



Jesid Mauricio Mejía Castro

Computer and Systems Engineer

I am a professional from the National University of Colombia with an ongoing master's degree in engineering and data analytics from the University of Bogotá Jorge Tadeo Lozano.

I have experience working with financial services companies and analytics agencies. My 4 years of experience encompass the usage of database tools like Oracle and SQL Server Databases, ETL development tools like Infosphere DataStage, Integration Services. Recently, I have been getting involved with The Spark ecosystem through Databricks and PySpark. My programming skills are linked with languages like Python and R. In my free time I experiment with Unix-like operating systems, the programming languages Rust, Go and Julia (still at beginner level) and free software.

I keep an obsessive learning spirit and I like to promote technical excellence and best practices in the final products. I'm very compromised with honesty and ethics in the workplace.

Personal data

Email jmauriciomejia@gmail.com
Mobile +351 915180671
Age 31

Education

- 2010–2017 **Bachelor's degree in Computer and Systems Engineering,**
Universidad Nacional de Colombia, Sede Bogotá
- 2020–2021 **Master's degree in engineering and data analytics,** *University of Bogotá Jorge Tadeo Lozano, Bogotá*
Ongoing

Experiencia Laboral

2016 – 2017 **Intern,** *Enel Codensa E.S.P,* Bogotá, Colombia



In this public service company I completed my internship with the Customer Ombudsman Office. My responsibilities were the introduction of new adjustments to the local database, the calculation and presentation KPIs, among others *in-house* developments, mainly with MS Access and VBA.

2018 – 2020 **Data Analyst - Data Engineer**, *Banco de Bogotá S.A.*, Bogotá, Colombia



Between 2018 and 2019 I worked with the Loans Division of the bank making new developments with the local database system (MS SQL Server 2012) and building reports (mainly in Excel) aimed to support the operation of credit studies for the bank customers. Between mid-2019 and early 2021 I worked in the Governance and Data Quality Office developing ETLs with IBM DataStage and Oracle PL/SQL through an on-premise data warehouse. My responsibilities were the measuring of the data quality from multiple transactional systems of the organization.

2021 **Data Warehouse Professional**, *Credivalores S.A.*, Bogotá, Colombia



Between February 2021 and May 2021 I worked in the Data Warehouse Department of this financial services company developing the information requirements of the other departments. Each member of the team received the initial requirements and our responsibility was to develop them from start to finish, from analysis and design to presentation of indicators and tests. SQL Server 2017 was used for the storage, while Integration Services was used as the ETL tool, Power BI was the visualization solution (it was a Microsoft-oriented environment). The Kimball data warehousing practices were heavily used during this period.

2021 **Junior Data Engineer**, *Ágata S.A.S*, Bogotá, Colombia



Since June 2021 I have been working at Data Analytics Agency (Ágata) created as an initiative of the Capital District and meant to be the catalyst for the development of a *smart city* in Bogotá. I briefly played the role of data quality manager (2 months) in this company. Later I was promoted to the position of Junior Data Engineer. My work at the Agency consists of preparing the data for the different projects that the Agency begins to contract with the various entities of the district. My activities include data profiling, data cleaning, data pipeline design and implementation and ETL development, dimensional analysis, data quality measurements, my main responsibility here is to deliver clean, useful and high-quality data to the data scientist of the company. different analytical requirements. All of the above must to be done under the data governance guidelines defined by the Agency. Considering itself as a *cloud-agnostic* organization, the Agency uses different clouds, specifically: Microsoft Azure, Google Cloud Platform and Amazon Web Services. In my particular case, I worked with Azure components (Azure Data Factory and Azure Databricks) and to a lesser extent, with some GCP components (BigQuery, Dataflow, etc.).

2022–2023 **Data Engineer**, *Promptly Analytics Health Services*, Oporto, Portugal



Since March 2022 until May 2022 I worked in a health analytics company. My work consisted on processing the clinical data coming from the products developed by the company and the different data integrations we had, so we can introduce this data in a our data warehouse to obtain insights oriented to improve the patients treatment. I used technologies like Apache Airflow and Pandas for ETL orchestration and implementation; and some AWS services like S3 for datalake storage with parquet files. Along with standard tools like Git, Gitlab, Docker, etc.

2023–now **Data Engineer**, *Continental*, Lousado, Portugal



Since July 2023 I started working in Continental. This company developed several products based on an IoT architecture where a sensor is installed inside the tires and this devices through a control unit are constantly uploading data to AWS S3. My work consists in maintaining the infrastructure definition through Infrastructure as Code with CDK. To store the data we use AWS S3 and to process the data we use AWS Glue and Python (PySpark and AWS Wrangler). To orchestrate the workflow we use Apache Airflow.

Skills

| | | | |
|--------------------------------|---|---------------------------------------|--|
| Oracle | Modeling, DDL, DML and PL/SQL | SQL Server | Modeling, DDL, DML and T-SQL |
| IBM Infosphere DataStage | Design and implementation of ETLs | SQL Server Integration Services | ETL Development |
| Python | General purpose programming. Data analysis and processing with PySpark and Pandas. Data visualization with matplotlib, among other packages modules | R | Basic programming, descriptive statistics, linear regression models. |
| Git | Collaborative work, experience with GitFlow | GNU/Linux | General administration tasks, familiarity with bash. |
| Amazon Web Services | Storage with AWS S3, ETLs with AWS Glue, infrastructure definition with CDK, and other services like AWS Lambda, AWS CodePipeline, etc. | Apache Airflow | Data pipeline orchestration |