In [27]:
```
import pandas as pd
import numpy as np


import seaborn as sns
import matplotlib.pyplot as plt
%matplotlib inline

import warnings
warnings.filterwarnings("ignore")
```

In [28]:
```
features=["Target","cap-shape","cap-surface","cap-color","bruises?","odor","gill-attachment","gill-spacing","gill-size",
          "gill-color","stalk-shape","stalk-root","stalk-surface-above-ring","stalk-surface-below-ring",
          "stalk-color-above-ring","stalk-color-below-ring","veil-type","veil-color","ring-number","ring-type",
          "spore-print-color","population","habitat"]

df = pd.read_csv('agaricus-lepiota.data',names=features)
df
```

Out[28]:

	Target	cap-shape	cap-surface	cap-color	bruises?	odor	gill-attachment	gill-spacing	gill-size	gill-color	...	stalk-surface-below-ring	stalk-color-above-ring	stalk-color-below-ring	veil-type	veil-color	ring-number	ring-type	spore-print-color	population	habitat
0	p	x	s	n	t	p	f	c	n	k	...	s	w	w	p	w	o	p	k	s	u
1	e	x	s	y	t	a	f	c	b	k	...	s	w	w	p	w	o	p	n	n	g
2	e	b	s	w	t	l	f	c	b	n	...	s	w	w	p	w	o	p	n	n	m
3	p	x	y	w	t	p	f	c	n	n	...	s	w	w	p	w	o	p	k	s	u
4	e	x	s	g	f	n	f	w	b	k	...	s	w	w	p	w	o	e	n	a	g
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
8119	e	k	s	n	f	n	a	c	b	y	...	s	o	o	p	o	o	p	b	c	l
8120	e	x	s	n	f	n	a	c	b	y	...	s	o	o	p	n	o	p	b	v	l
8121	e	f	s	n	f	n	a	c	b	n	...	s	o	o	p	o	o	p	b	c	l
8122	p	k	y	n	f	y	f	c	n	b	...	k	w	w	p	w	o	e	w	v	l
8123	e	x	s	n	f	n	a	c	b	y	...	s	o	o	p	o	o	p	o	c	l

8124 rows × 23 columns

In [29]:
```
df.to_csv('agaricus-lepiota.csv',sep=',',index=False)
```

In [ ]:
```
'''This data set includes descriptions of hypothetical samples corresponding to 23 species of gilled mushrooms in
the Agaricus and Lepiota Family (pp. 500-525).  Each species is identified as definitely edible, definitely
poisonous, or of unknown edibility and not recommended.  This latter class was combined with the poisonous one.
The Guide clearly states that there is no simple rule for determining the edibility of a mushroom;
no rule like ``leaflets three, let it be'' for Poisonous Oak and Ivy.'''
```

In [30]:
```
df['Target'].value_counts()
```

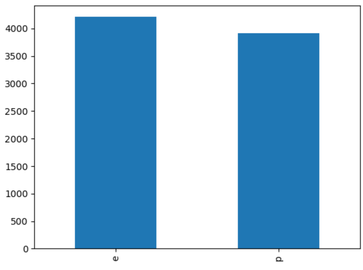
Out[30]:
```
e    4208
p    3916
Name: Target, dtype: int64
```

In [31]:
```
df['Target'].value_counts(normalize=True)
```

Out[31]:
```
e    0.517971
p    0.482029
Name: Target, dtype: float64
```

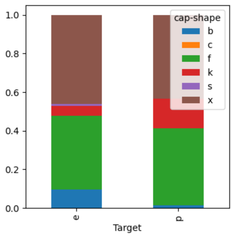
In [32]:
```
df['Target'].value_counts().plot.bar()
```

Out[32]:
```
<Axes: >
```



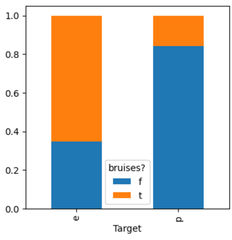
In [33]:
```
Target=pd.crosstab(df['Target'],df['cap-shape'])
Target.div(Target.sum(1).astype(float), axis=0).plot(kind="bar", stacked=True, figsize=(4,4))
```

Out[33]:
```
<Axes: xlabel='Target'>
```



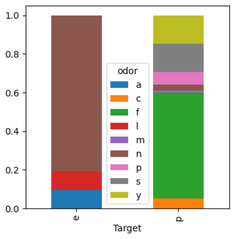
In [34]:
```
Target=pd.crosstab(df['Target'],df['bruises?'])
Target.div(Target.sum(1).astype(float), axis=0).plot(kind="bar", stacked=True, figsize=(4,4))
```

Out[34]:
```
<Axes: xlabel='Target'>
```



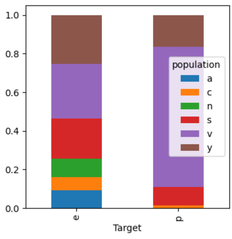
In [35]:
```
Target=pd.crosstab(df['Target'],df['odor'])
Target.div(Target.sum(1).astype(float), axis=0).plot(kind="bar", stacked=True, figsize=(4,4))
```

Out[35]:
```
<Axes: xlabel='Target'>
```



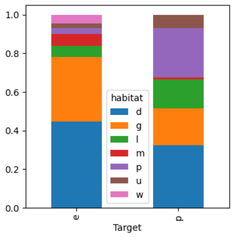
In [36]:
```
Target=pd.crosstab(df['Target'],df['population'])
Target.div(Target.sum(1).astype(float), axis=0).plot(kind="bar", stacked=True, figsize=(4,4))
```

Out[36]:
```
<Axes: xlabel='Target'>
```



In [37]:
```
Target=pd.crosstab(df['Target'],df['habitat'])
Target.div(Target.sum(1).astype(float), axis=0).plot(kind="bar", stacked=True, figsize=(4,4))
```

Out[37]:
```
<Axes: xlabel='Target'>
```



In [ ]:
```
```