Yield Curve Regime Clustering & Interest Rate Forecasting

Sajan Arora, Karan Badlani
Khoury College of Computer Sciences,
Northeastern University, Boston, MA, USA
arora.saj@northeastern.edu, badlani.k@northeastern.edu

November 19, 2025

1 Introduction

The U.S. Treasury yield curve summarizes interest rates across multiple bond maturities, such as 3-month, 2-year, 10-year, and 30-year. Its shape is a widely used indicator of macroeconomic conditions. A normal upward–sloping curve typically reflects expectations of growth and moderate inflation, whereas a flat or inverted curve—when short rates exceed long rates—has often preceded U.S. recessions.

Although practitioners commonly discuss the yield curve in terms of "normal" or "inverted", the curve in reality moves through a continuum of shapes. This motivates a data–driven characterization of yield curve regimes: typical shapes inferred directly from historical data rather than imposed a priori. In addition, the evolution of these regimes over time, and of spreads such as the 10-year minus 2-year yield, may carry predictive information about future economic conditions.

Project goal. Our goal is to (1) construct a weekly panel of U.S. Treasury yields across key maturities, (2) identify yield curve regimes via unsupervised clustering, and (3) explore simple forecasting models for yields and/or regime transitions built on this representation. This report documents our data collection, preprocessing, and exploratory data analysis (EDA), and outlines the modeling plan for the remainder of the project.

Research questions.

1. What distinct yield curve regimes emerge when we cluster weekly yield curve snapshots?
2. How do these regimes evolve over time, and how do they relate to known macroeconomic episodes?
3. Do spread dynamics (e.g., 10Y–3M, 10Y–2Y) provide useful early warning signals for regime changes or stress periods?

2 Data and Pre-processing

2.1 Data sources

We use publicly available U.S. Treasury constant-maturity yields from the Federal Reserve Economic Data (FRED) service of the Federal Reserve Bank of St. Louis.[1] The series provide fixed-maturity yields constructed from underlying Treasury securities. We work at weekly frequency to smooth daily noise while preserving medium-term dynamics.

Table 1 summarizes the maturities and FRED series codes currently used.

We download all series via the FRED API, align them to a common weekly date index, and restrict to the intersection of dates where all maturities are available. The final sample spans roughly 1981–2025, yielding about 2300 weekly observations per maturity.

[1] `https://fred.stlouisfed.org/`

Table 1: Treasury constant-maturity yield series used in this project.

Maturity	Description	FRED Code
3-month	Treasury bill	TB3MS
2-year	Treasury yield	DGS2
5-year	Treasury yield	DGS5
10-year	Treasury yield	DGS10

2.2 Baseline panel and spreads

We begin from a weekly panel of Treasury yields that we constructed in earlier steps and saved as `treasury_yields_weekly.csv`. For the current analysis, we focus on a baseline set of four maturities,

$$\text{COLS\_BASE} = \{y\_3m, y\_2y, y\_5y, y\_10y\},$$

which are mostly complete over the sample. Reading this file and dropping rows with missing values yields a baseline DataFrame `df_base` of shape $(2306, 4)$, covering the period from 4 September 1981 to 7 November 2025. We also construct spread series such as 10Y–3M and 10Y–2Y and save the cleaned baseline to `Data/interim/yields_baseline.csv` for reuse in downstream steps.

2.3 Pre-processing pipeline

Pre-processing follows the steps below, consistent with our data collection notebooks:

1. **Download and align:** All FRED series are downloaded separately and aligned on a common weekly date index.
2. **Select baseline maturities:** For modelling the yield curve shape, we focus on the four most complete maturities:
$$\text{COLS\_BASE} = \{y_{3m},\, y_{2y},\, y_{5y},\, y_{10y}\}.$$
These maturities span over 2,300 weekly observations with minimal missing data.
3. **Clean and filter:** Rows with missing values among the four baseline maturities are dropped, producing a cleaned panel of shape $(2306, 4)$ covering 1981–2025.
4. **Spread construction:** We compute key slope measures such as
$$s_{10y,3m} = y_{10y} - y_{3m}, \qquad s_{10y,2y} = y_{10y} - y_{2y},$$
which track inversion episodes. Additional shape features $(s_{2y,3m}, s_{5y,3m}, s_{10y,3m})$ are constructed for clustering.

The resulting baseline dataset is saved as `yields_baseline.csv` in the `Data/interim/` directory. It contains the four main maturities and derived spread features.

Figure 1 displays a sample of the cleaned weekly panel used for exploratory analysis and downstream clustering.

date	y_3m	y_1y	y_2y	y_5y	y_10y	y_30y	spr_10y_3m	spr_10y_2y
2025-10-24	3.95	3.57	3.46	3.58	4.00	4.57	0.05	0.54
2025-10-31	3.90	3.66	3.55	3.67	4.06	4.61	0.16	0.51
2025-11-07	3.95	3.67	3.59	3.71	4.12	4.70	0.17	0.53

Figure 1: Sample of the cleaned weekly dataset. Each row corresponds to one week; columns contain yields at different maturities and spread features

3 Exploratory Data Analysis

3.1 Correlation structure

To understand the joint behaviour of yields across maturities, we compute the correlation matrix of weekly yields. As shown in Figure 2, the correlations between maturities are very high (roughly 0.94–1.00). Short-term and long-term rates move closely together, consistent with a smooth yield curve that shifts and tilts over time rather than moving erratically at individual maturities.

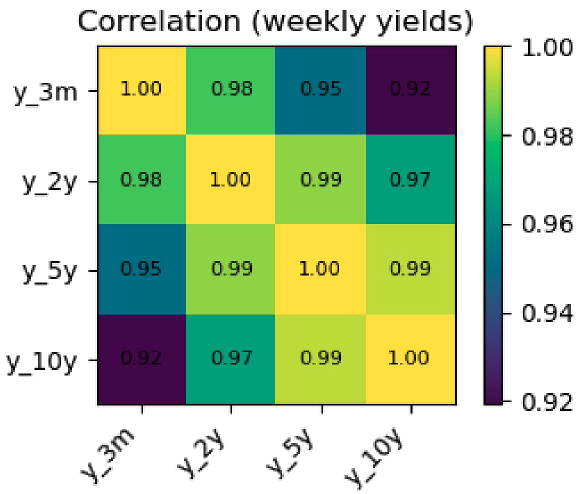

Figure 2: Correlation matrix of weekly yields across maturities. Strong positive correlations support the use of low-dimensional representations (e.g., principal components) for the yield curve.

The strong correlation structure justifies dimensionality reduction techniques such as principal component analysis (PCA), which can capture most of the variation in a small number of factors.

3.2 Yield level dynamics

Figure 3 plots the weekly yields for 3M, 2Y, 5Y, and 10Y maturities over the sample period. All maturities share similar long-run patterns: yields were high and volatile in the early 1980s, declined over subsequent decades, and exhibit several distinct interest-rate cycles associated with monetary tightening and easing episodes.

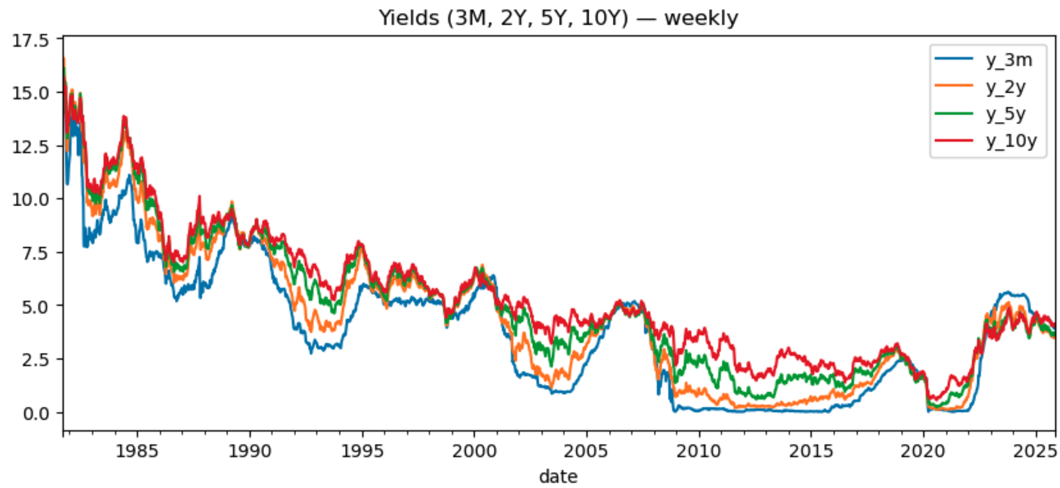

Figure 3: Weekly yields for selected maturities (3M, 2Y, 5Y, 10Y). All maturities show similar cyclical behaviour and a long-term downward trend since the early 1980s.

3.3 Spread dynamics and inversions

We pay particular attention to the 10Y–2Y and 10Y–3M spreads, which turn negative when the curve inverts. Figure 4 shows the 10Y–2Y spread over time. Periods where the spread is below zero correspond to inversions, many of which line up with subsequent recession periods documented in the macroeconomic literature.

Using the preprocessed dataset, the 10Y–3M spread is negative only in a minority of weeks. This confirms that inversions are relatively rare events but occur in distinct clusters.

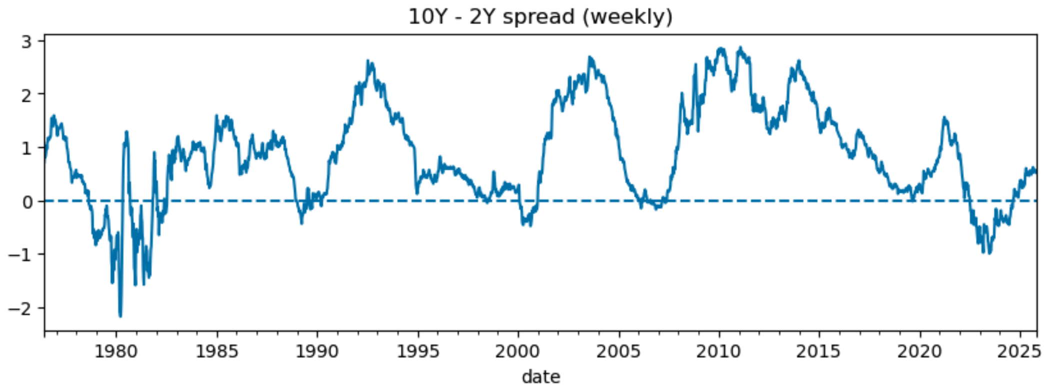

Figure 4: Weekly 10Y–2Y spread. Negative values (below the dashed line) indicate yield curve inversion and tend to appear before recessionary periods.

These EDA results validate the economic structure of the dataset and highlight the potential value of spreads as features for regime identification and forecasting.

3.4 Shape features and principal components

To isolate curve *shape* from level, we define spreads vs the 3M rate:

$$s\_2y, 3m = y\_2y - y\_3m, \quad s\_5y, 3m = y\_5y - y\_3m, \quad s\_10y, 3m = y\_10y - y\_3m.$$

The resulting `shape_df` has 2,306 weekly observations. Summary statistics show that the average slope is positive (e.g., $E[s\_10y, 3m] \approx 1.55$ percentage points), but the minimum values around -1.8 confirm that strong inversions occur.

We also run PCA on standardized raw yields. The first three components explain approximately 97.5%, 2.3% and 0.15% of the variance, respectively. The loading patterns match the classic level–slope–curvature interpretation: PC1 loads roughly equally on all maturities, PC2 contrasts short vs long rates (slope), and PC3 captures small curvature effects. We save the corresponding factor scores as `PC1_level`, `PC2_slope` and `PC3_curv` for use as features.

4 Methods: Clustering Yield Curve Regimes

This section describes the preliminary methods implemented so far. Forecasting model development will be completed in the final report.

4.1 K-means on slope-based shape features

To identify yield curve regimes, we cluster the shape-only features ($s\_2y, 3m, s\_5y, 3m, s\_10y, 3m$). After standardizing these three spreads, we run K-means with $k \in \{3, 4, 5\}$ and evaluate the silhouette score. The best value is obtained at $k = 3$ (silhouette ≈ 0.449, compared to 0.398 and 0.382 for $k = 4, 5$), so we retain three clusters for our initial regime definition.

We then invert the standardization to express cluster centroids back in spread units. Table 2 reports the typical spreads vs 3M for each cluster.

Table 2: Cluster centroids in spread space (percentage points).

Cluster	$s\_2y, 3m$	$s\_5y, 3m$	$s\_10y, 3m$
0 (normal)	0.56	1.06	1.56
1 (flat/inverted)	-0.20	-0.26	-0.13
2 (steep)	1.32	2.27	2.89

Cluster 0 corresponds to a **normal** positively sloped curve, cluster 2 to a **steep curve** with very large spreads, and cluster 1 to a **flat/inverted** configuration where long rates are only slightly above, or below, the 3M rate. Figure 5 visualizes the centroid shapes.

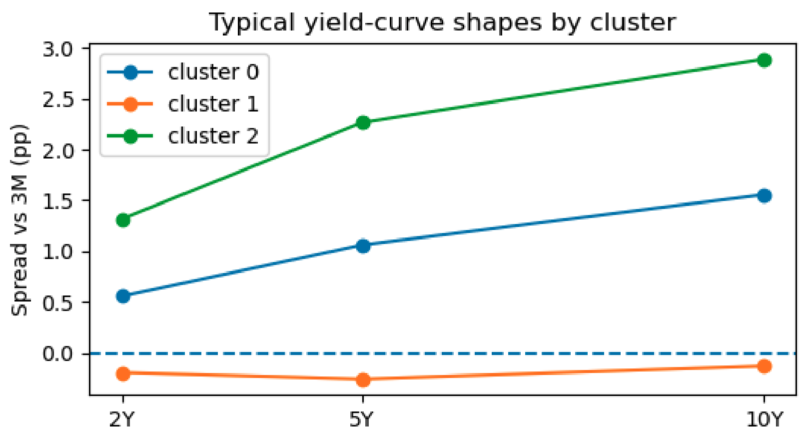

Figure 5: Typical yield curve shapes by cluster, plotted as spreads vs 3M at 2Y, 5Y and 10Y.

4.2 Regime labels, inversion rates and persistence

Using these clusters, we label each week as belonging to a `normal`, `steep` or `flat_inverted` regime based on the ordering of centroid slopes. Joining the cluster labels with the 10Y–3M spread allows us to compute inversion frequencies by regime. The flat/inverted cluster accounts for 541 weeks, with an inversion share of roughly 48%, whereas the normal and steep clusters are almost never inverted.

We also construct a regime timeline by shading the 10Y–3M spread series by cluster and compute a 3×3 transition matrix and duration statistics. All three regimes have similar average durations (around 19–22 weeks), but the maximum realized run lengths can be substantial (up to 160 weeks for flat/inverted and 394 weeks for normal), indicating that regimes can persist for long periods once entered.

5 Feature Engineering and Modelling Dataset

To support downstream forecasting tasks, we construct a rich weekly feature set:

- **Slope and curvature:** 10Y–3M and 10Y–2Y slopes, butterfly spreads such as $(y\_10y - 2y\_5y + y\_3m)$ as curvature proxies.
- **Volatility and momentum:** First differences, rolling realized volatility over 4, 13 and 26 weeks, moving averages and rate-of-change features for slopes and curvature.
- **Regime-aware features:** Current regime label, one-hot indicators for each regime, and a running count of *weeks in current regime.*
- **Factor-based features:** PCA factors `PC1_level`, `PC2_slope`, `PC3_curv` and their 13-week differences, volatility and moving averages.

- **Inversion timing features:** Distance to the inverted vs steep centroids, a continuous "regime score", weeks since last zero-crossing of the slope, and rolling counts of weeks spent inverted over the past 26 and 52 weeks.

We then lag all predictors by one week to avoid look-ahead bias and define a 12-week forecasting horizon $h = 12$. The target variables are:

- future 10Y yield level $y\_10y, t + h$,
- future slope $(y\_10y - y\_3m)\_t + h$,
- a binary indicator of whether the curve will be inverted in h weeks,
- the next regime label at horizon h.

Joining lagged predictors with these targets and dropping rows with missing values yields a modelling table `model_table_h12` with roughly 2,242 weekly observations and 25 columns. Before training, we apply a simple correlation filter to drop highly collinear predictors (absolute pairwise correlation > 0.95), removing three redundant features.

6 Preliminary Results

So far, we have moved beyond pure data collection and built several concrete artifacts:

- A cleaned baseline panel of weekly yields for 3M, 2Y, 5Y and 10Y maturities from 1981–2025, and spread series.
- Evidence that yields are highly correlated across maturities and that spreads display the expected inversion episodes.
- A three-regime clustering of yield curve shapes (normal, steep, flat/inverted), supported by silhouette analysis and interpretable centroids.
- Regime diagnostics, including a transition matrix and duration statistics, showing that regimes are persistent but occasionally switch for extended periods.
- A comprehensive feature set has been constructed—covering slope, curvature, volatility, momentum, regime indicators, and PCA factors—along with a 12-week forecasting target table that will serve as the foundation for regression and classification models in the next stage.

These results confirm that our pipeline successfully transforms raw FRED yields into a structured representation of yield curve regimes and a rich modelling dataset. The next phase of the project will focus on training and evaluating forecasting models that exploit these features.

7 Next Steps

For the remaining of the project, we will:

- train forecasting models for the 10Y yield and 10Y–3M slope using the engineered weekly features,
- build classification models to predict yield curve inversion and regime transitions at a 12-week horizon,
- evaluate performance using MAE/RMSE (regression) and accuracy/ROC-AUC (classification),
- compare models with and without regime-aware features to assess added value,
- prepare visual and quantitative results for the final project report.

References

[1] Federal Reserve Bank of St. Louis (FRED). Treasury Constant Maturity Rates. https://fred.stlouisfed.org/

[2] Estrella, A., & Mishkin, F. S. (1996). *The Yield Curve as a Predictor of U.S. Recessions*. Federal Reserve Bank of New York. https://web.archive.org/web/20150908020815/https://www.newyorkfed.org/research/current_issues/ci12-2/ci12-2.html

[3] Diebold, F. X., & Rudebusch, G. D. (2013). *Yield Curve Modeling and Forecasting*. NBER Working Paper No. 16714. https://www.nber.org/papers/w16714