

IMPLANTAÇÃO DA ANÁLISE DE DADOS EM PEQUENOS NEGÓCIOS

IMPLEMENTING DATA ANALYSIS IN SMALL BUSINESSES

Leonardo Fernando Pacheco ¹

André Gomes dos Santos²

Resumo: Este artigo tem como objetivo abordar o tema da análise de dados, fornecendo uma visão ampla sobre o conceito, as técnicas e os métodos de utilização, bem como exemplos de aplicações possíveis. Embora o campo do Data Science seja relevante para empresas e negócios de todos os portes, este artigo concentrará seus esforços em destacar a importância e a relevância específicas dessa disciplina para os pequenos negócios. Nesse sentido, será explorada uma linguagem acessível e de fácil compreensão, com o intuito de simplificar os métodos de aplicação e facilitar a assimilação desses conhecimentos por parte dos empreendedores e gestores de pequenas empresas. Serão abordados exemplos práticos de como a análise de dados pode ser implementada em diferentes áreas, como marketing, vendas, operações e tomada de decisão estratégica. Além disso, serão discutidos os benefícios diretos e indiretos que os pequenos negócios podem obter ao adotar uma abordagem baseada em dados.

Palavras-chave: Análise de Dados, Data Science, Pequenos Negócios, Métodos.

Abstract: This article aims to address the topic of data analysis, providing a comprehensive overview of the concept, techniques, and methods of utilization, as well as examples of possible applications.

1 Engenharia de produção pela Faculdade São Judas

2 Graduação em engenharia química, Graduação em matemática, Mestrado em Práticas Docentes, Doutorado em ensino de ciências e matemática



Although the field of Data Science is relevant to companies and businesses of all sizes, this article will focus its efforts on highlighting the specific importance and relevance of this discipline for small businesses. In this regard, an accessible and easily understandable language will be explored, with the intention of simplifying the methods of application and facilitating the assimilation of this knowledge by entrepreneurs and managers of small businesses. Practical examples of how data analysis can be implemented in different areas such as marketing, sales, operations, and strategic decision-making will be addressed. Additionally, the direct and indirect benefits that small businesses can obtain by adopting a data-driven approach will be discussed.

Keywords: Data Analysis, Data Science, Small Businesses, Methods.

Introdução

Nos últimos anos, a análise de dados tem se mostrado uma ferramenta poderosa para a tomada de decisões em empresas de diversos setores e tamanhos. No entanto, apesar dos benefícios, muitas pequenas empresas ainda não utilizam a análise de dados em suas operações diárias. Isso pode ser devido a uma falta de conhecimento sobre as ferramentas e técnicas disponíveis, ou a recursos financeiros limitados para investir em tecnologia e pessoal especializado em análise de dados.

De acordo com dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) do IBGE, em 2019, havia cerca de 19,3 milhões de empresas ativas no Brasil. Desse total, 99% eram micro e pequenas empresas (MPEs), que respondem por mais da metade dos empregos com carteira assinada no setor privado, com um total de 16,1 milhões de trabalhadores. Além disso, segundo o Portal do Empreendedor, há cerca de 11,2 milhões de microempreendedores individuais (MEIs) registrados no país.

No decorrer deste artigo vamos discutir a importância da coleta e análise de dados para a



otimização de processos produtivos, bem como os desafios e benefícios envolvidos nessa implantação em micro e pequenas empresas no Brasil.

Serão discutidas as ferramentas e técnicas disponíveis para a análise de dados, bem como as habilidades e recursos necessários para a implementação dessas ferramentas. Além disso, serão fornecidas dicas para pequenos negócios que desejam implementar a análise de dados em suas operações.

Este artigo é relevante para pequenos negócios que buscam manter sua competitividade no mercado atual, assim como para pesquisadores interessados em investigar os benefícios e desafios da implementação da análise de dados em pequenas empresas. A análise de dados é uma ferramenta poderosa capaz de transformar a maneira como as empresas operam e tomam decisões, e espera-se que este artigo inspire mais empresas a explorarem essa tecnologia.

Justificativa

A implementação da análise de dados em pequenos negócios é extremamente importante, pois permite que essas empresas se mantenham competitivas em um mercado cada vez mais tecnológico e orientado por dados. A análise de dados pode ajudar os pequenos negócios a identificar tendências, avaliar o desempenho, identificar oportunidades de crescimento e melhorar a eficiência operacional, permitindo que tomem decisões informadas e precisas com base em evidências concretas.

A coleta e análise de dados pode fornecer informações valiosas sobre o comportamento do cliente, permitindo que os pequenos negócios identifiquem as necessidades e preferências do cliente e ajustem sua oferta de produtos ou serviços de acordo. Além disso, pode ajudar os pequenos negócios a entender melhor a concorrência e o mercado em geral, permitindo que eles identifiquem oportunidades de diferenciação e inovação.

Outro benefício importante da implementação da ciência de dados em pequenos negócios



é a melhoria da eficiência operacional. Ao analisar dados relacionados a processos internos, como controle de estoque, gerenciamento de recursos humanos e gerenciamento de fluxo de caixa, os pequenos negócios podem identificar gargalos e ineficiências e implementar soluções para melhorar a produtividade e reduzir custos.

Apesar dos muitos benefícios da implementação da ciência de dados em pequenos negócios, muitos ainda não estão utilizando essa tecnologia em seu dia a dia, seja por falta de conhecimento ou recursos financeiros limitados. É importante destacar que a implementação da ciência de dados não precisa ser cara ou complicada. Existem várias ferramentas e técnicas disponíveis que são acessíveis e fáceis de usar, e muitas vezes é possível começar com dados que a empresa já possui.

Dado o cenário atual de mercado altamente competitivo e orientado por dados, é crucial que os pequenos negócios reconheçam a importância da implementação da ciência de dados em suas operações. Aqueles que se adaptam e implementam efetivamente essa tecnologia terão uma vantagem competitiva significativa sobre aqueles que não o fazem, e estarão mais bem posicionados para crescer e se destacar em um ambiente cada vez mais desafiador.

Objetivos

O objetivo deste trabalho é apresentar a importância da implementação da ciência de dados em pequenos negócios, abordando seus benefícios e desafios, bem como as ferramentas, técnicas e recursos necessários para sua implementação. Além disso, pretende-se fornecer dicas práticas para que pequenos negócios possam implementar a análise de dados em suas operações de forma eficiente e bem-sucedida. O trabalho visa contribuir para o entendimento dos gestores de pequenos negócios sobre a importância da análise de dados como uma ferramenta estratégica para a tomada de decisões e para a competitividade no mercado atual.

Para que o trabalho tenha um bom entendimento, dentro do tema citado anteriormente ire-



mos buscar os seguintes objetivos específicos:

- Proporcionar uma visão geral sobre a análise de dados.
- Apresentar suas principais técnicas e ferramentas disponíveis para a implementação em pequenos negócios.
- Identificar os principais benefícios e desafios da implementação.

O que é Data Science

O Data Science é uma área multidisciplinar que busca aplicar métodos científicos, processos e sistemas para extrair informações e conhecimentos a partir de grandes volumes de dados estruturados e não estruturados. Esta área tem sido impulsionada pela crescente disponibilidade de dados e pela necessidade de explorá-los de forma mais eficiente.

De acordo com Coelho (2017) Data Science ou Ciência de Dados é um estudo multidisciplinar com relação aos dados e demais informações inerentes à empresa e as visões que cercam um determinado assunto. Resumidamente é uma ciência que visa estudar as informações do processo de coleta, transformação em dados sólidos, posteriormente, a análise desses dados.

O Data Science tem como principal objetivo fornecer informações acionáveis para as organizações, a fim de melhorar sua eficiência, tomar decisões mais informadas e solucionar problemas de forma mais eficaz. Com a aplicação de métodos avançados de análise de dados, os profissionais de Data Science têm a capacidade de extrair informações úteis a partir de grandes conjuntos de dados, identificando padrões e tendências que muitas vezes não seriam detectados por meios convencionais.

Segundo PROVOST e FAWCET (2016) o objetivo primordial de Data Science, é o aprimoramento da tomada de decisão, uma vez que isso geralmente é interesse direto para os negócios. Pode-se afirmar que Data Science, é a ciência que estuda os dados por meio de análises, buscando encontrar ou descobrir informações relevantes que possam vir a resolver problemas, questionamentos e gerar

conhecimento.

Portanto simplificadamente podemos definir o Data Science como uma ciência que busca extrair informação, conhecimento e insight a partir das mais variadas fontes volumes de dados, por meio da coleta, processamento e análise.

Decisão orientada por dados

Como citado anteriormente, a maior parte das decisões em pequenos negócios são tomadas simplesmente por insights e a DOD (Decisão orientada por dados) tem justamente o objetivo contrário, procurando basear as decisões na análise de dados, em vez de apenas na intuição. Essa abordagem tem sido amplamente adotada em diversos setores, incluindo negócios, finanças, saúde, governo, entre outros. Ao tomar decisões baseadas em dados, as empresas e organizações podem reduzir o risco de falhas e melhorar seus resultados, aumentando a eficácia e eficiência das suas operações. Segundo PROVOST; FAWCETT. 2016 A DOD não é uma prática do tipo “tudo ou nada”, e diversas empresas a adaptam em maior ou menor grau.

As vantagens de se trabalhar com a decisão orientada de dados tem sido demonstrada conclusivamente. Um estudo feito por Erik Brynjolfsson e sua equipe do MIT e da Penns Wharton School demonstrou como a DOD afeta diretamente no desempenho das empresas com o uso de dados para tomada de decisões. Nesse estudo foi mostrado, estatisticamente, que quanto mais as empresas se orientavam através dos dados, mais produtiva ela é mesmo tendo de controlar uma série de possíveis fatores de confusão.

E as diferenças são expressivas. Um desvio padrão a mais na escala de DOD está associado a um aumento de 4%-6% na produtividade. Além disso a DOD também está ligada com o maior retorno sobre ativos, patrimônio líquido, utilização de ativos e valor de mercado e a relação parece casual (PROVOST; FAWCETT, 2016).



FURLAN e FILHO, (2005) afirmam que possuir a informação em mãos é uma ferramenta poderosa quando se precisa tomar uma decisão, tendo isso em vista, as empresas começaram a trabalhar mais em cima de seus dados.

É importante ressaltar que a decisão orientada por dados não significa que as decisões são tomadas automaticamente pelos dados, mas sim que as informações coletadas são usadas para orientar a tomada de decisões e complementar a experiência e conhecimento dos tomadores de decisão.

Aplicação nos negócios

A aplicação dos conceitos de Data Science se dá por meio da implantação de técnicas e ferramentas de análise de dados. Para implementar essas técnicas e ferramentas, é necessário ter acesso a dados relevantes, possuir habilidades e recursos para a análise desses dados, entender o contexto do problema, o público-alvo, as restrições e oportunidades envolvidas além disso é importante estar disposto a tomar decisões baseadas nas informações coletadas.

Algumas técnicas e ferramentas comuns utilizadas para a análise de dados em pequenos negócios incluem: análise de tendências, segmentação de clientes, análise de desempenho financeiro, análise de dados de vendas, análise de dados de redes sociais, entre outras. É importante lembrar que a aplicação dessas técnicas deve ser adaptada às necessidades específicas de cada negócio e setor.

Para implementar a análise de dados em pequenos negócios, é necessário ter uma equipe capacitada sendo necessário investir na capacitação de membros da equipe para trabalhar com essas ferramentas e técnicas ou contratar serviços de consultoria especializados em análise de dados.

Além disso, é importante ter uma infraestrutura de tecnologia adequada, como bancos de dados confiáveis, Ferramentas de visualização que permitem que as informações sejam apresentadas de forma clara e objetiva, facilitando a compreensão dos dados pelos usuários, Softwares de programação específicos, como por exemplo, a linguagens de programação Python, e soluções de arma-



zenamento em nuvem que pode ser uma alternativa viável, pois permite o acesso a recursos de alta capacidade sem a necessidade de investimentos em infraestrutura local.

É importante ressaltar que a escolha das tecnologias e ferramentas adequadas deve levar em consideração as características específicas do negócio, bem como o nível de conhecimento da equipe responsável pela implementação do projeto de Data Science. De acordo com PROVOST e FAWCET (2016) os dados e a capacidade que temos de extrair conhecimento útil através deles, devem ser considerados importantes ativos estratégicos. Portanto a aplicação dos conceitos abordados neste artigo se dá por meio da implantação de uma cultura de análise de dados nas pequenas empresas, o que pode levar a uma tomada de decisão mais informada, melhores resultados financeiros e maior eficiência operacional.

Definição de problema

A definição clara e precisa do problema é um dos passos mais importantes na implementação do data science, é uma etapa crucial no processo, pois é a partir dessa definição que será possível direcionar as atividades de coleta, processamento e análise de dados. A definição do problema pode ser resumida como o processo de identificar um desafio ou oportunidade de negócio que possa ser abordado a partir da análise de dados.

De acordo com PROVOST e FAWCET (2016), para definir um problema é necessário ter clareza do objetivo da análise, definir quais variáveis serão utilizadas, ter acesso aos dados e identificar quais perguntas precisam ser respondidas. É importante também ter em mente que a definição do problema pode ser um processo iterativo, ou seja, pode ser ajustada ao longo da análise de acordo com os resultados obtidos e novas informações disponíveis. Uma boa forma de definir o problema é por meio de perguntas-chave que ajudem a esclarecer os objetivos do negócio e a identificar a fonte do problema. Algumas perguntas que podem ser úteis nesse processo são: Qual é o objetivo principal

do negócio? Qual é o principal desafio enfrentado atualmente? Qual é a fonte desse desafio? Qual é o impacto do desafio nos resultados do negócio?

Portanto a definição do problema é uma etapa crítica no processo de data science e deve ser feita de forma clara e objetiva, com base em perguntas específicas a serem respondidas, de forma a guiar todo o processo de análise de dados e evitar análises sem rumo ou sem objetivo claro. Ela deve envolver a identificação do contexto, a definição do escopo, o estabelecimento de objetivos claros, a identificação das fontes de dados e a definição das métricas de sucesso. Uma definição cuidadosa do problema pode levar a análises de dados mais precisas e resultados mais relevantes para o negócio.

Uma vez que o problema é claramente definido, é possível aplicar técnicas de data science para analisar dados e identificar soluções para o problema.

Definição do dataset

O dataset é o conjunto de dados organizados para que seja possível identificar padrões, tendências e relações entre as variáveis do problema a ser tratado, deve ser construído a partir da definição do problema com qualidade e consistência para que a análise seja precisa e confiável, ele deve conter todas as variáveis do problema a ser tratado.

Estes dados normalmente são uma mescla de diversas bases de dados, tanto adquiridas dentro no próprio negócio, quanto externas como dados de mercado, valores de fornecedores e até mesmo dados de bases públicas, todos com o intuito de enriquecer os dados e tornar o processo mais assertivo além de ser bem definido.

Uma ferramenta comumente utilizada nesta etapa é o Microsoft Excel, pois a grande maioria das pequenas empresas já possuem o pacote do Microsoft Office instalado em seus computadores, se tratando de uma ferramenta versátil que pode ser usada para armazenar, organizar e analisar dados de maneira eficiente e econômica. Ele permite que as pequenas empresas gerenciem dados de clien-



tes, produtos, vendas e despesas com a ajuda de fórmulas, tabelas dinâmicas e gráficos, além disso, o Excel pode ser integrado com outras ferramentas, como softwares de contabilidade e ERP (Enterprise Resource Planning), o que torna o gerenciamento de dados de negócios ainda mais eficiente.

Provost e Fawcet (2016) ressaltam a importância da coleta e da documentação desses dados, para que os usuários possam entender como os dados foram coletados e organizados, bem como as limitações e possíveis vieses dos dados. Isso é importante para garantir a confiabilidade e a qualidade das análises e modelos construídos com o conjunto de dados.

Transformação e higienização de dados

Essas etapas envolvem a limpeza e preparação dos dados para que possam ser utilizados em análises mais avançadas.

A transformação de dados consiste na manipulação, organização e limpeza dos dados brutos coletados de diversas fontes, a fim de torná-los aptos para análise.

Os dados brutos, por si só, geralmente não apresentam informações claras e estruturadas, sendo necessário transformá-los em um formato que possa ser utilizado por ferramentas de análise e modelagem. A transformação de dados envolve diversas técnicas e ferramentas, que permitem a integração de dados de diferentes fontes, a limpeza de dados incorretos e a normalização de formatos.

A integração de dados é fundamental para o sucesso do processo de transformação de dados. Ela consiste em reunir os dados de diferentes fontes em um único arquivo ou base de dados, de modo que possam ser analisados de forma conjunta. A integração de dados pode ser realizada manualmente, através de copiar e colar, ou com o auxílio de ferramentas específicas de integração.

A normalização de dados é uma técnica que visa tornar os dados mais uniformes e facilitar a análise. Ela consiste em transformar os dados para um formato comum, que possa ser comparado e analisado de forma consistente. A normalização de dados pode ser realizada através de técnicas como



padronização, escalonamento e discretização.

Já a higienização de dados envolve a identificação e correção de erros e inconsistências. É comum que os dados coletados contenham informações duplicadas, ausência de valores ou valores incorretos, o que pode prejudicar a análise. Por isso, é necessário realizar a limpeza dos dados, através de técnicas de filtragem, remoção de duplicatas, substituição de valores ausentes e correção de valores incorretos.

A transformação e higienização de dados são essenciais para garantir que as análises realizadas sejam precisas e confiáveis. Sem essas etapas, os dados podem estar incompletos, inconsistentes ou imprecisos, o que pode levar a conclusões incorretas e decisões equivocadas.

Mineração de dados e modelagem

A mineração de dados é uma etapa importante do processo de Data Science, é o processo de explorar grandes conjuntos de dados para encontrar padrões, tendências e informações úteis que podem ser utilizadas para tomada de decisões ou para a construção de modelos preditivos, que são capazes de prever comportamentos futuros com base nos dados históricos disponíveis. Como por exemplo prever a demanda de um determinado produto ou serviço em um determinado período, ajudando a empresa a planejar melhor a produção, os estoques e as estratégias de marketing.

A aplicação de técnicas de modelagem, como regressão, clustering, árvores de decisão, redes neurais, entre outras dependem do problema em questão e das características dos dados disponíveis. Por exemplo, a regressão pode ser utilizada para prever valores numéricos, enquanto o clustering pode ser usado para identificar grupos de dados com características semelhantes. Sendo possível identificar os clientes com maior probabilidade de cancelamento de contrato, permitindo que a empresa adote medidas preventivas para reter esses clientes.

Sendo assim, podemos dizer que o objetivo da modelagem é identificar as inter-relações



entre todas as variáveis e gerar previsões, projeções de acontecimentos futuros de series temporais, análises de perfis, categorizar indivíduos similares, criar possíveis cenários, analisar a eficácia de um processo a ser implantado entre diversas possibilidades que essa modelagem de dados nos permite.

É importante ressaltar que a construção de um modelo de qualidade requer uma análise cuidadosa dos dados disponíveis, a escolha adequada de algoritmos e técnicas de modelagem, e a validação rigorosa do modelo com dados independentes. Por isso, a contratação de profissionais qualificados e o uso de ferramentas adequadas são fundamentais para garantir o sucesso da aplicação do Data Science em um negócio.

Transformação dos resultados obtidos em linguagem de negócio

Muitas vezes, os resultados obtidos a partir das análises de dados podem ser complexos e difíceis de entender para aqueles que não são especialistas na área.

Portanto, a transformação dos resultados em uma linguagem mais acessível e compreensível para os gestores e demais colaboradores da empresa é fundamental para que eles possam entender as conclusões obtidas, facilitando a tomada de decisão e a implementação de ações baseadas nas conclusões da análise.

Os resultados das projeções obtidos pela ciência de dados devem ser feitos em linguagem de negócio, portanto devem manter o foco na objetividade e na agilidade, demonstrando os KPI's (Key Performance Indicators) que sofreram impacto com o projeto e seus resultados.

Os Indicadores-Chave de Performance, em português. São medidas utilizadas para avaliar o desempenho de uma organização, equipe, processo ou indivíduo em relação a objetivos definidos. Os KPIs são importantes para ajudar a monitorar o progresso em direção às metas, identificar áreas que precisam de melhorias, tomar decisões informadas e comunicar o desempenho para partes interessadas internas e externas. Os KPIs podem ser quantitativos, como vendas, lucro, produtividade, ou



qualitativos, como satisfação do cliente, qualidade do produto, entre outros.

Muitas vezes, não temos os dados corretos para melhor tomar decisões e/ou o talento certo para melhor apoiar a tomada de decisão a partir dos dados. (PROVOST; FAWCETT, 2016).

Analisando esta citação, podemos perceber que a aplicação do Data Science como uma ferramenta de ajuda à tomada de decisão não pode somente se prender ao fato de possuir dados, mas necessita ser levado em consideração quem fara a análise desses dados, a pessoa deve ser capaz de estudar e explorar diversas fontes com o objetivo de encontrar respostas aos problemas que surgem, se aprimorando com novas tecnologias.

Por fim a utilização do Data Science deve ser realizada para análises descritivas, buscando compreender os eventos e situações do presente, análises preditivas com o intuito de prever cenários futuros com a identificação de padrões de comportamento nos dados, análises prescritivas que nos auxiliam em possíveis situações que possam vir a ocorrer e então a análise diagnostica com a proposta de buscar entender as causas de um evento e seus detalhes.

Revisão Bibliográfica

Kelleher e Tierney (2018) afirmam que o Data Science é importante porque permite a descoberta de padrões, relações e insights que podem ser difíceis de encontrar manualmente ou através de métodos tradicionais de análise de dados. Além disso, eles argumentam que a grande quantidade de dados disponíveis atualmente em diversas áreas, como negócios, ciência, medicina, entre outras, torna o uso do Data Science essencial para extrair informações valiosas e tomar decisões embasadas em dados. O Data Science também permite a criação de modelos preditivos que podem ser usados para prever tendências e comportamentos futuros, o que pode ser extremamente útil para empresas e organizações em diversos setores.



Materiais e Métodos

Com o objetivo de fornecer uma explicação abrangente e detalhada, desde a identificação do problema até a coleta e transformação dos dados em algo útil, a fim de solucionar um problema existente, preparar um projeto futuro ou agir preventivamente em relação a um possível evento futuro, são necessárias técnicas e ferramentas eficazes de análise de dados.

Buscando exemplificar a aplicação prática das técnicas e ferramentas de análise de dados, apresentamos um estudo de caso fictício envolvendo um restaurante que deseja aprimorar seu serviço. Nesse contexto, é importante ressaltar que, mesmo que algumas ideias aparentem ser promissoras, a tomada de decisão empresarial não deve se basear somente na intuição. Dessa forma, os proprietários do restaurante decidiram utilizar as técnicas e ferramentas de análise de dados apresentadas neste artigo com o objetivo de obter uma visão mais ampla e fundamentada sobre a situação atual do estabelecimento e, assim, embasar suas decisões em informações precisas e confiáveis.

Este exemplo destaca a importância da análise de dados na tomada de decisões empresariais. Essa abordagem pode ajudar a identificar oportunidades de melhoria e fornecer informações valiosas para os proprietários e gerentes de negócios. Dessa forma, a análise de dados se torna uma ferramenta poderosa e essencial para qualquer negócio que busca crescer e se destacar em um mercado altamente competitivo.

O procedimento escolhido é o estudo de caso que possibilita ao pesquisador trazer à tona dados e informações de modo a contribuir com o saber na área do conhecimento. Ele varia no formato e na forma de aplicação, conforme a área do saber, tendendo a ser utilizado em todas as áreas do conhecimento por fazerem uso de diversas técnicas de apoio.

“O estudo de caso é uma estratégia de pesquisa empírica que investiga um fenômeno contemporâneo dentro de seu contexto da vida real, especialmente quando os limites entre o fenômeno e o contexto não são claramente definidos” (Yin, 2014, p. 16).

Para realizar este estudo, foram seguidos os seguintes métodos:

- Identificar dos objetivos de negócio: foram identificados os principais objetivos de negócio da empresa, a fim de determinar as áreas em que a utilização de Big Data poderia ser mais útil;
- Coleta de dados: foram coletados dados de por meio de técnicas de estimativa, considerando diferentes fontes empresariais.
- Processamento de dados: os dados devem ser processados usando ferramentas de estatística, para gerenciar grandes volumes de dados e identificar padrões e tendências relevantes para os objetivos de negócio da empresa;
- Análise de dados: os dados a ser analisados utilizando técnicas de análise de Big Data, estatística e outras técnicas para obter insights valiosos para a empresa;
- Visualização de dados: Vamos utilizar ferramentas de visualização de dados para apresentar os insights obtidos a partir da análise de dados;
- Implementação de soluções: a partir dos insights obtidos, foram desenvolvidas soluções para melhorar os processos de negócio da empresa e aumentar sua eficiência e rentabilidade;
- Avaliação dos resultados: a eficácia das soluções implementadas será avaliada por meio de métricas relevantes, para determinar se elas foram bem-sucedidas em atingir os objetivos de negócio estabelecidos.

As etapas de condução do estudo de caso deste trabalho são apresentadas a seguir. Primeiramente é preciso definir uma estrutura conceitual teórica, essa estrutura foi construída com base no que foi analisado no estudo da literatura pesquisada, para isso foram estudadas inúmeras publicações como livros, artigos, dissertações e teses.

Na etapa de planejamento dos casos, vamos definir o objetivo do estudo de caso. O objetivo principal é verificar a eficácia da utilização do Data Science como ferramenta de otimização de tomada de decisão na garantia da receita.

Nessa mesma etapa, é preciso escolher os meios para coleta e análise de dados em um grande conjunto de dados. Para realizar a coleta dos dados foram realizadas técnicas de coleta de dados.

Na próxima fase, vamos definir os critérios para a seleção dos casos realizados. Os principais critérios de seleção estão relacionados aos tópicos desta pesquisa, ou seja, foram pesquisadas empresas que possuem problemas similares. Para isso, foram pesquisadas informações nos meios de comunicação e consultas a funcionários das empresas. Para selecionar o mercado estudado não foi excluído nenhum setor da economia, pois segundo Manyika et al. (2011) a big data analytics pode ser utilizado em diferentes setores da economia. Dessa forma entramos em contato com empresas do ramo de bebidas, varejista e alimentício.

Definindo o problema e o objetivo.

A definição do problema é o primeiro passo que devemos tomar, ela envolve identificar claramente qual é o desafio que precisa ser solucionado ou a questão que precisa ser respondida, a fim de orientar o processo de coleta, análise e interpretação dos dados.

Ao definir o problema de forma clara e específica, é possível direcionar melhor os esforços para a solução do problema. Por exemplo, o restaurante pode estar enfrentando uma queda nas vendas e precisa entender a causa para poder tomar medidas corretivas. Nesse caso, a definição do problema pode ser: “Qual é a causa da queda nas vendas do restaurante?”

Definir um objetivo é essencial para garantir que o processo de análise seja direcionado e focado nos resultados desejados. A definição clara dos objetivos pode ajudar a estabelecer prioridades, selecionar as fontes de dados adequadas e os métodos de análise mais apropriados para o problema em questão.

A fim de resolver o problema de quedas nas vendas a equipe responsável pelo projeto optou por realizar uma análise de público que frequenta o restaurante.



Identificar o público-alvo é um aspecto fundamental para garantir o sucesso do negócio. Saber quem são os clientes potenciais, suas preferências e necessidades pode direcionar as estratégias de marketing, adequar o cardápio, criar um ambiente atraente e melhorar a experiência do cliente.

Criar uma experiência mais agradável para o cliente também é importante para aumentar a fidelização dos clientes. Se o público-alvo são famílias com crianças, por exemplo, o restaurante pode oferecer jogos e atividades infantis para tornar a visita mais agradável para todos.

Sendo assim, temos a seguinte definição:

Problema a ser tratado:	Queda nas vendas
Objetivo traçado:	Identificação do público-alvo

Coleta de dados.

A coleta de dados dos clientes em um restaurante pode ser feita de várias maneiras, mas é importante garantir que a privacidade dos clientes seja respeitada e que os dados coletados sejam relevantes. No caso do nosso restaurante, nosso objetivo é identificar o público-alvo portanto nossas duas principais variáveis são a idade e o sexo, fatores que podem influenciar em diversas decisões de consumo, como por exemplo, na escolha dos produtos, nos hábitos de consumo, no poder de compra e nas preferências de mídia e canais de comunicação.

Para que essa prática não seja invasiva ou inconveniente para os clientes, a equipe do restaurante optou por realizar a coleta de maneira mais discreta, a fim de preservar a privacidade e a satisfação dos clientes.

Para atingir esse objetivo, foi decidido que a coleta de dados seria feita por meio dos garçons, que têm contato direto com cada cliente do estabelecimento. Dessa forma, os garçons foram instruí-

dos a estimar a idade dos clientes, sem que isso interfira no atendimento ou na experiência do cliente no restaurante.

Os dados recolhidos durante o período de almoço, foram inseridos em uma planilha como podemos ver a seguir:

IDADE DOS CLIENTES			
HOMENS		MULHERES	
19	19	18	19
20	28	20	28
30	29	25	29
19	26	19	23
23	24	22	39
24	20	28	20
18	24	18	24
25	18	21	18
19	25	19	25
24	21	19	19
27	19	23	19
20	31	22	31
18	43	18	22
25	36	25	40
21	23	21	20

Como visto anteriormente, a qualidade dos dados é um fator crucial para uma análise precisa e confiável. Porém, no caso em questão, a coleta de dados foi realizada de maneira mais discreta, por meio da estimativa da idade pelos funcionários do restaurante. Embora essa abordagem possa gerar

dados menos precisos, ela se mostrou viável para o objetivo específico de identificar a faixa etária dos clientes que frequentam o estabelecimento.

Preparo dos dados

Em nossa última planilha, foram registrados as idades e o sexo de cada cliente. No entanto, a maneira como os dados foram dispostos não é suficiente para fornecer informações relevantes e identificar padrões significativos para nossa análise.

Devemos lembrar que os dados devem ser estruturados para que sejam mais fáceis de entender e analisar. Isso pode ser feito por meio da criação de tabelas, gráficos ou outras visualizações que ajudem a identificar padrões e tendências nos dados.

Técnicas e métodos

Uma técnica simples e útil que podemos utilizar é a média aritmética, ela fornece o valor numérico que representa o centro de um conjunto de dados. Isso permite ter uma noção da tendência central dos dados, indicando um valor que pode ser considerado “típico” ou “representativo” do conjunto. No contexto da análise do público-alvo, a média pode ser utilizada para identificar a faixa etária mais comum entre os clientes, por exemplo.

O cálculo é simples, realizamos a soma de todos os dados e os dividimos pelo total de dados que temos, como mostra a seguinte fórmula:
$$\bar{x} = \frac{\sum x_i}{n}$$

Realizando essa conta temos um valor de média igual a 23,94 anos para homens e 23,13 anos para mulheres.

Para saber se a média apresentou um valor válido, é importante analisar a dispersão dos



dados em relação à média. Isso pode ser feito utilizando o desvio padrão ou a variância, que indicam o quão dispersos os dados estão em relação à média. Quanto menor for o desvio padrão ou a variância, mais próximos os dados estão da média e mais confiável é o valor médio como representativo do conjunto de dados. Por outro lado, se a dispersão dos dados for muito grande, a média pode não ser um valor representativo do conjunto de dados e pode ser mais adequado utilizar outra medida de tendência central, como a mediana.

Tendo isso em mente podemos prosseguir para o cálculo do desvio padrão.

O desvio padrão é uma medida simples e fácil de entender. Ele pode fornecer uma maneira intuitiva de entender a variabilidade dos dados, facilitando a interpretação dos resultados da análise o que torna essa técnica, o que é bem útil para o nosso caso.

A fórmula do desvio padrão é dada por:

$$DP = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n}}$$

Utilizando o Excel para realizar os cálculos de maneira mais fácil e rápida em nosso conjunto de dados, obtemos o seguinte resultado:

HOMENS		
NUM.	IDADE	$(x_i - \bar{x})^2$
1	18	35,204
2	18	35,204
3	18	35,204
4	19	24,338
5	19	24,338
6	19	24,338
7	19	24,338
8	19	24,338
9	20	15,471
10	20	15,471

MULHERES		
NUM.	IDADE	$(x_i - \bar{x})^2$
1	18	35,204
2	18	35,204
3	18	35,204
4	18	35,204
5	19	24,338
6	19	24,338
7	19	24,338
8	19	24,338
9	19	24,338
10	19	24,338



11	20	15,471
12	21	8,604
13	21	8,604
14	23	0,871
15	23	0,871
16	24	0,004
17	24	0,004
18	24	0,004
19	24	0,004
20	25	1,138
21	25	1,138
22	25	1,138
23	26	4,271
24	27	9,404
25	28	16,538
26	29	25,671
27	30	36,804
28	31	49,938
29	36	145,604
30	43	363,538

11	20	15,471
12	20	15,471
13	20	15,471
14	21	8,604
15	21	8,604
16	22	3,738
17	22	3,738
18	22	3,738
19	23	0,871
20	23	0,871
21	24	0,004
22	25	1,138
23	25	1,138
24	25	1,138
25	28	16,538
26	28	16,538
27	29	25,671
28	31	49,938
29	39	227,004
30	40	258,138

Média masculina: 23,933

Média feminina: 23,133

Variância masculina: 31,596

Variância feminina: 32,022

Desvio Padrão masculino: 5,621

Desvio Padrão feminino: 5,659



Como melhorar esses resultados? Uma técnica muito utilizada é a exclusão dos outliers que nada mais são, que valores atípicos em um conjunto de dados, eles se diferenciam significativamente dos outros valores. Podem ser valores muito altos ou muito baixos em relação à média do conjunto de dados e podem afetar negativamente a análise estatística, especialmente quando se deseja encontrar tendências ou padrões nos dados. É importante identificar e lidar com outliers para garantir a qualidade da análise estatística e a precisão dos resultados.

Outliers podem ser identificados usando o desvio padrão, especificamente por meio da regra empírica. A regra empírica afirma que cerca de 68% dos dados em uma distribuição normal estão dentro de um desvio padrão da média, cerca de 95% dos dados estão dentro de dois desvios padrão da média e quase todos os dados (99,7%) estão dentro de três desvios padrão da média.

Portanto, qualquer valor que esteja a mais de três desvios padrão da média pode ser considerado um outlier. Essa é apenas uma regra geral, e a decisão final sobre o que é ou não um outlier deve levar em consideração o contexto e a natureza dos dados em questão. Em alguns casos, é possível que um valor que pareça ser um outlier na verdade seja uma observação válida e precisa.

Sabendo disso vamos refazer nossos cálculos sem os outliers em nosso conjunto de dados.

HOMENS		
NUM.	IDADE	$(x_i - \bar{x})^2$
1	18	35,204
2	18	35,204
3	18	35,204
4	19	24,338
5	19	24,338
6	19	24,338
7	19	24,338

MULHERES		
NUM.	IDADE	$(x_i - \bar{x})^2$
1	18	35,204
2	18	35,204
3	18	35,204
4	18	35,204
5	19	24,338
6	19	24,338
7	19	24,338

8	19	24,338
9	20	15,471
10	20	15,471
11	20	15,471
12	21	8,604
13	21	8,604
14	23	0,871
15	23	0,871
16	24	0,004
17	24	0,004
18	24	0,004
19	24	0,004
20	25	1,138
21	25	1,138
22	25	1,138
23	26	4,271
24	27	9,404
25	28	16,538
26	29	25,671
27	30	36,804
28	31	49,938
29	36	145,604

8	19	24,338
9	19	24,338
10	19	24,338
11	20	15,471
12	20	15,471
13	20	15,471
14	21	8,604
15	21	8,604
16	22	3,738
17	22	3,738
18	22	3,738
19	23	0,871
20	23	0,871
21	24	0,004
22	25	1,138
23	25	1,138
24	25	1,138
25	28	16,538
26	28	16,538
27	29	25,671
28	31	49,938

Média masculina: 23,276

Média feminina: 21,964

Variância masculina: 20,149

Variância feminina: 16,983



Desvio Padrão masculino: 4,489

Desvio Padrão feminino: 4,121

Conclusão do caso

O que podemos concluir com os valores apresentados? O valor de desvio padrão sem os outliers é de cerca de 4,3. Isso significa que, em média, a idade dos clientes do restaurante varia em torno de 4,3 anos em relação à idade média da amostra. Como os outliers foram removidos, esse valor de desvio padrão é mais representativo da dispersão dos dados em torno da média. Em outras palavras, essa medida indica o quão próximas ou distantes as idades estão da média, fornecendo uma informação importante para a análise do público-alvo do restaurante e para a tomada de decisões estratégicas.

Ao analisar nossos resultados e interpretá-los em termos de negócios, é possível observar que o público jovem é o que mais frequenta os restaurantes. Com base nessa informação, a equipe responsável pelo projeto pode apresentar soluções mais precisas e eficazes para os proprietários do restaurante, soluções como:

Direcionamento das estratégias de marketing: A equipe pode desenvolver campanhas publicitárias e promocionais voltadas especificamente para jovens, utilizando plataformas e canais que atinjam esse público com mais eficácia, como redes sociais, influenciadores digitais, entre outros.

Adequação do cardápio: Com base nas preferências alimentares dos jovens, o cardápio pode ser adaptado para incluir opções mais saudáveis e modernas, bem como novas bebidas e coquetéis para atrair esse público.



Criação de um ambiente atraente: Os jovens são atraídos por ambientes modernos e acolhedores, além de uma experiência de cliente única. Assim, melhorias no ambiente do restaurante, incluindo decoração, iluminação, música e atendimento, podem aumentar o apelo do estabelecimento para o público jovem.

Aumento da fidelização dos clientes: Promoções como desconto para universitários ou filia- dos de um clube próximo da região, promoções de happy hour e descontos especiais em dias especí- ficos da semana podem atrair ainda mais os jovens e incentivar a fidelização desse público.

Finalizando nosso projeto que busca soluções para a queda nas vendas do restaurante, lem- brando que a escolha do plano de ação pode variar, algumas soluções necessitam de um maior capital investido, outras podem apresentar resultados somente a longo prazo, por isso a escolha deve ser tomada com cautela, pensando em todos intemperes que podem vir a ocorrer e sempre buscando a evolução constante, procurando novas soluções como analisar conjuntos diferentes de dados como por exemplo, a quantidade de dinheiro que o público jovem disponibiliza para gastar com a alimenta- ção, pode ser uma análise útil para ajustar os preços nos cardápio, o que faz da análise de dados um processo crucial para melhoria continua e a obtenção de soluções mais eficazes e duradouras.

Resultados e Discussão

Um estudo de caso publicado em 2016 sobre o uso do Data Science em uma pequena empre- sa é a Nuts.com, uma empresa familiar de venda online de frutas secas, nozes e outros produtos ali- mentícios. A empresa utilizou técnicas de Data Science para otimizar seus processos de embalagem e envio de produtos, melhorar a satisfação do cliente e aumentar as vendas.

A Nuts.com coletou dados de seus processos de embalagem e envio, incluindo tempos de processamento, volumes de pedidos, rotas de envio, entre outros. Com esses dados, a empresa utilizou técnicas de análise de dados para identificar gargalos no processo de envio e atrasos nas entregas.

A partir dessas análises, a empresa implementou mudanças em seus processos de embalagem e envio, como a adoção de rotas mais eficientes e a otimização do processo de empacotamento dos produtos. Além disso, a empresa utilizou técnicas de modelagem de dados para prever a demanda futura dos produtos e adaptar seus processos de produção e envio de acordo.

Com essas mudanças, a Nuts.com conseguiu reduzir seus tempos de processamento e atrasos nas entregas, melhorando a satisfação do cliente e aumentando as vendas. Além disso, a empresa também utilizou técnicas de mineração de dados para entender melhor o comportamento de compra dos clientes e desenvolver estratégias de marketing personalizadas para aumentar ainda mais as vendas.

Esse é um exemplo real de como uma pequena empresa, como a Nuts.com, pode utilizar técnicas de Data Science para otimizar seus processos e melhorar a satisfação do cliente, além de aumentar as vendas e a competitividade no mercado permitindo à empresa expandir seus negócios e se tornar uma líder de mercado no segmento de nozes e lanches saudáveis online.

Considerações Finais

Podemos concluir que a prática do Data Science pode ser descrita como uma combinação de métodos teóricos de coleta, análise e resultados, porém no mundo real, com problemas reais, temos que inovar, adaptar e saber utilizar as ferramentas a nosso favor. Raramente surgirão problemas que serão resolvidos com soluções básicas.



As pequenas empresas enfrentam muitos desafios em um mercado competitivo, e é essencial que elas utilizem as ferramentas disponíveis para se manterem competitivas. O Data Science é uma dessas ferramentas, por isso como aprendemos no decorrer deste artigo, devemos decompor nosso problema em subtarefas, isso torna nosso grande objetivo mais fácil de ser alcançado. Algumas tarefas serão mais fáceis de resolver, podendo até ser utilizadas ferramentas já existentes, porém em alguns momentos a solução pode parecer distante e essas ferramentas podem não ser por si só suficientes para chegarmos a um bom resultado e nesses momentos o pequeno empreendedor pode se destacar, com o que talvez seja a principal qualidade no empreendedorismo brasileiro, a criatividade.

Em resumo, o uso do Data Science pode oferecer muitas vantagens para pequenas empresas, desde a identificação de oportunidades de mercado até a melhoria da eficiência operacional e da experiência do cliente. Com o acesso crescente a ferramentas e recursos de Data Science, as pequenas empresas podem tirar proveito dessas vantagens para se manterem competitivas em um mercado em constante mudança.

Referências Bibliográficas

COELHO, Lucas. Ciência de Dados: O que é, Conceito e Definição. Disponível em: <https://www.cetax.com.br/blog/data-science-ou-ciencia-de-dados/>. Acesso em: 28 abr. 2023.

DHAR, Vasant. Data science and prediction. *Communications of the ACM*, v. 56, n. 12, p. 64-73, 2013.

FURLAN, Marco R.; FILHO, Educaro V. G. Uma proposta de aplicação de business intelligence no chão-de-fábrica. *Gestão & Produção*, v. 12, n. 1, p. 55-66, jan.-abr. 2005.

KELLEHER, John D.; TIERNEY, Brendan. Data Science: An Introduction. 2018.

MANYIKA, J. et al. Big data: The next frontier for innovation, competition, and productivity. McKinsey & Company: McKinsey Global Institute, 2011.

PROVOST, Foster; FAWCETT, Tom. Data Science for Business: What you need to know about data mining and data-analytic thinking. “ O’Reilly Media, Inc.”, 2013.

SEMPLE, Kirk. Data Science Is Changing How Nuts.com Runs Its Business. Forbes, 20 jun. 2016.

YIN, Robert. Case Study Research: design and methods . 5 ed. Thousand Oaks, CA: Sage , 2014.

