Table 1 : Description of Features:

Features		Description
Database Info	ID	Number of the record in the original database
Personal Information	Age	Age of the participant
	Gender	Male or Female
	Education	Education level of the participant
	Country	Country of origin of the participant
	Ethnicity	Ethnicity of the participant
Psychology	Nscore	Neuroticism [NEO-FFI-R]
	Escore	Extraversion [NEO-FFI-R]
	Oscore	Openness to experience [NEO-FFI-R]
	Ascore	Agreeableness [NEO-FFI-R]
	Cscore	Conscientiousness [NEO-FFI-R]
	Impulsive	Impulsiveness [BIS-11]
	SS	Sensation Seeking [ImpSS]
Drugs	Alcohol	Consumption of Alcohol
	Amphet	Consumption of Amphet
	Amyl	Consumption of Amyl
	Benzos	Consumption of Benzos
	Caff	Consumption of Caffeine
	Cannabis	Consumption of Cannabis
	Choc	Consumption of Chocolate
	Coke	Consumption of Cocaine
	Crack	Consumption of Crack
	Ecstasy	Consumption of Ecstasy
	Heroin	Consumption of Heroin
	Ketamine	Consumption of Ketamine
	Legalh	Consumption of Legalh
	LSD	Consumption of LSD
	Meth	Consumption of Meth
	Mushroom	Consumption of Mushroom

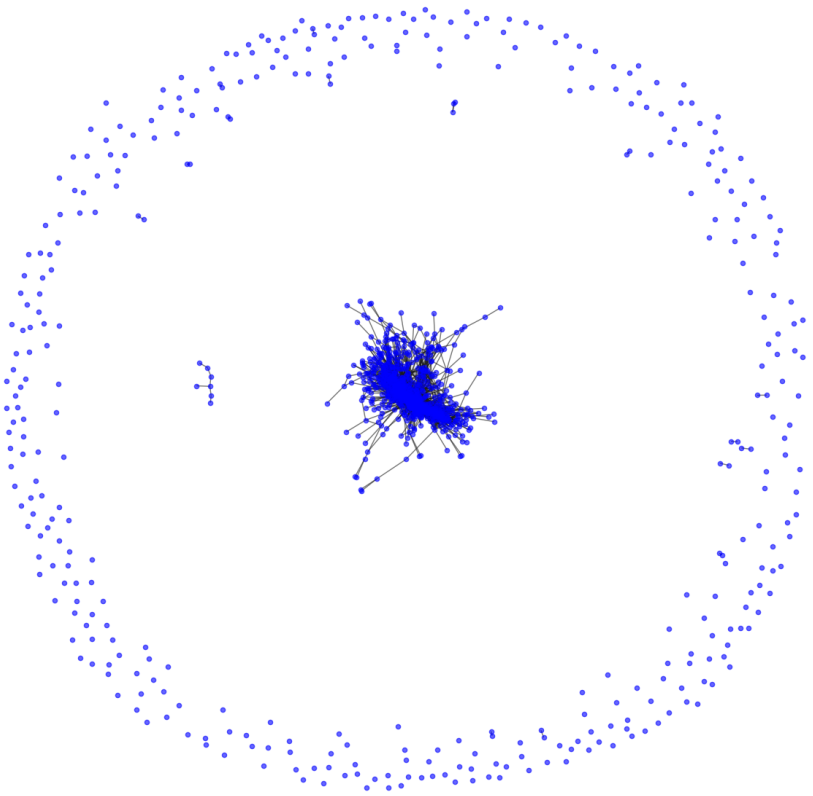
	Nicotine	Consumption of Nicotine
	VSA	Abuse of volatile substance consumption
	Semer	Consumption of Semer [Fictional Drug, to detect over-claimants]

Table 2 : Encoded Data Header (All numerical values)

	Age	Gender	Education	Country	Ethnicity	Nscore	Escore	Oscore	AScore	Cscore	...	Crack	Ecstasy	Heroin	Ketamine	Legalh	LSD	Meth	Mushrooms	Nicotine	VSA
0	1	1	8	5	6	-0.67825	1.93886	1.43533	0.76096	-0.14277	...	0	4	0	2	0	2	3	0	4	0
1	2	1	5	5	6	-0.46725	0.80523	-0.84732	-1.62090	-1.01450	...	0	0	0	0	0	0	0	1	0	0
2	0	0	7	5	6	-0.14882	-0.80615	-0.01928	0.59042	0.58489	...	0	0	0	2	0	0	0	0	2	0
3	2	0	8	5	6	0.73545	-1.63340	-0.45174	-0.30172	1.30612	...	0	1	0	0	1	0	0	2	2	0
4	5	0	3	1	6	-0.67825	-0.30033	-1.55521	2.03972	1.63088	...	0	0	0	0	0	0	0	0	6	0

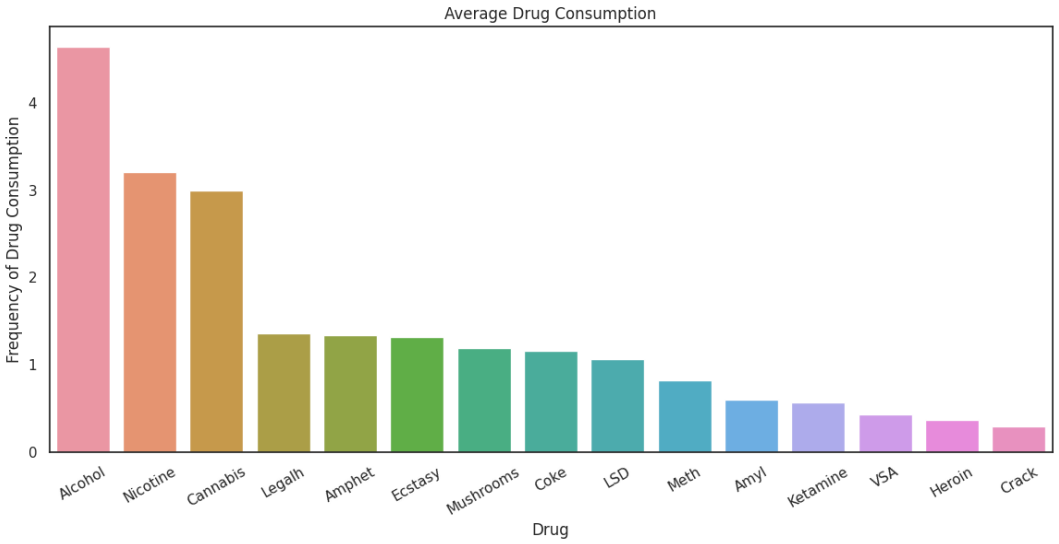
5 rows × 27 columns

a. Network Analysis of Drug Consumption Habits:

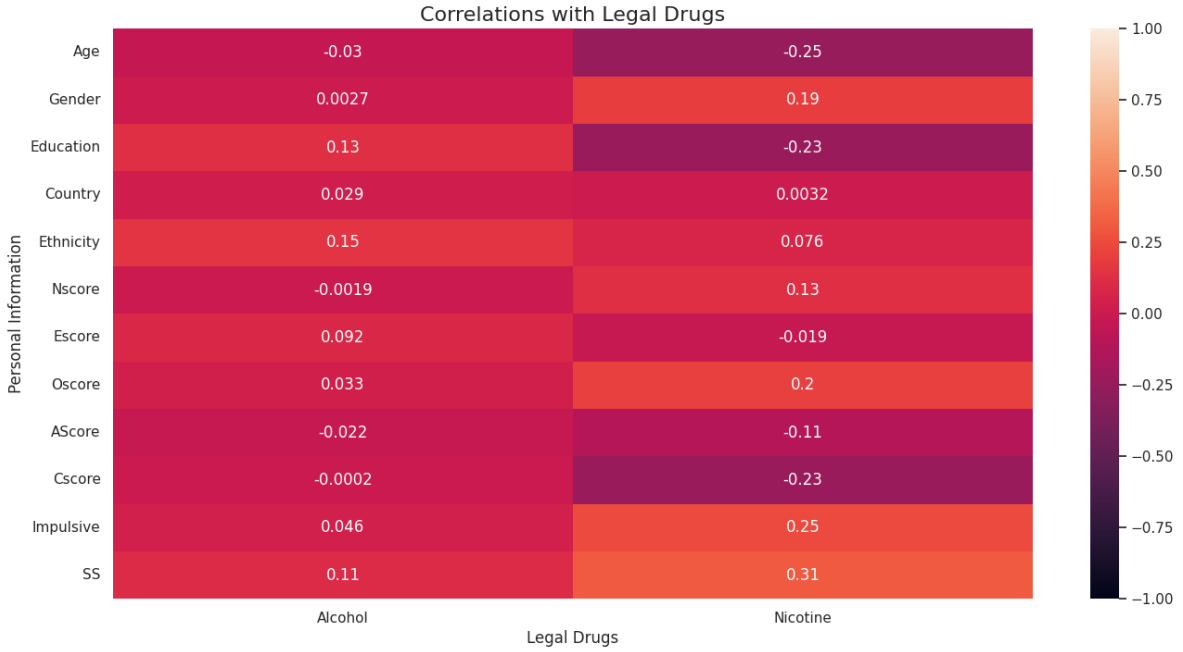


Graph 1 : Drug Consumption Similarity Network

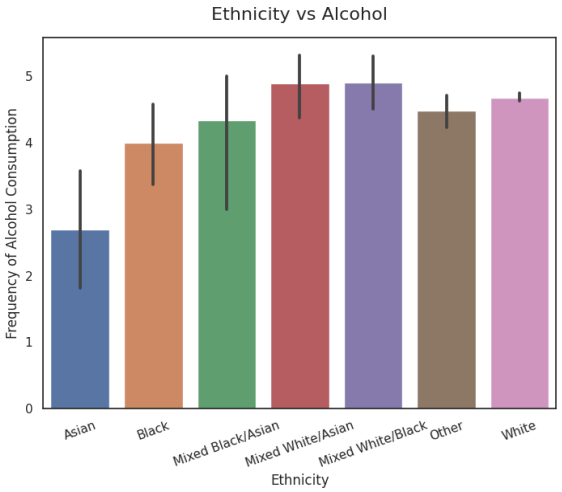
you can also see the visualization using “Gephi”: open the file”Drug_Graph.gephi”



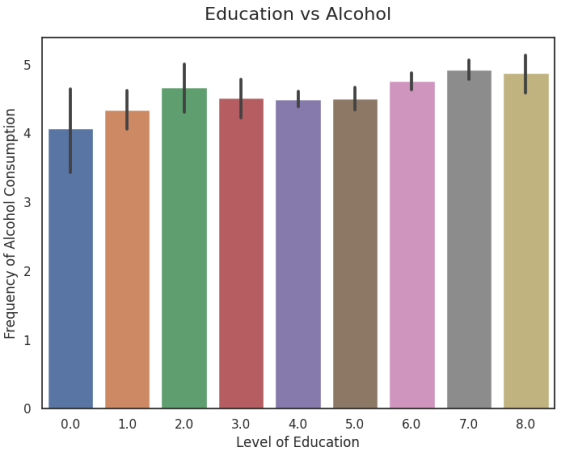
Graph 2 : Average Drug Consumption



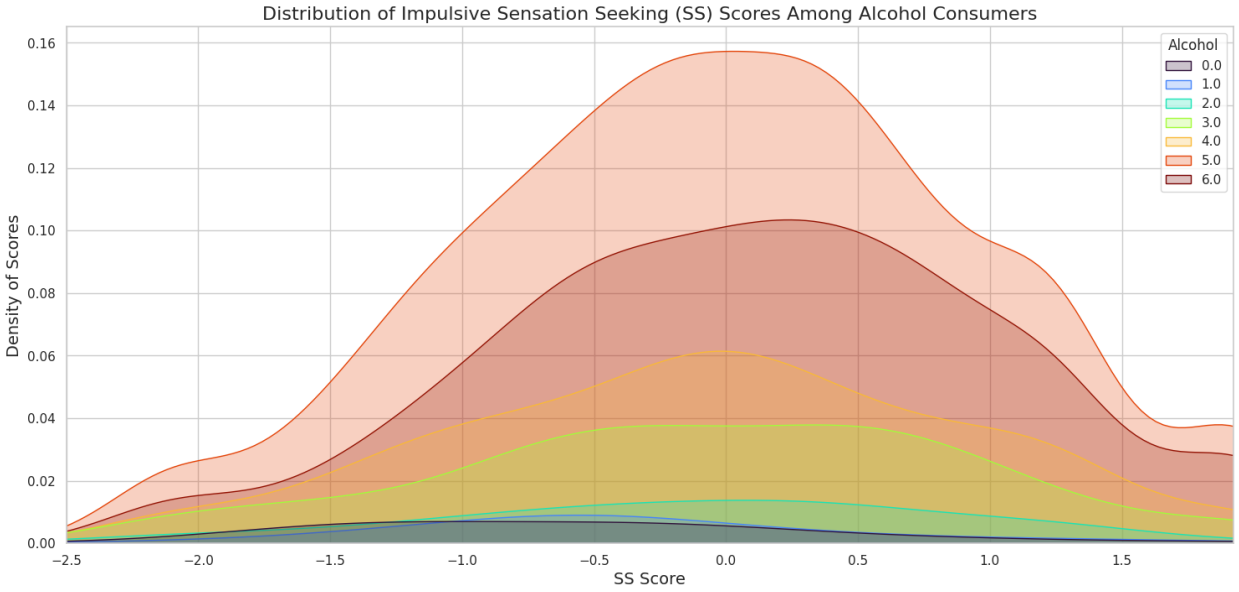
Graph 3 : Correlations with Legal Drugs



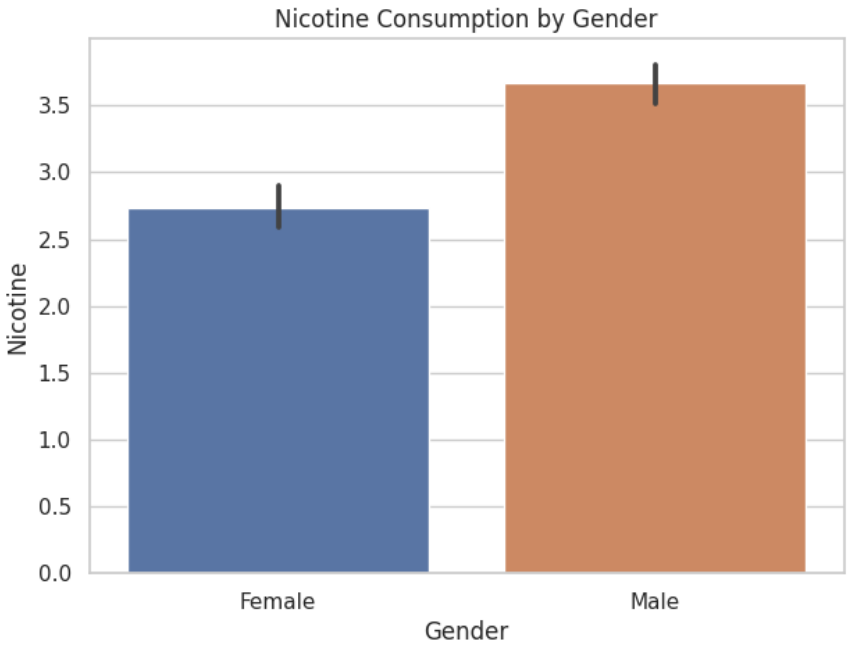
Graph 4 : Alcohol Consumption by Ethnicity



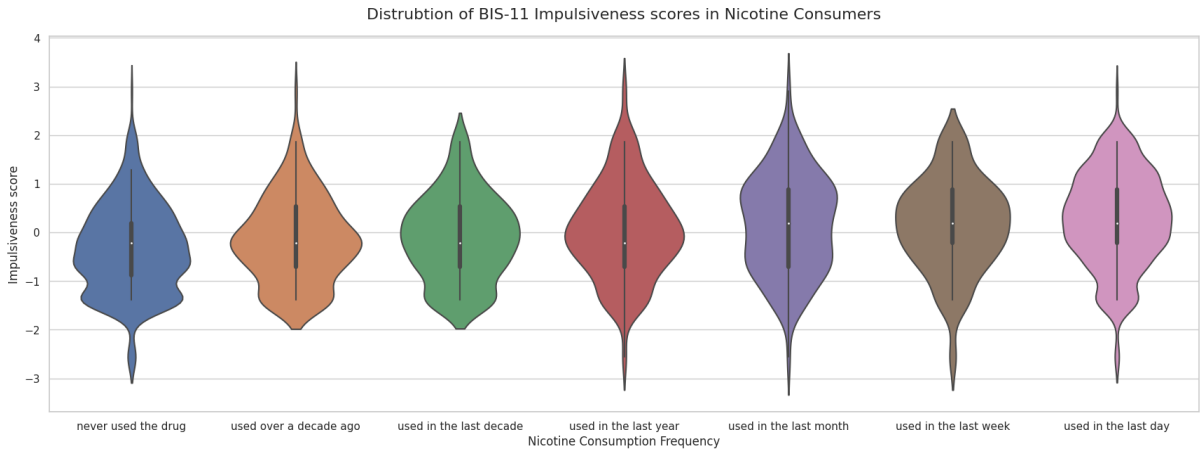
Graph 5: Alcohol Consumption by Ethnicity Education



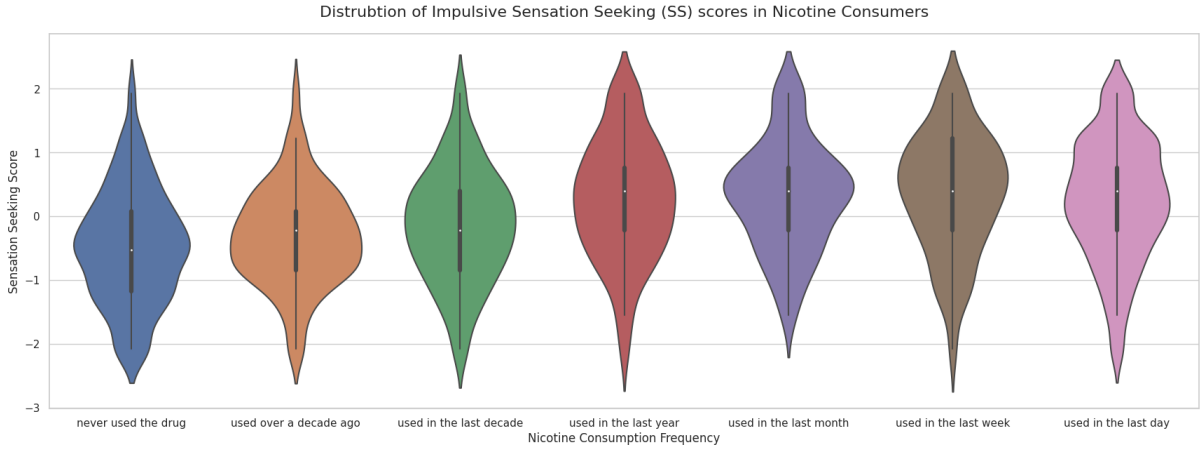
Graph 6 : Distribution of impulsive sensation seeking (SS) Scores among Alcohol Consumers



