Vaishnavi Samboji

San Jose, CA | +1 (669) 321-7390 | vaishnavi.samboji@sjsu.edu | linkedin | Github

EDUCATION

MS in Data Analytics
Jan 2024 - Dec 2025

San Jose State University • San Jose, CA

Coursework: Computational Programming for Data Analytics, Data Visualization, Big Data, Machine Learning Technologies, Database Systems for Analytics, Mathematical Methods for Data Analytics, Deep Learning Technologies, Generative Model Applications.

B.Tech in Electronics and Communication Engineering
Aug 2019 - May 2023

Malla Reddy Engineering College For Women • Hyderabad, India

SKILLS

Programming Languages: C, CSS, HTML5, Javascript, MySQL, PL/SQL, Python, SQL

Statistics: Descriptive Statistics, Hypothesis Testing, Linear Algebra, Probability, Regression Analysis, Time Series Analysis

Frameworks/Libraries: Flask, Matplotlib, NumPy, Pandas, PySpark, PyTorch, Scikit-Learn, Seaborn

Other: API, AWS, Data Cleaning, Data Wrangling, ETL, Excel, GCP, Git, Hadoop, JSON, Microsoft Power BI, MongoDB, NoSQL, Oracle Forms, Oracle Integration Cloud (OIC), Oracle Reports, Tableau, XML

EXPERIENCE

Application Development Associate
Jun 2023 - Dec 2023

Accenture • Pune, Maharashtra, India

- Developed foundational database skills in **Oracle SQL, PL/SQL, Oracle Forms, and Reports** via stream training.
- **Improved data handling** processes by leveraging expertise in **JSON, XML,** and **Oracle Integration Cloud**, ensuring seamless data integration for the Veritas project.
- **Engineered** and **deployed Oracle Integration Cloud (OIC) solutions**, optimizing ERP Finance module workflows and reducing processing time by **20%**.
- Collaborated with teams to analyze and design integration requirements, driving a **15% enhancement in system scalability and performance**.

PROJECTS

Netflix Recommendation, EDA and Visualization, *GitHub Link*
Dec 2024

- Developed a Netflix recommendation system using **Python**, **NLP**, and Kaggle datasets, achieving **90% accuracy** and improving efficiency by 25% with TF-IDF and cosine similarity.
- Conducted EDA and created visualizations with Matplotlib, Seaborn, and Plotly to uncover actionable insights.
- Validated the system with titles like F.R.I.E.N.D.S., **achieving 85% user satisfaction**.

Spotify Data Analysis, *GitHub Link*
Nov 2024

- Analyzed the Spotify Top 50 songs dataset to identify trends in genres, song popularity, and audience preferences by performing extensive data preprocessing to ensure accurate and clean analysis.
- **Developed machine learning models** to predict song popularity, leveraging insights visualized with Matplotlib and Seaborn to support **improved decision-making** and data-driven strategies.

Movie Catalog and Analytics System using MySQL and MongoDB, *GitHub Link*
May 2024

- Designed a scalable movie catalog and analytics system using MySQL and MongoDB, **migrating data to AWS** to enhance performance.
- Optimized SQL and executed advanced NoSQL queries for complex **data retrieval**, while **data denormalization** and preprocessing were performed using Python for seamless MongoDB integration.