

# Bruno Garcia Piato

## Cientista de Dados

33 anos, Jundiaí-SP, (11)98711.9967

**E-mail:** brunogpiato@gmail.com

**LinkedIn:** <https://www.linkedin.com/in/piatobruno/>

**Portfólio:** <https://brunopiato.github.io/portfolio>

### RESUMO

- Cientista de Dados Júnior em busca de uma oportunidade para utilizar meus conhecimentos e habilidades em análises descritivas, diagnósticas e preditivas para gerar *insights* acionáveis e auxiliar na inteligência e tomada de decisão da empresa.
- Larga bagagem científica em decorrência dos anos dedicados à pesquisa acadêmica na área de Biologia Evolutiva e Genética, aliada a dez anos de experiência como professor de biologia e ciências, garantindo sólidas habilidades analíticas, de comunicação e de aprendizado rápido.
- Desenvolvi quatro projetos de dados durante meus estudos voltados para a transição de carreira, empregando técnicas de compreensão de negócios, visualização de dados e aprendizado de máquina em R e Python usando dados fictícios ou públicos coletados por raspagem web.

### FORMAÇÕES

- *Bacharelado em Ciência de Dados* – UNIVESP, 2023-atual.
- *Mestrado em Biologia Evolutiva e Biodiversidade* – UFABC, 2015-2017.
- *Bacharelado e Licenciatura em Biologia* – UNESP, 2008-2012.

### EXPERIÊNCIA PROFISSIONAL

#### Produtor de Conteúdo Digital | Comunidade Data Science | jul/23 - atual

- Produção de conteúdo digital para o canal de YouTube da Comunidade Data Science envolvendo temas técnicos como o uso de Python, Git e GitHub, SQL, Streamlit, etc, e temas voltados para as carreiras de dados, como trajetórias de aprendizado e estudos, soft skills desejadas, casos famosos de uso de análise de dados, etc.
- Os vídeos podem ser encontrados no canal do YouTube da Comunidade Data Science (@ComunidadeDS), um dos maiores canais sobre carreiras de dados em português.

#### Cientista de Dados – Transição de Carreira | Comunidade Data Science | jan/23 – atual

- Desenvolvimento de projetos com o objetivo de aprimorar e consolidar os conhecimentos adquiridos nos estudos para a transição de carreira para a área de dados.
- O principal objetivo destes projetos é simular o ambiente real de desenvolvimento de soluções através do uso de dados publicamente disponíveis para elaborar as análises usando os princípios da metodologia ágil, como planejamento e aplicação de *sprints* de forma iterativa.
- Com isso pude empregar, desenvolver e consolidar o aprendizado acerca da resolução de problemas de negócios utilizando ferramentas como Python, R, SQL, Streamlit, Google Colabs e bibliotecas como plotly, scikit-learn, matplotlib, ggplot2, etc.
- Os projetos de dados desenvolvidos podem ser encontrados em meu portfólio de projetos ou em meu perfil do GitHub em <https://github.com/brunopiato/>.
- Os principais projetos desenvolvidos foram:

- **WebApp NBA-Déx:** a fim de prover aos gerentes gerais de times da NBA melhores informações para a avaliação e tomada de decisão quanto à contratação de novos talentos para seus times, desenvolvi um aplicativo web utilizando Python e Streamlit que facilita a visualização e comparação das principais estatísticas dos jogadores. O produto final foi um *WebApp*, que está disponível no link <https://nbadex.streamlit.app/>
- **Previsão de vendas das lojas Rossmann:** a fim de auxiliar o CFO da rede de drogarias *Rossmann* a escolher as melhores filiais em que investir para aprimoramentos de infraestrutura, desenvolvi análises preditivas e prescritivas utilizando algoritmos de aprendizado de máquina supervisionado para regressão por séries temporais para prever o faturamento de cada loja durante as seis semanas que se seguiram à análise.
- **Identificação de clientes insatisfeitos para as lojas InStyle:** utilizando as métricas obtidas na pesquisa de satisfação *Net Promoter Score* aplicada aos clientes das lojas *InStyle*, treinei algoritmos de aprendizado de máquina supervisionado para classificação a fim de prever o desfecho de satisfação de clientes e, desta forma, auxiliar os times de marketing e de pós-vendas a direcionar seus esforços a corrigir os pontos fracos da operação da empresa e a reaver os clientes insatisfeitos.
- **Ensaio sistemático de algoritmos de aprendizado de máquina:** a fim de avaliar, estudar e comparar o desempenho de diferentes algoritmos de *machine learning* sobre um mesmo conjunto de dados, realizei um estudo sistemático aplicando diferentes algoritmos, de cada um dos tipos, comparando suas principais métricas de performance. Para classificação utilizei os algoritmos K Nearest Neighbors (KNN), Árvores de Decisão, Florestas de Aleatoriedade e Regressão Logística; para regressão utilizei Regressão Linear (comum e regularizada), Regressão Polinomial (comum e regularizada), Árvores de Decisão, Florestas de Aleatoriedade e RANSAC; por fim, para clusterização utilizei os algoritmos K-Means e Propagação de Afinidade. O resultado final foi um *WebApp*, que pode ser acessado pelo link <https://machinelearningevaluator.streamlit.app/> além de um relatório depositado em <https://github.com/brunopiato/MachineLearningEssay>.

#### **Professor de Ciências e Biologia | Colégio Anglo Leonardo da Vinci | ago/13 - dez/22**

- Durante aproximadamente dez anos ministrei aulas de biologia e ciências para turmas de Ensino Fundamental (anos finais) e Ensino Médio.
- Pude desenvolver e ampliar minhas capacidades de comunicação e assertividade, planejamento e gerenciamento de tempo, flexibilidade e liderança.

#### **Pesquisador (Mestrado) | Laboratório de Sistemática e Diversidade | mai/15 - jul/17**

- Minha pesquisa de mestrado foi sobre a distribuição geográfica de variantes genéticas de duas espécies de seres vivos e o impacto das mudanças climáticas desde o Período Pleistoceno nessa distribuição. Para isso lancei mão da análise de dados genéticos e de algoritmos de aprendizado de máquina para estimar a distribuição geográfica das espécies há 21 mil anos e, assim, reconstruir sua história evolutiva.
- Nesse contexto usei ferramentas de obtenção de dados genéticos (BLAST) em bases de dados públicos (GenBank), análises estatísticas e análises multivariadas de genética populacional, algoritmos de aprendizado de máquina para reconstruir os requerimentos climáticos dos seres vivos com base em seus pontos de ocorrência e variáveis bioclimáticas (WorldClim), implementados em linguagem R.
- Conclui todas as disciplinas do programa de pós-graduação com conceito A, inclusive a qualificação e a defesa da dissertação.

## **OUTRAS EXPERIÊNCIAS**

### **Competidor nos 4° e 5° Hackdays | Comunidade Data Science | abr/23 e jul/23**

Particpei do 4° e do 5° Hackdays da Comunidade Data Science, tendo ficado em 5° lugar, de 23 times, neste último (julho de 2023). Neles pude participar de projetos de dados simulando o ambiente de trabalho real, realizando pesquisa para compreensão do problema de negócios, limpeza de dados, análises exploratórias, levantamento de hipóteses e premissas de trabalho, pré-processamentos de dados, implementação de algoritmos de aprendizado de máquina e ajuste fino de hiper-parâmetros.

## **INFORMAÇÕES COMPLEMENTARES**

- **Idioma:** inglês fluente
- **Linguagens de programação:** experiência com R e Python
- **Banco de Dados:** SQL, MySQL e SQLite
- **Engenharia de Dados:** Git, GitHub e Linux
- **Ferramentas de Visualização:** PowerBI e Streamlit
- **Aprendizado de Máquina (Machine Learning):** classificação, regressão e clusterização
- **Microsoft Office:** Excel, PowerPoint, Word e Outlook
- **Soft Skills:** Comunicação, alta capacidade de aprendizado, resolução de problemas, planejamento e gerenciamento de tempo, trabalho em equipe, flexibilidade e adaptabilidade.