

# **Big\_Contest 2022**

# 목차

## 1. 작업과정

- 1-1. 프로젝트 배경
- 1-2. 프로젝트 목표
- 1-3. 진행과정
- 1-4. 기술동향분석

## 2. 결과물

- 2-1. 분석 데이터 시각화
- 2-2. 상품 분석

# 1. 작업과정

The background features a series of concentric, wavy lines in shades of blue and green, creating a sense of depth and movement. The lines are more pronounced on the left side and fade towards the right.

프로젝트 배경



## 프로젝트 배경

요즘 대세 기술인 Machine Learning을 활용한 주제

공모전이라는 확실한 동기부여로 프로젝트의 효율성의 극대화  
공모전을 통해 빅데이터를 경험해볼 수 있는 기회

# 프로젝트 목표



# 프로젝트 목표

비대면 대출앱인 'Finda' 어플의  
실제 사용자 데이터를 통한  
대출신청 예측 분석



# 제공된 데이터

2022년 3~5월 데이터 제공

2022년 6월 예측이 목표

(개인정보는 가명화)

	user_id	bank_id	product_id	loan
0	154546	33	189	90
1	275138	31	224	310
2	834243	22	164	150
3	549899	13	262	700
4	273510	19	127	110
...	...	...	...	...
549660	778686	54	235	410
549661	442655	25	163	350

	credit_score	...	event3	event4	event5	event6
0	480.0	...	13.0	17.0	20.0	14.0
0	640.0	...	0.0	12.0	18.0	9.0
0	640.0	...	70.0	21.0	22.0	16.0
0	770.0	...	36.0	9.0	8.0	7.0
0	780.0	...	3.0	8.0	9.0	6.0
...	...	...	...	...	...	...
0	690.0	...	8.0	7.0	5.0	6.0
0	790.0	...	52.0	26.0	51.0	17.0
0	780.0	...	9.0	7.0	8.0	2.0





# 진행 과정

---

Data Cleaning & Labeling



Modeling



Cross Validation



Data visualization



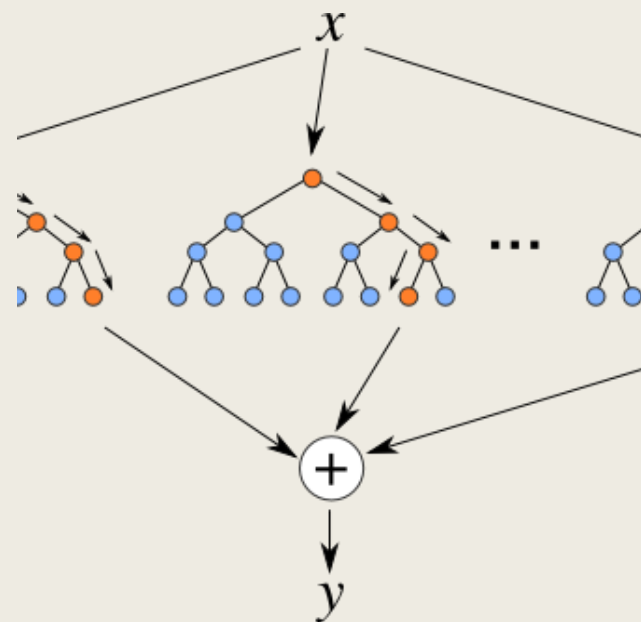
product planning

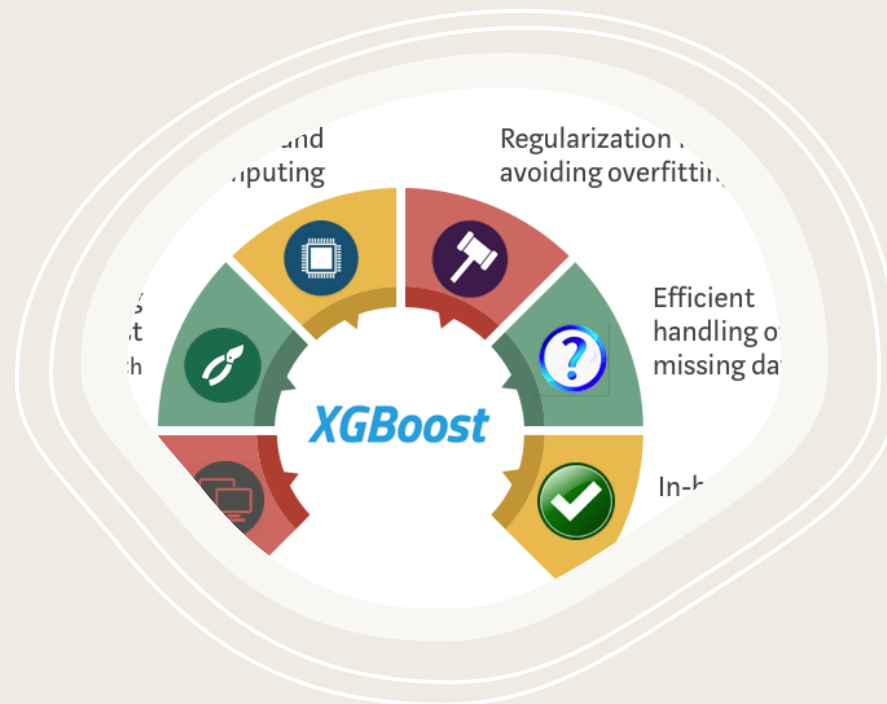


# 기술 동향 분석

## Random Forest

여러 개의 decision tree 형성 후 각 트리가 분류한 결과에서 투표를 실시, 가장 많이 득표한 결과를 최종 분류 결과로 선택합니다. 정확도면에서 LightGBM와 비슷했으나 Cross Validation 결과 다른 모델에 비해 성능이 떨어짐을 확인했습니다.





## 기술동향분석

# XGBoost

Gradient Boost 알고리즘을 병렬 학습이 지원되도록 구현한 것으로 성능이 뛰어나지만 모델 학습 시간이 오래 걸리며 높은 컴퓨팅 성능을 요구합니다.



## 기술동향분석

# LightGBM

기존의 다른 Tree기반 알고리즘은 Tree 구조가 수평적으로 확장되는 것(Level-Wise)에 비해

LightGBM은 수직적으로 확장 (Leaf-Wise)합니다.

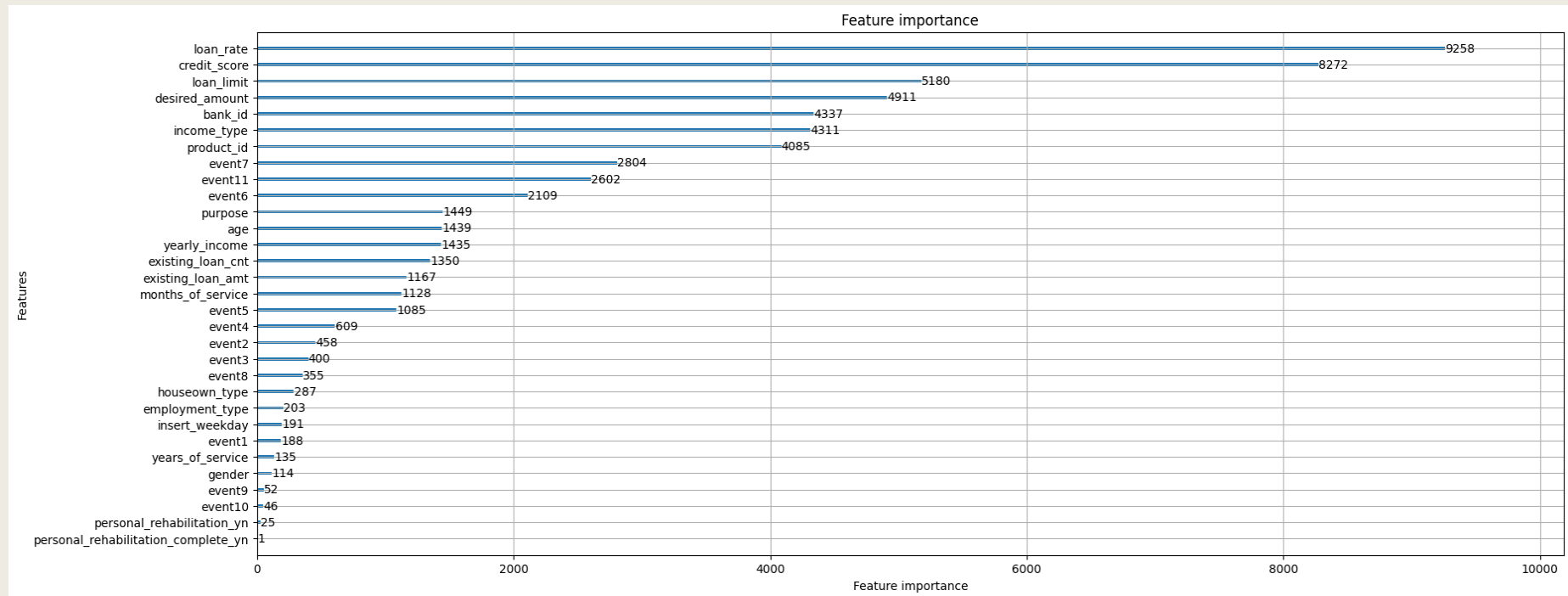
Random Forest 보다 높은 성능, XGBoost 보다 빠른 학습 및 분석이 가능해 선택했습니다.

## 2. 결과물

An abstract illustration featuring a dark blue circular background. In the center, a magnifying glass is held by a hand, focusing on a pie chart. To the left, there's a blue speech bubble and a document icon. Below the speech bubble is a blue box containing a line graph with orange and light blue areas. At the bottom, there's a bar chart with blue bars and a line graph with grey dots. The text '분석' (Analysis) and '데이터 시각화' (Data Visualization) is written in white Korean characters over the central part of the image.

# 분석 데이터 시각화

# Feature importance 기반 분석



- 상위권에 있는 Feature를 위주로 분석 및 시각화 하였습니다.



