BA KHADY
HUYNH GIA HOA

Professeur: Monsieur Marc-Michel Corsini

RAPPORT DU PROJET D'INTELLIGENCE ARTIFICIELLE

L3 MIASHS - GROUPE 3

2022 - 2023

I. Présentation du projet

On souhaite utiliser l'algorithme Apriori qui a été implémenté dans les classes Apriori et Arules pour classifier et extraire des données utiles qui pourraient aider la prise de décision. Cet algorithme est particulièrement utile pour identifier des règles de corrélation entre différents éléments d'un ensemble de données. En utilisant l'algorithme Apriori, il est possible de déterminer les associations les plus fréquentes entre des éléments spécifiques d'un ensemble de données. Cela peut être très utile dans un certain nombre de contextes, tels que l'analyse de données marketing, la recommandation de produits et la surveillance des transactions financières.

En utilisant l'algorithme Apriori, il est possible de classifier les données en fonction de leur fréquence et de leur pertinence pour une tâche spécifique. Les données peuvent être classées en fonction de leur importance, de leur popularité et de leur pertinence par rapport à un certain objectif. Cela permet de filtrer les données inutiles et de se concentrer sur les données qui sont les plus pertinentes pour la tâche en cours.

II. Présentation des différents fichiers

Les données sont stockées dans quatre répertoires différents:

D'une part, par exemple les données sur le panier de la ménagère contenues dans les fichiers "grocery", "Bread_Basket_DMS" et "Online_retail" sont analysées en utilisant des classes d'algorithmes afin de déterminer les habitudes d'achat des consommateurs et de découvrir les produits phares mais aussi ceux qui sont le plus souvent achetés ensemble pour mieux organiser les produits dans les magasins ou faire des campagnes de marketing vers un public visé.

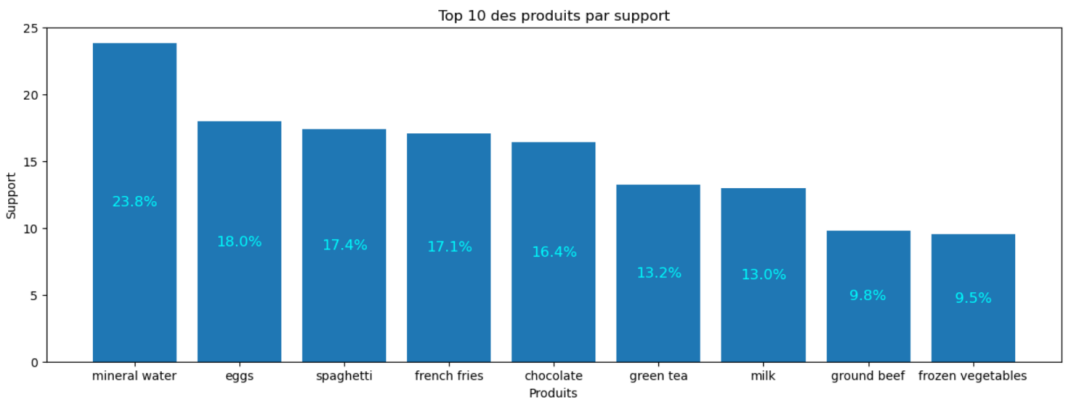
On cherchera en effet à déterminer l'influence qu'un produit peut avoir sur un autre dans ces trois fichiers.

D'autre part le fichier "mushrooms" est utilisé pour trouver des règles permettant de classifier les données et de prédire la classe d'un champignon donné. Cela permet d'identifier les caractéristiques des champignons comestibles et toxiques en fonction de leurs attributs physiques et de leurs caractéristiques. Cette analyse peut être utile pour la cueillette de champignons sauvages ou pour la production de champignons comestibles en culture.

A. Les données du fichier "grocery"

Ce fichier est constitué de 119 produits différents du quotidien dont 7901 transactions autrement dit 7901 de ventes ont effectué sur sur ces produits. On remarque cependant qu'il n'y a pas de colonnes.

Pour une bonne visibilité des données car celles-ci étant nombreuses, on a décidé de ne prendre en compte que ceux qui sont dans le top 10 du classement selon leurs fréquences d'achat (valeur de support) ce qui pourra d'ailleurs nous permettre de faire ultérieurement des analyses descriptives en amont avant tout autre traitement.



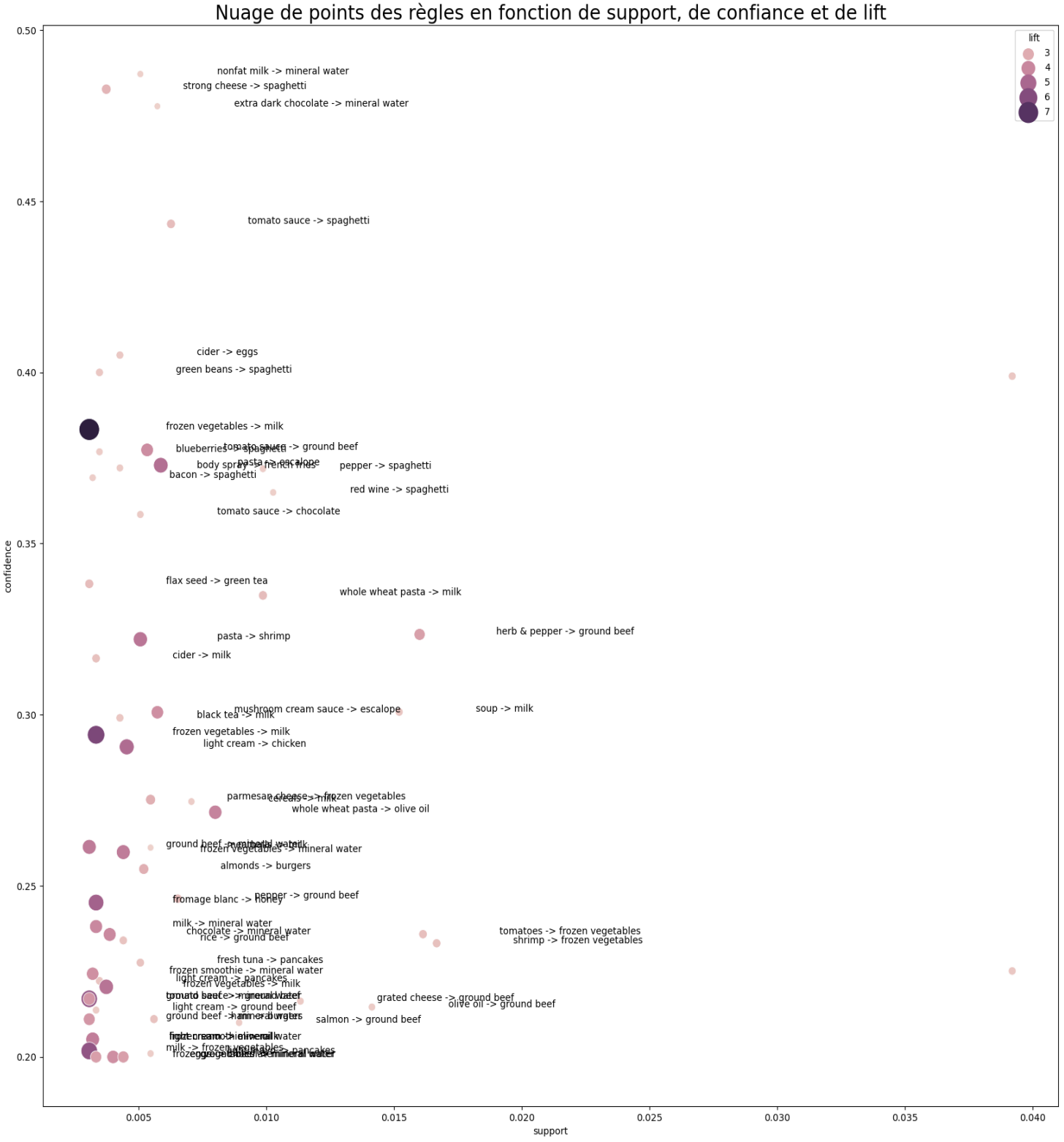
On remarque que l'eau minérale occupe 23,8% des transactions suivie des œufs(18 %), spaghetti (17.4%), french fries (17.1%). Les légumes congelées "frozen vegetables" arrivent en dernière place avec un pourcentage de 9.5%
En effet, on a filtré les données grâce à l'algorithme "apriori" avec min_support=0.003 (0.3% de 7501) qui signifie que le produit doit être présent dans au moins 22 transactions sur 7501 uniquement si nous considérons cet produit dans des produits fréquentées (qui ont la valeur de support supérieur à 0.003).

