

ROSHAN FAREED N

Aspiring Data Engineer | PySpark · Databricks · SQL · Azure

✉ roshanfareed53@gmail.com ☎ +91 86084 00363
📍 Alagankulam, Ramanathapuram, Tamil Nadu - 632512 🔗 linkedin.com/in/roshanfareed53
🐙 github.com/ROSHANFAREED 🔗 https://ROSHANFAREED.github.io

Professional Summary

Aspiring Data Engineer skilled in building scalable ETL/ELT pipelines using Databricks, PySpark, ADF and SQL. Experienced in data modelling, cleaning, transformation and developing dashboards for analytical insights. Completed 10+ end-to-end data engineering and analytics projects applying Medallion architecture, Delta Lake, and cloud data engineering concepts. Strong engineering foundation combined with analytical mindset to deliver reliable and efficient data solutions.

Core Skills

Databases & SQL: MySQL, SQL Server, Joins, Subqueries, Window Functions, Stored Procedures, Aggregations	Data Engineering & Cloud: Databricks (PySpark, SparkSQL), Delta Lake, ADF, Azure Synapse, Azure Fabric, DBT
Programming: Python (Pandas, NumPy), PySpark, Git/GitHub, Linux	Business Intelligence: Power BI (DAX, Drillthrough, KPI dashboards), Tableau
Data Handling & Analysis: ETL Concepts, Data Cleaning, Data Validation, Exploratory Data Analysis	Other Tools: Airflow, Kafka, Snowflake, BigQuery, Excel

Projects (GitHub Portfolio)

Retail Sales Data Engineering Pipeline using Databricks (Medallion Architecture, Tech: Databricks, PySpark, Delta Lake, Power BI, CSV/JSON, SparkSQL) • 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	11/2025 – 12/2025
E-commerce Online Retail Analytics, MySQL, SQL Server, Power BI • Implemented advanced DAX measures and drillthroughs, improving dashboard performance by 25%. • Delivered insights on purchase patterns and profitability drivers, leading to strategic recommendations.	

Uber Rides Analytics Dashboard (Power BI | Kaggle Dataset)

- Built an end-to-end **Uber rides analytics dashboard** using Power BI, analyzing 149K bookings across 2024 to extract customer, driver, and operational insights.
- Processed and transformed raw Kaggle data using Power Query; engineered time-based features and validated data quality (completeness, consistency, outliers).
- Designed interactive visualizations highlighting **peak demand hours, payment preferences, booking completion rates, customer vs. driver satisfaction**, and revenue trends.
- Delivered KPIs such as **Total Bookings, Total Revenue (\$51M), Distance Traveled (2.49M miles), Avg Customer Rating (4.40)**, supporting business decision-making.

Professional Experience

Quality Inspector, Pro Candour Automotive Pvt Ltd, Chennai

2019 – 2021

- Conducted end-to-end inspections ensuring 100% compliance with quality standards.
- 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, reducing defect rates by 15% through corrective actions.

Education

B.E. Mechanical Engineering, AMS College of Engineering, Anna University

2019

CGPA: 6.56/10

Key Achievements

- Completed **10+ 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.
- Successfully transitioned from Mechanical Engineering to Data Engineering through structured learning, hands-on projects, and cloud technology skills.

Certifications

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

Additional Projects:

Prison Data Analysis | Customer Segmentation | Financial KPI Dashboard and more
<https://github.com/ROSHANFAREED?tab=repositories>

Availability:

Immediate Joiner | Open to Relocation