In [1]:

```
import numpy as np
import pandas as pd
```

In [2]:

```
a=pd.read_csv(r"C:\Users\user\Downloads\uber - uber.csv")
```

To print the first 10000 rows

In [48]:

```
b=a.head(1000)
b
```

Out[48]:

	Unnamed: 0	key	fare_amount	pickup_datetime	pickup_longitude	pickup_latitude	dr
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	
...	...	...	...	...	...	...	
995	13439193	2011-05-04 6:39:00	5.7	2011-05-04 06:39:00 UTC	-73.969720	40.757577	
996	32405310	2011-11-23 20:43:20	8.1	2011-11-23 20:43:20 UTC	-73.993784	40.757054	
997	51612001	2010-01-11 20:58:00	8.5	2010-01-11 20:58:00 UTC	-73.972338	40.765078	
998	937243	2013-06-12 17:01:24	5.5	2013-06-12 17:01:24 UTC	-73.979054	40.784730	
999	47946613	2011-12-11 21:48:22	9.7	2011-12-11 21:48:22 UTC	-73.983675	40.729944	

1000 rows × 9 columns

To print the last 5 rows

In [4]:

```
a.tail(5)
```

Out[4]:

	Unnamed: 0	key	fare_amount	pickup_datetime	pickup_longitude	pickup_latitude
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

To print the size function

In [5]:

```
print(np.size(a))
```

```
1800000
```

To print shape function

In [6]:

```
print(np.shape(a))
```

```
(200000, 9)
```

To print the na function

In [7]:

```
pd.isna(a)
```

Out[7]:

	Unnamed: 0	key	fare_amount	pickup_datetime	pickup_longitude	pickup_latitude	d
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	
...	...	...	...	...	...	...	
199995	False	False	False	False	False	False	
199996	False	False	False	False	False	False	
199997	False	False	False	False	False	False	
199998	False	False	False	False	False	False	
199999	False	False	False	False	False	False	

200000 rows × 9 columns

To print the na function

In [8]:

```
pd.isna(a)
```

Out[8]:

	Unnamed: 0	key	fare_amount	pickup_datetime	pickup_longitude	pickup_latitude	d
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	
...	...	...	...	...	...	...	
199995	False	False	False	False	False	False	
199996	False	False	False	False	False	False	
199997	False	False	False	False	False	False	
199998	False	False	False	False	False	False	
199999	False	False	False	False	False	False	

200000 rows × 9 columns

To print the na function

In [9]:

```
pd.isna(a)
```

Out[9]:

	Unnamed: 0	key	fare_amount	pickup_datetime	pickup_longitude	pickup_latitude	d
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	
...	...	...	...	...	...	...	
199995	False	False	False	False	False	False	
199996	False	False	False	False	False	False	
199997	False	False	False	False	False	False	
199998	False	False	False	False	False	False	
199999	False	False	False	False	False	False	

200000 rows × 9 columns

To print the drop function

In [10]:

```
a.dropna()
```

Out[10]:

	Unnamed: 0	key	fare_amount	pickup_datetime	pickup_longitude	pickup_latitude
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

To print the fill function

In [11]:

```
a.fillna(value=10)
```

Out[11]:

	Unnamed: 0	key	fare_amount	pickup_datetime	pickup_longitude	pickup_latitude
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

To describe the function

In [12]:

```
a.describe()
```

Out[12]:

	Unnamed: 0	fare_amount	pickup_longitude	pickup_latitude	dropoff_longitude	dro
count	2.000000e+05	200000.000000	200000.000000	200000.000000	199999.000000	19
mean	2.771250e+07	11.359955	-72.527638	39.935885	-72.525292	
std	1.601382e+07	9.901776	11.437787	7.720539	13.117408	
min	1.000000e+00	-52.000000	-1340.648410	-74.015515	-3356.666300	
25%	1.382535e+07	6.000000	-73.992065	40.734796	-73.991407	
50%	2.774550e+07	8.500000	-73.981823	40.752592	-73.980093	
75%	4.155530e+07	12.500000	-73.967153	40.767158	-73.963659	
max	5.542357e+07	499.000000	57.418457	1644.421482	1153.572603	

In [13]:

```
conda install matplotlib
```

```
Collecting package metadata (current_repodata.json): ...working... done
Solving environment: ...working... done

# All requested packages already installed.


Note: you may need to restart the kernel to use updated packages.


==> WARNING: A newer version of conda exists. <==
  current version: 4.10.1
  latest version: 23.5.2

Please update conda by running

    $ conda update -n base -c defaults conda
```

In [14]:

```
import matplotlib.pyplot as pp
```

In [49]:

```
c=b[["pickup_latitude","dropoff_latitude"]]
c
```

Out[49]:

	pickup_latitude	dropoff_latitude
0	40.738354	40.723217
1	40.728225	40.750325
2	40.740770	40.772647
3	40.790844	40.803349
4	40.744085	40.761247
...	...	...
995	40.757577	40.766960
996	40.757054	40.775632
997	40.765078	40.783833
998	40.784730	40.775048
999	40.729944	40.758208

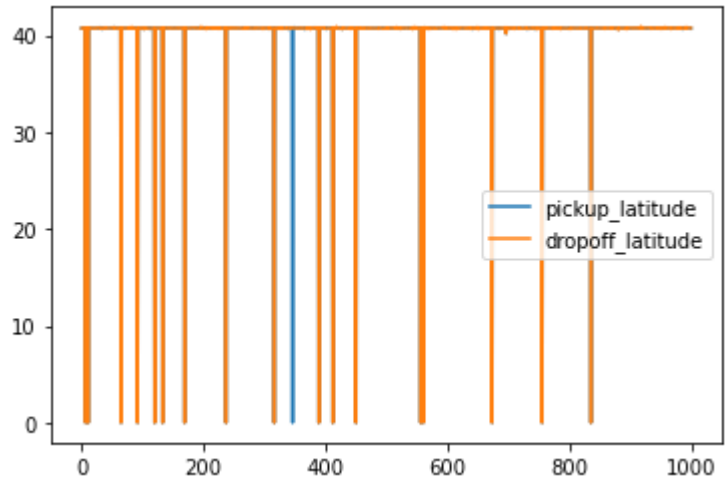
1000 rows × 2 columns

In [50]:

```
c.plot.line()
```

Out[50]:

```
<AxesSubplot:>
```

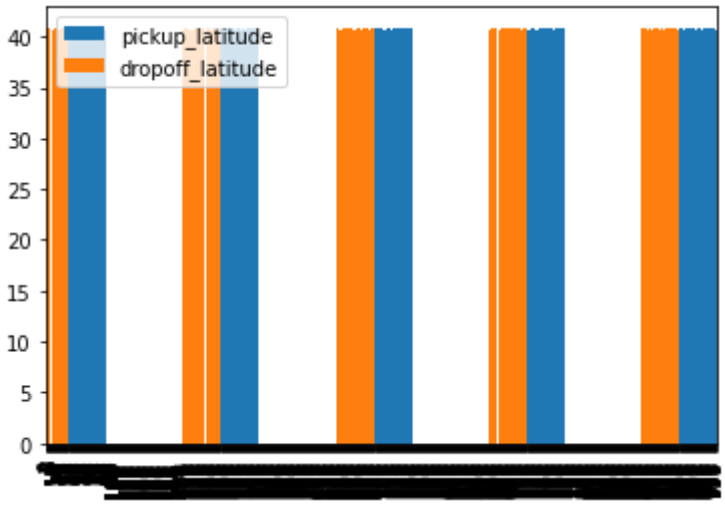


In [51]:

```
c.plot.bar()
```

Out[51]:

<AxesSubplot:>

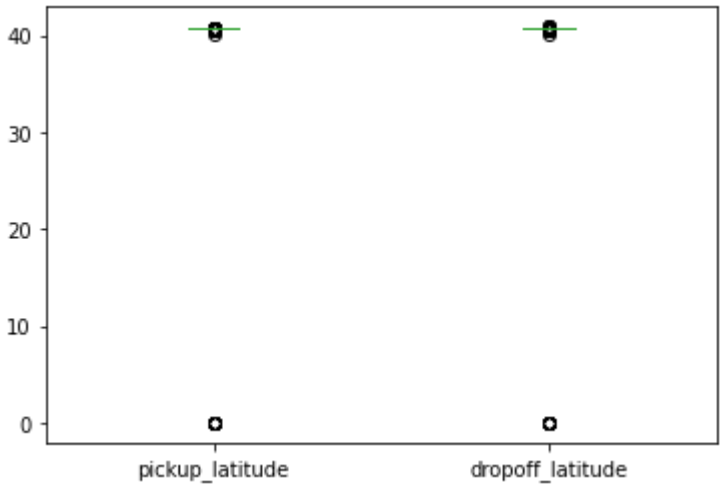


In [52]:

```
c.plot.box()
```

Out[52]:

