

# DTREG

## Predictive Modeling Software

[www.dtreg.com](http://www.dtreg.com)



## Time-Series Analysis and Forecasting

Phil Sherrod  
[phil@philsherrod.com](mailto:phil@philsherrod.com)

Copyright © 2014, all rights reserved  
This material may not be reproduced without permission

# Introduction to Time-Series Analysis

- A *time series* is a chronological sequence of observations on a particular variable. Usually the observations are taken at regular intervals (days, months, years).
- A time-series function has the form:  $Y_t = f(Y_{t-1}, Y_{t-2}, Y_{t-3}, \dots, Y_{t-n}) + e_t$   
Where  $Y_t$  is the value of  $Y$  at time  $t$ , and  $Y_{t-n}$  is the value of  $Y$   $n$  periods earlier.
- Time series have two components, (a) a cyclical pattern that repeats at regular intervals, and (b) a trend that increases or decreases the overall value over time. DTREG models both components.
- All of the function types that can handle continuous target values such as neural networks, gene expression programming, TreeBoost, etc., can be used as the function for a time-series model.
- DTREG presents several tables and graphs to assist in evaluating time-series models.

# Creating a Time-Series Model

- When you start creating a new model, DTREG will present you with a screen where you can select a normal predictive model or a time-series model. You also can select which variables to use and the number and type of lag values.

The screenshot shows the DTREG software interface for creating a time-series model. The interface is divided into several sections:

- Time series or normal predictive model:** Two radio buttons are present. "Generate a normal predictive model" is unselected, and "Generate a time series forecasting model" is selected.
- Type of model to build:** A dropdown menu is set to "Multilayer Perceptron".
- Range of lag values to generate:** Two input fields are shown: "Minimum lag: 1" and "Maximum lag: 12".
- Automatic removal of trend:** Three radio buttons are present: "None" (unselected), "Linear" (unselected), and "Automatic" (selected). There is also a checkbox for "Stabilize variance" which is unselected.
- Lag, moving average and other generated variables:** A table with columns: Variable, Lag, SMA, LMA, EMA, Delta, LTrend, and Slope. The first row is for "Passengers", with checkboxes for Lag (checked), SMA, LMA, EMA, Delta, LTrend, and Slope.
- Validation of forward predictions:** Two checkboxes are present: "Validate predictions for end of series" (checked) and "Print validation values and forecasts" (checked). There is also an input field for "Number of values to use for validation: 12".
- Forecast future values:** Two checkboxes are present: "Forecast future values beyond end of series" (checked) and "Print future forecast values" (checked). There is also an input field for "Number of values to forecast: 12" and a checkbox for "Write forecast to file" which is unselected.

# Select Variables

- Select the target variable whose values are being forecast, and select the predictor variables which usually include lagged values of the target variable.

Variables

Variable	Target	Predictor	Weight	Categorical	Character
Passengers	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Passengers_Lag_01	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Passengers_Lag_02	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Passengers_Lag_03	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Passengers_Lag_04	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Passengers_Lag_05	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Passengers_Lag_06	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Passengers_Lag_07	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Passengers_Lag_08	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Passengers_Lag_09	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Passengers_Lag_10	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Passengers_Lag_11	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>
Passengers_Lag_12	<input type="checkbox"/>	<input checked="" type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>	<input type="checkbox"/>

Predictor range

All predictors

Predictor coverage

Type range

All categorical

All continuous

All numeric

All character

All reset

Search

Report options

- ☒ Report summary of variables
- ☐ Report category statistics for categorical variables
- ☐ Report category statistics for continuous variables
- ☐ Report Min., Max., Mean for continuous variables

Surrogate variables for missing value imputation

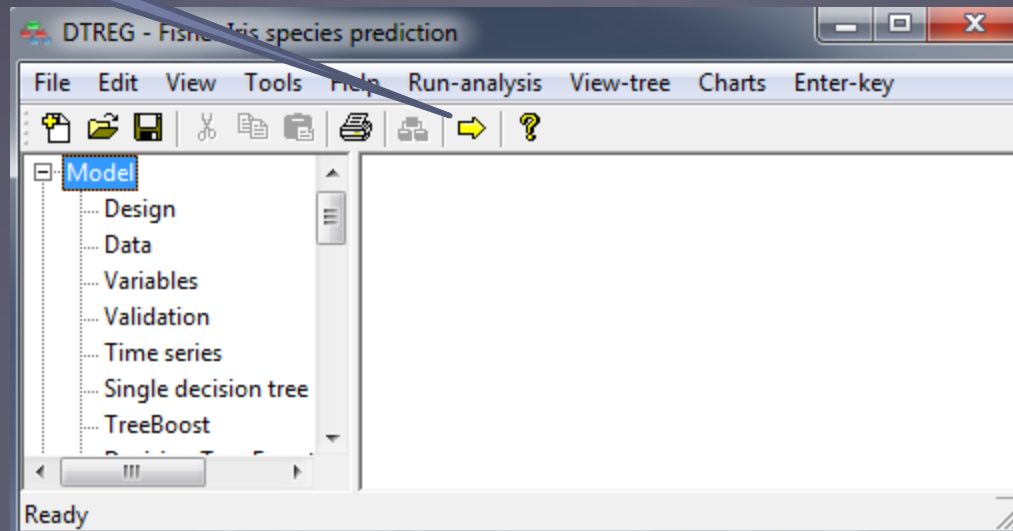
Number of surrogates to store: 5 Max. polynomial order: 1

Minimum surrogate association: 60 ☒ Report surrogate variables

# Tell DTREG to Train the Model

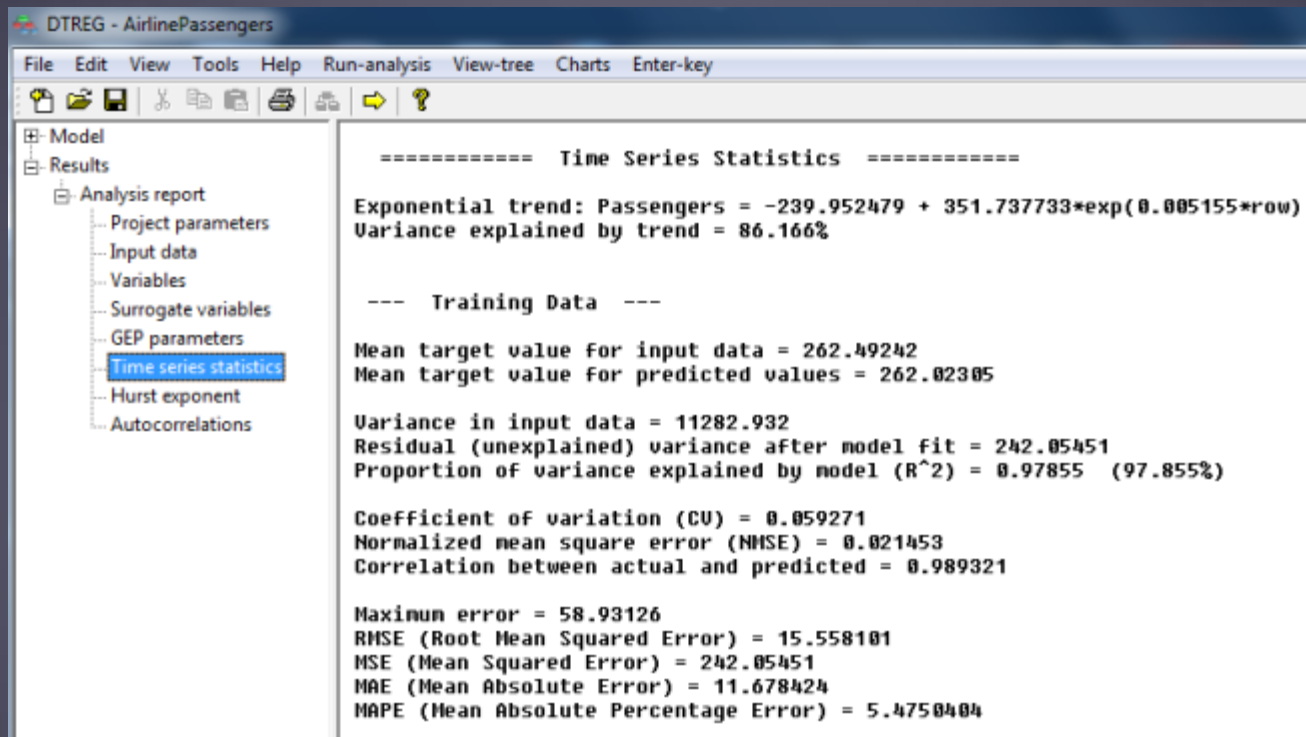
- Click the ➡ icon to start training the model.

Click to start  
training



# Examine Primary Time-Series Statistics

- Examine the primary time-series statics in the analysis report.
- The first section provides information about the trend it found.
- The next two sections provide quality of fit measures for the training and validation data.

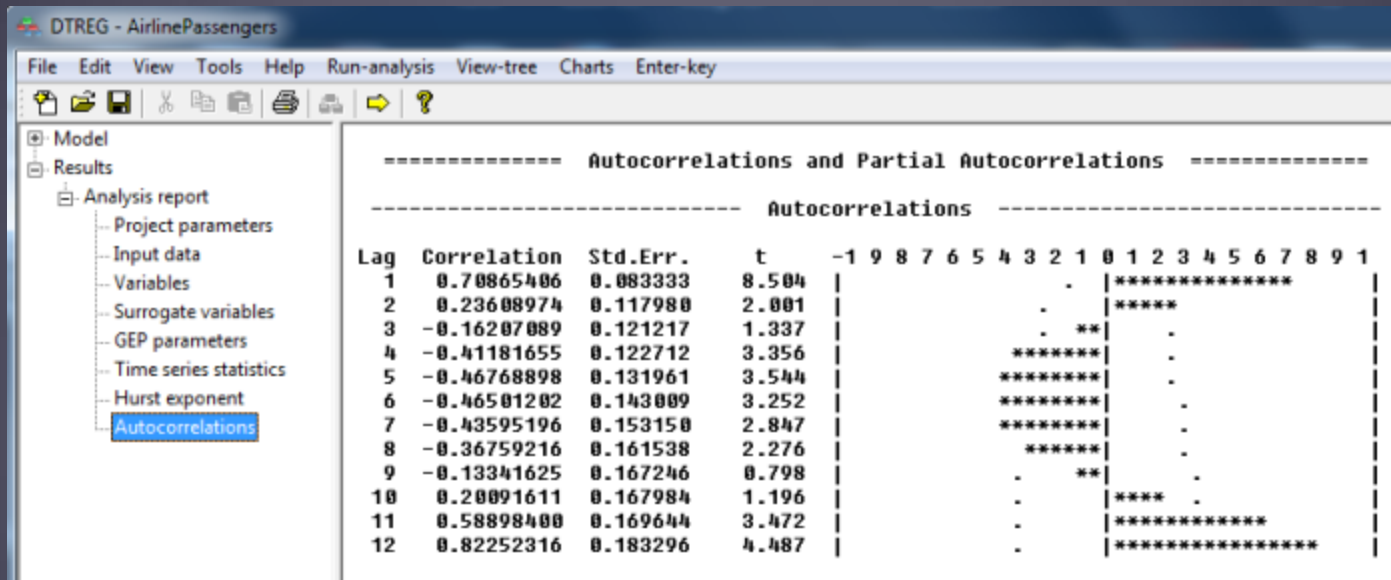


The screenshot shows the DTREG - AirlinePassengers application window. The left sidebar displays a tree view with 'Model' and 'Results' folders. Under 'Results', the 'Analysis report' folder is expanded, showing a list of report sections: 'Project parameters', 'Input data', 'Variables', 'Surrogate variables', 'GEP parameters', 'Time series statistics' (which is highlighted with a blue selection bar), 'Hurst exponent', and 'Autocorrelations'. The main window area displays the content of the 'Time series statistics' report. The report is titled '==== Time Series Statistics =====' and contains the following information:

```
==== Time Series Statistics =====  
  
