

Aplicação de técnicas de mineração de dados no desempenho operacional de concessionárias de ferrovias: Um estudo de caso do SIADE

Application of data mining techniques in the operational performance of railway concessionaires: A SIADE case study

Augusto Rodrigo Camblor Santos 

Fatec Santos
augusto.santos14@fatec.sp.gov.br

Lais Tebas Santana 

Fatec Santos
lais.santana3@fatec.sp.gov.br

Karina Barros Oliveira 

Fatec Santos
kbarros.o@hotmail.com

Samuel Alves Charadias 

Fatec Santos
samuel.charadias@fatec.sp.gov.br

Vagner dos Santos Macedo 

Fatec Santos
vagner.macedo2@fatec.sp.gov.br

RESUMO

A presente pesquisa objetiva a aplicação de técnicas de mineração de dados, a partir da metodologia CRISP-DM, no desempenho operacional de concessionárias, empregando-se dados históricos que contemplam, desde 2013 a 2023, a movimentação mensal de cargas nas ferrovias federais concedidas, oriundos do subsistema integrado do Sistema de Acompanhamento e Fiscalização do Transporte Ferroviário (SAFF), o Sistema de Acompanhamento do Desempenho Operacional das Concessionárias (SIADE). Sendo assim, propõe-se a utilização de mineração de dados para análise exploratória descritiva preliminar em linguagem de programação Python, apresentando-se, por meio de gráficos de colunas, a movimentação ferroviária brasileira em toneladas úteis e posterior modelagem preditiva em série temporais, com o emprego do método ARIMA (Autorregressivo Integrado de Médias Móveis), visando a identificação de eventuais tendências, sazonalidade e previsão de desempenho operacional das concessionárias ferroviárias contempladas nesse estudo. Verifica-se que, durante o período analisado, o minério de ferro manteve-se como o produto de maior movimentação em unidade de toneladas úteis (aproximadamente 381 milhões de TU, em 2023) e dentre as concessionárias, a MRS e a Vale apresentaram maior produção em TU em 2023. Ademais, obteve-se previsões que possibilitaram mensurar a expectativa de movimentação de carga ferroviária para cada concessionária no ano de 2024.

PALAVRAS-CHAVE: CRISP-DM; Mineração de dados; SIADE; Transporte ferroviário.

ABSTRACT

The current research aims to apply data mining techniques, based on the CRISP-DM methodology to the operational performance of concessionaires, using historical data that includes, from 2013 to 2023, the monthly movement of cargo on concession federal railways, originated from the integrated subsystem of the Rail Transport Monitoring and Inspection System (SAFF), the Concessionaires Operational Performance Monitoring System (SIADE). Therefore, it is proposed to use data mining for preliminary descriptive exploratory analysis in Python programming language, presenting, through column graphs, Brazilian railway movement in useful tons and subsequent predictive modeling in time series, with the use of the ARIMA (Autoregressive Integrated Moving Average) method, aiming to identify possible trends, seasonality and forecast the operational performance of the railway concessionaires included in this study. It can be seen that, during the period analyzed, iron ore remained the product with the largest movement in units of useful tons (approximately 381 million TU, in 2022) and among the concessionaires, MRS and Vale presented the highest production in TU in 2023. Furthermore, predictions were obtained, making it possible to measure the expected movement of railway cargo for each concessionaire for year 2024.

KEY-WORDS: CRISP-DM; Data mining; SIADE; Rail transport.

INTRODUÇÃO

As ferrovias representam um ponto fundamental na infraestrutura de transporte, desempenhando papel vital na movimentação de cargas em todo o mundo ao possibilitar o transporte de cargas em grande quantidade, com baixo custo de frete, maior segurança, economia e sustentabilidade, em comparação ao transporte rodoviário (DIAS, 2012; ALVES; RAMOS; SILVA, 2020).

No contexto das ferrovias federais concedidas, a eficiência operacional e o gerenciamento eficaz da movimentação de cargas são de importância crítica para atender às demandas crescentes da indústria e da economia, visto que a Malha Ferroviária Brasileira representa um dos principais PIB do país. Portanto, a concessão de operação e manutenção de ferrovias a empresas privadas demanda monitoramento e avaliação do desempenho operacional das concessionárias, garantindo que os objetivos de segurança, eficiência e produtividade sejam alcançados (ALVES; RAMOS; SILVA, 2020).

Dessa forma, o Sistema de Acompanhamento do Desempenho Operacional das Concessionárias (SIADE) surgiu como uma ferramenta valiosa para monitorar, coletar e analisar dados relacionados à operação ferroviária, oriundo do Ministério dos Transportes. Não obstante, com o aumento exponencial na quantidade de dados gerados por essas operações, há a necessidade de aplicar métodos avançados de análise, como a mineração de dados, a fim de extrair informações significativas e *insights* valiosos a partir desses registros, pois a mineração de dados oferece a capacidade de identificar padrões, tendências e anomalias nos dados de

movimentação de cargas, melhorando a tomada de decisões e otimizando os processos operacionais.

Assim, este estudo propôs-se a explorar a aplicação da mineração de dados, partindo da metodologia *Cross-Industry Standard Process for Data Mining* (CRISP-DM, traduzida para o português como Processo Padrão Intersetorial para Mineração de Dados) no contexto da movimentação de cargas nas ferrovias federais concedidas, utilizando o SIADE como estudo de caso. Assim, utilizou-se dados históricos que contemplam, desde 2013 a 2023, a movimentação mensal de cargas nas ferrovias federais concedidas, empregando-se, posteriormente, técnicas mineração de dados para análise descritiva preliminar em linguagem de programação Python, apresentada através de gráficos de colunas a movimentação ferroviária brasileira em toneladas úteis e posterior modelagem preditiva em série temporais, com o emprego do método ARIMA (*Autoregressive Integrated Moving Average*, sendo em português, Autorregressivo Integrado de Médias Móveis), visando possibilitar a identificação de eventuais tendências, sazonalidade e previsão de desempenho operacional das concessionárias ferroviárias contempladas nesse estudo, propondo-se a investigação quanto à relevância do emprego da mineração de dados no aprimoramento da gestão da movimentação de cargas, bem como incentivo à tomada de decisão embasada em dados.

Nas seções deste artigo, apresenta-se as metodologias e técnicas de mineração de dados aplicadas, os resultados obtidos e suas implicações para o setor ferroviário, visando contribuir para o aprimoramento contínuo das operações ferroviárias concedidas, destacando o potencial da mineração de dados como uma ferramenta estratégica para aprimorar a gestão e o desempenho operacional nesse contexto em evolução.

1. FUNDAMENTAÇÃO TEÓRICA

A fundamentação teórica do presente artigo circunda o conceito histórico das fontes de pesquisa que promoveram embasamento teórico inicial ao mesmo e motivaram a escolha do tema, além da ferramenta tecnológica utilizada para a elaboração das análises, especificamente, a ferramenta *Google Colab*.

1.1 MALHA FERROVIÁRIA BRASILEIRA

As ferrovias foram introduzidas no Brasil em 1854, quando a Estrada de Ferro Mauá inaugurou a primeira linha férrea no país, conectando o Rio de Janeiro a Petrópolis. Esse marco inicial desencadeou rápido processo de expansão ferroviária nas décadas seguintes, impulsionado pela necessidade de transportar produtos agrícolas, minerais e mercadorias entre as regiões produtoras e os portos para exportação (SGROTT; ZAGHENI, 2020; ALVES; RAMOS; SILVA, 2020).

O Ciclo do Café, iniciado no século XIX e com fim no século XX, na década de 1930, que propiciou grande desenvolvimento econômico no Brasil, ocasionou um grande impulso na construção de ferrovias. O transporte ferroviário se tornou vital para o escoamento da safra de café do interior de São Paulo para os portos do Rio de Janeiro, consolidando a importância das ferrovias como uma espinha dorsal do transporte de mercadorias (MILLIET, 1941).

Durante o século XX, as ferrovias foram dominantes no transporte de cargas. Contudo, o advento da rodovia e a falta de investimentos no setor ferroviário levaram a um declínio acentuado na rede ferroviária. Sendo assim, linhas foram desativadas, e a indústria ferroviária sofre, atualmente, com problemas de infraestrutura e falta de modernização (LANZA, 2020; SGROTT, 2019).

Todavia, há a Associação Nacional dos Transportadores Ferroviários (ANTF), que é uma instituição nacional sem fins lucrativos, criada para defender e promover o desenvolvimento e aprimoramento do transporte de carga por ferrovia no país, desde o processo de desestatização, realizado pelo Governo Federal em 1996 (ANTF, 2023).

Atualmente, a ANTF representa as operadoras ferroviárias responsáveis pelo transporte de carga em 14 malhas concedidas à iniciativa privada, cuja extensão abrange mais de 31 mil km e por onde milhões de toneladas de minério, soja, milho, açúcar, produtos siderúrgicos, entre outros, circulam anualmente (ANTF, 2023).

As ferrovias das empresas associadas à ANTF (Bamin, FTC, MRS, Rumo, TLSA, Vale e VLI) conectam o Quadrilátero Ferrífero, no sul de Minas Gerais, e outros centros de mineração e siderurgia, além dos maiores polos industriais e áreas agrícolas aos principais portos brasileiros, entre eles, os de Santos, no estado de São Paulo, de Itaqui, no Maranhão, Vitória, no Espírito Santo, e o do Rio de Janeiro (ANTF, 2023; GABRIEL, 2021).

Apesar de apresentar cerca de 31 mil km de malha ferroviária, no cenário internacional, o Brasil ainda apresenta baixa densidade da malha ao ser comparado a países de dimensões

continentais, como Estados Unidos da América, Canadá, Índia e China, com cerca de 293 mil km, 77 mil km, 108 mil km, 141 mil km, respectivamente (ANTF, 2023).

A fim de aumentar a competitividade do Brasil no mercado internacional, reduzir os custos logísticos e melhorar a integração entre as regiões, em 2019, o governo federal lançou o Programa de Parcerias de Investimentos (PPI), que incluiu um conjunto de projetos ferroviários, como o da Ferrovia Transnordestina, que ligará o Porto de Pecém, no Ceará, ao Porto de Suape, em Pernambuco; da Ferrovia de Integração Oeste-Leste (Fiol), que ligará o Porto de Ilhéus, na Bahia, ao município de Figueirópolis, em Tocantins, conectando-se então com a Ferrovia Norte-Sul e o restante do país; e da Ferrovia Norte-Sul, que ligará o Porto de Santos, em São Paulo, ao Porto de Barcarena, no Pará (BRASIL, 2023).

Além disso, há os Programas de Concessões e Renovações Ferroviárias, cuja unidade responsável é a Secretaria de Fomento e Parceria (SFPP). O Programa de Concessões Ferroviárias visa transferir a gestão da infraestrutura ferroviária para o setor privado, sendo assim, empresas privadas serão responsáveis por ampliar, manter e explorar as ferrovias brasileiras, objetivando-se o aprimoramento de eficiência e qualidade do serviço ferroviário, bem como redução dos custos logísticos. Já o Programa de Renovações Ferroviárias almeja renovar antecipadamente os contratos com as concessionárias ferroviárias atuais, possibilitando que as mesmas realizem investimentos em melhorias na infraestrutura ferroviária, como a modernização de equipamentos e a ampliação da capacidade de transporte, atraindo investimentos e modernizando a malha ferroviária brasileira (BRASIL, 2023).

1.2 CRISP-DM

O CRISP-DM, significa em língua inglesa "Cross-Industry Standard Process for Data Mining" (Processo Padrão Intersetorial para Mineração de Dados), e é uma metodologia amplamente reconhecida e utilizada para orientar o processo de mineração de dados, sendo essa uma ciência que usa diversas técnicas descritivas e preditivas – como classificação, regressão linear e não-linear, agrupamento, entre outras –, na junção e organização de grandes volumes de dados para extrair conhecimento desses, buscando-se padrões, associações, mudanças e anomalias relevantes. Dessa forma, a metodologia CRISP-DM fornece uma estrutura clara e consistente para supracitado processo, tendo sido desenvolvida para auxiliar profissionais e pesquisadores a planejar, executar e avaliar projetos de mineração de dados de forma sistemática e eficaz, sendo composta por seis fases interrelacionadas que ocorrem de forma

cíclica: compreensão do negócio, compreensão dos dados, preparação dos dados, modelagem, avaliação e implantação (BARBIERI, 2011 *apud* AMEIDA, 2017; CAMILO; SILVA, 2009; ROBERTO, 2023).

Na etapa de entendimento do negócio, propõe-se a compreensão das necessidades do negócio e o problema que a mineração de dados pode resolver (GAMA NETO, 2018).

No entendimento dos dados, ocorre a avaliação dos dados disponíveis, sua qualidade e como os mesmos podem ser usados para a eficaz resolução do problema (GAMA NETO, 2018).

Na etapa de preparação dos dados, os dados são preparados para a modelagem, realizando-se a limpeza, transformação e redução de dimensionalidade (GAMA NETO, 2018).

Na modelagem, aplica-se as técnicas de mineração de dados, propondo-se a criação de modelos para identificação de padrões e particularidades dos dados analisados (IBM, 2023; GAMA NETO, 2018).

Na avaliação do modelo, o desempenho dos modelos é avaliado, a fim de determinar qual deles é o melhor para a resolução do problema (GAMA NETO, 2018).

Na etapa de implementação, o modelo é implantado para produção, com planejamento e monitoramento e, eventual, revisão de projeto (IBM, 2023; GAMA NETO, 2018).

1.3 GOOGLE COLAB

O *Google Colaboratory*, comumente denominado *Google Colab*, é um serviço de nuvem gratuito, lançado em 2017 pela empresa multinacional americana especializada em tecnologia, Google, sendo proveniente do *Google Research*, setor da Google especializado no incentivo e disseminação de pesquisas científicas (GOOGLE, 2023).

O produto da Google possibilita a composição, execução e compartilhamento de código em Python no navegador, ideal para aprendizado de máquina, ensino e análise de dados, não sendo necessário o download e instalação de *softwares* para sua utilização, o que o torna um instigante *Integrated Development Environment* (IDE), traduzido para a língua portuguesa como Ambiente de Desenvolvimento Integrado, para programação na linguagem de programação de alto nível, Python (GOOGLE, 2023; GUNAWAN et al, 2020).

2. PROCEDIMENTOS METODOLÓGICOS

O desenvolvimento da temática utilizou a análise descritiva exploratória, sendo essa uma técnica estatística que busca descrever e resumir os principais aspectos de um conjunto de dados, visando a obtenção de informações relevantes e compreensíveis sobre as características, distribuição e padrões presentes nos dados, empregando-se a metodologia CRISP-DM para norteamento (MARCONI; LAKATOS, 2003; BUSSAB; MORETTIN, 2017).

Por conseguinte, para o alcance do objetivo proposto, o projeto foi desenvolvido em seis etapas: compreensão do negócio, compreensão dos dados, preparação dos dados, modelagem, avaliação e implantação.

A coleta de dados remete às etapas de compreensão e preparação dos dados, sendo os mesmos provenientes do SIADE, preconizado pela Agência Nacional de Transportes Terrestres (ANTT), do período de janeiro de 2013 a dezembro de 2023 (BRASIL, 2023a).

A transformação e processamento dos dados ocorreu no ambiente do Google Colab. Logo, realizou-se a integração dos dados ao ambiente integrado, Google Colab, sendo essa a ferramenta selecionada para a elaboração do algoritmo provisor dos gráficos pertinentes ao estudo, desde a extração dos mesmos da sua fonte, seguida pela transformação e o seu carregamento, cujo processo é conhecido por ETL (*Extract, Transform, Load*). Ademais, determinou-se os indicadores e cálculos pertinentes à pesquisa, personalizando-se formatos de dados e categorias, quando prudente, a fim de promover o enriquecimento dos dados para a extração de informações adicionais (LAGO; LAENNDER, 2018).

A elaboração de resumos numéricos e gráficos no Google Colab realizou-se com o auxílio das bibliotecas pandas e matplotlib, para manipulação de dados tabulares e criação de gráficos, e funções glob e os, para listagem de arquivos nos diretórios e obtenção de informações sobre os arquivos.

Para as fases de modelagem, avaliação dos modelos e implementação, empregou-se as técnicas de mineração preditivas convenientes para previsões de séries temporais, assim, utilizou-se a biblioteca sklearn para métricas de avaliação e a pmdarima para autoajuste do modelo ARIMA, sendo esse um modelo estatístico linear que combina dependências de autocorrelação e médias móveis, vastamente usado para modelar e prever séries temporais, além de sua extensão SARIMA (*Seasonal AutoRegressive Integrated Moving Average*, traduzido para o português como Autorregressivo Integrado de Médias Móveis com Sazonalidade) (KOTU; DESHPANDE, 2019).

Salienta-se que a autocorrelação é a dependência entre os valores de uma série temporal em diferentes momentos, enquanto as médias móveis representam a dependência entre os valores de uma série temporal e seus valores passados médios. (THE PENNSYLVANIA STATE UNIVERSITY, 2023)

Logo, dividiu-se os dados em treinamento e teste para validação cruzada, e a biblioteca pmdarima realizou o teste de diferentes combinações de valores para os parâmetros de ordem do modelo, sendo esses p (ordem autoregressiva), d (ordem de diferenciação) e q (ordem da média móvel), selecionando a que minimiza o AIC (Critério de Informação de Akaike) ou o BIC (Critério de Informação Bayesiano) que são medidas de qualidade de modelo estatístico que penalizam modelos superparametrizados (SMITH, 2023).

Dessa forma, com a previsão de séries temporais usando o modelo ARIMA, bem como o modelo SARIMA, que incorpora componentes sazonais, considerando-se padrões sazonais que se repetem em intervalos fixos de tempo, fez-se uma análise da série temporal, almejando a compreensão de eventuais tendências, sazonalidades e padrões, prevendo-se a movimentação de carga ferroviária de distintas concessionárias para o ano de 2024 (THE PENNSYLVANIA STATE UNIVERSITY, 2023).

