




# Michel Souza Santana

*Data Engineer*

 Amargosa, Bahia, Brazil

 michellss.py@gmail.com

 [linkedin.com/in/michel-souza-santana](https://www.linkedin.com/in/michel-souza-santana)

 WhatsApp: (47) 9 9665-8601

---

## Objetivo Profissional

Busco uma posição de Engenheiro de Dados Júnior/Pleno onde eu possa aplicar minhas habilidades técnicas e experiência em ambientes de cloud computing, como GCP e Azure, para resolver desafios e contribuir para o sucesso da equipe. Desejo continuar crescendo profissionalmente, ter vivência em projetos desafiadores e impulsionar melhorias contínuas nos processos de dados.

---

## Perfil Profissional

Engenheiro de Dados com mais de 2 anos de experiência na Dataside e Fleury, com um histórico em modelagem de dados, ETL/ELT e análise de dados. Experiência comprovada em projetos para clientes de grande porte como Minerva Foods, Grupo Jacto, Sisprime, Agricopel e Lavoro, demonstrando habilidades na otimização de processos e na implementação de soluções em ambientes de cloud computing (GCP e Azure). Proativo, com excelentes habilidades interpessoais.

---

## Portfólio

<https://www.michelsantana.com/>

---

## Competências

- **Ferramentas e Tecnologias:** Dataflow, Apache Beam, Databricks, Azure, DBT, Apache Airflow, Docker, PostgreSQL, Python, PySpark, Pandas, SQL, Data warehouse, Data Lake, Data Lakehouse, Big Data, GCP, BigQuery, Modelagem de dados, Git, GitHub, ETL, ELT, Apache Hop, Linux, Pentaho, MySQL.
  - **Áreas de Conhecimento:** Arquitetura de Data Warehouse, Análise de Big Data, Qualidade de Dados, Ciência de Dados, Cloud Computing, Modelagem de Dados para Data Warehouse, Análise de Dados nos Negócios, Engenharia de Dados.
-

## Experiência Profissional

### Data Engineer

*Dataside Consultoria em Dados e IA*

05/04/2023 - Presente

### Data Engineer

*Grupo Fleury Medicina, saúde e laboratórios*

07/10/2024 – Presente

- **Desenvolvimento e Implementação:** Implementação de pipelines de dados em grande escala para clientes como Minerva Foods e Grupo Jacto, utilizando Dataflow com Python/Apache Beam e Databricks com SparkSQL e PySpark. Processamento de grandes volumes de dados com mais de 1 milhão de linhas.
  - **Automação e Otimização:** Desenvolvimento de scripts para automação de processos de ingestão e transformação de dados, aumentando a eficiência das operações.
  - **Gerenciamento de Workflows:** Implementação e gerenciamento de workflows com Apache Airflow, garantindo agendamento e execução eficientes de processos.
  - **Armazenamento e Processamento:** Ingestão e processamento de dados no Google Cloud Storage, com uso de Dataflow. Transformação de dados provenientes de múltiplas fontes como SQL Server, S3, Oracle, MySQL, CSV e JSON.
  - **Modelagem e Análise:** Desenvolvimento de modelos de dados e tabelas otimizadas no BigQuery, aplicando melhores práticas em particionamento e clusterização. Modelagem Star Schema para suportar análises avançadas.
- 

## Formação Acadêmica

- **Análise e Desenvolvimento de Sistemas**  
Universidade Estácio de Sá
  - **Cientista de Dados**  
Escola DNC
- 

## Certificações e Cursos

- Bootcamp Engenharia de Dados: Construa um Projeto Real
- Academy Accreditation - Databricks Lakehouse Fundamentals
- Modelagem de Dados para Data Warehouse
- Python & MySQL • SQL (Advanced)
- Big Data Fundamentos
- Google Cloud Fundamentals: Core Infrastructure

- Digital Transformation with Google Cloud
  - Introdução ao Power BI
  - DevOps Essentials
  - Apache Spark Programming with Databricks
  - Certificação Google Cloud Associate Engineer (GCP)
  - Git e Github - Essencial
  - Introduction to Apache Spark Architecture
  - Engenharia de Dados com Google Dataflow e Apache Beam na GCP
  - Docker Essencial para o Desenvolvedor
  - Domine Apache Airflow: Pipelines de Engenharia de Dados
- 

## Soft Skills

- **Trabalho em Equipe:** Proativo e sempre pronto para colaborar.
  - **Comunicação:** Habilidade em transmitir conceitos técnicos de forma clara para diferentes audiências, incluindo times de negócios e técnicos.
  - **Resolução de Problemas:** Foco em solução de problemas críticos em ambientes de alta pressão.
  - **Aprendizado Contínuo:** Sempre atualizado com as tendências de engenharia de dados e tecnologias emergentes, investindo constantemente em treinamentos e certificações para expandir meus conhecimentos.
- 

## Idiomas

- Português (Nativo)
- Inglês (Básico/Intermediário)