

25-07-2023

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

```
In [37]: a=pd.read_csv(r"C:\Users\user\Downloads\6_Salesworkload1 - 6_Salesworkload1.csv")
a
```

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

7658 rows x 9 columns

```
In [67]: a=a.head(20)
a
```

12	10.2016	1.0	United Kingdom	88253.0	London (I)	12.0	Checkout	5825.097
13	10.2016	1.0	United Kingdom	88253.0	London (I)	16.0	Customer Services	3320.085
14	10.2016	1.0	United Kingdom	88253.0	London (I)	11.0	Delivery	0
15	10.2016	1.0	United Kingdom	88253.0	London (I)	17.0	others	2253.252
16	10.2016	1.0	United Kingdom	88253.0	London (I)	18.0	all	40086.486
17	10.2016	1.0	United Kingdom	38976.0	Manchester	1.0	Dry	2583.687
18	10.2016	1.0	United Kingdom	38976.0	Manchester	2.0	Frozen	5145.345
19	10.2016	1.0	United Kingdom	38976.0	Manchester	3.0	other	47.205

In [68]: `a.sum()`

```
Out[68]: MonthYear      10.201610.201610.201610.201610.201610.201610.2...
Time index                                20.0
Country      United KingdomUnited KingdomUnited KingdomUnit...
StoreID      1617229.0
City         London (I)London (I)London (I)London (I)London...
Dept_ID      167.0
Dept. Name    DryFrozenotherFishFruits & VegetablesMeatFoodC...
HoursOwn     3184.7641582.94147.2051623.8521759.1738270.316...
HoursLease    0.0
Sales units   15826520.0
Turnover     50044494.0
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 [69]: `a.mean()`

```
Out[69]: Time index      1.00
StoreID      80861.45
Dept_ID      8.35
HoursLease    0.00
Sales units   791326.00
Turnover     2502224.70
Customer     NaN
dtype: float64
```

In [70]: `a.median()`

```
Out[70]: MonthYear      10.2016
Time index      1.0000
StoreID      88253.0000
Dept_ID      7.5000
HoursOwn     2884.2255
HoursLease    0.0000
Sales units   263175.0000
Turnover     940450.5000
Customer     NaN
Area (m2)    1146.8400
dtype: float64
```

In [71]: `a.mode()`

Out[71]:

	MonthYear	Time index	Country	StoreID	City	Dept_ID	Dept. Name	HoursOwn	HoursLease	Sales units
0	10.2016	1.0	United Kingdom	88253.0	London (I)	1.0	Dry	47.205	0.0	0.0
1	NaN	NaN	NaN	NaN	NaN	2.0	Frozen	NaN	NaN	NaN
2	NaN	NaN	NaN	NaN	NaN	3.0	other	NaN	NaN	NaN

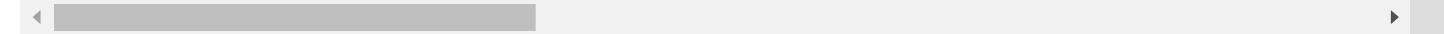
```
In [72]: a.describe()
```

Out[72]:

	Time index	StoreID	Dept_ID	HoursLease	Sales units	Turnover	Customer
count	20.0	20.000000	20.000000	20.0	2.000000e+01	2.000000e+01	0.0
mean	1.0	80861.450000	8.350000	0.0	7.913260e+05	2.502225e+06	NaN
std	0.0	18052.508149	5.751659	0.0	1.173690e+06	3.629374e+06	NaN
min	1.0	38976.000000	1.000000	0.0	0.000000e+00	0.000000e+00	NaN
25%	1.0	88253.000000	3.000000	0.0	5.817375e+04	3.196478e+05	NaN
50%	1.0	88253.000000	7.500000	0.0	2.631750e+05	9.404505e+05	NaN
75%	1.0	88253.000000	13.250000	0.0	8.439088e+05	2.428861e+06	NaN
max	1.0	88253.000000	18.000000	0.0	3.435790e+06	1.107044e+07	NaN

```
In [73]: a.cumsum()
```

17	10.201610.201610.201610.201610.201610.201610.2...	18.0	United KingdomUnited KingdomUnit...	1539277.0	(I)Lc (I)Lc (I)Lon
18	10.201610.201610.201610.201610.201610.201610.2...	19.0	United KingdomUnited KingdomUnit...	1578253.0	Lc (I)Lc (I)Lc (I)Lc (I)Lon
19	10.201610.201610.201610.201610.201610.201610.2...	20.0	United KingdomUnited KingdomUnit...	1617229.0	Lc (I)Lc (I)Lc (I)Lc (I)Lon



```
In [65]: a.count()
```

```
Out[65]: MonthYear      7658  
Time index      7650  
Country         7650  
StoreID         7650  
City            7650  
Dept_ID         7650  
Dept. Name      7650  
HoursOwn        7650  
HoursLease      7650  
Sales units     7650  
Turnover        7650  
Customer         0  
Area (m2)       7650  
Opening hours   7650  
dtype: int64
```

```
In [47]: a.min()
```

```
Out[47]: MonthYear      - - - -  
Time index           1.0  
StoreID            12227.0  
Dept_ID            1.0  
HoursLease          0.0  
Sales units         0.0  
Turnover            0.0  
Customer            NaN  
dtype: object
```

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

```
Out[48]: MonthYear      6.2017  
Time index           9.0  
StoreID            98422.0  
Dept_ID            18.0  
HoursLease         3984.0  
Sales units       11242955.0  
Turnover          42717390.0  
Customer            NaN  
dtype: object
```

```
In [49]: from numpy import cov
```

```
In [74]: cov(a['Sales units'])
```

```
Out[74]: array(1.37754829e+12)
```

```
In [75]: from scipy.stats import pearsonr  
pearsonr(a['Turnover'],a['StoreID'])
```

```
Out[75]: (0.06903160386857898, 0.7724395217668092)
```

```
In [78]: from scipy.stats import spearmanr  
spearmanr(a['StoreID'],a['Turnover'])
```

```
Out[78]: SpearmanrResult(correlation=-0.25604235255957397, pvalue=0.2758811450911547)
```