22/07/23 tony6

In [132]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as pp
```

In [134]:
```
x=pd.read_csv(r"C:\Users\user\Downloads\6_Salesworkload1.csv")
x
```

Out[134]:

	MonthYear	Time index	Country	StoreID	City	Dept_ID	Dept. Name	HoursOwn	HoursLea
0	10.2016	1.0	United Kingdom	88253.0	London (I)	1.0	Dry	3184.764	
1	10.2016	1.0	United Kingdom	88253.0	London (I)	2.0	Frozen	1582.941	
2	10.2016	1.0	United Kingdom	88253.0	London (I)	3.0	other	47.205	
3	10.2016	1.0	United Kingdom	88253.0	London (I)	4.0	Fish	1623.852	
4	10.2016	1.0	United Kingdom	88253.0	London (I)	5.0	Fruits & Vegetables	1759.173	
...	...	...	...	...	...	...	...	...	
7653	06.2017	9.0	Sweden	29650.0	Gothenburg	12.0	Checkout	6322.323	
7654	06.2017	9.0	Sweden	29650.0	Gothenburg	16.0	Customer Services	4270.479	
7655	06.2017	9.0	Sweden	29650.0	Gothenburg	11.0	Delivery	0	
7656	06.2017	9.0	Sweden	29650.0	Gothenburg	17.0	others	2224.929	
7657	06.2017	9.0	Sweden	29650.0	Gothenburg	18.0	all	39652.2	

7658 rows × 14 columns

In [135]:
```
x=x.head(500)
x
```

Out[135]:

	MonthYear	Time index	Country	StoreID	City	Dept_ID	Dept. Name	HoursOwn	HoursLease
0	10.2016	1.0	United Kingdom	88253.0	London (I)	1.0	Dry	3184.764	0.0
1	10.2016	1.0	United Kingdom	88253.0	London (I)	2.0	Frozen	1582.941	0.0
2	10.2016	1.0	United Kingdom	88253.0	London (I)	3.0	other	47.205	0.0
3	10.2016	1.0	United Kingdom	88253.0	London (I)	4.0	Fish	1623.852	0.0
4	10.2016	1.0	United Kingdom	88253.0	London (I)	5.0	Fruits & Vegetables	1759.173	0.0
...	...	...	...	...	...	...	...	...	...
495	10.2016	1.0	Italy	64983.0	Milano	3.0	other	47.205	0.0
496	10.2016	1.0	Italy	64983.0	Milano	4.0	Fish	2451.513	0.0
497	10.2016	1.0	Italy	64983.0	Milano	5.0	Fruits & Vegetables	1944.846	0.0
498	10.2016	1.0	Italy	64983.0	Milano	6.0	Meat	11980.629	122.0
499	10.2016	1.0	Italy	64983.0	Milano	13.0	Food	23665.44	122.0

500 rows × 14 columns

In [136]:
```
x. dtypes
```

Out[136]:
```
MonthYear         object
Time index       float64
Country           object
StoreID          float64
City              object
Dept_ID          float64
Dept. Name        object
HoursOwn          object
HoursLease       float64
Sales units      float64
Turnover         float64
Customer         float64
Area (m2)         object
Opening hours     object
dtype: object
```

In [137]:
```
x. dtypes
```

Out[137]:
```
MonthYear         object
Time index       float64
Country           object
StoreID          float64
City              object
Dept_ID          float64
Dept. Name        object
HoursOwn          object
HoursLease       float64
Sales units      float64
Turnover         float64
Customer         float64
Area (m2)         object
Opening hours     object
dtype: object
```

In [138]:
```
x.tail()
```

Out[138]:

	MonthYear	Time index	Country	StoreID	City	Dept_ID	Dept. Name	HoursOwn	HoursLease	
495	10.2016	1.0	Italy	64983.0	Milano	3.0	other	47.205	0.0	
496	10.2016	1.0	Italy	64983.0	Milano	4.0	Fish	2451.513	0.0	
497	10.2016	1.0	Italy	64983.0	Milano	5.0	Fruits & Vegetables	1944.846	0.0	
498	10.2016	1.0	Italy	64983.0	Milano	6.0	Meat	11980.629	122.0	2
499	10.2016	1.0	Italy	64983.0	Milano	13.0	Food	23665.44	122.0	4

In [139]:
```
x.columns
```

Out[139]:
```
Index(['MonthYear', 'Time index', 'Country', 'StoreID', 'City', 'Dept_ID',
       'Dept. Name', 'HoursOwn', 'HoursLease', 'Sales units', 'Turnover',
       'Customer', 'Area (m2)', 'Opening hours'],
      dtype='object')
```

In [140]:
```
x. index
```

Out[140]:
```
RangeIndex(start=0, stop=500, step=1)
```

In [143]:
```
x.describe()
```

Out[143]:

	Time index	StoreID	Dept_ID	HoursLease	Sales units	Turnover	Customer
count	500.0	500.000000	500.000000	500.000000	5.000000e+02	5.000000e+02	0.0
mean	1.0	57412.764000	9.406000	31.520000	9.397837e+05	3.153113e+06	NaN
std	0.0	32104.273482	5.350366	142.134408	1.486945e+06	5.165524e+06	NaN
min	1.0	15552.000000	1.000000	0.000000	0.000000e+00	0.000000e+00	NaN
25%	1.0	20891.000000	5.000000	0.000000	5.200250e+04	2.345122e+05	NaN
50%	1.0	71991.000000	9.000000	0.000000	2.555375e+05	7.053345e+05	NaN
75%	1.0	88253.000000	14.000000	0.000000	8.903900e+05	2.542147e+06	NaN
max	1.0	96857.000000	18.000000	1896.000000	7.476680e+06	2.571973e+07	NaN

In [144]:
```
x["Time index"]
```

Out[144]:
```
0      1.0
1      1.0
2      1.0
3      1.0
4      1.0
      ...
495    1.0
496    1.0
497    1.0
498    1.0
499    1.0
Name: Time index, Length: 500, dtype: float64
```

In [145]:
```
x.iloc[0:2]
```

Out[145]:

	MonthYear	Time index	Country	StoreID	City	Dept_ID	Dept. Name	HoursOwn	HoursLease	Sal un
0	10.2016	1.0	United Kingdom	88253.0	London (I)	1.0	Dry	3184.764	0.0	39856(
1	10.2016	1.0	United Kingdom	88253.0	London (I)	2.0	Frozen	1582.941	0.0	82725

In [146]:
```
x.loc[0:3]
```

Out[146]:

	MonthYear	Time index	Country	StoreID	City	Dept_ID	Dept. Name	HoursOwn	HoursLease	Sal un
0	10.2016	1.0	United Kingdom	88253.0	London (I)	1.0	Dry	3184.764	0.0	39856(
1	10.2016	1.0	United Kingdom	88253.0	London (I)	2.0	Frozen	1582.941	0.0	82725
2	10.2016	1.0	United Kingdom	88253.0	London (I)	3.0	other	47.205	0.0	43840(
3	10.2016	1.0	United Kingdom	88253.0	London (I)	4.0	Fish	1623.852	0.0	30942

In [148]:
```
x.loc["StoreID":"Dept_ID"]
```

Out[148]:

	MonthYear	Time index	Country	StoreID	City	Dept_ID	Dept. Name	HoursOwn	HoursLease	Sales units	Turn

In [149]:
```
x[x["Dept_ID"]<=2]
```

391	10.2016	1.0	Spain	20166.0	Madrid (II)	1.0	Dry	3134.412		
392	10.2016	1.0	Spain	20166.0	Madrid (II)	2.0	Frozen	2268.987		
408	10.2016	1.0	Spain	16927.0	Barcelona (I)	1.0	Dry	3656.814		
409	10.2016	1.0	Spain	16927.0	Barcelona (I)	2.0	Frozen	1913.376		
425	10.2016	1.0	Spain	96493.0	Barcelona (II)	1.0	Dry	4676.442		
426	10.2016	1.0	Spain	96493.0	Barcelona (II)	2.0	Frozen	2665.509		
442	10.2016	1.0	Spain	88750.0	Bilbao	1.0	Dry	4641.825		
443	10.2016	1.0	Spain	88750.0	Bilbao	2.0	Frozen	3191.058		
459	10.2016	1.0	Italy	78450.0	Rome (I)	1.0	Dry	4160.334		
460	10.2016	1.0	Italy	78450.0	Rome (I)	2.0	Frozen	2574.246		
476	10.2016	1.0	Italy	94153.0	Rome (II)	1.0	Dry	4336.566		1

In [150]:
```
x.fillna(value=5)
```

Out[150]:

	MonthYear	Time index	Country	StoreID	City	Dept_ID	Dept. Name	HoursOwn	HoursLease
0	10.2016	1.0	United Kingdom	88253.0	London (I)	1.0	Dry	3184.764	0.0
1	10.2016	1.0	United Kingdom	88253.0	London (I)	2.0	Frozen	1582.941	0.0
2	10.2016	1.0	United Kingdom	88253.0	London (I)	3.0	other	47.205	0.0
3	10.2016	1.0	United Kingdom	88253.0	London (I)	4.0	Fish	1623.852	0.0
4	10.2016	1.0	United Kingdom	88253.0	London (I)	5.0	Fruits & Vegetables	1759.173	0.0
...	...	...	...	...	...	...	...	...	...
495	10.2016	1.0	Italy	64983.0	Milano	3.0	other	47.205	0.0
496	10.2016	1.0	Italy	64983.0	Milano	4.0	Fish	2451.513	0.0
497	10.2016	1.0	Italy	64983.0	Milano	5.0	Fruits & Vegetables	1944.846	0.0
498	10.2016	1.0	Italy	64983.0	Milano	6.0	Meat	11980.629	122.0
499	10.2016	1.0	Italy	64983.0	Milano	13.0	Food	23665.44	122.0

500 rows × 14 columns

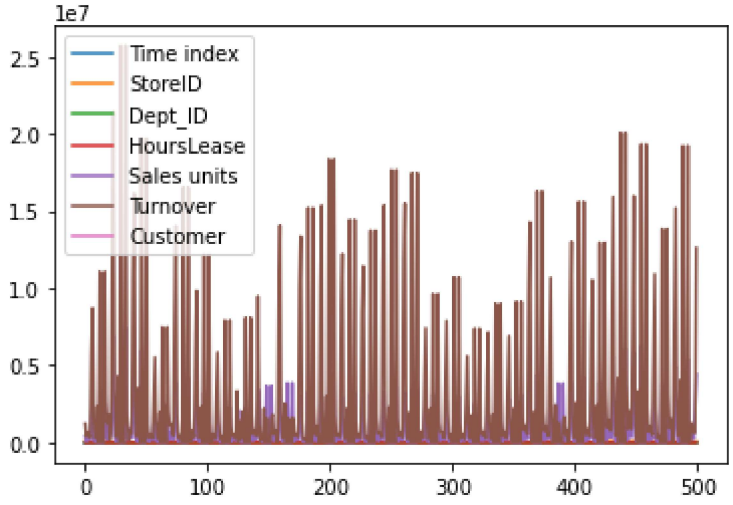
In [151]:
```
x.dropna()
```

Out[151]:

	MonthYear	Time index	Country	StoreID	City	Dept_ID	Dept. Name	HoursOwn	HoursLease	Sales units	Turn

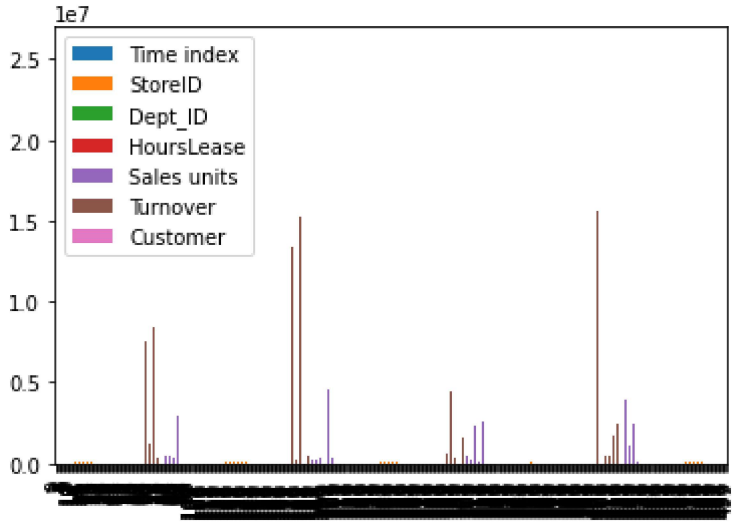
In [152]:
