

# Phil Salazar

+1 (707) 853-3757 • philsalazar@proton.me • linkedin.com/in/phil-salazar • phil-salazar.github.io

Experienced data scientist with a robust business acumen and full-stack technical background, including AWS, GCP, Snowflake, DBT, Python, SQL, and React. Extensive working knowledge of LLMs, ML, and NLP, with training in statistical analysis, econometrics, and corporate finance at Harvard.

## WORK EXPERIENCE

### Grupago

09/2024 - Present

#### Data Scientist

- **Developed and deployed a comprehensive machine learning pipeline in Python**, SageMaker, and Redshift, including ETL and a robust ML credit risk model whose predictions outperform actuals by 4%.
- **Leveraged deep learning in Keras and Tensorflow** to extract predictive features from unstructured data, improving precision and recall by 11% and 5%, respectively.
- **Engineered an Agentic AI system using Large Language Models (LLMs)**, automating key credit analysis processes and providing intelligent decision capabilities.

### Bankaya

05/2023 - 08/2023

#### Credit Risk Data Analyst • Internship

- **Implemented a new credit risk strategy** to mitigate delinquency and fraud, using Snowflake, SQL, and DBT. Trained and deployed ML model in AWS and Python using payments data, resulting in a significant improvement to overall credit margin by more than \$1M annually.
- **Managed key stakeholders and negotiations** with alternative data vendors in Spanish, securing favorable pricing terms relative to usage volume in production models.

### GoodLeap

11/2019 - 11/2021

#### Business Intelligence Data Analyst

- **Ensured on-time delivery for 95% of high-priority business analytics reporting**, managing cross-functional technical requirements from business stakeholders and C-suite executives in Agile workflow.
- **Significantly contributed to production codebase for full-stack BI platform**, developing complex SQL queries, Python ETL pipelines in Databricks, and designing visualizations in React for over 100 dashboards.
- **Ensured software quality and maintainability and data governance** by writing hundreds of tests for all frontend components and maintaining comprehensive documentation for pipelines, data models, and dashboards.

### Amazon

06/2017 - 09/2018

#### Retail Vendor Manager

- **Grew \$20M portfolio by 23%** year over year, analyzing sales trends through financial models in Excel to extract insights for category growth.
- **Reduced inventory lead time by 12%**, collaborating cross-functionally with supply chain analysts using SQL to find bottlenecks, resulting in over \$7M additional revenue.

EDUCATION

Master in Public Administration in International Economics

Harvard University 08/2022 - 05/2024

- **Coursework:** Advanced Statistics, Econometrics, Advanced Microeconomic Theory & Analysis, Advanced Practical Data Science (Harvard Engineering), Deep Learning (MIT Sloan), Corporate Finance (MIT Sloan), Financial Services (HBS)
- **Course Assistant:** Advanced Microeconomic Theory & Analysis I, II (Prof. Luis Armona, Prof. Jie Bai)
- **Research Assistant:** Randomized Control Trial (RCT): Economic Impact of Trade Fairs in Chile (Prof. Jie Bai)

Bachelor of Science in Engineering

Stanford University 08/2012 - 06/2016

- Managing Officer, Society of Latino Engineers (SOLE)

PROJECTS

Randomized Control Trial (RCT): Economic Impact of Trade Fairs in Chile 08/2023 - 05/2024

Professor Jie Bai, Harvard University

- **Designed a randomized control trial (RCT)** experiment and performed econometric and statistical analysis in R to evaluate the economic impact of trade fairs in Chile, demonstrating a 7% increase in firm revenue one year later for participating firms compared to non-participants.

Prescribe.ai 08/2023 - 12/2023

Advanced Practical Data Science, Harvard University

- **Developed app end-to-end** to advise medical professionals in Latin America. Managed Docker environments and model deployment in Google Cloud, trained and tuned LLM using Hugging Face embeddings and Weights & Biases, and built frontend in React.

SKILLS

Cloud Platforms & Infrastructure

- AWS
- Databricks
- Docker
- GCP
- Prefect
- Terraform

Machine Learning

- Deep Learning
- Keras
- NLP
- Tensorflow

Programming

- Github
- Pandas
- Python
- R
- React
- scikit-learn
- SQL
- Unix