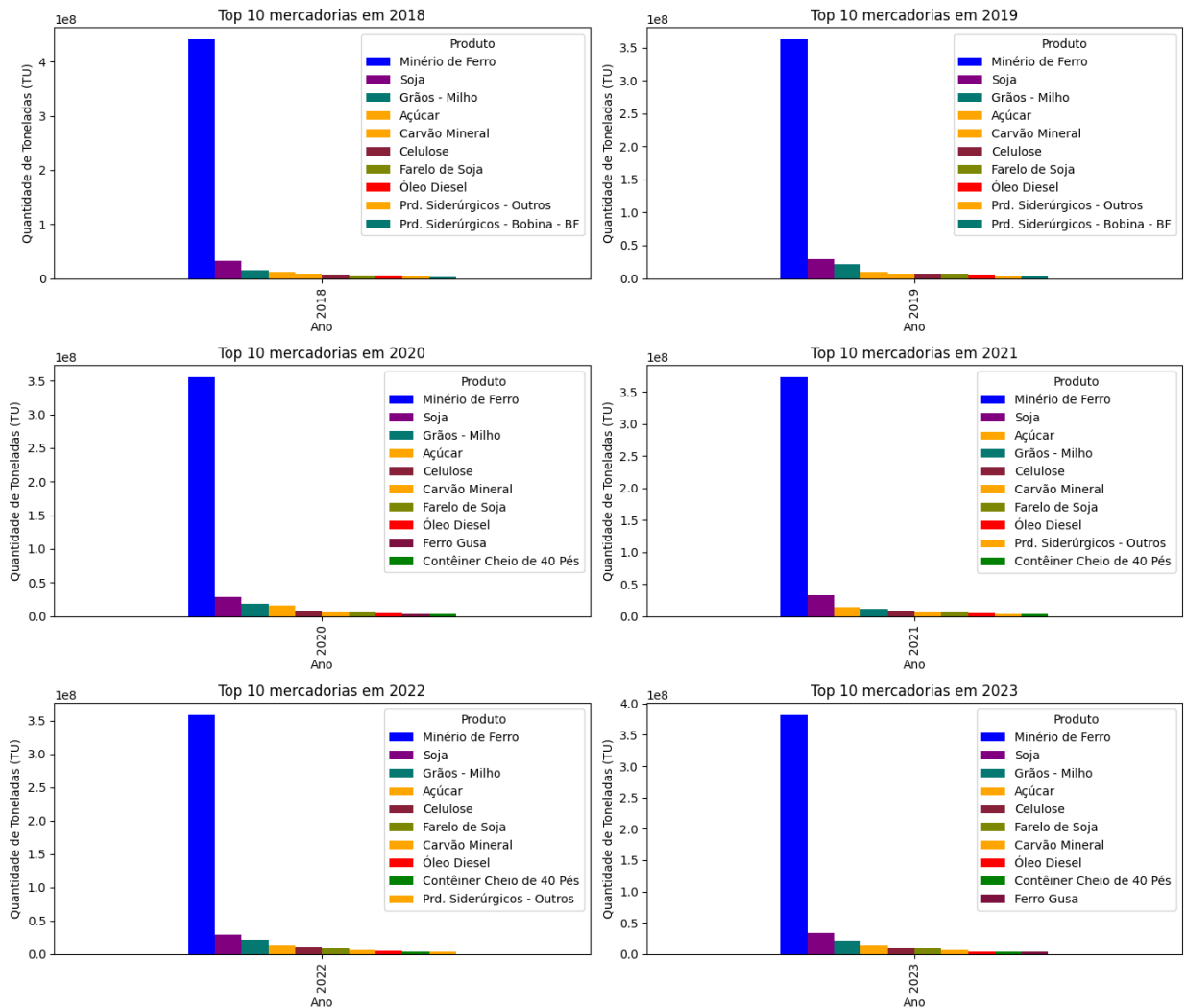
3. RESULTADOS E DISCUSSÃO

No período analisado, o ano com maior movimentação de carga total em toneladas úteis (TU) foi o ano de 2018 (571.621.071 TU), seguido pelos anos de 2017 (540.237.780 TU), 2023 (530.596.848 TU) e 2021 (506.796.406 TU).

Os gráficos elaborados durante a análise preliminar, permitiram a verificação da quantidade de toneladas por mercadoria transportada anualmente no período de 2013 a 2023, possibilitando a constatação de que o minério de ferro foi o produto mais transportado, destacando-se, em conjunto com soja, milho e açúcar, como os principais produtos em todo o período estudado.

Entretanto, na Figura 1, apresenta-se os gráficos de colunas da quantidade de TU por produto para os anos de 2018 a 2023, nos quais é possível confirmar a vasta discrepância entre a movimentação do minério de ferro e dos demais produtos, como soja, açúcar, milho, celulose e carvão.

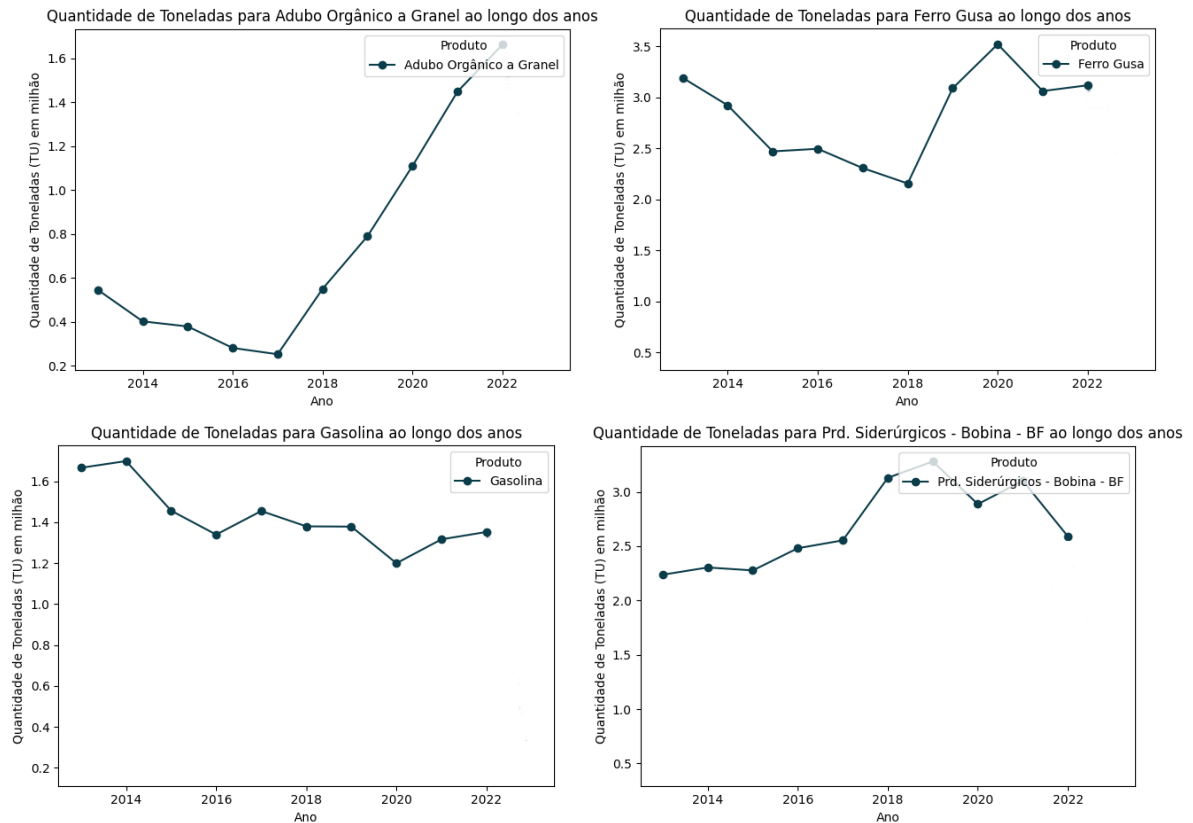
Figura 1 – Movimentação de cargas em TU por Produto no ano de 2018 a 2023



Fonte: Autores (2023)

Atesta-se, na Figura 2, que a movimentação dos produtos, em geral, teve um crescimento significativo ao longo dos anos compreendidos neste estudo e, a partir de 2020, período marcado pela pandemia de COVID-19, a movimentação de alguns produtos teve um aumento considerável durante o período da pandemia como adubo orgânico e ferro, enquanto outros tiveram redução em sua movimentação, como a gasolina.

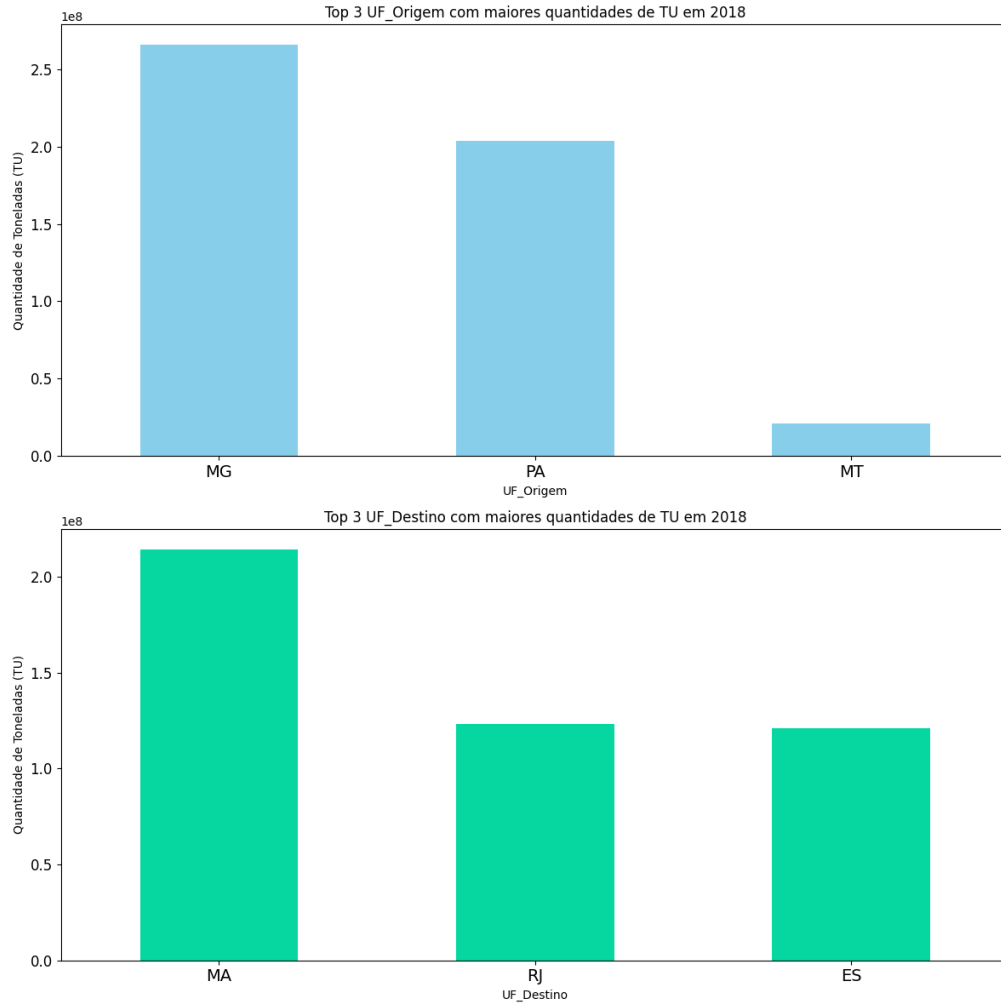
Figura 2 – Movimentação de cargas em TU para os produtos: adubo orgânico, ferro gusa, gasolina e bobina, ao longo dos anos



Fonte: Autores (2023)

Na Figura 3, apresenta-se gráficos congruentes da distribuição das toneladas entre distintos estados de origem e destino, propondo-se a análise acerca dos estados com transporte ferroviário desenvolvido e recorrência de destinos. Na origem, destacam-se os estados de Minas Gerais, Mato Grosso e Pará, enquanto no destino, destacam-se os estados de Maranhão, Espírito Santo e Rio de Janeiro.

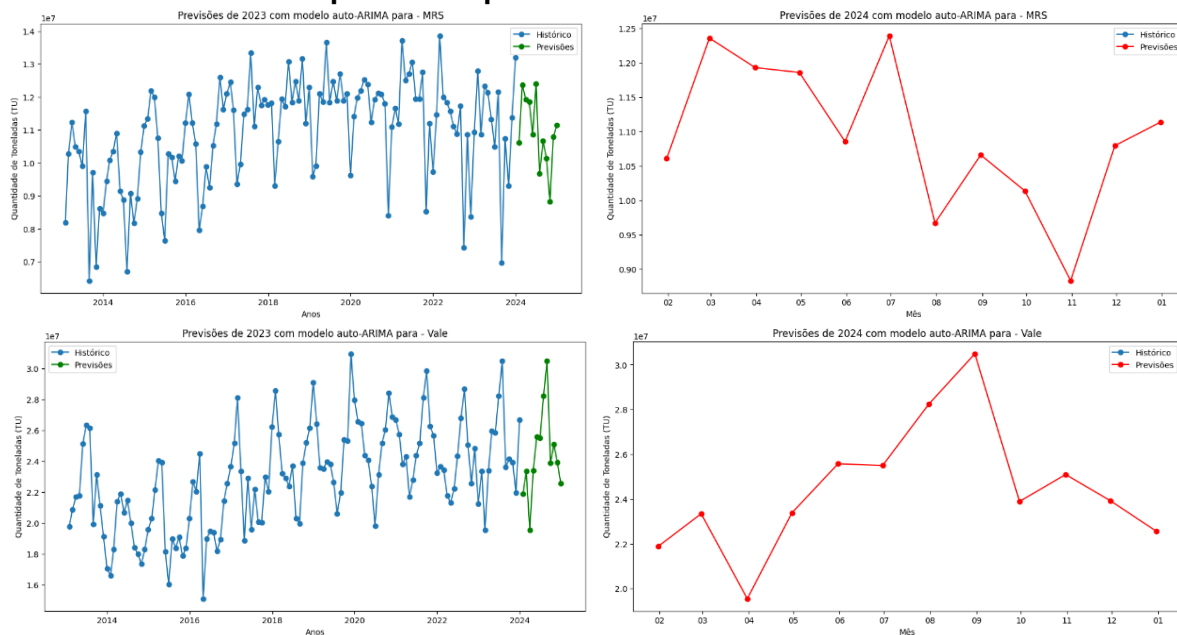
Figura 3 – Movimentação de cargas em TU entre Estados para o ano 2018



Fonte: Autores (2023)

Exemplos da previsão para a quantidade de toneladas úteis movimentada obtida para as empresas abordadas no estudo, podem ser observadas na Figura 4. Além disso, aliada ao gráfico de série temporal, verifica-se a tendência temporal e possibilidade de sazonalidade intrínseca aos meses de declínio ou aumento, justificáveis por decorrências climáticas ou acidentais, acentuada para a previsão da empresa Vale (293 milhões TU em 2024) que obteve acurácia superior a 91%.

Figura 4 – Previsão de movimentação de cargas em TU para as empresas MRS e Vale em 2024



Fonte: Autores (2023)

Em suma, as análises realizadas evidenciaram as vantagens da utilização das técnicas de mineração de dados no setor de transporte, especificamente no modal ferroviário, possibilitando reflexões e predições quanto às movimentações de cargas e mercadorias de maior relevância para o supracitado setor, considerando-se o território brasileiro.

4. CONSIDERAÇÕES FINAIS

Os dados provenientes da fonte de dados consultada apresentaram consistência, atualização sistemática e precisão, o que possibilitou uma análise exploratória descritiva eficiente e satisfatória da movimentação de carga ferroviária no Brasil, no período de 2013 a 2023, salientando-se a oportunidade de rastreabilidade para a base de dados, essencial para definição de metodologias e verificação qualitativa dos dados.

Sendo assim, a análise proporcionou embasamento quantitativo relativo à relevância da movimentação ferroviária de cargas no Brasil, bem como os produtos de maior preponderância para o setor no período analisado, sendo os mesmos o minério de ferro, a soja, o milho e o açúcar, além da percepção de que, dentre as empresas concessionárias associadas à ANTF, a Vale apresentou maior movimentação de carga, seguida pela MRS.

Por conseguinte, atesta-se a necessidade de implementação de políticas e ações que promovam o investimento na manutenção e modernização da malha ferroviária brasileira, de

modo a melhorar a eficiência, a qualidade e a competitividade do modal, visto que a ferrovia é um transporte eficiente e de baixo custo, capaz de contribuir para a redução dos custos logísticos e para a integração entre as regiões.

Em relação ao uso do autoajuste ARIMA para as previsões de movimentação de cargas, evidencia-se que o mesmo torna a análise de séries temporais mais acessível para usuários menos experientes em modelagem ARIMA. Contudo, é necessária a revisão dos resultados do autoajuste e verificação da adequação do modelo ajustado para a série temporal abordada.

Em suma, o artigo permitiu a verificação de que a aplicação de técnicas de mineração de dados é uma ferramenta promissora para o monitoramento do desempenho operacional de concessionárias ferroviárias, devido à possibilidade de obtenção de *insights* valiosos para a tomada de decisões, contribuindo para a melhoria da eficiência e da qualidade do serviço ferroviário brasileiro, a partir da análise de padrões e tendências no desempenho das concessionárias ferroviárias ao longo do tempo.

REFERÊNCIAS

ALMEIDA, J. O. **Maturidade dos processos de business intelligence na área de tecnologia da informação: um estudo de caso do Instituto Municipal de Administração Pública**. 2017. 112f. Dissertação (Mestrado em Administração) - UNIFACS, Salvador, 2017. Disponível em: <https://tede.unifacs.br/handle/tede/696>. Acesso em: 23 mar. 2024.

ALVES, E. G. C.; RAMOS, R. B. S.; SILVA, C. R. **Transporte ferroviário no Brasil: Desafios e oportunidades**. Revista Conecta, São Paulo, Brasil, v. 3, p. 15–27, 2020. Disponível em: <https://www.fatecrl.edu.br/revistaconecta/index.php/rc/article/view/18>. Acesso em: 06 set. 2023.

BARBIERI, C. **BI2 – Business Intelligence: modelagem e qualidade**. 1. ed. Rio de Janeiro: Elsevier, 2011.

BRASIL. Ações e Programas. **Ministério dos Transportes**, 2023. Disponível em: <https://www.gov.br/transportes/pt-br/acesso-a-informacao/acoes-e-programas>. Acesso em: 06 set. 2023.

BRASIL. Agência Nacional de Transportes Terrestres. **Portal de Dados Abertos: Sistema de Acompanhamento do Desempenho Operacional das Concessionárias - SIADE**. Brasília, 31 ago. 2023a. Disponível em: <https://dados.antt.gov.br/dataset/sistema-de-acompanhamento-do-desempenho-operacional-das-concessionarias-siade>. Acesso em: 31 ago. 2023.

BUSSAB, W. O. MORETTIN, P. A. **Estatística básica**. 9. ed. São Paulo: Saraiva, 2017.
CAMILO, C. O.; SILVA, J. C. **Mineração de dados: conceitos, tarefas, métodos e ferramentas**. Technical Report - RT-INF 001-09 - Instituto de Informática - Universidade Federal de Goiás (UFG), 2009. Disponível em: <https://ww2.inf.ufg.br/node/355>. Acesso em: 23 mar. 2024.

DIAS, M. A. **Logística, transporte e infraestrutura: armazenagem, operador logístico, gestão via TI, multimodal**. São Paulo: Atlas, 2012.

GABRIEL, L. C. A importância do transporte ferroviário de cargas no Brasil. **Revista do Clube Naval**, Rio de Janeiro, v. 4 n. 396, p. 56–61, 2021. Disponível em: <http://187.29.162.44/index.php/clubenaival/article/view/1872>. Acesso em: 06 set. 2023.

GAMA NETO, M. V. **O processo CRISP-DM aplicado na construção de uma solução para Análise de Risco de Crédito**. 2018. 42 f. Monografia (Graduação em Ciência da Computação) – Universidade Federal de Pernambuco, Recife, 2018.

GOOGLE. Colaboratory. **Google**, 2023. Disponível em: <https://research.google.com/colaboratory/intl/pt-BR/faq.html>. Acesso em: 05 out. 2023.

GUNAWAN, T. S. et al. Development of video-based emotion recognition using deep learning with Google Colab. **TELKOMNIKA (Telecommunication Computing Electronics and Control)**, v. 18, n. 5, p. 2463-2471, 2020. Disponível em: <http://doi.org/10.12928/telkomnika.v18i5.16717>. Acesso em: 09 out. 2023.

IBM. Visão geral da ajuda do CRISP-DM. **IBM**, 2023. Disponível em: <https://www.ibm.com/docs/pt-br/spss-modeler/18.4.0?topic=dm-crisp-help-overview>. Acesso em: 06 set. 2023.

KOTU, V.; DESHPANDE, B. Time Series Forecasting. *In*: KOTU, V.; DESHPANDE, B. (org.). **Data Science**. 2 ed. Cambridge: Morgan Kaufmann, 2019. p. 395-445.

LAGO, K.; LAENNDER, A. **Dominando o Power BI**. 1 ed. São Paulo: DATAB, 2018.

LANZA, J.F.R. **Ferrovias, mercado e políticas públicas: As shortlines como solução para o transporte ferroviário no Brasil**. São Paulo: Labrador, 2020.

MAPA FERROVIÁRIO. **ANTF**, 2023. Disponível em: <https://www.antf.org.br/mapa-ferroviario/>. Acesso em: 05 out. 2023.

MARCONI, M. A.; LAKATOS, E. M. **Fundamentos de metodologia científica**. 5. ed. São Paulo: Atlas, 2003.

MILLIET, S. **Roteiro do café e outros ensaios: contribuição para o estudo da história econômica e social do Brasil**. 3 ed. São Paulo: 1941. 211 p. (Coleção Departamento de Cultura, v. 25). Disponível em:

<https://bibliotecadigital.seade.gov.br/view/singlepage/index.php?pubcod=10011698&parte=1>. Acesso em: 14 set. 2023.

ROBERTO, C. **Crisp-DM: as 6 etapas da metodologia do future**. MBA USP/Esalq, 2023. Disponível em: <https://blog.mbauspesalq.com/2022/04/12/crisp-dm-as-6-etapas-da-metodologia-do-futuro/>. Acesso em: 17 out. 2023.

SGROTT, P. R. **Uma análise do cenário do transporte ferroviário de cargas**. 2019. Monografia (Graduação em Engenharia Ferroviária e Metroviária) – Centro Tecnológico de Joinville da Universidade Federal de Santa Catarina, Joinville, 2019. Disponível em: <https://repositorio.ufsc.br/handle/123456789/197312>. Acesso em: 06 set. 2023.

SGROTT, P. R.; ZAGHENI, E. S. S. **Uma análise do cenário do transporte ferroviário de cargas / An analysis of the rail freight scenario**. Brazilian Applied Science Review, [S. l.], v. 4, n. 6, p. 3931–3948, 2020. Disponível em: <https://ojs.brazilianjournals.com.br/ojs/index.php/BASR/article/view/21684>. Acesso em: 06 set. 2023.

SMITH, T. G. Understand p, d, and q. **Alkaline-ml**, 2023. Disponível em: https://alkaline-ml.com/pmdarima/tips_and_tricks.html#understand-p-d-and-q. Acesso em: 23 mar. 2024.

THE Pennsylvania State University. Seasonal Models. *In*: Pennsylvania State University. **Applied Time Series Analysis**. [Pennsylvania]: Penn State's Department of Statistics, 2023. Disponível em: <https://online.stat.psu.edu/stat510/lesson/4>. Acesso em: 02 nov. 2023.