In [ ]:

In [1]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
%matplotlib inline

import seaborn as sns
```

In [ ]:

In [2]:
```
titanic = pd.read_csv("titanic.csv")
titanic.shape
```

Out[2]:
```
(712, 7)
```

In [3]:
```
titanic.head()
```

Out[3]:

	pclass	gender	age	sibling	fare	embark_town	survived
0	3	male	22.0	1	7.2500	Southampton	no
1	1	female	38.0	1	71.2833	Cherbourg	yes
2	3	female	26.0	0	7.9250	Southampton	yes
3	1	female	35.0	1	53.1000	Southampton	yes
4	3	male	35.0	0	8.0500	Southampton	no

In [4]:
```
titanic.info()
```

```
<class 'pandas.core.frame.DataFrame'>
RangeIndex: 712 entries, 0 to 711
Data columns (total 7 columns):
pclass          712 non-null int64
gender          712 non-null object
age             712 non-null float64
sibling         712 non-null int64
fare            712 non-null float64
embark_town     712 non-null object
survived        712 non-null object
dtypes: float64(2), int64(2), object(3)
memory usage: 39.0+ KB
```

In [6]:
```
titanic.describe(include = "all")
```

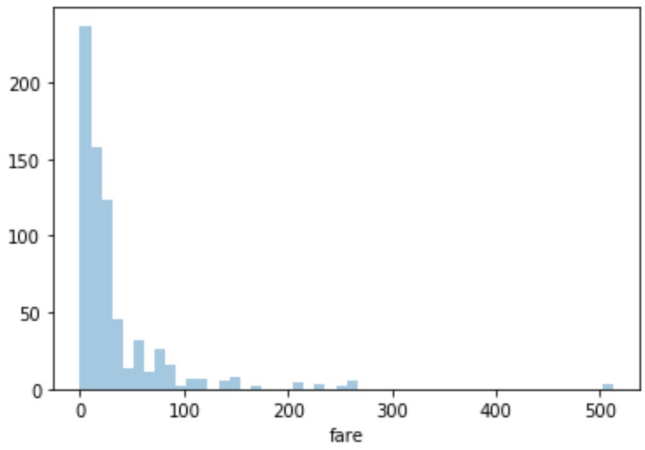
Out[6]:

	pclass	gender	age	sibling	fare	embark_town	survived
count	712.000000	712	712.000000	712.000000	712.000000	712	712
unique	NaN	2	NaN	NaN	NaN	3	2
top	NaN	male	NaN	NaN	NaN	Southampton	no
freq	NaN	453	NaN	NaN	NaN	554	424
mean	2.240169	NaN	29.642093	0.514045	34.567251	NaN	NaN
std	0.836854	NaN	14.492933	0.930692	52.938648	NaN	NaN
min	1.000000	NaN	0.420000	0.000000	0.000000	NaN	NaN
25%	1.000000	NaN	20.000000	0.000000	8.050000	NaN	NaN
50%	2.000000	NaN	28.000000	0.000000	15.645850	NaN	NaN
75%	3.000000	NaN	38.000000	1.000000	33.000000	NaN	NaN
max	3.000000	NaN	80.000000	5.000000	512.329200	NaN	NaN

In [ ]:
```
```

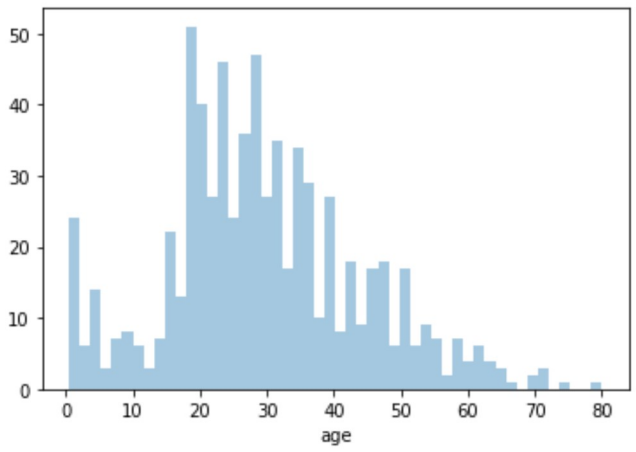
In [8]:
```
# distribution plot
sns.distplot(titanic['fare'], bins = 50, kde = False)
```

Out[8]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x201bbc808d0>
```



In [9]:
```
sns.distplot(titanic['age'], bins = 50, kde = False)
```

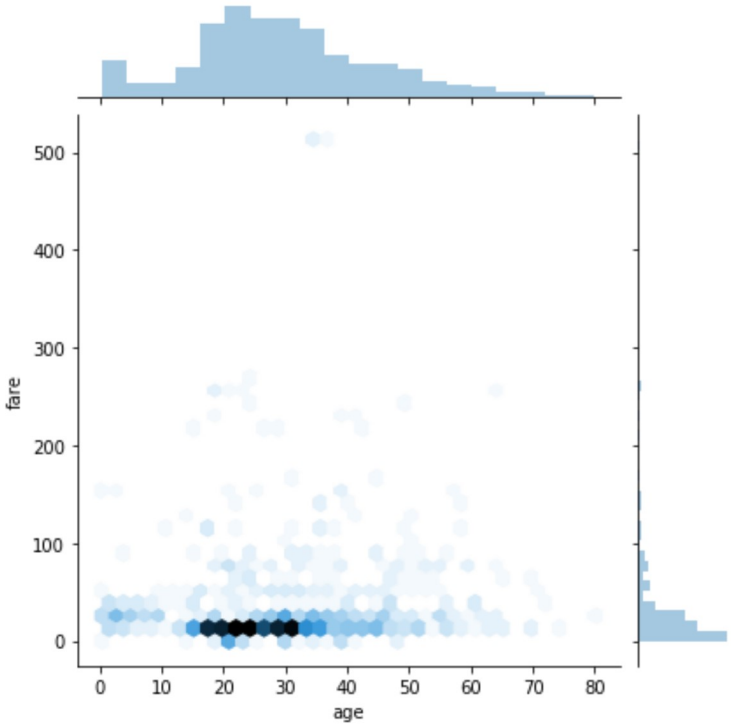
Out[9]:
```
<matplotlib.axes._subplots.AxesSubplot at 0x201bbd6ca58>
```



In [11]:
```
# compare numerical

sns.jointplot(x = titanic['age'], y = titanic['fare'], kind = 'hex')
```

Out[11]: <seaborn.axisgrid.JointGrid at 0x201beafb198>

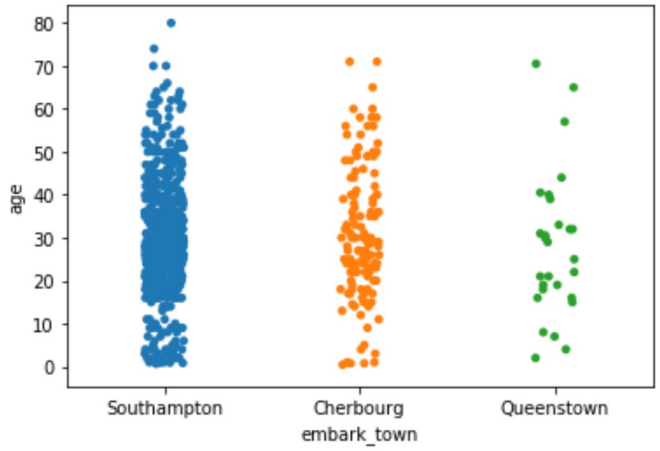


In [ ]:

In [ ]:

In [20]:
```
# categorical and numerical
# boxplot, stripplot, violinplot

sns.stripplot(x = titanic['embark_town'], y = titanic["age"])
plt.show()
```



In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]:

In [ ]: