Estatística Aplicada
EAD Ao Vivo

Tema da aula
Série Temporal - AR

LABDATA | FIA FUNDAÇÃO INSTITUTO DE ADMINISTRAÇÃO BUSINESS SCHOOL

16/12/2020

ESTATÍSTICA APLICADA EAD AO VIVO

Professora:
Drª Karin Ayumi Tamura

Coordenadores:
Profª Drª Alessandra de Ávila Montini
Profª Dr. Adolpho Walter Pimazoni Canton

Currículo - Prof.ª Drª Karin Ayumi Tamura

FORMAÇÃO ACADÊMICA | EXPERIÊNCIA PROFISSIONAL

Profª Dra.
Karin Ayumi Tamura

"Tenho duas paixões no meu trabalho: dados e pessoas. Voltar a lecionar no LABDATA FIA está sendo a realização de um sonho planejado desde a minha época de aluna de pós-graduação. Meu objetivo como professora é integrar a visão do mercado com as técnicas e tecnologias de análise de dados, por meio de uma atuação humanista no ensino aos alunos"

Contato: karin.tamura@fia.com.br

- **FORMAÇÃO ACADÊMICA:** Pós-doutora (2015), Doutora (2012), mestre (2007) e bacharel (2003) em Estatística pelo Instituto de Matemática e Estatística da USP, tendo como área de pesquisa modelos de regressão, análise multivariada de dados e algoritmos de *machine learning*.

- **ATUAÇÃO PROFISSINAL:** Foi *Head* de *Analytics* por 14 anos, e atualmente é Conselheira Executiva e *Head* de Inovação na *Marketdata Solutions*, uma empresa do grupo WPP, e Professora Doutora no LABDATA FIA.

- **HISTÓRICO:** Atuação no mercado por 17 anos, com experiência profissional no segmento bancário (Bradesco) e consultoria *(Marketdata Solutions)*. Atuou como docente em cursos de pós-graduação (2010-16) no LABDATA FIA e ABEMD. Especialista em Estatística e *Advanced Analytics* trabalhando em projetos de diversos segmentos do mercado. Participante de congressos nacionais e internacionais voltados a área de Estatística, Dados e Algoritmos de *Machine Learning*.

Projetos atendidos

FIA Business School

NOSSOS DIFERENCIAIS | QUEM SOMOS

BUSINESS SCHOOL
Graduação, pós-graduação, MBA, Pós- MBA, Mestrado Profissional, Curso In *Company* e EAD

CONSULTING
Consultoria personalizada que oferece soluções baseada em seu problema de negócio

RESEARCH
Atualização dos conhecimentos e do material didático oferecidos nas atividades de ensino

FUNDAÇÃO INSTITUTO DE ADMINISTRAÇÃO

Líder em Educação Executiva, referência de ensino nos cursos de graduação, pós-graduação e MBA, tendo excelência nos programas de educação. Uma das principais escolas de negócio do mundo, possuindo convênios internacionais com Universidades nos EUA, Europa e Ásia. +8.000 projetos de consultorias em organizações públicas e privadas.

NOTA MÁXIMA ENADE 2019 MELHOR ESCOLA DO BRASIL
Único curso de graduação em administração a receber as notas máximas

A primeira escola brasileira a ser finalista da maior competição de MBA do mundo

Única *Business School* brasileira a figurar no *ranking* LATAM

Signatária do Pacto Global da ONU

Membro fundador da ANAMBA - Associação Nacional MBAs

Credenciada pela AMBA - *Association of MBAs*

Credenciada ao *Executive MBA Council*

Filiada a AACSB - *Association to Advance Collegiate Schools of Business*

Filiada a EFMD - *European Foundation for Management Development*

Referência em cursos de MBA nas principais mídias de circulação

O **Laboratório de Análise de Dados** – LABDATA é um Centro de Excelência que atua nas áreas de ensino, pesquisa e consultoria em análise de informação utilizando técnicas de ***Big Data***, ***Analytics*** e **Inteligência Artificial**.

Profª Drª Alessandra Montini

O LABDATA é um dos pioneiros no lançamento dos cursos de *Big Data* e *Analytics* no Brasil

Os diretores foram professores de grandes especialistas do mercado

+10 anos de atuação

+1000 alunos formados

Docentes

- ➢ Sólida formação acadêmica: doutores e mestres em sua maioria
- ➢ Larga experiência de mercado na resolução de *cases*
- ➢ Participação em Congressos Nacionais e Internacionais
- ➢ Professor assistente que acompanha o aluno durante todo o curso

Estrutura

- ➢ 100% das aulas realizadas em laboratórios
- ➢ Computadores para uso individual durante as aulas
- ➢ 5 laboratórios de alta qualidade (investimento +R$2MM)
- ➢ 2 Unidades próximas a estação de metrô (com estacionamento)

Conteúdo Programático do Curso

21 AULAS AO VIVO COM PROFA. KARIN | 27 PLANTÕES AO VIVO COM PROF. STEPHAN, 7 LISTAS DE EXERCÍCIOS E EAD VIDEO AULA EM PYTHON

Dia	Mês	Aula	EAD Ao Vivo	Plantão Prof. Stephan
5	Agosto	Introdução ao Curso e Análise Exploratória de Dados	Aula Prof. Karin	06/ago
12	Agosto	Análise Exploratória de Dados	Aula Prof. Karin	13/ago
19	Agosto	Análise Exploratória de Dados - Introdução ao R	Aula Prof. Karin	20/ago
26	Agosto	Lista de Exercícios em Sala de Aula (19hs-23hs - com presença obrigatória)	-	27/ago
2	Setembro	Regressão Linear Simples	Aula Prof. Karin	03/set
9	Setembro	Regressão Linear Simples e Múltipla	Aula Prof. Karin	10/set
16	Setembro	Regressão Linear Simples e Múltipla	Aula Prof. Karin	17/set
23	Setembro	Lista de Exercícios em Sala de Aula (19hs-23hs - com presença obrigatória)	-	24/set
30	Setembro	Análise de Cluster	Aula Prof. Karin	01/out
7	Outubro	Análise de Cluster	Aula Prof. Karin	08/out
14	Outubro	Lista de Exercícios em Sala de Aula (19hs-23hs - com presença obrigatória)	-	15/out
21	Outubro	Arvore de Decisão	Aula Prof. Karin	22/out
28	Outubro	Lista de Exercícios em Sala de Aula (19hs-23hs - com presença obrigatória)	-	29/out
4	Novembro	Regressão Logística	Aula Prof. Karin	05/nov
11	Novembro	Regressão Logística	Aula Prof. Karin	11/nov
18	Novembro	Lista de Exercícios em Sala de Aula (19hs-23hs - com presença obrigatória)	-	19/nov
25	Novembro	estudo de caso	Aula Prof. Karin	26/nov
2	Dezembro	estudo de caso	Aula Prof. Karin	30/dez
9	Dezembro	estudo de caso	Aula Prof. Karin	10/dez
16	Dezembro	Análise de Série Temporal - modelo auto regressivo	Aula Prof. Karin	17/dez
23	Dezembro	Lista de Exercícios em Sala de Aula (Frequência Liberada - véspera Natal)	-	-
Recesso Escolar		EAD - INTRODUÇÃO AO PYTHON	EAD Video Aula (8 horas)	-
		EAD - INTRODUÇÃO AO PYTHON		-
6	Janeiro	Modelos estatísticos em Python	Aula Prof. Karin	07/jan
13	Janeiro	Modelos estatísticos em Python	Aula Prof. Karin	14/jan
20	Janeiro	Modelos estatísticos em Python	Aula Prof. Karin	20/jan
27	Janeiro	Introdução a Big Data - Aplicações de Machine Learning e Deep Learning	Aula Prof. Karin	28/jan
3	Fevereiro	Aplicações de Machine Learning	Aula Prof. Karin	04/fev
10	Fevereiro	Aplicações de Machine Learning	Aula Prof. Karin	11/fev
17	Fevereiro	Lista de Exercícios (Frequência Liberada - quarta de cinzas)	-	18/fev
24	Fevereiro	EXERCICIOS DE REVISÃO - EAD (19hs e 23hs - com presença obrigatória)	-	24/fev
3	Março	Prova (Plataforma On Line: 19hs e 23hs)	-	

Conteúdo da Aula

1. Introdução
   1. Componentes de uma Série Temporal
   2. Estacionariedade
   3. Metodologia de Box e Jenkins

2. Modelo Auto-regressivo: Ordens 1, 2, p e incompleto
   1. Identificação
   2. Estimação
   3. Previsão
3. Exercícios para casa
   - CASE: *Call Center*
   - CASE: Retorno de Ações
   - Exercício de Fixação

1. Introdução

Estrutura de dados de série histórica

Nesta aula trabalharemos com dados estruturados de forma longitudinal: séries de tempo univaridadas. Diferente da estrutura de seção transversal (*cross sectional*) trabalhadas até o momento.

Data	Compra	Venda
01/04/2020	5,2399	5,2404
02/04/2020	5,2645	5,2651
03/04/2020	5,2991	5,2997
06/04/2020	5,2465	5,2471
07/04/2020	5,2211	5,2217
08/04/2020	5,2117	5,2123
09/04/2020	5,0773	5,0779
13/04/2020	5,1818	5,1824
14/04/2020	5,1852	5,1858
15/04/2020	5,2573	5,2579
16/04/2020	5,2371	5,2377
17/04/2020	5,2567	5,2573
20/04/2020	5,2831	5,2837
22/04/2020	5,3841	5,3847
23/04/2020	5,4461	5,4467
...	...	...

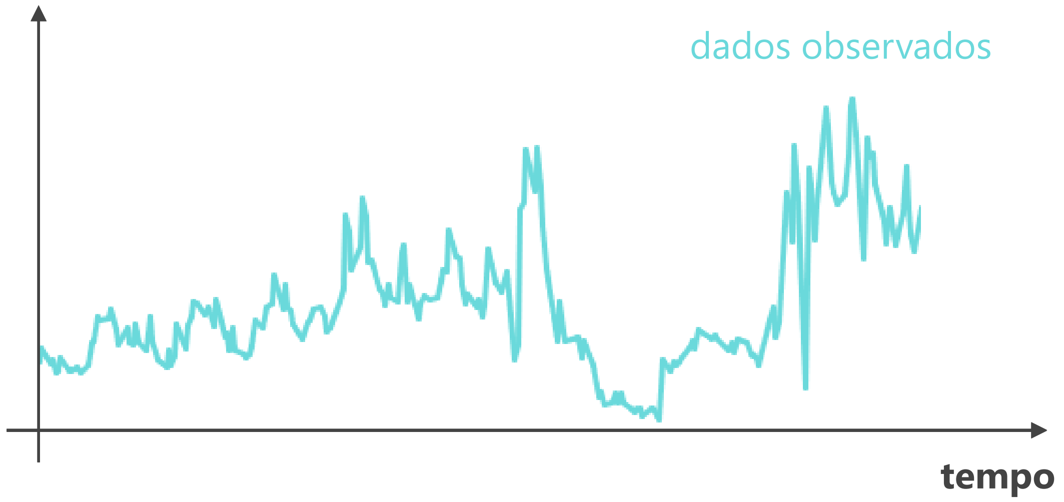
Valor de compra do dólar analisado diariamente nos últimos 14 anos.

Dados de séries históricas

São dados coletados ao longo de diversos períodos de tempo. Neste banco de dados, o valor do dólar foi analisado nos últimos 14 anos, na periodicidade diária.

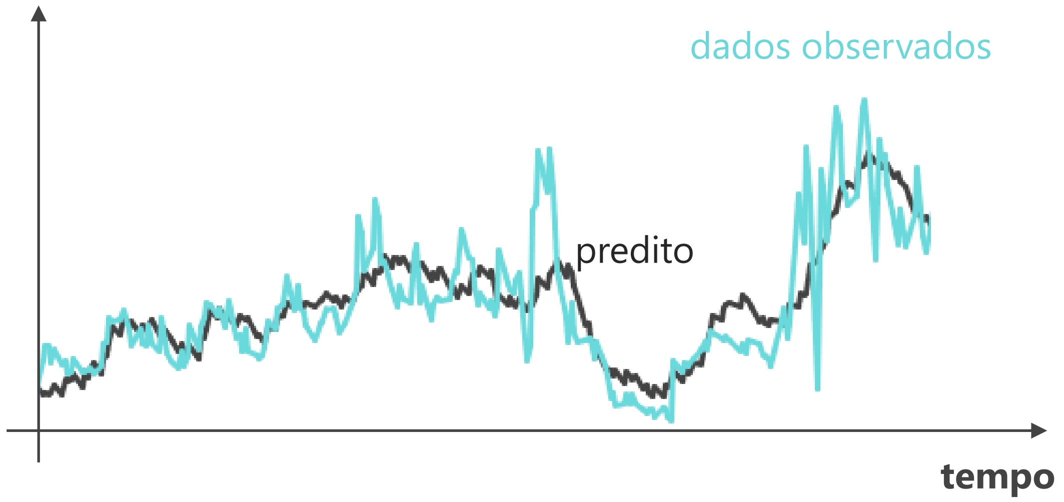
Componentes de uma Série Temporal

Série temporal: dados observados.



Componentes de uma Série Temporal

Série temporal: dados observados e predito pelo modelo.



As séries de tempo são analisadas com o objetivo de compreender o passado e projetar no futuro a mesma série de dados ao longo do tempo.

Uma característica importante das séries de tempo é que as observações vizinhas apresentam dependência entre si, e estamos interessados em modelar essa dependência.

Diferente dos modelos de regressão (dados com seção transversal) em que a ordem das observações apresentadas é irrelevante para análise de dados.

Objetivo

Na análise de série temporal observa-se a série em um período, obtém-se a equação de projeção e projeta-se os valores futuros com base no comportamento passado da série temporal.

https://www.logianalytics.com/logi-news/logi-predict-product-update-key-enhancements/

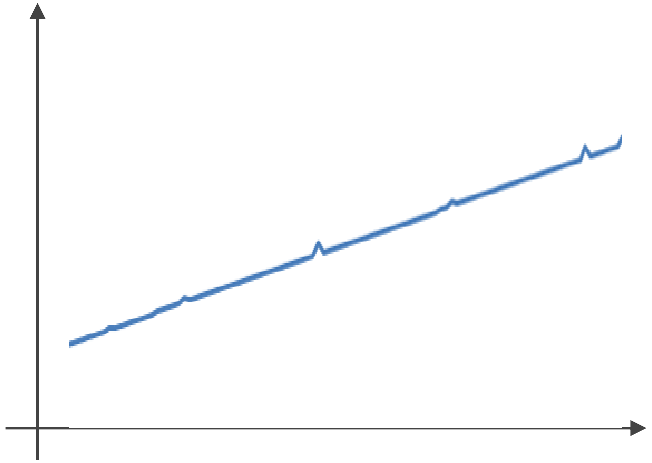
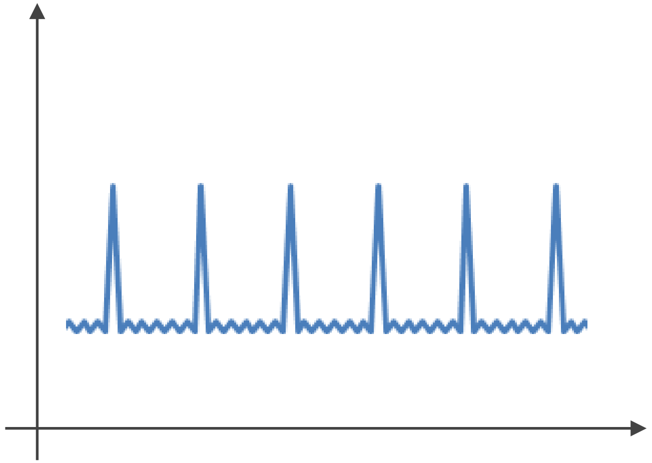
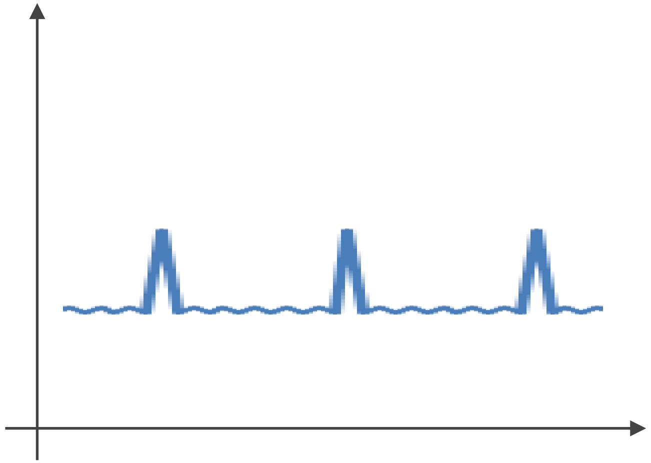
Aplicações

Diversas áreas da ciências têm utilizado amplamente modelos de série de tempo, como por exemplo: Economia, Administração, Marketing, Engenharia, Agronegócio, entre outros.

Exemplos de segmentos de negócios que utilizam tradicionalmente Séries Temporais:

- **Financeiro:** previsão de preços de ativos financeiros, previsão de faturamento da empresa, previsão de *budget* de clientes, etc.
- **Marketing:** previsão de aquisição de novos clientes, previsão de regaste de pontos de programas de *loyalty*, previsão de volume de viagens, etc.
- **Administrativo:** previsão de vendas de produtos, previsão de quantidade de funcionários da empresa, previsão de mobilidade urbana, etc.
- **Saúde:** previsão de demanda de leitos, previsão de quantidade de pacientes na UTI, etc.

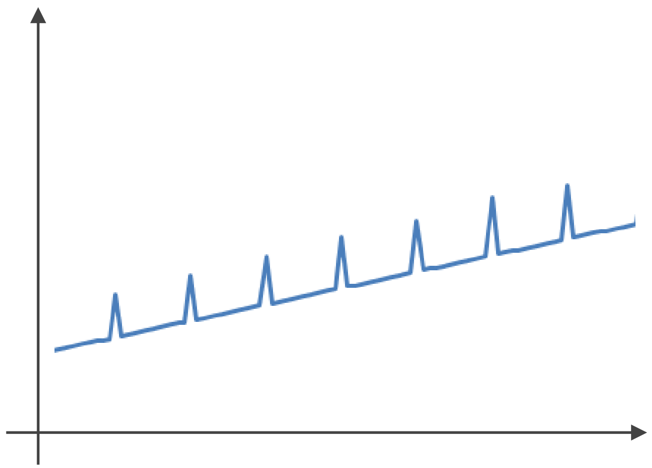
Componentes de uma Série Temporal

Tendência	Sazonalidade	Ciclo
		
Esta série temporal apresenta **tendência**, pois os valores crescem ao longo do tempo.	Esta série apresenta sazonalidade, pois tem picos em determinados **períodos do ano**. Ex.: vendas no natal, dias das mães, etc.	Esta série apresenta ciclo, pois tem um comportamento que se repete em um período **maior que um ano**.

Componentes de uma Série Temporal

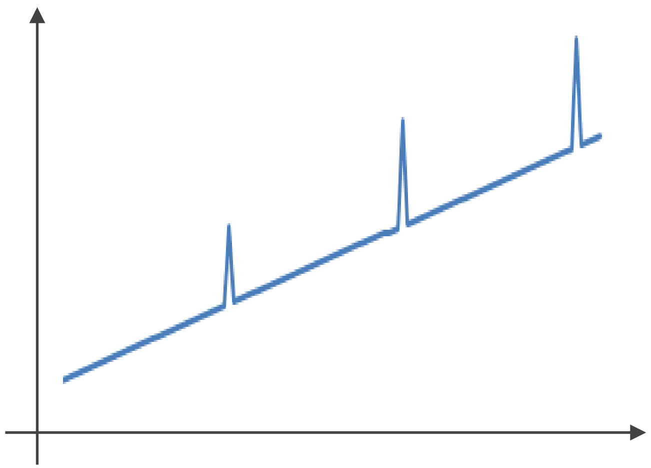
1.i. COMPONENTES | SÉRIE TEMPORAL

Tendência e Sazonalidade



Esta série temporal apresenta **tendência** e **sazonalidade,** pois os valores crescem ao longo do tempo e apresentam picos em todo mês de dezembro, por exemplo.

Tendência e Ciclo



Esta série temporal apresenta **tendência** de crescimento em **períodos cíclicos,** pois apresenta comportamentos que se repetem.

Google

Brasil 7 de julho de 2020

COVID-19: Relatório de mobilidade da comunidade

https://www.gstatic.com/covid19/mobility/2020-07-07_BR_Mobility_Report_pt-BR.pdf

State of São Paulo

Mudanças na mobilidade

Retail & recreation

-54% compared to baseline

Grocery & pharmacy

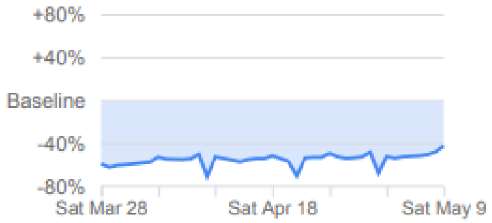
+10% compared to baseline

Parks

-49% compared to baseline

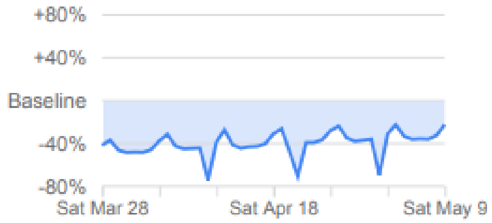
Transit stations

-42% compared to baseline



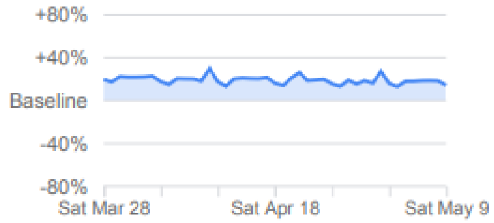
Workplace

-23% compared to baseline



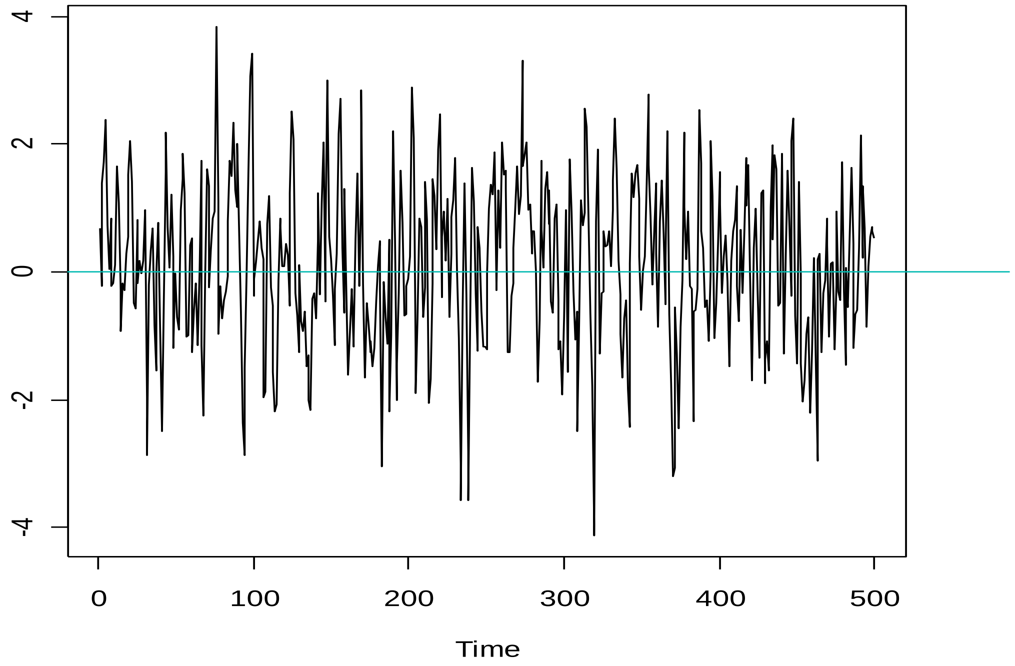
Residential

+15% compared to baseline



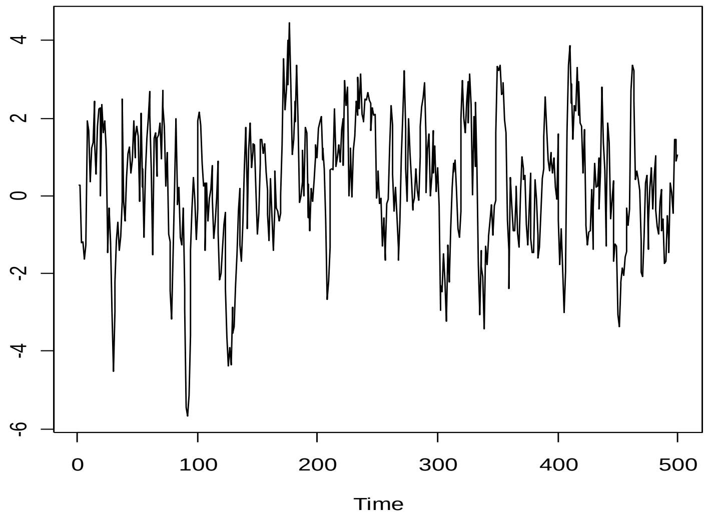
Séries Estacionárias

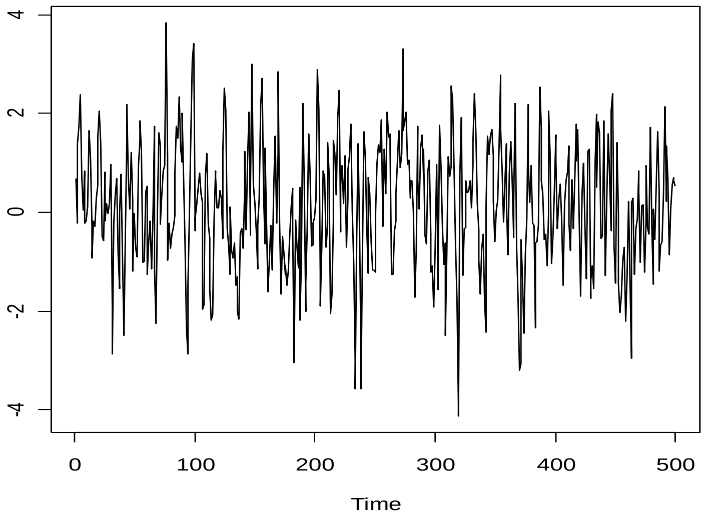
Uma série temporal é estacionária quando oscila ao redor de uma **média constante,** refletindo um comportamento estável e aleatório ao longo do tempo.

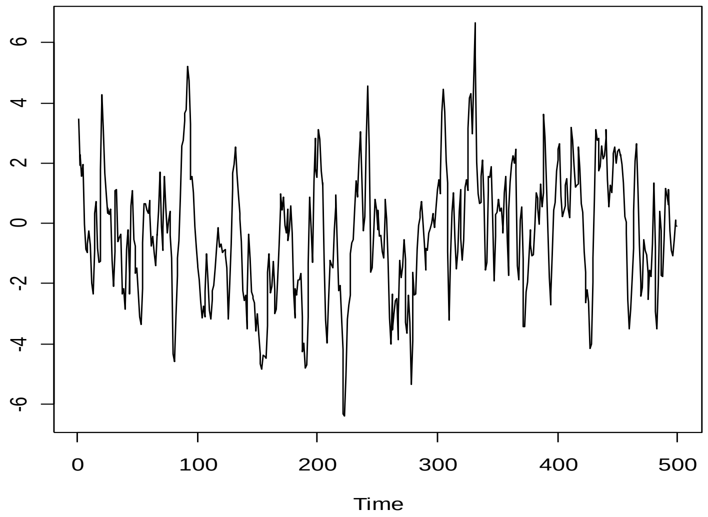


A média da série é constante

Exemplos de séries de tempo estacionárias:



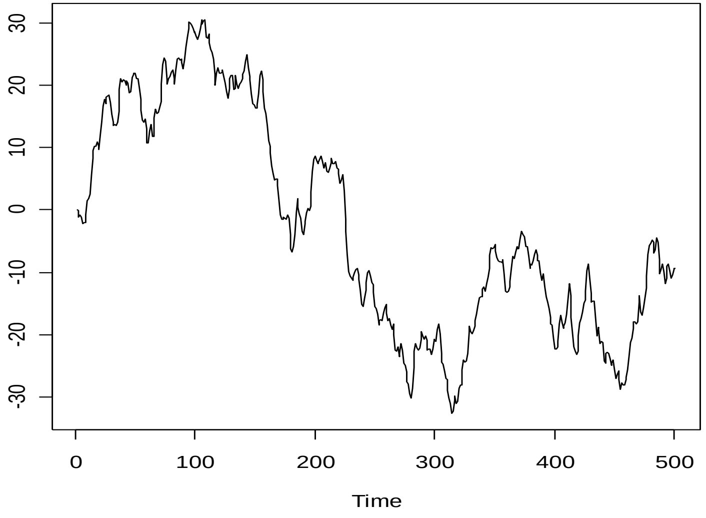


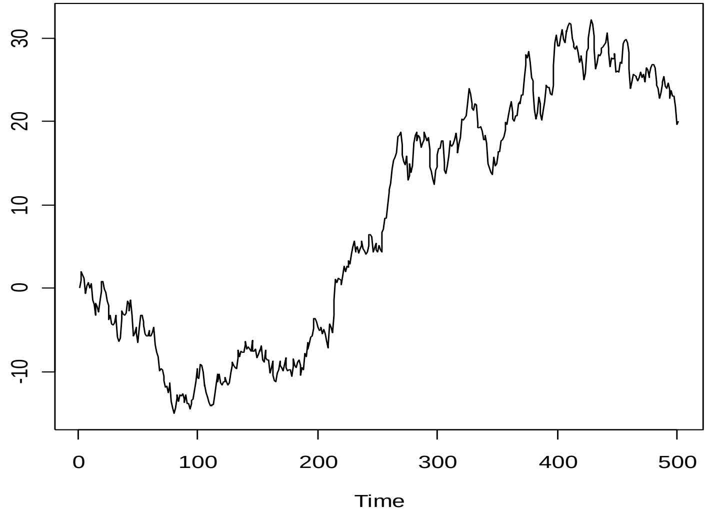


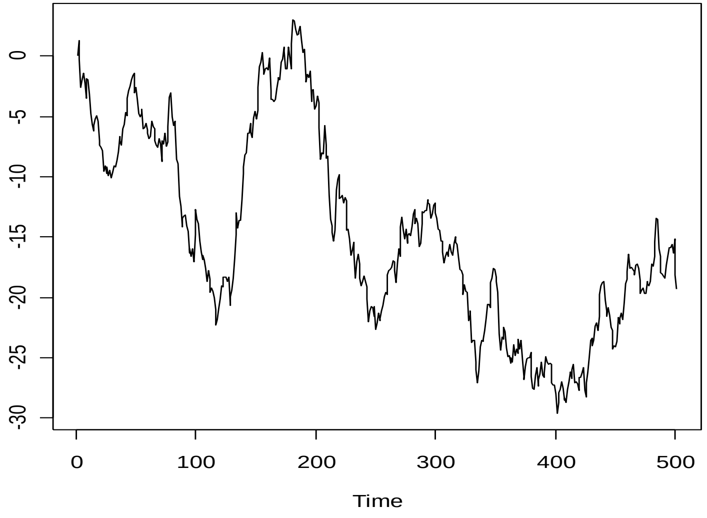
Quando uma série é estacionária, a média e a variância são constantes, então elas dependem da distância dos passos (*lags* de defasagens) que as separam.

