**Illicit Drug Use in a Community-Based Sample of Heterosexually Identified Emerging Adults**

Design Overview

In this study, participants between the ages of 18 to 25 years old and who identified as heterosexual took part in a cross- sectional survey. This survey was administered to 261 participants. The overarching goal of this study design was to look at substance use, health risks, and health promotion among these participants.

Measures

Demographics

For this study, demographic characteristics in this study were gender, race, US born, employment status, and education attainment. The only continuous variable was age and the categorical variables included: gender, race, US born, employment status, and education attainment. Age was measured on ratio scale and the categorical variables, except education attainment, were measured on a nominal scale. Educational attainment was computed on an ordinal scale. Among these variables, all the data was accounted for among the 261 participants.

The participants in this study aged from 18 to 25 years old, and the mean was 21.44 years old (SD= 2.22: Median: 21). Majority of the sample (83. 14%; N=217) of this study were females while 16.6%(N= 44) were males. Approximately half of the participants(45.21%; N=118) identified as White/Caucasian, 21.07% (N=55)marked themselves as Asian/Pacific Islander(API), 14.01%(N=37) were Black/African American, and 19.54%(N=51) identified as Mixed Ethnicity. Majority of the study-population(76.25%;199/261) stated that they were born in the United States and the rest(23.75%; 62/261) were not born in the United States. When it comes to employment status, during the time of this study,162/261(N=162;62.07%) participants stated that they were employed and 99 out of 261 sample(N=99; 37.93%) were unemployed. Among the population, 189 of the participants(72.41%;189/261) either currently or previously attended school, whereas the 72 of the sample (27.59%;72/261) never went to school at all.

Illicit Drug Use

In this survey, the participants were asked about their use of 13 drugs: Alcohol, Tobacco, Ecstasy, Ketamine, Methamphetamine, Marijuana, Cocaine Powder, GHB, Cocaine Crack, Heroin, Mushroom, LSD, and RX. Marijuana(N=172; 65.90%), Alcohol(N=158;60.53%), and Tobacco(N=98; 37.55%) were the top three used drugs. The rest of the drugs were used less among the 261 participants. Ecstasy was used by 19.54%(N=51), Ketamine was used by 5.36% (N=14),Methamphetamine was used by 4.60%(N=12), Cocaine Powder was used by 19.92%(N=52), GHB was used by 2.68%(N=7), by the sample. 3.45%(N=9) used Cocaine Brack, 2.68%(N=7) used Heroin, 19.54%(N=51) used Mushroom, 7.66%(N=20) used LSD, and 22.99% used RX.

Mental Health

In the study, the researchers looked at these mental health factors: therapy, depression, anxiety, and loneliness. Therapy is a categorical variable and was measured using a nominal scale while depression, anxiety, and loneliness were continuous variables, measured on a ratio scale. Although there was no missing data for therapy, depression, and anxiety, there was 48 missing data for loneliness. When asked about therapy experience, 85.44%(N=223; 223/261) of the sample stated that they had no previous experience in therapy, while the rest of the sample 14.56% (N=38,38/261) had previous therapy experience.

Summary

I used a computer programming software called R to analyze the different variables in this study. R studio was used to compute the mean, median, and frequencies are the categorical and continuous variables. The most used drug in this study was Marijuana and most of the population were white and males.

Appendix

The R code

data=read.csv('Data1.csv',header=T)

data=data[,-1]

data[1,2] #to see a specific value located in the X row and Y column, then use data[X,Y]

data[,2]# the second column in the data table

data[2:5,4] # values from the 2nd row to the 5th row at the 4th column

data[2,c(3,5)]#read the 3th and the 5th column value at the 2nd row

data[c(2,4,5),c(1,3,6)]#(2nd row&1st column), (2nd row&3rd column), etc. In total, 9 cells' values

dim(data)#give the dimensions of the data table

dim(data)[1] # give the number of rows

dim(data)[2]#give the number of rolumns

summary(data$RACE) #give specifically the summary about frequencies of difference racces

n\_race=length(data$RACE)#give the number of variables in the dataset related to race, and assign a new name to the data which is n\_race

n\_race

summary(data$RACE)/n\_race #relative frequency (including missing data)

data$RACE[1]=NA#change the first row in the race column to missing data

data$RACE[1]

summary(data$RACE)

n\_race\_noNA=length(na.omit(data$RACE))#to get the number of race without missing numbers by using "na.omit" function.

n\_race\_noNA

summary(data$RACE)/n\_race\_noNA

data$RACE[1]='API' #change the cell value back to the one when it is originally assigned

median(data$AGE)#calculate the median value of the age variable

median(data[,1])#calculate the median value of the first column

mean(data$LONELINESS) #if missing numbers exist, the mean of the variable will be NA

mean(data$LONELINESS, na.rm = TRUE) #calculate the stats summary after deleting the missing values

data=read.csv('Data1.csv',header=T)

table(data$LONELINESS)

range(data$AGE)

range(data$AGE)[2]-range(data$AGE)[1]

var(data$AGE)#variance

sd(data$AGE)#standard deviation

data=read.csv('Data1.csv',header=T)

data=data[,-1]