



De 11 a 14 - NOVEMBRO DE 2024

**Inteligência Artificial
na Gestão de Operações:**
Limitações e possibilidades



APLICAÇÃO DA CIÊNCIA DE DADOS NA ENGENHARIA DE PRODUÇÃO: TRANSFORMANDO DADOS EM DECISÕES ESTRATÉGICAS

EDEN DE OLIVEIRA SOUZA – edensouza02@ufu.br
UNIVERSIDADE FEDERAL DE UBERLÂNDIA – UFU

EUGÊNIO PACCELI COSTA – eugenio@ufu.br
UNIVERSIDADE FEDERAL DE UBERLÂNDIA - UFU

ÁREA: 6. ENGENHARIA ORGANIZACIONAL
SUBÁREA: 6.4 – GESTÃO DA INFORMAÇÃO

RESUMO: A ANÁLISE DE DADOS SE TORNOU CRUCIAL PARA IMPULSIONAR A INOVAÇÃO, ALINHANDO PROCESSOS ORGANIZACIONAIS COM DEMANDAS EXTERNAS E DIRECIONANDO DECISÕES ESTRATÉGICAS FUNDAMENTAIS. OS SISTEMAS DE GERENCIAMENTO DE BANCO DE DADOS (SGBD) DESEMPENHAM UM PAPEL FUNDAMENTAL NA ORGANIZAÇÃO E INTEGRIDADE DOS DADOS, ESPECIALMENTE EM ÁREAS QUE EXIGEM CONFIANÇA E PRECISÃO. ESTE ESTUDO ILUSTRA COMO A CIÊNCIA DE DADOS PODE FORTALECER A ENGENHARIA DE PRODUÇÃO, UTILIZANDO SQL PARA GERENCIAR INFORMAÇÕES E DESCOBRIR CORRELAÇÕES SIGNIFICATIVAS, CONVERTENDO DADOS BRUTOS EM INSIGHTS VALIOSOS. FOI REALIZADO UMA ANÁLISE DETALHADA DE UM BANCO DE DADOS RELACIONAL FORNECIDO PELO GOVERNO FEDERAL, QUE ABRANGE DADOS DO CENSO DEMOGRÁFICO DE 2010 A 2022 SOBRE CURSOS E INSTITUTOS DE ENSINO SUPERIOR (IES). COM FERRAMENTAS AVANÇADAS DE GERENCIAMENTO DE BANCO DE DADOS, A APLICAÇÃO DA CIÊNCIA DE DADOS PERMITE REVELAR TENDÊNCIAS SUTIS E RELAÇÕES NÃO ÓBVIAS QUE PODEM PASSAR DESPERCEBIDAS POR MÉTODOS CONVENCIONAIS. BANCOS DE DADOS RELACIONAIS DESEMPENHAM UM PAPEL CRUCIAL AO MANTER A CONSISTÊNCIA E EVITAR REDUNDÂNCIAS, PROPORCIONANDO INFORMAÇÕES PRECISAS E CONFIÁVEIS PARA ORIENTAR DECISÕES EMPRESARIAIS ESTRATÉGICAS. AO ADOTAR UMA ABORDAGEM ORIENTADA POR DADOS, AS ORGANIZAÇÕES PODEM COMPREENDER MELHOR OS INTER-RELACIONAMENTOS ENTRE DADOS, FACILITANDO PREVISÕES MAIS ROBUSTAS E CONFIÁVEIS PARA RESULTADOS FUTUROS.

PALAVRAS-CHAVES: CIÊNCIA DE DADOS; BANCO DE DADOS RELACIONAL; TOMADA DE DECISÕES BASEADA EM DADOS, ENGENHARIA DE PRODUÇÃO.

APLICAÇÃO DA CIÊNCIA DE DADOS NA ENGENHARIA DE PRODUÇÃO: TRANSFORMANDO DADOS EM DECISÕES ESTRATÉGICAS

ABSTRACT: DATA ANALYSIS HAS BECOME CRUCIAL FOR DRIVING INNOVATION, ALIGNING ORGANIZATIONAL PROCESSES WITH EXTERNAL DEMANDS, AND DIRECTING KEY STRATEGIC DECISIONS. DATABASE MANAGEMENT SYSTEMS (DBMS) PLAY A FUNDAMENTAL ROLE IN DATA ORGANIZATION AND INTEGRITY, ESPECIALLY IN AREAS THAT REQUIRE RELIABILITY AND ACCURACY. THIS STUDY ILLUSTRATES HOW DATA SCIENCE CAN STRENGTHEN PRODUCTION ENGINEERING BY USING SQL TO MANAGE INFORMATION AND DISCOVER MEANINGFUL CORRELATIONS, CONVERTING RAW DATA INTO VALUABLE INSIGHTS. A DETAILED ANALYSIS WAS CARRIED OUT ON A RELATIONAL DATABASE PROVIDED BY THE FEDERAL GOVERNMENT, COVERING DATA FROM THE DEMOGRAPHIC CENSUS FROM 2010 TO 2022 ON HIGHER EDUCATION COURSES AND INSTITUTES (HEIS). WITH ADVANCED DATABASE MANAGEMENT TOOLS, THE APPLICATION OF DATA SCIENCE ALLOWS FOR THE REVELATION OF SUBTLE TRENDS AND NON-OBVIOUS RELATIONSHIPS THAT MAY GO UNNOTICED BY CONVENTIONAL METHODS. RELATIONAL DATABASES PLAY A CRUCIAL ROLE IN MAINTAINING CONSISTENCY AND AVOIDING REDUNDANCY, PROVIDING ACCURATE AND RELIABLE INFORMATION TO GUIDE STRATEGIC BUSINESS DECISIONS. BY ADOPTING A DATA-DRIVEN APPROACH, ORGANIZATIONS CAN BETTER UNDERSTAND THE INTERRELATIONSHIPS BETWEEN DATA, FACILITATING MORE ROBUST AND RELIABLE PREDICTIONS FOR FUTURE OUTCOMES.

