

Trabalho Final – Programando IA com Python

1. Leia o arquivo lemonades.csv usando pandas
2. retorne as informações sobre as colunas e as principais estatísticas sobre o dataset.
3. Realize a limpeza do dataset:
 - a. verifique se existe valores nulos
 - b. verifique e remova, se existir, linhas duplicadas
 - c. trate os elementos faltantes (interpolação (data) e media)
4. Adicione uma coluna chamada "Sales" que contém o total de vendas de limão e laranja
5. Adicione uma coluna chamada "Revenue" que calcula o lucro (venda*preço)
6. Escreva uma função que retorne o lucro total
7. Escreva uma função que receba dois parâmetros, dataset e temp(int). Se temp for 1, ele retorna a máxima temperatura observada no conjunto de dados; se 0, retorna a média do período observado; se -1, retorna a temperatura mínima.
8. Escreva uma função que receba dois parâmetros (dataset, localização) e retorne o dataset com o preço do limão e laranja ajustados em 15% se a localização for 'Park' ou ajustados em 10% se a localização for 'Beach'.
9. Escreva uma função que receba o dataset como parâmetro e retorne o dataset com uma coluna a mais que mostre o rank de temperatura. O dataset retornado precisa estar ordenado de acordo com rank (descendente)
10. Imprima um gráfico de linha que retorne o lucro ao longo do tempo
11. Crie um scatter-plot leaflets x sales. Existe alguma correlação entre os dados? (Use a correlação de Pearson e explique o significado do valor obtido por ela)
12. Crie um histograma com 10 bins que mostre o lucro (revenue).

13. Crie um gráfico de linha que mostre o lucro por tipo de fruta ao longo do tempo.
14. Ajuste uma Regressão Linear aos dados para prever vendas usando cada uma das features abaixo:
 - a. Temperature
 - b. Leaflets
 - c. Price
 - d. Retorne um gráfico com a reta ajustada aos dados para cada feature e diga se ela é uma variável boa ou não para prever vendas.
15. Ajuste uma Regressão Linear aos dados para prever vendas usando as três features juntas. Comente o resultado obtido.
16. Generalize o algoritmo de Regressão Linear implementada em sala para que ele possa ser usado para regressão multivariada.
17. Aplique o algoritmo implementado na questão 16 no conjunto de dados `consumo_cerveja.csv` usando as mesmas features para treinamento.
 - a. Compare os resultados que você obteve com os resultados obtidos através da Scikit Learning.
 - b. Prediga o consumo usando seu modelo e o modelo da sklearn.
 - c. Calcule o MSE de ambos