Narcissistic Traits in the General Population: A Data Science Approach with the Big Five Personality Traits

---

Introduction

In recent years, there has been a surge of public interest in **narcissism**. Conversations in media, workplaces, and daily life increasingly reference narcissistic behavior, often blurring the line between **narcissistic traits** and **Narcissistic Personality Disorder (NPD)**. While casual use of the term suggests that narcissism is widespread, clinical diagnosis of NPD remains rare, with prevalence estimates of only **0.5–1% of the population**. This discrepancy raises an important question: **how common are narcissistic traits in the general population, and how do they relate to broader models of personality?**

To address this, we draw on the **DSM-5 diagnostic framework**, which defines NPD in terms of nine criteria (e.g., grandiosity, lack of empathy, manipulativeness), of which **five or more must be present**, alongside evidence of dysfunction. Rather than studying NPD as a clinical disorder, our focus is on the **trait level**: the everyday expressions of narcissism that may appear in otherwise typical individuals. Specifically, we connect these DSM-5 criteria to six **Big Five personality facets** where the "low" pole aligns with narcissistic tendencies: **Modesty, Sympathy, Altruism, Morality, Dutifulness, and Friendliness**. Their opposites — arrogance, lack of empathy, selfishness, manipulativeness, irresponsibility, and aloofness — represent the traits most often labeled as narcissistic in common discourse.

Using a dataset of over **400,000 participants** from Kaggle, assessed with the **IPIP-NEO-120**, we explore how often these traits appear in the general population, individually and in combination. By analyzing their **distribution, overlap, and co-occurrence**, we aim to clarify the difference between the **statistical normality of isolated narcissistic traits** and the **rarity of their convergence into a disorder-like profile**. In doing so, this project provides a **data science perspective** on narcissism: not as a clinical diagnosis, but as a measurable constellation of personality facets whose patterns can be quantified and interpreted at the population level.

Mapping DSM-5 NPD Criteria to Big Five Personality Traits

Personality researchers often interpret disorders like NPD as extreme constellations of normal trait dimensions. The DSM-5 criteria for NPD indeed correspond to maladaptive extremes on several of the **Big Five** trait domains. In particular, Narcissistic Personality Disorder is most strongly aligned with **very low Agreeableness**, high **Extraversion**, and certain facets of **Neuroticism**, along with some aspects of Conscientiousness:

- **Agreeableness:** Individuals with NPD characteristically exhibit *antagonism*, which is the inverse of agreeableness. They are described as arrogant, lacking empathy, manipulative, and entitled – all reflecting extremely low agreeableness. Low agreeableness (antagonism) is arguably the core of NPD's interpersonal problems. For example, an NPD person's exploitative and egocentric behaviors represent an absence of traits like modesty, altruism, or cooperativeness. Studies consistently find that narcissism (especially the grandiose type) correlates **negatively** with Agreeableness on Big Five measures. In fact, among the Big Five, **low agreeableness is the most robust trait correlate of narcissistic tendencies**, as also seen in the overlap between "Dissociality/Antagonism" traits and NPD in ICD/DSM trait models.
- **Extraversion:** Many features of DSM-5 NPD reflect high extraversion, particularly the *agentic* or socially bold facets of extraversion. Narcissistic individuals often appear confident, assertive, and attention-seeking. DSM criteria such as having a grandiose sense of self-importance and requiring excessive admiration entail a tendency to *exhibit oneself* and draw attention. In trait terms, this aligns with high extraversion facets like **assertiveness, dominance, and exhibitionism**. Narcissistic persons may come across as outgoing or even charismatic initially, since they actively seek social reinforcement (though their relationships remain superficial). Empirical research with trait inventories finds that narcissism (particularly *grandiose narcissism*) correlates positively with Extraversion. It's worth noting, however, that narcissists' extraversion is often more about **status and attention** than warmth – they may be socially bold but not necessarily empathic or affiliative.
- **Neuroticism:** The relationship between NPD and neuroticism is complex. On the surface, the DSM-5 portrayal of NPD (especially grandiose narcissism) does *not* emphasize anxiety or vulnerability – narcissists often insist on their superiority and may appear confident or emotionally stable. Some research on grandiose narcissistic traits (e.g. using the Narcissistic Personality Inventory in nonclinical samples) has found **lower Neuroticism** or no strong link to Neuroticism, alongside the high extraversion and low agreeableness profile. However, clinical descriptions and DSM text note that narcissistic individuals are **"very sensitive to injury"** and prone to feelings of shame or rage when their esteem is threatened. In other words, beneath the overt grandiosity, they often have fragile self-esteem and can react with intense **reactive anger, shame, or envy** – which are neurotic (negative affect) traits. Expert trait mappings (e.g. by Lynam & Widiger) indicate NPD includes facets of Neuroticism such as **reactive anger and shame** (and an excessive need for approval). Thus, **vulnerable narcissism** (a less overt form of NPD) is associated with high Neuroticism – including insecurity, hostility, and depressive feelings – whereas **grandiose narcissism** tends toward lower Neuroticism in day-to-day

demeanor, except when the person's ego is challenged. In summary, many narcissistic individuals have a **fluctuating self-esteem** and can be emotionally labile in the face of criticism (a neurotic tendency), even if they project confidence outwardly.

In addition to those three domains, some **Conscientiousness** facets appear in NPD as well. Narcissistic personalities might show **high achievement orientation or perfectionism** ("acclaim-seeking" in one description, or **narcissistic perfectionism** noted in ICD-11 trait research). They often have unreasonably high personal standards and ambitions (stemming from seeing oneself as exceptional), which relate to compulsivity and orderliness (a high Conscientiousness tendency) – though this may coexist with impulsivity in other areas due to entitlement (low Conscientiousness in terms of self-discipline). Narcissistic grandiosity can thus be a mix of **high drive and ambition** alongside poor empathy and arrogance.

To summarize, the DSM-5 criteria for NPD map onto an extreme **profile of Big Five traits**: predominantly **very low Agreeableness**, **high Extraversion** (especially assertive, attention-seeking aspects), and a mix of **Neuroticism-related tendencies** (e.g. anger and vulnerability to shame) that are often hidden by outward confidence. This trait profile has been borne out in personality research. For example, a 2002 analysis by Trull & McCrae noted that narcissism (as measured in personality inventories) corresponded to **low Agreeableness and high Extraversion** scores, with Neuroticism showing an inverse or minimal correlation in grandiose narcissism. Meanwhile, clinical theorists (e.g. Russ, Shedler, and others) have observed the undercurrents of emotional reactivity in NPD, linking it to **unstable self-esteem and anger** – a neurotic element. Thus, when viewing NPD through a five-factor lens, it can be conceptualized as a **syndrome of extreme antagonism and extraverted self-promotion, often underpinned by emotional volatility**. This dimensional perspective is explicitly acknowledged in modern diagnostic frameworks: the DSM-5's alternative trait model and the ICD-11 trait-domain model both illustrate that personality disorders like NPD are essentially maladaptive extremes of normal personality traits.

Method

For this study, we rely on the **DSM-5 criteria for Narcissistic Personality Disorder (NPD)** rather than the ICD-10. The DSM-5 is the most widely used reference in clinical psychology and psychiatry, and it provides **nine explicit diagnostic criteria** that describe narcissistic traits in detail. In contrast, the ICD-10 does not recognize NPD as a distinct category; it only mentions narcissistic features under "Other Specific Personality Disorders," offering no structured criteria. Because of this lack of specificity, the ICD-10 is less suitable for trait-level research. The DSM-5, by contrast, offers **clear, operationalized features**—such as grandiosity, entitlement, lack of empathy, and need for admiration—that can be systematically mapped onto personality trait dimensions. This makes it the most appropriate framework for our project, which focuses on analyzing narcissistic tendencies in the general population.

It is important to clarify that this work is not intended as a full clinical lecture on Narcissistic Personality Disorder. This is a **Data Science project**, and the purpose of describing the DSM-5 criteria is simply to establish a solid foundation for selecting traits that can be analyzed statistically. Our focus is not on diagnosing NPD, but on examining **narcissistic traits at a subclinical level** — the kinds of tendencies and behavioral patterns that may appear in the **general population** without reaching the severity of a personality disorder.

The goal is to identify **tendencies, commonalities, and facets** that align with narcissistic behavior, and to evaluate how frequently they appear in large personality datasets. This approach allows us to treat narcissism as a **spectrum of traits** rather than a categorical disorder, making it possible to explore correlations with the Big Five model in a rigorous, data-driven way.

For readers who wish to expand their understanding of NPD in its clinical sense, recommended resources include the **DSM-5-TR (American Psychiatric Association)**, the **ICD-11 (World Health Organization)**, and the **Mayo Clinic's overview of Narcissistic Personality Disorder**. Academic reviews by Elsa Ronningstam (Harvard/McLean Hospital) and studies compiled in the *Journal of Personality Disorders* also provide detailed insights. These resources offer the clinical depth that lies beyond the scope of this project.

Step 1. Identifying the Traits of Narcissism with NPD as the base.

Narcissistic Personality Disorder (DSM-5 Overview)

The **DSM-5** defines Narcissistic Personality Disorder (NPD) as a pervasive pattern of **grandiosity, need for admiration, and lack of empathy**. A diagnosis requires meeting at least **five of nine criteria**, which include:

1. **Grandiose self-importance** (exaggerating achievements, expecting recognition).
2. **Fantasies of success or power** (absorbed in unlimited achievement or beauty).
3. **Belief of being "special"** (only associates with high-status people).
4. **Excessive need for admiration** (seeking constant praise).
5. **Sense of entitlement** (expects special treatment).
6. **Exploitative behavior** (taking advantage of others).
7. **Lack of empathy** (inability/unwillingness to recognize others' needs).
8. **Envy** (either envying others or believing others envy them).
9. **Arrogant, haughty behavior** (condescending or superior attitudes).

While these define the clinical disorder, **many people in the general population display some of these traits without meeting diagnostic thresholds**.

Clinically, NPD is often associated with **fragile self-esteem**, **relationship conflicts**, and **comorbid conditions** such as depression, substance use, or other personality disorders. Individuals may function well outwardly but experience significant **interpersonal and occupational impairment** due to entitlement, arrogance, and empathy deficits.

Table 1. DSM-5 NPD Table of Criteria and Related Traits.

The table lists the **nine DSM-5 criteria for NPD** with the **traits and behaviors linked to each**, highlighting observable tendencies that can later be mapped to Big Five facets:

DSM-5 Diagnostic Criteria	Related Traits / Behaviors
1. Grandiose self-importance	Exaggerating achievements; expecting recognition; arrogance; boastfulness
2. Fantasies of unlimited success/power	Daydreaming about fame, brilliance, beauty, or ideal love; unrealistic goal setting
3. Belief of being "special" and unique	Associating only with high-status people; devaluing "ordinary" others; superiority complex
4. Requires excessive admiration	Constant need for praise; fishing for compliments; attention-seeking
5. Sense of entitlement	Expecting special treatment; impatience when denied privileges; entitlement-driven anger
6. Interpersonally exploitative	Manipulativeness; using others for personal gain; discarding people once "unuseful"
7. Lack of empathy	Indifference to others' feelings; inability/unwillingness to recognize needs; callousness
8. Envy of others / belief others envy them	Jealousy; resentment of others' success; presuming others are jealous of them
9. Arrogant, haughty behaviors/attitudes	Condescending speech; belittling others; pretentiousness; superiority in demeanor

Step 2. Define The Anchor Narcissistic Traits As Adjectives.

In this step, we expand the DSM-5 diagnostic criteria for narcissistic personality disorder into a working set of **anchor adjectives** that capture the observable traits most often linked to narcissism, such as *grandiose, arrogant, boastful, entitled, manipulative,* or *haughty*. These anchors serve as the practical language for our analysis, bridging the clinical descriptions to everyday personality tendencies. We then proceed to evaluate their correspondence with the **IPIP-NEO-120 facets**, which measure finer aspects of the Big Five model. Each facet will be

judged as either closely related, opposed, or neutral in relation to these narcissistic adjectives, providing a structured way to identify where narcissistic traits align with or diverge from established personality dimensions, to consider broader and better adjectives, a large language model was used.

Table 2. Anchor Adjectives for DSM-5 Narcissistic Traits

DSM-5 Criterion	**Anchor Adjectives (refined)**
Grandiose self-importance	grandiose, arrogant, boastful, conceited, self-important
Fantasies of unlimited success/power	vainglorious, glory-seeking, triumphalist, fantasizing, visionary
Belief of being “special”/unique	elitist, superior, snobbish, pretentious, status-conscious
Requires excessive admiration	attention-seeking, praise-hungry, exhibitionistic, showy, approval-seeking
Sense of entitlement	entitled, demanding, presumptuous, overbearing, expectant
Interpersonally exploitative	exploitative, manipulative, opportunistic, calculating, predatory
Lack of empathy	unempathetic, callous, indifferent, cold, unsympathetic
Envy / belief others envy them	envious, jealous, resentful, covetous, begrudging
Arrogant / haughty behavior	haughty, condescending, pretentious, supercilious, disdainful

Big Five Dimensions and Facets rated by their possible relation with the Narcissistic Traits.

In order to connect the DSM-5 narcissistic traits (operationalized as anchor adjectives) with the Big Five personality model, we evaluated each **IPIP-NEO-120 facet** against the adjective sets. Initially, we considered using an NLP-based approach (embedding vectors and cosine similarity) to identify proximity between trait descriptors. However, this method produced unstable results, often placing antonyms (e.g., *arrogant* and *humble*) near each other due to their shared linguistic context. Because our project aims for clarity, reliability, and interpretability rather than technical novelty, we **discarded vector similarity as the primary tool**.

Instead, we used a **language-based evaluation with a large language model (LLM)** to analyze the semantic relationship between each facet definition and the narcissistic anchor adjectives. Each facet was judged according to a standardized rubric ranging from **+2 (strongly related)** to **−2 (strongly opposed)**, with 0 indicating no clear relationship. This approach allowed us to capture nuanced relationships between constructs while ensuring consistency and transparency in the scoring process.

The result is a **facet–trait mapping matrix** that identifies which facets, when high or low, are most aligned with narcissistic tendencies. This mapping provides the conceptual bridge between clinical traits (DSM-5 criteria) and psychometric data (IPIP-NEO-120), making it possible to analyze the frequency and distribution of narcissistic tendencies in the general population through Big Five personality results.

Table 3. Narcissistic Traits ↔ IPIP-NEO-120 Facets Matrix

Domain	Facet	Anchors Group	Score	Rationale
Neuroticism	Anxiety	grandiose_self_importance	0	Anxiety involves fear and tension, not strongly related to grandiosity.
Neuroticism	Anxiety	excessive_admiration	0	Needing admiration is not the same as experiencing fear; neutral relationship.
Neuroticism	Anger	envy	+1	High Anger overlaps with resentment and bitterness found in envy/jealousy.
Neuroticism	Anger	sense_of_entitlement	+1	Feeling cheated aligns with entitlement-driven resentment.
Neuroticism	Depression	grandiose_self_importance	−1	Depression (low energy, discouragement) opposes inflated self-importance.
Neuroticism	Depression	fantasies_success_power	−1	Sadness and low drive oppose success/power preoccupation.
Neuroticism	Self-Consciousness	excessive_admiration	+1	Fear of rejection overlaps with admiration-seeking tendencies.
Neuroticism	Self-Consciousness	haughty_behavior	−1	Shyness/shame conflicts with arrogant presentation.
Neuroticism	Immoderation	exploitative_behavior	+1	Poor impulse control supports manipulative or opportunistic behaviors.
Neuroticism	Immoderation	excessive_admiration	+1	Craving and urges resonate with praise-hungry behaviors.
Neuroticism	Vulnerability	grandiose_self_importance	−1	Panic/helplessness opposes confident, grandiose presentation.
Neuroticism	Vulnerability	haughty_behavior	−1	Helplessness contrasts with arrogant, disdainful attitudes.
Extraversion	Friendliness	haughty_behavior	−1	Warmth conflicts with arrogant, condescending traits.
Extraversion	Friendliness	lack_of_empathy	−2	Friendly engagement strongly opposes callousness.
Extraversion	Gregariousness	excessive_admiration	+1	Enjoying crowds aligns with attention-seeking.
Extraversion	Assertiveness	excessive_admiration	+1	Leadership overlaps with admiration-seeking tendencies.
Extraversion	Assertiveness	grandiose_self_importance	+1	Commanding presence supports grandiose self-image.
Extraversion	Activity Level	fantasies_success_power	+1	Energetic activity resonates with ambitious drives.
Extraversion	Activity Level	grandiose_self_importance	+1	High energy aligns with self-importance.
Extraversion	Excitement-Seeking	excessive_admiration	+1	Thrill-seeking overlaps with exhibitionistic traits.
Extraversion	Excitement-Seeking	exploitative_behavior	+1	Risk-taking facilitates opportunistic behavior.
Extraversion	Cheerfulness	haughty_behavior	−1	Positive affect contrasts with disdainful traits.
Extraversion	Cheerfulness	lack_of_empathy	−1	Upbeat mood opposes callousness.
Openness	Imagination	fantasies_success_power	+1	Fantasy overlaps with preoccupation with success/power.
Openness	Imagination	grandiose_self_importance	+1	Elaborate fantasies reinforce self-importance.
Openness	Artistic Interests	excessive_admiration	0	Interest in art is neutral to admiration-seeking.
Openness	Emotionality	lack_of_empathy	−1	Emotional awareness opposes callousness.
Openness	Adventurousness	fantasies_success_power	+1	Novelty-seeking aligns with visionary fantasizing.
Openness	Intellect	grandiose_self_importance	+1	Love of ideas supports superiority/elitism.
Openness	Intellect	special_unique	+1	Abstract thinking fosters uniqueness beliefs.
Openness	Liberalism	special_unique	+1	Challenging authority supports elitist attitudes.
Openness	Liberalism	entitlement	+1	Hostility to rules overlaps with entitlement.
Agreeableness	Trust	exploitative_behavior	−1	High Trust conflicts with manipulativeness.
Agreeableness	Trust	lack_of_empathy	−1	Low Trust relates to cynicism/callousness.
Agreeableness	Morality	exploitative_behavior	−2	High Morality rejects deception; low Morality aligns with manipulation.
Agreeableness	Morality	haughty_behavior	−1	Low Morality supports arrogant traits.
Agreeableness	Altruism	lack_of_empathy	−2	Altruism opposes callousness.

