**Shashank Shashishekhar Reddy**

San Jose, California | [LinkedIn](https://www.linkedin.com/in/shnkreddy) | [shashankshashishekharreddy@gmail.com](mailto:shashankshashishekharreddy@gmail.com) | (510) 892-7191

**Summary**

Data Engineer with hands-on experience with designing, developing and maintaining Extract, Transform and Load (ETL) pipelines coming from diverse data sources, and creating dashboards tailored to end-user or team requirements, facilitating data-driven decision-making, and ensuring proficiency in version control and data lifecycle management.

**Work Experience**

**Data Engineer** | Kantar | Bengaluru, India Apr 2020 – Apr 2022

* Built data pipelines with Apache Airflow and Spark to collect ads and sales data, storing it in an Amazon S3 data lake.
* Used Apache Spark for data transformation and enrichment, ensuring reliability with schema validation and error handling.
* Implemented data quality checks to maintain integrity throughout ingestion, transformation, and storage phases.
* Performed EDA and created visualizations in Tableau and Excel to identify patterns, boosting holiday sales by 12%.
* Designed ROI models to increase sales by 34% by optimizing ad spend on high-performing channels.
* Mentored two junior team members, improving team productivity and meeting deadlines with positive stakeholder feedback.

**Data Engineer** | TheSparksFoundation| Bengaluru, India Jul 2019 - Mar 2020

* Engineered robust ETL pipelines to manage and process large volumes of data from multiple sources, enhancing data integration by 25%.
* Collaborated with cross-functional teams to optimize database performance, reducing data retrieval times by 30%.
* Designed and implemented data quality checks to ensure accuracy and reliability, significantly minimizing data discrepancies.
* Automated data transformation processes using Python and SQL, decreasing manual intervention by 40%.
* Developed dashboards and visualizations to track data metrics, providing real-time insights for stakeholders.

**Data Engineer Intern** | TheSparksFoundation | Bengaluru, India Apr 2019 – Jun 2019

* Supported the development of ETL processes, contributing to the seamless integration of new data sources.
* Conducted data analysis to identify patterns and trends, informing decision-making across departments.
* Assisted in optimizing SQL queries to improve data extraction efficiency, achieving a 20% reduction in processing time.
* Created scripts for automated data cleaning, enhancing data quality and accuracy.

**Skills**

Certificates: AWS Cloud Practitioner **|** Validation Code:QD8E074C82EQ1WCE

Programming Languages: C, C++, Java, Python, 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, Flume, Airflow, Kafka, Pandas, Scikit-Learn, Keras, PyTorch, Docker, Terraform.

**Academic Projects**

**Your Own Cabs Analysis |** [GitHub](https://github.com/shnkreddy98/YourOwnCabsAnalysis) Jan 2024 – May 2024

* Built data pipelines to analyze booking and clickstream data using Kafka, Spark, Hadoop, and Hive, driving actionable insights.
* Ingested clickstream data into Hadoop HDFS via Kafka and batch booking data into AWS RDS with Sqoop, increasing data throughput by 15%.
* Aggregated data using Spark and stored results in Hive tables to generate customer metrics like ride counts, conversion rates, and service KPIs.

**Spar Nord Bank ETL |** [GitHub](https://github.com/shnkreddy98/SparNordBankETL) Jan 2023 – May 2023

* Managed ETL for Spar Nord Bank’s transactional data from MySQL RDS to Amazon Redshift with Apache Sqoop, PySpark, S3, and Redshift.
* Conducted analysis on Redshift to identify usage patterns, optimizing ATM refilling and reducing refills by 20%.

**Education**

**Master of Science in Data Analytics** | San Jose State University, California, USA Jan 2023 – May 2024

**PG Diploma in Data Science** **Specialization in Data Engineering** | IIIT, Bengaluru, India Oct 2020 – Nov 2021

**Bachelor’s in Computer Science** | VTU, Bengaluru, India Aug 2016 – Aug 2020