

Projeto Final - Análise e Visualização de Dados

Step 1. Importar as bibliotecas necessárias

Step 2. Importar os dados deste [endereço](#)

Step 3. Atribua o csv a variável acidentes

	longitude	latitude	data de abertura	hora de abertura	bairro	endereco	c
0	-34.905705	-8.037728	01/01/2016	18:30	JAQUEIRA	AV RUI BARBOSA	EM FRENTE AO PARQUE D
1	-34.938068	-8.131177	04/01/2016	16:05	JORDÃO	RUADRALVAROFERRAZ	TERMINALDOÔNIBUSDEJO
2	-34.915617	-7.993612	06/01/2016	17:33	DOISUNIDOS	RUAVINTEEUMDEJUNHO343	PRÓXIMOAAASSEMBLÉIADEDEUSEA
3	-34.896884	-8.118186	07/01/2016	10:29	BOAVIAGEM	AVENGENHEIRODOMINGOSFERREIRA3333	EMFRENTEAGALERIA
4	-34.910203	-8.098421	07/01/2016	13:53	IMBIRIBEIRA	AVMARECHALMASCARENHASDEMORAES	EMBAIXODOSE

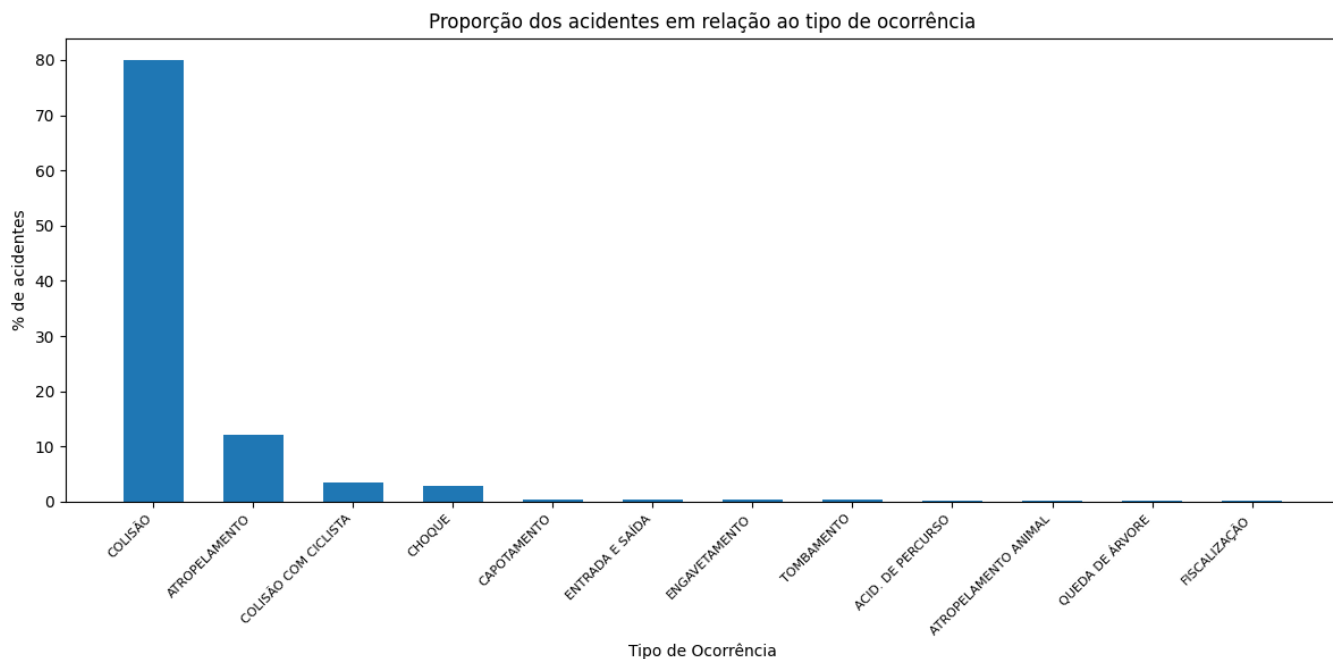
Step 4. Extraia a proporção dos acidentes em relação ao tipo de ocorrência

```
COLISÃO          79.119086
ATROPELAMENTO    11.990212
COLISÃO COM CICLISTA  3.425775
CHOQUE           2.773246
COLISÃOa         0.652529
CAPOTAMENTO      0.407830
ENTRADA E SAÍDA  0.326264
ENGAVETAMENTO    0.326264
TOMBAMENTO       0.326264
ACID. DE PERCURSO 0.163132
COLISÃO\t2016 13 050\t 0.163132
ATROPELAMENTO ANIMAL 0.081566
QUEDA DE ÁRVORE  0.081566
ATROPELAMENTOa   0.081566
FISCALIZAÇÃO    0.081566
Name: tipo de ocorrencia, dtype: float64
```

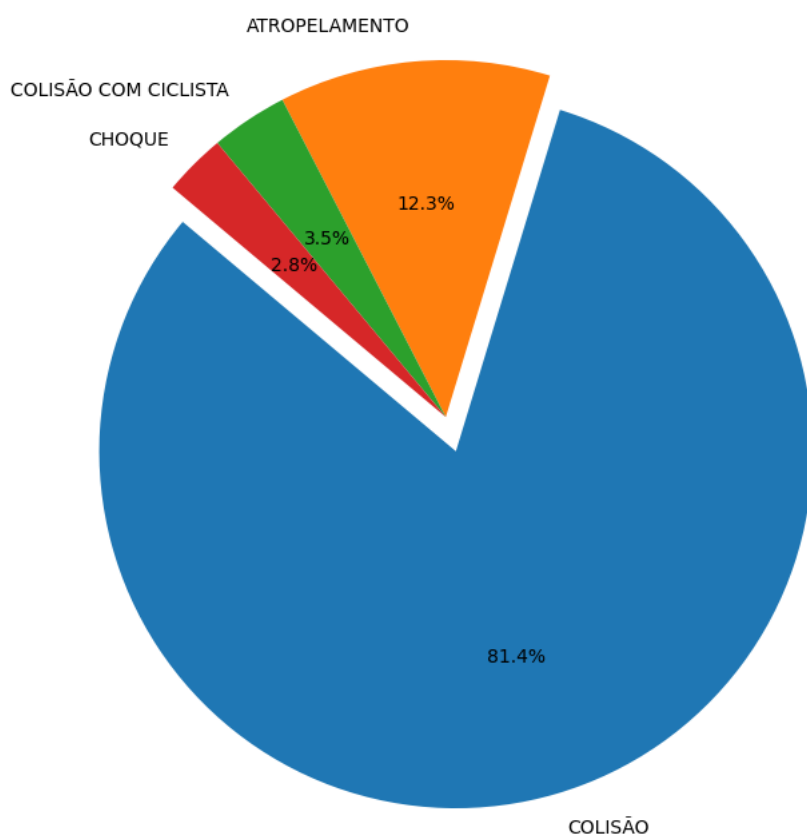
Step 5. Observe que há tipo de ocorrências que estão escritas erradas. Utilize str.replace() para agrupar os valores do mesmo tipo que podem estar com erro na escrita. No fim, serão 12 tipos de ocorrências.

```
TOTAL 12
array(['COLISÃO', 'CHOQUE', 'COLISÃO COM CICLISTA', 'ATROPELAMENTO',
      'CAPOTAMENTO', 'ATROPELAMENTO ANIMAL', 'ENTRADA E SAÍDA',
      'ENGAVETAMENTO', 'ACID. DE PERCURSO', 'QUEDA DE ÁRVORE',
      'TOMBAMENTO', 'FISCALIZAÇÃO'], dtype=object)
```

Step 6. Apresente um gráfico de barras contendo as informações da proporção dos acidentes em relação ao tipo de ocorrência (após tratamento)



- Step 7. Apresente um gráfico de pizza que mostre a porcentagem da proporção de acidentes por tipo de ocorrência apenas para os tipos de ocorrência que se repetiram 30 vezes ou mais. Na legenda deve estar o tipo de ocorrência. (Não se preocupar com textos sobrescritos)



- Step 8. Apresente a distribuição da quantidade de vítimas analisando o tipo de veículo.

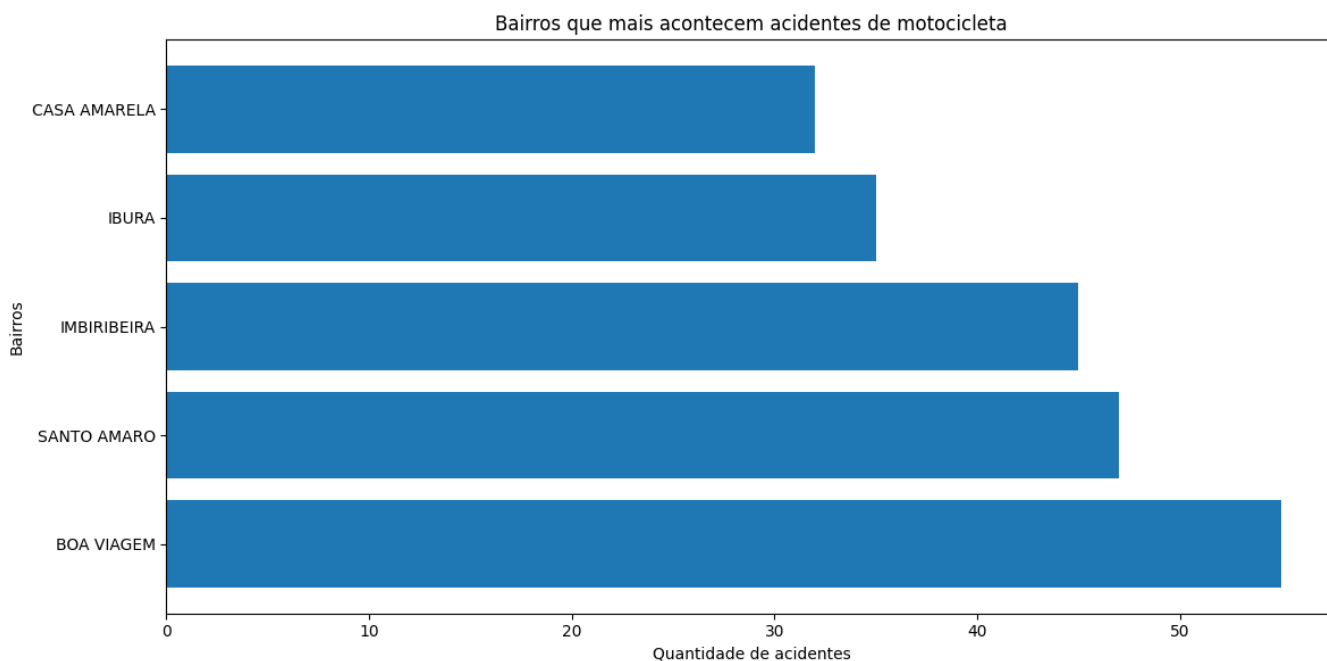
Antes disso, atenção: alguns valores podem estar errados tanto na coluna de vítimas quanto na coluna de tipos de veículo. Agrupe os valores do mesmo tipo de veículo, resultando em 6 tipos. Na coluna de vítimas, altere os valores que foram preenchidos errados por 0 (zero). Observe se as colunas estão preenchidas com as variáveis do tipo correto. (Para mudar o tipo de dados de uma coluna usa .astype())

```

tipo
Motocicleta    882
Automóvel      226
Pedestre       164
Ciclista        63
Outros          49
Ciclomotor      26
Name: quantidade de vitimas, dtype: int64

```

- ▼ Step 9. Encontrado o tipo de veículo com o maior número de vítimas, apresente um gráfico de barras horizontal que mostre os TOP 5 bairros em que aconteceram acidentes com esse tipo de veículo



- ▼ Step 10. Em qual mês teve mais vítimas? Quantas vítimas?

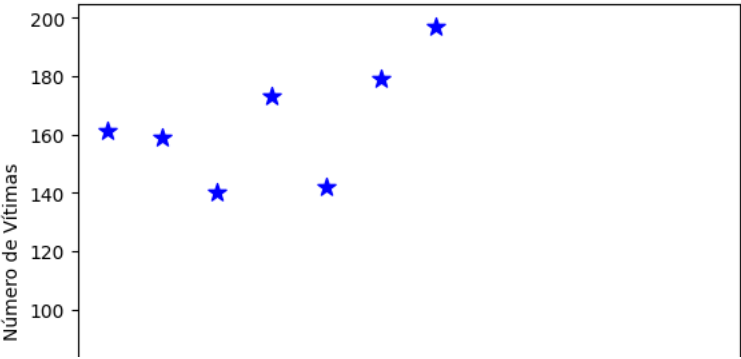
Atenção: Use `pd.to_datetime` para transformar uma coluna em formato data e use `dt.month_name()` para extrair o mês dessa coluna. É permitido criar novas colunas no dataframe.

```

mes
July          197
June          179
April         173
January       161
February      159
Name: quantidade de vitimas, dtype: int64

```

- ▼ Step 11. Insira um gráfico de dispersão que apresente no eixo x os meses e no eixo y o número de vítimas



▼ Step 12. Crie seu enunciado e seu gráfico.

Atenção: Deve ter manipulação de dados e plotagem de gráfico

