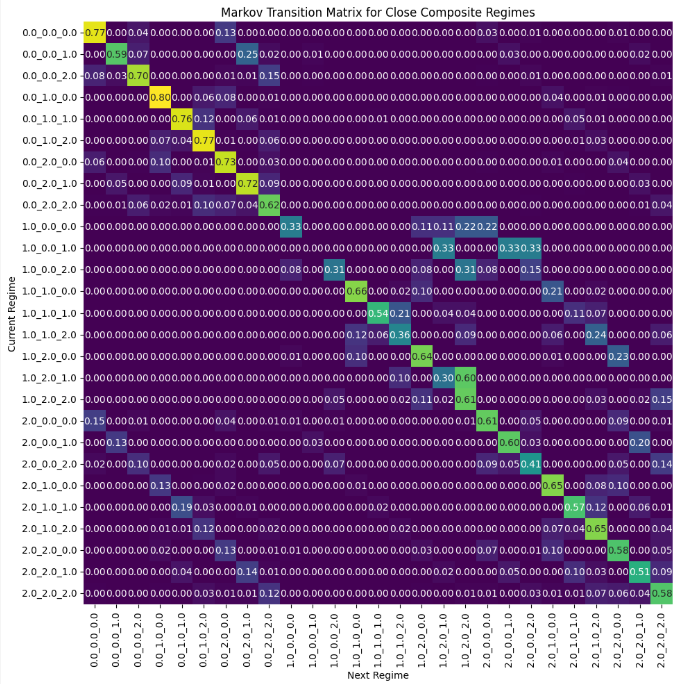
**VIX Regimes**

GARCH (volatility clustering), Permutation Entropy (chaos), Hurst’s Exponent (trend persistence), K-Means Clusters, and Markov Transitions.

[GuardsGuards/VIX-Regimes](https://github.com/GuardsGuards/VIX-Regimes/tree/main)

****

**Abstract**

April 4, 2025 marked a historic surge in the VIX to $45, an almost 200% markup from just over a week prior. Estimating the expected volatility of the S&P 500 Index, the VIX aggregates the price of 30-day SPX calls and puts over a wide number of strike prices. It is *the* augur of financial fear and — when concurrent with a one-day $5 trillion wipeout — loathing. And for any self-respecting risk manager, it should also be the object of occasional piety.

This study seeks to illume by classification, VIX regimes. It establishes the baseline regime, and by application of Markov transitions, assigns probabilities to the deterioration of one regime into another and/or the persistence of a regime onto itself. The regimes are characterized by three quantitative metrics: volatility bunching derived from generalized autoregressive conditional heteroskedasticity (GARCH), complexity assessed through Permutation Entropy (PE), and trend persistence evaluated via the Hurst Exponent (HE). These indicators were chosen for their immense explanatory power, exhibition of negligible multicollinearity, measured by way of the variance inflation factor test (VIF), and resilient cointegration, assessed through the Johansen test. Each indicator is then binned into three distinct classes using K-means clusters and concatenated for a composite regime.

Cursory implications for traders and portfolio managers are as follows: for high GARCH and persistent risk, hedge and reduce exposure; during high entropy states expect market chaos by considering straddles and flexible strategies; and while high Hurst regimes imply trending behavior, low Hurst’s are mean-reverting and investors should deploy contrarian/mean-reversion strategies. Intriguing are composite regimes that, while slightly less common, exhibit high relative stability and strong persistence from one day to the next. The combination of relative rarity and predictability should be especially interesting to market participants and warrants further evaluation.

**Indicators**

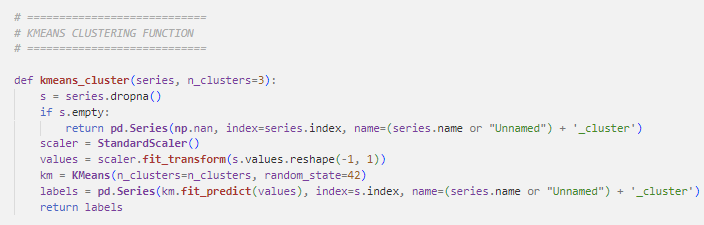
**Modeling Volatility, Complexity, and Persistence**

* **Volatility:** GARCH models are employed on price and return to assess volatility persistence and shock impacts. A low value (e.g., 0.0) suggests a relatively calm or mean‐reverting return volatility, while a high value (e.g., 2.0) would indicate persistent, high volatility.
* **Complexity:** Permutation entropy is used to gauge unpredictability, chaos, and disorder. Lower bins (e.g., 0.0) indicate lower entropy (i.e., more predictable or orderly behavior), while higher bins (e.g., 2.0) indicate greater complexity and unpredictability.
* **Trend Persistence:** Hurst exponents provide insight into the trending behavior or mean reversion properties of the yield curve. This reflects the Hurst exponent on returns, which indicates whether the return series is trending (values > 0.5) or mean-reverting (values < 0.5). After clustering, a low bin (e.g., 0.0) might represent a mean-reverting regime, and a higher bin (e.g., 2.0) could indicate trending behavior.

**Calculations**

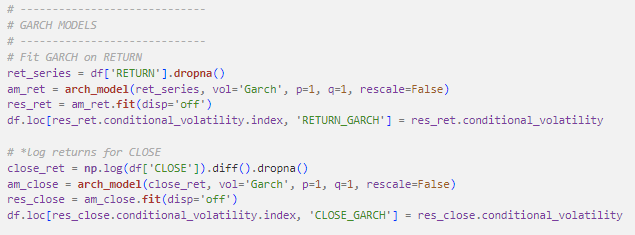
**K-Means Clustering:** is an unsupervised learning algorithm that partitions a dataset into K distinct clusters. The goal is to assign each data point to one of the clusters in such a way that the total intracluster variance is minimized (via the sum of squared distances (errors)), effectively creating clusters where data points are as close to each other as possible. The objective function for K-means is defined as:

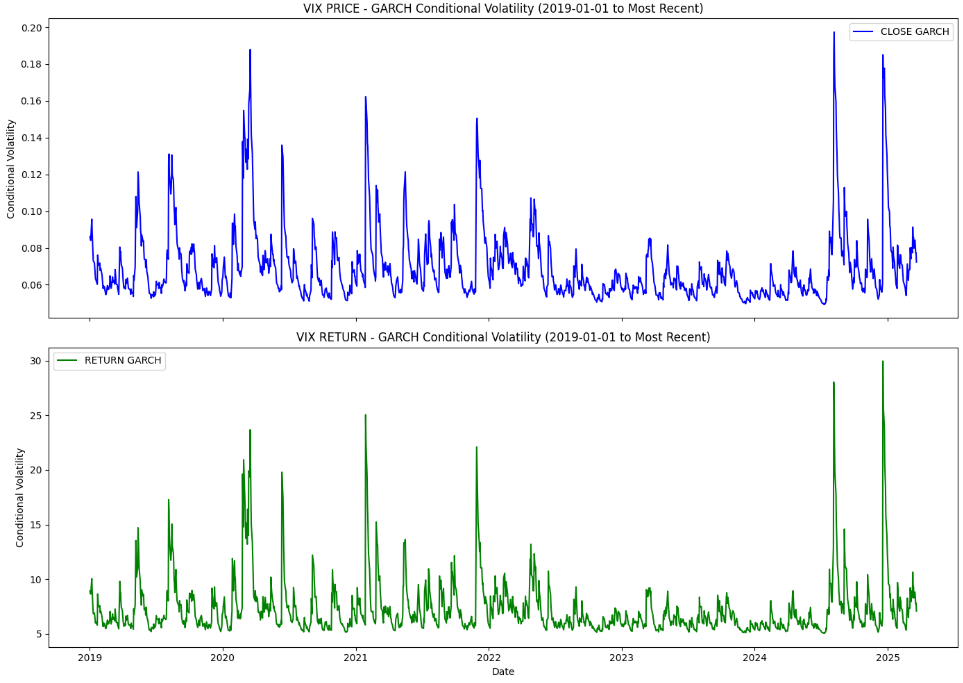
* + and
* is the squared distance between data point and the centroid



**GARCH (Generalized Autoregressive Conditional Heteroskedasticity):** captures the clustering of volatility over time, indicating that shocks to returns tend to persist. A standard GARCH(1,1) model is defined as:

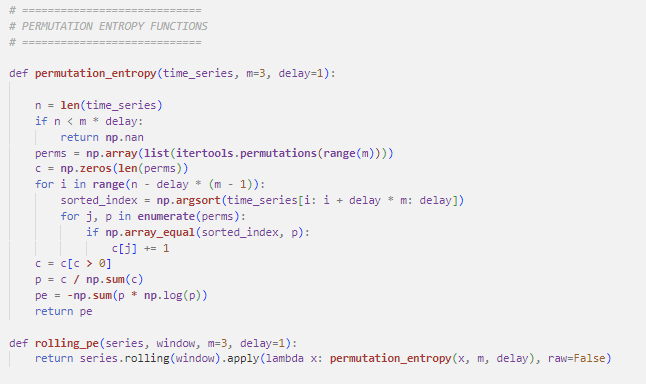
* + (constant term): This parameter represents the long-run average level of variance in the absence of recent shocks. It ensures that the conditional variance remains positive even when past shocks or volatilities are small. Typically, ω is required to be positive (ω > 0), and its magnitude influences the baseline volatility level in the model,
* α **(ARCH term):** The parameter α measures the immediate impact of new information (or shocks) on volatility. It multiplies the squared residual ​, so a large α indicates that recent shocks have a strong effect on current volatility. This term captures the “news effect,” where unexpected events lead to a spike in volatility. In practice, α is constrained to be non-negative (α ≥ 0).
* **(GARCH term):** The parameter β reflects the persistence of volatility from one period to the next by weighing the previous period’s conditional variance ​. A high β means that shocks to volatility die out slowly, indicating that volatility is highly persistent. As with α, β is generally constrained to be non-negative (β ≥ 0).

