

PEOPLE ANALYTICS

HR 현업이 데이터 분석 활용의 주체가 되는 법

HEARTCOUNT | 양승준 (sidney.yang@idk2.co.kr)



소비하는 사람

- 구매력 있는 사람이
- 더 많이
- 자주

People



일하는 사람

- 재능 있는 사람이
- 열심히 (잘)
- 사이 좋게

"People" Analytics: Two Types of People



- 양적 차이
- 찾고 자잘한
- 상관관계

무슨 차이?

KPI
의사결정
패턴



- 질적 차이
- 드물고 주요한
- 인과관계

People “Analytics”: What is Analytics?

Statistics

Small Data → Statistical Significance
가설 검증 · 일반화 · 인과관계
정책 결정 · 신약 효과

AI · ML

Big Data → Decision Automation
Performance · Prediction(Who/What)
자잘하고 빈번한 의사결정(개인화 · 추천)

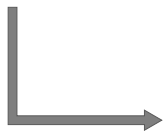
Analytics

Relevant Data → **Decision Intelligence**
Fact(What) · Explanation(Why/How)
기업내 중요하고 집단적인 의사결정

People Analytics: People 데이터의 쓸모를 궁리하는 일

데이터가 답할 수 있는 좋은 질문

Frame Real-World Problem into Data Problem



*뻔하지 않고 쓸모 있는 패턴 발견 · 공유
Non-Obvious & Useful Pattern

기술/도구가
해줄 수 있는 것



분석결과 수용 · 활용
Audience Acceptance & Intervention

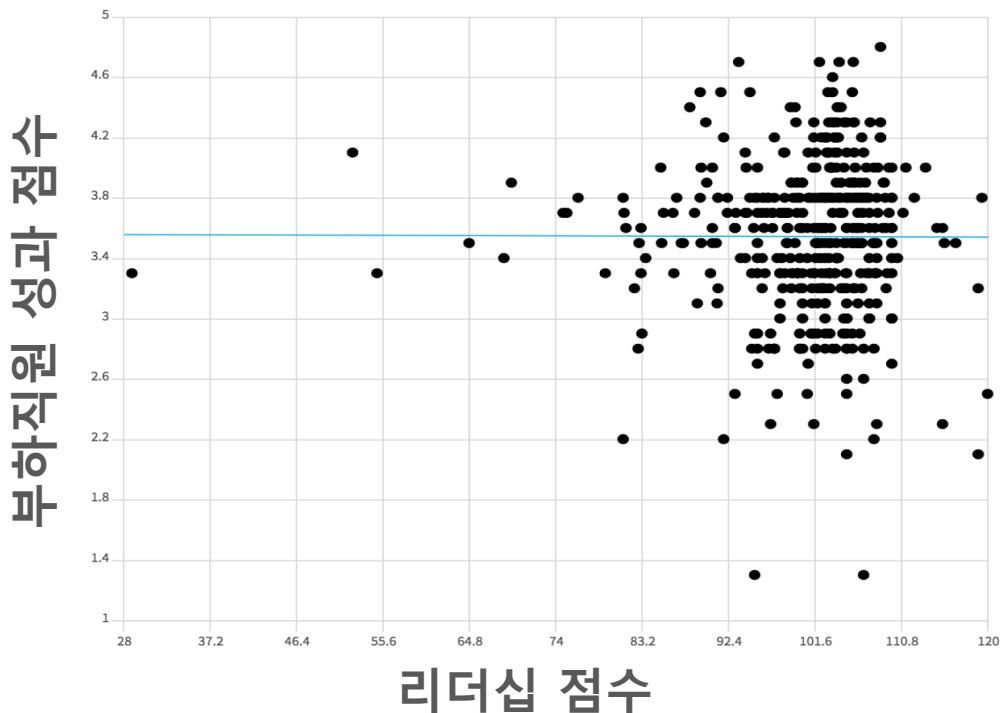
매니저 리더십 점수와 부하직원 성과 점수
X와 Y 사이에 관계(규칙·질서)가 있을까?

X_1 리더십.역량1	X_2 리더십.역량3	X_2 리더십.역량3	X_2 리더십.역량4	Y 부하.성과점수
85	55	99	65	4.5
65	76	46	92	3.5
92	77	92	56	3.2
56	36	56	46	3.5
76	54	76	92	4.8
...	...	...	...	...

리더십.점수와 성과.점수 간 상관관계: 빵

Y: 성과3년평균 X: 리더십_환경_리더십 (하위그룹: 없음)

상관계수: 0

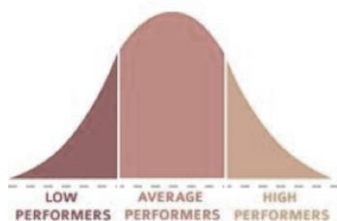


데이터에 답이 없는 질문

- **Data is Never Raw:** 기록·측정·선택·처리 과정에 사람이 개입
- **Provenance:** 데이터의 기원(수집 맥락) 이해

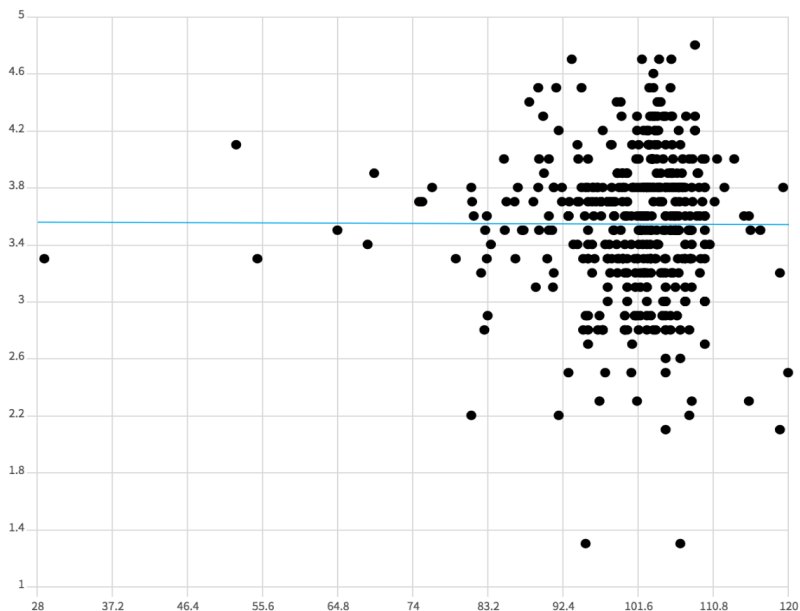
Y: 성과3년평균 X: 리더십_환경_리더십 (하위그룹: 없음)

상관계수: 0



리더십과 무관하게
모든 리더 부하직원의
성과점수는 고르게 분포

성과
점수
분포



리더십 점수

데이터에 답이 있는데도 못 찾는 경우

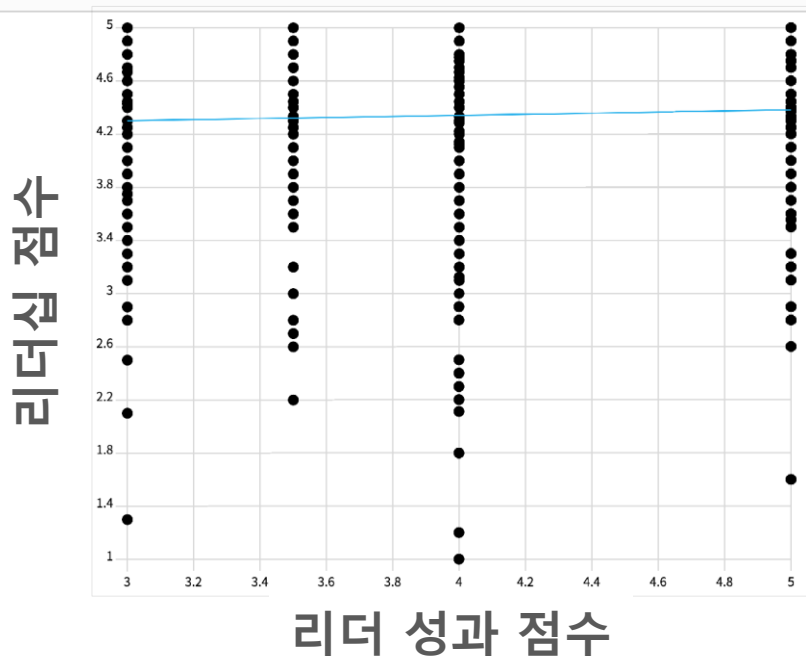
매니저 리더십 점수와 매니저 성과 점수
X와 Y 사이에 관계(규칙·질서)가 있을까?

X_1 리더십.역량1	X_2 리더십.역량3	X_2 리더십.역량3	X_2 리더십.역량4	Y 본인.성과점수
4	3	2	2	4.5
3	3	5	4	3.5
2	2	3	3	3.2
5	4	3	2	3.5
3	3	3	5	4.8
...	...	...	...	...

데이터에 답이 있는데도 못 찾는 경우

매니저 리더십 점수와 매니저 성과 점수
X와 Y 사이에 관계(규칙·질서)가 있을까?

