
How Education Level Influences Cannabis Consumptions

Claudia Tung, Will Everett, Alex Latendresse, Shuhan Wang, Cody Greene, Manu Lopez



Agenda

01 Problem &
Hypothesis

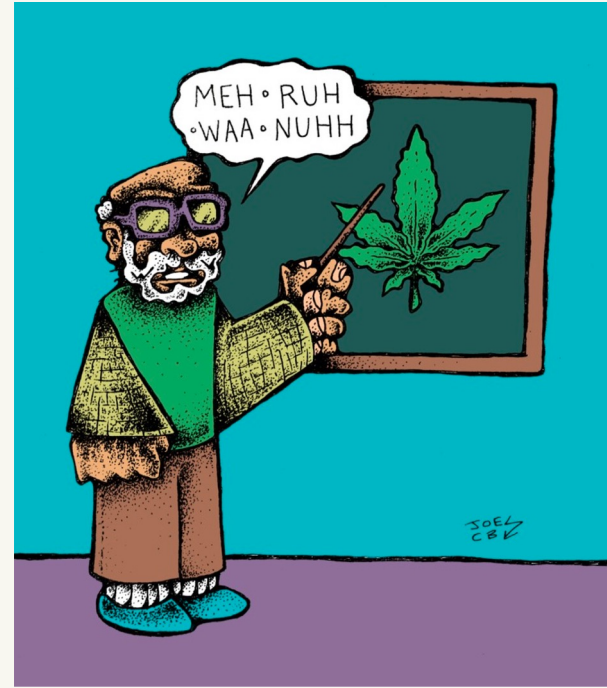
Data Driven
02 Approach & Ideal
Experiment

03 Data &
Exploratory Analysis

04 Models &
Conclusions

The Problem

- Importance of understanding how education affects cannabis consumption
 - Create policies around drug usage
 - Teach responsible drug consumption
 - Develop proper drug prevention programs
 - Improve rehab facilities



Hypothesis

H_0 : Not pursuing higher education **has no effect** on the likelihood that someone will consume cannabis within the next month.

H_A : Not pursuing higher education **has an effect** on the likelihood that someone will consume cannabis within the next month.

Why a Data Driven Approach

- Informs important business decisions, as potential strategies are backed up with evidence
 - Potential systematic **biases** present
 - Conclusions and implications may be affected/**incorrect**
- Awareness of our model, after best attempting to eliminate these biases, better informs us about potential solutions to drug usage trends

The logo features the words "DATA" and "DRIVEN" in a bold, sans-serif font. The word "DATA" is positioned above "DRIVEN". The letters of "DATA" are filled with a pattern of small, light blue binary code (0s and 1s). The word "DRIVEN" is solid white. The entire logo is set against a dark teal rectangular background.

DATA
DRIVEN

Ideal Experiment

- Policy makers force universities to **randomly admit** people so there can be two groups:
 - Educated
 - Not educated
- All countries listed have legalized cannabis
- Use surveys to understand people's demographics, personality and cannabis consumption
 - **Everyone told the truth**
- Separate the results by different educational levels and correlate that with cannabis consumption
- Everyone that used cannabis only used **recreationally**



Our Dataset

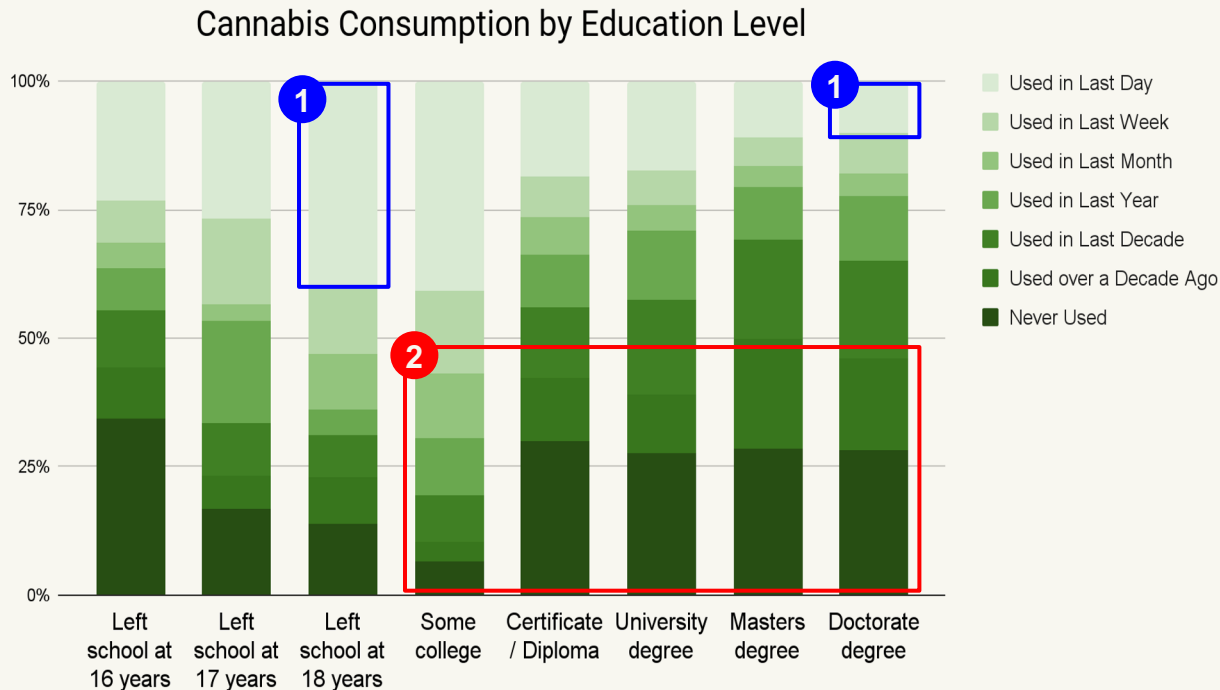
	Variables	Description
Dependent Variable	Cannabis	Level of cannabis consumption; <i>Never Used, Used over a Decade Ago, Used in Last Decade, Used in Last Year, Used in Last Month, Used in Last Week, Used in Last Day</i>
Independent Variables	Education	Highest education level obtained by participant
	Gender	Gender of participant
	Country	Country of origin of participant
	Ethnicity	Ethnicity of participant

Data source: https://www.kaggle.com/datasets/obeykhadija/drug-consumptions-uci?select=Drug_Consumption_Quantified.csv

Understanding the Data

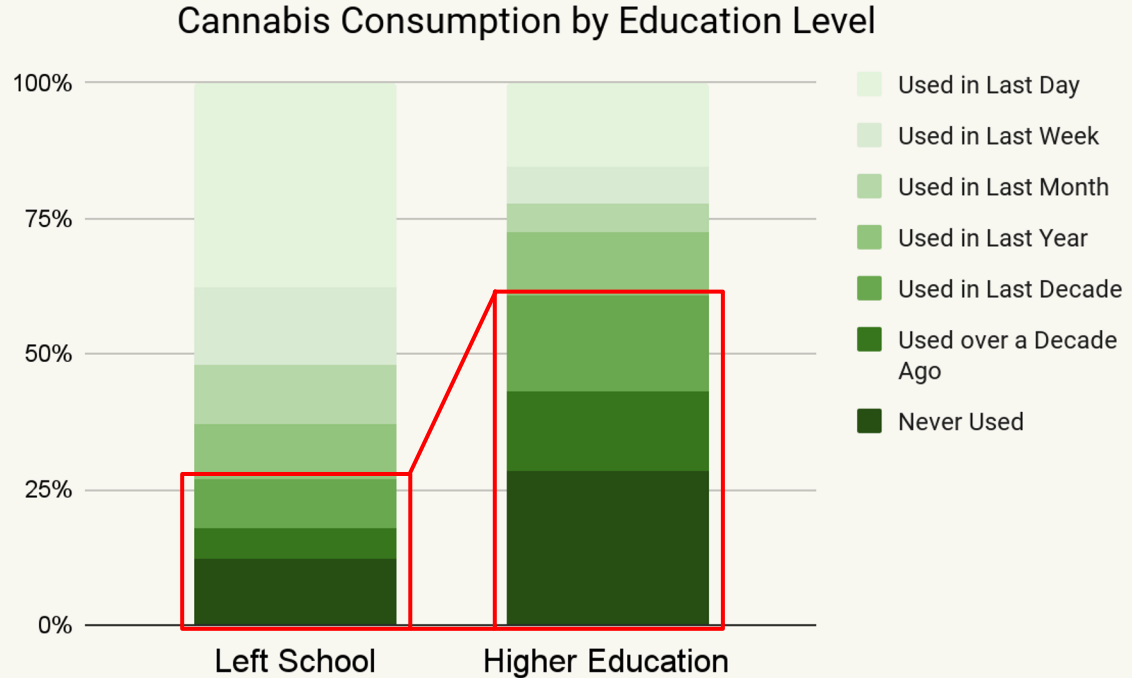
Findings

1. “Left school at 18 years” category has a significantly higher response rate of “Used in Last Day” when compared to “Doctorate degree” category.
2. Individuals that have completed a higher education seem to consume cannabis less regularly



Understanding the Data (cont.)

After grouping education level into these categories, it seems clear that “Higher Education” consumes cannabis less regularly.



Data Cleaning and Data Manipulation

01

Categorical Explanatory Variables

Created dummy variables to handle categorical explanatory data

02

Categorical Response Variable

Converted categorical response variable to binary data; 1 if consumed cannabis within past month, 0 if not

03

Unequal Distribution

Divided education levels into two buckets: *Left School* & *Higher Education*

Education Level	Count
Left school before 16 years	28
Left school at 17 years	30
Doctorate degree	89
Left school at 16 years	99
Left school at 18 years	100
Professional certificate/ diploma	269
Masters degree	283
University degree	480
Some college or university, no certificate or degree	506
Grand Total	1884



Education Level	Count
Higher Education	1121
Left School	763
Grand Total	1884

Results

Model 1

R-squared: 0.124
Adj. R-squared: 0.124
Prob (F-statistic) <0.000

Treatment	coeff	t	P> t
Intercept	0.2748	19.919	0.000
Education[T. Left School]	0.3543	16.348	0.000

Results

Model 2

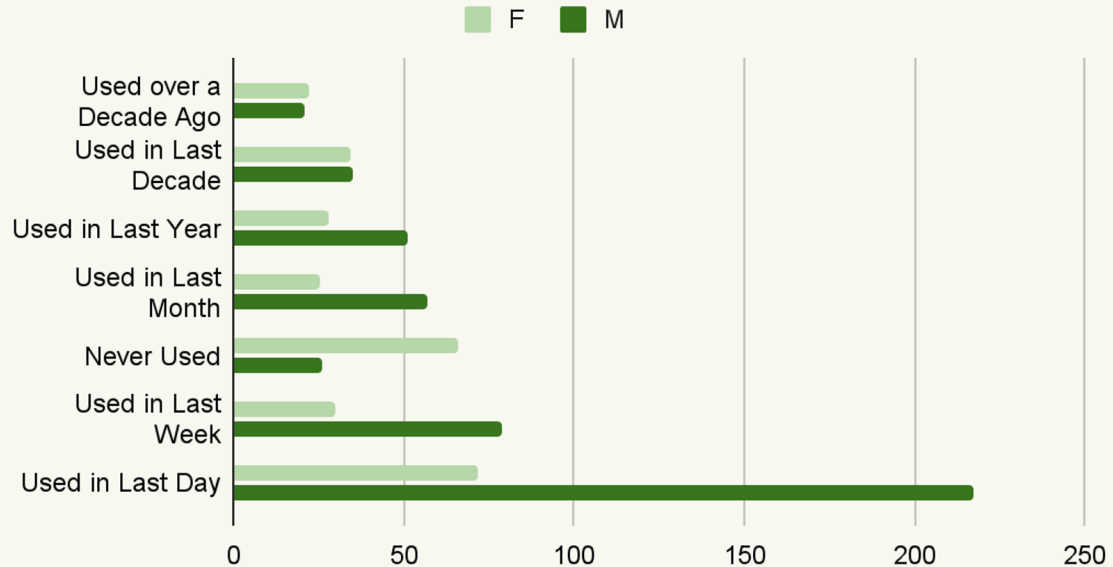
R-squared: 0.325
Adj. R-squared: 0.320
Prob (F-statistic) <0.000

