

Bootcamp DiversiData Tech Pan

Trabalho Prático

Módulo 1	Fundamentos do Big Data
----------	-------------------------

Objetivos

Exercitar os seguintes conceitos trabalhados no Módulo:

- Coleta de dados;
- Analisar e tratamento de dados;
- Criar visualização de dados;
- Implementar algoritmo de *Machine Learning*;
- Analisar resultados obtidos;
- Conhecimento teórico ministrado nas videoaulas.

Enunciado

Uma empresa do ramo imobiliário quer criar um modelo preditivo que utiliza os dados de valor do custo da diária e quantidade de pessoas. Ela possui uma base de dados histórica que contém a quantidade pessoas e o valor cobrado no custo da diária.

Diante disso, a empresa contratou uma consultoria de tecnologia para auxiliar no processo de predição. Os analistas perceberam que a melhor abordagem para esse problema é criar um algoritmo de *Machine Learning* de regressão

linear para prever o valor da diária do imóvel baseado na quantidade de pessoas que realizaram a diária.

Atividades

Para esta atividade, os alunos deverão criar um algoritmo de regressão linear para prever o valor da diária de um imóvel baseado no número de pessoas.

1. Criar um projeto no google drive;
2. Coletar e inserir o arquivo custo_diaria.csv na plataforma;
3. Analisar os dados coletados;
4. Avaliar a relação entre as variáveis;
5. Criar algoritmo de regressão linear;
6. Responder as questões teóricas e práticas do trabalho.

Dicas do professor:

1. Analisem com cuidado os dados através da representação gráfica;
2. Analisem bem o gráfico gerado e a disponibilização dos dados;
3. Antes de enviar as respostas verifiquem se o gabarito esteja correto;
4. Tenham atenção no que pede cada questão;
5. Os dados disponibilizados no *dataset* são fictícios. Ou seja, não tem relação com o mundo real.
6. O *dataset* utilizado no trabalho podem ser obtidos no link:

<https://github.com/ProfLeandroLessa/TP-M1-ABD>

Bom trabalho prático a todos!