

ANÁLISE DO AEROPORTO DE GUARULHOS: IMPACTO DAS PESQUISAS DE SATISFAÇÃO NA RENTABILIDADE NÃO AERONÁUTICA DOS AEROPORTOS – UM ESTUDO DE CASO

Guarulhos Airport Analysis: Assessing the Impact of Passenger Satisfaction on Airports' Non-Aeronautical Revenue – A Case Study

Gabriel Martins Gazaneo

FATEC, São José dos Campos, São Paulo

E-mail: gabriel.gazaneo@fatec.sp.gov.br

Marcus Vinicius do Nascimento

FATEC, São José dos Campos, São Paulo

E-mail: nascimento.mv@fatec.sp.gov.br

Luiz Antonio Tozi

FATEC, São José dos Campos, São Paulo

E-mail: luiz.tozi@fatec.sp.gov.br

RESUMO

O trabalho apresenta um estudo de caso com ênfase em analisar o nível de satisfação de passageiros que influenciam a margem de renda não aeronáutica do Aeroporto Internacional de Guarulhos. A pesquisa abrange o período de 2024 a 2024, dividido em dados trimestrais tanto das receitas não aeronáuticas quanto das notas e critérios avaliados. As técnicas utilizadas foram de análise estatística e inteligência artificial, como correlação de Pearson, regressão linear, análise de componentes principais (PCA) e o modelo de árvore de decisão M5P, com apoio de ferramentas como Python e Weka, para identificar os fatores de maior impacto sobre as receitas. Os resultados demonstraram correlações significativas entre determinados aspectos da experiência do passageiro e o desempenho financeiro do aeroporto. Fatores como o custo-benefício do estacionamento e a confiabilidade da inspeção de segurança apresentaram impacto positivo nas receitas, enquanto a qualidade do Wi-Fi e a limpeza dos sanitários mostraram influência negativa, possivelmente devido a ruídos ou mudanças comportamentais pós-pandemia. Conclui-se que o monitoramento recorrente dos serviços e o uso de modelos preditivos auxiliam nas decisões estratégicas.

Palavras-chave: Guarulhos, satisfação, renda, passageiros e modelo preditivo.

ABSTRACT

The study presents a case analysis focused on examining the level of passenger satisfaction that influences the non-aeronautical revenue margin of Guarulhos International Airport. The

research covers the period from 2024 to 2024, divided into quarterly data for both non-aeronautical revenues and the evaluated scores and criteria. The techniques used included statistical analysis and artificial intelligence, such as Pearson correlation, linear regression, principal component analysis (PCA), and the M5P decision tree model, supported by tools such as Python and Weka, to identify the factors with the greatest impact on revenue.

The results showed significant correlations between certain aspects of the passenger experience and the airport's financial performance. Factors such as the cost-benefit of parking and the reliability of security inspection had a positive impact on revenues, while the quality of Wi-Fi and restroom cleanliness showed a negative influence, possibly due to noise or post-pandemic behavioral changes. It is concluded that the continuous monitoring of services and the use of predictive models support strategic decision-making

Keywords: *Guarulhos, satisfaction, income, passengers, predictive model*

1. INTRODUÇÃO

As receitas não aeronáuticas são uma tendência crescente no mercado global de aeroportos, com previsão de crescimento significativo até 2032. A diversificação das receitas, reduzindo a dependência de tarifas aeronáuticas, permite que os aeroportos tenham maior controle sobre gastos e ofertas. Essa tendência é especialmente relevante para os aeroportos privados, que buscam maximizar seus lucros em um mercado em constante crescimento. Com base nisso é fundamental compreender como esses serviços podem ser otimizados e aprimorados. (Almeida e Borille, 2017)

A definição de serviços pode ser determinada como uma série de atividades intangíveis que, usualmente, ocorrem entre consumidores e empresas. Dessa forma, o aeroporto se enquadra como um serviço "industrial puro", pois seu núcleo é o serviço em si. O serviço também pode estar alinhado a uma aquisição. (Ardiansyah, Yuniawati e Ridwanudin, 2018)

A qualidade dos serviços pode ser definida como uma métrica para avaliar a gestão aeroportuária, visando proporcionar uma melhor experiência e satisfação aos passageiros. Nesse sentido, é importante definir o conceito de "Paisagem de Serviços", que abrange os elementos essenciais para o fluxo eficiente de passageiros e a conclusão eficaz do trajeto dentro do terminal.

Isso inclui desde a utilização de sanitários, serviços bancários e restaurantes até serviços pessoais, que envolvem a interação direta entre usuários e funcionários, exigindo a presença física do cliente. (Ardiansyah, Yuniawati e Ridwanudin, 2018)

Aplicando o conceito de "Paisagem de Serviços" nos aeroportos, cria-se ambientes acolhedores com restaurantes, salas de descanso e áreas VIP, melhorando a experiência dos passageiros. (Ardiansyah, Ridwanudin e Yuniawati 2018)

Com base nesses conceitos, é possível correlacionar a receita não aeronáutica como um catalisador principal para o objeto de estudo, que busca fazer uma comparação de causalidade

entre o nível de satisfação dos usuários.

Com essa perspectiva, é relevante destacar que, dentro do setor aeronáutico, especialmente em aeroportos, segundo Nascimento (2009), a avaliação da qualidade dos serviços oferecidos dentro do terminal aeroportuário é essencial para identificar as melhores práticas dentro do setor de serviços.

É fundamental considerar a “Pesquisa Nacional de Satisfação do Passageiro e Desempenho Aeroportuário”, que fornece dados essenciais sobre a avaliação dos principais aeroportos do Brasil, como base para efetuar esta pesquisa. Essa pesquisa é realizada pela Secretaria Nacional de Aviação Civil do Ministério de Portos e Aeroportos (2025).

Este projeto realiza um estudo, por meio da coleta e análise de dados, com o objetivo de criar um modelo preditivo com Inteligência Artificial que permita compreender e identificar os principais fatores de satisfação que podem ser correlacionados com o aumento total das receitas não aeronáuticas, com ênfase no Aeroporto Internacional de Guarulhos.

1.1 OBJETIVOS DO PROJETO

Os objetivos deste estudo seguem a premissa principal de realizar um estudo de caso sobre o impacto das receitas não aeronáuticas na paisagem de serviços dos aeroportos, correlacionando-os com o nível de satisfação dos usuários no Aeroporto Internacional de Guarulhos, no período de 2014 a 2024.

Os objetivos estabelecidos para esse projeto consistem em:

- i) Realizar uma análise quantitativa com base nos dados das pesquisas de satisfação dos aeroportos, obtidos através do Ministério dos Portos e Aeroportos, no período de 2014 a 2024, com ênfase no Aeroporto de Guarulhos.
- ii) Levantar informações sobre as receitas não aeronáuticas por trimestre do Aeroporto de Guarulhos no período de 2014 a 2024.
- iii) Correlacionar a eficiência dos setores melhor avaliados pelos usuários com seu impacto monetário na renda não aeronáutica dos aeroportos.
- vi) Criar um modelo preditivo baseado em inteligência artificial para identificar o peso dos fatores avaliados nas pesquisas de satisfação.
- v) Identificar os pontos e fatores de variação que podem ser foco de futuras otimizações para aprimorar a qualidade dos serviços e embasar decisões estratégicas.

2. TECNOLOGIAS UTILIZADAS

Este tópico busca apresentar as tecnologias adotadas para a realização dos objetivos necessários para o auxílio do estudo de caso, analisando o impacto das receitas não aeronáuticas sobre os serviços e a satisfação dos usuários no Aeroporto Internacional de Guarulhos entre 2014 e 2024. Essa análise envolve a integração de tecnologias de informação, logística e análise

gerencial, todos fundamentados em referências da literatura e alinhados com a orientação acadêmica.

2.1 TECNOLOGIAS DE LOGÍSTICA OU DE ANÁLISE GERENCIAL

A área de logística é ampla, e cada uma de suas vertentes conta com um conjunto específico de tecnologias. Essas tecnologias e conceitos, aplicados ao estudo de casos, permitem identificar e qualificar, de maneira objetiva, quais pontos e direções devem ser priorizados e aplicados para uma melhor definição dos resultados obtidos.

Segundo Cano, Salazar, Gómez-Montoya e Cortés (2021), Big Data Analytics trata-se de uma ferramenta analítica que aplica estatísticas avançadas a qualquer tipo de dado, identificando padrões de comportamento. Isso envolve o conceito de Big Data Analytics, que abrange uma grande variedade de informações que podem ser utilizadas. O uso de Big Data pode ser empregado como uma estratégia de visão e tendência para a cadeia de suprimentos, auxiliando na otimização e no carregamento do setor.

Segundo Pearson (1907), correlação de Pearson é uma medida estatística de associação linear entre duas variáveis quantificadas, com intuito de expressar até que ponto os dados de uma variável podem ser previstos a partir dos valores da outra, com base numa relação linear.

Segundo Smith (2002), a Análise de Componentes Principais (PCA) é uma forma de identificar padrões em dados, com o objetivo de realçar semelhanças ou diferenças. Serve principalmente para analisar dados, sendo utilizada principalmente para avaliar e agrupar informações.

Segundo Oliveira Filho (2002), a regressão linear tem como objetivo descobrir, por meio de um modelo matemático, a relação que pode existir entre duas ou mais variáveis. A partir de um conjunto de observações, as variáveis independentes têm o propósito de aprimorar a capacidade de previsão em comparação com a regressão linear simples.

Segundo Ardiansyah, Ridwanudin e Yuniawati (2018), o conceito de servicescape refere-se à aplicação de elementos espaciais como parte da funcionalidade do ambiente, garantindo que o fluxo de passageiros ocorra de maneira eficiente e coordenada com outros serviços que não estão diretamente relacionados às finalidades do percurso.

2.2 TECNOLOGIAS DA INFORMAÇÃO

As tecnologias da informação desempenharam um papel fundamental na realização deste estudo de caso, possibilitando a modelagem e o tratamento dos dados. O uso dessas tecnologias permitiu o cálculo eficiente das variáveis e a criação de modelos preditivos, essenciais para a

realização das análises finais.

O Python, com bibliotecas como Pandas, NumPy, Scikit-learn, TensorFlow e PyTorch, pode ser amplamente utilizado na análise de dados de pesquisas de satisfação em aeroportos e no processamento de receitas não aeronáuticas, possibilitando o desenvolvimento de modelos preditivos baseados em inteligência artificial.

LibreOffice é uma ferramenta útil para a gestão inicial dos dados, facilitando a organização de receitas aeroportuárias e resultados de pesquisas em planilhas, servindo como um ponto de partida para análises mais detalhadas em outras plataformas.

Weka (Waikato Environment for Knowledge Analysis) é um software de código aberto desenvolvido pela Universidade de Waikato, na Nova Zelândia. Ele auxilia na análise de dados e em aplicações de machine learning, permitindo o processamento, cálculo e correlação entre variáveis de forma prática e eficiente.

3 METODOLOGIA

Este tópico descreve, em detalhes, o processo de coleta, tratamento, integração e formatação dos dados para análise estatística e preditiva, com base na pesquisa de satisfação dos passageiros no período de 2014 a 2024 do Aeroporto Internacional de Guarulhos (GRU Airport). O principal objetivo é detalhar as possíveis correlações entre a percepção dos usuários e a performance financeira, expressa em receitas não aeronáuticas do aeroporto ao longo dos anos.

Os dados de satisfação dos passageiros foram obtidos a partir das pesquisas publicadas pelo Governo Federal do Brasil. Foram utilizadas duas séries temporais:

2013 a 2019: base com 38 critérios de avaliação.

2020 a 2024: base com 61 critérios de avaliação.

Ambas as séries foram baixadas em formato CSV diretamente do portal oficial.

Paralelamente, foram coletados os relatórios anuais de administração do GRU Airport (voltados para investidores), contendo os dados de receitas não aeronáuticas no período de 2014 a 2024. Importante ressaltar que a escolha pelo GRU Airport se deu pela consistência das publicações e pela disponibilidade de dados em ambos os conjuntos utilizados.

Tabela 1 - Evolução da Receita Não Tarifária Anual do GRU Airport (2014–2024)

Trimestre	Aeroporto	Receita (R\$ milhões)
2014-03-31	Aeroporto Internacional de Guarulhos (GRU)	176.1
2014-06-30	Aeroporto Internacional de Guarulhos (GRU)	241.6
2014-09-30	Aeroporto Internacional de Guarulhos (GRU)	251.9
2014-12-31	Aeroporto Internacional de Guarulhos (GRU)	241.1
2015-03-31	Aeroporto Internacional de Guarulhos (GRU)	240.5
2015-06-30	Aeroporto Internacional de Guarulhos (GRU)	239.6

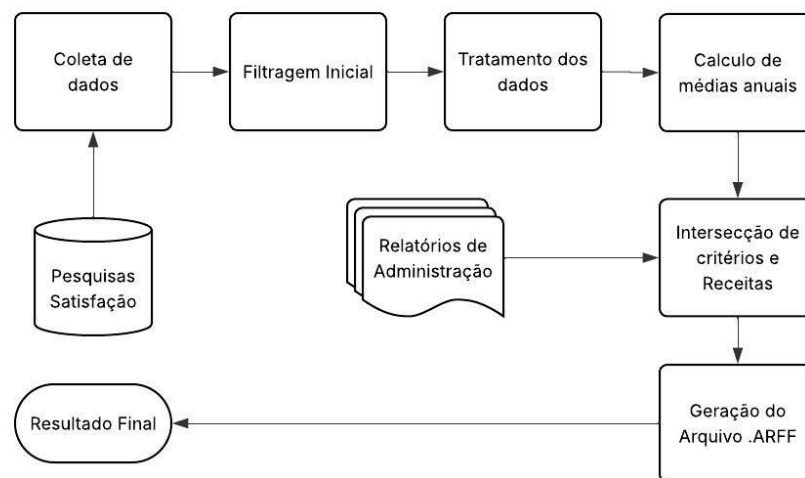
2015-09-30	Aeroporto Internacional de Guarulhos (GRU)	229.4
2015-12-31	Aeroporto Internacional de Guarulhos (GRU)	228.6
2016-03-31	Aeroporto Internacional de Guarulhos (GRU)	221.6
2016-06-30	Aeroporto Internacional de Guarulhos (GRU)	227.5
2016-09-30	Aeroporto Internacional de Guarulhos (GRU)	234.7
2016-12-31	Aeroporto Internacional de Guarulhos (GRU)	227.6
2017-03-31	Aeroporto Internacional de Guarulhos (GRU)	213.4
2017-06-30	Aeroporto Internacional de Guarulhos (GRU)	223.5
2017-09-30	Aeroporto Internacional de Guarulhos (GRU)	241.4
2017-12-31	Aeroporto Internacional de Guarulhos (GRU)	242.4
2018-03-31	Aeroporto Internacional de Guarulhos (GRU)	232.2
2018-06-30	Aeroporto Internacional de Guarulhos (GRU)	243.9
2018-09-30	Aeroporto Internacional de Guarulhos (GRU)	257.1
2018-12-31	Aeroporto Internacional de Guarulhos (GRU)	257.0
2019-03-31	Aeroporto Internacional de Guarulhos (GRU)	258.9
2019-06-30	Aeroporto Internacional de Guarulhos (GRU)	258.4
2019-09-30	Aeroporto Internacional de Guarulhos (GRU)	271.2
2019-12-31	Aeroporto Internacional de Guarulhos (GRU)	276.0
2020-03-31	Aeroporto Internacional de Guarulhos (GRU)	259.2
2021-03-31	Aeroporto Internacional de Guarulhos (GRU)	194.3
2021-06-30	Aeroporto Internacional de Guarulhos (GRU)	158.9
2021-09-30	Aeroporto Internacional de Guarulhos (GRU)	180.7
2022-03-31	Aeroporto Internacional de Guarulhos (GRU)	256.0
2022-06-30	Aeroporto Internacional de Guarulhos (GRU)	277.6
2022-09-30	Aeroporto Internacional de Guarulhos (GRU)	296.6
2022-12-31	Aeroporto Internacional de Guarulhos (GRU)	315.4
2023-12-31	Aeroporto Internacional de Guarulhos (GRU)	329.4
2023-12-31	Aeroporto Internacional de Guarulhos (GRU)	358.4
2023-12-31	Aeroporto Internacional de Guarulhos (GRU)	383.8
2023-12-31	Aeroporto Internacional de Guarulhos (GRU)	397.8
2024-03-31	Aeroporto Internacional de Guarulhos (GRU)	386.1
2024-06-30	Aeroporto Internacional de Guarulhos (GRU)	411.9
2024-09-30	Aeroporto Internacional de Guarulhos (GRU)	448.1
2024-12-31	Aeroporto Internacional de Guarulhos (GRU)	474.4

Fonte: Autor. 2025

É importante mencionar que, nesta série histórica, há ausência de informações referentes ao 2º, 3º e 4º trimestres dos anos de 2020 e 2021, especialmente a falta de dados do 4º trimestre. Segundo informações fornecidas pelo Ministério de Portos e Aeroportos, essa lacuna se deve ao impacto da pandemia de COVID-19, que afetou diretamente a coleta e o registro dessas informações.

O fluxograma da Figura 1 ilustra os processos realizados durante as etapas de coleta, tratamento, integração e preparação dos dados utilizados no desenvolvimento deste trabalho.

Figura 1. Fluxograma de tratamento de dados



Fonte: Autor. 2025

O primeiro passo consistiu na coleta dos dados, realizada por meio da importação de arquivos contendo os resultados das pesquisas de satisfação dos aeroportos, provenientes de fontes oficiais do governo. Esses arquivos continham informações detalhadas sobre diversos critérios de avaliação no período de 2014 a 2024.

Em seguida, foi realizada a etapa de filtragem e tratamento dos dados, utilizando a ferramenta Google Colab, com o auxílio de scripts desenvolvidos em Python. Inicialmente, os dados foram filtrados para manter apenas os registros correspondentes ao Aeroporto de Guarulhos, que é o foco da análise deste trabalho. Posteriormente, foram tratados os valores inconsistentes, como respostas ausentes ou registros com notas fora do intervalo válido (de 1 a 5). Após esse processo, foram calculadas as médias anuais para cada critério de satisfação, proporcionando uma visão consolidada do desempenho do aeroporto ao longo dos anos.

Adicionalmente, realizou-se a identificação e seleção dos critérios comuns entre as duas séries temporais disponíveis, garantindo que a comparação entre os períodos fosse realizada de forma consistente. As tabelas foram devidamente padronizadas, com ajustes nos formatos e

estruturas dos dados, a fim de viabilizar análises conjuntas.

Na etapa seguinte, foi feita a integração dos dados de satisfação com os dados de receita não aeronáutica do aeroporto, os quais foram importados dos relatórios de administração do Aeroporto de Guarulhos. Esse processo permitiu a construção de uma base de dados unificada, que consolidou, para cada ano, as médias dos critérios de satisfação (considerando apenas os critérios comuns entre os períodos analisados) e o respectivo valor total da receita não aeronáutica.

Por fim, os dados tratados, consolidados e padronizados foram organizados em um arquivo no formato .arff, compatível com o software Weka (Waikato Environment for Knowledge Analysis). Esse formato é específico para utilização no Weka, permitindo a aplicação de técnicas de mineração de dados e aprendizado de máquina (Figura 2).

Figura 2. Esquematização modelo .arff

```
@relation SBGR_trimestres_2014_2019
@attribute TRIMESTRE date "yyyy-MM-dd"
@attribute "FACILIDADE DE EMBARQUE / DESEMBARQUE NO MEIO-FIO" numeric
@attribute "DISPONIBILIDADE DE TRANSPORTE PÚBLICO PARA O AEROPORTO" numeric
@attribute "CORDIALIDADE E PRESTATIVIDADE DOS FUNCIONÁRIOS DO CHECK-IN" numeric
@attribute "CORDIALIDADE E PRESTATIVIDADE DOS FUNCIONÁRIOS DA INSPEÇÃO DE SEGURANÇA"
numeric
@attribute "TEMPO DE FILA NO CHECK-IN (AUTOATENDIMENTO)" numeric
@attribute "TEMPO DE FILA NO CHECK-IN (GUICHÊ)" numeric
@attribute "CONFIABILIDADE DA INSPEÇÃO DE SEGURANÇA" numeric
@attribute "TEMPO DE FILA NA EMIGRAÇÃO" numeric
@attribute "QUANTIDADE E QUALIDADE DE LANCHONETES/RESTAURANTES" numeric
@attribute "QUANTIDADE E QUALIDADE DE ESTABELECIMENTOS COMERCIAIS" numeric
@attribute "QUALIDADE DAS INSTALAÇÕES DE ESTACIONAMENTO DE VEÍCULOS" numeric
@attribute "CUSTO-BENEFÍCIO DO ESTACIONAMENTO" numeric
@attribute "QUALIDADE DA SINALIZAÇÃO DO AEROPORTO" numeric
@attribute "INTEGRIDADE DA BAGAGEM" numeric
@attribute "VELOCIDADE DA RESTITUIÇÃO DE BAGAGEM" numeric
@attribute "DISPONIBILIDADE DE ASSENTOS NA SALA DE EMBARQUE" numeric
@attribute "CONFORTO TÉRMICO DO AEROPORTO" numeric
@attribute "CONFORTO ACÚSTICO DO AEROPORTO" numeric
@attribute "DISPONIBILIDADE DE SANITÁRIOS" numeric
@attribute "LIMPEZA DOS SANITÁRIOS" numeric
@attribute "QUALIDADE DA INTERNET / WI-FI" numeric
@attribute "RECEITA NAO AERONAUTICA" numeric

@data
2014-03-31,2.35,2.46,3.62,3.83,3.49,3.09,3.77,2.86,3.43,3.28,2.19,1.66,4.04,3.28,3.1,?,3.85,3.83,4.12,4.02,2.65,176.1
2014-06-30,4.32,4.31,4.43,4.29,4.56,4.17,4.24,4.86,3.90,3.89,4.09,2.5,3.84,4.06,3.82,3.81,4.22,3.83,4.14,4.31,3.18,241.6
2014-09-30,4.35,4.05,4.44,4.44,4.65,4.17,4.39,4.74,3.95,4.04,4.12,3.05,4.08,4.29,4.13,4.43,4.33,4.21,4.36,4.36,290,251.9
2014-12-31,3.95,3.64,4.05,4.07,4.42,3.88,4.03,4.26,3.73,3.86,4.13,3.13,3.85,4.25,4.1,4.07,3.95,3.84,3.99,3.92,3.14,241.1
2015-03-31,4.3,3.87,4.70,4.47,4.53,3.99,4.44,4.61,4.08,4.31,4.36,3.1,4.25,4.35,4.08,4.35,4.45,4.33,4.33,4.21,3.07,240.5
2015-06-30,4.36,4.13,4.57,4.51,4.67,4.23,4.46,4.62,4.16,4.30,4.39,3.08,4.22,4.26,4.07,4.19,4.31,4.25,4.28,4.05,3.26,239.6
2015-09-
30,2.97,2.68,3.86,3.82,3.90,3.80,3.77,3.51,2.71,2.86,2.88,2.32,3.46,3.32,2.77,3.05,3.42,3.30,3.13,3.11,2.54,229.4,.....
```

Fonte: Autor. 2025

O arquivo utilizado está estruturado no formato ARFF (Attribute-Relation File Format),

que é amplamente utilizado em mineração de dados, especialmente em ferramentas como o Weka. Esse formato é composto por duas partes principais: cabeçalho e dados. O cabeçalho contém informações sobre o conjunto de dados e a definição de suas variáveis. Seus principais componentes são:

RELATION: Indica o nome do conjunto de dados.

ATTRIBUTE: Define cada variável presente no conjunto de dados, especificando seu nome e seu tipo de dado.

Os tipos de dados mais comuns são:

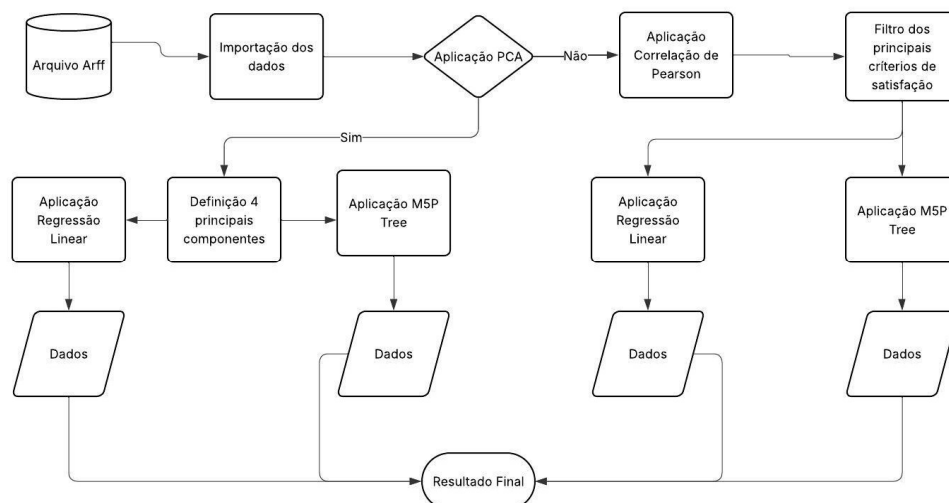
NUMERIC – Indica que a variável assume valores numéricos (inteiros ou decimais).

DATE "yyyy" – Define que a variável é uma data no formato de ano e trimestre (ex.: 2014-09-30).

{categoria1, categoria2, ...} – Representa variáveis categóricas, com valores pré-definidos.

Neste tópico são apresentados os resultados obtidos a partir do tratamento de dados realizado com o software Weka. São discutidos os diferentes modelos testados, seus desempenhos e uma análise sobre a correlação entre os critérios de nível de satisfação dos passageiros no Aeroporto de Guarulhos e suas respectivas receitas não aeronáuticas (Figura 3).

Figura 3. Fluxograma tratamento de dados através do Weka



Fonte: Autor .2025

Este fluxograma determina o processo feito para tratamento de dados através do Weka. Neste tratamento, foi seguida a seguinte ordem:

- I. Tratamento do arquivo ARFF – Pré-processamento dos dados no formato adequado.
- II. Escolha entre utilizar ou não o método PCA – Decisão sobre a aplicação da Análise de Componentes Principais.

Com PCA, foram calculados os principais componentes. Em seguida, foram aplicados tanto regressão linear quanto árvore de decisão M5P para análise. Sem PCA, foi feita a correlação de Pearson para identificar relações entre variáveis. Determinou-se os principais critérios com valores positivos (mais relevantes). Após essa etapa, o PCA foi aplicado para redução dimensional. Observação importante: Os resultados do PCA foram obtidos com centralização, pois o estudo compara variáveis em escalas diferentes:

- I. Uma variável em escala de 1 a 5 (ex: avaliação qualitativa).
- II. Outra variável em valores monetários (R\$ milhões) (ex.: receita).

4 RESULTADOS E DISCUSSÕES

Tabela 2. Comparação de Modelos Preditivos

Cenário	Modelo	Correlação (R)	Erro abs. Médio	Erro quadrático médio	Erro Rel. Absoluto	Erro Rel. Quadrático
PCA (com cent.)	Regressão Linear	0.6864	36.50	52.70	63.91%	70.22%
PCA (com cent.)	M5P	0.8867	25.28	36.34	44.25%	48.42%
Seleção (Corr)	Regressão Linear	0.6207	44.54	58.69	77.98%	78.20%
Seleção (Corr)	M5P	0.8518	30.40	39.08	53.22%	52.07%

Fonte: Autor .2025

4.1 CENÁRIO 1 REGRESSÃO LINEAR COM PCA:

$$\text{RECEITA_NAO_AERONAUTICA} = -0 \times (\text{componentes principais}) + 271.605$$

Logo, o modelo final não utilizou nenhuma variável ativa (os coeficientes foram arredondados para zero) e previu apenas a média, 271.605. Isso demonstra que, neste cenário, os critérios de avaliação acabaram selecionando um modelo que gera previsões pela média, com desempenho fraco (correlação de 0,6864). Portanto, não representa um modelo mais eficiente.

4.2 CENÁRIO 2 ÁRVORE DE DECISÃO (M5P) COM PCA:

$$\text{LM número 3: RECEITA_NAO_AERONAUTICA} = 252.9651$$

Para a partição correspondente, o modelo sempre prevê o valor 252.9651,

independentemente dos valores de entrada. A árvore gerou 4 modelos lineares, mas nas regiões onde os critérios não apresentaram variação suficiente na Receita Não Aeronáutica, foram atribuídas constantes em vez de modelos lineares (os nós finais são representados por LM, ou seja, modelos lineares ou constantes).

Dessa forma, a árvore particionou os dados principalmente com base no período de tempo, resultando em previsões constantes para cada segmento: 174.10, 230.01, 252.97 e 246.18.

Por isso, o modelo apresentou desempenho muito bom, com correlação $R \approx 0,887$ (Figura 4).

Figura 4 . Resultado correlação de Pearson

```
Ranked attributes:
0.7271  1 TRIMESTRE
0.5446  22 QUALIDADE_DA_INTERNET_WIFI
0.4933  13 CUSTO_BENEFICIO_DO_ESTACIONAMENTO
0.2873  2 FACILIDADE_DE_EMBARQUE_DESEMBARQUE_NO_MEIO_FIO
0.1809  3 DISPONIBILIDADE_DE_TRANSPORTE_PUBLICO_PARA_O_AEROPORTO
0.0803  8 CONFIABILIDADE_DA_INSPECAO_DE_SEGURANCA
0.0506  21 LIMPEZA_DOS_SANITARIOS
-0.1906  11 QUANTIDADE_E_QUALIDADE_DE_ESTABELECIMENTOS_COMERCIAIS
-0.1927  10 QUANTIDADE_E_QUALIDADE_DE_LANCHONETES_RESTAURANTES
-0.2246  17 DISPONIBILIDADE_DE_ASSENTOS_NA_SALA_DE_EMBARQUE
-0.2518  20 DISPONIBILIDADE_DE_SANITARIOS
-0.282  7 TEMPO_DE_FILA_NO_CHECK_IN_GUICHE
-0.3272  16 VELOCIDADE_DA_RESTITUICAO_DE_BAGAGEM
-0.3928  12 QUALIDADE_DAS_INSTALACOES_DE_ESTACIONAMENTO_DE_VEICULOS
-0.395  19 CONFORTO_ACUSTICO_DO_AEROPORTO
-0.4188  18 CONFORTO_TERMICO_DO_AEROPORTO
-0.4334  5 CORDIALIDADE_E_PRESTATIVIDADE_DOS_FUNCIONARIOS_DA_INSPECAO_DE_SEGURANCA
-0.4543  6 TEMPO_DE_FILA_NO_CHECK_IN_AUTOATENDIMENTO
-0.4665  4 CORDIALIDADE_E_PRESTATIVIDADE_DOS_FUNCIONARIOS_DO_CHECK_IN
-0.4731  9 TEMPO_DE_FILA_NA_EMIGRACAO
-0.4901  15 INTEGRIDADE_DA_BAGAGEM
-0.5939  14 QUALIDADE_DA_SINALIZACAO_DO_AEROPORTO
```

Selected attributes: 1,22,13,2,3,8,21,11,10,17,20,7,16,12,19,18,5,6,4,9,15,14 : 22

Fonte: Autor .2025

A imagem acima representa o resultado da correlação de Pearson obtido no software Weka. Neste caso, o software identificou que apenas os atributos 1, 22, 13, 2, 3, 8 e 21 apresentaram correlções positivas com a variável de interesse.

Dessa forma, esses atributos foram filtrados para serem avaliados nos modelos com foco em correlação, priorizando o cálculo dos principais fatores e seus critérios de impacto na Receita Não Aeronáutica.

4.3 CENÁRIO 3 – REGRESSÃO LINEAR COM CORRELAÇÃO PEARSON:

$$\text{RECEITA_NAO_AERONAUTICA} = 0 * \text{TRIMESTRE} + -44.2923 * \text{QUALIDADE_DA_INTERNET_WIFI} + -699.6559$$

Houve apenas uma variável considerada relevante, que foi a Qualidade da Internet Wi-Fi. No entanto, o resultado apresentou um coeficiente negativo, o que pode indicar inconsistência contextual. Além disso, o modelo obteve uma correlação fraca com a variável dependente, com $R \approx 0,4849$, o que reforça o baixo poder preditivo da regressão linear nesse cenário.

4.4 CENÁRIO 4 – ÁRVORE DE DECISÃO (M5P) COM CORRELAÇÃO PEARSON:

Este cenário conseguiu identificar padrões relevantes por meio de cinco regras distintas, baseadas em segmentações lógicas e interpretações compreensíveis. As regras evidenciaram diferenças significativas entre os trimestres anteriores e posteriores a 2020, possivelmente refletindo os impactos da pandemia de COVID-19 sobre a dinâmica aeroportuária (Figura 5).

Figura 5. Resultado das regras

1 - Rule 1 – Períodos antigos ($\text{TRIMESTRE} \leq 1534302000000$) =
 $\text{RECEITA_NAO_AERONAUTICA} = +11.69 * \text{TRANSPORTE_PÚBLICO} + 16.687 * \text{ESTACIONAMENTO} - 12.39 * \text{LIMPEZA} - 27.17 * \text{WIFI} - 101.50$

2 - Rule 2 – Períodos mais recentes ($\text{TRIMESTRE} > 1688223600000$) = **$-114.5 * \text{LIMPEZA} - 84.5 * \text{WIFI} - 3790.6$**

3 - Rule 3 – Wi-Fi ruim, segurança confiável :
 $\text{QUALIDADE_DA_INTERNET_WIFI} \leq 4.27$
 $\text{CONFIABILIDADE} > 4.33$
 $\text{RECEITA} = -47.1 * \text{WIFI} - 628.62$

4 - Baixa confiabilidade na inspeção:
 $\text{CONFIABILIDADE} \leq 4.35$

Fonte: Autor .2025

A figura acima apresenta cada uma das regras geradas pelo modelo MP5 na árvore de decisão. A seguir, descreve-se a interpretação de cada uma delas:

Regra 1 – Períodos antigos ($\text{TRIMESTRE} \leq 1534302000000$): A receita apresenta

aumento quando há boa avaliação do transporte público e do custo-benefício do estacionamento. Curiosamente, melhores avaliações de Wi-Fi e sanitários aparecem associadas a coeficientes negativos, o que pode indicar um efeito de época ou ruído nos dados. Esta regra abrange 18 instâncias (45%) e apresenta erro médio de 11,69%.

Regra 2 – Períodos mais recentes ($TRIMESTRE > 1688223600000$): Esta regra parece captar uma queda drástica na receita nos trimestres mais recentes, possivelmente associada à pandemia, crises econômicas ou obras no aeroporto. Abrange 8 instâncias (aproximadamente 20%).

Regra 3 – Wi-Fi ruim e segurança confiável: Indica que, mesmo quando a inspeção de segurança é considerada confiável, a receita é impactada negativamente se a qualidade do Wi-Fi for ruim.

Regra 4 – Baixa confiabilidade na inspeção: Mostra que a queda na confiabilidade da inspeção de segurança tem impacto fortemente negativo sobre a receita.

Regra 5 – Regra default: Quando nenhuma das regras anteriores é satisfeita, o modelo aplica uma previsão padrão (valor constante). Essa regra representa 3 instâncias (7,5%). cada uma delas. O modelo apontou efeitos significativos de variáveis como:

- a) Qualidade da Internet Wi-Fi
- b) Confiabilidade da inspeção de segurança
- c) Custo-benefício do estacionamento
- d) Limpeza dos sanitários

É importante observar que muitos coeficientes foram negativos, o que pode indicar mudanças no comportamento e nas expectativas dos passageiros, possivelmente relacionadas a restrições externas ao ambiente aeroportuário, como o fechamento temporário de atividades comerciais e operacionais.

Esse resultado sugere que métodos de análise em séries históricas podem ser viabilizados com a redução do número de perguntas nos questionários de avaliação, consolidando os critérios em menos sessões. De forma geral, o M5P demonstra que a receita não aeronáutica apresenta associação positiva com o custo-benefício do estacionamento e a confiabilidade da inspeção de segurança, enquanto fatores como qualidade do Wi-Fi e limpeza podem impactar negativamente. Evidenciando a necessidade de monitorar continuamente a percepção de qualidade dos serviços auxiliares com sua influência na performance financeira do aeroporto.

5 CONSIDERAÇÕES FINAIS

A partir da integração dos dados de níveis de satisfação dos passageiros, disponibilizados pelo Ministério dos Portos e Aeroportos, com os relatórios de receitas não aeronáuticas do GRU Airport, foi possível concluir, em uma base sólida, que a visualização, através de métodos estatísticos com análise preditiva, auxilia na identificação de padrões na percepção do usuário e no desempenho financeiro.

Os resultados implicam na evidência de que a relação entre satisfação e receita não aeronáutica é complexa e nem sempre pode ser linear. Entre os modelos avaliados, a árvore de decisão M5P com aplicação de PCA apresentou o melhor desempenho, com correlação próxima de 0,89 e erros relativos inferiores aos demais cenários testados, indicando maior capacidade preditiva em comparação à regressão linear. Além disso, a análise de correlação e as regras geradas pelo modelo destacaram a relevância das variáveis principais, como o custo-benefício do estacionamento, a confiabilidade da inspeção de segurança, a qualidade da internet Wi-Fi e a limpeza dos sanitários, que surgiram como fatores associados, de maneira positiva ou negativa, às variações na receita não aeronáutica ao longo da série histórica.

Essas definições reforçam o ponto principal de que a gestão orientada por dados pode apoiar decisões estratégicas voltadas à melhoria da paisagem de serviços e ao aumento da rentabilidade em atividades não aeronáuticas. Evidenciando essa tese, é necessário afirmar que determinados critérios de avaliação podem oferecer, através do estudo, base para priorização de investimentos em áreas com maior potencial de retorno, como estacionamento e processos de segurança, sem negligenciar aspectos que, embora apresentem coeficientes negativos em alguns cenários, podem refletir mudanças de comportamento e de expectativa dos passageiros, sobretudo em períodos marcados por eventos disruptivos, como a pandemia de COVID-19.

Por fim, é importante mencionar que tais pontos foram possíveis de serem notados mesmo com a concentração da análise em um único aeroporto e com a existência de lacunas na série temporal, especialmente entre 2020 e 2021, decorrentes do impacto da pandemia na coleta de informações, abrindo espaço para pesquisas futuras que aspirem complementar e fortalecer essa argumentação.

6 REFERÊNCIAS

ALMEIDA, Rafael de Araujo; BORILLE, Giovanna Miceli Ronzani. Investigação sobre a relação entre as características do terminal de passageiros aeroportuário e as receitas não aeronáuticas. 2017. Disponível em: <https://www.scielo.br>. Acesso em: 04 jul. 2025.

ARDIANSYAH, Fadjri Noor; YUNIAWATI, Yeni; RIDWANUDIN, Oce. The Influence of Airport Service Quality toward Passenger Satisfaction. In: INTERNATIONAL SEMINAR ON TOURISM (ISOT), 3., 2018. Anais... 2019.

BRASIL. Ministério dos Portos e Aeroportos. Pesquisa de Satisfação do Passageiro em Aeroportos. Disponível em: <https://dados.gov.br/dados/conjuntos-dados/pesquisa-de-satisfacao-do-passageiro-em-aeroportos>. Acesso em: 04 jul. 2025.

CANO, J. A.; SALAZAR, F.; GÓMEZ-MONTOYA, R. A.; CORTÉS, P. Tecnologias disruptivas e convencionais para o suporte aos processos logísticos: uma revisão da literatura. International Journal of Technology, v. 12, n. 3, p. 448–460, 2021.

OLIVEIRA FILHO, M. L. A Utilização da Regressão Linear Como Ferramenta Estratégica Para a Projeção dos Custos de Produção. In: CONGRESSO BRASILEIRO DE CUSTOS, 9., 2002, Natal. Anais... Natal: Associação Brasileira de Custos, 2002.

PEARSON, Karl. Mathematical contributions to the theory of evolution. XIII. On the theory of contingency and its relation to association and normal correlation. In: Drapers' Company Research Memoirs, Biometric Series, I. London: Dulau and Co., 1907

SMITH, L. I. A tutorial on principal components analysis. 26 fev. 2002. Disponível em: http://www.cs.otago.ac.nz/cosc453/student_tutorials/principal_components.pdf. Acesso em: 04 jul. 2025.