A no-framework pseudo-psychometric Web application

Submitted in fulfilment of a university assignment.

James Arrowsmith

This document provides a user guide and explains the methodology for converting questionnaire responses into a job suitability score and movie recommendations.

Chapter 1

User Guide

1.1 Source overview

Table 1.1 shows each source directory and file with a brief description.

Table 1.1: Source contents for this assignment.

Directory/file	**Description**
server.py	Implements a micro-framework to serve the website.
authentication.py	Implements flexible authentication (see Section 1.2.2).
analysis.py	Performs analysis and constructs a profile (see Section 2).
weights.json	Stores the weights for analysis (see Section 2).
server_utils.py	Provides helper functions specific to serving the website.
fetch_utils.py	Provides helper functions for fetching data so that the server may act as a client.
response_utils.py	Provides helper functions for sending responses.
auth.json	Stores the hashed login and attempts for IP blacklisting (see 1.2.2).
default_input.json	Stores the default form inputs to fill in missing inputs (only used when sending data via the command-line).
index.html	HTML for the index entry point (see Section 1.2.3).
psycho.html	HTML for the form.
404.html	HTML for the 404 entry point (see Section 1.2.3).
403.html	HTML sent IP is blacklisted, if enabled (see Section 1.2.2).
main.js	The front-end JavaScript for actioning URIs and displaying content.
style.css	Cascading style sheet for the website.
favion.ico, logo.png	Logo images for the website.
Dockerfile	Contains instructions for building the Docker image.
requirements.txt	Contains the modules not in Python's standard library.
data/	Directory generated at runtime containing input.json (user form inputs) and profile.json (the serialised profile).
images/	Directory generated at runtime containing images fetched from RESTful APIs.

1.2 Running and using the application

1.2.1 Starting the server

Pull the image from GCHR with

```
docker pull ghcr.io/jarrowsm/psychapp:latest
```

and run the server:

```
docker run -it -p PORT:8080 ghcr.io/jarrowsm/psychapp:latest
```

Further options may be specified when running the server (Figure 1.1). Once running, the server is accessible at `localhost:PORT`.

```
usage: server.py [-h] [-p PORT] [-a AUTH_ATTEMPTS] [--disable-ban] [--disable-auth] [--reset-auth]

HTTP server.

options:
  -h, --help            show this help message and exit
  -p, --port PORT       port to listen on (default 8080)
  -a, --auth-attempts AUTH_ATTEMPTS
                        maximum authentication attempts (default 3)
  --disable-ban         disable IP blacklisting while logging IPs
  --disable-auth        disable IP logging and blacklisting
  --reset-auth          reset IP logs and blacklist
```

Figure 1.1: Server usage.

1.2.2 Authentication

Basic HTTP authentication is enabled by default (Figure 1.2). The username and password are both `20005743`. The SHA256 hash is compared to the string in `auth.json` (Figure 1.3). By default, client IPs will be blacklisted after 3 failed attempts. Attempts are stored in `auth.json`. To revert this, restart the server passing `--reset-auth` or rerun the container. Figure 1.4 shows the 403 page for a banned IP. If Curl or Wget are used (based on the user agent header), a JSON error will be displayed instead (Figure 1.5). Attempting to access hidden files (e.g., `auth.json`) results in a 404 error.

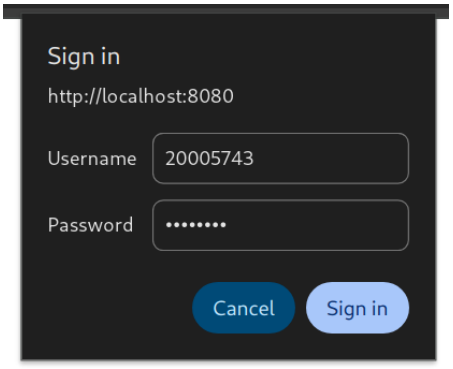


Figure 1.2: 401 Response.

Figure 1.3: Contents of `auth.json` after 3 failed attempts.

Figure 1.4: 403 page as shown in a browser.

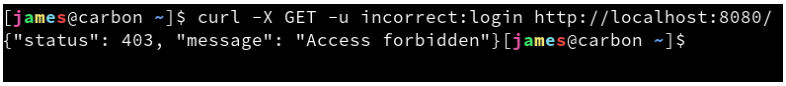


Figure 1.5: 403 JSON as shown in the command line.

1.2.3 Website navigation

Once authenticated, the index page (Figure 1.6), or 404 page (Figure 1.7) for invalid URIs serves as the entry point, following a single-page application scheme.

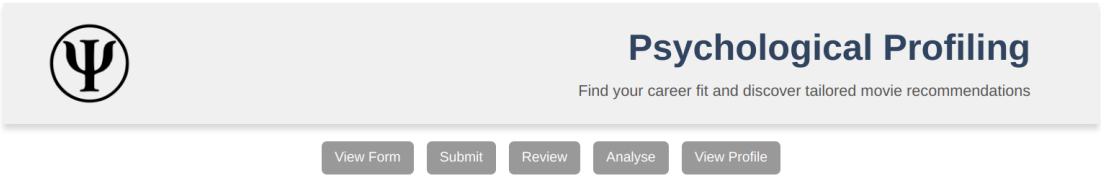


Figure 1.6: Index entry point.

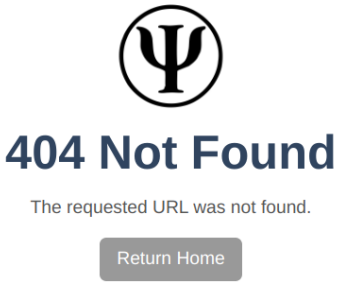


Figure 1.7: 404 entry point.

All buttons call functions in `main.js` that action back-end URIs. `main.js` is also responsible for formatting content for the browser. JSON serialisation is used. Figure 1.8 shows the application after clicking 'View Form' and filling in the fields. Figure 1.9 shows the message after 'Submit' which saves the data on the server at `data/input.json`. Figure 1.10 shows the locked form filled in with the data stored on the server after 'Review'. Figure 1.11 shows the message after 'Analyse', which analyses the data (see Section 2). Figure 1.12 shows the profile after clicking 'View Profile'.

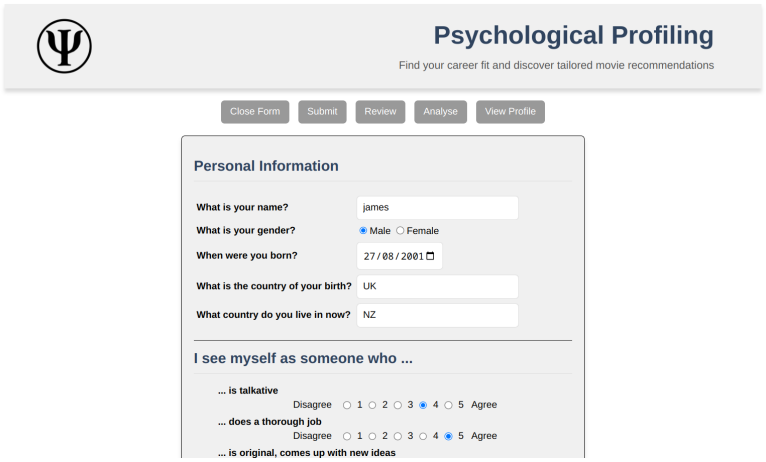


Figure 1.8: Application after clicking View Form and filling in the form.

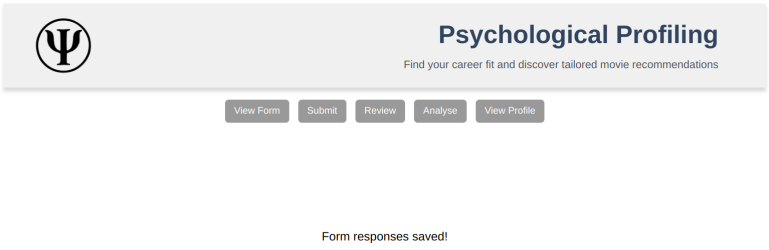


Figure 1.9: Application after clicking Submit.

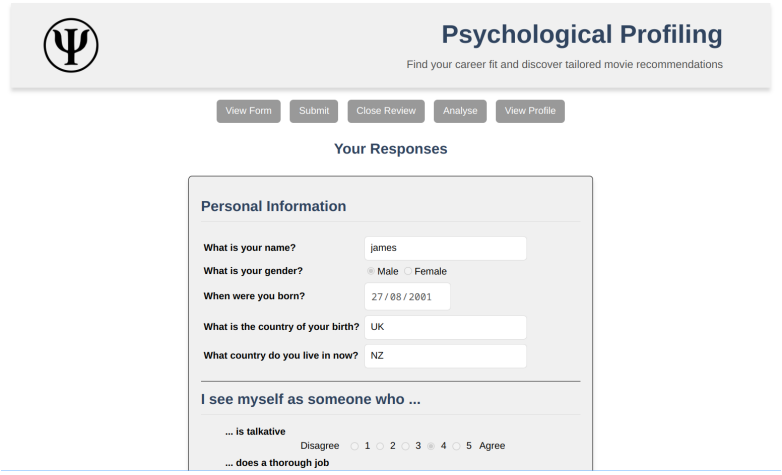


Figure 1.10: Application after clicking Review. Note that the fields cannot be modified.

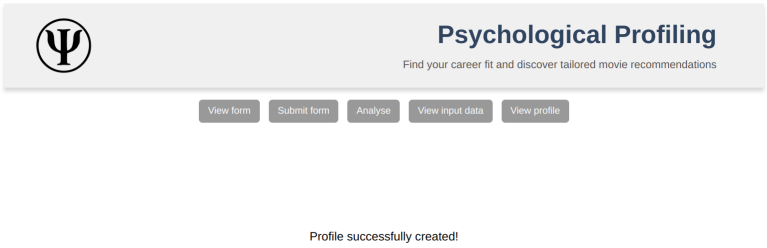
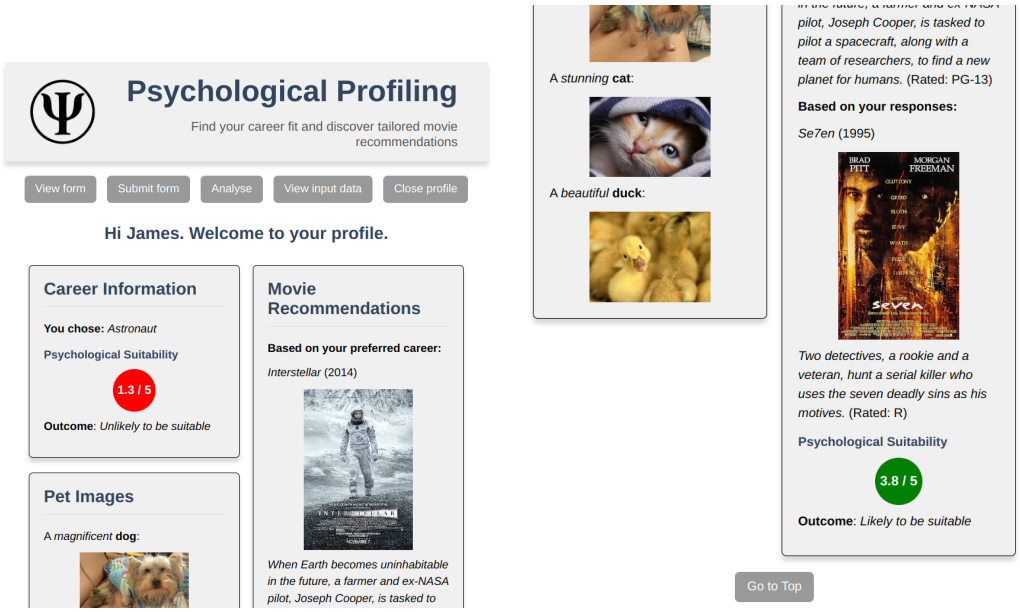


Figure 1.11: Application after clicking Analyse.



(a) Top of the page.

(b) Bottom of the page.

Figure 1.12: Page after clicking View Profile.

Errors may also be displayed when appropriate. On the command-line, these are presented as JSON strings. In a browser, error responses display pages as described above. Other errors are displayed and formatted similar to the messages. Figure 1.13 highlights a 400 error after the user attempted to submit a form that was already submitted. Other 400 errors are shown when the user attempts to re-analyse the data, or use buttons in an incorrect order (e.g., analyse before submitting, view profile before analysing). 500 errors are displayed when an exception occurred on the back-end, such as the analysis error in Figure 1.14. Such errors are not expected with a browser, but may occur on the command-line. Command-line compatibility is a late addition and, while possible to fully interact with the server in this way, this is not recommended due to the difficulty of correctly formatting inputs—particularly since JSON is the only serialisation format supported (no URL-encoding).

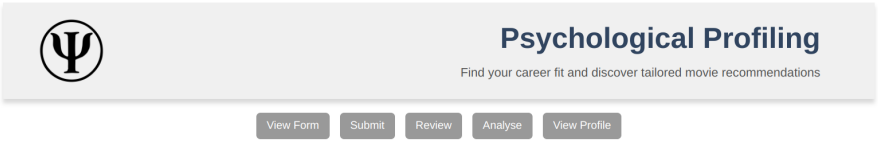


Figure 1.13: A 400 error after trying to submit an already-submitted form.

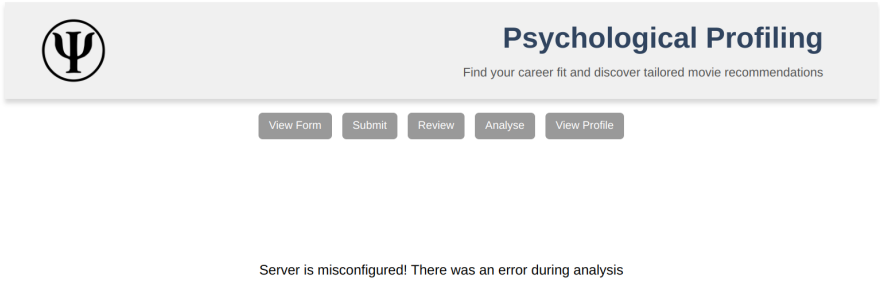


Figure 1.14: A 500 error after an Exception occurred during analysis.

Chapter 2

Analysis

In this chapter, I describe the method for deriving a job suitability score and movie recommendations from the form data.

2.1 The Big Five Personality Traits

In personality psychology, the "Big Five" personality dimensions are general texonomy of personality traits, derived from analyses of people's natural-language descriptions of themselves and others [4]. These characteristics are listed besides their descriptions in Table 2.1. These traits frequently appear in the literature, so may be used to inform the method in the present project. The questions and recommendations are not a rigorous psychometric assessment and are not to be taken seriously. Nevertheless, by categorising each question into these broad traits, the psychological literature can be used to inform their impact on the final scores.

Table 2.1: The "Big Five" personality traits.

Trait	**Description**
Openness (O)	imaginative/philosophical vs uncreative/unintellectual
Conscientiousness (C)	efficient/organised vs haphazard/careless
Extraversion (E)	bold/energetic vs shy/bashful
Agreeableness (A)	sympathetic/cooperative vs cold/competitive
Neuroticism (N)	moody/nervious vs relaxed/calm

2.2 Categorising questions into traits

The questions are categorised into one of the five traits as shown in Table 2.2 based my perception of which is most applicable.[1] A question is 'reversed' when a low score implies a greater level of the associated trait and vice versa.

[1] Certain questions could relate to multiple traits (e.g., the desire to work in a team implying both extraversion and agreeableness).

Table 2.2: Questions categorised according to the "Big Five" traits.

Question *I see myself as someone who...*	**Description**
...is talkative	Extraversion
...does a thorough job	Conscientiousness
...is original, comes up with new ideas	Openness
...is helpful, unselfish with others	Agreeableness
...can be somewhat careless	Conscientiousness (reversed)
...is relaxed, handles stress well	Neuroticism (reversed)
...is curious about many things	Openness
...is full of energy	Extraversion
...starts quarrels with others	Agreeableness (reversed)
...is a reliable worker	Conscientiousness
...is a deep thinker	Openness
...tends to be disorganized	Conscientiousness (reversed)
...worries a lot	Neuroticism
...tends to be quiet	Extraversion (reversed)
...tends to be lazy	Conscientiousness (reversed)
...sometimes shy	Extraversion (reversed)
...is sometimes rude to others	Agreeableness (reversed)
...tends to find fault with others	Agreeableness (reversed)
...gets nervous easily	Neuroticism
...likes to work in a team	Extraversion

2.3 Calculating a suitability score

Given a Likert scale response for question i, $r_i \in \{1, 2, \ldots, 5\}$, I account for reversals via $6 - r_i$ and map to $\{-2, -1, 0, 1, 2\}$ via $\mathbf{r} - 3$. Responses are weighted according to

$$s_i = w_c^t \cdot r_i \tag{2.1}$$

where w_c^t is the weight associated with trait $t \in \{O, C, E, A, N\}$ for career (or film) c. The methods of obtaining these weights are described in Section 2.3.1. Computing this for each of the 20 questions in Table 2.2 produces a set of weighted responses, $\mathbf{s} \in \mathbb{R}^{20}$. A final score, s_F is then computed as the sum of the weighted responses and normalised to $[0, 5]$:

$$s_F = 5 \left(\frac{\sum_{i=1}^{20} s_i - 20 \cdot s_{\min}}{20 \cdot (s_{\max} - s_{\min})} \right). \tag{2.2}$$

For $s_{\min}$ and $s_{\max}$, I determine the worst possible responses for a particular career or film, $\mathbf{r}_{\min}$, by setting each response to the smallest (largest) value for questions with positive (negative) associated weights. I compute $\mathbf{s}_{\min}$ following Equation 2.1, then set $s_{\min} = \sum \mathbf{s}_{\min}$ and $s_{\max} = -s_{\min}$ since responses were mapped to a range symmetrical about 0.

2.3.1 Computing trait weights

The weight for a trait is applied to all corresponding questions. Weights are informed by research on the career or film, where available. Due to normalisation, only the relative magnitudes and directions of weights are relevant, so weights may be sourced from diverse methodologies.

CEO of a large mega-corporation. Harrison et al. [3] propose a linguistic measure of CEOs' Big Five traits leveraging machine learning to make estimates from speech transcripts which they validate against a traditional method. They apply the model to a sample of 3,000 CEOs and find the following interactions between Big Five traits and firm performance: openness (-0.04), conscientiousness (-0.05), extraversion (0.02) and neuroticism (-0.02)—no effect found for agreeableness. Based on these results, I set

$$W_{\text{ceo}} = \{w^O, w^C, w^E, w^A, w^N\} = \{-0.4, -0.5, 0.2, 0, -0.2\}.$$

Astronaut. Rose et al. [10] collect effectiveness and personality data from 65 NASA astronauts and analyse their relationship. Nine effectiveness dimensions are consolidated into one supervisor-rated factor, with correlations reported for openness (-0.277), conscientiousness (-0.205), extraversion (-0.147), agreeableness (0.286) and neuroticism (-0.205). From these results, I set

$$W_{\text{astronaut}} = \{w^O, w^C, w^E, w^A, w^N\} = \{-0.277, -0.205, -0.147, 0.286, -0.205\}.$$

Medical doctor. Lievens et al. [5] analyse Big Five traits and academic performance of medical students. Taking third-year performance as dependent, their multiple regression analysis provides weights for openness (0.15), conscientiousness (0.19), extraversion (-0.04), agreeableness (-0.10), and neuroticism (0.03). However, these may not fully transfer to real-world job performance, where interpersonal skills are emphasised and indecisiveness is undesirable. Babaei et al. [1] report average scores for nurses and physicians (Table 2.3), grouping them on whether they had a medical error on record. Their results highlight extraversion and agreeableness as beneficial and neuroticism as detrimental. Their results are shown in Table 2.3.

Table 2.3: Personality factors between two groups of medical workers, derived from Figure 1 in [1].

Trait	**Without Medical Error**	**With Medical Error**
Openness	36.54	36.89
Conscientiousness	45.54	41.44
Agreeableness	43.26	39.1
Extraversion	42.06	38.2
Neuroticism	29.17	34.59

I normalise the differences to derive weights from this data. These are 4.1, 4.16, 3.86, -5.42, and -0.35, with an absolute sum of 17.89. The weights are then computed, $\frac{-0.35}{17.89} = -0.02$, $\frac{4.1}{17.89} = 0.229$, $\frac{3.86}{17.89} = 0.216$, $\frac{4.16}{17.89} = 0.233$, $\frac{-5.42}{17.89} = -0.303$. Compared to the student weights, conscientiousness is similarly beneficial, while extraversion agreeableness and neuroticism were more impactful, with the latter two inverted. Meanwhile, openness was less significant. I suspect these weights better reflect actual job performance compared to pre-clinical academic performance. Accordingly, I set

$$W_{\text{doctor}} = \{w^O, w^C, w^E, w^A, w^N\} = \{-0.02, 0.229, 0.216, 0.233, -0.303\}.$$

Fashion model. I was unable to find research on the Big Five personality traits of fashion models. Yet, assuming 'looks' are met, other aspects—including personality—might predict better or worse career outcomes. Mair [7] offers a relevant chapter on mental health. The author notes that models face stress and anxiety due to industry demands with risks of body dysmorphia,

eating disorders, and low self-esteem. Lower neuroticism could speculatively be beneficial, so I set this to -0.2. Modelling's transient nature may provoke financial instability and job insecurity, compounding these maladies. Elevated conscientiousness might improve resilience through better personal financial responsibility which is associated with this trait (alongside agreeableness and openness) [9], so I set this to 0.2. Additionally, models are often required to pose, perform at runway events, and have their image published, so higher extraversion might be helpful and I set this to 0.4. These activities, alongside the often need to travel alone [7], suggest that elevated openness may also help, so I set this to 0.2. The minimal autonomy [7] could be especially problematic for disagreeable individuals, so I set this to 0.3. Overall, I derive

$$W_{\text{model}} = \{w^O, w^C, w^E, w^A, w^N\} = \{0.2, 0.2, 0.4, 0.3, -0.2\}.$$

Rockstar. Gillespie and Myors [2] provide a profile (using z-scores) of rock musicians—openness: 1.35, conscientiousness: -0.75, extraversion: 0.24, agreeableness: -0.44, and neuroticism: 0.73. Their study defined rock musicians as "anyone engaged in playing rock or pop music either in a rehearsal, recording or performance situation," so the application of their findings to bonafide *rockstars* is approximate. Notwithstanding, the following values appear sensible:

$$W_{\text{rockstar}} = \{w^O, w^C, w^E, w^A, w^N\} = \{1.35, -0.75, 0.24, -0.44, 0.73\}.$$

Refuse collection operative. I am not aware of any study relating personality to refuse collection performance, but there is a thesis by Tichon [11] that analyses the impact of personality dimensions on the performance of delivery drivers. Both jobs involve routine tasks (driving repetitive routes) physical labour, early morning shifts, and interacting with local communities. I leverage this data in lieu of more relevant studies. Tichon [11] uses the Personal Style Inventory (PSI) [6] which extends the Big Five to include additional work-based constructs. PSI exhibits high levels of correlation with main facet scales of major Big Five inventories (e.g., 'Introversion' with Extraversion; 'Emotional Stability' with Neuroticism). I thus derive the weights from the R^2 Change between the personality variables and overall performance ratings: $w^E = -0.001$ (inverse of Introversion), $w^N = -0.051$ (inverse of Emotional Stability), $w^A = 0.018$. Aligning 'comfort with procedures' with Conscientiousness, $w^C = 0.174$, and inverting 'Preference for Long Tenure', $w^O = -0.022$. Hence,

$$W_{\text{garbage}} = \{w^O, w^C, w^E, w^A, w^N\} = \{-0.022, 0.174, -0.001, 0.018, -0.051\}.$$

Movie recommendations. I provide two movie recommendations—one corresponding to the preferred career, and another based on the psychological data. First, a film is associated with each career as shown in Table 2.5. Regarding the second, Monteiro et al. [8] explore individual personality differences as they relate to movie genre preferences. The authors introduce a five-factor 'MOVIE' (Melodrama, cOmic, Violent, Imaginative, and Exciting) framework. Shown in Table 2.5, I derive $W_{\text{film}} = \{w^O, w^C, w^E, w^A, w^N\}$ for five films from the standardised regression coefficients for the movie preferences which I associate with the example, provided in the study, for the genre with the highest correlation within each preference category. I generate s_F for each film using the methods described in Section 2.3 and recommend the film with the largest value.

Table 2.4: Films selected for preferred careers.

Career	**Film (Year)**
CEO	The Wolf of Wall Street (2013)
Astronaut	Interstellar (2014)
Medical Doctor	Contagion (2011)
Fashion Model	Zoolander (2001)
Rockstar	Zappa (2020)
Refuse Collection Operative	WALL-E (2008)

Table 2.5: Weights for selecting a film based on the psychological data.

Film (Year)	**Genre (Category)**	$\mathbf{W_{film}}$
City of Angels (1998)	Romance (M)	$\{0.0, 0.01, 0.06, 0.05, 0.09\}$
Angry Birds (2016)	Animation (O)	$\{-0.7, 0.0, 0.12, 0.07, -0.05\}$
Apocalypse Now (1979)	War (V)	$\{0.3, 0.0, 0.06, 0.05, -0.06\}$
The Lord of the Rings (2001)	Adventure (I)	$\{0.07, -0.07, 0.03, 0.12, 0.01\}$
Se7en (1995)	Mystery (E)	$\{0.14, -0.04, -0.09, 0.02, 0.12\}$

Bibliography

[1] Mansour Babaei, Mohammad Mohammadian, Masoud Abdollahi, and Ali Hatami. Relationship between big five personality factors, problem solving and medical errors. *Heliyon*, 4(9), 2018.

[2] Wayne Gillespie and Brett Myors. Personality of rock musicians. *Psychology of Music*, 28 (2):154–165, 2000.

[3] Joseph S Harrison, Gary R Thurgood, Steven Boivie, and Michael D Pfarrer. Measuring ceo personality: Developing, validating, and testing a linguistic tool. *Strategic Management Journal*, 40(8):1316–1330, 2019.

[4] Oliver P John. History, measurement, and conceptual elaboration of the big-five trait taxonomy. *Handbook of personality: Theory and research*, pages 35–82, 2021.

[5] Filip Lievens, Pol Coetsier, Filip De Fruyt, and Jan De Maeseneer. Medical students' personality characteristics and academic performance: A five-factor model perspective. *Medical education*, 36(11):1050–1056, 2002.

[6] John W Lounsbury, Lucy W Gibson, and Frances L Hamrick. The development and validation of a personological measure of work drive. *Journal of Business and Psychology*, 18: 427–451, 2004.

[7] Carolyn Mair. Mental health and fashion. In *The Psychology of Fashion*. Routledge, 2 edition, 2024. ISBN 9781032712550. URL `https://doi.org/10.4324/9781315620664`.

[8] Renan Pereira Monteiro, Gabriel Lins de Holanda Coelho, Tatiana Medeiros Costa Monteiro, and Carlos Eduardo Pimentel. The movie model: Assessing movie preferences and personality correlates. *Estudos de Psicologia (Natal)*, 27(2):146–154, 2022.

[9] Gokhan Ozer and Ummuhan Mutlu. The effects of personality traits on financial behaviour. *Journal of Business Economics and Finance*, 8(3):155–164, 2019.

[10] RM Rose, LF Fogg, RL Helmreich, and TJ McFadden. Psychological predictors of astronaut effectiveness. *Aviation, space, and environmental medicine*, 1994.

[11] Mark Andrew Tichon. *Personnel selection in the transportation sector: An investigation of personality traits in relation to the job performance of delivery drivers*. The University of Tennessee, 2005.