

Sebastian Daza

SENIOR DATA SCIENTIST · SENIOR DATA ANALYST ·

+34634491616 | ✉ sebastian.daza@gmail.com | 🌐 sdaza.com | 📷 sdaza | 📱 sdaza

Summary

I am a senior data scientist and analyst with practical experience designing, managing, supervising, and analyzing applied industry and social studies. I possess strong quantitative and qualitative skills, as well as a background in health, social, and behavioral sciences. I am excited to apply my knowledge and skills to research and influence people, and I enjoy working with interdisciplinary and multicultural teams.

Work Experience

Teladoc (Telemedicine)

Hybrid, Spain

SENIOR DATA SCIENTIST

Apr 2023 - Present

- Technical lead in data science, personalization, and causal inference (internal workshops and technical guidance)
- Led [uplift modeling](#) and incrementality testing of media campaigns with [geo-experiments](#) and [A/B tests](#)
- Created [experiment and causal inference package](#) for imbalance adjustment (IPW, IV), meta-analysis, and power analysis
- Developed a multi-touch attribution ([MTA](#)) model to evaluate investments in paid media and optimize spending
- Developed a life time value ([LTV](#)) model that helps inform marketing investments for optimal results
- Create [LLM](#) pipeline with [Azure OpenAI](#) to extract information from doctors' notes and content generation
- Created tools to process data efficiently with [PySpark](#) and [Databricks](#) (e.g., internal Python packages, CICD pipelines, Terraform)
- Automated the generation of metrics using [PySpark](#)
- Created [Tableau](#) dashboards that provide insights to stakeholders

Lookiero (Online Personal Styling Service)

Remote, Spain

SENIOR DATA SCIENTIST

Jun 2022 - Feb 2023

- Technical lead in data science and causal inference (internal workshops and technical guidance)
- Used statistical models to identify sources of product success (variance decomposition)
- [Causal estimation](#) of discount strategies. Simulation of scenarios. Increased the contribution margin more than 5%
- Developed a [hybrid recommendation](#) model using more than 200M feedbacks from users
- Contributed to define a framework to iterate neural network models in [Databricks](#) ([PySpark](#)), decreasing time in training

Institute of Economy, Geography and Demography, CCSH-CSIC

Madrid, Spain

SENIOR DATA RESEARCHER ANALYST

Dec 2019 - Apr 2022

- Led a [team of researchers](#) to study early conditions, delayed adult effects and morbidity, disability, and mortality
- Led [Scrum ceremonies](#) and tracked progress using [Notion](#)
- Authorship of 5+ academic papers and presentations in conferences
- Supervised and implemented complex data analysis on demography and [genomics](#) using [R](#) and [Java](#)
- Health modeling and simulation using [Anylogic](#)

Juntos Global (Conversational technology platform for financial users)

Bay area, California, US

DATA SCIENTIST

May 2018 - Dec 2019

- Led analytic and data science team for [banking user insights](#)
- Provided user insights through the design and analysis of [experiments](#) ([A/B testing](#)), improving banks' engagement
- Implemented [NLP](#) and [machine learning](#) models to classify banking users and topics and help to resolve their needs
- Worked closely with [Product Managers](#), [Software Engineers](#), and [Design team](#) to improve Juntos' banking engagement product
- Collaborated in the design and implementation of the data pipeline using [Python](#), [PySpark](#), and [Databricks](#)

Center for Studies in Demography and Ecology, UW-Madison

Madison, US

RESEARCH ASSISTANT

Jul 2014 - May 2018

- Communicated insights to different audiences in 5+ empirical research papers, and 10+ presentations at scientific conferences
- Led all aspects of 5+ research studies (funding, design, analysis, reporting)

Center for Social Studies, Catholic University

Santiago, Chile

RESEARCH DIRECTOR

May 2005 - Jul 2010

- Led several research projects in areas such as organizational studies (engagement surveys and culture analysis), political perception and government approval ratings, education (curriculum impact), tobacco and substance, and crime and delinquency
- A key project was the Survey of Social Cohesion in Latin America (ECosocial) that involved 7 countries: Chile, Argentina, Brazil, Colombia, Guatemala, Mexico and Peru, and 10,000 cases

Consulting

Kreuger Consulting, LLC

Texas, US

DATA SCIENTIST

May. 2021 - Jul. 2021

- Created an interactive data dashboard tool using R and Shiny, to visualize the output of a Covid-19 agent-based model
- Provided domain stakeholders with the ability to interact with and modify the results displayed

Personio

Madrid, Spain

DATA SCIENTIST

May. 2020 - Jul. 2020

- Created an interactive data dashboard tool using R and Shiny to assess recruitment performance and transition rates between interview states.
- The dashboard also simulates hiring scenarios to estimate the number of interviews and recruiters needed by position, role, and office.
- It improved the detection of bottlenecks and assessment of performance.

Applied Population Lab, UW-Madison

Madison, US

DATA SCIENTIST

Oct. 2016 - Jul. 2018

- Created R package to extract variables and statistics from the American Community Survey and Decennial US Census.
- The package automatized the extraction of data and improved the time for reporting.

Center for Social Studies, Catholic University

Santiago, Chile

DATA SCIENTIST

Jul. 2010 - May. 2018

- Designed samples and post-adjustments in large studies (organizational, public opinion, and educational studies)
- Provided methodological guidance (attrition, non-response adjustment)

Skills

Programming Python, R, Java

Data science ML, DL (Keras), Visualization (Dash, Shiny), PySpark, Petastorm, SQL, Redshift, PostgreSQL, Databricks

Stats/modeling Experimentation (AB testing), causal inference, hierarchical modeling, survival analysis, Bayesian statistics, simulation, agent-based modeling (ABM)

Survey design Sample design, weighting, non-response adjustments, bias analysis, technical reporting, methodological evaluation

Tools Notion, Trello, Jira, Confluence, Git, Latex

Languages English, Spanish

Teaching Prepared course materials on statistics and modeling, and provided feedback for over 200 students.

Software

experiment-utils Python package for analysis and design of experiments

acsr R package to extract variables and statistics from the American Community Survey and Decennial US Census

zotnote Python package to extract notes from Zotero (Reference Manager)

Education

University of Wisconsin-Madison

Madison, US

PHD IN DEMOGRAPHY AND SOCIOLOGY

The Data Incubator

Washington DC, US

DATA SCIENCE PROGRAM

Catholic University of Chile

Santiago, Chile

MASTER OF ARTS IN SOCIOLOGY