Treatment	coeff	t	P> t
Intercept	0.3095	3.185	0.001
Education[T. Left School]	0.1683	8.101	0.000
Gender[T.M]	0.1489	7.632	0.000
Ethnicity[T.Black]	0.0646	0.605	0.545
Ethnicity[T.Mixed-Black/Asian]	0.4949	1.989	0.047
Ethnicity[T.Mixed-White/Asian]	0.2947	2.391	0.017
Ethnicity[T.Mixed-White/Black]	0.1361	1.122	0.262
Ethnicity[T.Other]	0.2316	2.425	0.015
Ethnicity[T.White]	0.1604	1.989	0.047
Country[T.Canada]	-0.1272	-1.801	0.072
Country[T.New Zealand]	0.1436	0.755	0.450
Country[T.Other]	-0.0405	-0.604	0.546
Country[T.Republic of Ireland]	-0.0028	-0.026	0.979
Country[T.UK]	-0.3745	-6.567	0.000
Country[T.USA]	0.0708	1.214	0.225

Further Analysis

Cannabis Consumption by Gender (Left School Only)

When filtering for education level equal to "Left School", there appears to be a stark difference in cannabis consumption by gender

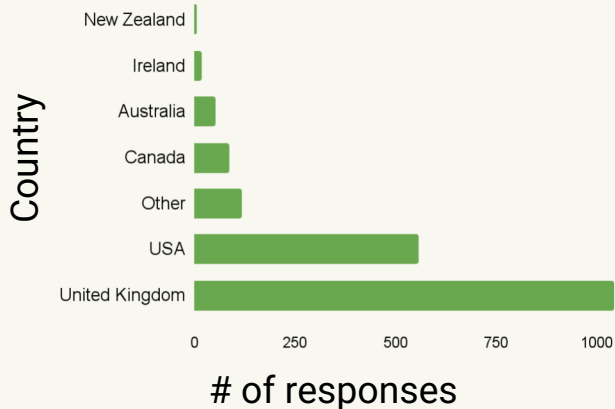


Conclusion

We **reject** the null hypothesis.

There is sufficient evidence to suggest that not pursuing higher education will increase the likelihood that a person consumes cannabis within the next month.

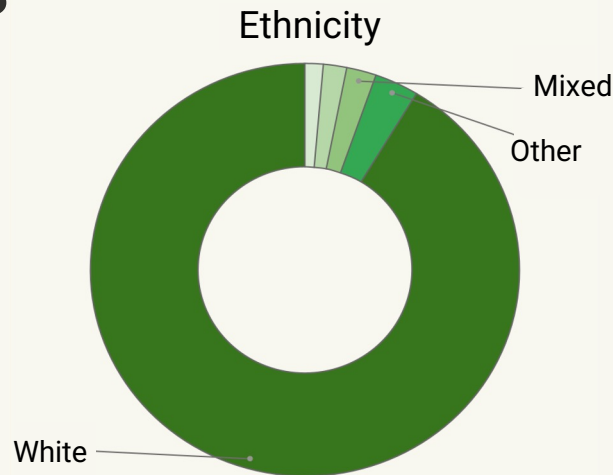
Improvements



More countries/ethnicities

Data on legality

Greater sample size



Results

Model 3

R-squared: 0.327
Adj. R-squared: 0.322
Prob (F-statistic) <0.000

Treatment	coeff	t	P> t
Intercept	0.3303	3.391	0.001
Education[T. Left School]	0.1143	3.827	0.000
Gender[T.M]	0.1099	4.407	0.000
Ethnicity[T.Black]	0.0590	0.554	0.580
Ethnicity[T.Mixed-Black/Asian]	0.4900	1.972	0.049
Ethnicity[T.Mixed-White/Asian]	0.2839	2.304	0.021
Ethnicity[T.Mixed-White/Black]	0.1123	1.009	0.313
Ethnicity[T.Other]	0.2315	2.428	0.015
Ethnicity[T.White]	0.1552	1.927	0.054
Country[T.Canada]	-0.1273	-1.805	0.071
Country[T.New Zealand]	0.1413	0.744	0.457
Country[T.Other]	-0.0401	-0.599	0.549
Country[T.Republic of Ireland]	0.0008	0.008	0.994
Country[T.UK]	-0.3743	-6.572	0.000
Country[T.USA]	0.0708	1.215	0.224
Education[T. Left School]:Gender[T.M]	0.0989	2.511	0.012

Based on the regression result, t-value for Education is 3.827 which is greater than the critical value 1.96 at the 95% confidence level

Thank you!

