

# Depression Classification

”

불안증, 졸음증 외 14가지 척도

발표시작 >



”

**01** 데이터 소개  
Data description  
& Preprocess

**02** 주제 선정 이유

**03** 데이터 분석 결과

**04** 결과 종합

**05** 한계점



”

# 데이터 소개

Data description & Preprocess

# 01

# “ 1.데이터 소개

멕시코 의대생 776명의 정신 건강 데이터

2014년 멕시코 사립 대학 4520명의 의대생 중 1200명을 무작위 샘플링. 의대생 776명 참여 동의(64.6% 응답).

- age	나이	- PHQ-9 우울증 척도
- gender	성별	- 0-4= no notable depressive symptoms
- height	키(m)	- 5-9= mild
- weight	무게(kg)	- 10-14= moderate
- thoughts_of_dropping_out	자퇴를 심각하게 생각하였는지 여부	- 15-19= moderately severe
- previous_depression_diagnosis	이전에 우울증 진단을 받은적 있는지 여부	- 20+= severe
- previous_depression_treatment	이전에 우울증 치료를 받은적 있는지 여부	- GAD-7 불안증 척도
- previous_anxiety_diagnosis	이전에 불안 진단을 받은적 있는지 여부	- 0-4= no notable anxiety symptoms
- previous_anxiety_treatment	이전에 불안 치료를 받은적 있는지 여부	- 5-9= mild
- bed_time	자가 보고한 평소 취침 시간	- 10-14= moderate
- wake_up_time	자가 보고한 평소 기상 시간	- 15+= severe
- reported_sleep_hours	자가 보고한 체감 수면 시간	- EPW Epworth 졸음증 척도
- times_week_nap	일주일에 낮잠을 자는 횟수	- 0-7: no notable sleepiness symptoms
- nap_duration	자가 보고한 낮잠 시간	- 8-9: mild
- weekly_study_hours	자가 보고한 주간 학습 시간	- 10-15: moderate
- grades	자가 보고한 학점 평균(1.0-10.0)	- 16-24: severe

# “ 1.데이터 소개

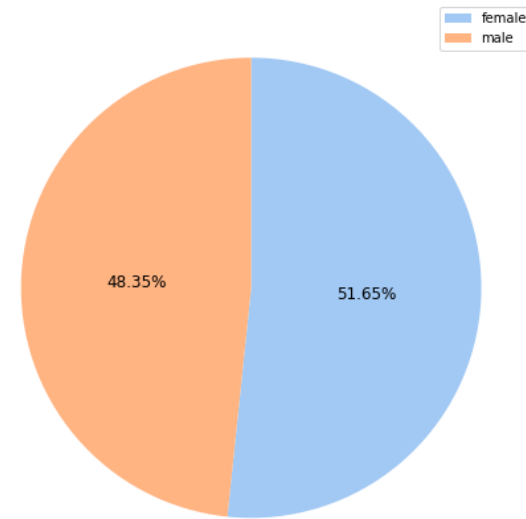
## 데이터 전처리

### 데이터 가공 및 불필요한 데이터 삭제

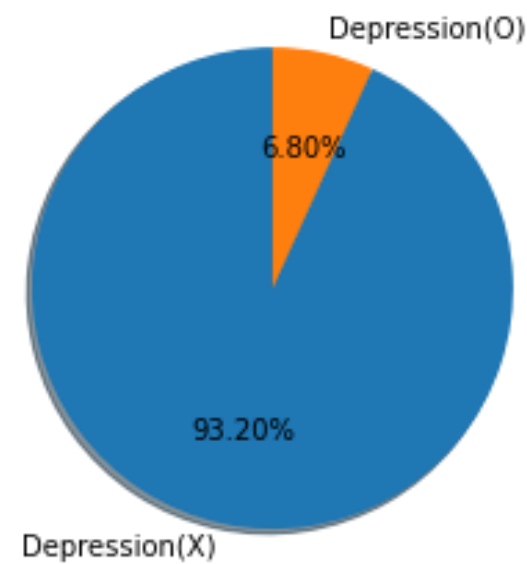
- age	나이		- PHQ-9 우울증 척도
- gender	성별		PHQ-9 문항 9개 합산 점수가 15점 이상 : 우울증 진단
- height	키(m)	→	bmi = weight / (height x height)
- weight	무게(kg)		
- thoughts_of_dropping_out	자퇴를 심각하게 생각하였는지 여부		
- previous_depression_diagnosis	이전에 우울증 진단을 받은적 있는지 여부		
- previous_depression_treatment	이전에 우울증 치료를 받은적 있는지 여부		
- previous_anxiety_diagnosis	이전에 불안 진단을 받은적 있는지 여부		
- previous_anxiety_treatment	이전에 불안 치료를 받은적 있는지 여부		
- bed_time	자가 보고한 평소 취침 시간	→	sleep_ratio = reported_sleep_hours / (wake_up_time - bed_time) 수면비 = 체감 수면시간 / 자가보고한 수면시간
- wake_up_time	자가 보고한 평소 기상 시간		
- reported_sleep_hours	자가 보고한 체감 수면 시간		
- times_week_nap	일주일에 낮잠을 자는 횟수		
- nap_duration	자가 보고한 낮잠 시간		
- weekly_study_hours	자가 보고한 주간 학습 시간		
- grades	자가 보고한 학점 평균(1.0-10.0)		