	lhs	rhs	lhs_support	rhs_support	support	confidence	lift	leverage	conviction	Rules
0	(1,)	(15,)	0.020397	0.087188	0.005199	0.254902	2.923577	0.003421	1.225089	almonds -> burgers
1	(1,)	(25,)	0.020397	0.163845	0.005999	0.294118	1.795099	0.002657	1.184553	almonds -> chocolate
2	(1,)	(37,)	0.020397	0.179709	0.006532	0.320261	1.782108	0.002867	1.206774	almonds -> eggs
3	(1,)	(43,)	0.020397	0.170911	0.004399	0.215686	1.261983	0.000913	1.057089	almonds -> french fries
4	(1,)	(54,)	0.020397	0.132116	0.005066	0.248366	1.879913	0.002371	1.154663	almonds -> green tea
...	...	...	...	...	...	...	...	...	...	...
300	(55, 71)	(72, 100)	0.021997	0.059725	0.004399	0.200000	3.348661	0.003086	1.175343	ground beef -> mineral water
301	(55, 81)	(72, 100)	0.014131	0.059725	0.003066	0.216981	3.632981	0.002222	1.200833	ground beef -> mineral water
302	(55, 82)	(72, 100)	0.014531	0.059725	0.003066	0.211009	3.532991	0.002198	1.191743	ground beef -> mineral water
303	(55, 108)	(72, 100)	0.011732	0.059725	0.003066	0.261364	4.376091	0.002366	1.272987	ground beef -> mineral water
304	(71, 108)	(72, 100)	0.013998	0.059725	0.003333	0.238095	3.986501	0.002497	1.234110	milk -> mineral water

305 rows × 10 columns

Ensuite on établit les règles à partir de l'ensemble des données filtrées
On n'utilise pas la confiance comme critère car l'une des limitations majeures de la mesure de "confidence" dans l'algorithme Apriori est que cette mesure peut être trompeuse lorsque la fréquence d'apparition d'un produit conséquent est très élevée dans l'ensemble de données. La "confiance" (ou "confidence" en anglais) est une mesure de la probabilité conditionnelle que deux produits apparaissent ensemble dans une transaction donnée, étant donné que l'un des produits apparaît également dans cette transaction. Plus précisément, la confiance d'une règle d'association A => B est définie comme le nombre de transactions contenant à la fois A et B, divisé par le nombre de transactions contenant A. Cela permet de mesurer à quel point l'apparition de A est corrélée avec celle de B.
Par exemple, on sait que "eau minérale" est très populaire dans une épicerie et qu'il apparaît dans 23.8 % des transactions. Si une règle d'association lait ⇒ eau minérale a une confidence d'environ 49 %, c'est-à-dire qu'il y a 49 % de chance d'acheter de l'eau minérale si on prend du lait. Cela pourrait sembler être une forte association entre les deux produits.Alors qu'en réalité, la plupart des transactions contenant l'eau minérale contiennent également "lait" simplement parce que l'eau minérale est présente dans 23.8% des transactions. Ainsi, la règle d'association lait ⇒ eau minérale n'a peut-être pas beaucoup de valeur prédictive et pourrait ne pas être très utile pour les analyses de marché.
Cependant, la métrique lift nous a permis de sélectionner les règles d'association les plus significatives et les plus utiles dans l' ensemble de données. Il est calculé comme le rapport entre la confiance de la règle et la fréquence de l'élément conséquent dans l'ensemble des données. Plus précisément, il mesure la probabilité que les éléments de la règle se produisent ensemble, comparée à ce que l'on pourrait attendre s'ils étaient indépendants les uns des autres. Si le lift est supérieur à 1 cela indique qu'il y a une corrélation positive entre les deux produits, et donc la présence de l'un augmente la probabilité de la présence de l'autre; le taux d'association entre les deux produits. Si le lift est égal à 1 cela indique que les

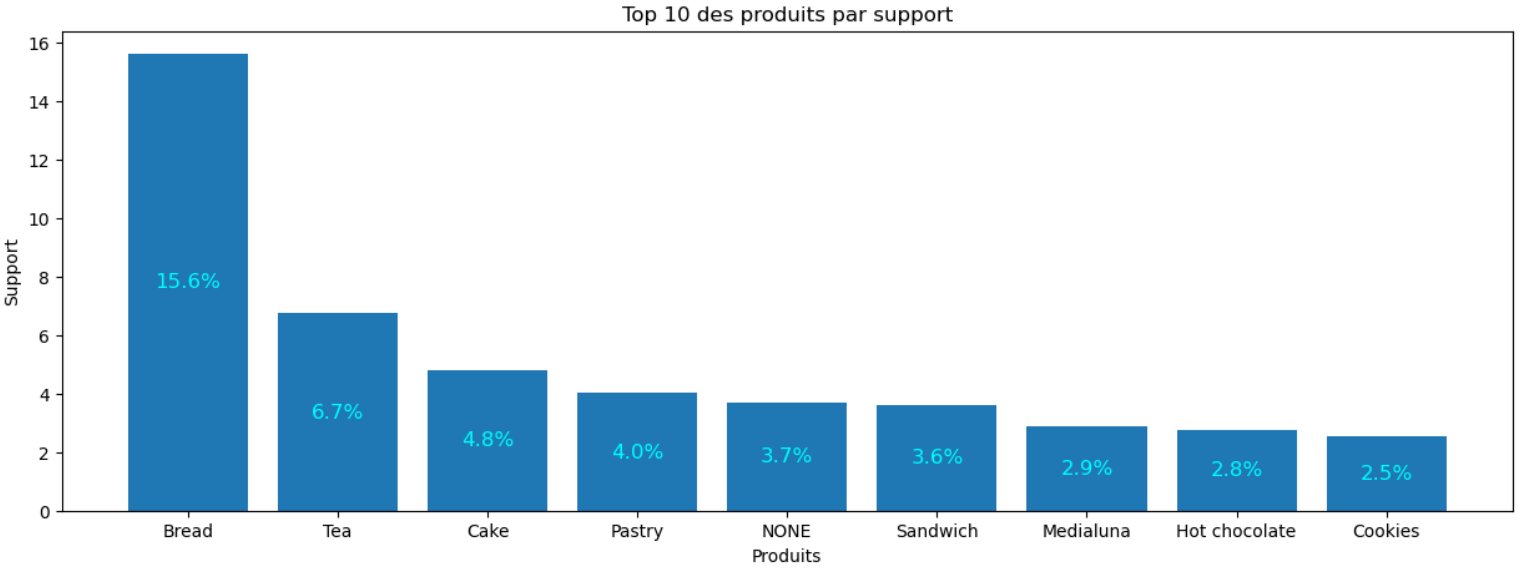
deux produits sont indépendants l'un de l'autre, tandis qu'un lift inférieur à 1 indique une corrélation négative entre les deux produits.
Pour cela on a filtré les règles d'associations dont les "lift" sont supérieurs à 4 en allant de light cream ⇒ chicken (environ 4.84) à ground beef ⇒ mineral water (environ 4.38)



Supposons que l'épicerie achète trop de sauce tomate et qu'elle craint que ses stocks s'épuisent si elle n'arrive pas à les vendre à temps. De plus, la marge bénéficiaire de la sauce tomate est si faible que l'épicerie ne peut se permettre d'accorder une remise promotionnelle sans réduire ses bénéfices. Une approche qui peut être proposée consiste alors à déterminer les produits qui stimulent les ventes de la sauce tomate ici c'est du bœuf haché et à offrir des réductions sur ces produits à la place.
Donc l'organisation la plus appropriée serait de mettre les produits qui sont le plus souvent achetés ensemble dans des rayons plus ou moins proches pour augmenter la visibilité.

