



Kirti Singh

DATA ENGINEER

CONTACT

- 91 8240283959
- singhkirti.mumbai@gmail.com
- Kolkata, India

SKILL MATRIX

Domain	Finance, Supply Chain, Retail
Cloud ETL Tools	Azure Data Factory, Azure Data Flows, Databricks, AWS Glue (Exploring)
Other ETL Tools	IBM Datastage Designer
Programming	Python, SQL, Unix Shell Scripting
Cloud Technologies	Azure, Snowflake(exploring), AWS (exploring)
Big Data	Apache Spark
Datawarehouse	Azure Synapse Analytics, Snowflake, Redshift (Exploring) Teradata
Databases	Azure SQL DB, SQL Server, Oracle
Visualization	Power BI(exploring)

EDUCATION

Bachelor of Technology in Electronics & Communication, West Bengal
University of Technology, India
2013-2017

HOBBIES

- Reading Books (*Latest book read: Attitude is Everything by Jeff Keller, Fav Quote: If it takes 5 steps to reach my goal, I will take those 5 steps*)
- Cooking various types of Indian cuisine

SUMMARY

Data Engineer with more than 5 years of experience in IT. Involved in end-to-end ETL projects by integrating data into Azure SQL Datawarehouse using Data Factory, developing Databricks notebook for big data processing, and developing reports and dashboards using Power BI.

EXPERIENCE

Consultant, Capgemini Technology Services India Ltd.

May 2021– Present

Finance & Reporting [Tools used: Data Factory, Data Lake Gen2, Databricks]

- Responsible for creating data pipelines using **Azure Data Factory** inside Azure Synapse Analytics
- Transformed data as per business rules such as deriving new columns, joins, aggregations using **Azure Databricks** and store in Curated zone.
- Created high quality summarized data inside Analytics zone so that it can be consumed by further downstream processes for model training and prediction.
- Used Azure Data Factory to orchestrate and schedule the pipeline.

Migration Project [Tools used: SQL Server, Data Factory, Data Flow, Data Lake Gen2]

- Created pipelines in Azure Data Factory to extract data from SQL Server and load into Azure Data Lake.
- Used ADF and Azure Notification services to ingest flat file data from data lake into Snowflake in near real time.
- Transformed data as per business rules such as deriving new columns, joins, aggregations using **Azure Data Flow** and store in Curated zone

Sr R&D Engineer, Hind Rectifiers Ltd.

July 2017 – April 2021

Optimizing Supply Chain [Tools used: IBM Datastage, Teradata, SQL Server]

- Responsible for loading of Supply Chain data, purchasing & suppliers, sales, accounts, human resources, CRM, Servicing, Logistics & GST related data from various sources into the database.
- Using **IBM Infosphere Datastage 9.1** as an ETL tool to extract data from sources systems & load the data into Teradata DB.
- Responsible for creating DataStage parallel jobs using different stages like the transformer, aggregator, sort, join, merge, lookup.
- Responsible for unit testing of ETL jobs on Development environment.
- Active participation in weekly meetings to review status.
- Interacted with Business Partners to understand the business requirements, Analyse and design the integration Approach.

Dispatch and Logistics [Tools used: IBM Datastage, Teradata, SQL Server]

- Performed ETL DataStage coding using Datasets, Sequential file, Transformer, Sort, Merge, Aggregator stages compiled, debugged and tested. Extensively used stages available to redesign DataStage jobs for performing the required integration.
- Used DataStage Designer tools for creating Parallel jobs and importing and exporting the jobs.
- Used DataStage Job Sequences to execute the jobs in a sequence order.
- Created DataStage Jobs to build business rules to load data.
- Used DataStage Director to run and monitor status of the jobs.
- Imported table/file definitions into the Data stage repository.
- Controlled jobs execution using sequencer, used notification activity to send email alerts.