

Suyog Kantaram Pote

Employer : Data Quarter Technologies

Phone no: +91-95614 44965

Email id : potesuyog99@gmail.com

Address : At- Shankarpur, Post- Katepimpalgaon, Tal- Gangapur, Dist-Aurangabad, MH-423701.

Objective

Looking for a long term career in technical role, starting with position of a Bigdata Developer. Experienced individual with strong statistical analytic capabilities and ability to work with a variety of data environment. Looking to apply extensive data scientist experience in a fast-paced and dynamic firm to create data collection models that generate insight.

Career Summary

- 6.3 years of Experience in different domains in IT industry from that 3.5 years' experience in the field of Big Data as a Big Data Developer
- Have experience in Hadoop Ecosystem, Pyspark , SQL, Cloud Services (AWS), MySQL database, Linux shell script.
- Have Experience building optimized big data pipeline using HDFS, Hive, Hbase, PySpark on cloud premises.
- Have Experience relational and dimensional database
- Collecting large amount of data from Amazon RDS (MySQL) using Apache Spark and load it into HIVE in HDFS analysis.
- Have a hands-on experience on Amazon web services mainly S3, EC2, EMR, IAM Roles, RDS, Cloud watch.
- Good understanding of Partitioning, bucketing concepts in Hive and designed both Internal and External tables in Hive to optimize performance.
- Good experience on Spark Architecture including Spark-Core, Spark-SQL,
- Experience in working on file compression techniques as well as various file formats like ORC, Parquet and AVRO.
- Have experience with job scheduling in Oozie
- Capable of processing large sets of structured, semi-structured data.
- Can easily co-operate, interact and work in a team.
- Strong problem solving & technical skills coupled with clear decision-making.
- Hands-on experience with EMR Apache Spark Hadoop technologies Good knowledge of Configuring and working on Multi-node clusters and distributed data processing framework Spark. Python and PySpark, Spark SQL.
- Excellent oral and written communication skills.
- Strong analytical, problem solving and investigative skills.
- A high level of intellectual curiosity, external perspective, and innovation interest

PROFESSIONAL EXPERIENCE:

Project Name: Health Care Analysis

The purpose of the Project is to ingest the hospital prescription data that is present on RDBMS cloud storage. 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 (SQL, Oracle) using Sqoop.
- Write Hive DDL to Create Hive Table for Optimize Query Performance.
- Ingest Flat Files like Delimited, Fixed Length, etc. into Hive Warehouse.
- Implement Data Validation, Quality Checks, Profiling.
- Involved in import data from various RDBMS into HDFS using Sqoop which includes
- Incremental Load to populate Hive External Table and vice-versa.
- Involved in import data from various RDBMS into Hive Tables which includes Queries using Sqoop.

Project Name: Data Migration into Data Warehouse

The purpose of the project is to do the ingestion of the data that is present in Mainframe system to Big Data Warehouse. It's Mainframe to Hadoop migration project. Data migration utility to migrate existing system into Hadoop (Big Data) from Mainframe system. History creation process of data warehouse application and perform aggregation operations on Hadoop Big Data platform.

Roles and Responsibilities:

- Import Data into the Hive from various Relational Databases (Oracle, MYSQL Server) into HDFS using Sqoop.
- Ingest Flat Files like Delimited, Fixed Length, etc. into Hive Warehouse.
- Involved in creating tables, loading data into hive tables and process it as per the business logic using hive queries.
- Designed both Managed and External Hive Tables and Defined static and dynamic partitions as per requirement for optimized performance on production datasets.
- Perform optimization techniques in order to enhance performance of Sqoop job.
- Worked with various HDFS file formats like ORC, RC, Avro, CSV.
- Load processed final data to AWS S3.
- Engage with BA to understand the requirements clearly.

Technical Skill Set

- **Big Data Ecosystems** : Hadoop, Hive, Sqoop, Oozie and Spark
- **Cloud Technologies** : AWS (EC2, RDS, S3, IAM and EMR)
- **Language** : Python, SQL, Linux Shell Scripting.
- **Database** : MYSQL, Oracle
- **Operating Systems** : Linux , Window

Educational Qualification

- Bachelor of Engineering from Savitribai Phule Pune University, Pune with 67.27%.

Declaration

I, Suyog Pote, hereby declare that the information contained herein is true and correct to the best of my knowledge and belief.