

KISHORE MS

MULTI CLOUD DATA ENGINEER - BUSINESS ANALYST
& DATA SCIENTIST

CONTACT_



+91 95005 66969



inbox.kishorems@gmail.com



[linkedin.com/in/kishore-m-s-ba56b4245](https://www.linkedin.com/in/kishore-m-s-ba56b4245)

EDUCATION_

Post Graduate Diploma in Management (PDGM) - Data Science and Analytics (Summa Cum Laude - Chairman's Merit List)
Great Lakes Institute of Management, Gurgaon
2022-2024

Bachelor of Commerce - Corporate Secretaryship
DG Vaishnav College, Chennai
2019-2022

TECH AND CLOUD SKILLS_

Cloud & Data Platforms:

- AWS - S3, Glue, Athena, Lambda, EC2, IAM, CloudWatch, Redshift, DMS, SNS, SQS, VPC, RDS, CloudFront, MSK, MWAA, Step Functions.

- AZURE- Data Factory, Azure Data Lake Storage Gen2 (ADLS Gen2), Databricks, Delta Lake, Unity Catalog, Azure SQL DB, Azure DevOps, Azure Synapse.

Data Engineering & Processing:

- Pyspark - RDDs, DataFrames, Advanced transformations.
- ETL/ELT Pipelines - Batch & Streaming
- Data Modeling - Dimensional, Star & Snowflake Schema
- Apache Kafka - Producer/Consumer, Topics, Stream Processing

Databases & Querying:

- SQL - Joins, Window functions, Optimization

Programming & Markup:

- Python, XML, HTML, CSS

Visualization & Analytics:

- Power BI, Kibana, Amazon QuickSight
- Exploratory Data Analysis (EDA) with Pandas, Numpy, Reporting & Dashboarding

Collaboration & Productivity:

- Office 365
- Jira, Confluence

Version Control & DevOps:

- Git

ABOUT_

Multi-Cloud Data Engineer & Business Analyst with 3.5 years of experience in AWS and Azure. Specialized in designing scalable, governed data ecosystems using Medallion, Lakehouse, CDC, and SCD. Proven ability in stakeholder engagement and requirements analysis to translate complex data into actionable insights that drive strategic decisions and BI reporting. Experienced in developing end-to-end ETL pipelines and coordinating with cross-functional teams to deliver timely, high-quality analytical solutions.

WORK EXPERIENCE_

Data Engineer - Tech Ops Associate (Catalog Lead)

AMAZON
2023 - Present

- Engineered and optimized multiple marketplace ETL pipelines using PySpark on the Lakehouse architecture and Delta Lake to transform competitor data into analytics-ready OLAP datasets.

- Managed product refresh cycles using Slowly Changing Dimension Type 2 (SCD2) for incremental selection updates, ensuring only newly surfaced selections are pushed to the persistent layer. This guarantees a consistent end-to-end design from OLTP→ETL→OLAP.

- Executed cross AWS data migration, ensuring validation without data loss using AWS DMS snapshots and improving global accessibility with AWS CloudFront caching. Collaborated with cross-functional teams using Agile practices and conducted quality audits/sprints to guarantee data integrity before OLAP integration.

- Built high-volume ETL workflows using the Medallion Architecture (Bronze, Silver, Gold) to ingest data from Nielsen, GfK, and Google Best Sellers. Consumed raw data via Kafka into the Bronze layer and performed automated KPI and reconciliation checks using AWS Lambda.

- Orchestrated complex ETL workflows using Airflow (MWAA) and Step Functions to automate pipelines. Applied Dimensional Modeling and Star Schema design in Redshift to optimize OLAP performance, while triggering SNS notifications to stakeholders upon successful ingestion and validation.

- Led a team of 8 L2 analysts across marketplaces, managing web scraping ETL configurations using advanced Git (stashing, cherry-picking) to ensure validated data delivery to the persistent OLTP database (source of truth).

- Drove significant Strategic Impact by mapping 850+ new brands (+15% diversity), which improved availability & satisfaction by 10%. Simultaneously reduced data errors by 18% and boosted team productivity by 25% ensuring high-quality pipelines for the end stakeholder consumption.

Data Analyst

NielsenIQ
2022 - 2023

- Collaborated with onshore sales teams and business stakeholders to gather requirements. Implemented data cleaning & normalization methods to process and transform large, denormalized datasets, supporting downstream ETL pipelines and OLAP reporting.

FUNCTIONAL SKILLS_

- Stakeholder Requirement Analysis & Engagement
- Requirements Elicitation & Documentation
- Cross-Functional Stakeholder Alignment
- Agile Project Management
- Data Governance & Compliance
- Strategic Reporting & Insights
- Master Data Management (MDM) Principles
- Cost Optimization (Cloud Resources)
- Data Storytelling
- Source-to-Target (S2T) Mapping
- Root Cause Analysis (RCA)
- Service Level Agreement (SLA) Management
- Disaster Recovery (DR) Planning

ACADEMIC PORTFOLIO PROJECTS_

(Click on any project to visit e-portfolio)

[Harnessing AI for Industry Transformation: A Case Study Approach](#)

[Predicting Salary and Location Preferences of Job Seekers in India](#)

[Developing & Implementing Digital Presence & Data-Driven Marketing Framework](#)

[Analysed datasets and made inferences by using economic model, logistic equation & cluster analysis](#)

[Conducting Data Analysis and Constructing Linear Regression Models with Interpretative Insights](#)

[Comprehensive Financial Analysis & Investment Decision: Evaluating Company Performance, Ratios, & Market Data](#)

Applied industry-standard data normalization techniques (1 - 4 NF) to streamline the data preparation process, significantly reducing processing errors and supporting downstream ETL pipelines.

Led the end-to-end data auditing process, implementing data deduplication and validation techniques to resolve anomalies & inconsistencies, boosting the reliability of stakeholder insights.

Utilized advanced analytics and predictive modeling to mine large datasets, generating forecasts and trends that enabled top-tier clients (e.g., Amazon, Walmart, Target, P&G) to make data-driven decisions.

Developed intuitive, visually appealing Power BI dashboards to present complex datasets, enhancing client understanding of market patterns and informing strategic product and marketing decisions.

HIGH IMPACT PROJECT PORTFOLIO_

(Available on [GitHub](#) - Click any project title to view the full repository with files and README screenshots)

Project 1 - Next Gen Azure Data Engineering Pipeline with CI/CD:

Architected a scalable Medallion pipeline leveraging Azure Data Factory, ADLS Gen2, and Databricks. Integrated Unity Catalog for enterprise governance and security, and deployed CI/CD workflows using Azure DevOps and Git branching strategies for automated releases and analytics-ready dimensional models.

Project 2 - Enterprise Azure Data Lakehouse with CDC, Watermarking & OLAP Star Schema:

Engineered a production-grade Azure Data Lakehouse supporting CDC, watermarking, and incremental loads. Implemented idempotent procedures, surrogate key, SCDs for OLAP-ready dimensional modeling, and adopted Delta Lake for ACID transactions and time-travel. Deployed CI/CD pipelines with Azure DevOps to deliver an optimized enterprise OLAP schema.

Project 3 - Automated API-Driven Data Pipeline :

Developed an automated end-to-end Bronze → Silver → Gold pipeline consuming REST APIs via Azure Data Factory. Applied PySpark transformations on Databricks, ensuring secure access and secret management via Service Principal authentication and Key Vault.

Project 4 - AWS Real-Time Financial Transactions Pipeline :

Built a low-latency, real-time ingestion and processing pipeline using Kafka, AWS DMS, Glue (PySpark), and RDS following a Medallion-like architecture. Implemented CDC replication and incremental processing, integrating AWS Lambda for automated KPI checks and SNS notifications for governance and real-time alerts.

Project 5 - Automated Financial Transactions ETL Pipeline:

Automated incremental data processing (Bronze → Silver → Gold) using AWS S3, Lambda, Step Functions, and Glue (PySpark). Ensured data quality, KPI validation, and reconciliation with DMS CDC, orchestrating the end-to-end workflow for dimensional modeling and aggregations in Gold.

AWARDS & HONORS_

- Chairman's Merit List for Academic Excellence Great Lakes Institute of Management.
- Best in Firing Award and Medal Holder Ministry of Defence, Govt. of India.
- Distinctive Graded Music Performer Trinity College of London.
- Rising Star Amazon.
- Pure Performer Amazon.
- NOC Star NielsenIQ.