**[paired\_FM.xlsx] paired t-test 검정 시행결과**

**1. 엑셀**



데이터 요약

평균: 변수 1의 평균은 65,438.2이고, 변수 2의 평균은 60,373.68입니다.

분산: 변수 1의 분산은 444,981,809.5이고, 변수 2의 분산은 469,441,784.6입니다.

관측수: 두 변수 모두 25개의 관측치가 있습니다.

피어슨 상관 계수: 두 변수 사이의 상관계수는 0.952입니다, 이는 매우 높은 양의 상관관계를 나타냅니다.

가설 평균차: 0, 이는 두 집단 간의 차이가 없다는 귀무가설을 설정한 것을 의미합니다.

t-검정 결과

자유도: 24

t 통계량: 3.809688414, 이 값은 t 분포를 사용하여 두 평균 간의 차이의 통계적 유의성을 검정하는 데 사용됩니다.

P(T<=t) 단측 검정: 0.000425543, 이 값은 귀무가설 하에서 표본 데이터가 관측된 t 값 또는 더 극단적인 값을 가질 확률을 의미합니다. 이 확률이 매우 낮기 때문에, 귀무가설을 기각할 수 있는 근거가 됩니다.

t 기각치 단측 검정: 1.71088208, 이는 귀무가설을 기각하기 위한 t 통계량의 최소값을 의미합니다. 관측된 t 값은 이 기각치보다 큽니다.

P(T<=t) 양측 검정: 0.000851086, 이는 변수 1이 변수 2보다 통계적으로 유의미하게 크거나 작을 확률을 의미하며, 이 경우에도 귀무가설을 기각할 수 있습니다.

t 기각치 양측 검정: 2.063898562, 양측 검정에서 사용되는 기각치로, 관측된 t 값은 이 기각치를 초과합니다.

해석

관측된 t 통계량이 양측 및 단측 검정의 기각치보다 높으며, p값이 매우 낮기 때문에 통계적으로 유의미한 차이가 있음을 확인할 수 있습니다. 따라서 변수 1과 변수 2 간에는 통계적으로 유의한 차이가 있다고 결론 지을 수 있습니다. 이 결과는 변수 1이 변수 2보다 평균적으로 높다는 것을 나타냅니다.

**2. R 프로그램**

Paired t-test

data: data$Finance and data$Marketing

t = 3.8097, df = 24, p-value = 0.0008511

alternative hypothesis: true mean difference is not equal to 0

95 percent confidence interval:

2320.816 7808.224

sample estimates:

mean difference

5064.52

검정 결과 요약

t 통계량: 3.8097. 이 값은 두 집단 간의 평균 차이가 귀무가설 하에서 관찰될 확률이 매우 낮음을 나타냅니다. t 통계량이 크면 클수록 귀무가설을 기각할 가능성이 높아집니다.

자유도 (df): 24. 이는 쌍을 이룬 관찰값이 25쌍이 있음을 나타내며, 자유도는 관찰값 수에서 1을 뺀 값입니다.

p-값: 0.0008511. 이 값은 귀무가설 하에서 관찰된 결과나 더 극단적인 결과가 발생할 확률입니다. p-값이 매우 낮기 때문에 (일반적으로 0.05 이하를 유의미하다고 판단) 우리는 귀무가설을 기각할 수 있습니다.

신뢰 구간

95% 신뢰 구간: [2320.816, 7808.224]. 이 구간은 두 집단 간의 차이가 통계적으로 유의미한 범위 내에서 실제로 얼마나 차이가 나는지를 나타냅니다. 이 구간은 0을 포함하지 않기 때문에, 통계적으로 유의미한 차이가 존재한다는 것을 나타냅니다.

평균 차이 추정치

평균 차이: 5064.52. 이는 Finance 전공자들의 연봉이 Marketing 전공자들보다 평균적으로 약 5064.52만큼 높다는 것을 의미합니다.

해석

이 통계 분석에 따르면, Finance 데이터와 Marketing 데이터 간에는 통계적으로 유의미한 차이가 존재합니다. p-값이 매우 낮고 (0.0008511), 95% 신뢰 구간이 0을 포함하지 않으므로, 우리는 Finance 집단이 Marketing 집단보다 통계적으로 유의하게 높은 연봉을 보인다는 결론을 내릴 수 있습니다. 이는 Finance 졸업자들이 Marketing 졸업자들보다 연봉이 더 높음을 시사합니다.

**3. 파이썬**

T-statistic: 3.80968841351656, P-value: 0.0008510859875589974

검정 결과 요약

T-통계량: 3.80968841351656. 이 값은 두 데이터 세트 간 평균 차이가 0이라는 귀무가설 하에서 얻을 수 있는 결과 중 매우 극단적인 값이라는 것을 나타냅니다. T-통계량이 클수록, 즉 더 극단적일수록 귀무가설이 기각될 가능성이 높아집니다.

P-값: 0.0008510859875589974. 이 값은 귀무가설이 참일 때, 관찰된 t-통계량이나 더 극단적인 값을 얻을 확률입니다. 일반적으로 p-값이 0.05 이하일 경우, 귀무가설을 기각하고 대립가설을 수용하는 것이 통상적입니다. 이 경우, p-값이 매우 낮으므로 통계적으로 유의미한 차이가 있다고 할 수 있습니다.

해석

이 결과는 두 집단 간에 통계적으로 유의미한 차이가 있음을 나타냅니다. T-통계량과 매우 낮은 p-값은 데이터 세트 간의 평균 차이가 우연히 발생했을 가능성이 매우 낮다는 것을 의미합니다. 즉, 실험 데이터는 귀무가설을 기각하고, 두 집단 간의 차이가 실제로 존재함을 지지합니다.