```
mk 25-07-23
```

In [45]:
```
import numpy as np
import pandas as pd
```

In [46]:
```
a=pd.read_csv(r"C:\Users\user\Downloads\6_Salesworkload1.csv")
a
```

Out[46]:

	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 [47]:
```
a=a.head(50)
```

In [48]:
```
a.sum()
```

Out[48]:
```
MonthYear        10.201610.201610.201610.201610.201610.201610.2...
Time index                                                    50.0
Country          United KingdomUnited KingdomUnited KingdomUnit...
StoreID                                                 2445245.0
City             London (I)London (I)London (I)London (I)London...
Dept_ID                                                     465.0
Dept. Name       DryFrozenotherFishFruits & VegetablesMeatFoodC...
HoursOwn         3184.7641582.94147.2051623.8521759.1738270.316...
HoursLease                                                 3048.0
Sales units                                              61592195.0
Turnover                                                203349924.0
Customer                                                        0.0
Area (m2)        953.04720.48966.721053.361053.3611735.1619865....
Opening hours    Type AType AType AType AType AType AType AType...
dtype: object
```

In [49]:
```
a.mean()
```

Out[49]:
```
Time index           1.00
StoreID          48904.90
Dept_ID              9.30
HoursLease          60.96
Sales units    1231843.90
Turnover       4066998.48
Customer              NaN
dtype: float64
```

In [50]:
```
a.median()
```

Out[50]:
```
MonthYear          10.2016
Time index          1.0000
StoreID         38976.0000
Dept_ID             9.0000
HoursOwn         3351.5550
HoursLease          0.0000
Sales units    309342.5000
Turnover       740052.0000
Customer               NaN
Area (m2)        1442.1000
dtype: float64
```

In [51]:
```
a.mode()
```

Out[51]:

	MonthYear	Time index	Country	StoreID	City	Dept_ID	Dept. Name	HoursOwn	HoursLease
0	10.2016	1.0	United Kingdom	38976.0	London (I)	1.0	Admin	0	0.0
1	NaN	NaN	NaN	88253.0	Manchester	2.0	Checkout	47.205	NaN
2	NaN	NaN	NaN	NaN	NaN	3.0	Clothing	NaN	NaN
3	NaN	NaN	NaN	NaN	NaN	4.0	Customer Services	NaN	NaN
4	NaN	NaN	NaN	NaN	NaN	5.0	Delivery	NaN	NaN
5	NaN	NaN	NaN	NaN	NaN	6.0	Dry	NaN	NaN
6	NaN	NaN	NaN	NaN	NaN	7.0	Fish	NaN	NaN
7	NaN	NaN	NaN	NaN	NaN	8.0	Food	NaN	NaN
8	NaN	NaN	NaN	NaN	NaN	9.0	Frozen	NaN	NaN
9	NaN	NaN	NaN	NaN	NaN	11.0	Fruits & Vegetables	NaN	NaN
10	NaN	NaN	NaN	NaN	NaN	12.0	Hardware	NaN	NaN
11	NaN	NaN	NaN	NaN	NaN	13.0	Household	NaN	NaN
12	NaN	NaN	NaN	NaN	NaN	14.0	Meat	NaN	NaN
13	NaN	NaN	NaN	NaN	NaN	15.0	Non Food	NaN	NaN
14	NaN	NaN	NaN	NaN	NaN	16.0	other	NaN	NaN
15	NaN	NaN	NaN	NaN	NaN	17.0	others	NaN	NaN

In [52]:
```
a.describe()
```

Out[52]:

	Time index	StoreID	Dept_ID	HoursLease	Sales units	Turnover	Customer
count	50.0	50.000000	50.000000	50.000000	5.000000e+01	5.000000e+01	0.0
mean	1.0	48904.900000	9.300000	60.960000	1.231844e+06	4.066998e+06	NaN
std	0.0	29839.520941	5.304022	213.640644	2.088301e+06	6.868434e+06	NaN
min	1.0	17647.000000	1.000000	0.000000	0.000000e+00	0.000000e+00	NaN
25%	1.0	17647.000000	5.000000	0.000000	5.504125e+04	1.477058e+05	NaN
50%	1.0	38976.000000	9.000000	0.000000	3.093425e+05	7.400520e+05	NaN
75%	1.0	88253.000000	14.000000	0.000000	9.128262e+05	3.521022e+06	NaN
max	1.0	88253.000000	18.000000	1152.000000	7.476680e+06	2.571973e+07	NaN

In [53]:
```
a.cumsum()
```

46	10.201610.201610.201610.201610.201610.201610.2...	47.0	United KingdomUnited KingdomUnited KingdomUnit...	2392304.0	(I)Lo (I)Lo (I)Lo (I)Lond
47	10.201610.201610.201610.201610.201610.201610.2...	48.0	United KingdomUnited KingdomUnited KingdomUnit...	2409951.0	Lo (I)Lo (I)Lo (I)Lo (I)Lond
48	10.201610.201610.201610.201610.201610.201610.2...	49.0	United KingdomUnited KingdomUnited KingdomUnit...	2427598.0	Lo (I)Lo (I)Lo (I)Lo (I)Lond

In [54]:
```
a.count()
```

Out[54]:
```
MonthYear        50
Time index       50
Country          50
StoreID          50
City             50
Dept_ID          50
Dept. Name       50
HoursOwn         50
HoursLease       50
Sales units      50
Turnover         50
Customer          0
Area (m2)        50
Opening hours    50
dtype: int64
```

In [55]:
```
a.min()
```

Out[55]:
```
MonthYear               10.2016
Time index                  1.0
Country          United Kingdom
StoreID                 17647.0
City                   Liverpool
Dept_ID                     1.0
Dept. Name                Admin
HoursOwn                      0
HoursLease                  0.0
Sales units                 0.0
Turnover                    0.0
Customer                    NaN
Area (m2)                     0
Opening hours            Type A
dtype: object
```

In [56]:
```
a.max()
```

Out[56]:
```
MonthYear               10.2016
Time index                  1.0
Country          United Kingdom
StoreID                 88253.0
City                 Manchester
Dept_ID                     18.0
Dept. Name                others
HoursOwn                8965.803
HoursLease                1152.0
Sales units             7476680.0
Turnover               25719732.0
Customer                    NaN
Area (m2)                987.24
Opening hours            Type A
dtype: object
```

In [57]:
```
from numpy import cov
```

In [58]:
```
cov(a['StoreID'])
```

Out[58]:
```
array(8.9039701e+08)
```

In [61]:
```
from scipy.stats import pearsonr
pearsonr(a['Turnover'],a['StoreID'])
```

Out[61]:
```
(-0.10648035357395985, 0.4617414817945737)
```

In [60]:
```
from scipy.stats import spearmanr
spearmanr(a['StoreID'],a['Turnover'])
```

Out[60]:
```
SpearmanrResult(correlation=-0.043899316897637874, pvalue=0.7621110891029952)
```

In [ ]: