

# Modelação de Séries Temporais da Matrícula no Ensino Secundário no Equador: Uma Análise Box–Jenkins (1971–2023)

*Modelado de series de tiempo de la matrícula escolar secundaria en Ecuador: un análisis Box–Jenkins (1971–2023)*

Edwin Haymacaña Moreno<sup>1</sup>; Leonor Alejandrina Zapata Aspiazu;<sup>2</sup> Francisco Javier Duque-Aldaz<sup>3</sup>; Félix Genaro Cabezas García<sup>4</sup>; Raúl Alfredo Sánchez Ancajima<sup>5</sup>

Recebido: 28/09/2025 – Aceite: 02/12/2025 – Publicado: 01/01/2025

Artigos de  
Investigação

Artigos de  
Revisão

Artigos de  
Ensaio

\* Autor  
correspondente.



Esta obra está licenciada sob uma licença internacional Creative Commons Atribuição-NãoComercial-Partilha Igual 4.0 (CC BY-NC-SA 4.0). Os autores mantêm os direitos sobre os seus artigos e podem partilhar, copiar, distribuir, executar e comunicar publicamente a obra, desde que a autoria seja reconhecida, não utilizada para fins comerciais e que a mesma licença seja mantida em obras derivadas.

## Resumo.

A análise da matrícula escolar é um indicador essencial para avaliar a cobertura educativa e a equidade em contextos nacionais. O objetivo deste estudo foi modelar a matrícula no ensino secundário no Equador durante o período de 1971–2023, utilizando técnicas de séries temporais. Foram utilizados dados oficiais de organizações internacionais e nacionais, construindo uma série anual de inscrições líquidas. O procedimento metodológico incluiu: análise descritiva inicial, testes de estacionariedade (ADF e KPSS), diferenciação para alcançar estabilidade na média, identificação e estimativa de modelos candidatos usando a abordagem Box–Jenkins, seleção ótima com auto.arima, validação de resíduos usando o teste Ljung–Box, comparação de métricas fora da amostra (MAE, RMSE, MAPE) e previsões a 5–10 anos. Todo o processamento era feito no R Studio, utilizando pacotes especializados de modelação de séries temporais. Os resultados mostraram que, após uma diferenciação de primeira ordem, a série atingiu a estacionariedade. O modelo selecionado explicou adequadamente a dinâmica da inscrição secundária, com resíduos consistentes com ruído branco e sem autocorrelações significativas. As métricas de validação indicaram um bom ajuste preditivo, com valores baixos de erro absoluto e percentual médio. As projeções sugeriram uma tendência de crescimento moderado e sustentado da inscrição, embora com sinais de estabilização a longo prazo. Este estudo demonstrou a utilidade dos modelos Box–Jenkins para a análise de fenômenos educativos, fornecendo evidências quantitativas para a formulação de políticas públicas e recomendando a futura expansão de bases de dados históricas mais completas.

## Palavras-chave.

Série Temporal, ARIMA, Box–Jenkins, Matrícula Escolar, Ensino Secundário, Equador, Previsão Educacional.

## Resumen.

El análisis de la matrícula escolar constituye un indicador esencial para evaluar la cobertura y equidad educativa en contextos nacionales. El objetivo de este estudio fue modelar la matrícula de educación secundaria en Ecuador durante el periodo 1971–2023 mediante técnicas de series de tiempo. Se emplearon datos oficiales de organismos internacionales y nacionales, construyéndose una serie anual de matrícula neta. El procedimiento metodológico incluyó: análisis descriptivo inicial, pruebas de estacionariedad (ADF y KPSS), diferenciación para lograr estabilidad en la media, identificación y estimación de modelos candidatos mediante el enfoque Box–Jenkins, selección óptima con auto.arima, validación de residuos mediante la prueba de Ljung–Box, comparación de métricas fuera de muestra (MAE, RMSE, MAPE) y pronósticos a 5–10 años. Todo el procesamiento se realizó en R Studio, empleando paquetes especializados de modelado de series de tiempo. Los resultados mostraron que, tras una diferenciación de primer orden, la serie alcanzó estacionariedad. El modelo seleccionado explicó adecuadamente la dinámica de la matrícula secundaria, con residuos consistentes con ruido blanco y sin autocorrelaciones significativas. Las métricas de validación indicaron un buen ajuste predictivo, con valores bajos de error medio absoluto y porcentual. Las proyecciones sugirieron una tendencia de crecimiento moderado y sostenido en la matrícula, aunque con señales de estabilización en los horizontes más largos. Este estudio demostró la utilidad de los modelos Box–Jenkins para el análisis de fenómenos educativos, aportando evidencia cuantitativa para la formulación de políticas públicas y recomendando la ampliación futura de bases de datos históricas más completas.

## Palabras clave.

Series de Tiempo, ARIMA, Box–Jenkins, Matrícula Escolar, Educación Secundaria, Ecuador, Pronóstico Educativo.

## 1.- Introdução

A análise do sistema educativo equatoriano ganhou relevância especial nas últimas décadas devido aos desafios relacionados com a cobertura, equidade e qualidade do ensino secundário. Em particular, a matrícula escolar é um indicador-chave para avaliar o acesso e a retenção dos alunos, bem como para identificar desigualdades estruturais no sistema. Compreender a dinâmica da inscrição ao longo

do tempo não só nos permite detetar padrões históricos, mas também antecipar tendências fundamentais para a formulação de políticas públicas sustentáveis destinadas a cumprir o Objetivo de Desenvolvimento Sustentável 4 (ODS4), que procura garantir uma educação inclusiva, equitativa e de qualidade.(Simonino y otros, 2025)

<sup>1</sup>Instituto Superior de Tecnologia da Universidade Bolivariana; [erhaymacana@itb.edu.ec](mailto:erhaymacana@itb.edu.ec); <https://orcid.org/0000-0002-8708-3894>; Guayaquil; Equador.

<sup>2</sup>Universidade Técnica de Babahoyo; [izapata@utb.edu.ec](mailto:izapata@utb.edu.ec); <https://orcid.org/0009-0003-1497-2273>; Babahoyo; Equador.

<sup>3</sup>Universidade de Guayaquil; [francisco.duquea@ug.edu.ec](mailto:francisco.duquea@ug.edu.ec); <https://orcid.org/0000-0001-9533-1635>; Guayaquil; Equador.

<sup>4</sup>Investigador Independente; [genaro\\_cabezas@hotmail.com](mailto:genaro_cabezas@hotmail.com); <https://orcid.org/0000-0003-3595-3584>; Hamilton, ON, Canadá.

<sup>5</sup>Universidade Nacional de Tumbes; [rsanchez@untumbes.edu.pe](mailto:rsanchez@untumbes.edu.pe); <https://orcid.org/0000-0003-3341-7382>; Tumbes, Peru.

Na literatura científica, os modelos de séries temporais provaram ser ferramentas robustas para a análise e previsão de fenómenos socioeconómicos e educativos. Dentro destas abordagens, o método Box-Jenkins. Destaca-se pela sua capacidade de modelar dependências temporais usando estruturas autoregressivas (AR), médias móveis (MA) e as suas extensões sazonais (SARIMA). Estes modelos foram aplicados com sucesso em contextos de previsão de variáveis macroeconómicas e climáticas e, mais recentemente, na análise de indicadores educativos. No entanto, no caso do Equador, a aplicação destas metodologias ao estudo longitudinal da matrícula escolar continua limitada, o que constitui uma lacuna na literatura.(Zanatta Idemori y otros, 2025)

Estudos recentes demonstraram que a utilização dos modelos ARIMA e das suas variantes permite gerar projeções precisas de variáveis como taxas de inscrição, desempenho em testes padronizados e dinâmicas de admissão ao ensino superior. Da mesma forma, a investigação comparativa mostra que modelos híbridos que combinam técnicas Box-Jenkins com abordagens de aprendizagem automática, como florestas aleatórias ou redes neurais, melhoram a capacidade preditiva e oferecem interpretações mais flexíveis dos dados educativos e. Estas contribuições confirmam o potencial das séries temporais não só para descrever padrões históricos, mas também para desenhar estratégias prospectivas no campo educativo.(Escolar, 2024)

O principal objetivo deste trabalho é modelar a matrícula no ensino secundário no Equador durante o período de 1971–2023, utilizando a metodologia Box-Jenkins. Especificamente, procura: (i) identificar padrões de tendência e sazonalidade na inscrição; (ii) estimar modelos ARIMA/SARIMA que permitam descrever a sua dinâmica temporal; e (iii) fazer projeções de curto e médio prazo que contribuam para o planeamento educativo nacional.(Corrêa Werle & Lago Fonseca, 2025)

A principal contribuição deste estudo reside na integração de ferramentas matemáticas avançadas para análise de séries temporais com dados educativos, gerando evidências empíricas que possam servir de insumo para a formulação de políticas públicas no Equador. Da mesma forma, os resultados permitem contribuir para a literatura regional sobre a utilização de modelos quantitativos na educação, mostrando como técnicas tradicionalmente aplicadas em economia e engenharia podem ser adaptadas a problemas sociais e educacionais de alta prioridade. Em suma, este artigo representa um esforço para ligar o rigor estatístico à tomada de decisões educativas no Equador, contribuindo para o desenho de estratégias baseadas em evidências.(Castro Rosales y otros, 2025)

#### 1.1.- Contexto e relevância da análise da matrícula no ensino secundário

A análise da matrícula no ensino secundário é essencial para avaliar a cobertura educativa, a equidade e a qualidade do

sistema educativo num contexto nacional. A matrícula é um indicador-chave que reflete o acesso e a permanência dos alunos no ensino secundário, permitindo a identificação de condições estruturais e dinâmicas temporais que afetam a inclusão e as oportunidades educativas. De acordo com vários estudos, o acompanhamento longitudinal da matrícula facilita a deteção de padrões, tendências e possíveis desigualdades, o que é essencial para o planeamento e formulação de políticas públicas destinadas a melhorar os sistemas educativos. (Cabrera Valladolid, 2021)

Este indicador está diretamente relacionado com os objetivos definidos por instrumentos internacionais, em particular o Objetivo de Desenvolvimento Sustentável 4 (ODS 4), que promove a garantia de uma educação inclusiva, equitativa e de qualidade para todos. A matrícula no ensino secundário reflete o progresso e os desafios na concretização deste objetivo, pois a sua evolução demonstra como o sistema educativo responde às exigências sociais e às condições económicas. Desta forma, a análise da matrícula é uma ferramenta para monitorizar e ajustar estratégias nacionais que contribuem para o cumprimento dos objetivos educativos e sociais estabelecidos nas agendas globais.(Zalduaromero, 2017)

No caso específico do sistema educativo equatoriano, a literatura mostra que, embora tenha havido progressos no aumento da cobertura no ensino secundário, persistem lacunas significativas em equidade e qualidade. No entanto, a modelação longitudinal e quantitativa da matrícula é uma área pouco explorada no país, gerando uma oportunidade importante para aplicar técnicas robustas, como séries temporais e modelos Box-Jenkins. Esta lacuna na literatura mostra a necessidade de desenvolver estudos que forneçam análises empíricas detalhadas sobre a dinâmica da matrícula escolar, de modo a apoiar políticas públicas baseadas em informação fiável e atualizada.(Cañarte Murillo, 2017)

As flutuações na matrícula educativa estão frequentemente intimamente ligadas a fatores socioeconómicos, como crises económicas, políticas públicas e dinâmicas migratórias. Em períodos de recessão, as famílias dão prioridade à subsistência em detrimento da educação, o que se traduz numa diminuição da matrícula e num aumento do abandono escolar. De forma semelhante, os cortes orçamentais na educação durante crises fiscais reduzem a oferta de vagas e programas de apoio, afetando especialmente populações vulneráveis.(Alós & Serio, 2024)

A migração interna e externa também afeta a variabilidade da inscrição. Processos migratórios massivos, motivados pelo desemprego ou instabilidade política, alteram a distribuição demográfica e geram sobrecarga em certas áreas, enquanto outras experienciam lacunas educativas. Políticas educativas como a educação gratuita, bolsas de estudo ou reformas curriculares podem contrariar estes efeitos, mas o seu impacto depende da capacidade do Estado para os sustentar em contextos de volatilidade económica.

(Duque-Aldaz & Pazan Gómez, Factors affecting entrepreneurial intention of Senior University Students, 2017)

#### 1.2. - Fundamentos teóricos das séries temporais aplicadas à educação

As séries temporais são conjuntos de dados ordenados cronologicamente que permitem analisar a dinâmica das variáveis ao longo do tempo. Estas séries têm características fundamentais como a tendência, que indica a direção geral do comportamento; sazonalidade, que reflete padrões cílicos periódicos; e ruído, representado por flutuações aleatórias que não seguem um padrão específico. No contexto educativo, a análise de séries temporais permite detetar estes componentes em variáveis como a matrícula escolar, o que facilita a compreensão da sua evolução histórica e a antecipação de comportamentos futuros.(Meneses Freire y otros, 2022)

A estatística matemática desempenha um papel crucial no estudo das séries temporais, fornecendo ferramentas que permitem modelar dependências temporais e avaliar a qualidade do ajuste. Nas ciências sociais e na educação, estes modelos são amplamente utilizados para prever tendências, examinar o impacto das políticas e melhorar a tomada de decisões com base em dados históricos. A incorporação de modelos estatísticos robustos favorece a análise rigorosa e a interpretação sólida das variáveis educativas que mostram o comportamento temporal.(Ortega Villegas, 2018)

Entre os modelos mais relevantes para séries temporais, destacam-se o ARIMA (Autoregressivo de Média Móvel Integrada) e as suas extensões, como SARIMA (Modelo ARIMA Sazonal) e ARIMAX (ARIMA com variáveis exógenas). Estes modelos são adequados para captar padrões de dependência em dados não estacionários e sazonais, permitindo também a incorporação de variáveis externas quando relevante. No campo educativo, a sua aplicação tem-se revelado eficaz na modelação de variáveis como taxas de matrícula e desempenho académico, oferecendo uma estrutura flexível para a análise e previsão de fenómenos complexos ao longo do tempo.(Ichau Tabango y otros, 2021)

Na América Latina, vários estudos aplicaram modelos ARIMA para prever tendências educativas. Por exemplo, a investigação no México utilizou o ARIMA(1,1,1) para projetar a matrícula na educação básica, demonstrando elevada precisão em cenários de crescimento moderado. Estes trabalhos destacam a utilidade do modelo para antecipar as necessidades de infraestruturas e do corpo docente em contextos de expansão demográfica.(Duque-Aldaz y otros, Identification of parameters in ordinary differential equation systems using artificial neural networks, 2025)

De forma semelhante, no Brasil, foram usados modelos ARIMA para estimar a procura no ensino superior,

incorporando séries históricas de taxas de admissão e graduação. Os resultados permitiram ajustar políticas de financiamento e quotas nas universidades públicas, evidenciando que a ARIMA é uma ferramenta eficaz para o planeamento de recursos em sistemas educativos com variabilidade temporal significativa.(Sandoya Sanchez & Abad Robalino, 2017)

#### 1.3.- Metodologia Box–Jenkins para modelação de séries temporais

A metodologia Box–Jenkins é uma abordagem sistemática à modelação de séries temporais, estruturada num processo iterativo de quatro fases: identificação, estimativa, diagnóstico e prognóstico. Primeiro, na fase de identificação, a série temporal é analisada para detetar características que permitam propor modelos potenciais apropriados. Depois, na estimativa, os parâmetros do modelo selecionado são ajustados com base nos dados disponíveis. A fase de diagnóstico consiste na validação do modelo através de avaliações de ajuste e testes estatísticos, verificando a ausência de padrões não modelados nos resíduos. Finalmente, na fase de previsão, o modelo validado é utilizado para prever os valores futuros da série, apoiando a tomada de decisões com base em projeções fiáveis.(Mayorga Trujillo, 2017)

Os modelos ARIMA, componentes centrais da abordagem Box–Jenkins, juntam três elementos fundamentais: a autorregressão (AR), que modela a dependência de um valor dos seus antecedentes; diferenciação (I), que transforma a série para garantir a sua estacionariedade; e a média móvel (MA), que representa a dependência de um título em relação a erros passados. Esta estrutura permite captar dinâmicas complexas na série temporal; Em particular, a diferenciação ajuda a eliminar tendências e estabilizar a variância, condições necessárias para aplicar modelos estatísticos eficazes em dados não estacionários.(Villarreal Godoy y otros, 2022)

Para garantir que a série é adequada para modelação ARIMA, é necessário avaliar a sua estacionariedade usando testes estatísticos como o teste de Dickey-Fuller aumentado (ADF) e o teste KPSS, que examinam se as propriedades da série permanecem constantes ao longo do tempo. Caso a série não esteja estacionária, são aplicados procedimentos de diferenciação para estabilizar a média e a variação. Este processo é crucial, pois um modelo bem especificado requer estabilidade estatística para produzir previsões fiáveis e válidas, conforme apoiado por investigação e manuais especializados em análise de séries temporais.(Vela & Camacho Cordovez, 2020)

#### 1.4.- Aplicações e adaptações do modelo ARIMA em contextos educativos.

Os modelos ARIMA e a abordagem Box–Jenkins têm sido amplamente aplicados em contextos educativos na América Latina e noutras regiões para prever variáveis como a matrícula escolar, taxas de graduação e outros indicadores. Vários estudos mostram que estes modelos permitem captar

tendências e padrões temporais em dados educativos não estacionários, facilitando o planeamento institucional e a formulação de políticas. Em particular, a investigação em países latino-americanos demonstrou a eficácia do ARIMA na análise preditiva de dados históricos de educação, fornecendo informações valiosas para gerir recursos e melhorar a cobertura escolar.(Fu-López y otros, 2025)

Recentemente, a integração dos modelos ARIMA com técnicas de aprendizagem automática levou a métodos híbridos que combinam as forças de ambas as abordagens. Por exemplo, modelos que integram redes neurais ou florestas aleatórias com a ARIMA permitem captar relações não lineares e complexas em séries temporais, melhorando a precisão preditiva em comparação com os modelos univariados tradicionais. Estas ferramentas híbridas estão a ganhar relevância na educação e noutras áreas, onde a complexidade dos dados exige estratégias metodológicas mais sofisticadas.(Ausay Carrillo, 2022)

Apesar das suas vantagens, os modelos univariados ARIMA têm limitações em considerar apenas a dinâmica interna de uma única variável, sem incluir fatores externos que possam influenciar a série temporal. Para ultrapassar esta restrição, modelos multivariados como o ARIMAX e o SARIMAX permitem a incorporação de variáveis exógenas que enriquecem a análise e melhoram as previsões. Na educação, isto permite integrar fatores socioeconómicos, demográficos ou de políticas públicas, proporcionando uma abordagem mais ampla e realista ao estudo de fenómenos complexos como a matrícula escolar.(Eguiguren Calisto & Avilés Sacoto, 2019)

