**Shashank Shashishekhar Reddy**

San Jose, California | (510) 892-7191 | shashankshashishekharreddy@gmail.com | https://www.linkedin.com/in/snkreddy

**Summary**

Results-driven Data Engineer with extensive experience in administering and enhancing data infrastructure, developing ETL pipelines, and supporting statistical data analysis for manufacturing environments. Skilled in creating efficient data models and facilitating communication across engineering teams. Proficient in Python, SQL, and data visualization tools to deliver actionable insights that improve production performance and drive decision-making.

**Work Experience**

|  |  |
| --- | --- |
| Kantar | Data Engineer | Dec 2024 - Dec 2024 |

• Administered and optimized existing data applications to improve production statistics including throughput and yield

• Developed ETL pipelines using Apache Airflow and Spark, supporting data analysis for customer-oriented inquiries

• Documented application functionality and provided training for end-users, facilitating a smooth user experience

|  |  |
| --- | --- |
| The Sparks Foundation | Data Engineer | Dec 2024 - Dec 2024 |

• Engineered robust ETL processes to manage and process vast volumes of data from multiple sources, significantly enhancing data integration

• Collaborated with engineering teams to support data inquiries related to product specifications and production processes, reducing data retrieval times by 30%

|  |  |
| --- | --- |
| The Sparks Foundation | Data Engineer Intern | Dec 2024 - Dec 2024 |

• Supported the development of data pipelines and conducted thorough analysis to identify trends that inform decision-making

• Assisted in optimizing SQL queries and automating data cleaning processes to ensure high data quality and reliability

**Skills**

Certificates: AWS Cloud Practitioner | Validation Code: QD8E074C82EQ1WCE Programming Languages: Python, C++, Java, R, SQL, Bash, JavaScript Cloud Services: AWS (EC2, S3, Lambda, Glue, Amazon Redshift), GCP, Azure Databases: MySQL, MongoDB, PostgreSQL, RDS, Elasticsearch. Tools: Hadoop, Spark, Airflow, Tableau, Pandas, Scikit-Learn, Keras, Docker, Terraform. Statistical Analysis: JMP, Origin..

**Academic Projects**

|  |  |
| --- | --- |
| Your Own Cabs Analysis | Dec 2024 - Dec 2024 |

• Developed data pipelines leveraging Kafka, Spark, and Hadoop to analyze booking and clickstream data

• Implemented data models that improved data throughput by 15% and provided critical insights for decision-making, enhancing operational efficiency

|  |  |
| --- | --- |
| Spar Nord Bank ETL | Dec 2024 - Dec 2024 |

• Managed the ETL process for transactional data from MySQL RDS to Amazon Redshift

• Conducted thorough data analysis to identify patterns in usage, optimizing ATM refueling processes and achieving a 20% reduction in unnecessary refills

**Education**

|  |  |
| --- | --- |
| Master of Science in Data Analytics | SJSU | Dec 2024 - Dec 2024 |

|  |  |
| --- | --- |
| PG Diploma in Data Science Specialization in Data Engineering | VTU | Dec 2024 - Dec 2024 |

|  |  |
| --- | --- |
| Bachelor’s in Computer Science | IIIT | Dec 2024 - Dec 2024 |