<AxesSubplot:>

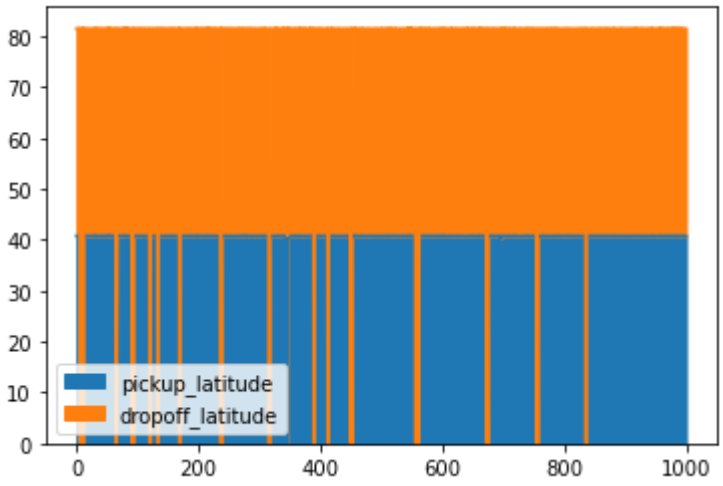


In [53]:

```
c.plot.area()
```

Out[53]:

```
<AxesSubplot:>
```

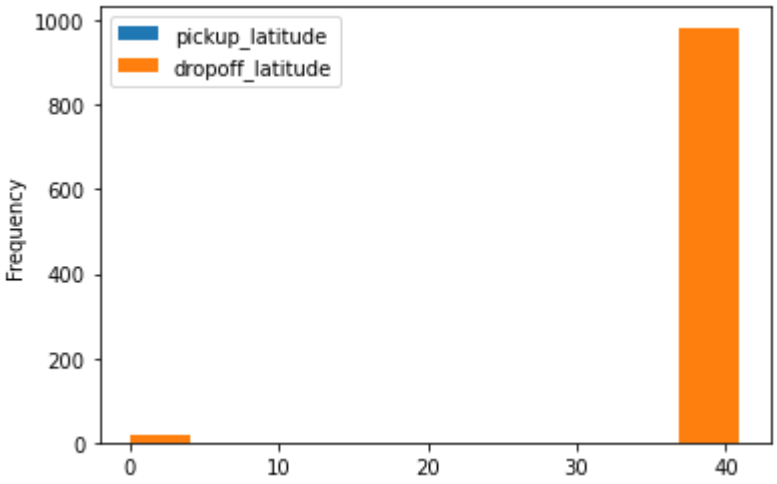


In [54]:

```
c.plot.hist()
```

Out[54]:

```
<AxesSubplot:ylabel='Frequency'>
```

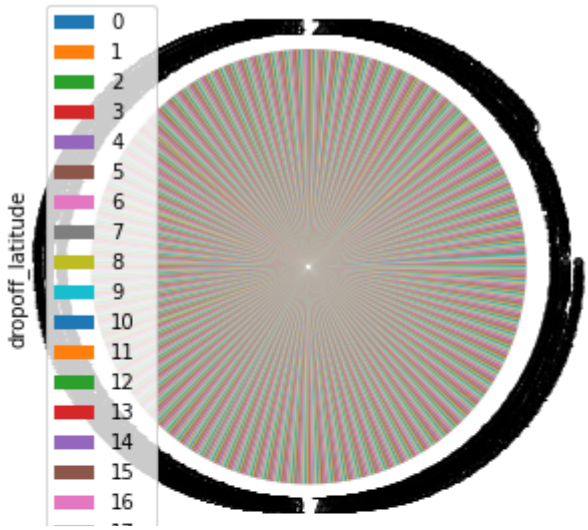


In [55]:

```
c.plot.pie(y='dropoff_latitude',figsize=(5,5))
```

Out[55]:

```
<AxesSubplot:ylabel='dropoff_latitude'>
```

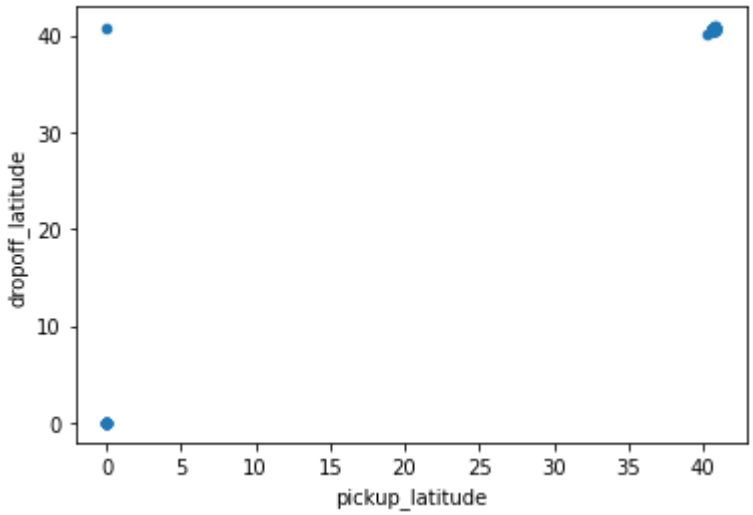


In [56]:

```
c.plot.scatter(x='pickup_latitude',y='dropoff_latitude')
```

Out[56]:

```
<AxesSubplot:xlabel='pickup_latitude', ylabel='dropoff_latitude'>
```



In [ ]: