

**UNIVERSIDADE ESTADUAL PAULISTA**  
**"JÚLIO DE MESQUITA FILHO"**  
**CAMPUS DE GUARATINGUETÁ**

**WILLIAN EIJI KANASHIRO**

**Modelo de Previsão SARIMA Aplicado na Indústria Automobilística Brasileira: Previsão de  
demanda de Veículos para Pessoas Com Deficiência e Taxistas**

Guaratinguetá

2019

**Willian Eiji Kanashiro**

**Modelo de Previsão SARIMA Aplicado na Indústria Automobilística Brasileira: Previsão de demanda de Veículos para Pessoas Com Deficiência e Taxistas**

Trabalho de Graduação apresentado ao Conselho de Curso de Graduação em Engenharia de Materiais da Faculdade de Engenharia do Campus de Guaratinguetá, Universidade Estadual Paulista, como parte dos requisitos para obtenção do diploma de Graduação em Engenharia de Materiais .

Orientador: Profº Dr. Aneirson Francisco da Silva

Guaratinguetá  
2019

K16m	<p>Kanashiro, Willian Eiji</p> <p>Modelo de previsão SARIMA aplicado na indústria automobilística brasileira: previsão de demanda de veículos para pessoas com deficiência e taxistas / Willian Eiji Kanashiro – Guaratinguetá, 2019.</p> <p>56 f. : il.</p> <p>Bibliografia: f. 55</p> <p>Trabalho de Graduação em Engenharia de Materiais – Universidade Estadual Paulista, Faculdade de Engenharia de Guaratinguetá, 2019.</p> <p>Orientador: Prof. Dr. Aneirson Francisco da Silva</p> <p>1. Indústria automobilística. 2. Deficientes físicos – Transporte. 3. Planejamento empresarial. I.Título.</p> <p>CDU 67.01</p>
------	--

Luciana Máximo

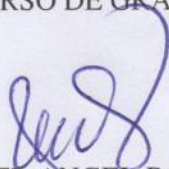
Bibliotecária CRB-8/3595

**UNIVERSIDADE ESTADUAL PAULISTA**  
**"JÚLIO DE MESQUITA FILHO"**  
**CAMPUS DE GUARATINGUETÁ**

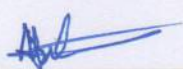
**WILLIAN EIJI KANASHIRO**

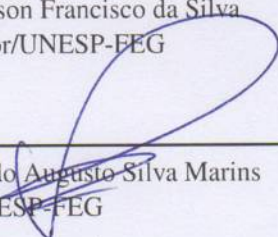
ESTE TRABALHO DE GRADUAÇÃO FOI JULGADO ADEQUADO COMO PARTE DO  
REQUISITO PARA A OBTENÇÃO DO DIPLOMA DE "GRADUANDO EM  
ENGENHARIA DE MATERIAIS "

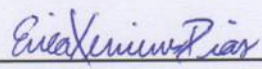
APROVADO EM SUA FORMA FINAL PELO CONSELHO DE CURSO DE GRADUAÇÃO EM  
ENGENHARIA DE MATERIAIS

  
Profº Dr. MIGUEL ANGEL RAMIREZ GIL  
Coordenador

**BANCA EXAMINADORA:**

  
Profº Dr. Aneirson Francisco da Silva  
Orientador/UNESP-FEG

  
Profº Dr. Fernando Augusto Silva Marins  
UNESP-FEG

  
Profº Dr. Erica Ximenes Dias  
Membro Externo

Novembro , 2019

## **DADOS CURRICULARES**

### **WILLIAN EIJI KANASHIRO**

**NASCIMENTO** 04/10/1992 - São Paulo / SP

**FILIAÇÃO** Jorge Kanashiro  
Regina Matiko Kanashiro

**2012 / 2019** Formação acadêmica (Bacharel em Engenharia de Materiais)  
Universidade Estadual Paulista "Júlio de Mesquita Filho"

À minha família e os meus amigos, da Republica Vamointão,  
que me apoiaram em todo esse tempo, pois sem eles,  
não estaria onde cheguei .

## RESUMO

A produção de automóveis no Brasil vem apresentando um crescimento muito significativo nos últimos anos devido à grande demanda no território nacional. Para suprir a necessidade deste tipo de mercado, é necessário um bom planejamento para que as montadoras mantenham a competitividade no mercado, e um dos setores de automóveis que vem apresentando uma grande expansão de demanda é a venda de veículos para Pessoas com Deficiência (PCD) e Taxistas. Para manter a competitividade neste mercado é importante a escolha de estratégias mais eficazes. Para isso, é importante a avaliação e estudo por meio de previsão de séries temporais. Este trabalho tem como objetivo aplicar o modelo de previsão de Médias Móveis Integradas autoregressivas Sazonais (SARIMA) em uma série histórica mensal referente ao faturamento de veículos para Pessoas com Deficiência e Taxistas entre o período de Fevereiro de 2011 a Julho de 2019. Foi possível obter o modelo SARIMA (2,2,4)x(0,2,2), sendo o mais adequado para a previsão.

**PALAVRAS-CHAVE:** SARIMA. Indústria automobilística Brasileira. Pessoa com deficiência. Táxi. Série Temporal.

## **ABSTRACT**

The automobile production in Brazil has been presenting a very significant growth in the last years due to the great demand in the national territory. To meet the need for this type of market, good planning is needed for automakers to remain competitive in the market, and one of the auto sectors that has been experiencing a major expansion in demand is the sale of vehicles for People with Disabilities (PWD) and taxi drivers. To remain competitive in this market, it is important to choose the most effective strategies. For this, it is important to evaluate and study through the prediction of time series. The objective of this paper is to apply the Seasonal Autoregressive Integrated Moving Average (SARIMA) forecasting model in a monthly historical series regarding the invoicing of vehicles for People with Disabilities and Taxi drivers from February 2011 to July 2019. It was possible to obtain the SARIMA (2,2,4) x (0,2,2) model, being the most suitable for the forecast.

**KEYWORDS:** SARIMA. Brazilian automobile industry. People with Disabilities. Taxi. Time series.

## LISTA DE ILUSTRAÇÕES

Figura 1	Número de carros 0 km vendidos com isenção para Pessoas com Deficiência (PCD) e familiares	13
Figura 2	Número de publicações encontradas as palavras-chaves “SARIMA” e “forecast” entre o período de 1990 a 2020.	15
Figura 3	Número de publicações entre países com mais publicações, localizados com as palavras chaves “SARIMA” e “forecast”, o período de 1990 a 2020.	16
Figura 4	Relação das áreas de conhecimento encontradas pelo relatório da base Scopus.	17
Figura 5	Série histórica de licenciamentos de automóveis entre 1957 até 2019 (ANFAVEA, 2019).	20
Figura 6	Série não estacionária em diferentes níveis e inclinações.	21
Figura 7	Gráfico da função de correlação, com uma confiança de 5%, para viscosidade obtida da Tabela I	23
Figura 8	Função de correlação para preço do estoque de mercado de alimentos	24
Figura 9	Representação de um processo estocástico com um conjunto de séries possíveis.	25
Figura 10	Exemplo de um ruído branco.	26
Figura 11	Representação de um círculo unitário.	27
Figura 12	Histórico de faturamento de veículos para pessoas com deficiência (PCD) e clientes Taxistas.	34
Figura 13	Série temporal após a primeira diferença	36
Figura 14	Série temporal após a segunda diferença	38
Figura 15	Periodograma para a série original	41
Figura 16	Periodograma para a série com uma diferença sazonal	42
Figura 17	Periodograma para série com duas diferenciações sazonais	43
Figura 18	Janela do gretl para incluir os <i>inputs</i> do modelo	45
Figura 19	Modelo ajustado	47
Figura 20	Resíduos do modelo ajustado	48
Figura 21	Desenvolvimento do modelo SARIMA (2,2,4)x(0,2,2) sobre a série temporal completa	49
Figura 22	Desenvolvimento do modelo Sarima(2,2,4)x(0,2,2) sobre a série temporal entre julho de 2017 até Setembro de 2019	50
Figura 23	Sobreposição entre a previsão e os dados reais para Agosto e Setembro de 2020	51
Figura 24	Previsão estendida para Fevereiro de 2020	53

## LISTA DE TABELAS

Tabela 1 – Tabela de leituras de viscosidades . . . . .	22
Tabela 2 – Teste Dickey-fuller para a série temporal não diferenciada . . . . .	35
Tabela 3 – Teste Dickey-fuller para a série temporal diferenciada . . . . .	37
Tabela 4 – Teste Dickey-fuller para a série temporal diferenciada pela segunda vez . . . . .	39
Tabela 5 – Correlograma da série temporal diferenciada pela segunda vez . . . . .	40
Tabela 6 – Correlograma da série temporal com duas diferenciações sazonais . . . . .	44
Tabela 7 – Avaliação do modelo estimado SARIMA (5, 2, 4) x (1, 2, 4) . . . . .	46
Tabela 8 – Avaliação do modelo ajustado SARIMA (2, 2, 4) x (0, 2, 2) . . . . .	48
Tabela 9 – Avaliação do modelo ajustado SARIMA (2, 2, 4) x (0, 2, 2) para a série reduzida . . . . .	50
Tabela 10 – Faturamento entre os meses de Agosto e de Setembro para o ano de 2019 . . . . .	51
Tabela 11 – Dados de previsão do modelo . . . . .	52
Tabela 12 – Histórico de faturamento para pessoas com deficiência e taxistas entre o período de Fevereiro de 2011 a Julho de 2019 . . . . .	57

## LISTA DE ABREVIATURAS E SIGLAS

ANFAVEA	- Associação Nacional dos Fabricantes de Veículos Automotores
ARIMA	- <i>Autoregressive Integrated Moving Average</i>
ADF	- <i>Augmented Dickey–Fuller test</i>
FAC	- Função de autocorrelação
FACP	- Função de autocorrelação parcial
ICMS	- Imposto sobre Operações relativas à Circulação de Mercadorias e sobre Prestações de Serviços
IPI	- Imposto sobre Produtos Industrializados
IPVA	- Imposto sobre Propriedades de Veículos Automotores
PCD	- Pessoa Com Deficiência
SARIMA	- <i>Seasonal Autoregressive Integrated Moving Average</i>
SUV	- <i>Sport Utility Vehicle</i>

## LISTA DE SÍMBOLOS

$\rho$	Letra Grega Ró
$\mu$	Letra Grega Mi
$\gamma$	Letra Grega Gama
$\varepsilon$	Letra Grega Épsilon
$\theta$	Letra Grega Teta
$\phi$	Letra Grega Fi
$\psi$	Letra Grega Psi
$\sigma$	Letra Grega Sigma

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO</b>	<b>12</b>
1.1	QUESTÕES DA PESQUISA E OBJETIVOS	13
1.1.1	Objetivos Gerais	13
1.1.2	Objetivos Específicos	14
1.2	DELIMITAÇÃO DA PESQUISA	14
1.3	JUSTIFICATIVA	14
1.4	MATERIAIS E MÉTODOS	17
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA</b>	<b>19</b>
2.1	A INDÚSTRIA AUTOMOBILÍSTICA NO BRASIL	19
2.2	SÉRIES TEMPORAIS	20
2.2.1	Processos Estocásticos	24
2.3	RUÍDO BRANCO	25
2.3.1	CÍRCULO UNITÁRIO	26
2.4	MÉDIAS MÓVEIS	28
2.4.1	MÉDIAS MÓVEIS DE ORDEM $q$ – MA( $q$ )	28
2.4.2	Processo autor regressivo de ordem $p$ – AR ( $p$ )	28
2.4.3	Modelos auto regressivos de média móveis – ARMA ( $p,q$ )	29
2.5	PROCESSOS NÃO ESTACIONÁRIOS	30
2.5.1	Construção do modelo ARIMA ( $p,d,q$ )	30
2.6	Processos Sazonais	31
<b>3</b>	<b>DESCRIÇÃO DO PROBLEMA E MODELAGEM</b>	<b>33</b>
3.1	MODELAMENTO AUTO REGRESSIVO	33
3.2	ESTUDO E ANÁLISE DA SÉRIE	33
3.3	– MODELO SARIMA ( $p, d, q$ ) X ( $P, D, Q$ ): AJUSTE E VALIDAÇÃO DO MODELO	39
3.4	A previsão	49
<b>4</b>	<b>CONCLUSÕES E RECOMENDAÇÕES PARA PESQUISAS FUTURAS</b>	<b>54</b>
4.1	Respostas das questões da pesquisa e recomendações para estudos futuros	54
4.2	recomendações para pesquisas futuras	54
	<b>REFERÊNCIAS</b>	<b>55</b>
	<b>ANEXO A – DADOS DO NÚMERO DE FATURAMENTO DE VEÍCULOS</b>	
	<b>PARA PESSOAS COM DEFICIÊNCIA E TAXÍSTAS ENTRE</b>	
	<b>O PERÍODO DE FEVEREIRO DE 2011 A JULHO DE 2019</b>	<b>57</b>

## 1 INTRODUÇÃO

De acordo com o Ministério da Indústria, Comércio Exterior e Serviços (MDIC), a indústria automobilística brasileira apresenta uma grande importância para a economia nacional, representando 22,0% do PIB Industrial e 4,0% do PIB total, em 2015, e estando em 10º lugar entre os maiores produtores do mundo, ficando atrás da China, Estados Unidos, Japão, Alemanha, Índia, Coreia do Sul, México, Espanha e Canadá (MDIC, 2019).

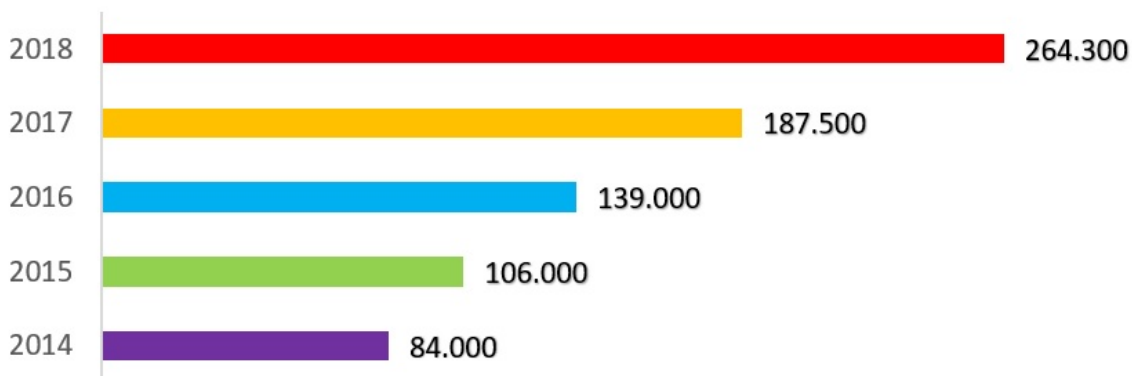
Por ser uma indústria de muita relevância para a economia nacional, desencadeia uma competitividade entre as montadoras atuantes no mercado brasileiro, e com isso, gerando um grande impacto na infraestrutura brasileira. Para isso, é importante que tenham um bom planejamento da produção com o uso de séries temporais para que as montadoras consigam tomar decisões em uma forma mais segura (MARTIN *et al.*, 2016), evitando diversos tipos de problemas na produção, estoque e transporte que poderiam custar milhões para as montadoras e empresas terceirizadas.

A indústria automobilística, por apresentar uma demanda muito grande, acaba trabalhando com outras empresas para suprir a demanda do mercado, ajudando a diminuir custos e atender a demanda. As empresas terceirizadas funcionam como fornecedoras de peças e acessórios essenciais para a produção de veículos automotores e peças reservas. Como são empresas que estão muito ligadas com o mercado automobilístico, precisam, também, apresentar um bom controle sobre a produção, usando, também, o método de séries temporais (VARGAS; CORTÉS, 2017).

Dentro da indústria automobilística brasileira, um setor vem chamando muita atenção pela sua demanda que vem crescendo, nos últimos 5 anos, que é a venda de automóveis para Pessoas com Deficiência (PCD) (REAÇÃO, 2019). Comparando os dados de 5 anos, entre 2014 e 2018, as vendas de automóveis para pessoas com deficiência triplicaram e isso vem chamando atenção do governo federal, pois as vendas de automóveis para pessoas com deficiência apresentam benefícios fiscais, como isenção de IPI, ICMS, IPVA e de rodízios municipais, fazendo com que o governo deixe de arrecadar bilhões de reais nesses últimos anos. A Figura 1 ilustra o crescimento deste tipo de vendas.

Figura 1 – Número de carros 0 km vendidos com isenção para Pessoas com Deficiência (PCD) e familiares

## Número de carros 0 km vendidos com isenção para pessoas com deficiência e familiares



Fonte: Revista Nacional de reabilitação Reação REAÇÃO (2019).

Martin *et al.* (2016) afirma que aplicação de métodos de previsão de demanda é um método muito importante para a escolha de estratégias quando se trata de demandas elevadas. Nisso, segundo Subramaniyan *et al.* (2018), é muito importante a aplicação do estudo de séries temporais em linhas de produção, pois em muitos casos, apresentam regiões onde ocorrem o chamado efeito de gargalo (*bottleneck*), que são processos que limitam a produtividade da linha como todo. Estes processos podem ser aperfeiçoados com a aplicação do estudo de previsão com a constante coleta de dados para o monitoramento da linha de produção, pois as regiões onde ocorrem o efeito gargalo podem mudar com o decorrer do tempo, ou seja, podem surgir novos problemas em outros processos da linha de produção.

Antes de qualquer tomada de decisão, é necessário o entendimento do cenário em questão para coleta de dados mais confiáveis (BERTRAND; FRANSOO, 2002). Segundo o autor, o levantamento por meio de um modelo de pesquisa quantitativa é uma ferramenta importante para a compreensão real de processos operacionais que influenciarão nas tomadas de decisões finais.

Chaiwuttisak (2018) afirma que usando a metodologia Box – Jenkins, em séries temporais mensais, com a aplicação do modelo Auto Regressivo de Médias Móveis Integradas (ARIMA), é uma forma de obter previsões de alta confiabilidade, no que seria o foco deste trabalho, com o uso de dados reais do mercado automobilístico brasileiro para explorar o mercado atual e contribuir com estudos futuros.

## 1.1 QUESTÕES DA PESQUISA E OBJETIVOS

### 1.1.1 Objetivos Gerais

Este trabalho tem como objetivo aplicar métodos auto regressivos sazonais para prever a demanda de veículos PCD de uma montadora.

### 1.1.2 Objetivos Específicos

- Como realizar a previsão da demanda de veículos PCD?
- O modelo de médias móveis integradas autoregressivas sazonais (SARIMA) gerará uma previsão confiável ?

## 1.2 DELIMITAÇÃO DA PESQUISA

Este tem como visão a aplicação de modelos auto-regressivos para a elaboração de previsão de demanda.

## 1.3 JUSTIFICATIVA

A aplicação de métodos de previsão auto – regressivos com séries temporais é vital para o bom planejamento de produção em mercados onde existe uma grade demanda. Este tipo de análise tem como princípio identificar a natureza do fenômeno observado e elaborar uma projeção para um futuro próximo, com a identificação de tais componentes: tendência, regularidade, sazonalidade e irregularidade (CHAIWUTTISAK, 2018). Todos os tipos de séries temporais podem conter todos ou alguns dos componentes citados.

Apesar de o setor automobilístico apresentar quantidades consideráveis de investimentos em diversos gêneros, muitos países em desenvolvimento apresentam indústrias onde a produção é norteadada pelo estado da economia e da demanda atual, não havendo um método de previsão dos mesmos (SUEPTAETRAKUN; TANGTHONG, 2018), podendo trazer grandes prejuízos e desperdícios de recursos importantes.

A metodologia de análise (*Autoregressive Integrated Moving Average - ARIMA*) possibilita uma melhor compreensão de eventos passados e com isso, proporciona a elaboração de modelos de previsão mais precisos (STORTI *et al.*, 2018). Segundo Liu (1992), existe diversos tipos de análise de séries temporais, mas o mais bem sucedido é o modelo Box-Jenkins ARIMA, mencionado anteriormente.

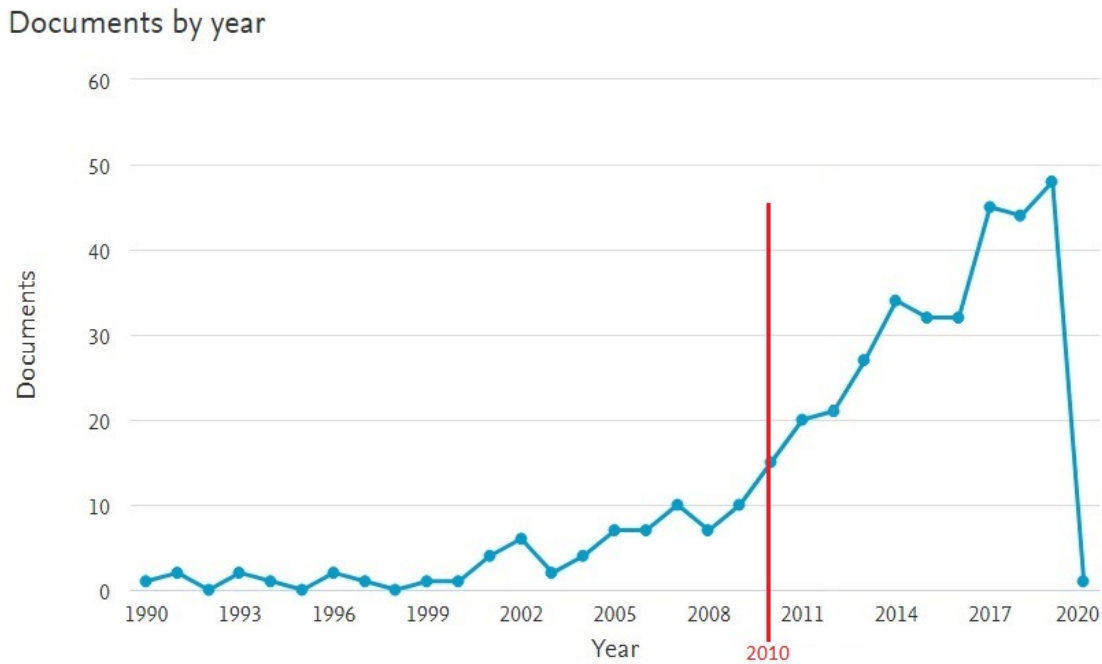
Liu (1992) cita os pontos mais relevantes pelo motivo de o modelo de Box-Jenkins ser o mais usado:

- Capacidade de representar grande quantidade de processos com modelos parcimoniosos;
- Possibilidade de expansão para modificar a modelagem seguindo a presença de fenômenos externos;
- Por ser um modelo com uma presença sólida nos últimos anos.

Storti *et al.* (2018) confirmam que a aplicação de métodos de previsão auto regressivo integrado de média móvel (*Sazonal Autoregressive Integrated Moving Average- SARIMA*) é muito válido para proporcionar ganhos em linhas de produção onde apresenta uma grande demanda no mercado, identificando problemas em potencial e diminuindo desperdício de recursos que poderiam correr com a não aplicação do modelo de BOX-Jenkins.

Foram realizadas algumas pesquisas na base Scopus no período de 1990 a 2020 usando as palavras-chaves “SARIMA” e “forecast” para verificar a importância do estudo de previsão por séries temporais, verificando os números de publicações existentes. A [Figura 2](#) apresenta o número de publicações por ano sobre o tema deste trabalho, e é possível observar o crescimento do interesse sobre o assunto a partir do ano de 2008. Entre o período de 2010 a 2019, foram publicados 318 estudos, sendo que já existem estudos a serem publicados no ano de 2020.

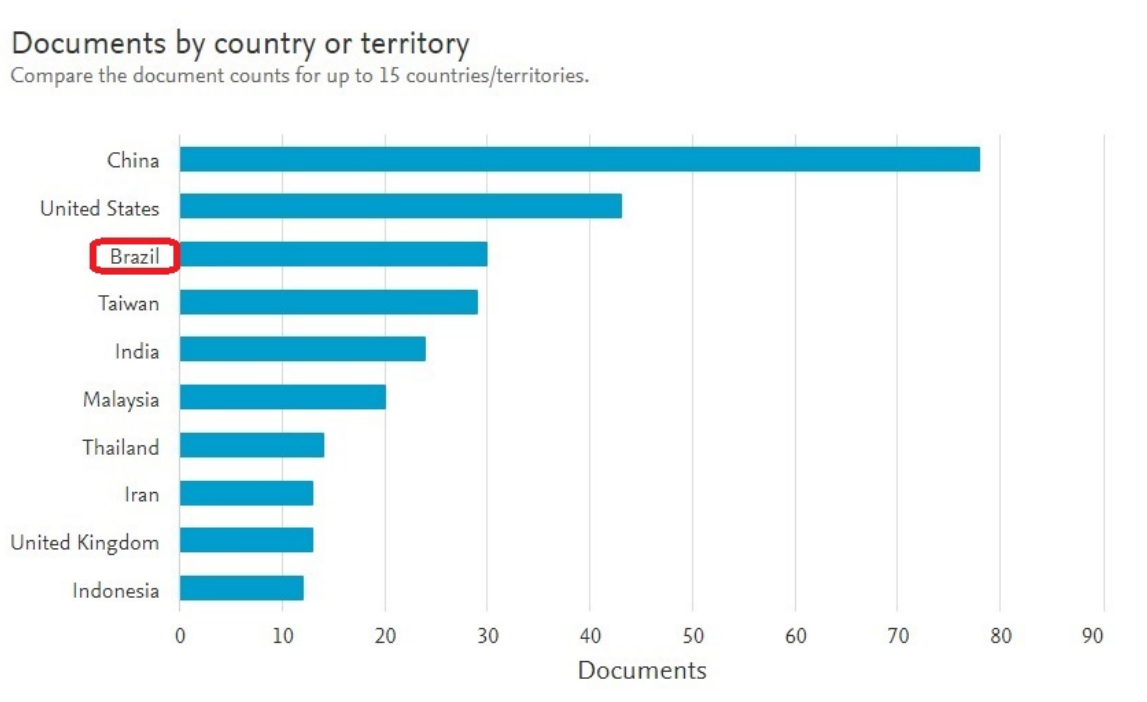
Figura 2 – Número de publicações encontradas as palavras-chaves “SARIMA” e “forecast” entre o período de 1990 a 2020.



Fonte: Adaptado pelo relatório de buscas do Scopus (2019)

Entre os países com maiores números de publicações a respeito do assunto, o Brasil está em terceiro lugar, com 30 estudos citados na base Scopus, publicados entre o período de 2001 a 2019, ficando atrás dos Estados Unidos, em segundo, com 43 estudos publicados, e da China, em primeiro, com 78 estudos publicados. A [Figura 3](#) apresenta um rank dos 10 primeiros países com maiores números de publicações sobre o assunto.

Figura 3 – Número de publicações entre países com mais publicações, localizados com as palavras chaves “SARIMA” e “forecast”, o período de 1990 a 2020.

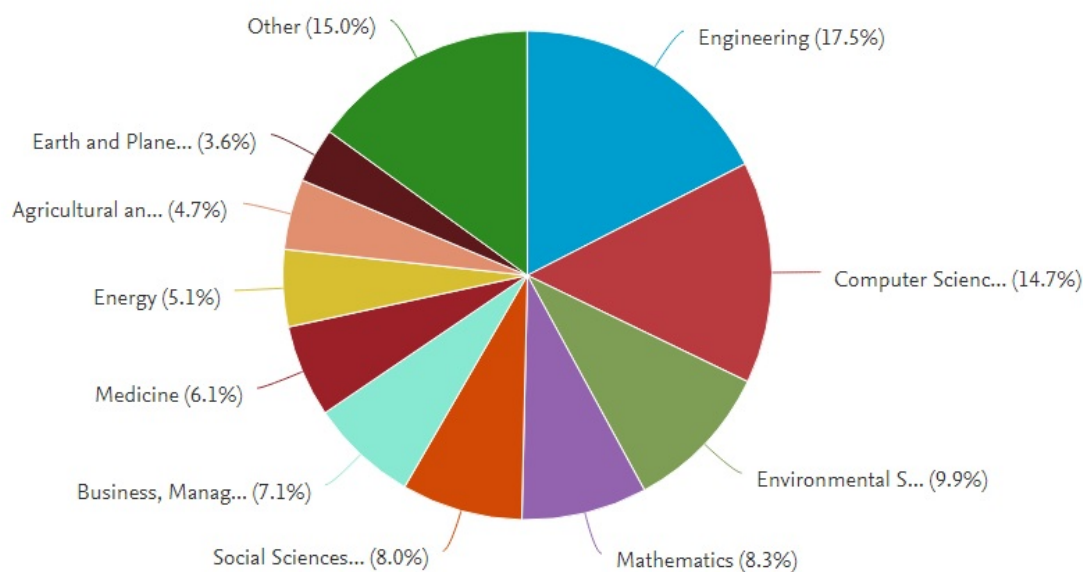


Fonte: Adaptado pelo relatório de buscas do Scopus (2019)

Por meio da **Figura 4** é possível ter uma visão das áreas de atuações referente aos estudos encontrados pelo relatório fornecido pela base Scopus, e a maior parte dos estudos está relacionada com Engenharia, representando 17,5% das publicações, seguida pela Ciência da Computação, em segundo lugar, representando 14,7% das publicações, e Ciência Ambiental, em terceiro, com 9,9% de representatividade nos estudos sobre o assunto. Com isso pode-se verificar a grande utilidade da ferramenta de previsão de demanda na Engenharia, podendo ser aplicada em muitos setores industriais e econômicos para resolução de diversos tipos de problemas.

Figura 4 – Relação das áreas de conhecimento encontradas pelo relatório da base Scopus.

#### Documents by subject area



Fonte: Adaptado pelo relatório de buscas do Scopus (2019)

## 1.4 MATERIAIS E MÉTODOS

Por ser um trabalho onde busca modelar um método de previsão para a resolução de problemas reais e melhorar a eficiência de operações existentes, de acordo com [Bertrand e Fransoo \(2002\)](#), este trabalho é classificado como uma pesquisa aplicada, pois procura buscar modelos que explique o comportamento de eventos reais e orientações para a tomada de decisões frente a um problema, com objetivos empíricos descritivos, pois visa identificar as relações de causa e efeito que podem existir em processos reais. A forma de abordar o problema é quantitativa, sendo o método de pesquisa a modelagem e simulação.

O modelo que será estudado é denominado SARIMA (Médias Móveis Integradas Auto Regressivas Sazonais), uma variação do método ARIMA (Médias Móveis Integradas Auto Regressivas). Os modelos a serem analisados serão testados e validados pela ferramenta GRETL, um software de código aberto.

A metodologia proposta por Box e Jenkins foi o ponto de partida para o surgimento de diversos tipos de modelos e métodos de previsão atualmente existentes, sendo assim, a metodologia desenvolvida pelos pesquisadores são comumente denominados de Box-Jenkins ARIMA ou Modelos de Box-Jenkins ([LIU, 1992](#)).

Box e Jenkins elaboraram um procedimento para a criação de modelos de séries temporais:

1. Identificação: analisar o comportamento da série temporal para buscar relações com modelos específicos;

2. Estimativa: parâmetros que podemos estimar para identificar o modelo correto com dados disponíveis;
3. Checagem final: examinar os modelos verificados e analisar os resíduos dos ajustes para verificar se os modelos são confiáveis para serem aplicados em prática.

Se os modelos propostos passarem dos três passos, os mesmos poderão ser usados para previsão de dados ou simplesmente usados para ter uma melhor compreensão dos fenômenos estudados (LIU, 1992).

## 2 FUNDAMENTAÇÃO TEÓRICA

### 2.1 A INDÚSTRIA AUTOMOBILÍSTICA NO BRASIL

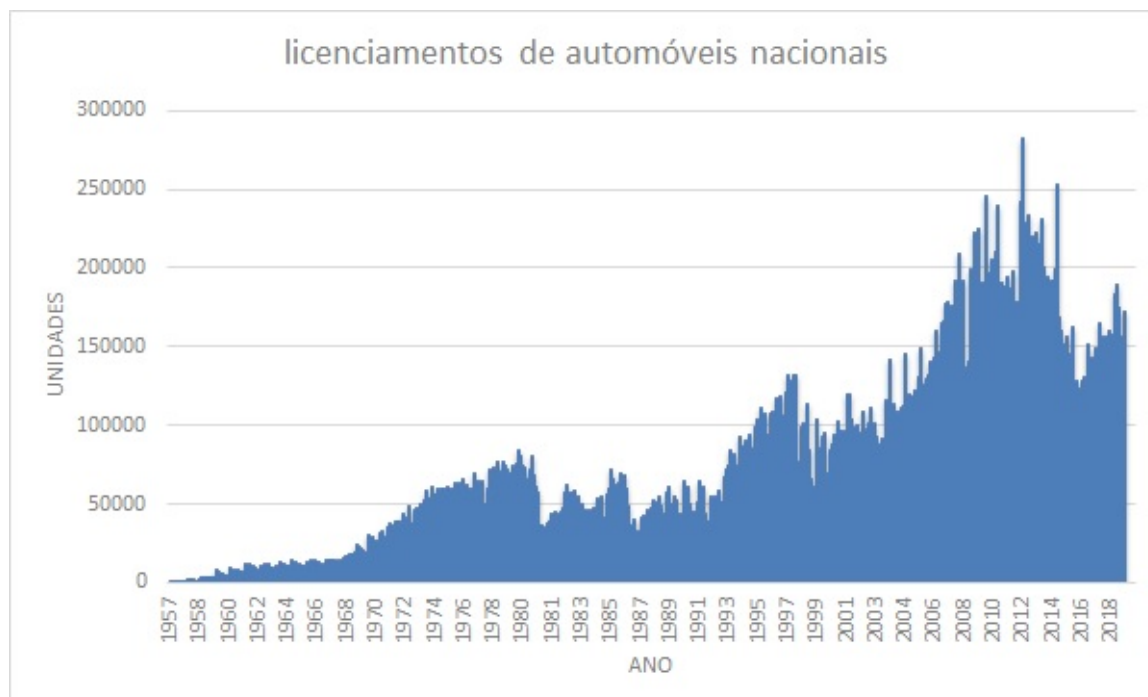
O Brasil teve o início de suas atividades do ramo automobilístico no início do século XX, onde os veículos eram trazidos para o Brasil por intermédio de agentes de importação (ANFAVEA, 2019). Na década de 1920, Ford e a Companhia Geral de Motores S.A., que logo depois teve o nome alterado para *General Motors of Brazil* S.A., se instalam no Brasil. Com o advento da Ford e da GM no Brasil, surgiu a oportunidade para outras marcas atuarem no mercado brasileiro, expandindo a produção nacional para caminhões e máquinas agrícolas.

O bom desempenho das montadoras, que iniciaram os seus negócios como importadoras, apenas, despertou um forte interesse de fortalecer a produção nacional de veículos, nisso, em 1952, o então presidente Getúlio Vargas, proibiu a importação de autopeças para estimular a produção nacional. Com isso, outras duas grandes montadoras se instalaram no Brasil, na mesma década, se estabelecendo como Volkswagen do Brasil Ltda. e a Mercedes – Bens do Brasil.

Outras montadoras se instalaram no Brasil, incentivadas pela grande expansão da produção de veículos automotores de diversas categorias como grandes marcas japonesas que se instalaram na década de 90 e nos anos 2000. Segundo a ANFAVEA (Associação Nacional dos Fabricantes de Veículos Automotores) o Brasil possui 26 montadoras instaladas, sendo a maior parte da produção direcionada para automóveis, seguido pela produção de comerciais leves, caminhões, ônibus e por último, máquinas agrícolas (ANFAVEA, 2019).

A **Figura 5** mostra uma série histórica de licenciamentos de veículos nacionais a partir de 1957, disponibilizados pela ANFAVEA:

Figura 5 – Série histórica de licenciamentos de automóveis entre 1957 até 2019 (ANFAVEA, 2019).



Fonte: Associação Nacional dos Fabricantes de Veículos Automotores, (2019).

Com a [Figura 5](#), nota-se como alguns eventos históricos afetaram a indústria automobilística brasileira, como a crise do petróleo de 1979 até 1981, o aumento do preço do barril do petróleo causou uma diminuição significativa na compra de automóveis dentro do período, até sofrer um crescimento considerável com o Plano Real em 1994, implantado durante o governo de Itamar Franco. A indústria automobilística brasileira teve que enfrentar mais outros desafios como a crise econômica dos Tigres Asiáticos, em 1997, e a mais recente crise de 2014, com um foco mais interno, ambos podendo ser notados na [Figura 5](#).

Uma nota publicada no anuário de 2019 da ANFAVEA o setor vem apresentando bons resultados e os anos seguintes se mostram promissores para o segmento. O ano de 2018 apresentou um aumento de 13,8% de licenciamentos em relação ao ano de 2017 (ANFAVEA, 2019), e para os próximos anos é esperado um incentivo na modernização da frota de automóveis e seguir com as etapas do Plano de Controle de Emissões, iniciada em 1986.

## 2.2 SÉRIES TEMPORAIS

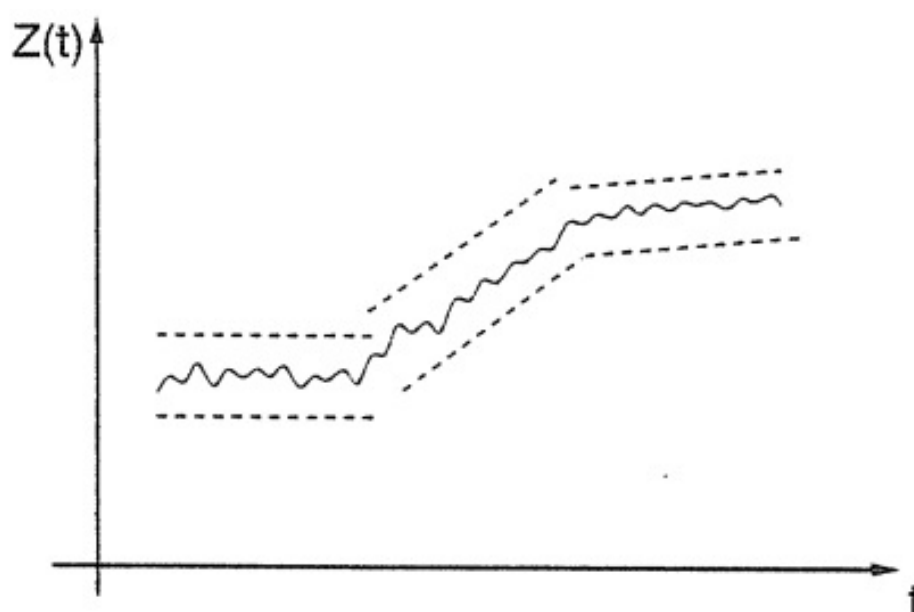
Segundo [Souza \(1981\)](#) série temporal é definida pela coleta de dados por meio de observações de um fenômeno, podendo ser natural ou não, onde são organizados seguindo uma unidade de medida de tempo, sendo cada observação organizada em intervalos de tempos idênticos. Ao organizar graficamente os dados coletados, pode-se dizer que a variável tempo é independente e as observações depende de diversas variáveis que apresentam influência sobre o fenômeno observado. O modo mais comum de se representar uma série é usando a notação  $Z$ , onde cada observação é coletada em um momento “ $t$ ” representada por  $Z_t$ , sendo  $t = 1, 2, 3, \dots, n$ .

Para iniciar uma análise em séries temporais, supõe-se que toda série temporal é estacionária (MORETTIN; TOLOI, 2006), ou seja, a série temporal varia em torno de uma média, mas grande parte das séries temporais não apresenta tal comportamento, pois há a influência de outros fatores que alteram o desenvolvimento da série temporal, fazendo com que a série apresente linhas de tendências crescentes ou decrescentes (tendência linear), séries não - estacionárias.

Séries Temporais podem apresentar um comportamento estacionário por um período muito longo, como por exemplo, um registro de temperaturas de uma cidade ao longo de décadas, ou pode ser estacionária por períodos curtos, alterando de nível muito frequentemente. Para esses tipos de séries temporais podem ser representados com muita precisão pelos métodos auto regressivos, ARIMA. Séries que se desenvolvem sobre um nível são chamados de homogêneos, podendo mudar de inclinação, mas sempre ao redor de um nível (MORETTIN; TOLOI, 2006).

A Figura 6 aborda um exemplo de uma série não estacionária:

Figura 6 – Série não estacionária em diferentes níveis e inclinações.



Fonte: Morettin e Toloi (2006).

Então podemos dizer que qualquer par de observações,  $y_t$  e  $y_{t+k}$ , uma seguida da outra, devem ser iguais para intervalos de tempos idênticos,  $t$  e  $t + k$ , sendo " $t$ " o momento da primeira observação e  $k$ , intervalo de tempo até a segunda observação. A Tabela I apresenta leituras de viscosidade por tempo:

Tabela 1 – Tabela de leituras de viscosidades .

Período	leitura	Período	leitura	Período	leitura	Período	leitura
1	86.7418	26	87.2397	51	85.5722	76	84.7053
2	85.3195	27	87.5219	52	83.7935	77	83.8168
3	84.7355	28	86.4992	53	84.3706	78	82.4171
4	85.1113	29	85.6050	54	83.3762	79	83.0420
5	85.1487	30	86.8293	55	84.9975	80	83.6993
6	84.4775	31	84.5004	56	84.3495	81	80.2033
7	84.6827	32	84.1844	57	85.3395	82	82.1413
8	84.6757	33	85.4563	58	86.0503	83	80.7961
9	86.3169	34	86.1511	59	84.8839	84	82.3241
10	88.0006	35	86.4142	60	85.4176	85	81.5316
11	86.2597	36	86.0468	61	84.2309	86	81.7280
12	85.8286	37	86.6642	62	83.5761	87	82.5375
13	35.7500	38	84.7289	63	84.1343	88	82.3877
14	84.4628	39	85.9523	64	82.6974	89	82.4159
15	84.6476	40	86.8473	65	83.5454	90	82.2102
16	84.5751	41	88.4250	66	86.4714	91	82.7673
17	82.2473	42	89.6481	67	86.2143	92	83.1234
18	83.3774	43	87.8566	68	87.0215	93	83.2203
19	83.5385	44	88.4997	69	86.6504	94	84.4510
20	85.1620	45	87.0622	70	85.7082	95	84.9145
21	83.7881	46	85.1973	71	86.1504	96	85.7609
22	84.0421	47	85.0767	72	85.8032	97	85.2302
23	84.1023	48	84.4362	73	85.6197	98	86.7312
24	84.8495	49	84.2112	74	84.2339	99	87.0048
25	87.6416	50	85.9952	75	83.5737	100	85.0572

Fonte: [Montgomery, Jennings e Kulahci \(2008\)](#).

Para verificar se a série da Tabela 1 é estacionária, é necessário a formulação a Função de Autocorrelação, da coleta dos coeficientes de correlação ( $\rho_k$ ), segundo a equação (1), para  $k = 1, 2, 3, \dots$

$$\rho_k = \frac{E[(y_t - \mu)(y_{t+k} - \mu)]}{\sqrt{E[(y_t - \mu)^2]E[(y_{t+k} - \mu)^2]}} = \frac{Cov(y_t, y_{t+k})}{Var(y_t)} = \frac{\gamma_k}{\gamma_0} \quad (1)$$

Contudo, é necessária a estimativa das função de covariância, dada pela equação (2) :

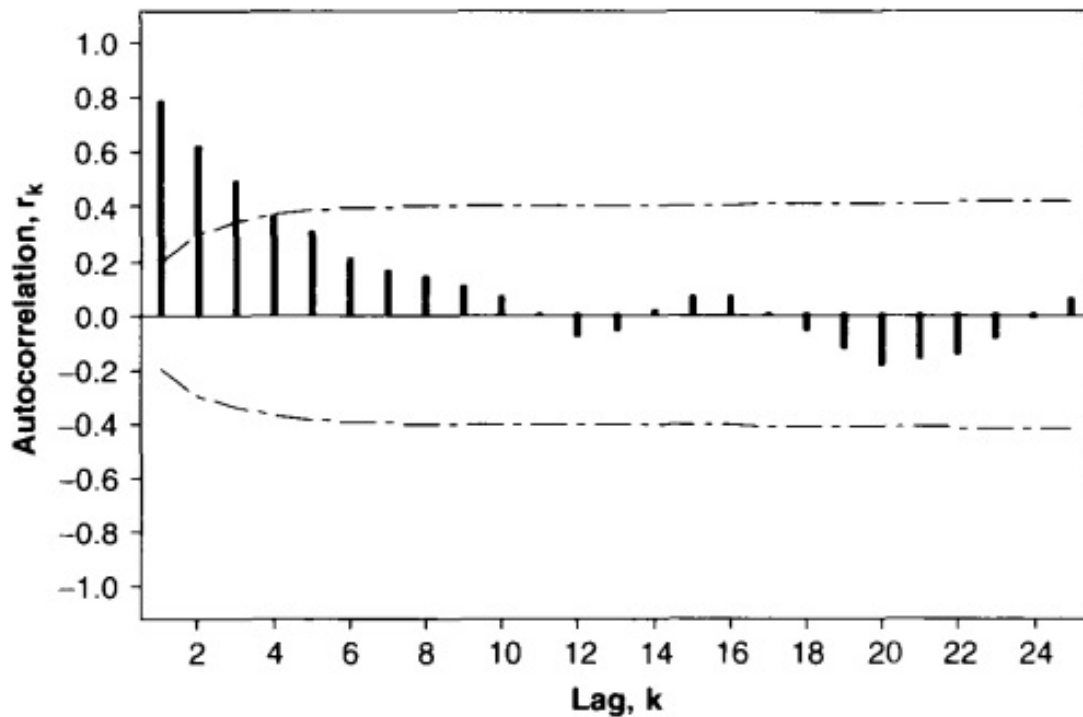
$$c_k = \hat{\gamma}_k = \frac{1}{T} \sum_{t=1}^{T-k} (y_t - \bar{y})(y_{t+k} - \bar{y}); \quad k = 0, 1, 2, \dots, K \quad (2)$$

E da função de autocorrelação, apresentada pela equação (3):

$$r_k = \hat{\rho}_k = \frac{c_k}{c_0}; \quad k = 0, 1, \dots, K \quad (3)$$

Aplicando esta metodologia, apresentada pelas equações (1), (2) e (3), para analisar os dados da Tabela da 1 obtém-se o seguinte gráfico, ilustrado pela Figura 7:

Figura 7 – Gráfico da função de correlação, com uma confiança de 5%, para viscosidade obtida da Tabela I

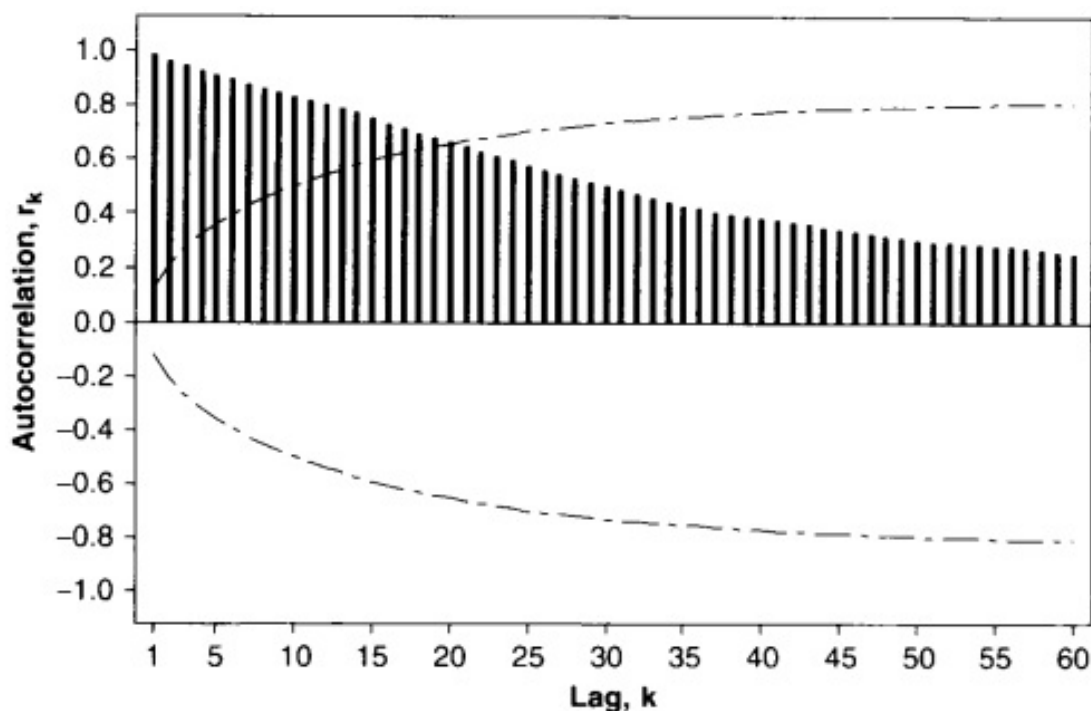


Fonte: [Montgomery, Jennings e Kulahci \(2008\)](#).

Cada intervalo de tempo,  $k$ , é chamado de *lag*. Pode-se observar que até o *lag* de número 11, as autocorrelações são positivas, mas a partir do *lag* de número 12, os coeficiente de autocorrelações cruzam o eixo zero apresentando um comportamento oscilatório, sobre o eixo zero, indicando estacionalidade ([MONTGOMERY; JENNINGS; KULAHCI, 2008](#)).

Para comparação, a [Figura 8](#) mostra um gráfico de uma função de autocorrelação de uma série não estacionária:

Figura 8 – Função de correlação para preço do estoque de mercado de alimentos



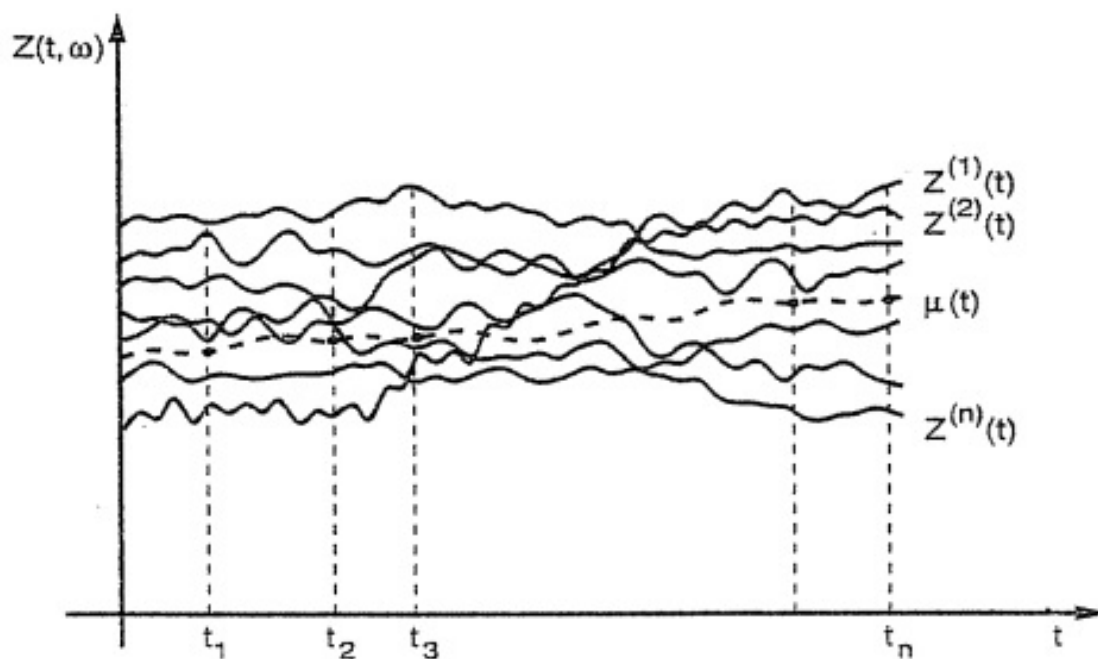
Fonte: [Montgomery, Jennings e Kulahci \(2008\)](#).

Todos os coeficientes de autocorrelação ilustrados pela [Figura 8](#) são positivos, decaindo aos poucos e tendendo a zero. Fenômeno muito característico de séries não estacionárias.

### 2.2.1 Processos Estocásticos

Processos Estocásticos são um conjunto de séries temporais possíveis dentro de uma variável aleatória ([BUENO, 2012](#)). O autor ainda afirma que uma variável aleatória pode ter diferentes comportamentos, ao longo do tempo, quando submetido em diferentes condições, podendo ter uma infinidade de combinações possíveis para um mesmo fenômeno observado. Segundo [Morettin e Toloí \(2006\)](#) processos estocásticos são processos controlados por leis probabilísticas, e se, com instrumentos apropriados, softwares adequados, é possível simular várias possibilidades de construções de séries temporais para o mesmo fenômeno, variável aleatória. A [Figura 9](#) ilustra um processo estocástico.

Figura 9 – Representação de um processo estocástico com um conjunto de séries possíveis.



Fonte: Morettin e Toloi (2006).

Cada  $Z^{(j)}(t)$ , sendo  $j = 1, 2, \dots, n$ , é uma representação de uma série em função de  $t$ , um avariável não aleatória, e se  $t$  for fixo,  $Z^{(j)}(t)$  é um número inteiro real. Todas  $Z^{(j)}(t)$  são representações de possíveis séries temporais sob diferentes condições e influências externas.

Para fazer uma estimativa de uma série temporal, supondo que haja estacionaridade, é necessário verificar as médias da série em cada instante de tempo. Se existe estacionaridade as médias deverão ser muito próximas e se esta condição for verificada, a série apresenta ergodicidade (BUENO, 2012). Isso implica que em qualquer intervalo de tempo, todos os valores tenderão a apresentar uma média constante sendo necessário um número muito baixo de análises para fazer estimativas do modelo. Morettin e Toloi (2006) apresentam três classificações para estas séries: Processos estocásticos podem ser:

- Estritamente estacionário: média ( $\mu$ ) e variância ( $\sigma^2$ ) são constantes;
- Fracamente estacionária: O modelo se desenvolve sobre uma média, mas há diferenças de níveis em certos intervalos de tempo, como visto na Figura 6.
- Gaussiano: determinada pelas médias e covariâncias.

Mas segundo Bueno (2012), a ergodicidade está restrita para séries fracamente estacionárias.

### 2.3 RUÍDO BRANCO

O chamado Ruído Branco é um fenômeno fundamental para a compreensão de eventos estocásticos (BUENO, 2012). Basicamente é uma sequência de dados ( $a_t$ ) onde a média gira em torno de zero, com

variância constante e as variáveis ( $a_t$ ) não sejam correlacionadas, para qualquer instante de tempo  $t$ , representados pelas equações (4), (5) e (6), respectivamente.

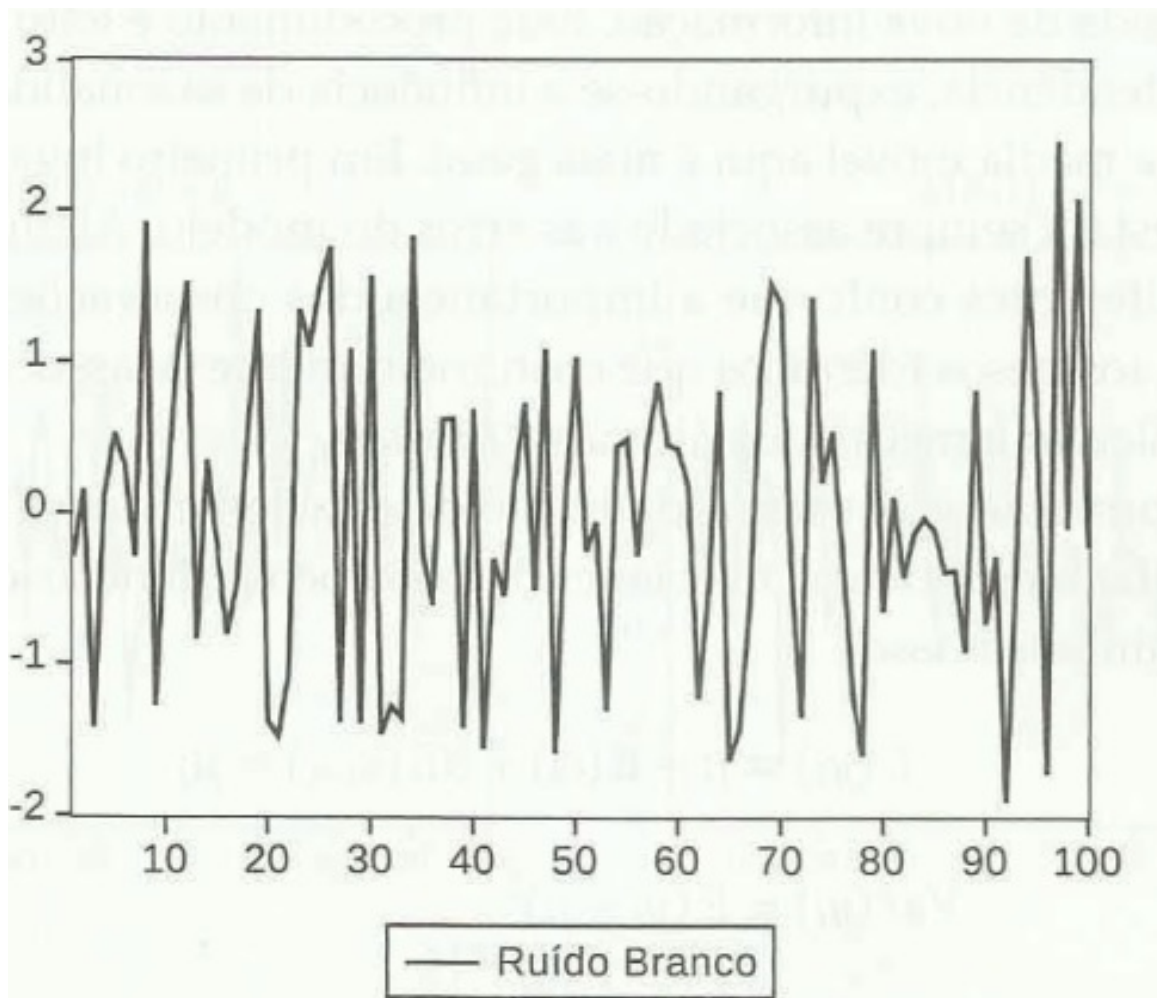
$$E(a_t) = 0, \forall t \quad (4)$$

$$E(a_t^2) = a_a^2, \forall t \quad (5)$$

$$E(a_t a_s) = 0, s \neq t \quad (6)$$

Este tipo de conceito foi primeiramente aplicado na área da Astronomia e Física, para determinar o posicionamento de corpos celestes em um determinado instante de tempo (MORETTIN; TOLOI, 2006).

Figura 10 – Exemplo de um ruído branco.



Fonte: Bueno (2012).

### 2.3.1 CÍRCULO UNITÁRIO

Imagine um círculo de raiz unitária sobre um plano cartesiano definido por um eixo vertical, eixo  $y$ , e um eixo horizontal, eixo  $x$ . Sendo o centro da circunferência estando sobre o ponto  $(0, 0)$  do plano.

Se definido um ponto genérico  $(x_0, y_0)$  dentro da circunferência unitária e for traçada uma reta da origem, ponto  $(0, 0)$  até o ponto genérico, a reta terá um ângulo  $\theta$  com o eixo  $x$ , então pode-se afirmar as condições das equações (7) e (8):

$$\sin(\theta) = y_0 \quad (7)$$

$$\cos(\theta) = x_0 \quad (8)$$

Se o comprimento da reta traçada for definido como  $R$ , obtêm-se as seguintes condições das equações (9) e (10):

$$\frac{y}{R} = \frac{y_0}{1} \Rightarrow y = R \sin \theta \quad (9)$$

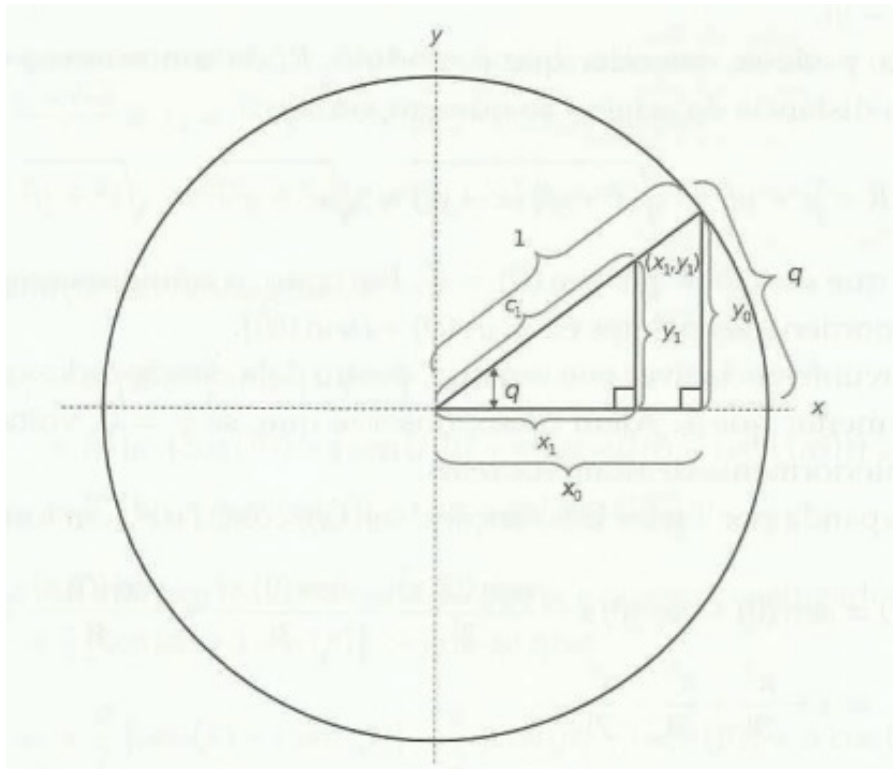
$$\frac{x}{R} = \frac{x_0}{1} \Rightarrow x = R \cos \theta \quad (10)$$

e por Pitágoras, representado pela equação (11).

$$R = \sqrt{x^2 + y^2} \quad (11)$$

A Figura 11 mostra uma representação gráfica do círculo unitário:

Figura 11 – Representação de um círculo unitário.



Fonte: Bueno (2012).

As raízes de um polinômio de modelos ARIMA que estiverem no interior do círculo unitário são denominadas de “raízes unitárias”, e se um polinômio tiver todas as suas raízes unitárias, é dito que o

sistema é estável (BUENO, 2012). Este tipo de conceito ajuda a visualizar se uma série é estacionário ou não, e é válido para sistemas de ordem dois ou superiores.

## 2.4 MÉDIAS MÓVEIS

### 2.4.1 MÉDIAS MÓVEIS DE ORDEM $q$ – MA( $q$ )

Para o início, a equação (12) representando um processo estocástico.

$$y_t = \mu + \varepsilon_t + \theta_{\varepsilon_{t-1}} \quad (12)$$

sendo  $\varepsilon_t$ , o ruído branco,  $\mu$  é a média da série e  $\theta_{\varepsilon_{t-1}}$ , um erro ocorrido em um período imediatamente anterior.

Observando a equação (12) verifica-se que  $y_t$  depende de  $\varepsilon_t$ , o ruído branco, e de um erro de um instante de tempo imediatamente anterior, por isso o nome de médias móveis de ordem 1. A função  $y_t$  é uma função estacionária, independentemente dos valores de seus pesos.

Agora é apresentado um caso mais geral para médias móveis, MA( $q$ ), representada pela equação (13).

$$y_t = \mu + \sum_{j=0}^q \theta_j \varepsilon_{t-j}, \quad \theta_0 = 1 \quad (13)$$

A equação (13) respeita a condição de estacionariedade e as correlações cruzadas são inexistentes (BUENO, 2012). Fazendo a análise da função, é possível notar que a auto correlação é existente para defasagens ( $j$ ) menores que  $q$  ( $j < q$ ). Quando  $j > q$ , a auto correlação desaparece e define-se que a função é uma MA ( $q$ ).

A equação (13) respeita a condição de estacionariedade e as correlações cruzadas são inexistentes. Fazendo-se a análise da função, é possível notar que a auto correlação é existente para defasagens menores que  $q$  ( $j < q$ ). Quando  $j > q$ , a auto correlação desaparece e define-se que a função é uma MA ( $q$ ).

### 2.4.2 Processo autor regressivo de ordem $p$ – AR ( $p$ )

Um processo autor regressivo de ordem  $q$  é representado pela seguinte equação (14).

$$y_t = c + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \dots + \phi_p y_{t-p} + \varepsilon_t \quad (14)$$

sendo  $\varepsilon_t$  é o ruído branco,  $c$ , uma constante, e  $\phi$ , coeficientes da equação que são identificados após a formulação do polinômio da série temporal.

No modelo representado pela equação (14), um número muito pequeno de eventos que ocorreram próximo do presente momento que terão influência no momento da observação, eventos muito antigos serão desconsiderados (MONTGOMERY; JENNINGS; KULAHCI, 2008).

Se as raízes do polinômio estiverem dentro do círculo unitário, o processo representado pela equação (14) é fracamente estacionária e podendo ser representada pela equação (15):

$$y_t = \mu + \psi(L)\varepsilon_t \quad (15)$$

sendo  $\mu$  representado pela equação (16) e  $\psi(L)$ , pela equação (17)

$$\mu = \frac{c}{1 - (\phi_1 + \phi_2 + \dots + \phi_p)} \quad (16)$$

$$\psi(L) = (1 - (\phi_1 L + \phi_2 L^2 + \dots + \phi_p L^p))^{-1} \quad (17)$$

Para cada instante de tempo serão obtidas polinômios com diferentes coeficientes, que podem ser determinados pelas chamadas Equações de Yule – Walker com o uso do algoritmo de Durbin-Levinson (MORETTIN; TOLOI, 2006).

### 2.4.3 Modelos auto regressivos de média móveis – ARMA (p,q)

O modelo auto regressivo de Médias Móveis (ARMA) é, resumidamente, a união dos dois modelos vistos anteriormente. Com a combinação dos dois modelos vistos, obtendo-se o seguinte modelo, ARMA (p,q), representado pela equação (18).

$$y_t = c + \sum_{i=1}^p \phi_i y_{t-i} + \sum_{j=0}^q \theta_j \varepsilon_{t-j} \quad (18)$$

Para a que o modelo ARMA seja válido, é necessário que a série analisada seja estável, ou seja, as raízes das equações dos modelos AR e MA esteja no interior do círculo unitário, e esta condição também é válida para verificação da invertibilidade (MONTGOMERY; JENNINGS; KULAHCI, 2008). Se as raízes dos modelos AR e MA estiverem no interior do círculo unitário, o modelo ARMA é invertível (BUENO, 2012). Isto quer dizer que, se o modelo é invertível, podendo escrever um modelo MA (q) como uma AR ( $\infty$ ), se as condições, ditas anteriormente, forem respeitadas.

Este tipo de modelagem, unindo os modelos auto regressivos e de médias móveis, é um modo mais adequados para o estudo de eventos que acontecem na vida real, pois consegue englobar eventos externos que podem influenciar na série temporal, o que não ocorre com aplicação de modelos puramente MA ou AR.

Os modelos de séries temporais podem ser separados em dois grupos: Modelos Paramétricos e Modelos Não – Paramétricos (MORETTIN; TOLOI, 2006).

Em Modelos paramétricos, a quantidade de variáveis envolvidas é finita e a análise é feita em função do tempo. Modelos mais comuns são: Modelos auto regressivos, ARMA, ARIMA modelos estruturais e modelos não – lineares. Os modelos não-paramétricos são analisados em função de suas autocovariâncias e transformadas de Fourier.

## 2.5 PROCESSOS NÃO ESTACIONÁRIOS

Série não estacionárias, diferentemente de séries estacionárias, não apresentam um comportamento que gira em torno de uma média. Tal comportamento é difícil de estimar em todas as transições da série. Mas existe um método para transformar uma série não estacionária em estacionária que seria fazer diferenciações até que a série mude de não estacionária para estacionária (BUENO, 2012), o correndo a estabilização da mesma.

Tais diferenciações ( $w_t$ ) são feitas seguindo a equação (19).

$$w_t = y_t - y_{t-1} \quad (19)$$

Em muitos casos, apenas uma diferenciação é necessária para a estabilização da série não estacionária, mas existem muitos casos onde são necessárias mais que uma diferenciação para encontrar alguma estacionariedade.

Geralmente séries não estacionárias apresentam os seguintes componentes:

$$Y_t = \text{tendência} + \text{componente estacionário} + \text{ruído}$$

A tendência pode ser de dois tipos: Tendência estacionária, onde a série desenvolve-se em torno de uma média, não apresentando valores muito distantes da média; Tendência estocástica, é dada pela diferenciação de séries para torná-las estacionárias, onde causa alterações aleatórias na série.

Cada diferenciação é denotada pela ordem “ $d$ ”, e as séries diferenciadas são chamadas de Séries Integradas (BUENO, 2012). Toda diferenciação acaba inserindo um “erro” dentro da série temporal, nisso, modelos ARMA que precisam ser diferenciadas para se tornarem estacionárias são chamadas de séries ARIMA (p,d,q).

### 2.5.1 Construção do modelo ARIMA (p,d,q)

Para a formulação de modelos ARIMA, é necessário seguir os seguintes passos para se ter um modelo ARIMA correto: 1 – verificar se a série histórica se encaixa no modelo ARIMA; 2 – identificar os parâmetros  $p,d,q$ ; 3 – verificar se o modelo é confiável e se é necessário aperfeiçoamentos (MONTGOMERY; JENNINGS; KULAHCI, 2008)). Os passos são melhores descritos.

**Identificação do modelo:** Antes do início da construção do modelo temporal, deve-se ter o entendimento a procedência dos dados da série temporal, para ter um breve conhecimento do comportamento da série. Com o conhecimento da origem dos dados, podemos direcionar a análise e validar o modelo já pré-determinado. Com o conhecimento da procedência dos dados, é necessário a plotagem dos dados em um gráfico cartesiano para obter uma visualização da série temporal para detectar estacionariedade ou se é necessário aplicar diferenciações para tornar a série estacionária, caso haja não – estacionariedade. Para uma série ser válida, deve conter pelo menos 60 observações. Em seguida, deve ser analisada as funções de autocorrelação.

**Estimativa de parâmetros:** Existem diversas metodologias para a determinação de parâmetros, mas como grande parte das séries em modelos ARIMA são não estacionárias, há uma maior complexidade

para a determinação dos mesmos, necessitando de uso de softwares especializados para aplicação de análises não lineares.

Verificação do Modelo: O último passo a ser seguido é a validação do modelo resultante dos dois primeiros passos para verificar se deve ser feito alguns ajustes para melhorar o modelo. Esta análise pode ser feita por resíduos do modelo ARMA pela equação (20).

$$\hat{\varepsilon}_t = y_t - \left( \sum_{i=1}^p \hat{\phi}_i y_{t-i} - \sum_{i=1}^q \hat{\theta}_i \hat{\varepsilon}_{t-i} \right) \quad (20)$$

Se o modelo estiver certo, sendo os índices  $d$  e  $p$  ajustados adequadamente, a equação (20) terá um comportamento de um ruído branco.

Fazendo uma verificação com a função de autocorrelação dos resíduos feita pela equação (20), o modelo será adequado se as autocorrelações do resíduo tiverem nenhuma estrutura identificável, ou seja, as autocorrelações para todos os *lags* são muito próximas de zero.

## 2.6 PROCESSOS SAZONAIS

Existe alguns casos onde séries temporais apresentam um comportamento periódico ao longo do tempo. A série exibe um desenvolvimento semelhante em intervalos de tempos iguais, podendo ser de caráter anual, mensal e semanal. Tal comportamento é chamado de sazonalidade, e o modelo de séries temporais com comportamento sazonal apresentam duas componentes, uma componente determinística e uma componente estocástica, que vem do modelo ARMA (MONTGOMERY; JENNINGS; KULAHCI, 2008). De uma forma geral, a função do modelo sazonal pode ser escrita de acordo com a equação (21).

$$y_t = S_t + N_t \quad (21)$$

Onde:  $S_t$  é a componente determinística com periodicidade  $s$  e  $N_t$  a componente estocástica.

Como no final de cada período  $s$  os valores de  $y_t$  são iguais, podemos escrever o seguinte:  $S_t = S_{t+s}$ . Com isso obtêm-se a equação (22).

$$S_t - S_{t-s} = (1 - B^s)S_t = 0 \quad (22)$$

Aplicando  $(1 - B^s)$  na equação (21) obtém-se a equação (23):

$$(1 - B^s)y_t = (1 - B^s)S_t + (1 - B^s)N_t \quad (23)$$

Supõe-se que  $(1 - B^s)y_t = w_t$ , uma notação para deixar a equação mais simples, e  $(1 - B^s)S_t = 0$ , com isso obtém-se a equação (24).

$$w_t = (1 - B^s)N_t \quad (24)$$

Como os modelos ARMA podem ser usados no modelo  $N_t$ , componente estocástica, pode-se afirmar seguinte condição da equação (25).

$$\phi(B)w_t = (1 - B^s)\Theta(B)\varepsilon_t \quad (25)$$

Onde  $\varepsilon_t$  é o ruído branco.

Considera-se  $S_t$  como um processo estocástico, mas a sazonalidade ainda influencia no processo  $S_t$ . Se diferenciações forem feitas, as autocorrelações serão muito altas. O modelo ARMA sazonal é representado pela equação (26):

$$(1 - \phi_1^*B^s - \phi_2^*B^{2s} - \dots - \phi_P^*B^{Ps})w_t = (1 - \theta_1^*B^s - \theta_2^*B^{2s} - \dots - \theta_Q^*B^{Qs})\varepsilon_t \quad (26)$$

O modelo apresentado pela equação (26) fica limitado para *lags* de  $s, 2s, \dots$ . Nisso obtêm-se um modelo mais geral sendo o modelo SARIMA  $(p, d, q) \times (P, D, Q)$  com periodicidade  $s$ , mostrado pela equação (27):

$$\Phi^*(B^s)\Phi(B)(1 - B)^d(1 - B^s)^D y_t = \delta + \Theta^*(B^s)\Theta(B)\varepsilon_t \quad (27)$$

Os modelos SARIMA apresentam duas componentes, uma não sazonal, representadas pelos coeficientes  $p, d$  e  $q$ , e uma sazonal, representadas pelos coeficientes  $P, D$  e  $Q$

### 3 DESCRIÇÃO DO PROBLEMA E MODELAGEM

Para este trabalho coletou-se um histórico de faturamento mensal, entre o período de Julho de 2013 a Agosto de 2019, pertencentes à montadora de veículos automotores Nissan do Brasil. Os dados de faturamento mensal estão na Tabela 12, presente no Anexo A, e foram analisados com o software GRETL, proporcionando uma melhor abordagem quantitativa para obtenção de um modelo de previsão mais adequado para a série temporal. Com os dados em mãos, possibilitou-se uma melhor visualização do histórico de faturamento da montadora, tendo uma ideia do que o modelo retornaria como previsão e contribuir com os estudos atuais sobre este assunto.

#### 3.1 MODELAMENTO AUTO REGRESSIVO

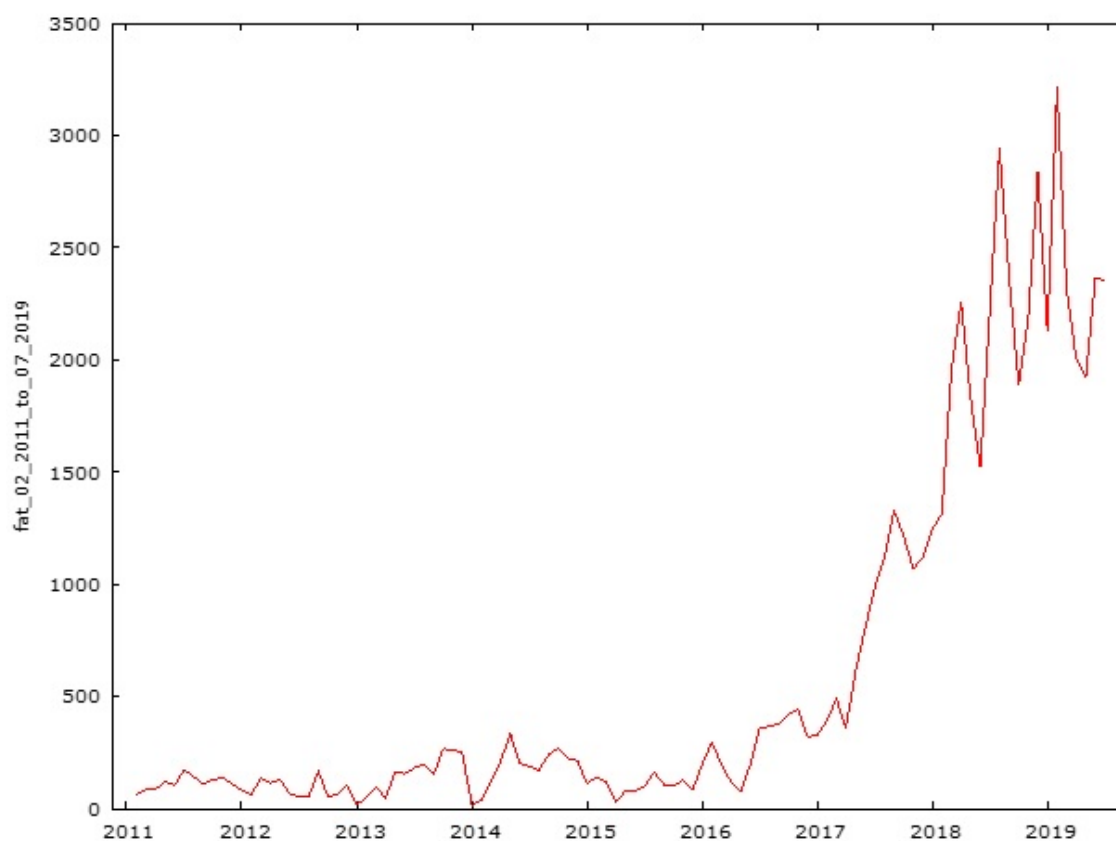
Como dito a seção 1.4, o modelamento auto regressivo de uma série temporal deve seguir os seguintes passos: identificação, estimação e verificação do modelo. Parte da identificação do modelo é influenciada com base das autocorrelações e autocorrelações parciais (MORETTIN; TOLOI, 2006). O cálculo dos fatores de autocorrelação já foram demonstradas na seção 2.2, e com este método de determinação dos fatores de autocorrelação (FAC) é possível determinar o modelo SARIMA (p, d, q) x (P, D, Q).

Para a identificação dos coeficientes  $p$ ,  $d$  e  $q$ , é preciso verificar se é necessário o uso de diferenciações para a estabilização da série, como explicado na seção 2.5. O número “ $d$ ” de diferenciações necessárias para a estabilização da série será o próprio coeficiente “ $d$ ” do modelo ARIMA. O número de diferenciações é verificado com os coeficientes de autocorrelação (FAC), quando estão convergindo para o valor zero, o Teste de Dickey – Fuller, ou ADF (*augmented Dickey–Fuller test*), é aplicado para a verificação de raízes unitárias do polinômio do modelo auto regressivo. Com as diferenciações necessárias aplicadas, a série pode ser trabalhada como modelo ARMA(p, q) para a determinação do restante dos coeficientes, que são estimadas com os coeficientes de auto correlação e autocorrelação parcial.

#### 3.2 ESTUDO E ANÁLISE DA SÉRIE

Seguindo a metodologia de análise proposta por Box e Jenkins (LIU, 1992), a primeira etapa será a análise da série temporal para a observar o desenvolvimento da série com o decorrer do tempo. Com isso, foi feito o gráfico da série temporal do faturamento de veículos PCD e Taxi, do período entre Fevereiro de 2011 a Julho de 2019, ilustrado pela Figura 12, por meio dos dados da Tabela 12, no Anexo A.

Figura 12 – Histórico de faturamento de veículos para pessoas com deficiência (PCD) e clientes Taxistas.



Fonte: Elaborado pelo programa Gretl

Observa-se que entre os anos de 2011 e 2016, a série temporal apresenta oscilações ao longo do período e, após o ano de 2016, apresentando tendência de crescimento, com certos decaimentos acompanhados por um crescimento no número de faturamento de veículos . Como observado a série temporal não é estacionária, e para trabalhar com previsão de séries deste tipo, é necessário a estabilização da série, ou seja, torna-la estacionária. Para verificar a estacionaridade da série temporal, foi feito o teste da Raiz unitária ou Teste de Dickey – Fuller aumentado para p-valores de 5%, pois apenas analisando visualmente o gráfico, não é possível obter conclusões sólidas a respeito da série temporal. Os resultados do primeiro teste ADF, para a série não diferenciada, de caráter demonstrativo, estão ilustrados pela Tabela 2:

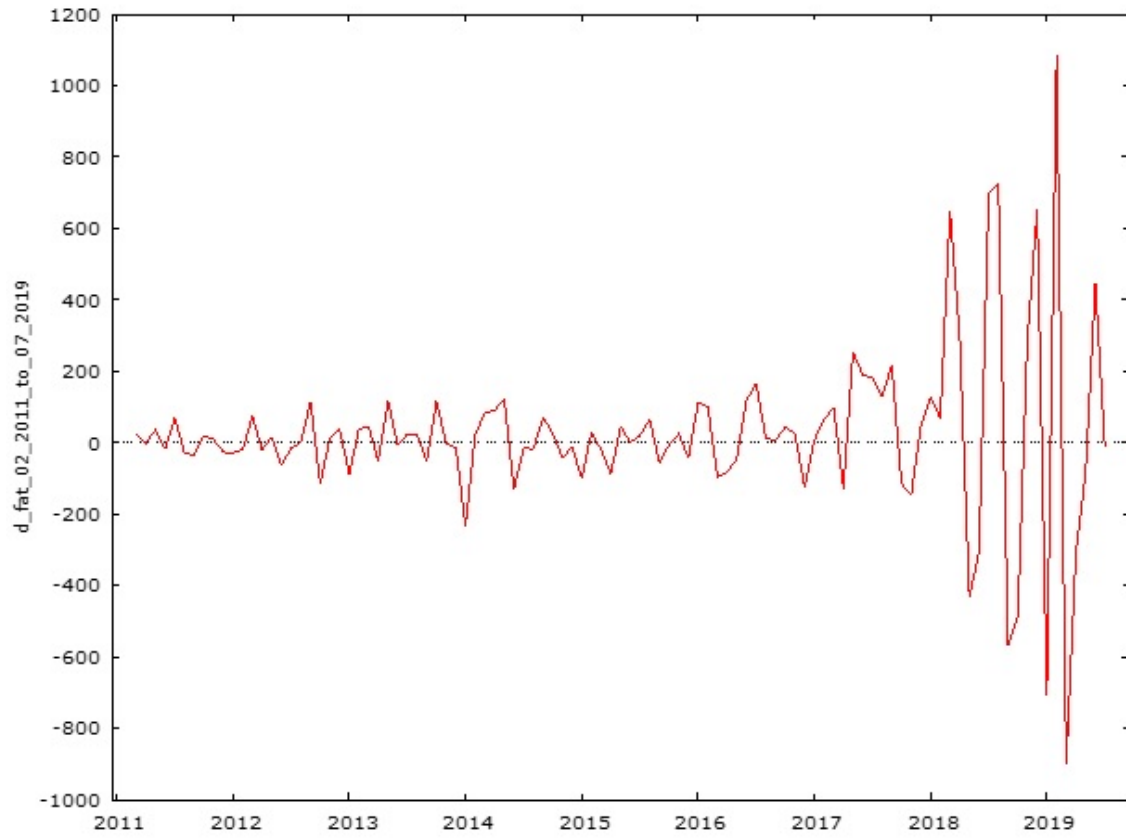
Tabela 2 – Teste Dickey-fuller para a série temporal não diferenciada

<p>Teste Aumentado de Dickey-Fuller para fat_02_2011_to_07_2019  testar para baixo a partir de 12 defasagens, critério AIC  tamanho da amostra: 90  hipótese nula de raiz unitária: <math>a = 1</math></p> <p>teste com constante  incluindo 11 defasagens de <math>(1-L)fat\_02\_2011\_to\_07\_2019</math>  modelo: <math>(1-L)y = b_0 + (a-1)*y(-1) + \dots + e</math>  valor estimado de <math>(a - 1)</math>: -0,182989  estatística de teste: <math>\tau\_c(1) = -2,88664</math>  p-valor assintótico <b>0,0469</b>  coeficiente de 1ª ordem para e: -0,004  diferenças defasadas: <math>F(11, 77) = 5,025 [0,0000]</math></p> <p>com constante e tendência  incluindo 11 defasagens de <math>(1-L)fat\_02\_2011\_to\_07\_2019</math>  modelo: <math>(1-L)y = b_0 + b_1*t + (a-1)*y(-1) + \dots + e</math>  valor estimado de <math>(a - 1)</math>: -0,228244  estatística de teste: <math>\tau\_ct(1) = -3,27722</math>  p-valor assintótico <b>0,07001</b>  coeficiente de 1ª ordem para e: -0,004  diferenças defasadas: <math>F(11, 76) = 4,477 [0,0000]</math></p> <p>Fonte: Elaborado pelo programa Gretl</p>
--

Com o primeiro teste de Dikey – Fuller aumentado, para a série não diferenciada, p-valore com constante de 0.0469 e p-valor com constante e tendência de 0.07001, destacados em negrito, mas visualmente a série apresenta uma tendência de crescimento e para criar um modelo de previsão é necessário a estabilização da série.

Para estabilização a série temporal e eliminação a tendência, a série deve ser diferenciada e o teste de Dikey – Fuller aumentado reaplicado, com o intuito de verificar a estacionaridade e tendência da mesma. A [Figura 13](#) ilustra a série diferenciada, e podemos observar a série oscila sobre o valor zero, dando indícios de estabilização, mas para validar a estacionaridade, deve-se aplicar novamente o teste de Dickey – Fuller aumentado.

Figura 13 – Série temporal após a primeira diferença



Fonte: Elaborado pelo programa Gretl

Os resultados do teste Dickey – Fuller aumentado para a série diferenciada estão ilustrados na Tabela 3:

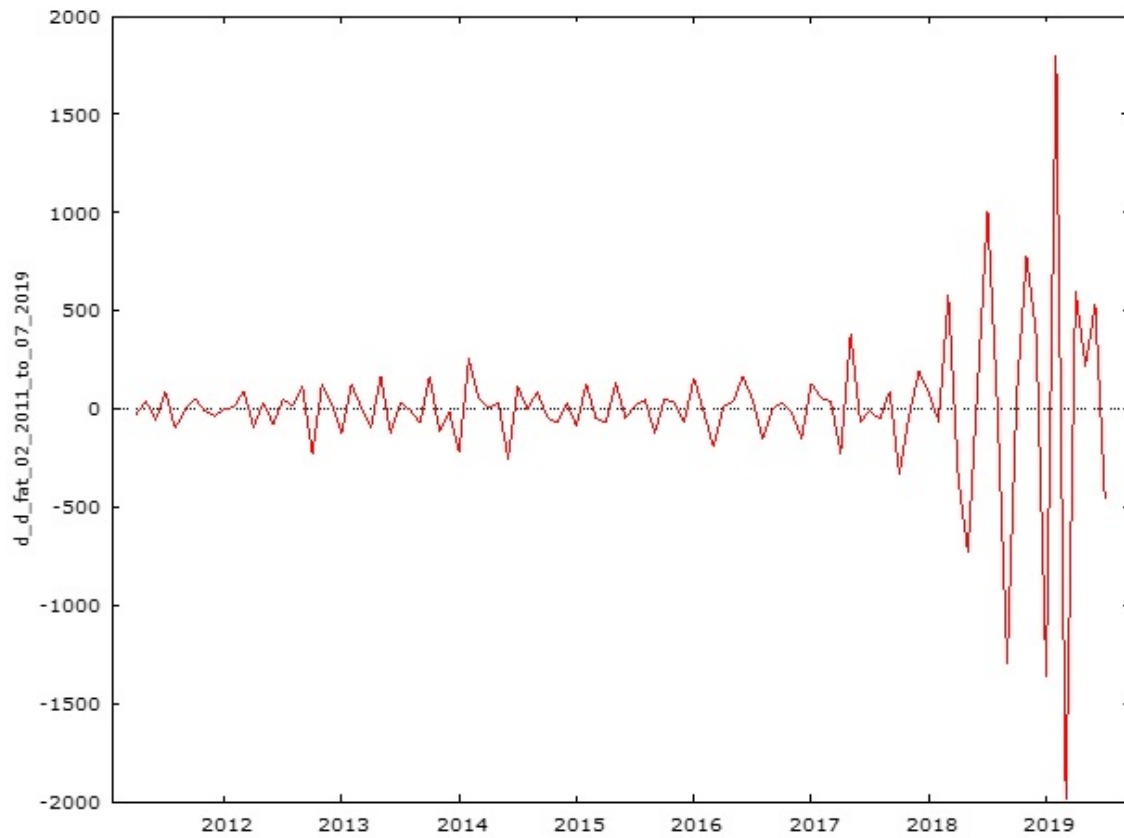
Tabela 3 – Teste Dickey-fuller para a série temporal diferenciada

<p>Teste Aumentado de Dickey-Fuller para <math>d\_fat\_02\_2011\_to\_07\_2019</math>  testar para baixo a partir de 12 defasagens, critério AIC  tamanho da amostra: 94  hipótese nula de raiz unitária: <math>a = 1</math>  teste com constante  incluindo 6 defasagens de <math>(1-L)d\_fat\_02\_2011\_to\_07\_2019</math>  modelo: <math>(1-L)y = b_0 + (a-1)*y(-1) + \dots + e</math>  valor estimado de <math>(a - 1)</math>: -1,19708  estatística de teste: <math>\tau_c(1) = -2,32736</math>  p-valor assintótico <b>0,1633</b>  coeficiente de 1ª ordem para e: 0,015  diferenças defasadas: <math>F(6, 86) = 5,549 [0,0001]</math></p> <p>com constante e tendência  incluindo 6 defasagens de <math>(1-L)d\_fat\_02\_2011\_to\_07\_2019</math>  modelo: <math>(1-L)y = b_0 + b_1*t + (a-1)*y(-1) + \dots + e</math>  valor estimado de <math>(a - 1)</math>: -1,53742  estatística de teste: <math>\tau_{ct}(1) = -2,34139</math>  p-valor assintótico <b>0,4108</b>  coeficiente de 1ª ordem para e: 0,016  diferenças defasadas: <math>F(6, 85) = 5,402 [0,0001]</math></p>
--

Fonte: Elaborado pelo Autor

Obtive-se p-valor de 0.1633 com constante e p-valor de 0.4108, destacados em negrito, os p-valores tiveram os seus valores incrementados ficando acima do valor de 5%. Com a primeira diferença não se obtive evidência de estacionaridade e a série continuou apresentando tendência, nisso, foi preciso uma segunda diferença da série, que consiste em aplicar mais uma diferença sobre a serie já diferenciada. Por meio da [Figura 14](#) podemos observar a nova série, com duas diferenciações, apresentando o mesmo comportamento da série com apenas uma diferenciação, com oscilação em torno do valor zero.

Figura 14 – Série temporal após a segunda diferença



Fonte: Elaborado pelo programa Gretl

Mesmo com a série duas vezes diferenciada, é necessário a validação pelo Teste Dickey – Fuller aumentado. A Tabela 4 apresenta os resultados retornados pelo teste.

Tabela 4 – Teste Dickey-fuller para a série temporal diferenciada pela segunda vez

<p>Teste Aumentado de Dickey-Fuller para <math>d\_d\_fat\_02\_2011\_to\_07\_2019</math>  testar para baixo a partir de 12 defasagens, critério AIC  tamanho da amostra: 90  hipótese nula de raiz unitária: <math>a = 1</math></p> <p>teste com constante  incluindo 9 defasagens de <math>(1-L)d\_d\_fat\_02\_2011\_to\_07\_2019</math>  modelo: <math>(1-L)y = b_0 + (a-1)*y(-1) + \dots + e</math>  valor estimado de <math>(a - 1)</math>: -13,1526  estatística de teste: <math>\tau_c(1) = -5,72369</math>  p-valor assintótico <b>5,377e-007</b>  coeficiente de 1ª ordem para e: 0,023  diferenças defasadas: <math>F(9, 79) = 15,655 [0,0000]</math></p> <p>com constante e tendência  incluindo 9 defasagens de <math>(1-L)d\_d\_fat\_02\_2011\_to\_07\_2019</math>  modelo: <math>(1-L)y = b_0 + b_1*t + (a-1)*y(-1) + \dots + e</math>  valor estimado de <math>(a - 1)</math>: -13,3589  estatística de teste: <math>\tau_{ct}(1) = -5,79659</math>  p-valor assintótico <b>3,551e-006</b>  coeficiente de 1ª ordem para e: 0,023  diferenças defasadas: <math>F(9, 78) = 15,800 [0,0000]</math></p>
---

Fonte: Elaborado pelo Autor

Obtive-se p-valor de  $5,377 \times 10^{-7}$  com constante e p-valor de  $3,551 \times 10^{-6}$  com constante e tendência. Ambos os p-valores estão abaixo de 5%, validando a estacionaridade da série e a eliminação da tendência. Como foi necessário duas diferenciações, para a estabilização da série, obtém-se o coeficiente  $d$  do nosso modelo ARIMA igual a 2, nisso  $d = 2$ . Com a série estabilizada, pode-se prosseguir para a identificação dos outros índices, ou filtros, do modelo ARIMA, com a análise da função de Autocorrelação (FAC) e da função de Autocorrelação Parcial (FACP) da série estabilizada.

### 3.3 – MODELO SARIMA (p, d, q) X (P, D, Q): AJUSTE E VALIDAÇÃO DO MODELO

Tomado como valor de referência  $c = \pm \frac{2}{\sqrt{n}}$ , sendo  $c$  o valor de referência e  $n$ , o número de observações. Para cada diferença executada, a amostra foi reduzida em uma unidade, como foram necessárias duas diferenciações, a amostra com 102 leituras fora reduzida para 100,  $n = 100$ , nisso, o valor de referência ficou em um valor de  $c = 0.2$ . O valor de referência será usado para definição dos filtros Auto Regressivo e de Médias Móveis,  $p$  e  $q$  respectivamente.

Para avaliação dos FAC e FACP, tomando como referência o valor de  $c$ , os coeficientes são analisados considerado os seus valores absolutos, ou seja, o módulo dos valores dos coeficientes. Para o filtro auto regressivo ( $p$ ) é definido pelo último  $lag$  com valor de FACP acima o valor de referência. O filtro de médias móveis ( $q$ ) é definido pelo último  $lag$  em que o valor de FAC está acima do valor de referência. Nisso, obtivemos os seguintes filtros:  $p = 6$  e  $q = 4$ , ambos destacados em negrito na

Tabela 5. Não são avaliados todos dos *lags* da série temporal, pois a partir dos *lags* selecionados para a definição dos filtros, os valores de FAC e FACP começam a tender a zero, não havendo a necessidade de ilustrar a função de autocorrelação para série toda.

Tabela 5 – Correlograma da série temporal diferenciada pela segunda vez

Função de autocorrelação para d_d_fat_02_2011_to_07_2019				
***, **, * indicam significância aos níveis de 1%, 5% e 10%				
usando erro padrão $1/T^{0,5}$				
Defas.	FAC	FACP	Estat.	Q [p-valor]
1	-0,5086***	-0,5086***	266,482	[0,000]
2	-0,0217	-0,3781***	266,971	[0,000]
3	-0,1248	-0,5230***	283,349	[0,000]
4	<b>0,2805***</b>	-0,2465**	366,937	[0,000]
5	-0,1975**	-0,3843***	408,810	[0,000]
6	0,0841	<b>-0,3890***</b>	416,477	[0,000]
7	0,0708	-0,1178	421,980	[0,000]
8	-0,0823	-0,0594	429,485	[0,000]
9	-0,1115	-0,1323	443,413	[0,000]
10	0,0909	-0,2291**	452,770	[0,000]

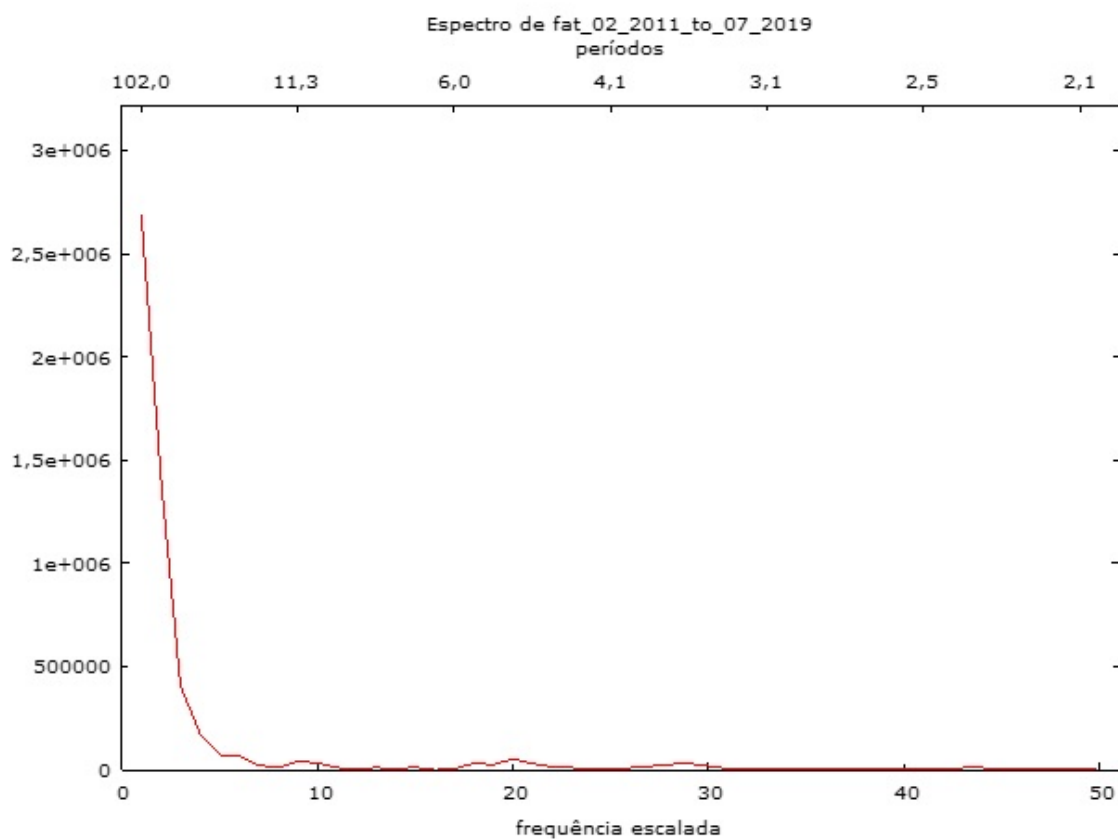
Fonte: Elaborado pelo Autor

Com os filtros definidos não sazonais, o modelo escolhido foi o SARIMA (6, 2, 4) x (P, D, Q), mas segundo a literatura, os filtros “*p*” e “*q*” nunca serão acima de 5, com isso, o define-se *q* como 5, consequentemente o modelo SARIMA (5, 2, 4) x (P, D, Q) como modelo estimado.

Com os coeficientes *p*, *d*, e *q* definidos, parte-se para a identificação da porção sazonal do modelo, identificados como *P*, *D* e *Q*. O processo de identificação dos filtros sazonais é muito semelhante com o processo de identificação dos filtros não sazonais.

Antes de fazer a análise dos FAC e FACP, é necessário observar o comportamento do periodograma da série temporal. A Figura 15 ilustra o periodograma da série temporal não diferenciada e é notável que o gráfico inicia com um valor muito acentuado que decai rapidamente, tendendo a zero, sendo necessário ajustar a série para que este comportamento seja alterado, com a aplicação de diferenciações sazonais para a “quebra” do pico ilustrado pela Figura 15.

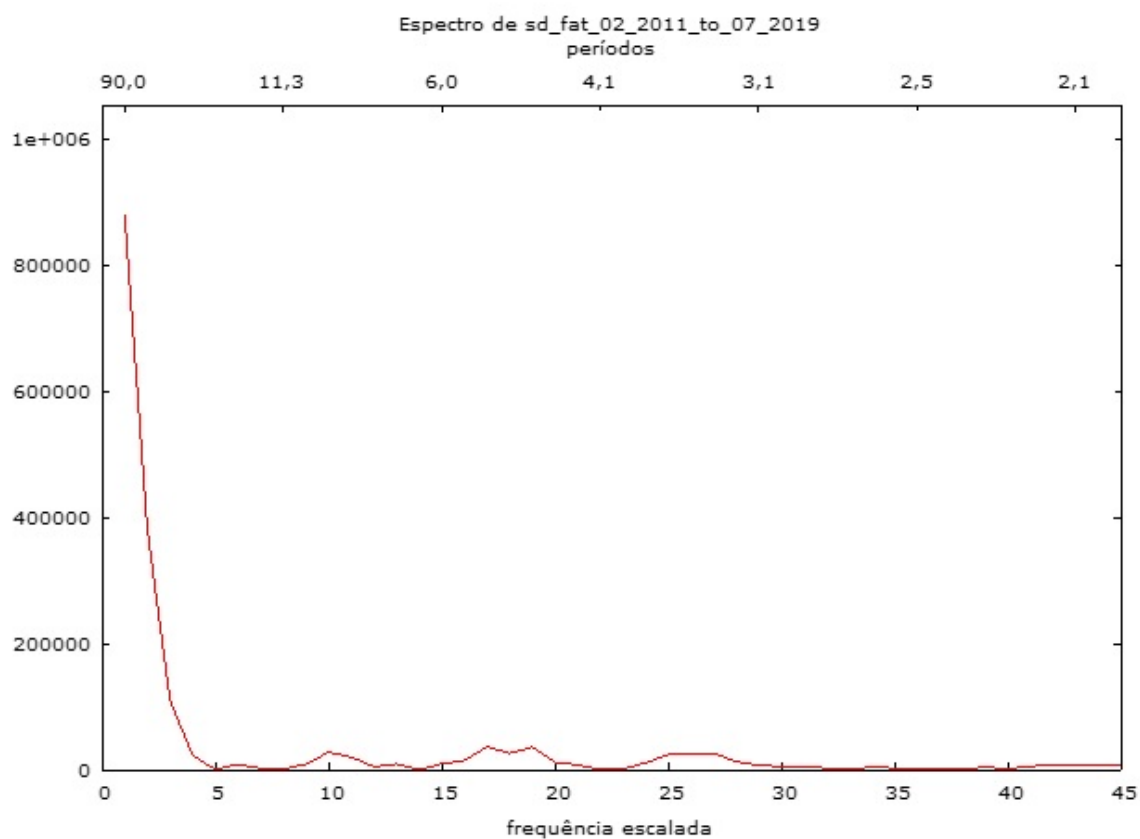
Figura 15 – Periodograma para a série original



Fonte: Elaborado pelo programa Gretl

A [Figura 16](#) ilustra o periodograma para série com apenas uma diferença sazonal, e ainda é possível observar o pico mencionado anteriormente, nisso, será necessário mais uma diferenciação sazonal para avaliação da série.

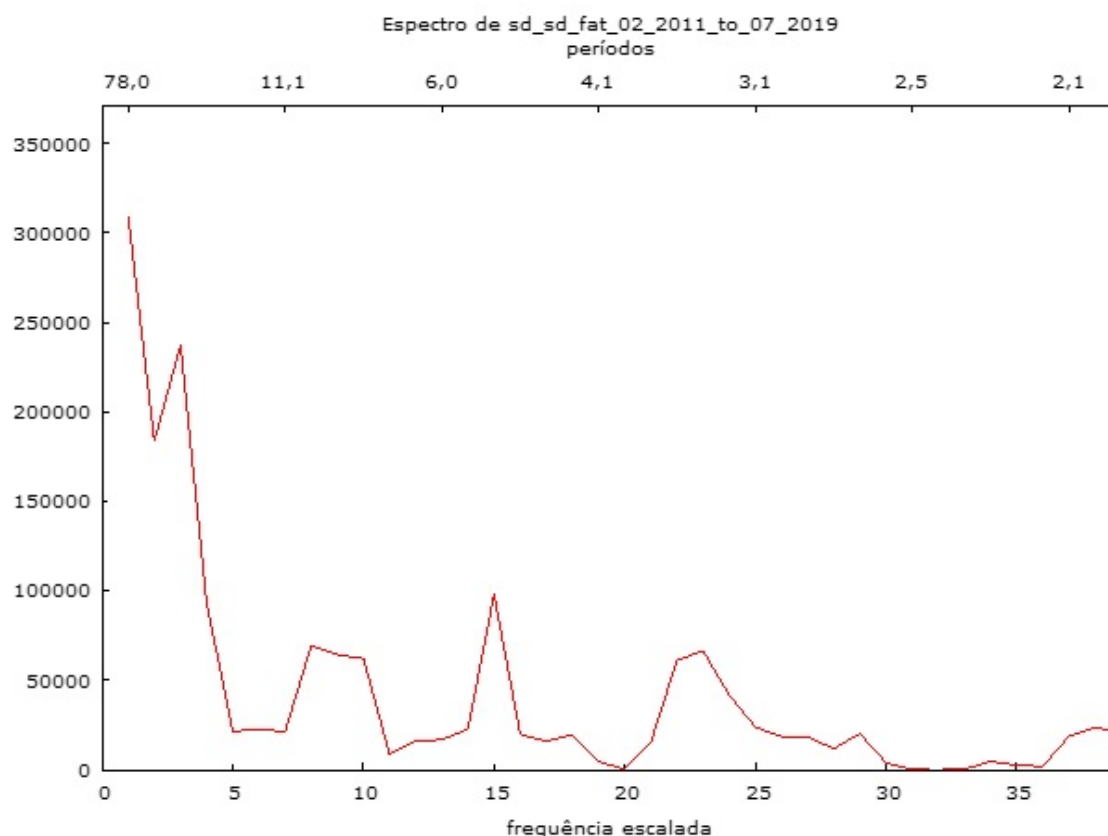
Figura 16 – Periodograma para a série com uma diferença sazonal



Fonte: Elaborado pelo programa Gretl

A [Figura 17](#) apresenta o periodograma para a série com duas diferenciações. Segundo a literatura, o número máximo de diferenciações possíveis em uma série temporal é até duas, passando disso, a análise estará inválida. Seguindo este raciocínio, define-se  $D=2$ .

Figura 17 – Periodograma para série com duas diferenciações sazonais



Fonte: Elaborado pelo programa Gretl

Com o filtro  $D$  definido, pode-se partir para a avaliação dos FAC e FACP da série com diferenciações sazonais, para a identificação dos filtros  $P$  e  $Q$ . A Tabela 6 apresenta os valores dos FAC e FACP para a série ajustada. Para cada diferenciação sazonal, o número de dados da série é reduzido em 12 unidade, como foram necessárias duas diferenciações sazonais, a amostra foi reduzida em 24 ficando com  $n' = 78$ . Usando novamente a fórmula  $c = \pm \frac{2}{\sqrt{n}}$  para calcular o novo valor de referência, obtém-se  $c' = 0.2264$ . Seguindo a mesma análise para os filtros  $q$  e  $p$ , demonstrados anteriormente para estimação filtros não sazonais, e usando o valor da nova referência  $c'$  obtém-se:  $P = 1$  e  $Q = 4$ , destacados em negrito, ficando com o modelo SARIMA (5, 2, 4) x (1, 2, 4). Mas antes aplicar o modelo, é necessário a validação, do mesmo, antes de qualquer tomada de decisões.

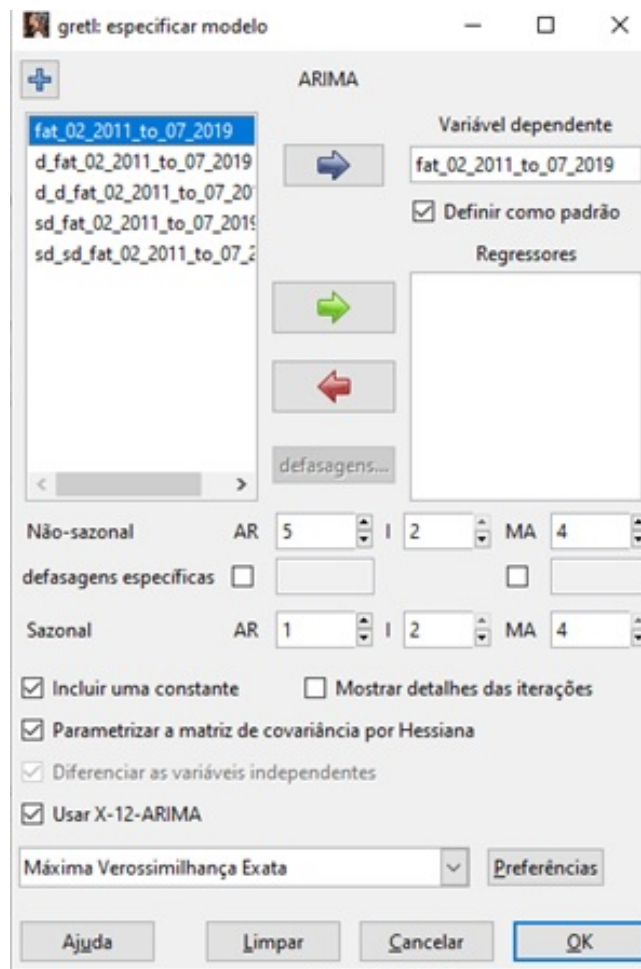
Tabela 6 – Correlograma da série temporal com duas diferenciações sazonais

Função de autocorrelação para d_d_fat_02_2011_to_07_2019 ***, **, * indicam significância aos níveis de 1%, 5% e 10% usando erro padrão $1/T^{0,5}$				
Defas.	FAC	FACP	Estat.	Q [p-valor]
1	0,5538***	<b>0,5538***</b>	248,565	[0,000]
2	0,3096***	0,0042	327,274	[0,000]
3	0,3238***	0,2174*	414,495	[0,000]
4	<b>0,2614**</b>	-0,0023	472,092	[0,000]
5	0,0882	-0,1283	478,737	[0,000]
6	0,2300**	0,2900**	524,602	[0,000]
7	0,1380	-0,2141*	541,349	[0,000]
8	0,0477	0,0710	543,376	[0,000]
9	0,0784	0,0141	548,929	[0,000]
10	0,1558	0,0608	571,213	[0,000]
8	0,0477	0,0710	543,376	[0,000]
9	0,0784	0,0141	548,929	[0,000]
10	0,1558	0,0608	571,213	[0,000]

Fonte: Elaborado pelo Autor

A [Figura 18](#) ilustra o modelo inserido no programa gretl, para validação, e com os resultados da avaliação do modelo inserido, ilustrados pela Tabela [7](#), é possível concluir se o modelo estimado é adequado ou não.

O modelo SARIMA (5, 2, 4) x (1, 2, 4) retornou muitos p-valores não significativos, destacados em negrito, acima de 5%, sendo necessário o ajuste do modelo. Com valores não significativos, o modelo não trará previsões confiáveis, nisso, é necessário o ajuste do modelo.

Figura 18 – Janela do gretl para incluir os *inputs* do modelo

Fonte: Elaborado pelo programa Gretl

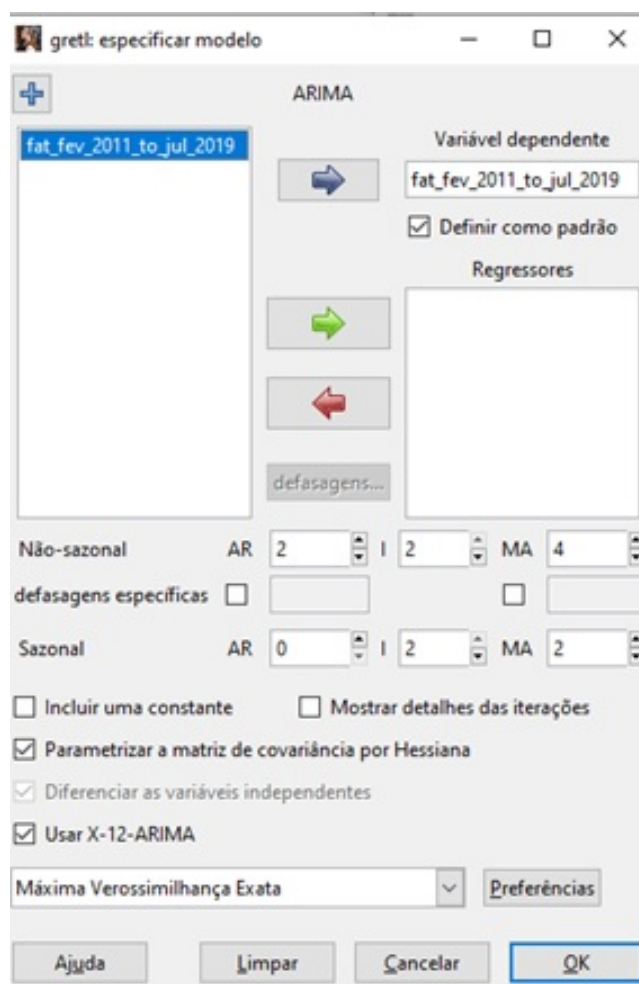
Tabela 7 – Avaliação do modelo estimado SARIMA (5, 2, 4) x (1, 2, 4)

Modelo 2: ARIMA, usando as observações 2013:04-2019:07 (T = 76)					
Estimado usando X-12-ARIMA (Máxima verossimilhança exata)					
Variável dependente: $(1-L)\hat{2}(1-Ls)\hat{2}$ fat_fev_2011_to_jul_2019					
	coeficiente	erro padrão	z	p-valor	
const	-1,53767	0,973616	-1,579	<b>0,1143</b>	
phi_1	-0,801710	0,473388	-1,694	<b>0,0903</b>	*
phi_2	-0,183863	0,377297	-0,4873	<b>0,6260</b>	
phi_3	-0,208526	0,205427	-1,015	<b>0,3101</b>	
phi_4	0,00989711	0,213715	0,04631	<b>0,9631</b>	
phi_5	-0,105157	0,200730	-0,5239	<b>0,6004</b>	
Phi_1	-0,995438	0,0414401	-24,02	1,67e-127	***
theta_1	-0,644627	0,482516	-1,336	<b>0,1816</b>	
theta_2	-0,818317	0,456572	-1,792	0,0731	*
theta_3	0,383279	0,496854	0,7714	<b>0,4405</b>	
textbf theta_4	0,326433	0,449771	0,7258	<b>0,4680</b>	
Theta_1	-0,346934	0,190914	-1,817	<b>0,0692</b>	*
Theta_2	-0,782189	0,192841	-4,056	4,99e-05	***
Theta_3	0,545608	0,178915	3,050	0,0023	***
Theta_4	0,0299938	0,185147	0,1620	<b>0,8713</b>	
Média var. dependente -15,17105 D.P. var. dependente 709,6670					
Média de inovações -7,182969 D.P. das inovações 258,5450					
Log da verossimilhança -544,4550 Critério de Akaike 1120,910					
Critério de Schwarz 1158,202 Critério Hannan-Quinn 1135,814					
		Real	Imaginária	Módulo	Frequência
AR					
Raiz	1	-10,349	0,0000	10,349	0,5000
Raiz	2	-0,6295	14,643	15,939	0,3146
Raiz	3	-0,6295	-14,643	15,939	-0,3146
Raiz	4	11,940	14,804	19,019	0,1420
Raiz	5	11,940	-14,804	19,019	-0,1420
AR (sazonal)					
Raiz	1	-10,046	0,0000	10,046	0,5000
MA					
Raiz	1	10,022	0,3303	10,553	0,0507
Raiz	2	10,022	-0,3303	10,553	-0,0507
Raiz	3	-15,893	0,4745	16,586	0,4538
Raiz	4	-15,893	-0,4745	16,586	-0,4538
MA (sazonal)					
Raiz	1	11,603	0,5782	12,964	0,0736
Raiz	2	11,603	-0,5782	12,964	-0,0736
Raiz	3	-10,177	0,0000	10,177	0,5000
Raiz	4	-194,935	0,0000	194,935	0,5000

Fonte: Elaborado pelo Autor

Com os ajustes feitos, avaliando os p-valores retornados pelo *Gretl*, chegou-se ao modelo SARIMA (2, 2, 4) X (0, 2, 2). Os p-valores dos encontrados são significativos, abaixo de 5%, sendo assim, o modelo SARIMA (2, 2, 4) X (0, 2, 2) é o mais adequado para a previsão. O novo modelo está ilustrado pela [Figura 19](#) e os resultados da avaliação estão contidos na [Tabela 8](#).

Figura 19 – Modelo ajustado



Fonte: Elaborado pelo programa Gretl

Tabela 8 – Avaliação do modelo ajustado SARIMA (2, 2, 4) x (0, 2, 2)

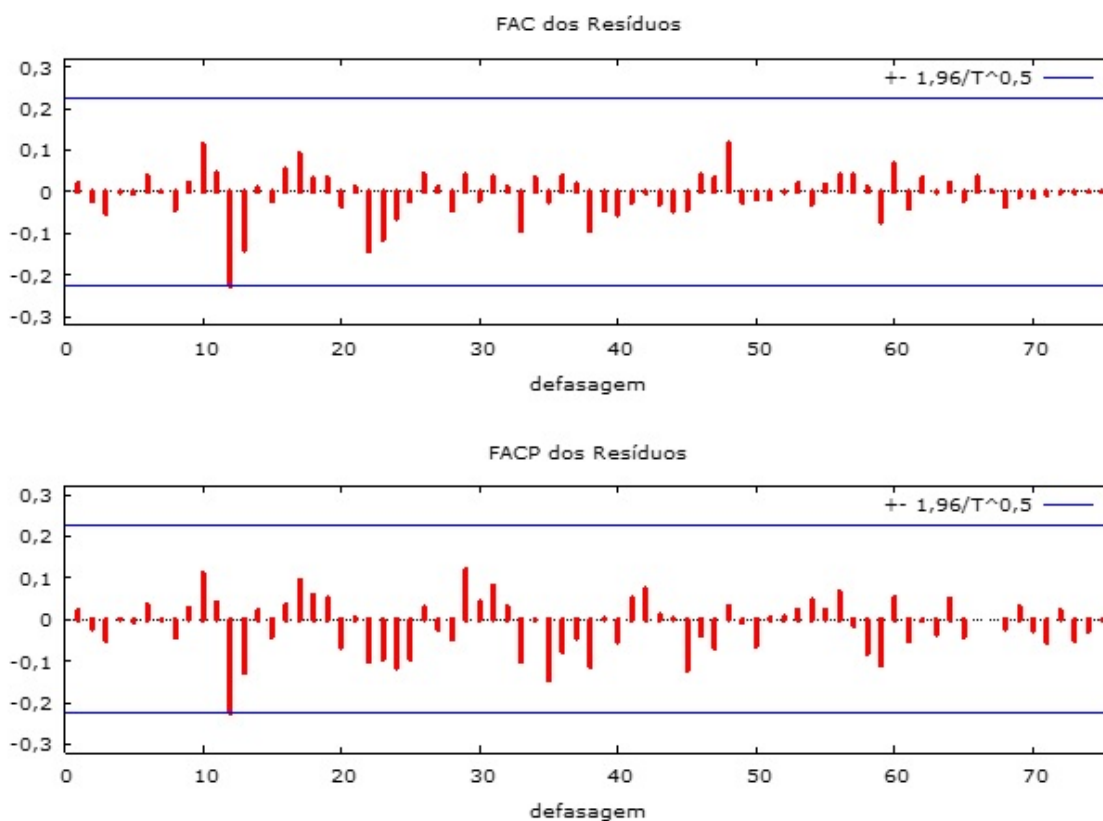
Modelo 3: ARIMA, usando as observações 2013:04-2019:07 (T = 76)  
 Estimado usando X-12-ARIMA (Máxima verossimilhança exata)  
 Variável dependente:  $(1-L)\hat{2}(1-Ls)\hat{2}$  fat\_fev\_2011\_to\_jul\_2019

coeficiente	erro	padrão	z	p-valor	
phi_1	-1,18472	0,0848925	-13,96	<b>2,91e-044</b>	***
phi_2	-0,594972	0,0840253	-7,081	<b>1,43e-012</b>	***
theta_1	-0,235044	0,0605057	-3,885	<b>0,0001</b>	***
theta_2	-0,957750	0,0612685	-15,63	<b>4,41e-055</b>	***
theta_3	-0,371660	0,0617052	-6,023	<b>1,71e-09</b>	***
theta_4	0,913577	0,0552002	16,55	<b>1,59e-061</b>	***
Theta_1	-1,43490	0,121169	-11,84	<b>2,36e-032</b>	***
Theta_2	0,779635	0,111105	7,017	<b>2,27e-012</b>	***

Fonte: Elaborado pelo Autor

Fazendo uma avaliação do modelo SARIMA (2, 2, 2) X (0, 2, 2) por meio dos resíduos exibidos pela [Figura 20](#), é possível verificar que os resíduos estão dentro da região de significância, sendo apenas no lag 12, tanto no FAC quanto FACP, estando no limite da região de significância, indicando que o modelo poderá ser usado.

Figura 20 – Resíduos do modelo ajustado

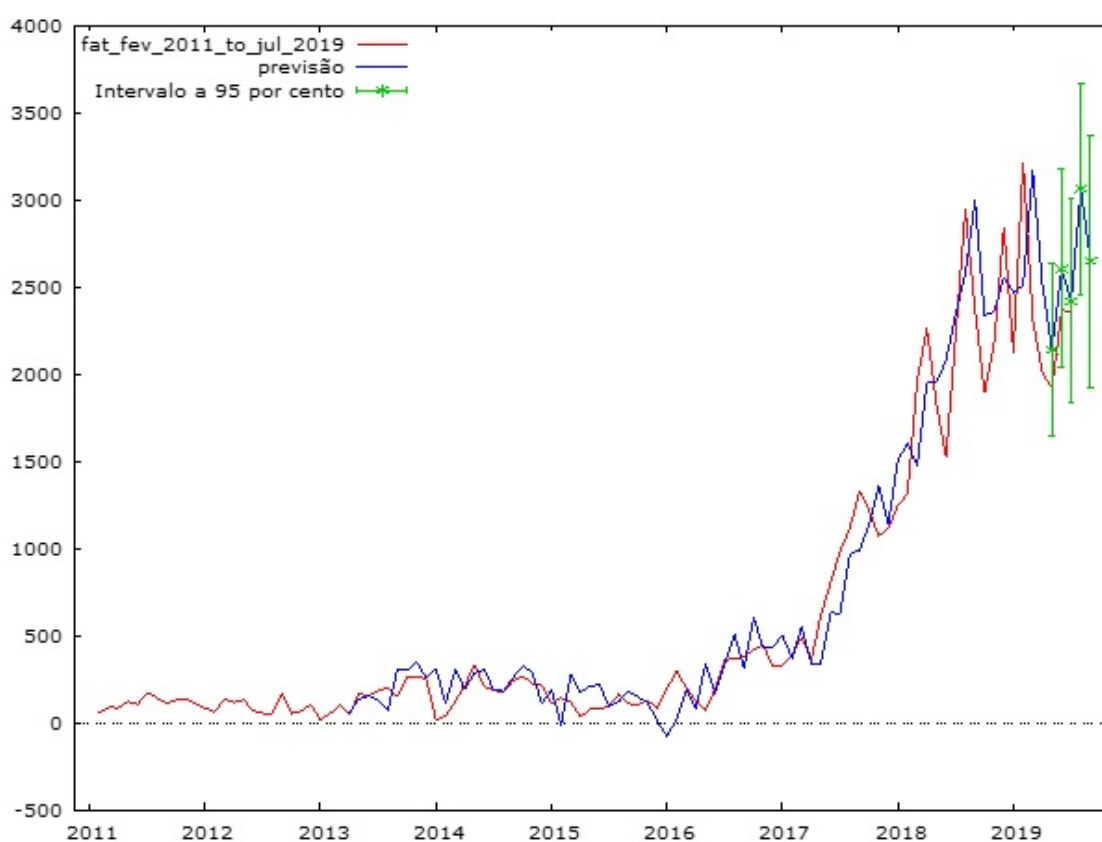


Fonte: Elaborado pelo programa Gretl

### 3.4 A PREVISÃO

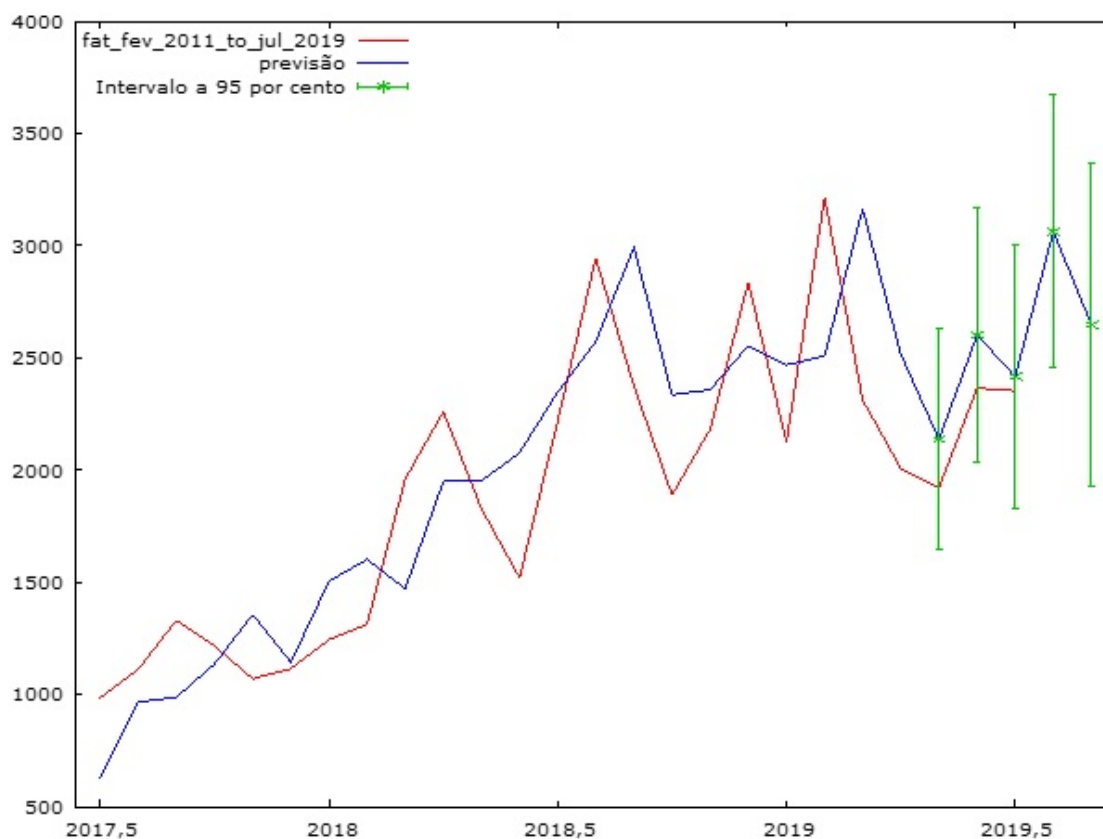
A [Figura 21](#) possibilita a visualização gráfica do desenvolvimento da previsão retornada pelo modelo ajustado, em azul, sobre a série temporal, em vermelho, e pode-se observar que o modelo apresentou uma boa aderência à série temporal. Para a avaliação da previsão da série temporal, a amostra foi reduzida em três períodos, de Julho de 2019 para Abril de 2019, isso serve para verificar se não há ruídos, para isso, os valores dos períodos de Maio, junho e Julho de 2019, devem estar dentro do intervalo de confiança do modelo. Os resultados para esta análise podem ser visualizados pelas [Figura 21](#), onde ilustra a série completa, e a [Figura 22](#), onde ilustra valores entre o período de Julho de 2017 até setembro de 2019, para uma visualização mais ampliada dos períodos mais recentes.

Figura 21 – Desenvolvimento do modelo SARIMA (2,2,4)x(0,2,2) sobre a série temporal completa



Fonte: Elaborado pelo programa Gretl

Figura 22 – Desenvolvimento do modelo Sarima(2,2,4)x(0,2,2) sobre a série temporal entre julho de 2017 até Setembro de 2019



Fonte: Elaborado pelo programa Gretl

Mesmo com a redução da amostra em três períodos, os p-valores do modelo continuam significativos, destacados em negrito na Tabela 9, indicando a validação do modelo SARIMA (2,2,4)x(0,2,2) para a série reduzida.

