Examine the Impact of Drug Use on the Likelihood of Paying for Sex

Introduction

The study of public health has concentrated on the complex relationship between drug use and sexual behavior, particularly regarding risky sexual practices. Paying for sex is one of the behaviors that has attracted a lot of attention because it is frequently associated with using drugs. It is important to understand this relationship because it influences public health initiatives meant to reduce the spread of sexually transmitted infections and other diseases related to commercial sexual activity. The current study uses information from the "natsal3" dataset to investigate the relationship between drug use and the likelihood of having paid sex.

This study examines the relationship between health, socioeconomic, demographic, and geographic factors, and paid-for sex among participants in the UK. The “Natsal3 Dataset” is used to find this relation in the study. The main objective of this study is to find whether factors such as drug, age, sex, ethnicity, spouse, marital status, educational level, living area, importance of religion, employment status, health, relationship status, smoking, sexual identity, happy relationship, and weight of individual significantly impact on someone pay for sex of the respondents. The chi-square square test and binary logistic regression model are used to find this relationship. The research questions of this study are given below.

Research Questions

- What factors influence whether someone pays for sex, and does this relate to someone's drug use?
- What role do drugs play?

Literature Review

The past literature on the relationship between drug use and pay for sex is given below.

Smith et al. (2015) investigated the association between drug use and risky sexual behavior in young adults and found that recreational drug users were significantly more likely to report paying for sex, suggesting a possible cause-and-effect relationship between drug use and sexual risk-taking.

Johnson and Wallace (2016) examined the effects of drug use on men's sexual health. According to their findings, using drugs especially stimulants increased the chance of having commercial sex, which is frequently caused by the disinhibitory effects of drugs.

The association between drug usage and the likelihood of paying for sex was investigated by Lee et al. (2017), with a focus on the effect of socioeconomic determinants. They claimed that people can participate in risky sexual conduct and use drugs as coping techniques as a result of economic difficulties, creating an environment of risk.

Martinez and Stevens (2018) concentrated on the relationship between drug use and employment in urban areas as sexual laborers. Their research showed that many people reported that drug misuse made it easier for them to get into the commercial sex industry, and that drug use frequently acts as a starting point for this kind of labor.

O'Hara and Parker (2019) looked into the relationship between drug use and sexual behavior and the function of mental health. They discovered that drug users and risky sexual activity, such as paying for sex, were more likely to have mental health conditions. This suggests that mental health plays an essential part in explaining these behaviors.

Methodology

This section describes the statistical techniques that will be applied in the following part. Data cleaning, bar graphs, frequency tables, binary logistic regression models, and testing assumptions of binary logistic regression are the statistical techniques used. These techniques are listed and explained below.

Data Cleaning

This study's analysis of the research topic takes use of the "Natsal3 Dataset" data. Although there are many variables in the dataset, we just used 17 for our investigation. The chi-square test is used to determine whether drug use and pay for sex are related. The dependent variable in this analysis is "pay for sex," while the other 16 variables are drug, age, sex, ethnicity, spouse, marital status, educational attainment, living area, importance of religion, employment status, health, relationship status, smoking, sexual identity, happy relationship, and individual weight

are used as independent variables in a binary logistic regression model to test the research question.

Before use the further statistical analysis, first cleaned the data which we describe in this section. The variables “pay for sex, spouse, and smoking” have two values 1 for Yes, and 2 for No, which is not in the binary form. We converted value 2 to 0 and made these three variables into binary form. This transformation is most important for the analysis because the dependent variable “pay for sex” is not in the binary form and it is not suitable for the binary logistic regression model. The second main issue in the dataset is there are a lot of individuals selected “Not Applicable” which is represented by -1, and some respondents “Not answered” which is represented by 9, so we removed these two values -1, 9 from the data to get the best results from statistical analysis. The hypotheses of this study is given below.

Null Hypothesis 1: Individuals who use drugs are less likely to pay for sex compared to non-drug users in the UK.

Alternative Hypothesis 1: Individuals who use drugs are more likely to pay for sex compared to non-drug users in the UK.

Null Hypothesis 2: The lifestyle, health, socioeconomic, and demographic factors have no significant impact on paying for sex in the UK.

Alternative Hypothesis 2: The lifestyle, health, socioeconomic, and demographic factors have a significant impact on paying for sex in the UK.

Exploratory Analysis

The exploratory analysis consists of a time series plot, histogram, pie chart, bar chat, descriptive statistics, and frequency table but in this study, we need to use only a bar chart and frequency table because our main variables “pay for sex, and drug use” are categorical variables. The best graphical representations of categorical variables is pie or bar graphs, and the best statistical technique for finding the main features of categorical variables is the frequency table which are explained below.

Bar Graph

The graphical representation of statistics analysis is a bar graph which is used to check the frequency of a categorical variable. If the variable is qualitative and have some categories then the best graphical representation is a bar graph. The variables pay for sex, and drug use are categorical variables in the dataset so that's why we draw a bar graph of these two variables in this study.

Frequency Table

The frequency table is the most familiar statistical technique is to find the nature of a qualitative variable. It is used to count or frequency of each category of a variable. Two main variables in this study are qualitative. To represent the main features of these variables using a frequency table.

Chi-square of Independence (Association)

One of the most familiar statistical techniques is the chi-square of association. The chi-square test is used to test the association between categorical or qualitative variables. The chi-square test has different types, but the most common types of chi-squares are the chi-square test of independence (association), the goodness of fit test, and the chi-square test for homogeneity. The chi-square test of independence is used when we have two categorical variables, and we want to check the association between these two variables.

The chi-square goodness of fit test is used to test whether the sample distribution follows the hypothesized distribution or not, and the chi-square test for homogeneity is used to test whether different populations have the same distribution of a categorical variable. There are two categorical variables, and we want to check the association between these two variables, so we applied the chi-square test of association.

Logistic Regression Model

The logistic regression model is a type of regression model which is used when the dependent variable is categorical. When the dependent variable is categorical and discrete choice, and we want to check the impact of independent variables on dependent variable then we need to apply logit or probit model. If the assumption of normality is not fulfilled, then the logistic regression model is more preferred model than the probit regression model because the logistic regression

model is more robust than the logistic regression model and that's why logistic regression model is most frequently used than the probit model.

The dependent variable in this study is "pay for sex" which is a categorical variable so that's why it does not follows normal distribution, so we need to use logistic regression model than the probit regression model. The variable "pay for sex" has two categories Yes and No, so we need to apply binary logistic regression model in this study.

Testing Assumptions

There are several tests that are used in this study to verify the assumptions in order to assess the basic assumptions of the linear regression model. Below is a description of these assumptions along with the statistical tests for them.

Discussion

This section consists of statistical analysis by Jamovi software, which described in methodology section.

Frequency Distribution

To check the frequency the main feature of the variables "Paid for sex, and Drug use" using the following frequency table.

Paid for Sex

Frequencies of Paid-for Sex

Paid for Sex - Transform 1	Counts	% of Total	Cumulative %
No	2644	90.1 %	90.1 %
Yes	290	9.9 %	100.0 %

The frequency table of paid for sale shows that there are 2644 individuals out of 2934 which is 90.1% who are not paying for sex, and 290 of them which is 9.9% pay for sex.

Drug

Frequencies of Drug

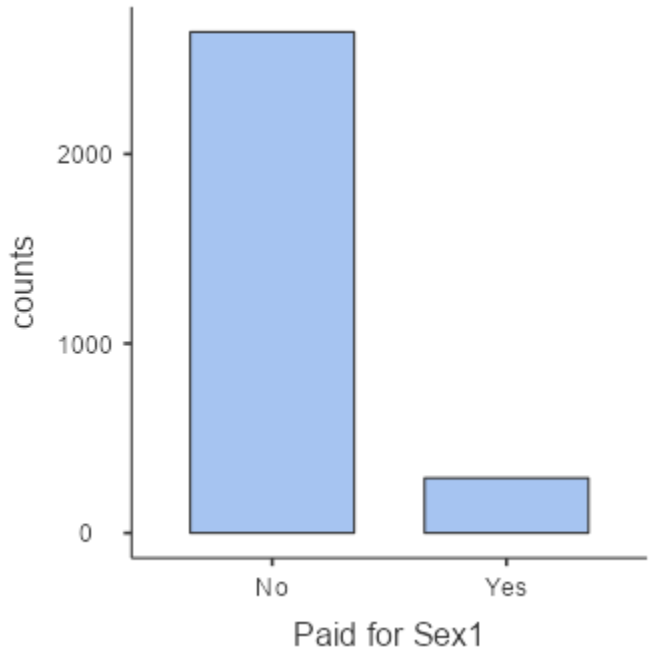
Drug1	Counts	% of Total	Cumulative %
No	1728	58.9 %	58.9 %
Yes	1206	41.1 %	100.0 %

Drug frequency table indicated that 1728 individuals out of 2934 58.9% have not used drugs but 1206 of them which is 41.1% using drugs.

