

An Introduction to AutoRegressive Integrated Moving Average (ARIMA) Models

By Paul Bakker
info@cathexisconsulting.ca



124 Merton St., Suite 502
Toronto, Ontario M4S 2Z2
Telephone: (416) 469-9954
Fax: (416) 469-8487
E-mail: info@cathexisconsulting.ca



Introduction

- ARIMA models can help you determine attribution with high level of confidence
- We'll cover:
 - When to use ARIMA
 - Long vs. short time series designs
 - ARIMA vs. linear regression
 - Weaknesses, challenges, and going for gold

When to use ARIMA



When to use ARIMA

- 1. Random assignment is not possible or appropriate**
- 2. There is no appropriate comparison group**
- 3. Want to determine attribution with high level of confidence**



When to use ARIMA

Example:

- Cathexis posted a job posting on CES's website on January 22, 2009.
- We want to find out if the posting was an effective way of attracting candidates to Cathexis.

How do we measure the post's effectiveness?



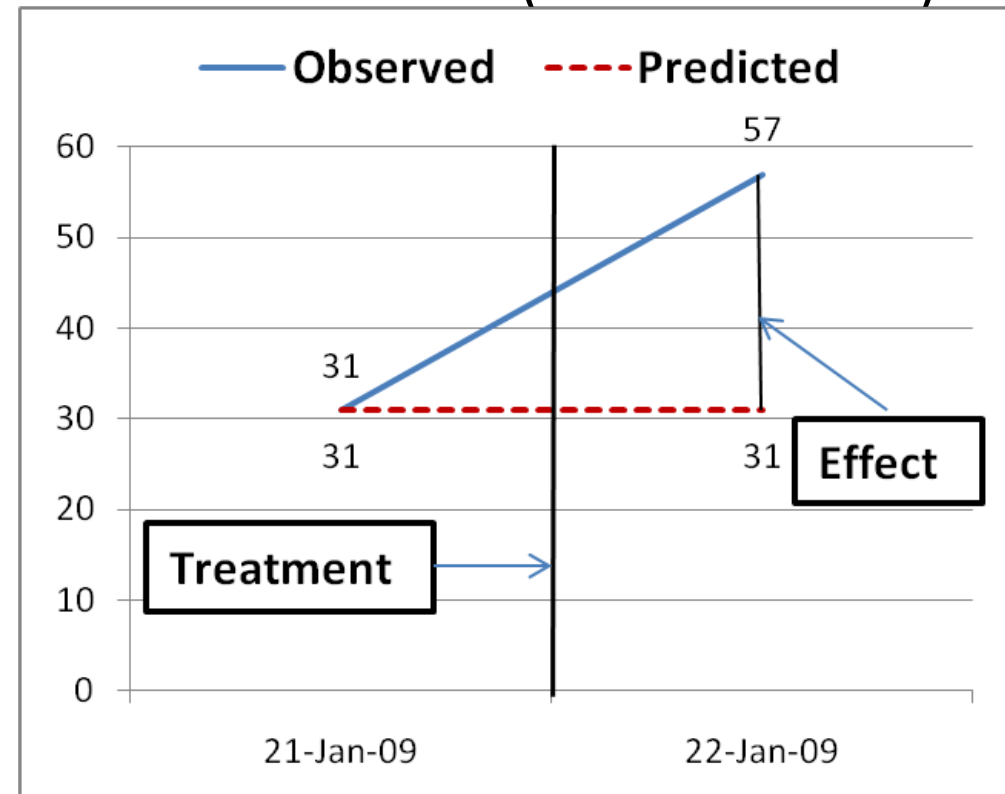
ARIMA's Improvements on Short Pre- Post Designs

ARIMA's Improvements on Short Pre-Post Designs

The Issue of Maturation (a series normal trend)

Were visits normally increasing or decreasing before CES post?

Basic Pre-Post (2 observations)

