

Juan Sebastián Barbosa

+57 316 794 6510
jsbarbosacoy@hotmail.com
www.github.com/jsbarbosa

Data team lead with a background in natural sciences, technology and credit risk assessment. Accomplished physicist and chemist with a robust track record of leveraging data-driven insights to propel organizational success. Proficiency in problem-solving across diverse domains and a keen interest in continuous learning. Extensive experience in harnessing the power of data to inform decision-making and drive business outcomes. Specialized skills in data analysis, interpretation, and the development of data tools for impactful solutions. Skilled in cloud computing with expertise in utilizing AWS services to build scalable and resilient solutions, optimize data storage and retrieval processes, and implement secure and cost-effective cloud infrastructure strategies.

Experience

- October 2022–June 2024 **Head of Data, Foodology**
Foodology creates and operates virtual restaurant brands, created especially for delivery. <https://www.foodology.com.co/>
- Shaped Foodology's data strategy leveraging data assets to drive business growth and innovation
 - Built a culture of data-driven decision-making by democratization access to data by launching **Apache Superset** and leveraging drag-and-drop features to extract insights by all kind of teams
 - Collaborate cross-functionally with various teams to identify opportunities for leveraging data to drive business value by leading a group of 4 data analysts, 3 data engineers and a data scientist
 - Directed the development and implementation of **Apache Airflow** to orchestrate more than 100 automated ETL processes, decreasing down times due to failures to almost zero in the last months and allowing the team to proactively monitor the performance of the jobs
 - Reduced infrastructure cost by 30 %, with twice as much information by designing and implementing a robust data infrastructure, ensuring scalability, reliability, and security
 - Helped reduce stock outs in Brazilian stores by a factor of 4x by allowing the supply team to change from a reactive to a proactive methodology by providing better and faster insights through the day
 - Received the Rocket recognition twice for top performance, got awarded a top leader recognition (9.85 / 10 Leadership Feedback) and increased the data team eNPS score from 50 to 88
- July 2019–October 2022 **Data scientist, Zinobe - Tangelo Latam**
Technology company specializing in alternative credit. <https://www.tangelolatam.com/>
- Studied the behavior of clients with consumer credit loans, and built a statistical model for automatic approval of 6,000 loans **increasing the coverage of credit facilities for 4,000 low income Colombians during the Covid pandemic**
 - Developed an scoring system for **Merchant Cash Advance using Clip terminals** that lead to about 10,000 disbursements
 - Built a decision support system, using descriptive statistical analysis and predictive models to assess small and medium-sized enterprises with limited information to reduce the credit committee evaluation time
 - Coordinator and developer of the **psychometrics project**, a project that increased approval rates of customers by measuring the willingness to pay a loan by using a behavioral test developed using **Django**
 - Designed and deployed the **system for portfolio (credit) assignation to debt investors** using **Django**
- January 2017–January 2021 **Software Developer, Tausand Electronics**
Scientific instrumentation development company. <https://www.tausand.com/>
- Designer and developer of the **Abacus** software to be used with the laboratory equipment
 - Built the interface to deal with the low level communication with the Abacus device, so that customer could import a library and customice the behavior of an experiment

- January 2018–June 2019 **Research assistant**, *Universidad de los Andes, Centro de Microscopía*
<https://investigaciones.uniandes.edu.co/es/microscopia/>
- Guided investigators on image analysis techniques to automate the extraction of information using computer vision to count, and measure relevant properties of the image
 - Developed the customer relationship management software ([Microbill](#), [/microbill](#), [Python](#)) used to quote and keep track of the customer balance of the services used in the research facility
 - Developed the Atomic Force Microscope data extraction software ([Igor Extractor](#), [/SPM](#), [Python](#)) to help customers extract their data without the need of specialized software in an easy to read format
- August 2016–December 2017 **Teaching assistant of Computational Methods**, *Universidad de los Andes, Departamento de Física*
- Developed [astrohut](#) ([/astrohut](#)) a library for n-body gravitational simulations with [Python](#), to help physics students understand the fundamental concepts of numerical simulations
 - Developed [rippleTank](#) ([/rippleTank](#)) a library for simulating wave propagation in ripple tanks, it includes different boundary conditions and a framework to build obstacles in the tank. Build with [Python](#)
 - Published [a set of computer simulations describing a variety of simple and complex physical systems](#) to motivate college and school students to develop programming abilities

Education

Majors

- 2013–2019 **Physicist**, *Universidad de los Andes, Colombia*
- GPA: 4.01/5.00
 - Dissertation on [Orbits of black holes in galactic triaxial potentials](#), a simulation of astronomical physical systems
 - Teaching assistant of the Computational Methods class for three consecutive semesters
- 2014–2019 **Chemist**, *Universidad de los Andes, Colombia*
- GPA: 4.01/5.00
 - Dissertation on the [Installation and Calibration of a 2277 Thermal Activity Monitor microcalorimeter](#)

Certifications

- 2022–2022 **Nanodegree: Data Architect**, *Udacity*
- Foundations of Data Architecture were used to build a [HR database](#)
 - Designed and deployed a [data warehouse](#) ([Redshift](#)) with YELP and climate datasets in order to analyze the effects the weather has on customer reviews of restaurants
 - Built a Data Governance project implementing data quality validations and master data management
- 2021–2021 **Nanodegree: Data Engineering**, *Udacity*
- Designed and deployed a [NoSQL ETL pipeline for a music streaming app](#) using [Python](#) and [Apache Cassandra](#) for storage
 - Created an [ETL pipeline](#) that extracts data from [S3](#), stages in [Redshift](#), and transforms data into a set of dimensional tables
 - Created a [data lake using an ETL pipeline in Spark](#) that reads data from [S3](#), processes the data into analytics tables, and loads them back into [S3](#)
 - Deployed an [Airflow pipeline](#) to load immigration and demographics data into a data lake hosted in [S3](#) using parquet files, partitioned by date and state, to study the relation of immigration and demographic data in the US

2020–2020 **Data Science For All**, *Correlation One* - MinTIC, Colombia, 375 hours

- Graduated with Honors
- Collaborated with a team of students to [build a system](#) for the [IGAC](#), in which physical soil properties were used to predict chemical properties of the soil without the use of specialized equipment

2020–2020 **Scrum Foundation Professional Certificate**, *CertiProf*

- Studied the foundational principles of the [Scrum - Agile](#) methodology

Extracurricular activities

May 2019 **Participant**, *Machine Learning for Quantum Matter and Technology*, Décima Escuela de Física Matemática

https://matematicas.uniandes.edu.co/~cursillo_gr/escuela2019/index.php

April 2019 **Collaborator**, *Sympy*

<https://github.com/sympy/sympy/blob/master/AUTHORS>

- Collaborated in the development of the library [SymPy](#) for the symbolic calculation in [Python](#)

October 2018 **Exhibitor**, *Expociencia, Uniandes en Corferias*

<https://medicina.uniandes.edu.co/index.php/es/facultad/noticias/448-expociencia-2018>

- Science disseminator for light, and electron microscopy techniques for the attending public
- Presenter of demonstration experiments with microscopes from Centro de Microscopía of Universidad de los Andes

April 2017 **Organizer**, *Python Bootcamp Uniandes, Universidad de los Andes*

<https://pythonbootcampuniandes.github.io>

- Developed and presented the solutions to the exercises proposed during the workshops
- Helped attendees with the fundamentals of [Python](#) and the use [Pandas](#) for basic data analysis

Languages

- **English:** Fluent (C1). IELTS certificate (7.0/9.0)
- **Spanish:** Native speaker

Tools and Technologies

Databases

SQL (Family): Proficient
Mongo: Proficient
Cassandra: Intermediate

Programming languages

Python: Proficient
C: Intermediate
Java: Basic
R: Basic

Others

Apache Spark: Basic
Git: Proficient
Linux: Proficient
Excel: Intermediate

Data Analytics

Apache Superset: Proficient
Chartio: Proficient
Looker: Intermediate
Quicksight: Intermediate
Tableau: Basic
Power BI: Basic

Data Science

Scikit-learn: Proficient
TensorFlow: Basic

Data Engineering

Apache Airflow: Proficient
Docker: Proficient
AWS Services: Proficient
Google Cloud: Intermediate