Agreeableness	Altruism	exploitative_behavior	−1	Low Altruism supports opportunism.
Agreeableness	Cooperation	entitlement	−1	Low Cooperation aligns with entitlement.
Agreeableness	Cooperation	haughty_behavior	−1	Low Cooperation fits arrogance/disdain.
Agreeableness	Modesty	haughty_behavior	−2	Modesty opposes arrogance.
Agreeableness	Modesty	grandiose_self_importance	−2	Modesty opposes grandiosity.
Agreeableness	Sympathy	lack_of_empathy	−2	Sympathy opposes indifference/callousness.
Agreeableness	Sympathy	envy	−1	Low Sympathy can foster resentment/envy.
Conscientiousness	Self-Efficacy	grandiose_self_importance	+1	Confidence overlaps with self-importance.
Conscientiousness	Orderliness	exploitative_behavior	−1	Order opposes opportunistic manipulation.
Conscientiousness	Dutifulness	exploitative_behavior	−2	Dutifulness opposes manipulation; low scores align with exploitation.
Conscientiousness	Dutifulness	entitlement	−1	Low Dutifulness overlaps with entitlement.
Conscientiousness	Achievement-Striving	fantasies_success_power	+1	High drive aligns with success fantasies.
Conscientiousness	Achievement-Striving	grandiose_self_importance	+1	Achievement orientation reinforces self-importance.
Conscientiousness	Self-Discipline	excessive_admiration	−1	Self-discipline opposes impulsive attention-seeking.
Conscientiousness	Cautiousness	exploitative_behavior	−1	Deliberation opposes opportunism/manipulation.
Conscientiousness	Cautiousness	excessive_admiration	−1	Cautiousness resists impulsive showy behavior.

Analyzing the facets and selecting the strongest facets as the best indicator for analysis (±2)

To maximize **construct validity** and **specificity**, the primary analysis uses only the facets with a **strong theoretical link** to narcissistic traits (rubric **±2**). Moderate indicators (±1) are common across many non-narcissistic profiles and therefore inflate **false positives**—especially in large, heterogeneous samples with **multimodal** score distributions. By restricting the core signal to Modesty, Sympathy, Altruism, Morality (Agreeableness) and Dutifulness (Conscientiousness)—plus the inverse Friendliness marker—we privilege facets that most directly capture **arrogance/grandiosity, callousness, self-centeredness, manipulativeness, and rule-disregard**.

This choice also improves **interpretability** and **robustness**. In multimodal data, “high/low” cannot be anchored to a single mean; relying on weaker associations would make thresholds arbitrary and unstable. Strong facets, evaluated with **percentile-based** thresholds and aligned direction (low = more narcissistic for inverse markers), yield a cleaner, defensible **Core Narcissistic Alignment Index** that is less sensitive to distribution shape and subgroup composition.

Moderate indicators are retained as **secondary context** only—used to describe profiles **after** the core index is elevated (e.g., high Assertiveness or Achievement-Striving at extreme percentiles). This preserves nuance without diluting the primary signal, and keeps the analysis firmly about **trait alignment**, not diagnosis.

Step 3: From Theoretical Indicators to Empirical Scales

In our analysis we identified **moderate and strong indicators** of narcissistic traits across Big Five facets, with rubric scores of ±2 representing the strongest theoretical relationships. However, a challenge emerges: our rubric alone does not tell us what facet scores should be considered “high” or “low.” For example, if Modesty is rated −2, narcissistic tendencies appear at

the *low* end of the scale, but what range of scores actually counts as “low” in practice? Similarly, facets rated ±1 may become stronger indicators when individuals score at the very extremes of the scale.

Because the IPIP-NEO-120 reports facet scores on a 0–100 scale derived from multiple 0–5 Likert items, scores are **bounded** and classical outliers are not expected. “High” and “low” therefore need to be defined **empirically** (via the sample’s distributions and percentiles), not by arbitrary cutoffs. However, because each facet score is an **average across several items** and then rescaled, an exact **0 or 100** implies endorsing the extreme response on **every** item in that facet—a statistically rare pattern that often reflects response style (straight-lining) rather than true trait extremity. Accordingly, we will **flag 0 or 100 as potential data-quality anomalies** (not automatically errors), review them, and run sensitivity analyses with and without these cases.

To resolve this, we will carry out an **exploratory data analysis (EDA)** for all facets. This includes calculating descriptive statistics (mean, median, standard deviation, min, max), visualizing score distributions, and determining percentile thresholds (e.g., 25th and 75th percentiles, or stricter 10th and 90th). We will also standardize scores using z-scores to allow comparisons across facets with different spreads. By combining these empirical thresholds with our theoretical rubric (±2/±1), we will generate **adjusted indicator values** that reflect both the strength of the theoretical link and the actual extremity of an individual’s score in the dataset.

This approach ensures that our analysis is both **theoretically anchored** and **data-driven**, allowing us to identify which facets truly signal narcissistic tendencies in the general population.

STEP 4. EDA in Python

Exploratory Data Analysis. We will determine the mean and the standard deviation of each facet in order to establish what are scores considered high and low, specially for those facets that were highly or moderately related to Narcissism.

You can see the Jupyter Notebook for the comments on each basic step for setting up the DB.

1. Cleaning the data and setting up the dataframe.

Standard practices for importing the data and setting up the dataframe.

2. Age filter (pre-processing step)

As a first data-quality step, we excluded respondents under 18 years of age. Personality trait norms and interpretability for the IPIP-NEO are calibrated primarily on adult samples; including minors would mix developmental effects with our target construct and could bias facet distributions. Therefore, all records with age < 18 were removed prior to any analyses (descriptives, percentiles, correlations, or index construction). We report the number and proportion of excluded cases in the sample characteristics table and, as a robustness check,

verified that key results are unchanged when the under-18 filter is not applied. Our database shape after excluding those cases is: 339711, 38

3. Flagging extreme endpoints (0 or 100):

Rationale: on IPIP-NEO facets (averages of multiple 0–5 items rescaled to 0–100), an exact **0** or **100** implies endorsing the extreme option on **every** item—statistically rare and often a response-style artifact (straight-lining).

We verified whether any IPIP-NEO facet scores hit the scale endpoints (exact **0** or **100**), which can indicate uniform extreme responding across items. **No endpoint values were found** in our dataset. Accordingly, no facet-level winsorization or exclusions were required for this step, and we proceeded with the full sample for subsequent analyses.

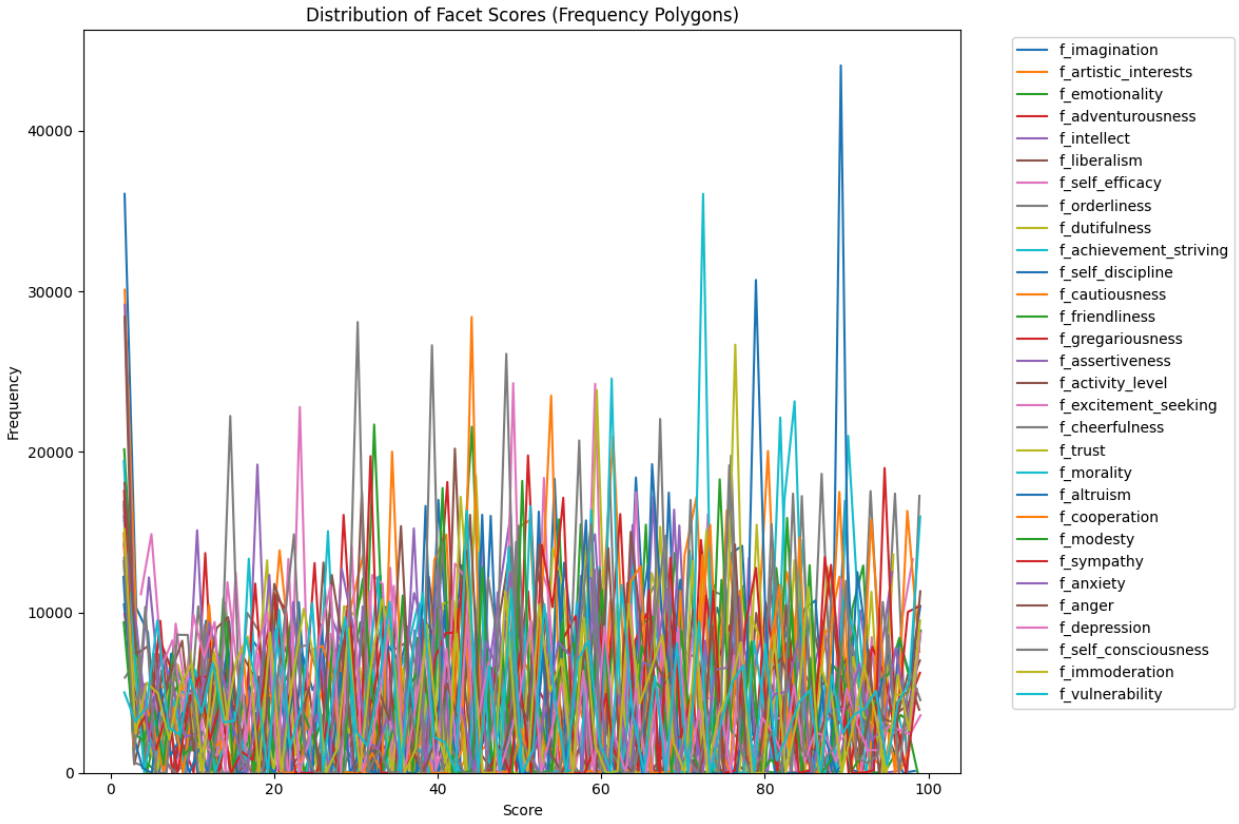
4. Plotting

Upon inspecting the facet and domain score distributions (0–100 scale), we observed **multimodal** patterns rather than single bell curves. This is expected with Likert-derived (0–5) item averages rescaled to 0–100 in a large, heterogeneous sample. Multiple peaks likely arise from:

- **Population subgroups** (e.g., age, gender, culture/language) that cluster at different ranges,
- **Ceiling/floor response preferences** on certain traits, and
- **Item-keying/rescaling artifacts** that create discrete “clumps” after aggregation.

Implication: Assuming a single normal distribution for each facet would be inappropriate. Relying on parametric assumptions (e.g., z-scores that presume unimodality) can **distort the identification of “extremes”** and **bias correlations**.

Figure 1. Distribution of Facet Scores



5. Plotting only the 6 strongest indicators according to step 2.

Primary indicators (±2 only)

These are the facets with **the clearest, theory-consistent signal**. For narcissistic alignment, we look for **low** scores on these (because high opposes narcissism):

- **Agreeableness → Modesty (−2)** → low Modesty = arrogance/grandiosity
- **Agreeableness → Sympathy (−2)** → low Sympathy = callousness/lack of empathy
- **Agreeableness → Altruism (−2)** → low Altruism = self-centeredness
- **Agreeableness → Morality (−2)** → low Morality = manipulativeness/deception
- **Conscientiousness → Dutifulness (−2)** → low Dutifulness = rule-disregard/exploitation
- **Extraversion → Friendliness (−2)** (vs lack of empathy) → low Friendliness = interpersonal coldness

One of the findings was that altruism is the facet with the highest frequency within high scores, followed by morality. We can also see that Agreableness has four of the 6 strongly related facets, which can indicate that a low agreableness score might be an indicator of narcissistic traits.

Figure 2. Distribution of Selected Facet Scores - Frequency Polygons.

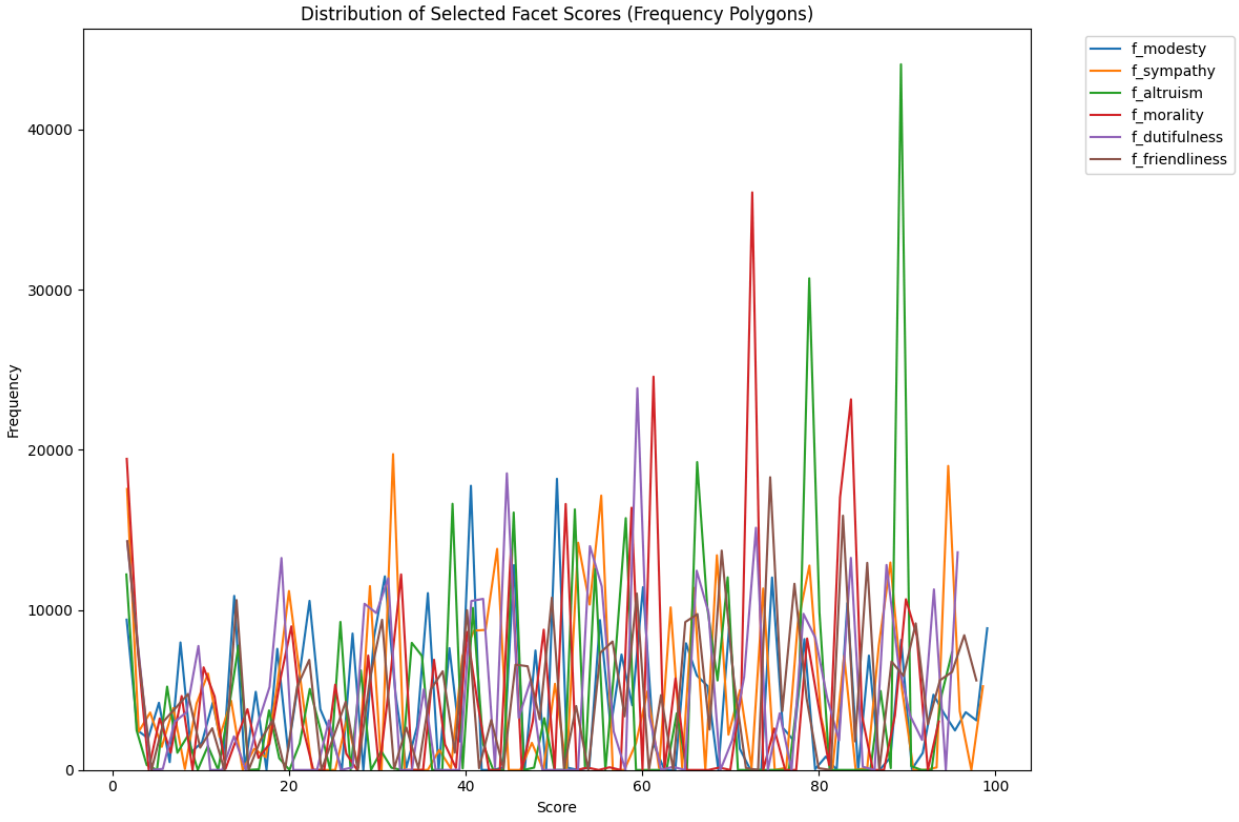


Figure 3. Distribution of Selected Facet Scores - KDE Plots

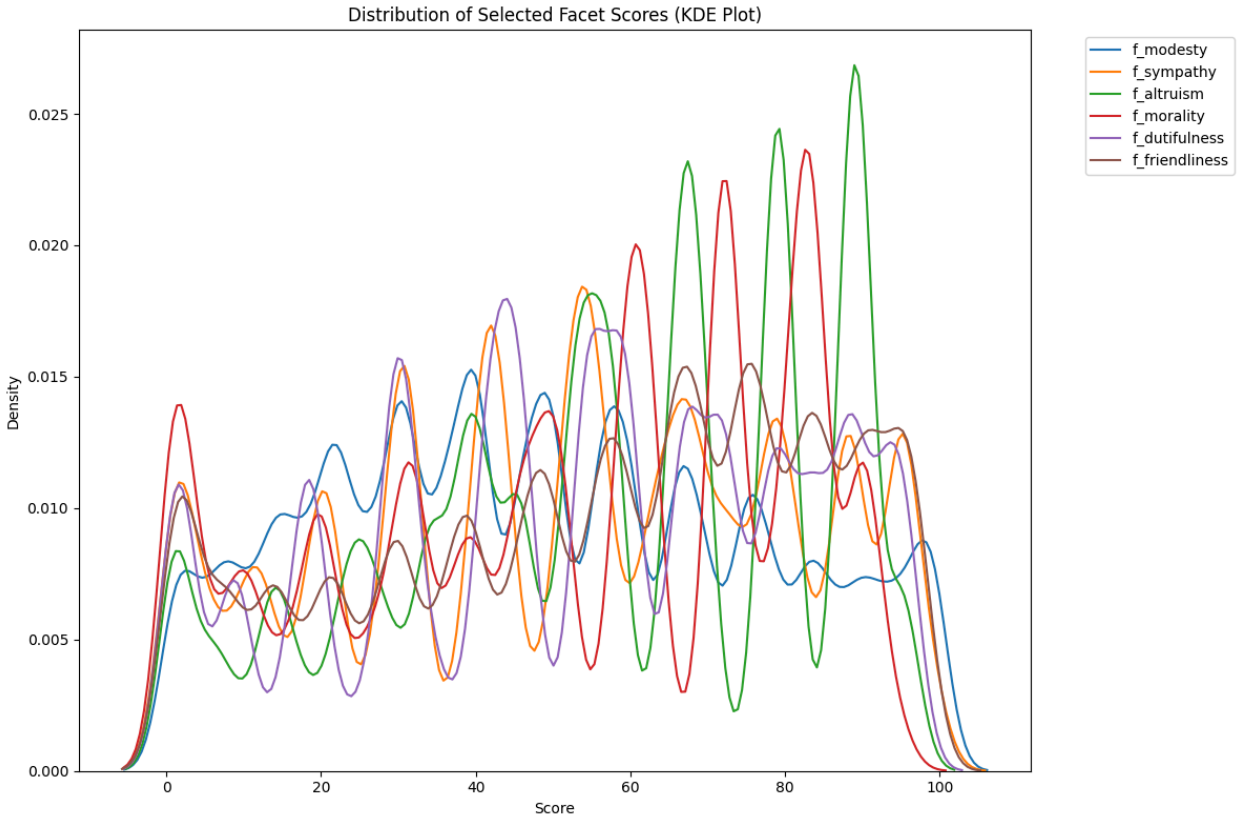
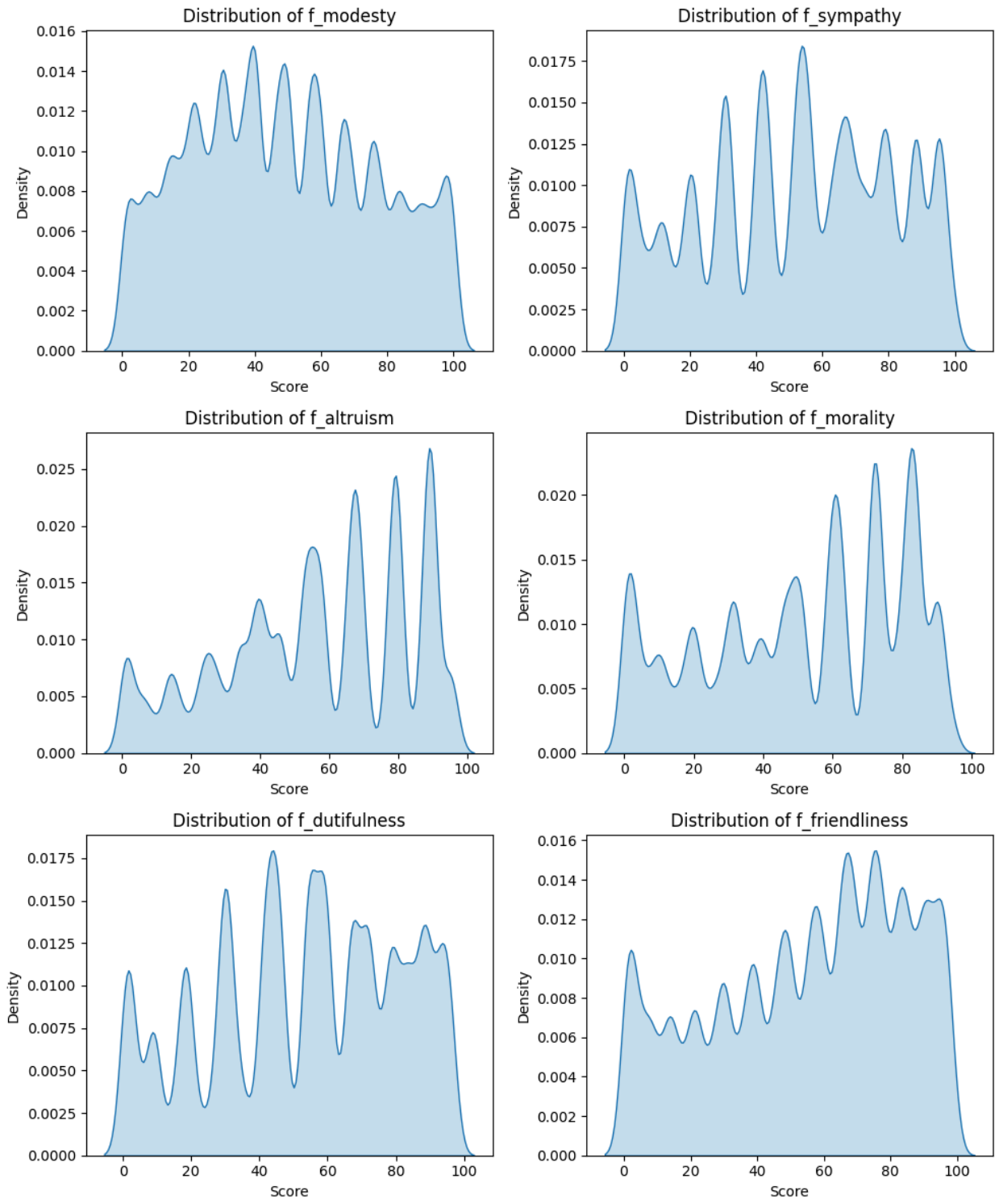
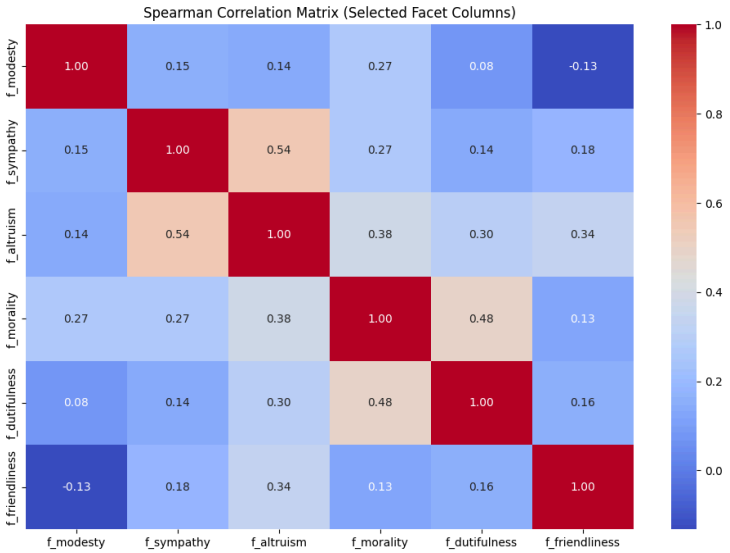


Figure 4. Individual KDE Plots per Selected Facet.



From the plots it seems there's a correlation between some of them, because our variables are categorical and multimodal we will use Spearman's p.

Figure 5. Spearman Correlation Matrix.



Because of the strong visual similarities, we also calculated the skewedness, confirming the similarities:

Skewness of f_modesty: 0.1372051480331306

Skewness of f_sympathy: -0.17197947732724284

Skewness of f_altruism: -0.3997916872620112

Skewness of f_morality: -0.3379083396829463

Skewness of f_dutifulness: -0.2647420275496774

Skewness of f_friendliness: -0.36927620475064843

Remember that:

- **Skewness near 0:** Indicates a more symmetrical distribution, where the left and right sides are roughly mirror images.
- **Positive skewness:** Means the tail of the distribution is longer on the right side. The bulk of the data is concentrated on the left.
- **Negative skewness:** Means the tail of the distribution is longer on the left side. The bulk of the data is concentrated on the right.

Looking at the skewness values and the corresponding KDE plots:

- **f_modesty (Skewness: 0.137):** This value is close to 0, indicating a relatively symmetrical distribution. The KDE plot for f_modesty should appear somewhat balanced around its center.
- **f_sympathy (Skewness: -0.172):** This value is slightly negative. The KDE plot should show a slightly longer tail on the left side, with the peak shifted slightly towards the right.
- **f_altruism (Skewness: -0.399):** This is a more pronounced negative skewness. The KDE plot for f_altruism should clearly show the distribution skewed towards the left, with a longer tail extending to the left.
- **f_morality (Skewness: -0.338):** Similar to f_altruism, this negative skewness suggests a left-skewed distribution in the KDE plot, with the bulk of the data on the right.
- **f_dutifulness (Skewness: -0.265):** Another negative skewness value, indicating a left-skewed distribution in the KDE plot.
- **f_friendliness (Skewness: -0.369):** This negative skewness points to a left-skewed distribution in the KDE plot, with a longer tail on the left.

In summary, the skewness values confirm what we visually observed in the KDE plots: f_modesty is roughly symmetrical, while the other selected facets show varying degrees of negative skew, meaning more people tend to score higher on these traits than lower.

What this means for our analysis

- **Low scores on five facets are rarer** (because of negative skew). When they *do* occur, they're more diagnostic. A **percentile approach** should handle this—"low" (≤P10/P20) automatically adapts to each facet's shape.

- **Low Modesty will be more common** than low Sympathy/Altruism/etc., so Modesty could dominate if we're not careful.

Skewness confirms the design choices—**rank-based thresholds and a multi-facet view**. Low scores on Sympathy/Altruism/Morality/Dutifulness/Friendliness are **rarer and more telling**; low Modesty is **more common** and should be interpreted in context rather than alone.

Median/mode/percentiles calculations

Our facet scores are **bounded** (0–100), derived from **categorical Likert items (0–5)**, and the **empirical distributions are multimodal**. Under these conditions, **parametric summaries** that assume an approximately normal, unimodal distribution (e.g., **mean** and **standard deviation**) can be **misleading**: the mean is pulled by uneven mode weights, SD overstates/understates spread when there are multiple peaks, and z-scores implicitly impose a single-bell-curve model that the data do not follow.

Accordingly, we adopt **non-parametric, shape-robust summaries**:

- **Median** (50th percentile) as the measure of central tendency,
- **Mode(s)** to acknowledge multiple high-frequency regions,
- **Percentiles** (e.g., P10/P25/P75/P90) to define “low” and “high” without assuming normality,
- **Rank-based associations** (Spearman’s ρ) instead of Pearson correlations.

Table 4. Mode, Median and Percentiles per Facet.

Facet	Mode	Median	P25	P50	P75
modesty	40.29	47.84	26.63	47.84	69.65
sympathy	55.94	53.76	31.27	53.76	77.22
altruism	89.45	57.65	38.69	57.65	79.05
morality	83.41	58.96	29.01	58.96	72.42
dutifulness	44.96	54.89	31.36	54.89	78.31
friendliness	68.55	59.19	30.54	59.19	78.80

Figure 6. Boxplots.

Generating boxplots:

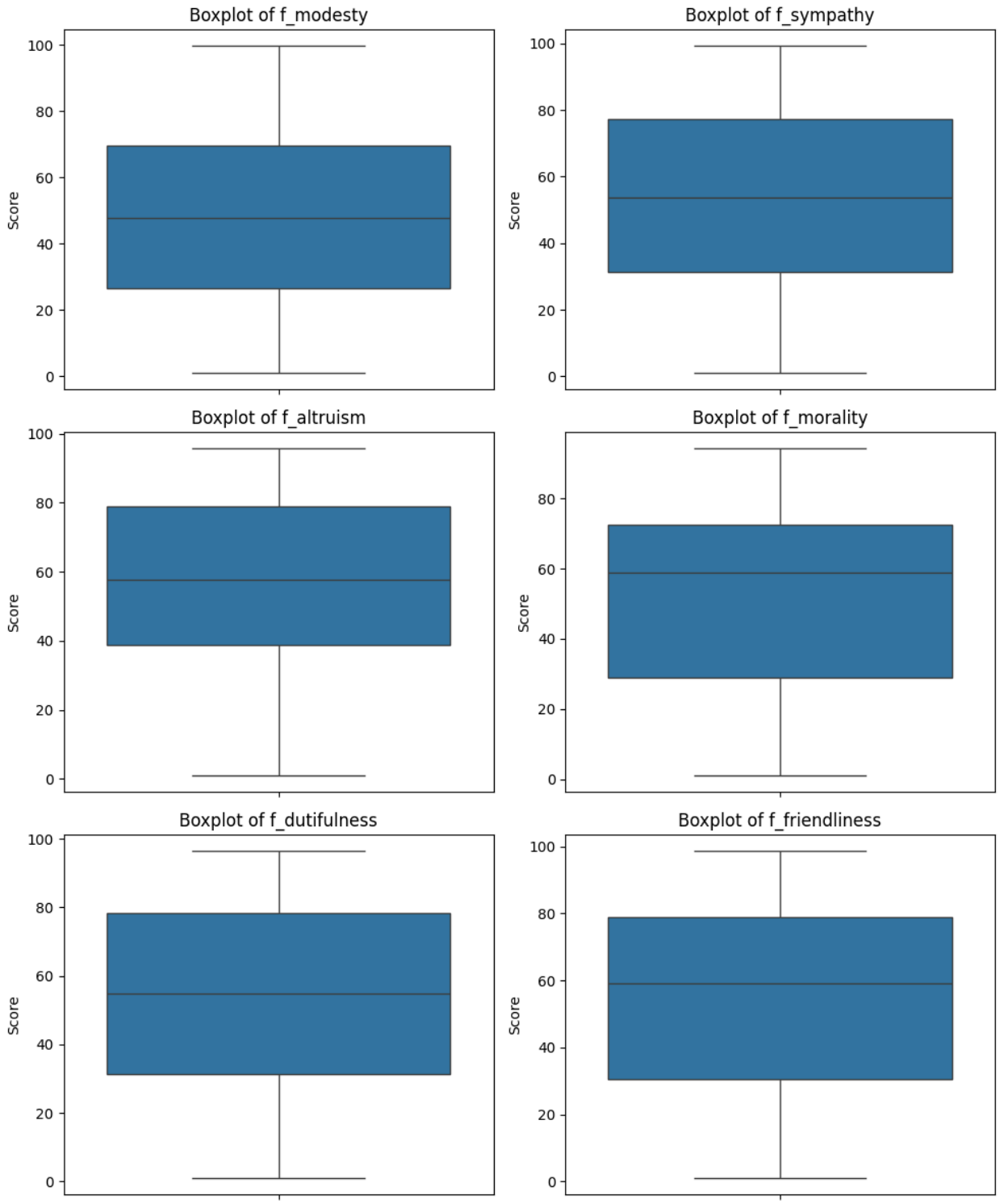
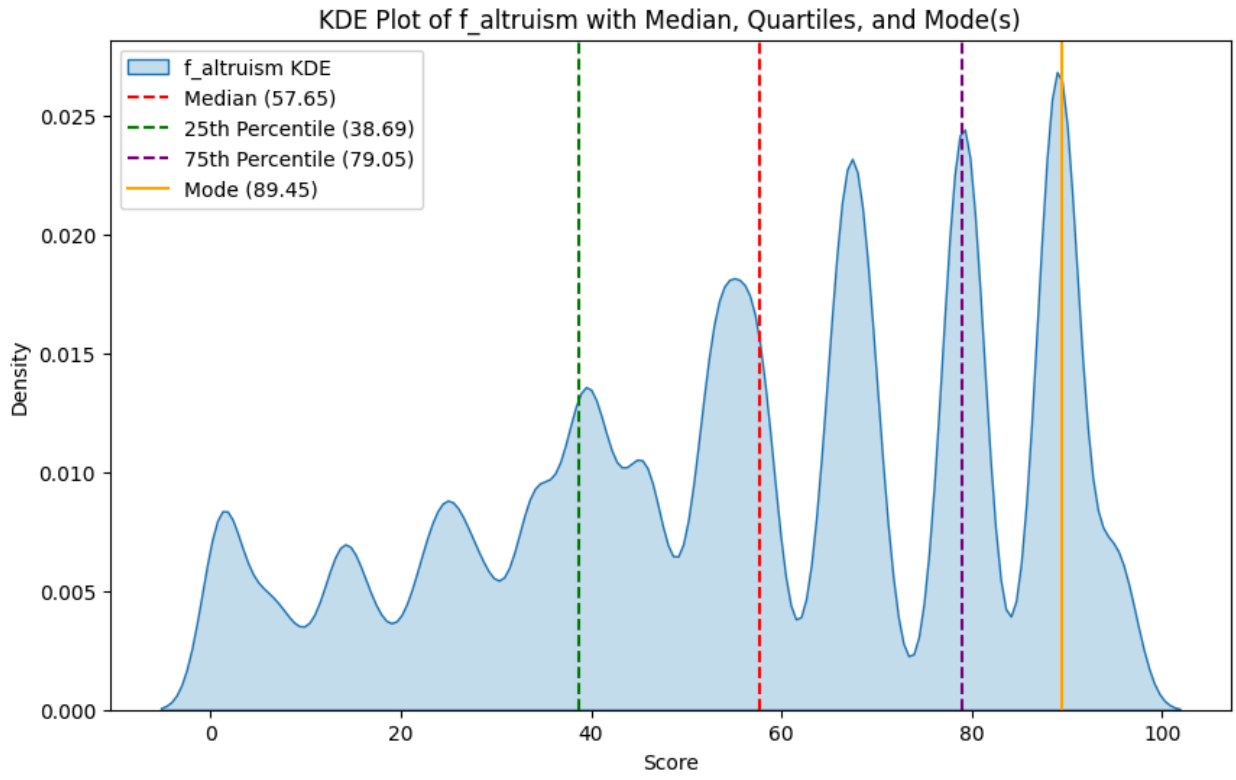
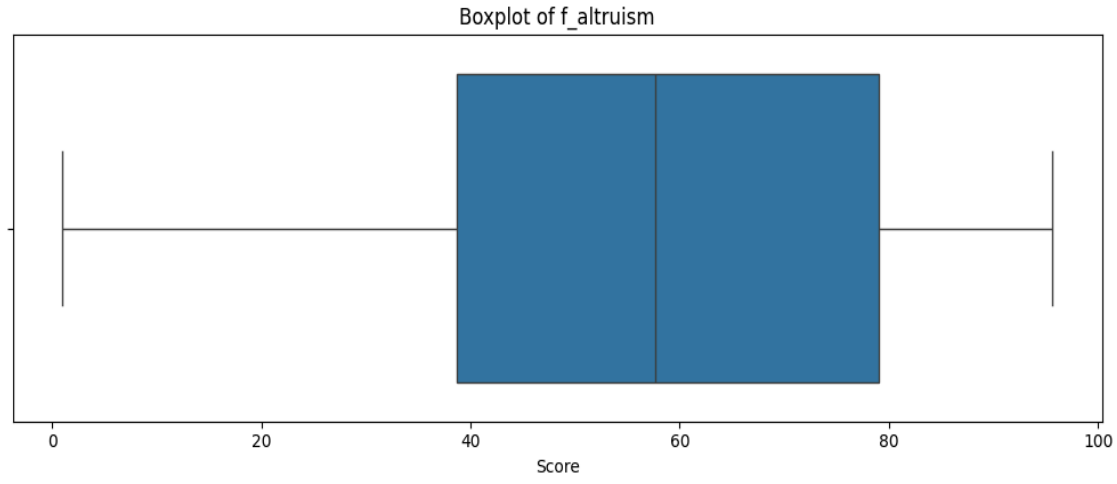


Figure 7. KDE and Boxpot for Altruism.