Uma série temporal <u>não é estacionária</u> quando:

- A média não é constante.
- A série apresenta tendência.







Metodologia de Box e Jenkins

Box e Jenkins (1976) propuseram uma metodologia para ajuste de **modelos de séries temporais estacionários** que é composta pelas seguintes etapas:

i. Identificação
ii. Estimação
iii. Previsão

Teste de Estacionariedade

Para o ajuste de uma série temporal por meio da metodologia de Box e Jenkins é necessário verificar se a série é estacionária, pois a metodologia de Box e Jenkins é aplicada quando a série é estacionária.

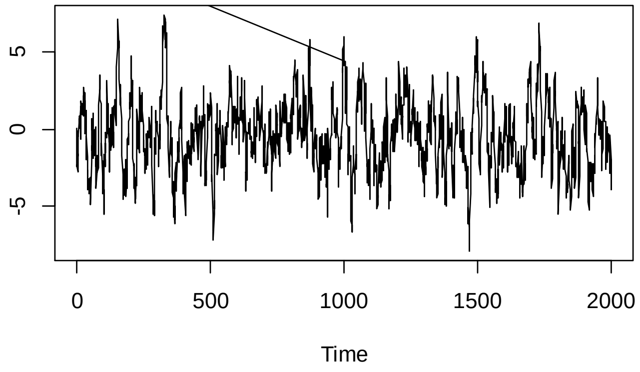
Apesar de ser possível visualizar graficamente a estacionariedade da série, pode-se utilizar o **teste de estacionariedade de Dickey-Fuller Aumentado** (Dickey & Fuller, 1979) , com as seguintes hipóteses:

H_0: a série não é estacionária

H_1: a série é estacionária

Regra de decisão: Quando o nível descritivo é < 0,10 rejeitamos H_0, ou seja, há evidência de que a série é estacionária.

Esta série temporal parece estacionária, pois oscila em torno de uma média constante.



Modelos para Séries Estacionárias

A literatura de modelos de séries ao longo do tempo apresenta uma diversidade de famílias de modelos que podem ser utilizados para ajuste de séries estacionárias.

Na aula de hoje, será apresentado o **Modelo Auto-regressivo (AR)**, amplamente utilizado no mercado, em que a predição futura depende do comportamento passado da própria série, de um ou vários períodos históricos.

2. Modelo Auto-regressivo

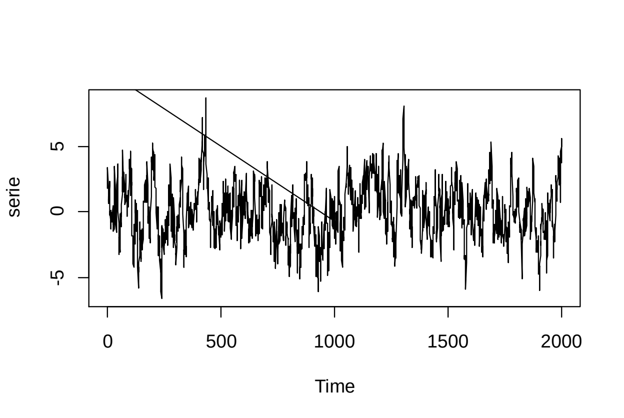
Modelo Auto-regressivo

2. MODELO AUTO-REGRESSIVO | SÉRIE TEMPORAL

Uma série temporal segue um processo auto-regressivo quando o valor da série no tempo t depende do que aconteceu em, por exemplo, t-1, t-2, t-3, t-4, etc.

Em um modelo auto-regressivo, a previsão de valores futuros da série é realizada utilizando uma **combinação linear de valores do passado** da mesma série de dados.

O termo "auto-regressivo" indica a regressão da variável resposta contra ela mesma no passado.



Exercício: AR(1)

Uma série temporal segue um processo auto-regressivo de ordem 1 ou AR(1) quando o valor da série no tempo t depende do valor da série em t-1.

$$\hat{Y}_t = 24 + 0{,}823Y_{t-1}$$

$\hat{Y}_t$: valor previsto da série no instante t

Y_{t-1} : valor da série no instante t-1

Seja Y, a variável quantidade de chamadas em uma central de atendimento de uma empresa de serviços que dá suporte para seus usuários, quando o serviço apresenta problemas de funcionamento. A variável Y é modelada na periodicidade diária.

O gestor da área gostaria de saber qual será a quantidade de chamadas previstas para amanhã, dado que hoje o dia fechou com 321 chamadas.

Modelo Auto-regressivo de Ordem 1 – AR(1)

Uma série temporal segue um processo auto-regressivo de ordem 1 ou AR(1) quando o valor da série no tempo t depende do valor da série em t-1.

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \varepsilon_t$$

Y_t : valor da série no instante t

Y_{t-1} : valor da série no instante t-1

ϕ_0: parâmetro do modelo

ϕ_1: parâmetro do modelo associado ao Y_{t-1}

ε_t : erro aleatório no instante t

Suposição: o erro é uma variável aleatória que possui média zero e variância σ^2.

Exercício: AR(2)

Uma série temporal segue um processo auto-regressivo de ordem 2 ou AR(2) quando o valor da série no tempo t depende do valor da série em t-1 e t-2.

$$\hat{Y}_t = 2{,}13 + 0{,}796Y_{t-1} + 0{,}235Y_{t-2}$$

$\hat{Y}_t$: valor previsto da série no instante t

Y_{t-1} : valor da série no instante t-1

Y_{t-2} : valor da série no instante t-2

Seja Y, a variável quantidade de vendas de um determinado item de um *e-commerce*. A variável Y é modelada na periodicidade mensal.

O dono do *e-commerce* gostaria de saber qual será a projeção de vendas para o mês corrente, dado que no mês passado ele vendeu 456 itens e mês retrasado 524.

Modelo Auto-regressivo de Ordem 2 – AR(2)

Uma série temporal segue um processo auto-regressivo de ordem 2 ou AR(2) quando o valor da série no tempo t depende do valor da série em t-1 e t-2.

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \varepsilon_t$$

Y_t : valor da série no instante t

Y_{t-1} : valor da série no instante t-1

Y_{t-2} : valor da série no instante t-2

ϕ_0 : parâmetro do modelo

ϕ_1 : parâmetro do modelo associado ao Y_{t-1}

ϕ_2 : parâmetro do modelo associado ao Y_{t-2}

ε_t : erro aleatório no instante t

Suposição: o erro é uma variável aleatória que possui média zero e variância σ^2.

Modelo Auto-regressivo de Ordem p – AR(p)

Uma série temporal segue um processo auto-regressivo de ordem p ou AR(p) quando o valor da série no tempo t depende do valor da série em t-1, t-2 até t-p.

$$Y_t = \phi_0 + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \ldots + + \phi_p Y_{t-p} + \varepsilon_t$$

Y_t : valor da série no instante t

$Y_{t-1}, Y_{t-2}, \ldots, Y_{t-p}$: valor da série no instante t-1, t-2 e t-p, respectivamente

ϕ_0 : parâmetro do modelo

ϕ_1 : parâmetro do modelo associado ao Y_{t-1}

ϕ_2 : parâmetro do modelo associado ao Y_{t-2}

$\vdots$

ϕ_p : parâmetro do modelo associado ao Y_{t-p}

ε_t : erro aleatório no instante t

Suposição: o erro é uma variável aleatória que possui média zero e variância σ^2.

Exercício: AR Incompleto - Ordens (2,5)

Uma série temporal segue um processo auto-regressivo de ordem incompleta, quando o valor da série no tempo t depende do valor da série em tempos passados específicos (não necessariamente com *lags* de defasagens consecutivas).

$$\hat{Y}_t = 14{,}7 + 0{,}791 Y_{t-2} + 0{,}125 Y_{t-5}$$

$\hat{Y}_t$: valor previsto da série no instante t

Y_{t-2} : valor da série no instante t-2

Y_{t-5} : valor da série no instante t-5

Seja Y, a variável quantidade de frutas que não foram vendidas (em um dia) de um grande Hortifruti. A variável Y é modelada na periodicidade diária.

O gerente do Hortifruti, com o objetivo de evitar desperdício, gostaria de saber qual será a projeção de sobra de frutas hoje, tendo em mãos os dados da última semana:

- **Ontem: 324**
- **Antes de ontem: 245**
- **t-3: 345**
- **t-4: 279**
- **t-5: 319**
- **t-6: 351**

Metodologia de Box e Jenkins

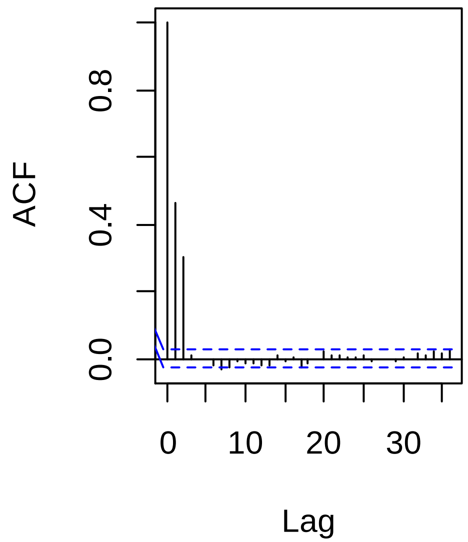
Uma vez confirmado pelo teste de Dickey-Fuller Aumentado que a série é estacionária, pode-se utilizar a metodologia de Box e Jenkins, que consiste na realização das seguintes fases:

i. Identificação
ii. Estimação
iii. Previsão

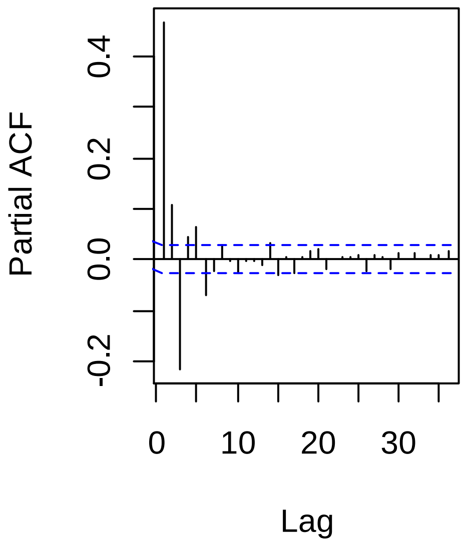
Para identificar se a série temporal segue um modelo Auto-regressivo, deve-se analisar as funções de autocorrelação (ACF – *autocorrelation function*) e autocorrelação parcial (PACF – *partial autocorrelation function*).

Esta é uma análise exploratória da série de dados para investigar as ordens a serem sugeridas para o modelo AR de maneira a se obter uma estrutura de modelo parcimoniosa.

Função Autocorrelação



Função de Autocorrelação Parcial



Para o modelo ser auto-regressivo, pela análise da ACF, várias autocorrelações devem apresentar valores diferentes de zero.

Note que a **ACF na *lag* zero será igual a um**, pois é a correlação da série no mesmo instante t contra ela mesma, ou seja, sem *lag* de defasagem.

Além disso, pela análise da PACF, as autocorrelações diferentes de ZERO dão um indicativo da ordem do AR que deve ser testada. No exemplo ao lado, AR(6).

2.1 Fase de Identificação

Para identificar se a série temporal segue um modelo Auto-regressivo deve-se analisar as funções de autocorrelação (ACF – *autocorrelation function*) e autocorrelação parcial (PACF – partial *autocorrelation function*).

O entendimento da autocorrelação e da autocorrelação parcial é:

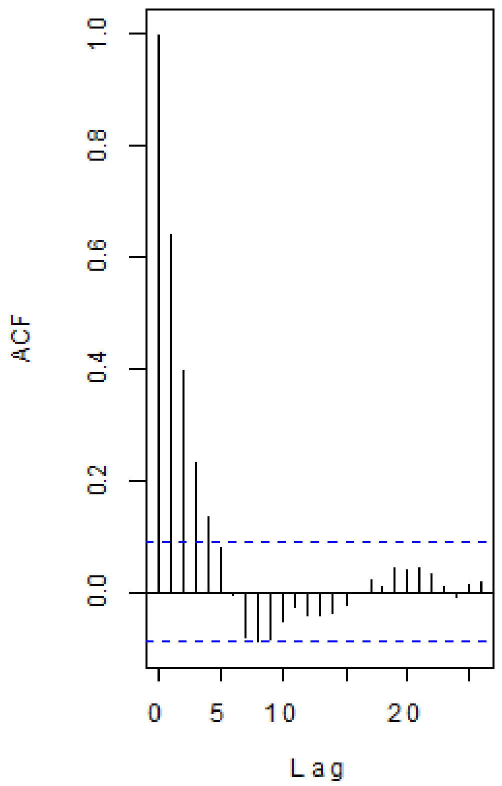
Autocorrelação

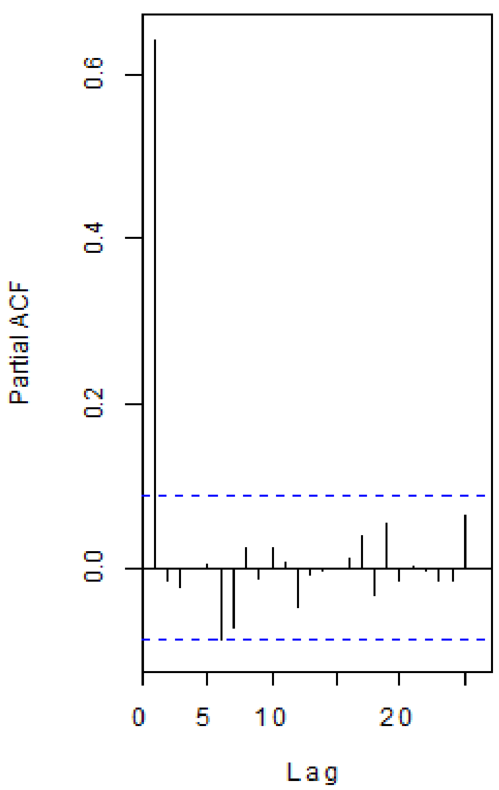
A autocorrelação mede a correlação linear entre as observações da série em diferentes períodos de tempos, ou seja, as correlações são calculadas nas *lags* defasadas em 1, 2, 3, 4, ... períodos de tempo. Denominamos essa defasagem como *lag* 1, *lag* 2, *lag* 3, *lag* 4, etc. Na *lag* 0 a autocorrelação será sempre igual a 1.

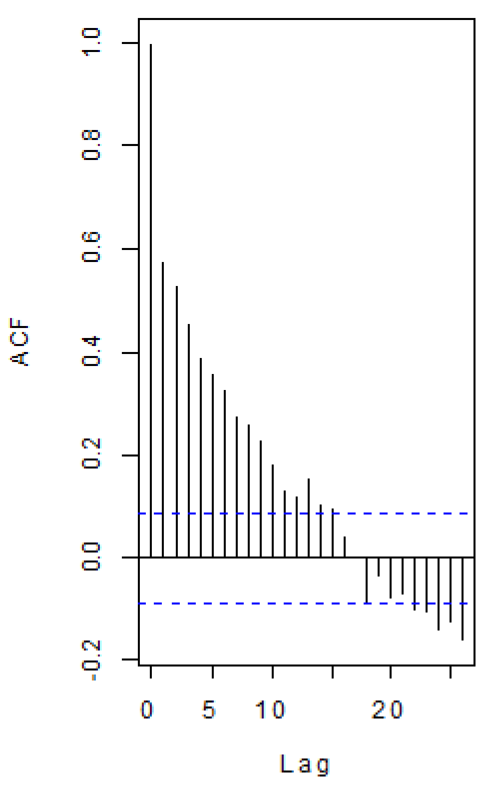
Autocorrelação Parcial

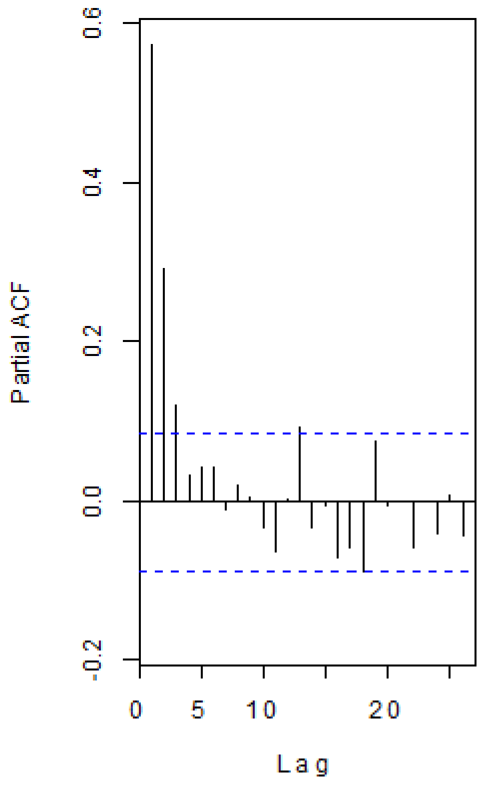
A função de autocorrelação parcial é útil para determinar a ordem de um processo AR.
Por exemplo, um modelo será AR(1) se todas as autocorrelações parciais das ordens >1 forem iguais a zero. No geral, uma AR(k) terá sua a autocorrelação parcial igual a zero depois da *lag* k.

Exercício: Identifique pelas ACF e ACFP se seriam modelos AR e qual sua ordem

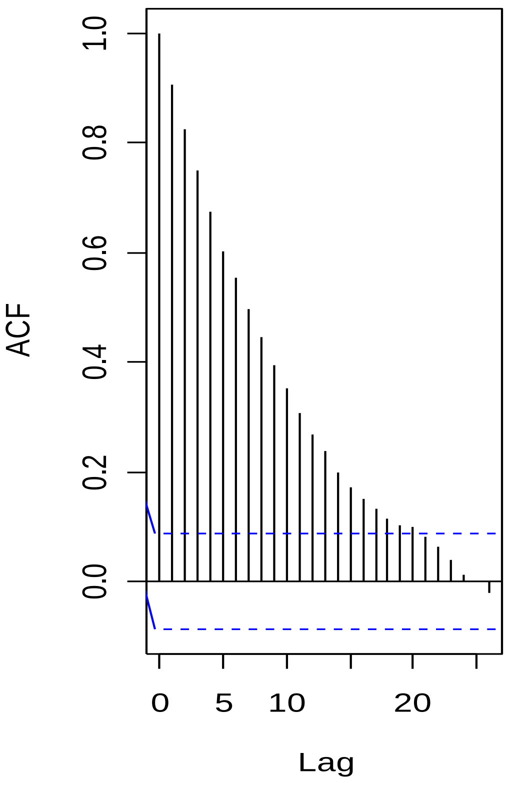


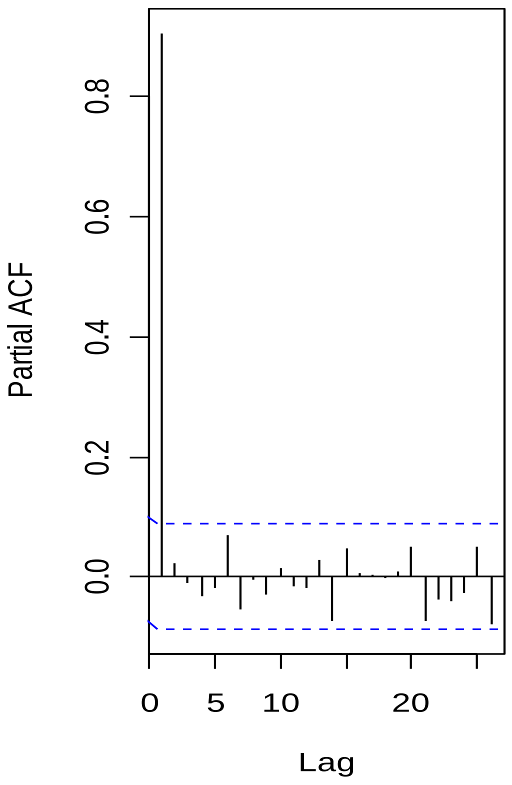


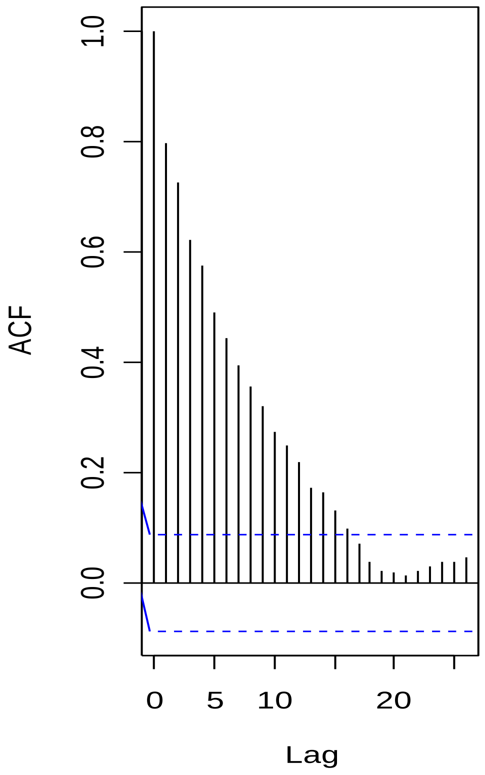


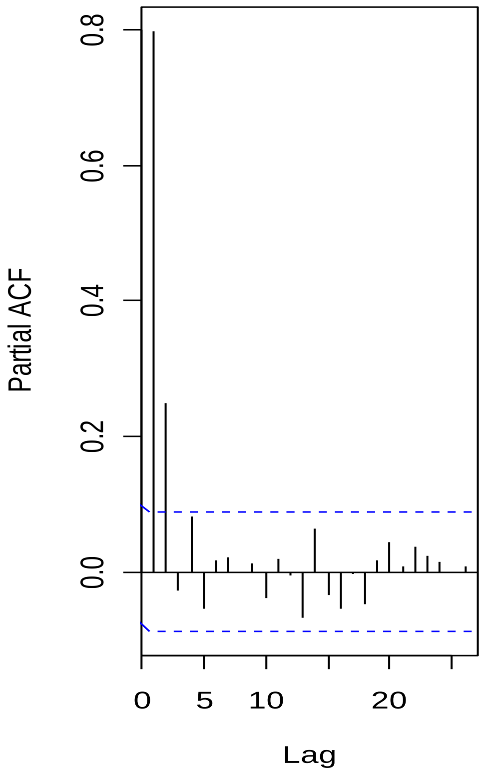


Exercício: Identifique pelas ACF e ACFP se seriam modelos AR e qual sua ordem









Uma vez identificada a ordem do modelo AR a ser testada, o próximo passo é estimar os parâmetros do modelo:

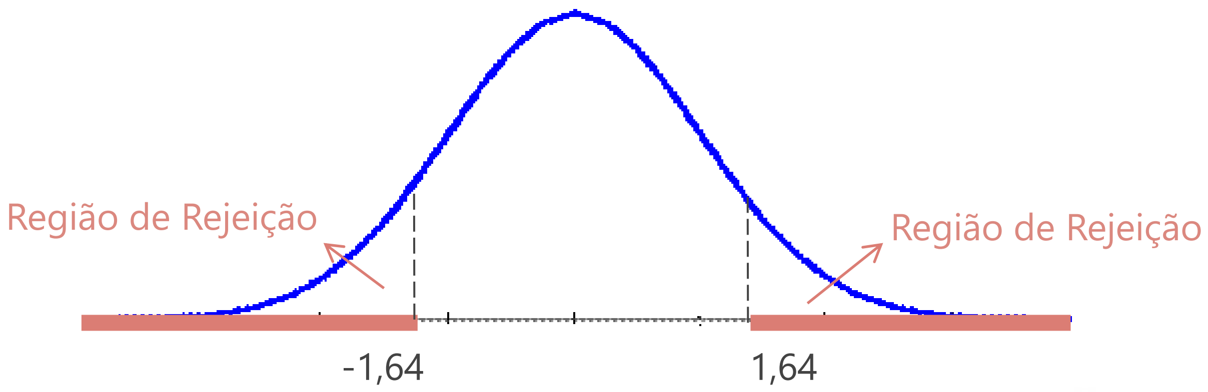
$$\hat{Y}_t = \hat{\phi}_0 + \hat{\phi}_1 Y_{t-1} + \hat{\phi}_2 Y_{t-2} + \ldots + + \hat{\phi}_p Y_{t-p}$$

$$\hat{\phi}_0 = intercepto\left(1 - \hat{\phi}_1 - \hat{\phi}_2 - \hat{\phi}_2 \ldots - \hat{\phi}_p\right)$$

Teste de hipótese para cada um dos parâmetros i=0,1, ..., p:

$$H_0: \phi_i = 0$$
$$H_1: \phi_i \neq 0$$

Considerando 90 % de confiança da Distribuição Normal:



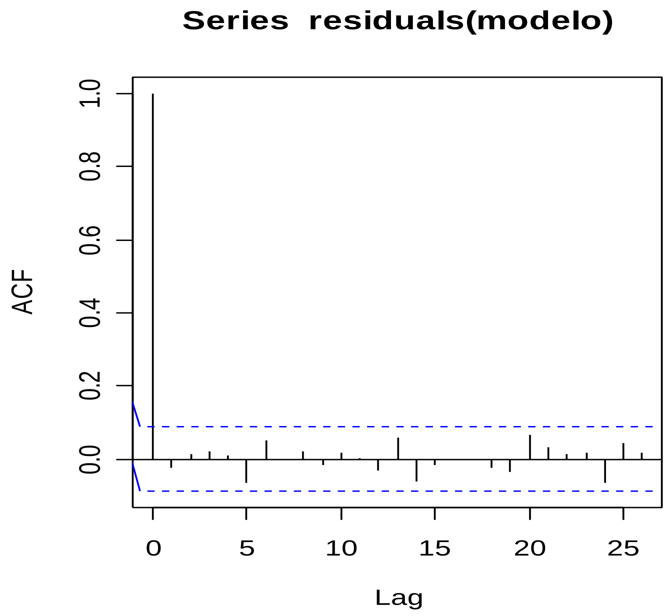
Quando o p-valor do teste estiver abaixo do nível descritivo adotado de 10% ou quando a estatística do teste estiver na região de rejeição, deve-se rejeitar H_0.

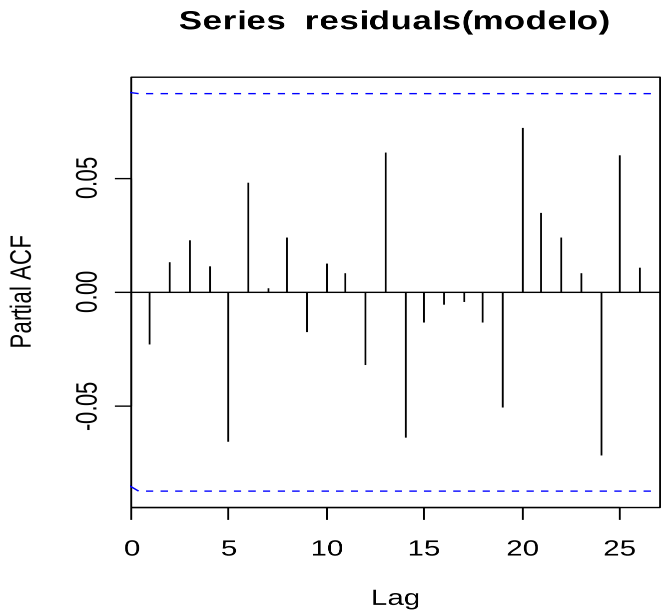
Análise dos Resíduos: Funções de Autocorrelação e Autocorrelação Parcial

2.ii. ESTIMAÇÃO | METODOLOGIA DE BOX E JENKINS

Para avaliar o ajuste adequado do modelo, estudamos as funções de autocorrelação e autocorrelação parcial dos **resíduos**.

Nota: A autocorrelação de ordem 0 sempre é 1, pois é a correlação da série de resíduos contra ela mesma sem nenhuma *lag* de defasagem.

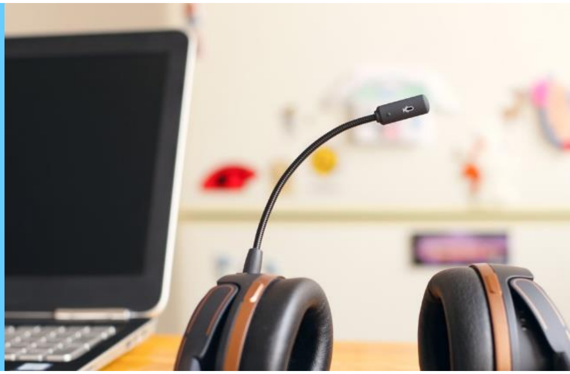




Há evidência de que todas as autocorrelações e todas as autocorrelações parciais são iguais a zero, pois estão dentro da linha pontilhada azul, portanto o modelo está adequado.

Case *Call Center*: Realize as fase de identificação e estimação da série

Com o objetivo de planejar a escala do time de *call center*, uma empresa de serviços gostaria de realizar a previsão da quantidade de recebimento de chamadas na próxima semana. A empresa possui o histórico da quantidade de chamadas recebidas (dias úteis) dos últimos 2 anos.



(a) Faça a análise exploratória da série, e comente seu comportamento.

(b) Construa o gráfico descritivo da série.

(c) Teste se a série é estacionária.

(d) Apresente graficamente a função de autocorrelação (ACF) e a função de autocorrelação parcial (PACF) e discuta sobre a ordem a ser testada no modelo AR.

(e) Ajuste um modelo auto-regressivo AR e avalie a significância dos parâmetros.

(f) Apresente graficamente a função de autocorrelação (ACF) e a função de autocorrelação parcial (PACF) dos resíduos do modelo e discuta se a ordem do AR adotada está adequada.

(g) Uma vez o modelo bem ajustado, apresente a equação do modelo.

(h) Faça a previsão para a próxima semana (dias úteis) e comente os resultados.

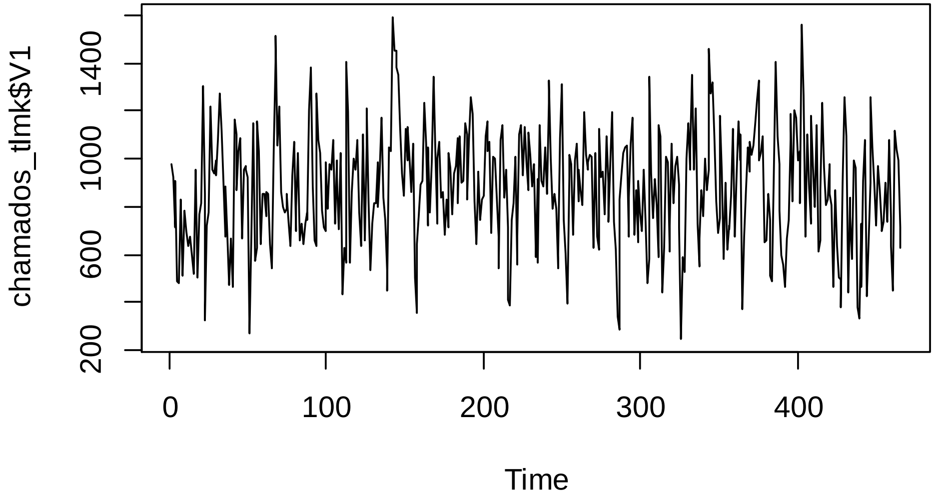
Vamos fazer juntos?

R Studio

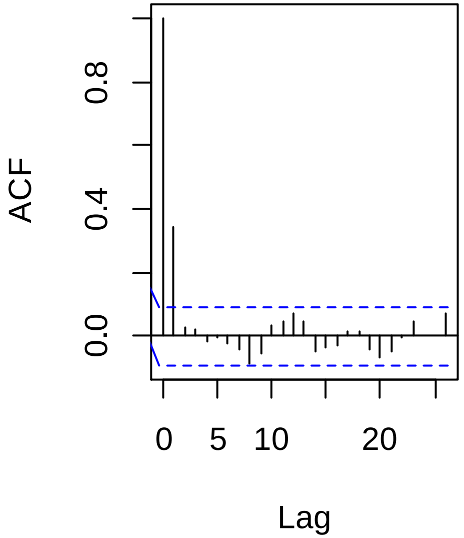
Case *Call Center*: Análise Exploratória

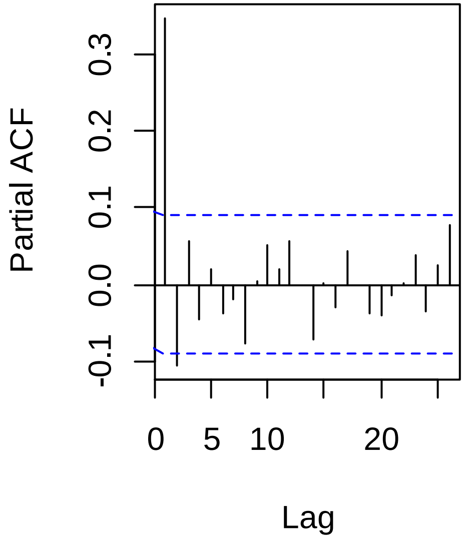
2.ii. ESTIMAÇÃO | METODOLOGIA DE BOX E JENKINS

Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
249.0	711.0	884.0	877.5	1036.5	1591.0



Augmented Dickey-Fuller Test data:
p-value = 0.01
alternative hypothesis: stationary

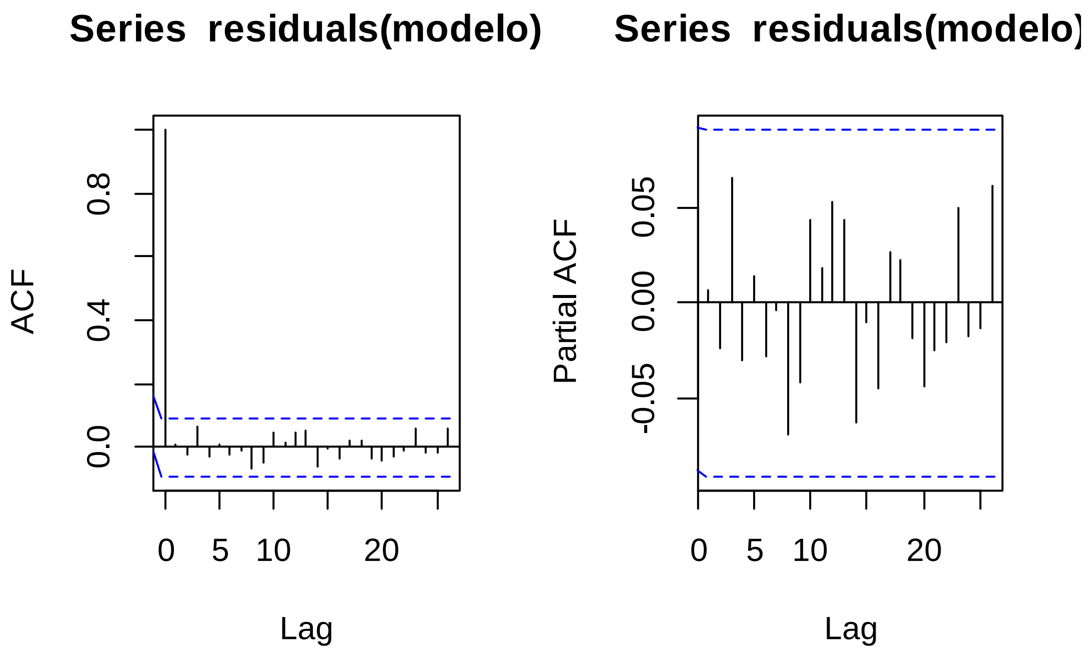




Case *Call Center*: Ajuste do modelo no R

	Estimate	Std. Error	z value	Pr(>|z|)
Ar1	0.381796	0.046054	8.2902	< 2e-16 ***
Ar2	-0.103981	0.046044	-2.2583	0.02393 *
Intercept	877.396897	14.282585	61.4312	< 2e-16 ***

--- Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1



ϕ_2

```
modelo <- arima(chamados_tlmk$V1, order = c(2,0,0), fixed = c(NA,NA,NA))
```

ϕ_1

Adiciona o intercepto

A base de dados está armazenada na variável série

O valor 2 sinaliza que será ajustado um modelo AR(2)

Uma vez o modelo bem ajustado, pode-se obter por meio da equação do modelo a previsão N passos para frente.

$$\hat{Y}_t = \hat{\phi}_0 + \hat{\phi}_1 Y_{t-1} + \hat{\phi}_2 Y_{t-2} + \dots + + \hat{\phi}_p Y_{t-p}$$

$$\hat{\phi}_0 = intercepto\left(1 - \hat{\phi}_1 - \hat{\phi}_2 - \hat{\phi}_2 \dots - \hat{\phi}_p\right)$$

Equação do modelo:

$$\hat{Y}_t = \hat{\phi}_0 + \hat{\phi}_1 Y_{t-1} + \hat{\phi}_2 Y_{t-2}$$

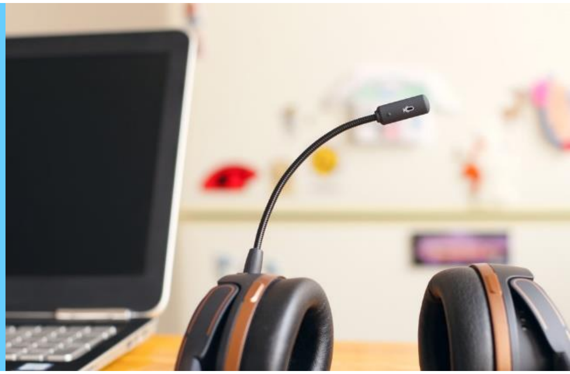
$$\hat{\phi}_0 = intercepto\left(1 - \hat{\phi}_1 - \hat{\phi}_2\right)$$

Projeção N passos para frente:

Com base na equação do modelo, utilizando os dados das *lags* passadas é possível prever passos para frente, ou seja, fazer projeção para o futuro de dados ainda não conhecidos.

Case *Call Center*: Realize as fase de identificação e estimação da série

Com o objetivo de planejar a escala do time de *call center*, uma empresa de serviços gostaria de realizar a previsão da quantidade de recebimento de chamadas na próxima semana. A empresa possui o histórico da quantidade de chamadas recebidas (dias úteis) dos últimos 2 anos.



(a) Faça a análise exploratória da série, e comente seu comportamento.

(b) Construa o gráfico descritivo da série.

(c) Teste se a série é estacionária.

(d) Apresente graficamente a função de autocorrelação (ACF) e a função de autocorrelação parcial (PACF) e discuta sobre a ordem a ser testada no modelo AR.

(e) Ajuste um modelo auto-regressivo AR e avalie a significância dos parâmetros.

(f) Apresente graficamente a função de autocorrelação (ACF) e a função de autocorrelação parcial (PACF) dos resíduos do modelo e discuta se a ordem do AR adotada está adequada.

(g) Uma vez o modelo bem ajustado, apresente a equação do modelo.

(h) Faça a previsão para a próxima semana (dias úteis) e comente os resultados.

Case *Call Center*: Ajuste do modelo no R

Uma vez o modelo bem ajustado, pode-se obter por meio da equação do modelo a previsão 5 passos para frente.

$$\hat{Y}_t = \hat{\phi}_0 + \hat{\phi}_1 Y_{t-1} + \hat{\phi}_2 Y_{t-2}$$

$$\hat{\phi}_0 = intercepto \cdot (1 - \hat{\phi}_1 - \hat{\phi}_2)$$

Equação do modelo AR(2):

$$\hat{Y}_t = 633{,}64 + 0{,}3818 * Y_{t-1} - 0{,}1040 * Y_{t-2}$$

$$\hat{Y}_{467} = 633{,}64 + 0{,}3818 * 627 - 0{,}1040 * 711 = 799{,}1$$

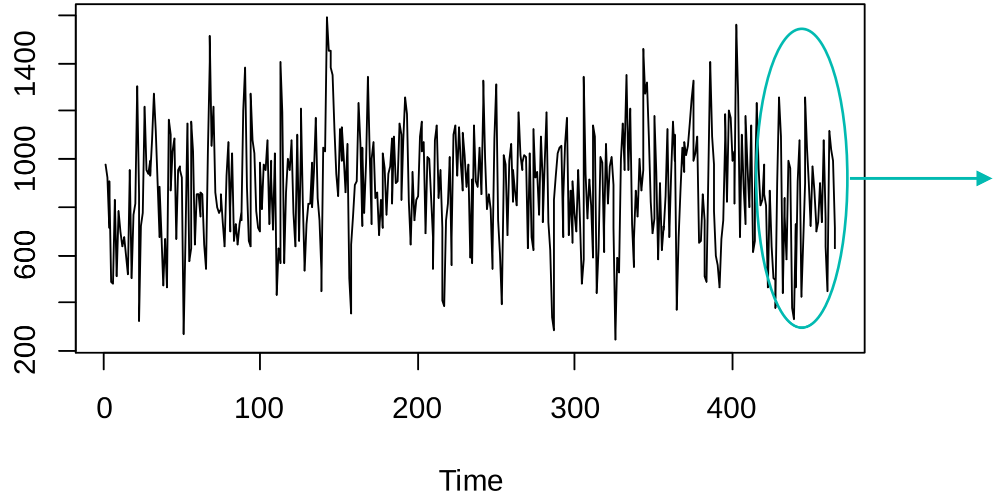
$$\hat{Y}_{468} = 633{,}64 + 0{,}3818 * 799{,}1 - 0{,}1040 * 627 = 873{,}6$$

.....

