Tomás Capretto

Doctorate Candidate

IMASL/CONICET

✉ tomicapretto@gmail.com | ⌂ tcapretto.netlify.com | ⊙ tomicapretto

Statistician interested in R, Python, R Shiny and anything that improves how we do data analysis

Education

Universidad Nacional de Rosario   *Rosario, Argentina*

Doctorate in Statistics   *2019-2024 (expected)*

- I work under the supervision of PhD. Osvaldo Martin on Bayesian statistics and exploratory data analysis.
- Studied density estimators and implemented a new alternative for the Python package ArviZ.This implementation was then translated to R and included in bayesplot package.
- Currently focused on Bambi, a Python library that makes Bayesian model building easier.

Universidad Nacional de Rosario   *Rosario, Argentina*

Bachelor in Statistics   *2014-2019*

- The career gave me a solid background in foundations of statistics as well as the ability to interpret results in an analytical manner.

Research experience

Universidad Nacional de Rosario   *Rosario, Argentina*

Research assistant   *Apr 2018 - Mar 2019*

- **Project:** Statistical methods in official statistics.
- Developed my undergraduate thesis *Estimation in complex survey samples in presence of nonresponse*.

Universidad Nacional de Rosario   *Rosario, Argentina*

Research assistant   *Apr 2016 - Mar 2017*

- **Project:** Sampling designs based on models for populations with spatial variability.
- As a complement of a course in sampling techniques, I dedicated most of the time to the study of new sampling methods that incorporate spatial variability.

Work experience

AlixPartners   *Buenos Aires, Argentina*

Consultant   *Nov 2018 - Jan 2019*

- Data manipulation and analysis using R, Python, SQL and Tableau.
- Time series forecasting.
- Task automation.

Instituto Nacional de Estadística y Censos   *Buenos Aires, Argentina*

Statistical assistant   *May 2018 - Oct 2018*

- I worked on the study of non-response, its classification and treatment.
- Contributed to methodological manuals and internal reports.
- Evaluation of methods to impute missing data.

Universidad Nacional de Rosario   *Rosario, Argentina*

Statistical assistant   *Jul 2016 - Sep 2016*

- Survey sampling study to study characteristics of members from a professional college in Santa Fe province.
- Participated in the strategy to collect the data.
- Worked on the consistency of the data bases and the generation of internal reports.
- Worked on telephone surveys and cases follow-up.

Training experience

Chegg Tutors
Remote
ONLINE TUTOR
Aug 2018 - May 2019
- Tutored Statistics, R and SAS Programming, and related topics in English.
- I workd with both undergraduate and graduate students from a very wide range of disciplines.

Latinhire
Remote
ONLINE TUTOR
Nov 2015 - Nov 2018
- Tutored Statistics both in English and Spanish.
- Worked with both high-school and undergraduate students.

Scholarships, awards and distinctions

CONICET
San Luis, Argentina
PhD SCHOLARSHIP IN STRATEGIC TOPICS
2019 - 2024
- **Topic:** Exploratory analysis of Bayesian methods.
- Full-time position dedicated to study and research.

Grupo Argentino de Biometría
Rosario, Argentina
1ST PRIZE IN THE XIII YOUNG BIOMETRISTS CONTEST
Oct 2017
- I worked with David Presutti, Agronomic Engineering student.
- We used generalized mixed linear models to solve the proposed problem.

Comisión Fulbright Argentina
Austin, United States
FULBRIGHT SCHOLAR
Jan 2017 - Feb 2017
- I made a stay in the University of Texas at Austin thanks to the scholarship Friends of Fulbright 2017.
- The activities involved english training, cultural exchange, volunteering and I also attended classes in Statistics.

Side projects

Out of curiosity or fun, I also dedicate part of my time to projects not directly related to my doctorate.

- `latex2r`: An R package to translate math formulas written in LaTeX to R code.
- `shinymath`: Math inputs in R Shiny apps. This package gives a new input that is based on Mathquill and returns R code.
- `formulae`: A WIP library that enables using the popular R formulas in Python. While there is `patsy` out there, this implementation also handles group specific effects (a.k.a. random effects), among other features.