

# Student Homework Sheet — Stage 10b

---

## Assignment

Create lag and rolling features on a time series you gather (or generated from the homework-starter notebook) (in your project you will have an opportunity to something like this to your project dataset), then fit **either**:

- A basic **time series model** to forecast next-step returns, **or**
- A basic **classifier** to predict next-step direction (up/down).

## Requirements

1. Implement at least **two** new features from: `lag_k`, `rolling_mean`, `rolling_std`, `rolling_min/max`, `momentum`, or `zscore`.
2. Use a **time-aware split** for time series, or standard split for non-time classification.
3. Build a **sklearn Pipeline** (preprocessing → model).
4. Evaluate:
  - If forecasting: MAE/RMSE and a prediction vs truth plot.
  - If classification: accuracy, precision, recall, F1, confusion matrix.
5. Write a short **interpretation**: what works, what fails, where assumptions might break.

Chain statement: **In the lecture, we learned** how to create lag/rolling features, build pipelines, and evaluate with appropriate metrics. **Now, you will adapt** those patterns **to your own dataset** to accomplish a validated time series or classification baseline.

## Step-by-Step

1. Load/prepare your dataset (ensure DateTime index).
2. Engineer features (avoid leakage: only use past info).
3. Create target:
  - Forecasting: next-step return `y = ret.shift(-1)` (drop tail).
  - Classification: `y_up = (ret.shift(-1) > 0).astype(int)`.
4. Split data:
  - Time series: most recent 20–30% as test (or `TimeSeriesSplit`).
  - Classification (non-time): standard `train_test_split`.
5. Build pipeline:
  - Example: `Pipeline([('scaler', StandardScaler()), ('clf', LogisticRegression())])`

6. Fit, predict, evaluate; make plots.

7. Write markdown interpretation.

## Rubric (100 pts)

- (20) Feature engineering correctness (no leakage)
- (20) Pipeline implementation
- (20) Appropriate split strategy
- (20) Metrics & plots correctness
- (20) Interpretation quality (risk-aware reasoning)

## Submission

- Save as `notebooks/modeling_<team>.ipynb`
- Commit by next session.

## Example Expectations

- Clean notebook with sections: Data → Features → Target → Split → Pipeline → Metrics → Interpretation.
- At least one figure showing results.