Seasonal Time Series Model Comparison for Nonstationary Passenger Flight Data

By

Theresa Moore

Submitted in Partial Fulfillment of the Requirements

For the Degree of

Master of Science

In the

Mathematics

Program

Youngstown State University

December, 2007

Seasonal Time Series Model Comparison for

Nonstationary Passenger Flight Data

Theresa Moore

I hereby release this thesis to the public. I understand that this thesis will be made available from the OhioLINK ETD Center and the Maag Library Circulation Desk for public access. I also authorize the University or other individuals to make copies of this thesis as needed for scholarly research.

Signature:

Theresa Moore, Student Date

Approvals:

Dr. G. Jay Kerns, Thesis Advisor Date

Dr. G. Andy Chang, Committee Member Date

Dr. Gary Stanek, Committee Member Date

Peter J. Kasvinsky, Dean of School of Graduate Studies & Research Date

Abstract

The objective of this paper is to analyze the number of passengers flying a sample of three airlines before and after 9/11 to discover whether there has been a recovery. The three airlines were modeled using simple linear regression and time series analysis. Dummy variables and trigonometric functions were used to mimic the seasonal variation and additive decomposition was used to remove the seasonal component and model the trend. The additive decomposition quadratic models were deemed the best fits. From the quadratic models is concluded that the three airlines chosen for this paper have recovered from the effects of 9/11.

Table of Contents

List of Tables

List of Output

List of Figures

1. Introduction

The United States airline industry suffered a tragic blow due to the terrorist attacks on September 11, 2001. Not only did the temporary shut-down of the commercial aviation system contribute to a decline in the number of passengers on domestic flights, but the attacks caused many passengers to reduce or avoid air travel, in fear of the risk associated with flying. Similarly, following September 11, 2001, many businesses temporarily froze all but the most essential travel for their employees.[1] This tragic date has affected the airline industry in many different ways. Have the effects on the number of passengers traveling domestically endured? The purpose of this paper is to analyze the number of people traveling on domestic flights before and after September 11, 2001 and examine whether or not the numbers have recovered.

According to a brief issued by the Bureau of Transportation Statistics, "In the August preceding 9/11, the airline industry experienced what was then a record high in the number of airline passengers for a given month when 65.4 million travelers took to the air. After 9/11, that number trailed off dramatically, and it took nearly 3 years, until July 2004, for the industry to match and finally surpass the pre 9/11 levels".[2] This paper investigates the statistical models of number of passengers on domestic U.S. flights pre 9/11 and Post 9/11 to show if the numbers are in fact at or above that of the pre 9/11 years.

The question here is important because the American public needs to know if the terrorist attacks are still affecting the way we live our lives. If so, then the terrorists are winning and the public needs to change their perspective and not be terrorized. If not, then we as Americans have shown the world that nothing can keep us down.

---

[1] Ito, H., and D. Lee (2003): "Assessing the Impact of the September 11 Terrorist Attacks on U.S. Airline Demand," http://www.brown.edu/Departments/Economics/Papers/2003/2003-16_paper.pdf

[2] Bureau of Transportation Statistics (2005): "Airline Travel Since 9/11", Brief #13 issued Dec 2005.

2. Overview

The data in this paper were collected from the Bureau of Transportation. The data were collected on three of the top airlines as indicated by the amount of market share that these airlines represented at time of collection. The data spanned a ten-year period beginning January of 1996 through October of 2006, which was the latest data available at the time of collection. The three airlines chosen for this paper are, Continental airlines with 7.6% market share, Delta airlines with 11.5% market share and American airlines with 15.6% market share.

The interest is the comparison of the pre 9/11 data to the post 9/11 data. Thus, two timelines are used for each airline. The first timeline begins January 1996 and ends August 2001. The second begins October 2001, the month following September 11, and ends October 2006.

The six timelines were modeled using simple linear regression, time series regression using dummy variables to model seasonal pattern, then trigonometric functions to model seasonal pattern, and finally additive decomposition to remove seasonal pattern. The best models were chosen and different aspects were compared.

For the simple linear regression models, actual post 9/11 values were compared to confidence and prediction intervals. With the dummy variable models, regression coefficient intervals were formed and compared between pre and post data. Also, fitted values of the post 9/11 data were compared to that of August 2001. For the additive decomposition models, the regression equations were compared to see when, if ever, they would be equal.

The findings of this paper are that the three airlines sampled have recovered to their pre 9/11 passenger numbers.

2.1 Acknowledgements

I would like to thank Dr. Kerns for all his help and encouragement.

3. Methodology

3.1 Data Collection

The data for the analysis were obtained from the Bureau of Transportation Statistics. Data were downloaded from yearly Tables entitled T-100 Domestic Market (U. S. Carriers) for the years January 1996 through October 2006, which were the latest data available at the time. These data cited passengers for individual flights within each month of the year. The data were downloaded to Microsoft Access databases per year. Three individual airlines were chosen for the analysis. The three airlines chosen for paper are: Continental airlines with 7.6% market share, Delta airlines with 11.5% market share, and American airlines with 15.6% market share. These three airlines were chosen randomly from among the top ten airlines in reference to market share; at the time the data were collected. For each of the three airlines, the data were filtered from the yearly databases and imported into separate Microsoft Access Databases. Consequently, there were three Microsoft Access databases, each containing all ten years of data for its respective airline. Each MS Access database was imported into SPSS. The number of passengers were totaled by month for each year and transferred to separate Microsoft Excel files for each airline. The statistical program R was used to analyze the data from the Microsoft Excel files.

3.2 Models

In this section is explained the models used to analyze the data.

3.2.1 Simple Linear Regression

The first method used to analyze the data was a simple linear regression model, with number of passengers as the response variable versus months as the predictor variable. The data were plotted in a scatter diagram and it was tentatively decided that there was an approximate linear association between the two variables.

The regression model is of the form:

$$y_i = \beta_0 + \beta_1 x + \varepsilon_i \qquad i = 1,2,\ldots,n \tag{1}$$

where

β_0 = y-intercept

β_1 = slope of the line

ε_i = error term.

In a regression model we make four assumptions:

1. At any given value of the independent variable, the population of potential error terms has a mean equal to zero.
2. At any given value of the independent variable, the population of potential error terms has a variance not dependent on the independent variable. That is the different populations of potential error terms corresponding to different values of the independent variables have equal variances. This is the constant variance assumption.
3. At any given value of the independent variable, the population of potential error terms is normally distributed. This is the normality assumption.
4. The independence assumption states that at any given value of the independent variable, in the population of potential error terms, each is independent of the other.

The regression assumptions very rarely hold exactly in any problem. Regression results are not extremely sensitive to mild violations of the assumptions.

The true values of the regression parameters B_0, B_1, are not known. Therefore, using the n observed values of the indicator variable x,

$x_1, x_2, \ldots x_n$

and the n observed values of the response variable y,

$y_1, y_2, \ldots, y_n$

the least squares point estimates b_0 and b_1 of β_0 and β_1 are calculated with the below equations. These equations are derived from calculus techniques and have been shown to minimize the value of the sum of squared residuals.

The least squares point estimate of the slope β_1 is:

$$b_1 = \frac{SS_{xy}}{SS_{xx}} \qquad (2)$$

where

$$SS_{xy} = \sum (x_i - \bar{x})(y_i - \bar{y}) = \sum x_i y_i - \frac{(\sum x_i)(\sum y_i)}{n} \qquad (3)$$

and

$$SS_{xx} = \sum (x_i - \bar{x})^2 = \sum x_i^2 - \frac{(\sum x_i)^2}{n} \qquad (4)$$

and the least squares point estimate of the y-intercept β_0 is:

$$b_0 = \bar{y} - b_1 \bar{x} \qquad (5)$$

where

$$\bar{y} = \frac{\sum y_i}{n} \quad \text{and} \quad \bar{x} = \frac{\sum x_i}{n}. \qquad (6)$$

The point estimates of the regression model become $\hat{y} = b_0 + b_1 x$, which is the estimate of the mean value of the dependent variable when the value of the independent value is x_0, and $\hat{y}$ may be used to predict future observations.

When an individual value is predicted, the error term is predicted to be zero due to the assumption that the population of all error terms is normally distributed with a mean equal to zero. Since it is also assumed that successive error terms are independent with distribution symmetric about 0, each error term has a 50% chance of being positive, and the same chance of being negative. Thus it is reasonable to predict any particular error term to be zero.

The point estimate $\hat{y}$ is very rarely exactly the mean value of *y* when the *x* value is equal to x_0 or a particular individual value of *y* when the *x* value is equal to x_0. Therefore, confidence intervals for the mean value of *y* and prediction intervals for an individual value of *y* are calculated. To find these intervals the distance value is used. The distance value is a measure of the distance between the value x_0 and $\bar{x}$, and is calculated as:

$$\text{Distance Value } (DV) = \frac{1}{n} + \frac{(x_0 - \bar{x})^2}{SS_{xx}}. \qquad (7)$$

If regression assumptions hold, then the population of all possible $\hat{y}$ values is normally distributed with mean $\mu_{y|x_0}$ and standard deviation:

$$\sigma_{\hat{y}} = \sigma \sqrt{DV}. \qquad (8)$$

The point estimate of $\sigma_{\hat{y}}$ is:

$$s_{\hat{y}} = s\sqrt{DV} \tag{9}$$

this is called the standard error of the estimate $\hat{y}$. The 100(1 - α) % confidence interval for the mean value of y when the independent variable is x_0 is thus calculated:

$$\left[\hat{y} \pm t_{[\alpha/2]}^{(n-2)} s\sqrt{DV}\right] \ . \tag{10}$$

To develop the prediction interval, the prediction error y - $\hat{y}$ is considered. If the regression assumptions hold, the population of all prediction errors is normally distributed with mean 0 and standard deviation:

$$\sigma_{(y-\hat{y})} = \sigma\sqrt{1+DV} \ . \tag{11}$$

The point estimate of $\sigma_{(y-\hat{y})}$ is:

$$s_{(y-\hat{y})} = s\sqrt{1+DV} \ , \tag{12}$$

this is the standard error of the prediction error. Thus a 100(1 - σ) % prediction interval for an individual value of y when the independent variable is x_0 is:

$$\left[\hat{y} \pm t_{[\alpha/2]}^{(n-2)} s\sqrt{1+DV}\right]. \tag{13}$$

Using the confidence and prediction intervals for the pre 9/11 data, the post 9/11 data for specific months were compared to see where they fell in relation to the pre 9/11 data. Values and intervals were predicted for September 2001, July 2002, May 2003, March 2004, January 2005, November 2005, and September 2006.

In the case of Continental Airlines, the majority of the actual data fell below both confidence intervals and prediction intervals. The American Airlines actual data seemed to fall mostly above confidence intervals and within prediction intervals. The majority of Delta Airlines actual data fell below both confidence intervals and prediction intervals.

3.2.2 Dummy Variables

The second method used to analyze the data was to use a time series modeling the seasonal variation with dummy variables. The seasonal variation seemed to be constant over time, so the model used was of the form:

$$y_t = TR_t + SN_t + \varepsilon_t \tag{14}$$

where

y_t = the observed value of the time series in time period t

TR_t = the trend in time period t

SN_t = the seasonal factor in time period t

ε_t = the error term in time period t.

Assuming the error term satisfies the usual regression assumptions, the data can be represented by an average level that changes over time according to the equation:

$$\mu_t = TR_t + SN_t . \tag{15}$$

Furthermore, using estimates, the model that estimates y_t is of the form:

$$\hat{y}_t = tr_t + sn_t . \tag{16}$$

Assuming there are L seasons (months) per year, the seasonal factor SN_t, using dummy variables is expressed as:

$$SN_t = \beta_{s1} x_{s1,t} + \beta_{s2} x_{s2,t} + \ldots + \beta_{s(L-1)} x_{s(L-1),t} \tag{17}$$

where $x_{s1,t}, x_{s2,t}, \ldots, x_{s(L-1),t}$ are dummy variables defined as:

$x_{s1,t} = \{1$ if time period t is season 1, 0 otherwise

$x_{s2,t} = \{1$ if time period t is season 2, 0 otherwise

$\vdots$

$x_{s(L-1),t} = \{1$ if time period t is season L-1, 0 otherwise.

The models for each timeline were examined for possible outliers. An outlier is an observation that is separated from the rest of the data. Outliers may be influential in that they may cause aspects of the model to change substantially if removed. An observation may be an outlier with respect to its *y* value and/or its *x* value.

Models were visually assessed for possible outliers using the diagnostic plots: Normal Q-Q plot, Scale Location plot, and Residuals vs. Fitted values plot. If the plots singled out particular observations, these were held to be possible outliers.

Residual vs. Leverage plots were used to visually detect influential outlying observations. Following that, the use of the Cook's Distance measure was employed to be certain. Cook's D depends on both residuals and leverage.

A residual is defined as:

$$e = \hat{y} - y = \text{(observed value of } y - \text{predicted value of } y\text{)} \qquad (18)$$

where the predicted value of y is calculated using the least squares prediction equation:

$$\hat{y} = b_0 + b_1 x. \qquad (19)$$

The linear regression model $y = \beta_0 + \beta_1 x + \varepsilon$ implies that the error term ε is given by the equation $\varepsilon = y - (\beta_0 + \beta_1 x)$. Since $\hat{y}$ is the point estimate of $\beta_0 + \beta_1 x$, the residual $e = y - \hat{y}$ is the point estimate of the error term.

The leverage value, h_{ii}, for an observation is a distance value, and is used to calculate a prediction interval for the y value of the observation. This value is a measure of the distance between an observation's x value and the center of the experimental region. A leverage value for an observation is large if it is greater than twice the average of all of the leverage values. Twice this average is equal to $2(k+1)/n$, where k is equal to the number of independent variables.

All data, with and without potential outliers, were fitted with dummy variables. Using the adjusted R^2 statistic, the AIC statistic, and the F statistic, the best of the pre and post 9/11 models were chosen to run comparisons.

Using these models' estimated parameters, two times the standard error were added and subtracted to form regression coefficient intervals. The intervals of the pre 9/11 data were compared to the post 9/11 intervals. In the case of Continental Airlines, the majority of the post 9/11 intervals were below that of the pre 9/11 intervals. In the case of American airlines, the majority of the post 9/11 intervals overlap that of the pre 9/11 intervals. In the case of Delta Airlines, the majority of the post 9/11 intervals overlap that of the pre 9/11 intervals.

Using these same models, observations of the fitted values were made. The post 9/11 fitted values were compared to the fitted value for August 2001. In the case of Continental Airlines, the post fitted values never reached that of August 2001. In the case of American Airlines, 39 of the 58 values reached or exceeded that of August 2001. In the case of Delta Airlines, the post 9/11 fitted values never reached that of August 2001.

3.2.3 Trigonometric Functions

A second attempt to model seasonal variation was made using trigonometric functions. Trigonometric models for constant seasonal variation were used. The three different trigonometric modes used were that of two sets:

$$\begin{aligned} y_t &= \beta_0 + \beta_1 t + \beta_2 \sin\left(\frac{2\pi t}{L}\right) + \beta_3 \cos\left(\frac{2\pi t}{L}\right) + \beta_4 \sin\left(\frac{2\pi t}{L}\right) \\ &+ \beta_5 \cos\left(\frac{2\pi t}{L}\right) + \varepsilon_t \end{aligned} \tag{20}$$

three sets:

$$\begin{aligned} y_t &= \beta_0 + \beta_1 t + \beta_2 \sin\left(\frac{2\pi t}{L}\right) + \beta_3 \cos\left(\frac{2\pi t}{L}\right) + \beta_4 \sin\left(\frac{2\pi t}{L}\right) + \beta_5 \cos\left(\frac{2\pi t}{L}\right) + \\ &\beta_6 \sin\left(\frac{2\pi t}{L}\right) + \beta_7 \cos\left(\frac{2\pi t}{L}\right) + \varepsilon_t \end{aligned} \tag{21}$$

and four sets:

$$\begin{aligned} y_t &= \beta_0 + \beta_1 t + \beta_2 \sin\left(\frac{2\pi t}{L}\right) + \beta_3 \cos\left(\frac{2\pi t}{L}\right) \beta_4 \sin\left(\frac{2\pi t}{L}\right) + \beta_5 \cos\left(\frac{2\pi t}{L}\right) + \\ &\beta_6 \sin\left(\frac{2\pi t}{L}\right) + \beta_7 \cos\left(\frac{2\pi t}{L}\right) \beta_8 \sin\left(\frac{2\pi t}{L}\right) + \beta_9 \cos\left(\frac{2\pi t}{L}\right) + \varepsilon_t \end{aligned} \tag{22}$$

where, L is the number of months in a year and the sine cosine pairs model the seasonal pattern. Visual assessment of the fitted models and values of the AIC test statistic proved dummy variables to be a more useful model; therefore the trigonometric models were abandoned.

3.2.4 Additive Decomposition

The final attempt at modeling the seasonal pattern was to use additive decomposition. The data seemed to display a constant seasonal variation warranting the use of the additive decomposition model of the form:

$$y_t = TR_t + SN_t + CL_t + IR_t \tag{23}$$

where

y_t = the observed value of the time series in time period t

TR_t = the trend component in time period t

SN_t = the seasonal component in time period t

CL_t = the cyclical component in time period t

IR_t = the irregular component in time period t

To obtain point estimates tr_t, sn_t, cl_t, and ir_t of the previous components, the first step is to calculate the centered moving averages, CMA_t, which is regarded as an estimate of $TR_t + CL_t$. Since the additive decomposition model implies that $SN_t + IR_t = y_t - (TR_t + CL_t)$, the estimate $sn_t + ir_t$, of $SN_t + IR_t$ is:

$$sn_t + ir_t = y_y - (tr_t + cl_t) = y_t - CMA_t . \tag{24}$$

To obtain sn_t, the values of $sn_t + ir_t$ are grouped by like months. For each month, the average $\bar{sn}_t$ is calculated. To obtain the monthly factors, the $\bar{sn}_t$ values are normalized so they sum to zero. Normalization is accomplished by subtracting the quantity $\sum_{t=1}^{L} \bar{sn}_t \big/ L$ from each of the $\bar{sn}_t$ values. The estimate of SN_t is defined to be:

$$sn_t = \bar{sn}_t - \left(\sum_{t=1}^{L} \bar{sn}_t \Big/ L \right). \tag{25}$$

Next is to calculate the deseasonalized observation in time period t:

$$d_t = y_t - sn_t . \tag{26}$$

This step removes the seasonality from the data allowing for a better estimate of the trend. The estimate tr_t of the trend TR_t is obtained by fitting a regression equation to the deseasonalized data. The regression equations used to estimate the trend were a linear equation, $TR_t = \beta_0 + \beta_1 t$ a quadratic equation, $TR_t = \beta_0 + \beta_1 t + \beta_2 t^2$, and a cubic equation, $TR_t = \beta_0 + \beta_1 t + \beta_2 t^2 + \beta_3 t^3$.

The additive decomposition model also implies that $CL_t + IR_t = y_t - TR_t - SN_t$, thus the estimate of $CL_t + IR_t$ is computed as:

$$cl_t + ir_t = y_t - tr_t - sn_t . \tag{27}$$

In order to average out ir_t, a three-period moving average of the $cl_t + ir_t$ value is computed. Thus the estimate of CL_t is computed as:

$$cl_t = \frac{(cl_{t-1} + ir_{t-1}) + (cl_t + ir_t) + (cl_{t+1} + ir_{t+1})}{3}. \quad (28)$$

Finally, the estimate of IR_t is calculated as:

$$ir_t = (cl_t + ir_t) - cl_t. \quad (29)$$

The linear, quadratic and cubic regression models of the deseasonalized data were examined for potential outliers. Models were then made for the original data and also the data discarding the possible outliers and all models were compared. Using AIC and ANOVA, the best models were chosen and the regression equations for the pre 9/11 data were compared to that of the post 9/11 data. In the case of Continental Airlines, pre and post quadratic models were used and were equal in October 2005. In the case of American airlines the pre and post quadratic equations were used and were equal in August 2007. In the case of Delta Airlines the pre and post quadratic equations were used and were equal in August 2003.

3.2.5 Box-Cox Transformation

A Box-Cox transformation of a response variable is used to make a linear model more appropriate to the data. It can be used as an attempt to impose linearity, reduce skewness, or stabilize residual variance.

When attempting a linear fit on a dataset, an appropriate transformation of the response variable can be performed to maximize the correlation between the predictor and response variables. The Box-Cox transformation finds the maximum likelihood power transformation of the response variable in a regression model. The transformation is defined as:

$$T(y_t) = (y_t^{\lambda} - 1)/\lambda \quad (30)$$

where:

$y_t =$ response variable

$\lambda =$ transformation parameter

If $\lambda = 0$ the natural log of the data is taken.

The linear regression models of the deseasonalized data were subjected to Box-Cox transformations. In the case of Continental Airlines, $\lambda \approx 4$ for the pre 9/11 data and $\lambda \approx -2$ for the post 9/11 modified data. For American Airlines, $\lambda \approx -2$ for both pre 9/11 data and post 9/11 modified data. For Delta Airlines $\lambda \approx 10.5$ for the pre 9/11 data and for the post 9/11 data $\lambda \approx 6$ is suggested.

3.3 Diagnostic Tests for Assumptions

In this section is explained the diagnostics used to test the assumption.

3.3.1 Constant Variance

Constant variance is the assumption that the spread of residuals is constant over time. In order to visually test the validity of the assumption the residuals must be standardized then plotted against the index or fitted values. To standardize the residuals the standard error of e_i is estimated with $s_{e_i} = \sqrt{MSE(1-h_{ii})}$ and the studentized residual r_i is defined by:

$$r_i = \frac{e_i}{\sqrt{MSE(1-h_{ii})}} \tag{31}$$

where

e_i = ordered residuals

MSE = mean squared error

h_{ii} = leverage value

For constant variance the sign of residual is not needed, so $\sqrt{|r_i|}$ is plotted against the index or fitted values. When plotted against time or $\hat{y}$ it is expected that a constant band would be displayed indicating no change over time of magnitude of observed distance from the line. If a funneling in appears, this indicates a decrease, and a fanning out indicates an increase.

To statistically test the assumption, the Breusch-Pagan test is used. The test statistic is found by regressing squared residuals against x_i and obtaining regression sum of squares, SSR^*. The test statistic is calculated as:

$$B = \frac{SSR^*}{2} \cdot \left(\frac{n}{SSE}\right)^2 \tag{32}$$

where

SSE = sum of squared residuals.

The significance of this test statistic is then used to test the hypotheses:

H_0 : variance equal for all observations

H_a : variance not the same for all observations.

3.3.2 Normality

The normality assumption states that the population of error terms is normally distributed. To visually validate this assumption the order statistics of the sample are plotted against quantiles from an $N(0,1)$ distribution. If the normality assumption holds, the points would be randomly scattered around the straight-line display.

The Shapiro-Wilk normality test is used to statistically test the assumption. The test statistic is based on the statistic:

$$W = \frac{(\sum_{i=1}^{n} a_i e_{(i)})^2}{\sum_{j=1}^{n} e^2} \tag{33}$$

where

$e_{(i)}$ = ordered residuals

a_i = constants generated from means, variances & covariances of the order statistics of sample size n.

The significance of the statistic is used to test the hypotheses:

H_0 : the residuals are normally distributed

H_a : the residuals are not normally distributed

3.3.3 Independence

Independence is the assumption that each and every error term is statistically independent of each and every other error term. To visually assess the assumption the residuals versus fitted values are plotted. Departures from independence are viewed as correlation amongst residuals. Positive correlation is exhibited by positive residuals followed by positive residuals, and negative residuals followed by negative residuals. This causes cyclical features in residual plots with sequences of positive residuals being followed by negative sequences. Negative correlation is the instance where positive residuals are followed by negative ones followed by positive ones. This is visually associated as an alternating pattern in residual plots.

To statistically test the assumption the Durbin-Watson test is employed. The test statistic is based on the statistic:

$$D = \frac{\sum_{i=1}^{n} (e_i - e_{i=1})^2}{\sum_{j=1}^{n} e_j^2} . \tag{34}$$

The significance of the statistic is used to test the hypotheses:

H_0 : residuals are not correlated

H_a : residuals are correlated.

3.4 Model Comparisons

In this section is explained the tests used to compare models for best fit.

3.4.1 Akaike's Information Criterion

Akaike's Information Criterion (AIC) is a statistical measure of goodness of fit of a model to the data. The formula is:

$$\text{AIC} = 2k - 2 \ln L \tag{35}$$

where

k = number of parameters

L = likelihood function.

According to this criterion, a better model has a lower AIC.

3.4.2 Adjusted R^2

The multiple coefficient of determination, R^2, is a measure of the usefulness of a model. For the regression model:

1. Total Variation (TV) = $\sum (y_i - \bar{y})^2$ (36)

2. Explained Variation (EV)= $\sum (\hat{y}_i - \bar{y})^2$ (37)

3. Unexplained Variation (UV)= $\sum (y_i - \hat{y}_i)^2$ (38)

4. Total Variation = Explained Variation + Unexplained Variation

5. The Multiple Coefficient of Determination is:

$$R^2 = \frac{EV}{TV} \tag{39}$$

6. R^2 is the proportion of the total variation in the n observed values of the dependent variable that is explained by the overall regression model.

7. Multiple Correlation Coefficient = R = $\sqrt{R^2}$ (40)

Adding independent variables reduces the unexplained variation. For this reason, the multiple coefficient of determination improves with the addition of independent variables. To avoid overestimating the importance of the independent variables, a calculation of the adjusted multiple coefficient of determination is recommended. The adjusted multiple coefficient of determination (adjusted R^2) is:

$$\bar{R}^2 = \left(R^2 - \frac{k}{n-1} \right) \left(\frac{n-1}{n-(k+1)} \right) \tag{41}$$

where

k = number of independent variables.

The closer the adjusted R^2 is to 1, the more useful the model.

3.4.3 F-statistic

The F-statistic is a measure used to test the significance of a model. The test is a way of testing the hypotheses:

H_0: $\beta_1 = \beta_2 = ...\beta_k = 0$

H_a: at least one of $\beta_1, \beta_2, ..., \beta_k$ does not equal 0, where k is number of parameters. The F-statistic is defined as:

$$F(\text{model}) = \frac{EV / k}{UV / [n - (k+1)]}. \tag{42}$$

When F is large, the null hypothesis is rejected, meaning the model is significant.

3.4.4 ANOVA

Analysis of variance (ANOVA) is used to compare the different additive decomposition regression models in order to choose the best fit. ANOVA is used to compare parameters of a full model to the nested models to see if subsequent parameters are significant. When a linear model is compared to a quadratic model, the parameter B_2 is tested for significance. When comparing a quadratic model to a cubic model the B_3 parameter is tested for significance. The null hypothesis would be of the form:

$H_0 : B_i = 0$,

where i is the parameter whose significance is being tested. The test is based on the statistic:

$$F = \frac{(SYY - RSS)/(p-1)}{RSS/(n-p)}, \tag{43}$$

where

SSY is the residual sum of squares for the reduced model

RSS is the residual sum of squares for the full model

n is the number of parameter in the full model

p is the number of parameters in the reduced model.

When the p-value is small, we reject the null hypothesis and conclude that the parameter being tested is of significance and the larger model is used.

4. Analysis

4.1 Simple Linear Regression

This section describes the simple linear regression model of each airline. In this section is shown the confidence and prediction intervals that were constructed and how the original data were compared to the intervals.

4.1.1 Continental Airlines

Continental Airlines held 7.6% of the market share of domestic airlines as of 2006. The number of passengers per month for the years starting January 1996 (month 1) through August 1991 (month 69) is held as the *y* variable, while the month is held as the *x* variable. The data were plotted as a scatter-plot diagram (Figure 1).

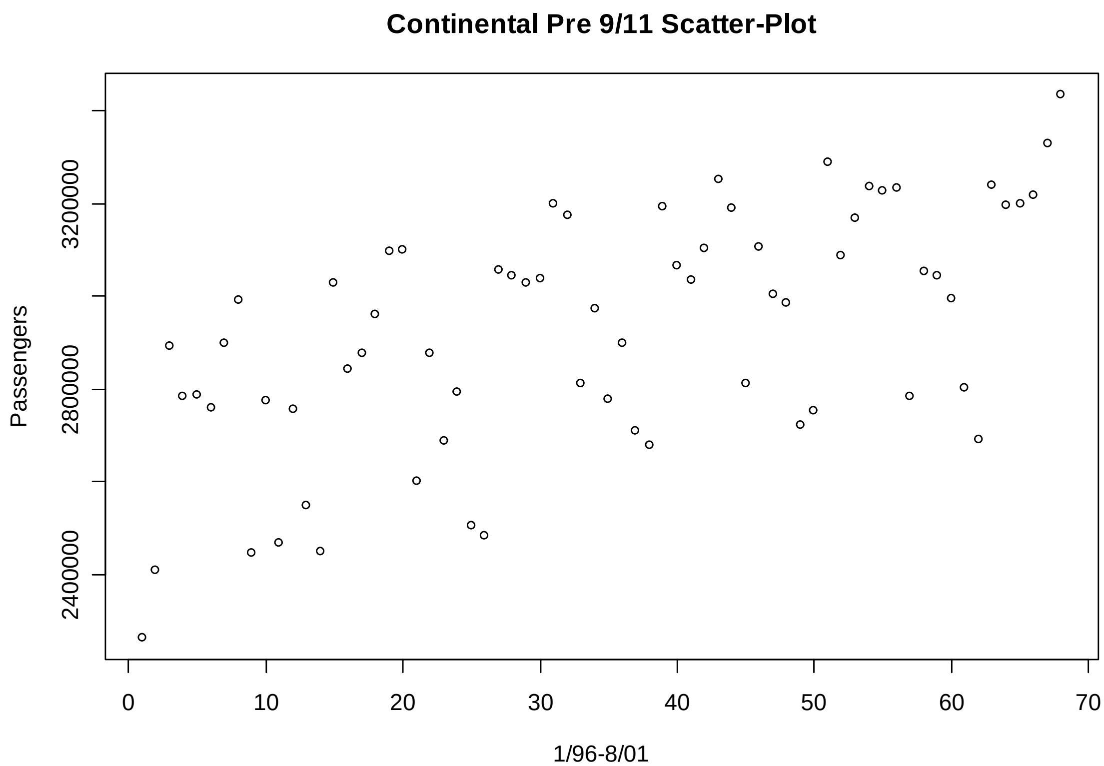

Figure 1: Continental Airlines Pre 9/11 Scatter-plot.

It was decided there appeared to be a straight-line relationship between the variables.

The data were then fit with a simple linear regression model. In the summary of the model it is shown by the intercept and the parameter, Monthsb4, that the equation of the regression line is $\hat{y} = 2655634 + 7803x$. The F-statistic of 37.11 with a p-value 6.422e-08, which is very close to zero, implies that there is a significant relationship between the month and number of passengers (Output 1).

```
Call:
lm(formula = Passengersbf ~ Monthsb4)

Residuals:
    Min      1Q  Median      3Q     Max 
-449125 -163181   50863  152715  300371 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  2655634      50842  52.233  < 2e-16 ***
Monthsb4        7803       1281   6.092 6.42e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Residual standard error: 207300 on 66 degrees of freedom
Multiple R-Squared: 0.3599,     Adjusted R-squared: 0.3502 
F-statistic: 37.11 on 1 and 66 DF,  p-value: 6.422e-08
```

Output 1: Summary on Continental Airlines Pre 9/11 Simple Linear Regression Model.

The Shapiro-Wilk test shows that there is strong evidence to reject the normality assumption.

```
 Shapiro-Wilk normality test
data:  resid(passb4.lm) 
W = 0.9346, p-value = 0.001473
```

By the Breusch-Pagan test it is seen the evidence is not strong enough to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  passb4.lm 
BP = 0.8285, df = 1, p-value = 0.3627
```

The Durbin-Watson test indicates there is strong evidence that the data are not independent.

```
  Durbin-Watson test
data:  passb4.lm
DW = 1.4077, p-value = 0.008337
alternative hypothesis: true autocorelation is not 0
```

Confidence and prediction bands for the data were calculated. The confidence bands are seen as dashed lines and the prediction bands are dotted. (Figure 2).

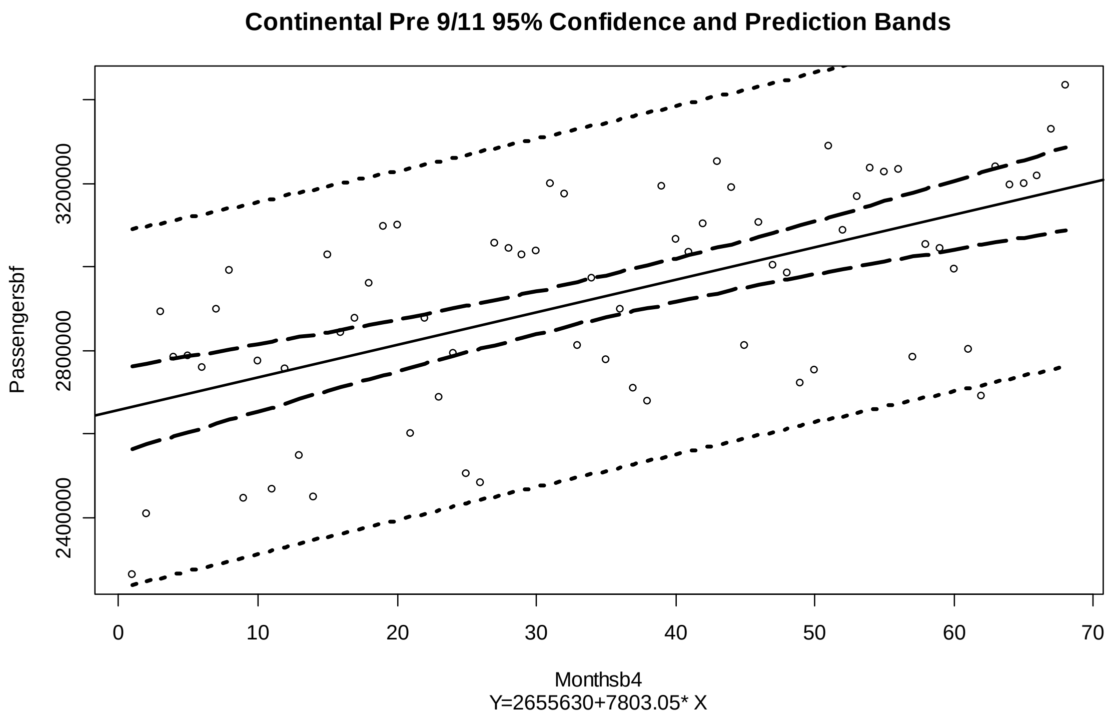

Figure 2: Confidence and Prediction Bands for Continental Airlines Pre 9/11.

From the regression equation predictions were made for September 2001 (69), July 2002 (79), May 2003 (89), March 2004 (99), January 2005 (109), November 2005 (119), and September 2006 (129). The, actual number of passengers, were then compared to the confidence and prediction intervals (Table 1).

	Month	fit	lwr.CI	upr.CI	fit.1	lwr.PI	upr.PI	Actual
1	69	3194045	3092534	3295555	3194045	2767850	3620239	1836826
2	79	3272075	3147692	3396458	3272075	2839862	3704289	2877723
3	89	3350106	3201963	3498248	3350106	2910465	3789746	2628289
4	99	3428136	3255715	3600558	3428136	2979732	3876541	2762806
5	109	3506167	3309138	3703195	3506167	3047737	3964596	2391679
6	119	3584197	3362343	3806052	3584197	3114562	4053832	2805470
7	129	3662228	3415395	3909061	3662228	3180290	4144166	2629849

Table 1: Confidence and Prediction Intervals for Continental Airlines Pre 9/11.

4.1.2 American Airlines

American Airlines with 15.6% of the market share in 2006 was then similarly analyzed. The scatter-plot diagram (Figure 3) showed an apparent straight-line relationship between month and number of passengers.

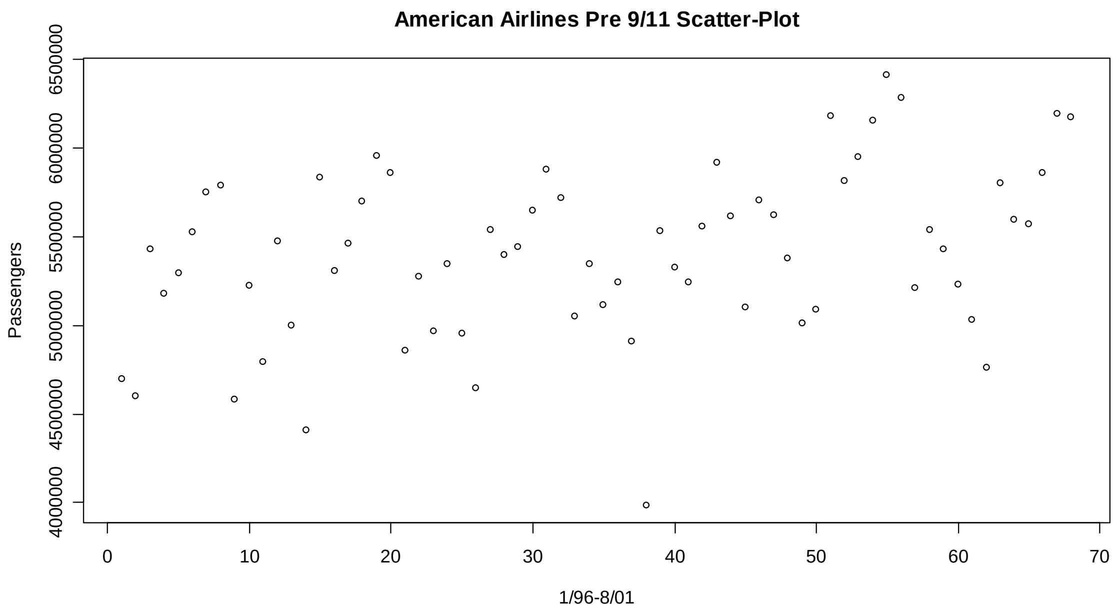

Figure 3: American Airlines Pre 9/11 Scatter-plot.

The data were then fit with a simple linear regression model.

```
Call:
lm(formula = Passengersbf ~ Monthsb4)

Residuals:
     Min       1Q   Median       3Q      Max
-1449362  -348891    45138   307206   826794

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  5103262     110170  46.322  < 2e-16 ***
Monthsb4        8736       2776   3.147  0.00247 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 449200 on 66 degrees of freedom
Multiple R-Squared: 0.1305,     Adjusted R-squared: 0.1173
F-statistic: 9.906 on 1 and 66 DF,  p-value: 0.002473
```

Output 2: Summary on American Airlines Pre 9/11 Simple Linear Regression Model

The F-statistic is 9.906 with a p-value of .002473 < .05 (Output 2). Again, there is strong evidence of a relationship between month and number of passengers. The regression equation for American Airlines became $\hat{y} = 5103262 + 8736x$ (Output 2).

The Shapiro-Wilk test shows there is not enough evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(passb4.lm)
W = 0.9752, p-value = 0.1924
```

The Breusch-Pagan test shows there is not enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  passb4.lm
BP = 0.0842, df = 1, p-value = 0.7716
```

The Durbin-Watson test gives enough evidence to reject the assumption of constant variance.

```
  Durbin-Watson test
data:  passb4.lm
DW = 1.384, p-value = 0.006073
alternative hypothesis: true autocorelation is not 0
```

Confidence and prediction bands for the data were calculated. The confidence bands are seen as dashed lines and the prediction bands are dotted (Figure 4).

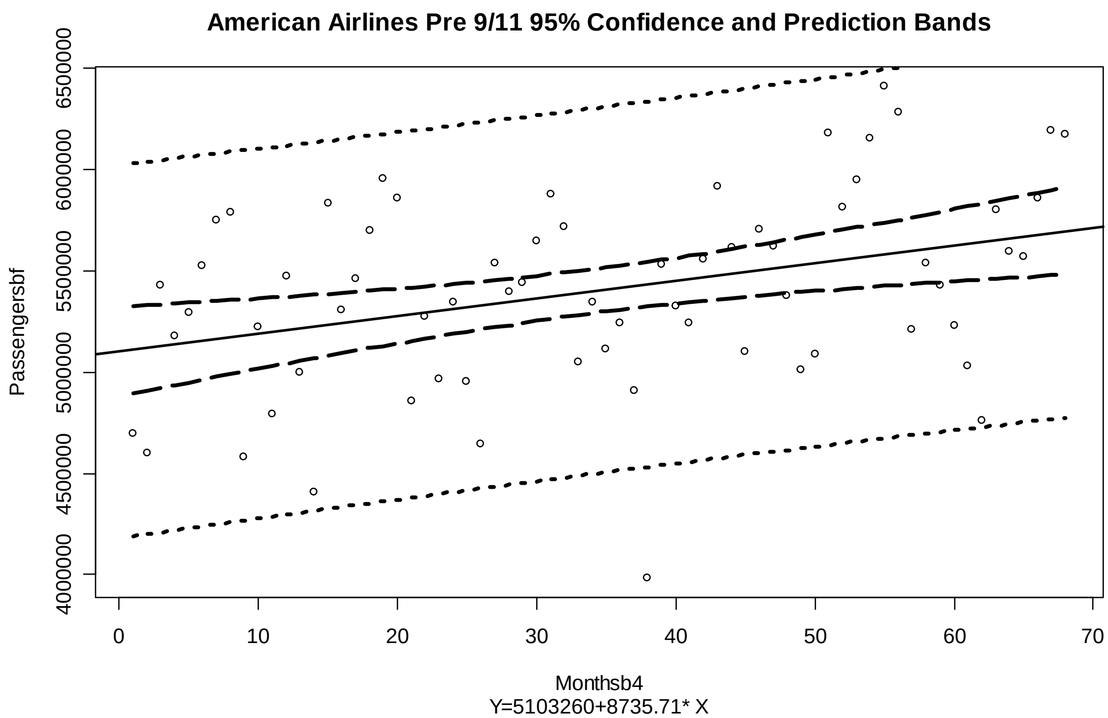

Figure 4: American Airlines Pre 9/11 Confidence and Prediction Bands.

From the regression equation predictions were made for September 2001 (69), July 2002 (79), May 2003 (89), March 2004 (99), January 2005 (109), November 2005 (119), and September 2006 (129). The, actual number of passengers, were then compared to the confidence and prediction intervals (Table 3).

	Month	fit	lwr.CI	upr.CI	fit.1	lwr.PI	upr.PI	Actual
1	69	5706026	5486065	5925987	5706026	4782511	6629540	3251642
2	79	5793383	5523858	6062907	5793383	4856825	6729940	7368482
3	89	5880740	5559732	6201747	5880740	4928089	6833390	6218797
4	99	5968097	5594479	6341714	5968097	4996455	6939738	6425910
5	109	6055454	5628516	6482392	6055454	5062089	7048818	5784397
6	119	6142811	5662078	6623544	6142811	5125166	7160456	6338404
7	129	6230168	5695309	6765027	6230168	5185864	7274472	5676574

Table 2: American Airlines Pre 9/11 Confidence and Prediction Intervals.

It was noticed that with this airline, the majority of the actual data are above confidence bands and within prediction intervals.

4.1.3 Delta Airlines

Delta Airlines held 11.5% of the market share as of 2006. Delta was analyzed in a similar manner. The scatter-plot shows an apparent straight-line relationship between month and number of passengers (Figure 5).

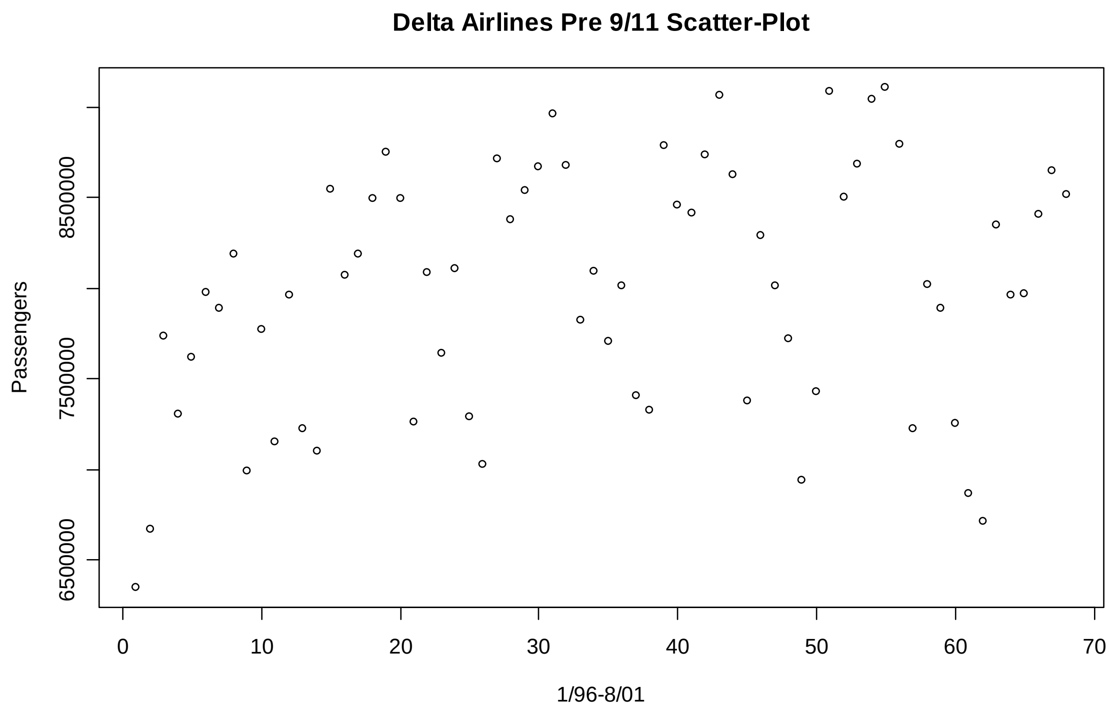

Figure 5: Delta Airlines Pre 9/11 Scatter-plot.

The data were then fit with a simple linear regression model.

```
Call:
lm(formula = Passengersbf ~ Monthsb4)

Residuals:
     Min       1Q   Median       3Q      Max
-1528168  -571753   118993   536447  1007430

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  7670285     161154  47.596   <2e-16 ***
Monthsb4        9148       4060   2.253   0.0276 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 657100 on 66 degrees of freedom
Multiple R-Squared: 0.07142,    Adjusted R-squared: 0.05735
F-statistic: 5.077 on 1 and 66 DF,  p-value: 0.02758
```

Output 3: Summary on Delta Airlines Pre 9/11 Simple Linear Regression Model.

The F-statistic of 5.077 with a p-value of .02758 < .05 indicates evidence of a significant relationship between month and number of passengers (Output 3). The equation for the regression line from this model was $\hat{y} = 7670285 + 9148x$ (Output 3).

The Shapiro-Wilk test shows there is strong evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(passb4.lm)
W = 0.9624, p-value = 0.03852
```

The Breusch-Pagan test signifies there is not enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  passb4.lm
BP = 0.9592, df = 1, p-value = 0.3274
```

The Durbin-Watson test indicates that there is enough evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  passb4.lm
DW = 1.1258, p-value = 7.974e-05
alternative hypothesis: true autocorelation is not 0
```

Confidence and prediction bands for the data were calculated. The confidence bands are seen as dashed lines and the prediction bands are dotted (Figure 6).

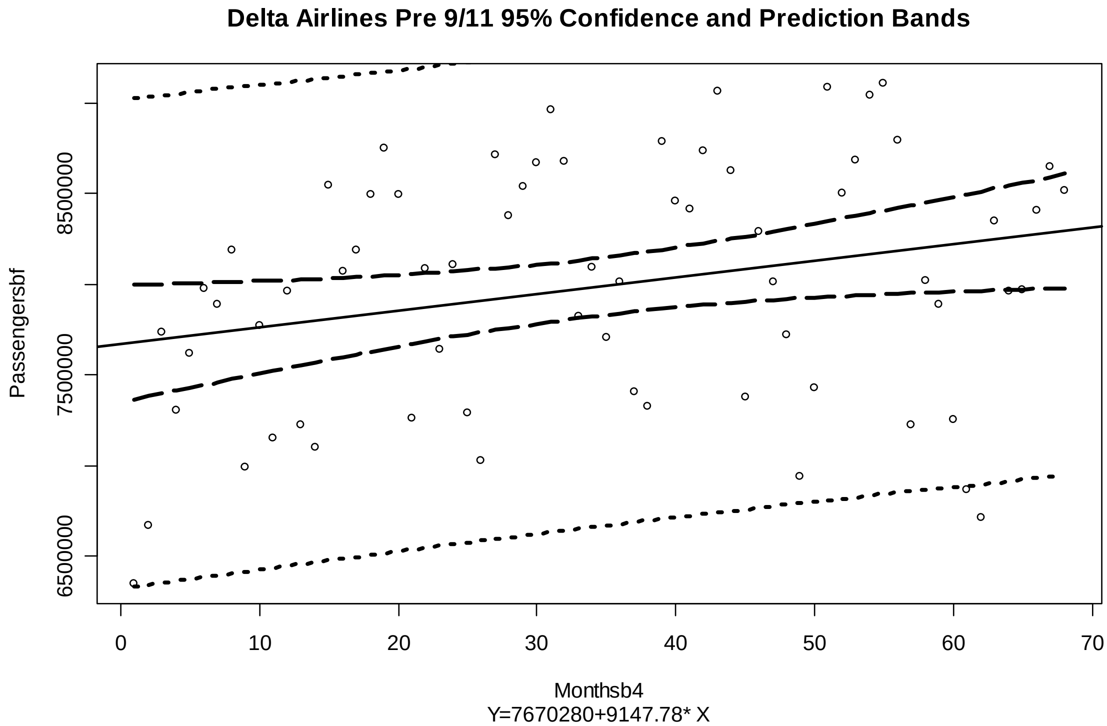

Figure 6: Delta Airlines Pre 9/11 Confidence and Prediction Bands.

From the regression equation predictions were made for September 2001 (69), July 2002 (79), May 2003 (89), March 2004 (99), January 2005 (109), November 2005 (119), and September 2006 (129). The, actual number of passengers, were then compared to the confidence and prediction intervals (Table 3).

	Month	fit	lwr.CI	upr.CI	fit.1	lwr.PI	upr.PI	Actual
1	69	8301481	7979727	8623236	8301481	6950583	9652380	4651514
2	79	8392959	7998704	8787214	8392959	7022982	9762936	7700152
3	89	8484437	8014874	8954000	8484437	7090919	9877954	6337258
4	99	8575915	8029394	9122435	8575915	7154617	9997212	6947233
5	109	8667392	8042876	9291909	8667392	7214319	10120466	6186057
6	119	8758870	8055663	9462077	8758870	7270280	10247460	5742877
7	129	8850348	8067967	9632729	8850348	7322761	10377935	5676574

Table 3: Delta Airlines Pre 9/11 Confidence and Prediction Intervals.

The actual numbers for Delta fall below that of the predicted values.

Both Continental Airlines and Delta Airlines actual passenger numbers for the months predicted with the simple linear regression model using the pre 9/11 data fell below the confidence and prediction intervals. This indicates that the airlines have not yet achieved the pre 9/11 passenger numbers. American Airlines actual numbers, for the majority, are above confidence intervals and within prediction intervals. This indicates that not much has changed according to pre 9/11 forecasting.

4.2 Dummy Variables

This section describes the technique of modeling the seasonal variation of each airline with dummy variables. In this section is shown how comparisons were made between the pre 9/11 and post 9/11 parameter intervals and the fitted values.

4.2.1 Continental Airlines

It was decided to model the seasonal variation in Continental Airlines with dummy variables (Figure 7).

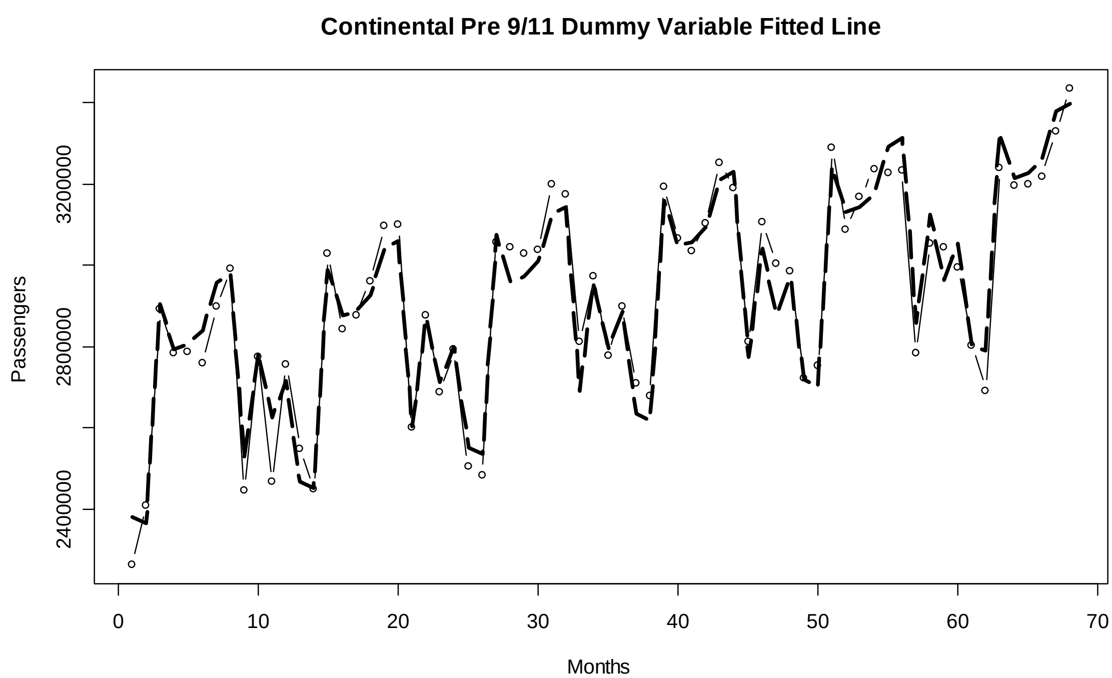

Figure 7: Continental Airlines Pre 9/11 Data Modeled with Dummy Variables.

A summary of the model indicates that with an *F*-statistic of 89.45 and an associated p-value <2.2e-16, there is very strong evidence that the number of passengers and the month are related (Output 4). The regression equation for the Continental Data with seasonal variation modeled with dummy variables (Output 4) is:

$$\begin{aligned}\hat{y}_t &= 2373901.3 + 7043.4t - 22313.9t_{s2} + 510071.4t_{s3} + 389920.5t_{s4} \\ &+ 395062.8t_{s5} + 424554.5t_{s6} + 532435.0t_{s7} + 545091.7t_{s8} + 8036.7t_{s9} \\ &+ 342868.3t_{s10} + 175106.3t_{s11} + 257199.1t_{s12}\end{aligned} \tag{44}$$

where

t_{s2} = 1 if time period 2 is season 2, 0 otherwise
t_{s3} = 1 if time period 3 is season 3, 0 otherwise
t_{s4} = 1 if time period 4 is season 4, 0 otherwise
t_{s5} = 1 if time period 5 is season 5, 0 otherwise
t_{s6} = 1 if time period 6 is season 6, 0 otherwise
t_{s7} = 1 if time period 7 is season 7, 0 otherwise
t_{s8} = 1 if time period 8 is season 8, 0 otherwise
t_{s9} = 1 if time period 9 is season 9, 0 otherwise
t_{s10} = 1 if time period 10 is season 10, 0 otherwise
t_{s11} = 1 if time period 11 is season 11, 0 otherwise
t_{s12} = 1 if time period 12 is season 12, 0 otherwise

```
Call:
lm(formula = Passengersbf ~ Monthsb4 + period)

Residuals:
    Min      1Q  Median      3Q     Max
-160942  -39923    1505   39734  124695

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 2373901.3    28287.9  83.919  < 2e-16 ***
Monthsb4       7043.4      389.3  18.094  < 2e-16 ***
periodFeb    -22313.9    36184.6  -0.617   0.5400
periodMar    510071.4    36190.9  14.094  < 2e-16 ***
periodApr    389920.5    36201.4  10.771 3.65e-15 ***
periodMay    395062.8    36216.0  10.909 2.26e-15 ***
periodJun    424554.5    36234.8  11.717  < 2e-16 ***
periodJul    532435.0    36257.8  14.685  < 2e-16 ***
periodAug    545091.7    36285.0  15.023  < 2e-16 ***
periodSep     83036.7    37956.5   2.188   0.0330 *
periodOct    342868.3    37966.5   9.031 1.89e-12 ***
periodNov    175106.3    37980.5   4.610 2.44e-05 ***
periodDec    257199.1    37998.4   6.769 9.02e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 62670 on 55 degrees of freedom
Multiple R-Squared: 0.9513,     Adjusted R-squared: 0.9406
F-statistic: 89.45 on 12 and 55 DF,  p-value: < 2.2e-16
```

Output 4: Summary on Continental Airlines Pre 9/11 Dummy Variable Model.

The Shapiro-Wilk test indicates there is not enough evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(passb4dum.lm)
W = 0.9901, p-value = 0.8696
```

The Breusch-Pagan test indicates there is not enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  passb4dum.lm
BP = 17.6, df = 12, p-value = 0.1284
```

The Durbin-Watson test indicates there is enough evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  passb4dum.lm
DW = 1.4409, p-value = 0.02969
alternative hypothesis: true autocorelation is not 0
```

The data were then tested for possible outliers using residual plots. In each plot observation numbers 11, 33, and 47 are singled out (Figure 8). These observations are deemed to be possible outliers. The Cook's Distance plot is used to check for influential outliers (Figure 8).

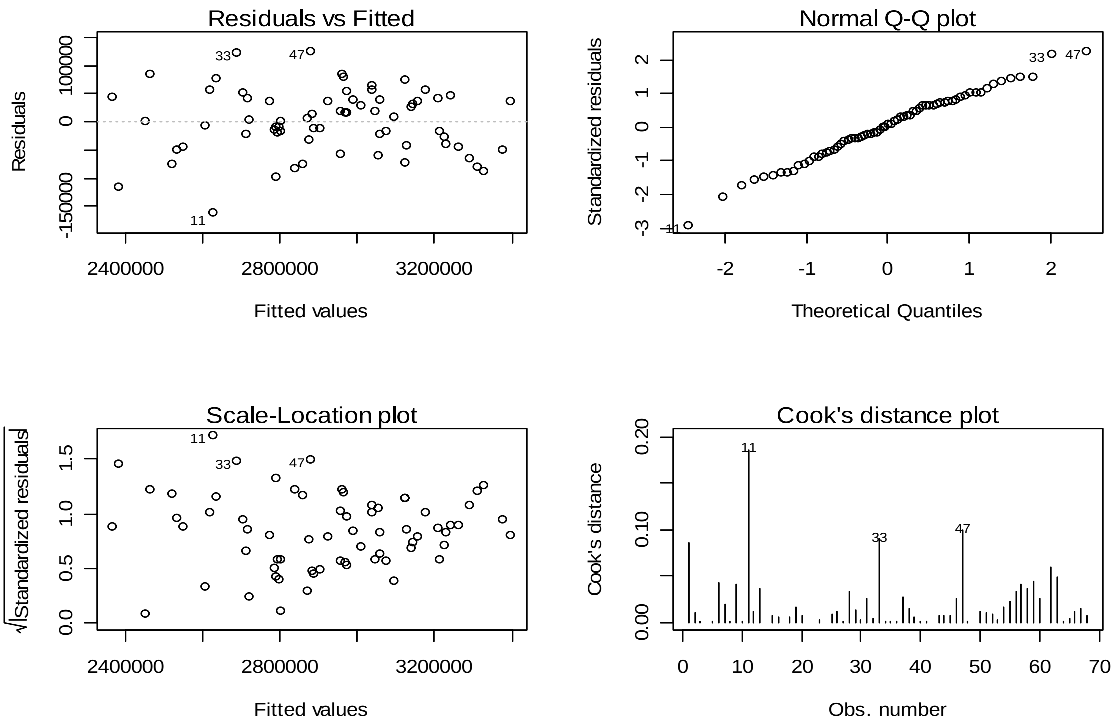

Figure 8: Continental Airlines Pre 9/11 Dummy Variable Residual Plots.

According to the plot and the Cook's D test, there are no influential outliers.

```
cooksD=cooks.distance(passb4dum.lm)
f0.50=qf(0.5, df1=5,df2=24)
 cooksD[which(cooksD>f0.50)]
named numeric(0)
```

It was decided to model the data excluding the possible outliers.

```
Call:
lm(formula = Passengersbf ~ Monthsb4 + period)

Residuals:
    Min      1Q  Median      3Q     Max
-417265  -87038    8316  117193  366265

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2527083      82725  30.548  < 2e-16 ***
Monthsb4        8134       1201   6.770 1.15e-08 ***
periodFeb     200170     104469   1.916  0.06086 .
periodMar     249536     104490   2.388  0.02060 *
periodApr     254034     104525   2.430  0.01857 *
periodMay     304911     104573   2.916  0.00522 **
periodJun     234008     109568   2.136  0.03743 *
periodJul     227709     109561   2.078  0.04263 *
periodAug     154877     109568   1.414  0.16346
periodSep      19771     109588   0.180  0.85753
periodOct     -27097     109621  -0.247  0.80573
periodNov     -71402     109667  -0.651  0.51786
periodDec      31675     109726   0.289  0.77398
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 180900 on 52 degrees of freedom
Multiple R-Squared: 0.5945,     Adjusted R-squared: 0.5009
F-statistic: 6.352 on 12 and 52 DF,  p-value: 9.251e-07
```

Output 5: Continental Airlines Pre 9/11 Dummy Variable Modified Model.

According to the *F*-statistic with a p-value 9.251e-07, the number of passengers is related to the month (Output 5). The regression equation (Output 5) for the Continental Airlines modified data is:

$$\hat{y}_t = 2527083 + 8134t + 200170t_{s2} + 249536t_{s3} + 254034t_{s4} + 304911t_{s5} + 234008t_{s6} + 227709t_{s7} + 154877t_{s8} + 19771t_{s9} - 27097t_{s10} - 71402t_{s11} + 31675t_{s12} \quad (45)$$

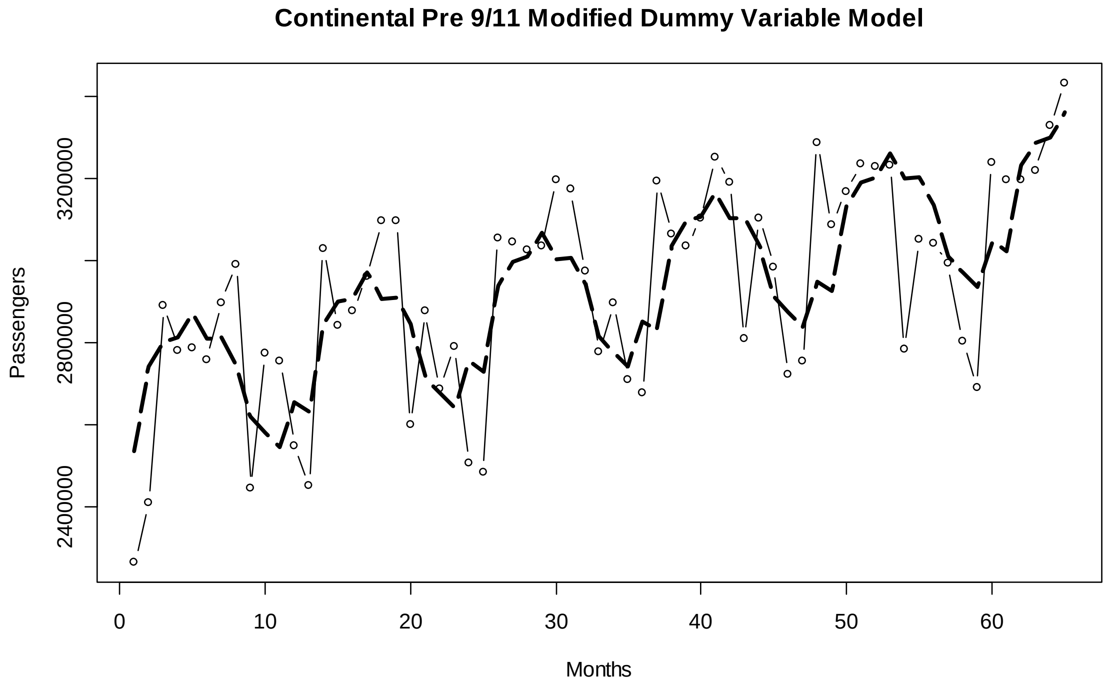

Figure 9: Continental Airline Pre 9/11 Dummy Variable Modified Model.

The Shapiro-Wilk test does not give enough evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(passb4dum.lm)
W = 0.9853, p-value = 0.6345
```

The Breusch-Pagan test does not give enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  passb4dum.lm
BP = 20.0141, df = 12, p-value = 0.06682
```

The Durbin-Watson test gives enough evidence to reject the independence assumption.

```
   Durbin-Watson test
data:  passb4dum.lm
DW = 1.8997, p-value = 0.6959
alternative hypothesis: true autocorelation is not 0
```

Visual examination (Figure 9) shows the fit is not as nice as our previous model (Figure 7). Checking the Akaike's Information Criterion (AIC) for both models (Output 6) and comparing them, it is discovered that the criterion for the model with the original data is slightly better than the model with the modified data. Thus further analysis was run with the original model.

```
AIC(passb4dum.lm)
[1] 1708.754
AIC(passb4dum.lm)
[1] 1771.723
```

Output 6: AIC Comparison for Continental Airlines Pre 9/11 Original and Modified Data.

The focus was then shifted to the data after 9/11, October 1991 (month 70) through October 2006 (month 130). The seasonal variation was modeled with dummy variables (Figure 10) and a summary was run.

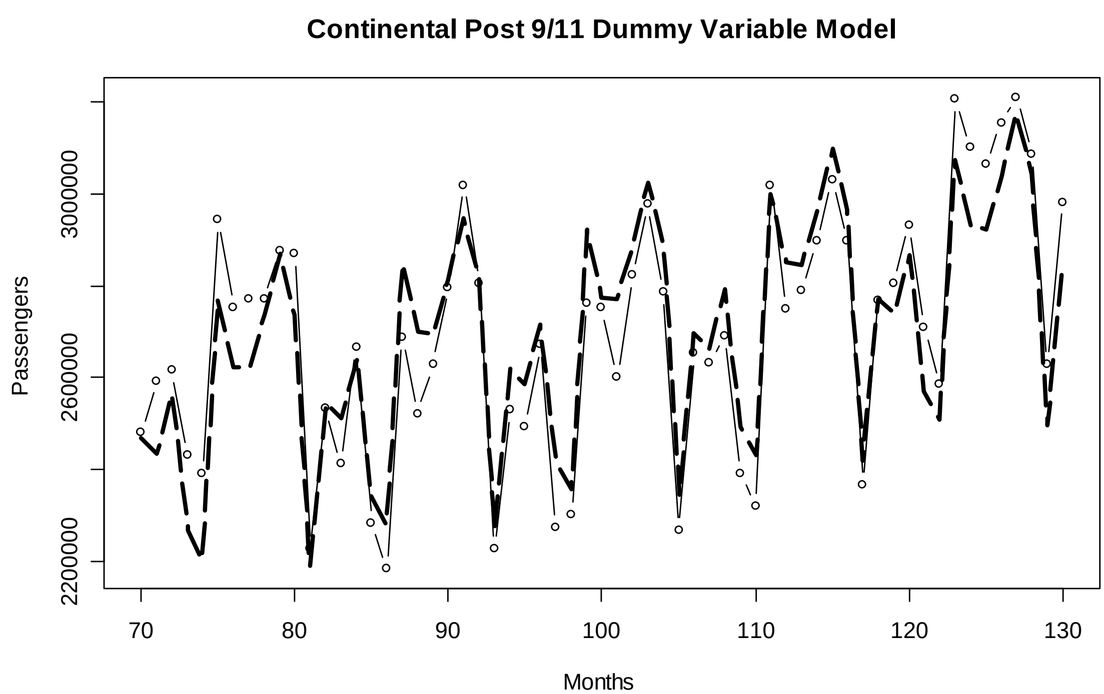

Figure 10: Continental Post 9/11 Data Modeled with Dummy Variables.

According to the *F*- statistic 23.61 with corresponding p-value 4.106e-16 indicates there is a relationship between month and number of passengers (Output 7). The regression equation for the Continental Airlines data post 9/11 is (Output 7):

$$\hat{y}_t = 1803171.7 + 6342.2t - 68792.6t_{s2} + 492511.3t_{s3} + 338644.5t_{s4} + 327354.2t_{s5} + 438785.0t_{s6} + 567468.3t_{s7} + 427359.7t_{s8} - 125313.4t_{s9} + 221134.8t_{s10} + 181716.1t_{s11} + 304587.8t_{s12} \quad (46)$$

```
Call:
lm(formula = passengersafter ~ monthsafter + period)

Residuals:
    Min      1Q  Median      3Q     Max
-177957  -69092  -15459   72523  187959

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 1803171.7    95451.2  18.891  < 2e-16 ***
monthsafter    6342.2      834.6   7.599 8.92e-10 ***
periodFeb    -68792.6    71524.8  -0.962  0.34097
periodMar    492511.3    71539.4   6.884 1.11e-08 ***
periodApr    338644.5    71563.7   4.732 1.99e-05 ***
periodMay    327354.2    71597.8   4.572 3.40e-05 ***
periodJun    438785.0    71641.6   6.125 1.62e-07 ***
periodJul    567468.3    71695.0   7.915 2.96e-10 ***
periodAug    427359.7    71758.1   5.956 2.93e-07 ***
periodSep   -125313.4    71830.9  -1.745  0.08746 .
periodOct    221134.8    68520.9   3.227  0.00225 **
periodNov    181716.1    71539.4   2.540  0.01437 *
periodDec    304587.8    71524.8   4.258 9.53e-05 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 113100 on 48 degrees of freedom
Multiple R-Squared: 0.8551,     Adjusted R-squared: 0.8189
F-statistic: 23.61 on 12 and 48 DF,  p-value: 4.106e-16
```

Output 7: Continental Airlines Post 9/11 Data Modeled with Dummy Variables.

The Shapiro-Wilk test showed there was evidence to reject the normality assumption.

```
   Shapiro-Wilk normality test
data:  resid(passafterdum.lm)
W = 0.9504, p-value = 0.01501
```

The Breusch-Pagan test showed evidence of rejecting the constant variance assumption.

```
   studentized Breusch-Pagan test
data:  passafterdum.lm
BP = 21.5924, df = 12, p-value = 0.04235
```

The Durbin-Watson test indicates there is enough evidence to reject the independence assumption.

```
   Durbin-Watson test
data:  passafterdum.lm
DW = 0.5296, p-value = 3.551e-10
alternative hypothesis: true autocorelation is not 0
```

It was then decided to check the post 9/11 data for possible outliers. The residual plots used to check for possible outliers singled out observation numbers 19, 5, and 55 (Figure 11). It is reasonable to assume that after such a shocking event as 9/11, the public would have an initial shock recovery period. It was decided that the most recent months after 9/11 corresponding to possible outliers, would be removed as a block of possible outliers, representing this initial shock recovery period. In the Continental data, observation 5 was singled out as a possible outlier. Therefore, observations 1 through 5 were omitted.

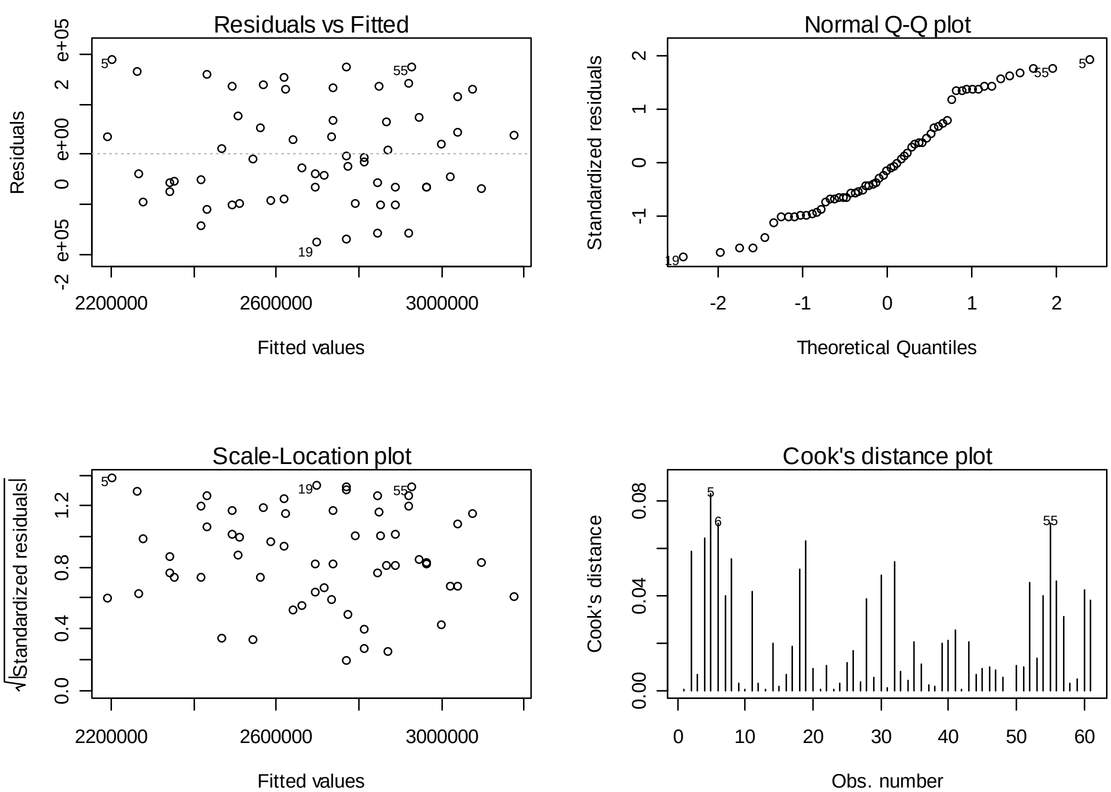

Figure 11: Continental Airlines Post 9/11 Dummy Variable Residual Plots.

The Leverage plot (Figure 11) and Cooks' D statistic showed there did not exist any influential outliers.

```
cooksD=cooks.distance(passafterdum.lm)
f0.50=qf(0.5, df1=5, df2=24)
cooksD[which(cooksD>f0.50)]
named numeric(0)
```

Once the block of possible outliers was removed, the analysis was run again. The summary (Output 8) showed that with an *F*-statistic of 26.26 with a p-value of 6.411e-16 there is evidence of a significant relationship between number of passengers and month. The regression equation for Continental Airlines dummy variable modified model is (Output 8):

$$\begin{aligned}\hat{y}_t &= 2156890.2 + 7546.0t + 127479.4t_{s2} - 13833.1t_{s3} - 567710.1t_{s4} - 224699.4t_{s5} \\ &- 295303.1t_{s6} - 147479.4t_{s7} - 481395.2t_{s8} - 556975.2t_{s9} + 57337.9t_{s10} \\ &- 97732.7t_{s11} - 110227.0t_{s12}\end{aligned} \tag{47}$$

```
Call:
lm(formula = Passengersafter ~ Monthsafter + perioda)

Residuals:
    Min      1Q  Median      3Q     Max
-169978  -72012   -8374   68715  201850

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 2156890.2    97573.1  22.105  < 2e-16 ***
Monthsafter    7546.0      880.2   8.573 7.42e-11 ***
periodaFeb   127479.4    66806.8   1.908 0.063059 .
periodaMar   -13833.1    66824.2  -0.207 0.836982
periodaApr  -567710.1    66853.2  -8.492 9.64e-11 ***
periodaMay  -224699.4    66893.7  -3.359 0.001647 **
periodaJun  -295303.1    70858.7  -4.167 0.000146 ***
periodaJul  -147479.4    70853.2  -2.081 0.043379 *
periodaAug  -481395.2    70858.7  -6.794 2.56e-08 ***
periodaSep  -556975.2    70875.1  -7.859 7.53e-10 ***
periodaOct    57337.9    66853.2   0.858 0.395830
periodaNov   -97732.7    66824.2  -1.463 0.150866
periodaDec  -110227.0    66806.8  -1.650 0.106243
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 105600 on 43 degrees of freedom
Multiple R-Squared: 0.8799,     Adjusted R-squared: 0.8464
F-statistic: 26.26 on 12 and 43 DF,  p-value: 6.411e-16
```

Output 8: Summary on Continental Airlines Post 9/11 Dummy Variable Modified Model.

The plot (Figure 12) shows a slightly better fit to the data than the original model (Figure 10).

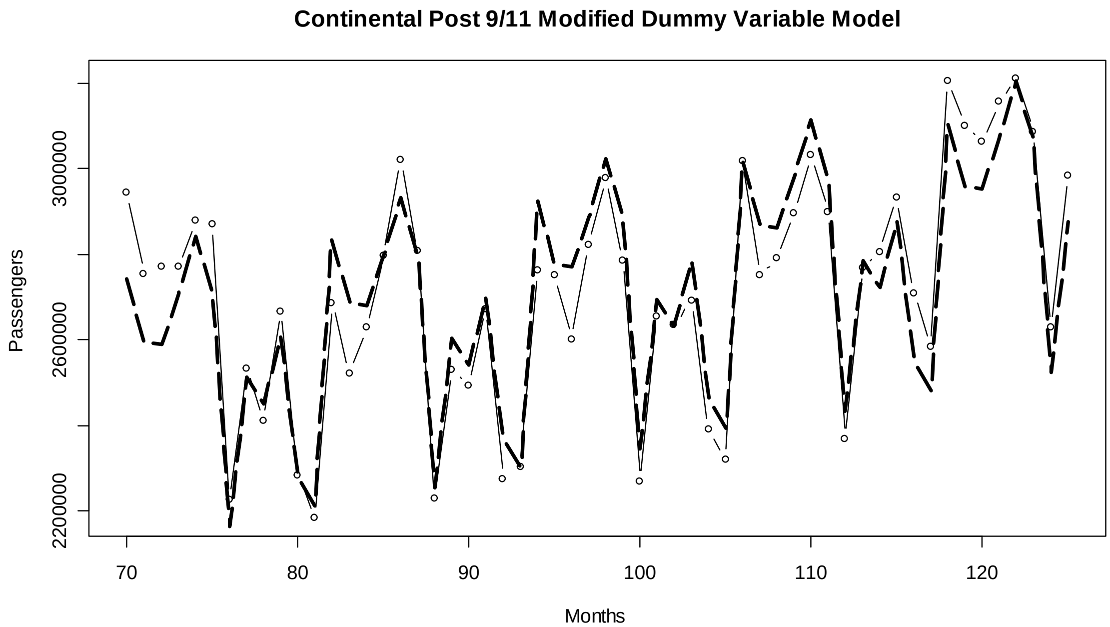

Figure 12: Continental Airlines Post 9/11 Modified Dummy Variable Plot.

The Shapiro-Wilk test shows there is not enough evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(passafterdum.lm)
W = 0.9704, p-value = 0.1825
```

The Breusch-Pagan test shows there is enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  passafterdum.lm
BP = 21.482, df = 12, p-value = 0.04375
```

The Durbin-Watson test shows there is enough evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  passafterdum.lm
DW = 0.5966, p-value = 2.542e-08
alternative hypothesis: true autocorelation is not 0
```

Comparing the AIC of the two Continental post 9/11 models, it is seen that the modified model is slightly better. Thus it is decided to use this model for a comparison.

```
AIC(passafterdum.lm)
[1] 1606.067
AIC(passafterdum.lm)
[1] 1467.702
```

Output 9:AIC Comparison of Continental Airlines Post 9/11 Original and Modified Dummy Variable Models.

The pre 9/11 model with all data and the post 9/11 modified model were compared. Using the regression coefficient estimates, intervals were created using $\pm 2 \cdot$ standard deviation. The intervals for the pre 9/11 data were compared to the post 9/11 intervals (Table 4). The majority of the post 9/11 intervals were below that of the pre 9/11 intervals.

Prediction Interval Comparison

Continental Pre 9/11 Intervals			**Continental Post 9/11 Intervals**			
Estimate	-2*se	+2*se	Estimate	-2*se	+2*se	
2373901.30	2317325.50	2430477.10	2156890.20	1961744.00	2352036.40	overlap
7043.40	6264.80	7822.00	7546.00	5785.60	9306.40	overlap
-22313.90	-94683.10	50055.30	127479.40	-6134.20	261093.00	overlap
510071.40	**437689.60**	**582453.20**	-13833.10	**-147481.50**	**119815.30**	**below**
389920.50	**317517.70**	**462323.30**	-567710.10	**-701416.50**	**-434003.70**	**below**
395062.80	**322630.80**	**467494.80**	-224699.40	**-358486.80**	**-90912.00**	**below**
424554.50	**352084.90**	**497024.10**	-295303.10	**-437020.50**	**-153585.70**	**below**
532435.00	**459919.40**	**604950.60**	-147479.40	**-289185.80**	**-5773.00**	**below**
545091.70	**472521.70**	**617661.70**	-481395.20	**-623112.60**	**-339677.80**	**below**
83036.70	**7123.70**	**158949.70**	-556975.20	**-698725.40**	**-415225.00**	**below**
342868.30	**266935.30**	**418801.30**	57337.90	**-76368.50**	**191044.30**	**below**
175106.30	**99145.30**	**251067.30**	-97732.40	**-231380.80**	**35916.00**	**below**
257199.10	**181202.30**	**333195.90**	-110227.00	**-243840.60**	**23386.60**	**below**

Table 4: Continental Airlines Pre and Post 9/11 Regression Coefficient Interval Comparison.

Another comparison made was that of the fitted values. The fitted values of the post 9/11 data were compared to that of the pre 9/11 data (Tables 5,6). It was seen that the post 9/11 fitted values lied below that of the pre 9/11 data.

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
2380945	2365674	2905103	2791995	2804181	2840716	2955640	2975340	2520329	2787204
11	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
2626485	2715621	2465465	2450195	2989624	2876516	2888702	2925237	3040161	3059861
21	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
2604849	2871724	2711006	2800142	2549986	2534716	3074144	2961037	2973222	3009758
31	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**
3124681	3144382	2689370	2956245	2795526	2884663	2634507	2619236	3158665	3045558
41	**42**	**43**	**44**	**45**	**46**	**47**	**48**	**49**	**50**
3057743	3094278	3209202	3228902	2773891	3040766	2880047	2969183	2719028	2703757
51	**52**	**53**	**54**	**55**	**56**	**57**	**58**	**59**	**60**
3243186	3130078	3142264	3178799	3293723	3313423	2858411	3125286	2964568	3053704
61	**62**	**63**	**64**	**65**	**66**	**67**	**68**		
2803548	2788278	3327707	3214599	3226785	3263320	3378244	**3397944**		

Table 5: Continental Airlines Pre 9/11 Fitted Values

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
2742451	2594927	2589978	2707751	2842777	2709010	2162679	2513236	2450178	2605548
11	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
2279178	2211144	2833004	2685479	2680531	2798304	2933329	2799563	2253232	2603789
21	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
2540731	2696101	2369731	2301697	2923556	2776032	2771083	2888856	3023882	2890115
31	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**
2343784	2694341	2631284	2786653	2460284	2392250	3014109	2866584	2861636	2979409
41	**42**	**43**	**44**	**45**	**46**	**47**	**48**	**49**	**50**
3114434	2980668	2434337	2784894	2721836	2877206	2550836	2482802	3104661	2957137
51	**52**	**53**	**54**	**55**	**56**				
2952188	3069961	3204987	3071220	2524889	2875446				

Table 6: Continental Airlines Post 9/11 Modified Data Fitted Values.

4.2.2 American Airlines

American Airlines seasonal variation was then modeled with dummy variables (Figure 13).

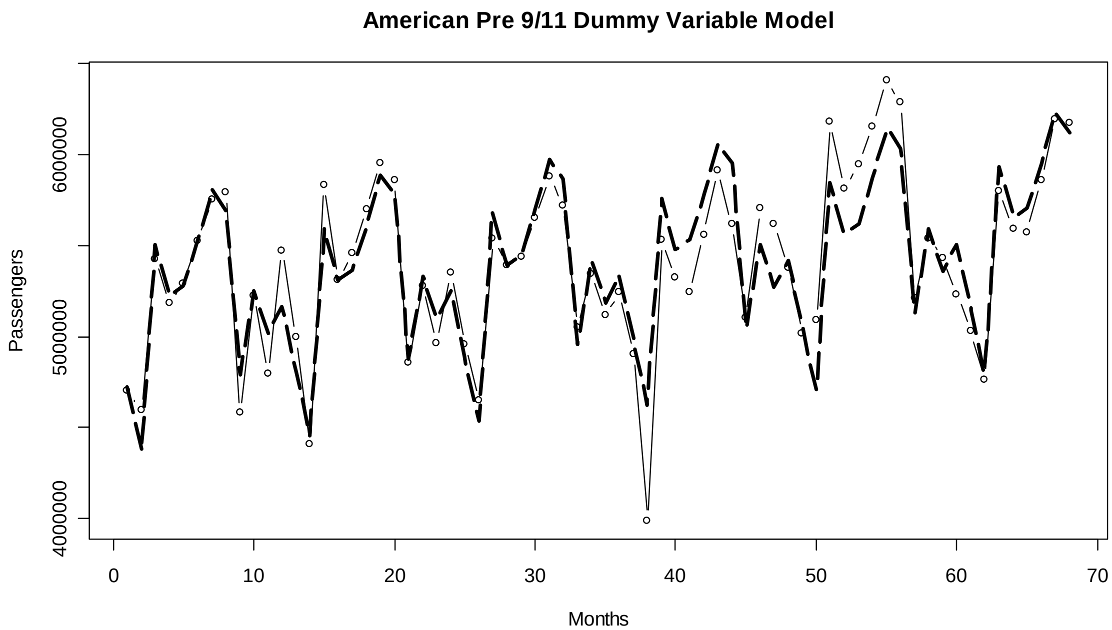

Figure 13: American Airlines Pre 9/11 Data Modeled with Dummy Variables.

A summary of the data indicates that with an F statistic of 26.71 and associated p-value 2.2e-16 there is a strong relationship between number of passengers and months (Output 10). The regression equation (Output 10) is of the form:

$$\hat{y}_t = 4713669 + 7107t - 359725t_{s2} + 771555t_{s3} + 481706t_{s4} + 530778t_{s5} + 773228t_{s6} + 1041421t_{s7} + 925148t_{s8} + 12098t_{s9} + 463890t_{s10} + 22999t_{s11} + 366694t_{s12} \tag{48}$$

```
Call:
lm(formula = Passengersbf ~ Monthsb4 + period)

Residuals:
    Min      1Q  Median      3Q     Max 
-638156  -94616  -28884   90426  379892 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  4713669      91172  51.701  < 2e-16 ***
Monthsb4        7107       1255   5.665 5.56e-07 ***
periodFeb    -359725     116623  -3.085 0.003187 ** 
periodMar     771555     116643   6.615 1.61e-08 ***
periodApr     481706     116677   4.129 0.000125 ***
periodMay     530778     116724   4.547 3.03e-05 ***
periodJun     773228     116785   6.621 1.57e-08 ***
periodJul    1041421     116859   8.912 2.93e-12 ***
periodAug     925148     116947   7.911 1.22e-10 ***
periodSep      12098     122334   0.099 0.921580    
periodOct     463890     122366   3.791 0.000375 ***
periodNov     222999     122411   1.822 0.073937 .  
periodDec     366694     122469   2.994 0.004118 ** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Residual standard error: 202000 on 55 degrees of freedom
Multiple R-Squared: 0.8535,     Adjusted R-squared: 0.8216 
F-statistic: 26.71 on 12 and 55 DF,  p-value: < 2.2e-16
```

Output 10: Summary on American Airlines Pre 9/11 Dummy Variable Model.

According to the Shapiro-Wilk test there is evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(passb4dum.lm) 
W = 0.9602, p-value = 0.02897
```

The Breusch-Pagan test indicates that there is not enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  passb4dum.lm 
BP = 14.2964, df = 12, p-value = 0.2822
```

The Durbin-Watson test reveals that there is strong evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  passb4dum.lm 
DW = 1.096, p-value = 0.0002823
alternative hypothesis: true autocorelation is not 0
```

The American Airlines pre 9/11 data fit with dummy variables was then checked for possible outliers using the residual plots. According to the residual plots (Figure 14), observation numbers 38, 47, and 50 are singled out. These were deemed to be possible outliers. According to the leverage plot (Figure 14), and the corresponding Cook's D statistic:

```
cooksD=cooks.distance(passb4dum.lm)
f0.50=qf(0.5, df1=5,df2=24)
cooksD[which(cooksD>f0.50)]
named numeric(0)
```

There weren't any influential outliers.

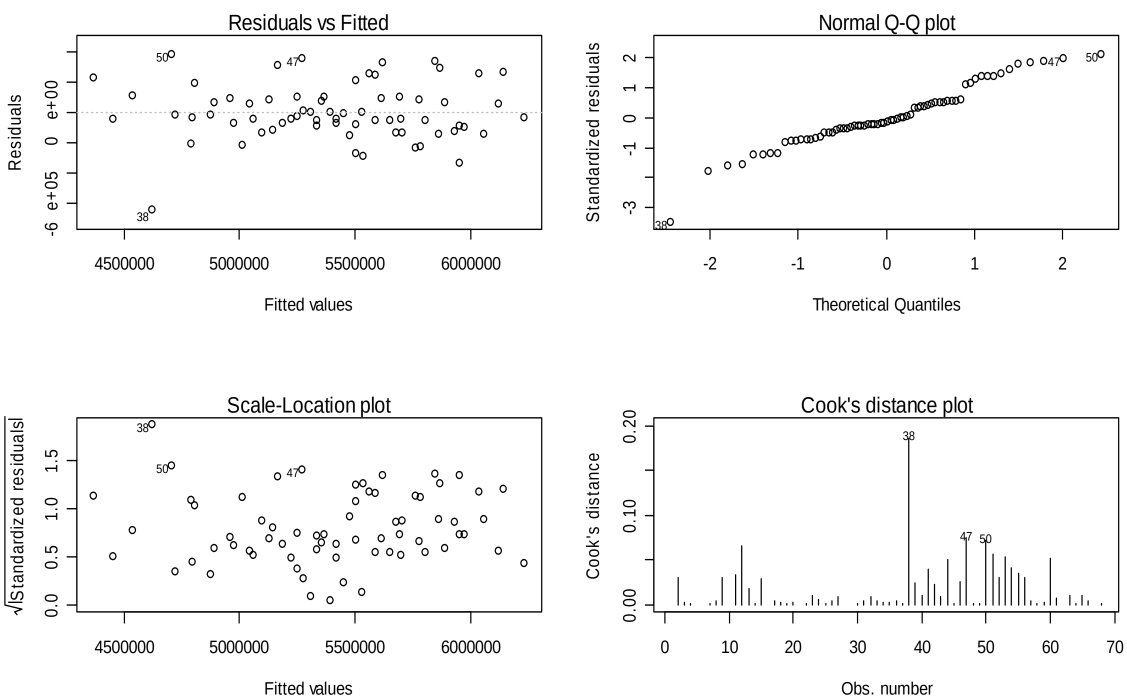

Figure 14: American Airlines Pre 9/11 Dummy Variable Model Residual Plots.

The analyses on the American Airlines pre 9/11 data were then repeated excluding the potential outliers (Figure 15).

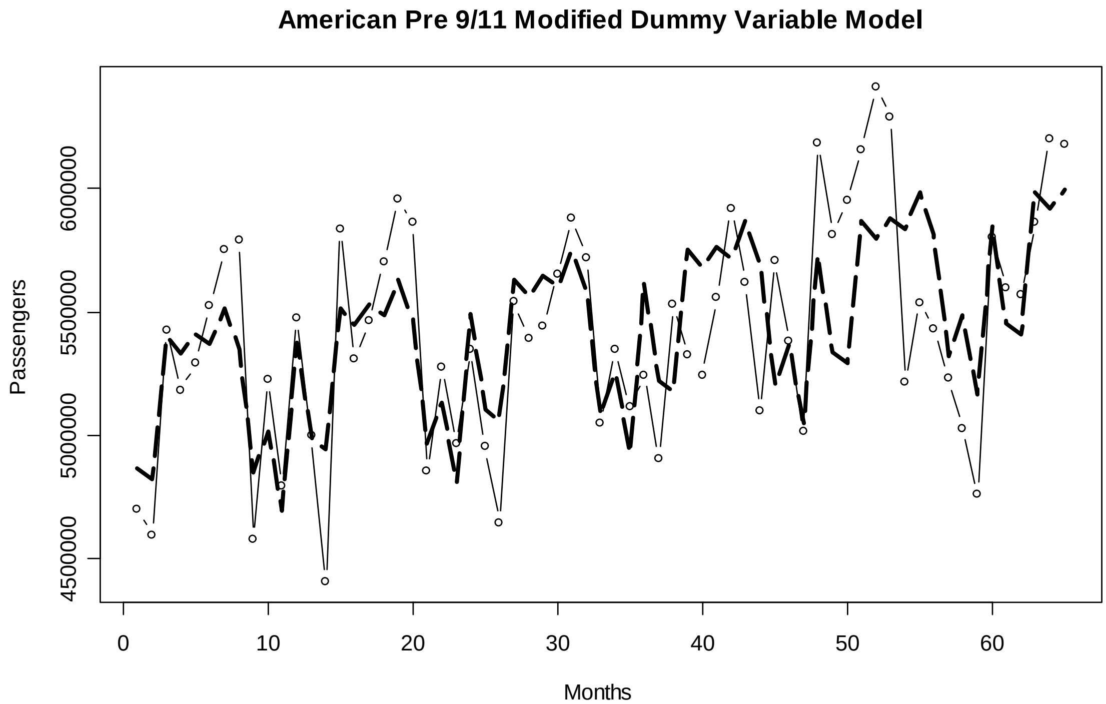

Figure 15: American Airlines Pre 9/11 Modified Data Modeled with Dummy Variables.

According to the F statistic of 5.235 and corresponding p-value 1.111e-05 (Output 11), there is a strong relationship between number of passengers and months. The regression equation for the American Airlines dummy variable modified model is of the form (Output 11):

$$\begin{aligned} \hat{y}_t &= 4860027 + 9731t - 54201t_{s2} + 509888t_{s3} + 431978t_{s4} + 502675t_{s5} \\ &+ 449467t_{s6} + 586561t_{s7} + 409253t_{s8} - 95968t_{s9} + 60708t_{s10} - 270432t_{s11} \\ &+ 400884t_{s12} \end{aligned} \tag{49}$$

```
Call:
lm(formula = Passengersbf ~ Monthsb4 + period)

Residuals:
    Min      1Q  Median      3Q     Max 
-621060 -201876   10903  197281  656327 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  4860027     154579  31.440  < 2e-16 ***
Monthsb4        9731       2245   4.335 6.71e-05 ***
periodFeb     -54201     195212  -0.278   0.7824    
periodMar     509888     195250   2.611   0.0118 *  
periodApr     431978     195315   2.212   0.0314 *  
periodMay     502675     195405   2.572   0.0130 *  
periodJun     449467     204738   2.195   0.0326 *  
periodJul     586561     204726   2.865   0.0060 ** 
periodAug     409253     204738   1.999   0.0509 .  
periodSep     -95968     204775  -0.469   0.6413    
periodOct      60708     204837   0.296   0.7681    
periodNov    -270432     204923  -1.320   0.1927    
periodDec     400884     205034   1.955   0.0559 .  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1 

Residual standard error: 338100 on 52 degrees of freedom
Multiple R-Squared: 0.5471,     Adjusted R-squared: 0.4426 
F-statistic: 5.235 on 12 and 52 DF,  p-value: 1.111e-05
```

Output 11: Summary on American Airlines Modified Dummy Variable Model.

The Shapiro-Wilk test does not give enough evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(passb4dum.lm) 
W = 0.9863, p-value = 0.6928
```

The Breusch-Pagan test does not give enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  passb4dum.lm 
BP = 18.4562, df = 12, p-value = 0.1025
```

The Durbin-Watson test gives enough evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  passb4dum.lm 
DW = 1.462, p-value = 0.04186
alternative hypothesis: true autocorelation is not 0
```

Visual observation of Figures 13 and 15 would indicate the prior a better fit. The AIC of each model was turned to for confirmation.

```
AIC(passb4dum.lm)
[1] 1867.917
AIC(passb4dum.lm)
[1] 1852.998
```

The latter of the two, which is the modified model, is slightly better. Therefore, any further analysis was done with the modified model.

Turning to the post 9/11 data for American Airlines, the model was fit with dummy variables (Figure 16).

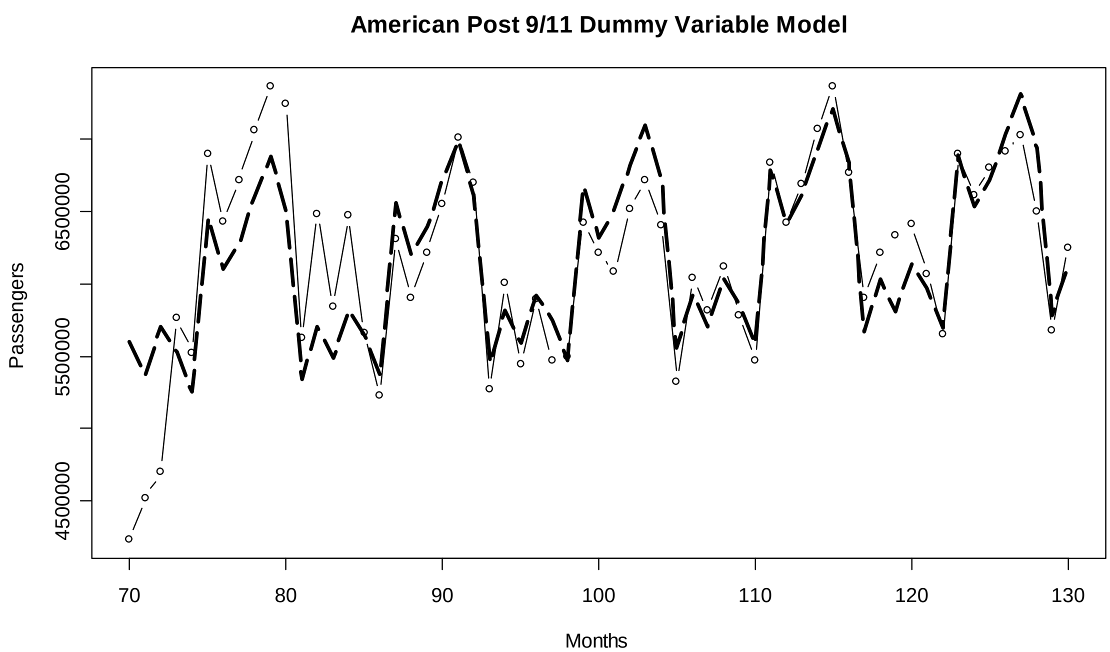

Figure 16: American Airlines Post 9/11 Data Modeled with Dummy Variables.

The summary Output (Output 12) with an *F* statistic of 9.06 and corresponding p-value 9.69e-09 indicates a strong relationship between number of passengers and month. The regression equation for the American Airlines post 9/11 data is of the form (Output 12):

$$\begin{aligned}\hat{y}_t &= 4872318 + 9068t - 286425t_{s2} + 904193t_{s3} + 537281t_{s4} + 714954t_{s5} \\ &+ 1025981t_{s6} + 1292819t_{s7} + 910632t_{s8} - 263110t_{s9} + 93016t_{s10} - 139620t_{s11} \\ &+ 180909t_{s12}\end{aligned} \tag{50}$$

```
Call:
lm(formula = Passengersafter ~ Monthsafter + perioda)

Residuals:
     Min       1Q   Median       3Q      Max
-1368447  -175500    17666   186730   778792

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  4872318     353384  13.788  < 2e-16 ***
Monthsafter     9068       3090   2.935  0.00511 **
periodaFeb   -286425     264802  -1.082  0.28481
periodaMar    904193     264857   3.414  0.00131 **
periodaApr    537281     264947   2.028  0.04814 *
periodaMay    714954     265073   2.697  0.00962 **
periodaJun   1025981     265235   3.868  0.00033 ***
periodaJul   1292819     265433   4.871 1.25e-05 ***
periodaAug    910632     265666   3.428  0.00126 **
periodaSep   -263110     265936  -0.989  0.32744
periodaOct     93016     253681   0.367  0.71548
periodaNov   -139620     264857  -0.527  0.60052
periodaDec    180909     264802   0.683  0.49777
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 418700 on 48 degrees of freedom
Multiple R-Squared: 0.6937,     Adjusted R-squared: 0.6172
F-statistic:  9.06 on 12 and 48 DF,  p-value: 9.69e-09
```

Output 12: Summary on American Airlines Post 9/11 Data Modeled with Dummy Variables.

The Shapiro-Wilk test shows there is evidence enough to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(passafterdum.lm)
W = 0.9385, p-value = 0.004239
```

The Breusch-Pagan test indicates that there is enough evidence to reject the constant variance assumption.

```
    studentized Breusch-Pagan test
data:  passafterdum.lm
BP = 24.1388, df = 12, p-value = 0.01947
```

The Durbin-Watson test indicates that there is strong evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  passafterdum.lm
DW = 0.56, p-value = 1.213e-09
alternative hypothesis: true autocorelation is not 0
```

The model is then checked for possible outliers using residual plots. Observations 1, 2, and 3 were all singled out and therefore counted as a possible outlier block of the initial shock recovery period (Figure 17).

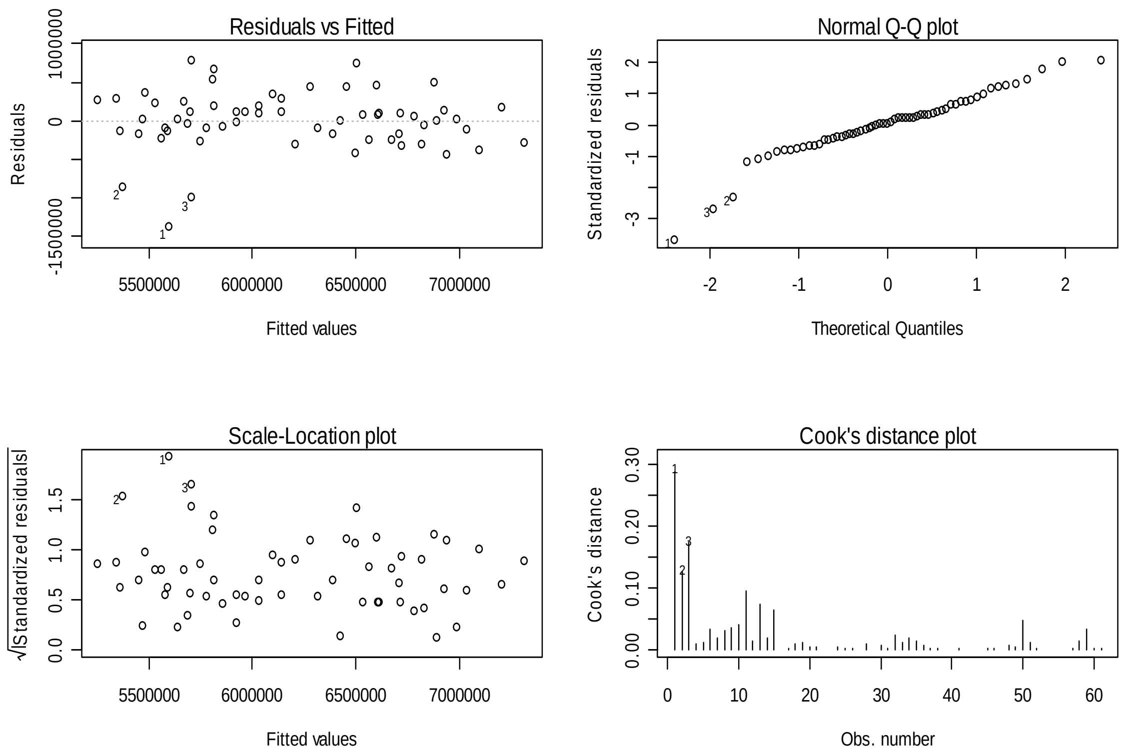

Figure 17: American Airlines Post 9/11 Dummy Variable Residual Plots.

Cook’s Distance statistic did not indicate any influential outliers.

```
cooksD=cooks.distance(passafterdum.lm)
f0.50=qf(0.5,df1=5,df2=24)
cooksD[which(cooksD>f0.50)]
named numeric (0)
```

The Cook’s Distance plot (Figure 17) also indicated the absence of influential outliers.

The analysis on the American Airlines post 9/11 data excluding the first three months was then run and plotted (Figure 18). The summary (Output 13) shows that with an F statistic of 16.96 and corresponding p-value 6.892e-13 there is strong evidence that number of passengers is related to month. Also, the regression equation for the American Airlines post 9/11 modified data is of the form:

$$\hat{y}_t = 6079950 + 2437t + 184304t_{s2} + 501961t_{s3} + 775429t_{s4} + 399873t_{s5} \\ - 767239t_{s6} - 130793t_{s7} - 456425t_{s8} - 92278t_{s9} - 5557172t_{s10} - 836967t_{s11} \\ + 360282t_{s12} \tag{51}$$

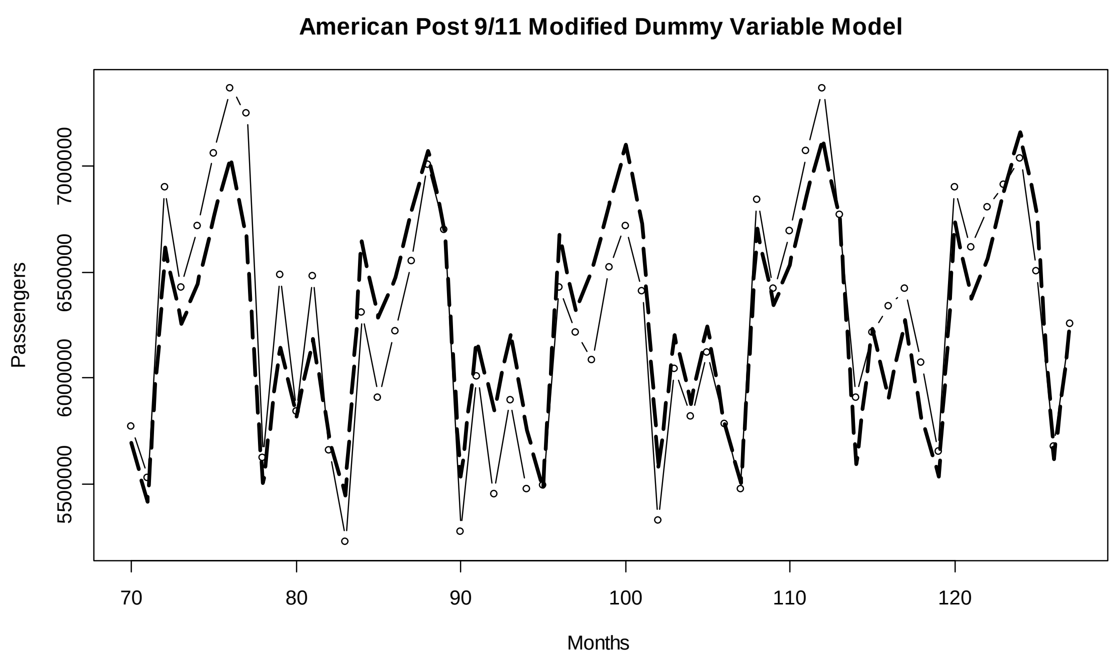

Figure 18: American Airlines Post 9/11 Modified Dummy Variable Model.

```
Call:
lm(formula = Passengersafter ~ Monthsafter + perioda)

Residuals:
    Min      1Q  Median      3Q     Max
-421629 -230829    9434  169642  580127

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  6079950     241186  25.209  < 2e-16 ***
Monthsafter     2437       2151   1.133  0.26309
periodaFeb    184304     171201   1.077  0.28743
periodaMar    501961     171241   2.931  0.00529 **
periodaApr    775429     171309   4.527 4.37e-05 ***
periodaMay    399873     171403   2.333  0.02418 *
periodaJun   -767239     171524  -4.473 5.19e-05 ***
periodaJul   -130793     171673  -0.762  0.45011
periodaAug   -456425     181584  -2.514  0.01560 *
periodaSep    -92278     181622  -0.508  0.61388
periodaOct   -557172     171309  -3.252  0.00217 **
periodaNov   -836967     171241  -4.888 1.34e-05 ***
periodaDec    360282     171201   2.104  0.04096 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 270700 on 45 degrees of freedom
Multiple R-Squared: 0.8189,     Adjusted R-squared: 0.7707
F-statistic: 16.96 on 12 and 45 DF,  p-value: 6.892e-13
```

Output 13: Summary of American Airlines Post 9/11 Modified Data.

The Shapiro-Wilk test does not give enough evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(passafterdum.lm)
W = 0.9721, p-value = 0.201
```

The Breusch-Pagan test does not give enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  passafterdum.lm
BP = 12.79, df = 12, p-value = 0.3845
```

The Durbin-Watson test gives sufficient evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  passafterdum.lm
DW = 0.6354, p-value = 4.505e-08
alternative hypothesis: true autocorelation is not 0
```

The AIC statistics for the post 9/11 models were then compared.

```
AIC(passafterdum.lm)
[1] 1765.758
AIC(passafterdum.lm)
[1] 1628.882
```

The latter statistic representing the post 9/11 modified data was deemed the better fit. Visual examination (Figures 16 & 18) agrees with this finding. Therefore, the American Airlines post 9/11 modified data were used for comparison.

The pre and post 9/11 modified models were compared. Using the regression coefficient estimates, intervals were created using $\pm 2 \cdot$ standard deviation. The intervals for the pre 9/11 data were compared to the post 9/11 intervals (Table 7). The majority of the post 9/11 intervals were overlapped that of the pre 9/11 intervals.

Regression Coefficient Interval Comparison

American Pre 9/11 Intervals			**American Post 9/11 Intervals**			
Estimate	-2*se	+2*se	Estimate	-2*se	+2*se	
4860027.00	4550869.00	5169185.00	6079950.00	5597578.00	6562322.00	above
9731.00	**5241.00**	**14221.00**	2437.00	**-1865.00**	**6739.00**	**overlap**
-54201.00	**-444625.00**	**336223.00**	184304.00	**-158098.00**	**526706.00**	**overlap**
509888.00	**119388.00**	**900388.00**	501961.00	**159479.00**	**844443.00**	**overlap**
431978.00	**41348.00**	**822608.00**	775429.00	**432811.00**	**1118047.00**	**overlap**
502675.00	**111865.00**	**893485.00**	399873.00	**57067.00**	**742679.00**	**overlap**
449467.00	39991.00	858943.00	-767239.00	-1110287.00	-424191.00	below
586561.00	**177109.00**	**996013.00**	-130793.00	**-474139.00**	**212553.00**	**overlap**
409253.00	-223.00	818729.00	-456426.00	-819594.00	-93258.00	below
-95968.00	**-505518.00**	**313582.00**	-92278.00	**-455522.00**	**270966.00**	**overlap**
60708.00	**-348966.00**	**470382.00**	-557172.00	**-899790.00**	**-214554.00**	**overlap**
-270432.00	**-680278.00**	**139414.00**	-836967.00	**-1179449.00**	**-494485.00**	**overlap**
400884.00	**-9184.00**	**810952.00**	360282.00	**17880.00**	**702684.00**	**overlap**

Table 7: American Airlines Pre and Post 9/11 Regression Coefficient Intervals.

Another comparison made was that of the fitted values. The fitted values of the post 9/11 data were compared to that of the pre 9/11 data (Tables 8,9). The majority of the post 9/11 fitted values were above that of the pre 9/11 fitted value for August 2001.

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
4869758	4825289	5399108	5330930	5411358	5367881	5514706	5347130	4851640	5018048
11	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
4696639	5377685	4986533	4942064	5515883	5447705	5528133	5484656	5631481	5463905
21	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
4968415	5134823	4813414	5494461	5103309	5058839	5632659	5564480	5644909	5601432
31	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**
5748257	5580680	5085191	5251598	4930189	5611236	5220084	5175614	5749434	5681256
41	**42**	**43**	**44**	**45**	**46**	**47**	**48**	**49**	**50**
5761684	5718207	5865032	5697456	5201966	5368374	5046965	5728011	5336859	5292390
51	**52**	**53**	**54**	**55**	**56**	**57**	**58**	**59**	**60**
5866209	5798031	5878459	5834982	5981807	5814231	5318741	5485149	5163740	5844787
61	**62**	**63**	**64**	**65**					
5453634	5409165	5982985	5914806	**5995234**					

Table 8: American Airlines Pre 9/11 Modified Model Fitted Values.

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
5693390	5416032	**6615718**	**6257874**	**6444614**	**6764709**	**7040614**	**6667495**	5502821	**6141704**
11	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
5818510	**6185094**	5722637	5445280	**6644966**	**6287121**	**6473862**	**6793957**	**7069862**	**6696743**
21	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
5532069	**6170952**	5847757	**6214342**	5751885	5474527	**6674214**	**6316369**	**6503110**	**6823205**
31	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**
7099110	**6725991**	5561317	**6200199**	5877005	**6243590**	5781132	5503775	**6703461**	**6345616**
41	**42**	**43**	**44**	**45**	**46**	**47**	**48**	**49**	**50**
6532357	**6852452**	**7128357**	**6755238**	5590564	**6229447**	5906252	**6272837**	5810380	5533022
51	**52**	**53**	**54**	**55**	**56**	**57**	**58**		
6732709	**6374864**	**6561605**	**6881700**	**7157605**	**6784486**	5619812	**6258694**		

Table 9: American Airlines Post 9/11 Modified Model Fitted Values.

4.2.3 Delta Airlines

Delta Airlines pre 9/11 data were then subjected to the dummy variable model for the seasonal variation (Figure 19).

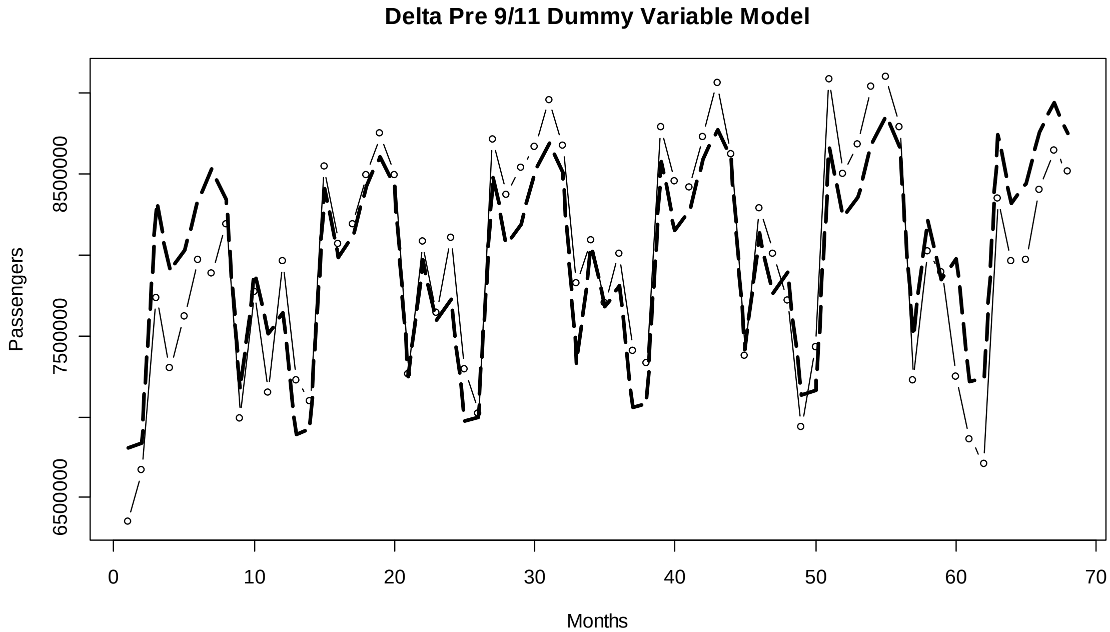

Figure 19: Delta Airlines Pre 9/11 Data Modeled with Dummy Variables.

The summary data (Output 14) reveals that with an *F* statistic of 18.21 and corresponding p-value 5.214e-15 there is a strong relationship between number of passengers and month. The regression equation for the Delta Airlines dummy variable model is of the form:

$$\begin{aligned}\hat{y}_t &= 6799448 + 6900t + 22260t_{s2} + 1510049t_{s3} + 1077616t_{s4} 1194792t_{s5} \\ &+ 1505153t_{s6} 1681583t_{s7} + 1486924t_{s8} + 307356t_{s9} + 1017206t_{s10} + 638529t_{s11} \\ &+ 760737t_{s12}\end{aligned} \quad (52)$$

```
Call:
lm(formula = Passengersbf ~ Monthsb4 + period)

Residuals:
    Min      1Q  Median      3Q     Max
-725421 -209335   73194  245159  488962

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  6799448     151222  44.963  < 2e-16 ***
Monthsb4        6900       2081   3.316 0.001622 **
periodFeb      22260     193436   0.115 0.908805
periodMar    1510049     193470   7.805 1.81e-10 ***
periodApr    1077616     193526   5.568 7.93e-07 ***
periodMay    1194792     193604   6.171 8.50e-08 ***
periodJun    1505153     193705   7.770 2.07e-10 ***
periodJul    1681583     193828   8.676 7.01e-12 ***
periodAug    1486924     193973   7.666 3.06e-10 ***
periodSep     307356     202909   1.515 0.135561
periodOct    1017206     202962   5.012 5.94e-06 ***
periodNov     638529     203037   3.145 0.002679 **
periodDec     760737     203132   3.745 0.000434 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 335000 on 55 degrees of freedom
Multiple R-Squared: 0.7989,     Adjusted R-squared: 0.755
F-statistic: 18.21 on 12 and 55 DF,  p-value: 5.214e-15
```

Output 14: Summary of Delta Airlines Pre 9/11 Dummy Variable Model.

The Shapiro-Wilk test indicates there is strong evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(passb4dum.lm)
W = 0.9352, p-value = 0.001582
```

The Breusch-Pagan test does not indicate strong enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  passb4dum.lm
BP = 12.1269, df = 12, p-value = 0.4355
```

The Durbin-Watson test gives enough evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  passb4dum.lm
DW = 0.5797, p-value = 2.053e-10
alternative hypothesis: true autocorelation is not 0
```

The data were then tested for possible outliers using residual plots. The plots all singled out observations 4, 7, and 60 (Figure 20). These observations were deemed possible outliers. Using a leverage plot, the data were tested for influential outliers; the plot (Figure 20) indicated none.

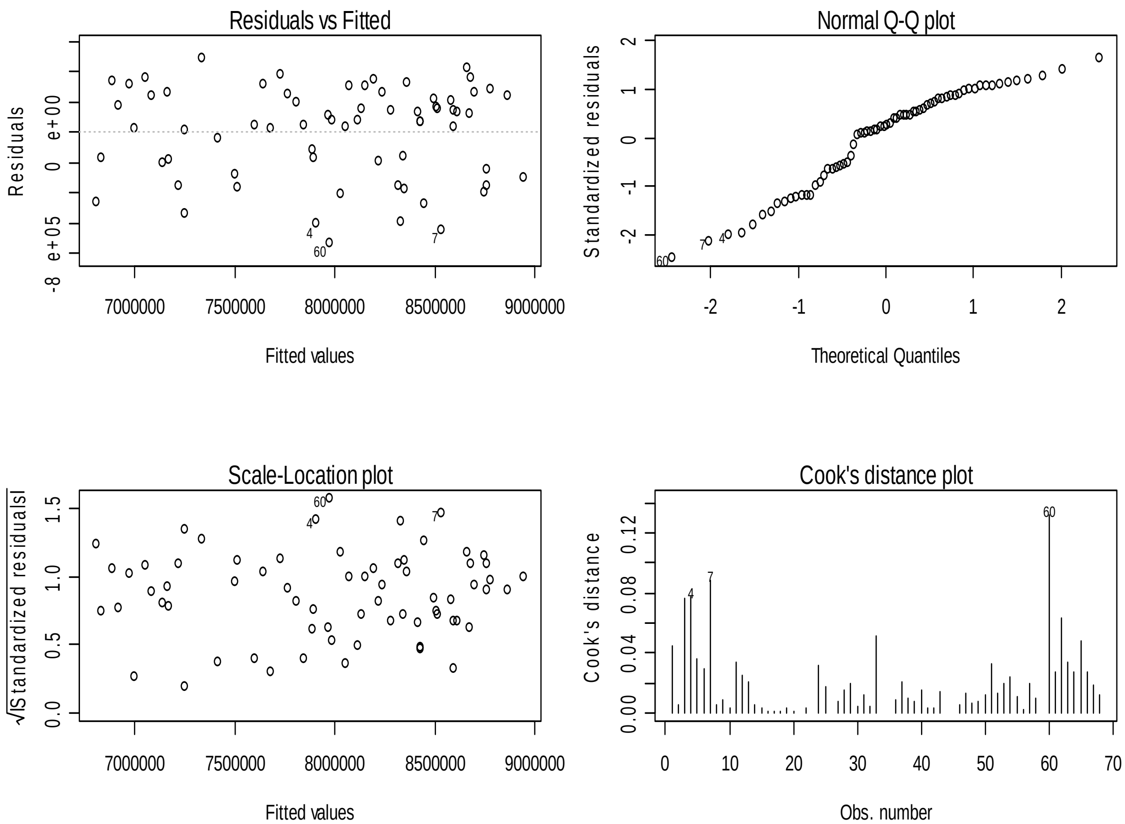

Figure 20: Delta Airlines Pre 9/11 Dummy Variable Model Residual Plots.

The Cook's Distance test was then run to confirm the absence of possible outliers. It confirmed the previous findings.

```
cooksD=cooks.distance(passb4dum.lm)
f0.50=qf(0.5, df1=5,df2=24)
cooksD[which(cooksD>f0.50)]
named numeric(0)
```

The analyses on the Delta Airlines pre 9/11 modified data were then applied (Figure 21).

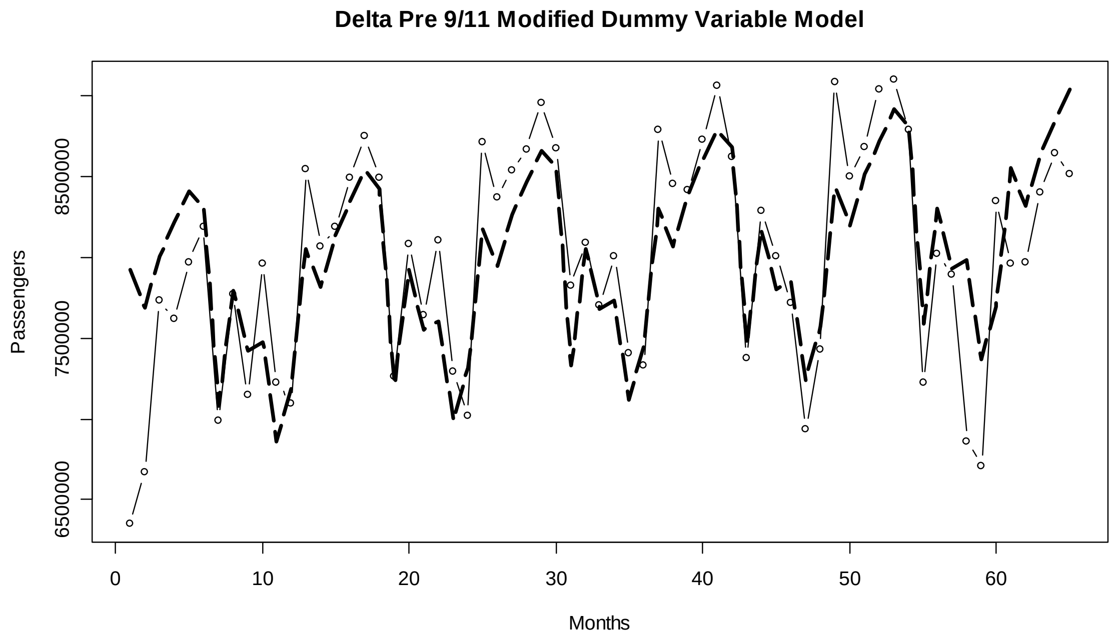

Figure 21: Delta Airlines Pre 9/11 Modified Data Dummy Variable Plot.

The summary data (Output 15) showed an *F* statistic of 7.205 with a corresponding p-value of 1.582e-07 indicating a relationship between number of passengers and month. The regression equation for the Delta Airlines pre 9/11 modified data is of the form:

$$\begin{aligned}\hat{y}_t &= 7914021 + 10554t - 244784t_{s2} + 66042t_{s3} + 261460t_{s4} + 445044t_{s5} \\ &+ 324633t_{s6} - 906666t_{s7} - 200470t_{s8} - 582799t_{s9} - 541294t_{s10} - 1168179t_{s11} \\ &- 848731t_{s12}\end{aligned} \tag{53}$$

```
Call:
lm(formula = Passengersbf ~ Monthsb4 + period)

Residuals:
     Min       1Q   Median       3Q      Max
-1574761  -203231    54911   277884   654574

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  7914021     211544  37.411  < 2e-16 ***
Monthsb4       10554       3072   3.435 0.001172 **
periodFeb    -244784     267150  -0.916 0.363750
periodMar      66042     267203   0.247 0.805758
periodApr     261460     267291   0.978 0.332512
periodMay     445044     267415   1.664 0.102078
periodJun     324633     280187   1.159 0.251901
periodJul    -906666     280170  -3.236 0.002110 **
periodAug    -200470     280187  -0.715 0.477510
periodSep    -582799     280238  -2.080 0.042504 *
periodOct    -541294     280322  -1.931 0.058946 .
periodNov   -1168179     280440  -4.166 0.000117 ***
periodDec    -848731     280591  -3.025 0.003860 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 462700 on 52 degrees of freedom
Multiple R-Squared: 0.6244,     Adjusted R-squared: 0.5378
F-statistic: 7.205 on 12 and 52 DF,  p-value: 1.582e-07
```

Output 15: Summary of Delta Airlines Pre 9/11 Modified Dummy Variable Model.

The Shapiro-Wilk test indicates enough evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(passb4dum.lm)
W = 0.9192, p-value = 0.0004123
```

The Breusch-Pagan test does not give enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  passb4dum.lm
BP = 18.8938, df = 12, p-value = 0.09112
```

The Durbin-Watson test indicates there is enough evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  passb4dum.lm
DW = 0.9989, p-value = 7.213e-05
alternative hypothesis: true autocorelation is not 0
```

A visual comparison of Figures 19 and 21 was deemed inconclusive. The AIC of both the pre 9/11 original data and the pre 9/11 modified data were relied upon.

```
AIC(passb4dum.lm)
[1] 1936.733
AIC(passb4dum.lm)
[1] 1893.782
```

The decision was made that the latter model was a better fit and would henceforth be used for further analysis.

Attention was then turned to the Delta Airlines post 9/11 data. The data were then fit with a dummy variable model representing the seasonal variation (Figure 22).

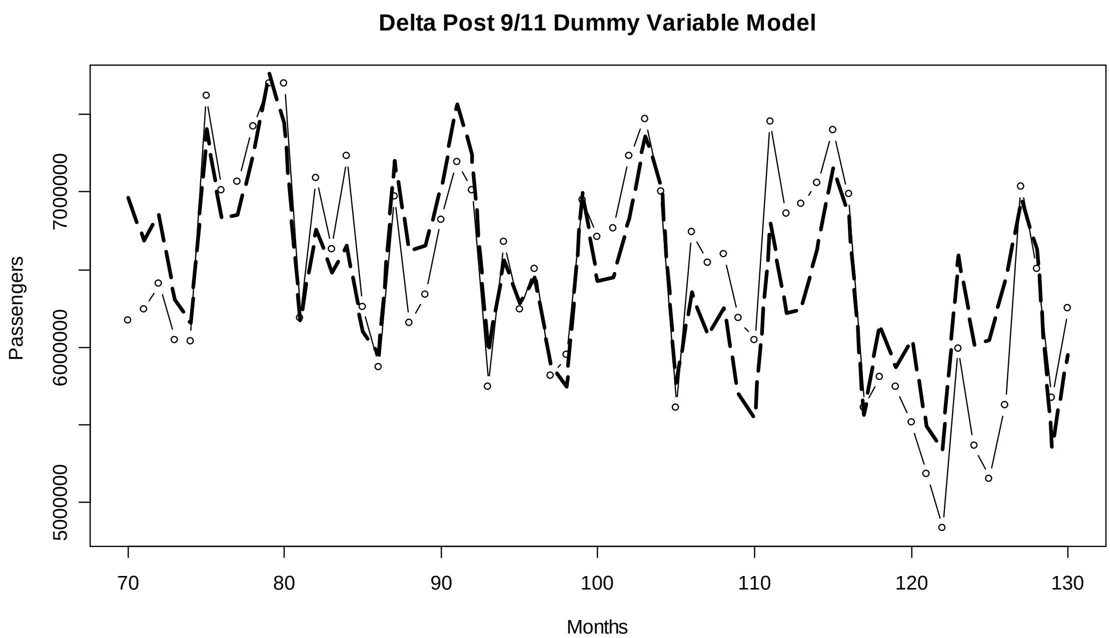

Figure 22: Delta Airlines Post 9/11 Dummy Variables Model.

The summary data (Output 16) indicates with an F statistic of 9.113 and corresponding p-value 8.875e-09 that there is of a relationship between number of passengers and month. The regression equation for the Delta Airlines post 9/11 data dummy variable model is of the form:

$$\hat{y}_t = 7539592 - 16924t - 134663_{t2} + 1134599t_{s3} + 574685t_{s4} + 617786t_{s5} + 1017892t_{s6} + 1561392t_{s7} + 1258841t_{s8} + 3010t_{s9} + 609238t_{s10} + 348569t_{s11} + 536623t_{s12} \quad (54)$$

```
Call:
lm(formula = Passengersafter ~ Monthsafter + perioda)

Residuals:
    Min      1Q  Median      3Q     Max
-891418 -231983   48532  289496  677136

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  7539592     359131  20.994  < 2e-16 ***
Monthsafter   -16924       3140  -5.390 2.11e-06 ***
periodaFeb   -134663     269109  -0.500 0.619077
periodaMar   1134599     269164   4.215 0.000110 ***
periodaApr    574685     269255   2.134 0.037949 *
periodaMay    617786     269383   2.293 0.026251 *
periodaJun   1017892     269548   3.776 0.000439 ***
periodaJul   1561392     269749   5.788 5.27e-07 ***
periodaAug   1258841     269987   4.663 2.52e-05 ***
periodaSep      3010     270260   0.011 0.991159
periodaOct    609238     257807   2.363 0.022218 *
periodaNov    348569     269164   1.295 0.201513
periodaDec    536623     269109   1.994 0.051842 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 425500 on 48 degrees of freedom
Multiple R-Squared: 0.695,      Adjusted R-squared: 0.6187
F-statistic: 9.113 on 12 and 48 DF,  p-value: 8.875e-09
```

Output 16: Summary of Delta Airlines Post 9/11 Dummy Variable Model.

The Shapiro-Wilk test does not give enough evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(passafterdum.lm)
W = 0.9796, p-value = 0.3992
```

The Breusch-Pagan test does not give enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  passafterdum.lm
BP = 19.2637, df = 12, p-value = 0.08236
```

The Durbin-Watson test indicates that there is enough evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  passafterdum.lm
DW = 0.4361, p-value = 4.491e-12
alternative hypothesis: true autocorelation is not 0
```

The Delta Airlines post 9/11 data were then tested for a possible outlier block representing an initial shock recovery period using residual plots. The only observation singled out by all plots (Figure 23) that would correspond to this period was observation 1. The Cook's Distance plot (Figure 23) did not indicate any influential outliers. Thus the Cook's Distance test was relied upon and confirmed this indication.

```
cooksD=cooks.distance(passafterdum.lm)
f0.50=qf(0.5,df1=5,df2=24)
cooksD[which(cooksD>f0.50)]
named numeric(0)
```

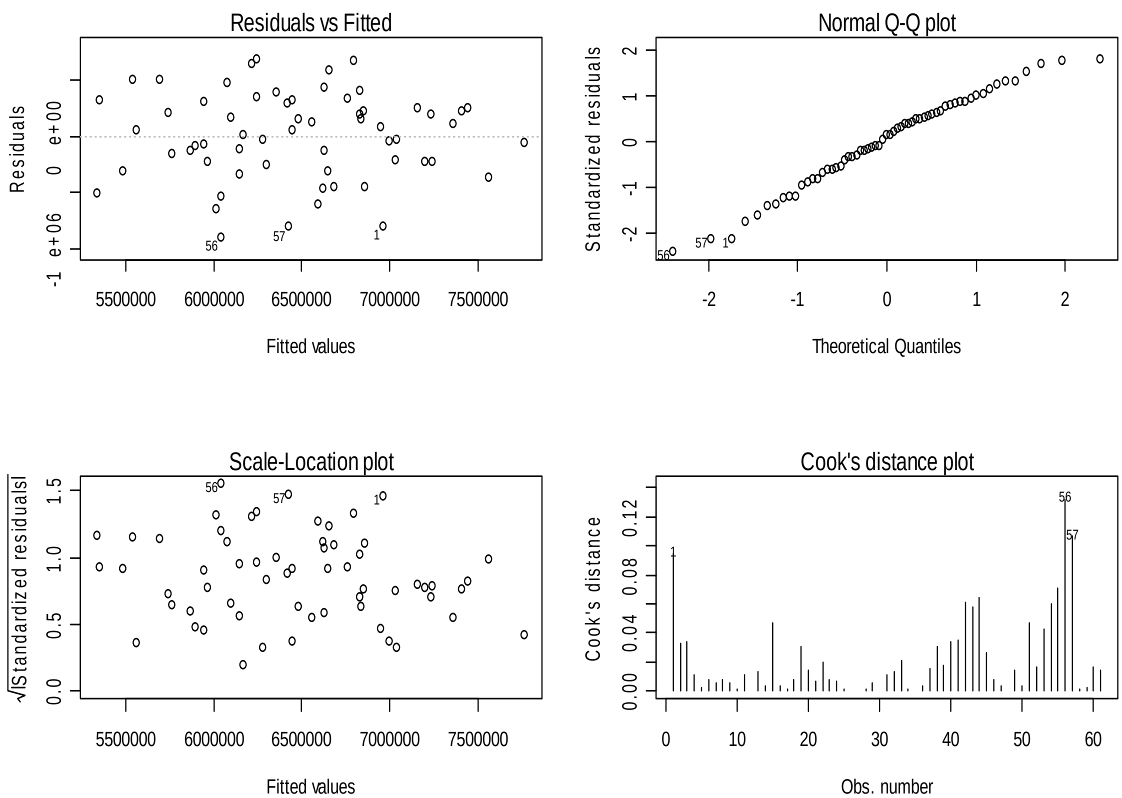

Figure 23: Delta Airlines Post 9/11 Dummy Variable Residual Plots.

Observation 1 was then excluded and the analysis was applied to the data. This new data set was fit with dummy variables (Figure 24).

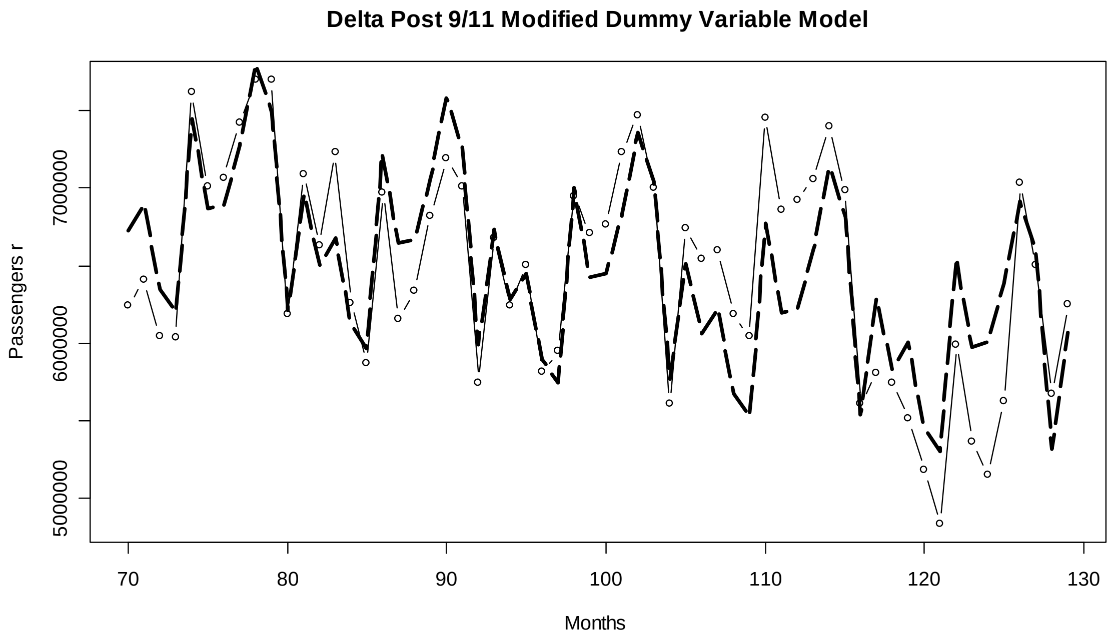

Figure 24: Delta Airlines Post 9/11 Modified Dummy Variable Plot.

The summary data (Output 17), with an *F* statistic of 10.19 and corresponding p-value 1.922e-09 indicates that there is a relationship between number of passengers and month. The regression equation for Delta Airlines post 9/11 modified data is of the form:

$$\hat{y}_t = 7548831 - 18582t + 1270920t_{s2} + 712664t_{s3} + 757423t_{s4} + 1159187t_{s5} + 1704345t_{s6} + 1403452t_{s7} + 149279t_{s8} + 916334t_{s9} + 478258t_{s10} + 667970t_{s11} + 133005t_{s12} \quad (55)$$

```
Call:
lm(formula = Passengersafter ~ Monthsafter + perioda)

Residuals:
    Min      1Q  Median      3Q     Max
-851626 -248598   12820  214542  697032

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  7548831     353339  21.364  < 2e-16 ***
Monthsafter   -18582       3115  -5.965 3.04e-07 ***
periodaFeb   1270920     259012   4.907 1.15e-05 ***
periodaMar    712664     259068   2.751 0.008414 **
periodaApr    757423     259162   2.923 0.005323 **
periodaMay   1159187     259293   4.471 4.91e-05 ***
periodaJun   1704345     259462   6.569 3.69e-08 ***
periodaJul   1403452     259667   5.405 2.11e-06 ***
periodaAug    149279     259910   0.574 0.568470
periodaSep    916334     260190   3.522 0.000965 ***
periodaOct    478258     259162   1.845 0.071285 .
periodaNov    667970     259068   2.578 0.013120 *
periodaDec    133005     259012   0.514 0.610001
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 409500 on 47 degrees of freedom
Multiple R-Squared: 0.7224,     Adjusted R-squared: 0.6515
F-statistic: 10.19 on 12 and 47 DF,  p-value: 1.922e-09
```

Output 17: Summary of Delta Airlines post 9/11 Modified Dummy Variable Model.

The Shapiro-Wilk test does not indicate enough evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(passafterdum.lm)
W = 0.9839, p-value = 0.613
```

The Breusch-Pagan test indicates that there is sufficient evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  passafterdum.lm
BP = 25.4832, df = 12, p-value = 0.01269
```

The Durbin-Watson test indicates enough evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  passafterdum.lm
DW = 0.4633, p-value = 2.98e-11
alternative hypothesis: true autocorelation is not 0
```

To decide which post 9/11 model for Delta Airlines was the better fit, Figures 22 and 24 were visually compared. The two models being so close, it was necessary to rely on the AIC for comparison. The AIC indicated that the latter value for the post 9/11 modified data was a better fit.

```
AIC(passafterdum.lm)
[1] 1767.726
AIC(passafterdum.lm)
[1] 1734.345
```

Therefore, this model would be used for further comparison.

The pre 9/11 modified model and the post 9/11 modified model were compared. Using the regression coefficient estimates, intervals were created using $\pm 2 \cdot$ standard deviation. The intervals for the pre 9/11 data were compared to the post 9/11 intervals (Table 10). The majority of the post 9/11 intervals were overlapped that of the pre 9/11 intervals.

Regression Coefficient Interval Comparison

Delta Pre 9/11 Intervals			**Delta Post 9/11 Intervals**			
Estimate	-2*se	+2*se	Estimate	-2*se	+2*se	
7914021.00	**7490933.00**	**8337109.00**	7548831.00	**6842153.00**	**8255509.00**	**overlap**
10554.00	4410.00	16698.00	-18582.00	-24812.00	-12352.00	below
-244784.00	-779084.00	289516.00	1270920.00	752896.00	1788944.00	above
66042.00	**-468364.00**	**600448.00**	712664.00	**194528.00**	**1230800.00**	**overlap**
261460.00	**-273122.00**	**796042.00**	757423.00	**239099.00**	**1275747.00**	**overlap**
445044.00	**-89786.00**	**979874.00**	1159187.00	**640601.00**	**1677773.00**	**overlap**
324633.00	-235741.00	885007.00	1704345.00	1185421.00	2223269.00	above
-906666.00	-1467006.00	-346326.00	1403452.00	884118.00	1922786.00	above
-200470.00	**-760844.00**	**359904.00**	149279.00	**-370541.00**	**669099.00**	**overlap**
-582799.00	-1143275.00	-22323.00	916334.00	395954.00	1436714.00	above
-541294.00	**-1101938.00**	**19350.00**	478258.00	**-40066.00**	**996582.00**	**overlap**
-1168179.00	-1729059.00	-607299.00	667970.00	149834.00	1186106.00	above
-848731.00	**-1409913.00**	**-287549.00**	133005.00	**-385019.00**	**651029.00**	**overlap**

Table 10: Delta Airlines Pre and Post 9/11 Regression Coefficient Interval Comparison.

Another comparison made was that of the fitted values. The fitted values of the post 9/11 data were compared to that of the pre 9/11 data (Tables 11,12). The fitted values of the post 9/11 model did not reach or exceed that of the August 2001 fitted value.

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
7924575	7690345	8011724	8217696	8411833	8301976	7081231	7797980	7426204	7478262
11	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
6861931	7191933	8051217	7816987	8138366	8344338	8538475	8428618	7207873	7924622
21	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
7552846	7604905	6988573	7318575	8177860	7943630	8265008	8470980	8665118	8555261
31	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**
7334515	8051265	7679489	7731547	7115216	7445218	8304502	8070272	8391651	8597623
41	**42**	**43**	**44**	**45**	**46**	**47**	**48**	**49**	**50**
8791760	8681903	7461158	8177907	7806131	7858190	7241858	7571860	8431145	8196914
51	**52**	**53**	**54**	**55**	**56**	**57**	**58**	**59**	**60**
8518293	8724265	8918403	8808545	7587800	8304550	7932774	7984832	7368501	7698502
61	**62**	**63**	**64**	**65**					
8557787	8323557	8644936	8850908	**9045045**					

Table 11: Delta Airlines Pre 9/11 Modified Data Fitted Values.

1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
6726352	6897482	6343936	6192349	7444686	6867848	6894026	7277208	7803784	7484309
11	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
6211554	6960027	6503369	6674499	6120952	5969365	7221703	6644865	6671042	7054225
21	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	**30**
7580801	7261325	5988571	6737043	6280385	6451515	5897969	5746382	6998720	6421882
31	**32**	**33**	**34**	**35**	**36**	**37**	**38**	**39**	**40**
6448059	6831241	7357817	7038342	5765587	6514060	6057402	6228532	5674985	5523398
41	**42**	**43**	**44**	**45**	**46**	**47**	**48**	**49**	**50**
6775736	6198898	6225075	6608258	7134834	6815359	5542604	6291077	5834418	6005549
51	**52**	**53**	**54**	**55**	**56**	**57**	**58**	**59**	**60**
5452002	5300415	6552753	5975915	6002092	6385274	6911850	6592375	5319621	6068093

Table 12: Delta Airlines Post 9/11 Modified Data Fitted Values.

The Continental Airlines regression coefficient intervals for the post 9/11 data lie below those of the pre 9/11 data. Also, the fitted values of the post 9/11 data never reach that of the fitted value for August 2001, the month preceding 9/11. This indicates that Continental Airlines is not boarding the number of passengers that it did before 9/11. American Airlines regression coefficient intervals for the post 9/11 data overlap most of the pre 9/11 intervals. Of the 58 post 9/11 fitted values, 39 of them exceed that of August 2001. This can be interpreted in such a way as to say 9/11 did not have much effect on the passenger numbers traveling this airline. The Delta Airlines regression intervals for the post 9/11 data, for the most part, overlap that of the pre 9/11 intervals. The fitted values for the post 9/11 model do not reach or exceed that of August 2001. This indicates that 9/11 had a negative effect on the number of passengers traveling this airline.

4.3 Trigonometric Functions

In this section trigonometric functions are explored in modeling the seasonal variation of the airline data. Each airline was modeled with trigonometric functions then compared to their corresponding dummy variable models to see which was the better fit.

4.3.1 Continental Airlines

The first attempt at modeling the seasonal variation was to use two sets of trigonometric functions (Figure 25).

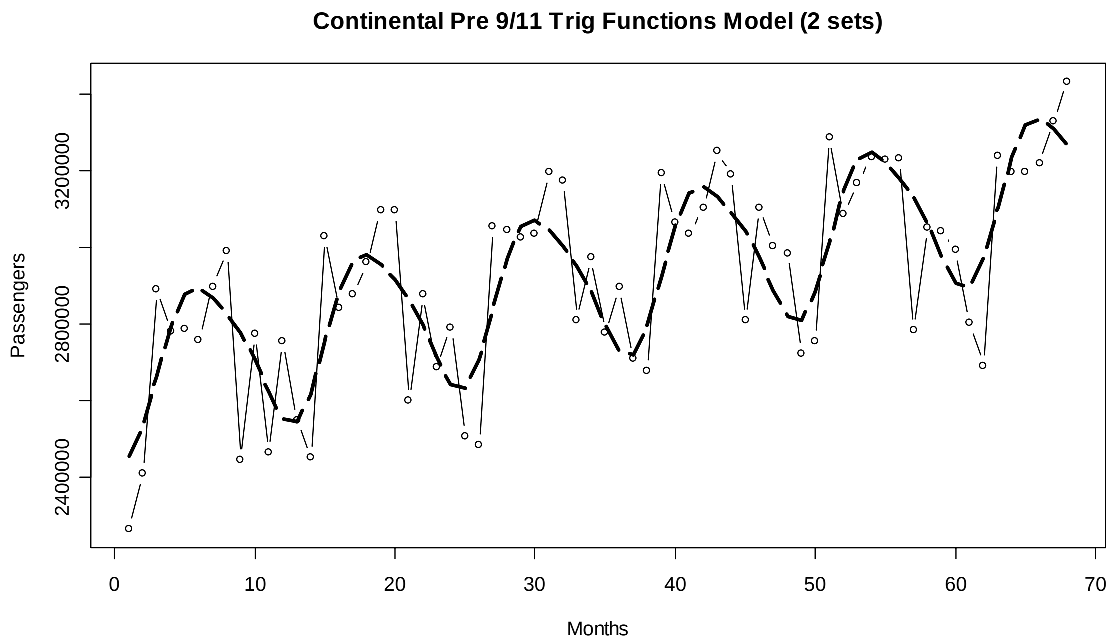

Figure 25: Continental Airlines Pre 9/11 Modeled with 2 Sets of Trigonometric Functions.

The summary data (Output 18) indicates that with an *F* statistic of 24.74 and associated p-value 1.300e-13 there is a strong relationship between number of passengers and month. The regression equation for the Continental Airlines pre 9/11 data modeled with two sets of trigonometric functions is of the form:

$$\hat{y}_t = 2665899.7 + 7372.1t - 35199.4\sin\left(\frac{2\pi t}{12}\right) - 192426.8\cos\left(\frac{2\pi t}{12}\right) - 34149.0\sin\left(\frac{4\pi t}{12}\right) - 8873.6\cos\left(\frac{4\pi t}{12}\right) \quad (56)$$

```
Call:
lm(formula = Passengersbf ~ Monthsb4 + I(sin(2 * pi * Monthsb4/12)) +
    I(cos(2 * pi * Monthsb4/12)) + I(sin(4 * pi * Monthsb4/12)) +
    I(cos(4 * pi * Monthsb4/12)))

Residuals:
    Min      1Q  Median      3Q     Max
-347110 -109503  -10823  121461  279010

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                  2665899.7    38008.7  70.139  < 2e-16 ***
Monthsb4                        7372.1      957.6   7.699 1.32e-10 ***
I(sin(2 * pi * Monthsb4/12))  -35199.4    26604.3  -1.323    0.191
I(cos(2 * pi * Monthsb4/12)) -192426.8    26604.3  -7.233 8.49e-10 ***
I(sin(4 * pi * Monthsb4/12))  -34149.0    26352.5  -1.296    0.200
I(cos(4 * pi * Monthsb4/12))   -8873.6    26734.4  -0.332    0.741
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 154500 on 62 degrees of freedom
Multiple R-Squared: 0.6662,     Adjusted R-squared: 0.6392
F-statistic: 24.74 on 5 and 62 DF,  p-value: 1.300e-13
```

Output 18: Summary of Continental Airlines Pre 9/11 Modeled with 2 Sets of Trigonometric Functions

The Shapiro-Wilk test does not indicate that there is enough evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(passb4trig.lm)
W = 0.9828, p-value = 0.4701
```

The Breusch-Pagan test indicates that there is enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  passb4trig.lm
BP = 21.1294, df = 5, p-value = 0.0007657
```

The Durbin-Watson test indicates that there is enough evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  passb4trig.lm
DW = 2.453, p-value = 0.1743
alternative hypothesis: true autocorelation is not 0
```

The visual examination of the plot (Figure 25) proved a poor model. A second attempt to model the data was made using three sets of trigonometric functions. Upon visual examination (Figure 26), the plot seemed a much better fit.

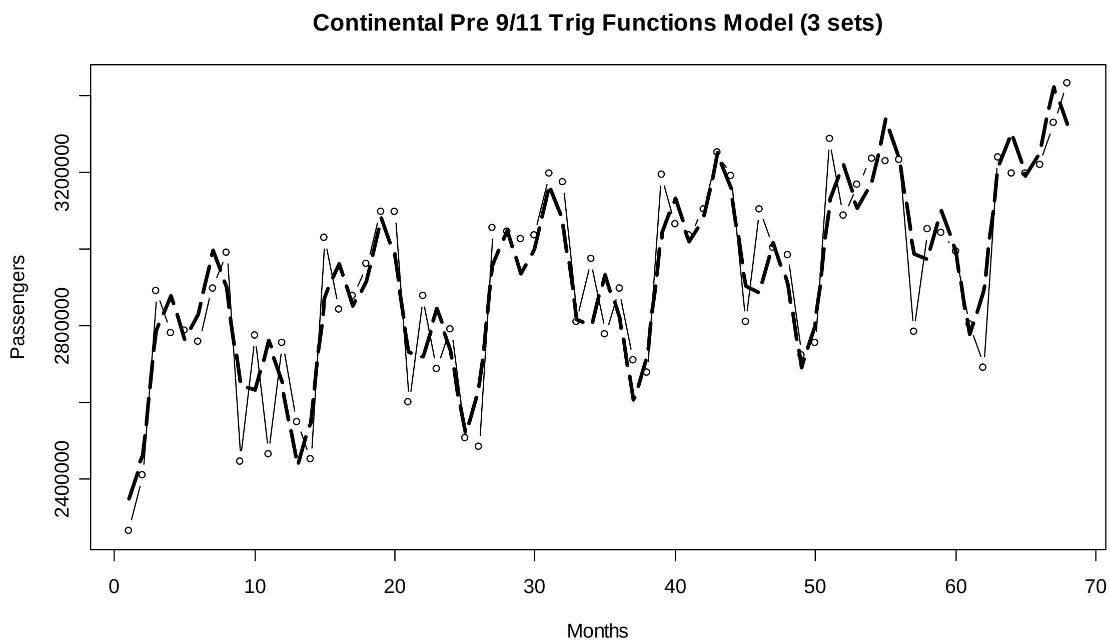

Figure 26: Continental Airlines Pre 9/11 Modeled with 3 Sets of Trigonometric Functions.

The summary data with an *F* statistic of 40.99 and associated p-value 2.2e-16 indicates a strong relationship between passenger numbers and months (Output 19). The regression equation for Continental Airlines pre 9/11 data modeled with three sets of trigonometric functions is of the form:

$$\hat{y}_t = 2674807.0 + 7115.4t - 31631.2\sin\left(\frac{2\pi t}{12}\right) - 188252.7\cos\left(\frac{2\pi t}{12}\right) \\ - 35339.3\sin\left(\frac{4\pi t}{12}\right) + 617.1\cos\left(\frac{4\pi t}{12}\right) - 123389.5\sin\left(\frac{6\pi t}{12}\right) + 76901.0\cos\left(\frac{6\pi t}{12}\right) \quad (57)$$

```
Call:
lm(formula = Passengersbf ~ Monthsb4 + I(sin(2 * pi * Monthsb4/12)) +
    I(cos(2 * pi * Monthsb4/12)) + I(sin(4 * pi * Monthsb4/12)) +
    I(cos(4 * pi * Monthsb4/12)) + I(sin(6 * pi * Monthsb4/12)) +
    I(cos(6 * pi * Monthsb4/12)))

Residuals:
    Min      1Q  Median      3Q     Max
-294621  -87692   14712   82859  215744

Coefficients:
                              Estimate Std. Error t value Pr(>|t|)
(Intercept)                  2674807.0    27836.6  96.089  < 2e-16 ***
Monthsb4                        7115.4      701.5  10.143 1.25e-14 ***
I(sin(2 * pi * Monthsb4/12))  -31631.2    19473.8  -1.624 0.109554
I(cos(2 * pi * Monthsb4/12)) -188252.7    19473.8  -9.667 7.58e-14 ***
I(sin(4 * pi * Monthsb4/12))  -35339.3    19294.2  -1.832 0.071977 .
I(cos(4 * pi * Monthsb4/12))     617.1    19604.3   0.031 0.974993
I(sin(6 * pi * Monthsb4/12)) -123389.5    19431.4  -6.350 3.17e-08 ***
I(cos(6 * pi * Monthsb4/12))   76901.0    19431.4   3.958 0.000203 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 113000 on 60 degrees of freedom
Multiple R-Squared: 0.827,      Adjusted R-squared: 0.8069
F-statistic: 40.99 on 7 and 60 DF,  p-value: < 2.2e-16
```

Output 19: Summary of Continental Airlines Pre 9/11 Modeled with 3 Sets of Trigonometric Functions.

The Shapiro-Wilk test does not give enough evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(passb4trig2.lm)
W = 0.9815, p-value = 0.4076
```

The Breusch-Pagan test indicates that there is enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  passb4trig2.lm
BP = 22.1668, df = 7, p-value = 0.002378
```

The Durbin-Watson test indicates that there is enough evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  passb4trig2.lm
DW = 3.0207, p-value = 0.00013
alternative hypothesis: true autocorelation is not 0
```

A third attempt was made at modeling the seasonal variation with four sets of trig functions. Visually the plots for three and four sets did not seem much different (Figures 26,27), but both seemed much better than two sets (Figure 25).

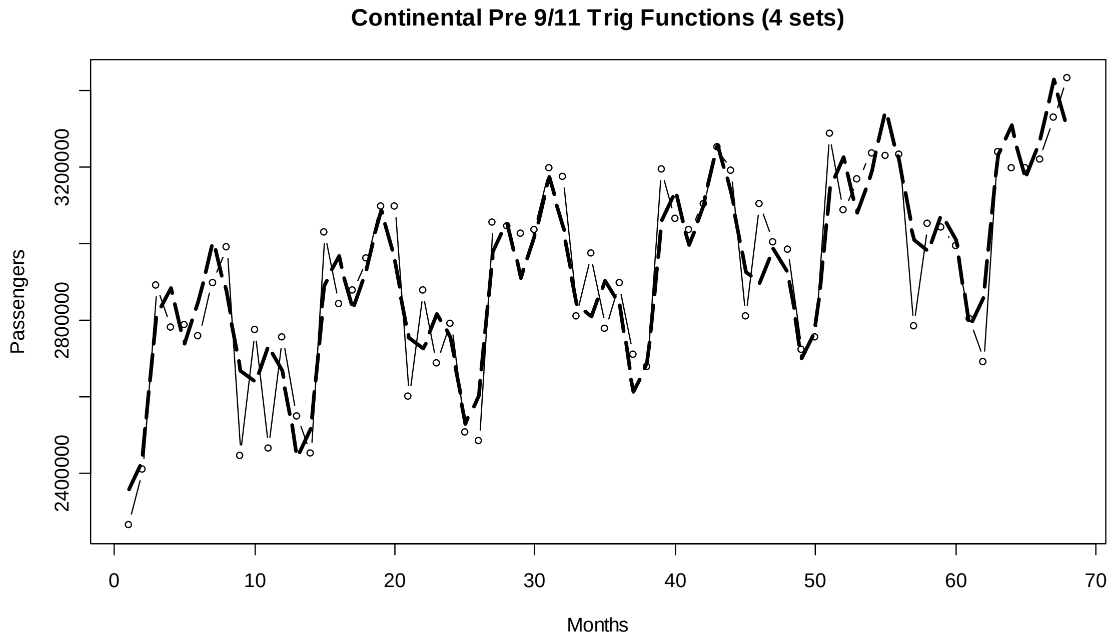

Figure 27: Continental Airlines Pre 9/11 Modeled with 4 Sets of Trigonometric Functions.

The summary data (Output 20) indicates that with an *F* statistic of 32.06 and associated p-value 2.2e-16 there is a strong relationship between number of passengers and month. The regression equation for Continental Airlines pre 9/11 data fit with four sets of trigonometric functions is of the form:

$$\hat{y}_t = 2674552.59 + 7131.08t - 31997.56\sin\left(\frac{2\pi t}{12}\right) - 188262.08\cos\left(\frac{2\pi t}{12}\right)$$
$$- 34835.08\sin\left(\frac{4\pi t}{12}\right) + 55.26\cos\left(\frac{4\pi t}{12}\right) - 121995.20\sin\left(\frac{6\pi t}{12}\right) + \quad (58)$$
$$77241.27\cos\left(\frac{6\pi t}{12}\right) + 18807.68\sin\left(\frac{8\pi t}{12}\right) + 19410.56\cos\left(\frac{8\pi t}{12}\right)$$

```
Call:
lm(formula = Passengersbf ~ Monthsb4 + I(sin(2 * pi * Monthsb4/12)) +
    I(cos(2 * pi * Monthsb4/12)) + I(sin(4 * pi * Monthsb4/12)) +
    I(cos(4 * pi * Monthsb4/12)) + I(sin(6 * pi * Monthsb4/12)) +
    I(cos(6 * pi * Monthsb4/12)) + I(sin(8 * pi * Monthsb4/12)) +
    I(cos(8 * pi * Monthsb4/12)))

Residuals:
    Min      1Q  Median      3Q     Max
-266608  -90857   15593   81301  208880

Coefficients:
                               Estimate Std. Error t value Pr(>|t|)
(Intercept)                  2674552.59   27859.12  96.003  < 2e-16 ***
Monthsb4                        7131.08     702.11  10.157 1.74e-14 ***
I(sin(2 * pi * Monthsb4/12))  -31997.56   19486.39  -1.642 0.105992
I(cos(2 * pi * Monthsb4/12)) -188262.08   19486.39  -9.661 1.10e-13 ***
I(sin(4 * pi * Monthsb4/12))  -34835.08   19311.69  -1.804 0.076451 .
I(cos(4 * pi * Monthsb4/12))      55.26   19623.71   0.003 0.997763
I(sin(6 * pi * Monthsb4/12)) -121995.20   19469.37  -6.266 4.94e-08 ***
I(cos(6 * pi * Monthsb4/12))   77241.27   19469.37   3.967 0.000202 ***
I(sin(8 * pi * Monthsb4/12))   18807.68   19280.81   0.975 0.333383
I(cos(8 * pi * Monthsb4/12))   19410.56   19590.69   0.991 0.325896
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 113100 on 58 degrees of freedom
Multiple R-Squared: 0.8326,     Adjusted R-squared: 0.8067
F-statistic: 32.06 on 9 and 58 DF,  p-value: < 2.2e-16
```

Output 20: Summary of Continental Airlines Pre 9/11 Modeled with 4 Sets of Trigonometric Functions.

The Shapiro-Wilk test indicates that there is not enough evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(passb4trig3.lm)
W = 0.9752, p-value = 0.1922
```

The Breusch-Pagan test indicates that there is enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  passb4trig3.lm
BP = 21.9736, df = 9, p-value = 0.008963
```

The Durbin-Watson test indicates that there is enough evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  passb4trig3.lm
DW = 3.0092, p-value = 0.0001666
alternative hypothesis: true autocorelation is not 0
```

Noticing that the models with three and four sets of trigonometric functions were much better than that of two sets, but not much different from each other, it was decided to compare all three models to the previous dummy variable model.

Visual examination of the plots (Figure 28) reveals that the dummy variable model has much sharper corners in association with the observations, where the trigonometric function models seem to curve around the observations. The AIC of all four models was turned to for determination of the best fit. The AIC for the dummy variable model is smaller that that of all three trig function models (Output 21), therefore the dummy variable model was deemed best fit.

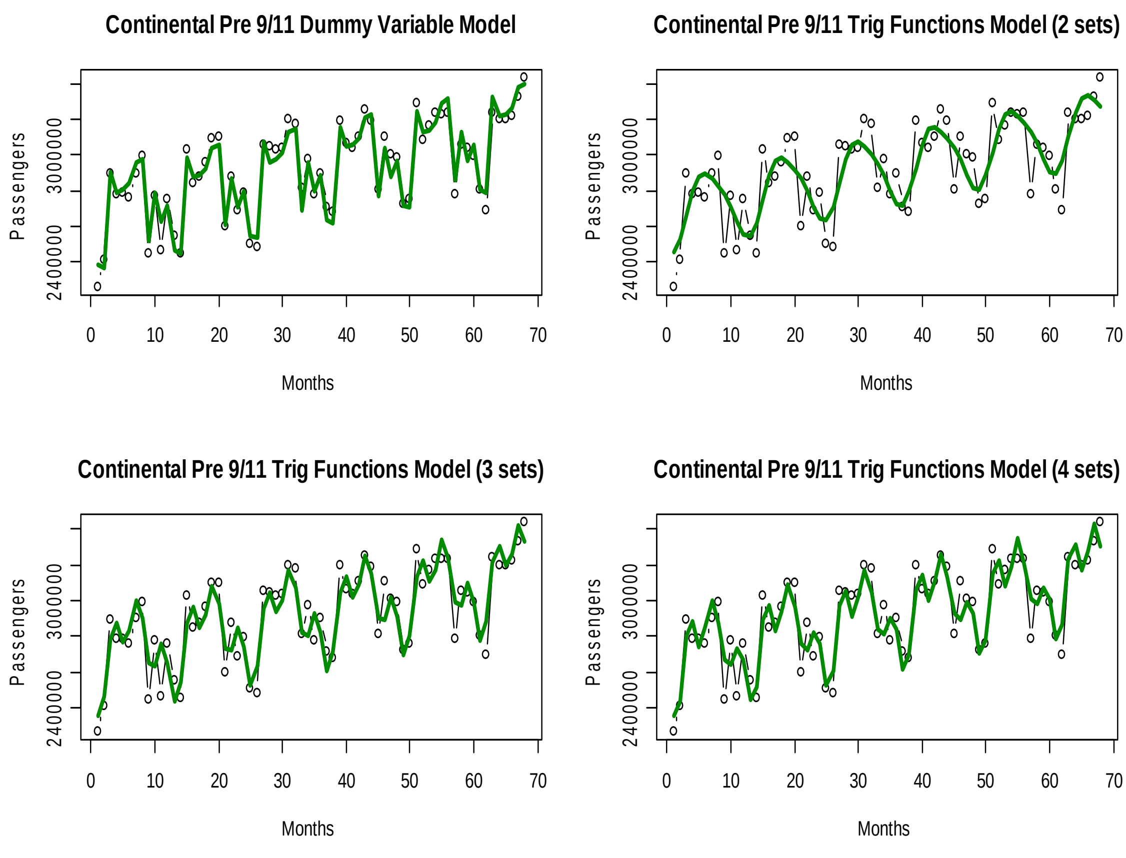

Figure 28: Continental Airlines Pre 9/11 Dummy Variable and Trigonometric Functions Comparison.

```
AIC(passb4dum.lm)
[1] 1708.754
AIC(passb4trig.lm)
[1] 1825.595
AIC(passb4trig2.lm)
[1] 1784.876
AIC(passb4trig3.lm)
[1] 1786.646
```

Output 21: AIC Comparison for Continental Airlines Pre 9/11 Dummy Variable and Trigonometric Functions Models.

Attention was then turned to the post 9/11 data. Similar analyses and comparisons were made. The first model was that of the Continental Airlines post 9/11 data fit with two sets of trigonometric functions (Figure 29).

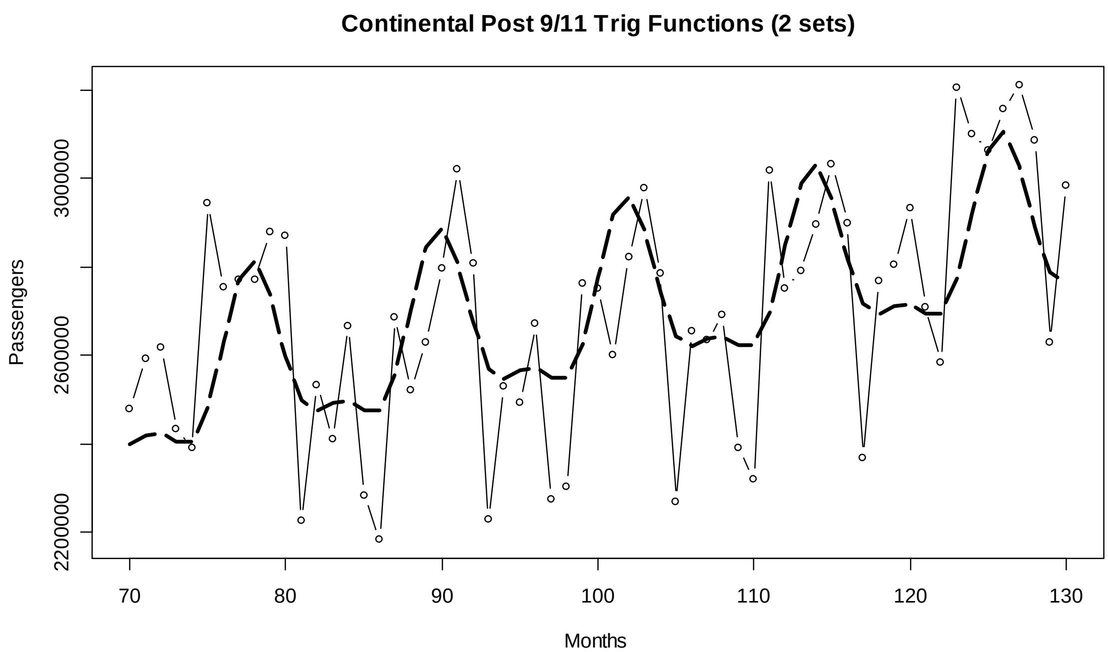

Figure 29: Continental Airlines Post 9/11 Modeled with 2 Sets of Trigonometric Functions.

The summary data (Output 22) indicates that with an *F* statistic 9.626 with associated p-value 1.218e-06 there is a strong relationship between number of passengers and month. The regression equation for Continental Airlines post 9/11 data modeled with two sets of trigonometric functions is of the form:

$$\hat{y}_t = 2089732 + 6077t + 10485\sin\left(\frac{2\pi t}{12}\right) - 176027\cos\left(\frac{2\pi t}{12}\right) - 22570\sin\left(\frac{4\pi t}{12}\right) + 74219\cos\left(\frac{4\pi t}{12}\right) \quad (59)$$

```
Call:
lm(formula = passengersafter ~ monthsafter + I(sin(2 * pi *
monthsafter/12)) +
    I(cos(2 * pi * monthsafter/12)) + I(sin(4 * pi * monthsafter/12)) +
    I(cos(4 * pi * monthsafter/12)))

Residuals:
    Min      1Q  Median      3Q     Max
-375551 -136883   28056  135179  462524

Coefficients:
                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                      2089732     151655  13.779  < 2e-16 ***
monthsafter                         6077       1494   4.067 0.000153 ***
I(sin(2 * pi * monthsafter/12))    10485      36683   0.286 0.776086
I(cos(2 * pi * monthsafter/12))  -176027      37178  -4.735 1.58e-05 ***
I(sin(4 * pi * monthsafter/12))   -22570      36599  -0.617 0.539986
I(cos(4 * pi * monthsafter/12))    74219      36931   2.010 0.049386 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 202700 on 55 degrees of freedom
Multiple R-Squared: 0.4667,     Adjusted R-squared: 0.4182
F-statistic: 9.626 on 5 and 55 DF,  p-value: 1.218e-06
```

Output 22: Summary of Continental Airlines Post 9/11 Modeled with 2 Sets of Trigonometric Functions.

The Shapiro-Wilk test indicates there is not enough evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(passaftertrig.lm)
W = 0.9773, p-value = 0.3153
```

The Breusch-Pagan test indicates that there is enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  passaftertrig.lm
BP = 13.7913, df = 5, p-value = 0.01699
```

The Durbin-Watson test indicates that there is enough evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  passaftertrig.lm
DW = 1.992, p-value = 0.5768
alternative hypothesis: true autocorelation is not 0
```

The second attempt at modeling the Continental Airlines post 9/11 data with trigonometric functions was to use three sets. Visual examination of the plots (Figures 29,30) show that three sets is a much better fit.

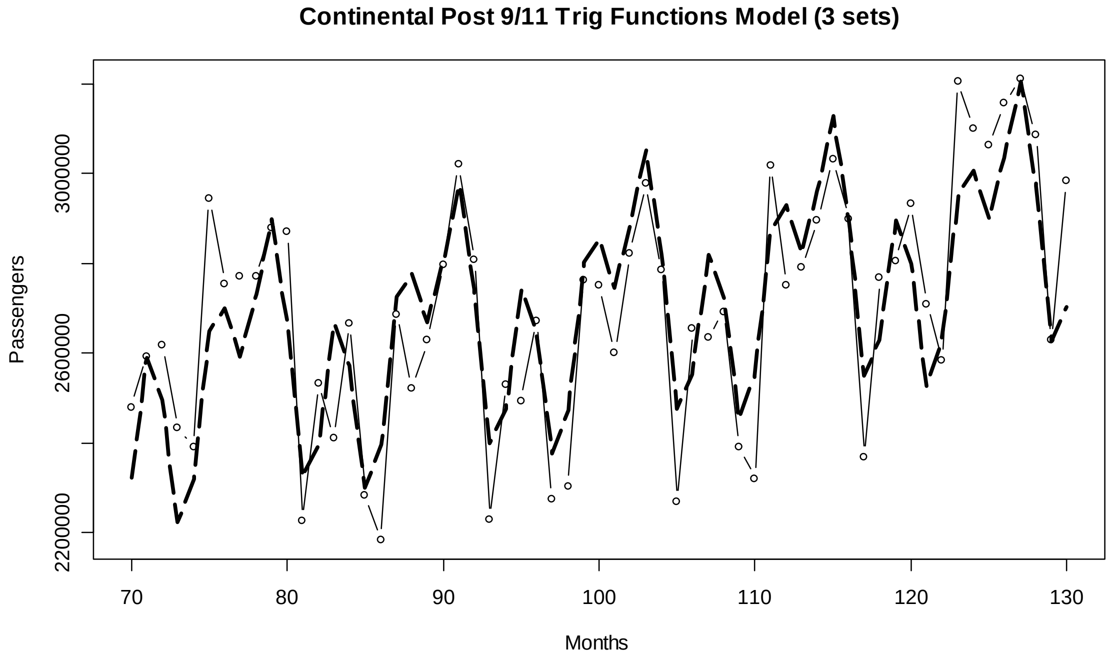

Figure 30: Continental Airlines Post 9/11 Modeled with 3 Sets Trigonometric Functions.

The summary data (Output 23) indicates that with an *F* statistic 19.48 and associated p-value 1.287e-12 there is strong evidence of a relationship between number of passengers and month. The regression equation for the Continental Airlines post 9/11 data modeled with three sets of trigonometric functions is of the form:

$$\hat{y}_t = 2062636 + 6360t + 9011\sin\left(\frac{2\pi t}{12}\right) - 173959\cos\left(\frac{2\pi t}{12}\right)$$
$$-24816\sin\left(\frac{4\pi t}{12}\right) + 73487\cos\left(\frac{4\pi t}{12}\right) - 173293\sin\left(\frac{6\pi t}{12}\right) + 75076\cos\left(\frac{6\pi t}{12}\right) \qquad (60)$$

```
Call:
lm(formula = passengersafter ~ monthsafter + I(sin(2 * pi *
monthsafter/12)) +
    I(cos(2 * pi * monthsafter/12)) + I(sin(4 * pi * monthsafter/12)) +
    I(cos(4 * pi * monthsafter/12)) + I(sin(6 * pi * monthsafter/12)) +
    I(cos(6 * pi * monthsafter/12)))

Residuals:
    Min      1Q  Median      3Q     Max
-255208  -95548  -10075  113816  295880

Coefficients:
                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                      2062636     112007  18.415  < 2e-16 ***
monthsafter                         6360       1104   5.763 4.31e-07 ***
I(sin(2 * pi * monthsafter/12))     9011      27080   0.333  0.74062
I(cos(2 * pi * monthsafter/12))  -173959      27440  -6.340 5.19e-08 ***
I(sin(4 * pi * monthsafter/12))   -24816      27019  -0.918  0.36254
I(cos(4 * pi * monthsafter/12))    73487      27258   2.696  0.00938 **
I(sin(6 * pi * monthsafter/12))  -173293      27332  -6.340 5.18e-08 ***
I(cos(6 * pi * monthsafter/12))    75076      26899   2.791  0.00729 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 149600 on 53 degrees of freedom
Multiple R-Squared: 0.7201,     Adjusted R-squared: 0.6831
F-statistic: 19.48 on 7 and 53 DF,  p-value: 1.287e-12
```

Output 23: Summary of Continental Airlines Post 9/11 Modeled with 3 Sets of Trigonometric Functions.

The Shapiro-Wilk test indicates that there is not enough evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(passaftertrig2.lm)
W = 0.9795, p-value = 0.3967
```

The Breusch-Pagan test does not give enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  passaftertrig2.lm
BP = 11.1084, df = 7, p-value = 0.1340
```

The Durbin-Watson test does not give enough evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  passaftertrig2.lm
DW = 1.9008, p-value = 0.3709
alternative hypothesis: true autocorelation is not
```

A third attempt at modeling the Continental Airlines post 9/11 data was made using four sets of trigonometric functions. Visual examination of the plots (Figures 30,31) does not seem to distinguish between the two models, but both are better fits than the model with two sets of trigonometric functions (Figure 29).

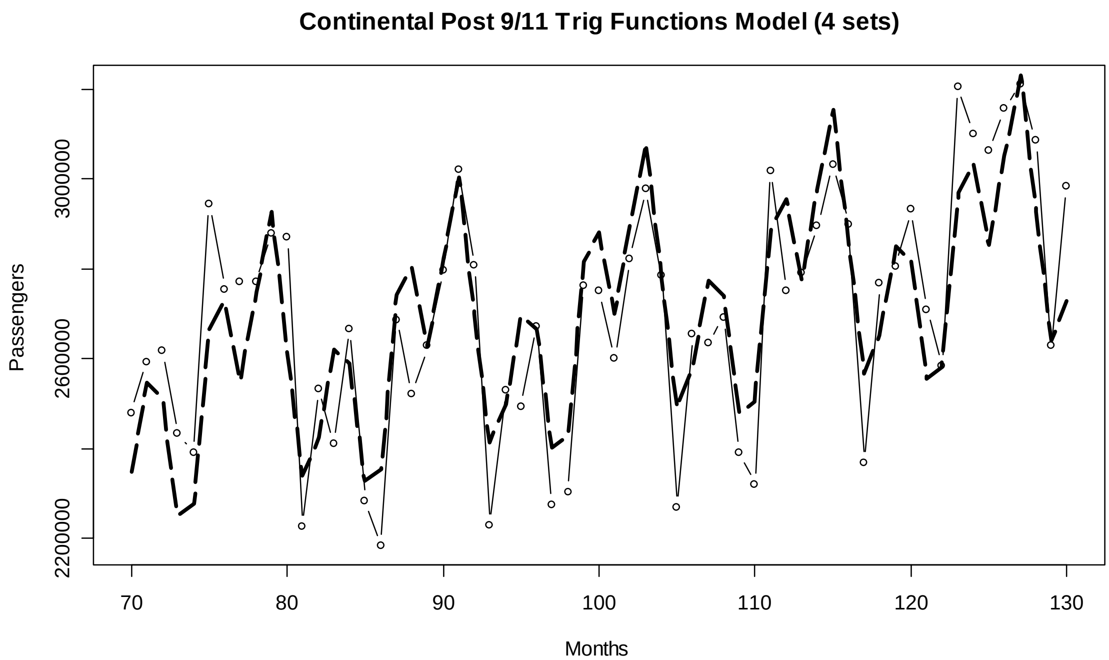

Figure 31: Continental Airlines Post 9/11 Modeled with 4 Sets of Trigonometric Functions.

The summary data (Output 24) indicates that with an *F* statistic of 15.65 and associated p-value 7.311e-12 there is strong evidence of a relationship between number of passengers and month. The regression equation for Continental Airlines post 9/11 data modeled with four sets of trigonometric functions is of the form:

$$\begin{aligned}\hat{y}_t &= 2065539 + 6327t + 9649\sin\left(\frac{2\pi t}{12}\right) - 174470\cos\left(\frac{2\pi t}{12}\right)\\ &- 24088\sin\left(\frac{4\pi t}{12}\right) + 73841\cos\left(\frac{4\pi t}{12}\right) - 173260\sin\left(\frac{6\pi t}{12}\right) + 75884\cos\left(\frac{6\pi t}{12}\right)\\ &+ 40964\sin\left(\frac{8\pi t}{12}\right) + 16849\cos\left(\frac{8\pi t}{12}\right)\end{aligned} \tag{61}$$

```
Call:
lm(formula = passengersafter ~ monthsafter + I(sin(2 * pi *
monthsafter/12)) +
    I(cos(2 * pi * monthsafter/12)) + I(sin(4 * pi * monthsafter/12)) +
    I(cos(4 * pi * monthsafter/12)) + I(sin(6 * pi * monthsafter/12)) +
    I(cos(6 * pi * monthsafter/12)) + I(sin(8 * pi * monthsafter/12)) +
    I(cos(8 * pi * monthsafter/12)))

Residuals:
    Min      1Q  Median      3Q     Max
-282767 -101069    3862  105951  278358

Coefficients:
                                Estimate Std. Error t value Pr(>|t|)
(Intercept)                      2065539     111311  18.556  < 2e-16 ***
monthsafter                         6327       1097   5.769 4.71e-07 ***
I(sin(2 * pi * monthsafter/12))     9649      26914   0.359  0.72144
I(cos(2 * pi * monthsafter/12))  -174470      27265  -6.399 4.84e-08 ***
I(sin(4 * pi * monthsafter/12))   -24088      26853  -0.897  0.37392
I(cos(4 * pi * monthsafter/12))    73841      27084   2.726  0.00876 **
I(sin(6 * pi * monthsafter/12))  -173260      27155  -6.381 5.18e-08 ***
I(cos(6 * pi * monthsafter/12))    75884      26736   2.838  0.00650 **
I(sin(8 * pi * monthsafter/12))    40964      26838   1.526  0.13310
I(cos(8 * pi * monthsafter/12))    16849      27040   0.623  0.53598
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 148600 on 51 degrees of freedom
Multiple R-Squared: 0.7342,     Adjusted R-squared: 0.6872
F-statistic: 15.65 on 9 and 51 DF,  p-value: 7.311e-12
```

Output 24: Summary of Continental Airlines Post 9/11 Modeled with 4 Sets of Trigonometric Functions.

The Shapiro-Wilk test does not indicate evidence enough to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(passaftertrig3.lm)
W = 0.9814, p-value = 0.481
```

The Breusch-Pagan test does not give enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  passaftertrig3.lm
BP = 9.1272, df = 9, p-value = 0.4256
```

The Durbin-Watson test indicates that there is strong evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  passaftertrig3.lm
DW = 1.8782, p-value = 0.4042
alternative hypothesis: true autocorelation is not 0
```

Investigation of the three trigonometric function models was similar to that of the pre 9/11 findings. Thus, it was decided to compare these models with that of the dummy variable model. Upon visual examination, the dummy variable model better fit the data than any of the trig function models (Figure 32). The AIC for the dummy variable model was smallest (output 25) and confirmed the visual findings, the dummy variable model is the best fit to the Continental Airlines post 9/11 data.

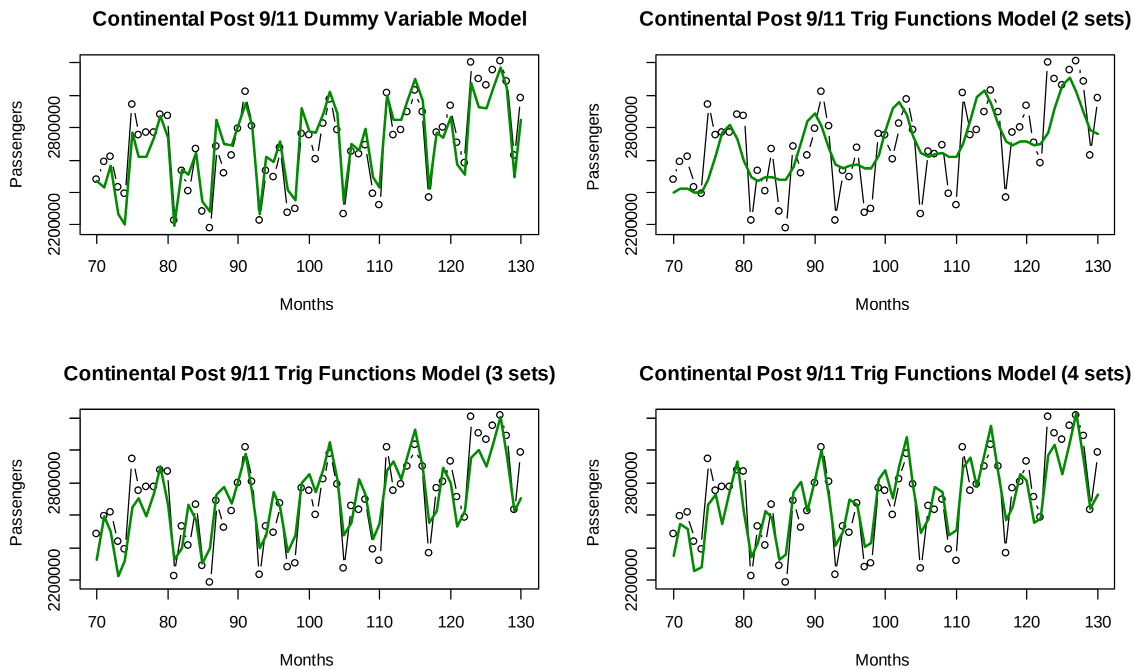

Figure 32: Continental Airlines Post 9/11 Dummy Variable and Trigonometric Models Comparison.

```
AIC(passafterdum.lm)
[1] 1606.067
AIC(passaftertrig.lm)
[1] 1671.565
AIC(passaftertrig2.lm)
[1] 1636.236
AIC(passaftertrig3.lm)
[1] 1637.097
```

Output 25: AIC Comparison for Continental Airlines Post 9/11 Dummy Variable and Trigonometric Functions Models.

Similar comparisons were made for both American Airlines and Delta airlines with the same outcome. Therefore, the trigonometric models were abandoned and no further analyses were made using them.

4.4 Additive Decomposition

In this section additive decomposition was used to model the airline data. The data were decomposed and the seasonal component was removed. The remaining deseasonalized data were then fit with linear, quadratic, and cubic models to see which model fit best and comparisons were made.

4.4.1 Continental Airlines

The Continental Airlines pre 9/11 data were made into a times series and plotted (Figure 33).

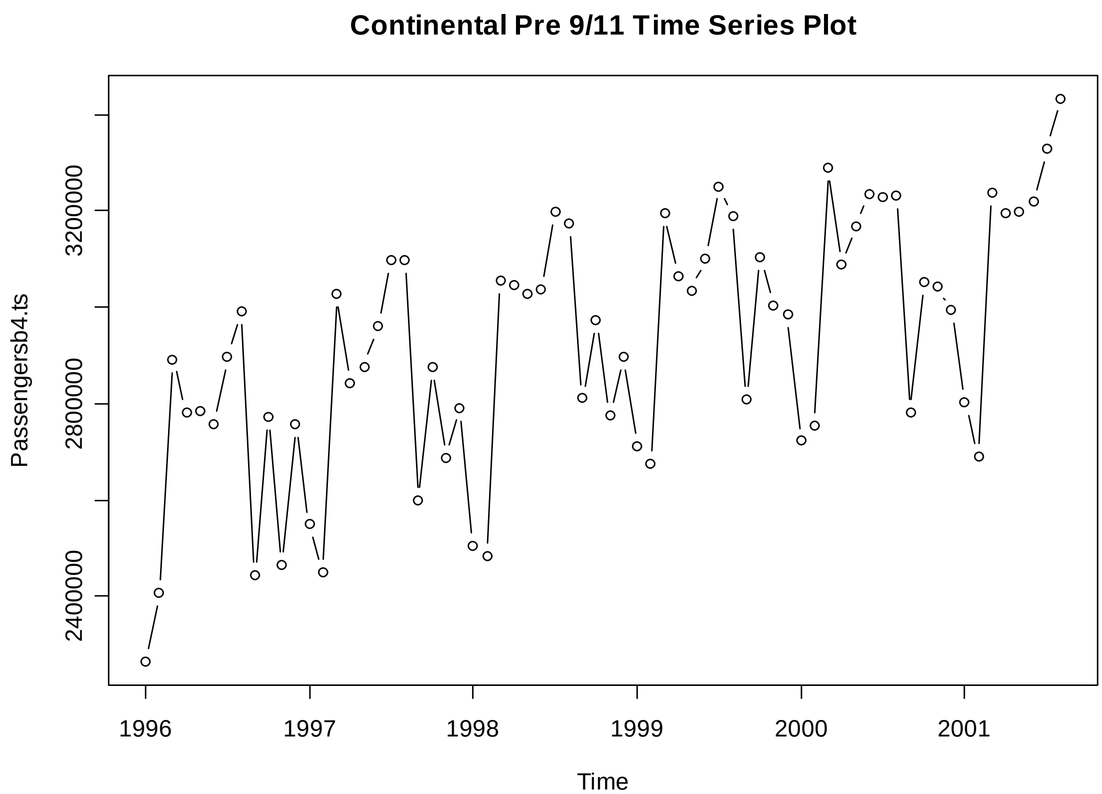

Figure 33: Continental Airlines Pre 9/11 Time Series Plot.

It was then decomposed into its seasonal, trend, and random components (Figure 34). The seasonal component is easily spotted; it looks like a series of capital and lowercase M's.

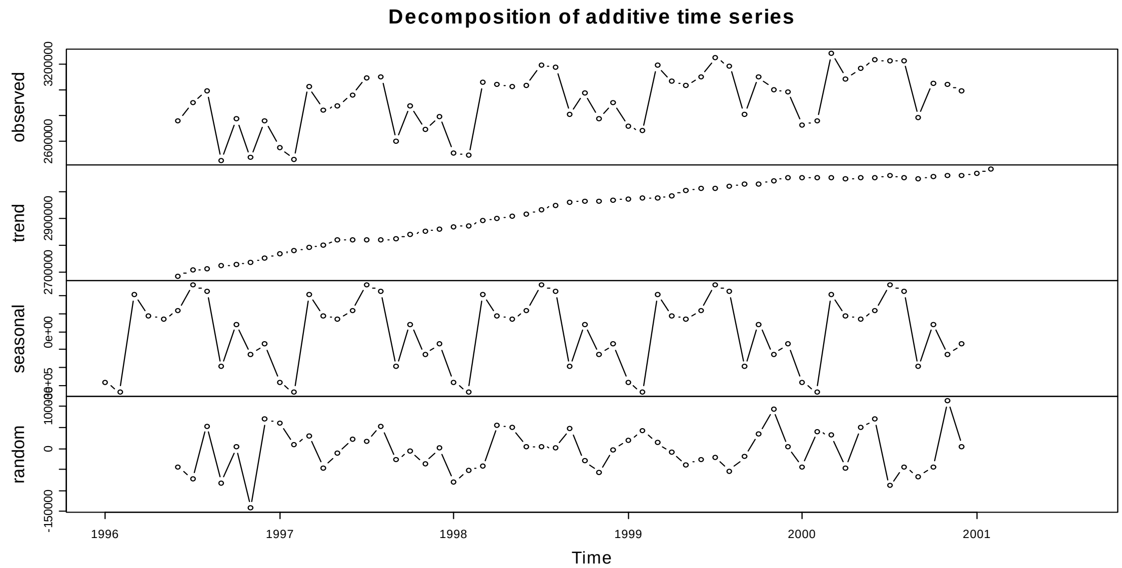

Figure 34: Continental Airlines Pre 9/11 Time Series Decomposition Plot.

The seasonal component was removed and the data were plotted (Figure 35).

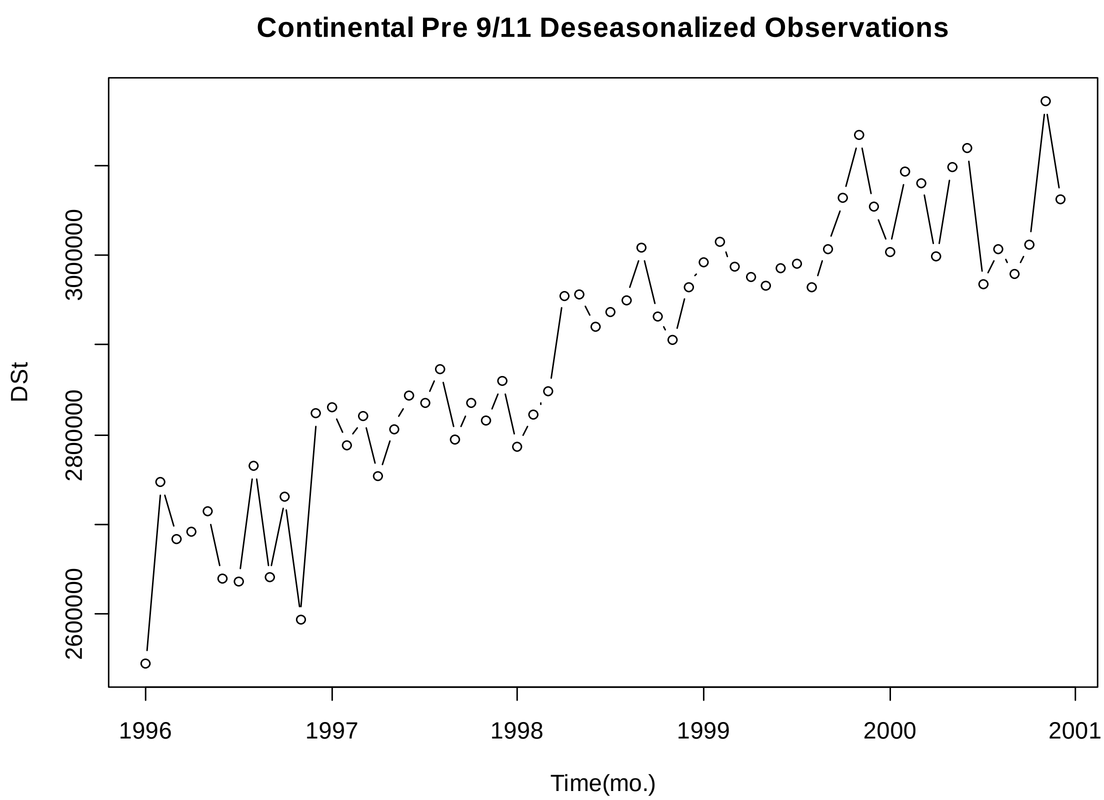

Figure 35: Continental Airlines Pre 9/11 Deseasonalized Plot.

The deseasonalized time series was fit firstly to a linear model. With an *F* statistic of 299.3 and corresponding p-value 2.2e-16 (Output 26) the variables have a strong relationship. The regression equation for the Continental Airlines pre 9/11 time series linear model is of the form:

$$\hat{y}_t = 2664366.3 + 7605.9t \qquad (62)$$

```
Call:
lm(formula = DSt ~ times.ds)

Residuals:
    Min      1Q  Median      3Q     Max
-154063  -32523   10700   42472  111324

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 2664366.3    15419.4   172.8   <2e-16 ***
times.ds       7605.9      439.6    17.3   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 58970 on 58 degrees of freedom
Multiple R-Squared: 0.8377,     Adjusted R-squared: 0.8349
F-statistic: 299.3 on 1 and 58 DF,  p-value: < 2.2e-16
```

Output 26: Summary of Continental Airlines Pre 9/11 Linear Time Series Model.

Looking at Figure 36, the linear regression equation seems to be a good fit to the time series data. The observations all seem to fluctuate around the regression line.

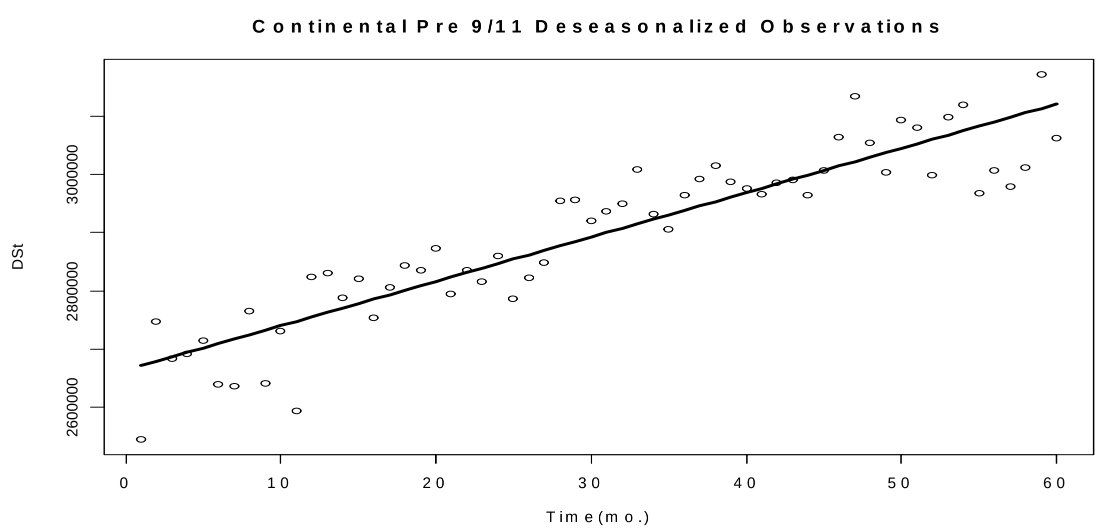

Figure 36: Continental Airlines Pre 9/11 Time Series Model with Linear Regression Line.

The Shapiro-Wilk test does indicate that there is slight evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(DSt.lm)
W = 0.9604, p-value = 0.04949
```

The Breusch-Pagan test does not give enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  DSt.lm
BP = 9e-04, df = 1, p-value = 0.9765
```

The Durbin-Watson test gives sufficient evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  DSt.lm
DW = 1.6191, p-value = 0.05125
alternative hypothesis: true autocorrelation is greater than 0
```

A second attempt at modeling the Continental Airlines pre 9/11 deseasonalized time series was to use a quadratic model. With an F statistic of 177.3 and corresponding p-value 2.2e-16, there is strong evidence of a relationship between number of passengers and month (Output 27). The regression equation for Continental Airlines pre 9/11 deseasonalized time series quadratic model is of the form:

$$\hat{y}_t = 2612114.45 + 12662.52t - 82.90t^2 \tag{63}$$

```
Call:
lm(formula = DSt ~ times.ds + I(times.ds^2))

Residuals:
    Min      1Q  Median      3Q     Max 
-147403  -44382    2967   41266  109717 

Coefficients:
                Estimate Std. Error t value Pr(>|t|)    
(Intercept)   2612114.45   22009.39 118.682  < 2e-16 ***
times.ds        12662.52    1664.83   7.606 3.11e-10 ***
I(times.ds^2)     -82.90      26.45  -3.134  0.00273 ** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 54940 on 57 degrees of freedom
Multiple R-Squared: 0.8615,     Adjusted R-squared: 0.8567 
F-statistic: 177.3 on 2 and 57 DF,  p-value: < 2.2e-16
```

Output 27: Summary of Continental Airlines Pre 9/11 Time Series Quadratic Model.

Plotting the quadratic regression equation with the time series data, there seemed a slight curvature that had not been noticed previously (Figure 37). The data seemed to fluctuate around the regression line as they did with the linear model.

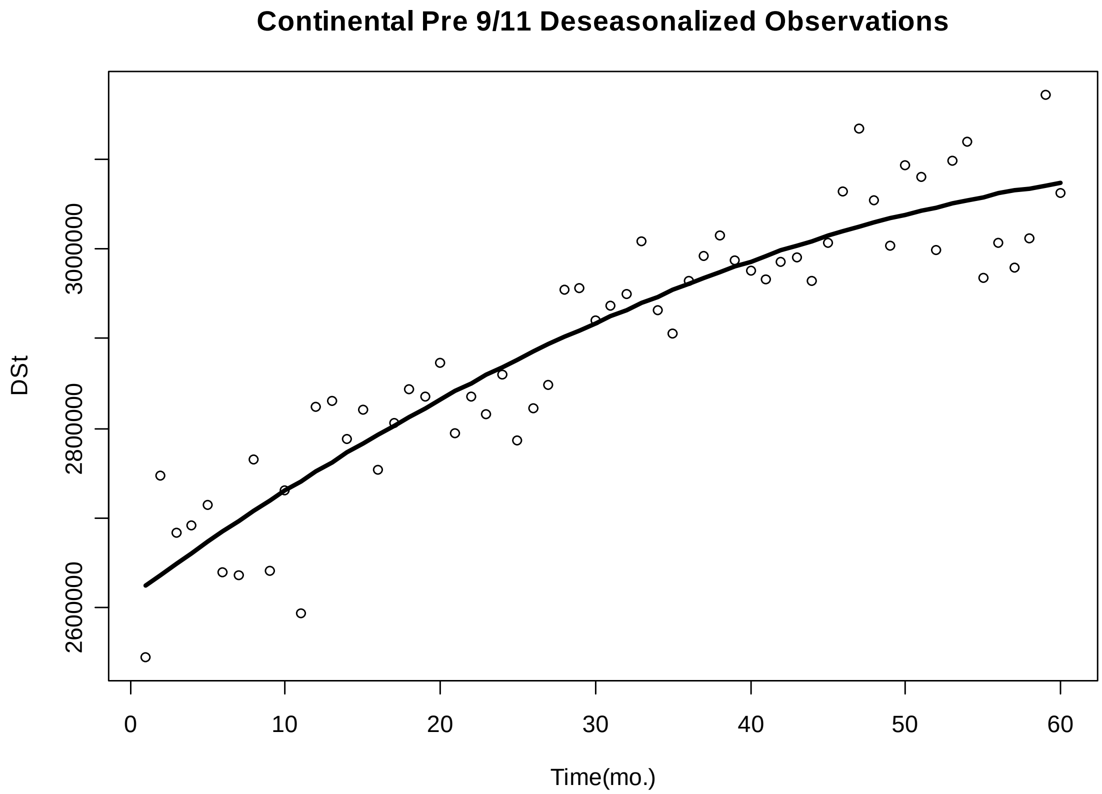

Figure 37: Continental Airlines Pre 9/11 Time Series Model with Quadratic Regression Line.

The Shapiro-Wilk test does not give sufficient evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(DStquad.lm)
W = 0.9854, p-value = 0.6915
```

The Breusch-Pagan test gives sufficient evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  DStquad.lm
BP = 7.6543, df = 2, p-value = 0.02177
```

The Durbin-Watson test gives enough evidence to reject the independence assumption.

```
Durbin-Watson test
data:  DStquad.lm
DW = 1.89, p-value = 0.2406
alternative hypothesis: true autocorrelation is greater than 0
```

A last attempt at modeling the Continental Airlines pre 9/11 deseasonalized data was made using a cubic model. The summary data (Output 28) shows that with an *F* statistic of 118.5 and corresponding p-value 2.2e-16 the number of passengers and the month are strongly related. The regression equation for the Continental Airlines pre 9/11 deseasonalized time series cubic model is of the form:

$$\hat{y}_t = 2633000 + 8799t + 74.16t^2 - 1.716t^3 \quad (64)$$

```
Call:
lm(formula = DSt ~ times.ds + I(times.ds^2) + I(times.ds^3))

Residuals:
    Min      1Q  Median      3Q     Max 
-142068  -39567    3717   37348  114402 

Coefficients:
                Estimate Std. Error t value Pr(>|t|)    
(Intercept)    2.633e+06  3.025e+04  87.025   <2e-16 ***
times.ds       8.799e+03  4.259e+03   2.066   0.0435 *  
I(times.ds^2)  7.416e+01  1.615e+02   0.459   0.6480    
I(times.ds^3) -1.716e+00  1.742e+00  -0.986   0.3286    
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Residual standard error: 54960 on 56 degrees of freedom
Multiple R-Squared: 0.8639,     Adjusted R-squared: 0.8566 
F-statistic: 118.5 on 3 and 56 DF,  p-value: < 2.2e-16
```

Output 28: Summary of Continental Airlines Pre 9/11 Time Series Cubic Model.

Visual examination of the deaseasonalized time series plot (Figure 38) with the cubic regression equation shows a similarity in the curvature to the quadratic model (Figure 37).

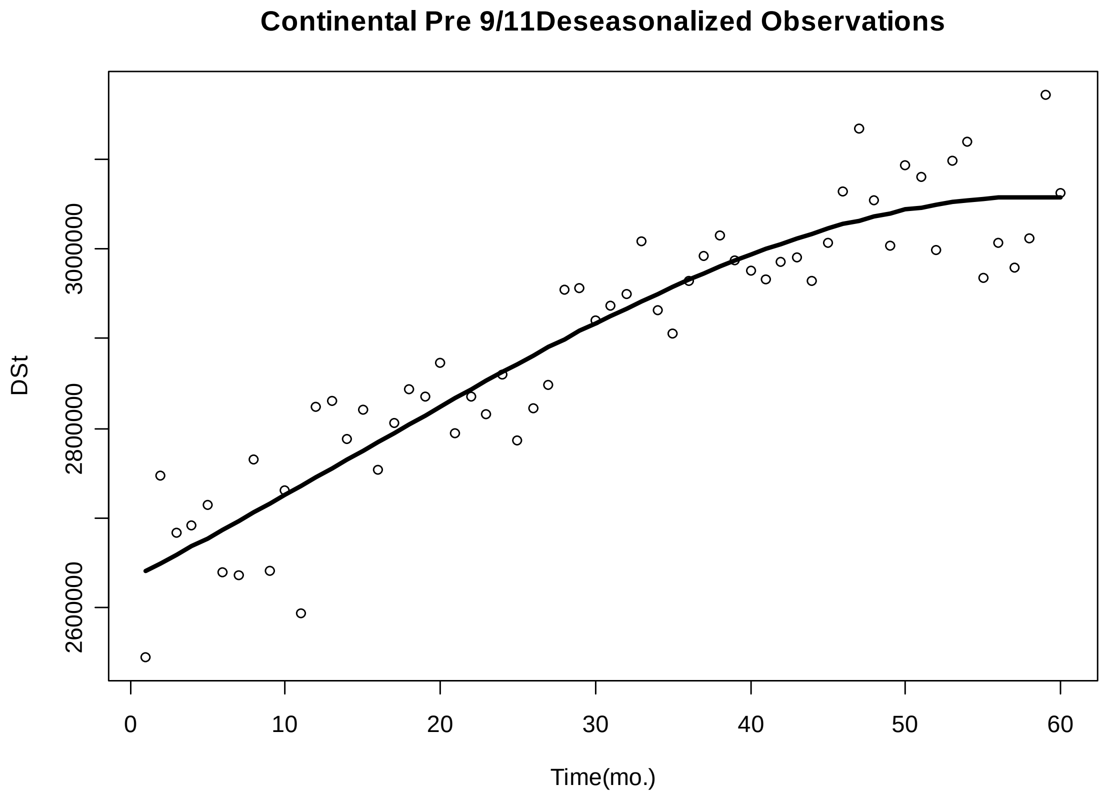

Figure 38: Continental Airlines Pre 9/11 Time Series Model with Cubic Regression Line.

The Shapiro-Wilk test does not give enough evidence to reject the normality assumption.

```
Shapiro-Wilk normality test
data:  resid(DStcubic.lm)
W = 0.9912, p-value = 0.9444
```

The Breusch-Pagan test gives sufficient evidence to reject the constant variance assumption.

```
studentized Breusch-Pagan test
data:  DStcubic.lm
BP = 8.6894, df = 3, p-value = 0.03372
```

The Durbin-Watson test indicates there is sufficient evidence to reject the independence assumption.

```
Durbin-Watson test
data:  DStcubic.lm
DW = 1.9265, p-value = 0.2402
alternative hypothesis: true autocorrelation is greater than 0
```

It was decided to compare the AIC of the three models to see which was the better fit. Output 29 shows that the quadratic model has the lowest AIC and is therefore the better fit. ANOVA was then turned to for confirmation. Output 30 shows that the quadratic model's squared parameter is significant when compared to the linear model, but the cubic model's cubed parameter is of little significance when compared to the quadratic model. Therefore, the quadratic model was deemed the best fit and any further analysis would be done with it.

```
AIC(DSt.lm)
[1] 1492.421
AIC( DStquad.lm)
[1] 1484.884
AIC(DStcubic.lm)
[1] 1485.852
```

Output 29: AIC Comparison for Continental Airlines Pre 9/11 Time Series Linear, Quadratic, and Cubic Models.

```
Analysis of Variance Table

Model 1: DSt ~ times.ds
Model 2: DSt ~ times.ds + I(times.ds^2)
  Res.Df        RSS Df  Sum of Sq      F   Pr(>F)
1     58 2.0172e+11
2     57 1.7208e+11  1 2.9644e+10 9.8197 0.002727 **
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Analysis of Variance Table

Model 1: DSt ~ times.ds + I(times.ds^2)
Model 2: DSt ~ times.ds + I(times.ds^2) + I(times.ds^3)
  Res.Df        RSS Df  Sum of Sq      F Pr(>F)
1     57 1.7208e+11
2     56 1.6914e+11  1 2.9340e+09 0.9714 0.3286
```

Output 30: ANOVA Comparison of Continental Airlines Pre 9/11 Time Series Linear, Quadratic, and Cubic Models.

Attention was then turned to the post 9/11 data to find the best model to use for comparison. The data were made into a time series and plotted (Figure 39).

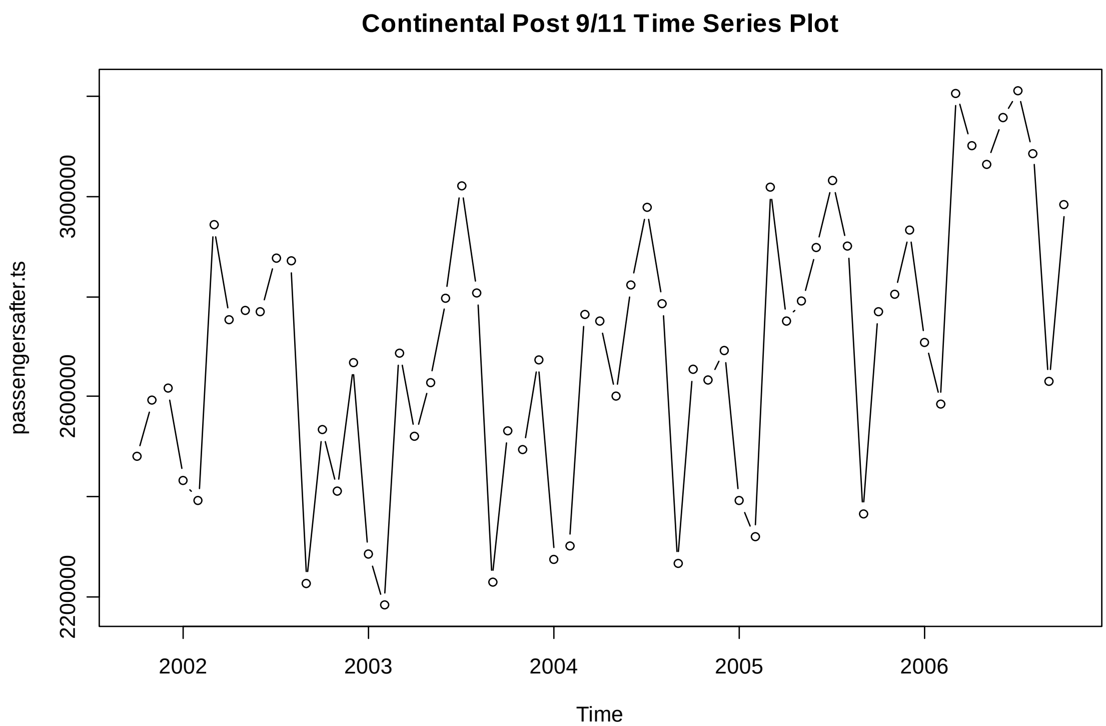

Figure 39: Continental Airlines Post 9/11 Time Series Plot.

The time series was decomposed into its seasonal, trend, and random components (Figure 40). The seasonal component is easily again, looks like a series of capital and lower case M's.

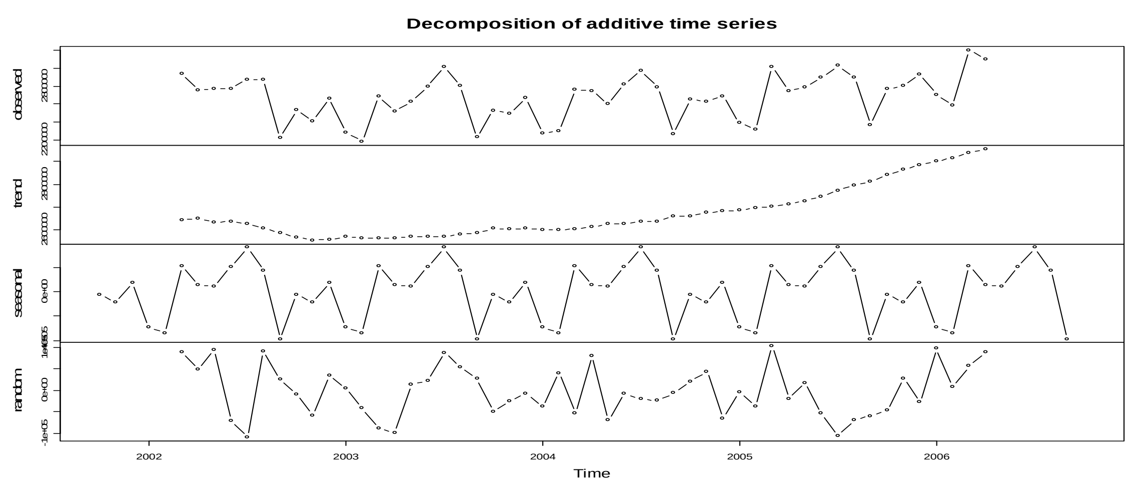

Figure 40: Continental Airlines Post 9/11 Time Series Decomposition Plot.

The seasonal component was removed and the data were plotted (Figure 41).

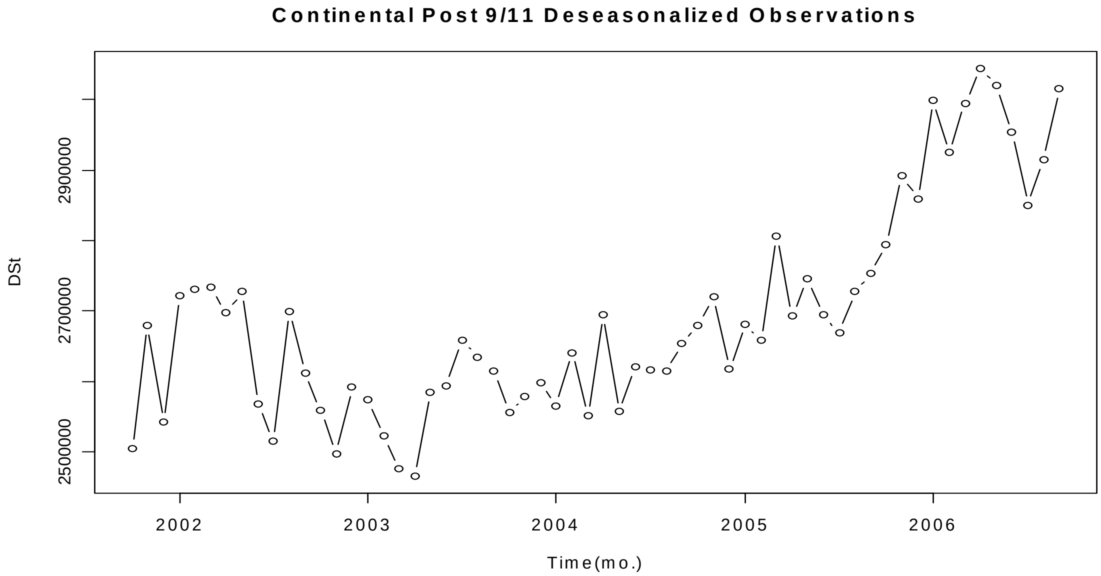

Figure 41: Continental Airlines Post 9/11 Time Series Deseasonalized Plot.

The deseasonalized time series was then fitted with a linear model. The *F* statistic 60.62 and corresponding p-value 1.402e-10 indicates a strong relationship between the number of passengers and time. The regression equation for Continental Airlines post 9/11 deseasonalized linear time series model is of the form:

$$\hat{y}_t = 2087266 + 6065t \tag{65}$$

```
Call:
lm(formula = DSt ~ times.ds)

Residuals:
    Min      1Q  Median      3Q     Max
-156132  -84503  -26686   58697  205134

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  2087266      78679  26.529  < 2e-16 ***
times.ds        6065        779   7.786 1.40e-10 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 104500 on 58 degrees of freedom
Multiple R-Squared: 0.511,      Adjusted R-squared: 0.5026
F-statistic: 60.62 on 1 and 58 DF,  p-value: 1.402e-10
```

Output 31: Summary of Continental Airlines Post 9/11 Time Series Linear Model.

The deseasonalized data were then plotted with the linear regression equation (Figure 42). Upon examination it was found that the line was not a good fit. The observations at either end lied above the line, while the middle observation lied below.

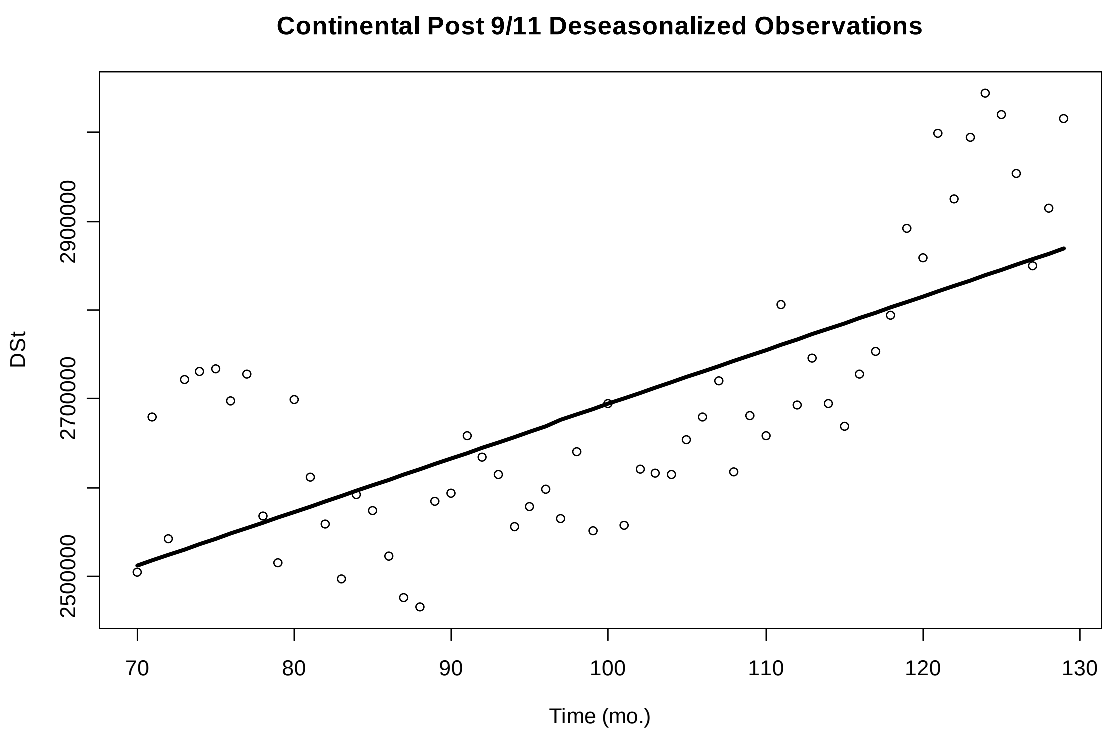

Figure 42: Continental Airlines Post 9/11 Time Series Plot with Linear Regression Line.

The Shapiro-Wilk test indicates that there is sufficient evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(DSt.lm)
W = 0.9168, p-value = 0.0005721
```

The Breusch-Pagan test indicates that there is not sufficient evidence to reject the constant variance assumption.

```
   studentized Breusch-Pagan test
data:  DSt.lm
BP = 0.2313, df = 1, p-value = 0.6305
```

The Durbin-Watson test indicates that there is sufficient evidence to reject the independence assumption.

```
 Durbin-Watson test
data:  DSt.lm
DW = 0.6163, p-value = 6.082e-11
alternative hypothesis: true autocorrelation is greater than 0
```

The data were then subjected to a quadratic model. The *F* statistic of 100.9 with corresponding p-value of < 2.2e-16 (Output 32) indicates that the number of passengers is strongly related to month. The regression equation for the Continental Airlines post 9/11 deseasonalized quadratic model is of the form:

$$\hat{y}_t = 481464.18 - 50469.06t + 284.09t^2 \tag{66}$$

```
Call:
lm(formula = DSt ~ times.ds + I(times.ds^2))

Residuals:
    Min      1Q  Median      3Q     Max 
-169394  -47397   -3246   53933  131366 

Coefficients:
                Estimate Std. Error t value Pr(>|t|)    
(Intercept)   4814646.18  331283.84  14.533  < 2e-16 ***
times.ds       -50469.06    6798.16  -7.424 6.25e-10 ***
I(times.ds^2)     284.09      34.06   8.341 1.87e-11 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Residual standard error: 70740 on 57 degrees of freedom
Multiple R-Squared: 0.7798,     Adjusted R-squared: 0.7721 
F-statistic: 100.9 on 2 and 57 DF,  p-value: < 2.2e-16 
```

Output 32: Summary of Continental Airlines Post 9/11 Time Series Quadratic Model.

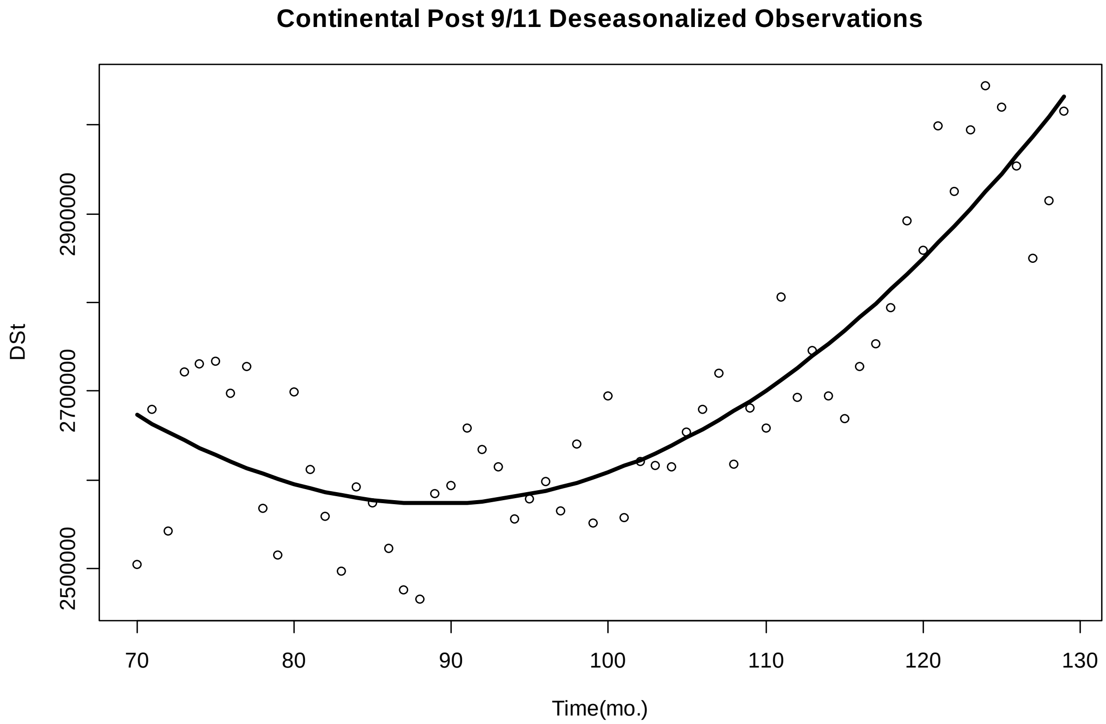

Figure 43: Continental Airlines Post 9/11 Time Series Plot with Quadratic Regression Line.

The quadratic regression equation (Figure 43) seems a much better fit to the data. Observations seem to encase the quadratic line.

The Shapiro-Wilk test does not give enough evidence to reject the normality assumption.

```
   Shapiro-Wilk normality test
data:  resid(DStquad.lm)
W = 0.9838, p-value = 0.6067
```

The Breusch-Pagan test gives sufficient evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  DStquad.lm
BP = 15.9975, df = 2, p-value = 0.0003359
```

The Durbin-Watson test gives sufficient evidence to reject the independence assumption.

```
 Durbin-Watson test
data:  DStquad.lm
DW = 1.3714, p-value = 0.002599
alternative hypothesis: true autocorrelation is greater than 0
```

A final attempt to model the post 9/11 deseasonalized time series was made using a cubic model. The F statistic of 66.11 with corresponding p-value <2.2e-16 indicates a strong relationship between number of passengers and time (Output 33). The regression equation was of the form:

$$\hat{y}_t = 4737000 - 48030t + 259.2t^2 + .08349t^3 \qquad (67)$$

```
Call:
lm(formula = DSt ~ times.ds + I(times.ds^2) + I(times.ds^3))

Residuals:
    Min      1Q  Median      3Q     Max 
-168579  -47147   -3390   54084  131505 

Coefficients:
                Estimate Std. Error t value Pr(>|t|)  
(Intercept)    4.737e+06  2.133e+06   2.221   0.0304 *
times.ds      -4.803e+04  6.630e+04  -0.724   0.4718  
I(times.ds^2)  2.592e+02  6.759e+02   0.383   0.7029  
I(times.ds^3)  8.349e-02  2.262e+00   0.037   0.9707  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1 

Residual standard error: 71370 on 56 degrees of freedom
Multiple R-Squared: 0.7798,     Adjusted R-squared: 0.768 
F-statistic: 66.11 on 3 and 56 DF,  p-value: < 2.2e-16
```

Output 33: Summary on Continental Airlines Post 9/11 Time Series Cubic Model.

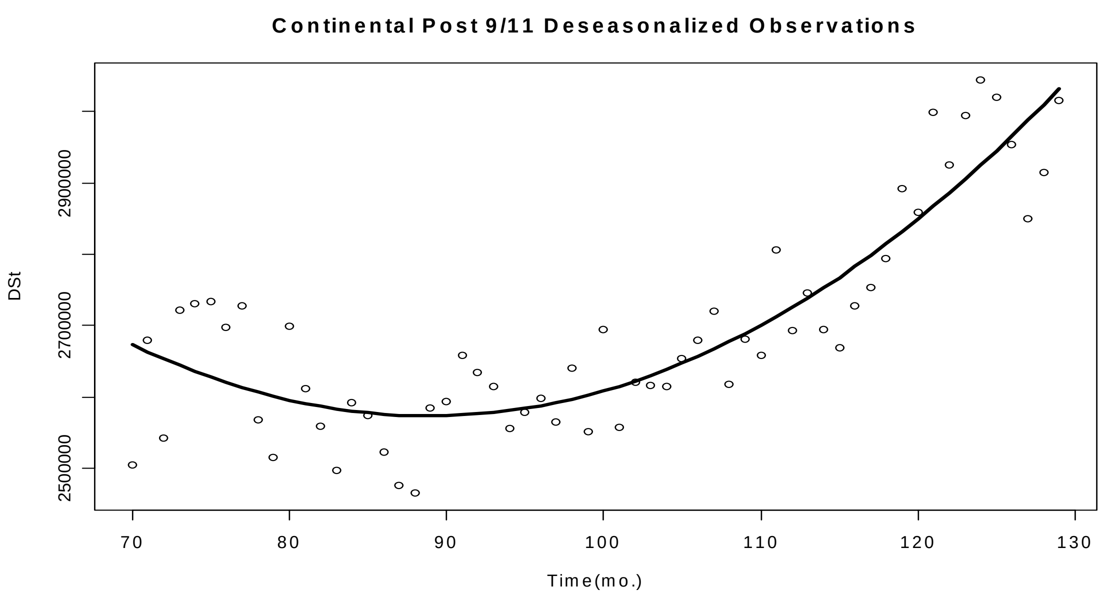

Figure 44: Continental Airlines Post 9/11 Time Series with Cubic Regression Line.

Upon visual examination of the cubic regression line with the deseasonalized data (Figure 44), it is seen that the model much resembles that of the quadratic model (Figure 43).

The Shapiro-Wilk test does not indicate that there is sufficient evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(DStcubic.lm)
W = 0.9836, p-value = 0.5963
```

The Breusch-Pagan test gives sufficient evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  DStcubic.lm
BP = 16.6188, df = 3, p-value = 0.0008465
```

The Durbin-Watson test gives enough evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  DStcubic.lm
DW = 1.3712, p-value = 0.001613
alternative hypothesis: true autocorrelation is greater than 0
```

The three models were compared to see which was the best fit. Output 34 shows the AIC comparison of the three models. The quadratic model has the lowest AIC indicating that this is the best fit.

```
AIC (DSt.lm)
[1] 1561.076
AIC(DStquad.lm)
[1] 1515.208
AIC(DStcubic.lm)
[1] 1517.206
```

Output 34: AIC Comparison of Continental Airlines Post 9/11 Time Series Linear, Quadratic, and Cubic Models.

ANOVA was turned to for confirmation. Output 35 shows the ANOVA comparisons between the linear and quadratic model, also the quadratic and cubic models. The quadratic model's parameter for t^2 is significant, while the cubic model's parameter for t^3 is not. Thus the ANOVA agrees with AIC in that the quadratic model is the best fit.

```
Analysis of Variance Table
Model 1: DSt ~ times.ds
Model 2: DSt ~ times.ds + I(times.ds^2)
  Res.Df        RSS Df  Sum of Sq      F    Pr(>F)
1     58 6.3342e+11
2     57 2.8524e+11  1 3.4818e+11 69.577 1.869e-11 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Analysis of Variance Table
Model 1: DSt ~ times.ds + I(times.ds^2)
Model 2: DSt ~ times.ds + I(times.ds^2) + I(times.ds^3)
  Res.Df        RSS Df  Sum of Sq      F Pr(>F)
1     57 2.8524e+11
2     56 2.8523e+11  1 6.9422e+06 0.0014 0.9707
```

Output 35: ANOVA Comparison of Continental Airlines Post 9/11 Time Series Linear, Quadratic, and Cubic Models.

All post 9/11 models were tested for potential and influential outliers signifying an initial shock recovery period as in previous analysis. Using the residual and leverage plots and Cook's Distance statistics on each of the models, singled out observations were counted as potential outliers.

For the linear model, (Figure 45) observations 4 and 5 are singled out and close to the incident, which would indicate an initial shock recovery period. The Cook's D statistic did not indicate any influential outliers (Output 36).

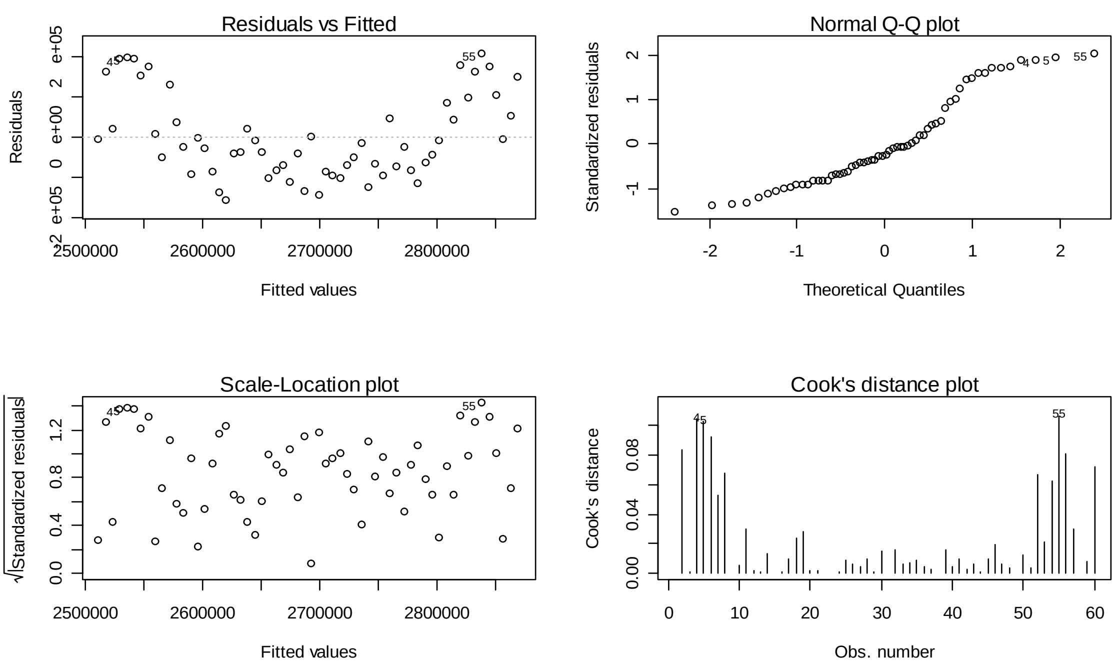

Figure 45: Continental Airlines Post 9/11Time Series Linear Model Residual Plots.

```
cooksD=cooks.distance(DSt.lm)
f0.50=qf(0.5,df1=5,df2=24)
cooksD[which(cooksD>f0.50)]
named numeric(0)
```

Output 36: Cook's D Statistic on Continental Airlines Post 9/11 Time Series Linear Model.

For the quadratic model, the only observation that stood out corresponding to an initial recovery period was observation one (Figure 46). The Cook's D test did not indicate any influential outliers (Output 37).

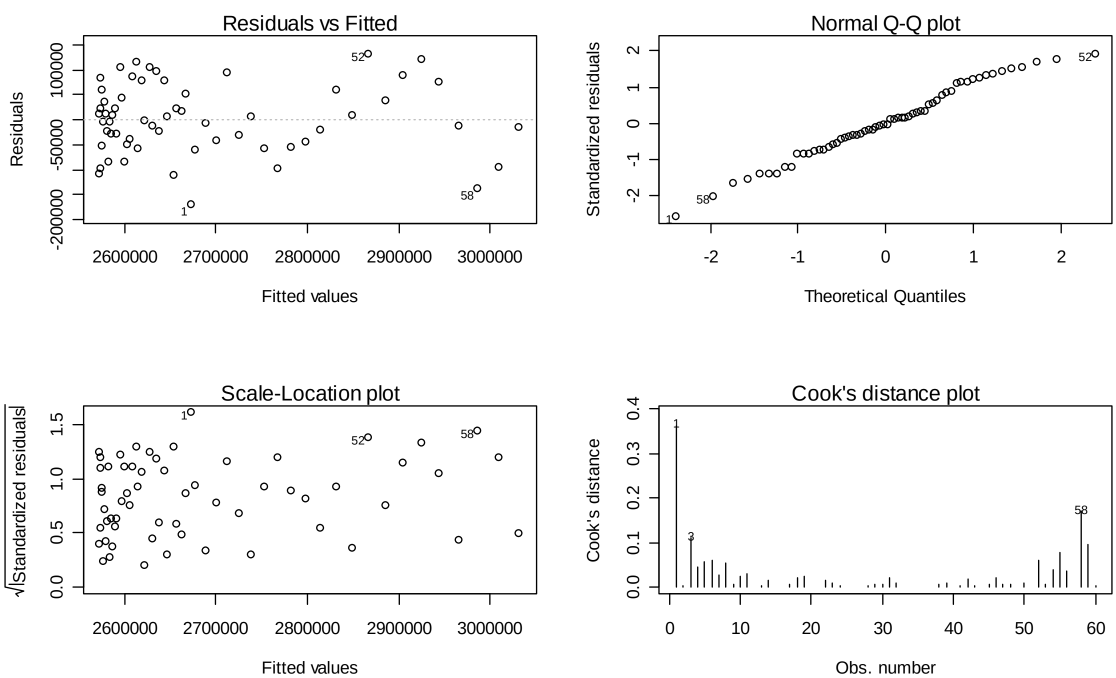

Figure 46: Continental Airlines Post 9/11 Time Series Quadratic Model Residual Plots.

```
cooksD=cooks.distance(DStquad.lm)
f0.50=qf(0.5, df1=5, df2=24)
cooksD[which(cooksD>f0.50)]
named numeric(0)
```

Output 37: Cook's D Statistic on Continental Airlines Post 9/11 Time Series Quadratic Model.

The residual plots for the cubic model all singled out observation one, which is close to 9/11 (Figure 47). The Cook's D test did not indicate any influential observations (Output 38).

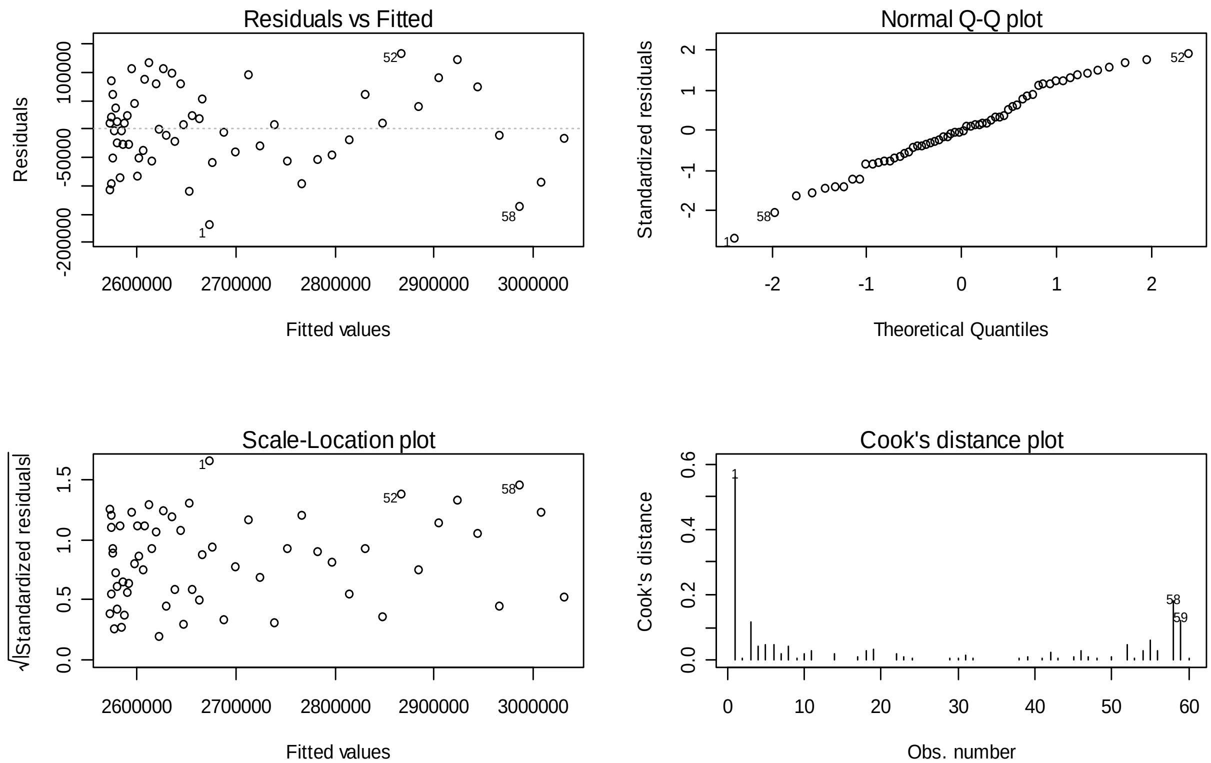

Figure 47: Continental Airlines Post 9/11 Time Series Cubic Model Residual Plots.

```
cooksD=cooks.distance(DStcubic.lm)
f0.50=qf(0.5, df1=5, df2=24)
cooksD[which(cooksD>f0.50)]
named numeric(0)
```

Output 38: Cook’s D Statistic on Continental Airlines Post 9/11 Time Series Cubic Model.

It was decided that the first five months after 9/11 would be removed as an initial recovery period and the analysis on the new post 9/11 dataset would be run again. The dataset was made into a time series and plotted (Figure 48).

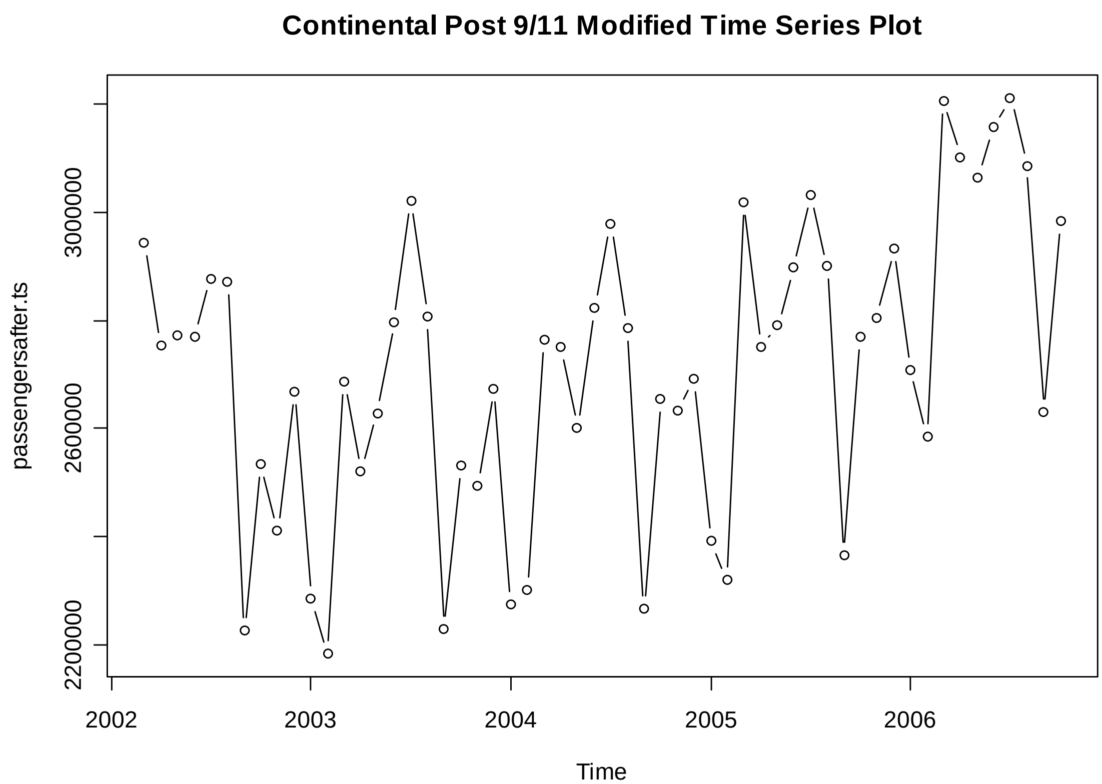

Figure 48: Continental Airlines Post 9/11 Modified Time Series Plot.

The time series was decomposed into its trend, seasonal, and random components (Figure 49).

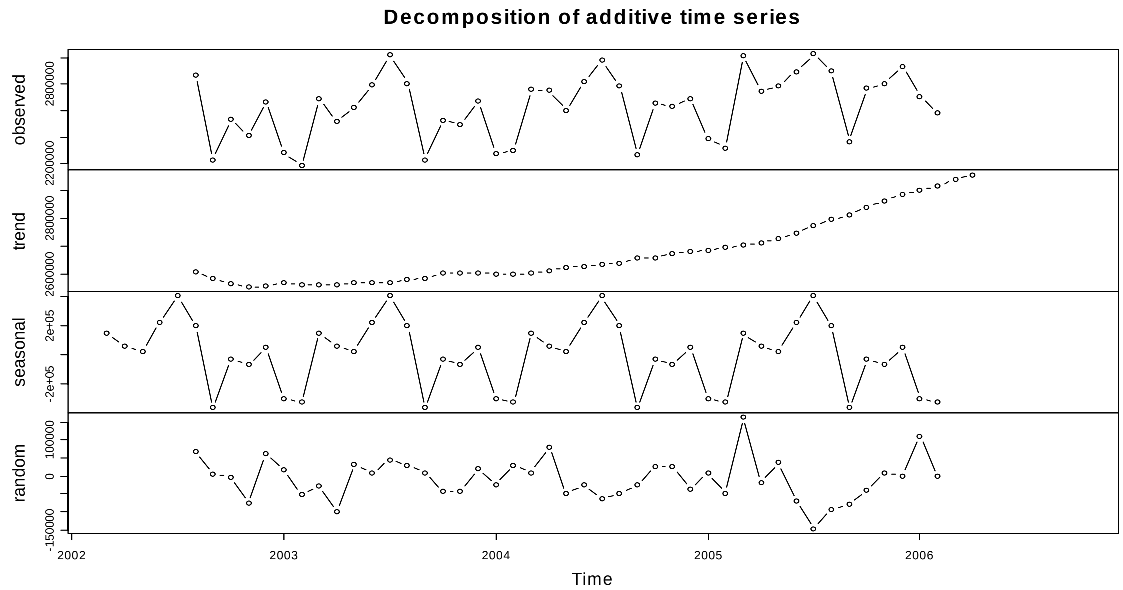

Figure 49: Continental Airlines Post 9/11 Modified Time Series Decomposition Plot.

The seasonal component was removed and the time series plotted (Figure 50).

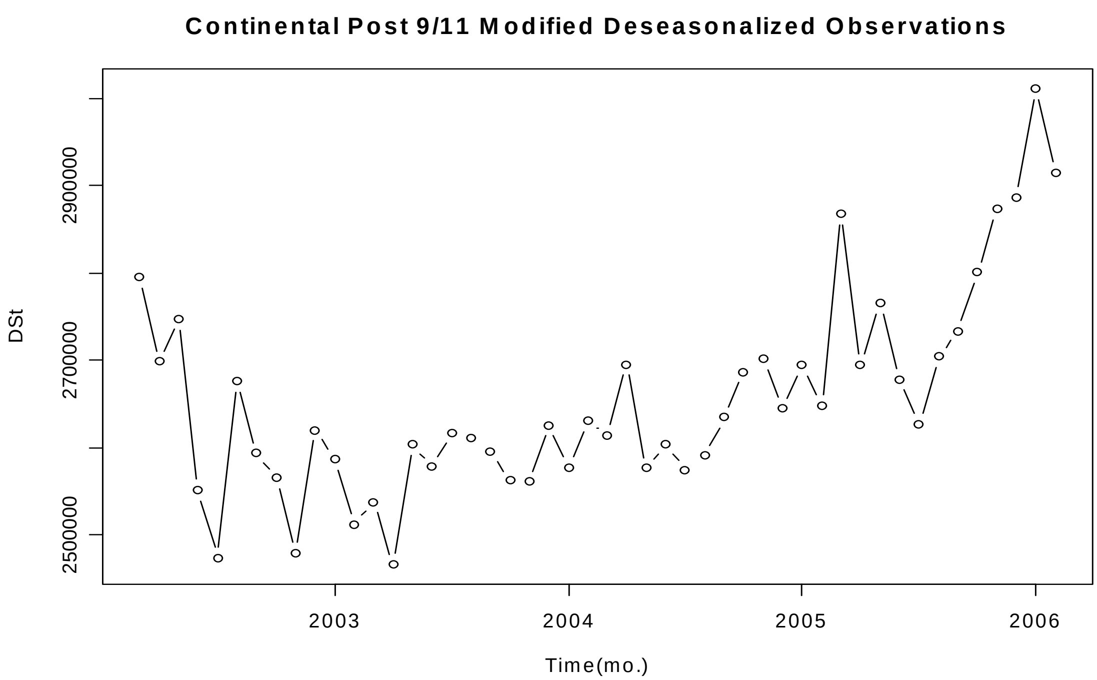

Figure 50: Continental Airlines Post 9/11 Modified Time Series Deseasonalized Plot.

The deseasonalized time series was fit with a linear model. The *F* statistic of 27 and corresponding p-value 4.539e-06 (Output 39) indicates a strong relationship between passenger numbers and time. The Continental Airlines linear regression equation for the deseasonalized time series post 9/11 excluding the first five months is of the form:

$$\hat{y}_t = 2179313.8 + 5095.8t \tag{68}$$

```
Call:
lm(formula = DSt ~ times.ds)

Residuals:
    Min      1Q  Median      3Q     Max 
-136543  -67061  -17120   32739  258822 

Coefficients:
             Estimate Std. Error t value Pr(>|t|)    
(Intercept) 2179313.8    92695.7  23.510  < 2e-16 ***
times.ds       5095.8      980.7   5.196 4.54e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1 

Residual standard error: 94130 on 46 degrees of freedom
Multiple R-Squared: 0.3699,     Adjusted R-squared: 0.3562 
F-statistic:    27 on 1 and 46 DF,  p-value: 4.539e-06
```

Output 39: Summary of Continental Airlines Post 9/11 Modified Time Series Linear Model.

The data plot of the time series and the linear regression equation does not seem a good fit (Figure 51). Observations at either end do not seem to coincide with the linear regression line.

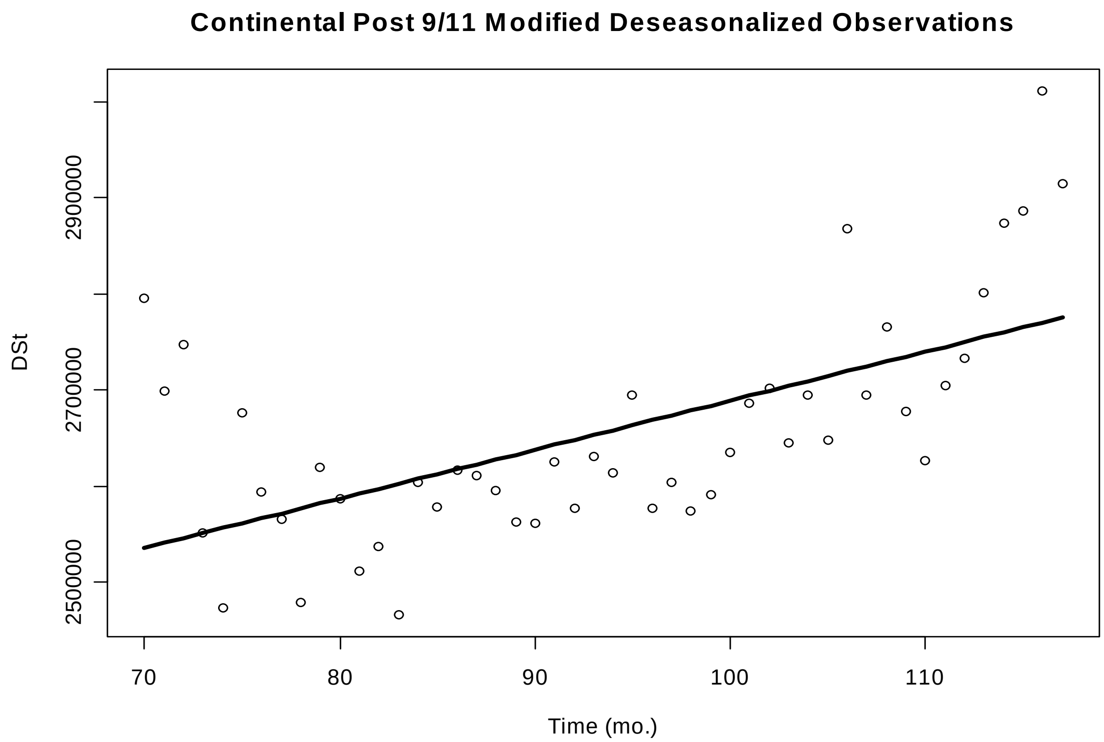

Figure 51: Continental Airlines Post 9/11 Modified Time Series with Linear Regression Line.

The Shapiro-Wilk test indicates that there is enough evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(DSt.lm)
W = 0.8904, p-value = 0.0003119
```

The Breusch-Pagan test does not give enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  DSt.lm
BP = 0.2972, df = 1, p-value = 0.5856
```

The Durbin-Watson test indicates that there is enough evidence to reject the independence assumption.

```
Durbin-Watson test
data:  DSt.lm
DW = 0.8521, p-value = 2.372e-06
alternative hypothesis: true autocorrelation is greater than 0
```

The data were then fit with a quadratic model. With an *F* statistic of 46374 and a corresponding p-value of 1.038e-11 there is strong evidence of a relationship between number of passengers and time (Output 40). The regression equation for the Continental Airlines post 9/11 time series excluding the first 5 months quadratic model is of the form:

$$\hat{y}_t = 5375495 - 64806.3t + 373.8t^2 \tag{69}$$

```
Call:
lm(formula = DSt ~ times.ds + I(times.ds^2))

Residuals:
    Min      1Q  Median      3Q     Max
-153318  -37035    2161   40408  162031

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)   5375495.0   496251.0  10.832 4.01e-14 ***
times.ds       -64806.3    10776.5  -6.014 2.98e-07 ***
I(times.ds^2)     373.8       57.5   6.501 5.62e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 68340 on 45 degrees of freedom
Multiple R-Squared: 0.675,      Adjusted R-squared: 0.6606
F-statistic: 46.74 on 2 and 45 DF,  p-value: 1.038e-11
```

Output 40: Summary on Continental Airlines Post 9/11 Modified Time Series Quadratic Model.

The plot of the time series with the quadratic regression equation (Figure 78) seems to fit quite well.

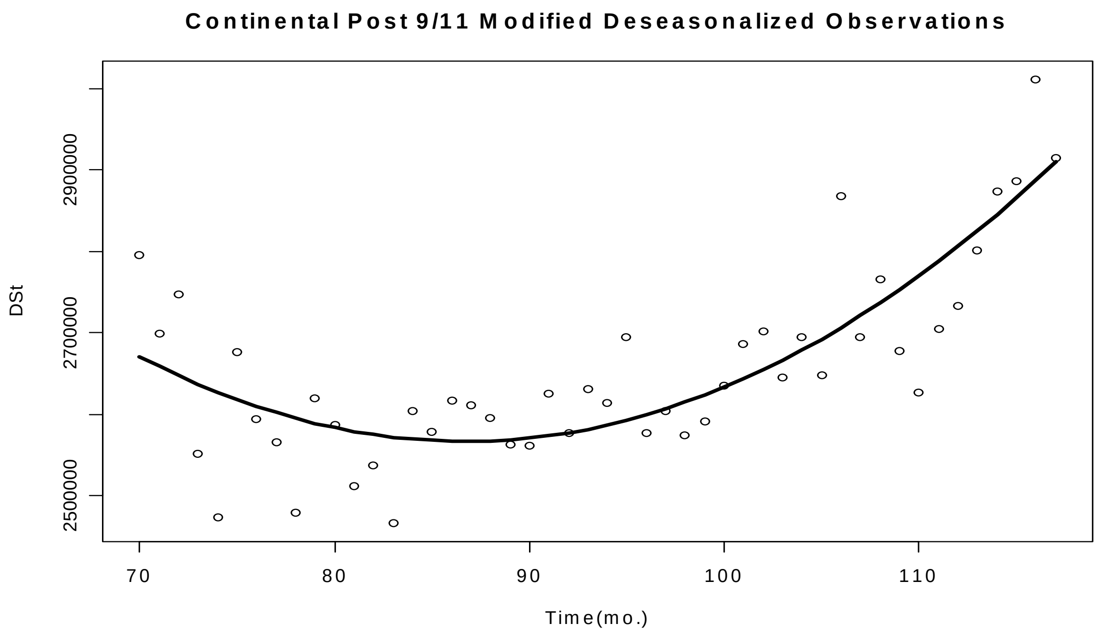

Figure 52: Continental Airlines Post 9/11 Modified Time Series Model with Quadratic Regression Line.

The Shapiro-Wilk test does not give sufficient evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(DStquad.lm)
W = 0.9829, p-value = 0.7008
```

The Breusch-Pagan test does not give sufficient evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  DStquad.lm
BP = 5.5474, df = 2, p-value = 0.06243
```

The Durbin-Watson test gives sufficient evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  DStquad.lm
DW = 1.6074, p-value = 0.04451
alternative hypothesis: true autocorrelation is greater than 0
```

A final attempt was made at modeling the time series data using a cubic model. The *F* statistic of 30.79 with corresponding p-value 6.98e-11 indicates a strong relationship between number of passengers and time (Output 41). The cubic regression equation for the Continental Airlines post 9/11 time series data excluding the first 5 months is of the form:

$$\hat{y}_t = 7466000 - 133700t + 1121t^2 - 2.662t^3 \tag{70}$$

```
Call:
lm(formula = DSt ~ times.ds + I(times.ds^2) + I(times.ds^3))

Residuals:
    Min      1Q  Median      3Q     Max
-155135  -34557    5010   38069  155742

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    7.466e+06  3.779e+06   1.976   0.0545 .
times.ds      -1.337e+05  1.240e+05  -1.079   0.2866
I(times.ds^2)  1.121e+03  1.339e+03   0.837   0.4073
I(times.ds^3) -2.662e+00  4.771e+00  -0.558   0.5796
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 68870 on 44 degrees of freedom
Multiple R-Squared: 0.6773,     Adjusted R-squared: 0.6553
F-statistic: 30.79 on 3 and 44 DF,  p-value: 6.98e-11
```

Output 41: Summary of Continental Airlines Post 9/11 Modified Time Series Cubic Model.

Visual examination of the post 9/11 time series with the cubic regression line showed a similarity between it and the quadratic model (Figures 52 & 53).

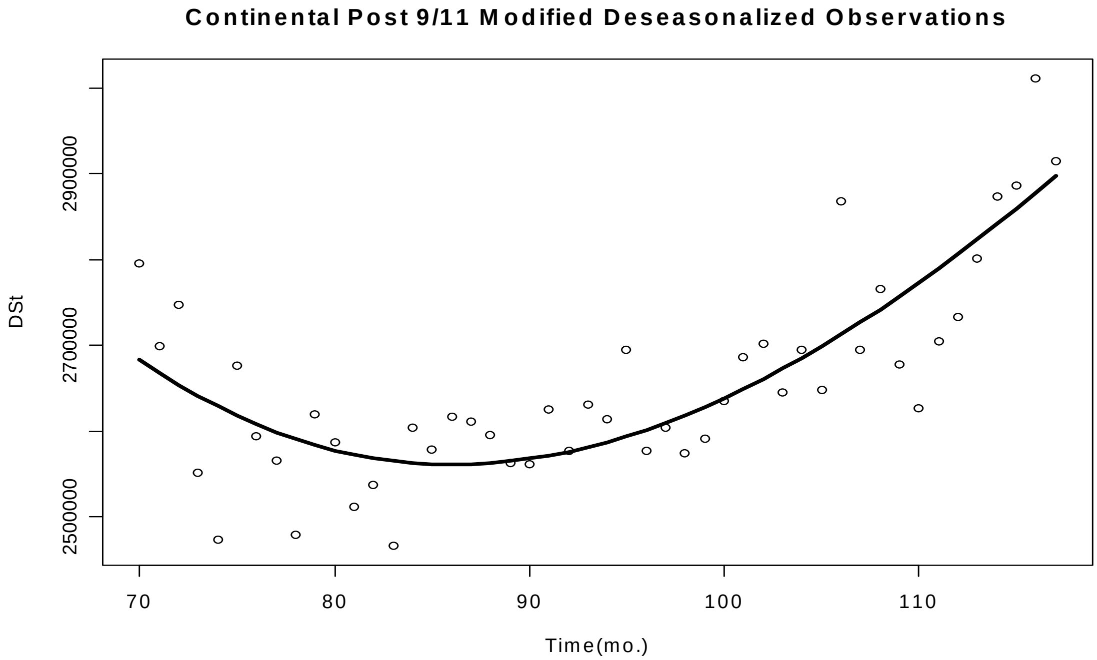

Figure 53: Continental Airlines Post 9/11 Modified Time Series with Cubic Regression Line.

The Shapiro-Wilk test does not give enough evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(DStcubic.lm)
W = 0.9854, p-value = 0.8077
```

The Breusch-Pagan test does not give enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  DStcubic.lm
BP = 6.3558, df = 3, p-value = 0.09552
```

The Durbin-Watson test gives sufficient evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  DStcubic.lm
DW = 1.6149, p-value = 0.03273
alternative hypothesis: true autocorrelation is greater than 0
```

To determine which model of the three was the best fit, the AIC of each was compared. Output 42 shows that the AIC for the quadratic model is the smallest indicating that the quadratic model is the best fit to the post 9/11 Continental Airlines time series excluding the first 5 months. The ANOVA (Output 43) comparison agrees with these findings.

```
AIC (DSt.lm)
[1] 1239.605
AIC(DStquad.lm)
[1] 1209.818
AIC(DStcubic.lm)
[1] 1211.479
```

Output 42: AIC Comparison of Continental Airlines Post 9/11 Modified Time Series Linear, Quadratic, and Cubic Models.

```
Analysis of Variance Table

Model 1: DSt ~ times.ds
Model 2: DSt ~ times.ds + I(times.ds^2)
  Res.Df        RSS Df  Sum of Sq      F    Pr(>F)
1     46 4.0755e+11
2     45 2.1017e+11  1 1.9737e+11 42.259 5.624e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Analysis of Variance Table

Model 1: DSt ~ times.ds + I(times.ds^2)
Model 2: DSt ~ times.ds + I(times.ds^2) + I(times.ds^3)
  Res.Df        RSS Df  Sum of Sq      F Pr(>F)
1     45 2.1017e+11
2     44 2.0870e+11  1 1.4770e+09 0.3114 0.5796
```

Output 43: ANOVA Comparison of Continental Airlines Post 9/11 Modified Time Series Linear, Quadratic, and Cubic Models.

According to AIC (Output 34 & 42), the quadratic model for the Continental Airlines post 9/11 data excluding the first five months is the best fit for the post 9/11 data. Further analysis would be done with this model.

The pre 9/11 quadratic time series original model and the post 9/11 quadratic time series modified model were then plotted together and compared. The quadratic regression equations were set equal and solved for time to see when they crossed. Figure 54 shows the comparison and the point at which they cross. The vertical line represents the end of

October 2005. The dashed line represents the pre 9/11 regression equation extrapolated into the future had 9/11 not happened. The solid line represents the post 9/11 excluding the first five months regression equation. The point, at which they crossed, October 2005, can be interpreted as the point of recovery from 9/11.

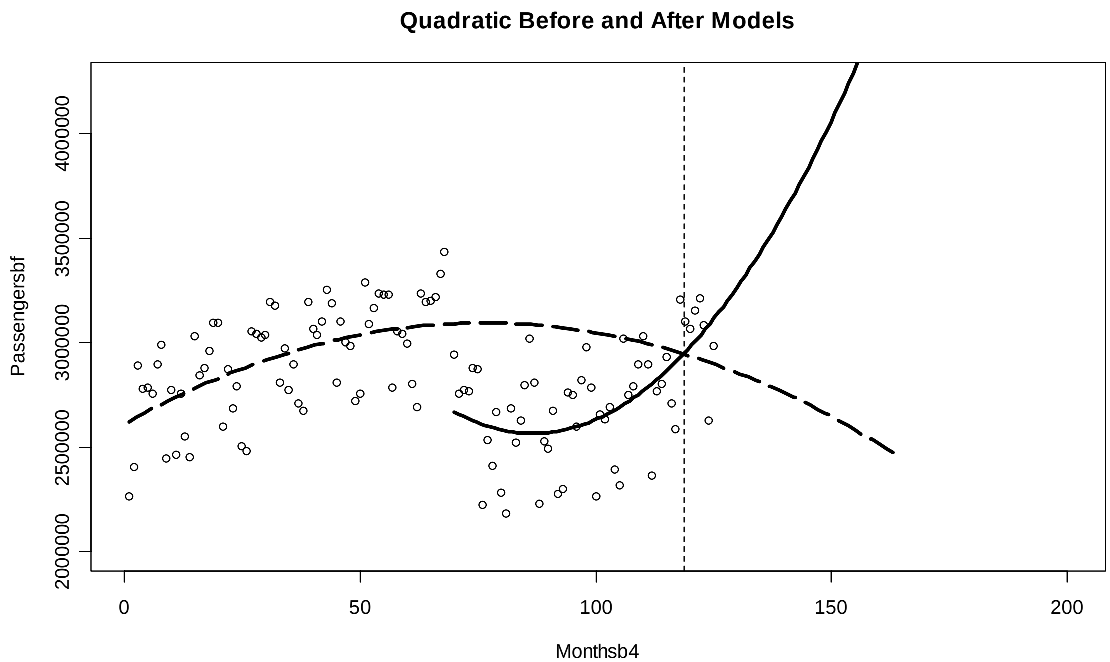

Figure 54: Continental Airlines Pre 9/11 and Post 9/11 Modified Time Series with Quadratic Regression Lines.

4.4.2 American Airlines

The American Airlines data were turned into a time series and plotted (Figure 55). The data were then decomposed and plotted (Figure 56) into their trend, seasonal, and random components. The seasonal component was removed and the time series plotted (Figure 57).

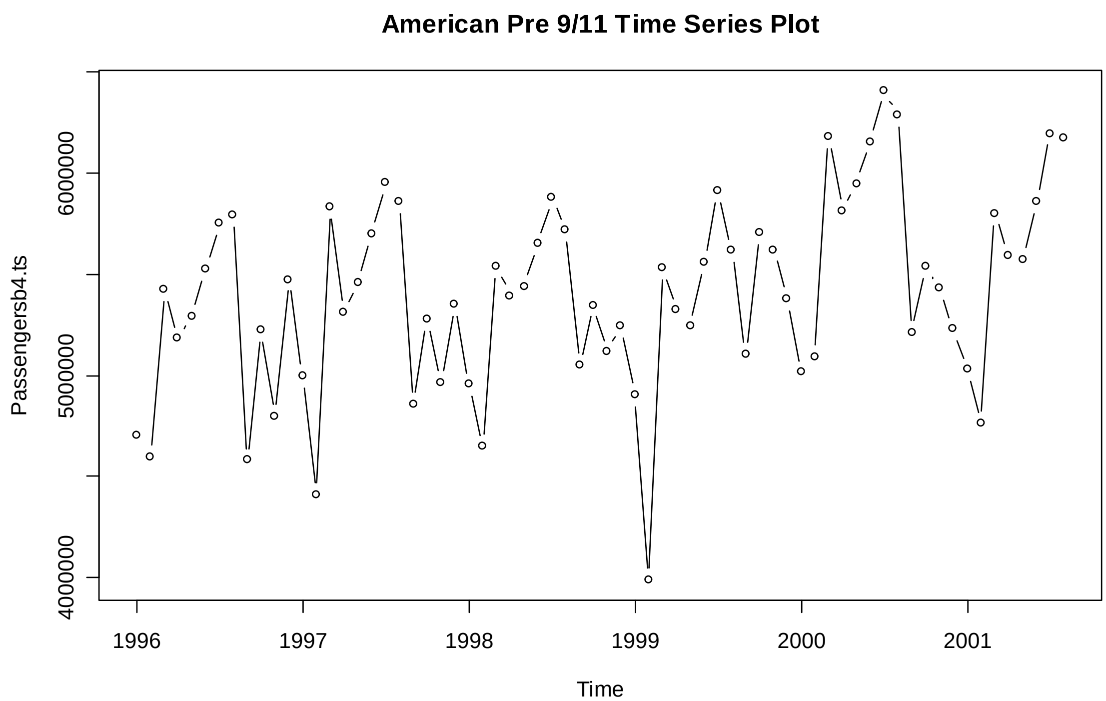

Figure 55: American Airlines Pre 9/11 Time Series Plot.

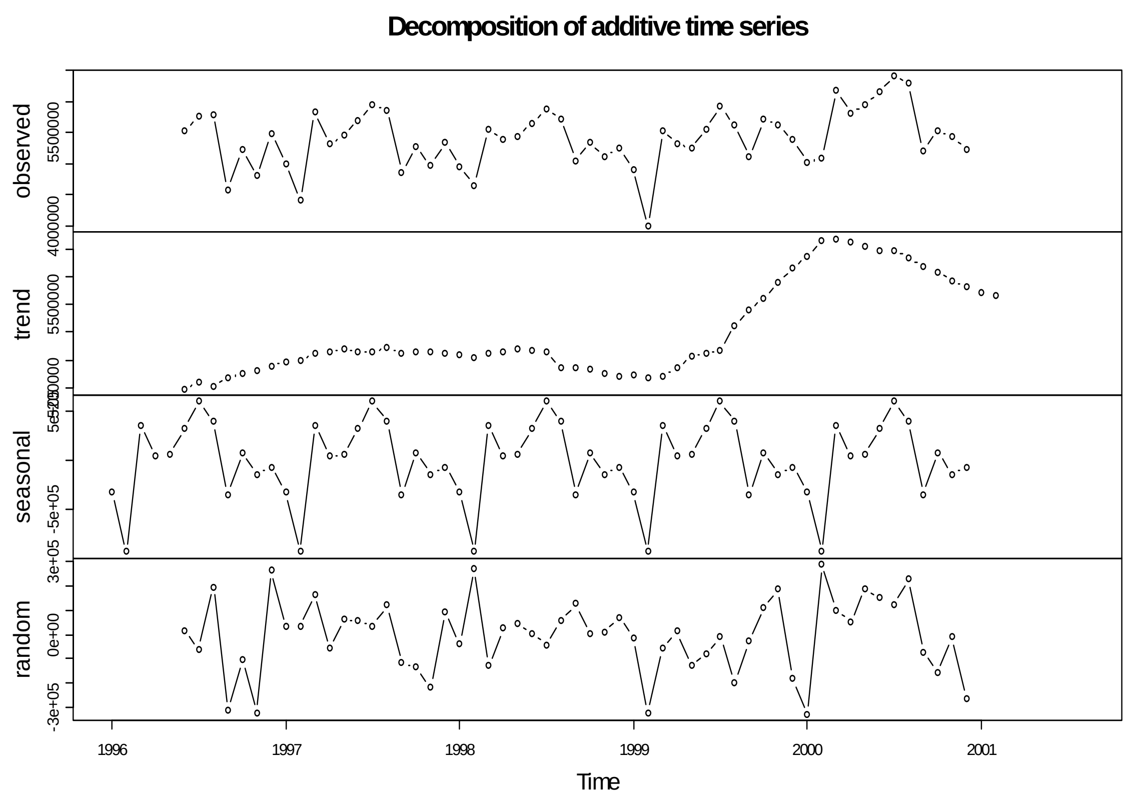

Figure 56: American Airlines Pre 9/11 Time Series Decomposition Plot.

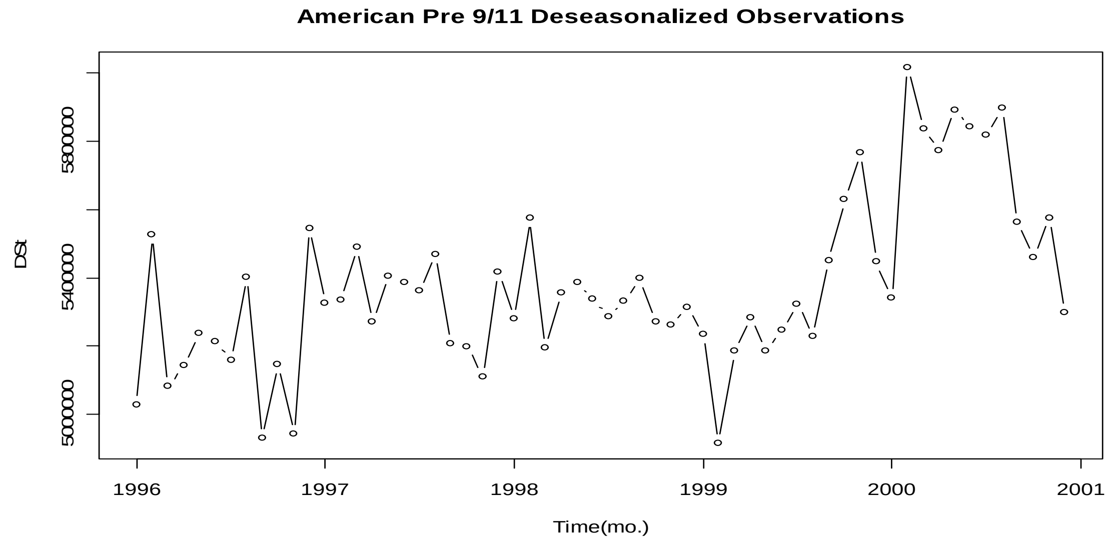

Figure 57: American Airlines Pre 9/11 Deseasonalized Time Series Plot.

The first attempt at modeling the data was to use a linear model. With an F statistic of 27.25 with corresponding p-value 2.515e-06 indicates a strong relationship between number of passengers and time (Output 44). The regression equation for American Airlines pre 9/11 time series linear model is of the form:

$$\hat{y}_t = 5132985 + 7945t \tag{71}$$

```
Call:
lm(formula = DSt ~ times.ds)

Residuals:
    Min      1Q  Median      3Q     Max 
-521109 -137421  -28284  133680  486881 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  5132985      53380   96.16  < 2e-16 ***
times.ds        7945       1522    5.22 2.51e-06 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1 

Residual standard error: 204200 on 58 degrees of freedom
Multiple R-Squared: 0.3197,     Adjusted R-squared: 0.3079 
F-statistic: 27.25 on 1 and 58 DF,  p-value: 2.515e-06
```

Output 44: Summary of American Airlines Pre 9/11 Time Series Linear Model.

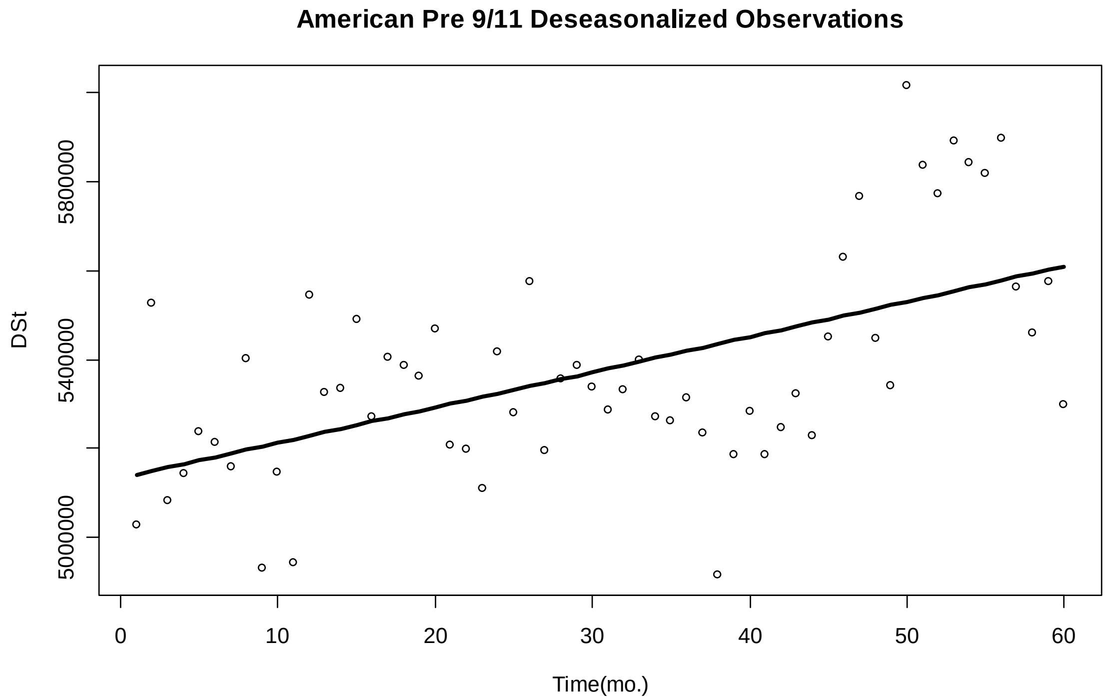

Figure 58: American Airlines Pre 9/11 Time Series with Linear Regression Line.

Visual examination of the time series with the linear regression equation (Figure 58) shows that the linear model is not a good fit.

The Shapiro-Wilk test does not give sufficient evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(DSt.lm)
W = 0.9823, p-value = 0.5324
```

The Breusch-Pagan test does not give sufficient evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  DSt.lm
BP = 3.7822, df = 1, p-value = 0.0518
```

The Durbin-Watson test indicates that there is sufficient evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  DSt.lm
DW = 1.2363, p-value = 0.0006182
alternative hypothesis: true autocorrelation is greater than 0
```

A second attempt at modeling the time series was made using a quadratic model. The F statistic of 16 along with the corresponding p-value 3.048e-06 (Output 45) indicates a relationship between number of passengers and time. The American Airlines quadratic regression equation for the pre 9/11 data is of the form:

$$\hat{y}_t = 5247276.9 - 3115.2t + 181.3t^2 \tag{72}$$

```
Call:
lm(formula = DSt ~ times.ds + I(times.ds^2))

Residuals:
    Min      1Q  Median      3Q     Max 
-476927 -126649  -22866  147128  472315 

Coefficients:
               Estimate Std. Error t value Pr(>|t|)    
(Intercept)   5247276.9    80040.1  65.558   <2e-16 ***
times.ds        -3115.2     6054.4  -0.515   0.6089    
I(times.ds^2)     181.3       96.2   1.885   0.0646 .  
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1 

Residual standard error: 199800 on 57 degrees of freedom
Multiple R-Squared: 0.3596,     Adjusted R-squared: 0.3371 
F-statistic:    16 on 2 and 57 DF,  p-value: 3.048e-06
```

Output 45: Summary of American Airlines Pre 9/11 Time Series Quadratic Model.

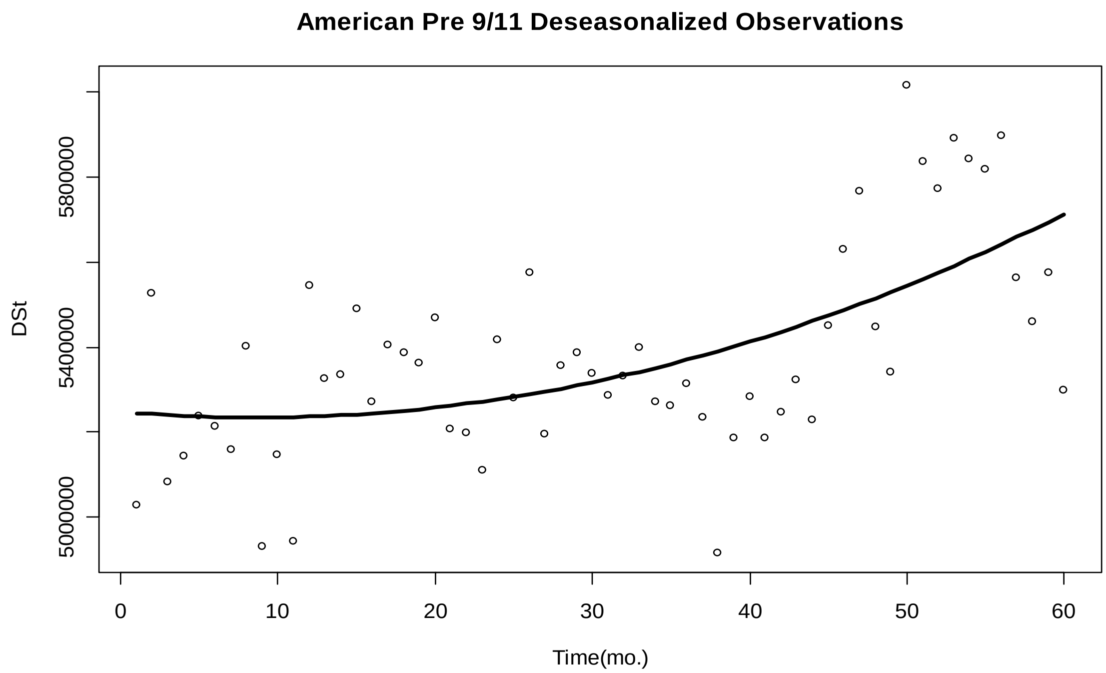

Figure 59: American Airlines Pre 9/11 Time Series with Quadratic Regression Line.

The time series plot (Figure 59) with the quadratic regression equation reveals a closer fit than did the linear model (Figure 58).

The Shapiro-Wilk test indicates that there is not sufficient evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(DStquad.lm)
W = 0.9854, p-value = 0.6908
```

The Breusch-Pagan test gives sufficient evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  DStquad.lm
BP = 6.627, df = 2, p-value = 0.03639
```

The Durbin-Watson test indicates there is sufficient evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  DStquad.lm
DW = 1.3184, p-value = 0.001263
alternative hypothesis: true autocorrelation is greater than 0
```

A final attempt to model the time series was made using a cubic model. With an *F* statistic of 10.73 and corresponding p-value 1.100e-05 indicates a strong relationship between number of passengers and time (Output 46). The American Airlines regression equation for the cubic model is of the form:

$$\hat{y}_t = 5192000 + 7331t - 243.3t^2 + 4.640t^3 \quad (73)$$

```
Call:
lm(formula = DSt ~ times.ds + I(times.ds^2) + I(times.ds^3))

Residuals:
    Min      1Q  Median      3Q     Max
-460104 -119471  -15543  143454  486738

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    5.192e+06  1.104e+05  47.013   <2e-16 ***
times.ds       7.331e+03  1.555e+04   0.471    0.639
I(times.ds^2) -2.433e+02  5.897e+02  -0.413    0.682
I(times.ds^3)  4.640e+00  6.358e+00   0.730    0.469
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 200600 on 56 degrees of freedom
Multiple R-Squared: 0.3656,     Adjusted R-squared: 0.3316
F-statistic: 10.76 on 3 and 56 DF,  p-value: 1.100e-05
```

Output 46: Summary of American Airlines Pre 9/11 Time Series Cubic Model.

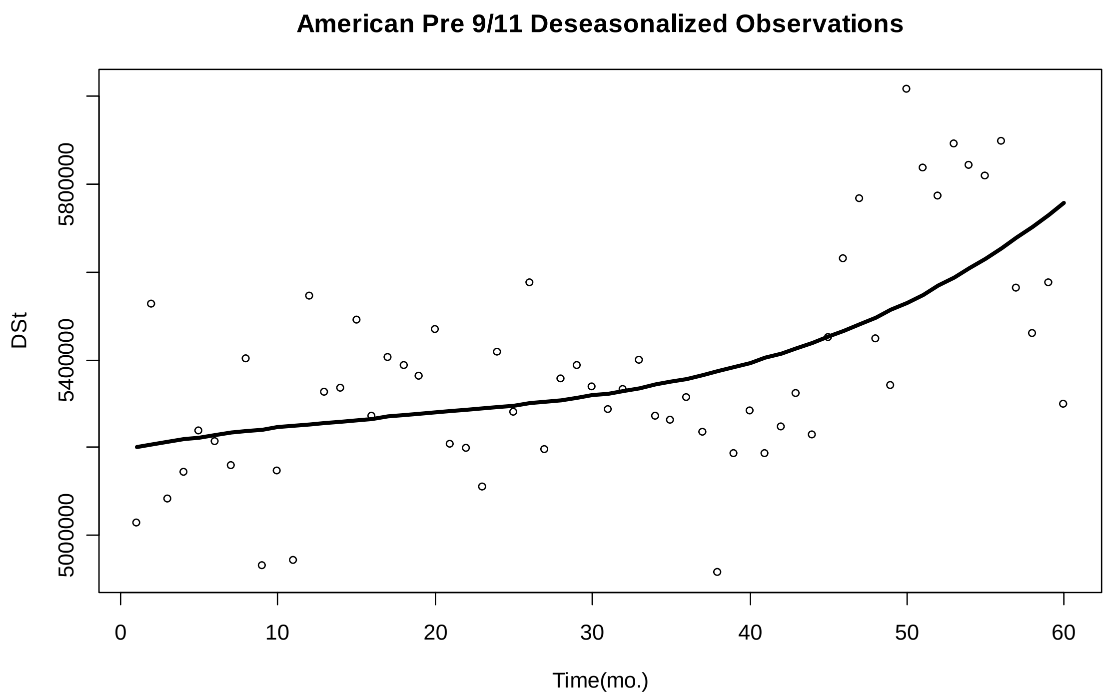

Figure 60: American Airlines Pre 9/11 Time Series with Cubic Regression Line.

Visual examination of the cubic regression equation with the time series data (Figure 60) does not seem much different than that of the quadratic model (Figure 59).

The Shapiro-Wilk test does not give sufficient evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(DStcubic.lm)
W = 0.9848, p-value = 0.6589
```

The Breusch-Pagan test indicates there is sufficient evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  DStcubic.lm
BP = 9.0197, df = 3, p-value = 0.02903
```

The Durbin-Watson test gives sufficient evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  DStcubic.lm
DW = 1.3331, p-value = 0.0009414
alternative hypothesis: true autocorrelation is greater than 0
```

The AIC of each model was compared to decide which was the best fit. Output 47 shows that the quadratic model has the smallest AIC and is the better fit.

```
AIC(DSt.lm)
[1] 1641.439
AIC( DStquad.lm)
[1] 1639.811
AIC(DStcubic.lm)
[1] 1641.243
```

Output 47: AIC Comparison of American Airlines Pre 9/11 Time Series Linear, Quadratic, and Cubic Models.

ANOVA was then used to compare the models to confirm previous findings. Output 48 shows that the t^2 parameter of the quadratic model is significant when compared to the linear model, but the t^3 parameter of the cubic model is not significant when compared to the quadratic model. Thus the previous findings were verified and any further analysis was done with the quadratic model.

```
Analysis of Variance Table

Model 1: DSt ~ times.ds
Model 2: DSt ~ times.ds + I(times.ds^2)
  Res.Df        RSS Df  Sum of Sq      F  Pr(>F)
1     58 2.4176e+12
2     57 2.2757e+12  1 1.4183e+11 3.5524 0.06456 .
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
Analysis of Variance Table

Model 1: DSt ~ times.ds + I(times.ds^2)
Model 2: DSt ~ times.ds + I(times.ds^2) + I(times.ds^3)
  Res.Df        RSS Df  Sum of Sq      F Pr(>F)
1     57 2.2757e+12
2     56 2.2543e+12  1 2.1444e+10 0.5327 0.4685
```

Output 48: ANOVA Comparison of American Airlines Pre 9/11 Time Series Linear, Quadratic, and Cubic Models.

Attention was then turned to the American Airlines post 9/11 data. The data were turned into a time series and plotted (Figure 61). The time series was then decomposed into its trend, seasonal and random components and plotted (Figure 62). The seasonal component was removed and the time series plotted (Figure 63).

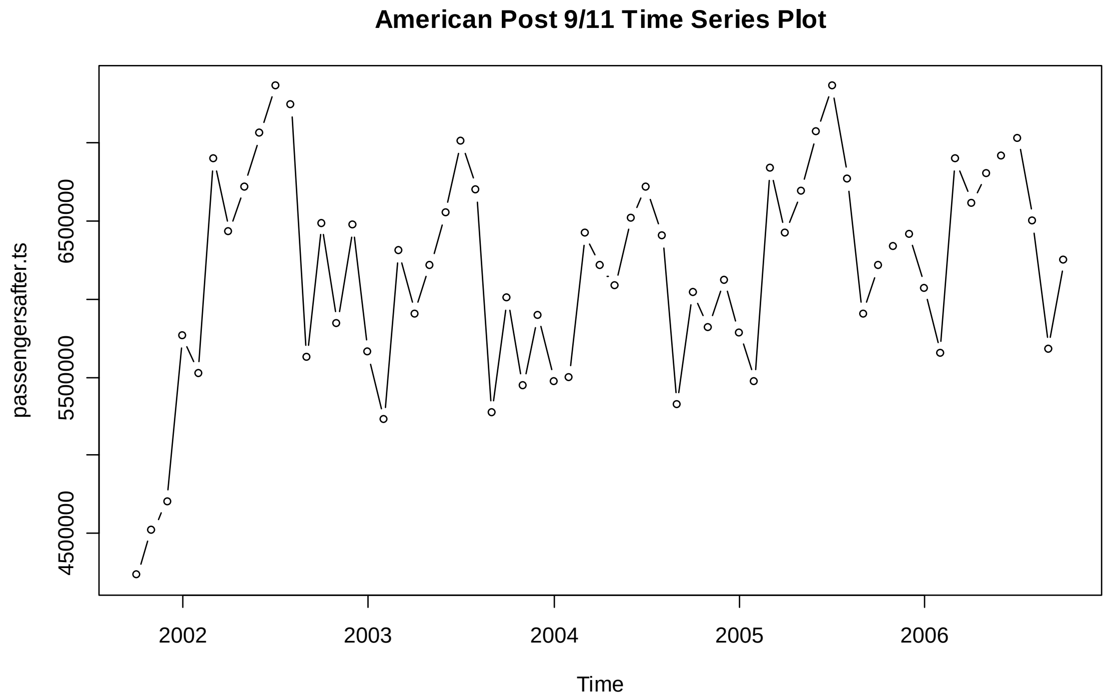

Figure 61: American Airlines Post 9/11 Time Series Plot.

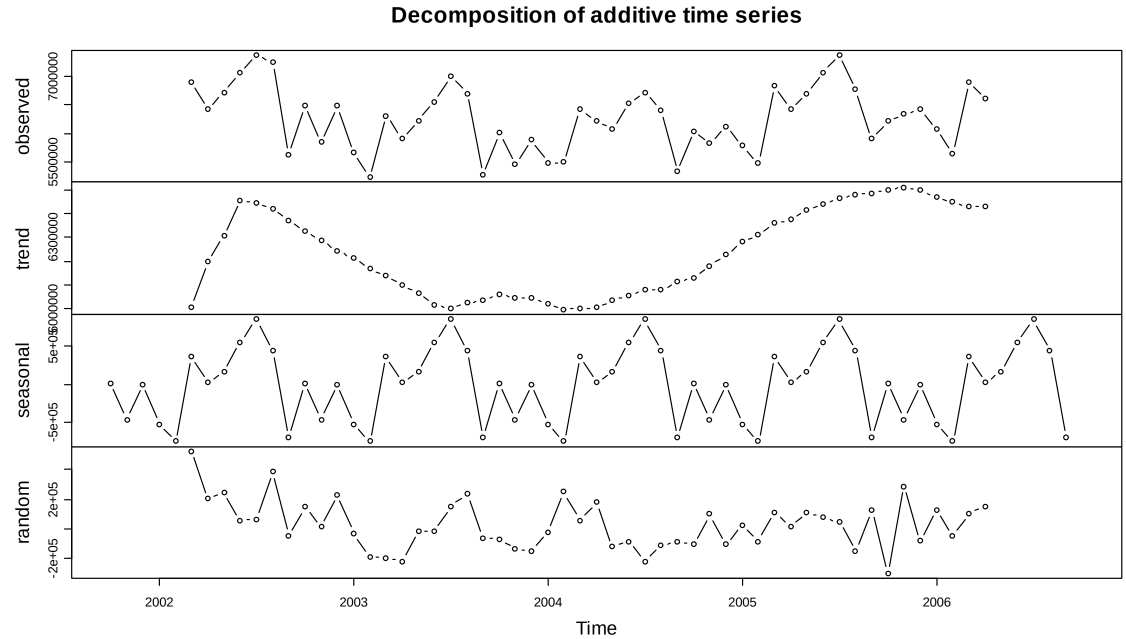

Figure 62: American Airlines Post 9/11 Time Series Decomposition Plot.

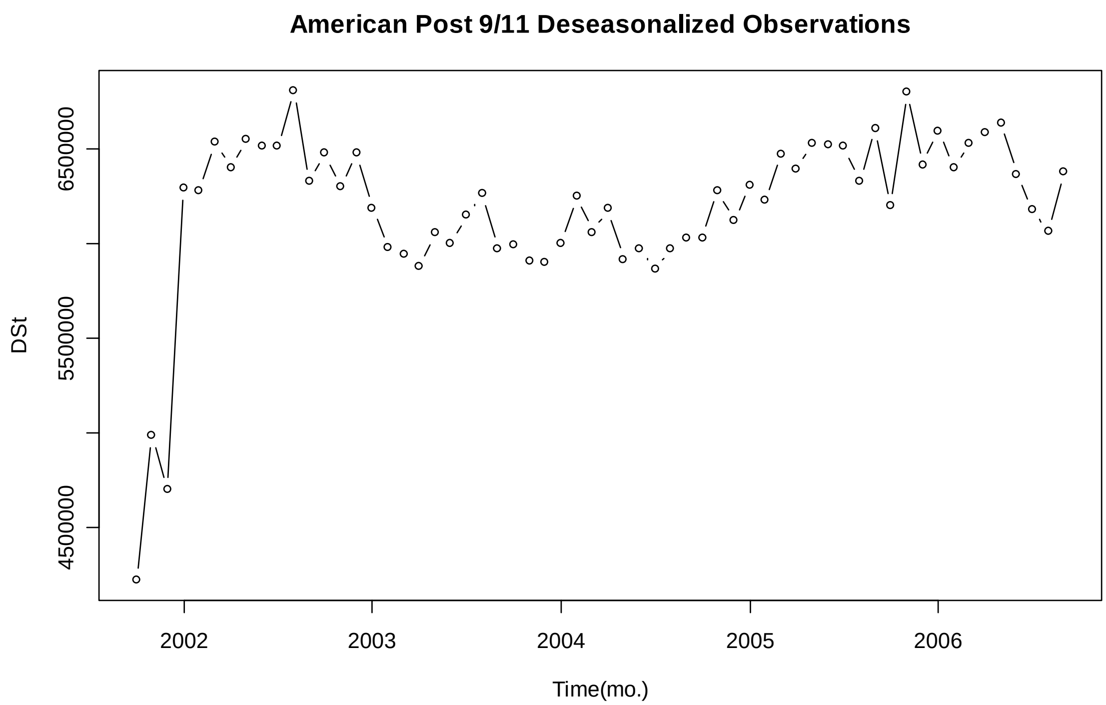

Figure 63: American Airlines Post 9/11 Deseasonalized Time Series Plot.

The first attempt made to model the time series was to use a linear model. With an *F* statistic of 10.5 with corresponding p-value .001975 there is slight evidence of a relationship between number of passengers and time (Output 49). The regression equation for the American Airlines post 9/11 time series linear model is of the form:

$$\hat{y}_t = 5205463 + 9864t \tag{74}$$

```
Call:
lm(formula = DSt ~ times.ds)

Residuals:
     Min       1Q   Median       3Q      Max 
-1678862  -165737     5688   200381   812046 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  5205463     307386  16.935  < 2e-16 ***
times.ds        9864       3044   3.241  0.00198 ** 
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Residual standard error: 408300 on 58 degrees of freedom
Multiple R-Squared: 0.1533,     Adjusted R-squared: 0.1387 
F-statistic:  10.5 on 1 and 58 DF,  p-value: 0.001975
```

Output 49: Summary of American Airlines Post 9/11 Time Series Linear Model.

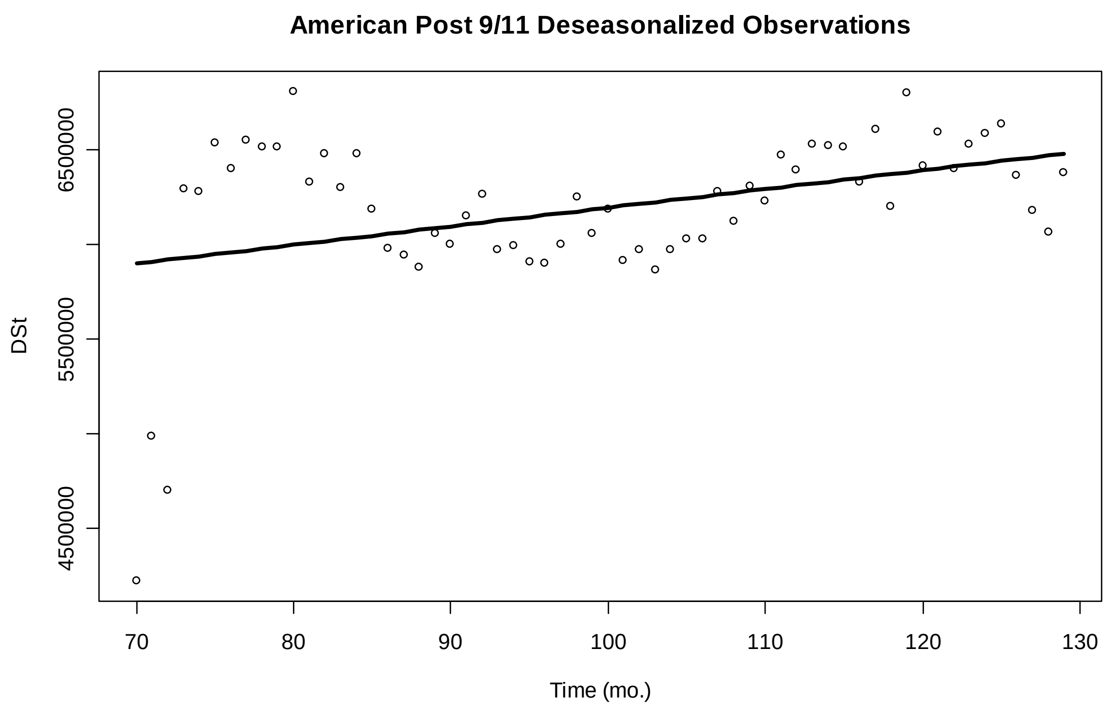

Figure 64: American Airlines Post 9/11 Time Series with Linear Regression Line.

The time series with the linear regression equation (Figure 64) does not seem like a good fit.

The Shapiro-Wilk test gives sufficient evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(DSt.lm)
W = 0.8789, p-value = 2.429e-05
```

The Breusch-Pagan test gives sufficient evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  DSt.lm
BP = 11.9289, df = 1, p-value = 0.0005527
```

The Durbin-Watson test gives sufficient evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  DSt.lm
DW = 0.5675, p-value = 6.896e-12
alternative hypothesis: true autocorrelation is greater than 0
```

A second attempt at modeling the American Airlines post 9/11 time series was made using a quadratic model. With an *F* statistic of 5.486 with corresponding p-value .006625 indicates a relationship between number of passengers and time (Output 50). The regression equation for American Airlines post 9/11time series quadratic model is of the form:

$$\hat{y}_t = 3802011.9 + 38955.5t - 146.2t^2 \tag{75}$$

```
Call:
lm(formula = DSt ~ times.ds + I(times.ds^2))

Residuals:
     Min       1Q   Median       3Q      Max
-1595486  -193564   -10442   226102   823790

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)   3802011.9  1919484.7   1.981   0.0525 .
times.ds        38955.5    39389.1   0.989   0.3268
I(times.ds^2)    -146.2      197.3  -0.741   0.4619
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 409900 on 57 degrees of freedom
Multiple R-Squared: 0.1614,     Adjusted R-squared: 0.132
F-statistic: 5.486 on 2 and 57 DF,  p-value: 0.006625
```
Output 50: Summary of American Airlines Post 9/11 Time Series Quadratic Regression Model.

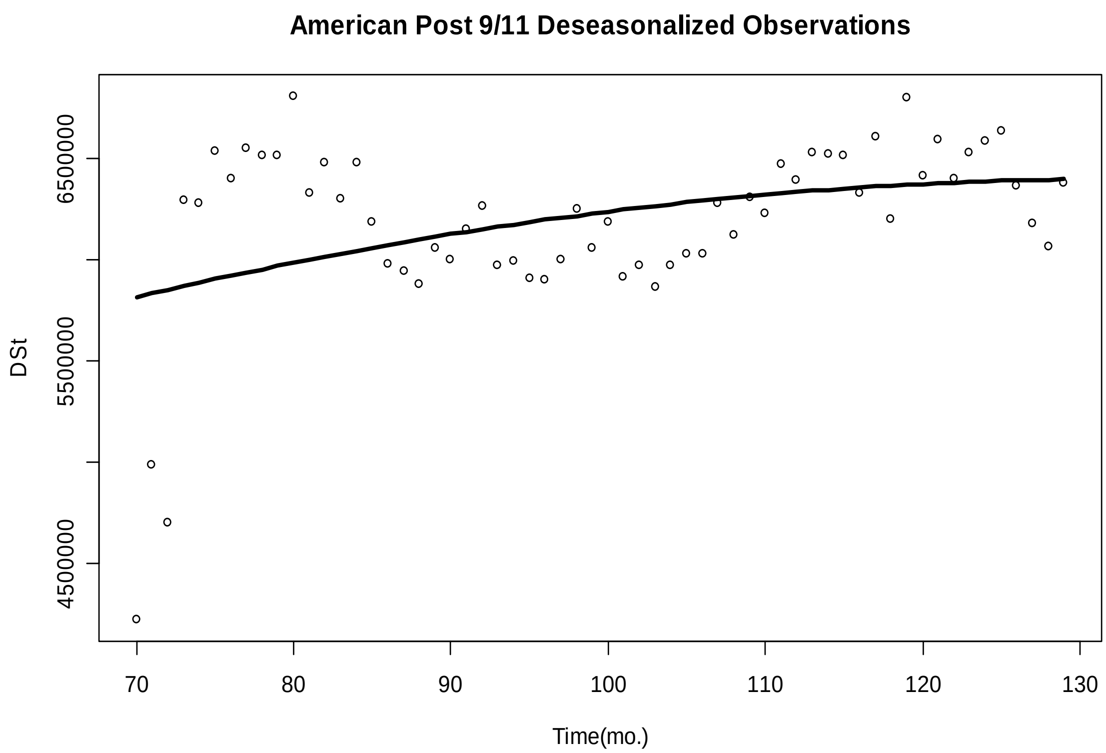

Figure 65: American Airlines Post 9/11 Time Series Plot with Quadratic Regression Line.

Visual examination of the post 9/11 time series model with the quadratic regression (Figure 65) line indicates a slight improvement in fit from the linear model (Figure 64).

The Shapiro-Wilk test indicates sufficient evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(DStquad.lm)
W = 0.9093, p-value = 0.0002913
```

The Breusch-Pagan test indicates sufficient evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  DStquad.lm
BP = 26.0653, df = 2, p-value = 2.188e-06
```

The Durbin-Watson test gives sufficient evidence to reject the independence assumption.

```
   Durbin-Watson test
data:  DStquad.lm
DW = 0.5699, p-value = 2.368e-12
alternative hypothesis: true autocorrelation is greater than 0
```

A final attempt at modeling the American Airlines post 9/11 time series was made using a cubic model. The *F* statistic of 5.623 with corresponding p-value .001928 shows a relationship between number of passengers and time (Output 51). The regression equation for the American Airlines time series cubic model is of the form:

$$\hat{y}_t = -22600000 + 865600t - 8608t^2 + 28.35t^3 \tag{76}$$

```
Call:
lm(formula = DSt ~ times.ds + I(times.ds^2) + I(times.ds^3))

Residuals:
     Min       1Q   Median       3Q      Max
-1319006  -247395     7543   282494   735673

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)   -2.260e+07  1.183e+07  -1.911   0.0612 .
times.ds       8.656e+05  3.678e+05   2.354   0.0221 *
I(times.ds^2) -8.608e+03  3.749e+03  -2.296   0.0254 *
I(times.ds^3)  2.835e+01  1.254e+01   2.260   0.0277 *
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 395900 on 56 degrees of freedom
Multiple R-Squared: 0.2315,     Adjusted R-squared: 0.1903
F-statistic: 5.623 on 3 and 56 DF,  p-value: 0.001928
```

Output 51: Summary of American Airlines Post 9/11 Time Series Cubic Model.

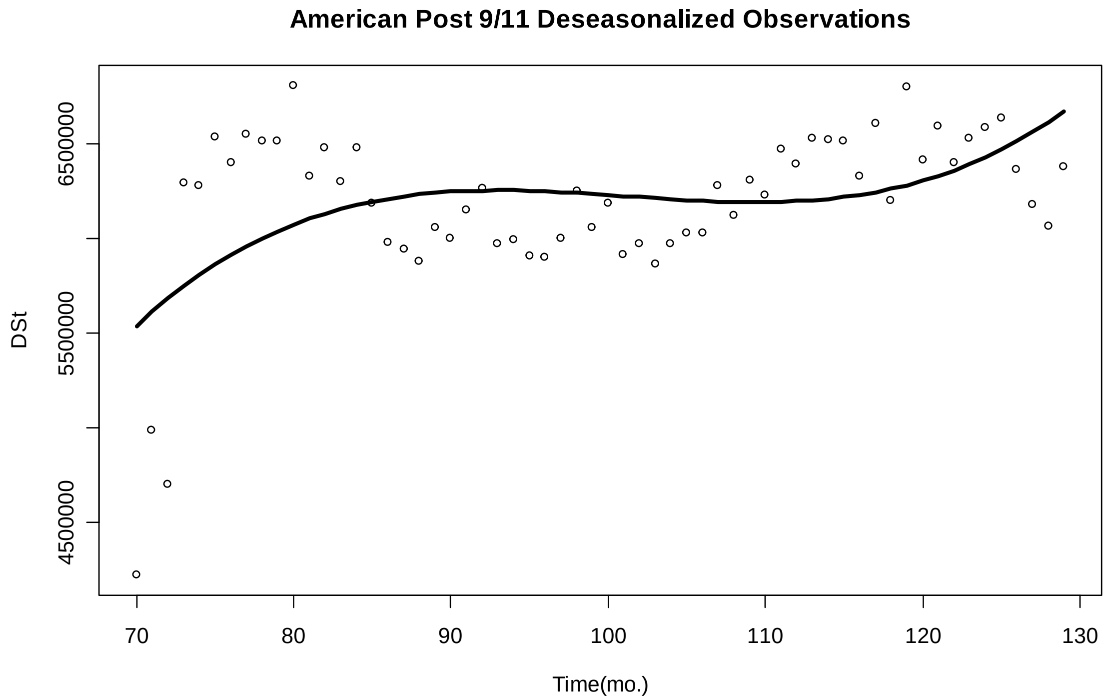

Figure 66: American Airlines Post 9/11 Time Series with Cubic Regression Line.

The cubic regression line (Figure 66) does not seem to improve the fit from the quadratic model (Figure 65). The S curvature seems backwards from the majority of the data.

The Shapiro-Wilk test gives sufficient evidence to reject the normality assumption.

```
   Shapiro-Wilk normality test
data:  resid(DStcubic.lm)
W = 0.9596, p-value = 0.04478
```

The Breusch-Pagan test gives sufficient evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  DStcubic.lm
BP = 35.4027, df = 3, p-value = 1.002e-07
```

The Durbin-Watson test indicates that there is sufficient evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  DStcubic.lm
DW = 0.6037, p-value = 3.387e-12
alternative hypothesis: true autocorrelation is greater than 0
```

Using the AIC statistics the models were compared to determine which was the best fit. Output 52 indicates that the cubic model best fits the time series.

```
AIC (DSt.lm)
[1] 1724.603
AIC(DStquad.lm)
[1] 1726.028
AIC(DStcubic.lm)
[1] 1722.791
```

Output 52: AIC Comparison of American Airlines Post 9/11 Time Series Linear, Quadratic, and Cubic Models.

ANOVA comparisons of the three models were made to confirm the previous findings. Output 53 shows that the t^2 parameter of the quadratic model is not significant in comparison to the linear model, but the t^3 parameter of the cubic model is significant when compared to the quadratic model. Thus previous findings were confirmed.

```
Analysis of Variance Table
Model 1: DSt ~ times.ds
Model 2: DSt ~ times.ds + I(times.ds^2)
  Res.Df        RSS Df  Sum of Sq      F Pr(>F)
1     58 9.6681e+12
2     57 9.5759e+12  1 9.2194e+10 0.5488 0.4619


Analysis of Variance Table
Model 1: DSt ~ times.ds + I(times.ds^2)
Model 2: DSt ~ times.ds + I(times.ds^2) + I(times.ds^3)
  Res.Df        RSS Df  Sum of Sq      F  Pr(>F)
1     57 9.5759e+12
2     56 8.7755e+12  1 8.0037e+11 5.1074 0.02773 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Output 53: ANOVA Comparison of American Airlines Time Series Linear, Quadratic, and Cubic Models.

As in previous analysis, the post 9/11 models were examined for outliers using the residual plots and influential outliers using the leverage plot and Cook's Distance statistic. For the linear mode the residual plots (Figure 67) all singled out observations 1, 2, and 3. The leverage plot (Figure 67) and Cook's D test (Output 54) did not indicate any influential outliers.

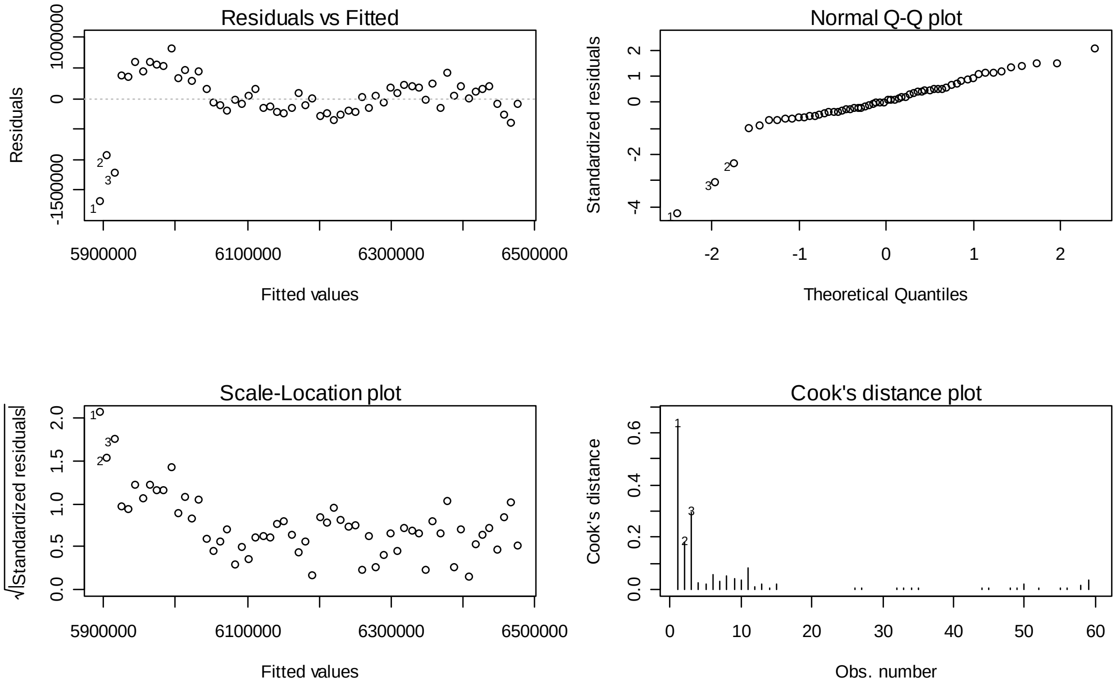

Figure 67: American Airlines Post 9/11 Time Series Linear Model Residual Plots.

```
cooksD=cooks.distance(DSt.lm)
f0.50=qf(0.5,df1=5,df2=24)
cooksD[which(cooksD>f0.50)]
named numeric(0)
```

Output 54: Cook's D Statistic of American Airlines Post 9/11 Time Series Linear Model.

For the quadratic model (Figure 68) all residual plots singled out observations 1, 2, and 3. The leverage plot (Figure 68) and Cook's D test (Output 55) indicated that observation 1 was an influential outlier.

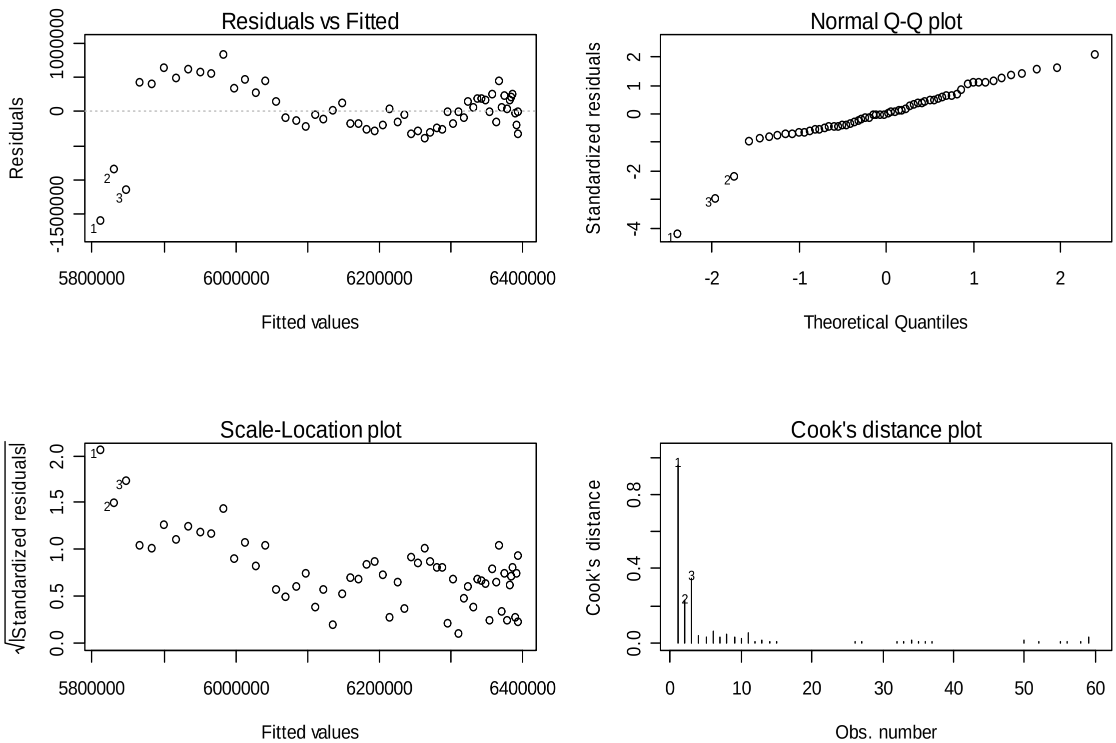

Figure 68: American Airlines Post 9/11 Time Series Quadratic Model Residual Plots.

```
cooksD=cooks.distance(DStquad.lm)
f0.50=qf(0.5, df1=5, df2=24)
cooksD[which(cooksD>f0.50)]
         1
0.9599551
```

Output 55: Cook's D Statistic on American Airlines Post 9/11 Time Series Quadratic Model.

For the cubic model (Figure 69) observations 1, 3, and 11 were singled out. The leverage plot (Figure 69) and Cook's D test (Output 56) both indicated observation 1 as an influential outlier

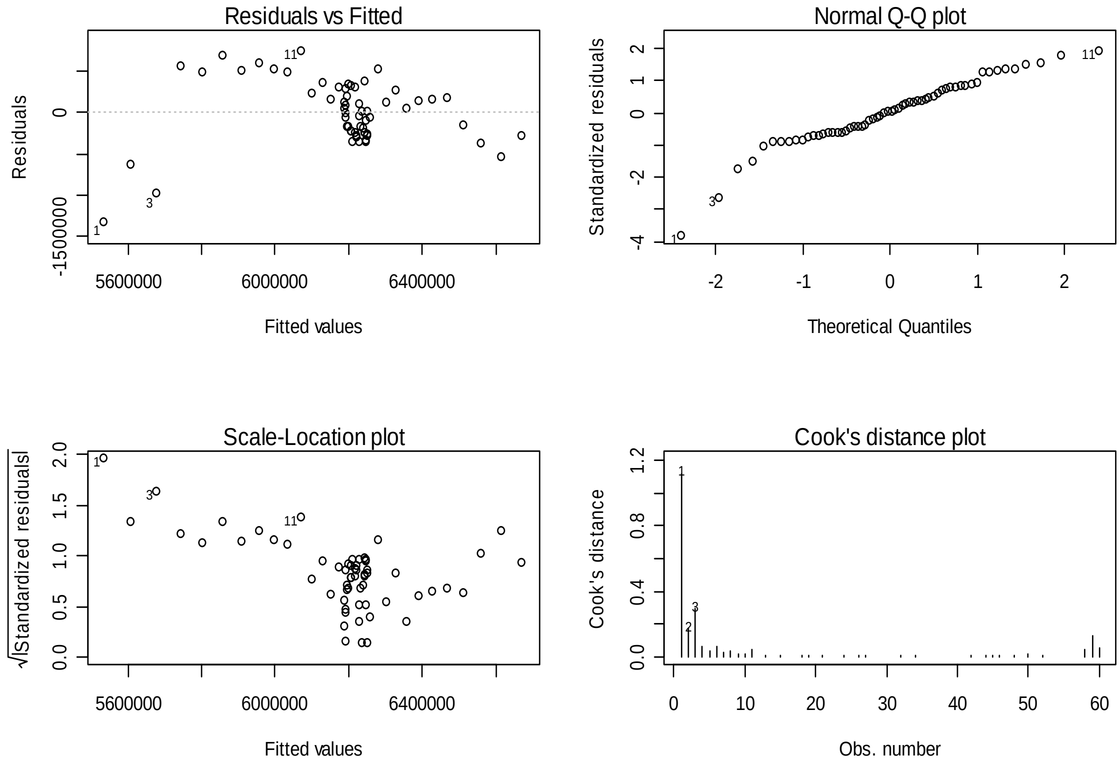

Figure 69: American Airlines Post 9/11 Time Series Cubic Model Residual Plots.

```
cooksD=cooks.distance(DStcubic.lm)
f0.50=qf(0.5, df1=5, df2=24)
cooksD[which(cooksD>f0.50)]
       1
1.121718
```

Output 56: Cook's D Statistic on American Airlines Post 9/11 Time Series Cubic Model.

Observations 1, 2, and 3 were determined to be the initial shock recovery period after the incident. These observations were removed from the post 9/11 dataset and the analysis was run again.

The American Airlines post 9/11 modified data were turned into a time series and plotted (Figure 75). The model was decomposed into its trend, seasonal, and random components and plotted (Figure 76). The seasonal component was removed and the time series was plotted (Figure 77).

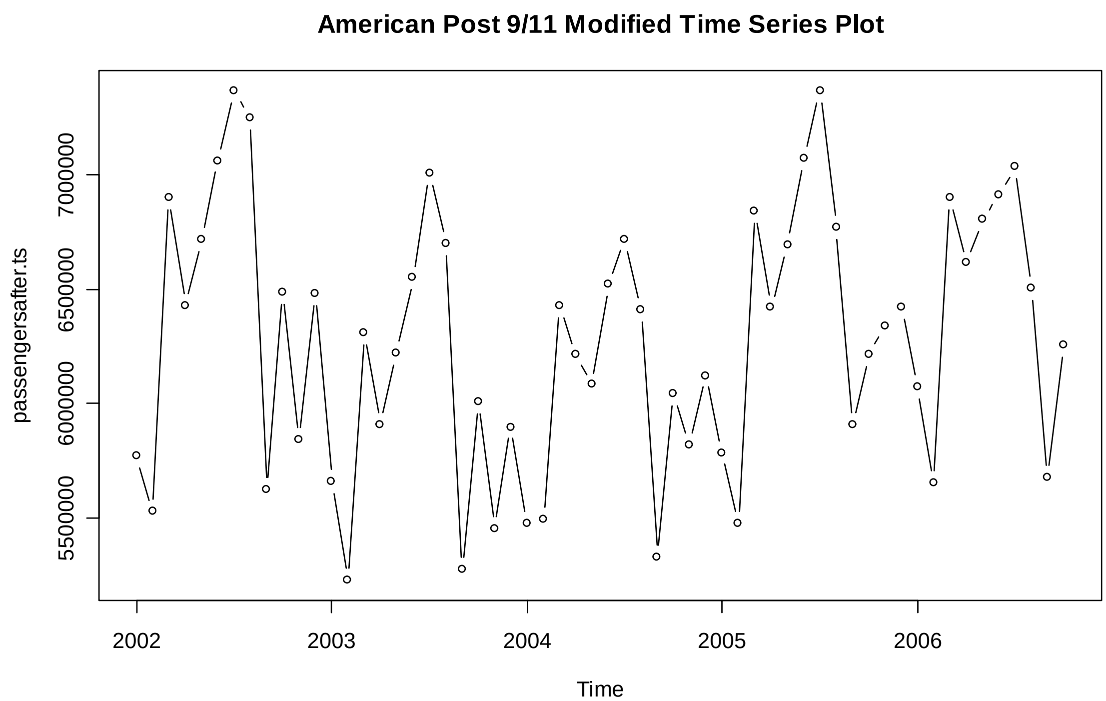

Figure 70: American Airlines Post 9/11 Modified Time Series Plot.

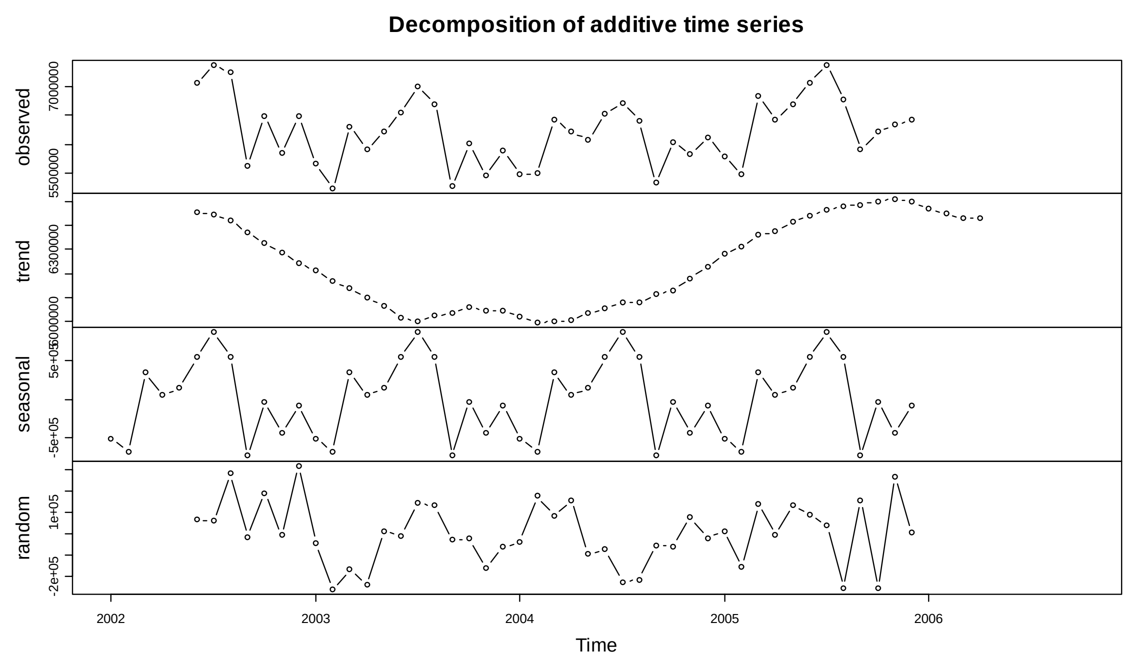

Figure 71: American Airlines Post 9/11 Modified Time Series Decomposition Plot.

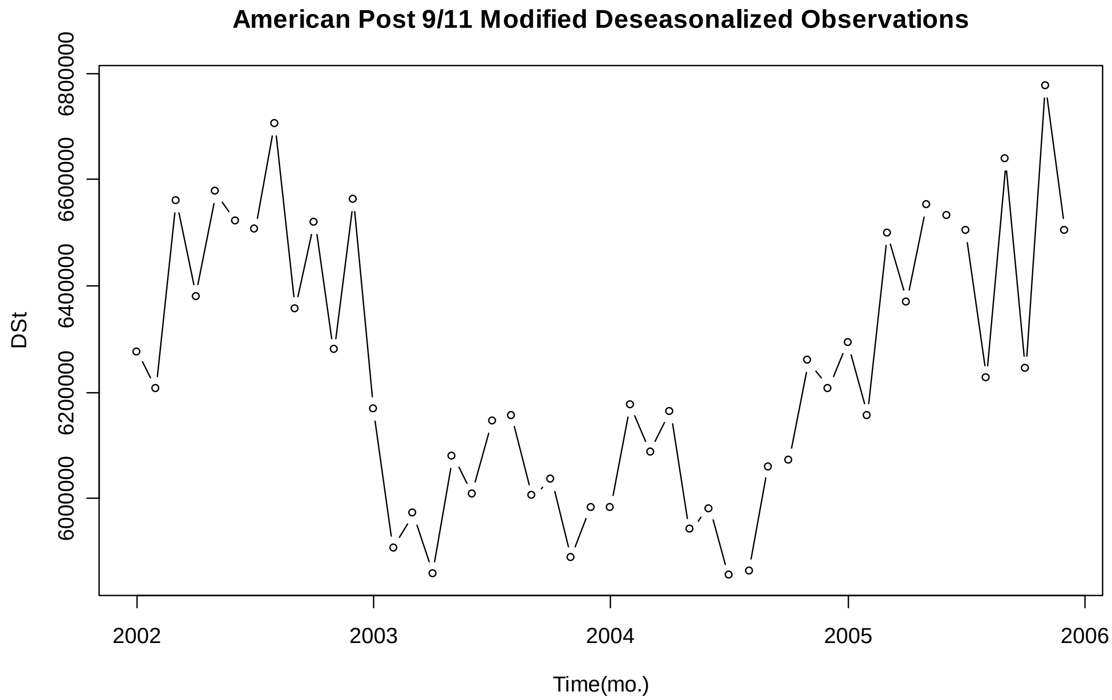

Figure 72: American Airlines Post 9/11 Modified Time Series Deseasonalized Plot.

The first attempt at modeling the time series was done with a linear model. With an F statistic of .057 and associated p-value .8124 (Output 57) there does not seem to be a relationship between number of passengers and month. The American Airlines post 9/11 modified time series linear regression equation is of the form:

$$\hat{y}_t = 6182346.2 + 633.6t \tag{77}$$

```
Call:
lm(formula = DSt ~ times.ds)

Residuals:
    Min      1Q  Median      3Q     Max
-390776 -210942  -34131  249283  520721

Coefficients:
             Estimate Std. Error t value Pr(>|t|)
(Intercept) 6182346.2   250841.0  24.646   <2e-16 ***
times.ds        633.6     2653.8   0.239    0.812
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 254700 on 46 degrees of freedom
Multiple R-Squared: 0.001238,   Adjusted R-squared: -0.02047
F-statistic: 0.057 on 1 and 46 DF,  p-value: 0.8124
```

Output 57: Summary on American Airlines Modified Time Series Linear Model.

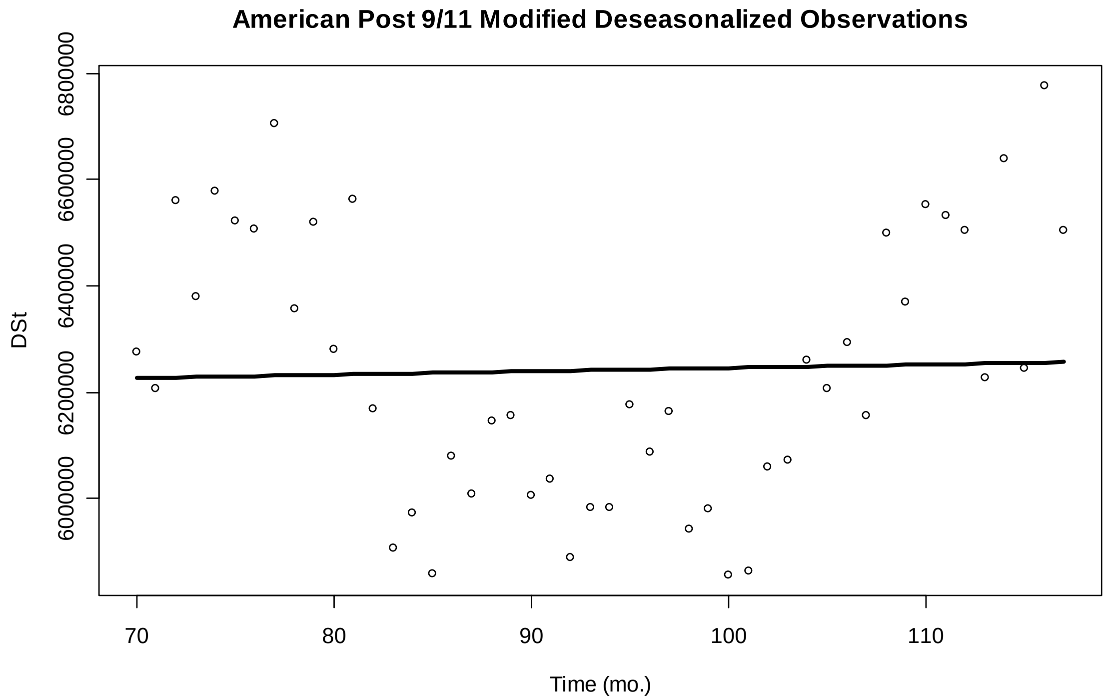

Figure 73: American Airlines Post 9/11 Modified Time Series with Linear Regression Line.

The time series plot with the linear regression line is not a good fit (Figure 73).

The Shapiro-Wilk test does not give sufficient evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(DSt.lm)
W = 0.9527, p-value = 0.0514
```

The Breusch-Pagan test does not give sufficient evidence to reject the constant variance assumption.

```
   studentized Breusch-Pagan test
data:  DSt.lm
BP = 0.0122, df = 1, p-value = 0.9121
```

The Durbin-Watson test gives enough evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  DSt.lm
DW = 0.6845, p-value = 2.919e-08
alternative hypothesis: true autocorrelation is greater than 0
```

The second attempt at modeling the time series was made using a quadratic model. With an F statistic of 24.89 and associated p-value 5.272e-08 there is a strong relationship between number of passengers and time (Output 58). The quadratic regression equation for the American Airlines post 9/11 time series modified model is of the form:

$$\hat{y}_t = 15184038.2 - 196237.9t + 1052.8t^2 \tag{78}$$

```
Call:
lm(formula = DSt ~ times.ds + I(times.ds^2))

Residuals:
    Min      1Q  Median      3Q     Max 
-350743 -101291    9369  114075  388270 

Coefficients:
                Estimate Std. Error t value Pr(>|t|)    
(Intercept)   15184038.2  1289363.7  11.776 2.44e-15 ***
times.ds       -196237.9    27999.7  -7.009 9.90e-09 ***
I(times.ds^2)     1052.8      149.4   7.047 8.69e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Residual standard error: 177600 on 45 degrees of freedom
Multiple R-Squared: 0.5252,     Adjusted R-squared: 0.5041 
F-statistic: 24.89 on 2 and 45 DF,  p-value: 5.272e-08
```

Output 58: Summary of American Airlines Post 9/11 Modified Time Series Quadratic Model.

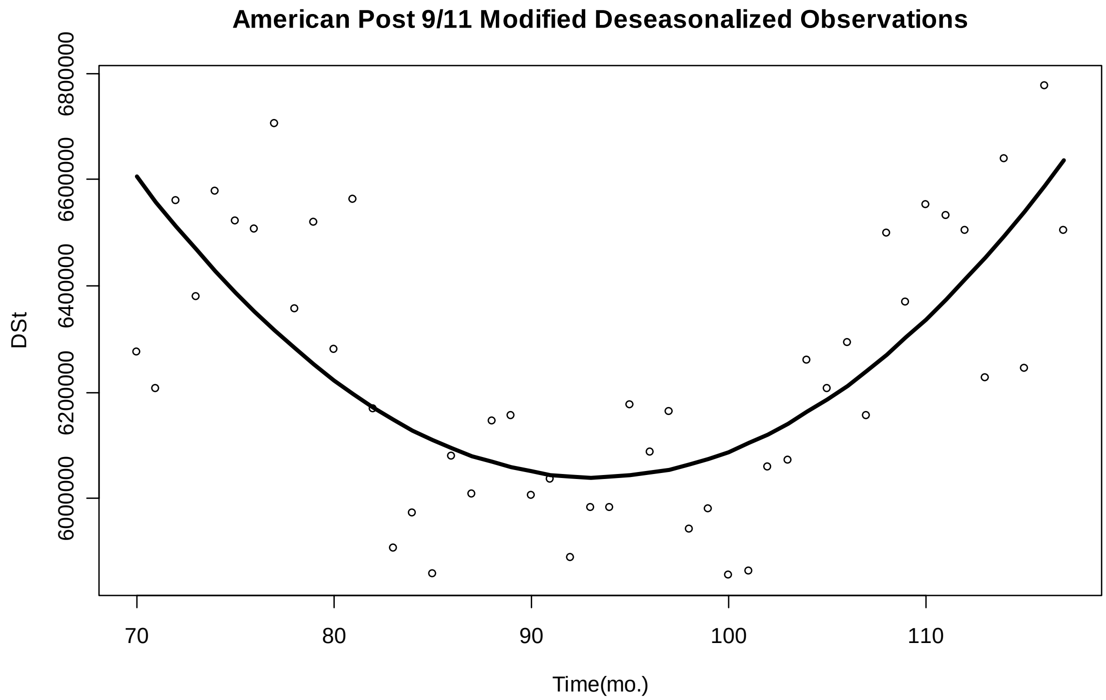

Figure 74: American Airlines Post 9/11 Modified Time Series with Quadratic Regression Line.

The time series plot (Figure 74) with the quadratic regression line indicates a nice fit.

The Shapiro-Wilk test does not give sufficient evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(DStquad.lm)
W = 0.984, p-value = 0.7504
```

The Breusch-Pagan test gives sufficient evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  DStquad.lm
BP = 8.7718, df = 2, p-value = 0.01245
```

The Durbin-Watson test gives sufficient evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  DStquad.lm
DW = 1.4457, p-value = 0.01092
alternative hypothesis: true autocorrelation is greater than 0
```

A last attempt at modeling the time series was made using a cubic model. With an F statistic of 17.15 and associated p-value 1.601e-07 (Output 59) there is a strong relationship between number of passengers and time. The cubic regression equation for the American Airlines time series modified model is of the form:

$$\hat{y}_t = 4099958.49 + 169149.34t - 2907.23t^2 + 14.12t^3 \tag{79}$$

```
Call:
lm(formula = DSt ~ times.ds + I(times.ds^2) + I(times.ds^3))

Residuals:
    Min      1Q  Median      3Q     Max
-331047  -99908    1802  126027  371265

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)   4099958.49 9707629.94   0.422    0.675
times.ds       169149.34  318422.28   0.531    0.598
I(times.ds^2)   -2907.23    3440.97  -0.845    0.403
I(times.ds^3)      14.12      12.26   1.152    0.256

Residual standard error: 176900 on 44 degrees of freedom
Multiple R-Squared: 0.5391,     Adjusted R-squared: 0.5076
F-statistic: 17.15 on 3 and 44 DF,  p-value: 1.601e-07
```

Output 59: Summary of American Airlines Post 9/11 Modified Time Series Cubic Model.

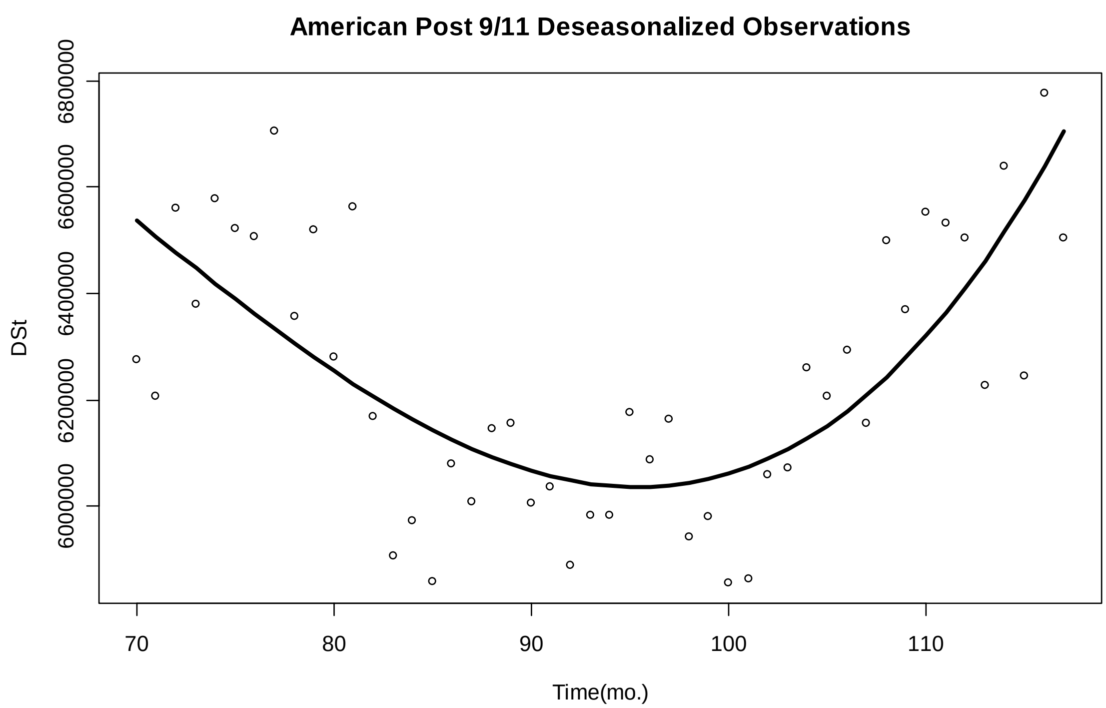

Figure 75: American Airlines Post 9/11 Modified Time Series with Cubic Regression Line.

The time series fit with the cubic regression line (Figure 75) does not seem much different that that of the quadratic model (Figure 74).

The Shapiro-Wilk test does not give sufficient evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(DStcubic.lm)
W = 0.9755, p-value = 0.4081
```

The Breusch-Pagan test does not give sufficient evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  DStcubic.lm
BP = 7.059, df = 3, p-value = 0.07004
```

The Durbin-Watson test indicates that there is sufficient evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  DStcubic.lm
DW = 1.4857, p-value = 0.01028
alternative hypothesis: true autocorrelation is greater than 0
```

A comparison of the AIC of all three models was done to determine which was the best fit. Output 60 shows that the AIC for the quadratic model is smallest, and thus the best fit.

```
AIC (DSt.lm)
[1] 1335.172
AIC(DStquad.lm)
[1] 1301.481
AIC(DStcubic.lm)
[1] 1302.055
```

Output 60: AIC Comparison of American Airlines Post 9/11 Modified Time Series Linear, Quadratic, and Cubic Models.

ANOVA comparisons were then made of all three models. Output 61 shows that the t^2 parameter of the quadratic model is significant in comparison to the linear model. Also, the t^3 parameter of the cubic model is not significant when compared to the quadratic model. Thus the quadratic model was deemed the best fit.

```
Analysis of Variance Table

Model 1: DSt ~ times.ds
Model 2: DSt ~ times.ds + I(times.ds^2)
  Res.Df        RSS Df  Sum of Sq      F    Pr(>F)
1     46 2.9844e+12
2     45 1.4188e+12  1 1.5656e+12 49.655 8.691e-09 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

anova(DStquad.lm,DStcubic.lm)

```
Analysis of Variance Table

Model 1: DSt ~ times.ds + I(times.ds^2)
Model 2: DSt ~ times.ds + I(times.ds^2) + I(times.ds^3)
  Res.Df        RSS Df  Sum of Sq      F Pr(>F)
1     45 1.4188e+12
2     44 1.3773e+12  1 4.1535e+10 1.3269 0.2556
```

Output 61: ANOVA Comparison of American Airlines Post 9/11 Time Series Linear, Quadratic, and Cubic Models.

Comparing the AIC for the post 9/11 quadratic model that included all data (Output 52) to the post 9/11 quadratic modified model (Output 60), it is determined that the post 9/11 modified data quadratic model is the best fit.

The regression equations for the pre 9/11 quadratic model and post 9/11 modified quadratic model were then set equal and solved for time. Figure 76 shows the pre 9/11 and post 9/11 quadratic models with crossing point which represents August 2007. This indicates that not until August of 2007 did American Airlines recover from the attack on September 11, 2001.

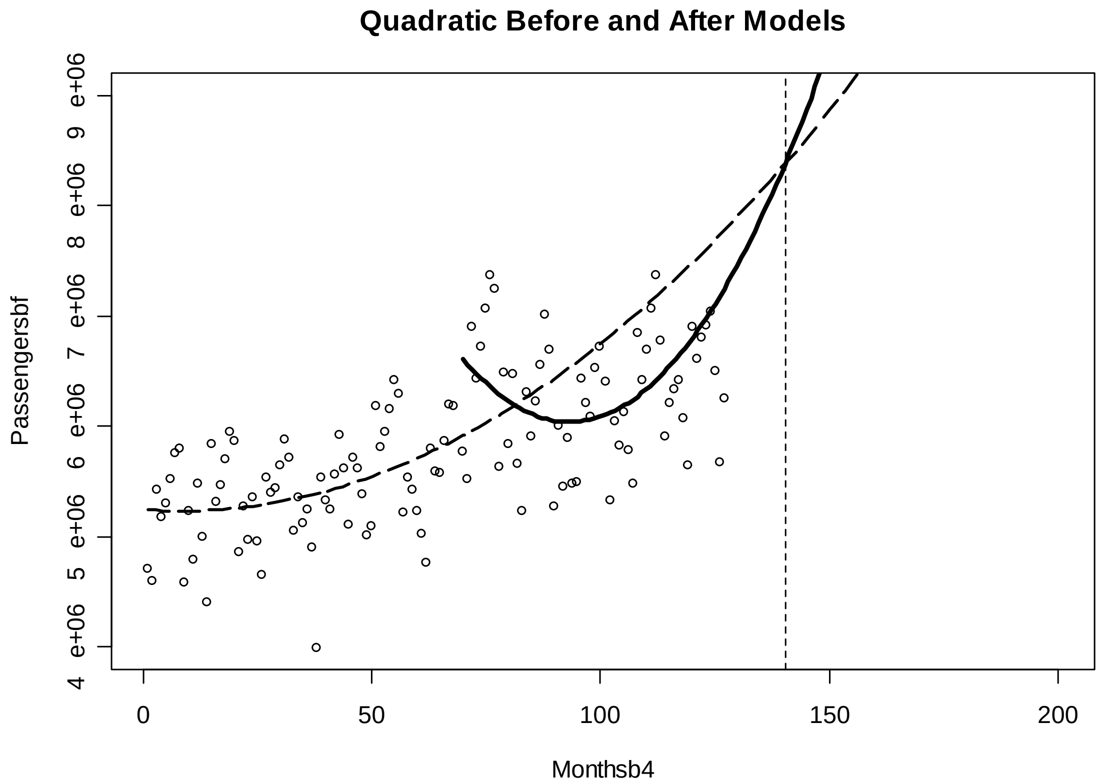

Figure 76: American Airlines Pre 9/11 and Post 9/11 Modified Time Series with Quadratic Regression Lines.

4.4.3 Delta Airlines

Delta Airlines pre 9/11 data were turned into a time series and plotted (Figure 77). The time series was decomposed into its trend, seasonal and random components (Figure 78). The seasonal component was removed and deseasonalized time series plotted (Figure 79).

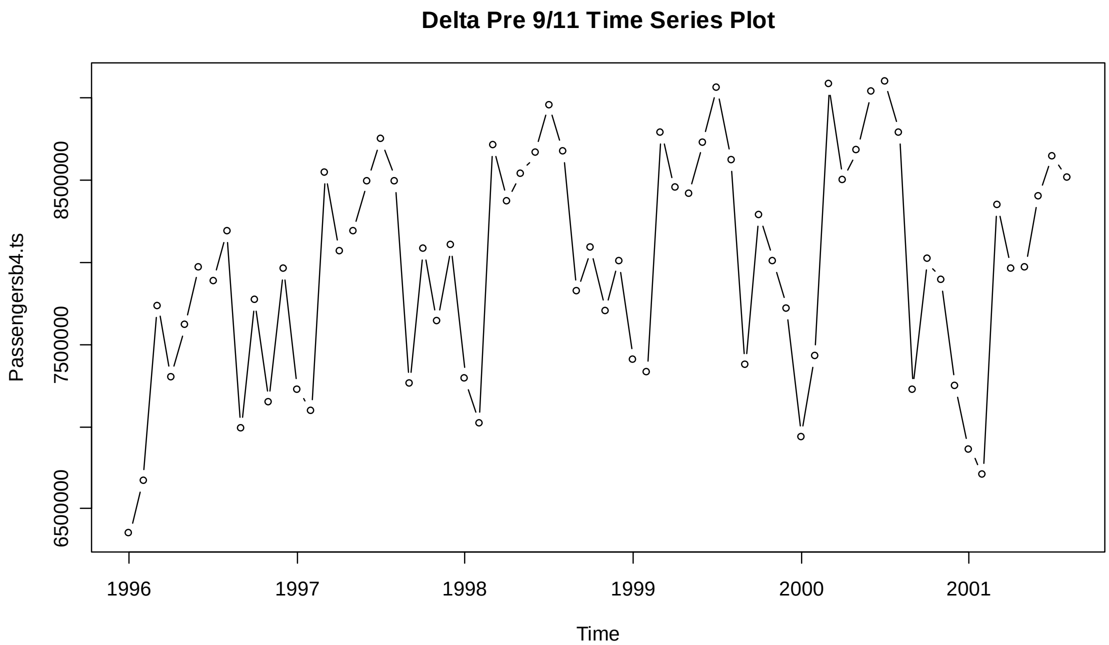

Figure 77: Delta Airlines Pre 9/11 Time Series Plot.

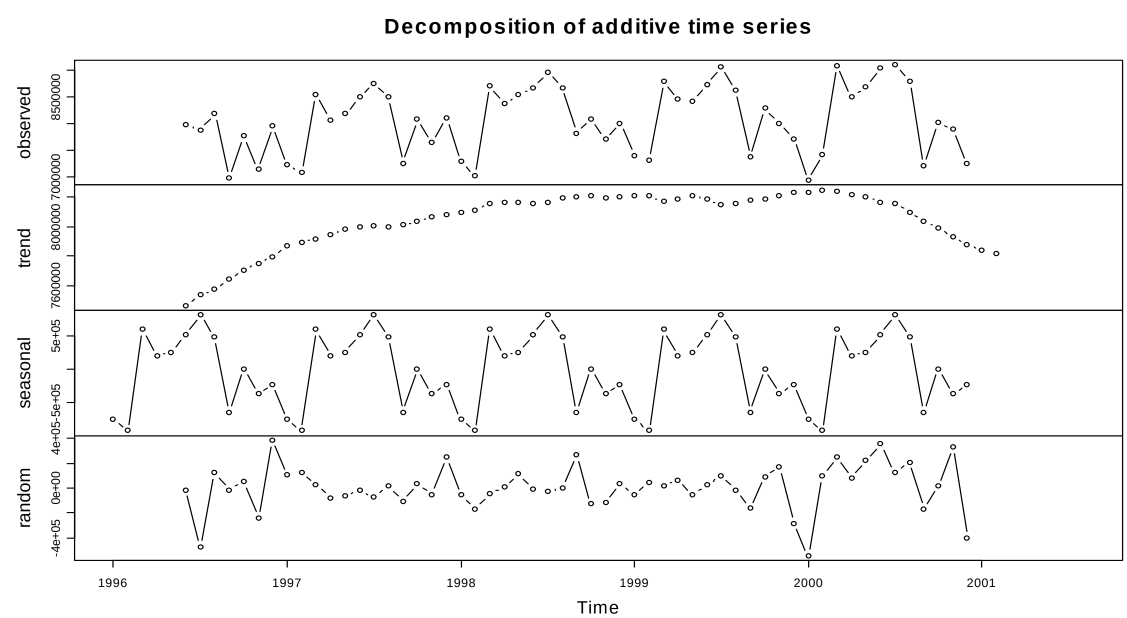

Figure 78: Delta Airlines Pre 9/11 Time Series Decomposition Plot.

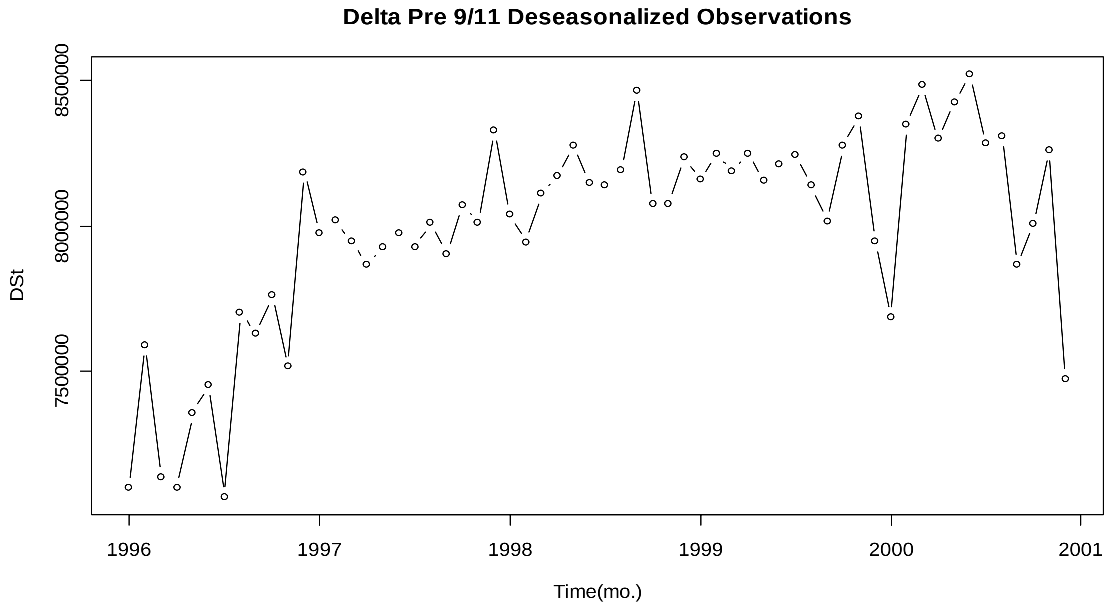

Figure 79: Delta Airlines Pre 9/11 Deseasonalized Time Series Plot.

The first attempt at modeling the Delta Airlines pre 9/11 deseasonalized data was to try and fit it with a linear model. With an *F* statistic of 42.02 and associated p-value 2.159e-08, there is a strong relationship between number of passengers and time (Output 62). The linear regression equation for the Delta Airlines pre 9/11 deseasonalized time series is of the form:

$$\hat{y}_t = 7592920 + 13139t \tag{80}$$

```
Call:
lm(formula = DSt ~ times.ds)

Residuals:
    Min      1Q  Median      3Q     Max
-907210  -44978   78226  158414  438786

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  7592920      71093 106.802  < 2e-16 ***
times.ds       13139       2027   6.482 2.16e-08 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 271900 on 58 degrees of freedom
Multiple R-Squared: 0.4201,     Adjusted R-squared: 0.4101
F-statistic: 42.02 on 1 and 58 DF,  p-value: 2.159e-08
```

Output 62: Summary of Delta Airlines Pre 9/11 Time Series Linear Model.

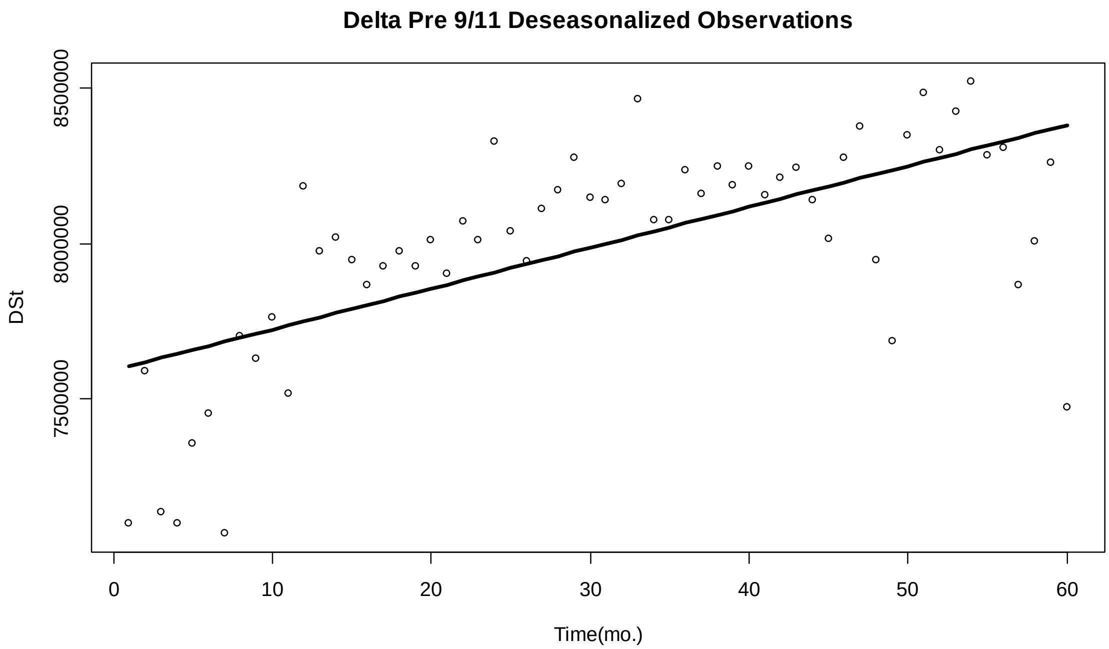

Figure 80: Delta Airlines Pre 9/11 Time Series with Linear Regression Line.

The regression line (Figure 80) does not seem to fit the data accurately.

The Shapiro-Wilk test indicates that there is enough evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(DSt.lm)
W = 0.8799, p-value = 2.622e-05
```

The Breusch-Pagan test does not give sufficient evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  DSt.lm
BP = 0.0084, df = 1, p-value = 0.9271
```

The Durbin-Watson test indicates there is enough evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  DSt.lm
DW = 0.9541, p-value = 2.753e-06
alternative hypothesis:true autocorrelation is greater than 0
```

The second attempt at modeling the Delta Airlines pre 9/11 data was to use a quadratic model. With an *F* statistic of 61.92 and associated p-value 5.119e-15, there is strong evidence of a relationship between number of passengers and time (Output 63). The quadratic regression equation for the Delta Airlines pre 9/11 time series model is of the form:

$$\hat{y}_t = 7168326.42 + 54228.98t - 673.60t^2 \tag{81}$$

```
Call:
lm(formula = DSt ~ times.ds + I(times.ds^2))

Residuals:
    Min      1Q  Median      3Q     Max
-523032 -104707  -11327  134966  463150

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)   7168326.42   81002.31  88.495  < 2e-16 ***
times.ds        54228.98    6127.16   8.851 2.71e-12 ***
I(times.ds^2)    -673.60      97.36  -6.919 4.34e-09 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Residual standard error: 202200 on 57 degrees of freedom
Multiple R-Squared: 0.6848,     Adjusted R-squared: 0.6738
F-statistic: 61.92 on 2 and 57 DF,  p-value: 5.119e-15
```

Output 63: Summary of Delta Airlines Pre 9/11 Time Series Quadratic Model.

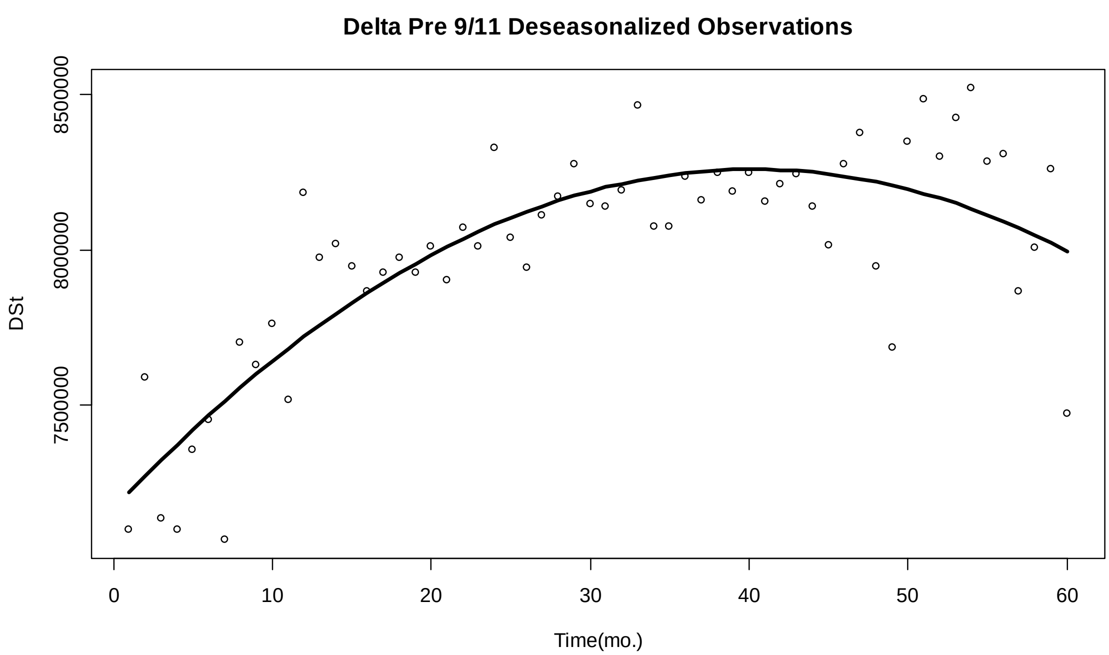

Figure 81: Delta Airlines Pre 9/11 Time Series with Quadratic Regression Line.

The quadratic regression line plotted with the time series data (Figure 81) seems a good fit to the observations.

The Shapiro-Wilk test does not give enough evidence to reject the normality assumption.

```
    Shapiro-Wilk normality test
data:  resid(DStquad.lm)
W = 0.974, p-value = 0.2277
```

The Breusch-Pagan test gives sufficient evidence to reject the constant variance assumption.

```
    studentized Breusch-Pagan test
data:  DStquad.lm
BP = 11.1629, df = 2, p-value = 0.003767
```

The Durbin-Watson test gives sufficient evidence to reject the independence assumption.

```
   Durbin-Watson test
data:  DStquad.lm
DW = 1.7196, p-value = 0.0846
alternative hypothesis: true autocorrelation is greater than 0
```

The final attempt made at modeling the Delta Airlines pre 9/11 data was to use a cubic model. With an F statistic of 41.54 and associated p-value 2.917e-14, there is a strong relationship between number of passengers and time (Output 64). The cubic regression equation for the Delta Airlines pre 9/11 time series data is of the form:

$$\hat{y}_t = 7095000 + 68120t - 1238t^2 + 6.170t^3 \tag{82}$$

```
Call:
lm(formula = DSt ~ times.ds + I(times.ds^2) + I(times.ds^3))

Residuals:
    Min      1Q  Median      3Q     Max 
-583207  -77231   -3552  140580  440618 

Coefficients:
                Estimate Std. Error t value Pr(>|t|)    
(Intercept)    7.095e+06  1.114e+05  63.701  < 2e-16 ***
times.ds       6.812e+04  1.568e+04   4.344 5.95e-05 ***
I(times.ds^2) -1.238e+03  5.947e+02  -2.082   0.0419 *  
I(times.ds^3)  6.170e+00  6.412e+00   0.962   0.3400    
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1 

Residual standard error: 202300 on 56 degrees of freedom
Multiple R-Squared: 0.6899,     Adjusted R-squared: 0.6733 
F-statistic: 41.54 on 3 and 56 DF,  p-value: 2.917e-14
```
Output 64: Summary of Delta Airlines Pre 9/11 Time Series Cubic Model.

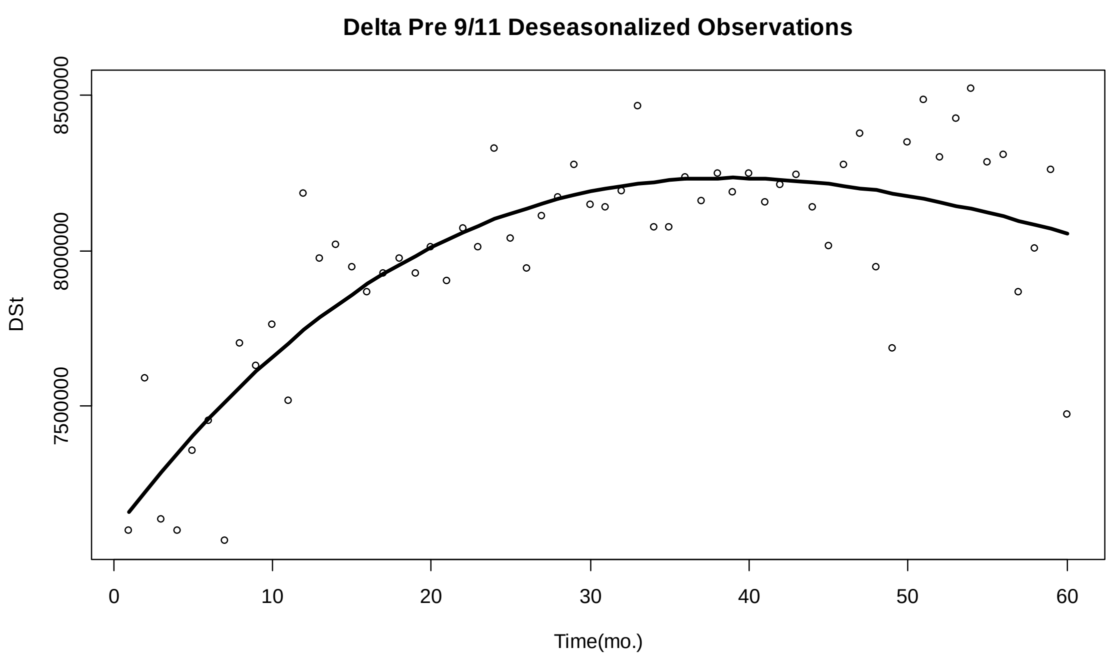

Figure 82: Delta Airlines Pre 9/11 Time Series with Cubic Regression Line.

The cubic regression line plotted with the data (Figure 82) does not seem to be very different than the quadratic model (Figure 81).

The Shapiro-Wilk test does not give sufficient evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(DStcubic.lm)
W = 0.9657, p-value = 0.08942
```

The Breusch-Pagan test gives sufficient evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  DStcubic.lm
BP = 13.1207, df = 3, p-value = 0.004383
```

The Durbin-Watson test gives sufficient evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  DStcubic.lm
DW = 1.7534, p-value = 0.0828
alternative hypothesis: true autocorrelation is greater than 0
```

To decide which model was the best fit, a comparison of the AIC statistic of each of the models was made. Output 65 shows that the AIC for the quadratic model is the smallest and thus the best fit.

```
AIC(DSt.lm)
[1] 1675.825
AIC( DStquad.lm)
[1] 1641.245
AIC(DStcubic.lm)
[1] 1642.261
```

Output 65: AIC Comparison of Delta Airlines Pre 9/11 Time Series Linear, Quadratic, and Cubic Models.

An ANOVA comparison of the models was used to confirm these findings. Output 66 shows that the t^2 parameter of the quadratic equation is significant when compared to the linear equation. Also, the t^3 parameter of the cubic model is of little significance when compared to the quadratic model. Therefore the previous findings are confirmed and any further analysis would be done with the quadratic model for the Delta Airlines pre 9/11 data.

```
Analysis of Variance Table

Model 1: DSt ~ times.ds
Model 2: DSt ~ times.ds + I(times.ds^2)
  Res.Df        RSS Df  Sum of Sq     F    Pr(>F)
1     58 4.2882e+12
2     57 2.3308e+12  1 1.9574e+12 47.87 4.344e-09 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

Analysis of Variance Table

Model 1: DSt ~ times.ds + I(times.ds^2)
Model 2: DSt ~ times.ds + I(times.ds^2) + I(times.ds^3)
  Res.Df        RSS Df  Sum of Sq     F Pr(>F)
1     57 2.3308e+12
2     56 2.2929e+12  1 3.7913e+10 0.926 0.3400
```

Output 66: ANOVA Comparison of Delta Airlines Pre 9/11 Time Series Linear, Quadratic, and Cubic Models.

The Delta Airlines post 9/11 data were then turned into a time series and plotted (Figure 83).

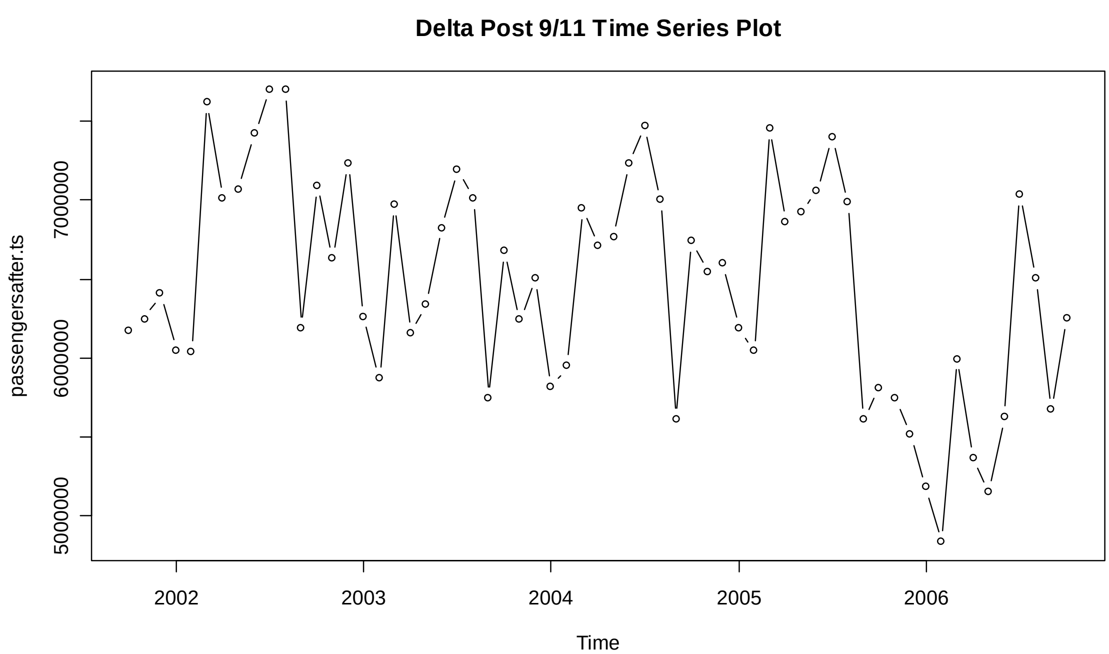

Figure 83: Delta Airlines Post 9/11 Time Series Plot.

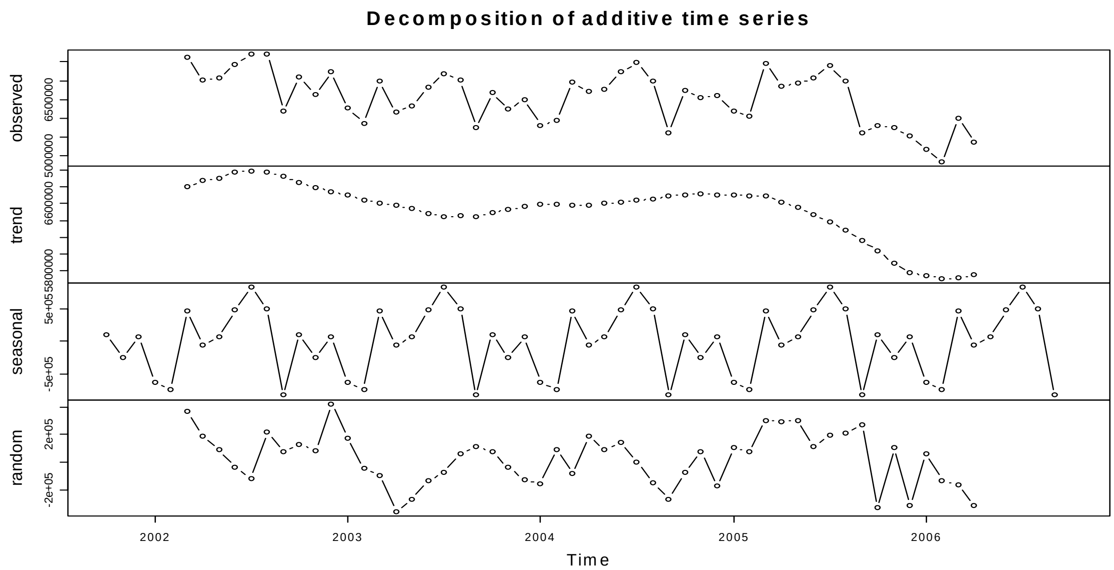

Figure 84: Delta Airlines Post 9/11 Time Series Decomposition Plot.

The time series was then decomposed into its trend, random, and seasonal components and plotted (Figure 84). The seasonal component was removed and the deseasonalized time series was plotted (Figure 85).

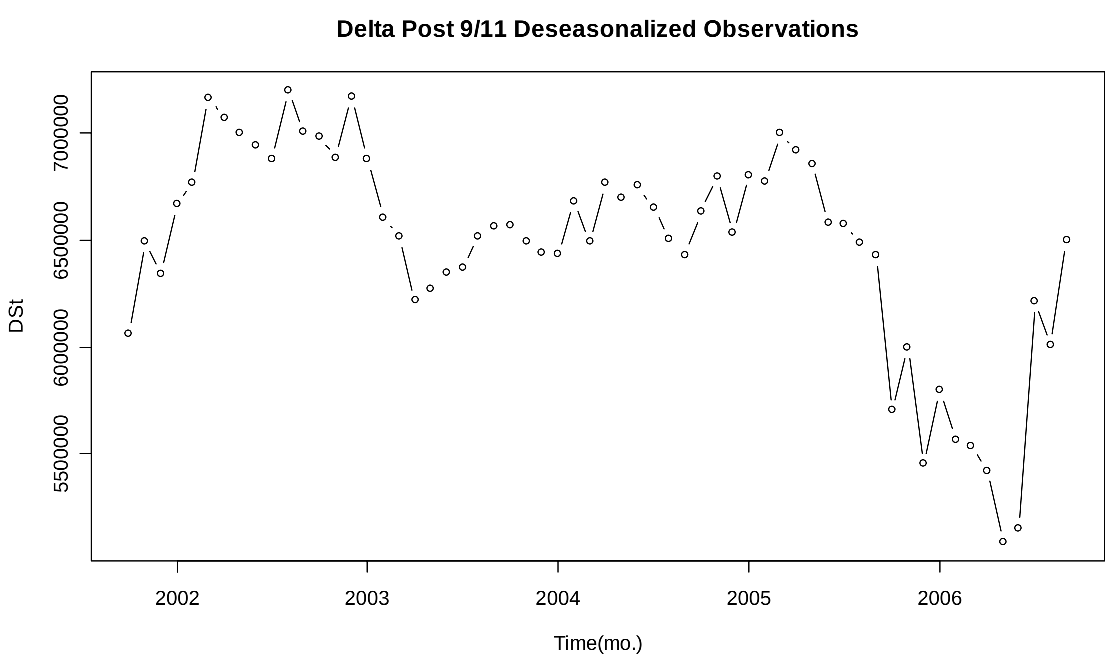

Figure 85: Delta Airlines Post 9/11 Time Series Deseasonalized Plot.

The time series was then fit with a linear model. With an *F* statistic of 30.77 and associated p-value 7.502e-07 there is a strong relationship between passenger numbers and time (Output 67). The linear regression equation for the Delta Airlines post 9/11 time series model is of the form:

$$\hat{y}_t = 8125729 - 16561t \qquad (83)$$

```
Call:
lm(formula = DSt ~ times.ds)

Residuals:
    Min      1Q  Median      3Q     Max
-969869 -248234   82508  266369  713006

Coefficients:
            Estimate Std. Error t value Pr(>|t|)
(Intercept)  8125729     301522  26.949  < 2e-16 ***
times.ds      -16561       2986  -5.547  7.5e-07 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 400500 on 58 degrees of freedom
Multiple R-Squared: 0.3466,     Adjusted R-squared: 0.3354
F-statistic: 30.77 on 1 and 58 DF,  p-value: 7.502e-07
```

Output 67: Summary of Delta Airlines Post 9/11 Time Series Linear Model.

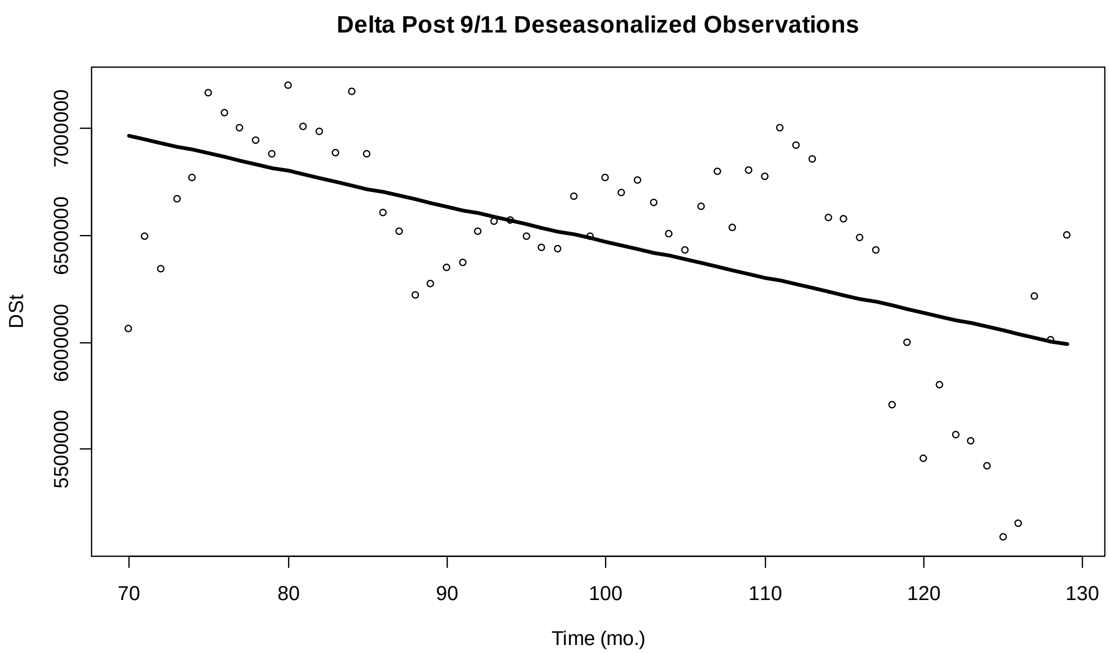

Figure 86: Delta Airlines Post 9/11 Time Series with Linear Regression Line.

The linear regression line does not seem a good fit to the observations.

The Shapiro-Wilk test does not give sufficient evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(DSt.lm)
W = 0.961, p-value = 0.05244
```

The Breusch-Pagan test does indicate that there is enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  DSt.lm
BP = 5.3058, df = 1, p-value = 0.02125
```

The Durbin-Watson test indicates that there is enough evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  DSt.lm
DW = 0.4593, p-value = 2.187e-14
alternative hypothesis: true autocorrelation is greater than 0
```

A second attempt at modeling the Delta Airlines post 9/11 time series was made using a quadratic model. With an *F* statistic of 24.62 and corresponding p-value 1.961e-08, there is a strong relationship between number of passengers and time (Output 68). The form of the quadratic regression equation for Delta Airlines post 9/11 time series is:

$$\hat{y}_t = 2163225 + 107032.4t - 621.1t^2 \tag{84}$$

```
Call:
lm(formula = DSt ~ times.ds + I(times.ds^2))

Residuals:
    Min      1Q  Median      3Q     Max
-752287 -259074   52628  281679  861832

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)   2163225.0  1714369.4   1.262 0.212153
times.ds       107032.4    35180.0   3.042 0.003545 **
I(times.ds^2)    -621.1      176.3  -3.524 0.000846 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 366100 on 57 degrees of freedom
Multiple R-Squared: 0.4635,     Adjusted R-squared: 0.4447
F-statistic: 24.62 on 2 and 57 DF,  p-value: 1.961e-08
```
Output 68: Summary of Delta Airlines Post 9/11 Time Series Quadratic Model.

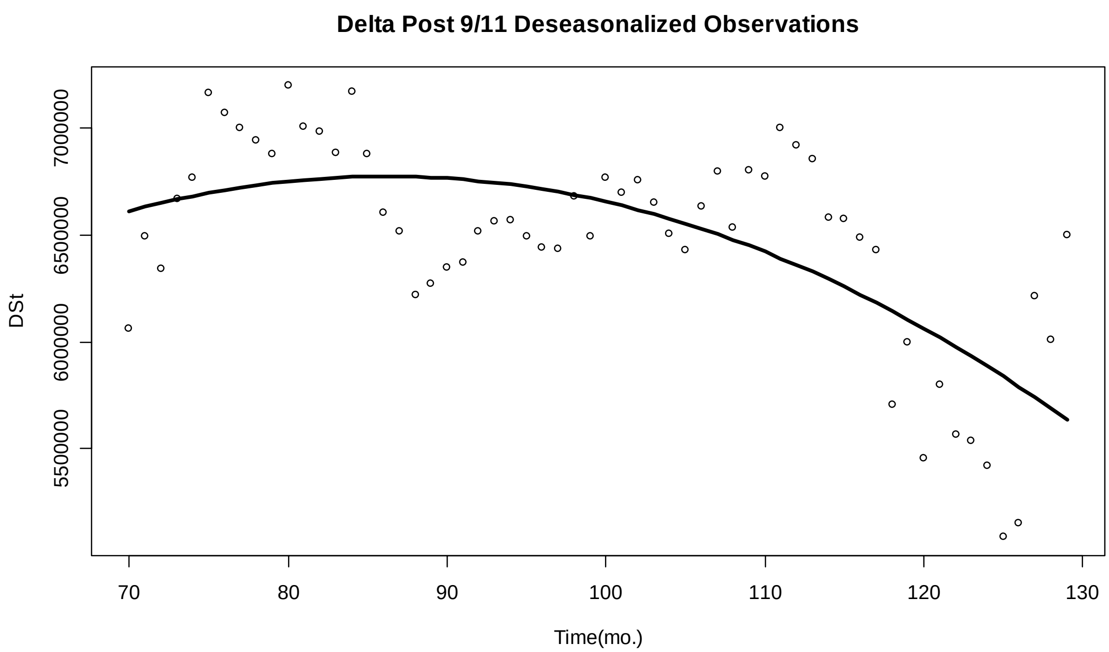

Figure 87: Delta Airlines Post 9/11 Time Series with Quadratic Regression Line.

The plot of the time series and quadratic line proves to be a good fit (Figure 87).

The Shapiro-Wilk test does not give enough evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(DStquad.lm)
W = 0.9843, p-value = 0.6359
```

The Breusch-Pagan test gives sufficient evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  DStquad.lm
BP = 19.2193, df = 2, p-value = 6.708e-05
```

The Durbin-Watson test gives enough evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  DStquad.lm
DW = 0.5589, p-value = 1.385e-12
alternative hypothesis: true autocorrelation is greater than 0
```

A final attempt made at modeling the Delta Airlines post 9/11 time series was made using a cubic model. With an *F* statistic of 16.24 and associated p-value 1.032e-07 there is a relationship between number of passengers and time (Output 69). The regression equation for the Delta Airlines post 9/11 time series cubic model is of the form:

$$\hat{y}_t = 6716000 - 35500t + 837.9t^2 - 4.888t^3 \tag{85}$$

```
Call:
lm(formula = DSt ~ times.ds + I(times.ds^2) + I(times.ds^3))

Residuals:
    Min      1Q  Median      3Q     Max 
-738502 -239876   38407  267310  909500 

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    6.716e+06  1.102e+07   0.609    0.545
times.ds      -3.550e+04  3.426e+05  -0.104    0.918
I(times.ds^2)  8.379e+02  3.493e+03   0.240    0.811
I(times.ds^3) -4.888e+00  1.169e+01  -0.418    0.677

Residual standard error: 368800 on 56 degrees of freedom
Multiple R-Squared: 0.4652,     Adjusted R-squared: 0.4365 
F-statistic: 16.24 on 3 and 56 DF,  p-value: 1.032e-07
```

Output 69: Summary of Delta Airlines Post 9/11 Time Series Cubic Model.

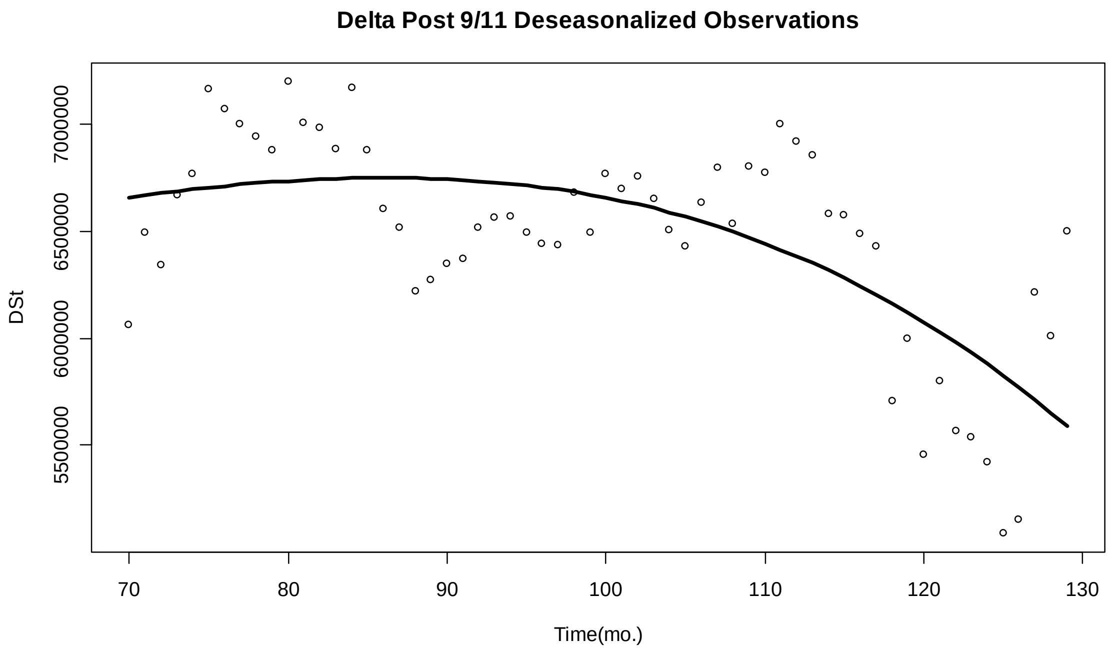

Figure 88: Delta Airlines Post 9/11 Time Series with Cubic Regression Line.

Figure 88 shows the cubic regression line with the time series data. The fit is good, but similar to the quadratic model (Figure 87).

The Shapiro-Wilk test does not indicate that there is enough evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(DStcubic.lm)
W = 0.9835, p-value = 0.5951
```

The Breusch-Pagan test indicates that there is enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  DStcubic.lm
BP = 23.605, df = 3, p-value = 3.02e-05
```

The Durbin-Watson test indicates that there is enough evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  DStcubic.lm
DW = 0.5647, p-value = 5.084e-13
alternative hypothesis: true autocorrelation is greater than 0
```

In order to decide which of the Delta Airlines post 9/11 time series models was the best fit the AIC comparison was employed. Output 71 reveals that the AIC for the quadratic model is smallest, thus the best fit.

```
AIC (DSt.lm)
[1] 1722.292
AIC(DStquad.lm)
[1] 1712.467
AIC(DStcubic.lm)
[1] 1714.280
```

Output 70: AIC Comparison of Delta Airlines Post 9/11 Time Series Linear, Quadratic, and Cubic Models.

An ANOVA comparison was made to confirm these findings. In Output 71 is seen that the t^2 term of the quadratic model is significant when compared to the linear mode. Also, the t^3 term of the cubic model is of no significance when compared to the quadratic model. Thus, the quadratic model is the best fit for the post 9/11 Delta Airlines time series and this model would be used for further comparisons.

```
Analysis of Variance Table
Model 1: DSt ~ times.ds
Model 2: DSt ~ times.ds + I(times.ds^2)
  Res.Df        RSS Df  Sum of Sq      F    Pr(>F)
1     58 9.3027e+12
2     57 7.6387e+12  1 1.6640e+12 12.417 0.0008461 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

 Analysis of Variance Table
Model 1: DSt ~ times.ds + I(times.ds^2)
Model 2: DSt ~ times.ds + I(times.ds^2) + I(times.ds^3)
  Res.Df        RSS Df  Sum of Sq     F Pr(>F)
1     57 7.6387e+12
2     56 7.6149e+12  1 2.3791e+10 0.175 0.6773
```

Output 71: ANOVA Comparison of Delta Airlines Post 9/11 Time Series Linear, Quadratic, and Cubic Models.

The post 9/11 models were then checked for possible outliers using the residual plots and Cook's D test as in previous analysis. For the Delta Airlines post 9/11 linear model, (Figure 89) observation 1 is the only observation that meets such criteria. According to the leverage plot (Figure 89) and Cook's D test (Output 72) there do not exist any influential outliers.

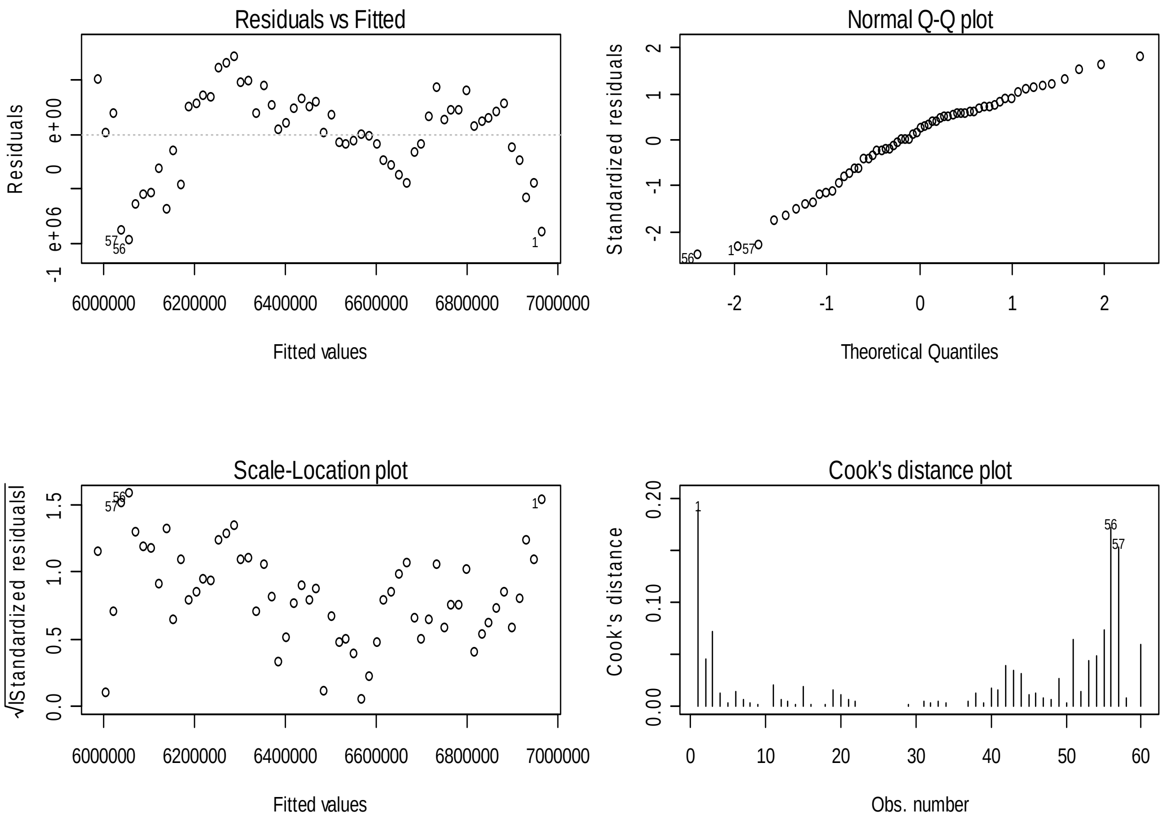

Figure 89: Delta Airlines Post 9/11 Time Series Linear Model Residual Plots.

```
cooksD=cooks.distance(DSt.lm)
f0.50=qf(0.5,df1=5,df2=24)
cooksD[which(cooksD>f0.50)]
named numeric(0)
```

Output 72: Cook's D Statistic on Delta Airlines Post 9/11 Time Series Linear Model.

For the quadratic model (Figures 90) there are no observations that meet the criteria, and according to Cook's D (Output 73) and leverage plot (90) there are no influential outliers.

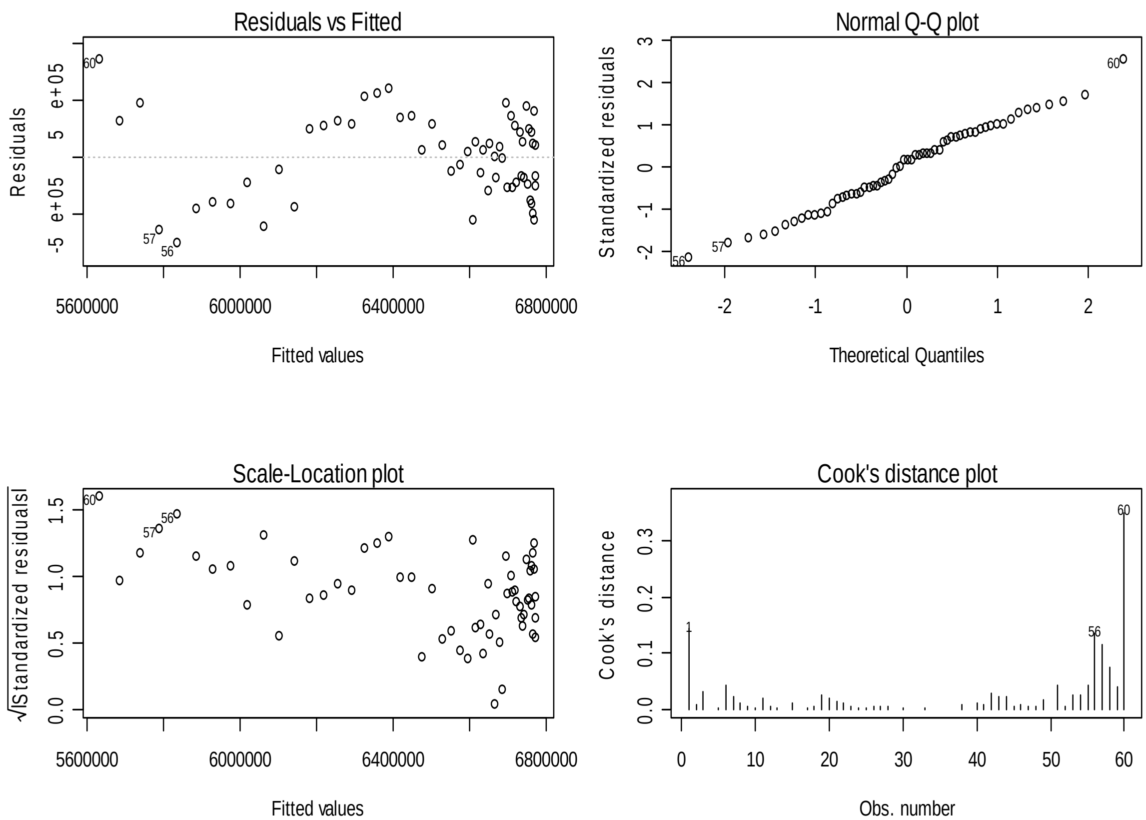

Figure 90: Delta Airlines Post 9/11 Time Series Quadratic Model Residual Plots.

```
cooksD=cooks.distance(DStquad.lm)
f0.50=qf(0.5, df1=5, df2=24)
cooksD[which(cooksD>f0.50)]
named numeric(0)
```

Output 73: Cook's D Statistic on Delta Airlines Post 9/11 Time Series Quadratic Model.

For the cubic model (Figures 91) observation 1 is the only one singled out that fits the criteria, and Cook's D (Output 74) and the leverage plot (91) do not indicate any influential outliers.

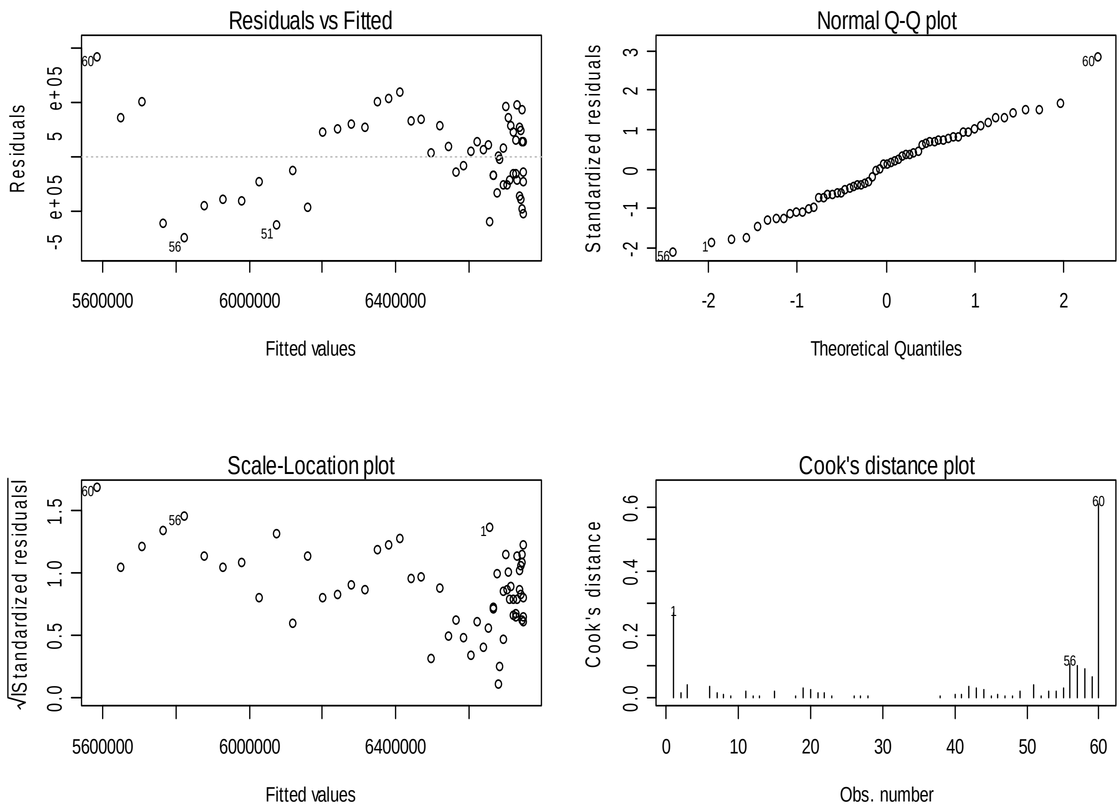

Figure 91: Delta Airlines Post 9/11 Time Series Cubic Model Residual Plots.

```
cooksD=cooks.distance(DStcubic.lm)
f0.50=qf(0.5, df1=5, df2=24)
cooksD[which(cooksD>f0.50)]
named numeric(0)
```

Output 74: Cook's D Statistic on Delta Airlines Post 9/11 Time Series Cubic Model.

It was decided that observation one would be the only one removed as an initial shock recovery period and the analysis would be repeated on the data excluding this observation. The data were made into a time series and plotted (Figure 92).

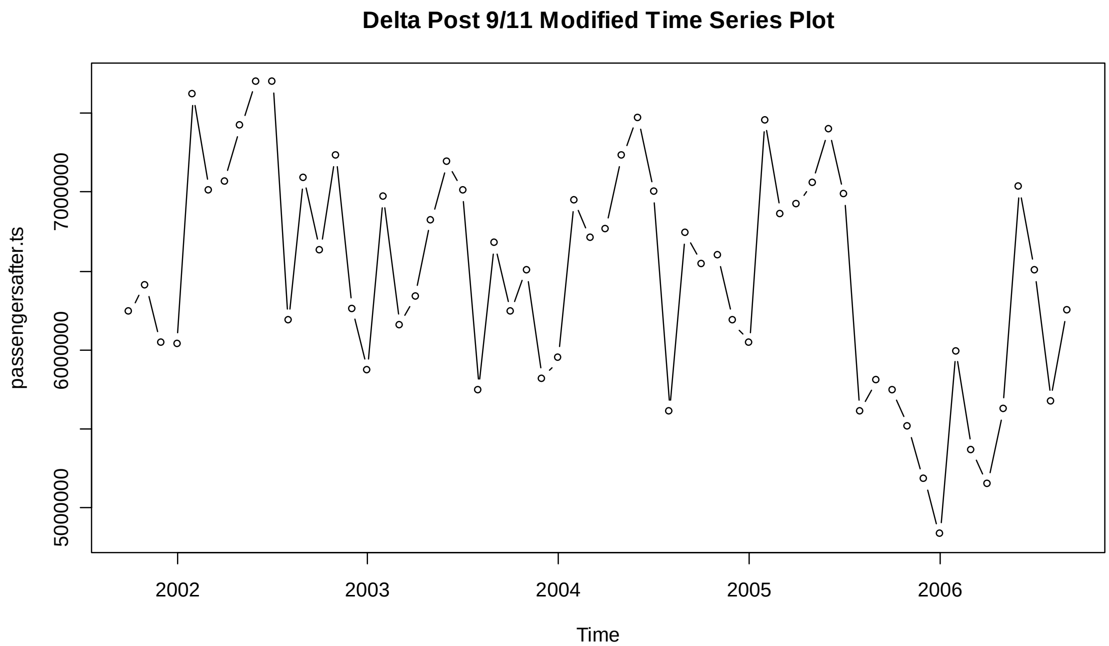

Figure 92: Delta Airlines Post 9/11 Modified Time Series Plot.

The time series was decomposed into its trend, seasonal, and random components (Figure 93).

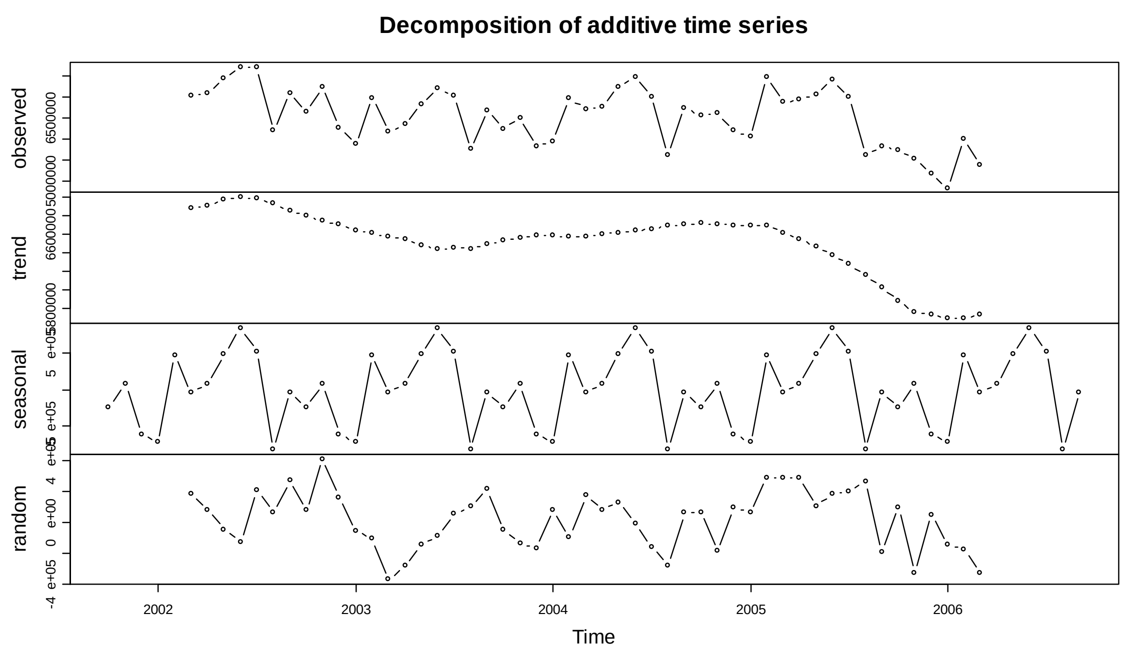

Figure 93: Delta Airlines Post 9/11 Modified Time Series Decomposition Plot.

The seasonal component was then removed and the time series once again was plotted (Figure 94).

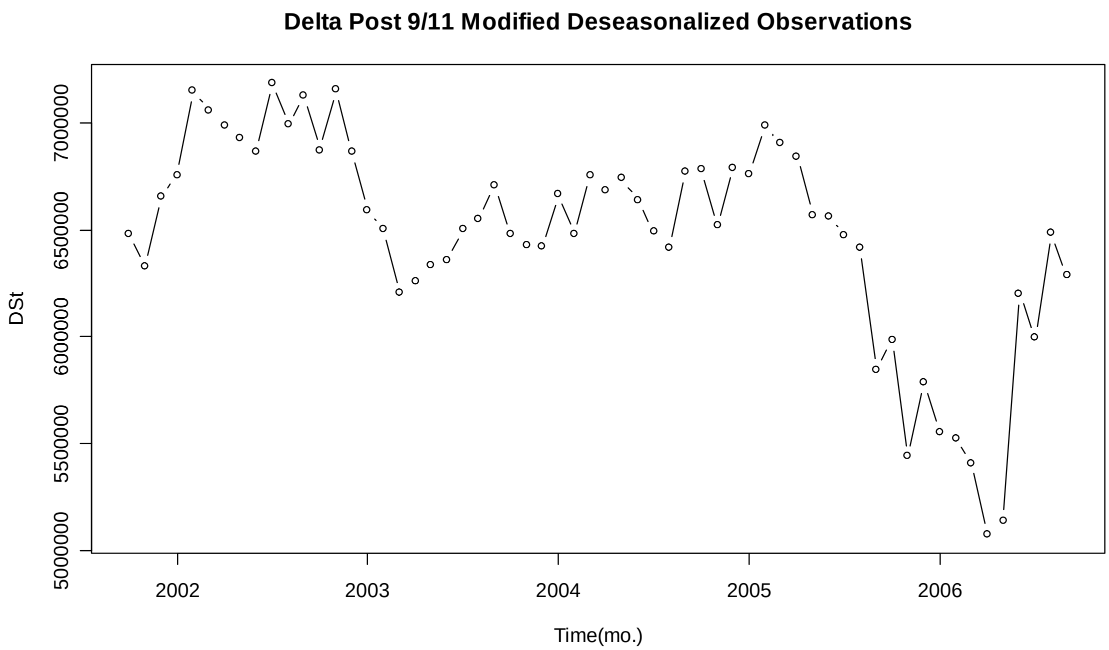

Figure 94: Delta Airlines Post 9/11 Modified Time Series Deseasonalized Plot.

The first model that was attempted on the modified time series was a linear model. With an *F* statistic of 37.44 and associated p-value of 8.67e-08 the summary (Output 75) shows that there is a strong relationship between number of passengers and time. The linear regression equation for Delta Airlines post 9/11 modified time series is of the form:

$$\hat{y}_t = 8228184 - 17576t \tag{86}$$

```
Call:
lm(formula = DSt ~ times.ds)

Residuals:
    Min      1Q  Median      3Q     Max 
-975732 -230270   79892  274277  692927 

Coefficients:
            Estimate Std. Error t value Pr(>|t|)    
(Intercept)  8228184     290104  28.363  < 2e-16 ***
times.ds      -17576       2872  -6.119 8.67e-08 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1 

Residual standard error: 385300 on 58 degrees of freedom
Multiple R-Squared: 0.3923,     Adjusted R-squared: 0.3818 
F-statistic: 37.44 on 1 and 58 DF,  p-value: 8.668e-08
```

Output 75: Summary of Delta Airlines Post 9/11 Modified Time Series Linear Model.

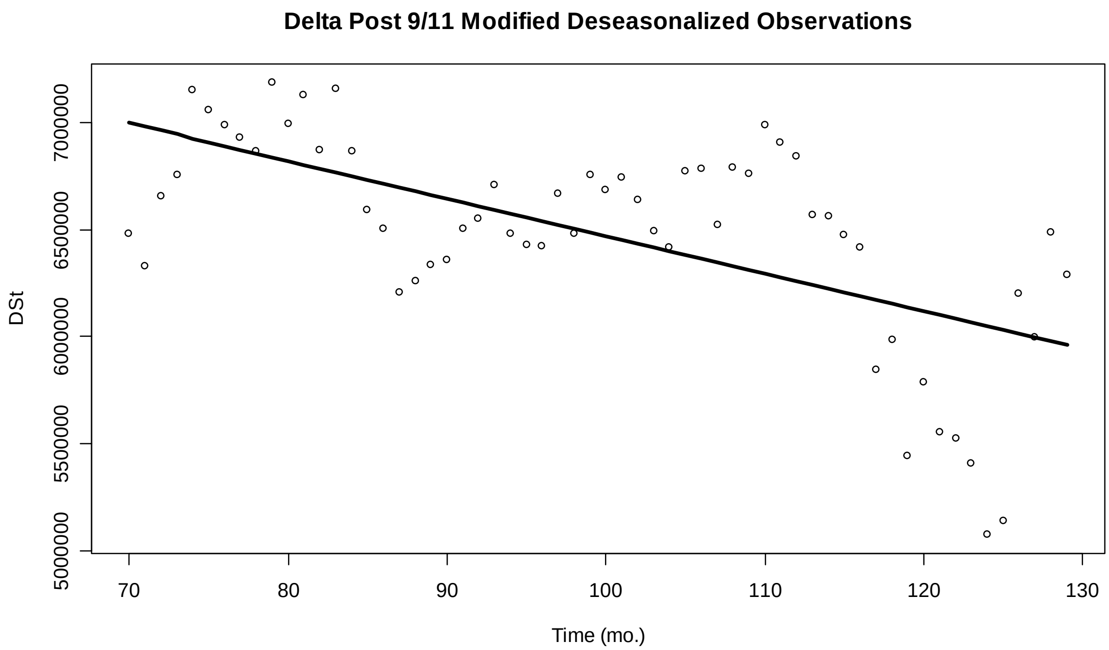

Figure 95: Delta Airlines Post 9/11 Modified Time Series with Linear Regression Line.

Figure 95 reveals that the linear regression equation is decreasing corresponding to the data, but there could be a better fit.

The Shapiro-Wilk test does not indicate that there is enough evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(DSt.lm) 
W = 0.9685, p-value = 0.1228
```

The Breusch-Pagan test reveals that there is sufficient evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  DSt.lm
BP = 9.038, df = 1, p-value = 0.002644
```

The Durbin-Watson test indicates that there is sufficient evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  DSt.lm
DW = 0.4692, p-value = 3.949e-14
alternative hypothesis: true autocorrelation is greater than 0
```

A second attempt made at modeling the post 9/11 modified data was made with a quadratic model. The summary (Output 76) shows that with an *F* statistic of 24.03 and corresponding p-value 2.703e-08 there is a strong relationship between number of passengers and time. The quadratic regression equation for the Delta Airlines post 9/11 modified data is of the form:

$$\hat{y}_t = 3787412 + 74473.9t - 462.6t^2 \quad (87)$$

```
Call:
lm(formula = DSt ~ times.ds + I(times.ds^2))

Residuals:
    Min      1Q  Median      3Q     Max
-836809 -255471   62719  275999  742799

Coefficients:
               Estimate Std. Error t value Pr(>|t|)
(Intercept)   3787412.0  1719948.4   2.202   0.0317 *
times.ds        74473.9    35294.5   2.110   0.0393 *
I(times.ds^2)    -462.6      176.8  -2.616   0.0114 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 367300 on 57 degrees of freedom
Multiple R-Squared: 0.4574,     Adjusted R-squared: 0.4384
F-statistic: 24.03 on 2 and 57 DF,  p-value: 2.703e-08
```

Output 76: Summary of Delta Airlines Post 9/11 Modified Time Series Quadratic Model.

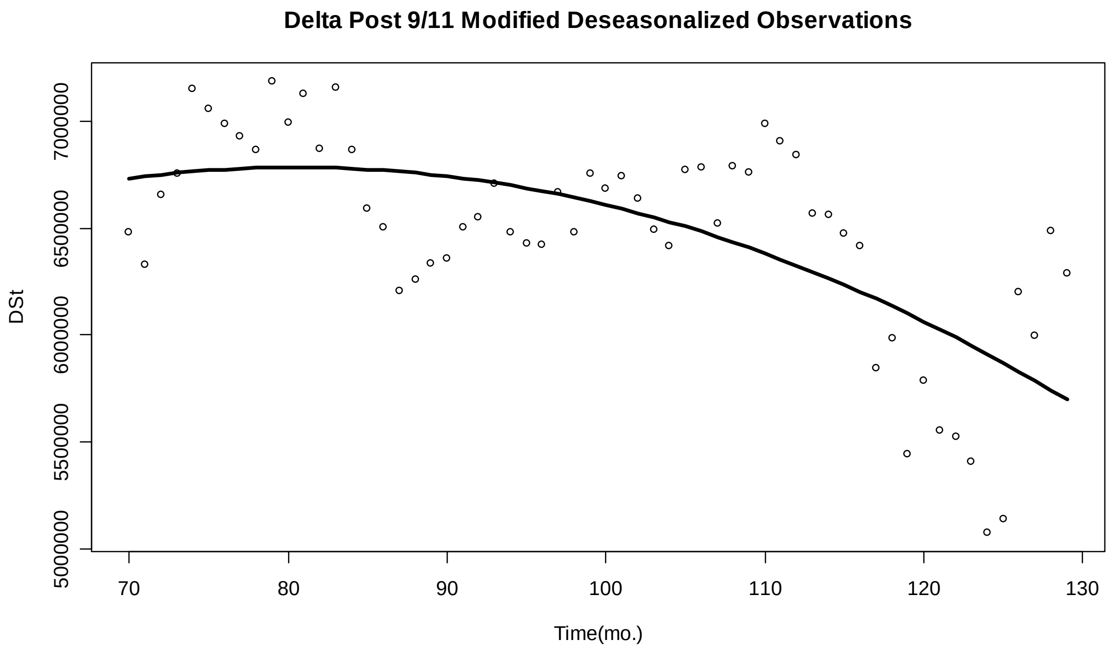

Figure 96: Delta Airlines Post 9/11 Modified Time Series with Quadratic Regression Line.

The quadratic regression line with the time series modified data (Figure 96) reveals a better fit than the linear model (Figure 95).

The Shapiro-Wilk test does not give enough evidence to reject the normality assumption.

```
  Shapiro-Wilk normality test
data:  resid(DStquad.lm)
W = 0.9835, p-value = 0.5912
```

The Breusch-Pagan test does give enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  DStquad.lm
BP = 20.5023, df = 2, p-value = 3.532e-05
```

The Durbin-Watson test indicates that there is sufficient evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  DStquad.lm
DW = 0.5261, p-value = 2.568e-13
alternative hypothesis: true autocorrelation is greater than 0
```

A final attempt at modeling the modified post 9/11 Delta Airlines data was made using a cubic model. The summary (Output 77) shows that with an *F* statistic of 15.83 and associated p-value 1.428e-07 there is a strong relationship between number of passengers and time. The cubic regression equation is of the form:

$$\hat{y}_t = 7627000 - 55120t + 864t^2 - 4.444t^3 \quad (88)$$

```
Call:
lm(formula = DSt ~ times.ds + I(times.ds^2) + I(times.ds^3))

Residuals:
    Min      1Q  Median      3Q     Max
-830211 -246721   51699  257626  777326

Coefficients:
                Estimate Std. Error t value Pr(>|t|)
(Intercept)    7.927e+06  1.106e+07   0.717    0.476
times.ds      -5.512e+04  3.438e+05  -0.160    0.873
I(times.ds^2)  8.640e+02  3.505e+03   0.247    0.806
I(times.ds^3) -4.444e+00  1.173e+01  -0.379    0.706

Residual standard error: 370100 on 56 degrees of freedom
Multiple R-Squared: 0.4588,     Adjusted R-squared: 0.4298
F-statistic: 15.83 on 3 and 56 DF,  p-value: 1.428e-07
```

Output 77: Summary of Delta Airlines Post 9/11 Modified Time Series Cubic Model.

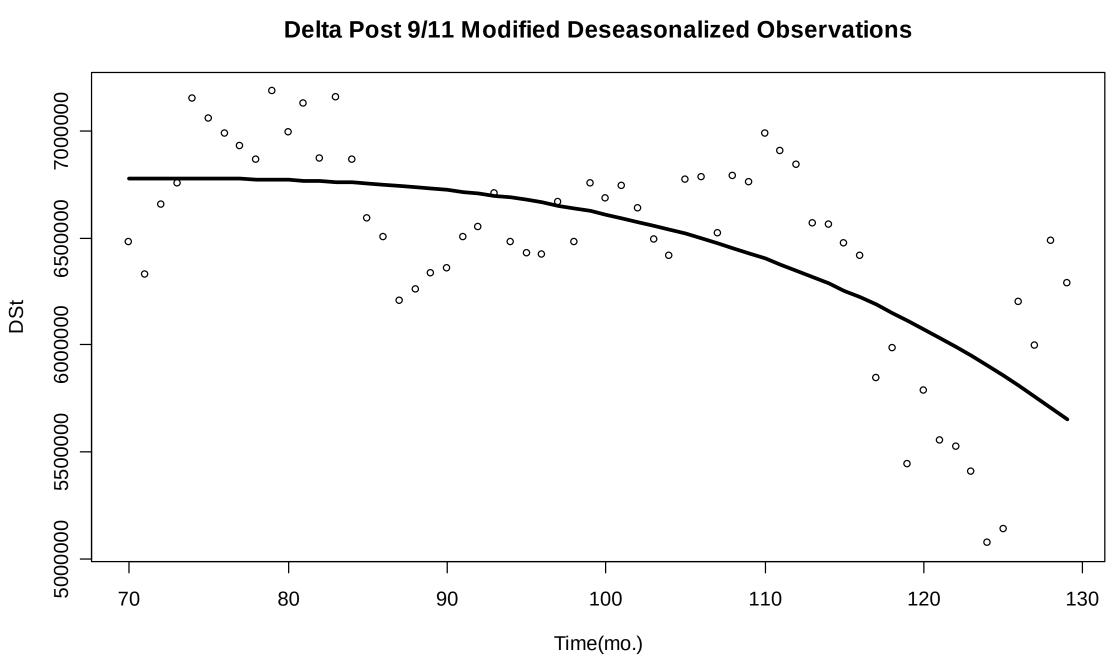

Figure 97: Delta Airlines Post 9/11 Modified Time Series with Cubic Regression Line.

The cubic regression curve plotted with the modified time series (Figure 97) does not seem much different from the quadratic one (Figure 96).

The Shapiro-Wilk test does not indicate that there is sufficient evidence to reject the normality assumption.

```
Shapiro-Wilk normality test
data:  resid(DStcubic.lm)
W = 0.986, p-value = 0.7218
```

The Breusch-Pagan test indicates that there is enough evidence to reject the constant variance assumption.

```
  studentized Breusch-Pagan test
data:  DStcubic.lm
BP = 24.9561, df = 3, p-value = 1.577e-05
```

The Durbin-Watson test indicates that there is enough evidence to reject the independence assumption.

```
  Durbin-Watson test
data:  DStcubic.lm
DW = 0.5296, p-value = 8.013e-14
alternative hypothesis: true autocorrelation is greater than 0
```

To decide which model best fit the Delta Airlines post 9/11 modified data time series the AIC of each model was relied upon. According to Output 78, The AIC for the quadratic model is the smallest, and thus the best fit.

```
AIC (DSt.lm)
[1] 1717.659
AIC(DStquad.lm)
[1] 1712.857
AIC(DStcubic.lm)
[1] 1714.703
```

Output 78: AIC Comparison of Delta Airlines Post 9/11 Modified Time Series Linear, Quadratic, and Cubic Models.

An ANOVA comparison was made for the three models to verify these findings. According to Output 79, the t^2 term of the quadratic model is significant when compared to the linear model, also, the t^3 parameter of the cubic model is of no significance when compared to the quadratic model. Thus the quadratic model is the best fit and all comparisons would be made with it.

```
Analysis of Variance Table
Model 1: DSt ~ times.ds
Model 2: DSt ~ times.ds + I(times.ds^2)
  Res.Df        RSS Df  Sum of Sq      F  Pr(>F)
1     58 8.6115e+12
2     57 7.6885e+12  1 9.2305e+11 6.8432 0.01137 *
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Analysis of Variance Table
Model 1: DSt ~ times.ds + I(times.ds^2)
Model 2: DSt ~ times.ds + I(times.ds^2) + I(times.ds^3)
  Res.Df        RSS Df  Sum of Sq      F Pr(>F)
1     57 7.6885e+12
2     56 7.6688e+12  1 1.9670e+10 0.1436 0.7061
```

Output 79: ANOVA Comparison of Delta Airlines Post 9/11 Modified Time Series Linear, Quadratic, and Cubic Models.

The two quadratic models for the Delta Airlines post 9/11 time series were compared. The model that included all observations had an AIC of 1712.467 and the AIC for the modified data was 1712.857. Although very slight, the AIC for the model including all observations is smaller, thus the better of the two and would be used in further comparisons.

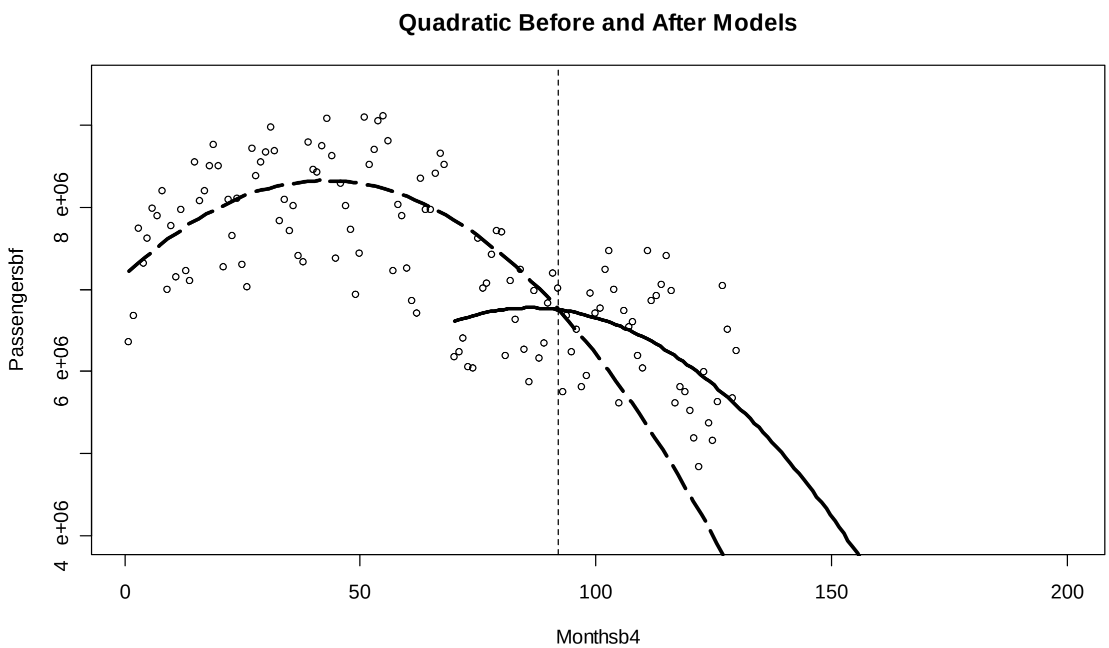

Figure 98: Delta Airlines Pre 9/11 and Post 9/11 Original Time Series with Quadratic Regression Lines.

The pre 9/11 Delta Airlines quadratic model and the post 9/11 quadratic model were set equal and solved for time. The time the two models are equal is August 2003. Figure 98 shows the pre 9/11 quadratic regression line stretched into the future, representing what could have happened if 9/11 never took place, along with the post 9/11 quadratic model and the crossing point. This point can be interpreted as the time of recover from the effects of the incident.

According to the additive decomposition method all three airlines have recovered. Continental Airlines recovered in October 2005, American Airlines more recently recovered in August of 2007, and Delta Airlines in August of 2003. Looking at the models in Figures 54, 76, and 98, it is seen that the three airlines have acted in extremely different manners. Continental, which seemed to be on a decline, turned around after September 2001 and is currently on an incline. American Airlines, which was on an incline, dipped and then recovered with an even greater slope. Delta Airlines started out on a decline continued with its decline at almost exactly the same rate.

4.5 Box-Cox Transformations

In this section is demonstrated the Box-Cox transformation models for each of the deseasonalized linear models.

4.5.1 Continental Airlines

A Box-Cox transformation was performed on the response variable for each of Continental Airlines deseasonalized linear models. For the pre 9/11 data Figure 99 suggests $(\hat{y}_t^4 - 1)/4 = 2664366.3 + 7605.9t$ would be a more appropriate fit and for the post 9/11 modified data Figure 100 shows $(\hat{y}_t^{-2} - 1)/-2 = 2179313.8 + 5095.8t$ should be utilized.

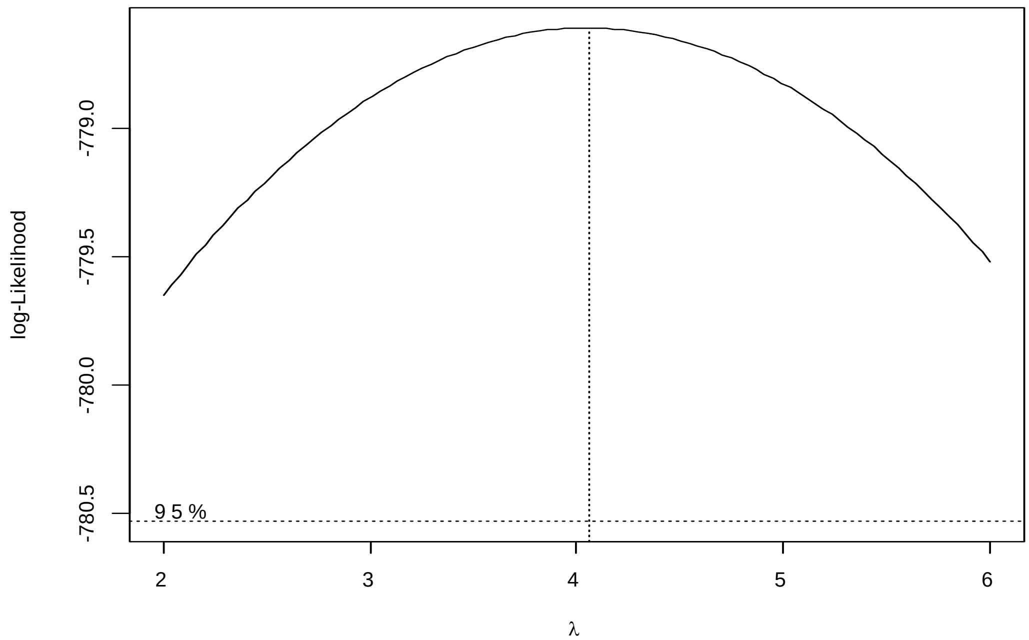

Figure 99: Continental Airlines Pre 9/11 Box-Cox Transformation Plot

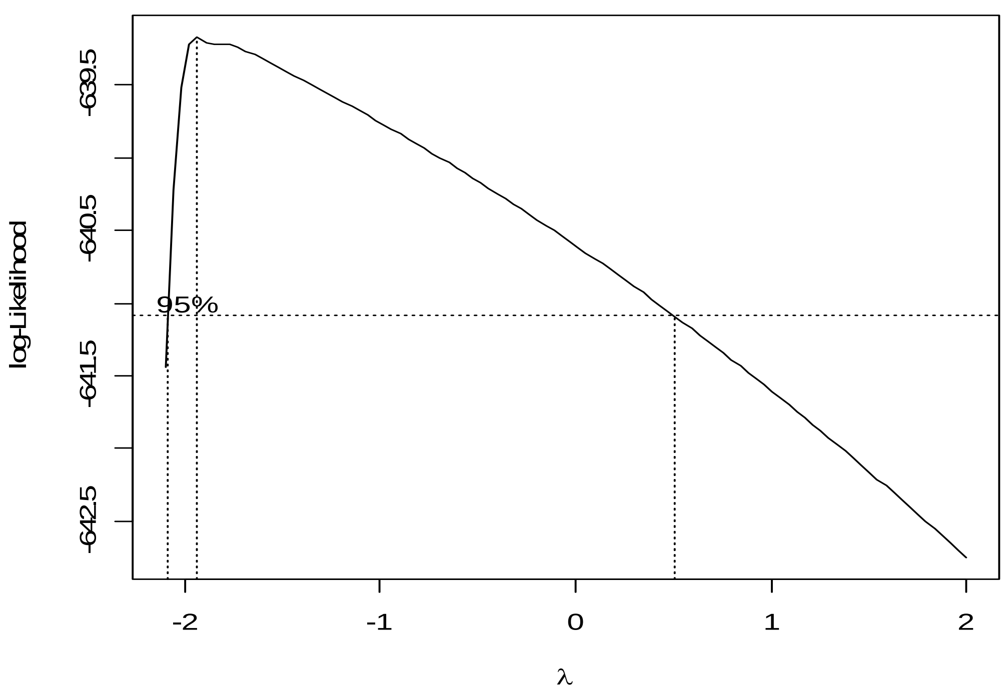

Figure 100: Continental Post 9/11 Modified Data Box-Cox Transformation Plot.

4.5.2 American Airlines

When the Box-Cox transformation was applied to American Airlines Figures 101 and 102 show that for the pre 9/11 data $(\hat{y}_t^{-2} - 1)/-2 = 5132985 + 7945t$ would be a more appropriate model and for the post 9/11 modified model $(\hat{y}_t^{-2} - 1)/-2 = 6182346.2 + 633.6t$ would be a better fit.

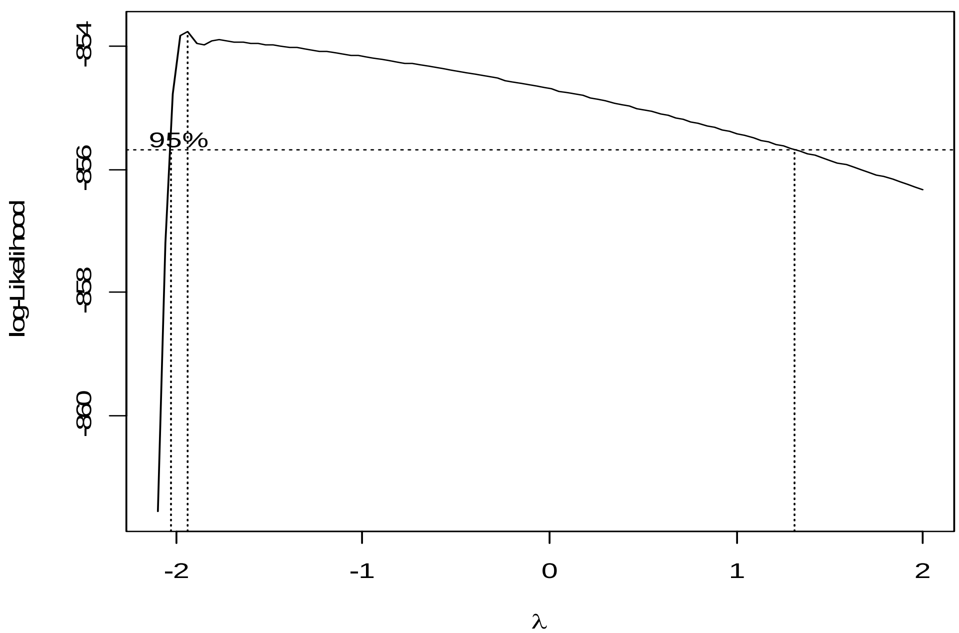

Figure 101: American Airlines Pre 9/11 Box-Cox Transformation Plot.

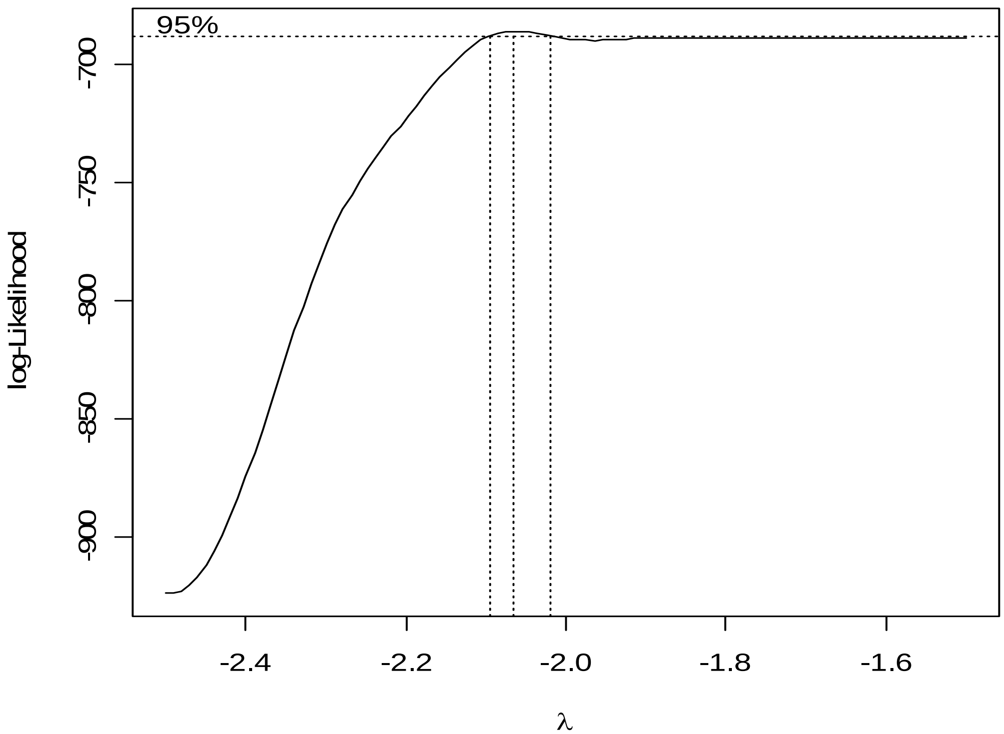

Figure 102: American Airlines Post 9/11 Modified Data Box-Cox Transformation Plot.

4.5.3 Delta Airlines

When applying the Box-Cox transformation on the Delta Airlines deseasonalized linear models, it was found that $(\hat{y}_t^{10.5} - 1)/10.5 = 7592920 + 13139t$ would better fit the pre 9/11 data and $(\hat{y}_t^{6} - 1)/6 = 8125729 - 16561t$ would be more appropriate for the post 9/11 data (Figures 103 & 104).

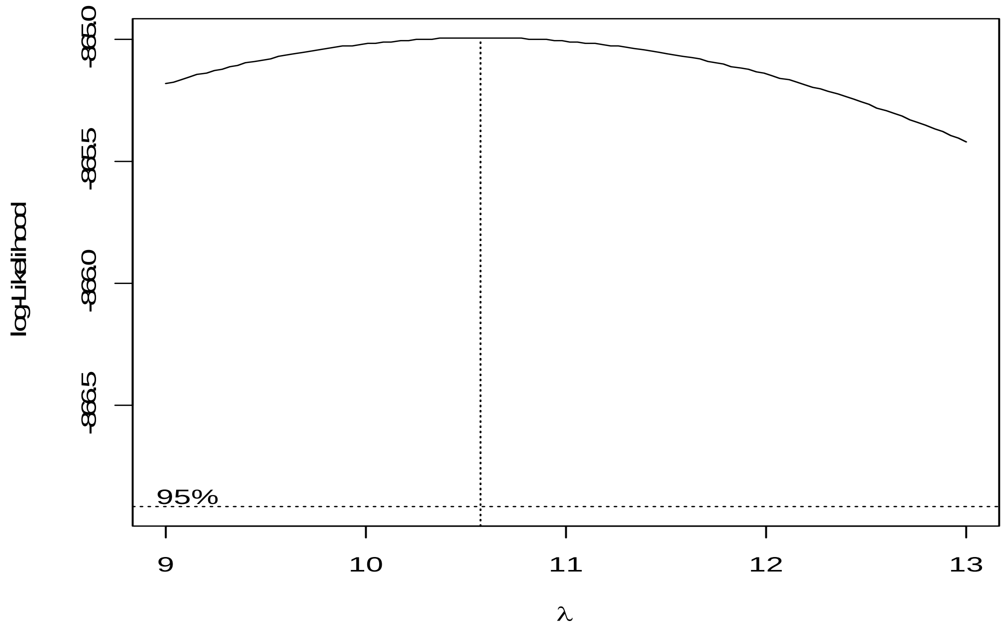

Figure 103: Delta Airlines Pre 9/11 Box-Cox Transformation Plot.

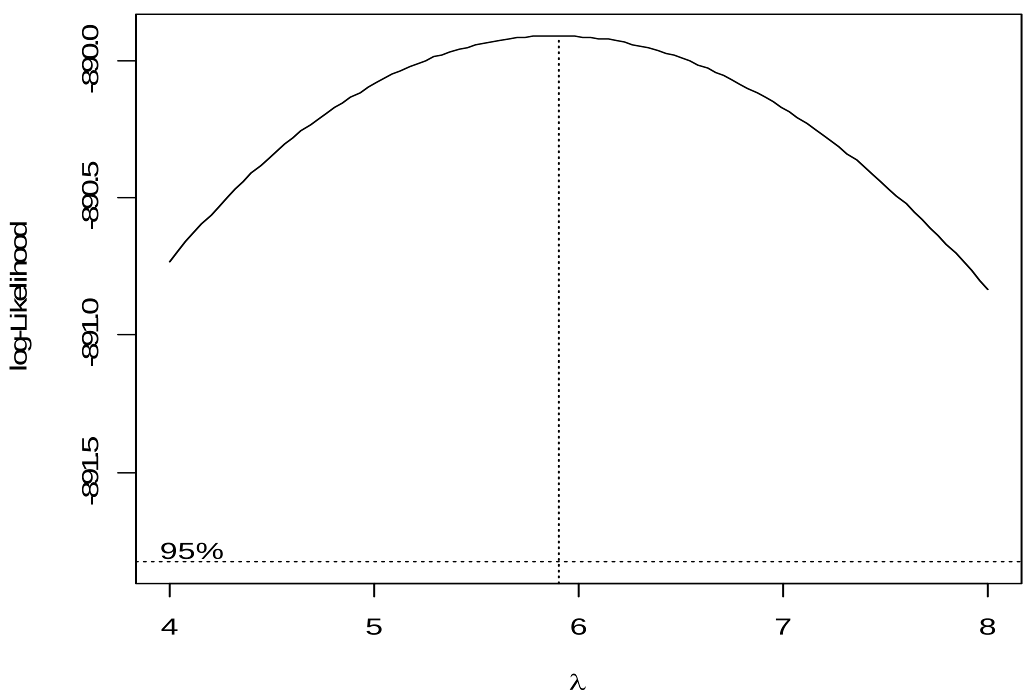

Figure 104: Delta Airlines Post 9/11 Box-Cox Transformation Plot.

5. Discussion

On September 11, 2001 four airplanes were hijacked. American Airlines Flight 11 was the first. It was crashed into the World Trade Center's North tower. It contained 73 passengers and 11 crewmembers.[3] United Airlines Flight 175 was the second airliner hijacked. It was crashed into the South Tower of the World Trade Center. It contained 56 passengers and 9 crewmembers.[4] American Airlines Flight 77 was the third airliner hijacked. It was crashed into the Pentagon killing 64 people on the plane and 125 in the buildings.[5] United Airlines Flight 93 was the fourth plane hijacked. Due to passenger insurrection, it did not find its target and crashed in an empty field outside Shanksville, PA, about 150 miles northwest of Washington D.C. It carried 37 passengers and 7 crewmembers.[6] There were a total of 2,974 fatalities, not including the 19 hijackers: 246 on the airplanes, 2,603 in the World Trade Center and the surrounding area, and 125 at the Pentagon.[7]

The major focus of this paper was to discover if the number of people traveling domestically has recovered from the September 11, 2001 terrorist attacks. This day had a dramatic affect on the U.S. airline industry. All civilian airplane traffic was grounded until Thursday, September 13, 2001.[8] All domestic airlines saw a drop in number of passengers in the days and months to follow. Now it is six years later and one wonders if the American people are still under the influence of terror.

The three airlines were modeled with a simple linear regression model. With this model predictions were made for September 2001, July 2002, May 2003, March 2004, January 2005, November 2005, and September 2006. These predictions were then

---

[3] http://en.wikipedia.org/wiki/American_Airlines_Flight_11

[4] http://en.wikipedia.org/wiki/United_Airlines_Flight_175

[5] http://en.wikipedia.org/wiki/American_Airlines_flight_77

[6] http://en.wikipedia.org/wiki/United_Airlines_Flight_93

[7] http://en.wikipedia.org/wiki/September_11,_2001_attacks#Fatalities

[8] http://en.wikipedia.org/wiki/Closings_and_cancellations_following_the_September_11%2C_2001_attacks #Travel_effects

compared to the actual numbers of passengers for those months. From Table 1 it can be seen that the only prediction interval that covered the actual number was that of July 2002 and the rest were below the actual numbers. This suggests that Continental Airlines is operating below its pre 9/11 passenger capacity. From Table 2 it can be seen that the majority of the actual numbers are within that of the prediction intervals. This suggests that American Airlines is operating as usual. From Table 3 it can be seen that the actual numbers are below that of the prediction intervals. This suggests that Delta Airlines is seating fewer passengers than it had pre 9/11.

The three airlines data were then fit with dummy variables to account for seasonal variation. With these models were calculated regression parameter intervals for both pre and post 9/11 data. The post 9/11 intervals were compared to the pre 9/11 intervals for each airline. The post 9/11 intervals for Continental Airlines were mostly below that of the pre 9/11 intervals (Table 4). This adds to the evidence that Continental is not doing as well now than it had pre 9/11. In Table 7 it is seen that the post 9/11 regression intervals for American Airlines are mostly overlapping that of the pre 9/11 intervals. This again suggests it is business as usual for American Airlines. The intervals for Delta Airlines post 9/11 data seem to overlap most of the pre 9/11 intervals (Table 10). This suggests that Delta is operating as usual.

Also using the dummy variable models, the fitted values for pre and post 9/11 data were compared. The Continental Airlines fitted values post 9/11 (Table 6) lie below that of the pre 9/11 fitted values (Table 5). This again suggests that Continental is still not up to pre 9/11 passenger numbers. American Airlines fitted values post 9/11are mostly above that of the pre 9/11 fitted values (Table 8 & 9). This suggests it is business as usual for American Airlines. Tables 11, and 12 show that the Delta Airlines post 9/11 fitted values are below that of pre 9/11 fitted values, suggesting that Delta is not seating as many passengers as it had pre 9/11.

The last model used to analyze the data was an additive decomposition model. This model was used to remove the seasonal variation in the data, and fit a model to the trend. In all airlines the quadratic model won out for both pre and post 9/11 data. The pre and post quadratic models were equated and solved for time to see when they crossed. In the case of Continental Airlines, Figure 54 shows the pre 9/11 quadratic regression line meets the post 9/11 quadratic regression line in October 2005. This suggests that as of October 2005 Continental Airlines reached the levels that would have been expected had 9/11 never happened. According to this model Continental had been on a decline and after 9/11 it dropped even further only to turn around and surpass its previous levels. Figure 76 shows the two quadratic models for American Airlines. American seemed to be on a positive trend before 9/11. Afterward, numbers dropped then began climbing again only to reach the pre 9/11 trend in August 2007, from there the numbers are expected to climb even further. Delta Airlines had illustrated a downward trend before 9/11. After the incident it continued on a downward trend meeting the pre 9/11 trend in August of 2003 (Figure 98). This suggests that 9/11 only reduced the intercept but the airline continued on the same path.

The additive decomposition models proved to the best fits to the data for all three airlines by AIC comparison (Appendix B). According to these models there was a drop in number of passengers for all three airlines, then Continental and American both picked up pace and overcame the incident, but Delta continued on the same downward trend.

It is understandable that fear of flying and added security measures which made flying more difficult was the cause of the decline in number of people traveling by air. As time went on, people became more confident and took to the skies once again. Continental and American Airlines are both benefiting from this recovery, but Delta is not. It could be that Delta did not have the financial solvency that Continental and American did to withstand the time of turmoil and grow once the fear had ceased.

A brief issued by the Bureau of Transportation Statistics reports that airlines cut employment and reduced the number of available seats to keep costs down while the public avoided air travel. Network air carriers shifted seat capacity to international flights as a reaction to the increasing dominance of low cost air carriers.[9]

Continental Airlines was one to bounce back more quickly than its competitors. The day following the airspace ban, Continental was flying a complete new schedule, well ahead of its competitors. This accomplishment was greatly attributed to the CrewSolver program designed by CALEB Technologies that Continental had adopted

---

[9] Bureau of Transportation Statistics (2005): “Airline Travel Since 9/11”, Brief #13 issued Dec 2005.

late 2000. This is an operations research based decision support program that could generate near optimal crew recovery solutions. This system was designed to quickly and efficiently get an airlines flight crews in place to fly following a major disruption to operations. By the end of 2001, continental had saved $40 million in recovery costs due to this valuable asset.[10]

Along with cutting costs American Airlines instilled an electronic ticketing program. Self-service ticketing kiosks were installed with an average check in time of 85 seconds. They've also added checking in at home on the Internet at aa.com and checking in by phone. One can check in at curbside with wireless kinds of kiosk devices. American has focused on smoothing out the travel experience to enhance its recover.[11]

Delta Airlines still struggles with recovery. On Dec. 31 2003 Delta's CEO Leo Mullin stepped down as the company's chief executive. There is speculation his departure is due to his incapability of setting the company on a profitable course.[12] On Wednesday September 14, 2005 Delta filed bankruptcy protection.[13] On Tuesday December 19, 2006 Delta filed a five-year reorganization business plan to emerge from bankruptcy as a stand-alone company. The plan was set in opposition to a hostile takeover bid.[14] As of November 2007 Delta and United Airlines are discussing a merger.[15] According to the merger the "United" name would be retained, Delta may be no more.

---

[10] http://lionhrtpub.com/orms/orms-6-02/fredelman.html

[11] http://www.eweek.com/article2/0,1895,1503743,00.asp

[12] http://www.usatoday.com/money/biztravel/2003-11-25-delta-cover_x.htm

[13] http://www.msnbc.msn.com/id/9317550/

[14] http://www.foxnews.com/story/0,2933,237377,00.html

[15] http://www.denverpost.com/airlines/ci_7460498

Of the three airlines analyzed in this paper, two are headed for positive recovery. The third, Delta, may not be in existence much longer. The recovery of Continental and American shows that the public is no longer being terrorized and has ultimately returned to normal life travel.

The analysis in this paper was limited. Three airlines were chosen from the airlines in operation. It is suggested that a larger sample be analyzed. Three tests were administered and conclusions drawn. It is recommended that more tests be performed before an accurate conclusion can be made. The Box-Cox transformations also add to the shortcomings of this paper. It is suggested that the transformations be investigated further and analysis performed with the suggested models.

6. Conclusion

“recover \ri-‘kəv-ər\ *vb* **2**: to regain normal health, poise, or status”
-The Merriam-Webster Dictionary. 11th ed. 1974

Airline industry recovery can be considered in many different terms. Trend, revenue, number of available seats, number of empty seats and number of flights are all ways in which to measure recovery. According to BreakingNews.iol.ie the aviation industry made a convincing recovery with 27.4 million flights in 2005 which was higher than it had been since 2001. According to MIT Global Airline Industry Program, in the United States the airline industry posted cumulative net losses of over $40 billion from 2001 to 2005, and only in 2006 was it able to return to the black with a total net profit of just over $3 billion. In this paper recovery is based on the post 9/11 regression equation crossing the pre 9/11 regression equation that has been extended into the future, thought to represent the path of the airline had the September 11th attacks never occurred, thus, a return to normal status.

Conclusions of this paper are drawn from the additive decomposition models, which displayed the best fits to the airline data collected. Continental, American, and Delta airlines have all recovered to the pre 9/11 point. The fear caused by the 9/11 terrorist attacks has subsided and no longer affects the American public’s choice of traveling by air.

7. References

1. Bowerman, Bruce L., Richard T. O'Connell, and Anne B. Koehler, Forecasting, Time Series, and Regression: An Applied Approach. 4th ed. Belmont: Thomson Brooks/Cole, 2005.

2. Bureau of Transportation Statistics (2005): "Airline Travel Since 9/11", Brief #13 issued Dec 2005.

3. http://en.wikipedia.org/wiki/American_Airlines_Flight_11

4. http://en.wikipedia.org/wiki/United_Airlines_Flight_175

5. http://en.wikipedia.org/wiki/American_Airlines_flight_77

6. http://en.wikipedia.org/wiki/United_Airlines_Flight_93

7. http://en.wikipedia.org/wiki/September_11,_2001_attacks#Fatalities

8. http://en.wikipedia.org/wiki/
Closings_and_cancellations_following_the_September_11%2C_2001_attacks#Travel_effects

9. http://en.wikipedia.org/wiki/Box-Cox_transformation

10. http://breakingnews.iol.ie/news/ireland/print/cwqlojojcwql/

11. http://cran.r-project.org/doc/contrib/Faraway-PRA.pdf

12. http://lionhrtpub.com/orms/orms-6-02/fredelman.html

13. http://web.mit.edu/airlines/www/the-airline-industry/the-airline-industry.htm

14. http://www.denverpost.com/airlines/ci_7460498

15. http://www.eweek.com/article2/0,1895,1503743,00.asp

16. http://www.foxnews.com/story/0,2933,237377,00.html

17. http://www.itl.nist.gov/div898/handbook/eda/section3/boxcoxli.htm

18. http://www.math.yorku.ca/SCS/sasmac/boxcox.html

19. http://www.msnbc.msn.com/id/9317550/

20. http://www.usatoday.com/money/biztravel/2003-11-25-delta-cover_x.htm

21. Ito, H., and D. Lee (2003): "Assessing the Impact of the September 11 Terrorist Attacks on U.S. Airline Demand,"
http://www.brown.edu/Departments/Economics/Papers/2003/2003-16_paper.pdf

22. Kerns, Dr. G. J., Applied Regression and Time Series, manuscript.

23. The Merriam-Webster Dictionary. 11th ed. 1974

Appendix A: Code

Italics indicate label names and variable names changed throughout the paper.

Import data from excel file

Independent Variable = `scan()`
Dependent Variable = `scan()`

Simple Scatter-Plot

```
plot( Data Name~Data Name,
xlab = "Label Name",
ylab = "Label Name",
main = "Title " )
```

Fit Regression Line to Linear Model

```
Linear Model Name = lm( Dependent Var~Independent Var )
```

Summarize Model

```
summary( Model Name )
```

Graph Fitted Line with Confidence Bands and Prediction Intervals For Linear Model

```
a=signif( Model Name$coefficients[1], digits=6)
b=signif( Model Name$coefficients[2], digits=6)
```

Draw Scatterplot

```
plot( Data Name~Data Name,
main = "Title",
sub = paste( "Y=",a,"+",b,"* X",sep="" ))
```

Draw Regression Line

```
abline(Model Name$coef,lty=1,lwd=2)
```

Sort Points to Calculate Confidence Bands

```
xy = data.frame( Independent Var=sort( Independent Var ))
```

Calculate Confidence Bands at Those Points

```
yhat.conf = predict( Model Name,newdata = xy,
interval = "confidence")
yhat.pred = predict( Model Name,newdata = xy,
interval = "prediction" )
```

Save the Confidence Band Calculations

```
ci.c= data.frame(lower = yhat.conf[,"lwr"],
upper = yhat.conf[,"upr"])
ci.p = data.frame(lower = yhat.pred[,"lwr"],
upper = yhat.pred[,"upr"])
```

Draw the Confidence Bands

```
lines(xy$Independent Var,ci.c$lower,lty=5,lwd=3)
lines(xy$Independent Var,ci.c$upper,lty=5,lwd=3)
```

Draw the Prediction Intervals

```
lines(xy$ Independent Var,ci.p$lower,
         lty=3,lwd=3)
lines(xy$ Independent Var,ci.p$upper,
         lty=3,lwd=3)
```

Make predictions for Linear Model

Import Actual Data to Compare to Predictions

```
actual = scan()
```

Make Predictions and Table with Actual Data

```
predict.here=c(69,79,89,99,109,119,129)
p.f.t = data.frame(Independent Var =predict.here)
p.conf = predict(Model Name,newdata=p.f.t,
int="confidence",
level=0.95)
p.pred = predict(Model Name,newdata=p.f.t,
int="prediction",
level=0.95)
z = data.frame(p.f.t,p.conf,p.pred,actual)
names(z) =
c("Month","fit","lwr.CI","upr.CI","fit","lwr.PI","upr.PI","
Actual")
attach(z)
data.frame(Month,fit,lwr.CI,upr.CI,fit,lwr.PI,upr.PI,Actual
)
     detach(z)
```

Add period for Dummy Variables

```
period=scan()
period = factor( period, levels = 1:12)

levels(period) = c( "Jan", "Feb",
                    "Mar", "Apr",
                    "May", "Jun",
                    "Jul", "Aug",
                    "Sep", "Oct",
                    "Nov", "Dec")
```

Plot Data

```
 plot( as.vector(Independent Var), as.vector( Dependent Var),
 type ="b",
 main = "Title",
 xlab = "Label",
 ylab = "Label")
```

Fit Data with Dummy Variables

```
Model Name = lm( Dependent Var ~ Independent Var + period )
```

Make fitted line plot

```
plot(   as.vector(Independent Var), as.vector(Dependent Var),
        type = "b",
    main = "Title",
    xlab = "Label",
    ylab = "Lable")
    lines( predict(Model Name), lty = 1, lwd = 3)
```

Fit linear model with 2 sets trig

```
Model Name = lm( Dependent Var ~ Independent Var
              + I(sin(2*pi* Independent Var /12))
              + I(cos(2*pi* Independent Var /12))
              + I(sin(4*pi* Independent Var /12))
              + I(cos(4*pi* Independent Var /12))
               )
```

Fit Linear Model with 3 sets of trig

```
  Model Name = lm(Dependent Var ~ Independent Var
              + I(sin(2*pi* Independent Var /12))
              + I(cos(2*pi* Independent Var /12))
              + I(sin(4*pi* Independent Var /12))
              + I(cos(4*pi* Independent Var /12))
              + I(sin(6*pi* Independent Var /12))
              + I(cos(6*pi* Independent Var /12))
               )
```

Fit linear model with 4 sets of trig

```
Model Name = lm(Dependent Var ~    Independent Var
                + I(sin(2*pi* Independent Var /12))
                + I(cos(2*pi* Independent Var /12))
                + I(sin(4*pi* Independent Var /12))
                + I(cos(4*pi* Independent Var /12))
                + I(sin(6*pi* Independent Var /12))
                + I(cos(6*pi* Independent Var /12))
                + I(sin(8*pi* Independent Var /12))
    + I(cos(8*pi* Independent Var /12))
          )
```

Compare dummy and trig models

Make side-by-side plots

```
par( mfrow = c(2,2))

plot( as.vector(Independent Var), as.vector(Dependent Var),
    type = "b",
    main = "Title (Dummy Variable Model)",
    xlab = "Label",
    ylab = " Label ")
lines( predict(Model Name), lty = 1, lwd = 2, col =
      "green4")

plot( as.vector(Independent Var), as.vector(Dependent Var),
   type = "b",
   main = "Title (2 sets Trig Fns)",
   xlab = " Label ",
   ylab = " Label ")
lines( predict(Model Name), lty = 1, lwd = 2, col =
      "green4")

plot( as.vector(Independent Var), as.vector(Dependent Var),
    type = "b",
    main = "Title (3 sets Trig Fns)",
    xlab = " Label ",
    ylab = " Label ")
lines( predict(Model Name), lty = 1, lwd = 2, col =
       "green4")
```

```
plot( as.vector(Independent Var), as.vector(Dependent Var),
    type = "b",
    main = "Title (4 sets Trig Fns)",
    xlab = " Label ",
    ylab = " Label ")
lines( predict(Model Name), lty = 1, lwd = 2, col =
       "green4")

par(mfrow = c(1,1))
```

Find AIC

```
AIC(Model Name)
```

Create Residual Plots

```
par(mfrow = c(2,2))
plot(Model Name)
par(mfrow = c(1,1))
```

Diagnostic Tests

```
library(lmtest)
shapiro.test( resid( Model Name ) )
bptest( Model Name )
dwtest( Model Name ,alternative = "two.sided")

cooksD=cooks.distance(Model Name)
f0.50=qf(0.5, df1=5,df2=24)
cooksD[which(cooksD>f0.50)]
```

Analysis of Variance Comparisons

```
anova(Model Name, Model Name)
```

Create Time Series

```
Model Name = ts( Dependent Variable,
        start = c(year, # month),
        freq = 12)
```

Plot the Time Series

```
plot( Model Name,
     type = "b",
     main = "Title")
```

Decompose the Time Series

```
Model Name= decompose( Model Name,
          type = "additive",
          filter = rep(1/12, 12))
```

Save the Results

```
CMAt  = Model Name$trend              # centered moving averages
SNt  = Model Name$seasonal            # seasonal component
IRt  = Model Name$random              # irregular component
seas.Figure = Model Name$Figure       # est. seasonal Figure
```

Plot Components on Same Graph

```
  plot( Model Name , type = "b")
```

Deseasonalize the Observations

```
DSt  = Model Name - SNt
```

Look at the Scatterplot

```
plot( DSt,
    xlab = "Time(mo.)",
    ylab = "DSt",
    type = "b",
    main = "Deseasonalized Observations")
```

Fit With linear model

```
times.ds = 1:length(DSt)

DSt.lm = lm( DSt ~ times.ds )
```

Look at the Results

```
plot(   as.vector(DSt)~times.ds,
   xlab = "Time(mo.)",
   ylab = "DSt",
   main = "Label")
lines( predict( Model Name ),lty = 1, lwd = 3 )
```

Fit With a Quadratic Model

```
DStquad.lm = lm( DSt ~ times.ds+I(times.ds^2) )
```

Fit With a Cubic Model

```
DStcubic.lm = lm( DSt ~
              times.ds+I(times.ds^2)+I(times.ds^3) )
```

Plot Quadratic Pre and Post Models Together

```
plot( Independent Variable pre 9/11,
   Dependent Var Pre 9/11,
   xlim = c(1, 200),
   ylim = c(Dependent Variable Minimum, Dependent Variable Maximum),

   )

 par(add=TRUE)

points(Independent Variable post 9/11,
   Independent Variable post 9/11
   )

curve( Quadratic Equation for pre 9/11 Model,
    from = 1,
    to = 165,
    lty = 5,
    lwd = 3,
    add = TRUE)

curve( Quadratic Equation for post 9/11 Model,
    from = 70,
    to = 165,
    lwd = 3,
    add = TRUE)
abline( v = Time of Crossing,   lty = 2)
```

Box-Cox Transformations

```
boxcox(Model Name, lambda = seq(#, #, 1/10), plotit = TRUE,
       interp = TRUE, eps = 1/50, xlab =
expression(lambda),
       ylab = "log-Likelihood")
```

Appendix B: Model Comparison

Continental	Adj Rsq	F stat	AIC	DW test	BP test	Shapiro	Cook's	Anova
Dummy Variables								
Before Data	9.4060E-01	89.450000	1708.754000	2.9690E-02	1.2840E-01	0.869600	0.000000	
minus outs	5.5090E-01	6.352000	1771.723000	6.9590E-01	6.6820E-02	0.634500	0.000000	
After Data	8.1890E-01	23.610000	1606.067000	3.5510E-10	4.2350E-02	0.015010	0.000000	
minus outs	8.4640E-01	26.260000	1467.702000	2.5420E-10	4.3750E-02	0.182500	0.000000	
Trig Functions	Adj Rsq	F stat	AIC	DW test	BP test	Shapiro	Cook's	Anova
Before Data								
2 sets	6.3920E-01	24.740000	1825.595000	1.7430E-01	7.6570E-04	0.470100		
3 sets	8.0690E-01	40.990000	1784.876000	1.3000E-04	2.3780E-03	0.407600		
4 sets	8.0670E-01	32.060000	1786.646000	1.6660E-04	8.9630E-03	0.192200		
After Data								
2 sets	4.1820E-01	9.626000	1671.560000	5.7680E-01	1.6990E-02	0.315300		
3 sets	6.8310E-01	19.480000	1636.236000	3.7090E-01	1.3400E-01	0.396700		
4 sets	6.8720E-01	15.650000	1637.097000	4.0420E-01	4.2560E-01	0.481000		
TIME SERIES	Adj Rsq	F stat	AIC	DW test	BP test	Shapiro	Cook's	Anova
Before Data								
Linear model	8.3490E-01	299.300000	1492.421000	5.1250E-02	9.7650E-01	0.049490	0.000000	
Quad Model	8.5670E-01	177.300000	1484.884000	2.4060E-01	2.1770E-02	0.691500	0.000000	2.727000E-03
Cubic Model	8.5660E-01	118.500000	1485.852000	2.4020E-01	3.3720E-02	0.944400	0.000000	3.286000E-01
After Data	Adj Rsq	F stat	AIC	DW test	BP test	Shapiro	Cook's	Anova
Linear Model	5.0260E-01	60.620000	1561.076000	6.0820E-11	6.3050E-01	0.000572	0.000000	
Quad Model	7.7210E-01	100.900000	1515.208000	2.5990E-03	3.3590E-04	0.606700	0.000000	1.869000E-11
Cubic Model	7.6800E-01	66.110000	1517.206000	1.6130E-03	8.4650E-04	0.596300	0.000000	9.707000E-01
after 1st 5 mos								
Linear Model	5.6320E-01	27.000000	1239.605000	2.3720E-10	5.8560E-01	0.000312	0.000000	

Quad Model	6.6060E-01	46.740000	1209.818000	4.4510E-02	6.2430E-02	0.700800	0.000000	5.624000E-08
cubic	6.5530E-01	30.790000	1211.479000	3.2730E-02	9.5520E-02	0.807700	0.000000	5.796000E-01

American Airlines	Adj Rsq	F stat	AIC	DW test	BP test	Shapiro	Cook's	Anova
Dummy Variables								
Before Data	8.2160E-01	26.710000	1867.917000	2.8230E-03	2.8220E-01	0.028970	0.000000	
minus outs	4.4260E-01	5.235000	1852.998000	4.1860E-02	1.0250E-01	0.692800	0.000000	
After Data	6.1750E-01	9.060000	1765.758000	1.2130E-10	1.9470E-02	0.004236	0.000000	
minus outs	7.7070E-01	16.960000	1628.882000	4.5050E-10	3.8450E-01	0.201000	0.000000	

Trig Functions	Adj Rsq	F stat	AIC	DW test	BP test	Shapiro	Cook's	Anova
Before Data								
2 sets	4.6280E-01	12.550000	1937.003000	6.7650E-01	3.1120E-02	0.561300		
3 sets	6.5830E-01	19.440000	1908.009000	2.8710E-01	2.2780E-01	0.259000		
4 sets	6.7350E-01	16.350000	1906.620000	3.6970E-01	6.2000E-01	0.268100		
After Data								
2 sets	4.5680E-01	11.090000	1781.400000	2.2790E-02	4.4350E-03	0.890800		
3 sets	5.4420E-01	11.230000	1772.447000	4.4260E-03	1.3750E-03	0.339300		
4 sets	5.4470E-01	8.976000	1774.027000	3.8440E-03	7.8880E-03	0.272700		

Time Series	Adj Rsq	F stat	AIC	DW test	BP test	Shapiro	Cook's	Anova
Before Data								
Linear Model	3.0790E-01	27.250000	1641.439000	6.1820E-04	5.1800E-02	0.532400	0.000000	
Quad Model	3.3710E-01	16.000000	1639.811000	1.2630E-03	3.6390E-02	0.690800	0.000000	6.456000E-02
Cubic Model	3.3160E-01	10.760000	1641.243000	9.4140E-04	2.9030E-02	0.658900	0.000000	4.685000E-01

After Data	Adj Rsq	F stat	AIC	DW test	BP test	Shapiro	Cook's	Anova
Linear Model	1.3870E-01	10.500000	1724.603000	6.8960E-12	5.5270E-04	0.000024	0.000000	
Quad Model	1.3200E-01	5.486000	1726.028000	2.3680E-12	2.1880E-06	0.000291	1;.9599551	4.619000E-01

Cubic Model	1.9030E-01	5.623000	1722.791000	3.3870E-12	1.0020E-07	0.044780	1;1.121718	2.773000E-02
after 1st 3 mos								
Linerar Model	-.02047	0.057000	1335.172000	2.9190E-10	9.1210E-01	0.051400	0.000000	
Quad Model	5.0410E-01	24.890000	1301.481000	1.0920E-02	1.2450E-02	0.750400	0.000000	8.691000E-09
Cubic Model	5.0760E-01	17.150000	1302.055000	1.0280E-02	7.0040E-02	0.408100	0.000000	2.556000E-01
Delta Airlines	Adj Rsq	F stat	AIC	DW test	BP test	Shapiro	Cook's	Anova
Dummy Variables								
Before Data	5.2410E-10	18.210000	1936.733000	2.0530E-10	4.3550E-01	0.001582	0.000000	
minus outs	5.3780E-01	7.205000	1893.782000	7.2130E-05	9.1120E-02	0.000412	0.000000	
After Data	6.1870E-01	9.113000	1767.726000	4.4910E-12	8.2360E-02	0.399200	0.000000	
minus outs	6.5150E-01	10.190000	1734.345000	2.9800E-11	1.2690E-02	0.613000	0.000000	
Trig Functions								
Before Data								
2 sets	4.4570E-01	11.780000	1986.390000	2.1170E-01	8.9370E-02	0.424500		
3 sets	6.3650E-01	17.760000	1959.470000	1.5270E-01	5.8160E-01	0.172500		
4 sets	6.3860E-01	14.150000	1960.784000	1.4940E-01	9.0340E-01	0.025960		
After Data								
2 sets	3.4010E-01	7.185000	1795.487000	2.9470E-02	2.1360E-01	0.500900		
3 sets	5.3210E-01	10.750000	1776.257000	1.3710E-03	1.1940E-01	0.551700		
4 sets	5.2890E-01	8.485000	1778.321000	1.5360E-03	2.7660E-01	0.796100		
Time Series	Adj Rsq	F stat	AIC	DW test	BP test	Shapiro	Cook's	Anova
Before Data								
Linear Model	4.1010E-01	42.020000	1675.825000	2.7530E-06	9.2710E-01	0.000026	0.000000	
Quad Model	6.7380E-01	61.920000	1641.245000	8.4600E-02	3.7670E-03	0.227700	0.000000	4.344000E-09
Cubic Model	6.7330E-01	41.540000	1642.261000	8.2800E-02	4.3830E-03	0.089420	0.000000	2.437000E-08

After Data	Adj Rsq	F stat	AIC	DW test	BP test	Shapiro	Cook's	Anova
Linear Model	3.3540E-01	30.770000	1722.292000	2.1870E-14	2.1250E-02	0.052440	0.000000	
Quad Model	4.4700E-01	24.620000	1712.467000	1.3850E-12	6.7080E-05	0.635900	0.000000	8.461000E-04
Cubic Model	4.3650E-01	16.240000	1714.280000	5.0840E-13	3.0200E-05	0.595100	0.000000	6.773000E-01
after 1st mo								
Linear Model	3.8180E-01	37.440000	1717.659000	3.9490E-14	2.6440E-03	0.122800	0.000000	
Quad Model	4.3840E-01	24.030000	1712.857000	2.5680E-13	3.5320E-05	0.591200	0.000000	1.137000E-02
Cubic Model	4.2980E-01	15.830000	1714.703000	8.0130E-14	1.5770E-05	0.721800	0.000000	7.061000E-01

Appendix C: Data Sets

Continental Airlines

Year	Period	Passengers	Months	Year	Period	Passengers	Months
1996	1	2263393	1	1999	9	2809849	45
1996	2	2408092	2	1999	10	3104524	46
1996	3	2892031	3	1999	11	3004742	47
1996	4	2782114	4	1999	12	2985688	48
1996	5	2785627	5	2000	1	2722353	49
1996	6	2757916	6	2000	2	2753811	50
1996	7	2897983	7	2000	3	3288287	51
1996	8	2990899	8	2000	4	3088453	52
1996	9	2444203	9	2000	5	3168629	53
1996	10	2773420	10	2000	6	3235860	54
1996	11	2465543	11	2000	7	3228168	55
1996	12	2755931	12	2000	8	3232101	56
1997	1	2548843	13	2000	9	2783075	57
1997	2	2449879	14	2000	10	3053317	58
1997	3	3029166	15	2000	11	3043285	59
1997	4	2843453	16	2000	12	2993479	60
1997	5	2877189	17	2001	1	2802902	61
1997	6	2960261	18	2001	2	2690298	62
1997	7	3097152	19	2001	3	3238694	63
1997	8	3098441	20	2001	4	3196241	64
1997	9	2598839	21	2001	5	3198848	65
1997	10	2876237	22	2001	6	3218545	66
1997	11	2687255	23	2001	7	3329105	67
1997	12	2791398	24	2001	8	3433692	68
1998	1	2505533	25	2001	9	1836826	69
1998	2	2483183	26	2001	10	2479425	70
1998	3	3055936	27	2001	11	2592933	71
1998	4	3044904	28	2001	12	2617523	72
1998	5	3027110	29	2002	1	2431774	73
1998	6	3037127	30	2002	2	2391658	74
1998	7	3197900	31	2002	3	2944301	75
1998	8	3175172	32	2002	4	2754391	76
1998	9	2810884	33	2002	5	2772405	77
1998	10	2973727	34	2002	6	2770761	78
1998	11	2776807	35	2002	7	2877723	79
1998	12	2896817	36	2002	8	2871211	80
1999	1	2710455	37	2002	9	2226421	81
1999	2	2676593	38	2002	10	2533451	82
1999	3	3194314	39	2002	11	2411641	83
1999	4	3064618	40	2002	12	2667298	84
1999	5	3035494	41	2003	1	2284065	85
1999	6	3102399	42	2003	2	2182741	86
1999	7	3251343	43	2003	3	2686997	87
1999	8	3189547	44	2003	4	2521969	88

year	period	passengers	Months
2003	5	2628289	89
2003	6	2797292	90
2003	7	3020299	91
2003	8	2806609	92
2003	9	2228328	93
2003	10	2530875	94
2003	11	2493280	95
2003	12	2672575	96
2004	1	2275091	97
2004	2	2301375	98
2004	3	2762806	99
2004	4	2751538	100
2004	5	2601105	101
2004	6	2822954	102
2004	7	2978741	103
2004	8	2786578	104
2004	9	2267568	105
2004	10	2654994	106
2004	11	2633638	107
2004	12	2692307	108
2005	1	2391679	109
2005	2	2318987	110
2005	3	3017625	111
2005	4	2750674	112
2005	5	2789924	113
2005	6	2897037	114
2005	7	3030896	115
2005	8	2899686	116
2005	9	2366756	117
2005	10	2769157	118
2005	11	2805470	119
2005	12	2933328	120
2006	1	2709194	121
2006	2	2584790	122
2006	3	3206052	123
2006	4	3101586	124
2006	5	3063694	125
2006	6	3156238	126
2006	7	3211750	127
2006	8	3086493	128
2006	9	2629849	129
2006	10	2983229	130

American Airlines

year	period	passengers	Months	year	period	passengers	Months
1996	1	4699990	1	1999	9	5102026	45
1996	2	4597904	2	1999	10	5708910	46
1996	3	5426538	3	1999	11	5621010	47
1996	4	5181626	4	1999	12	5379277	48
1996	5	5292983	5	2000	1	5014368	49
1996	6	5526231	6	2000	2	5089190	50
1996	7	5751187	7	2000	3	6183512	51
1996	8	5791567	8	2000	4	5814650	52
1996	9	4579327	9	2000	5	5948717	53
1996	10	5224778	10	2000	6	6157160	54
1996	11	4795804	11	2000	7	6410520	55
1996	12	5475811	12	2000	8	6286932	56
1997	1	4999404	13	2000	9	5213922	57
1997	2	4407407	14	2000	10	5538563	58
1997	3	5836300	15	2000	11	5429690	59
1997	4	5310512	16	2000	12	5231487	60
1997	5	5461785	17	2001	1	5030187	61
1997	6	5700848	18	2001	2	4760199	62
1997	7	5954308	19	2001	3	5802343	63
1997	8	5859698	20	2001	4	5596152	64
1997	9	4857325	21	2001	5	5571864	65
1997	10	5275200	22	2001	6	5861793	66
1997	11	4964962	23	2001	7	6197348	67
1997	12	5349133	24	2001	8	6176936	68
1998	1	4953995	25	2001	9	3251642	69
1998	2	4647667	26	2001	10	4231626	70
1998	3	5540049	27	2001	11	4521123	71
1998	4	5394716	28	2001	12	4702768	72
1998	5	5441333	29	2002	1	5767838	73
1998	6	5650669	30	2002	2	5526501	74
1998	7	5879461	31	2002	3	6898499	75
1998	8	5720420	32	2002	4	6429205	76
1998	9	5048905	33	2002	5	6717971	77
1998	10	5348549	34	2002	6	7061784	78
1998	11	5115614	35	2002	7	7368482	79
1998	12	5245381	36	2002	8	7247622	80
1999	1	4905986	37	2002	9	5624627	81
1999	2	3985857	38	2002	10	6487677	82
1999	3	5529802	39	2002	11	5840822	83
1999	4	5324438	40	2002	12	6477632	84
1999	5	5242486	41	2003	1	5660534	85
1999	6	5559808	42	2003	2	5225578	86
1999	7	5915488	43	2003	3	6310554	87
1999	8	5617765	44	2003	4	5907580	88

year	period	passengers	Months
2003	5	6218797	89
2003	6	6548796	90
2003	7	7007963	91
2003	8	6699904	92
2003	9	5271735	93
2003	10	6004427	94
2003	11	5449401	95
2003	12	5897565	96
2004	1	5475307	97
2004	2	5495032	98
2004	3	6425910	99
2004	4	6211663	100
2004	5	6081481	101
2004	6	6520999	102
2004	7	6718150	103
2004	8	6407856	104
2004	9	5326779	105
2004	10	6041460	106
2004	11	5820897	107
2004	12	6122148	108
2005	1	5784397	109
2005	2	5475154	110
2005	3	6838822	111
2005	4	6419180	112
2005	5	6692714	113
2005	6	7070922	114
2005	7	7367486	115
2005	8	6770842	116
2005	9	5906868	117
2005	10	6213660	118
2005	11	6338404	119
2005	12	6418518	120
2006	1	6071348	121
2006	2	5650371	122
2006	3	6897283	123
2006	4	6614216	124
2006	5	6804585	125
2006	6	6913522	126
2006	7	7033467	127
2006	8	6503729	128
2006	9	5676574	129
2006	10	6253772	130

Delta Airlines

Year	Period	Passengers	Months	Year	Period	Passengers	Months
1996	1	6349814	1	1999	9	7375299	45
1996	2	6669471	2	1999	10	8287283	46
1996	3	7737882	3	1999	11	8008989	47
1996	4	7305044	4	1999	12	7720908	48
1996	5	7617605	5	2000	1	6938581	49
1996	6	7973500	6	2000	2	7427618	50
1996	7	7888452	7	2000	3	9085719	51
1996	8	8188367	8	2000	4	8503033	52
1996	9	6987007	9	2000	5	8686854	53
1996	10	7773685	10	2000	6	9041929	54
1996	11	7147461	11	2000	7	9104110	55
1996	12	7959943	12	2000	8	8792019	56
1997	1	7226700	13	2000	9	7224010	57
1997	2	7096898	14	2000	10	8021386	58
1997	3	8547915	15	2000	11	7891007	59
1997	4	8071289	16	2000	12	7248784	60
1997	5	8186801	17	2001	1	6863536	61
1997	6	8496901	18	2001	2	6709279	62
1997	7	8749921	19	2001	3	8349728	63
1997	8	8495512	20	2001	4	7961634	64
1997	9	7262784	21	2001	5	7968860	65
1997	10	8084686	22	2001	6	8404049	66
1997	11	7642422	23	2001	7	8647677	67
1997	12	8103918	24	2001	8	8515205	68
1998	1	7292261	25	2001	9	4651514	69
1998	2	7023277	26	2001	10	6168312	70
1998	3	8713187	27	2001	11	6241481	71
1998	4	8374600	28	2001	12	6407136	72
1998	5	8538588	29	2002	1	6047842	73
1998	6	8668770	30	2002	2	6037664	74
1998	7	8961296	31	2002	3	7621980	75
1998	8	8677041	32	2002	4	7012477	76
1998	9	7823477	33	2002	5	7066533	77
1998	10	8089284	34	2002	6	7420655	78
1998	11	7707565	35	2002	7	7700152	79
1998	12	8009431	36	2002	8	7695594	80
1999	1	7409258	37	2002	9	6184794	81
1999	2	7328567	38	2002	10	7092859	82
1999	3	8788817	39	2002	11	6632686	83
1999	4	8454452	40	2002	12	7233741	84
1999	5	8415804	41	2003	1	6261054	85
1999	6	8732928	42	2003	2	5871287	86
1999	7	9066602	43	2003	3	6974199	87
1999	8	8623364	44	2003	4	6160049	88

Year	Period	Passengers	Months
2003	5	6337258	89
2003	6	6823463	90
2003	7	7190969	91
2003	8	7009446	92
2003	9	5745634	93
2003	10	6675188	94
2003	11	6240097	95
2003	12	6503916	96
2004	1	5814214	97
2004	2	5947669	98
2004	3	6947233	99
2004	4	6711378	100
2004	5	6763930	101
2004	6	7230138	102
2004	7	7469377	103
2004	8	6999817	104
2004	9	5609903	105
2004	10	6738373	106
2004	11	6544785	107
2004	12	6596761	108
2005	1	6186057	109
2005	2	6041572	110
2005	3	7456338	111
2005	4	6860192	112
2005	5	6922107	113
2005	6	7056723	114
2005	7	7395121	115
2005	8	6983124	116
2005	9	5611032	117
2005	10	5810108	118
2005	11	5742877	119
2005	12	5516023	120
2006	1	5180677	121
2006	2	4833717	122
2006	3	5993848	123
2006	4	5365312	124
2006	5	5150466	125
2006	6	5625227	126
2006	7	7033467	127
2006	8	6503729	128
2006	9	5676574	129
2006	10	6253772	130