```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import seaborn as sns

data = pd.read_csv('mushrooms.csv')
dt = pd.DataFrame(data)
dt
```

```
     class cap-shape cap-surface cap-color bruises odor gill-
attachment  \
0        p         x           s         n       t    p
f
1        e         x           s         y       t    a
f
2        e         b           s         w       t    l
f
3        p         x           y         w       t    p
f
4        e         x           s         g       f    n
f
...    ...       ...         ...       ...     ...  ...             ..
.
8119     e         k           s         n       f    n
a
8120     e         x           s         n       f    n
a
8121     e         f           s         n       f    n
a
8122     p         k           y         n       f    y
f
8123     e         x           s         n       f    n
a

     gill-spacing gill-size gill-color  ... stalk-surface-below-
ring  \
0               c         n          k  ...                        s

1               c         b          k  ...                        s

2               c         b          n  ...                        s

3               c         n          n  ...                        s

4               w         b          k  ...                        s

...           ...       ...        ...  ...                      ...

8119            c         b          y  ...                        s
```

```
8120           c         b          y  ...                        s

8121           c         b          n  ...                        s

8122           c         n          b  ...                        k

8123           c         b          y  ...                        s


     stalk-color-above-ring stalk-color-below-ring veil-type veil-
color  \
0                         w                       w         p
w
1                         w                       w         p
w
2                         w                       w         p
w
3                         w                       w         p
w
4                         w                       w         p
w
...                     ...                     ...       ...        ..
.
8119                      o                       o         p
o
8120                      o                       o         p
n
8121                      o                       o         p
o
8122                      w                       w         p
w
8123                      o                       o         p
o

     ring-number ring-type spore-print-color population habitat
0              o         p                 k          s       u
1              o         p                 n          n       g
2              o         p                 n          n       m
3              o         p                 k          s       u
4              o         e                 n          a       g
...          ...       ...               ...        ...     ...
8119           o         p                 b          c       l
8120           o         p                 b          v       l
8121           o         p                 b          c       l
8122           o         e                 w          v       l
8123           o         p                 o          c       l

[8124 rows x 23 columns]

dt.duplicated().sum()
```

```
0
```

```
dt.isnull().sum()
```

```
class                       0
cap-shape                   0
cap-surface                 0
cap-color                   0
bruises                     0
odor                        0
gill-attachment             0
gill-spacing                0
gill-size                   0
gill-color                  0
stalk-shape                 0
stalk-root                  0
stalk-surface-above-ring    0
stalk-surface-below-ring    0
stalk-color-above-ring      0
stalk-color-below-ring      0
veil-type                   0
veil-color                  0
ring-number                 0
ring-type                   0
spore-print-color           0
population                  0
habitat                     0
dtype: int64
```

```
dt.info()
```

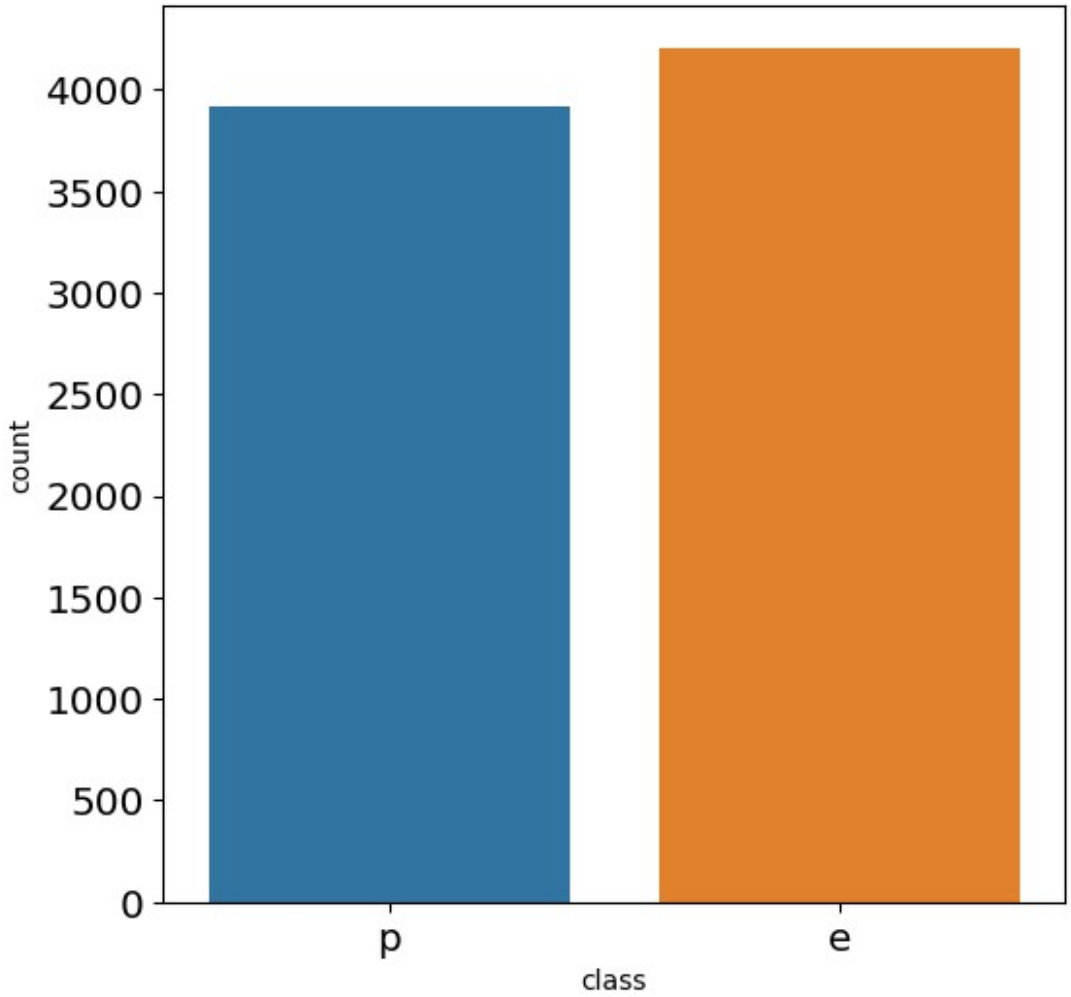
```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 8124 entries, 0 to 8123
Data columns (total 23 columns):
 #   Column                    Non-Null Count  Dtype
---  ------                    --------------  -----
 0   class                     8124 non-null   object
 1   cap-shape                 8124 non-null   object
 2   cap-surface               8124 non-null   object
 3   cap-color                 8124 non-null   object
 4   bruises                   8124 non-null   object
 5   odor                      8124 non-null   object
 6   gill-attachment           8124 non-null   object
 7   gill-spacing              8124 non-null   object
 8   gill-size                 8124 non-null   object
 9   gill-color                8124 non-null   object
 10  stalk-shape               8124 non-null   object
 11  stalk-root                8124 non-null   object
 12  stalk-surface-above-ring  8124 non-null   object
 13  stalk-surface-below-ring  8124 non-null   object
 14  stalk-color-above-ring    8124 non-null   object
```

```
 15  stalk-color-below-ring    8124 non-null   object
 16  veil-type                 8124 non-null   object
 17  veil-color                8124 non-null   object
 18  ring-number               8124 non-null   object
 19  ring-type                 8124 non-null   object
 20  spore-print-color         8124 non-null   object
 21  population                8124 non-null   object
 22  habitat                   8124 non-null   object
dtypes: object(23)
memory usage: 1.4+ MB

dt.describe()

       class cap-shape cap-surface cap-color bruises  odor gill-
attachment  \
count   8124      8124        8124      8124    8124  8124
8124
unique     2         6           4        10       2     9
2
top        e         x           y         n       f     n
f
freq    4208      3656        3244      2284    4748  3528
7914

       gill-spacing gill-size gill-color  ... stalk-surface-below-ring
\
count          8124      8124       8124  ...                      8124

unique            2         2         12  ...                         4

top               c         b          b  ...                         s

freq           6812      5612       1728  ...                      4936


       stalk-color-above-ring stalk-color-below-ring veil-type veil-
color  \
count                    8124                    8124      8124
8124
unique                      9                       9         1
4
top                         w                       w         p
w
freq                     4464                    4384      8124
7924

       ring-number ring-type spore-print-color population habitat
count         8124      8124              8124       8124    8124
unique           3         5                 9          6       7
top              o         p                 w          v       d
```

```
freq          7488      3968                2388       4040     3148

[4 rows x 23 columns]
```

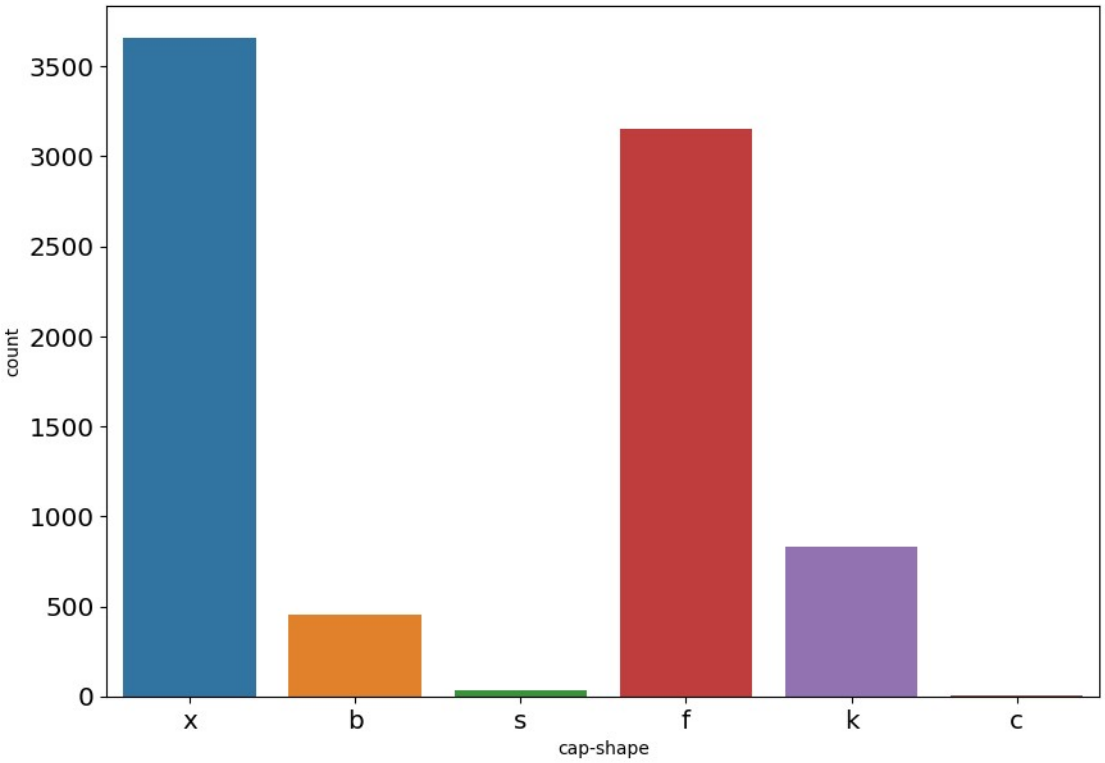
```
plt.figure(figsize = (6, 6))
mushrooms = ['poisonous','edible']
count = [len(dt[dt['class'] == 'p']), len(dt[dt['class'] == 'e'])]
colors = sns.color_palette()[1 : 10]

sns.countplot(data = dt,x = 'class')
```

```
<Axes: xlabel='class', ylabel='count'>
```



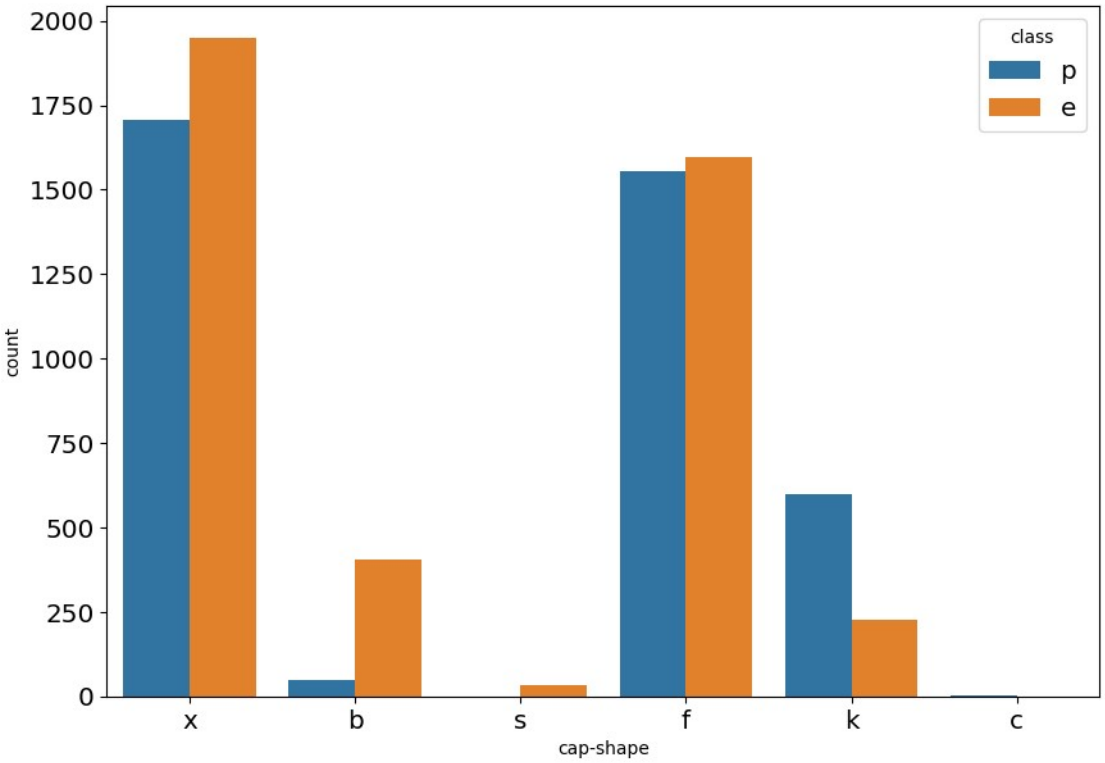

```
fig, ax = plt.subplots(figsize=(10,7))
sns.countplot(data=dt,x='cap-shape')
```

```
<Axes: xlabel='cap-shape', ylabel='count'>
```

