Descrição do trabalho:

Você deve gerar 3 arquivos:

1. trabalho.docx com:

* fonte dos dados (endereço completo de onde a base está)
* descrição da base de dados, isto é, o que representam as colunas
* Perguntas a serem respondidas – Todas devem ser numeradas

1. templateTrabalho.py com:

* Para cada pergunta do docx

print("\n-----------------------------------------------------")

print("\n *número e texto da questão:* \n")

print("\n-----------------------------------------------------")

1. gabaritoTrabalho.py com:

* Para cada pergunta do docx

print("\n-----------------------------------------------------")

print("\n *número e texto da questão*: \n")

print( *sua resposta para a questão*)

print("\n-----------------------------------------------------")

Na sua proposta deve constar, no mínimo:

1. Uma concatenação ( em qualquer eixo) (FEITO)
2. Duas estratégias distintas de preenchimento de valores ausentes

Por ex: substituir por 0 e substituir pelo valor médio do agrupamento (FEITO)

1. Uma estratégia para substituição de valores

Por exemplo, df[col].replace ( {1:'UM',2:’DOIS'}) (FEITO)

1. Criar Categorias em função do valor de uma coluna ( pd.cut) (FEITO)
2. Três Filtros (FEITO)
3. Duas Tabelas de Frequência (FEITO)
4. Dois gráficos (FEITO)
5. Medidas de Sumarização: (FEITO)
   1. Geral (FEITO)

Grupos simples (uma coluna/series/categoria) (FEITO)

* 1. Grupos estruturados ( mais de uma coluna/series/categoria) (FEITO)

1. Quatro cruzamentos de colunas (FEITO)
   1. simples (1 X 1) (FEITO)
   2. estruturados ( nXm) (FEITO)