In [1]:
```
import numpy as np
import pandas as pd
from numpy import linalg as lg
from numpy import cov
```

# 5 Data Sets

In [2]:
```
a=pd.read_csv("Sales.csv")
a
```

Out[2]:

| | MonthYear | Time index | Country | StoreID | City | Dept_ID | Dept. Name | HoursOwn | HoursLeas |
|---|---|---|---|---|---|---|---|---|---|
| **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. |
| **...** | ... | ... | ... | ... | ... | ... | ... | ... | . |
| **7653** | 06.2017 | 9.0 | Sweden | 29650.0 | Gothenburg | 12.0 | Checkout | 6322.323 | 0. |
| **7654** | 06.2017 | 9.0 | Sweden | 29650.0 | Gothenburg | 16.0 | Customer Services | 4270.479 | 0. |
| **7655** | 06.2017 | 9.0 | Sweden | 29650.0 | Gothenburg | 11.0 | Delivery | 0 | 0. |
| **7656** | 06.2017 | 9.0 | Sweden | 29650.0 | Gothenburg | 17.0 | others | 2224.929 | 0. |
| **7657** | 06.2017 | 9.0 | Sweden | 29650.0 | Gothenburg | 18.0 | all | 39652.2 | 0. |

7658 rows × 14 columns

# a) Find mean, median, mode and describe

In [3]:
```
print(a.mean)
print("Median:")
print(a.median)
print("Mode:")
print(a.mode)
print("Describe")
print(a.describe())
```

```
<bound method NDFrame._add_numeric_operations.<locals>.mean of      MonthYear  Time
index         Country  StoreID        City  Dept_ID  \
0      10.2016         1.0  United Kingdom  88253.0  London (I)      1.0
1      10.2016         1.0  United Kingdom  88253.0  London (I)      2.0
2      10.2016         1.0  United Kingdom  88253.0  London (I)      3.0
```

```
3      10.2016         1.0  United Kingdom  88253.0  London (I)      4.0
4      10.2016         1.0  United Kingdom  88253.0  London (I)      5.0
...        ...         ...             ...      ...         ...      ...
7653   06.2017         9.0          Sweden  29650.0  Gothenburg     12.0
7654   06.2017         9.0          Sweden  29650.0  Gothenburg     16.0
7655   06.2017         9.0          Sweden  29650.0  Gothenburg     11.0
7656   06.2017         9.0          Sweden  29650.0  Gothenburg     17.0
7657   06.2017         9.0          Sweden  29650.0  Gothenburg     18.0

                Dept. Name  HoursOwn  HoursLease  Sales units    Turnover  \
0                      Dry  3184.764         0.0     398560.0   1226244.0
1                   Frozen  1582.941         0.0      82725.0    387810.0
2                    other    47.205         0.0     438400.0    654657.0
3                     Fish  1623.852         0.0     309425.0    499434.0
4     Fruits & Vegetables  1759.173         0.0     165515.0    329397.0
...                    ...       ...         ...          ...         ...
7653             Checkout  6322.323         0.0    3886530.0  14538825.0
7654    Customer Services  4270.479         0.0        245.0         0.0
7655             Delivery         0         0.0          0.0         0.0
7656               others  2224.929         0.0        245.0         0.0
7657                  all   39652.2         0.0    3886530.0  15056214.0

      Customer Area (m2) Opening hours
0          NaN    953.04        Type A
1          NaN    720.48        Type A
2          NaN    966.72        Type A
3          NaN   1053.36        Type A
4          NaN   1053.36        Type A
...        ...       ...           ...
7653       NaN       #NV        Type A
7654       NaN       #NV        Type A
7655       NaN       #NV        Type A
7656       NaN       #NV        Type A
7657       NaN       #NV        Type A

[7658 rows x 14 columns]>
Median:
<bound method NDFrame._add_numeric_operations.<locals>.median of      MonthYear  Tim
e index         Country  StoreID        City  Dept_ID  \
0      10.2016         1.0  United Kingdom  88253.0  London (I)      1.0
1      10.2016         1.0  United Kingdom  88253.0  London (I)      2.0
2      10.2016         1.0  United Kingdom  88253.0  London (I)      3.0
3      10.2016         1.0  United Kingdom  88253.0  London (I)      4.0
4      10.2016         1.0  United Kingdom  88253.0  London (I)      5.0
...        ...         ...             ...      ...         ...      ...
7653   06.2017         9.0          Sweden  29650.0  Gothenburg     12.0
7654   06.2017         9.0          Sweden  29650.0  Gothenburg     16.0
7655   06.2017         9.0          Sweden  29650.0  Gothenburg     11.0
7656   06.2017         9.0          Sweden  29650.0  Gothenburg     17.0
7657   06.2017         9.0          Sweden  29650.0  Gothenburg     18.0

                Dept. Name  HoursOwn  HoursLease  Sales units    Turnover  \
