About

Hello! My name is Luís Assunção and I'm a Data Scientist and Statistician.

My work is to help people make decisions under uncertainty through experimentation, causal inference and probabilistic modeling.

I also enjoy hiking, music and woodworking. I live with my partner and our ginger cat in Belo Horizonte, Brazil.

Reach me at assuncaolfi@gmail.com. If you're reading the PDF version of this document, I encourge you to read it in my website instead.

Email Website GitHub LinkedIn

Curriculum Vitae

Education

B.S in Statistics

Federal University of Minas Gerais (UFMG) | Belo Horizonte, MG - Brazil | 2017 - 2021

- Research and reproducible monograph (in portuguese with an abstract in english) on exponential random graphs applied to epidemiology
- Co-author of Frequency and burden of neurological manifestations upon hospital presentation in COVID-19 patients: Findings from a large Brazilian cohort

Experience

Hotmart

Staff Data Scientist | April 2020 - present

To help creators live their passions, I worked in:

- Developing an in-house AB testing framework with optional stopping
- Consulting for and developing randomized controlled trials
- Estimating causal effects in non-randomized experiments
- Hierarchical modeling of pricing elasticity for digital products
- Hidden state modeling of evergreen vs launching sales strategies
- Topification and latent variable modeling for creators and products

Oper

Data Scientist | Oct 2018 - March 2020

For companies such as AB InBev and GTB to understand and make decisions from data, I worked in:

- Spatial modeling of pricing elasticity for beverages
- Multi-touchpoint attribution modeling

IRIS

Intern | 2015 - 2017

To defend and promote public policies that advance human rights in the digital matters, I worked in:

- Collecting, wrangling and describing data

Examples

Blog

Some blog posts:

- Picking a fantasy football team: In this post, I delve into the data for the 2022 season of a brazilian fantasy football league; formulate a mixed integer linear program to pick the optimal team; and present initial concepts for forecasting player scores using mixed effects linear models.
- Parametric non-monotonic models: In this post, I analyse some function alternatives for modeling non-monotonic relationships with interpretable parameters, and compare these models using a a cognitive decline dataset.

Click the **</> CODE** link on the upper right of any post to see its full code. Common tools used in these posts: Python, `polars`, `pymc` and `pulp`.

Repositories

- site: My website and blog post codes
- mldc2020: 7th place solution to the Mercado Libre Data Challenge 2020
- rstanbtm: Biterm Topic Model implementation in Stan
- qlm: Generate predictive SQL queries from linear models in R
- tophat: Scheduled script to fetch and save fantasy football data

Other stuff

- Pod e Dev podcast episode, where I talk (in portuguese) about the challenges in pricing digital products and causal assumptions we made to overcome these challenges in our model at Hotmart. We also discuss good and bad use cases for large language models, as well as how models with 2 parameters can be as useful as models with 200 million parameters.