

# Saurabh Ghumnar

[saurabhghumnar@gmail.com](mailto:saurabhghumnar@gmail.com) | 447-902-6357 | Champaign, IL | [www.linkedin.com/in/saurabhghumnar12/](https://www.linkedin.com/in/saurabhghumnar12/) | [Portfolio](#)

## SUMMARY

Results-driven Data Engineer with 3+ years of experience designing and implementing scalable ETL pipelines and cloud data infrastructure. Proven track record of improving data processing efficiency by 30%, reducing pipeline latency by 28%, and enabling real-time analytics for business-critical applications. Expert in Python, SQL, AWS, and Apache Airflow with hands-on experience building production-grade data architecture supporting 1M+ daily events. M.S. in Technology Management graduate from UIUC, with a focus on Big Data Infrastructures.

## EDUCATION

### Gies College of Business, University of Illinois Urbana Champaign

Champaign, IL

*Master of Science in Technology Management (STEM Designated)*

Aug 2025

*Specialization: Business Data Analytics - GPA: 4.00/4.00*

*Relevant Coursework: Big Data Infrastructures, Data Science and Analytics, Database Management Systems, Data Storytelling.*

### Savitribai Phule Pune University

Pune, MH

*Bachelor of Engineering, Computer Engineering - GPA: 3.74/4.00*

May 2021

## PROFESSIONAL EXPERIENCE

### QSTN LLC (Business Practicum) | Data Consultant

May 2025 - Present

- Increased platform scalability by 37% user growth and 40% engagement improvement by architecting real-time data pipelines and implementing automated ETL processes using Python, AWS Lambda, and Kinesis to process 500K+ daily events across 3 data sources with 99.9% uptime.
- Enhanced system reliability by 50% and accelerated deployment by 3 weeks by implementing automated data validation and monitoring frameworks using Apache Airflow and Prometheus to orchestrate 25+ data pipeline workflows with self-healing capabilities.

### Persistent Systems | Data Engineer II

May 2021 - May 2024

- Achieved 25% performance improvement and 99% system reliability by designing and implementing enterprise-scale ETL pipelines for McAfee's Data Loss Prevention platform using Apache Spark, Kafka, and AWS Redshift to process high daily data volumes with automated monitoring and alerting systems.
- Reduced data processing errors by 33% and improved pipeline reliability by implementing automated data quality monitoring and observability frameworks using Apache Airflow to establish real-time alerting and self-healing pipeline capabilities across 15+ data sources.
- Improved data availability from 82% to 94% and reduced query response times by 60% by optimizing data warehouse architecture and implementing automated ETL pipelines using Snowflake, dbt, and Python to enable real-time analytics and self-service data access for cross-functional teams.
- Contributed to 17% sales growth and 40% faster feature deployment by building automated data processing pipelines and implementing real-time analytics infrastructure using Python, SQL, and AWS services to enable data-driven development and A/B testing capabilities.

## KEY PROJECTS

### FinOps and Cloud Cost Optimization using Snowflake

Feb 2025 - May 2025

- Conducted comprehensive business requirements analysis and stakeholder management for cloud cost optimization, achieving 27% cost reduction and 40% improvement in resource utilization by creating detailed business cases, implementing FinOps best practices, and developing executive-level analytics dashboards.

### Marketing Offers using AWS

Nov 2024 - Jan 2025

- Increased conversion rates by 25% and enabled real-time personalization by building end-to-end streaming data architecture processing 1M+ daily events using AWS Kinesis, Lambda, and DynamoDB with Apache Kafka for real-time feature serving and ML model inference pipelines.

## SKILLS AND COMPETENCIES

- Data Engineering:** ETL / ELT Processes, Pipeline Architecture, Data Modeling and Validation, Data Warehouses, Database Architecture, Query Optimization, Real-time Processing.
- Technical Skills:** SQL, NoSQL, Python, C++, AWS, Microsoft Azure, Google Cloud Platform, Docker, Hadoop, Hive, EMR, Snowflake, Spark, Airflow, Kafka, Cassandra, Tableau, Power BI, Git, Jira, Jenkins.
- Certifications:** Certified Associate in Project Management (CAPM), Certified Data Engineer, Microsoft Azure Fundamentals, Microsoft Python Programming, SQL Fundamentals, Google AI Essentials.