B. Les données du fichier "Bread_Basket_DMS"

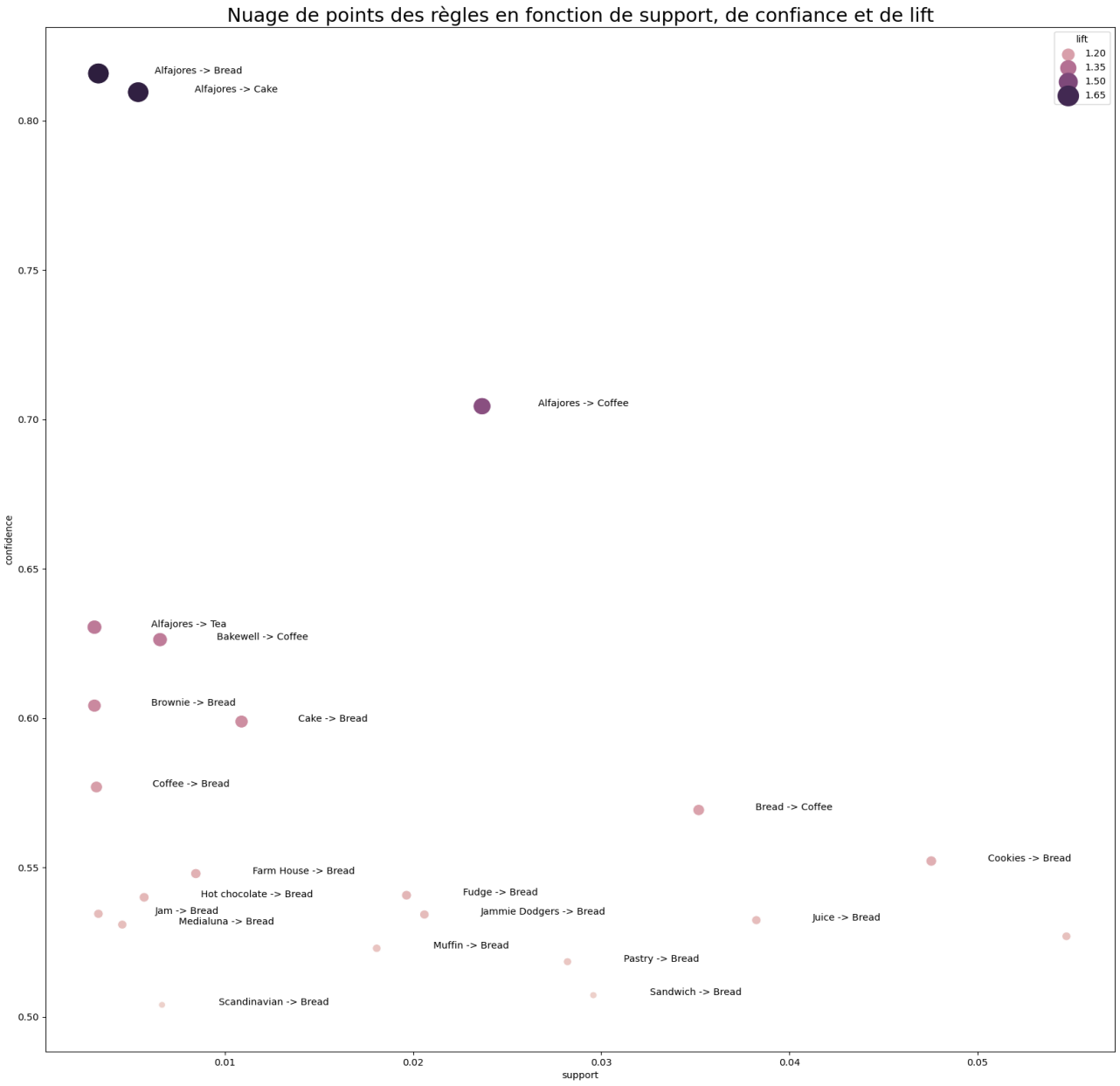
Ce fichier contient les informations sur les transactions, les items (les dates et les temps des transactions constituant les deux premières colonnes ont été enlevé car elles ne sont pas intéressantes dans notre analyse) et chaque ligne ne contient qu'un seul achat.
On compte 9465 transactions sur 94 items. Comme dans le cas précédent, on choisit de représenter le top 10 des produits par support d'abord.



On remarque que Bread occupe 15.6% des transactions suivie de tea (6.7 %), "Cake" (4.8%), Pastry (4%).Les cookies arrivent en dernière place avec un pourcentage de 2.5%
En effet, on a filtré les données grâce à l'algorithme "apriori" avec min_support=0.003 (0.3% de 9465) qui signifie que le produit doit être présent dans au moins 28 transactions sur 9465 uniquement si nous considérons cet produit dans des produits fréquentées (qui ont la valeur de support supérieure à 0.003).
Cependant, on peut noter que tous les items sauf "Bread" prennent moins de 7% des transactions donc on peut filtrer les données selon les "confidence" supérieures à 0.4 autrement dit selon la règle A ⇒ B il y a 40% de chance que l'item B soit acheté si l'item A est pris d'abord.

	lhs	rhs	lhs_support	rhs_support	support	confidence	lift	leverage	conviction	Rules
0	(36,)	(24,)	0.004015	0.478394	0.003275	0.815789	1.705267	0.001355	2.831575	Alfajores -> Bread
1	(53,)	(24,)	0.006656	0.478394	0.005388	0.809524	1.692169	0.002204	2.738431	Alfajores -> Cake
2	(88,)	(24,)	0.033597	0.478394	0.023666	0.704403	1.472431	0.007593	1.764582	Alfajores -> Coffee
3	(83,)	(24,)	0.004860	0.478394	0.003064	0.630435	1.317815	0.000739	1.411404	Alfajores -> Tea
4	(73,)	(24,)	0.010460	0.478394	0.006550	0.626263	1.309094	0.001547	1.395648	Bakewell -> Coffee
5	(8,)	(24,)	0.005071	0.478394	0.003064	0.604167	1.262906	0.000638	1.317741	Brownie -> Bread
6	(80,)	(24,)	0.018172	0.478394	0.010882	0.598837	1.251766	0.002189	1.300235	Cake -> Bread
7	(93,)	(24,)	0.005494	0.478394	0.003170	0.576923	1.205958	0.000541	1.232887	Coffee -> Bread
8	(56,)	(24,)	0.061807	0.478394	0.035182	0.569231	1.189878	0.005614	1.210871	Bread -> Coffee
9	(66,)	(24,)	0.086107	0.478394	0.047544	0.552147	1.154168	0.006351	1.164682	Cookies -> Bread

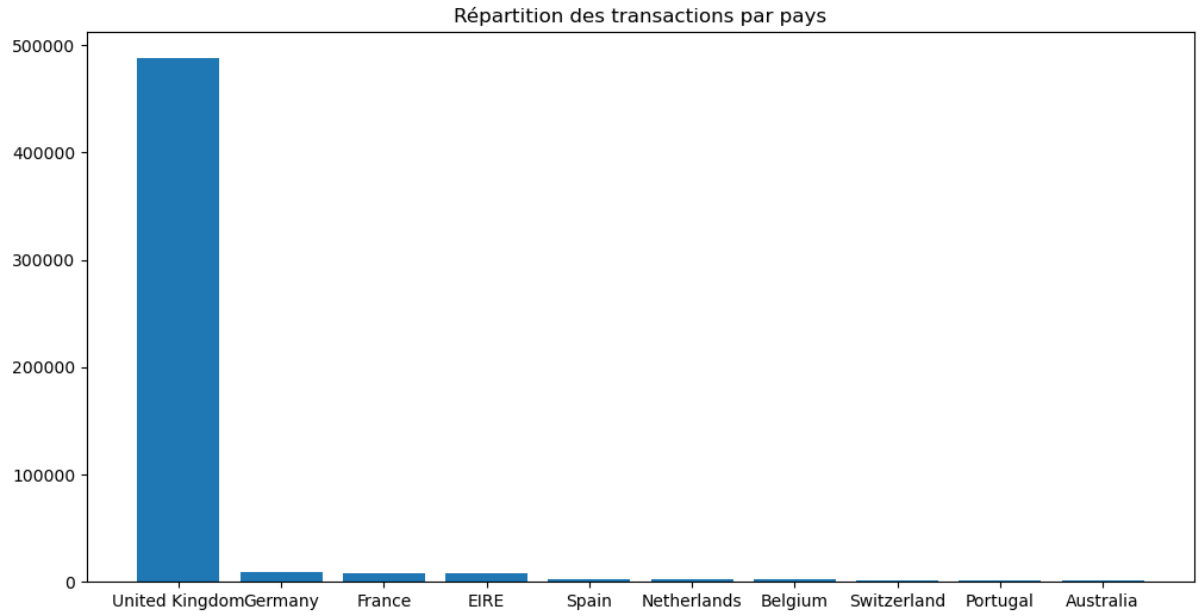
“Confidence” est classée par ordre décroissant dans le tableau ci-dessus.



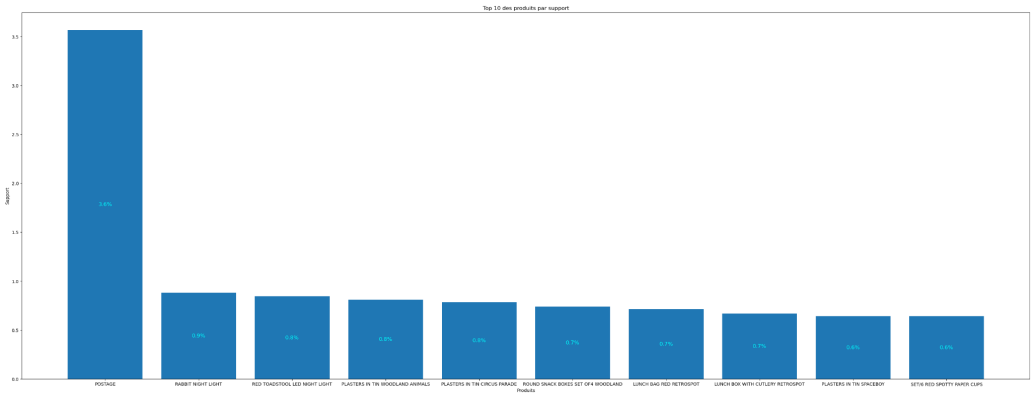
On peut remarquer que “Bread” est associé à beaucoup d’autres produits donc cela fait qu’il est l’un des produits phares de la boulangerie. Notons également qu’on n’a pas une différence significative entre les valeurs de lift du fait que le calcul de lift prend en compte la popularité du produit mais aussi rappelons que, dans le top 10, les autres produits sauf “Bread” occupent moins de 7% des transactions. Donc on peut en déduire que c’est la vente de “Bread” qui influence le plus la commercialisation des autres produits associés.
Ainsi nous proposons que des remises soient plus appliquées sur le produit phare pour augmenter la vente des autres avec lesquels il est le plus souvent acheté.

