

# SANTOSH KUMAR GOUDA

[s.k.gouda21@gmail.com](mailto:s.k.gouda21@gmail.com) | +91-7619264934 | <https://linkedin.com/in/sgouda0412> | <https://github.com/sgouda0412>

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 analytics solutions at scale

## EXPERIENCE

---

### Data Engineer | Sears Global Technologies India Private Limited | Pune | November 2022 – Present

- 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 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, PostgreSQL, MongoDB

### Data Engineer | Enkash | Bangalore | May 2022 – September 2022

- Built ETL pipelines to migrate data from 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 Modelling, Data Warehousing, Data Lakes, Data Pipelines

### Data Engineer | Tech Mahindra | Bangalore | May 2021 – May 2022

- 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, ETL, Data Modelling, Data Warehousing, Data Lakes

### Data Engineer | Alten India | Bangalore | September 2020 – April 2021

- Migrated multiple SISS 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 EL 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.
- Python, PySpark, SQL, AzureBlob Storage, Azure Data Factory, Data Warehousing, ETL, Azure Databricks.

### Data Analytics Engineer | Bristlecone | Bangalore | November 2019 – September 2020

- Implemented a systematic approach to plan, create, and maintain data architectures, ensuring alignment with business requirements.
- Created efficient ETL processed for data ingestion into Azure Data Lakes, processing over 500000 records daily resulted in a 30 % improvement in data availability and accessibility for analytics teams.
- Python, PySpark, SQL, AzureBlob Storage, Azure Data Factory, Data Warehousing, ETL, Azure Databricks.

## **Data Analyst | Accenture Solutions | Bangalore | January 2016 – September 2018**

- 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 stock outs by 50% during peak seasons.
- Python, SQL, Excel, Power BI, Data Visualization

## **TECHNICAL SKILLS**

---

**Cloud Technologies:** AWS(S3, IAM, RDS, Glue, Redshift, Lambda, Athena, CloudWatch, EMR, Quicksight), Azure(Blob Storage, ADF, Azure Databricks, Synapses, Azure Develops)

**Data Solutions:** Data pipelines, Data Warehousing, Data Lake, ETL, Data Modelling

**Languages:** Python, SQL, PySpark, Data Structures and Algorithms

**Databases and Tools:** Oracle, MySQL

**Version Control/Reporting Tools:** Git, JIRA

**Scheduling:** Airflow

**Platforms:** Windows, Linux.

**Soft skills:** Communication, Business Acumen, Critical thinking, Lifelong Learner

## **EDUCATION**

---

Completed Bachelor of Technology in Computer Science and Engineering with a CGPA of 7.74 from National Institute of Technology, Nagaland in 2011 - 2015

## **CERTIFICATIONS**

---

**Associate Cloud Engineer, Google | September 2023 – September 2026**