Relatório UFO Data

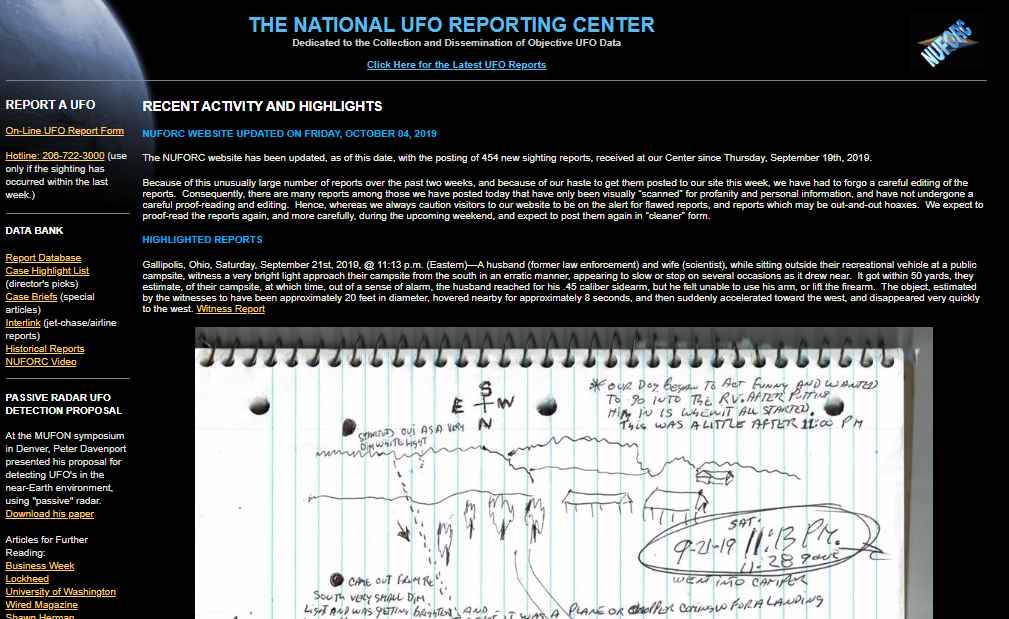
Levi Alves de Freitas Junior

**Problema**

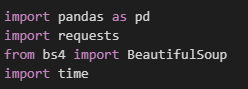
Criar um script para realizar requisições à um site utilizando conhecimentos de Web Scraping,o objetivo dessa etapa foi obter todos os dados entre os períodos de setembro 1997 e agosto de 2017, ao longo do relatório serão demonstrados as etapas até chegar ao resultado inicial que é de criar um arquivo CSV. Para visualizar todos os dados gerados ao longo desses anos.

A ferramenta utilizada foi o Google Colaboratory <https://colab.research.google.com/> . Dentro da Ferramenta utilizamos a linguagem Python e suas bibliotecas para análise de dados.

Foi utilizado o site <http://www.nuforc.org/> para obtermos todos os dados sobre Ovnis, a página inicial do site até o momento dessa postagem, encontra-se na figura abaixo.



Para a primeira parte precisamos importar as bibliotecas para requisitar a página, digite no Google Colab os seguintes comandos:



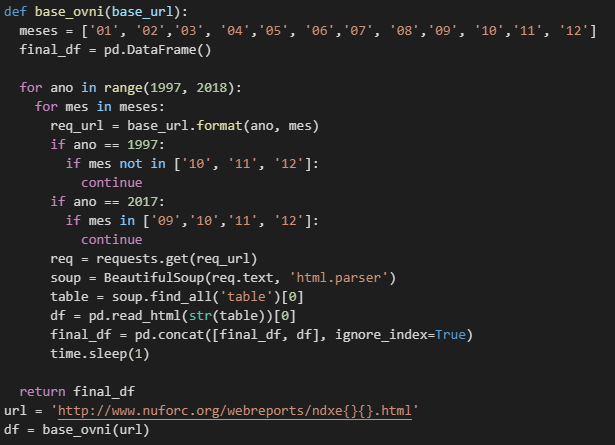
Pandas – Para armazenar, limpar e salvar os dados em tabelas.

Requests – Para realizar requisições http.

BeaultifulSoup – Para extrair dados em arquivos HTML e XML.

Time – Para realizar funções que utilizam tempo. No exemplo do projeto é utilizado o time para a espera entre as requisições ao site Nuforc.

Para a segunda etapa vamos criar uma função que passa como parâmetros (entre parênteses) a url do site, definimos os meses que nosso código irá utilizar, foi criado um for para navegar nos anos com início e fim e um for para os meses.



Utilizou-se if no ano de 1997 para começar dos meses à partir de setembro e um if para terminar no ano de 2017 no mês de agosto. Foi feito o request, criarmos um DataFrame com os dados e concatenamos ao nosso final\_df e damos uma pausa a cada iteração do for, para que não sejamos bloqueado ao fazer as requisições.

E por fim criamos um novo arquivo.csv no Colab para testarmos nossa requisição. Com o seguinte comando.



O projeto encontra-se no GitHub – para eventual consulta. <https://github.com/LeviAFJunior/Analise_Dados_Base_ovni.git>

**Referências**

WebScraping - <https://medium.com/data-hackers/como-fazer-web-scraping-em-python-23c9d465a37f>

Pandas - <https://pandas.pydata.org/pandas-docs/stable/>