Final Report Draft (Structure with Integrated Outputs)

---

1. Executive Summary

- Traders' performance is influenced by **sentiment regimes**, but the relationship is nonlinear.
- **Fear & Greed regimes** drive higher risk-taking and anomalies.
- **Volatility and momentum** are stronger predictors of profitability than sentiment alone.
- Risk analysis shows ~9.7% of trades are anomalous, concentrated in high trade-count and high-volatility clusters.
- Portfolio backtesting suggests **volatility-based strategies are more stable**, while sentiment signals provide opportunistic edges during extreme regimes.

---

2. Introduction

Objective:
Exploring the relationship between trader performance and market sentiment, uncovering hidden patterns, and deriving insights for smarter strategies.

Methodology:

- EDA & Time Series
- Behavioral & Predictive ML models
- Advanced ML (LightGBM, SHAP)
- Risk & Anomaly detection
- Portfolio Backtests .

---

3. Exploratory Data Analysis

Sentiment Impact

- Distribution shows **Neutral sentiment** dominates trading volume.

- **Extreme Fear & Extreme Greed** regimes correspond to wider spreads in PnL outcomes.
- Profitability is higher in Greed phases, but **risk of sharp losses also spikes**.

Time Series Analysis

- Market sentiment and trader PnL exhibit **co-movement**, but with lag effects.
- Shifts in sentiment regimes (Fear → Neutral → Greed) often precede performance swings.
- Hidden cyclicality detected in **PnL volatility**, aligned with sentiment regime transitions.

---

4. Advanced Behavioral Analytics

- **Overtrading tendency**: Clear escalation in trade count during Fear & Greed.
- **Risk asymmetry**: Losses during Fear are disproportionately higher than gains in Greed.
- **Behavioral clustering**:
  - Conservative cluster (low trades, stable returns)
  - Opportunistic cluster (moderate trades, higher win rate)
  - Risk-seeking cluster (high trades, volatile PnL).

---

5. Machine Learning Predictive Models

- Regression & classification models applied to predict PnL from features.
- Performance moderate, highlighting difficulty of direct prediction.
- Key features: **volatility, leverage, and momentum indicators**.

---

6. Advanced ML Models

Model Comparison:

- XGBoost $R^2 \sim -1.8$ (poor fit)
- LightGBM $R^2 \sim 0.37$ (better but limited predictive power).

Feature Importance (LightGBM):

- **pnl_momentum**, **pnl_ma_7**, **pnl_volatility**, **trade_count** = top drivers.
- Sentiment features are weak alone, but useful in interaction terms.

SHAP Analysis:

- High **pnl_momentum** increases both upside and downside risk.
- Trade count has nonlinear effects → small increases in benign, high levels linked with anomalies.

---

7. Risk & Anomaly Detection

- **9.66% anomalies detected** (~20,395 cases).
- Anomalous trades:
  - Mean pnl_volatility ~6.9 vs ~1.0 for normal trades.
  - Trade count ~2,078 vs ~442 normal.
- **Correlation heatmap**:
  - Overtrading ↔ anomalies (0.54) strongest link.
- Distribution: losses heavily skewed (fat tail).
- Temporal analysis: **spikes in anomaly rate precede high loss periods**.

---

8. Portfolio Optimization & Strategy Backtests

- **Sentiment-driven strategy**:
  - Mean return ~36K, but high volatility and negative median = inconsistent.
- **Volatility-driven strategy**:
  - Mean return ~63K, narrower downside, better stability.
- **Recommendation**:
  - Use a volatility **strategy as a baseline**.
  - Overlay **sentiment triggers** during extremes for opportunistic trades.

---

9. Key Insights & Hidden Patterns

1. Sentiment vs Performance

- **Fear & Greed regimes** strongly influence trading behavior.
- In **Fear phases**, traders increase trade counts but suffer **disproportionately higher losses**.
- In **Greed phases**, profitability rises but so does **volatility**, creating unstable outcomes.
- **Neutral sentiment** seems safer but still hides clusters of anomalies.

2. Hidden Behavioral Patterns

- **Overtrading** is the biggest behavioral risk → directly linked with anomalies and high losses.
- **Conservative traders** (low trade count, steady PnL) outperform risk-seekers over time.
- Cyclical patterns: spikes in sentiment volatility tend to precede performance swings.

3. Machine Learning Insights

- **Volatility and momentum indicators** are the strongest predictors of PnL — not sentiment alone.
- **Trade count** has a nonlinear effect: moderate = healthy, extreme = anomaly risk.
- Sentiment features add value when combined with PnL/volatility → best as a **modifier/filter**, not a direct predictor.
- ML performance is modest, showing the market remains **partly unpredictable**.

4. Risk & Anomaly Detection

- **~9.7% of trades are anomalous**, with far higher volatility and trade counts.
- Anomalies explain a large share of catastrophic losses.
- Correlation analysis: **overtrading** ↔ **anomalies (0.54)** = strongest driver of hidden risk.
- Monitoring **anomaly rate over time** can serve as an **early warning signal** for loss spikes.

5. Portfolio Strategy Development

- **Volatility-driven strategies**: more stable, higher mean returns (~63K).
- **Sentiment-driven strategies**: opportunistic but inconsistent — better as a tactical overlay.
- **Best approach** = **hybrid strategy**:
  - Use volatility as a **core signal**.
  - Layer sentiment triggers during extreme regimes (Extreme Fear/Greed).
  - Apply **strict trade count/leverage caps** to suppress anomaly-driven losses.

10. Conclusion & Recommendations

- **Volatility-driven strategies** provide stable core performance.
- **Sentiment overlays are valuable** during extreme Fear/Greed for tactical positioning.
- **Risk controls**: cap trade count/leverage to suppress anomaly-driven blowups.
- **Smarter trading strategies**:
  1. Anchor portfolio allocation on volatility signals.
  2. Apply sentiment filters only in extremes.
  3. Monitor anomaly risk rate as a live indicator.
  4. Use ML-extracted signals (momentum, moving averages) for tactical timing.