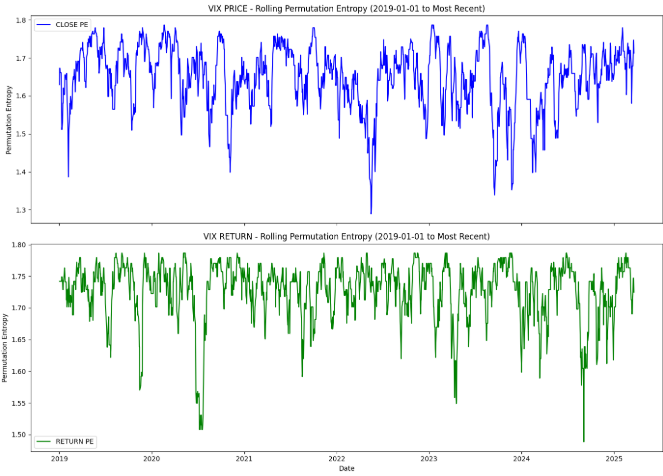


****

**Permutation Entropy (PE):** measures the complexity or unpredictability of a time series by analyzing the order of values in a series (with respect to their magnitudes).

* + For example, if d =3 and τ = 1, you would look at patterns formed by sequences like [,]. Each vector is then mapped to a permutation pattern based on the relative ordering of its elements (e.g., whether is the smallest, the middle, or the largest):
* there are d! possible distinct ordinal patterns, and
* is the probability of each permutation pattern. It is bounded between 0 and 1, where 0 indicates a completely predictable dynamic and 1 indicates a completely stochastic dynamic (i.e. more disorder and less predictability), which helps to identify regimes where market behavior is more chaotic.



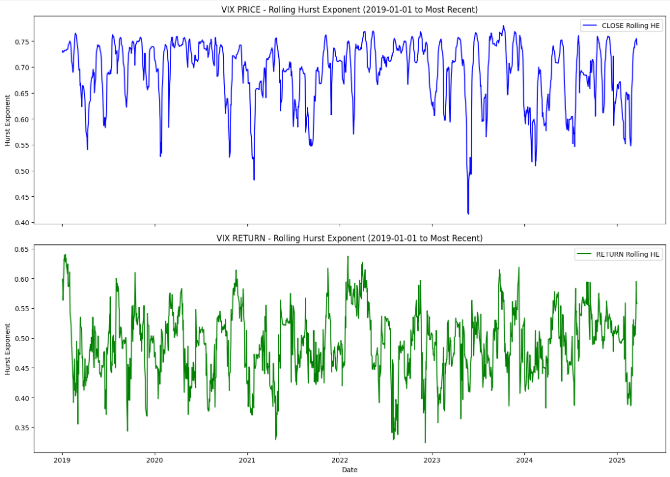
****

**Hurst’s Exponent:** quantifies the tendency of a time series to either persist (trend) or revert to the mean. It is derived from the scaled range analysis:

* + means that the expected range grows proportionally to n raised to the power H.

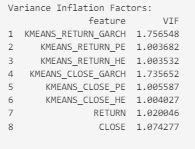
An value above 0.5 indicates persistence (trending behavior), while a value around 0.5 suggests randomness, and below 0.5 suggests anti-persistence.

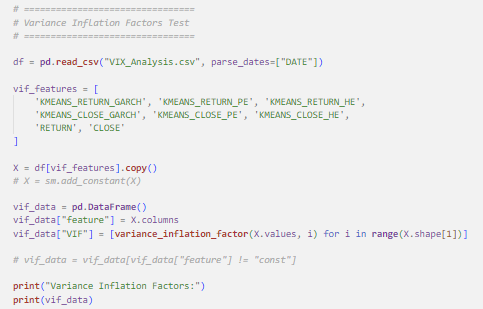


****

**Variance Inflation Factor (VIF):** quantifies the extent of multicollinearity among the predictors in a regression model. For a given predictor , the VIF is defined as:

* + where ​ is the coefficient of determination obtained by regressing on all the other predictors. A higher VIF indicates greater collinearity, suggesting that the predictor is highly correlated with the others, which can inflate the variance of the estimated coefficients. Typically, VIF values greater than 10 are cause for concern, however, values greater than 5 still deserve scrutiny.





**Cointegration**

Cointegration is a statistical property of a collection of time series variables that indicates the existence of a long-run equilibrium relationship among them, despite being individually non-stationary, given by:

The rank of the matrix determines the number of cointegrating relationships:

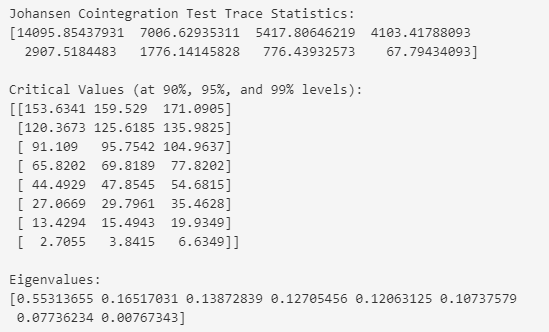
* + If rank () = 0, there is no cointegration.
  + If 0 < rank () = r < k, there are r cointegrating vectors, meaning that there exist r linear combinations of the variables that are stationary.

The Johansen cointegration test utilizes the trace statistic, which is given by:

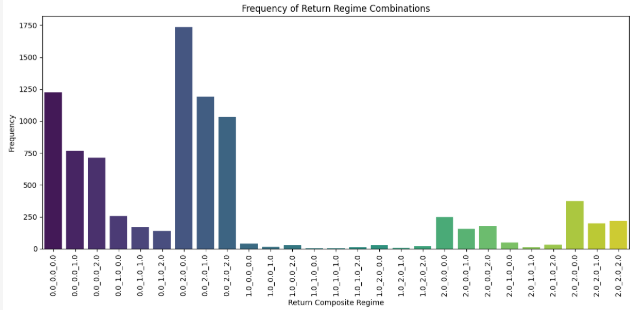
* and

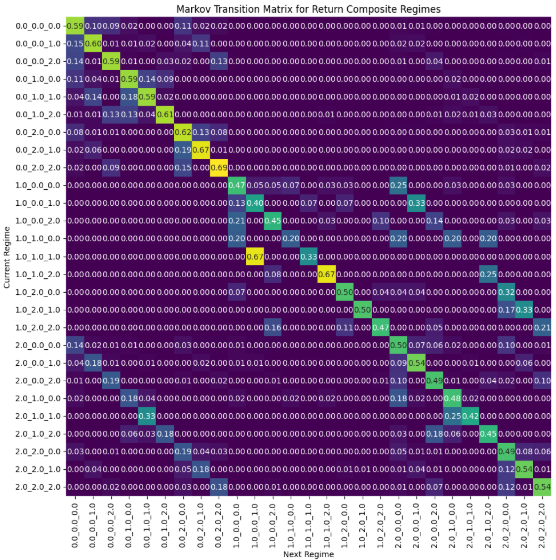
This trace statistic is compared against critical values to determine whether to reject the null hypothesis of no cointegration.

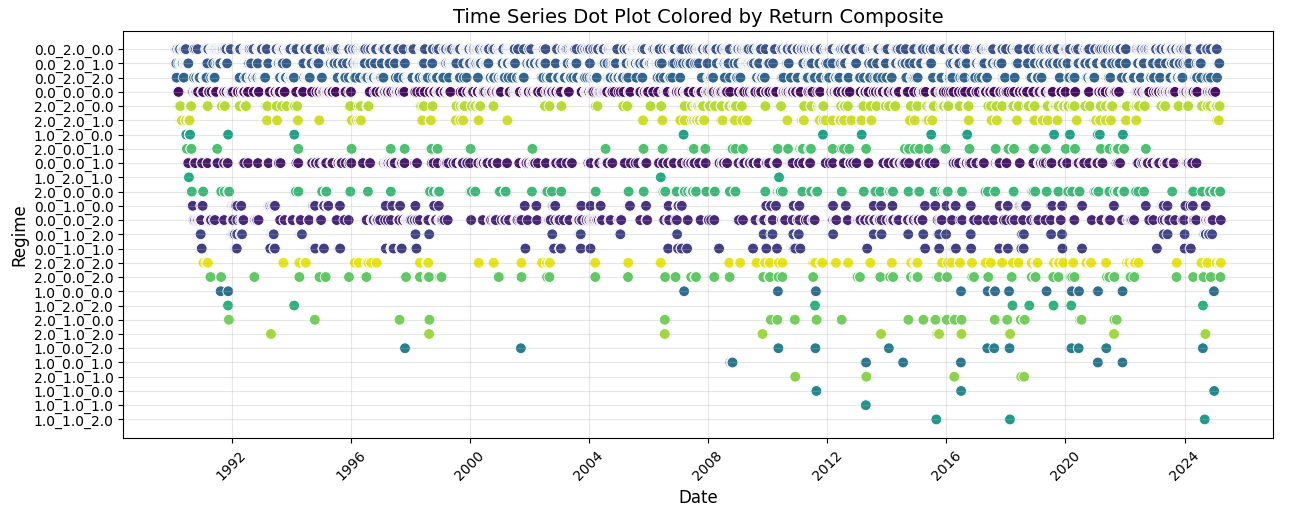


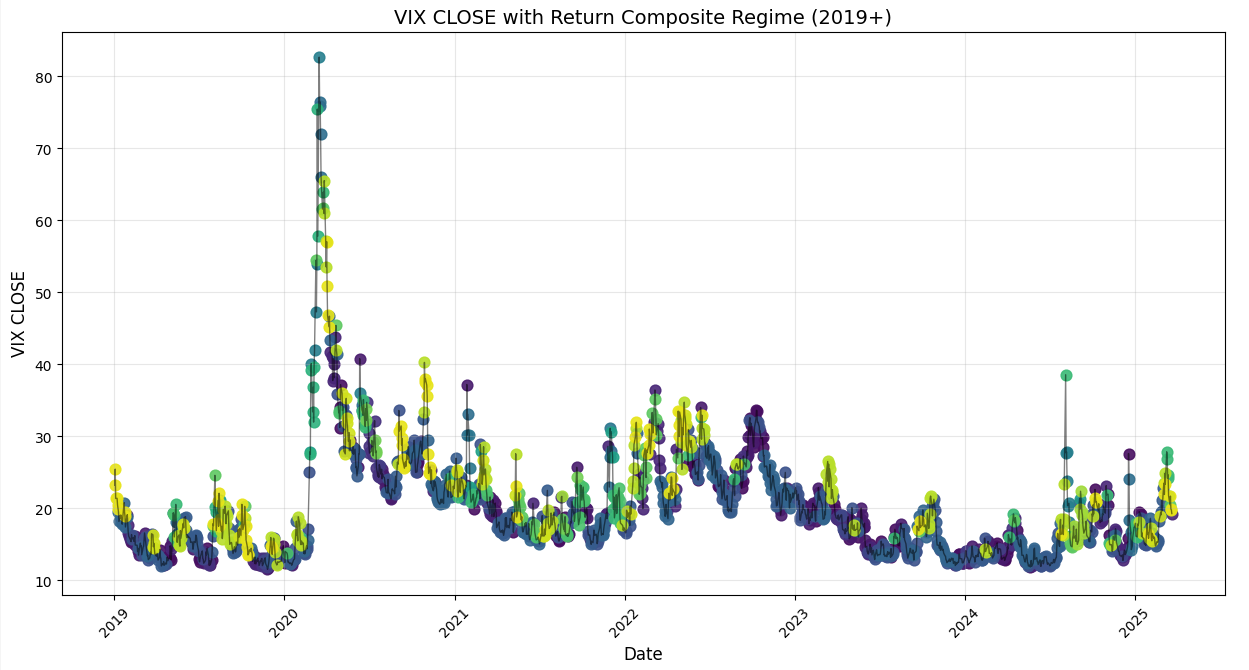


**Return Frequency Distribution Insights**

****

****

****

****

**1. Regime 0.0\_2.0\_0.0 (1,734 occurrences)**

* **Interpretation:**
  + **GARCH (0.0):** VIX returns are most often found in a state where the volatility clustering remains at its lowest level.
  + **PE (2.0):** There is a high level of price effect influencing the returns.
  + **HE (0.0):** There is little impact from extreme historical movements.
* **Stability:** Moderate, **0.62**
* **Implication for VIX:** This is the dominant, “baseline” state, suggesting that for a large portion of time, the VIX behaves in a relatively calm manner with only moderate price-driven changes and few extreme shocks.

**2. Regime 0.0\_0.0\_0.0 (1,226 occurrences)**

* **Interpretation:**
  + **GARCH (0.0):** Low volatility clustering.
  + **PE (0.0):** Minimal price effect influence.
  + **HE (0.0):** Low impact of historical extremes.
* **Stability:** Moderate, **0.59**
* **Implication for VIX:** This regime represents a state of deep calm — essentially, an “all low” scenario. VIX returns are subdued, suggesting extended periods of market complacency or stability with little reaction to price or extreme events.

**3. Regime 0.0\_2.0\_1.0 (1,192 occurrences)**

* **Interpretation:**
  + **GARCH (0.0):** Continues to be in a low volatility regime.
  + **PE (2.0):** Now at the highest price effect.
  + **HE (1.0):** Moderate influence from historical extremes.
* **Stability:** Second highest, **0.67**
* **Implication for VIX:** Even with a calm underlying volatility (GARCH), the pronounced HE factor implies that, at times, the VIX exhibits significant jumps or spikes driven by extreme market moves. This state could correspond to periods when isolated shocks (perhaps after a news event or market anomaly) cause temporary surges in volatility.

**4. Regime 0.0\_2.0\_2.0 (1,032 occurrences)**

* **Interpretation:**
  + **GARCH (0.0):** Remains low, maintaining the calm baseline in volatility.
  + **PE (2.0):** Now at the highest level, suggesting a resounding price effect.
  + **HE (2.0):** A high level of influence from historical extremes.
* **Stability:** Highest, **0.69**
* **Implication for VIX:** Here, the overall market conditions are calm except for the occurrence of strong, sporadic shocks. This state might capture instances where the market, generally quiet, is jolted by rare but significant event

**5. Regime 0.0\_0.0\_1.0 (796 occurrences)**

* **Interpretation:**
  + **GARCH (0.0) & PE (0.0):** Both at the lowest level.
  + **HE (1.0):** A moderate influence from past extremes.
* **Stability:** Moderate, **0.60**
* **Implication for VIX:** This is similar to the previous regime but with a less pronounced effect from historical extremes. It suggests that while the market remains largely in a low-reactivity state, there are occasional moderate deviations from past extremes.

**Return Frequency Distribution Conclusions**

**Dominance of a “0” GARCH Component**

**Low Volatility Clustering:**

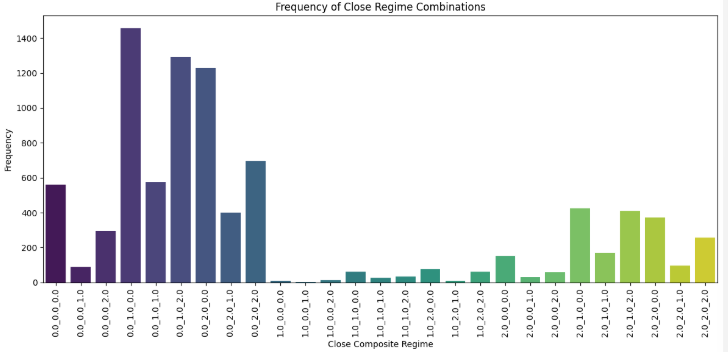
* Most popular states exhibit a “0” in the GARCH position. This indicates that VIX returns are generally characterized by low volatility clustering, meaning that past volatility doesn’t strongly persist into the future. In practical terms, VIX returns tend to revert quickly instead of sustaining prolonged periods of high volatility.
* **Frequent Calm Periods:** The market is usually in a regime where volatility isn’t self-reinforcing. Sharp spikes in volatility tend to be isolated rather than part of a sustained turbulent period.
* **Short-lived Shocks:** When shocks occur, their impact on volatility clustering dissipates quickly. The market either quickly absorbs the shock or reverts back to normalcy.

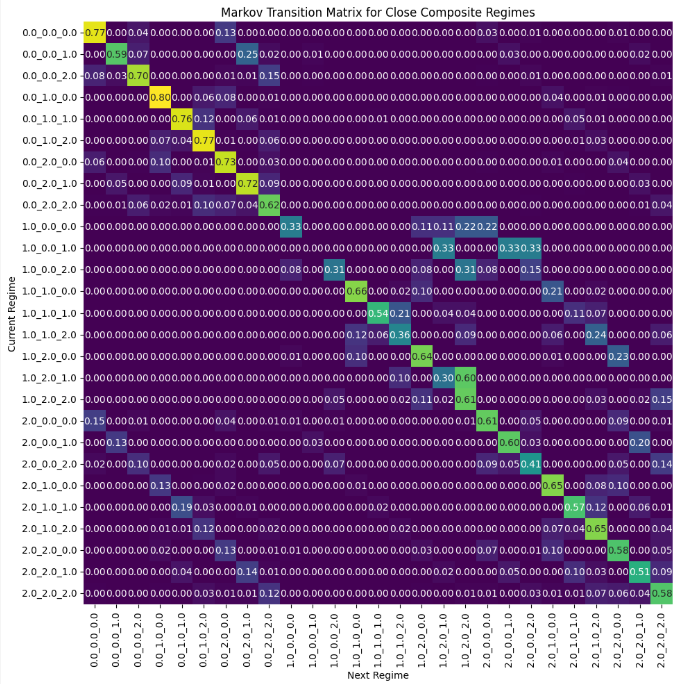
**Uncommon, Extreme Conditions**

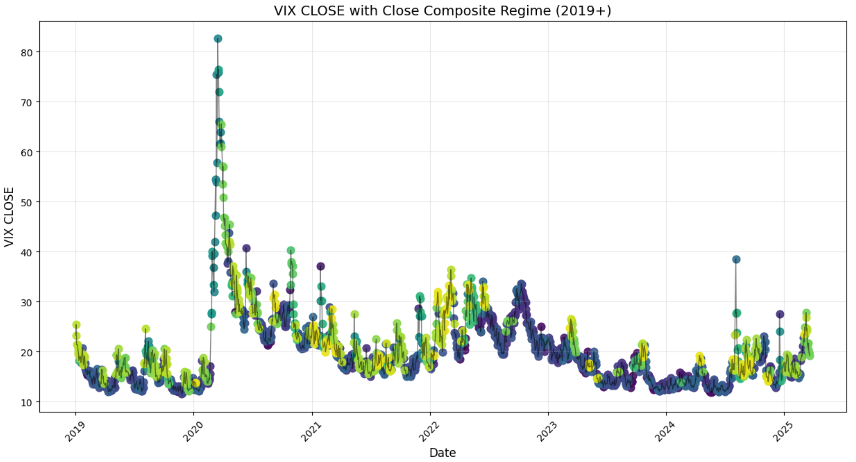
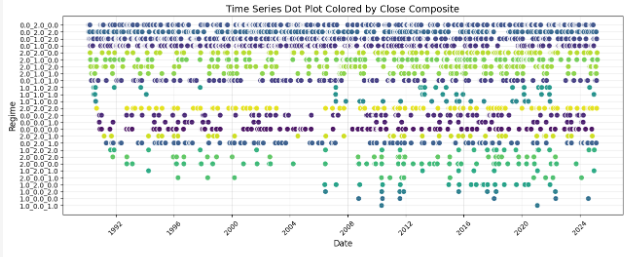
**High Volatility Clustering:**

* Regimes starting with “2” (high volatility clustering), especially when combined with a “2” in the PE component, are notably rare. These states likely occur during periods of market stress or extraordinary events.
* **Persistent Volatility:** Once volatility is triggered in these regimes, it tends to persist longer than in common states.
* **Extreme Price Reactions:** The market shows strong reactions to price movements, potentially amplifying the impact of shocks.

**Price Frequency Distribution Insights**

****

****

****

**1. Regime 0.0\_1.0\_0.0 (1,456 occurrences)**

* **Interpretation:**
  + **GARCH (0.0):** Low volatility clustering.
  + **PE (1.0):** Moderate price effect.
  + **HE (0.0):** Little to no influence from historical extremes.
* **Stability:** Highest, **0.80**
* **Implication for Close Prices:** In this state, the closing behavior is primarily driven by current price movements without the added influence of past extremes. It reflects a scenario of typical market conditions with a clean price signal and minimal memory of past shocks.

**2. Regime 0.0\_1.0\_2.0 (1,293 occurrences)**

* **Interpretation:**
  + **GARCH (0.0):** Indicates a low level of volatility clustering in the close prices.
  + **PE (1.0):** Suggests a moderate influence from price movements.
  + **HE (2.0):** Shows that past extreme events are having a moderate impact.
* **Stability:** Second highest, **0.77**
* **Implication for Close Prices:** This is the most common state, implying that on most days the market close is determined by a balanced mix of moderate price influences and the legacy of previous extremes — all while the overall volatility persistence remains low.

**3. Regime 0.0\_2.0\_0.0 (1,231 occurrences)**

* **Interpretation:**
  + **GARCH (0.0):** Low volatility clustering.
  + **PE (2.0):** A high level of price-driven influence.
  + **HE (0.0):** Low historical extreme impact.
* **Stability:** High, **0.73**
* **Implication for Close Prices:** In this state, the market is reacting strongly to current price movements while still being moderately affected by past extremes. The low volatility clustering suggests that the high price effect is not part of a prolonged turbulent phase, but rather a short‐term price shock influencing the close.

**4. Regime 0.0\_2.0\_2.0 (697 occurrences)**

* **Interpretation:**
  + **GARCH (0.0):** Very low volatility clustering.
  + **PE (2.0) & HE (2.0):** Very high price effect and influence from past extremes.
* **Stability:** Low, **0.62**
* **Implication for Close Prices:** This represents a baseline or “all low” state — a calm market environment where closing prices are not significantly driven by recent price changes or memory of past extremes, reflecting extended periods of market stability.

**5. Regime 0.0\_0.0\_0.0 (559 occurrences)**

* **Interpretation:**
  + **GARCH (0.0), PE (0.0), HE (0.0):** Very low volatility clustering, negligible price effect, and no discernible influence from past extremes.
* **Stability:** High, **0.77**
* **Implication for Close Prices:** This represents a baseline or “all low” state — a calm market environment where closing prices are not significantly driven by recent price changes or memory of past extremes, reflecting extended periods of market stability.

**Price Frequency Distribution Conclusions**

**Dominance of GARCH = 0 States**

* **Calm Conditions:** This indicates that for many closing days, the market exhibits minimal volatility clustering. The impact of past volatility quickly fades, suggesting that any shocks are short-lived.
* **Market Implication:** The market is typically in a stable regime where volatility does not self-reinforce, leading to rapid reversion back to normal levels after any transient disturbances.

**Dominance of GARCH = 2 States**

* **Persistent Volatility:** Here, closing day prices experience significant volatility clustering. Once volatility is triggered, it tends to persist, reflecting periods of ongoing market turbulence.
* **Market Implication:** These conditions likely correspond to times of market stress or instability where volatility is self-reinforcing, leading to prolonged periods of elevated risk.

**Scarcity of GARCH = 1 States**

* **Transitional/Moderate Regime:**  The relative rarity of GARCH = 1 indicates that the market doesn’t often reside in a transitional state between calm and turbulent conditions. Instead, it tends to switch directly to either a very calm state (GARCH = 0) or a highly volatile state (GARCH = 2)

**Bimodal Behavior:** The bimodal distribution suggests that the closing prices tend to oscillate between two distinct regimes:

* **Stable Regime (GARCH = 0):** Characterized by rapid volatility reversion and minimal persistence.
* **Turbulent Regime (GARCH = 2):** Marked by sustained volatility clustering and extended periods of instability.
* For traders and risk managers, this dichotomy is crucial. The stable, common state implies a normal, less risky market environment, whereas the turbulent state signals conditions that require heightened caution and potentially more conservative risk strategies.

**Implications for Traders and Portfolio Managers**

**Market Regime Identification:**

* **Frequent Regimes as "Normal" States:** The most common composite regimes reflect the market’s typical behavior, it represents a "normal" state, the baseline regime, where the market exhibits moderate disorder but tends to revert to the mean
* **Rare Regimes as Signals of Transition or Stress:** Regimes that occur infrequently — especially those with high values in the GARCH or Hurst components — might signal that the market is entering a stressed state. For example, a regime like "2.0\_1.0\_0.0" may indicate a sudden shift towards persistent high volatility, prompting a reevaluation of risk exposure or hedging strategies.

**Risk Management & Strategy Adjustment:**

* **Volatility Persistence (GARCH):** When the GARCH component is high (e.g., bin 2.0), it suggests that volatility is more persistent, which can lead to prolonged periods of market stress. In such cases, risk managers might tighten stop-loss levels, reduce position sizes, or increase hedging.
* **Market Chaos (Permutation Entropy):** A higher permutation entropy bin indicates greater disorder in the return series. This could mean that the market is less predictable. Traders might adopt more cautious, flexible strategies or favor options strategies that benefit from uncertainty.
* **Trend Persistence (Hurst Exponent):** A high Hurst exponent bin (after clustering) may indicate that the market is trending, which could be favorable for trend-following strategies. Conversely, a low Hurst bin suggests mean reversion, implying that contrarian or mean-reversion strategies might perform better.
* **Market Dynamics:** The transitions from all “0.0” states to those with moderate or high values in the second and third digits signal periods of market adjustment. For instance, regimes moving from 0.0\_1.0\_0.0 to 0.0\_1.0\_2.0 highlight moments where the market, despite a calm GARCH state, starts to reflect more dramatic historical events — possibly corresponding to emerging risks or corrections.

**Conclusion**

**Dynamic Strategy Allocation:** If the composite regime shifts from the most frequent state (normal) to a rare state (e.g., high GARCH persistence), it might signal an impending period of elevated risk. Strategies might then dynamically adjust — such as shifting from aggressive positions to defensive ones or from momentum strategies to volatility-based hedging.

**Research & Further Analysis:**

* **Statistical Analysis:** Correlate these composites with subsequent market returns, volatility, options prices, or Greeks to see if certain regimes reliably predict auxiliary metrics.
* **Forecasting Integration:** These composite regime indicators can be integrated into machine learning models, such as Long Short-Term Memory (LSTM) networks to forecast regime changes. This helps in proactive risk management and dynamically adjusting trading