

Ciência de Dados em Políticas Públicas e Saúde

Como um governo decide onde construir um novo hospital ou como combater uma epidemia?

A resposta está na **Ciência de Dados**, que permite analisar grandes volumes de informações para tomar decisões estratégicas.

Nesta aula, vamos explorar como os dados podem transformar a gestão pública e a saúde, tornando-as mais eficientes e baseadas em evidências.



O Mapa da Sociedade: Ciência de Dados em Políticas Públicas

↳ O que é Ciência de Dados em Políticas Públicas?

É a aplicação de técnicas analíticas avançadas para transformar dados brutos em **insights acionáveis** que podem informar decisões governamentais e políticas públicas.

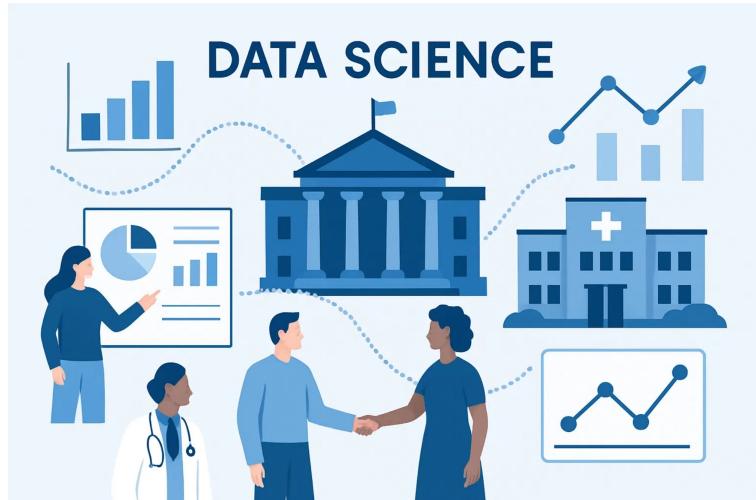
⚙️ Como os dados ajudam os governos?

🔍 **Entender a sociedade:** Mapear necessidades, comportamentos e tendências da população.

☰ **Monitorar serviços:** Acompanhar a eficiência e eficácia de serviços de saúde, educação e segurança.

◎ **Criar políticas eficazes:** Desenvolver intervenções baseadas em evidências e direcionadas a problemas específicos.

⚖️ **Distribuir recursos:** Alocar recursos limitados de forma mais justa e eficiente.



Exemplo Prático:

Análise de dados de criminalidade para identificar áreas prioritárias para policiamento, iluminação pública e programas sociais, resultando em redução de crimes em áreas específicas.

Ciência de Dados em Saúde Pública

的心 Applicações na Saúde Pública

A ciência de dados está transformando a saúde pública, permitindo abordagens mais **preventivas, precisas e eficientes** para proteger e melhorar a saúde da população.

📍 Rastreamento de Doenças

Identificação de padrões de propagação de doenças e fatores de risco em tempo real.

🕒 Monitoramento de Surtos

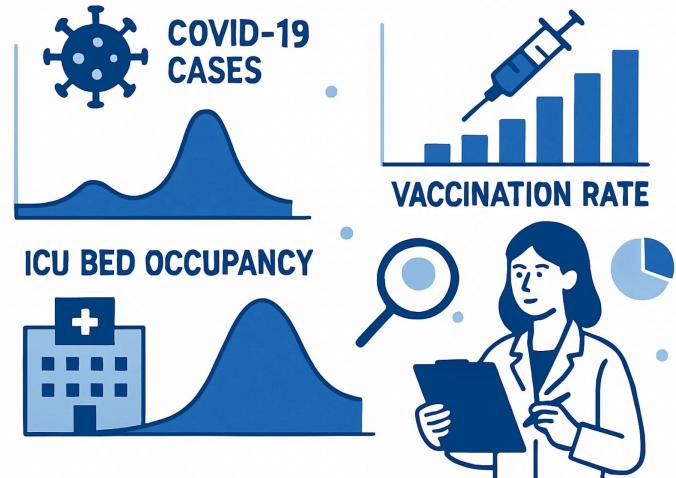
Detecção precoce e previsão da evolução de epidemias para intervenções rápidas.

📊 Otimização de Recursos

Alocação eficiente de leitos hospitalares, equipamentos e profissionais de saúde.

📝 Avaliação de Campanhas

Mensuração da eficácia de campanhas de vacinação e programas de saúde pública.



"A análise de dados em saúde pública não é apenas sobre números, mas sobre transformar informações em ações que salvam vidas."

Principais Fontes de Dados Públicos no Brasil

DATASUS

O Departamento de Informática do SUS é uma das **maiores fontes de dados de saúde do mundo**. Contém informações sobre o sistema de saúde brasileiro.

Exemplos de datasets:

SINASC: Nascidos Vivos

SIH: Internações Hospitalares

IBGE

O Instituto Brasileiro de Geografia e Estatística fornece **dados demográficos e socioeconômicos** essenciais para políticas públicas.

Exemplos de datasets:

Censo Demográfico

PNS: Pesquisa Nacional de Saúde

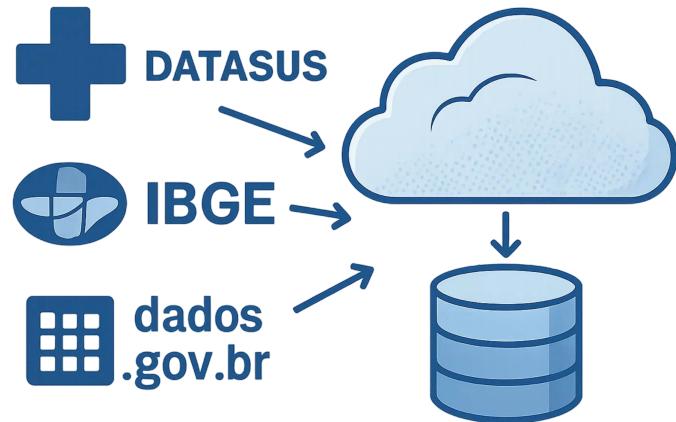
dados.gov.br

O portal de dados abertos do governo federal **reúne datasets de diversas áreas**, incluindo políticas sociais e econômicas.

Exemplos de datasets:

Dados do INEP sobre educação

Programas sociais



Dica:

Ao trabalhar com dados públicos, sempre verifique a data da última atualização e a metodologia de coleta.

Atividade A: Exploração de Dados Públicos

Exploração de Dados Públicos

Instruções

1. Divida-se em grupos de 3-4 pessoas
2. Cada grupo deve selecionar e baixar um dataset de uma das fontes mencionadas (DATASUS, IBGE, ou dados.gov.br) que seja relevante para saúde ou políticas públicas

Exemplos de Datasets:

 **DATASUS:** Dados de nascidos vivos

 **IBGE:** Taxas de analfabetismo

 **dados.gov.br:** Estatísticas de vacinação

 **DATASUS:** Internações hospitalares

3. Faça um reconhecimento inicial do dataset: que tipo de informações ele contém? Qual período ele cobre? Quais variáveis estão disponíveis?



Atividade B: Carregamento e Limpeza de Dados

Análise Exploratória em Python

Nesta primeira etapa, vamos trabalhar com os dados baixados das fontes públicas, começando pelo carregamento e limpeza básica.

→ 1. Carregamento do Dataset

```
import pandas as pd # Importando a biblioteca pandas
df = pd.read_csv("nome_do_arquivo.csv") # Carregando o arquivo CSV
df.head() # Visualizando as primeiras linhas
```

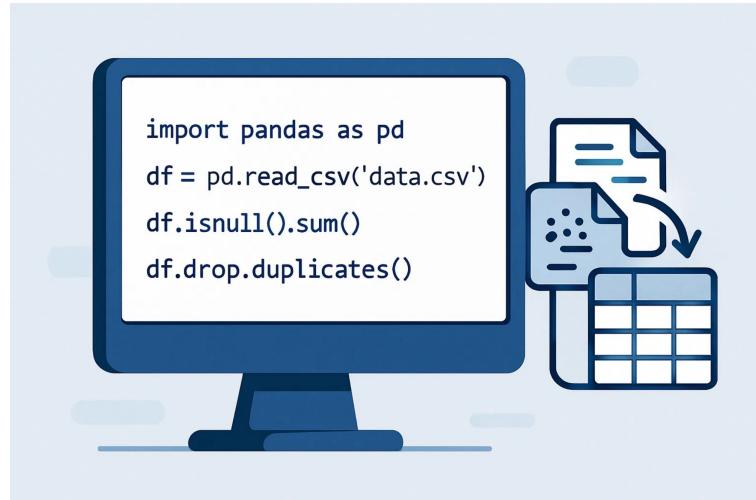
→ 2. Limpeza Rápida dos Dados

```
# Verificando valores nulos  valores_nulos = df.isnull().sum()
print(valores_nulos)
# Removendo dados duplicados  df = df.drop_duplicates()
```



Dica:

Para dados numéricos, considere preencher valores nulos com a média usando `mean()` em vez de removê-los.
`df['coluna'].fillna(df['coluna'].mean())`



Atividade B: Cálculo de Estatísticas Descritivas

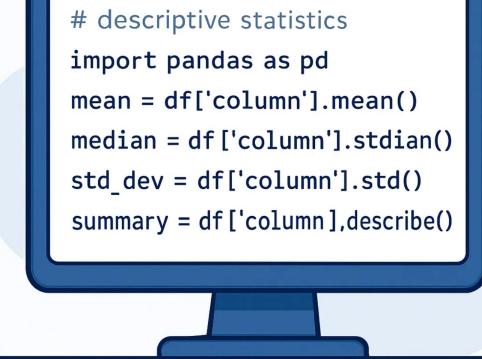
Análise Exploratória em Python

Estatísticas Descritivas Básicas

```
# Calculando estatísticas básicas  
media = df['idade'].mean()  
mediana = df['idade'].median()  
desvio_padrao = df['idade'].std()  
minimo = df['idade'].min()  
maximo = df['idade'].max()  
resumo = df.describe() # Resumo completo
```

- 1 Use o método `.mean()` para calcular a média de uma coluna numérica, como idade ou número de casos.
- 2 Calcule a `mediana` com `.median()` para entender o valor central, especialmente útil quando há valores extremos.
- 3 Obtenha o `desvio padrão` com `.std()` para medir a dispersão dos dados em relação à média.
- 4 Use o método `.describe()` para obter um resumo completo com contagem, média, desvio padrão, mínimo, quartis e máximo.

```
# descriptive statistics  
import pandas as pd  
mean = df['column'].mean()  
median = df['column'].stdian()  
std_dev = df['column'].std()  
summary = df['column'].describe()
```



Dica:

Compare diferentes variáveis ou grupos para identificar disparidades. Por exemplo, compare a média de idade entre diferentes regiões ou a taxa de vacinação entre diferentes grupos demográficos.

Atividade B: Criação de Visualizações Básicas

☰ Visualizando os Dados

Após calcular estatísticas descritivas, vamos criar visualizações para entender melhor os padrões nos dados:

☰ Gráfico de Barras

Uso: Comparar categorias (casos por região)

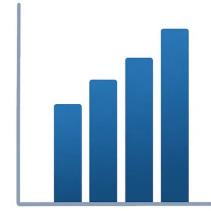
```
import matplotlib.pyplot as plt
plt.figure(figsize=(8, 5))
df['regiao'].value_counts().plot(kind='bar')
plt.title('Casos por Região')
plt.xlabel('Região')
plt.ylabel('Casos')
plt.show()
```

↳ Gráfico de Linha

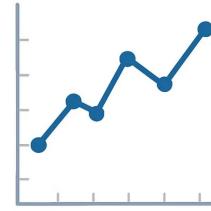
Uso: Evolução temporal (casos por mês)

```
# Agrupando por data
dados_tempo = df.groupby('data')['casos'].sum()
plt.figure(figsize=(8, 5))
dados_tempo.plot(kind='line') plt.title('Evolução de Casos')
plt.xlabel('Data')
plt.ylabel('Casos')
plt.show()
```

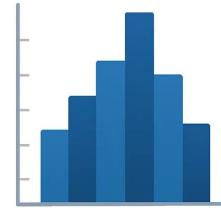
BAR CHART



LINE CHART



HISTOGRAM



↳ Histograma

Uso: Distribuição de variáveis numéricas (distribuição de idades)

```
plt.figure(figsize=(8, 5))
plt.hist(df['idade'], bins=15, edgecolor='black', alpha=0.7)
plt.title('Distribuição de Idades')
plt.xlabel('Idade')
plt.ylabel('Frequência')
plt.show()
```

💡 Dicas para Visualizações Eficazes:

Inclua títulos e rótulos de eixos claros

Escolha cores acessíveis para pessoas com daltonismo

Mantenha os gráficos simples e focados na mensagem principal

Atividade C: Estudo de Caso COVID-19

Estudo de Caso Específico

Nesta atividade, vamos analisar um exemplo concreto: a pandemia de COVID-19. Utilizaremos dados sobre a pandemia para discutir os desafios de trabalhar com esses dados.

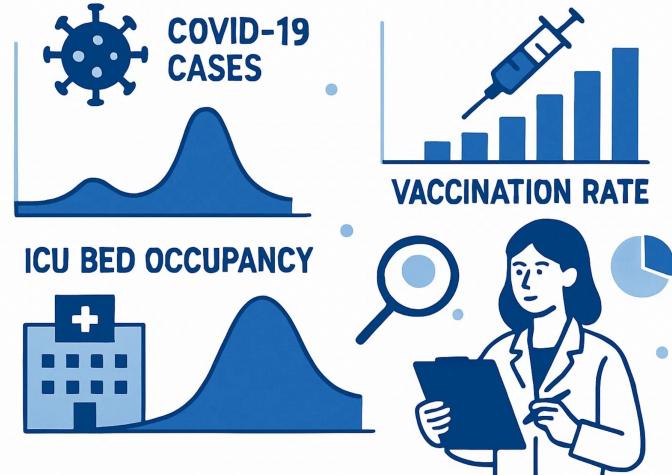
* Dados da COVID-19 para Análise

-  Casos por dia e região
-  Taxa de vacinação
-  Ocupação de leitos UTI
-  Índice de isolamento

Discussão em Grupo

⚠ Desafios dos Dados

-  **Dados Faltantes:** Como lidar quando um estado não reporta dados por alguns dias?
-  **Confiabilidade:** Como avaliar se os dados foram coletados uniformemente?
-  **Tomada de Decisão:** Como usar dados para decisões sobre medidas sanitárias?
Cada grupo deve discutir esses desafios e propor possíveis soluções. Ao final, compartilhem suas conclusões com a turma.



Atividade D: Elaboração de Relatório/Resumo

Elaboração de Relatório Baseado em Dados

Nesta atividade, cada grupo deve elaborar um relatório conciso baseado nas análises realizadas nas atividades anteriores. O relatório deve ser estruturado e comunicar claramente os insights obtidos dos dados.

Estrutura do Relatório

- 1 Título e Objetivo:** Defina claramente o que está sendo analisado e qual o propósito da análise.
- 2 Descrição dos Dados:** Explique a fonte dos dados, o período coberto e as principais variáveis analisadas.
- 3 Análise Estatística:** Apresente as principais estatísticas descritivas e visualizações criadas, explicando o que elas revelam.
- 4 Insights Principais:** Destaque de 3 a 5 descobertas importantes que emergiram da análise dos dados.
- 5 Limitações:** Discuta as limitações dos dados ou da análise que podem afetar as conclusões.
- 6 Recomendações:** Sugira ações ou políticas baseadas nos insights obtidos.



Dica:

Pense como um tomador de decisão: que informações seriam mais úteis para alguém que precisa decidir sobre alocação de recursos ou implementação de políticas públicas?

- Seja conciso e objetivo, evitando jargões técnicos desnecessários
- Inclua visualizações que apoiem suas conclusões
- Conecte os dados às implicações para políticas públicas ou saúde
- Considere diferentes audiências que poderiam utilizar seu relatório

Atividade E: Apresentação e Feedback

Compartilhando Descobertas

Apresentação dos Resultados

Cada grupo terá 5 minutos para apresentar suas análises, insights e recomendações baseadas nos dados públicos que exploraram.

- 1 Contextualize os dados:** Explique brevemente a fonte dos dados e por que eles são relevantes para políticas públicas ou saúde.
- 2 Mostre visualizações-chave:** Apresente 2-3 gráficos que melhor ilustram os padrões ou tendências que você identificou.
- 3 Compartilhe insights:** Quais foram as descobertas mais surpreendentes ou significativas? O que esses dados revelam sobre a realidade?
- 4 Proponha recomendações:** Com base nos dados, que ações ou políticas poderiam ser implementadas para melhorar a situação?

Feedback Construtivo

Após cada apresentação, os outros grupos terão 2 minutos para oferecer feedback construtivo.



💡 Orientações para Feedback Eficaz

- ✓ Seja específico sobre o que funcionou bem e o que poderia ser melhorado
- ✓ Faça perguntas que ajudem a aprofundar a análise
- ✓ Sugira outras perspectivas ou variáveis que poderiam ser consideradas
- ✓ Comente sobre a clareza das visualizações e a força dos argumentos

Reflexão Ética e Social

Dimensões Éticas da Ciência de Dados

A ciência de dados aplicada a políticas públicas e saúde traz consigo importantes responsabilidades éticas e sociais. Ao final desta aula, é fundamental refletirmos sobre o impacto dessas tecnologias na sociedade.

Perguntas para Discussão:

- ① **Privacidade:** Como equilibrar o uso de dados pessoais de saúde para o bem público com o direito à privacidade dos cidadãos?
- ② **Equidade:** Como garantir que algoritmos e análises de dados não perpetuem ou ampliem desigualdades sociais existentes?
- ③ **Transparência:** Qual o nível adequado de transparência que governos devem ter ao utilizar dados para tomar decisões que afetam a população?
- ④ **Responsabilidade:** Quem deve ser responsabilizado quando decisões baseadas em dados levam a consequências negativas não previstas?



"Com grande poder vem grande responsabilidade. Como cientistas de dados, temos o poder de influenciar políticas que afetam milhões de vidas - nossa responsabilidade ética é proporcional a esse impacto."