may ask are: "Can I make it across the street before that car comes?"; "How much food and drink will I need for the party?"; "Will I get the job?"; "When should I leave to make it to class, work, the bank, the interview, ..., on time?" To make these forecasts, they may take into account two kinds of information. One relates to the current factors or conditions. The other regards past experience in a similar situation. Sometimes, they will rely more on one than the other, depending on which approach seems more relevant at the time [1].

Forecasting for business purposes involves similar approaches. In business, however, more formal methods are used to make forecasts and to assess forecast accuracy. Forecasts are the basis for budgeting and planning for capacity, sales, production, inventory, personnel, purchasing, and more. Forecasts play an important role in the planning process because they enable managers to anticipate the future so they can plan accordingly. Forecasts affect decisions and activities throughout an organization (e.g., accounting, finance, human resources, marketing, information systems, operations, and product/service design). In most of these uses of forecasts, decisions in one area have consequences in other areas. Therefore, it is important for managers in different areas to coordinate decisions. For example, marketing decisions on pricing and promotion affect demand which, in turn, generates requirements for operations.

A business that has rapidly increased its presence in the U.S. economy is the service industry. In the last half century, the growth in service workers has risen from fifty percent to nearly ninety percent [2]. This rise was caused by a strong demand for output combined with a slow growth in productivity. The result of this rise has caused an increase in service functions within the manufacturing sector [3]. Due to this rise, the economic impacts of white-collar service activities such as research, supply chain management and planning, and information management have begun to overshadow the impacts of blue-collar shop floor production.

Many service organizations market their services via telephone calls to their customers. Some examples include telecommunications companies selling telephone service, airlines providing flight reservations, and financial service companies selling credit cards. The last decade has seen rapid advances in telecommunications technology in an increasingly deregulated and competitive marketplace. While the technology has developed rapidly to deliver a disparate range of services to the consumer and to organizations, an equally important change has been the privatization and deregulation of the service providers. This has led to the emergence of new markets. These new markets are typically composed of new consumers taking up a product or service for the first time, established users changing their usage patterns, users of competing services shifting to the alternative service, and those exiting from a particular segment of the market altogether [4].

In such a competitive environment, it has been recognized that a critical factor in business success is being able to respond rapidly to customer requests. One way to do so in a telephone call center is to answer all calls immediately [5]. However, this may be an unrealistic objective, especially when demand is highly unpredictable. At the same time, the costs of staffing telephone call centers have become a substantial part of a business expense. Thus, it is essential to efficiently manage telephone call centers so that customer requests are met without excess staffing. Therefore, it is necessary to accurately forecast call volume in advance so that the necessary service requested by the customer can be provided in a timely fashion.

This paper shows how to accomplish this by first selecting candidate forecasting methods based on a time series plot of the call center transaction data. The candidate forecasting methods are then applied to the call center transaction data with the help of Minitab [6] and Excel [7]. The methodology to do this is presented in detail so that it can be followed by other service industries to forecast their call center transaction volume when their time series plots follow a similar pattern. The results are compared with each other using the mean square error (MSE). The

MSE, which is explained in more detail in section 3.1, measures the difference between actual and forecasted values. Better forecasting methods for a particular application are indicated by smaller MSE values. The MSE is a widely used and accepted tool to compare forecasting methods, which justifies its legitimacy for this application. This paper concludes with a discussion of different approaches to forecast call center transaction volume. An earlier and abbreviated version of this paper (without the ARIMA modeling) is in the 2006 Institute of Industrial Engineers Annual Conference proceedings [8].

II. Choice of Forecasting Methodology

Call centers have been a major telecom growth area during the 1990s [4]. Various forecasting techniques have been applied to determine staffing requirements at call centers. These include Box-Jenkins [9, 10], ARIMA (autoregressive integrated moving average) [11, 12, 13], and Winter's method [14].

One way to identify appropriate forecasting techniques is to construct a time series plot of the available data. A time series is a time-ordered sequence of observations taken at regular intervals over a period of time (e.g., hourly, daily, weekly, monthly, quarterly, annually). The call center data considered in this paper consists of call volume data collected hourly for one week (Monday-Friday) in each of several different months. The time series plot of this data for one week of October is in Figure 1. Figure 1 indicates seasonal patterns are the most dominant

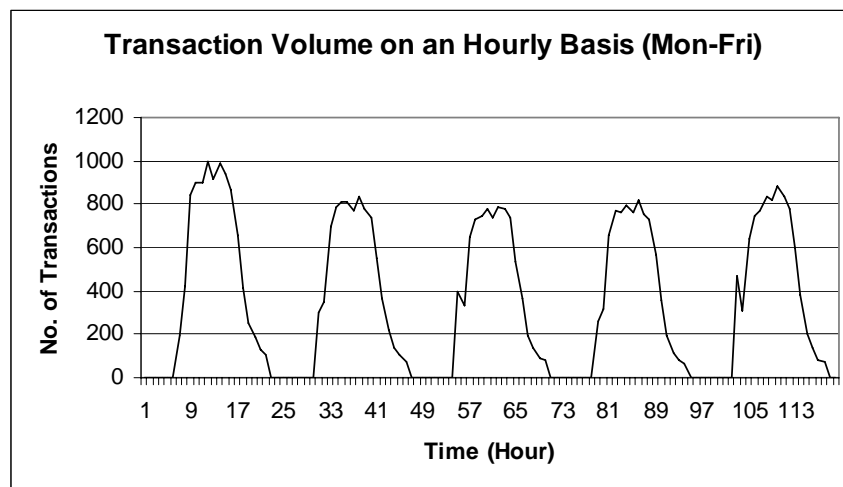


Figure 1: Time series plot of transaction volume on an hourly basis

component of this time series. Plots of the other months showed a similar pattern. Seasonal variations are periodic patterns in a time series that complete themselves within a certain period of time and repeat on a periodic basis. Therefore, those time series forecasting models which account for seasonality are candidates for modeling future call volume. These include decomposition [15], Winter's three parameter smoothing [16], and ARIMA models [17]. Also, Figure 1 indicates that the magnitude of the seasonal fluctuations does not vary with the level of the series, thus indicating the use of additive decomposition and additive Winter's methods [15].

It should be noted it may not be the case that seasonal patterns are the most dominant component of the time series for other call center data. Other components that can dominate are trend, cycle, and irregular fluctuations. Trend refers to the upward or downward movement that characterizes a time series over a period of time. Cycle refers to recurring up and down movements around trend levels. Irregular fluctuations are erratic movements in a time series that follow no recognizable or regular pattern. Such movements represent what is "left over" in a time series after trend, cycle, and seasonal variations have been taken into account.

III. Forecasting Methodology Applied

The call center data was stored in an Excel [7] spreadsheet named "Hourly_Based_Data". Some of the data from this spreadsheet is displayed in the Appendix. The columns of this spreadsheet contained data for one whole week (Monday-Friday). Thus, each column had 120 hourly-based observations. The first four days of data (Monday-Thursday, the first 96 observations) were used in the model building. The data from the last day (Friday, the last 24 observations) were used for measuring the forecasting accuracy and errors. A similar procedure of historical data usage was undertaken to forecast incoming calls to customer centers at AT&T [12]. The built forecasting models were used to generate forecasted values for each of the 24 hours of Friday. These forecasted values were compared to the actual Friday values (preserved as previously explained) using the mean square error (MSE) in order to measure the accuracy of the forecasting models.

