Estudo de caso 3: Siga seu próprio caminho para o estudo de caso

# Introdução

Bem-vindo ao seu próprio estudo de caso! Este é diferente de todos os outros casos deste curso, pois lhe dá a oportunidade de escolher uma questão de negócios que seja de seu interesse particular. Como resultado, encontrar conjuntos de dados e realizar análises será mais complexo e exigirá a tomada de muitas decisões por conta própria.

Ao final desta lição, seu objetivo é ter um estudo de caso pronto para um portfólio. Para ter certeza de que você está preparado para o sucesso, é recomendado que conclua primeiro um dos estudos de caso mais estruturados. Além disso, leia o cenário incluído e siga o roteiro do estudo de caso, que detalha as etapas do processo de análise de dados: **perguntar**, **preparar**, **processar**, **analisar**, **compartilhar** e **agir**.

Quando seu estudo de caso estiver concluído, faça o download do pacote e consulte os detalhes do seu trabalho a qualquer momento. Então, quando iniciar sua busca por emprego, seu estudo de caso será uma maneira tangível de demonstrar seus conhecimentos e habilidades para potenciais empregadores.

# Cenário

Você é um analista de dados júnior que trabalha para um consultor de inteligência de negócios. Considerando que está desempenhando sua função há seis meses, seu chefe sente que você está pronto para ter mais responsabilidades. Ele solicitou que você lidere um projeto para um novo cliente – o trabalho envolverá vários aspectos, desde a definição da tarefa de negócios até a apresentação de suas recomendações baseadas em dados. Você deverá escolher o tópico, fazer as perguntas certas, identificar um novo conjunto de dados e garantir sua integridade, conduzir análises, criar visualizações de dados atrativas e preparar uma apresentação.

# Perguntar

Cinco perguntas guiarão seu estudo de caso:

1. Que tipo de empresa seu cliente representa e o que estão pedindo para você realizar?
2. Quais são os principais fatores envolvidos na tarefa de negócios que você está investigando?
3. Que tipo de dados será adequado para sua análise?
4. Onde você obterá esses dados?
5. Quem é seu público-alvo e quais materiais o ajudarão a apresentá-los de maneira eficaz?

Você criará um relatório com as entregas a seguir:

1. Uma declaração clara da tarefa de negócios que você selecionou para investigar
2. Uma descrição de todas as fontes de dados usadas
3. Documentação de qualquer limpeza ou manipulação de dados
4. Um resumo da sua análise
5. Como justificar visualizações e descobertas-chave
6. Com base nas suas descobertas, uma lista de entregas adicionais que você considera útil incluir para exploração adicional
7. Seus principais *insights* chave com base em sua análise

Observação: Concluir este estudo de caso em uma semana é uma meta adequada.

Para começar, disponibilize tempo para discutir algumas ideias. Uma apresentação eficaz deve durar cerca de 30 minutos, portanto, certifique-se de que sua tarefa de negócios não seja muito simples ou muito complexa. Abaixo estão três exemplos de tarefas de negócios que podem ser usadas como inspiração.

* **Classificar as equipes esportivas de uma universidade**: Imagine que uma empresa queira fazer recomendações para jogadores de vários times esportivos. Explore os dados dos últimos anos, considere quais equipes estão classificadas entre as cinco primeiras no início de sua temporada e como as classificações mudam ao longo do tempo. Quais equipes demonstram potencial para apoios?
* **Expandir o mercado de aluguel por temporada**: Imagine que uma empresa de administração de imóveis deseja determinar se deve expandir para uma nova área geográfica. Use dados públicos de Airbnb para responder a perguntas sobre o mercado de aluguel por temporada em uma área específica. Investigue como bairros ou comodidades influenciam os preços de Airbnb, quais anúncios são alugados com mais frequência e onde os superanfitriões estão localizados.
* **Considerar os padrões de compra do consumidor e os padrões climáticos:** O clima pode influenciar o interesse do consumidor em uma variedade de produtos. O clima mais frio aumenta a demanda por lenços e sopas? O clima mais quente impulsiona as vendas de ventiladores e chinelos? E o clima mais agressivo aumenta o consumo de água engarrafada e baterias? Use dados públicos da Administração Oceânica e Atmosférica Nacional para descobrir padrões de compra relacionados ao clima.

Observação: Esses exemplos não têm uma conexão imediata com os conjuntos de dados para não influenciar sua escolha de tarefas de negócios com base nos dados.

Use o Roteiro de estudo de caso a seguir como um guia:

|  |  |  |
| --- | --- | --- |
| **Roteiro do estudo de caso - Perguntar** | | |
| **Perguntas de orientação** |  |  |
| * Qual tema você está explorando? |  |  |
| * Qual é o problema que você está tentando resolver? |  |  |
| * Quais métricas você usará para medir seus dados a fim de alcançar | seu | objetivo? |
| * Quem são as partes interessadas? |  |  |
| * Quem é seu público-alvo? |  |  |
| * Como seus *insights* podem ajudar seu cliente a tomar decisões? |  |  |
| **Principais tarefas** |  |  |
| 1. Identificar a tarefa de negócios |  |  |
| 2. Determinar as principais partes interessadas |  |  |
| 3. Escolher um conjunto de dados |  |  |

|  |
| --- |
| 4. Estabelecer métricas |
| **Entrega**  Uma declaração clara da tarefa de negócios que você selecionou para investigar |

# Preparar

Encontre dados adequados para sua análise em qualquer conjunto de dados confiável. Existem vários conjuntos de dados públicos que podem ser usados. Você já encontrou alguns deles ao longo deste programa:

* [World Happiness Report](https://www.kaggle.com/unsdsn/world-happiness) (disponibilizado pela Sustainable Development Solutions Network sob uma [licença CC0](https://creativecommons.org/publicdomain/zero/1.0/)): Dados regionais sobre satisfação geral que contêm informações interessantes sobre a relação entre felicidade, dinheiro, saúde e muitas outras métricas.
* [Preços de abacate](https://www.kaggle.com/neuromusic/avocado-prices) (disponibilizado por [Justin Kiggins](https://www.kaggle.com/neuromusic) sob uma [licença CC0](https://creativecommons.org/publicdomain/zero/1.0/)): Dados históricos sobre preços e vendas de abacate em supermercados nos Estados Unidos.
* [Conjunto de dados de filmes](https://www.kaggle.com/rounakbanik/the-movies-dataset) (disponibilizado por [Rounak Banik](https://www.kaggle.com/rounakbanik) sob uma [licença CC0](https://creativecommons.org/publicdomain/zero/1.0/)): Metadados sobre 45.000 filmes, com pontos de dados incluindo elenco, equipe, palavras-chave da trama, orçamento, receita, classificações, datas de lançamento, idiomas, empresas de produção e muito mais.
* [Os 50 livros mais vendidos da Amazon](https://www.kaggle.com/sootersaalu/amazon-top-50-bestselling-books-2009-2019) (disponibilizado por [Souter Saalu](https://www.kaggle.com/sootersaalu) sob uma [licença CC0](https://creativecommons.org/publicdomain/zero/1.0/)): Dados sobre os livros mais vendidos da Amazon de 2009 a 2019, categorizados por ficção e não ficção.

Depois de decidir sobre um conjunto de dados, reúna e armazene-o adequadamente, limpe os dados e certifique-se de que sejam confiáveis e livres de erros. Documente seu processo, pois os clientes geralmente pedem para ver dados brutos e limpos. Agora, prepare seus dados para análise usando o seguinte roteiro do estudo de caso como guia:

|  |
| --- |
| **Roteiro do estudo de caso - Preparar** |
| **Perguntas de orientação**   * Onde seus dados estão localizados? * Como os dados são organizados? |

|  |
| --- |
| * Existem problemas com viés ou credibilidade nesses dados? [Seus dados são confiáveis, originais, abrangentes, atuais e incluem a fonte?](https://www.coursera.org/learn/data-preparation/lecture/lHirM/what-is-bad-data) * Como você está lidando com o licenciamento, a privacidade, a segurança e a acessibilidade? * Como você verificou a integridade dos dados? * Como isso o ajuda a responder à sua pergunta? * Há algum problema com os dados? |
| **Chave**   * Fazer o download dos dados e armazená-los adequadamente. * Identificar como estão organizados. * Classificar e filtrar os dados. * Determinar a credibilidade dos dados. |
| **Entrega**  Uma descrição de todas as fontes de dados usadas |

# Processar

Em seguida, processe seus dados para análise usando o seguinte roteiro do estudo de caso como guia:

|  |
| --- |
| **Roteiro do estudo de caso - Processar** |
| **Perguntas de orientação**   * Quais ferramentas você está escolhendo e por quê? * Você garantiu a integridade dos seus dados? * Que medidas foram tomadas para garantir que seus dados estejam limpos? * Como você pode verificar se seus dados estão limpos e prontos para análise? * Você documentou seu processo de limpeza para poder revisar e compartilhar esses resultados? |
| **Principais tarefas**   * Verificar os dados em busca de erros. |

|  |
| --- |
| * Escolher suas ferramentas. * Transformar os dados para que você possa trabalhar com eles de forma eficaz. * Documentar o processo de limpeza. |
| **Entrega**  Documentação de qualquer limpeza ou manipulação de dados |

## Siga as etapas a seguir:

1. Faça o download do conjunto de dados.
2. Crie uma pasta em sua área de trabalho ou no Drive para armazenar os arquivos. Use as devidas convenções de nomenclatura
3. de arquivo. Crie subpastas para arquivos .CSV ou .XLS.
4. Siga estas instruções para Excel (a) ou Planilhas Google (b):
   1. Inicie o Excel, abra seu arquivo e escolha Salvar como um arquivo de pasta de trabalho do Excel. Coloque-o na subpasta criada para arquivos .XLS.
   2. Abra cada arquivo .CSV no Planilhas Google e salve-o na subpasta
5. correta. Abra sua planilha e rotule adequadamente as colunas.
6. Prossiga para a etapa de análise.

Se desejar, continue trabalhando com os dados para se familiarizar melhor e talvez até identificar novas abordagens para responder às perguntas de negócios.

# Analisar

Agora que seus dados estão armazenados adequadamente e preparados para análise, comece a colocá-los em prática. Use o Roteiro de estudo de caso a seguir como um guia:

|  |
| --- |
| **Roteiro do estudo de caso - Analisar** |
| **Perguntas de orientação** |

|  |
| --- |
| * Como você deve organizar seus dados para realizar análises sobre eles? * Seus dados foram formatados corretamente? * Que surpresas você descobriu nos dados? * Que tendências ou relações você encontrou nos dados? * Como esses *insights* o ajudarão a responder às suas perguntas de negócios? |
| **Principais tarefas**   * Agregar seus dados para que sejam úteis e acessíveis. * Organizar e formatar seus dados. * Realizar cálculos. * Documentar seus cálculos para acompanhar suas etapas de análise. * Identificar tendências e relações. |
| **Entrega**  Um resumo da sua análise |

## Siga estas etapas se usar planilhas

Abra seu aplicativo de planilhas e conclua as etapas a seguir:

1. Quando relevante, uniformize as colunas e combine-as em uma única planilha.
2. Limpe e transforme seus dados para se preparar para uma análise.
3. Faça uma análise descritiva.
4. Execute alguns cálculos para ter uma noção melhor do layout dos dados.
5. Crie uma tabela dinâmica para calcular e visualizar os dados rapidamente.
6. Quando estiver trabalhando com várias planilhas individuais, mescle-as usando a ferramenta escolhida para realizar sua análise final, seja uma planilha, um banco de dados e SQL ou RStudio.
7. Exporte um arquivo de resumo para análise posterior.

## Siga estas etapas para usar SQL

Abra a ferramenta SQL de sua escolha e conclua as etapas a seguir:

1. Importe seus dados.
2. Explore seus dados, talvez observando o número total de linhas, valores distintos, valores máximos, mínimos ou médios.
3. Quando relevante, use instruções JOIN para combinar seus dados relevantes em uma tabela.
4. Crie algumas estatísticas resumidas.
5. Investigue tendências interessantes e salve essas informações em uma tabela.

## Siga estas etapas para usar R

Abra o RStudio e [siga este script](https://docs.google.com/document/d/1gUs7-pu4iCHH3PTtkC1pMvHfmyQGu0hQBG5wvZOzZkA/copy) para se inspirar. Observação: O roteiro é de outro estudo de caso, mas fornece instruções claras que podem nortear sua própria análise. Em seguida, conclua as etapas a seguir:

1. Importe seus dados.
2. Uniformize as colunas e mescle-as em um único data frame.
3. Limpe e adicione dados para se preparar para a análise.
4. Faça uma análise descritiva.
5. Exporte um arquivo de resumo para análise posterior.

# Compartilhar

Agora que você realizou sua análise e obteve *insights* sobre seus dados, crie visualizações para compartilhar suas descobertas. Você apresentará ao seu cliente e outras partes interessadas, portanto, os visuais devem ser sofisticados e polidos para comunicar seus *insights* efetivamente. Use o Roteiro de estudo de caso a seguir como um guia:

## Roteiro do estudo de caso - Compartilhar perguntas de orientação

* + Você conseguiu responder à pergunta de negócios?
  + Que história seus dados contam?
  + Como suas descobertas se relacionam com sua pergunta original?
  + Quem é seu público-alvo? Qual é a melhor maneira de se comunicar com esse público?
  + A visualização de dados pode ajudá-lo a compartilhar suas descobertas?

|  |
| --- |
| * Sua apresentação é acessível ao seu público-alvo? |
| **Principais tarefas**   * Determinar a melhor maneira de compartilhar suas descobertas. * Criar visualizações de dados eficazes. * Apresentar suas descobertas. * Garantir que seu trabalho seja acessível. |
| **Entrega**  Como justificar visualizações e descobertas-chave |

## Siga as etapas a seguir:

1. Pegue um papel e uma caneta e esboce algumas ideias de como você visualizará os dados.
2. Depois de escolher um formato visual, abra a ferramenta de sua preferência para criar sua visualização. Use um software de apresentação, como PowerPoint ou Google Slides; seu programa de planilha; Tableau; ou R.
3. Crie sua visualização de dados, lembrando que o contraste deve ser usado para chamar a atenção do seu público-alvo para os *insights* mais importantes. Use princípios artísticos, incluindo tamanho, cor e formato.
4. Garanta um significado claro por meio do uso adequado de elementos comuns, como títulos, subtítulos e rótulos.
5. Filtre sua visualização de dados aplicando profunda atenção aos detalhes.

# Agir

Agora que terminou de criar suas visualizações, aja de acordo com suas descobertas. Organize as entregas que você criou, incluindo seus principais *insights* de alto nível com base em sua análise. Use o Roteiro de estudo de caso a seguir como um guia:

|  |
| --- |
| **Roteiro do estudo de caso - Agir** |
| **Perguntas de orientação** |
| * Qual é a sua conclusão final com base na sua análise? |

|  |
| --- |
| * Como sua equipe e sua empresa podem aplicar seus *insights*? * Que próximos passos você ou suas partes interessadas podem seguir com base em suas descobertas? * Existem dados adicionais que você poderia usar para expandir suas descobertas? |
| **Principais tarefas**   * Criar seu portfólio. * Adicionar seu estudo de caso. * Praticar a apresentação do seu estudo de caso a um amigo ou membro da família. |
| **Entrega**  Seus principais *insights* chave com base em sua análise  Com base nas suas descobertas, uma lista de entregas adicionais que você considera útil incluir para exploração adicional |

# Finalização

Parabéns por concluir seu próprio estudo de caso! Se desejar, conclua um dos outros estudos de caso para continuar ampliando seu portfólio. Boa sorte na sua busca por emprego!