SHADCHIKA SHANMUGARASA

DATA ENGINEERING INTERN

CONTACT

0766572148

shadchika20@gmail.com

shadchika

shadchika

Dehiwala, colombo

shadchika

EDUCATION

BACHELOR OF SCIENCE IN COMPUTER SCIENCE

SLIIT 2023 – 2026 (Expected)

- Relevant Coursework: Data Analysis, Machine Learning, Predictive Analytics, Data Visualization

TECHNICAL SKILLS

- Programming: Python (Pandas, NumPy, Scikit-Learn, TensorFlow, Keras), Java, C++, C#, JavaScript
- Data Analysis & Visualization: SQL, Pandas, NumPy, Matplotlib, Seaborn, Power BI
- Machine Learning & AI: Scikit-Learn, TensorFlow, Pattern Recognition, Predictive Modeling
- Databases: MySQL, PostgreSQL, MongoDB, SQLAlchemy
- Cloud: AWS (S3, Lambda, EC2), Azure (basic knowledge)
- Web Development: React.js, Node.js, Spring Boot
- Version Control & CI/CD: Git, GitHub, Jenkins, GitLab

SOFT SKILLS

- Problem-Solving
- Team Collaboration
- Adaptability
- Communication
- Time Management
- Creativity

CERTIFICATIONS

- SQL for Data Analysis – Coursera
- Machine Learning Specialization – Coursera
- Data Visualization – freeCodeCamp
- Selenium in Java – Test Automation University
- Power BI for Beginners – Udemy

PROFILE

Motivated and detail-oriented aspiring Data Operations Engineer with hands-on experience in software engineering, full-stack development, and database systems. Proficient in data modeling, SQL, ETL processes, data querying, and cloud tools like AWS. Skilled in Python, Power BI, and modern database management systems. Known for strong analytical thinking, problem-solving skills, and a passion for building reliable, scalable data workflows. Familiar with Agile methodologies and remote team environments. Eager to contribute to data infrastructure, automate pipelines, and support efficient data operations.

PROJECTS

Netflix Movie & TV Show Data Analysis

2025

Python, Pandas, Matplotlib, Seaborn | GitHub Repository

- Collected, cleaned, and preprocessed Netflix's movie and TV show dataset, handling missing values, reclassifying ratings, and formatting date fields for time-series analysis.
- Designed and implemented efficient data querying pipelines using pandas to extract insights and generate detailed reports.
- Built a structured data pipeline and developed interactive visualizations (bar plots, heatmaps, pie charts, line graphs) to present key patterns clearly.
- Applied data management best practices and delivered findings through an organized, reproducible Jupyter Notebook.

Walmart Sales ETL Project

2025

Python, Pandas, PostgreSQL, SQLAlchemy | GitHub Repository

- Configured Kaggle API locally, automated dataset download, and loaded Walmart sales data into Pandas for processing.
- Cleaned, transformed, and engineered new features in the dataset to prepare it for storage.
- Exported the refined dataset into a PostgreSQL database using SQLAlchemy for efficient querying and analysis.
- Developed and tested the entire ETL pipeline in Jupyter Notebook within VS Code, ensuring end-to-end data integration.

Spotify Data Pipeline Project using AWS

2025

Tools: AWS (S3, Glue, Athena, QuickSight), Python, Pandas

- 🔗 Processed Data: Google Drive
- 📌 Project Post: LinkedIn Showcase
- Cleaned and merged Spotify data from Kaggle using Python and Pandas.
- Built an end-to-end AWS ETL pipeline: S3 for storage, Glue for transformation, Athena for querying, and QuickSight for data visualization.

REFERENCES

MS. DILUSHINIE FERNANDO
SENIOR LECTURER UNIVERSITY OF SLIIT ,
DILUSHINIE.F@SLIIT.LK

DR. YASAS JAYAWEERA
SENIOR LECTURER UNIVERSITY OF SLIIT ,
YASAS.J@SLIIT.LK