

PROFILE

I am a Software developer with 2+ years of experience working in Big data projects. I have been experience working on Big data Technologies, Hadoop Architecture, HDFS, Data Engineering, PYSPARK, Impala, Hive, Hue, Python, SQOOP, AWS S3, Azure Storage Explorer, Tableau, SQL Developer 19C & Unix Shell Scripting.

EXPERIENCE SUMMARY

- 2+ years of working experience at IQVIA.
- Experienced in Big data, Hadoop, HDFS, Hive, PYSPARK, AWS S3 and Azure Storage technologies.
- Experience in Big Data projects.
- Experience in Sqooing data from Database to Hadoop environment.
- Experience in writing the Unix Shell Scripting.
- Experienced in working with Database Schema, Data Architecture, Database design and development, Data marts.
- Experienced developing PL/SQL scripts for Data Extraction from Datawarehouse.
- Experienced in working with Analyzed and designed data flow diagram for Onboarding the supplier to system and aggregated data to XPO.
- Experienced in Design and development of application and tool solutions to identify and address the business issues.
- Experienced in working in a complete project cycle like Development, Testing and Production support in project.

Directly involved in developing PYSPARK Scripts, Hive Query, Cloudera Data Science Workbench Projects developments, PL/SQL scripts and solving tickets of all priority.

PROFESSIONAL EXPERIENCE

IQVIA

Project –

RWD (Real World Data)
(June 2019 - Present)

Project Description –

RWD works for diseases pharmacy to provide insights to Big data.

Key Responsibility –

- Developing the project as per the business requirement in CDSW (Cloudera Data Science Workbench).
- Developing PYSPARK scripts in CDSW Workbench for Data Extraction.
- Developing Impala and Hive scripts for data Extraction from Hue.
- Developing Unix shell scripts for Moving files from Cloud Storage to Linux servers and to various cloud locations.
- Handling Git repository for scripts updates.
- Handling non-production and production defects, analyzing, and providing solutions for it.

TRAINING EXPERIENCE

HPES

Project –

Airlines data Analytics
(May 2018 – June 2018)

Project Description –

This project was based on Big Data technologies for developing clusters Hadoop Distributed File System (HDFS) and faster processing of Airlines data through MapReduce framework. Hive Query language was used for analyzing the huge volume of data.

Key Responsibility –

- Set up the projects through CDSW workbench
- Developing PYSPARK scripts for creating data frames and extract the data as per client requirements.
- Developing Unix Shell Scripts.
- Developing the Hive Scripts for QUERY the data from database.
- Involved in all these big data technologies like Apache Spark, SQOOP etc.

Tools and Technologies -

- Cloudera Data Science Workbench,
- HUE,
- Cloudberry,
- Microsoft Azure storage Explorer,
- Tableau,
- Windows PowerShell,
- Toad for Oracle, Oracle client 19c,
- SQL Developer,
- Putty,
- WinSCP/FileZilla

Programming Language –

- C
- C++
- JAVA
- PYTHON

EDUCATION

BACHELOR OF TECHNOLOGY IN INFORMATION TECHNOLOGY,

COCHIN UNIVERSITY OF SCIENCE AND TECHNOLOGY, KOCHI

FIRST CLASS ACHIEVED, 2019

Class 12,

RAJARAM SHAH COLLEGE, RAXAUL

66.8 PERCENT ACHIEVED, 2013

Class 10,

HAZARIMAL HIGH SCHOOL, RAXAUL

74 PERCENT ACHIEVED, 2011

GENERAL INFORMATION

Date of Birth:	20 October 1995
Father's name:	MD GANI
Marital Status:	Single
Citizenship:	Indian
Languages (Read, write, and speak):	English, Hindi
Contact No:	+91- 7736267843