# 데이터 시각화

---

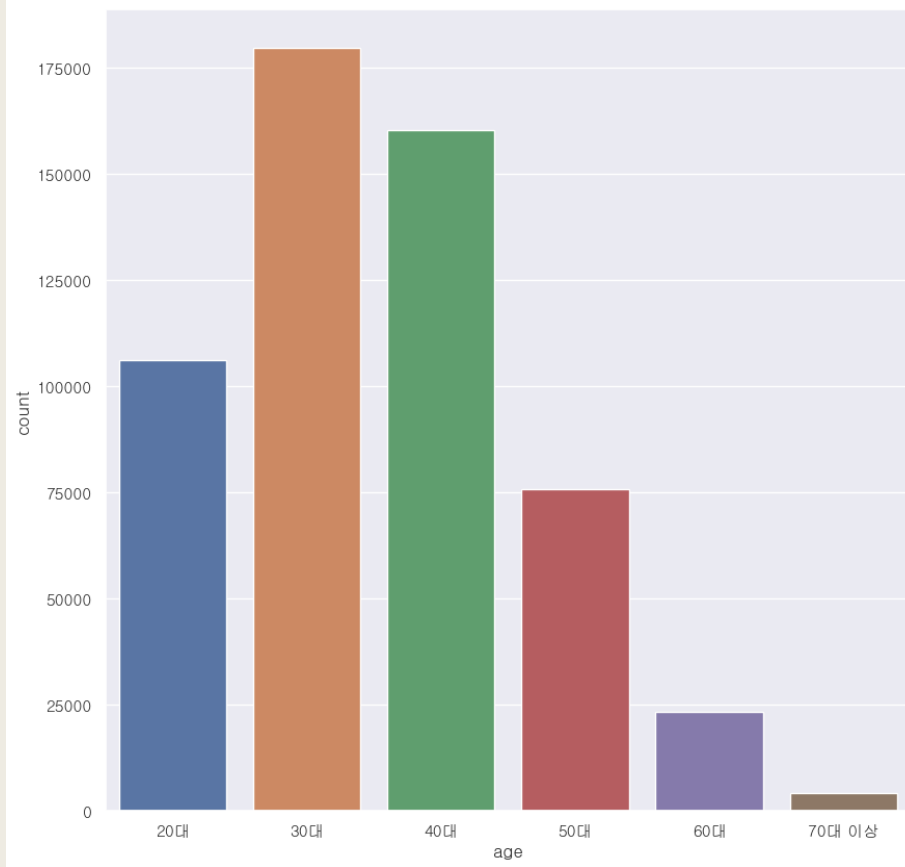
- 신청자 특성 분석
- 상품 분석
- 로그 데이터 분석



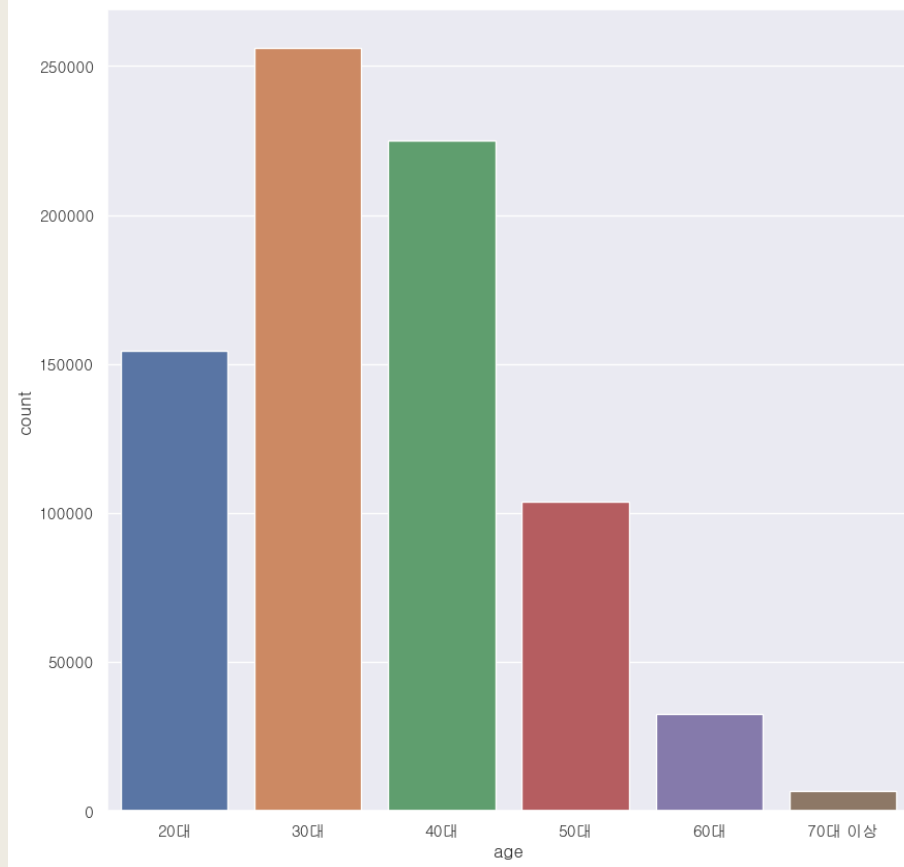
# 신청자 특성 분석

Seaborn 라이브러리를 활용한

# 신청자 특성 분석 (나이)



3~5월 실제 데이터



6월 예측 데이터

- 20 ~ 40대가 주 고객층임을 알 수 있음.

# 신청자 특성 분석 (신용등급)

등급	신용평가회사	신용자구분
1등급	942~1000점	우량 신용자
2등급	891~ 941점	
3등급	832~890점	
4등급	768~831점	보통 신용자
5등급	698~767점	
6등급	630~697점	
7등급	530~629점	저신용자
8등급	454~529점	
9등급	335~453점	
10등급	0~334점	

‘올크레딧’기준 신용등급평가표



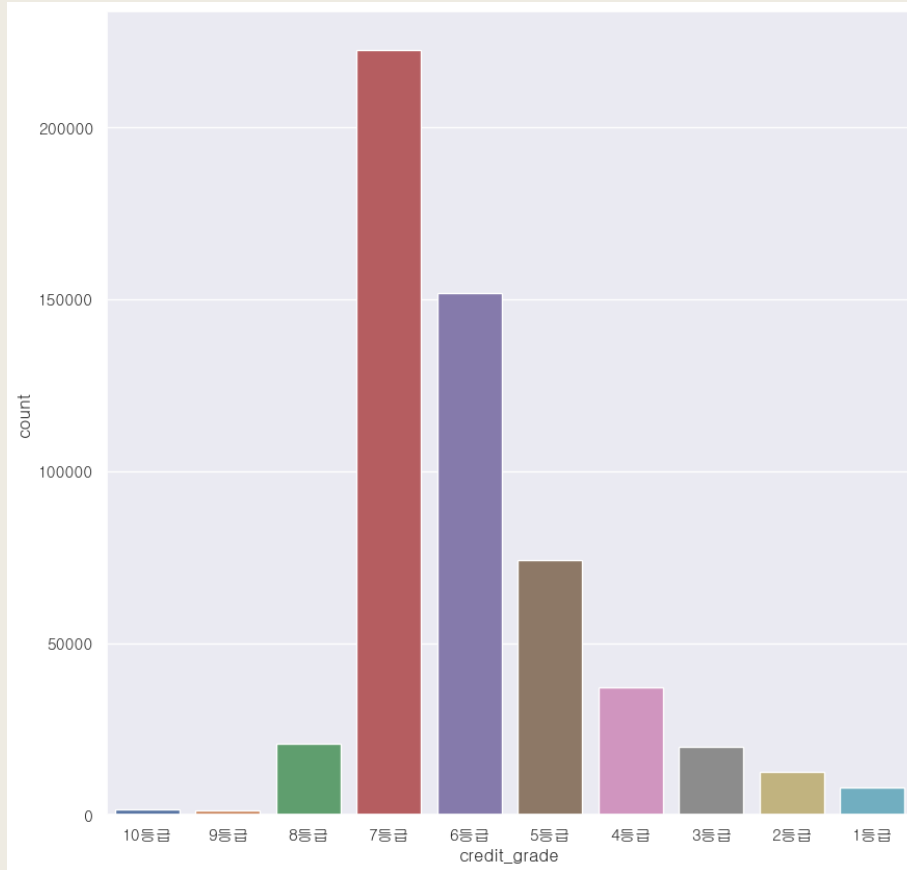
신용등급제도는

폐지된 제도이지만

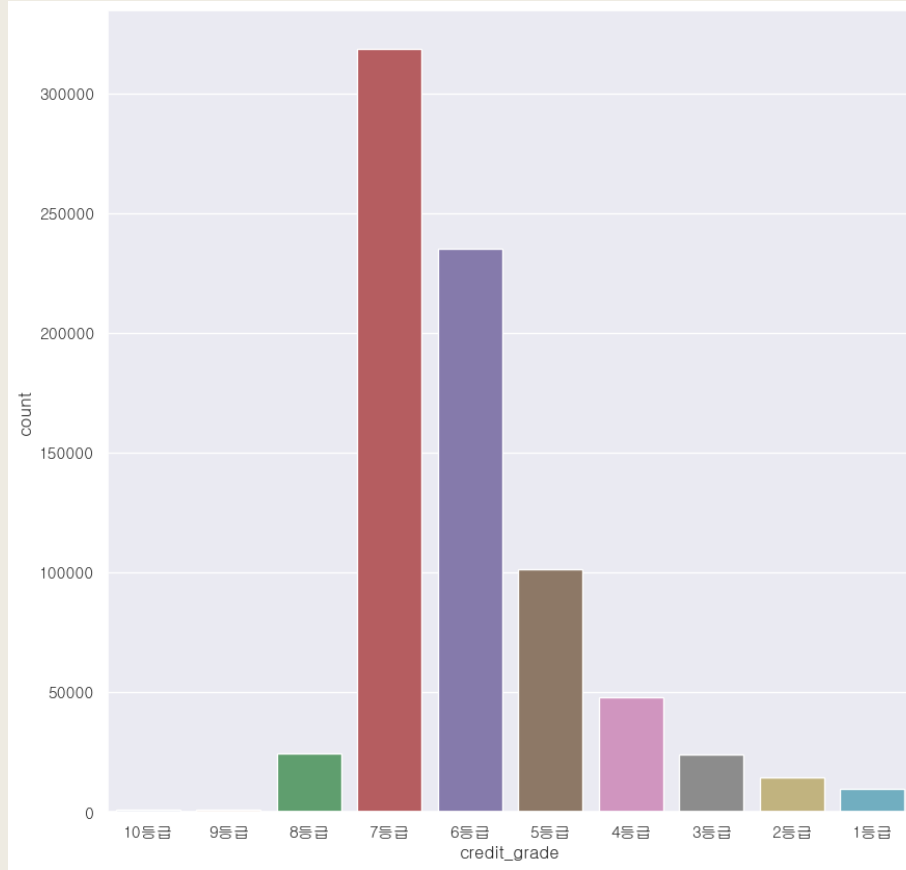
시각화 편의성을 위해

사용하였습니다.

# 신청자 특성 분석 (신용등급)



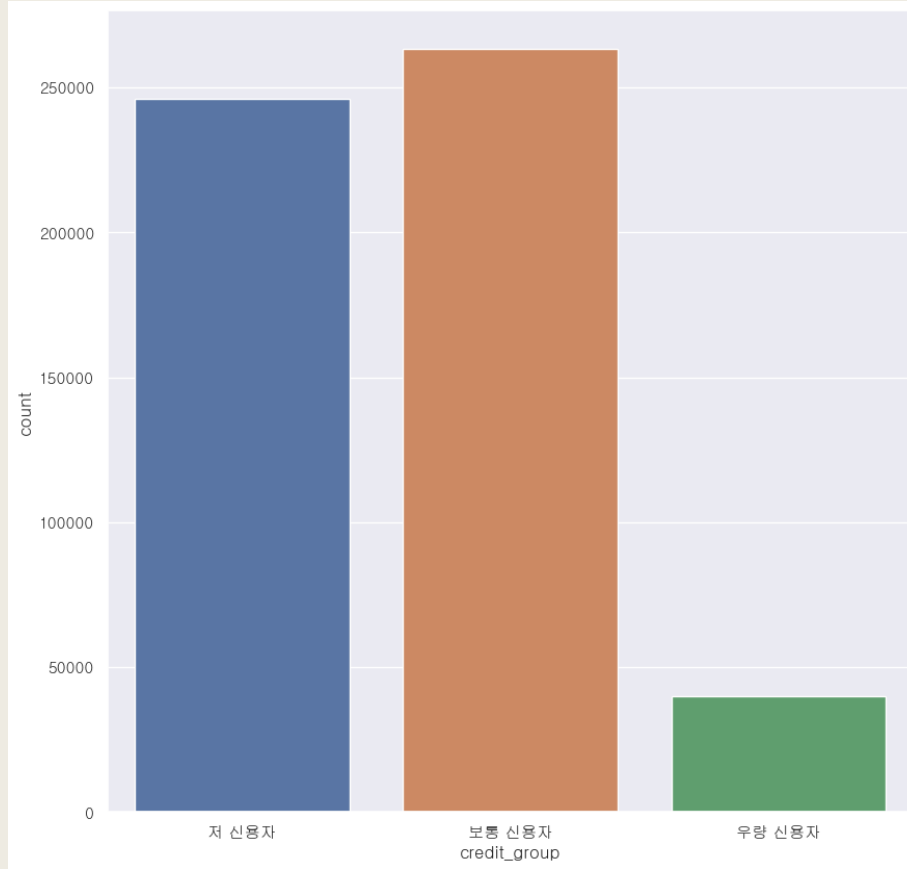
3~5월 실제 데이터



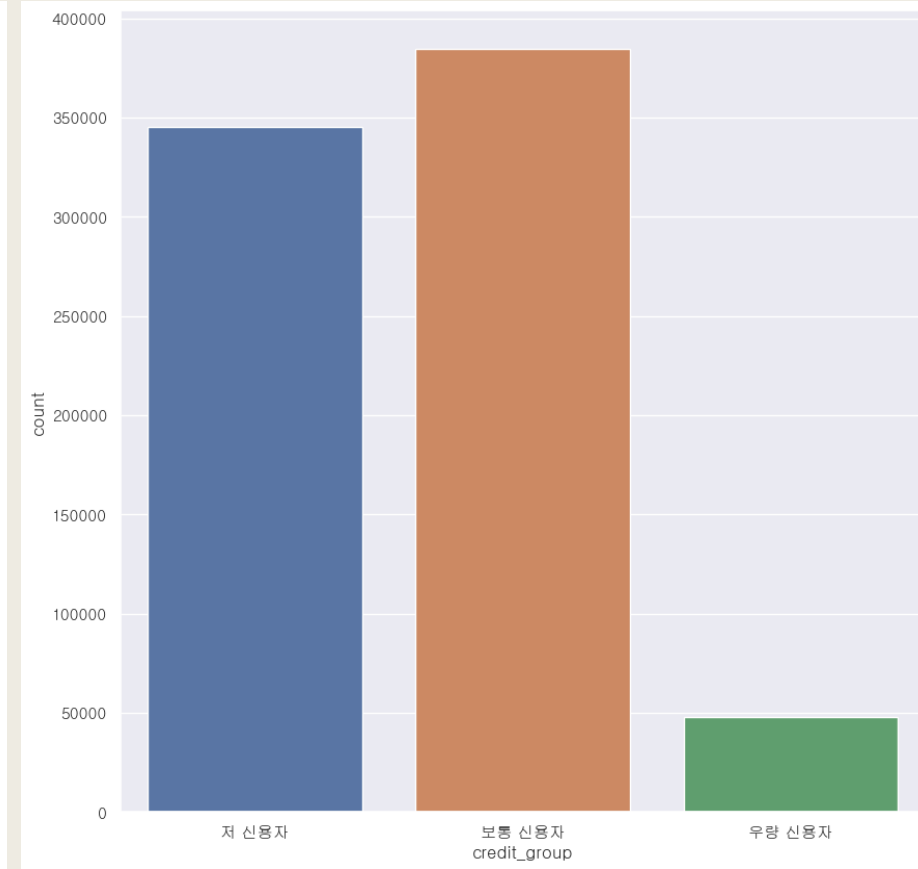
6월 예측 데이터

- 7등급에 해당하는 사람이 가장 많음.

# 신청자 특성 분석 (신용등급)



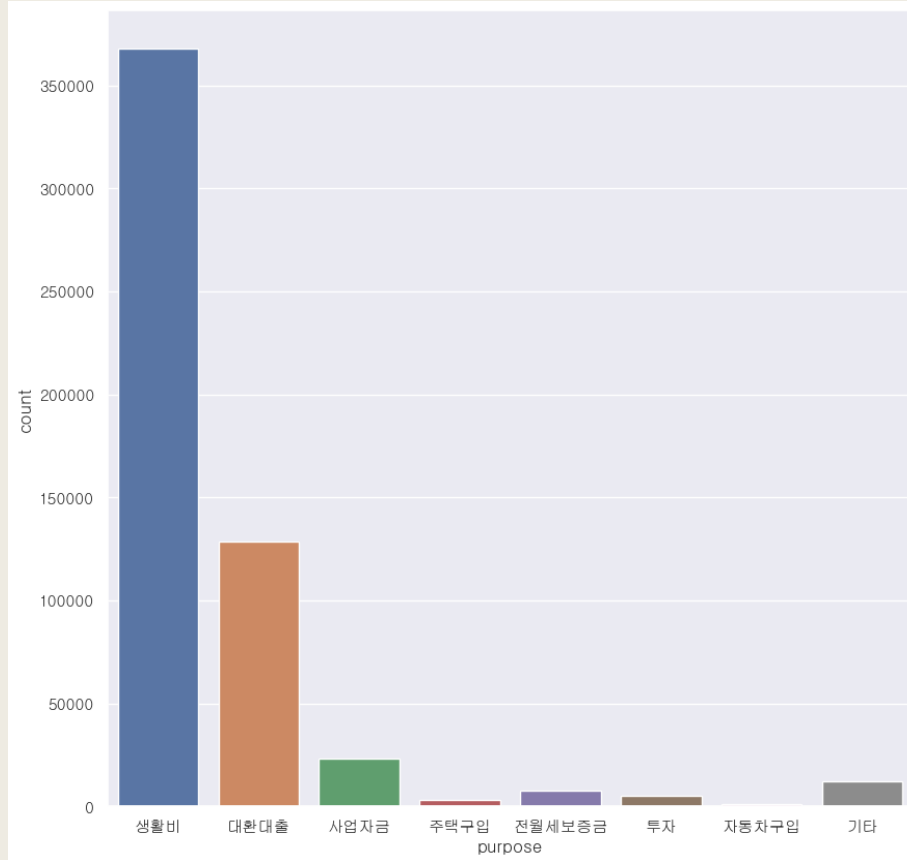
3~5월 실제 데이터



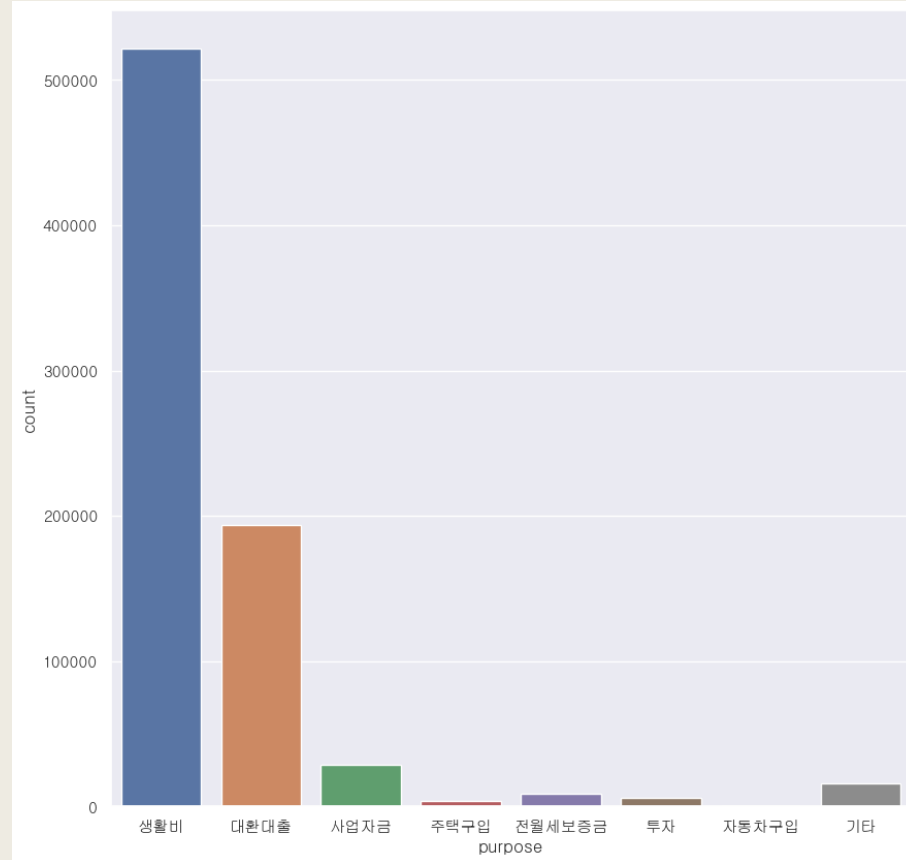
6월 예측 데이터

- 대부분의 사람들이 저 신용자와 보통 신용자에 속함.

