Eduardo Silva Coqueiro

Junior Data Scientist

Brazil | +55 (77) 98160-9066| edu.coqueiro1@gmail.com
https://www.linkedin.com/in/eduardocoqueiro/?locale=en_US/ | https://github.com/Edu-png | https://dataguy.my.canva.site/eduardo-coqueiro

SUMMARY

Biotechnology undergraduate with experience in Data Science, Bioinformatics, and Process Automation. Seeking a position as a Data Scientist to apply technical and analytical skills in innovative projects and data-driven solutions. Despite not having a formal education in data science, I have extensive participation in related projects and am committed to continuously improving my knowledge.

PROFESSIONAL EXPERIENCE

SUZANO Papel & Celulose.

About: Suzano, a leader in the paper and pulp industry, reported a net revenue of BRL 44.4 billion in 2023, reflecting its strong market position and innovative approach.

Mucuri, BA

Role: Research & Development Intern (Plant Genetic Improvement and Data) March, 2023 – Present

- I developed interactive dashboards in Microsoft Power BI for data visualization and analysis, specifically for analyzing the regrowth of materials. This facilitated quicker decision-making, potentially saving the company approximately USD 19,951,030 per year.
- I used the R language as a statistical tool for analyzing large volumes of experimental data, applying modeling and inference techniques. The R scripts save weeks of work, allowing researchers more time for other tasks.
- I created scripts in R language for data processing and automation, increasing the efficiency of collection and analysis operations that were previously done manually via Excel, saving around 18 hours in data allocation.
- I contributed to the planning and execution of experiments in the genetic improvement of eucalyptus, integrating complex data into understandable visualizations via the R programming language.
- I collaborated with multidisciplinary teams to implement innovative technological solutions in research projects and in the seedling nursery of the Bahia Hybridization Center. The nursery produces 35 million eucalyptus seedlings a year.

Key Soft Skills: Collaboration, Communication, Problem-Solving, Adaptability, Initiative.

CORACTIUM.

About: CORACTIUM is a small scientific education startup supported by innovation grants, operates with a team of 8 dedicated employees. The company focuses on leveraging cutting-edge technology to enhance educational processes, driving innovation and excellence in the scientific education sector.

Rio de Janeiro, RJ (remote job)

Role: Junior Data Scientist / Python Backend Developer January 2024 – Present

- I develop and maintain automation scripts in Python, optimizing internal processes and increasing operational efficiency, saving hours for the company's processes.
- I create interactive dashboards in Power BI to visualize sales-related data, helping to make strategic decisions.
- I have designed and implemented software solutions using Python (Django), meeting the specific needs of projects, and improving the functionality of systems.
- I use A/B tests to make strategic decisions for the company, making it easier to close contracts with various institutions.
- I trained a machine learning model to improve our contacts with customers by mailing, our main form of contact.
- I use SQL to administer and manipulate company databases.

Role: Assistant Programmer October 2022 – January 2024

- I developed APIs and automations to integrate and improve the efficiency of the company's educational systems.
- I have done web scraping with Python (Beautiful Soup4) to collect and process data, contributing to the creation of robust and up-to-date databases.
- I prototyped software platforms, collaborating closely with stakeholders to ensure the technical and functional viability of projects.

Key Soft Skills: Analytical Thinking, Creativity, Time Management, Continuous Learning, Resilience, Teamwork.

Federal University of Bahia (UFBA) Vitória da Conquista, BA
Role: Bioinformatics and Data Science Lab Intern Julho 2022 – Novembro 2022

- I developed an innovative tool using R to analyze protein surfaces, making it easier to identify possible pockets (regions with greater affinity for binding compounds), enabling more rational drug development.
- I implemented analysis algorithms for protein mapping based on data from PDB (Protein Data Bank) files.
- I created graphical representations of the protein conformations, allowing detailed visualization and interpretation of the analysis results.
- I collaborated with the research team to integrate the tool into bioinformatics workflows, improving the efficiency and accuracy of the analyses.

Key Soft Skills: Curiosity, Empathy, Organization, Leadership, Teamwork.

EDUCATION

Federal University of Bahia (UFBA) Vitória da Conquista, BA
Biotechnology March 2018 – December 2023

- Participated in research, extension, and teaching projects.
- Served as vice-coordinator of the National Biotechnology Students League (LINA).
- Active member of the board of Primer Jr. (Junior Enterprise of the campus).
- Student representative of the course in the academic center.
- Interned in the bioinformatics lab, developing tools for bioinformaticians.

Interned in the biochemistry lab, conducting research projects in protein studies.

Alura, Online Remote
Data Science Training Program March 2023 - Present

- Completed comprehensive coursework in Data Science, covering key topics such as Python programming, data analysis, machine learning, and data visualization.
- Engaged in practical projects and exercises, applying theoretical knowledge to real-world data problems and enhancing technical skills.
- Collaborated with peers on group projects and discussions, fostering teamwork and effective communication in a virtual learning environment.
- Continuously updated knowledge with the latest trends and tools in Data Science, ensuring a strong foundation and staying current with industry developments.

PROJECTS

Spotify Data Analysis Project: *https://github.com/Edu-png/Spotify-Data-Project*
In this project, I explored a Spotify database available on Kaggle, using Python and its libraries seaborn, matplotlib, pandas and numpy. During the analysis, I investigated correlations between various variables, allowing me to observe trends in musical patterns over the years. It was a valuable opportunity to examine how different elements influence users' musical preferences and how these preferences have evolved over time. Later, this data gave rise to a power BI, which used HTML and chatGPT to generate a visual and informative dashboard.

CNPEM Data Scraper: *https://github.com/Edu-png/CNPEM_scrapper*
I developed a scrapper using Python (BeautifulSoup4) to extract data from the CNPEM website in order to map vacancies of interest to me. This project initially emerged as a personal initiative, aimed at identifying opportunities aligned with my professional interests. Later, this idea evolved into a broader project, integrated into Coractium, a startup of which I am a member. This initiative has become part of a larger effort to automate the collection and analysis of information relevant to our business, providing a competitive advantage through strategic data.

YouTube Data Scraping and API Analysis: *https://github.com/Edu-png/Youtube_data*
This script allows YouTube channel metrics to be analyzed on an individual basis, focusing on specific videos to generate relevant insights. Using libraries such as Pandas and Seaborn, and with the help of ChatGPT to enhance the code, it was possible to extract detailed information and create informative visualizations to better understand the performance of YouTube videos and channels.

Simple Sentiment Analysis in Python: *https://github.com/Edu-png/Simple-Sentiment-Text-Analysis-in-Python*

To carry out a basic text sentiment analysis in Python, I used libraries and functionalities - NLTK (Natural Language Toolkit), TextBlob and Newspaper3k - that allow me to work with natural language processing, including support for the Portuguese language. I explored different sentiment analysis methods, such as text polarity analysis, to understand the emotions expressed in different types of texts in Portuguese.

Protein Mapping Tool for Pocket Identification: *https://github.com/Edu-png/Protein_map*
During my internship at UFBA, I developed an innovative tool in the R programming language to analyze protein surfaces, facilitating the identification of protein pockets. Using analysis algorithms and graphical visualizations, I integrated the tool into bioinformatics workflows, optimizing the efficiency and precision of the analyses, thus contributing to the rational development of drugs. This was one of my first projects with data, so there's still a lot to improve.

4 Years of E-sports: League of Legends: *https://github.com/Edu-png/LOL-4-years--of-esport-and-Prediction-30min--84-Acc--RNN-.git*
This ongoing project, developed as a hobby and with the aim of furthering my studies in machine learning, aims to analyze data from the League of Legends world championships between 2015 and 2017. The data includes matches from the NALCS, EULCS, LCK, LMS and CBLoL leagues, as well as the World Championship and Mid-Season Invitational tournaments. Using machine learning models in Python, the project aims to make predictions and extract valuable insights into the performance of teams and players over the course of matches.

Plant Disease Detection System using AI: *(in progress)*
This ongoing project, developed to further my studies in machine learning, focuses on accurately identifying plant diseases using advanced machine learning techniques. Utilizing a dataset with approximately 87,000 RGB images of healthy and diseased plant leaves, categorized into 38 classes, the project aims to build a robust detection system.

Movie Recommendation System: *(in progress)*
This ongoing project aims to develop a highly effective movie recommendation system using advanced machine learning techniques using the Python language. Combining comprehensive datasets from TMDB and MovieLens, I am exploring different recommendation algorithms, from simple popularity-based models to customized collaborative filtering systems.

A/B Testing in Mobile Games: *(in progress)*
Ongoing project focused on A/B testing in mobile games. Here the goal is to understand how different game elements affect the user experience and game performance, with the focus being to deepen my knowledge of A/B testing using Python. In addition, collecting and analyzing data to optimize the interface, game mechanics and monetization models with the goal of improving user retention and engagement.

SKILLS

Technical Skills: Proficient in SQL, Python (Pandas, Numpy, Matplotlib, Seaborn, Scikit-Learn), R, Power BI (ETL, Modeling, Storytelling, DAX, M), Excel, and PowerPoint.
Languages: Fluent in Portuguese, Intermediate in Spanish, Basic in Japanese.
Certifications and Training: Complete Data Science training (Alura) and Google Data Analytics Professional Certificate (PT) (Coursera).