

Python ML: Linear Regression, Ridge & Lasso

Learn to build and tune linear regression models in Python, from data preparation through diagnostics to Ridge and Lasso regularization.

2 Days

DURATION

6

MODULES

16

TOPICS

14

HANDS-ON ACTIVITIES

COURSE OVERVIEW

This course begins with Python and pandas basics for handling data, then covers framing regression problems and evaluating models with metrics such as MAE, RMSE and R^2 . It moves into simple and multiple linear regression, including feature scaling and encoding, before addressing model diagnostics through residual plots and outlier handling. The course then introduces feature engineering and polynomial regression for capturing non-linearity, finishing with regularization techniques such as Ridge, Lasso and Elastic Net alongside cross-validation for hyperparameter selection.

WHO THIS COURSE IS FOR

Data Analysts

Machine Learning Engineers

Python Developers

Data Scientists

PREREQUISITES

- Knowledge of Python programming

LEARNING OBJECTIVES

- Apply pandas to clean, filter and summarise tabular data
- Explain the difference between regression and classification tasks
- Evaluate model performance using MAE, RMSE and R^2
- Build simple and multiple linear regression models
- Interpret regression coefficients and encode categorical features
- Identify overfitting and underfitting from residual plots
- Develop polynomial and interaction features for non-linear patterns
- Apply Ridge and Lasso regularization with cross-validation

6 Modules, 16 Topics

DAY 1

01**3 Activities**

Python & Data Basics (for ML)

TOPICS

- **Python quick tour:** variables, lists/dicts, functions, importing libraries
- **Jupyter workflow:** cells, markdown, plotting
- **pandas essentials:** read_csv, filtering, joins/merges, groupby, missing values

02**Activity**

Problem Framing & Metrics

TOPICS

- Regression vs classification; targets and features
- Train/test split; data leakage basics
- **Metrics:** MAE, RMSE, R^2 (intuition and when to use)

03**5 Activities**

Simple & Multiple Linear Regression

TOPICS

- Linear regression idea; assumptions (at a basic level)
- One-feature vs multi-feature models; coefficient meaning (direction & magnitude)
- Feature scaling when needed; one-hot encoding for simple categories (e.g., crude_type)

DAY 2

04**Activity**

Model Diagnostics (Basics)

TOPICS

- Residual plots; under/over-fitting intuition
- Simple outlier handling strategies (cap, remove, or log-transform target/feature)

05

3 Activities

Feature Engineering & Polynomial Regression

TOPICS

- Creating simple derived features: ratios, moving averages, interactions
- Polynomial features (degree 2) for gentle non-linearity

06

Activity

Regularization & Model Selection

TOPICS

- Why regularize: bias–variance trade-off
- Ridge vs Lasso vs Elastic Net (basic intuition)
- Cross-validation & GridSearchCV; picking hyperparameters simply

HANDS-ON ACTIVITIES

Practise What You Learn

#	MODULE	ACTIVITY
01	Python & Data Basics (for ML)	LAB 1 (REFINERY—BASIC) ■ Load a small CSV of daily refinery operations ■ Columns: date, crude_feed_tph, column_temp_C, ambient_temp_C, steam_tph ■ Tasks: handle missing values, create a clean DataFrame, quick summary stats and plots (line/scatter)
02	Problem Framing & Metrics	Lab 2: Split Day-1 dataset into train/test; compute MAE/RMSE/R ² baselines (mean predictor)
03	Simple & Multiple Linear Regression	LAB 3 ■ Use Case A — Predict Steam Usage ■ Target: steam_tph ■ Features: crude_feed_tph, column_temp_C, ambient_temp_C ■ Fit LinearRegression, evaluate on test set, interpret coefficients (e.g., higher feed ⇒ more steam) ■ Stretch: Add crude_type (3 categories) via one-hot encoding; re-evaluate
04	Model Diagnostics (Basics)	Lab 4: Make a residual plot for Use Case A; check if any single variable dominates error

#	MODULE	ACTIVITY
		LAB 5 (REFINERY)
05	Feature Engineering & Polynomial Regression	■ Use Case B — Predict Diesel Output ■ Dataset with columns: crude_feed_tph, column_temp_C, reflux_ratio, diesel_output_tph ■ Add polynomial/interaction features (e.g., feed×temp, temp²), compare to plain linear model
06	Regularization & Model Selection	Lab 6: Refit Use Case B with Ridge and Lasso; choose alpha via cross-validation; compare metrics

WHAT IS INCLUDED

Everything You Need to Succeed

- Koenig course material for every module
- Delivery by certified instructors and subject matter experts
- Hands-on exercises throughout the course
- Course completion certificate from Koenig

ABOUT KOENIG

Who We Are

Established in 1993, Koenig Solutions is a globally recognised training provider with 35+ technology partnerships, 5,000+ training programmes and a network of certified instructors and subject matter experts, supporting individuals and enterprise teams worldwide.

FLEXIBLE DELIVERY

Learn the Way That Suits You

Classroom Training

Live Online Classes (ILO)

Fly-Me-A-Trainer (FMAT)

Flexi (Self-Paced)

1-on-1 Training

Ready to Get Started?

Talk to a learning advisor about schedules, group training and customised delivery.

Visit koenig-solutions.com

✉ info@koenig-solutions.com