Defining “Low” Facet Scores and Opposite Traits

To determine when a facet reflects the *opposite* pole of its construct (and therefore aligns with narcissistic tendencies), we follow conventions from validated Big Five instruments and personality research. Both the **IPIP-NEO** and the **NEO PI-R/NEO-PI-3** manuals classify **the lowest 20% of scores** as “low” and label them with the facet’s opposite trait (e.g., Modesty → Arrogance) . Research studies likewise often define “very low” trait groups as the bottom **10–20%** of the distribution . Accordingly, in this project, we define **low = ≤20th percentile**. Scores in this range are interpreted as expressing the opposite trait descriptors listed below.

- **Modesty:** Low scorers are described as **conceited or arrogant**, believing themselves superior to others.
- **Sympathy (Tender-Mindedness):** Low scorers are **hard-headed, unsentimental, and indifferent** to others’ suffering.
- **Altruism:** Low scorers are **self-centered or egocentric**, reluctant to help or concerned mainly with their own needs.
- **Morality (Straightforwardness):** Low scorers are **manipulative or deceitful**, comfortable bending rules or flattery to achieve their aims.
- **Dutifulness:** Low scorers are **undependable or irresponsible**, neglecting rules, promises, or obligations.
- **Friendliness (Warmth):** Low scorers are **aloof, distant, or unaffectionate**, showing formality rather than genuine warmth toward others.

Thus, in our analysis, an individual facet is considered **“narcissistically aligned”** if it falls within the **lowest 20% of the sample distribution**. These low-end labels (arrogant, indifferent, selfish, manipulative, irresponsible, aloof) provide the trait-opposite anchors that operationalize narcissism within the Big Five framework.

Defining Low Facet Scores in Practice

Following conventions from the **NEO PI-R** and **IPIP-NEO**, we classify the **lowest quintile (≤20th percentile)** of each facet distribution as “low.” This avoids arbitrary raw cutoffs (e.g., 20 on a 0–100 scale) and instead uses the **empirical sample distribution**. For our dataset, the 20th percentile thresholds for the six strongly related facets are:

Table 5. 20th Percentile Cutoff and Opposite Labels

Facet	20th Percentile Cutoff	Low-End Label (Opposite Pole)
Modesty	21.87	Arrogant / Conceited
Sympathy	21.65	Cold / Unsympathetic
Altruism	33.58	Selfish / Egocentric
Morality	20.04	Manipulative / Deceitful
Dutifulness	28.95	Irresponsible / Undependable
Friendliness	26.46	Aloof / Distant

Table 6. Percentage of population scoring lower than the lowest quintile for each selected facet:

Modesty	19.06%
Sympathy	18.86%
Altruism	19.69%
Morality	17.89%
Dutifulness	18.78%
Friendliness	19.55%

Interpreting Low Facet Prevalence

By defining “low” scores as the **lowest quintile (≤20th percentile)**, we operationalize the idea that roughly one in five people will fall into the “opposite pole” for each facet. In our dataset, the actual percentages range between **~18% and ~19.7%**, slightly below the ideal 20%. This small discrepancy arises because the facet scores are derived from **categorical Likert items** rescaled to 0–100, which creates clusters of tied values at specific points. When a percentile cut is applied to such distributions, the boundary may slice through a cluster of identical scores, yielding slightly fewer or more respondents below the threshold.

The key point is that, conceptually, **about 20% of the population expresses <u>at least one</u> of the low-end trait** for each of these facets. In practical terms:

- ~20% of individuals score low on **Modesty**, reflecting *arrogance or conceit*.
- ~20% score low on **Sympathy**, reflecting *indifference or lack of empathy*.
- ~20% score low on **Altruism**, reflecting *self-centeredness*.
- ~20% score low on **Morality**, reflecting *manipulativeness or deceitfulness*.
- ~20% score low on **Dutifulness**, reflecting *irresponsibility or unreliability*.
- ~20% score low on **Friendliness**, reflecting *aloofness or detachment*.

Thus, when we describe “low” facets as indicators of narcissistic alignment, it means that **approximately one-fifth of the population naturally falls into these categories** for each facet, even in a general, non-clinical sample. These individuals do not necessarily meet criteria for narcissistic personality disorder, but they express *trait-level opposites* that conceptually overlap with narcissistic tendencies.

Interpreting Low Scores Across Facets

It is important to clarify that a **low score in one facet** simply reflects the opposite pole of that *specific trait*. For example, an individual low in Modesty may appear **arrogant**, while another person low in Sympathy may seem **indifferent to others’ feelings**. By design, **about 20% of the population will fall in the lowest quintile for each facet**, but these groups are not the same people. Someone who is low in Modesty is not necessarily also low in Altruism, Morality, or the other facets.

This helps explain why narcissistic traits often appear so visible in everyday life: many people show *some* behaviors that resemble narcissism (arrogance, aloofness, selfishness, etc.), but these isolated traits do **not** imply that the person is narcissistic overall. Narcissistic Personality Disorder (NPD) involves a **pattern across multiple facets simultaneously**, not just one.

In fact, only individuals who score **consistently low across several of the six strongly related facets**—for example, showing arrogance (low Modesty), selfishness (low Altruism), manipulativeness (low Morality), and aloofness (low Friendliness) together—would begin to approximate a broader narcissistic profile. The proportion of people meeting that pattern is expected to be **much lower than 20%**.

Later in our analysis, we will quantify this explicitly by calculating how many individuals score in the lowest quintile for **more than one facet at once**. This will allow us to distinguish between the common occurrence of isolated traits and the rare convergence of multiple low facets that could indicate a more pronounced narcissistic alignment.

Illustrative Examples

- **Person A: Low in one facet only**
  Imagine someone who scores in the **lowest 20% on Modesty**, but average on Sympathy, Altruism, Morality, Dutifulness, and Friendliness. This person may come across as **arrogant or boastful**, yet they still show empathy, reliability, honesty, and warmth in other areas. Such a profile reflects a *single trait tendency*, not a narcissistic personality.

- **Person B: Low in multiple facets**
  Another individual might score **low on Modesty, Altruism, and Friendliness**. This combination could present as **arrogant, self-centered, and aloof**, a pattern that others might recognize as strongly narcissistic in everyday interactions. Still, unless this pattern extends consistently across most or all relevant facets, it does not indicate Narcissistic Personality Disorder.

- **Person C: Low across all six facets**
  A very rare case would be someone scoring **low on all six strongly related facets**: arrogant (low Modesty), unsympathetic (low Sympathy), selfish (low Altruism), manipulative (low Morality), irresponsible (low Dutifulness), and aloof (low Friendliness). This constellation of traits represents a **broad, pervasive alignment with narcissistic tendencies**, which might approximate what clinical definitions of Narcissistic Personality Disorder capture. However, the proportion of individuals showing such a profile in a general population dataset is expected to be **far below 20%**.

PERCENTAGE OF POPULATION SCORING BELOW THE LOWEST QUINTILE FOR COMBINATIONS OF SELECTED FACETS:

--- Combinations of low scores on 2 facets ---

- f_modesty, f_sympathy: 5.34%
- f_modesty, f_altruism: 5.35%
- f_modesty, f_morality: 6.33%
- f_modesty, f_dutifulness: 4.60%
- f_modesty, f_friendliness: 3.06%
- f_sympathy, f_altruism: 9.44%
- f_sympathy, f_morality: 5.69%
- f_sympathy, f_dutifulness: 4.87%
- f_sympathy, f_friendliness: 5.22%
- f_altruism, f_morality: 7.10%
- f_altruism, f_dutifulness: 6.60%
- f_altruism, f_friendliness: 7.16%
- f_morality, f_dutifulness: 8.31%
- f_morality, f_friendliness: 4.49%
- f_dutifulness, f_friendliness: 5.14%

--- Combinations of low scores on 3 facets ---

- f_modesty, f_sympathy, f_altruism: 3.07%
- f_modesty, f_sympathy, f_morality: 2.42%
- f_modesty, f_sympathy, f_dutifulness: 1.61%
- f_modesty, f_sympathy, f_friendliness: 1.23%
- f_modesty, f_altruism, f_morality: 2.76%

- f_modesty, f_altruism, f_dutifulness: 2.02%
- f_modesty, f_altruism, f_friendliness: 1.58%
- f_modesty, f_morality, f_dutifulness: 2.87%
- f_modesty, f_morality, f_friendliness: 1.25%
- f_modesty, f_dutifulness, f_friendliness: 1.00%
- f_sympathy, f_altruism, f_morality: 3.82%
- f_sympathy, f_altruism, f_dutifulness: 3.24%
- f_sympathy, f_altruism, f_friendliness: 3.59%
- f_sympathy, f_morality, f_dutifulness: 2.81%
- f_sympathy, f_morality, f_friendliness: 1.79%
- f_sympathy, f_dutifulness, f_friendliness: 1.68%
- f_altruism, f_morality, f_dutifulness: 3.91%
- f_altruism, f_morality, f_friendliness: 2.61%
- f_altruism, f_dutifulness, f_friendliness: 2.63%
- f_morality, f_dutifulness, f_friendliness: 2.44%

--- Combinations of low scores on 4 facets ---

- f_modesty, f_sympathy, f_altruism, f_morality: 1.69%
- f_modesty, f_sympathy, f_altruism, f_dutifulness: 1.16%
- f_modesty, f_sympathy, f_altruism, f_friendliness: 0.95%
- f_modesty, f_sympathy, f_morality, f_dutifulness: 1.17%
- f_modesty, f_sympathy, f_morality, f_friendliness: 0.62%
- f_modesty, f_sympathy, f_dutifulness, f_friendliness: 0.46%
- f_modesty, f_altruism, f_morality, f_dutifulness: 1.49%

- f_modesty, f_altruism, f_morality, f_friendliness: 0.82%

- f_modesty, f_altruism, f_dutifulness, f_friendliness: 0.64%

- f_modesty, f_morality, f_dutifulness, f_friendliness: 0.66%

- f_sympathy, f_altruism, f_morality, f_dutifulness: 2.09%

- f_sympathy, f_altruism, f_morality, f_friendliness: 1.47%

- f_sympathy, f_altruism, f_dutifulness, f_friendliness: 1.35%

- f_sympathy, f_morality, f_dutifulness, f_friendliness: 1.01%

- f_altruism, f_morality, f_dutifulness, f_friendliness: 1.52%

--- Combinations of low scores on 5 facets ---

- f_modesty, f_sympathy, f_altruism, f_morality, f_dutifulness: 0.90%

- f_modesty, f_sympathy, f_altruism, f_morality, f_friendliness: 0.53%

- f_modesty, f_sympathy, f_altruism, f_dutifulness, f_friendliness: 0.39%

- f_modesty, f_sympathy, f_morality, f_dutifulness, f_friendliness: 0.34%

- f_modesty, f_altruism, f_morality, f_dutifulness, f_friendliness: 0.47%

- f_sympathy, f_altruism, f_morality, f_dutifulness, f_friendliness: 0.86%

--- Combinations of low scores on all 6 facets ---

- f_modesty, f_sympathy, f_altruism, f_morality, f_dutifulness, f_friendliness: 0.30%

Overlap of Low Facet Scores

Because each facet defines “low” as the **lowest 20% of the population**, we would expect about one in five people to show the opposite pole of any given trait. However, the groups of low scorers on different facets are not identical. Our analysis of co-occurrence shows that while many people express one narcissistically aligned trait, far fewer express multiple traits at once.

- **Two-facet combinations:** Between **3–9% of the population** score low on *two facets simultaneously*. For example, **9.44%** scored low in both **Sympathy and Altruism**, suggesting that *indifference to others’ feelings* often co-occurs with *self-centeredness*.

Other pairs, such as **Modesty + Friendliness**, were much rarer (~3%).

- **Three-facet combinations:** These drop to roughly **1–4%**, depending on the traits. The strongest overlap was **Altruism + Morality + Dutifulness (3.9%)**, consistent with a small subgroup showing *self-centeredness, manipulativeness, and irresponsibility* together.

- **Four-facet combinations:** Now below **2%** across all trait sets. Even the most frequent cluster, **Sympathy + Altruism + Morality + Dutifulness**, appeared in just **2.1%** of the sample.

- **Five-facet combinations:** Rare, between **0.3–0.9%**, with the highest at **0.9%** for **Modesty + Sympathy + Altruism + Morality + Dutifulness**.

- **All six facets:** Only **0.3% of the population** scored low on *every* facet (arrogant, unsympathetic, selfish, manipulative, irresponsible, and aloof). This extremely small subgroup might approximate the type of broad pattern seen in Narcissistic Personality Disorder.

Interpretation.
These results show why narcissistic *traits* are commonly observed (each trait touches ~20% of the population), but true narcissistic *profiles* are rare. Isolated traits such as arrogance or aloofness are relatively widespread, but the **convergence of multiple low facets shrinks dramatically**: from ~20% for one trait, to ~5–10% for two, ~2–4% for three, under 2% for four, under 1% for five, and a fraction of a percent for all six. This gradient supports the distinction between *trait-level narcissism* in the general population and *disorder-level narcissism* that requires a pervasive, multi-faceted pattern.

Five-Facet Combinations and DSM-5 Parallels

Our data show that **five-facet combinations are extremely rare**, occurring in only **0.3–0.9% of the sample**. The most frequent cluster (0.9%) was the combination of **low Modesty, low Sympathy, low Altruism, low Morality, and low Dutifulness**. In descriptive terms, this profile corresponds to individuals who are simultaneously **arrogant, unsympathetic, selfish, manipulative, and irresponsible**.

This finding resonates with the **DSM-5 framework**, which specifies that a diagnosis of Narcissistic Personality Disorder requires **five or more out of nine diagnostic criteria**, plus evidence of functional impairment. While our approach is trait-based and non-diagnostic, the fact that a five-facet constellation appears in about **1% of the population** provides a clear conceptual parallel: a person must show **multiple, co-occurring low traits** to approximate the clinical definition of narcissism.

Trait-Level Indicators vs. DSM-5 Diagnosis

It is equally important to emphasize that our work analyzes **trait tendencies, not disorders**. In our project, we identified six “opposite facets” (arrogant, unsympathetic, selfish, manipulative, irresponsible, and aloof) that conceptually align with narcissistic traits. Individuals who score low on one or two of these facets simply display *isolated characteristics* that are common in the general population and do not imply Narcissistic Personality Disorder.

By contrast, **true NPD requires more than just low scores**. According to DSM-5, the diagnosis depends not only on a sufficient number of traits (five or more of the nine listed criteria), but also on **pervasiveness across contexts** and the presence of **clinically significant distress or impairment**. Our results illustrate this gap: while ~20% of the population scores low on any given facet, only **0.3% scored low on all six simultaneously**, a figure that approximates the reported prevalence of NPD (~0.5–1%).

Thus, our measure provides **trait-level insight into narcissistic tendencies**, but it cannot substitute for the clinical thresholds that distinguish everyday personality variation from personality disorder.

Conclusions

This project set out to examine how **narcissistic traits** appear in the general population by linking them to **Big Five personality facets**. Using the IPIP-NEO dataset, we identified six facets whose low expressions conceptually align with narcissistic tendencies: **Modesty, Sympathy, Altruism, Morality, Dutifulness, and Friendliness**. Their opposites — arrogance, lack of empathy, selfishness, manipulativeness, irresponsibility, and aloofness — capture the core behaviors that people often describe as “narcissistic.”

Our analysis showed that **about 20% of the population** scores low on any given facet, confirming that isolated narcissistic traits are **common and widespread** in everyday life. However, the **co-occurrence of multiple low facets drops sharply**:

- Low on two or three facets: relatively uncommon (3–9% and 1–4% respectively).
- Low on four or five facets: rare (<2% and <1%).
- Low on all six facets: extremely rare (0.3%).

These findings clarify an important distinction: while many people exhibit one or two narcissistic traits, very few display the **broad, pervasive pattern** that characterizes Narcissistic Personality Disorder (NPD). The **DSM-5** requires **five or more diagnostic criteria** plus evidence of **functional impairment**; our five-facet combinations (~0.3–0.9%) and six-facet group (~0.3%)

mirror the **estimated prevalence of NPD (0.5–1%)**. This convergence underscores the value of our trait-based approach while reinforcing that it is **not diagnostic**.

In conclusion, narcissistic traits are **statistically normal at the individual facet level**, which explains why they are so frequently observed in social life. What distinguishes clinical narcissism is not the presence of isolated traits, but their **breadth, co-occurrence, and impact on functioning**. By analyzing narcissistic tendencies through the lens of Big Five facets, this project provides both a **quantitative perspective on their distribution** and a framework for understanding the **gap between traits and disorder**.