Salesworkload Import labary

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

Import dataset

In [2]:
```
data=pd.read_csv(r"c:\Users\user\Downloads\Salesworkload.csv")
data
```

Out[2]:

| | 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** | 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 | |

7658 rows × 14 columns

Print head first 20 rows

In [12]:
```
data.head(20)
```

Out[12]:

| | 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.( |
| **1** | 10.2016 | 1.0 | United Kingdom | 88253.0 | London (I) | 2.0 | Frozen | 1582.941 | 0.( |
| **2** | 10.2016 | 1.0 | United Kingdom | 88253.0 | London (I) | 3.0 | other | 47.205 | 0.( |
| **3** | 10.2016 | 1.0 | United Kingdom | 88253.0 | London (I) | 4.0 | Fish | 1623.852 | 0.( |
| **4** | 10.2016 | 1.0 | United Kingdom | 88253.0 | London (I) | 5.0 | Fruits & Vegetables | 1759.173 | 0.( |
| **5** | 10.2016 | 1.0 | United Kingdom | 88253.0 | London (I) | 6.0 | Meat | 8270.316 | 0.( |
| **6** | 10.2016 | 1.0 | United Kingdom | 88253.0 | London (I) | 13.0 | Food | 16468.251 | 0.( |
| **7** | 10.2016 | 1.0 | United Kingdom | 88253.0 | London (I) | 7.0 | Clothing | 4698.471 | 0.( |
| **8** | 10.2016 | 1.0 | United Kingdom | 88253.0 | London (I) | 8.0 | Household | 1183.272 | 0.( |
| **9** | 10.2016 | 1.0 | United Kingdom | 88253.0 | London (I) | 9.0 | Hardware | 2029.815 | 0.( |
| **10** | 10.2016 | 1.0 | United Kingdom | 88253.0 | London (I) | 14.0 | Non Food | 7911.558 | 0.( |
| **11** | 10.2016 | 1.0 | United Kingdom | 88253.0 | London (I) | 15.0 | Admin | 4308.243 | 0.( |
| **12** | 10.2016 | 1.0 | United Kingdom | 88253.0 | London (I) | 12.0 | Checkout | 5825.097 | 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.( |
| **16** | 10.2016 | 1.0 | United Kingdom | 88253.0 | London (I) | 18.0 | all | 40086.486 | 0.( |
| **17** | 10.2016 | 1.0 | United Kingdom | 38976.0 | Manchester | 1.0 | Dry | 2583.687 | 0.( |
| **18** | 10.2016 | 1.0 | United Kingdom | 38976.0 | Manchester | 2.0 | Frozen | 5145.345 | 0.( |
| **19** | 10.2016 | 1.0 | United Kingdom | 38976.0 | Manchester | 3.0 | other | 47.205 | 0.( |

Print tail last 7 rows

In [13]:

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

Out[13]:

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

To print statistical data

In [14]:

```
data.describe()
```

Out[14]:

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

To print rows and coloum

In [15]:

```
np.shape(data)
```

Out[15]:

```
(7658, 14)
```

To print no. of elements

In [16]:

```
np.size(data)
```

Out[16]:

```
107212
```

To print missing values

In [17]:

```
data.isna()
```

Out[17]:

| | MonthYear | Time index | Country | StoreID | City | Dept_ID | Dept. Name | HoursOwn | HoursLease | Sales units |
|---|---|---|---|---|---|---|---|---|---|---|
| **0** | False | False | False | False | False | False | False | False | False | False |
| **1** | False | False | False | False | False | False | False | False | False | False |
| **2** | False | False | False | False | False | False | False | False | False | False |
| **3** | False | False | False | False | False | False | False | False | False | False |
| **4** | False | False | False | False | False | False | False | False | False | False |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... | ... | ... |
| **7653** | False | False | False | False | False | False | False | False | False | False |
| **7654** | False | False | False | False | False | False | False | False | False | False |
| **7655** | False | False | False | False | False | False | False | False | False | False |
| **7656** | False | False | False | False | False | False | False | False | False | False |
| **7657** | False | False | False | False | False | False | False | False | False | False |

7658 rows × 14 columns

Filla a value 60 in missing place

```
In [18]: data.fillna(value=60)
```

Out[18]:

| | 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** | 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 | |

7658 rows × 14 columns

```
In [19]: import matplotlib.pyplot as pp
In [20]: dd=data[['HoursOwn','Turnover']]
         dd
```

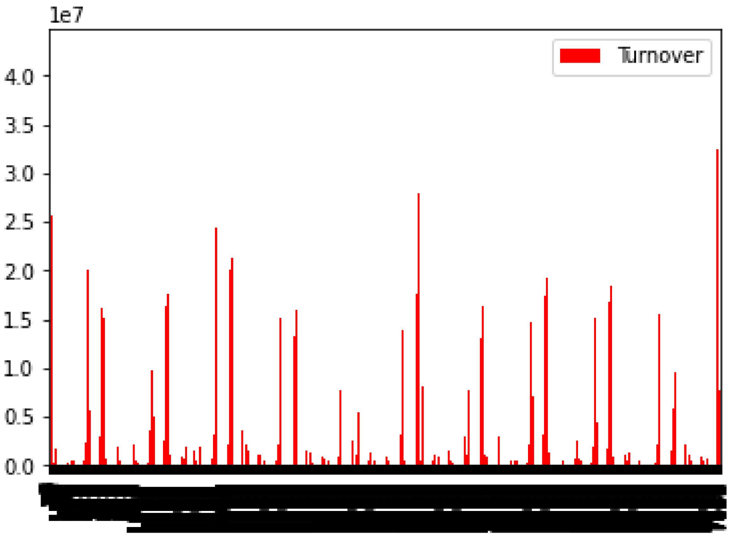
Out[20]:

| | HoursOwn | Turnover |
|---|---|---|
| **0** | 3184.764 | 1226244.0 |
| **1** | 1582.941 | 387810.0 |
| **2** | 47.205 | 654657.0 |
| **3** | 1623.852 | 499434.0 |
| **4** | 1759.173 | 329397.0 |
| **...** | ... | ... |
| **7653** | 6322.323 | 14538825.0 |
| **7654** | 4270.479 | 0.0 |
| **7655** | 0 | 0.0 |
| **7656** | 2224.929 | 0.0 |
| **7657** | 39652.2 | 15056214.0 |

7658 rows × 2 columns

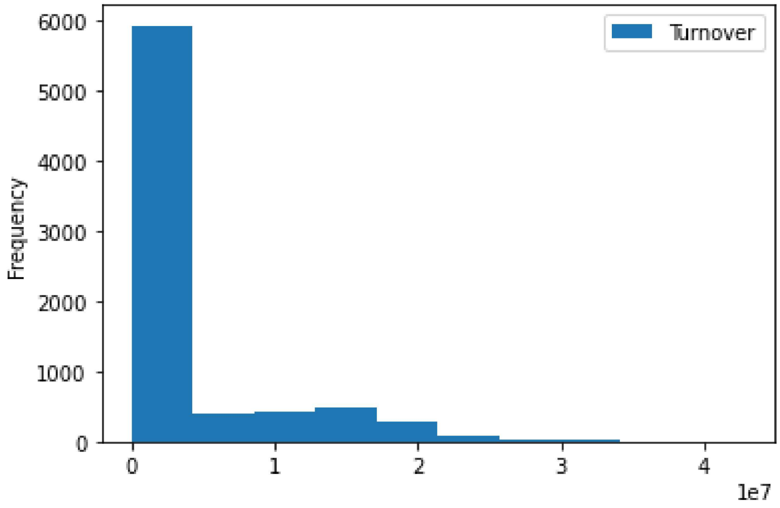
```
In [21]: dd.plot.bar(color='r')
```

```
Out[21]: <AxesSubplot:>
```



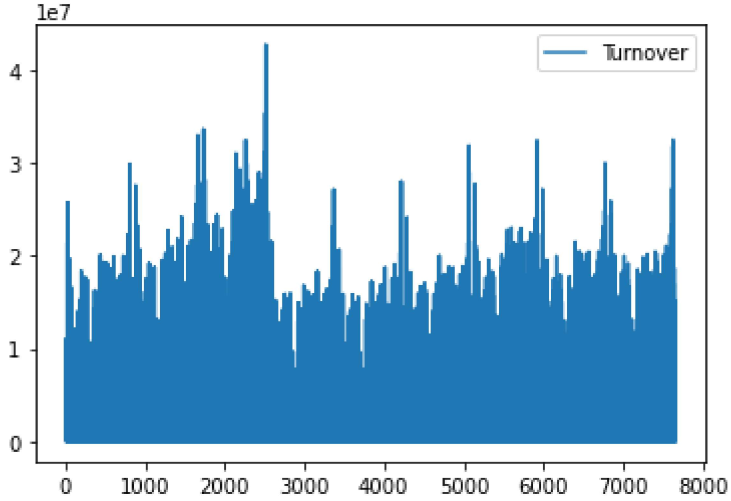
In [25]:
```
dd.plot.hist()
```

Out[25]:
```
<AxesSubplot:ylabel='Frequency'>
```



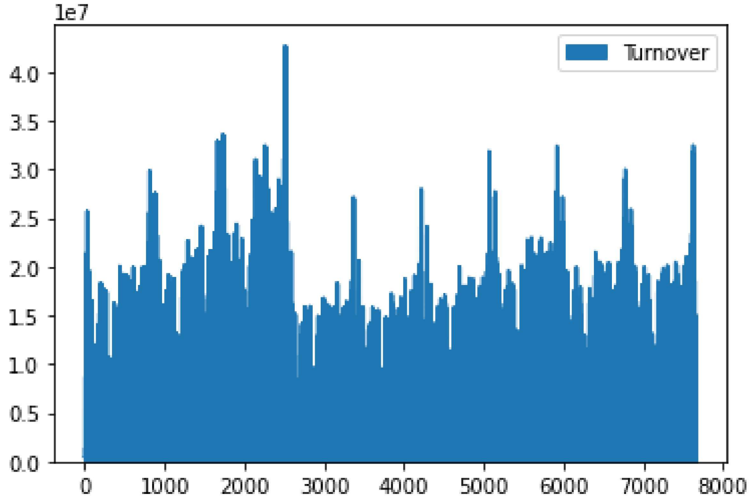
In [26]:
```
dd.plot.line()
```

Out[26]:
```
<AxesSubplot:>
```



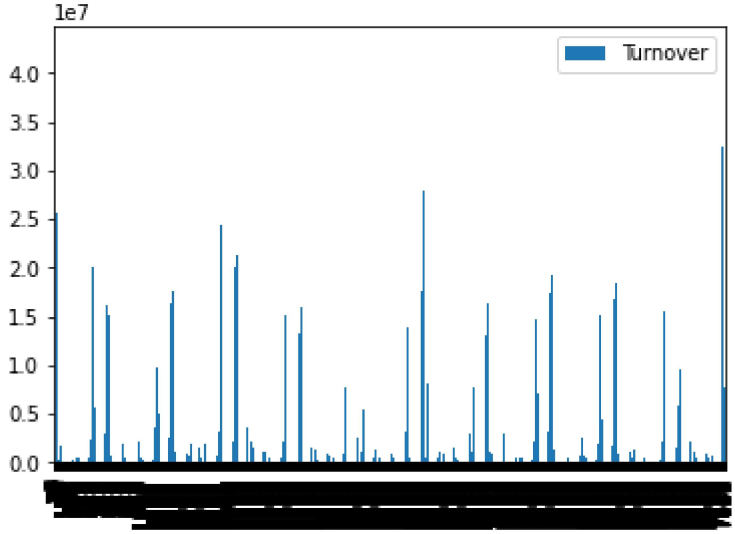
```
In [27]: dd.plot.area()
```

```
Out[27]: <AxesSubplot:>
```



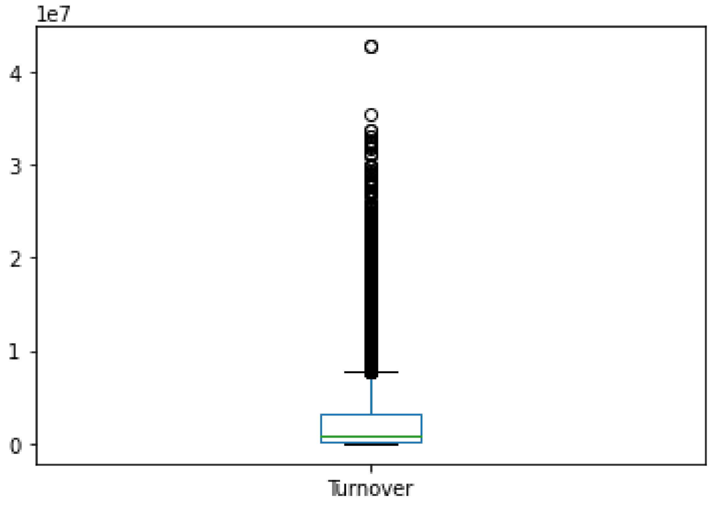
```
In [28]: dd.plot.bar()
```

```
Out[28]: <AxesSubplot:>
```



```
In [31]: dd.plot.box()
```

```
Out[31]: <AxesSubplot:>
```



```
In [ ]:
```