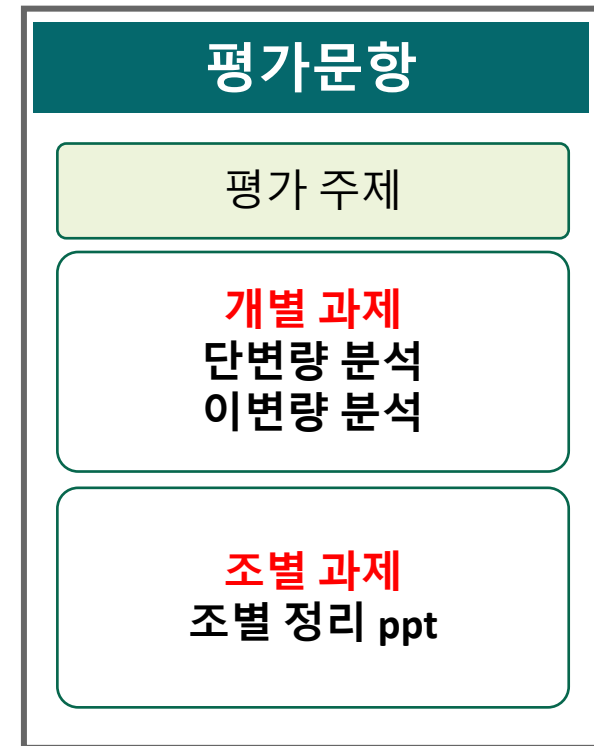
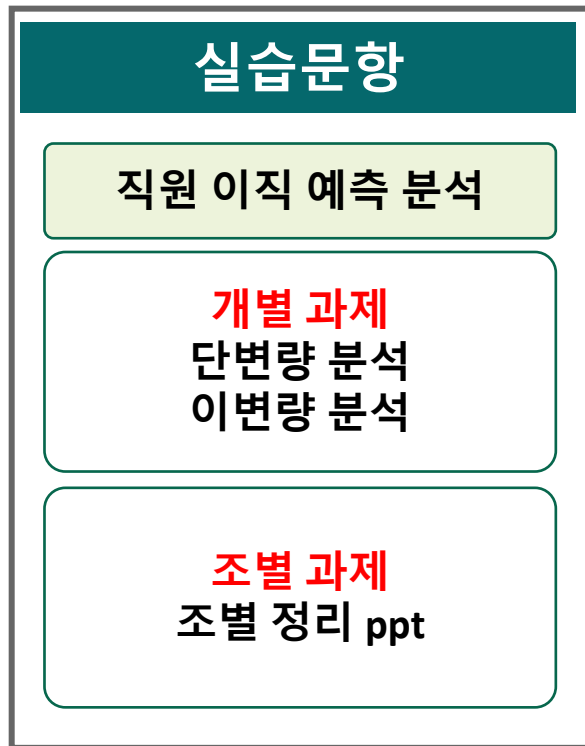


KT AIVLE School

# 미니프로젝트\_데이터처리 실습과제

# 순서



# 제출 결과물

✓ 매일 아래와 같은 결과물을 제출합니다.

✓ 개별 결과물 :

- 단변량, 이변량 분석 결과 (ipynb)

AI 000반 00조\_개별과제\_성명.ipynb

✓ 조별 결과물 :

- 조별 취합 및 정리 ppt

AI 000반 00조\_조별과제.ppt

✓ 실습문항 : 개별과제도 조별로 모여서 토의하며 진행.

✓ 평가문항 : 개별과제는 개인이 수행. 조별과제만 모여서 수행.

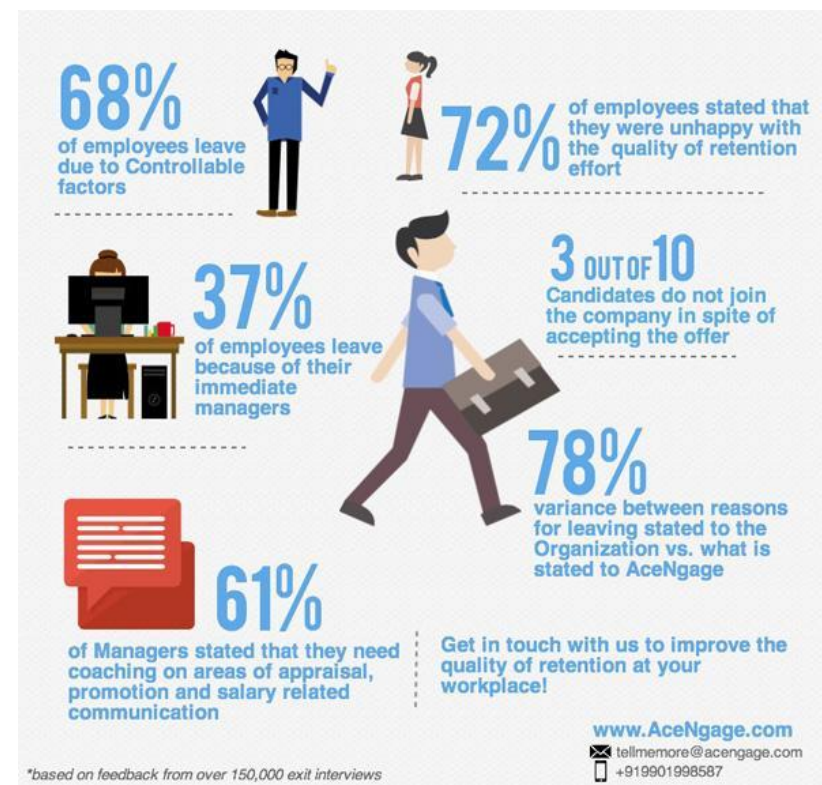
# 실습문항

# 주제 #1 : 직원 이직 예측 분석

✓ 직원들의 이직 원인을 분석하고자 합니다.

✓ 변수설명

변수 명	내용	type
Attrition	이직여부, Yes , No	범주
Age	나이	숫자
Department	현 부서	범주
DistanceFromHome	집-직장 거리(마일)	숫자
EmployeeNumber	사번	
EnvironmentSatisfaction	근무환경에 대한 만족도(범주)	범주
Gender	성별	범주
JobRole	직무	범주
JobSatisfaction	직무 만족도	범주
MaritalStatus	결혼상태	범주
MonthlyIncome	월급	숫자
NumCompaniesWorked	현재까지 근무한 회사 수	숫자
OverTime	야근여부	범주
PercentSalaryHike	전년대비 급여인상률(%)	숫자
RelationshipSatisfaction	동료와의 관계 만족도	범주
TotalWorkingYears	총 경력 연수	숫자
TrainingTimesLastYear	전년 교육훈련 횟수	숫자
WorkLifeBalance	워라벨. 일-삶 균형도	범주
YearsAtCompany	현직장 근무 연수	숫자



# 과제 #1 : 가설수립

## ✓(초기)가설 수립 $x \rightarrow y$

- 주어진 데이터셋과 기본 상식을 기반으로 가설을 도출합니다.
- 이미 데이터가 갖춰진 상태이기에 주어진 상황에서만 도출을 해 봅시다.
- 두개의 변수 조합으로도 가설 수립 가능합니다.
- 최소한 5개 이상 도출하시오.
  - 단  $x$ 변수 중 숫자형 혹은 범주형 각각 적어도 2개 이상씩은 도출해야 합니다.

## 과제 #2 : 단변량 분석

✓ 개별 변수의 분포를 살펴보고 다음 사항을 파악해 봅시다.

- 변수의 비즈니스 의미
- 숫자? 범주?
- NaN 존재 여부 및 조치 방안
- 기초통계량
- 분포
- 위 정보로부터 파악한 내용(비즈니스!) 정리
- 추가 분석해볼 사항 도출

## 과제 #3 : 이변량 분석

### ✓ $x$ 와 $y$ 의 관계를 살펴봅시다.

- 적절한 시각화를 수행하시오.
- 적절한 수치화를 수행하시오.
- 시각화와 수치화 결과를 판단하시오.
- 위 정보로부터 파악한 내용(비즈니스!) 정리
- 추가 분석해볼 사항 도출

### ✓ $Y$ 와의 관계를 아래 3가지 그룹으로 정리해봅시다.

다분히 주관적 판단입니다. 정답은 없습니다.

- $y$ 와 강한 관계의  $x$
- $y$ 와 중간 관계의  $x$
- $y$ 와 약한 관계의  $x$



kt

AIVLE

AIVLE  
Let's make it possible