# 신청자 특성 분석 (대출 목적)



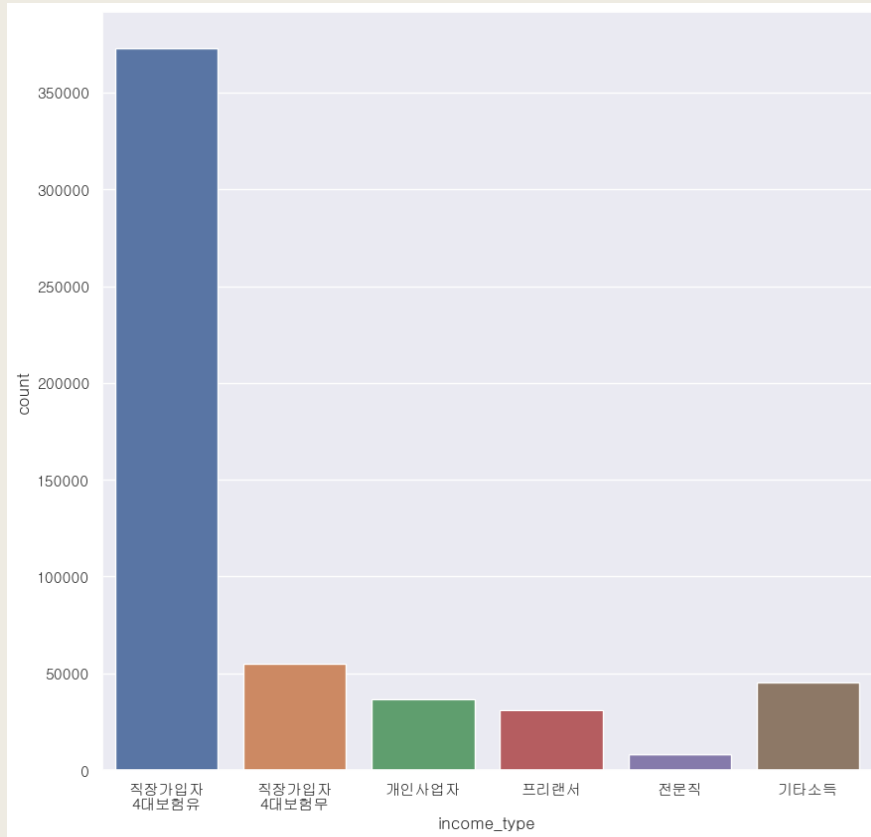
3~5월 실제 데이터



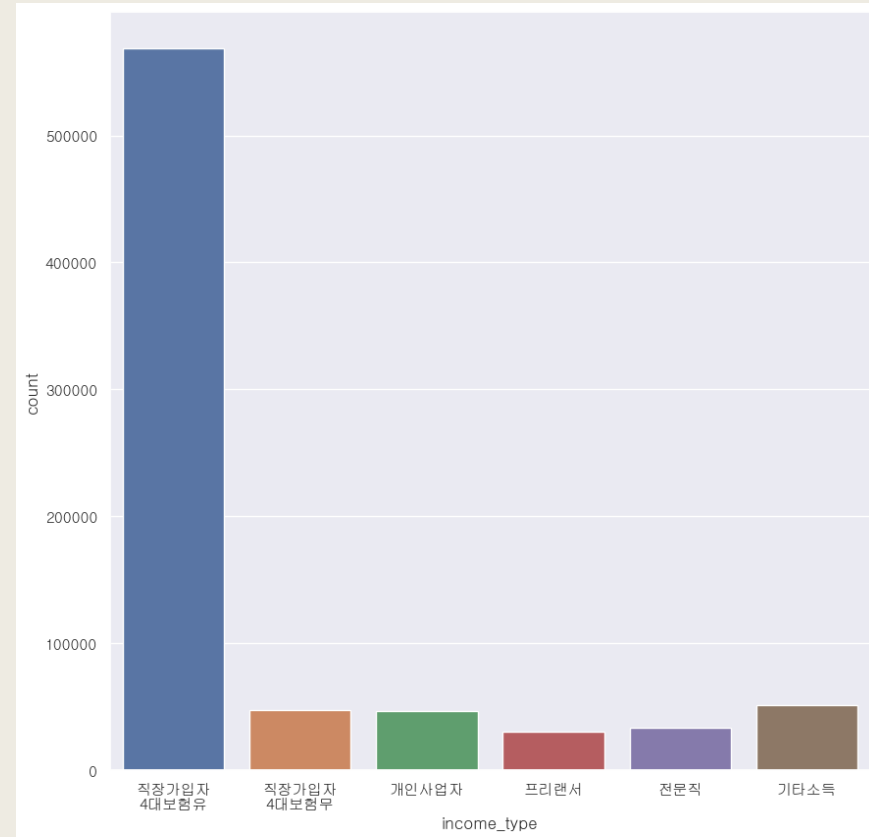
6월 예측 데이터

- 과반 이상을 생활비가 차지, 그 뒤로는 대환대출, 사업자금 등이 따름.

# 신청자 특성 분석 (근로 형태)



3~5월 실제 데이터

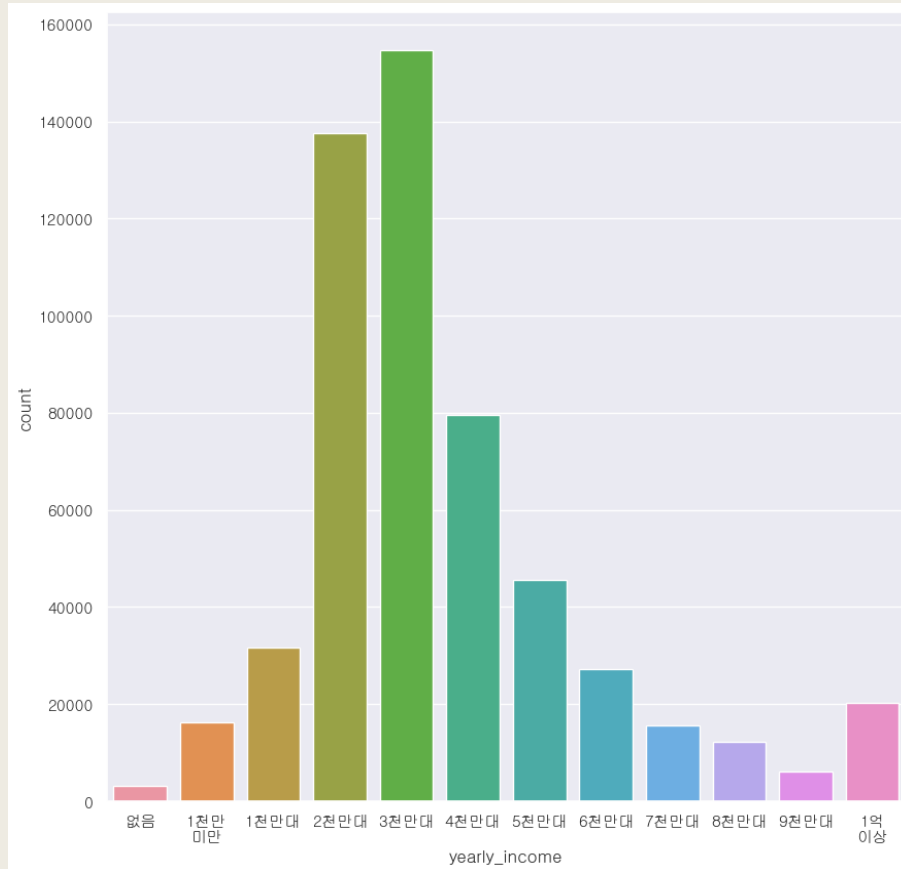


6월 예측 데이터

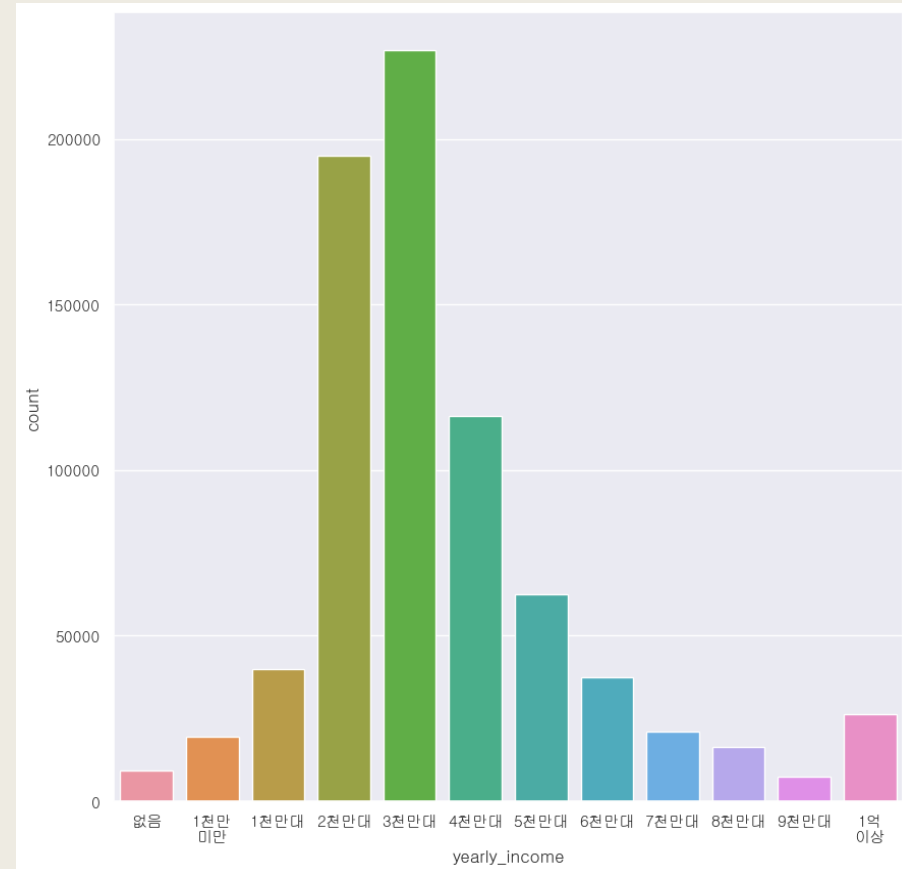
- 대부분이 4대보험에 적용되는 정규직임을 알 수 있음.



# 신청자 특성 분석 (연 소득)



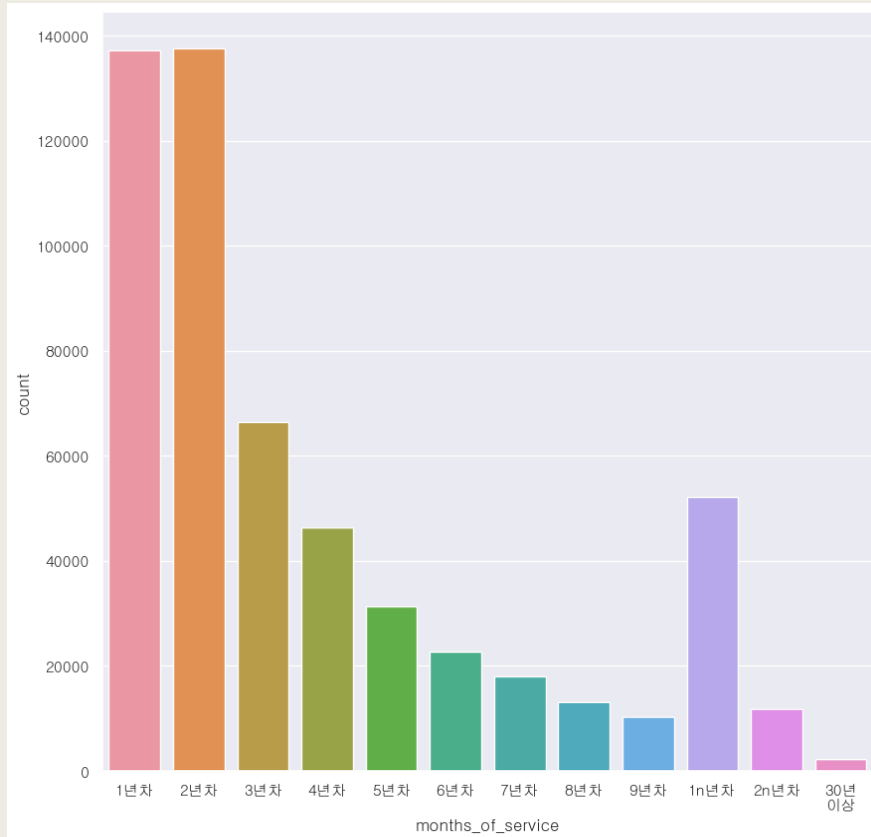
3~5월 실제 데이터



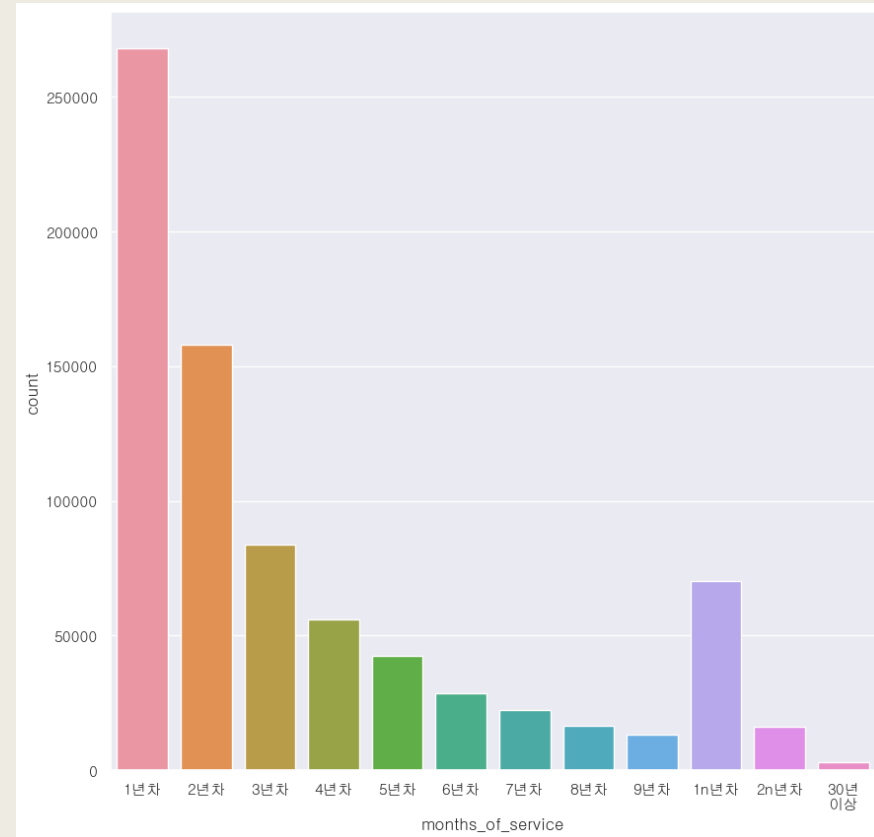
6월 예측 데이터

- 과반수 이상이 대한민국 평균연봉인 4327만원에 못 미침.

# 신청자 특성 분석 (근로 기간)



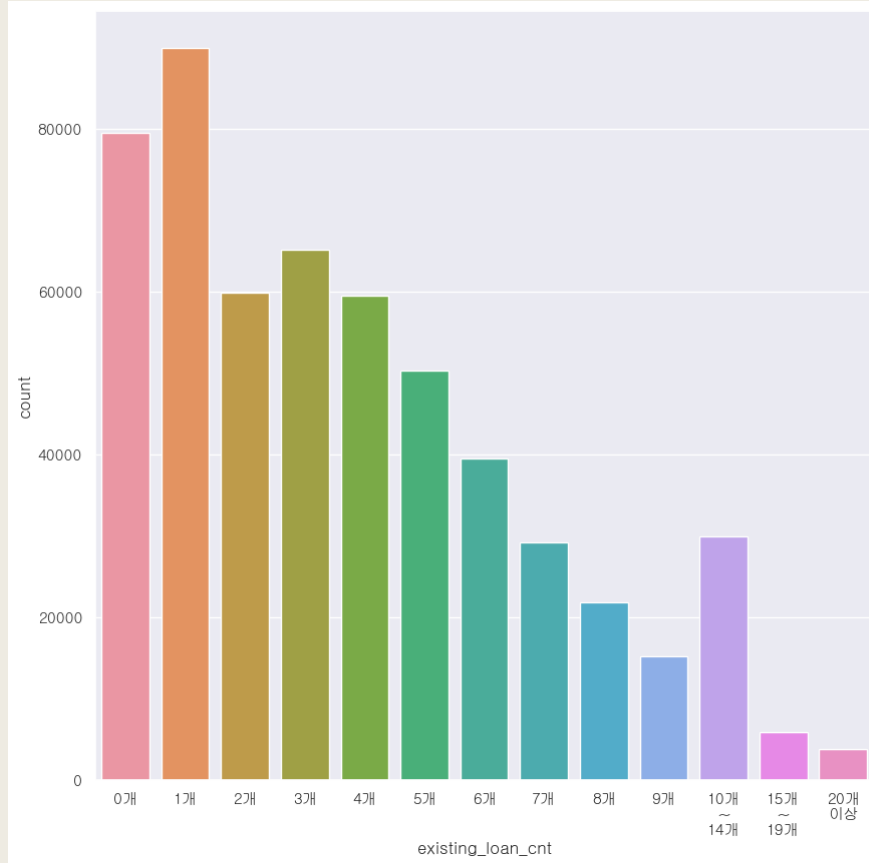
3~5월 실제 데이터



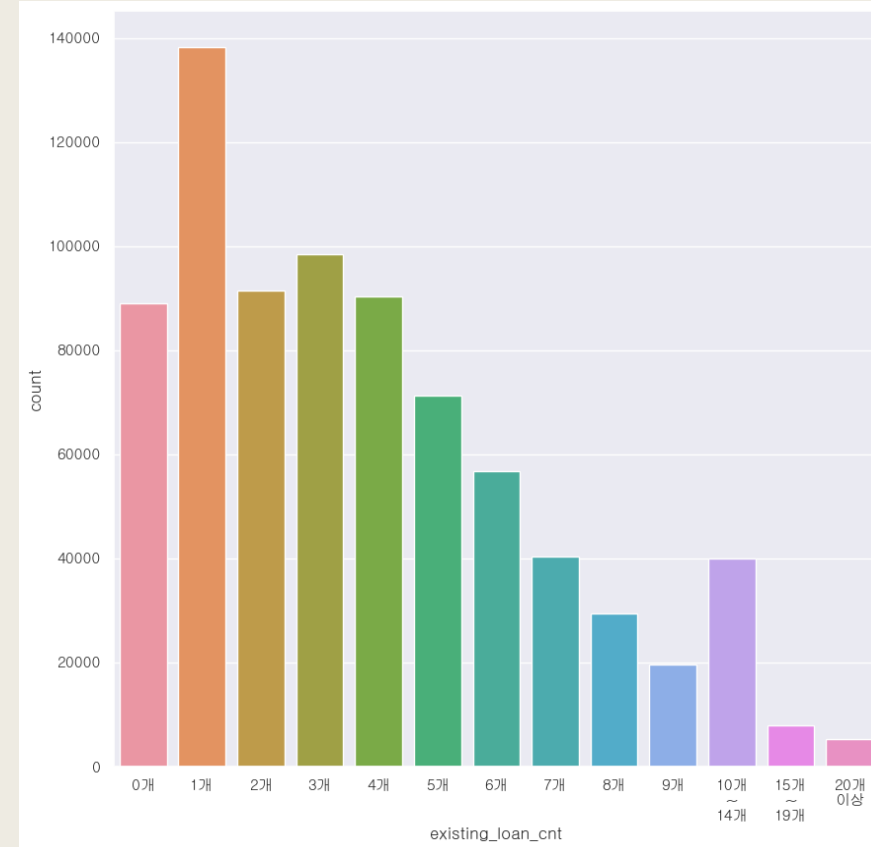
6월 예측 데이터

- 사회초년생이 대부분임을 알 수 있음.

# 신청자 특성 분석 (기대출수)



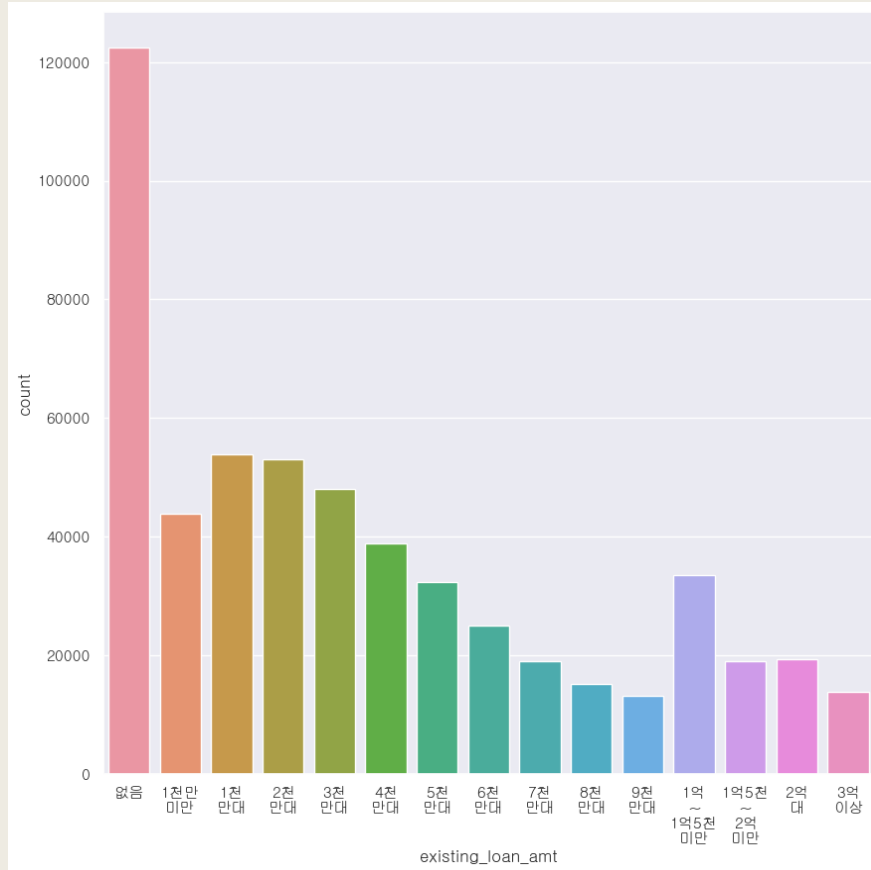
3~5월 실제 데이터



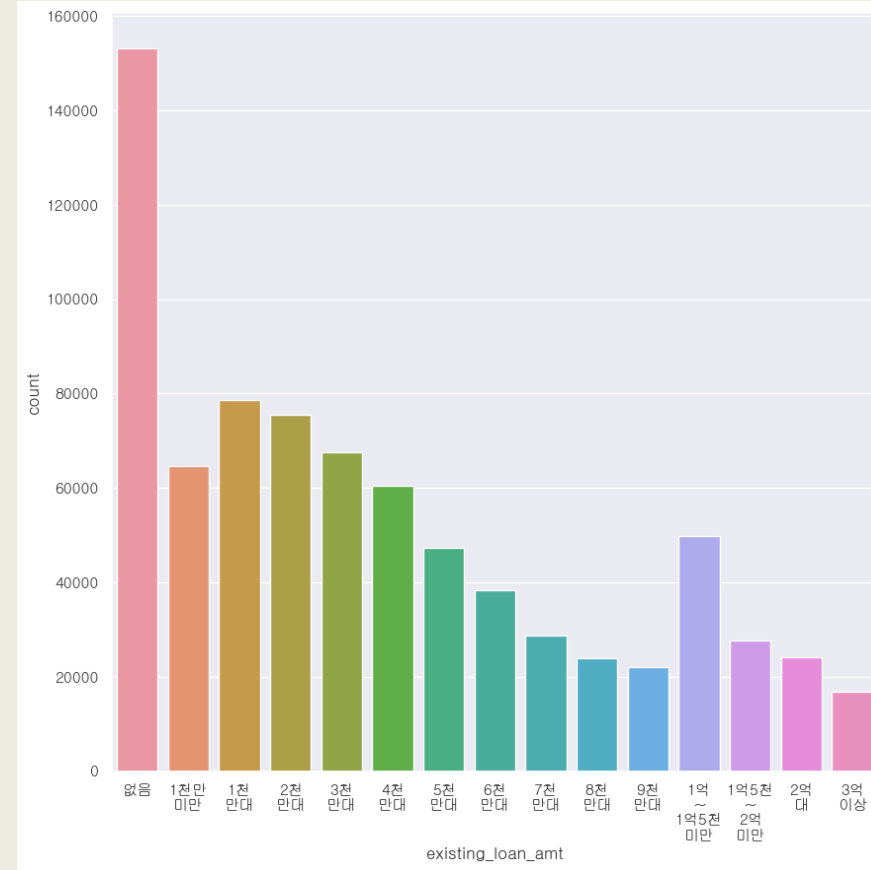
6월 예측 데이터

- 과반 이상이 대출금이 남아있음에도 새로운 대출을 신청함.

