Laboratory Activities for Week 4: Seaborn (SNS)

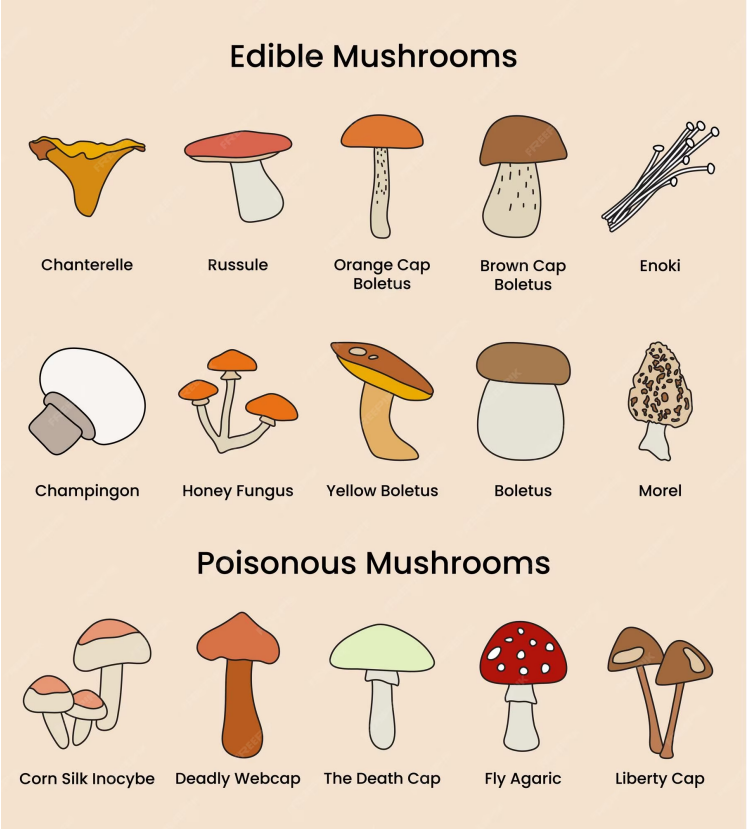
SC310005 Artificial Intelligence
Khon Kaen Business School

(10 Points) The Mushroom Classification dataset utilizing Seaborn

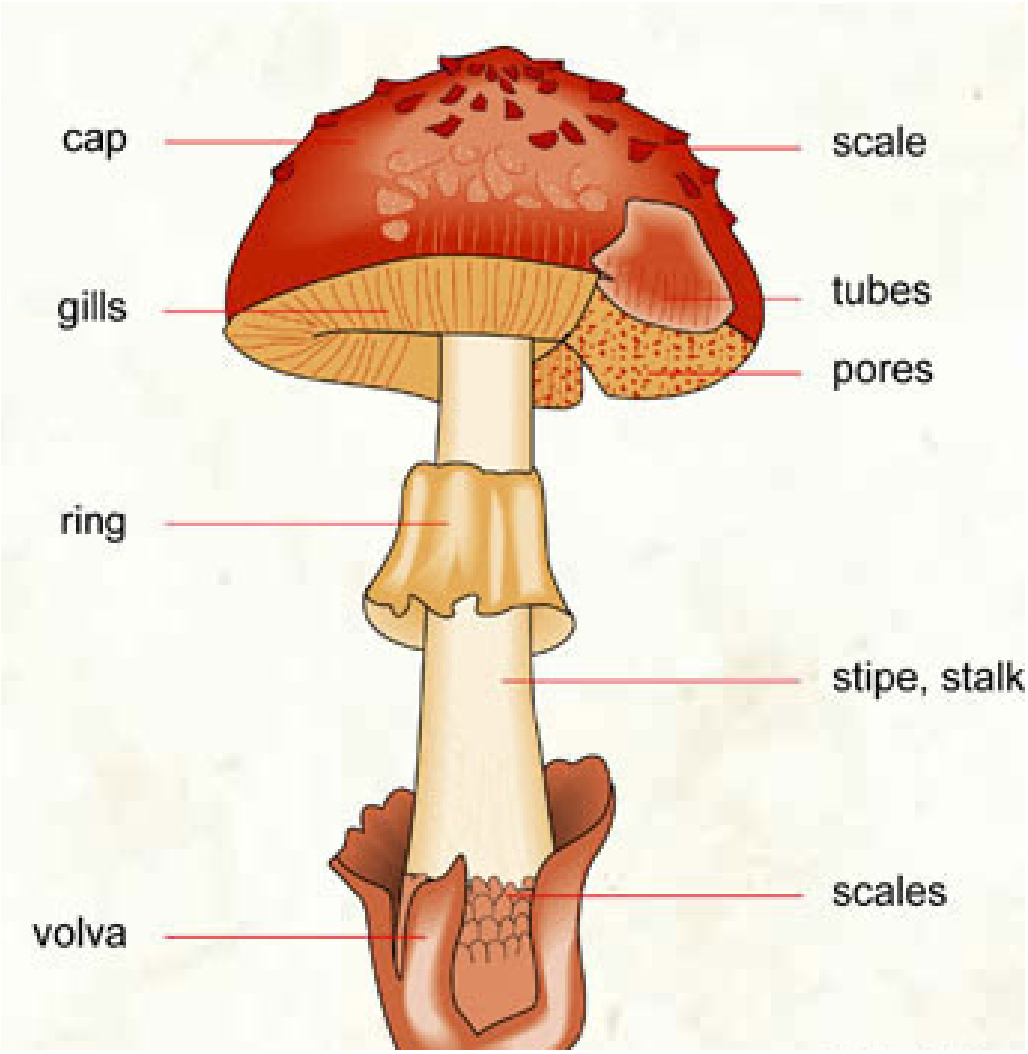
---

Introduction:
The Mushroom Classification dataset comprises categorical attributes of mushrooms, such as cap shape, cap color, odor, etc., and their classification as either edible or poisonous. The analysis aims to understand the distribution of different features, examine correlations between attributes, and derive meaningful insights to distinguish edible from poisonous mushrooms.

Dataset:
https://raw.githubusercontent.com/kaopanboonyuen/SC310005_ArtificialIntelligence_2023s1/main/dataset/mushrooms_dataset.csv



Q1	Visualize the distribution of classes using a count plot in Seaborn.
Q2	Create a histogram of cap sizes ('cap-surface').
Q3	Generate a count plot to display the count of different stalk shapes.
Q4	Create a count plot to visualize the distribution of gill sizes.
Q5	Visualize the count of different cap shapes specifically for edible mushrooms.
Q6	Visualize the count of different cap colors specifically for poisonous mushrooms.
Q7	Visualize a Seaborn countplot with a new variable 'cap-surface-color' by combining 'cap-surface' and 'cap-color'.
Q8	Generate a histogram in Seaborn to visualize the distribution of spore print colors.
Q9	Visualize a count plot with a new variable, 'habitat-population', by combining 'habitat' and 'population'.
Q10 (Extra Point)	Determining the Most Impactful Variable in Identifying Poisonous Mushrooms **Task Description:** (1) You are tasked with identifying the variable that has the most significant effect on separating poisonous mushrooms from edible ones using the Mushroom Classification dataset. Utilize Seaborn visualization techniques to explore the relationship between mushroom characteristics and their edibility. **Analysis and Conclusion:** (2) Examine the plots to determine the variable where the 'poisonous' and 'edible' classes are most distinctly separated. Provide reasoning and evidence from the Seaborn graphs to identify the variable that has the most significant effect on distinguishing between poisonous and edible mushrooms.



Appendix:

Attribute Information: (classes: edible=e, poisonous=p)

cap-shape: bell=b,conical=c,convex=x,flat=f, knobbed=k,sunken=s

cap-surface: fibrous=f,grooves=g,scaly=y,smooth=s

cap-color:
brown=n,buff=b,cinnamon=c,gray=g,green=r,pink=p,purple=u,red=e,white=w,yellow=y

bruises: bruises=t,no=f

odor: almond=a,anise=l,creosote=c,fishy=y,foul=f,musty=m,none=n,pungent=p,spicy=s

gill-attachment: attached=a,descending=d,free=f,notched=n

gill-spacing: close=c,crowded=w,distant=d

gill-size: broad=b,narrow=n

gill-color: black=k,brown=n,buff=b,chocolate=h,gray=g,
green=r,orange=o,pink=p,purple=u,red=e,white=w,yellow=y

stalk-shape: enlarging=e,tapering=t

stalk-root: bulbous=b,club=c,cup=u,equal=e,rhizomorphs=z,rooted=r,missing=?

stalk-surface-above-ring: fibrous=f,scaly=y,silky=k,smooth=s

stalk-surface-below-ring: fibrous=f,scaly=y,silky=k,smooth=s

stalk-color-above-ring:
brown=n,buff=b,cinnamon=c,gray=g,orange=o,pink=p,red=e,white=w,yellow=y

stalk-color-below-ring:
brown=n,buff=b,cinnamon=c,gray=g,orange=o,pink=p,red=e,white=w,yellow=y

veil-type: partial=p,universal=u

veil-color: brown=n,orange=o,white=w,yellow=y

ring-number: none=n,one=o,two=t

ring-type: cobwebby=c,evanescent=e,flaring=f,large=l,none=n,pendant=p,sheathing=s,zone=z

spore-print-color:
black=k,brown=n,buff=b,chocolate=h,green=r,orange=o,purple=u,white=w,yellow=y

population: abundant=a,clustered=c,numerous=n,scattered=s,several=v,solitary=y

habitat: grasses=g,leaves=l,meadows=m,paths=p,urban=u,waste=w,woods=d