

# O Mestre dos Dados

## Ascensão à Excelência



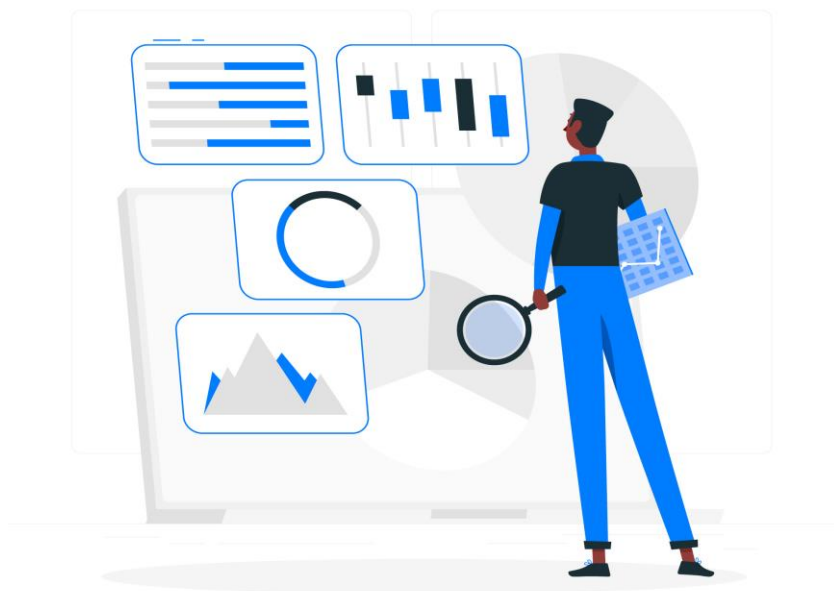
Um guia com as melhores ferramentas

*Isadora Bagatini*

# Análise de Dados

Guia com as principais ferramentas e exemplos práticos.

Os dados estão em todos os lugares e, com o boom da tecnologia, eles se tornaram parte fundamental para tomadas de decisão inteligentes. Seja para entender as preferências dos clientes ou prever tendências futuras, a análise de dados é uma habilidade que pode te levar longe. Nesse ebook iremos demonstrar algumas ferramentas extremamente úteis que te ajudam a manipular os dados e aprender com eles.



01

# Python: Um Superpoder em Análise de Dados



# Python

Python é como um canivete suíço para análise de dados. Você pode fazer praticamente de tudo com as bibliotecas certas, como Pandas para manipular dados e Matplotlib para criar gráficos.

Aqui está um exemplo básico de como carregar dados de venda em CSV e fazer algumas análises:

```
import pandas as pd

# Carregando os dados de vendas de um arquivo CSV
dados = pd.read_csv('vendas.csv')

# Mostrando as primeiras linhas para ter uma noção dos dados
print(dados.head())

# Calculando a média das vendas
media_vendas = dados['vendas'].mean()
print(f'Média das vendas: {media_vendas}')

# Criando um gráfico de barras das vendas por mês
import matplotlib.pyplot as plt

dados_por_mes = dados.groupby('mes')['vendas'].sum()
dados_por_mes.plot(kind='bar')
plt.title('Vendas por Mês')
plt.xlabel('Mês')
plt.ylabel('Vendas')
plt.show()
```

02

# R: Estatísticas e Gráficos Sem Dor de Cabeça



# R

R é uma linguagem super popular para quem curte estatísticas e visualização de dados. Com pacotes como dplyr e ggplot2, trabalhar com dados fica bem mais tranquilo.

Vamos supor que você tenha um data frame com dados de uma pesquisa. Aqui vai um exemplo de como calcular a média de uma variável e fazer um gráfico de dispersão:

```
# Carregando pacotes
library(dplyr)
library(ggplot2)

# Carregando dados de um data frame
dados <- read.csv('pesquisa.csv')

# Calculando a média de uma variável
media_var <- mean(dados$variavel)
print(paste('Média da variável:', media_var))

# Fazendo um gráfico de dispersão
ggplot(dados, aes(x=variavel_x, y=variavel_y)) +
  geom_point() +
  labs(title='Gráfico de Dispersão', x='Variável X', y='Variável Y')
```



# 03


## **SQL: O Básico para Manipular Dados**



# SQL

SQL (Structured Query Language) é a linguagem padrão para bancos de dados relacionais. Saber SQL é essencial para extrair dados e se comunicar com bancos de dados.

Suponha que você queira saber as vendas totais por produto de um banco de dados. Aqui está um código SQL que faz isso:



```
-- Pegando vendas totais por produto
SELECT produto, SUM(vendas) AS vendas_totais
FROM vendas
GROUP BY produto
```



# AGRADECIMENTOS



# Obrigada por ler até aqui!!!

Esse ebook foi gerado por IA e diagramado por humano.

O conteúdo foi criado apenas para fins didáticos, como desafio proposto no módulo de Introdução à Engenharia de Prompts com ChatGPT, do Bootcamp Santander 2024 - IA para Devs em parceria com a plataforma DIO. O foco do desafio foi a utilização de IA para gerar imagens, conteúdos e ideias criativas, e foi revisado e verificado por humano, portanto, está sujeito a erros e não deve ser utilizado como base de aprendizagem.

## Sobre a autora (humana):



<https://github.com/IsahBag>