In [1]:
```
import numpy as np
import pandas as pd
```

In [2]:
```
x=pd.read_csv(r"C:\Users\user\Downloads\uber - uber.csv")
x
```

Out[2]:

	Unnamed: 0	key	fare_amount	pickup_datetime	pickup_longitude	pickup_latitude	dro
0	24238194	2015-05-07 19:52:06	7.5	2015-05-07 19:52:06 UTC	-73.999817	40.738354	
1	27835199	2009-07-17 20:04:56	7.7	2009-07-17 20:04:56 UTC	-73.994355	40.728225	
2	44984355	2009-08-24 21:45:00	12.9	2009-08-24 21:45:00 UTC	-74.005043	40.740770	
3	25894730	2009-06-26 8:22:21	5.3	2009-06-26 08:22:21 UTC	-73.976124	40.790844	
4	17610152	2014-08-28 17:47:00	16.0	2014-08-28 17:47:00 UTC	-73.925023	40.744085	
...	...	...	...	...	...	...	
199995	42598914	2012-10-28 10:49:00	3.0	2012-10-28 10:49:00 UTC	-73.987042	40.739367	
199996	16382965	2014-03-14 1:09:00	7.5	2014-03-14 01:09:00 UTC	-73.984722	40.736837	
199997	27804658	2009-06-29 0:42:00	30.9	2009-06-29 00:42:00 UTC	-73.986017	40.756487	
199998	20259894	2015-05-20 14:56:25	14.5	2015-05-20 14:56:25 UTC	-73.997124	40.725452	
199999	11951496	2010-05-15 4:08:00	14.1	2010-05-15 04:08:00 UTC	-73.984395	40.720077	

200000 rows × 9 columns

In [ ]:
```
```

In [6]:
```
x=x.head(10)
```

In [7]:
```
x.tail(5)
```

Out[7]:

	Unnamed: 0	key	fare_amount	pickup_datetime	pickup_longitude	pickup_latitude	dropoff_
5	44470845	2011-02-12 2:27:09	4.9	2011-02-12 02:27:09 UTC	-73.969019	40.755910	-7
6	48725865	2014-10-12 7:04:00	24.5	2014-10-12 07:04:00 UTC	-73.961447	40.693965	-7
7	44195482	2012-12-11 13:52:00	2.5	2012-12-11 13:52:00 UTC	0.000000	0.000000	
8	15822268	2012-02-17 9:32:00	9.7	2012-02-17 09:32:00 UTC	-73.975187	40.745767	-7
9	50611056	2012-03-29 19:06:00	12.5	2012-03-29 19:06:00 UTC	-74.001065	40.741787	-7

In [8]:
```
x.describe()
```

Out[8]:

	Unnamed: 0	fare_amount	pickup_longitude	pickup_latitude	dropoff_longitude	dropoff_la
count	1.000000e+01	10.000000	10.000000	10.000000	10.000000	10.0
mean	3.443881e+07	10.350000	-66.580708	36.667971	-66.570116	36.6
std	1.342943e+07	6.460693	23.394088	12.883834	23.390384	12.8
min	1.582227e+07	2.500000	-74.005043	0.000000	-74.002720	0.0
25%	2.465233e+07	5.850000	-73.998451	40.730757	-73.989303	40.7
50%	3.601534e+07	8.700000	-73.975656	40.741278	-73.967168	40.7
75%	4.485598e+07	12.800000	-73.963340	40.745346	-73.962684	40.7
max	5.061106e+07	24.500000	0.000000	40.790844	0.000000	40.8

In [ ]:
```
x.index
```

In [13]:
```
x.dtypes
```

Out[13]:
```
Unnamed: 0           int64
key                 object
fare_amount        float64
pickup_datetime     object
pickup_longitude   float64
pickup_latitude    float64
dropoff_longitude  float64
dropoff_latitude   float64
passenger_count      int64
dtype: object
```

