

1.IMPORTING LIBRARIES

```
In [1]: import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
```

2.Importing dataset

```
In [2]: data=pd.read_csv(r"C:\Users\user\Downloads\6_Salesworkload1 - 6_Salesworkload1.csv")
data
```

```
Out[2]:
```

	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
...
7653	6.2017	9.0	Sweden	29650.0	Gothenburg	12.0	Checkout	6322.323	0.0 3
7654	6.2017	9.0	Sweden	29650.0	Gothenburg	16.0	Customer Services	4270.479	0.0
7655	6.2017	9.0	Sweden	29650.0	Gothenburg	11.0	Delivery	0	0.0
7656	6.2017	9.0	Sweden	29650.0	Gothenburg	17.0	others	2224.929	0.0
7657	6.2017	9.0	Sweden	29650.0	Gothenburg	18.0	all	39652.2	0.0 3

7658 rows × 14 columns



3.head

```
In [3]: data.head(8)
```

Out[3]:

	MonthYear	Time index	Country	StoreID	City	Dept_ID	Dept. Name	HoursOwn	HoursLease	Sale unit
0	10.2016	1.0	United Kingdom	88253.0	London (I)	1.0	Dry	3184.764	0.0	398560.0
1	10.2016	1.0	United Kingdom	88253.0	London (I)	2.0	Frozen	1582.941	0.0	82725.0
2	10.2016	1.0	United Kingdom	88253.0	London (I)	3.0	other	47.205	0.0	438400.0
3	10.2016	1.0	United Kingdom	88253.0	London (I)	4.0	Fish	1623.852	0.0	309425.0
4	10.2016	1.0	United Kingdom	88253.0	London (I)	5.0	Fruits & Vegetables	1759.173	0.0	165515.0
5	10.2016	1.0	United Kingdom	88253.0	London (I)	6.0	Meat	8270.316	0.0	1713310.0
6	10.2016	1.0	United Kingdom	88253.0	London (I)	13.0	Food	16468.251	0.0	3107935.0
7	10.2016	1.0	United Kingdom	88253.0	London (I)	7.0	Clothing	4698.471	0.0	213680.0

4.tail

In [4]:

```
data.tail(7)
```

Out[4]:

	MonthYear	Time index	Country	StoreID	City	Dept_ID	Dept. Name	HoursOwn	HoursLease	
7651	6.2017	9.0	Sweden	29650.0	Gothenburg	14.0	Non Food	6498.555	0.0	3
7652	6.2017	9.0	Sweden	29650.0	Gothenburg	15.0	Admin	3433.377	0.0	
7653	6.2017	9.0	Sweden	29650.0	Gothenburg	12.0	Checkout	6322.323	0.0	38
7654	6.2017	9.0	Sweden	29650.0	Gothenburg	16.0	Customer Services	4270.479	0.0	
7655	6.2017	9.0	Sweden	29650.0	Gothenburg	11.0	Delivery	0	0.0	
7656	6.2017	9.0	Sweden	29650.0	Gothenburg	17.0	others	2224.929	0.0	
7657	6.2017	9.0	Sweden	29650.0	Gothenburg	18.0	all	39652.2	0.0	38

5.describe()

In [5]:

```
data.describe()
```

Out[5]:

	Time index	StoreID	Dept_ID	HoursLease	Sales units	Turnover	Customer
count	7650.000000	7650.000000	7650.000000	7650.000000	7.650000e+03	7.650000e+03	0.0
mean	5.000000	61995.220000	9.470588	22.036078	1.076471e+06	3.721393e+06	NaN
std	2.582158	29924.581631	5.337429	133.299513	1.728113e+06	6.003380e+06	NaN
min	1.000000	12227.000000	1.000000	0.000000	0.000000e+00	0.000000e+00	NaN
25%	3.000000	29650.000000	5.000000	0.000000	5.457125e+04	2.726798e+05	NaN
50%	5.000000	75400.500000	9.000000	0.000000	2.932300e+05	9.319575e+05	NaN
75%	7.000000	87703.000000	14.000000	0.000000	9.175075e+05	3.264432e+06	NaN
max	9.000000	98422.000000	18.000000	3984.000000	1.124296e+07	4.271739e+07	NaN

6.shape()

In [6]:

np.shape(data)

Out[6]: (7658, 14)

7.size()

In [7]:

np.size(data)

Out[7]: 107212

8.isna()

In [8]:

data.isna()

Out[8]:

	MonthYear	Time index	Country	StoreID	City	Dept_ID	Dept. Name	HoursOwn	HoursLease	Sales units	Turno
0	False	False	False	False	False	False	False	False	False	False	F
1	False	False	False	False	False	False	False	False	False	False	F
2	False	False	False	False	False	False	False	False	False	False	F
3	False	False	False	False	False	False	False	False	False	False	F
4	False	False	False	False	False	False	False	False	False	False	F
...	
7653	False	False	False	False	False	False	False	False	False	False	F
7654	False	False	False	False	False	False	False	False	False	False	F
7655	False	False	False	False	False	False	False	False	False	False	F

	MonthYear	Time index	Country	StoreID	City	Dept_ID	Dept. Name	HoursOwn	HoursLease	Sales units	Turnover
	7656	False	False	False	False	False	False	False	False	False	F
	7657	False	False	False	False	False	False	False	False	False	F

7658 rows × 14 columns

9.dropna

In [9]:

data.dropna()

Out[9]:

	MonthYear	Time index	Country	StoreID	City	Dept_ID	Dept. Name	HoursOwn	HoursLease	Sales units	Turnover
<div><div></div></div>											

10.selecting specific column

In [10]:

da=data[["Country","Dept_ID"]]
da

Out[10]:

	Country	Dept_ID
0	United Kingdom	1.0
1	United Kingdom	2.0
2	United Kingdom	3.0
3	United Kingdom	4.0
4	United Kingdom	5.0
...
7653	Sweden	12.0
7654	Sweden	16.0
7655	Sweden	11.0
7656	Sweden	17.0
7657	Sweden	18.0

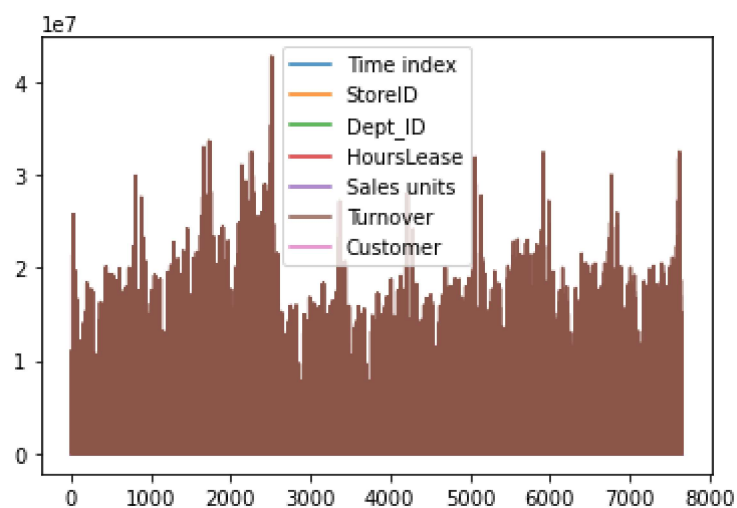
7658 rows × 2 columns

11.line plot

In [11]:

data.plot.line()

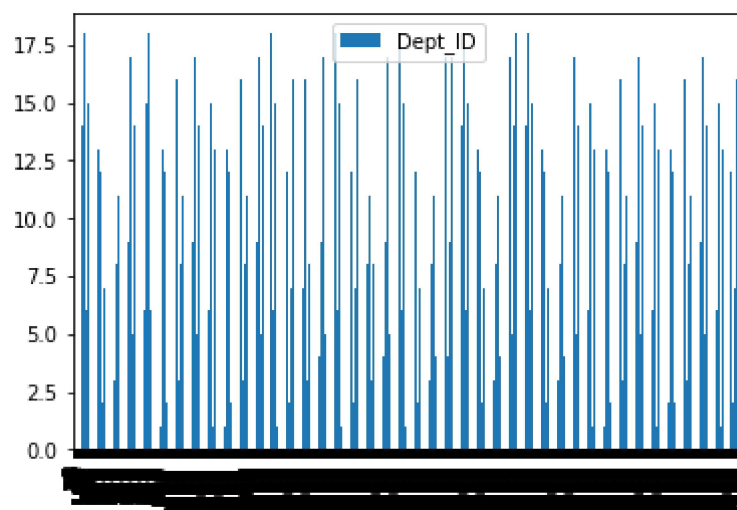
Out[11]: <AxesSubplot:>



12.bar plot

In [12]: `da.plot.bar()`

Out[12]: <AxesSubplot:>



13.area plot

In [13]: `data.plot.area()`

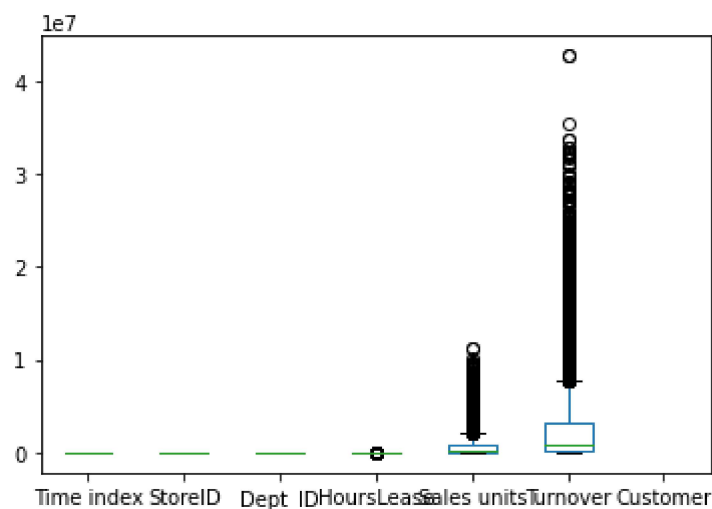
Out[13]: <AxesSubplot:>



14. box plot

```
In [14]: data.plot.box()
```

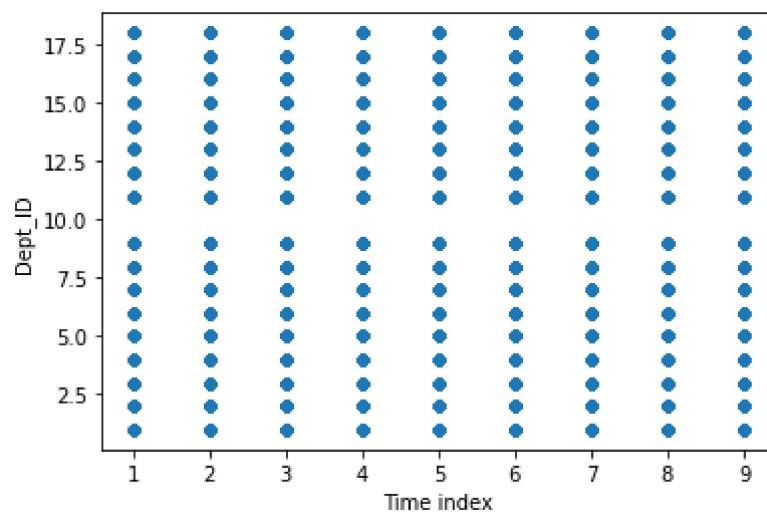
```
Out[14]: <AxesSubplot:>
```



15. scatter plot

```
In [15]: data.plot.scatter("Time index", "Dept_ID")
```

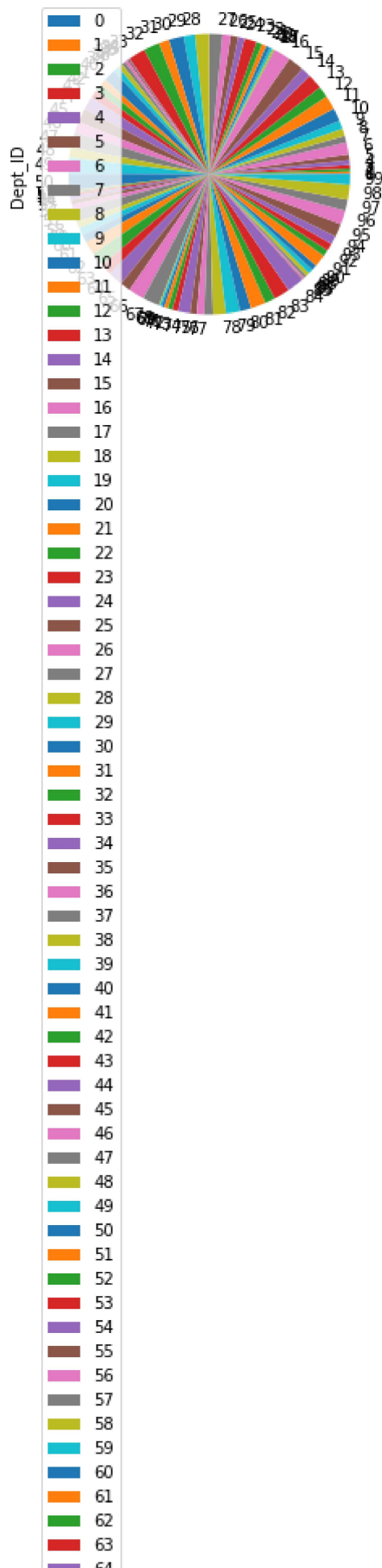
```
Out[15]: <AxesSubplot:xlabel='Time index', ylabel='Dept_ID'>
```

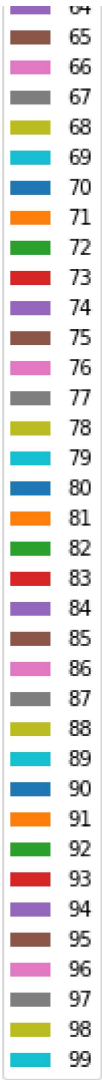


16.pie plot

```
In [16]: da=data.head(100)
         da.plot.pie(y="Dept_ID")
```

```
Out[16]: <AxesSubplot:ylabel='Dept_ID'>
```

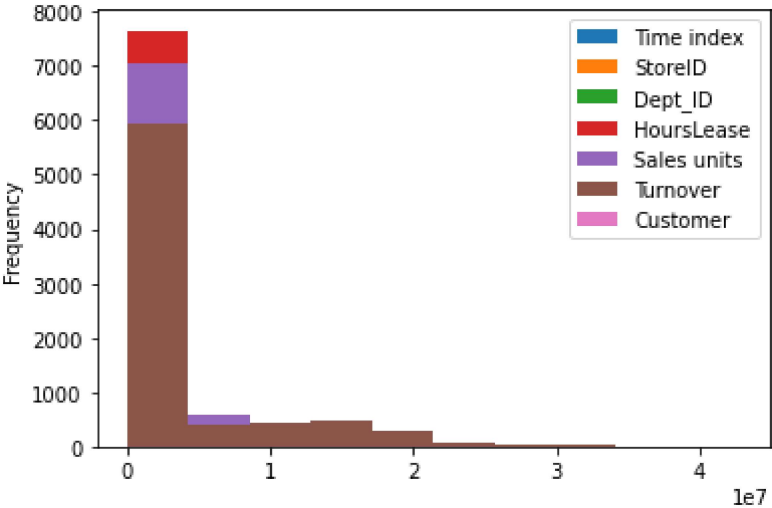




17.histogram

```
In [17]: data.plot.hist()
```

Out[17]: <AxesSubplot:ylabel='Frequency'>



In []: