Shubham Mishra

Data Engineer

✉ shubham.mishra@gmail.com 📞 9876543210 in LinkedIn ⊙ Github 📍 Bengalore Karnataka

---

Profile

3+ years of experience in various projects related to Data Modeling, Data Analysis, Design, and Development for Data warehousing environments. Practical understanding of Data modeling concepts like Start-Schema Modeling. Snowflake Schema Modeling, Fact and Dimension tables. Comprehensive knowledge and experience in process improvement, normalization/de-normalization, data extraction, data cleansing, and data manipulation. Exceptional troubleshooting skills with ETL technologies.

Professional Experience

05/2020 – present
Bangalore, India

Data Engineer
Wipro

Salesforce to Redshift Ingestion - Migration from Informatica to Native AWS

Tech Stack – Salesforce, Informatica, S3, Lambda, Glue, AppFlow, Redshift, SNS

- Crafted generic scalable Native AWS solution for Salesforce to Redshift ingestion
- It helped to move ingestion pipelines from third-party tool Informatica and saved cost for the heavy license fee
- This generic framework helped other business units for smooth ingestion of newly onboarded Salesforce objects into a Redshift Data lake

Dehleez - Report Scheduling Tool

Tech Stack – Python, JavaScript, Django, Azkaban, Docker, Hive, Ajax, Bootstrap, REST API, DataDog

- Enhanced Paytm's proprietary report scheduling tool which is used by business users working on data analysis where they can schedule their reports by writing HIVE/MySQL/Cassandra queries and report output in various formats
- Diff Checker - Admins can check the difference between queries before approving reports
- Time slot picker to schedule a report - The user can see scheduled reports for the next 4 hours from the intended schedule time and can pick the slot accordingly
- Dump report output into the S3 bucket - The user can take a dump of report output into the AWS S3 bucket
- Cassandra Connector - User can schedule reports having Cassandra query panels in addition to HIVE/MySQL

| 01/2019 – 05/2020 | **Data Engineer** |
| Mumbai, India | *Capgemini* |

Länsförsäkringar, or literally County Insurance, is a Swedish group of customer-owned insurance companies. The group consists of 23 independent companies, one in each of the counties of Sweden (pre-1998 division), that cooperate under a common brand name.

Role: SQL and SSIS Developer & SQL Support
- Troubleshoot live bugs and resolved them on production.
- To fulfill the requirement of altering existing objects.
- Performance tuning using system views & Execution Plan.
- Highly involved in development activities as per requirement from client
- Taking ownership of technical issues, and Handling issues
- Proper documentation on development and Issues handled providing the Root cause of the problem.

Skills

Programing Language:
Python, Java Scala, C++, and Linux Scripting

Web Development Framework
Django, Flask

Databases
MySQL, Elastic Search, Cassandra, HBase, Hive, SQL Server

Distributed Framework
Hadoop, Spark, Kafka, BigQuery

Dashboarding & Monitoring Tools
Grafana, Kibana, Power BI

Workflow Management
Airflow, IBM Tivoli

AWS Services
S3 bucket, EMR, Lambda, Cloudwatch, DynamoDB, Redshift, MWAA, Glue, SNS, SQS

Build Tools & Containers
Docker, Kubernetes

Code Management and Documentation
Git, BitBucket, Azure Develops, Circle CI, Travis CI

Education

| 06/2015 – 06/2018 | **Master of Computer Applications** |
| Prayagraj, India | *National Institute of Technology Allahabad (MNNIT), U.P* |

| 06/2012 – 06/2015 | **Bachelor in Computer Science** |
| Mumbai, India | *Mumbai University* |

Certificates

- AWS certified Database specialty
- AWS certified Solution Architect
- Azure Data Scientist
- Azure Fundamental