

# GURUDAS JADHAV

Mumbai | +91 8104843224 | gurudas.jadhav.work@gmail.com

## SUMMARY

Passionate **Data Engineer** with hands-on experience in **Python, SQL, PySpark, and Databricks** for developing and maintaining scalable **ETL pipelines**. Proficient in **data wrangling, data modeling**, and workflow orchestration using cloud platforms like **AWS** and **Databricks**. **Strong analytical** mindset with a passion for transforming raw data into actionable insights to support data-driven decisions..

## TECHNICAL SKILLS

- **Languages:** Python, SQL, Java
- **Cloud Platforms:** AWS (S3, Glue, Athena, IAM), Azure Databricks
- **Databases/Warehouses:** Snowflake, Microsoft SQL Server, MySQL, MongoDB
- **Frameworks & Tools:** PySpark, Pandas, NumPy, Git/GitHub, Power BI, Matplotlib, Seaborn

**Keywords:** Cloud Data Warehousing, Data Pipelines, ELT/ETL, Serverless ETL, Performance Tuning, Data Governance, Big Data, PySpark, Snowflake Development, Workflow Automation, Version Control

## WORK EXPERIENCE

**Operational Application (Data Domain) Intern** SULZER INDIA PVT. LTD | Airoli, Navi Mumbai *Feb 2025–May 2025*

- **Performance Tuning:** Accelerated data retrieval speed by **60%** by designing and performance-tuning over **15 stored procedures and 10 Common Table Expressions (CTEs)** in **Microsoft SQL Server**.
- **ELT Pipeline Development:** Developed and optimized scalable **ELT pipelines** utilizing **Snowflake** to ingest, transform, and manage high-volume structured and semi-structured data sources.
- **Data Quality & Governance:** Implemented stringent **data quality checks** and **workflow automation** to enforce data integrity and **Data Governance** across production datasets, ensuring data reliability.
- **Database Administration:** Managed critical database administration tasks, including user access management and security protocol enforcement, ensuring continuous data availability.

## DATA ENGINEERING PROJECTS

**VEHICLE SALES DATA LAKE & PIPELINE – AWS (S3 / Glue / Athena / QuickSight)** | Dec 2024

- **Cloud Architecture:** Built a functional **Data Lake architecture** using **AWS S3** for raw storage and **AWS Glue** for serverless **ETL processing** and cataloging.
- **Query Optimization:** Utilized **AWS Athena** for complex analytical queries, delivering query results **50% faster** than traditional data reporting methods.
- **Access Control:** Managed secure resource access and **Data Governance** using **AWS IAM** policies and created interactive sales dashboards in **QuickSight**.

**TRAFFIC VIOLATION ETL PIPELINE – Databricks / PySpark / Python** | March 2025

- **Big Data Processing:** Engineered an end-to-end **ETL pipeline** leveraging **PySpark** running on **Azure Databricks** to efficiently process raw traffic violation data.
- Pipeline performed comprehensive data cleaning and transformation logic before generating customized output files tailored for three different traffic departments, **improving reporting accuracy and accessibility**.

**DIWALI SALES ANALYSIS & REPORTING – Python / Pandas / Matplotlib** | Aug 2024

- Performed comprehensive sales data analysis using **Python libraries (Pandas, Matplotlib, Seaborn)** to identify top-selling products and market trends.
- Delivered data-driven insights that **improved targeting efficiency** and identified high-value customer segments for marketing teams..

## EDUCATION

**Master's of Computer Application (MCA)**

GURU NANAK INSTITUTE OF MANAGEMENT STUDIES - 71.62%

**Bachelor's of Computer Application (BCA)**

SAHYOG COLLEGE - 84.48%

## CERTIFICATIONS

- Python for Data Science (NPTEL), Big Data Computing (NPTEL) | Jul 2024 - Sep 2024
- Python for Data Analysis, SQL for Data Science, MongoDB Tutorial ,NumPy, Pandas, Matplotlib | Aug 2024
- SQL Basic (5 Star Gold Badge In HackerRank) | March 2025