

INTRODUÇÃO AO DATA SCIENCE

Descritivo da **atividade prática extra** do curso

SUMÁRIO

1. OBSERVAÇÃO	1
2. ATIVIDADE PRÁTICA	1
2.1. INTRODUÇÃO	1
2.2. SOLICITAÇÃO	2
3. ENTREGA	2
4. AVALIAÇÃO	3

1. OBSERVAÇÃO

Esta atividade prática extra é opcional e não faz parte da certificação do curso **Introdução ao Data Science (IP 20h A)**. Caso o estudante tenha como objetivo apenas obter o certificado de conclusão do curso, não é necessário realizar esta atividade prática extra, bastando concluir normalmente a atividade avaliativa final.

A atividade prática extra possui dois objetivos principais:

1. Proporcionar ao estudante uma oportunidade adicional de praticar e consolidar os conhecimentos adquiridos ao longo do curso;
2. Possibilitar a participação em um processo de seleção para os cursos de qualificação profissional síncronos de 300 horas, como próxima etapa da 'Carreira Tech', do programa SCTEC.

Importante: caso o(a) estudante opte por realizar esta atividade prática extra, ela **deverá ser concluída antes da realização da atividade avaliativa final do curso**, uma vez que, após o envio da avaliação final, não será mais possível submeter a atividade prática extra.

2. ATIVIDADE PRÁTICA

2.1. INTRODUÇÃO

Santa Catarina destaca-se como um dos pólos nacionais de crescimento nas áreas de Tecnologia da Informação e Ciência de Dados, o que amplia a demanda por profissionais capazes de analisar informações de forma estruturada e objetiva. Nesse contexto, a análise de dados torna-se uma competência essencial para a compreensão de fenômenos sociais, econômicos e operacionais, bem como para a tomada de decisões baseadas em evidências.

Assim, esta atividade prática tem como finalidade estimular o estudante a aplicar conceitos fundamentais de Data Science por meio da análise de um conjunto de dados real, utilizando técnicas de exploração e interpretação de informações como base para a geração de conhecimento a partir dos dados.

2.2. SOLICITAÇÃO

O estudante deverá utilizar a base de dados pública do Titanic, em formato CSV, com o objetivo de realizar uma Análise Exploratória de Dados (AED). A atividade consiste em importar, organizar e analisar o conjunto de dados, buscando compreender o comportamento geral das informações e identificar padrões, relações entre variáveis e possíveis fatores associados à sobrevivência dos passageiros. A partir desse processo, espera-se a obtenção de insights relevantes, como estatísticas descritivas, distribuições, comparações entre grupos e análises exploratórias das variáveis disponíveis.

A base de dados pública do Titanic, em formato CSV, está disponível neste endereço eletrônico: https://drive.google.com/file/d/11HptTxJbUMRG16xpC39fcliba_-Z_J9d/

3. ENTREGA

A entrega deverá ser realizada **dentro do período de vigência do curso**, conforme datas estabelecidas no cronograma oficial, por meio de um **único arquivo compactado**, nos formatos .zip ou .rar, com **tamanho máximo de 20 MB**, contendo todos os arquivos necessários para o funcionamento do projeto, incluindo: código-fonte, conjunto de dados utilizado, insights extraídos e arquivo de documentação.

O projeto deverá ser desenvolvido de forma estruturada, contemplando as etapas de importação dos dados, compreensão do conjunto de informações, preparação e tratamento dos dados, incluindo a verificação e o tratamento de valores nulos ou duplicados, ajustes de tipagem e organização das colunas.

A análise deverá envolver a compreensão das variáveis, a aplicação de filtros, ordenações e agrupamentos (GroupBy), bem como a construção de pelo menos uma visualização gráfica que represente a distribuição dos dados ou a relação entre variáveis relevantes, utilizando a linguagem Python e bibliotecas adequadas ao contexto da análise exploratória.

Todas as informações do projeto deverão estar claramente documentadas em um arquivo de texto, que poderá estar nos formatos .txt, .doc, .docx, .pdf ou README.md, e que deverá ser incluído dentro da pasta compactada juntamente com o código da aplicação. A documentação deverá apresentar a descrição das etapas realizadas, as principais decisões tomadas durante o tratamento dos dados e os principais insights obtidos a partir da análise

exploratória. Caso o aluno deseje, o código também poderá ser versionado no GitHub, desde que o arquivo compactado submetido na tarefa contenha todos os arquivos necessários para a avaliação.

4. AVALIAÇÃO

A avaliação do desafio será realizada pelo SENAI/SC LAB365 em momento posterior à entrega da solução, seguindo os critérios presentes na tabela abaixo.

Para a correção, o score terá uma variação de 0 (zero) a 100 (cem) como nota mínima e máxima, respectivamente. A distribuição do peso da avaliação prática extra se dará da seguinte forma:

- Documentação: 30% do score final
- Linguagem de programação: 30% do score final
- Codificação: 40% do score final

Critérios de Avaliação da Atividade Prática Extra				
Nº	Critério	50 pontos	70 pontos	100 pontos
1	Quantidade mínima de caracteres na documentação do projeto. Peso: 30%	A documentação do projeto contém menos do que 400 caracteres.	A documentação do projeto contém entre 400 e 800 caracteres.	A documentação do projeto contém entre mais de 800 caracteres.
Nº	Critério	70 pontos	100 pontos	
2	Linguagens e arquivos utilizados no projeto. Peso: 30%	Foram utilizados Python e Fonte de Dados durante a realização do projeto.	Foram utilizados Python, Fonte de Dados e Visualizações Geradas durante a realização do projeto.	
Nº	Critério	0 pontos	100 pontos	
3	Desenvolvimento e codificação do projeto. Peso: 40%	O arquivo submetido não contém um projeto compatível com a proposta do desafio, está corrompido ou não pode ser executado devido a erros de desenvolvimento.	O arquivo submetido contém um projeto válido, desenvolvido na linguagem solicitada, que pode ser executado normalmente, sem erros críticos de execução.	