

Niharika Tellakula

Georgia, USA | +1 737 377 8191 | ntellakula99@gmail.com

Summary

- Having 2+ years of experience in handling Data Warehousing and Data Engineering projects in Banking, Finance, and Retail Industry.
- Evaluating technology stack for building Analytics solutions on cloud by doing research and finding right strategies, tools for building end to end analytics solutions and help designing technology roadmap for Data Ingestion, Data lakes, Data processing and Visualization.
- Good knowledge on Hadoop Architecture and its ecosystem.
- Experience in handling different file formats (**AVRO, ORC and PARQUET**) in Spark.
- Hands-on experience with AWS S3, IAM, Lambda, EMR, Glue Analytics and ATHENA.
- Have good experience in technical designs that will meet system objectives and minimize the impact on operations.
- Experience on migrating on Premises ETL process to Cloud.
- Solid experience in **SDLC, Debugging** and **Documentation**.
- Strong **analytical, problem solving & organizational abilities**.
- Experience in optimizing Hive SQL queries and Spark Jobs.
- Experience with creation of technical document for Functional Requirement, Impact Analysis, Technical Design documents, Data Flow Diagram with MS Visio.
- Having experience in delivering the highly complex project with Agile and Scrum methodology.
- Quick learner and up-to-date with industry trends, Excellent written and oral communications, analytical and problem-solving skills and good team player, Ability to work independently and well-organized.

Experience

DATA ENGINEER | 08/2021 – 07/2022

Project: AIM (Analytical Information Management) FCA (Regulatory Body of Bank of England, UK)

Domain: Banking

- Performed import of data from RDBMS to Amazon S3 using Talend ETL and then importing to HDFS to process the huge amounts of Data.
- Worked on Data Lake and Staging Area where data will be staged into AVRO file formats.
- Data modeling is done using the Data Vault technique.
- Design, development, testing, troubleshooting and debugging of the application.
- Assisting in integration testing, system testing, User Acceptance Test & implementation activities.
- Implementing project plans within deadlines.
- Involved in working on the Data Analysis, Data quality and Data Profiling for handling the business that helped the business team.
- JSON Parsing integration
- Worked on the EDH layer (Enterprise Data Hub) for doing hive analytics.
- SFTP (Secured File Transfer Protocol)
- Developed the modules and changes based on Client's requirements and production support.

Technologies: Python, PySpark, Hadoop, Hive, Shell Scripting, SQL, HBASE, AWS S3, AWS EMR, AWS Glue Analytics, TALEND (Injection Tool), AWS IAM, AWS LAMBDA.

- Worked on Sales data which consists of data related to laptop sales across multiple locations in India and USA.
- Importing huge amounts of retail data into HDFS from various sources and processing it CLOUDERA.
- Performed **Import and Export** of data into **HDFS and Hive** using **Sqoop** and managed data within the environment.
- Involved in **creating Hive tables, data loading** and writing **hive queries**.
- Was responsible for **Optimizing Hive** queries that helped in saving Cost to the project.
- **Managed Hive Tables** and created child tables based on partitions.
- Used spark transformations while moving the data to **STAGING layer**.
- Worked on different file formats like Text file, Sequence file, Avro, Json, Parquet, Orc, Csv, Tsv, Xml, and Custom delimited file formats.
- Involved in working on the Data Analysis, Data Quality and data profiling for handling the business that helped the Business team.
- Hands on experience with Hadoop Ecosystem components like **MapReduce (Processing)**, **HDFS (Storage)**, **YARN**, **Sqoop**, **Hive**, **Hbase** for data storage and analysis.

Technologies: Python, PySpark, Hadoop, Hive, Shell Scripting, SQL, HBASE, AWS S3, AWS EMR, AWS ATHENA, CLOUDERA SQOOP.

Skills

- **Programming Languages:** Python, Java
- **Database Management:** MySQL, Oracle
- **No SQL Databases:** HBase
- **Hadoop ecosystem:** Hadoop, Sqoop, Hive, HDFS, Spark-SQL
- **AWS:** S3, EMR, Lambda, Glue Analytics, ATHENA, IAM
- **Microsoft Office Tools:** MS Word, MS Excel, MS PowerPoint, MS Outlook
- **IDE & Tools:** Git, Maven, JIRA, Jenkins, Oozie, Pycharm

Education

PITTSBURG STATE UNIVERSITY – Pittsburg, Kansas | Master of Information Technology – Jan 2023 – May 2024

Shree Uma Degree College, (Affiliated to OSMANIA UNIVERSITY, HYDERABAD) | Bachelor of Computers and Commerce – June 2017 – May 2020