

# Suraj korishetti

Belgaum, Karnataka, India surajkorishetti111@gmail.com 6363498124 in/suraj-korishetti-333a44258 github.com/Suri6363498/Suri

## SUMMARY

Dynamic Azure Data Engineer intern with expertise in big data processing, ETL development, and cloud-based data integration. Skilled in Azure Data Factory, Azure Databricks, PySpark, ADLS, Delta Lake, and Azure SQL, with a proven track record of designing scalable pipelines, optimizing Spark jobs, and ensuring data quality. Adept at end-to-end solution design, automating deployments via CI/CD, and delivering high-volume, high-accuracy datasets for analytics and business intelligence.

## EXPERIENCE

### Microsoft Power Bi Data Analyst Associate Internship

ICT Academy June 2024 - October 2024, Belagavi

- Designed and implemented Azure Data Factory pipelines to ingest and transform **1M+ records** from SQL, Oracle, SAP, and ADLS into Azure Data Lake, reducing data processing time by **30%**.
- Developed and optimized Delta Live Tables in Azure Databricks using PySpark and Spark SQL, improving transformation efficiency by **25%** and reducing runtime errors by **1**.
- Designed and implemented Azure Data Factory pipelines to ingest and transform **1M+ records** from SQL, Oracle, SAP, and ADLS into Azure Data Lake, reducing data processing time by **30%**.
- Automated CI/CD deployments for ADF pipelines, achieving **100% deployment consistency** across environments and reducing manual intervention by **80%**.
- Built ETL workflows with Linked Services/Datasets to integrate data from **5+ sources** into Blob Storage, Data Lake, and Azure SQL, enabling faster analytics delivery.
- Implemented robust data quality checks, improving data accuracy to **99%** and ensuring compliance with business and technical rules.
- Scheduled and monitored **20+** production jobs, resolving pipeline issues within SLA and improving system uptime by **10%**.

## PROJECT

### Azure Data Engineering Pipeline

Freelance • June 2025 - July 2025

- Engineered Azure Data Factory pipelines to **extract, transform, and load (ETL)** over **1M+ records** from SQL, Oracle, SAP, and ADLS into Azure Data Lake, reducing ingestion time by **30%**.
- Developed and optimized Delta Live Tables in Azure Databricks using PySpark and Spark SQL, improving transformation efficiency by **25%**.
- Automated deployments via CI/CD pipelines, ensuring **100%** deployment accuracy across environments.
- Performed large-scale data validation, reconciliation, and optimization to achieve **99% accuracy**.
- Mentored junior engineers on **big data best practices** in Azure and PySpark.

### Retail Data Engineering Pipeline

Microsoft Learn • June 2025 - July 2025

- Designed and implemented an end-to-end data pipeline to process retail sales data from multiple sources using Azure Data Factory & Databricks
- Automated ETL processes to clean, transform, and load over **1M+ records** monthly, improving data accuracy by **25%**.
- Performed large-scale data validation, reconciliation, and optimization to achieve **99% accuracy**.
- Mentored junior engineers on **big data best practices** in Azure and PySpark.

### Sales Forecasting Dashboard

self-made • June 2025 - July 2025

- Built an interactive Power BI dashboard to forecast sales trends using historical data and time series models.
- Improved forecast accuracy by **15%** and enabled stakeholders to make data-driven inventory decisions.

## EDUCATION

### Bachelor of Engineering

SGBIT Belagavi • Belgaum • 2024 • 6.85

## CERTIFICATIONS

### Deloitte-Data Analytics Job Simulation

Forage • 2025

- Analyzed client data to identify key risks, applied strategic frameworks to develop solutions, and prepared a presentation to communicate recommendations.

### Microsoft Power BI Data Analyst Associate

ICT Academy • 2023

- Assessed client data to uncover key risks, applied strategic frameworks to design solutions, and delivered recommendations through a structured presentation.

## AWARDS & HONORS

### KIMO Edge

KIMO • 2023

- Certificate of 1st prizewinner, Secured the 1st position with a score of 99/120.

## SKILLS

Programming & Querying: SQL, Python, R, SAS, MATLAB.

Big Data & Cloud: Azure Data Factory, Azure Databricks, Azure Data Lake Storage, Delta Lake, Apache Spark, Hadoop

ETL & Data Integration: Data ingestion, transformation, and orchestration, Delta Live Tables, CI/CD pipelines, workflow automation

Data Quality & Optimization: Data validation, reconciliation, performance tuning, data governance

Visualization & Reporting: Power BI, Excel

Tools & DevOps: GitHub, Azure DevOps