

# Sebastian Daza

DATA SCIENTIST · DATA ANALYST · RESEARCHER · DEMOGRAPHER

+34634491616 | ✉ sebastian.daza@gmail.com | 🏠 sdaza.com | 📷 sdaza | 📺 sdaza

## Summary

Data scientist and analyst with hands-on experience designing, managing, supervising, and analyzing applied industry studies. Strong applied quantitative and qualitative skills and social and behavioral background. I am eager to apply my knowledge and skills in researching and impacting people around us and working with interdisciplinary and multicultural teams. For the moment, living in Madrid, Spain.

## Work Experience

### Lookiero

Remote, Spain

SENIOR DATA SCIENTIST

June. 2022 - PRESENT

- Technical lead in data science and causal inference
- Define a hybrid recommendation model using more than 200M feedbacks from users
- Use statistical models to identify sources of product success (variance decomposition) and the causal effect of discount strategies
- Simulation of scenarios
- Contribute to define a framework to iterate neural network models in [Databricks \(PySpark\)](#)

### Institute of Economy, Geography and Demography, CESH-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](#)
- Supervised and implemented complex data analysis on demography and [genomics](#) using [R](#) and [Java](#)
- Health modeling and simulation using [Anylogic](#)

### Juntos Global

Bay area, California, US

SENIOR DATA SCIENTIST

May. 2018 - Dec. 2019

- Led analytic and data science team for [user insights](#)
- Provided user insights through the design and analysis of [experiments \(A/B testing\)](#)
- Implemented [NLP](#) and [machine learning](#) models to classify users and topics
- Worked closely with [Product Managers](#), [Software Engineers](#), and [Design team](#) to improve Juntos' 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, climate and pulse checks, 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

### Center for Justice and Society Studies, Catholic University

Santiago, Chile

RESEARCH AFFILIATE

Mar. 2016 - PRESENT

- Co-investigator of the project Women Re-entry Study
- Provide methodological advice for the design of samples and experiments

### 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

DATA SCIENTIST

Madrid, Spain

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.

## Applied Population Lab, UW-Madison

DATA SCIENTIST

Madison, US

Oct, 2016 - Jul, 2018

- Created R package to extract variables and statistics from the American Community Survey and Decennial US Census.

## Center for Social Studies, Catholic University

DATA SCIENTIST

Santiago, Chile

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

Machine Learning, NLP, Deep Learning (Keras), Visualization, Dashboard (Dash, Shiny), pySpark, Petastorm, SQL, PostgreSQL, Databricks

### Modeling

Causal inference, AB testing, hierarchical modeling, survival analysis, Bayesian statistics, simulation, agent-based modeling (ABM)

### Tools

Notion, Trello, Jira, Confluence, Git, Latex

### Languages

English, Spanish

### Teaching

Delivered over \$400k in statistical and modeling instructional services. Prepared course materials and provided written feedback for 200+ students.

## Software

### acsr

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

### zotnote

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

## Education

### University of Wisconsin-Madison

PHD IN DEMOGRAPHY (QUANTITATIVE SCIENCES)

Madison, US

### Catholic University of Chile

MASTER OF ARTS IN SOCIOLOGY

Santiago, Chile

## Certifications

2017 **The Data Incubator (Data Science)**, Capstone Project: Tracking sentiment of Congress members' tweets

Washington DC, US

2021 **Genomics for Social Scientists - National Institute of Aging**, University of Michigan

Michigan, US

2016 **Agent-Based and Hybrid Modeling Bootcamp**, University of Saskatchewan

Saskatoon, Canada

2015 **Agent-based Model Analysis**, Technische Universität Dresden

Dresden, Germany

2014 **Agent-based Modeling and Simulation**, Max Planck Institute for Demographic Research

Rostock, Germany

2011 **Social Network Analysis with Actor-based Models**, Summer Institutes in Statistics, University of Kansas

Kansas, US