

Abhilash Bagde

Data Engineer
LinkedIn | GitHub

Phone : (737) -396-6402
Email : bagde.abhilash77@gmail.com
Place : Austin, Texas

Professional Summary

Experienced and results-driven **Data Engineer with 3+ years** of hands-on experience **designing, building, and optimizing scalable ETL/ELT pipelines**, data warehouses, and analytical **data models** using **Python, SQL, Snowflake**, and dbt. Proficient in data **modeling, pipeline orchestration**, data quality frameworks, and cloud platforms including **AWS** and **GCP** to deliver reliable, high-performance data solutions. **Currently pursuing a Master of Science in Business Analytics (MSBA) at The University of Texas at Austin**, strengthening expertise in advanced analytics, **data-driven decision-making**, and business strategy. Proven ability to collaborate with cross-functional teams to standardize **KPIs, implement Infrastructure-as-Code practices**, and enable self-service analytics. Strong experience in data integration, real-time streaming using **Kafka**, and **data visualization** with Amazon QuickSight to drive actionable insights and operational efficiency across multiple domains.

Technical Skills

• Programming Languages	: Python, SQL	• Big Data Technologies	: Apache Kafka, Apache Hive, Apache HBase, Apache Flink
• Data Processing	: Apache Spark, Pandas	• Database Management	: MySQL, PostgreSQL, Snowflake, Google BigQuery, Amazon Redshift
• Streaming	: Apache Kafka		
• Testing & Monitoring	: Data Quality Validation	• Cloud Platforms	: AWS (S3, Lambda, Redshift, RDS), Google Cloud Platform
• Version Control & CI/CD	: Git, Docker, Airflow	• Data Pipeline	: Data Lake, Star & Snowflake Schema, Data Pipeline Orchestration
• Data Warehousing & ETL	: Apache Spark, AWS Glue, Airflow	• Security & Compliance	: Role-Based Access Control (RBAC), Data Encryption, GDPR

Professional Experience

DISCO Gurgaon, India	Business Intelligence Analyst	July 2023 – Nov 2024
----------------------	-------------------------------	----------------------

Roles and Responsibilities

- Designed, developed, and optimized scalable **ELT/ETL pipelines using Python, SQL, and Snowflake**, ensuring reliable ingestion, transformation, and delivery of analytics-ready datasets while reducing end-to-end data latency by 20%.
- Built and maintained scalable data warehouses** and analytical data models by integrating **multiple data sources**, implementing best practices in fact/dimension modeling and **star/snowflake schemas**, and ensuring optimized query performance and historical tracking.
- Implemented robust data quality, validation, and governance frameworks using **DBT tests and SQL checks**, proactively identifying anomalies, enforcing schema integrity, and ensuring accuracy, integrity, and security of data across pipelines.
- Standardized **enterprise-wide KPIs** by creating **DBT incremental models**, tests, and documentation, achieving metric consistency and improving weekly reporting performance by 25%.
- Engineered and maintained **real-time data streaming** pipelines using **Apache Kafka and Spark Streaming**, enabling near-instant processing and analytics of high-volume event data across multiple platforms.
- Designed and **optimized cloud-based data lakes** and storage architectures on **AWS S3, Redshift, and GCP BigQuery**, enabling scalable, cost-efficient storage and fast retrieval for analytics and machine learning workflows.
- Developed and enforced security and compliance measures including **role-based access control (RBAC), data encryption, and GDPR compliance**, ensuring sensitive data protection and regulatory adherence across all pipelines and warehouses.
- Developed **feature-ready** analytical models and **visualizations, including cohort analysis**, rolling engagement metrics, and time-series aggregations, enabling actionable insights across multiple products and domains.
- Optimized data pipeline and warehouse performance by **applying partitioning, indexing, and query tuning strategies** in **Snowflake and AWS** environments, improving system efficiency and reducing operational costs.
- Collaborated cross-functionally with **data scientists, engineers, and platform teams to design**, implement, and maintain end-to-end data architectures and pipelines, supporting advanced modeling use cases and reliable analytics.
- Enabled **self-service analytics** and executive reporting by delivering **curated datasets and optimized views** in **Amazon QuickSight**, improving accessibility and decision-making for non-technical stakeholders.
- Maintained comprehensive documentation and **data lineage tracking** for pipelines, models, and metrics, ensuring transparency, compliance, and simplified onboarding for cross-functional teams.
- Contributed to **Infrastructure-as-Code (IaC)** initiatives by reviewing **Terraform-based Snowflake configurations**, enhancing warehouse sizing, role management, and cost-performance efficiency.

Merkle Pune, India	Business Intelligence Analyst	April 2022 – July 2023
--------------------	-------------------------------	------------------------

Roles and Responsibilities

- Designed, trained, and validated Gradient Boosting and **Logistic Regression models using Python and Scikit-learn** to optimize marketing spend, driving a 10–15% improvement in ROI.
- Applied **NLP techniques (TF-IDF)** to analyze and classify ad **creatives, improving targeting relevance** and increasing click-through rate (CTR) by 9%.
- Integrated and processed large-scale marketing datasets** in Snowflake and automated creative taxonomy workflows, reducing manual effort by 60% and improving operational efficiency. Designed and executed **A/B testing frameworks** to evaluate model effectiveness, measure performance uplift, and deliver data-driven insights to support strategic marketing decisions.

VVDN Technologies	Marketing Executive Tech Author	Feb 2020 – July 2021
-------------------	-----------------------------------	----------------------

Roles and Responsibilities

- Analyzed end-to-end business **processes, systems, and workflows** to **identify inefficiencies**, conduct feasibility and impact assessments, and drive data-backed optimization initiatives aligned with strategic objectives.
- Collaborated with stakeholders to **elicit, define, and document** business requirements, translating them into detailed functional specifications, user stories, use cases, and process flows to enable effective solution delivery.
- Performed advanced **data analysis** and visualization to uncover **trends and actionable insights, monitored KPIs** to measure solution effectiveness, and recommended continuous improvement opportunities.
- Partnered with cross-functional teams to define **SOPs and best practices, led UAT activities** to ensure solution readiness and adoption, and introduced innovative ideas through research on industry trends and emerging technologies.

Education

- UT Austin course is MSBA 2026 May
- Master of Business Administration in Business Analysis & Marketing GPA: 7.17 / 10.00 | Graduated: July 2020
- Bachelor of Technology in Computer Science Lovely Professional University, Jalandhar, Punjab