Aaryen DSouza

437-966-6119 | dsouza96@uwindsor.ca | linkedin.com/in/aaryendsouza | github.com/aaryen-dsouza | Portfolio

Technical Skills

Languages: Java, JavaScript, TypeScript, Python, C, SQL, HTML/CSS
Databases: Oracle SQL, PostgreSQL, MySQL, MongoDB, PL/SQL
Data Manipulation & Visualization: Pandas, NumPy, Matplotlib, Seaborn, Tableau, Power BI, (Learning) PySpark
Frameworks: React, Django, Redux, Node.js, Express, WordPress, Material-UI, Bootstrap, Tailwind CSS
Dev Tools: Git, GitHub, Jupyter Notebook, Google Collab, VS Code, DataGrip, PyCharm, IDEA, Postman, Jira
Other Skills: Windows, Linux/Unix (Ubuntu, CentOS), Shell Scripting, Microsoft Excel, Google Sheets

Professional Experience

Software Engineer | *Oracle SQL, MS Excel, Python, Bash/Shell Scripting, Java, Linux, Jira* July 2022 – Mar 2024
Vermont Information Processing (Acquired Vistaar US in July 2023) *Mumbai, India*

- Ensured data integrity by conducting data cleaning, preprocessing and validations through **Oracle SQL** and **MS Excel**, resolving 95% of data quality issues within 24 hours, and providing detailed analysis reports to stakeholders.
- Reduced issue resolution time by 40% for 100+ customer-reported issues by performing in-depth root cause analysis (RCA) on legacy app server and Oracle database, leading to faster issue identification and customer satisfaction.
- Automated ETL workflows by implementing **Bash/Shell** scripts, optimizing data workflows, and reducing manual interventions by 20%, which led to improved data onboarding efficiency.

Projects

Predictive Caching for Web Pages | *Python, Pandas, Matplotlib, Statsmodels, Scikit-learn, Git* Nov 2024
GitHub

- Developed a hybrid AI-driven caching system combining **ARIMA** for time-series forecasting and LLM for contextual analysis, improving cache efficiency and reducing latency.
- Improved data quality by automating data preprocessing, normalization, and missing value imputation using **Python** and **Pandas**, ensuring high-quality inputs for predictive models.
- Conducted exploratory data analysis (EDA) on web traffic data, identifying trends such as average daily and monthly page views, and visualized key insights using **Matplotlib**
- Built and optimized ARIMA models to forecast future page views, achieving high accuracy with RMSE values of **4.41** for ARIMA and **4.67** for Auto-ARIMA, and generated forecasts for the next 30 days.

Stress Level Prediction Using Machine Learning | *Python, Pandas, Matplotlib, Seaborn, Scikit-learn, Git* Nov 2024
GitHub

- Developed a stress prediction model for a mobile app by analyzing user behavior metrics, optimizing and training ML models on digital wellbeing data including screen time, app usage, and social media activity.
- Improved data quality by implementing data preprocessing, encoding, and normalization using **Python** and **Pandas**, ensuring high-quality inputs for predictive modeling.
- Trained and enhanced model performance by optimizing **Random Forest** and **XGBoost** hyperparameters, increasing classification, increasing accuracy to **77%**.

Bus Pass Scanner | *React, NodeJs, Express, MongoDB, Material UI, Git* Mar 2022
GitHub

- Designed and developed a cross-platform bus pass authentication application as part of a 3-member team, enhancing security and efficiency in college transportation.
- Implemented a dynamic, responsive UI using **React** and **Material UI**, improving accessibility and user experience across devices.
- Integrated **JWT**-based authentication, reducing unauthorized access attempts and strengthening the overall security of the system.

Education

University of Windsor Windsor, ON
Master of Applied Computing *May 2024 – Present*

University of Mumbai Mumbai, India
Bachelor of Engineering - Computer Engineering *Aug 2018 – May 2022*