**import pandas as pd**

**import numpy as np**

**I.Pandas Library**

1. df = pd.read\_csv('link\_file',sep='\t') -> read dataset

2.head() -> read row head

3.shape ->view row,column

4.info() -> view information file data

5.columns / list(df.columns) -> read columns

6.index -> read index (start,stop,step)

6.describe() -> return someone value : count , mean ,min,max ....

7.describe(include="all") -> return someone value: mean,top,unique ,....

8.df.loc[df.name\_column == 'Dieu kien']

9.df.loc[(df.quantity == 2) & (df.item\_name == "Nantucket Nectar"),['order\_id','quantity','item\_name']]

10.iloc[vtristart:vitriend] -> return row can get

11.dtype -> check type

12.df.item\_price = df.item\_price.apply(lambda x : float(x.replace('$',''))) -> covert delete redundancy

13: df['total\_price'] = df['quantity'] \* df['item\_price'] ->head() -> return new column

14.revenue = df['total\_price'].sum() -> return sum of total\_price

15. a = df.groupby('item\_name')['quantity'].sum() -> return groupby

a.sort\_values(ascending = False).head(5) -> return 5 most-item used sorted

16.df.item\_name.value\_counts() -> value\_counts() -> return item difference

17.count() -> return dem vd: df.item\_name.value\_counts().count()

18.df.item\_name.nunique() -> nunique() -> return dem difference

II.Numpy Library

**import numpy as np**

1. a = np.array([4.14,4.2,2,3]) -> creat array
2. type(a) -> check type
3. np.array([1.2,3,4],dtype='float32') -> convert -> float
4. a2 = np.array([[1,2,3],[3,4,5]]) -> Creat array 2 chieu
5. shape -> change shape
6. ndim -> check so chieu
7. size
8. np.zeros([2,4],dtype=int)
9. np.ones([2,3],dtype=int)
10. np.arange(0,10,2,dtype=int)
11. np.full([3,5],6.9)
12. np.linspace(0.1,5)
13. np.random.random((4,4))
14. np.random.seed(0) -> creat random array ko bi thay doi gia tri qua cac lan random

np.random.random((4,4))

1. normal('mean','do lech chuan','size')
2. np.random.randint(0,10,[3,5]) | x1= np.random.randint(20,size=6)
3. np.random.rand(3,4)
4. x.T -> row <-> clumn
5. np.concatenate((x,y)) -> noi 2 array
6. vstack -> noi theo chieu doc vd: np.vstack((x,grid))
7. hstack -> noi theo chieu ngang
8. split() -> tach array
9. **Statistics - xac suat thong ke**
10. Std ()-> do lech chuan
11. Var() -> phuong sai
12. sqrt(np.var() -> do lech chuan can bac 2
13. argsort() -> return index cua cac element khi sort
14. np.sort(MatA,axis=0) #sort column
15. np.sort(MatA,axis=1) #sort row

# Linear Algebra - Dai do tuyen tinh

1. A.dot(B) -> matric A nhan Matrix B
2. A@B -> A nhan B tuong tu note :

**32 Example :**

np.random.seed(0) # tao random ko change value

sales\_amouts = np.random.randint(20,size=(5,3))

**#creat dataFrame : index la kieu tuong tu row , columns**

import pandas as pd

weekly\_sales = pd.DataFrame(sales\_amouts,index = ['Mon','Tues','Wed','Thurs','Fril'],

columns=['Almond Butter','Peanut Butter','Cashew Butter'])

weekly\_sales

#**create a price array tao them gia tri gia ca**

prices = np.array([10,8,12])

butter\_prices = pd.DataFrame(prices.reshape(1,3),index=['Price'],

columns=['Almond Butter','Peanut Butter','Cashew Butter'])

**#Lay dataFarme price \* dataFrame weekly\_sales theo dang matrix**

total\_prices = weekly\_sales.dot(butter\_prices.T)

#**them clumn totalPrice vao dataFrame weekly\_sales**

weekly\_sales['Total Price'] = total\_prices

weekly\_sales

III.Matplotlib

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

1. Hien thi

plt.style.available

plt.style.use('seaborn-whitegrid')

# backgroud for graph

plt.plot(); #display || plt.plot(x,y,color='red');