

PREPARAÇÃO DOS DADOS

#1) carregar a base de dados: base_carros_usados.csv

#2) exibir a estrutura das variáveis (colunas).

#3) alterar a descrição da coluna (renomear as colunas) para o idioma português brasileiro

#não utilizar espaços nem acentos

#exibir a estrutura das variáveis após a alteração

#4) detalhar o dicionário dos dados (fazer uma tabela com o nome de cada coluna, tipo e a descrição dos dados que a mesma contém)

Coluna(variável)	Tipo da variável	Descrição
Nome_variavel	Qualitativa nominal Qualitativa categórica Quantitativa discreta Quantitativa continua	Que dado está sendo armazenado

#5) exibir as 5 primeiras linhas do dataframe, use head(5)

#6) exibir as 5 últimas linhas do dataframe, use tail (5)

#7) verificar se a base de dados contém NaN e tratá-los.

#8) excluindo colunas desnecessárias, justifique a exclusão.

#9) excluindo todas as linhas dos carros da marca 'Tata'

#reseta o índice (isso é necessário devido as exclusões das linhas)

#10) alterando dados das informações da coluna 'tipo_vendedor':

De 'Individual' de 'particular'

De 'Trustmark Dealer' de 'concessionária'

De 'Dealer' de 'distribuidora'

#11) crie uma nova coluna denominada ipva e calcule:

Para carros com menos de 20 anos: o ipva é 4% do valor do carro

Carros com 20 anos ou mais é isento de ipva

#12) crie novas colunas que julgue pertinente ao seu estudo, justifique.

#13) gravar o novo dataframe, com nome “carros_usados_atualizado.csv”

ANALISANDO OS DADOS

- #1) em um novo arquivo python abra a base de dados: “carros_usados_atualizado.csv”
- #2) faça uma tabela de frequência para cada variável categórica representações gráficas e interpretações dos resultados.
- #3) faça agrupamentos, gráficos e interpretações
- #4) faça tabelas de contagens, gráficos e interpretações
- #5) correlacione variáveis e interprete o resultado.

Importante fazer gráficos adequados aos dados.