



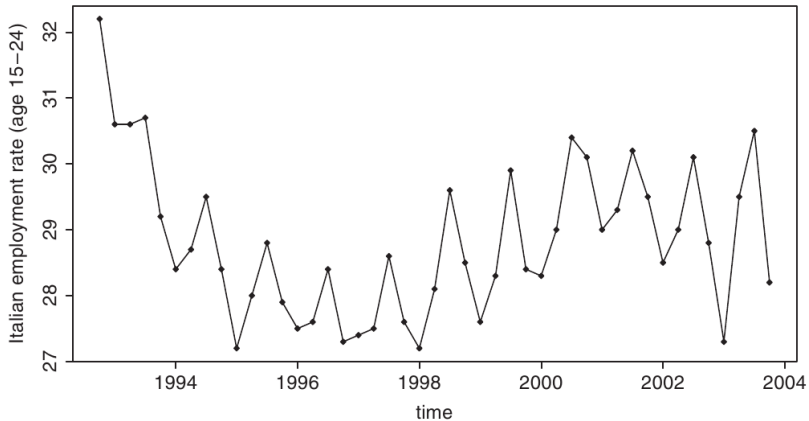
Aula 2

Séries Temporais - Parte I - Definições, Avaliação e Componentes

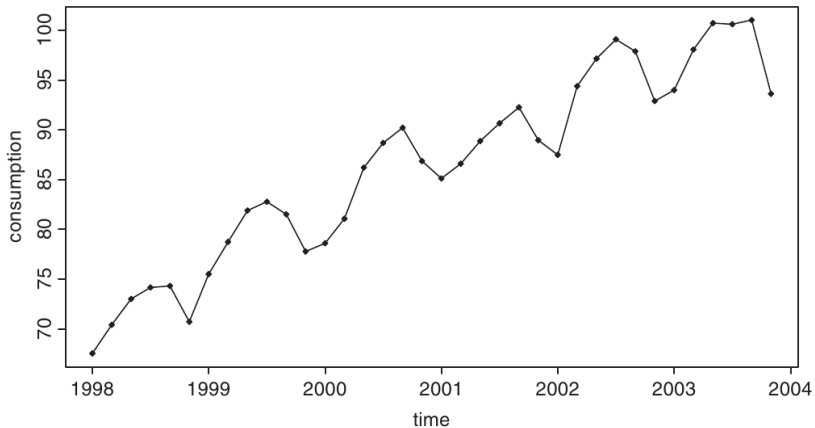
Introdução

- Série Temporal é um conjunto de observações sobre uma variável alvo (ou de interesse) ordenadas no tempo
- Os períodos de tempo de uma série temporal são regulares, isto é, as unidades de tempo das observações variam em períodos uniformes
- **Exs:**
 - De hora em hora
 - De minuto em minuto
 - De ano em ano
 - ...

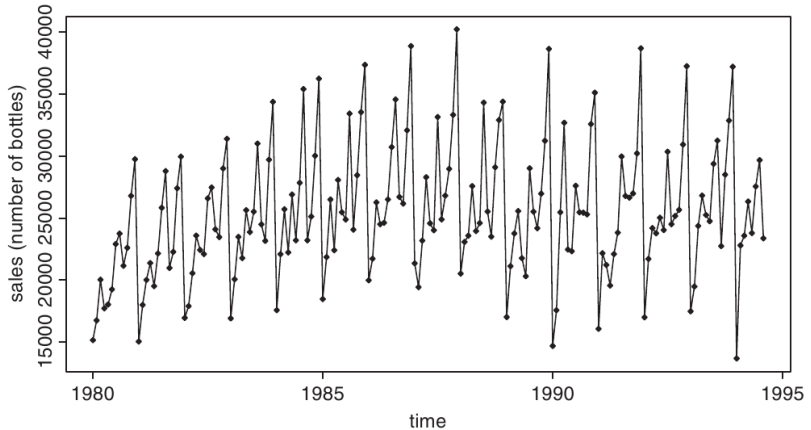
Introdução



Introdução



Introdução



Introdução

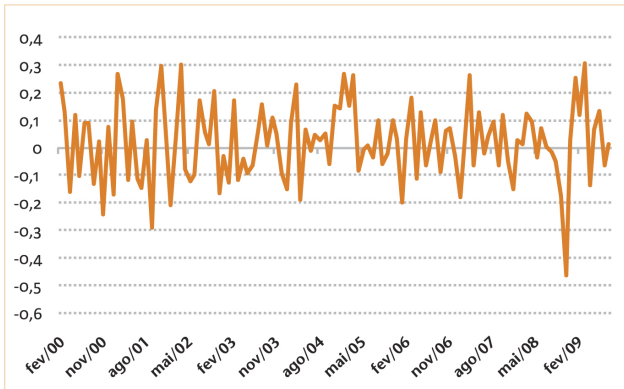


Figura 3. Variação das vendas de automóveis no Brasil

http://www.ilos.com.br/web/wp-content/uploads/2011/02/2011_02_imagem_04.jpg

Introdução

- Vários conjuntos de dados do mundo real são registrados em um formato de série temporal
- Exs:
 - Variação da temperatura a cada hora do dia
 - Vendas mensais de uma empresa
 - Valores de fechamento diários do IBOVESPA
 - Resultado de um eletroencefalograma
 - ...

Introdução



Introdução

- Uma das **principais tarefas** quando se tem dados representados em séries temporais é a **tarefa preditiva**, isto é, **prever o valor futuro (regressão)** ou **prever se haverá uma alta ou uma baixa** no valor futuro em relação ao valor atual (**classificação**)
- O foco da disciplina será na **previsão do valor futuro**
- Exemplos de aplicações com predição de valor futuro
 - Previsão da quantidade de vendas de produtos
 - Indicadores de valores de ações
 - Aumento da temperatura ou devastação em ecossistemas
 - Previsão de demanda de energia elétrica
 - ...

Introdução

Para que se possa prever o valor futuro, é preciso **identificar padrões** não aleatórios **da variável de interesse** na série temporal, para que a observação do **comportamento passado** (padrões) possam permitir fazer **previsões sobre o futuro**, **orientando a tomada de decisões**

Definição

- Uma série temporal é uma sequência $\{y_t\}$ de valores que podem ser mensurados períodos de tempo específicos t
- Normalmente os valores de t são discretos e correspondem a sequência naturais e uniformes no horizonte de tempo
 - Minutos
 - Horas
 - Dias
 - Semanas
 - Meses
 - Anos
 - ...

Modelos Preditivos

- Os modelos preditivos analisam as observações de uma série temporal $\{y_t\}$ para descobrir padrões regulares que ocorreram no passado de forma a prever o provável comportamento da série no futuro
- Suponha que t é o índice do período de tempo atual (ou último período), e que k (tamanho da janela) correspondende aos últimos valores conhecidos ou últimos valores que devem ser utilizados para fazer a predição

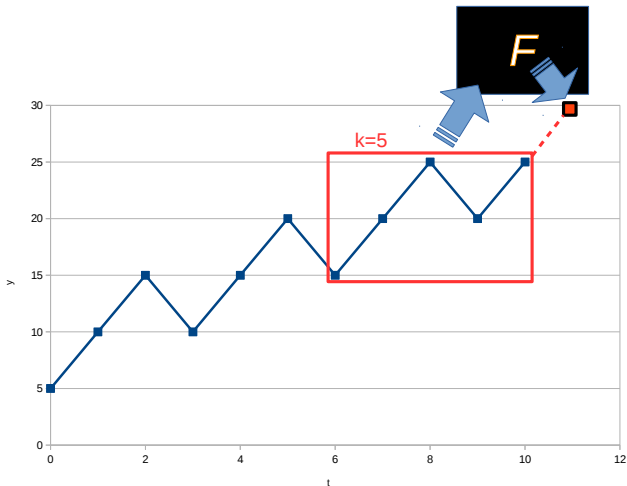
Modelos Preditivos

- Portanto, o objetivo de um modelo preditivo é

$$\hat{f}_{t+1} = F(y_t, y_{t-1}, \dots, y_{t-k+1})$$

- Para desenvolver um modelo preditivo, portanto, precisamos selecionar uma função F que melhor realiza as previsões dos valores futuros de acordo com a série temporal que queremos analisar