# “ 1.데이터 소개

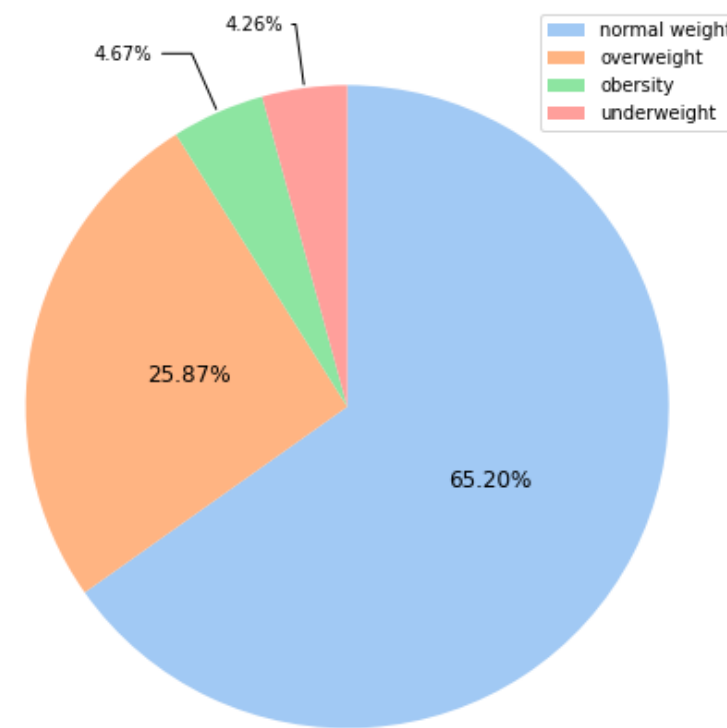
성별 : 여자 391명 / 남자 366명



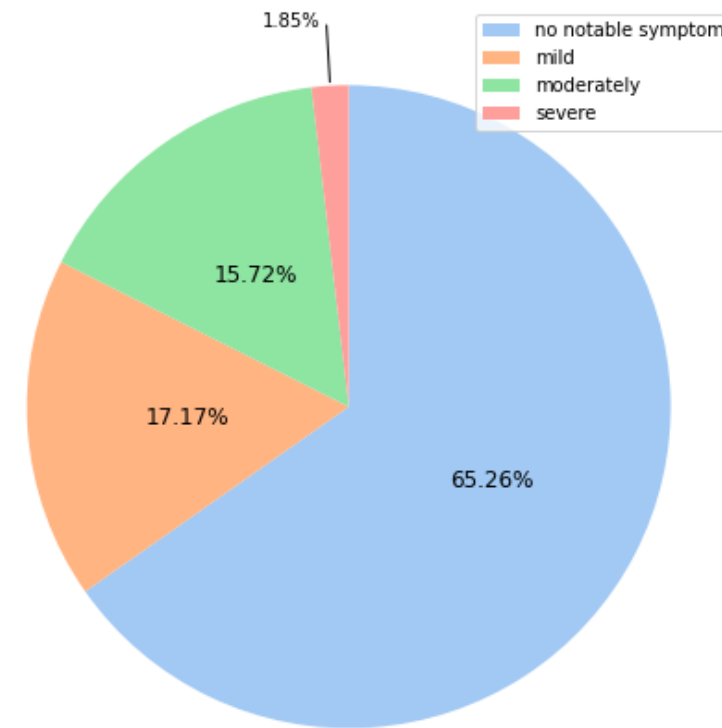
우울증: 無 713명 / 有 52명



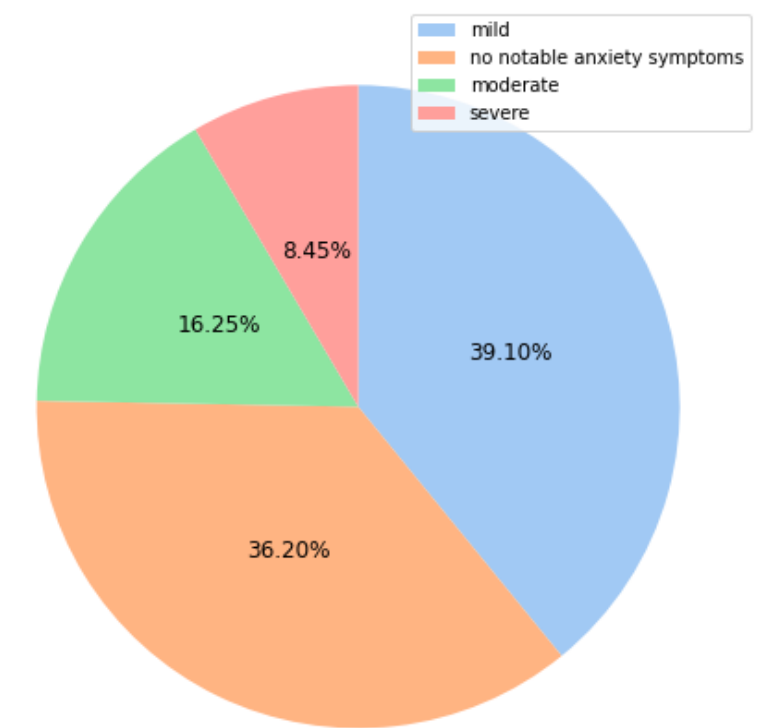
bmi :  
저체중 32명  
정상체중 489명  
과체중 194명  
비만 35명



졸음증:  
증상없음 494명  
정상범주 130명  
경증 119명  
중증 14명



불안증:  
증상없음 296명  
정상범주 274명  
경증 123명  
중증 64명



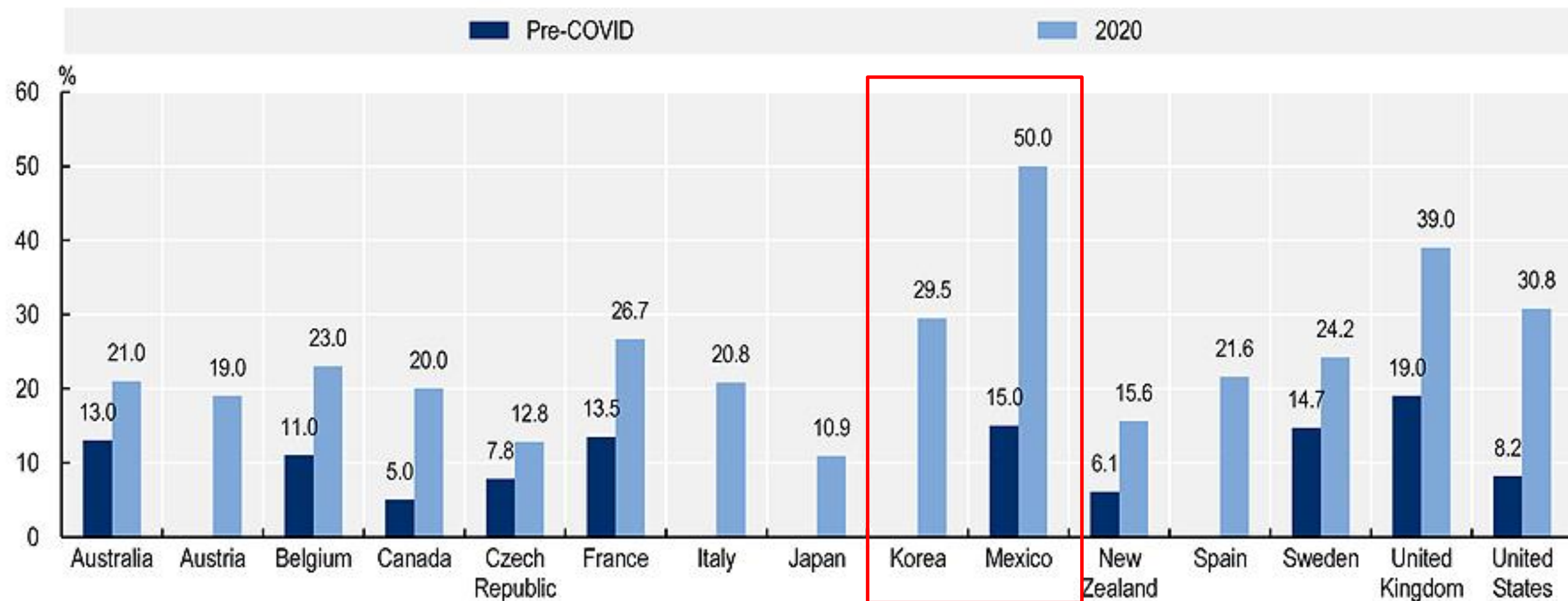


”

주제 선정 이유

02

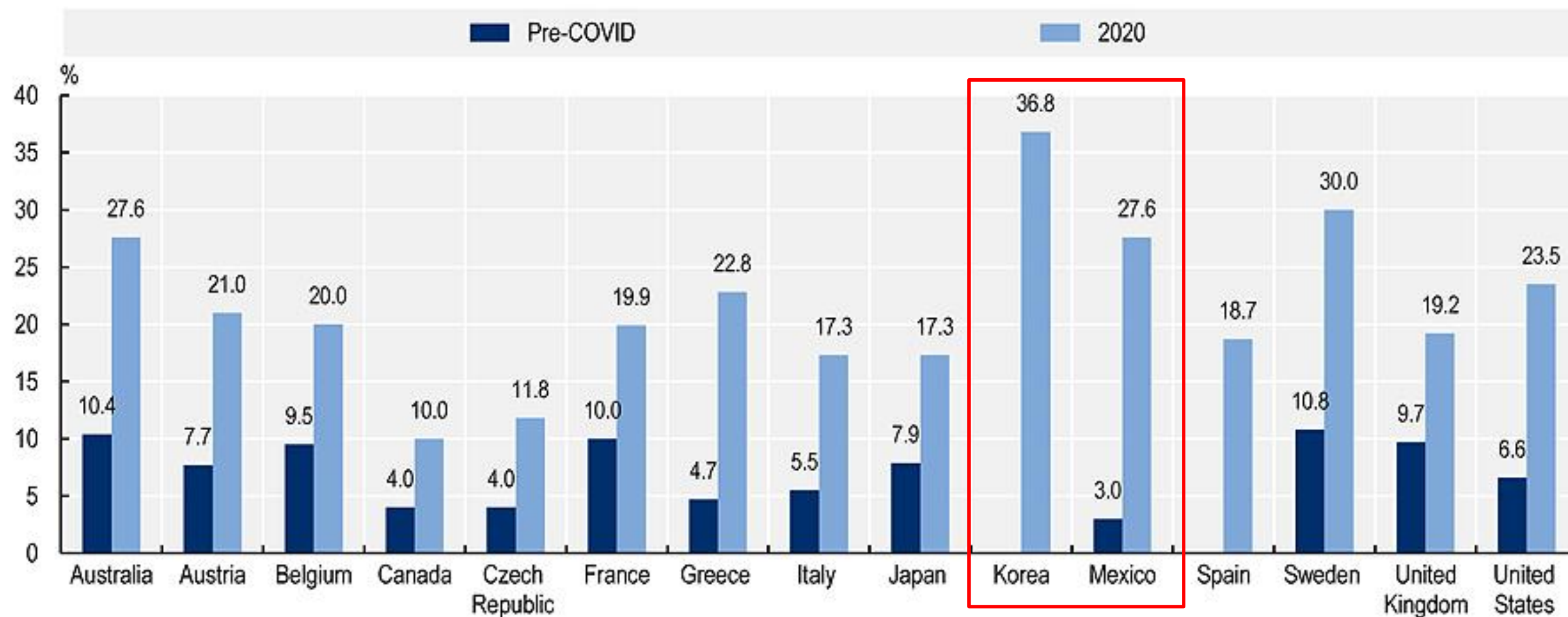
## “ 2.주제 선정 이유



OECD 회원국 불안증상 유병률 2020



## “ 2.주제 선정 이유



OECD 회원국 우울증상 유병률 2020



Q. 불안증이 있는 사람은 **Yes..?**  
반드시 우울증도 있을까요? **No..?**



**YES**

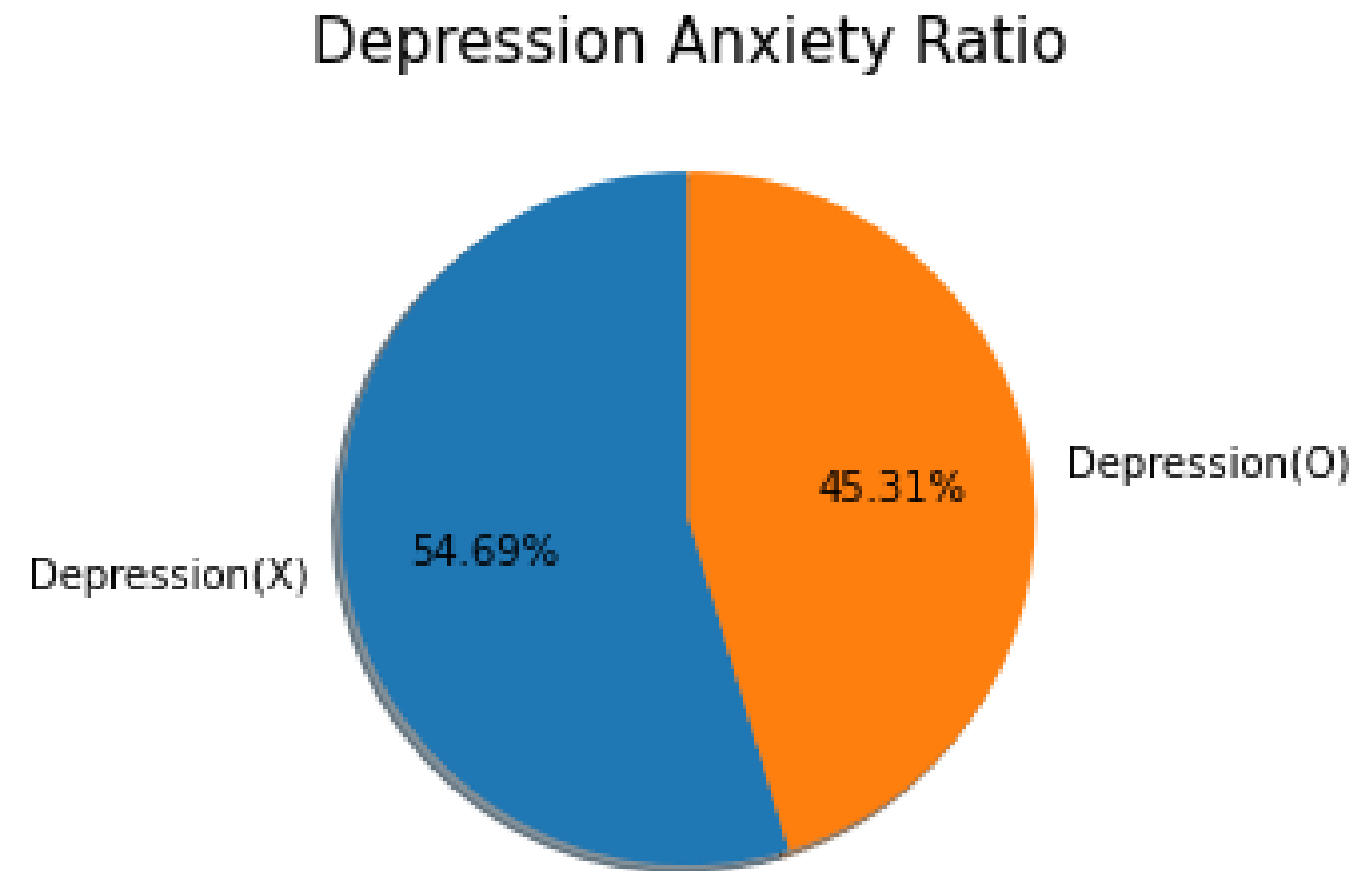


**NO**

## “ 2.주제 선정 이유

불안증 설문에 응답한 765명 中  
불안증이 있는 사람은 64명(12%)

불안증이 있는 사람 中  
우울증이 있는 사람은 29명(45%)





## “ 2.주제 선정 이유

어떤 특성을 가진 사람들이  
우울증을 가지고 있을까..?



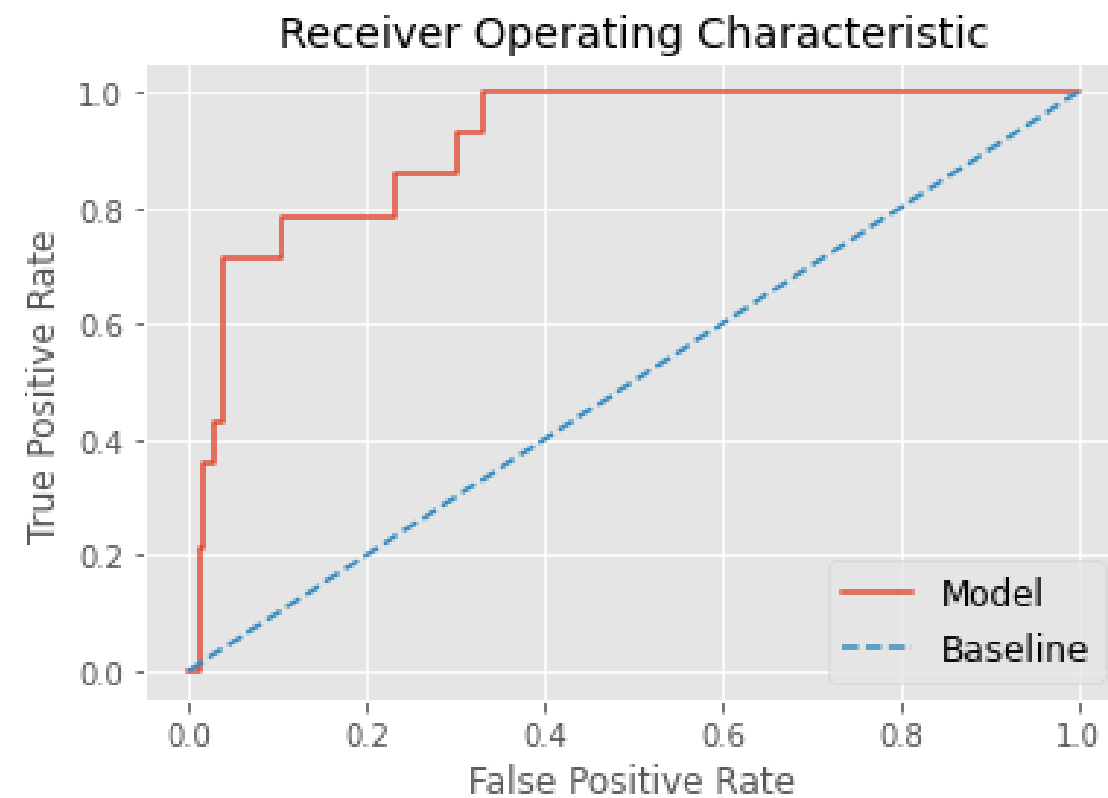
”

데이터 분석 결과

03

## “ 3.데이터 분석 결과

모델의 신뢰



Val AUC: 0.913  
f1 score : 0.625  
accuracy\_score : 0.939

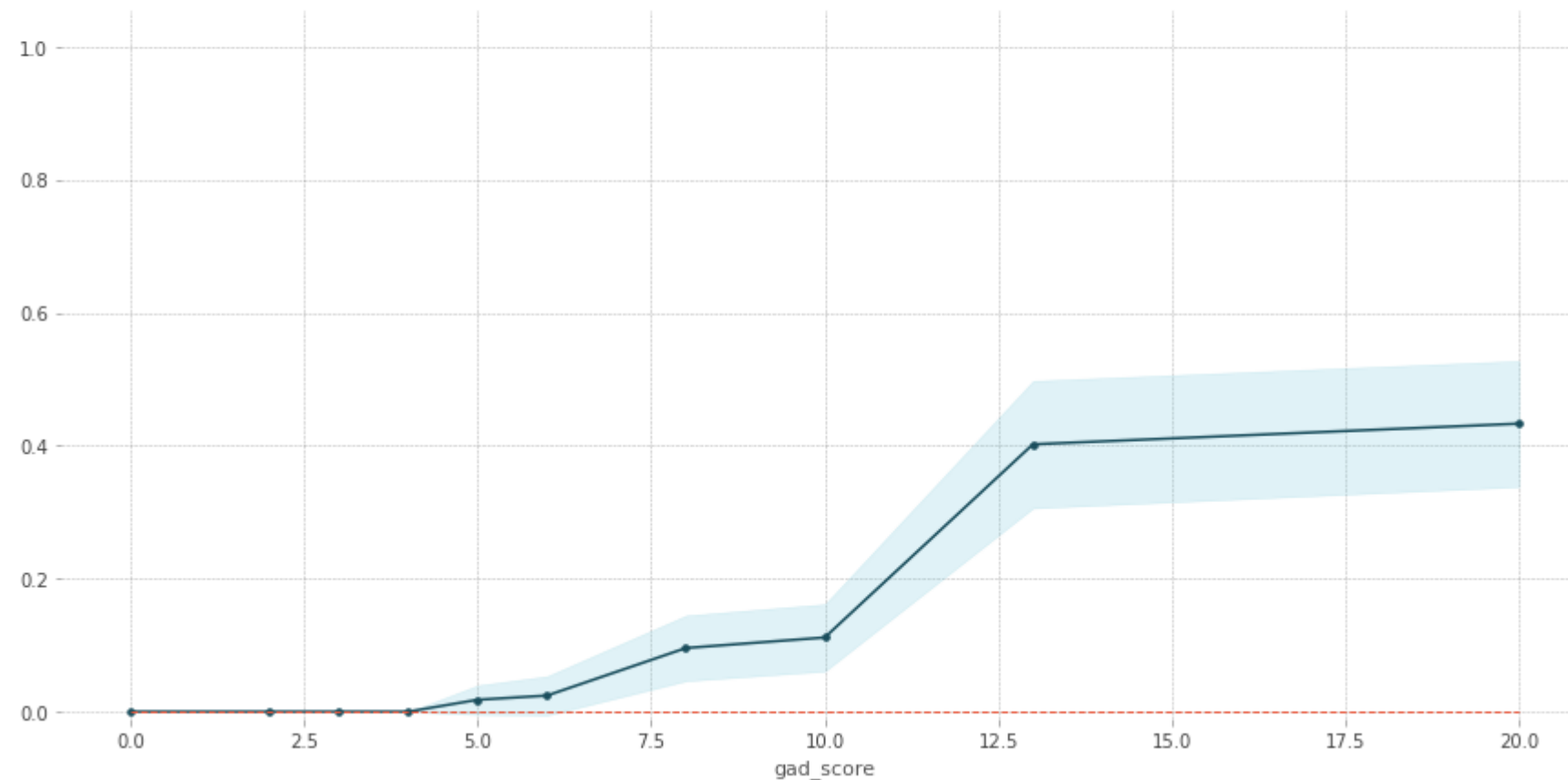
accuracy\_score > baseline accuracy score  
0.939 > 0.931

