Aarav Mehta

Senior Data Engineer

aarav.mehta@email.com — +91-9XXXXXXXXX — LinkedIn — GitHub

Professional Summary

Senior Data Engineer with 7+ years of experience designing, building, and scaling enterprise-grade data platforms. Proven expertise in batch and cloud-native data pipelines, data warehousing, and analytics enablement. Strong background in Python, SQL optimization, Airflow orchestration, and AWS-based architectures. Experienced in working with cross-functional analytics, BI, and data science teams.

Core Technical Skills

- **Languages:** Python, SQL
- **Big Data:** PySpark, Spark SQL
- **Orchestration:** Apache Airflow
- **Cloud:** AWS (S3, Glue, Redshift, EC2, IAM)
- **Databases:** PostgreSQL, MySQL
- **Warehousing:** Star & Snowflake Schema
- **Data Modeling:** Dimensional Modeling, KPI Layers
- **Data Quality:** Validation, Reconciliation, Monitoring
- **Version Control:** Git, CI/CD Pipelines

Professional Experience

XYZ Technologies Pvt. Ltd.   *Bangalore, India*
Senior Data Engineer   **Jan 2020 – Present**

- Led the design and development of large-scale batch data pipelines processing **800M+ records daily** from multiple transactional and third-party data sources
- Architected an enterprise data warehouse supporting finance, sales, marketing, and operations analytics
- Designed and implemented **dimensional data models** (fact and dimension tables) to enable fast BI queries
- Built Python-based ETL frameworks with reusable transformation logic and configuration-driven pipelines
- Orchestrated pipelines using Apache Airflow with dependency management, retries, SLA monitoring, and alerting
- Optimized complex SQL queries and table partitioning strategies, reducing dashboard latency by **45%**
- Implemented data validation, reconciliation checks, and anomaly detection to ensure data accuracy and consistency
- Collaborated closely with data analysts, BI developers, and stakeholders to define business metrics and KPIs
- Mentored junior data engineers and conducted design and code reviews

DataSoft Solutions   *Pune, India*
Data Engineer   **Jul 2017 – Dec 2019**

- Developed end-to-end ETL pipelines using Python and SQL for structured and semi-structured data
- Ingested data from relational databases, APIs, and flat files into centralized analytics systems
- Automated batch workflows using Airflow, improving pipeline reliability and observability
- Supported ad-hoc reporting and analytical data requests from business teams
- Worked on data migration projects from legacy systems to modern data platforms

Selected Projects

Enterprise Data Warehouse Modernization

- Designed a centralized data warehouse consolidating data from 15+ source systems
- Implemented star schema models optimized for BI and reporting workloads
- Enabled self-service analytics for business users, reducing ad-hoc data requests

Data Quality & Monitoring Framework

- Built automated data quality checks covering nulls, duplicates, schema drift, and reconciliation
- Implemented alerting and monitoring dashboards for pipeline health
- Improved overall data trust and reduced production incidents

Education

Bachelor of Technology (B.Tech) in Computer Science

ABC Institute of Technology   2017

Certifications

- AWS Certified Data Analytics – Specialty
- Apache Airflow Fundamentals

Additional Information

- Experience working in Agile / Scrum environments
- Strong collaboration with analytics, BI, and leadership teams
- Comfortable owning data systems end-to-end in production