

**TECNOLOGIA EM SISTEMAS PARA INTERNET**

**Daniel Evangelista Pereira**

**Ribson Coelho Cardoch Valdés**

**Douglas Seidi Shibata**

**RELATÓRIO DE PRÁTICA INTEGRADA**

**DE**

**CIÊNCIA DE DADOS E INTELIGÊNCIA ARTIFICIAL**

01/10/2020

**Brasília - DF**

**2020**

**Sumário**

[**1. Objetivos**](#_r578hjw9oxco) **3**

[**2. Descrição do problema**](#_jsmvxwf8smme) **4**

[**3. Desenvolvimento**](#_ybi6lms96bcw) **5**

[3.1 Código implementado](#_tcrscypicfq5) 5

[**4. Considerações Finais**](#_sp9srgapjf6s) **12**

[**Referências**](#_5hyd2pq8xv2g) **13**

# 1. Objetivos

O Objetivo desta sprint é acrescentar novas variáveis para que os dados fiquem de uma forma mais acessível para a análise, e por isso será destrinchado a coluna Date/time para que seja mais fácil a análise com o desagrupamento do dia e da hora.

# 2. Descrição do problema

Desagrupar a data da hora e depois separar o dia por mês, dia, e dia da semana.

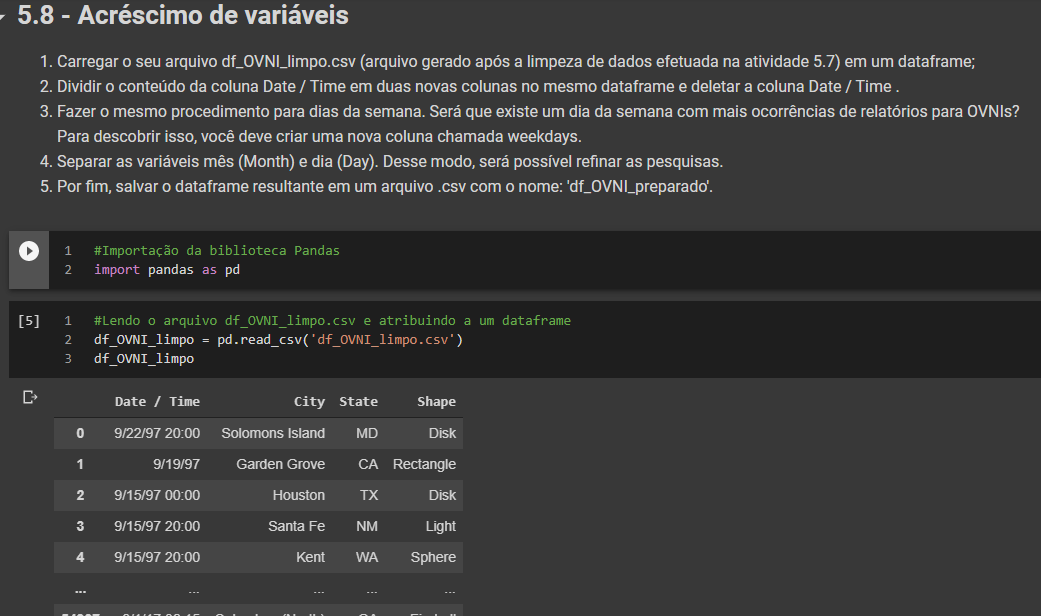
# 3. Desenvolvimento

As tecnologias utilizadas para a elaboração desta fase de acréscimo de variáveis do projeto, será a linguagem python e algumas bibliotecas para ajudar no desenvolvimento, com a biblioteca pandas , por meio do notebook e o ambiente de desenvolvimento Google Colab.