Exponential trend: Passengers = -239.952479 + 351.737733*exp(0.005155*row)  
Variance explained by trend = 86.166%  
  
--- Training Data ---  
  
Mean target value for input data = 262.49242  
Mean target value for predicted values = 262.02305  
  
Variance in input data = 11282.932  
Residual (unexplained) variance after model fit = 242.05451  
Proportion of variance explained by model (R^2) = 0.97855 (97.855%)  
  
Coefficient of variation (CV) = 0.059271  
Normalized mean square error (NMSE) = 0.021453  
Correlation between actual and predicted = 0.989321  
  
Maximum error = 58.93126  
RMSE (Root Mean Squared Error) = 15.558101  
MSE (Mean Squared Error) = 242.05451  
MAE (Mean Absolute Error) = 11.678424  
MAPE (Mean Absolute Percentage Error) = 5.4750404
```

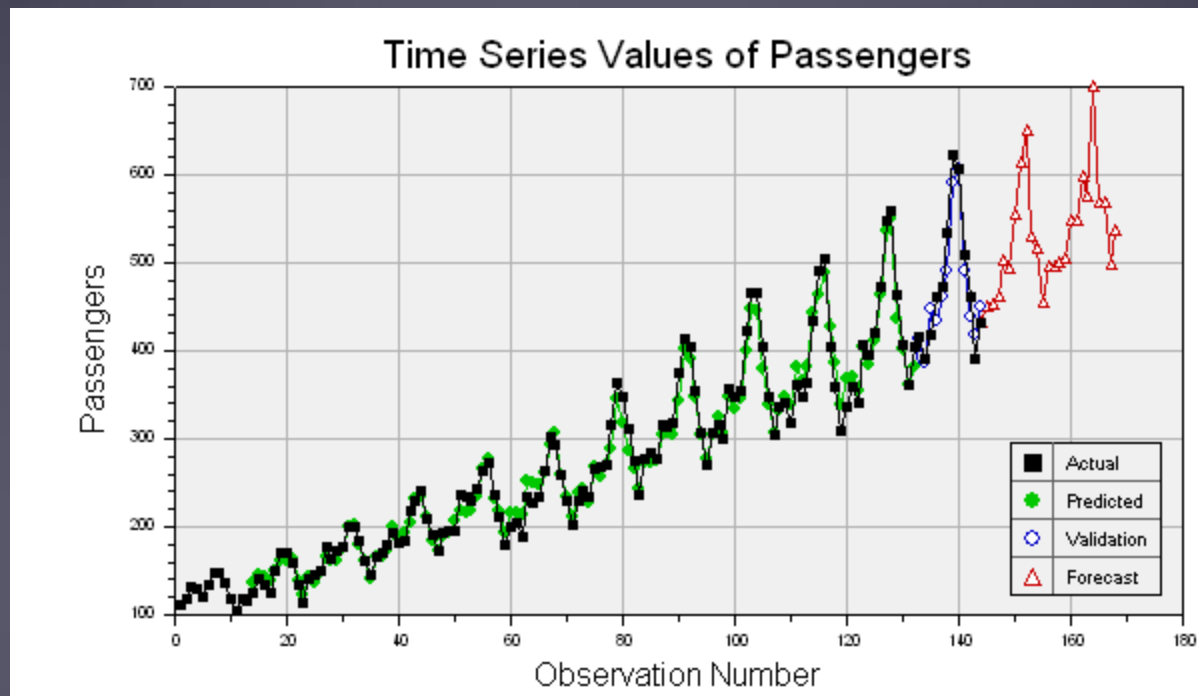
# Examine the Autocorrelations

- The Autocorrelation Table shows relative correlation between the target variable and lagged values of the target variable.
- Depending on the phase of a period within a cycle, the autocorrelation value may be positive or negative.