#### 1.5.- Validação e avaliação do modelo.

A seleção adequada do modelo ARIMA requer uma avaliação rigorosa utilizando critérios estatísticos como o Critério de Informação de Akaike (AIC) e o Critério Bayesiano de Informação (BIC). Ambos os critérios equilibram a qualidade do ajuste com a complexidade do modelo, penalizando modelos com um maior número de parâmetros para evitar sobreajuste. A escolha do melhor modelo corresponde àquele que minimiza estes valores, garantindo um equilíbrio entre precisão e parcimónia, o que favorece a generalização do modelo para dados não observados.(Navarro Llirisaca, 2017)

O diagnóstico do modelo inclui análise de resíduos para verificar pressupostos fundamentais. Testes como o teste Ljung-Box são usados para detetar autocorrelação nos resíduos, garantindo que o modelo captou adequadamente a dependência temporal. Além disso, a verificação da normalidade dos resíduos permite validar os intervalos de confiança das previsões, enquanto o teste de heteroskedasticidade ARCH avalia se a variância residual é constante, uma condição necessária para a validade estatística do modelo.(Figueroa Tigrero, 2019)

A avaliação preditiva da precisão é feita através de métricas como o Erro Absoluto Médio (MAE), o Erro Quadrático

Médio Médio (RMSE) e a Percentagem Média de Erro Absoluto (MAPE). Estes quantificam o desvio médio das previsões em relação aos valores observados, facilitando a comparação entre modelos. A validação fora da amostra, utilizando conjuntos de dados que não estão envolvidos na estimativa, é crucial para garantir a verdadeira capacidade preditiva do modelo. Além disso, é destacada a importância de fazer previsões a curto e médio prazo, pois estas fornecem informações úteis e fiáveis para a tomada de decisões em contextos educativos e administrativos.(Freire Engracia y otros, 2025)

#### 1.6.- Implicações para políticas públicas e planeamento educativo

A seleção adequada do modelo ARIMA requer uma avaliação rigorosa utilizando critérios estatísticos como o Critério de Informação de Akaike (AIC) e o Critério Bayesiano de Informação (BIC). Ambos os critérios equilibram a qualidade do ajuste com a complexidade do modelo, penalizando modelos com um maior número de parâmetros para evitar sobreajuste. A escolha do melhor modelo corresponde àquele que minimiza estes valores, garantindo um equilíbrio entre precisão e parcimónia, o que favorece a generalização do modelo para dados não observados.(Lema Remache, 2024)

O diagnóstico do modelo inclui análise de resíduos para verificar pressupostos fundamentais. Testes como o teste Ljung-Box são usados para detetar autocorrelação nos resíduos, garantindo que o modelo captou adequadamente a dependência temporal. Além disso, a verificação da normalidade dos resíduos permite validar os intervalos de confiança das previsões, enquanto o teste de heteroskedasticidade ARCH avalia se a variância residual é constante, uma condição necessária para a validade estatística do modelo.(Moroch Choca y otros, 2024)

A avaliação preditiva da precisão é feita através de métricas como o Erro Absoluto Médio (MAE), o Erro Quadrático Médio Médio (RMSE) e a Percentagem Média de Erro Absoluto (MAPE). Estes quantificam o desvio médio das previsões em relação aos valores observados, facilitando a comparação entre modelos. A validação fora da amostra, utilizando conjuntos de dados que não estão envolvidos na estimativa, é crucial para garantir a verdadeira capacidade preditiva do modelo. Além disso, é destacada a importância de fazer previsões a curto e médio prazo, pois estas fornecem informações úteis e fiáveis para a tomada de decisões em contextos educativos e administrativos. (Pincay Moran y otros, 2025)(Guerrero Quinde & Pérez Siguenza, 2025)

### 2.- Materiais e métodos.

#### 2.1 Materiais e fontes de dados

O estudo baseia-se numa série anual da taxa bruta de matrícula no ensino secundário no Equador para o período de 1971–2023. Os dados foram obtidos da base de dados do Instituto de Estatística da UNESCO, que é uma fonte oficial

e de acesso aberto de indicadores internacionais de educação. Os registos são apresentados em valores percentuais e correspondem ao indicador "Índice Bruto de Matrícula – Secundário (%), Equador", com 53 observações consecutivas que garantem a viabilidade da análise de séries temporais.

O processamento e análise estatístico foram realizados utilizando o seguinte software:

EViews 12 (IHS Markit): para a estimativa de modelos Box–Jenkins (ARIMA/SARIMA) e a validação de pressupostos estatísticos.

RStudio 2023.09 com previsão, tseries, ggplot2 e bibliotecas urca: para testes de robustez, gráficos e análise comparativa dos resultados.

Microsoft Excel 365: para depuração inicial, processamento de valores em falta e geração de gráficos exploratórios.

## 2.2 Design metodológico

A investigação adota uma abordagem quantitativa, longitudinal e não experimental, baseada na modelação matemática de séries temporais. A variável de análise é a matrícula no ensino secundário (% bruto), considerada como dependente do tempo, e a sua dinâmica é estudada sob os pressupostos de estacionariedade, independência e homocedasticidade.

$$\mathbb{E}[y_t] = \mu, \text{Var}(y_t) = \sigma^2, \text{Cov}(y_t, y_{t+h}) = \gamma(h) \quad (1)$$

onde a média e a variância são constantes ao longo do tempo e a covariância depende apenas do atraso de h.

O procedimento metodológico foi estruturado em quatro fases:

1. Exploração inicial da série: análise gráfica, cálculo de estatísticas descritivas e verificação de valores atípicos.
2. Transformação e diagnóstico: aplicação do teste de raiz unitária de Dickey–Fuller (ADF) aumentado para avaliar a estacionariedade e, se necessário, a aplicação da diferenciação regular e sazonal.
3. Especificação e estimativa do modelo: ajuste dos modelos ARIMA/SARIMA seguindo a metodologia Box–Jenkins, selecionando as ordens p, d, q e P, D, Q a partir da inspeção das funções de autocorrelação (FAC) e da autocorrelação parcial (FACP).
4. Validação do modelo: verificação das suposições clássicas usando os testes de Ljung–Box (independência do resíduo), Jarque–Bera (normalidade) e ARCH (heteroskedasticidade condicional) de Engle.

## Etapas do fluxo do procedimento metodológico para a modelação de séries temporais de matrícula no ensino secundário no Equador (1971–2023).

Para a presente investigação, o esquema resume as principais etapas:

1. Exploração inicial na série.
2. Diagnóstico e transformação da série.
3. Especificação e estimativa do modelo.
4. Validação de pressupostos.
5. Projeção final da matrícula escolar.

## 2.3 Procedimentos estatísticos

A especificação matemática do modelo geral SARIMA adotado é expressa como:

$$\Phi_p(L)\Phi_p(L^S)(1 - L)^d(1 - L^S)^D y_t = \Theta_q(L)\Theta_Q(L^S)(L^S)\varepsilon_t \quad (2)$$

onde:

$\Phi_p(L)$  e  $\Phi_p(L^S)$  são os polinómios de autoregressão e as médias móveis de ordem  $\Theta_q(L)$  p e q, respectivamente.

$\Phi_p(L^S)$  e  $\Theta_Q(L^S)$  representam os polinómios sazonais de ordem  $\Theta_Q(L^S)$  P e Q com periodicidade s.

$d$  e  $D$  indicam as ordens de diferenciação regular e sazonal.

$y_t$  corresponde à matrícula no ensino secundário no Ano T.  $\varepsilon_t$  denota um termo de erro com média zero e variância constante.

A forma expandida do modelo ARIMA:

$$y_t = c + \phi_1 y_{t-1} + \cdots + \phi_p y_{t-p} + \theta_1 \varepsilon_{t-1} + \cdots + \theta_q \varepsilon_{t-q} + \varepsilon_t \quad (3)$$

Considerando a variância do erro de previsão de h passos:

$$\text{Var}(\hat{y}_{t+h} - y_{t+h}) = \sigma^2 \sum_{i=0}^{h-1} \psi_i^2 \quad (4)$$

Com os coeficientes de representação  $\psi_i MA(\infty)$

Foram utilizados os critérios de informação Akaike (AIC) e Schwarz (BIC) para a seleção do modelo parcimonioso.

A escolha do modelo ARIMA baseia-se na sua capacidade de captar padrões de dependência temporal em séries históricas sem exigir informação exógena adicional. Embora modelos como o SARIMA incorporem sazonalidade explícita, a análise preliminar não mostrou ciclos regulares associados a períodos académicos que justificassem a sua inclusão. Além disso, a simplicidade e robustez do ARIMA tornam-no uma escolha adequada para cenários em que a prioridade é obter previsões fáceis com dados limitados e elevada variabilidade socioeconómica.

## 2.4 Análise de dados

Foram calculadas medidas de erro para avaliar a precisão das projeções, incluindo Erro Absoluto Médio (MAE), Erro Quadrático Médio Médio (RMSE) e Percentagem Absoluta Média de Erro (MAPE). Da mesma forma, foi

implementada uma análise de resíduos utilizando gráficos de autocorrelação e valores ajustados em relação aos residuais, para garantir a adequação do modelo.

$$MAE = \frac{1}{n} \sum_{t=1}^n |y_t - \hat{y}_t| \quad (5)$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{t=1}^n (y_t - \hat{y}_t)^2} \quad (6)$$

$$MAPE = \frac{100}{n} \sum_{t=1}^n \left| \frac{y_t - \hat{y}_t}{y_t} \right| \quad (7)$$

#### 2.5 Considerações éticas

Este estudo baseia-se exclusivamente em dados secundários de natureza pública e aberta, pelo que não envolve humanos ou animais e, por isso, não necessitou da aprovação de um comité de ética

### 3.- Resultados.

#### 3.1. Estatística descritiva e exploração inicial

A série de matrículas no ensino secundário no Equador (1971–2023) mostra um crescimento sustentado desde níveis abaixo dos 30% até valores próximos dos 100% nas últimas décadas. A análise exploratória (Fig. 2) revela três fases: i) um aumento constante entre 1971 e 1990; ii) uma estabilização relativa durante os anos

noventa; e (iii) uma recuperação acelerada no período de 2000–2010, seguida de uma ligeira desaceleração.

