

TECNOLOGIA EM SISTEMAS PARA INTERNET

**Emanuelle Ferreira da Silva
Lucas Matheus Silva
Neander Wendel Nobre Teixeira**

**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

O projeto se encontra na fase de acréscimo de variáveis, em que o principal objetivo é desagrupar as variáveis date/time e acrescentar novas variáveis como sight datetime, weekdays, day, month da tabela da fase anterior.

2. Descrição do problema

O problema a ser desenvolvido nessa etapa do projeto consiste no acréscimo de novas variáveis como sight date(data), time(hora), weekdays(dia da semana), day(dia), month(mês) com o objetivo de organizar os dados de acordo com as suas categorias, além de ter trabalhando com a conversão da variável date/ time e o seu desagrupado.

3. Desenvolvimento

O desenvolvimento desta etapa consiste no acréscimo de novas variáveis na tabela. Utilizando o ambiente de desenvolvimento Google Colab e a biblioteca pandas e algumas funções como drop e map para o manuseio e o acréscimo das novas variáveis.

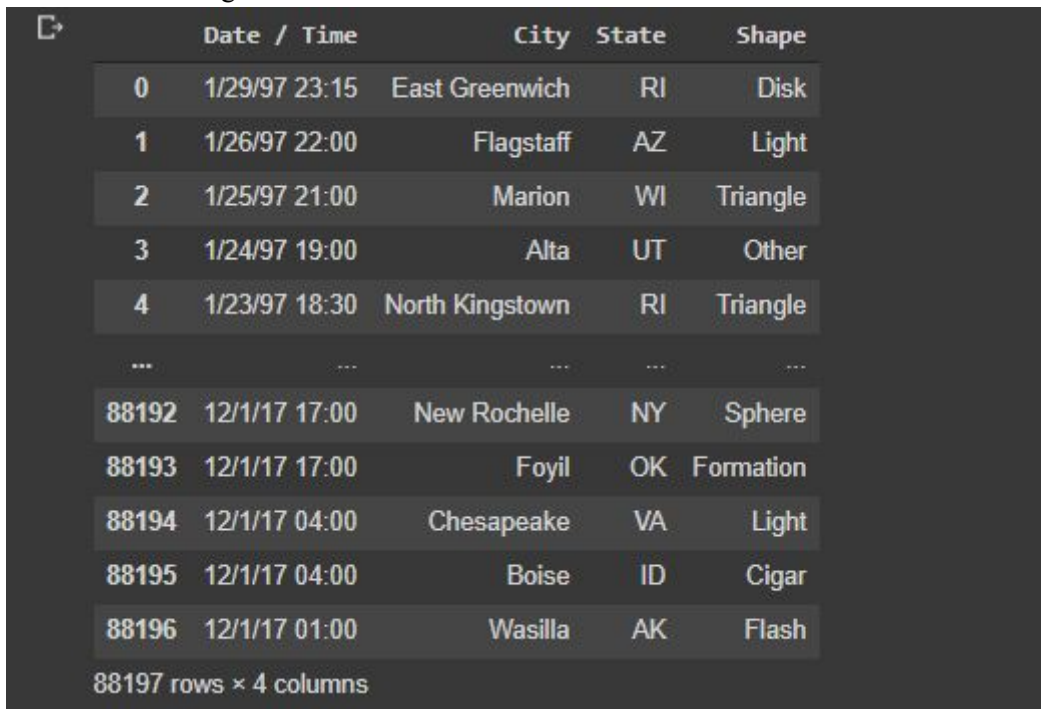
3.1 Código implementado

#5.8 Acréscimo de Variáveis

- 1- Carregando o arquivo df_OVNI_limpo gerado na atividade 5.7 após a limpeza dos dados em um novo data frame chamada acres_variavel.

```
3 acres_variavel = pd.read_csv("df_OVNI_limpo.csv")
4 acres_variavel
```

Resultado do código acima.



	Date / Time	City	State	Shape
0	1/29/97 23:15	East Greenwich	RI	Disk
1	1/26/97 22:00	Flagstaff	AZ	Light
2	1/25/97 21:00	Marion	WI	Triangle
3	1/24/97 19:00	Alta	UT	Other
4	1/23/97 18:30	North Kingstown	RI	Triangle
...
88192	12/1/17 17:00	New Rochelle	NY	Sphere
88193	12/1/17 17:00	Foyil	OK	Formation
88194	12/1/17 04:00	Chesapeake	VA	Light
88195	12/1/17 04:00	Boise	ID	Cigar
88196	12/1/17 01:00	Wasilla	AK	Flash

88197 rows x 4 columns

- 2- Convertendo a coluna date/time por meio do pd.to_datetime, para logo em seguida divida-lá em duas colunas.

```
2
3 acres_variavel["Date / Time"] = pd.to_datetime(acres_variavel["Date / Time"])
```

- 3- Criando duas novas variáveis sight_date e sight_time e cessado os dados pormeio da at

```
2 acres_variavel["sight_Date"], acres_variavel["sight_Time"] = (acres_variavel["Date / Time"].dt.date, acres_variavel["Date / Time"].dt.time)
3 acres_variavel
```

Resultado do código acima.

	Date / Time	City	State	Shape	sight_Date	sight_Time
0	1997-01-29 23:15:00	East Greenwich	RI	Disk	1997-01-29	23:15:00
1	1997-01-26 22:00:00	Flagstaff	AZ	Light	1997-01-26	22:00:00
2	1997-01-25 21:00:00	Marion	WI	Triangle	1997-01-25	21:00:00
3	1997-01-24 19:00:00	Alta	UT	Other	1997-01-24	19:00:00
4	1997-01-23 18:30:00	North Kingstown	RI	Triangle	1997-01-23	18:30:00
...
88192	2017-12-01 17:00:00	New Rochelle	NY	Sphere	2017-12-01	17:00:00
88193	2017-12-01 17:00:00	Foyil	OK	Formation	2017-12-01	17:00:00
88194	2017-12-01 04:00:00	Chesapeake	VA	Light	2017-12-01	04:00:00
88195	2017-12-01 04:00:00	Boise	ID	Cigar	2017-12-01	04:00:00
88196	2017-12-01 01:00:00	Wasilla	AK	Flash	2017-12-01	01:00:00

88197 rows × 6 columns

4-Removendo a coluna date/time da tabela por meio da função drop.

```
3 df_OVNI_preparado = acres_variavel.drop(columns=['Date / Time'],axis=1)
4 df_OVNI_preparado
5
```

Resultado do código acima com duas colunas.

	City	State	Shape	sight_Date	sight_Time
0	East Greenwich	RI	Disk	1997-01-29	23:15:00
1	Flagstaff	AZ	Light	1997-01-26	22:00:00
2	Marion	WI	Triangle	1997-01-25	21:00:00
3	Alta	UT	Other	1997-01-24	19:00:00
4	North Kingstown	RI	Triangle	1997-01-23	18:30:00
...
88192	New Rochelle	NY	Sphere	2017-12-01	17:00:00
88193	Foyil	OK	Formation	2017-12-01	17:00:00
88194	Chesapeake	VA	Light	2017-12-01	04:00:00
88195	Boise	ID	Cigar	2017-12-01	04:00:00
88196	Wasilla	AK	Flash	2017-12-01	01:00:00

88197 rows × 5 columns

5-Criando uma nova variável para os dias da semana

```
df_OVNI_preparado["sight Weekdays"] = (acres_variavel["Date / Time"].dt.weekday)
df_OVNI_preparado
```

Resultado do código acima com as três colunas como date, time e weekdays.

	City	State	Shape	sight_Date	sight_Time	sight_Weekdays
0	East Greenwich	RI	Disk	1997-01-29	23:15:00	2
1	Flagstaff	AZ	Light	1997-01-26	22:00:00	6
2	Marion	WI	Triangle	1997-01-25	21:00:00	5
3	Alta	UT	Other	1997-01-24	19:00:00	4
4	North Kingstown	RI	Triangle	1997-01-23	18:30:00	3
...
88192	New Rochelle	NY	Sphere	2017-12-01	17:00:00	4
88193	Foyil	OK	Formation	2017-12-01	17:00:00	4
88194	Chesapeake	VA	Light	2017-12-01	04:00:00	4
88195	Boise	ID	Cigar	2017-12-01	04:00:00	4
88196	Wasilla	AK	Flash	2017-12-01	01:00:00	4

88197 rows × 6 columns

6- Adicionando os nomes dos dias da semana na coluna sight_weekdays

```

2 dia_semana={0:'Segunda-Feira',1:'Terça-Feira',2:'Quarta-Feira',3:'Quinta-Feira',4:'SextaFeira',5:'Sábado',6:'Domingo',}
3 df_OVNI_preparado["sight_Weekdays"] = acres_variavel["Date / Time"].dt.weekday.map(dia_semana)
4 df_OVNI_preparado

```

Resultado do código acima com as variáveis date, time e weekdays criadas.

	City	State	Shape	sight_Date	sight_Time	sight_Weekdays
0	East Greenwich	RI	Disk	1997-01-29	23:15:00	Quarta-Feira
1	Flagstaff	AZ	Light	1997-01-26	22:00:00	Domingo
2	Marion	WI	Triangle	1997-01-25	21:00:00	Sábado
3	Alta	UT	Other	1997-01-24	19:00:00	SextaFeira
4	North Kingstown	RI	Triangle	1997-01-23	18:30:00	Quinta-Feira
...
88192	New Rochelle	NY	Sphere	2017-12-01	17:00:00	SextaFeira
88193	Foyil	OK	Formation	2017-12-01	17:00:00	SextaFeira
88194	Chesapeake	VA	Light	2017-12-01	04:00:00	SextaFeira
88195	Boise	ID	Cigar	2017-12-01	04:00:00	SextaFeira
88196	Wasilla	AK	Flash	2017-12-01	01:00:00	SextaFeira

88197 rows × 6 columns

7-Criando as variáveis sight_day e month para dia e mês.

```

1 df_OVNI_preparado["sight_Day"],df_OVNI_preparado["sight_Month"] = (acres_variavel["Date / Time"].dt.day, acres_variavel["Date / Time"].dt.month)
2 df_OVNI_preparado

```

Resultado do código acima com as variáveis criadas.

	City	State	Shape	sight_Date	sight_Time	sight_Weekdays	sight_Day	sight_Month
0	East Greenwich	RI	Disk	1997-01-29	23:15:00	Quarta-Feira	29	1
1	Flagstaff	AZ	Light	1997-01-26	22:00:00	Domingo	26	1
2	Marion	WI	Triangle	1997-01-25	21:00:00	Sábado	25	1
3	Alta	UT	Other	1997-01-24	19:00:00	SextaFeira	24	1
4	North Kingstown	RI	Triangle	1997-01-23	18:30:00	Quinta-Feira	23	1
...
88192	New Rochelle	NY	Sphere	2017-12-01	17:00:00	SextaFeira	1	12
88193	Foyil	OK	Formation	2017-12-01	17:00:00	SextaFeira	1	12
88194	Chesapeake	VA	Light	2017-12-01	04:00:00	SextaFeira	1	12
88195	Boise	ID	Cigar	2017-12-01	04:00:00	SextaFeira	1	12
88196	Wasilla	AK	Flash	2017-12-01	01:00:00	SextaFeira	1	12

88197 rows × 8 columns

8-Salvando o data frame final em um arquivo csv.

```
1 df_OVNI_preparado.to_csv("df_OVNI_preparado.csv", index=False)
```

9-Lendo o arquivo df_OVNI_preparado.csv por meio da função read_csv em um novo dataframe.

```
novo = pd.read_csv("df_OVNI_preparado.csv")
novo
```

Resultado do código acima com as variáveis criadas.

	City	State	Shape	sight_Date	sight_Time	sight_Weekdays	sight_Day	sight_Month
0	East Greenwich	RI	Disk	1997-01-29	23:15:00	Quarta-Feira	29	1
1	Flagstaff	AZ	Light	1997-01-26	22:00:00	Domingo	26	1
2	Marion	WI	Triangle	1997-01-25	21:00:00	Sábado	25	1
3	Alta	UT	Other	1997-01-24	19:00:00	SextaFeira	24	1
4	North Kingstown	RI	Triangle	1997-01-23	18:30:00	Quinta-Feira	23	1
...
88192	New Rochelle	NY	Sphere	2017-12-01	17:00:00	SextaFeira	1	12
88193	Foyil	OK	Formation	2017-12-01	17:00:00	SextaFeira	1	12
88194	Chesapeake	VA	Light	2017-12-01	04:00:00	SextaFeira	1	12
88195	Boise	ID	Cigar	2017-12-01	04:00:00	SextaFeira	1	12
88196	Wasilla	AK	Flash	2017-12-01	01:00:00	SextaFeira	1	12

88197 rows × 8 columns

4. Considerações Finais

Por meio dessa etapa do projeto podemos concluir que é muito importante realizamos a criação de novas variáveis na tabela para a melhor análise e compreensão dos dados.

Referências

FIGUEIREDO, Vinicius. Seus Primeiros Passos com Data Scientist: Introdução ao Pandas. **Data Hackers**. São Paulo, 30 de maio de 2018. Disponível em:

<<https://medium.com/data-hackers/uma-introdu%C3%A7%C3%A3o-simples-ao-pandas-1e15ea37fa1>>. Acesso em: 08 de Set. de 2020.

FAMETHEMES. Biblioteca seaborn com o matplotlib. **Vooo - Insights**. c2020

Disponível em: <<https://www.vooo.pro/insights/biblioteca-seaborn-com-o-matplotlib/>>. Acesso em: 26 de set. de 2020

MATHEUS, Yuri. Matplotlib uma biblioteca Python para gerar gráficos interessantes. **alura. São Paulo**, 27 de Novembro de 2018. Disponível em:

<<https://www.alura.com.br/artigos/criando-graficos-no-python-com-a-matplotlib>>. Acesso em: 25 de set. de 2020.