

# José Marcello Lopes de oliveira

## Arquiteto e/ou Engenheiro de Dados

Resido em São Paulo Capital região de Interlagos

+55 11 94028.8272 - [jm.lope@live.com](mailto:jm.lope@live.com) - [linkedin.com/in/josemarcellolopes](https://www.linkedin.com/in/josemarcellolopes)

**39 anos de experiência** em arquitetura e engenharia de dados, projeto e implemento a arquitetura de dados, visando produtos e serviços necessários, modelagem de dados e processos, governança, segurança (LGPD e GDPR) e escalabilidade, construções (codificação) e implementação de pipelines para dados estruturados, não estruturados, streaming e IoT, implementação e utilização de IA Generativa com linguagem natural para análise e predição dos dados.

Especialista em **AWS, Azure, GCP e Databricks**, com amplo conhecimento em bancos de dados **SQL, NoSQL** e bem como em **Python, Spark**. Sólida experiência em modelagem de dados **OLTP, OLAP e BigData**, implementação de soluções escaláveis, otimizadas a performance e baixo custo.

### Hard Skills

Sistemas Operacionais	Windows, Linux (Unix e derivados) e IOs
Clouds	AWS, Azure, GCP e <b>Databricks</b>
Bancos de Dados SQL	Oracle, SQL Server, MySQL, Postgres, DB2, Informix, Sybase, Teradata, Hive, RedShift e Aurora
Bancos de Dados NoSQL	MongoDB, DynamoDB, CosmosDB, SimpleDB, Redis, MemCached, ElastiCache, HBase e Zim
Big Data	Hadoop, HDFS, Sqoop, Nifi, Kafka e Airflow
Linguagens	Python, PySpark, SQL, T-SQL, PL-SQL e Shell Script
Inteligência Artificial	GPT, LLMs, NLP, LangChain, Hugging Face, TensorFlow, PyTorch, Transformers e Genie.
Modelagem	Big Data, OLTP, OLAP, ETL e ELT
Metodologias de Trabalho	Agile, Kanban, Scrum, DevOps, DataOps e CI/CD
Plataformas e Softwares DEV	Git, GitHub, GitFlow, Docker e Kubernetes

Formação Superior e Pós em Análise de Sistemas pela UNIP concluído no ano de 1993.

Dezembro/2023 Atual	<p><b>Grupo Petrópolis</b></p> <p>Uma das maiores cervejarias do Brasil, fabricando marcas renomadas como Itaipava, Crystal e Petra. A empresa atua na produção e distribuição de bebidas alcoólicas e não alcoólicas, mantendo um forte compromisso com inovação e qualidade no setor de bebidas.</p> <p><b>Arquiteto e Engenheiro de Dados</b></p> <p>AWS, Azure, Databricks, Python, PySpark, Airbyte e Airflow.</p> <p>Definição e implementação da arquitetura de dados utilizando AWS, Azure e Databricks.</p> <p>Implementações de processos de ELT e captura de dados, como também orquestração de processos de carga utilizando Airbyte e Airflow</p> <p>Desenvolvimento de pipelines envolvendo as três camadas (bronze, Silver e Gold), utilizando Python e PySpark, com latência batch e streaming.</p> <p>Definição e implementação da governança de dados utilizando Unity Catalog, LGPD e GDPR.</p> <p>Implementação de IA Generativa para análise dos dados, utilizando linguagem natural NLP.</p>
Janeiro/2021 Dezembro/2023	<p><b>Intellibrand by Ascential (Flywhell)</b></p> <p>Especializada em análise de dados e soluções de business intelligence para grandes corporações globais, incluindo Coca-Cola, Unilever e Pepsico.</p> <p><b>Engenheiro de Dados</b></p> <p>AWS, GCP, Databricks, Python, PySpark, Glue e Airflow.</p> <p>Desenvolvimento de pipelines envolvendo as três camadas (bronze, Silver e Gold), utilizando Python e PySpark, com</p>

	<p>latência batch e escalável para captura de milhões até bilhões de dados dia.</p> <p>Modelagem de dados voltado a Big Data para as soluções de captura e BI.</p> <p>Orquestração de processos de captura e transformação dos dados utilizando Airflow e Glue.</p>
<p>Novembro/2019 Janeiro/2021</p>	<p><b>Hospital Sírio-Libanês</b></p> <p>Uma das principais instituições de saúde da América Latina, reconhecida por seu padrão de excelência em atendimento médico, pesquisa e inovação em saúde digital.</p> <p><b>Engenheiro de Dados</b></p> <p>AWS, Azure, Databricks, Python, PySpark e Airflow.</p> <p>Definição de estrutura dos dados para processamento do sistema de detecção de possibilidades do paciente contrair sepse.</p> <p>Desenvolvimento de pipelines utilizando Python e PySpark para captura de dados estruturados, não estruturados e IoT.</p> <p>Orquestração do processo de ELT utilizando Airflow.</p> <p>Interação com a equipe de Cientistas de Dados para geração de datasets voltados para pesquisa das doenças.</p>
<p>As demais experiências antes de novembro de 2019 se encontram mais detalhadas no meu LinkedIn no endereço <a href="https://www.linkedin.com/in/josemarcellolopes">linkedin.com/in/josemarcellolopes</a></p>	