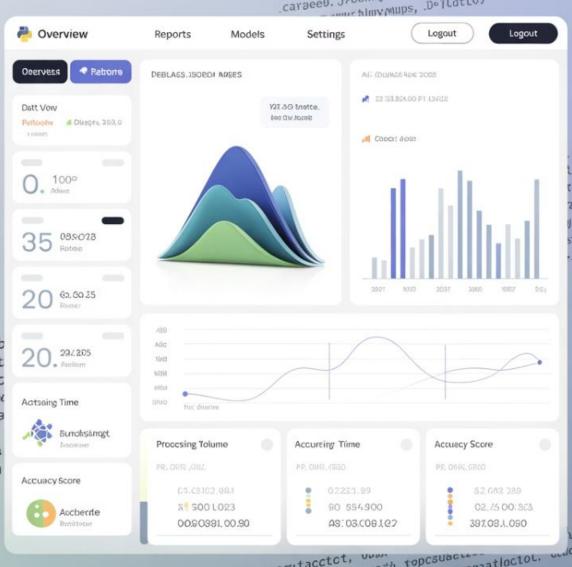
Mini Projeto de Análise de Dados

"Pythen, d'rnta" eryoetiye): Mp, t : 61 taarna, loave ongoosceslua, siyaca.ctuo, prisectiong tulwine Jactupso)

('Pumpuces)
oasuocyennyteosavitis, nža. coto ananenčii) dt
eoocugr /jyvol, is iocorten cane ovotsavook u
eoocugr /jyvol, is iocorten cane ovotsavook u
caraeeb. Jrounna lotenasekhiporiila, oecsuunt
caraeeb. Jrounna lotenasekhiporiila,



(j)geoogitacctot, vian topcsubcteibcastalloaupt)

v. "bcastalloaupt)

v. cgoccpsilluguatuep, abaun roatoduml shazear. a "rva" b topcsubctetu. cgoccpsilluguatuep, abaun roatoduml shazear.

v. cgoccpsilluguatuep, abaun roatoduml shazear.

Objetivo de Aprendizagem



Integração de Conhecimentos Estatísticos

Para interpretação precisa e crítica da validade das conclusões.



Realização de Análise Exploratória de Dados (AED)

E criação de visualizações relevantes com um olhar crítico.



Aplicação dos Fundamentos de Programação Python

E bibliotecas essenciais
(Pandas, NumPy) para
manipulação e processamento
de dados.



Estrutura da Análise e Comunicação

Clara e persuasiva dos resultados para fundamentar recomendações estratégicas.



Síntese das Habilidades

Python, Estatística, Comunicação para resolver problemas práticos e gerar inteligência acionável.

Fontes de Dados Disponíveis

Kaggle - Sales Data

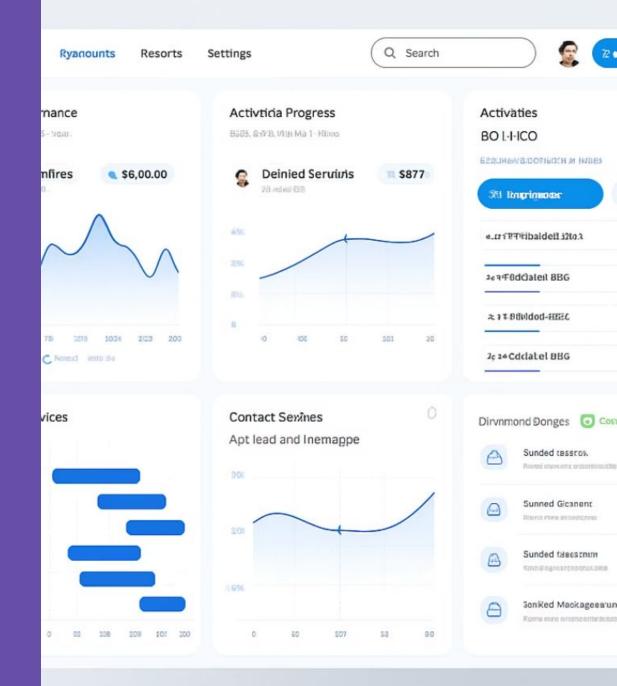
Conjunto de dados de vendas para análise de comportamento de compra e tendências comerciais.
Link: Kaggle - Sales Data

Student Performance Data Set

Conjunto de dados sobre desempenho acadêmico de estudantes.

Link: Student Performance

Data Set



Etapas Iniciais do Projeto

Importação de Bibliotecas

Importe pandas, numpy e matplotlib.pyplot.

Leitura e Exploração Inicial da Base

- Carregue o arquivo CSV escolhido usando pandas.
- Mostre as 5 primeiras linhas.

Mostre o nome das colunas, a quantidade de linhas/colunas e tipos de dados detectados pelo pandas (.dtypes).

Explique em Markdown os principais tipos de dados encontrados (str, int, float, ...).

Listas, Dicionários e Tuplas

Extraia uma coluna de interesse em uma **lista** Python e faça um loop imprimindo informações sobre cada item.

Monte um **dicionário** relacionando dois campos importantes da base (ex: produto:categoria, filme:gênero, aluno:disciplina) e exiba 3 pares.

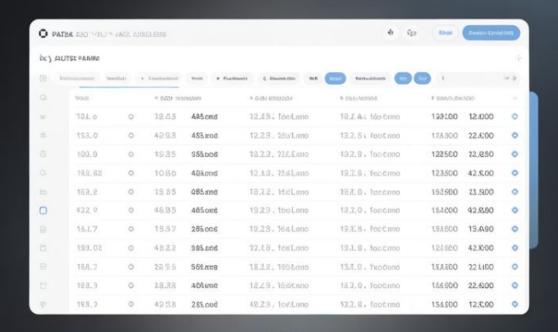
Crie uma **tupla** contendo três informações distintas de uma linha e imprima o resultado.

Pandas

Pyhoe Data! Districord

Data Insights Dashboard

Load Data



Sample Data import and exporation code

DIFFETTIND -BALE EDECRRIDWER)

RCPANUS TUDT ESDOO

EDGCHRISHER SSOURD ENALESTATIONN)

ACATE USYDIOPEESITEELRE)

RE 0877 I LATEACSBALAL, EJRONOUN

PIL, BLIL LIOTI-FEACSAPINICATY

CITESTI "PE CNATRINOTUMA"

MET AWWOEP LOCKETSS

CHARGET: SMITHS OVERDS
CLMMOYORES - "BGG ALCUS
COOGGETH UNDERMEREMELACTILE BOILING!
AREA (FRECT) - REJUSTE OG. SOEMMO"
OREC (FRECT) - REJUSTE OG. SOEMMO"

SHORPATTISCOCLREBISERELCANSTARSORMETT)

MO"NISCABRS BASSCOROTTGCOTHEN HUPCOESRIAL PARAY. B. TASUCOHNTPAUL

CURCONT SUVEMBILET!;

SUMGRIT S. TERO. TLAIR "BPFA/1)...

O"EC ONI "TOEDITAO SNGTCLOLOLES).

STOORTOTTE LALL ENSOLEBI, TTOOOTE

POI-ED. GDIAMS O. SUVE) | STRAURLASCONTERST. GENEL. BLISLAGLENTEPO WRM(O'''I SLIAGE ADRELORNTO (UNITY) "> LO JEDERS,

CHTAVP. L.ATOCCORSI: PBIIF "GRONSUI"SLGATCO BESSTJ "TSGROCSF ARTIG.TUMAL SOJEDHEIOT FEGGSTCADTTLOHEA) LOGALOTPA IMMMNI.

MATRI)

Estruturas Condicionais e Laços

Escolha uma coluna **numérica** de interesse na base e utilize uma **estrutura condicional (if/elif/else)** para imprimir uma mensagem de acordo com o valor de um elemento.

Use um **laço for** para calcular a soma dos primeiros 5 valores dessa coluna e mostre o resultado.

Empregue um **laço while** para encontrar o primeiro valor da mesma coluna que atenda a uma determinada condição e exiba esse valor quando encontrado.

Operadores Aritméticos e Manipulação de Dados

- Realize operações matemáticas (soma, subtração, multiplicação ou divisão) entre colunas ou valores de interesse da base.
- Crie uma nova coluna no DataFrame aplicando um cálculo simples.

NumPy e Arrays Numéricos

Crie um array NumPy a partir de uma coluna numérica da base.

- Realize pelo menos duas operações vetorizadas com esse array.
- Mostre como acessar, modificar e exibir partes do array (fatiamento/slicing ou indexação).
- Mostre como fazer uma operação agregada, como soma ou média, sobre o array.

Manipulação de Dados com Pandas

Vamos explorar algumas técnicas essenciais para acessar e transformar seus dados:

Seleção de Dados

Selecione linhas/colunas específicas usando .loc, .iloc ou filtrando por condição.

Análise Agregada

Use algum método de pandas para contar ou agrupar informações simples (ex: .value_counts(), .groupby()), mostrando o resultado em Markdown.

2

Agregações Simples

Utilize métodos como .value_counts() para contagem, e df.groupby('coluna').mean() para agrupar e calcular médias. Exiba os resultados em formato Markdown:

Visualização de Dados







Gráfico de Linha

Ilustrando tendências/séries temporais quando possível.

Gráfico de Barras

Comparando categorias relevantes.

Gráfico de Dispersão

Mostrando a relação entre duas variáveis numéricas.

Para cada gráfico, escreva em Markdown um comentário sobre o que ele mostra. Se algum dos gráficos não for aplicável à sua base, crie (ou simule) dados para mostrar que compreendeu o recurso.

Relatório e Insights



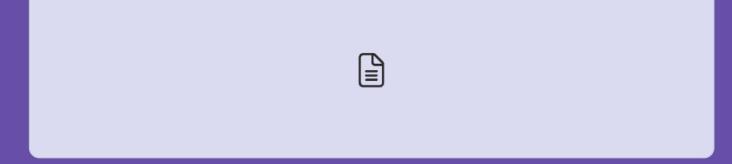
Descobertas

Após os usos de listas, dicionários, tuplas, operações, agrupamentos e gráficos, escreva em Markdown sempre que descobrir algo interessante, curioso ou surpreendente.



Análise

Interprete os resultados encontrados e explique seu significado no contexto dos dados.



Relatório Final

Inclua um **relatório final** em Markdown com suas conclusões e recomendações baseadas na análise.