

Universidade Federal de Minas Gerais
Instituto de Ciências Exatas
Departamento de Estatística
Curso de Especialização em Estatística

Pais adolescentes e jovens: o que a distribuição dos dados revela para uma regional de Belo Horizonte entre 2013 e 2018?

Carla Gomes Franco

Belo Horizonte
2020

Carla Gomes Franco

Pais adolescentes e jovens: o que a distribuição dos dados revela para uma regional de Belo Horizonte entre 2013 e 2018?

Monografia de especialização em estatística apresentada ao Instituto de Ciências Exatas da Universidade Federal de Minas Gerais, como requisito parcial à obtenção do título de Especialista em Estatística.

Orientadora: Prof^a. Ela Mercedes Medrano

Belo Horizonte
2020

2020, Carla Gomes Franco
Todos os direitos reservados

Ficha catalográfica elaborada pela bibliotecária Belkiz Inez Rezende
Costa CRB 6ª Região nº 1510

Franco, Carla Gomes

F825p Pais adolescentes e jovens: o que a distribuição dos dados revela para uma regional de Belo Horizonte entre 2013 e 2018? / Carla Gomes Franco. — Belo Horizonte, 2020.

96.f.. il.; 29 cm.

Monografia (especialização) - Universidade Federal de Minas Gerais – Departamento de Estatística.

Orientadora: Ela Mercedes Medrano de Toscano.

1. Estatística. 2. Paternidade. 3. Adolescência. 4. Modelo autorregressivo integrado de médias móveis. 5. Modelo Holt-Winters. I. Orientadora. II. Título.

CDU 519.2 (043)



Universidade Federal de Minas Gerais
Instituto de Ciências Exatas
Departamento de Estatística
Programa de Pós-Graduação / Especialização
Av. Pres. Antônio Carlos, 6627 - Pampulha
31270-901 – Belo Horizonte – MG

E-mail: pgest@ufmg.br
Tel: 3409-5923 – FAX: 3409-5924

ATA DO 210ª. TRABALHO DE FIM DE CURSO DE ESPECIALIZAÇÃO EM ESTATÍSTICA DE CARLA GOMES FRANCO.

Aos dois dias do mês de julho de 2020, às 13:00 horas, com utilização de recursos de videoconferência a distância, reuniram-se os professores abaixo relacionados, formando a Comissão Examinadora homologada pela Comissão do Curso de Especialização em Estatística, para julgar a apresentação do trabalho de fim de curso da aluna **Carla Gomes Franco**, intitulado: “Pais adolescentes e jovens: o que a distribuição dos dados revela para uma regional de Belo Horizonte entre 2013 e 2018?”, como requisito para obtenção do Grau de Especialista em Estatística. Abrindo a sessão, a Presidente da Comissão, Professora Ela Mercedes Medrano de Toscano – Orientadora, após dar conhecimento aos presentes do teor das normas regulamentares, passou a palavra à candidata para apresentação de seu trabalho. Seguiu-se a arguição pelos examinadores com a respectiva defesa da candidata. Após a defesa, os membros da banca examinadora reuniram-se sem a presença da candidata e do público, para julgamento e expedição do resultado final. Foi atribuída a seguinte indicação: a candidata foi considerada Aprovada condicional às modificações sugeridas pela banca examinadora no prazo de 30 dias a partir da data de hoje por unanimidade. O resultado final foi comunicado publicamente à candidata pela Presidente da Comissão. Nada mais havendo a tratar, a Presidente encerrou a reunião e lavrou a presente Ata, que será assinada por todos os membros participantes da banca examinadora. Belo Horizonte, 02 de julho de 2020.

Profa. Ela Mercedes Medrano de Toscano (Orientadora)
Departamento de Estatística / UFMG

Profa. Paula de Miranda Ribeiro
FCE - DPTO Demografia / UFMG

Profa. Sueli Aparecida Mingoti
Departamento de Estatística/UFMG

A todos os pais, em especial àqueles que se dedicam a um novo padrão de paternidade, independente da idade.

Resumo

A paternidade tem sido cada vez mais debatida e apresentada entre as mídias sociais como uma perspectiva de uma mudança de paradigma no cumprimento de tal papel. No entanto, a paternidade entre adolescentes e jovens ainda é pouco discutida, estudada e portanto, pouco conhecida. **Objetivo:** Estudar a distribuição dos dados que podem revelar sobre fecundidade masculina, portanto entre homens que foram pais entre os 15 a 24 anos e encontrar modelos que acompanhem a variabilidade dos dados, a fim de fazer previsões. **Dados:** A base de dados para o trabalho originou-se dos registros civis de nascimentos realizados em uma unidade de Cartório de Belo Horizonte, localizado em uma região de Belo Horizonte – MG, no período de 2013 a 2018. **Método:** Foi utilizada a metodologia de séries temporais incluindo os modelos SARIMA e Holt-Winters. **Resultados:** Os resultados revelam que a distribuição por grupo etário do pai, se assemelha o padrão da distribuição observado entre as mulheres, com a “barriga” apontando a ocorrência de maior número de casos entre as idades mais jovens. **Considerações:** Nota-se uma tendência de redução do número de casos para todos os grupos com destaque para aqueles pais adolescentes e jovens. Os modelos conseguiram acompanhar bem a variabilidade dos dados, e as previsões calculadas mostraram-se satisfatórias em relação ao número de pais adolescentes e jovens observados.

Palavras-chave: paternidade; adolescência; Modelo SARIMA; Modelo Holt-Winters
Aditivo

Abstract

Fatherhood has been increasingly debated and presented among social media as a perspective of paradigm change in fulfilling such role. However, paternity among teenagers and in the youth less discussed, studied and therefore little known. Objectives: To study the distribution of data that can reveal male fertility, therefore among men who were parents between the ages of 15 and 24, and to find models that follow the variability of the data, in order to make predictions. Data: The database for the work originated from the civil birth registrations found in a unit of the Registry Office of Belo Horizonte, located in a region of Belo Horizonte - MG, in the period from 2013 to 2018. Methodology: this work used time series forecast with the SARIMA and Holt-Winters models. Results: The results reveal that the distribution by the father's age group, is similar to the pattern of distribution observed among women, with the "belly" pointing to the occurrence of a higher number of cases among the younger ages. Considerations: There is a tendency to reduce the number of cases in all groups, especially those adolescent and young parents. The models were well succeeded and followed the variability of data, and the calculated predictions proved to be in line with the number of teenagers and young parents observed.

Keywords: Fatherhood; teenagers; SARIMA models; Holt-Winters Additive.

Sumário

AGRADECIMENTOS	10
LISTA DE ILUSTRAÇÕES	11
LISTA DE TABELAS	8
LISTA DE SIGLAS	8
1 - INTRODUÇÃO	16
1.1 - Justificativa.....	18
1.2 - Objetivos	22
1.2.1 - Objetivos Específicos.....	23
2 - METODOLOGIA	23
2.1 - Método de Decomposição.....	29
2.2 - Modelo ARIMA	31
2.3 - Alisamento Exponencial.....	32
2.4 Medidas de Acurácia	33
3 - ANÁLISE DOS DADOS.....	34
3.1 - Descrição dos Registros	34
3.1.1 - Descrição da variável idade dos pais.....	44
3.2 – Descrição dos Grupos Etários de 15 a 19 anos e de 20 a 24 anos.....	46
4 - APLICAÇÃO DOS MODELOS DE SÉRIE TEMPORAL.....	51
4.1 - Ajuste dos modelos para os registros civis de filhos de pais entre 15 e 19 anos	51
4.1.1 – Modelagem Através do Método de Decomposição Clássica	51
4.1.2 - Modelagem Através do Método ARIMA.....	56
4.1.3 - Modelagem Através do Método de Alisamento Exponencial - Holt.....	62
4.1.4 - Modelagem Através do Método de Alisamento Exponencial – Holt-Winter	65
4.2 - Ajuste dos modelos para os registros civis de filhos de pais entre 20 e 24 anos	68
4.2.1 – Modelagem Através do Método de Decomposição Clássica	68

4.2.2 - Modelagem Através do Método ARIMA.....	72
4.2.3 - Modelagem Através do Método de Alisamento Exponencial – Holt.....	79
4.2.4 - Modelagem Através do Método de Alisamento Exponencial – Holt-Winters.....	82
4.3 - Comparação e Análise dos Modelos.....	85
5 - CONSIDERAÇÕES FINAIS	87
REFERÊNCIAS BIBLIOGRÁFICAS	90
ANEXO 1 - DISTRIBUIÇÃO DE REGISTROS POR IDADE DO PAI	92
ANEXO 2 - DISTRIBUIÇÃO DAS IDADES DOS PAIS POR ANO	94

Agradecimentos

Agradeço o apoio de todos que vivenciaram, sentiram e vibraram neste percurso comigo:

Em especial à minha orientadora, que aderiu ao tema e a causa de prontidão e se dedicou, pacientemente a me ajudar, a entender os dados e realidade dos mesmos! Sua colaboração foi muito grande e motivadora, e suas contribuições foram importantíssimas e cuidadosas!

Agradeço imensamente ao Antônio que, pela sua dedicação no trabalho estabeleceu uma relação de confiança com o representante do cartório, que gentilmente, cedeu os dados para este trabalho. À ele, Robson, e toda a equipe do Cartório de Registro Civil e Notas de Venda Nova que se dedicou a me ajudar no dia-a-dia dos registros, meu muitíssimo obrigada pela presteza e disposição! Pude entender e compreender os bastidores dos dados do sistema de estatísticas vitais, no qual se inserem os registros civis.

Agradeço em especial a minha mãe e minha irmã gêmea, sem justificativas extras, apenas por ocuparem tal lugar em minha vida, indicando apoio incondicional!

Ao professor Roberto Quinino, do departamento de estatística da UFMG que se dispôs a me ouvir em momentos difíceis; à professora Sueli Mingoti, também do departamento, pelo apoio e incentivo demonstrados muito antes do meu ingresso no curso e pela composição da banca examinadora; à professora Paula Miranda-Ribeiro, do Cedeplar, que me acompanhou em trabalho com a mesma temática, porém com abordagem diferente, agradeço por me dar a honra do aceite em compor a banca; aos colaboradores do departamento de estatística por toda simpatia e colaboração em todas as etapas do curso.

Agradeço muito a Ana Julia Camara, doutoranda em estatística pela UFMG, pelos esclarecimentos técnicos em momentos pontuais e muito muito importantes!

Aos colegas de sala, especialmente Ana Cristina, Sara e Mirele pelas trocas e motivação.

Aos colegas de profissão por onde passei, que se dispuseram à compreensão quando precisei! .

Lista de Ilustrações

Figura 1 - Significado de ser pai.....	21
Figura 2 - Representação de um processo estocástico.....	27
Figura 3 - Distribuição dos registros de nascimentos segundo sexo, 2013-2018	35
Figura 4 - Distribuição dos registros de nascimentos segundo idade do pai, 2013-2018.....	35
Figura 5 - Distribuição dos registros de nascimentos segundo ano, 2014-2017	37
Figura 6 - Distribuição dos registros de nascimentos segundo ano, por grupo etário do pai, 2013-2018	38
Figura 7 - Distribuição percentual dos registros de nascimento segundo ano, por grupo etário do pai, 2013-2018	39
Figura 8 - Variação percentual entre os anos, considerando o início do período, jul/2013.....	42
Figura 9 - Diagrama de caixa segundo ano de nascimento.....	43
Figura 10 - Diagrama de caixa segundo mês de nascimento	44
Figura 11 - Distribuição dos registros mensais em perspectiva histórica.	44
Figura 12 - Histograma da idade do pai para o total de registros (n: 41675)	45
Figura 13 - Distribuição dos registros de nascimentos segundo sexo, no grupo etário de pais de 15 a 19 anos, 2013-2018.....	46
Figura 14 - Distribuição dos registros de nascimentos segundo sexo, no grupo etário de pais de 20 a 24 anos, 2013-2018.....	46
Figura 15 - Distribuição dos registros de nascimentos segundo ano, por grupo etário de interesse... ..	47
Figura 16 - Variação percentual entre os anos de análise para os grupos etários de interesse, 2013-2018.....	48
Figura 17 - Variação percentual entre os anos, para os grupos etários de interesse, considerando o início do período, jul/2013.....	48
Figura 18 - Distribuição dos registros mensais em perspectiva histórica por grupo etário de interesse.	49
Figura 19 - Diagrama de caixa dos registros de nascimentos para os para os grupos etários de 15 a 19 anos e de 20 a 24 anos.....	50
Figura 20 - Diagrama dos registros segundo ano de nascimento para o grupo etário de 15 a 19 anos.	50
Figura 21 - Diagrama dos registros segundo ano de nascimento para o grupo etário de 20 a 24 anos.	50
Figura 22 - Diagrama dos registros segundo mês de nascimento para o grupo etário de 15 a 19 anos.	51
Figura 23 - Diagrama dos registros segundo mês de nascimento para o grupo etário de 20 a 24 anos.	51
Figura 24 - Análise de tendência para a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo: equação (27), 2013-2018.....	52
Figura 25 - Gráficos dos resíduos da análise de tendência para a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo: equação (27), 2013-2018.....	53
Figura 26 - Análise de decomposição (tendência e sazonalidade) para registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equação (28).	54
Figura 27 - Análise dos índices sazonais para a série registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equação (28).	55
Figura 28 - Análise de sazonalidade para a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equação (28).	55

Figura 29 – Gráficos dos resíduos da análise de decomposição (tendência e sazonalidade) para a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equação (28).	56
Figura 30 - Função de autocorrelação da série de registros de filhos de pais entre 15 e 19 anos – dados mensais com a diferença aplicada (d=1).	56
Figura 31 - Função de autocorrelação parcial da série de registros de filhos de pais entre 15 e 19 anos – dados mensais com a diferença aplicada (d=1).	57
Figura 32 – Correlograma do vetor de resíduos do modelo ajustado para a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equação (30).	58
Figura 33 – Dados ajustados segundo o modelo ARIMA (2, 1, 0) para a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equação (30).	59
Figura 34 - Gráficos dos resíduos do modelo ARIMA (2, 1, 0) para a série de registros de filhos de pais entre 15 e 19 anos.....	59
<i>Figura 35 – Correlograma do vetor de resíduos do modelo ARIMA (2, 1, 0) da série de registros de filhos de pais entre 15 e 19 anos, n=48.</i>	<i>61</i>
Figura 36 - Previsão para a série de registros de filhos de pais entre 15 e 19 anos, modelo ARIMA (2, 1, 0), n=48.....	62
Figura 37 - Análise de suavização (nível e tendência) para a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equações (32 e 33), n=60.....	63
Figura 38 - Gráficos dos resíduos do modelo Holt para a a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equações (32 e 33), n=60.	63
Figura 39 - Análise de suavização (nível e tendência) para a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equações (34 e 35), n=48.....	65
Figura 40 - Previsão para a série de registros de filhos de pais entre 15 e 19 anos, modelo Holt, n=48.	65
Figura 41 - Análise de suavização (nível, tendência e sazonalidade) para a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equações (36, 37 e 38), n=60.....	66
Figura 42 - Análise de suavização (nível, tendência e sazonalidade) para a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equações (36, 37 e 38), n=48.....	67
Figura 43 - Previsão para a série de registros de filhos de pais entre 15 e 19 anos, modelo Holt-Winter, n=48.	68
Figura 44 - Análise de tendência para a série registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo: equação (39), 2013-2018.....	69
Figura 45 - Gráficos dos resíduos da análise de tendência para a série de registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo: equação (39), 2013-2018.	69
Figura 46 - Análise de decomposição (tendência e sazonalidade) para a série de registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo ajustado: equação (40).	70
Figura 47 - Análise dos índices sazonais para a série de registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo ajustado: equação (40).	71
Figura 48 - Análise de sazonalidade para a serie de registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo ajustado: equação (40).	72
Figura 49 - Gráficos dos resíduos da análise de decomposição (tendência e sazonalidade) para a série de registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo ajustado: equação (40).	72
Figura 50 - Função de autocorrelação da série de registros de filhos de pais entre 20 e 24 anos.....	73
Figura 51 - Função de autocorrelação parcial da série registros de filhos de pais entre 20 e 24 anos.	73
Figura 52 - Correlograma do vetor de resíduos do modelo ajustado para a série de registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo ajustado: equação (41), n=60.	75

Figura 53 – Dados ajustados segundo o modelo ARIMA (0, 1, 1) para a série de registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo ajustado: equação (41), n=60.....	75
Figura 54 - Gráficos dos resíduos do modelo ARIMA (2, 1, 0) para a série de registros de filhos de pais entre 15 e 19 anos, n=60.....	76
Figura 55 - Correlograma do vetor de resíduos do modelo ARIMA (0, 1, 1) da série de registros de filhos de pais entre 20 e 24 anos, n=48.....	77
Figura 56 - Previsão para a série de registros de filhos de pais entre 20 e 24 anos, modelo ARIMA (0, 1, 1), n=48.....	79
Figura 57 - Análise de suavização (nível e tendência) para a série de registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo ajustado: equações (43 e 44), n=60.....	80
Figura 58 - Gráficos dos resíduos do modelo Holt para a a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equações (42 e 44), n=60.	80
Figura 59 - Análise de suavização (nível e tendência) para a série de registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo ajustado: equações (45 e 46), n=48.....	82
Figura 60 - Previsão para a série de registros de filhos de pais entre 20 e 249 anos, modelo Holt, n=48.....	82
Figura 61 - Análise de suavização (nível, tendência e sazonalidade) para a série de registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo ajustado: equações (47, 48 e 49), n=60.....	83
Figura 62 - Análise de suavização (nível, tendência e sazonalidade) para a série de registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo ajustado: equações (40, 41 e 42), n=48.....	84
Figura 63 - Previsão para a série de registros de nascimentos de filhos de pais entre 20 e 24 anos, modelo Holt-Winter, n=48.	85

Lista de Tabelas

Tabela 1 - Distribuição dos registros de nascimentos segundo grupo etário do pai, 2013-2018	36
Tabela 2 - Distribuição dos registros de nascimentos segundo local de nascimento e grupo etário do pai, 2013-2018	37
Tabela 3 - Distribuição de registros por ano de nascimento, segundo grupo etário do pai	40
Tabela 4 - Distribuição de registros por ano e mês de nascimento e variação percentual	41

Lista de Siglas

ARIMA – Modelo Autorregressivo Integrado de Médias Móveis

DataSUS – Departamento de Informática do Sistema Único de Saúde do Brasil

DHS – Demographic and Health Survey

FAC – Função de Autocorrelação

FACP – Função de Autocorrelação Parcial

MAD – Erro Absoluto Médio ou Desvio Absoluto Médio

MAPE – Erro Percentual Absoluto Médio

MPE – Erro Percentual Médio

MSE – Erro Quadrático Médio ou Desvio Quadrático Médio

PNAISH – Política Nacional de Atenção Integral à Saúde do Homem

SINASC – Sistema de Informações sobre Nascidos Vivos

UPA – Unidade de Pronto Atendimento

1 - Introdução

Este trabalho se presa a estudar um tema que tem se tornado cada vez mais debatido entre as mídias sociais: a paternidade. Do ponto de vista prático, a bandeira da paternidade tem sido levantada no sentido de despertar entre os homens um novo exercício de paternidade: aquela que acolhe mais e não somente sustenta financeira e materialmente, rompendo com os padrões consolidados de “mero” papel mantenedor. Do ponto de vista acadêmico, a visibilidade dos pais vai de encontro à esse novo formato, mas também coloca em discussão a revisão de estereótipos criados no que diz respeito à fecundidade na adolescência. Ou seja, paternidade na adolescência também é assunto discutível, sobretudo no que diz respeito aos impactos vivenciados por tais indivíduos. O trabalho buscou analisar os dados sobre paternidade, observado através dos registros civis obtidos junto a uma unidade de cartório de Belo Horizonte. Tal análise correspondeu às etapas de limpeza dos dados; análise descritiva; aplicação dos modelos de série temporal, e; verificação do melhor modelo.

Quando se fala em fecundidade, dificilmente a figura paterna – o homem enquanto cuidador – surge no imaginário de quem ouve o termo. No âmbito dos estudos demográficos, entre o tripé que sustenta a área – fecundidade, mortalidade e migração – a análise da fecundidade é considerada mais complexa que a análise da mortalidade, sobretudo pelo fato de a *“human fertility involves two individuals of opposite sexes”* (PRESTON *et al.*, 2001, p. 92). Considerando, também, que a fecundidade possui limitações fisiológicas, do ponto de vista da fertilidade, onde nem todos os indivíduos – sejam homens ou mulheres – são capazes ou sujeitos a reproduzirem, a escolha por apenas um sexo para a definição das técnicas de medição, limita os riscos de erros. Além disso, por conhecer os limites de reprodução feminina, que se configura entre a menarca e a menopausa, os autores compreendem que a idade seja um importante fator no dimensionamento da análise da fecundidade feminina, a definindo como o padrão de estudos sobre fecundidade. Assim consolidou-se o cálculo da fecundidade feminina como padrão nos estudos sobre fecundidade na demografia.

É compreensível, portanto, que quando se trata de estudos sobre fecundidade, os homens constituam-se uma “minoridade negligenciada” segundo a definição de

Coleman *apud* Zhang (2011). Isso se evidencia pela pouca abordagem sobre as perspectivas de fecundidade masculina no que tange aos métodos de contracepção; às decisões de reprodução; ao desejo de ter filhos; ao envolvimento nos cuidados diários com estes, e demais temas que podem relacionar a fecundidade e a paternidade. Oliveira *et al.* (2009) afirmam que a “ênfase quase exclusiva na mulher nos estudos sobre reprodução reflete a construção naturalizada do gênero feminino” (p. 276), levando a uma definição restrita de papéis dentro e fora do domicílio nas diversas atividades.

Embora a mulher seja tratada como a principal protagonista no processo de concepção, geração e criação de filhos e, desta forma, ser o foco dos estudos sobre fecundidade, aos poucos os homens vêm sendo incluídos nas preocupações dos pesquisadores, uma vez que fazem parte não apenas da concepção, mas também do cuidado e envolvimento futuro com a criança. Lyra (1997) ao se debruçar sobre o problema da invisibilidade dos pais nos estudos sobre fecundidade, identificou o não lugar na sociedade brasileira por parte dos pais adolescentes, fundamentalmente devido ao fato de “o filho ser considerado como da mãe e o adolescente [pai] ser percebido apenas como filho.” (p. 104), reforçando a tradição dos estudos sob esta lógica. O autor propõe uma metodologia de implantação de espaços que acolham, eduquem que conscientizem os pais, sobretudo, os adolescentes, dando visibilidade às figuras: do homem, do pai e do adolescente.

Em relação ao grupo etário sobre o qual se insere este trabalho, - adolescência e juventude – vários estudos sobre fecundidade têm sido realizados (CAMARANO, 1998; HEILBORN, 1999; LOYOLA, 1999; LONGO, 2002; CABRAL, 2005; RODRIGUEZ; 2005; ALMEIDA *et al.*, 2006; BASSI, 2008; FONTOURA e PINERO, 2009; CHACHAM, 2012; CHAGAS, 2012). Entre os estudos, destacam-se aqueles que enfatizam a manutenção das elevadas taxas de fecundidade neste grupo etário por algumas décadas (no que tange à fecundidade feminina). Destacam-se também, aqueles que reforçam que a gravidez nesta fase da vida é vista como um problema não apenas individual, onde, por muito tempo discutiu-se os riscos para a saúde da mãe e da criança, mas também social, uma vez que as mães adolescentes são, em sua maioria, pobres, e que, em contrapartida, buscam com essa trajetória, uma alternativa para alcançar algum *status* socialmente reconhecido (PANTELIDES,

2004). De modo geral as pesquisas assumem, mais uma vez, a ênfase nas mães adolescentes, negligenciando os pais adolescentes.

No que diz respeito à região de estudo, trata-se de uma escolha por conveniência, uma vez que os registros civis de nascimento foram obtidos apenas junto a uma unidade de cartório civil, localizado, portanto, na Regional de Venda Nova, em Belo Horizonte. Segundo dados do Censo demográfico de 2010, a regional abriga uma população de 262.183 moradores, distribuídos entre 36 bairros – o menor número entre as demais regionais, ocupando uma área de 29,27 km² do município (Belo Horizonte, 2018 - <https://prefeitura.pbh.gov.br/venda-nova> - acesso em 10/12/18).

1.1 - Justificativa

Tradicionalmente, como já mencionado, as taxas de fecundidade são calculadas considerando apenas as informações da mãe/mulher, uma vez que estas, podem afirmar, com precisão, sobre a sua parturição. Tal fato se justifica na perspectiva de objetividade e menor complexidade para tal. Desta forma, consolidou-se um paradigma que dificilmente será rompido, por exigir precisão das informações coletadas para a elaboração das taxas de fecundidade.

Do ponto de vista biológico, metodológico, teórico e social, Zhang (2011) compreende que a fecundidade masculina não seja medida apontando as seguintes razões: 1) Quanto ao aspecto biológico, é sabido que as mulheres têm seu período reprodutivo mais facilmente identificável (puberdade, menopausa, duração da gravidez, amamentação), o que as condicionam a uma restrição biológica já conhecida e, portanto, consolidada em várias áreas da ciência; 2) Em relação aos aspectos metodológicos, a não evidência do homem nos estudos de fecundidade pode ser percebida em todas as etapas, a saber: quando na coleta dos dados, aborda-se, na maioria das vezes, as mulheres que estão em casa e respondem as entrevistas. Ademais o modelo clássico de estimativa da fecundidade é o modelo de um só sexo, que segundo Wood *apud* Zhang (2011) o uso deste modelo tornou-se uma questão de convenção. De modo geral, sabe-se que os dados disponíveis sobre fecundidade referem-se à fecundidade feminina; 3) O terceiro aspecto diz respeito à tradição teórica da demografia. Entre as teorias que explicam as mudanças na fecundidade, poucas levam em consideração ou incluem os homens, não considerando as decisões do pai sobre a fecundidade. Além disso, tais teorias

não consideram as mudanças sociais relativas ao cumprimento do papel do homem dentro do domicílio (ZHANG, 2011); 4) Por fim, nas abordagens sociológicas tradicionais, os estudos de família, por exemplo, ainda destacam os papéis estigmatizados do homem e da mulher no domicílio, onde os homens aparecem, principalmente, como mantenedores, para prover bens materiais, e as mulheres, como as responsáveis pelo trabalho doméstico e pelas crianças (WATKINS, 1993).

Falar da fecundidade masculina, portanto, implica em falar de uma série de fatores que envolvam o homem em atividades ou assuntos que, por ora são tratados apenas como assuntos femininos ou na perspectiva das mulheres (MEDRADO e LYRA, 1999). Ou seja, não falar dos pais, dos homens, do masculino no que diz respeito à fecundidade é o equivalente a não os considerar, em um lugar que eles possam ocupar, configurando-se como a constituição de um não lugar na sociedade (LYRA, 1997).

Zhang (2011) se debruçou sobre os estudos de fecundidade e identificou que a maioria dos estudos pesquisados tratava a fecundidade masculina na perspectiva biológica ou médica, discutindo questões como fatores biológicos; comportamentais e naturais que afetam a reprodução masculina ou a saúde reprodutiva; métodos contraceptivos masculinos, e; causas de infertilidade. Alguns destes estudos, realizados no campo da demografia e da sociologia, se fundamentam em uma das principais fontes de dados da área, a *Demographic and Health Surveys* (DHS) e abordam sobre o papel do homem na redução da fecundidade e no planejamento familiar em alguns países da África (LAMPTEY *et al.*, 1978; MAHARAJ, 2001; PETR-NUSTAS, 1999 *apud* ZHANG, 2011).

Tais estudos demonstram que, em alguns países da África, os homens têm grandes conhecimentos em métodos contraceptivos e são muito envolvidos no planejamento familiar, desenvolvendo um papel decisivo na tomada de decisão sobre a fecundidade. Segundo Zhang (2011) “men’s motivations of contraceptive use and their fertility preferences eclipse those of their wives.” (p. 12). O que chama a atenção, já que o desejo masculino é o que prevalece e o que mensuramos é dado pelo desejo feminino, omitido neste caso por questões de imposições culturais, sobre as quais prevalecem a decisão masculina.

Trindade e Menandro (2002), em uma abordagem qualitativa, encontraram que, entre oito jovens entrevistados de classe média e baixa, há relatos que demonstram claramente a falta de conhecimento sobre métodos contraceptivos; falta de compromisso quanto à prevenção de doenças sexualmente transmissíveis; bem como atribuição da responsabilidade sobre a gravidez às parceiras. O estudo também abordou questões sobre a forma como os pais receberam a notícia de gravidez, a qual envolveu reações de susto, medo, revolta ou vergonha, indicando a falta de planejamento para o evento. As autoras investigaram, também, sobre as possíveis mudanças provocadas pelo evento da gravidez/paternidade, sobre as quais os entrevistados apontaram tanto mudanças negativas, quanto positivas. As mudanças negativas referem-se, para os jovens de classe baixa, à mudanças nos planos, pois pretendiam se manter no Exército após o ingresso no serviço militar; já para os de classe média, a mudança nos planos se refere aos estudos, pois pretendiam entrar no ensino superior e ingressarem-se no mercado de trabalho mais tardiamente. Além disso, os jovens citaram, como aspecto negativo, a perda de liberdade após o filho. As mudanças positivas referem-se ao aumento da responsabilidade, ao amadurecimento e melhor relacionamento com o filho. Por fim, um último aspecto abordado pelas autoras diz respeito ao significado da paternidade e maternidade na perspectiva dos pais entrevistados. O sustento alcançado como fruto do trabalho é o primeiro aspecto mencionado pelos entrevistados como significado de paternidade. A presença do pai, o envolvimento na educação, o acompanhamento, e demonstrações de carinho, amor e atenção foram outras questões apresentadas pelos entrevistados como significado de paternidade. Já à maternidade os entrevistados atribuíram, principalmente, o significado de cuidar, dar amor e carinho. Dois dos entrevistados afirmaram ter, o pai e a mãe, o mesmo papel no desenvolvimento dos filhos. Em estudo recente, Franco (2016) abordou, também a partir de análise qualitativa, questões semelhantes a estas e identificou que o significado de paternidade perpassa, por entre outras questões, a responsabilidade, que significa, para os mesmos entrevistados, ser adulto indicando portanto, que os pais entrevistados naquele trabalho, mesmo em idades que correspondem às fases da adolescência e juventude, transitaram para a fase adulta (Figura 1).

idade entre 20 e 59 anos ao sistema de saúde. Para isso, a PNAISH se desenvolve através de cinco eixos temáticos, dos quais, um, corrobora com os interesses deste trabalho. Os eixos são os seguintes:

- Acesso e acolhimento;
- Saúde sexual e saúde reprodutiva;
- Paternidade e cuidado;
- Doenças prevalentes na população masculina, e;
- Prevenção de violências e acidentes.

A motivação conceitual para a elaboração da PNAISH ampara-se na questão de gênero, destacando os lugares ocupados por homens e mulheres na sociedade, no qual não se estimula à prática, pelo homem, do afeto, no sentido da paternidade e do cuidado consigo próprio. Uma vez que a prática do cuidado com a saúde é atribuída e direcionada (como foi por muito tempo, nos serviços públicos de saúde) à mulher, no âmbito das decisões pessoais, a fecundidade foi, por muito tempo, e ainda é, por muitos, atribuída apenas à mulher. Neste sentido, a Política Nacional de Atenção Integral à Saúde do Homem, busca colaborar com a quebra de tal paradigma além de, fundamentalmente, estimular o acesso aos serviços de saúde pelo homem, buscando uma naturalização do cuidado com a saúde por parte deste público. Desta forma, este trabalho procura colaborar com essa perspectiva no sentido de tirar apenas da mulher, a responsabilização pela fecundidade.

Com os estudos apresentados aqui, fica clara a diversidade de situações já abordadas e ainda passíveis de serem abordadas nos estudos sobre pais adolescentes e jovens, bem como sobre o envolvimento destes no processo de fecundidade. Entretanto, é notória a insuficiência de dados para uma compreensão mais ampla do fenômeno da paternidade na adolescência e juventude. Nesse sentido, nos dedicaremos a estimar o número de casos de ocorrência de nascimentos de filhos de pais entre 15 a 24 anos, para um grupo de população que acessa uma unidade de cartório de Belo Horizonte para tais registros.

1.2 - Objetivos

O objetivo deste trabalho é entender a distribuição dos dados que dizem sobre a fecundidade masculina entre homens que tiveram filhos com idade entre 15 a 24

anos entre os anos de 2013 e 2018. A análise se dará através do número de registros de nascimentos, realizados segundo a informação dos pais deste grupo etário, que será analisado como: adolescentes (15 a 19 anos) e jovens (20 a 24 anos). Os dados são provenientes de uma unidade de registro civil do município, que tem como jurisdição os registros de nascimento de dois hospitais do município de Belo Horizonte, mas que, no entanto, recebe demandas para registros de nascimentos ocorridos em outros hospitais, bem como, em outros municípios. Inicialmente foi feita uma análise descritiva dos dados para a compreensão de sua distribuição e, posteriormente, uma análise da distribuição ao longo do tempo, com a proposta de previsão dos registros de nascimento entre o perfil etário de interesse. Os resultados sugerem que, conforme observa-se entre as mulheres, a fecundidade de adolescentes e jovens do sexo masculino, também está se reduzindo.

1.2.1 - Objetivos Específicos

Quanto aos objetivos específicos, a partir dos dados coletados será possível os seguintes êxitos:

- Apresentar a tendência de registros de nascimentos para cada grupo etário (15 a 19 anos e 20 a 24 anos);
- Prever o número de registros que ocorrerão na unidade de cartório sobre a qual os dados foram obtidos.

2 - Metodologia

Para alcançar o objetivo do trabalho, optou-se por tentar obter os dados sobre registros de nascimentos através de uma das unidades de cartório de registro civil do município de Belo Horizonte. Por questões de relacionamento pessoal, a unidade escolhida foi aquela onde a pesquisadora possuía facilidade de acesso por conhecer um de seus representantes legais. Assim, escolheu-se o Cartório de Registro Civil e de Notas de Venda Nova – daqui em diante mencionado como Cartório, localizado em uma, das nove regionais administrativas de Belo Horizonte, a Regional Venda Nova, sobre o qual, conseguiu-se obter os dados, que prontamente foram cedidos pelo representante legal da unidade.

Além dos dados desta unidade de Cartório, foram analisados, também, dados do Sistema de Informações sobre Nascidos Vivos (SINASC) do Departamento de Informática do Sistema de Saúde do Brasil (DataSUS), para os anos de 2013 e a 2017. Os dados do SINASC foram obtidos a partir do portal da instituição (<http://www2.datasus.gov.br/DATASUS/index.php?area=0901&item=1&acao=28>).

Os dados obtidos junto ao Cartório correspondiam aos registros realizados entre o período de 01/01/2013 a 17/07/2018, totalizando 51.060 casos. Importa destacar que nem todo registrado, aqui entendido pelo indivíduo que nasceu, corresponde ao nascimento neste período. Isto se deve ao fato de haver, entre os dados, registros tardios e registros de adoção, indicando, portanto, datas de nascimento anteriores ao período analisado. Correspondem, portanto, a casos em que o registro é realizado para oficializar papéis civis ou regularizar situações civis dos indivíduos que estão passando por tais situações, e, portanto, feitos posteriormente à data de nascimento. Tais registros, que correspondiam a 0,99% do universo analisado foram excluídos. Além disso, consta nesta base de dados, casos sem informações do pai (17,36%), o que não permitiria alcançar os objetivos propostos neste trabalho, e que, portanto, foram também desconsiderados. Optou-se, também, por trabalhar apenas com registros de pessoas nascidas em Minas Gerais, excluindo-se os demais estados¹ que correspondiam a 0,031%. Por fim, preferiu-se trabalhar com o período de cinco anos inteiros, selecionando casos a partir de 01/07/2013 até 30/06/2018, que resultou em um número total de registros de nascimentos de 41.675 casos para a execução deste trabalho.

O trabalho se apropriou da metodologia de séries temporais que se constitui em realizar uma previsão de casos de nascimentos registrados através do registro civil no Cartório de interesse, entre pais que tinham, à época do registro, entre 15 e 24 anos, utilizando como base amostral e de validação o período entre julho de 2013 e junho de 2018.

Uma série temporal corresponde a qualquer conjunto de dados disposto segundo o tempo – seja ele qual for – minuto, hora, dia, semana, mês, ano etc – de forma ordenada. Esse conjunto pode ser contínuo, a exemplo de registros de velocidade

¹ Constam nos registros de nascimentos do Cartório de Venda Nova, no período em análise, os seguintes estados, além de Minas Gerais: Bahia; Espírito Santo; Goiás; Maranhão; Pará; Rio de Janeiro, e; Sergipe.

do vento em um determinado local, ou discreto, como no caso dos dados deste trabalho que correspondem ao número de nascimentos por dia/mês registrados em uma unidade cartorial de Belo Horizonte durante o período de cinco anos, ou seja, uma contagem – de nascimentos.

Uma série temporal pode ser expressa da seguinte forma:

$$Y(t) = [Y_{1(t)}, Y_{2(t)}, \dots, Y_{r(t)}], \quad (1)$$

onde cada componente de Y , denota uma variável e distintas dimensões de tempo, ou seja, r corresponde ao número de variáveis no tempo t .

No caso deste trabalho em que se observa apenas uma variável (registros de nascimentos), a distribuição da série equivale à seguinte expressão:

$$Y(t) = [Y_{1(t)}], \quad (2)$$

onde $Y_{1(t)}$ é o número de registros nascimentos no tempo t .

Dentre os objetivos que a análise de série temporal permite, este trabalho buscou realizar previsões² de valores futuros de registros de nascimentos de filhos de pais entre 15 e 24 anos. A previsão aqui estabelecida é dada por:

$$Y(t+h) = \hat{Y}_t(h), \quad (3)$$

onde $Y(t+h)$ é a variável de interesse originada no tempo t com horizonte de previsão no tempo h .

Além disso, buscou-se apresentar como a série se distribui, observando tendências, ciclos e variações (caso existam). Ao apresentar a distribuição da série, é possível refletir sobre as seguintes perguntas: o número de registros entre o público de interesse na unidade de cartório analisada irá diminuir ou aumentar? Os registros apresentam distribuição constante? O número de registros apresentam sazonalidade? Se sim, qual a regularidade e frequência?

² Segundo Morettin e Tolo (2006), previsão indica “ver uma coisa antes que ela exista.” (p. 7). Além disso, não se trata de uma previsão simplesmente, mas de previsão de mínimos quadrados de um determinado valor futuro baseado em combinações lineares de valores de referência, ou seja, do passado.

Desta forma, um conjunto de dados temporais pode ser estacionário ou não estacionário, e, em ambos os casos, apresentarem alguma tendência (manutenção, queda ou aumento) ou ainda, sazonalidade. A estacionariedade, que indica a manutenção da distribuição dos dados ao longo do tempo, caracteriza-se por ser um dos princípios da análise de série temporal, indicando que os dados se distribuem ao longo do tempo em torno de uma média constante e variância conhecida – configura-se como uma estacionariedade fraca – e que, portanto, apresentam-se distribuídos equilibradamente e de modo estável do ponto de vista estatístico. Já aquelas séries que apresentam tendência de aumento ou de queda, não se distribuem em torno de uma média, mas de uma reta, que pode ser positiva ou negativa. Há ainda os casos em que a não estacionariedade apresenta-se de modo espontâneo, como no caso do surgimento de bactérias. Por fim, as séries que apresentam sazonalidade correspondem àquelas que apresentam padrão de repetição no dado intervalo de tempo, seja, dia; semana; mês; ano. Esta sazonalidade configura-se cíclica, apresentando repetições regulares, portanto, sazonalidades determinísticas, em que pode ser prevista pela observação dos registros anteriores. Há também sazonalidades não regulares, portanto, que não apresentam repetições determinadas, ou sazonalidades não determinísticas.

Caso um conjunto de dados não cumpra com o suposto de estacionariedade, é necessário uma transformação para torná-lo estacionário. Basicamente, há dois procedimentos que podem ser utilizados, a saber: aplicando diferenças sucessivas na série original, buscando obter uma série estacionária em relação à média, e/ou aplicando transformações que estabilizam a variância, no caso de não estacionariedade em relação a variância da série, muito comum em dados econômicos. Nesse caso, algumas transformações não lineares utilizadas são a logarítmica ou a transformação de Box-Cox.

As transformações implicam na relação do instante $t \leftrightarrow t - 1$. As razões para a transformação consistem em: Estabilizar a variância dos dados; tornar o efeito sazonal aditivo; e obter uma distribuição mais simétrica, próxima de uma distribuição Normal. No entanto, vários autores concluíram que a transformação não melhora a qualidade da precisão, gerando, portanto, viés que demanda consideração e tratamento (NELSON, 1976; MAKRIDAKIS e HIBON, 1979; GRANGER e

NWEBOLD, 1976). As diferenças, por sua vez, consistem em estimar os parâmetros postos em:

$$Z_t^{(\lambda)} = \begin{cases} \frac{Z_t^{(\lambda)} - c}{\lambda}, & \text{Se } \lambda \neq 0, \\ \log Y_t, & \text{Se } \lambda = 0, \end{cases}$$

onde se dá a transformação de Box-Cox (1964), na qual os parâmetros λ e c serão estimados.

Conforme Morettin e Tolo (2006, p. 63) a correção da sazonalidade (S_t) em uma série (Y_t) compreende na obtenção das estimativas da sazonalidade (\hat{S}_t) a partir da sazonalidade observada (S_t) e na subtração da série estimada de (Y_t). A correção é dada por:

$$Y_t^{SA} = Z_t - \hat{S}_t. \quad (4)$$

A série com o ajuste de sazonalidade será dada por:

$$Y_t^{SA} = Y_t | \hat{S}_t. \quad (5)$$

Todo e qualquer modelo utilizado para estudar um conjunto de séries temporais configura-se como um processo estocástico, e que, portanto, é controlado pelas leis probabilísticas. Deste modo, tais modelos pertencem à família $Y = \{Y(t), t \in T\}$, onde $Y(t)$ é uma variável aleatória para cada $t \in T$. A Figura 2 apresenta a exemplificação de um processo estocástico em perspectiva histórica:

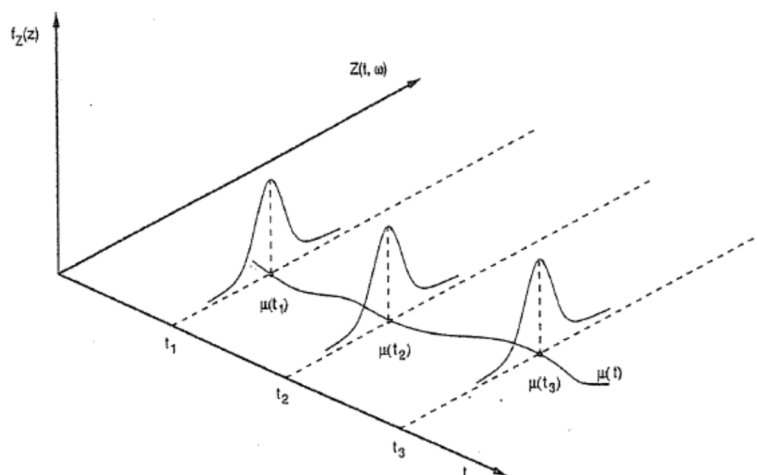


Figura 2 - Representação de um processo estocástico.
Fonte: Morettin e Tolo, 2006.

O processo estocástico configura-se como uma família de variáveis aleatórias, inseridas, supostamente, em um mesmo espaço de probabilidade, ou seja, onde a função densidade de probabilidade é a mesma para todo $t \in T$. Logo, se a função de probabilidade não varia, pode-se dizer que o processo é estritamente estacionário.

Quando se pretende analisar um conjunto de dados através dos modelos de séries temporais, é necessário estabelecer algumas suposições, para além daquela já mencionada anteriormente sobre estacionariedade. São elas:

- Estacionariedade ou não estacionariedade, conforme a independência, ou não, relativamente à origem dos dados;
- Normalidade (Gaussiano) ou não normalidade, conforme a função distribuição de probabilidade (f.d.p.) que dá característica ao processo;
- Markovianos ou não Markovianos, conforme a independência dos valores do processo em um determinado instante, considerando os valores em instantes precedentes.

Um processo estacionário se caracteriza por ter todas as distribuições permanentes sob as variações do tempo, onde a média $\mu(t)$ e a variância $V(t)$ são constantes equivalendo-se à:

$$\mu(t) = \mu \text{ e } V(t) = \sigma^2. \quad (6)$$

Além disso, em um processo estacionário, supomos que a média seja constante. Conforme Morettin e Tolo (2006, p. 24), em um processo estocástico, interessa-nos caracterizar apenas um pequeno “número de distribuições finito-dimensionais, ou de momentos, restringindo-nos aos momentos de primeira e segunda ordens”. Desta forma, tem-se que um processo estocástico é fracamente estacionário ou estacionário de segunda ordem se atender as seguintes condições:

1. $E\{Y(t)\} = \mu(t) = \mu$, constante, para todo $t \in T$;
2. $E\{Y^2(t)\} < \infty$, para todo $t \in T$, sendo este considerado de segunda ordem;
3. $\tau(t_1, t_2) = Cov\{Y(t_1), Y(t_2)\}$, é uma função de $|t_1 - t_2|$ (MORETTIN e TOLOI, 2006, p. 24).

Um processo estocástico se diz Gaussiano se, para qualquer conjunto de dados em qualquer tempo (t_1, t_2, \dots, t_n) , as variáveis aleatórias $(Y_{1(t)}, Y_{2(t)}, \dots, Y_{r(t)})$ tem distribuição Normal n-variada. Logo, se ele for Gaussiano ou Normal, será determinado pelas médias e covariâncias (MORETTIN e TOLOI, 2006, p. 24).

A função de autocorrelação (FAC) mede a dependência linear entre os valores da série no instante t e $t+j$, ou seja, em diferentes períodos. Os valores da FAC resutam do quanto um valor é influenciado pelo valor já observado. Desta forma, como meio de avaliação, retorna o correlograma, que apresenta os coeficientes de autocorrelação em função do atraso desta dependência. Uma vez verificada as dependências entre os valores da série e removidas, obtemos a autocorrelação condicional, ou parcial (FACP) que avalia os resíduos da série.

Após a transformação dos dados buscando torná-la estacionária, observa-se os correlogramas da FAC e da FACP para a identificação dos modelos que podem ser aplicados e que podem acompanhar a variabilidade dos dados. Procede-se com a estimação e verificação dos modelos que exigem: 1 – a verificação dos coeficientes estimados que devem ser estatisticamente diferentes de zero; 2 – a observação de que os mesmos cumpram as propriedades de estacionariedade e de invertibilidade; 3 – a verificação do vetor de resíduos que deve se configurar como um processo de ruído branco. Tal constatação se dá pela verificação das autocorrelações da FAC e da FACP do vetor de resíduos que devem ser não significativas. Além disso, é necessário verificar, também, se a distribuição de probabilidade dos resíduos se aproxima a uma distribuição Normal.

2.1 - Método de Decomposição

A decomposição das séries temporais começou a ser estudada nos finais do século passado e desenvolveu-se nos princípios deste século. Este método consiste numa descrição matemática dos movimentos presentes nas séries, os quais após serem identificados, podem ser utilizados para facilitar a interpretação e a projeção da série temporal. Os movimentos, ou componentes, que se distinguem nas séries temporais são usualmente classificados em quatro tipos: tendência, ciclo, sazonalidade e termo aleatório.

A tendência pode ser entendida como o movimento persistente dos dados em uma dada direção. O ciclo, por sua vez, indica o movimento oscilatório em torno da tendência, podendo este ser verificado por períodos prolongados. A sazonalidade indica a distribuição regular assumido pela série em algum subperíodo. Já o termo aleatório leva em consideração movimentos esporádicos e irregulares presentes na série. A combinação destes componentes pode ser feita mediante a forma aditiva ou multiplicativa.

Cada observação, Y_t , é vista como resultante da combinação dos quatro elementos (tendência, ciclo, sazonalidade e termo aleatório) expressando-se na forma genérica:

$$Y_t = f(T_t, C_t, S_t, a_t) \quad t = 1, 2, 3, \dots, n. \quad (7)$$

Se admitimos um modelo multiplicativo, em que as diferentes componentes são interdependentes, temos:

$$Y_t = T_t \times C_t \times S_t \times a_t \quad t = 1, 2, 3, \dots, n. \quad (8)$$

Um modelo aditivo, em que se admite uma independência da distribuição, temos:

$$Y_t = T_t + C_t + S_t + a_t \quad t = 1, 2, 3, \dots, n. \quad (9)$$

Para um modelo, log-aditivo, temos:

$$\text{Log}Y_t = \text{Log}T_t + \text{Log}C_t + \text{Log}S_t + \text{Log}a_t \quad t = 1, 2, 3, \dots, n. \quad (10)$$

Assim, para uma distribuição onde a tendência e a sazonalidade não estejam presentes, é dado por:

$$Y_t = \mu + \alpha_t. \quad (11)$$

Já uma distribuição com a tendência, em que a componente sazonal não esteja presente é representada por:

$$Y_t = \beta_0 + \beta_1 t + a_t, \quad (12)$$

onde o a_t representa o ruído branco no instante de tempo t .

2.2 - Modelo ARIMA

O modelo ARIMA corresponde à metodologia proposta por George Box e Gwilym Jenkins³ para a análise de modelos não paramétricos, consolidada na década de 1970. A proposta considera a possibilidade de trabalhar com processos autorregressivos integrados e de médias móveis, gerando um modelo **Autorregressivo Integrado de Médias Móveis de ordem (p,d,q) - ARIMA(p,d,q)**. Onde o componente AR(p) indica um processo auto-regressivo de ordem p ; o componente MA(q) indica um processo de médias móveis de ordem q . Em se tratando de “processos lineares não estacionários homogêneos” que se “constituem uma generalização dos processos lineares estacionários que supõem que o mecanismo gerador da série produz erros auto-correlacionados tornando-a não estacionária em nível e/ou em inclinação” e que portanto demandam aplicações de diferenças (MORETTIN e TOLOI, 2006, p. 36), tem-se a componente I(d), que indica um modelo auto-regressivo integrado e de médias móveis de ordem d .

A definição do modelo é dada por processos iterativos em que se configuram os seguintes estágios: especificação; identificação, baseada nas análises de autocorrelação e de autocorrelação parcial; estimação dos parâmetros; verificação do modelo ajustado através da análise dos resíduos (MORETTIN e TOLOI, 2006).

Consideremos então uma série Y_t não estacionária, onde após a realização de “ d ” diferenciações não sazonais, esta série se torna estacionária, apresentando componente autorregressivo de ordem p e componente média móvel de ordem q resultando no modelo ARIMA (p, d, q).

A equação do modelo usando operador de retardo é representada como segue:

$$\Phi_p(B)(1-B)^d Y_t = \Theta_q(B)a_t, \quad (13)$$

onde $\Phi_p(B) = (1 - \phi_1 B - \phi_2 B^2 - \dots - \phi_p B^p)$ é o polinômio autorregressivo de ordem p ; $\Theta_q(B) = (1 - \theta_1 B - \theta_2 B^2 - \dots - \theta_q B^q)$ é o polinômio de médias móveis

³ Para mais detalhes ver BOX, G. E. P., JENKINS, G. M. Time Series Analysis: Forecasting and Control. Holden-Day, 1970; BOX, G. E. P., JENKINS, G. M. e REINSEL, G. Time Series Analysis: Forecasting and Control. Englewood Cliffs: Prentice Hall, 1994.

de ordem q com raízes fora do círculo unitário e sem raízes comuns; e $\{a_t\}$ é um processo de ruído branco com média zero $E(a_t)=0$ e variância constante $\text{Var}(a_t)=\sigma^2$ (MORETTIN e TOLOI, 2006).

2.3 - Alisamento Exponencial

O método alisamento exponencial pondera os valores utilizados na previsão de acordo com o tempo, eliminando possíveis ruídos que possam haver na série observada. O que significa dizer que o alisamento exponencial permite atribuir um maior peso em valores mais recentes no tempo, atribuindo a estes maior importância na determinação dos valores previstos (WANKE *et al.*, 2006). Essa determinação é dada pelos valores de constantes que serão distintos conforme a distribuição da série (estacionária, com tendência, sazonal etc.) (MORETTIN e TOLOI, 2006).

Este método é conhecido com esta denominação em virtude de aplicar um conjunto de pesos desiguais aos valores passados da série, sendo que tais pesos decaem de forma exponencial da mais recente até a mais distante observação (MORETTIN e TOLOI, 2006).

Os modelos de alisamento exponencial são originários da década de 40, com os estudos realizados por Goodbell Brown. Na década seguinte, Charles Carter Holt, independentemente de Brown, desenvolveu métodos diferentes de alisamento exponencial, publicando seu trabalho em 1957. No início da década de 1960, Peter Winters testou os modelos de Holt utilizando dados empíricos, tornando o modelo conhecido como Holt-Winters. As equações de alisamento são dadas a seguir:

$$L_t = \alpha Y_t + (1-\alpha)[L_{t-1} + T_{t-1}], \quad (14)$$

$$T_t = \gamma [L_t - L_{t-1}] + (1-\gamma)T_{t-1}, \quad (15)$$

onde α é a constante de alisamento do nível da série, e γ a constante de alisamento da tendência da série; Y_t é o nível da série; T_t é a tendência da série.

As previsões h passos à frente da série Y_t , com base nas observações até o instante de tempo t é dada por:

$$\hat{Y}_t = Y_{(t-1)} + h\hat{T}_{(t-1)}. \quad (16)$$

O modelo de suavização sazonal de Holt-Winter, que considera a componente sazonal apresentada na série, com período s , pode ser aplicado considerando o fator sazonal como aditivo. A seguir as equações de alisamento:

$$L_t = \alpha(Y_t - S_{t-s}) + (1 - \alpha)(L_t - T_{t-1}); \quad (17)$$

$$T_t = \gamma(L_t - L_{t-1}) + (1 - \gamma)T_{t-1}; \quad (18)$$

$$S_t = \delta(Y_t - L_t) + (1 - \delta)S_{t-s}. \quad (19)$$

Respectivamente as equações acima apresentam as formas de estimação do nível, da tendência e do fator sazonal, onde a constante de alisamento do nível da série é α ; a constante de alisamento da tendência é γ , e; a constante de alisamento do fator sazonal é δ . As previsões h passos à frente da série Y_t , com base nas observações até o instante de tempo t é dada por:

$$\hat{Y}_t = L_{t-1} + T_{t-1} + S_{t-s}. \quad (20)$$

A escolha das constantes de alisamento envolve a minimização da soma de erros de predição ao quadrado.

2.4 Medidas de Acurácia

A definição do melhor modelo de previsão será dada pela comparação das estatísticas dos erros de previsão. A previsão de h registros de nascimentos à frente com origem no instante t é representada por:

$$Y_t(h), h \geq 1. \quad (21)$$

Portanto, a avaliação do modelo se dá através do erro de previsão verificado para h registros de nascimentos à frente, que é dada por:

$$e_{T+h} = Y_T(h) - Y_{T+h}, h \geq 1, \quad (22)$$

onde e_{T+h} é o erro de previsão; e T é o dado no tempo que origina a previsão.

O melhor modelo será aquele que apresentar o menor valor em alguma das estatísticas que avaliam os erros de previsão, quais sejam:

Erro Percentual Médio (MPE):

$$\text{MPE} = \frac{1}{m} \sum_{t=1}^m \left(\frac{y_t - \hat{y}_t}{y_t} \right) * 100, \hat{y}_t \neq 0. \quad (23)$$

Erro Percentual Absoluto Médio (MAPE):

$$\text{MAPE} = \frac{1}{m} \sum_{t=1}^m \left| \frac{y_t - \hat{y}_t}{y_t} \right| * 100, \hat{y}_t \neq 0. \quad (24)$$

Erro Absoluto Médio ou Desvio Absoluto Médio (MAD):

$$\text{MAD} = \frac{1}{m} \sum_{t=1}^m |y_t - \hat{y}_t|. \quad (25)$$

Erro Quadrático Médio ou Desvio Quadrático Médio (MSE):

$$\text{MSE} = \frac{1}{m} \sum_{t=1}^m (y_t - \hat{y}_t)^2. \quad (26)$$

3 - Análise dos Dados

A análise dos dados será baseada no universo de 41.675 registros de nascimentos e será exposta a seguir.

3.1 - Descrição dos Registros

A proporção dos registros segundo sexo corresponde ao que se observa nos dados do SINASC, que nos diz sobre a fecundidade feminina. Entre o mesmo período, os dados do SINASC apontaram que 51,2% dos registros de nascimentos do estado de Minas Gerais correspondiam ao sexo masculino. Os dados obtidos junto ao Cartório (Figura 3) também apontaram maioria dos nascimentos, no mesmo período, para o sexo masculino (51%). Ao avaliar a razão de sexo, verificamos que ambos os grupos etários apresentam uma razão entre o percentual de nascimentos do sexo masculino e feminino acima de 1, indicando maior população do sexo masculino. Entre os registros do Cartório, a razão entre os percentuais de registros do sexo masculino e

feminino indica uma razão igual a 1,050, supondo que, pelo menos para esta região de Belo Horizonte, há uma possibilidade de reprodução da razão entre nascimentos em relação ao sexo.

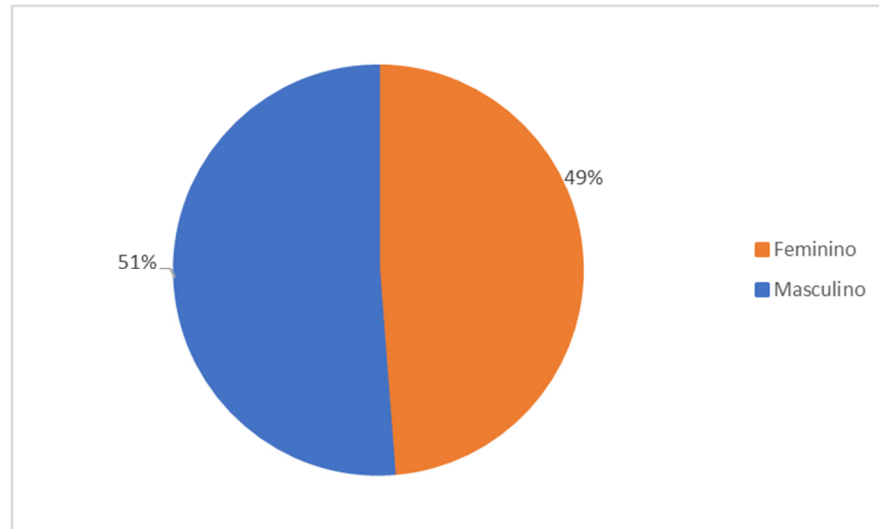


Figura 3 - Distribuição dos registros de nascimentos segundo sexo, 2013-2018
Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

No que tange à idade do pai, a maior concentração de registros é observada entre as idades de 25 a 29 anos, com 24,14% dos registros, seguido pelo grupo de 30 a 34 anos, com 23,35% e pelo grupo de 20 a 24 anos, que representa quase um quinto dos registros (19,06%). Nos casos de idade abaixo de 16 anos, é necessário, em algumas situações, o testamento jurídico (Figura 4 e Tabela 1).

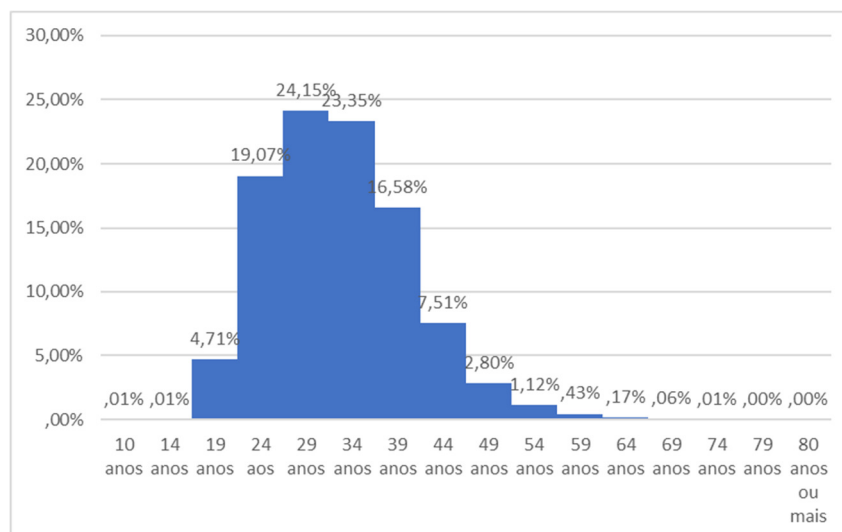


Figura 4 - Distribuição dos registros de nascimentos segundo idade do pai, 2013-2018
Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Tabela 1 - Distribuição dos registros de nascimentos segundo grupo etário do pai, 2013-2018

Grupo Etário	Número de Registros	%
De 10 a 14 anos	7	0,017
De 15 a 19 anos	1962	4,708
De 20 a 24 anos	7948	19,071
De 25 a 29 anos	10066	24,154
De 30 a 34 anos	9734	23,357
De 35 a 39 anos	6910	16,581
De 40 a 44 anos	3129	7,508
De 45 a 49 anos	1168	2,803
De 50 a 54 anos	468	1,123
De 55 a 59 anos	181	0,434
De 60 a 64 anos	71	0,170
De 65 a 69 anos	24	0,058
De 70 a 82 anos	7	0,017
Total	41675	100,00

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Quanto ao local de nascimento, os registros apontam ocorrências em hospitais, majoritariamente (99,62%). Além disso, registros de nascimentos em domicílio (0,24%); em trânsito (0,07%); em via pública (0,01%); em Unidade de Pronto Atendimento (0,002%) e em Centro de Saúde (0,005%) surgem, porém com menos frequência. Em apenas três casos de registro de nascimento, o local de ocorrência do mesmo não foi informado.

Ao avaliarmos o local de nascimento por idade do pai observamos que os registros de nascimentos ocorridos em hospitais (99,66%) ocorrem com mais frequência entre os pais com idade entre 25 a 29 anos, com 24,15% de casos. Em segundo lugar, o maior número de registros ocorre entre aqueles pais com idade entre 30 a 34 anos (23,35%), seguido pelos pais com idade entre 20 a 24 anos (19,08%). Os registros de nascimentos ocorridos nos demais locais (em domicílio; em trânsito; em via pública; em Unidade de Pronto Atendimento – UPA, e; em Centro de Saúde) correspondem a apenas 0,34% dos dados (Tabela 2).

Tabela 2 - Distribuição dos registros de nascimentos segundo local de nascimento e grupo etário do pai, 2013-2018

Grupo Etário	Hospitais	%	Em outros locais	%	Total	%
De 11 a 14 anos	7	0,017	0	0,000	7	0,017
De 15 a 19 anos	1960	4,719	2	1,399	1962	4,708
De 20 a 24 anos	7925	19,082	23	16,084	7948	19,071
De 25 a 29 anos	10032	24,155	34	23,776	10066	24,154
De 30 a 34 anos	9701	23,358	33	23,077	9734	23,357
De 35 a 39 anos	6882	16,570	28	19,580	6910	16,581
De 40 a 44 anos	3115	7,500	14	9,790	3129	7,508
De 45 a 49 anos	1164	2,803	4	2,797	1168	2,803
De 50 a 54 anos	464	1,117	4	2,797	468	1,123
De 55 a 59 anos	181	0,436	0	0,000	181	0,434
De 60 a 64 anos	71	0,171	0	0,000	71	0,170
De 65 a 69 anos	23	0,055	1	0,699	24	0,058
De 70 a 74 anos	7	0,017	0	0,000	4	0,017
Total	45532	99,656	143	0,343	41675	100%

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Quanto à distribuição dos registros por ano de nascimento, os dados apontam que o ano em que houve maior número de registros, considerando aqueles anos em que foi obtido os registros do ano integral, foi o ano de 2014, com 23,4% dos casos, seguido pelo ano de 2015 e 2016. O ano de 2013 apresenta apenas os registros do segundo semestre, com 10,67% de casos, enquanto o ano de 2018, do primeiro semestre, com 8,5% de registro. Acerca dos valores absolutos, é possível perceber uma tendência de queda dos registros na Figura 5, abaixo.

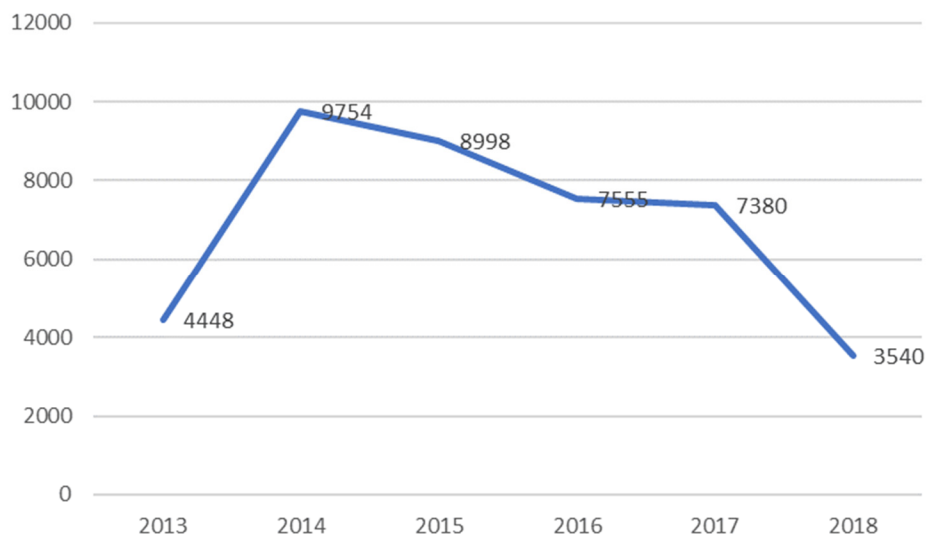


Figura 5 - Distribuição dos registros de nascimentos segundo ano, 2014-2017

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

O padrão de distribuição dos dados analisados se assemelha àquele observado entre as mulheres, com a barriga apontando a ocorrência de maior número de casos

entre as idades mais jovens. Observamos também que nos anos de 2013 até 2016 o padrão de distribuição geral, permanece. Ou seja, o maior número de casos nestes anos ocorreu entre a população de pais com idade entre 25 a 29 anos de idade, seguido pelos pais com idade entre 30 a 34 e 20 a 24 anos. No entanto, para os anos de 2017 e 2018 a distribuição é diferente: nestes anos, observa-se uma tendência ao adiamento da fecundidade por parte dos pais, uma vez que o maior número de registros ocorre no grupo etário de pais com idade entre 30 a 34 anos. Em linhas gerais, nota-se uma tendência de redução para todos os grupos com destaque para aqueles de 20 a 39 anos. Embora haja uma mudança nos picos, observado pela Figura 6, em termos percentuais, a proporção de registros por ano apresenta valores próximos.

Por fim, a figura que melhor representa e pode justificar alguma tendência dos registros corresponde à Figura 6, em que se observa a redução dos registros ano a ano.

A Figura 6 bem como a Tabela 3, abaixo, apresentam a distribuição absoluta dos registros por ano de nascimento, segundo grupo etário do pai. A Figura 7 apresenta a distribuição percentual dos registros.

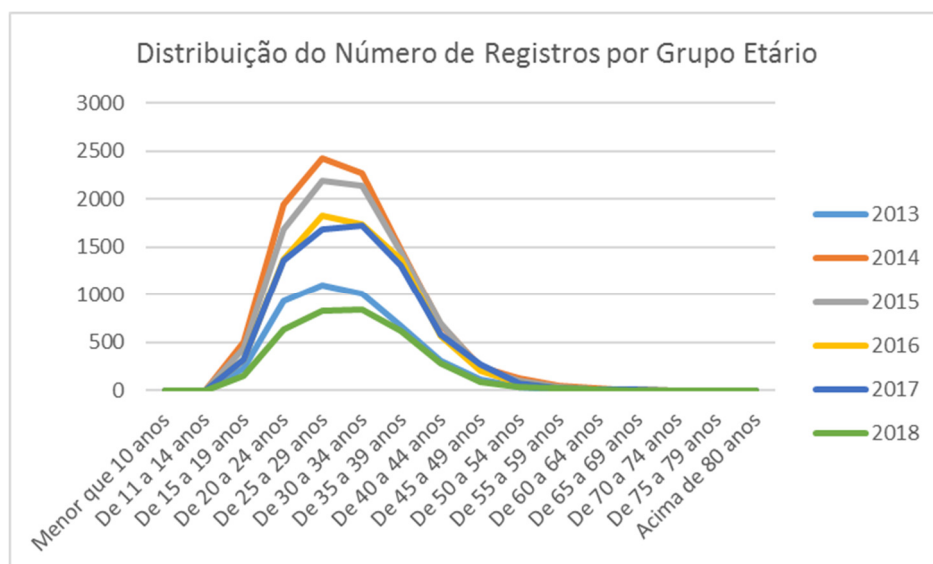


Figura 6 - Distribuição dos registros de nascimentos segundo ano, por grupo etário do pai, 2013-2018
 Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

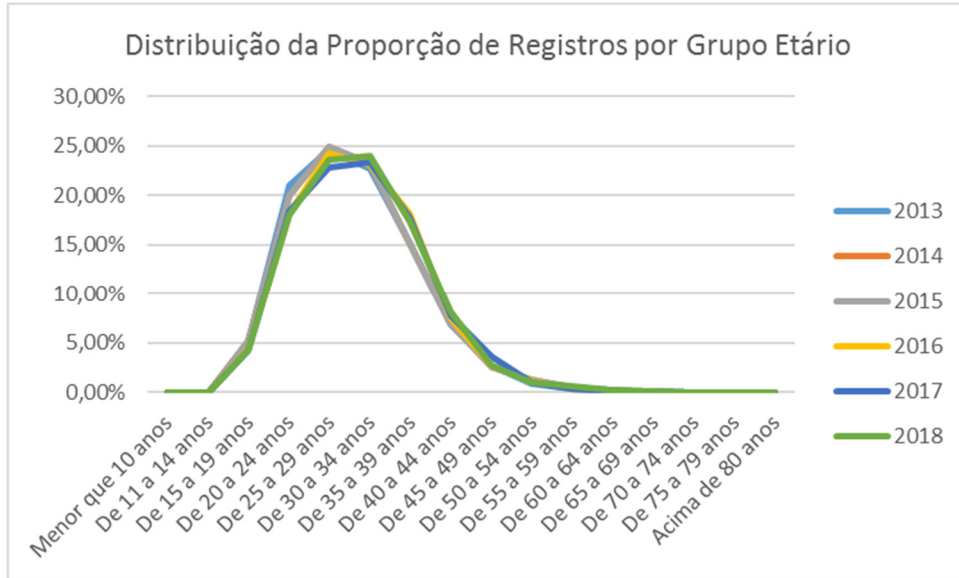


Figura 7 - Distribuição percentual dos registros de nascimento segundo ano, por grupo etário do pai, 2013-2018
 Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Tabela 3 - Distribuição de registros por ano de nascimento, segundo grupo etário do pai.

Estado/Ano	2013		2014		2015		2016		2017		2018		Totais	
	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%	Freq.	%
De 10 a 14 anos	2	0,000	2	0,000	0	0,000	2	0,000	1	0,000	0	0,000	7	0,000
De 15 a 19 anos	231	0,052	500	0,051	437	0,049	328	0,043	317	0,043	149	0,020	1962	0,047
De 20 a 24 anos	936	0,210	1951	0,200	1687	0,187	1374	0,182	1361	0,184	639	0,087	7948	0,191
De 25 a 29 anos	1098	0,247	2425	0,249	2197	0,244	1830	0,242	1681	0,228	835	0,113	10066	0,242
De 30 a 34 anos	1008	0,227	2264	0,232	2144	0,238	1743	0,231	1727	0,234	848	0,115	9734	0,234
De 35 a 39 anos	674	0,152	1483	0,152	1452	0,161	1370	0,181	1315	0,178	616	0,083	6910	0,166
De 40 a 44 anos	308	0,069	680	0,070	703	0,078	572	0,076	577	0,078	289	0,039	3129	0,075
De 45 a 49 anos	121	0,027	248	0,025	227	0,025	207	0,027	271	0,037	94	0,013	1168	0,028
De 50 a 54 anos	41	0,009	130	0,013	99	0,011	83	0,011	80	0,011	35	0,005	468	0,011
De 55 a 59 anos	14	0,003	47	0,005	37	0,004	29	0,004	30	0,004	24	0,003	181	0,004
De 60 a 64 anos	10	0,002	21	0,002	11	0,001	9	0,001	13	0,002	7	0,001	71	0,002
De 65 a 69 anos	4	0,001	2	0,000	2	0,000	6	0,001	6	0,001	4	0,001	24	0,001
De 70 a 82 anos	1	0,000	1	0,000	2	0,000	2	0,000	1	0,000	0	0,000	7	0,000
Totais	4448	0,107	9754	0,234	8998	0,216	7555	0,181	7380	0,177	3540	0,480	41675	1,000

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Conforme observado na Tabela 4 o volume total dos dados registram aumento expressivo no ano de 2014, em relação ao total de registros realizados no ano de 2013, justificado pela redução do período comparativo, uma vez que os dados de 2013 correspondem apenas aos registros do segundo semestre. No ano de 2018 observamos queda justificada pelo mesmo motivo. Ao observarmos os dados entre os anos de 2014 e 2017, observamos queda em todos os anos tendo sido a maior observada em 2016, representando 16,04% a menos no número de registros em relação ao ano anterior. A distribuição dos registros por mês revelam que, para os anos de 2014 a 2017, os meses que ocorrem os maiores números de registros de nascimentos são do primeiro semestre. A distribuição anual de registros foi, em média, 6946 casos. O máximo de registros de nascimentos foi de 9754, no ano de 2014, e o mínimo, de 3540, em 2018 (Tabela 4).

Tabela 4 - Distribuição de registros por ano e mês de nascimento e variação percentual

Mês/ano	2013	2014	2015	2016	2017	2018
Janeiro	-	931	800	696	606	597
Fevereiro	-	848	763	605	610	546
Março	-	845	850	712	748	616
Abril	-	839	817	699	652	656
Maio	-	830	861	709	685	568
Junho	-	836	773	659	678	557
Julho	611	794	682	633	583	-
Agosto	688	780	678	584	584	-
Setembro	764	763	719	633	575	-
Outubro	810	778	674	582	527	-
Novembro	790	749	683	512	586	-
Dezembro	785	761	698	531	546	-
Totais	4448	9754	8998	7555	7380	3540
Variação em relação ao ano anterior (%)	-	-	-7,75%	-16,04%	-2,32%	-

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Ao observarmos o período de julho a junho, pode-se perceber claramente a redução no número de registros, com destaque para o período entre julho de 2015 e junho de 2016, que apresentou redução de 13,44% em relação ao mesmo período anterior. Em contrapartida, o período que houve menor redução foi o período entre julho de 2014 a junho de 2015, que apresentou uma redução de apenas 0,92% em relação ao período anterior, ou seja, entre julho de 2013 e junho de 2014 (Figura 8).

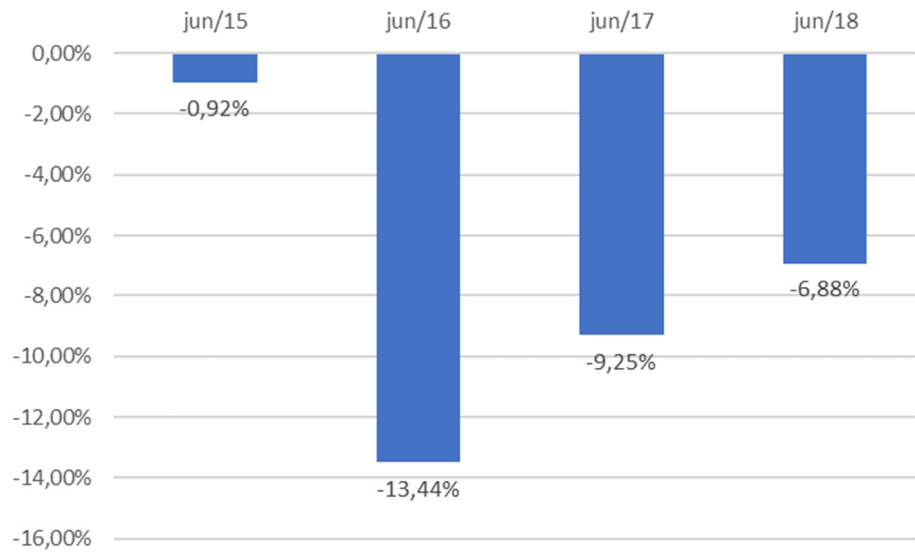


Figura 8 - Variação percentual entre os anos, considerando o início do período, jul/2013.
Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

No Quadro 1 e na Figura 9 apresenta-se os resultados da análise descritiva do número de registros ao longo dos anos no período de 2013 a 2018. No período entre 2014 e 2017 se observa uma redução no número médio de registros. A variabilidade em torno do valor médio mensal é maior nos anos 2013, 2016 e 2017 sendo que o coeficiente de variação chega a 10% (Quadro 1). No ano de 2014 se deu o maior número de registros mensal com a menor variabilidade (CV=6,39%). Entre os anos de 2015 e 2017 o número de registros mensal apresenta queda com maior variabilidade (Figura 9). Não há evidências de casos atípicos em nenhum dos anos analisados.

Quadro 1 - Estatística descritiva dos registros anuais

Ano	Média	DP	CV%	Mínimo	Q1	Mediana	Q3	Máximo
2013	741,3	76,6	10,33	611	668,8	774,5	795,0	810
2014	812,8	51,9	6,39	749	766,8	812,0	843,5	931
2015	749,8	69,9	9,32	674	682,3	741,0	812,8	861
2016	629,6	68,6	10,90	512	582,5	633,0	698,3	712
2017	615,0	63,9	10,39	527	577,0	596,0	671,5	748
2018	590,0	41,4	7,02	546	554,3	582,5	626,0	656

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.
Nota: para os anos de 2013 e 2018, tem-se apenas os registros de 6 meses.

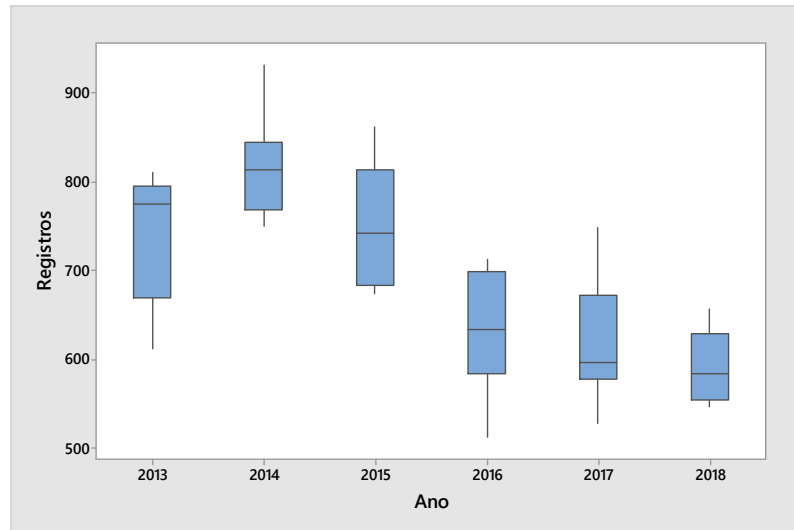


Figura 9 - Diagrama de caixa segundo ano de nascimento

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Quanto a distribuição mês a mês vemos que o mês como maior número de registros é o mês de janeiro, com 931 registros. Já o mês como menor número de registros é o mês novembro, com 512 registros (Quadro 2). Na Figura 10, observamos que não há valores discrepantes ao longo dos meses. Os meses de fevereiro, abril e julho são os que apresentam maior assimetria positiva, indicando maior concentração de registros no início do mês. Em contrapartida, o mês de setembro, apresenta-se com assimetria negativa. O mês com maior variabilidade dos dados é janeiro, seguido dos meses de fevereiro e outubro (Figura 10). A distribuição do número de registros mensal (Figura 11) apresenta uma tendência de queda, indicando não estacionariedade da série mensal. Considerando o valor médio de 694,8 registros/mês, não é possível perceber a distribuição de valores em torno deste valor, indicando a necessidade de aplicar uma correção no caso de ajuste de modelos de séries temporais.

Quadro 2 - Estatística descritiva dos registros mensais

Mês	Média	DP	CV%	Mínimo	Q1	Mediana	Q3	Máximo
Janeiro	726,0	140,9	19,4	597	601,5	696,0	865,5	931
Fevereiro	674,4	125,9	18,7	546	575,5	610,0	805,5	848
Março	754,2	97,9	13,0	616	664,0	748,0	847,5	850
Abril	732,6	89,4	12,2	652	654,0	699,0	828,0	839
Mai	730,6	118,2	16,2	568	626,5	709,0	845,5	861
Junho	700,6	107,7	15,4	557	608,0	678,0	804,5	836
Julho	660,6	82,9	12,5	583	597,0	633,0	738,0	794
Agosto	662,8	82,9	12,5	584	584,0	678,0	734,0	780
Setembro	690,8	83,8	12,1	575	604,0	719,0	763,5	764
Outubro	674,2	121,8	18,1	527	554,5	674,0	794,0	810
Novembro	664,0	114,7	17,3	512	549,0	683,0	769,5	790
Dezembro	664,2	119,2	17,9	531	538,5	698,0	773,0	785

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

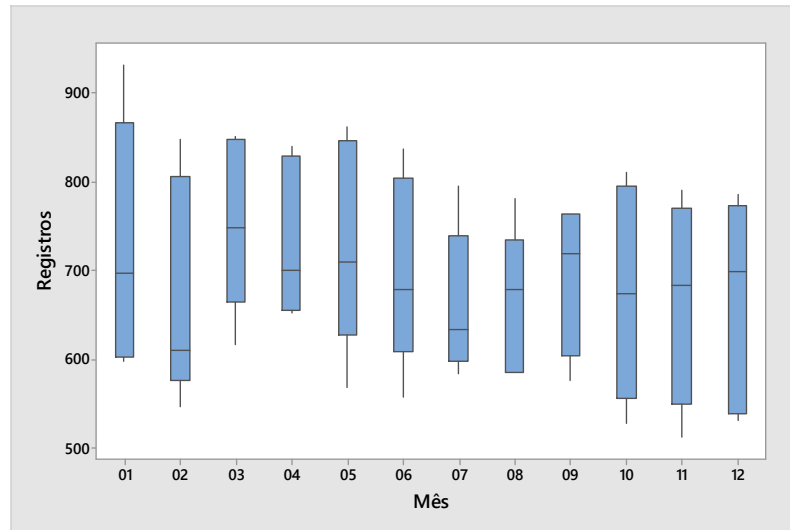


Figura 10 - Diagrama de caixa segundo mês de nascimento

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

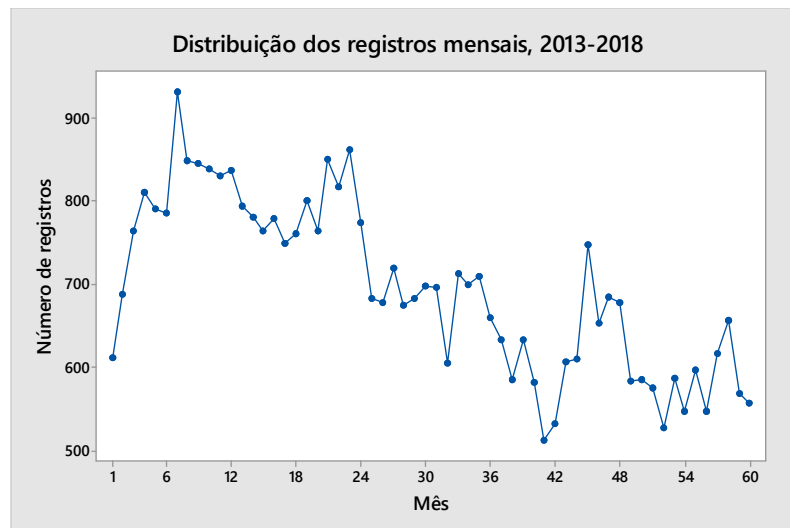


Figura 11 - Distribuição dos registros mensais em perspectiva histórica.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

3.1.1 - Descrição da variável idade dos pais

No que tange à variável de interesse (idade) a média de idade dos pais que registram filhos na unidade de cartório em análise é de 30 anos com desvio padrão em torno da média de 7,7 anos. A idade mínima é de 10 anos e a máxima de 82 anos (Quadro 3). 25% dos pais que registram seus filhos nesta unidade de cartório tem abaixo de 25 anos, enquanto que 25% tem idade acima de 35 anos. Ao comparar os registros do cartório com os dados do SINASC, registra-se uma diferença na idade média de 8 anos a menos, ou seja, a idade média das mães está entre 27 anos, com idade mínima tendo ocorrência aos 9 anos, e máxima, aos 99 anos, com registros de 1 e 3 casos, respectivamente entre os dados do SINASC.

Um quarto das mães tem idade abaixo de 22 anos, um valor inferior ao dos homens. Além disso, 25% das mães têm idade acima de 32 anos indicando portanto, que as mulheres têm filhos mais jovens que os homens, entre os dados analisados. Acerca das idades mínimas e máximas é possível supor, que podem se tratar, no caso de idades máximas, de adoção. No caso de idades mínimas, possível erro de registro.

Quadro 3 - Estatística descritiva da idade dos pais

N	Média	SE Média	DP	Mínimo	Q1	Mediana	Q3	Máximo
41675	30,582	0,0378	7,709	10,000	25,000	30,00	35,000	82,000

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

A distribuição da idade apresenta-se com assimetria positiva, com cauda longa à direita.

A Figura 12 apresenta a distribuição dos registros segundo a idade do pai e no Anexo 1 encontra-se a frequência de registros por idade do pai.

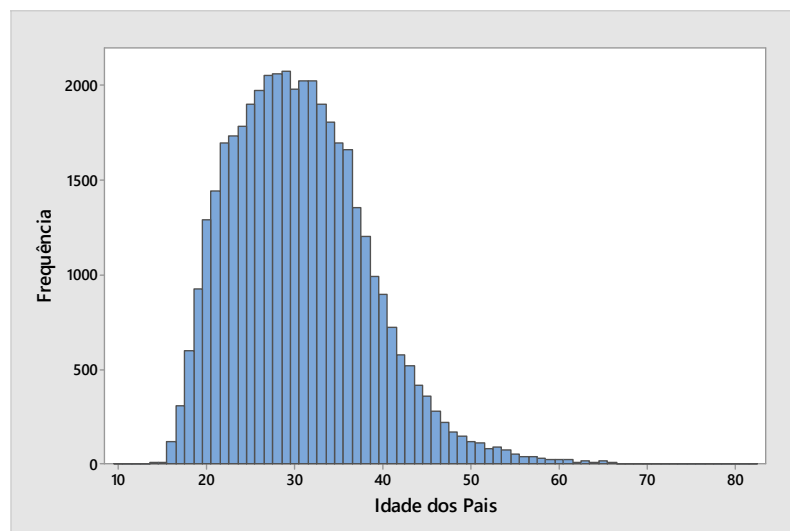


Figura 12 - Histograma da idade do pai para o total de registros (n: 41675)

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Quanto à distribuição da idade ao longo dos anos, observamos que a média de idade dos pais que registram os filhos na unidade de cartório em análise se mantém em torno dos 30 anos. O ano de 2016 apresentou a maior variabilidade de idades, com idade mínima de 10 anos e máxima, de 82. O ano em que apresentou a menor média de idade foi em 2013, e o ano em que se observou a maior, foi em 2017 (Quadro 4). O Anexo 2 apresenta os histogramas da distribuição das idades por ano.

