# Raj Sarode

Irving, TX | rajsarode05@gmail.com | 617-653-0918 | LinkedIn | GitHub | Portfolio

#### **EDUCATION**

## Northeastern University, Boston, MA

Sep 2022 – May 2024

Master of Science in Information Systems - Concentration in Data Engineering

**GPA:** 3.84/4.0

Relevant Coursework: Designing Data Architecture for Business Intelligence, Data Management and Database Design

#### University of Mumbai, Mumbai, India

Aug 2018 - May 2022

Bachelor of Engineering in Computer Engineering

**GPA:** 3.5/4.0

Relevant Coursework: Data Warehouse and Mining, Big Data Analytics, Database Management Systems

#### PROFESSIONAL EXPERIENCE

Utsavi Trading Co. | Mumbai, India

## Data Engineer Intern

Feb 2022 - Jul 2022

- Employed SQL and Python for advanced querying and data cleansing, achieving a significant reduction in data discrepancies
- Optimized data processing with **SQL Server** to boost accuracy in financial reporting, revealing **20%** more cost-saving opportunities
- Streamlined **ETL** workflows and improved risk analysis efficiency for pharmaceutical clients using **Alteryx**, resulting in a **15**% revenue growth and a **60**% reduction in manual operations
- Implemented Tableau dashboards for financial analysis on a vast inventory of 10,000+ usage metrics, enhancing the data quality

#### **Data Operations Intern**

May 2021 - Aug 2021

- Conducted data analysis of **50+** pharmaceutical clients, employing **Python** and **SQL** to optimize sales and distribution strategies, leading to a **25%** increase in company revenue
- Executed **ETL** using **SQL**, and **Alteryx** to integrate terabytes of data, reducing data processing time by **40**% and increasing efficiency for the analytics team
- Transformed data by integrating 10K records with Excel VLOOKUP and creating Pivot tables for real-time sales trend analysis
- Engineered a dynamic inventory dashboard with Power BI, reducing overstock by 50% and optimizing warehouse space by 20%
- Leveraged SQL to analyze 1M rows of shipping data and visualized delay patterns in Power BI, leading to a 30% reduction in delays

#### The Sparks Foundation | Mumbai, India

#### **Data Science and Business Analytics Intern**

Feb 2021 - Mar 2021

- Developed a classification model for predicting heart disease using ML and analyzed the results on the basis of age, sex, etc
- Optimized hyperparameters using grid search, resulting in a 5% improvement in model accuracy, achieving a final accuracy of 86%
- Performed 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

- Designed IMDB data warehouse model using ER studio; created a schema of DW model in MS SQL for loading data
- Integrated, cleaned, mapped, and loaded around 30 million records in the IMDB data warehouse using Alteryx
- Analyzed 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

- Conducted in-depth analysis of **25M+** unclean liquor sales records from lowa, identifying key trends and insights
- Developed a dimensional model in Alteryx, reducing data processing time by 50% and achieving 99.5% accuracy
- Created 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

- Extracted NYC-MVC data from BigQuery and ingested in Azure Data Studio for data validation and cleaning utilizing SQL
- Performed data integration through Alteryx ETL processes, resulting in successful ingestion into the data model
- Generated 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

## **LEADERSHIP EXPERIENCE AND RECOGNITIONS**

• Graduate Teaching Assistant (Prompt Engineering & Data Science)

• Treasurer (Computer Society of India)

Jan 2024 - May 2024 Aug 2021 - Jul 2022