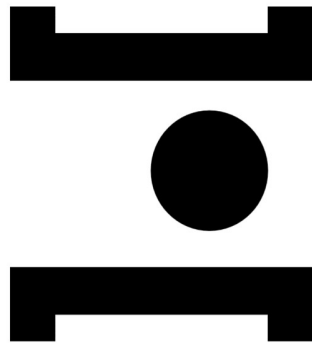


INSTITUTO POLITÉCNICO DE SANTARÉM

Escola Superior de Gestão e Tecnologia



**POLITÉCNICO
DE SANTARÉM**

**Padrões principais do Abandono Escolar no Ensino Superior: Estudo
de caso no *IPSantarém***

Relatório de Estágio

Mestrado em Informática Aplicada

Guilherme Ribeiro de Oliveira

Orientação:

Filipe Montez Coelho Madeira (PhD)

dezembro, 2025

Breve Nota Biográfica

O candidato licenciou-se em Informática na Escola Superior de Gestão e Tecnologia de Santarém em 2023. Nesse ano ingressou no Mestrado em Informática Aplicada e em 2025 trabalha como técnico de Informática no *CIIPS* (Centro de Informática do Instituto Politécnico de Santarém).

Agradecimentos

Aos meus pais, pelo apoio que me deram na minha vida e percurso acadêmico. Agradeço-lhes a sua paciência e motivação mostrada ao longo deste trabalho.

Ao meu orientador, Prof. Filipe Madeira, por todo o apoio. Agradeço-lhe pela sua boa disposição, conselhos e total disponibilidade que revelou em todas as fases do trabalho e, acima de tudo, pela confiança que depositou em mim.

Ao Prof. José Maurício e aos meus colegas no *CIIPS* pelo apoio, paciência e contributo que me deram ao longo do estágio.

Acrónimos/Lista de Siglas

API – Application Programming Interface.

BI – Business Intelligence.

CIIPS – Centro de Informática do Instituto Politécnico de Santarém.

CPLP – Comunidade dos Países de Língua Portuguesa.

DAX – Data Analysis Expressions.

ESA – Escola Superior Agrária.

ESDRM – Escola Superior de Desporto de Rio Maior.

ESE – Escola Superior de Educação.

ESGT – Escola Superior de Gestão e Tecnologia.

ESS – Escola Superior de Saúde.

ETL – Extract, Transform and Load.

IA – Inteligência Artificial.

IPSantarém – Instituto Politécnico de Santarém.

IPVC – Instituto Politécnico de Viana do Castelo.

SIGARRA – Sistema de Informação para Gestão Agregada dos Recursos e dos Registos Académicos.

SQL – Structured Query Language.

Resumo

O abandono escolar no ensino superior é um problema complexo que afeta os estudantes, as instituições de ensino e a sociedade (Sosu & Pheunpha, 2019). Muitos estudantes enfrentam dificuldades que os levam a interromper os estudos, seja por razões económicas, falta de adaptação ao ambiente académico ou outros fatores pessoais e institucionais.

Este estudo, de natureza quantitativa e baseado num estudo de caso, centra-se na análise dos padrões de abandono no Instituto Politécnico de Santarém (*IPSantarém*), recorrendo aos dados administrativos provenientes do Sistema de Informação para Gestão Agregada dos Recursos e dos Registos Académicos (*SIGARRA*). Através de processos de *ETL* (*Extract, Transform and Load*) e da utilização de ferramentas de *Business Intelligence* (*BI*), nomeadamente o *Power BI*, foram desenvolvidas visualizações e análises que permitem identificar tendências, perfis de risco e variáveis associadas ao abandono. Foram ainda testadas abordagens analíticas complementares, incorporando técnicas de Inteligência Artificial (*IA*) para apoiar a previsão de tendências.

Os resultados obtidos através da análise dos dados do *SIGARRA* e das visualizações construídas no *Power BI* permitem identificar padrões claros de abandono no *IPSantarém*. Observou-se que o abandono tende a concentrar-se em determinados ciclos de estudo, cursos e faixas etárias sendo também influenciado por outros fatores tais como o facto de o estudante estar ou não deslocado. A incorporação de técnicas de Inteligência Artificial permitiu aprofundar a compreensão das relações entre variáveis, contribuindo para a identificação de perfis de risco e reforçando a utilidade de abordagens preditivas.

Com base nos resultados obtidos, recomenda-se o desenvolvimento de estratégias institucionais direcionadas aos grupos de maior risco, nomeadamente os estudantes deslocados, os que apresentam baixo desempenho académico e os que “desaparecem” do instituto sem anular a matrícula. A criação de programas de acolhimento mais estruturados, acompanhamento pedagógico contínuo e sistemas de alerta precoce suportados por modelos preditivos de IA poderão contribuir para a redução das taxas de abandono. Sugere-se ainda que o *IPSantarém* invista na melhoria contínua da recolha e consolidação dos dados administrativos, de forma a permitir análises mais robustas e o desenvolvimento de mecanismos automáticos de monitorização.

Palavras-chave: Abandono escolar; Perfis de risco; Ensino superior, *Business Intelligence*; *ETL*; Inteligência Artificial; Análise preditiva.

Abstract

Dropout in higher education is a complex problem that affects students, educational institutions and society (Sosu & Pheunpha, 2019). Many students face difficulties that lead to interrupting their studies, whether for economic reasons, lack of adaptation to the academic environment or other personal and institutional factors.

This quantitative case study focuses on analyzing dropout patterns at the Polytechnic Institute of Santarém (IPSantarém), using administrative data from the Information System for Aggregated Management of Academic Resources and Records (SIGARRA). Through ETL (Extract, Transform and Load) processes and the use of Business Intelligence (BI) tools, namely Power BI, visualizations and analyses were developed to identify trends, risk profiles, and variables associated with dropout. Complementary analytical approaches were also tested, incorporating Artificial Intelligence (AI) techniques to support trend forecasting.

The results obtained through the analysis of SIGARRA data and the visualizations built in Power BI allow for the identification of clear dropout patterns at IPSantarém. It was observed that dropout tends to concentrate in certain study cycles, courses and age groups, and is also influenced by factors such as whether the student is displaced. The incorporation of Artificial Intelligence techniques allowed for a deeper understanding of the relationships between variables, contributing to the identification of risk profiles and reinforcing the usefulness of predictive approaches.

Based on the results obtained, it is recommended that institutional strategies be developed targeting the highest-risk groups, namely displaced students, those with low academic performance, and those who leave the institute without canceling their enrollment. The creation of more structured welcoming programs, continuous

pedagogical support, and early warning systems supported by AI predictive models could contribute to reducing dropout rates. It is also suggested that IPSantarém invest in the continuous improvement of the collection and consolidation of administrative data, in order to allow for more robust analyses and the development of automatic monitoring mechanisms.

Keywords: School dropout; Risk profiles; Higher education; Business Intelligence; Power BI; ETL; Artificial Intelligence; Predictive analytics.

Índice

Breve Nota Biográfica.....	2
Agradecimentos	3
Acrónimos/Lista de Siglas.....	4
Resumo.....	5
<i>Abstract</i>	7
Índice.....	9
Índice de Figuras.....	12
1. Introdução.....	14
2. Revisão de Literatura.....	18
2.1. Abandono Escolar.....	18
2.2. Estado da Arte.....	22
2.2.1. Abandono Escolar.....	22
2.2.2. Aplicações de <i>BI</i> e <i>IA</i> na educação.....	26
2.2.3. Estudos Semelhantes em Portugal ou na Europa	27
3. Metodologia.....	30
3.1. Introdução	30
3.2. Estratégia de investigação.....	30
3.3. Recolha de dados.....	31
3.4. Modelação dos Dados	32
3.5. Visualização e Ferramentas Analíticas.....	35
3.5.1. Formulação dos <i>Dashboards</i> em <i>Power BI</i>	35
3.5.2. Visualização de Dados	37
3.5.3. <i>Python/ IA/ Modelos de Previsão</i>	39

3.5.4.	Considerações éticas e proteção de dados	39
3.5.5.	<i>Justificações Técnicas</i>	40
4.	Resultados	42
4.1.	Introdução	42
4.2.	Resultados Globais do Abandono Escolar	42
4.2.1.	Indicadores Gerais	42
4.2.2.	Perfil Agregado do Abandono	44
4.2.3.	Relações Gerais entre Variáveis	45
4.3.	Resultados por Variáveis Demográficas	46
4.3.1.	Idade	46
4.3.2.	Gênero	47
4.3.3.	Nacionalidade	48
4.3.4.	Estatuto Académico	50
4.3.5.	Interações Demográfico-Académicas	52
4.4.	Resultados por Variáveis Académicas	55
4.4.1.	Escola/Unidade Orgânica	55
4.4.2.	Grau Académico	56
4.4.3.	Curso	57
4.4.4.	Ano Curricular	58
4.4.5.	Tipo de Anulação	59
4.4.6.	Motivo de Anulação	62
4.5.	Tendências Temporais	63
4.5.1.	Evolução Anual	63
4.5.2.	Distribuição Mensal	64
4.5.3.	Diferença entre Pedido e Aceitação	65
4.5.4.	Padrões Sazonais	67

4.6.	Resultados Complementares e Explorações Específicas	69
4.6.1.	Distribuição Geográfica	69
4.6.2.	Estudantes Não Inscritos no Ano Seguinte.....	70
4.7.	Modelos Analíticos e Preditivos	71
4.7.1.	Modelo <i>ARIMA</i> – Tendência Agregada.....	71
4.7.2.	Impacto Institucional	72
4.8.	Síntese Final dos Resultados	73
5.	Discussão	74
5.1.	Interpretação dos Resultados	74
5.2.	Relação com a Literatura.....	76
5.3.	Implicações Práticas e Teóricas.....	77
5.4.	Limitações do Estudo.....	78
5.5.	Sugestões para Trabalhos Futuros	79
5.6.	Conclusão da Discussão	79
6.	Conclusão e Recomendações	81
6.1.	Síntese das Conclusões	81
6.1.1.	Cumprimento dos Objetivos	81
6.1.2.	Principais Descobertas	82
6.2.	Contribuições e Valor do Estudo	82
6.3.	Recomendações para o <i>IPSantarém</i>	83
6.4.	Sugestões para Futuras Pesquisas.....	83
6.5.	Considerações Finais.....	84
	Referências	85

Índice de Figuras

Figura 1 - Indicadores Abandono Escolar.....	43
Figura 2 - Número Total de Estudantes e Cursos por Grau Académico.....	44
Figura 3 - Tabela Abandono Escolar.....	45
Figura 4 - Resumo Abandono Escolar	46
Figura 5 - Abandono por Idade	47
Figura 6 - Abandono por Género	48
Figura 7 - Abandono por Nacionalidade	49
Figura 8 - Abandono por CPLP	50
Figura 9 - Abandono por Trabalhador-Estudante	51
Figura 10 - Abandono por Deslocado	51
Figura 11 - Abandono por Bolsa de Estudo.....	52
Figura 12 - Abandono por Idade x Escola.....	53
Figura 13 - Abandono por Género x Curso	54
Figura 14 - Abandono por Grau x Nacionalidade.....	55
Figura 15 - Abandono por Escola	56
Figura 16 - Abandono por Grau Académico.....	57
Figura 17 - Abandono e Total Estudantes por Curso	58
Figura 18 - Abandono por Ano Curricular	59
Figura 19 - Abandono por Tipo de Anulação	60
Figura 20 - Abandono por Tipo Anulação x Ano	61
Figura 21 - Abandono por Motivo.....	62
Figura 22 - Abandono por Ano.....	63
Figura 23 - Abandono Percentual por Ano	64
Figura 24 - Abandono por Mês	65
Figura 25 - Diferença entre Pedido e Inserção de Anulação em Meses	66
Figura 26 - Abandono por Pedido e Inserção x Ano	67
Figura 27 - Abandono por Mês x Ano	69
Figura 28 - Abandono por Concelho e Distrito	69
Figura 29 - Abandono por Top 5 Distritos e Concelhos	70

Figura 30 - Abandono por Não Inscrito.....	71
Figura 31 - Abandono por Ano (Modelo de Previsão)	72
Figura 32 - Abandono por Motivos (Modelo de Impacto).....	73

1. Introdução

O abandono escolar no ensino superior é um dos problemas mais preocupantes para as instituições de ensino, afetando diretamente a sua qualidade e a trajetória de vida dos estudantes. Trata-se de um fenómeno multidimensional, influenciado por fatores económicos, académicos, psicológicos e institucionais, que afeta tanto o desenvolvimento individual dos estudantes como a sustentabilidade e qualidade das instituições. Quando um estudante desiste do curso, as razões que levam a essa decisão podem ser várias, desde dificuldades financeiras e falta de apoio académico até problemas de adaptação e desmotivação.

O abandono escolar no ensino superior continua a representar um desafio significativo para as instituições académicas, incluindo o Instituto Politécnico de Santarém (*IPSantarém*). Apesar da disponibilidade de dados administrativos que poderiam ser explorados para compreender melhor os padrões de desistência, estes têm sido subutilizados, dificultando a identificação de perfis de risco, tendências temporais e causas institucionais que influenciam a decisão dos estudantes de abandonar os estudos.

Este trabalho pretende contribuir para a compreensão do fenómeno do abandono escolar, explorando os fatores que influenciam essa decisão e propondo medidas para a minimização deste fenómeno. A relevância deste estudo reside na possibilidade de gerar conhecimento útil e aplicável a outras instituições, para implementarem estratégias de intervenção mais eficazes e personalizadas. Este estudo foca-se especificamente no abandono escolar no Instituto Politécnico de Santarém, analisando dados provenientes do instituto e utilizando ferramentas de análise para ajudar na investigação.

Este estudo distingue-se pela utilização de dados administrativos institucionais, analisados através de ferramentas de *Business Intelligence* e modelos analíticos exploratórios, com o objetivo de apoiar a monitorização e compreensão do abandono escolar no ensino superior, no contexto do ensino politécnico.

Assim o problema central que orienta este estudo é o seguinte:

Quais são os principais padrões do abandono escolar no *IPSantarém* e de que forma as ferramentas de Business Intelligence e Inteligência Artificial podem ser usadas para apoiar a tomada de decisão institucional para a mitigação deste fenómeno?

Para atingir o objetivo, este estudo procura:

1. Analisar os dados administrativos do *IPSantarém*, de modo a identificar fatores e padrões associados ao abandono escolar;
2. Desenvolver visualizações no *Power BI* que facilitem a interpretação clara e intuitiva dos dados analisados;
3. Explorar a utilização de técnicas de Inteligência Artificial (*IA*) para a construção de modelos preditivos exploratórios e para a análise do impacto institucional dos motivos de abandono;
4. Propor recomendações práticas e estratégias de mitigação com base nos padrões identificados.

De forma complementar, o estudo procura responder às seguintes questões de investigação:

- Q1: Que padrões e fatores se encontram associados ao abandono escolar no *IPSantarém*, considerando características demográficas, académicas e institucionais dos estudantes?

- Q2: Em que medida os dados administrativos disponíveis permitem identificar tendências temporais e estimar cenários futuros relacionados com o abandono escolar?
- Q3: Como podem ferramentas de *Business Intelligence (BI)* e abordagens analíticas em Inteligência Artificial apoiar a monitorização, compreensão e prevenção do abandono escolar no *IPSantarém*?

Metodologicamente, o estudo assume a forma de um **estudo de caso** aplicado, de **natureza quantitativa e exploratória**, centrado no Instituto Politécnico de Santarém. A investigação baseia-se na análise de dados secundários provenientes do Sistema de Informação para Gestão Agregada dos Recursos e dos Registos Académicos (*SIGARRA*), que forneceu informação administrativa essencial sobre o percurso académico dos estudantes. Para o tratamento e exploração dos dados, foram utilizadas ferramentas de *Business Intelligence*, com destaque para o *Power BI*, onde foram realizados processos de *ETL*, construção de visualizações e análise de indicadores. Complementarmente, recorreu-se à aplicação de modelos exploratórios em *Python*, permitindo aprofundar padrões, relações entre variáveis e potenciais tendências associadas ao abandono escolar.

O estudo apresenta uma relevância significativa a nível científico, institucional, social e educacional. A nível científico contribui para a compreensão do fenómeno do abandono escolar no ensino superior, aprofundando a identificação de padrões, fatores de risco e dinâmicas que influenciam a permanência dos estudantes. A nível institucional, o estudo oferece ferramentas práticas que podem apoiar a tomada de decisão, permitindo às escolas e ao *IPSantarém* desenvolver estratégias mais eficazes de monitorização, prevenção e intervenção precoce. Finalmente, a pesquisa tem valor social e educativo, ao contribuir para a promoção do sucesso académico, reforçando a igualdade de oportunidades e apoiando a criação de políticas que diminuam as desigualdades que afetam o percurso dos estudantes.

O relatório está organizado em cinco capítulos:

1. **Revisão de Literatura:** neste capítulo será feita uma introdução ao tema Abandono Escolar no Ensino Superior, a sua definição, fatores, consequências tanto pessoais como institucionais, formas de possível mitigação, os valores de abandono em Portugal, o estado da arte, onde serão mostrados modelos teóricos e empíricos utilizados para estudar o tema do Abandono Escolar, as aplicações que *Business Intelligence* e Inteligência Artificial podem ter na educação e estudos parecidos em Portugal e na Europa;
2. **Metodologia:** este capítulo explica o tipo de investigação realizada, as etapas do processo, (*ETL*, modelação, visualização), o que é o *BI* e o *Power BI*, de que maneira este foi utilizado no estudo, descreve os dados usados e diversas justificações técnicas;
3. **Resultados:** neste capítulo são apresentados os *dashboards* utilizados no estudo, e os padrões identificados;
4. **Discussão:** este capítulo tem como objetivo discutir os resultados alcançados no capítulo anterior, encontrando possíveis soluções ao Abandono Escolar e possíveis pontos de foco;
5. **Conclusão:** resume os principais contributos, limitações e recomendações para investigação futura e aplicação institucional.

2.Revisão de Literatura

2.1. Abandono Escolar

O abandono escolar é um fenômeno complexo que tem afetado diversos sistemas educacionais e estudantes em todo o mundo, com consequências significativas a nível individual, institucional e social. Em termos gerais, o abandono escolar envolve a desistência dos estudos antes da conclusão de um determinado ciclo de ensino, comprometendo o desenvolvimento acadêmico e profissional dos estudantes.

Apesar de parecer um problema em conceito simples, o abandono escolar é influenciado por uma multiplicidade de fatores que tornam a sua compreensão e mitigação particularmente exigentes, tais como:

- **Fatores Socioeconômicos:** dificuldades financeiras (Gairín et al., 2014; Núñez-Naranjo, 2024), podendo haver necessidade de trabalhar para ajudar a família, pagar as propinas e em alguns casos pagar casa ou quarto.
- **Fatores Pedagógicos:** metodologia de ensino desmotivante (Gairín et al., 2014), currículo descontextualizado (Araque et al., 2009; Glomo-Narzoles & Palermo, 2021), falta de apoio institucional (Glomo-Narzoles & Palermo, 2021; Urbina-Nájera et al., 2020).
- **Fatores Pessoais e Familiares:** problemas de saúde (Glomo-Narzoles & Palermo, 2021), conflitos familiares (Rodríguez-Pineda & Zamora-Araya, 2021), falta de motivação (Bernardo et al., 2016; Casanova et al., 2018) e problemas de integração (Bernardo et al., 2016).
- **Fatores Institucionais:** infraestrutura inadequada, falta de professores qualificados, ambiente escolar hostil devido a histórico sociodemográfico e características pessoais (Kehm et al., 2019).

Em síntese, a literatura converge na compreensão do abandono escolar como um fenómeno multidimensional e cumulativo, resultante da interação entre fatores pessoais, socioeconómicos, pedagógicos e institucionais. Estes fatores raramente atuam de forma isolada, reforçando-se mutuamente ao longo do percurso académico dos estudantes. Esta natureza complexa dificulta a identificação de causas únicas e evidencia a necessidade de abordagens analíticas integradas, capazes de considerar simultaneamente múltiplas dimensões e a sua evolução temporal, particularmente no contexto do ensino superior.

A nível das instituições de ensino e da sociedade o abandono escolar tem um impacto significativo no seu desenvolvimento, tendo consequências como:

- **Redução dos Índices de Sucesso Académico:** escolas e universidades enfrentam uma diminuição nas taxas de conclusão dos cursos, comprometendo a qualidade do ensino oferecido (Gomme & Gilbert, 1984; Scott et al., 2008).
- **Impacto Económico:** o governo pode diminuir o financiamento institucional face a diminuição de retenção de estudantes (Simpson, 2005).
- **Problemas no planeamento de estratégias de retenção:** o abandono escolar influencia como as instituições planeiam as suas estratégias, devido a terem de considerar as consequências financeiras nas suas práticas de retenção à medida que competem por estudantes e financiamento (Scott et al., 2008; Simpson, 2005).

O abandono escolar tem um impacto significativo na vida futura dos estudantes, incluindo:

- **Dificuldades no Mercado de Trabalho:** estudantes sem formação académica adequada enfrentam maiores dificuldades para conseguir emprego e recebem salários inferiores em comparação com os graduados (Torres et al., 2024).

- **Problemas Psicológicos e Sociais:** a falta de escolarização pode estar associada a um aumento de problemas autoestima, depressão e dificuldades de integração social (Torres et al., 2024).
- **Maior Vulnerabilidade a Comportamentos de Risco:** estudantes que desistem podem enfrentar consequências pessoais, sociais e económicas, incluindo potenciais impactos na autoestima e no consumo de álcool (McCaul et al., 1992).