# 신청자 특성 분석 (기대출금액)



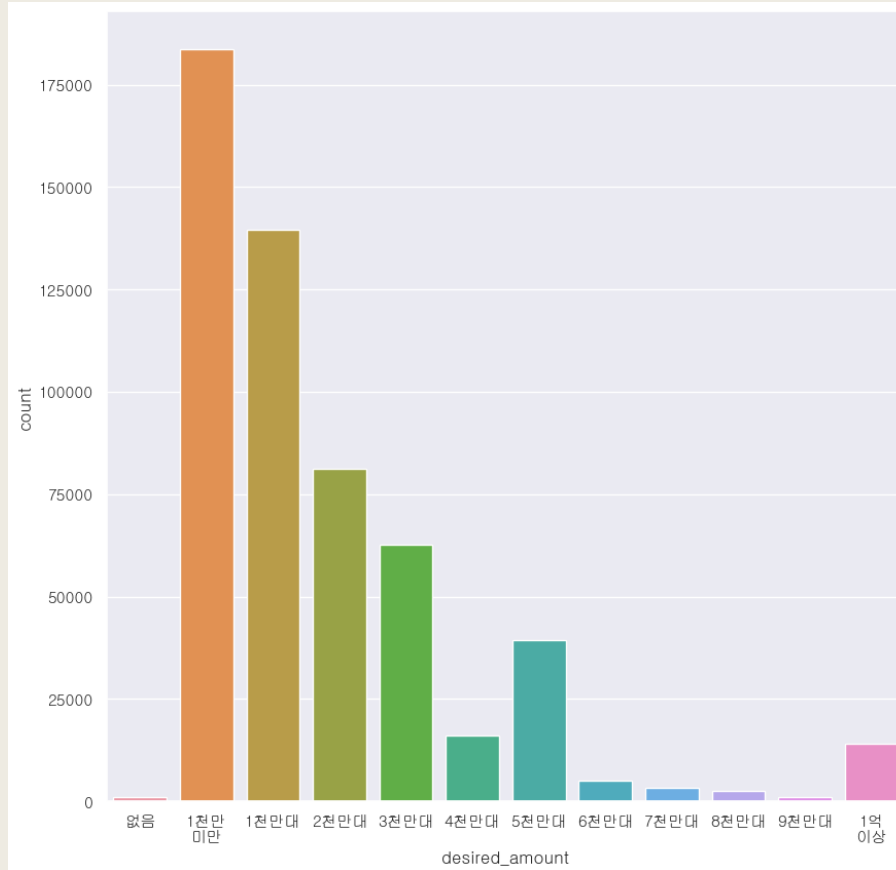
3~5월 실제 데이터



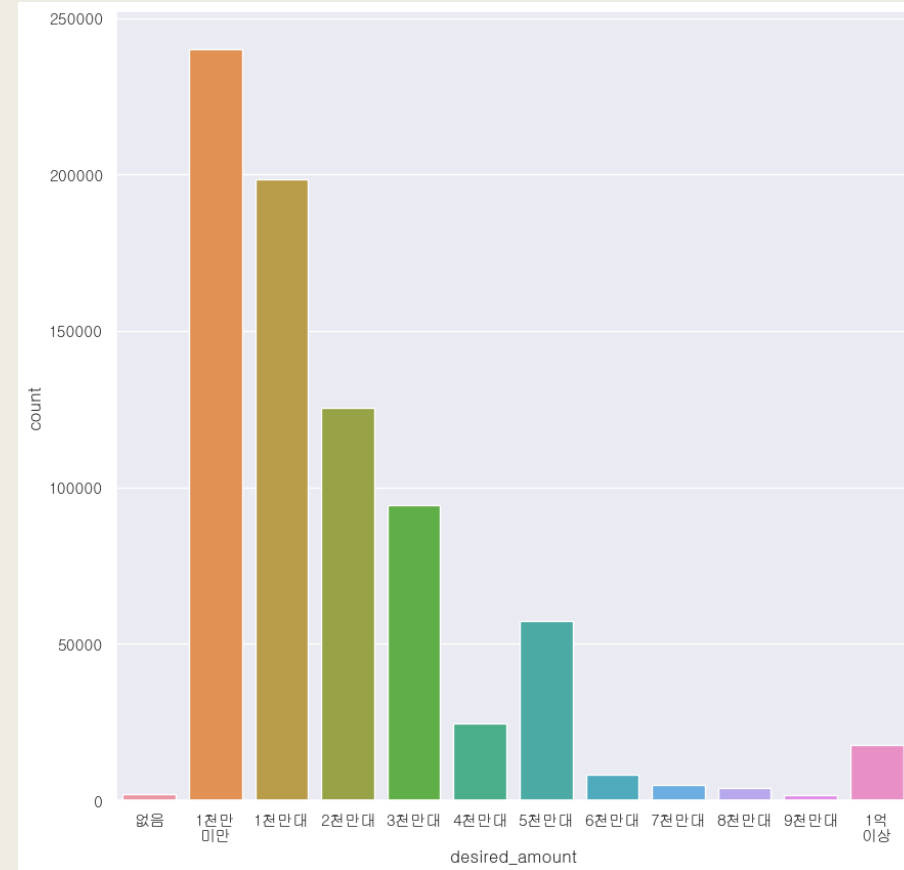
6월 예측 데이터

- 과반 이상이 대출금이 남아있음에도 새로운 대출을 신청함.

# 신청자 특성 분석 (희망 대출 금액)



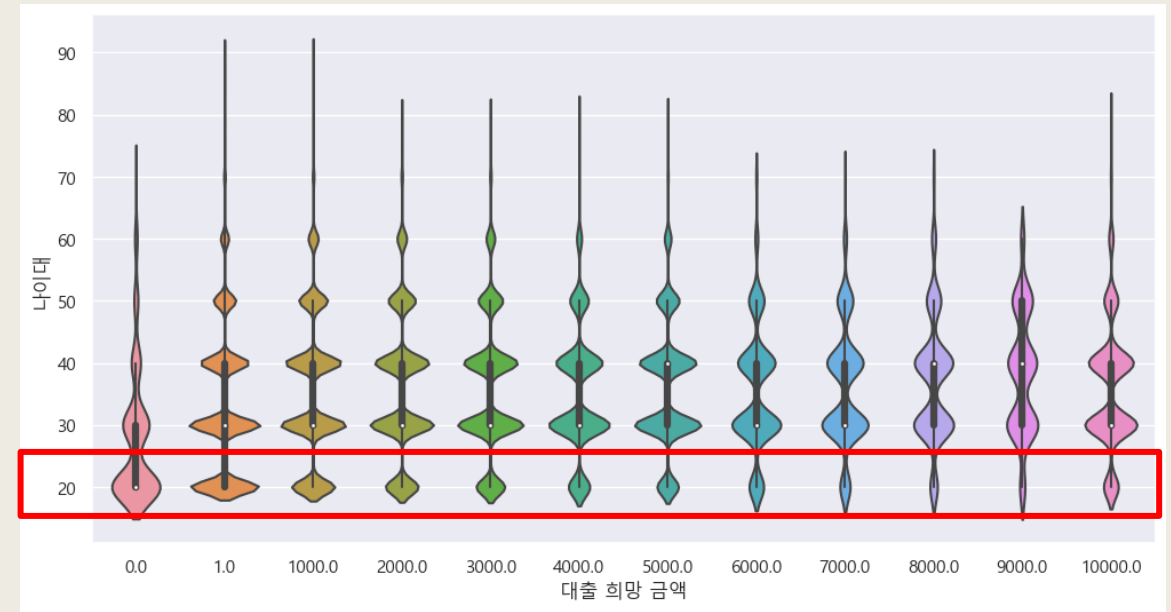
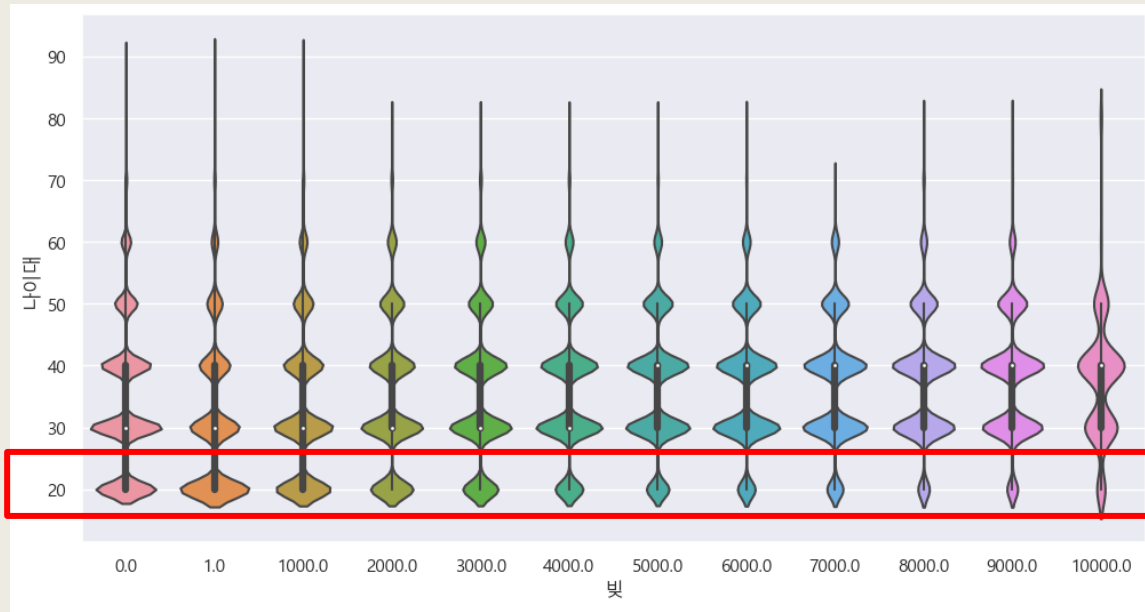
3~5월 실제 데이터



6월 예측 데이터

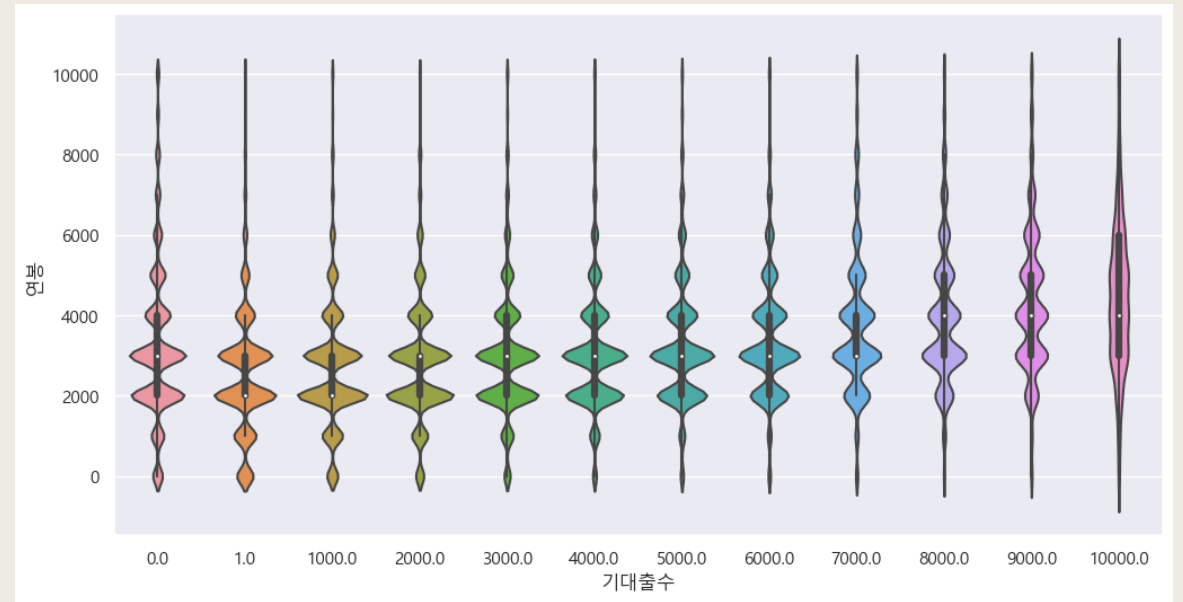
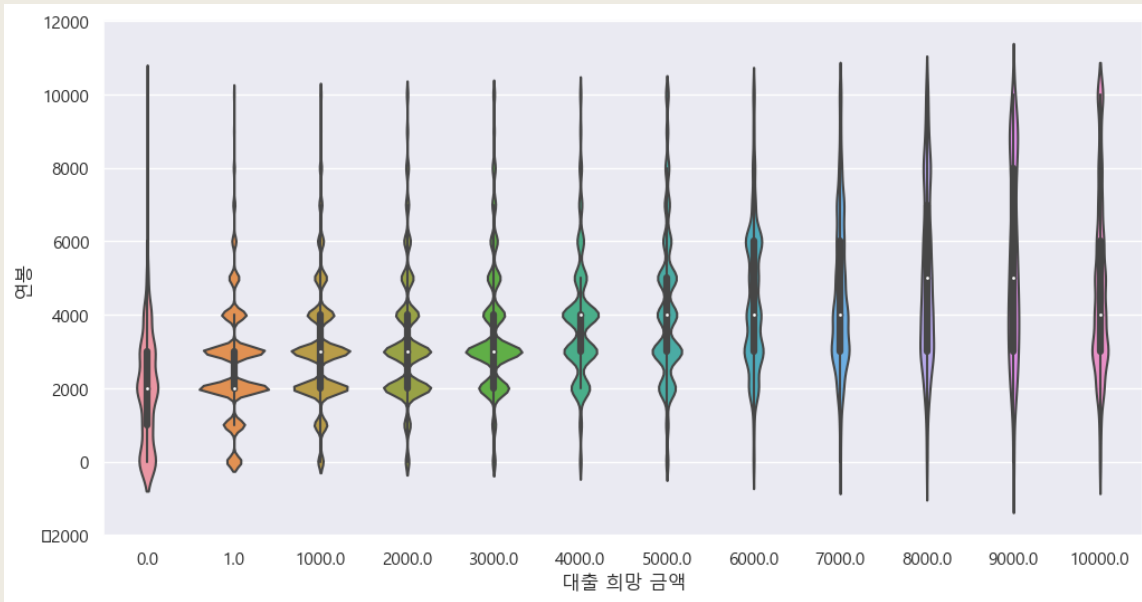
- 1천만원 미만의 소액을 희망하는 사람이 주를 이룸.

