

---

# JUAN DANIEL GARCIA BELMAN

## Data Engineer | ETL & Cloud Data Solutions

Celaya, Querétaro, México | +52 (461) 173-3822 | danielgb331@outlook.com

LinkedIn: juan-daniel-garcia-belman | GitHub: Daniel-jcVv | Portfolio: danieljcvv-portfolio.vercel.app

---

### PROFESSIONAL SUMMARY

Data Engineer with **3+ years** of experience building ETL/ELT pipelines and data solutions for Fortune 500 clients (**Telcel, Citi Banamex**) at **NTT DATA**. Specialized in **Python (PySpark, Pandas), SQL, Apache Airflow, and multi-cloud platforms (Azure, AWS, GCP)** for scalable data processing. Experienced in data warehouse development, pipeline automation, data quality frameworks, and data governance across enterprise data lakes and warehouses. Proven track record in optimizing ETL operations, managing metadata governance, and delivering data integration solutions for banking and telecommunications sectors.

---

### TECHNICAL SKILLS

**Programming & ETL:** Python (PySpark, Pandas), SQL (T-SQL, PostgreSQL, Oracle), Apache Airflow, Apache Kafka, Shell Scripting

**Cloud Platforms:** Azure (Data Factory, ADLS, Synapse), AWS (Glue, S3, RDS, Lambda), GCP (BigQuery, DataProc, Cloud Storage)

**Databases & Storage:** MS SQL Server, PostgreSQL, Oracle, Apache Hive, Teradata, Delta Lake

**Data Engineering:** Medallion Architecture, Data Warehousing, Dimensional Modeling (Star/Snowflake), Incremental Loading, Data Quality, Data Governance (DAMA-DMBOK principles)

**DevOps & Tools:** Docker, Git, Jenkins, CI/CD, JIRA/Confluence, IBM DataStage, Ab Initio

---

### PROFESSIONAL EXPERIENCE

#### NTT DATA - Data Pipeline Operations Engineer | Telcel Client

MAY 2024 - APRIL 2025

Managed ETL pipeline operations and monitoring in IBM DataStage, ensuring **99%+ uptime** for production data workflows processing **millions of daily telecommunications records**

Performed data validation and quality checks using **SQL (Oracle)**, implementing automated data quality rules that identified and resolved data integrity issues

Coordinated secure data file transfers from remote servers to enterprise data lakes using **SFTP protocols** for ETL pipeline ingestion

Automated operational workflows using **Bash scripting and Python** for file management, log monitoring, and reporting

#### NTT DATA - Data & Metadata Governance Analyst | Citi Banamex Client

FEBRUARY 2023 - MAY 2024

Defined and standardized naming conventions for data attributes across **Data Lakes, Data Marts, and Data Warehouses** using **Hive and Teradata**, following DAMA-DMBOK governance principles

Executed data lineage and field mapping strategies ensuring data integrity and traceability across multiple enterprise banking systems

Managed change requests through **JIRA/Confluence**, facilitating communication between ETL development teams and business stakeholders

## NTT DATA - Data Integration & Business Analyst | Citi Banamex Client

FEBRUARY 2022 - FEBRUARY 2023

Created source-to-target mapping documentation for banking data reconciliations using **SQL and Python**

Developed **Python scripts (OpenPyXL)** for automated validation of monthly attributes in banking reconciliation processes

Collaborated with ETL teams using **Ab Initio** for Data Warehouse migration projects

Provided analytical support using **IBM Cognos Analytics** to generate strategic reports for business stakeholders

---

## SELECTED PROJECTS

### Insurance Analytics Platform - Medallion Architecture

[github.com/Daniel-jcVv/insurance-analytics-medallion](https://github.com/Daniel-jcVv/insurance-analytics-medallion)

End-to-end analytics platform with Bronze/Silver/Gold layers using Microsoft Fabric and PySpark for insurance data processing and Power BI dashboards

*Microsoft Fabric, PySpark, Power BI, Medallion Architecture*

### Enterprise Data Warehouse - Medallion Architecture

[github.com/Daniel-jcVv/modern-datawarehouse-project](https://github.com/Daniel-jcVv/modern-datawarehouse-project)

Production-ready data warehouse with Bronze/Silver/Gold layers processing CRM/ERP records using SQL Server, Python ETL, and Docker containerization

*SQL Server, Python, Docker, Medallion Architecture, Delta Lake*

### Data Migration Pipeline with File Tracking

[github.com/Daniel-jcVv/pipeline-data-migration-tracker](https://github.com/Daniel-jcVv/pipeline-data-migration-tracker)

Automated SFTP-to-Fabric migration pipeline with SQLite-based tracking system preventing duplicate file processing

*Microsoft Fabric, Python, SQLite, SFTP*

---

► View complete portfolio with **6+ data engineering projects** at [danieljcvv-portfolio.vercel.app](https://danieljcvv-portfolio.vercel.app)

---

## PROFESSIONAL CERTIFICATIONS

- **IBM Data Warehousing Engineer** - IBM (Coursera) | June 2024

Credential: [coursera.org/account/accomplishments/specialization/87W6HBAU3QDQ](https://coursera.org/account/accomplishments/specialization/87W6HBAU3QDQ)

- **ETL and Data Pipelines with Shell, Airflow and Kafka** - IBM (Coursera) | January 2023

Credential: [coursera.org/account/accomplishments/verify/QMZNBTZS4QMU](https://coursera.org/account/accomplishments/verify/QMZNBTZS4QMU)

- **Google Data Analytics** - Google (Coursera) | June 2022

Credential: [coursera.org/account/accomplishments/professional-cert/WLKVK7PBLVVU](https://coursera.org/account/accomplishments/professional-cert/WLKVK7PBLVVU)

- **Java Back-End Developer** - Oracle Next Education | July 2024

---

## EDUCATION

**Manufacturing Technologies Engineering** - Universidad Politécnica de Guanajuato

JUNE 2015

---

## LANGUAGES

**English:** Professional Working Proficiency (B2) - Technical documentation, professional communication, Agile ceremonies

**Spanish:** Native Speaker