

dnc



# **DATA CLEANING E DATA WRANGLING**

**EXERCÍCIO**



Data Science & Machine Learning

## EXERCÍCIO

# LIMPEZA E ESTRUTURAÇÃO

Coloque o conhecimento  
em prática!



# DATA WRANGLING E DATA CLEANING - PARTE 1

---

Se você fosse gerente em uma empresa, o que você faria de diferente caso descobrisse que alguns de seus funcionários estão pensando em sair? Daria um aumento, buscaria contratar outro ou tentaria ver como melhorar? De toda forma, ter essa informação seria útil e evitaria muitos problemas!

Neste exercício simularemos a etapa de Data Preparation através da limpeza e estruturação dos dados que servirão de insumo para a construção do modelo preditivo mencionado acima.

## INSTRUÇÕES:

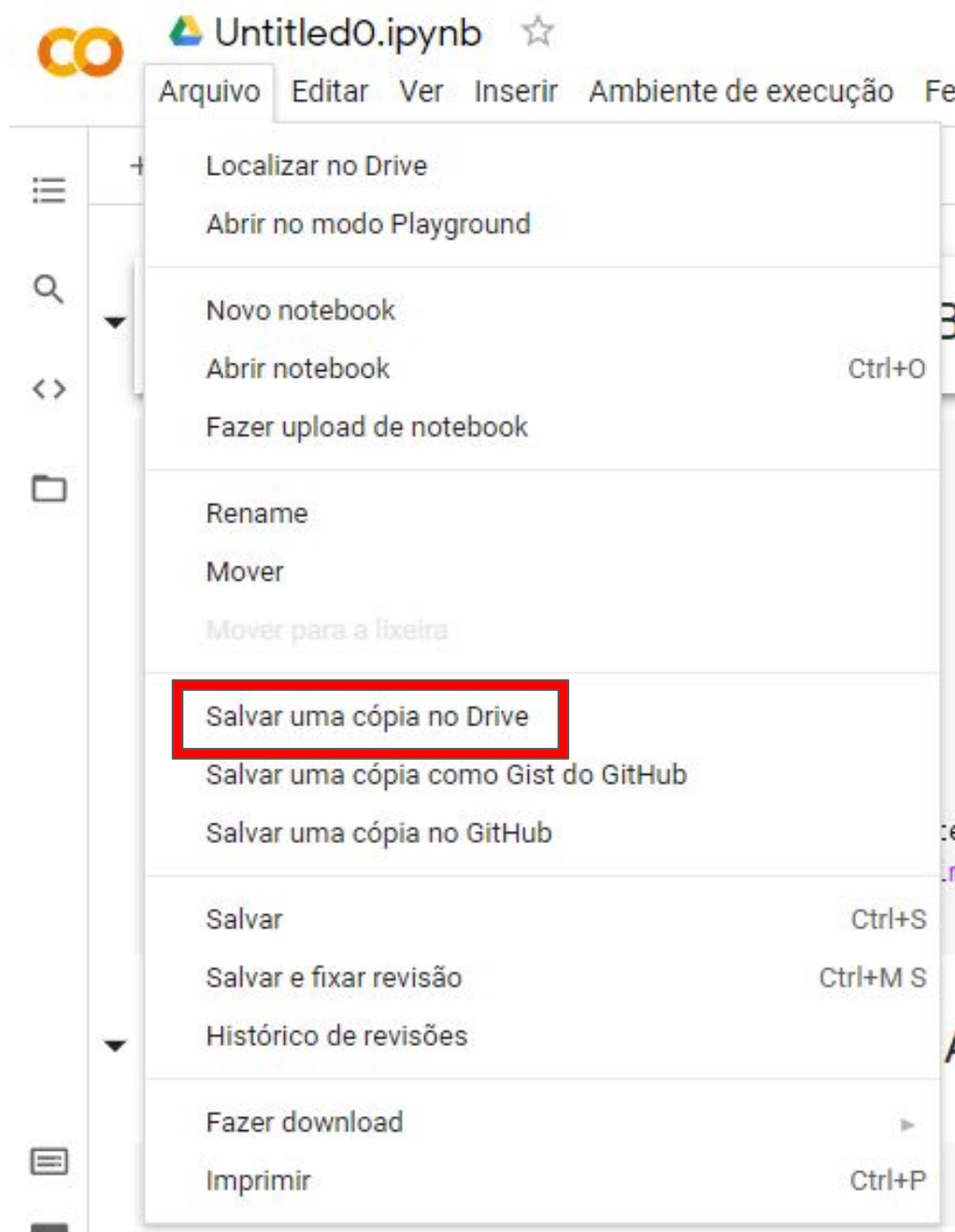
- Faça o download da base de dados anexado no exercício ou através deste [LINK](#).
- Acesse o Collab gabarito neste [LINK](#) e faça uma cópia para seu Collab pessoal caso desejar.



# REGRESSÃO PARTE 1

---

Como Fazer uma cópia para meu Collab pessoal?





# LIMPEZA E ESTRUTURAÇÃO PARTE 1

---

## ETAPA 1: Carregamento de Dados

Importe as bibliotecas necessárias para fazer a análise.

Carregue a base de dados disponibilizada no link do slide 3.

## ETAPA 2: Entenda seus dados

Interprete as features do dataset para se familiarizar com as informações que podem ser extraídas dessa base e faça a limpeza dos dados utilizando as técnicas aprendidas em aula.

As perguntas abaixo podem auxiliar na construção da resolução.

“Qual a sua variável target?”

“Quais as colunas do seu dataset que podem te ajudar a prever a informação desejada?”

“Existem outras informações (talvez na internet) que poderiam te ajudar a prever?”

“Como os 4 datasets se relacionam? Quais as chaves de cada um?”

“Qual o tipo de variável que você tem?”

“Qual a faixa de valores de cada uma delas?”

dnc



Data Science & Machine Learning

**#HARDWORK**

Esperamos que o material  
tenha sido útil para você!  
Foco nos estudos!