

2022

KT AIVLE School

# [미니 프로젝트] 회사 직원 이직 여부 분석

수도권 2반 3조(이세빈, 조혜빈, 차선희, 차혜돈, 최기호)

“

# Target

”

회사 직원의 이직에 영향을 주는 요인은 무엇인가?

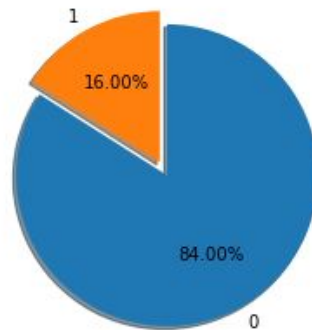
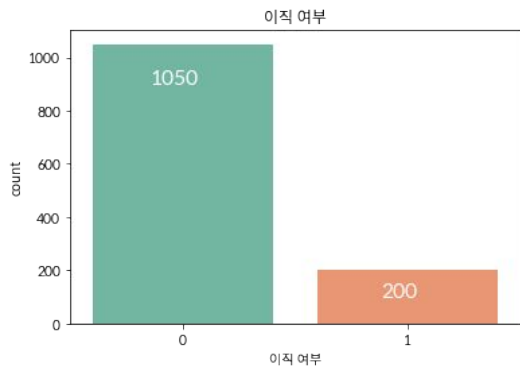
## 가설

### 우리 회사 직원의 이직에 영향을 주는 요인

- ① 나이는 이직 여부에 영향을 미칠 것이다.
- ② 집-직장 거리는 이직 여부에 영향을 미칠 것이다.
- ③ 결혼 상태는 이직 여부에 영향을 미칠 것이다.
- ④ 월급은 이직 여부에 영향을 미칠 것이다.
- ⑤ 야근 여부는 이직 여부에 영향을 미칠 것이다.
- ⑥ 총 경력 연수는 이직 여부에 영향을 미칠 것이다.

## 01. Target

### Attrition(이직 여부)



- ✓ 회사 직원들의 이직 여부를 파악했을 때, 잔류(84%)가 이직(16%)보다 많음.
- ✓ 다른 변수와의 관계를 파악할 필요 있음.

“

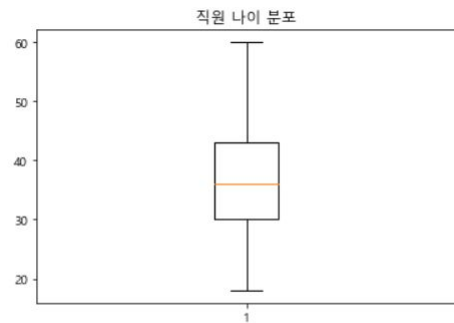
Age

”

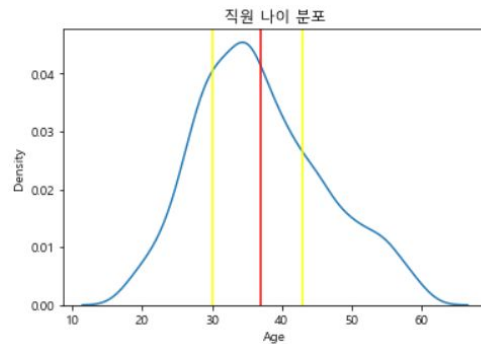
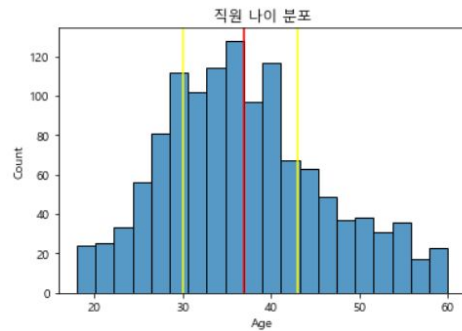
나이가 이직 여부에 영향을 미칠 것

## 02. Age(나이)

### 단변량 분석



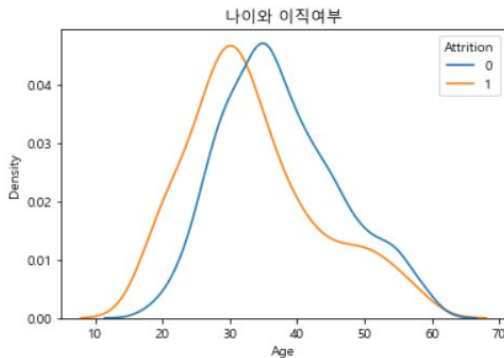
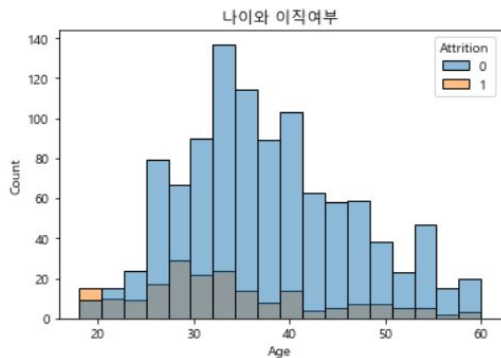
[30. 18.]  
[43. 60.]



- ✓ 30~43에 직원에 50%가량이 분포해 있고 60세 이하의 직원이 75%이다.
- ✓ 직원들의 대부분은 30~40대 초반인 것으로 추정되어지며 이 나이대 직원들의 이직여부에 대한 조사가 중요할 것으로 보인다.

## 02. Age(나이)

### 이변량 분석



```
Optimization terminated successfully.  
Current function value: 0.424660  
Iterations 5  
Age 4.667821e-100  
dtype: float64
```

- ✓ 그래프상 20대 초반 이전을 제외하고는 이직하는 사람이 적다.
- ✓ 20대 초반 이전 직원을 따로 분류하여 분석해 볼 필요가 있다.
- ✓ 나이에 따라 이직여부 분포에 차이가 있는것으로 보인다.
- ✓ 검정결과 p-value값이 0에 가까워 나이에 따른 이직여부에는 차이가 있다.

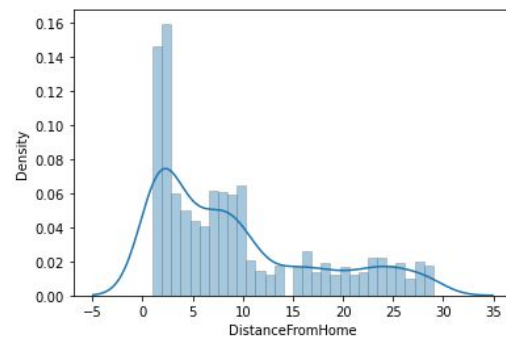
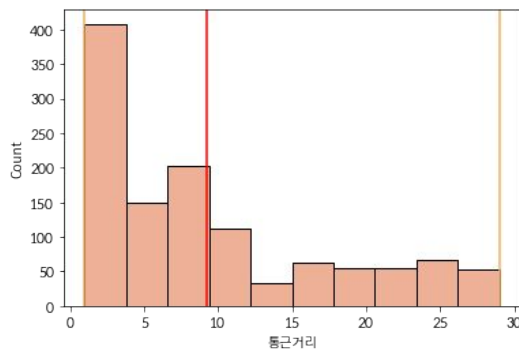
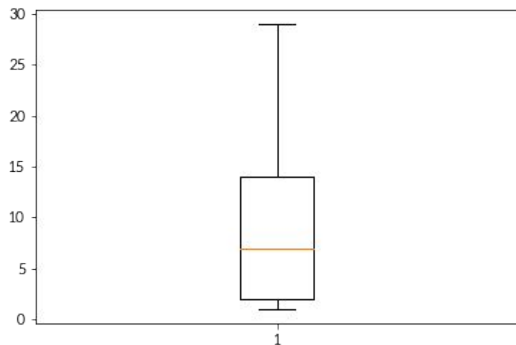
“ DistanceFromHome ”

집-직장 거리가 이직 여부에 영향을 미칠 것



### 03. DistanceFromHome(집-직장 거리)

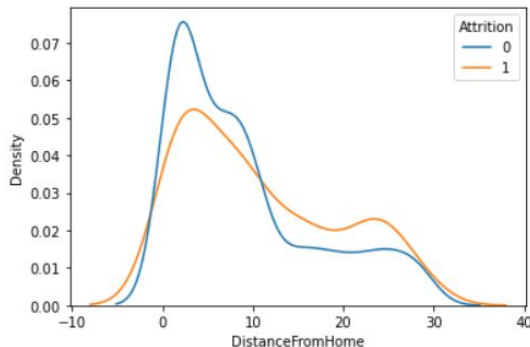
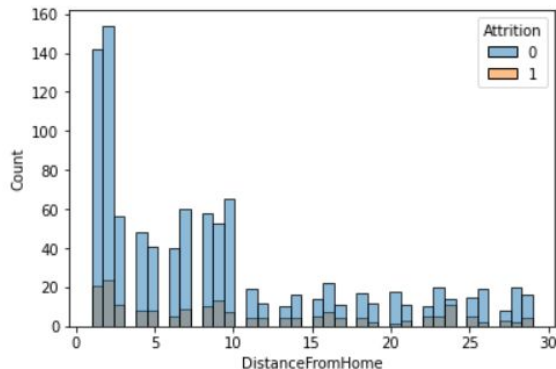
#### 단변량 분석



- ✓ 약 25% 직원들이 2마일 이내에 살고 있다.
- ✓ 그래프상 15마일 이하, 이상인 2그룹으로 나눠 분석해볼 수 있겠다.
- ✓ 직장과의 거리가 먼 25%의 직원들의 이직여부와 그렇지 않은 직원들의 이직여부를 비교해 볼 필요가 있다.
- ✓ 결측치 제외: 결측치가 전체 데이터의 약 5%미만이며, 큰 영향을 미치지 않을 것으로 판단하여 제거하였다.

### 03. DistanceFromHome(집-직장 거리)

#### 이변량 분석



Optimization terminated successfully.  
Current function value: 0.562055  
Iterations 6  
집-직장 거리 5.890702e-48  
dtype: float64

- ✓ 그래프상 집-회사 거리에 따른 이직여부에는 유의미한 차이가 있어 보인다.
- ✓ 검정 결과 집-회사 거리에 따른 이직여부에는 유의미한 차이가 있다 판단되고, 교차지점을 확인하여 분석 해볼만 하다.
- ✓ 집-회사 거리가 10마일 이상인 직원들의 이직 비율이 높기 때문에 집 회사 거리가 먼 직원들을 위해 방안을 고안해야 함

“

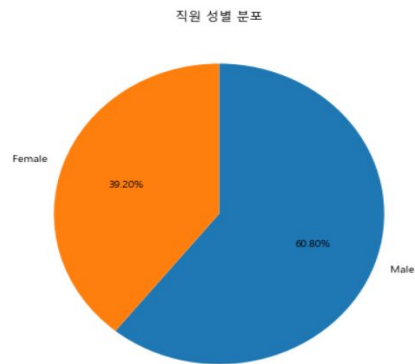
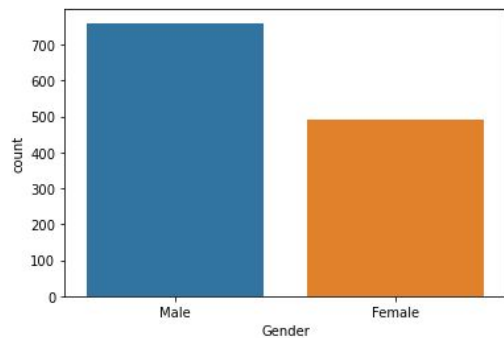
# Gender

”

성별이 이직 여부에 영향을 미칠 것

## 04. Gender(성별)

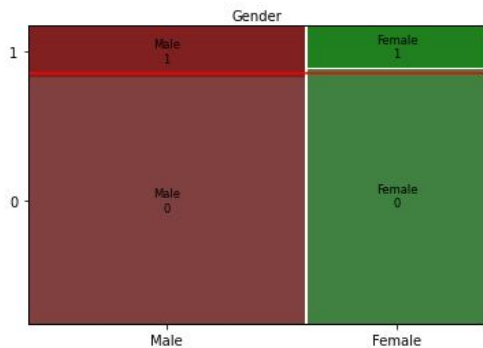
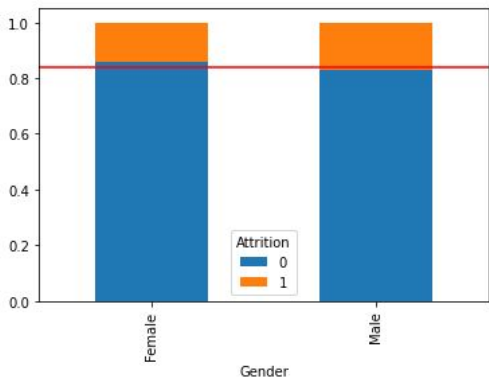
### 단변량 분석



✓ 여성에 비해 남성 직원의 수가 더 많은 것을 확인할 수 있다.