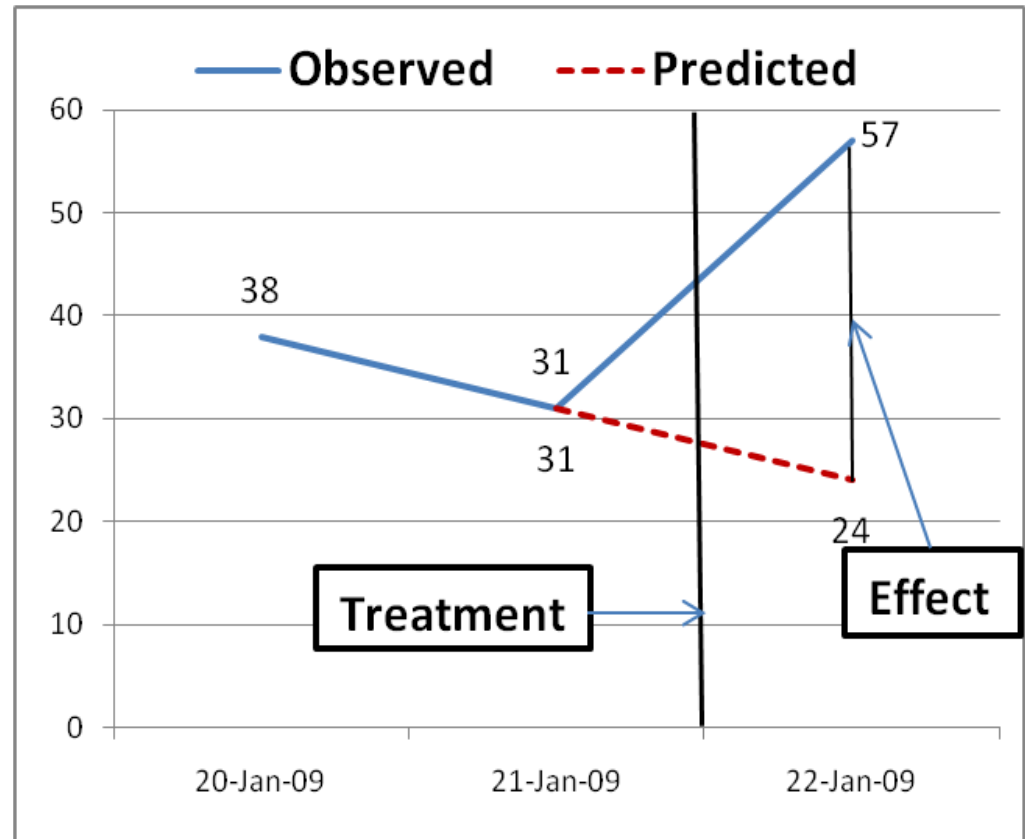


ARIMA's Improvements on Short Pre-Post Designs

Begin to Consider Trend

But Periodic Changes in Trend May be Normal

1 Additional Pre-Observation

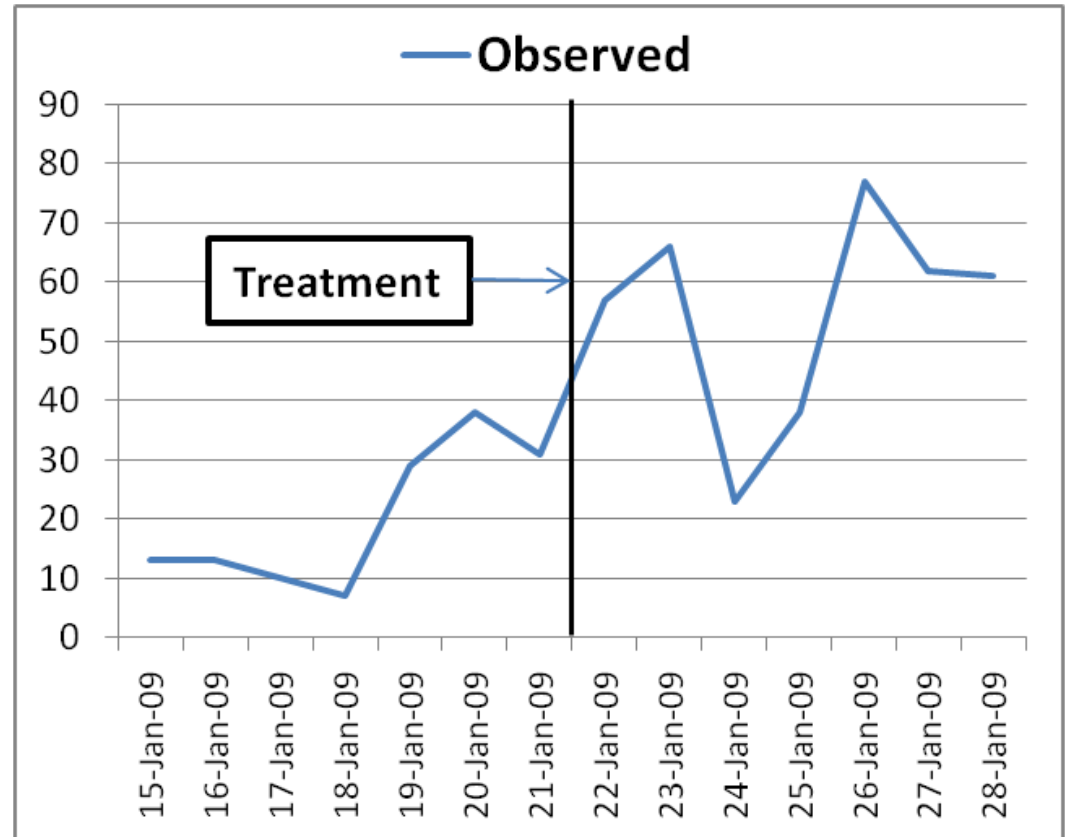


ARIMA's Improvements on Short Pre-Post Designs

Clear idea of trend and possible effect

Q's remain: Dip normal? Do pre-post trend continue?

1 Week Pre & 1 Week Post

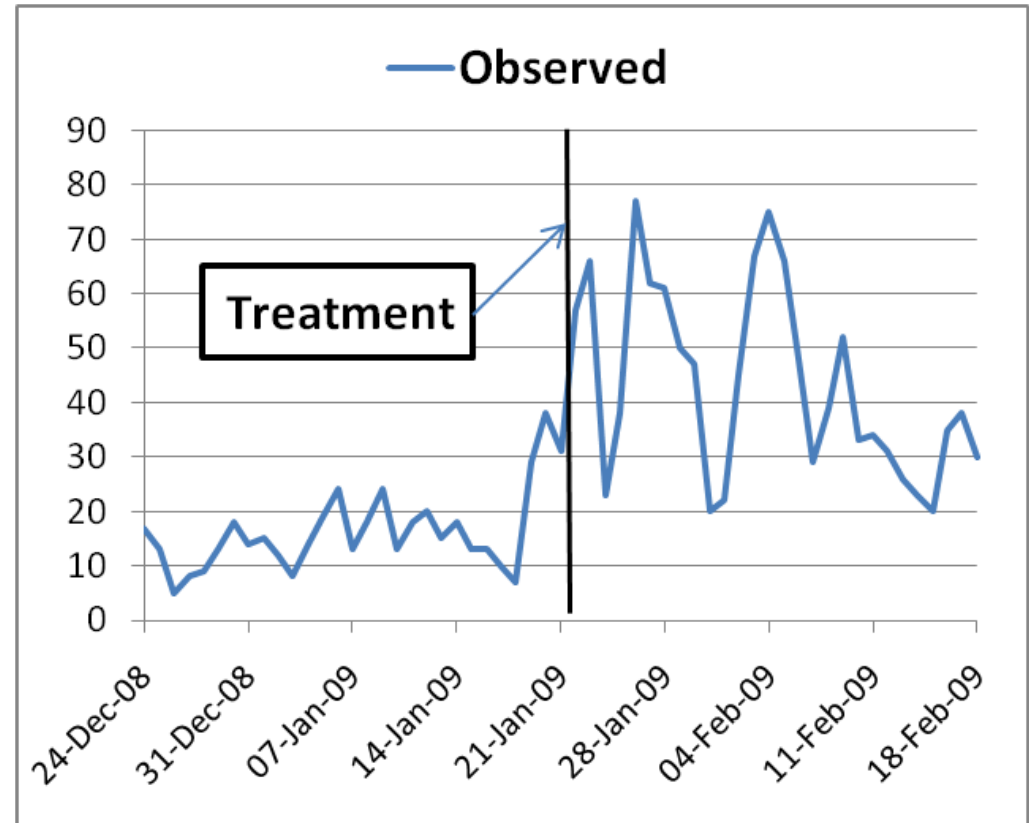


ARIMA's Improvements on Short Pre-Post Designs

Clear difference between pre and post trends

can take 100 observations to properly model all patterns in a series

1 Month Pre & 1 Month Post



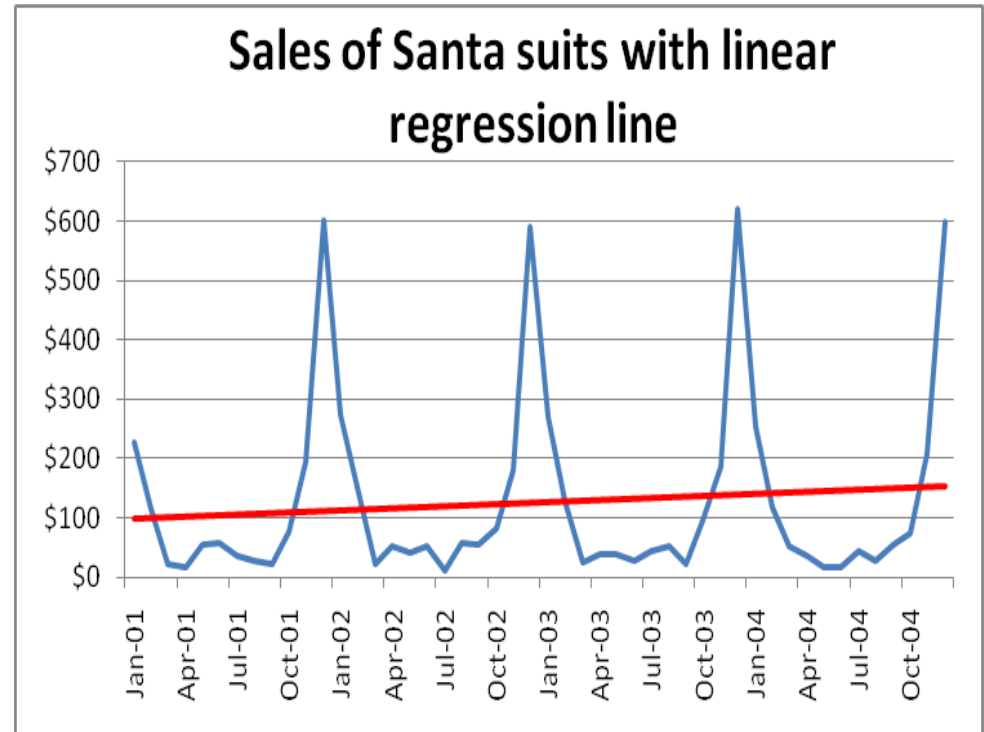
ARIMA's Improvements on Linear Regression

ARIMA's Improvements on Linear Regression

The Problem of Autocorrelation

Example: Outliers

- Sales in Dec. > 2 SD from the regression line.
