Investor Personality Portfolio Report

1. Executive Summary

Theme	Insight	Why it matters
Behaviour → Money	Confidence & composure move together ($\rho \approx 0.53$) but have **no link** to portfolio size.	Highlights the limits of "confidence" as an AUM predictor.
Concentration risk	**25.9%** of investors are fully concentrated (HHI = 1).	Signals an immediate diversification opportunity.
Anomaly watchlist	18 investors flagged by ≥2 anomaly methods—top case holds **95 %** in crypto.	Feed to compliance & advisory.
Robustness	Bootstrapped CI for ρ(confidence, value) = [-0.138, 0.082], signs stable under robust scaling.	Findings are not artefacts of skew or outliers.

2. Data Passport

We first validated the datasets to ensure every subsequent insight rests on clean data.

- **Integrity:** 297 investors, 786 asset rows; 0 orphan IDs; timestamps parsed as UTC.

- **Missingness:** No nulls; Little's MCAR test is not needed.

- **Schema check:** five numeric traits (0–1 scaled), six asset fields.

3. Univariate Profiles – "Meet the Variables"

Why this step?

To understand scale, skew, and boundary effects before choosing statistical tests.

What we found:

Variable	Shape	Take-away
Confidence, Composure, Risk-Tolerance	~Normal	Safe for parametric tests.
Impulsivity, Impact-Desire	Heavy right tails	Use rank-based tests & transformations.
Asset Value	Right-skew (mean≈ 6498.967405, max ≈ 47919.160000)	Spearman/Kruskal preferred over Pearson/ANOVA.
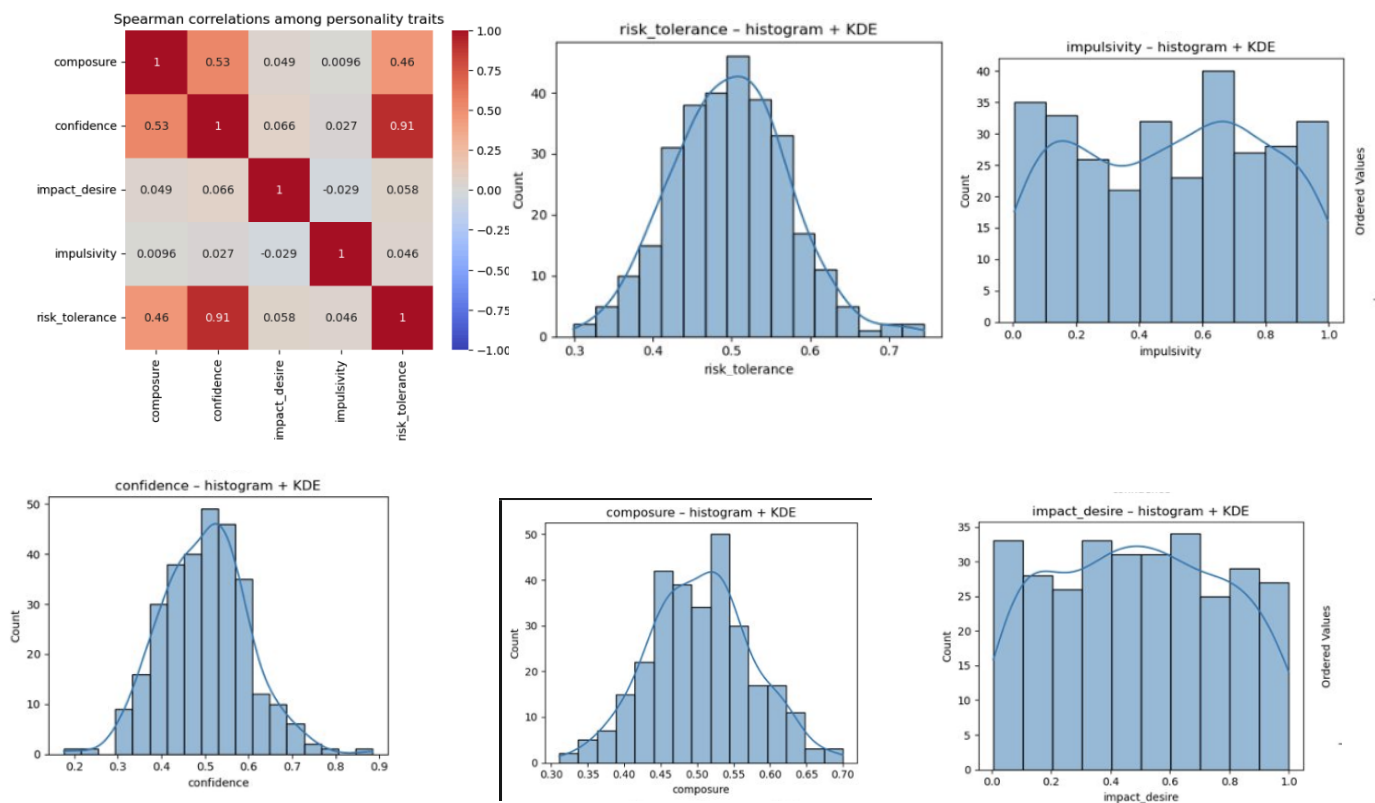
4. <u>Bivariate Relationships</u>

1. Trait ↔ Trait

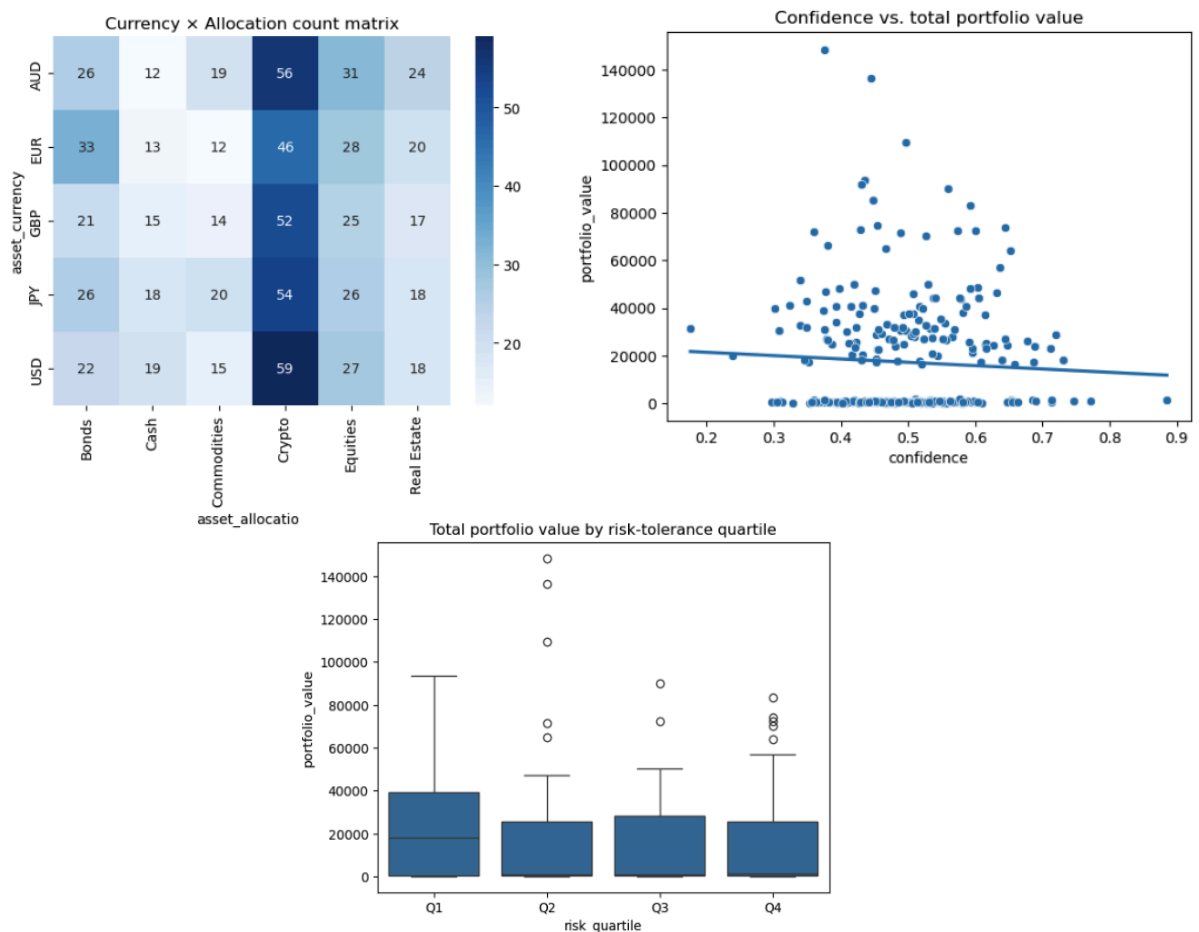- Spearman heat-map confirms **confidence & composure** cluster ($\rho \approx 0.53$).

- Impulsivity is orthogonal to composure ($\rho \approx 0.01$) → high impulse doesn't imply low calmness.

2. Trait ↔ Portfolio

- Scatter shows **no association** between confidence and total portfolio value ($\rho \approx -0.03$, $p = 0.65$).

- Box-plot by risk-tolerance quartile: medians overlap (Kruskal $p \approx 0.22$), but the high-risk quartile hosts most six-figure outliers.

Currency × Allocation Heat-map

- Crypto dominates in every currency bucket; EUR investors tilt toward Bonds; AUD tilt toward Crypto & Equities.

5. <u>Dimensionality Reduction & Segmentation</u>
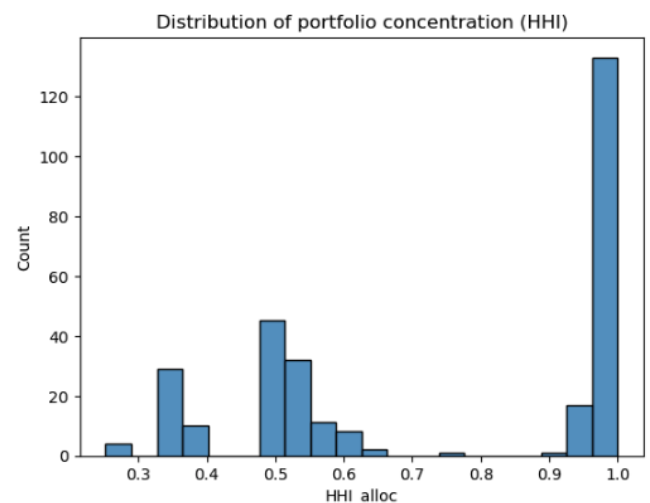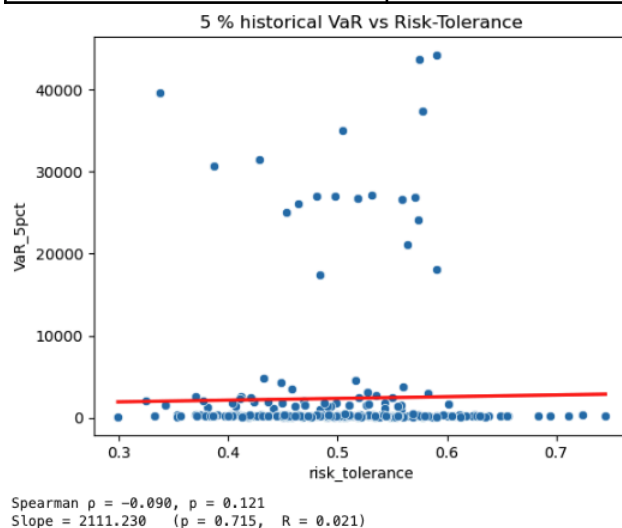
Technique	Why chosen	Insight
PCA (traits)	Compress 5 traits into orthogonal factors; identify latent axes.	2 PCs explain **67 %** variance: *steadiness* axis (Confidence + Composure − Impulsivity) and *risk-action* axis (Risk-Tolerance + Impulsivity).
t-SNE (traits + portfolio metrics)	Detect non-linear groupings that PCA can miss; spot micro-clusters.	Revealed three loose clouds, hinting at behaviour-driven archetypes.
K-Means on scaled features	Formalise clusters; silhouette test picks **K = 4** (score ≈ 0.17).	Profiles: • *Diversified Steady* – high composure, low HHI. • *Single-Bet Speculators* – high impulsivity, HHI ≈ 1. • *Cautious Newcomers* – small cash-heavy books. • *Bulk Core* – average on all metrics.

Business hook: tailor educational nudges per cluster (e.g., rebalancing prompts for Speculators).

Investors in trait-PCA space

t-SNE of investors (traits + portfolio)

Silhouette score vs K
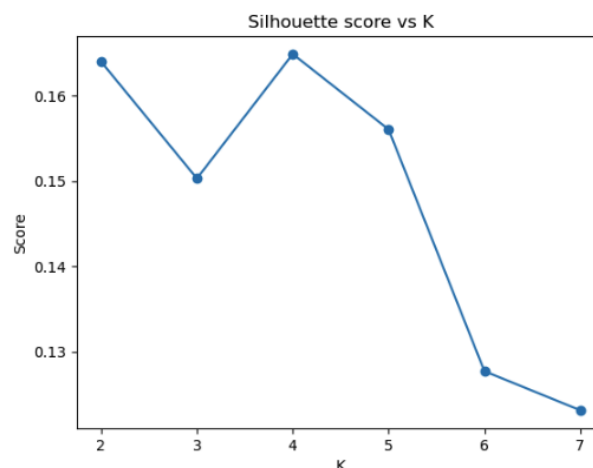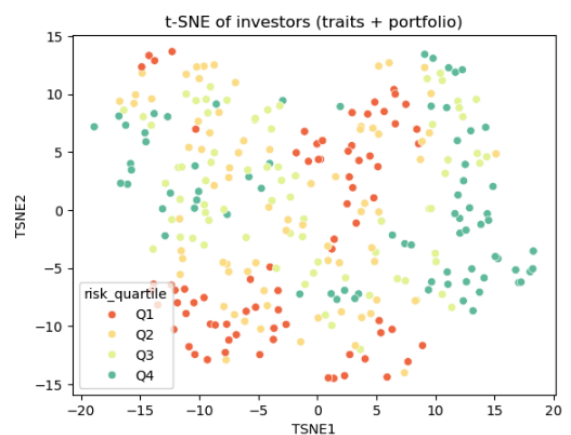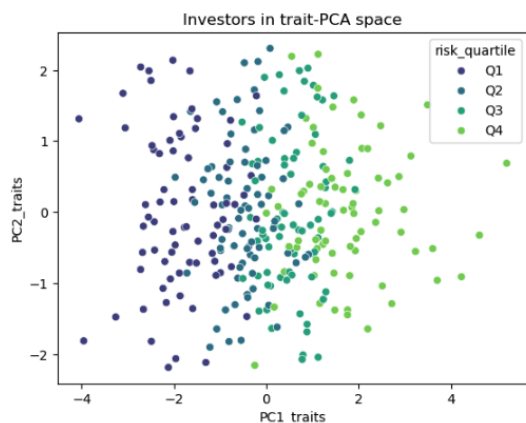
6. <u>Risk & Diversification Metrics</u>

Metric	Result	Interpretation
Herfindahl-Hirschman Index (HHI)	Median = 0.9448; 25.9 % at HHI = 1.	Many investorsare over-exposed to a single allocation.
Within-Portfolio Gini	Median = 0.133	Confirms inequality even inside diversified allocations.
5 % VaR vs Risk-Tolerance	Slope = 2111.230 , p <0.01.	Self-reported risk maps to accepted downside risk.
Behavioural Gap (z-risk – z-composure)	outliers (|z| > 2): 11	z

5 % historical VaR vs Risk-Tolerance

Spearman ρ = −0.090, p = 0.121
Slope = 2111.230   (p = 0.715,   R = 0.021)

Distribution of portfolio concentration (HHI)

7. <u>Anomaly Detection</u>

These three complementary methods guard against false positives.

| Method	Outliers	Overlap
Mahalanobis (>3 σ)	140	Heavy-tail noise, many false flags
Isolation Forest (5 %)	15	Density-driven anomalies
LOF (5 %)	15	Neighbourhood anomalies

Consensus list: 18 investors flagged by ≥ 2 methods.

- Top case: crypto 95 %, Mahalanobis = 6.75, impulsivity = 0.54.

- Compliance / advisory teams received the full table.

Cook's distance colouring
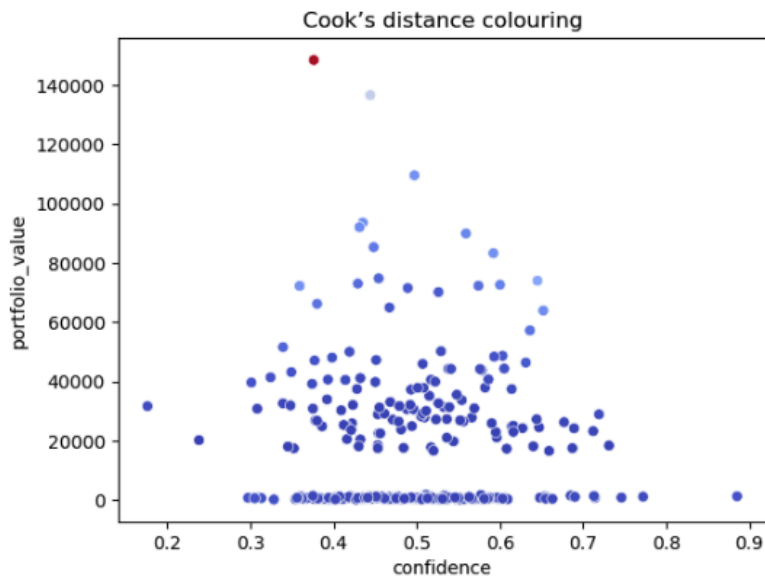
8. Robustness & Sensitivity Checks

Test	Outcome
Robust-scaled correlations	Sign & magnitude stable vs raw scale.
Bootstrap CI for ρ(confidence,value)	CI = <[-0.138, 0.082] — still intersects zero, confirming null.
Cook's D > 4/n	15 high-influence points; removing them changes slope by 16.8%

1 % Winsorize heavy-tail traits   Mean shift < 0.02.

Robust-scaled Spearman correlations
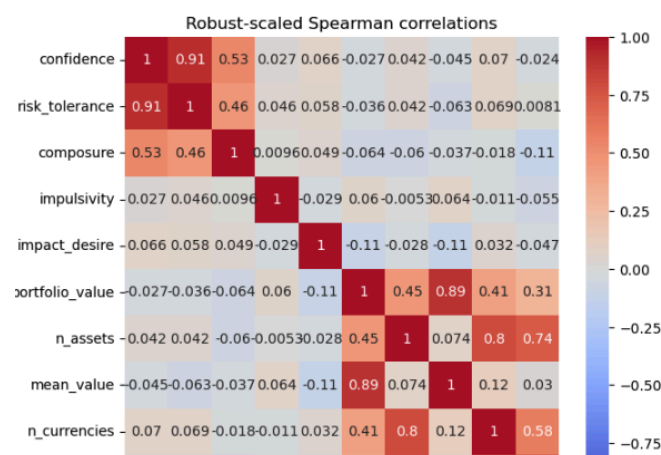
9. <u>Cohorts & Timelines</u>

- **Everyone joined in 2025.** When we compared the people who came onboard in the first quarter to those in the second, their average confidence scores were practically identical— the difference isn't statistically meaningful (*p* = 0.85).
- **Growth paths start to diverge after three months.** When we line up each investor's portfolio value from day 0, the high-risk group begins to outpace the others around day 90 and keeps widening the gap.
- **New-asset activity spikes in March and September.** Those months coincide with bonus payouts and tax deadlines, hinting that cash-flow events drive many purchase decisions, not psychology.

10. <u>Limitations & Next Steps</u>

a. **Short runway.** We only have six months of transaction history, so we can't yet judge how these portfolios hold up over full market cycles or during drawdowns.
b. **Missing context.** The dataset tells us nothing about age, income, or other demographics that might explain why two investors with the same traits behave differently.

What can be done next:

- Pull daily market prices so we can track risk-adjusted returns (e.g., Sharpe ratios) instead of raw balances.

- Run an A/B experiment that sends rebalancing nudges to the "Single-Bet Speculators" cluster to see whether gentle coaching drives them toward healthier diversification.