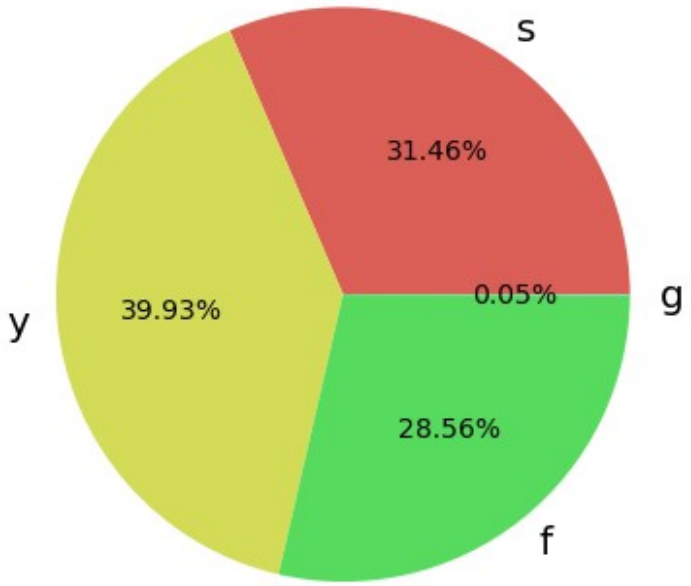



```
fig, ax = plt.subplots(figsize = (10, 7))
sns.countplot(data = dt, x = 'cap-shape', hue = 'class')
```

```
<Axes: xlabel='cap-shape', ylabel='count'>
```

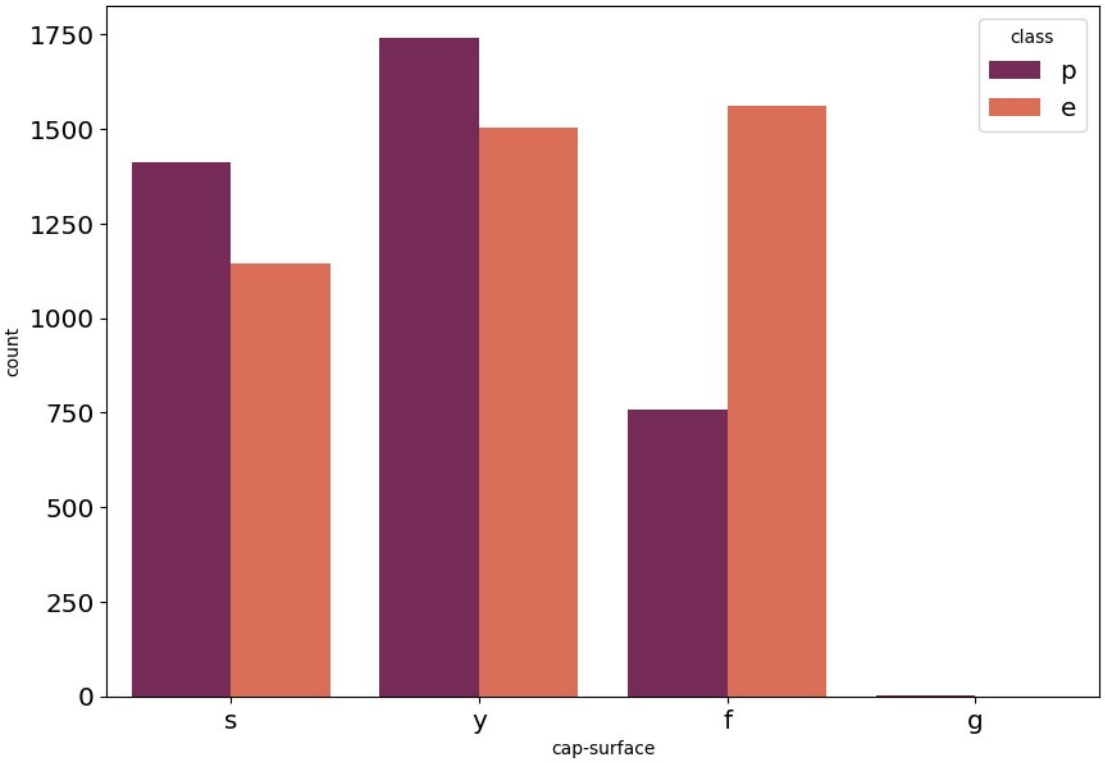



```
mushrooms = ['s','y','f','g']
count = [len(dt[dt['cap-surface'] == 's']), len(dt[dt['cap-surface']
== 'y']), len(dt[dt['cap-surface'] == 'f']), len(dt[dt['cap-surface']
== 'g'])]
colors = sns.color_palette("hls")
plt.pie(count, labels = mushrooms, colors = colors, autopct = '%1.2f%
%')
plt.show()
```



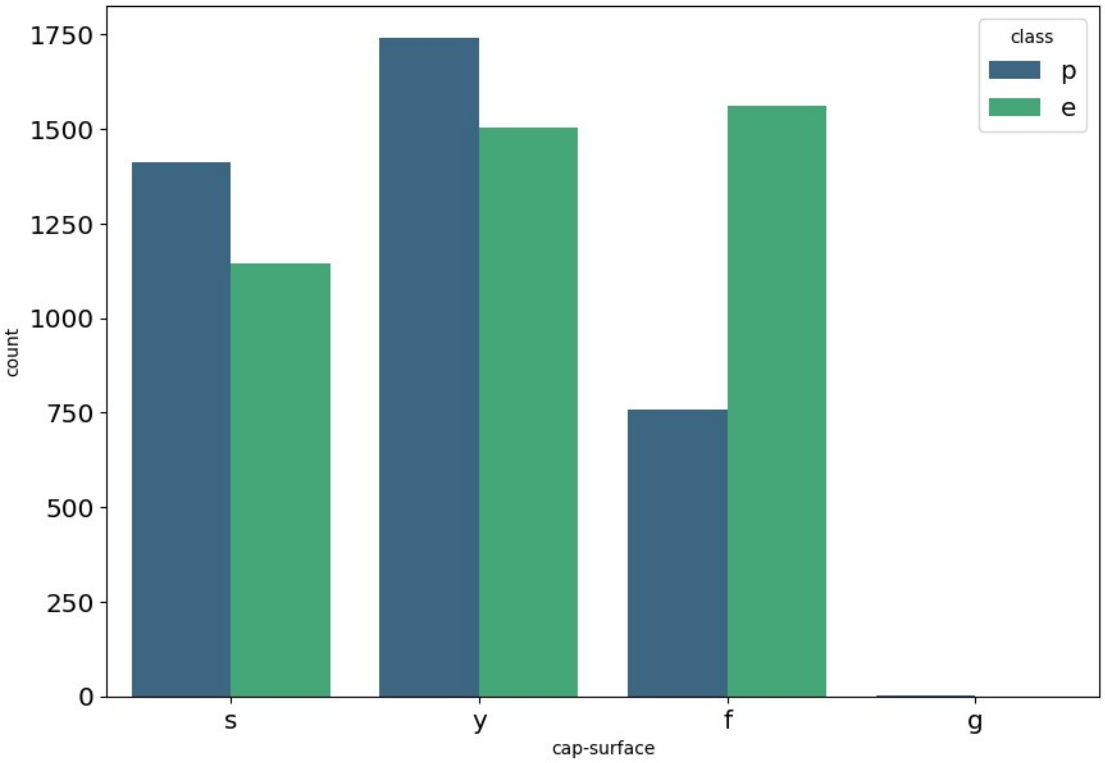

```
fig, ax = plt.subplots(figsize = (10, 7))
sns.countplot(data = dt, x = 'cap-surface', hue = 'class', palette =
'rocket')
```

```
<Axes: xlabel='cap-surface', ylabel='count'>
```



