

공분산과 상관계수 코드

```
import numpy as np

H=np.array([180,170,160])
W=np.array([70,65,50])

H_mean=np.mean(H)
W_mean=np.mean(W)

cov=((H[0]-H_mean)*(W[0]-W_mean)+(H[1]-H_mean)*(W[1]-W_mean)+(H[2]-H_mean)*(W[2]-W_mean))/3
cov # 공분산

66.66666666666667

H_std=np.std(H)
W_std=np.std(W)

Rel=cov/(H_std*W_std)
Rel # 상관계수

0.9607689228305227
```

공분산은 2개의 상관관계를 구하는 것이다.

여기에서는 키와 몸무게의 상관관계를 구하는 것이다.

공분산을 구하면 covariance 공식=(x1-x의평균)*(y1-y의평균)+(x2-x의평균)*(y2-y의평균),...,/데이터의개수

66.66666666666667 이 나왔다. 이거에대해서 수치가 감이 안잡히니까 스케일링후 익숙하게 고쳐서 상관계수를 구합니다.

절대값이 높으면 높을수록 높은상관관계를 가진다고 보여줍니다.

1에가까울수록 양의상관관계 -1에가까울수록 음의 상관관계 0에가까울수록 상관관계없음 독립관계입니다 .

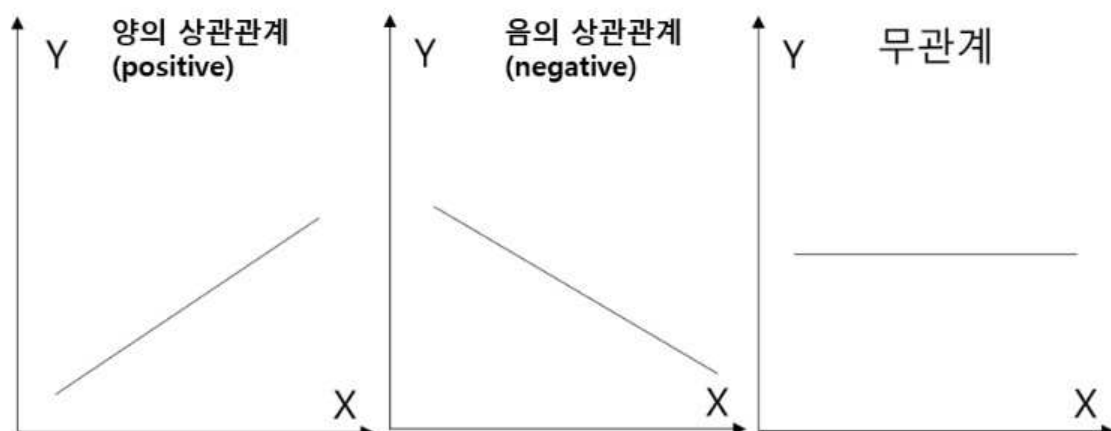
여기 결과에서 보이는 것은 양의상관관계이기 때문에 키가클수록 몸무게가 크다.라고 보여줍니다. 또한나 만약에 -1에가까웠다면 키가클수록 몸무게는 작아지는결과가나옵니다

보통 1에 가까움 → 강한 양의 상관 관계 (X 증가 시 Y도 증가)

0에 가까움 → 상관 관계 없음

-1에 가까움 → 강한 음의 상관 관계 (X 증가 시 Y 감소)

이렇게 보여줍니다.



<상관 관계 그래프>