



Recent Experience

Lead Data Scientist, Fitbod

2021-present

- Fitbod is an app that creates workouts tailored for you. My team and I improve Fitbod's workouts, help design and analyze A/B tests, do R&D, and help the company make sense of data from a wide variety of sources. As team lead I also prioritize projects, assess impact, evangelize a data-driven culture, recruit, and participate in Fitbod's overall business planning.

Data Scientist, Government of Brazil

2014-2021

- I trained a model that helps the Comptroller-General of Brazil choose which municipalities it should audit each year. The total auditable budget was US\$ 3.6 billion in 2018 alone.
- I co-developed and maintained CEREBRO, an app that mines government procurement data looking for evidence of bid rigging. CEREBRO uses a mix of machine learning, time series econometrics, and graph theory. CEREBRO screens all non-classified federal government bidding processes in Brazil. The total auditable budget was US\$ 5.9 billion in 2018 alone.
- I co-developed Trinity, an app that uses graph theory to find corruption networks.
- I trained a model that estimates the price of government-owned real estate properties, to help the government auction them off.
- I created a bot that automates the background check of people appointed for senior government positions.
- I mined the database of covid economic relief to find cases of fraud.
- I helped train and mentor data scientists in multiple government agencies.

Adjunct Professor, several universities

2016-2018

- I have taught statistics and machine learning at a number of private and public universities in Brasília: UnB (machine learning; graduate level), IESB (machine learning; graduate level), IDP (statistics; undergraduate level).

Profile

I have nine years of experience with high-impact, end-to-end data science projects, in different industries and countries.

I also have peer-reviewed publications in top econometrics journals and experience teaching statistics and machine learning at private and public universities.

You can find my public projects and publications at
<https://thiagomarzagao.com>

Main Skills

Machine learning

Statistics

Python

SQL

AWS

GCP

Education & Research

Ph.D. in Political Science

Ohio State University

2010-2014

I took lots of statistics and econometrics classes and in my dissertation research I used natural language processing to create the first machine-coded democracy index.

Online Coursework

Udacity: Nanodegree Deep Reinforcement Learning (2018). Coursera: Deep Learning Specialization (2018), Machine Learning (2013), Introduction to Data Science (2013).