

Week 2: Data Visualization + Advanced ML Schedule

Overall Goals for the Week:

- Master data visualization (Matplotlib, Seaborn, Plotly)
 - Deep dive into supervised ML algorithms (regression, classification, evaluation)
 - Hands-on with real datasets (from Kaggle/UCI)
 - Complete **2 ML projects** (one classification, one regression)
 - Learn **model evaluation, hyperparameter tuning, and cross-validation**
 - Push all work to GitHub
-

Day 1 – Visualization Mastery + EDA

- **1 hour:** Learn advanced plots using Seaborn & Plotly (heatmaps, pairplots, violin plots, interactive charts)
- **1.5 hours:** Perform full EDA (Exploratory Data Analysis) on a dataset (e.g., Titanic or Iris)
- **30 mins:** Study a good EDA Kaggle notebook
- **30 mins:** Push visualizations to GitHub with proper README

Tools: Seaborn, Plotly, Pandas

Day 2 – Regression Deep Dive

- **30 mins:** Learn Linear & Ridge/Lasso Regression theory
 - **1 hour:** Code Linear, Ridge, and Lasso regression from scratch and using Scikit-learn
 - **1 hour:** Evaluate models (R^2 , MAE, MSE), learn residual analysis
 - **1 hour:** Work on a regression project (e.g., House Price Prediction)
-

Day 3 – Classification Algorithms

- **30 mins:** Learn Logistic Regression, KNN, Decision Trees theory

- **1.5 hours:** Implement classifiers and evaluate using confusion matrix, precision, recall, F1-score
 - **1 hour:** Start classification project (e.g., Breast Cancer Detection or Titanic)
-

Day 4 – Model Evaluation + Tuning

- **30 mins:** Learn about train-test split, cross-validation, bias-variance
 - **1.5 hours:** Practice cross-validation & hyperparameter tuning (GridSearchCV, RandomizedSearchCV)
 - **1 hour:** Tune your previous classification/regression models
-

Day 5 – Real-World Project (1)

- **2.5 hours:** Complete and polish **Project 1: Regression** (e.g., House Prices with full EDA, model training, tuning, results)
 - **30 mins:** Write a project README, upload to GitHub
-

Day 6 – Real-World Project (2)

- **2.5 hours:** Complete and polish **Project 2: Classification** (e.g., Heart Disease or Customer Churn Prediction)
 - **30 mins:** Document & upload to GitHub
-

Day 7 – Model Comparison + Feature Engineering

- **1 hour:** Learn and apply feature selection techniques (correlation, SelectKBest, Recursive Feature Elimination)
 - **1.5 hours:** Compare ML models (SVM, RF, Gradient Boosting)
 - **1 hour:** Visualize feature importances and model comparisons
-

Tools & Libraries to Focus on:

- **Pandas, NumPy**
- **Seaborn, Matplotlib, Plotly**
- **Scikit-learn**
- (Optional: XGBoost, LightGBM for advanced learners)