Module 4 — Hypothesis-Driven Exploratory Analysis

Turning Structured Data Into Decision Evidence

Module Purpose (Why This Exists)

Module 4 answers a critical question:

*Given a well-framed decision, defined KPIs, and a disciplined data model — what evidence does the data provide **for or against** deploying this EA under standardized rules?*

This is **not** open-ended EDA.
This is **decision-constrained exploration**.

Your task is not to "find insights."
Your task is to **test beliefs** that already exist.

That distinction is what separates:

- Analysts → from → decision partners

What Changes in This Module

Earlier Modules	Module 4
Structure first	Evidence first
Definitions	Validation
"What should we measure?"	"Is it actually true?"
Design correctness	Decision risk

You are no longer designing analytics.
You are **challenging assumptions**.

Core Analytical Principle (Very Important)

Exploratory Analysis without hypotheses is just wandering.

In professional analytics environments, *every exploration* is anchored to:

- A belief
- A risk
- Or a decision criterion

Inputs You Will Use

From prior modules:

- **Decision Framing (Module 1)**
    - Go / No-Go deployment decision
- **KPIs (Module 2)**
    - Risk %, Drawdown, Expectancy, Goal Hit Rate
- **Data Model (Module 3)**
    - Trade-level and Daily-level facts

From data:

- **Phase 1 data** (pre-standardisation)
- **Phase 2 data** (post-standardisation)

These are **two regimes**, not one dataset.

Module 4 Deliverable

You will produce a:

Hypothesis-Driven Exploratory Analysis Report

This is a document — not code, not charts alone.

Think of it as:

- A *scientific lab notebook*
- Written for business stakeholders

Required Structure of the Artefact

You must follow this structure exactly.
Deviation weakens decision traceability.

---

1. Analytical Context

Briefly restate:

- The decision being supported
- The execution regime under evaluation
- The dataset phase being analyzed (Phase 1 or Phase 2)

This prevents hindsight bias.

---

2. Hypothesis Register (Core of the Module)

You will define **4–6 hypotheses**, each formatted like this:

H# — Hypothesis Statement

"If the EA is executed under standardized rules, then [measurable outcome] should occur, because [business logic]."

Example:

If risk is fixed at 5% per trade, then maximum daily drawdown should remain within X%, because trade exposure is capped and trade frequency is constrained.

Each hypothesis must reference:

- A KPI
- A grain (trade or day)
- A phase (pre or post)

---

3. Exploratory Tests (Not Conclusions Yet)

For each hypothesis:

- What metrics were examined?
- What distributions or comparisons were inspected?
- Why these views were chosen

You may include:

- Summary statistics
- Distribution descriptions
- Outliers and anomalies

Do **not** optimize yet.
Do **not** conclude yet.

---

4. Observations (Evidence, Not Opinion)

This is where discipline matters.

Use language like:

- "The data shows…"
- "The distribution indicates…"
- "In X% of cases…"

Avoid:

- "This proves…"
- "Clearly…"
- "Obviously…"

This is evidence gathering, not verdict delivery.

---

5. Decision Implications (Preliminary)

For each hypothesis:

- Does evidence **support**, **challenge**, or **fail to test** it?
- What decision risk does this introduce or reduce?

This section connects analytics back to Module 1.

What You Are Explicitly NOT Doing Yet

- ❌ Final recommendations
- ❌ Strategy optimization
- ❌ Dashboard storytelling
- ❌ Overfitting interpretations

Those belong to **Module 5 & 6**.

Teaching Note: What Good Looks Like

A strong Module 4 submission:

- Is uncomfortable — because it may expose flaws
- Challenges your own assumptions
- Makes uncertainty visible
- Uses restraint in conclusions

Senior analysts are trusted **not** because they are confident, but because they are careful.

Common Failure Modes (Avoid These)

- Jumping to profitability conclusions
- Mixing Phase 1 and Phase 2 results
- Ignoring losing streaks because net profit is positive
- Letting charts “tell stories” without hypotheses

If you avoid these, you are operating at a **professional level**.

Your Next Action

1. Create a document titled:
   “Hypothesis-Driven Exploratory Analysis — Module 4”

2. Start with **3–4 hypotheses only**
   (More than that dilutes focus.)

3. Use **Phase 2 data first**
   (Phase 1 will be used for contrast, not validation.)