

# Guillermo Lezama

(412) 897-0180 | Trenton, NJ (open to relocate) | guillermo.lezama@pitt.edu | [LinkedIn](#) | [Github](#) | [Personal Website](#)

## PROFESSIONAL SUMMARY

Ph.D. candidate in Economics with 10+ years in quantitative roles, specializing in data science, NLP, LLMs, causal inference, and large-scale machine learning in cloud environments (AWS). Experience includes developing LLM pipelines (OpenAI, RAG), and scalable ML systems with PySpark and SageMaker. Experienced in machine learning, large-scale data analysis, and AI-driven text classification. Strong communicator with a record of translating complex data into business and policy insights. Seeking NLP, product analytics, or data science roles to apply these skills.

## SKILLS AND TRAINING

**Programming, Cloud & Data Science:** Python, SQL, Stata, R, PySpark, AWS (SageMaker, S3, EMR)

**LLMs & GenAI:** GPT-4, OpenAI, RAG, fine-tuning

**NLP/ML Frameworks:** Scikit-learn, TensorFlow, spaCy, HuggingFace

**Machine Learning & NLP:** OpenAI API, Prompt Engineering, Naïve Bayes, Random Forests, K-Means, Text Classification

**Causal Inference & Experimentation:** A/B Testing, Diff-in-Diff, RDD, IV, Monte Carlo Simulations

**Data Visualization:** Tableau, ArcGIS, Streamlit, Matplotlib

**Languages:** Spanish (native/fluent), English (fluent)

**Certificates:** [Summer Institute in Computational Social Science](#), [Erdos Institute Data Science Boot Camp](#)

## EXPERIENCE

Amazon, Seattle, WA

May – August 2024

### Economist Intern, Prime Economics

- Built a causal inference model in PySpark, analyzing billions of transactions to optimize logistics.
- Enhanced Amazon's Shipping Investment Simulator to model peak-period demand shifts for better capacity planning.
- Leveraged AWS (SageMaker, S3, EMR) for scalable economic modeling.
- Presented insights to economists and stakeholders which helped contribute to economical decisions for the company.

Búsqueda (newspaper), Montevideo, Uruguay

November – December 2024

### Data Analyst ([Dashboard 1](#) | [Dashboard 2](#) | [Github Repository](#) | [Slides](#))

- Developed and deployed interactive dashboards with Streamlit & Heroku which helped increase website traffic and engagement during national elections.
- Automated processing of 20 years of election data (2004–2024) using Python which enabled scalable voter behavior analysis.
- Built data visualizations of electoral trends which simplified public access to complex election data.

Universidad de la República, Montevideo, Uruguay

August 2024– Present

### Instructor of Exploratory Data Analysis ([Github Repository](#))

- Taught industry-relevant Python and SQL for data science class, covering NLP, clustering, and time series forecasting (20 students)
- Mentored two students on applied projects, helping them build machine learning pipelines to be used in business.

University of Pittsburgh, Pittsburgh, PA

August 2019 – Present

### Research Assistant and Instructor

- Built and optimized an OpenAI-powered NLP validation system, achieving consistent classifications with human annotations (correlation > 0.5), enabling scalable narrative labeling while minimizing researcher bias.
- Designed and taught a fully asynchronous undergraduate economics course, producing 80+ instructional videos and assessments to improve access and engagement across institutions.
- Conducted applied research in NLP, econometrics, and causal inference, contributing to peer-reviewed publications, international conference presentations, and securing multiple research grants.

The Summer Institute in Computational Social Science, Montevideo, Uruguay

January – July 2024

### Co-Organizer

- Secured funding (8000 USD) from the Templeton World Charity Foundation to host a two-week Spanish-Language Institute for 15 graduate students and professionals in computational social science.
- Coordinated logistics and programming for lectures, workshops, and research projects on text as data, web scraping, digital experiments, and political polarization.

Legislators' Office, Montevideo, Uruguay

February 2015 – February 2017

### Economic Consultant to a Member of the Congress

- Synthesized information by crafting presentations and briefs that simplified economic concepts for non-specialist politicians, utilizing macroeconomic and budget data to enhance understanding and decision-making.
- Received recognition as one of the "top 5" in a Trade Policy Essay competition.

Universidad de la República, Montevideo, Uruguay

February 2017 – February 2019

### Academic Advisor to the Dean, Facultad de Ciencias Sociales

- Led automation of contract renewal reporting, streamlined faculty hiring evaluations, and reduced administrative workload for 100 professors.
- Coordinated two high-profile research conferences, achieving record attendance (~2000 people) and collaborated with UNESCO to support academic dissemination.
- Oversaw the editorial process for a book publication including managing timelines, revisions, and coordinated with contributors and publishers to maintain the momentum of ensuring the book was published in a timely manner.
- Facilitated negotiations with students, alumni, and faculty on institutional consensus on strategic decisions.
- Assisted budget advisor and accounting office with financial planning and analysis for faculty operations.

**PUBLICATIONS AND DATA PROJECTS**

- Automating Large-Scale Political Text Analysis with OpenAI (LLMs, NLP & Causal Inference)** (co-authored) ([Dissertation Chapter Draft](#) - submitted | [Published Paper](#))
- Engineered and fine-tuned OpenAI API prompts for 136K political statements, achieving 84% accuracy.
  - Validated LLM classifications against human annotations, showing ChatGPT could reduce labeling costs from \$12,000 to \$197 per 122,000 texts.
- Information about Corruption and Politicians' Proposals (NLP & Causal Inference)** ([Job Market Paper](#) - submitted | [Github Repository](#) | [Slides for NLP part](#))
- Analyzed 13,344 political manifestos using ML & NLP, finding that incumbents reduced focus on exposed policy issues and adopted more populist rhetoric.
  - Presented findings at the World Bank-IFS-ODI Tax Conference, shaping discussions on how transparency influences electoral strategy and political discourse.
- Extreme Weather Events and Household Environmentally Friendly Consumption** (co-authored)
- Merged and cleaned efficiently over 15M Nielsen Consumer Panel data points to analyze eco-friendly consumption metrics and assess extreme weather impacts on consumption.
- Testing Qualitative Effects with Experimenter Demand (A/B Testing & Causal Inference)** (co-authored) ([Draft](#))
- Designed and ran controlled A/B experiments to test experimenter demand bias in surveys, providing insights for business and behavioral research.
  - Developed Monte Carlo simulations in Python, ensuring statistical power for robust experiment design.
- Export Restrictions and Global Food Prices** (co-authored) ([Working Paper](#))
- Built a database of trade restrictions and estimated their effects on global prices using a disaggregated gravity model.
- Sanctions Analytics: Financial Impact of OFAC Enforcement**
- Used NLP to analyze OFAC enforcement data and merged with CRSP stock returns to study firm-level reputation loss.
  - Demonstrated how public enforcement signals can shape market behavior showcasing relevancy for compliance risk and regulatory impact.

**EDUCATION**

|   |                     |
|---|---------------------|
| University of Pittsburgh, Pittsburgh, PA            | Expected: June 2025 |
| <b>PhD Economics</b>                                |                     |
| University of Pittsburgh, Pittsburgh, PA            | 2020                |
| <b>Master of Arts in Economics</b>                  |                     |
| Universidad de la República (Uruguay)               | 2019                |
| <b>Master of Science in International Economics</b> |                     |
| Universidad de la República (Uruguay),              | 2013                |
| <b>Bachelor of Arts in Economics</b>                |                     |