COOLSTATSBLOG

Economics and statistics in plain English

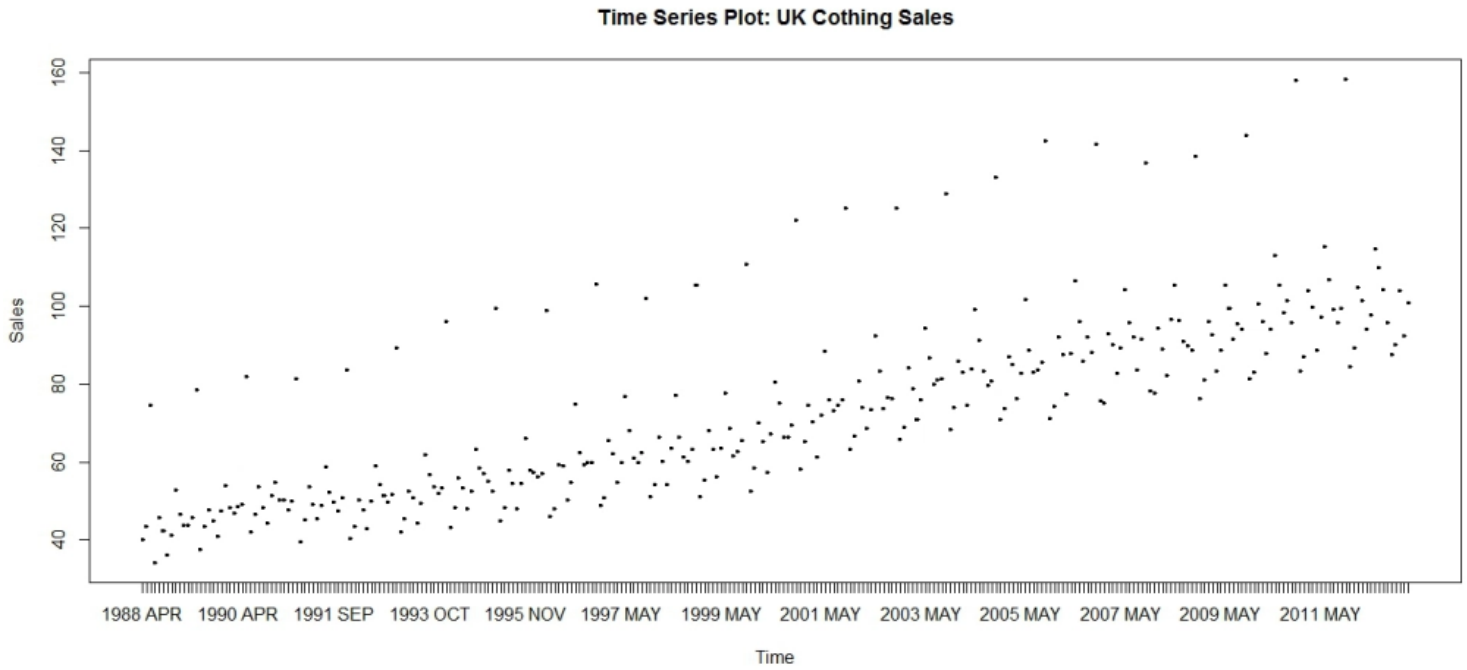
HOW TO USE AUTOCORRELATION FUNCTION (ACF) TO DETERMINE SEASONALITY

Posted on August 11, 2013 by Abbas Keshvani under Time Series

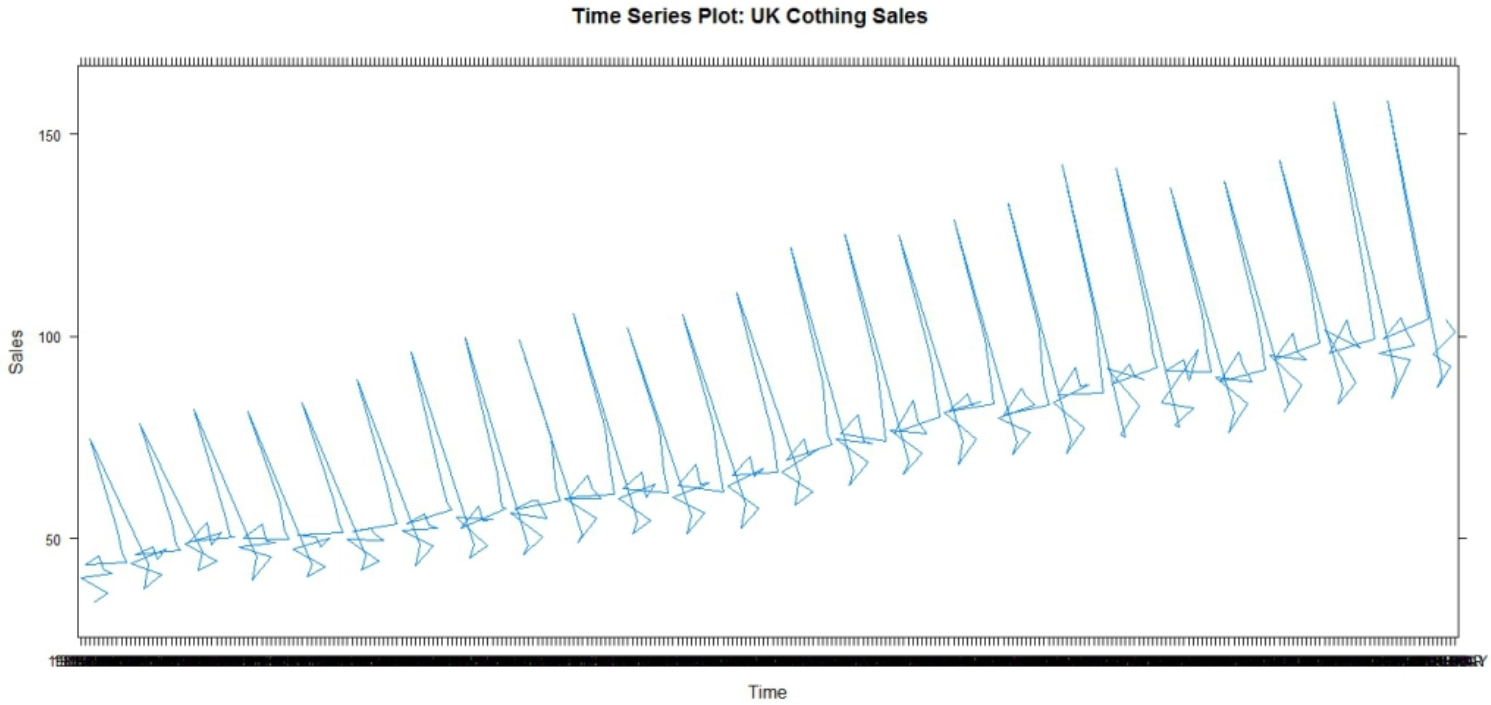
In my previous post, I wrote about using the autocorrelation function (ACF) to determine if a timeseries is stationary. Now, let us use the ACF to determine **seasonality**. This is a relatively straightforward procedure.

Firstly, seasonality in a timeseries refers to predictable and recurring trends and patterns over a period of time, normally a year. An example of a seasonal timeseries is retail data, which sees spikes in sales during holiday seasons like Christmas. Another seasonal timeseries is box office data, which sees a spike in sales of movie tickets over the summer season. Yet another example is sales of Hallmark cards, which spike in February for Valentine's Day.

The below graphs show sales of clothing in the UK, and how these sales follow seasonal trends, spiking in the holiday season:



