

**SUBRAHMANYA BHAGATH GADAMSETTY**  
**DATA ENGINEER**

Phone: 469-288-3832 | Email: [bhagath.g@careernb.com](mailto:bhagath.g@careernb.com) | LinkedIn: [subrahmanya-b-g](#)

## **SUMMARY**

Seasoned Data Engineer with 5+ years of expertise in Data Extraction, Transformation, Modeling, Mining, Data Visualization, and System Design in Healthcare, Financial, and Real Estate Domains. Skilled in AWS and Azure public cloud services along with Big Data tools and Databases. Expert in Spark applications using Spark-SQL in Databricks for data extraction, transformation, and aggregation from multiple file formats for transforming the data to uncover insights into customer usage patterns. Well-versed in ensuring data governance, compliance, and security.

## **WORK EXPERIENCE**

### **DXC Technology, VA | AWS Data Engineer**

*Jan 2023 – Present*

- Generating monthly reports to include month-end revenue and profit, variance analysis of actual data to budget and forecasting, and project status. Improved MySQL data extraction by 30% with PySpark, enhancing model training by 20%.
- Integrating Python with AWS Lambda for serverless data processing, reducing infrastructure costs by 15%.
- Streamlined report generation via Power BI and Tableau dashboards, saving 4 weekly hours and achieving 30% faster reports.
- Enhancing Kafka-driven data streaming with Spark, cutting latency by 50% and boosting processing speed by 40%.
- Architected data modeling on AWS Glue and Python, improving ML accuracy by 15% by reducing redundancy.
- Implementing CI/CD for data pipelines using AWS Data Pipeline and Airflow, cutting deployment time by 20%.
- Utilized Python for automated Snowflake cluster monitoring, achieving a 25% increase in query concurrency.
- Establishing scripts to load data to hive from HDFS and was involved in ingesting data into the Data Warehouse using various data loading techniques.
- Exploited Pandas and NumPy libraries in Python for efficient data manipulation and analysis, enabling data-driven decision-making.

### **Zensar Technologies, India | Azure Data Engineer**

*Aug 2017 – Jul 2021*

- Revolutionized deployment workflows through the integration of PowerShell and Azure Cloud Shell, optimizing operations and slashing deployment time by 25%, leading to significant cost savings of \$250,000 and improved project delivery timelines.
- Developed and deployed advanced data transfer algorithms linking Azure with on-premises from T-SQL, MS SQL resulting in a remarkable 95% increase in data transfer speed and a 50% reduction in errors.

## **Highlighted Projects:**

### **Digital Payments Report**

- Utilized scheduling of Batch pipelines by creating triggers in Azure Data Factory, ensuring timely and automated data processing.
- Configured Spark executor memory and tuned Spark jobs using PySpark, decreasing job execution time by 40% and increasing processing efficiency by 20%.
- Operated Postman for API testing, reducing testing time of customer data by 40% and increasing coverage by 30%.
- Established Python scripts for Snowflake data loading, cutting the loading period by 65% and reducing ingestion errors by 80%.
- Streamlined Kafka for real-time data streaming, reducing latency and improving processing speed of customer transactions.
- Enhanced data governance by implementing Apache Atlas, ensuring compliance and reducing security risks by 40%.

### **Property Asset Management Report**

- Architected Tableau dashboards and Power BI on Azure for data processing and visualization, increasing insights accuracy by 40% and decision-making speed by 45%.
- Automated CI/CD pipelines with Azure DevOps, reducing deployment process by 50% and increasing frequency by 75%.
- Collaborated with stakeholders, increasing project completion rate by 20% and improving alignment with business objectives by 15%. Evolved data models, increasing processing speed by 30% and reducing processing errors by 15%.

## **TECHNICAL SKILLS**

- **Programming Language:** Scala, Python, SQL, Java, C#.
- **Big Data Ecosystem:** Spark, Hadoop/HDFS, MapReduce, HIVE, HBase, Sqoop, Oozie, Kafka, Airflow.
- **Cloud Technologies:** AWS (EC2, Amazon Redshift, IAM, Kinesis, Data Pipelines), Azure (Azure SQL DW, Azure DevOps, Azure Data Lake, Azure Data Warehouse, Azure Synapse Analytics, Azure Security, Azure Data Factory, Azure Data Bricks, IaC)
- **Packages:** NumPy, Pandas, Matplotlib, SciPy, Scikit-learn, Seaborn, TensorFlow
- **Analytics and Visualization Tools:** Tableau, Power BI, Talend, GitHub, Git, SSIS, SSRS, Google Analytics, KPI'
- **Data Processing/ Version Control:** PySpark, Jenkins, Kubernetes, Docker, PaaS, SaaS, SAS, GitHub, Git.
- **ETL and Database:** MS SQL, MongoDB, MySQL, Oracle, Informatica, Snowflake, Terraform, Cosmos DB, NoSQL, DB2, PL/SQL.

## **EDUCATION**

### **Master of Science in Information Systems**

Northwest Missouri State University, MO

*Dec 2022*

### **Bachelor of Technology in Electrical Engineering**

Jawaharlal Nehru Technological University Hyderabad, India

*May 2017*

## **CERTIFICATIONS**

- **AWS Cloud Developer Associate – Amazon Web Services**
- **Azure Data Engineer Associate – DP 203 – Microsoft Azure**