Helio Ribeiro – Senior Data Engineer

helioribeiropro@gmail.com | LinkedIn | GitHub | Portfolio |

SUMMARY

- Google Cloud Professional Data Engineer with 5+ years of experience in designing and implementing data pipelines.
- Skilled with Machine Learning algorithms and Generative AI / NLP for text and code-based solutions.
- Demonstrated ability to work cross-functionally with stakeholders to deliver high-quality data solutions.
- Excellent communicator fluent in English (C2), Portuguese (Native) and Spanish (C2). Currently studying Dutch (A2).

TECHNOLOGIES AND LANGUAGES

Languages: Python, R, SQL.

Technologies: Google Cloud Platform (Cloud Storage, Cloud Functions, Cloud Run, BigQuery, Cloud SQL, Pub/Sub, Dataflow, Dataproc, Data Fusion, Vertex AI, Dialogflow, Looker, Dataplex), Azure (Blob, ADLS, Azure Data Factory, Azure Data Studio, Event Hub, Functions, Fabric), AWS (S3, Athena, Redshift, EMR, Glue, RDS, ECR, Managed Airflow, Lambda, Step Functions), LangSmith (LangChain, Evals), Apache Spark (NLP, ML, SQL and Streaming), Apache Kafka, LLMs (Gemini, OpenAI), RAGs, ChromaDB, Pinecone, Airflow, Snowflake, Databricks, dbt, Jenkins, Terraform, Git, Power BI, Tableau, ETL / ELT, MySQL, PostgreSQL, Docker, Excel / Google Sheets, CI/CD, DevOps, MLOps, REST API, FastAPI, Flask, n8n, Agile.

EMPLOYMENT & EXPERIENCE

Senior Data Engineer   Eindhoven, The Netherlands
Tech Seals   *May 2025 – Present*

Senior Data Engineer   Sao Paulo, Brazil
Minsait   *Jun 2024 – May 2025*
- Supervised a team of four Data Engineers.
- **Migrated terabytes** to AWS + Snowflake with PySpark and SQL, **saving up to 1.5MM USD / year** on data processing.
- **Achieved cost reduction** of LLM-based chatbots by **132.9x per message** and a **6.8x reduction in time-to-market**.
- Deployed **AI chatbot handling 120k monthly messages,** with an integrated data pipeline on **Microsoft Fabric**.

Data Engineer   Sao Paulo, Brazil
uCloud, the Global Google Partner   *Nov 2023 – Jun 2024*
- Deployed **AI Agents handling ~15k messages/month**, including voice capabilities, reducing call center workload.
- Built **10+ Looker reports for financial and CRM analytics**, **saving ~12 management hours/week**.
- **Accelerated NLP workflow by over 1000x by integrating Gemini**, transforming tasks from days into minutes.

Data Engineer   Remote
Genova   *Jun 2020 – Nov 2023*
- Deployed **APIs, Text 2 SQL solutions**, Python functions on AWS Lambda and GCP Cloud Functions for data tasks.
- **Built 15+ dashboards on PowerBI** and PowerQuery, mostly saving uncountable management hours on reports.

EDUCATION

UMC – Universidade de Mogi das Cruzes   Mogi das Cruzes, Sao Paulo – Brazil
M. Sc. Biomedical Engineering   *Aug 2024 – Present (Remote since May 2025)*
- **RESEARCH:** Coronary Artery Disease (CAD) prediction and prevention by use of **Machine Learning**.

ESPM – Escola Superior de Propaganda e Marketing   Sao Paulo, Sao Paulo - Brazil
MBA - Marketing   *Mar 2022 – May 2025*

UMC – Universidade de Mogi das Cruzes   Mogi das Cruzes, Sao Paulo - Brazil
B. Sc. Electrical Engineering   *Feb 2014 - Jun 2021*