Clothing Sales in the UK



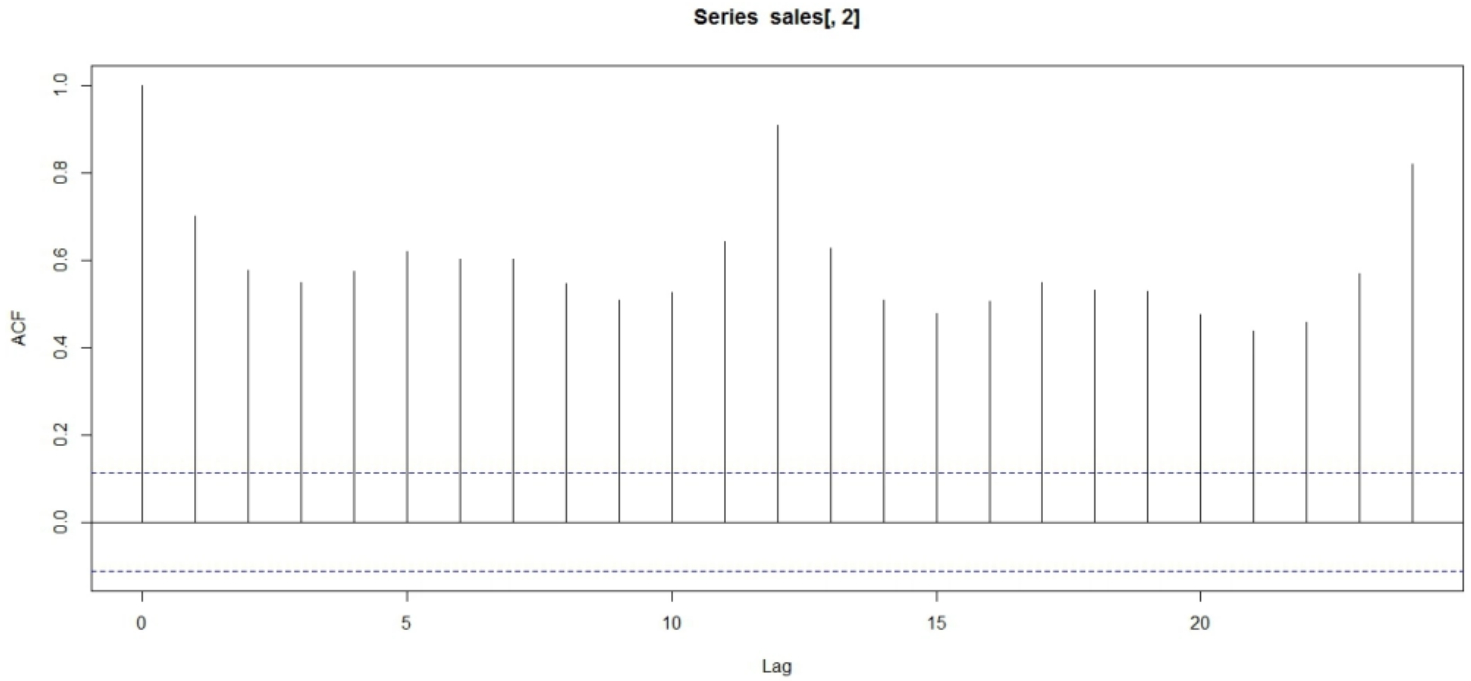
Clothing Sales in the UK: line graph

Note the spikes in sales, which obediently occur every December, in time for Christmas. This is evident in the trail of December plot points (Graph 1), which hover significantly above the sales data for other months, and also in the actual spikes of the line graph (Graph 2).

The above is a simple example of a seasonal timeseries. However, timeseries are not always simply seasonal. For example, a SARMA process comprises of seasonal, autoregressive, and moving average components, hence the acronym. This will not look as obviously seasonal, as the AR and MA processes may overlap with the

seasonal process. Thus, a simple timeseries plot, as shown above, will not allow us to appreciate and identify the seasonal element in the series.

Thus, it may be advisable to use an autocorrelation function to determine seasonality. In the case of seasonality, we will observe an ACF as below:



ACF of UK clothing sales data

Note that the ACF shows an oscillation, indicative of a seasonal series. Note the peaks occur at lags of 12 months, because April 2011 correlates with April 2012, and 24 months, because April 2011 correlates with April 2013, and so on.

The above analyses were conducted on R. Credits to data.gov.uk and the Office of National Statistics, UK for the data.

Abbas Keshvani

Related

How to use the Autocorreation Function (ACF)?
In "Time Series"

Daily/monthly/yearly tallies for your data
In "R"

Using AIC to Test ARIMA Models
In "Time Series"

This entry was tagged ACF, Autocorrelation, Seasonality, Time series. Bookmark the permalink.

← How to use the Autocorreation Function (ACF)?

Using AIC to Test ARIMA Models →

11 thoughts on "How to Use Autocorrelation Function (ACF) to Determine Seasonality"

Anonymous

January 28, 2015 at 8:00 pm

Excellent commentary on seasonality using ACF.

Reply

Abbas Keshvani

February 8, 2015 at 6:53 am

Thank you!

Reply

Pingback: How to use the Autocorreation Function (ACF)? | CoolStatsBlog

Jay

March 31, 2016 at 11:36 am

Dear Abbas,

Would I please be able to use your ACF function plots for seasonality and stationarity for my dissertation? Just as a point of reference.

Thanks,
Jay

Reply

Abbas Keshvani

June 4, 2016 at 4:51 am

Sure

Reply

Hugo Barbier

June 20, 2016 at 8:13 am

Excellent explanations! But I wonder how you can get the information of the peiriod, here 12 month, without read it on the ACF graph? Can you help me? I have a time seires with an unknow period….

Reply

Abbas Keshvani

September 2, 2016 at 4:23 pm

Hi Hugo, look at your ACF. The period is the first spike that stands out.

I know my clothing data has a period of 12, since the ACF spikes at the 12th lag.

Reply

maz

June 22, 2016 at 12:26 am

Thank you…good explanation!!!

Reply

Abbas Keshvani

September 2, 2016 at 4:17 pm

Thanks!

Reply

Vikrant

September 7, 2016 at 10:52 am

Dear Abbas,

Thanks for explaining usage of ACF to find seasonality. I have two questions

1) If there exists no seasonality in data, how can it be identified? Since ACF will still give spikes may be with smaller range. What should be cutoff value or threshold value for considering the spike as significant from others. Also does it guarantee presence/ absence of seasonality.?

2) In real world, datasets would have noise and hence while predicting the period using ACF might take that as well into consideration, because it might affect in computing period.

Thanks and Regards,

-Vikrant

Reply

Abbas Keshvani

September 10, 2016 at 11:07 am

Thanks Vikrant.

1. An ACF spike at lags of 4 means that there is a strong correlation between every 4th point. It indicates seasonality. But if the correlation is strong for all points, then it means that the series is just non-stationary or random.

2. ACF tells you the correlation between every single point and its k-th lag. Noise cannot affect the entire data series to show a false correlation.

Reply

Leave a Reply

Enter your comment here...

JUST SO YOU KNOW

This blog was written in my personal capacity. The opinions expressed here are my own and do not reflect the views of my employer.

SYNDICATED FEED

R blog

Stats Blogs

SEARCH ON COOLSTATSBLOG

Search …

Search