

Tomás Capretto

STATISTICIAN

✉ tomicapretto@gmail.com | 🌐 tomicapretto.github.io | 📷 tomicapretto

Education

Universidad Nacional de Rosario

DOCTORATE IN STATISTICS

Rosario, Argentina

2019-2024 (expected)

- I develop software to make it easier to do Bayesian statistics with Dr. Osvaldo Martin.

Universidad Nacional de Rosario

BACHELORS IN STATISTICS

Rosario, Argentina

2014-2019

- GPA: **9.17**.

Work experience

Grupo San Cristóbal

DATA SCIENTIST

Rosario, Argentina

Oct 2021 - Present

- Data analysis and development of statistical models.
- Improvements in automated decision-making processes.

NumFOCUS - Google Summer of Code 2021

DEVELOPER

Remote

Jun 2021 - Aug 2021

- Implemented a new method to obtain prior distributions.
- Added new families of generalized linear models.
- Several internal changes to improve the performance and readability of the code.

Independent

FREELANCE

Remote

Aug 2020 - Dec 2020

- Development of Shiny apps to conduct data analysis and automate tasks.
- Statistical modeling and data analysis.

AlixPartners

CONSULTANT

Buenos Aires, Argentina

Nov 2018 - Jan 2019

- Time series forecasting and task automation.

Instituto Nacional de Estadística y Censos

STATISTICAL ASSISTANT

Buenos Aires, Argentina

May 2018 - Oct 2018

- I worked on the estimation in the presence of non-response and conducted simulation studies to evaluate imputation methods.

Skills

R

ADVANCED

- I have more than 5 years of experience with **R** and a pretty good understanding of how it works as a language. I write clean and efficient **R** code.
- I can comfortably work with the **tidyverse** or other frameworks. I know how to use different OOP systems (S3 and R6).
- I am very well versed with **ggplot2** and associated libraries. I contributed several articles to The R Gallery.
- I have been using **Shiny** for three years now. It started as a means to make apps in **R** and ended up being my front door to web development.

Python

ADVANCED

- I am fluent with data science libraries such as **NumPy**, **SciPy**, **Pandas**, and **Matplotlib**. I contributed several articles to The Python Gallery.
- I am very proficient with probabilistic machine learning libraries such as **PyMC**.

Others

MISC SKILLS AND INTERESTS

- Fluent in English (C1 level). I work with both **Linux** and **Windows** and I use **git** on a daily basis. I'm very well versed with RMarkdown and **LaTeX**.
- I'm deeply interested in web technologies. I am familiar with (and have used) **Javascript**, **HTML**, **CSS**, **D3**, **React**, and **Next**.
- I have experience building **Shiny** apps that use **Python** through the **reticulate** package.
- I'm very curious and I'm constantly looking to learn new things. To give an example, I wanted to understand how programming languages worked so much that I ended up reading the book Crafting Interpreters and wrote my own interpreter in **Java**.

Software

Almost all of my activity in software development is on my [Github profile](#). Below I include some of the projects that stand out the most or that I most enjoyed doing. By clicking on the name of the project you go to the repository.

PYTHON

ArviZ

CONTRIBUTOR

- A library for exploratory analysis of Bayesian models with Python. It is currently used in more than 1700 OSS projects and 60 Python libraries.
- My [main contribution](#) was the implementation from scratch of a more precise and efficient method for estimating density functions.

Bambi

DEVELOPER

- A high-level interface to construct and fit Bayesian models in **Python**.
- I'm the main contributor in terms of commits and additions since 2020.

formulae

AUTHOR

- An implementation of statistical formulas to specify mixed models in **Python** in a similar way to the **lme4** package in **R**.
- It is an alternative to **Patsy** that we use to specify models in **Bambi**.

R

latex2r

AUTHOR

- Implements an interpreter that translates \LaTeX math equations into **R** code.

shinymath

AUTHOR

- Implements a new math input for **Shiny**.
- Uses **latex2r** under the hood to return expressions that can be understood by **R**.

R SHINY

samplingdistributions

AUTHOR

- Explore sampling distributions of several statistics computed with samples taken from arbitrary mixtures of probability distributions.

playwithregression

AUTHOR

- A Shiny application to play with simple linear regression fit via Least Squares method.
- The main interactive visualisation is produced with **D3** and made available via **r2d3**.

ifuns

AUTHOR

- Visualise and tune univariate mathematical functions.
- This project was created to show the functionality and versatility of `pkg shinymath`.

whackamole

AUTHOR

- A simple Whack-A-Mole game based on the original version made by [Wes Bos](#) as part of his [Javascript 30](#) challenge.

rbingo

AUTHOR

- An entire system to manage and run a Bingo game made in **Shiny**.
- Create games, administer vendors and sales, play the game, and obtain automated reports.
- This app was created for a club in my community and thus is written in Spanish.
- The version they use is a desktop app created with **Electron**.