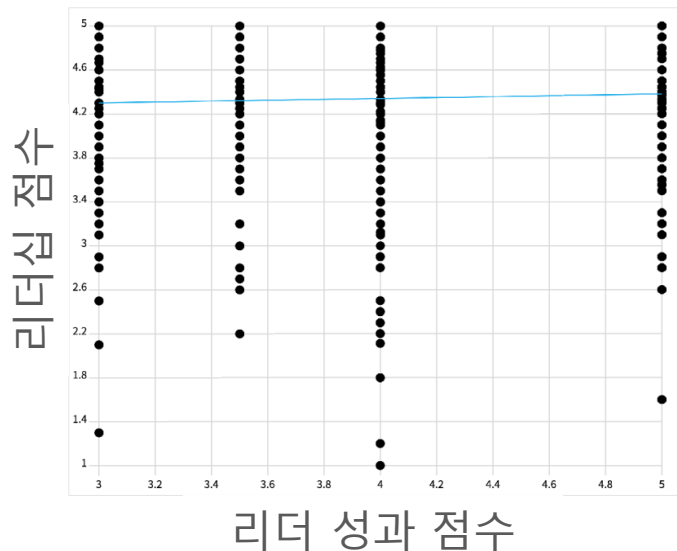
X와 Y 간 상관관계(0.06)가 매우 약함



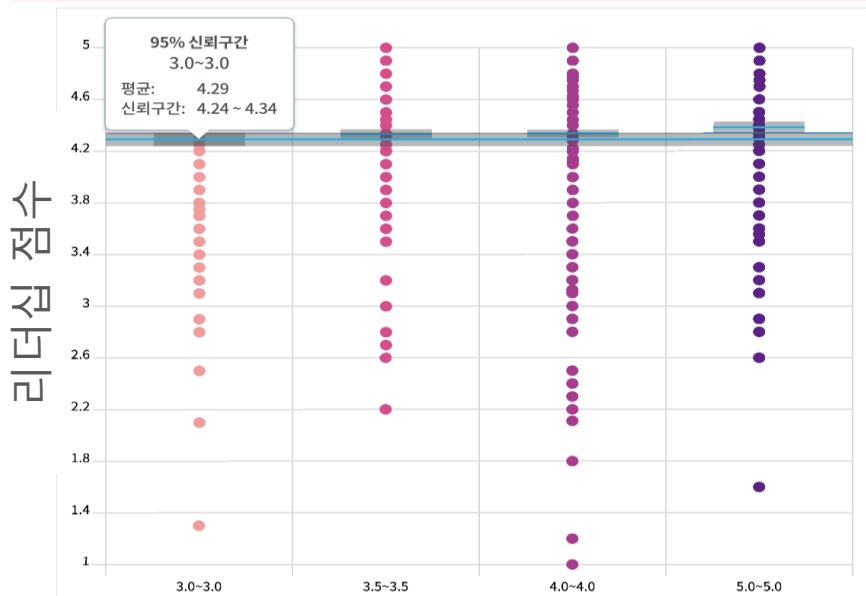
데이터에 답이 있는데도 못 찾는 경우 - Data Literacy!

성과 점수를 범주(등급)로 처리 → 집단 간 유의미한 점수 차이

X를 숫자로 처리
상관관계(0.06)가 매우 약함



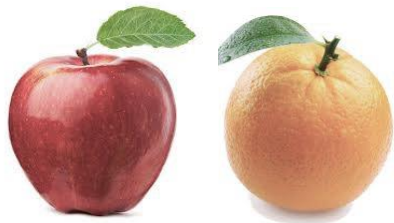
X를 범주로 처리
C(3점)와 S(5점) 간 유의미한 점수 차이



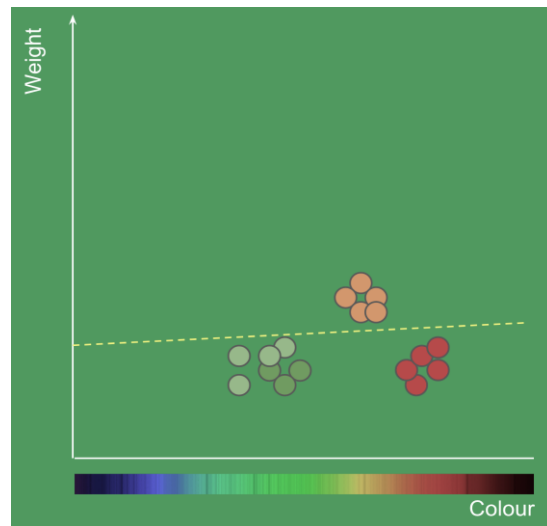
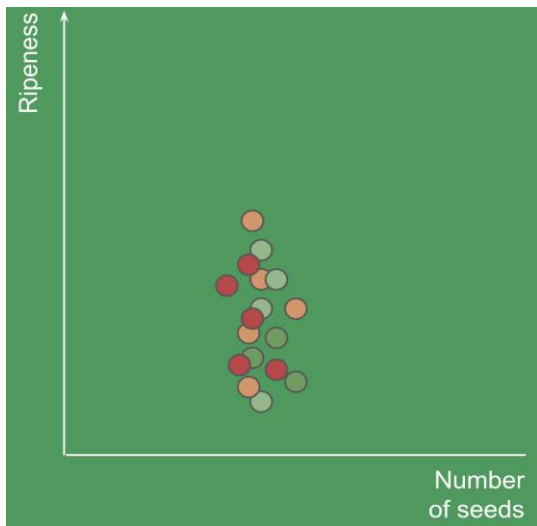
리더 성과 등급

질문과 관련성이 높은 데이터(Features)

Feature: Y(궁금한 것)를 설명하거나 분류(예측)하는데 유용한 속성

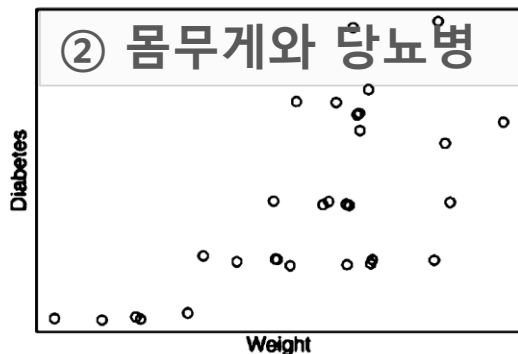
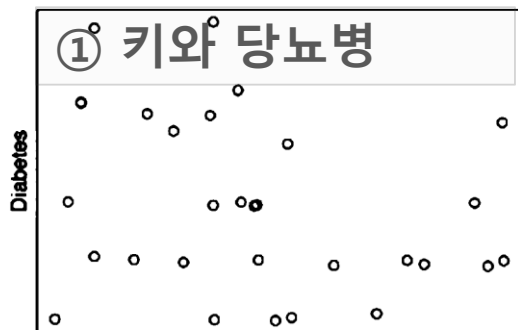


	X_1 Ripeness	X_2 # of Seeds	X_3 Weight (g)	X_4 Color	Y Fruit Type
	0.56	5	320	Orange	Orange
	0.61	6	280	Red	Apple

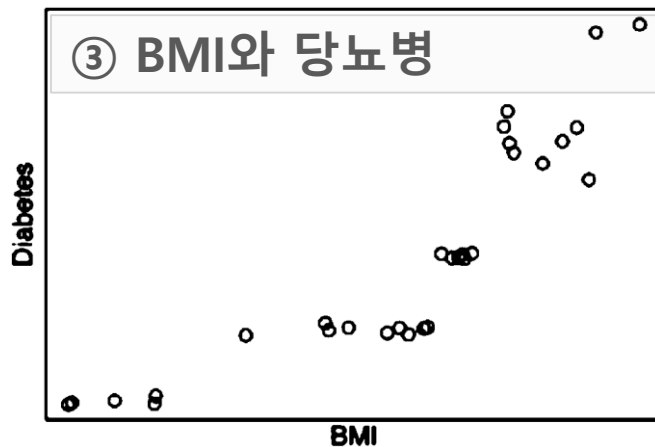


Feature Engineering

Y를 더 잘 설명하기 위해 기존 변수를 창의적으로 가공하는 일



- 비만도 더 잘 설명할 수 있는 새로운 변수 가공
- *BMI(Body Mass Index) = kg/m^2



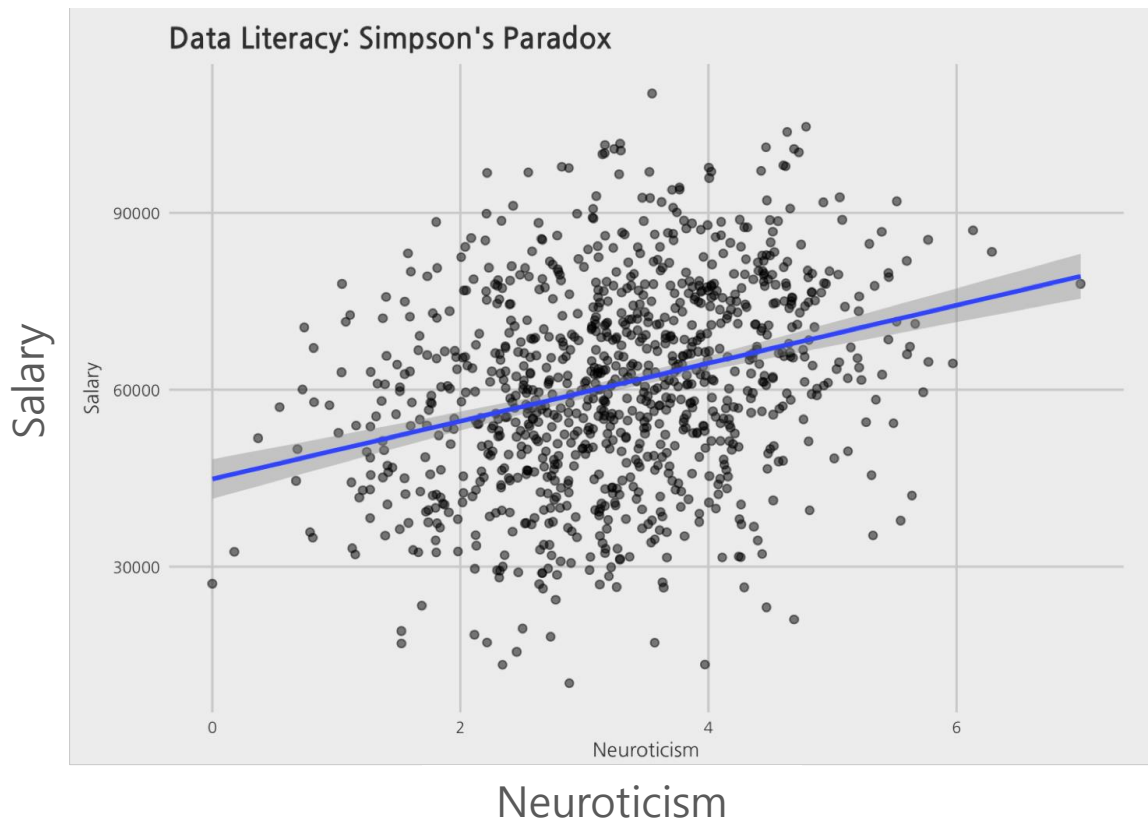
*발명한 사람의 이름을 따서
Quetelet Index라고도 함



Feature Engineering for People Data

Feature Engineering Examples	설명
Number of jobs per unit of tenure	보직변경횟수 / 재직기간
Difference between employee and manager age	부서장과의 나이 차이
Years with current manager	현 부서장과 함께 근무한 기간
Percentage change between last two performance appraisal scores	이전 성과 점수와의 차이
Relative Comparable Salary Growth Ratio	동직급 전체평균 대비 연봉 상승률
Absolute and Relative Time Since Last Promotion	마지막 승진 이후 절대적/상대적 경과 기간

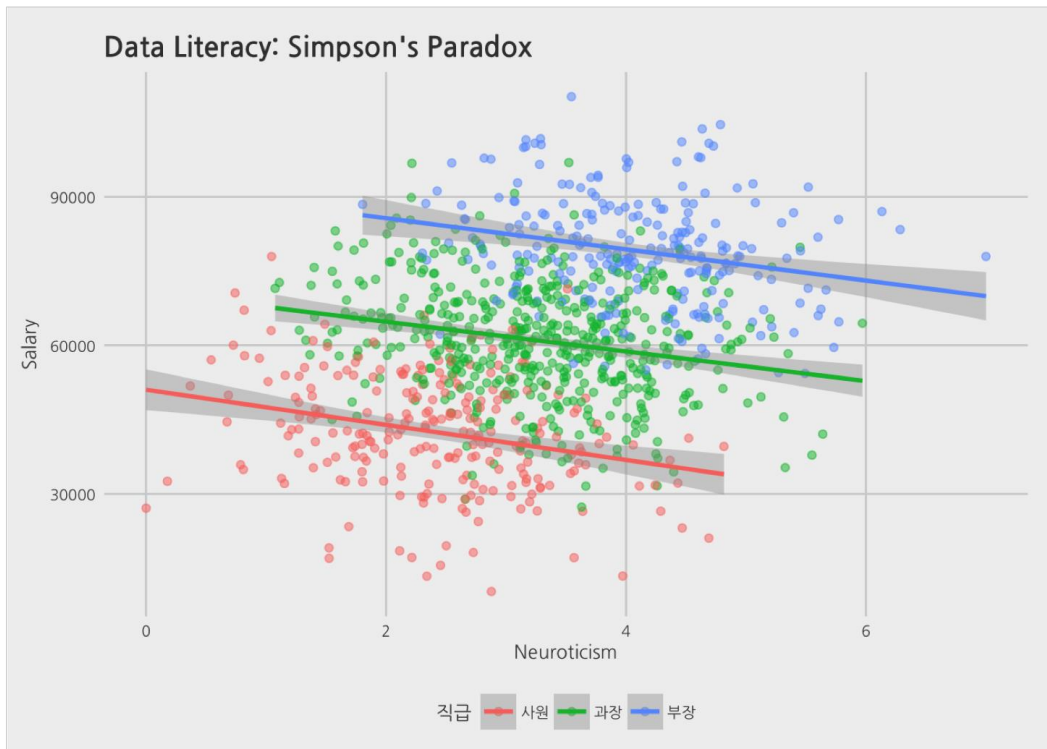
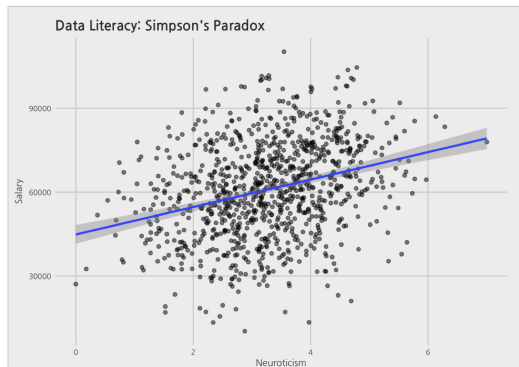
연봉과 까칠함과의 관계



뻔하지 않은 결과 - 제한된·익숙한 관점 탈피

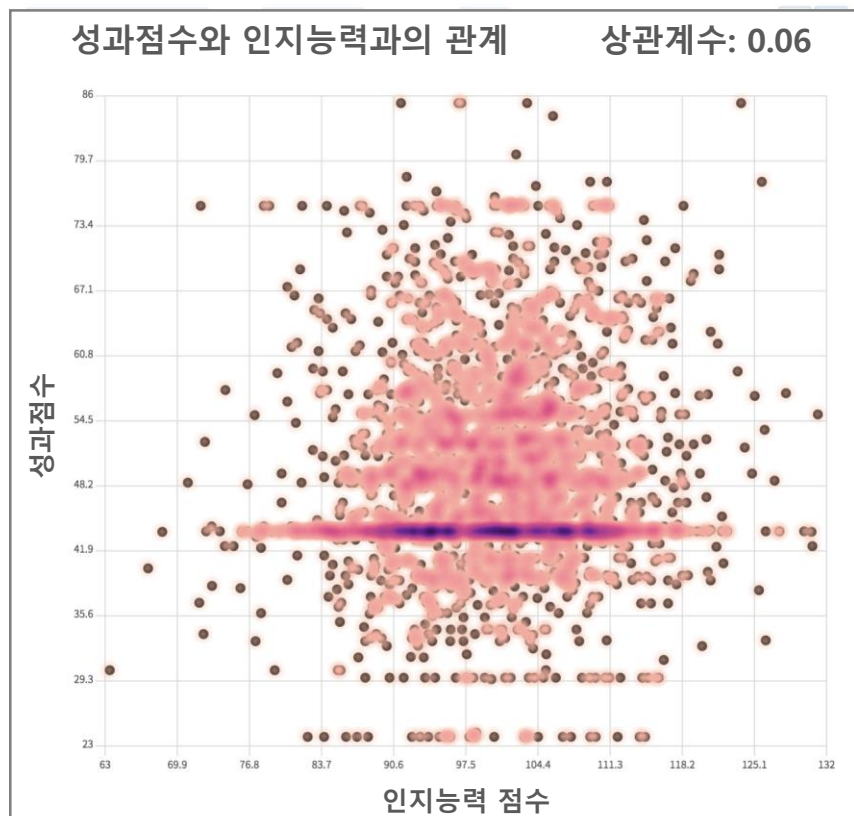
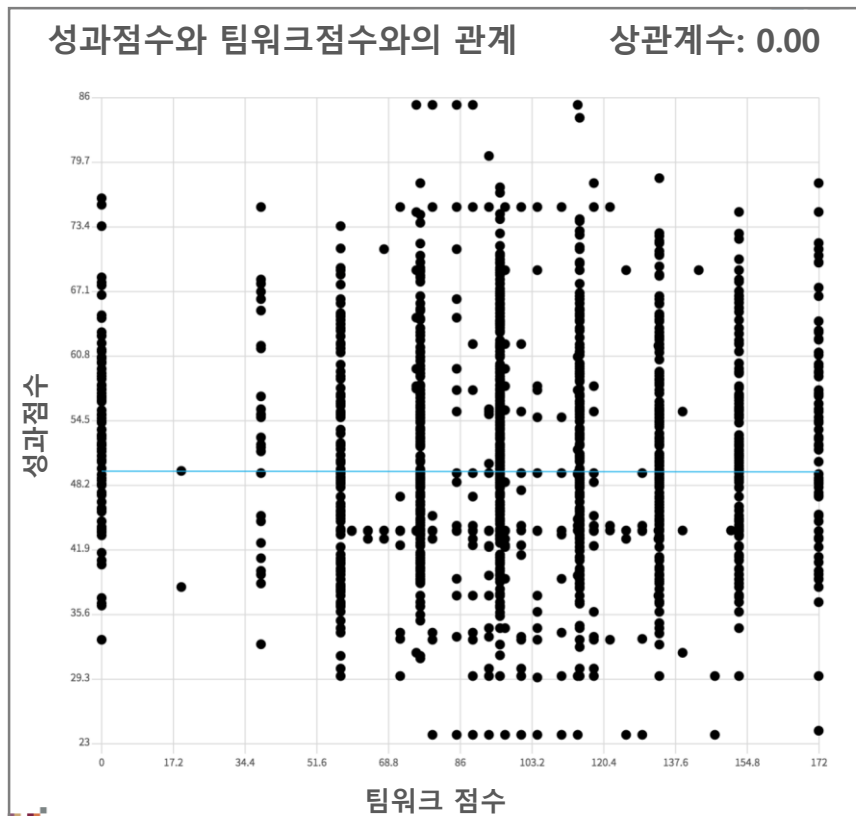
새로운 관점(Dimension) 추가

연봉과 까칠함과의 관계 → 직급별 연봉과 까칠함과의 관계



쓸모 있는 결과

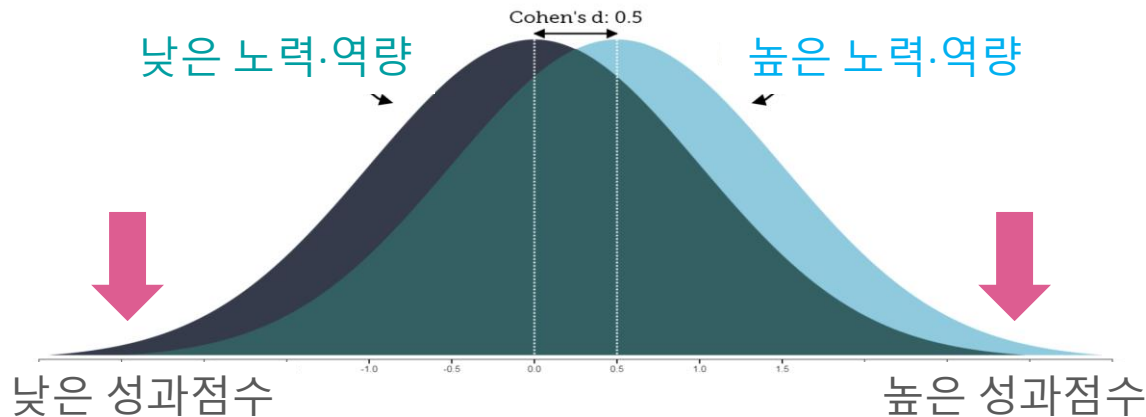
신입사원 성과점수와 인적성점수 간 선형적 관계 없음



쓸모 있는 결과

Performance에서 Noise와 Signal을 어떻게 구분할까?

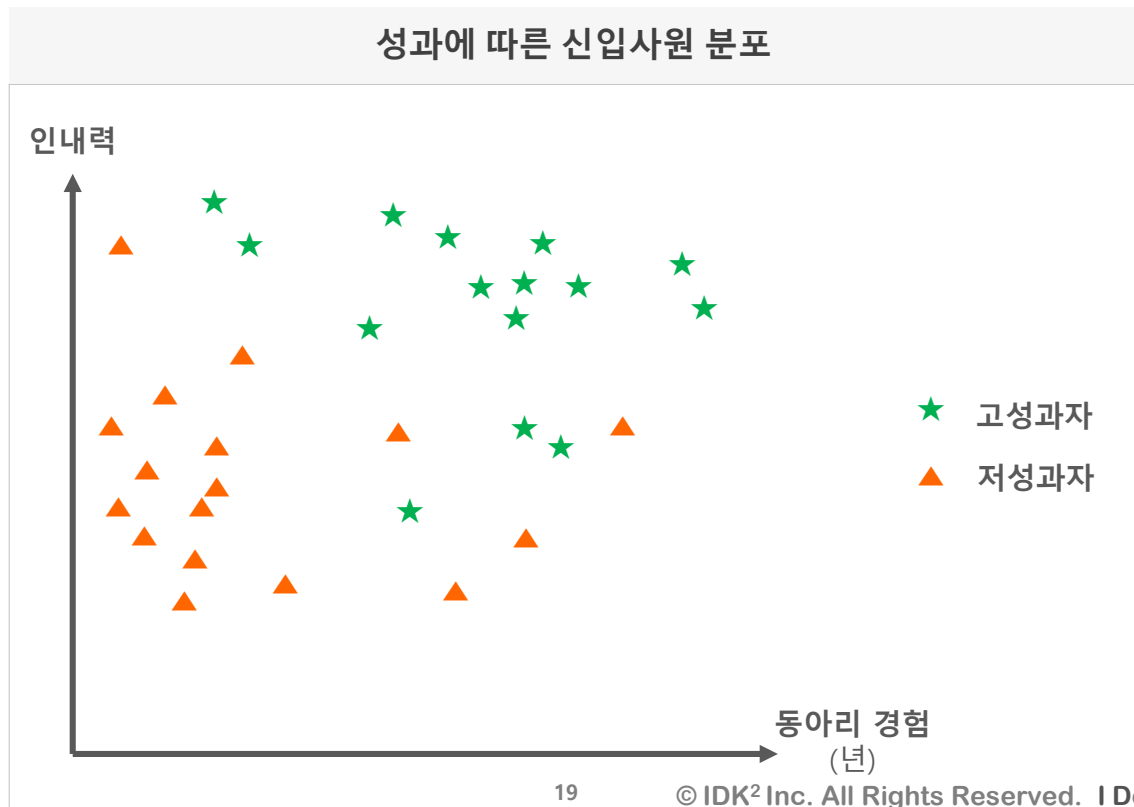
운의 영향으로 높은 노력·역량에도 불구하고, 낮은 성과 점수를 받거나
낮은 노력·역량에도 불구하고, 높은 성과 점수를 받을 수 있음



Compare The Extremes: 극단의 두 집단 비교, 운의 영향 최소화

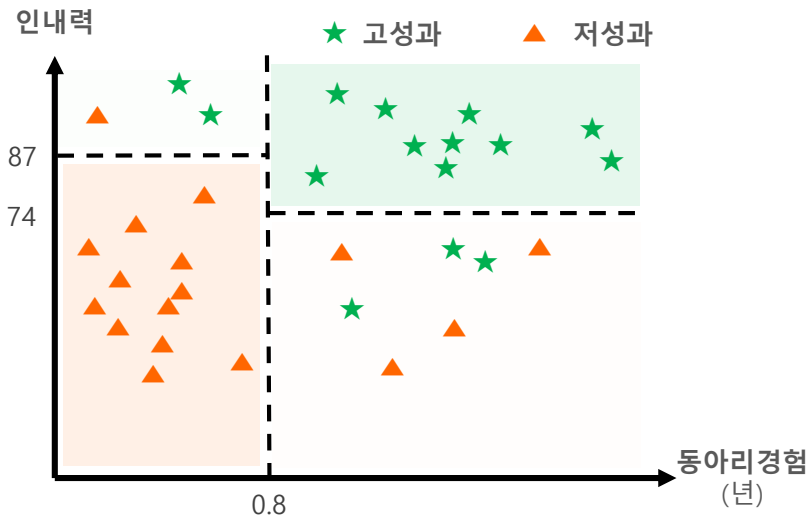
Decision Tree: (양 극단의) 두 집단을 분류하는 규칙을 찾는 방법

- **Purity:** *엔트로피를 최소화하도록(= 끼리끼리 모이도록) 공간 구획
- **Homogeneity:** 동질적 집단이 밀집한 세그먼트의 논리적 규칙 찾기

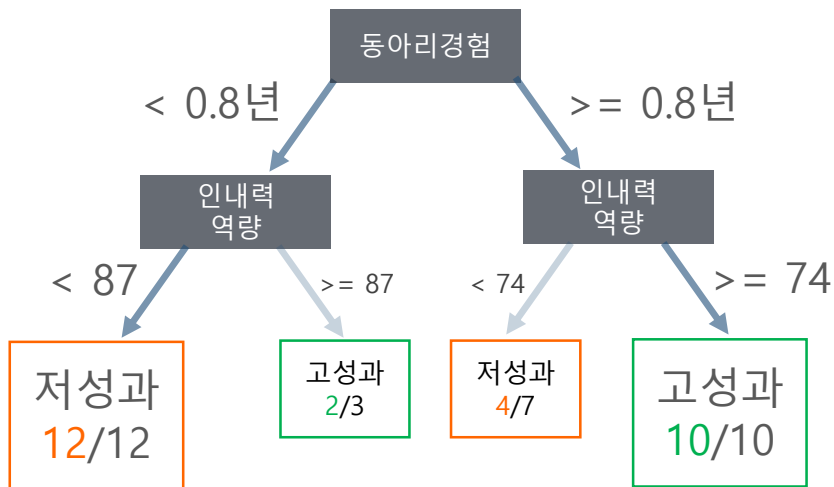


Decision Tree: (양 극단의) 두 집단을 분류하는 규칙을 찾는 방법

성과에 따른 신입사원 분포



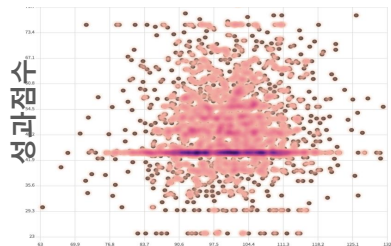
고성과 신입사원 분류모형



분류 규칙	확률
(동아리경험 < 0.8년) and (인내력역량 < 87) = 저성과	100% (12/12)
(동아리경험 >= 0.8년) and (인내력역량 >= 74) = 고성과	100% (10/10)

쓸모 있는 결과 - 의사결정과 행동에 활용할 수 있는 패턴

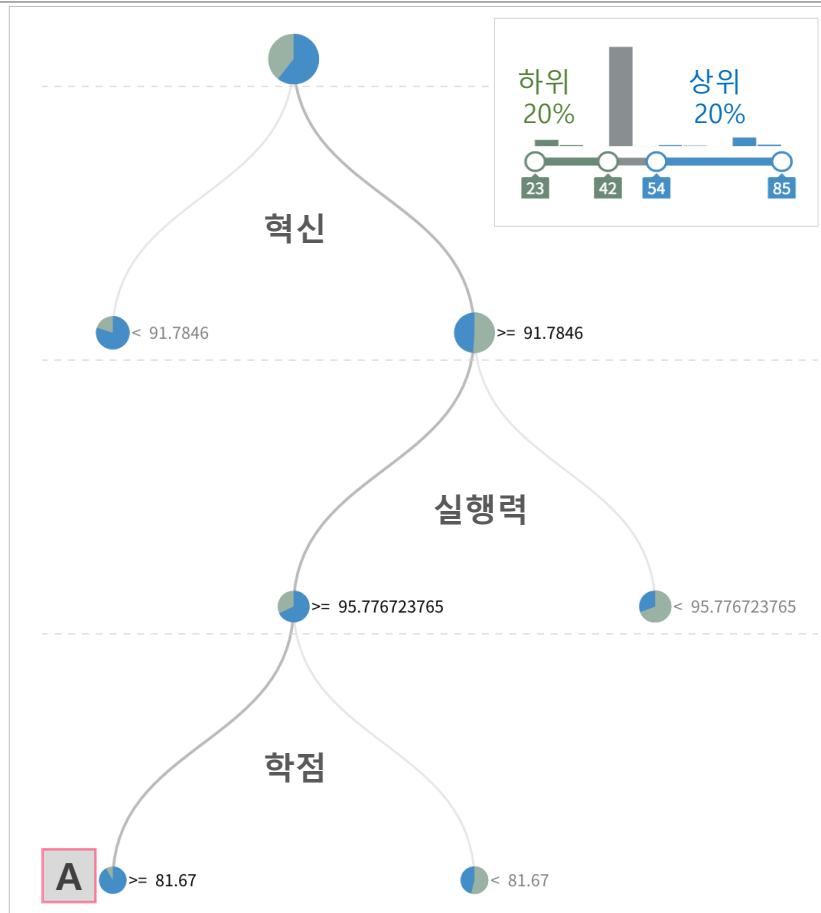
상관계수: 0.06



성과점수 양극단
(상위20% vs. 하위20%)
집단을 비교한 결과

A 혁신, 실행력, 학점 동시에 높으면:

- 고성과자 비율: 92%가 됨
- 고성과 분류 규칙: [혁신 ≥ 91.7]
& [실행력 ≥ 95.7] & [학점 ≥ 92]



People Analytics이 실용적인 일이 되려면

- 의견(믿음)을 바꾸려 하기보다 의사결정과 행동에 활용할 수 있는 통제 가능한 패턴 찾기
- 통제 가능한 패턴 → 개인보다 일하는 환경에 집중
- 측정할 수 없으면 묻자
- 공부에 순서 없음; 그때 그때 Google에 물어보면서 배우기