

# LUCAS MORAES

Sou um profissional que transita entre a ciência e a análise de dados. Minha experiência em ambientes diversos (academia, governo e iniciativa privada) exigiram de mim boas habilidades comunicativas e comportamentais. Tenho utilizado ferramentas de data science e boas práticas de pesquisa por cerca de 10 anos.

## EXPERIÊNCIA PROFISSIONAL RECENTE

Presente  
|  
2022

### ● Senior Data Analyst

PicPay

- Suporte estatístico e experimental para o desenvolvimento do modelo de segmentação comportamental de clientes da empresa (GMM).
- Estruturação e análise de dados de experimentos (testes A/B) referentes à home page do aplicativo do PicPay, com o objetivo de aumentar a conversão.
- Machine Learning supervisionado para o desenvolvimento de modelos de propensão.
- Desenvolvimento de pipelines analíticos utilizando GitHub + Databricks + Airflow para a consolidação de tabelas personalizadas no data lake.

2022  
|  
2018

### ● Consultor independente (Data Science & Analytics)

Freelancer

- Compilação, Limpeza, Análise Exploratória e Modelagem de dados para relatórios e/ou projetos de pesquisa.
- Concepção de desenhos experimentais e hipóteses para a resolução de perguntas com rigor científico, estatístico e metodológico (e.g. amostragem e testes estatísticos).
- Desenvolvimento de modelos de machine learning para análises preditivas ou de correlação.

2022  
|  
2021

### ● Data Scientist

Melhor envio

- Desenvolvimento ponta-a-ponta do modelo de segmentação de clientes da empresa (K-prototypes).
- Análises estatísticas para o monitoramento em tempo real da performance/atividade dos clientes, utilizado na tomada de decisão de equipes cross.
- Machine Learning supervisionado para a previsão e identificação de elementos ligados à propensão ao churn.

## Contato & Portfolio

 [lucasmoraes.org](http://lucasmoraes.org)

 [moraessaur@gmail.com](mailto:moraessaur@gmail.com)

 [lucasmoraes.blog](http://lucasmoraes.blog)

 GitHub: @moraessaur

 LinkedIn

 Download do CV em pdf.



## EDUCAÇÃO

2018  
|  
2016



### MsC, Genética

Universidade Federal do Rio de Janeiro

- Clusterização hierárquica e análise de dendrogramas para a identificação de linhagens evolutivamente distintas, integrando dados biológicos, geográficos e moleculares, não estruturados.
- Orientação: Carlos Guerra Schrago.

2012  
|  
2007



### BsC, Genética

Universidade Federal do Rio de Janeiro

- Estimativa filogenéticas/topológicas de cetáceos utilizando inferências bayesiana e de máxima verossimilhança para a parametrização de clusterização hierárquica.
- Publicação: Phylogenetic Status and Timescale for the Diversification of Steno and Sotalia Dolphins. PLOS ONE. <https://doi.org/10.1371/journal.pone.0028297>
- Orientação: Carlos Guerra Schrago.



## MAIS SOBRE MIM



### Skills & Stack

- Minha atuação está mais voltada à modelagem estatística (que inclui Machine Learning) para descrever comportamentos através da interpretação de resultados, desenvolvimento de modelos preditivos (e.g. churn ou propensão) e segmentação de clientes usando ML não supervisionado, seja para produção ou análises ad hoc.
- Modelagem estatística, análise de dados e teste de hipóteses são considerados conhecimentos básicos e corriqueiros em minha área de formação, o que me permite transitar em diferentes áreas técnicas (modelagem, testes A/B ou visualização de dados).
- Prefiro programar em R, mas tenho proficiência em Python, SQL e Spark (tanto Pyspark quanto SparkR). Trabalho confortavelmente com git, github e Linux e também tenho experiência trabalhando no ambiente do Data Bricks (e notebooks em geral) e AWS (e.g. Redshift e S3).
- Trabalhei em muitos projetos de pesquisa com colaboradores de backgrounds e níveis de senioridade distintos, alguns de ponta-a-ponta. Sou treinado desde a graduação a explicar conhecimento técnico para público não técnico.

## Capacitação técnica



R



Python



Spark



SQL



Estatística



Machine Learning



Data Viz



Inglês fluente



Soft Skills