



Andres Felipe Castaño Morales

DATA ENGINEER

✉ andres95fc@gmail.com

📍 Medellín, Colombia

☎ +57 3105392277

Profile

Experienced and dedicated Data Engineer with 3+ years of experience developing and maintaining ETL data pipelines, including the automation and orchestration of different processes in a Hadoop ecosystem. Also with experience working along data scientists building machine learning models for banking risk management.

Employment History

Data Engineer, Bancolombia (Specialized Analytics Department), Medellín

January 2022 — Present

- Co-developed a Python package for measuring Data Quality into an orchestrated flow with Python which is now being used widely across the organization along with monitoring reports for bad quality data identified by the package to generate warnings in the process.
- Creating coherent Logical Data Models that help guide important client business decisions using Data Warehousing for an OLAP system.
- Preparing and presenting reports, analysis and presentations to various stakeholders including executives using Power BI.
- Teaching Python, SQL and best practices to support continuous process automation for data ingestion and data pipeline workflows across the department using Azure DevOps platform, explaining how to implement different pipelines for CI/CD stages.

Data Engineer, Bancolombia (Risk Analytics Department), Medellín

December 2020 — January 2022

- Worked along data scientists developing automated ETL data pipelines (using Spark with Pyspark, Hive, Python and Impala SQL) that periodically build datasets needed for machine learning models such as customer risk scores and income estimation, ensuring high quality data.
- Created automated orchestrators capable of running periodically machine learning models using Azure DevOps under CI/CD methodology, acting as an administrator of the different repositories and main responsible in the acceptance of new features and fixes of the orchestrated packages.
- Developed fixes and new features using OOP, Pandas, Matplotlib, Seaborn, for python tools created in the department which were useful for backtesting purposes, data visualization, among others.

Junior Data Engineer, Bancolombia (SMEs and Enterprises Analytics Department), Medellín

May 2019 — December 2020

Links

[LinkedIn](#)

Skills

SQL

Python

ETL

Microsoft Power BI

Data Analysis

Data Warehousing

Data Modeling

Machine Learning

Git

Apache Spark

Microsoft Excel

Azure DevOps

REST APIs

NoSQL

Airflow

AWS

GCP

Languages

English

Spanish; Castilian

- Built automated ETL data pipelines (Extracted data from different sources such as csv files, webpages using webscraping and datalake, transformed data according to business rules and quality criteria, and then load it with specific data structure into a platform called Visor) that used financial and demographic data to determine the best suitable products for the customers to obtain such as credit cards, insurances, loans, etc.
- Created an orchestrator capable of running several ETL data pipelines periodically using a library called Luigi.

Education

Bachelor in Telecommunications Engineering, University Of Antioquia, Medellín

April 2013 — December 2020

References

Santiago Machado from Bancolombia

+57 3136679679

Yorman Quiceno Giraldo from Liberty Mutual Insurance

+57 3117798942