

Charu Aggarwal

A Big Data Engineer with 3+ years of work experience in development, architecting and unit testing. Proficient in various AWS and Azure cloud services such as RDS, Glue, Lambda, S3, EC2, QuickSight, Athena, ADF and analysis service in integration with Python and Spark.

House No 38B, Sector 9/2,
Hanumangarh Jn. Rajasthan
+91-7073197731
aggarwalcharu061@gmail.com

EXPERIENCE

KPMG India, Gurugram – Data Engineer

March 2021 – PRESENT

- ETL from SAP cubes to Azure Data warehouse using ADF and created data model using Analysis service.
- Generated various kinds of MIS and finance reports using PowerBI to be consumed by users for analysis.
- Wrote stored procedure to transform the data in the warehouse to be consumed in a model for analysis.
- Created logic apps to control the functioning of models and Azure services.
- Documentation of the end to end process of Audit & Publishing of reports and data flow.

Capgemini Technology Services India Limited, Hyderabad— Big Data Developer

June 2018 - March-2021

- Generated various kinds of knowledge reports using Power BI and Tableau based on Business specification.
- Worked with AWS Glue using PySpark and python shell for data cleaning and flattening of data to be consumed by Athena using Data Catalog.
- Developed Lambda functions using python and scheduled through Cloudwatch as per conditions provided.
- Developed scripts to schedule Cloudera jobs as per conditions provided.
- Wrote scripts to manage files and clean old or junk files in Cloudera.
- Wrote complex SQL queries, views, triggers to be consumed in scripts.
- Collaborate closely with other team members to plan, design and develop robust solutions.
- Wrote maintainable, extensible and modular code in a team environment.
- Consulted regularly with client about issues via IT Works

CERTIFICATIONS

Microsoft Certified: Data Analyst DA-100

Microsoft Certified: Azure Fundamentals AZ 900

Microsoft Certified: Azure AI Fundamentals AI 900

Microsoft Certified: Azure Data Fundamentals DP 900

AWS Certified: Cloud Practitioner

SKILLS

Fluent in Python, PySpark

PowerBI

Unix, Linux

SQL

AWS Services (Glue, Lambda, Athena, RDS, Redshift, S3, EC2)

Azure Services (Azure Data Factory, Azure SQL Database, Analysis service)

AWARDS

Rising Star for 2020.

LANGUAGES

English, Hindi

PROJECTS

Migration to AWS RDS – *Glue, Lambda, RDS, Spark, Python*

This framework intends to pick up data from S3 on a weekly basis and transform it using Glue(PySpark) to populate the respective RDS PostgreSQL databases and tables.

Stock Market Analysis – *Glue, Lambda, S3, QuickSight, Spark*

The purpose of this framework is to showcase which all companies are performing well as per their share prices in the market using Glue for data cleansing and QuickSight to generate reports.

Financial Analysis – *Glue, Athena*

The purpose of this framework is to create databases and tables using Data Catalog & Crawlers using Glue and to query the end results in Athena.

Access Control Application – *Cloudera, HUE*

This allows the users to add a new user to Cloudera and HUE or to update the details for a particular user.

On Demand Database – *SQL, Glue*

Responsible for creating the structure of tables and stored procedures to perform the backend tasks.

Job Scheduling – *Unix, Sql, Hql*

Wrote SQL, HQL and shell scripts to schedule jobs on Cloudera edge-nodes. The jobs were required to run on particular days of week and time.

Job Creation & Scheduling – *AWS Lambda, S3 & Cloudwatch*

Wrote Lambda functions to perform specific tasks with integrating to S3 and RDS and scheduled them using Cloudwatch to run on particular days and time.

EDUCATION

Bachelor of Technology — *Computer Engineering*

Poornima University, Aug 2014 - June 2018