Introduction

The exploratory data analysis (EDA) was structured to answer one central question:
To what extent does expressed risk tolerance align with actual financial behavior?

To do this, I analyzed:

- The **distribution and range of risk tolerance** scores
- The **structure of GBP asset holdings**, in line with task requirements
- The **relationship** between risk tolerance and financial engagement
- Behavioral mismatches and emerging patterns in **currency diversity** and **asset type preference**

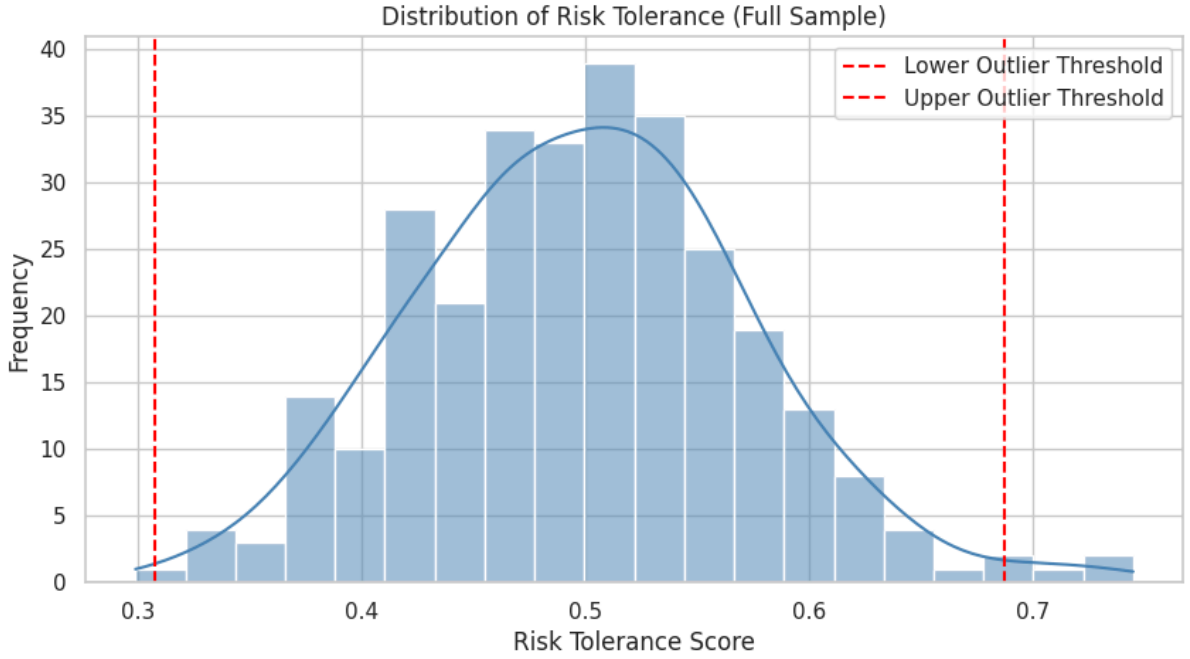
The analysis follows a narrative-driven structure with behavioral framing, enabling us to uncover not just what people do, but **how their actions reflect or diverge from their self-perception**.

We'll proceed through five parts, asking five questions to guide our analysis:
1. *What kinds of investors are we dealing with?*
2. *What do they hold — and how much?*
3. *Do their behaviors match their beliefs?*
4. *What patterns of diversification or currency preference emerge?*
5. *What drives portfolio complexity?*

As a bonus, I added a spotlight section on outliers that I thought would be interesting to study.

Question 1: What kinds of investors are we dealing with?

This section aims to understand the distribution and extremes of investor risk tolerance across all 297 individuals — this lays the behavioral foundation of the population we're analyzing, before linking it to financial activity.



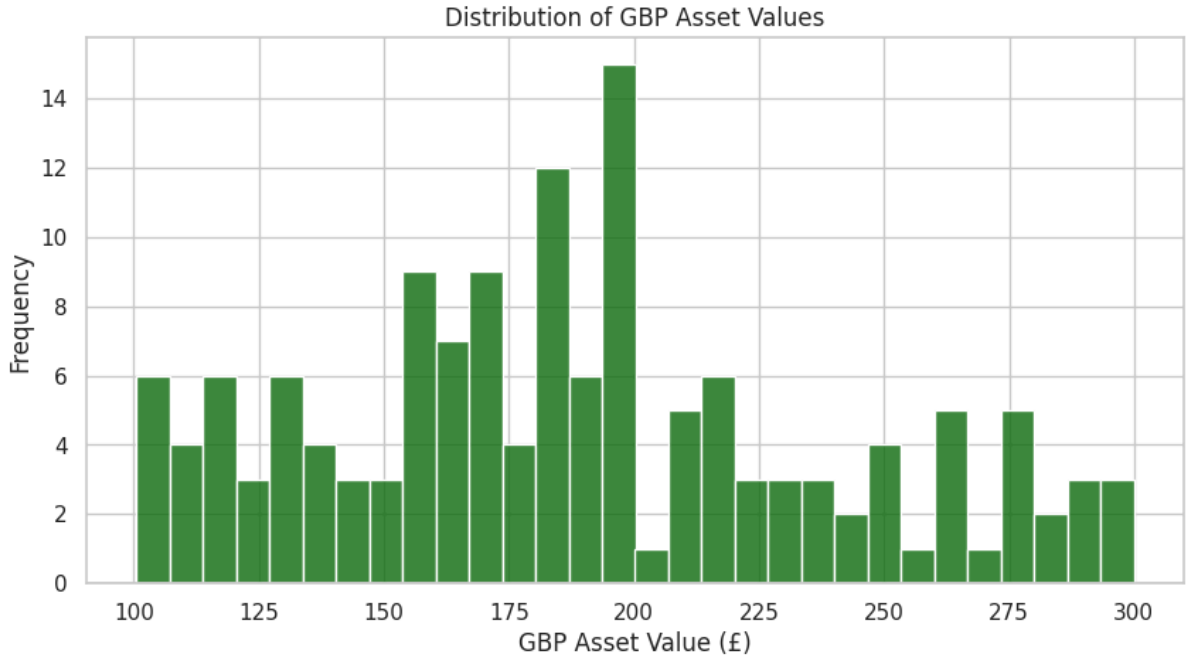
Some key statistics (the corresponding code is in the main analysis file):
1. Mean risk tolerance: 0.499
2. Standard deviation: 0.074
3. Median risk tolerance (visually inspected): 0.509
3. Range: 0.299 – 0.745
4. Interquartile Range (IQR): 0.450 – 0.545
5. Outliers: 5 individuals (outside [~0.34, ~0.66])

From the visual of the distribution:
1. The distribution is approximately normal, centered around 0.5 — indicating a moderately risk-averse population.
2. The lack of extreme skew suggests balanced sampling, but the outliers merit further investigation (later in our outlier spotlight).

A total of 5 behavioral outliers were identified based on risk tolerance (using 1.5×IQR). These individuals were retained in all analyses, as they represent plausible real-world profiles and their statistical influence was negligible. Additionally, no missing values or structural issues were found during the initial data health check, confirming the datasets were analysis-ready.

Question 2: What do they hold — and how much?

This section aims to understand how GBP asset values are distributed and whether certain individuals hold unusually high or unusually low GBP-denominated wealth.



Some key statistics:
1. Total GBP Assets Analyzed: 144
2. Mean Value: GBP 188.30
3. Median: GBP 185.69
4. IQR: GBP 153.88 – GBP 218.47
5. Maximum: GBP 299.99 (per-asset), GBP 542.86 (per-person)

6. Financial Outliers - Individuals with total GBP assets ≥ £450.45 (top 5%) were flagged. 7 individuals exceeded this threshold — indicating high-concentration GBP holdings. For example, _id = 134 holds £542.86 — the highest. These individuals were retained in all analyses, as they represent plausible real-world profiles and their statistical influence was negligible. Additionally, no missing values or structural issues were found during the initial data health check, confirming the datasets were analysis-ready.

Here are some insights I could derive from the observed distribution:
1. GBP asset values are tightly clustered with minimal spread — suggesting standardized holdings.
2. The top 5% are meaningful financial outliers, despite relatively small nominal values — these individuals likely warrant closer behavioral scrutiny.

This distribution sets the stage for analyzing behavioral-financial mismatches, which we explore next.

Question 3: Do their behaviours match their beliefs?

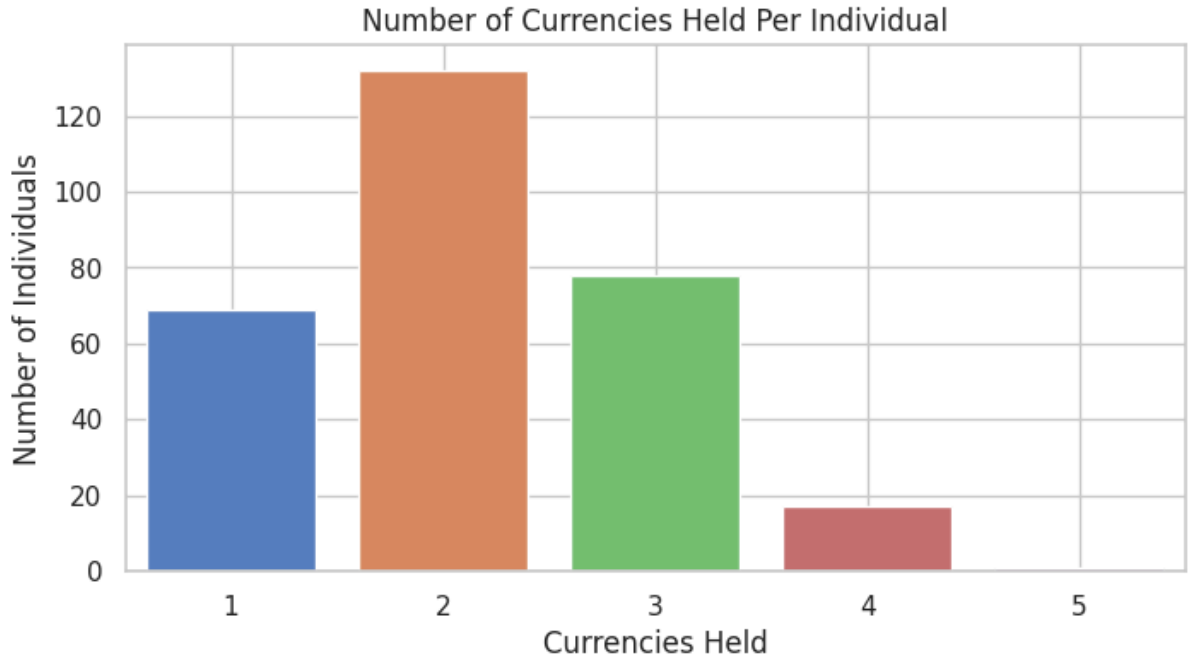
This sections aim to identify individuals whose self-reported risk tolerance diverges from their financial engagement via GBP holdings. We're not measuring riskiness of investment, but rather **degree of commitment or exposure to GBP-denominated assets within the context provided**. These mismatches highlight potential behavioral contradictions — such as **overconfidence**, **hidden constraints**, or **unexpressed strategy**.
A. *Mismatch Group 1: Expressed Risk Appetite, Low GBP Exposure*
There are 7 individuals with high risk tolerance (≥ 0.60) but below-median GBP holdings. For example: _id = 210 has a very high risk tolerance (0.745) but only £162.31 in GBP assets. Some possible interpretations are that their risk appetite is expressed in non-GBP assets or non-financial domains; that there could be a behavioral aspiration vs. access gap (possibly due to some structural limitations i.e., they lack the capital, tools, or access to act accordingly); or that there is a preference for riskier instruments which are outside the scope of GBP.

B. *Mismatch Group 2: Expressed Caution, High GBP Exposure*
behavioral anchoring in perceived safety of GBP instruments
There are 5 individuals with low risk tolerance (≤ 0.40) but upper-quartile GBP holdings. For example: _id = 266 reports a risk tolerance of 0.400, yet holds £405.05 in GBP assets. Some possible interpretations is that they have a preference for safe asset classes despite high allocation; that they are conservative individuals who invest heavily but cautiously; or that there is a behavioral anchoring in the perceived safety of GBP instruments i.e., they invest heavily not due to risk appetite, but due to trust in the stability of familiar, domestic financial products.

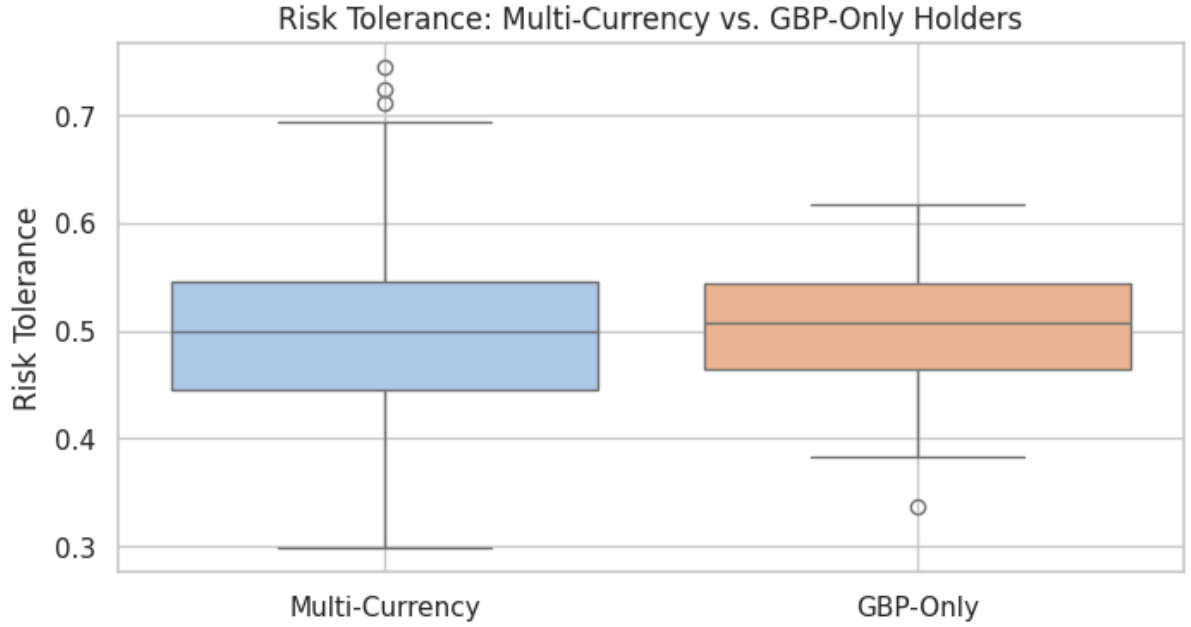
In my opinion, these mismatches are not "errors," but rather behavioral tensions that can inform:
1. Personalization of advice or nudges
2. Profile segmentation
3. Future qualitative research

Question 4: What patterns of diversification or currency preference emerge?

This section aims to explore how individuals structure their portfolios beyond value, by currency diversity and asset allocation types. This adds context to both behavioral types and outliers, checking whether these patterns reflect behavioral traits like risk tolerance.

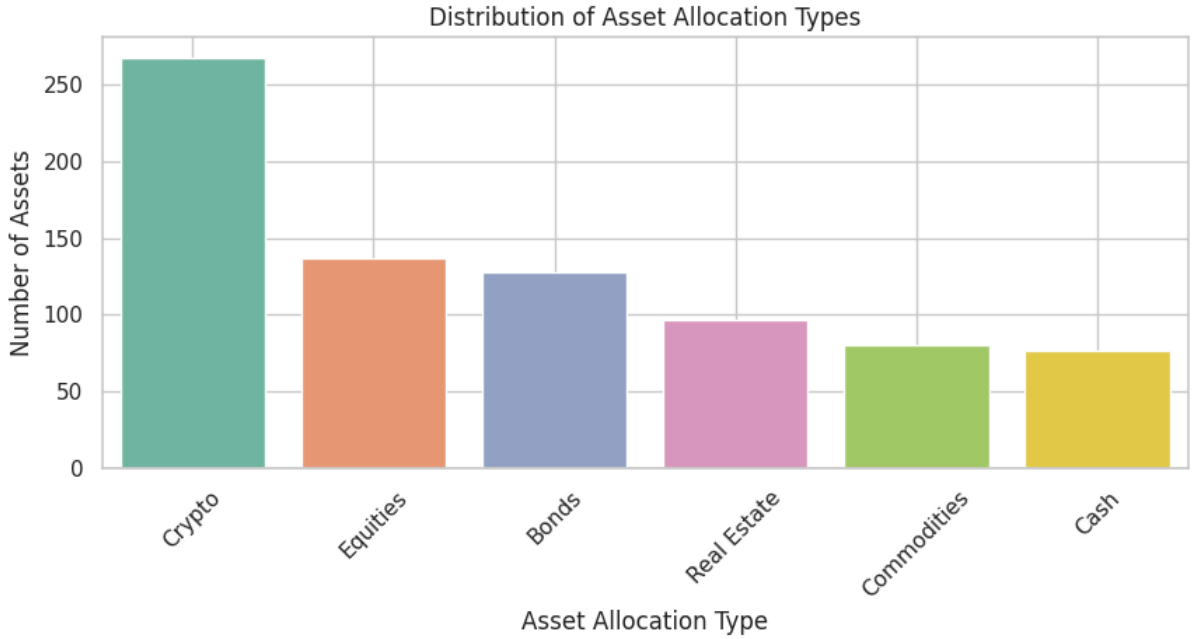


Most individuals hold 2 currencies, followed by 3. Few hold 4, and apparently none hold 5 — indicating limited global diversification in this dataset.



Risk tolerance spread is broader among multi-currency holders than GBP-only individuals. The multi-currency group reading indicates higher variance and several high-risk outliers. GBP-only holders show a narrower risk range, a slightly higher median risk tolerance (visual only, no significance test run).

A possible interpretation is that multi-currency investors span a wider behavioral spectrum (from very cautious to very bold), possibly reflecting financial literacy or geographic spread. GBP-only investors may be more behaviorally homogenous or risk-aware within limited options.



Crypto is the most commonly held asset class, followed by Equities and Bonds. Cash, Commodities, and Real Estate are least represented. Crypto's dominance suggests:
a. Either a skewed dataset (e.g., sourced from crypto-engaged users)
b. Or a real shift toward alternative assets among participants

This contrasts with the overall conservative GBP value distribution, possibly revealing segmented risk behavior: low GBP exposure, but high crypto participation. So, one insight is that investors differ not just in how much they hold, but how they allocate and diversify. Currency and asset choice reflect latent behavioral dimensions not captured by risk tolerance alone.

Question 5: What drives portfolio complexity?

This section investigates whether individuals who express higher risk tolerance also engage in **more complex or volatile portfolio behaviors**, specifically:
1. Holding multiple currencies
2. Allocating more of their assets to crypto

Hypothesis 1: Risk-Tolerant Individuals Hold More Currencies
Rationale: Multi-currency holdings may indicate openness to complexity, exposure, or risk preference.
Result: r = 0.10; A weak positive relationship, suggesting that risk-tolerant individuals may be slightly more likely to diversify across currencies, but the effect is minimal.
Hypothesis 2: Risk-Tolerant Individuals Allocate More to Crypto
Rationale: Crypto is generally perceived as a high-volatility, high-risk asset class.
Result: r = 0.014; Virtually no relationship. Individuals who allocate heavily to crypto do not appear more risk-tolerant than others, based on their self-reported scores.

While crypto is widely held in this dataset, it does not appear to be concentrated among high-risk individuals. This suggests:
1. Cultural or generational drivers may play a bigger role in crypto adoption than personality traits alone.
2. Some individuals may treat crypto as a default or accessible asset, not necessarily a risky one.
The weak relationship with multi-currency behavior implies that holding multiple currencies may reflect access, geography, or product availability rather than risk preference alone.

Bonus Spotlight Section: Outliers

This section highlights individuals who stand out as **behavioral or financial outliers**, based on the patterns uncovered in previous analyses. These cases are not necessarily errors or anomalies; they are **behaviorally interesting profiles** that may represent edge cases, tensions, or hidden strategies.

We spotlight three types:
1. *High Risk, Low Exposure*: Despite the highest risk score in the dataset, this individual holds only median-level GBP assets. It's possible that they may take risk elsewhere (e.g., crypto or non-GBP currencies), or face barriers to acting on their intent, a classic intention–action gap.
2. *Low Risk, High Exposure*: These 5 individuals report low risk tolerance but have some of the highest GBP holdings in the dataset. Their holdings may be in stable or passive asset classes (e.g., bonds), or they may reflect financial conservatism with high allocation, indicating nuanced risk postures. These 25+ individuals allocate 100% (or near-100%) of their assets to crypto while reporting moderate, average risk tolerance (~0.45–0.55). These individuals treat crypto not as a speculative asset, but possibly as a default or normalized investment class. This decoupling of crypto from high-risk perception is a behaviorally significant pattern — it may reflect generational norms, platform defaults, or social proof.
3. *Crypto-Rich, Behaviorally Average*: These 25+ individuals allocate 100% (or near-100%) of their assets to crypto while reporting moderate, average risk tolerance (~0.45–0.55). These individuals treat crypto not as a speculative asset, but possibly as a default or normalized investment class. This decoupling of crypto from high-risk perception is a behaviorally significant pattern: it may reflect generational norms, platform defaults, or social proof.

The presence of these outliers underscores a key lesson in behavioral science: People don't just act on what they say they believe. Their behavior reflects context, access, identity, and cognitive framing, not just their preferences. These profiles are ideal candidates for deeper segmentation, interviews, or product experimentation.

Conclusion

Actual portfolio decisions (especially around crypto and currency exposure) are likely shaped by a combination of contextual constraints, experience, confidence, impulsivity, and possibly peer or trend influence. This analysis indicates that risk tolerance is a necessary but insufficient predictor of complex financial behavior.