Modelos Preditivos

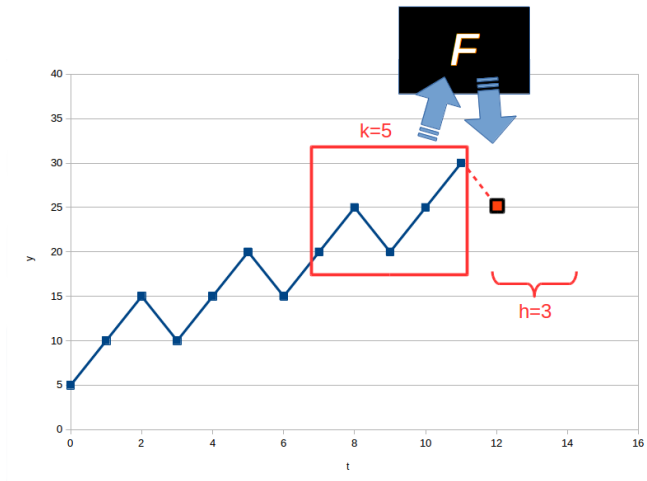


Modelos Preditivos

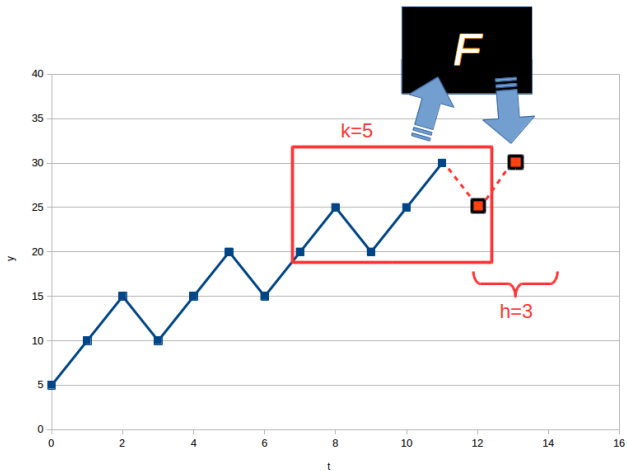
- Em algumas situações, pode-se querer gerar previsões para um período à frente do período $t + 1$
 - **Ex:** prever as vendas de um produto para um ano inteiro
- Para isso é feito um **processo recursivo de previsões** dos períodos anteriores até se obter a previsão do período $t + h$, na qual h é o período de tempo desejado e $h \geq 1$

$$\hat{f}_{t+h} = F(\hat{f}_{t+h-1}, \hat{f}_{t+h-2}, \dots, \hat{f}_{t+1}, y_t, t_{t-1}, \dots, y_{t-k+1})$$

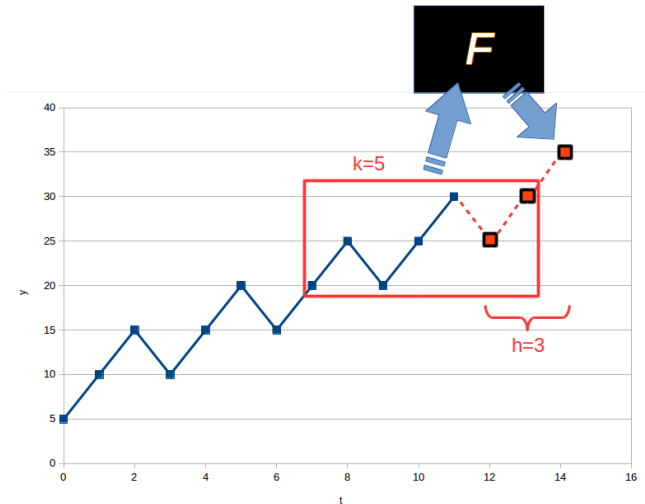
Modelos Preditivos



Modelos Preditivos



Modelos Preditivos

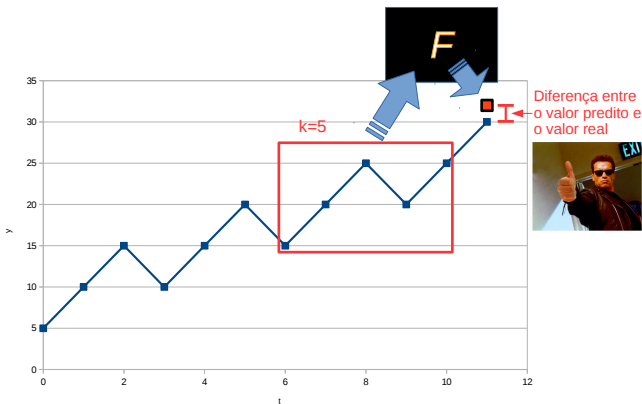


Avaliação de Modelos de Séries Temporais

Para escolher um modelo de previsão apropriado e para analisar o quão bem a previsão de valores irá se comportar na prática, **é necessário analisar o modelo de previsão**

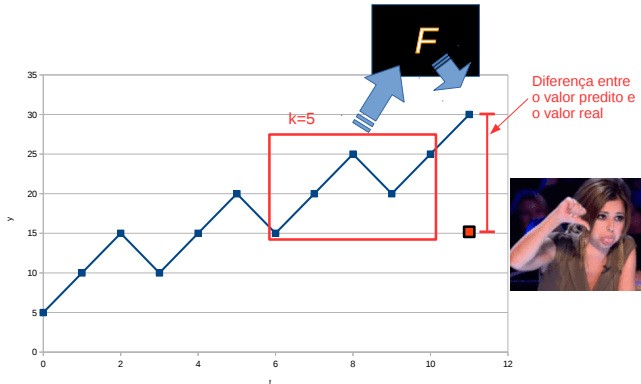
Avaliação de Modelos de Séries Temporais

- Intuitivamente, as medidas de avaliação irão avaliar o quanto o valor predito é diferente do valor real



Avaliação de Modelos de Séries Temporais

- Intuitivamente, as medidas de avaliação irão avaliar o quanto o valor predito é diferente do valor real



Erro de Predição

- Dadas observações y_t de uma série temporal e as previsões correspondentes f_t para os últimos k períodos de tempo, o **Erro de Predição** do período t , $t = 1, 2, \dots, k$ é definido como

$$e_t = y_t - f_t$$

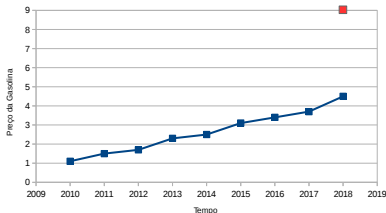
Erro de Predição Percentual

- Já o **Erro de Predição Percentual** no período t é dado por:

$$e_t^P = \frac{y_t - f_t}{y_t} \times 100$$

- A predição de erro percentual é independente da escala na qual as observações são medidas e, portanto, é mais “real”, especialmente para comparar os erro gerados em diferentes séries temporais

Erro de Predição



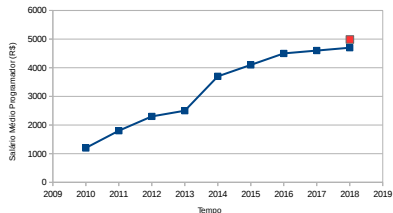
Legenda

■ Valor Real

■ Valor Predito

$$e_t^p = 4.5 - 9 = -4.5$$

$$e_t^p = \frac{4.5 - 9}{4.5} \times 100 = 100\%$$



$$e_t = 4700 - 5000 = -300$$

$$e_t^p = \frac{4700 - 5000}{4700} \times 100 = 6\%$$

Medidas de Distorção

- Índices de distorção são utilizados para considerar as k avaliações de um modelo
- **Erro Médio** (*Mean Error (ME)*):

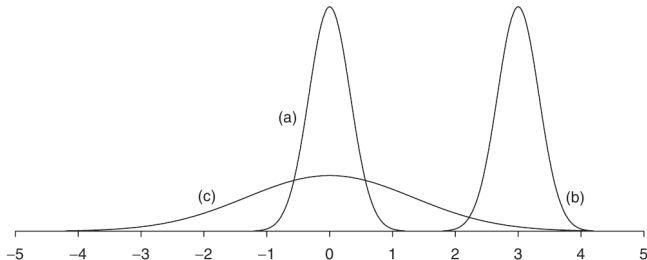
$$ME = \frac{\sum_{t=1}^k e_t}{k}$$

- **Erro Médio Percentual** (*Mean Percentage Error*):

$$MPE = \frac{\sum_{t=1}^k e_t^P}{k}$$

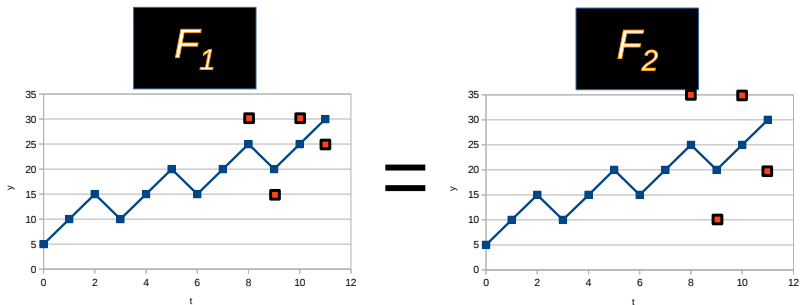
- Idealmente um modelo bom seria o modelo cujo ME ou MPE seja próximo a 0

Problemas Medidas Distorção



- O erro médio de **a** e **c** é 0 e o erro médio de **b** é 3
- Entretanto, apesar da distorção ser nula, os modelos **a** e **c** não são equivalentes, uma vez que a gerar erros menores comparados com **c** → **a** seria um modelo preferido

Problemas Medidas Distorção



Medidas de Dispersão

- Os problemas das medidas de distorção podem ser solucionados com medidas de dispersão
- Desvio Médio Absoluto** (*Mean Absolute Deviation - MAD*)

$$MAD = \frac{\sum_{t=1}^k |e_t|}{k} = \frac{\sum_{t=1}^k |y_t - f_t|}{k}$$

- Erro Percentual Absoluto Médio** (*Mean Absolute Percentage Error - MAPE*)

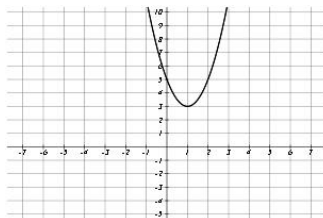
$$MAPE = \frac{\sum_{t=1}^k |e_t^P|}{k}$$

Medidas de Dispersão

- **Erro Quadrático Médio** (Mean Squared Error (MSE))

$$MSE = \frac{\sum_{t=1}^k e_t^2}{k} = \frac{\sum_{t=1}^k (y_t - f_t)^2}{k}$$

- O MSE é uma função quadrática
- Funções quadráticas são mais fáceis de serem otimizadas



Medidas de Dispersão

- Erros de Desvio Padrão (*Standar Deviation of Errors - SDE*)

$$SDE = \sqrt{\frac{\sum_{t=1}^k e_t^2}{k}} = \sqrt{\frac{\sum_{t=1}^k e_t^2}{k}}$$

- O MSE tende amplificar o efeito de grandes erros
- Portanto, o SDE é introduzido para obter uma medida na mesma escala das observações originais

Componentes de uma Série Temporal

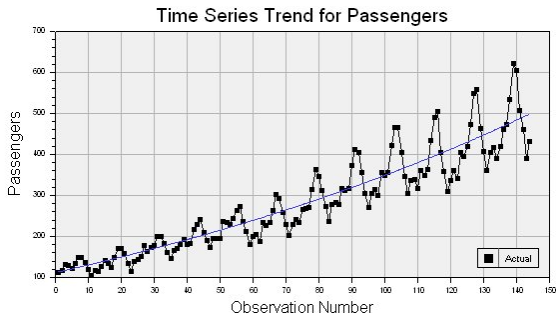
- Pode-se distinguir na série temporal, componentes que podem ser úteis nas previsões
- Esses componentes são:
 - **Tendência** (*trend*)
 - **Sazonalidade** (*seasonality*)
 - **Ruído aleatório** (*random noise*)

Tendência

- Um componente de tendência descreve um comportamento médio ao longo do tempo
- Esse comportamento pode ser
 - Crescente
 - Decrescente
 - Estacionário

Tendência

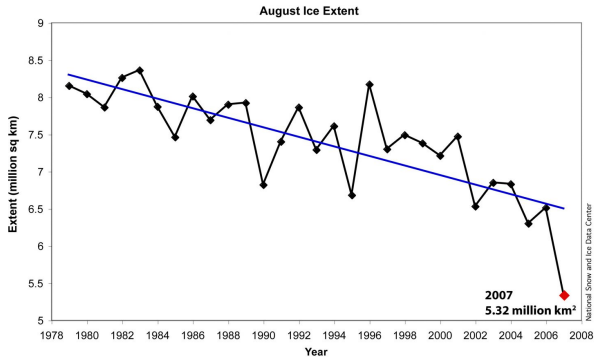
Exemplo de tendência crescente



https://www.dtrek.com/uploaded/pageimg/TsTrend_1.jpg

Tendência

Exemplo de tendência decrescente

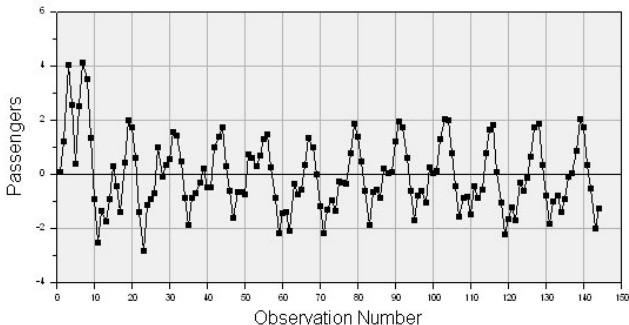


http://nsidc.org/sites/nsidc.org/files/images/asina/20070904_augtrend.jpg

Tendência

Exemplo de tendência estacionária

Transformed Time Series for Passengers



https://www.dtreng.com/uploaded/pageimg/TsTrendAndAmplitude_1.jpg

Média Móvel

- A média móvel (*moving average*) $m_t(h)$ com parâmetro h em um tempo t é definido como é média aritmética de h observações consecutivas das séries temporais $\{y_t\}$
- Pode ser utilizada para identificar tendência, removendo da série flutuações aleatória e componentes de sazonalidade
- É possível computar diferentes valores da média móvel, dependendo da posição do índice t na sequencia das h observações consecutivas utilizadas para calcular a média

-

Média Móvel

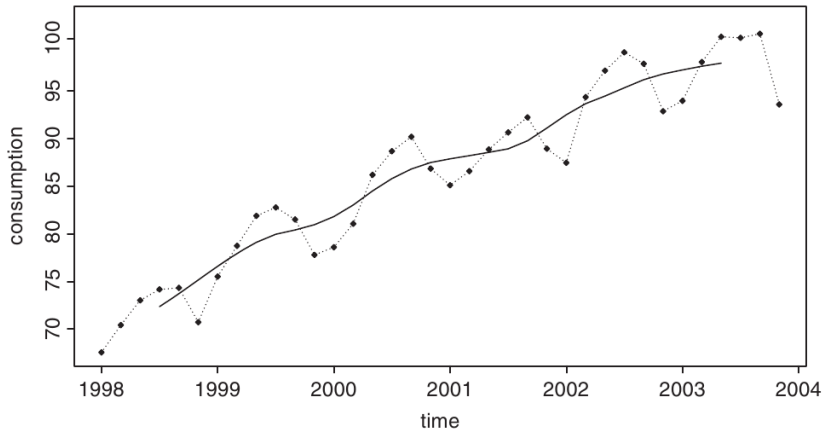
- Média móvel centrada (*centered moving average*), na qual t é o ponto médio das observações, e assumindo que h é ímpar é dada por

$$m_t(h) = \frac{y_{t+(h-1)/2} + y_{t+(h-1)/2-1} \cdots y_{t-(h-1)/2}}{h}$$

- Média móvel não centrada (*noncentered moving average*)

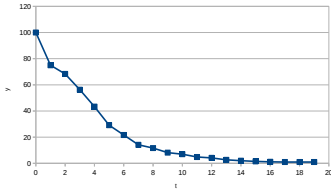
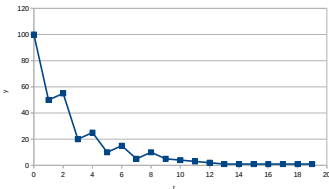
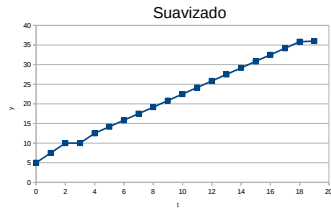
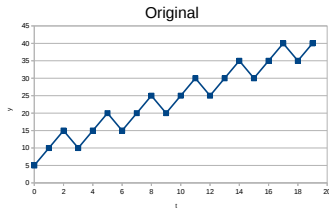
$$m_t(h) = \frac{y_t + y_{t-1} + \cdots + y_{t-(h+1)}}{h}$$

Média Movel Centrada



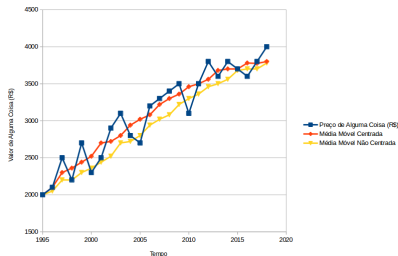
Média Móvel

Exemplo de suavização com média móvel considerando não centrada $h = 5$



Comparativa entre média móvel centrada e não centrada com $h = 5$

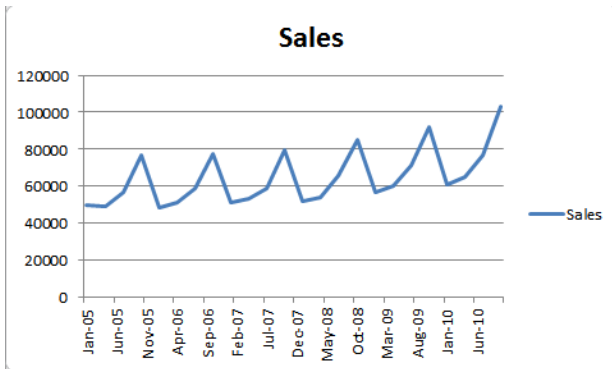
Tempo	Preço de Alguma Coisa (R\$)	Média Móvel Centrada	Média Móvel Não Centrada	$h=5$
1995	2000	2000	2000	
1996	2100	2100	2050	
1997	2500	2300	2200	
1998	2200	2360	2200	
1999	2700	2440	2300	
2000	2300	2520	2360	
2001	2500	2700	2440	
2002	2900	2720	2520	
2003	3100	2800	2700	
2004	2800	2940	2720	
2005	2700	3020	2800	
2006	3200	3080	2940	
2007	3300	3220	3020	
2008	3400	3300	3080	
2009	3500	3360	3220	
2010	3100	3460	3300	
2011	3500	3500	3360	
2012	3900	3560	3460	
2013	3600	3680	3500	
2014	3900	3700	3560	
2015	3700	3700	3680	
2016	3600	3780	3700	
2017	3800	3775	3700	
2018	4000	3800	3780	



Sazonalidade

- Sazonalidade corresponde a variações nos valores da série temporal as quais se repetem ao longo de tempo e em intervalos uniformes
- São também conhecidas por ciclos naturais

