

Raj Sarode

San Jose, CA | rajsarode08@gmail.com | 617-653-0918 | [LinkedIn](#) | [GitHub](#) | [Portfolio](#)

EDUCATION

Northeastern University , Boston, MA <i>Master of Science in Information Systems - Concentration in Data Engineering</i> Relevant Coursework: Designing Data Architecture for Business Intelligence, Data Management and Database Design	Sep 2022 – May 2024 GPA: 3.84/4.0
University of Mumbai , Mumbai, India <i>Bachelor of Engineering in Computer Engineering</i> Relevant Coursework: Data Warehouse and Mining, Big Data Analytics, Database Management Systems	Aug 2018 – May 2022 GPA: 3.5/4.0

PROFESSIONAL EXPERIENCE

CREWASIS Brooklyn, New York Data Science Intern	Jul 2024 – Present
<ul style="list-style-type: none">Integrated ETL with BeautifulSoup for social media & news data, generating pickle files and NLP insights on 2M+ data pointsGenerated clusters to analyze relationships in textual data from various brands, improving accuracy by 60% using LSI with PythonCollaborated with the Marketing team to create Power BI reports, boosting campaign analysis and cutting reporting time by 15%	
Utsavi Trading Co. Mumbai, India Data Engineer Intern	Feb 2022 – Jul 2022
<ul style="list-style-type: none">Employed MS SQL and Python for advanced querying and data cleansing, achieving a significant reduction in data discrepanciesOptimized data processing with SQL Server to boost accuracy in financial reporting, revealing 20% more cost-saving opportunitiesStreamlined ETL workflows and improved risk analysis efficiency for pharmaceutical clients using Alteryx, resulting in a 15% revenue growth and a 60% reduction in manual operationsAutomated monthly reporting with advanced Excel (VLOOKUP, Pivot tables), reducing data compilation and analysis time by 35%Implemented Power BI dashboards for financial analysis of 10,000+ metrics, enhancing data quality and saving 18 hours/week	
Data Operations Intern	May 2021 – Aug 2021
<ul style="list-style-type: none">Conducted data analysis for pharmaceutical clients with 1M+ monthly active users, employing Python and MySQL to optimize sales and distribution strategies, leading to a 25% increase in company revenueExecuted ETL using MySQL and Alteryx to integrate terabytes of data, reducing data processing time by 40% and increasing efficiency for the analytics teamTransformed data by integrating 10,000 records with Excel VLOOKUP and creating Pivot tables for real-time sales trend analysisEngineered a dynamic inventory dashboard with Tableau, reducing overstock by 50% and optimizing warehouse space by 10%Analyzed 5M+ rows of shipping data via SQL and visualized shipping delays in Tableau, resulting in a 30% drop in late deliveries	
The Sparks Foundation Mumbai, India Data Science and Business Analytics Intern	Feb 2021 – Mar 2021
<ul style="list-style-type: none">Developed a classification model for predicting heart disease using ML and analyzed the results on the basis of age, sex, etcPerformed exploratory data analysis and worked on predicting the results using KNN model and achieved an accuracy of 86%Created visualizations like scatter plots, heatmaps, and box plots to show correlations between symptoms and heart disease	

PROJECTS

IMDB Project (Hadoop, MS SQL, Talend, Alteryx, Tableau)	Jan 2024 – Mar 2024
<ul style="list-style-type: none">Designed IMDB data warehouse model using ER studio; created a schema of DW model in MS SQL for loading dataIntegrated, cleaned, mapped, and loaded around 30 million records in the IMDB data warehouse using AlteryxAnalyzed data by writing complex queries, performed data profiling, developed visualizations, and dashboards on Tableau	
IOWA Liquor Sales (Alteryx, SQL, Power BI)	Sep 2023 – Dec 2023
<ul style="list-style-type: none">Conducted in-depth analysis of 25M+ unclean liquor sales records from Iowa, identifying key trends and insightsDeveloped a dimensional model in Alteryx, reducing data processing time by 50% and achieving 99.5% accuracyCreated data visualizations in Power BI, improving KPI understanding by 70% and decision-making efficiency by 30%	
Collision Scope: NYC Traffic Dynamics (SQL, BigQuery, Azure Data Studio, Talend, Power BI, Tableau)	Feb 2023 – May 2023
<ul style="list-style-type: none">Extracted NYC-MVC data from BigQuery and ingested in Azure Data Studio for data validation and cleaning utilizing SQLPerformed data integration through Alteryx ETL processes, resulting in successful ingestion into the data modelGenerated advanced dashboards in Tableau and Power BI for NYC traffic, focusing on collision trends in high-risk areas	

SKILLS

Programming	Python, SQL, R, JavaScript
Database Systems	MySQL, PostgreSQL, Microsoft SQL Server, BigQuery
Visualization Tools	Tableau, Microsoft Power BI, SSRS, MS Excel, Matplotlib, Seaborn
Big Data	Alteryx, Talend Open Studio, DBT, ER Studio
Cloud Services	Amazon Web Services (AWS), Microsoft Azure, Snowflake, Google Cloud Platform (GCP)
Certifications	AWS Certified Cloud Practitioner, Google Analytics, Alteryx Designer Core