

Cristian Villamil



cavillamil98@gmail.com

+33 6 82 51 80 93

Data Analysis & Data Science Consultant

Mathematician and Data Scientist with experience in **medical data analysis, MRI-based classification, statistical modeling, and scientific research**. I have worked as a **Research Assistant at Universidad El Bosque**, contributing to interdisciplinary projects in neurology, environmental epidemiology, and public health.

Currently, I support a neurologist as a **freelance Data Scientist**, developing tools for **MRI preprocessing, clinical data interpretation, and automated analysis pipelines**.

My undergraduate thesis involved building a **machine learning classification model for Multiple Sclerosis using MRI data**, integrating feature engineering, mathematical modeling, and supervised learning.

Skilled in Python, R, SQL, geospatial analysis, and machine learning. Native Spanish speaker, B2 English, and learning French (A2).

Research Experience

Research Assistant — Universidad El Bosque

Engineering Faculty & Public Health Research Group
2023 – 2024

- Processed and analyzed NASA satellite datasets (temperature, air quality, NO₂, ozone) to identify correlations with **mortality and environmental health outcomes**.
- Used Python, GeoPandas, and spatial APIs to build geospatial pipelines and reproducible scientific workflows.
- Supported interdisciplinary research involving public health, environmental engineering, and applied mathematics.
- Contributed to data cleaning, statistical modeling, visualizations, and academic documentation.

Undergraduate Thesis

“Machine Learning Classification Model for Multiple Sclerosis Using MRI Data”

Bachelor of Mathematics — Universidad El Bosque, 2024

- Built a **machine learning classifier** to distinguish MRI patterns associated with Multiple Sclerosis.
- Performed image preprocessing (normalization, intensity adjustments, feature extraction).

- Applied dimensionality reduction, statistical modeling, and supervised learning techniques.
- Developed reproducible pipelines using Python and standard scientific libraries.

Profesional Experience

Neurology Data Science Consultant

Independent Project — Client: Neurologist
2023 – 2025

- Implemented statistical and machine learning models for **MRI datasets and clinical patient information**.
- Designed automated MRI-processing workflows, improving analysis efficiency by **20%**.
- Conducted data cleaning, feature extraction, and exploratory analysis for neurological datasets.
- Translated analytical findings into actionable insights for medical decision support.

Scientific & Data Science Projects

Environmental Data Correlation Project

Client: Universidad El Bosque Research Group
2024

- Processed NASA satellite imagery and geospatial datasets via APIs and ETL pipelines.
- Identified environmental variables associated with mortality (temperature, air quality).
- Developed geospatial models using Python, GeoPandas, rasterio, and spatial statistics.

Student Success Predictive Modeling — Diginexa / Universidad El Bosque
2024

- Built automated pipelines to assess academic performance from student grade datasets.
- Designed predictive models that increased early detection of at-risk students by **30%**.
- Improved reporting and intervention processes through data-driven insights.

Growth Specialist — Dataico
March 2025 – September 2025

- Applied quantitative analysis to optimize activation, retention, and funnel performance.
- Increased onboarding AHA-moment from 49% to 59%.
- Generated dashboards and behavioral reports for strategic decisions.

Additional Competencies

- **Technical Skills:** Python, R, SQL, ETL processes, Machine Learning, Data Visualization (Power BI).
- **Project Management:** Leadership, team coordination, agile methodologies.
- **Languages:** Spanish (Native), English (Advanced), French (Intermediate, currently learning in France).
- **Compliance & Regulations:** Experience handling personal data according to local regulations and data protection standards.

Core Skills

- Data Analysis / Data Science
- Statistical Modeling / Inferential Statistics
- Machine Learning / Neural Networks
- Programming Languages: Python, R, SQL
- ETL Processes / Data Extraction & Transformation
- Data Visualization: Tableau, Power BI
- Image Data Processing: Medical Imaging Analysis
- Satellite Imagery Analysis
- Predictive Modeling
- Process Optimization
- Automated Data Pipelines
- Project Management / Team Leadership
- Multidisciplinary Team Collaboration
- Agile Methodologies
- Data Compliance / Data Protection Standards

Languages

- English (B2)
- French (A2)
- Spanish (C2 – Native)

Education

- **Bachelor of Mathematics**, Universidad El Bosque