2021 ASA DataFest

Topic

presented by Tomato303

Jaekang Lee, Alexander MacKenzie, Chang Sun, Tzuying Yu

ASA DataFest

UNIVERSITY OF TORONTO

1. Introduction

1.1 Background of the data:

- The RADARS® System Survey of Non-Medical Use of Prescription Drugs Program of Canada launched in 2017 is used in this report
- Total of 10,007 online participants

1.2 Aim:

- Understand the non-medical use of prescription drugs
- Predict the potential pattern of misusing prescription drugs to prevent their diversion into illicit channels.

among Canadian population.

2. Data Exploration

2.1 Assumptions:

- We assumed all surveys were answered truthfully and recorded without errors
- We do not use the weighted variable in the report

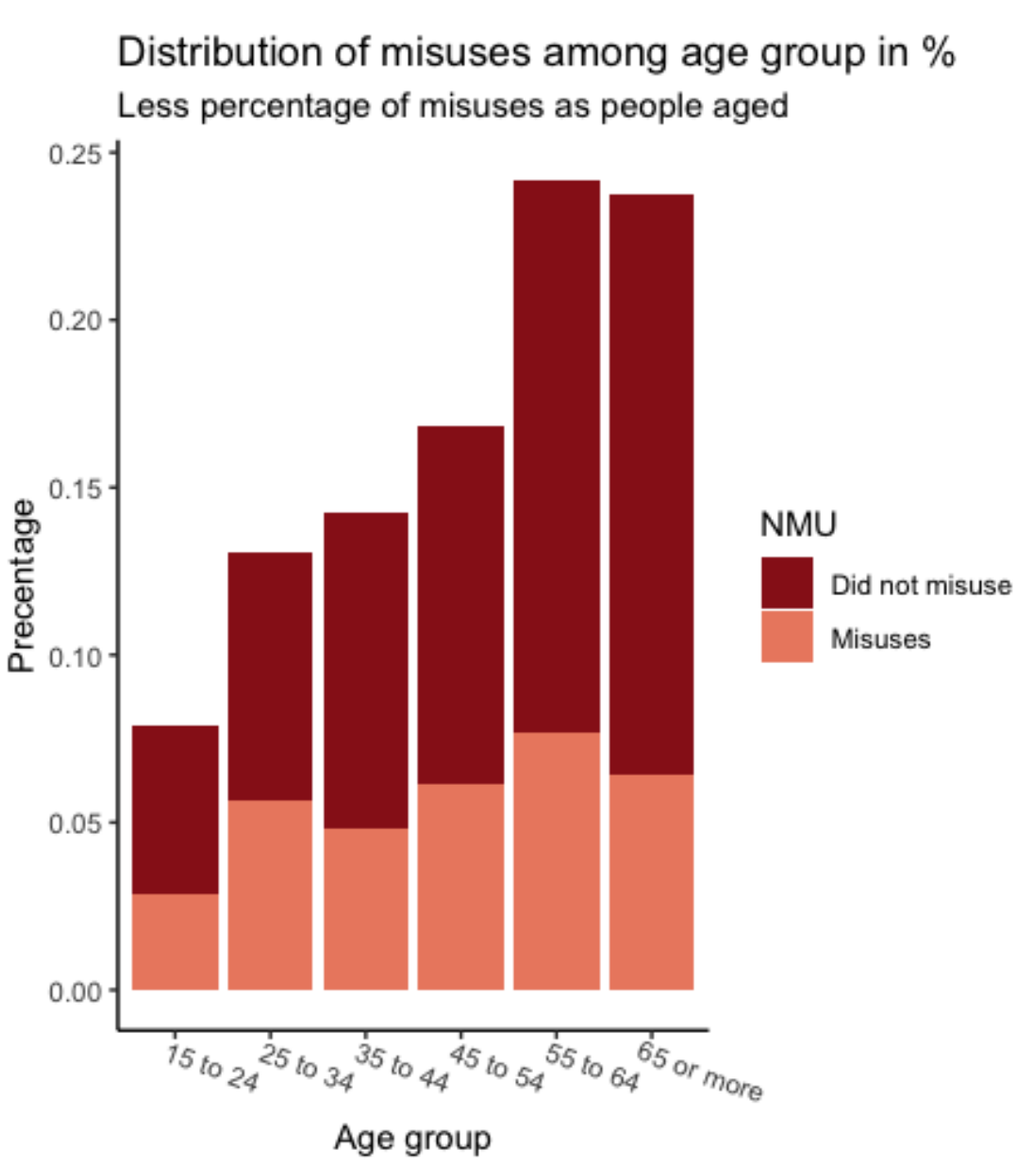
2.2 Data Wrangling:

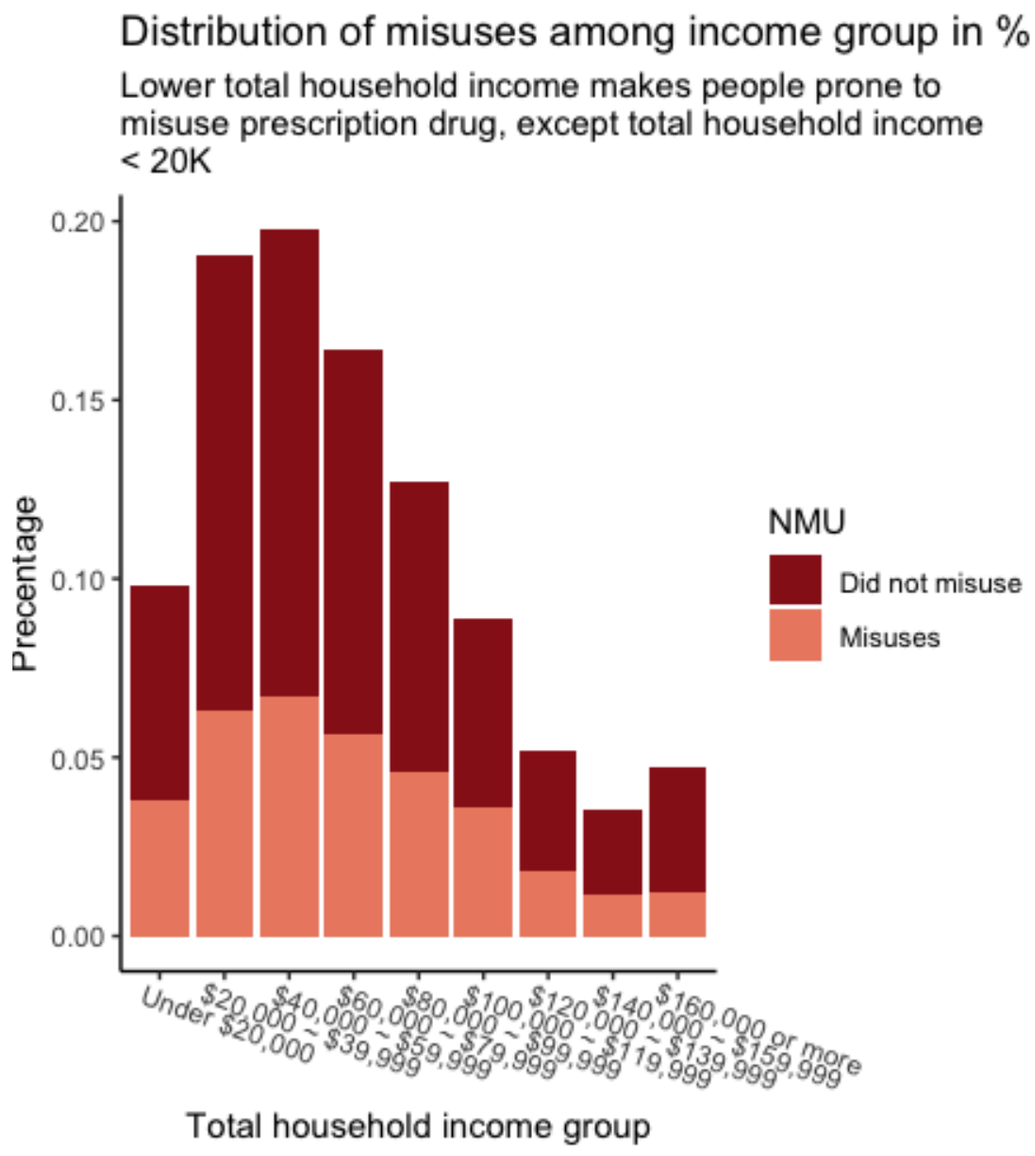
- We define non-medical use of prescription drug (NMU) as 1 if respondents have at least once answered “Yes” to any kind of Non-medical use survey question. 0 Otherwise

- We fill the missing value with best logical answers.
  For example, variables on pregnancy for all males were ‘NA’. We fill the missing value with 0.

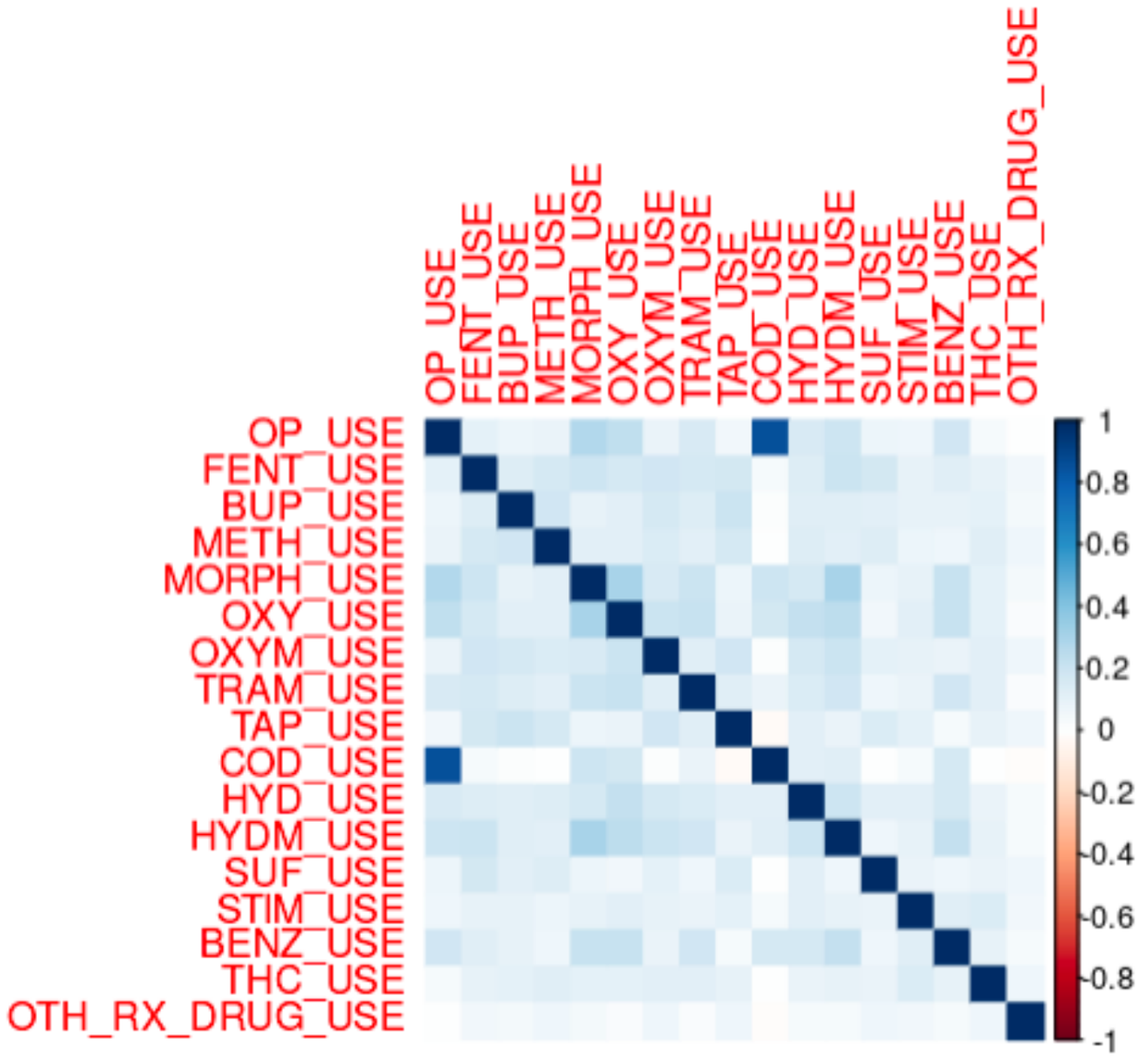
2.3 Visualizations:

- ## 2.3.1 Demographic features and the NMU

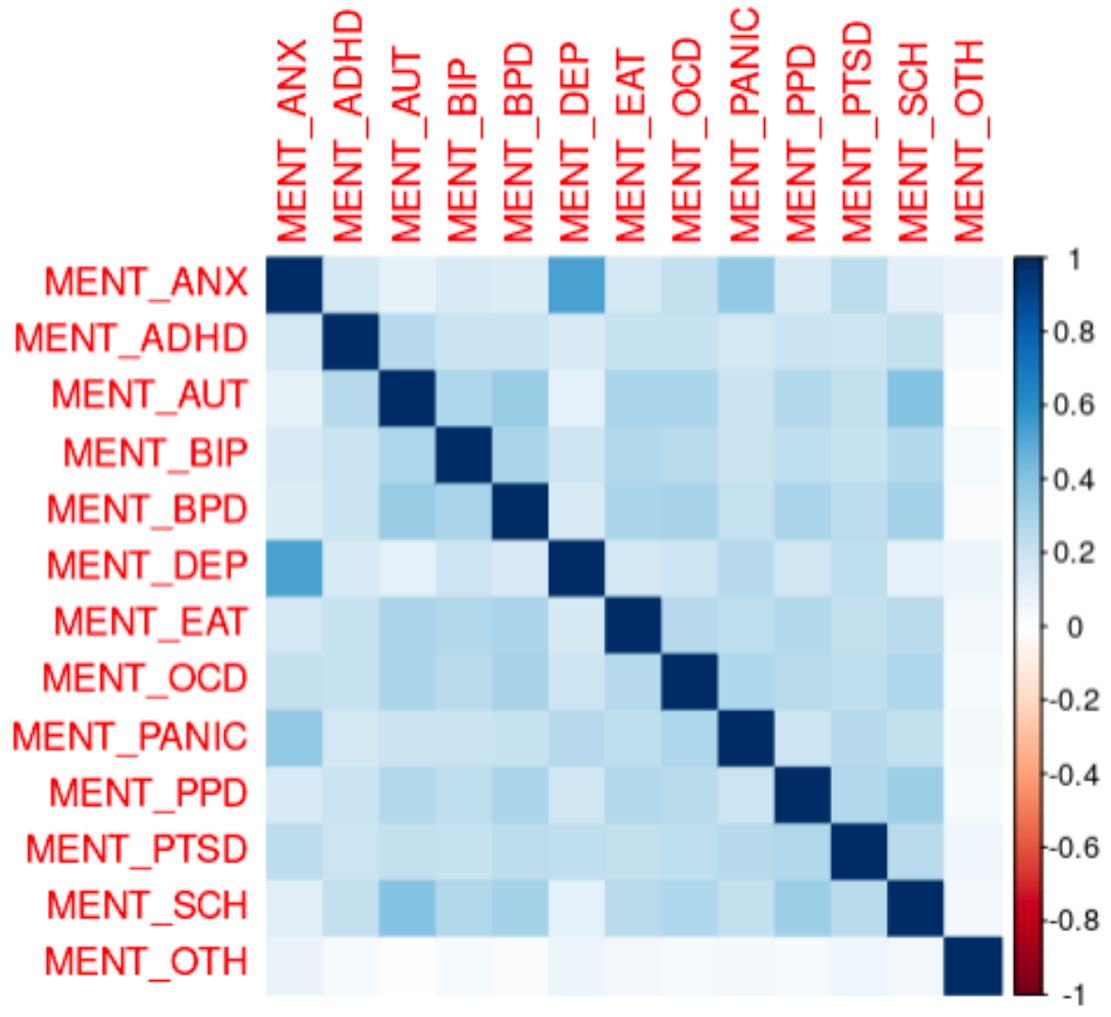




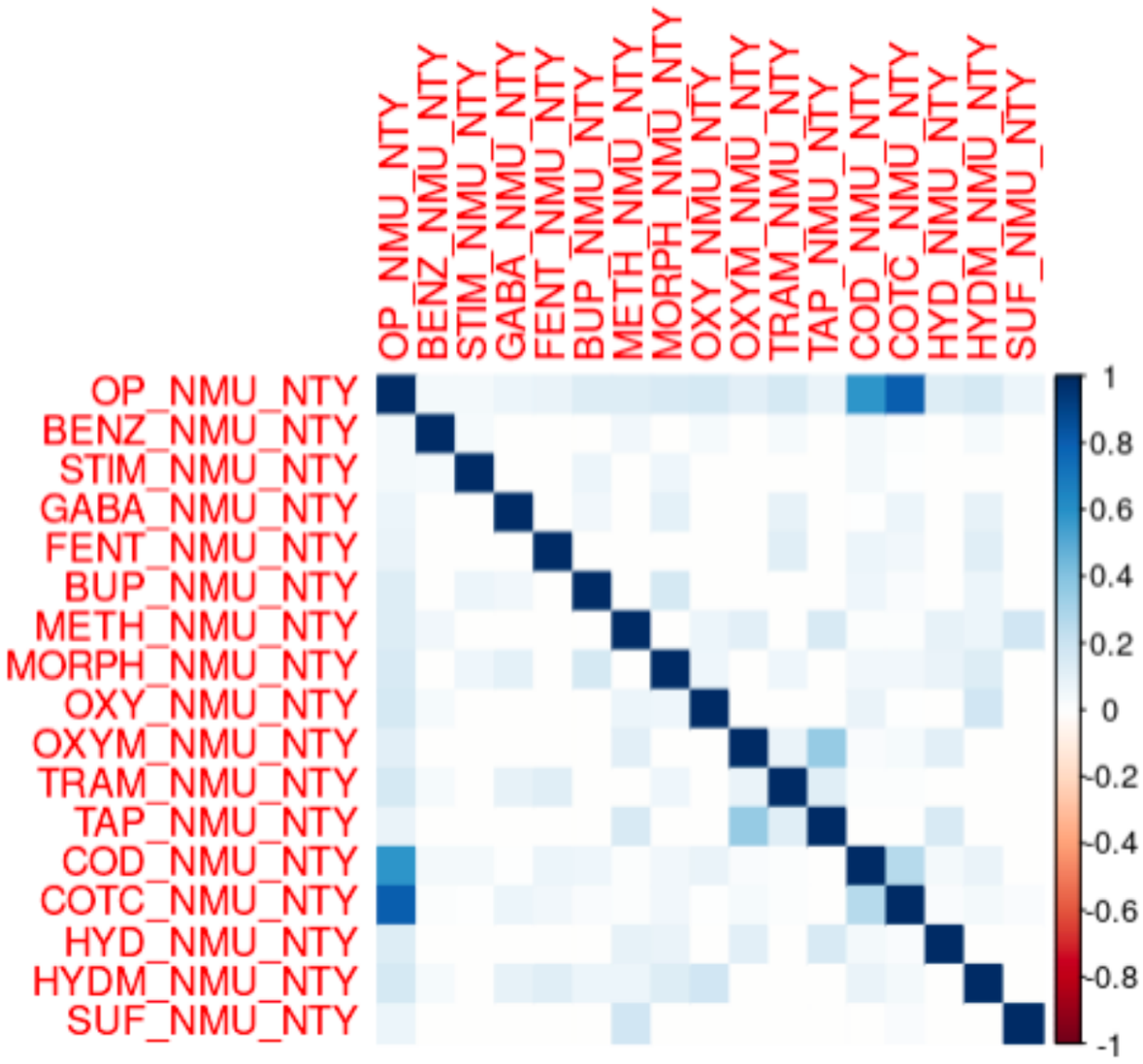
- **2.3.2 Substance use**



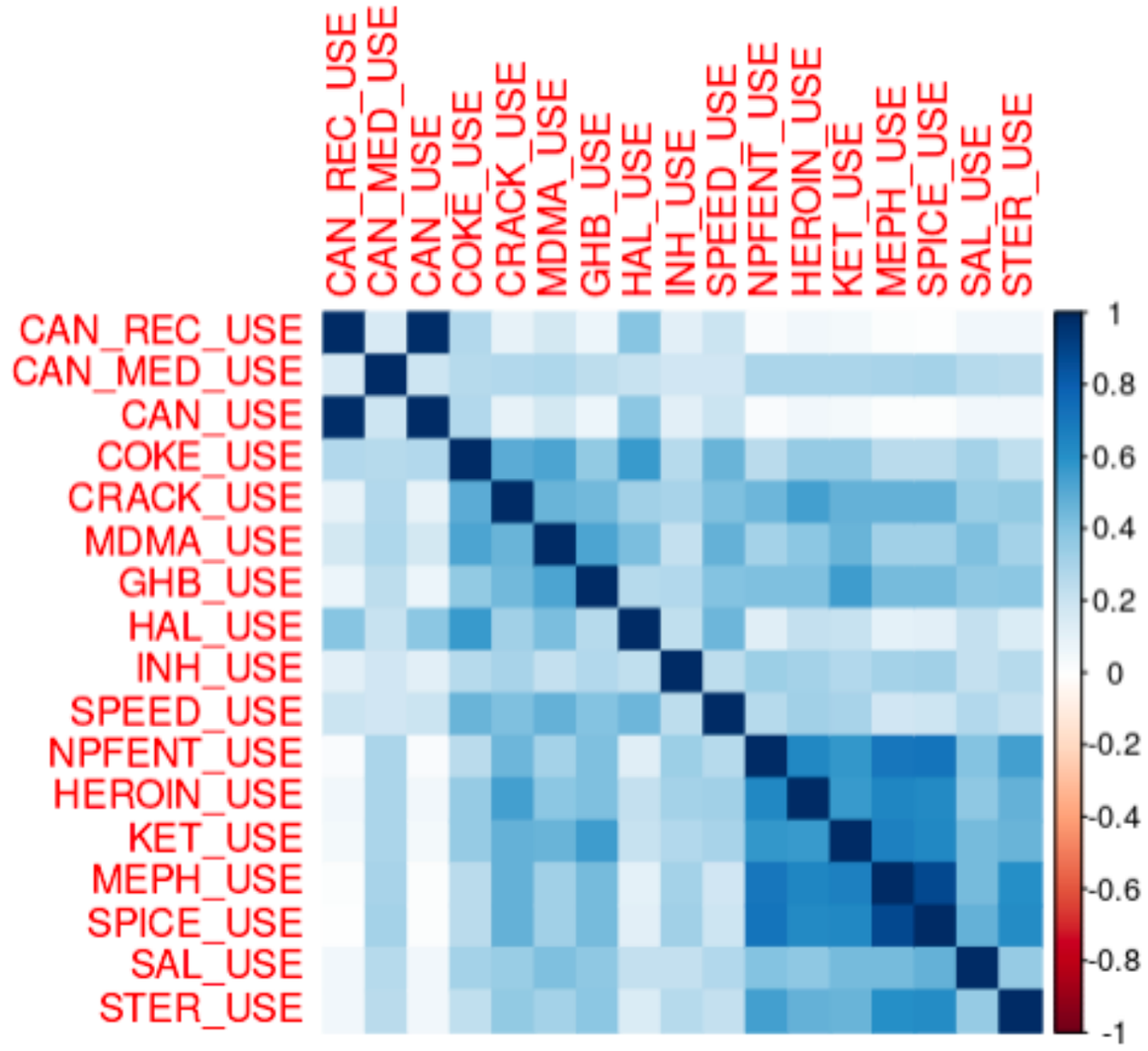
- **2.3.3 Mental disorders**

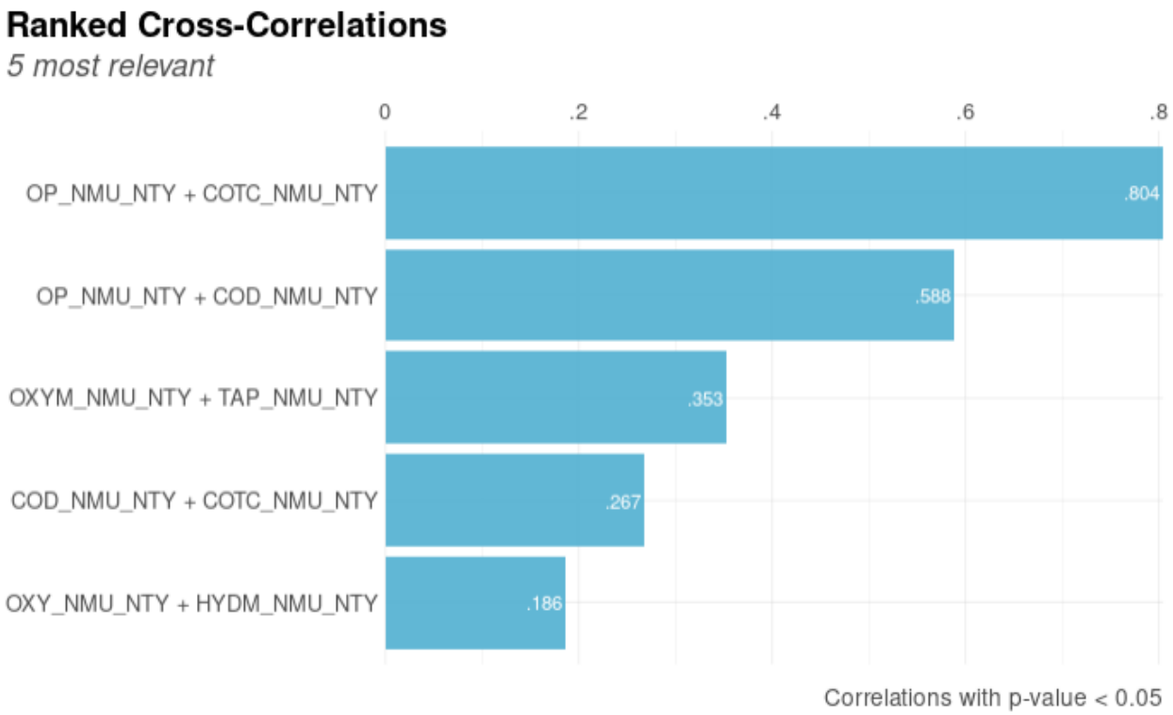


- **2.3.4 Substance NMU 90 days**



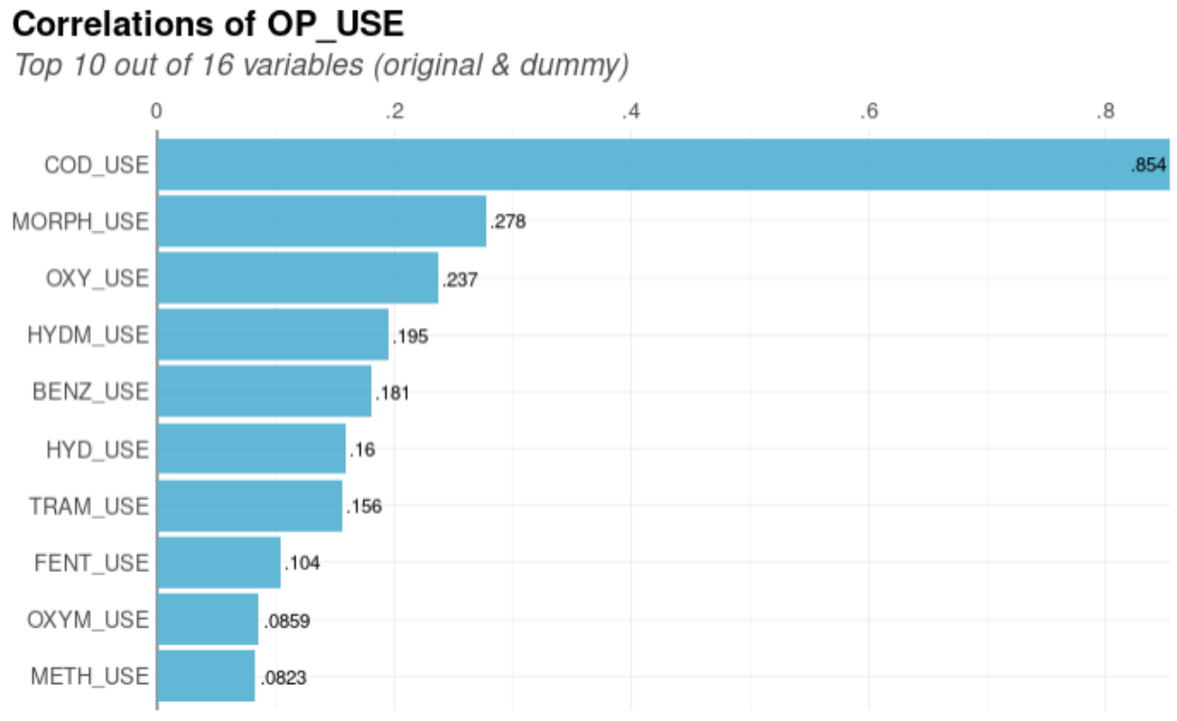
- **2.3.5 Frequency of Substance use**





2.3.6 Ranked Correlation between substance nonmedical use

2.3.7 Ranked correlation of Opioid use

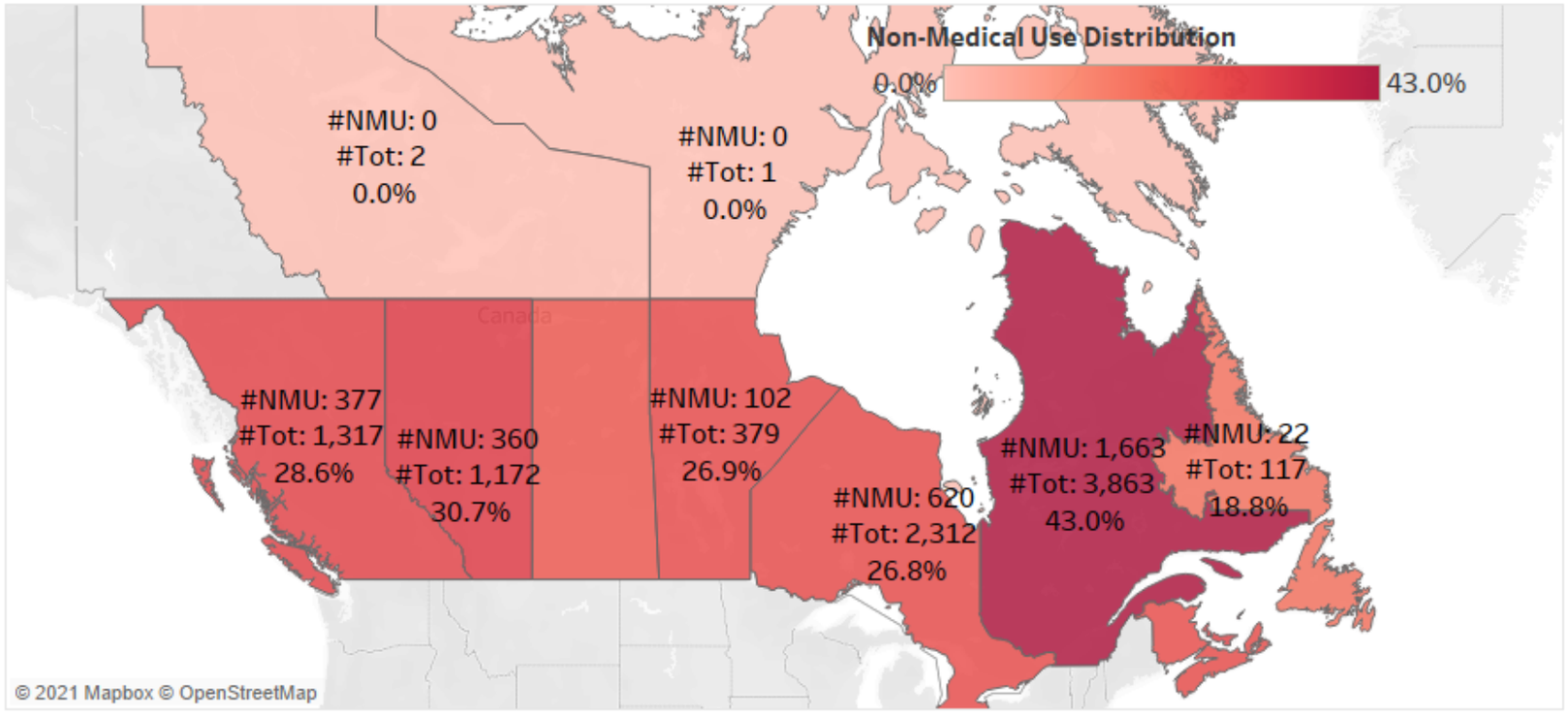


3. Research Questions

- **3.1 What types of social groups best explains non-medical use?**

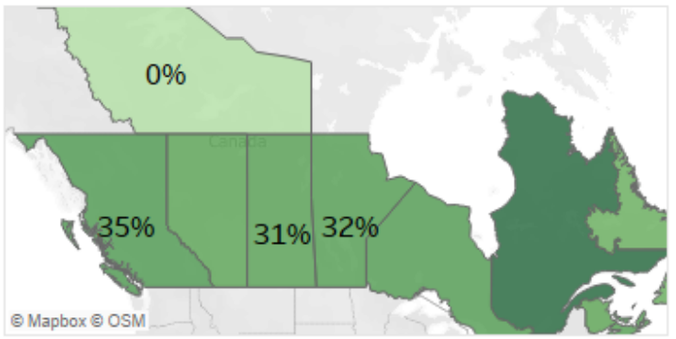
2021 ASA DataFest

Non-Medical Used Drugs in Canada Distribution

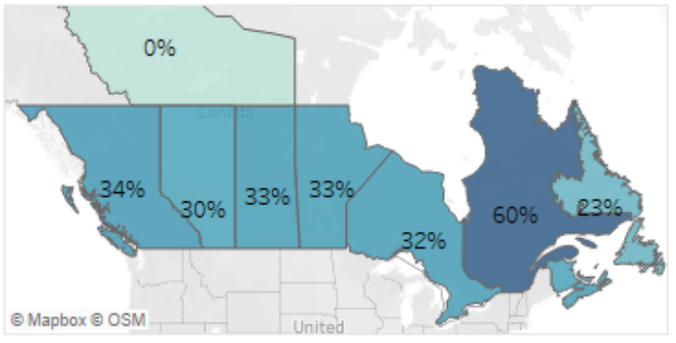


Top Three NMU Drugs

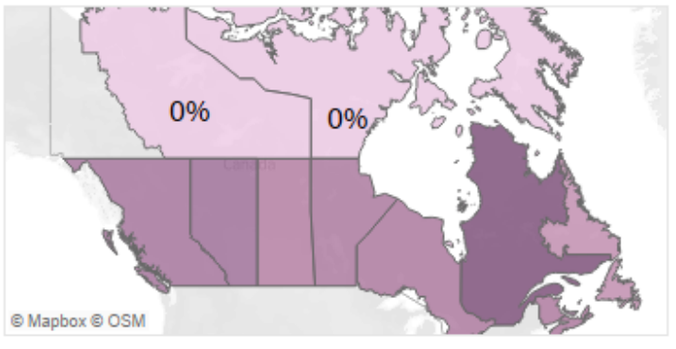
1. Opioid API NMU Dist.



2. Codeine NMU Dist.



3. Cocaine NMU Dist.

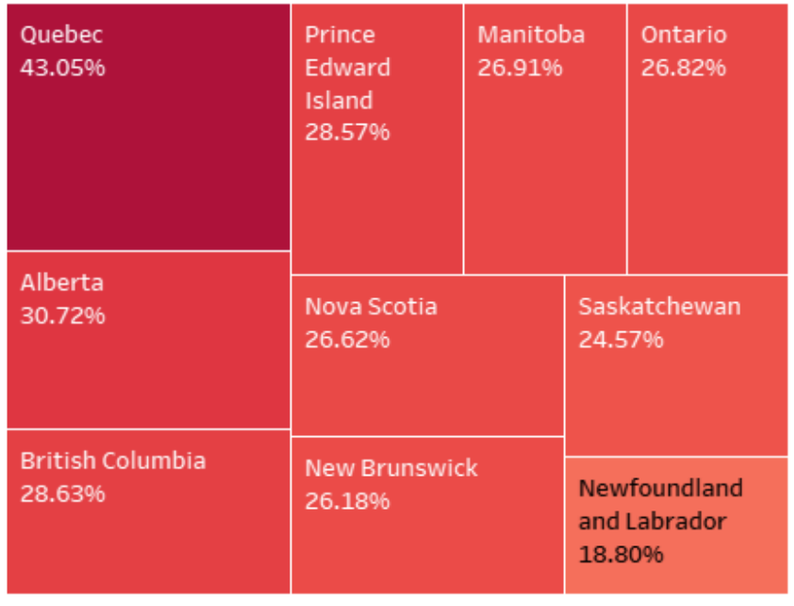


NMU | OP 0% 57%

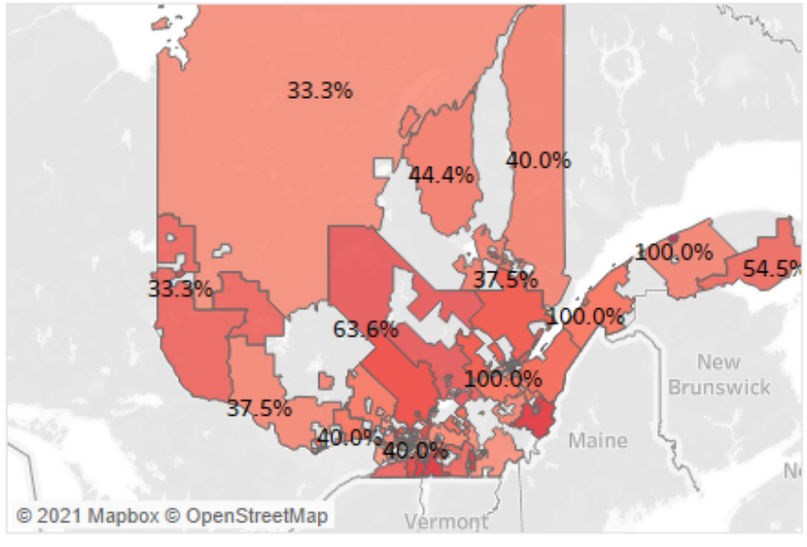
NMU | COD 0% 60%

NMU|Coke 0% 43%

NMU Distribution by Regions



Quebec: Most Severe NMU



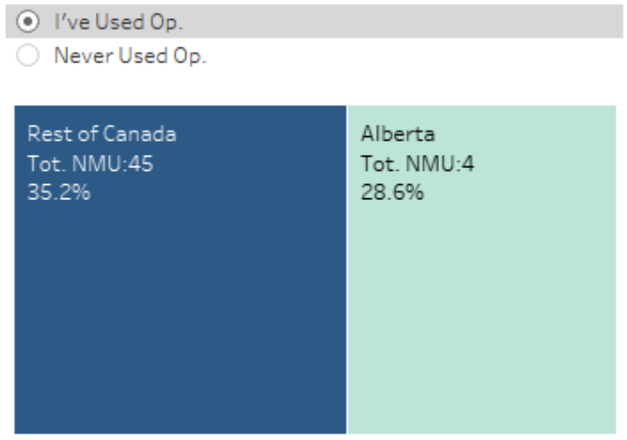
Avg. NMU 0.0% 100.0%

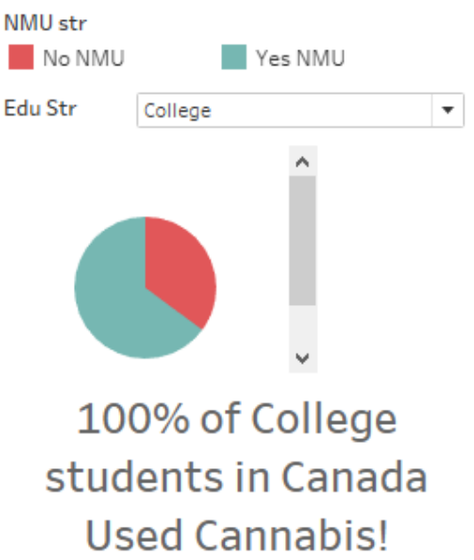
2021 ASA DataFest

Top NMU Social Groups

Group 2. Alberta College Students who adores Cannabis but stays away from Op.&Cod.

College NMU Alberta vs Canada given Op Use

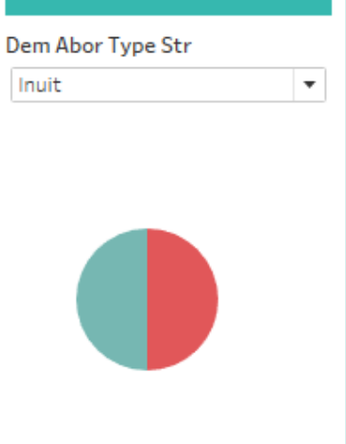




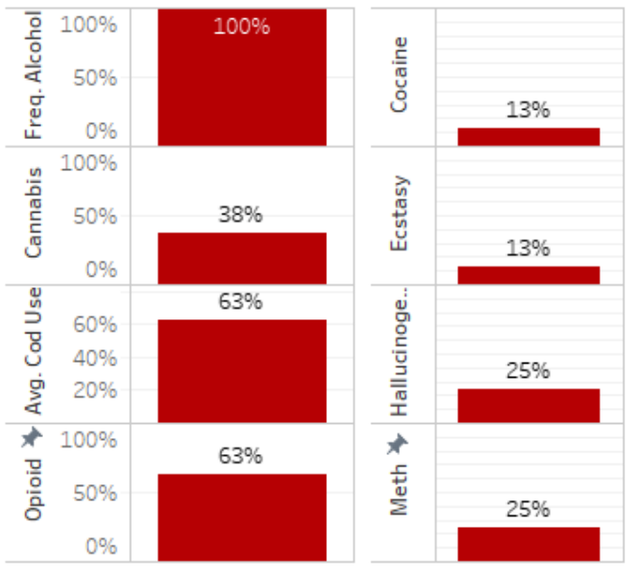
100% of College students in Canada Used Cannabis!

Group 3. Inuits who likes to try all types of drugs and frequently use OP. & Cod.

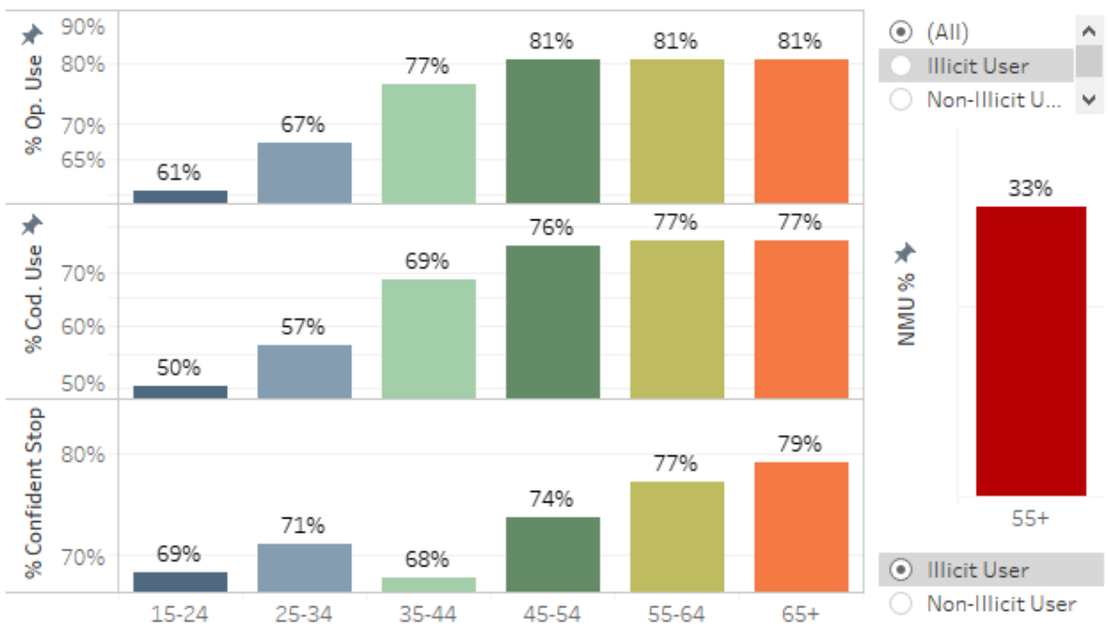
NMU dist. over Inuits



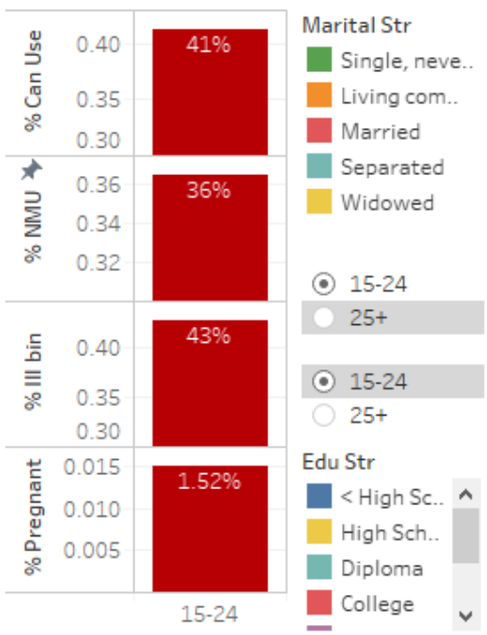
Inuits and their Usage of different types of drugs %

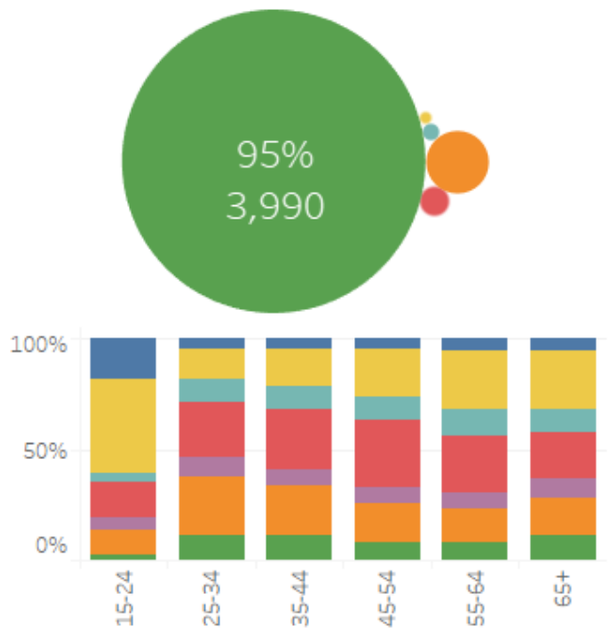


Group 4. Middle-Aged(55+) Illicit drug users who are confident that they can stop anytime.



Group 6. **Single and never married young people (15-24 years old). Mostly High School Graduates.**



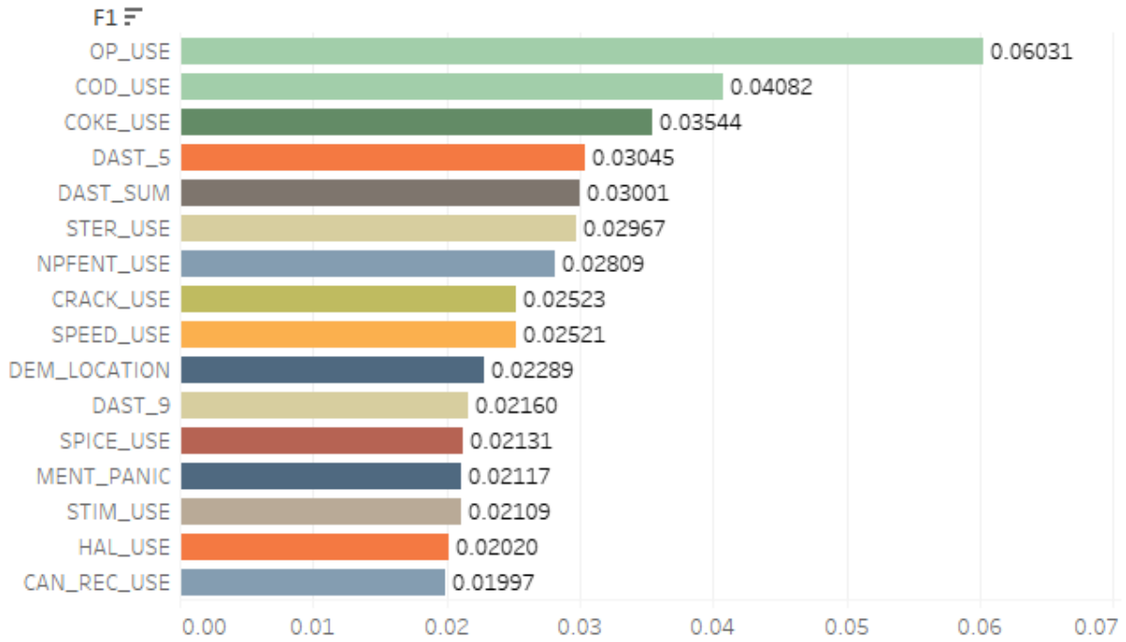


- **3.2 What variables explains NMU variable?**

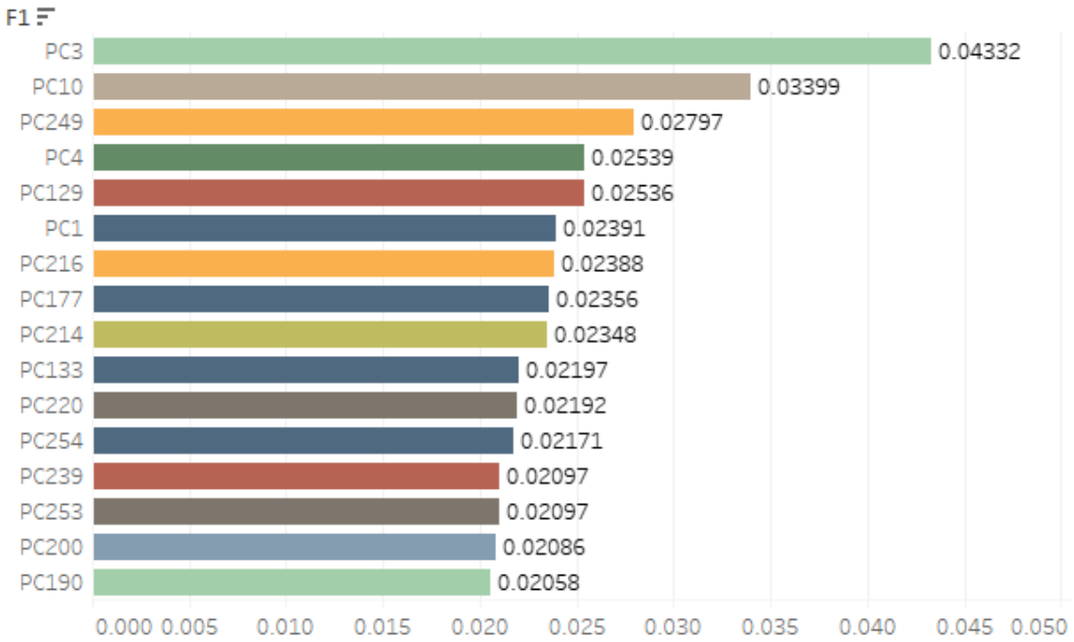
2021 ASA DataFest

Finding Social Groups with Principal Component Analysis and MI Scores

Top Mutual Information of 'NMU'