## 04. Gender(성별)

### 이변량 분석



카이제곱통계량 1.9782495940105371  
p-value 0.15957492990528155  
기대빈도  
[[411.6 638.4]  
[ 78.4 121.6]]

✓ 시각화를 통해서도 성별은 이직여부에 영향을 미치지 않는 것으로 보인다.

✓ 가설검정을 통해 p-value가 유의수준인 0.05에 비해 큰 값인 약 0.16으로 나온 것을 확인할 수 있으므로 성별과 이직여부는 관련이 없는 것으로 보인다.

“

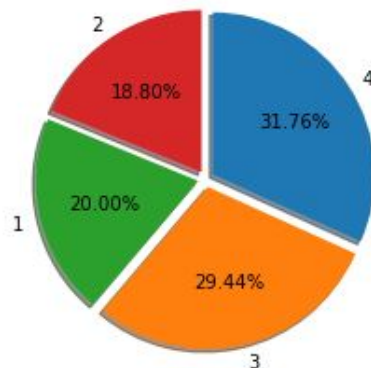
# JobSatisfaction

”

직무 만족도가 이직 여부에 영향을 미칠 것

## 05. JobSatisfaction(직무 만족도)

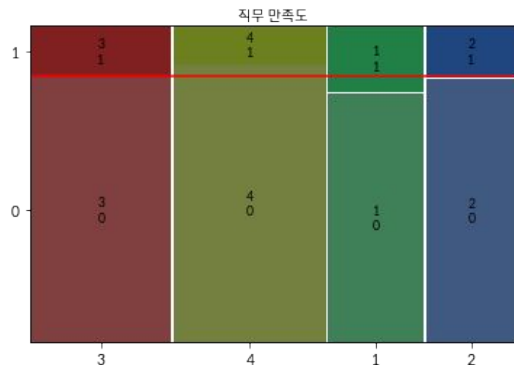
### 단변량 분석



- ✓ 직무 만족도 평가에 부정적인(1.Low) 평가를 한 직원이 전체의 20% (250명)의 수치로 긍정적인 평가를 한 직원에 비해 적은 편이긴 하지만 20% 라는 수치는 적은 편이라고는 할 수 없다.
- ✓ 직무 만족도를 위해 회사 내규에 따라 직무 변경 등 직원들의 편의를 보장해 줄 필요가 있다.
- ✓ 부정적인 평가를 한 직원들의 이직 여부와 긍정적인 평가를 한 직원들의 이직 여부를 비교 분석해 볼 필요가 있다.
- ✓ 만족 이상이 60%로 절반 이상이다. -> 이직률이 낮을 것으로 예상된다.

## 05. JobSatisfaction(직무 만족도)

### 이변량 분석



카이제곱통계량 : 9.63151465436614  
p-value : 0.02197264936930621443

✓ 가설검정을 통해 p-value가 약 0.02로 유의수준인 0.05보다 작지만 보수적으로 보면 0.01보다 높아 직무 만족도와 이직 여부는 관계가 있다고 할 수 없다.

✓ 직무 만족도에서 부정적인 평가를 한 직원에서의 이직이 가장 많이 일어난다.



“

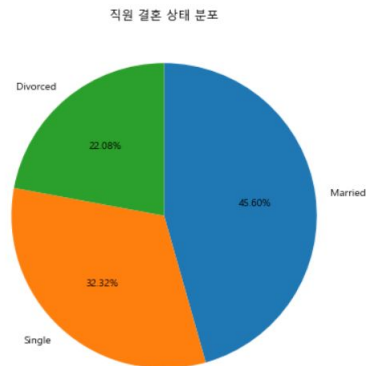
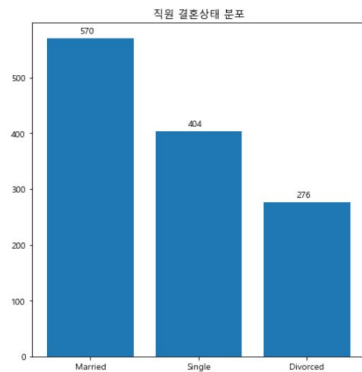
# MaritalStatus

”

결혼상태가 이직 여부에 영향을 미칠 것

## 06. MaritalStatus(결혼상태)

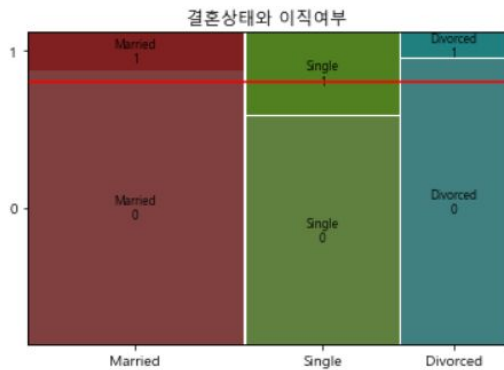
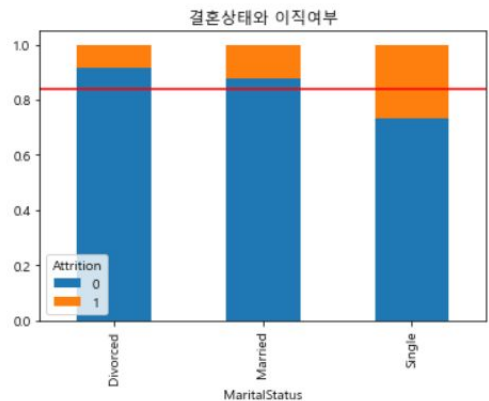
### 단변량 분석



✓ 싱글과 이혼한 직원이 과반수 이상을 차지하고 있다.

## 06. MaritalStatus(결혼상태)

### 이변량 분석



Attrition	0	1
MaritalStatus		
Divorced	253	23
Married	501	69
Single	296	108

---

카이제곱 통계량 : 53.12935546374186  
P-value : 2.9047385436211595e-12  
기대빈도 : [[231.84 44.16]  
[478.8 91.2]  
[339.36 64.64]]

- ✓ 그래프상 결혼상태에 따른 이직여부에 유의미한 차이가 있어보인다.
- ✓ 검정결과 p-value값이 충분히 작아 결혼상태에 따른 이직여부에는 차이가 있다.

“

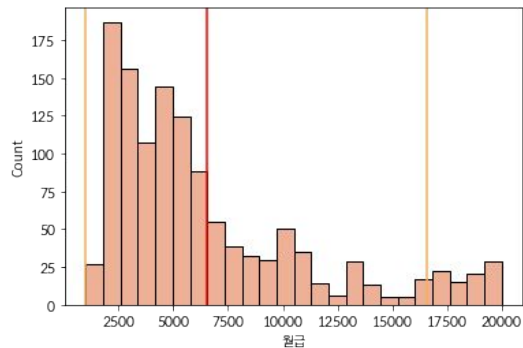
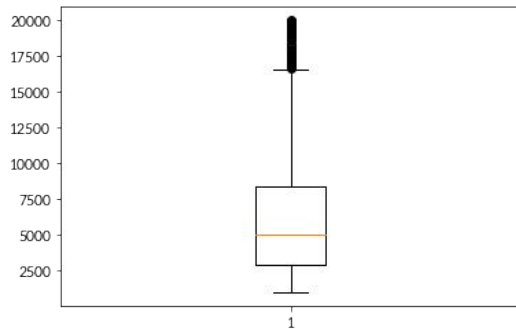
# MonthlyIncome

”

월급이 이직 여부에 영향을 미칠 것

## 07. MonthlyIncome(월급)

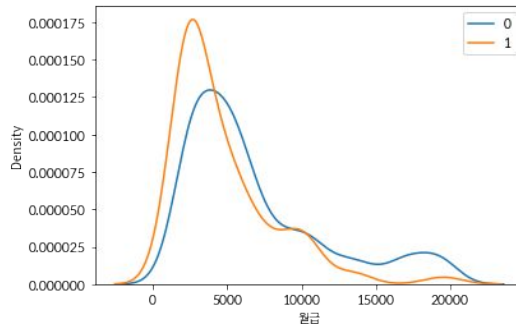
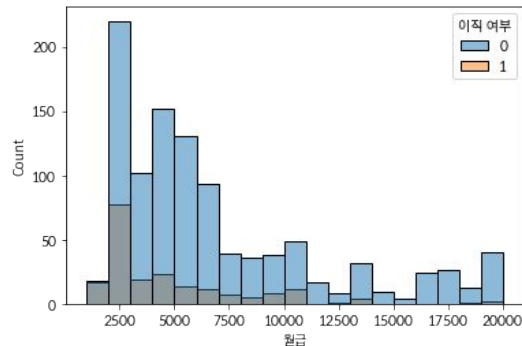
### 단변량 분석



- ✓ 4988달러 이하의 월급을 전 직원에 50%가 받고 있는 부분과 그래프가 왼쪽으로 치우쳐 있는것으로 볼때 월급 격차가 심하다
- ✓ 이상치가 존재한다(임원급 예상)
- ✓ 그래프상 평균월급(약 6,500달러) 이상, 이하 2분류로 나눠 분석해볼 수 있겠다.

## 07. MonthlyIncome(월급)

### 이변량 분석



Optimization terminated successfully.  
Current function value: 0.437401  
Iterations 6  
월급 1.922717e-71  
dtype: float64

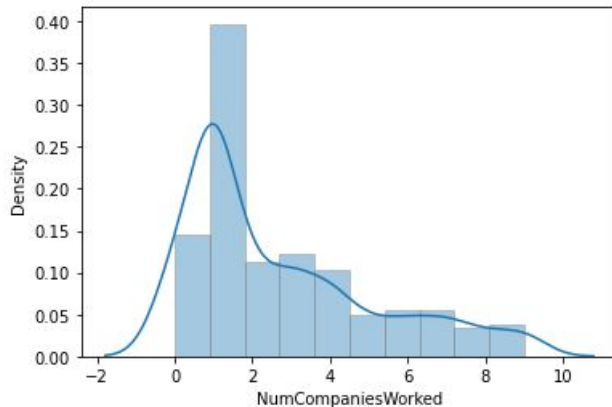
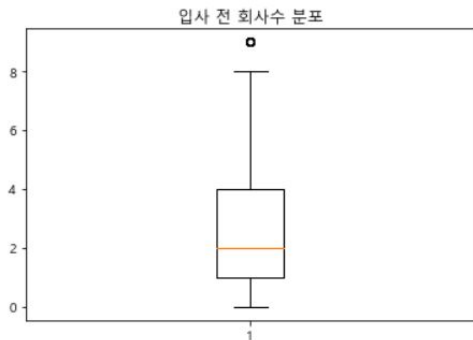
- ✓ 그래프상 직원월급에 따른 이직여부 분포에 차이가 보인다.
- ✓ 월급이 5000달러 이하인 직원들의 이직 분포가 높은편이다.
- ✓ 검정결과 p-value값이 0에 충분히 가까워 직원월급에 따른 이직여부에는 유의미한 차이가 있다.
- ✓ 신입 직원이 월급이 5,000달러 이하의 직원들에 해당할 확률이 크다. 신입 직원의 이직을 막을 방안을 고안해야 한다.

“ NumCompaniesWorked ”

입사 전 근무한 회사 수가 이직 여부에 영향을 미칠 것

## 08. NumCompaniesWorked(입사 전 근무한 회사 수)

### 단변량 분석

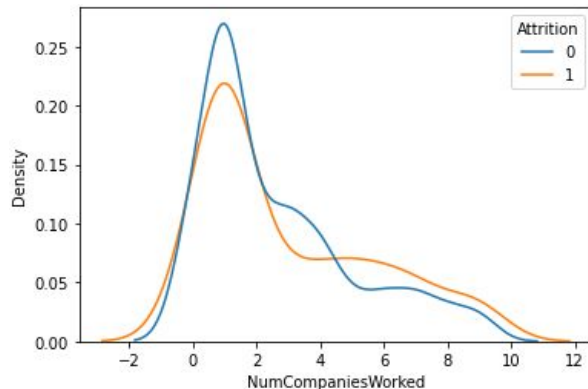
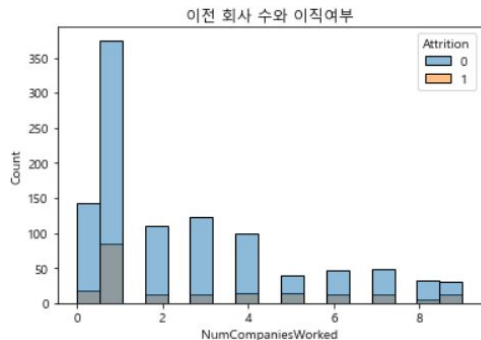


- ✓ 입사 전 회사 수가 많은 직원은 회사를 자주 옮긴다는 것으로 간주하고 이직 회수가 많은 직원들의 현재 이직 여부를 검토해 볼 필요가 있다.
- ✓ 입사 전 회사수가 8개이상인 직원은 이상치로 검토해볼 필요가 있다.
- ✓ 그래프가 왼쪽으로 치우쳐있고 50%의 직원들은 2개이하의 회사에서 근무하고 입사하였다.
- ✓ 총 경력 연수를 고려하여 입사 전 근무한 회사 수를 통해 이직 횟수가 많은지 파악할 필요가 있다.
- ✓ 그래프가 전체적으로 왼쪽으로 치우쳐있어 결측치를 최빈값으로 대체하여 진행하였다.



## 08. NumCompaniesWorked(입사 전 근무한 회사 수)

### 이변량 분석



```
Optimization terminated successfully.  
Current function value: 0.564409  
Iterations 6  
NumCompaniesWorked 7.457295e-46  
dtype: float64
```

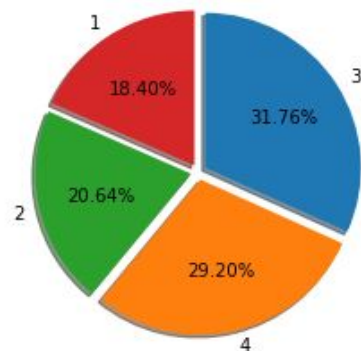
- ✓ 전체적으로 이직하지 않은 직원들의 수가 이직한 직원들의 수 보다 많은 부분을 차지하지만 밀도 그래프에서 입사 전 근무한 회사 수가 4개 이상의 직원들의 이직 비율이 더 큰 부분을 차지하고 있다.
- ✓ 그래프상 이전 회사 수에 따른 이직여부에 차이가 어느정도 있으나 많지 않은 것으로 보인다.
- ✓ 검정결과 이전 회사수에 따른 이직여부에 차이가 있다.(약한관계)

# “ Relationship Satisfaction ”

동료와의 관계가 이직 여부에 영향을 미칠 것

## 09. Relationship Satisfaction(동료와의 관계)

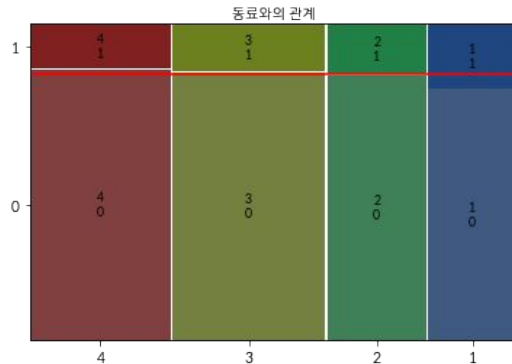
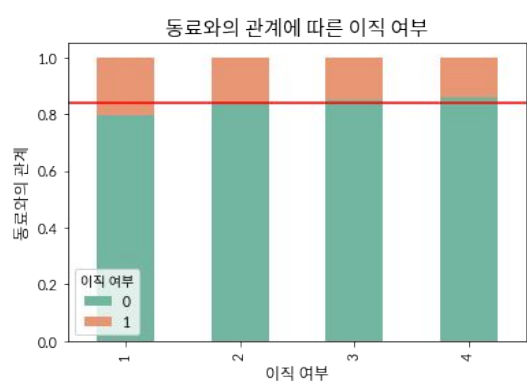
### 단변량 분석



- ✓ 동료와의 관계에서 긍정적인 범주가 높게 나온 것을 알 수 있다.
- ✓ 동료와의 관계에 부정적인(1.Low) 평가를 한 직원이 18.4%(230명)의 수치로 긍정적인 평가를 한 직원에 비해 적은 편이긴 하지만 간과할 만큼 적은 수치라고 생각하면 안 될 것이다.
- ✓ 동료와의 관계를 위해 회사 내에 회사 동료들 간의 커뮤니케이션을 할 수 있는 장소 설치하는 등 여러 방안들을 고려해 볼 필요가 있다.
- ✓ 동료와의 관계와 직무만족도의 관계가 서로 비슷한 것으로 보아 묶어 해석할 여지가 있다.

## 09. Relationship Satisfaction(동료와의 관계)

### 이변량 분석



카이제곱통계량 : 4.435126158518535  
p-value : 0.2181503949975081

- ✓ 동료와의 관계에서 이직률은 평균과 비슷한 수준인 것을 확인할 수 있으므로 이직 여부와 크게 관련이 없어 보인다.
- ✓ 가설검정을 통해 p-value가 약 0.2로 유의수준인 0.05보다 큰 것을 확인할 수 있으며 동료와의 관계는 이직 여부와 관련이 없다고 할 수 있다.
- ✓ 각 부서의 이직 여부를 파악하여 각 부서별 동료와의 관계가 이직 여부에 영향을 미치는지 파악할 필요가 있다.