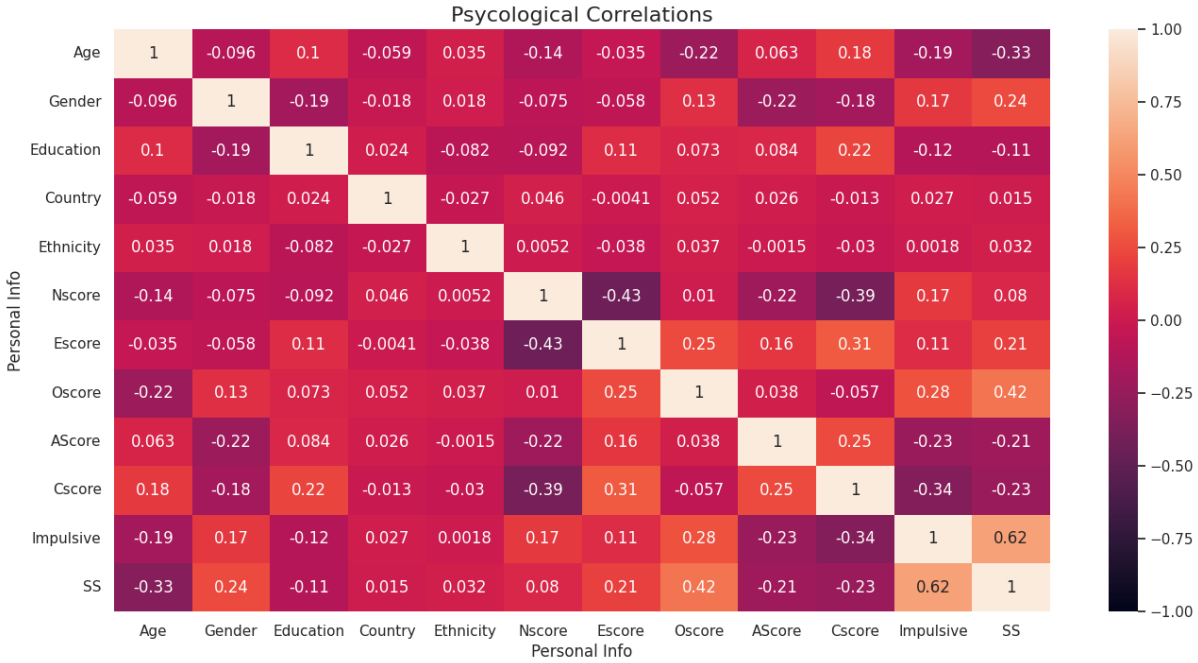
Graph 7 :Nicotine Consumption by Gender



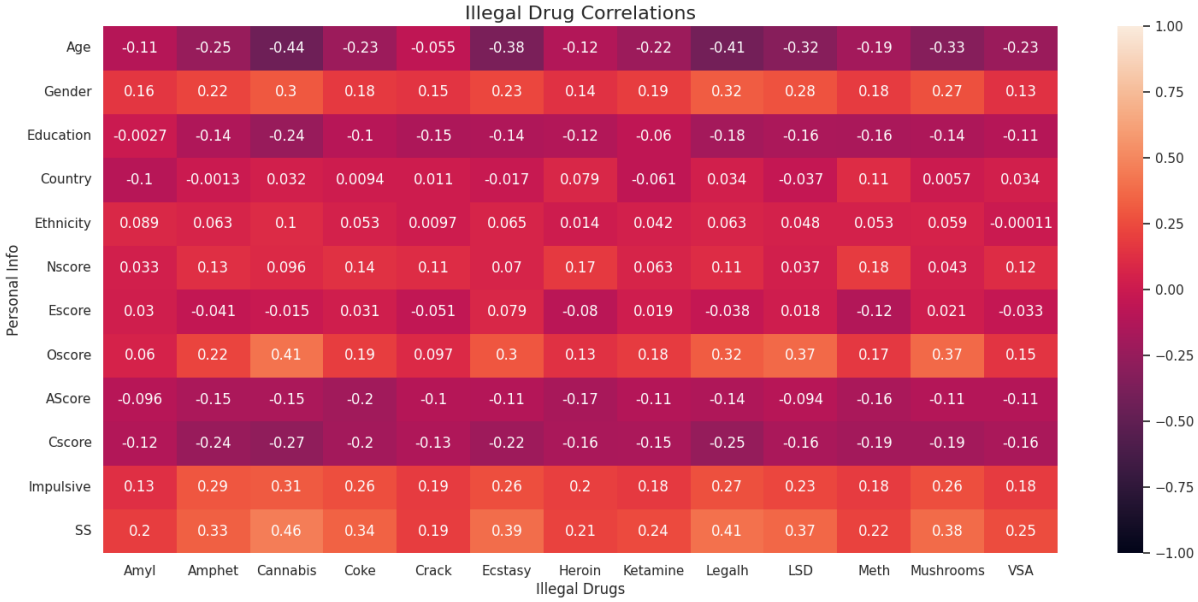
Graph 8 : Distribution of BIS-11 Impulsiveness scores in Nicotine Consumers



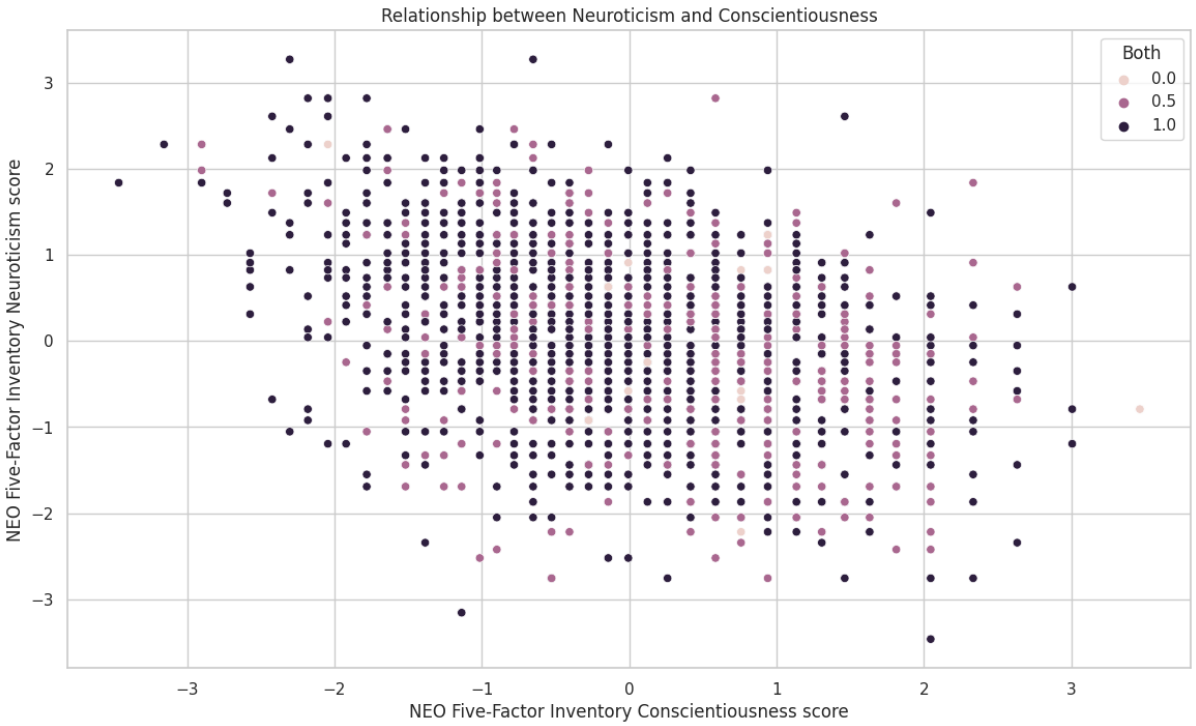
Graph 9 : Distribution of Impulsive Sensation Seeking (SS) scores in Nicotine Consumers



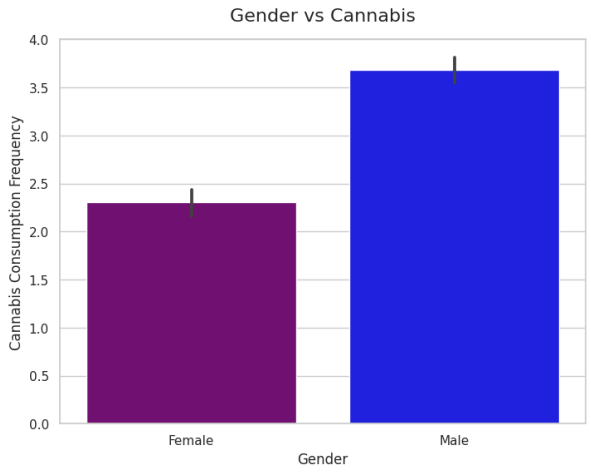
Graph 10 : Psychological Correlations



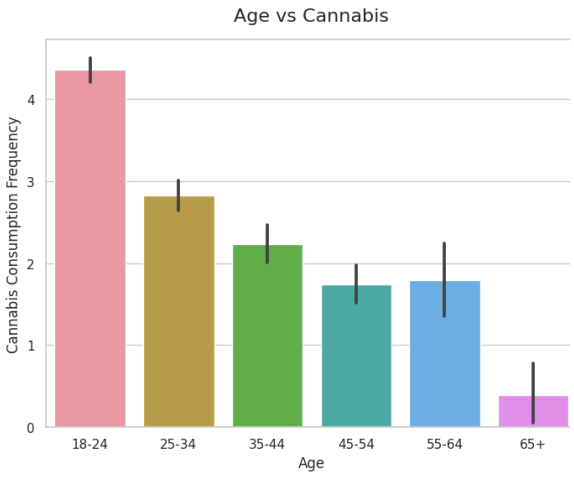
Graph 11 : Illegal Drugs Correlations



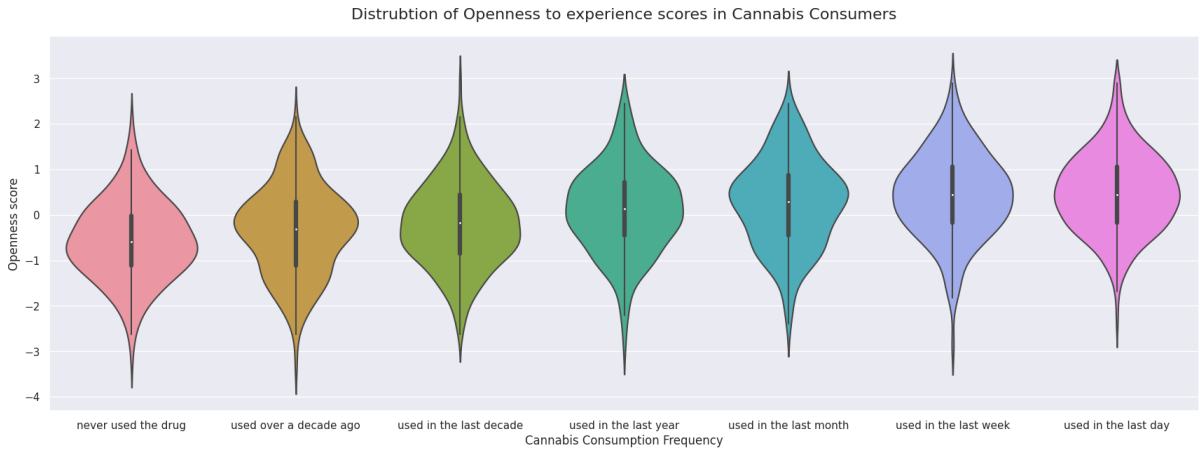
Graph 12 : Relationship between Neuroticism and Conscientiousness



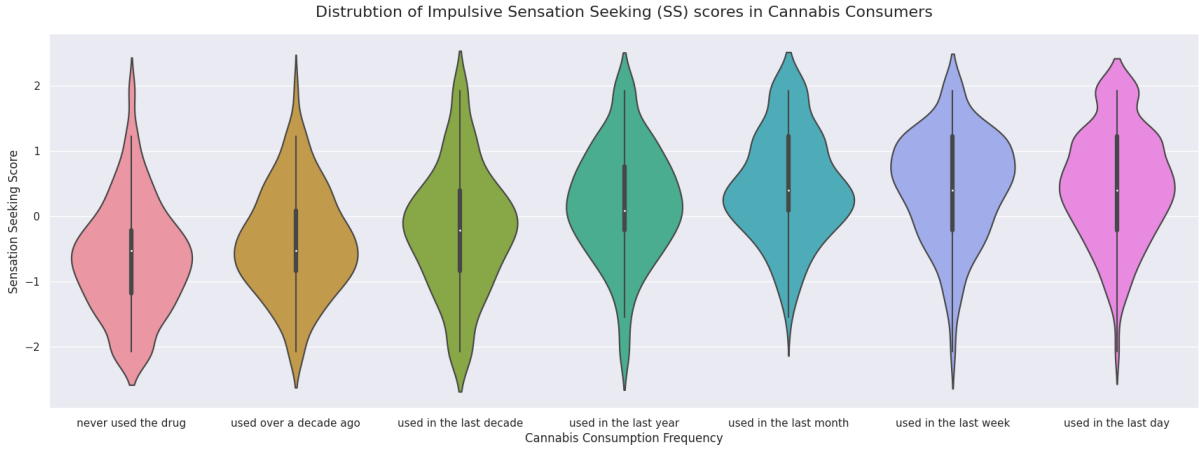
Graph 13 : Cannabis Consumption by Gender



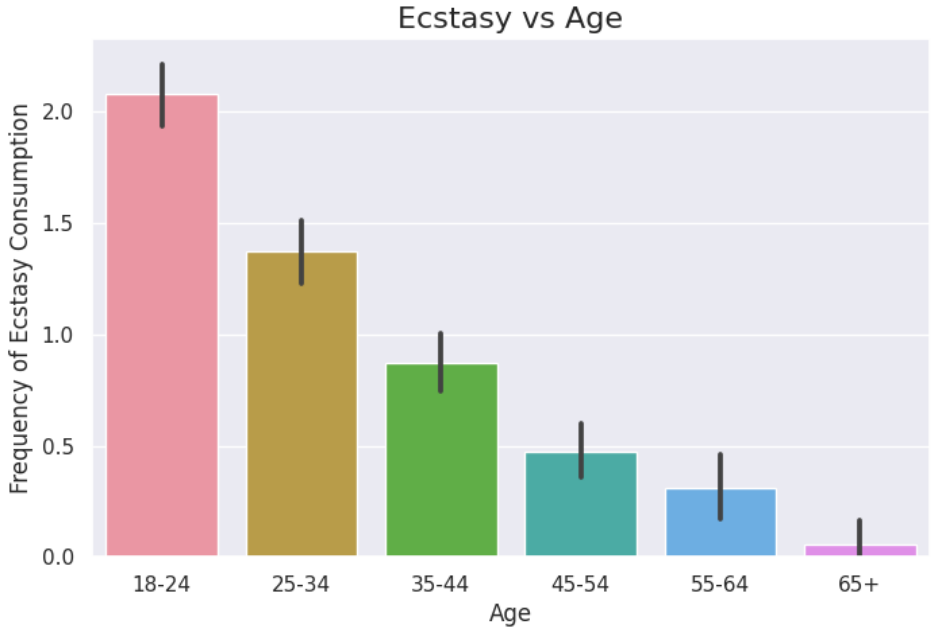
Graph 14 : Cannabis Consumption by Age



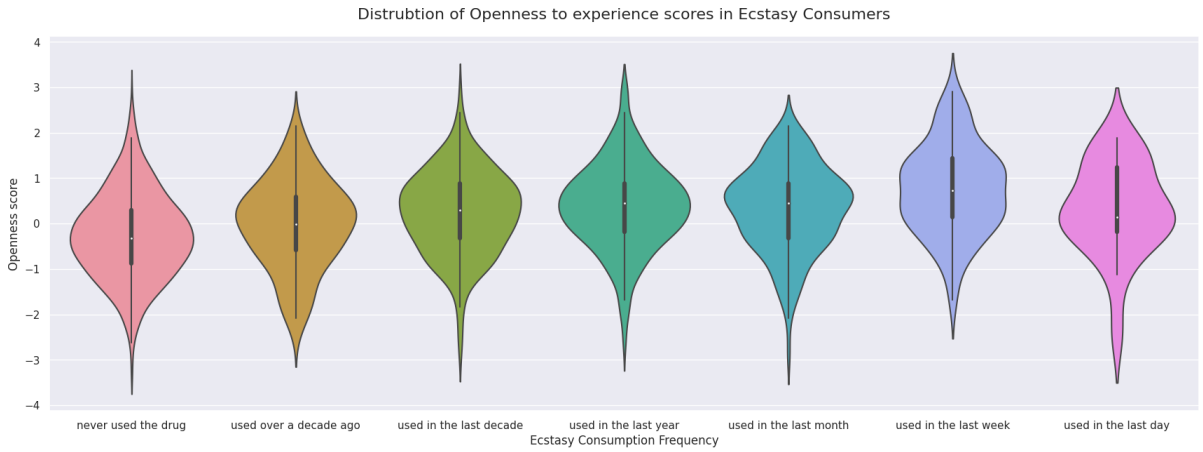
Graph 15 : Distribution of Openness to experience scores in Cannabis Consumers



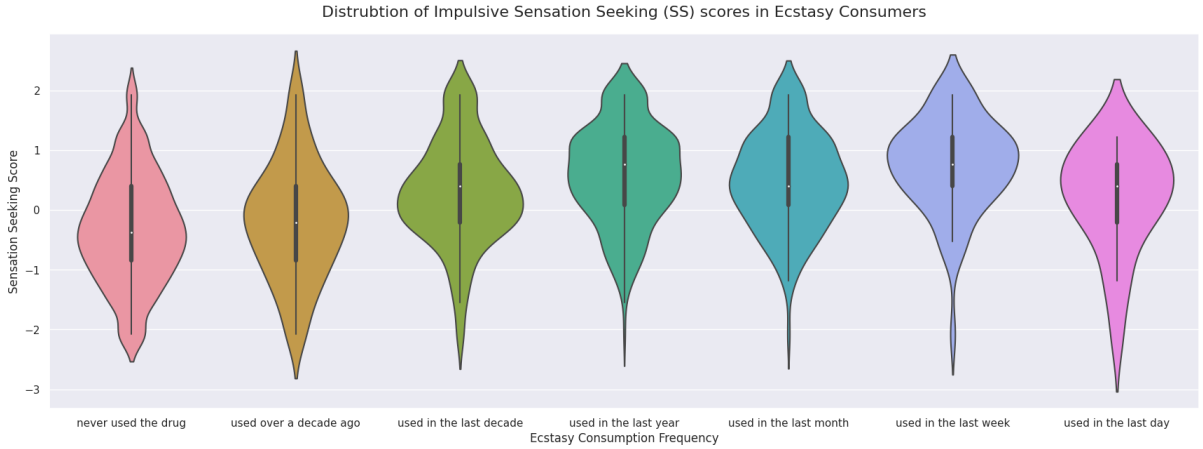
Graph 16 : Distribution of Impulsive Sensation Seeking (SS) scores in Cannabis Consumers



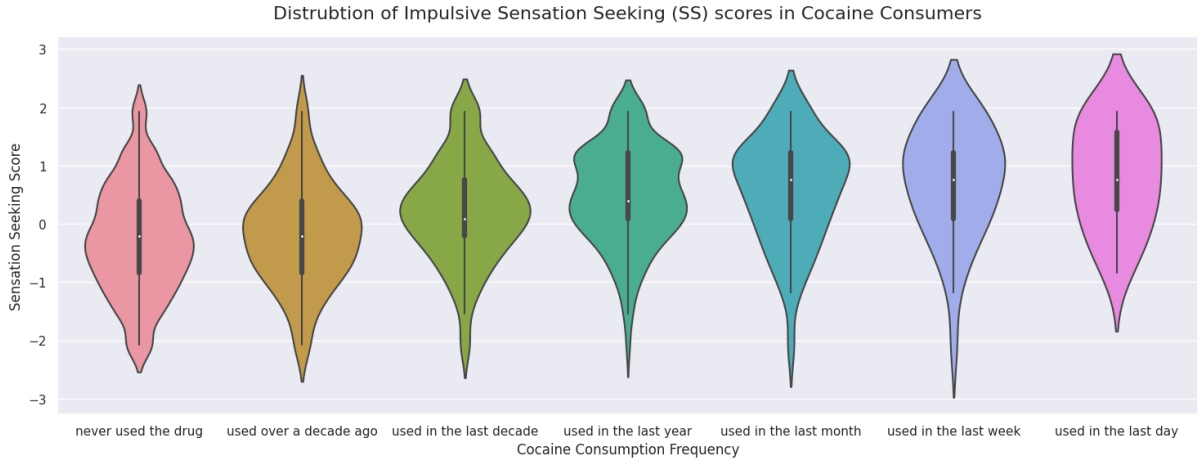
Graph 17 : Ecstasy Consumption by Age



Graph 18 : Distribution of Openness to experience scores in Ecstasy Consumers



Graph 19 : Distribution of Impulsive Sensation Seeking (SS) scores in Ecstasy Consumers



Graph 20 : Distribution of Impulsive Sensation Seeking (SS) scores in Cocaine Consumers

2. Naive Bayes

```
Accuracy Score de [ Alcohol ] : 0.9414893617021277
F1 Score de [ Alcohol ] : 0.9695290858725762
--------------------
Accuracy Score de [ Nicotine ] : 0.6888297872340425
F1 Score de [ Nicotine ] : 0.7771428571428572
--------------------
Accuracy Score de [ Amyl ] : 0.7340425531914894
F1 Score de [ Amyl ] : 0.23076923076923075
--------------------
Accuracy Score de [ Amphet ] : 0.7287234042553191
F1 Score de [ Amphet ] : 0.6506849315068494
--------------------
Accuracy Score de [ Cannabis ] : 0.7686170212765957
F1 Score de [ Cannabis ] : 0.8277227722772278
--------------------
Accuracy Score de [ Coke ] : 0.6914893617021277
F1 Score de [ Coke ] : 0.5999999999999999
--------------------
Accuracy Score de [ Crack ] : 0.8643617021276596
F1 Score de [ Crack ] : 0.2608695652173913
--------------------
Accuracy Score de [ Ecstasy ] : 0.7154255319148937
F1 Score de [ Ecstasy ] : 0.6603174603174603
--------------------
Accuracy Score de [ Heroin ] : 0.848404255319149
F1 Score de [ Heroin ] : 0.27848101265822783
--------------------
Accuracy Score de [ Ketamine ] : 0.7553191489361702
F1 Score de [ Ketamine ] : 0.3134328358208955
--------------------
Accuracy Score de [ Legalh ] : 0.7154255319148937
F1 Score de [ Legalh ] : 0.6514657980456026
--------------------
Accuracy Score de [ LSD ] : 0.726063829787234
F1 Score de [ LSD ] : 0.5654008438818565
--------------------
Accuracy Score de [ Meth ] : 0.7127659574468085
F1 Score de [ Meth ] : 0.425531914893617
--------------------
Accuracy Score de [ Mushrooms ] : 0.6968085106382979
F1 Score de [ Mushrooms ] : 0.6174496644295302
--------------------
Accuracy Score de [ VSA ] : 0.848404255319149
F1 Score de [ VSA ] : 0.3132530120481928
--------------------
```

Figure 1 : Naive Bayes Accuracy Scores

3. Random Forest Classifier

```
Accuracy Score de [ Alcohol ] : 0.9601063829787234
F1 Score de [ Alcohol ] : 0.9796472184531885
--------------------
Accuracy Score de [ Nicotine ] : 0.7021276595744681
F1 Score de [ Nicotine ] : 0.7902621722846442
--------------------
Accuracy Score de [ Amyl ] : 0.8085106382978723
F1 Score de [ Amyl ] : 0.1
--------------------
Accuracy Score de [ Amphet ] : 0.6914893617021277
F1 Score de [ Amphet ] : 0.532258064516129
--------------------
Accuracy Score de [ Cannabis ] : 0.7952127659574468
F1 Score de [ Cannabis ] : 0.846307385229541
--------------------
Accuracy Score de [ Coke ] : 0.6968085106382979
F1 Score de [ Coke ] : 0.5169491525423728
--------------------
Accuracy Score de [ Crack ] : 0.8936170212765957
F1 Score de [ Crack ] : 0.047619047619047616
--------------------
Accuracy Score de [ Ecstasy ] : 0.7127659574468085
F1 Score de [ Ecstasy ] : 0.6142857142857143
--------------------
Accuracy Score de [ Heroin ] : 0.8696808510638298
F1 Score de [ Heroin ] : 0.0392156862745098
--------------------
Accuracy Score de [ Ketamine ] : 0.7925531914893617
F1 Score de [ Ketamine ] : 0.04878048780487805
--------------------
Accuracy Score de [ Legalh ] : 0.7579787234042553
F1 Score de [ Legalh ] : 0.6829268292682927
--------------------
Accuracy Score de [ LSD ] : 0.7712765957446809
F1 Score de [ LSD ] : 0.5376344086021505
--------------------
Accuracy Score de [ Meth ] : 0.7872340425531915
F1 Score de [ Meth ] : 0.3442622950819672
--------------------
Accuracy Score de [ Mushrooms ] : 0.7579787234042553
F1 Score de [ Mushrooms ] : 0.6513409961685823
--------------------
Accuracy Score de [ VSA ] : 0.851063829787234
F1 Score de [ VSA ] : 0.0
--------------------
```

Figure 2 : Random Forest Accuracy Scores

4. SVM Classifier

```
Accuracy Score de [ Alcohol ] : 0.9627659574468085
F1 Score de [ Alcohol ] : 0.9810298102981029
----------
Accuracy Score de [ Nicotine ] : 0.7047872340425532
F1 Score de [ Nicotine ] : 0.8146911519198665
----------
Accuracy Score de [ Amyl ] : 0.824468085106383
F1 Score de [ Amyl ] : 0.0
----------
Accuracy Score de [ Amphet ] : 0.7287234042553191
F1 Score de [ Amphet ] : 0.5641025641025641
----------
Accuracy Score de [ Cannabis ] : 0.7712765957446809
F1 Score de [ Cannabis ] : 0.8371212121212123
----------
Accuracy Score de [ Coke ] : 0.675531914893617
F1 Score de [ Coke ] : 0.37755102040816324
----------
Accuracy Score de [ Crack ] : 0.8909574468085106
F1 Score de [ Crack ] : 0.0
----------
Accuracy Score de [ Ecstasy ] : 0.7420212765957447
F1 Score de [ Ecstasy ] : 0.6472727272727272
```

```
Accuracy Score de [ Heroin ] : 0.8723404255319149
F1 Score de [ Heroin ] : 0.0
----------
Accuracy Score de [ Ketamine ] : 0.8031914893617021
F1 Score de [ Ketamine ] : 0.0
----------
Accuracy Score de [ Legalh ] : 0.7712765957446809
F1 Score de [ Legalh ] : 0.6950354609929078
----------
Accuracy Score de [ LSD ] : 0.7686170212765957
F1 Score de [ LSD ] : 0.47272727272727266
----------
Accuracy Score de [ Meth ] : 0.7925531914893617
F1 Score de [ Meth ] : 0.0
----------
Accuracy Score de [ Mushrooms ] : 0.7792553191489362
F1 Score de [ Mushrooms ] : 0.6795366795366794
----------
Accuracy Score de [ VSA ] : 0.8643617021276596
F1 Score de [ VSA ] : 0.0
```

Figure 3 : SVM Accuracy Scores

5. Logistic Regression

```
Accuracy Score de [ Alcohol ] : 0.9601063829787234
F1 Score de [ Alcohol ] : 0.9796472184531885
----------
Accuracy Score de [ Nicotine ] : 0.6968085106382979
F1 Score de [ Nicotine ] : 0.7964285714285714
----------
Accuracy Score de [ Amyl ] : 0.8191489361702128
F1 Score de [ Amyl ] : 0.028571428571428574
----------
Accuracy Score de [ Amphet ] : 0.7101063829787234
F1 Score de [ Amphet ] : 0.5240174672489083
----------
Accuracy Score de [ Cannabis ] : 0.75
F1 Score de [ Cannabis ] : 0.8233082706766918
----------
Accuracy Score de [ Coke ] : 0.7047872340425532
F1 Score de [ Coke ] : 0.4884792626728111
----------
Accuracy Score de [ Crack ] : 0.8909574468085106
F1 Score de [ Crack ] : 0.0
----------
Accuracy Score de [ Ecstasy ] : 0.726063829787234
F1 Score de [ Ecstasy ] : 0.6113207547169812
```

```
Accuracy Score de [ Heroin ] : 0.8696808510638298
F1 Score de [ Heroin ] : 0.0
----------
Accuracy Score de [ Ketamine ] : 0.8058510638297872
F1 Score de [ Ketamine ] : 0.0759493670886076
----------
Accuracy Score de [ Legalh ] : 0.7446808510638298
F1 Score de [ Legalh ] : 0.6571428571428571
----------
Accuracy Score de [ LSD ] : 0.7579787234042553
F1 Score de [ LSD ] : 0.46783625730994155
----------
Accuracy Score de [ Meth ] : 0.7872340425531915
F1 Score de [ Meth ] : 0.2156862745098039
----------
Accuracy Score de [ Mushrooms ] : 0.7127659574468085
F1 Score de [ Mushrooms ] : 0.5304347826086957
----------
Accuracy Score de [ VSA ] : 0.8590425531914894
F1 Score de [ VSA ] : 0.03636363636363636
```

Figure 4 : Logistic Regression Accuracy Scores

6. Logistic Regression avec GridSearch

```
tuned_parameters = [{"solver" : ['liblinear'],"multi_class" : ['ovr'],
        "class_weight" : [None, 'balanced'],"max_iter" : [100,200,500,800, 1600]},
        {"solver" : ['newton-cg'],"multi_class" : ['ovr','multinomial'],
        "class_weight" : [None, 'balanced'],"max_iter" : [100,200,500,800, 1600]}]
model_cv = GridSearchCV(model, tuned_parameters, cv=10,scoring='f1')
for drug in legal_drugs:
    predict_drug_usage(model_cv, drug)
    print('---------------------')
for drug in illegal_drugs:
    predict_drug_usage(model_cv, drug)
    print('---------------------')
```

```
Accuracy Score de [ Alcohol ] : 0.9796472184531885
F1 Score de [ Alcohol ] : 0.9796472184531885
---------------------
Accuracy Score de [ Nicotine ] : 0.7964285714285714
F1 Score de [ Nicotine ] : 0.7964285714285714
---------------------
Accuracy Score de [ Amyl ] : 0.3700440528634361
F1 Score de [ Amyl ] : 0.3700440528634361
---------------------
Accuracy Score de [ Amphet ] : 0.6449511400651466
F1 Score de [ Amphet ] : 0.6449511400651466
---------------------
Accuracy Score de [ Cannabis ] : 0.8233082706766918
F1 Score de [ Cannabis ] : 0.8233082706766918
---------------------
Accuracy Score de [ Coke ] : 0.6143790849673203
F1 Score de [ Coke ] : 0.6143790849673203
---------------------
Accuracy Score de [ Crack ] : 0.3492063492063492
F1 Score de [ Crack ] : 0.3492063492063492
---------------------
Accuracy Score de [ Ecstasy ] : 0.6624203821656051
F1 Score de [ Ecstasy ] : 0.6624203821656051
```

```
Accuracy Score de [ Heroin ] : 0.3837209302325581
F1 Score de [ Heroin ] : 0.3837209302325581
---------------------
Accuracy Score de [ Ketamine ] : 0.43555555555555553
F1 Score de [ Ketamine ] : 0.43555555555555553
---------------------
Accuracy Score de [ Legalh ] : 0.6945337620578779
F1 Score de [ Legalh ] : 0.6945337620578779
---------------------
Accuracy Score de [ LSD ] : 0.6067415730337079
F1 Score de [ LSD ] : 0.6067415730337079
---------------------
Accuracy Score de [ Meth ] : 0.4622222222222222
F1 Score de [ Meth ] : 0.4622222222222222
---------------------
Accuracy Score de [ Mushrooms ] : 0.6775244299674267
F1 Score de [ Mushrooms ] : 0.6775244299674267
---------------------
Accuracy Score de [ VSA ] : 0.37288135593220334
F1 Score de [ VSA ] : 0.37288135593220334
```

Figure 5 : Logistic Regression with Grid Search Accuracy Scores