

**Mahesh Babu**  
**Senior Data Engineer**

Email: [Ko.mahesh111@gmail.com](mailto:Ko.mahesh111@gmail.com)

Mobile: 7337080825

#### Experience Summary

- A dynamic professional with 4+ years of experience in Software Development and implementation in **Big Data, Hadoop** and **Spark** related technologies
- Having experience in Hadoop and its components like **HDFS, Pig, Hive, Sqoop and Map Reduce**.
- Working experience on implementation of **Spark core, Spark SQL** and Spark Streaming
- Involved in Design and Development of technical specifications using Hadoop-Eco System tools.
- Having good exposure in **Python** programming.
- Having Experience on Amazon web services (AWS) like **S3, Athena** and **EMR**
- Implemented Spark using **Scala** and Spark SQL for faster processing of data
- Having Experience on DeltaLake
- Written complex queries with the usage of Hive Analytical functions
- Hands on experience in writing hive scripts using HiveQL
- Experience in writing Pig scripts using Pig Latin Language to process data
- Hands on experience in integrating Hive with Spark to perform HQL in spark.
- Hands on experience No-SQL database like **vertica** and Druid.
- Experience in handling different file formats like JSON, Parquet, Avro, Text files, XML and Flat file format
- Experience in Relational Databases like MySQL, Oracle.

#### Skills Summary

Domain	Retail Enterprise Analysis
Programming Languages	Python, and SQL
Operating System / ERP Version	Windows 7
Tools / DB / Packages / Framework / ERP Components	Eclipse, Jira, GlitHub
Hardware Platforms	Intel Series

#### Professional Experience:

- Worked as a Senior Data Engineer in LTIMINDTREE from Oct 2021 to Dec 2022.
- Worked as a Data Engineer in Lemnisk from Oct 2018 to Oct 2021.

#### Project 2

Project Name	HSBC Bank	Team Size	12
Start Date	Nov 2021	End Date	Dec 2022

Project Description	In HSBC worked for the credit card monitoring team in which it will deal with the information of card transactions. For the region like UK, USA, ASP (Asia Pacific Countries).
Role & Contribution	<ul style="list-style-type: none"> <li>Developed end to end data pipeline using python spark and hive to easily flow of data from source to target layer</li> <li>Monitored the daily running pyspark jobs in Control-M.</li> <li>Will analyse and debug the log to find out error.</li> <li><b>Will report to the respective team regarding the failure of jobs</b></li> <li>Involved in data ingestion from different sources and applied transformations</li> <li>Involved in writing Test cases to test the pipelines</li> <li>Worked with spark for computations and processing</li> </ul>
Technology & Tools	SQL, HIVE, PYTHON

Work Experience			
Project 1			
Project Name	Retail Enterprise Analysis	Team Size	5
Start Date	Oct 2018	End Date	Oct 2021
Project Description	<p>The purpose of the project is to store terabytes of different departmental information generated by Safeway Market Intelligence team and extract meaning information out of it. The solution is based on the open source Big Data s/w Hadoop .The data will be stored in S3 and processed using Hive jobs. Extract various reports out of the processed information and Export the information for further processing. This project is mainly for the re-designing of the current existing system which is running in the new cloud solution technology called Hadoop which can able to process large data sets in order to meet the client requirements.</p>		
Role & Contribution	<ul style="list-style-type: none"> <li>Involved in developing the Reports using Hive</li> <li>Created Hive tables to store the processed results in a tabular format</li> </ul>		

	<ul style="list-style-type: none"><li>Created Hive Managed and External tables with partition and bucket in Hive and loaded data in to Hives.</li><li>Developed data queries using HiveQL and optimized the Hive queries.</li><li>Created Vertica tables to load large volume of data.</li><li>Worked on import and export data from MySQL and Oracle into HIVE using Sqoop.</li><li>Performed data ingestion, entity resolution and ran ad-hoc queries using HDFS, Hive and Pig.</li><li>Fetching reports using Hive depends on client requirement.</li><li>Developed DF's, Case classes for required input data and performed the data transformations using spark core.</li><li>Implemented Spark RDD Transformations and Actions.</li><li>Used Hive Queries in Spark-SQL for analysis and processing the data</li></ul>		
Technology & Tools	Eclipse, Jira, GlitHub		
Other Experience			
Title	Universal softech hub pvt ltd		
Start Date	Oct 2018	End Date	Oct 2021
Role & Contribution	<ul style="list-style-type: none"><li>Involved in developing the Reports using Hive</li><li>Created Hive tables to store the processed results in a tabular format</li><li>Created Hive Managed and External tables with partition and bucket in Hive and loaded data in to Hives.</li><li>Developed data queries using HiveQL and optimized the Hive queries.</li><li>Created Vertica tables to load large volume of data.</li><li>Worked on import and export data from MySQL and Oracle into HIVE using Sqoop.</li><li>Performed data ingestion, entity resolution and ran ad-hoc queries using HDFS, Hive and Pig.</li><li>Fetching reports using Hive depends on client requirement.</li></ul>		

	<ul style="list-style-type: none"> <li>▪ Developed DF's, Case classes for required input data and performed the data transformations using spark core.</li> <li>▪ Implemented Spark RDD Transformations and Actions.</li> <li>▪ Used Hive Queries in Spark-SQL for analysis and processing the data</li> </ul>
Educational Qualification	
Education & Credentials	B.E Mech