Tabela 9 – Avaliação do modelo ajustado SARIMA (2, 2, 4) x (0, 2, 2) para a série reduzida

Modelo 13: ARIMA, usando as observações 2013:04-2019:04 (T = 73) Estimado usando X-12-ARIMA (Máxima verossimilhança exata) Variável dependente: $(1-L)\hat{2}(1-Ls)\hat{2}$ fat_fev_2011_to_jul_2019					
coeficiente	erro	padrão	z	p-valor	
phi_1	-1,18445	0,0969313	-12,22	<b>2,45e-034</b>	***
phi_2	-0,621844	0,118474	-5,249	<b>1,53e-07</b>	***
theta_1	-0,235684	0,0629118	-3,746	<b>0,0002</b>	***
theta_2	-0,947166	0,0657725	-14,40	<b>5,13e-047</b>	***
theta_3	-0,380073	0,0696298	-5,458	<b>4,80e-08</b>	***
theta_4	0,908370	0,0570420	15,92	<b>4,28e-057</b>	***
Theta_1	-1,44405	0,125650	-11,49	<b>1,44e-030</b>	***
Theta_2	0,773754	0,114663	6,748	<b>1,50e-011</b>	***

Fonte: Elaborado pelo Autor

Os números de faturamento de Agosto e Setembro de 2019 não foram incluídos na análise, para

serem usados como referência na previsão fornecida modelo SARIMA (2, 2, 4) x (0, 2, 2). Com isso tivemos os números de faturamento para estes dois meses estão contido na Tabela 10:

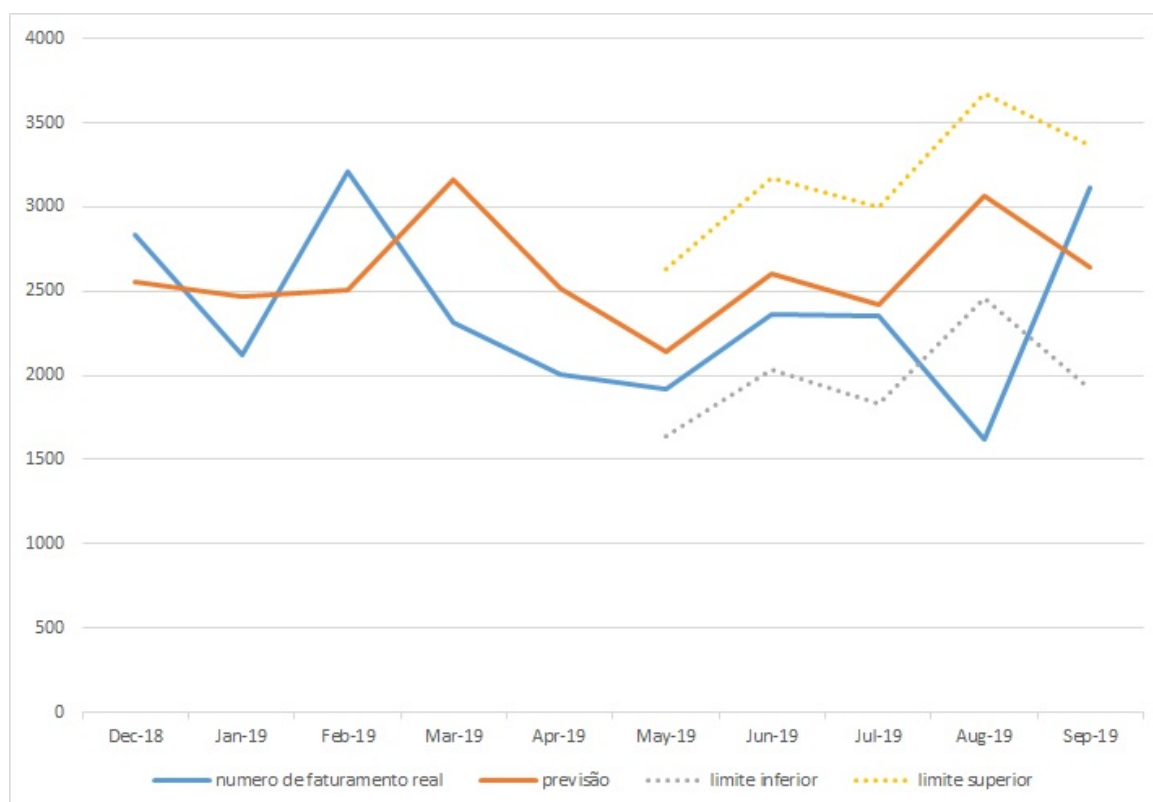
Tabela 10 – Faturamento entre os meses de Agosto e de Setembro para o ano de 2019

Mês	Número de faturamento
Ago/2019	1.621
Set/2019	3.113

Fonte: Elaborado pelo Autor

Uma observação importante a fazer é que o número de faturamento real de Agosto está abaixo do valor previsto e do intervalo de confiança do modelo, isso ocorreu pois a fábrica da montadora entrou em férias coletivas por 15 dias, afetando muito negativamente o número de faturamento do mês, como pode ser observado por meio da Figura 23 e dos dados da previsão do modelo apresentados pela Tabela 11, no entanto, supondo que as operações da fábrica estivessem normais durante o mês de Agosto, o número de faturamento estaria dentro do intervalo de confiança, pois o número de faturamento obtido em Setembro, quando a fábrica normalizou suas atividades, está dentro do intervalo de confiança, validando o modelo apesar do número de Agosto estar muito abaixo do previsto.

Figura 23 – Sobreposição entre a previsão e os dados reais para Agosto e Setembro de 2020



Fonte: Elaborado pelo Autor

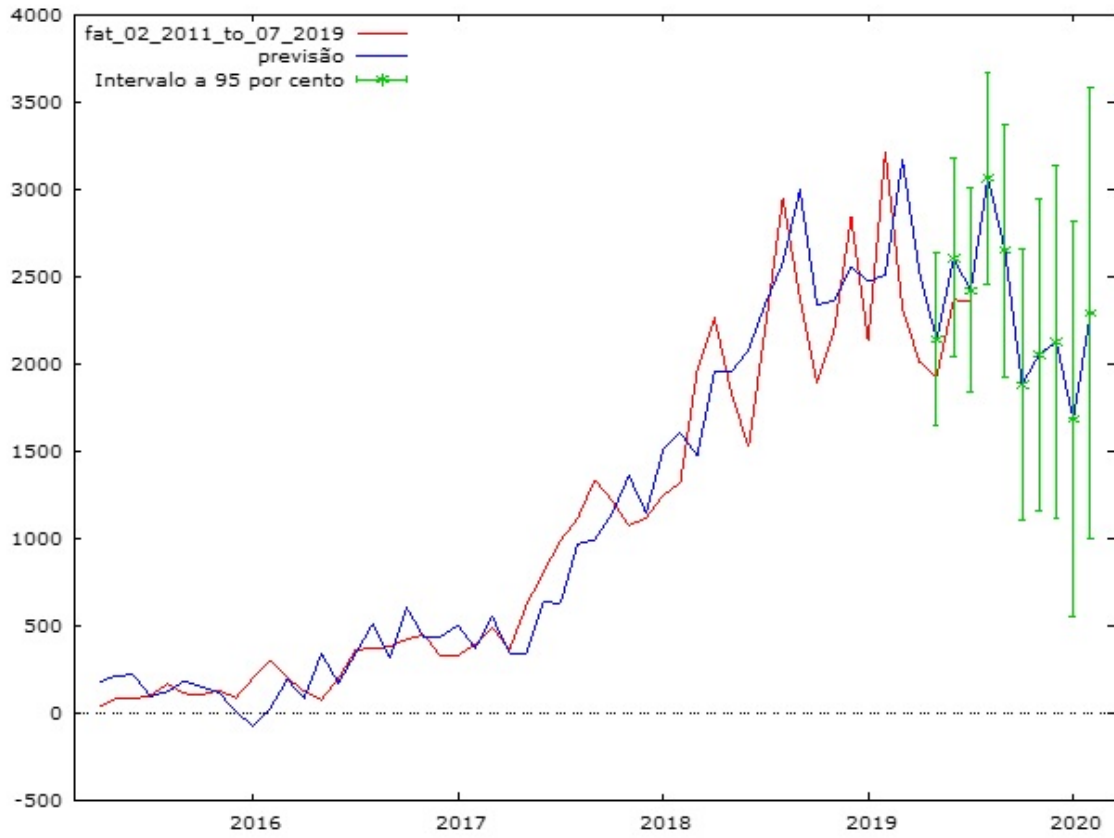
Tabela 11 – Dados de previsão do modelo

Para intervalos de confiança de 95%, $z(0,025) = 1,96$					
	fat_fev_2011_to_~	previsão	erro padrão	intervalo a 95%	
				limite inferior	limite superior
Dec-18	2836,00	2554,24			
Jan-19	2126,00	2467,96			
Feb-19	3213,00	2507,86			
Mar-19	2312,00	3163,82			
Apr-19	2007,00	2516,20			
May-19	1921,00	2136,31	251,146	1644,08	- 2628,55
Jun-19	2365,00	2601,48	290,315	2032,48	- 3170,49
Jul-19	2353,00	2415,96	298,290	1831,33	- 3000,60
Aug-19		<b>3062,05</b>	<b>309,917</b>	<b>2454,63</b>	- <b>3669,48</b>
Sep-19		<b>2646,81</b>	<b>369,038</b>	<b>1923,51</b>	- <b>3370,11</b>

Fonte: Elaborado pelo Autor

A [Figura 24](#) apresenta uma previsão mais prolongada, para fevereiro de 2020, e é notável que a quando mais longa for a previsão, maior será o intervalo de confiança, a 95%, pois a série é influenciada por diversos fatores externos de caráter econômico, geográfico e social, não sendo possível efetuar uma previsão confiável para um intervalo de tempo muito distante do presente. Mas em relação a previsões para períodos próximo do presente, é um método muito eficaz para a escolha de estratégias econômicas e produtivas, estando sempre preparado para possíveis problemas futuros.

Figura 24 – Previsão estendida para Fevereiro de 2020



Fonte: Elaborado pelo programa Gretl

## 4 CONCLUSÕES E RECOMENDAÇÕES PARA PESQUISAS FUTURAS

### 4.1 RESPOSTAS DAS QUESTÕES DA PESQUISA E RECOMENDAÇÕES PARA ESTUDOS FUTUROS

O objetivo principal deste trabalho foi trabalhar com uma série histórica de faturamento mensal de veículos destinados a Pessoas com Deficiência e Taxistas, para apresentar um modelo de previsão da demanda deste tipo de mercado. O tema abordado analisado de uma forma quantitativa, aplicando o método de pesquisa o modelo de previsão Box - Jenkins.

O questionário desta pesquisa foi respondida e o objetivo atendido pois, aplicando a metodologia de Box- Jenkins, foi possível obter um modelo de previsão adequada para a série histórica apresentada pela Tabela 12, no anexo A. O modelo SARIMA (2,2,4) X (0,2,2) demonstrou uma boa aderência sobre a série histórica, como ilustrado pelas Figura 21 e Figura 22. No ano de 2016, quando ocorreu o lançamento do Kicks, modelo SUV da Nissan, houve um crescimento muito acentuado no faturamento de veículos na categoria PCD e TÁXI. Se um estudo na previsão de demanda tivesse ocorrido em períodos anteriores ao crescimento repentino, o modelo retornaria uma previsão com crescimento mais suavizado para o mesmo período. Para crescimentos com este caráter, é aconselhável trabalhar com previsões curtas até obter um banco de dados, referente ao crescimento, suficientemente grande para gerar uma previsão confiável.

Conclui-se que a aplicação do método de previsão proposta por Box e Jenkins é muito adequado para a previsão de demanda para se ter uma estimativa confiável do faturamento para clientes PCD e Taxista, sendo uma ferramenta com grande potencial para ser aplicada em outros tipos de indústrias para se obter uma visão de cenários futuros.

### 4.2 RECOMENDAÇÕES PARA PESQUISAS FUTURAS

- Realizar um estudo com dados de outras montadoras de veículos destinados ao mesmo segmento apresnetado neste trabalho;
- Trabalhar com outros tipos de previsão para ter meios de comparação e verificar quais são os modelos mais adequados para série temporal em questão;
- Estudar outras categorias de veículos para verificar se os modelos de previsões apresentam comportamentos semelhantes em relação a previsão fornecida e os índices auto regressivos e de médias móveis.

## REFERÊNCIAS

- BERTRAND, J. W. M.; FRANSOO, J. C. Operations management research methodologies using quantitative modeling. **International Journal of Operations & Production Management**, Howard House, Wagon Ln, Bingley, Reino Unido, v. 22, n. 2, p. 241 – 265, 2002. ISSN 0144-3577. Disponível em: <https://www.emerald.com/insight/content/doi/10.1108/01443570210414338/full/html>, Acesso em: 20 Abr. 2019.
- BUENO, R. L. S. **Econometria de Séries Temporais**. 2nd, revised, updated. ed. São Paulo, SP, Brasil: Cengage Learning, 2012. ISBN 9788522111572.
- CHAIWUTTISAK, P. Forecasting export value in the automobile industry. **2018 5th International Conference on Business and Industrial Research (ICBIR)**, Bangkok, Tailândia, número do volume, n. número da página, p. 95 – 99, 2018. Disponível em: <https://ieeexplore.ieee.org/document/8391173>, Acesso em: 28 Abr. 2019.
- LIU, L. **Forecasting and Time Series Analysis Using SCA Statistical System**. Chicago, Illinois, EUA: Scientific Computing Associates ©corp, 1992. v. 1. Disponível em: [https://www.researchgate.net/publication/246213178\\_Forecasting\\_and\\_Time\\_Series\\_Analysis\\_Using\\_the\\_SCA\\_Statistical\\_System](https://www.researchgate.net/publication/246213178_Forecasting_and_Time_Series_Analysis_Using_the_SCA_Statistical_System), Acesso em: 30 Ago. 2019.
- MARTIN, A. C. *et al.* Análise de séries temporais para previsão da evolução do número de automóveis no município de Joinville. **Revista Espacios**, Colinas de Bello Monte, Caracas, Venezuela, v. 37, n. 06, p. 29 – 44, 2016. ISSN 0798-1015. Disponível em: <https://www.revistaespacios.com/a16v37n06/16370629.html>, Acesso em: 07 Abr. 2019.
- MDIC. **Setor Automotivo**. 2019. Disponível em: <http://www.mdic.gov.br/index.php/competitividade-industrial/setor-automotivo>, Acesso em: 07 Abr. 2019.
- MONTGOMERY, D. C.; JENNINGS, C. L.; KULAHCI, M. **Introduction to Time Series Analysis and Forecasting (Wiley Series in Probability and Statistics)**. Illinois, E.U.A.: Wiley Interscience, 2008. ISBN 9780471653974.
- MORETTIN, P.; TOLOI, M. C. **Análise de Séries Temporais**. 2. ed. São Paulo, SP, Brasil: Edgard Blucher, 2006. ISBN 8521203896,9788521203896.
- REAÇÃO: Revista nacional de reabilitação. 126. ed. São Paulo, SP, Brasil: Comunicação & Marketing, 2019.
- SOUZA, R. C. Metodologias para a análise e previsão de séries temporais univariadas e multivariadas. **Brazilian Review of Econometrics**, Rio de Janeiro, RJ, Brasil, v. 1, n. 2, p. 78–105, 1981. ISSN 1980-2447. Disponível em: <http://bibliotecadigital.fgv.br/ojs/index.php/bre/article/view/3167>.
- STORTI, E. *et al.* Customized knowledge discovery in databases methodology for the control of assembly systems. **Machines**, St. Alban-Anlage, Basel, Suíça, v. 6, n. 4, 2018. ISSN 2075-1702. Disponível em: <http://www.mdpi.com/2075-1702/6/4/45>, Acesso em: 05 Mai. 2019.
- SUBRAMANIYAN, M. *et al.* A data-driven algorithm to predict throughput bottlenecks in a production system based on active periods of the machines. **Computers & Industrial Engineering**, Langford Lane, Oxford, Reino Unido, v. 125, n. 533, p. 533 – 544, 2018. ISSN 0360-8352. Disponível em: <https://www.sciencedirect.com/science/article/pii/S0360835218301608?via%3Dihub>, Acesso em: 20 Abr. 2019.

SUEPTAETRAKUN, W.; TANGTHONG, S. Determinants of collaboration and innovation towards organization performance within thailand's automotive parts industry. **Journal of Business & Retail Management Research**, Harrow, Middlesex, Reino Unido, v. 13, 12 2018. ISSN 1751-8202. Disponível em: <https://jbrmr.com/details&cid=447>. Acesso em: 20 Abr. 2019.

VARGAS, C. A. G.; CORTÉS, M. E. Automobile spare-parts forecasting: A comparative study of time series methods. **International Journal of Automotive and Mechanical Engineering**, v. 14, n. 3898, p. 3898 – 3914, 2017. ISSN 2180-1606. Disponível em: [http://ijame.ump.edu.my/images/Vol\\_14\\_Issue1/7\\_vargas%20and%20corts.pdf](http://ijame.ump.edu.my/images/Vol_14_Issue1/7_vargas%20and%20corts.pdf). Acesso em: 29 Abr. 2019.

**ANEXO A – DADOS DO NÚMERO DE FATURAMENTO DE VEÍCULOS PARA  
PESSOAS COM DEFICIÊNCIA E TAXÍSTAS ENTRE O PERÍODO DE FEVEREIRO DE  
2011 A JULHO DE 2019**

Tabela 12 – Histórico de faturamento para pessoas com deficiência e taxistas entre o período de Fevereiro de 2011 a Julho de 2019

periodo	nº de faturamento	periodo	nº de faturamento	periodo	nº de faturamento
Feb/2011	60	dez/2013	249	out/2016	420
Mar/2011	87	jan/2014	15	nov/2016	445
Apr/2011	85	fev/2014	39	dez/2016	320
May/2011	122	mar/2014	122	jan/2017	327
Jun/2011	104	abr/2014	212	fev/2017	390
Jul/2011	173	mai/2014	334	mar/2017	489
Aug/2011	143	jun/2014	204	abr/2017	360
Sep/2011	109	jul/2014	189	mai/2017	614
Oct/2011	129	ago/2014	171	jun/2017	804
Nov/2011	138	set/2014	241	jul/2017	985
Dec/2011	111	out/2014	268	ago/2017	1113
Jan/2012	82	nov/2014	225	set/2017	1331
Feb/2012	62	dez/2014	213	out/2017	1216
Mar/2012	136	jan/2015	112	nov/2017	1070
Apr/2012	116	fev/2015	141	dez/2017	1115
May/2012	131	mar/2015	120	jan/2018	1245
Jun/2012	68	abr/2015	32	fev/2018	1313
Jul/2012	53	mai/2015	78	mar/2018	1960
Aug/2012	53	jun/2015	78	abr/2018	2260
Sep/2012	168	jul/2015	98	mai/2018	1828
Oct/2012	53	ago/2015	164	jun/2018	1520
Nov/2012	65	set/2015	108	jul/2018	2219
Dec/2012	103	out/2015	102	ago/2018	2944
Jan/2013	15	nov/2015	127	set/2018	2374
Feb/2013	53	dez/2015	83	out/2018	1890
Mar/2013	98	jan/2016	196	nov/2018	2182
Apr/2013	46	fev/2016	296	dez/2018	2836
May/2013	164	mar/2016	201	jan/2019	2126
Jun/2013	156	abr/2016	119	fev/2019	3213
Jul/2013	181	mai/2016	73	mar/2019	2312
ago/2013	201	jun/2016	192	abr/2019	2007
set/2013	150	Jul/2016	357	mai/2019	1921
out/2013	266	ago/2016	368	Jun/2019	2365
nov/2013	264	set/2016	376	Jul/2019	2353

Fonte: dados fornecidos pela nissan do Brasil