Quadro 4 - Estatística descritiva da idade dos pais - por ano de registro

Ano	N	Média	DP	Mínimo	Q1	Mediana	Q3	Máximo	Assimetria	Curtose
2013	4448	30,054	7,670	10	24,000	29,000	35,000	71,000	0,78	1,03
2014	9754	30,275	7,765	10	24,000	29,000	35,000	82,000	0,77	0,92
2015	8998	30,503	7,591	15	25,000	30,000	35,000	73,000	0,65	0,66
2016	7555	30,796	7,646	14	25,000	30,000	36,000	82,000	0,65	0,86
2017	7380	31,009	7,805	10	25,000	30,000	36,000	71,000	0,62	0,57
2018	3540	30,945	7,753	16	25,000	30,000	36,000	67,000	0,70	0,91

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

3.2 – Descrição dos Grupos Etários de 15 a 19 anos e de 20 a 24 anos

Entre o grupo de interesse deste trabalho, os registros de nascimento no período em análise totalizam 9910 casos, correspondendo a 23,77% do total de 41675 registros considerando-se todos os grupos etários. A distribuição de registros segundo sexo do registrado corresponde ao total de 9910 nascimentos e mostra que: entre o grupo de 15 a 19 anos, os registros se distribuem igualmente entre sexo (Figura 13). Já entre o grupo de 20 a 24 anos, a maioria dos registros (51,84%) é do sexo masculino (Figura 14).

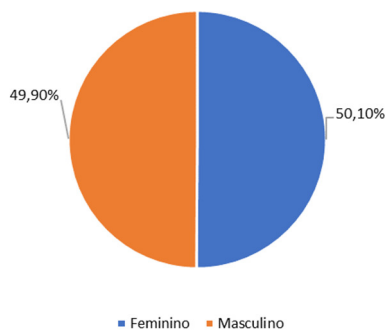


Figura 13 - Distribuição dos registros de nascimentos segundo sexo, no grupo etário de pais de 15 a 19 anos, 2013-2018

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

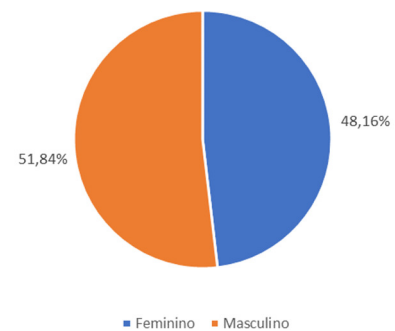


Figura 14 - Distribuição dos registros de nascimentos segundo sexo, no grupo etário de pais de 20 a 24 anos, 2013-2018.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Entre os locais de nascimento, a maioria (99,74%) ocorreu em hospitais. O restante ocorreu disperso em domicílio (0,19%); em trânsito (0,04%); em via pública (0,01%) e em um caso, não houve informação sobre o local de nascimento.

A distribuição anual dos registros de nascimentos é apresentada na Figura 15. Como se pode ver, a maior ocorrência se dá entre a população com idade entre 20 a 24 anos, em todos os anos. No grupo etário de 15 a 19 anos, o maior número de registros ocorreu no ano de 2014. Para ambos os grupos etários, a partir de 2014 o

número de registros apresenta queda. Acerca do registro para pais com idade abaixo de 16 anos, vale destacar a obrigatoriedade de estarem presentes seus representantes no momento do registro de paternidade.

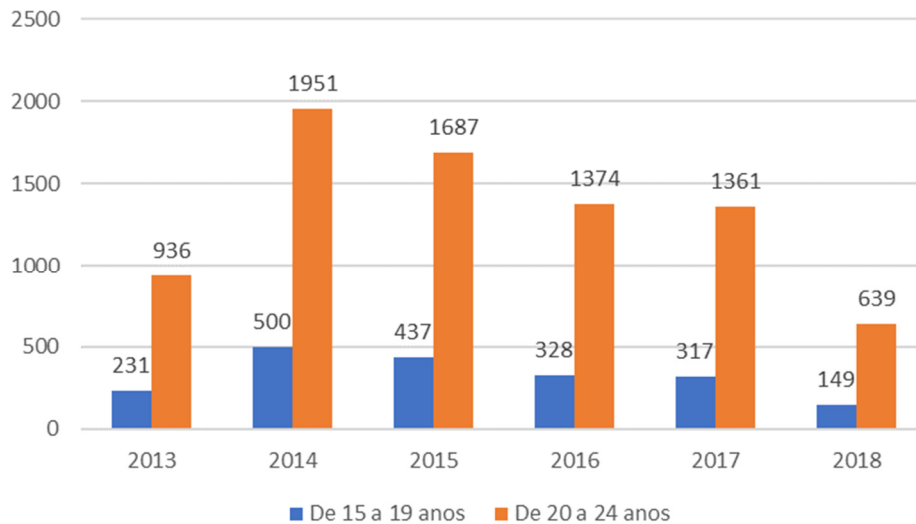


Figura 15 - Distribuição dos registros de nascimentos segundo ano, por grupo etário de interesse
 Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

No que tange à distribuição segundo ano, observa-se redução para ambos os grupos etários, com destaque para o grupo de 15 a 19 anos que registrou queda de 0,25% no ano de 2016 em relação ao ano de 2015. Apesar de se destacar em números absolutos, o grupo etário de 20 a 24 anos registrou queda no mesmo período, porém em menor percentual, de apenas 0,19% (Figura 16). Considerando o início do período (jul/2013), a maior variação percentual é observada entre o grupo de 15 a 19 anos, no período de julho /2015 a jun/2016, com queda de 21,11%. Neste mesmo período observou-se, também, a maior queda para o grupo de 20 a 24 anos (18,43%) (Figura 17).

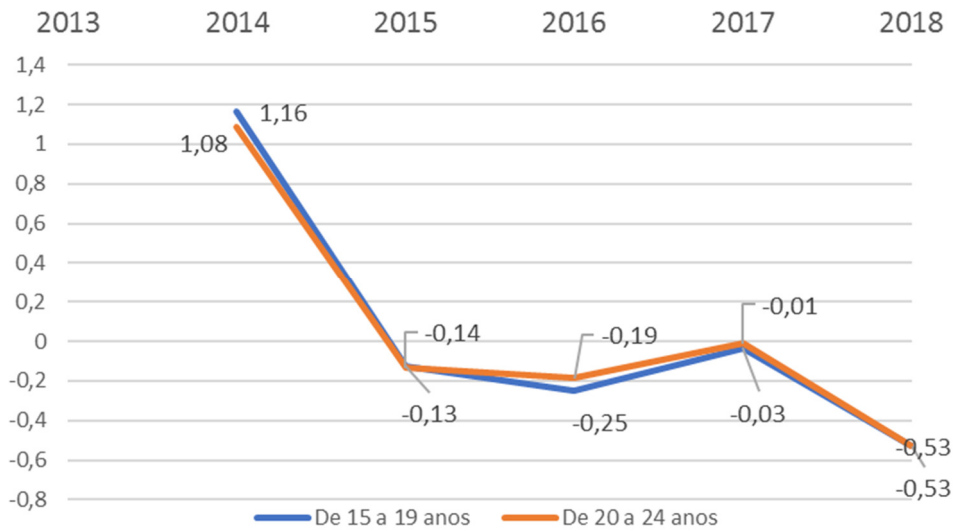


Figura 16 - Variação percentual entre os anos de análise para os grupos etários de interesse, 2013-2018.
Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

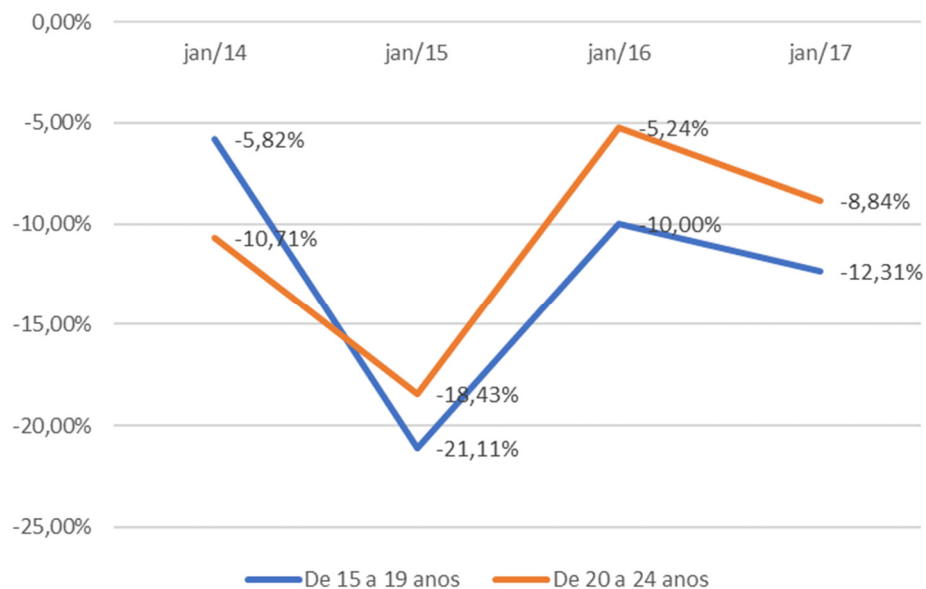


Figura 17 - Variação percentual entre os anos, para os grupos etários de interesse, considerando o início do período, jul/2013
Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Do ponto de vista da distribuição histórica, ambas as séries apresentam distribuição sazonal com tendência de queda ao longo dos anos. Observa-se uma tendência mais evidente para a série de registros do grupo etário de pais de 20 a 24 anos. Para ambas as séries é possível afirmar que não apresentam distribuição sazonal regular, cíclico, tão pouco estacionário (Figura 18).

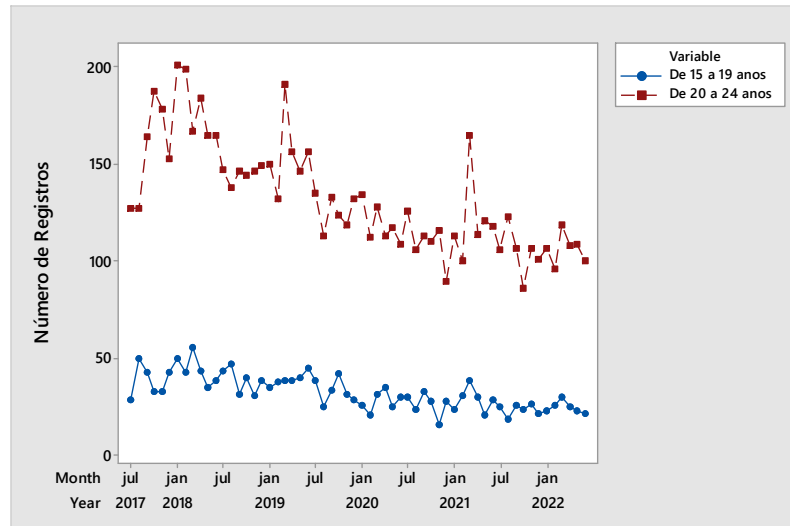


Figura 18 - Distribuição dos registros mensais em perspectiva histórica por grupo etário de interesse.
Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Quanto à distribuição dos registros de nascimento para o grupo de 15 a 19 anos ao longo dos anos em análise, o maior número de ocorrências observadas foi de 500 registros de nascimentos, no ano de 2014. Já o mínimo, em 2018, com 149. A média, ao longo dos anos foi de 327 registros. O quartil 1, ou seja, 25% dos registros estão abaixo de 210 casos, enquanto o quartil 3, acima de 452,8 casos. Já para o grupo de 20 a 24 anos, o maior número de ocorrências observadas foi de 1951 registros de nascimentos. O mínimo foi de 639 casos. A média, ao longo dos anos foi de 1325 registros. O quartil 1, ou seja, 25% dos registros estão abaixo de 826 casos, enquanto o quartil 3, acima de 1753 casos (Quadro 5 - Estatística descritiva do número de registros para os grupos etários).

Quadro 5 - Estatística descritiva do número de registros para os grupos etários de 15 a 19 anos e de 20 a 24 anos.

Grupo etário	Média	DP	Mínimo	Q1	Mediana	Q3	Máximo
15 a 19 anos	327	52,6	149	210,5	322,5	452,8	500
20 a 24 anos	1325	479	639	862	1369	1753	1951

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Os dados entre o grupo etário de pais de 15 a 19 anos apresentam-se dispersos ao longo dos anos, sem valores atípicos. A metade dos dados (50%) se distribui entre os quartis 1 e 3 (Figura 19). Os dados entre o grupo etário de pais de 20 a 24 anos apresentam-se dispersos ao longo dos anos, sem valores atípicos.

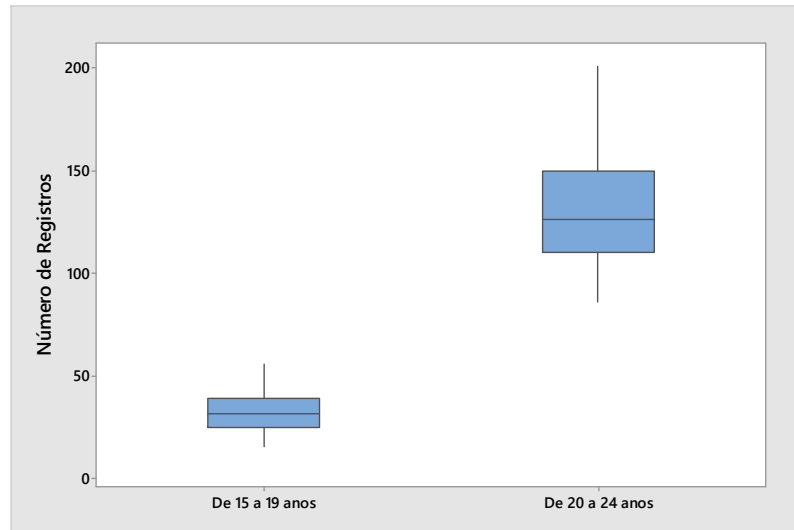


Figura 19 - Diagrama de caixa dos registros de nascimentos para os para os grupos etários de 15 a 19 anos e de 20 a 24 anos.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Ao observamos a distribuição dos registros ano a ano, verificamos para o grupo de 15 a 19 anos que a maior variabilidade é observada no ano de 2014. Já para o grupo de 20 a 24 anos, a maior variabilidade é observada em 2013. Ambas a séries apresentam uma tendência de redução com possibilidade de estabilidade. Para o grupo de 15 a 19 anos não se observa valores atípicos (Figura 20). Já para o grupo etário de 20 a 24 anos, registra-se valor atípico no ano de 2017 (Figura 21). Na distribuição mensal se observa o não alinhamento das medianas, permitindo a identificação de variação sazonal (Figura 22 e Figura 23).

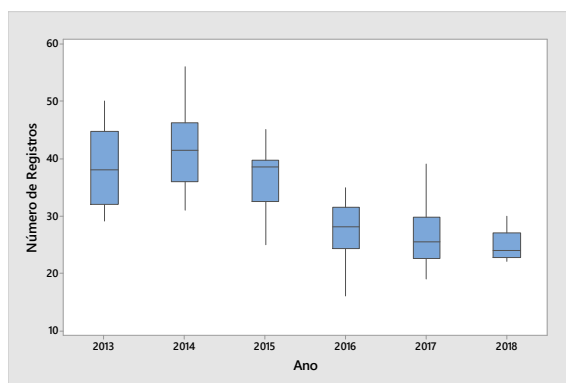


Figura 20 - Diagrama dos registros segundo ano de nascimento para o grupo etário de 15 a 19 anos.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

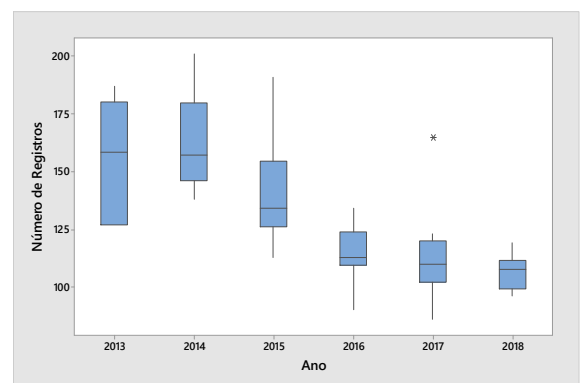


Figura 21 - Diagrama dos registros segundo ano de nascimento para o grupo etário de 20 a 24 anos.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

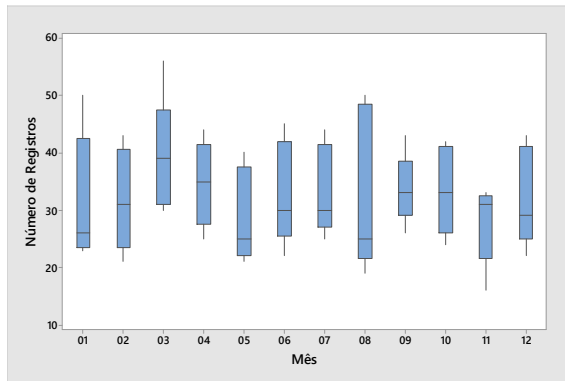


Figura 22 - Diagrama dos registros segundo mês de nascimento para o grupo etário de 15 a 19 anos.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

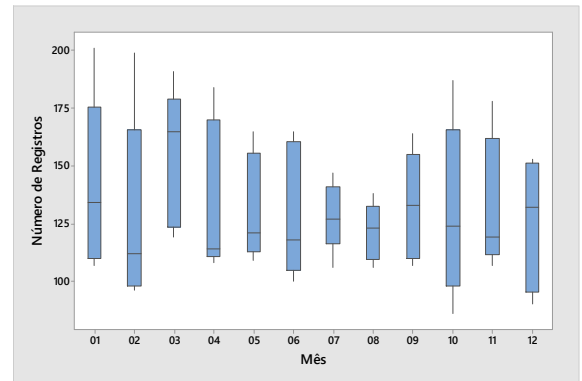


Figura 23 - Diagrama dos registros segundo mês de nascimento para o grupo etário de 20 a 24 anos.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

4 - Aplicação dos Modelos de Série Temporal

Neste capítulo serão apresentados os modelos de série temporal aplicado aos dados em análise. As séries de dados correspondem então aos registros de filhos de pais entre 15 a 19 anos e entre 20 a 24 anos coletados em uma unidade de cartório de Belo Horizonte que se referem aos registros ocorridos entre os anos de 2013 e 2018.

Os dados serão ajustados aos métodos de decomposição; ARIMA e suavização exponencial. Destaca-se no entanto que o método de decomposição permite descrever adequadamente a tendência e a sazonalidade da distribuição da série, mas não é o mais adequado para fazer previsões de valores futuros de séries temporais (MORETTIN e TOLOI, 2006). Sendo assim, seus resultados não serão considerados para a comparação dos modelos que serão apresentados a seguir, no entanto poderão evidenciar o comportamento dos dados no que diz respeito aos movimentos de tendência e sazonalidade.

4.1 - Ajuste dos modelos para os registros civis de filhos de pais entre 15 e 19 anos

Nesta seção serão apresentados os resultados dos modelos aplicados aos registros de nascimentos de filhos de pais entre 15 e 19 anos.

4.1.1 – Modelagem Através do Método de Decomposição Clássica

Como podemos observar no capítulo anterior, a distribuição dos registros de nascimento entre pais de 15 a 19 anos apresenta uma leve tendência de queda ao

longo dos anos. Após aplicar a análise de tendência, observamos, através da equação (27) de tendência que a taxa de queda dos dados é de 0,364712. Já o valor de intercepto (α_0) indica que o número de registros de nascimentos para pais deste grupo etário varia em torno de 43,82 registros mensais, independente do mês. Desta forma, quando t for igual a zero, o número de registros será de 43,82, e, na medida em que o t aumentar, o número de registros decrescerá a uma taxa de 0,364712. A Figura 24 mostra a tendência estimada (de acordo com o método descrito na seção 2), e os valores reais observados da série para todo o período analisado, bem como os erros do ajuste.

$$\text{Equação de tendência: } Y_t = 43,82 - 0,364712 * t, \quad (27)$$

onde Y_t é o número de registros mensais no instante t e t é o tempo previsto.

A análise do vetor de resíduos do modelo de tendência para esta série de dados apresenta variação em torno de zero, embora não apresente uma distribuição próxima da normal (Figura 25).

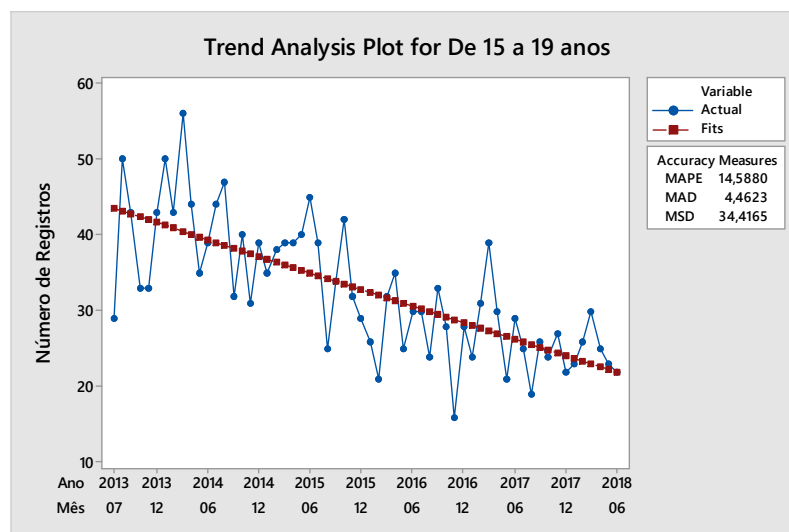


Figura 24 - Análise de tendência para a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo: equação (27), 2013-2018.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

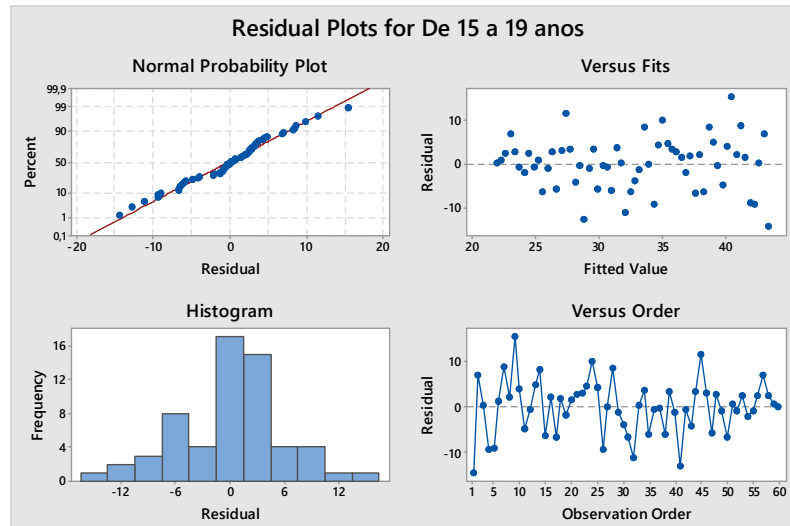


Figura 25 - Gráficos dos resíduos da análise de tendência para a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo: equação (27), 2013-2018.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Considerando a tendência e a variabilidade sazonal da série, aplicamos o método de decomposição, conforme explicitado no item 2.1, para reavaliar a distribuição dos registros de nascimentos quanto aos aspectos de tendência e sazonalidade. A seguir, apresentam-se os resultados obtidos a partir da decomposição aplicada e da equação de tendência ajustada.

Ao compararmos a equação de tendência (27) com a do modelo atual (28), observamos que o intercepto (α_0) aumenta mesmo que pouco, aumentando portanto, a média em torno da qual os valores de registros mensais se distribuem. No entanto, mesmo que seja pequena a diferença percebida, o coeficiente de variação (β_0) indica maior queda.

$$\text{Equação de tendência ajustada: } Y_t = 44,11 - 0,374194 * t. \quad (28)$$

O índice sazonal resultante da análise de decomposição indica que os meses de fevereiro, maio, julho e novembro, entre este grupo etário, são os que mais sofrem variação negativa. Já os meses de janeiro, outubro e setembro, as variações são positivas. Junho também sofre variação negativa, porém menor que o observado nos demais meses (Quadro 6). A Figura 26 apresenta os valores da série real; da série ajustada, e a tendência decrescente da série. Observa-se aqui que a taxa de queda é maior que o modelo anterior (Figura 24) que considera apenas a tendência.

Quadro 6 - Índices Sazonais do Modelo de Decomposição ajustado aos registros de filhos de pais entre 15 e 19 anos.

Período	Índice
1	3,00347
2	-5,20486
3	0,39931
4	0,77431
5	-3,89236
6	-0,99653
7	-3,55903
8	0,46181
9	7,37847
10	3,12847
11	-4,12153
12	2,62847

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

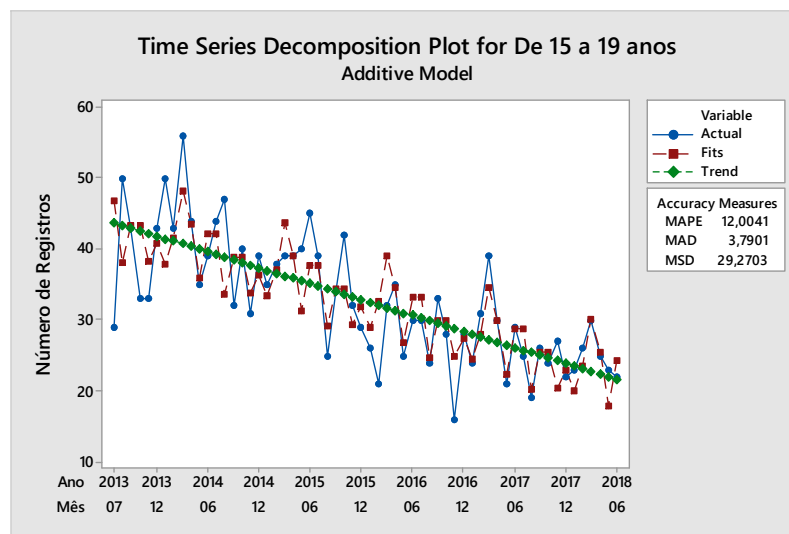


Figura 26 - Análise de decomposição (tendência e sazonalidade) para registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equação (28).

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

A Figura 27 nos mostra a representação gráfica dos índices sazonais, na qual percebe-se a descrição anterior acerca dos meses que refletem variação positiva e negativa. Em termos percentuais, os meses que mais variaram foram os meses de janeiro e fevereiro, e o mês que menos variou, foi outubro. A possível hipótese para esta variação pode ser atribuída à falta de planejamento observada entre os pais jovens, não indicando um padrão de concepção e posterior registro. Os dados representados no boxplot desta figura indicam que, após o ajuste do modelo, o mês com o maior número de registros é setembro, evidenciando uma concentração do momento de concepção coincidindo com o período de férias (janeiro).

A Figura 28 apresenta a distribuição dos dados: sem a tendência; com o ajuste de sazonalidade, e; então, aplicado o modelo, indicando que o modelo se ajusta bem à série.

Quando analisadas as estatísticas dos erros (MAPE; MAD e MSE), este modelo, que considera a tendência e a sazonalidade, apresenta-se melhor ajustado à série de dados, e portanto, com melhores coeficientes e menores valores. O MAPE mostrou-se 2,5839 menor; o MAD, 0,6722; o MSE, 5,1462. Além disso, os valores dos resíduos da decomposição da série de registros de nascimentos do grupo etário de pais entre 15 e 19 anos variam em torno de zero, apresentando uma distribuição próxima da Normal (Figura 29).

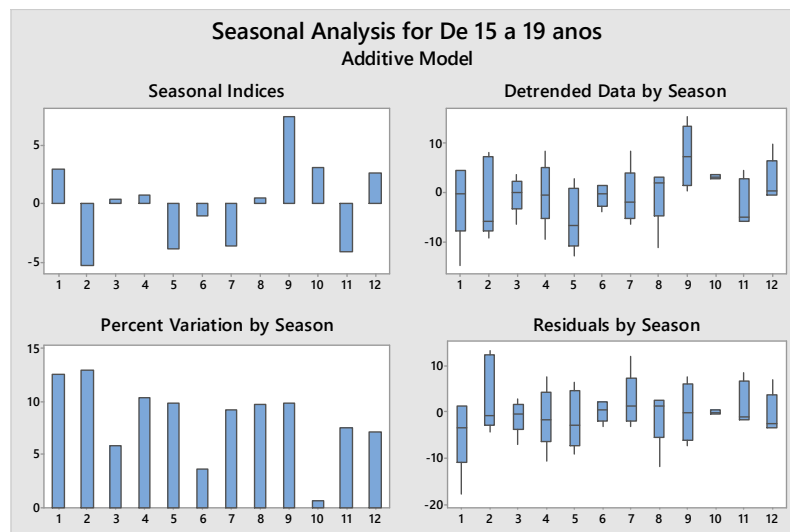


Figura 27 - Análise dos índices sazonais para a série registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equação (28).

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova

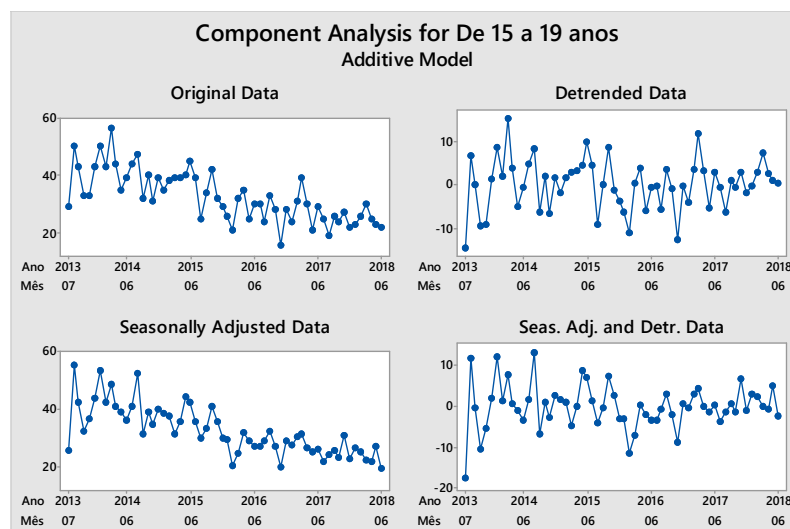


Figura 28 - Análise de sazonalidade para a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equação (28).

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova

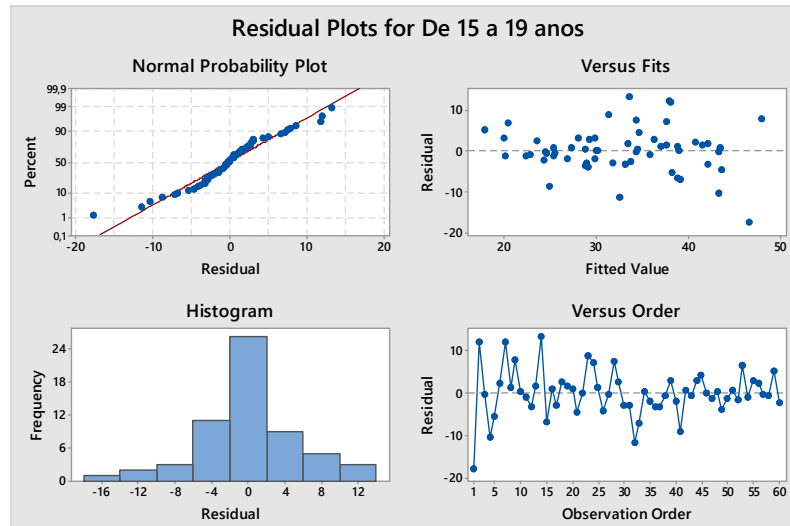


Figura 29 – Gráficos dos resíduos da análise de decomposição (tendência e sazonalidade) para a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equação (28).

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

4.1.2 - Modelagem Através do Método ARIMA

Como visto no item anterior, a série de 15 a 19 anos apresenta uma tendência de declínio, portanto, não é estacionária. Ao aplicar uma diferença $d=1$ observa-se que a série se torna estacionária. Considerando um nível de significância de 5%, o correlograma da FAC (Figura 30) indica um modelo auto regressivo e a FACP (Figura 31) indica uma correlação de ordem 2 pois apresenta duas autocorrelações significativas. Assim, o modelo ARIMA (2,1,0) pode ser ajustado à série de registros de nascimentos de filhos de pais entre 15 a 19 anos.

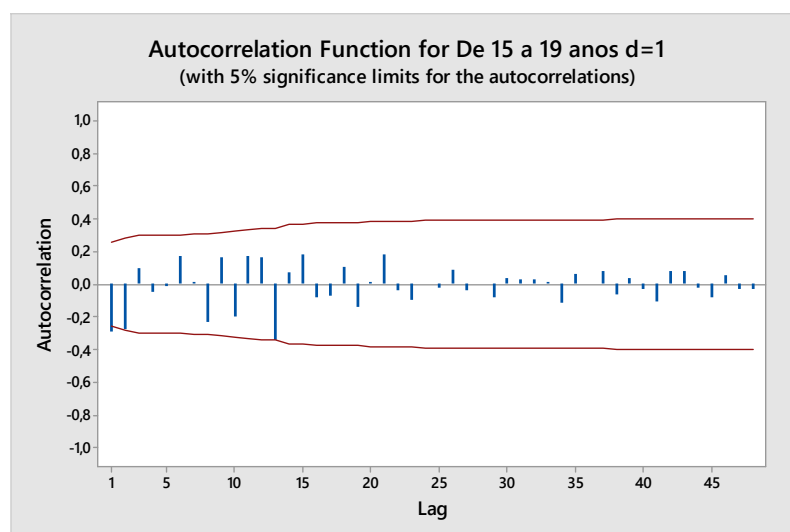


Figura 30 - Função de autocorrelação da série de registros de filhos de pais entre 15 e 19 anos – dados mensais com a diferença aplicada ($d=1$).

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

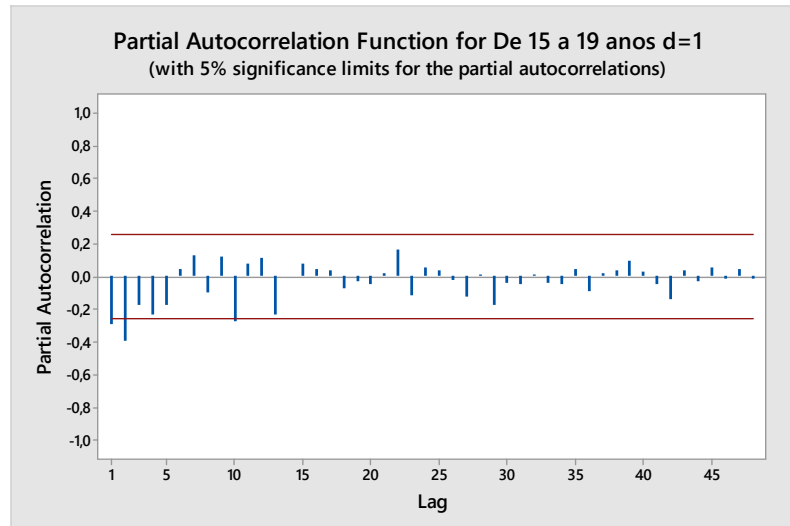


Figura 31 - Função de autocorrelação parcial da série de registros de filhos de pais entre 15 e 19 anos – dados mensais com a diferença aplicada ($d=1$).

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

As estimativas dos parâmetros do modelo ARIMA (2, 1, 0) ajustado utilizando-se todos os meses do período ($n=60$ observações), foram obtidas após 9 iterações e encontram-se no Quadro 7 e a equação de ajuste apresenta-se abaixo (30):

$$\hat{Y}_t = (1 - \hat{\theta}_1)Y_{t-1} + (\hat{\theta}_2 - \hat{\theta}_1)Y_{t-2} - \hat{\theta}_2 Y_{t-3} + \hat{\theta}_0 + a_t. \quad (29)$$

$$\hat{Y}_t = 1,4672Y_{t-1} + 0,0113Y_{t-2} + 0,4559Y_{t-3} + a_t. \quad (30)$$

Quadro 7 - Estimativas e parâmetros do modelo ARIMA (2, 1, 0) para a série de registros de filhos de pais entre 15 e 19 anos, $n=60$.

Estimates at each iteration					
Iteration	SSE	Parameters			
0	3909,23	0,100	0,100		
1	3297,76	-0,050	-0,044		
2	2868,66	-0,200	-0,189		
3	2624,39	-0,350	-0,336		
4	2568,11	-0,439	-0,424		
5	2564,50	-0,460	-0,447		
6	2564,25	-0,465	-0,454		
7	2564,24	-0,467	-0,455		
8	2564,23	-0,467	-0,456		
9	2564,23	-0,467	-0,456		
Relative change in each estimate less than 0,0010					
Final Estimates of Parameters					
Type	Coef	SE Coef	T	P	
AR 1	-0,4672	0,1171	-3,99	0,000	
AR 2	-0,4559	0,1165	-3,91	0,000	
Differencing: 1 regular difference					
Number of observations: Original series 60, after differencing 59					
Residuals: SS = 2474,44 (backforecasts excluded)					
MS = 43,41 DF = 57					
Modified Box-Pierce (Ljung-Box) Chi-Square statistic					
Lag	12	24	36	48	
Chi-Square	19,9	31,2	39,1	52,7	
DF	10	22	34	46	
P-Value	0,030	0,093	0,250	0,230	

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Os termos auto-regressivo tem um p-valor menor que o nível de significância de 0,05, indicando que os coeficientes são estatisticamente significativos e podem ser mantidos no modelo. Os valores do correlogramas dos resíduos (Figura 32), a 5% de significância, mostram um comportamento similar ao de um ruído branco, embora a correlação na defasagem 3 esteja um pouco acima do limite inferior do intervalo de confiança e na defasagem de ordem 8 esteja bem próxima desse limite. O p-valor = 0,03 do teste da estatística qui-quadrado de Ljung-Box indica que há significância no conjunto das 12 primeiras autocorrelações dos resíduos (a 5% de significância). Mesmo assim, pode-se considerar que os resíduos estão próximos de uma série ruído branco. Os resíduos se distribuem ao redor de zero e com uma variância estimada igual a 43,41. Os valores dos resíduos do modelo ARIMA (2, 1, 0) variam em torno de zero, embora não apresentem uma distribuição próxima da Normal (Figura 34).

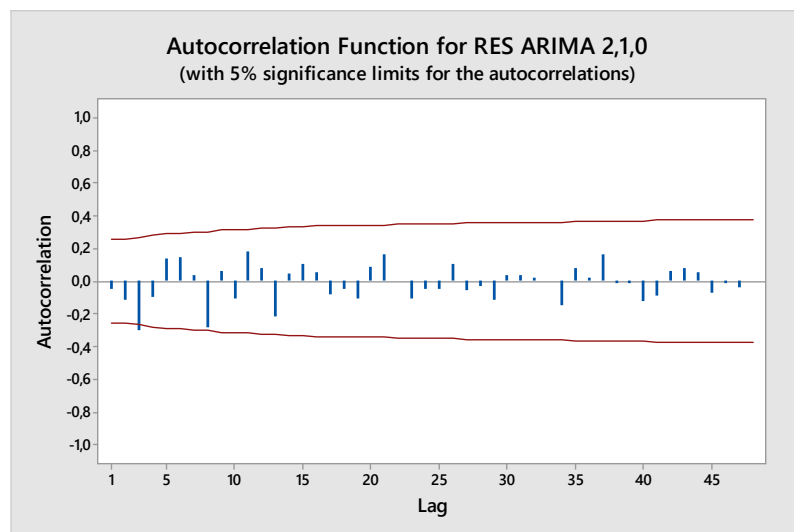


Figura 32 – Correlograma do vetor de resíduos do modelo ajustado para a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equação (30).

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

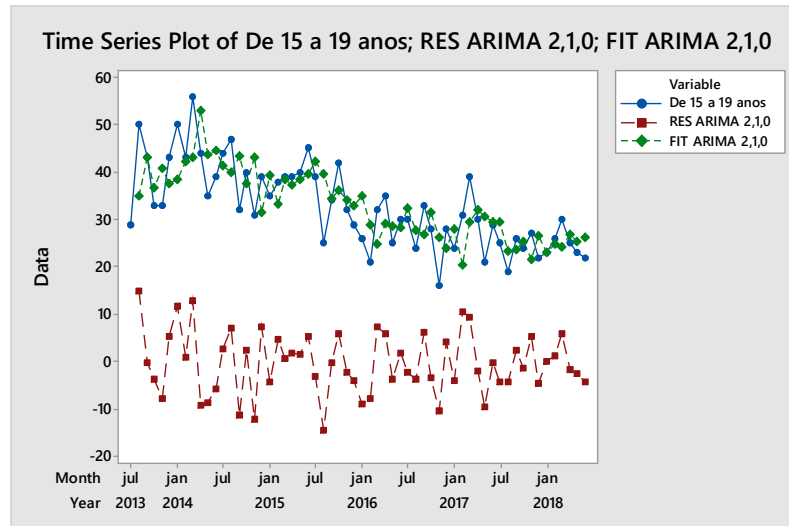


Figura 33 – Dados ajustados segundo o modelo ARIMA (2, 1, 0) para a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equação (30).

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

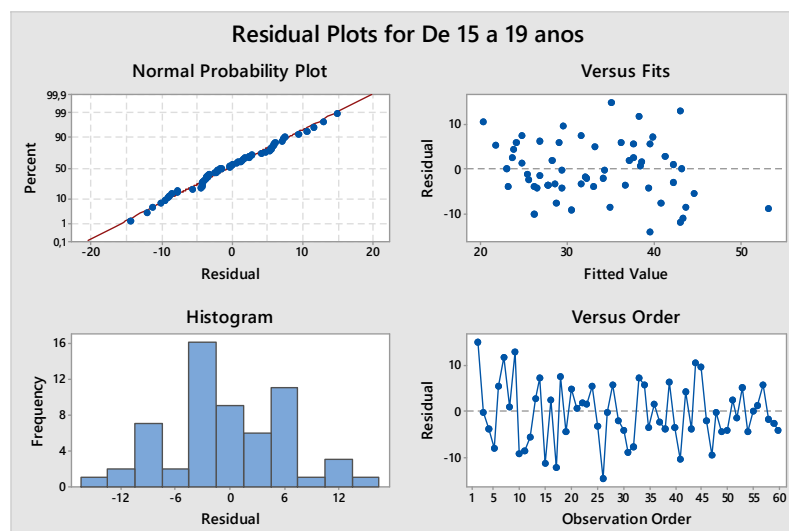


Figura 34 - Gráficos dos resíduos do modelo ARIMA (2, 1, 0) para a série de registros de filhos de pais entre 15 e 19 anos.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Uma vez verificado que o modelo ARIMA (2, 1, 0) se ajusta bem à série de dados de registros de filhos de pais entre 15 e 19 anos optou-se por realizar a previsão de valores futuros da série. A finalidade do ajuste deste modelo não é fazer uma comparação como o modelo resultante da equação (30), uma vez que aquele foi ajustado para todo o período ($n=60$ observações) e este apenas para o período amostral, que compreende ao período de julho de 2013 a junho de 2017 ($n=48$). Este modelo, permite, portanto que se faça as previsões de valores futuros para o período de um ano, considerando a origem de previsão no tempo $t=48$ (julho/2017) usando o modelo ARIMA (2, 1, 0).

A seguir a equação do modelo estimado para o ajuste de previsão da série observada:

$$Y_t = 1,4599Y_{t-1} - 0,0153Y_{t-2} - 0,4752Y_{t-3} + a_t. \quad (31)$$

O Quadro 8 nos mostra os resultados do ajuste do modelo ARIMA (2, 1, 0) com n=48 observações. Tais resultados mostram-se coerentes com os observados para o modelo ajustado com n=60 em termos de significância dos coeficientes ao nível de 5%. Os coeficientes são significativos e os resíduos são próximos de uma série ruído branco. A variância dos resíduos é maior (51,47) o que é natural, já que o tamanho da amostra é menor. O gráfico do correlograma dos resíduos do modelo de previsão (n=48) (Figura 35), mostra que a função de autocorrelação é semelhante ao que fora observado no modelo ajustado ao período completo (n=60).

Quadro 8 - Estimativas e parâmetros a partir do modelo ARIMA (2, 1, 0) para a série de registros de filhos de pais entre 15 e 19 anos, n=48.

Estimates at each iteration					
Relative change in each estimate less than 0,0010					
Final Estimates of Parameters					
Type		Coef	SE Coef	T	P
AR	1	-0,4599	0,1316	-3,50	0,001
AR	2	-0,4752	0,1322	-3,59	0,001
Differencing: 1 regular difference					
Number of observations: Original series 48, after differencing 47					
Residuals: SS = 2316,04 (backforecasts excluded)					
MS = 51,47 DF = 45					
Modified Box-Pierce (Ljung-Box) Chi-Square statistic					
Lag		12	24	36	48
Chi-Square		18,0	37,3	54,1	*
DF		10	22	34	*
P-Value		0,055	0,022	0,016	*

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

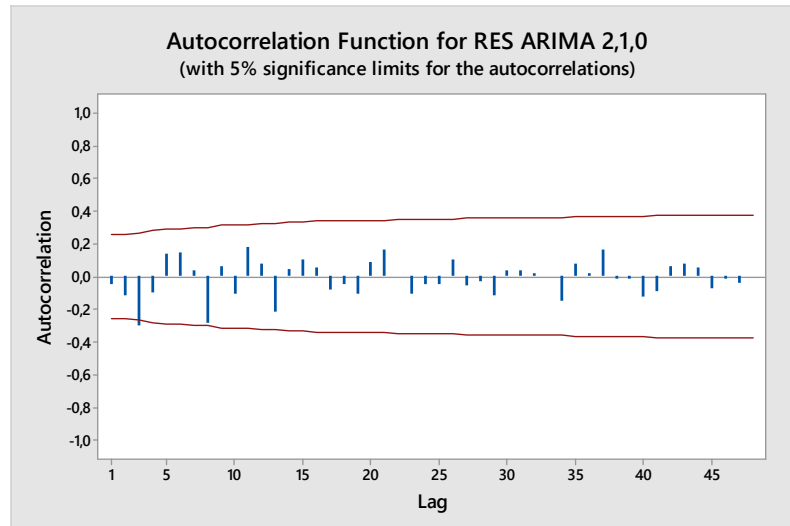


Figura 35 – Correlograma do vetor de resíduos do modelo ARIMA (2, 1, 0) da série de registros de filhos de pais entre 15 e 19 anos, $n=48$.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

No Quadro 9 é possível comparar os valores resultantes da previsão através da indicação de previsões (*Forecast*) com os valores observados no período de validação que correspondem ao observado nos registros reais (*Actual*). As previsões também podem ser vistas na Figura 36. Destaca-se que para ter um bom ajuste de previsão a partir de um modelo ARIMA, é desejável que o número de observações para o período amostral seja maior que 48.

Quadro 9 – Previsões, limites inferiores e superiores a partir do modelo ARIMA (2, 1, 0) e valores reais de registros de filhos de pais entre 15 e 19 anos, $n=48$.

Previsões com origem no período 48 (Forecasts from period 48)				
95% Limits				
Period	Forecast	Lower	Upper	Actual
49	29,5980	15,5339	43,6620	25
50	25,5211	9,5368	41,5054	19
51	27,1118	10,6617	43,5620	26
52	28,3177	9,7243	46,9112	24
53	27,0072	6,6733	47,3411	27
54	27,0368	5,7903	48,2833	22
55	27,6460	5,1741	50,1179	23
56	27,3517	3,5743	51,1292	26
57	27,1976	2,4078	51,9874	30
58	27,4083	1,6090	53,2077	25
59	27,3847	0,5326	54,2368	23
60	27,2954	-0,5166	55,1074	22

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

A Figura 36 mostra que os dados reais, do período de validação, estão dentro da estimativa intervalar (95% de confiança). As estimativas pontuais estão variando acima do observado no período de validação, exceto no mês de março. Observa-se grande amplitude dos intervalos de confiança especialmente para os meses mais distantes da origem ($t=48$ – julho/2017).

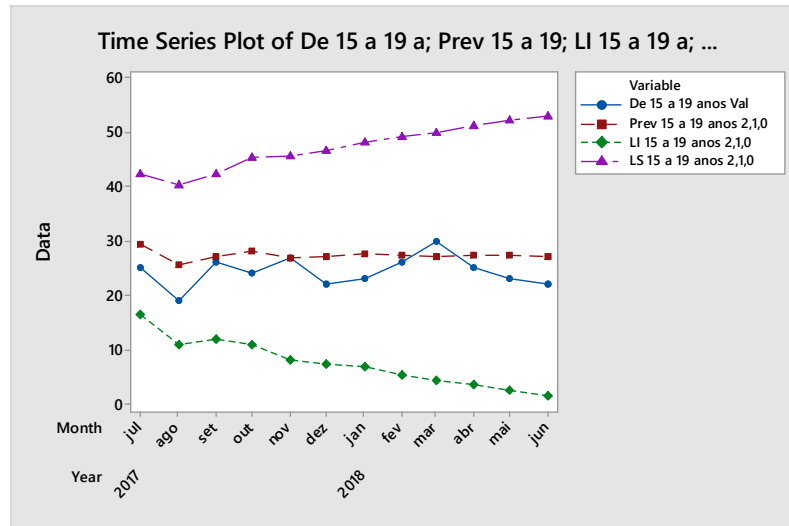


Figura 36 - Previsão para a série de registros de filhos de pais entre 15 e 19 anos, modelo ARIMA (2, 1, 0), $n=48$.
 Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

4.1.3 - Modelagem Através do Método de Alisamento Exponencial - Holt

Ao observar que a série de registros de nascimentos no grupo etário de pais entre 15 e 19 anos apresenta tendência optou-se por aplicar o método de suavização exponencial de Holt (ver item 2.3), em que se determina as constantes de suavização para o nível e para a tendência da série (α e γ).

As equações que resultam na suavização da tendência e do nível da série são dadas por:

$$L_t = 0,6883Y_t + 0,3117[L_{t-1} + Y_t]. \quad (32)$$

$$T_t = 0,03976[L_t - L_{t-1}] + 0,96024T_{t-1}. \quad (33)$$

Onde L_t nos dá a suavização do nível da série e T_t , da tendência da série.

A Figura 37 apresenta a série de registros de nascimentos no grupo etário de pais entre 15 e 19 anos real e a suavizada (*Smoothed*) após o ajuste do modelo Holt $n=60$ que se ajustam bem. Os valores dos resíduos do modelo Holt variam em torno de zero com registro de um valor desalinhado, embora não apresentem uma distribuição próxima da Normal (Figura 38).

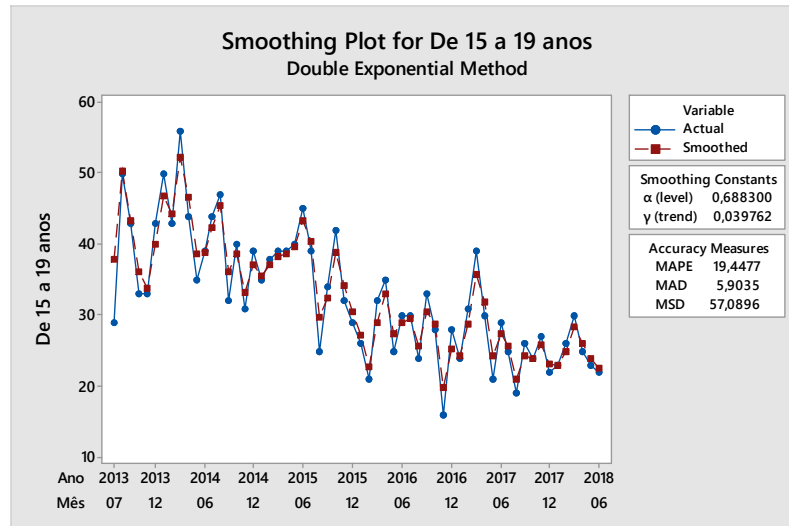


Figura 37 - Análise de suavização (nível e tendência) para a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equações (32 e 33), $n=60$.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

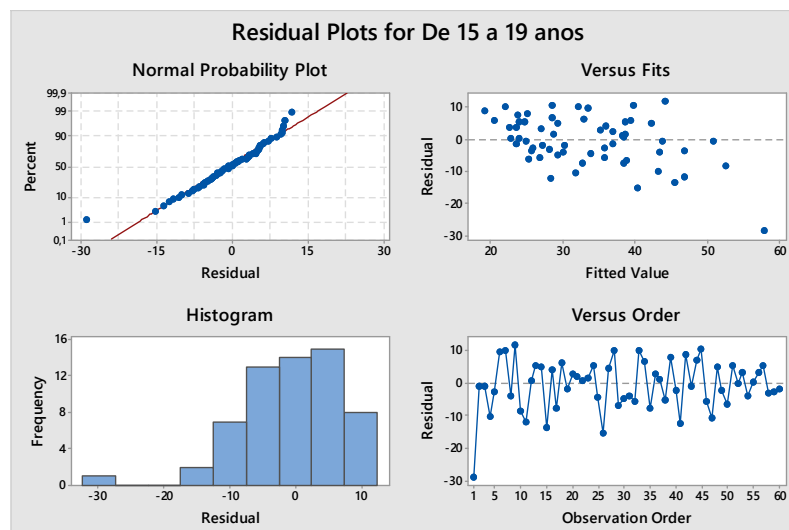


Figura 38 - Gráficos dos resíduos do modelo Holt para a a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equações (32 e 33), $n=60$.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Como se observa que os dados ajustados ao modelo Holt se ajustam bem à série real de registros de nascimentos no grupo etário de pais entre 15 e 19 anos optou-se por realizar a previsão de valores futuros da série utilizando este modelo. A finalidade do ajuste deste modelo não é fazer uma comparação como o modelo resultante das equações (31 e 32), uma vez que aquele fora ajustado para todo o período ($n=60$ observações) e o que se segue, apenas para o período amostral, que compreende ao período de julho de 2013 a junho de 2017 ($n=48$). Este modelo, permite, portanto que se faça as previsões de valores futuros para o período de um ano, considerando a origem de previsão no tempo $t=48$ (julho/2017).

A seguir as equações do modelo estimado para o ajuste de previsão da série observada bem como os valores previstos, os limites e os valores reais (Quadro 10):

$$L_t = 0,5721Y_t + 0,4279[L_{t-1} + Y_t]. \quad (34)$$

$$T_t = 0,0607[L_t - L_{t-1}] + 0,9393T_{t-1}. \quad (35)$$

Quadro 10 - Previsões, limites inferiores e superiores a partir do modelo Holt e valores reais de registros de filhos de pais entre 15 e 19 anos, $n=48$.

Length	48			
Smoothing Constants				
Alpha (level)	0,572079			
Gamma (trend)	0,060721			
Forecasts				
Period	Forecast	Lower	Upper	Actual
49	27,1453	11,6603	42,6302	25
50	26,8954	8,6923	45,0986	19
51	26,6456	5,4532	47,8380	26
52	26,3958	2,0426	50,7489	24
53	26,1459	-1,4806	53,7725	27
54	25,8961	-5,0808	56,8730	22
55	25,6463	-8,7355	60,0281	23
56	25,3964	-12,4300	63,2229	26
57	25,1466	-16,1543	66,4475	30
58	24,8968	-19,9014	69,6950	25
59	24,6469	-23,6665	72,9604	23
60	24,3971	-27,4458	76,2400	22

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

A Figura 39 mostra os dados reais e os dados suavizados para o período amostral. A Figura 40 mostra que os dados reais, correspondentes ao período de validação, estão dentro da estimativa intervalar (95% de confiança). As estimativas pontuais estão variando acima do observado no período de validação, exceto no mês de março. Observa-se grande amplitude dos intervalos de confiança especialmente para os meses mais distantes da origem (julho/2017).

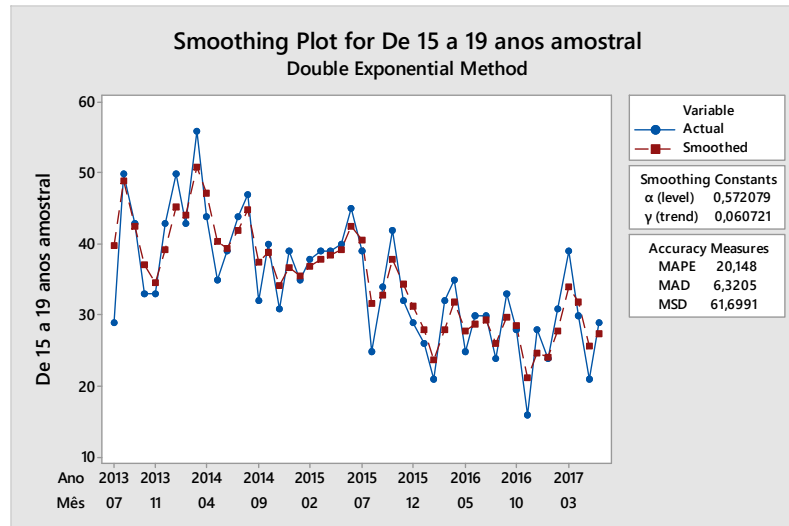


Figura 39 - Análise de suavização (nível e tendência) para a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equações (34 e 35), $n=48$.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

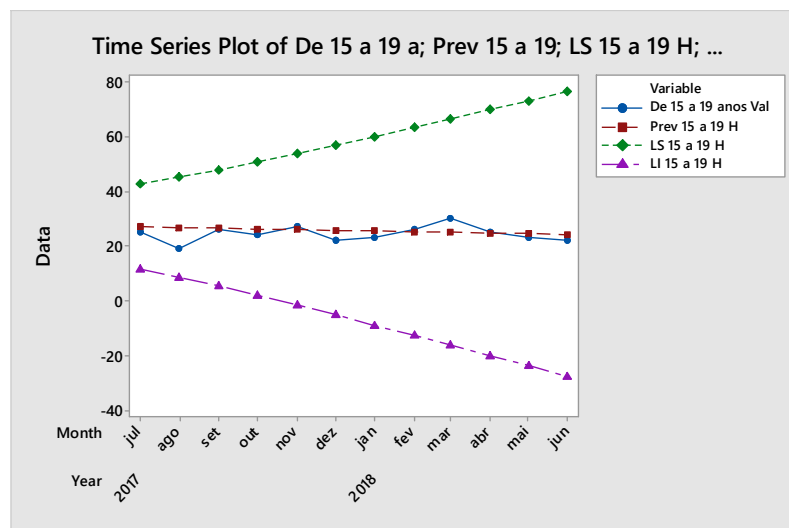


Figura 40 - Previsão para a série de registros de filhos de pais entre 15 e 19 anos, modelo Holt, $n=48$.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

4.1.4 - Modelagem Através do Método de Alisamento Exponencial – Holt-Winter

Considerando a sazonalidade observada no item 4.1.1, optou-se por aplicar o método de Holt-Winter, que considera, além do nível e da tendência, a sazonalidade (ver item 2.3), determinando, então, as constantes de suavização para o nível, para a tendência e para a sazonalidade da série (α , γ e δ).

As equações que resultam na suavização da tendência, do nível e da tendência da série são dadas por:

$$L_t = 0,2(Y_t - S_{t-s}) + 0,8(L_t - T_{t-1}); \quad (36)$$

$$T_t = 0,2(L_t - L_{t-1}) + 0,8T_{t-1}; \quad (37)$$

$$S_t = 0,2(Y_t - L_t) + 0,8S_{t-s}. \quad (38)$$

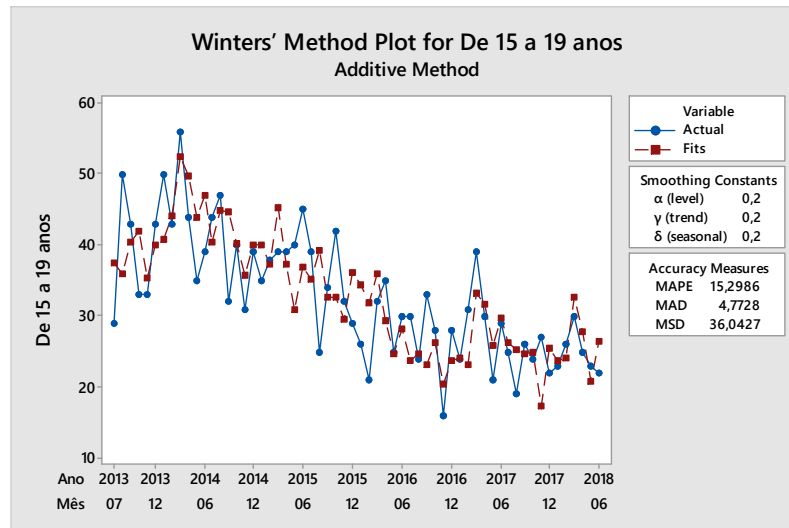


Figura 41 - Análise de suavização (nível, tendência e sazonalidade) para a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equações (36, 37 e 38), $n=60$.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Como se observa que os dados ajustados ao modelo Holt-Winter se ajustam bem à série real de registros de nascimentos no grupo etário de pais entre 15 e 19 anos optou-se por realizar a previsão de valores futuros da série utilizando este modelo. A finalidade do ajuste deste modelo não é fazer uma comparação como o modelo resultante das equações (36, 37 e 38), uma vez que aquele fora ajustado para todo o período ($n=60$ observações) e o que se segue, apenas para o período amostal, que compreende ao período de julho de 2013 a junho de 2017 ($n=48$). Este modelo, permite, portanto que se faça as previsões de valores futuros para o período de um ano, considerando a origem de previsão no tempo $t=48$ (julho/2017).

A seguir o modelo estimado para o ajuste de previsão da série bem como os valores previstos, os limites e os valores reais (Quadro 11):

Quadro 11 - Previsões, limites inferiores e superiores a partir do modelo Holt-Winter $n=48$ e valores reais de registros de filhos de pais entre 15 e 19 anos.

Additive Method				
Data	De 15 a 19 anos			
Length	48			
Smoothing Constants				
Alpha (level)	0,2			
Gamma (trend)	0,2			
Delta (seasonal)	0,2			
Forecasts				
Period	Forecast	Lower	Upper	Actual
49	26,3230	13,8292	38,8169	25
50	26,2966	13,6070	38,9862	19
51	26,5387	13,6308	39,4465	26
52	26,9619	13,8144	40,1094	24
53	18,8899	5,4825	32,2972	27
54	25,8860	12,1996	39,5723	22
55	24,8372	10,8539	38,8204	23
56	25,1462	10,8492	39,4433	26
57	33,6203	18,9938	48,2468	30
58	29,4969	14,5262	44,4677	25
59	22,8934	7,5648	38,2220	23
60	28,8855	13,1862	44,5849	22

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

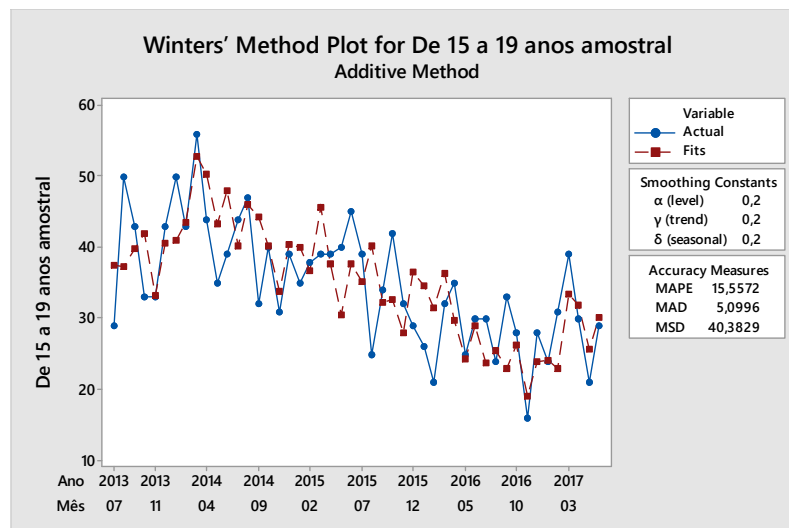


Figura 42 - Análise de suavização (nível, tendência e sazonalidade) para a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equações (36, 37 e 38), $n=48$.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Os ajustes acompanham os dados relativamente de perto, com padrão sazonal e tendência estáveis no final do período de registro, podendo concluir que no período usado para a validação do modelo, esse produziu previsões satisfatórias (Figura 43).

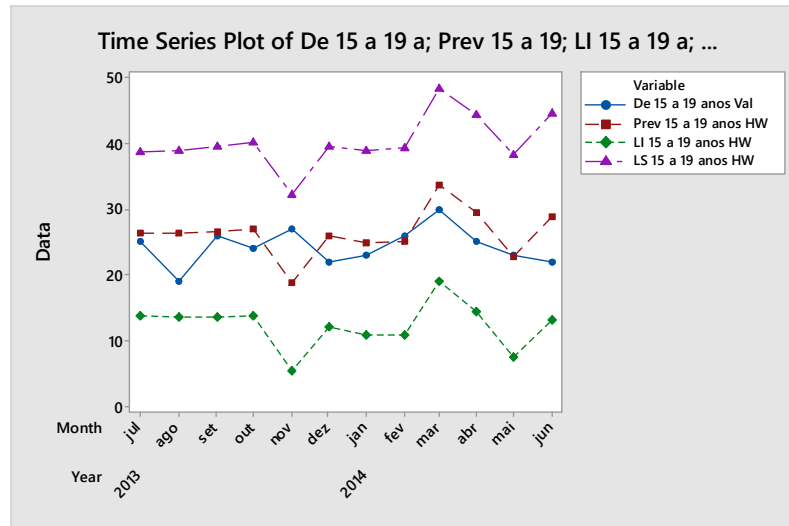


Figura 43 - Previsão para a série de registros de filhos de pais entre 15 e 19 anos, modelo Holt-Winter, $n=48$.
Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

4.2 - Ajuste dos modelos para os registros civis de filhos de pais entre 20 e 24 anos

Nesta seção serão apresentados os resultados dos modelos aplicados aos registros de nascimentos de filhos de pais entre 20 e 24 anos.

4.2.1 – Modelagem Através do Método de Decomposição Clássica

Como visto na seção 3, a série de registros de nascimentos do grupo etário de pais de 20 a 24 anos também apresenta sazonalidade e tendência. Ao aplicarmos a análise de tendência, observamos, através da equação de estimação da tendência que os dados apresentam uma taxa de queda de 1,23534 (equação 39), portanto, maior que o observado no grupo etário de 15 a 19 anos. O valor do intercepto (α_0) indica que o número de registros de filhos de pais deste grupo etário varia em torno de 170,18 registros/mês. A Figura 44 mostra a tendência para o período completo de observações e os erros do ajuste.

$$\text{Equação de Tendência: } Y_t = 170,18 - 1,23534 * t. \quad (39)$$

Onde Y_t é o número de registros mensais no instante t e t é o tempo previsto.

A análise do vetor de resíduos do modelo de tendência para esta série de dados apresenta variação em torno de zero com distribuição próxima da normal (Figura 45).

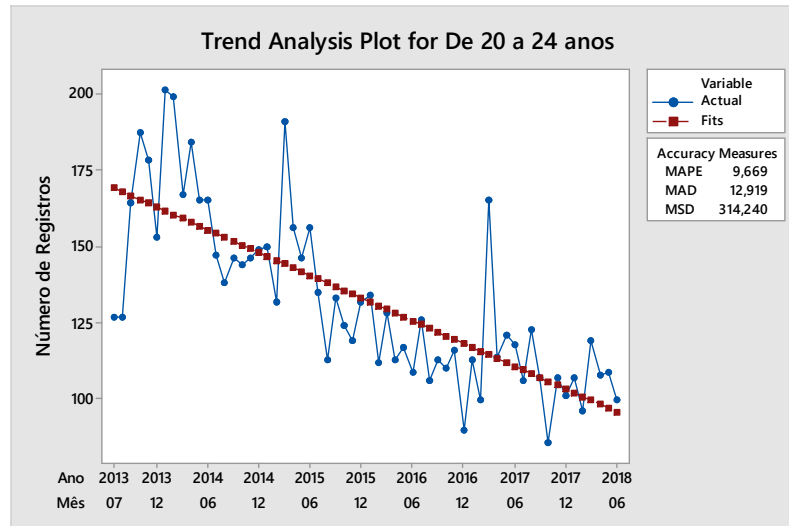


Figura 44 - Análise de tendência para a série registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo: equação (39), 2013-2018.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

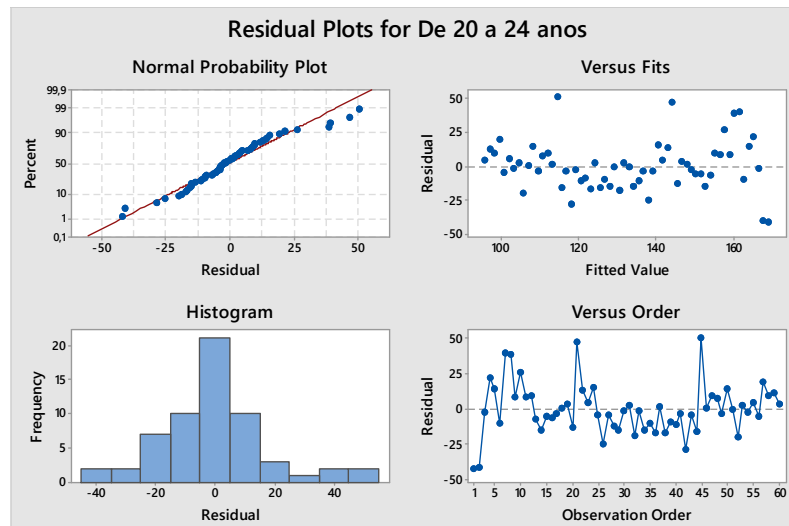


Figura 45 - Gráficos dos resíduos da análise de tendência para a série de registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo: equação (39), 2013-2018.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Considerando a tendência e a variabilidade sazonal da série, aplicamos o método de decomposição para reavaliar a distribuição. A seguir, apresentam-se os resultados obtidos a partir da decomposição aplicada e da equação de tendência ajustada.

Ao compararmos a equação estimada de tendência dos dados reais (39) com a do modelo (40), observamos que o intercepto aumenta para 172,04 e o coeficiente para 1,29640.

$$\text{Equação de tendência ajustada: } Y_t = 172,04 - 1,29640 * t. \quad (40)$$

O índice sazonal resultante da decomposição indica que apenas os meses de julho e de setembro até dezembro não sofrem variação negativa. Entre estes meses, o que mais varia positivamente é o mês de setembro. Já entre os meses com variação negativa, o mês de fevereiro é o que apresenta maior variação, seguido do mês de agosto (Quadro 12). A Figura 46 apresenta os valores da série real, da série ajustada, e da tendência da série, apontando a manutenção da queda.

Quadro 12 - Índices Sazonais do Modelo de Decomposição ajustado para a série de registros de filhos de pais entre 20 e 24 anos.

Period	Index
1	-5,7587
2	-11,1128
3	-2,5295
4	-7,7378
5	-2,8628
6	-3,1545
7	6,5122
8	-11,9670
9	26,4288
10	6,1163
11	2,1372
12	3,9288

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

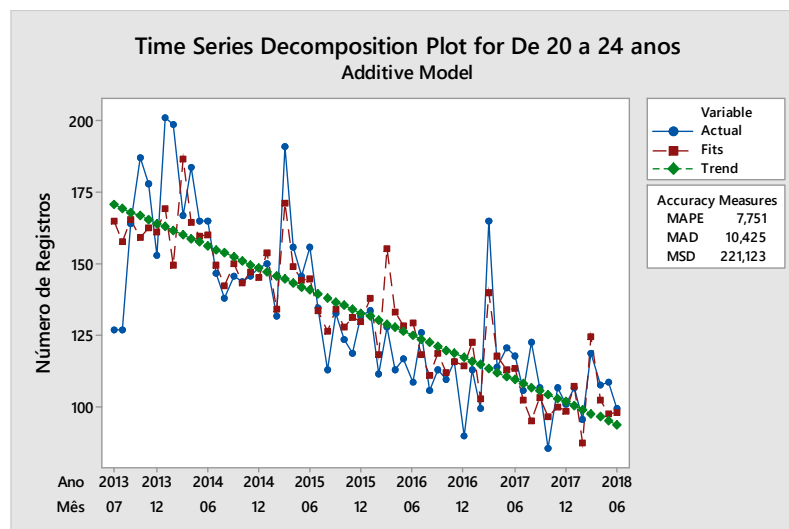


Figura 46 - Análise de decomposição (tendência e sazonalidade) para a série de registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo ajustado: equação (40).

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

A Figura 47 mostra a representação gráfica dos índices sazonais, na qual percebe-se a descrição anterior acerca dos meses que refletem variação positiva e negativa. Os meses de julho, e de setembro até dezembro apresentam índices sazonais positivos. Em termos percentuais, os meses que mais variaram foram os meses de agosto e setembro, e o mês que menos variou, março. A partir do boxplot é possível perceber a mesma distribuição observada no grupo etário mais jovem: o maior

número de registros, bem como a maior variabilidade, é observada em setembro, levando à mesma hipótese, de que há maior disposição para a concepção no período de férias (janeiro) para registro em setembro, no momento do nascimento. O mês de março é o que apresenta a menor variabilidade (Figura 47). A Figura 48 apresenta a distribuição dos dados originais; sem a tendência; com o ajuste de sazonalidade, e; então, aplicado o modelo da equação (38), indicando que o modelo se ajusta bem à série. O vetor de resíduos da decomposição desta série varia em torno de zero, mas não apresenta distribuição próxima da normal (Figura 49).

Entre os modelos aqui apresentados, observamos uma diferença considerável do valor da estatística de erro MSE, que representa a real variabilidade sazonal dos dados ao longo dos meses. O modelo que considera tendência e sazonalidade apresenta-se melhor, pois os valores das medidas de avaliação dos resíduos são menores. Os valores de MAPE, MAD e MSE reduziram, respectivamente: 1,918; 2,494; 93,117.

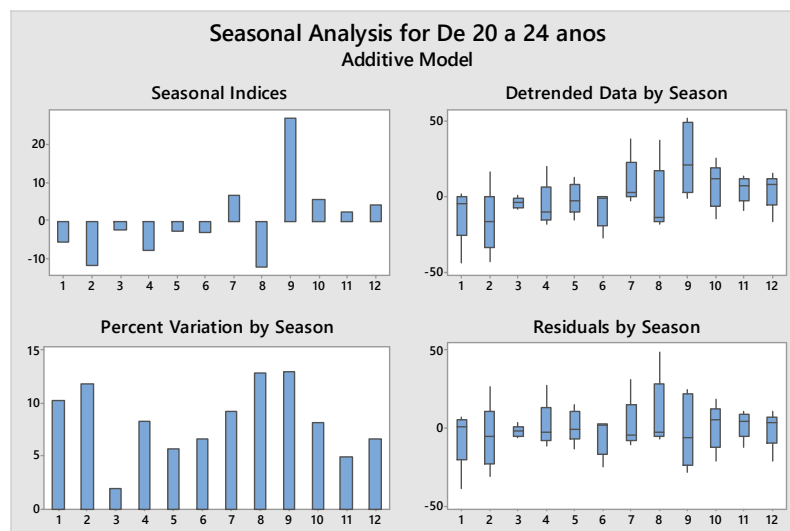


Figura 47 - Análise dos índices sazonais para a série de registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo ajustado: equação (40).

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

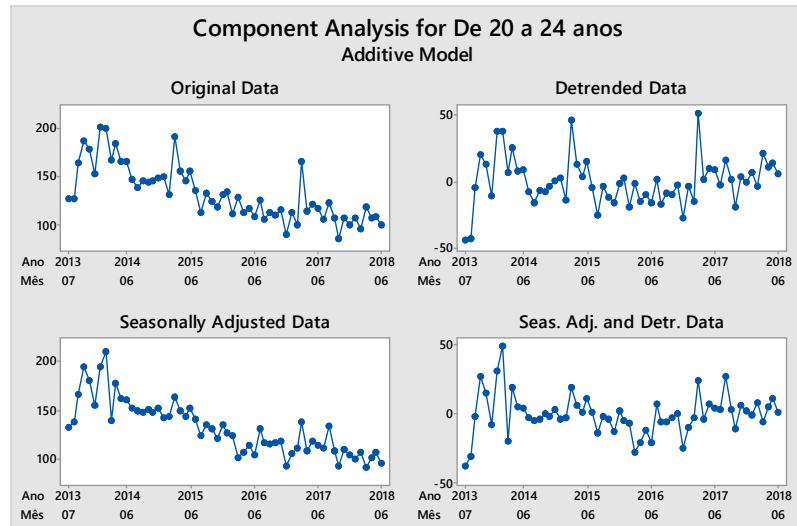


Figura 48 - Análise de sazonalidade para a série de registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo ajustado: equação (40).

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

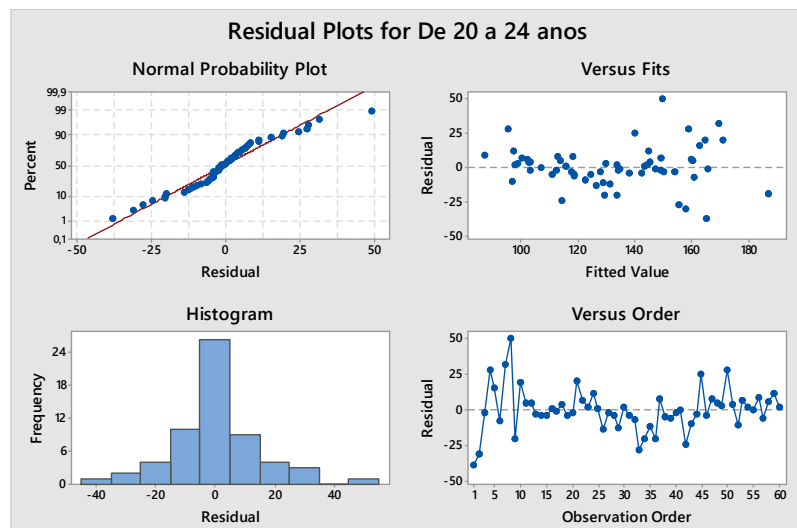


Figura 49 - Gráficos dos resíduos da análise de decomposição (tendência e sazonalidade) para a série de registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo ajustado: equação (40).

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

4.2.2 - Modelagem Através do Método ARIMA

Considerando que a série de registros de nascimentos de filhos de pais do grupo etário de 20 a 24 anos, apresenta uma tendência, podemos afirmar que a mesma é não estacionária. Após a aplicação de uma diferenciação simples ($d=1$), a série se torna estacionária. Os correlogramas da série diferenciada (Figura 50 e Figura 51) mostram que o comportamento das correlações sugere um modelo ARIMA (0,1,1) a ser ajustado aos dados da série.

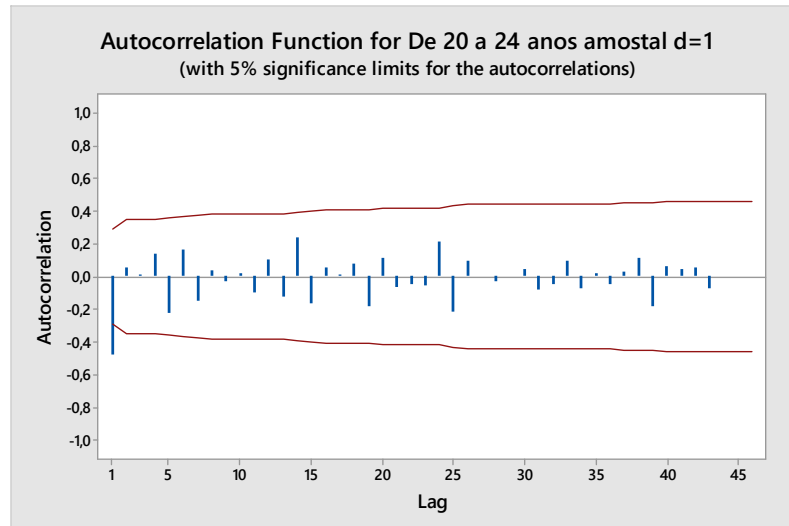


Figura 50 - Função de autocorrelação da série de registros de filhos de pais entre 20 e 24 anos.
Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

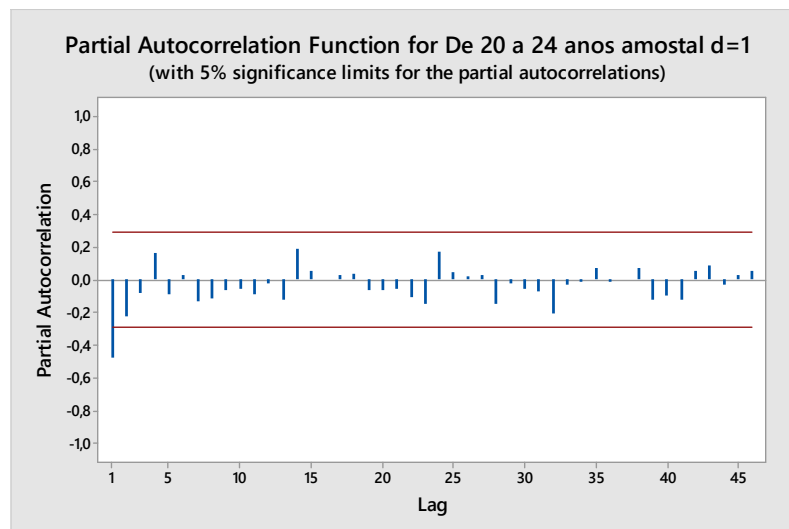


Figura 51 - Função de autocorrelação parcial da série registros de filhos de pais entre 20 e 24 anos.
Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

As estimativas dos parâmetros do modelo ARIMA (0, 1, 1) ajustado utilizando-se todos os meses do período (n=60 observações), foram obtidas após 8 interações e encontram-se no Quadro 13:

$$Y_t = Y_{t-1} + a_t - 0,5892a_{t-1}. \quad (41)$$

Quadro 13 - Estimativas e parâmetros do modelo ARIMA (0, 1, 1) da série de registros de filhos de pais entre 20 e 24 anos, n=60.

Estimates at each iteration					
Iteration	SSE	Parameters			
0	24222,6	0,100			
1	21698,4	0,250			
2	20087,2	0,400			
3	19369,5	0,550			
4	19339,7	0,576			
5	19336,5	0,584			
6	19336,1	0,588			
7	19336,0	0,589			
8	19336,0	0,589			
Relative change in each estimate less than 0,0010					
Final Estimates of Parameters					
Type	Coef	SE Coef	T	P	
MA 1	0,5892	0,1054	5,59	0,000	
Differencing: 1 regular difference					
Number of observations: Original series 60, after differencing 59					
Residuals: SS = 19073,7 (backforecasts excluded)					
MS = 328,9 DF = 58					
Modified Box-Pierce (Ljung-Box) Chi-Square statistic					
Lag	12	24	36	48	
Chi-Square	9,6	26,0	36,4	52,9	
DF	11	23	35	47	
P-Value	0,565	0,303	0,403	0,257	

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

O termo auto-regressivo tem um p-valor menor que o nível de significância de 0,05, indicando que o coeficiente é estatisticamente significativo e pode ser mantido no modelo. Os valores do correlogramas dos resíduos (Figura 52), a 5% de significância, mostram um comportamento similar ao de um ruído branco, pois não há nenhuma correlação nas defasagens próxima dos limites. O p-valor = 0,565 do teste da estatística qui-quadrado de Ljung-Box indica que há significância no conjunto das 12 primeiras autocorrelações dos resíduos (a 5% de significância). Os resíduos se distribuem ao redor de zero e com uma variância estimada igual a 328,9.

Os valores dos resíduos do modelo ARIMA (0, 1, 1) se distribuem em torno de zero e apresentam distribuição assimétrica (Figura 54).

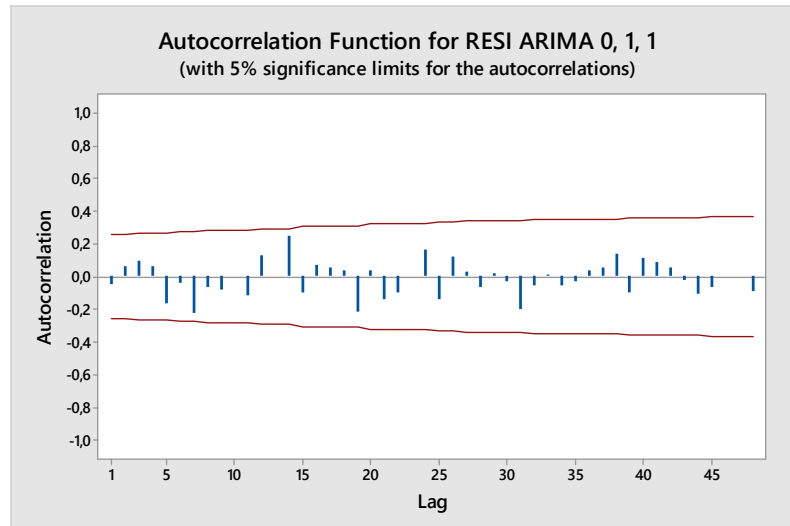


Figura 52 - Correlograma do vetor de resíduos do modelo ajustado para a série de registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo ajustado: equação (41), $n=60$.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

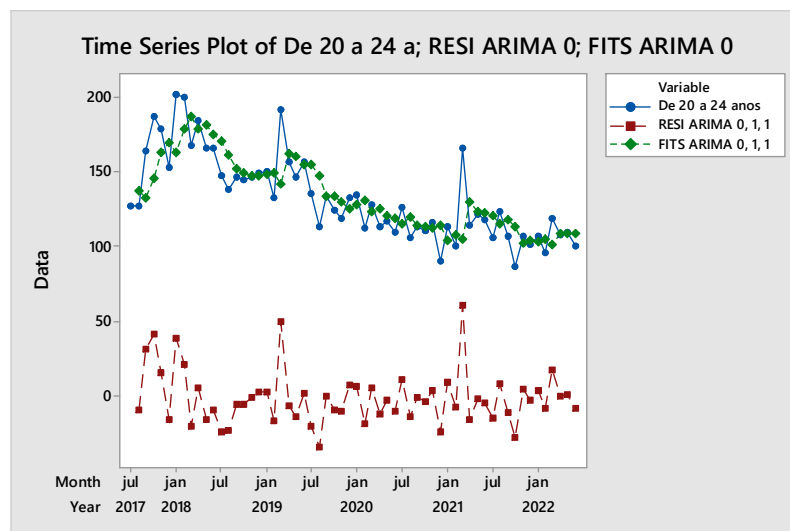


Figura 53 – Dados ajustados segundo o modelo ARIMA (0, 1, 1) para a série de registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo ajustado: equação (41), $n=60$.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

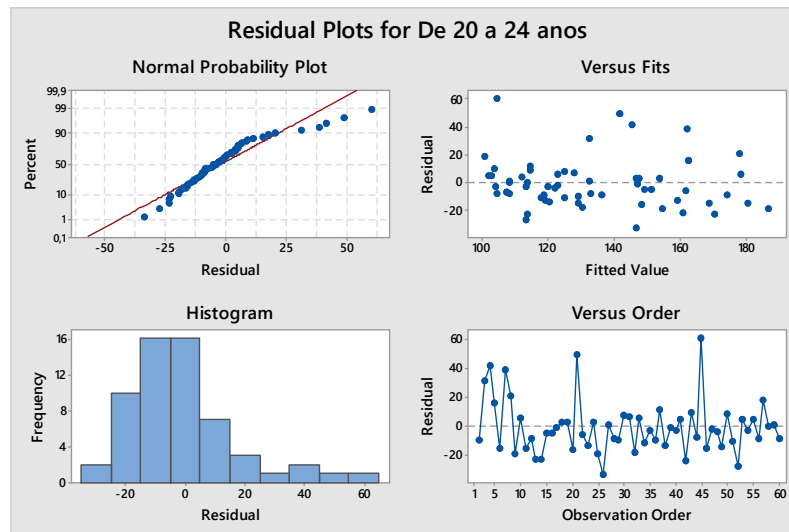


Figura 54 - Gráficos dos resíduos do modelo ARIMA (2, 1, 0) para a série de registros de filhos de pais entre 15 e 19 anos, n=60.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Uma vez verificado que o modelo ARIMA (0, 1, 1) se ajusta bem à série de dados de registros de filhos de pais entre 20 a 24 anos optou-se por realizar a previsão de valores futuros da série. A finalidade do ajuste deste modelo não é fazer uma comparação como o modelo resultante da equação (41), uma vez que aquele foi ajustado para todo o período (n=60 observações) e este apenas para o período amostral, que compreende ao período de julho de 2013 a junho de 2017 (n=48). Este modelo, permite, portanto que se faça as previsões de valores futuros para o período de um ano, considerando a origem de previsão no tempo $t=48$ (julho/2017) usando o modelo ARIMA (0, 1, 1).

A seguir a equação do modelo estimado para o ajuste de previsão da série observada:

$$Y_t = Y_{t-1} + a_t - 0,5801a_{t-1}. \quad (42)$$

O Quadro 8 nos mostra os resultados do ajuste do modelo ARIMA (0, 1, 1) com n=48 observações. Tais resultados mostram-se coerentes com os observados para o modelo ARIMA (0, 1, 1) com n=60 em termos de significância dos coeficientes ao nível de 5%. O p-valor do coeficiente de médias móveis é menor que o nível de significância de 0,05 indicando que o coeficiente de médias móveis é estatisticamente significativo e que pode ser mantido no modelo. Os p-valores da estatística qui-quadrado de Ljung-Box indicam que a suposição de resíduos independentes foi respeitada. Os valores dos resíduos se distribuem em torno de

média zero com variância igual a 378,5, portanto maior que o observado no modelo com o $n=60$, o que é natural, já que o tamanho da amostra é menor ($n=48$). O gráfico do correlograma dos resíduos do modelo de previsão ($n=48$) (Figura 35), mostra que a função de autocorrelação é semelhante ao que fora observado no modelo ajustado ao período completo ($n=60$).

Quadro 14 - Estimativas e parâmetros a partir do modelo ARIMA (0, 1, 1) da série de registros de filhos de pais entre 20 e 24 anos, $n=48$.

Estimates at each iteration					
Iteration	SSE	Parameters			
0	21962,1	0,100			
1	19709,6	0,250			
2	18284,5	0,400			
3	17674,6	0,550			
4	17658,3	0,570			
5	17656,6	0,577			
6	17656,3	0,579			
7	17656,3	0,580			
8	17656,3	0,580			

Relative change in each estimate less than 0,0010
 Final Estimates of Parameters

Type	Coef	SE Coef	T	P
MA 1	0,5801	0,1192	4,87	0,000

Differencing: 1 regular difference
 Number of observations: Original series 48, after differencing 47
 Residuals: SS = 17410,8 (backforecasts excluded)
 MS = 378,5 DF = 46

Modified Box-Pierce (Ljung-Box) Chi-Square statistic				
Lag	12	24	36	48
Chi-Square	6,9	19,5	26,3	*
DF	11	23	35	*
P-Value	0,804	0,674	0,854	*

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

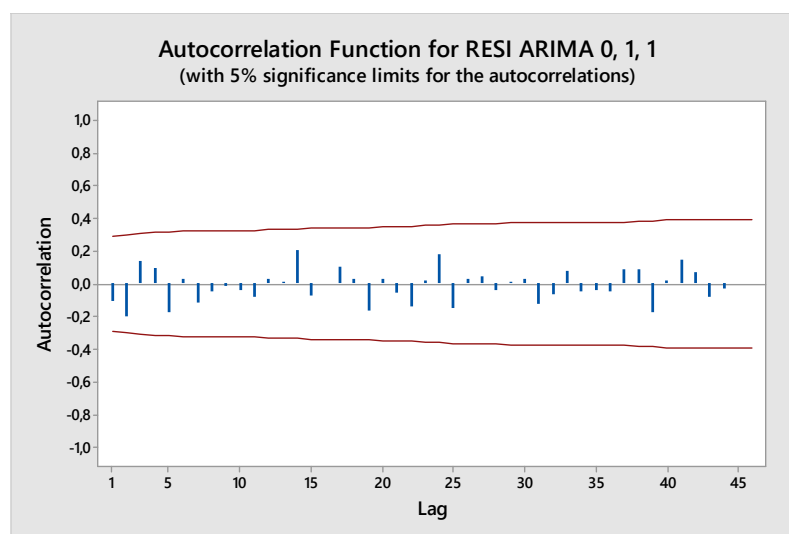


Figura 55 - Correlograma do vetor de resíduos do modelo ARIMA (0, 1, 1) da série de registros de filhos de pais entre 20 e 24 anos, $n=48$.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

No Quadro 15 é possível comparar os valores resultantes da previsão através da indicação de previsões (*Forecast*) com os valores observados no período de validação que correspondem ao observado nos registros reais (*Actual*). As previsões também podem ser vistas na Figura 56. Destaca-se que para ter um bom ajuste de previsão a partir de um modelo ARIMA, é desejável que o número de observações para o período amostral seja maior que 48.

Quadro 15 - Previsões, limites inferiores e superiores a partir do modelo ARIMA (0, 1, 1) e valores reais de registros de filhos de pais entre 20 e 24 anos, n=48.

Previsões com origem no período 48 (Forecasts from period 48)				
95% Limits				
Period	Forecast	Lower	Upper	Actual
49	120,481	82,341	158,620	106
50	120,481	79,115	161,846	123
51	120,481	76,123	164,838	107
52	120,481	73,320	167,641	86
53	120,481	70,675	170,286	107
54	120,481	68,163	172,798	101
55	120,481	65,767	175,194	107
56	120,481	63,471	177,490	96
57	120,481	61,264	179,697	119
58	120,481	59,137	181,825	108
59	120,481	57,081	183,881	109
60	120,481	55,089	185,872	100

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

A Figura 56 mostra as previsões para os dados reais para o período de validação, onde se observa que o número de registros real está dentro dos limites da estimação intervalar (95% de confiança). As previsões pontuais estão em torno de 120 registros por mês, que se estão acima do número de registros do período de validação, exceto nos meses de agosto e de março. Observa-se grande amplitude dos intervalos de confiança especialmente para os meses mais distantes da origem (t=48, julho/2017).

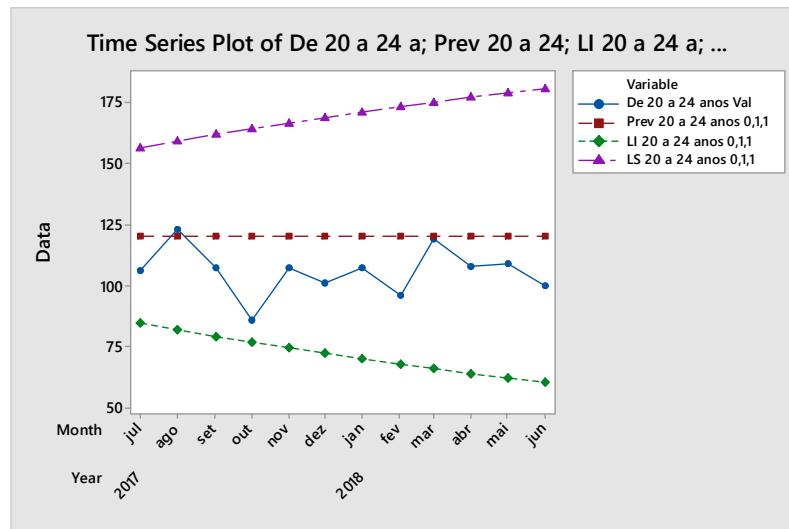


Figura 56 - Previsão para a série de registros de filhos de pais entre 20 e 24 anos, modelo ARIMA (0, 1, 1), n=48.
Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

4.2.3 - Modelagem Através do Método de Alisamento Exponencial – Holt

Ao observar que a série de registros de nascimentos no grupo etário de pais entre 20 e 24 anos apresenta tendência optou-se por aplicar o método de suavização exponencial de Holt (ver item 2.3), em que se determina as constantes de suavização α e γ .

As equações que resultam na suavização da tendência e do nível da série são dadas por:

$$L_t = 0,5449Y_t + 0,4550[L_{t-1} + Y_t], \quad (43)$$

$$T_t = 0,0370[L_t - L_{t-1}] + 0,963T_{t-1}. \quad (44)$$

Onde L_t nos dá a suavização do nível da série e T_t , da tendência da série.

A Figura 57 apresenta a série de registros de nascimentos no grupo etário de pais entre 20 e 24 anos real e a suavizada (*Smoothed*) após o ajuste do modelo Holt n=60 que se ajustam bem. Os valores dos resíduos do modelo Holt variam em torno de zero com registro de alguns valores acima de 40, embora não apresentem uma distribuição próxima da normal (Figura 58).

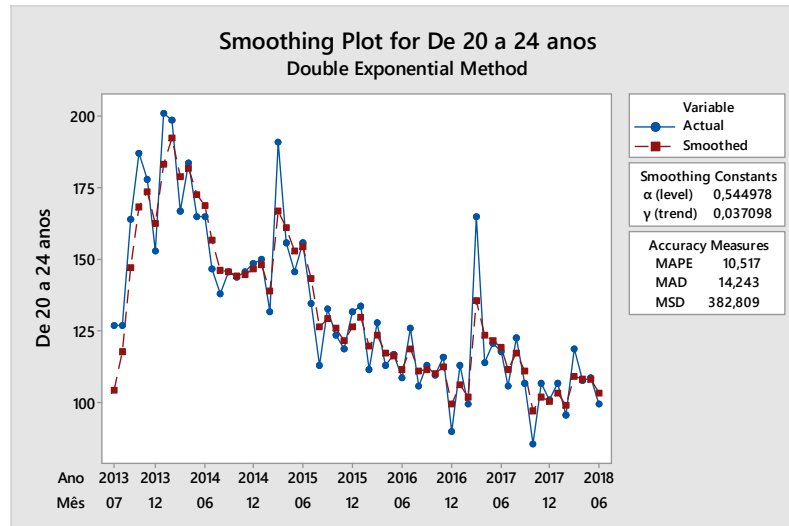


Figura 57 - Análise de suavização (nível e tendência) para a série de registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo ajustado: equações (43 e 44), $n=60$.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

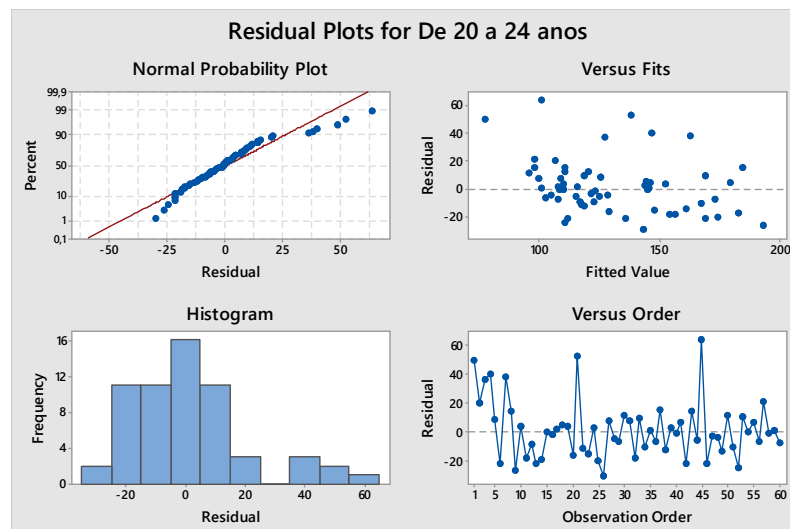


Figura 58 - Gráficos dos resíduos do modelo Holt para a a série de registros de filhos de pais entre 15 e 19 anos - dados mensais do modelo ajustado: equações (42 e 44), $n=60$.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Como se observa que os dados ajustados ao modelo Holt se ajustam bem à série real de registros de nascimentos no grupo etário de pais entre 20 e 24 anos optou-se por realizar a previsão de valores futuros da série utilizando este modelo. A finalidade do ajuste deste modelo não é fazer uma comparação como o modelo resultante das equações (42 e 44), uma vez que aquele fora ajustado para todo o período ($n=60$ observações) e o que se segue, apenas para o período amostral, que compreende ao período de julho de 2013 a junho de 2017 ($n=48$). Este modelo, permite, portanto que se faça as previsões de valores futuros para o período de um ano, considerando a origem de previsão no tempo $t=48$ (julho/2017).

A seguir as equações do modelo estimado para o ajuste de previsão da série observada bem como os valores previstos, os limites e os valores reais (Quadro 16):

$$L_t = 0,4711Y_t + 0,5289[L_{t-1} + Y_t]; \quad (45)$$

$$T_t = 0,0755[L_t - L_{t-1}] + 0,9245T_{t-1}. \quad (46)$$

Quadro 16 - Previsões, limites inferiores e superiores a partir do modelo Holt e valores reais de registros de filhos de pais entre 20 e 24 anos, n=48.

Length 48				
Smoothing Constants				
Alpha (level) 0,471118				
Gamma (trend) 0,075486				
Forecasts				
Period	Forecast	Lower	Upper	Actual
49	120,099	83,1662	157,033	106
50	119,877	78,4941	161,261	123
51	119,655	73,3544	165,956	107
52	119,433	67,8807	170,986	86
53	119,211	62,1652	176,258	107
54	118,990	56,2714	181,708	101
55	118,768	50,2436	187,291	107
56	118,546	44,1130	192,978	96
57	118,324	37,9024	198,745	119
58	118,102	31,6284	204,575	108
59	117,880	25,3034	210,456	109
60	117,658	18,9369	216,378	100

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

A Figura 59 mostra os dados reais e os dados suavizados para o período amostral. A Figura 60 mostra que os dados reais, correspondentes ao período de validação, estão dentro da estimativa intervalar (95% de confiança). As estimativas pontuais estão variando acima do observado no período de validação, exceto no mês de março. Observa-se grande amplitude dos intervalos de confiança especialmente para os meses mais distantes da origem (julho/2017).

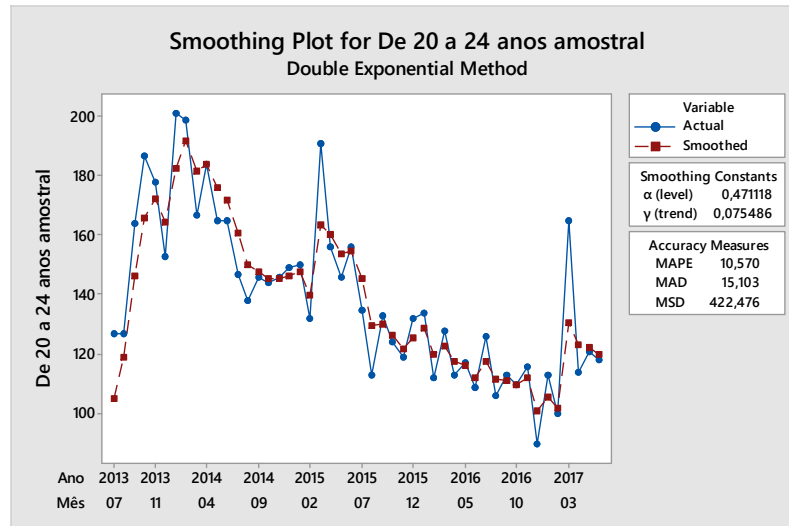


Figura 59 - Análise de suavização (nível e tendência) para a série de registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo ajustado: equações (45 e 46), $n=48$.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

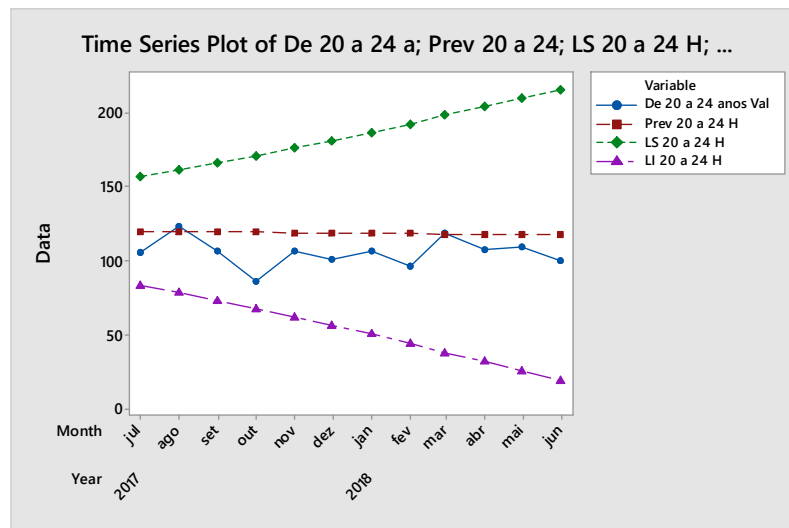


Figura 60 - Previsão para a série de registros de filhos de pais entre 20 e 24 anos, modelo Holt, $n=48$.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

4.2.4 - Modelagem Através do Método de Alisamento Exponencial – Holt-Winters

Considerando a sazonalidade observada no item 4.2.1, optou-se por aplicar o método de Holt-Winter, que considera, além do nível e da tendência, a sazonalidade (ver item 2.3), determinando, então, as constantes de suavização para o nível, para a tendência e para a sazonalidade da série (α , γ e δ).

As equações que resultam na suavização da tendência, do nível e da tendência da série são dadas por:

$$L_t = 0,2(Y_t - S_{t-s}) + 0,8(L_t - T_{t-1}); \quad (47)$$

$$T_t = 0,2(L_t - L_{t-1}) + 0,8T_{t-1}; \quad (48)$$

$$S_t = 0,2(Y_t - L_t) + 0,8S_{t-s}. \quad (49)$$

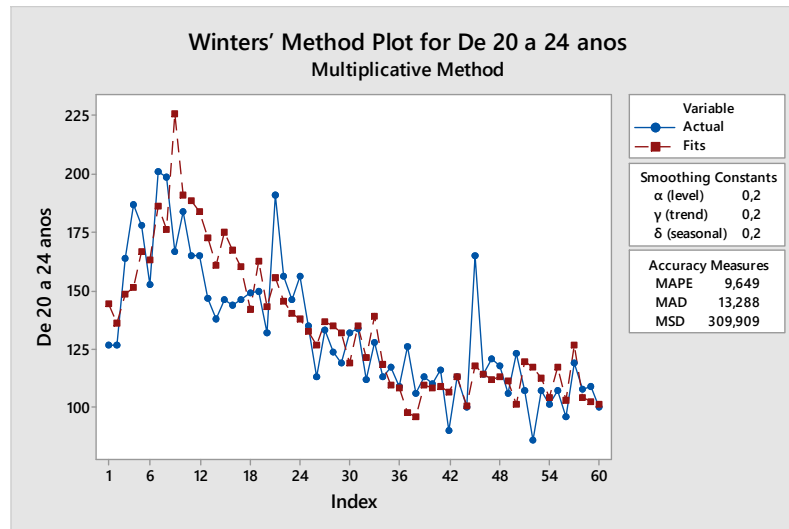


Figura 61 - Análise de suavização (nível, tendência e sazonalidade) para a série de registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo ajustado: equações (47, 48 e 49), $n=60$.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Como se observa que os dados ajustados ao modelo Holt-Winter se ajustam bem à série real de registros de nascimentos no grupo etário de pais entre 20 e 24 anos optou-se por realizar a previsão de valores futuros da série utilizando este modelo. A finalidade do ajuste deste modelo não é fazer uma comparação como o modelo resultante das equações (47, 48 e 49), uma vez que aquele fora ajustado para todo o período ($n=60$ observações) e o que se segue, apenas para o período amostral, que compreende ao período de julho de 2013 a junho de 2017 ($n=48$). Este modelo, permite, portanto que se faça as previsões de valores futuros para o período de um ano, considerando a origem de previsão no tempo $t=48$ (julho/2017). A seguir o modelo estimado para o ajuste de previsão da série bem como os valores previstos, os limites e os valores reais (Quadro 17):

Quadro 17 - Previsões, limites inferiores e superiores a partir do modelo Holt-Winter $n=48$ e valores reais de registros de filhos de pais entre 20 e 24 anos.

Additive Method				
Data	De 20 a 24 anos			
Length	48			
Smoothing Constants				
Alpha (level)	0,2			
Gamma (trend)	0,2			
Delta (seasonal)	0,2			
Forecasts				
Period	Forecast	Lower	Upper	Actual
49	104,419	70,153	138,685	
50	91,139	56,336	125,941	
51	109,029	73,628	144,430	
52	111,595	75,536	147,653	
53	112,132	75,361	148,904	
54	104,802	67,266	142,338	
55	124,205	85,854	162,555	
56	111,970	72,759	151,181	
57	143,599	103,484	183,714	
58	121,477	80,418	162,535	
59	119,811	77,771	161,852	
60	120,698	77,641	163,755	

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

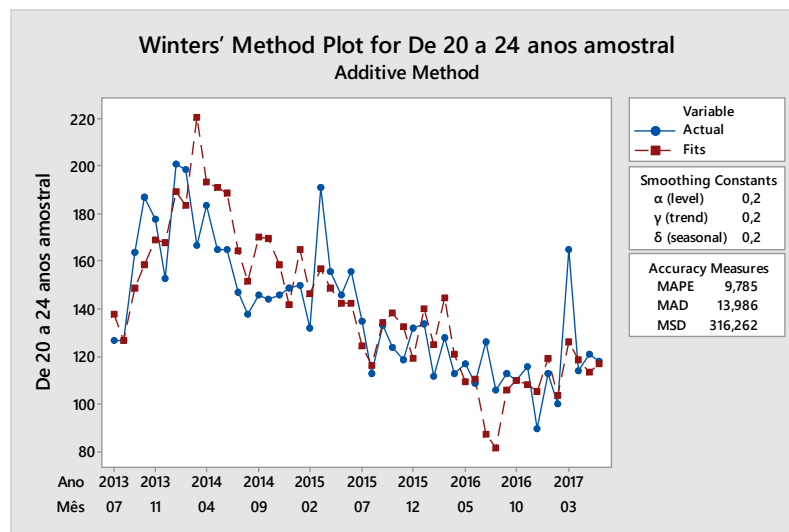


Figura 62 - Análise de suavização (nível, tendência e sazonalidade) para a série de registros de filhos de pais entre 20 e 24 anos - dados mensais do modelo ajustado: equações (40, 41 e 42), $n=48$.

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

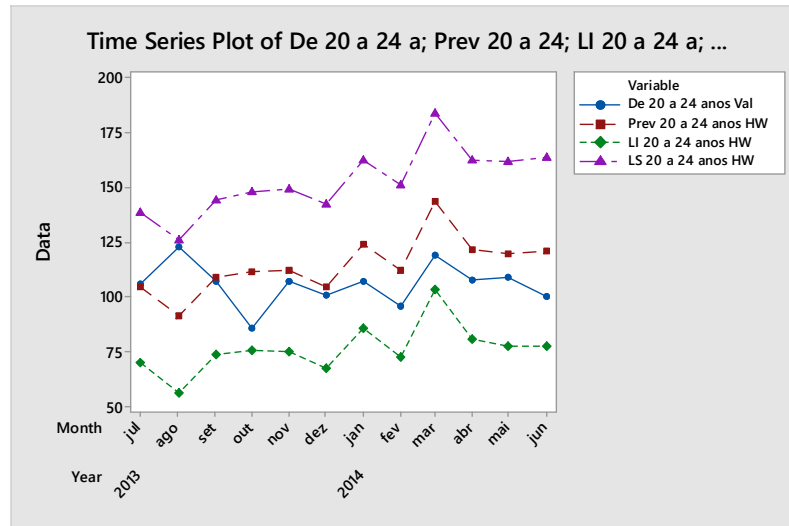


Figura 63 - Previsão para a série de registros de nascimentos de filhos de pais entre 20 e 24 anos, modelo Holt-Winter, $n=48$.
 Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

4.3 - Comparação e Análise dos Modelos

Como estabelecido no capítulo 2, seção 2.4, o procedimento de análise dos modelos ajustados no capítulo 4 será o de análise das estatísticas de erros. Desta forma, compreendendo que o objetivo foi avaliar as previsões realizadas para cada modelo, optou-se por avaliar apenas as estatísticas dos erros dos modelos de previsão, portanto, com o $n=48$.

Vale destacar que o modelo de decomposição, portanto, não possibilitou a realização de previsão, e por isso não consta na comparação abaixo.

As estatísticas dos erros devem ser as menores possíveis, considerando, portanto, que os modelos para os ajustes dos dados (amostral) e para as previsões (validação) grantam as melhores aplicações.

A partir da análise das estatísticas dos erros, é possível observar que o melhor modelo ajustado à série de 15 a 19 anos, para o período amostral é o modelo Holt-Winter, que considera então, além nível, a tendência e a sazonalidade da série. Já para o período de validação, o modelo Holt, pois apresenta os menores valores para a maioria das estatísticas de erros (MAPE, MAE, MSE). O MAPE e o MAE do modelo Holt-Winter caem, respectivamente, 23% e 19% em relação ao modelo Holt, que apresenta os maiores valores destas estatísticas no período amostral. Já o MSE, cai 38% em relação ao modelo ARIMA 2, 1, 0 no mesmo período. No período de validação, ou seja, aquele que avalia as previsões realizadas com a série de

dados originais, o modelo Holt apresentou-se melhor em todas as estatísticas, tendo sido a maior diferença observada no MSE deste modelo em relação ao modelo Holt-Winter, que apresentou o maior valor. A redução representou uma queda de mais da metade (52%) no valor desta estatística. Já o MSE, o MAPE e o MAE reduziram, respectivamente 47%; 28% 29% em relação às mesmas estatísticas do modelo ARIMA (2, 1, 0) que apresentaram os maiores valores.

Para a série de 20 a 24 anos, o mesmo resultado fica evidente através das estatísticas de erros no período amostral, em que o melhor modelo é o Holt-Winter, que apresentou a maior redução no MSE, resultando em queda de 56% no valor da estatística apresentada no modelo Holt-Winter em relação ao modelo ARIMA (0, 1, 1) que apresentou as maiores estatísticas. No período de validação, ambos os modelos (Holt- e Holt-Winter) apresentaram as menores estatísticas para dois erros. De tal forma que a queda observada nas estatísticas MSE e MAPE do modelo Holt-Winter que apresentou os menores valores foram, respectivamente de 38% e de 9%, em relação às estatísticas apresentadas no modelo ARIMA (0, 1, 1) que apresentou os maiores valores. Já os erros MAE e MSE apresentaram queda de 9% e 15% respectivamente, em relação aos valores observados também no modelo ARIMA (0, 1, 1), portanto, menores quedas em relação às estatísticas MSE e MAPE, que podem qualificar este modelo (Holt-Winter) como o melhor para a realização de previsões.

O Quadro 18 e o Quadro 19 apresentam as estatísticas dos erros para as séries de registros nascimentos dos dois grupos etários avaliados.

Quadro 18 - Avaliação dos modelos segundo as estatísticas dos erros para a série de registros de filhos de pais entre 15 e 19 anos – períodos amostral e de validação.

Estatísticas	ARIMA 2, 1, 0	AED-Holt	AED-HWinters
Erros de Previsão no período amostral			
Erro Percentual Médio – MPE	-2,1059	-5,3882	-3,017
Erro Percentual Absoluto Médio - MAPE	19,7395	20,1480	15,5572
Erro Absoluto Médio – MAE	6,3087	6,3205	5,0996
Erro Quadrado Médio – MSE	65,7718	61,6991	40,3829
Erros de Previsão no período de validação			
Erro Percentual Médio – MPE	-13,8938	-7,3165	-9,277
Erro Percentual Absoluto Médio MAPE	15,4507	10,9957	14,908
Erro Absoluto Médio – MAE	3,5401	2,5069	3,493
Erro Quadrado Médio – MSE	16,1039	10,6998	19,148

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Quadro 19 - Avaliação dos modelos segundo as estatísticas dos erros para a série de registros de filhos de pais entre 20 e 24 anos – períodos amostral e de validação.

Estatísticas	ARIMA 0, 1, 1	AED-Holt	AED-HWinters
Erros de Previsão no período amostral			
Erro Percentual Médio – MPE	0,68042	0,0086	-1,3979
Erro Percentual Absoluto Médio - MAPE	11,7971	10,5698	9,7847
Erro Absoluto Médio – MAE	16,6088	15,1028	13,9863
Erro Quadrado Médio – MSE	723,343	422,4756	316,2624
Erros de Previsão no período de validação			
Erro Percentual Médio – MPE	-14,8424	-13,1946	-9,1087
Erro Percentual Absoluto Médio MAPE	15,18374	13,7472	13,6745
Erro Absoluto Médio – MAE	15,15083	13,6791	14,3966
Erro Quadrado Médio – MSE	303,1899	254,9913	300,0571

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

5 - Considerações Finais

De acordo com a distribuição dos dados, foi possível observar que os registros civis segundo sexo é coerente com o observado na distribuição da população, na qual se observa, a partir dos dados do SINASC, no qual a maioria da população de nascidos vivos é masculina. A maior concentração dos registros aqui observados ocorre entre os pais com idades de 25 a 34 anos, com a mesma tendência, também, observada entre as mulheres, em que a barriga da curva se dá entre as idades mais jovens. Quanto ao local de nascimento, a maior ocorrência se dá em hospitais e nos demais lugares (domicílio, trânsito, via pública, UPA e centro de saúde) os registros do cartório civil não passam de 0,5%.

A distribuição dos dados por ano revela uma queda observada entre os anos de 2014 a 2017, dentre os quais, o que apresentou a maior queda relativa foi o ano de 2016, em relação ao total observado em 2015 (-16,04%). O ano, entre aqueles com registros em todos os meses, que teve o maior número de registro foi o ano de 2014, com 9754 registros de nascimento.

A média de registros por ano foi igual a 6946 casos. O máximo observado foi de 9754 e o mínimo, de 3540. A menor variabilidade foi observada no ano de 2014, quando se registrou o maior número de registros. Quanto aos meses, a média de registros observada foi de 694. Os meses com a maior variabilidade foram janeiro, fevereiro e outubro. Observou-se uma variação de registros ao longo dos meses, mas não foi possível constatar, na literatura as razões para tal, sendo passível, portanto, de ampliação de estudos para compreensão da distribuição de registros civis ao longo ano e de suas variações entre os meses.

Quanto à idade, a média observada entre os pais que registram seus filhos nesta unidade de cartório foi de 30 anos e se manteve ao longo de todos os anos observados. A idade dos pais variou entre 10 e 82 anos e possui distribuição assimétrica positiva. Em comparação com os dados do SINASC, a média de idade dos pais é menor oito anos, indicando que as mulheres têm filhos mais jovens que os homens. A média de idade dos pais se mantém em torno dos 30 anos ao longo de todos os anos observados.

A distribuição de registros por grupo etário de interesse revela uma maior quantidade de registros entre os pais com idade entre 20 e 24 anos. No entanto, para ambos os grupos (15 a 19 e 20 a 24) observou-se queda ao longo dos anos. Embora a maior queda, em termos absolutos tenha sido observada entre o grupo mais jovem (de 15 a 19 anos), a partir da distribuição de dados pelo gráfico de série temporal, é no grupo de maior idade a maior evidência a esta tendência. Em ambos os grupos não foi observado distribuição sazonal, cíclica ou estacionária. Por esta razão é possível supor que o fenômeno observado (registros de nascimentos de filhos de pais entre 15 e 24 anos) pode ter sofrido redução, passível de verificação devido ao intervalo de tempo entre este estudo e os dados.

Ao avaliar a tendência, para ambos os grupos etários, o modelo que considera a tendência e a sazonalidade apresenta-se melhor, sobretudo para o grupo de pais entre 20 e 24 anos. As observações do período de validação estão dentro da estimativa intervalar, para ambos grupos. Assim como temos observado para a fecundidade feminina, é possível que a fecundidade masculina, observada aqui através dos registros civis, também passe por queda.

Entre os modelos, o que mais se ajustou às séries de dados foi distinto segundo grupo e período de análise. Para o grupo mais velho (de 20 a 24 anos) o melhor modelo é o de alisamento exponencial duplo, no qual se aplicou o modelo Holt-Winter, no caso do período amostral. Já no caso do período de validação deste grupo, o melhor modelo compreende o alisamento exponencial duplo Holt. O mesmo se observa para o grupo mais jovem: o melhor modelo no período de análise se distingue do modelo no período de validação. Neste sentido, o melhor modelo para o ajuste de dados de registros civis de filhos dos grupos etários de 15 a 19 anos e de 20 a 24 anos compreende-se o modelo Holt-Winter, que considera, além do nível da

série, a tendência e a sazonalidade. Já para a realização de predições, o modelo que considera o nível e a tendência (Holt) se adequa melhor.

Em relação aos estudos da fecundidade masculina, faz-se necessário ampliá-los com o objetivo de avaliar como se dão os registros de fecundidade masculina em relação ao da fecundidade feminina. Além disso, entre os grupos de idade extremos é interessante, também, verificar, se em ambas as extremidades, a distribuição apresenta queda além de discutir as implicações de fecundidade em grupos extremos, tais como, a influência na trajetória de vida, como na transição para a vida adulta, no caso dos jovens; ou a transferência intergeracional, no caso dos mais velhos.

Destaca-se a importância de aplicação de um conjunto maior de dados de mais dados para obtenção de melhores ajustes de modelos. Seria de importante contribuição a colaboração de mais unidades de cartório o que permitiria a análise de dados sobre paternidade no município de Belo Horizonte, para além de uma única regional administrativa.

Por fim, ressalta-se a possibilidade de aplicação de tal estudo, uma vez com dados atualizados, no planejamento de serviços não somente na estrutura do (s) cartório (s), em termos de planejamento e adequação de rotinas e de recursos, mas também, no planejamento de políticas públicas setoriais, direcionadas ao público etário estudado bem como, ao público de nascidos/registrados, no que tange à campanhas de vacinação; envolvimento paterno nos cuidados diários com o (s) filho (s); e outras.

Referências Bibliográficas

1. ALMEIDA, M. C. C. *et al.* School trajectory and teenage pregnancy in three Brazilian state capitals. *Cadernos de Saúde Pública*. Rio de Janeiro, 22 (7), p. 1397-1409, jul, 2006.
2. BASSI, C. Exposição à maternidade precoce e estratos sociais das adolescentes brasileiras: justificativas via determinantes próximos das taxas de fecundidade. Instituto de Pesquisa Econômica Aplicada. Brasília: 2008.
3. Belo Horizonte, 2018 - <https://prefeitura.pbh.gov.br/venda-nova> - acesso em 10/12/18).
4. BRANDÃO, E. R. Gravidez na adolescência: um balanço bibliográfico. In: HEILBORN, M. L. *et al.* (ORGs). O aprendizado da sexualidade: reprodução e trajetórias sociais de jovens brasileiros. Rio de Janeiro, Garamond e Fiocruz, 2006.
5. BOX, G. E. P., JENKINS, G. M. *Time Series Analysis: Forecasting and Control*. Holden-Day, 1970.
6. BOX, G. E. P., JENKINS, G. M. e REINSEL, G. *Time Series Analysis: Forecasting and Control*. Englewood Cliffs: Prentice Hall, 1994.
7. CABRAL, C. S. Gravidez na adolescência: Negociações na família. In: HEILBORN, M. L. *et al.* Sexualidade, família e ethos religioso. Rio de Janeiro: Gramond, 2005.
8. CAMARANO, A. A. Fecundidade e anticoncepção da população jovem. In: *Jovens acontecendo na trilha das políticas públicas*. Brasília: CNPD, 1998. v.1, p. 109-133.
9. CHACHAM, A. S. *et al.* Autonomia, gênero e gravidez na adolescência: uma análise comparativa da experiência de adolescentes e mulheres jovens provenientes de camadas médias e populares em Belo Horizonte. *Revista Brasileira de Estudos de População*. V. 29, n. 2, p. 389-407, jul./dez. 2012.
10. CHAGAS, A. C. M. A. Ao encontro do infinito curto: maternidade, escola e trabalho entre jovens de baixa renda. Dissertação de Mestrado em Ciências Sociais. Belo Horizonte, PUC Minas, 2012.
11. FONTOURA, N. O. e PINEIRO, L. S. Síndrome de juno: gravidez, juventude e políticas públicas. In: CASTRO, J. A. *et al.* (Orgs). *Juventude e políticas sociais no Brasil*. Brasília, IPEA, 2009.
12. FRANCO, Carla Gomes. *Agora é com os pais! [manuscrito]: fecundidade na adolescência e juventude na perspectiva masculina em Belo Horizonte e Região Metropolitana*. Dissertação de Mestrado em Demografia. Belo Horizonte, UFMG/FACE/Cedeplar, 2016.
13. GRANGER, C. W. J. e NEWBOLD, J. P. Forecasting Transformed Series. *Journal of the Royal Statistical Society. Series B*, 38, 189-203, 1976.
14. HEILBORN, M. L. (Org.) *Sexualidade: o olhar das ciências sociais*. Rio de Janeiro: Jorge Zahar Editor, 1999.
15. LEVANDOWSKI, D. C. Paternidade na adolescência: uma breve revisão da literatura internacional. *Estudos de Psicologia*. 6(2), 195-209, 2001.
16. LONGO, L. A. F. de B. Juventude e contracepção: um estudo dos fatores que influenciam o comportamento contraceptivo das jovens brasileiras de 15 a 24 anos. *Revista Brasileira de Estudos de População*. V. 19, n. 2, p. 229-247, jul./dez. 2002.

17. LOYOLA, M. A. A sexualidade como objeto de estudo das ciências humanas. In: HEILBORN, M. L. (Org.) Sexualidade: o olhar das ciências sociais. Rio de Janeiro: Jorge Zahar Editor, 1999.
18. LYRA, J. Paternidade adolescente: uma proposta de intervenção. Dissertação de Mestrado. Pontifícia Universidade Católica de São Paulo. São Paulo, 1997.
19. MAKRIDAKIS, S. e HILBON, M. Accuracy of Forecasting: Na empirical investigation. *Journal of the Royal Statistical Society. Series A*, 142, 97-145, 1979.
20. MEDRADO, B. e LYRA, J. A adolescência “desprevenida” e a paternidade na adolescência: uma abordagem geracional e de gênero. In: SCHOR, N. MOTA, M. S. F. T. e BRANCO, V. C. (Orgs), *Cadernos de Juventude, saúde e desenvolvimento*. Brasília. Ministério da Saúde, Secretaria de Políticas de Saúde, V. 1, p. 230-248, 1999.
21. MINISTÉRIO DA SAÚDE. Secretaria de Vigilância Sanitária. Sistema de Informações Sobre Nascidos Vivos – SINASC.
22. MORETTIN, Pedro A., TOLOI, Clélia M. C. *Análise de séries temporais*. 2ª ed. São Paulo, Egard Blucher, 2006.
23. NELSON, H. L. *The Use of Box-Cox Transformations in Economic Time Series Analysis: Na empirical study*. Doctoral Dissertation, Univ. of California, San Diego, 1976.
24. OLIVEIRA, M.C., BILAC, E.D., MUSZTAT, M. Homens e anticoncepção: um estudo sobre duas gerações masculinas das “camadas médias” paulistas. In: MIRANDA-RIBEIRO, P. e SIMÃO, A. B. (Orgs). *Qualificando os números: estudos sobre saúde sexual e reprodutiva no Brasil*. 2ª ed. Belo Horizonte: ABEP: UNFPA, 2009.
25. PANTELIDES, E. A. Aspectos sociales del embarazo y la fecundidade adolescente em América Latina. *Notas de Población* 78: 7-33, 2004.
26. PRESTON, S. H. *et al.* *Demography: measuring and modeling population processes*. Blackwell Publisher, Massachusetts, 2001.
27. RODRIGUEZ, J. Reproducción en la adolescencia: el caso de Chile y sus implicaciones de política. *Revista de la CEPAL*. 86, Ago, 2005.
28. TRINDADE, Z. A. e MENANDRO, M. C. S. Pais Adolescentes: vivência e significação. *Estudos de Psicologia*. 7(1), p.15-23, 2002.
29. UTIAMADA, M. R. P. A paternidade na adolescência: um estudo a partir da visão dos pais adolescentes do ambulatório de pré-natal do Hospital de Clínicas de Londrina. *Anais do I Simpósio sobre Estudos de gênero e Políticas Públicas*. Universidade Estadual de Londrina. 2010.
30. WANKE, P.; JULIANELLI, L. *Previsão de vendas: processos organizacionais e métodos quantitativos e qualitativos*. São Paulo: Atlas, 2006.
31. WATKINS, S. C. If All We Knew About Women was What We Read in Demography, What Would We Know? *Demography*. Vol. 30, Nº. 4, p. 551-577, Nov., 1993.
32. ZHANG, L. Male Fertility and Determinants. *The Springer Series on Demographic Methods and Population Analysis* 27. 2011. p. 201.

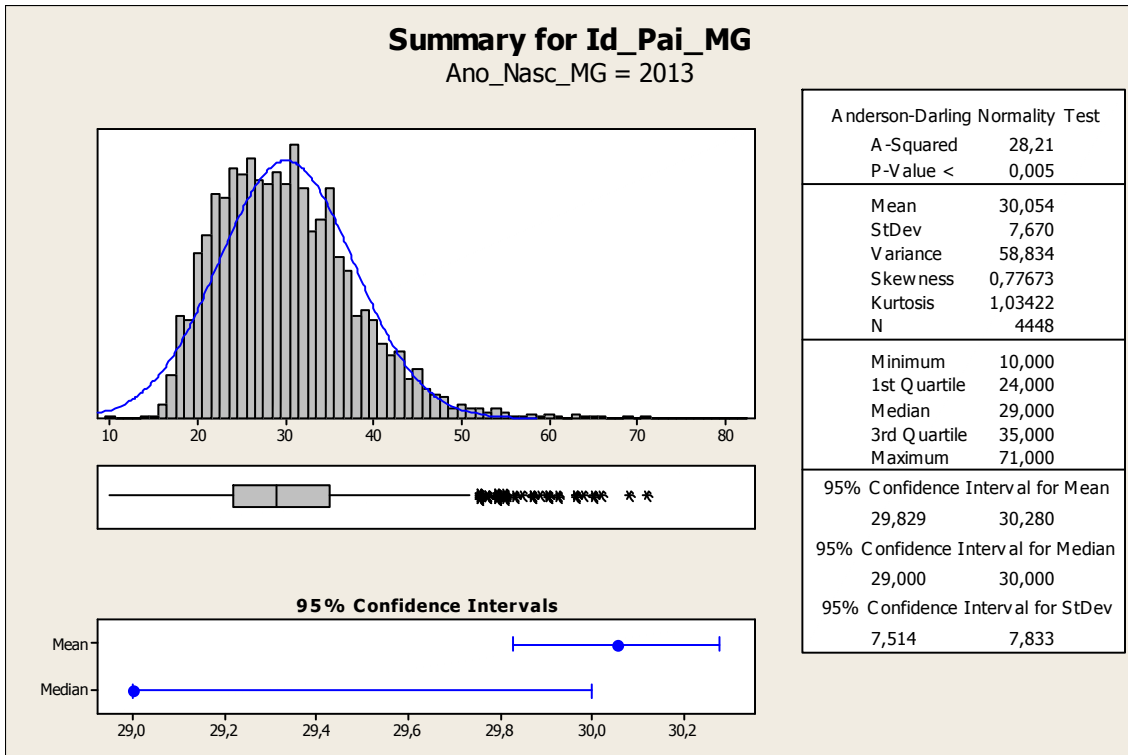
Anexo 1 - Distribuição de registros por idade do pai

Id_Pai_MG	Count	Percent
10	3	0,01
14	4	0,01
15	10	0,02
16	117	0,28
17	307	0,74
18	600	1,44
19	928	2,23
20	1287	3,09
21	1445	3,47
22	1698	4,07
23	1737	4,17
24	1781	4,27
25	1903	4,57
26	1972	4,73
27	2052	4,92
28	2064	4,95
29	2075	4,98
30	1978	4,75
31	2024	4,86
32	2027	4,86
33	1898	4,55
34	1807	4,34
35	1699	4,08
36	1664	3,99
37	1354	3,25
38	1203	2,89
39	990	2,38
40	899	2,16
41	718	1,72
42	579	1,39
43	515	1,24
44	418	1
45	357	0,86
46	275	0,66
47	219	0,53
48	170	0,41
49	147	0,35
50	119	0,29
51	111	0,27
52	78	0,19
53	84	0,2
54	76	0,18

Id_Pai_MG	Count	Percent
55	50	0,12
56	40	0,1
57	38	0,09
58	28	0,07
59	25	0,06
60	22	0,05
61	22	0,05
62	8	0,02
63	12	0,03
64	7	0,02
65	11	0,03
66	6	0,01
67	3	0,01
68	1	0
69	3	0,01
70	1	0
71	2	0
73	1	0
77	1	0
82	2	0
Total	41675	100

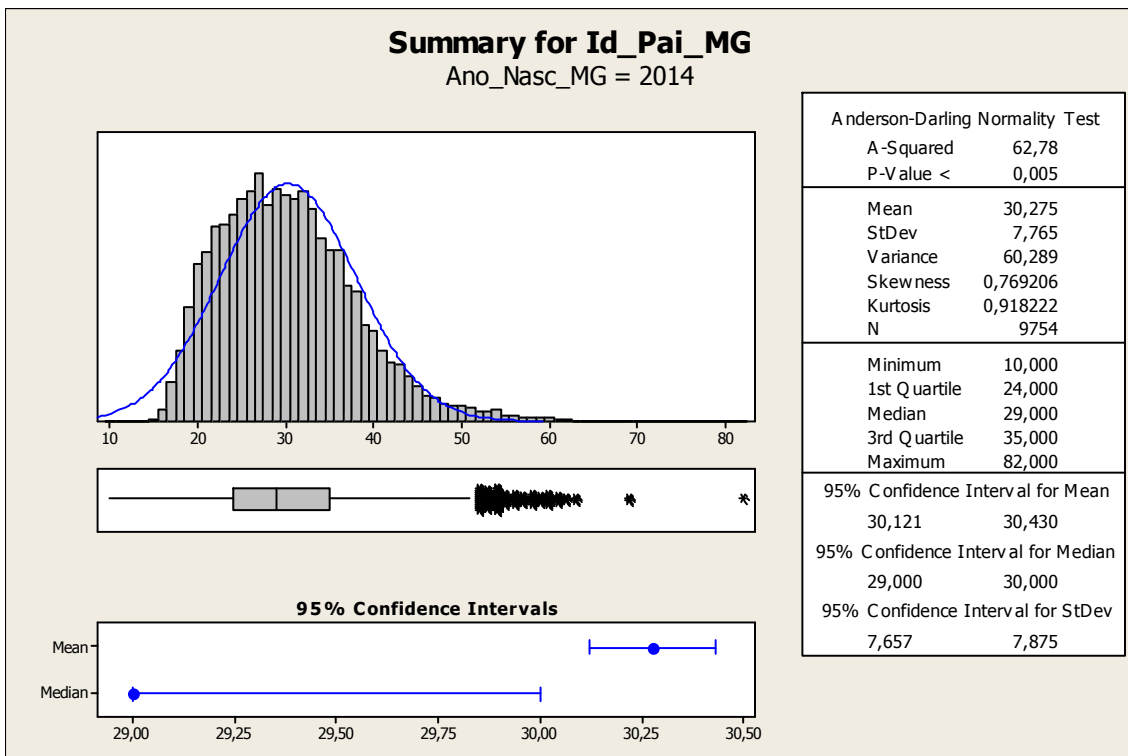
Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.

Anexo 2 - Distribuição das idades dos pais por ano



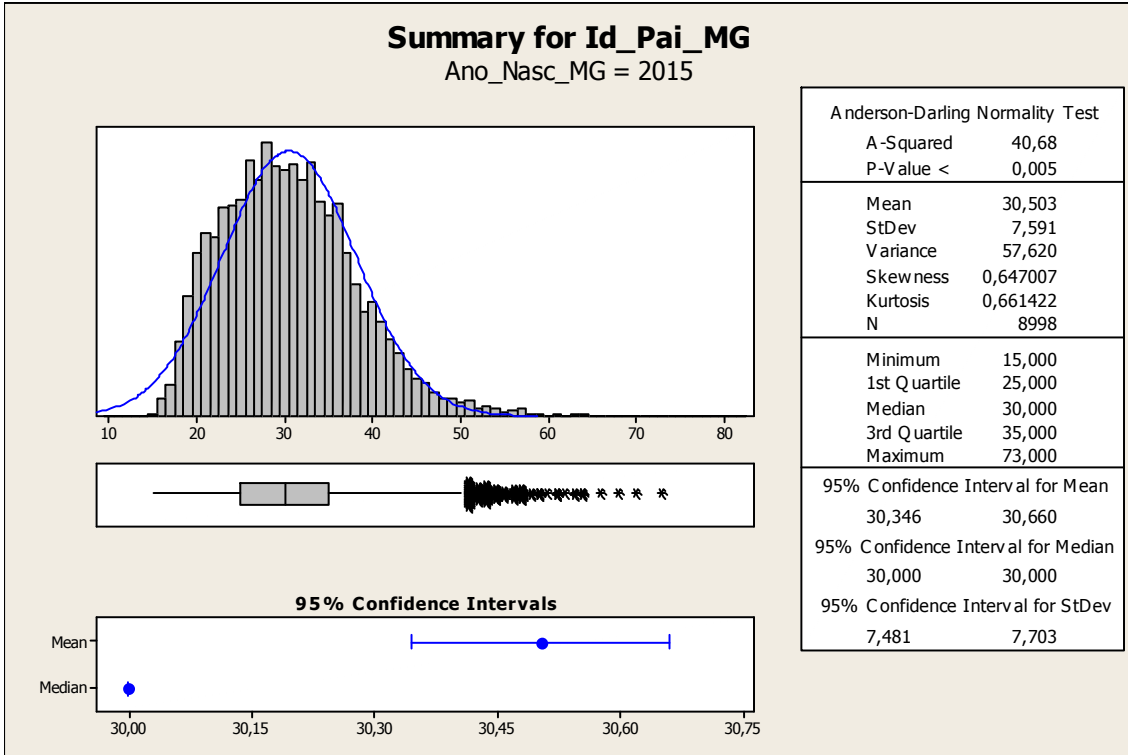
Distribuição da idade dos pais e estatística descritiva para o ano de 2013

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.



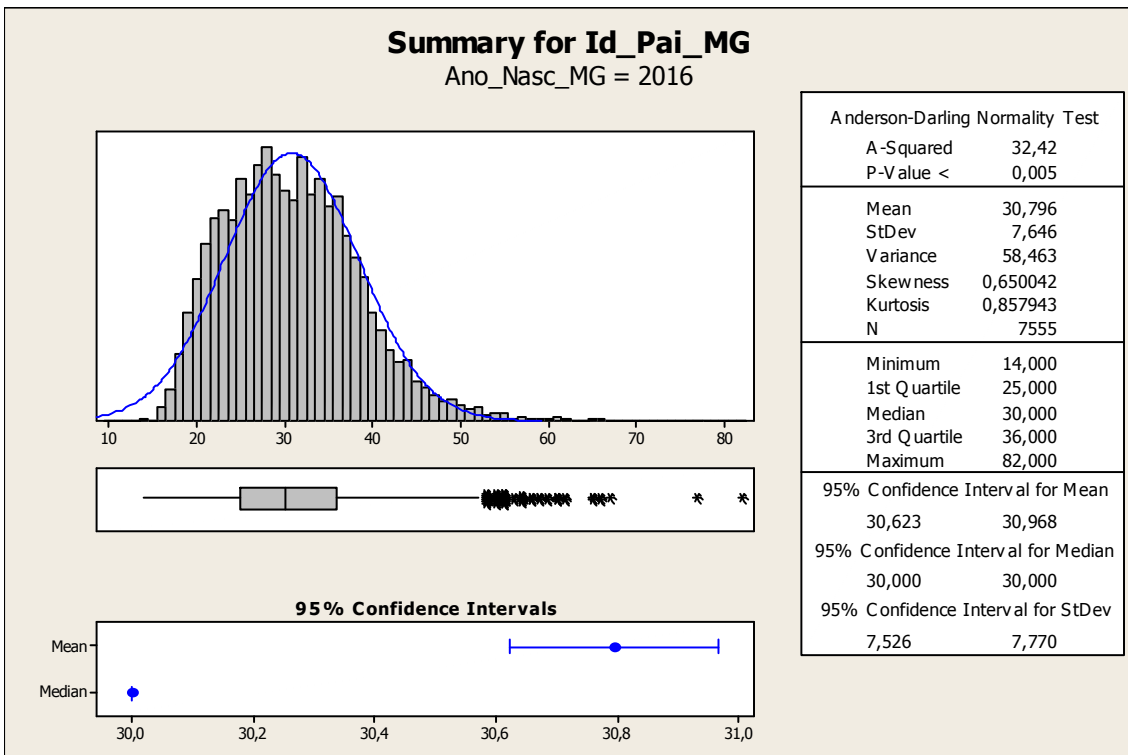
Distribuição da idade dos pais e estatística descritiva para o ano de 2014

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.



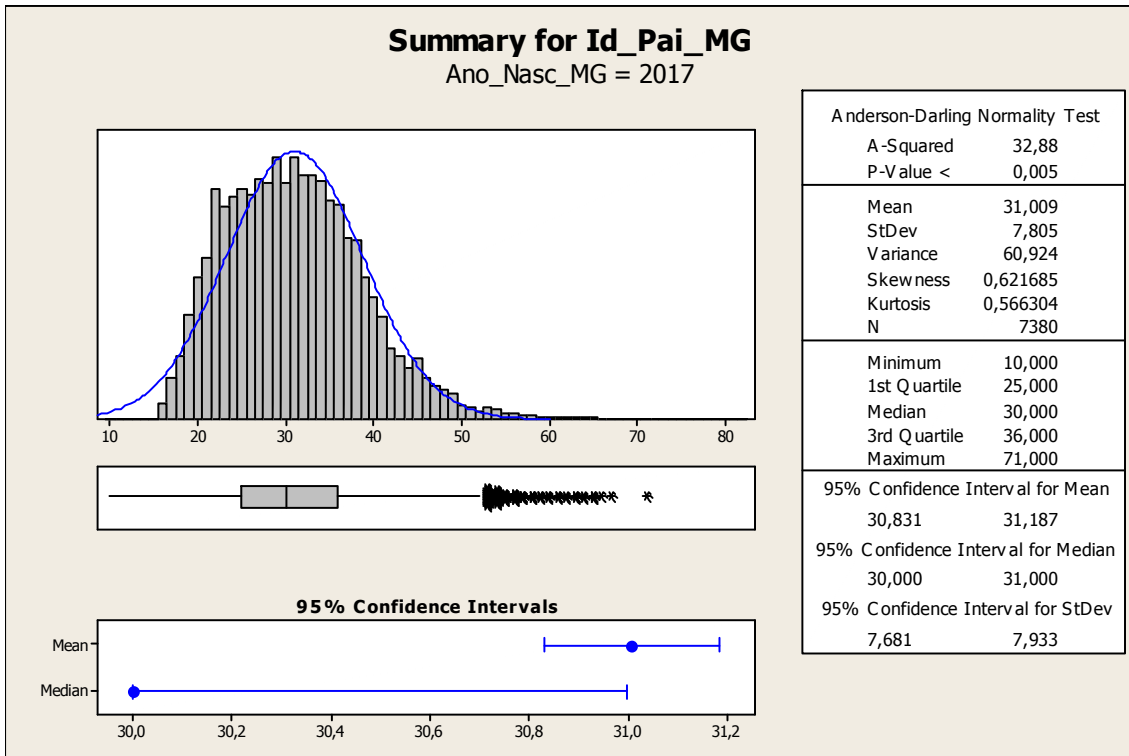
Distribuição da idade dos pais e estatística descritiva para o ano de 2015

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.



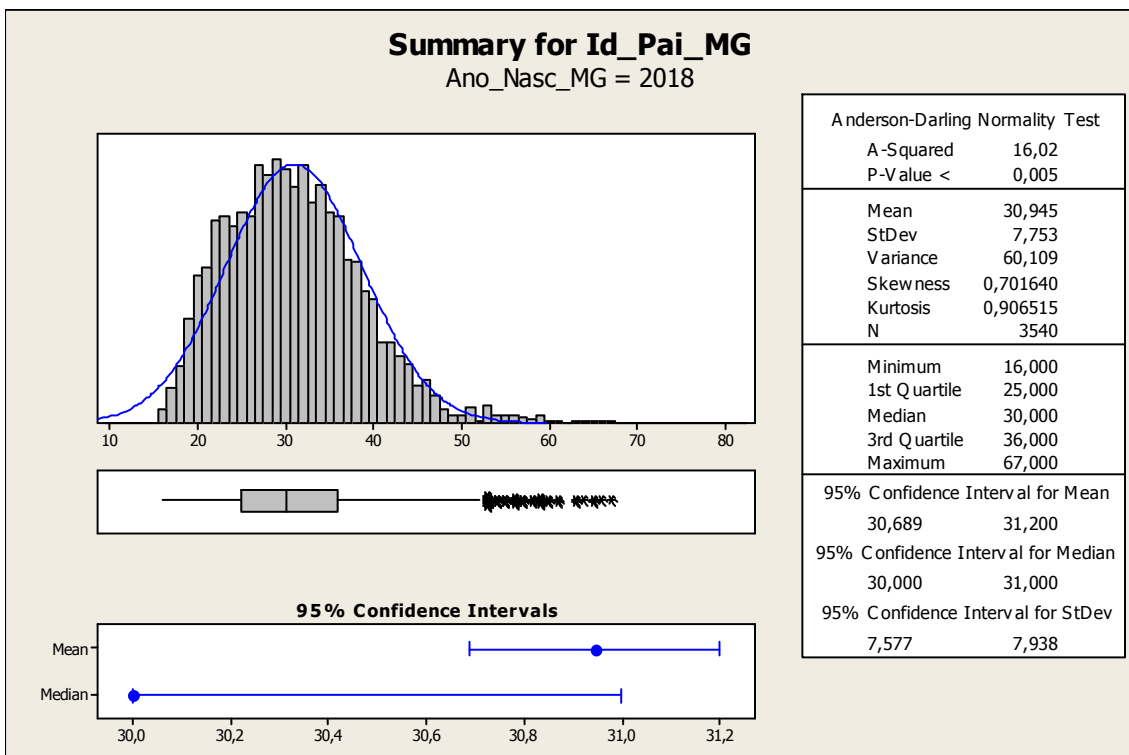
Distribuição da idade dos pais e estatística descritiva para o ano de 2016

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.



Distribuição da idade dos pais e estatística descritiva para o ano de 2017

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.



Distribuição da idade dos pais e estatística descritiva para o ano de 2018

Fonte: Elaboração própria com base nos dados do Cartório de Registro Civil de Venda Nova.