EXAMINING PSYCHEDELICS USE AND MENTAL HEALTH CONDITIONS

knowledge, use, and outcomes

Data Analysis by
Lisa Girard
Thinkful Data Science Flex Student
December 19, 2022

Capstone II: Supervised Learning

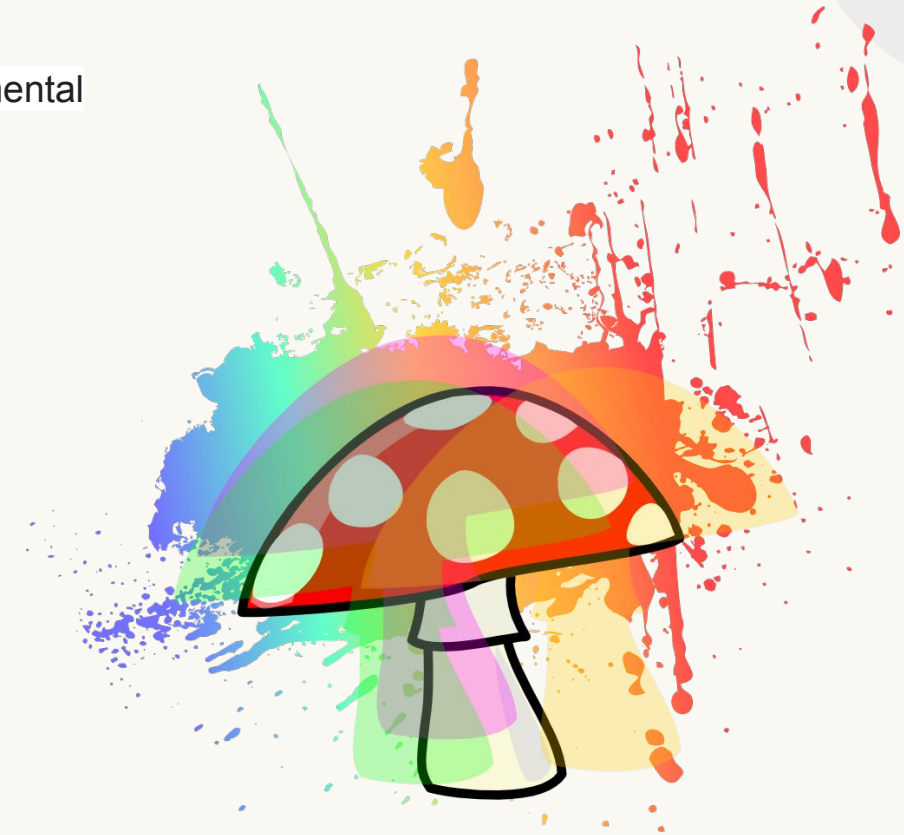
PSYCHEDELIC RENAISSANCE

- **Mental health care modalities** uprooted by psychedelic “plant-based medicines”, produce altered states of consciousness, some lab manufactured
- US research halted early 1970s, Schedule I substance “high potential for abuse and no medical value”
- 2000 - Johns Hopkins granted research regulatory approval
- Last 20 years - new studies, data analysis, and campaigns to decriminalize, regulate, and facilitate Psychedelic-Assisted Therapy (PAT)
- New methods, modalities, and mindsets for crucial mental health tools and therapeutics

PSYCHEDELIC MEDICINES

Known psychedelic substances being considered for mental health therapeutic modalities:

- Psychedelic mushrooms (psilocybin)
- Ayahuasca
- DMT
- 5-MEODMT
- Ketamine
- LSD
- MDMA
- Peyote/mescaline



MENTAL HEALTH CONDITIONS

Psychedelics currently being tested in treating:

- Trauma
- Addiction
- Anxiety
- Depression
- Chronic Pain
- Death diagnoses (cancer)

AN INTRODUCTION TO THE DATA

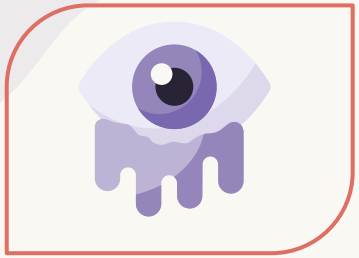
DATA ANALYSIS

DATA SOURCE:
Psychedelic Mushrooms in the USA: Knowledge, Patterns of Use, and Association With Health Outcomes

Downloadable raw data set

STUDY OBJECTIVE:
"The objective of this study was to assess knowledge about psychedelic mushrooms (PM) among American adults and explore associations between PM use and various general and self-reported mental health outcome measures."

FINDING DATA & CORRELATION

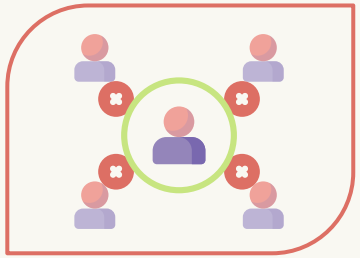


FINDING DATA

Finding raw data regarding psychedelics use is difficult due to criminalization/legal issues, DEA drug classification, stringent research trial exclusion criteria, stigma, and financial barriers.

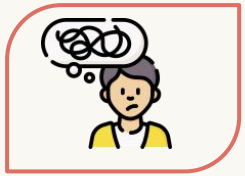
ASSOCIATION vs CORRELATION & CAUSATION

This is an exercise in association-relationship / correlation ONLY.
This data analysis is not an examination or exercise in "causation and correlation". The intention of this data analysis is to determine prediction based on association and relationship and **not** to determine or predict mental health states *causing* psychedelic use.



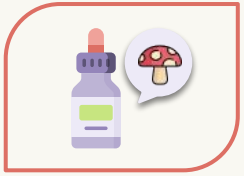
DATA EXPLORATIONS

What is the data telling us?



MHCs

Examine pre-existing **Mental Health Conditions (MHC)** and supposed recreational psychedelic use



SELF-MEDICATE

Are people with certain MHCs knowingly or unknowingly using psychedelics as a form of self-medication for healing and/or treatment?

FACTORS

- DEPRESSION
- ANXIETY
- CHRONIC PAIN
- MIGRAINES
- INSOMNIA
- RECENT PRIOR KNOWLEDGE
- GAD >= 10
- PHQ >= 10
- MCS12
- CCI

OBJECTIVE:

To predict psychedelics users based on certain mental health conditions and/or recent prior knowledge of psychedelic therapeutic use

DATA ANALYSIS:
WHAT DID I FIND?

EXPLORATORY DATA ANALYSIS

First Run
features: depression, anxiety, pain, migraine, insomnia, GAD >=10, PHQ >=10, MCS12, CCI Score
target: just psychedelic mushroom (PM) use

ACTIONS

- Checked for nulls, data types, dupes, outliers, correlations
- Removed irrelevant features
- Reviewed stats & distributions
- Scaled continuous variables
- Reviewed shape of data set(s)

DISCOVERIES

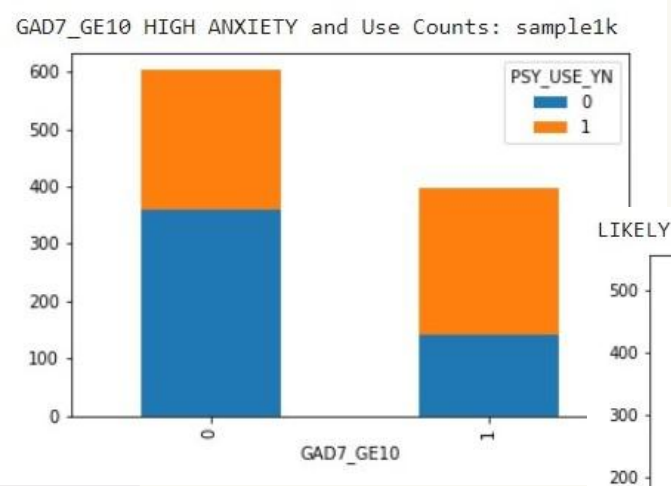
- Fairly clean data with no nulls/dupes, easy-to-manage outliers, and expected stats/distributions
- 7,139 rows with 122 mushroom-only users → **Imbalanced data set**