# Time-Series Chart

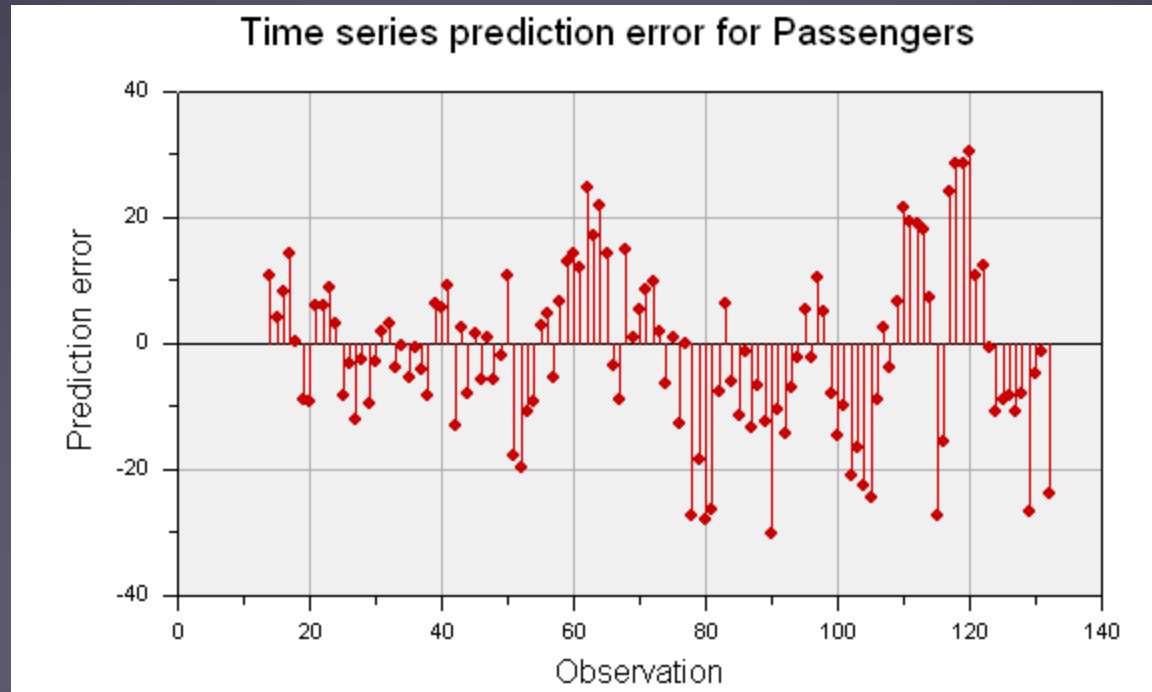
The Time-Series Chart shows the actual values, the predicted values and the forecast future values.





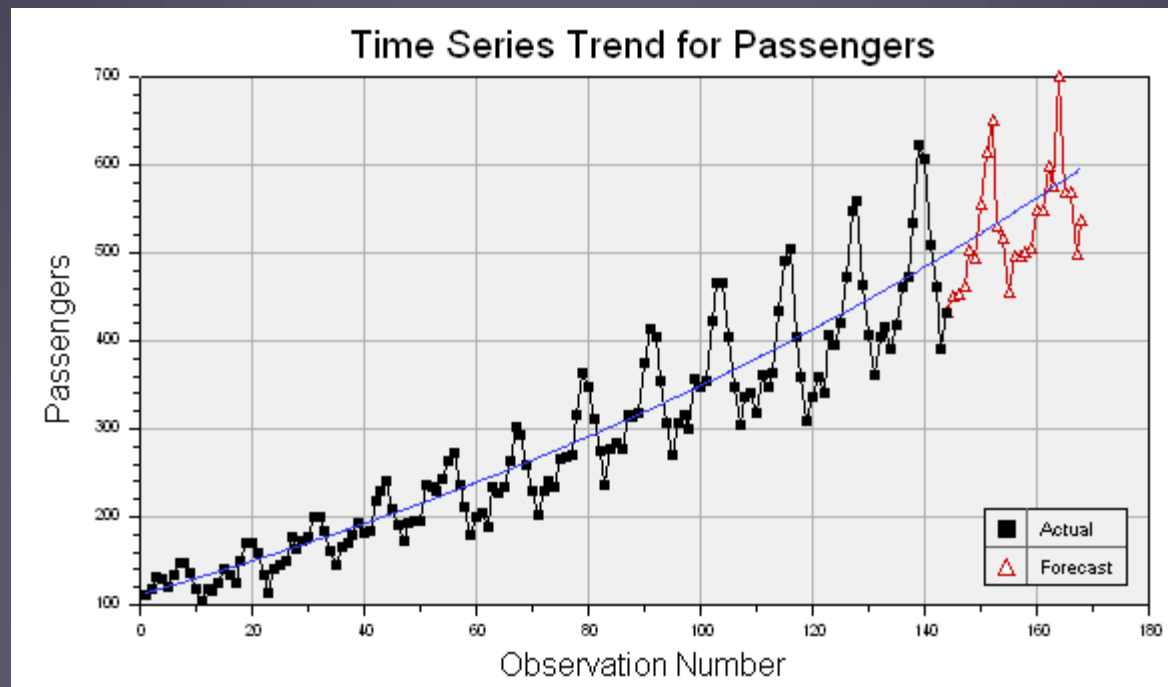
# Time-Series Residual Chart

The Time-Series Residual Chart shows the residual (error) between the predicted values and the actual values over time.



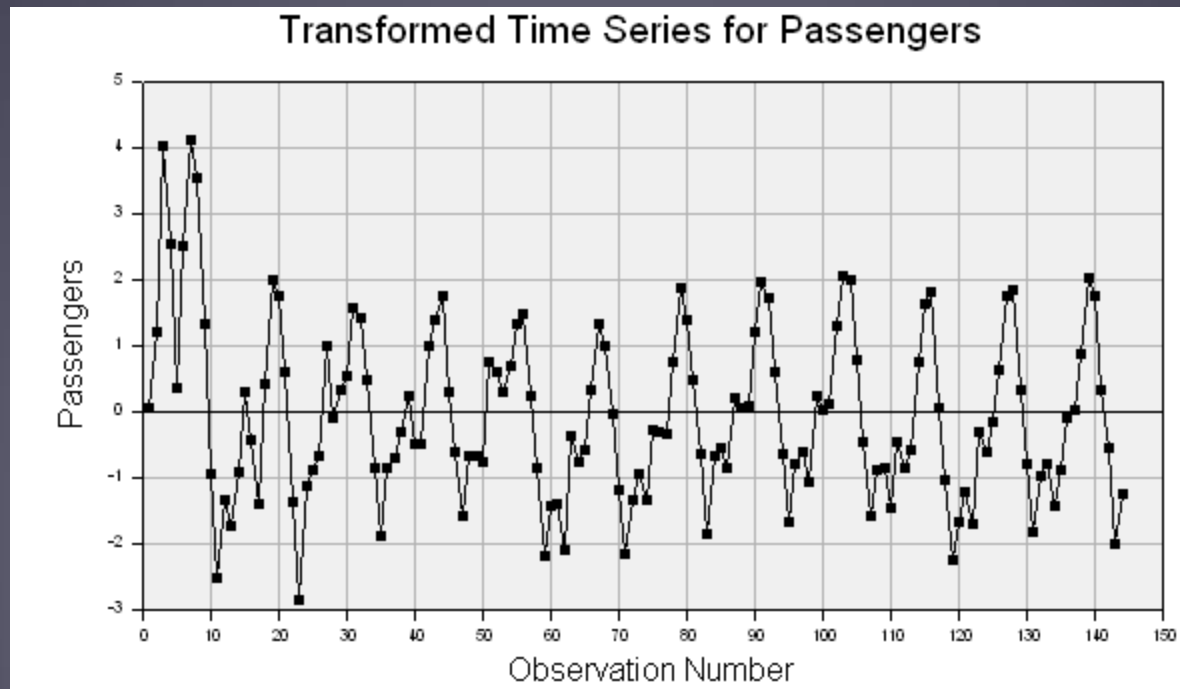
# Time-Series Trend Chart

The Time-Series Trend Chart shows the computed trend line fitted to a set of time series points. When performing a time-series analysis, DTREG computes the trend and subtracts it from the data points before looking for cyclical patterns.



# Time-Series Transformed Chart

The Time-Series Transformed Chart shows the time-series data points after the trend line has been subtracted. When performing a time-series analysis, DTREG computes the trend and subtracts it from the data points before looking for cyclical patterns.



# End of Tutorial

- This completes the Time-Series DTREG training tutorial