KEYWORDS: DATA SCIENCE; RELATIONAL DATABASE; DATA-BASE DECISION MAKING; PRODUCTION ENGINEERING.

1. INTRODUÇÃO

De acordo com Sant'Ana (2016, p. 116-142), a possibilidade de acessar uma grande quantidade de dados está em crescimento nos últimos anos, devido ao aumento nas formas de coletar, armazenar e recuperar os dados. Todo esse conjunto de informação armazenada e escalável está sendo chamado de *Big Data*.

As tecnologias emergentes estão impactando todos os negócios, e assim os dados se tornam um insumo que poderá trazer competitividade. Então, a necessidade de se realizar a análise de dados se torna um processo de inovação, procurando por oportunidades que se identificam com os objetivos da empresa, gerando novas ideias, norteando as tomadas de decisões e integrando os dados com os processos dentro da própria corporação. As tomadas de decisão estratégicas visam alinhar os processos organizacionais da empresa com as exigências exteriores. (LI E GRIFFIN, 2023). Porém, necessita-se de um sistema que gerencie esse banco de dados.

Para Muniz *et. al* (2015), em um Sistema de Gerenciamento de Banco de Dados (SGBD) as informações estão relacionadas, mantidas organizadas e seguindo uma estrutura pré-definida, sendo cada vez mais utilizadas nas áreas em que a confiança e integridade dos dados precisam ser mantidas, o que se faz necessário o uso de banco de dados relacionais para cumprir esse objetivo.

O presente trabalho visa mostrar como a Ciência de Dados pode auxiliar as áreas da Engenharia de Produção a nortear suas tomadas de decisões, desenvolvendo *insights* por meio de banco de dados relacional, utilizando um sistema para gerenciamento com uso de SQL (*Structured Query Language*) como ferramenta de gestão de informação para acessar os dados e identificar correlações. Assim, transformando dados em informações que mostram como alguns conjuntos de dados se comportam tendo uma visão diferente do convencional e como podem ser utilizados para decidir o caminho a ser seguido.

2. REFERENCIAL TEÓRICO

A seguir, será contextualizado alguns dos termos que são muito utilizados na Ciência de dados:

- **Big Data**

De acordo com Rautenberg e Carmo (2019), o termo é originado da alta produção de dados, caracterizando assim o alto volume e uma caracterização distinta, e que se necessita uma forma não convencional de processamento, e assim indicando seu dinamismo e complexidade. Os autores também destacam que os profissionais necessitam de habilidades diversas para manipular este conjunto de dados e visualizar as informações extraídas, finalizando assim com tomadas de decisões baseadas nas informações obtidas.

Segundo Chang e Grady (2019), *Big Data* precisa possuir o que é definido como os 4 Vs, que são:

- **Volume:** Conjunto extenso de dados para ser analisado e armazenado;
- **Velocidade:** A taxa do fluxo dos dados é elevada, pois é necessário processar um alto conjunto de dados em pouco tempo;
- **Variedade:** Significa que a análise precisa ser feita em muitos repositórios com domínios e tipos diferentes de dados;
- **Variabilidade:** Essa característica significa que todo o formato citado anteriormente pode sofrer qualquer tipo de variação, mudando a sua estrutura por completo.

De acordo com Oracle (2024), o processo que as organizações orientadas a dados usam para realizar a coleta, gerar relatórios, realizar análises e tomar decisões em um grande conjunto de dados é o ETL (Extract, Transform, and Load). Este processo combina a limpeza e organização de dados de múltiplas fontes provenientes de diversos bancos de dados (IBM, 2024).

- **Dados, informação e conhecimento**

De acordo com Peixe *et al.* (2019, p. 22-23), o dado é definido como a circunstância de algo antes de ser organizado, agregado ou relacionado, sendo considerado a matéria-prima para a obtenção da informação. Já a informação consiste na configuração desses dados de forma que as pessoas possam entender e compreender. O conhecimento se dá quando ocorre a reflexão, buscando sintetizar e contextualizar a informação.

Por outro lado, Laudon e Laudon (2014, p. 13) consideram o dado como o fato ocorrido, que foi somente coletado e armazenado sem sofrer nenhuma transformação, análise e agregação de forma a relacionar entre si, proporcionando o entendimento do que ele

representa. Já a informação é a modelagem dos dados de forma a se ter significado, sendo assim o processo de transformação da matéria-prima (dado). O conhecimento é a utilização da informação gerada pelo processo para tomada de decisões.

- **Banco de dados**

Segundo Laudon e Laudon (2014, p. 183), um banco de dados é um conjunto de dados que se relacionam entre si. Os principais tipos são:

- **Relacional:** Segundo os mesmos autores, organizam os dados em tabelas que relacionam colunas e linhas. Cada coluna representa um atributo e só pode receber os tipos de dados referente àquele atributo. Uma característica é o uso do SQL (*Structured Query Language*) (MARQUESONE, 2016, p. 40).

De acordo com o Oracle (2024), um banco de dados relacional possui dados que estão relacionados entre si, em que cada linha possui uma chave que a identifica, e as colunas são atributos daqueles dados. Cada valor presente na tabela é único para cada atributo. Ainda acrescenta que tal modelo traz benefícios para o gerenciamento dos dados seguindo o modelo ACID sendo eles:

- **Atomicidade:** Significa que a alteração deve ser feita como um todo;
- **Consistência dos dados:** Os dados devem ser consistentes em várias instâncias. Ou seja, toda alteração deve ser previsível; caso contrário, deve-se retornar a alguma versão anterior, garantindo a integridade;
- **Isolamento:** Cada relação deve-se ser feita independente, então cada mudança se faz invisível a outro usuário até que ela seja completada;
- **Durabilidade:** Garante que as transações se tornem permanente quando forem completadas

Com tais características, pode-se garantir a integridade e confiabilidade presente nos bancos de dados relacionais, e assim fazer relações que geram *insights* valiosos para as organizações.

- **Não relacional:** É utilizado quando se tem a necessidade de uma performance superior e é preciso lidar com escalabilidade e esse tipo de banco de dados é focado a lidar com grandes volumes de dados, ele encarrega-se como estruturas não convencionais e armazenagem de estruturas dinâmicas. . Uma característica é o uso do NoSQL (*Not only SQL*) (OLIVEIRA *et al.* 2018).

- **SQL Server**

De acordo com o Google (2024), SQL SERVER é um RDBMS (*Relational*

Database Management System) desenvolvido pela *Microsoft*. É estruturado em tabelas baseado em linhas e colunas, podendo conectar dados de diferentes tabelas sem a necessidade de armazenar os mesmos dados várias vezes no banco de dados. Sendo conhecido por ter uma alta disponibilidade, desempenho eficiente em cargas de trabalho elevadas e uma fácil integração com aplicativos. Este *Software* foi escolhido para a realização da análise dos dados deste artigo.

- **Data Science como suporte à decisão**

Para Chang e Grady (2019), *Data Science* se refere à gestão e execução do processamento de dados de forma ampla, contendo diferentes métodos para se aproveitar o máximo dos dados e resolver problemas. Sendo o profissional que possui um grande entendimento das necessidades do negócio, ter habilidades analíticas, de *softwares* e de sistemas de engenharia para conseguir manipular o processo por completo no ciclo de vida analítico. Os *insights* que ele produz podem guiar melhorias nas estratégias da empresa, permitindo que adotem a estratégia MD³ (*Make Data-Driven Decisions*) (AFSHARIAN, 2024). A tomada de decisão baseada em dados promove a excelência operacional, conseguindo desenvolver novos produtos, serviços e modelos de negócios que realmente agreguem valor aos clientes e fornecedores, e assim estreitando ainda mais o relacionamento entre eles (Laudon e Laudon, 2014, p. 11-12).

Chang e Grady (2019) afirmam que para atingir resultados no processo de *Data Science*, é necessário que se tenha características importantes, são elas:

- **Veracidade:** refere-se à precisão dos dados, em que a sua qualidade pode impactar nos resultados das análises. Levando em conta o que se entende por *Garbage-in, Garbage-out*, que entradas ruins geram conclusões ruins;
- **Validade:** É a característica de que os dados podem ser usados para responder a um problema específico. Ou seja, a utilização de dados que ainda são válidos para aquela análise específica;
- **Volatilidade:** Para analisar acúmulos de dados, deve-se entender as mudanças na produção dos dados para poder corrigir esses desvios, permitindo assim uma análise histórica de todos os dados;
- **Visualização:** A visualização é um importante passo para que as pessoas entendam os dados, as análises ou os resultados. Sendo elas:
 - **Visualização exploratória:** Refere-se à exploração de todo o conjunto de dados original, para se entender como os elementos

presentes nesse banco de dados se relacionam entre si.

- **Visualização avaliativa:** Trata da avaliação e entendimento da performance e uma análise em particular ou *machine learning*. Esse tipo de visualização é implementado quando se tem um alto volume de dados e se necessita testar alguma consulta ou algoritmo antes de implementar no todo. Essa fragmentação é chamada de *Small Data*.
- **Visualização Explicativa:** Corresponde à apresentação de dados complexos em uma forma de fácil compreensão para os tomadores de decisão. Termo conhecido como a criação de *Dashboards*.
- **Valor:** É a medida que se usa para se definir o ganho com a implementação do processo de análise de dados ou do potencial que aquele conjunto de dados terá ao ser analisado.

3. MÉTODOS

No presente artigo, foi desenvolvida a análise de um banco de dados relacional aberto disponibilizado pelo governo federal. Esses dados são referentes ao Censo Demográfico dos anos de 2010 até 2022, nele consta as situações de cursos e IES (Institutos de Ensino Superior) separados em duas tabelas para cada ano. Com o auxílio de uma ferramenta de gerenciamento de banco de dados, foi possível filtrar os dados do Censo, possibilitando tirar informações e fazer relações sobre eles.

4. ANÁLISE DE DADOS E RESULTADOS

Foi desenvolvida uma sentença SQL com o objetivo de identificar a existência, o advento ou o desaparecimento de cursos durante todo o período do banco de dados acessado. A consulta será mostrada na Figura 1 a seguir:

FIGURA 1 - Consulta para identificar cursos ausentes no banco de dados.

```
--- Cursos que estão presentes no banco de dados
SELECT distinct(CURSO_2011)
FROM (
    SELECT DISTINCT db_2011.NO_CINE_ROTULO AS CURSO_2011
    FROM dbo.MICRODADOS_CADASTRO_CURSOS_2011 AS db_2011
    LEFT JOIN dbo.MICRODADOS_CADASTRO_CURSOS_2010 AS db_2010
    ON db_2010.NO_CINE_ROTULO = db_2011.NO_CINE_ROTULO
    WHERE db_2010.NO_CINE_ROTULO IS NULL
) AS A

--- Cursos que desapareceram do banco de dados
SELECT distinct(CURSO_2010)
FROM (
    SELECT DISTINCT db_2010.NO_CINE_ROTULO AS CURSO_2010
    FROM dbo.MICRODADOS_CADASTRO_CURSOS_2010 AS db_2010
    LEFT JOIN dbo.MICRODADOS_CADASTRO_CURSOS_2011 AS db_2011
    ON db_2010.NO_CINE_ROTULO = db_2011.NO_CINE_ROTULO
    WHERE db_2011.NO_CINE_ROTULO IS NULL
) AS A
```

Fonte: Autoria Própria (2024).

A consulta tem o objetivo de encontrar os cursos que estão presentes na tabela de um ano e ausente no ano subsequente. Ao se analisar as consultas, notou-se que quatro cursos se comportam de maneira peculiar, aparecendo (ou já estava presente) e desaparecem ao decorrer dos anos. Para o desenvolvimento do artigo, foi escolhido dois deles, com o objetivo de se aprofundar em seu comportamento e das IES que os oferecem.

4.1 Engenharia Portuária

Para obter os dados apenas desse curso, foram filtrados no BD, do qual foi separada a quantidade de ingressantes e concluintes do curso em questão, a consulta foi estendida para todos os dados referentes, como é mostrado na Figura 2:

FIGURA 2 - Consulta que filtra as vagas, concluintes e ingressantes do curso ao decorrer dos anos.

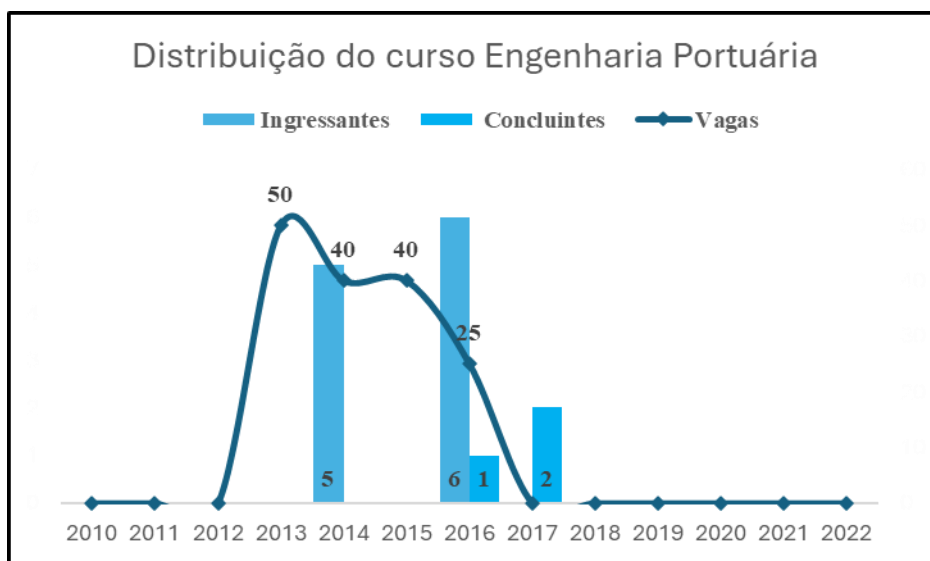
```
SELECT
    ANO, CURSO, QTD_VAGAS, QTD_ING, QTD_CONC
FROM (
    SELECT
        NO_CINE_ROTULO AS Curso,
        SUM(COALESCE([QT_ING], 0)) AS QTD_ING,
        NU_ANO_CENSO AS ANO,
        SUM(COALESCE([QT_CONC], 0)) AS QTD_CONC,
        SUM(COALESCE([QT_VG_TOTAL], 0)) AS QTD_VAGAS
    FROM
        dbo.MICRODADOS_CADASTRO_CURSOS_2010
    WHERE
        QT_ING < 99999999 and NO_CINE_ROTULO
        like 'Engenharia portuária'
    GROUP BY
        NO_CINE_ROTULO,
        NU_ANO_CENSO
    UNION
```

Fonte: Autoria Própria (2024).

Ao analisar o curso de Engenharia Portuária, verificou-se que este é oferecido, exclusivamente por uma instituição privada, no modelo presencial. Foram obtidos dados

referentes às vagas para o curso ao longo de quatro anos, conforme a distribuição apresentada no Gráfico 1, exibido a seguir:

GRÁFICO 1 - Fluxo de alunos no curso de Engenharia Portuária.



Fonte: Autoria Própria (2024).

Houve registro de vagas para o curso nos anos de 2013 a 2016, porém houve registro de ingressantes somente entre 2014 e 2016. Apenas 11 alunos ingressaram, correspondendo à 7% do total de vagas ofertadas. A última oferta do curso ocorreu em 2017. Dos ingressantes, apenas 3 alunos chegaram a a concluir o curso, sendo aproximadamente 30% dos ingressantes.

Para analisar a razão da falta de ausência, foi estudada a instituição que o ofertava o curso. Para isso, foi utilizado a consulta que é mostrada na Figura 3:

FIGURA 3 - Consulta para identificação da IES que ofertou o curso.

```
SELECT *
FROM (
    SELECT
        dbc.NU_ANO_CENSO,
        dbc.NO_CINE_ROTULO AS CURSO,
        dbi.NO_IES as Nome_univ,
        dbc.TP_REDE AS Rede_ensi, --1 público e 2 privado
        dbi.SG_IES as codigo_fac,
        TP_MODALIDADE_ENSINO as Tipo_ensino --1 Presencial e 2 EAD
    FROM
        dbo.MICRODADOS_CADASTRO_CURSOS_2013 as dbc
    INNER JOIN
        dbo.MICRODADOS_CADASTRO_IES_2013 as dbi
        on dbc.CO_IES = dbi.CO_IES
    WHERE
        DBC.NO_CINE_ROTULO
        LIKE 'Engenharia portuária'
    UNION
```

Fonte: Autoria Própria (2024).

A partir do código da IES, foi possível fazer uma análise da estrutura da entidade com relação a seus vínculos empregatícios, quantidade de alunos por ano e cursos ofertados. Na Figura 4, está apresentada a sentença utilizada para obter as informações sobre os profissionais colaboradores da IES em estudo.

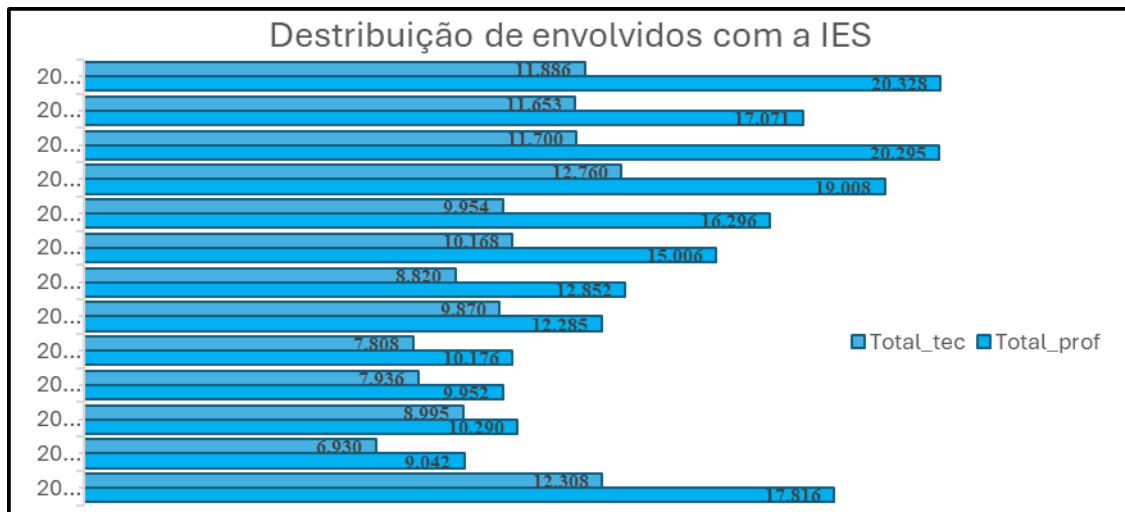
FIGURA 4 - Consulta para obter informações sobre profissionais colaboradores da IES.

```
SELECT *
FROM (
    SELECT dbc.NU_ANO_CENSO,
        NO_MANTENEDORA AS Razão_social,
        dbi.NO_IES as Nome_univ,
        dbi.SG_IES as codigo_fac,
        SUM(COALESCE([QT_MAT],0)) AS QUANT_TOTA_ALUN,
        SUM(COALESCE([QT_TEC_TOTAL],0)) AS QUANT_TEC_ADM_FUNC,
        SUM(COALESCE([QT_DOC_TOTAL],0)) AS NT_PROF,
        dbc.CO_IES AS COD_UNIV,
        SUM(COALESCE([QT_CURSO],0)) AS QUANT_CURSOS
    FROM dbo.MICRODADOS_CADASTRO_CURSOS_2010 as dbc
    INNER JOIN dbo.MICRODADOS_CADASTRO_IES_2010 as dbi
    ON dbc.CO_IES = dbi.CO_IES
    WHERE dbi.NO_IES LIKE
    GROUP BY dbc.NU_ANO_CENSO,
        dbi.NO_IES, dbi.SG_IES, dbc.CO_IES, NO_MANTENEDORA
    HAVING dbc.CO_IES = 227
    UNION
```

Fonte: Autoria Própria (2024).

Com os resultados obtidos com a consulta, foi elaborado o Gráfico 2.

GRÁFICO 2 - Distribuição de técnicos/administrativos e professores com o passar do tempo.

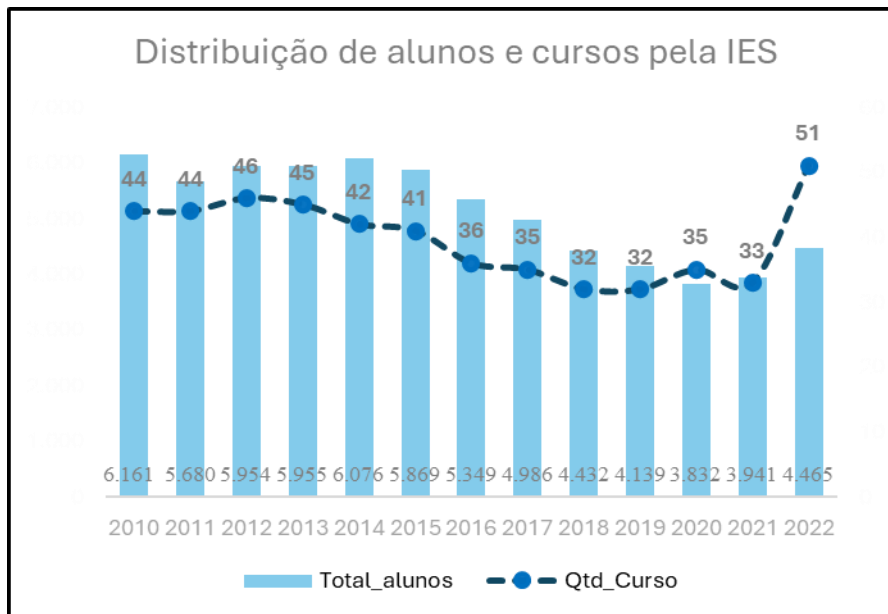


Fonte: Autoria Própria (2024).

Pode-se observar que a quantidade de professores sofreu um decréscimo desde o ano de 2013 até 2018. Após esse período, teve uma quantidade média de 9.000 a 10.000 professores. O decréscimo na quantidade de docentes pode ter sido o resultado de desativação de cursos que não foram tão procurados por alunos. Com relação aos técnicos/administrativos, não ocorreu uma queda tão acentuada, mas é possível ver que seus números diminuíram no período anterior, o que pode ser uma consequência no decréscimo de cursos ofertados.

Relacionando a quantidade de alunos que possuem vínculo com a IES e a quantidade de cursos ofertados pela mesma, obteve-se o Gráfico 3.

GRÁFICO 3 - Distribuição de alunos e quantidade de cursos distribuídos na IES com o passar do tempo.



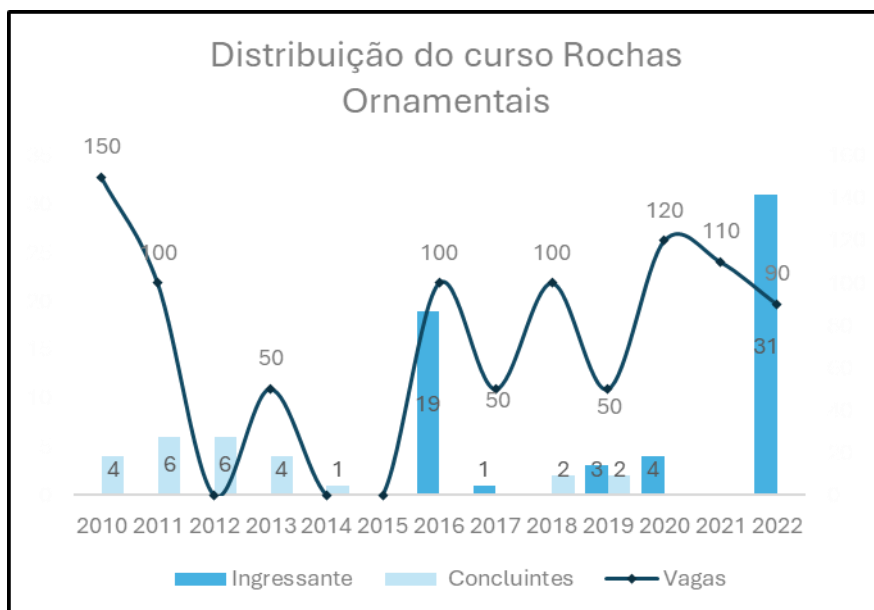
Fonte: Autoria Própria (2024).

Pode-se notar que ocorreu uma queda da quantidade de cursos no mesmo período citado anteriormente, o que reforça a ideia de corte de profissionais por esse motivo, pois a IES atua no modelo presencial (resultado alcançado utilizando a sentença presente na Figura 3). Porém, é notado que a quantidade de matrículas só sentiu um decréscimo a partir do ano de 2015, começando a ocorrer desaparecimento mais acentuados de cursos, dando a entender que poderiam estar realizando uma reformulação nas suas grades de ensino, o que manteve uma média de 5.000 alunos pós 2014, que resultou uma queda da média anterior de 17%. Entretanto, mesmo com a queda ocasionada, ainda foi possível atrair alunos. Mas deve-se levar em conta que curso de Engenharia Portuária teve seu primeiro registro em 2012, o que se pode notar no gráfico o aumento no número de cursos, provavelmente com o objetivo de atrair mais alunos.

4.2 Rochas Ornamentais

Utilizando a mesma consulta da Figura 2, foi possível analisar que o curso de Rochas Ornamentais é oferecido exclusivamente por uma instituição privada, no modelo presencial. Foram obtidos dados referentes às vagas para o curso ao longo de quatro anos, conforme a distribuição apresentada no Gráfico 4.

GRÁFICO 4 - Distribuição do curso de Rochas Ornamentais.



Fonte: Autoria Própria (2024).

O curso divulga vagas recorrentes durante quase todo o período analisado, porém apresenta uma grande variação, e no ano de 2015 ocorreu a ausência de vagas, ingressantes e concluintes, dando a entender que o curso ficou desativado neste período.

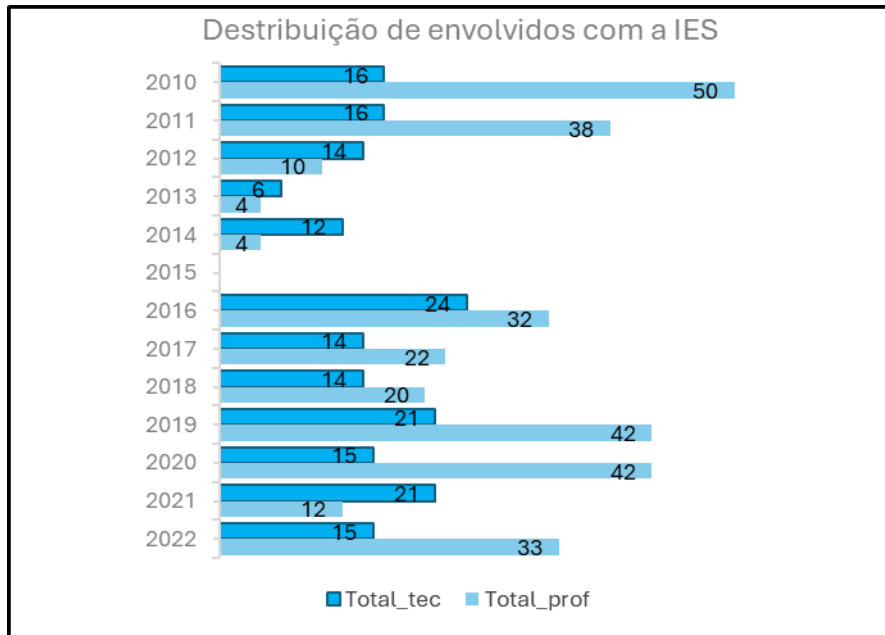
Também se nota que durante o período, o curso só recebeu ingressantes consideráveis nos anos de 2016 e 2022, mas ficando longe da quantidade total de vagas liberadas. Além do que nos anos de 2021 não ocorreu ingressantes e nem concluintes, ocasionando nesse ano uma grade totalmente vazia de alunos. Além do que, em todo o período analisado, somente 9% em média das vagas ofertadas ao todo foram preenchidas, o que pode ser uma causa de não interesse de alunos no curso ou na IES que possa estar ofertando a vaga.

Para descobrir as entidades que ofertam esses cursos durante esse período, foi utilizada a consulta a apresentada na Figura 3, trocando o campo onde está escrito 'Engenharia portuária' por 'Rochas Ornamentais', e assim mostrando que no ano de 2010 duas entidades ofertaram o curso.

4.2.1 ANÁLISE DA IES OFERTANTE DO CURSO

Seguindo a consulta realizada na Figura 4, a partir do código da IES, foi possível fazer uma análise da estrutura da entidade com relação a seus vínculos empregatícios, quantidade de alunos por ano e cursos ofertados. Sendo possível a criação do Gráfico.

GRÁFICO 5 - Distribuição de técnicos/administrativos e professores com o passar do tempo.



Fonte: Autoria Própria (2024).

Se nota que ocorre uma variação grande de profissionais no decorrer dos anos, porém entre 2011 à 2014 ocorre uma redução gradativa no número de colaboradores, até chegar no ano de 2015 não há mais registros da IES, dando a entender que o curso Rochas Ornamentais pode ter sofrido uma desativação temporária, já que no ano anterior só um aluno concluiu o curso o que pode ter deixado a mesma sem estudantes, mas que voltou de maneira acentuada no ano de 2016, atraindo novos alunos para esse curso e até aumentando o seu quadro de colaboradores. No Gráfico 6, é mostrado a variação da quantidade de cursos e alunos vinculados a essa IES:

GRÁFICO 6 - Distribuição de alunos e quantidade de cursos oferecido pela IES.



Fonte: Autoria Própria (2024).

Com o gráfico, nota-se o decréscimo acentuado na quantidade de alunos que se matricularam na IES após o ano de 2011, enquanto que a quantidade de cursos se mantém

constante, podendo ter direcionado a entidade a optar por fechar temporariamente os seus serviços em 2015. Mesmo aumentando a quantidade de cursos no ano de 2019, não resultou em acréscimo de matrículas, podendo ser relacionado que os cursos ofertados não atraíram uma quantidade relevante de alunos. Uma falta de previsibilidade nas matrículas ocasiona receitas flutuantes, dificultando projetos a mais longo prazo e assim podendo trazer problemas financeiros a IES.

5. CONSIDERAÇÕES FINAIS

Como foi visto com o uso da Ciência de Dados, pode-se descobrir tendências e fazer relações entre dados, que antes não eram explícitos utilizando o sistema convencional de análise. Devido ao banco de dados relacional manter os dados consistentes e não possuindo redundâncias, geram informações valiosas para as empresas, que podem nortear tomadas de decisões mais assertivas, diminuindo assim custos de erros estratégicos e podendo aumentar os ganhos, já que as decisões tomadas são mais fundamentadas.

Ainda a tomada de decisão com base no uso de banco de dados, permite visualizar o problema de maneira não convencional, entendendo todos os relacionamentos que os dados possuem e podendo realizar previsões de resultados baseado em seus comportamentos. E assim esperar resultados das decisões tomadas de maneira mais fidedigna.

REFERÊNCIAS

AFSHARIAN, Mohsen. **Data science essentials in business administration: A multidisciplinary perspective**. Decision Analytics Journal, 2024.

CHANG, Wo.; GRADY, Nancy. **NIST Big Data Interoperability Framework: Volume 1**, NIST Pubs 2019.

Google cloud. PostgreSQL x SQL Server: **Quais são as principais diferenças?** Disponível em: < <https://cloud.google.com/learn/postgresql-vs-sql?hl=pt-br/> > Acessado em: 19 de maio de 2024.

IBM. **What is ETL?** Disponível em: < <https://www.ibm.com/topics/etl> > Acessado em: 19 de maio de 2024.

LAUDON, Kenneth C.; LAUDON, Jane P. **Sistemas de Informação Gerenciais**. 11^o edição. São Paulo: Pearson Education do Brasil, 2014.

LI, Keyao; GRIFFIN, Mark A. **Unpacking human systems in data science innovations: Key innovator perspectives**. Technovation, v. 128, p. 102869, 2023.

MARQUESONE, Rosangela. **Big Data: Técnicas e tecnologias para extração de valor dos dados**. [s.l.] Editora Casa do Código, 2016.

MUNIZ, Emanuel Bonfante; BOTACIM, Renato Sousa; XAVIER, Bruno Missi, ALBUQUERQUE, Marcelo Schuster **Uma proposta de auditoria em sistemas JAVA com Hibernate e Aspectj em banco de dados com suporte a SQL**. REINPEC, v. 1, n. 1, p. 82–99, 5 maio 2015.

OLIVEIRA, Moacir de Souza; MELO, Nicole de Farias; SANTOS, Leandro Colevati dos; OLIVEIRA, Wellington Pinto de. **Banco de dados NO-SQL X Banco de dados SQL: Estudo de Desempenho em Grandes Massas**. South American Development Society Journal. 2018.

ORACLE. **O que é um banco de dados relacional?** Disponível em: < <https://www.oracle.com/br/database/what-is-a-relational-database/> >. Acessado em: 30 de abril de 2024.

ORACLE. **O que é ETL?** Disponível em: < <https://www.oracle.com/br/integration/what-is-etl/> > Acessado em: 19 de maio de 2024.

PEIXE, Adriana Maria Miguel; LAUER, Eduardo; PINTO, José Simão de Paula; CORECHA, Josias Farias; BALSAN, Jorge; BENCK, Larissa Lourenço Nunes. CYGAN CZUK, Mariçyn de Souza; SANTOS, Michel Cesar dos, **Gestão da informação: Temas e Abordagens. Programa de Pós-Graduação em Gestão de TI (PPGGI)**, 2019.

SANT'ANA, Ricardo César Gonçalves. **Ciclo de vida dos dados: uma perspectiva a partir da ciência da informação**. Informação & Informação, v. 21, n. 2, p. 116–142, 2016.