Data Science Study Plan (22 Sept – 31 Oct 2025)

Week 1 (22–26 Sept): Data Handling with Pandas & NumPy

Goal: Master data manipulation.

- **Mon 22 Sept** → NumPy refresher: arrays, indexing, operations.
- **Tue 23 Sept** → Pandas Series & DataFrames (create, inspect).
- **Wed 24 Sept** → Import/export data (CSV, Excel, JSON).
- **Thu 25 Sept** → Data cleaning: missing values, duplicates, types.
- **Fri 26 Sept** → Data wrangling: filtering, groupby, merge, pivot.

Week 2 (29 Sept – 3 Oct): Data Visualization

Goal: Tell stories with data.

- **Mon 29 Sept** → Matplotlib basics (line, bar, scatter).
- **Tue 30 Sept** → Advanced Matplotlib (subplots, styling).
- **Wed 1 Oct** → Seaborn basics (histogram, boxplot, scatterplot).
- **Thu 2 Oct** → Seaborn advanced (heatmaps, pairplot, catplot).
- **Fri 3 Oct** → Visualization project (dataset: sales, weather, or COVID-19).

Week 3 (6–10 Oct): Statistics for Data Science

Goal: Build statistical foundation.

- **Mon 6 Oct** → Descriptive stats (mean, median, variance, std).
- **Tue 7 Oct** → Probability distributions (normal, binomial, uniform).
- **Wed 8 Oct** → Inferential stats (sampling, confidence intervals).
- **Thu 9 Oct** → Hypothesis testing (t-test, chi-square).
- **Fri 10 Oct** → Stats project (analyze survey/real dataset with insights).

Week 4 (13–17 Oct): Exploratory Data Analysis (EDA)

Goal: Learn real-world data analysis.

- **Mon 13 Oct** → EDA process overview.
- **Tue 14 Oct** → Handling categorical vs numerical data.
- **Wed 15 Oct** → Feature engineering (encoding, scaling).
- **Thu 16 Oct** → Correlations & relationships.
- **Fri 17 Oct** → EDA project (Titanic / Netflix / Sales dataset).

Week 5 (20–24 Oct): Machine Learning Foundations

Goal: Learn supervised ML basics.

- **Mon 20 Oct** → ML workflow: train/test split, overfitting, metrics.
- **Tue 21 Oct** → Regression (Linear Regression).
- **Wed 22 Oct** → Classification (Logistic Regression).
- **Thu 23 Oct** → KNN & Decision Trees.
- **Fri 24 Oct** → ML mini project (predict housing prices/classify customers).

Week 6 (27–31 Oct): Advanced ML & Final Project

Goal: Practice advanced methods & complete a portfolio project.

- **Mon 27 Oct** → Ensemble models (Random Forest, Gradient Boosting).
- **Tue 28 Oct** → Unsupervised learning (Clustering – KMeans, PCA).
- **Wed 29 Oct** → Model evaluation (confusion matrix, ROC, precision/recall).
- **Thu 30 Oct** → Final Project (end-to-end: cleaning → EDA → ML model).
- **Fri 31 Oct** → Final Project Presentation (insights + visualization).