

DATA ANALYTICS
VISUALIZAÇÃO DE DADOS

AULA 05

SUMÁRIO

O QUE VEM POR AÍ?	3
CONHEÇA SOBRE O ASSUNTO	3
HANDS ON	4
O QUE VOCÊ VIU NESTA AULA?	4
REFERÊNCIAS	5

EMSE

O QUE VEM POR AÍ?

Até aqui você teve uma noção de como carregar fontes de dados de maneiras distintas, viu algumas manipulações base para sua jornada, conheceu a biblioteca Seaborn do Python, aprendeu algumas funções que combinam Matplotlib e Seaborn e a manipular datas com uma biblioteca nativa do Python, o Datetime.

Essa aula vai lhe trazer uma outra noção de modificações dentro de um Dataframe, manipulando datas por meio do próprio Pandas.

A base de dados você encontra aqui <https://github.com/alura-tech/pos-datascience-introducao-a-visualizacao/archive/refs/heads/dados.zip>.

Agora vamos conhecer mais sobre este maravilhoso mundo dos dados!

CONHEÇA SOBRE O ASSUNTO

Um modo eficaz e legal para usarmos quando queremos converter uma coluna de data para um formato correto, utilizando a biblioteca Pandas, é utilizar a função `to_datetime()`, que nos permite passar dentro da função a coluna da qual queremos fazer a conversão específica.

Podemos visualizar por meio do comando:

```
pd.to_datetime(mensal_aberto["dia_mes_ano"])
```

Esse exemplo foi tirado da nossa videoaula sobre manipulação de datas. Para melhor entendimento do tema, assista a videoaula!

Veja que, na verdade, essa videoaula é uma continuação da videoaula anterior, onde nós estamos ampliando as possibilidades na manipulação de datas.

Outro ponto importante é que, nesta aula aprendemos a como criar colunas desmembrando-as por ano, mês e dia, por meio do atributo 'dt'.

HANDS ON

Agora, chegou o momento de ver, na prática, como começar a importar os nossos dados e trabalhar com eles via programação. O ambiente utilizado é o Google Colab, e as bases de dados que foram disponibilizadas no início do material. A ideia é não se limitar apenas ao código explícito no hands on, então, é sempre bom procurar a documentação das bibliotecas, explorar novas funcionalidades e muito mais! Temos um exemplo muito próximo do que foi feito na videoaula, então assista para ver mais sobre o tema, combinado?!

O QUE VOCÊ VIU NESTA AULA?

- Como fazer uma comparação ano a ano dos custos do SUS usando o DateTime.
- Como analisar e manipular os dados para criar análises mais avançadas.

Material completo: notebook_completo <<https://github.com/alura-tech/pos-datascience-introducao-a-visualizacao/archive/refs/heads/aula5.zip>>.

Daqui em diante, é importante que você replique os conhecimentos adquiridos para fortalecer ainda mais suas bases e conhecimentos, já que um bom (ou uma boa) cientista de dados não é somente aquele(a) que é uma enciclopédia humana, mas sim aquele(a) que sabe ler um problema e atuar com eficácia.

IMPORTANTE: não esqueça de praticar com o desafio da disciplina, para que assim você possa aprimorar os seus conhecimentos!

Você não está sozinho(a) nesta jornada! Te esperamos no Discord e nas lives com os professores e professoras especialistas, onde você poderá tirar dúvidas, compartilhar conhecimentos e estabelecer conexões!

REFERÊNCIAS

DOCUMENTAÇÃO PANDAS. <<https://pandas.pydata.org/>>. Acesso em: 09 fev 2023.

GOOGLE COLAB. <<https://colab.research.google.com/>>. Acesso em: 09 fev 2023.

EMASP

PALAVRAS-CHAVE

Python. Pandas. Dataframe.

EMAP

The background is a dark blue field filled with numerous small, light blue dots, resembling a starry sky. Overlaid on this are several large, wavy, translucent lines in shades of blue and yellow. A vertical line on the left side features a small 'x' mark. A circle in the upper left contains the number '7'. A hexagon is located in the bottom right corner. The text 'POSTECH' is centered in a white, outlined font.

POSTECH