0                      Dry  3184.764         0.0     398560.0   1226244.0
1                   Frozen  1582.941         0.0      82725.0    387810.0
2                    other    47.205         0.0     438400.0    654657.0
3                     Fish  1623.852         0.0     309425.0    499434.0
4     Fruits & Vegetables  1759.173         0.0     165515.0    329397.0
...                    ...       ...         ...          ...         ...
7653             Checkout  6322.323         0.0    3886530.0  14538825.0
7654    Customer Services  4270.479         0.0        245.0         0.0
7655             Delivery         0         0.0          0.0         0.0
7656               others  2224.929         0.0        245.0         0.0
7657                  all   39652.2         0.0    3886530.0  15056214.0

      Customer Area (m2) Opening hours
0          NaN    953.04        Type A
1          NaN    720.48        Type A
2          NaN    966.72        Type A
3          NaN   1053.36        Type A
```

```
4          NaN   1053.36        Type A
...        ...       ...           ...
7653       NaN       #NV        Type A
7654       NaN       #NV        Type A
7655       NaN       #NV        Type A
7656       NaN       #NV        Type A
7657       NaN       #NV        Type A

[7658 rows x 14 columns]>
Mode:
<bound method DataFrame.mode of      MonthYear  Time index         Country  StoreID
City  Dept_ID  \
0      10.2016         1.0  United Kingdom  88253.0  London (I)      1.0
1      10.2016         1.0  United Kingdom  88253.0  London (I)      2.0
2      10.2016         1.0  United Kingdom  88253.0  London (I)      3.0
3      10.2016         1.0  United Kingdom  88253.0  London (I)      4.0
4      10.2016         1.0  United Kingdom  88253.0  London (I)      5.0
...        ...         ...             ...      ...         ...      ...
7653   06.2017         9.0          Sweden  29650.0  Gothenburg     12.0
7654   06.2017         9.0          Sweden  29650.0  Gothenburg     16.0
7655   06.2017         9.0          Sweden  29650.0  Gothenburg     11.0
7656   06.2017         9.0          Sweden  29650.0  Gothenburg     17.0
7657   06.2017         9.0          Sweden  29650.0  Gothenburg     18.0

               Dept. Name  HoursOwn  HoursLease  Sales units    Turnover  \
0                     Dry  3184.764         0.0     398560.0   1226244.0
1                  Frozen  1582.941         0.0      82725.0    387810.0
2                   other    47.205         0.0     438400.0    654657.0
3                    Fish  1623.852         0.0     309425.0    499434.0
4     Fruits & Vegetables  1759.173         0.0     165515.0    329397.0
...                    ...       ...         ...          ...         ...
7653             Checkout  6322.323         0.0    3886530.0  14538825.0
7654    Customer Services  4270.479         0.0        245.0         0.0
7655             Delivery         0         0.0          0.0         0.0
7656               others  2224.929         0.0        245.0         0.0
7657                  all   39652.2         0.0    3886530.0  15056214.0

      Customer Area (m2) Opening hours
0          NaN    953.04        Type A
1          NaN    720.48        Type A
2          NaN    966.72        Type A
3          NaN   1053.36        Type A
4          NaN   1053.36        Type A
...        ...       ...           ...
7653       NaN       #NV        Type A
7654       NaN       #NV        Type A
7655       NaN       #NV        Type A
7656       NaN       #NV        Type A
7657       NaN       #NV        Type A

[7658 rows x 14 columns]>
Describe
        Time index       StoreID      Dept_ID   HoursLease   Sales units  \
count  7650.000000   7650.000000  7650.000000  7650.000000  7.650000e+03
mean      5.000000  61995.220000     9.470588    22.036078  1.076471e+06
std       2.582158  29924.581631     5.337429   133.299513  1.728113e+06
min       1.000000  12227.000000     1.000000     0.000000  0.000000e+00
25%       3.000000  29650.000000     5.000000     0.000000  5.457125e+04
50%       5.000000  75400.500000     9.000000     0.000000  2.932300e+05
75%       7.000000  87703.000000    14.000000     0.000000  9.175075e+05
max       9.000000  98422.000000    18.000000  3984.000000  1.124296e+07

            Turnover  Customer
count  7.650000e+03       0.0
mean   3.721393e+06       NaN
std    6.003380e+06       NaN
min    0.000000e+00       NaN
25%    2.726798e+05       NaN
50%    9.319575e+05       NaN
```

```
75%     3.264432e+06       NaN
max     4.271739e+07       NaN
```

# b) Find sum(), cumsum(), count, min and max values

In [4]:
```
print(a.sum())
```

```
MonthYear      10.201610.201610.201610.201610.201610.201610.2...
Time index                                                38250.0
StoreID                                                474263433.0
Dept_ID                                                    72450.0
HoursLease                                                168576.0
Sales units                                            8235000965.0
Turnover                                              28468656015.0
Customer                                                     0.0
dtype: object
```

In [10]:
```
b=a.head(90)
print(b.cumsum())
```

```
                                           MonthYear  Time index  \
0                                            10.2016         1.0
1                                       10.201610.2016         2.0
2                                  10.201610.201610.2016         3.0
3                             10.201610.201610.201610.2016         4.0
4                        10.201610.201610.201610.201610.2016         5.0
..                                                      ...         ...
85  10.201610.201610.201610.201610.201610.201610.2...        86.0
86  10.201610.201610.201610.201610.201610.201610.2...        87.0
87  10.201610.201610.201610.201610.201610.201610.2...        88.0
88  10.201610.201610.201610.201610.201610.201610.2...        89.0
89  10.201610.201610.201610.201610.201610.201610.2...        90.0

                                              Country    StoreID  \
0                                      United Kingdom    88253.0
1                          United KingdomUnited Kingdom   176506.0
2           United KingdomUnited KingdomUnited Kingdom   264759.0
3   United KingdomUnited KingdomUnited KingdomUnit...   353012.0
4   United KingdomUnited KingdomUnited KingdomUnit...   441265.0
..                                                ...        ...
85  United KingdomUnited KingdomUnited KingdomUnit...  4114822.0
86  United KingdomUnited KingdomUnited KingdomUnit...  4133630.0
87  United KingdomUnited KingdomUnited KingdomUnit...  4152438.0
88  United KingdomUnited KingdomUnited KingdomUnit...  4171246.0
89  United KingdomUnited KingdomUnited KingdomUnit...  4190054.0

                                                  City  Dept_ID  \
0                                           London (I)      1.0
1                                 London (I)London (I)      3.0
2                       London (I)London (I)London (I)      6.0
3             London (I)London (I)London (I)London (I)     10.0
4   London (I)London (I)London (I)London (I)London...     15.0
..                                                ...      ...
85  London (I)London (I)London (I)London (I)London...    806.0
86  London (I)London (I)London (I)London (I)London...    808.0
87  London (I)London (I)London (I)London (I)London...    811.0
88  London (I)London (I)London (I)London (I)London...    815.0
89  London (I)London (I)London (I)London (I)London...    820.0

                                         Dept. Name  \
0                                               Dry
1                                         DryFrozen
2                                    DryFrozenother
```

```
3                                DryFrozenotherFish
4             DryFrozenotherFishFruits & Vegetables
..                                                ...
85  DryFrozenotherFishFruits & VegetablesMeatFoodC...
86  DryFrozenotherFishFruits & VegetablesMeatFoodC...
87  DryFrozenotherFishFruits & VegetablesMeatFoodC...
88  DryFrozenotherFishFruits & VegetablesMeatFoodC...
89  DryFrozenotherFishFruits & VegetablesMeatFoodC...

                                           HoursOwn  HoursLease  \
0                                           3184.764         0.0
1                                     3184.7641582.941         0.0
2                                  3184.7641582.94147.205         0.0
3                         3184.7641582.94147.2051623.852         0.0
4                3184.7641582.94147.2051623.8521759.173         0.0
..                                                ...         ...
85  3184.7641582.94147.2051623.8521759.1738270.316...      6288.0
86  3184.7641582.94147.2051623.8521759.1738270.316...      6288.0
87  3184.7641582.94147.2051623.8521759.1738270.316...      6288.0
88  3184.7641582.94147.2051623.8521759.1738270.316...      6288.0
89  3184.7641582.94147.2051623.8521759.1738270.316...      6288.0

    Sales units     Turnover  Customer  \
0      398560.0    1226244.0       NaN
1      481285.0    1614054.0       NaN
2      919685.0    2268711.0       NaN
3     1229110.0    2768145.0       NaN
4     1394625.0    3097542.0       NaN
..          ...          ...       ...
85   96145695.0  322567602.0       NaN
86   96238930.0  322624806.0       NaN
87   96728045.0  323392302.0       NaN
88   97071970.0  324047673.0       NaN
89   97227435.0  324399396.0       NaN

                                             Area (m2)  \
0                                              953.04
1                                        953.04720.48
2                                     953.04720.48966.72
3                            953.04720.48966.721053.36
4                   953.04720.48966.721053.361053.36
..                                                ...
85  953.04720.48966.721053.361053.3611735.1619865....
86  953.04720.48966.721053.361053.3611735.1619865....
87  953.04720.48966.721053.361053.3611735.1619865....
88  953.04720.48966.721053.361053.3611735.1619865....
89  953.04720.48966.721053.361053.3611735.1619865....

                                        Opening hours
0                                               Type A
1                                         Type AType A
2                                   Type AType AType A
3                             Type AType AType AType A
4                   Type AType AType AType AType A
..                                                ...
85  Type AType AType AType AType AType AType AType...
86  Type AType AType AType AType AType AType AType...
87  Type AType AType AType AType AType AType AType...
88  Type AType AType AType AType AType AType AType...
89  Type AType AType AType AType AType AType AType...

[90 rows x 14 columns]
```

In [6]:

```
print(a.count())
print(a.min())
print(a.max())
```

```
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
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
MonthYear         12.2016
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
```

# c) Find covariance and correlation (spearman and pearsons)

In [7]:
```
print(a.cov())
```

```
              Time index       StoreID       Dept_ID    HoursLease  \
Time index   6.667538e+00  2.094019e-10 -1.430561e-14 -3.438253e+01
StoreID      2.094019e-10  8.954806e+08  1.814186e-11 -4.798921e+04
Dept_ID     -1.430561e-14  1.814186e-11  2.848815e+01  9.613363e+01
HoursLease  -3.438253e+01 -4.798921e+04  9.613363e+01  1.776876e+04
Sales units -1.552564e+05  6.869293e+08  2.645877e+06  4.927821e+07
Turnover    -3.152632e+05  1.653862e+09  9.630443e+06  1.495198e+08
Customer              NaN           NaN           NaN           NaN

              Sales units      Turnover  Customer
Time index  -1.552564e+05 -3.152632e+05       NaN
StoreID      6.869293e+08  1.653862e+09       NaN
Dept_ID      2.645877e+06  9.630443e+06       NaN
HoursLease   4.927821e+07  1.495198e+08       NaN
Sales units  2.986375e+12  9.828553e+12       NaN
Turnover     9.828553e+12  3.604057e+13       NaN
Customer              NaN           NaN       NaN
```

In [19]:
```
from scipy.stats import pearsonr
print(pearsonr(b['StoreID'],b['MonthYear']))
```

```
(nan, nan)
```

In [11]:
```
from scipy.stats import spearmanr
print(spearmanr(a['Time index'],a['MonthYear']))
```

```
SpearmanrResult(correlation=-0.3601703830417159, pvalue=2.5854375801641448e-233)
```

In [ ]: