

# 조직 내 성과 분석과 향후 전략 제안

HR 데이터를 토대로 한 고성과자 및 저성과자 특징 조사

## 프로젝트 개요

**우리 회사의 우수 성과자들은 어떤 특징을 가지고 있을까 ?**

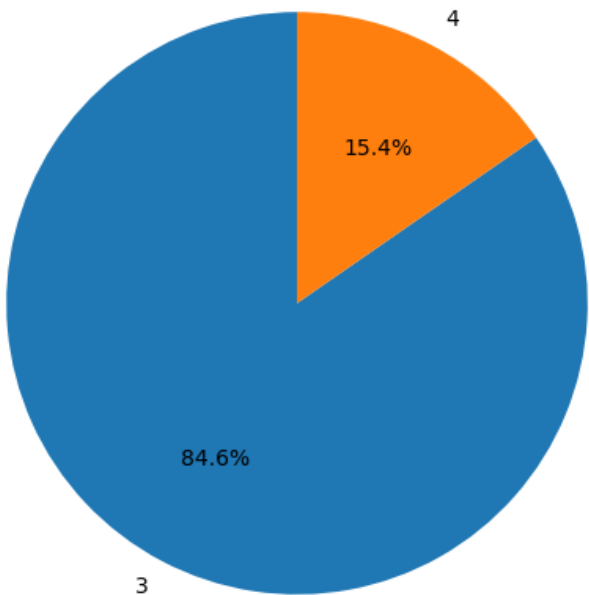
# 프로젝트 개요

## 프로젝트 목표

### 업무 성과 분석을 통한 성과 향상에 대한 솔루션 제공

#### HR Dataset

Performance Rating Distribution

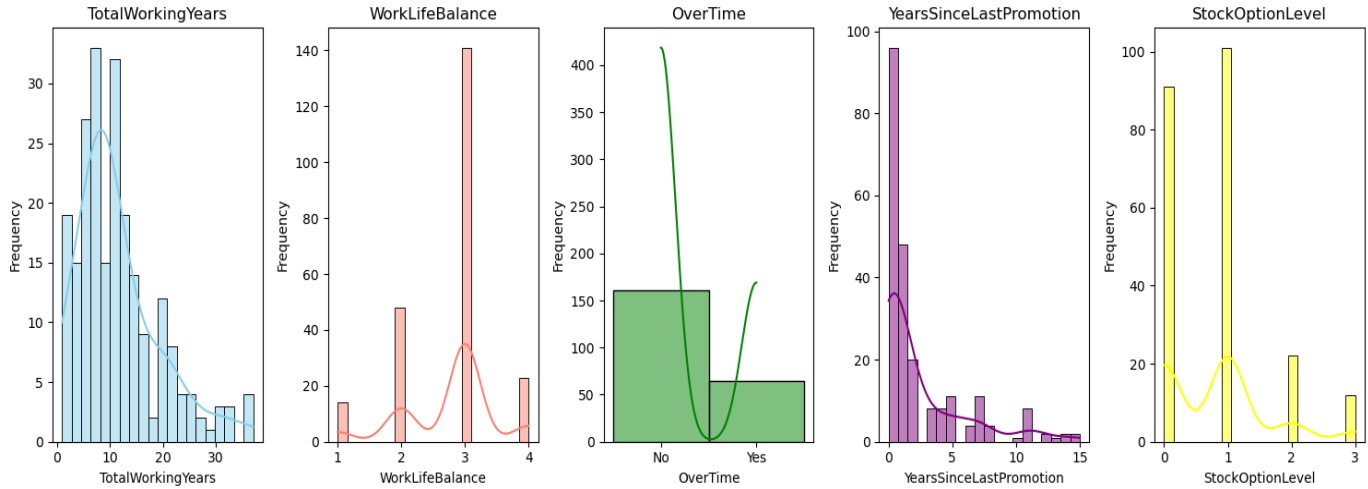


1470명의 직원 데이터 중  
업무 성과(PerformanceRating)분포

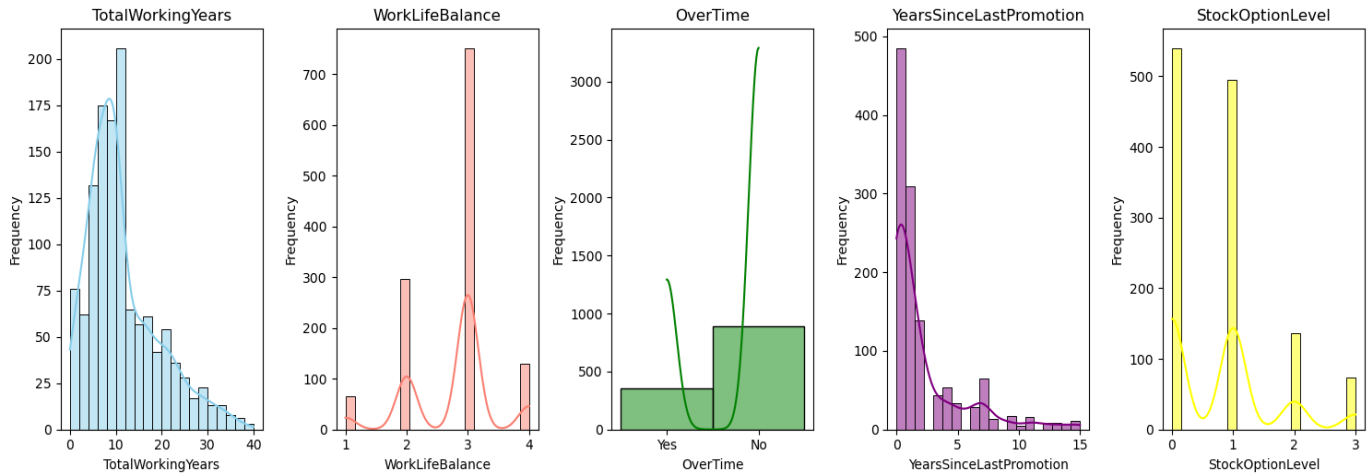
3(Excellent)	84.6%
4(Outstanding)	15.45%
2(Good)	0%
1(Low)	0%

# 단순 비교 : 주요 요인별

PerformanceRating 4

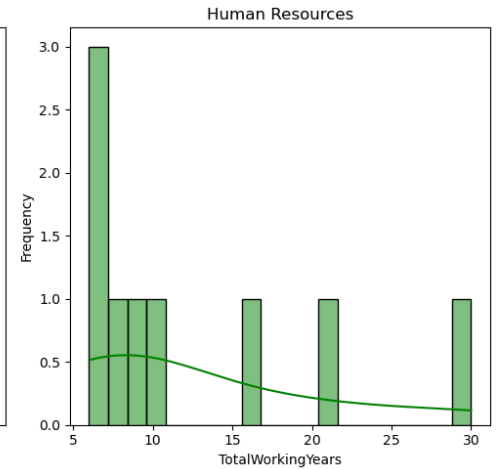
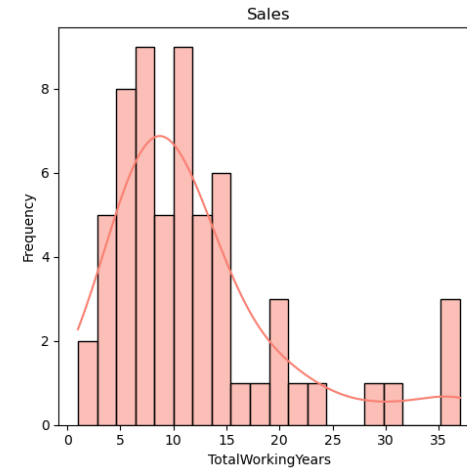
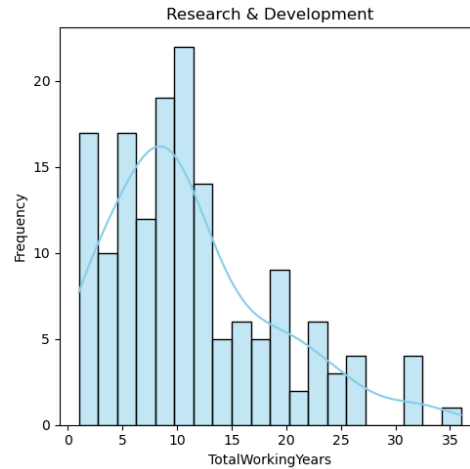


PerformanceRating 3

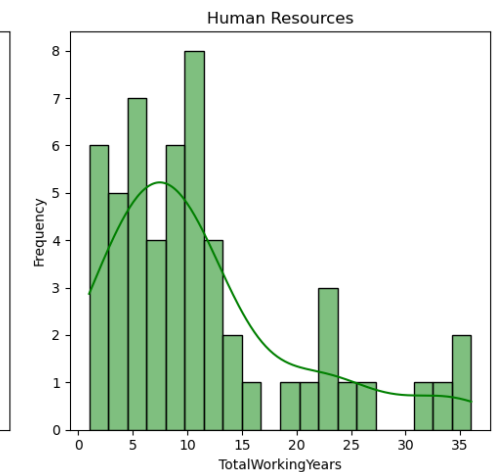
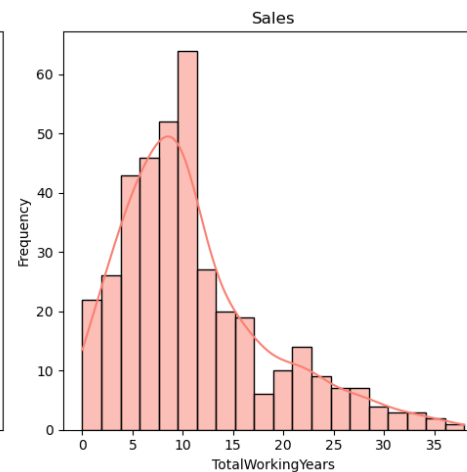
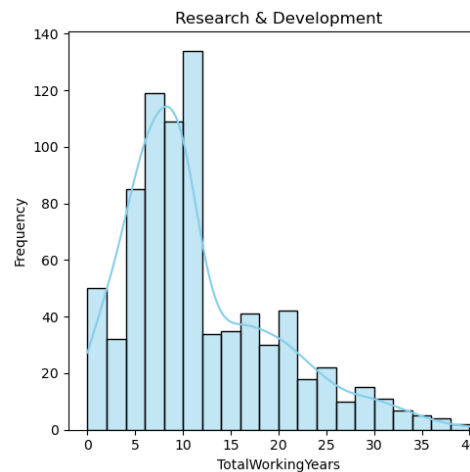


# 단순 비교 : Department(부서)별

PerformanceRating 4



PerformanceRating 3



**업무 성과별 단순 비교 시, 눈에 띄는 차이를 찾아내지 못함**

**업무 성과 점수가 1,2인 데이터의 부재로 인한  
데이터의 신뢰성 부족**

## 문제 정의



지표에 나오지 않는 "진짜 " 저성과자, 고성과자를 나누어 볼 수 없을까 ?

또한, 이들의 특징에 대해 알아볼 수 있을까 ?

1. 저성과자, 고성과자를 재정의 후 지표적 차이 분석
2. 주요 요인에 대한 상관 분석

# 데이터 전처리

## 가용 데이터 :HR DataSet

	Age	Attrition	BusinessTravel	DailyRate	Department	DistanceFromHome	Education	EducationField	EmployeeCount	EmployeeNumber	EnvironmentSatisfaction	Gender	HourlyRate	JobInvolvement	JobLevel	JobRole	JobSatisfaction	M
0	41	Yes	Travel_Rarely	1102	Sales	1	2	Life Sciences	1	1	2	Female	94	3	2	Sales Executive	4	
1	49	No	Travel_Frequently	279	Research & Development	8	1	Life Sciences	1	2	3	Male	61	2	2	Research Scientist	2	
2	37	Yes	Travel_Rarely	1373	Research & Development	2	2	Other	1	4	4	Male	92	2	1	Laboratory Technician	3	
3	33	No	Travel_Frequently	1392	Research & Development	3	4	Life Sciences	1	5	4	Female	56	3	1	Research Scientist	3	
4	27	No	Travel_Rarely	591	Research & Development	2	1	Medical	1	7	1	Male	40	3	1	Laboratory Technician	2	
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
1465	36	No	Travel_Frequently	884	Research & Development	23	2	Medical	1	2061	3	Male	41	4	2	Laboratory Technician	4	
1466	39	No	Travel_Rarely	613	Research & Development	6	1	Medical	1	2062	4	Male	42	2	3	Healthcare Representative	1	
1467	27	No	Travel_Rarely	155	Research & Development	4	3	Life Sciences	1	2064	2	Male	87	4	2	Manufacturing Director	2	
1468	49	No	Travel_Frequently	1023	Sales	2	3	Medical	1	2065	4	Male	63	2	2	Sales Executive	2	
1469	34	No	Travel_Rarely	628	Research & Development	8	3	Medical	1	2068	2	Male	82	4	2	Laboratory Technician	3	

1470 rows × 35 columns

직원의 특징에 대한 34개의 독립변수, 1개의 종속변수(Attrition)를 가지고 있는 Dataset



# 데이터 전처리

## 저성과자의 특성

〈표 1〉 성과자 유형별 분석을 통한 저성과자의 특성

	고성과자 (high performer)	중간성과자 (adequate)	저성과자 (low performer)
전반적인 능력 수준	- 연봉 수준에서 상위 10%에 드는 자들	- 연봉 수준에서 65%~89% 백분위 수준에 해당	- 연봉 수준에서 65%이하에 해당
문제 해결	- 'quick study' - 복잡한 문제를 신속 해결	- 영리함(smart) - 평균적인 통찰력	- 복잡한 상황을 다루는데 어려움을 겪음
리더십	- 필요한 변화 실행 - 적응력이 뛰어나고 조직을 고무시킬 능력 있음	- 중도적인 변화 선호 - 부하들의 미온적 태도	- 현상유지 - 진실성 부족으로 사람들이 따르지 않음
열정	- 매우 높은 열정 수준 - 주당 55시간 이상 근무 - 성공에 사로잡힘	- 동기 부여 - 때때로 에너지가 넘침 - 주당 50-54시간 근무	- 헌신적 - 모순되는 pace 보임 - 주당 40-49시간 근무
자원의 풍부함	- 장애물을 헤쳐 나가는 방법을 찾는데서 뛰어난 능력을 지님 - can-do의 태도	- 때때로 해결방법 구함 - 다른 이들이 해결할 때까지 기다림	- 방해에 부딪힘 - boss에게 계속해서 책임을 떠넘김
Topgrading	- 고성과자와 잠재력 있는 직원 을 선택함 - 만성적인 저성과자들을 재배 치함	- 몇몇의 고성과자를 선택함	- 대부분 저성과자를 선택함 - 평범한 사람을 목인함
코칭	- 각각의 팀 구성원이 뛰어난 성과를 내고 경력을 쌓을 수 있도록 성공적으로 상담해줌	- 해마다 성과 검토를 수행함 - 부분적으로 코칭이 이루어짐	- 접근하기 어렵고, 혹평하고, 칭찬에 인색함 - 피드백이 늦거나 제대로 이루어지지 않음
팀 구성	- 목표에 집중하고, 협동적이고, 결과지향적인 팀을 구성함	- 팀워크를 원할지 모르나, 조성하지는 않음	- 타인의 에너지를 고갈시킴 - 시너지를 막는 행동들
진로 기록	- 꾸준히 뛰어난 성과를 냄	- 일부 핵심 지지자들의 기 대를 만족시킴	- 우발적으로 기대를 충족시킴
성실성	- 빈틈없이 완벽함	- 대개 정직함	- 규칙을 어김
의사소통	- 말하고 쓰는 기술이 매우 뛰 어남	- 말하고 쓰는 기술이 보통	- 평범함

출처: Smart(2010)

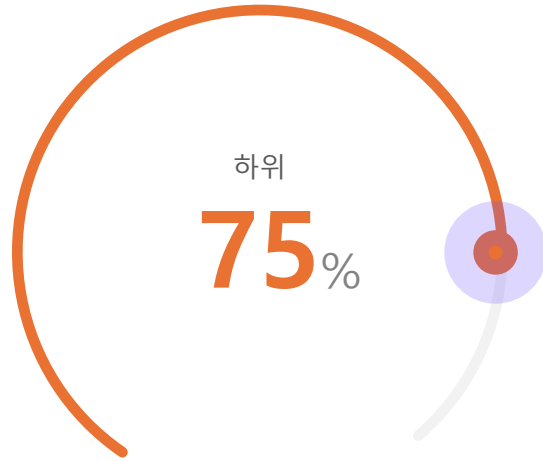
문제 해결 능력이 낮음

대인관계 능력이 부족함

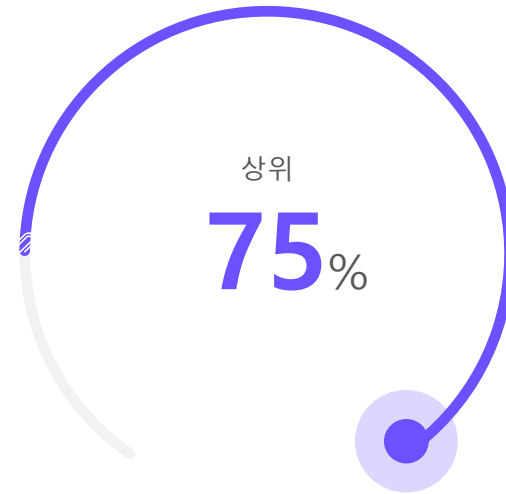
업무 만족도, 참여도가 낮음

# 데이터 전처리

## 데이터 재정의



**Low Performer**



**High Performer**

PerformanceRating이 30이며, RelationshipSatisfaction, EnvironmentSatisfaction, JobInvolvement 이 하위 75%인 직원들 = **Real Low Performer**,  
PerformanceRating이 40이며, RelationshipSatisfaction, EnvironmentSatisfaction, JobInvolvement 이 상위 75%인 직원들 = **Real High Performer**로 정의.

# 데이터 전처리

## LabelEncoding, One-HotEncoding

결혼 여부, 전공, 성별 등 범주화 되어있는 변수 가공 필요

순서가 중요한 범주변수는 LabelEncoding, 순서가 중요하지 않은 범주변수는 OneHotEncoding처리

LabelEncoding = Education, JobLevel

OneHotEncoding = Attrition, Department, BusinessTravel, EducationField, Gender, JobRole, MaritalStatus, Over18

	Age	DailyRate	DistanceFromHome	Education	EmployeeCount	EmployeeNumber	EnvironmentSatisfaction	HourlyRate	JobInvolvement	JobLevel	JobSatisfaction	MonthlyIncome	MonthlyRate	NumCompaniesWorked	C
	1	49	279	8	0	1	2	3	61	2	1	2	5130	24907	1
	7	30	1358	24	0	1	11	4	67	3	0	3	2693	13335	1
	8	38	216	23	2	1	12	4	44	2	2	3	9526	8787	0
	21	36	1218	9	3	1	27	3	82	2	0	1	3407	6986	7
	44	30	721	1	1	1	57	3	58	3	1	4	4011	10781	1
	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
	1394	32	1373	5	3	1	1966	4	56	2	1	4	9679	10138	8
	1398	33	1303	7	1	1	1970	4	36	3	1	3	5968	18079	1
	1405	31	697	10	2	1	1979	3	40	3	2	3	11031	26862	4
	1446	34	704	28	2	1	2035	4	95	2	1	3	6712	8978	1
	1467	27	155	4	2	1	2064	2	87	4	1	2	6142	5174	1

128 rows × 56 columns

Raw 35개 > 56개로 Encoding완료

### 주요 요인

경력기간(TotalWorkingYears),  
일과 생활의 균형 정도(WorkLiffeBalance)  
마지막 승진(YearsSinceLastPromotion)  
규정외 노동시간(Overtime)  
스톡옵션 정도(StockOptionLevel)

각각의 요인에 대한 상관 분석과 저성과자와 고성과자의 차이를 비교해  
성과를 내는 요인에 대해 분석해 보기로 함

**상관계수는 +- 0.5 이상을 기준으로 측정**

# 상관분석

## 경력기간(TotalWorkingYears)

### Real High Performer

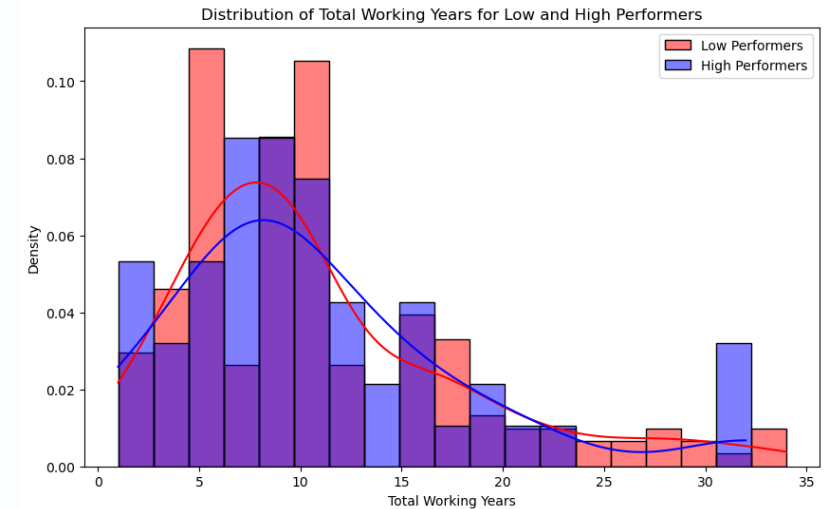
```
TotalWorkingYears와 상관관계가 0.5 이상이거나 -0.5 이하인 변수들:  
Age          0.703568  
JobLevel      0.782594  
MonthlyIncome 0.792687  
TotalWorkingYears 1.000000  
YearsAtCompany 0.509688  
JobRole_Manager 0.502228  
Name: TotalWorkingYears, dtype: float64
```

### Real Low Performer

```
TotalWorkingYears와 상관관계가 0.5 이상이거나 -0.5 이하인 변수들:  
Age          0.675704  
JobLevel      0.791588  
MonthlyIncome 0.775093  
TotalWorkingYears 1.000000  
YearsAtCompany 0.656271  
JobRole_Manager 0.559635  
Name: TotalWorkingYears, dtype: float64
```

나이, 직업수준, 월급, 근속기간이 강한 상관관계  
를 가지는 것으로 나타남  
특이점으로 직무 중 Manager의 경력기간이 높은  
것을 알 수 있음

### 두 그룹간 차이 (KDE Graph, Histogram)



대체로 저년차(5년이하)에 LowPerformer의 비율이 높으며  
시간이 지날수록 HighPerformer비율이 늘어나는  
현상을 볼 수 있음  
5-7년차 구간에 비율이 크게 상반됨을 보임

# 상관분석

## 일과 생활의 균형 정도(WorkLifeBalance)

### Real High Performer

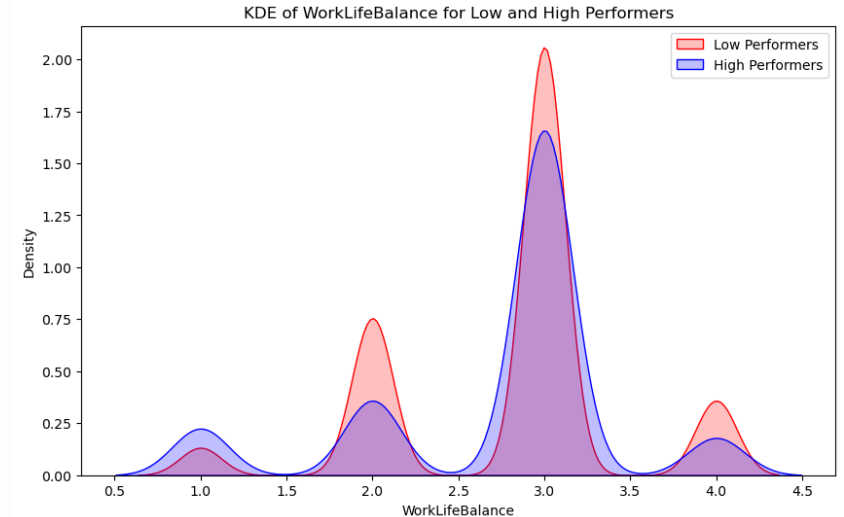
```
WorkLifeBalance와 상관관계가 0.5 이상이거나 -0.5 이하인 변수들:  
WorkLifeBalance    1.0  
Name: WorkLifeBalance, dtype: float64
```

### Real Low Performer

```
WorkLifeBalance와 상관관계가 0.5 이상이거나 -0.5 이하인 변수들:  
WorkLifeBalance    1.0  
Name: WorkLifeBalance, dtype: float64
```

워라벨과 강한 상관관계를 가지는 요인은  
발견되지 않음

### 두 그룹간 차이 (KDE Graph)



대체로 LowPerformer가 HighPerformer보다  
워라벨이 높다고 측정됨

# 상관분석

## 마지막 승진(YearsSinceLastPromotion)

### Real High Performer

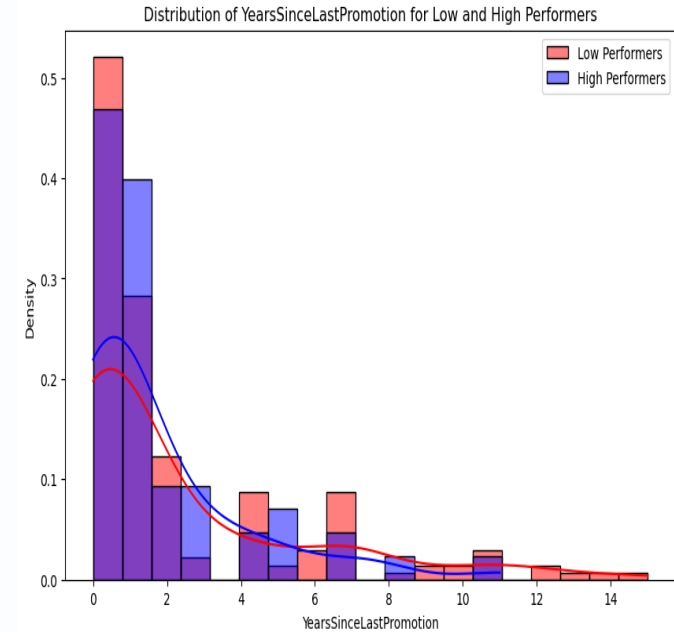
```
YearsSinceLastPromotion와 상관관계가 0.5 이상이거나 -0.5 이하인 변수들:  
YearsAtCompany      0.674485  
YearsInCurrentRole   0.612198  
YearsSinceLastPromotion  1.000000  
YearsWithCurrManager 0.577849  
Name: YearsSinceLastPromotion, dtype: float64
```

### Real Low Performer

```
YearsSinceLastPromotion와 상관관계가 0.5 이상이거나 -0.5 이하인 변수들:  
YearsAtCompany      0.560225  
YearsInCurrentRole   0.581730  
YearsSinceLastPromotion  1.000000  
Name: YearsSinceLastPromotion, dtype: float64
```

공통적으로 근속 연수, 현재 역할의 연수가 마지막 승진과 상관관계를 보임  
특이점으로 HighPerfomer은 현재 관리자와 보낸 시간과 상관관계를 가짐

### 두 그룹간 차이 (KDE Graph, Histogram)



승진하고 1~2년 사이에 perform이 향상되는 현상이 보이며, 승진이 지체된 직원들은 대체로 LowPerformance을 내고 있음

# 상관분석

## 규정외 노동시간(Overtime)

### Real High Performer

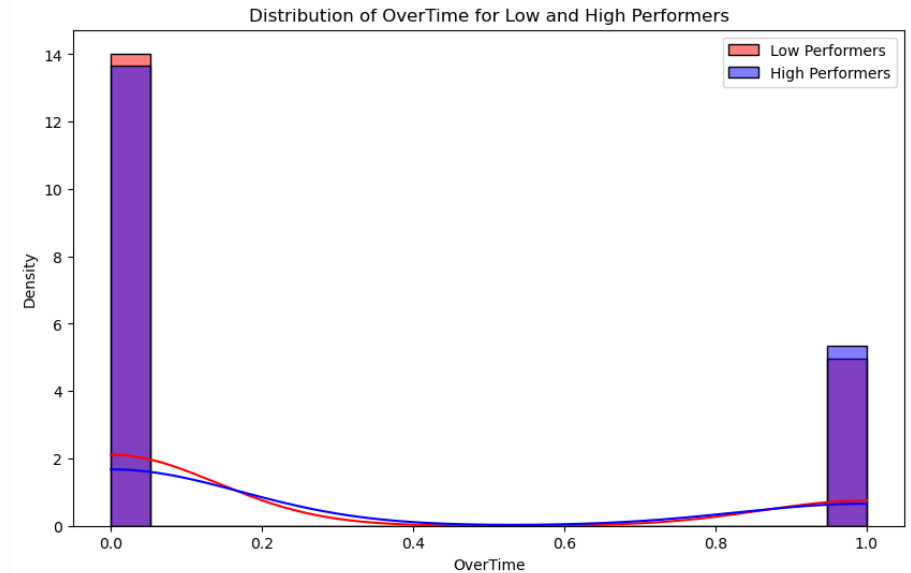
```
OverTime와 상관관계가 0.5 이상이거나 -0.5 이하인 변수들:  
OverTime    1.0  
Name: OverTime, dtype: float64
```

### Real Low Performer

```
OverTime와 상관관계가 0.5 이상이거나 -0.5 이하인 변수들:  
OverTime    1.0  
Name: OverTime, dtype: float64
```

OverTime은 다른 요인들과 별다른 상관 관계를 가지지 않는 것으로 나타남

### 두 그룹간 차이 (KDE Graph, Histogram)



OverTime은 다른 요인들과 별다른 상관 관계를 가지지 않는 것으로 나타남



# 상관분석

## 스톡옵션 정도(StockOptionLevel)

### Real High Performer

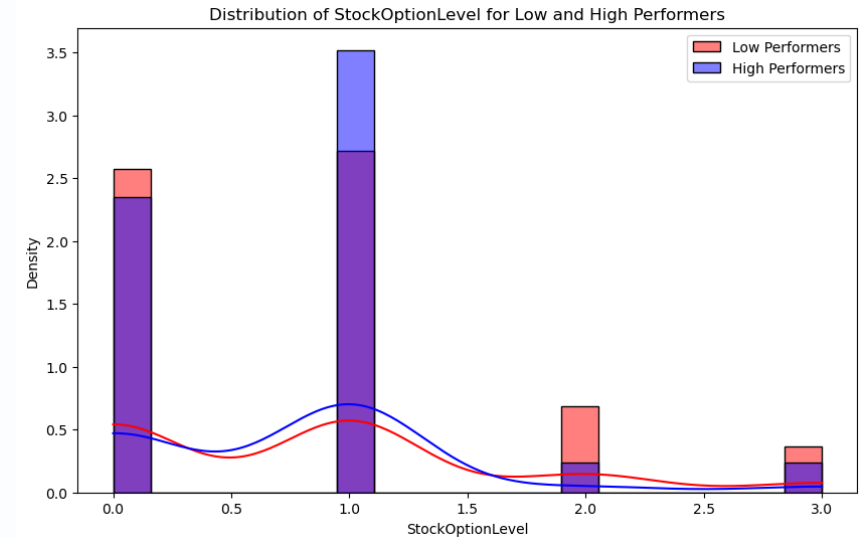
```
StockOptionLevel와 상관관계가 0.5 이상이거나 -0.5 이하인 변수들:  
StockOptionLevel      1.000000  
MaritalStatus_Single  -0.685085  
Name: StockOptionLevel, dtype: float64
```

### Real Low Performer

```
StockOptionLevel와 상관관계가 0.5 이상이거나 -0.5 이하인 변수들:  
StockOptionLevel      1.000000  
MaritalStatus_Single  -0.627576  
Name: StockOptionLevel, dtype: float64
```

Single일 경우, 스톡옵션의 정도가 낮은 경향을 보임

### 두 그룹간 차이 (KDE Graph, Histogram)



스톡옵션의 정도가 높을수록, 오히려  
LowPerformance를 내는 현상이 나타남

# 결론

**스톡옵션이 높은 저성과자의 경우 - 계약직으로의 직간 전환, 정리해고, 권고사직등의 조치가 필요함**

**장기간 승진하지 못한 저성과자에 경우 - 새로운 직급부여**

**이외로도,**

**사수 재배치, 역량개발 프로그램 기회제공 등이 저성과자들의 퍼포먼스를 향상시킬 수 있을것으로 보여짐**