

A EXTRAÇÃO DE DADOS, PARA *BUSINESS INTELLIGENCE* (BI): UM CASE DO IFRJ

Christiano Roofln Menezes de Oliveira
<https://orcid.org/0009-0006-4520-0331>

Marcus Vinicius Freire Junior
<https://orcid.org/0009-0008-3132-0142>

Ricardo Esteves Kneipp
<https://orcid.org/0000-0003-4899-1731>

Shirley Nunes Costa Santos
<https://orcid.org/0009-0003-5859-4446>

RESUMO

O presente artigo propõe a extração de dados da plataforma Nilo Peçanha para a elaboração de um Dashboard, utilizando a ferramenta Power BI por meio de um processo ETL. O objetivo central é oferecer uma abordagem aplicada aos conhecimentos teóricos, visando aprimorar as práticas existentes de extração de dados para Business Intelligence (BI). A pesquisa é caracterizada como exploratória, voltada para a investigação de oportunidades de otimização e inovação. A metodologia adotada é de caráter misto, integrando elementos qualitativos e quantitativos. Além disso, inclui uma revisão da literatura especializada no assunto, explorando teorias e conceitos relevantes. A base de dados utilizada consiste em informações coletadas na plataforma Nilo Peçanha, sendo apresentados os resultados por meio de um estudo de caso.

Palavras-chave: Extração de Dados. *Business Intelligence*. Plataforma Nilo Peçanha. Dashboard.

1 INTRODUÇÃO

No contexto da evolução tecnológica, marcado pelo influxo massivo de dados e elevado grau de conectividade, informações fornecidas pelo Serviço Federal de Processamento de Dados (SERPRO) para o ano de 2023 indicam que a humanidade gerou aproximadamente 44 zettabytes de dados até 2020. Salienta-se que essa produção de dados mantém uma trajetória de crescimento exponencial. Diante desse cenário, torna-se evidente a necessidade de implementação de uma estratégia que possibilite o tratamento eficiente e eficaz dessas informações.

Diante desse cenário, emergiu a Business Intelligence (BI) como uma solução promissora. Este artigo tem como propósito demonstrar a importância da ETL (Extract, Transform and Load - Extrair, Transformar e Carregar), ilustrando-a por meio de um dashboard elaborado na ferramenta PowerBI da Microsoft, com os dados obtidos da Plataforma Nilo Peçanha. ETL é uma abordagem alternativa, embora relacionada, projetada para jogar o processamento para o banco de dados, de modo a aprimorar a performance (Batista *et al.*, 2012).

Dessa forma, o objetivo principal deste artigo é evidenciar a aplicabilidade e eficácia da ferramenta Power BI na extração de dados da plataforma Nilo Peçanha, por meio de um processo ETL, para aprimorar as práticas de BI. Considerando a experiência dos autores e a literatura revisada, busca-se destacar como o Power BI se destaca em

termos de eficiência, eficácia, facilidade de manejo, acessibilidade e segurança dos dados, consolidando sua posição proeminente tanto no ambiente de negócios quanto no meio acadêmico.

2 DESENVOLVIMENTO

2.1 A Extração de Dados

Comumente conhecida como ETL, esta estratégia visa facilitar a utilização de informações armazenadas em diversos bancos de dados, levando em consideração aspectos como tecnologia, simplicidade, tamanho e complexidade. Cada fase do processo possui particularidades e importâncias distintas. A máxima relevância do modelo reside na congruência dos dados analisados e na veracidade das informações resultantes após a aplicação de cada etapa.

Na etapa de Extração, a fase inicial, os dados são consolidados e organizados em um formato uniforme, reduzindo erros mediante sua padronização. Na Transformação, ocorre a limpeza, ajuste e higienização dos dados, visando disponibilizar ao usuário apenas informações essenciais. O Carregamento, última etapa, consiste no depósito das informações já organizadas em um novo ambiente Warehouse, essencialmente as mesmas informações com tratamentos adequados e prontas para serem utilizadas de maneira eficiente e eficaz.

2.2 BI (*BUSINESS INTELLIGENCE*)

Com o avanço de diversas tecnologias, surgiu o conceito de BI, que, de maneira resumida, representa um conjunto de conceitos e metodologias destinadas a apoiar a tomada de decisões por meio da utilização de dados provenientes de uma organização. Em termos mais detalhados, o termo possui diversas definições, como a proposta por Angeloni e Reis (2006, p.3), que o descrevem como:

O conceito de Business Intelligence envolve a compreensão de que se trata de inteligência aplicada aos negócios, composta por um conjunto de metodologias de gestão implementadas por meio de ferramentas de software. Sua função principal é proporcionar melhorias nos processos decisórios gerenciais e da alta administração, integrando em um único local todas as informações necessárias para o processo decisório.

Destaca-se que o objetivo do Business Intelligence é converter dados em conhecimento, sustentando o processo decisório para gerar vantagens competitivas.

Segundo Batista (2004 *apud* Angeloni; Reis, 2006, p.5):

Essas ferramentas podem oferecer uma visão sistêmica do negócio e auxiliar na distribuição uniforme dos dados entre os usuários, tendo como objetivo principal transformar grandes volumes de dados em informações de qualidade para a tomada de decisões. Através delas, é possível realizar cruzamento de dados, visualizar informações em diversas dimensões e analisar os principais indicadores de desempenho empresarial.

2.3 A Importância da ETL para BI

Para discutir a importância da ETL para BI, é necessário abordar a integração desses dois conteúdos, apesar de sua interdependência. A organização é o elemento central dessa união. Ao planejar uma estratégia de utilização do BI, é crucial que os dados estejam bem organizados para atender ao propósito do projeto. A criação de

hierarquias, a categorização de segmentos distintos e a identificação das particularidades de cada informação são fundamentais para alcançar a resposta desejada.

Além disso, outros aspectos ressaltam a importância da ETL para o BI. De maneira prática e benéfica, destacam-se as Oportunidades de Acesso, proporcionadas pela rapidez na integração e acesso aos dados. A Redução da Possibilidade de Erros é obtida pela ausência de manipulação manual nos dados armazenados. Adicionalmente, a consistência e qualidade são aprimoradas, garantindo maior precisão, confiabilidade e qualidade, pois os dados já foram submetidos a processos de refinamento.

2.4 A Plataforma Nilo Peçanha

Inaugurada em janeiro de 2018, a Plataforma Nilo Peçanha é um repositório abrangente de dados da Rede Federal de Educação Profissional, Científica e Tecnológica. Seu propósito principal é aprimorar e subsidiar a avaliação dos processos educacionais, por meio de uma plataforma colaborativa que envolve membros da Rede de Coleta, Validação e Disseminação das Estatísticas (Revalide), igualmente estabelecida em conjunto com a plataforma.

A finalidade é consolidar dados relativos ao corpo docente, discente, pessoal técnico-administrativo e despesas financeiras das unidades da Rede Federal. Segundo a Secretaria de Educação Profissional e Tecnológica - SETEC (2019), ela confere maior confiabilidade às informações estatísticas oficiais da Rede Federal. Conforme expresso por Gustavo Henrique Moraes (portal mec.gov.br), a Plataforma Nilo Peçanha é o ambiente virtual destinado à coleta, validação e divulgação das estatísticas oficiais da Rede Federal.

2.5 A Utilização do Power BI

O Power BI, lançado pela Microsoft em 2015, representa uma ferramenta de análise de dados destinada à criação de relatórios e dashboards interativos provenientes de diversas fontes de dados. Esta aplicação, considerada uma evolução do Excel, incorpora funcionalidades adicionais, como Power Query, Power Pivot e Power View. Integrado à plataforma de BI da Microsoft, o Power BI busca transformar dados dispersos em informações coesas e visualmente envolventes. Conforme destacado por Turban (2009), os objetivos fundamentais do BI incluem facilitar o acesso interativo aos dados.

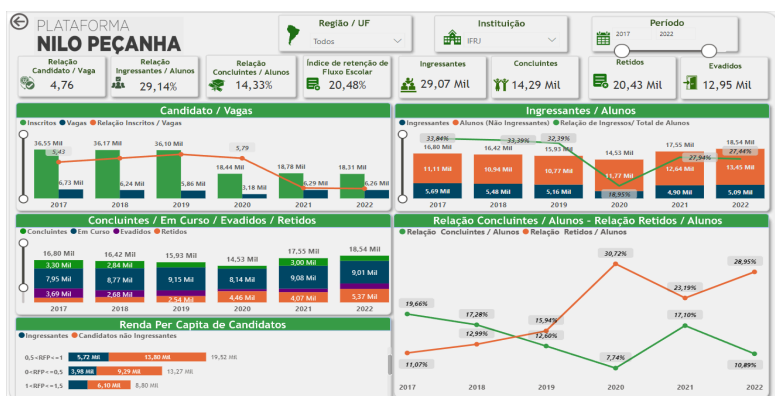
O Power BI possibilita a coleta, organização, tratamento e análise eficientes de dados de negócios, oferecendo a capacidade de conectar-se e extrair dados de diversos softwares, criar cálculos, indicadores, previsões e facilitar tomadas de decisão. Além disso, viabiliza o compartilhamento de insights dentro da organização ou a incorporação desses dados em aplicativos ou sites. A base de dados da Plataforma Nilo Peçanha, utilizada para criar indicadores no Power BI, foi obtida por meio do módulo extrator PNP. Este módulo possibilita o download de planilhas em formato CSV, divididas nas categorias "Indicadores de Gestão", "Dados Orçamentários" e "Dados de Gestão de Pessoas". Ao conectar essas planilhas à ferramenta Power BI, foi possível relacioná-las por meio dos dados de Ano, Região, Estado e Instituição.

2.6 Dashboard

Com o propósito de fornecer uma visão de fácil compreensão sobre os dados coletados após o processo ETL e sua transformação para BI, identifica-se que a inclusão de um Dashboard gerado por meio do Power BI seria uma adição complementar significativa ao estudo. Dessa forma identificou-se, em relação aos dados do IFRJ, aquilo que se destacou como mais relevante, levando em consideração a agilidade e a interatividade proporcionadas pela ferramenta.

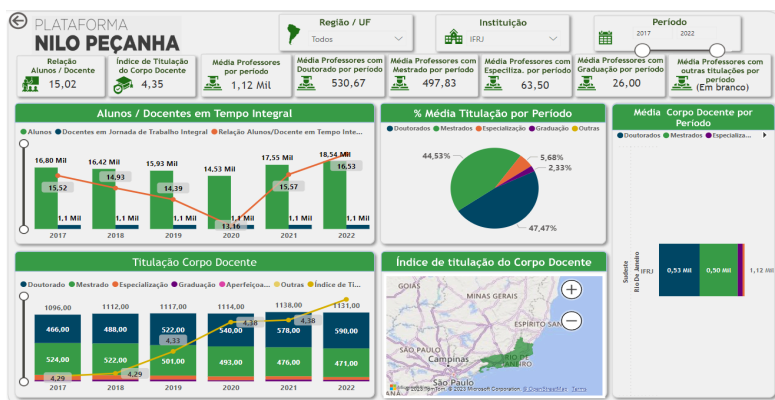
No **Painel 1** é possível analisar os dados Acadêmicos do corpo docente. Como, quantificação da taxa de retenção do fluxo escolar em relação ao total de alunos e a quantificação da taxa de ingressos em relação ao total de alunos. No **Painel 2** é possível analisar dados como, Número de Alunos por Docente em Tempo Integral. No **Painel 3** é possível analisar dados como os gastos por aluno, por região e para o País.

Painel 1: Acadêmico – Alunos (Filtrado Instituição IFRJ)



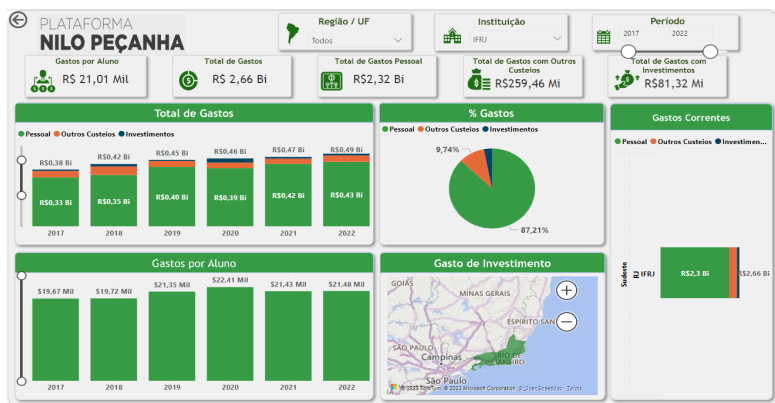
Fonte: Autoria própria (2023).

Painel 2: Acadêmico – Docente (Filtrado Instituição IFRJ)



Fonte: Autoria própria (2023).

Painel 3: Gastos / Custos (Filtrado Instituição IFRJ)



Fonte: Autoria própria (2023).

3 CONCLUSÃO

Com base nos dados fornecidos pela Plataforma Nilo Peçanha, e com o intuito de apresentar determinados resultados obtidos pelo Instituto Federal do Rio de Janeiro (IFRJ), a conclusão acerca da significativa importância do repositório de dados é alcançada.

A análise abrange não apenas os métodos utilizados na extração desses dados, mas também a abordagem específica direcionada ao seu tratamento, bem como a identificação do público-alvo que demonstra interesse nos resultados obtidos.

Destaca-se que a utilização da ferramenta para análise de dados do IFRJ não apenas simplifica o processo de extração e tratamento, mas também viabiliza uma tomada de decisão mais assertiva e ágil, pela interatividade proporcionada pela ferramenta.

A capacidade de analisar e interpretar os dados de maneira eficiente oferece uma base sólida para decisões informadas e estratégicas, ampliando a eficácia das ações a serem empreendidas.

REFERÊNCIAS

- ANGELONI, Maria T.; REIS, Eduardo S. Business Intelligence como Tecnologia de Suporte à Definição de estratégias para melhoria da qualidade do ensino. In: **Encontro da ANPAD**. 2006 v.1.p. 16 páginas.
- BATISTA, C. F.; SOUZA, E. P. R.; CORREIA NETO, J. S.; DORNELAS, J. S. Proposta de data mart para análise de faturamento de empresa de varejo utilizando software livre. **Revista Brasileira de Administração Científica**. Aracaju, v. 3, n. 2, p. 163-180, set. 2012.
- BATISTA E. O. **Sistemas de informação**. São Paulo: Saraiva, 2004.
- BRASIL. Ministério da Educação. **Plataforma reunirá dados sobre o ensino técnico e tecnológico**. Disponível em: <http://portal.mec.gov.br/PlataformaNiloPeCAanha>. Acesso em: 26 nov. 2023.
- Secretaria de Educação Profissional e Tecnológica. **Relatório Anual de análise dos indicadores de gestão das Instituições Federais de Educação Profissional, Científica e Tecnológica (Exercício 2018)**. Ministério da Educação, Brasil. 2019.
- TURBAN, Efraim et al. **Business Intelligence**: um enfoque gerencial para a inteligência do negócio. Bookman Editora, 2009.