

# Kishor S

7092840901 | [kishorsenthilkumar104@gmail.com](mailto:kishorsenthilkumar104@gmail.com) | [linkedin.com/in/kishor-s-5b4332274](https://www.linkedin.com/in/kishor-s-5b4332274) | [github.com/Kishorsenthilkumar](https://github.com/Kishorsenthilkumar)

## PROFESSIONAL SUMMARY

---

Early-career Data Engineer skilled in designing scalable data solutions using Apache Spark, Airflow, and Hadoop. Built efficient ETL pipelines and real-time streaming with Kafka for improved data processing. Strong coding abilities in Python, Java, and SQL on cloud platforms. Looking to use this expertise to support robust data infrastructure and enable smarter business insights.

## EDUCATION

---

### Panimalar Engineering College

*Bachelor of Technology, Artificial Intelligence and Data Science* (GPA: 8.1/10)

2022 - 2026

### Narayana Junior College

*HSC* (GPA: 8)

2021 - 2022

### Narayana Co School

*SSLC* (GPA: 9)

2019 - 2020

## PROJECTS

---

### Real-Time Industrial IoT Damage Monitoring Pipeline

- Built an end-to-end real-time IoT data pipeline to monitor JCB construction machinery damage using phone accelerometer as a real IoT device sending data via MQTT protocol to Azure IoT Hub
- Implemented PySpark Structured Streaming to consume data from Azure Event Hub, apply damage detection logic (vibration, temperature, strain thresholds) and write processed records to PostgreSQL using `foreachBatch`
- Configured Azure IoT Hub device registry with SAS token authentication and TCP-SSL encryption (port 8883) for secure device-to-cloud communication
- Designed a time-series PostgreSQL schema with optimized data types (`SMALLINT` for flags, `DECIMAL` for sensor values) and built live Grafana dashboard for real-time damage visualization and alerting

### Redis Server Implementation from Scratch

- Built a Redis-compatible TCP server in Python using `asyncio`, implementing the RESP (Redis Serialization Protocol) wire protocol from scratch without any Redis libraries
- Built master-replica replication from first principles - `REPLCONF` handshake, `PSYNC` negotiation, RDB snapshot transfer, and real-time write propagation to replicas
- Implemented core data structures - strings with TTL expiry, linked lists (`LPUSH/RPUSH/LPOP`), and append-only streams (`XADD/XRANGE/XREAD`) with millisecond-precision auto-generated IDs

### Real Estate Price Monitoring Lakehouse Pipeline

- Built an end-to-end real estate data pipeline to ingest listings and track historical price changes
- Implemented an SCD Type 2 model to maintain historical price records
- Optimized incremental updates using Delta Lake transaction logs.
- Automated the pipeline using Airflow and implemented a Bronze-Silver-Gold lakehouse architecture

## TECHNICAL SKILLS

---

- Programming Languages:** Python, Java, SQL, Shell Script
- Big Data Technologies:** Apache Hadoop, PySpark Structured Streaming, Apache Kafka, Apache Hive, Airflow (ETL)
- Backend:** FastAPI, REST APIs, Pydantic, MongoDB
- Version Control:** Git, GitHub
- Cloud Platforms:** Snowflake, Azure Synapse, Azure Data Lake, Azure IOT Hub, Docker, BigQuery
- Remote Access:** PUTTY, WinSCP, SSH
- AI Orchestration Frameworks:** LangChain (agents, chains, retrieval-augmented generation), LangFlow (low-code LLM workflows), Model Context Protocol(MCP))
- Data Modeling:** Data Normalization, Slowly changing Dimension, Star Schema, Snowflake schema

## INTERNSHIP

---

### AICL | *Full Stack Web Development*

Oct 2024 - Nov 2024

- Led integration of the MySQL database with the application using SQLAlchemy and SSH tunnels, improving data reliability and reducing sync errors by 30%.
- Developed and optimized RESTful APIs with Python Flask to enable reliable data exchange between the frontend and PostgreSQL, improving system performance and reducing response times by 30%.
- Utilizing Git for Version control and Debugging code to reduce deployment Errors

## Certifications

---

- Udemy-Hands on Practice With Apache Airflow
- Nasscom-Data Science For Beginners
- Oracle-Cloud Infrastructure and AI Foundations
- Udemy-Data Engineering Fundamentals