Sebastian David Pinzon Zambrano

Valledupar, Cesar · (57) 318-6087347 · sebastianpinzon1223@gmail.com · LinkedIn · Github

Experience

2023-2024 Imperio del Plastico   **Valledupar, Cesar**
Data Scientist

- Conceived and deployed advanced data analytics solutions using Python, R, SQL, and Excel, streamlining financial and operational workflows to reduce costs by 12%.
- Engineered predictive models and time series forecasting tools that refined forecast accuracy by 15%, empowering strategic planning and executive decision-making.
- Constructed robust ETL pipelines and simplified data cleaning frameworks that transformed raw data into actionable insights, boosting operational efficiency in accounting and finance by 63%.
- Directed the development of intelligent workflows and executed reporting systems with Python and SQL, increasing team productivity and reducing manual effort by 20%.
- Launched a profitable grill manufacturing startup in under three months by overseeing business model design, comprehensive market research, and full-scale operational execution, attaining profitability within the first quarter.
- Rolled out an integrated online store and last-mile delivery system by optimizing logistics with Python and managing backend data operations with SQL, laying the groundwork for a scalable, data-driven growth strategy.
- Formulated and executed targeted strategies based on statistical analysis and KPI dashboards, achieving sustained monthly revenue growth of 13% and stronger market positioning.

Education

2019-2023 National University of Colombia   **Manizales, Caldas**
Computer systems administration

- Studies postponed for family and financial reasons.

Certifications

2025 DeepLearning.AI   **Coursera**
DeepLearning.AI TensorFlow Developer

2025 DeepLearning.AI   **Coursera**
Deep Learning

2025 Amazon   **Coursera**
AWS Cloud Solutions Architect

2025 Google   **Coursera**
Google Professional Data Analytics Certificate

2025 Duke University   **Coursera**
MLOps | Machine Learning Operations

2025 Stanford University   **Coursera**
Machine Learning

2025 University of Michigan   **Coursera**
Applied Data Science with Python

2025 Johns Hopkins University   **Coursera**
Data Science: Fundamentals Using R

2025 University of Alberta   **Coursera**
Software Design and Architecture

2025 Imperial College London   **Coursera**
Specialization in Mathematics for Machine Learning

2025 University of Michigan   **Coursera**
PostgreSQL for Everybody

2024 IBM   **Coursera**
IBM Professional Data Science Certificate

Projects

Sales Prediction for "Favorita Grocery" (Kaggle)   *2025 – In Progress*
Data Science | Time Series | Reproducibility

I deployed a Docker container with PostgreSQL and TimescaleDB to optimize temporal data storage and ensure maximum reproducibility; versioned datasets and preprocessing pipelines with DVC linked to Google Drive (`dvc pull`), while GitHub Actions validates the Docker environment build; leveraged Apache Spark for distributed data cleaning and generation of temporal features (lags, seasonal and hierarchical aggregates); benchmarked and fine-tuned advanced models LightGBM, a WaveNet-style CNN (PyTorch/TensorFlow), GraphDeepAR (GluonTS/MXNet), and the Autoformer Transformer directly in Kaggle notebooks; and documented the entire architecture with UML

diagrams in Modelio, complemented by a practical step-by-step guide on GitHub.

[GitHub]   [Kaggle]

Customer Segmentation in Shopping Malls: Analysis and Marketing Strategies   *2025*
Data Science | Unsupervised Learning, Cluster Validation, Consumer Analytics
Applied advanced clustering techniques to segment customers using demographic and behavioral data, validated model robustness with multiple indices, and identified five distinct customer profiles that informed targeted marketing strategies, resulting in a 20% acceleration of campaign conversion rates.
[GitHub] | [Documentation]

Skills

Programming Languages: Python, C++, C, SQL, R, JavaScript, HTML, CSS, LaTeX.
Analysis Tools: Excel, Tableau, PowerBI, Scikit-learn, TensorFlow, Pandas, Matplotlib.
Databases and Big Data: MongoDB, PostgreSQL, Oracle, Spark, Hadoop.
Cloud Computing: AWS, Google Cloud, Azure.
Soft Skills: Leadership, Project Management, Effective Communication.

Honors & awards

2018 Secretary of Education   **Valledupar, Cesar**
Best Saber 11 Tests

About me

A cheerful, resilient and persistent professional, I am a Data Scientist & ML Engineer with more than a year of experience transforming data into tangible results (12% cost reduction and 15% forecast accuracy improvement). I am proficient in Python, SQL, R and cloud environments (AWS, GCP), and translate technical findings into high impact business strategies. My aspiration is to become a great leader capable of guiding organizations to success, driving multidisciplinary teams with vision, passion and collaboration.