- Considered outliers.
- But normal part of series's trend.





Weaknesses, Challenges, and Going for Gold



Weaknesses, Challenges, and Going for Gold

Threats to Internal Validity

- Major Threat is **History**
- **Instrumentation** (can be ruled out)
- **Selection**

Threats can be limited to gain strong confidence in attribution

- Consider our CES posting example



Weaknesses, Challenges, and Going for Gold

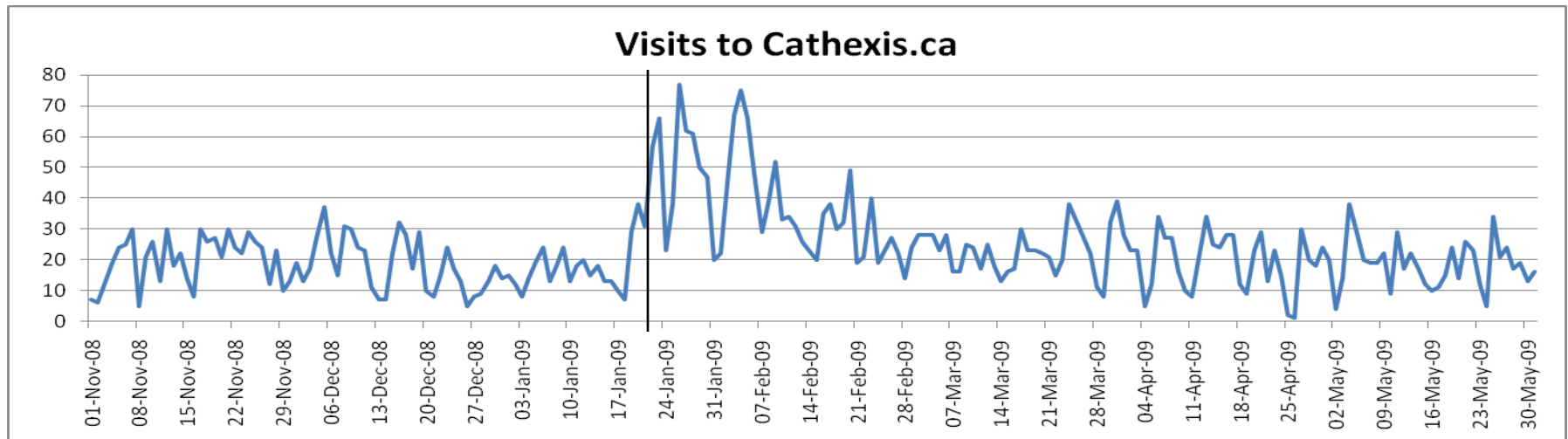
Practice challenges:

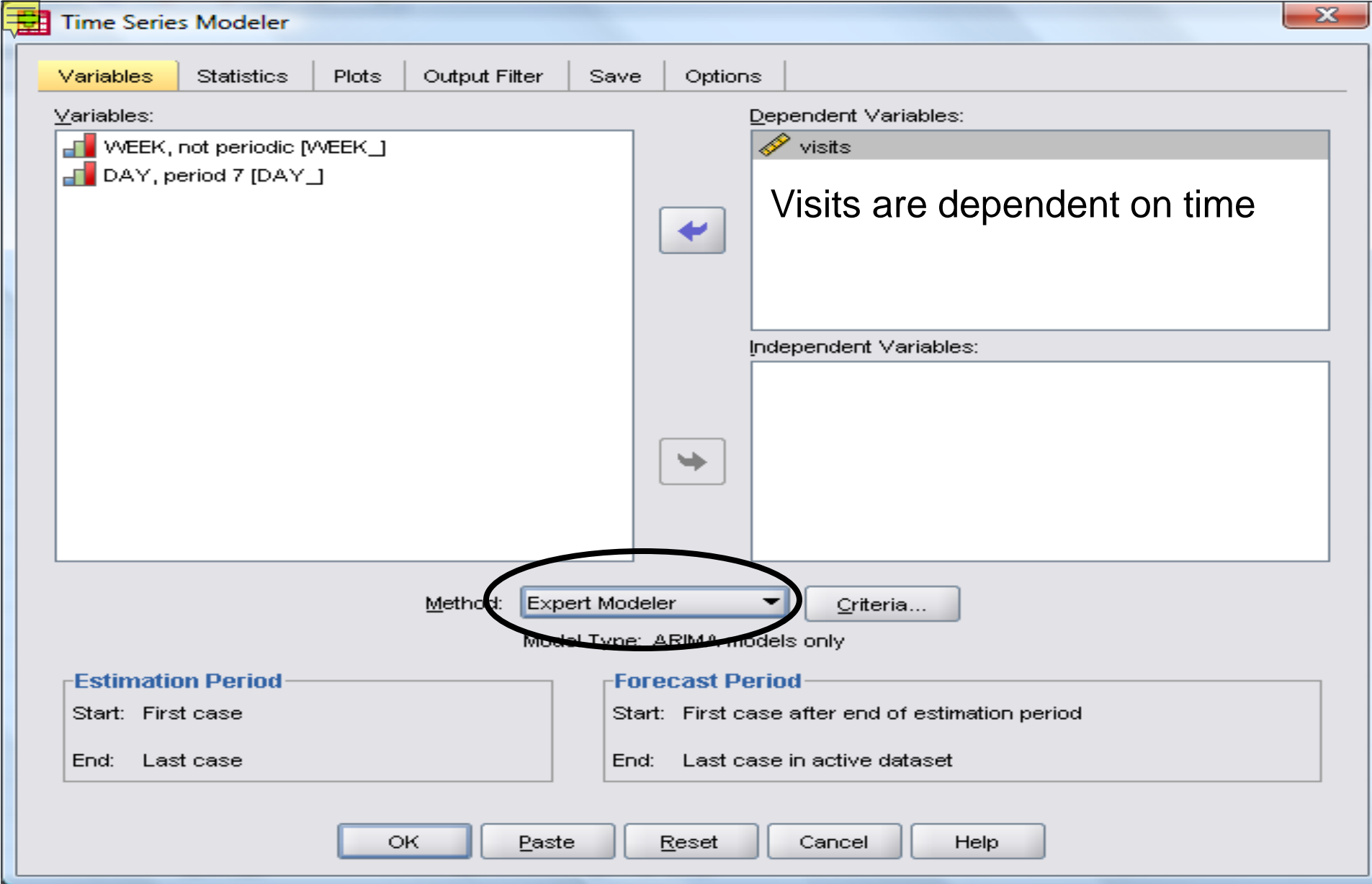
- Collecting all that data!!!!
- Ensuring timely implementation.
- Ability to conduct ARIMA analysis

ARIMA Analysis Using SPSS's Expert Modeler.

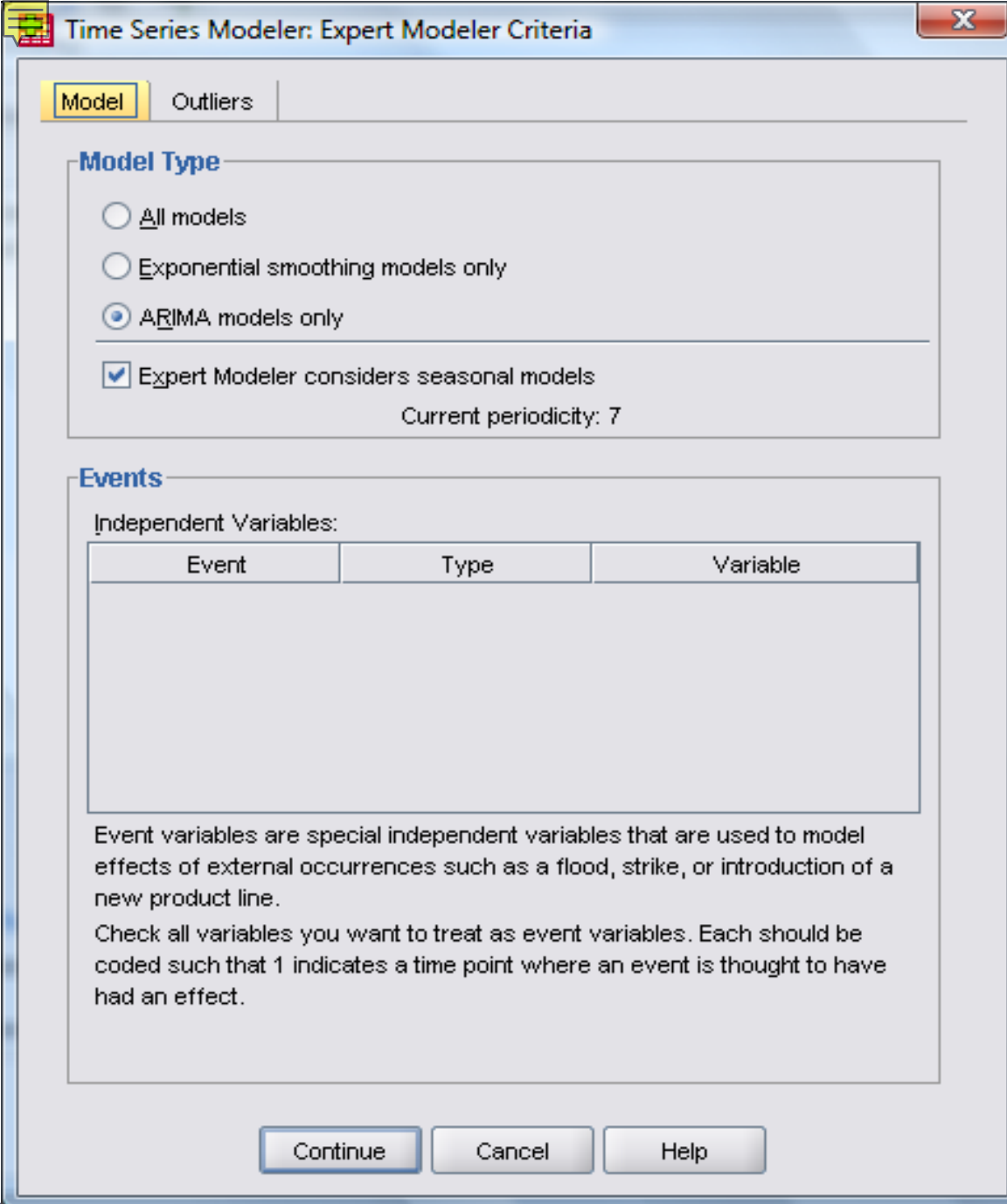
Back to our website example

An ARIMA model will allow to know when and by how much visits increased with given level (95%) of confidence.

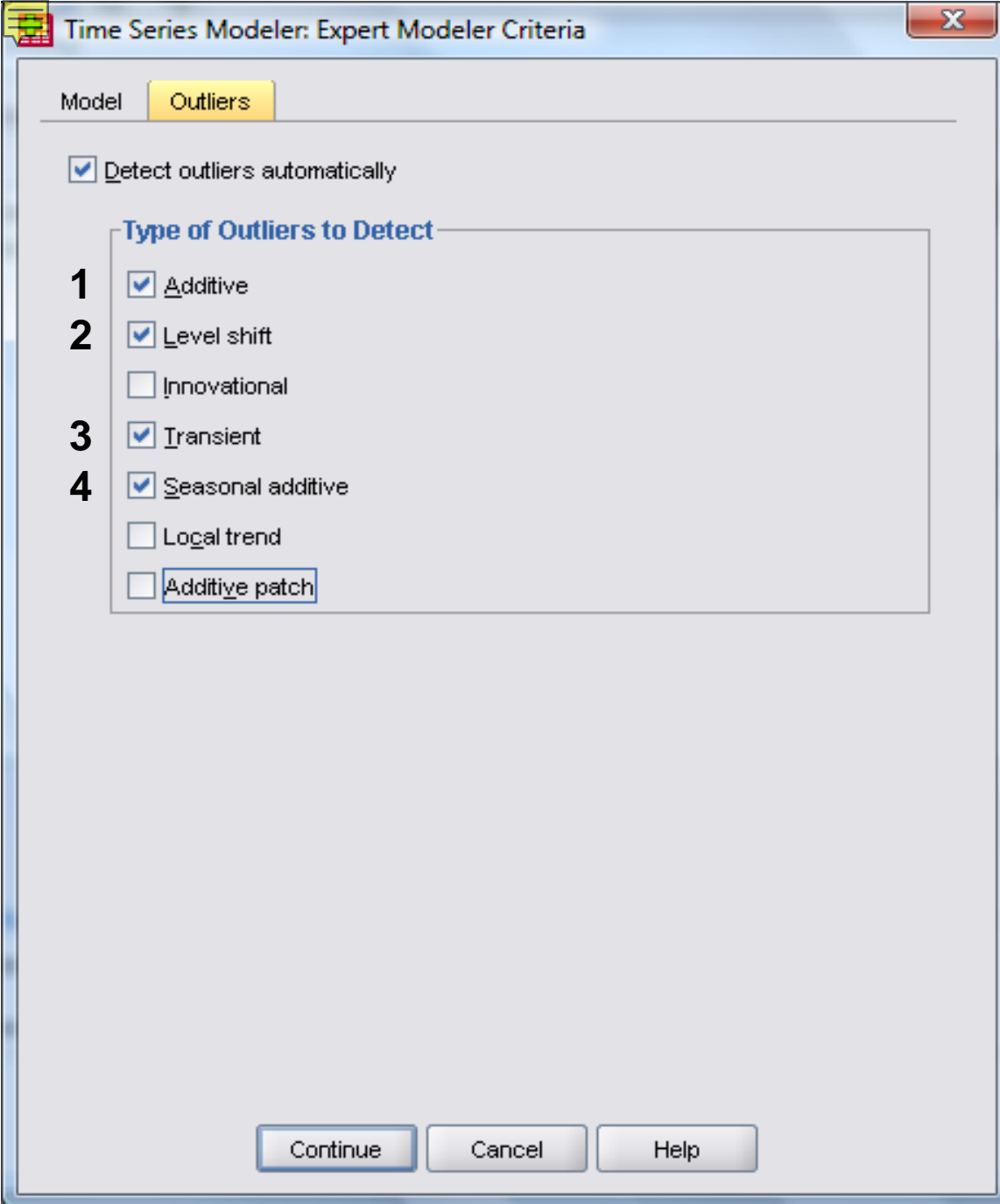




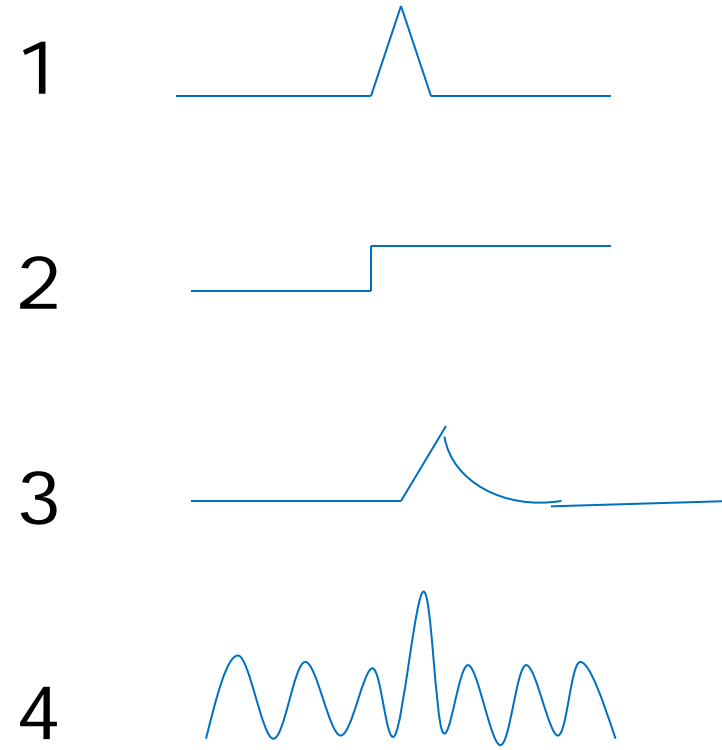
- Use SPSS's Expert Modeler



- Expert Modeler to only consider ARIMA.
- Expert Modeler to consider seasonality



- Automatically detect outliers.
- I.e. significant changes in the series.





ARIMA Analysis Using SPSS's Expert Modeler.

Overall Fit of the model

Stationary R-squared: The higher the better.

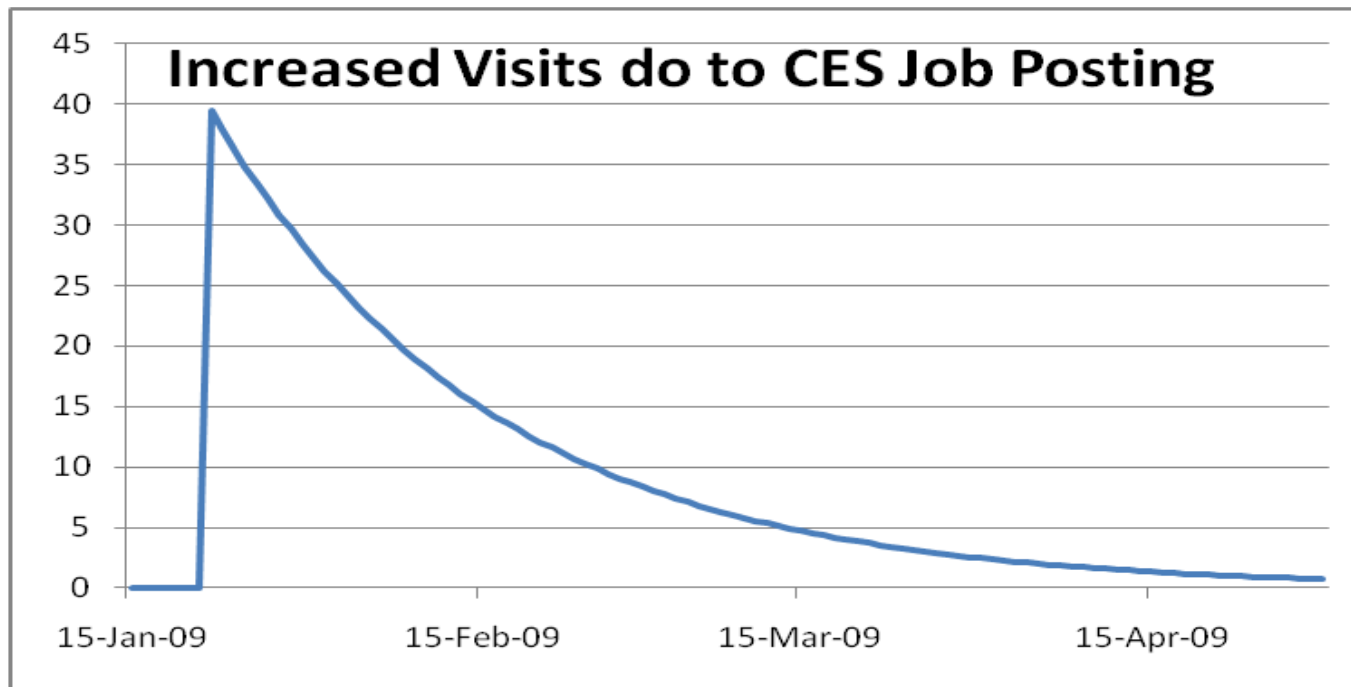
Ljung-Box Q: If non-significant then model's error is random.

For our example: Stationary R-squared = .695 &
Ljung-Box Q = .162

So far, a good model

ARIMA Analysis Using SPSS's Expert Modeler.

The model determines Jan. 22 (day of CES post) to be a transient outlier with a magnitude of 39.4 with a decay factor of 96%



Summary

- We'll covered:
 - When to use ARIMA
 - Long vs. short time series designs
 - ARIMA vs. linear regression
 - Weaknesses, challenges, and going for gold
- ARIMA models make it is possible to make claims of attribution with a high level of confidence without random assignment or comparison groups