22/07/2023

In [196]:
```
import numpy as np
import pandas as pd
import matplotlib.pyplot as pp
```

In [197]:
```
x=pd.read_csv(r"C:\Users\user\Downloads\6_Salesworkload1 - 6_Salesworkload1.csv
x
```

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	6.2017	9.0	Sweden	29650.0	Gothenburg	12.0	Checkout	6322.323
7654	6.2017	9.0	Sweden	29650.0	Gothenburg	16.0	Customer Services	4270.479
7655	6.2017	9.0	Sweden	29650.0	Gothenburg	11.0	Delivery	0
7656	6.2017	9.0	Sweden	29650.0	Gothenburg	17.0	others	2224.929
7657	6.2017	9.0	Sweden	29650.0	Gothenburg	18.0	all	39652.2

In [198]:
```
x=x.head(400)
```

In [199]:
```
x.dtypes
```

Out[199]:
```
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 [200]:
```
x.head()
```

Out[200]:

	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	39
1	10.2016	1.0	United Kingdom	88253.0	London (I)	2.0	Frozen	1582.941	0.0	8
2	10.2016	1.0	United Kingdom	88253.0	London (I)	3.0	other	47.205	0.0	43
3	10.2016	1.0	United Kingdom	88253.0	London (I)	4.0	Fish	1623.852	0.0	30
4	10.2016	1.0	United Kingdom	88253.0	London (I)	5.0	Fruits & Vegetables	1759.173	0.0	16

In [201]:
```
x.tail()
```

Out[201]:

	MonthYear	Time index	Country	StoreID	City	Dept_ID	Dept. Name	HoursOwn	HoursLease
395	10.2016	1.0	Spain	20166.0	Madrid (II)	5.0	Fruits & Vegetables	1985.757	0.0
396	10.2016	1.0	Spain	20166.0	Madrid (II)	6.0	Meat	7577.976	0.0
397	10.2016	1.0	Spain	20166.0	Madrid (II)	13.0	Food	17273.883	0.0
398	10.2016	1.0	Spain	20166.0	Madrid (II)	7.0	Clothing	6989.487	0.0
399	10.2016	1.0	Spain	20166.0	Madrid (II)	8.0	Household	62.94	0.0

In [202]:
```
x.columns
```

Out[202]:
```
Index(['MonthYear', 'Time index', 'Country', 'StoreID', 'City', 'Dept_ID',
       'Dept. Name', 'HoursOwn', 'HoursLease', 'Sales units', 'Turnover',
       'Customer', 'Area (m2)', 'Opening hours'],
      dtype='object')
```

In [203]:
```
x.index
```

Out[203]:
```
RangeIndex(start=0, stop=400, step=1)
```

In [204]:
```
x.describe()
```

Out[204]:

	Time index	StoreID	Dept_ID	HoursLease	Sales units	Turnover	Customer
count	400.0	400.000000	400.000000	400.00000	4.000000e+02	4.000000e+02	0.0
mean	1.0	54297.580000	9.380000	32.44000	8.702047e+05	2.954216e+06	NaN
std	0.0	31580.605275	5.341998	151.65006	1.379226e+06	4.942532e+06	NaN
min	1.0	15552.000000	1.000000	0.00000	0.000000e+00	0.000000e+00	NaN
25%	1.0	20891.000000	5.000000	0.00000	4.799375e+04	2.326808e+05	NaN
50%	1.0	45583.000000	9.000000	0.00000	2.356100e+05	5.990580e+05	NaN
75%	1.0	87703.000000	14.000000	0.00000	7.768538e+05	2.237584e+06	NaN
max	1.0	96857.000000	18.000000	1896.00000	7.476680e+06	2.571973e+07	NaN

In [205]:
```
x["StoreID"]
```

Out[205]:
```
0      88253.0
1      88253.0
2      88253.0
3      88253.0
4      88253.0
        ...
395    20166.0
396    20166.0
397    20166.0
398    20166.0
399    20166.0
Name: StoreID, Length: 400, dtype: float64
```

In [206]:
```
x[0:2]
```

Out[206]:

	MonthYear	Time index	Country	StoreID	City	Dept_ID	Dept. Name	HoursOwn	HoursLease	Sa 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	8272

In [207]:
```
x.iloc[0:2]
```

Out[207]:

	MonthYear	Time index	Country	StoreID	City	Dept_ID	Dept. Name	HoursOwn	HoursLease	Sa 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	8272

In [208]:
```
x.loc[0:3]
```

Out[208]:

	MonthYear	Time index	Country	StoreID	City	Dept_ID	Dept. Name	HoursOwn	HoursLease	Sa 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	8272
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 [209]:
```
x.loc[" Dept_ID":"Sales units"]
```

Out[209]:

	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
...	...	...	...	...	...	...	...	...	...
395	10.2016	1.0	Spain	20166.0	Madrid (II)	5.0	Fruits & Vegetables	1985.757	0.0
396	10.2016	1.0	Spain	20166.0	Madrid (II)	6.0	Meat	7577.976	0.0
397	10.2016	1.0	Spain	20166.0	Madrid (II)	13.0	Food	17273.883	0.0
398	10.2016	1.0	Spain	20166.0	Madrid (II)	7.0	Clothing	6989.487	0.0
399	10.2016	1.0	Spain	20166.0	Madrid (II)	8.0	Household	62.94	0.0

400 rows × 14 columns

In [210]:
```
x[x["Turnover"]<=2]
```

Out[210]:

	MonthYear	Time index	Country	StoreID	City	Dept_ID	Dept. Name	HoursOwn	HoursLeas
11	10.2016	1.0	United Kingdom	88253.0	London (I)	15.0	Admin	4308.243	0.
13	10.2016	1.0	United Kingdom	88253.0	London (I)	16.0	Customer Services	3320.085	0.
14	10.2016	1.0	United Kingdom	88253.0	London (I)	11.0	Delivery	0	0.
15	10.2016	1.0	United Kingdom	88253.0	London (I)	17.0	others	2253.252	0.
28	10.2016	1.0	United Kingdom	38976.0	Manchester	15.0	Admin	6967.458	0.
...	...	...	...	...	...	...	...	...	.
372	10.2016	1.0	Denmark	19000.0	Aalborg (II)	17.0	others	2070.726	368.
385	10.2016	1.0	Spain	88994.0	Madrid (I)	15.0	Admin	3194.205	0.
387	10.2016	1.0	Spain	88994.0	Madrid (I)	16.0	Customer Services	2187.165	0.
388	10.2016	1.0	Spain	88994.0	Madrid (I)	11.0	Delivery	0	0.
389	10.2016	1.0	Spain	88994.0	Madrid (I)	17.0	others	1796.937	0.

73 rows × 14 columns

In [211]:
```
x.fillna(value=5)
```

Out[211]:

	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
...	...	...	...	...	...	...	...	...	...
395	10.2016	1.0	Spain	20166.0	Madrid (II)	5.0	Fruits & Vegetables	1985.757	0.0
396	10.2016	1.0	Spain	20166.0	Madrid (II)	6.0	Meat	7577.976	0.0
397	10.2016	1.0	Spain	20166.0	Madrid (II)	13.0	Food	17273.883	0.0
398	10.2016	1.0	Spain	20166.0	Madrid (II)	7.0	Clothing	6989.487	0.0
399	10.2016	1.0	Spain	20166.0	Madrid (II)	8.0	Household	62.94	0.0

400 rows × 14 columns

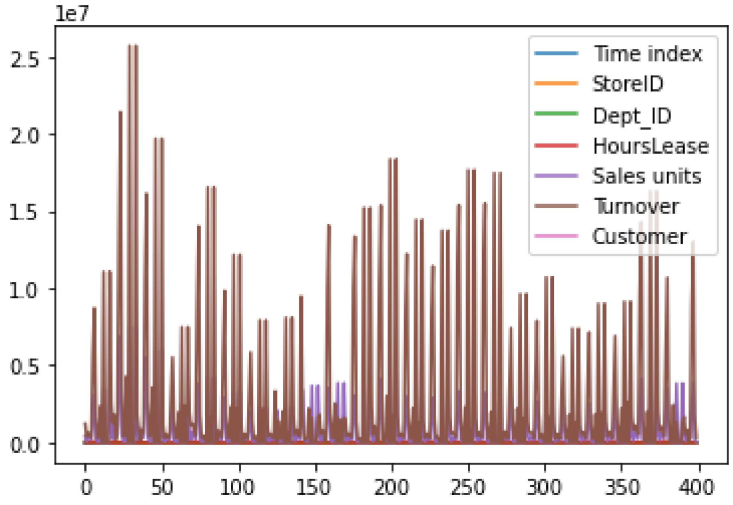
In [212]:
```
x.dropna()
```

Out[212]:

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

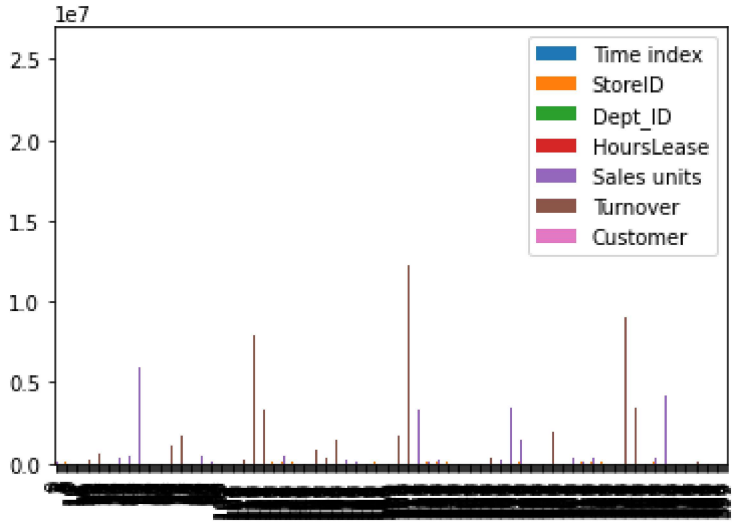
In [213]:
```
x.plot.line()
```

Out[213]: <AxesSubplot:>

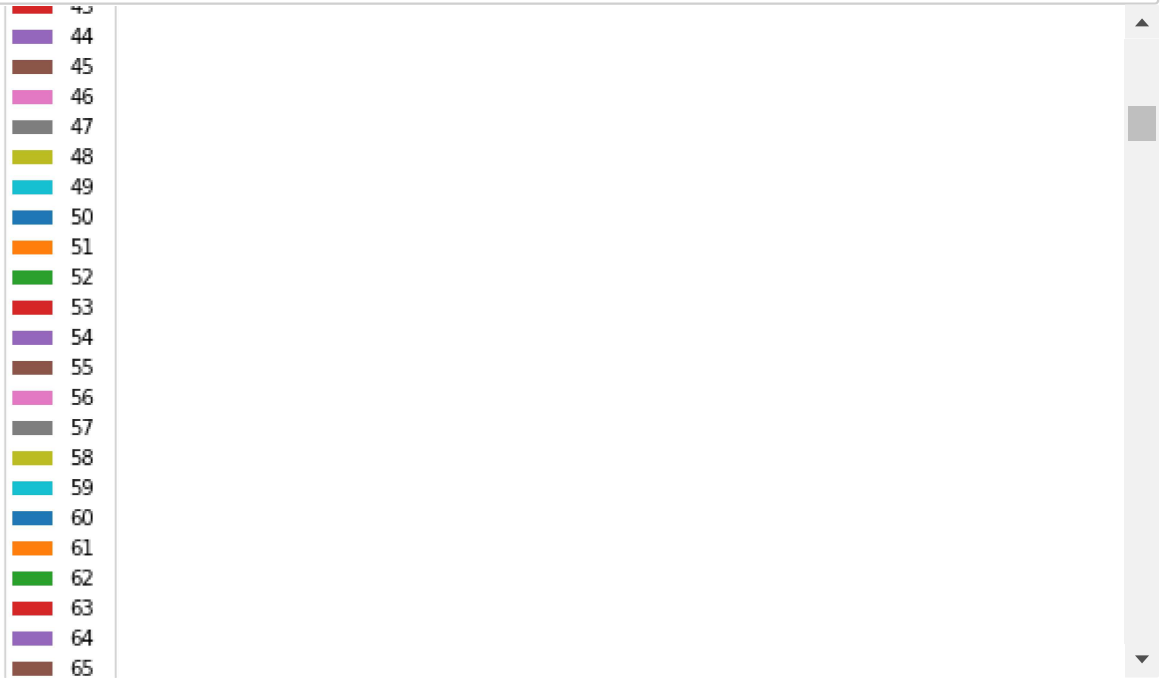


In [218]:
```
x.plot.bar()
```

Out[218]: <AxesSubplot:>

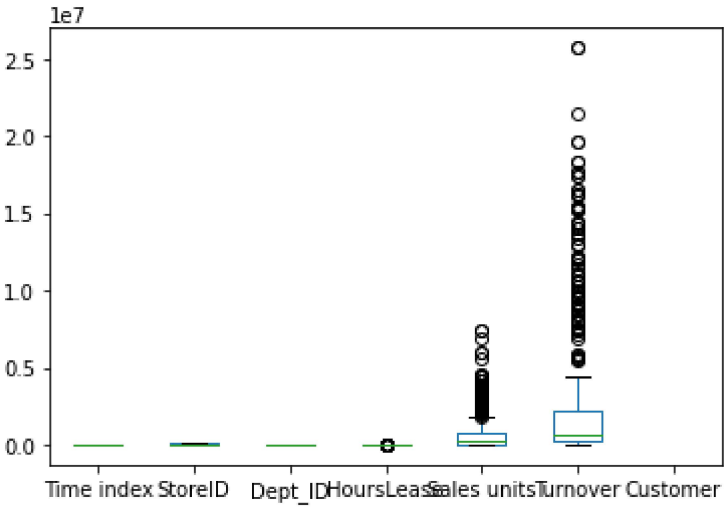


In [215]:
```
x.plot.pie(y='StoreID')
```



In [216]:
```
x.plot.box()
```

Out[216]: <AxesSubplot:>



In [217]:
```
x.plot.scatter(x='Time index',y='Turnover')
```

Out[217]: <AxesSubplot:xlabel='Time index', ylabel='Turnover'>