

SUDARSHAN JADHAV

Email: jsudarshan97@yahoo.com

Phone No: 8830380985

Address: Maliwada, Aurangabad 431002



Objective

I am seeking a challenging position to utilize my skills and abilities in an industry that offers professional growth while being resourceful, innovative, and flexible.

Professional Summary

- Having a total of 3 years and 6 months of experience in the field of Big Data, currently working as a **Data Engineer** with **MclanSys Solutions**.
- Expertise in Designing and Architecting Hadoop Applications and recommending the right solutions and technologies for the applications.
- Having good expertise in Hadoop tools like HDFS, Yarn, Hive, Sqoop, impala, and Spark.
- Experience in Apache Oozie Job Scheduling.
- Hands-on experience in various big data application phases like data ingestion, processing, and analytics.
- Expertise in Designing and Architecting Hadoop Applications and recommending the right solutions and technologies for the applications.
- Knowing of transferring data from RDBMS to HDFS and Hive table using Sqoop.
- Data Validation after Data Ingestion & Quality Checks.
- Writing DDL to create a hive table for optimized performance.
- Very Good understanding of Partitioning, and bucketing concepts in Hive and designed both Internal and External tables in Hive to optimize performance.
- Experience in basic Linux Scripting.
- Effectively made use of Table Functions, Indexes, Table Partitioning Collections, Views, and Query rewrite for effective usage.

Technical Skills

Big Data Ecosystem:	HDFS, YARN, Hive, Sqoop, Spark-Core, and Spark-SQL.
Distribution Package:	Cloudera CDH (5.13.1)
Language:	Python-basic
Management Tool:	Zookeeper.
Database:	MySQL
Scripting Language:	Shell Scripting
Cloud:	AWS (RDS, S3, EC2, IAM, EMR, Cloud Watch)
Platforms:	Windows, Linux

Organization Scan

- | | |
|-----------------------|-----------------------------|
| ➤ Organization | - MclanSys Solutions |
| ➤ Designation | - Data Engineer |
| ➤ Duration | - May-2019-Present |

Work Experience

Project: Health Insurance Data Analysis

Duration: Dec 2020- Present

The purpose of the Project is to ingest the Health Insurance data that is present on cloud storage to a Data Hub modeled as per an LDM. It's Cloud to hive ingestion Project in Real Time. Data Migration Utility to migrate existing system into Data hub which was built on top of Hadoop.

Roles and Responsibilities:

- Import Data into the Hive from various Relational Databases (Oracle, MySQL) using Sqoop.
- Write Hive DDL to Create Hive Table for Optimize Query Performance.
- Engage with BA to understand the requirements clearly.
- Define all the possible Test Cases along with the Test Data.
- Comes up with missing scenarios and gets them clarified with BA.
- Implement Data Validation, Quality Checks.
- Perform UAT of Big Data implementation.
- Involved in Collecting Business Requirements from Business Users, and translating them into Technical Design (Data Pipelines and ETL workflows).
- Involved in importing data from various RDBMS into HDFS using Sqoop and Pyspark, including Incremental Load to populate Hive External Table and vice-versa.
- Designed both Managed and External Hive Tables and Defined static and dynamic partitions as per requirement for optimized performance on production datasets.
- Worked with various File Formats like Text Files, Sequence Files, ORC Files, and Avro Files.
- Written Hive Queries for Data Analysis to meet the business requirements.

Environment: Hadoop, Hive, PySpark, AWS

Project #2: Enterprise Data warehouse Analysis:

Duration: May 2019-Dec 2020

Roles & Responsibilities:

- Understand the client's requirements for developing the project.
- Connect That Created MySql instance using workbench.
- Importing all data and tables into RDS using MySql DB.
- Load the in Hive for further process by using Sqoop service.
- Then create Sqoop jobs for importing and exporting data.
- Load the data from RDBMS to HDFS for processing.
- Extract the data into S3 Bucket in CSV File format.
- Perform optimization techniques to enhance the performance the of Hive job.
- Responsible for scheduling and monitoring the job.
- Raising the incident to a different team to support the project.
- Attend daily meetings and calls for requirements.
- Create the report using Tableau.

Environment: HDFS, Apache Sqoop, MySQL, Hive, AWS

Educational Qualification

- Bachelor of Engineering from PES College of Engineering, Aurangabad.
- SSC from Deogiri Vidyalay Daulatabad.

Personal Information

Date of Birth : 25th May 1997
Languages Known: English, Hindi, and Marathi
Marital Status : Single

I hereby declare that all the information mentioned above is true and to the best of my knowledge.

Date:

Place:

S.A.Jadhav