

Marie-Alix Honoré
Yasmina Bendaoud



**Analyse d'un dataset sur la consommation de
drogue parmi 1885 sujets**



Sommaire

- 1) Présentation du dataset
- 2) Problématique
- 3) Modification des variables
- 4) Feature Engineering
- 5) Modèles de prédiction
- 6) Regard critique sur les résultats et conclusion

1) Présentation du dataset

- Récolté sur le site UCI – Machine Learning Repository
- Date de 2016
- Réuni des données numériques et non numériques sur 1885 personnes interrogées
- Classifie les consommateurs de différentes drogues en fonction de leur personnalité
- Offre des données à l'international
- 32 attributs pour chaque personne
- Chaque ligne du dataset correspond à la consommation de différentes drogues selon la personnalité et les caractéristiques d'un individu.

Data Set Characteristics:	Multivariate	Number of Instances:	1885	Area:	Social
Attribute Characteristics:	Real	Number of Attributes:	32	Date Donated	2016-10-17
Associated Tasks:	Classification	Missing Values?	N/A	Number of Web Hits:	141503

Vue sur les attributs :

- **ID** : nombre référence de la personne
- **Age** : âge du participant
- **Gender** : sexe
- **Education** : niveau d'éducation
- **Country** : pays de résidence
- **Ethnicity** : origine du participant
- **Nscore*** : Névrosisme (Anxiété, Colère-Hostilité, Dépression, Timidité sociale, Impulsivité et Vulnérabilité)
- **Escore*** : Extraversion (Chaleur, Grégarité, Assertivité, Activité, Recherche de sensations et Émotions positives)
- **Oscore*** : Ouverture (Ouverture aux rêveries, à l'esthétique, aux sentiments, aux actions, aux idées et aux valeurs)
- **Ascore*** : Agréabilité (Confiance, Droiture, Altruisme, Compliance, Modestie, Sensibilité)
- **Cscore*** : Conscience (Compétence, Ordre, Sens du devoir, Recherche de réussite, Autodiscipline, Délibération)
- **Impulsive** : impulsivité d'après le test BIS-11
- **SensationSeeking** : recherche d'attention mesuré par ImpSS

**d'après le test de personnalité "NEO-FFI-R"*

Vue sur les attributs (suite) :

- **Alcohol** : consommation d'alcool
- **Amphet** : consommation d'amphetamines
- **Amyl** : consommation d'amyl nitrite
- **Benzos** : consommation de benzodiazepine
- **Caff** : consommation de cafeine
- **Cannabis** : consommation de cannabis
- **Choc** : consommation de chocolat
- **Coke** : consommation de cocaine
- **Crack** : consommation de crack
- **Ecstasy** : consommation d'ecstasy
- **Heroin** : consommation d'héroïne
- **Ketamine** : consommation de ketamine
- **Legalh** : consommation de drogue psychoactive
- **LSD** : consommation de LSD
- **Meth** : consommation de methadone
- **Mushrooms** : consommation de champignons magiques
- **Nicotine** : consommation de nicotine
- **Semer** : consommation de la drogue "Semeron"
- **VSA** : "volatile substance abuse", consommation abusive de substance volatile

Les consommations de drogues sont échelonnées de CL0 (consommation la plus faible) à CL6 (la plus élevée)

Droque et personnalité

Plusieurs attributs proviennent de tests de personnalités reconnus en psychologie.

NEO-FFI-R

NEO-FFI-R

Extraversion

Tendance optimiste à se focaliser sur les aspects positifs de la réalité. Tendance à éprouver fréquemment et intensément des émotions positives (joie, enthousiasme). Tendance à rechercher les contacts (sociabilité), l'activité (énergie), les stimulations et à y trouver du plaisir.

Névrosisme

Tendance à se focaliser sur les aspects négatifs de la réalité, à les grossir et à les ruminer. Tendance à éprouver fréquemment et intensément des émotions négatives (peur, anxiété, tristesse, honte...)

Ouverture à l'expérience

Tendance à rechercher la nouveauté dans différents secteurs (idées, actions, valeurs, méthodes,...)
Tendance à la curiosité, à l'originalité, à rechercher des cadres et des perspectives nouvelles

Agréabilité

Tendance à la bienveillance, l'altruisme, la compréhension des autres. Tendance à être attentif au vécu, aux intérêts et au point de vue des autres. Recherche des relations harmonieuses, consensuelles

Caractère Conscientieux

Tendance à se fixer spontanément - pour soi-même- des objectifs élevés et à les poursuivre de manière ordonnée, organisée et persévérante.

BIS-11

Second-order Factors
Attentional:
Motor:
Nonplanning:

First-order Factors
Attention:
Motor:
Self-Control:
Cognitive Complexity:
Perseverance:
Cognitive Instability:

ImpSS

C'est une échelle qui permet de mesurer la recherche d'attention d'une personne

Ce dataset récolte des données caractéristiques des personnes dont leur personnalité

2) Problématique

Ces données soulèvent la question suivante :

Ce dataset permet-il de relever des résultats réalistes sur la consommation de drogue dans le monde ?

3) Modification des variables

Nous avons supprimé la colonne "Id" qui ne sera pas utilisée

Au départ, les données étaient des flottants pour tous les attributs sauf pour les drogues mesurées par des classes (string)

Les drogues sont classées ainsi :

CL0 : jamais consommée

CL1 : il y a plus de 10 ans

CL2 : la décennie dernière

CL3 : l'année dernière

CL4 : le mois dernier

CL5 : la semaine dernière

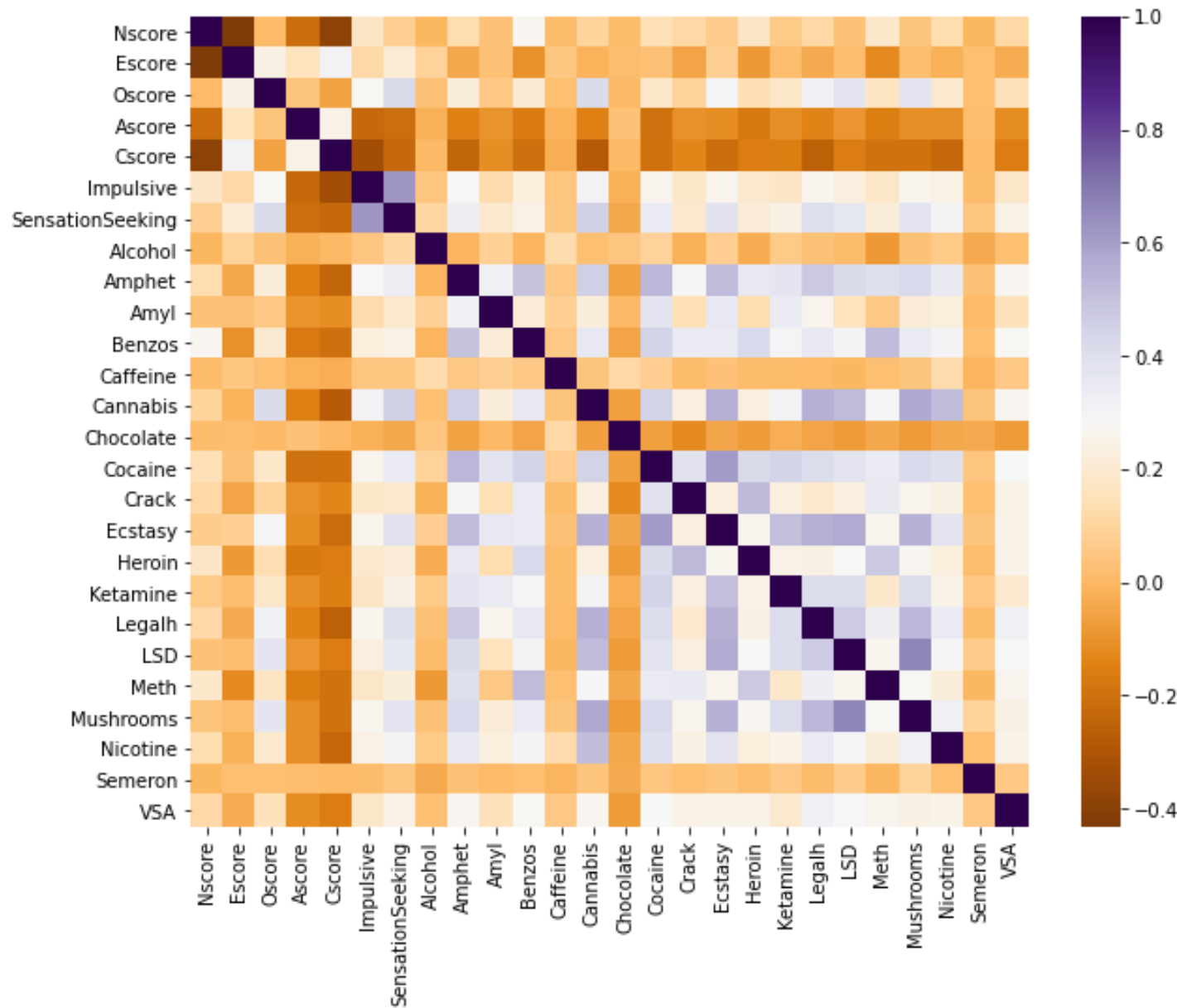
CL6 : le jour dernier

Nous avons donc décidé de :

- Changer le type de toutes les colonnes de Age jusqu'à Ethnicity pour avoir des strings plutôt que des valeurs chiffrées pas très parlantes
- Changer toutes les colonnes consommation en mettant un classement de 0 à 6 pour pouvoir traiter ces données par la suite par des calculs et des algorithmes de classification.

4) Feature engineering et Visualisations

Recherche de corrélations : Heatmap



Ce heatmap nous montre plusieurs choses :

- Certaines consommations de drogues sont corrélées : Cannabis, Cocaine, Ecstasy, ce qui montre qu'on peut créer des patterns de consommation.
- Certains traits de personnalité influent sur la prise de stupéfiant: Oscore, Ascore et Cscore sont corrélées négativement.

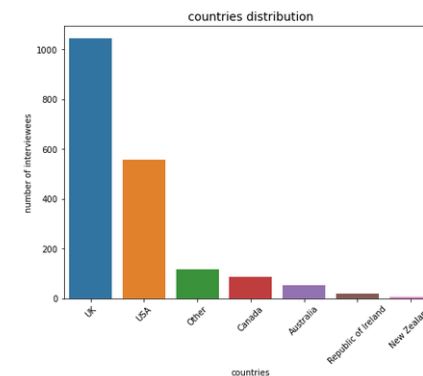
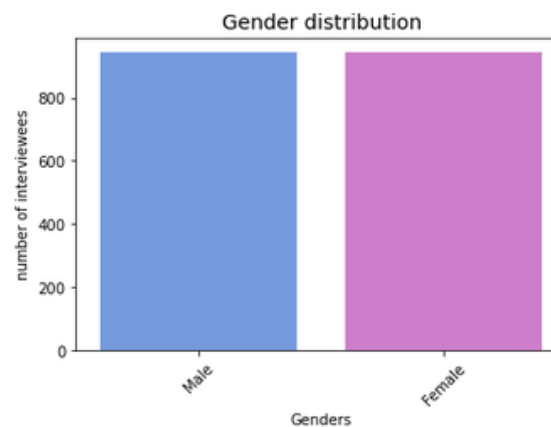
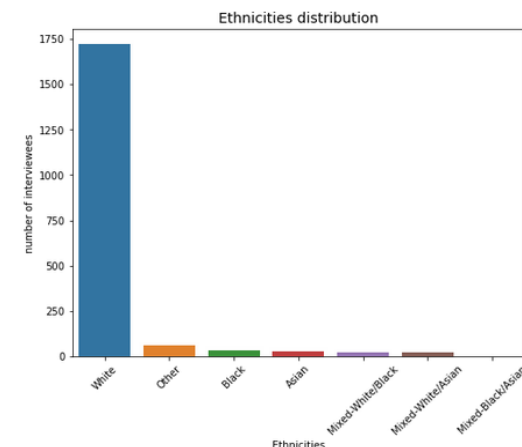
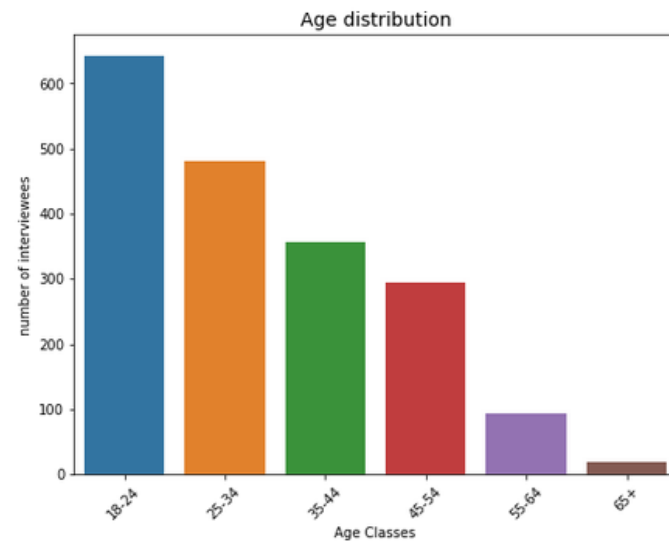
Ces premiers résultats sont assez intéressants : les personnes ayant des tendances à certaines consommations ne sont pas les plus ouvertes aux autres, les plus agréables ou les plus conscients du monde et d'eux-mêmes

Origine des personnes interrogées

Ensuite, nous voulons étudier les données pour en savoir plus sur les personnes interrogées : combien de personnes pour chaque tranche d'âge, origine, sexe et pays d'origine.

- L'âge le plus représenté dans le dataset est les 18-24 ans suivi par les 25-34 ans
- Les personnes blanches sont sur-représentées par rapport aux autres
- Les personnes provenant de l'Angleterre (UK) sont sur-représentés également
- Il y a une très bonne répartition par sexe (autant d'hommes que de femmes)

Les données n'étant pas réparties de manière homogène, on peut dire que les résultats peuvent donner un aperçu mais pas une idée réelle des choses.

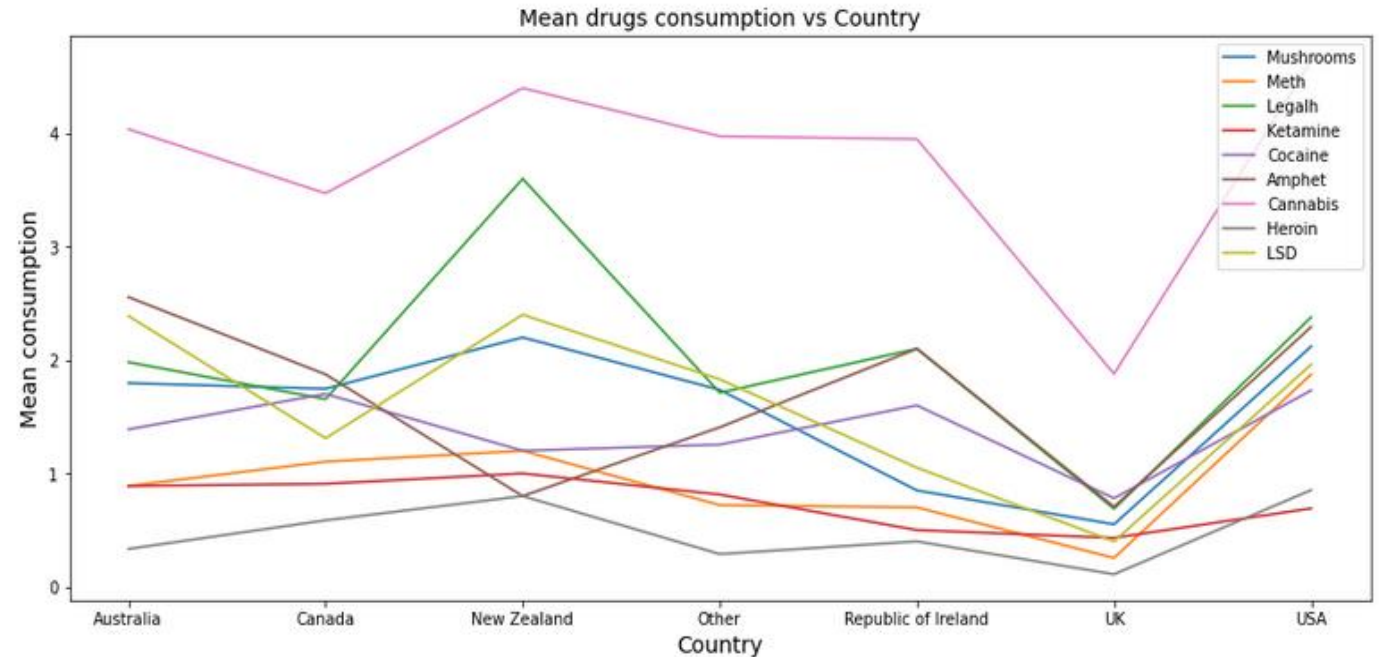


Origines des personnes interrogées

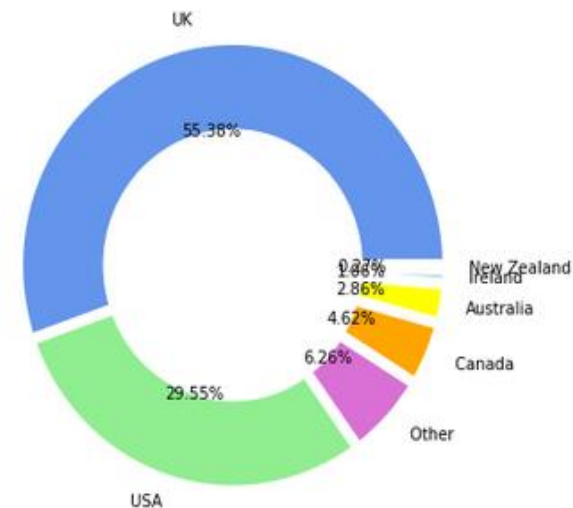
Nous voulons étudier la moyenne de consommation de plusieurs drogues en fonction du pays de résidence.

Comme la Nouvelle Zélande ou l'Irlande sont sous représentés dans les données, nous ne pouvons pas faire de constat là dessus (3 personnes de New Zeland dans le dataset).

On peut par contre comparer les USA et le UK et on voit qu'il y a une plus grande consommation des drogues aux USA, surtout pour la cannabis. On peut tout de même constater un pic de Legalh pour New Zealand.

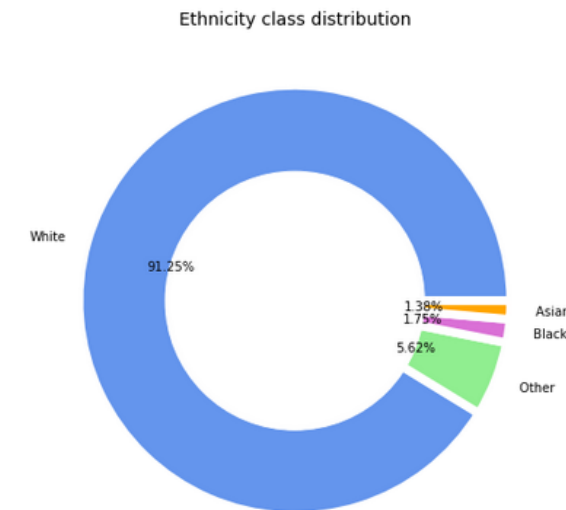
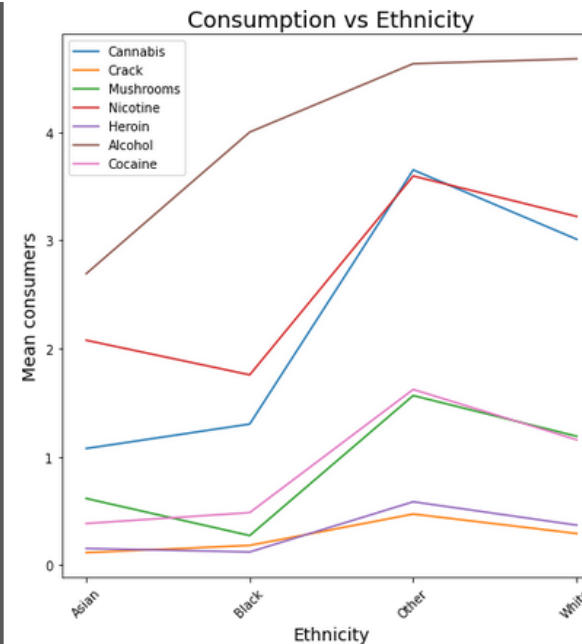
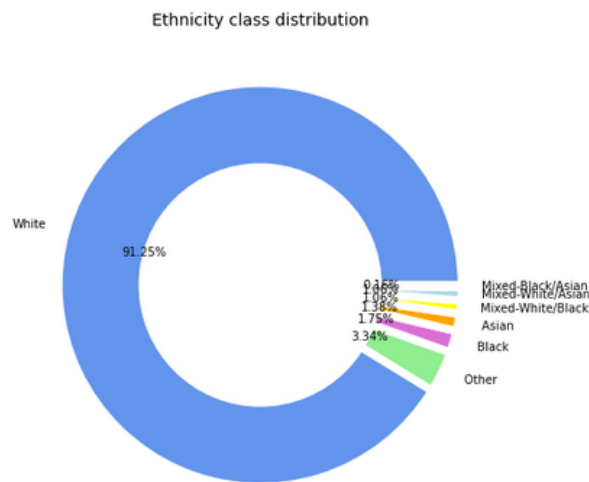
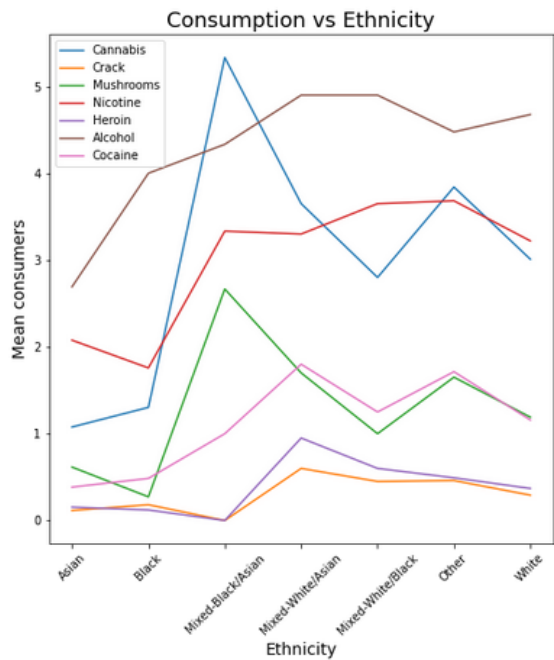


Countries distribution





Consommation en fonction de l'origine

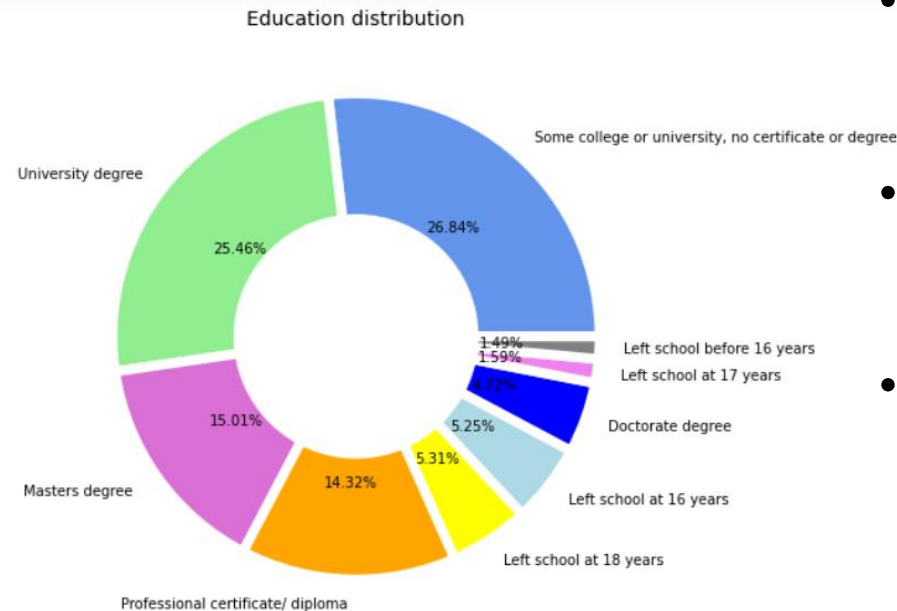
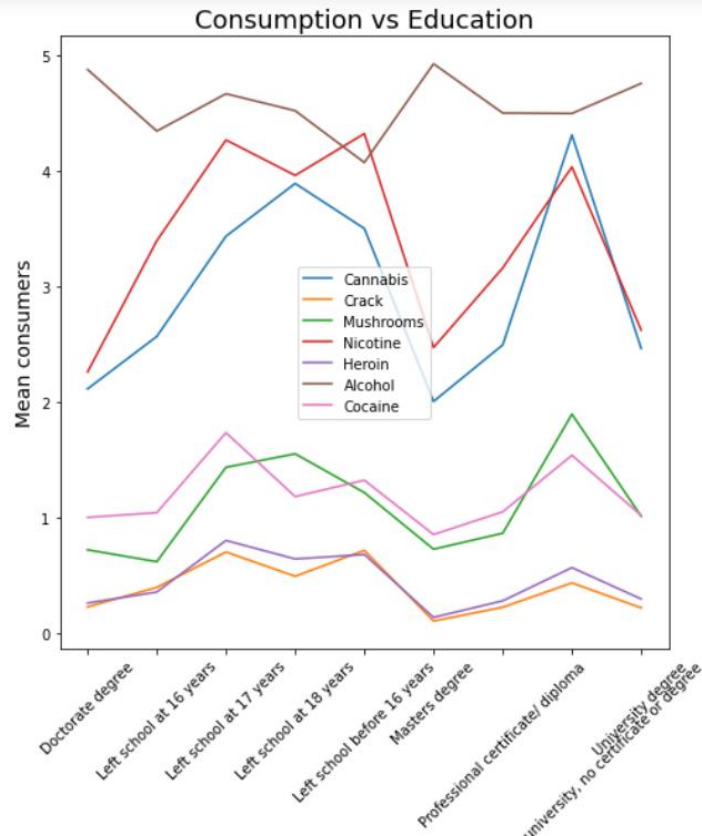


Certaines valeurs de variable étant sous-représentées. Ici, pour ethnicity, on pourrait croire que la population mixed-Black/Asian consomme plus de Mushrooms et de Cannabis que la moyenne. Ceci est dû à son effectif de 0,16% du dataset.

On décide ici de regrouper toutes les catégories d'ethnie en faible effectif dans la variable "Other" ce qui nous conduit à 4 catégories d'ethnicité.

