

Shashank Shashishekhar Reddy

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

Summary

Data Engineer with extensive experience in administering and enhancing data technologies and applications to support decision-making and operational efficiency. Proficient in data analysis, developing ETL pipelines, and implementing data-driven solutions for diverse industries. Exceptional skills in software development, statistical modeling, and data visualization to drive production improvements and customer satisfaction.

Work Experience

Data Engineer | Kantar

Apr 2020 - Apr 2022

- Administered and optimized data applications to enhance data infrastructure for advertising and sales insights
- Implemented Apache Spark and Airflow for robust data ingestion and transformation, integrating production statistics effectively
- Developed models to monitor throughput and yield, enabling significant quality and operational improvements
- Spearheaded data quality initiatives, significantly reducing discrepancies and ensuring accurate reporting
- Collaborated with cross-functional teams to deliver meaningful visualizations and analytical reports, increasing operational efficiency by 12%

Data Engineer | The Sparks Foundation

Jul 2019 - Mar 2020

- Engineered comprehensive ETL solutions to streamline data collection from various sources, enhancing data processing efficiency
- Partnered with engineering teams to conduct in-depth data analysis, supporting product development and quality assurance initiatives
- Developed and documented data application procedures and functionality to facilitate end-user training and support
- Automated data transformation processes with Python and SQL, significantly optimizing workflow and accuracy, yielding a 30% reduction in retrieval times

Data Engineer Intern | The Sparks Foundation

Apr 2019 - Jun 2019

- Assisted in the administration and enhancement of ETL processes, improving integration of diverse datasets
- Conducted data quality checks and analysis to inform engineering decisions and processes
- Optimized SQL queries, resulting in faster data extraction and processing capabilities
- Supported the development of documentation for operational processes related to data applications

Skills

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

Academic Projects

Your Own Cabs Analysis

Dec 2024 - May 2024

- Developed sophisticated data pipelines to analyze booking and clickstream data using advanced tools like Kafka, Spark, Hadoop, and Hive
- Ingested diverse datasets into a consolidated repository, enhancing processing speeds by 15%
- Leveraged data aggregation techniques in Spark to derive actionable customer metrics and service KPIs, bolstering data-driven decision-making

Spar Nord Bank ETL

Jan 2024 - Apr 2024

- Led the ETL process for Spar Nord Bank's transactional data, facilitating seamless data transfer from MySQL RDS to Amazon Redshift
- Utilized PySpark and Sqoop for data transformation and analysis, identifying key usage patterns that resulted in a 20% decrease in operational costs related to ATM refills

Education

Master of Science in Data Analytics | SJSU

Jan 2023 - Dec 2024

PG Diploma in Data Science | IIIT

Aug 2020 - Sep 2021

Bachelor's in Computer Science | VTU

Aug 2016 - Aug 2020