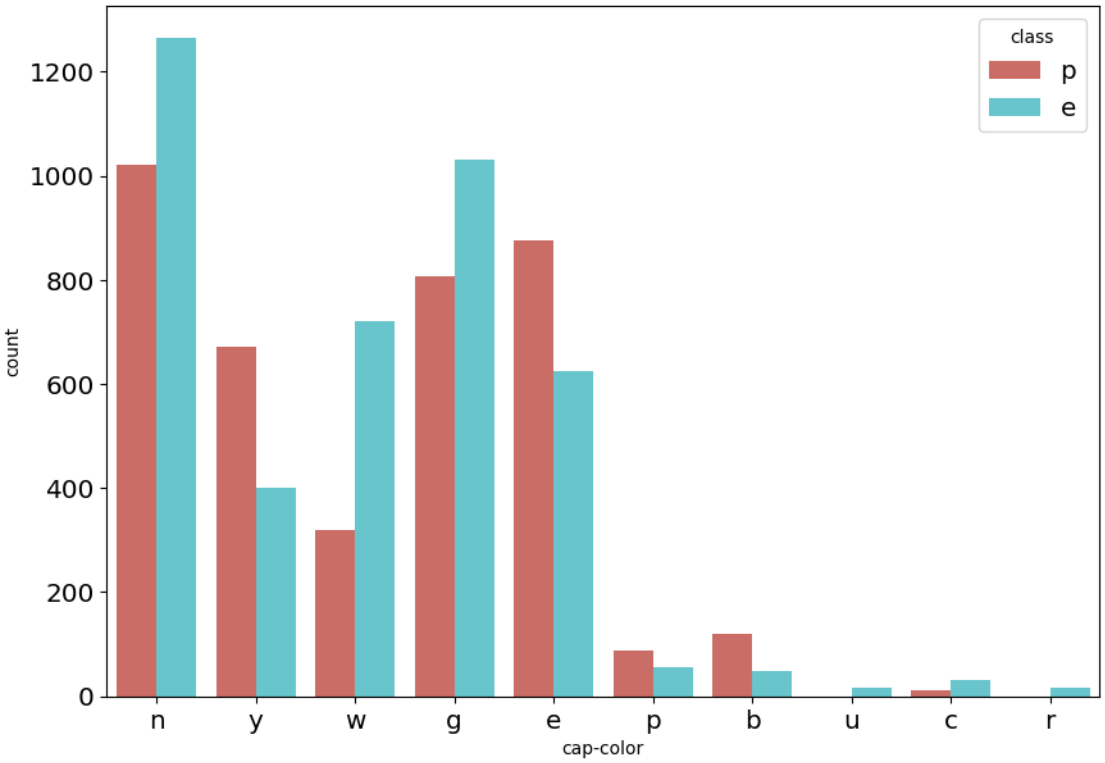

```
fig, ax = plt.subplots(figsize = (10, 7))
sns.countplot(data = dt, x = 'cap-surface', hue = 'class', palette =
'viridis')
```

```
<Axes: xlabel='cap-surface', ylabel='count'>
```




```
fig, ax = plt.subplots(figsize = (10, 7))
sns.countplot(data = dt, x = 'cap-color', hue = 'class', palette =
'hls')
```

```
<Axes: xlabel='cap-color', ylabel='count'>
```

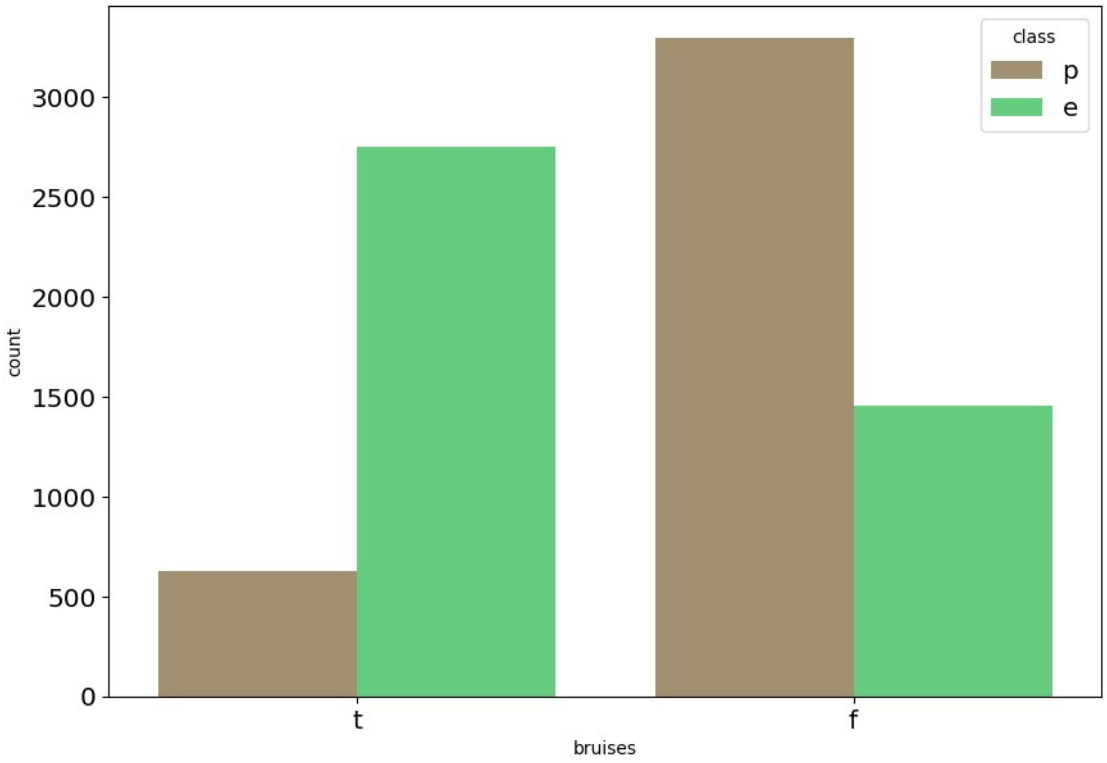



```
dt.groupby('cap-color')['class'].value_counts()
```

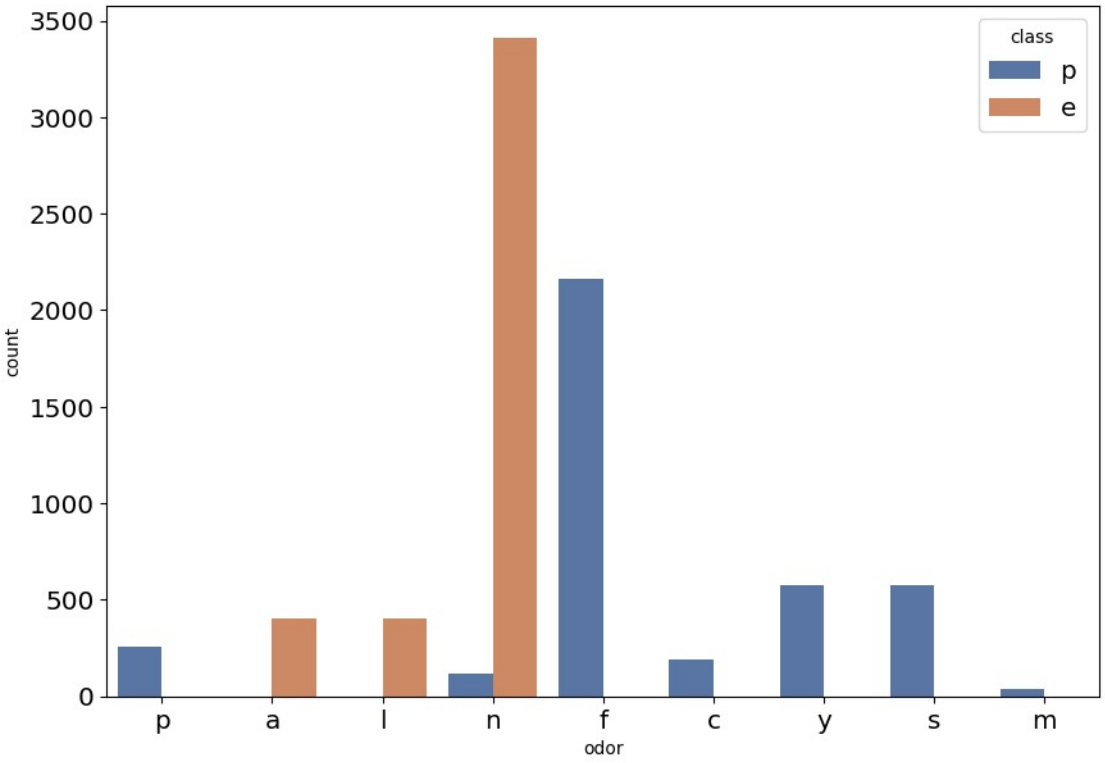
```
cap-color  class
b          p          120
           e           48
c          e           32
           p           12
e          p          876
           e          624
g          e         1032
           p          808
n          e         1264
           p         1020
p          p           88
           e           56
r          e           16
u          e           16
w          e          720
           p          320
y          p          672
           e          400
Name: class, dtype: int64
```

```
fig, ax = plt.subplots(figsize = (10, 7))
sns.countplot(data = dt, x = 'bruises', hue = 'class', palette =
'terrain_r')
```

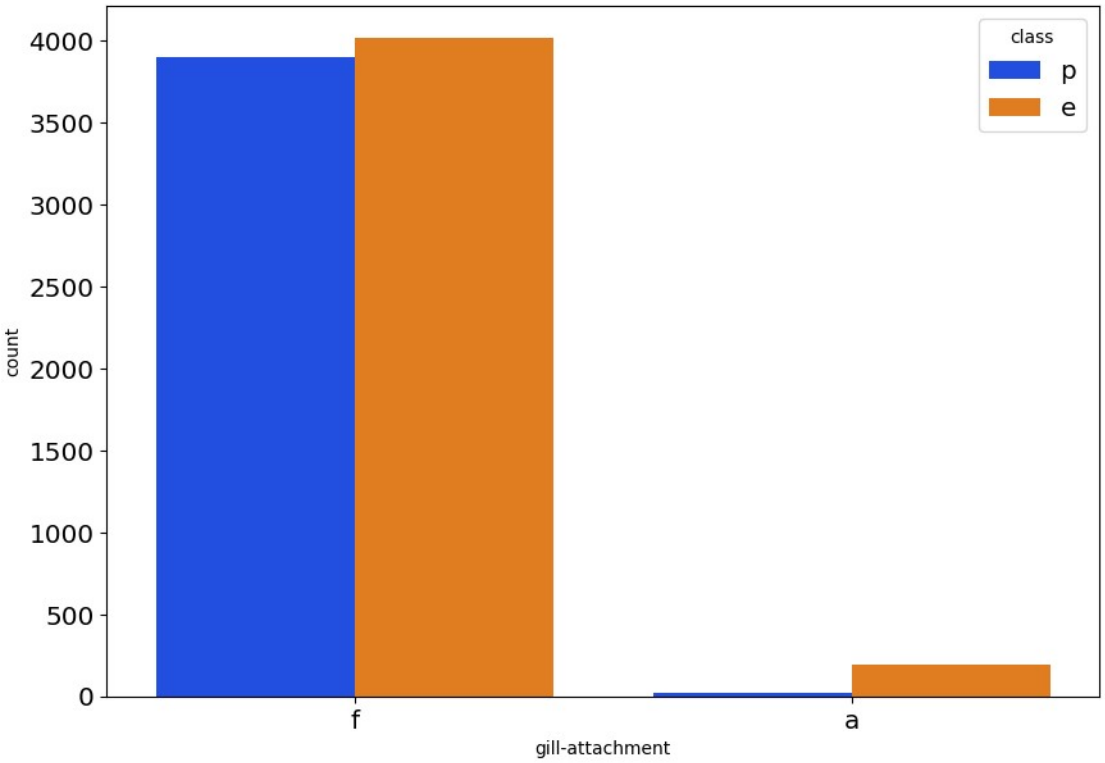
```
<Axes: xlabel='bruises', ylabel='count'>
```



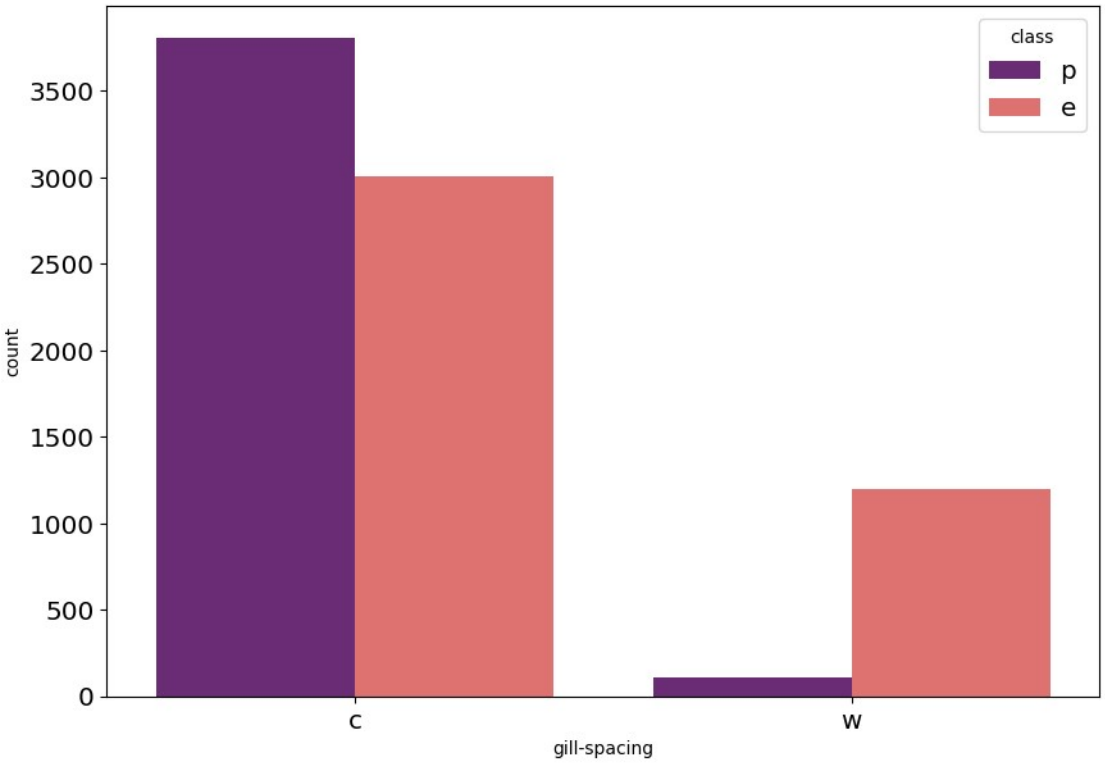

```
fig, ax = plt.subplots(figsize = (10, 7))
sns.countplot(data = dt, x = 'odor', hue = 'class', palette = 'deep')
plt.show()
```