In [9]:
```
x['key']
```

Out[9]:
```
0    2015-05-07 19:52:06
1    2009-07-17 20:04:56
2    2009-08-24 21:45:00
3     2009-06-26 8:22:21
4    2014-08-28 17:47:00
5     2011-02-12 2:27:09
6     2014-10-12 7:04:00
7    2012-12-11 13:52:00
8     2012-02-17 9:32:00
9    2012-03-29 19:06:00
Name: key, dtype: object
```

In [10]:
```
x.loc[1:8]
```

Out[10]:

	Unnamed: 0	key	fare_amount	pickup_datetime	pickup_longitude	pickup_latitude	dropoff_
1	27835199	2009-07-17 20:04:56	7.7	2009-07-17 20:04:56 UTC	-73.994355	40.728225	-7
2	44984355	2009-08-24 21:45:00	12.9	2009-08-24 21:45:00 UTC	-74.005043	40.740770	-7
3	25894730	2009-06-26 8:22:21	5.3	2009-06-26 08:22:21 UTC	-73.976124	40.790844	-7
4	17610152	2014-08-28 17:47:00	16.0	2014-08-28 17:47:00 UTC	-73.925023	40.744085	-7
5	44470845	2011-02-12 2:27:09	4.9	2011-02-12 02:27:09 UTC	-73.969019	40.755910	-7
6	48725865	2014-10-12 7:04:00	24.5	2014-10-12 07:04:00 UTC	-73.961447	40.693965	-7
7	44195482	2012-12-11 13:52:00	2.5	2012-12-11 13:52:00 UTC	0.000000	0.000000	
8	15822268	2012-02-17 9:32:00	9.7	2012-02-17 09:32:00 UTC	-73.975187	40.745767	-7

In [11]:
```
x.iloc[1:7]
```

Out[11]:

	Unnamed: 0	key	fare_amount	pickup_datetime	pickup_longitude	pickup_latitude	dropoff_
1	27835199	2009-07-17 20:04:56	7.7	2009-07-17 20:04:56 UTC	-73.994355	40.728225	-7
2	44984355	2009-08-24 21:45:00	12.9	2009-08-24 21:45:00 UTC	-74.005043	40.740770	-7
3	25894730	2009-06-26 8:22:21	5.3	2009-06-26 08:22:21 UTC	-73.976124	40.790844	-7
4	17610152	2014-08-28 17:47:00	16.0	2014-08-28 17:47:00 UTC	-73.925023	40.744085	-7
5	44470845	2011-02-12 2:27:09	4.9	2011-02-12 02:27:09 UTC	-73.969019	40.755910	-7
6	48725865	2014-10-12 7:04:00	24.5	2014-10-12 07:04:00 UTC	-73.961447	40.693965	-7

In [12]:
```
x.isna()
```

Out[12]:

	Unnamed: 0	key	fare_amount	pickup_datetime	pickup_longitude	pickup_latitude	dropoff_lon
0	False	False	False	False	False	False	
1	False	False	False	False	False	False	
2	False	False	False	False	False	False	
3	False	False	False	False	False	False	
4	False	False	False	False	False	False	
5	False	False	False	False	False	False	
6	False	False	False	False	False	False	
7	False	False	False	False	False	False	
8	False	False	False	False	False	False	
9	False	False	False	False	False	False	

In [13]:
```
x.fillna(value=1000)
```

Out[13]:

	Unnamed: 0	key	fare_amount	pickup_datetime	pickup_longitude	pickup_latitude	dropoff_
0	24238194	2015-05-07 19:52:06	7.5	2015-05-07 19:52:06 UTC	-73.999817	40.738354	-7
1	27835199	2009-07-17 20:04:56	7.7	2009-07-17 20:04:56 UTC	-73.994355	40.728225	-7
2	44984355	2009-08-24 21:45:00	12.9	2009-08-24 21:45:00 UTC	-74.005043	40.740770	-7
3	25894730	2009-06-26 8:22:21	5.3	2009-06-26 08:22:21 UTC	-73.976124	40.790844	-7
4	17610152	2014-08-28 17:47:00	16.0	2014-08-28 17:47:00 UTC	-73.925023	40.744085	-7
5	44470845	2011-02-12 2:27:09	4.9	2011-02-12 02:27:09 UTC	-73.969019	40.755910	-7
6	48725865	2014-10-12 7:04:00	24.5	2014-10-12 07:04:00 UTC	-73.961447	40.693965	-7
7	44195482	2012-12-11 13:52:00	2.5	2012-12-11 13:52:00 UTC	0.000000	0.000000	
8	15822268	2012-02-17 9:32:00	9.7	2012-02-17 09:32:00 UTC	-73.975187	40.745767	-7
9	50611056	2012-03-29 19:06:00	12.5	2012-03-29 19:06:00 UTC	-74.001065	40.741787	-7

In [19]:
```
x.dropna()
```

Out[19]:

	Unnamed: 0	key	fare_amount	pickup_datetime	pickup_longitude	pickup_latitude	drc
0	24238194	2015-05-07 19:52:06	7.5	2015-05-07 19:52:06 UTC	-73.999817	40.738354	
1	27835199	2009-07-17 20:04:56	7.7	2009-07-17 20:04:56 UTC	-73.994355	40.728225	
2	44984355	2009-08-24 21:45:00	12.9	2009-08-24 21:45:00 UTC	-74.005043	40.740770	
3	25894730	2009-06-26 8:22:21	5.3	2009-06-26 08:22:21 UTC	-73.976124	40.790844	
4	17610152	2014-08-28 17:47:00	16.0	2014-08-28 17:47:00 UTC	-73.925023	40.744085	
...	...	...	...	...	...	...	...
199995	42598914	2012-10-28 10:49:00	3.0	2012-10-28 10:49:00 UTC	-73.987042	40.739367	
199996	16382965	2014-03-14 1:09:00	7.5	2014-03-14 01:09:00 UTC	-73.984722	40.736837	
199997	27804658	2009-06-29 0:42:00	30.9	2009-06-29 00:42:00 UTC	-73.986017	40.756487	
199998	20259894	2015-05-20 14:56:25	14.5	2015-05-20 14:56:25 UTC	-73.997124	40.725452	
199999	11951496	2010-05-15 4:08:00	14.1	2010-05-15 04:08:00 UTC	-73.984395	40.720077	

199999 rows × 9 columns

In [14]:
```
x.dropna(axis=1,how='any')
```

Out[14]:

	Unnamed: 0	key	fare_amount	pickup_datetime	pickup_longitude	pickup_latitude	dropoff_
0	24238194	2015-05-07 19:52:06	7.5	2015-05-07 19:52:06 UTC	-73.999817	40.738354	-7
1	27835199	2009-07-17 20:04:56	7.7	2009-07-17 20:04:56 UTC	-73.994355	40.728225	-7
2	44984355	2009-08-24 21:45:00	12.9	2009-08-24 21:45:00 UTC	-74.005043	40.740770	-7
3	25894730	2009-06-26 8:22:21	5.3	2009-06-26 08:22:21 UTC	-73.976124	40.790844	-7
4	17610152	2014-08-28 17:47:00	16.0	2014-08-28 17:47:00 UTC	-73.925023	40.744085	-7
5	44470845	2011-02-12 2:27:09	4.9	2011-02-12 02:27:09 UTC	-73.969019	40.755910	-7
6	48725865	2014-10-12 7:04:00	24.5	2014-10-12 07:04:00 UTC	-73.961447	40.693965	-7
7	44195482	2012-12-11 13:52:00	2.5	2012-12-11 13:52:00 UTC	0.000000	0.000000	
8	15822268	2012-02-17 9:32:00	9.7	2012-02-17 09:32:00 UTC	-73.975187	40.745767	-7
9	50611056	2012-03-29 19:06:00	12.5	2012-03-29 19:06:00 UTC	-74.001065	40.741787	-7

In [15]:
```
x.columns
```

Out[15]:
```
Index(['Unnamed: 0', 'key', 'fare_amount', 'pickup_datetime',
       'pickup_longitude', 'pickup_latitude', 'dropoff_longitude',
       'dropoff_latitude', 'passenger_count'],
      dtype='object')
```

In [16]:
```
x=x[['fare_amount','passenger_count']]
x
```

Out[16]:

	fare_amount	passenger_count
0	7.5	1
1	7.7	1
2	12.9	1
3	5.3	3
4	16.0	5
5	4.9	1
6	24.5	5
7	2.5	1
8	9.7	1
9	12.5	1

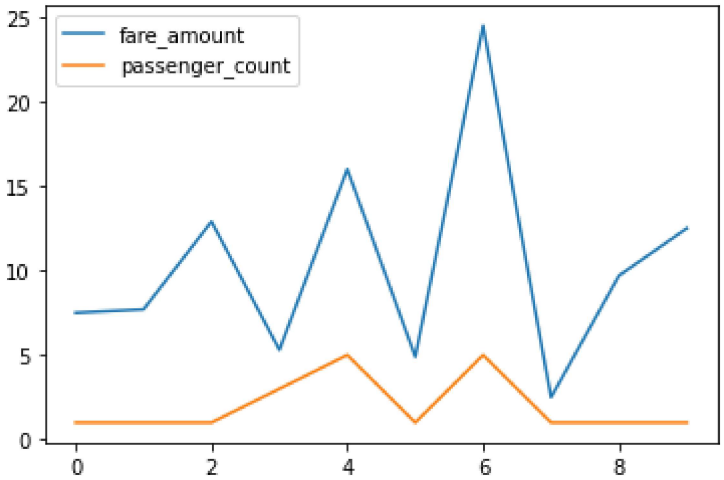
In [17]:
```
x.columns
```

Out[17]:
```
Index(['fare_amount', 'passenger_count'], dtype='object')
```

In [ ]:
```
import matplotlib.pyplot as pp
```

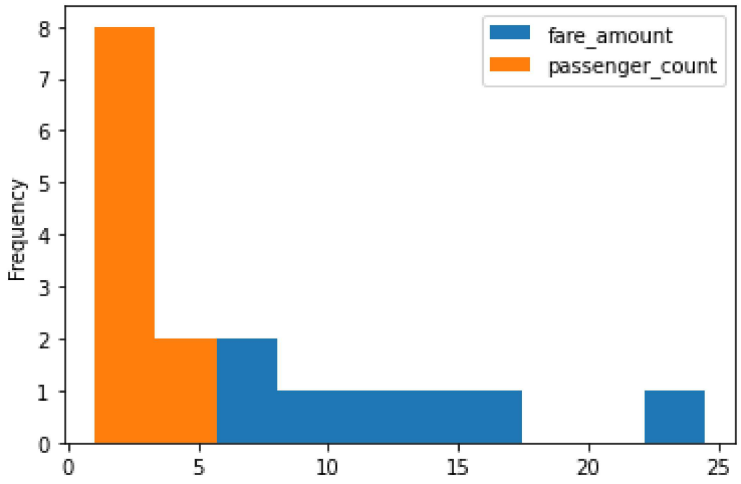
In [19]:
```
x.plot.line()
```

Out[19]:
```
<AxesSubplot:>
```



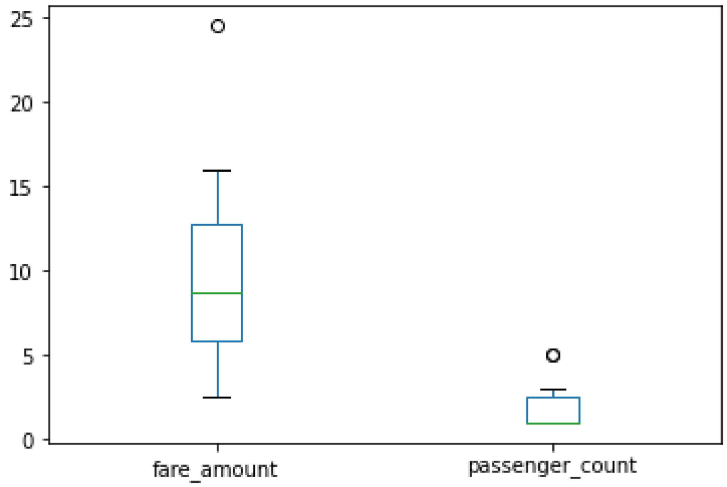
In [20]:
```
x.plot.hist()
```

Out[20]: <AxesSubplot:ylabel='Frequency'>



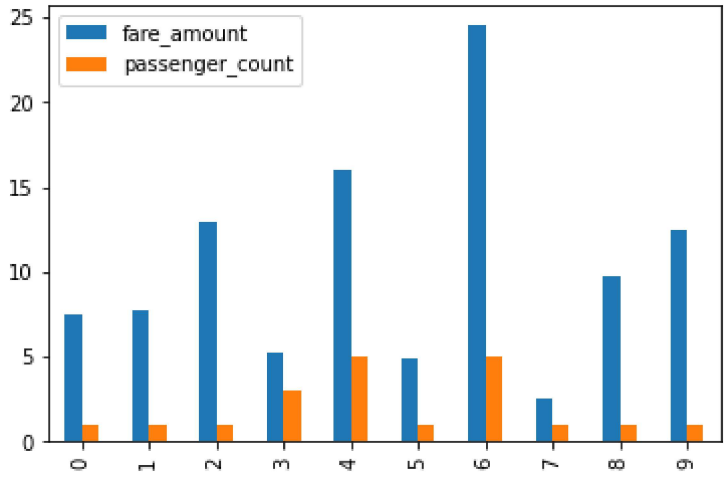
In [21]:
```
x.plot.box()
```

Out[21]: <AxesSubplot:>



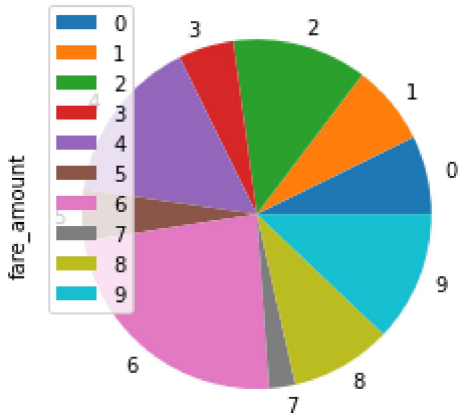
In [22]:
```
x.plot.bar()
```

Out[22]: <AxesSubplot:>



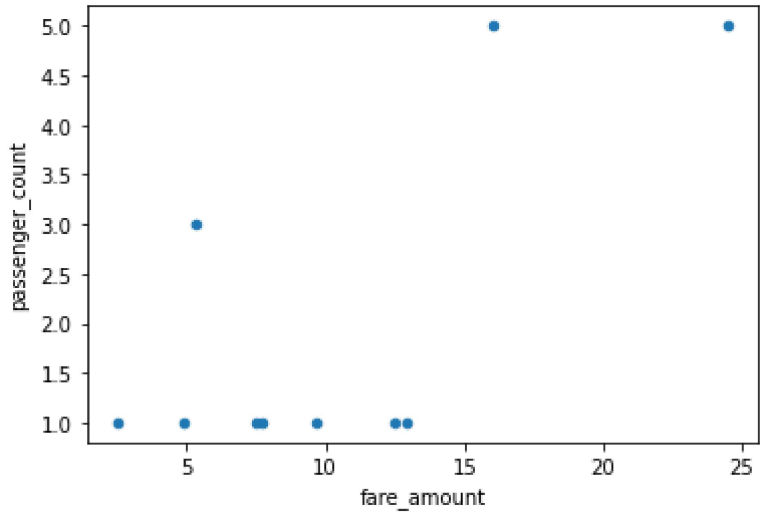
In [23]:
```
x.plot.pie(y='fare_amount')
```

Out[23]: <AxesSubplot:ylabel='fare_amount'>



In [24]:
```
x.plot.scatter(x='fare_amount',y='passenger_count')
```

Out[24]:
```
<AxesSubplot:xlabel='fare_amount', ylabel='passenger_count'>
```



In [ ]: