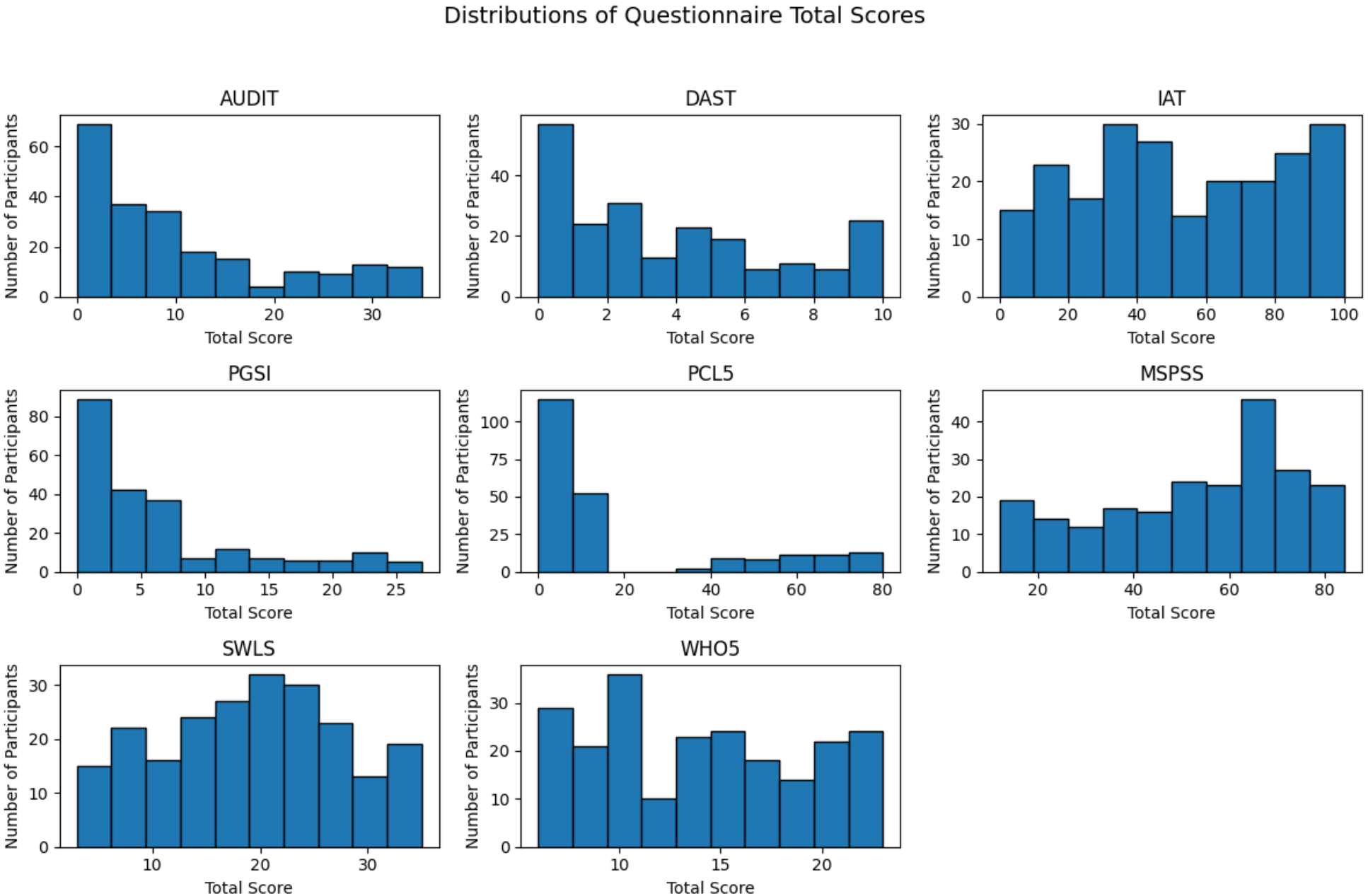
EDA Figures

Group 13

Data analysis and preparation


Distributions of Questionnaire Total Scores
AUDIT
DAST
IAT
PGSI
PCL5
MSPSS
SWLS
WHO5
Number of Participants
Total Score

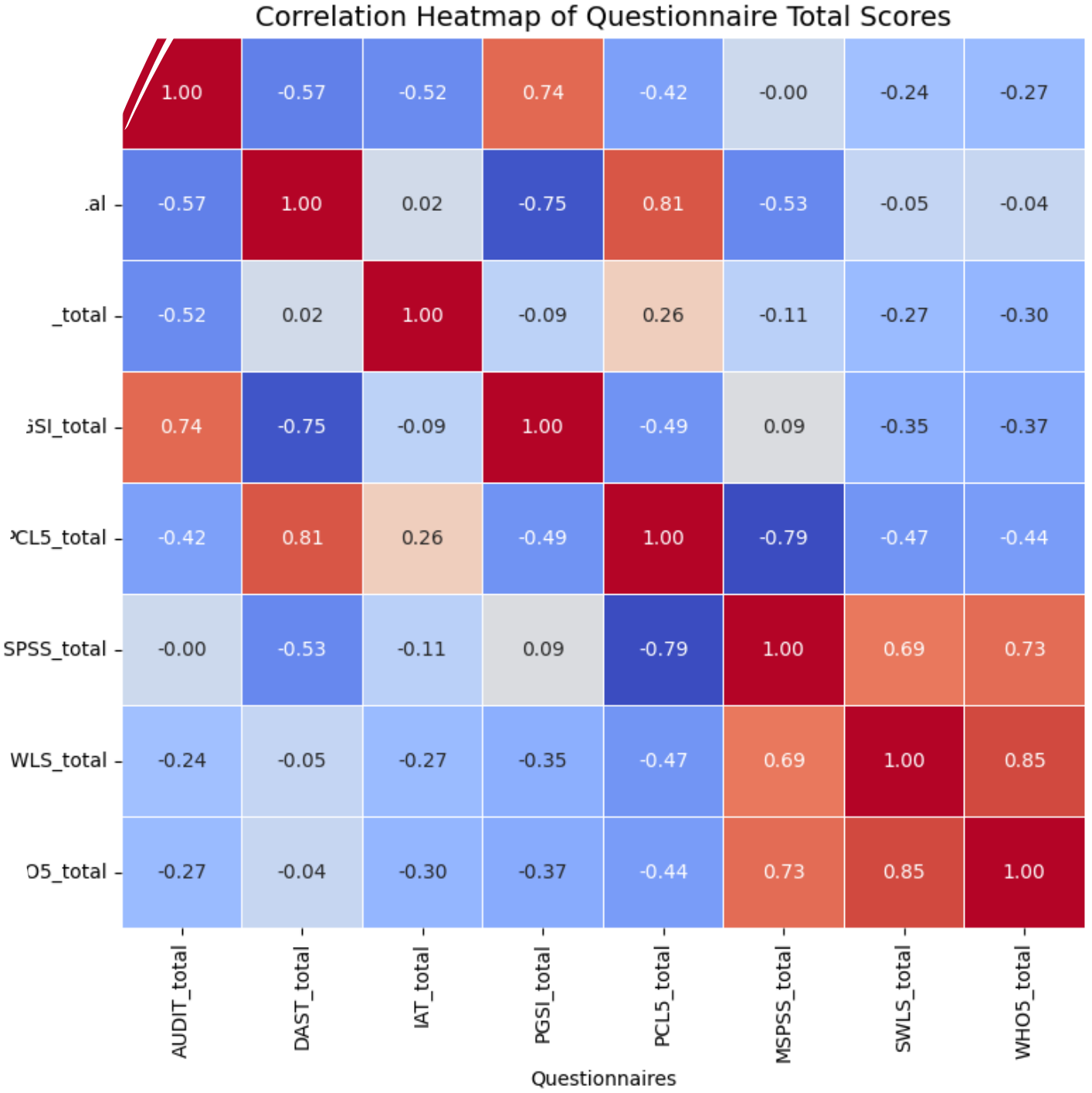
Heatmap of correlations

Co-occurrence of addictions: Alcohol and gambling problems tend to appear together (shared behavioral tendencies).

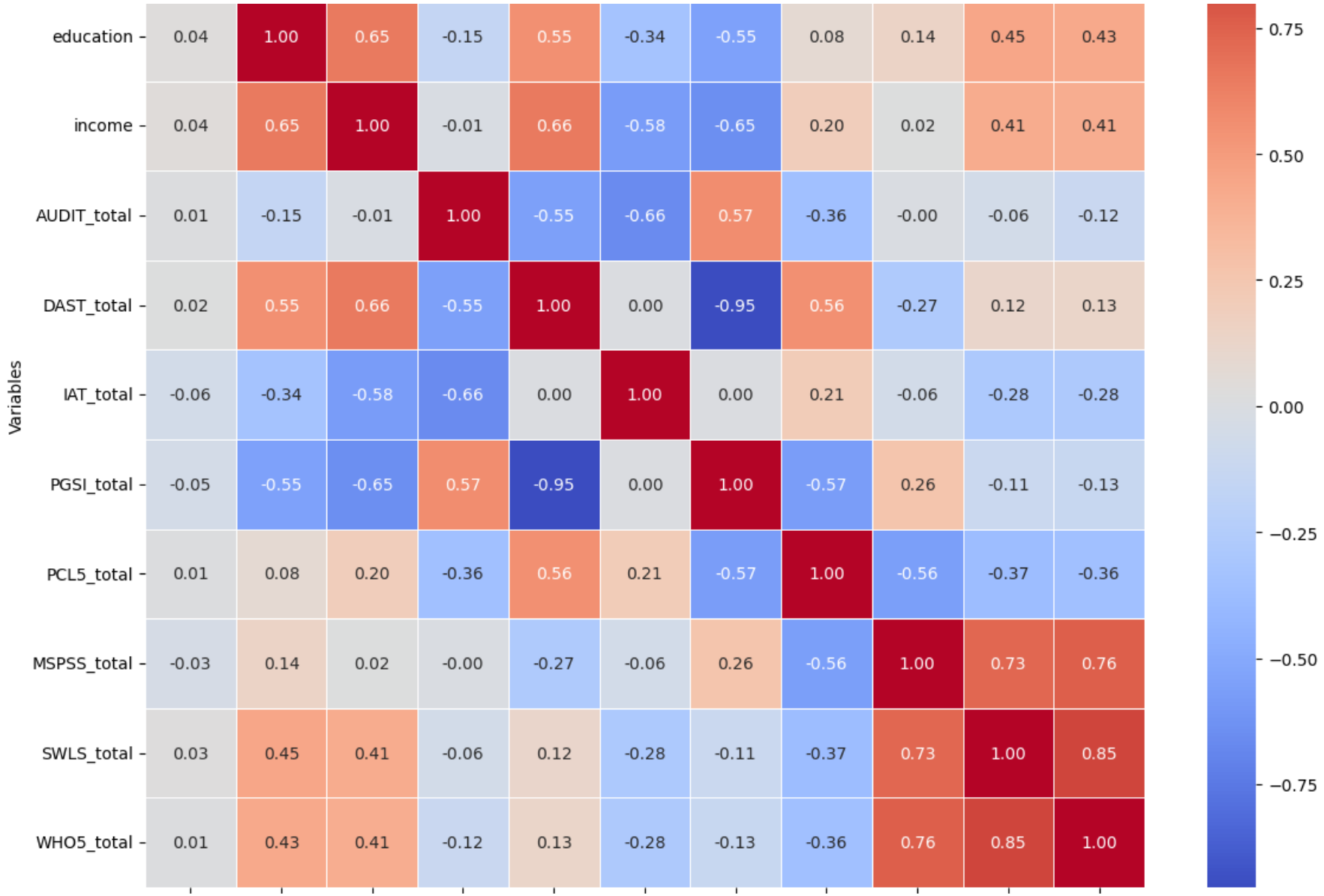
Different addiction profiles: Drug use (DAST) behaves differently from alcohol/gambling, possibly indicating a separate risk group.

Mental-health context: PTSD symptoms (PCL5) correlate positively with substance-use problems and negatively with social support and well-being.

Protective factors: High social support (MSPSS) and life satisfaction (SWLS/WHO5) move together and counterbalance addiction and trauma scores.

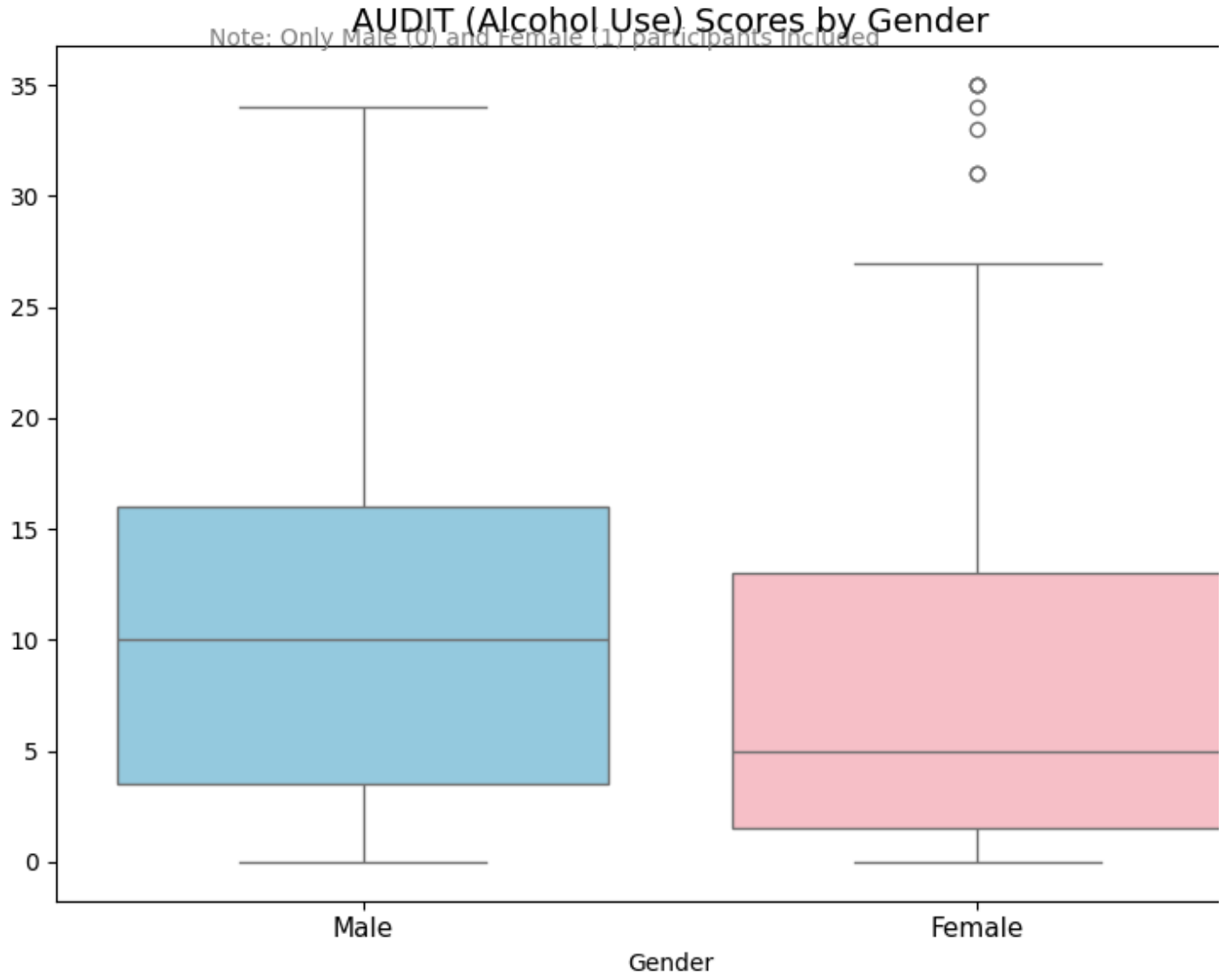


Correlation Heatmap (Excluding Gender and Marital)

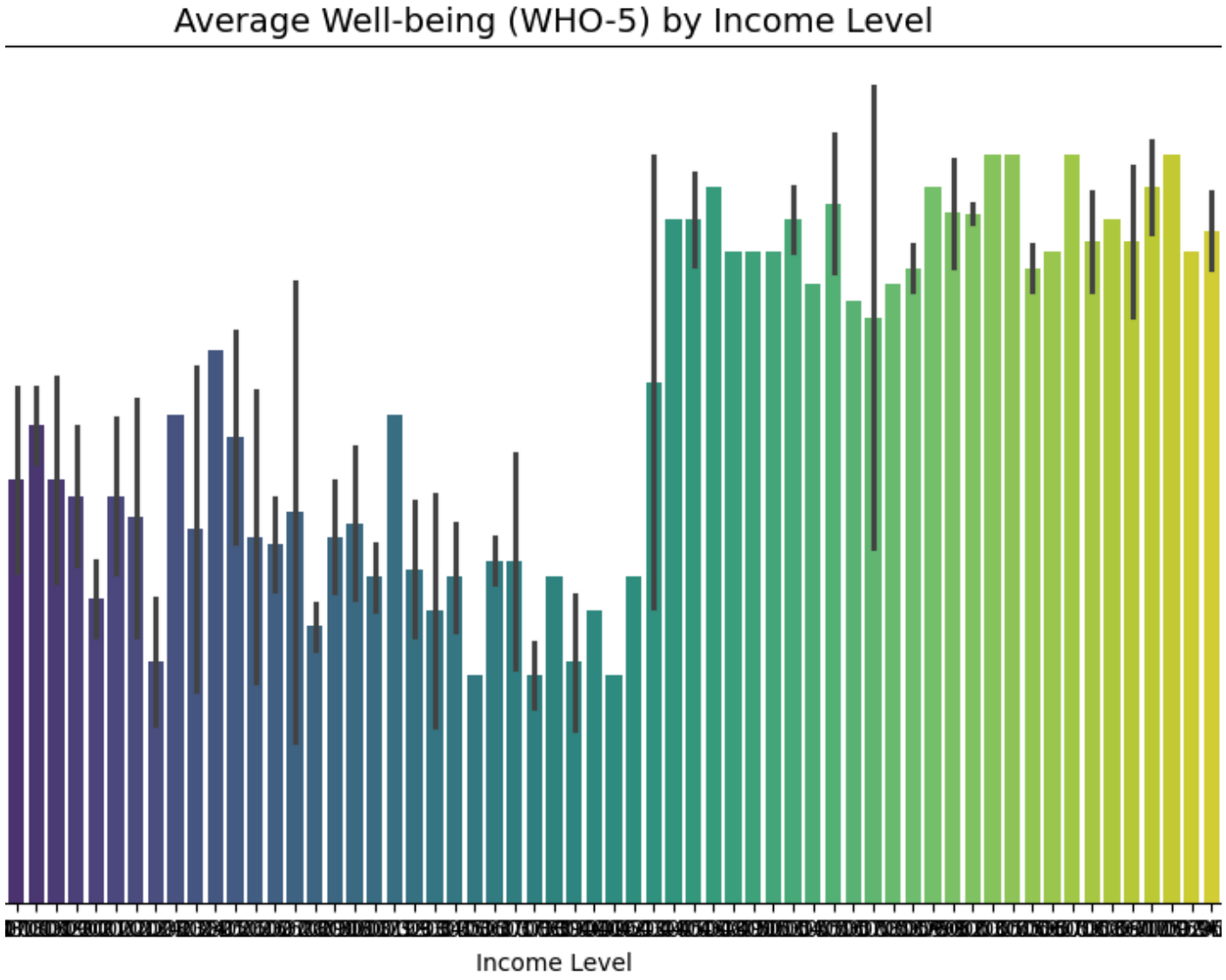


Alcohol use by gender

- The median for males is higher → on average, men report more alcohol use than women.
- The spread of male scores is also wider → greater variation among men (some drink very little, some very heavily).
- The female group has a lower median and smaller variation, but still a few outliers (high-risk drinkers).



- You can clearly see a positive trend: as income level increases, the average WHO-5 score tends to rise as well.
- This means participants with higher income generally report better emotional well-being and life satisfaction.
- Lower-income participants show lower and more variable well-being scores, suggesting more instability in reported happiness.
- The relationship isn’t perfectly linear (some fluctuations), but the overall trend is upward.

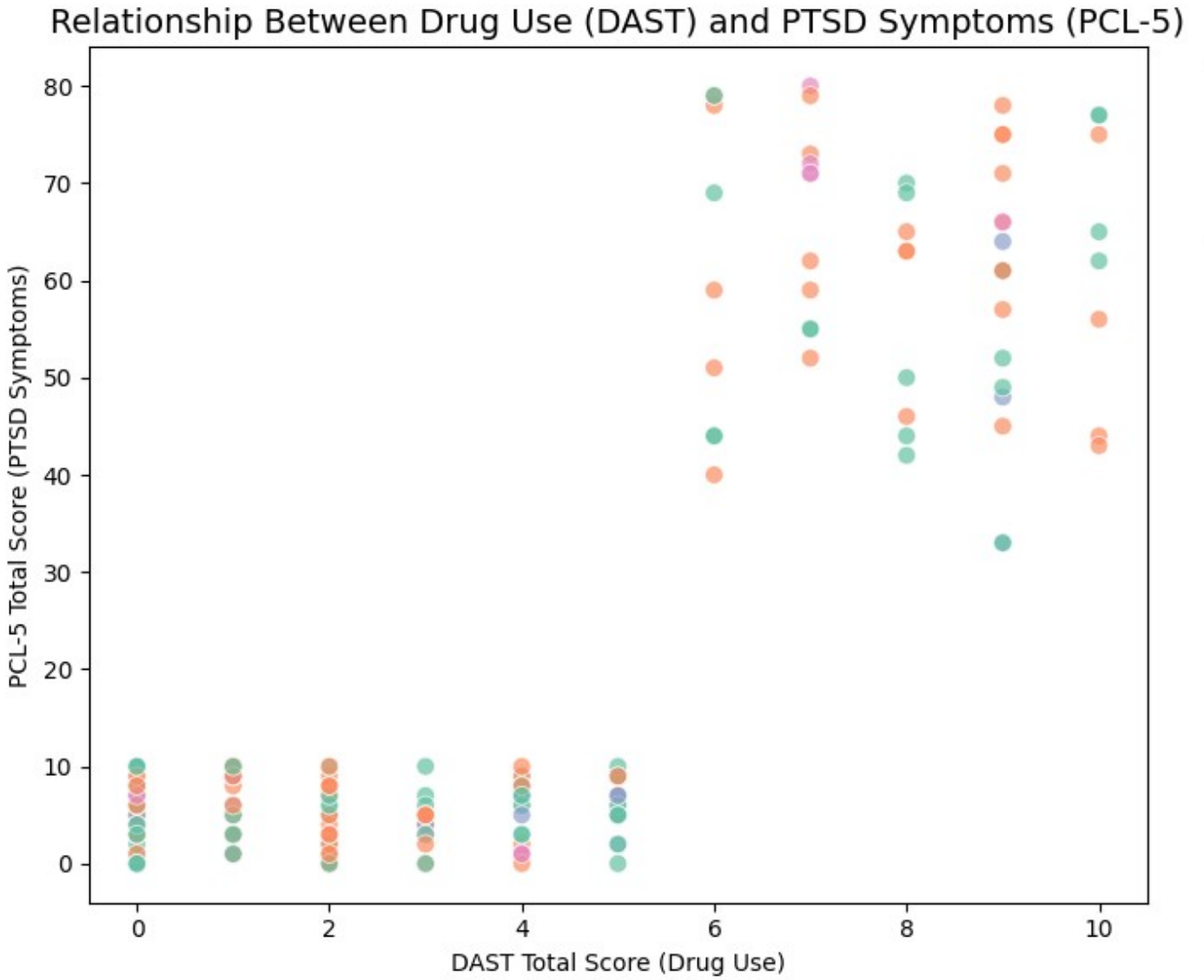


Each point = one participant.

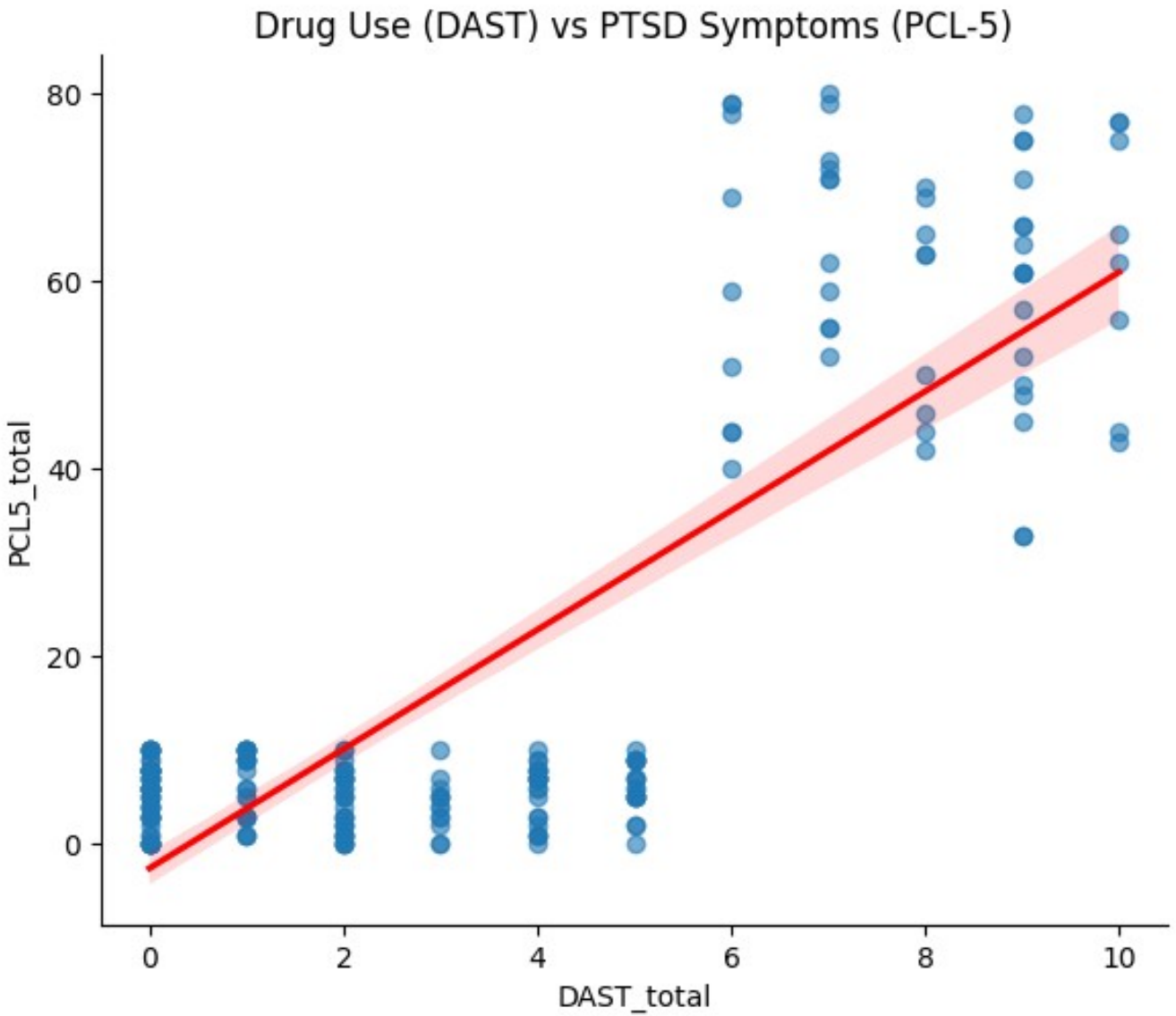
X-axis: level of drug use.

Y-axis: level of PTSD symptoms.

Color: participant's gender (optional for visual separation).

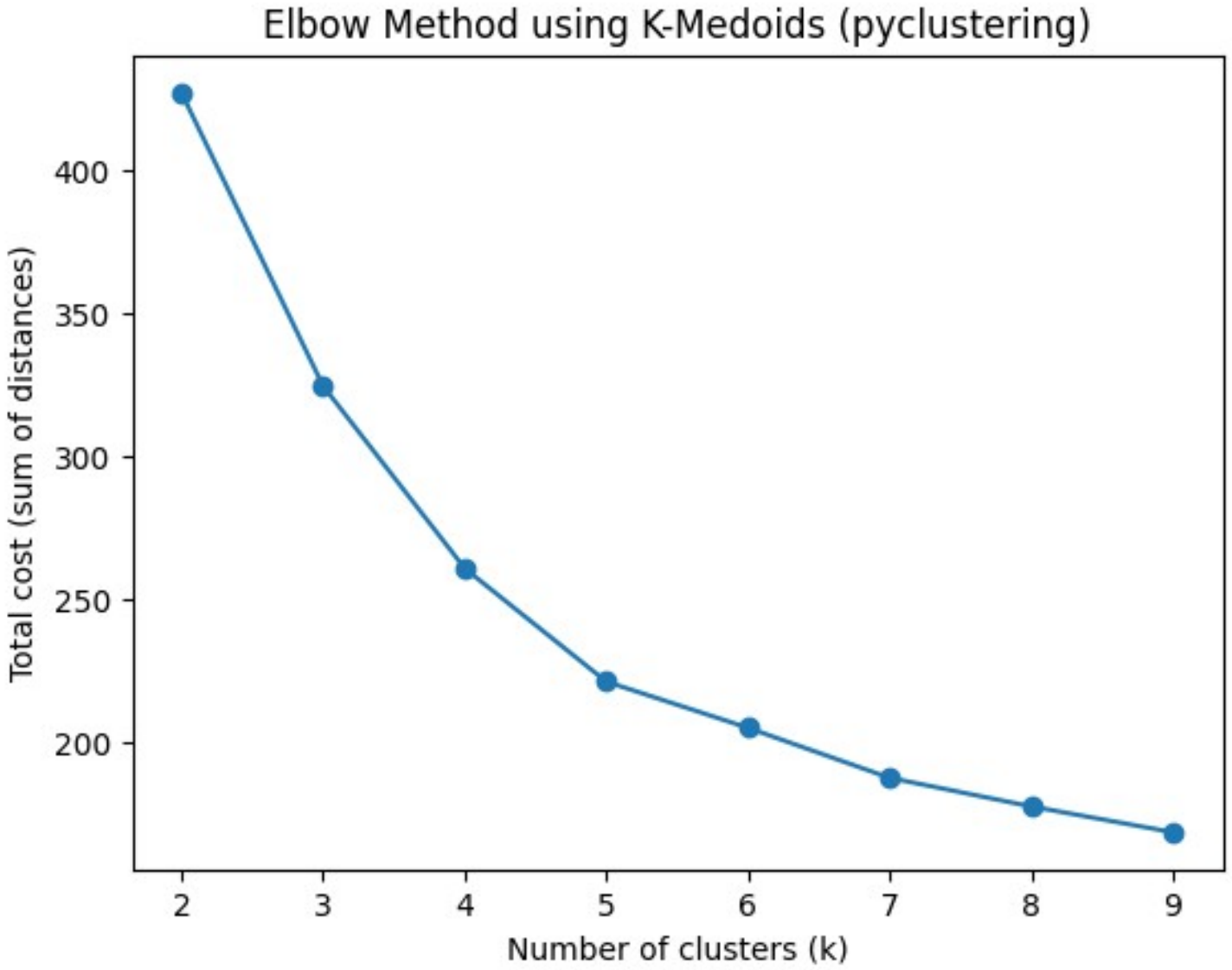


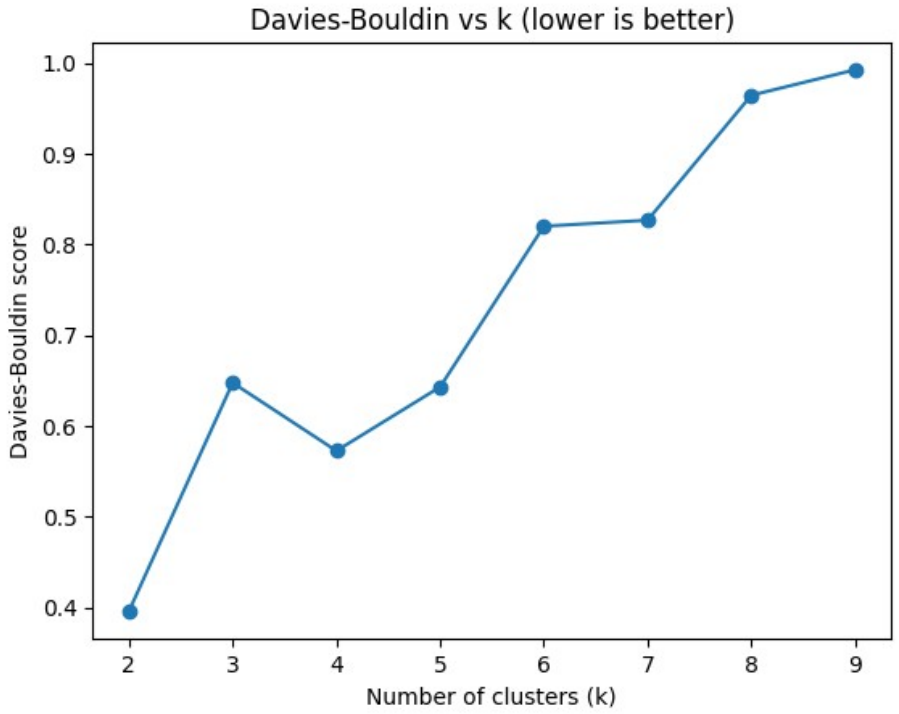
- Figure demonstrates a strong positive relationship between drug use and PTSD symptoms. Participants with higher DAST scores consistently show elevated PCL-5 totals, indicating that greater substance use is associated with more severe trauma-related distress. This pattern aligns with evidence of comorbidity between substance use disorders and post-traumatic stress symptoms.

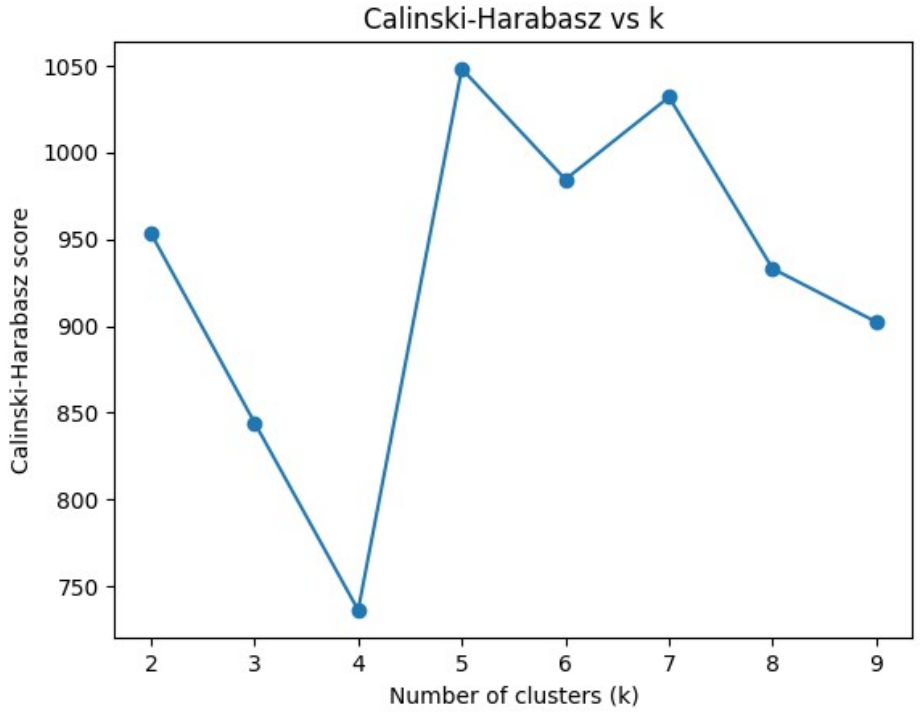


Choosing the correct number of clusters

- The elbow plot for the K-medoids solutions shows a steep decrease in total cost from k = 2 to k = 5, after which the curve flattens and additional clusters only marginally reduce the cost. This pattern suggests that a solution with around 4–5 clusters captures most of the structure in the data, while larger k values provide little improvement in model fit.

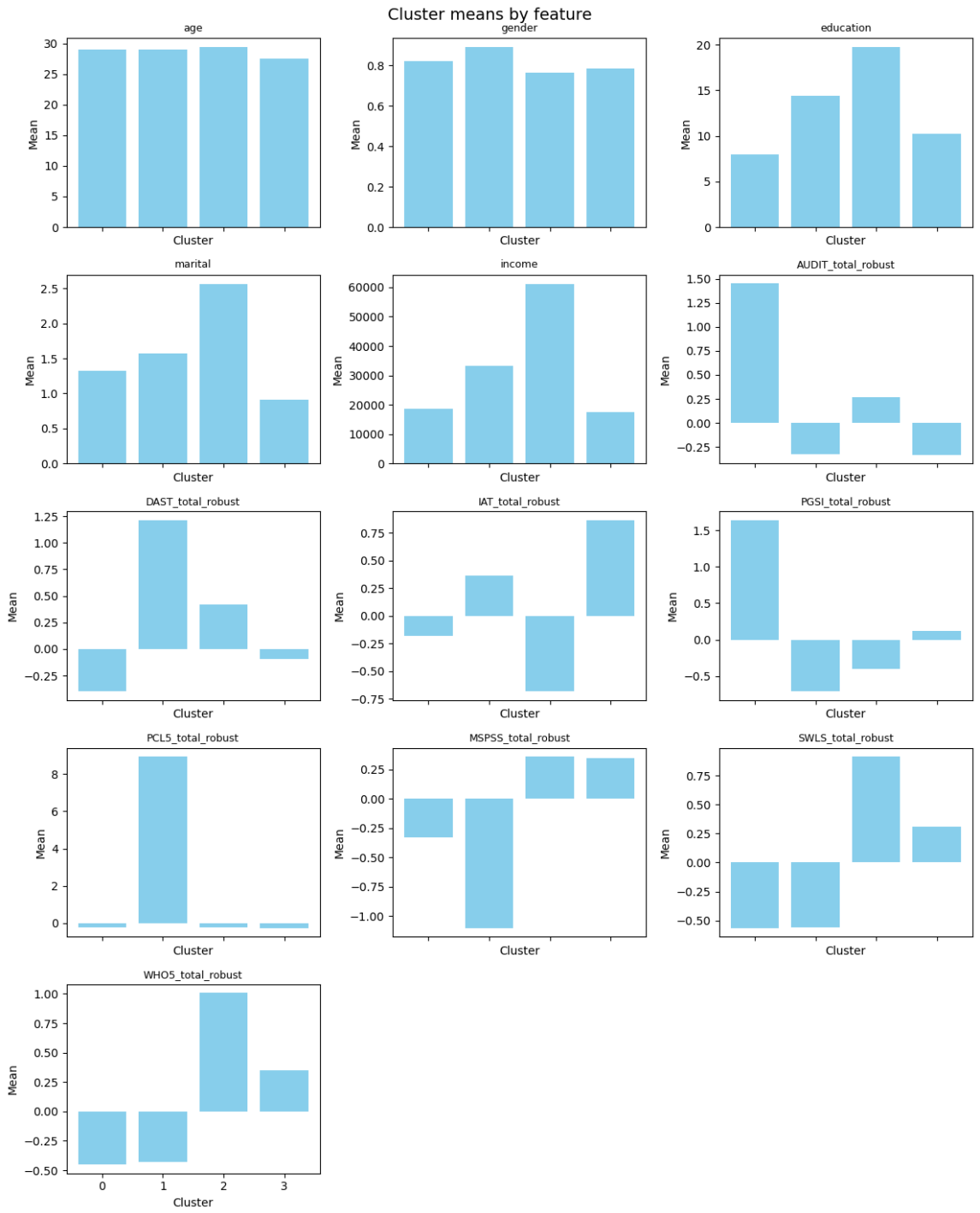




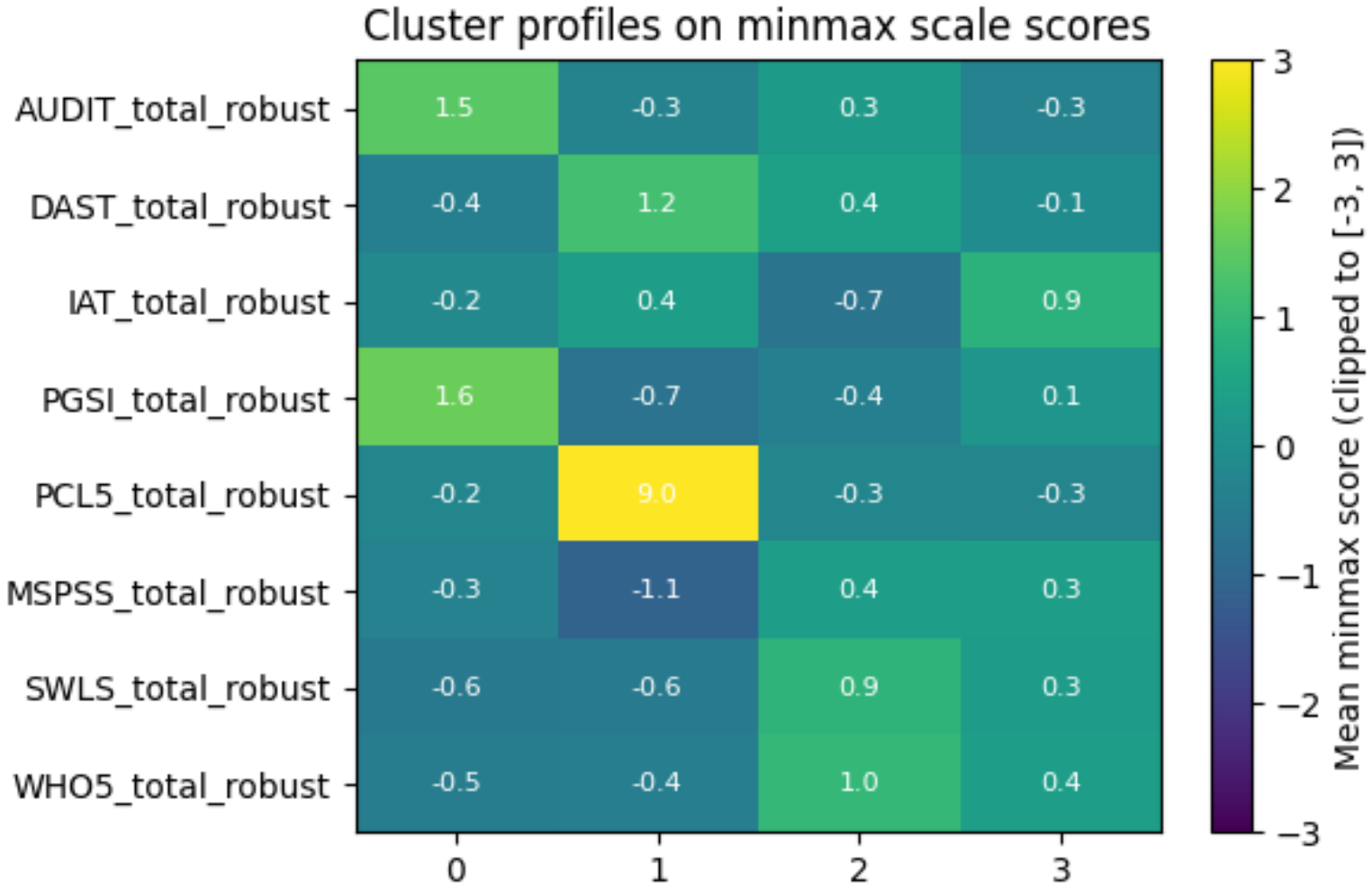


- The internal validity indices showed mixed but informative patterns across different values of k. The Calinski–Harabasz (CH) index was relatively high for k = 2 and increased again for solutions with 5–8 clusters, with local peaks at k = 5 and k = 8, suggesting that more granular partitions can improve between-cluster separation. In contrast, the Davies–Bouldin (DB) index was lowest for k = 2 and increased for larger k, with comparatively favourable values around k = 4–5. Taken together with the elbow plot and the interpretability of the resulting groups, we focused on a 4–5 cluster solution and ultimately retained the 4-cluster model for further analyses.

- These plots show the average value of every feature for each cluster so we can quickly see how the groups differ. It’s basically a visual check to confirm which variables separate the clusters the most (e.g., AUDIT, PGSI, PCL-5, etc.).



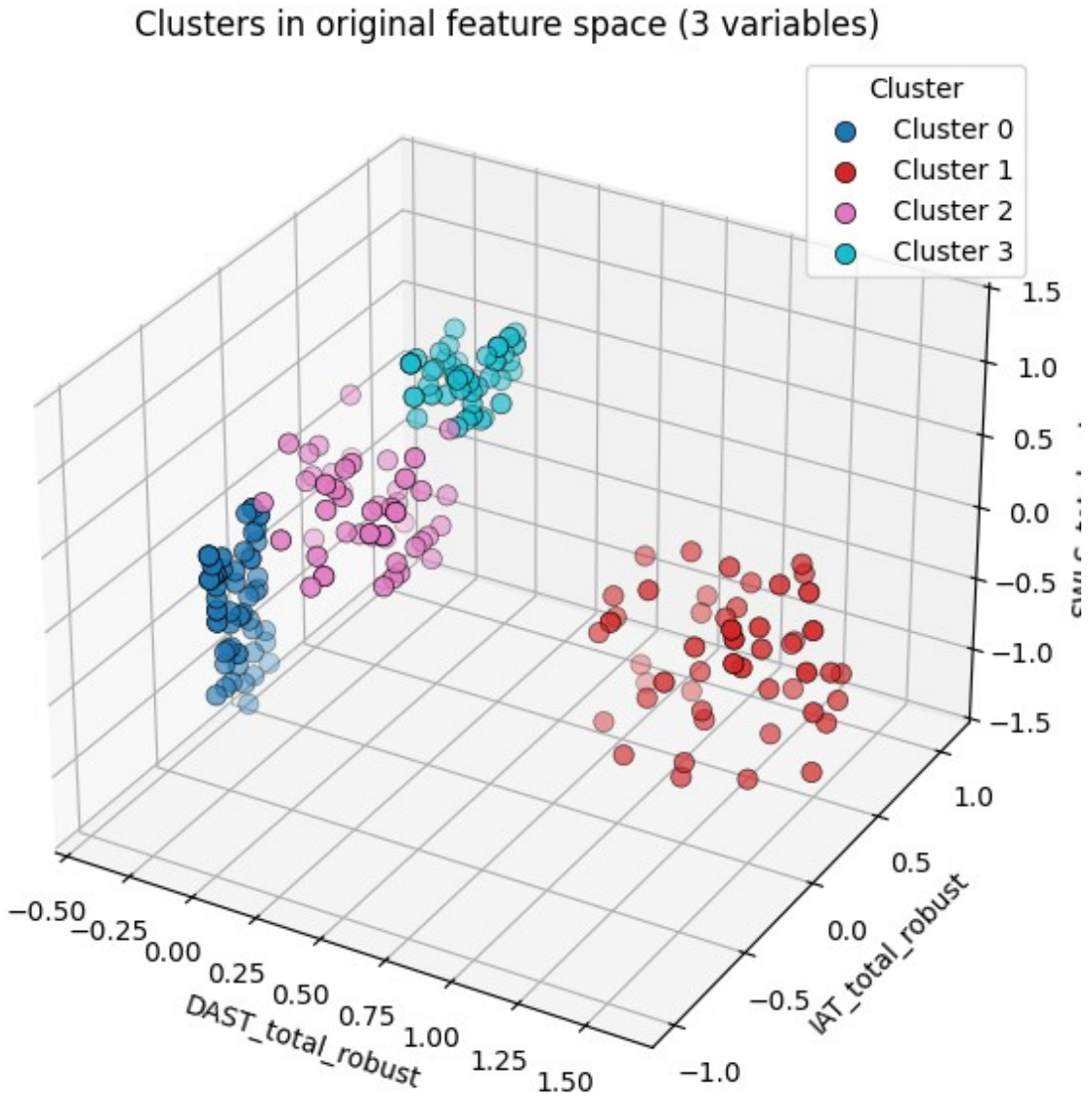
- The heatmap displays mean robust scale scores for each of the four clusters. Positive values indicate scores above the sample median, while negative values indicate scores below it. Cluster 2 shows clearly elevated AUDIT and PGSI scores, indicating higher alcohol use and gambling problems. Cluster 3 stands out with an extremely high mean PCL-5 score and the lowest perceived social support, suggesting a group with pronounced post-traumatic stress symptoms and low support. Cluster 1 has the highest life satisfaction and well-being (SWLS, WHO-5) and relatively low risk scores, whereas Cluster 0 is characterized by elevated internet-use scores (IAT) with relatively average levels on the other



3D representation of clusters

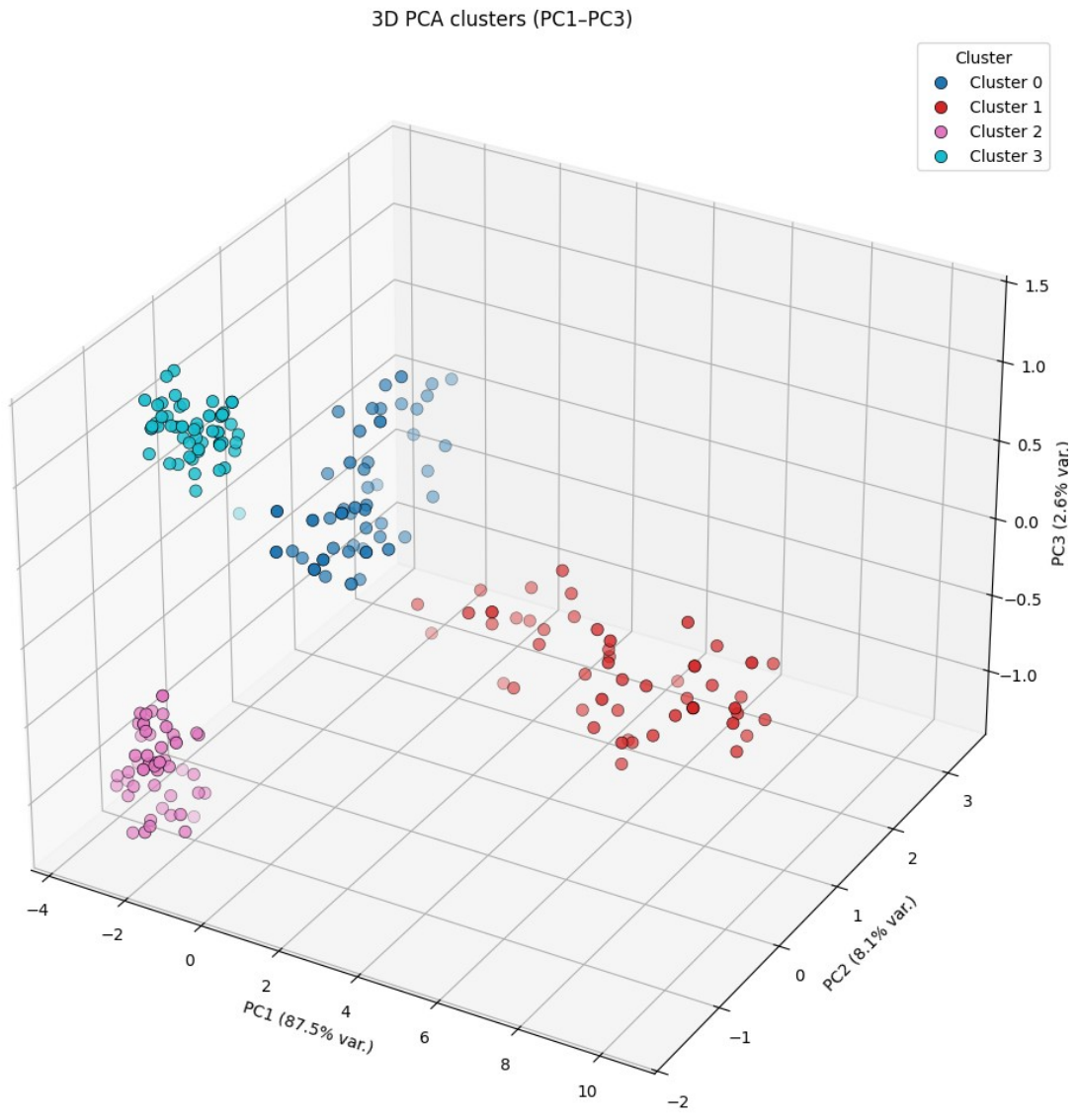
- Here were taken into consideration the top 3 features measured by their RF importance score

DAST_total_robust 0.164924
IAT_total_robust 0.156281
SWLS_total_robust 0.139267



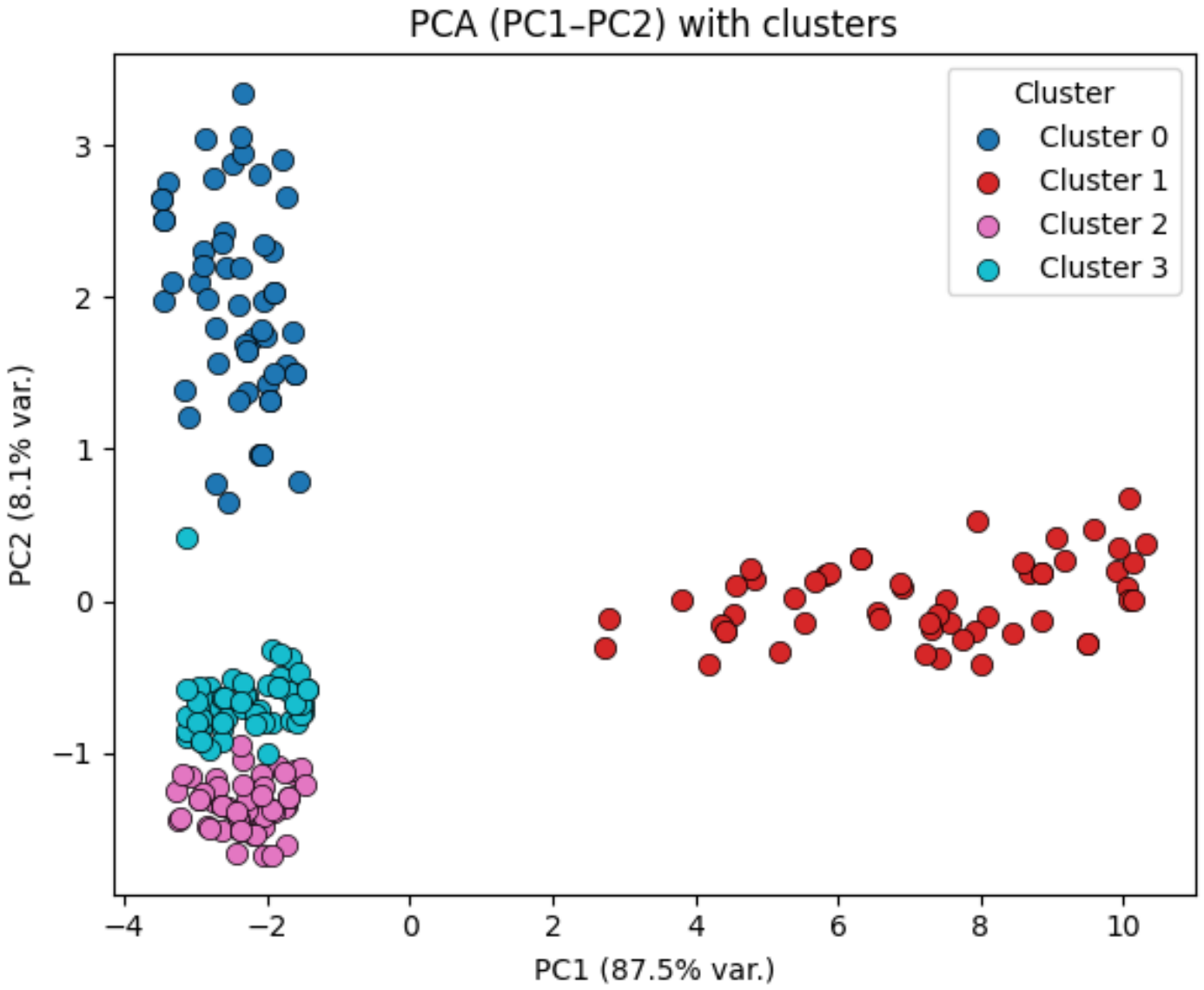
3D PCA representation

- **PC1**: Distress/Addiction vs. Well-being
- **PC2**: Behavioral Risk vs. Life Satisfaction
- **PC3**: Digital Dependency



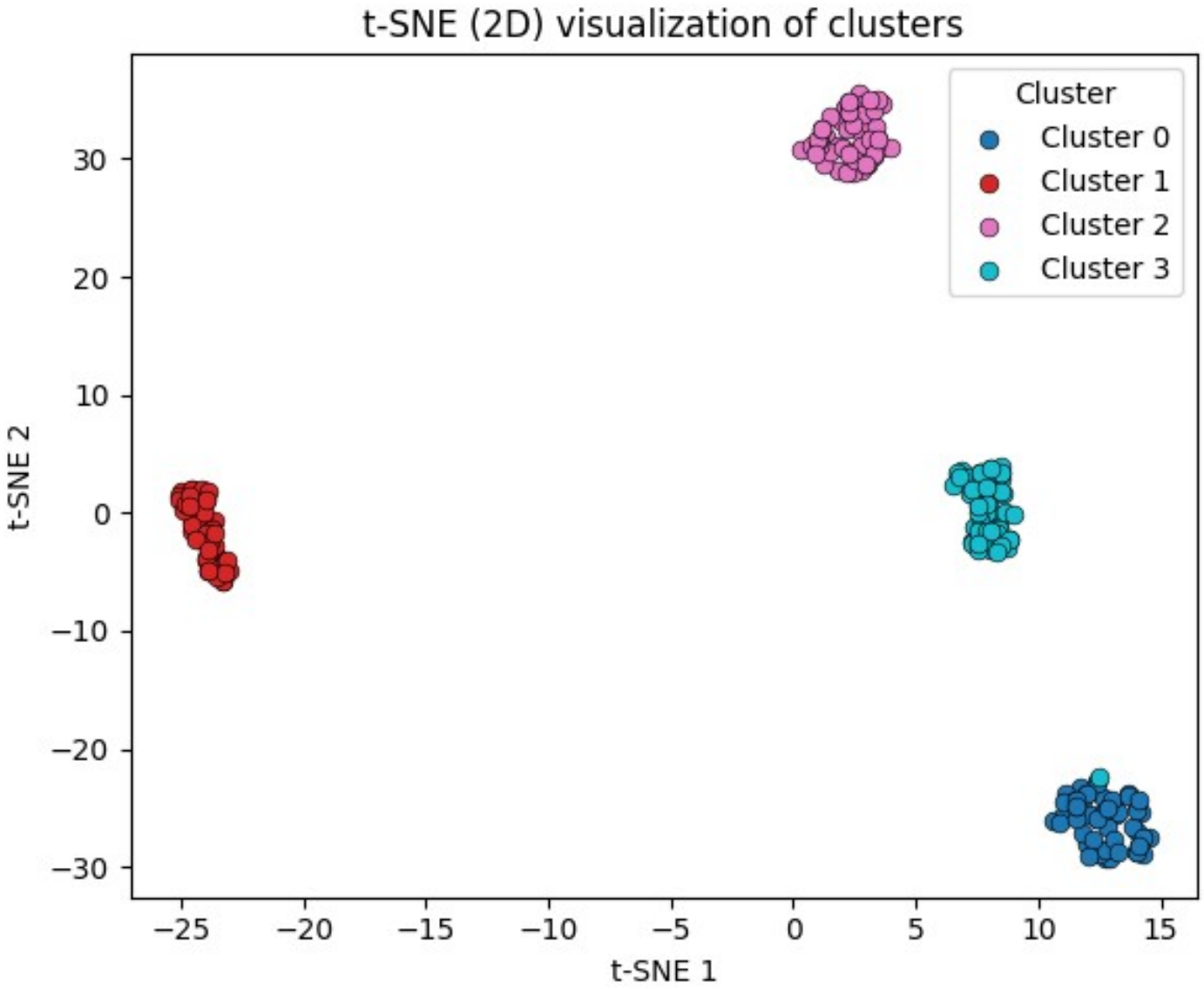
2D PCA representation

- Kaiser (λ>1): suggested number of components = 2
- Number of components needed for ≥90% variance = 2
- Cumulative explained variance: [0.875 0.955 0.981 0.988 0.993 0.996 0.998 1.] Using n_components = 2 for plotting/interpretation.



T-SNE representation

- We additionally explored non-linear embeddings using t-SNE. The two-dimensional t-SNE map showed a very clear separation between the four k-medoids clusters, whereas a three-dimensional t-SNE solution did not provide additional interpretability and therefore is not reported in detail.



Persona Cards

Cluster 0

Working Adult Facing Alcohol and Gambling Stress

""I just want to relax after work, but lately it feels like alcohol and gambling are running my life."

Responsible | Avoidant

Stressed | Pragmatic

Name: Mia/Michael

Age and gender: 28, female/male

Education and Job: Finished elementary school; works in a physically demanding, low-paid job

Family: Married, often with children; wants to "provide" but feels financial pressure

Location: Lives in a modest flat in a smaller town or suburb

Character: Hard-working and loyal, but easily overwhelmed by stress; tends to push feelings aside and "keep going"

Bio:
Mia/Michael works hard, often in shifts, and comes home exhausted. Alcohol and occasional gambling have become a routine way to unwind and escape worries about money and responsibilities. At the same time, arguments with a partner and growing financial stress are starting to feel scary. This persona would never call themselves "an addict" and instinctively resists moralising messages, but a part of them is worried about losing control and hurting the people they care about.

Goals:

- **Key drivers and priorities:**
  - Keep the family financially afloat and avoid conflicts at home
  - Have some form of "deserved fun" after work
  - Quietly check whether alcohol and gambling might already be a problem

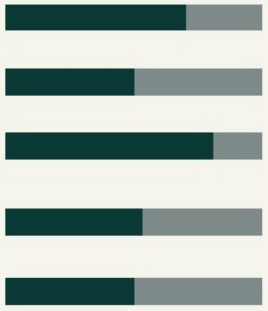
Frustrations:

- Frequent money worries after nights of drinking or gambling
- Arguments with partner about spending and time away from home
- Feels ashamed and doesn't like the word "addiction" – fears being judged

How the game should adress him/her

- Use a down-to-earth, non-judgmental tone that talks about money, family and responsibility
- Emphasise that asking for help is responsible, not weak

Motivation



Cluster 1

Overwhelmed Young Adult Coping with Trauma and Drugs

"I'm exhausted and numb most of the time... I just need something to switch my brain off."

Sensitive | Anxious | Self-critical | Withdrawn

Name: Marko/Maria

Age and gender: 29, male/female

Education and Job: Finished high school; unstable or demanding job with irregular hours

Family: Often in a relationship or married, but relationships can be tense and emotionally draining

Location: Rents a small flat or shares accommodation in a city

Character: Deep, emotional, easily overwhelmed; struggles to trust others and often feels alone even when surrounded by people

Bio:
Marko/Maria has a history of stressful or traumatic experiences and often feels emotionally overloaded and numb at the same time. Drugs are used as a way to sleep, calm down or escape intrusive thoughts, rather than to "have fun". Alcohol and gambling are less central, but the combination of high distress, low support and substance use creates a high-risk situation. This persona is suspicious of authority and easily feels blamed or judged, yet quietly wishes that someone could offer real, non-judgmental help.

Goals:

- Key drivers and priorities:
  - Stop feeling constantly on edge or haunted by past experiences
  - Find ways to cope that don't make things even worse
  - Understand whether their current drug use is actually dangerous

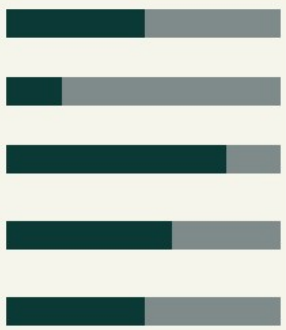
How the game should adress him/her

- Use a warm, validating tone, explicitly avoiding blame or moralising
- Offer information about support networks and help-seeking without pressure

Frustrations:

- Feels that nobody truly understands how bad the anxiety and memories are
- Uses substances to calm down, then feels guilty or ashamed
- Has tried self-help content but quickly loses hope or motivation

Motivation



Cluster 2

High-Functioning Professional Curious About Risk

"I'm fine, but I'm curious how things can quietly get out of control."

Bio:
Luka/Lea has a stable, satisfying life: good job, supportive partner and friends, and overall high well-being. Alcohol is mostly part of social life, and there may be occasional drug use at parties, but without clear negative consequences so far. This persona is not seeking treatment, but is curious and prevention-oriented. They are interested in understanding psychological mechanisms of addiction and in learning how to support others who might be struggling.

Goals:

- Key drivers and priorities:
  - Maintain a good quality of life and mental health
  - Understand how "normal" social drinking and experimentation can become risky
  - Learn how to recognise warning signs in themselves and others

Curious | Ambitious | Reflective | Social

Name: Luka/Lea

Age and gender: Early 30s, male/female

Education and Job: Master's degree; works in a skilled or professional job with solid career prospects

Family: In a stable relationship/married; maintains close contact with family and friends

Location: Lives in a comfortable flat in an urban area

Character: Responsible, empathetic and sociable; likes to understand complex topics

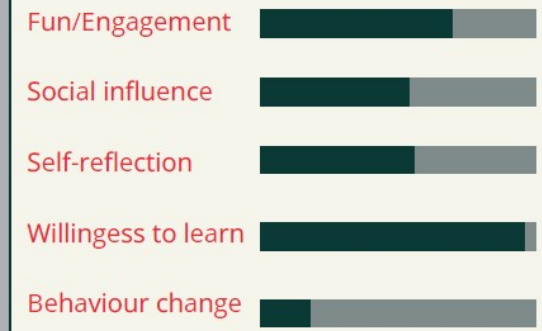
How the game should adress him/her

- Short, high-quality puzzles
- Clear psychoeducation linking each puzzle to real-life mechanisms (denial, peer pressure, stress, support)
- Emphasise how they can support friends

Frustrations:

- Worries about friends who drink or party more heavily
- Feels unsure how to talk about alcohol or drugs without sounding judgmental
- Has limited time and dislikes long, unstructured content

Motivation



Cluster 3

Low-Income, Digitally Engaged Young Adult

"I don't drink much, I just spend way too much time online."

Bio:
Sara/Alex spends a lot of time online: social media, videos, chats, sometimes games. Substance use (alcohol, drugs, gambling) is relatively low, but the internet has become the main way to cope with boredom, loneliness and stress. Well-being is not terrible, but there is a sense of drifting and procrastination. This persona does not strongly identify with addiction problems, yet some patterns (late-night scrolling, lost time, irregular sleep) already affect their functioning.

Goals:
- Key drivers and priorities:
  - Feel more in control of their time and daily routine
  - Avoid future problems with alcohol or other substances
  - Use games and digital tools in a way that is fun but not overwhelming

Friendly | Laid-back | Easily distracted | Sensitive

Name: Sara/Alex

Age and gender: Mid to late 20s, male/female

Education and Job: Finished middle school; works in low-paid or unstable jobs

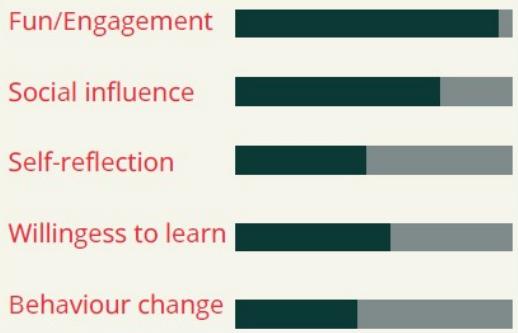
Family: Often single; maintains contact with family and a few close friends

Location: Lives with parents, roommates, or in a small rented flat

Character: Easy-going and kind, avoids conflict; struggles with motivation,tends to escape into online worlds

How the game should adress him/her

- Use simple, intuitive interaction and visually clear puzzles
- Connect the story to online life, peer pressure and coping with boredom or stress
- Connect the story to online life, peer pressure and coping with boredom or stress

Frustrations:
- Feels stuck in low-paid jobs and uncertain about the future
- Notices that screen time often replaces sleep or real-life activities
- Sometimes feels misunderstood or judged as "lazy"

Motivation