Sumit Kumar Saxena

Email: sumitsaxena266@gmail.com

Phone: +91-8909512064

Professional Exp:

1. **6+** years of experience in software **Developing, Debugging and process improvement.**
2. Experienced in **Python, PySpark ,MySQL ,Redshift**
3. Experienced in **AWS services** like **Glue/Redshift/S3/Athena/BOTO3/EC2.**
4. Experienced in data analysis and using dataframes through **SparkSQL**.
5. Experienced in working with **Spark RDDs** and **Map/Reduce** functions using **PySpark.**
6. Experienced in data migration to **AWS S3 from SAS9.2. as well as from S3 to redshift.**
7. Good knowledge in Relational Database model.
8. Experienced in working with various Python Integrated Development Environments like IDLE and jupyter-python.
9. Experienced in handling large volume of data and data migration/export.
10. Played role of Data Analyst and Data profiling activities to established strong communication between all the project stake holders.
11. Possess good analytical and debugging skill.
12. Experience of working at client place, work orientation according to client requirements and facing /handling clients on regular basis.
13. Worked regularly on Tables, Procedures, Packages, Functions, Collections, and Shell scripting and server management.
14. Strengths include ability in meeting deadlines and deliverables while achieving excellence and quality of work, quick learning of new concepts & techniques.
15. Dedicated to produce professional work with highest quality and creativity.
16. Strong understanding of the Software work environment with good analytical skills.
17. Possess excellent debugging skills .
18. Experienced in Bulk Data migration through unix shell scripting .
19. Developed and enhanced PL/SQL programs and Unix Shell scripts.

Areas of Expertise:

- **ETL tools:AWS-Glue,Google Cloud(Big Query-ETL)**
- **AWS SERVICES:GLUE,EC2,BOTO3,ATHENA,SYSTEM MANAGER,IAM USER**
- **IDE:pycharm,eclipse,jupyter**
- **Big Data Ecosystems:PySpark, Hadoop,Map-reduce, HDFS, Hive, Sqoop,Flume**
- **Programming Languages:C,python,core java ,SQL,Hive-QL**
- **Scripting Languages:Python**
- **Databases:My SQL,Hbase,Redshift**
- **Platforms:Windows(2000/XP), Linux, UBUNTU,CENTOS**

Company	Designation	DURATION
DATATOBIZ	Senior Data Engineer	April 2020-till now
DIVERSE LYNX PVT LTD	Senior Software Developer	March 2019-March 2020
Morling Global Pvt Ltd	Software Developer	Oct 2014– march 2019
Century papers ltd	Trainee	Sep 2012 – Sep 2013

Professional Experience:

1-Transsight(Migrating the data from S3 to Redshit)

Client   : Samtrans USA
Role   : Python,Pandas and PL/SQL Developer,AWS BOTO3,Redshift,S3,EC2,IamUser

Responsibilities:
- Read and analyze data using **pandas ,BOTO3 and s3**.
- File Handling using **Pyandas** and creating **dataframes** from file.
- Working with **SQL** and write interactive queries in dataframes .
- Developing and maintaining **PL/SQL** Functions/Procedures.
- Using **PL/SQL analytical functions** for data analysis and creating reports.
- Involved in requirement gathering, understanding current system.
- Responsible to **make data pipelines** to validate tables from s3.
- Using operations of **python and aws services** and fixing the code according to condition.
- Data viualisation and building and querying tables using **AWS redshift** and maintaining redshift objects in and mart,detail etc.
- Creating tables in **aws redshift.**
- Analyzing data to identify issue of **data error.**

2-Athena Replateforming(Migrating the data from sas9.2 to Aws-hadoop)

Client   : RBS , United Kingdom
Role   : Python,PySpark and PL/SQL Developer,AWS Glue,S3,EC2,Hive,

Responsibilities:
- Read and analyze data using **Pyspark**.
- File Handling using **PySpark** and creating **dataframes** from file.
- Working with **SparkSQL** and write interactive queries in dataframes .
- Performed transformations and actions on Spark RDDs using **PySpark**.
- Exposure in Map/Reduce functions through PySpark.
- Developing and maintaining **PL/SQL** Functions/Procedures.
- Using **PL/SQL analytical functions** for data analysis and creating reports.
- Involved in requirement gathering, understanding current system.
- Responsible to **make a data validation tool** to validate tables from sas and hive.
- Using operations of **pyspark** and fixing the code according to condition.
- Converting sas codes into **pyspark** codes.
- Data viualisation and building and querying tables using **AWS Athena** and maintaining Athena objects in staging and mart,detail etc.
- Creating tables in **hive,Using athena and aws services.**
- Analyzing data to identify issue of **data error.**

3) Alpha:Analyzing data to identify usage of data
Client   : BT, United Kingdom
Role   :AWS Glue, Python, PySpark Developer

Responsibilities:
- Read and analyze data using **Pyspark** in **AWS Glue** job.
- Data extraction,Transformation and load using **pyspark** and **sparksql** in **AWS Glue**.
- Handling and managing large **dataframes**.
- Working with **SparkSQL** and write interactive queries in **dataframes** .
- Performed transformations and actions on Spark RDDs using **PySpark**.
- Exposure in Map/Reduce functions through **PySpark**.
- Using Python Pandas library to handle file validations and transformations in python windows environment.
- Data viualisation and building and querying tables using **AWS Athena** and maintaining Athena objects in staging and mart.

Education:
- B.Tech (ECE) from Uttarakhan Technical university in 2012 with 69.96%.
- 12th from GIC,Lalkuan,Uttarakhand in 2007 with 72.05%.
- 10th from GIC,Lalkuan,Uttarakhand in 2005 with 73.11%.

Personal Details:

- Date of Birth   :   May5, 1990.

- Nationality   :   Indian.
- Language Known   :   Hindi, English.
- Permanent Address   : A6-91,25acor,lalkuan,naintal,uttarakhand

Declaration:

I hereby declare that the above-mentioned information is correct up to my knowledge and I bear the responsibility for the correctness of the above-mentioned particulars.

(Sumit Kumar Saxena)

`