

# Thiago Miranda

📍 Dublin, Ireland ✉ thiagofm.pa@gmail.com ☎ +353834174030 🌐 in/thiagofmiranda

---

## SUMMARY

---

**Data Scientist with 6 years of experience**, applying skills in Python, R, machine learning, and statistics to drive business growth. Developed machine learning models for scoring, ranking, and content recommendations, achieving an 11% increase in paying user conversions, and a 30% boost in job applications via automated CV submissions. Collaborated with cross-functional teams and utilised tools like AWS SageMaker, Tableau, and Power BI to deliver actionable insights and enhance user engagement and conversion strategies.

---

## SKILLS

---

**Tech Skills:** R, Python, Jupyter, SQL (MySQL, SQL Server, Presto), ETL process, Google Analytics, Git / GitHub, AWS (SageMaker, S3, EC2, Lambda, QuickSight), Docker, BI/Visualization tools, Power BI, Tableau, SAS, JIRA, MS Office, agile methodologies.

**Data Science Skills:** machine learning models (XGboost, random forest, clustering, classification techniques, deep learning), modelling (feature engineering, evaluation, monitoring), natural language processing (NLP), large language models (LLMs, transformers), latent trait models (IRT).

**Data Analytics Skills:** data wrangling, data cleaning, data visualisation, statistical analysis, hypothesis testing, A/B testing, storytelling.

**Soft Skills:** Committed, ability to learn, decision making, problem-solving, innovative, data-driven, communication, collaborative, critical thinking, business acumen, adaptability, flexibility.

---

## EXPERIENCE

---

### Data Scientist

Catho

May 2019 - April 2024, São Paulo

- **Developed and supported machine learning models** for tasks such as scoring, ranking, and content recommendations, using Python, R, and AWS tools.
- Created a user conversion propensity model to increase lead conversions, leveraging Python and AWS SageMaker.
- **Increased paying user conversions by 11%** by identifying key features based on free user behaviours and recommending targeted features for paid users.
- **Boosted job applications from paying users by 30%** by automating CV submissions through a job-CV scoring model.
- Led A/B testing initiatives for product teams, optimising pricing strategies, business models, and new product launches with a focus on enhancing customer experience.
- **Built a data framework that automated the end-to-end reporting process**, from data extraction to delivering actionable business insights, using tools like Tableau, Power BI, and Miro to communicate insights clearly to stakeholders.
- Streamlined data analysis and reporting workflows by integrating diverse data sources, including MySQL, AWS S3, and Google Analytics.
- Performed in-depth **quantitative analysis across multiple data sources** to uncover marketplace insights and identify growth opportunities.
- **Collaborated with cross-functional teams**, including product, engineering, and marketing, to ensure data-driven decision-making.

## Data Analyst

Avalia Educational

August 2018 - April 2019, São Paulo

- **Conducted Item Response Theory (IRT) analyses** to evaluate item characteristics and test reliability.
- Designed and maintained databases for storing assessment data, ensuring data integrity and scalability.
- Collaborated with psychometricians and educational experts to refine assessment models and improve test accuracy.
- **Automated data extraction and processing pipelines**, reducing manual workload and improving efficiency.
- Provided insights and recommendations based on data trends and IRT model outputs to enhance decision-making.
- **Presented data findings and technical results** to both technical and non-technical stakeholders.

## Data Analyst

Vunesp Foundation

December 2017 - June 2018, São Paulo

- Conducted Item Response Theory (IRT) analysis to estimate student abilities and evaluate item performance for 3 large-scale assessments in São Paulo and Pará in Brazil.
- **Developed and delivered detailed educational performance reports for 645 cities in São Paulo**, offering schools actionable insights into student achievement and areas for improvement.
- Utilised statistical expertise to clean, analyse, and present results from large-scale student databases using MySQL and R, contributing to educational assessments post-degree.
- **Presented key findings to stakeholders**, including government officials and educational institutions, ensuring clear communication of assessment outcomes and recommendations.

---

## EDUCATION

### Master of Science in Statistics

Minor in Latent Trait Models • University of São Paulo • São Paulo - Brazil • 2025

### English Exchange

Ibat • Dublin - Ireland • 2024

### Bachelor of Science in Statistics

Minor in Latent Trait Models • Federal University of Pará • Pará - Brazil • 2017

- Scientific initiation scholarship.

---

## PROJECTS

### Computerised adaptive testing tool

University of São Paulo • April 2023 - Present

- Created a CAT API using R for IRT and CDM models, integrated with Django (Python) for user interaction and item bank management.
  - Implemented unidimensional IRT models, established a database for item storage, and developed an administrative interface for managing tests.
  - Tested the API with IRT models in a classroom setting, using data from educational assessments for validation.
  - Future activities: implement multidimensional IRT and CDM models, conduct real-world evaluations, and finalise the documentation and academic publications.
-