**يتوفر اثنين DATA SETS :**

1. **mushrooms.csv**
2. **Student\_Marks.csv**

**يتم استخدام مجموعة البيانات الثانية في المصنف Linear regresion حيث أنها مجموعة بيانات رقمية تتكون من 100 سطر وثلاث أعمدة :**

1. **عدد المواد المسجلة لدى الطالب number\_courses**
2. **عدد ساعات الدراسة time\_study**
3. **درجة الطالب من مئة بناء على السمتين السابقتين Marks**

**يتم استخدام مجموعة البيانات الاولى في المصنف Decision tree حيث أنها مجموعة بيانات نصية تم عمل المعالجة اللازمة عليها لتتوافق مع هذا المصنف .**

تتكون من 23 عمود و 8124 سطر .

العمود الأول هو العمود الهدف الذي سيصنف الفطر هل هو سام ام صالح للاكل ، وباقي الاعمدة تمثل السمات .

مواصفات الأعمدة :

1. الفئات: صالحة للأكل = e ، سامة = p
2. شكل الغطاء: جرس = b ، مخروطي = c ، محدب = x ، مسطح = f ، مقبض = k ، غائر = s
3. سطح الغطاء: ليفي = f ، أخاديد = g ، متقشرة = y ، أملس = s
4. غطاء اللون: بني = n ، برتقالي = b ، قرفة = c ، رمادي = g ، أخضر = r ، وردي = p ، بنفسجي = u ، أحمر = e ، أبيض = w ، أصفر = y
5. كدمات: كدمات = t ، لا = f
6. الرائحة: لوز = a ، يانسون = l ، كريوزوت = c ، مريب = y ، كريه = f ، متعفن = m ، لا شيء = n ، لاذع = p ، حار = s
7. مرفق الخيشومية: مرفق = a ، تنازلي = d ، حر = f ، مسنن = n
8. تباعد الخياشيم: قريب = c ، مزدحم = w ، بعيد = d
9. حجم الخيشومية: عريض = b ، ضيق = n
10. لون الخيشومية: أسود = k ، بني = n ، برتقالي = b ، شوكولاتة = h ، رمادي = g ، أخضر = r ، برتقالي = o ، وردي = p ، أرجواني = u ، أحمر = e ، أبيض = w ، أصفر =y
11. شكل الساق: تكبير = e ، مستدق = t
12. ساق الجذر: بصلي = b ، مضرب = c ، كوب = u ، يساوي = e ، جذور الشكل = z ، جذر = ص ، مفقود =؟
13. ساق السطح فوق الحلقة: ليفي = f ، متقشر = y ، حريري = k ، أملس = s
14. ساق السطح تحت الحلقة: ليفي = f ، متقشر = y ، حريري = k ، أملس = s
15. ساق اللون فوق الحلقة: بني = n ، برتقالي = ب ، قرفة = ج ، رمادي = ز ، برتقالي = س ، وردي = ص ، أحمر = هـ ، أبيض = w ، أصفر = ص
16. لون ساق أسفل الحلقة: بني = n ، برتقالي = ب ، قرفة = ج ، رمادي = ز ، برتقالي = س ، وردي = ص ، أحمر = هـ ، أبيض = w ، أصفر = ص
17. نوع الحجاب: جزئي = p ، عالمي = u
18. لون الحجاب: بني = n ، برتقالي = o ، أبيض = w ، أصفر = y
19. رقم الحلقة: لا شيء = n ، واحد = o ، اثنان = t
20. نوع الحلقة: cobwebby = c ، زائل = e ، احتراق = f ، كبير = l ، لا شيء = n ، قلادة = p ، غلاف = s ، منطقة = z
21. لون الطباعة البوغ: أسود = ك ، بني = ن ، برتقالي = ب ، شوكولاتة = ح ، أخضر = ص ، برتقالي = س ، أرجواني = ش ، أبيض = ث ، أصفر = ص
22. عدد السكان: وفير = أ ، متجمع = ج ، عدد كبير = ن ، متناثرة = ق ، عدة = ت ، انفرادي = ص
23. الموائل: الأعشاب = g ، الأوراق = l ، المروج = m ، المسارات = p ، المناطق الحضرية = u ، النفايات = w ، الأخشاب = d