**Rakshitha Janumpally** *Senior Data Engineer*

 rakshitha93j@gmail.com  United States

 6196073749

 [www.linkedin.com/in/rakshitha93](http://www.linkedin.com/in/rakshitha93)

# SUMMARY

With a career spanning nine years in data engineering and analytics, I bring expertise in the design and enhancement of robust data systems. Proficient across various databases, ETL tools, and cloud services, my focus is on elevating data reliability and quality across a diverse range of industries.

# PROFESSIONAL EXPERIENCE

2022/02 – present Boston, MA, United States

2019/11 – 2022/02

Deerfield, Illinios, US

2018/04 – 2019/11

Chicago, Illinois, US

## SENIOR DATA ENGINEER

*Rxsense*

* Developed and implemented data pipelines using Apache Beam and dbt for efficient data management.
* Improved AWS RDS performance, ensuring robust database management and optimization.
* Led initiatives for AWS Data Pipeline projects, streamlining and automating data workflows.
* Integrated Amazon MSK to enhance data streaming and processing capabilities.
* Guided AWS DynamoDB implementations, providing scalable and reliable NoSQL solutions.
* Orchestrated Kubernetes and AWS ECS environments to enhance deployment processes.
* Utilized Docker for containerization, ensuring application consistency and portability.
* Created real-time data processing solutions using Apache Flink, optimizing data lake management with AWS Lake Formation. Collaborated on healthcare projects, demonstrating extensive knowledge of data and regulatory requirements.

## CLOUD ENGINEER

*Walgreens*

* Led database management with AWS RDS and Amazon Aurora, ensuring operational efficiency.
* Leveraged AWS DynamoDB for high-performance NoSQL database services.
* Coordinated smooth data transfer and transformation through AWS Data Pipeline.
* Employed AWS ECS for scalable container orchestration.
* Executed database schema conversions with the AWS Schema Conversion Tool.
* Automated infrastructure management through Terraform.
* Ensured swift application deployment through Docker.
* Implemented Apache Kafka for real-time data processing and managed data streams with Amazon Kinesis.
* Administered relational databases, notably PostgreSQL.
* Managed cloud-centric NoSQL database services utilizing MongoDB Atlas.
* Automated ETL processes utilizing AWS Glue.
* Managed extensive data processing at scale through Hadoop YARN.
* Streamlined workflow orchestration with Apache Airflow.

## ETL DEVELOPER

*Gray Matter Analytics*

* Employed Informatica for robust ETL processes.
* Orchestrated cloud-based ETL using Azure Data Factory.
* Processed extensive events through Azure Event Hubs.
* Analyzed substantial datasets with Azure Data Lake Analytics.
* Administered databases using Azure SQL Database.
* Automated workflows with Apache Airflow and managed container orchestration through Kubernetes.
* Implemented infrastructure as code automation using Terraform.
* Oversaw globally distributed databases utilizing Azure CosmosDB.
* Deployed containerized applications with Azure AKS.
* Processed large-scale data with Azure HDInsight.
* Managed comprehensive data warehousing using Azure SQL Data Warehouse.
* Implemented ETL processes through Talend.

2016/04 – 2017/12

Pune, Maharastra, India

2014/12 – 2016/04

Chennai, Tamilnadu, India

## DATA ANALYST

*Zensar*

* Used Python (pandas, NumPy) to create data analysis solutions.
* Managed and analyzed data with Amazon Redshift.
* Produced dynamic visualizations and reports using Power BI.
* Automated data workflows with Apache Airflow.
* Maintained and analyzed data through Jupiter Notebooks.
* Simplified CI/CD processes using Jenkins.
* Developed and deployed data-driven solutions, analyzed and visualized data for actionable insights, optimized data processing techniques, and collaborated with cross-functional teams for data solutions. Pioneered innovative data analysis methodologies.

## SQL DEVELOPER

*CSS Corp Pvt Ltd*

* Managed and improved Oracle databases for better performance.
* Created complex SQL queries and scripts for database tasks.
* Used Informatica PowerCenter for ETL processes.
* Improved data streaming with Apache Kafka.
* Analyzed data deeply using R for important business decisions.
* Automated tasks efficiently with Jenkins.
* Made interactive dashboards and reports using Tableau.
* Put applications into containerized environments using Docker.

# SKILLS

## DATABASES

Oracle Database, Amazon Redshift, Azure SQL Database, AWS RDS, PostgreSQL, MongoDB Atlas,Google Cloud SQL.

## ETL TOOLS

Informatica PowerCenter, Azure Data Factory, Talend, AWS Glue.

## CLOUD SERVICES

AWS, Azure, Google Cloud Platform.

## PROGRAMMING LANGUAGE

Python(Pandas, NumPy), PL/SQL.

## DEVOPS

Jenkins, Docker, Kubernetes, Terraform.

## DATA STREAMING

Apache Kafka, Amazon Kinesis, Google Pub/Sub.

## BIG DATA TECHNOLOGIES

Hadoop YARN, Apache Airflow, Apache Beam, Google Dataproc.

## DATA ANALYTICS

R, Tableau, Power BI, Google Data Studio.

# EDUCATION

2010/07 – 2014/07 **Bachelor's of Technology in Computer Science and Engineering**

*OSMANIA UNIVERSITY*