

Robustness & Scenario Validation Roadmap

This document defines the next testing scenarios beyond the initial full-year demo run. None of these tests have been executed yet. This is a planning scaffold.

Why this matters: A single backtest period can be lucky. Robustness requires the strategy to perform across different market regimes, cost assumptions, data conditions, and time periods.

Proposed Test Matrix

1. Historical Market Regime Tests

Scenario	Period	Regime Characteristics	Status
2020 crash & recovery	2020-02 to 2020-08	Flash crash, V-shaped recovery, high volatility	Not run
2022 bear market	2022-01 to 2022-12	Sustained drawdown, rate hikes, tech sell-off	Not run
2023 bull market	2023-01 to 2023-12	Low vol grind-up, narrow leadership	Not run
2024 demo (baseline)	2024-01 to 2024-12	Current demo artifact	✓ Complete

2. Synthetic Regime Tests

Scenario	Description	Status
High-vol regime	$\sigma = 25\text{--}30\%$ annual, no trend	Not run
Flat/choppy regime	Zero drift, high mean-reversion	Not run
Low-liquidity regime	Wider spreads, higher slippage model	Not run

3. Data Quality Stress Tests

Scenario	Description	Status
Missing data simulation	5–10% of bars randomly removed	Not run
Stale price injection	Bars frozen for 1–3 days mid-period	Not run
Corporate action stress	Multiple splits and dividends	Not run

4. Cost / Execution Stress Tests

Scenario	Description	Status
Higher slippage	2× base slippage model	Not run
Delayed execution	Orders filled at T+1 instead of T	Not run
Partial fill simulation	50% fill rate on orders	Not run

5. Out-of-Sample Validation

Scenario	Description	Status
Walk-forward test	Rolling 3-month IS / 1-month OOS windows	Not run

Out-of-sample period	Hold out 2025 data entirely	Not run

6. Statistical Robustness

Scenario	Description	Status
Monte Carlo bootstrap	1,000 synthetic paths from same parameters	Not run
Parameter sensitivity	±20% on key strategy parameters	Not run
Correlation stress	Reduce platform/benchmark correlation to 0.3	Not run

Recommended Priority Order

1. **2022 bear market** — most adversarial to momentum-based strategies
2. **Walk-forward test** — standard OOS validity check
3. **Higher slippage** — most common point of failure in live deployment
4. **Monte Carlo bootstrap** — quantify confidence intervals on reported Sharpe

Implementation Notes

- Each scenario should produce its own artifact JSON with a unique `fixture_name`
- The visualization runner accepts any artifact: `python -m visualization.run_all --artifact <path>`
- Scenario artifacts should be stored under `artifacts/`

platform/backtests/

- This roadmap should be updated as scenarios are completed

This file is a planning scaffold generated by the Autonomous Trading Platform visualization package. No scenario runs are implied to have been completed.