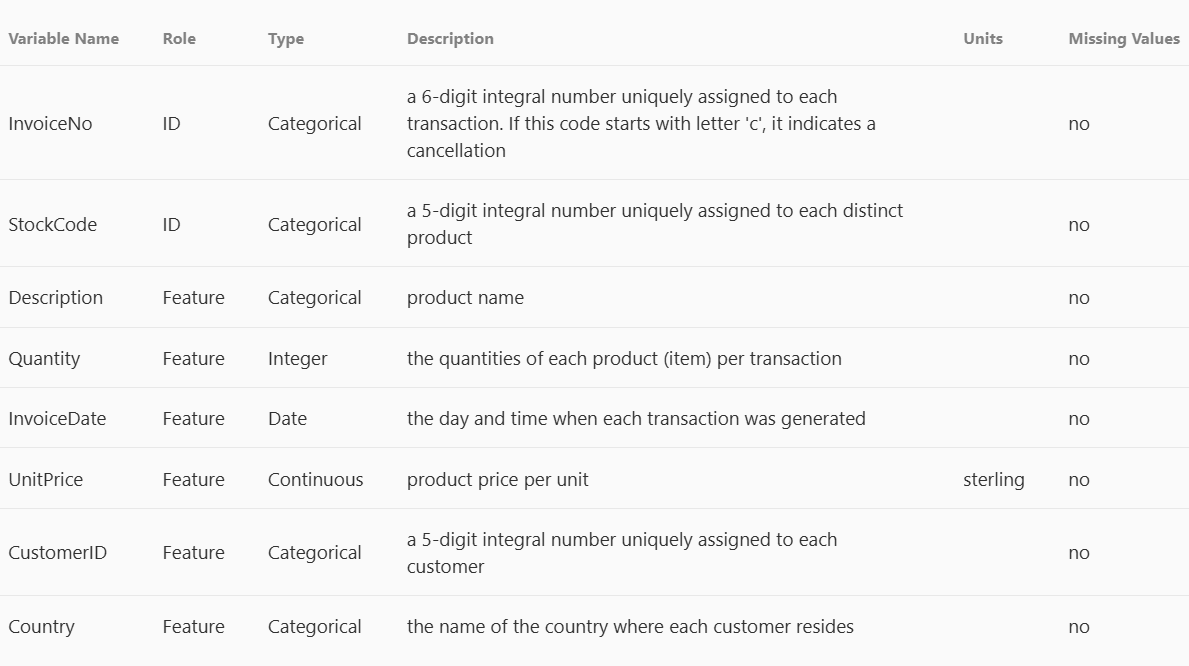
1 - Quais são as colunas ou variáveis do dataset? (Ex.: ID do cliente, categoria do produto, quantidade vendida, preço, data da venda)



2 - Existem dados faltantes? Em quais colunas? Como eles podem afetar a análise?

Sim, a database apresenta inconsistência em colunas como descrição, valor unitário, customerID e quantidade (negativa). O que impede uma análise fiel e previsões de futuras vendas, já que não há um histórico aprofundado.

3 - Quais variáveis parecem ter maior impacto para prever o comportamento de compra? (Ex.: frequência de compra, valor gasto, categorias preferidas)

País, preço unitário, quantidade e descrição

4 - Existem colunas com dados categóricos que precisem ser convertidos para análise (ex.: sexo, região, método de pagamento)?

País

5 - Quais tipos de variáveis temos? Numéricas, categóricas, temporais?

Númericas, categóricas e temporais.

6 - Quem são os clientes que mais compram e com que frequência?

Os três clientes que mais compram são (quantidade e id respectivamente):

|  |  |
| --- | --- |
| 80995 | 16446 |
| 74215 | 12346 |
| 12540 | 13256 |

16446 comprou 4 vezes, 2 em maio e 2 em dezembro.

12346 comprou 2 vezes em janeiro.

12346 comprou 1 vez em novembro.

Todos são da Inglaterra.

7 - Quais produtos costumam ser adquiridos juntos, formando combinações frequentes?