

Jorge Cimentada

Contact information cimentadaj@gmail.com
[Google Scholar](#)
Madrid, Spain
Date of birth: 1st of March, 1991

cimentadaj.github.io
Github: [github/cimentadaj](https://github.com/cimentadaj)
Phone: +34 697 382 009

Research Interests Social Mobility, Educational Policy, R programming, Statistical computing, Statistics

Education **Ph.D., in Quantitative Sociology** 2015/2019
Pompeu Fabra University, Barcelona, Spain
Fields: Stratification, Educational stratification
Advisors: Gøsta Esping Andersen, Bruno Arpino

Awarded **best PhD paper Prize (2017)** for the paper "Does curricular tracking explain SES gaps? An international comparison of the SES achievement gaps from 2000 to 2015"

My research involved collapsing data on over 3 million students to study educational policies and its impact on inequality. The key takeaway is that inequality depends heavily on the level of vocational education of the country.

M.S., in Research in Sociology and Demography 2014/2015
Pompeu Fabra University, Barcelona Spain
Dissertation: Improving cognitive and non-cognitive abilities: Impact evaluation of the AGE program in Mexico
Advisors: Gøsta Esping Andersen, Bruno Arpino
* Awarded best Master Thesis: featured [here](#).

B.A., Market Research 2009/2013
Pontificia Universidad Católica Madre y Maestra,
Santo Domingo, Dominican Republic.

Work Experience *Data Scientist* 2018 -
Kernel Analytics

Working full time as a Data Scientist at Kernel Analytics, among the best consultancy firms specialized in advanced analytics. My work involves querying 500GB datasets with Impala, Hadoop and Hive and applying state of the art algorithms to understand complex patterns with Python and R. Currently, I

predict customer churn and client segmentation.

R developer & Data scientist

2017 - 2018

Research and Expertise Centre for Survey Methodology

Projects: European Social Survey (ESS)-ERIC.

Principal Investigator: Wiebke Weber

Designed and implemented a Shiny application to download Spanish translated data from the European Social Survey using my R package *essurvey*. I automated the translations using the Google Translate API and text mining techniques. Also developed a second application to adjust model predictions for their measurement error. As part of the project, I developed the *sqpr* R library, a high-level interface to the SQP API where I implemented several algorithms that correct for measurement error.

Data Scientist

2015 - 2017

Pompeu Fabra University

Projects: Stratified Family Dynamics: Polarizing trends in couple behaviour and parenting

Principal Investigator: Gøsta Esping-Andersen

As the main data scientists for the project Stratified Family Dynamics, I performed statistical analysis to study mobility patterns between social classes. I exercised the complete research pipeline of accessing dirty data, cleaning, analyzing noisy records and reporting. In one particular case, I used several classification algorithms to estimate the propensity of mobility, finally landing on Latent Discriminant Analysis because the predictive classes were stable and had a very low out-of-sample error.

Statistical consulting

2014

Population Services International

Project: Measuring the impact of preventive HIV campaigns in the Military forces of the Dominican Republic

I designed, implemented and analyzed a randomized controlled trial funded by USAID to evaluate whether a communications campaign in the armed military forces of the Dominican Republic increased knowledge of sexually transmitted diseases. Result showed an increase of HIV awareness by 7% in the population of armed forces. This led to a renovation of funding from the Department of Defense of the United States and USAID. I also collaborated on other projects within PSI working as a statistical consultant on projects on public health and public policy.

Quantitative analyst

2013

Newlink Research

Projects: Public opinion polls, Electoral turnouts, Market analysis
Principal investigator: Eduardo Gamarra

My main job was statistical consulting within the company. I worked on projects related to customer churn, banking, retail and public opinion/political forecasts. The latter involved live-analysis of opinion polls in combination with the Miami Headquarters.

Teaching/Talks	<p>Pompeu Fabra University <i>Department of Political and Social Sciences</i> Teaching Assistant, Statistics 101 for Political Science, May-2015/Jul-2015</p> <p><i>Research and Expertise Centre for Survey Methodology</i></p> <ul style="list-style-type: none">- Introduction to R programming and its strengths, April 2016 [Presentation]- Getting familiar with R objects, April 2016 [Course]- Introduction to the tidyverse and functional programming, March 2017 [Course]- Introduction to R - Barcelona Summer School in Survey Methodology, Jun-Jul 2017 [Course] <p><i>Invited talks</i></p> <ul style="list-style-type: none">- Introduction to the tidyverse - PyData Conference, May-2017 [Course]- The essurvey package - University of Southampton, January 2018- Downloading European data with essurvey, University Carlos III of Madrid, September 2018
Academic skills	<p>Past training:</p> <ul style="list-style-type: none">- Introduction to Bayesian Inference and Stan, Newcastle University - December 7-8, 2017- Max Planck Institute for Demographic Research Visualization Retreat (awarded full stipend) - June 26-28, 2017- EERA Summer School: Methods and Methodology in Educational Research, Linz University, July 10-15, 2016- Probability Sampling Methods, Pompeu Fabra University, June 27 - July 1, 2016- Analysing Comparative Longitudinal Survey Data with Multilevel Modelling, Pompeu Fabra University, July 6-7, 2016- Methodological seminar on Causal Inference, Pompeu Fabra University, May 18-20, 2015- Protecting Human Research Participants, National Institute of Health (NIH) of the United States, May, 2013.
Extra-curricular	<ul style="list-style-type: none">- Co-organizer of the RECSM RonR! meeting group and seminars- Adaptation of the Randomization guide from the Abdul Latif Jameel Poverty Action Lab (J-PAL) to R. [Link]- PhD Representative years 2017-2018
Languages and Skills	<p>Spanish(native), English (fluent)</p> <p>R, Python, Impala, Hadoop, SQL, Git, Bash, Airflow, Regular Expressions, \LaTeX</p> <p>Microsoft Excel, PowerBI</p>

Software

Github ranking: R (<http://git-awards.com/users/cimentadaj>)

3 out of 126 in Madrid

13 out of 482 in Spain

R packages:

opendataes: Interact with the datos.gob.es API [[Github](#)]

essurvey: Download data from the European Social Survey [[CRAN](#), [Package website](#)]

perccalc: Estimate continuous distributions from categorical variables [[CRAN](#), [Github](#)]

coauthornetwork: Explore your network of coauthors from your Google Scholar profile [[Github](#)]

cimentadaj: Personal package [[Github](#)]

Contributor to other R packages

corr: Exploring correlations in a ‘tidy’ format. [[Github](#)]

broom: Convert statistical analysis objects into tidy format [[Github](#)]

scholar: Extract citation data from Google Scholar. [[Github](#)]

Professional Memberships

Comparative and International Education Society

Foundation for Open Access Statistics

Published Papers

Esping-Andersen, G. & Cimentada, J. (2018). Ability and mobility: The relative influence of skills and social origin on social mobility. *Social Science Research*, Volume 75, Pages 13-31. [[Replication files](#)] [[Open access](#)]

Working Papers

Cimentada, J. & Kluesner, S. (2018), Exploring the Demographic History of Populations with Enhanced Lexis Surfaces, submitted to a special issue of *Demographic Research*. [[Replication files](#)]

Cimentada, J. (2018), Does curricular tracking explain global SES gaps? an international comparison of the SES achievement gaps from 2000 to 2015. <https://doi.org/10.31235/osf.io/3mvxa> [[Replication files](#)]