

Pedro Duarte Faria

Economista e Engenheiro de Dados

Belo Horizonte, Brasil | Telefone: +55 (31) 98512-4294 | Email: pedropark99@gmail.com

[Website pessoal](#) | [GitHub](#) | [Linkedin](#) | [Twitter](#)

1 Educação

Universidade Federal de Ouro Preto - UFOP

B.A. Economia, Fevereiro 2022

Universidade Federal de Minas Gerais - UFMG

Intercâmbio acadêmico - Economia, Fevereiro 2019 - Novembro 2020.

2 Experiências profissionais

Engenheiro de Dados - [Take Blip](#), Fevereiro 2023 - *presente*

Eu desenvolvo e mantenho pipelines de dados para processos de ingest e ETL, com o objetivo de fornecer infraestrutura de dados para o departamento de Customer Success da empresa. Para isso, utilizo principalmente **Python**, **Apache Spark**, **Databricks** e **Azure Data Factory**. Eu também desenvolvo novas soluções de dados para a empresa que aprimoram processos internos e que tragam melhorias de vida para os meus colegas de trabalho.

Analista de dados - [Take Blip](#), Maio 2021 - Fevereiro 2023

Eu ajudo clientes a tomarem melhores decisões com dados, e a desenvolverem os seus negócios com o nosso produto (plataforma Blip). Eu também desenvolvo muitas pipelines de dados (com **pyspark** e SQL) e dashboards (com Power BI). Meus clientes incluem a [Sodexo](#), [Banco BMG](#) e o [Banco Pan](#);

Estágio no Departamento de Estatísticas e Informações - [Fundação João Pinheiro](#), Junho 2019 - Março 2021

Eu ajudei vários pesquisadores a automatizarem suas rotinas e pipelines de dados com a linguagem R. Esse estágio teve um impacto muito grande sobre o meu livro e meus outros projetos pessoais.

3 Projetos pessoais

3.1 Projetos de livro

- [Introduction to **pyspark**](#). Primeira edição. Um livro introdutório (em inglês) sobre o pacote de **Python** chamado **pyspark**.
- [Introdução à Linguagem R: seus fundamentos e sua prática](#). Quarta edição. Um livro introdutório (em português) sobre a linguagem de programação **R**.

3.2 Projetos open-source mais recentes

- **figma**: um pacote de R que fornece um cliente para a API da plataforma Figma.
- **spark_map**: um pacote de Python que te ajuda a aplicar funções sobre várias colunas de um Spark DataFrame.
- **ggfunnel**: um pacote de R que te ajuda a produzir gráficos de funil semelhantes aos disponíveis no Power BI.