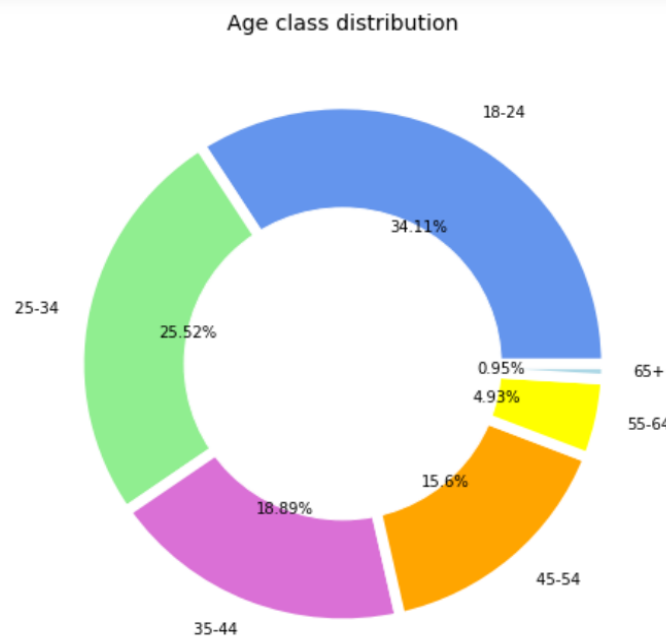
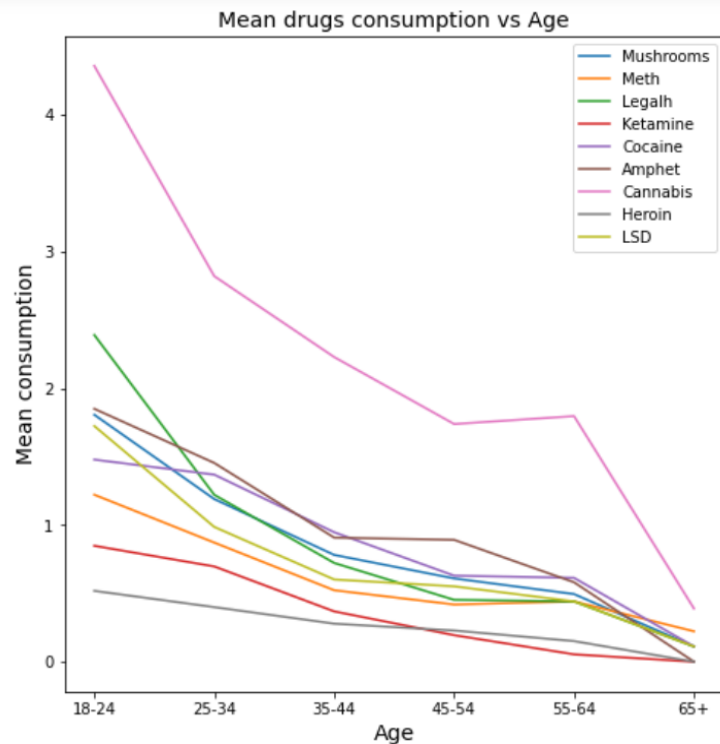
Les drogues les plus consommées toute ethnicité confondues sont l'alcool en numéro 1 puis la nicotine, c'est bien représentatif de la réalité.

Consommation en fonction du niveau d'études



- L'alcool est le plus consommé chez les personnes ayant quitté l'école avant 16 ans.
- La nicotine est le plus consommé chez les personnes ayant quitté l'école à 18 ans .
- Le cannabis est moins consommé chez les personnes ayant quitté l'école à 16 ans et le plus chez les personnes ayant un diplôme.
- Le crack est plus consommé par les personnes ayant fait le moins d'études.

Consommation en fonction de l'âge



- En prenant en compte la sur-représentation des plus jeunes, on voit globalement que la consommation de drogues diminue avec l'âge
- La drogue la plus consommée chez les 18-24 ans est le Cannabis. C'est aussi la drogue qui a la plus forte baisse de consommation.
- Pour les 45-54 ans, il y a une très légère hausse de la consommation de Cannabis et cocaïne.

Consommation moyenne de chaque substance addictive

Nous avons au préalable modifier les variables des substances addictives, passant de "CL0" à 0 et "CL1" à 1,...
Nous pouvons remarquer que pour la majorité des drogues l'effectif le plus important est la population n'en ayant jamais consommé excepté pour le cannabis où il y a plus de personnes qui en ont consommé dans les derniers 24h plutôt que jamais consommé.

0 : Never Used

1: Used over a Decade Ago

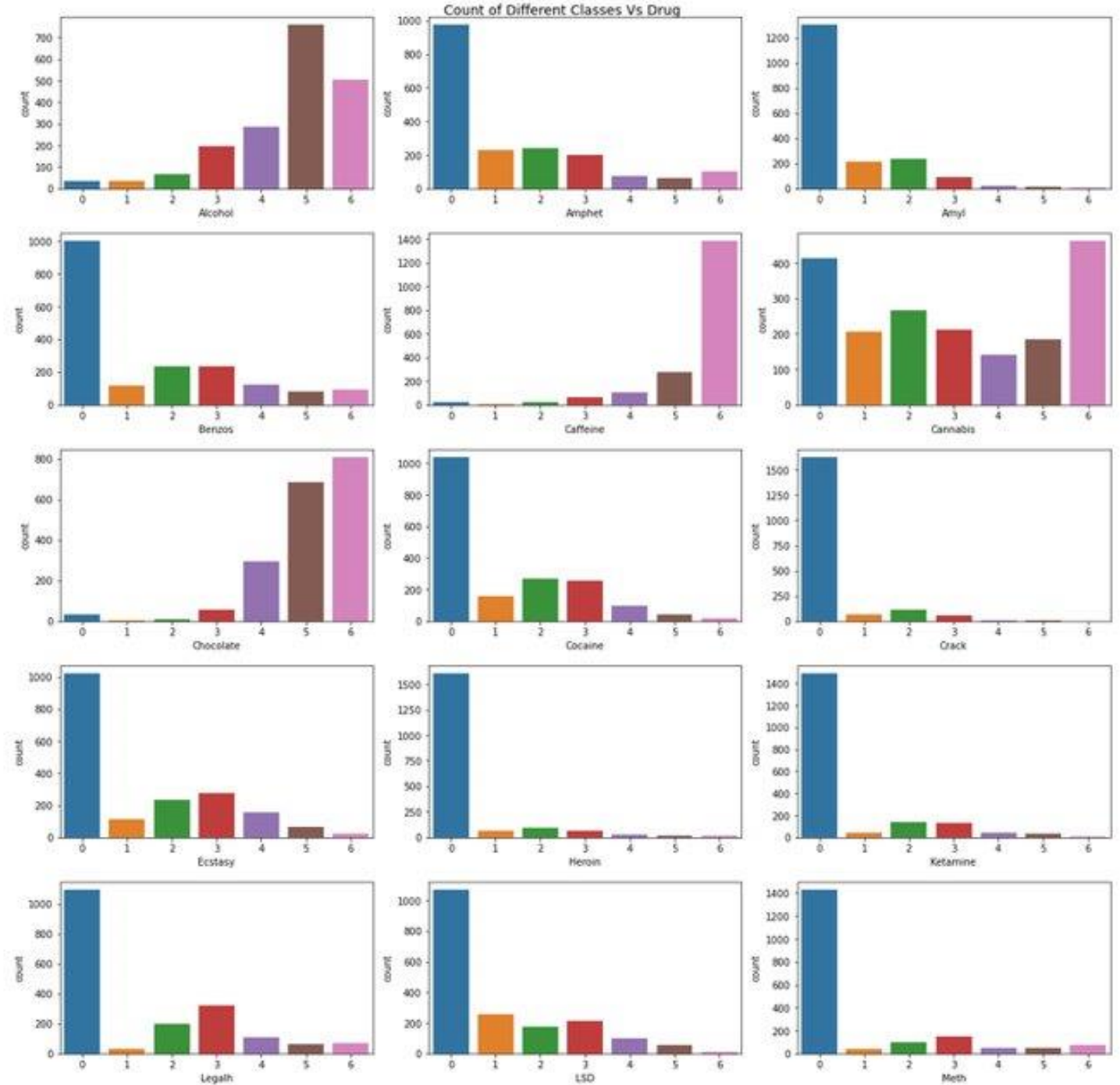
2 : Used in Last Decade

3 : Used in Last Year

4 : Used in Last Month

5 : Used in Last Week

6 : Used in Last Day

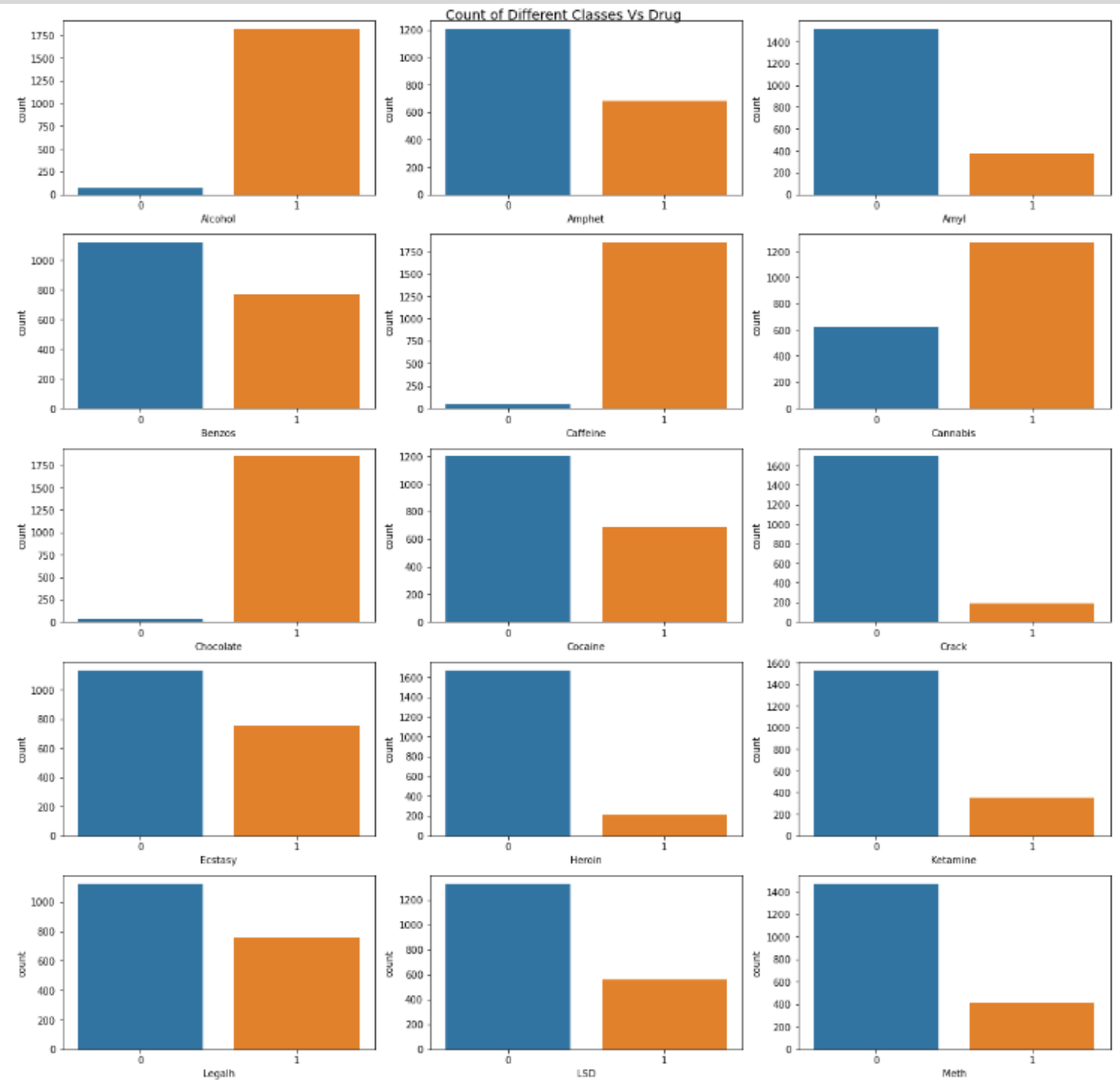


Consommation moyenne de chaque substance addictive

Nous avons reproduit les mêmes histogrammes mais cette fois-ci, nous avons modifié la variable telle que suggéré dans le descriptif du dataset : *"Never Used", "Used over a Decade Ago" form class "Non-user" and all other classes form class "User"*.

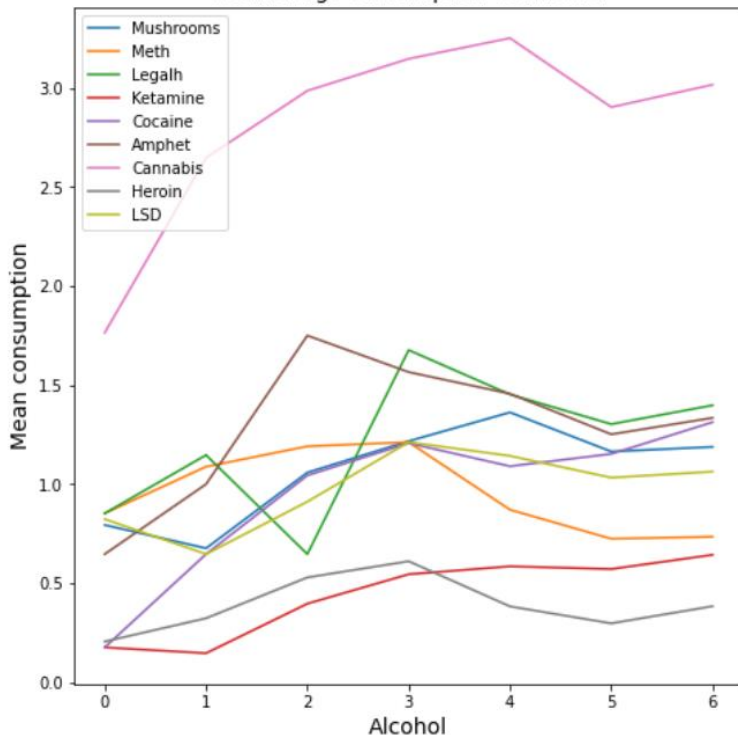
0 : Non-User

1 : User

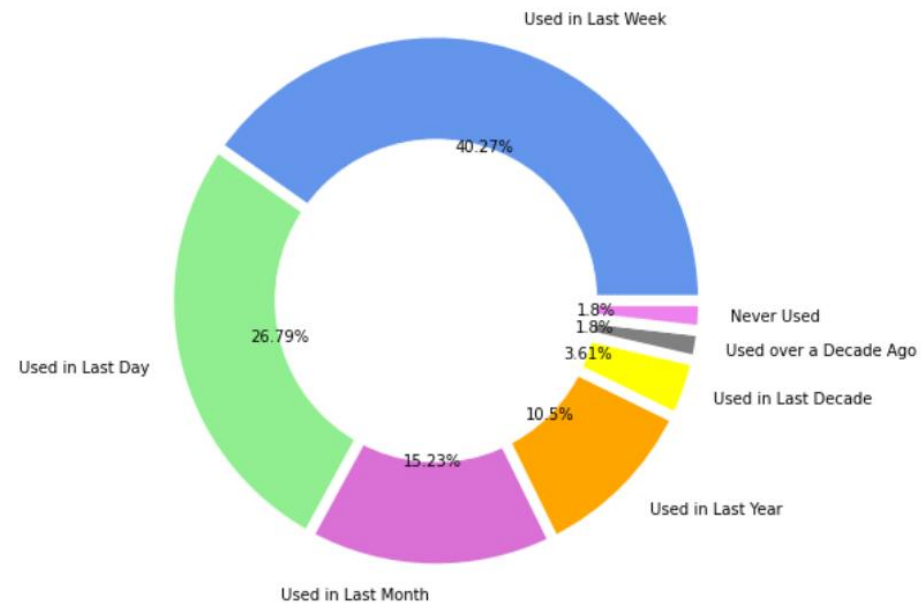


Consommation en fonction de l'alcool

Mean drugs consumption vs Alcohol

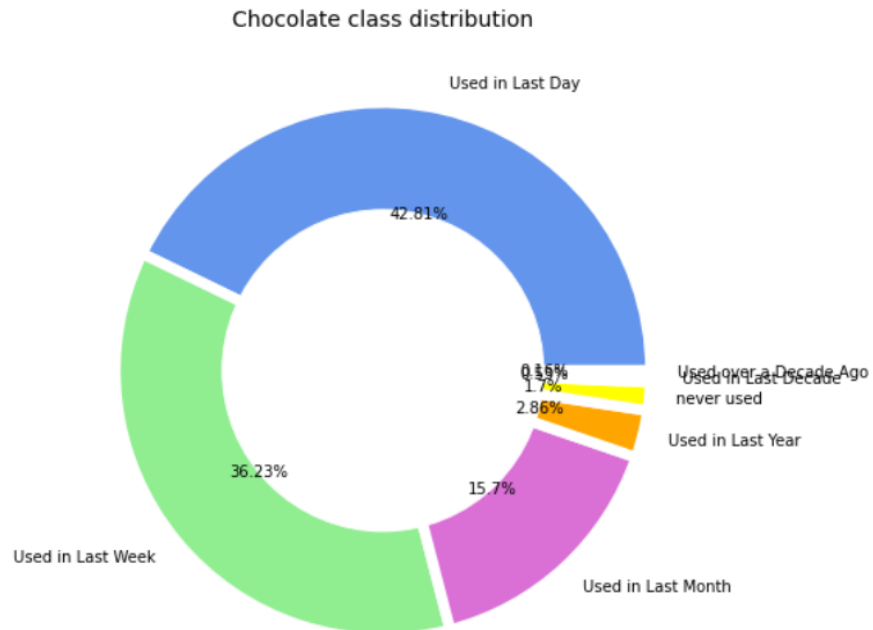
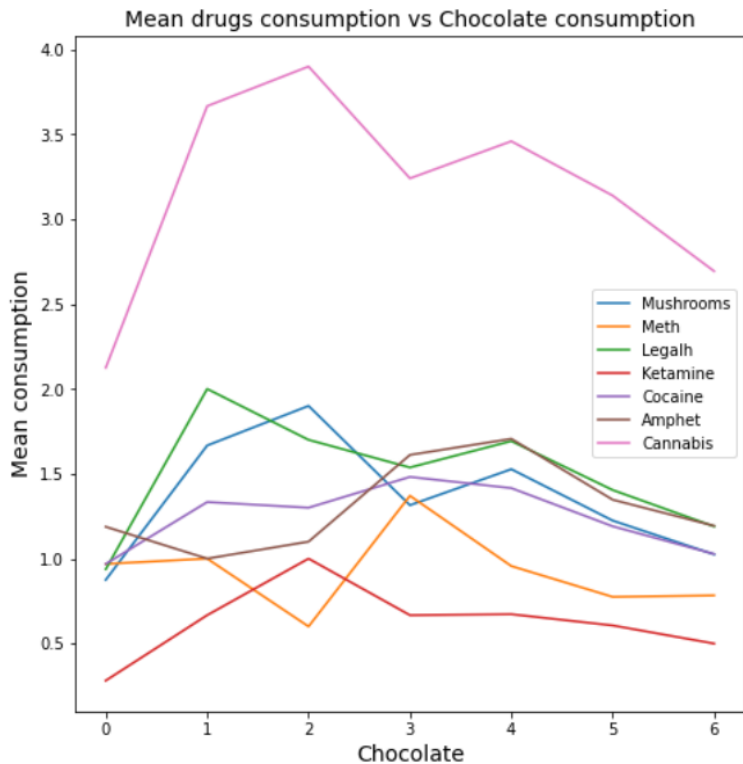


Proportion of alcohol consumption of surveyed drinkers



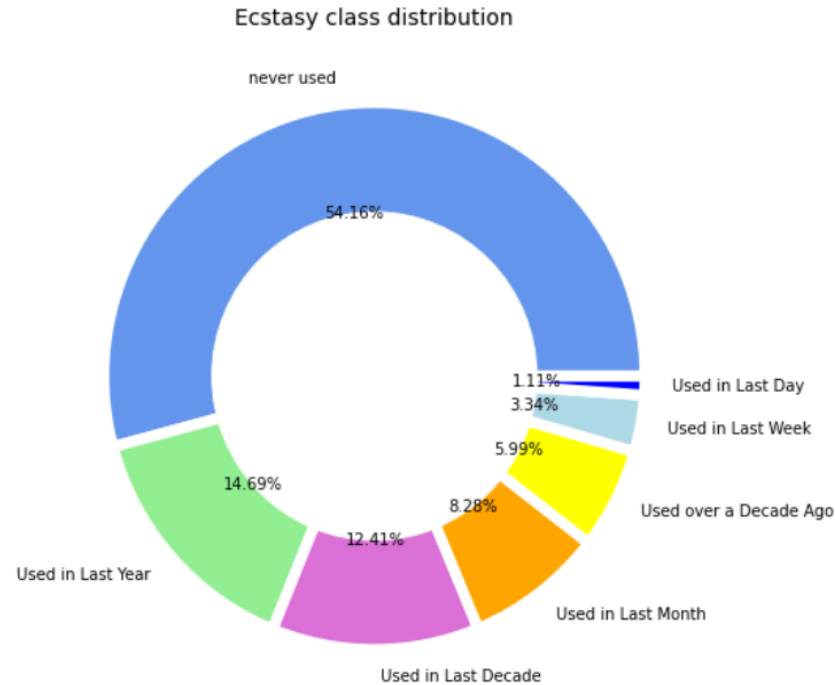
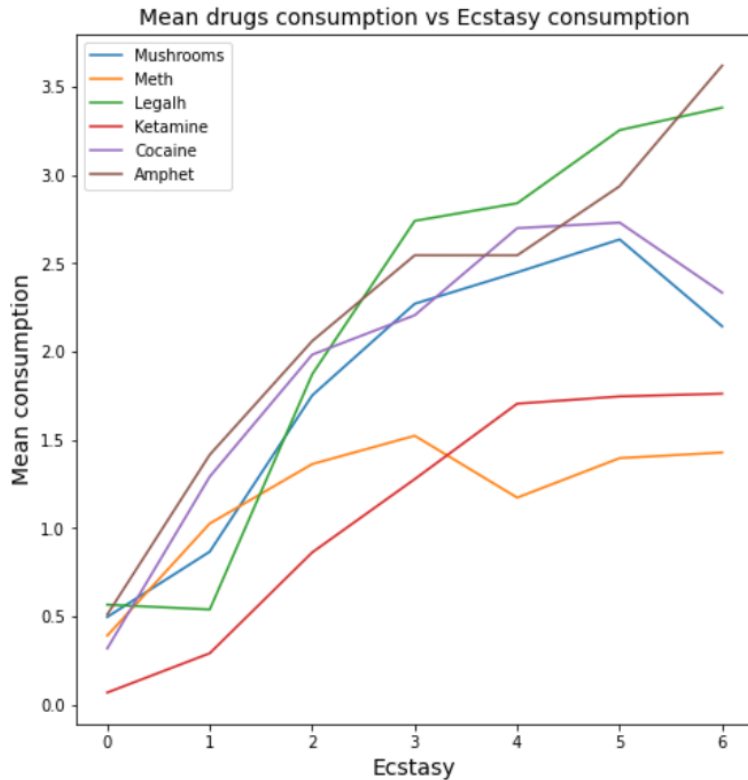
- On avait remarqué des liens entre consommation de certaines drogues dans le heat map, on souhaite donc explorer ces corrélations.
- On remarque ici que plus de 67% des personnes interrogées ont consommé de l'alcool au cours de cette dernière semaine

Consommation en fonction du chocolat



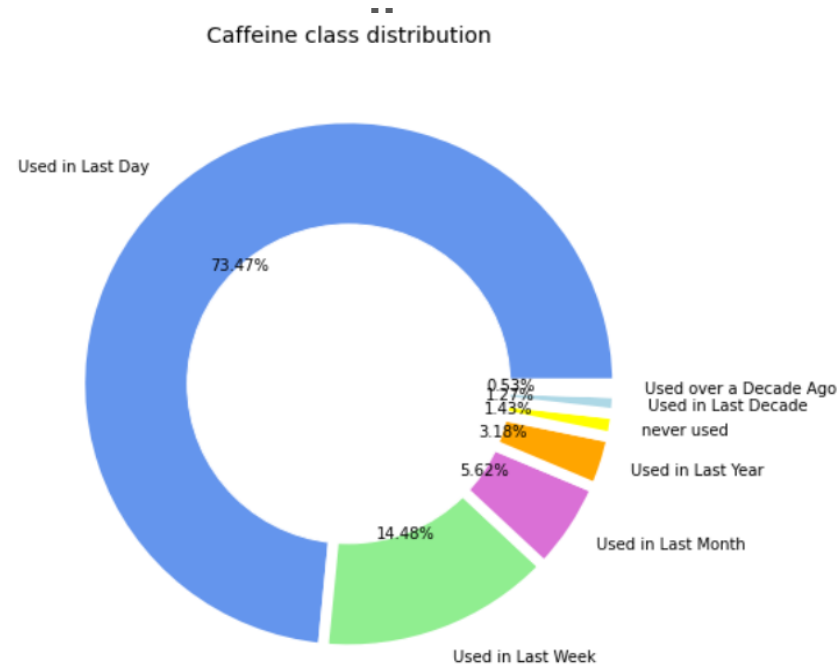
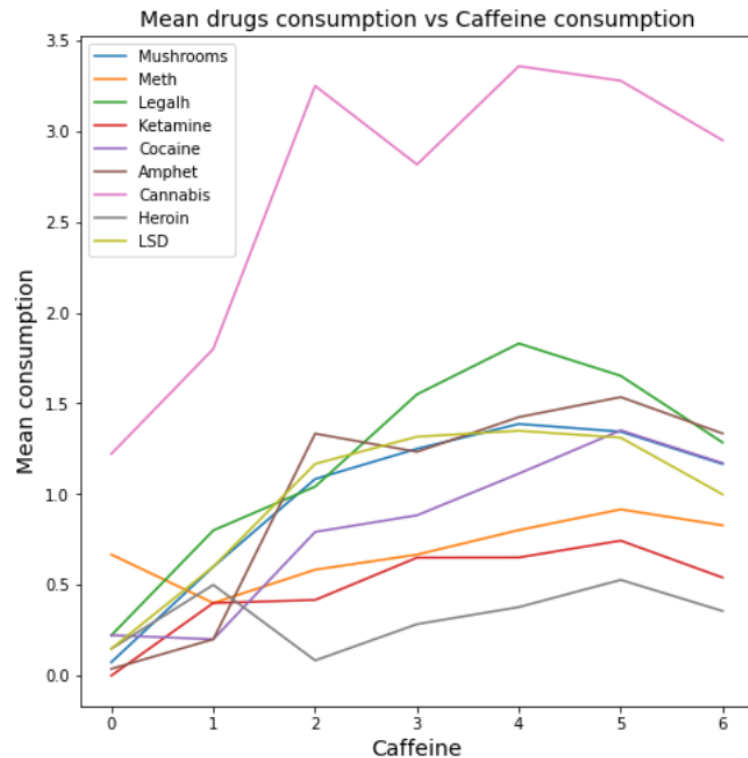
- Corrélation négative entre la consommation de chocolat et la consommation de drogue
- Plus les personnes interrogées mangent du chocolat, moins elles consomment de drogue.
- En moyenne, les personnes qui ont mangé du chocolat la veille ont consommé de la drogue dans la décennie. On peut être expliquer ce comportement par le fait que le chocolat selon plusieurs études libère de la dopamine dans le cerveau, permettant de se sentir plus heureux et donc de s'écarter d'autres drogues

Consommation en fonction de l'Ecstasy



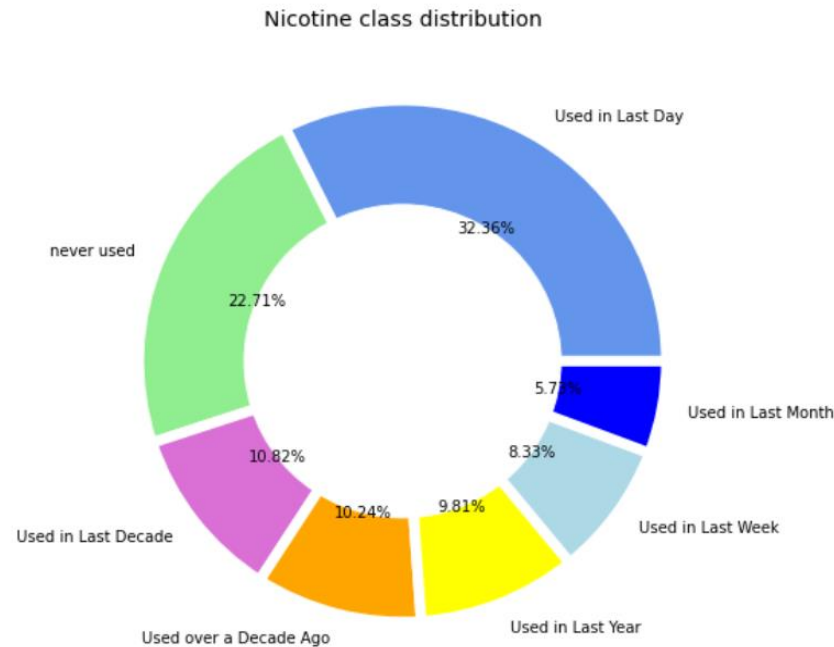
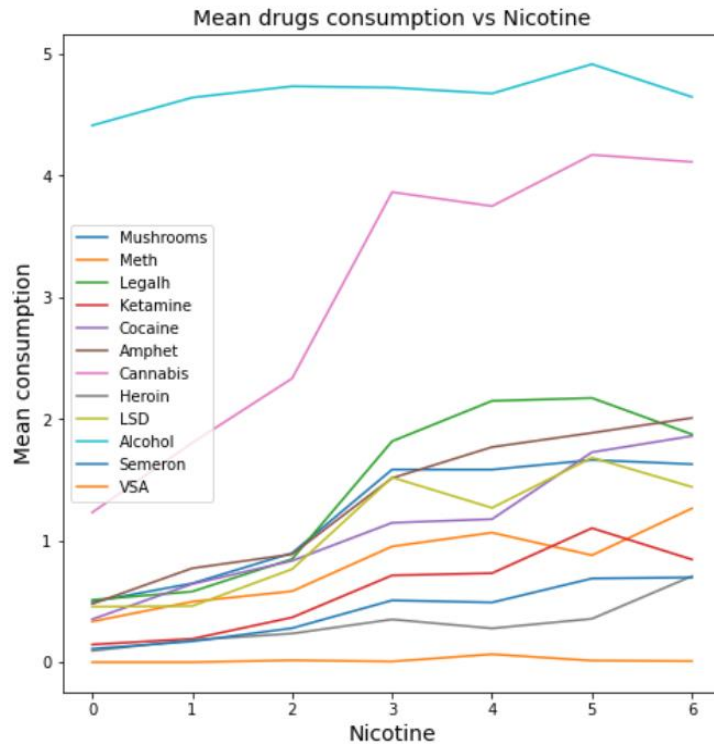
- La grande majorité (68%) des personnes interrogées n'ont jamais utilisé d'Ecstasy ou alors l'année dernière.
- On remarque que plus la consommation est forte (consommé dans les 24h, dans la semaine ou dans le mois) plus la consommation de d'autres drogues suivent cette tendance.

Consommation en fonction de Caffeine

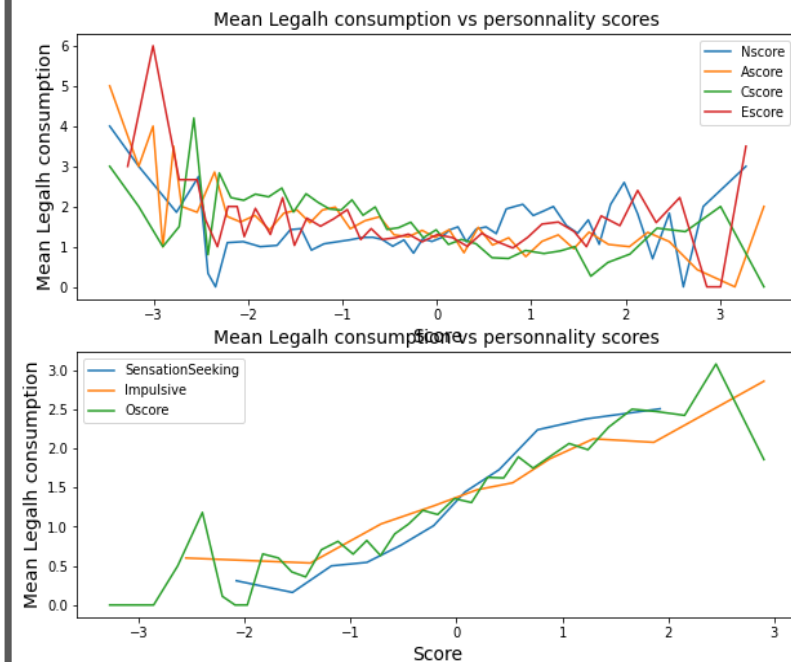
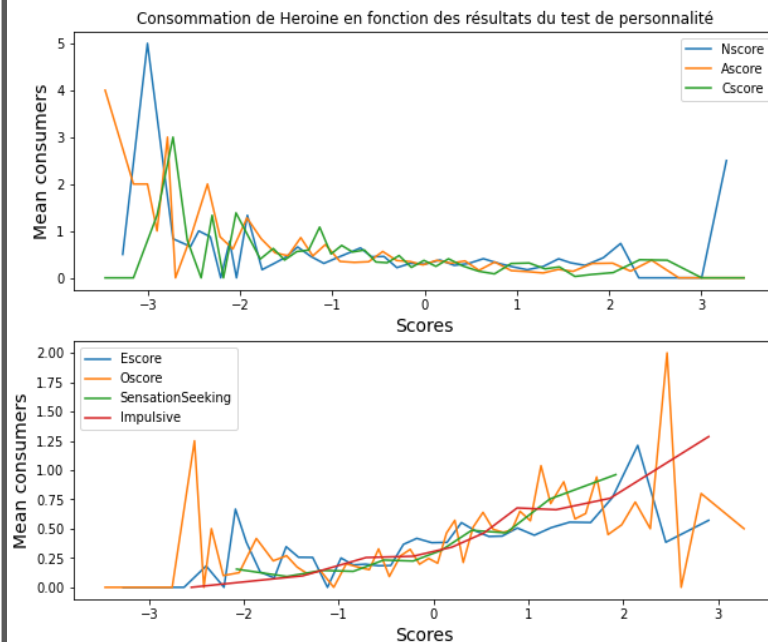
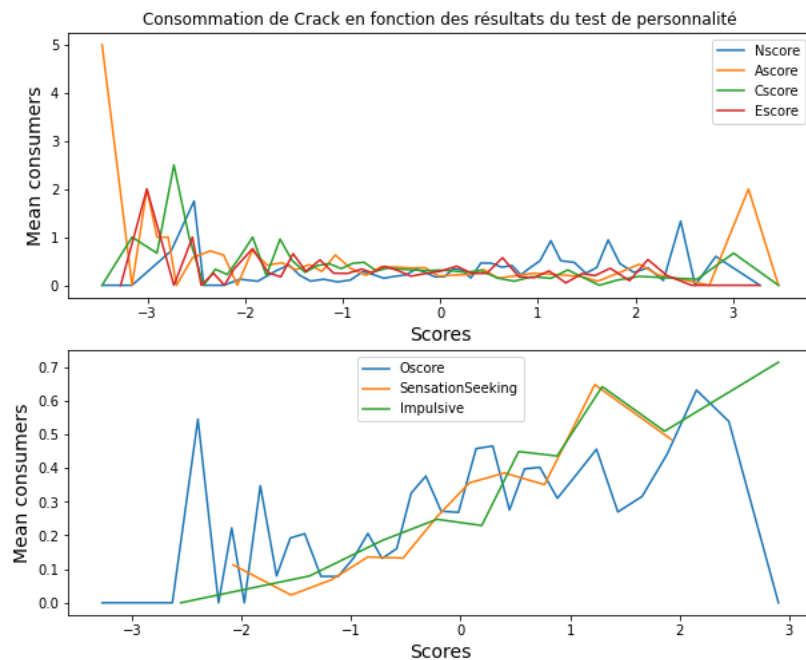


- La grande majorité (87%) des personnes interrogées ont consommé de la caffeine entre la semaine d'avant et le jour d'avant.
- On remarque un lien entre une consommation régulière de caféine et un risque accru de consommation de drogues.
- La drogue la plus associée à la caffeine est le cannabis.

Consommation en fonction de Nicotine



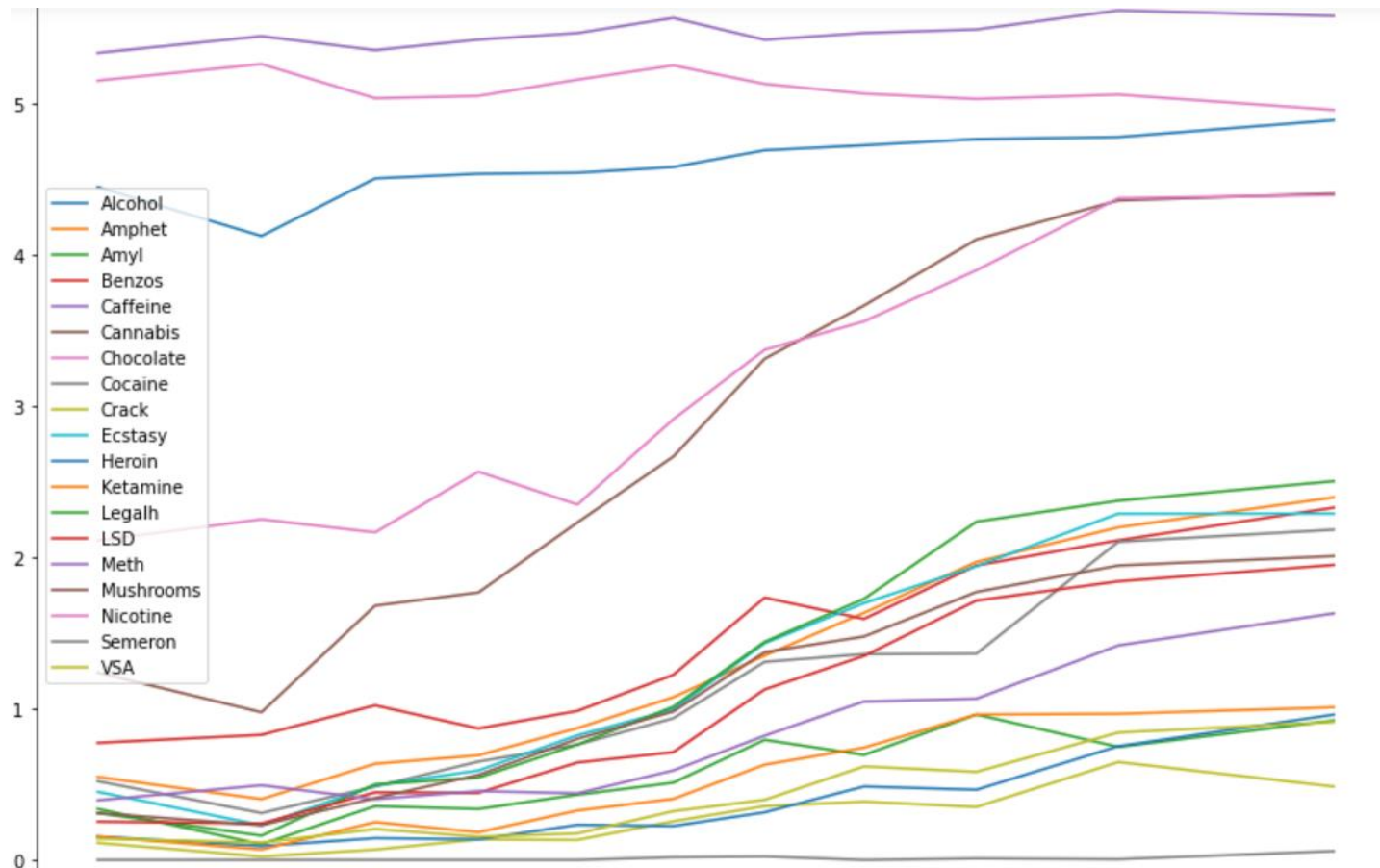
- Forte corrélation entre la nicotine et les autres drogues notamment l'alcool et le cannabis.
- La consommation est assez répartie : presque autant de personnes en ayant consommé le jour dernier que jamais.