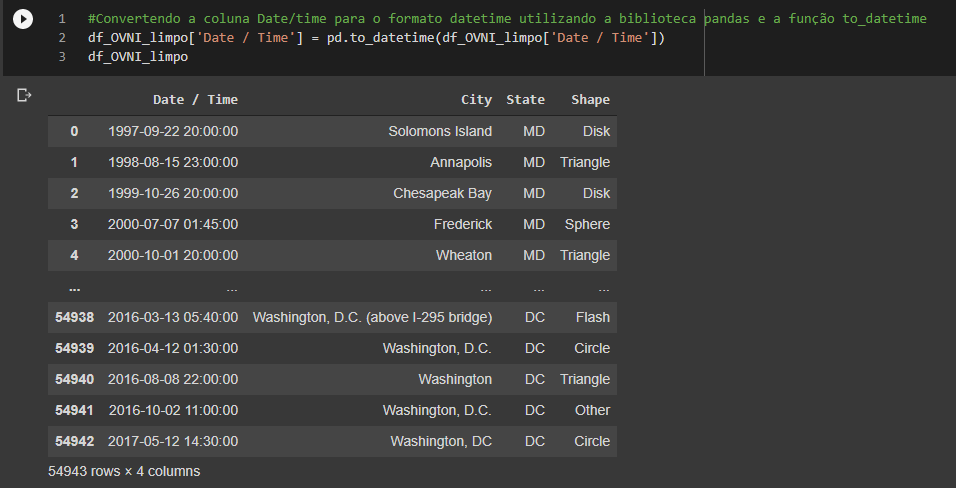
## 3.1 Código implementado

5.8 Acréscimo de variáveis

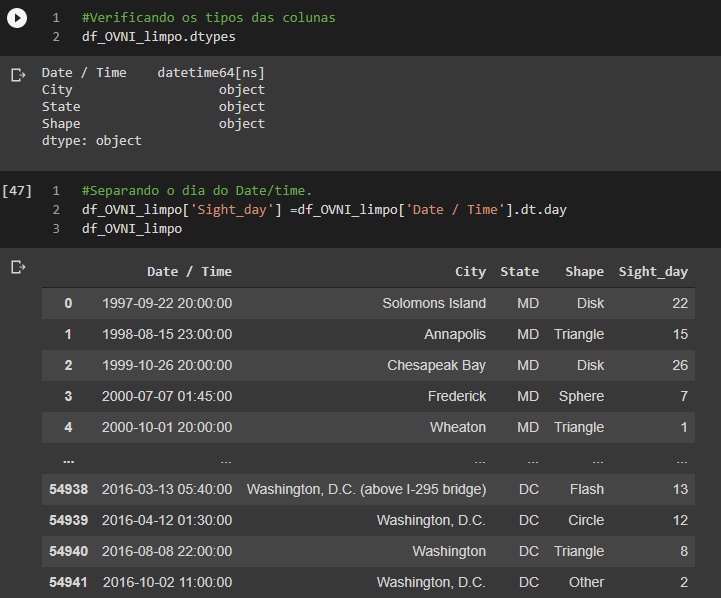
Realizando a importação da biblioteca e lendo o csv df\_OVNI\_limpo que foi gerado na fase anterior que está sendo atribuindo a variável df\_OVNI\_limpo que será o Dataframe que será trabalhado



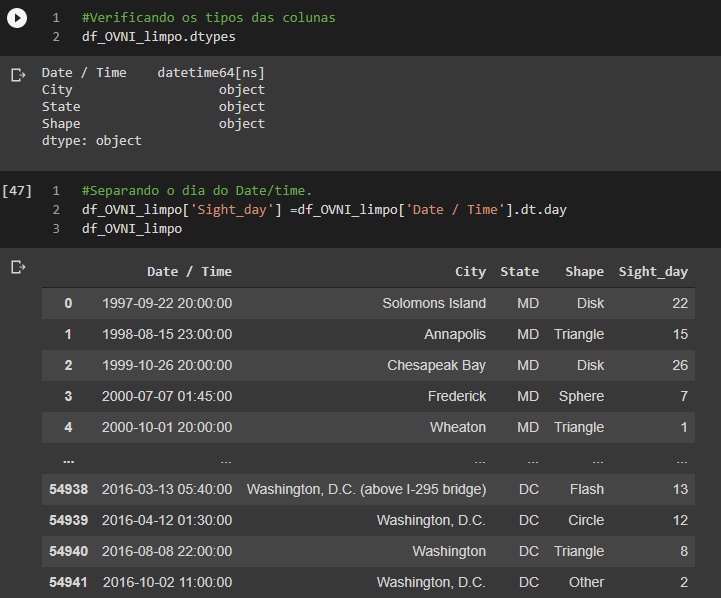
Convertendo Date / Time para o tipo datetime. Utilizando a função do pandas to\_datetime.



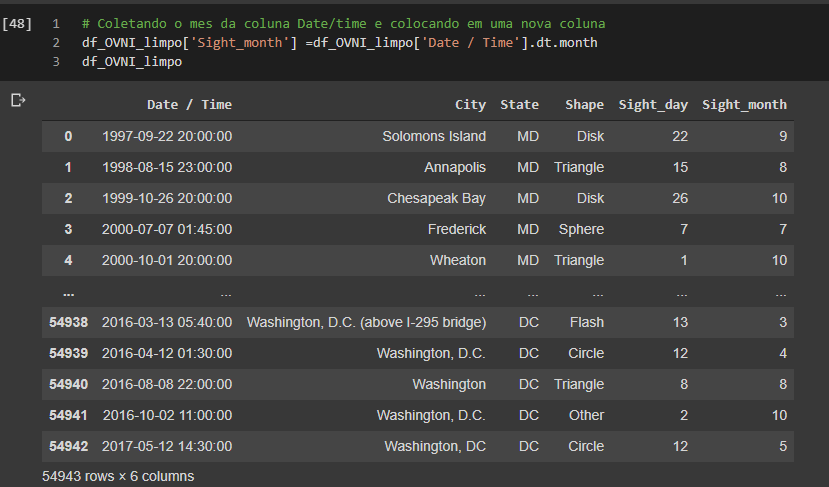
Verificando o tipos das colunas do dataframe . Utilizando o dtypes.



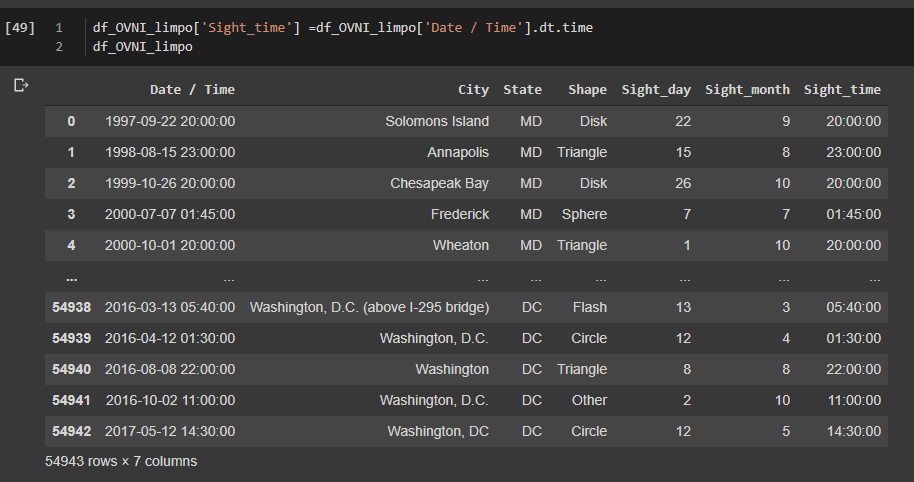
Declarando uma nova coluna chamada Sight\_day e coletando o dia a partir da coluna Date/time.



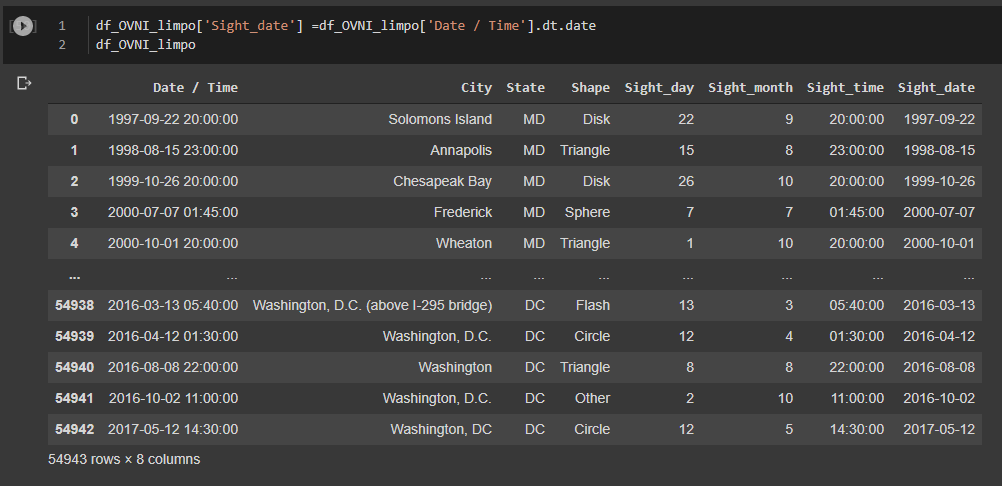
Declarando uma nova coluna chamada sight\_month e atribuindo o valor da coluna Date/time coletando apenas o mês da ocorrência



