Upasnasingh A Chauhan

DATA ENGINEER

📱 +91 8956464996

✉ upsingh9223@gmail.com

📍 Amravati, IN

PROFESSIONAL EXPERIENCE

Data Engineer   Aug 2019 – Present

At Bitwise Global Pune

Project 1: Product Analysis & Marketing   Oct 2021 - Present

The purpose of this project is to clean and validate the data of different category of products sent by upstream using PySpark and also store the validated data into Redshift

Roles and Responsibilities:

- Imported required data from source to target (Redshift) location using Glue jobs.
- Involved in monitoring and handling Glue jobs.
- Writing required logic to load as well as extract necessary data and writing queries that will run internally.
- Import data from S3 raw bucket and do the cleaning and masking and again move to S3 cleanse bucket by using glue.
- Extracting files according to customer's requirement at desired S3 location.
- Worked with various HDFS file formats like CSV, Parquet, RC, ORC .
- Written logic to extract the data according to customers requirement.

Project 2: Data Migration   March 2019 – Oct 2021

Objective of the project is to do ingestion of data that is present in source to Data Lake. Extracting data and perform Cleansing on the data using pyspark and loading data into S3 cleansed bucket. After validation load it into Redshift.

Roles and Responsibilities:

- Engage with BA to understand the requirements clearly
- Understand data migration needs and requirement
- Analyze and organize raw data
- Involved in data cleaning and day to day monitoring if any document is missing
- Efficiently loading data into Spark
- Extract data from different sources S3, jdbc , RDS
- Implement Data Validation
- Perform cleaning according to requirement
- Copy data into S3 bucket and load data into Redshift
- Data Reconciliation to ensure good Data Quality

EDUCATION

Sant Gadge Baba Amravati University   Jun '12 – Aug '14

Master in Applied Electronics │ Amravati, IN

Sant Gadge Baba Amravati University, formerly known as Amravati University is a NAAC Re-Accredited with Grade "A" by NAAC University of Amravati, Maharashtra.

- **Post Graduated with First Class with Distinction**

HVPM Amravati   Jun '09 – May '12

Graduate in Computer Science │Amravati, IN

Degree College of Physical Education, Hanuman Vyayam Prasarak Mandal Amravati is NAAC "A" GRADE MULTIFACULTY AUTONOMOUS COLLEGE.

- **Graduated with First Class with Distinction**

SUMMARY

Enthusiastic Engineer with 3.3 years on job experience in IT Industry as data engineer highly skilled in bigdata domains including Hadoop Ecosystem, PySpark, Cloud Services (AWS), SQL database such as MySQL and ORACLE, NoSQL Databases such as HBase. Well acquainted with pipelines, Data structures, data management and processing systems. Strong desire to move forward , face new challenges and expand my skill set.

KEY SKILLS

- Data Ingestion
- Data Cleaning
- Project Management Tools
- Data Processing
- Data Pipelines
- Performance Optimization
- Critical Thinking
- Distributed Systems

TECHNOLOGY PROFICIENCY

- **Operating Systems**– Linux(Ubuntu) and Windows
- **Databases** – Oracle and MySQL
- **Programming Language** – Python
- **Bigdata Tools** –Hadoop, Apache Spark, Cassandra.
- **Hadoop Ecosystem** – HDFS, Hive, Hbase
- **Cloud Technologies** – AWS (RDS, S3, EC2, EMR, IAM, Athena, Glue)
- **Data Warehouses** – Redshift
- **IDE'S** – PyCharm, Jupyter Notebook
- **Miscellaneous** – Jira, Git, Github