Corrélation entre la consommation de Cannabis et les tests *NEO-FFI-R*, *BIS-11* et *ImpSS*

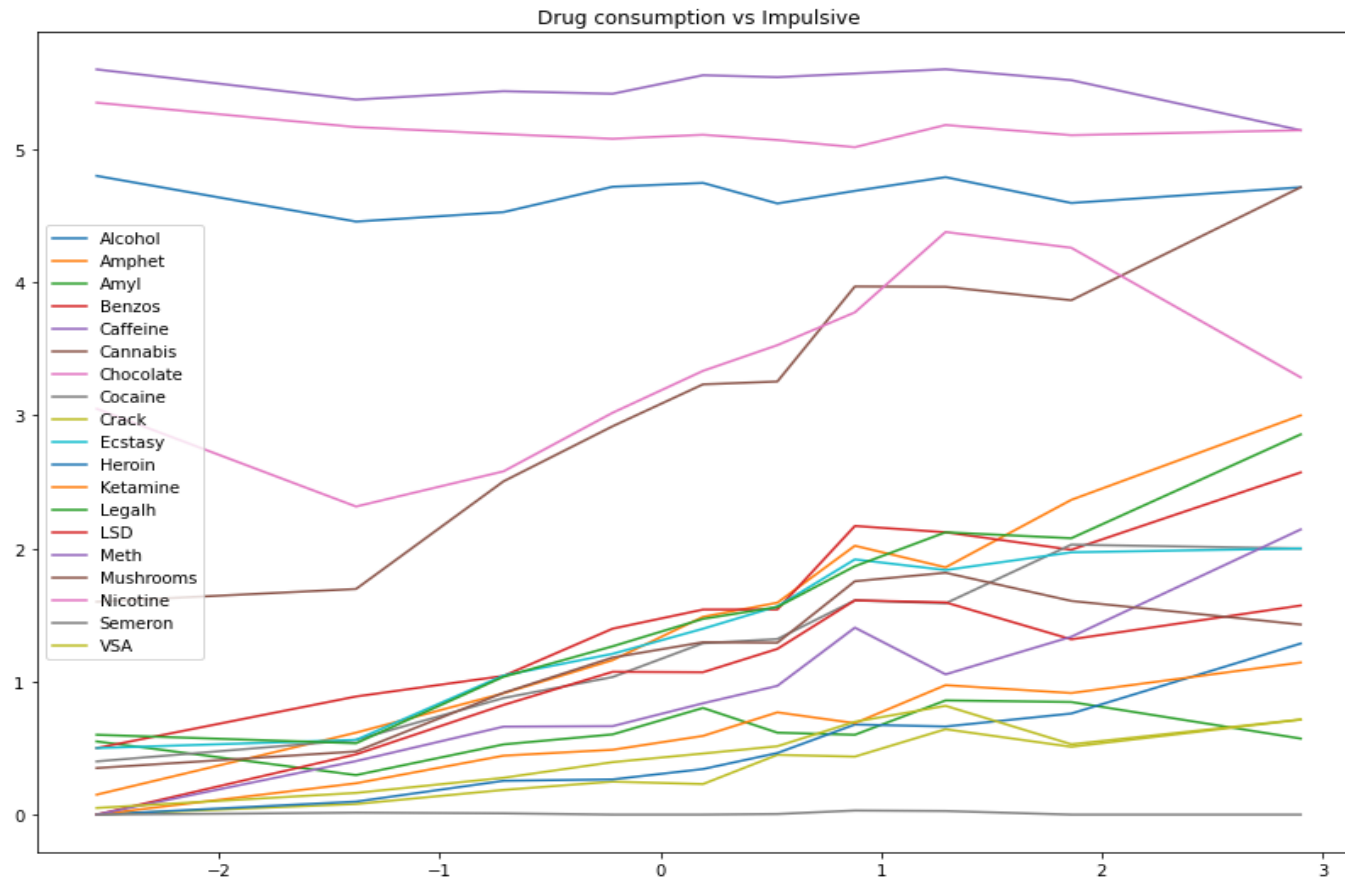
Nous avons obtenu des corrélations différentes entre consommation et trait de personnalité selon les drogues.

Consommation en fonction de la recherche d'attention



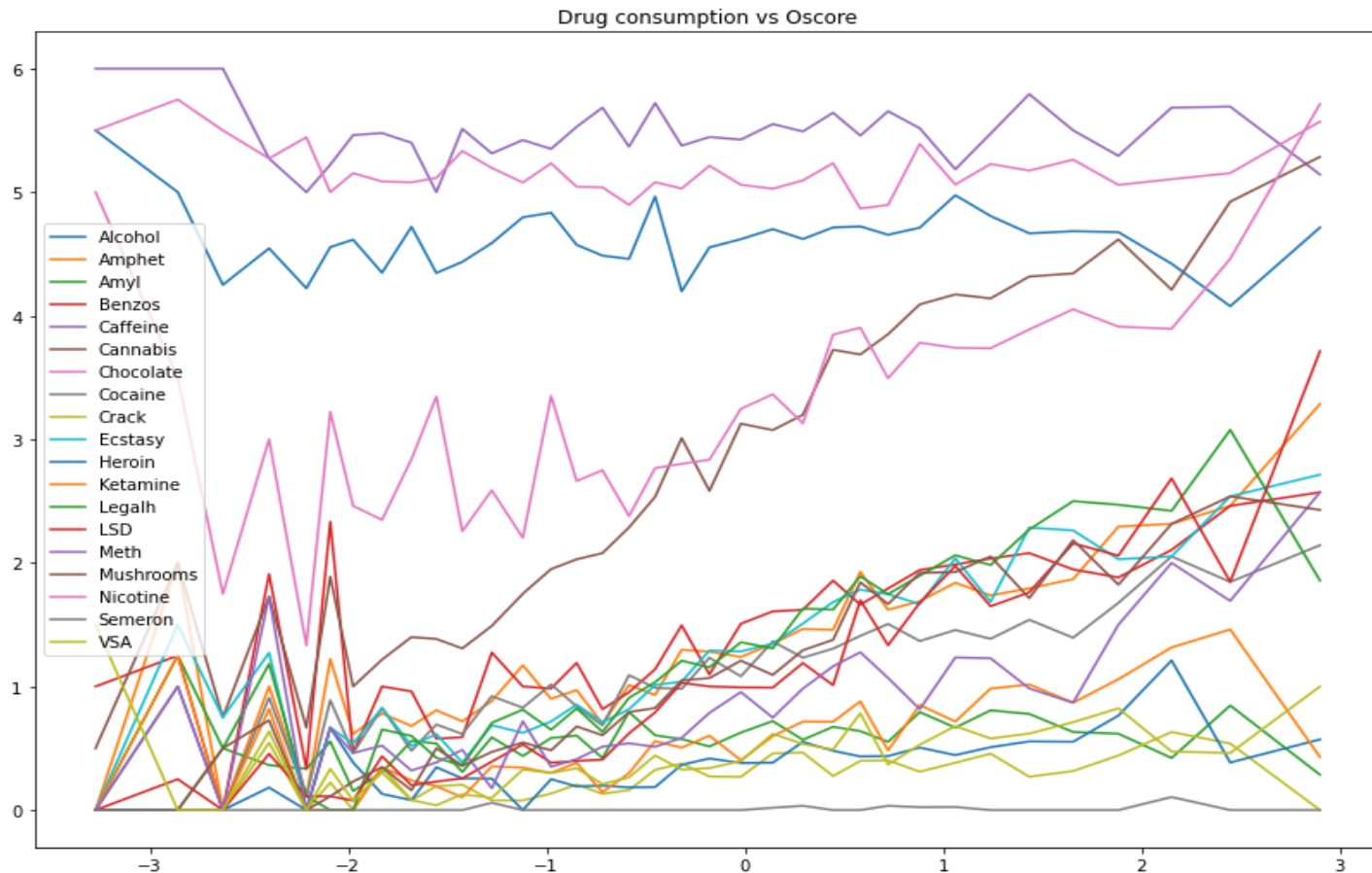
- Globalement, légère hausse des consommations de drogues quand le besoin de recherche d'attention augmente
- Sauf pour la consommation de chocolat et caféine qui reste constante
- La hausse la plus significative est pour la cocaïne.

Consommation en fonction de l'impulsivité



- Globalement, pour la plupart des drogues il y a une hausse quand l'impulsivité augmente.
- La consommation de caféine, de chocolat et d'alcool ne varie globalement pas avec l'impulsivité

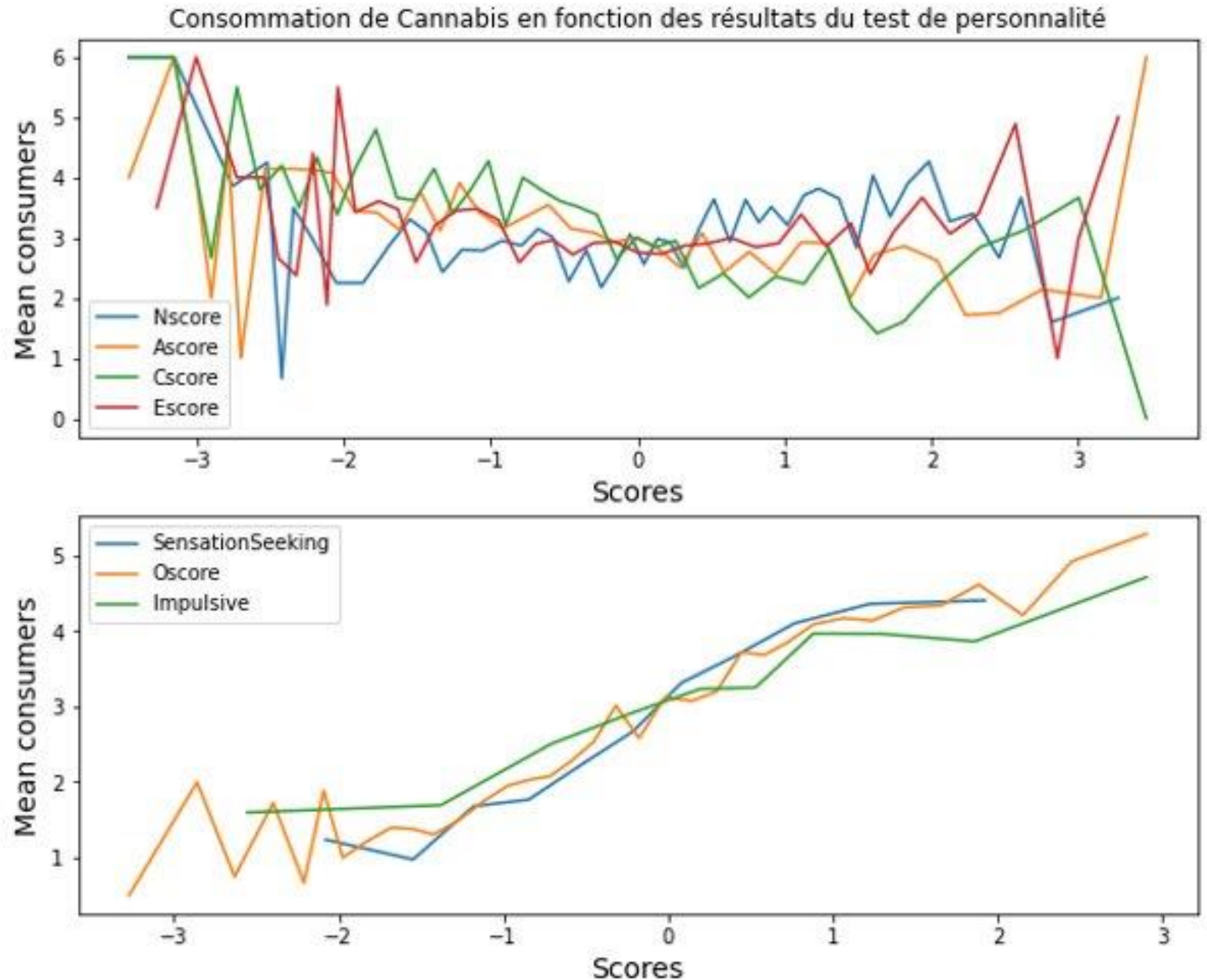
Consommation en fonction de l'Oscore



- Oscore (Ouverture): curiosité intellectuelle, indépendance de jugement, créativité.
- Il y a un lien fort entre Oscore et Consommation de Cannabis, Ecstasy, Legalh, Ketamine, Meth et plusieurs autres drogues
- Il n'a que très peu d'impact sur le sémeron

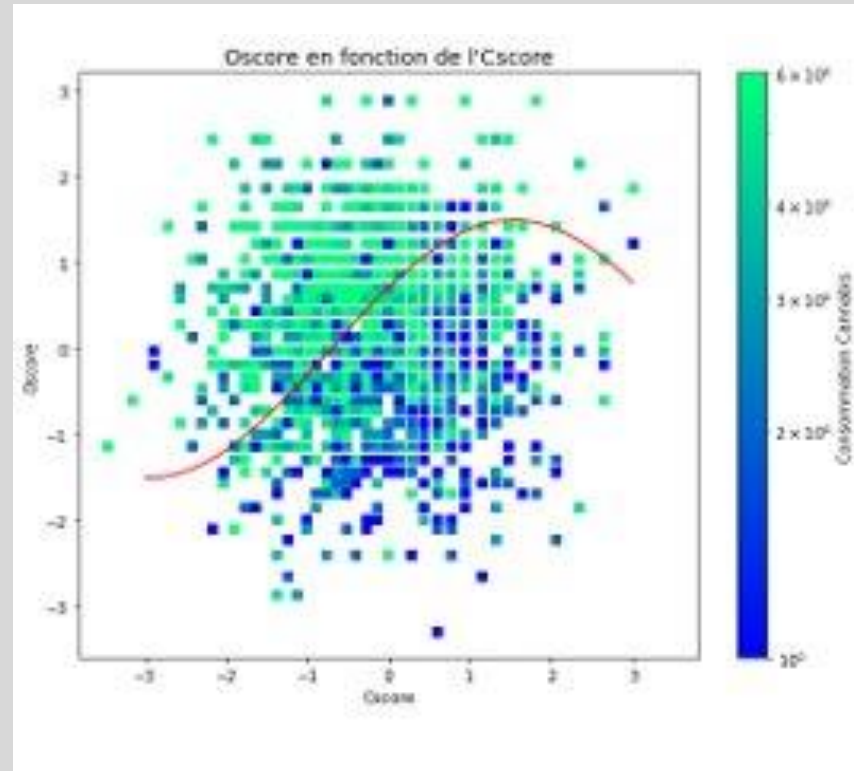
Corrélation entre la consommation de Cannabis et les tests **NEO-FFI-R**, **BIS-11** et **ImpSS**

- Nous avons divisé le graphe en deux sous-graphe afin de trouver des corrélations.
- Corrélation positive flagrante entre SensationSeeking, Oscore Et Impulsive pour le cannabis



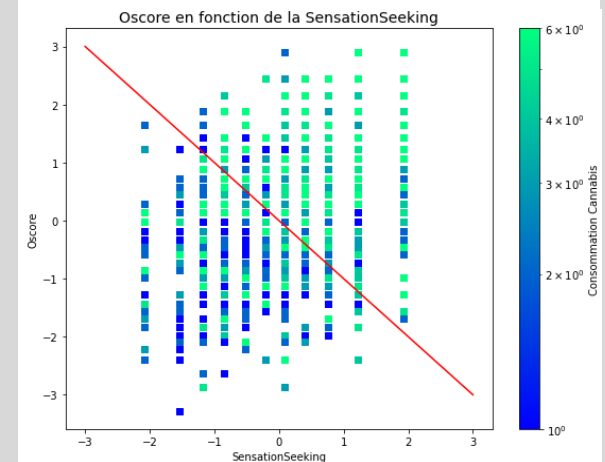
Mise en évidence d'une classification

Nous avons voulu mettre en évidence la classification possible des consommateurs de Cannabis en fonction de leurs traits de personnalité



Oscore et
SensationSeeking
en fonction de la
consommation
de Cannabis

Oscore et Cscore en
fonction de la
consommation de
Cannabis




5) Modèles de prédiction



Classification 1 : Consommation d'alcool

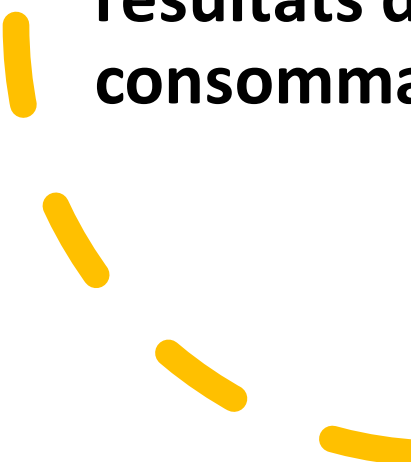
Pour pouvoir réaliser un modèle de deep learning avec Scikit-learn, il faut choisir le bon modèle adapté aux données et au problème.

C'est donc un problème de **classification**.



La colonne qui sera notre cible pour la prédiction sera la colonne "Alcohol".

C'est la première substance sujette à une addiction dans le monde et il serait intéressant de faire une classification pour savoir si **l'on peut prédire la consommation d'alcool d'une personne en connaissant les résultats des tests de personnalités la concernant ainsi que sa consommation d'autres substances.**



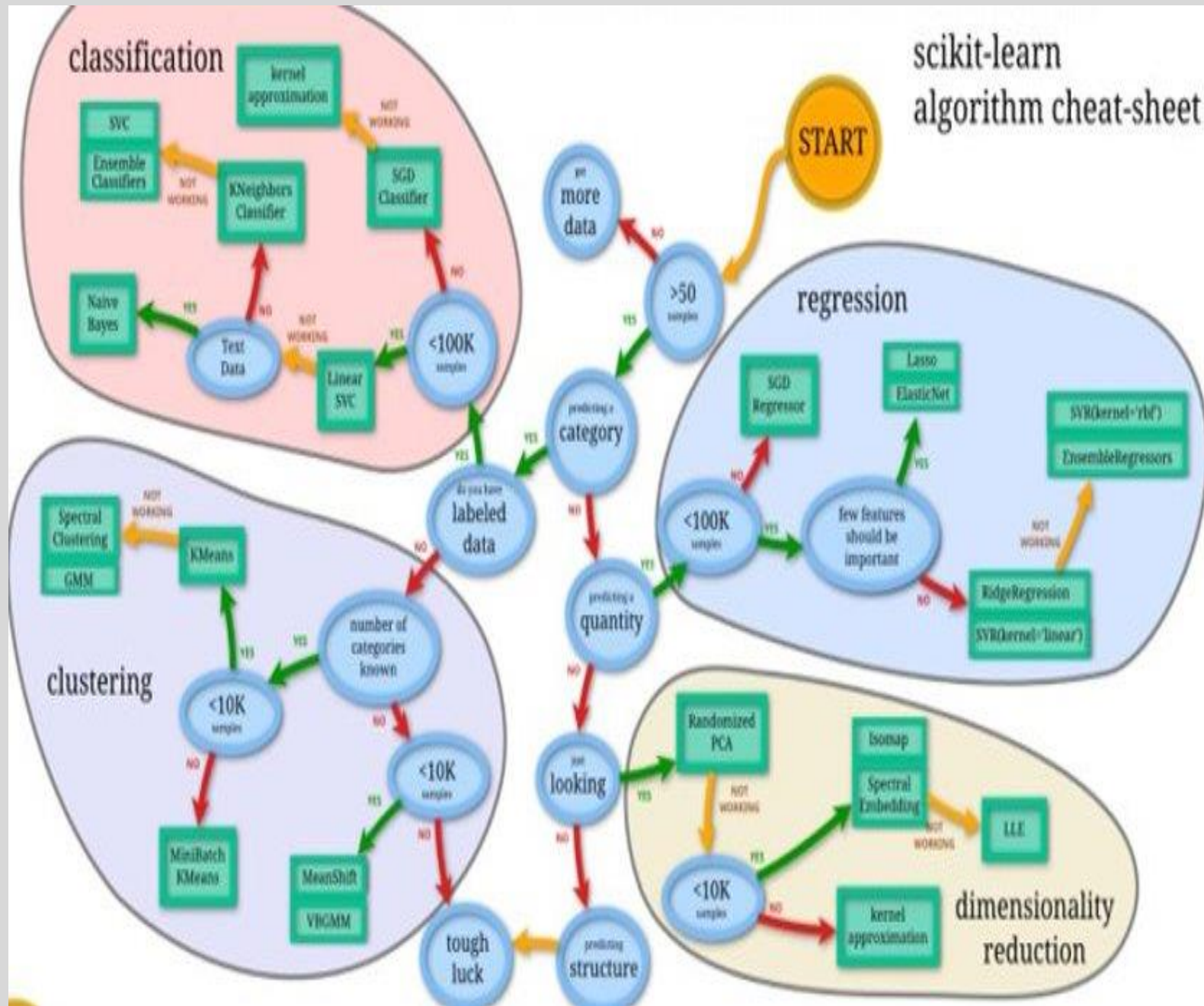
Choix des modèles

Pour choisir certains de nos modèles, nous nous sommes appuyés sur le schéma ci-contre trouvé sur le site de scikit-learn.

Ce schéma propose quelques modèles adaptés à notre taille de données, à leur nature etc...

Nous réaliserons les modèles suivants :

- KNN
- SVC
- Logistic Regression
- Random Forest
- Decision Tree



Résultats du modèle Logistic Regression

	precision	recall	f1-score	support
0	0.00	0.00	0.00	7
1	0.00	0.00	0.00	7
2	0.00	0.00	0.00	14
3	0.25	0.08	0.12	39
4	0.14	0.02	0.03	57
5	0.42	0.86	0.56	152
6	0.26	0.10	0.14	101
accuracy			0.38	377
macro avg	0.15	0.15	0.12	377
weighted avg	0.29	0.38	0.28	377

- On peut voir la précision pour chaque classe de consommation de 0 à 6.
- La précision globale du modèle est de 38 %.

La précision (accuracy) de la méthode Logistic est 0.38.

Résultats du modèle KNN

	precision	recall	f1-score	support
0	0.00	0.00	0.00	7
1	0.00	0.00	0.00	7
2	0.12	0.14	0.13	14
3	0.17	0.15	0.16	39
4	0.23	0.21	0.22	57
5	0.44	0.59	0.50	152
6	0.25	0.16	0.20	101
accuracy			0.33	377
macro avg	0.17	0.18	0.17	377
weighted avg	0.30	0.33	0.31	377

- On peut voir la précision pour chaque classe de consommation de 0 à 6.
- La précision globale du modèle est de 33 %.

La précision (accuracy) de la méthode KNN est 0.33.

Résultats du modèle Random Forest

	precision	recall	f1-score	support
0	0.00	0.00	0.00	7
1	0.00	0.00	0.00	7
2	0.00	0.00	0.00	14
3	0.25	0.03	0.05	39
4	0.14	0.02	0.03	57
5	0.41	0.74	0.53	152
6	0.27	0.25	0.26	101
accuracy			0.37	377
macro avg	0.15	0.15	0.12	377
weighted avg	0.29	0.37	0.29	377

- On peut voir la précision pour chaque classe de consommation de 0 à 6.
- La précision globale du modèle est de 37 %.

La précision (accuracy) de la méthode Random Forest est 0.37.

Résultats du modèle SVC

	precision	recall	f1-score	support
0	0.00	0.00	0.00	7
1	0.00	0.00	0.00	7
2	0.00	0.00	0.00	14
3	0.00	0.00	0.00	39
4	0.00	0.00	0.00	57
5	0.40	1.00	0.57	152
6	0.00	0.00	0.00	101
accuracy			0.40	377
macro avg	0.06	0.14	0.08	377
weighted avg	0.16	0.40	0.23	377

- On peut voir la précision pour chaque classe de consommation de 0 à 6.
- La précision globale du modèle est de 40 %.

La précision (accuracy) de la méthode SVC est 0.40.

Résultats du modèle Decision Tree

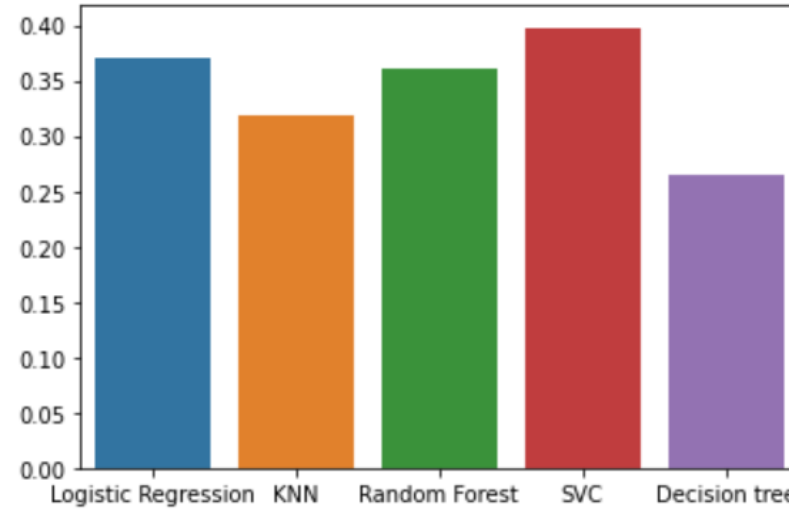
	precision	recall	f1-score	support
0	0.00	0.00	0.00	7
1	0.00	0.00	0.00	7
2	0.07	0.07	0.07	14
3	0.16	0.15	0.16	39
4	0.22	0.19	0.21	57
5	0.37	0.37	0.37	152
6	0.28	0.31	0.30	101
accuracy			0.28	377
macro avg	0.16	0.16	0.16	377
weighted avg	0.28	0.28	0.28	377

- On peut voir la précision pour chaque classe de consommation de 0 à 6.
- La précision globale du modèle est de 28 %.

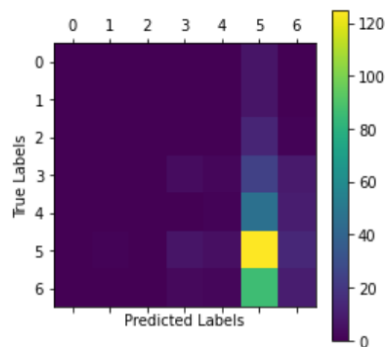
La précision (accuracy) de la méthode Decision Tree est 0.28.

Performances et matrices de confusion

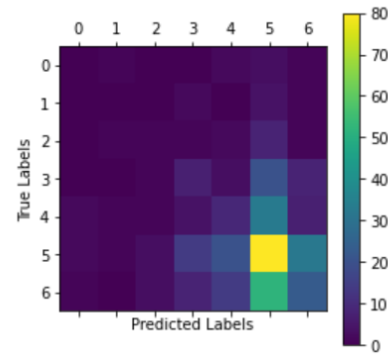
Performance en fonction du modèle



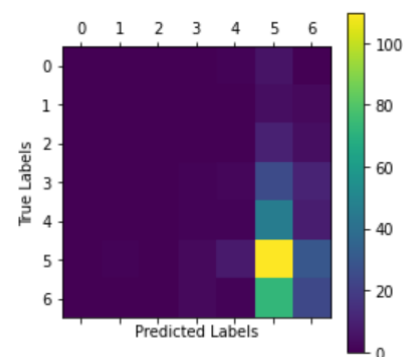
Confusion matrix of Logistic Regression (Accuracy of 0.37)



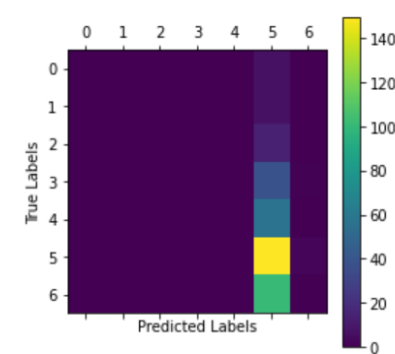
Confusion matrix of KNearest Neighbors (Accuracy of 0.32)



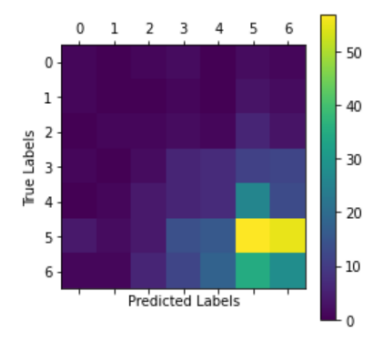
Confusion matrix of Random Forest (Accuracy of 0.36)



Confusion matrix of SVC Méthod (Accuracy of 0.40)



Confusion matrix of Decision Tree (Accuracy of 0.27)



Remarques sur les résultats

- Le modèle avec 6 classes donne une précision assez faible
- Le modèle qui donne la meilleure précision est SCV mais il est suivi de très près par Logistic Regression
- Nous avons fait varier les hyper-paramètres des modèles KNN, SVR et Logistic Regression pour trouver des valeurs optimales car il y a beaucoup de possibilités de paramètres qui peuvent donner des précisions différentes.

Les résultats sont dans le Notebook Jupyter.

- Finalement, il est difficile de prévoir quelle est la consommation d'alcool d'une personne basée sur ces informations.

Classification 2 :
consommateur de
Cannabis ou non ?

Nous allons maintenant faire une classification binaire parmi :

- une classe *Consommateur* de Cannabis
- une classe *Non Consommateur* de Cannabis

On fera la classification comme suggéré sur le site UCI :

- La personne n'a jamais consommé de Cannabis ou elle en a consommé il y a plus de 10 ans -> 0 (Non Consommateur)
- Toutes les autres possibilités : -> 1 (Consommateur)

On va maintenant voir si on peut prédire la consommation de Cannabis grâce aux données sur la personnalité

Résultats du modèle Logistic Regression

	precision	recall	f1-score	support
0	0.72	0.49	0.58	124
1	0.78	0.91	0.84	253
accuracy			0.77	377
macro avg	0.75	0.70	0.71	377
weighted avg	0.76	0.77	0.76	377

- On peut voir la précision pour les deux classes
- La précision globale du modèle est de 77 %, ce qui est bien meilleur que la classification précédente

La précision (accuracy) de la méthode Logistic est 0.77.

Résultats du modèle KNN

	precision	recall	f1-score	support
0	0.59	0.50	0.54	124
1	0.77	0.83	0.80	253
accuracy			0.72	377
macro avg	0.68	0.67	0.67	377
weighted avg	0.71	0.72	0.71	377

- On peut voir la précision pour les deux classes
- La précision globale du modèle est de 72 %

La précision (accuracy) de la méthode KNN est 0.72.

Résultats du modèle Random Forest

	precision	recall	f1-score	support
0	0.61	0.51	0.55	124
1	0.78	0.84	0.81	253
accuracy			0.73	377
macro avg	0.69	0.67	0.68	377
weighted avg	0.72	0.73	0.72	377

- On peut voir la précision pour les deux classes
- La précision globale du modèle est de 73 %

La précision (accuracy) de la méthode Random Forest est 0.73.

Résultats du modèle SVC

	precision	recall	f1-score	support
0	0.65	0.52	0.58	124
1	0.79	0.86	0.82	253
accuracy			0.75	377
macro avg	0.72	0.69	0.70	377
weighted avg	0.74	0.75	0.74	377

On peut voir la précision pour les deux classes.

La précision globale du modèle est de 75 %

La précision (accuracy) de la méthode SVC est 0.75.

Résultats du modèle Decision Tree

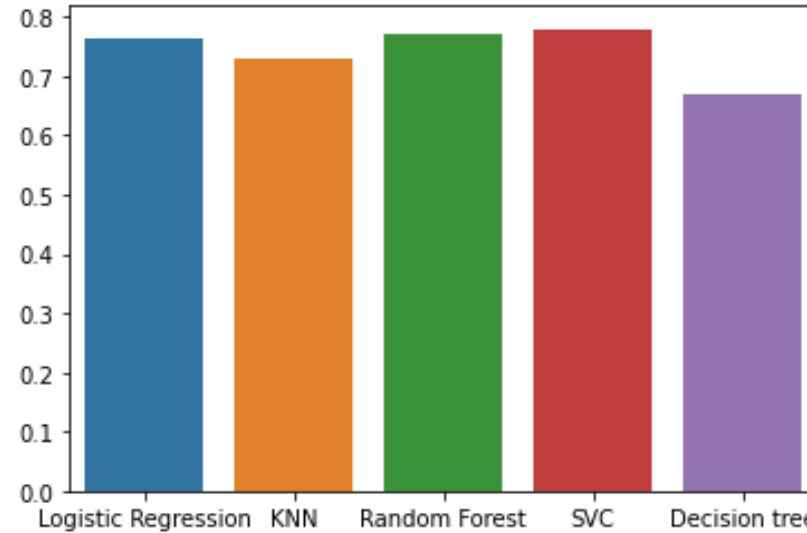
	precision	recall	f1-score	support
0	0.50	0.56	0.53	124
1	0.77	0.73	0.75	253
accuracy			0.67	377
macro avg	0.64	0.65	0.64	377
weighted avg	0.68	0.67	0.68	377

- On peut voir la précision pour les deux classes.
- La précision globale du modèle est de 67 %

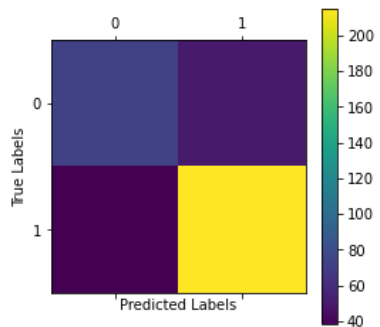
La précision (accuracy) de la méthode Decision Tree est 0.67.

Performances et matrices de confusion

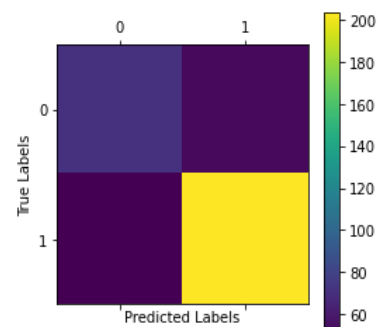
Performance en fonction du modèle



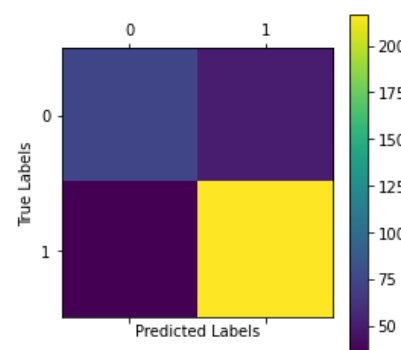
Confusion matrix of Logistic Regression (Accuracy of 0.76)



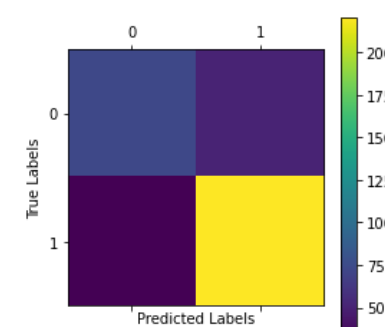
Confusion matrix of KNearest Neighbors (Accuracy of 0.73)



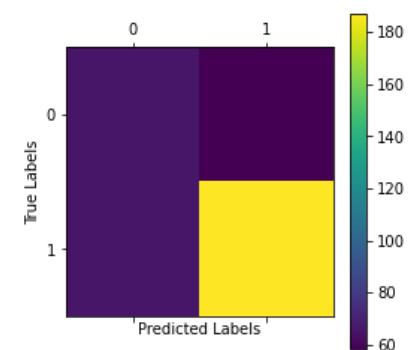
Confusion matrix of Random Forest (Accuracy of 0.77)



Confusion matrix of SVC Méthod (Accuracy of 0.78)



Confusion matrix of Decision tree (Accuracy of 0.67)



Remarques sur les résultats

- Le modèle avec 2 classes donne une précision satisfaisante.
- Les modèles qui donnent les meilleures précisions sont SVC et Random Forest et Logistic Regression.
- Nous avons fait varier les hyper-paramètres des modèles KNN, SVR et Logistic Regression et les résultats sont dans le Notebook Jupyter.
- Finalement, il est tout à fait possible de prévoir si une personne est consommatrice ou non de cannabis selon sa personnalité.

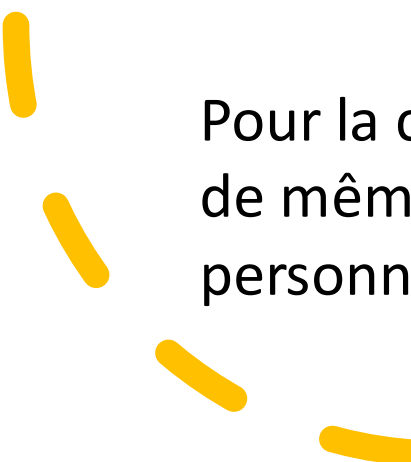
6) Regard critique sur les résultats et conclusion



Il est compliqué de savoir avec notre dataset de savoir quelle mesure influe sur quelle mesure.

Est-ce la personnalité qui augmente le risque de consommation de drogue ou est-ce la consommation de drogues qui modifie notre comportement/personnalité ?

Nous ne pouvons pas tirer de conclusion pour la plupart des minorités du dataset. Les seules données en quantité suffisante sont celles d'un Homme ou d'une femme blanche vivant au Royaume-Uni ayant entre 18 et 54 ans.

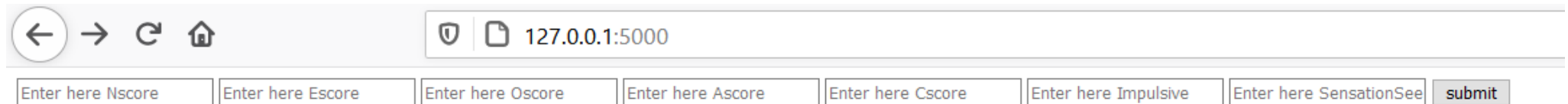


Pour la classification, nous avons pu voir qu'avec ces données nous pouvons tout de même avoir une certaine idée de la consommation de Cannabis grâce à la personnalité de la personne interrogée.

7) API Flask - modèle SVC pour
la classification 2

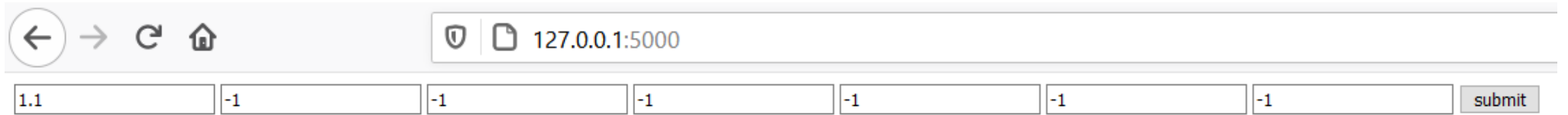
Résultats de l'API pour des valeurs de personnalités:

La personne peut saisir ses données de personnalités (Nscore, Escore, Oscore, Ascore, Cscore, Impulsive et SensationSeeking)



← → ↺ 🏠 127.0.0.1:5000

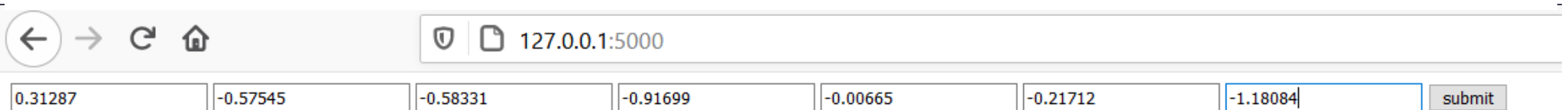
Enter here Nscore Enter here Escore Enter here Oscore Enter here Ascore Enter here Cscore Enter here Impulsive Enter here SensationSee submit



← → ↺ 🏠 127.0.0.1:5000

1.1 -1 -1 -1 -1 -1 -1 submit

The prediction is User :/



← → ↺ 🏠 127.0.0.1:5000

0.31287 -0.57545 -0.58331 -0.91699 -0.00665 -0.21712 -1.18084 submit

The prediction is Non-User :)