C. Les données du fichier "Online_retail"

Ce fichier est constitué des données sur une entreprise en ligne regroupées en huit colonnes dont “InvoiceNo” qui représente le numéro de facture. Un numéro entier de 6 chiffres est attribué de manière unique à chaque transaction. Si ce code commence par la lettre "c", cela indique une annulation de la transaction, “StockCode” qui représente le code produit. Un numéro entier de 5 chiffres est attribué de manière unique à chaque produit distinct, “Description” c’est-à-dire le nom du produit et “country” le nom du pays où réside le client.
On a enlevé les transactions qui ont été annulées et les colonnes concernant les quantités, la date et l’heure des factures, les prix unitaire ainsi que l’identifiant des clients. Ainsi on a représenté les données restantes dans le graphique ci-dessous représentant les transactions en fonction du pays.



On remarque en général un très grand nombre de transactions au Royaume Uni au moment où celles-ci tournent dans les environs de 9000 et moins dans les 9 autres pays pris comme exemple dans le tracé du graphique. Cependant dans le reste de notre analyse on ne tiendra pas compte des ventes au Royaume Uni.
Afin de limiter la taille de l'ensemble des données, on a décidé de ne prendre en compte que des données en France. Ainsi on compte 1563 items avec 392 transactions.
Le deuxième graphique représente le top 10 des produits les plus vendus en France.



On remarque que les pourcentages ne sont pas aussi grands que cela. Bien que faible, Postage regroupe un pourcentage de 3.6 % les autres sont moins de 1%.
Par la suite on ajouté une colonne tid qui représente l'identifiant pour chaque transaction.

InvoiceNo	StockCode	Description	Country	tid
536370	22728	ALARM CLOCK BAKELIKE PINK	France	1
536370	22727	ALARM CLOCK BAKELIKE RED	France	1
536370	22726	ALARM CLOCK BAKELIKE GREEN	France	1
536370	21724	PANDA AND BUNNIES STICKER SHEET	France	1
536370	21883	STARS GIFT TAPE	France	1
...	...	...	...	...
581587	22613	PACK OF 20 SPACEBOY NAPKINS	France	392
581587	22899	CHILDREN'S APRON DOLLY GIRL	France	392
581587	23254	CHILDRENS CUTLERY DOLLY GIRL	France	392
581587	23255	CHILDRENS CUTLERY CIRCUS PARADE	France	392
581587	22138	BAKING SET 9 PIECE RETROSPOT	France	392

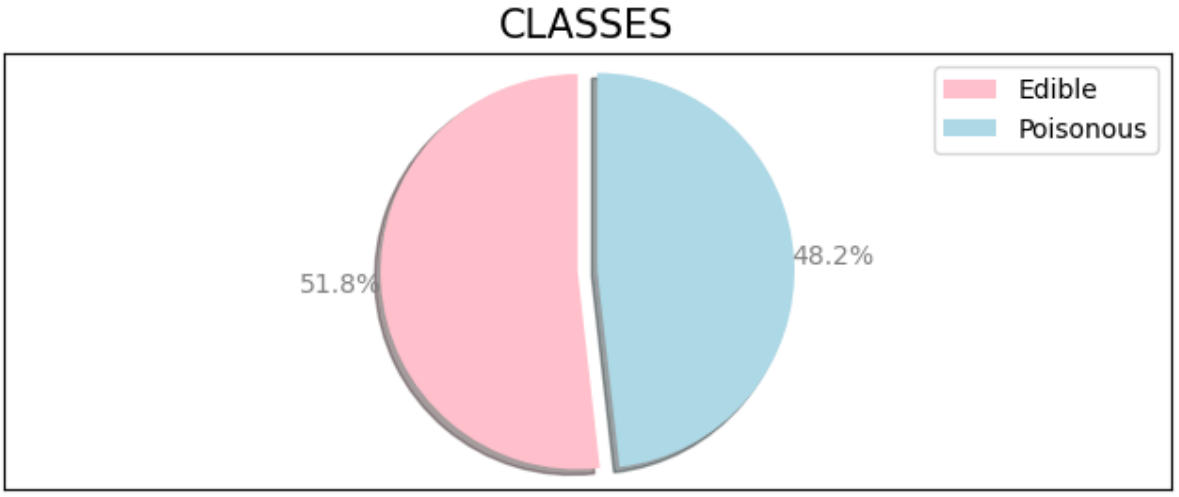
On a pris comme valeur minimum de support 10%.Le tableau suivant représente les 12 règles qui ont des valeurs de support supérieures à 10%

	lhs	rhs	lhs_support	rhs_support	support	confidence	lift	leverage	conviction	Rules
0	(698,)	(975,)	0.125000	0.765306	0.104592	0.836735	1.093333	0.008929	1.437500	LUNCH BAG APPLE DESIGN -> POSTAGE
1	(704,)	(975,)	0.153061	0.765306	0.122449	0.800000	1.045333	0.005310	1.173469	LUNCH BAG RED RETROSPOT -> POSTAGE
2	(711,)	(975,)	0.117347	0.765306	0.102041	0.869565	1.136232	0.012234	1.799320	LUNCH BAG WOODLAND -> POSTAGE
3	(713,)	(975,)	0.142857	0.765306	0.114796	0.803571	1.050000	0.005466	1.194805	LUNCH BOX WITH CUTLERY RETROSPOT -> POSTAGE
4	(957,)	(952,)	0.170918	0.168367	0.102041	0.597015	3.545907	0.073264	2.063681	PLASTERS IN TIN WOODLAND ANIMALS -> PLASTERS I...
5	(952,)	(957,)	0.168367	0.170918	0.102041	0.606061	3.545907	0.073264	2.104592	PLASTERS IN TIN CIRCUS PARADE -> PLASTERS IN T...
6	(952,)	(975,)	0.168367	0.765306	0.147959	0.878788	1.148283	0.019107	1.936224	PLASTERS IN TIN CIRCUS PARADE -> POSTAGE
7	(957,)	(954,)	0.170918	0.137755	0.104592	0.611940	4.442233	0.081047	2.221939	PLASTERS IN TIN WOODLAND ANIMALS -> PLASTERS I...
8	(954,)	(957,)	0.137755	0.170918	0.104592	0.759259	4.442233	0.081047	3.443878	PLASTERS IN TIN SPACEBOY -> PLASTERS IN TIN WO...
9	(954,)	(975,)	0.137755	0.765306	0.114796	0.833333	1.088889	0.009371	1.408163	PLASTERS IN TIN SPACEBOY -> POSTAGE
10	(957,)	(975,)	0.170918	0.765306	0.137755	0.805970	1.053134	0.006950	1.209576	PLASTERS IN TIN WOODLAND ANIMALS -> POSTAGE
11	(984,)	(975,)	0.188776	0.765306	0.165816	0.878378	1.147748	0.021345	1.929705	RABBIT NIGHT LIGHT -> POSTAGE
12	(975,)	(984,)	0.765306	0.188776	0.165816	0.216667	1.147748	0.021345	1.035606	POSTAGE -> RABBIT NIGHT LIGHT

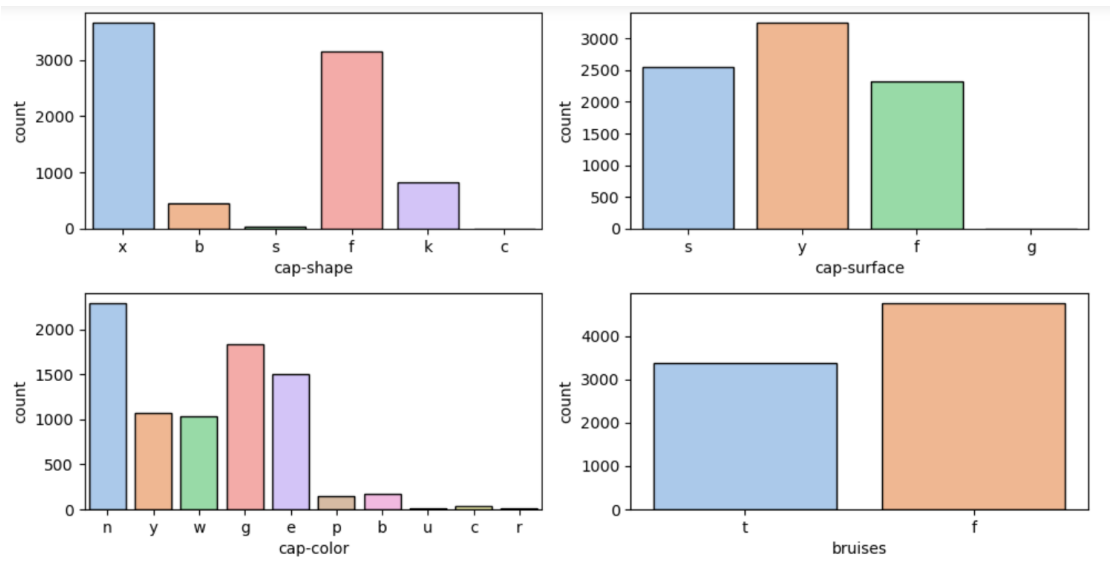
Après avoir filtré les règles selon le lift supérieur à 4, on obtient un plus grand lift lorsque les produits “SET/6 RED SPOTTY PAPER CUPS” (A) et “SET/6 RED SPOTTY PAPER PLATES” (B) sont achetés ensemble cependant on a une plus grande confiance lorsque l’achat de B favorise aussi la commercialisation de A.

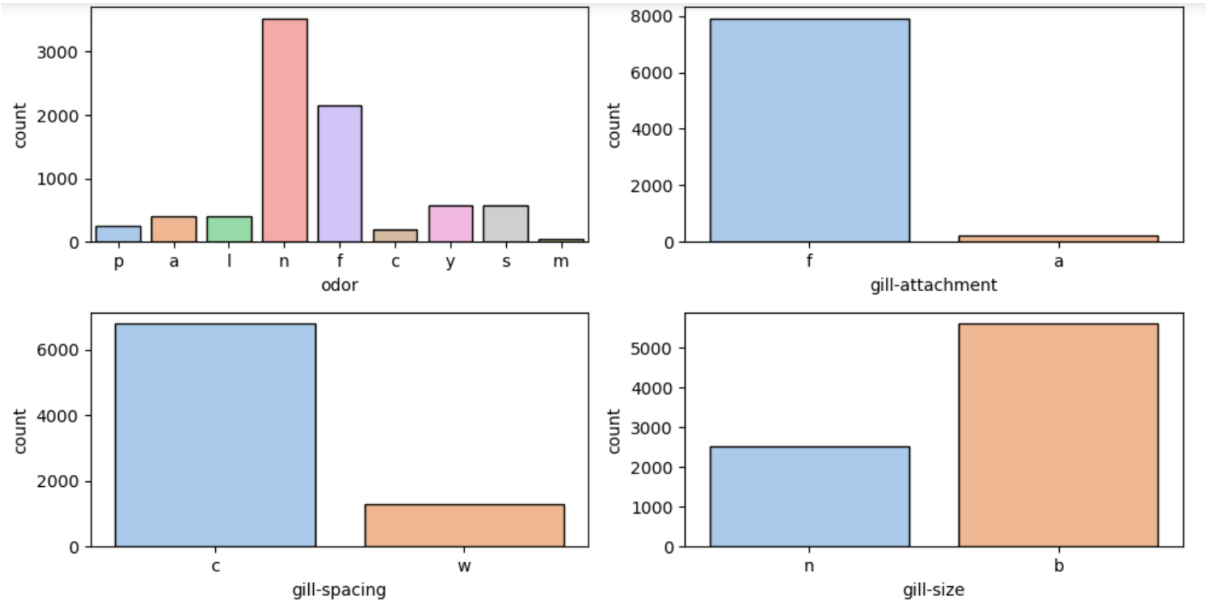
D. Les données du fichier "mushrooms"

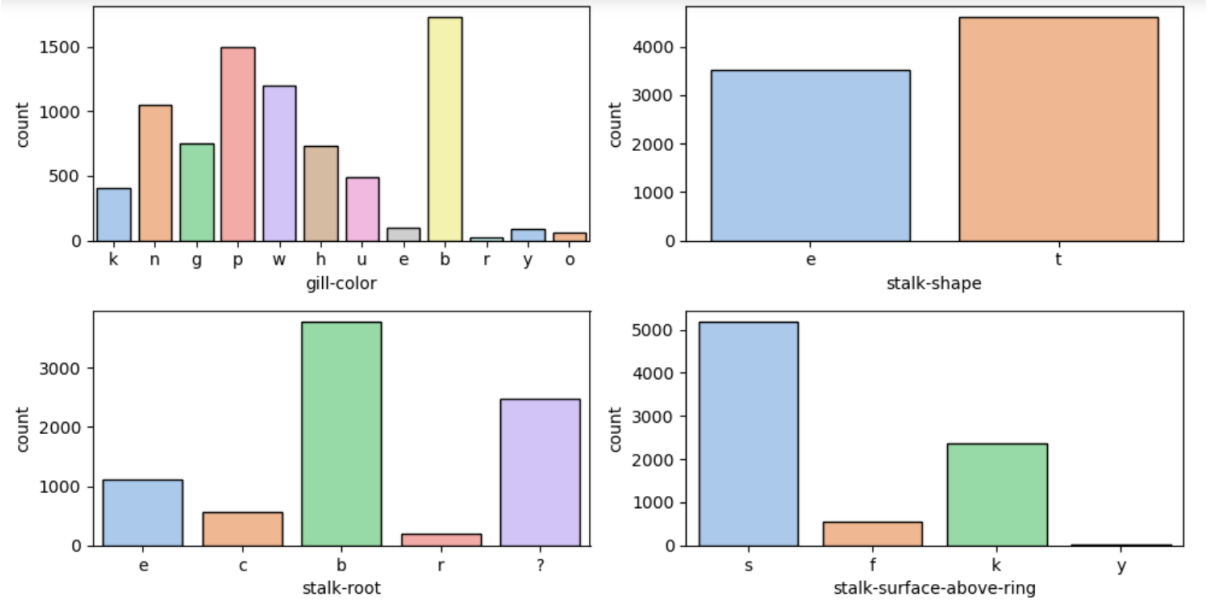
Cet ensemble de données comprend des descriptions d'échantillons hypothétiques correspondant à 23 espèces de champignons à branchies des familles Agaricus et Lepiota, tirées de The Audubon Society Field Guide to North American Mushrooms (1981). Chaque espèce est identifiée comme certainement comestible, certainement toxique, ou de comestibilité inconnue et déconseillée. Cette dernière catégorie a été combinée avec celle des champignons vénéneux. D’après The Audubon Society Field Guide to North American Mushrooms (1981), il n'existe pas de règle simple pour déterminer la comestibilité d'un champignon ; pas de règle du type "trois feuilles, que ce soit" pour le chêne et le lierre vénéneux.
Le graphe ci-dessous représente les pourcentages des deux populations “Edible” étant comestibles et “Poisonous” toxiques. On ne note pas une très grande différence de pourcentage entre ces deux espèces.

