# GLEICE CHAVES

+1 (747) 220-4099 | gleice.schaves@gmail.com | Los Angeles, CA, USA | linkedin.com/in/gleiceschaves

## EDUCATION

**University of Southern California**													**August 2024 - May 2026**
*Master's, Data Science*
- Database Systems, Machine Learning, Data Mining, Research Methods, Applied Natural Language Processing

**Faculdade de Informática e Administração Paulista**									**January 2019 - December 2022**
*Bachelor's, Business Analytics*																*GPA: 3.7*
- Data Modeling, Data Warehousing, NoSQL/SQL, BI, Cloud Services, Data Governance, Business Development, Python

## PROFESSIONAL EXPERIENCE

**University of Southern California**													**Los Angeles, CA**
*Data Engineer Intern*																	*March 2025 - Present*
- Enhanced data insight accuracy by 30% through the implementation of an end-to-end NLP Q&A layer on a Power BI/SSAS tabular model, by exposing KPI/measure metadata and synonyms to Copilot, optimizing response quality for a 100-question benchmark.
- Standardized KPI definitions and documentation across 300+ tables by owning the enterprise BI/tabular semantic model utilizing Power BI/SSAS, Tabular Editor, and DAX, while authoring hardened SQL Server views to encode business rules.
- Improved query performance by 25% by diagnosing bottlenecks in the data model using DAX Studio, VertiPaq Analyzer, and SQL Profiler/Query Store to eliminate CROSSJOIN and bi-directional filter hotspots.
- Streamlined reporting consistency across models by designing and maintaining advanced SQL CTEs and parameterized views, which codified business rules for dimensions, hierarchies, and period-to-date calculations using reusable patterns.

**Diagnostics of America**														**Sao Paulo, Brazil**
*Data Engineer & Analytics*															*June 2021 - December 2022*
- Reduced billing process time from 20 days to hours by designing and deploying a federated data lake and warehouse using AWS (S3/Athena), Databricks, and BigQuery for 350 hospitals and labs.
- Enhanced data reliability and simplified analytics by developing pipelines with Alteryx and SQL, incorporating incremental loads, RI/outlier/duplicate checks, and lineage/audit logs.
- Delivered decision-ready reporting for C-suite by creating and delivering Power BI dashboards for service-line and patient-level costing, optimizing DAX/SQL models and queries, and establishing daily refresh SLAs.
- Standardized data schemas and pipelines across multiple entities to facilitate seamless data integration and analytics for approximately 350 hospitals and labs.
- Improved analytics capabilities within the organization by leveraging SQL and data visualization tools to identify gaps and recommend solutions for business intelligence.

**Link School of Business**														**Sao Paulo, Brazil**
*Business Analyst*																	*December 2019 - May 2022*
- Enhanced grant allocation efficiency by 30% through the development of a scoring model, executing market research using Excel with Power Query and Pivot Tables to identify high-potential startups.
- Maximized return on ad spend (ROAS) by reallocating marketing budget based on established ROI models, utilizing Excel for monthly forecast and pacing reports across channels including search, social, display, and email.
- Increased insights into marketing performance by creating comprehensive reports and data visualizations, leveraging Excel for monthly forecasting and analysis to adjust strategies based on CPA targets.

## PROJECTS & OUTSIDE EXPERIENCE

**ChatDB: Natural-Language Database Query System - USC**								**Los Angeles, CA, USA**
																					*2025 - Present*
- Built a Streamlit + LangChain text-to-SQL/Mongo app turns English into parameterized PostgreSQL SQL and MongoDB aggregations with schema introspection, LangSmith tracing/evaluation and optional Ollama local LLM

**Human-AI Trust Calibration - PI: Gale Lucas - USC Institute**							**Los Angeles, CA, USA**
*Research Assistant*																	*March 2025 - Present*
- Co-authored annotation codebooks; performed thematic coding and qualitative analysis; authored methods and documented inter-rater reliability (Cohen's k) to inform calibrated human–agent teaming

**RAG assistant for finance KPIs - USC**												**Los Angeles, CA, USA**
*Research Assistant*
- Built a semantic-aware RAG assistant for finance KPIs by exporting 300+ DAX measures from Tabular Editor, embedding in Chroma (sentence-transformers) with retrieval evaluation and dashboards

## SKILLS

**Skills:** Python, LLM, SQL, NoSQL, Machine Learning, Natural Language Processing (NLP), Apache Spark, AWS, Databricks, Power BI, R, Pytorch, Scikit-learn, PostgreSQL, MongoDB, Tableau, LangChain, MongoDB, Statistical Analysis, Statistics, ETL, ELT, MATLAB, Hadoop, data analytics, data pipeline, Java, Apache Hive, Tuning, Hadoop, Microsoft Azure, GCP, troubleshooting, pipelines, predictive models, Data Models, Data Mining, Spark, data warehouse, Scala, data solutions, business performance, SAP
**Languages:** Spanish, Portuguese, English