1.

H0 : B0=0

H1: B0!=0

TV와 radio 에서는 p-value가 0.0001로 H0를 기각하는 것이 합당하고 sales와 상관성이 있는 것으로 유추할 수 있음.

newspaper에서는 p-value가 0.8599로 H0를 기각할 수 없음. Sales와 상관성이 없는 것으로 유추 가능.

2.

KNN classifier: 새로운 데이터가 주어졌을 때 X에 가까운 K개의 A 와 B 중에서 근처에 더 많은 클래스로 판단하는 것

KNN regression methods: 새로운 데이터가 주어졌을 때 X에 가까운 K개의 데이터 값을 평균내서 값을 예측

3.

Salary = B0+ B1\*gpa + B2\*iq + B3\*gender + B4\*gpa\*iq + B5\*gpa\*gender

Y = 50 + 20\*gpa + 0.07\*iq + 35\*gender + 0.01\*gpa\*iq -10\*gpa\*gender

(a)

E[Y | gender = 1] = 50+20\*gpa+0.07\*iq+35+0.01\*gpa\*iq-10\*gpa = 85+10\*gpa+0.07\*iq+0.01\*gpq\*iq

E[Y | gender = 0] = 50+20\*gpa+0.07\*iq+0.01\*gpa\*iq = 50+20\*gpa+0.07\*iq+0.01\*gpa\*iq

E[Y | gender = 1] - E[Y | gender = 0] = 35-10gpa

3번이 정답, gpa가 3.5보다 크면 남성이 수입이 높다.

(b)

E[Y | gender = 1] = 50+20\*gpa+0.07\*iq+35+0.01\*gpa\*iq-10\*gpa = 85+10\*gpa+0.07\*iq+0.01\*gpa\*iq

85+10\*4+0.07\*110+0.01\*110\*4=137100

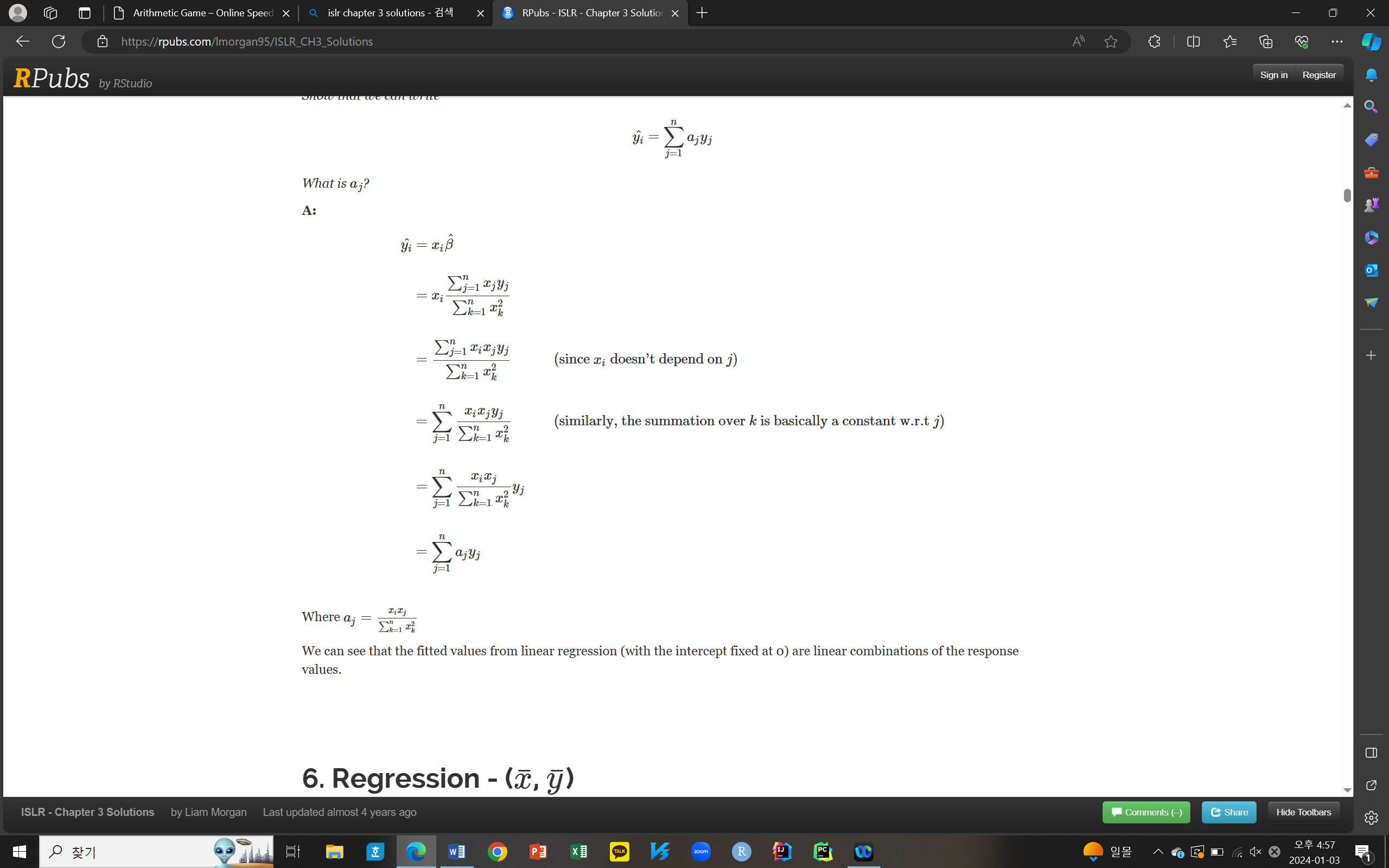
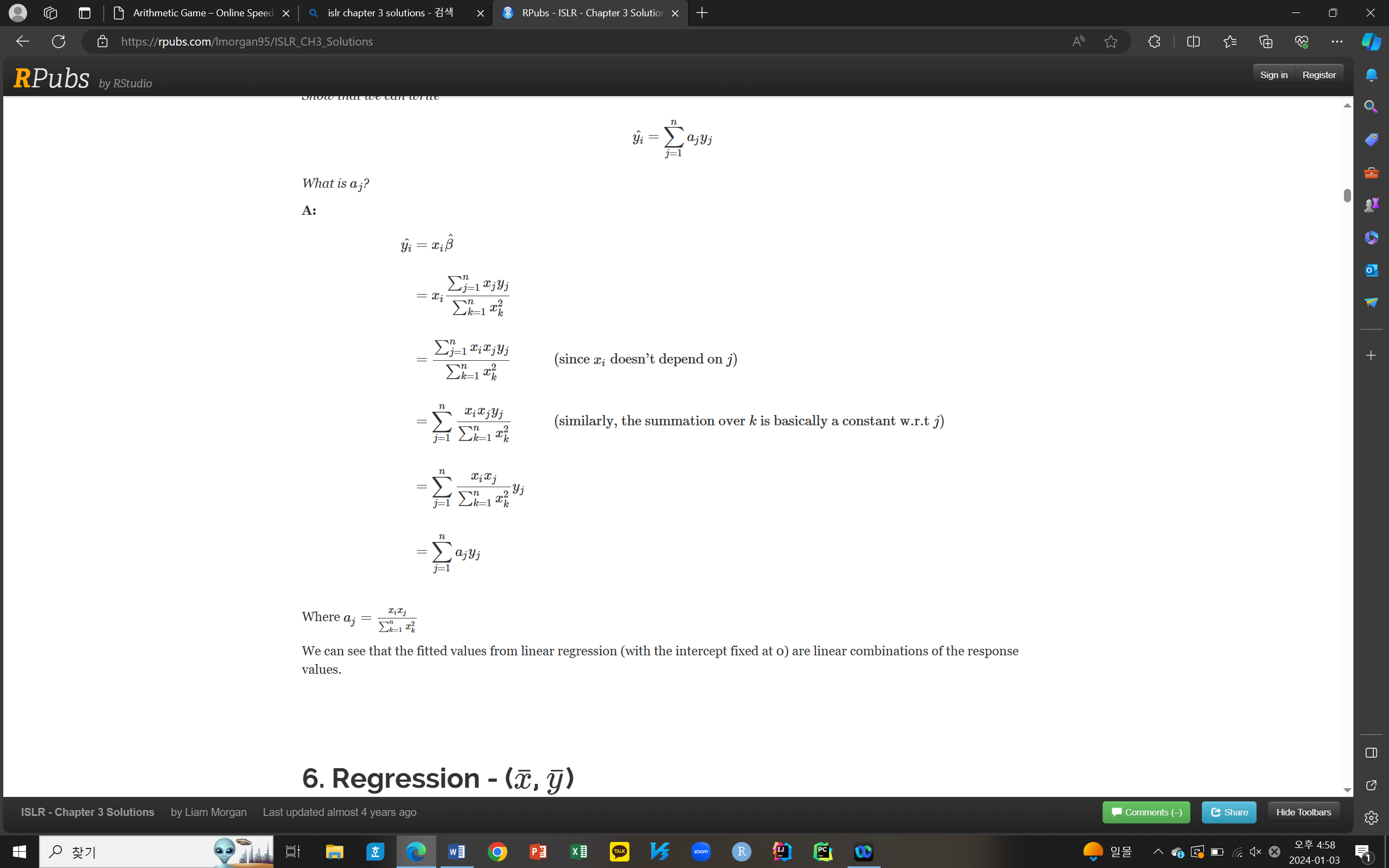
(c)

거짓. Gpa와 iq가 충분히 클 때 이는 유의미할 수 있다.

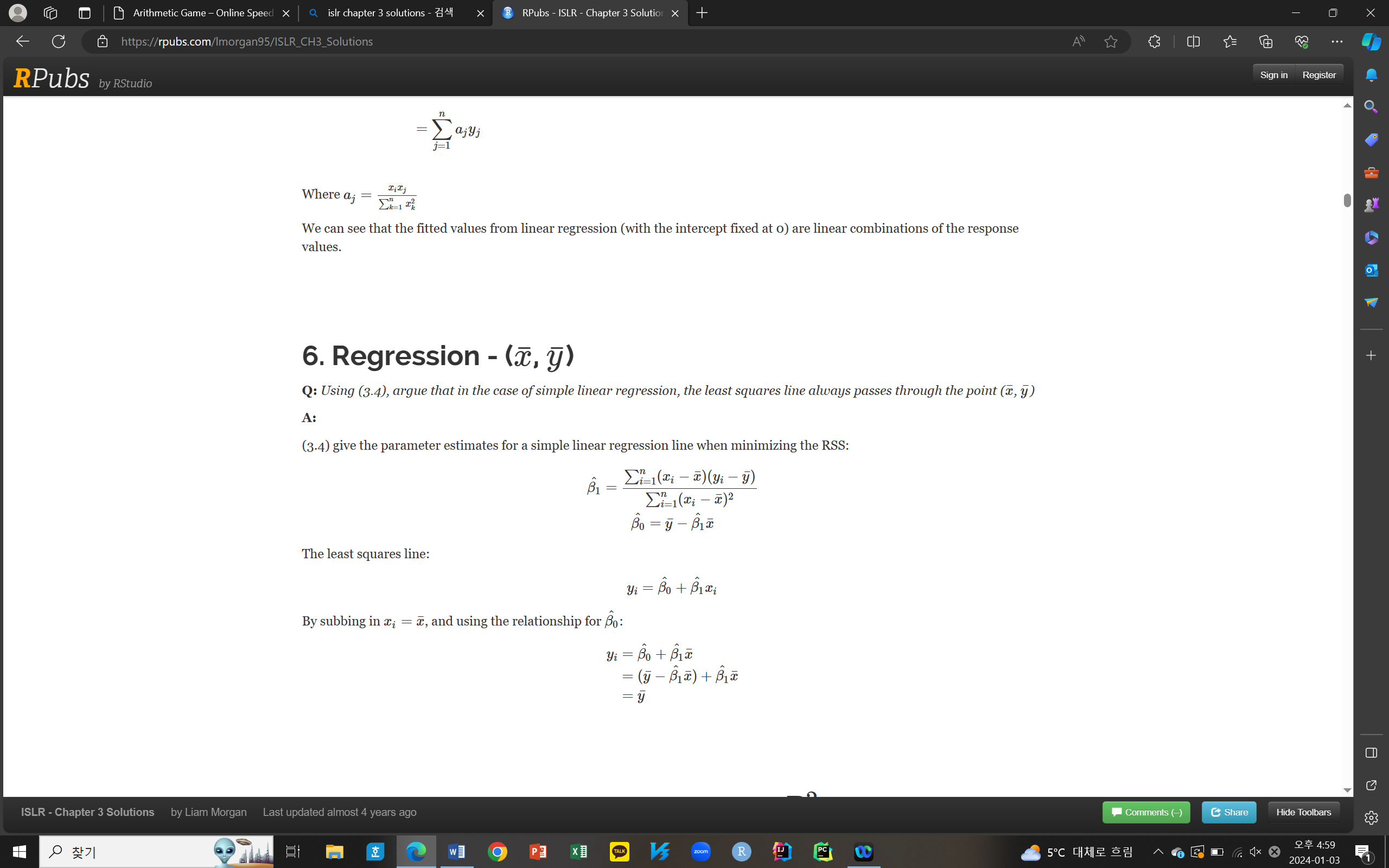
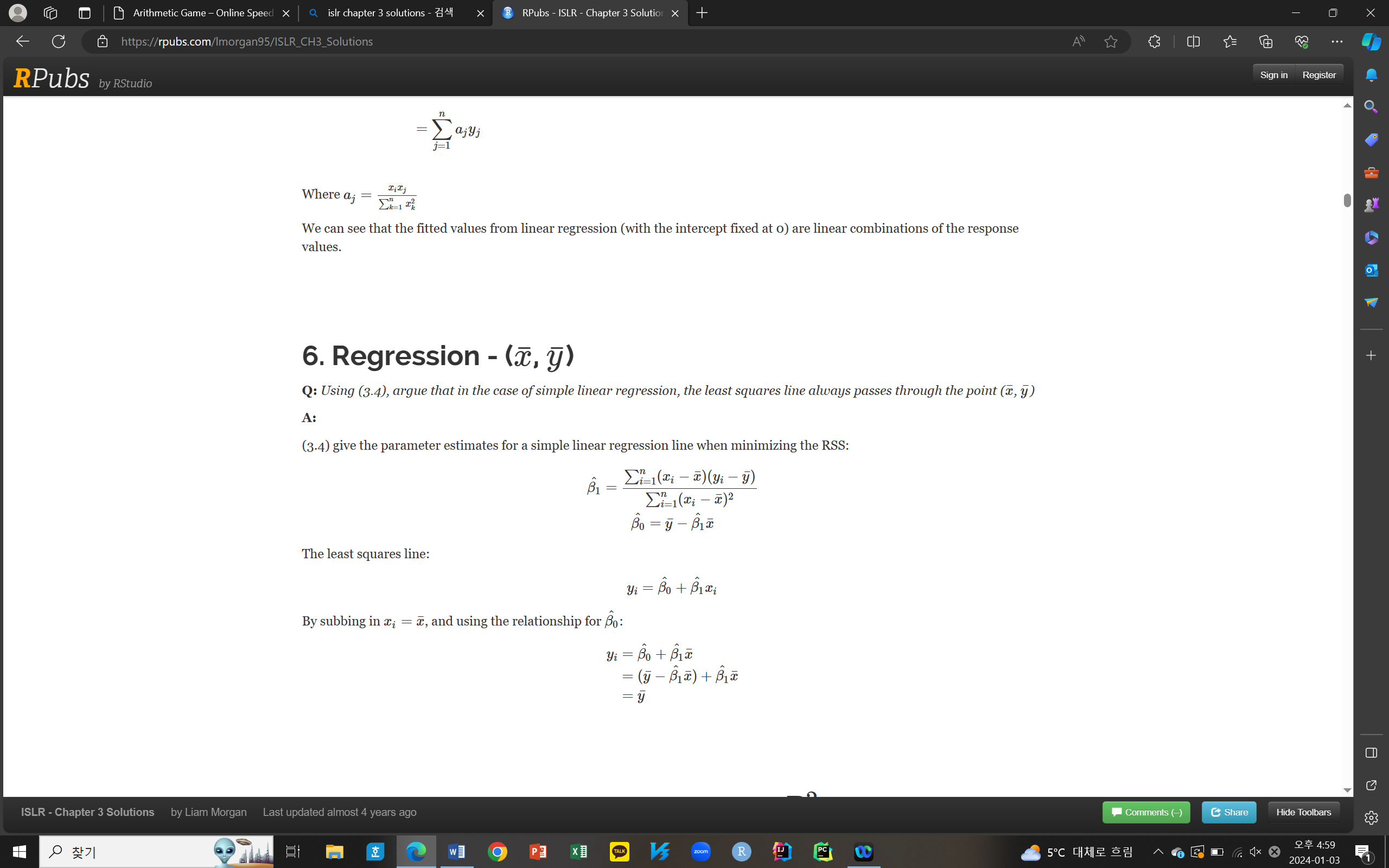
4.

(a)동일할 것으로 예상.

5.

6.



7.

R^2 = (TSS-RSS)/TSS

TSS = sum{(y - bar{y})^2} = sum{y^2}

Cor(X,Y) = sum(xy)/sqrt{sum(y^2)sum(x^2)}

RSS = sum(y^2)\*sum(x^2)-(sum(xy))^2/sum((x)^2)

R^2 = Cor(X,Y)

위 식을 정리하면 결론을 얻을 수 있다.