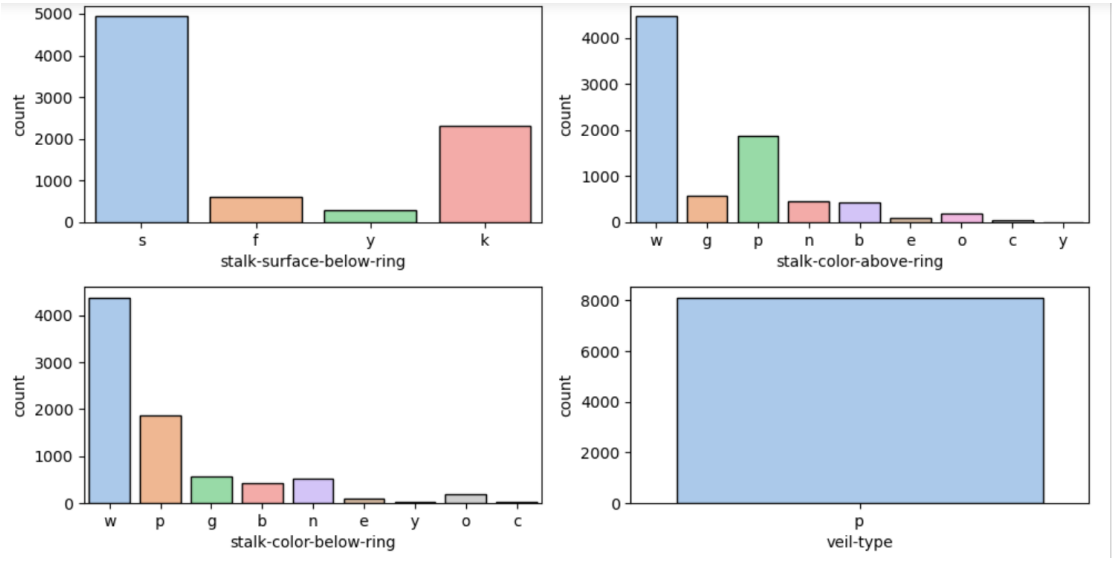


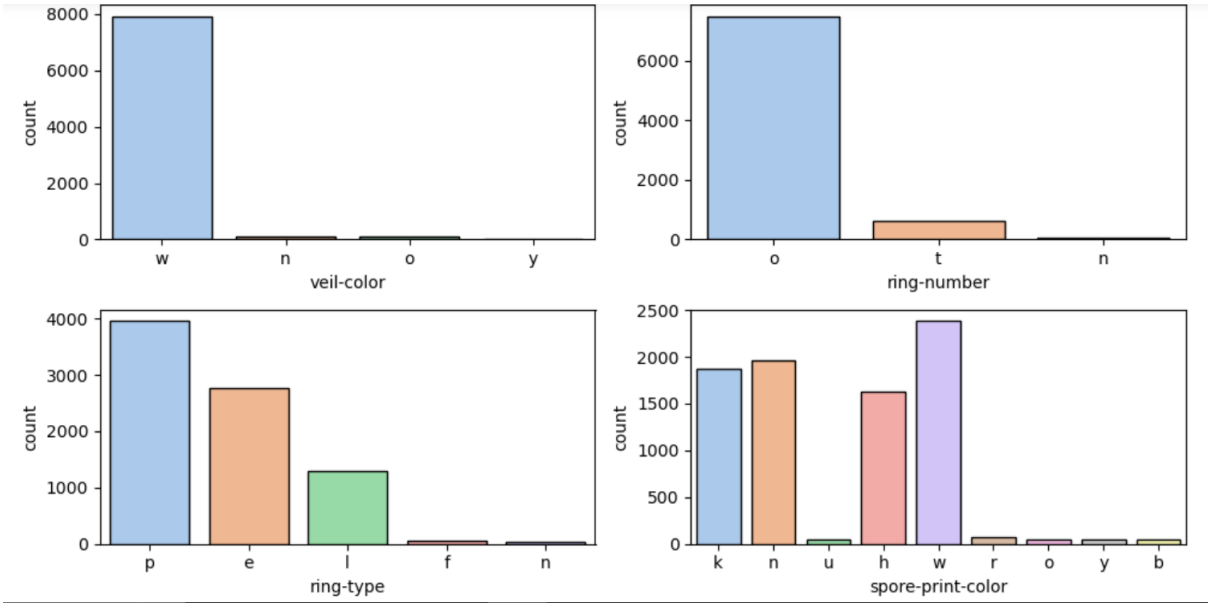
Les graphiques ci-dessous représentent les cartes de comptoir pour les attributs de ces deux espèces:




count
odor
p a l n f c y s m
gill-attachment
f a
gill-spacing
c w
gill-size
n b

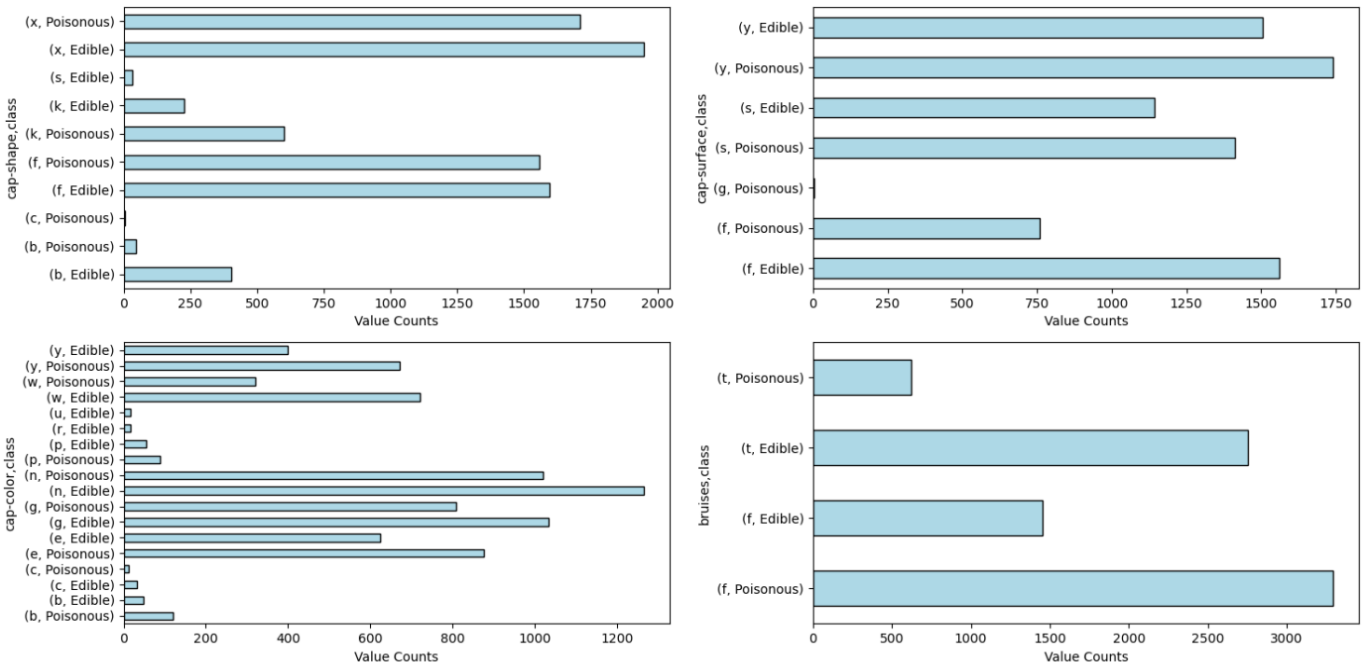

count
gill-color
k n g p w h u e b r y o
stalk-shape
e t
stalk-root
e c b r ?
stalk-surface-above-ring
s f k y

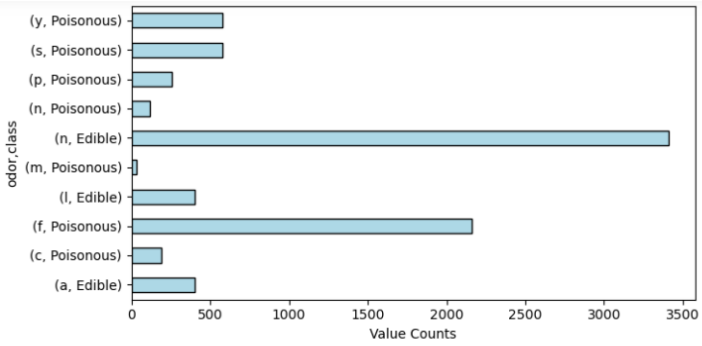

count
stalk-surface-below-ring
s f y k
stalk-color-above-ring
w g p n b e o c y
stalk-color-below-ring
w p g b n e y o c
veil-type
p

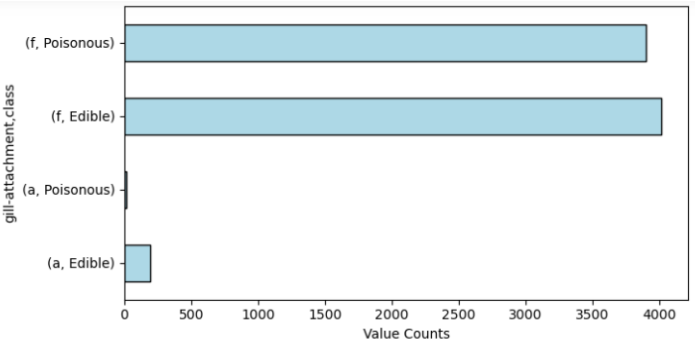


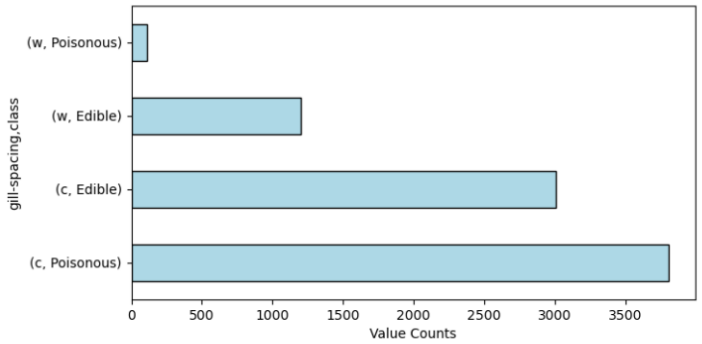
Pour plus de clarté on a représenté les attributs selon leurs classes: Edible ou Poisonous.
On remarque qu'entre autres:

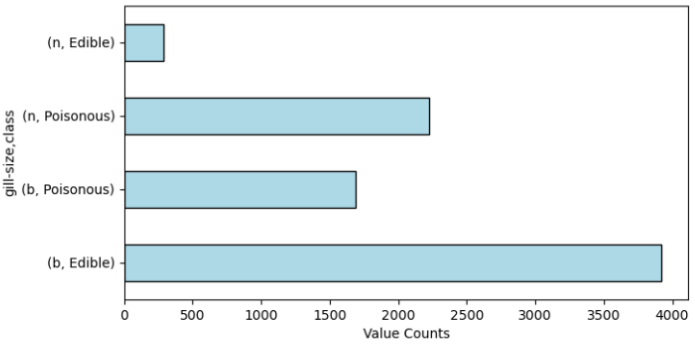
- Les champignons ayant une forme en chapeau conique ne sont pas toxiques.
- Les champignons ayant une surface en chapeau rainuré ne sont pas toxiques.
- Les champignons ayant un chapeau coloré en violet et en vert sont toxiques
- Les champignons ayant une odeur de moisi sont comestibles
- …

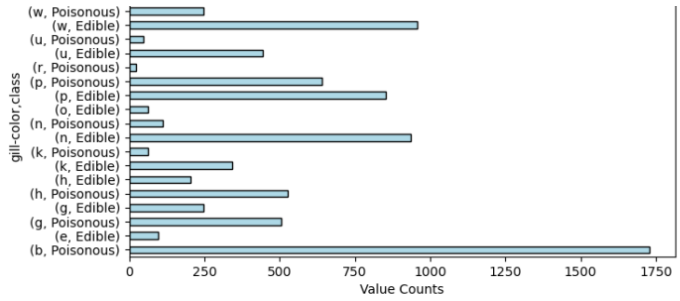


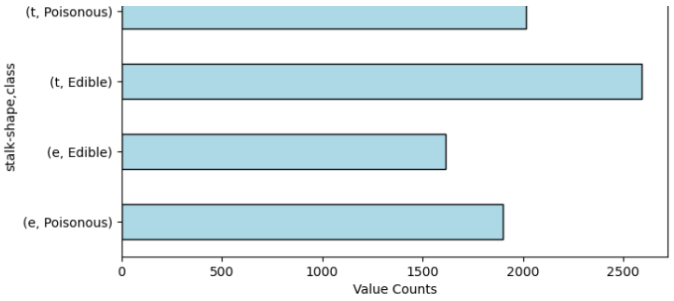

odor,class
(y, Poisonous)
(s, Poisonous)
(p, Poisonous)
(n, Poisonous)
(n, Edible)
(m, Poisonous)
(l, Edible)
(f, Poisonous)
(c, Poisonous)
(a, Edible)
0 500 1000 1500 2000 2500 3000 3500
Value Counts

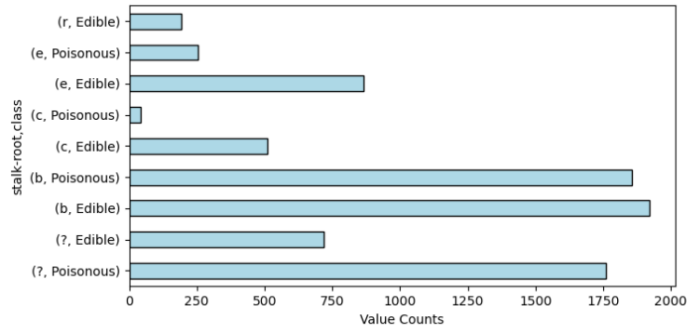

gill-attachment,class
(f, Poisonous)
(f, Edible)
(a, Poisonous)
(a, Edible)
0 500 1000 1500 2000 2500 3000 3500 4000
Value Counts

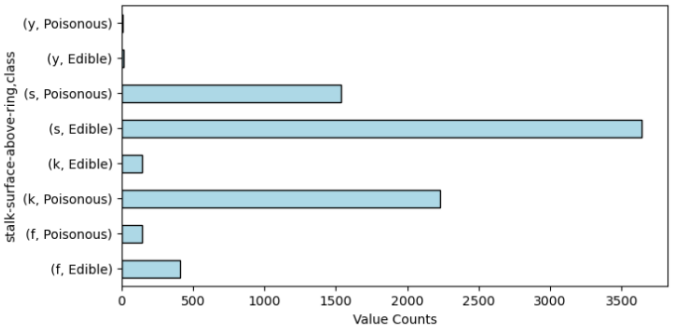

gill-spacing,class
(w, Poisonous)
(w, Edible)
(c, Edible)
(c, Poisonous)
0 500 1000 1500 2000 2500 3000 3500
Value Counts

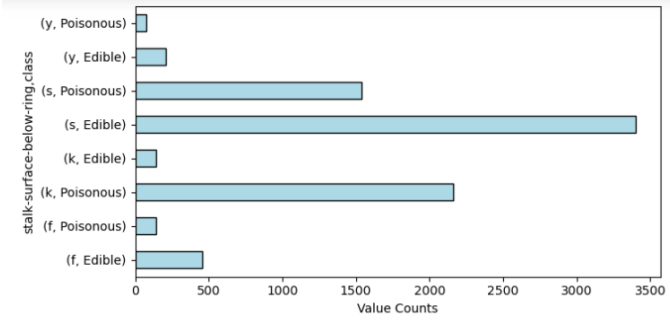

gill-size,class
(n, Edible)
(n, Poisonous)
(b, Poisonous)
(b, Edible)
0 500 1000 1500 2000 2500 3000 3500 4000
Value Counts

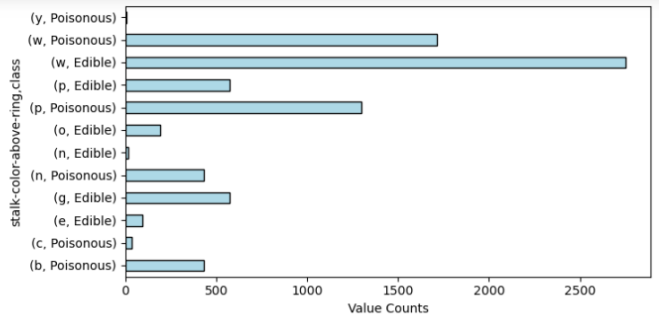

gill-color,class
(w, Poisonous)
(w, Edible)
(u, Poisonous)
(u, Edible)
(r, Poisonous)
(p, Poisonous)
(p, Edible)
(o, Edible)
(n, Poisonous)
(n, Edible)
(k, Poisonous)
(k, Edible)
(h, Edible)
(h, Poisonous)
(g, Edible)
(g, Poisonous)
(e, Edible)
(b, Poisonous)
0 250 500 750 1000 1250 1500 1750
Value Counts

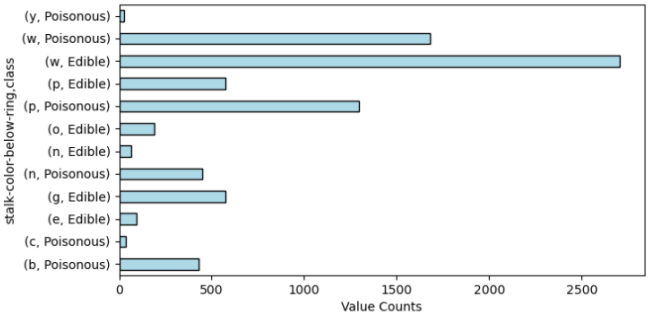

stalk-shape,class
(t, Poisonous)
(t, Edible)
(e, Edible)
(e, Poisonous)
0 500 1000 1500 2000 2500
Value Counts

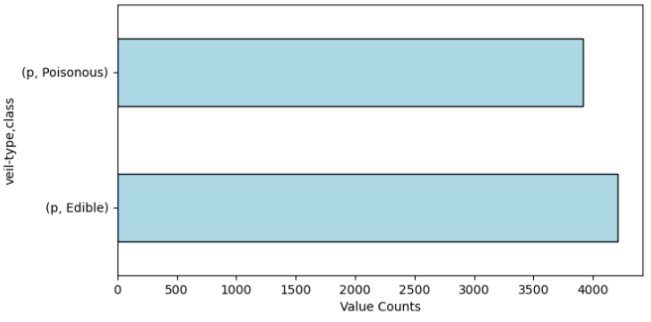

stalk-root,class
(r, Edible)
(e, Poisonous)
(e, Edible)
(c, Poisonous)
(c, Edible)
(b, Poisonous)
(b, Edible)
(?, Edible)
(?, Poisonous)
0 250 500 750 1000 1250 1500 1750 2000
Value Counts

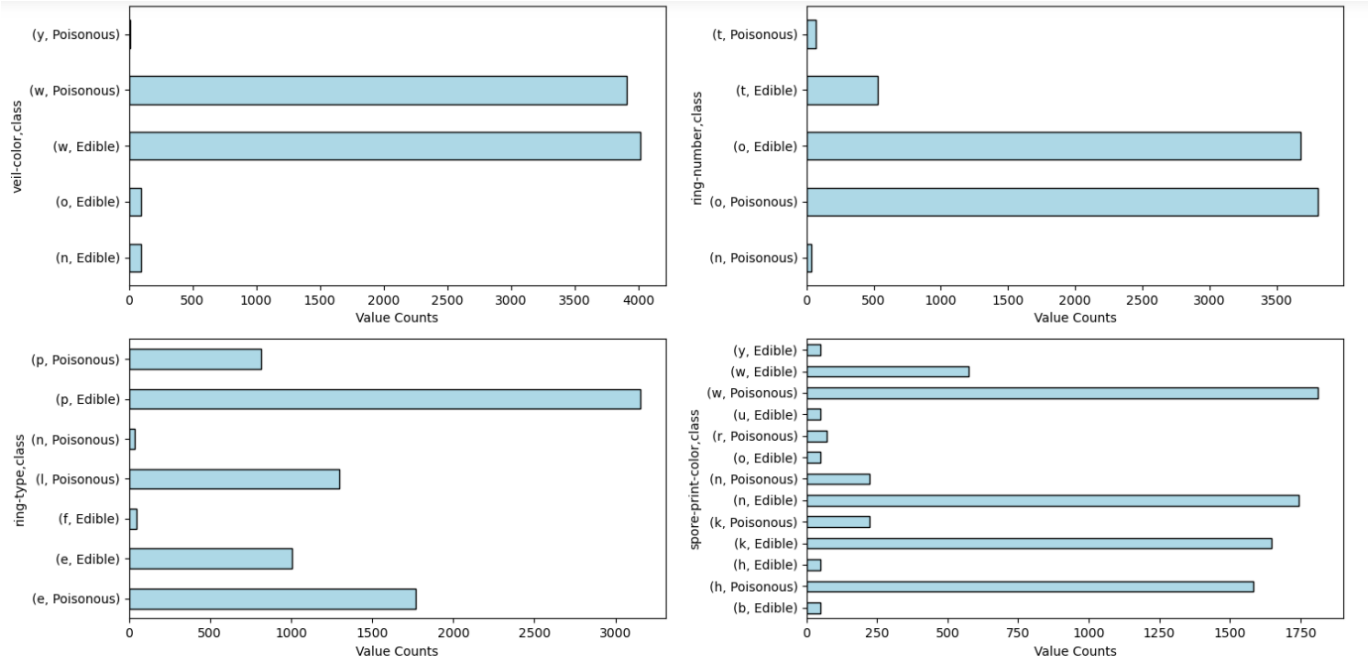

stalk-surface-above-ring,class
(y, Poisonous)
(y, Edible)
(s, Poisonous)
(s, Edible)
(k, Edible)
(k, Poisonous)
(f, Poisonous)
(f, Edible)
0 500 1000 1500 2000 2500 3000 3500
Value Counts


stalk-surface-below-ring,class
(y, Poisonous)
(y, Edible)
(s, Poisonous)
(s, Edible)
(k, Edible)
(k, Poisonous)
(f, Poisonous)
(f, Edible)
0 500 1000 1500 2000 2500 3000 3500
Value Counts


stalk-color-above-ring,class
(y, Poisonous)
(w, Poisonous)
(w, Edible)
(p, Edible)
(p, Poisonous)
(o, Edible)
(n, Edible)
(n, Poisonous)
(g, Edible)
(e, Edible)
(c, Poisonous)
(b, Poisonous)
0 500 1000 1500 2000 2500
Value Counts


stalk-color-below-ring,class
(y, Poisonous)
(w, Poisonous)
(w, Edible)
(p, Edible)
(p, Poisonous)
(o, Edible)
(n, Edible)
(n, Poisonous)
(g, Edible)
(e, Edible)
(c, Poisonous)
(b, Poisonous)
0 500 1000 1500 2000 2500
Value Counts


veil-type,class
(p, Poisonous)
(p, Edible)
0 500 1000 1500 2000 2500 3000 3500 4000
Value Counts



Pour répondre aux questions posées,

- **Les itemsets ayant un support >= 80% sont:**
  Itemsets fréquents de taille 1: [('gill-spacing_c',), ('gill-spacing_w',), ('veil-color_n',), ('veil-color_y',), ('ring-number_t',)]

  Itemsets fréquents de taille 2 : [('gill-spacing_c', 'gill-spacing_w'), ('gill-spacing_c', 'veil-color_n'), ('gill-spacing_c', 'veil-color_y'), ('gill-spacing_c', 'ring-number_t'), ('gill-spacing_w', 'veil-color_n'), ('gill-spacing_w', 'veil-color_y'), ('veil-color_n', 'veil-color_y'), ('veil-color_n', 'ring-number_t'), ('veil-color_y', 'ring-number_t')]

  Itemsets fréquents de taille 3: [('gill-spacing_c', 'gill-spacing_w', 'veil-color_n'), ('gill-spacing_c', 'gill-spacing_w', 'veil-color_y'), ('gill-spacing_c', 'veil-color_n', 'veil-color_y'), ('gill-spacing_c', 'veil-color_n', 'ring-number_t'), ('gill-spacing_c', 'veil-color_y', 'ring-number_t'), ('gill-spacing_w', 'veil-color_n', 'veil-color_y'), ('veil-color_n', 'veil-color_y', 'ring-number_t')]

  Itemsets fréquents de taille 4 : [('gill-spacing_c', 'gill-spacing_w', 'veil-color_n', 'veil-color_y'), ('gill-spacing_c', 'veil-color_n', 'veil-color_y', 'ring-number_t')]

- **Les règles ayant une confiance >= 90% sont classées sur le tableau ci-dessous.**
  On a gardé les règles d'associations des attributs utiles (c'est-à-dire plus de 90% de confiance) pour prédire les attributs et ou les classes des champignons.
  Par exemple, la règle "veil-color_n" -> "gill-spacing_c" avec une confiance de 0,974 signifie que lorsqu'un champignon a un voile de couleur brune (veil-color_n), il a également une faible espacement des lamelles (gill-spacing_c) avec une probabilité de 97,4%. En d'autres termes, si vous trouvez un champignon avec un voile brun, il y

a de fortes chances qu'il ait également un faible espacement entre les lamelles. Cependant, il est important de noter que cette règle est basée sur les données qui ont été utilisées pour la construire et qu'elle peut ne pas être généralisable à toutes les situations.

lhs	rhs	lhs_support	rhs_support	support	confidence	lift	leverage	conviction	Rules
(36,)	(35,)	0.838503	0.974151	0.812654	0.969172	0.994889	-0.004175	0.838503	gill-spacing_w -> gill-spacing_c
(85,)	(35,)	1.000000	0.974151	0.974151	0.974151	1.000000	0.000000	1.000000	veil-color_n -> gill-spacing_c
(35,)	(85,)	0.974151	1.000000	0.974151	1.000000	1.000000	0.000000	inf	gill-spacing_c -> veil-color_n
(88,)	(35,)	0.975382	0.974151	0.973166	0.997728	1.024203	0.022997	11.379452	veil-color_y -> gill-spacing_c
(35,)	(88,)	0.974151	0.975382	0.973166	0.998989	1.024203	0.022997	24.353767	gill-spacing_c -> veil-color_y
(91,)	(35,)	0.921713	0.974151	0.898080	0.974359	1.000214	0.000192	1.008124	ring-number_t -> gill-spacing_c
(35,)	(91,)	0.974151	0.921713	0.898080	0.921911	1.000214	0.000192	1.002524	gill-spacing_c -> ring-number_t
(36,)	(85,)	0.838503	1.000000	0.838503	1.000000	1.000000	0.000000	inf	gill-spacing_w -> veil-color_n
(36,)	(88,)	0.838503	0.975382	0.814870	0.971814	0.996343	-0.002991	0.873441	gill-spacing_w -> veil-color_y
(88,)	(85,)	0.975382	1.000000	0.975382	1.000000	1.000000	0.000000	inf	veil-color_y -> veil-color_n
(85,)	(88,)	1.000000	0.975382	0.975382	0.975382	1.000000	0.000000	1.000000	veil-color_n -> veil-color_y
(91,)	(85,)	0.921713	1.000000	0.921713	1.000000	1.000000	0.000000	inf	ring-number_t -> veil-color_n
(85,)	(91,)	1.000000	0.921713	0.921713	0.921713	1.000000	0.000000	1.000000	veil-color_n -> ring-number_t
(91,)	(88,)	0.921713	0.975382	0.897095	0.973291	0.997856	-0.001927	0.921713	ring-number_t -> veil-color_y
(88,)	(91,)	0.975382	0.921713	0.897095	0.919738	0.997856	-0.001927	0.975382	veil-color_y -> ring-number_t
(36,)	(35, 85)	0.838503	0.974151	0.812654	0.969172	0.994889	-0.004175	0.838503	gill-spacing_w -> gill-spacing_c
(36,)	(35, 88)	0.838503	0.973166	0.812654	0.969172	0.995896	-0.003349	0.870446	gill-spacing_w -> gill-spacing_c
(88,)	(35, 85)	0.975382	0.974151	0.973166	0.997728	1.024203	0.022997	11.379452	veil-color_y -> gill-spacing_c