

MEET NIRAV ZAVERI

+1 (214) 836-7590 meet.zaveri29@gmail.com [meetzaveri](https://www.linkedin.com/in/meetzaveri)

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: Python, SQL (MySQL, PostgreSQL, PySpark, Teradata), NoSQL (MongoDB), R, SAS

Data Engineering: ETL/ELT Pipelines, Apache Spark, Hive, Apache Airflow, DataStage, Apache NiFi, Data Modeling, Data Warehousing, Data Governance, Data Quality

Cloud & Big Data: GCP (BigQuery, Dataflow), AWS (S3, Lambda, Glue, Redshift), 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 Analyst | SQL, Python, Tableau, Airflow, DataStage, Git, Teradata, 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 IBM InfoSphere DataStage, 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
- Orchestrated a successful proof of concept for migrating Teradata data products to Google BigQuery with a 97% success rate, influencing strategic cloud adoption and enhancing stakeholder collaboration

Accenture

January 2021 – June 2022

Data Engineer Associate | SQL, Python, Excel, UNIX Shell Scripting

Mumbai, India

- Built data quality frameworks using UNIX Shell scripts and SQL validation checks, reducing data ingestion errors by 98% and deployed alerting in Splunk using Python for 20+ SQL batches, cutting incident detection time by 85%
- Monitored 100+ production jobs and reports processing 500M+ records daily, ensuring 99.9% uptime, additionally managed 50+ change deployments, reducing failures by 30% through governance best practices
- Resolved 200+ high-priority client incidents, maintaining 99.5% data accuracy in 15+ HR analytics, payroll, and postal products Business Object reports and QlikView dashboards improving stakeholder communication
- Optimized workload and performance tuning, cutting query execution time by 40%. Ensured 100% GDPR compliance by overseeing database maintenance and backup strategies, reducing downtime by 15%

Projects

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

Big Data Pipeline - Truck Fleet Data Analysis | Hadoop, HDFS, Hive, Impala, SAS

April 2024

- Engineered a robust ETL pipeline using Hadoop and HDFS, optimizing data ingestion, transformation, and storage. Enhanced data accessibility and reduced query response time by 40%, for logistics and risk factor calculation of drivers
- Utilized Hive and Impala for data processing and analysis on a 5GB dataset, creating a predictive model in SAS that validated the correlation between increased mileage and reduced risk, leading to a 15% drop in traffic violations