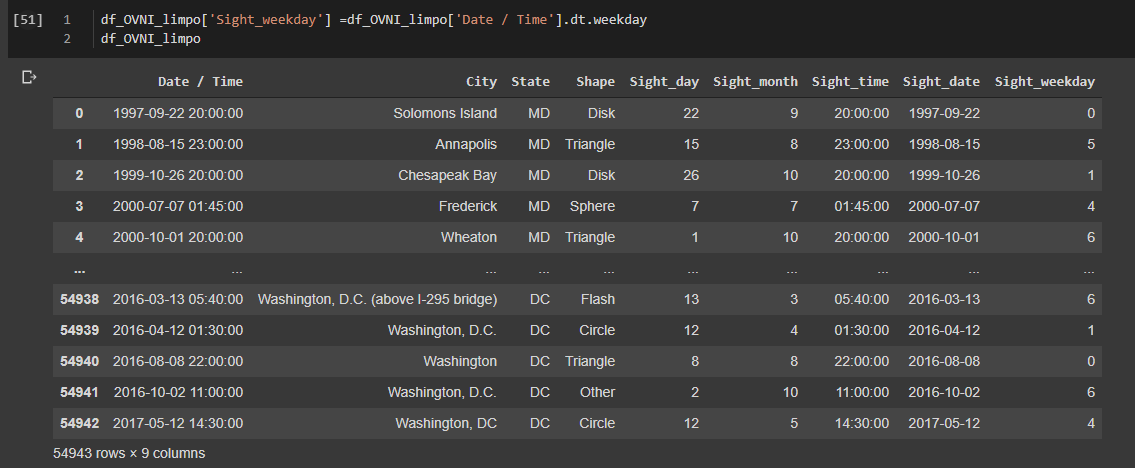
Declarando uma nova coluna chamada sight\_time e coletando a hora a partir da coluna date /time



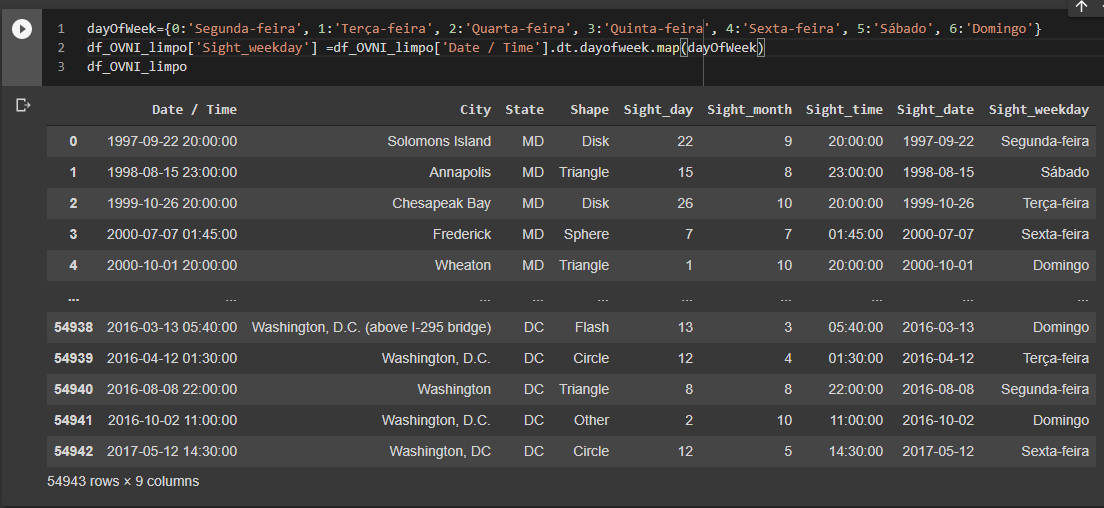
Coletando a data da coluna Date/time e atribundo este valor a uma nova coluna chamada sight\_date



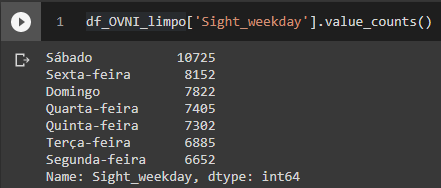
Coletando o dia da semana que ocorreu a ocorrência, criando uma nova coluna e atribuindo o valor de 0 a 6 indicando o dia da semana do fato, onde 0 é segunda feira, 1 é terça feira, 2 é quarta-feira, 3 é quinta-feira, 4- Sexta-feira, 5- Sábado e 6 - Domingo



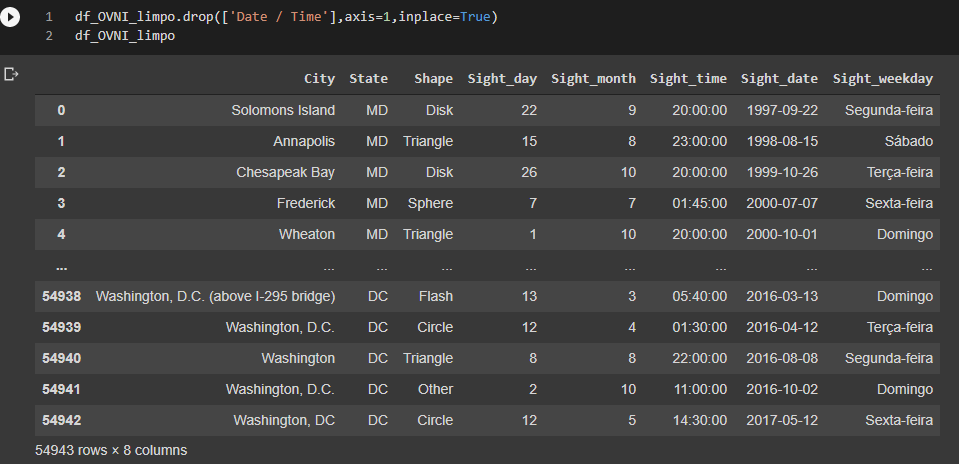
Para ficar mais fácil a interpretação do dia da semana, foi criado um dicionário atribuindo os valores do dia da semana por extenso e foi colocado na coluna de sight\_weekday utilizando a função map que percorreu os valores e atribui os valores por extenso.



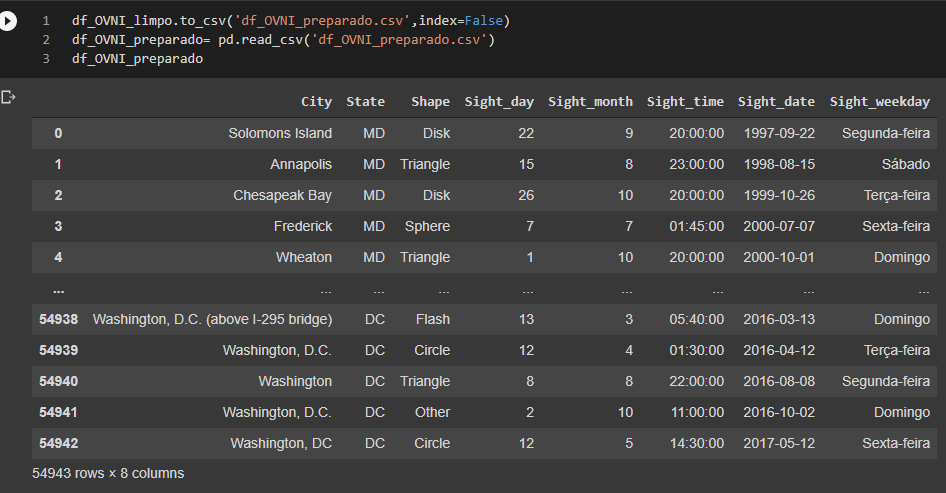
E após colocado os nomes dos dias da semana, foi possível saber qual foi o dia da semana que houve mais ocorrências



Excluindo a coluna Date/time



Utilizando a função to\_csv para gerar um novo csv com o dataframe dos ovnis já preparados com o dia, mês, hora, data e dia da semana.



# 4. Considerações Finais

Após feita a etapa de acréscimo a variáveis pode-se concluir que para futuras análises o dataframe ficou mais dividido para a coleta de informações.

# Referências

Little Bob Tables, Num day to Name day With Pandas.Stack overflow. 1º de Setembro de 2018. Disponível em <<https://stackoverflow.com/questions/29096381/num-day-to-name-day-with-pandas/>> Acesso em 30 de Setembro de 2020.