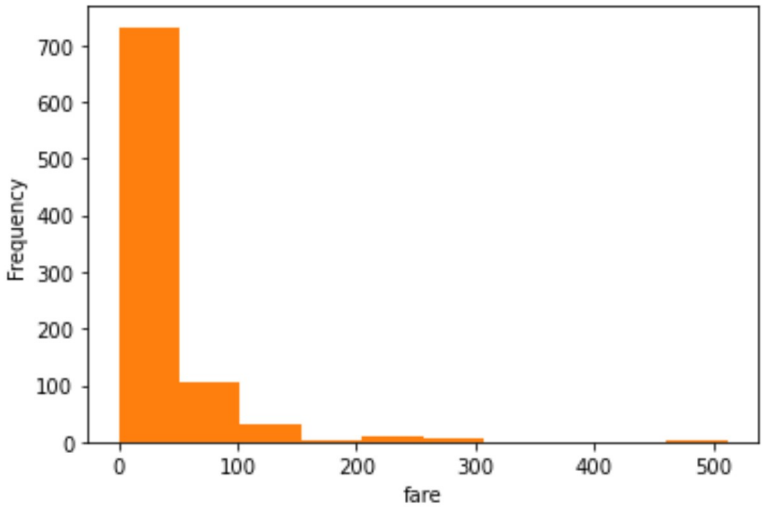
In [7]:

```
#!/usr/bin/env python
# coding: utf-8

# Project 3 - Titanic Information

## SDS348 Spring 2021

### Armando Marquez am8245

# ### Describe Data



#import pandas as pd
import pandas as pd
#import numpy as np
import numpy as np

#import matplot.pyplot as plt
import matplotlib.pyplot as plt
#import seaborn as sns
import seaborn as sns
#import dataset
dataset = sns.load_dataset('titanic')
#viewdataset
print(dataset.head())

# find out the number of rows and columns of data
print(dataset.shape)

#This has 891 rows and 15 columns.

# Select variables (with `filter`), filter only one color (with `query`) and
print(dataset.filter(['survived', 'age']) .query('survived == "1"') .agg(['mea

# Select variables (with `filter`), filter only one color (with `query`) and
print(dataset.filter(['sex', 'survived']) .query('sex == "male"') .agg(['mean

#Of the males that survived, the mean was 0.188908 and the std was 0.391775.



# Create a histogram
dataset['fare'].plot(kind = "hist")
plt.xlabel('fare') # add a label

#Of those that survived, the highest group was 20 year olds. There were also
```

```
   survived  pclass     sex   age  sibsp  parch     fare embarked  class  \
0         0       3    male  22.0      1      0   7.2500        S  Third
1         1       1  female  38.0      1      0  71.2833        C  First
2         1       3  female  26.0      0      0   7.9250        S  Third
3         1       1  female  35.0      1      0  53.1000        S  First
4         0       3    male  35.0      0      0   8.0500        S  Third

     who  adult_male deck  embark_town alive  alone
0    man        True  NaN  Southampton    no  False
1  woman       False    C    Cherbourg   yes  False
2  woman       False  NaN  Southampton   yes   True
```

```
3  woman       False    C  Southampton   yes  False
4    man        True  NaN  Southampton    no   True
(891, 15)
      survived        age
mean       1.0  28.343690
std        0.0  14.950952
      survived
mean  0.188908
std   0.391775
AxesSubplot(0.125,0.125;0.775x0.755)
Text(0.5, 0, 'age')
```

Out[7]: `Text(0.5, 0, 'fare')`



In [8]:
```
# Create a histogram
dataset['age'].plot(kind = "hist")
plt.xlabel('age') # add a label)
```

Out[8]: `Text(0.5, 0, 'age')`