Nandini Ethirajulu

Raleigh, NC | 5102039719 | nandinie@buffalo.edu | Portfolio: https://nandini-1312.github.io
LinkedIn: www.linkedin.com/in/nandini-ethirajulu | GitHub: https://github.com/Nandini-1312

EDUCATION

University at Buffalo, The State University of New York — Buffalo, NY
Master of Science – Data Science and Applications

EXPERIENCE

Local Grown Salads — Remote, United States
Data Engineer — February 2025 - Present

- Designed and implemented data models for MRP systems in Odoo, integrating them with PostgreSQL for efficient storage and retrieval, leveraging Python to process data from 5+ subsystems.
- Built and optimized Python-based ETL pipelines, handling ~1TB of monthly data from farm operations, MRP systems, and external sources, ensuring data consistency and accuracy.
- Collaborated with cross-functional teams to discuss subsystem interactions, ensuring seamless data flow, integration, and centralized storage in PostgreSQL within the LGS ecosystem.
- Enhanced pipeline performance using Python, reducing data sync times by 24%, enabling real-time decision-making for farm management, inventory, and order processing, while integrating farmer functionality with Odoo MRP/ERP systems for streamlined operations.

Deloitte Consulting — Chennai, TN, India
Data Analyst — June 2021 - July 2023

- Engineered 300+ SSIS ETL packages to replicate and stage source tables for Tennessee State's Benefits system, implementing initial full loads, incremental updates, and optimizing performance through indexing, views, and structured data models.
- Developed a master ETL package to orchestrate child package executions, log execution details, and enable targeted restarts for failed tasks, reducing ETL downtime by 90%.
- Designed Oracle PL/SQL procedures utilizing advanced SQL techniques like partitioning, CTEs, complex joins, subqueries, and unions to process data for 200+ client reports, incorporating robust validation to resolve data inconsistencies.
- Implemented a data warehouse structure by transforming staging data into reporting tables, applying data modeling principles to support Tableau visualizations for transactional, eligibility, and benefit issuance data.
- Created fixed-length CSV files for federal submissions to the Administration for Children and Families (ACF), encapsulating state-level family and customer data derived from business logic.
- Optimized ETL workflows by 43% through advanced data cleaning, preprocessing, and redundancy elimination, ensuring seamless data integration and transformation.
- Collaborated on Tableau dashboards to deliver detailed and summarized views of Eligibility Determination, Benefits Issuance and Claims data, enhancing reporting accuracy and boosting user satisfaction by 29%.

Deloitte Consulting — Chennai, TN, India
Associate Analyst — August 2019 - May 2021

- Extracted, preprocessed, and standardized data from diverse sources for a Legacy System Conversion initiative for the State of Connecticut, achieving 98% data accuracy. Optimized SSIS ETL workflows and procedural scripts, reducing processing time by 50% and minimizing data loss to under 0.2%.
- Migrated transformed data to target databases using Salesforce's bulk loading utility, enhancing operational efficiency and post-conversion reporting. Led user acceptance testing (UAT), system validation, and data reconciliation post-migration to ensure data accuracy, verify against the source system, and confirm alignment with business requirements and ensuring the integrity of the converted data.

PROJECTS

Analysis of Patient Satisfaction in healthcare – A Multiple Regression Approach Using SAS — November 2024 – December 2024

- Conducted multiple regression analysis in SAS to predict patient satisfaction, achieving an R-squared value of 97.81%, explaining 97.81% satisfaction variation. Key predictors identified were average staff visits (40% satisfaction variation), number of nurses (7%), and patient age (negative correlation with satisfaction).
- Model successfully predicted satisfaction scores for new observations, with a 95% prediction interval for a sample patient to be between 34.13 and 41.29.

Loan Repayment Prediction Analysis — November 2024 – December 2024

- Preprocessed 37,408 records, applied Principal Component Analysis (PCA) to retain 87% variance, and addressed class imbalance using SMOTE for balanced training in predicting successful repayment or default risk.
- Built logistic regression models achieving 72.64% balanced accuracy post-SMOTE, with an AUC of 0.78 and a well-calibrated ROC curve, effectively identifying repayment outcomes and minimizing misclassification.

DBMS Project — September 2023 – December 2023

- Implemented a normalized Oracle SQL database for banking operations, using data modeling and an ER diagram to define relationships between a fact table and 9-dimension tables, while applying integrity constraints for data consistency.
- Developed optimized SQL queries for transactional and analytical purposes, leveraging indexing, partitioning, and complex joins for efficient data retrieval and reporting.

TECHNICAL SKILLS

Programming Languages:	Python, R, Oracle PL/SQL, Java, SAS, Snowflake, HTML, CSS, JavaScript
Data Analysis and Visualizations:	Machine Learning, Statistical Data Mining, SSIS, Tableau, Power BI
Database Management:	Oracle, MS SQL Server, My SQL, Excel, PostgreSQL
Tools:	Visual Studio, R Studio, Microsoft Office Suite, GitHub, Bitbucket, Jira, Odoo
Domains:	Benefits and Claims Data, Data Warehousing, Loan Default Prediction, Agile Framework