# Projeto de prática Integrada de ciência de dados, inteligência artificial e machine learning

# 

## 

## Relatório de avistamento de Objetos Voadores Não Identificados.

## Sprint 2 → Acréscimo de Variáveis

Curso: Tecnologia em sistemas para internet

Estudantes:

Brenda Lopes Miranda Teixeira

Mateus Gomes da Silva Fonteles

Rickson Queiroz Marques de Souza

Samuel Araújo Lopes

Professores

Fábio Henrique

Diego Queiroz

Ana Régia

Brasília, agosto de 2021

## Sumário

[1. Objetivos](#_10388dqmfo5m) 3

[2. Descrição do problema](#_5ux8aubh88f4) 4

[3. Desenvolvimento](#_wkz1qes89g9u) 5

[3.1. Código implementado](#_q0vrp2dio7y5) 6

[4. Considerações Finais](#_t4rgjanj1qvz) 8

[Referências](#_wvqxb1nnry4r) 9

## 1. Objetivos

Nesta etapa do trabalho, com nosso arquivo OVNIS\_limpo.csv, será executada a seguinte sequência de tarefas:

1. Dividir o conteúdo da coluna Data/Time em duas novas colunas
2. Deletar a coluna Date/Time
3. Dividir a tabela Date para extrair colunas com os dias da semana, mês e dias.
4. Salvar o *dataframe* resultante em um arquivo de nome ‘df\_OVNI\_preparado.csv’

## 2. Descrição do problema

Após havermos procedido com a limpeza do nosso dataset, vamos dividir as colunas de data e hora, para melhor realizar operações com as mesmas, assim como separar a data em diversos níveis de granularidades, que permitirão realizar análises mais detalhadas dos relatos, podendo separá-los por dias da semana ou do mês. por exemplo, e saber em que dias da semana ou períodos do mês os avistamentos são mais frequentes.

## 3. Desenvolvimento

Este trabalho está sendo desenvolvido usando um Script Python por ser uma linguagem orientada a objetos é bastante maleável, o grupo está utilizando a plataforma Google Colaboratory, assim todos podem modificar e acrescentar o código quando necessário.

## 3.1. Código implementado

* Importação de bibliotecas e carregamento de arquivo:
* Para esta etapa do projeto somente será necessária a importação do Pandas.

#Importando Bibliotecas

import pandas as pd

#Carregando o arquivo csv

df\_OVNI\_limpo = pd.read\_csv('df\_OVNI\_limpo.csv')

df\_OVNI\_limpo

* Preparação da tabela:

#Gerando a coluna Date/time para o formato datetime

df\_OVNI\_limpo['Date / Time'] = pd.to\_datetime(df\_OVNIS\_limpo['Date / Time'])

df\_OVNI\_limpo

#Chamando o tipo

df\_OVNI\_limpo.dtypes

* Criando as colunas:

#Dia

df\_OVNI\_limpo['Sight\_day'] =df\_OVNI\_limpo['Date / Time'].dt.day

df\_OVNI\_limpo

#Mes

df\_OVNI\_limpo['Sight\_month'] =df\_OVNI\_limpo['Date / Time'].dt.month

df\_OVNI\_limpo

#Tempo

df\_OVNI\_limpo['Sight\_time'] =df\_OVNI\_limpo['Date / Time'].dt.time

df\_OVNI\_limpo

#Data

df\_OVNI\_limpo['Sight\_date'] =df\_OVNI\_limpo['Date / Time'].dt.date

df\_OVNI\_limpo

#Fim de semana

df\_OVNI\_limpo['Sight\_weekday'] =df\_OVNI\_limpo['Date / Time'].dt.weekday

df\_OVNI\_limpo

#Dias da semana

dayOfWeek={0:'Segunda-feira', 1:'Terça-feira', 2:'Quarta-feira', 3:'Quinta-feira', 4:'Sexta-feira', 5:'Sábado', 6:'Domingo'}

df\_OVNI\_limpo['Sight\_weekday'] =df\_OVNI\_limpo['Date / Time'].dt.dayofweek.map(dayOfWeek)

df\_OVNI\_limpo

df\_OVNI\_limpo['Sight\_weekday'].value\_counts()

df\_OVNI\_limpo.drop(['Date / Time'],axis=1,inplace=True)

df\_OVNI\_limpo

* Armazenando as modificações no arquivo ‘df\_OVNI\_preparado.csv’.

#Armazenando no novo CSV 'df\_OVNIS\_preparado'

df\_OVNI\_limpo.to\_csv('df\_OVNI\_preparado.csv',index=False)

df\_OVNI\_preparado= pd.read\_csv('df\_OVNI\_preparado.csv')

df\_OVNI\_preparado

## 4. Considerações Finais

Partimos de uma coluna data com a qual não se podia trabalhar adequadamente ou realizar cálculos, por conter informações de mais de um tipo, para uma definição bem específica e segmentada do tempo dos avistamentos, com a possibilidade de aplicar agora mais filtros e observar os dados pelo ponto de vista do tempo de forma mais clara.

## Referências

* Dividir data (dia, mês, ano) em novas colunas - DataFrame Pandas. 2021. Disponível em:
  + <https://pt.stackoverflow.com/questions/428236/dividir-data-dia-m%C3%AAs-ano-em-novas-colunas-dataframe-pandas>