

Machine Learning Programming

PROG8245

Lab2 - Data Collection and Pre-processing

Fasalu Rahman Kottaparambu

Student ID: 8991782

Date: May 28, 2025

GitHub Repository: <https://github.com/anonduke/Machine-Learning-Programming/tree/main/Week3/Lab2>

Project Source: This submission includes a Jupyter Notebook (Lab2_final.ipynb) implementing a 12-step data engineering pipeline for a synthetic e-commerce dataset. The repository contains the notebook, synthetic data (data/synthetic_sales.csv), city lookup (data/city_lookup.csv), README, and requirements.txt.