

SANTOSH KUMAR GOUDA

Bangalore

✉ s.k.gouda21@gmail.com ☎ +91-7619264934 🐙 [Github](#) 💼 [Linkedin](#) 🏠 [LeetCode](#) 🧑‍🔬 [GeeksforGeeks](#)

More than 8+ years of experience as a Data Engineer, specializing in building scalable data pipelines that process millions of records to deliver actionable insights, features, and analytical solutions at scale.

SKILLS

Cloud Platforms: AWS (S3, IAM, RDS, Glue, Redshift, Lambda, Athena, CloudWatch, EMR, QuickSight), Azure (Blob Storage, ADF, Databricks, DevOps) – 3.5 years' experience

Data Solutions: Data Pipelines, Data Warehousing, Data Lakes, ETL

Programming/Query Languages: Python (Proficient), SQL (Intermediate), PySpark (Intermediate), Airflow, Data Structures & Algorithms – 4 years' experience

Version Control/Tools: Git, JIRA – 4 years' experience

Soft Skills: Communication, Business Acumen, Critical Thinking, Lifelong Learner

WORK EXPERIENCE

Sears Gopal Technologies India Private Limited, Pune

Nov 2022 - Present

Data Analytic Tech Lead

- Designed ETL pipelines to transfer data from on-premises OLTP systems (MySQL, PostgreSQL, MongoDB) and cloud storage (SFTP, S3) to Snowflake, increasing data ingestion speed by 30% and achieving 95% data accuracy for Home Appliance and Tool repair analytics.
- Automated the ETL workflow with Airflow, reducing manual intervention by 40% and maintaining a 99% on-time data delivery rate, ensuring near real-time data availability for business insights.
- Crafted SQL queries to address over 200 ad-hoc requests monthly, achieving a 98% stakeholder satisfaction rate for accuracy and speed.
- Python, S3, SQL, Snowflake, Airflow, Data Migration, SFTP, MySQL.

Enkash, Bangalore

May 2022 - sept 2022

Data Engineer

- Built ETL pipelines to migrate data from an on-premises MySQL OLTP system to a Redshift data warehouse for a financial corporate cards company, reducing data processing time by 40% and enhancing analytics capabilities for over 200 users.
- Created tailored AWS Glue scripts to streamline data extraction and transformation workflows using advanced PySpark algorithms and optimized SQL queries; reduced processing time by over 25% per job on average.
- Constructed an analytics dashboard in AWS QuickSight, enabling business users to visualize monthly, weekly, and quarterly sales and marketing performance metrics.
- Python, SQL, AWS (S3, IAM, Athena, RDS, Glue, Redshift, EMR, QuickSight), PySpark, ETL, Data Modeling, Data Warehousing, Data Lakes, Data Pipelines

Tech Mahindra, Bangalore

May 2021 - May 2022

Data Engineer

- Designed a critical data pipeline that processed over 5 TB of data by consolidating information from multiple disparate-sources into a single destination, enabling swift data analysis for reliable business insights.
- Automated the manual data reconciliation process across different sources, resulting in a 90% increase in data accuracy and efficiency.
- Streamlined the ETL process, reducing data wrangling time by up to 40% through automation.
- Enhanced data retrieval efficiency by approximately 30% through query optimization and indexing techniques.
- Python, AWS (S3, Athena, IAM, Glue, Redshift, QuickSight), PySpark, SQL, ETL, Data Modeling, DataWarehousing, Data Lakes.

Alten India, Bangalore

Sept 2020 - Apr 2021

Data Engineer

- Migrated multiple SSIS packages from SQL Server to Azure Databricks, enhancing data processing capabilities and reducing ETL execution time by 35%, leading to faster data availability for analytics.
- Developed Azure Data Factory pipelines utilizing Linked Services and Datasets to ETL data from Azure SQL, Blob Storage, and Azure SQL Data Warehouse, improving data integration efficiency by 50% and supporting real-time analytics for business stakeholders

- Optimized notebooks in SQL, PySpark, and Python, streamlining data workflows and reducing processing times by 40%, which increased overall productivity in data analysis across teams.
- Python, PySpark, SQL, Azure Blob Storage, ADF, Data Warehousing, ETL, Azure Databricks

Bristlecone, Bangalore

Nov 2019 - Sept 2020

Data Analytics Engineer

- Implemented a systematic approach to plan, create, and maintain data architectures, ensuring alignment with business requirements.
- Created efficient ETL processes for data ingestion into Azure Data Lakes, processing over 500,000 records daily, which resulted in a 30% improvement in data availability and accessibility for analytics teams.
- Python, PySpark, SQL, Azure Blob Storage, ADF, Data Warehousing, ETL, Azure Databricks.

Accenture Solutions, Bangalore

Jan 2016 - Sept 2018

Data Analyst

- Implemented automated data extraction tools that streamlined the gathering of information from both primary and secondary sources, resulting in a 35% reduction in time spent on manual data acquisition tasks
- Diagnosed and rectified more than 50 coding errors in data extraction scripts, maintaining optimal system performance while safeguarding the integrity of vital financial records used by analysts daily.
- Implemented statistical forecasting techniques on 1M+ data points, improving demand planning accuracy by 30% and facilitating streamlined inventory management that reduced stockouts by 50% during peak seasons..
- Python, SQL, Excel, Power BI, Data Visualization.

EDUCATION

National Institute of Technology Nagaland

B.Tech in Computer Science and Engineering

May 2011 - Jun 2015

CGPA: 7.74/10