

과제연습문제

**** 아래의 내용을 파이썬 코드로 작성하여 그 코드와 결과를 제출하시오**

1. 아래의 표는 건강검진 대상자 12명에 대한 신장과 체중의 원 데이터와 평균보정 데이터를 보여주고 있다.

(1) 각각의 데이터에서 신장과 체중의 평균을 구하시오.

(2) 각각의 데이터에서 신장과 체중의 분산을 구하시오.

(3) 두 가지 데이터에 대해 평균과 분산의 차이를 비교하시오.

관측대상	원 데이터		평균보정 데이터	
	신장(X_1)	체중(X_2)	신장(X_1)	체중(X_2)
1	166.1732	55.6023	-1.2778	-10.3130
2	186.7266	87.3961	19.2756	21.4807
3	159.2694	55.2475	-8.1816	-10.6679
4	174.4309	71.6564	6.9799	5.7410
5	172.2638	65.7194	4.8128	-0.1959
6	148.988	53.0763	-18.463	-12.8391
7	159.2009	48.4501	-8.2501	-17.4653
8	171.779	70.6938	4.328	4.77839
9	163.1725	64.6079	-4.2785	-1.3075
10	165.4728	71.5016	-1.9782	5.5862
11	172.9416	71.4369	5.4906	5.5215
12	168.9933	75.5963	1.5423	9.6809

2. 앞의 표를 바탕으로 표준화 데이터를 만들고 평균과 분산을 계산하시오.

3. 앞의 표를 바탕으로 신장과 체중의 두 변수간의 공분산(covariance) 행렬을 만들어 보시오.

4. 아래의 표는 건강검진 대상자 15명에 대한 신장과 체중 데이터를 보여주고 있다.

(1) 신장과 체중에 대한 각각의 히스토그램을 그려 보시오

(2) 신장과 체중에 대한 상위 및 하위 사분위수(upper and lower quartile)를 각각 찾아 보시오.

(3) 히스토그램을 바탕으로 신장과 체중의 이상점(outlier)이 있는지를 확인해 보시오.

	신장	체중
1	166	56
2	186	87
3	159	55
4	174	72
5	172	66
6	149	53
7	159	48
8	204	139
9	163	65
10	165	72
11	173	71
12	184	76
13	178	80
14	175	75
15	180	81

5. 아래의 표는 가공기계의 고장이 발생하는 원인을 주요 요인별로 분석한 데이터이다.

발생 요인	기계 결함	부속품 불량	전기장치 불량	작동 오류
빈도	24	48	33	18

(1) 요인별 고장 발생 패턴을 Bar Chart로 도식화하시오.

(2) 요인별 고장 발생 패턴을 Pareto Chart로 도식화하시오.

(3) 요인별 고장 발생 패턴을 Pie Chart로 도식화하시오.