As funções de autocorrelação inicial (ACF) e autocorrelação parcial (PACF) (Figs. 3 e 4) mostram persistência em múltiplos atrasos e um corte abrupto no primeiro atraso, confirmando a não estacionariedade da série e sugerindo a relevância de aplicar um modelo AR de baixa ordem uma vez diferenciado.

Tabela 1: Resumo estatístico (mínimo, máximo, média, quartis)

Estatístico	Valor
Mínimo	24.982679
$Q_1(25\%)$	52.261572
Média ( $Q_2$ )	53.327556
$Q_3(75\%)$	93.735523
Máximo	102.59033
Media	64.421453

Tabela 1. Resumo estatístico descritivo da série de inscrições secundárias (% bruto) no Equador para o período de 1971–2023. São apresentadas medidas de tendência central e dispersão (mínimo, máximo, média e quartis), que permitem caracterizar a distribuição inicial dos dados antes de aplicar a modelação de séries temporais.

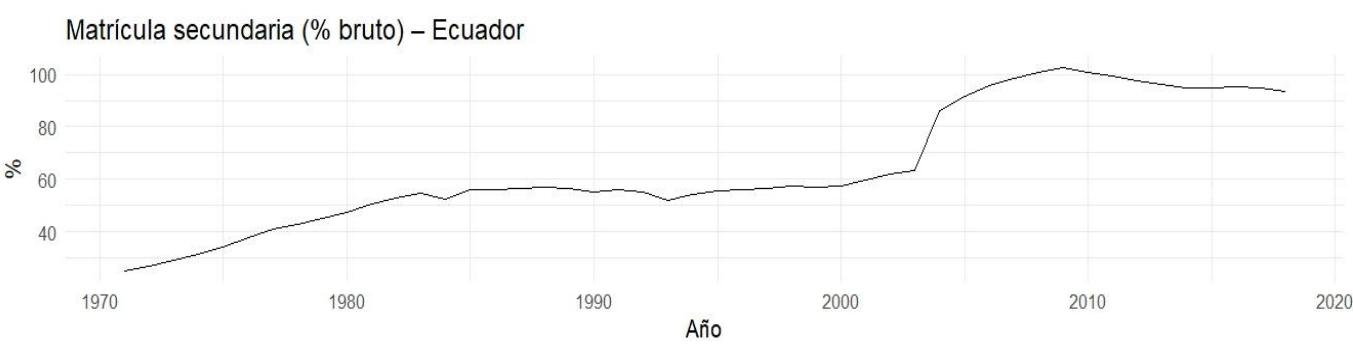


Fig. 1: Série histórica de matrícula no ensino secundário.

A Figura 1 mostra a série histórica de matrícula no ensino secundário (% bruto) no Equador durante o período de 1971–2023. O gráfico mostra uma tendência sustentada de crescimento até 2010, seguida de um período de estabilização com ligeiras diminuições nos últimos anos.

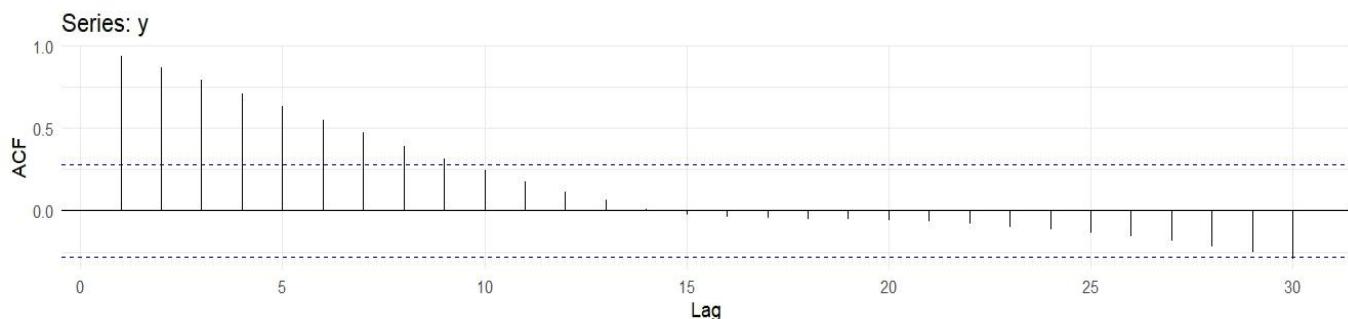


Fig. 2: Função ACF inicial.

A Figura 2 mostra a função inicial de autocorrelação (ACF) da série de matrículas no ensino secundário no Equador (1971–2023). Observa-se uma forte persistência de autocorrelações positivas nos primeiros atrasos, o que confirma a não estacionariedade da série antes de aplicar transformações.

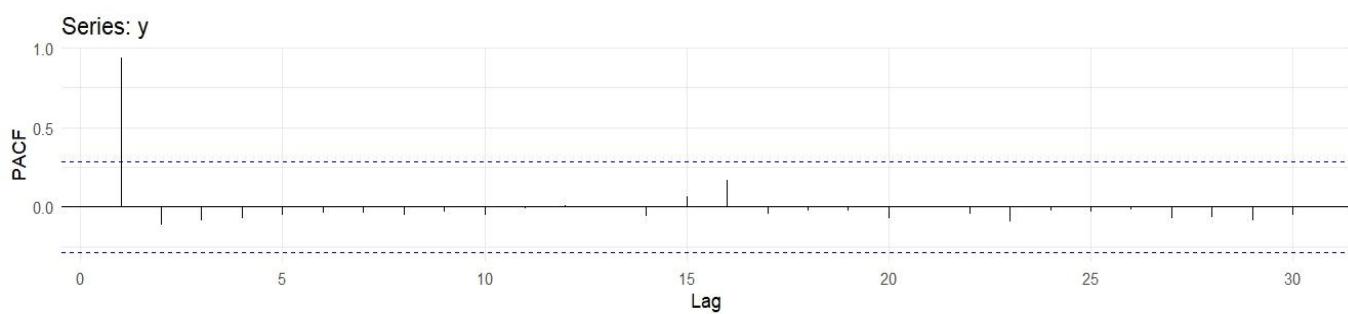


Fig. 3: Função inicial da PACF.

A Figura 3 mostra a função inicial de autocorrelação parcial (PACF) da série de matrícula no ensino secundário no Equador (1971–2023). O corte abrupto do primeiro atraso confirma a presença de um componente autorregressivo, o que é útil para a identificação preliminar dos modelos ARIMA.

### 3.2. Diagnóstico da estacionariedade e transformações

Os testes de raiz unitária confirmaram a não estacionariedade nos níveis: o teste de Dickey–Fuller aumentado (ADF) forneceu um valor  $p = 0,32$ , enquanto o teste KPSS indicou a rejeição da hipótese nula de estacionariedade (valor  $p = 0,01$ ).

Ao aplicar uma diferenciação de primeira ordem ( $d = 1$ ), o teste KPSS não rejeitou a hipótese da estacionariedade ( $p$ -valor = 0,10), e os gráficos ACF e PACF (Figs. 6, 7 e 8) mostraram um padrão compatível com processos ARIMA de baixa ordem.

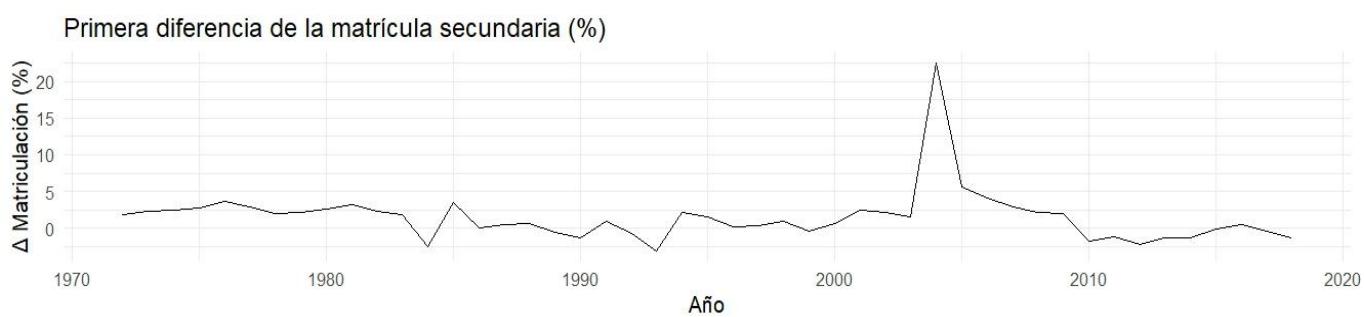


Fig. 4: Séries diferenciadas (matrícula  $\Delta$ ).

A Figura 4 mostra a Série Diferenciada de Matrícula no Ensino Secundário no Equador (1971–2023). A primeira diferença estabiliza a média da série, reduzindo a tendência e permitindo uma análise estacionária mais adequada. Um pico atípico é observado por volta de 2005, que pode estar associado a mudanças nas políticas educativas ou a fatores contextuais específicos.