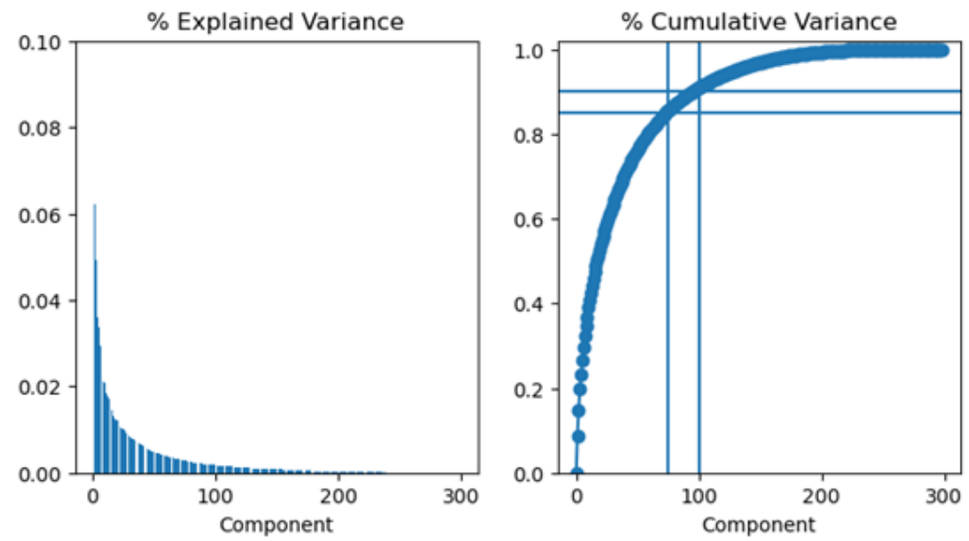


Top Mutual Information of 'NMU' with Principal Components

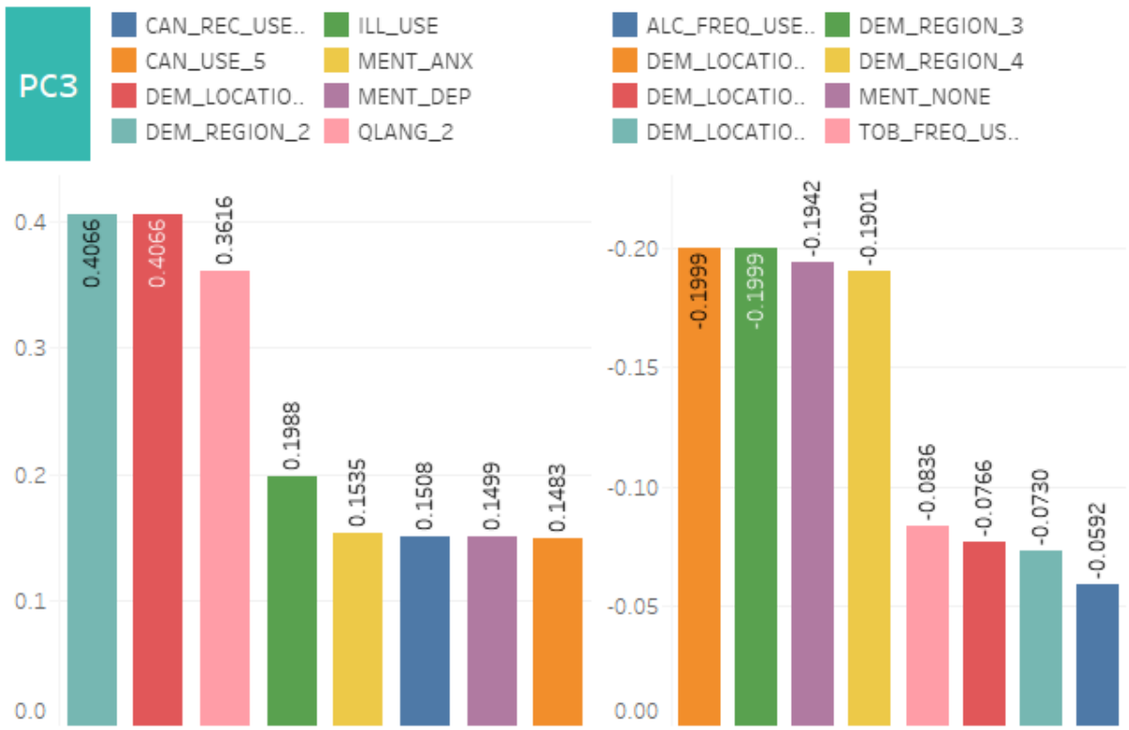


There are more than 4.9e+55 combination of groups in total!

Principal Component Variance Capture Plots



PC3



3.2.2 Fitted Model: Top two NMU drugs via MI score

3.2.2.1 The coefficient table for the model for NMU based on opioid and codeine use

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-3.423	0.117	-29.277	0.000
OP_USE	2.425	0.148	16.369	0.000
COD_USE	0.576	0.096	5.982	0.000
NPFENT5	0.603	0.469	1.284	0.199
NPFENT4	3.141	1.201	2.615	0.009
NPFENT3	1.259	0.915	1.375	0.169
NPFENT2	0.780	0.572	1.364	0.173
GHB5	0.836	0.187	4.459	0.000
GHB4	0.941	0.684	1.376	0.169
GHB3	2.041	0.593	3.440	0.001
GHB2	0.535	0.673	0.795	0.427
HEROIN5	0.688	0.254	2.714	0.007
HEROIN4	4.815	1.333	3.614	0.000
HEROIN3	1.966	0.745	2.638	0.008
HEROIN2	1.604	0.596	2.692	0.007
STIM_USE	1.192	0.107	11.120	0.000

Result:

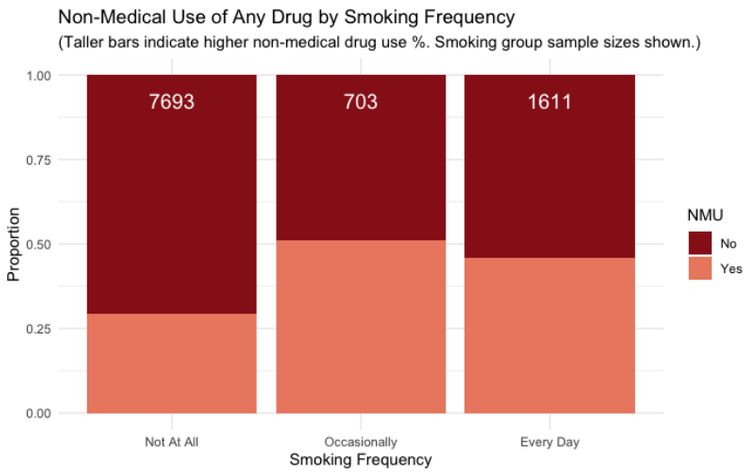
1030.6% increase in odds of misusing drugs between individuals who've never used prescription opioids to individuals who have

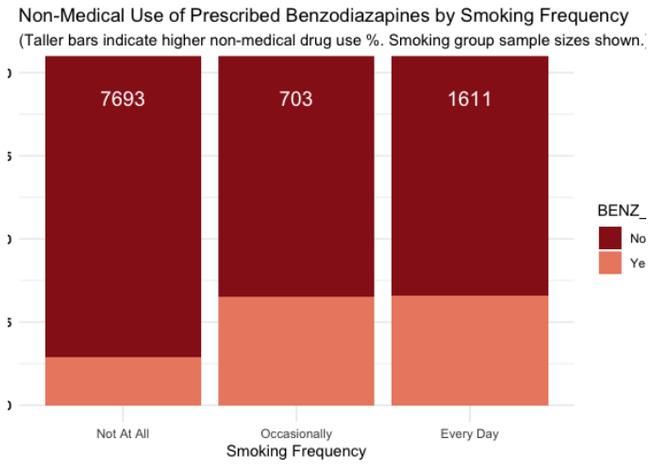
Odds of lifetime codeine users having misused any drug is 1.78 times the odds of lifetime non-codeine users having misused drugs

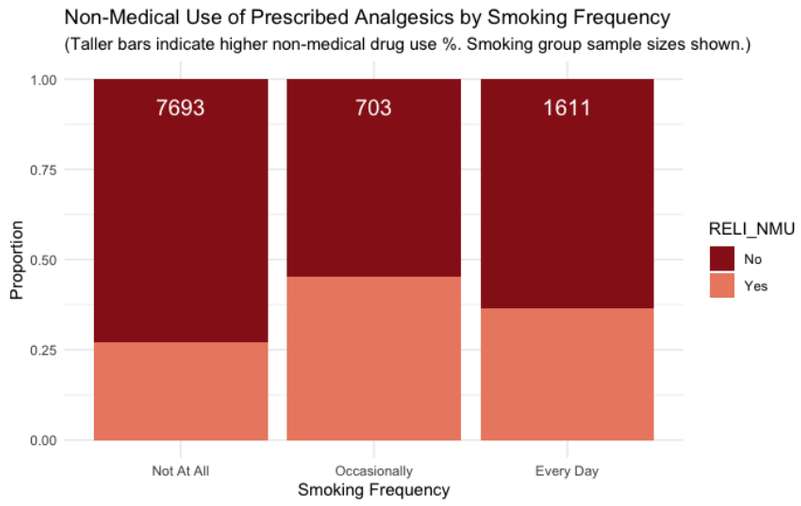
- **3.3 How does frequency of smoking and drinking impact drug use, specifically drug misuse and drug use recency?**

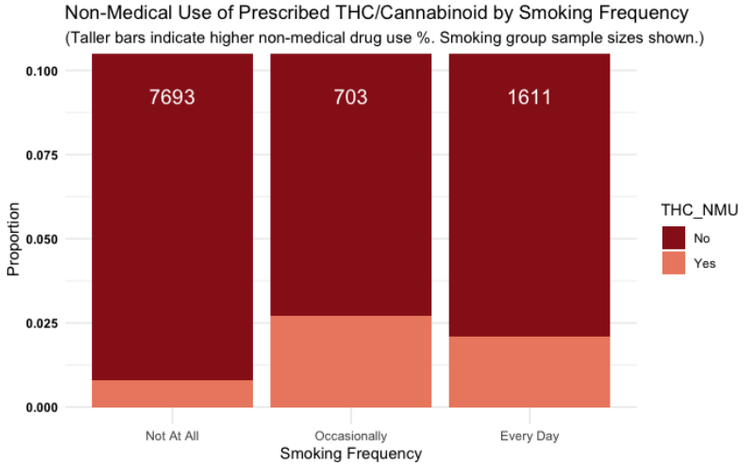
3.3.1 Visualization

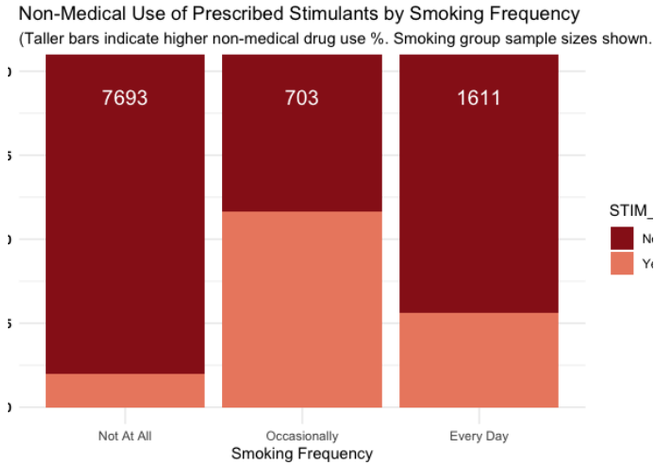
3.3.1.1 Smoking I



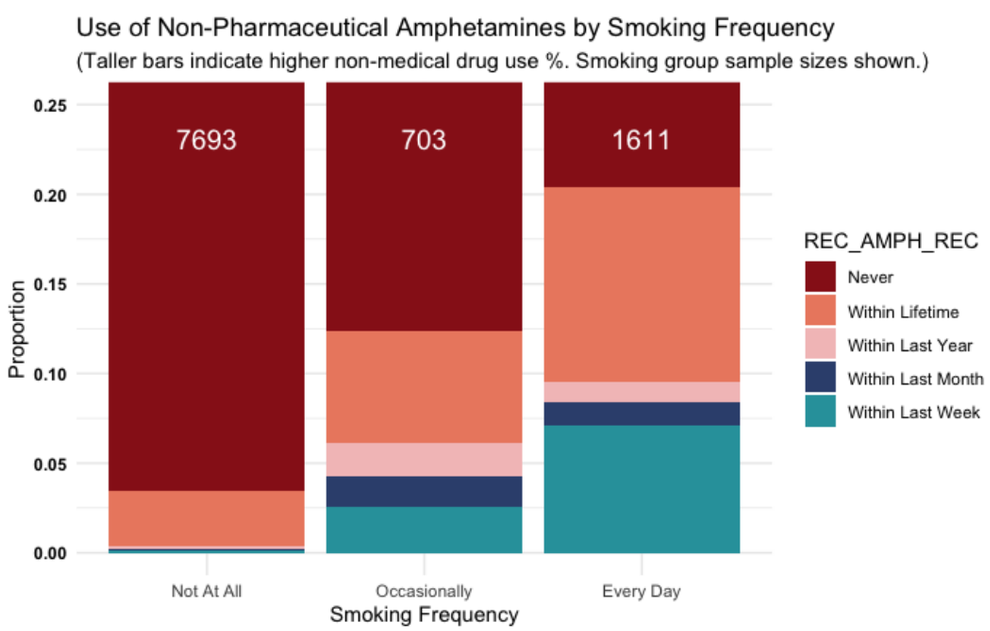


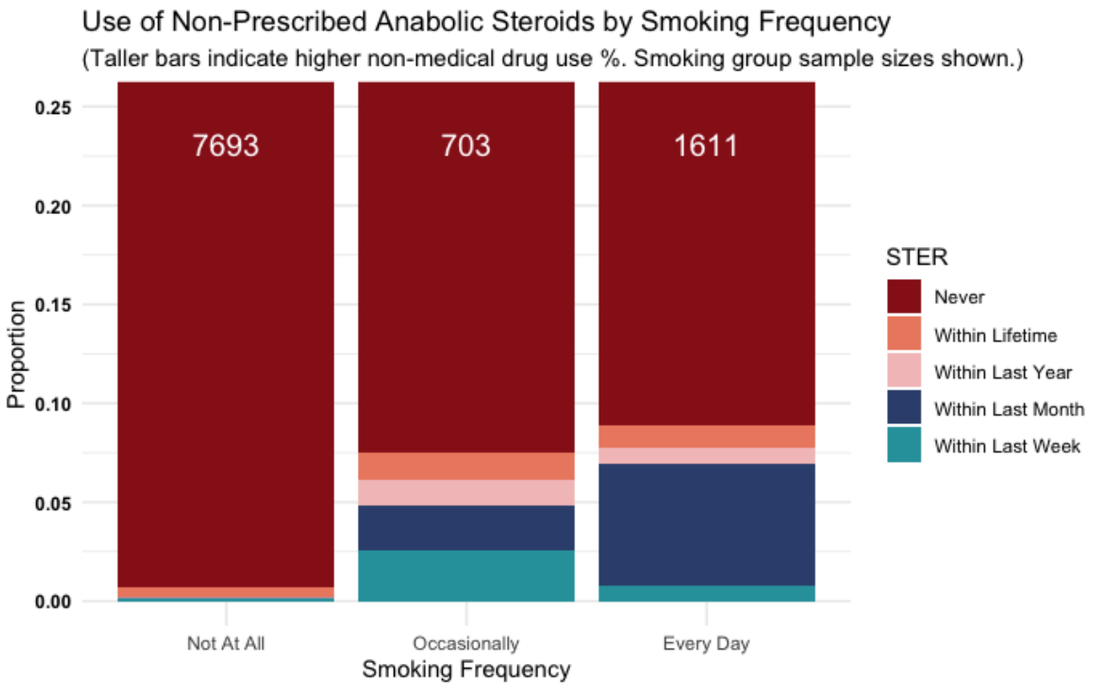


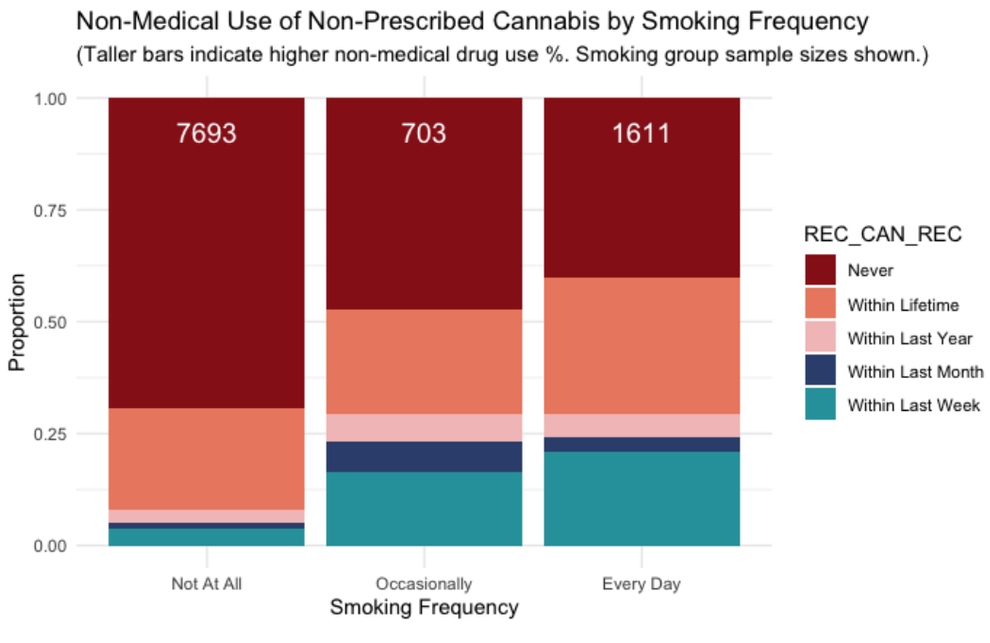


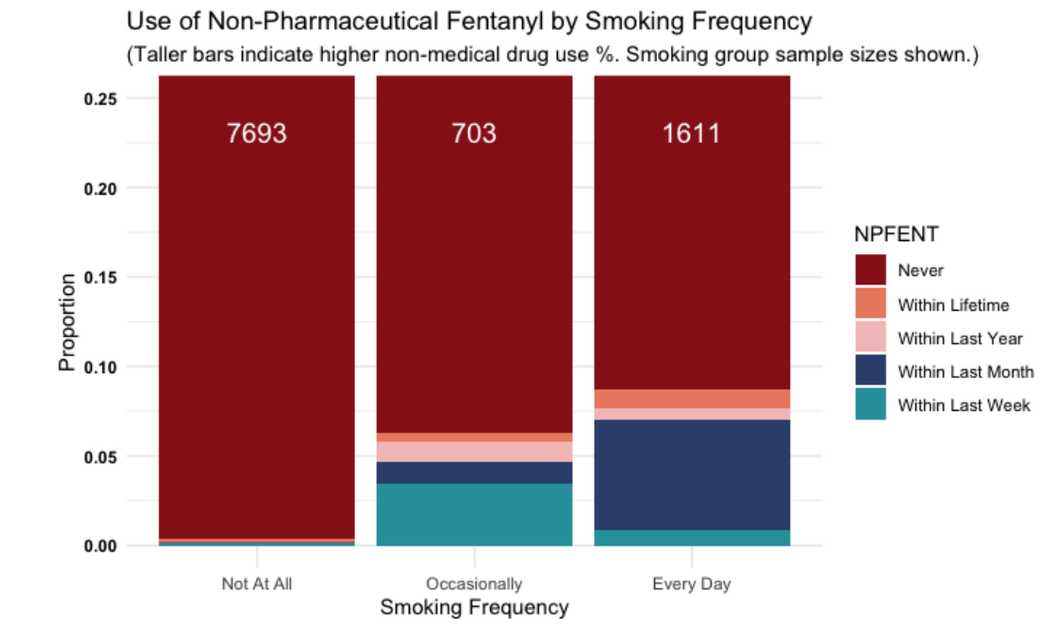


3.3.1.2 Smoking II

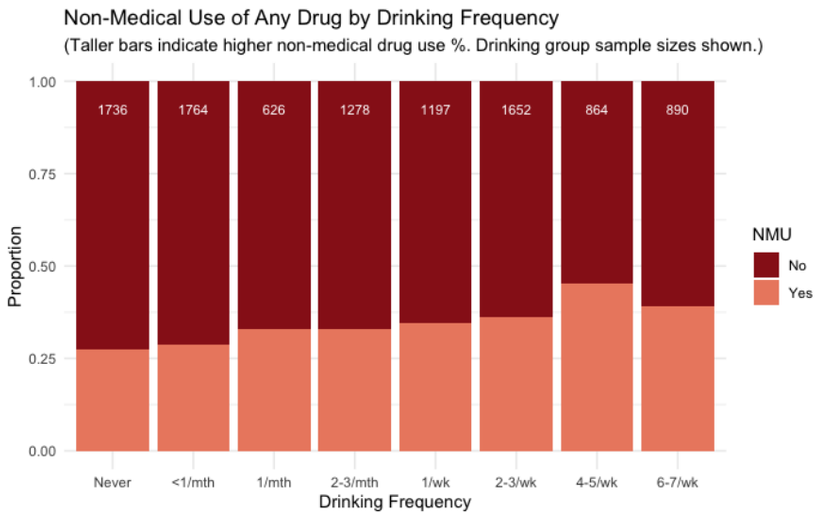


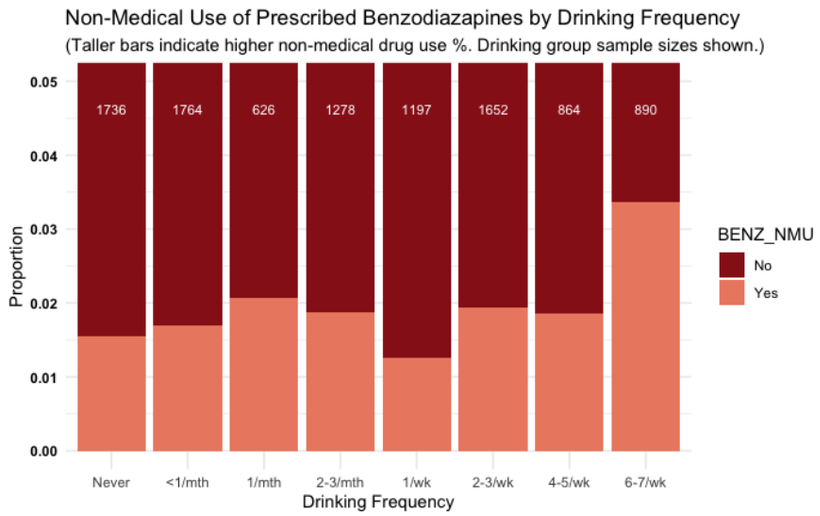


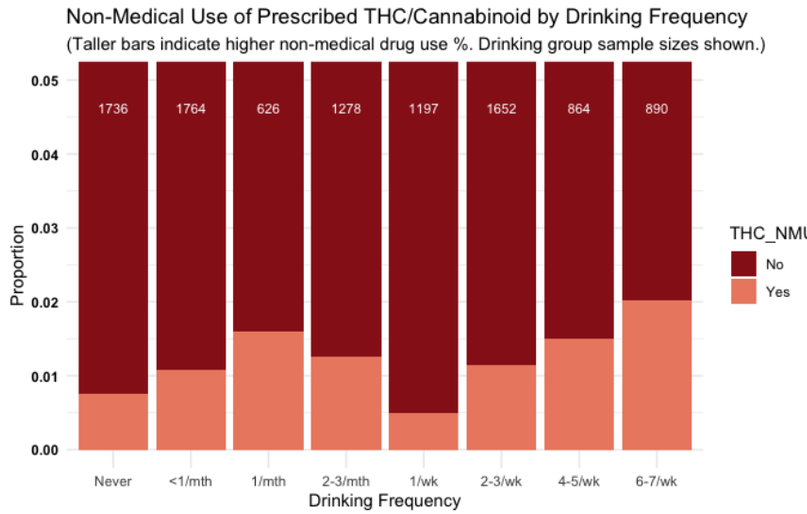


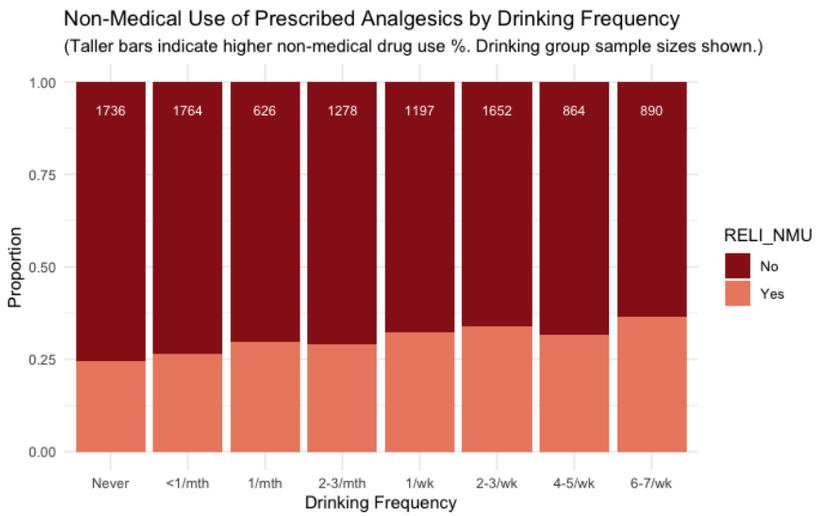


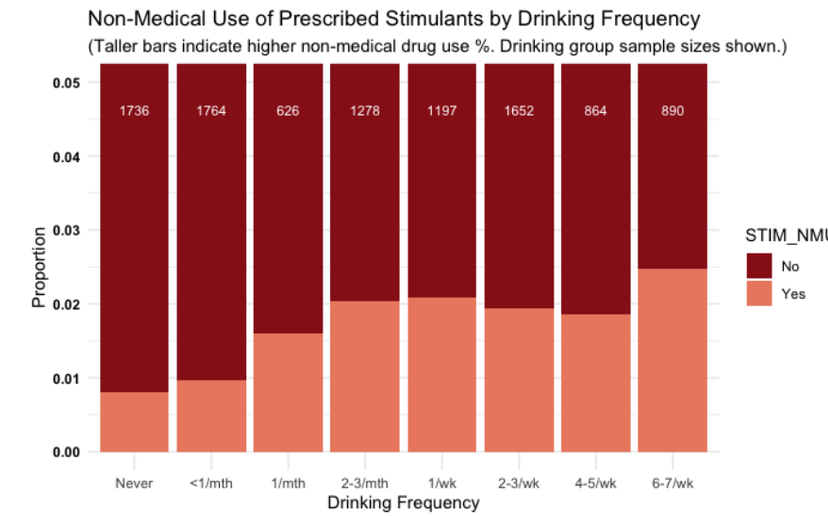
3.3.1.3 Drinking I



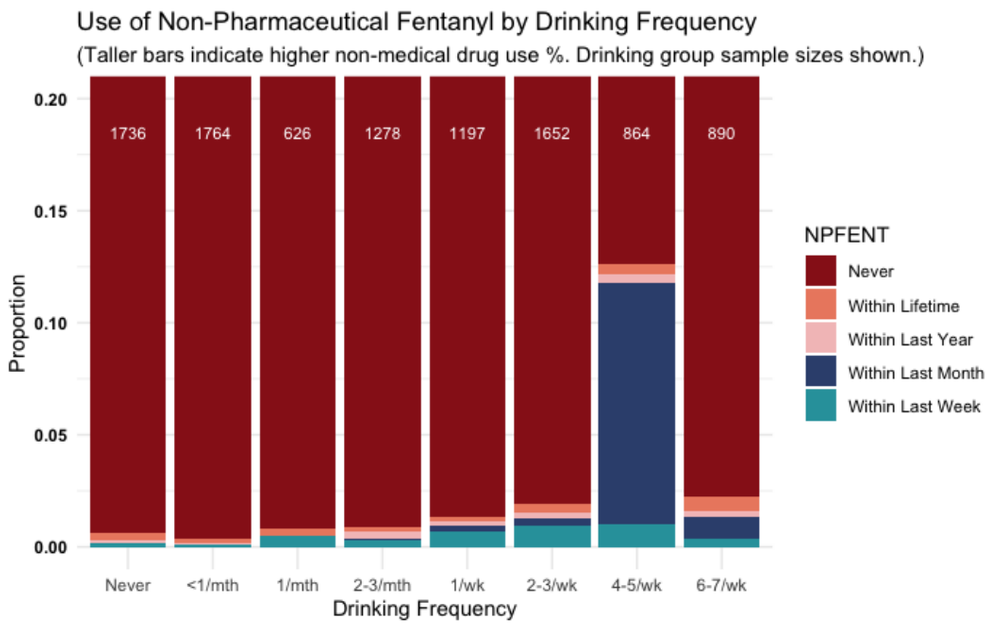


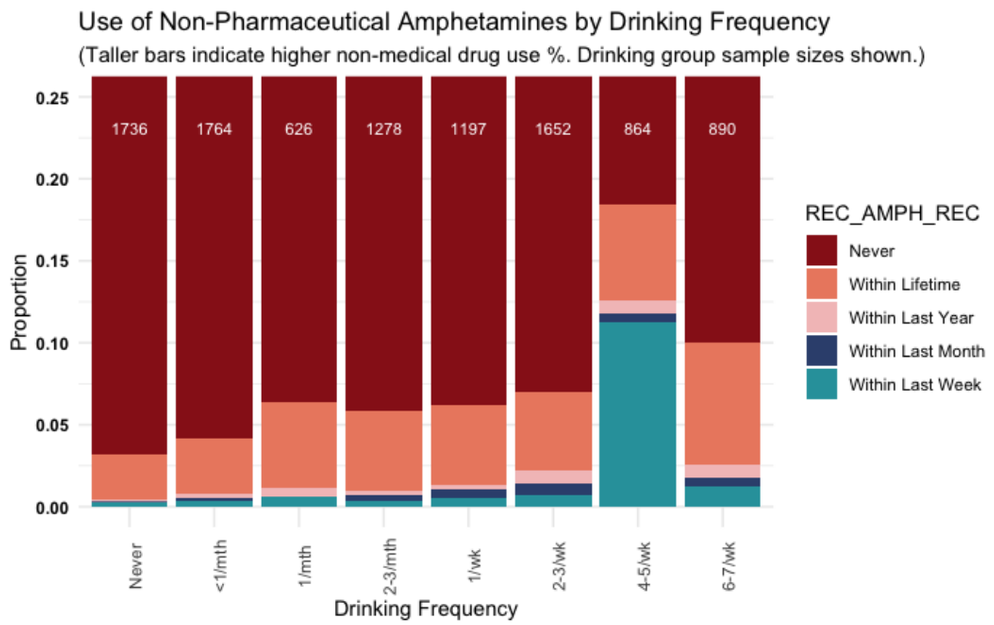


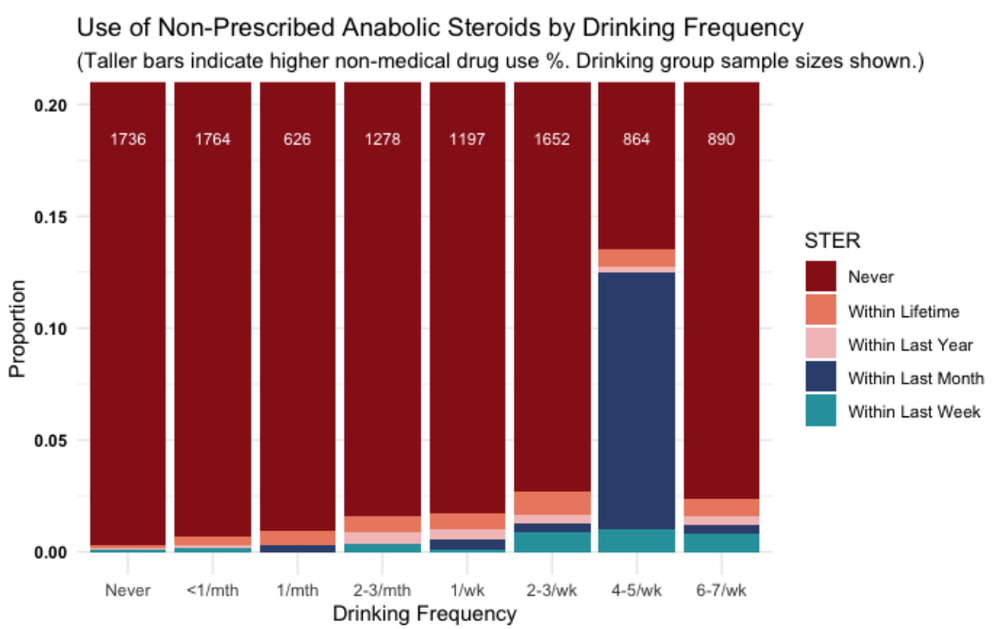


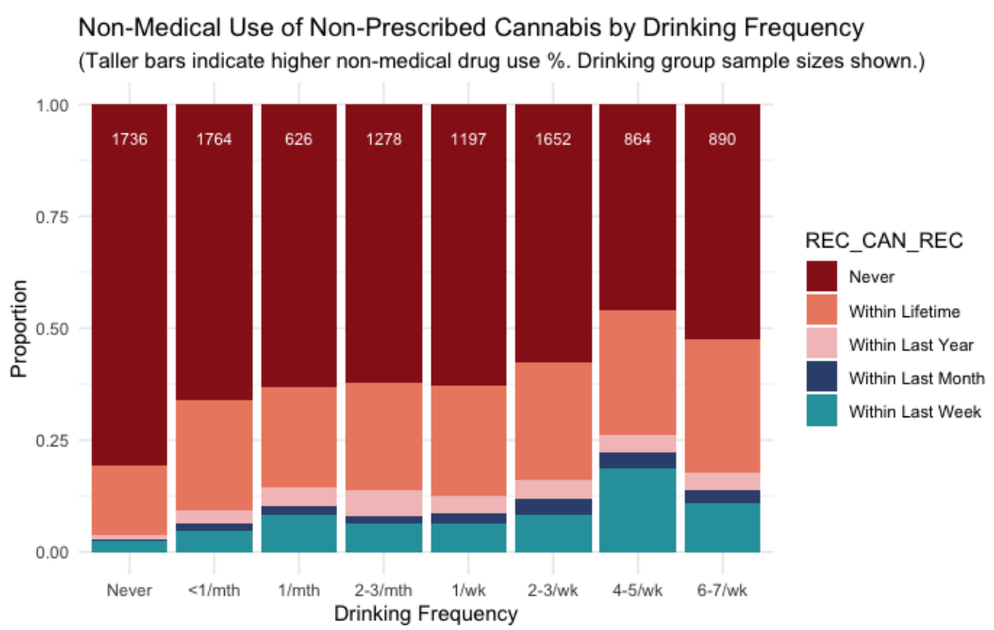


3.3.1.4 Drinking II

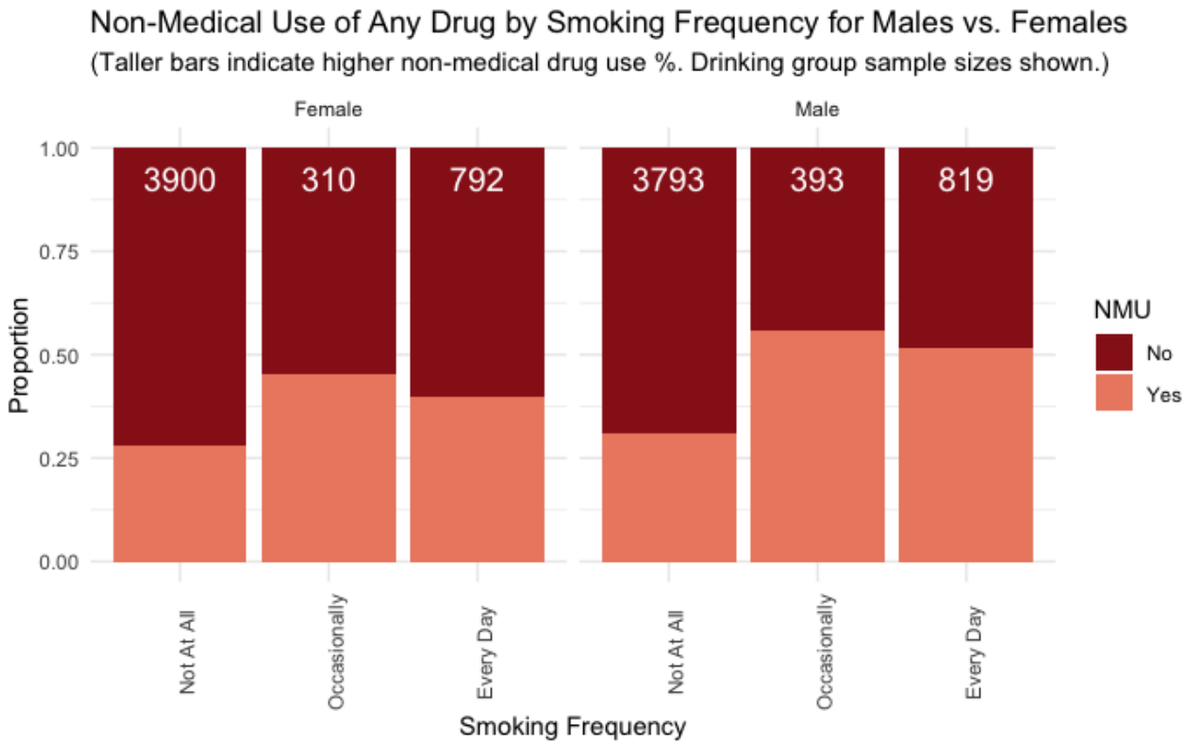


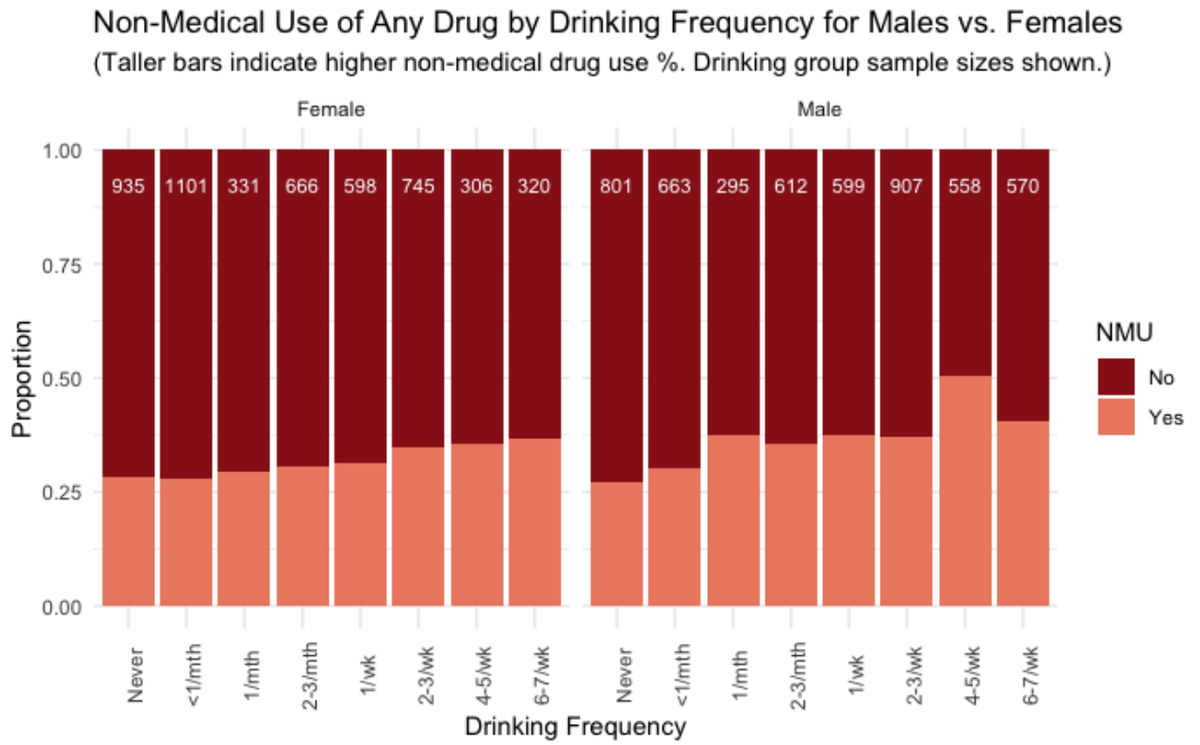




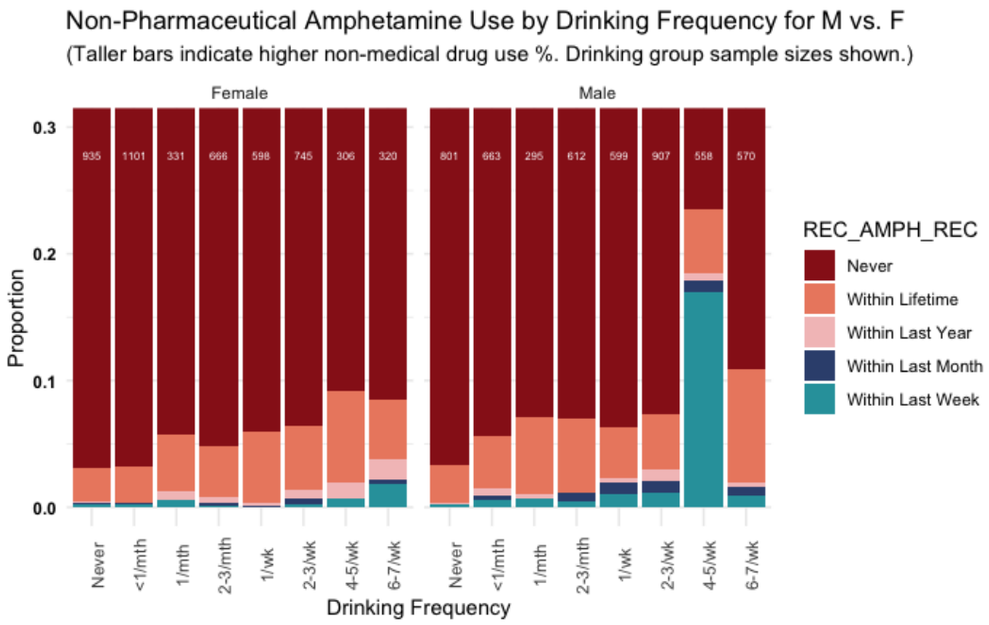


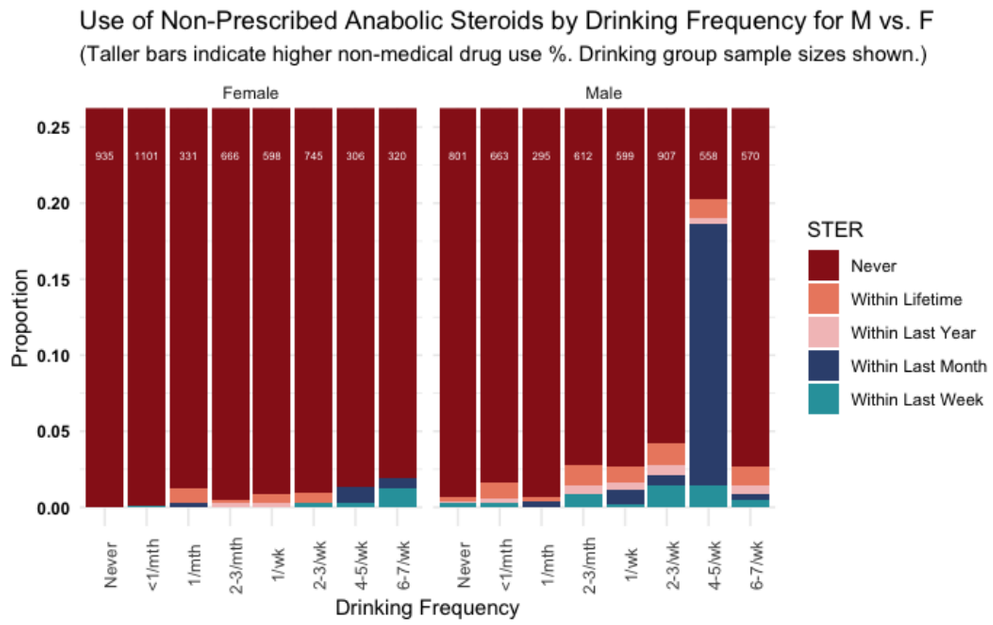
3.3.1.5 Smoking Frequency for Males vs. Females

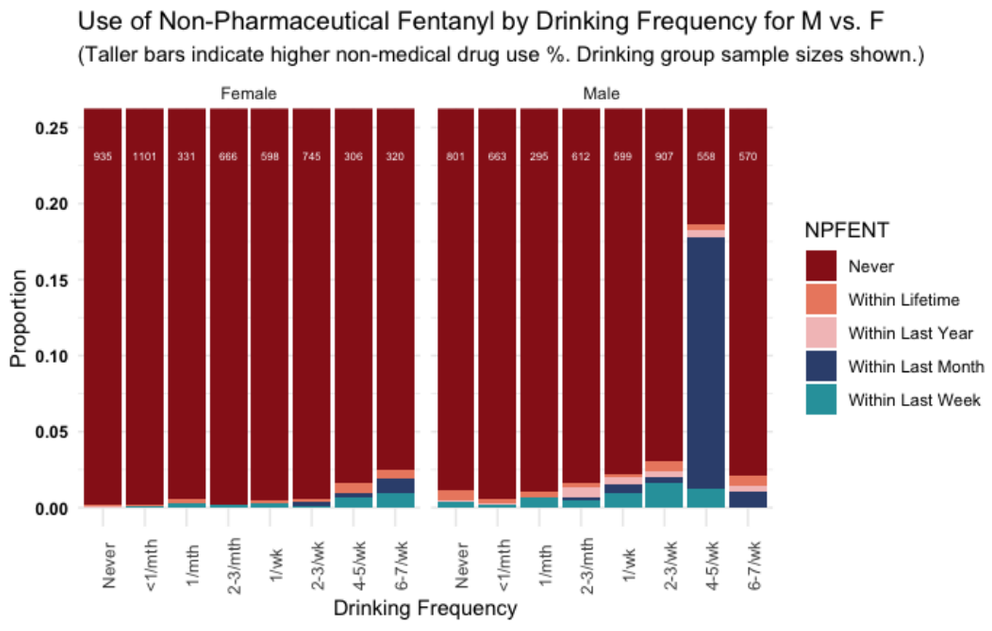




3.3.1.6 Drinking Frequency for Males vs. Females







3.3.1.7 The coefficient table for the model for NMU based on smoking and drinking

	Estimate	Std. Error	z value	Pr(>|z|)
(Intercept)	-1.086	0.055	-19.761	0.000
TOB_FREQ2	0.851	0.080	10.589	0.000
TOB_FREQ1	0.674	0.056	11.942	0.000
ALC_FREQ7	0.017	0.076	0.220	0.826
ALC_FREQ6	0.207	0.102	2.034	0.042
ALC_FREQ5	0.210	0.081	2.591	0.010
ALC_FREQ4	0.253	0.082	3.078	0.002
ALC_FREQ3	0.315	0.075	4.177	0.000
ALC_FREQ2	0.630	0.088	7.120	0.000
ALC_FREQ1	0.420	0.088	4.744	0.000

- The odds of an occasional smoker misusing any drug in their lifetime is 2.34 times the odds of a non-smoker misusing any drug. The odds of a daily smoker misusing any drug in their lifetime is 1.96 times the odds of a non-smoker misusing any drug.

- Similarly, the odds of a person drinking 4-5 times per week misusing any drug in their lifetime is 1.88 times the odds of a non-drinker misusing any drug. The odds of a daily drinker misusing any drug in their lifetime is 1.52 times the odds of a non-drinker misusing any drug.

4. Conclusion

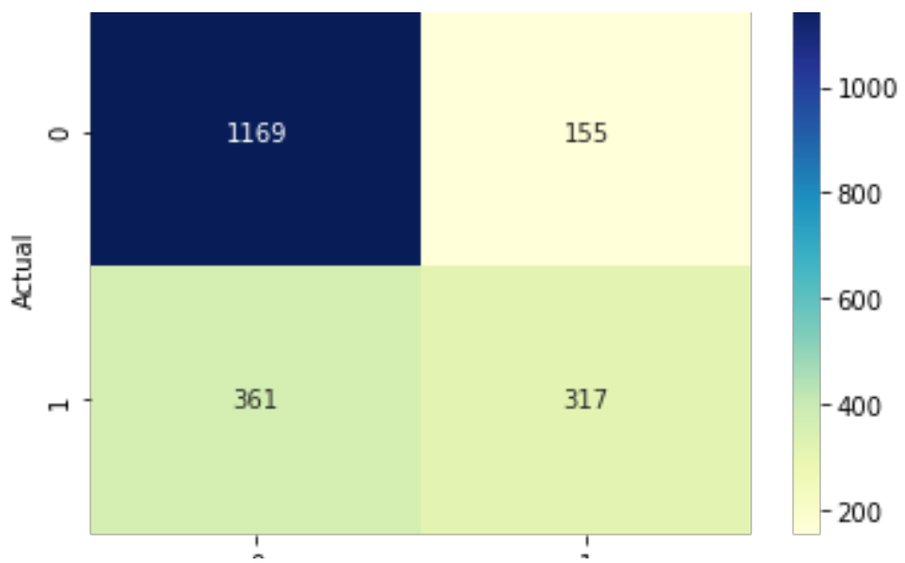
4.1 Key findings:

- **Opioid, Codeine, Cocaine were top three NMU drugs(in order) and Quebec was top NMU social group**

- **Increased frequency of smoking and drinking is associated with an increased rate of drug misuse, as well as more recent use of drugs across all drug groups**

4.2 Recommendations & Limitations

- **Difficulty in building certain models (random forest) due to existence of only binary and categorical variables. Causes bias towards heavier proportions.**
- **No way to know beyond the set of answers. (Ex. Reason for NMU was removed by the data provider)**
- **No other timeline make it hard to make strong claims.**



Confusion matrix of a random forest classifier giving poor result on test set when NMU = 1

Thank you for watching!

Reported by
Jaekang Lee, Alexander MacKenzie, Chang Sun, Tzuying Yu
2021. May. 02

ASA DataFest
University of Toronto