ROSHAN FAREED N

Data Engineer & Data Analyst | Azure · Databricks(PySpark)· SQL · Python · Power BI

✉ roshanfareed53@gmail.com   📞 +91 86084 00363

📍 Alagankulam, Ramanathapuram, Tamil Nadu - 632512   in linkedin.com/in/roshanfareed53

github.com/ROSHANFAREED   🔗 https://ROSHANFAREED.github.io

Professional Summary

Data Engineer and Analyst with hands-on experience designing and implementing end-to-end data pipelines and analytics solutions using Azure Data Factory, Databricks (PySpark), ADLS Gen2, Azure SQL, and SQL Server. Strong in ETL development, data transformation, and analytics-ready data modeling, with the ability to deliver reliable datasets for reporting and decision-making. Experienced in building interactive dashboards and business insights using Power BI, Tableau, SQL, and advanced Excel. Proven ability to work with structured and semi-structured data, apply scalable data architecture practices, and support data-driven outcomes, with a foundational understanding of machine learning concepts.

Core Skills

Data Engineering & Cloud (Azure):
- Azure Data Factory (ADF), Azure Databricks (PySpark, SparkSQL), ADLS Gen2, Azure SQL Database, Azure Synapse (exposure), Microsoft Fabric (exposure)

Programming & Data Processing:
- Python (Pandas, NumPy), PySpark, SQL, Spark SQL

Version Control & Dev Practices:
- Git, GitHub, Basic CI/CD concepts, Secure coding practices

Databases & Data Modeling:
- MySQL, SQL Server, Joins, Subqueries, Window Functions, Stored Procedures, Aggregations

Business Intelligence:
- Power BI (DAX, Data Modeling, Dashboards), Tableau, Advanced Excel

Analytical & Professional Skills:
- Analytical Thinking, Data Storytelling, Problem Solving, Stakeholder Communication, Documentation & Reporting.

Projects (GitHub Portfolio)

COVID-19 Azure Data Engineering Pipeline (Flagship Project),
Tech Stack: Azure Data Factory, Databricks (PySpark), ADLS Gen2, Azure SQL, Power BI, GitHub
- Designed and implemented an end-to-end Azure Data Engineering pipeline ingesting large public COVID-19 datasets from APIs and blob storage using Azure Data Factory.
- Built a scalable data lake architecture in ADLS Gen2 with **Raw and Clean zones**, following industry-standard data lake best practices.
- Performed data transformation, cleansing, and aggregations using **ADF Mapping Data Flows and Databricks (PySpark)** to prepare analytics-ready datasets.
- Loaded curated datasets into **Azure SQL Database** to support downstream reporting and analytical workloads.
- Developed interactive **Power BI dashboards** to analyze COVID-19 cases, hospital admissions, testing trends, and population-level insights.
- Implemented secure pipeline practices by externalizing secrets and structuring the solution for **CI/CD readiness**.

Retail Sales Data Engineering Pipeline (Medallion Architecture),
Tech Stack: Databricks, PySpark, Delta Lake, SparkSQL, Power BI, CSV/JSON
- Built an end-to-end ETL pipeline using Bronze–Silver–Gold layers to process raw retail data.
- Performed data cleaning, schema enforcement, null handling, and multi-format date parsing using PySpark.
- Created Gold-level aggregates including **Total Sales by Category, Order Trends, AOV, and Avg Price per Item**.
- Designed an interactive analytics dashboard visualizing customer behavior and revenue insights.
- Implemented transformation logic using Delta Lake for optimized storage and incremental updates

E-commerce Online Retail Analytics, *MySQL, SQL Server, Power BI, DAX*

- Designed analytical data models in **MySQL and SQL Server** and wrote **complex SQL (joins, subqueries, window functions)** to analyze sales and customer behavior.
- Built **optimized Power BI dashboards** using advanced **DAX, drill-throughs, and interactive visuals** to deliver actionable insights.

Professional Experience

Quality Inspector, *Pro Candour Automotive Pvt Ltd, Chennai*   2019 – 2021
- Conducted end-to-end quality inspections across production lines, ensuring **100% compliance** with defined quality standards and specifications.
- Developed Excel-based reporting system that reduced manual reporting time by 40%.
- Built visual dashboards for weekly review meetings, enabling management to make quicker decisions.
- Collaborated with production and engineering teams to analyze defect trends and implement corrective actions, contributing to a **15% reduction in defect rates**.

Education

B.E. Mechanical Engineering, *AMS College of Engineering, Anna University*   2019
CGPA: 6.56/10

Key Achievements

- Completed **end-to-end data engineering and analytics projects**, demonstrating strong understanding of ETL pipelines, PySpark transformations, SQL, and dashboard development.
- Automated reporting workflows using Excel & Python, reducing manual effort and improving accuracy.
- Improved quality control in automotive operations by **reducing defect rates by 15%** through data-driven insights.

Certifications

Microsoft PL-300 (Power BI Data Analyst)	**Microsoft Power BI Analyst Certificate – Coursera**	**Google Data Analytics Professional Certificate – Coursera**
– In Progress	(Completed)	(Completed)

Availability:

Immediate Joiner | Open to Relocation

portfolio : https://ROSHANFAREED.github.io