



Cannabis Data Science #71

June 22nd, 2022



"I believe my parents live forever in the code they bequeathed to me, and I am never lonely."

- Fictitious Quote

Scientific Process

- Hypothesis
- 2 Literature Review
- Methodology
- Oata
- Results
- Conclusion

1. Hypothesis

Hypothesis - A cannabis consumer's <u>personality</u> affects the <u>chemotype</u> of cannabis that the consumer consumes on <u>average</u>.

2. Literature Review

Economics

 Revealed Preferences - A theory that preferences of consumers can be revealed by their purchasing habits.

Psychology

- The Big 5 Personality Traits
 - Openness
 - Conscientiousness
 - Extraversion
 - Agreeableness
 - Neuroticism

3. Methodology - Ordinary Least Squares

Estimate the relationship between the Big 5 personality traits and the β -pinene to D-limonene ratio. Test

$$r_i = \beta_0 + \beta_o O_i + \beta_c C_i + \beta_e E_i + \beta_a A_i + \beta_n N_i,$$

where

- r_i is the average β -pinene to D-limonene for user i;
- O_i, C_i, E_i, A_i, N_i are the predicted personality traits for user i on a scale of 0 to 1.

4. Data - Natural Language Processing to Predict Personality Traits

Collected data - Consumer purchases merged with lab results.

Synthetic data - Attempt to detect the Big 5 personality traits with natural language processing (NLP).

- The idea is that the user reveals their personality with the language that they choose to use.
- A scale of 0 to 1 will be used for each trait.

The **lexical hypothesis** asserts

- Personality traits that are important to a group of people will eventually become a part of that group's language;
- Important personality traits are likely to be encoded into a single word.

5. Results, conclusions, and future research

Advanced Cannabis Analytics

Thursday, June 30th, 2022 4:20pm EST | 1:20 PM PST



Thank you for coming.

Insights of the Day

• The words that you write are powerful.

What is on your mind for next week?