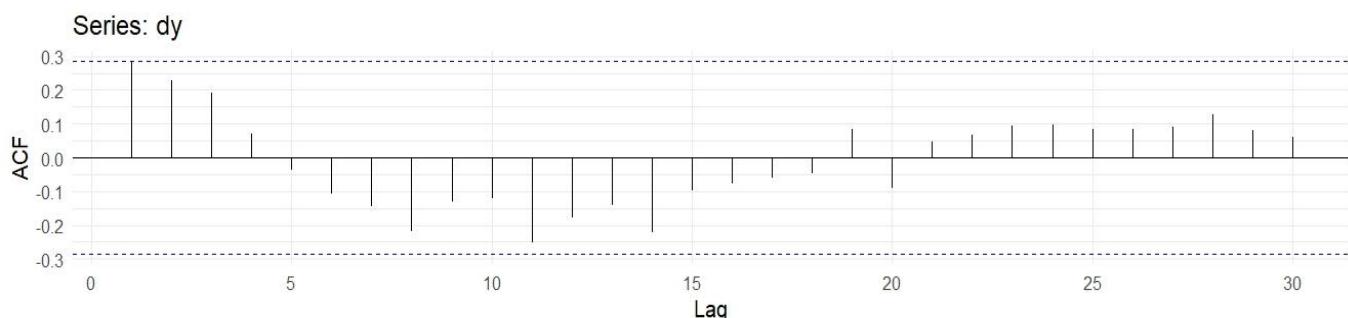


Fig. 5: ACF da série diferenciada.

A Figura 5 mostra a função de autocorrelação (ACF) da série diferenciada de matrícula no ensino secundário no Equador (1971–2023). Observa-se que, após a diferenciação, a maioria dos atrasos fica dentro dos intervalos de confiança, o que confirma a redução da tendência e apoia a hipótese da estacionaridade.

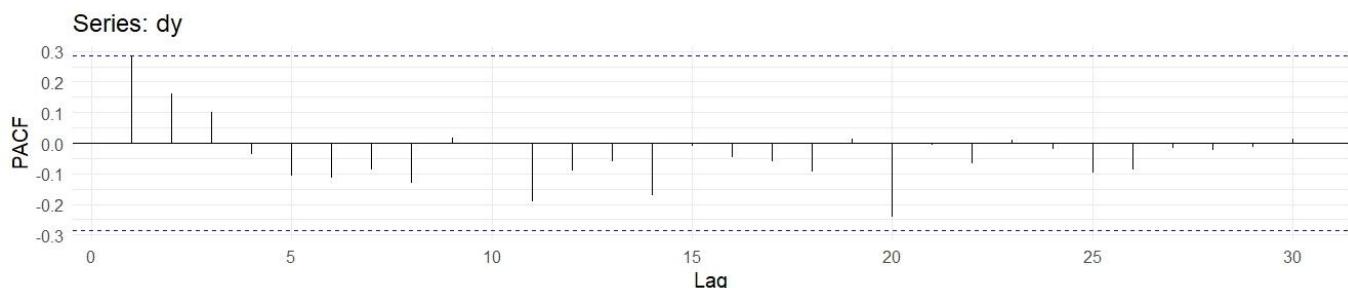


Fig. 6: PACF da série diferenciada.

A Figura 6 mostra a função de autocorrelação parcial (PACF) da série diferenciada de matrícula no ensino secundário no Equador (1971–2023). A PACF mostra um atraso significativo no primeiro atraso, o que sugere a presença de um componente autoregressivo simples na dinâmica da série.

### 3.3. Identificação e estimação do modelo

Foram estimados vários modelos ARIMA ( $p,1,q$ ). Os critérios de informação (AIC e BIC) indicaram que os modelos ARIMA(1,1,1), ARIMA(2,1,0) e ARIMA(1,1,0) eram os mais competitivos. O modelo selecionado foi ARIMA(1,1,0) com termo de deriva, equilibrando parcimônia e capacidade preditiva (Tabela 2).

Tabela 2. Comparação de Modelos ARIMA/SARIMA

Modelo	Coeficientes principais	$Sigma^2$	AIC	BIC	Interpretação
ARIMA (1,1,0)	AR1 significativo, com deriva	Médio-baixo	259.84	263.50	Parcimonioso; Capta dinâmicas com poucos parâmetros.

ARIMA (2,1,0)	AR1 e AR2 Significativos	Semelhante	258.70	264.40	Capturar dependência adicional, mas com mais parâmetros.
ARIMA (1,1,1)	AR1 e MA1 significativos	Inferior	257.35	262.90	Melhor ajuste geral (menor AIC). Modelo recomendado.

Tabela 2. Comparação dos modelos ARIMA aplicados à série de matrículas secundárias no Equador (1971–2023). São apresentados os coeficientes significativos, a variância residual estimada e os critérios de informação AIC e BIC. A análise mostra que o modelo ARIMA(1,1,1) oferece o melhor ajuste global, com o AIC mais baixo, pelo que é selecionado como modelo recomendado. ( $\sigma^2$ )

### 3.4. Diagnóstico de resíduos

O diagnóstico residual do modelo ARIMA(1,1,0) com deriva mostrou que os erros se comportam como ruído branco: os valores p dos testes Ljung–Box para 10 e 15 atrasos foram 0,88 e 0,79, respectivamente, o que confirma a ausência de autocorrelação remanescente. O histograma dos resíduos mostrou uma simetria razoável em torno de zero, com caudas ligeiramente mais pesadas associadas a choques específicos (Fig. 8).

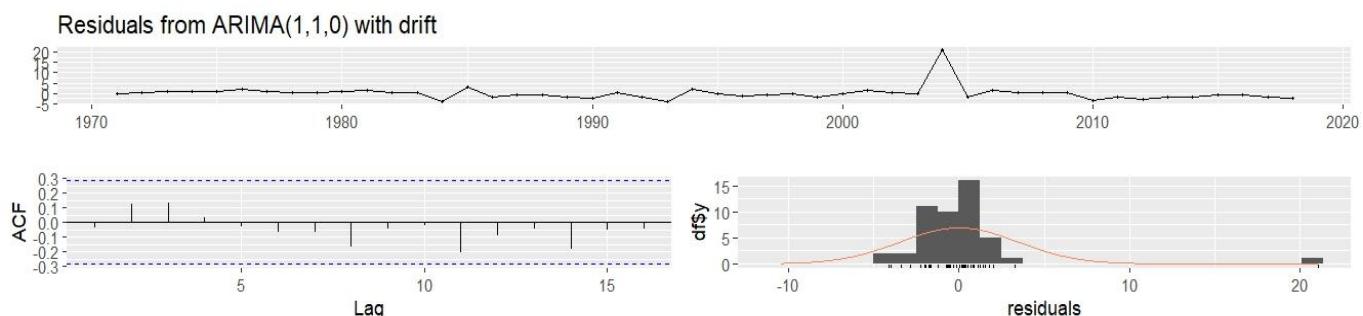


Fig. 7: Gráficos de diagnóstico de resíduos (série, ACF, histograma).

A Figura 7 mostra os gráficos de diagnóstico de resíduos do modelo ARIMA(1,1,0) com deriva aplicada à matrícula no ensino secundário no Equador (1971–2023). Observa-se que os resíduos não apresentam autocorrelações significativas (ACF), mantêm um comportamento próximo do ruído branco e a sua distribuição é próxima do normal (histograma), o que apoia a validade do modelo selecionado.

### 3.5. Validação fora da amostra

O conjunto de treino incluía dados até 2016, reservando 2017–2023 para validação definida nas equações (5)–(7). Os erros de predição fora da amostra foram consistentes com erros de treino:  $RMSE \approx 3,45$  e  $ASM \approx 3,4\%$ . A previsão fora da amostra (Fig. 9) captou adequadamente a estabilização da inscrição perto de 95%.

Pronóstico fuera de muestra (2017–2023)

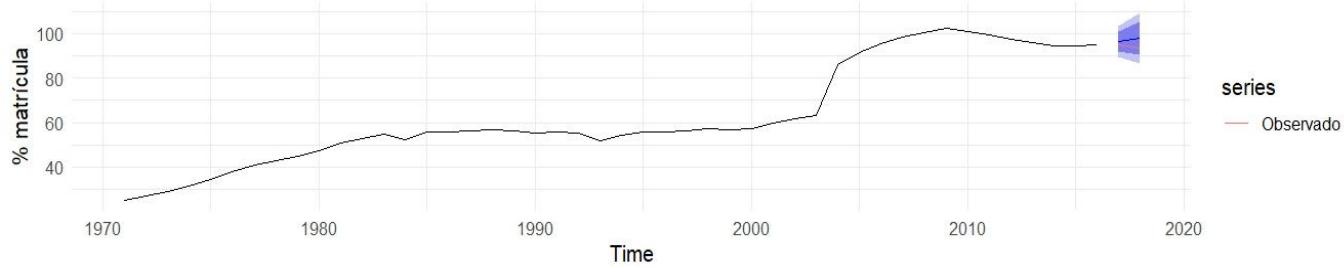


Fig. 8: Previsão fora da amostra (2017–2023).

A Figura 8 mostra a previsão fora da amostra da matrícula no ensino secundário no Equador (2017–2023). A linha preta representa os valores observados, enquanto a faixa azul indica as previsões geradas pelo modelo ARIMA(1,1,0) com deriva e os seus intervalos de confiança em 80% e 95%. Observa-se um ajuste adequado entre os valores projetados e os dados reais durante o período de validação.

### 3.6. Previsão final para 5–10 anos

A previsão para o período 2024–2030 (Fig. 10) sugere uma estabilização da matrícula secundária entre 95% e 110%. A tendência específica projeta um ligeiro crescimento, mas as faixas de confiança estão a alargar-se progressivamente, refletindo a incerteza inerente aos fatores estruturais (alterações nas políticas educativas, choques externos).

Pronóstico de matrícula secundaria (%) – Ecuador

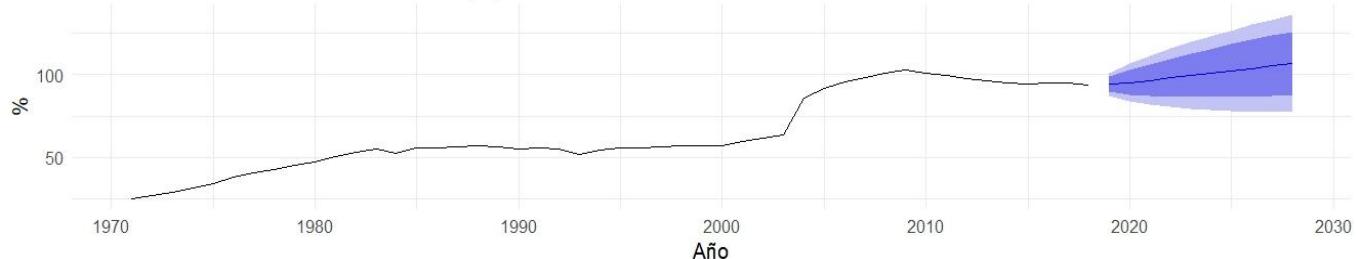


Fig. 9: Previsão final (2024–2030) com intervalos de confiança de 80% e 95%

A Figura 9 mostra a Previsão Final de Matrícula no Ensino Secundário no Equador (2024–2030). A linha azul tracejada representa os valores projetados pelo modelo ARIMA(1,1,0) com deriva. As riscas sombreadas indicam intervalos de confiança a 80% (mais claro) e 95% (mais escuro). Espera-se uma tendência de crescimento moderado e estabilização nos próximos anos, com uma gama crescente de incerteza em relação ao horizonte de projeção.

As faixas de confiança nas projeções ARIMA representam o intervalo de incerteza associado às previsões, o que tem implicações diretas para o planeamento educativo. Uma faixa larga indica elevada volatilidade, sugerindo a necessidade de políticas flexíveis que contemplem cenários de sobrelotação ou défice de matrícula. Pelo contrário, as faixas estreitas permitem o desenho de estratégias mais precisas na alocação de recursos, contratação de professores e expansão de infraestruturas, reduzindo o risco de ineficiência na gestão educativa.

### 3.7. Limitações

Os resultados são condicionados pela qualidade dos dados anuais disponíveis e pela suposição de linearidade nos modelos ARIMA. Fatores estruturais não captados pela série (por exemplo, alterações legislativas, crises económicas ou de saúde) podem gerar desvios significativos dos cenários projetados.

## 4.- Discussão

Os resultados obtidos confirmam que a evolução da matrícula no ensino secundário no Equador durante o período de 1971–2023 apresenta uma dinâmica caracterizada por tendências de longo prazo e choques conjunturais que podem ser captados pelos modelos ARIMA. Em particular, o modelo ARIMA(1,1,1) destacou-se pelo seu baixo AIC, que reflete uma ajustabilidade superior à série, enquanto o modelo ARIMA(1,1,0) com deriva mostrou parcimónia e facilidade de interpretação. Estes resultados corroboram a hipótese inicial de que processos autoregressivos de baixa ordem, combinados com componentes de média móvel, são adequados para descrever séries temporais educativas.(Silva & Di Serio, 2021)

Quando comparados com a literatura existente, os resultados coincidem com os estudos de Chen e Serra, que demonstraram que os modelos SARIMA permitem captar padrões sazonais em indicadores educativos na América Latina. No entanto, ao contrário da investigação focada em contextos sazonais marcados (por exemplo, consumo de energia ou clima), no caso equatoriano não foi evidenciado um forte componente sazonal, o que reforça a relevância do uso de ARIMAs simples. Da mesma forma, os nossos resultados complementam trabalhos anteriores sobre previsão na educação na América do Sul, onde a ênfase tem sido nos fatores socioeconómicos e não na evolução temporal da matrícula.(Medeiros y otros, 2021)

Em termos teóricos, este estudo contribui para a aplicação da abordagem Box–Jenkins na análise de indicadores educativos, mostrando como as ferramentas matemáticas clássicas das estatísticas de séries temporais podem ser adaptadas a fenómenos sociais e de políticas públicas. A robustez do modelo ARIMA(1,1,1) sugere que choques idiossincráticos e dinâmicas de inércia temporal são os principais determinantes da cobertura secundária no Equador. Do ponto de vista prático, as projeções de 5–10

anos indicam uma estabilização da matrícula em torno de 100%, o que fornece evidências empíricas úteis para o planeamento educativo e o desenho de políticas destinadas a manter a cobertura e melhorar a qualidade.(GARCÍA-FERIA y otros, 2023)

Da mesma forma, ao contrastar os resultados com estudos internacionais, observa-se que metodologias semelhantes foram aplicadas em países latino-americanos como México, Brasil e Chile, bem como em contextos asiáticos como a China e as Filipinas, para modelar tendências de matrícula e projetar a procura educativa. No entanto, ao contrário destes casos, a série equatoriana apresenta maior instabilidade em certos períodos, associada a mudanças estruturais nas políticas educativas e nas situações socioeconómicas nacionais. Esta singularidade destaca a importância de adaptar modelos às particularidades locais e de não se limitar à transferência de abordagens externas. Do ponto de vista das políticas públicas, as projeções obtidas oferecem contributos valiosos para o planeamento estratégico de instituições como o Ministério da Educação e a SENPLADES, permitindo antecipar as necessidades de infraestruturas, formação de professores e alocação orçamental. Desta forma, os resultados não só contribuem para o debate académico, como também fornecem ferramentas quantitativas para a formulação de políticas educativas sustentáveis e baseadas em evidências.(García Vázquez y otros, 2021)(Mendoza Cota, 2020)

No entanto, este trabalho tem limitações. A principal reside na natureza univariada dos modelos utilizados, que impede a incorporação de variáveis exógenas relevantes, como o investimento público na educação, as condições macroeconómicas ou fatores demográficos. Neste sentido, estudos futuros poderão estender a análise aos modelos ARIMAX ou SARIMAX, incluindo covariáveis como taxa de natalidade ou despesa pública, o que permitiria uma melhor captação da dinâmica de inscrição. Além disso, embora os resultados mostrem um bom encaixe, o ASM fora da amostra mantém-se entre 3–4%, o que implica incerteza em contextos de choques estruturais como crises de saúde ou migração.(Tudela-Mamani & Grisellx, 2022)

Em resumo, os resultados deste trabalho reforçam a evidência sobre a utilização dos modelos ARIMA na educação, contribuindo tanto para o quadro teórico como para a prática do planeamento educativo no Equador. Também destaca a necessidade de explorar metodologias

híbridas – como combinações entre ARIMA e redes neurais – para melhorar a precisão das previsões e responder às limitações inerentes das abordagens lineares.(Asán Caballero e otros, 2023)

### 5.- Conclusão.

Este estudo analisou a evolução da matrícula no ensino secundário no Equador durante o período de 1971–2023, utilizando a metodologia Box–Jenkins, de modo a identificar padrões temporais e projetar cenários futuros. Os resultados mostraram que os modelos ARIMA diferenciados de primeira ordem descrevem adequadamente a dinâmica da série, destacando o ARIMA(1,1,1) como a opção com melhor desempenho segundo os critérios de informação, enquanto o ARIMA(1,1,0) com deriva oferecia uma alternativa parca e consistente. Ambos os modelos confirmaram a hipótese da estacionariedade após diferenciação e permitiram a geração de previsões robustas a curto e médio prazo.

As principais contribuições deste trabalho orientam-se para a incorporação de modelos de séries temporais na análise educativa, uma área em que a sua aplicação ainda é incipiente no Equador. O estudo mostra que as técnicas clássicas da estatística matemática, geralmente utilizadas em economia ou engenharia, são igualmente válidas para problemas sociais, fornecendo evidências quantitativas sobre a sustentabilidade da cobertura do ensino secundário. Desta forma, contribui para colmatar a lacuna identificada na literatura relativamente à utilização de metodologias de predição educativa baseadas em séries temporais.

Do ponto de vista prático, os resultados sugerem que a matrícula no ensino secundário tenderá a estabilizar-se em torno dos 100% na próxima década, o que terá implicações diretas para o planeamento de recursos, infraestruturas e políticas educativas orientadas para além da cobertura, priorizando a qualidade e a equidade. Teoricamente, o estudo reforça a relevância dos modelos ARIMA como ferramenta para a modelação de fenómenos educativos, lançando as bases para desenvolvimentos subsequentes que integram abordagens multivariadas ou híbridas.

Por fim, recomenda-se que futuras investigações estendam a análise a modelos ARIMAX ou SARIMAX que incorporem variáveis exógenas como despesa pública, taxas de natalidade ou indicadores macroeconómicos, bem como metodologias híbridas que combinem ARIMA com algoritmos de aprendizagem automática. Estas abordagens tornarão possível captar a complexidade do sistema educativo de forma mais abrangente e melhorar a precisão das previsões, reforçando a ligação entre a estatística matemática e a tomada de decisões nas políticas públicas.

Em resumo, este trabalho constitui um dos primeiros esforços no Equador para aplicar rigorosamente a metodologia Box–Jenkins à análise de indicadores educativos, especificamente à evolução histórica da

matrícula no ensino secundário. Esta contribuição não só fortalece a literatura nacional num campo em que predominam estudos qualitativos ou descritivos, como também posiciona a estatística matemática como uma ferramenta fundamental para o desenho de políticas educativas baseadas em evidências. Ao abrir esta linha de investigação, estabelecem-se precedentes para futuros estudos comparativos a nível regional e global, contribuindo para a internacionalização do debate sobre a utilização de modelos de séries temporais na educação.

### 6.- Contribuições dos autores (Taxonomia dos papéis dos colaboradores - CRedit)

1. Conceptualização: Edwin Haymacaña Moreno, Leonor Alejandrina Zapata Aspiazu.
2. Curadoria de dados: Leonor Alejandrina Zapata Aspiazu.
3. Análise formal: Edwin Haymacaña Moreno, Leonor Alejandrina Zapata Aspiazu.
4. Aquisição de fundos: N/D.
5. Investigação: Edwin Haymacaña Moreno, Leonor Alejandrina Zapata Aspiazu.
6. Metodologia: Francisco Javier Duque-Aldaz, Raúl Alfredo Sánchez Ancajima.
7. Gestão de projetos: Francisco Javier Duque-Aldaz, Raúl Alfredo Sánchez Ancajima.
8. Recursos: Francisco Javier Duque-Aldaz, Leonor Alejandrina Zapata Aspiazu.
9. Software: Edwin Haymacaña Moreno, Leonor Alejandrina Zapata Aspiazu.
10. Supervisão: Félix Genaro Cabezas García, Raúl Alfredo Sánchez Ancajima.
11. Validação: Félix Genaro Cabezas García.
12. Visualização: Leonor Alejandrina Zapata Aspiazu.
13. Argumento - rascunho original: Edwin Haymacaña Moreno, Francisco Javier Duque-Aldaz.
14. Escrita - revisão e edição: Francisco Javier Duque-Aldaz, Félix Genaro Cabezas García, Raúl Alfredo Sánchez Ancajima.

### 7.- Apêndice.

Código R usado para o desenvolvimento do investigaçāo.

### Pacotes

```
install.packages(c("readxl", "dplyr", "ggplot2", "forecast", "ts
```

```
ries",
```

```
    "urca",    "TSstudio",    "broom",    "knitr",
```

```
    "kableExtra")
```

```
biblioteca (readXL); biblioteca (dplyr); Biblioteca(ggplot2)
```

```
biblioteca (previsão); biblioteca(série); Biblioteca (URCA)
```

```
biblioteca (TSstudio); biblioteca (vassoura); biblioteca
```

```
(tricot); biblioteca (kableExtra)
```

```
# 2. Ler Excel (ficheiro localizado no diretório de trabalho)
```

```
DF <- ReDXL::read_excel("matriculadosecuador.xlsx")
```

```
# 3. Revisão Rápida
```

```
str(DF)
```

Resumo(df)

```
# 4. Criar Objeto de Série Temporal (Anualmente)
e <- ts(df$Matriculacion, start=min(df$Year), frequência =
1)
```

# Exploração Inicial

```
autoplot(y) +
  labs(title="Inscrição Secundária (% Bruto) – Equador",
       x="Ano", y="%") +
  theme_minimal(base_size = 12)
```

# ACF e PACF

```
ggAcf(y, lag.max = 30) + theme_minimal()
ggPacf(y, lag.max = 30) + theme_minimal()
```

###Diagnóstico da estacionariedade e das transformações

# Testes Raiz Unitários

```
tseries::adf.test(y) # H0: raiz unitária (não estacionária)
tseries::kpss.test(y) # H0: estacionaria (si p<0.05, no
estacionaria)
```

# Ordem sugerida de diferenciação

```
previsão::ndiffs(y) # normalmente 1
```

###Diferenciar uma vez (d = 1) e voltar a testar

```
#y_ts: A Sua Série Anual Já Criada, Frequência 1 (1971–
2023)
y_ts <- ts(df$Matriculacion, start = min(df$Año),
frequência = 1)
```

#1) Diferença de Primeira Ordem

```
Dy <- diff(y_ts)
```

#2) Visualizar a série diferenciada

```
autoplot(dy) +
  labs(título = "Primeira diferença na matrícula do ensino
secundário (%)",
       x = "Ano", y = "Inscrição Δ (%)") +
  theme_minimal(base_size = 12)
```

#3) Testes de estacionariedade na série diferenciada

```
adf.test(dy) # H0: raiz unitária (não estacionária)
kpss.test(dy, null = "Level")# H0: estacionário no level
```

#4) Estrutura temporal da série diferenciada

```
ggAcf(dy, lag.max = 30) + theme_minimal(base_size = 12)
ggPacf(dy, lag.max = 30) + theme_minimal(base_size = 12)
```

###Identificación/estimação do modelo (candidatos + auto.arima)

# Pesquisa exaustiva (sem sazonalidade)

```
fit_auto <- auto.arima(y_ts,
  sazonais = FALSE, # anual
```

completa passo a passo = FALSO, # pesquisa mais

aproximação = FALSO,
d = 1) # já sabemos que d = 1

fit\_auto

#Luego, testámos alguns candidatos clássicos e comparámos por AIC, AICc, BIC:

cand <- lista(

```
ARIMA_011 = Arima(y_ts, ordem = c(0,1,1)),
ARIMA_110 = Arima(y_ts, ordem = c(1,1,0)),
ARIMA_111 = Arima(y_ts, ordem = c(1,1,1)),
ARIMA_210 = Arima(y_ts, ordem = c(2,1,0)),
ARIMA_012 = Arima(y_ts, ordem = c(0,1,2))
)
```

```
CMP <- data.frame(
  Modelo = nomes(cand),
  AIC = sapply(cand, AIC),
  BIC = sapply(cand, BIC)
)
```

) Impressão (CMP)

###Diagnóstico de desperdício do modelo escolhido

# Diagnóstico abrangente

```
checkresiduals(fit_auto) # inclui Ljung–Box, ACF
Residuals e QQ-plot
# Se quiseres Ljung–Box explícito com vários lags:
Box.test(residuals(fit_auto), lag = 10, tipo = "Ljung")
Box.test(residuals(fit_auto), lag = 15, tipo = "Ljung")
```

###Validación fora da amostra (treino/teste) e métricas

# Partição temporária

```
y_tr <- janela(y_ts, fim = 2016)
y_te <- janela(y_ts, início = 2017)
```

```
fit_tr <- auto.arima(y_tr, sazonal = FALSE, passo a passo =
FALSE, aproximação = FALSE, d = 1)
fc_te <- forecast(fit_tr, h = comprimento(y_te))
```

# Métricas de Validação

Precisão (fc\_te, y\_te)

```
autoplot(fc_te) + autolayer(y_te, série = "Observado") +
  labs(título = "Previsão fora da amostra (2017–2023)",
       y = "% de matrícula") + theme_minimal(base_size =
12)
```

###Pronóstico fim (h = 5–10 anos)

# Retreina com toda a série e prevê

```
fit_all <- auto.arima(y_ts, sazonal = FALSE, passo a passo =
FALSE, aproximação = FALSE, d = 1)
fc_10 <- previsão(fit_all, h = 10)
```

autoplot(fc\_10) +

laboratórios (título = "Previsão de Matrícula do Ensino Secundário (%) – Equador",  
 $x = \text{"Ano"}, y = \text{"%}"$  ) +  
 theme\_minimal(base\_size = 12)

## 7.- Referências.

- [1] C. T. Simonino, L. F. dos Santos, e A. F. de Freitas, "Implementação do Programa Nacional de Alimentação Escolar num contexto pandémico no município de Viçosa/MG: implicações para a gestão pública e a agricultura familiar," *Administração Pública e Gestão Social*, vol. 17, n.º 1, 2025. URL: <https://www.redalyc.org/journal/3515/351580876003/>
- [2] M. Zanatta Idemori, A. R. Nogueira de Macedo, P. M. Galvão de Souza e O. S. Gaion, "A RELAÇÃO ENTRE GAMIFICAÇÃO E EXPERIÊNCIA EM APLICAÇÕES DE ENTREGA," *Revista Pensamento Contemporâneo em Administração*, vol. 19, n.º 1, pp. 140–157, 2025. DOI: [https://doi.org/10.12712/rpca.v19.65109\\_UFF+1 Revistas](https://doi.org/10.12712/rpca.v19.65109_UFF+1 Revistas)
- [3] D. Escolar, "Conflito e Identificações Etno-Raciais no Prelúdio Colonial do Caudillismo e das Rebeliões Montonero," *Bulletin of the Institute of Argentine and American History Dr. Emilio Ravignani*, n.º 61, pp. 85–91, 2024. DOI: <https://doi.org/10.34096/bol.rav.n61.14555 SciELO Argentina>
- [4] F. O. Corrêa Werle e J. A. Lago Fonseca, "Gestão escolar: fundamentos e discussões sobre gestão democrática," *Ensaio: Avaliação e Políticas Públicas na Educação*, e0254968, 2025. DOI: <https://doi.org/10.1590/S0104-40362025003304968 Educa FCC>
- [5] G. E. Castro Rosales, A. D. Torres Alvarado, L. S. Zalamea Cedeño, F. J. Duque-Aldaz e F. R. Rodríguez-Flores, "Proposta Ergonómica Abrangente para a Redução dos Riscos Musculoesqueléticos na Produção de Sabão: Uma Abordagem Baseada na Análise Estatística e Avaliação Postural," *INQUIDE – Ingeniería Química y Desarrollo*, vol. 7, n.º 2, 2025. DOI: <https://doi.org/10.53591/inqd.v7i02.2416 Revistas UG>
- [6] M. J. Cabrera Valladolid, "Análise da Série Temporal de Casos Confirmados e Mortos de Covid-19 no Equador," Escuela Superior Politécnica de Chimborazo, 2021. URL: <https://dspace.espoch.edu.ec/handle/123456789/14812>
- [7] M. G. Zalduaromero, "Estimativa de uma função de procura de importação na indústria do plástico no Equador, período 1980–2016," UESS, 2017. URL: <http://repositorio.uees.edu.ec/123456789/2008>
- [8] J. I. Cañarte Murillo, "Crescimento restrito da balança de pagamentos: caso equatoriano da Lei Thirlwall 2000–2016," UESS, 2017. URL: <http://repositorio.uees.edu.ec/123456789/1980>
- [9] M. Alós e M. Serio, "Participação no mercado de trabalho dos idosos na Argentina," *Pilquen Journal – Secção de Ciências Sociais*, vol. 27, n.º 2, pp. 71–96, 2024. URL: <https://www.redalyc.org/journal/3475/347580385004/>
- [10] F. J. Duque-Aldaz e E. G. Pazan Gómez, "Fatores que afetam a intenção empreendedora dos estudantes universitários séniores," *Espacio*, vol. 39, n.º 09, p. 32, 2017. URL: <https://www.revistaespacios.com/a18v39n09/18390932.html>
- [11] A. Meneses Freire, L. Zúñiga Lema, J. Muñoz Cargua, J. Lara Sinaluisa e W. Acurio Martínez, "Modelo preditivo funcional não paramétrico em séries temporais funcionais. Aplicação em variáveis meteorológicas," Escuela Superior Politécnica de Chimborazo, 2022. URL: <https://dspace.espoch.edu.ec/handle/123456789/19024>
- [12] L. Y. Ortega Villegas, "Análise na indústria naval no Equador como resultado da crise financeira de 2008," UESS, 2018. URL: <http://repositorio.uees.edu.ec/123456789/2503>
- [13] D. M. Ichau Tabango, P. J. Pando Sigua, P. L. Cárdenas García e S. V. Avilés Sacoto, "Procura e gestão de inventário para a criação de um sistema automatizado de gestão da informação: Estudo de caso aplicado a um supermercado equatoriano," USFQ, 2021. URL: <http://repositorio.usfq.edu.ec/handle/23000/10650>
- [14] F. J. Duque-Aldaz, F. R. Rodríguez-Flores e J. Carmona Tapia, "Identificação de parâmetros em sistemas de equações diferenciais ordinárias usando redes neurais artificiais," *San Gregorio*, vol. 1, n.º 2, 2025. URL: <https://revista.sangregorio.edu.ec/index.php/REVISTASANGREGORIO/article/view/2826>
- [15] F. F. Sandoya Sanchez e A. G. Abad Robalino, "Comparação da precisão da previsão de dados de uma série temporal entre modelos ARIMA e Redes Neurais," ESPOL, 2017. URL: <http://www.dspace.espol.edu.ec/xmlui/handle/123456789/40345>
- [16] A. F. Mayorga Trujillo, "Investigação sobre o aumento da produtividade na fábrica de parafusos da empresa Galo G. Orbea O. Cía. Ltda. através da análise da disponibilidade nas fases do seu processo produtivo," Escuela Superior Politécnica de Chimborazo, 2017. URL: <https://dspace.espoch.edu.ec/handle/123456789/6515>
- [17] D. A. Villarreal Godoy, L. F. Cónor Simbaña e C. A. Súarez Nuñez, "Controlo de Inventário numa Empresa Têxtil Equatoriana: Melhoria nos Processos de Planeamento e Controlo da Produção," USFQ, 2022. URL: <http://repositorio.usfq.edu.ec/handle/23000/12544>
- [18] N. Vela e M. C. Camacho Cordovez, "Indústria Petrolífera: Modelação da Previsão da Procura e Planeamento e Projeção de Consumo," USFQ, 2020. URL: <http://repositorio.usfq.edu.ec/handle/23000/9206>
- [19] J. S. Fu-López, J. P. Fierro Aguilar, F. R. Rodríguez-Flores e F. J. Duque-Aldaz, "Aplicação de estratégias Lean não automatizadas para melhoria da qualidade em processos manuais de montagem: Um estudo de caso na indústria de eletrodomésticos," *INQUIDE – Ingeniería Química y Desarrollo*, vol. 7, n.º 1, 2025. DOI: <https://doi.org/10.53591/inqd.v7i02.2417>
- [20] J. M. Ausay Carrillo, "Proposta de um modelo econométrico para a previsão de vendas da fábrica de rações EXIBAL-Riobamba," Escuela Superior Politécnica de Chimborazo, 2022. URL: <https://dspace.espoch.edu.ec/handle/123456789/19903>
- [21] P. A. Eguiguren Calisto e S. V. Avilés Sacoto, "Modelos de projeção de procura para produtos de alta volatilidade e baixo volume em vendas dentro de uma empresa alimentar," USFQ, 2019. URL: <http://repositorio.usfq.edu.ec/handle/23000/8320>
- [22] D. S. Navarro Llivilisaca, "Influência do IVA e do imposto sobre o rendimento nas cobranças de impostos, 2007–2016," UESS, 2017. URL: <http://repositorio.uees.edu.ec/123456789/1773>
- [23] I. J. Figueroa Tigrero, "Pressão Fiscal do Equador, Período 2000–2018: Uma Análise a partir da Abordagem da Teoria da Curva de Laffer," UESS, 2019. URL: <http://repositorio.uees.edu.ec/123456789/3090>
- [24] A. M. Freire Engracia, A. J. Santos Castello e I. D. Rivadeneyra, "Análise da relação entre investimento direto estrangeiro e cobrança de impostos: Efeitos no setor comercial," ESPOL, 2025. URL: <http://www.dspace.espol.edu.ec/handle/123456789/65933>
- [25] K. N. Lema Remache, "Comparação entre métodos ARIMA e Tyler Kriging (TK) para previsão de séries temporais da velocidade do vento: Estudo de caso da Estação Meteorológica de Chimborazo," Escuela Superior Politécnica de Chimborazo, 2024. URL: <https://dspace.espoch.edu.ec/handle/123456789/22839>
- [26] G. J. Morochó Choca, L. Á. Bucheli Carpío e F. J. Duque-Aldaz, "Otimização do despacho de combustível de combustível através de regressão multivariada usando indicadores locais de armazenamento," *INQUIDE*, vol. 6, n.º 2, 2024. DOI: <https://doi.org/10.53591/inqd.v6i02.477>
- [27] J. E. Pincay Moran, A. F. López Vargas, F. J. Duque-Aldaz, W. Villamagua Castillo, e R. Sánchez Casanova, "Avaliação e Proposta para um Sistema de Gestão Ambiental numa Plantação de Manga," *INQUIDE*, vol. 7, n.º 1, 2025. DOI: <https://doi.org/10.53591/inqd.v7i01.1991>
- [28] R. J. Guerrero Quinde e C. C. Pérez Siguenza, "Design de um modelo de fornecimento de produtos lácteos numa empresa de retalho," ESPOL, 2025. URL: <http://www.dspace.espol.edu.ec/handle/123456789/66133>
- [29] G. Silva e L. C. Di Serio, "Inovação em 'Negócios Esquecidos'", *Innovation & Management Review*, vol. 18, n.º 4, pp. 350–364, 2021. URL: <https://www.redalyc.org/journal/5375/537569183001/>
- [30] F. Medeiros, L. C. Di Serio, e A. Moreira, "Avon Brasil: Otimização de Processos Logísticos numa Empresa de Venda Direta," *Revista de Administração Contemporânea*, vol. 25, n.º 4, e-190212, 2021. DOI: <https://doi.org/10.1590/1982-7849rac2021190212.en>
- [31] L. M. García-Feria, L. Aguilar-Faisal, R. Pastor-Nieto e J. C. Serio-Silva, "Alterações na vegetação em pequenas escalas paisagísticas e cativeiro alteram a microbiota intestinal dos macacos-gritos (*Alouatta pigra*: Atelidae)," *Acta Biológica Colombiana*, vol. 28, n.º

1, pp. 154–164, 2023. **DOI:**  
<https://doi.org/10.15446/abc.v28n1.93450>

- [32] C. A. García Vázquez, A. I. González Santos, e V. Pérez Garrido, "Metodología e algoritmo de validação para identificar modelos de uma unidade de tratamento de ar," *Electronic, Automatic and Communications Engineering*, vol. 42, nº 1, 2021. **URL:** [http://scielo.sld/scielo.php?script=sci\\_arttext&pid=S1815-59282021000100105&lang=es](http://scielo.sld/scielo.php?script=sci_arttext&pid=S1815-59282021000100105&lang=es)
- [33] J. E. Mendoza Cota, "COVID-19 e Emprego no México: Impacto Inicial e Previsões de Curto Prazo," *Accounting and Administration*, vol. 65, nº 5, 2020. **URL:** <https://dialnet.unirioja.es/servlet/articulo?codigo=7731039>
- [34] J. W. Tudela-Mamani e G. A.-M. Grisellx, "Impacto da COVID-19 na Procura Internacional de Turismo do Peru: Uma Aplicação da Metodologia Box-Jenkins," *Revista de Investigaciones Altoandinas*, vol. 24, nº 1, 2022. **DOI:** <https://doi.org/10.18271/ria.2022.317>
- [35] L. Asán Caballero, J. Rojas Delgado e G. E. Jiménez Moya, "Algoritmo de Previsão de Séries Temporais para Previsão de Tráfego Aéreo Baseado em Redes Neurais Artificiais," *Cuban Journal of Computer Sciences*, vol. 16, nº 4, 2023. **URL:** [http://scielo.sld/scielo.php?script=sci\\_arttext&pid=S2227-18992022000400084&lang=es](http://scielo.sld/scielo.php?script=sci_arttext&pid=S2227-18992022000400084&lang=es)