“

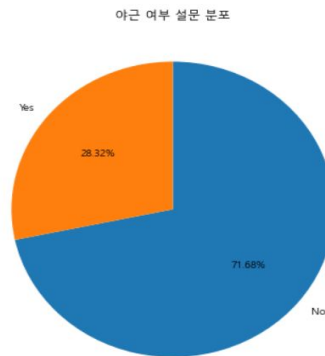
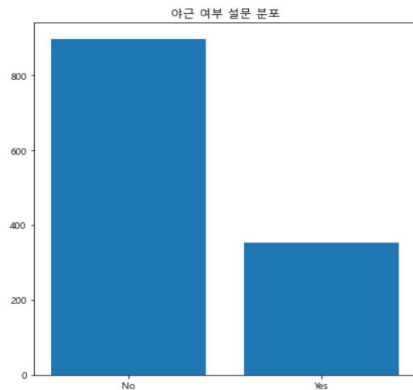
# OverTime

”

야근 여부가 이직 여부에 영향을 미칠 것

## 10. OverTime(야근여부)

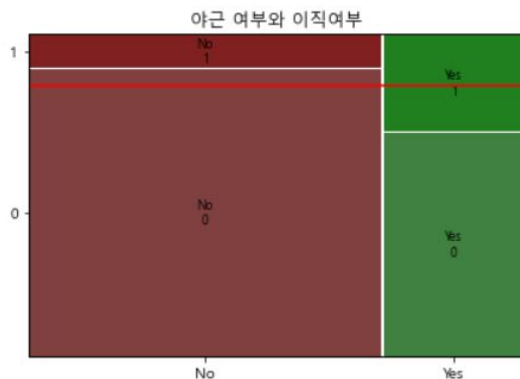
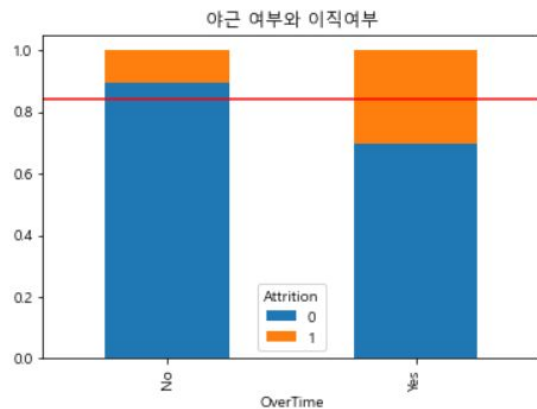
### 단변량 분석



- ✓ 필요한 야근에는 반드시 합당한 보상이 필요하고 야근을 지양하는 회사라면 특정 시간 이후에 자동으로 컴퓨터의 전원이 꺼지는 등 시스템의 도입을 고려해 볼 필요가 있다.
- ✓ 7:3 비율로 야근비율이 낮다. 회사 만족도에 대한 이유가 될 수 있다.

## 10. OverTime(야근여부)

### 이변량 분석



Attrition	0	1
OverTime		
No	803	93
Yes	247	107

---

카이제곱 통계량 : 72.89605419432075  
P-value : 1.366611233098835e-17  
기대빈도 : [[752.64 143.36]  
[297.36 56.64]]

- ✓ 그래프상 야근여부에 따른 이직여부는 극명한 차이를 보이고 있다.
- ✓ 검정결과 p-value가 충분히 낮아 야근여부에 따른 이직여부에는 차이가 있다

“

PercentSalaryHike

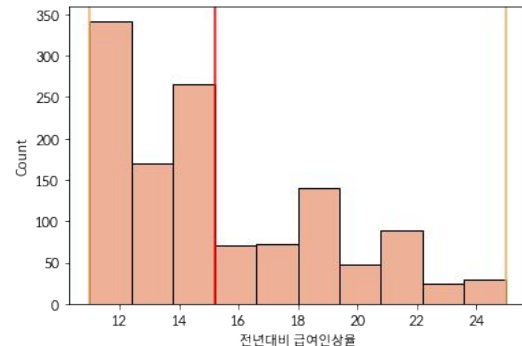
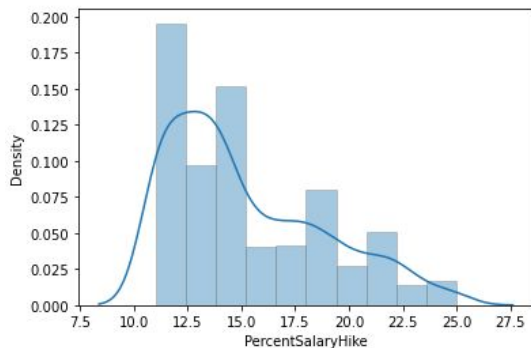
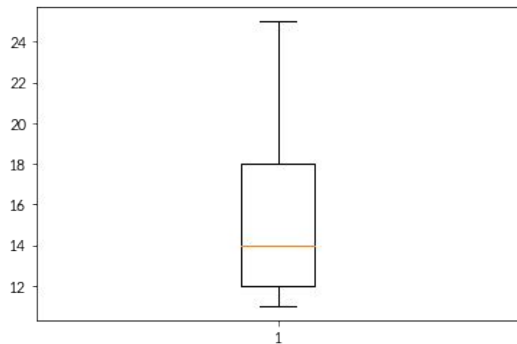
”

전년대비 급여 인상률이 이직 여부에 영향을 미칠 것



## 11. PercentSalaryHike(전년대비 급여인상율)

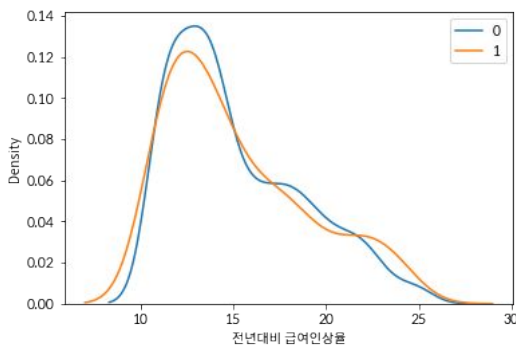
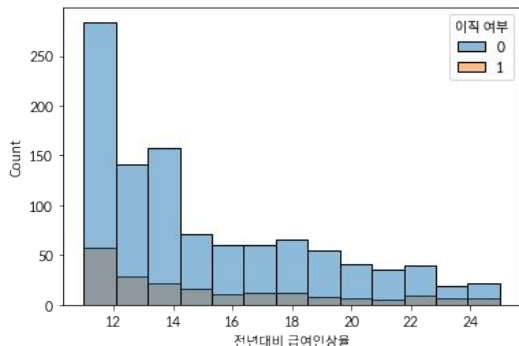
### 단변량 분석



- ✓ 평균(약 15%) 이하에 많은 직원들이 분포하기 때문에 평균 이상, 이하 2분류로 구분하여 분석해 볼 만하다.
- ✓ 직원 75%의 급여는 최소 11%에서 18% 정도 오르게 되었다.
- ✓ 이상치는 없다.

## 11. PercentSalaryHike(전년대비 급여인상율)

### 이변량 분석



Optimization terminated successfully.  
Current function value: 0.450230  
Iterations 5  
전년대비 급여인상율 3.979995e-95  
dtype: float64

- ✓ 인상률이 약 10% 이하인 직원의 이직률이 더 높은 것을 확인할 수 있다.
- ✓ 인상률이 약 20% ~ 25%인 직원들의 이직률도 잔류하는 직원에 비해 높은 것을 확인할 수 있었는데 이는 다른 회사의 스카우트 제의 및 새로운 도전을 했을 가능성이 있다.
- ✓ 급여 인상율 변화에 따른 이직여부에는 차이가 많아보이지 않다.
- ✓ 검정결과 p-value값이 충분히 작아 급여 인상율에 따른 이직여부에는 차이가 있다.(약한관계)

“

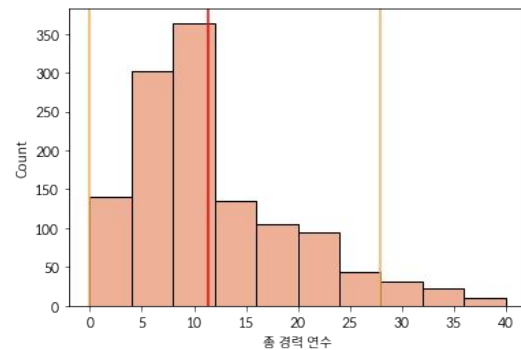
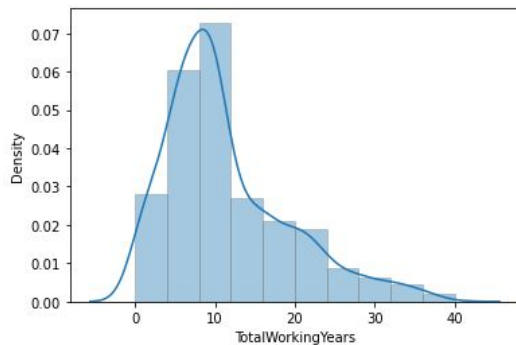
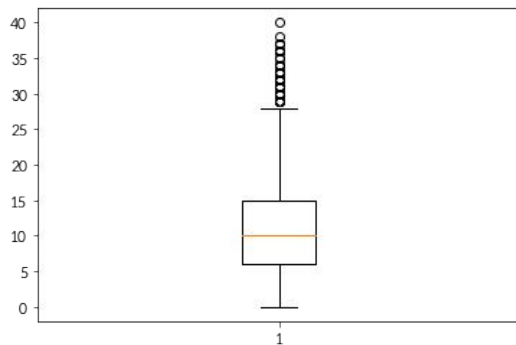
# TotalWorkingYears

”

총 경력 연수가 이직 여부에 영향을 미칠 것

## 12. TotalWorkingYears(총 경력 연수)

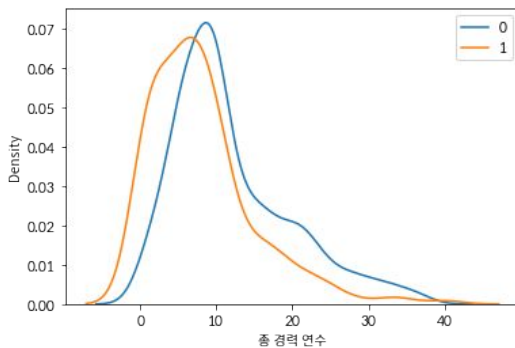
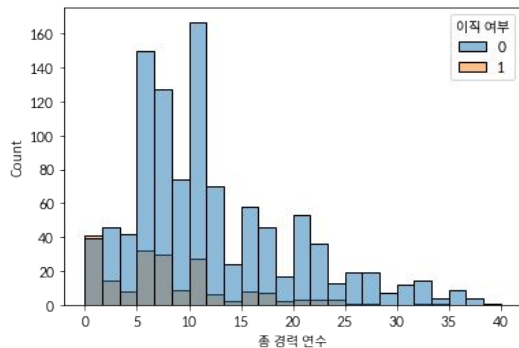
### 단변량 분석



- ✓ 총 경력 연수가 28년 이상에서 이상치가 발생하여 이상치에 대한 대책이 필요하다.
- ✓ 총 경력 연수의 최댓값은 40년이고 최솟값은 0년이다.
- ✓ 총 경력 연수가 높은 직원들은 회사의 문화, 환경 등에 잘 적응하여 이직할 확률을 낮을 것이라고 예상한다.

## 12. TotalWorkingYears(총 경력 연수)

### 이변량 분석



Optimization terminated successfully.  
Current function value: 0.432906  
Iterations 6  
총 경력 연수 1.867338e-74  
dtype: float64

- ✓ 그래프 상 총 경력 연수에 따른 이직 여부를 파악했을 때, 관계가 있고 그 차이가 많아 보인다.
- ✓ 검정결과, p-value 값이 유의수준(0.05)보다 작기 때문에 총 경력 연수에 따른 이직 여부에는 차이가 있다.
- ✓ 1~2년차 신입 직원들의 이탈이 보인다. 신입 직원들의 조기퇴사율을 낮추기 위한 방안을 고안할 필요 있다.
- ✓ 입사 후 1년 정도 후부터 이직율이 낮는데, 그 요인을 파악해볼 필요가 있다.

“

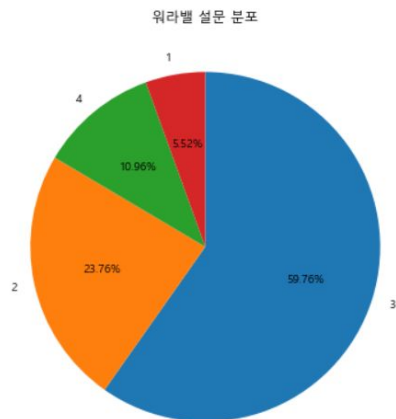
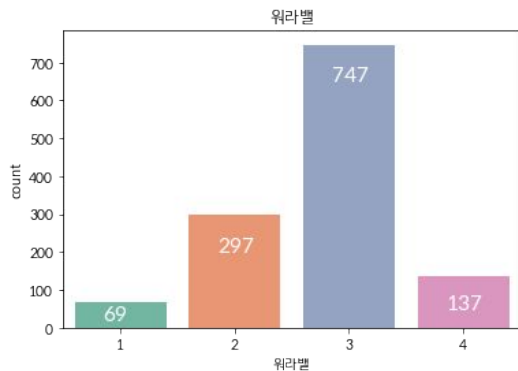
# WorkLifeBalance

”

워라밸이 이직 여부에 영향을 미칠 것

### 13. WorkLifeBalance(워라밸. 일-삶 균형도)

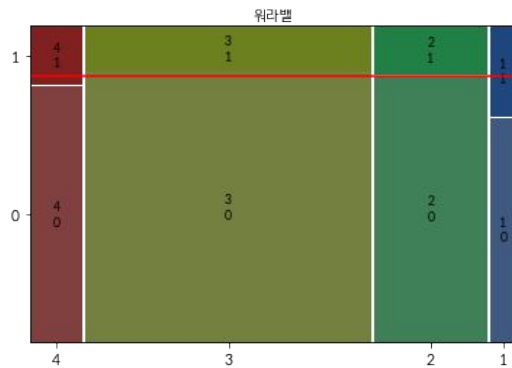
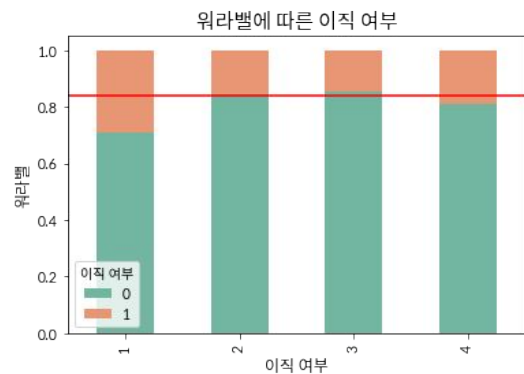
#### 단변량 분석



- ✓ 워라밸이 bad라고 답변한 설문은 전체의 5.52%로 굉장히 적은 편이다.
- ✓ better로 대답한 직원이 전체의 60%가량을 차지한다.
- ✓ 각 범주별로 나누어 이직 여부에 영향을 미치는지 확인해볼 필요가 있다.
- ✓ 전반적으로 회사만족도가 높아 이직률이 낮을 것이라 예상한다. 또한 건강한 회사일 확률이 높다.

### 13. WorkLifeBalance(워라밸. 일-삶 균형도)

#### 이변량 분석



카이제곱통계량 : 10.940810822342664  
p-value : 0.012050277390233675

- ✓ 검정 결과, p-value 값이 0.01로 유의수준(0.05)보다 작게 나왔으므로 약한 관계가 있다고 볼 수 있음.
- ✓ 그래프상 부정적인 평가한 사람들의 이직여부가 나머지와 유의미한 차이를 보였으며 좋고 나쁨으로 나누어 분석해 볼 필요가 있다.



“

## 관계 정리

”

회사 직원의 이직에 영향을 주는 요인은 무엇인가?

## 14. 관계 정리

### ✓ 강한 관계의 변수

- Age (나이)
- DistanceFromHome(집-직장 거리)
- MaritalStatus(결혼상태)
- MonthlyIncome(월급)
- OverTime(야근여부)
- TotalWorkingYears(총 경력 연수)

### ✓ 약한 관계의 변수

- WorkLifeBalance(워라밸)
- JobSatisfaction(직무 만족도)
- NumCompaniesWorked(입사 전 근무한 회사 수)
- PercentSalaryHike(전년대비 급여인상율)

### ✓ 관계 없음

- Gender(성별)
- RelationshipSatisfaction (동료와의 관계)

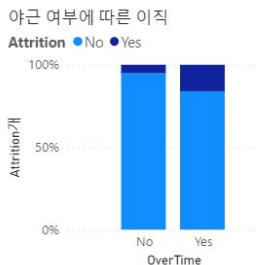
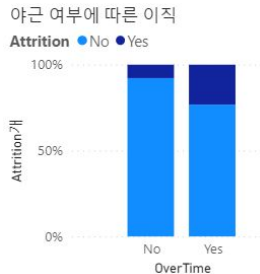
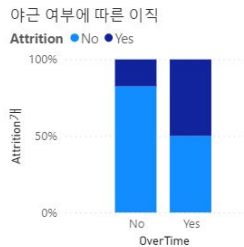
“

# 비즈니스 인사이트

”

회사 직원의 이직에 영향을 주는 요인은 무엇인가?

## 15. 비즈니스 인사이트



✓ 결혼하지 않은 싱글 직원이 야근을 많이 하고 이직률 또한 높다.

✓ 싱글에 비해 기혼자들의 이직율이 낮은 것으로 볼때 가족 친화적인 기업으로 추측된다.