```
x.plot.line()
```

Out[152]: <AxesSubplot:>



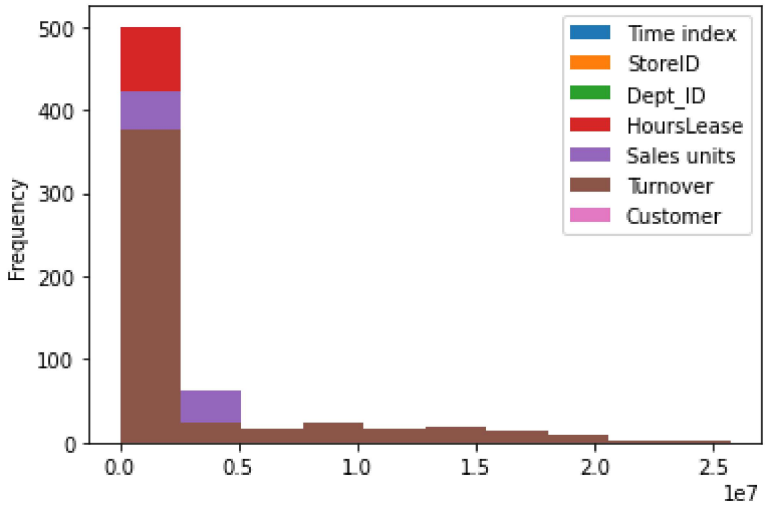
In [153]:
```
x.plot.bar()
```

Out[153]: <AxesSubplot:>



In [154]:
```
x.plot.hist()
```

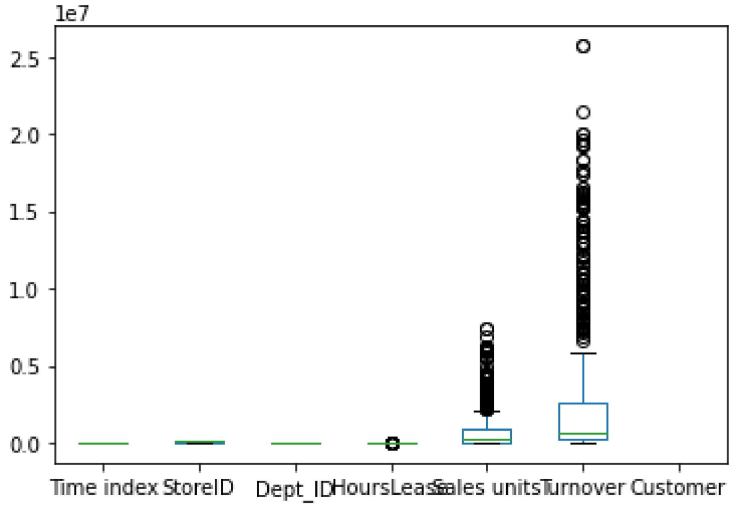
Out[154]: <AxesSubplot:ylabel='Frequency'>



In [156]:
```
x.plot.pie(y='HoursLease')
```

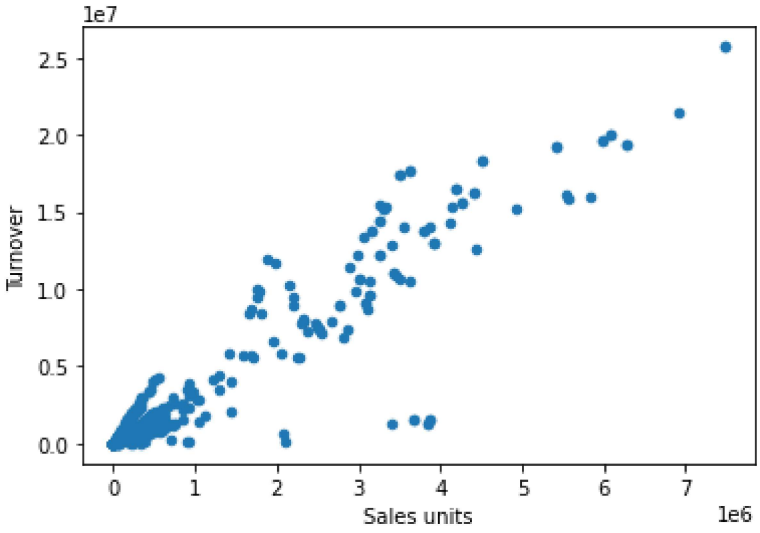
In [157]:
```
x.plot.box()
```

Out[157]: <AxesSubplot:>



In [158]:
```
x.plot.scatter(x="Sales units",y='Turnover')
```

Out[158]: <AxesSubplot:xlabel='Sales units', ylabel='Turnover'>



In [ ]:

In [ ]: