**Relatório - (Dado mongo-DB)**

Para esta atividade o problema é fazer um armazenamento utilizando ferramentas adequada, no caso, um gerenciador de banco de dados. Para utilizar esse banco de dados temaos duas opções:

1º : Fazer a instalação local no computador; ou

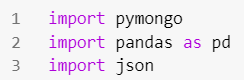
2º : Utilizar um servidor que forneça o serviço de armazenamento com o Mongo-DB.

No nosso caso vamos utilizar a 2ª opção!

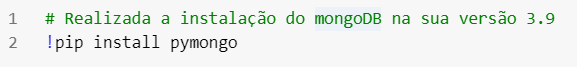
Nós daremos continuidade com os dados tratados na atividade ‘5.18 Dados preparados’

1ª PARTE:

Como start vamos importar as bibliotecas necessária.



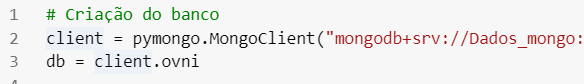
Antes de quaquer coisa vemos que importamos o “ pymongo ”.Ele ainda não esta instalado então vamos instala-lo pelo colab



2ª PARTE:

Agora vamos criar o banco de dados com o chamado OVIN: Assim podemos futuramente acessar o nosso banco.

**\*Este link está na criação do nosso banco**



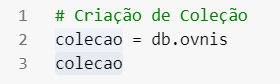
Saída:





3ª PARTE:

Em seguida vamos criar uma coleção com o nome (‘ovin’):

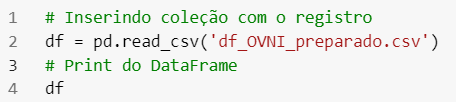


Saída:



4ª PARTE:

Inserindo na coleção criada todos os registros da base de dados com o DataFrame (df\_OVNI\_preparado) na formatação CSV



Saída:



E como complemento deletar dados não necessários do DataFrame:

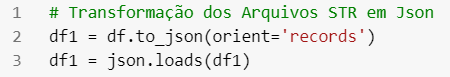


Saída:

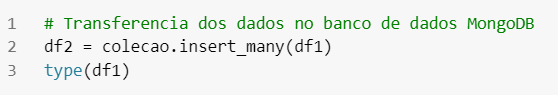


5ª PARTE:

Transformação do formato STR em JSON



Transferência de múltiplos dados para o banco e verificando o tipo de dado.



6ª PARTE:

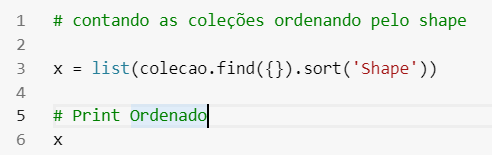
Agora vamos contar todas as ocorrências:



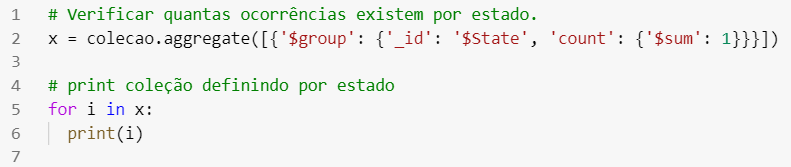
Saída:



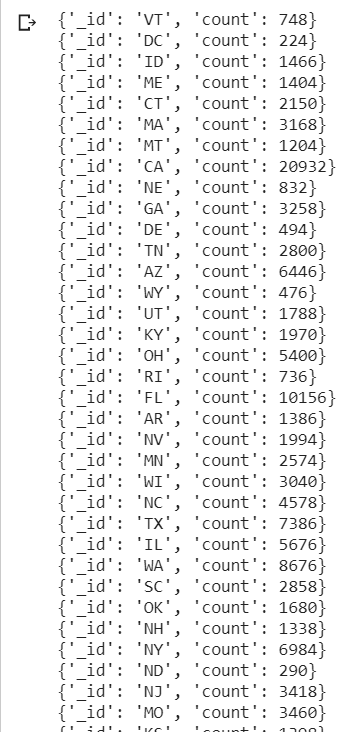
Vamos fazer lista das coleções ordenado pelo critério ‘Shape’.



Em seguida vamos fazer a busca da quantidade de ocorrências por estado

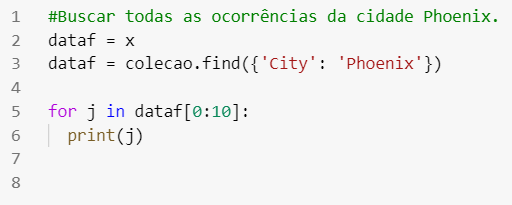


Saída:



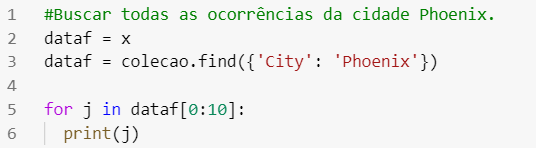
7ª PARTE:

Para termos uma busca mais especifica vamos acessar a coluna ‘City’ e pesquisar as ocorrências na cidade Phoenix.

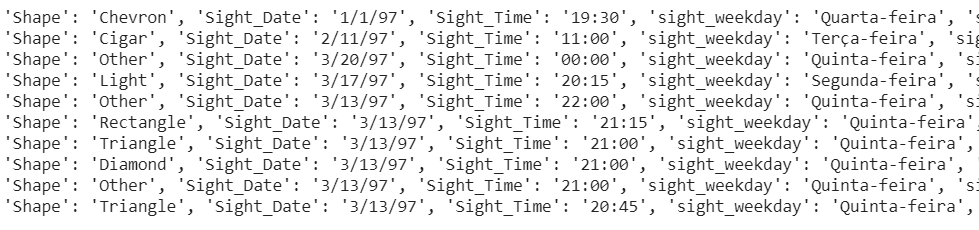


8ª PARTE:

Agora vamos printar as 10 primeiras posições dos relatos da cidade:

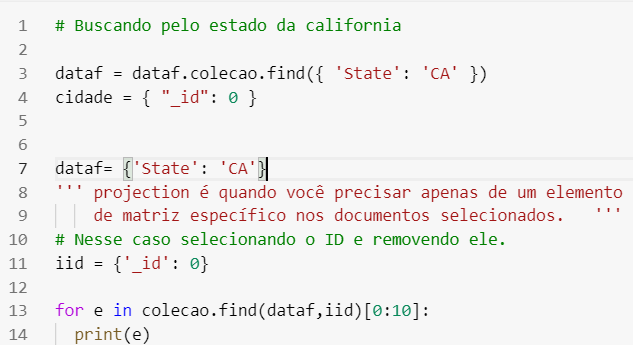


Saída:



9ª PARTE:

Para finalizarmos o processo vamos pesquisar pelo estado da Califórnia = ‘CA’. Com uma condição ocultando o ID dos registros.



E por fim tirando um print das 10 primeiras na lista

Saída:

LINK do projeto: https://github.com/samuelflopes/DataScience.git

FIM