Praveen kumar Thumati

Data Engineer (AWS & Azure)

Data Engineer with two and half years of experience in Cloud and ETL Process, self-reliant with the ability to perform analysis and make recommendations.

praveenkumart27@gmail.com ✉

+44 7469299025 📱

29 Ragstone Road, Slough, United Kingdom 📍

linkedin.com/in/ praveen-kumar-t in

TECHNICAL SKILLS

Azure	Azure Data Factory, Azure Data Bricks, Azure Data Lake Service, Azure SQL Server, Logic Apps, Synapse Analytics, Azure Devops, Azure Key Vault, Azure VM, Azure Analysis Service
AWS	AWS Glue, Athena, Redshift, EC2, S3,Lambda, Cloud Watch
Programming	Python, PySpark, SQL
Data science	Machine Learning, Deep Learning, Data Analysis, Natural Language processing
Other	Linux/Unix, Autosys

WORK EXPERIENCE

Associate Consultant

Capgemini Technology Services India

06/2018 - 12/2020   *Bangalore, India*

Summary

- 2.5 Years of experience on cloud platform (Azure & AWS).
- Proficient in Azure services (Data Factory, Data Bricks, SQL Server, Logic Apps, Synapse Analytics, Analysis Service, Virtual Machine etc).
- Proficient in AWS services (Glue, Athena, EC2, S3, Cloudwatch, Lambda, Redshift, VPC, IAM etc).
- Building Data Factory pipelines for data migration and data movement. Testing and scheduling of pipelines.
- Writing Spark code for Data pre-processing and data movement between layers.
- Built an automated notification system for Data Quality issues using Logic apps and SQL server.
- Built Automation script for Cube level data validations in python.
- Built a top layer of all data sources called Semantic layer using PySpark.

PROJECTS

ITBA - Smart Analytics (Atlas Copco) (05/2019 - 12/2020)

- Worked on ADF pipelines creation & scheduling, used most of the activities in ADF.
- Developed DataBricks notebooks using Python, Spark, Spark SQL for transformations and data injestions and automated testing.
- Writing stored procedures in SQL Server for transformations.
- Implemented Azure DevOps, create build & release of pipeline for ADF.
- Building logic apps flows for alerts and also worked on Key Vault.
- Worked on Data modelling dimensions, fact modelling.
- Worked on adding columns and making changed in Azure Analysis service.
- Interacting with client on daily basis for requirement gathering and delivery.
- **Platform:** Azure Data Factory, Databricks, Python, Pyspark, Logic Apps, Azure SQL Server, Key Vault, Analysis service, Azure Devops, SQL.

PROJECTS

Customer Success center (The Coca-Cola Company) (09/2018 - 04/2019)
- Programming with Python and Pyspark for ETL jobs in AWS Glue.
- Building Athena tables and creating CLI commands for automatic creation of tables and views.
- Responsible for Building semantic layer.
- Responsible for extracting data from Postgres database to AWS EC2 and From Ec2 to S3. From S3 the transformation glue jobs runs.
- Responsible for Placing Semantic scripts in EC2 server and scheduling with dependency using Autosys.
- **Platform:** Python, Pyspark, AWS (Glue, Athena, S3, EC2, lambda, DynamoDB, SES), Autosys, Linux, Postgres, SQL.

CERTIFICATES

Microsoft Certified: Azure Data Engineer Associate (10/2020 - 10/2022)
Credentials ID: H548-2489

AWS Certified Developer – Associate (05/2019 - 05/2022)
Credentials ID: NLRNN0YKB2BQ1R5M

Microsoft Certified: Azure Data Fundamentals (10/2020 - Present)
Credentials ID: H549-0482

Microsoft Certified: Azure Fundamentals (08/2020 - Present)
Credential ID: H503-0350

Oracle Cloud Infrastructure Foundations 2020 Certified Associate (09/2020 - 03/2022)
Credentials ID: 277574504OCIBF2020

EDUCATION

MSc. in Data Science and Analytics
Royal Holloway University of London
01/2021 - Present   *Egham, UK*

BTech. in Electronics & Communication
Sri Venkateswara College of Engineering
08/2014 - 05/2018   *Tirupati, India*

AWARDS

XtraMile (01/2020 - 04/2020)
Capgemini
- Award for outstanding performance in first Quarter 2020