

Victor Ramos Pasqualini

(11) 99735-6148 - victor.pasqualini@outlook.com – 27/06/2002

linkedin.com/in/victor-pasqualini

github.com/VictorPasqualini

Resumo

Engenheiro de Dados com 5+ anos de experiência construindo pipelines robustos em arquiteturas cloud. Especializado em migração de sistemas legados, processamento ETL/ELT com Spark, e infraestrutura de dados em AWS/GCP. Expertise em tecnologias open source (Spark, Kafka, Airflow) com foco em qualidade, integridade e disponibilidade dos dados.

Experiência

ago 2025 – jul 2026 Especialista de Dados

Serasa Experian pela ACT Digital

Liderou migração complexa de pipeline mainframe (Cobol/DB2) para AWS processando ingestão massiva de arquivos posicionais diários. Implementado validação e qualidade de dados com geração de tabelas consolidadas (bronze/silver/gold) para ML e produtos diversos. Stack: Spark Scala/Python, DuckDB, Pandas, Polars, PySpark, EMR, EC2, Lambda, SQS, MSK, MWA, S3, Iceberg, DeltaLake, Athena, RDS/Aurora, DynamoDB, DocumentDB, GlueCatalog. Observabilidade com Datadog, Grafana e CloudWatch. Deploy em produção com dualrunning entre ambientes. Integrações via EventBridge, APIs REST e IaC com CloudFormation.

abr 2025 - mai 2026 Engenheiro de Dados III

Vert Capital pela Stech Soluções

Mantive ecossistema Databricks com DLT, UnityCatalog e DeltaLake garantindo SLA 90% de disponibilidade. Otimizei pipelines batch PySpark reduzindo tempo de processamento em 40% através de query optimization. Implementei novo pipeline end-to-end utilizando PySpark, Kafka e APIs. Stack: Databricks, DLT, UnityCatalog, PySpark, DeltaLake, Kafka, PostgreSQL, Django, Metabase, PowerBI. Colaborei com time de Analytics para confiabilidade de dados.

nov 2024 - mai 2025 Engenheiro de Dados II

Stech Soluções

Desenvolvido pipeline GCP end-to-end focado em ingestão de dados de múltiplas fontes. Implementado validações de qualidade de dados com DuckDB em Cloud Functions. Stack: GCP (Cloud Storage, BigQuery, Compute Engine, Cloud Functions, Composer), Python, DuckDB. Automatizado processos com Composer (Apache Airflow) reduzindo intervenções manuais.

ago 2022 - ago 2025 Engenheiro de Dados I

SPC Brasil

Arquiteto de Data Lake em Hadoop/AWS (20+ TB diários) com liderança de migração on-premise para AWS. Projetos de qualidade de dados, geração de leads, pipelines de crédito e alimentação de ML. Processamento incremental de arquivos XML/CSV on-premise e em AWS. Stack: Spark (Scala/PySpark/Java), Hadoop, Airflow, Kafka, NiFi, CtrlM, RabbitMQ, AWS (EC2, EMR, Glue, S3, Athena, CloudWatch), Hive, Impala, Sqoop, Terraform. Dados: MySQL, MongoDB, Oracle, DuckDB. BI: PowerBI, Elasticsearch. IaC com Terraform, orquestração robusta com 100+ jobs diários, APIs REST, CloudWatch monitoring.

set 2021 - ago 2022 Estagiário em Engenharia de Dados SPC Brasil

Manutenção do Data Lake em Hadoop com limpeza de dados e compactação de partições. Desenvolvimento de scripts Python/Scala para validação de integridade e reprocessamento histórico. Suporte à infraestrutura de Spark em ambientes staging e produção.

Formação Acadêmica

2025 – 2026 Inbix Academy

MBA - IA para Inovação

2024 – 2025 FIAP

MBA - Engenharia de Dados

2020 - 2022 FATEC São Paulo

Tecnólogo, Análise e Desenvolvimento de Sistemas

2017 - 2019 ETEC Jardim Ângela

Técnico de Informática Integrado ao Médio

Idiomas

Inglês (nível avançado)

Tecnologias Conhecidas

Linguagens • Python (Pyspark, Pandas, Polars, DuckDB), Scala, Java, SQL, Cobol, Rust

ETL/ELT • Spark (Batch/Streaming), Glue, EMR, EC2 (spot), Lambda, Sqoop, DataFusion Comet, Databricks, Snowflake, Compute Engine, Cloud Functions, Incremental, CDC

Mensageria • Kafka, SQS, MSK, RabbitMQ, Redpanda

Orquestração • Airflow, MWA, EventBridge, Step Functions, Composer, CtrlM, NiFi, Databricks DLT, Databricks Workflows

Storage e Lakehouse • S3, GlueCatalog, Iceberg, Hadoop, Kudu, BigQuery, Cloud Storage, DeltaLake, UnityCatalog, Arquitetura Medalhão

Banco de Dados • RDS/Aurora, PostgreSQL, MySQL, DynamoDB, MongoDB, DocumentDB, Oracle, DB2

Infra • Terraform, Terragrunt, CloudFormation

Analytics e Observabilidade • Datadog, Grafana, Metabase, Power BI, BigQuery, Hive, Impala, Elasticsearch, CloudWatch, Athena

Outros • CI/CD (Cockpit), git (Gitlab, Bitbucket), Linux (Shell Script, Bash Script)

Certificações

MongoDB SI Associate (2025)