Valor obtido na predição do passo anterior

Projeção 5 passos para frente:

	Point Forecast	Lo 80	Hi 80	Lo 95	Hi 95
467	799.0986	513.5981	1084.599	362.4632	1235.734
468	873.5395	567.9382	1179.141	406.1626	1340.916
469	884.0657	578.2316	1189.900	416.3328	1351.799
470	880.3441	574.4349	1186.253	412.4963	1348.192
471	877.8287	571.8955	1183.762	409.9443	1345.713



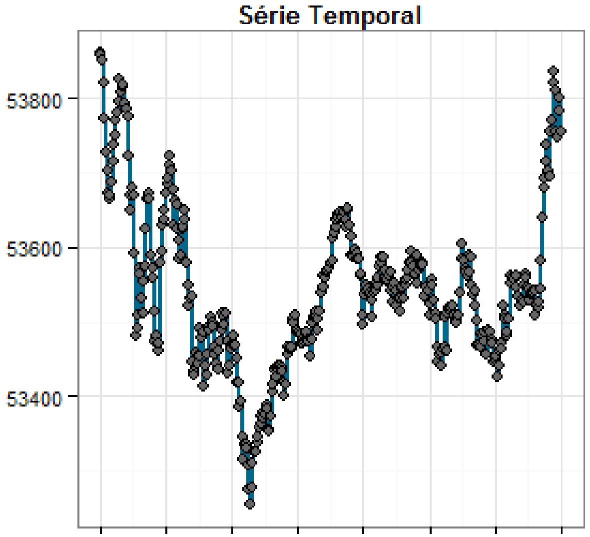
657
902
1256
1049
897
722
972
869
720
699
734
899
734
1076
753
638
445
1118
1039
995
711
627

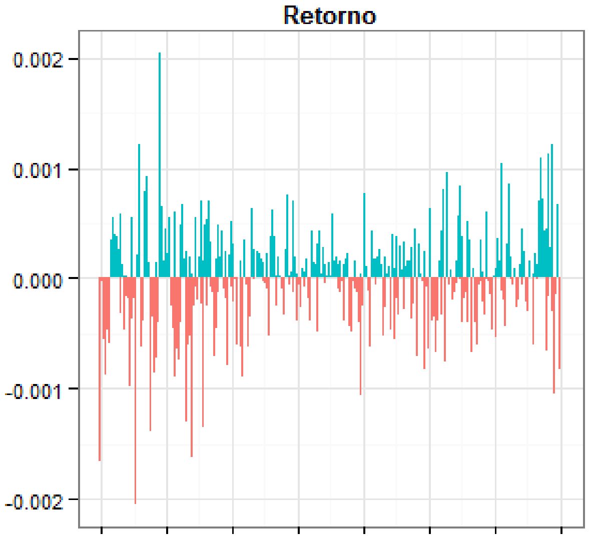
Um dos objetivos no mercado financeiro é avaliação do risco do preço dos ativos (Morettin, 2008).

O risco é frequentemente medido em termos das **variações de preços dos ativos**, denominado **Retorno**. Seja P_t o preço de um ativo no tempo t, a taxa de retorno R_t, é definida por:

$$R_t = \frac{P_t - P_{t-1}}{P_{t-1}}$$

Exemplo: Ações da Petrobrás medidas a cada minuto, com um total de 401 observações:





No mercado, é comum trabalhar com os retornos ao invés do preço dos ativos, que além de serem livres de escala, possuem propriedades interessantes como **obtenção da estacionariedade**, possibilitando, por exemplo, o uso de modelos AR.

http://www.portalaction.com.br/series-temporais/51-retornos

Case Retorno Ações

2. MODELO AUTO-REGRESSIVO | METODOLOGIA DE BOX E JENKINS

Um investidor deseja saber se aconselhável investir em um determinado tipo de ação na próxima semana. Ele tem disponível os dados históricos dos últimos 2 anos. Realize uma análise desta série de retornos e avalie a possibilidade do ajuste de um modelo AR. Identifique sua ordem e realize a previsão para a próxima semana.

(a) Faça a análise exploratória da série, e comente os números na visão de negócios.

(b) Construa o gráfico descritivo da série e comente seu comportamento.

(c) Teste se a série é estacionária. Comente sobre a regra de decisão.

(d) Apresente graficamente a função de autocorrelação (ACF) e a função de autocorrelação parcial (PACF) e discuta sobre a ordem a ser testada no modelo AR.

(e) Ajuste um modelo auto-regressivo AR e avalie a significância dos parâmetros.

(f) Apresente graficamente a função de autocorrelação (ACF) e a função de autocorrelação parcial (PACF) dos resíduos do modelo e discuta se a ordem do AR adotada está adequada.

(g) Uma vez o modelo bem ajustado, apresente a equação do modelo.

(h) Após a próxima semana ter passado, os dados observados foram os seguintes: 0.103906254, 0.188157433, -0.080349801, 0.11583838, 0.121133841. Compare sua predição com os resultados observados. Discuta se o modelo foi bem ajustado.

Modelo AR Incompleto: Ajuste no R

ϕ_2 ϕ_5

```
modelo <- arima(retorno$V1, order = c(5,0,0),  fixed = c(NA,NA,NA,0,NA,0))
```

ϕ_1 ϕ_3 retira o intercepto

A base de dados está armazenada na variável série

O valor 5 sinaliza que será ajustado um modelo AR(5)

ϕ_4

retira a ordem 4

3. Exercícios para casa

3. Exercícios para casa

DATA DE ENTREGA 17/01/2021 | EXERCÍCIOS-CASE

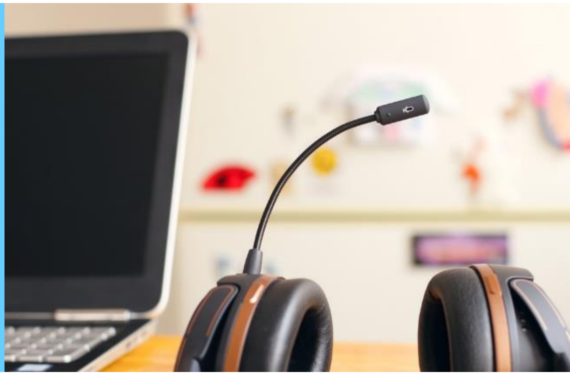
CASE: *Call Center* (5,0 pontos)
CASE: Retorno de Ações (5,0 pontos)

Instruções importantes:

- A lista vale nota (0-10) e deve ser entregue até 17/01/2021. Lista entregue até 24/01/2021 valerá 80% da nota. Posteriormente, não será mais aceita a lista para correção. Não serão aceitas listas parciais.
- O exercício será considerado como "realizado", quando tiver, além das análises, a interpretação do resultados.
- Soluções técnicas "elegantes e mais completas" serão considerados como ponto extra para o aluno (+0,5 na lista geral).
- Caso o aluno tire nota >10, considerando os pontos extras, os pontos extras poderão ser acumulados para listas seguintes, sendo a média geral de todas as listas realizadas no curso, com valor máximo igual a 10.

BOM ESTUDO ☺

Com o objetivo de planejar a escala do time de *call center*, uma empresa de serviços gostaria de realizar a previsão da quantidade de recebimento de chamadas na próxima semana. A empresa possui o histórico da quantidade de chamadas recebidas (dias úteis) dos últimos 2 anos.



Os outputs em R já foram gerados durante a aula. Escreva a interpretação dos resultados dos itens (a)-(h) e responda a pergunta de negócios em (i).

(a) Faça a análise exploratória da série, e comente os números na visão de negócios.
(b) Construa o gráfico descritivo da série e comente seu comportamento.
(c) Teste se a série é estacionária. Comente sobre a regra de decisão.
(d) Apresente graficamente a função de autocorrelação (ACF) e a função de autocorrelação parcial (PACF) e discuta sobre a ordem a ser testada no modelo AR.
(e) Ajuste um modelo auto-regressivo AR e avalie a significância dos parâmetros.
(f) Apresente graficamente a função de autocorrelação (ACF) e a função de autocorrelação parcial (PACF) dos resíduos do modelo e discuta se a ordem do AR adotada está adequada.
(g) Uma vez o modelo bem ajustado, apresente a equação do modelo.
(h) Faça a previsão para a próxima semana (dias úteis) e comente os resultados da estimativa pontual.
(i) Discuta na visão de negócios se você recomendaria fazer a previsão para a próxima semana ou se o ideal seria fornecer a previsão em um período diferente (maior ou menor do que 5 dias).

Case Retorno Ações

MODELO AUTO-REGRESSIVO | EXERCÍCIOS PARA CASA

Um investidor deseja saber se aconselhável investir em um determinado tipo de ação na próxima semana. Ele tem disponível os dados históricos dos últimos 2 anos. Realize uma análise desta série de retornos e avalie a possibilidade do ajuste de um modelo AR. Identifique sua ordem e realize a previsão para a próxima semana.

Os outputs em R já foram gerados durante a aula. Escreva a interpretação dos resultados dos itens (a)-(h) e responda a pergunta de negócios em (i).

(a) Faça a análise exploratória da série, e comente os números na visão de negócios.
(b) Construa o gráfico descritivo da série e comente seu comportamento.
(c) Teste se a série é estacionária. Comente sobre a regra de decisão.
(d) Apresente graficamente a função de autocorrelação (ACF) e a função de autocorrelação parcial (PACF) e discuta sobre a ordem a ser testada no modelo AR.
(e) Ajuste um modelo auto-regressivo AR e avalie a significância dos parâmetros.
(f) Apresente graficamente a função de autocorrelação (ACF) e a função de autocorrelação parcial (PACF) dos resíduos do modelo e discuta se a ordem do AR adotada está adequada.
(g) Uma vez o modelo bem ajustado, apresente a equação do modelo.
(h) Após a próxima semana ter passado, os dados observados foram os seguintes: 0.103906254, 0.188157433, -0.080349801, 0.11583838, 0.121133841. Compare sua predição com os resultados observados. Discuta se o modelo foi bem ajustado.
(i) Discuta na visão de negócios se você recomendaria para um investidor realizar o investimento no ativo em questão.

Referências

- Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control*. San Francisco: Holden-Day.
- Dickey, D. & Fuller, W. A. (1979). *Distribution of the estimates for autoregressive time series with a unit root.* Journal of the American Statistical Association, 74(2), 427-431.
- Morettin, P. A., & Toloi, C. M. de C. (2004). *Análise de séries temporais*. São Paulo: Edgard Blucher.
- Morettin, P. A. (2008). *Econometria financeira: um curso em séries temporais financeiras*. São Paulo: Edgard Blücher.