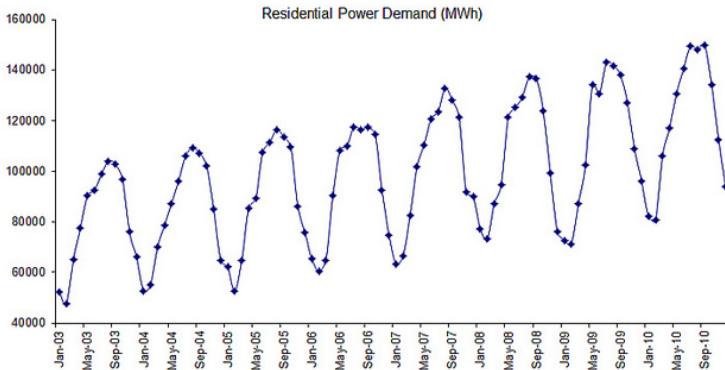
Sazonalidade



<http://cdn2.hubspot.net/hub/64283/file-15465931-png/images/>

[time-series-data-best-suited-for-data-driven-forecasting-methods-resized-600.png](#)

Sazonalidade

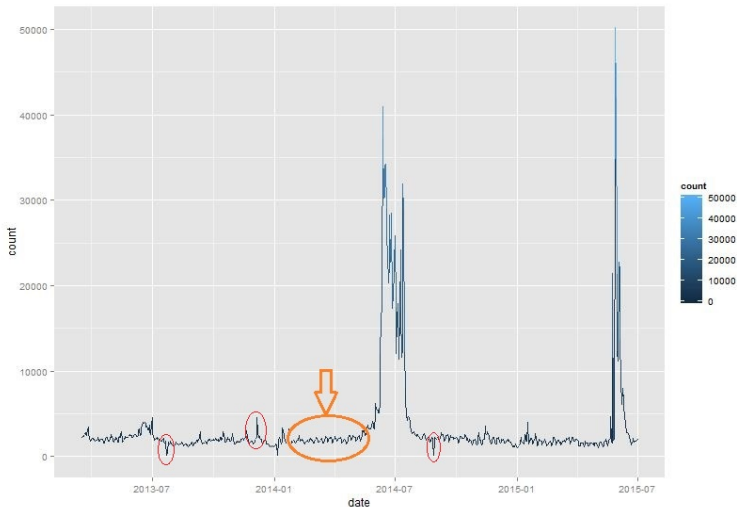


http://farm8.staticflickr.com/7108/7593920758_1a84d6f191_z.jpg

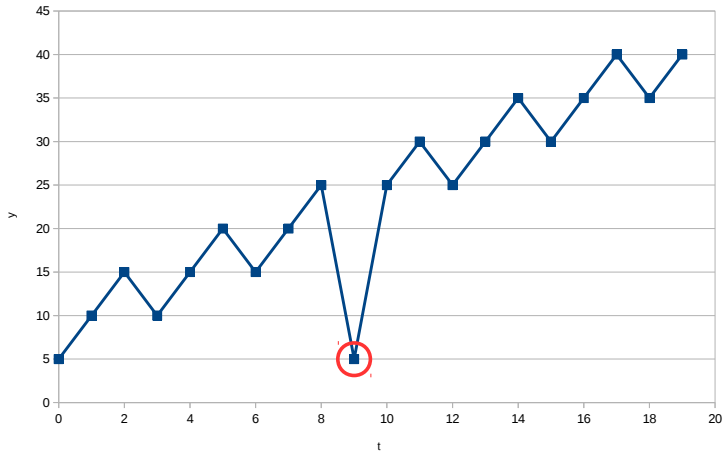
Ruído Aleatório

- É uma variação irregular nos valores de uma série temporal
- Valor que foge a uma tendência ou sazonalidade da série temporal

Ruído Aleatório



Ruído Aleatório



Material Complementar

● 4 - ANÁLISE DE SÉRIES TEMPORAIS

<http://www.inf.ufsc.br/~marcelo.menezes.reis/Cap4.pdf>

Imagem do Dia

