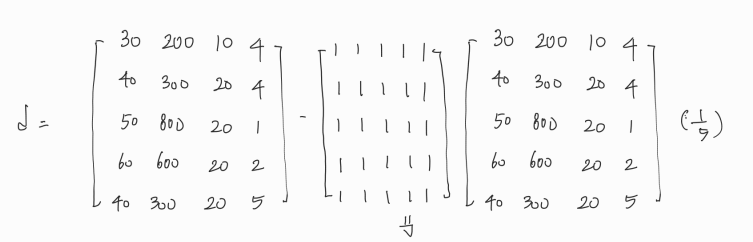
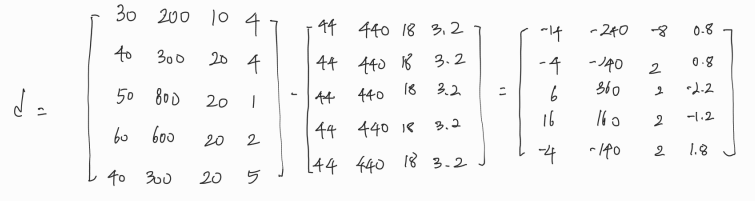
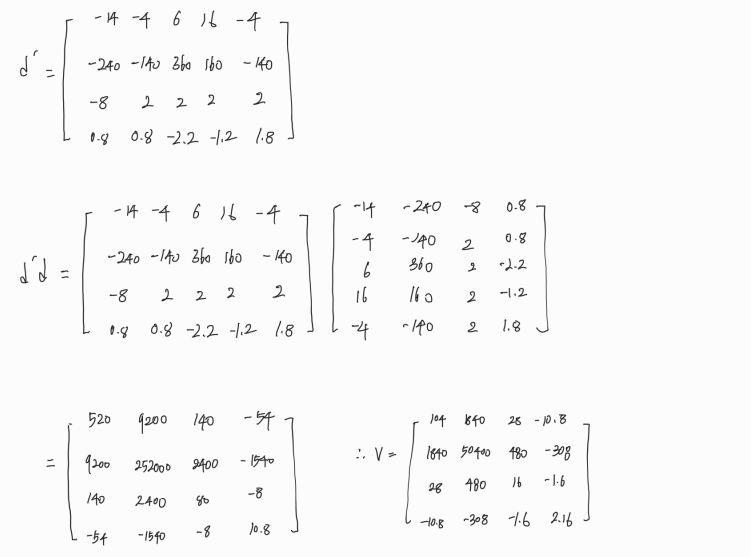
Exercise 1







Exercise 2

import numpy as np

import pandas as pd

import seaborn as sns

import matplotlib.pyplot as plt

np.set\_printoptions(precision=6, suppress=True)

# Set dataset array

dataset = np.array([(30,40,50,60,40),(200,300,800,600,300),(10,20,20,20,20),(4,4,1,2,5)])

# population covariance matrix

pcmatrix = np.cov(dataset,bias=True)

print("Population Covariance")

print(pcmatrix)

print()

# Seaborn visualizing population covariance matrix

sns.heatmap(pcmatrix,annot=True,fmt='g')

plt.show()

# sample covariance matrix

scmatrix = np.cov(dataset,bias=False)

print("Sample Covariance")

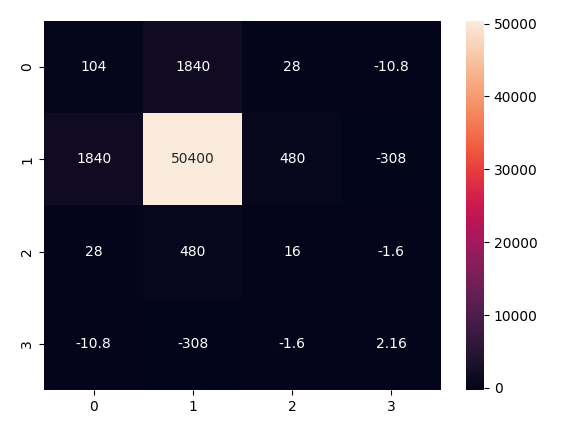
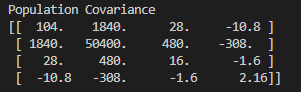
print(scmatrix)