Os impactos do abandono escolar, descritos de forma consistente em diferentes contextos nacionais e internacionais, evidenciam que este fenómeno ultrapassa a esfera individual do estudante, assumindo uma dimensão estrutural para as instituições de ensino e para a sociedade. A perda de capital humano, o aumento da ineficiência institucional e as consequências sociais associadas reforçam a importância de estratégias sustentadas por evidência empírica, particularmente em instituições de ensino superior que enfrentam desafios crescentes ao nível da retenção e do sucesso académico.

Para haver mitigação do abandono escolar é necessária uma abordagem integrada que envolva a atuação de diferentes atores: instituições de ensino, docentes, serviços de apoio ao estudante, governo e os próprios estudantes. Entre os métodos mais eficazes destacam-se:

- a) **Intervenção Precoce e Monitorização de Indicadores:** A criação de sistemas de alerta precoce baseados em dados (como presença em aulas, notas, uso de plataformas digitais) permite identificar estudantes em risco e intervir antes que a situação se torne irreversível. Ferramentas baseadas em Inteligência Artificial e modelos preditivos têm ganho destaque nesta área.
- b) **Apoio Psicopedagógico:** Serviços de apoio psicológico, sessões de *coaching* académico, e programas de tutoria são essenciais para acompanhar estudantes que demonstram sinais de desmotivação, ansiedade ou dificuldades de adaptação.

- c) **Apoio Financeiro e Logístico:** A atribuição de bolsas de estudo, a existência de alojamento universitário acessível e programas de apoio ao transporte e alimentação são medidas que permitem aliviar o peso financeiro sobre os estudantes.
- d) **Flexibilização Curricular e Adaptação Pedagógica:** A possibilidade de realizar os cursos em regime parcial, a criação de unidades curriculares de nivelamento e a diversificação dos métodos de avaliação podem promover maior inclusão e adaptação às diferentes realidades dos estudantes.
- e) **Promoção do Sentido de Pertença e Envolvimento Estudantil:** Projetos extracurriculares, grupos de estudo, atividades culturais e desportivas contribuem para o bem-estar dos estudantes e fortalecem os laços com a comunidade académica, diminuindo o sentimento de isolamento.

A literatura sugere que a mitigação eficaz do abandono escolar requer uma abordagem integrada e sistémica, combinando intervenções pedagógicas, psicossociais, financeiras e institucionais. Neste contexto, ganha relevância a utilização de dados institucionais e de sistemas de monitorização contínua, que permitem identificar precocemente sinais de risco e apoiar intervenções mais direcionadas. A articulação entre abordagens humanizadas e soluções tecnológicas surge, assim, como um elemento central para o desenvolvimento de políticas de retenção mais eficazes no ensino superior.

Segundo os dados divulgados no Portal *Infocursos*, a percentagem de estudantes que ingressaram numa Licenciatura e não concluíram o primeiro ano registou uma diminuição no setor público durante a época letiva de 2023/2024, quando comparada com os anos anteriores. Em 2023/2024, a taxa de abandono situou-se nos 11,4% (RTP, 2025), uma redução face aos 11,6% verificados em 2021/2022 (RTP, 2024) e aos 11,73% em 2022/2023 (DN/Lusa, 2024). Esta tendência decrescente pode indicar uma ligeira melhoria nas condições de integração e retenção dos estudantes no ensino superior público, embora os valores ainda revelem uma proporção significativa de estudantes que abandonam o curso logo no primeiro ano. Estes dados sublinham a importância de

continuar a investir em políticas de apoio ao sucesso acadêmico, especialmente nos primeiros meses do percurso universitário, fase crítica na adaptação dos estudantes ao novo ambiente educativo.

Deste modo, torna-se evidente que o abandono escolar é um problema multifacetado que requer ações conjuntas por parte das instituições de ensino, governos e sociedade. Políticas de prevenção, apoio pedagógico e inclusão social são fundamentais para reduzir as taxas de desistência e minimizar as suas consequências. A implementação de programas de apoio psicopedagógico e estratégias de ensino mais dinâmicas pode ser uma solução eficaz para enfrentar este desafio.

2.2. Estado da Arte

2.2.1. Abandono Escolar

Como mencionado antes, o abandono escolar é um problema complexo e multidimensional, tornando-o num problema que pode ser estudado tanto por modelos teóricos e modelos empíricos. Estes modelos oferecem contributos relevantes, mas também apresentam limitações que abrem espaço para novas abordagens, nomeadamente através de ferramentas de *Business Intelligence* e Inteligência Artificial.

2.2.1.1. Modelos Teóricos

Os modelos teóricos visam explicar as razões dinâmicas do abandono, destacando fatores individuais, institucionais e contextuais, chegando cada um ao seu conceito de abandono escolar, tais como:

- **Modelo de Integração do Estudante – Tinto, 1975:**

- Considera o abandono como resultado da falta de integração social e acadêmica do estudante na instituição;
- O processo é visto como dinâmico, envolvendo interações entre características pessoais do estudante, características institucionais e o contexto social;
- A integração social (relações com pares) e a integração acadêmica (desempenho, identificação com o curso) são decisivas para a permanência (Casanova J. , 2018; Tinto V. , 1975; Tinto V. , 2010);
- **Modelo do Atrito do Estudante – *Bean*, 1980:**
 - Inspirado em Tinto, mas com ênfase em processos sociocognitivos, como autoeficácia, autoconceito e identidade;
 - Destaca a importância dos fatores ambientais e familiares, além do tempo e envolvimento do estudante nas atividades da instituição (Bean J. , 2001; Bean J. , 1980; Casanova J. , 2018);
- **Modelo Integrado de Retenção do Estudante – *Nora e Castañeda*, 1992:**
 - Integra elementos dos modelos de *Tinto* e *Bean*;
 - Considera fatores como apoio financeiro, suporte familiar e de pares, desempenho acadêmico, integração social e intenção declarada de permanecer ou abandonar (Cabrera et al., 1992; Casanova J. , 2018);
- **Modelo de Envolvimento – *Alexander Astin*, 1984:**
 - À semelhança do *Tinto*, o envolvimento do estudante é entendido como o investimento de energia física e psicológica, variando conforme as tarefas e o contexto institucional (Astin, 1984; Casanova J. , 2018).

Apesar da sua relevância conceptual, os modelos teóricos do abandono escolar apresentam limitações quando aplicados a contextos institucionais específicos e a dados administrativos em larga escala. Estes modelos privilegiam frequentemente dimensões subjetivas, como a integração social ou a motivação individual, que nem sempre se encontram disponíveis nos sistemas de informação académicos. Assim, embora forneçam um enquadramento interpretativo fundamental, revelam-se insuficientes para

suportar análises operacionais e preditivas baseadas em dados institucionais, abrindo espaço para abordagens complementares assentes em *Business Intelligence* e Inteligência Artificial.

2.2.1.2. Modelos Empíricos

Os modelos empíricos buscam operacionalizar e testar as hipóteses levantadas pelos modelos teóricos, geralmente por meio de instrumentos de pesquisa quantitativos e qualitativos:

- **Questionário de Motivos de Abandono do Ensino Superior (QMA-ES):**
 - Instrumento validado em diferentes países (Brasil, Espanha, Portugal), avalia, em escala *Likert*, a importância de diferentes fatores para o abandono, agrupados em várias dimensões (financeira, acadêmica, institucional, pessoal, etc) (Almeida & Niemba, 2022).
- **Estudos de Caso e Perfis Sociológicos:**
 - Pesquisas qualitativas, como entrevistas e estudos de caso, permitem traçar perfis de estudantes desistentes e identificar trajetórias e projetos de vida incompatíveis com a permanência no ensino superior;
 - Tais estudos evidenciam a influência de fatores econômicos, familiares, expectativas profissionais e políticas públicas (Da Silva Araújo, 2018; Ditutala, 2015).
- **Análises Quantitativas:**
 - Aplicação de inquéritos e análise estatística de dados institucionais para identificar padrões de evasão, perfis de risco e fatores preditores, como rendimento acadêmico, idade, situação financeira e integração social (Da Silva Araújo, 2018).

Os estudos empíricos frequentemente apontam como causas principais do abandono:

- Dificuldades financeiras e contexto socioeconómico familiar (Da Silva Araújo, 2018; Ditutala, 2015);
- Falta de integração institucional e social (Casanova J. , 2018; Ditutala, 2015);
- Baixa adaptação ao curso e insatisfação com a escolha académica (Ditutala, 2015);
- Projetos pessoais e profissionais incompatíveis com a permanência (Da Silva Araújo, 2018);
- Mudança de instituição ou de curso (Ditutala, 2015);
- Condições do mercado de trabalho e desvalorização do diploma (Da Silva Araújo, 2018).

Embora fundamentais, estes métodos revelam limitações: baseiam-se em instrumentos muitas vezes pontuais (inquéritos, entrevistas) e pouco adaptados a uma monitorização contínua e preditiva. Ou seja, ajudam a explicar o passado, mas não a antecipar de forma eficaz o risco futuro de abandono. Aqui as ferramentas de *BI* e *IA* podem acrescentar valor, permitindo identificar padrões ocultos em grandes volumes de dados e apoiar decisões em tempo real.

Os modelos empíricos constituem um contributo essencial para a validação das abordagens teóricas, permitindo identificar padrões e fatores associados ao abandono escolar com base em dados recolhidos junto dos estudantes. No entanto, a dependência de instrumentos pontuais, como inquéritos e entrevistas, limita a capacidade de monitorização contínua e de antecipação do risco de abandono. Esta limitação evidencia a necessidade de explorar dados administrativos longitudinais e abordagens analíticas mais avançadas, capazes de apoiar decisões institucionais em tempo útil.

2.2.2. Aplicações de *BI* e *IA* na educação

2.2.2.1. Aplicações de *BI* na Educação

O *Business Intelligence* tem vindo a ser aplicado na educação em diferentes frentes:

- **Monitorização de desempenho académico:** o *BI* permite analisar resultados dos alunos em tempo real, identificar padrões de desempenho e antecipar dificuldades, possibilitando intervenções pedagógicas mais eficazes (Aranha, 2022; Felix, 2024; Da Silva Atila & Ecléa, 2024);
- **Personalização do ensino:** com dados detalhados, é possível adaptar o currículo e os métodos de ensino às necessidades individuais dos estudantes, aumentando a retenção e o engajamento (Aranha, 2022; Felix, 2024; Da Silva Atila & Ecléa, 2024);
- **Gestão eficiente de recursos:** *BI* auxilia na alocação de recursos financeiros, humanos e materiais, otimizando custos e melhorando a eficiência administrativa das instituições (Aranha, 2022; Felix, 2024);
- **Tomada de decisão baseada em dados:** relatórios e *dashboards* fornecem informações precisas para gestores, facilitando decisões estratégicas sobre matrículas, turmas, infraestrutura e políticas educacionais (Felix, 2024; Da Silva Atila & Ecléa, 2024);
- **Análise preditiva:** *BI* pode ser utilizado para prever tendências, como evasão escolar, necessidades futuras de formação e mudanças no perfil dos alunos, permitindo ações proativas (Da Silva Atila & Ecléa, 2024)

No entanto, em Portugal, o uso de *BI* ainda é incipiente e pouco sistematizado, restringindo-se maioritariamente a *dashboards* de desempenho e relatórios administrativos, sem grande aprofundamento na previsão de abandono escolar.

2.2.2.2. Aplicações de IA na Educação

A Inteligência Artificial, pode ser utilizada para:

- **Avaliação automatizada e feedback imediato:** softwares de IA corrigem provas e trabalhos rapidamente, fornecendo análises detalhadas do desempenho e identificando áreas de melhoria (Ferreira F. , 2024);
- **Desenvolvimento de habilidades:** plataformas com IA estimulam competências como resolução de problemas, lógica e criatividade ajustando o nível de desafio conforme o progresso do aluno (Ferreira F. , 2024; Fernando Giannini, 2025);
- **Inclusão e acessibilidade:** Ferramentas de reconhecimento de voz, tutores virtuais e recursos multimídia tornam o ensino mais acessível para alunos com necessidades especiais (Fernando Giannini, 2025).

Apesar do reconhecido potencial do BI e da IA no contexto educativo, a literatura indica que a sua aplicação no ensino superior permanece, em muitos casos, limitada a análises descritivas e relatórios estáticos. A exploração de abordagens preditivas e de apoio à decisão continua a ser incipiente, particularmente no contexto português, reforçando a relevância de estudos aplicados que demonstrem o valor destas tecnologias na compreensão e mitigação de fenómenos complexos como o abandono escolar.

2.2.3. Estudos Semelhantes em Portugal ou na Europa

A nível nacional, os estudos concentram-se em três áreas principais:

1. **Desempenho Académico e Conquistas:** pesquisadores apontam que baixo desempenho, especialmente no primeiro ano, bem como reprovações em disciplinas consideradas menos exigentes e a matrícula em cursos que não eram

a primeira escolha, vinculam-se decisivamente ao abandono (De Sousa et al., 2018);

- 2. Adaptação e Integração do Estudante:** relatos consistentemente associam dificuldades de adaptação acadêmica e social – incluindo a inadequação entre vocação e curso escolhido – a um risco elevado de evasão (Ferreira et al., 2022);
- 3. Fatores Institucionais e de Suporte:** Estudos evidenciam que elementos como flexibilidade de horários, serviços de apoio e a estrutura do currículo influenciam a permanência dos alunos (Almeida et al., 2008).

Vários estudos na Europa definem o abandono escolar no ensino superior como um fenómeno multifatorial no qual os principais são:

1. Fatores Académicos:

- a. Dificuldade do curso, falta de apoio académico, adaptação académica e desempenho académico são fatores importantes no abandono (Andreu, 2008; Kehm et al., 2019);
- b. A não escolha da área preferida ou a orientação profissional inadequada (Andreu, 2008).

2. Fatores Socioeconómicos:

- a. A incompatibilidade entre trabalho e estudo é um fator de risco, especialmente para estudantes que trabalham mais de 15 a 17 horas por semana (Andreu, 2008);
- b. Dificuldades económicas e falta de apoio financeiro eram temas recorrentes (Cabrera et al., 2014).

3. Fatores Institucionais:

- a. Apoio institucional, a orientação e os métodos de ensino foram cada vez mais reconhecidos como essenciais em vários estudos sobre o abandono escolar (Andreu, 2008);
- b. O papel da mentoria académica (Pérez & Aguilar, 2017) e a responsabilidade institucional pela retenção de estudantes (Gil, 2014).

Embora consistentes, estes estudos tendem a limitar-se a análises estatísticas clássicas. São raras as investigações que utilizam ferramentas de *BI* ou *IA* para transformar grandes volumes de dados académicos em conhecimento preditivo.

Apesar de existirem numerosos estudos sobre o abandono escolar no ensino superior, em Portugal as investigações que recorrem a *Business Intelligence* e Inteligência Artificial ainda são escassas. Um exemplo é o estudo “*Predicting and Preventing School Dropout with Business Intelligence: Insights from a Systematic Review*” (Córdova-Esparza et al., 2025) que combinou a técnicas de *machine learning*, *data mining*, e *Power BI*, para identificar padrões de abandono e propor intervenções. Contudo, abordagens deste tipo continuam a ser exceção.

Em síntese, embora a literatura nacional e europeia ofereça contributos relevantes para a compreensão do abandono escolar no ensino superior, persistem lacunas significativas no recurso sistemático a dados administrativos e a ferramentas de *BI* e *IA* para a análise deste fenómeno. Em particular, no contexto português, são escassos os estudos que integrem abordagens analíticas avançadas com uma orientação aplicada à tomada de decisão institucional. O presente estudo procura responder a esta lacuna, propondo uma análise baseada em dados administrativos do *IPSantarém*, com recurso a metodologias analíticas e preditivas, contribuindo para uma abordagem mais robusta, integrada e orientada para a ação.

3. Metodologia

3.1. Introdução

A metodologia adotada neste estudo foi delineada para garantir uma análise rigorosa, objetiva e orientada para a compreensão dos padrões associados ao abandono escolar no *IPSantarém*. Este capítulo apresenta a estratégia de investigação utilizada, os procedimentos adotados, o processo de recolha e tratamento de dados, os dados utilizados, bem como as tecnologias e modelos aplicados. Adicionalmente são discutidas as limitações metodológicas que condicionaram o desenvolvimento da investigação e que devem ser consideradas na interpretação dos resultados.

3.2. Estratégia de investigação

Este estudo segue uma abordagem quantitativa, descritiva e exploratória, estruturada como um estudo de caso centrado no *IPSantarém*.

A opção por um estudo de caso justifica-se pela necessidade de analisar o fenómeno do abandono escolar num contexto institucional específico, utilizando dados reais provenientes de sistemas administrativos. Esta estratégia permite identificar padrões, tendências e relações relevantes sem recorrer a métodos experimentais ou inferenciais complexos, alinhando-se assim com a natureza aplicada deste trabalho e com os objetivos do projeto 3C – Conectar, Colaborar e Crescer (Ficha de Projeto, 2019).

O Projeto 3C: Conectar, Colaborar e Crescer visa fomentar o sucesso académico, diminuir o insucesso e prevenir o abandono escolar dos estudantes do Instituto Politécnico de Santarém, através de uma abordagem integrada, holística e inclusiva, que

envolve toda a comunidade educativa. Com uma duração prevista de três anos, esta iniciativa contempla a implementação de diversas ações de capacitação, momentos de encontro, bem como a criação de conteúdos e ferramentas tecnológicas. Estas medidas pretendem assegurar, a longo prazo, uma intervenção inovadora e sustentável, capaz de apoiar a identificação precoce de situações de risco de abandono, desenvolver estratégias eficazes de resposta e promover o sucesso académico e pessoal dos estudantes do *IPSantarém*.

A abordagem quantitativa foi selecionada porque os dados disponíveis no *SIGARRA* são de natureza numérica, longitudinal e administrativa, permitindo uma análise sistemática do comportamento dos estudantes ao longo dos anos. Esta estratégia metodológica é suportada pela literatura que destaca o potencial de dados administrativos e de técnicas de visualização analítica para compreender o abandono escolar e apoiar decisões institucionais (Ficha de Projeto, 2019).

O estudo não pretende estabelecer relações causais, mas identificar padrões e associações relevantes para suporte à decisão.

3.3. Recolha de dados

Os dados utilizados neste estudo foram disponibilizados pelo *IPSantarém* e extraídos do sistema *SIGARRA*, correspondendo a dados administrativos estruturados relativos ao percurso académico dos estudantes. A extração foi realizada através de uma *query SQL*, sendo os dados posteriormente importados e integrados no *Power BI* numa tabela principal.

A tabela, denominada por **Estudantes**, inclui os seguintes campos:

- Ano Letivo
- Ano Curricular;
- Ano Letivo da Matrícula;
- Código do Curso;
- Estado da Bolsa;
- Concelho;
- Nome do Curso;
- Data da Anulação (data do pedido da anulação);
- Data da Inserção da Anulação (data que a anulação foi aceite);
- Data de Nascimento do Estudante;
- Deslocado Estudante;
- Distrito;
- Código da Escola;
- Nome da Escola;
- Sigla do Estado Atual;
- Estado Atual;
- Género;
- Grau Académico;
- Motivo da Anulação;
- Nacionalidade;
- Nome do Estudante;
- Número do Estudante;
- Observações;
- Regime de Ingresso;
- Tipo de Estudante;
- Trabalhador Estudante.

3.4. Modelação dos Dados

A modelação no contexto deste projeto não segue uma estrutura de modelação dimensional, uma vez que o objetivo principal é a análise e previsão de dados baseando-se numa única fonte principal. Por isso, optou-se por um método mais simples e objetivo, com o objetivo de facilitar o entendimento e tornar a análise mais eficiente, evitando complexidade desnecessária.

Entretanto ao longo do desenvolvimento dos gráficos foi necessário criar colunas e medidas na tabela, sendo estas:

- **Colunas:**

- Índice (coluna criada no *Power BI* com o intuito de cada linha receber um valor único);
- Bolsa de Estudo (vai buscar os valores da coluna estado da bolsa, como os valores estão em siglas a nova coluna vai designar o nome correspondente à sigla para melhor identificação);
- Concelho Ajustado (esta coluna usa os valores da coluna concelho e adiciona antes o texto “Portugal – “para facilitar na localização do concelho na visualização mapa);
- Distrito Ajustado (esta coluna usa os valores da coluna distrito e adiciona antes o texto “Portugal – “para facilitar a localização do distrito na visualização mapa);
- Deslocado (esta coluna utiliza os dados da coluna deslocado e define se o estudante é deslocado, ou seja, “sim” ou “não”, esta coluna foi necessária pois a coluna “estudante deslocado”, mostra se o estudante é deslocado com “S” ou “N”);
- Diferença Meses (uma coluna que calcula a diferença em meses entre a data que o estudante pede a anulação e a data que é aceite);
- Diferença Dias (uma coluna que calcula a diferença em dias entre a data que o estudante pede a anulação e a data que é aceite);
- Escola Sigla (uma coluna que vai buscar os valores da coluna escola e define a sigla de cada escola);

- Idade (busca a data de nascimento e atribui uma idade ao estudante);
- Mês Anulação (coluna que diz o nome do mês que o aluno pediu a anulação);
- Nº Mês de Anulação (coluna que mostra o número do mês da data de anulação);
- Nº Mês de Inserção (coluna que mostra o número do mês da data de anulação);
- Ordenar Mês (coluna criada com o intuito de ordenar o mês de anulação em visualizações);
- Nome Trabalhador Estudante (uma coluna que vai buscar os valores da coluna Trabalhador Estudante, estes valores estão em siglas a nova coluna vai designar o nome correspondente à sigla para melhor identificação);
- Repetições Nome Anulações (uma coluna criada para descobrir nas anulações quantas vezes aparece o mesmo estudante);
- Alunos que Repetem (esta coluna utiliza a coluna Repetições Nome Anulações para descobrir quantas estudantes anularam mais do que uma vez).

- **Medidas**

- Total de Anulações (conta o número de estudantes no Abandono Escolar);
- Total Inscritos (Mostra o número total de estudantes);
- Percentagem de Anulações (mostra a percentagem de estudantes que anularam nos dados);
- Percentagem de Anulações por Ano (mostra a percentagem de estudantes que anularam por ano);
- Total Não Inscritos (Mostra o total de estudantes que se inscreveram no instituto, mas que no ano seguinte não estão inscritos);
- Percentagem de Não Inscritos (mostra a percentagem de estudantes não inscritos de acordo com o ano);
- Diferença em Meses do Pedido de Anulação e o Pedido de inserção de Anulação (utiliza o *datediff* para calcular a diferença mensal entre o mês do pedido de anulação e o mês de inserção da anulação).

A medida Total Inscritos foi criada de maneira a caso o utilizador filtre o estado atual para a “Anulada a Matrícula” e “Anulada a Inscrição”, de forma isolada, o valor não altere, esta escolha foi feita devido a existir nos *dashboards* um indicador de estudantes que abandonaram o instituto e para calcular a percentagem de anulações.

Estas medidas foram colocadas numa tabela denominada **Medidas Tabela**. Esta tabela tem o único propósito de guardar as medidas criadas, separando as medidas das colunas, organizando os dados, sendo considerado uma boa prática no *Power BI*.

3.5. Visualização e Ferramentas Analíticas

3.5.1. Formulação dos *Dashboards* em *Power BI*

Para realizar a análise do abandono escolar foi necessário a utilização de ferramentas tecnológicas que permitam a recolha, tratamento e visualização de dados de forma eficaz.

A ferramenta utilizada neste estudo foi o *Power BI*, uma ferramenta robusta de *Business Intelligence*, conjunto de processos, tecnologias e ferramentas que permitem a recolha, integração, análise e apresentação de dados com o objetivo de apoiar a tomada de decisão, desenvolvida pela *Microsoft* que permite:

- **Conexão com diversas fontes de dados:** Pode integrar-se com bases de dados *SQL*, ficheiros *Excel*, *APIs*, *Google Analytics* e muitas outras fontes;
- **Transformação e modelação de dados:** A ferramenta *Power Query* permite limpar, transformar e combinar dados de diferentes fontes;

- **Utilização de DAX (*Data Analsys Expressions*):** Uma linguagem de expressão que permite a criação de medidas e cálculos complexos para análises aprofundadas;
- **Publicação e partilha de relatórios:** Através do *Power BI Service*, plataforma baseada em nuvem disponível para utilizadores com o pacote *Power BI Pro* e *Power BI Premium*, os relatórios podem ser partilhados com outros utilizadores de forma segura;
- **Automatização de atualização de dados:** Permite agendar atualizações automáticas, garantindo que os relatórios refletem os dados mais recentes.

No contexto deste estudo sobre o abandono escolar, o *Power BI* desempenhou um papel fundamental nos seguintes aspetos:

- **Conectar-se ao banco de dados** e extrair os dados relevantes sobre o abandono escolar;
- **Transformar e limpar os dados**, garantindo que estejam estruturados de forma adequada para análise;
- **Criar *dashboards* interativos**, permitindo a visualização de padrões e tendências no abandono escolar;
- **Gerar relatórios dinâmicos**, que facilitam a tomada de decisões e a identificação de fatores críticos relacionados ao fenómeno.

Foram criados no *Power BI* um total de 32 *dashboards*. Cada *dashboard* foi desenvolvido com o objetivo de facilitar a compreensão e interpretação dos dados relacionados com o abandono escolar, através de uma representação visual clara, intuitiva e interativa. Estes painéis permitem não só visualizar rapidamente padrões e tendências, como também explorar a informação de forma detalhada e personalizada, adaptando a análise às necessidades de diferentes utilizadores.

Para isso foram criadas segmentações e filtros que permitem analisar os dados por várias dimensões, tais como escola, curso, gênero, nacionalidade, idade, tipo de estudante ou situação de trabalhador-estudante. Esta capacidade de segmentação é fundamental para identificar grupos mais vulneráveis ao abandono escolar, compreender as causas associadas e apoiar a tomada de decisões fundamentadas por parte das instituições de ensino. Além disso, os *dashboards* foram estruturados para promover a leitura progressiva da informação, permitindo que o utilizador inicie a análise com uma visão geral e, a partir daí, aprofunde o estudo em áreas específicas com maior incidência de anulações.

Todos os *dashboards* que possuem um gráfico ou visualização das anulações dos estudantes foram criados com um filtro pré-definido do estado atual dos estudantes, definindo para mostrar apenas estudantes com o estado atual “Anulada a Matrícula” e “Anulada a Inscrição”.

3.5.2. Visualização de Dados

A visualização de dados é a fase final de um processo de *BI*, onde os dados tratados e modelados são convertidos em representações gráficas compreensíveis, com o objetivo de facilitar a interpretação e apoiar a tomada de decisão. No projeto sobre o abandono escolar no ensino superior, esta fase permitiu transformar informações complexas em relatórios e *dashboards* interativos, acessíveis a diferentes tipos de utilizadores.

A construção inicial dos relatórios foi realizada em ambiente *desktop*, utilizando uma ferramenta de *BI* com capacidade de modelação e visualização local. Esta abordagem oferece várias vantagens:

- **Desenvolvimento local rápido:** permite testar transformações, medidas e visualizações de forma imediata.

- **Interface intuitiva:** facilita a criação de gráficos, segmentações e indicadores, mesmo para utilizadores com menor experiência técnica.
- **Ligação a múltiplas fontes de dados:** possibilita a integração de várias tabelas num mesmo projeto, com controlo total sobre o modelo de dados.

Após a criação dos relatórios, é comum a necessidade de os partilhar com outros utilizadores, seja com docentes, direção de escola ou serviços administrativos. O processo de partilha pode ocorrer de várias formas:

- **Exportação em PDF ou PowerPoint,** para apresentação estática dos dados.
- **Exportação do ficheiro do projeto,** para que outro utilizador o abra em ambiente desktop.
- **Publicação em plataforma online (quando aplicável),** para visualização em tempo real através de navegadores web, com interatividade.

No contexto do projeto, os ficheiros foram criados localmente, mas a estrutura foi pensada de forma a partilhar os ficheiros com as pessoas envolvidas no projeto, seja na *cloud*/ em nuvem ou partilha direta de ficheiros.

Para garantir que os dados apresentados nos relatórios estejam sempre atualizados, foi prevista/ parametrizada a possibilidade de atualização periódica dos dados através de atualizações automáticas agendadas, permitindo definir horários para atualização, assegurando que os *dashboards* apresentam sempre os dados mais recentes.

Os relatórios estão criados de forma a que quando forem atualizados novos resultados, os gráficos atualizam após a atualização dos dados não sendo necessário modificar nada. No entanto, em caso de necessidade de adicionar novos campos ou filtros, é possível editar estes relatórios de acordo o que é pedido.

3.5.3. Python/ IA/ Modelos de Previsão

A integração da Inteligência Artificial com o *Power BI* abre novas possibilidades para a análise preditiva e prescritiva no ensino superior. Através da aplicação de técnicas de IA, é possível ir além da simples observação do que aconteceu, e poder prever o que poderá acontecer e recomendar ações adequadas. Alguns exemplos de aplicações incluem:

- **Modelos preditivos de abandono escolar**, que identificam estudantes com maior probabilidade de desistência com base em dados históricos;
- **Agrupação de perfis de estudantes**, agrupando indivíduos com características semelhantes (idade, percurso acadêmico, estatuto social) para melhor segmentação das estratégias de apoio;
- **Análise de causas prováveis**, que atribui pesos a diferentes fatores no risco de abandono, como desempenho, distância geográfica, ou apoio financeiro;
- **Sistemas de alerta precoce**, baseados em algoritmos que sinalizam comportamentos ou padrões que merecem intervenção.

Estes modelos contribuem para tomar decisões automatizadas e personalizada, permitindo às instituições atuar preventivamente na retenção dos seus estudantes.

No contexto do projeto foi utilizado um modelo preditivo de abandono escolar, para prever o número de alunos que poderão abandonar o instituto nos próximos 6 anos, e foi criado em *Python* um gráfico para descobrir qual o motivo de abandono no instituto tem um maior nível de impacto institucionalmente.

3.5.4. Considerações éticas e proteção de dados

Os dados utilizados são de natureza administrativa e foram disponibilizados pelo *IPSantarém* para fins de análise institucional. O tratamento foi realizado com foco em finalidade específica, minimização de dados e análise em formato agregado, não sendo produzidas inferências ao nível individual nem divulgada informação identificável. Sempre que aplicável, foram adotadas práticas de pseudonimização/ anonimização e respeitados os princípios do *RGPD*, garantindo confidencialidade e acesso restrito aos dados durante o projeto.

Adicionalmente, os dados sensíveis dos estudantes foram anonimizados nas visualizações apresentadas neste relatório, nomeadamente os nomes próprios, que foram substituídos de forma a apenas ser apresentada a primeira letra do primeiro e do último nome, impedindo a identificação direta dos indivíduos e reforçando a proteção da sua privacidade.

3.5.5. Justificações Técnicas

As tecnologias e metodologias utilizadas foram escolhidas pela sua eficiência, facilidade de uso e integração já existente no instituto. Tais como:

1. **Power BI:** o *Power BI* é uma ferramenta que permite integrar múltiplas fontes de dados, possui capacidades de visualização interativa, é fácil de usar e fácil de partilhar, permitindo construir relatórios dinâmicos e intuitivos. Esta ferramenta foi escolhida pelo instituto para realizar o projeto;
2. **Python:** a linguagem *Python* foi utilizada para criar gráficos personalizados no *Power BI*, utilizando *IA* para implementar análises e modelos de previsão exploratórios, recorrendo a bibliotecas como *pandas*, *matplotlib*, *seaborn*, entre outros;

3. **Modelo ARIMA:** a escolha do modelo *ARIMA* para a previsão baseou-se na natureza sequencial dos dados e na sua capacidade de encontrar tendências e padrões.
4. **Visualizações:** as visualizações utilizadas para apresentar os dados nos *dashboards* foram escolhidas por serem de fácil interpretação e leitura. Embora existam outras visualizações que também se ajustariam bem aos dados, optou-se por formatos que garantissem clareza, consistência visual e coerência com o objetivo exploratório do projeto.

4. Resultados

4.1. Introdução

O presente capítulo apresenta e analisa os resultados obtidos a partir dos *dashboards* e visualizações desenvolvidos em *Power BI*, permitindo uma compreensão estruturada e aprofundada do fenómeno do abandono escolar no *IPSantarém*. O objetivo é apresentar evidência empírica que responda às questões de investigação, mostrando de que forma os dados disponíveis revelam padrões, tendências temporais e fatores associados ao abandono. A análise dos resultados foi organizada segundo uma lógica progressiva de níveis, iniciando-se numa perspetiva global e avançando gradualmente para dimensões mais específicas: dados gerais do abandono, características demográficas, indicadores académicos, evolução temporal, análises complementares e, por fim, resultados derivados de abordagens analíticas exploratórias. Esta estrutura permite uma leitura coerente e cumulativa, facilitando a interpretação dos dados e a identificação dos aspetos mais relevantes para apoiar decisões institucionais e orientar estratégias futuras.

4.2. Resultados Globais do Abandono Escolar

4.2.1. Indicadores Gerais

A análise por indicadores globais permite caracterizar, de forma agregada, o abandono escolar no *IPSantarém*. A Figura 1 apresenta o universo de estudantes inscritos e o número de anulações de matrícula/inscrição registadas no período analisado, bem como a respetiva percentagem. Estes indicadores constituem a base de referência para a leitura das análises subsequentes, permitindo enquadrar a dimensão do fenómeno no conjunto da população estudantil.

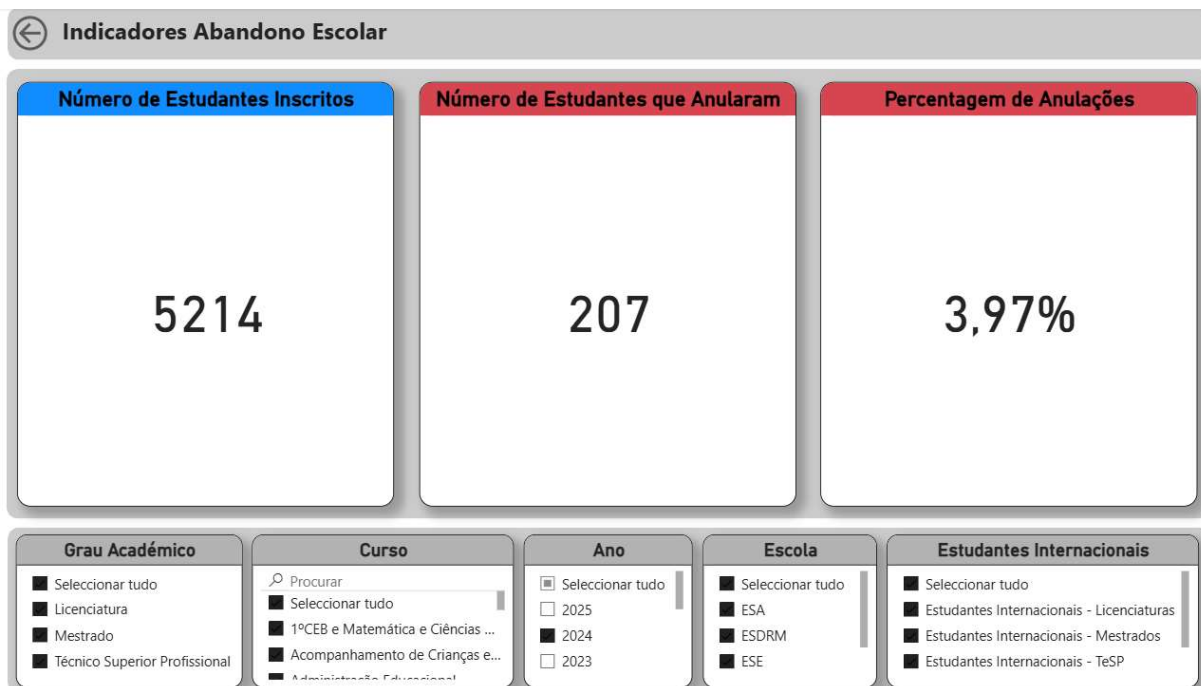


Figura 1 - Indicadores Abandono Escolar

A análise segmentada pelo número total de estudantes e total de cursos por grau académico, apresentado na figura 2, permite estabelecer a dimensão do instituto, servindo de referência para os cálculos subsequentes.

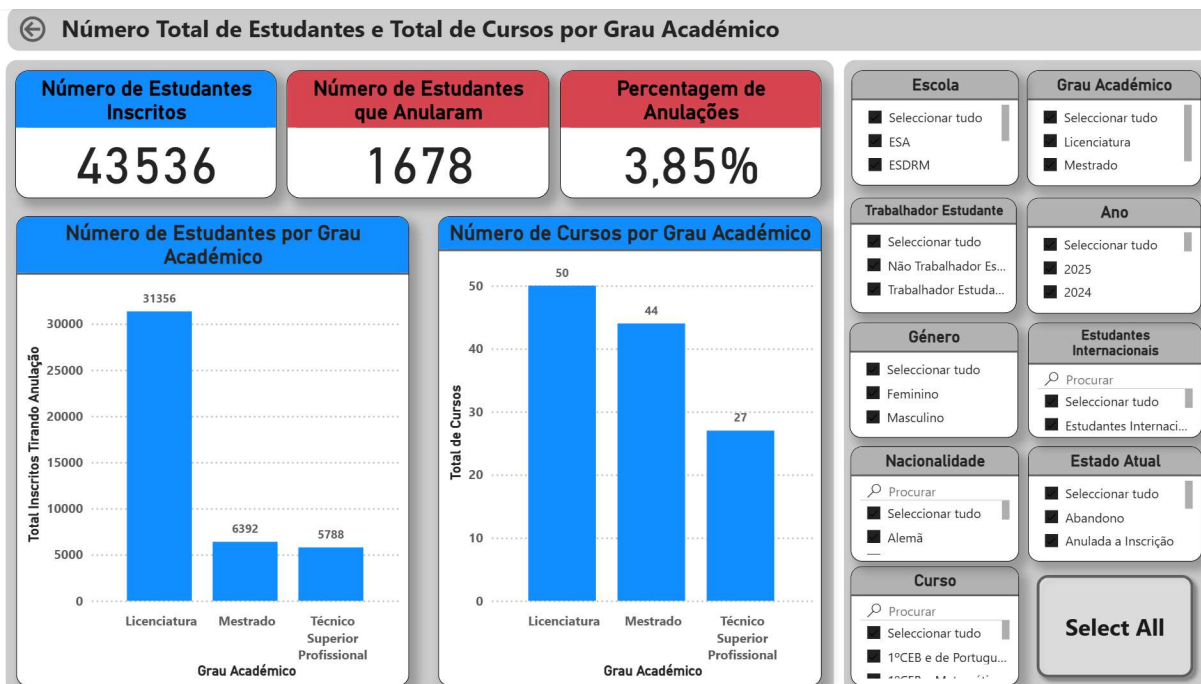


Figura 2 - Número Total de Estudantes e Cursos por Grau Académico

4.2.2. Perfil Agregado do Abandono

A análise segmentada por perfis globais de abandono permite identificar padrões transversais que caracterizam os estudantes que anulam a matrícula no *IPSantarém*. Na figura 3, emergem grupos mais frequentes associados ao abandono, destacando tendências comuns no comportamento académico. A figura 3 evidencia ainda as combinações de variáveis que surgem de forma recorrente, género, idade, curso ou escola, revelando perfis agregados que contribuem para a compreensão do fenómeno numa perspetiva ampla.



Figura 3 - Tabela Abandono Escolar

4.2.3. Relações Gerais entre Variáveis

A análise segmentada pelas relações gerais entre variáveis demonstra como diferentes dimensões académicas e demográficas se articulam no contexto do abandono escolar. Conforme ilustrado na figura 4, as comparações apresentadas no painel Resumo evidenciam padrões de associação entre variáveis-chave, permitindo observar tendências comuns entre grupos específicos de estudantes. A figura 4 evidencia como determinadas combinações de características produzem variações relevantes em termos descritivos nos níveis de abandono, revelando interações que ajudam a contextualizar os resultados globais.

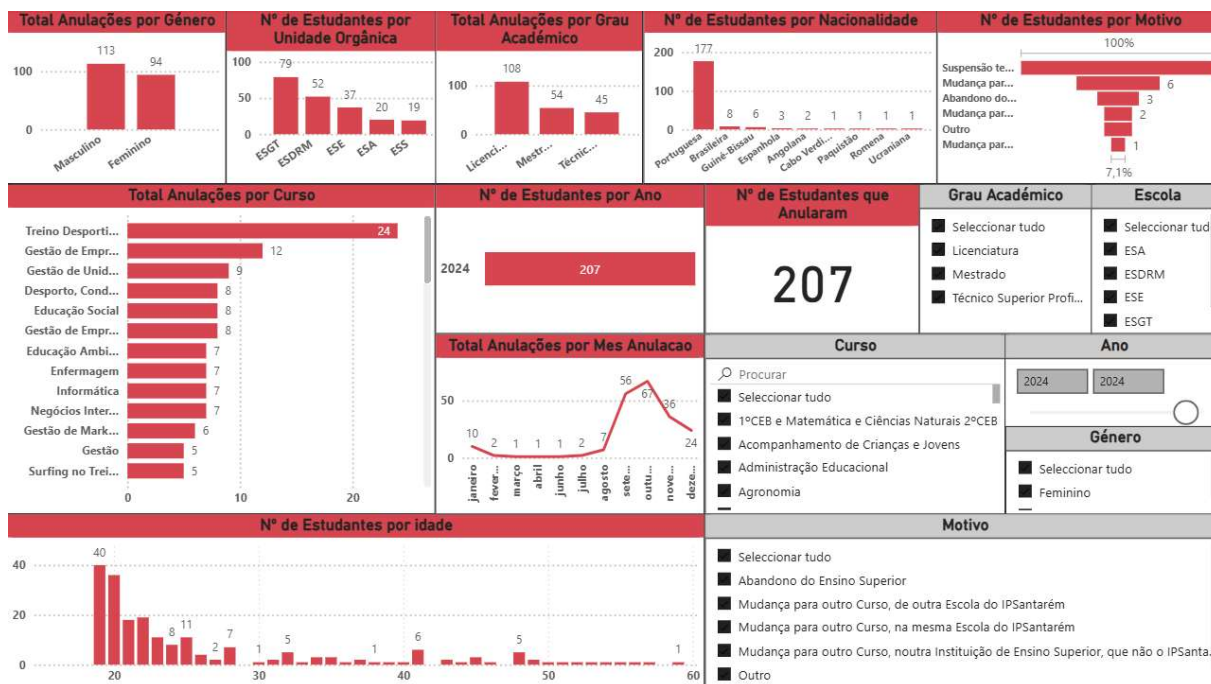


Figura 4 - Resumo Abandono Escolar

4.3. Resultados por Variáveis Demográficas

4.3.1. Idade

A análise segmentada por idade evidencia diferenças observáveis nos padrões de abandono entre faixas etárias. Conforme representado na figura 5, a distribuição por classes etárias revela que determinados grupos concentram maior proporção de anulações. A figura 5 evidencia ainda contrastes entre os diferentes grupos de idade, mostrando variações significativas no comportamento académico, especialmente entre estudantes mais jovens e aqueles com percursos académicos tardios. Estas observações reforçam a importância da idade enquanto fator demográfico associado ao abandono.

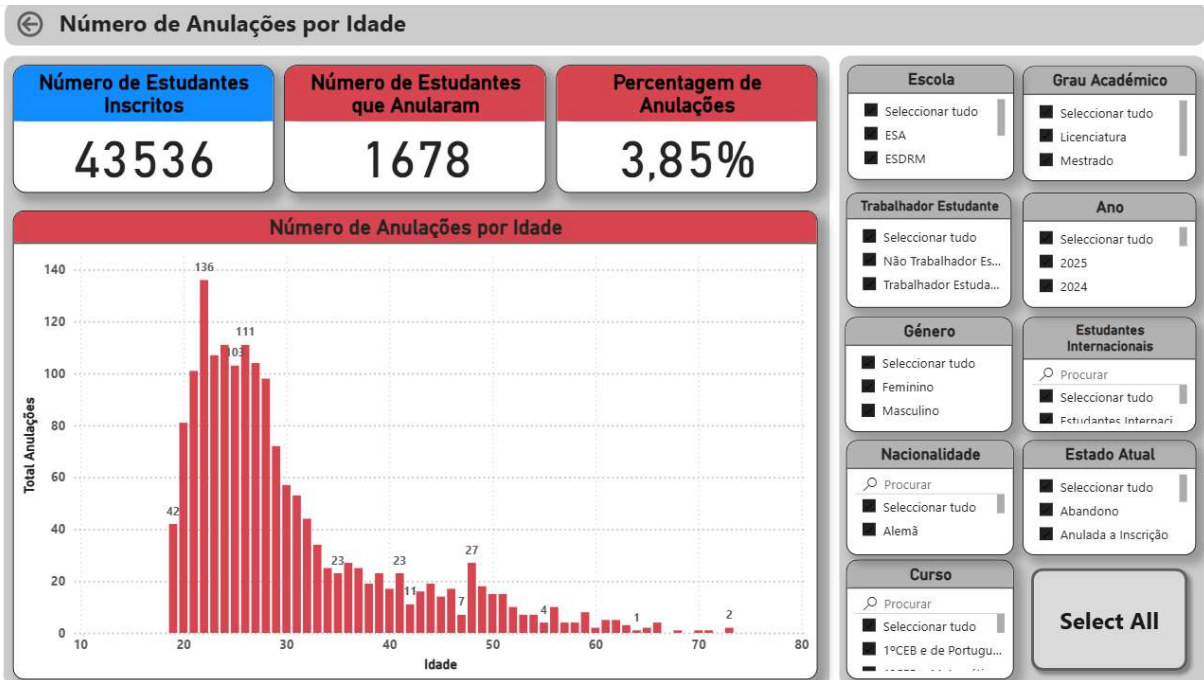


Figura 5 - Abandono por Idade

4.3.2. Género

A análise segmentada por género indica diferenças assinaláveis na participação e no abandono entre estudantes masculinos e femininos. Na figura 6, observa-se a distribuição por género revela padrões distintos na proporção de anulações. Estes resultados contribuem para uma compreensão mais aprofundada do papel do género no abandono escolar.

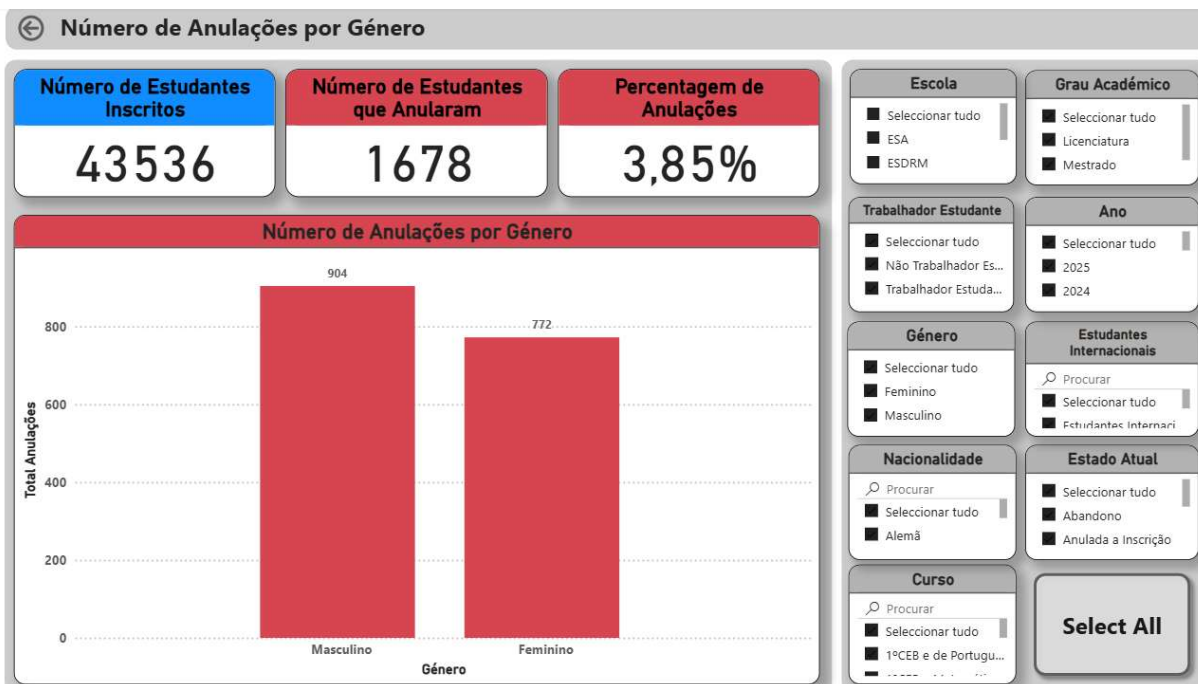


Figura 6 - Abandono por Género

4.3.3. Nacionalidade

A análise segmentada por nacionalidade mostra diferenças observáveis nos níveis de abandono entre estudantes nacionais e internacionais. Na figura 7, a distribuição indica que as anulações por nacionalidade evidenciam a predominância de determinados grupos, permitindo traçar um retrato mais granular do fenómeno.

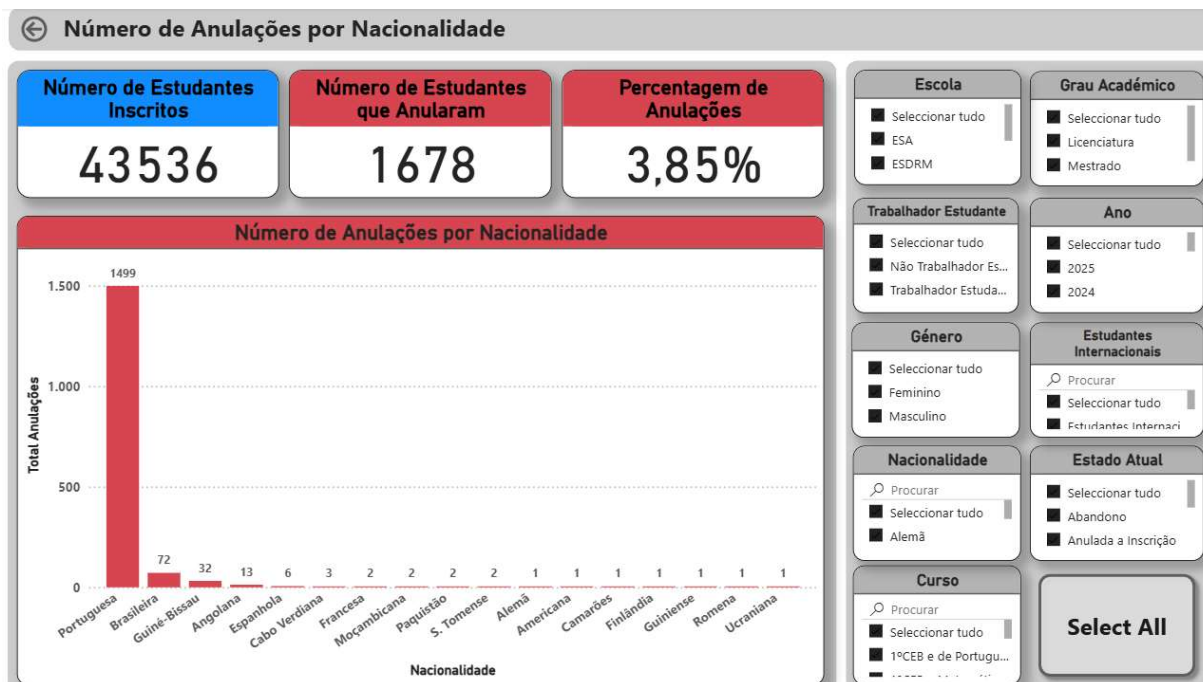


Figura 7 - Abandono por Nacionalidade

A figura 8 evidencia ainda a comparação entre estudantes portugueses e internacionais, revelando variações que se tornam mais marcadas quando observadas em subgrupos específicos como o CPLP. Estas distinções reforçam a relevância do contexto cultural e geográfico na interpretação dos resultados.

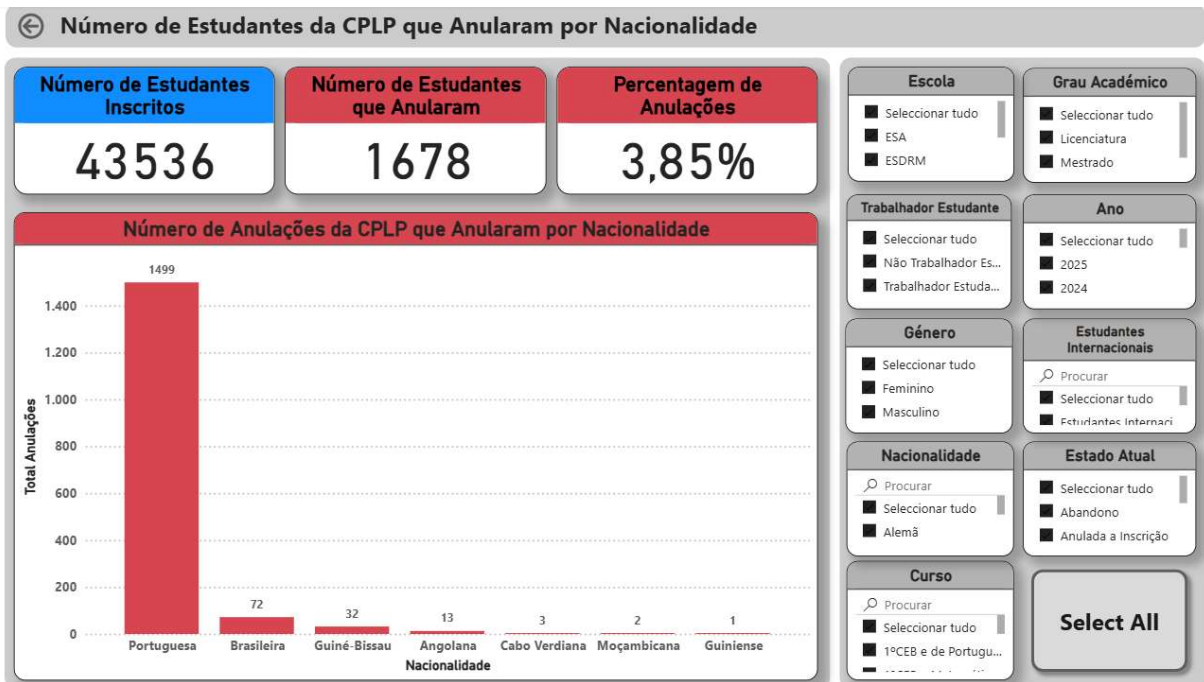


Figura 8 - Abandono por CPLP

4.3.4. Estatuto Académico

A análise segmentada por estatuto académico revela padrões diferenciados entre estudantes com características específicas, como trabalhador-estudante, deslocado ou se possui bolsa de estudo. Verifica-se nas figuras 9, 10 e 11, que estes grupos apresentam níveis de anulação que divergem do perfil global, permitindo identificar vulnerabilidades particulares associadas às suas condições académicas e pessoais. Estas observações reforçam o impacto que fatores externos à vida académica podem ter no abandono escolar.

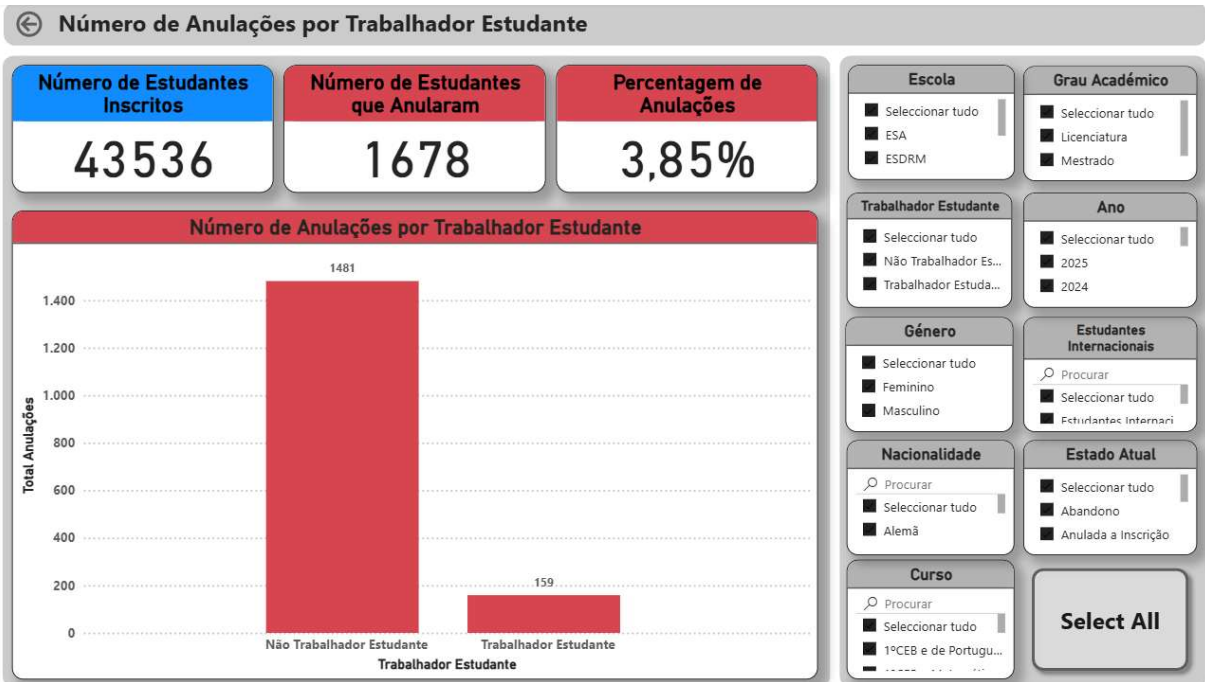


Figura 9 - Abandono por Trabalhador-Estudante

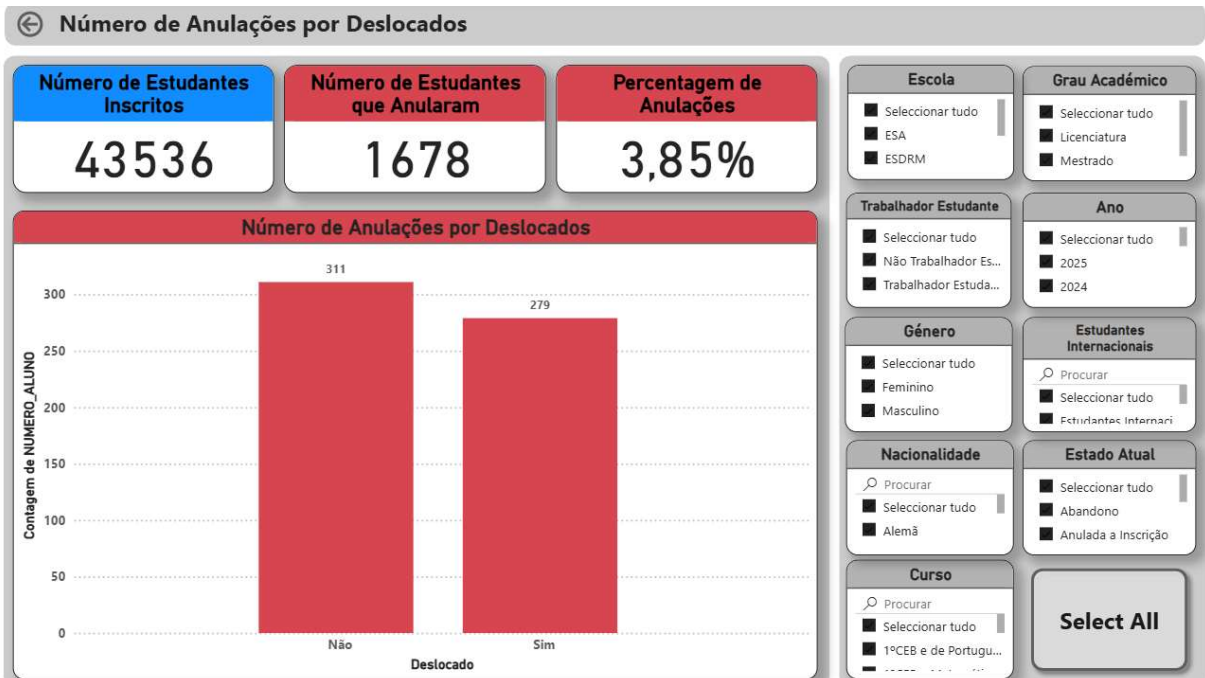


Figura 10 - Abandono por Deslocado

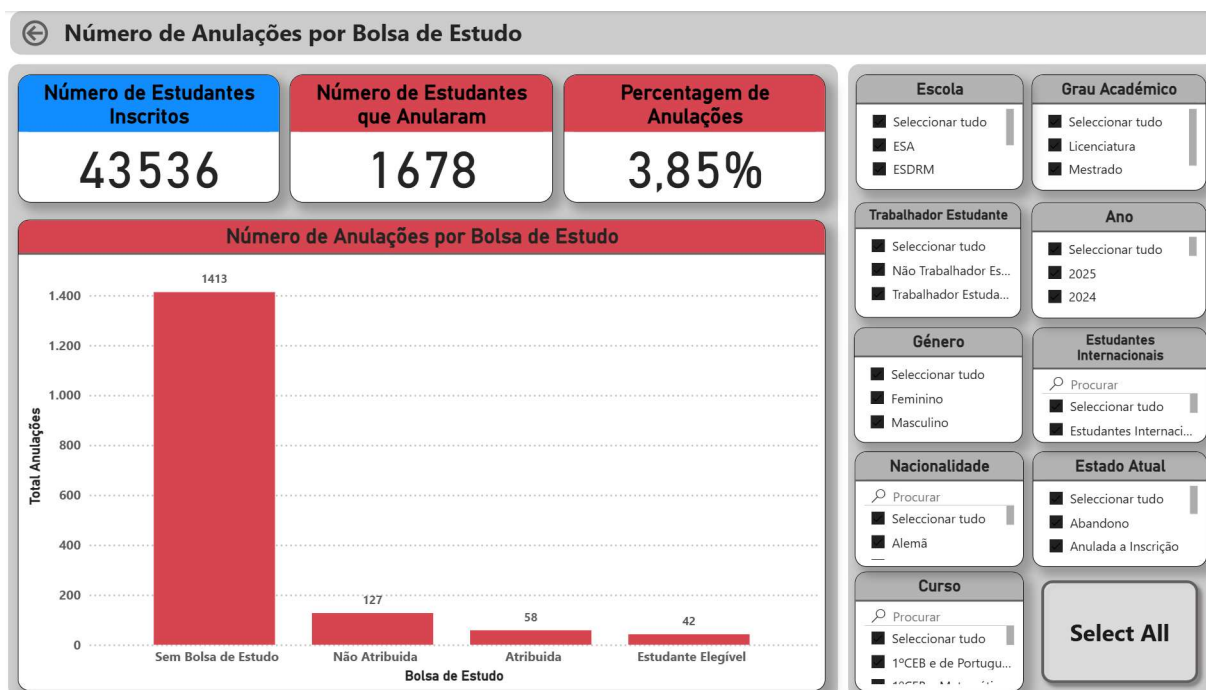


Figura 11 - Abandono por Bolsa de Estudo

4.3.5. Interações Demográfico-Académicas

A análise segmentada pelas interações entre variáveis demográficas e académicas permite observar padrões combinados que não são visíveis nas análises isoladas. Conforme representado na figura 12, cruzamento como “Idade x Escola” mostram variações notáveis na distribuição das anulações entre diferentes instituições do *IPSantarém*.

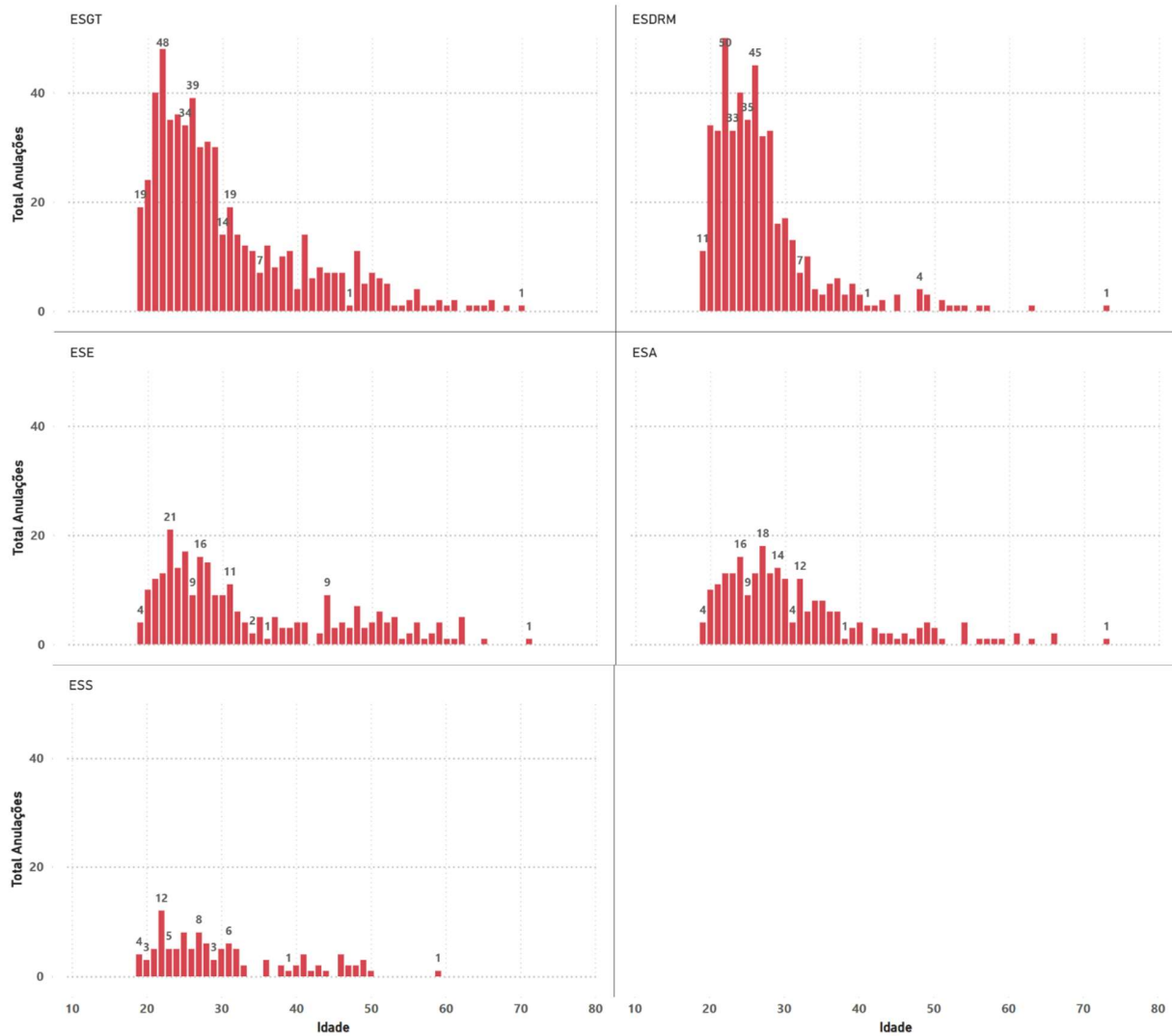


Figura 12 - Abandono por Idade x Escola

De forma semelhante, a interação “Género x Curso” na figura 13 evidencia tendências específicas que realçam diferenças entre áreas científicas.



Figura 13 - Abandono por Género x Curso

A figura 14, relativa à interação “Nacionalidade x Grau Académico”, mostra contrastes na participação e no abandono entre estudantes nacionais e internacionais em diferentes níveis de ensino. Estas interações aprofundam a compreensão das dinâmicas do abandono, revelando que o fenómeno resulta frequentemente da combinação de múltiplos fatores.

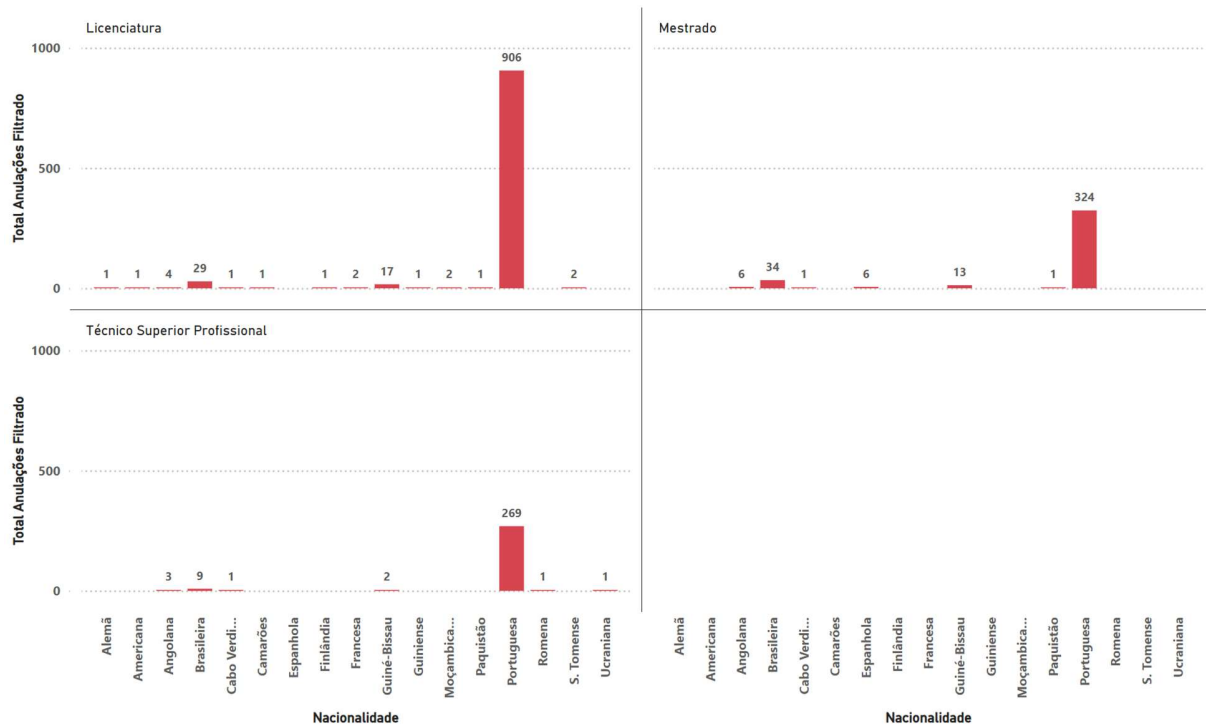


Figura 14 - Abandono por Grau x Nacionalidade

4.4. Resultados por Variáveis Académicas

4.4.1. Escola/Unidade Orgânica

A análise segmentada por escola evidencia diferenças observáveis na distribuição do abandono entre as várias unidades orgânicas do *IPSantarém*. Conforme representado na figura 15, a comparação entre escolas mostra variações relevantes em termos descritivos no número absoluto de anulações.

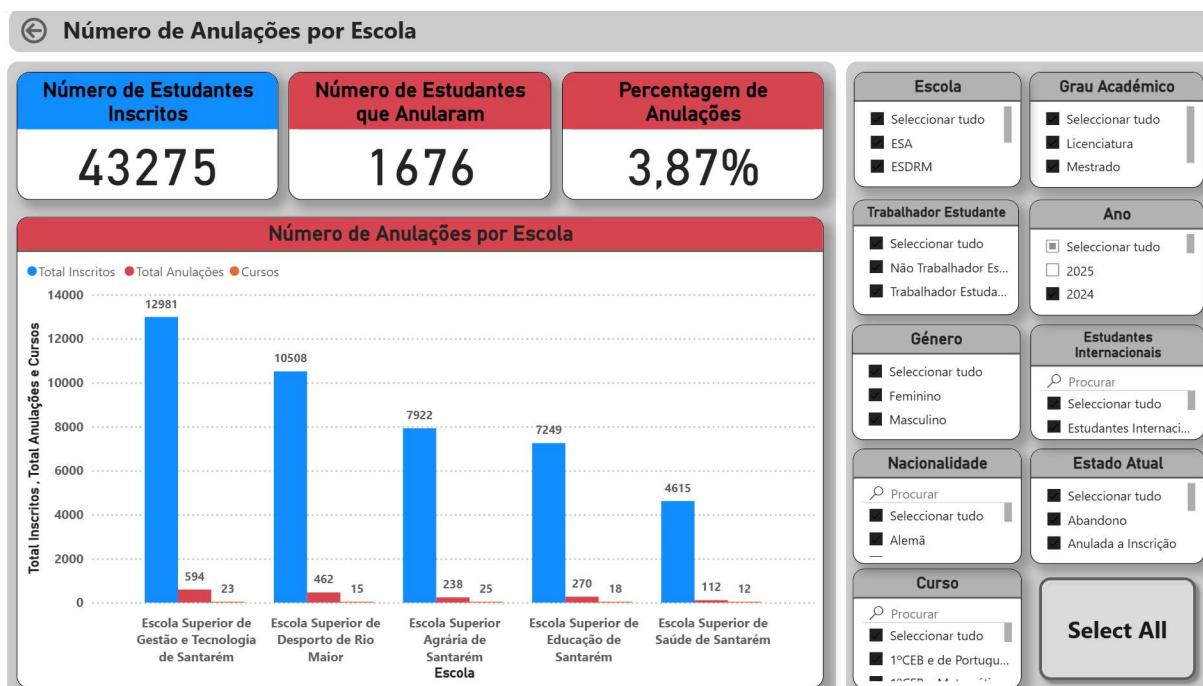


Figura 15 - Abandono por Escola

4.4.2. Grau Académico

A análise segmentada por grau académico revela diferenças observáveis entre licenciaturas, mestrados e técnicos superiores profissionais. Na figura 16 as percentagens de abandono variam de forma consistente entre os diferentes níveis de ensino, evidenciando padrões distintos de retenção e progressão.

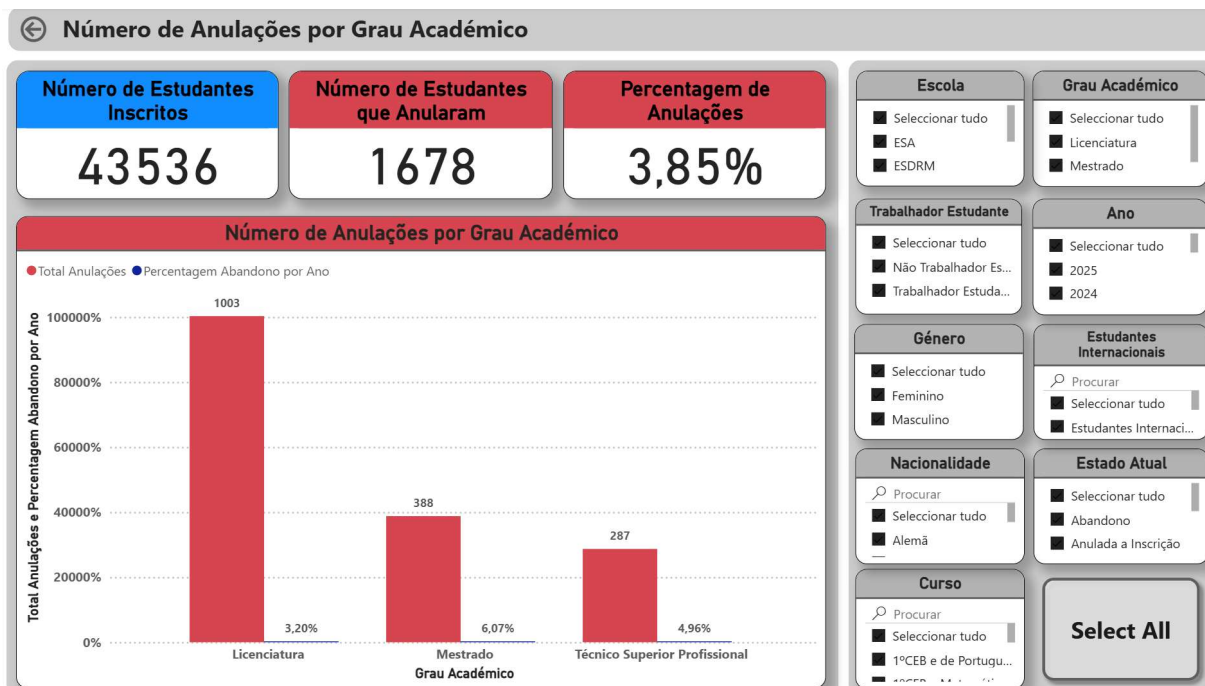


Figura 16 - Abandono por Grau Académico

4.4.3. Curso

A análise segmentada por curso permite identificar programas com níveis de abandono mais elevados, bem como cursos que apresentam valores proporcionalmente mais reduzidos. Conforme representado na figura 17, há padrões que parecem relacionados com a dimensão dos cursos. Estas observações permitem delinear um retrato mais granular das variações existentes entre ofertas formativas.

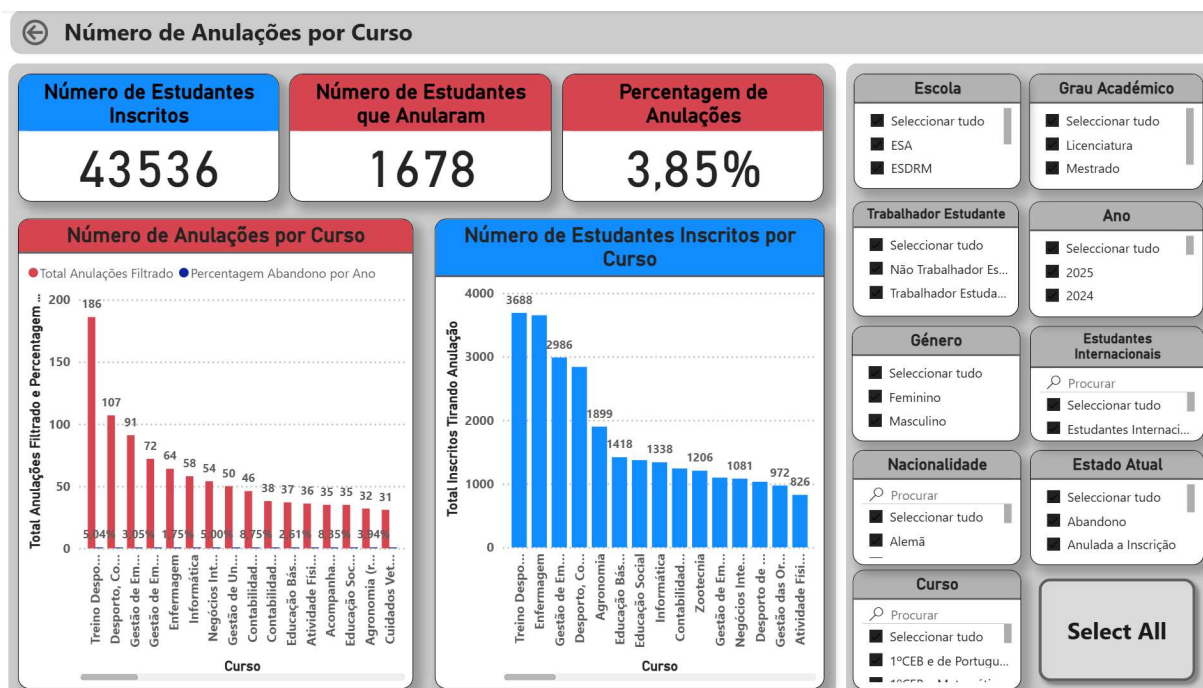


Figura 17 - Abandono e Total Estudantes por Curso

4.4.4. Ano Curricular

A análise segmentada por ano curricular mostra diferenças visíveis no comportamento de abandono ao longo do percurso académico. Conforme representado na figura 18, o ano com maior incidência de anulações tende a concentrar uma proporção baixa do número de inscritos, evidenciando um padrão de abandono precoce.

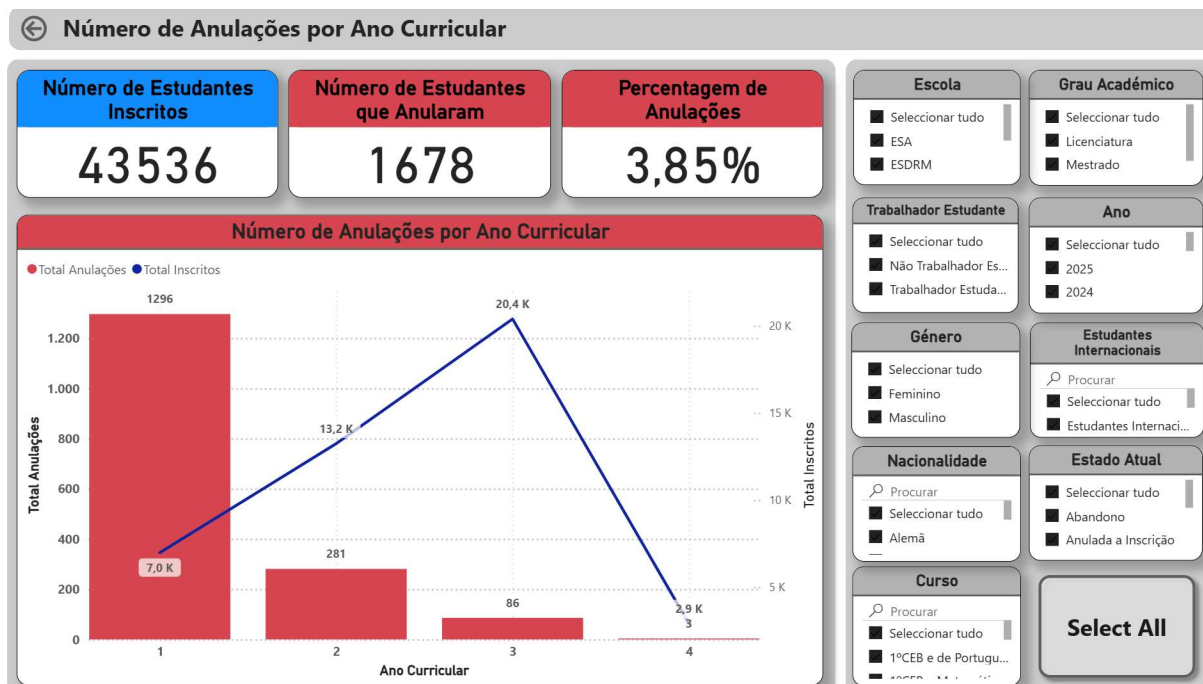


Figura 18 - Abandono por Ano Curricular

4.4.5. Tipo de Anulação

A análise segmentada por tipo de anulação distingue claramente os padrões associados à anulação de matrícula e de inscrição. Observa-se na figura 19 que estas categorias apresentam distribuições distintas, refletindo diferenças nos momentos em que os estudantes interrompem formalmente o processo académico.

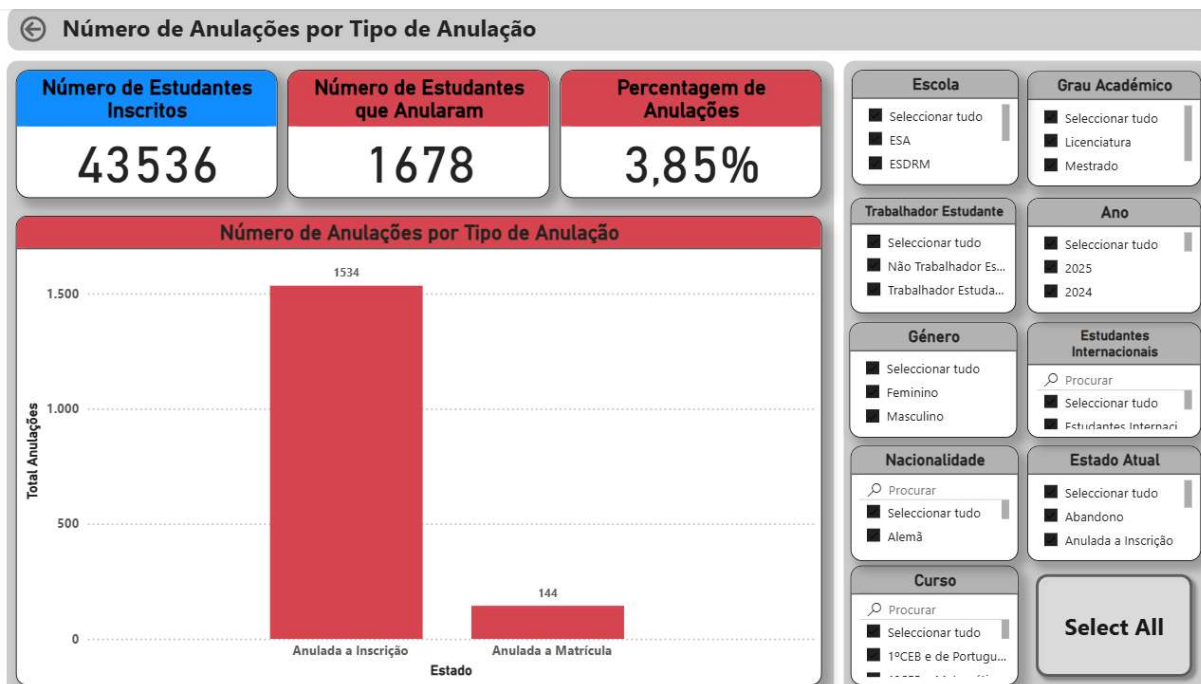


Figura 19 - Abandono por Tipo de Anulação

A figura 20 evidencia ainda a interação entre tipo de anulação e ano, mostrando os padrões que se repetem ao longo do ano.

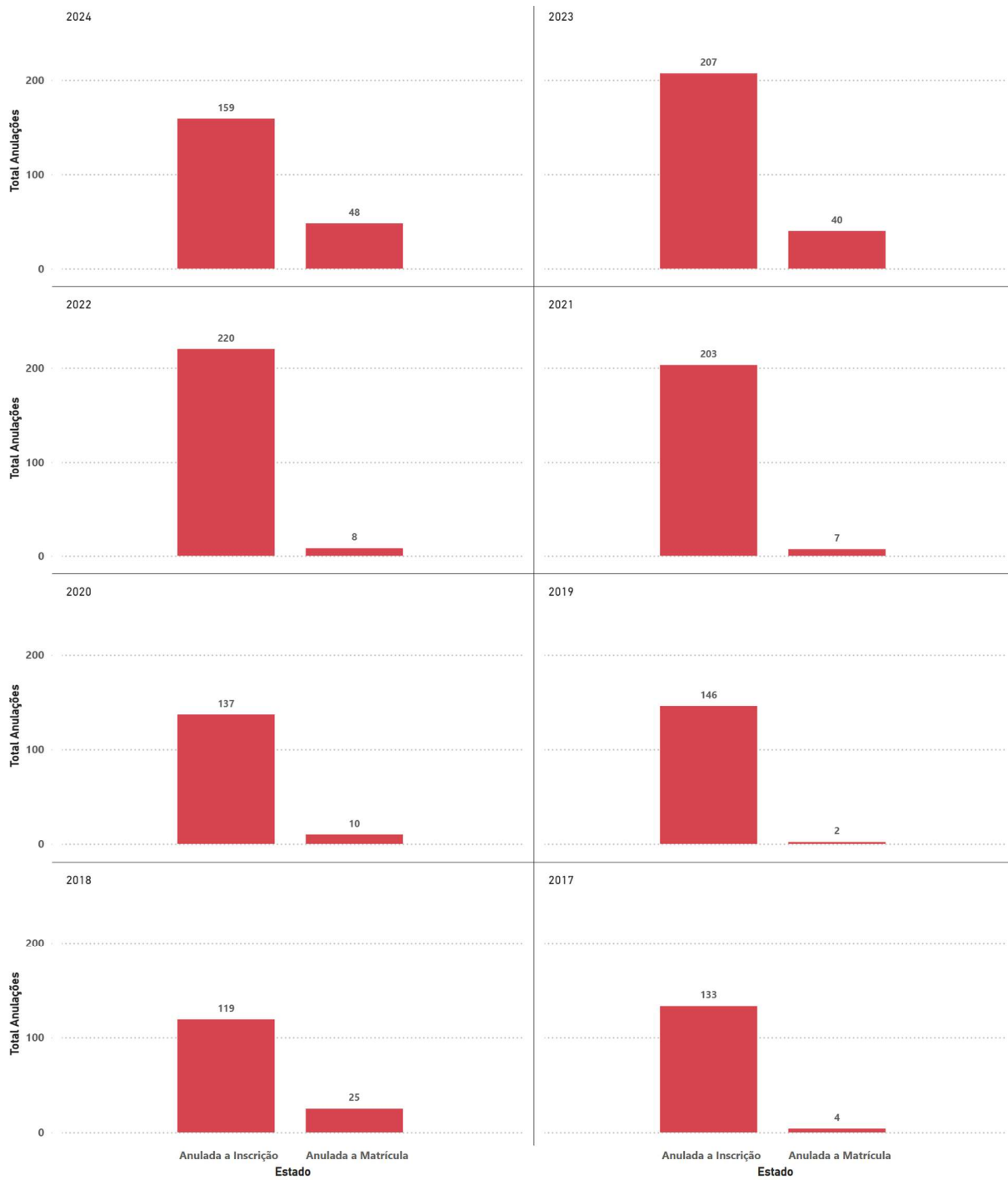


Figura 20 - Abandono por Tipo Anulação x Ano

4.4.6. Motivo de Anulação

A análise segmentada pelos motivos de anulação revela uma limitação significativa relacionada com a reduzida quantidade de estudantes que registam esta informação no sistema académico. Conforme representado na figura 21, apenas uma fração dos alunos que anulam a matrícula ou inscrição indica explicitamente o motivo, o que resulta numa distribuição pouco expressiva. Ainda assim, a figura 21 evidencia que, dentro do conjunto limitado de registos existentes, surgem alguns motivos mais frequentes, embora sem volume suficiente para permitir uma caracterização robusta ou comparações detalhadas entre cursos, escolas ou anos. Os resultados devem, por isso, ser interpretados com cautela, refletindo apenas os padrões observados nos poucos casos em que o motivo foi declarado.

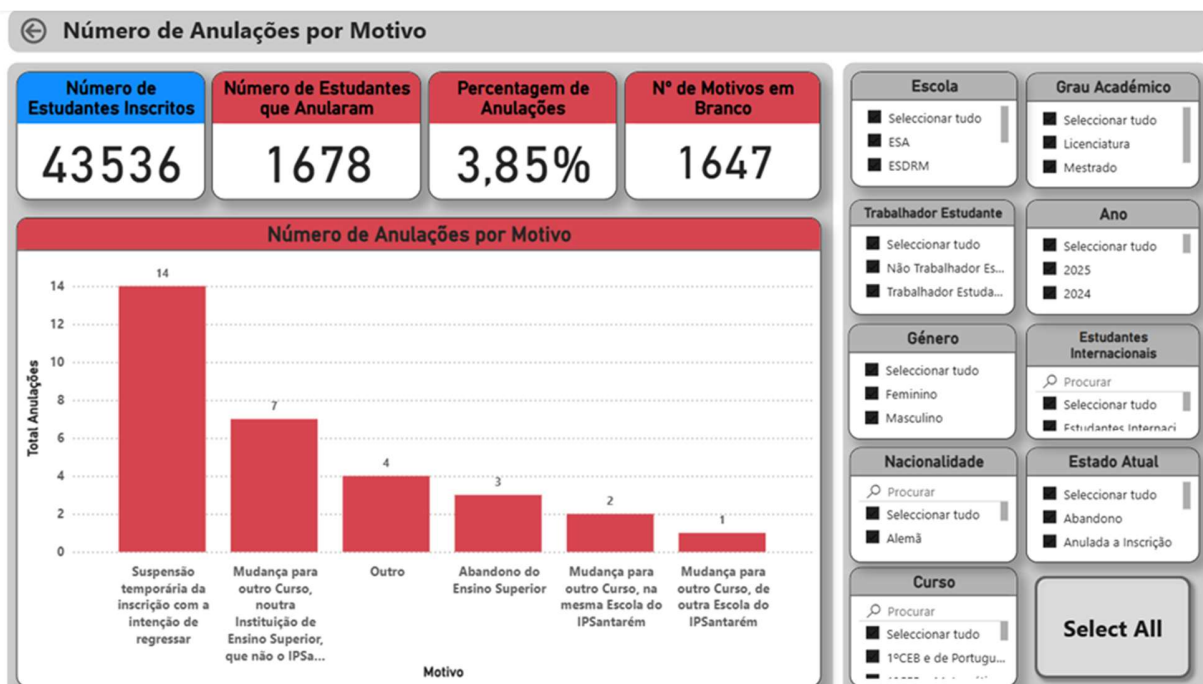


Figura 21 - Abandono por Motivo

4.5. Tendências Temporais

4.5.1. Evolução Anual

A análise segmentada pela evolução anual das anulações entre 2010 e 2024 permite observar tendências gerais ao longo do período estudado. Conforme representado nas figuras 22 e 23, verifica-se uma variação marcada no número total de anulações e percentagem de anulações entre anos, com uma tendência aparente de crescimento ao longo dos anos. Estes padrões anuais constituem um enquadramento importante para contextualizar as análises subsequentes.



Figura 22 - Abandono por Ano

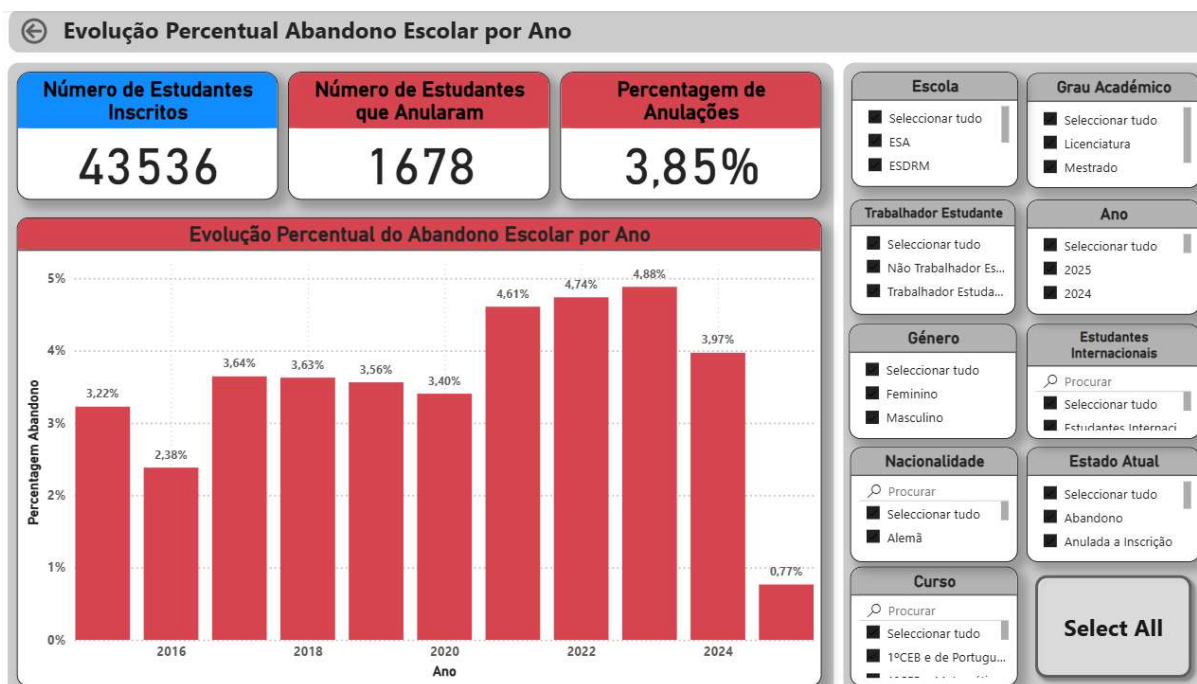


Figura 23 - Abandono Percentual por Ano

4.5.2. Distribuição Mensal

A análise segmentada pela distribuição mensal das anulações evidencia meses com maior incidência deste fenómeno. Conforme ilustrado na figura 24, verifica-se que certos meses concentram uma proporção superior de desistências, enquanto outros revelam volumes mais baixos ao longo do calendário académico.

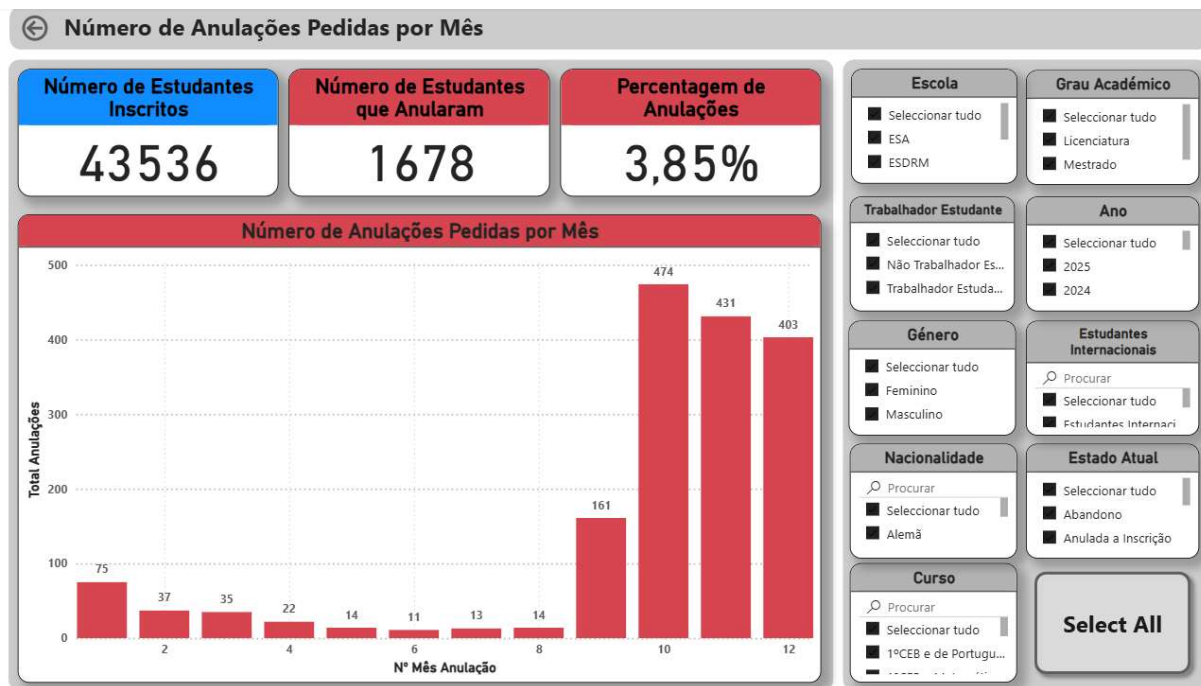


Figura 24 - Abandono por Mês

4.5.3. Diferença entre Pedido e Aceitação

A análise do intervalo temporal entre o pedido de anulação e a sua aceitação administrativa permite identificar dinâmicas relevantes do processo. Conforme mostrado na figura 25, a diferença em meses apresenta variação considerável nos tempos de processamento, refletindo diferenças entre pedidos e respetiva validação administrativa.

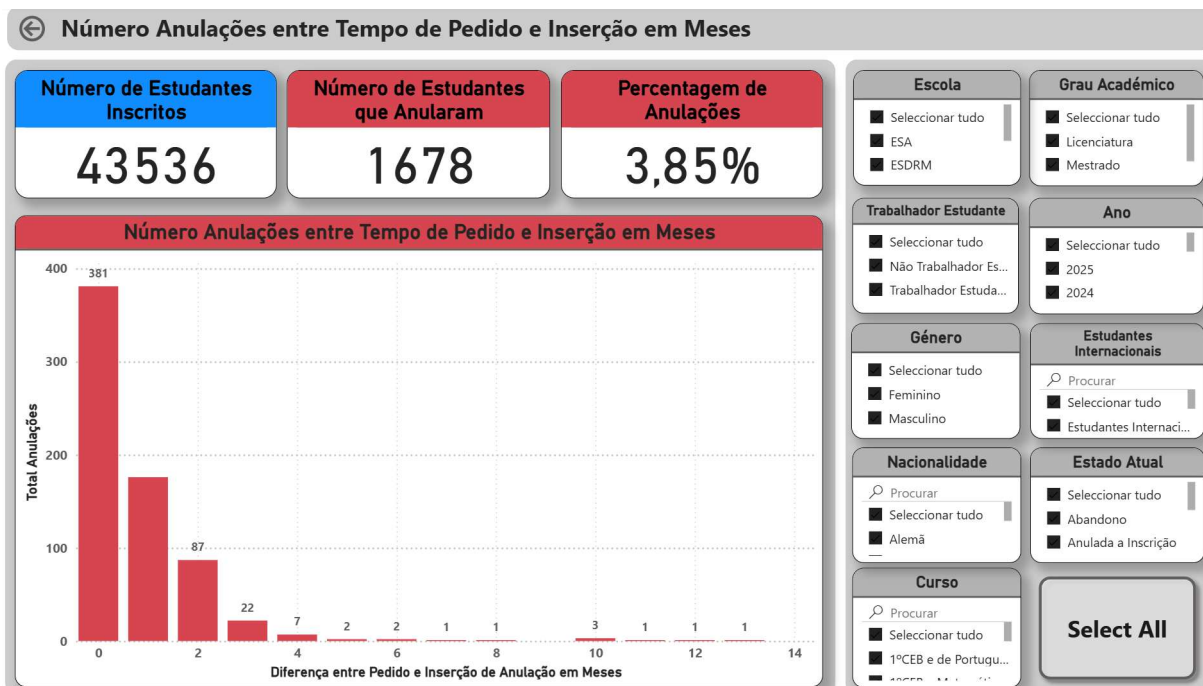


Figura 25 - Diferença entre Pedido e Inserção de Anulação em Meses

A evolução anual desta diferença, representada na figura 26, indica oscilações que permitem observar períodos em que o processamento foi mais célere e outros em que se registaram demoras superiores. Estas variações evidenciam apenas o comportamento administrativo observado, sem interpretação adicional.

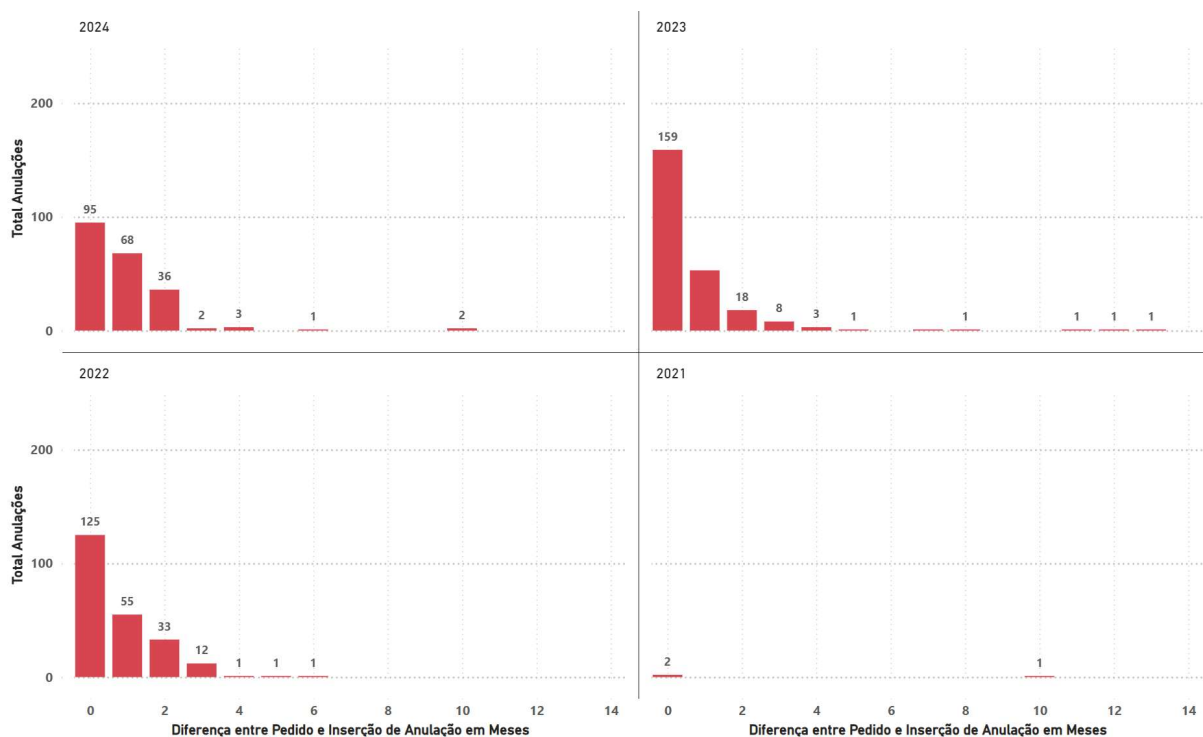


Figura 26 - Abandono por Pedido e Inserção x Ano

4.5.4. Padrões Sazonais

A análise segmentada por padrões sazonais evidencia variações cíclicas ao longo do ano acadêmico. Conforme apresentado na figura 27, observa-se um aumento do número de anulações nos meses iniciais do ano letivo, seguido de uma redução nos períodos subsequentes. A figura 27 destaca ainda meses com maior incidência, nos quais o volume de anulações atinge valores mais elevados de forma recorrente. Estas flutuações, observadas ao longo dos vários anos analisados, revelam padrões temporais consistentes que complementam a interpretação das análises anuais e mensais.

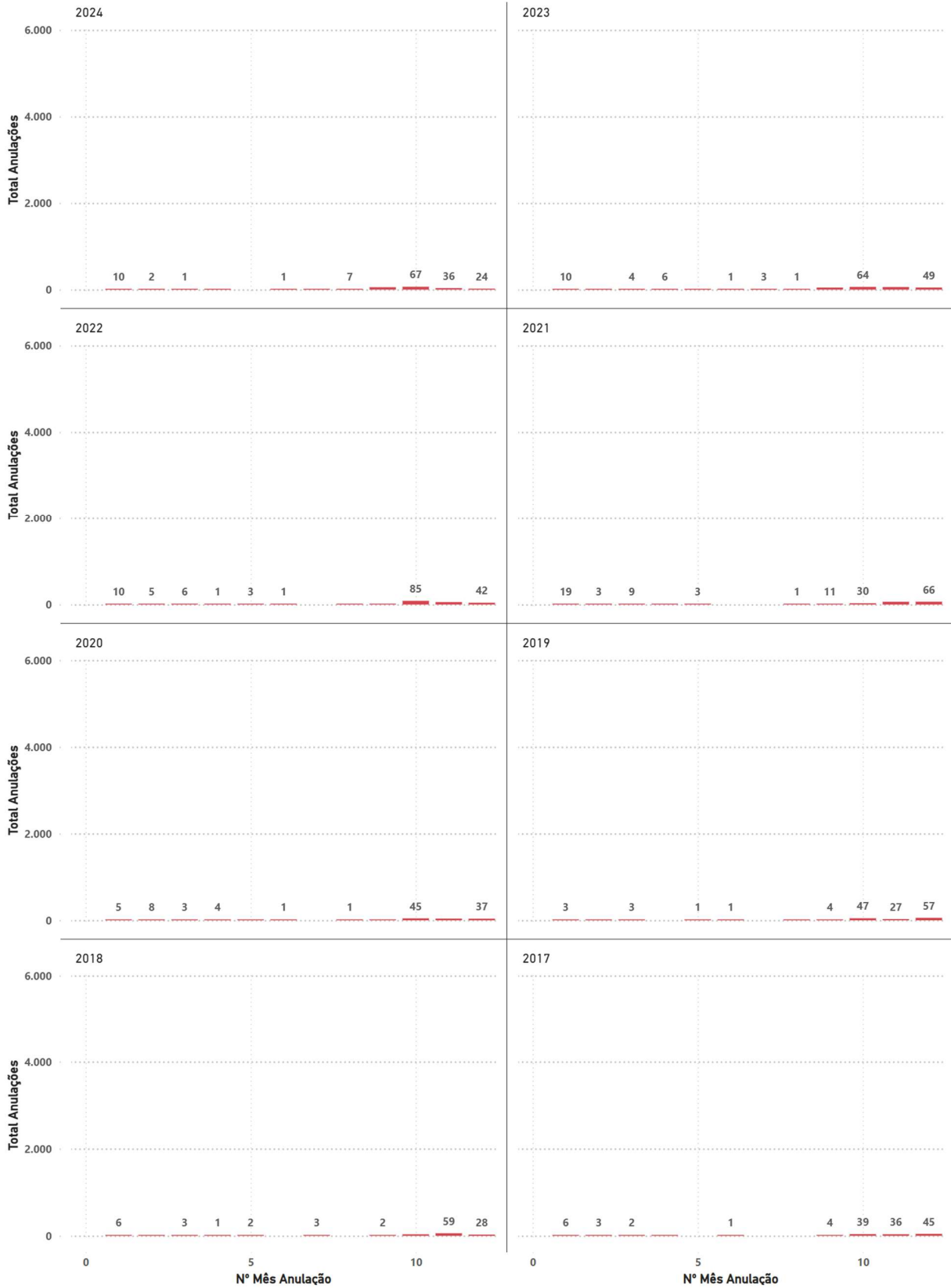


Figura 27 - Abandono por Mês x Ano

4.6. Resultados Complementares e Explorações Específicas

4.6.1. Distribuição Geográfica

A análise segmentada pela distribuição geográfica dos estudantes evidencia padrões diferenciados entre os distritos e concelhos de origem. Na figura 28, observa-se que determinados distritos concentram o maior número de estudantes que anularam, enquanto outros apresentam valores significativamente inferiores. A distribuição por concelhos, ilustrada na figura 28, mostra um padrão semelhante, destacando áreas com maior concentração de anulações e outras com presença residual. Estas observações permitem caracterizar a dispersão territorial dos estudantes envolvidos no fenómeno.

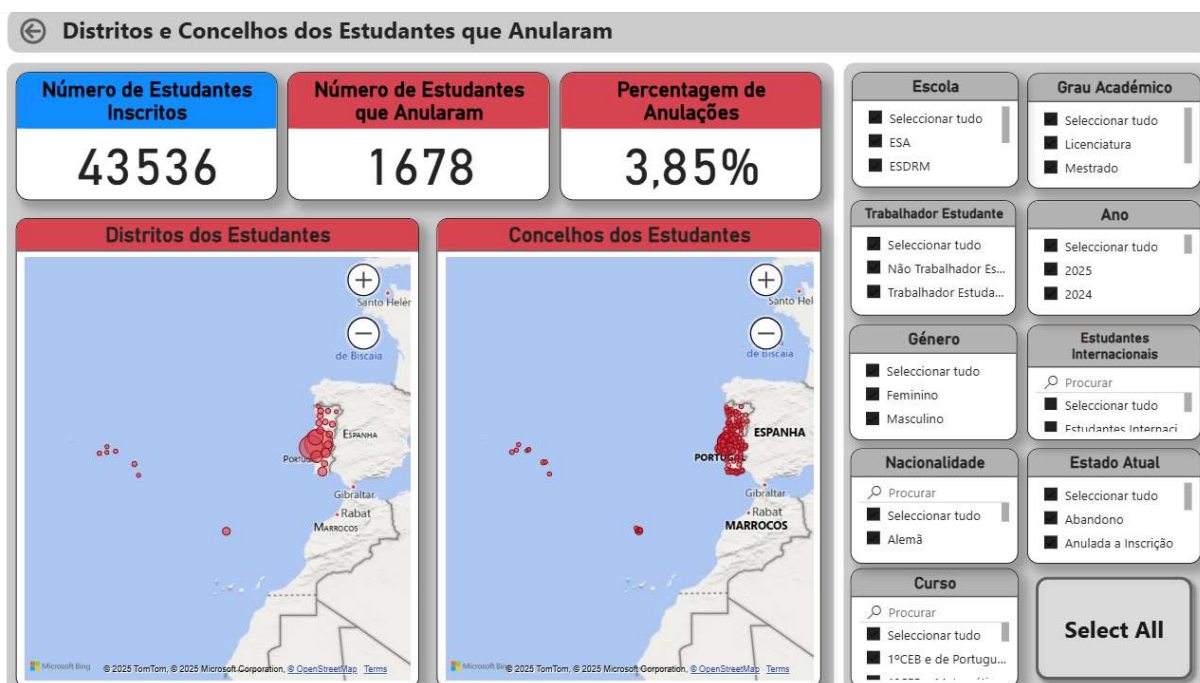


Figura 28 - Abandono por Concelho e Distrito

A figura 29 ilustra os 5 distritos e concelhos onde houve mais abandono no instituto.

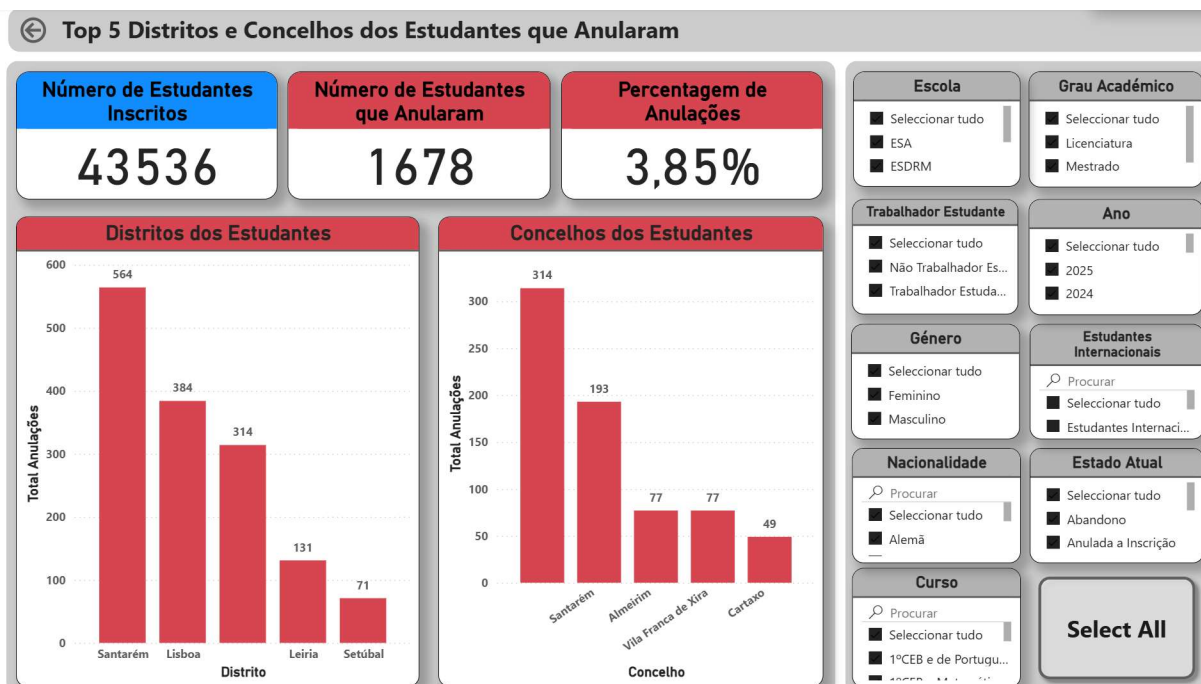


Figura 29 - Abandono por Top 5 Distritos e Concelhos

4.6.2. Estudantes Não Inscritos no Ano Seguinte

A análise segmentada pelos estudantes que não se inscrevem no ano letivo seguinte evidencia a percentagem global deste grupo, conforme representado na figura 30. A figura 30 permite observar individualmente os estudantes identificados como não inscritos, abrangendo variáveis como idade, ano curricular, género, nacionalidade, escola, curso e ano letivo. Esta distribuição detalhada torna possível caracterizar o perfil agregado dos estudantes que interrompem a continuidade académica, evidenciando padrões visíveis entre diferentes grupos demográficos e académicos. Ainda que

descritiva, esta análise destaca diferenças estruturais na composição do conjunto de não inscritos, sem inferir causas ou justificar comportamentos.

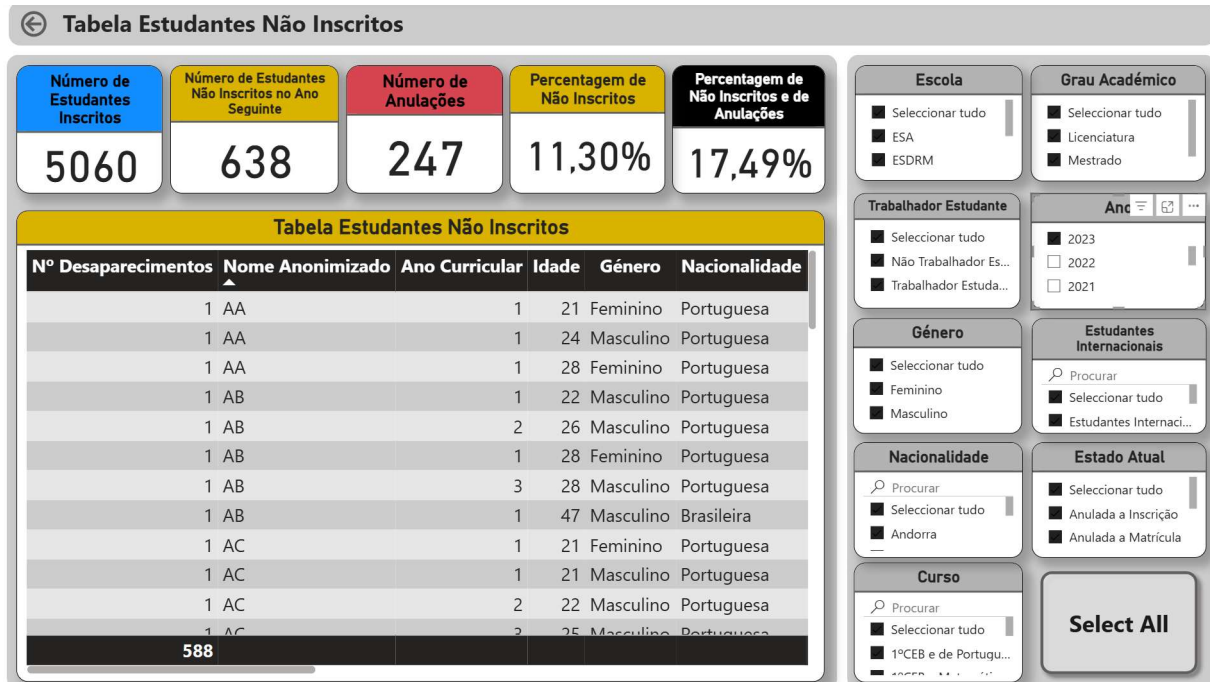


Figura 30 - Abandono por Não Inscrito

4.7. Modelos Analíticos e Preditivos

4.7.1. Modelo ARIMA – Tendência Agregada

A análise temporal das anulações foi complementada com a construção de um modelo ARIMA e as bibliotecas *seaborn*, *pandas* e *matplotlib* aplicado à série histórica agregada, apresentado na Figura 31. Após preparação e agregação dos dados, o modelo foi utilizado com finalidade exploratória, visando descrever o comportamento histórico e projetar a tendência global do número de anulações nos anos seguintes. A biblioteca *seaborn* foi utilizada para facilitar a visualização exploratória e interpretativa dos dados, especialmente em análises temporais, a biblioteca *pandas* foi utilizada para a

preparação, transformação e estruturação dos dados e a biblioteca *matplotlib* foi utilizada como base para a construção e personalização do gráfico. Os resultados devem ser interpretados como apoio indicativo à decisão, dependente da estabilidade dos padrões observados no período analisado.

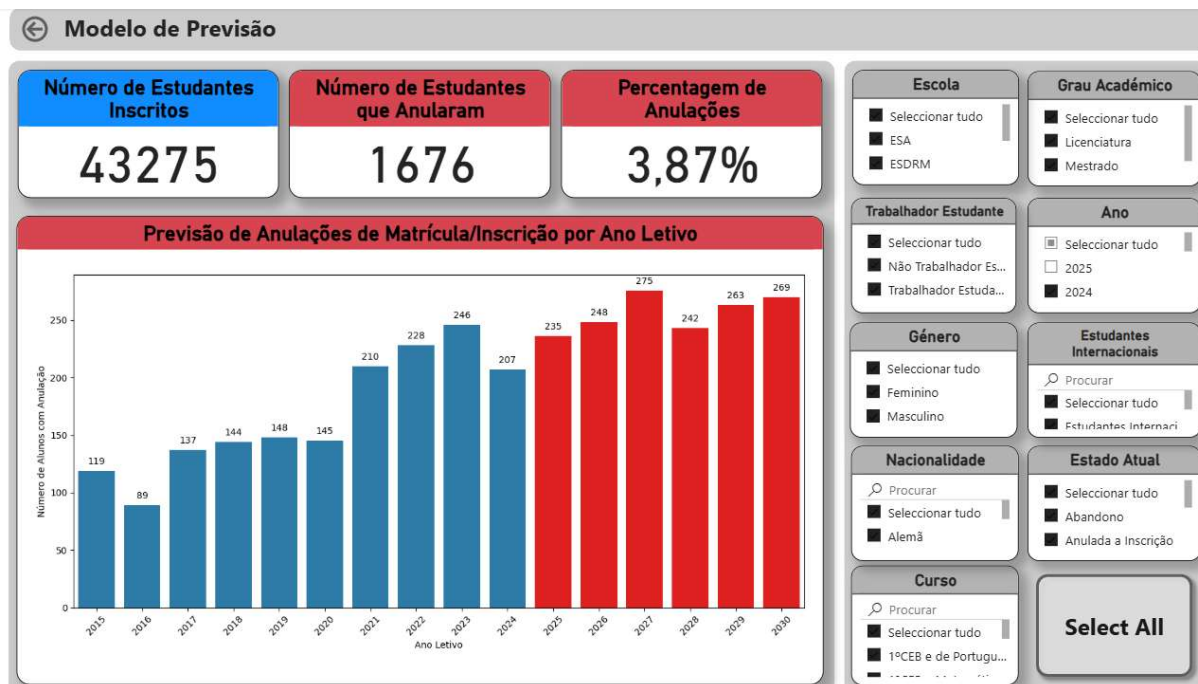


Figura 31 - Abandono por Ano (Modelo de Previsão)

4.7.2. Impacto Institucional

A análise de impacto institucional foi complementada com uma visualização desenvolvida em *Python*, apresentada na Figura 32, com o objetivo de discutir a relevância dos motivos de anulação enquanto suporte à definição de políticas institucionais. Utilizando as bibliotecas *pandas* e *matplotlib*, a biblioteca *pandas* para a preparação, transformação e estruturação dos dados e a biblioteca *matplotlib* foi utilizada para a construção e personalização do gráfico. Verifica-se, contudo, uma limitação relevante: embora exista o campo “motivo de anulação”, este não é preenchido na

maioria dos casos, restringindo a capacidade de caracterizar causas de forma robusta e de desenhar intervenções preventivas mais direcionadas.

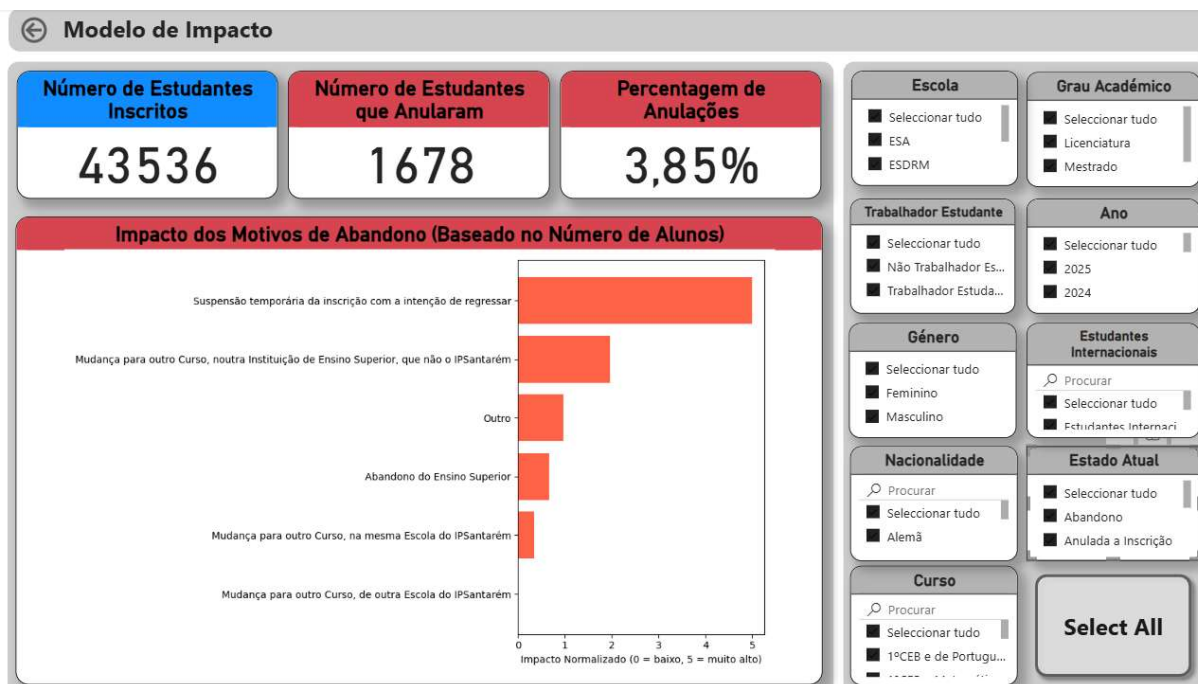


Figura 32 - Abandono por Motivos (Modelo de Impacto)

4.8. Síntese Final dos Resultados

A análise global realizada ao fenómeno de abandono no *IPSantarém* revela um conjunto de padrões consistentes que emergem ao longo das dimensões demográficas, académicas, temporais e complementares exploradas nos *dashboards*.

Em síntese, os resultados indicam padrões regulares e observáveis ao longo de variáveis demográficas, académicas e temporais, complementados por análises específicas e por uma componente preditiva exploratória. Esta evidência fornece uma visão estruturada do fenómeno no *IPSantarém* e serve de base ao capítulo seguinte, onde os resultados

são interpretados à luz da literatura e discutidas implicações institucionais e pistas para investigação futura.

5. Discussão

5.1. Interpretação dos Resultados

A análise dos dados permitiu identificar vários padrões relevantes que respondem às questões de investigação formuladas. A evolução anual do abandono escolar, na figura 22, mostra uma tendência de crescimento ao longo da última década, ainda que 2024 apresente uma ligeira redução face aos três anos anteriores. Apesar disso, registou-se o maior número de anulações de matrícula (48 estudantes), revelando dinâmicas distintas entre o abandono total e a suspensão temporária, como representado na figura 18.

Os resultados apresentados na figura 23 também evidenciam que a percentagem anual de abandono tem aumentado de forma consistente, indicando que o crescimento no número de estudantes inscritos não tem acompanhamento proporcionalmente o aumento dos estudantes que abandonam. O modelo preditivo, representado na figura 31, reforçam esta conclusão, ao estimarem um crescimento contínuo do abandono entre 2025 e 2030.

No plano demográfico, apresentado na figura 5, os estudantes entre os 21 e os 28 anos representam o grupo com maior incidência e abandono, correspondendo a mais de metade dos casos registados. A análise por idade mostra também que os picos de abandono ocorrem em idades que coincidem com o momento de maior esforço académico ou maior exigência de decisão sobre o percurso profissional. A análise

temporal, na figura 24, destaca ainda que a maioria dos abandonos ocorre entre outubro e dezembro, este padrão pode surgir devido ao instituto não obrigar o estudante a pagar o resto das propinas no caso de este anular a matrícula antes de janeiro, ou porque o estudante teve dificuldades na integração académica, ou o curso/instituto não cumpriu as expectativas do estudante.

No ponto de vista académico, a licenciatura concentra a maioria dos abandonos, o que é coerente com o maior volume de estudantes neste grau académico, apresentado na figura 16. Na figura 17, apesar do curso Treino Desportivo apresentar o maior número de absoluto de desistências e de estudantes inscritos, é a Escola Superior de Gestão e Tecnologia de Santarém (*ESGT*) que regista o maior volume total de abandono, como representado na figura 15, algo esperado pois esta tem mais estudantes inscritos e mais cursos.

Na figura 9, é revelado que os estudantes que não possuem estatuto de trabalhador-estudante abandonam em maior número, indicando que o abandono não está associado exclusivamente à conciliação entre trabalho e estudo. Comparando a estudos feitos no Instituto Politécnico de Viana do Castelo (*IPVC*) ((in)Sucesso e Abandono Académico no INSTITUTO POLITÉCNICO DE VIANA DO CASTELO, 2023), chega-se a uma conclusão semelhante, sendo que a maioria de estudantes que abandona o curso e o instituto não são trabalhadores estudantes.

Na figura 6, verifica-se que o número de anulações é maior no género masculino do que o feminino, algo consistente em outros ensinos como no IPVC ((in)Sucesso e Abandono Académico no INSTITUTO POLITÉCNICO DE VIANA DO CASTELO, 2023) e na Universidade do Minho (Almeida et al., 2008), onde o género masculino foi constantemente o género com maior número de anulações.

Dos resultados obtidos, o mais surpreendente foi o da figura 30, revelando que o valor de estudantes não inscritos (ou seja, estudantes que se inscrevem num curso, mas não o concluem, nem anulam matrícula ou inscrição) no ano 2023, é maior do que o número de estudantes que anularam matrícula e inscrição juntos.

Os resultados obtidos em comparação com estudos prévios sobre o abandono no ensino superior português, são consistentes com os padrões já documentados. No entanto, este estudo contribui com uma visão mais detalhada e contextualizada do *IPSantarém*, ao combinar dados reais com técnicas preditivas e *dashboards* interativos.

5.2. Relação com a Literatura

A análise realizada encontra um enquadramento direto nos modelos teóricos apresentados na revisão da literatura, nomeadamente os modelos de *Tinto*, *Bean*, *Nora & Castañeda* e *Astin*. Embora o conjunto de dados do *IPSantarém* utilizado neste estudo não inclua variáveis subjetivas como integração social, autoeficácia, suporte familiar ou motivos económicos, frequentemente destacados na literatura, os resultados obtidos são consistentes com várias premissas desses modelos.

De acordo com *Tinto* e *Astin* (Casanova J. , 2018), a integração académica é um dos principais determinantes da persistência. No contexto deste estudo, a integração não é medida diretamente, mas a análise dos padrões de abandono por curso, ciclo de estudos e ano sugere a presença de desafios académicos e de adaptação, particularmente evidentes nas taxas superiores de abandono em determinados cursos e nos primeiros anos. O que reforça a ideia de que o desempenho e a adequação entre estudante e curso continuam a ser fatores críticos, mesmo quando não são explicitamente declarados como motivo de saída.

O modelo do Atrito de *Bean* e o modelo integrado de *Nora & Castañeda* (1992) (Casanova J. , 2018) enfatizam dimensões sociocognitivas, emocionais e contextuais. Porém, estas dimensões não estão presentes nos dados administrativos analisados. Esta ausência de motivos subjetivos, financeiros ou sociofamiliares limita a capacidade de replicar os resultados da literatura empírica, que frequentemente identifica fatores económicos, expectativas profissionais e contexto familiar como causas predominantes do abandono. A comparação evidencia, assim, uma diferença metodológica essencial: enquanto os estudos empíricos tradicionais se apoiam em instrumentos qualitativos ou inquéritos, este trabalho baseia-se exclusivamente em dados institucionais.

Contudo apesar destas diferenças, os modelos empíricos e teóricos ajudam a contextualizar os padrões indentificadores nos *dashboards*, por exemplo, a maior incidência de abandono em determinados cursos e faixas etárias pode refletir, ainda que de forma indireta, questões de adaptação, expectativas desalinhadas ou dificuldades académicas, dimensões amplamente referidas por (Da Silva Araújo, 2018; Casanova J. , 2018; Ditutala, 2015). Assim, embora os dados não permitam inferir os motivos subjetivos, os padrões observados correspondem a tendências amplamente documentadas na literatura.

Por fim, a utilização de ferramentas analíticas e preditivas, como o modelo *ARIMA* para previsão de tendência agregada, introduz uma abordagem distinta da literatura tradicional, que raramente utiliza métodos de previsão temporal baseados em dados administrativos. Assim, este estudo contribui metodologicamente para o campo, demonstrado como *BI* e *IA* podem complementar modelos teóricos e empíricos, permitindo monitorizar o fenómeno em tempo real e antecipar evoluções futuras.

5.3. Implicações Práticas e Teóricas

Em termos práticos, os resultados permitem propor melhorias concretas ao nível das políticas institucionais do *IPSantarém*. A identificação dos cursos com maior risco fornece uma base objetiva para intervenções específicas, como programas de tutoria, reforço de apoio pedagógico ou revisão curricular. A análise da distribuição por idade e género sugere igualmente a necessidade de estratégias diferenciadas para públicos distintos, incluindo estudantes trabalhadores, deslocados ou mais velhos.

A previsão agregada para os próximos anos reforça a urgência de medidas institucionais de prevenção, dado que as tendências mostram probabilidade de aumento do abandono se nada for alterado. Estas projeções podem apoiar a tomada de decisão na gestão de recursos, planeamento de vagas, estratégias de retenção e alocação de apoio psicológico e social.

Do ponto de vista teórico, este estudo evidencia como a articulação entre análise ***data-driven*** e modelos clássicos de abandono pode enriquecer a compreensão do fenómeno. Embora modelos de previsão ao nível individual sejam frequentemente discutidos na literatura, **este trabalho privilegia uma abordagem agregada**, mais adequada aos dados administrativos disponíveis e particularmente útil para a tomada de decisão organizacional.

5.4. Limitações do Estudo

Apesar da robustez metodológica, o estudo apresenta algumas limitações. Muitos campos da base de dados, como motivos e deslocados, não são preenchidos pela grande maioria de estudantes que abandona o ensino, afetando a análise detalhada.

A maior limitação neste estudo foi o campo motivos, além dos estudantes quando anulam a matrícula ou inscrição não preencherem o campo, as escolhas dentro do campo não ajudam a chegar a uma solução do motivo do estudante ter abandonado o ensino, pois são poucas escolhas e muito genéricas.

Do ponto de preditivo, o modelo *ARIMA* funciona apenas para séries temporais agregadas, não sendo adequado para prever abandonos individuais, e depende da estabilidade histórica dos dados, que pode ser alterada por eventos inesperados, como crises económicas ou mudanças legislativas.

Outra limitação diz respeito à ausência de variáveis socioeconómicas diretas, como rendimento familiar, que poderiam melhorar significativamente a capacidade explicativa dos modelos.

5.5. Sugestões para Trabalhos Futuros

Para trabalhos futuros, recomenda-se a inclusão de variáveis socioeconómicas e psicossociais para enriquecer os modelos preditivos e compreender melhor as causas do abandono. Aplicar técnicas de *Machine Learning* supervisionado poderia fornecer previsões individuais mais robustas, algo que ultrapassa o escopo do modelo *ARIMA*.

Recomenda-se a reformulação do campo ‘motivo de anulação’, com categorias mais específicas e informativas.

5.6. Conclusão da Discussão

Em síntese, este capítulo demonstrou que os resultados obtidos não apenas respondem às questões de investigação, como também contribuem de forma relevante para a compreensão do abandono escolar no *IPSantarém*. A articulação entre dados administrativos, visualizações intuitivas, modelos analíticos e previsão temporal permitiu construir um diagnóstico rigoroso e aprofundado.

A discussão estabeleceu ainda as implicações práticas e teóricas do estudo, refletiu sobre as suas limitações e delineou caminhos promissores para investigação futura. Esta síntese final serve de ponte para o capítulo seguinte, dedicado às conclusões gerais e recomendações institucionais.

6. Conclusão e Recomendações

6.1. Síntese das Conclusões

6.1.1. Cumprimento dos Objetivos

O objetivo geral deste estudo, analisar e identificar padrões do abandono escolar no *IPSantarém*, foi alcançado através da recolha, tratamento e análise sistemática de dados administrativos institucionais, complementada pela utilização de ferramentas de *Business Intelligence* e técnicas analíticas em *Python*. Este processo permitiu construir uma visão estruturada do fenómeno, evidenciando perfis de risco, tendências históricas e dinâmicas associadas ao abandono escolar.

De forma alinhada com os objetivos específicos, o estudo permitiu:

- Caracterizar o abandono por variáveis demográficas, académicas e temporais;
- Identificar padrões e segmentos mais vulneráveis, incluindo cursos, ciclos de estudo e perfis de estudantes com maior incidência de anulação;
- Desenvolver *dashboards* interativos em *Power BI*, facilitando análises exploratórias multidimensionais;
- Explorar abordagens analíticas e preditivas de natureza agregada, recorrendo a técnicas de *IA* para estimar tendências futuras;
- Fornecer indicadores relevantes para a gestão institucional, reforçando a capacidade de monitorização e apoio à decisão.

A baixa taxa de preenchimento do campo “motivo de anulação” constitui uma limitação relevante, condicionando a robustez dessa componente analítica. Ainda assim, os *dashboards* foram preparados de forma a acomodar análises futuras mais completas, caso a instituição venha a recolher esta informação de modo mais sistemático.

6.1.2.Principais Descobertas

A análise realizada permitiu identificar um conjunto de tendências relevantes para compreender o abandono no *IPSantarém*. Destacam-se:

- **Perfis demográficos vulneráveis**, nomeadamente diferenças entre grupos etários, géneros, nacionalidades e estatutos académicos;
- **Padrões académicos significativos**, com variações entre escolas, cursos, graus de ensino e anos curriculares, evidenciando áreas de maior incidência;
- **Tendências temporais marcadas**, tanto ao nível anual como sazonal, reforçadas pelos ciclos típicos de início de ano letivo;
- **Resultados complementares**, como a percentagem de estudantes não inscritos no ano seguinte;

Os *dashboards* foram preparados para permitir análises futuras mais completas, caso a instituição venha a recolher esta informação de forma sistemática.

6.2. Contribuições e Valor do Estudo

Este estudo contribui de forma significativa para o conhecimento institucional ao:

- Introduzir uma abordagem integrada de *BI* e *IA* aplicada aos dados académicos do ensino superior;
- Oferecer ao *IPSantarém* um conjunto de ferramentas visuais e analíticas que permitem monitorizar o abandono escolar com maior precisão;
- Demonstrar a importância do tratamento e qualidade dos dados para análises fiáveis e para a construção de modelos preditivos futuros;

6.3. Recomendações para o IPSantarém

Com base nos resultados alcançados, apresentam-se as seguintes recomendações práticas:

1. Reforçar o preenchimento do motivo de anulação, incentivar os estudantes a preencher o campo de forma completa e fidedigna.
2. Implementar ações de intervenção precoce, direcionadas aos perfis identificados como de maior risco.
3. Aprofundar o acompanhamento dos estudantes que não se reinscrevem, analisando de forma mais detalhada trajetórias académicas interrompidas.
4. Desenvolver mecanismos institucionais de contacto com estudantes que não se reinscrevem, nomeadamente através de inquéritos ou contactos formais, com o objetivo de compreender os motivos da interrupção do percurso académico e apoiar estratégias de retenção
5. Promover a utilização institucional dos *dashboards*, garantindo que escolas, coordenadores e serviços académicos usam estes indicadores para ações estratégicas.

6.4. Sugestões para Futuras Pesquisas

O trabalho abre diversas possibilidades de aprofundamento, entre as quais:

- Desenvolvimento de modelos preditivos individuais de risco, recorrendo a técnicas avançadas de *Machine Learning*;
- Integração de variáveis subjetivas, como satisfação académica, integração social ou motivação;

- Estudos longitudinais sobre estudantes que “desaparecem” do instituto sem formalizar anulação;
- Expansão dos *dashboards* com mais fontes de dados e métricas complementares;
- Investigação comparativa com outras instituições do ensino politécnico ou universitário.

6.5. Considerações Finais

Este projeto demonstra o potencial estratégico da análise de dados na compreensão do abandono escolar no ensino superior. A integração de *Business Intelligence* e Inteligência Artificial permitiu não apenas identificar padrões relevantes, mas também desenvolver ferramentas práticas de apoio à gestão académica e à definição de políticas institucionais. Com a continuidade da recolha de dados e o reforço das capacidades analíticas, o *IPSantarém* dispõe agora de uma base sólida para melhorar a retenção estudantil e reforçar a qualidade do ensino.

Referências

- (in) *Sucesso e Abandono Académico no INSTITUTO POLITÉCNICO DE VIANA DO CASTELO*. (April de 2023). Obtido de ipvc: <https://www.ipvc.pt/wp-content/uploads/2023/11/inSucesso-e-Abandono-Academico-no-IPVC-2023.pdf>
- Almeida, L., & Niemba, A. (4 de March de 2022). *Questionário de motivos de abandono: validação junto de estudantes do 1º ano do ensino superior em Angola*. Obtido de <https://hdl.handle.net/1822/80515>
- Almeida, L., Vasconcelos, R., & Mendes, T. (2008). *O abandono dos estudantes no ensino superior: um estudo na Universidade do Minho*. Obtido de <https://api.semanticscholar.org/CorpusID:145518030>
- Andreu, M. (2008). Los abandonos universitarios: retos ante el espacio Europeo de Educación Superior. *Estudios sobre Educación*, 15, 101-121. <https://doi.org/10.15581/004.15.23447>
- Aranha, A. (2022). *Business Intelligence: desafios e aplicações na educação - Bedu Tech*. Obtido de Bedu Tech - Tecnologia para Educação: <https://bedu.tech/business-intelligence-desafios-e-aplicacoes-na-educacao/>
- Araque, F., Roldán, C., & Salguero, A. (2009). Factors influencing university drop out rates. *Computers & Education*, 53(3), 563-574. <https://doi.org/10.1016/j.compedu.2009.03.013>
- Astin, A. (1984). Student Involvement : A Developmental Theory for Higher Education. *Journal of College Student Development*, 25(4), 297-308.
- Bean, J. (1980). Dropouts and turnover: The synthesis and test of a causal model of student attrition. *Research in Higher Education*, 12(2), 155-187. <https://doi.org/10.1007/bf00976194>

- Bean, J. (2001). The Psychology Underlying Successful Retention Practices. *Journal of College Student Retention Research Theory & Practice*, 3(1), 73-89. <https://doi.org/10.2190/6r55-4b30-28xg-l8u0>
- Bernardo, A., Esteban, M., Fernández, E., Cervero, A., Tuero, E., & Solano, P. (2016). Comparison of personal, social and academic variables related to university drop-out and persistence. *Frontiers in Psychology*, 7, 1610. <https://doi.org/10.3389/fpsyg.2016.01610>
- Cabrera, A., Nora, A., & Castañeda, M. (1992). The role of finances in the persistence process: A structural model. *Research in Higher Education*, 33(5), 571-593.
- Cabrera, L., Bethencourt, J., Pérez, P., & Afonso, M. (2014). El problema del abandono de los estudios universitarios. *RELIEVE - Revista Electrónica de Investigación y Evaluación Educativa*, 12(2). <https://doi.org/10.7203/relieve.12.2.4226>
- Casanova, J. (2018). ABANDONO NO ENSINO SUPERIOR: MODELOS TEÓRICOS, EVIDÊNCIAS EMPÍRICAS e MEDIDAS DE INTERVENÇÃO. *Educação*, 28(57), 05-22. <https://doi.org/10.18675/1981-8106.vol28.n57.p05-22>
- Casanova, J., Cervero, A., Núñez, J., Almeida, L., & Bernardo, A. (2018). Factors that determine the persistence and dropout of university students. *Psicothema*, 4(30), 408-414. <https://doi.org/10.7334/psicothema2018.155>
- Córdova-Esparza, D.-M., Terven, J., Romero-González, J.-A., Córdova-Esparza, K.-E., López-Martínez, R.-E., García-Ramírez, T., & Chaparro-Sánchez, R. (2025). Predicting and Preventing School Dropout with Business Intelligence: Insights from a Systematic Review. *Information*, 16(4), 326. <https://doi.org/10.3390/info16040326>
- Da Silva Araújo, T. (2018). *O abandono escolar no ensino superior: trajetos e projetos: uma análise sociológica*. Obtido de <https://repositorium.sdum.uminho.pt/handle/1822/54835>
- Da Silva Atila, B., & Ecléa, P. (7 de July de 2024). TRANSFORMANDO a EDUCAÇÃO: UTILIZANDO DATA WAREHOUSE e BUSINESS INTELLIGENCE PARA APRIMORAR a TOMADA DE DECISÕES EDUCACIONAIS. <https://doi.org/10.5281/zenodo.10569346>

- De Sousa, A.-R., De Oliveira, C., & Borges, J. (2018). Utilização do sucesso acadêmico para prever o abandono escolar de estudantes do ensino superior: um caso de estudo. *Educação e Pesquisa*, 44. <https://doi.org/10.1590/s1678-4634201844180590>
- Ditutala, D. (2015). *Abandono escolar no ensino superior: estudo de caso do instituto superior politécnico metropolitano de Angola*. Obtido de <https://repositorio.iscte-iul.pt/handle/10071/10697>
- DN/Lusa. (3 de March de 2024). *Abandono está a aumentar no ensino superior público, sobretudo no interior*. Obtido de Diário de Notícias: <https://www.dn.pt/sociedade/abandono-esta-a-aumentar-no-ensino-superior-publico-sobretudo-no-interior>
- Felix, A. (9 de January de 2024). *Business Intelligence na Educação: Como Elevar a Gestão Escolar*. Obtido de Proximo Nivel: <https://www.proximonivel.com.br/post/business-intelligence-na-educacao-como-elevar-a-gestao-escolar>
- Fernando Giannini. (6 de January de 2025). *43 exemplos de inteligência artificial aplicados na educação*. Obtido de Fernando Giannini: <https://fernandogiannini.com.br/43-exemplos-de-inteligencia-artificial-aplicados-na-educacao>
- Ferreira, F. (25 de November de 2024). *Inteligência artificial no ambiente escolar: estratégias para utilizar de maneira saudável*. Obtido de Proesc: <https://www.proesc.com/blog/inteligencia-artificial-no-ambiente-escolar-estrategias-para-utilizar-de-maneira-saudavel>
- Ferreira, M., Cardoso, A., Campos, S., Duarte, J., Guiné, R., & Pereira, A. (2022). Variáveis de contexto pessoal e académico como predictoras do abandono escolar. *Revista de Enfermagem Referência*, 6(1). <https://doi.org/10.12707/rv21134>
- Ficha de Projeto*. (29 de July de 2019). Obtido de IPSantarem: <https://www.ipsantarem.pt/wp-content/uploads/2024/08/3C.pdf>

- Gairín, J., Triado, X., Feixas, M., Figuera, P., Aparicio Chueca, F., & Torrado, M. (2014). Student dropout rates in Catalan universities: profile and motives for disengagement. *Quality in Higher Education*, 20(2), 165-182. <https://doi.org/10.1080/13538322.2014.925230>
- Gil, A. (2014). *El abandono académico: análisis y propuestas paliativas. Dos proyectos de la Universidad Politécnica de Madrid*. Obtido de Dialnet: <https://dialnet.unirioja.es/servlet/articulo?codigo=5994492>
- Glomo-Narzoles, D., & Palermo, D. (2021). Examining the academic and Non-Academic correlates of student retention. *International Journal of Publication and Social Studies*, 6(1), 34-31. <https://doi.org/10.18488/journal.135.2021.61.34.41>
- Gomme, I., & Gilbert, S. (1984). Paying the Cost: Some observations on the problem of postsecondary student attrition. *Canadian Journal of Higher Education*, 14(3), 95-100. <https://doi.org/10.47678/cjhe.v14i3.182947>
- Kehm, B., Larsen, M., & Sommersel, H. (2019). Student dropout from universities in Europe: A review of empirical literature. *The Hungarian Educational Research Journal*, 9(2), 147-164. <https://doi.org/10.1556/063.9.2019.1.18>
- McCaul, E., Donaldson Jr., G., Coladarci, T., & Davis, W. (1992). Consequences of Dropping out of School: Findings from High School and beyond. *The Journal of Educational Research*, 85(4), 198-207. <https://doi.org/http://www.jstor.org/stable/27540476>
- Núñez-Naranjo, A. (2024). Analysis of the determinant factors in university dropout: a case study of Ecuador. *Frontiers in Education*, 9. <https://doi.org/10.3389/educ.2024.1444534>
- Pérez, P., & Aguilar, D. (2017). ESTUDIOS SOBRE DESERCIÓN ACADÉMICA y MEDIDAS ORIENTADORAS DE PREVENCIÓN EN LA UNIVERSIDAD DE LA LAGUNA (ESPAÑA). *revistaparadigma.com.br*. <https://doi.org/10.37618/PARADIGMA.1011-2251.2017.p48-71.id600>

- Rodríguez-Pineda, M., & Zamora-Araya, J. (2021). Abandono temprano en estudiantes universitarios: un estudio de cohorte sobre sus posibles causas. *Uniciencia*, 35(1), 19-37. <https://doi.org/10.15359/ru.35-1.2>
- RTP. (20 de June de 2024). *Abandono escolar no ensino superior aumentou*. Obtido de RTP Noticias: https://www.rtp.pt/noticias/pais/abandono-escolar-no-ensino-superior-aumentou_v1580568
- RTP. (27 de June de 2025). *Ensino Superior. Taxa de abandono é maior no setor privado*. Obtido de RTP Noticias: https://www.rtp.pt/noticias/economia/ensino-superior-taxa-de-abandono-e-maior-no-setor-privado_v1665162
- Scott, G., Shah, M., Grebennikov, L., & Singh, H. (2008). *Improving Student Retention: A University of Western Sydney case study*. Obtido de <https://eric.ed.gov/?id=EJ1055597>
- Simpson, O. (2005). The costs and benefits of student retention for students, institutions and governments. *The Open University*, 2(3), 34-43. <https://doi.org/https://oro.open.ac.uk/6761/>
- Sosu, E., & Pheunpha, P. (2019). Trajectory of University Dropout: Investigating the cumulative effect of academic vulnerability and proximity to family support. *Frontiers in Education*, 4. <https://doi.org/10.3389/feduc.2019.00006>
- Tinto, V. (1975). Dropout from Higher Education: A Theoretical Synthesis of Recent Research. *Review of Educational Research*, 45(1), 89-125. <https://doi.org/10.3102/00346543045001089>
- Tinto, V. (2010). From theory to Action: Exploring the institutional conditions for student retention. *Higher education*, 51-89. https://doi.org/10.1007/978-90-481-8598-6_2
- Torres, J., Santiago, A., Izaguirre, J., Garduza, S., García, M., & Alejandro, G. (2024). Multilayer fuzzy inference system for predicting the risk of dropping out of school at the high school level. *IEEE Access*, 12, 137523-137532. <https://doi.org/10.1109/access.2024.3425548>

Urbina-Nájera, A., Camino-Hampshire, J., & Barbosa, R. (2020). Deserción escolar universitaria: Patrones para prevenirla aplicando minería de datos educativa. *RELIEVE - Revista Electrónica de Investigación y Evaluación Educativa*, 26(1).
<https://doi.org/10.7203/relieve.26.1.16061>