이 모델은 신뢰 할 수 있다!

# 3.데이터 분석 결과

## 불안증과 우울증의 관계

PDP for feature "gad\_score"  
Number of unique grid points: 10

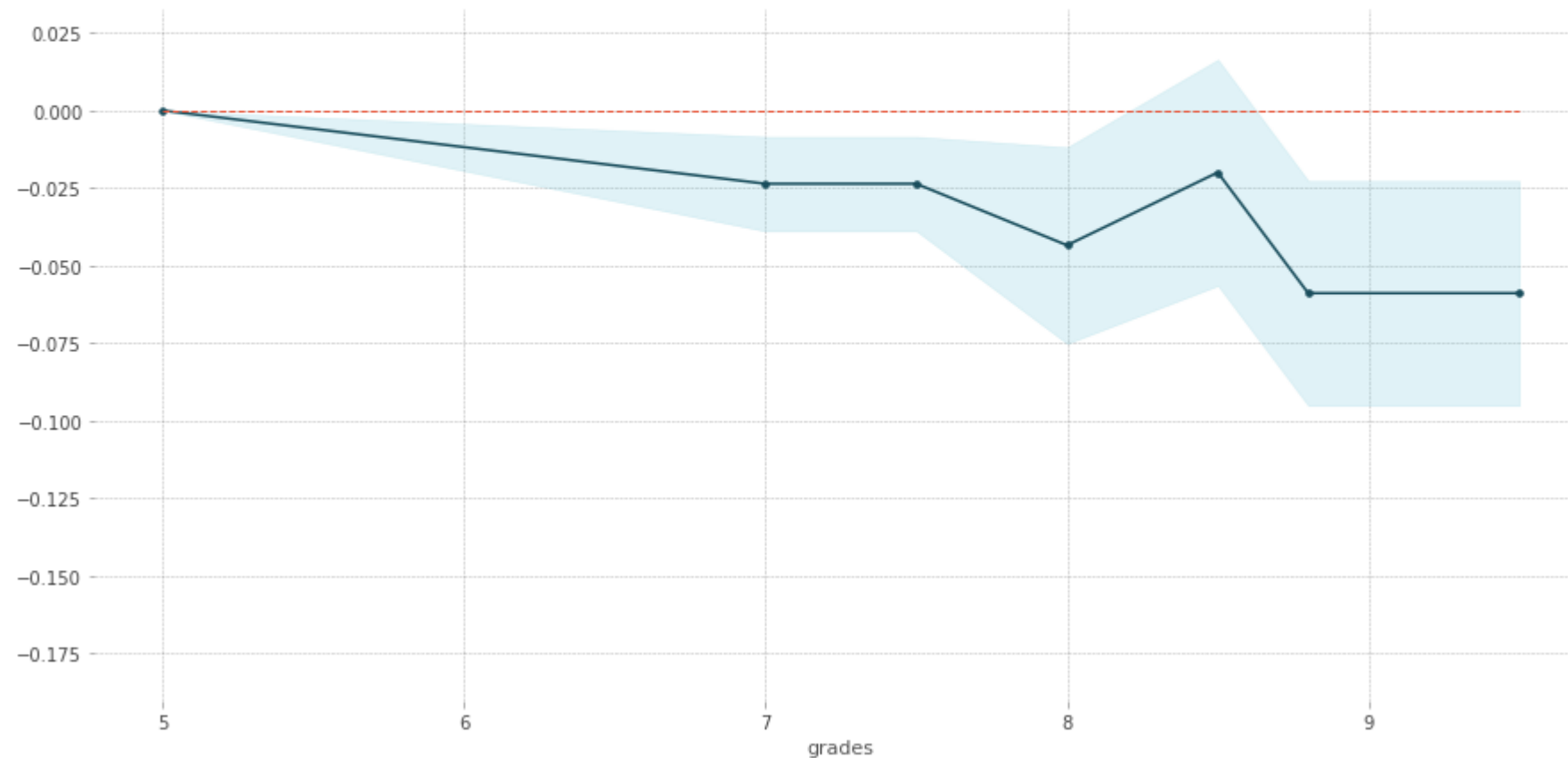


불안점수와 우울증은  
정적 상관을 보인다.

# 3.데이터 분석 결과

## 성적과 우울증의 관계

PDP for feature "grades"  
Number of unique grid points: 7



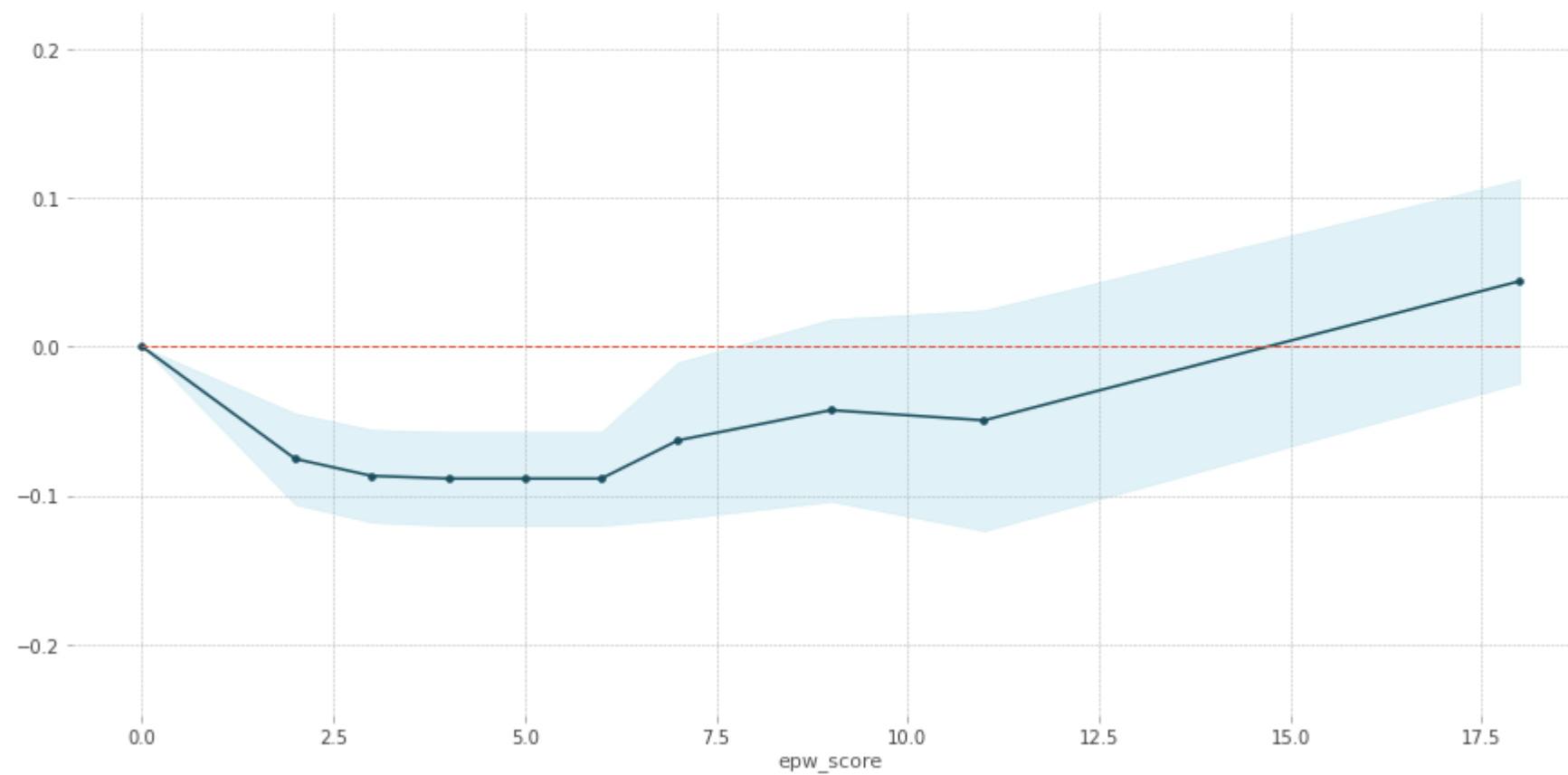
성적과 우울증은  
부적 상관을 보인다.



# 3.데이터 분석 결과

## 졸음증과 우울증의 관계

PDP for feature "epw\_score"  
Number of unique grid points: 10



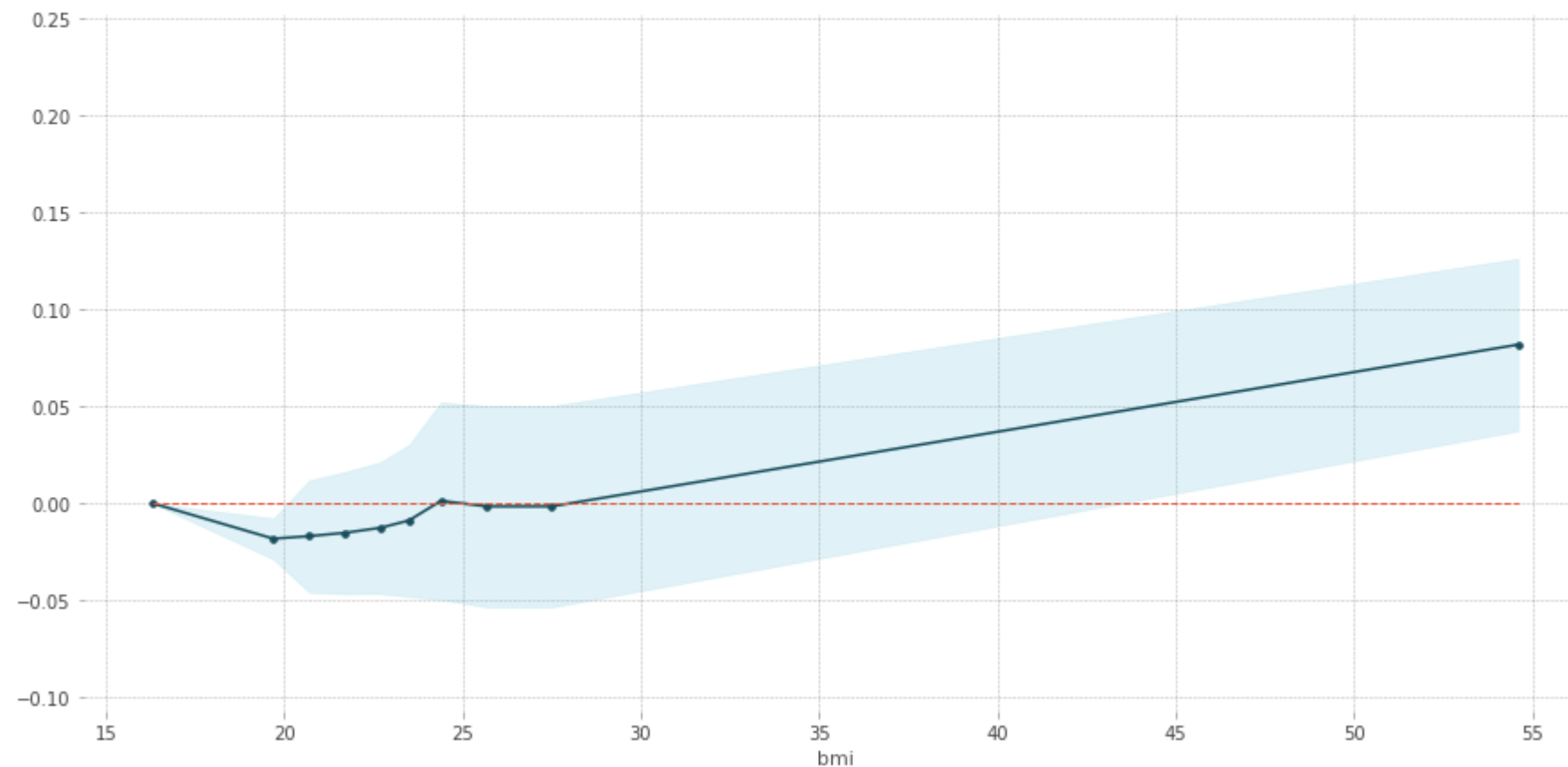
졸음증 중증 점수인  
15점 이상부터 우울증과  
정적 상관을 보인다.

## 3.데이터 분석 결과

### bmi점수와 우울증의 관계

PDP for feature "bmi"

Number of unique grid points: 10



과체중은 우울증과  
정적 상관을 보인다.

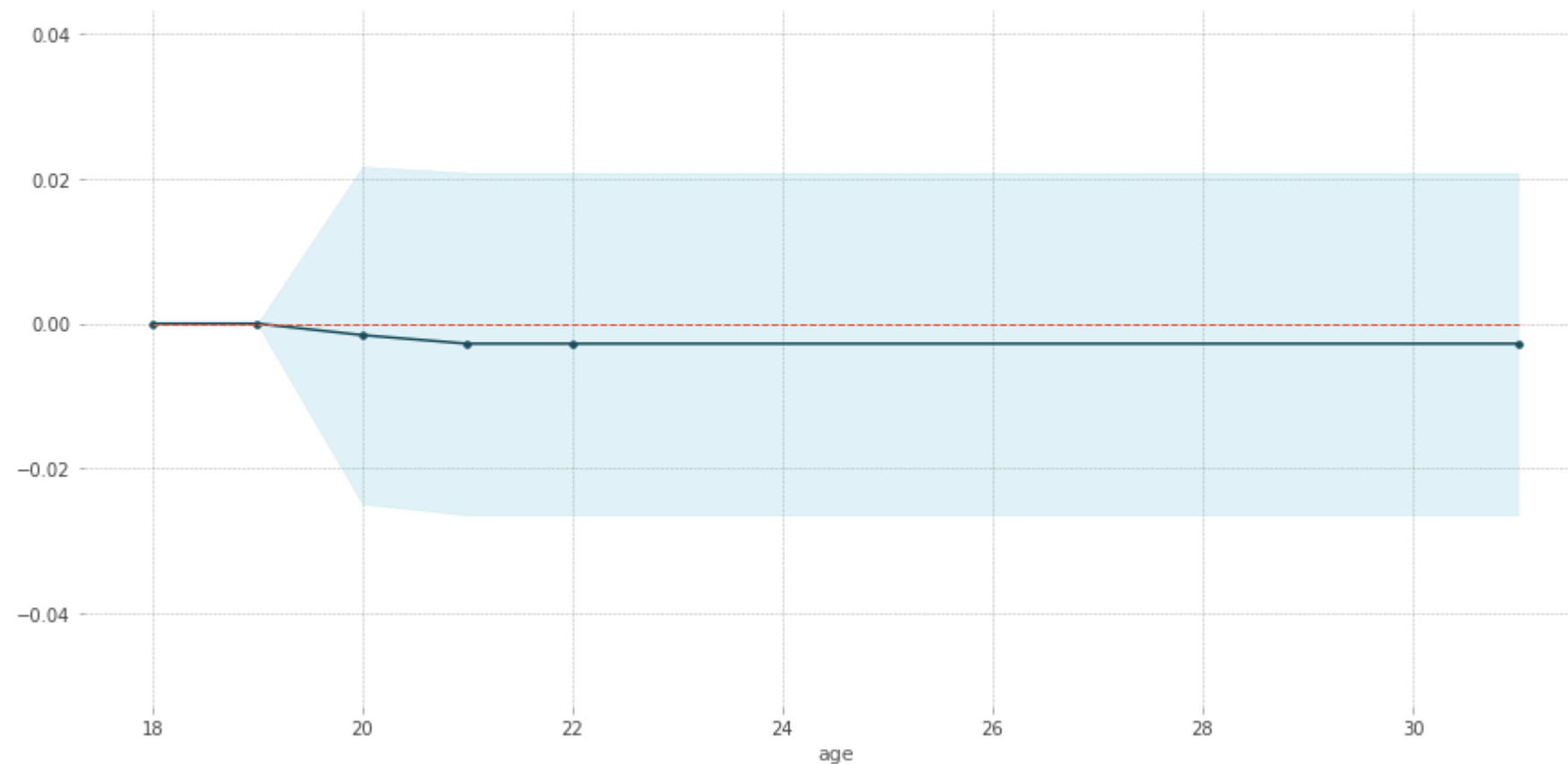
저체중 18.5 미만  
정상 18.5 - 24.9  
과체중 25 - 29.9  
비만 30 이상

멕시코 bmi 기준 : <https://www.bariatricmexico.com/bmi-calculator.html>

# 3.데이터 분석 결과

## 연령과 우울증의 관계

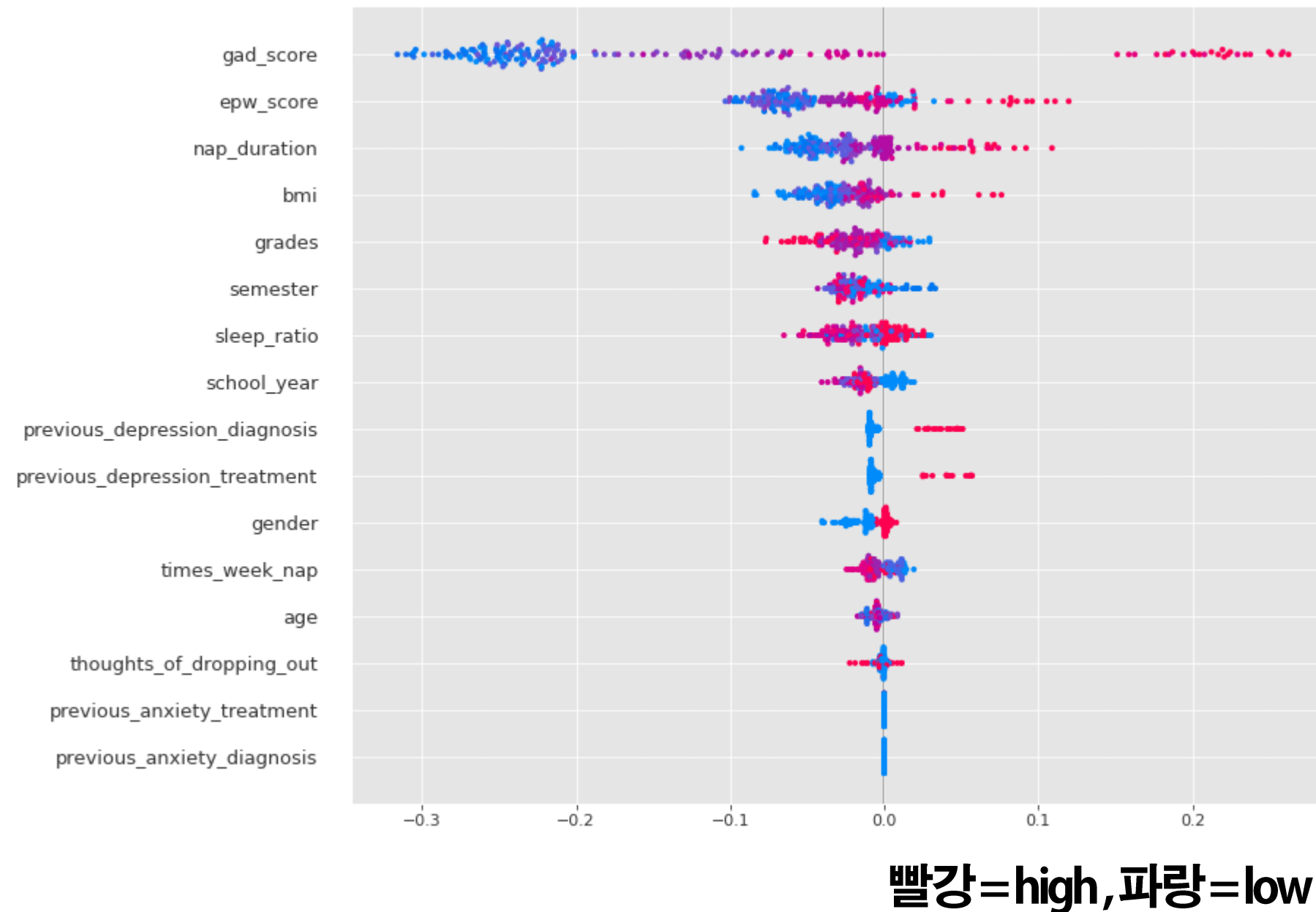
PDP for feature "age"  
Number of unique grid points: 6



본 데이터에서 연령은 우울증과  
**유의한 상관성이 없는 것으로**  
확인했다.

## 3.데이터 분석 결과

전체 특성들이 우울증을 분류하는데 어떤 영향을 미쳤는가?

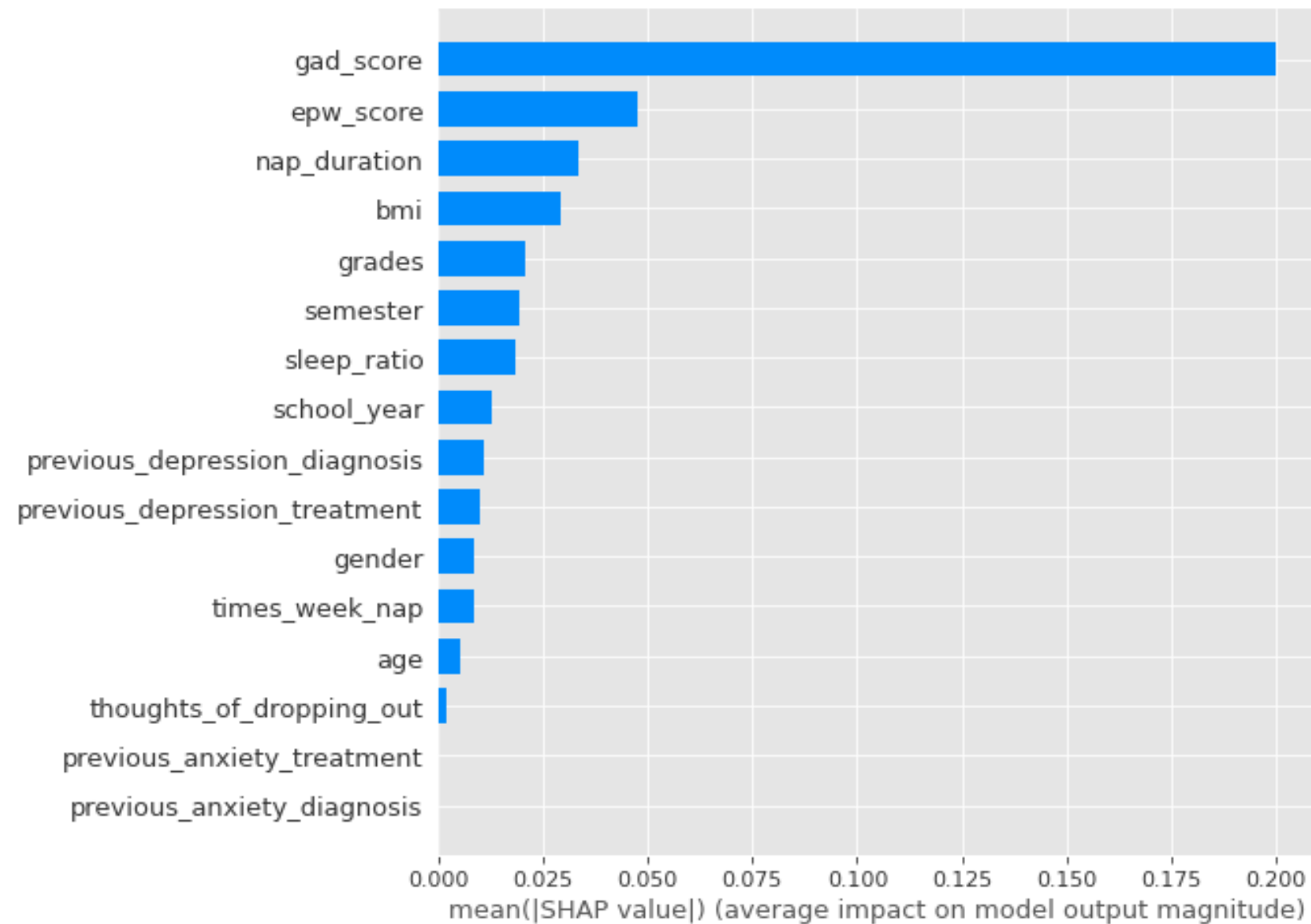


불안증, 졸음증, 낮잠시간, bmi는  
우울증과 **정적상관**

학점, 학기, 학년은  
우울증과 **부적상관**

## “ 3.데이터 분석 결과

각특성이모델에미치는영향도



우울증을 분류하는데  
가장 영향을 많이 준 것은  
불안증 점수이다.

졸음증 점수, 낮잠 시간, bmi가  
다음으로 영향을 주었다.



”

결과 종합

04

## “ 4.결론

우울증	
불안증	정적 상관
졸음증	정적 상관
과체중	정적 상관
성적	부적 상관
연령	유의한 상관 없음

우울증을 분류하는데  
가장 많은 영향을 준 것은  
불안증이다.



”

한계점

05





## “ 5.한계점

상관관계는 확인할 수 있었지만,  
인과관계는 확인할 수 없다.

인과관계를 확인하기 위해서는  
후속 연구가 필요할 것 같다.



# 감사합니다.

”

[https://github.com/DavidPark96/bootcamp\\_project](https://github.com/DavidPark96/bootcamp_project)

**발표 끝!**