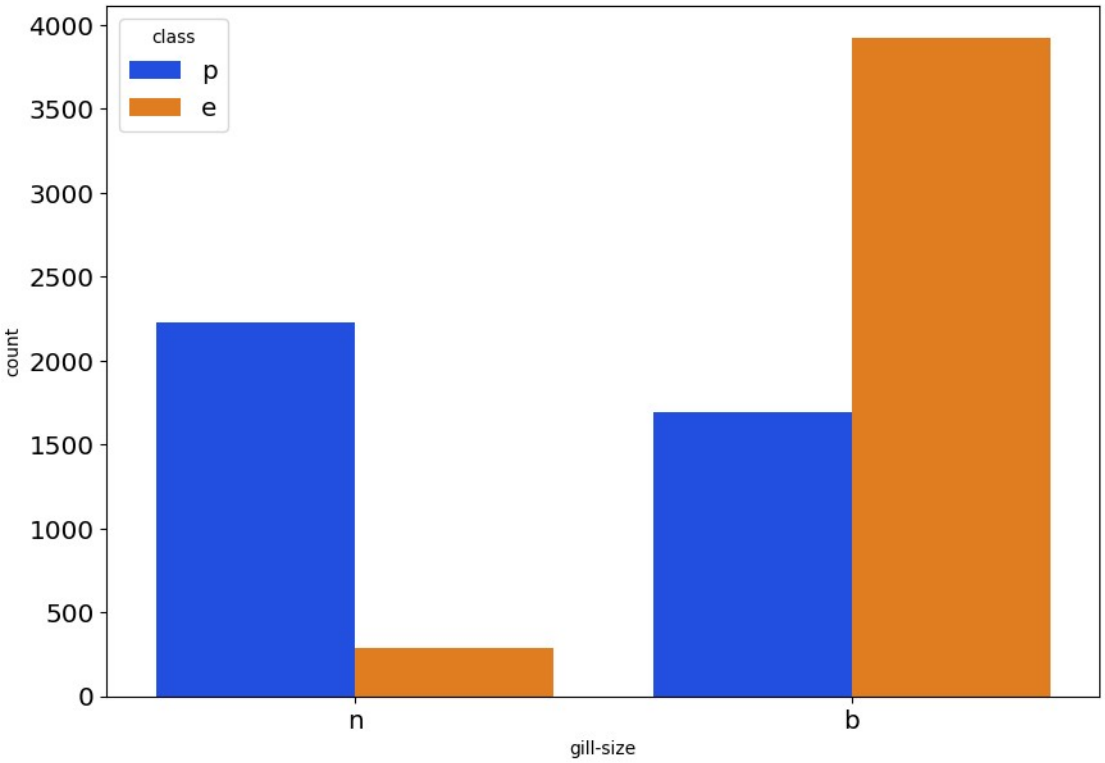

```
fig, ax = plt.subplots(figsize = (10, 7))
sns.countplot(data = dt, x = 'gill-attachment', hue = 'class', palette
= 'bright')
plt.show()
```



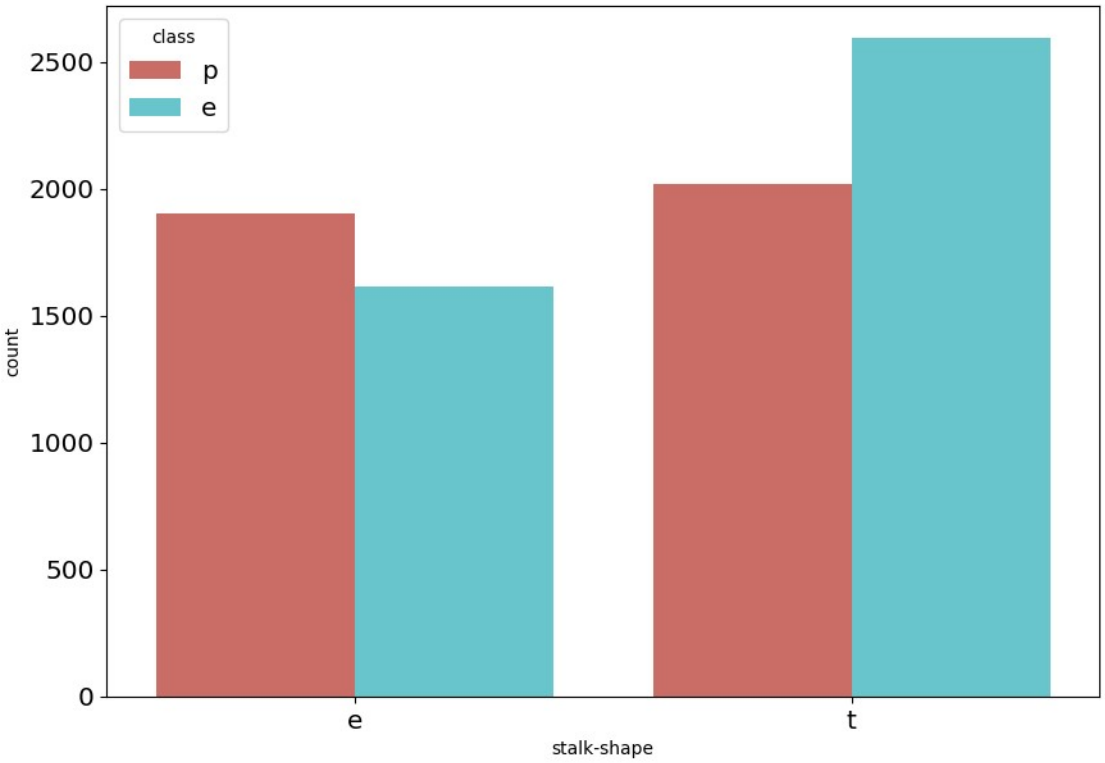

```
fig, ax = plt.subplots(figsize = (10, 7))
sns.countplot(data = dt, x = 'gill-spacing', hue = 'class', palette =
'magma')
plt.show()
```



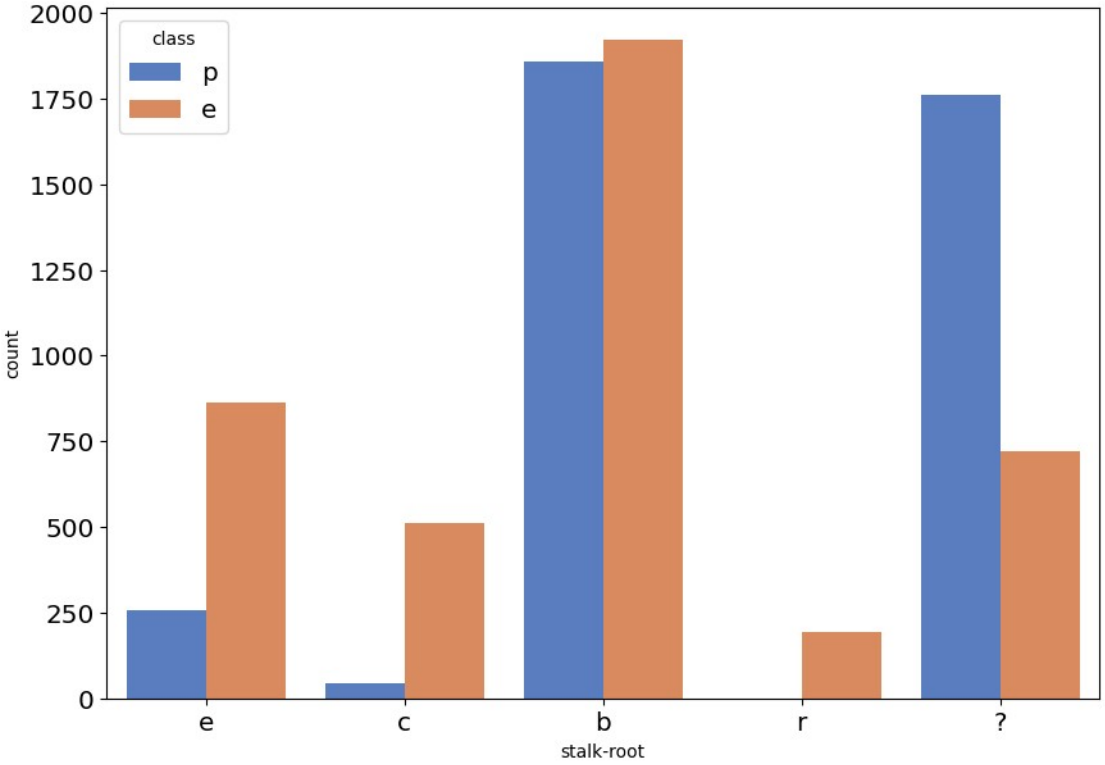

```
fig, ax = plt.subplots(figsize = (10, 7))
sns.countplot(data = dt, x = 'gill-size', hue = 'class', palette =
'bright')
plt.show()
```



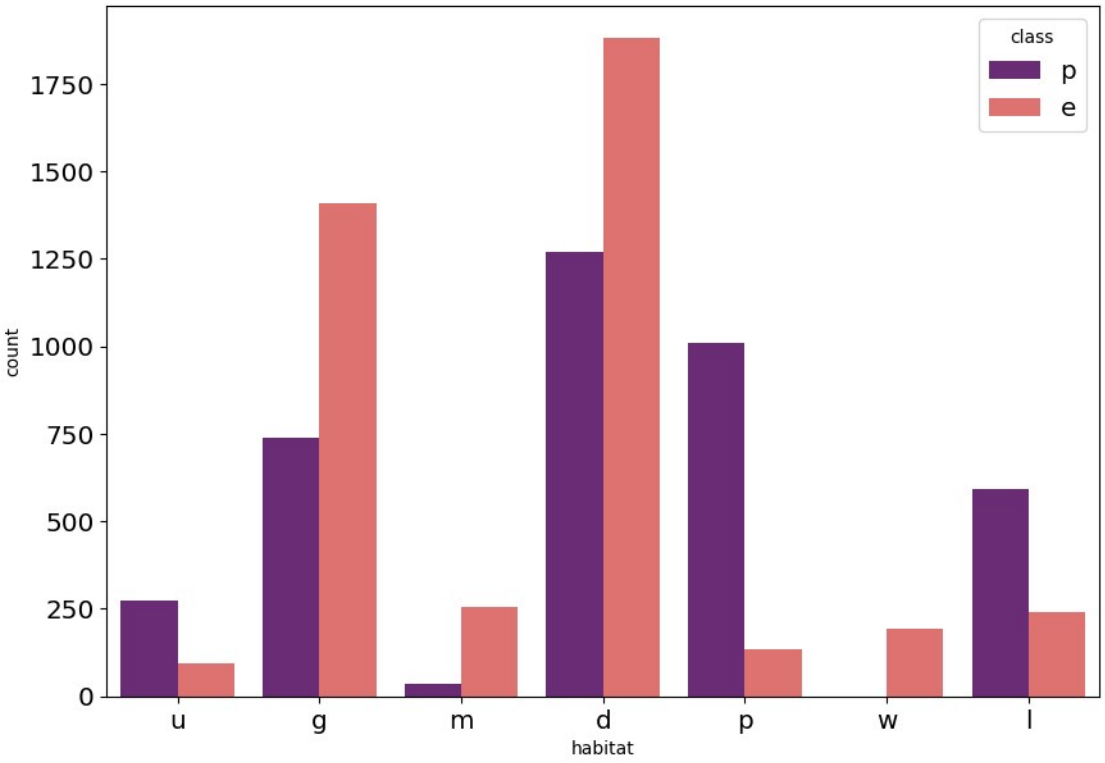

```
fig, ax = plt.subplots(figsize = (10, 7))
sns.countplot(data = dt, x = 'stalk-shape', hue = 'class', palette =
'hls')
plt.show()
```




```
fig, ax = plt.subplots(figsize = (10, 7))
sns.countplot(data = dt, x = 'stalk-root', hue = 'class', palette =
'muted')
plt.show()
```

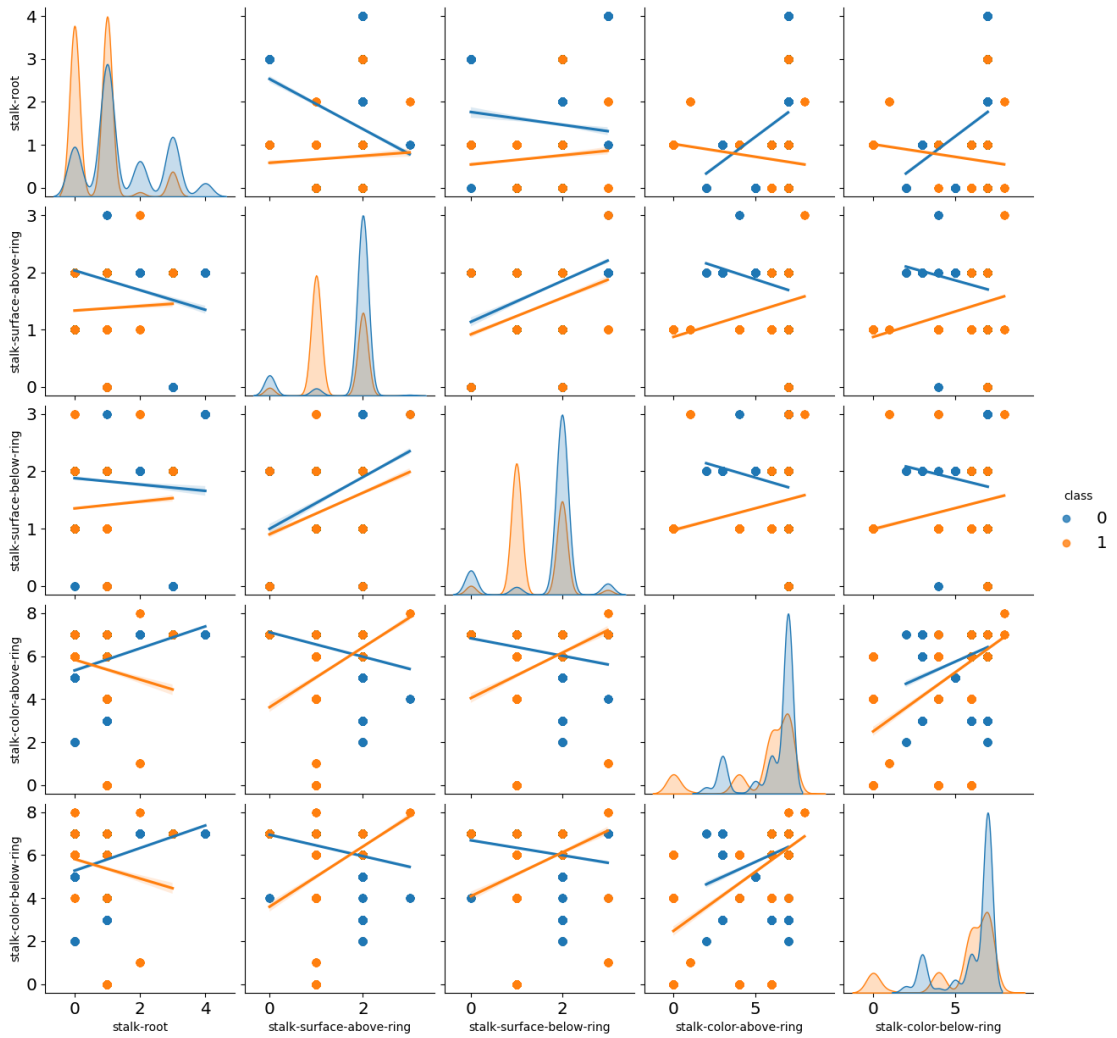



```
fig, ax = plt.subplots(figsize = (10, 7))
sns.countplot(data = dt,x = 'habitat', hue = 'class', palette =
'magma')
plt.show()
```




```
stalk_cats = ['class', 'stalk-root', 'stalk-surface-above-ring',
'stalk-surface-below-ring',
              'stalk-color-above-ring', 'stalk-color-below-ring']
data_cats = data_encoded[stalk_cats]
sns.pairplot(data_cats, hue='class', kind='reg')
```

```
<seaborn.axisgrid.PairGrid at 0x7fdfcce58820>
```

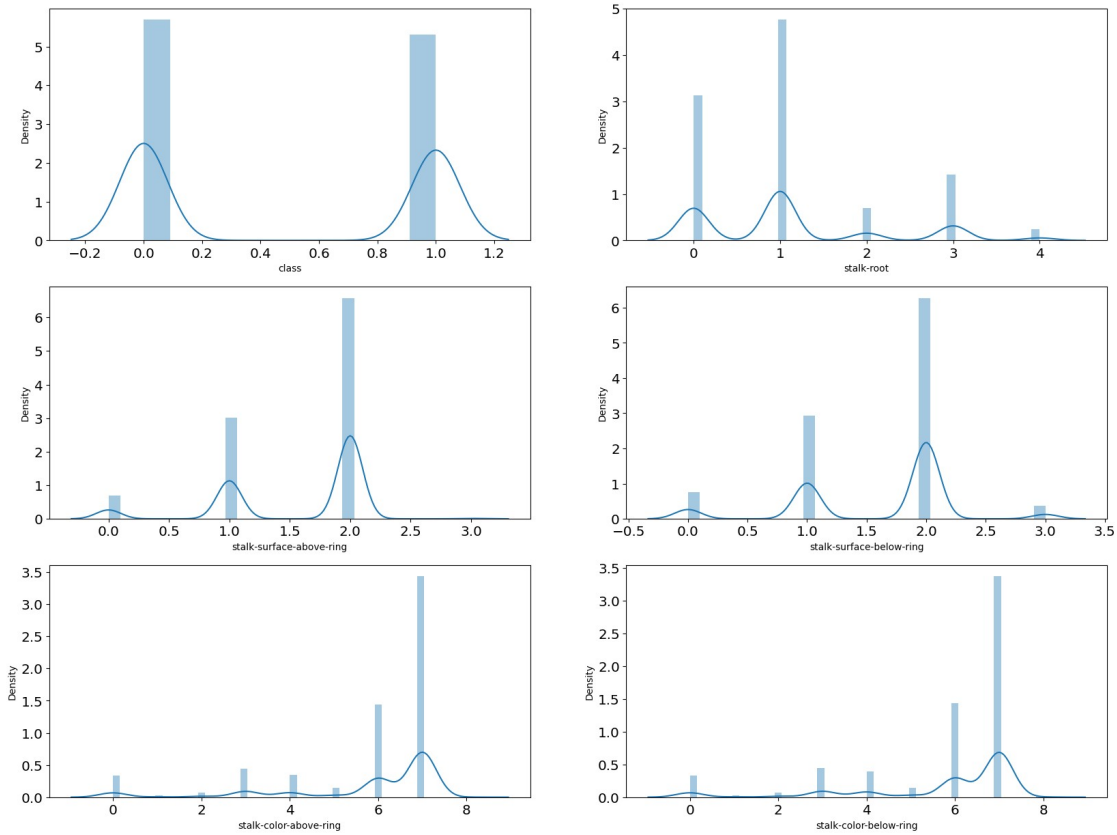



```
fig, ax = plt.subplots(3, 2, figsize=(20, 15))
for i, axis in enumerate(ax.flat):
    sns.distplot(data_cats.iloc[:, i], ax=axis)
```

```
/var/folders/nz/v4c3dsmn1293s89vq32wmzq00000gn/T/
ipykernel_14222/1659700951.py:3: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn
v0.14.0.

Please adapt your code to use either `displot` (a figure-level
function with
similar flexibility) or `histplot` (an axes-level function for
histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(data_cats.iloc[:, i], ax=axis)
```

```
/var/folders/nz/v4c3dsmn1293s89vq32wmzq00000gn/T/ipykernel_14222/16597
00951.py:3: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn
v0.14.0.

Please adapt your code to use either `displot` (a figure-level
function with
similar flexibility) or `histplot` (an axes-level function for
histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(data_cats.iloc[:, i], ax=axis)
/var/folders/nz/v4c3dsmn1293s89vq32wmzq00000gn/T/ipykernel_14222/16597
00951.py:3: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn
v0.14.0.

Please adapt your code to use either `displot` (a figure-level
function with
similar flexibility) or `histplot` (an axes-level function for
histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(data_cats.iloc[:, i], ax=axis)
/var/folders/nz/v4c3dsmn1293s89vq32wmzq00000gn/T/ipykernel_14222/16597
00951.py:3: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn
v0.14.0.

Please adapt your code to use either `displot` (a figure-level
function with
similar flexibility) or `histplot` (an axes-level function for
histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(data_cats.iloc[:, i], ax=axis)
/var/folders/nz/v4c3dsmn1293s89vq32wmzq00000gn/T/ipykernel_14222/16597
00951.py:3: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn
```

```
v0.14.0.

Please adapt your code to use either `displot` (a figure-level
function with
similar flexibility) or `histplot` (an axes-level function for
histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(data_cats.iloc[:, i], ax=axis)
/var/folders/nz/v4c3dsmn1293s89vq32wmzq00000gn/T/ipykernel_14222/16597
00951.py:3: UserWarning:

`distplot` is a deprecated function and will be removed in seaborn
v0.14.0.

Please adapt your code to use either `displot` (a figure-level
function with
similar flexibility) or `histplot` (an axes-level function for
histograms).

For a guide to updating your code to use the new functions, please see
https://gist.github.com/mwaskom/de44147ed2974457ad6372750bbe5751

  sns.distplot(data_cats.iloc[:, i], ax=axis)
```




```
columns=data.columns
names=enames+pnames
f,ax = plt.subplots(figsize = (12,16))
sns.barplot(x=erates,y=columns,color='lime',alpha =
0.8,label='Edible')
sns.barplot(x=prates,y=columns,color='green',alpha =
0.4,label='Poisonous')
#for p in ax.patches:
    #ax.annotate('{}'.format(p.get_width()), (p.get_y()+0.1,
p.get_height()+1))
sayac=0
itere=0
for p in ax.patches:
    width=p.get_width()
    #print(p.get_width())
    if sayac%2==0:
        a=5
        clr = 'black'
    else:
        a=25
        clr = 'blue'
    k=names[itere].capitalize()
    plt.text(a, p.get_y()+0.55*p.get_height(),'{}:
{:1.2f}'.format(k,width),color='black', va='center')
    itere=itere+1
```

```
    sayac=sayac+1
ax.legend(loc='lower right',frameon = True)     # legendlarin
gorunurlugu
ax.set(xlabel='Rates', ylabel='Columns',title = "Most frequently
encountered attributes")
plt.show()
```

```
---------------------------------------------------------------------------
-----
NameError                                 Traceback (most recent call
last)
Cell In[163], line 2
      1 columns=data.columns
----> 2 names=enames+pnames
      3 f,ax = plt.subplots(figsize = (12,16))
      4 sns.barplot(x=erates,y=columns,color='lime',alpha =
0.8,label='Edible')

NameError: name 'enames' is not defined
```