3.1 Additive Decomposition Model

The following sequence of steps was undertaken in order to use Minitab [6] to generate forecasts using the Additive Decomposition Method. The results of these steps are shown in Figure 2.

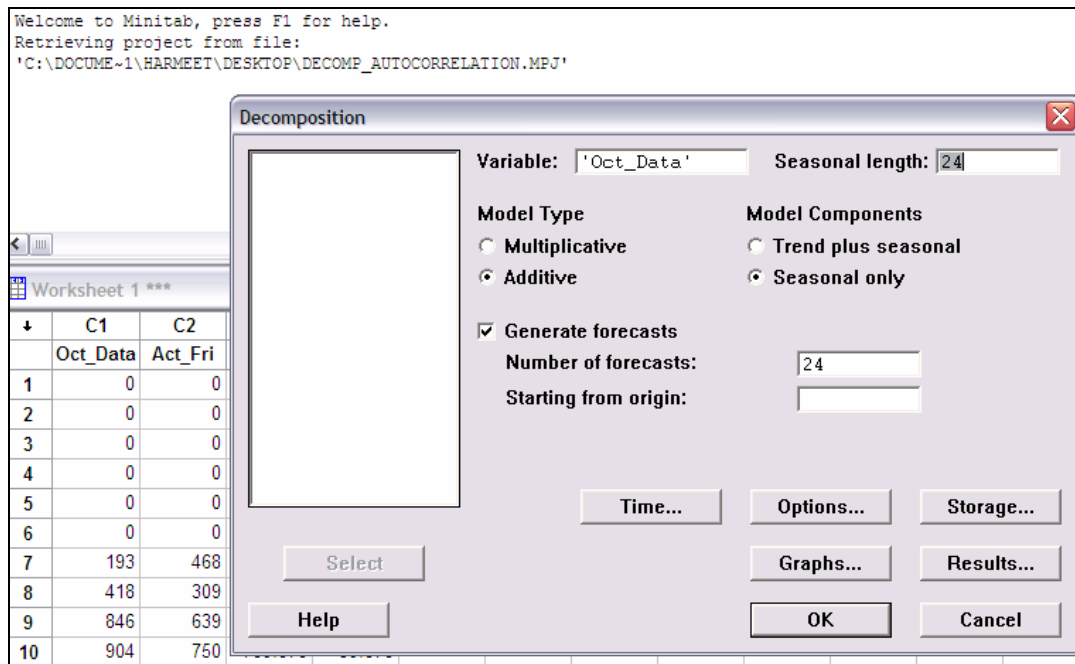


Figure 2: Selection of options in Minitab [6] to perform additive decomposition

First, the data from the first column of the Excel [7] spreadsheet "Hourly_Based_Data" was copied and pasted into a column named "Oct_Data" of the Minitab [6] worksheet. The data must be in column form. The zeros represent hours in which no calls were received. Next, the last 24 observations were cut from the "Oct_Data" column and pasted into a second column named "Act_Fri". As explained in the introduction to Section III, this data was used to evaluate forecasting errors. The menu options Stat/Time Series/Decomposition were chosen and the information shown in the "Decomposition" window in Figure 2 was entered. The value 24 in "Seasonal length" was for a 24 hour day. The value 24 in "Number of forecasts" generated forecasts for each hour of Friday. "Storage" was chosen, "Forecasts" was checked in the resulting window, and "OK" was selected in each dialog box. Forecasts for Friday were generated and stored in a column form in the Minitab [6] worksheet. This column was named "Fore_Fri". The resulting Minitab [6] worksheet is shown in Figure 3.

| | C1 | C2 | C3 | C4 | C5 | C6 | C7 | C8 | C9 | C10 | C11 | C12 | C13 | C14 | C15 | C16 | C17 | C18 |
|----|----------|---------|----------|----|----|----|----|----|----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| | Oct_Data | Act_Fri | Fore_Fri | | | | | | | | | | | | | | | |
| 1 | 0 | 0 | 15.294 | | | | | | | | | | | | | | | |
| 2 | 0 | 0 | 17.128 | | | | | | | | | | | | | | | |
| 3 | 0 | 0 | 18.211 | | | | | | | | | | | | | | | |
| 4 | 0 | 0 | 18.398 | | | | | | | | | | | | | | | |
| 5 | 0 | 0 | 18.961 | | | | | | | | | | | | | | | |
| 6 | 0 | 0 | 19.398 | | | | | | | | | | | | | | | |
| 7 | 193 | 468 | 306.357 | | | | | | | | | | | | | | | |
| 8 | 418 | 309 | 354.294 | | | | | | | | | | | | | | | |
| 9 | 846 | 639 | 682.398 | | | | | | | | | | | | | | | |
| 10 | 904 | 750 | 789.878 | | | | | | | | | | | | | | | |
| 11 | 902 | 769 | 782.128 | | | | | | | | | | | | | | | |
| 12 | 999 | 835 | 815.128 | | | | | | | | | | | | | | | |
| 13 | 916 | 816 | 782.419 | | | | | | | | | | | | | | | |
| 14 | 992 | 881 | 848.419 | | | | | | | | | | | | | | | |
| 15 | 942 | 834 | 795.419 | | | | | | | | | | | | | | | |
| 16 | 869 | 781 | 754.419 | | | | | | | | | | | | | | | |
| 17 | 657 | 604 | 562.419 | | | | | | | | | | | | | | | |
| 18 | 415 | 382 | 375.419 | | | | | | | | | | | | | | | |
| 19 | 254 | 201 | 219.315 | | | | | | | | | | | | | | | |
| 20 | 183 | 139 | 139.732 | | | | | | | | | | | | | | | |
| 21 | 127 | 82 | 109.065 | | | | | | | | | | | | | | | |
| 22 | 103 | 75 | 83.232 | | | | | | | | | | | | | | | |
| 23 | 0 | 0 | 11.753 | | | | | | | | | | | | | | | |
| 24 | 0 | 0 | 13.815 | | | | | | | | | | | | | | | |
| 25 | 0 | | | | | | | | | | | | | | | | | |

Figure 3: Minitab [6] worksheet showing the forecasted values "Fore_Fri"

In order to analyze the forecasting accuracy of this model, the mean square error (MSE) in Equation (1) was calculated using the Calc/Calculator menu options in Minitab [6].

$$(1) \text{MSE} = \left(\sum (\text{Actual} - \text{Forecast})^2 \right) / n$$

"Actual" and "Forecast" in Equation (1) correspond to the Minitab [6] columns "Act_Fri" and "Fore_Fri", respectively, in Figure 3. The value for n was 24. The resulting MSE was 1741.0.

In order for the MSE results to be valid, the difference between the actual and forecasted values (i.e., the error series) must be random. This was checked by analyzing the autocorrelation of the error series. The following sequence of steps was performed in Minitab [6] to do this. The menu options Stat/Time Series/Autocorrelation were chosen. In "Series", the name of the column containing the error series was entered. Upon selecting "OK", the autocorrelation plot in Figure

4 was generated. The x-axis represents the time lags. A lag is defined as the difference in time between an observation and a previous observation (e.g., Y_{t-k} lags Y_t by k periods) [15]. The y-axis is the magnitude of the autocorrelation between the errors at the corresponding lags. The two dashed lines are the 95% confidence limits for the correlations. Since none of the vertical spikes goes beyond the confidence limits, there is no significant autocorrelation between observations at any lag.

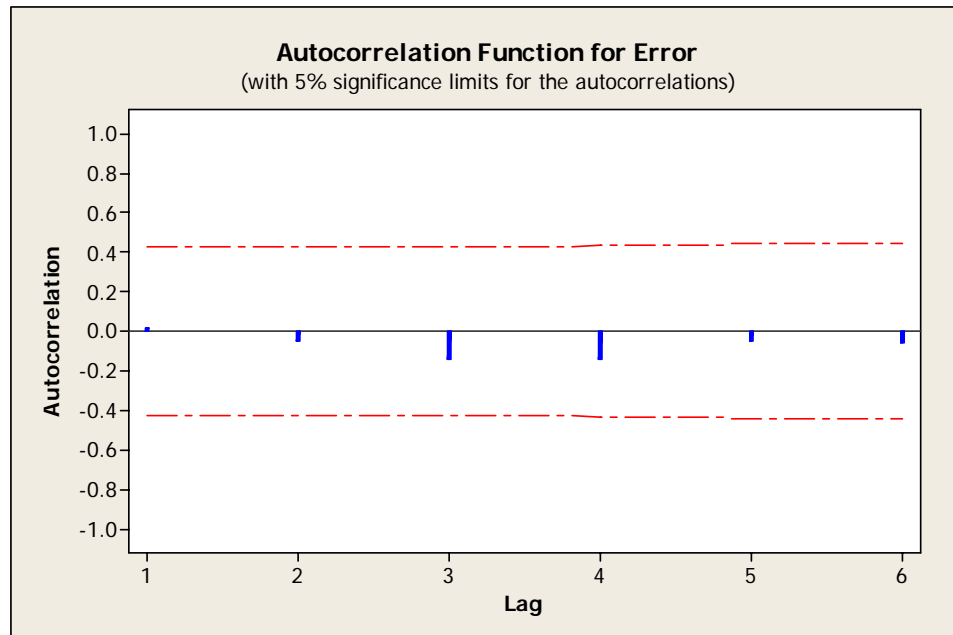


Figure 4: Autocorrelation function plot of forecasting errors from additive decomposition

3.2 Additive Winter's Method

This methodology was divided into three stages. Stage one involved initializing the method by calculating initial estimates of the level, trend, and seasonality components. In stage two, optimal values of the smoothing parameters α , β , and γ were obtained by minimizing the MSE of the data used for building the model. The smoothing parameters α , β , and γ are associated with Additive Winter's Method, not the MSE. The MSE was used to obtain optimal values for the smoothing parameters α , β , and γ . By optimal it is meant that the MSE is minimized using those values for the smoothing parameters α , β , and γ . This helps make Additive Winter's Method a better forecasting method for the particular application. The last stage was for generating forecasts for Friday based upon information obtained in the previous two stages.

3.2.1. Stage 1

The following sequence of steps was used along with Excel [7] to obtain initial estimates of the components of Winter's Additive method. The results of these steps are shown in Figure 5.

| | A | B | C | D | E | F | G | H | I | J | K | L | M |
|----|-----|-------|--------|-------|---------|----------|---------|----------|----------|----------|----------|----------|-----------|
| 1 | | | | | | | | | | | | | |
| 2 | | | | | | α | β | γ | | | | | |
| 3 | | | | s= | 24 | 0.535415 | 0 | 0.95 | | | | | |
| 4 | | | | | | | | | | | | | |
| 5 | Day | Y_t | Period | L_s | | b_s | S_i | L_t | b_t | S_t | F_t | Error | Error_Sqr |
| 6 | 1 | 0 | 1 | 405 | 0 | -2.40451 | -405 | | | -405 | | | |
| 7 | 1 | 0 | 2 | | 0 | | -405 | | | -405 | | | |
| 8 | 1 | 0 | 3 | | 0 | | -405 | | | -405 | | | |
| 9 | 1 | 0 | 4 | | 0 | | -405 | | | -405 | | | |
| 10 | 1 | 0 | 5 | | 0 | | -405 | | | -405 | | | |
| 11 | 1 | 0 | 6 | | 0 | | -405 | | | -405 | | | |
| 12 | 1 | 193 | 7 | | 4.5417 | | -212 | | | -212 | | | |
| 13 | 1 | 418 | 8 | | -2.875 | | 13 | | | 13 | | | |
| 14 | 1 | 846 | 9 | | -6.0833 | | 441 | | | 441 | | | |
| 15 | 1 | 904 | 10 | | -4.9167 | | 499 | | | 499 | | | |
| 16 | 1 | 902 | 11 | | -3.875 | | 497 | | | 497 | | | |
| 17 | 1 | 999 | 12 | | -7.75 | | 594 | | | 594 | | | |
| 18 | 1 | 916 | 13 | | -5.9583 | | 511 | | | 511 | | | |
| 19 | 1 | 992 | 14 | | -6.375 | | 587 | | | 587 | | | |
| 20 | 1 | 942 | 15 | | -6.875 | | 537 | | | 537 | | | |
| 21 | 1 | 869 | 16 | | -5.3333 | | 464 | | | 464 | | | |
| 22 | 1 | 657 | 17 | | -4.3333 | | 252 | | | 252 | | | |
| 23 | 1 | 415 | 18 | | -2.0417 | | 10 | | | 10 | | | |
| 24 | 1 | 254 | 19 | | -1.5417 | | -151 | | | -151 | | | |
| 25 | 1 | 183 | 20 | | -2.0417 | | -222 | | | -222 | | | |
| 26 | 1 | 127 | 21 | | -1.0417 | | -278 | | | -278 | | | |
| 27 | 1 | 103 | 22 | | -1.2083 | | -302 | | | -302 | | | |
| 28 | 1 | 0 | 23 | | 0 | | -405 | | | -405 | | | |
| 29 | 1 | 0 | 24 | | 0 | | -405 | 405 | -2.40451 | -405 | | | |
| 30 | 2 | 0 | 25 | | | | | 403.8829 | -2.40451 | -403.939 | -2.40451 | 2.404514 | 5.781687 |

Figure 5: Calculations of the initial estimates for level, trend, and seasonality

A new Excel [7] file was named "Winter's_Additive_Algorithm" and the first sheet in the file was named "Oct_Data". This name should not be confused with the "Oct_Data" column in Figures 2 and 3. In fact, column " Y_t " in Figure 5 is the same as the "Oct_Data" column in Figures 2 and 3. The "Day" column contains the number of the day ("1" for Monday, "2" for Tuesday, etc.). It should be recalled that there are 24 hourly observations for each day. The " Y_t " column contains the data from the first column of the Excel [7] spreadsheet "Hourly_Based_Data". The "Period" column contains the time period of the observation in which it was recorded. An initial estimate of the level component was obtained by calculating the average of the first 24 observations in column " Y_t " (i.e., the first season – Monday) using Equation (2) and storing the resulting value of 405 in column " L_s ".

$$(2) L_s = (1/s) \times (Y_1 + Y_2 + \dots + Y_s)$$

The variable "s" in Equation (2) is the length of the seasonality. In this case, s is equal to 24 hours. Equation (2) is a moving average of order s and therefore eliminates the seasonality in the data. An initial estimate of the trend component was obtained by using data from two complete seasons (i.e., the first 48 observations in column " Y_t " (2s periods) – Monday and Tuesday) along with Equation (3).

$$(3) b_s = (1/s) \times [(Y_{s+1} - Y_1)/s + (Y_{s+2} - Y_2)/s + \dots + (Y_{s+s} - Y_s)/s]$$

The resulting value of -2.40 was stored in column "b_s". Each of the terms in Equation (3) is an initial estimate of the trend over one complete season. The initial estimate of b_s was taken as the average of s such terms. Initial estimates of the 24 seasonal indices were calculated by subtracting the initial estimate of the level component in column "L_s" from each of the first 24 observations in column "Y_t". The resulting estimated seasonal indices were stored in column "S_i".

3.2.2. Stage 2

The following sequence of steps was used in order to obtain optimal values of the smoothing parameters α , β , and γ . Continuing with the "Oct_Data" spreadsheet in Figure 5, the columns beginning immediately to the right of column "S_i" were named "L_t" (represents the level of the series), "b_t" (denotes the trend), "S_t" (the seasonal component), "F_t" (the forecast), "Error", and "Error_Sqr". Also, at the top of the spreadsheet, the cells F2, G2, and H2 were named " α ", " β ", and " γ ", respectively. Arbitrary values between zero and one were entered for these three parameters in cells F3, G3, and H3, respectively, in order to initialize the Excel [7] Solver add-in. This initialization was necessary to generate optimal values of these parameters. The range of values for these three parameters is between zero and one [18]. The first value in column "L_t" was set equal to the initial estimate of L_s and was placed in row 29 to correspond to the row in which the "Period" column was equal to 24. The first value in column "b_t" was set equal to the initial estimate of b_s and was also placed in row 29. In the "S_t" column, the first 24 values were set equal to the values obtained in column "S_i". The equations for calculating L_t, b_t, S_t, and F_t in Equations (4)-(7), respectively, were entered into their respective columns starting in row 30 and continuing until row 101.

$$(4) L_t = \alpha \times (Y_t - S_{t-s}) + (1 - \alpha) \times (L_{t-1} + b_{t-1})$$

$$(5) b_t = \beta \times (L_t - L_{t-1}) + (1 - \beta) \times b_{t-1}$$

$$(6) S_t = \gamma \times (Y_t - L_t) + (1 - \gamma) \times S_{t-s}$$

$$(7) F_{t+m} = L_t + b_t \times m + S_{t-s+m}$$

In Equation (7), "F_{t+m}" is the forecast for m periods ahead. The Excel [7] versions of Equations (4)-(7) in row 30 are Equations (8)-(11), respectively:

$$(8) (\text{cell H30}) = \$F\$3 * (B30 - J6) + (1 - \$F\$3) * (H29 + I29)$$

$$(9) (\text{cell I30}) = \$G\$3 * (H30 - H29) + (1 - \$G\$3) * I29$$

$$(10) (\text{cell J30}) = \$H\$3 * (B30 - H30) + (1 - \$H\$3) * J6$$

$$(11) (\text{cell K30}) = H29 + (I29 * 1) + J6$$

The next step was the error calculation. Two types of error calculations were performed. In the first error calculation, the forecasted value F_t was subtracted from the actual value Y_t . This calculation was performed and stored in the "Error" column beginning in row 30 and ending in row 101. The second error calculation was performed in order to calculate the MSE. It involved squaring the value that was stored in the "Error" column. The resulting squared errors were stored in the "Error_Sqr" column from row 30 until row 101. The MSE was calculated by taking the average of the "Error_Sqr" column values in rows 30 to 101. This calculation was stored in row 101 of column "O".

The final step in this stage was to obtain optimal values of the smoothing parameters α , β , and γ . This was performed by utilizing Excel's [7] Solver add-in. In order to do this, "Tools" was selected from the spreadsheet toolbar and "Solver" was chosen from the drop down menu. Upon double-clicking on "Solver", a window named "Solver Parameters" opened up. In "Set Target Cell:", cell O101 (which contains the MSE of the test data) was chosen. In "Equal to:", "Min" was chosen. In "By Changing Cells:", cells F3, G3, and H3 (which contain the initial values of the smoothing parameters α , β , and γ , respectively) were chosen. In "Subject to the Constraints:", six constraints (two for each parameter) were added to bound the optimal values of the smoothing parameters between zero and one. After adding these constraints, "Solve" was selected. Solver generated new optimal values of the smoothing parameters by minimizing the MSE. It stored these new values where they were kept earlier in cells F3, G3, and H3.

3.2.3. Stage 3

Equation (7), the previously estimated values of the components L_t , b_t , and S_t , and the smoothing parameters α , β , and γ were used to generate the forecasts with "m" increasing from one to 24 in increments of one. This allowed for the generation of forecasts for every hour of Friday by utilizing data from Monday through Thursday of the same week. In the spreadsheet "Oct_Data", the forecasting equation in cell K102 (for hour one of Friday) had the same value for "m" ($m=1$) as the forecasting equation in cell K101 (for hour 24 of Thursday). All of the following cells incremented the value of "m" by one, resulting in cell K125 having the value of "m" equal to 24. After obtaining the forecasts, errors were calculated by subtracting the forecasted value from the actual value. The results were stored in the "Error" column. Squared errors were calculated by squaring the errors from the "Error" column. The MSE was calculated by taking the average of the values in cells M102 to M125. The resulting MSE was 3098.7. Minitab [6] was used to analyze the autocorrelations of the forecasting errors using an autocorrelation plot like that in Figure 4. This plot indicated no significant autocorrelation between observations at any lag.

3.3 ARIMA Model

This iterative approach involved using Minitab [6] to identify the potential models, estimate the parameters, check model adequacy, and generate forecasts. Since Figure 1 indicated the data contained a strong seasonal pattern (meaning a nonstationary time series), the first step was to take the seasonal difference of the data series and store those differences for further analysis. This work is shown in Figure 6. First, the data from the first column of the Excel [7] spreadsheet

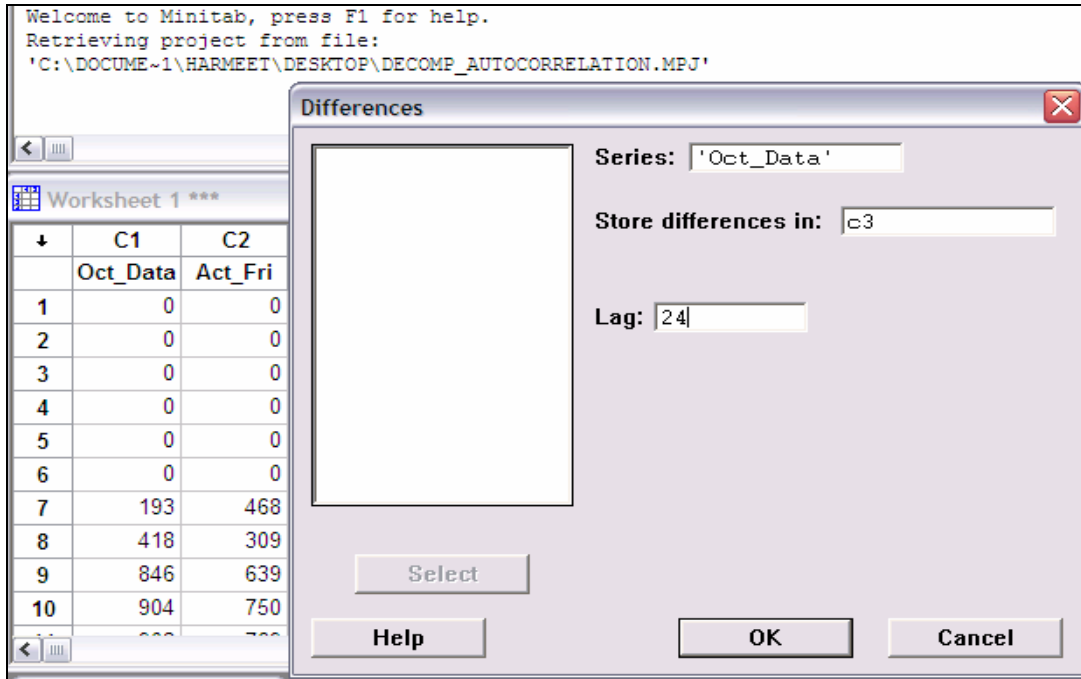


Figure 6: Selection of options in Minitab [6] to perform differencing

"Hourly_Based_Data" was copied and pasted into a column named "Oct_Data" of the Minitab [6] worksheet. Next, the last 24 observations were cut from the "Oct_Data" column and pasted into a second column named "Act_Fri". This left 96 observations in the "Oct_Data" column (observations from Monday through Thursday). As explained in the introduction to Section III, the "Act_Fri" data were used to evaluate forecasting errors. The menu options Stat/Time Series/Differences were chosen, resulting in the "Differences" window in Figure 6, which also shows the options that were chosen. The lag was 24 because the seasonal length was 24. Upon selecting "OK", the differences were calculated and stored in column "C3" of the Minitab [6] worksheet.

The next step was to study the autocorrelation function (ACF) and the partial autocorrelation function (PACF) of the differenced series in order to identify appropriate autoregressive or moving average model terms. The ACF was generated in the same manner as in Figure 4. The resulting autocorrelation plot is in Figure 7, which shows positive and significant spikes at lags one, two, and three, with subsequent positive autocorrelations that follow an exponentially decaying pattern.

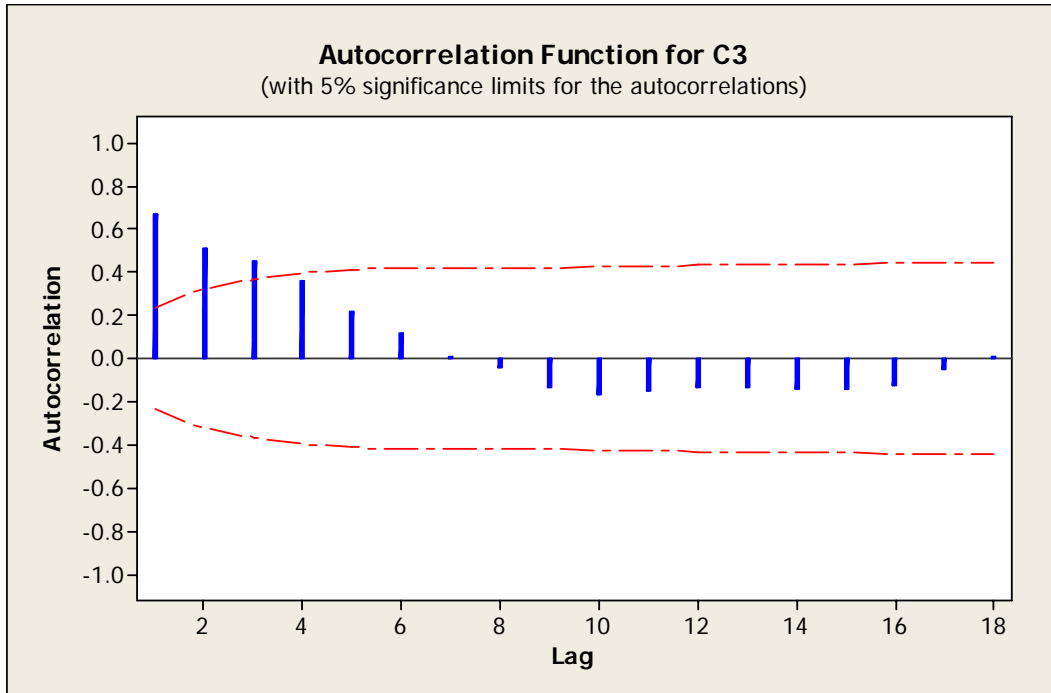


Figure 7: Plot of the ACF for the seasonally differenced series

The PACF was generated by choosing the menu options Stat/Time Series/Partial Autocorrelation. In this window, "C3" was entered for "Series". Upon selecting "OK", Minitab [6] generated a plot of the PACF with 95% confidence limits for the correlations in a graph window. This graphical output is shown in Figure 8. This plot shows one significant spike at lag

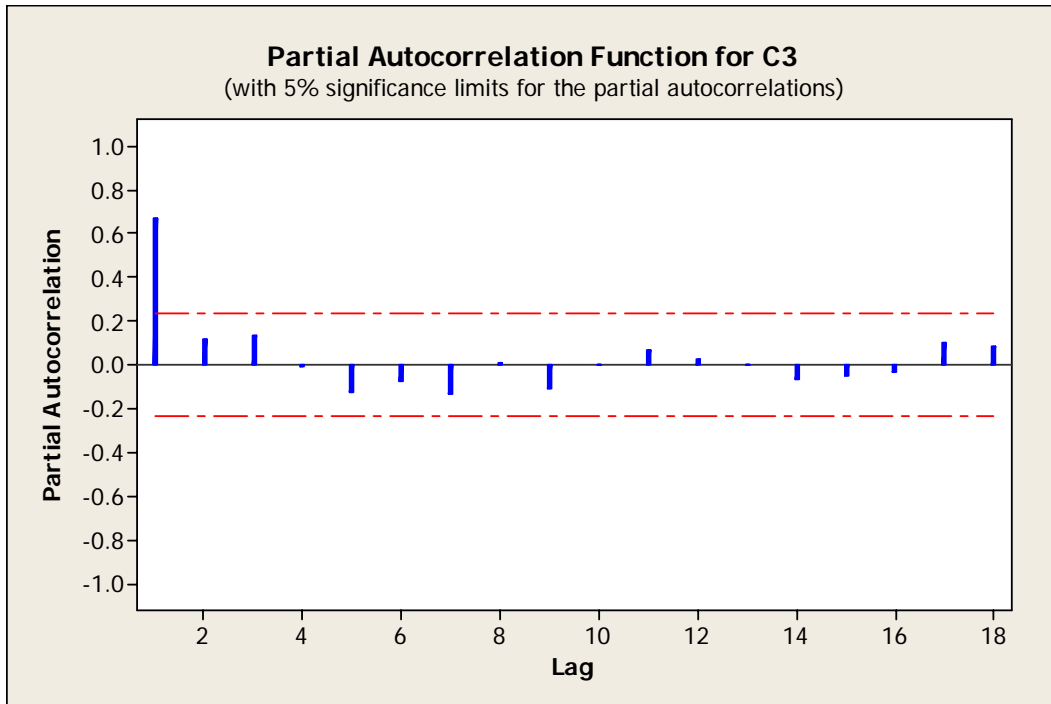


Figure 8: Plot of the PACF for the seasonally differenced series

one with the remaining spikes being insignificant. The results in Figures 7 and 8 are neither bad nor good. They provide information about the ARIMA forecasting model being built. In this case, Figures 7 and 8 together have patterns typical of an autoregressive process of order one, which means observation Y_t depends on Y_{t-1} [15].

The next steps in this modeling were fitting the ARIMA model identified in the previous steps, checking for the adequacy of the fitted model, and generating forecasts based upon the fitted model. The work performed to fit the potential ARIMA model is shown in Figure 9. The menu

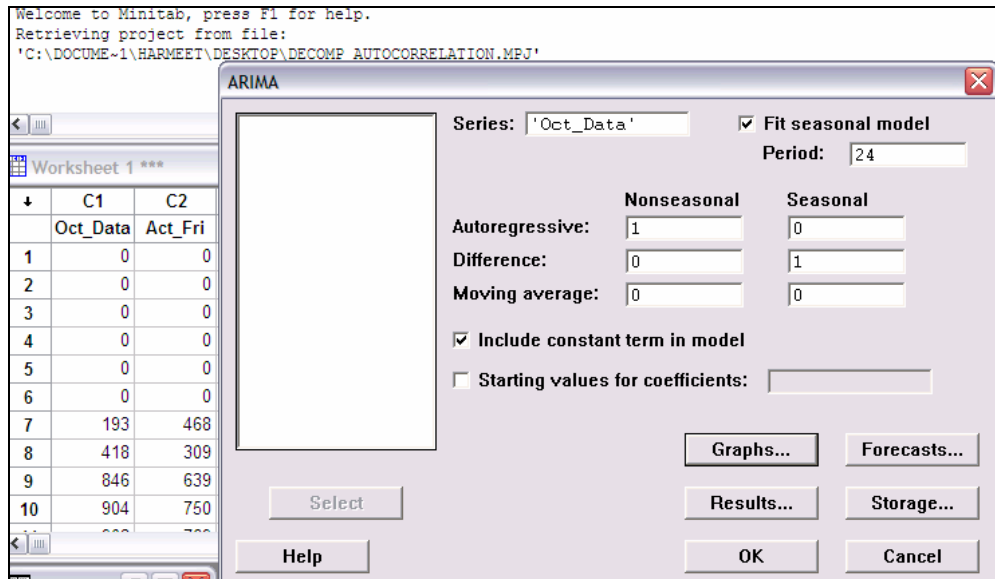


Figure 9: Selection of options in Minitab [6] to fit an ARIMA model

options Stat/Time Series/ARIMA were chosen, resulting in the "ARIMA" window in Figure 9, which also shows the chosen options. The seasonal difference was one because differencing was performed only once. In "Graphs", "ACF of residuals" and "PACF of residuals" were chosen. Upon selecting "OK" in each dialog box, the ACF and PACF plots of the residuals were generated. Since these plots had no significant spikes, it was concluded that there was no significant autocorrelation remaining in the residuals (i.e., the process was random). This was also verified by the results displayed in Minitab's [6] Session window (see Figure 10). The

| Final Estimates of Parameters | | | | | | |
|---|-------|---------|--------|-------|-------|--|
| Type | Coef | SE Coef | T | P | | |
| | AR1 | 0.7243 | 0.0818 | 8.85 | 0.000 | |
| Differencing: 0 regular, 1 seasonal of order 24 | | | | | | |
| Number of observations: Original series 96, after differencing 72 | | | | | | |
| Residuals: SS = 119603 (backforecasts excluded) | | | | | | |
| MS = 1685 DF = 71 | | | | | | |
| Modified Box-Pierce (Ljung-Box) Chi-Square statistic | | | | | | |
| Lag | 12 | 24 | 36 | 48 | | |
| Chi-Square | 6.0 | 10.4 | 12.5 | 35.9 | | |
| DF | 11 | 23 | 35 | 47 | | |
| P-Value | 0.875 | 0.989 | 1.000 | 0.882 | | |

Figure 10: Minitab [6] output generated by fitting an AR(1) model

Ljung-Box test statistics (indicated by "Chi-Square" in Figure 10) gave non-significant p-values, indicating uncorrelated errors. A p-value is a statistical result related to a test statistic. In this case, it is associated with the ARIMA modeling, not the MSE. A p-value is used to determine if a statistical hypothesis can be rejected or not. It can range in value from zero to one. Non-significant p-values usually range from slightly above 0.10 to one and indicate a failure to reject a statistical hypothesis. Significant p-values usually range from zero to slightly above 0.10 and indicate a rejection of a statistical hypothesis. In this case, the statistical hypothesis is that the errors are uncorrelated.

Also, the adequacy of the fitted ARIMA model was supported by the test of significance for the model parameter, which was AR(1) in this case (i.e., autoregressive of order one). The AR(1) parameter had a test statistic t-value of 8.85 and a corresponding p-value of zero (to the accuracy shown), indicating the associated parameter could be judged as significantly different from zero. The statistical hypothesis in this case was that the associated parameter was equal to zero.

In order to analyze the autocorrelation of the error series, the ACF of the error series was generated in the same manner as in Figure 4. Since this plot had no significant spikes, it was concluded there was no significant autocorrelation between the errors. The resulting MSE from fitting an ARIMA model to the test data was 2687.0.

IV. Results and Conclusions

The MSE values for the three forecasting methods are summarized as follows: Additive Decomposition, 1741.0; Additive Winter's, 3098.7; and ARIMA, 2687.0. For this application, additive decomposition is the best forecasting technique among the three because its MSE is the lowest.

Forecasted transaction volume may not be the only information needed to appropriately staff call centers. Depending on the organization, other issues may need to be considered. These include how the organization defines quality of customer service, statistics on lengths of calls, available budget for staffing the call center, pay earned by the call center staffers, and the staffing capacity of the call center.

Depending on how the call center data is collected and organized, several different approaches could have been used to forecast call center transaction volume. For example, instead of using the first four days of hourly call center data for one week for model building, the averages over several weeks of each hour's call center transaction volume for the first four days could have been used. One study built forecasting models based on call center data from the same source as this paper's call center data that provided six monthly, twelve weekly, and twelve hourly periods from one transaction location [19]. The nature of this data was different from the hourly data considered in this paper, which resulted in different choices of forecasting techniques such as exponential smoothing, autoregression, and moving average. By building forecasting models from data collected and organized in different ways, a service organization can determine which forecasting technique results in the lowest MSE. It then can implement the associated data collection and organization scheme in the collection of future call center transaction data.

References

- [1] Stevenson, W. and Hojati, M. *Operations Management*, McGraw-Hill Publishing Company, New York, 2004.
- [2] Brown, S.A. *Total Quality Service: How Organizations Use it to Create a Competitive Advantage*, Prentice Hall Canada Inc., Scarborough, Ontario, 1992.
- [3] Inman, R.P. *Managing the Service Economy: Prospects and Problems*, Cambridge University Press, New York, 1985.
- [4] Fildes, R. and Kumar, V. "Telecommunications Demand Forecasting – a Review," *International Journal of Forecasting*, 18(4), 489-522, 2002.
- [5] Whitt, W. "Dynamic Staffing in a Telephone Call Center Aiming to Immediately Answer All Calls," *Operations Research Letters*, 24, 205-212, 1999.
- [6] Minitab. Release 14.1, Minitab Inc., 2003.
- [7] Excel. Microsoft Office Professional Edition, Microsoft Corporation, 2003.
- [8] Singh, H., Elam, M.E., Moynihan, G.P., and Batson, R.G. "Forecasting Methods for Service Industry Call Centers," *Proceedings of the Institute of Industrial Engineers Annual Conference*, 2006.
- [9] Thompson, H.E. and Tiao, G.C. "Analysis of Telephone Data: A Case Study of Forecasting Seasonal Time Series," *The Bell Journal of Economics and Management Science*, 2(2), 515-541, 1971.
- [10] Buffa, E.S., Cosgrove, M.J., and Luce, B.J. "An Integrated Work Shift Scheduling System," *Decision Sciences*, 7(4), 620-630, 1976.

- [11] Andrews, B.H. and Cunningham, S.M. "L.L. Bean Improves Call-Center Forecasting," *Interfaces*, 25(6), 1-13, 1995.
- [12] Bianchi, L., Jarrett, J., and Hanumara, R.C. "Improving Forecasting for Telemarketing Centers by ARIMA modeling with Intervention," *International Journal of Forecasting*, 14(4), 497-504, 1998.
- [13] Tych, W., Pedregal, D.J., Young, P.C., and Davies, J. "An Unobserved Component Model for Multi-Rate Forecasting of Telephone Call Demand: The Design of a Forecasting Support System," *International Journal of Forecasting*, 18(4), 673-695, 2002.
- [14] Klungle, R. and Maluchnik, J. "Call Center Forecasting at AAA Michigan," *The Journal of Business Forecasting: Methods and Systems*, 16(4), 8-13, 1997.
- [15] Makridakis, S., Wheelwright, S.C., and Hyndman, R.J. *Forecasting: Methods and Applications*, John Wiley and Sons, Inc., New York, 1998.
- [16] Chase, C.W., Jr. "Selecting the Appropriate Forecasting Method," *The Journal of Business Forecasting: Methods and Systems*, 16(3), 23-29, 1997.
- [17] Box, G.E.P. and Jenkins, G.M. *Time Series Analysis: Forecasting and Control*, Holden-Day, Oakland, CA, 1976.
- [18] Gardner, E.S., Jr. and McKenzie, E. "Seasonal Exponential Smoothing with Damped Trends," *Management Science*, 35(3), 372-376, 1989.
- [19] Turner, C., Jr. *A Methodology for Forecasting Transaction Volumes in Large Service Organizations*, M.S. thesis, University of Alabama, Tuscaloosa, AL, 1997.

Appendix

| Day | Time Period | Oct Data | Nov Data | Dec Data | Jan Data | Feb Data | Mar Data |
|--------|-------------|----------|----------|----------|----------|----------|----------|
| Monday | 1 | 0 | 4 | 4 | 3 | 3 | 3 |
| | 2 | 0 | 4 | 4 | 2 | 3 | 3 |
| | 3 | 0 | 4 | 4 | 2 | 2 | 2 |
| | 4 | 0 | 2 | 3 | 2 | 3 | 3 |
| | 5 | 0 | 110 | 351 | 47 | 7 | 71 |
| | 6 | 0 | 214 | 191 | 414 | 344 | 333 |
| | 7 | 193 | 243 | 97 | 242 | 166 | 156 |
| | 8 | 418 | 433 | 358 | 390 | 363 | 345 |
| | 9 | 846 | 774 | 677 | 747 | 772 | 768 |
| | 10 | 904 | 887 | 758 | 887 | 887 | 831 |
| | 11 | 902 | 863 | 762 | 895 | 890 | 826 |
| | 12 | 999 | 935 | 852 | 971 | 971 | 917 |

| | | | | | | | |
|-----------|----|-----|-----|-----|-----|-----|-----|
| | 13 | 916 | 926 | 813 | 939 | 931 | 881 |
| | 14 | 992 | 958 | 841 | 913 | 909 | 884 |
| | 15 | 942 | 937 | 816 | 868 | 884 | 871 |
| | 16 | 869 | 827 | 728 | 788 | 815 | 802 |
| | 17 | 657 | 634 | 530 | 614 | 629 | 617 |
| | 18 | 415 | 395 | 327 | 425 | 437 | 396 |
| | 19 | 254 | 250 | 207 | 292 | 324 | 282 |
| | 20 | 183 | 192 | 141 | 214 | 228 | 208 |
| | 21 | 127 | 127 | 111 | 128 | 137 | 128 |
| | 22 | 103 | 95 | 84 | 101 | 99 | 97 |
| | 23 | 0 | 12 | 22 | 16 | 26 | 16 |
| | 24 | 0 | 0 | 18 | 0 | 17 | 16 |
| Tuesday | 25 | 0 | 0 | 12 | 0 | 8 | 6 |
| | 26 | 0 | 0 | 1 | 3 | 2 | 2 |
| | 27 | 0 | 1 | 3 | 6 | 3 | 3 |
| | 28 | 0 | 2 | 3 | 5 | 4 | 3 |
| | 29 | 0 | 6 | 7 | 40 | 70 | 6 |
| | 30 | 0 | 40 | 199 | 265 | 139 | 94 |
| | 31 | 302 | 318 | 95 | 207 | 177 | 200 |
| | 32 | 349 | 339 | 350 | 325 | 303 | 329 |
| | 33 | 700 | 648 | 543 | 722 | 672 | 659 |
| | 34 | 786 | 751 | 602 | 769 | 765 | 766 |
| | 35 | 809 | 761 | 633 | 805 | 761 | 777 |
| | 36 | 813 | 805 | 680 | 859 | 828 | 791 |
| | 37 | 773 | 769 | 665 | 875 | 811 | 752 |
| | 38 | 839 | 807 | 702 | 836 | 786 | 774 |
| | 39 | 777 | 794 | 689 | 860 | 840 | 789 |
| | 40 | 741 | 730 | 634 | 774 | 755 | 737 |
| | 41 | 553 | 531 | 447 | 572 | 551 | 555 |
| | 42 | 366 | 329 | 286 | 368 | 368 | 344 |
| | 43 | 217 | 205 | 169 | 230 | 254 | 224 |
| | 44 | 134 | 142 | 114 | 158 | 164 | 159 |
| | 45 | 102 | 96 | 84 | 95 | 93 | 91 |
| | 46 | 74 | 80 | 60 | 71 | 73 | 74 |
| | 47 | 0 | 12 | 14 | 12 | 15 | 13 |
| | 48 | 0 | 12 | 11 | 0 | 19 | 15 |
| Wednesday | 49 | 0 | 7 | 9 | 0 | 11 | 8 |
| | 50 | 0 | 0 | 2 | 1 | 2 | 2 |
| | 51 | 0 | 1 | 2 | 2 | 3 | 3 |
| | 52 | 0 | 2 | 3 | 2 | 3 | 3 |

| | | | | | | | |
|----------|----|-----|-----|-----|-----|-----|-----|
| | 53 | 0 | 109 | 6 | 63 | 100 | 86 |
| | 54 | 0 | 128 | 158 | 287 | 177 | 164 |
| | 55 | 398 | 213 | 224 | 232 | 178 | 222 |
| | 56 | 334 | 299 | 310 | 313 | 290 | 304 |
| | 57 | 651 | 594 | 602 | 635 | 638 | 560 |
| | 58 | 731 | 711 | 622 | 723 | 745 | 670 |
| | 59 | 743 | 726 | 630 | 751 | 716 | 658 |
| | 60 | 780 | 760 | 674 | 806 | 778 | 731 |
| | 61 | 735 | 732 | 642 | 790 | 752 | 700 |
| | 62 | 789 | 742 | 697 | 819 | 767 | 718 |
| | 63 | 775 | 754 | 706 | 826 | 769 | 735 |
| | 64 | 734 | 689 | 650 | 769 | 692 | 683 |
| | 65 | 533 | 511 | 453 | 560 | 501 | 499 |
| | 66 | 365 | 332 | 300 | 375 | 334 | 315 |
| | 67 | 196 | 191 | 176 | 242 | 222 | 228 |
| | 68 | 134 | 127 | 125 | 174 | 156 | 161 |
| | 69 | 93 | 90 | 84 | 106 | 88 | 91 |
| | 70 | 80 | 66 | 72 | 82 | 82 | 78 |
| | 71 | 0 | 9 | 12 | 13 | 13 | 21 |
| | 72 | 0 | 0 | 36 | 3 | 13 | 11 |
| Thursday | 73 | 0 | 0 | 47 | 0 | 7 | 6 |
| | 74 | 0 | 0 | 4 | 1 | 1 | 1 |
| | 75 | 0 | 1 | 13 | 3 | 2 | 2 |
| | 76 | 0 | 2 | 14 | 2 | 2 | 3 |
| | 77 | 0 | 45 | 16 | 5 | 6 | 34 |
| | 78 | 0 | 31 | 207 | 164 | 167 | 164 |
| | 79 | 259 | 163 | 113 | 145 | 133 | 126 |
| | 80 | 317 | 246 | 266 | 281 | 276 | 274 |
| | 81 | 660 | 480 | 510 | 601 | 580 | 641 |
| | 82 | 767 | 547 | 628 | 687 | 690 | 740 |
| | 83 | 759 | 554 | 607 | 668 | 659 | 729 |
| | 84 | 792 | 571 | 662 | 751 | 733 | 783 |
| | 85 | 766 | 557 | 640 | 745 | 734 | 767 |
| | 86 | 817 | 584 | 673 | 769 | 726 | 771 |
| | 87 | 755 | 425 | 649 | 763 | 708 | 768 |
| | 88 | 730 | 476 | 588 | 698 | 642 | 692 |
| | 89 | 566 | 273 | 423 | 549 | 497 | 551 |
| | 90 | 360 | 166 | 268 | 367 | 317 | 370 |
| | 91 | 193 | 119 | 166 | 238 | 212 | 259 |
| | 92 | 115 | 74 | 109 | 153 | 145 | 168 |

| | | | | | | | |
|--------|-----|-----|-----|-----|-----|-----|-----|
| | 93 | 82 | 50 | 78 | 99 | 90 | 98 |
| | 94 | 68 | 50 | 67 | 68 | 63 | 72 |
| | 95 | 0 | 6 | 11 | 12 | 13 | 23 |
| | 96 | 0 | 0 | 9 | 15 | 12 | 31 |
| Friday | 97 | 0 | 1 | 7 | 7 | 5 | 0 |
| | 98 | 0 | 2 | 1 | 1 | 1 | 3 |
| | 99 | 0 | 1 | 2 | 2 | 2 | 3 |
| | 100 | 0 | 3 | 3 | 2 | 3 | 3 |
| | 101 | 0 | 9 | 6 | 7 | 6 | 33 |
| | 102 | 0 | 60 | 180 | 413 | 280 | 283 |
| | 103 | 468 | 465 | 184 | 326 | 365 | 161 |
| | 104 | 309 | 425 | 247 | 338 | 254 | 359 |
| | 105 | 639 | 615 | 474 | 652 | 587 | 603 |
| | 106 | 750 | 685 | 577 | 755 | 675 | 699 |
| | 107 | 769 | 659 | 584 | 758 | 710 | 679 |
| | 108 | 835 | 752 | 611 | 806 | 732 | 731 |
| | 109 | 816 | 749 | 611 | 838 | 726 | 728 |
| | 110 | 881 | 754 | 609 | 842 | 740 | 729 |
| | 111 | 834 | 734 | 589 | 849 | 750 | 701 |
| | 112 | 781 | 688 | 545 | 794 | 693 | 697 |
| | 113 | 604 | 509 | 403 | 576 | 506 | 558 |
| | 114 | 382 | 327 | 242 | 359 | 325 | 389 |
| | 115 | 201 | 185 | 147 | 232 | 202 | 255 |
| | 116 | 139 | 135 | 99 | 147 | 141 | 154 |
| | 117 | 82 | 94 | 73 | 71 | 88 | 81 |
| | 118 | 75 | 75 | 60 | 56 | 64 | 63 |
| | 119 | 0 | 10 | 11 | 9 | 11 | 14 |
| | 120 | 0 | 0 | 7 | 0 | 10 | 4 |