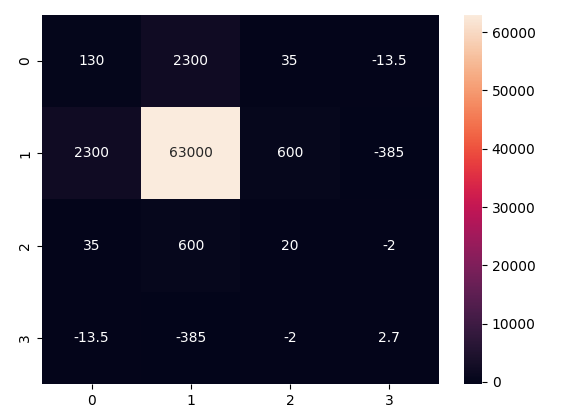
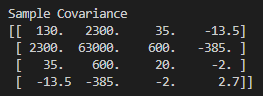
print()

# Seaborn visualizing sample covariance matrix

sns.heatmap(scmatrix,annot=True,fmt='g')

plt.show()





Homework

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Spends(x) | Income(y) |  |  |  |
| 2400 | 41200 | 98880000 | 5760000 | 1697440000 |
| 2650 | 50100 | 132765000 | 7022500 | 2510010000 |
| 2350 | 52000 | 122200000 | 5522500 | 2704000000 |
| 4950 | 66000 | 326700000 | 24502500 | 4356000000 |
| 3100 | 44500 | 137950000 | 9610000 | 1980250000 |
| 2500 | 37700 | 94250000 | 6250000 | 1421290000 |
| 5406 | 73500 | 375291000 | 26071236 | 5402250000 |
| 3100 | 37500 | 116250000 | 9610000 | 1406250000 |
| 2900 | 56700 | 164430000 | 8410000 | 3214890000 |
| 1750 | 35600 | 62300000 | 3062500 | 1267360000 |
|  |  |  |  |  |
| Σx | Σy | Σ | Σ | Σ |
| 30806 | 494800 | 1631016000 | 105821236 | 25959740000 |
|  |  |  |  |  |

M = 10 \* 1631016000 – 30806 \* 494800 / 10 \* 105821236 – 30806 \* 30806 = 9.774034574448894

B = 494800 \* 105821236 – 30806 \* 1631016000 / 10 \* 105821236 – 30806 \* 30806 = 19370.109089952737

Y = 9.774034574448894X + 19370.109089952737

**Input 3500**

9.77403457448894 \* 3500 + 19370.109089952737 = 53579.23010052387

**Input 5300**

9.77403457448894 \* 5300 + 19370.109089952737 = 71172.49233453188

import pandas as pd

import numpy as np

import matplotlib.pyplot as plt

import seaborn as sns

from sklearn import linear\_model

from sklearn import preprocessing

# set array x and y

x = np.array([2400, 2650, 2350, 4950, 3100, 2500, 5106, 3100, 2900, 1750])

y = np.array([41200, 50100, 52000, 66000, 44500, 37700, 73500, 37500, 56700, 35600])

# make dataframe

df = pd.DataFrame({'Spends(x)': x, 'Income(y)': y})

# find regression line and disply

E=linear\_model.LinearRegression()

E.fit(x[:,np.newaxis],y)

px=np.array([x.min()-1,x.max()+1])

py=E.predict(px[:,np.newaxis])

plt.scatter(x, y)

plt.plot(px,py,color="r")

plt.title('Linear Regression', fontsize=20)

plt.show()

