

MOHAMMAD SAIFI

+91 7209538634 | saifi.live | [LinkedIn](#) | [GitHub](#) | saifimd1234@gmail.com | Jamshedpur, JH

EDUCATION

PG DIPLOMA IN BIG DATA ANALYTICS – Centre for Development of Advanced Computing, Pune	03/2024 – 09/2024
BACHELOR OF ENGINEERING – RVS College of Engineering and Technology, Jamshedpur	06/2020 – 07/2023
DIPLOMA – Al-Kabir Polytechnic, Jamshedpur	06/2015 – 07/2018
MATRICULATE – Kerala Public School, Jamshedpur	03/2008 – 03/2015

SKILLS

- SQL (SQL Server, MySQL, PostgreSQL, BigQuery)
- Python (Pandas, NumPy, Scikit-Learn, SciPy, Matplotlib, Seaborn)
- Tableau (Data Visualization, Dashboard creation, Data Blending, Storytelling)
- HTML, CSS, JavaScript
- Git and Github (Version Control, Push, Pull, Branching, Merging, Collaborative Workflow)
- AWS (IAM, ECR, EC2, S3, Basic Cloud Computing)
- CI/CD Pipelines (GitHub Actions, Workflow Automation)
- Advance Excel (VLOOKUP/XLOOKUP, conditional formatting, complex IF functions, pivot tables, DAX calculations, Power Pivot)
- Linux (Debian)
- MLOps, Dagshub
- Docker(Containerization, Image Creation, Docker Compos

PROJECTS

MODERN DATA WAREHOUSE & ETL PIPELINE DEVELOPMENT

| PG Diploma Project | July 2024 | [View Project](#)

Objective: Design a modern SQL data warehouse using **Medallion Architecture** (Bronze, Silver, Gold layers) to consolidate sales data from multiple sources (ERP, CRM).

- Developed ETL pipelines using T-SQL, optimizing batch processing, incremental loads, and full refresh mechanisms.
- Revamped data transformation processes using stored procedures and functions, which enabled data scientists to process 5x more data for machine learning and business intelligence.
- Engineered a star schema in the Gold Layer, optimizing fact and dimension tables for analytical querying and performance.
- Conducted query optimization, indexing strategies, and partitioning to improve SQL query performance and reduce execution time.
- Ensured data integrity and governance, implementing validation checks, constraints, and documentation for seamless scalability.

Technical Stack: Python, scikit-learn, Pandas, Numpy, AWS (ECR, EC2, S3), Git, GitHubActions, MLOps, Dagshub, Docker, FastAPI

DATA-DRIVEN INSIGHTS GENERATION | BUSINESS INTELLIGENCE & VISUALIZATION

| Personal Project | July 2024 | [View Project](#)

Objective: To drive data-driven decision-making by performing **exploratory data analysis (EDA)** on large datasets, identifying revenue drivers, optimizing reporting, and delivering actionable insights through business intelligence visualizations.

- Conducted end-to-end EDA to uncover trends in 295 products and 18K+ customers, resolving data quality issues (e.g., deduplication via DISTINCT, outlier detection).
- Created actionable dashboards (Power BI/Tableau) using KPIs like customer lifetime value (CLV), average order value (AOV), and product revenue segmentation.
- Reduced report generation time by 50% via SQL-driven automation; delivered 10+ insights on customer/product behavior.
- Automated part-to-whole comparisons (e.g., regional sales contribution) using SUM() OVER() for real-time reporting.
- Performed cohort analysis to segment customers by recency, frequency, and monetary (RFM) metrics, identifying 20% of VIP customers driving 60% of revenue.

Technical Stack: SQL, Excel, Statistical Analysis, Tableau, Data Visualization

EXPERIENCE

BLUE BELLS HIGH SCHOOL, DIMNA — SENIOR MATHS TEACHER

03/2019 – 03/2020

- Pioneered a mentorship program pairing advanced students with peers struggling in logical reasoning; as the only mentor, contributed towards a 90% improvement rate on weekly quizzes.
- Designed and delivered interactive lessons incorporating real-world examples and collaborative activities, fostering deeper understanding and improving student participation.

CERTIFICATION

- HackerRank SQL (Basic and Intermediate)
- HackerRank Python (Basic)
- Essential Business Communication Skills – LinkedIn Learning Professional Certificate
- Data Visualisation: Empowering Business with Effective Insights Offered by Tata Group on Forage

ACHIEVEMENTS

- Earned LeetCode Top 50 SQL Badge by solving complex SQL problems, showcasing advanced query optimization skills.
- Secured 2nd place in State Powerlifting Championship, demonstrating discipline and perseverance.
- Earned Gold Badge in SQL on HackerRank, demonstrating advanced proficiency in writing complex and solving real-world data problems.