

TECNOLOGIA EM SISTEMAS PARA INTERNET

**Kenia Duarte
Lucas dos Santos
Ana Amélia da Conceição**

**RELATÓRIO DE PRÁTICA INTEGRADA
DE
CIÊNCIA DE DADOS E INTELIGÊNCIA ARTIFICIAL**

Brasília - DF

03/10/2020

Sumário

1. Objetivos	3
2. Descrição do problema	4
3. Desenvolvimento	5
3.1 Código implementado	5
4. Considerações Finais	6
Referências	7

1. Objetivos

Este trabalho tem por objetivo investigar dados de vinte anos, entre os anos 1997 e 2017. Coletar esses dados entre os meses de setembro de 1997 e setembro de 2017. Os dados serão salvos em um DataFrame e exportado para em um arquivo .CSV.

2. Descrição do problema

Para esta coleta temos que descobrir fatos interessantes das observações de OVINIS, que são os objetos voadores não identificados. Objeto ou luz visto no céu, que não consegue ser identificado. Essa coleta será feita dentro de um período de vinte anos, de setembro de 1997 a agosto de 2017. O desafio é fazer a extração de dados sendo de forma tabular, porque para analisar bem os dados eles precisam se tornar tabelas.

3. Desenvolvimento

Nesse passo do projeto é o acréscimo de variáveis. Muitas vezes, os dados se apresentam ordenados de uma maneira não funcional para a realização de pesquisas. Por exemplo, em nosso *dataset* temos uma variável *Date...Time*, em que hora e data estão agrupadas. Iremos separá-las, gerando duas variáveis diferentes entre outras modificações.

Foi feita uma sequência de tarefas:

1. carregar o arquivo *df_OVNI_limpo.csv* em um dataframe;
2. dividir o conteúdo da coluna *Date / Time* em duas novas colunas no mesmo dataframe e exclusão da coluna *Date / Time* e dias da semana, criando um arquivo *weekdays* para verificar um dia com maior ocorrência;
3. separar as variáveis mês (*Month*) e dia (*Day*) sendo possível refinar as pesquisas;
4. salvar o dataframe resultante em um arquivo *.csv* com o nome: *'df_OVNI_preparado'*.

3.1 Código implementado

4. Considerações Finais

Os dados que foram analisados são muito interessantes. Uma maneira melhor de analisar e compreender cada um dos dados foi selecionar tabela. Foi possível separar os dados em tabelas, separando as variáveis em dia e mês para verificar o dia com maior ocorrência.

Referências

Vooco – Insights. Data Science. Python. Gestão. **Guia para iniciantes de Web Scraping em Python (usando BeautifulSoup).**

<https://www.vooco.pro/insights/guia-para-iniciantes-de-web-scraping-em-python-usando-beautifulsoup/> - acessado em 11/09/2020.

5.8 - Acréscimo de variáveis.

https://canvas.instructure.com/courses/1839379/assignments/13790565?module_item_id=28164090