MEET ZAVERI

📞 +1 (214) 836-7590 ✉ meet.zaveri29@gmail.com in meetzaveri 🌐 meetzaveri29.github.io

Education

University of Texas at Dallas, *Texas*   **August 2023 – May 2025**

Master of Science in Information Technology and Management   *GPA: 3.97/4.0*

Coursework: Business Data Warehousing, Big Data, Applied Machine Learning, Database Foundations, OOP with Python

Achievements: Business Analytics and Data Mining Graduate Certificate, Jungchan Hsieh Fellowship

University of Mumbai, *India*   **August 2016 – October 2020**

Bachelor of Technology in Electronics Engineering   *GPA: 3.33/4.0*

Skills

Programming: SQL, Python, PySpark, NoSQL, R, SAS

Data Engineering & Databases: ETL/ELT Pipelines, Apache Spark, Apache Airflow, Informatica IICS, MySQL, PostgreSQL, MS SQL Server, Teradata, MongoDB, Data Modeling, Data Warehousing, Data Governance, Data Quality

Cloud & Big Data: Azure (Data Factory, Synapse, ADLS), GCP (BigQuery, Dataflow), Hadoop (HDFS, Sqoop, Impala)

Data Visualization & Management: Tableau, Power BI, Matplotlib, Seaborn, ggplot2, JIRA, Confluence, Agile & Scrum

Certifications: Databricks Fundamentals Accreditation, Teradata Vantage Associate 2.3

Experience

Accenture   **July 2022 – July 2023**

Data Engineer *| SQL, Python, Tableau, Airflow, Informatica, Git, Teradata, GCP BigQuery*   *Mumbai, India*

- Architected and built end-to-end ETL pipelines for 3+ revenue management applications handling orders and scans data using SQL and UNIX Shell scripts, improving reporting of Postal revenue KPIs efficiently by 20%
- Implemented CI/CD pipelines with Git, Bitbucket, and Jenkins, reducing deployment time and eliminating $80,000/month in revenue leakage for a Package Scanning Interoperability project with cross-functional collaboration
- Improved ETL workflows for fleet data with Apache Airflow and Informatica IICS, reducing refresh times for 2+ Tableau dashboard from 2 hours to 10 minutes, enhancing scalability and analytics
- Enhanced query performance by fine-tuning Teradata SQL queries and BTEQ scripts processing over 1.5TB of daily operational and financial data, resulting in a 30% increase in efficiency and reduced execution time
- Spearheaded a proof-of-concept for migrating Teradata data products to Google BigQuery, achieving a 97% success rate and influencing strategic cloud adoption

Accenture   **January 2021 – June 2022**

Associate Data Engineer *| SQL, Python, Shell Scripts, Teradata, Splunk, QlikView*   *Mumbai, India*

- Designed and implemented robust data quality frameworks using UNIX shell scripts and SQL-based validation rules, resulting in a 98% reduction in ingestion errors and streamlined root cause analysis for 20+ data pipelines
- Automated batch monitoring and incident alerting with Python and Splunk for 100+ SQL workflows processing over 500M records daily, improving anomaly detection speed by 85% and supporting 99.9% SLA adherence
- Led L2 and L3 support for pipelines and analytics layers, handling 80+ high-severity data incidents across HR, payroll, and postal domains maintaining 99.5% data integrity across QlikView and SAP Business Objects reports
- Enforced GDPR-compliant data lifecycle management policies through structured backups and archival routines, reducing system downtime by 15% and ensuring audit readiness

Projects

Modern Data Platform *| Azure Data Factory, ADLS, RESTful API, Databricks, Synapse, Power BI*   **August 2025**

- Built an ELT solution on Azure to unify sales and customer data, automating 3+ years of data from 10+ sources via a metadata-driven Azure Data Factory pipeline with Medallion Architecture, cutting prep time by 85%
- Processed 18K+ customer records with PySpark in Azure Databricks, delivering a gold-layer warehouse in Synapse that reduced Power BI reporting latency from hours to minutes

Healthcare Data Warehouse *| PostgreSQL, Python, ETL, Tableau, Apache Airflow*   **October 2024**

- Designed a healthcare data warehouse using PostgreSQL, integrating survey, workforce, and expenditure data through Python (NumPy,Pandas) pipelines orchestrated with Apache Airflow
- Generated interactive Tableau dashboards to answer 5 key business questions, identifying workforce shortages and high healthcare expenditure areas, enhancing data-driven decision making

Truck Fleet Data Analysis *| Hadoop, HDFS, Hive, Impala, SAS, Statistical Analysis, Regression*   **April 2024**

- Built a scalable ETL pipeline with Hadoop and HDFS to process 5GB+ of structured fleet telemetry data daily, simulating high-volume operations and cutting query response times by 40%
- Developed predictive model in SAS to identify risk factors linked to mileage, achieving a 15% drop in traffic violations, and applied optimization techniques to improve fleet operations