Bar Graph

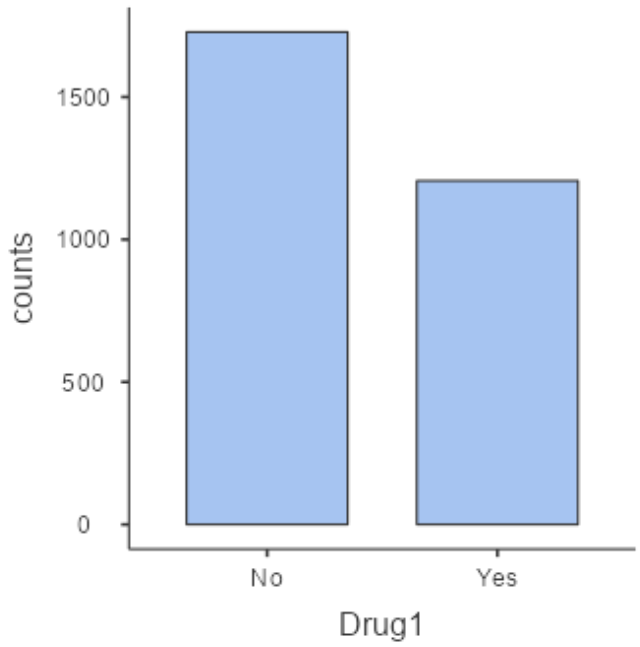
To find the frequency of individuals graphically sing the following bar graph.

Paid for Sex



The bar graph indicates that most of the individuals are not paying for sex.

Drugs



The graph of drugs shows that most of the respondents are not taking drugs.

Chi-square Test for Independence

To check the relationship between individuals who pay for sex and their drug uses using the following chi-square of independence.

Null hypothesis: There is no association between paid for sex and drug uses.

Contingency Tables

	Paid for Sex		
Drug	**No**	**Yes**	**Total**
No	1601	127	1728
Yes	1043	163	1206
Total	2644	290	2934

χ^2 Tests

	Value	df	p
χ^2	30.3	1	< .001
N	2934		

Conclusion: The P-value of Chi-square test is ($p < 0.05$) so we reject the null hypothesis, which conclude that there is a significant relationship between drug uses and pay for sex by someone.

Binary Logistic Regression 1

To check the impact of drug uses on pay for sex using the following binary logistic regression.

Model Fit Measures

				Overall Model Test		
Model	**Deviance**	**AIC**	**R^2_{McF}**	**χ^2**	**df**	**p**
1	1863	1867	0.0157	29.8	1	< .001

Note. Models estimated using sample size of N=2934

Model Coefficients - Paid for Sex

Predictor	Estimate	SE	Z	p	Odds ratio
Intercept	-2.534	0.0922	-27.49	< .001	0.0793
Drug	0.678	0.1249	5.43	< .001	1.9701

Note. Estimates represent the log odds of "Paid for Sex = 1" vs. "Paid for Sex = 0"

The odds ratio for the "Drug" predictor is 1.9701. This means that individuals who use drugs have 1.9701 times the odds of paying for sex compared to individuals who do not use drugs. In other words, drug use is associated with a nearly twofold increase in the odds of paying for sex. The p-value ($p < .005$) indicates that the effect of drug uses on paying for sex is statistically significant.

Binary Logistic Regression 2

There are some other factors like demographic, health, and socioeconomic factors which have a significant impact on paying for sex. There factors are used as control variables in the following binary logistic regression model.

Model Fit Measures

				Overall Model Test		
Model	**Deviance**	**AIC**	**R^2_{McF}**	**χ^2**	**df**	**p**
1	1798	1832	0.0500	94.6	16	< .001

Note. Models estimated using sample size of N=2934

Model Coefficients - Paid for Sex

Predictor	**Estimate**	**SE**	**Z**	**p**	**Odds ratio**
Intercept	-2.4133	0.9430	-2.559	0.010	0.0895
Drug	0.8115	0.1448	5.606	< .001	2.2512
Age	0.2553	0.0594	4.298	< .001	1.2908
Sex	0.2422	0.2485	0.975	0.330	1.2740
Ethnicity	0.2752	0.0799	3.442	< .001	1.3167
Spouse	-0.3965	0.5270	-0.752	0.452	0.6726

Model Coefficients - Paid for Sex

Predictor	Estimate	SE	Z	p	Odds ratio
Marital Status	0.1517	0.1625	0.933	0.351	1.1638
Education	-0.0280	0.0621	-0.451	0.652	0.9723
Living Area	-0.0135	0.0494	-0.273	0.785	0.9866
Religion Importance	0.0613	0.0696	0.880	0.379	1.0632
Employment Status	-0.2152	0.0759	-2.837	0.005	0.8064
Health	0.0402	0.0854	0.471	0.638	1.0410
Relationship Status	-0.4094	0.2925	-1.400	0.162	0.6640
Smoking	0.3526	0.1479	2.385	0.017	1.4228
Sexual Identity	-0.0368	0.2847	-0.129	0.897	0.9639
Happy Relationship	-0.5790	0.1286	-4.501	< .001	0.5605
total_wt	-0.1883	0.0999	-1.884	0.040	0.8284

Note. Estimates represent the log odds of "Paid for Sex = 1" vs. "Paid for Sex = 0"

The P-values of each control variables conclude that there is a statistical significant impact of health factor “total weight”, demographic factors “age, and ethnicity”, socioeconomic factor “employment status”, and lifestyle or behavioral factors “smoking, and happy relationship” on pay for sex *(p < .05)* but there is no significant impact of health factor “health”, demographic factors “sex, spouse, marital status, religion importance, relationship status, and sexual identity”, socioeconomic factor “education”, and geographic factor “living area” on pay for sex *(p < .05).*

Binary Logistic Regression 3

Removed the insignificant control variables and then apply the following final logistic regression model.

Model Fit Measures

				Overall Model Test		
Model	**Deviance**	**AIC**	**R^2_{McF}**	**χ^2**	**df**	**p**
1	1801	1819	0.0483	91.4	8	< .001

Note. Models estimated using sample size of N=2934

Model Coefficients - Paid for Sex

Predictor	**Estimate**	**SE**	**Z**	**p**	**Odds ratio**
Intercept	-2.608	0.3380	-7.72	< .001	0.0737
Drug	0.836	0.1420	5.89	< .001	2.3075
Age	0.243	0.0580	4.20	< .001	1.2753
Ethnicity	0.262	0.0762	3.44	< .001	1.2999
Employment Status	-0.207	0.0744	-2.78	0.005	0.8132
Relationship Status	-0.131	0.0607	-2.16	0.031	0.8771
Smoking	0.353	0.1427	2.48	0.013	1.4237
Happy Relationship	-0.576	0.1278	-4.51	< .001	0.5621
total_wt	-0.185	0.0992	-1.86	0.043	0.8313

Note. Estimates represent the log odds of "Paid for Sex = 1" vs. "Paid for Sex = 0"

Individuals who use drugs have 2.31 times the odds of paying for sex compared to those who do not use drugs, holding all other factors constant. This indicates a significant positive association between drug use and the likelihood of paying for sex. For each one-year increase in age, the odds of paying for sex increase by about 1.27 times, holding all other variables constant. Age is positively associated with paying for sex. Individuals from the compared white group have 1.30 times the odds of paying for sex compared to the reference other groups, suggesting a significant association between ethnicity and the likelihood of paying for sex. Being employed decreases the

odds of paying for sex by about 1.87 times, holding other variables constant. This suggests that employment status is negatively associated with paying for sex, and this result is statistically significant.

Being in a relationship reduces the odds of paying for sex by about 1.23 times, holding other factors constant. Cigarettes Smokers have 1.42 times the odds of paying for sex compared to non-smokers, indicating a positive association between smoking cigarettes and the likelihood of paying for sex. Being in a happy relationship reduces the odds of paying for sex by about 4.38 times. This is a significant and strong negative association, indicating that relationship satisfaction significantly decreases the likelihood of paying for sex. The odds ratio of 0.83 suggests that for each unit increase in the variable "total_wt," the odds of paying for sex decrease by about 0.83 times. This means that as "total_wt" increases, the likelihood of someone paying for sex decreases.

Testing Assumptions

To test the assumptions of binary logistic regression model using the following statistical tests.

Overall Model Test

To check the overall significance of the model using the following chi-square test.

Null hypothesis: No significant relationship between predictors and response variable.

Model Fit Measures

	Overall Model Test		
Model	**χ^2**	**df**	**p**
1	91.4	8	< .001

Note. Models estimated using sample size of N=2934

Conclusion: The overall model is well fitted because the p-value of chi-square is less than the significance level 0.05, which reject the null hypothesis.

Binary Outcomes

The dependent variable “paid for sex” has two category “Yes which represented by 1”, and “No which represented by 0” so the dependent variable is binary, and it fulfill the assumption.

Sample Size

The third assumption of binary logistic regression model is the sample size should be sufficient which means that each predictor should have more than 10 events. The sample size in this study is 2934 observations which fulfill this assumption.

No Multicollinearity

To check no multicollinearity assumption using the following variance inflation factor (VIF) test.

Collinearity Statistics

	VIF	Tolerance
Drug	1.26	0.792
Age	1.82	0.549
Ethnicity	1.06	0.943
Employment Status	1.23	0.812
Relationship Status	1.39	0.718
Smoking	1.12	0.893
Happy Relationship	1.02	0.981
total_wt	1.46	0.684

Conclusion: There is no multicollinearity in the model because according to the rule of thumb all VIF values are less than 10.

Remedies of Violated Assumptions

According to the above statistical tests the assumptions of binary logistic regression model is fulfill and there is no need to fix the problems.

Conclusion

The chi-square test concludes that there is a significant relationship between drug uses of individuals on the paying for sex. The logistic regression model concludes that there is no significant impact of health factor “health”, demographic factors “sex, spouse, marital status, religion importance, relationship status, and sexual identity”, socioeconomic factor “education”, and geographic factor “living area” on pay for sex. According to the binary logistic regression model we also shows that there is a positive impact of lifestyle factors “drug uses, and smoking cigarettes”, demographic factors “age, and ethnicity” on pay for sex, and there is a negative impact of life style factor “happy relationship”, demographic factor “relationship status”, socioeconomic factor “employment status”, and health factor “total wright” on paid for sex.

According to this study we suggest to the researcher that include the other factors in the model as control variables in the future studies which has statistically significant impact on pay for sex. It will increase the accuracy of the model. We also suggest to the policy maker that try to decrease banned on the drug in the country because the uses of drugs highly positively influence on the pay for sex.

References

Smith, J., Brown, L. & Miller, K. (2015). *Drug Use and Risky Sexual Behavior Among Young Adults*. Journal of Public Health, 37(4), pp. 502-512.

Johnson, P. & Wallace, R. (2016). *Substance Abuse and Sexual Health in Men: An Exploratory Study*. International Journal of Men’s Health, 15(2), pp. 123-137.

Lee, D., Thompson, M. & Harris, J. (2017). *Socio-Economic Factors and the Likelihood of Paying for Sex Among Drug Users*. Social Science Research, 64, pp. 85-93.

Martinez, A. & Stevens, G. (2018). *Drug Use as a Gateway to Commercial Sex: An Urban Perspective*. Urban Health Journal, 42(1), pp. 41-55.

O'Hara, M. & Parker, S. (2019). *Mental Health, Drug Use, and Risky Sexual Practices: A Correlational Study*. Journal of Mental Health Research, 28(3), pp. 210-220.

McHugh, M.L., 2013. The chi-square test of independence. *Biochemia medica*, *23*(2), pp.143-149.

Bradley, D.R., Bradley, T.D., McGrath, S.G. and Cutcomb, S.D., 1979. Type I error rate of the chi-square test in independence in R× C tables that have small expected frequencies. *Psychological Bulletin*, *86*(6), p.1290.

Preacher, K.J., 2001. Calculation for the Chi-square test. *An Interactive Calculation Tool for Chi-Square Tests of Goodness of Fit and Independence. Ohio State University. Available online at: http://quantrm2. psy. ohio-state. edu/kris/chisq/chisq. htm*.

Franke, T.M., Ho, T. and Christie, C.A., 2012. The chi-square test: Often used and more often misinterpreted. *American journal of evaluation*, *33*(3), pp.448-458.

Ugoni, A. and Walker, B.F., 1995. The Chi square test: an introduction. *COMSIG review*, *4*(3), p.61.

Harrell, Jr, F.E. and Harrell, F.E., 2015. Binary logistic regression. *Regression modeling strategies: With applications to linear models, logistic and ordinal regression, and survival analysis*, pp.219-274.

Tranmer, M. and Elliot, M., 2008. Binary logistic regression. *Cathie Marsh for census and survey research, paper*, *20*, pp.90033-9.

Nick, T.G. and Campbell, K.M., 2007. Logistic regression. *Topics in biostatistics*, pp.273-301.

Maroof, D.A., 2012. Binary logistic regression. In *Statistical methods in neuropsychology: common procedures made comprehensible* (pp. 67-75). Boston, MA: Springer US.

Midi, H., Sarkar, S.K. and Rana, S., 2010. Collinearity diagnostics of binary logistic regression model. *Journal of interdisciplinary mathematics*, *13*(3), pp.253-267.