

# Ricardo Morsoleto

Desenvolvedor Python | Data Analytics • Passos, Minas Gerais, Brasil  
Contato: morsoleto.dev@gmail.com • +55 (17) 98115-8281  
LinkedIn: in/morsoletodev • GitHub: <https://github.com/morsoletodev>

## RESUMO

Cientista de Dados com experiência no desenvolvimento de pipelines de dados e aplicação de técnicas de machine learning para análise de dados. Possuo conhecimentos em Python, PostgreSQL, Pandas e algoritmos de mineração de regras de associação, como o Apriori. Atuei em projetos que envolvem desde a ingestão e transformação de dados até o deploy de resultados. Comprometido em criar soluções flexíveis e eficientes que atendam às necessidades do domínio de aplicação, garantindo qualidade e prontidão dos dados para análise.

## TECH STACK

**Programação:** Python (Pandas, Scikit-learn, Keras, TensorFlow), SQL (PostgreSQL)  
**Ferramentas:** Visualização de dados (Matplotlib, Seaborn, Power BI), Pipeline de dados  
**Conhecimentos em IA/ML:** Machine Learning (Classificação, Regressão, Clusterização, Regras de Associação), Deep Learning  
**DevOps:** Git, GitHub, Poetry, Docker  
**Soft Skills:** Comunicação Ativa, Resolução de Problemas, Trabalho em equipe  
**Idiomas:** Português (Nativo), Inglês (Fluente)

## EXPERIÊNCIA ACADÊMICA

Bolsista de Pesquisa (CNPq): Cientista de Dados Apr 2024 – Aug 2025

- Projeto: Uso da inteligência artificial para a classificação de risco de mulheres negras nas assistências pré-natal e parto no contexto brasileiro
- Atuei como Cientista de Dados, sendo responsável diretamente pela Coleta dos Dados, Implementação de Modelos Preditivos e Análise Estatística dos Resultados.
- Desenvolvi pipelines ETL para lidar com os dados ruidosos do DATASUS. Tornado possível a análise dos bancos SIM e SINASC referentes aos anos de 2012 a 2023.
- Desenvolvi também modelos supervisionados (XGBoost, KNN, RL e etc.) e não supervisionados (Kmeans, PCA, Apriori e etc.) para analisar os dados coletados, sendo possível obter um panorama abrangente da saúde da mulher e do recém nascido. As análises criadas foram publicadas jornais nacionais e internacionais.
- Link para implementações públicas: <https://github.com/GOPAD-Datasus>

## EDUCAÇÃO

Instituto Federal do Sul de Minas, Passos, Minas Gerais, Brasil Feb 2022 – Dec 2025

- Bacharelado em Ciência da Computação
- Índice de Rendimento Acadêmico: 9.3/10

## PROJETOS

**Pipeline de Dados para a Saúde Pública Brasileira**

- Desenvolvimento de um pipeline de dados em Python para processar e padronizar dados do Sistema Único de Saúde (SUS), utilizando arquivos YAML para configuração e integração com PostgreSQL.
- Tecnologias: Python, PostgreSQL, PyYAML, Pandas.
- Resultado: Solução flexível que permite ingestão em banco de dados, geração de arquivos Parquet e adição simplificada de novas etapas de processamento.
- Repositório: <https://github.com/GOPAD-Datasus/ETL-DataSUS>

**Mineração de Regras de Associação em Dados de Saúde Materno-Infantil (SINASC)**

- Aplicação do algoritmo Efficient Apriori para identificar fatores maternos associados a desfechos pré-natais adversos.
- Colaboração com equipe de saúde para garantir que as necessidades do domínio fossem atendidas.
- Tecnologias: Python, Machine Learning, Pandas.
- Publicação relacionada: Congresso Internacional DOHaD (2025).
- Repositório: [https://github.com/GOPAD-Datasus/Paper\\_E-Dohad-Apriori](https://github.com/GOPAD-Datasus/Paper_E-Dohad-Apriori)

**Análise de Riscos Pré-Natais com Mineração de Regras de Associação**

- Mineração de regras de associação em dados do SINASC para identificação de fatores de risco pré-natais.
- Trabalho publicado em livro acadêmico pela Springer (2025).
- Tecnologias: Python, Machine Learning.
- Repositório: [https://github.com/GOPAD-Datasus/Paper\\_B-ICISMED-Apriori](https://github.com/GOPAD-Datasus/Paper_B-ICISMED-Apriori)
- Link da publicação: <https://link.springer.com/book/9783031996986>