Raj Sarode

Sunnyvale, CA | 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

CREWASIS | Brooklyn, New York
Business Analyst Intern   **Jul 2024 – Present**
- Improved analysis efficiency by **20%** by scraping and visualizing insights from textual data of **50+** consumer brands using **Python**
- Generated clusters to analyze relationships in textual data from various brands, improving accuracy by **60%** using **LSI** with **Python**
- Enhanced a **SQL Server** database with **60K** customer records for retail clients using SQL queries to achieve a **12%** revenue increase
- Streamlined data retrieval time by **30%** by optimizing database management with **MySQL** and **SQL Server** for quicker data access

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 **Azure SQL** to boost accuracy in financial reporting, revealing **15%** more cost-saving opportunities
- Streamlined **ETL** workflows using **Alteryx** to enhance risk analysis, incorporating data integrity checks to ensure quality and consistency, achieving a **60%** reduction in manual operations
- Automated monthly reporting with advanced **Excel (VLOOKUP, Pivot tables)**, reducing data compilation and analysis time by **40%**
- Engineered financial dashboard using **Power BI** to visualize KPIs facilitating a **25%** boost in decision-making for the PMO team

Data Operations Intern   **May 2021 – Aug 2021**
- Analyzed data for pharmaceutical clients with **1M+** monthly active users using **MS Excel** and **SQL**, using the insights to guide sales and distribution strategies, which increased company revenue by **20%**
- Modified **SQL** queries in **Azure Data Studio** for data retrieval of 1M+ records, reducing overall data transformation time by **40%**
- Spearheaded data modeling and validation of **1M+** healthcare records, executing **ETL** using **SQL** and **Alteryx** to integrate data
- Created comparative **Power BI** dashboard, illustrating variations in annual premiums and claims across diverse insurance products
- Examined KPIs and generated **Power BI** reports leveraging **DAX**, boosting product performance by 10% across different channels

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 IMDB data warehouse using **Alteryx**
- Analyzed data by writing complex queries, performed data profiling, developed visualizations, and dashboards on **Tableau**

Open Weather Cloud Analytics (AWS, Snowflake, Python, Sreamlit)   **Sep 2023 – Dec 2023**
- Built a **Python/AWS** ETL pipeline hosted on **EC2** to ingest data from **100+** US cities scraped from weather API into **Snowflake**
- Illustrated data visualizations to identify relationships between AQI, temperature, and humidity using **matplotlib, seaborn, plotly**
- Formulated an interactive **Streamlit** application to analyze relationships between climatic factors impacting major cities

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, Snowflake
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, Azure Data Factory, Azure Databricks
Certifications	AWS Certified Cloud Practitioner, Google Analytics, Alteryx Designer Core

LEADERSHIP EXPERIENCE AND RECOGNITIONS

- DATA Club - Active Member   **Sep 2022 – Present**
- Computer Society of India - Treasurer   **Jul 2021 – June 2022**