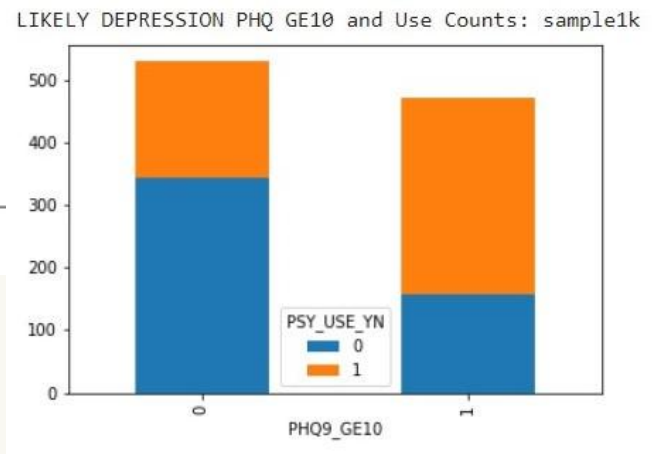
NEXT STEPS:

- resample data set to mitigate imbalanced data
- revised target to include “all psychedelic use” (not just mushrooms)

RESULTS: MH CONDITIONS

features: depression, anxiety, pain, migraine, insomnia, GAD >=10, PHQ >=10, MCS12, CCI Score
target: all psychedelics use
sample size: 1000 (500 users, 500 non-users, randomly generated, manually created)
model method: General Logistic Regression (GLR) from sklearn for Python





PRE-MODELING REVIEW:

Before any predictive modeling - higher percentage of use in people with higher evaluative scores on the GAD7 and PHQ9 tests.

Higher GAD scores indicate moderate-severe anxiety.

Higher PHQ9 scores indicate moderate-severe depression.

RESULTS: MH CONDITIONS

features: depression, anxiety, pain, migraine, insomnia, GAD >=10, PHQ >=10, MCS12, CCI Score
target: all psychedelics use
sample size: 1000 (500 users, 500 non-users, randomly generated, manually created)
model method: General Logistic Regression (GLR) from sklearn for Python

General Logistic Regression (GLR) Results

	sample1k GLR
ACCURACY	0.67
PRECISION	0.7
RECALL	0.66
F1	0.68
XVAL	0.52
BALANCE	BAL

67%
prediction accuracy

This model can predict psychedelic use with 67% accuracy for those people with certain mental health conditions (MHCs) and higher evaluative testing scores.

RESULTS: MH CONDITIONS

Imbalanced Data Mitigation: implemented 5 synthetic over/under resampling methods and modeled with GLR

- SMOTE
- RUS (Random Under Sample)
- ROS (Random Over Sample)
- Tomek
- Near Miss

For synthetic resampling, Accuracy score evaluation should be avoided and, instead, consider recall or precision score.

SYNTHETIC RESAMPLES: about the same across all scores

MANUAL SAMPLE1k: performs better for most scores, especially recall/precision

General Logistic Regression (GLR) Results for Synthetic Resampling

	sample1k GLR	RF	SMOTE	RUS	ROS	TOMEK	NEARMISS	KNN	SVC	SMOTE-BENCH	WHOLE LR
ACCURACY	0.67	0.63	0.68	0.67	0.68	0.68	0.5	0.6	0.66	0.93	0.92
PRECISION	0.7	0.65	0.13	0.13	0.14	0.14	0.1	0.58	0.14	0.0	0.0
RECALL	0.66	0.66	0.61	0.64	0.62	0.62	0.71	0.62	0.7	0.0	0.0
F1	0.68	0.66	0.22	0.22	0.66	0.22	0.17	0.6	0.24	0.0	0.0
XVAL	0.52	0.45	0.56	0.46	0.52	0.52	0.56	0.49	0.63	0.93	did not run
BALANCE	BAL	BAL	SYNTH BAL	SYNTH BAL	SYNTH BAL	SYNTH BAL	SYNTH BAL	BAL	IMBAL	IMBAL	IMBAL

AIMING FOR IMPROVEMENT

67%

prediction accuracy for use with just mental health conditions

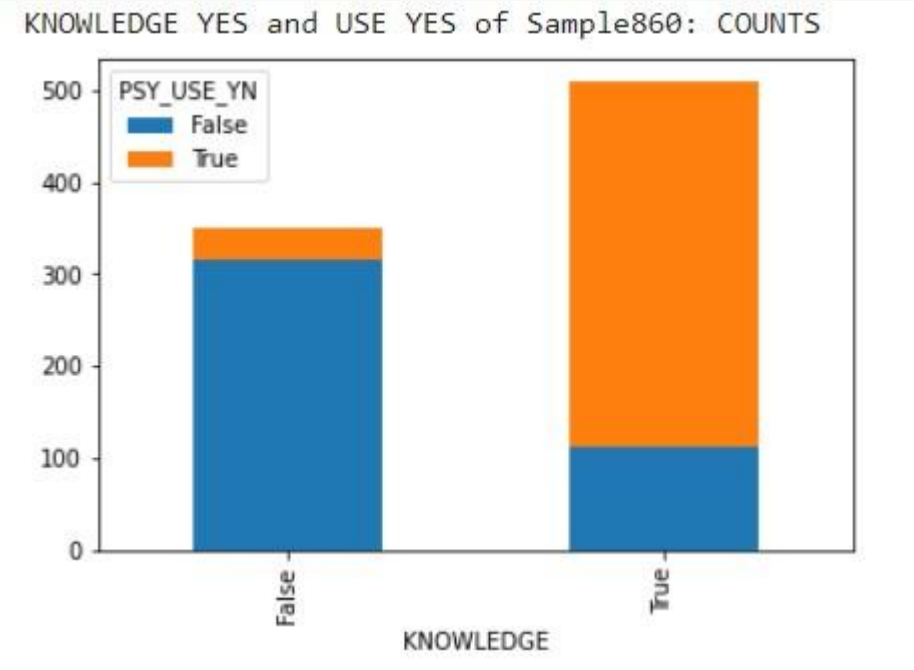
Model improvement strategy:

- Revised feature set to include recent prior knowledge of psychedelics as a therapeutic mental health modality

- Tuned random manual sample to include only “yes, heard of” and “no, have not heard of” answers to recent prior knowledge question (sample size reduced to 860)

RESULTS: MHCs & KNOWLEDGE

features: depression, anxiety, pain, migraine, insomnia, GAD >=10, PHQ >=10, MCS12, CCI Score, knowledge
target: all psychedelics use
sample size: 860 (430 users, 430 non-users, randomly generated)
model method: General Logistic Regression (GLR) from sklearn for Python



PRE-MODELING REVIEW:

Before any predictive modeling - higher percentage of use in people with recent prior knowledge of psychedelics as a mental health therapeutic.

Knowledge question:
"In the last 6 months heard more than usual about the positive uses of psychedelic drugs for mental health issues (depression, PTSD, addiction, etc.)"

RESULTS: MHCs & KNOWLEDGE

features: depression, anxiety, pain, migraine, insomnia, GAD >=10, PHQ >=10, MCS12, CCI Score, knowledge
target: all psychedelics use
sample size: 860 (430 users, 430 non-users, randomly generated)
model method: General Logistic Regression (GLR) from sklearn for Python

General Logistic Regression (GLR) Results

	WHOLE LR	860 LR
ACC	0.92	0.85
PREC	0.33	0.79
RECALL	0.06	0.94
F1	0.1	0.86
XVAL	did not run	0.76
BALANCE?	IMBAL	BAL

85%

prediction accuracy

This model can predict psychedelic use with 85% accuracy for those people with certain MHCs, higher evaluative testing scores, and prior knowledge of psychedelics as a therapeutic tool.

RESULTS: MHCs & KNOWLEDGE

MANUAL SAMPLE860: performs better for all scores

SYNTHETIC RESAMPLES: lower scores all around

OTHER MODELS:

- RANDOM FOREST
- KNN
- PENALIZED SVC

Predictive Modeling Results

	WHOLE LR	860 LR	SMOTE-BENCH	SMOTE	RUS	ROS	TOMEK	NEARMISS	RF	KNN	SVC
ACC	0.92	0.85	0.93	0.77	0.78	0.76	0.76	0.63	0.83	0.77	0.75
PREC	0.33	0.79	0.38	0.22	0.23	0.21	0.21	0.13	0.8	0.72	0.2
RECALL	0.06	0.94	0.02	0.81	0.82	0.84	0.84	0.73	0.88	0.86	0.81
F1	0.1	0.86	0.04	0.35	0.36	0.84	0.34	0.23	0.84	0.78	0.33
XVAL	did not run	0.76	0.91	0.72	0.67	0.69	0.69	0.67	0.8	0.69	0.73
BALANCE?	IMBAL	BAL	IMBAL	SYNTH BAL	SYNTH BAL	SYNTH BAL	SYNTH BAL	SYNTH BAL	BAL	BAL	IMBAL

PREDICTION COMPARISON

manual random samples compared to a few synthetic samples in relation to knowledge feature exclusion/inclusion

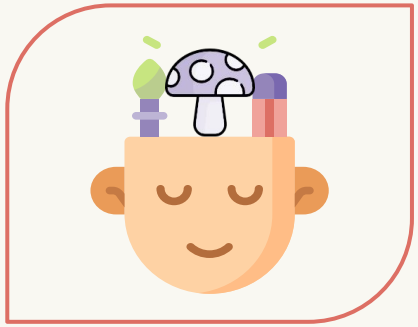
	NO KNOWLEDGE	KNOWLEDGE
sample1K	67%	✕
sample 860	✕	85%
RUS*	64%	82%
SMOTE*	61%	81%

* using Recall scores


CONCLUSION &
RECOMMENDATIONS

WHAT'S THE WHY?

why process this data?



- can predict with 85% accuracy that those with certain MHCs and knowledge are using psychedelics

- **establishes a connection** to recreational use as a form of self-medicated mental health care

The **connection between psychedelics and mental health** can advance science about:

- how these compounds work in relationship to mental health
- how to use them safely
- how to regulate and optimize

“KNOWLEDGE IS POWER”

The increase in use and predictability with Knowledge points to:

Knowledge can lead to use →
Use can lead to healing, treatment, relief.

KNOWLEDGE IS POWER

- Removes stigma and hesitation after 70+ years of fear-mongering, criminalization, and misinformation
- Fosters peace-of-mind regarding non-traditional options beyond current mental health approaches like talk therapy, psych medication, and relationships with those beholden to corporate influence
- Allows for the evolution of decriminalization and schedule reclassification, opening the door to more research
- Allows for statutory regulation and licensure to be established (ie Colorado Prop 122)
- Can shorten treatment periods from years and months to weeks and days
- Increases research funding and testing

RECOMMENDATIONS

- Process remaining data in this study about taking psychedelics with intention regarding mental health status
- Facilitate collection of post-use data regarding state of MHCs
- Future data collection including more mental health-related questions such as:
  - If currently/recently on psych meds
  - Seeing or seen in last year a mental health professional
  - Recent hospitalization for MHC
  - Suicidal ideation/attempts
  - Recent confirmed mental health diagnosis

RECOMMENDATIONS

- Mental health professionals, researchers, clinicians, data scientists, manufacturers, and legal professionals need to:
  - continue **collecting, processing and publishing data** regarding research, trials, and outcomes
  - continue **story-telling the results** of psychedelics as a mental health therapeutic tool
- **Positive mental health education campaigns** can advance the science and **change methods, modalities, and mindsets** related to mental health care.

FINAL CREDITS

DATA CALCULATIONS

- **Data processed by Lisa Girard in Jupyter Collab Notebook with Python**
- **Downloadable raw data set**
- Presentation template by Slidesgo, icons by Flaticon, infographics & some images by Freepik.

CONTACT INFO

http://lisagirard.com
lisaraygirard @ gmail.com

IMAGE SOURCES

- Speedy McVroom
- Gerd Altmann
- Rushikesh Gaikwad
- Gordon Johnson
- j4p4n