AKSHAT JAIN
Analyst - Data Engineer at Evalueserve

Email: jainakshat404@gmail.com | Mobile: +91 7417540681 | LinkedIn: www.linkedin.com/in/akshat-jain-dataeng

PROFESSIONAL SUMMARY

- Having 3 years of professional experience in IT which includes work experience in **Hadoop and Big Data** technologies.
- Strong experience with big data processing using Hadoop technologies **Map Reduce, Apache Spark and Hive**.
- Good knowledge in streaming applications using **Apache Kafka**.
- Hands on experience with data acquisition into Hadoop cluster using **Sqoop** and **Flume**.
- Experience in analysing data using **Spark SQL**, **HIVEQL and Spark/Python**.
- Experience in **EMR, Lambda Functions, AWS Glue, AWS Athena, AWS EC2, AWS S3** and other services of the AWS family.
- Worked with Relational Databases such as **MySQL and Oracle**.
- In-depth knowledge of **Snowflake Database, Schema and Table** structures.
- Good understanding/Knowledge of **Elastic Search**.
- Good knowledge of processing large sets of **structured**, **semi-structured** and **unstructured** data and supporting systems application architecture.
- Good in writing **SQL queries** and database analysis for good performance.
- Familiar with Data Architecture including Data Ingestion **Pipeline** Design, Hadoop Information Architecture, **Data Processing** and **Data Modeling**.
- Excellent understanding of all aspects of the Software Development Life Cycle with timely delivery against aggressive deadlines and ability to adapt a new environment/technology and work on new projects.
- Effective in working independently and collaboratively in teams.

CORE COMPETENCIES

- Big Data Processing
- Data Warehousing
- Data Governance
- Data Modeling & Design
- Data Architecture

TECHNICAL SKILLS

- **Big Data Technologies:** HDFS, YARN, Map Reduce, Spark, Sqoop, Hive, Flume and Kafka

- **Database/DW:** MySQL, Athena, Redshift, Snowflake, HBase and Elasticsearch

- **Languages:** Python, Scala, SQL, HiveQL

- **Data Formats:** JSON, CSV, ORC and Parquet

- **Development Tools:** PyCharm, Jupyter

- **Project Management Tool:** GIT

- **Operating Systems:** Linux, Windows

- **Cloud Platform:** AWS

PROFESSIONAL EXPERIENCE

Evalueserve
Business Analyst (Mar 2020 – Present) Gurugram, India

Baxter: The purpose of this project is to develop an automated Quarterly Sales Operations Review dashboard for APAC by leveraging the existing design and available data and to improve user (Territory and Region Managers) experience through enhanced visualization and by adding actionable insights.

Role and Responsibilities:

- Responsible for **Data Ingestion, Data Cleaning, Data Anonymization and Data Transformation**.
- Worked on creating **Hive managed** and **external tables** based on the requirements.
- Implemented **Partitioning** and **bucketing** on Hive table for better performance.
- Developed Map Reduce Job in **Scala spark** for data cleaning and pre-processing.
- Involved in optimization of existing algorithms in Hadoop using **Spark Context, Spark-SQL, Data Frames** and Pairs RDD's.
- Running Analytics **Pipeline** for Aggregating and joining the multiple daily ingestion tables to meet the stakeholder's requirements.
- Scheduled the **jobs** and transferring final output data into **Redshift** for reporting purpose.

Nat It Solved Pvt Ltd
Big Data Engineer (Dec 2018 – Feb 2020) Kolkata, India

Promethium: This project deal with the frictionless data management that helps data teams discover, prepare, query and visualize data from multiple data sources in minutes instead of months without requiring data to be migrated to a new platform. It is out of the conventional process and doctrine around analytics and big data.

Role and Responsibilities:

- Involved in import data from various RDBMS into HDFS using **Sqoop** which includes **Incremental Load** to populate Hive External Table.
- Involved in developing **Python** Scripts, **Hive** Scripts (**HQL**) and **Elasticsearch** queries to process and analyze the data, also developed **PySpark** Codes using **SparkSQL** and **Data frames**.
- Worked and learned a great deal from **AWS** cloud services like **EC2, RDS** and **S3**.
- Involved in writing **Restful APIs** and **CRUD** codes.
- Experience working in onshore/offshore model and ensuring smooth transition on daily basis.

EDUCATION AND TRAININGS
- Big Data and Hadoop Training of 6 Months from KVCH Pvt. Ltd. (Noida), INDIA.
- Bachelor of Technology from JP Institute of Eng. & Tech. (AKTU), INDIA in 2018.

DECLARATION
I hereby declare that the above-mentioned information is true and best of my knowledge.
.
Date: (Akshat Jain)