Milestone 3: Pattern Analysis Framework and Strategic Integration

Building upon the multi-scale pattern recognition system from Milestone 2, I have completed the pattern analysis framework that transforms clustering insights into trading strategy components. This milestone delivers analytical tools to bridge pattern identification and betting decisions, while also revealing statistical properties in the game's random number generation that warrant further investigation.

Pattern Analysis Framework Implementation (Task 5)

The pattern analysis framework represents the culmination of our clustering efforts, providing sophisticated tools to extract trading insights from the identified patterns. Each component serves a specific strategic purpose in our quest to build a winning competition strategy.

Representative Pattern Extraction (Medoids)

I've implemented a medoid extraction system that identifies the most representative round from each cluster. Medoids are actual game rounds rather than synthetic averages, making them directly interpretable. The parallel implementation reduces computation time for the expanded 462-round dataset.

For the optimized 5-cluster solution, the medoids represent distinct patterns:

- **Cluster 0 (round_4048)**: Moderate winners averaging 5.25x
- **Cluster 1 (round_4233)**: Low performers at 1.30x (95% crash rate)
- **Cluster 2 (round_10338)**: "Moon Shot" patterns reaching 31.51x
- **Cluster 3 (round_4127)**: "Steady Climbers" averaging 3.30x
- **Cluster 4 (round_0312)**: Marginal patterns at 1.81x (75% crash rate)

These medoids provide reference patterns, though their practical application requires careful consideration of the high crash rates observed.

Barycenter Average Computation

To complement the medoids, I've developed barycenter computation that creates synthetic average patterns for each cluster. Using fixed-length resampling (50 ticks), these barycenters provide smooth, normalized representations ideal for machine learning features.

The barycenter analysis reveals the underlying "shape" of each pattern type, abstracting away individual round noise to expose core behavioral characteristics. This dual approach - real medoids for interpretation and synthetic barycenters for computation - provides both intuitive understanding and mathematical precision.

Pattern Continuation Tendency Analysis

The continuation tendency analysis examines post-peak behavior across clusters. These metrics quantify average decline rates after patterns reach their maximum multiplier:

- **Cluster 0**: Slope of -0.39 (moderate decline after peak)
- **Cluster 1**: Slope of -0.07 (minimal decline, already low multipliers)
- **Cluster 2**: Slope of -4.77 (severe crash after moon shot peaks)
- **Cluster 3**: Slope of -0.12 (gradual decline)
- **Cluster 4**: Slope of -0.12 (similar gradual decline)

The extreme negative slope for Cluster 2 (-4.77) indicates these "moon shot" patterns experience catastrophic crashes after peaking, requiring immediate exit detection. The more moderate slopes in other clusters suggest varying degrees of manageability in exit timing.

Length-Based Distribution Analysis

The pattern distribution across length strata reveals game mechanics insights:

Short rounds (<50 ticks): 157 total - distributed across all clusters

Medium rounds (50-200): 137 total - slight concentration in Clusters 0,1

Long rounds (200-500): 116 total - fairly balanced distribution

Extra-long rounds (500+): 52 total - concentrated in Clusters 0,2

Notable findings:

- The instant crash rate has increased to 21.6% in this larger dataset
- Short rounds show no clear cluster preference, suggesting early behavior is less predictable
- Extra-long survivors tend toward either moderate wins (Cluster 0) or moon shots (Cluster 2)

This distribution indicates that survival time correlates with final outcome magnitude, though the relationship is complex given the varied crash rates across clusters.

Statistical Analysis: Autocorrelation Findings

During pattern analysis, I conducted additional statistical investigations on the game's random number generation. This supplementary analysis revealed measurable properties that merit documentation, though their practical exploitation faces significant challenges.

Drift Autocorrelation Findings

Analysis of over 10,000 rounds revealed significant autocorrelation in drift patterns:

- **Lag-1 autocorrelation**: 0.224 (consecutive rounds show correlated drift)
- **Lag-5**: 0.30 (correlation persists across 5 rounds)
- **Lag-10**: 0.29 (still significant at 10-round distance)

These values, which should theoretically be near zero in a truly random system, suggest the game's RNG exhibits memory effects. Consecutive rounds tend to have similar early-tick momentum characteristics.

Drift-Outcome Correlation

Further investigation revealed a positive correlation (0.20) between early-tick drift and final multiplier outcomes. Rounds with high positive drift in the first 20 ticks show:

- **Top 20% drift rounds**: 70-74% reach 2x multiplier
- **Bottom 20% drift rounds**: Only 35-40% reach 2x
- **Baseline**: 57% reach 2x

This represents a potential 15-17% edge if properly exploited.

Strategic Implications of Statistical Findings

While the autocorrelation patterns are statistically significant, practical application faces substantial obstacles:

1. **Early crash rate**: With 21.6% of rounds crashing before 20 ticks, any observation-based strategy faces automatic losses on these rounds
2. **Observation requirements**: Drift calculation requires surviving the initial observation period
3. **Execution constraints**: The 0.2-second decision window limits computational complexity

These findings suggest the game's RNG has measurable patterns, but whether they can be profitably exploited remains uncertain. The autocorrelation will be incorporated as additional features in our upcoming genetic algorithm optimization, where systematic

Strategic Framework Integration

The pattern analysis framework provides strategic guidance, though the high crash rates across most clusters present challenges:

Cluster-Based Approach:

1. **Cluster 2 (Moon Shots)**: 27% crash rate before 2x - most favorable for aggressive strategies
2. **Cluster 0 (Moderate Winners)**: 44% crash rate - acceptable risk for standard positions
3. **Cluster 3 (Steady Climbers)**: 52% crash rate - marginal opportunity
4. **Clusters 1 & 4**: 75-95% crash rates - generally avoid

Practical Considerations:

- The exit zones identified (mostly starting near 1.0x) suggest entry points rather than exits, indicating further refinement needed
- High overall crash rates (only Cluster 2 offers <50% crash probability)
- The 21.6% instant crash rate significantly impacts any strategy

Competition Strategy Elements:

1. Focus on Cluster 2 identification for best risk/reward
2. Conservative position sizing given high failure rates
3. Multi-bot approach to spread risk across attempts
4. Quick exit triggers given steep post-peak declines

Conclusion and Next Steps

Milestone 3 has completed the pattern analysis framework, providing tools to analyze clustering results and extract strategic insights. The expanded 462-round dataset has revealed both opportunities and challenges in developing profitable strategies.

Key findings:

- Comprehensive pattern analysis toolkit implemented (medoids, barycenters, continuation metrics)
- Five distinct clusters identified, though most show unfavorable crash rates
- Statistically significant autocorrelation detected (0.224), though practical application remains uncertain
- Higher instant crash rate (21.6%) than initially observed

- Clear documentation and framework for future development

Challenges identified:

- Only one cluster (Moon Shots) shows crash rates below 50%
- Exit zone calculations require refinement
- Early crash problem limits observation-based strategies
- Overall win rates appear lower than expected game parameters

Moving into Milestone 4, genetic algorithms will optimize across this complex parameter space. The framework will test various combinations of cluster predictions, drift indicators, and risk management rules to find viable strategies despite the challenging win rate environment. Given the high crash rates observed, optimization will need to focus on identifying the rare profitable opportunities while minimizing exposure to the predominant losing patterns.