**Python Insights**

Em projetos de dados normalmente é usado o tipo de arquivo ipynd.

Tratamento de dados explicado de uma forma mais simples é resolver o problema de uma base de dados, corrigir os erros.

A biblioteca pandas é a que permite trabalhar com base de dados, ela é a principal ferramenta do Python para trabalhar com base de dados.

O Plotly é uma ferramenta do Python que ajuda a criar gráficos dinâmicos.

Toda biblioteca no Python se quiser pode se dar um apelido pra ela, por exemplo o panda pode ser apelidado como pd.

O comando display é a mesma coisa que o print porém a forma como ele mostra a tabela é muito mais organizada. O display só existe nos arquivos ipynb.

Os três principais problemas que tem que corrigir em qualquer base de dados são informações inúteis, informações no formato errado e informações vazias.

Dataframe nada mais é do que uma tabela no Python.

Linha no Python se chama index.

O método dropna exclui as linhas com informações vazias na tabela.

Esse código vai contar os valores dentro da coluna cancelou. Colocar o display e parênteses na frente de tabela.

Nomalize é um termo que a galera de estatística usa que significa calcular a proporção, ou seja, colocar o valor numa escala de 0 á 1, nada mais é do que calcular o percentual.

Para usar gráfico no Python primeiro tem que criar o gráfico e depois exibir o gráfico.

Não existe limite de tamanho para arquivo que está sendo manipulado, o limite é a capacidade do computador.