

Shahanaj .M

Q CONNEQT Business Solutions LTD

mullashainu786@gmail.com

Ph no: 9380588235

ExperienceSummary

Overall,3.5 years of experience in IT and strong technical skills in SQL, Python, Hadoop, Spark, PySpark (Around using Big Data & Hadoop development using HDFS, Hive, MapReduce, Sqoop, HBase etc.) and Lambda, AWS Athena, Glue, EMR, EC2, Databricks. including in designing, developing applications using Hadoop Ecosystem components.

- Relevant, 3.0 years of experience as a data engineer with strong emphasis on Design, Development, Implementation, Deployment of applications.
- Good experience in Hadoop ecosystem& Big-Data Analytics.
- Experienced in Spark ecosystem components like Spark batch, Spark streaming, Spark SQL, Spark-core. Good experience in Spark, PySpark, AWS, data bricks
- Experienced in using HDFS, MapReduce, Hive, Pig, Sqoop, Cloudera Manager. Good knowledge of Hadoop Architecture and Yarn Architecture.
- Good knowledge of Spark Architecture and AWS.
- Good knowledge and Understanding of Spark-standalone-Architecture and Spark-Yarn Architecture.
- Experienced in working Hive data warehouse system, developing data pipeline, implementing complex business logic and optimizing business logic.
- Good knowledge of AWS, S3, EMR.
- Experienced Hadoop Developer, have a strong background with file distribution systems in a bigdata. Understands the complex processing needs of big data and have experience in developing codes and modules to address those needs.
- Knowledge of Map Reduce programs to perform analysis and MR unit jobs for testing
- Excellent communication, organizational skills, attention to detail & the ability to meet tight deadlines.

Skills Summary

Domain	Telecom,Retail,Banking
Programming Languages	Python
Operating System / ERPVersion	Windows10,Ubuntu(Linux)
Tools / DB /Packages /Framework / ERPComponents	Spark, PySpark, Python, Hive, HDFS,Sqoop, MapReduce, Linux, HBase, Pig, AWS (Athena, Glue,Airflow, DynamoDB)

Professional summary	Worked in HGS from Nov' 2018 to Mar'2020. Worked in Haya systems India pvt.ltd from Apr'2020 to June 2021. Worked in QConneqt business solutions ltd from Sep'2021 to Aug'2022.
-----------------------------	---

WorkExperience			
Project2			
ProjectName	P & G	TeamSize	12
StartDate	Sep2021	End Date	Aug'2022
ProjectDescription	<p>P&G Business such as P&G Products, enterprise customers and a huge global Products, data analytics is a critical component of P&G business. P&G collects huge amount of usage information about all their products, services and sites such as call records, wireless location, network traffic data, product and device specific information and other similar data and that will be consumed for Business Analytics and build Big Data Solutions for various P&G Business units. The collected data is aggregated and analyses for adhere legal compliance, business and marketing to we improve the services, measure and analyze the usage and to help to find services and advertising more relevant to the customers</p>		
Role &Contribution	<ul style="list-style-type: none"> ▪ Data Reading from Mount Path to Spark for Processing data ▪ Understand business story and analyses functional specifications& Design feasible solutions. ▪ Based on business logic at applying cleaning the data, applying transformations and analytics with Spark. ▪ Developed code to write the logic data into Athena External. Involved in configuring and creating Airflow schedule. ▪ Involved in setting up the properties for Apache Airflow. ▪ Involved in various log files to fix the issues during unit testing and performance tuning& Involved in developing the spark. ▪ Tuning &Optimizing Involved in developing the PySpark 		
Technology &Tools	PySpark, Python, Spark, Hadoop Hive,AWS, S3, Athena, Glue, lambda, Data Bricks		

WorkExperience			
Project2			
ProjectName	Toronto-Dominion bank	TeamSize	8
StartDate	Nov 2018	End Date	June2021
ProjectDescription	<p>RRTS Business like RRTS (ora-std) MTG, PCL, CRD, customers and a finance, data analytics is a critical component of Toronto-Dominion bank-business. It collects huge amount of usage information about all their products, services and sites such as Complete Cards Business, product and device specific information and other similar data and that will be consumed for Business Analytics and build Big Data Solutions for various Toronto Dominion Business units. The collected data is aggregated and analyzed for adhere legal compliance, business and marketing to improve the services, measure and analyze the usage and to help to find services and advertising more relevant to the customers.</p>		
Role &Contribution	<ul style="list-style-type: none"> ▪ Injecting data From Oracle to Hive Tables Using PySpark. ▪ Loading data from Hive tables to PySpark processing framework. ▪ Understand business story and analyze functional specifications & design feasible solutions. ▪ Based on business logic at applying cleaning the data, applying transformations and analytics with PySpark. Involved in configuring and creating crown tab schedule. ▪ Developed code to write the logic data into Hive External. ▪ Involved in various log files to fix the issues during unit testing and performance tuning Involved in developing the spark.. 		
Technology &Tools	Spark, PySpark, Python, Hive, HDFS, Sqoop, MapReduce, SQL, Tableau, Jenkins, Git bash, Bit bucket, Crown tab, ArcadiaData, JMS, Shell Scripting, Athena, S3, Glue, Airflow, Linux.		

EducationalQualification	
Education summary	Diploma Computer Science in 2018