# 신청자 특성 분석 (대출 금액 - 나이)



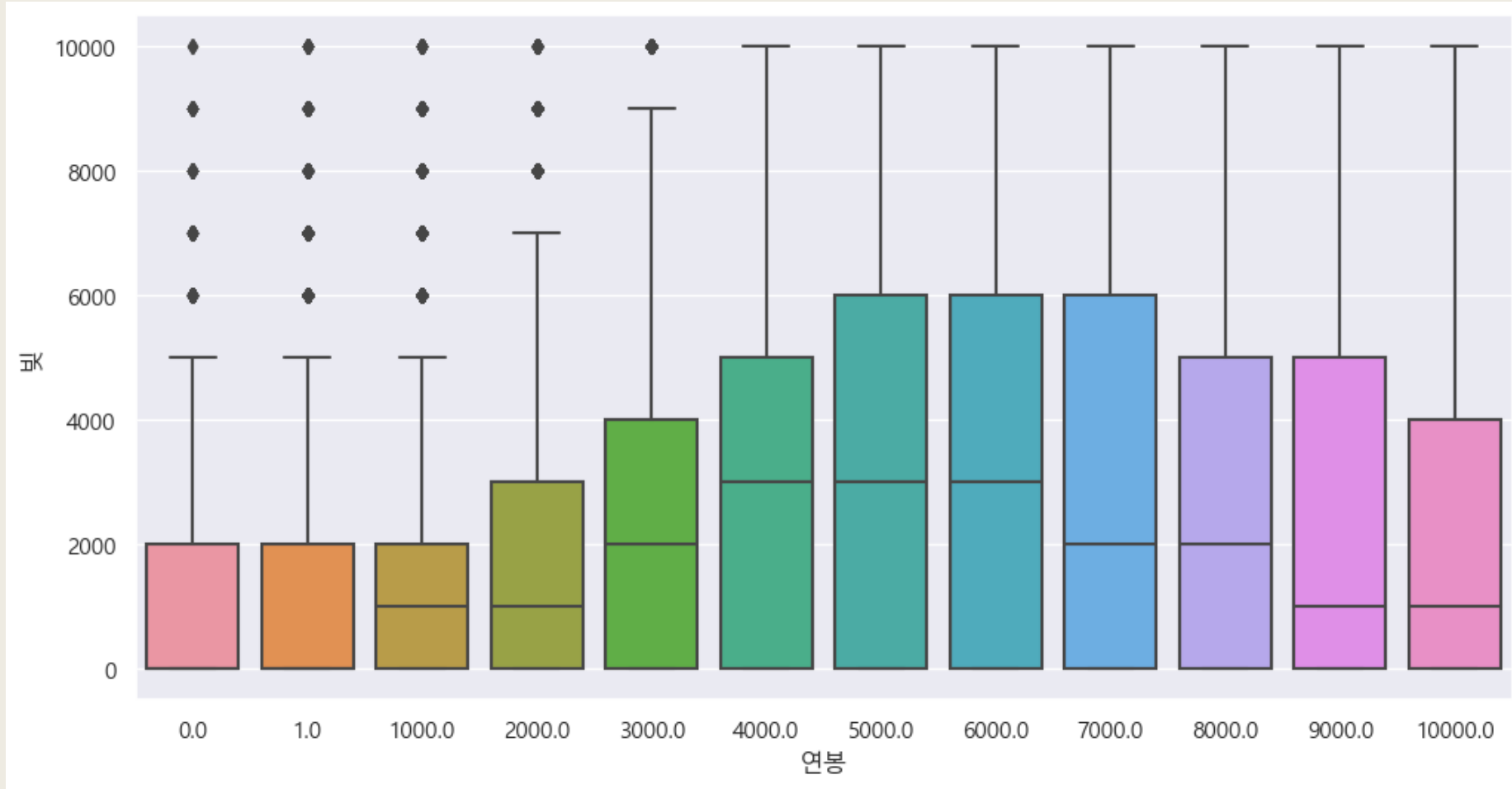
- 기대대출금액 중 소액은 20대가 상당수 차지하며  
대출 희망 금액 또한 20대가 소액을 상당수 차지함.

# 신청자 특성 분석 (대출 금액 - 연 수입)



- 대출 희망 금액 또는 기대대출금액이 높을 수록 연 수입이 높은 추세를 보임.  
(충분조건)

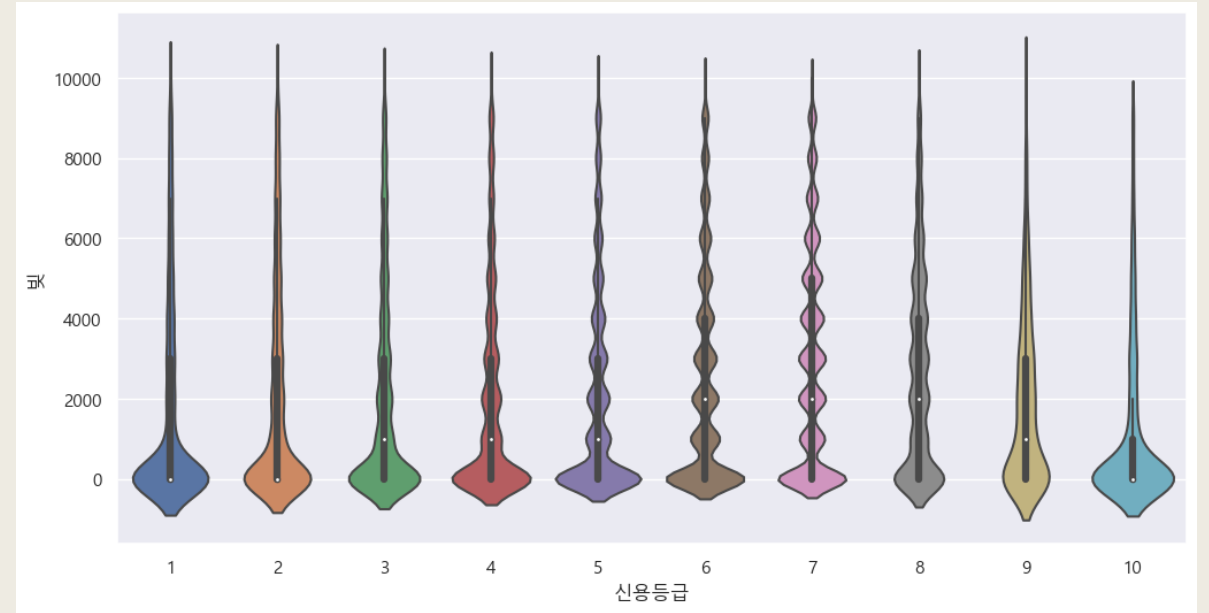
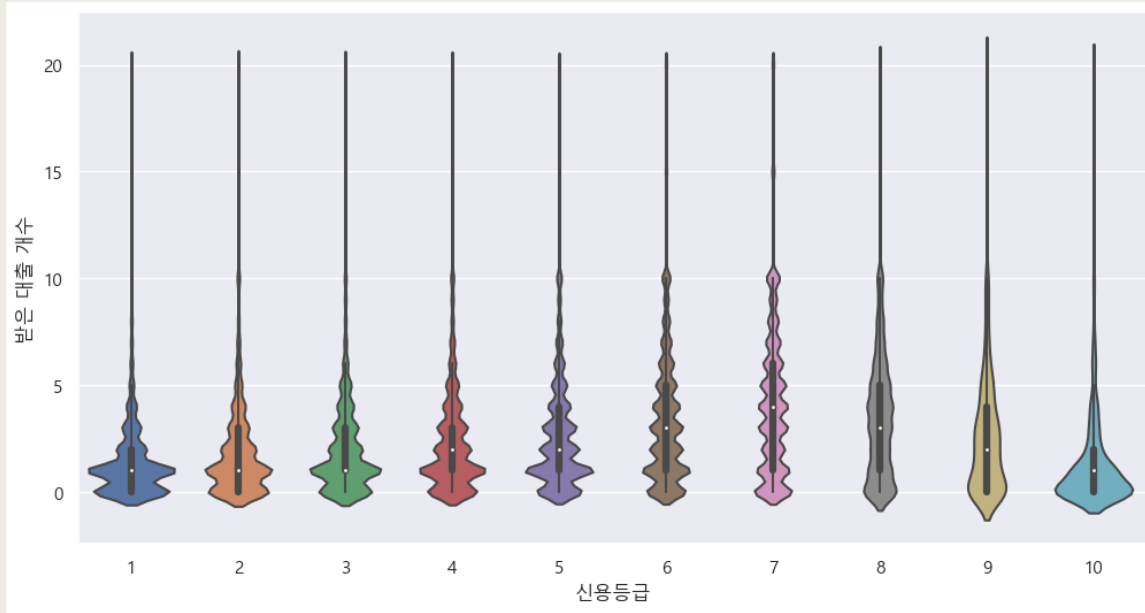
# 신청자 특성 분석 (연 수입 - 기대출금액)



- 연 수입이 5000~6000만원 대의 사람들이 기대출금액이 가장 높음.

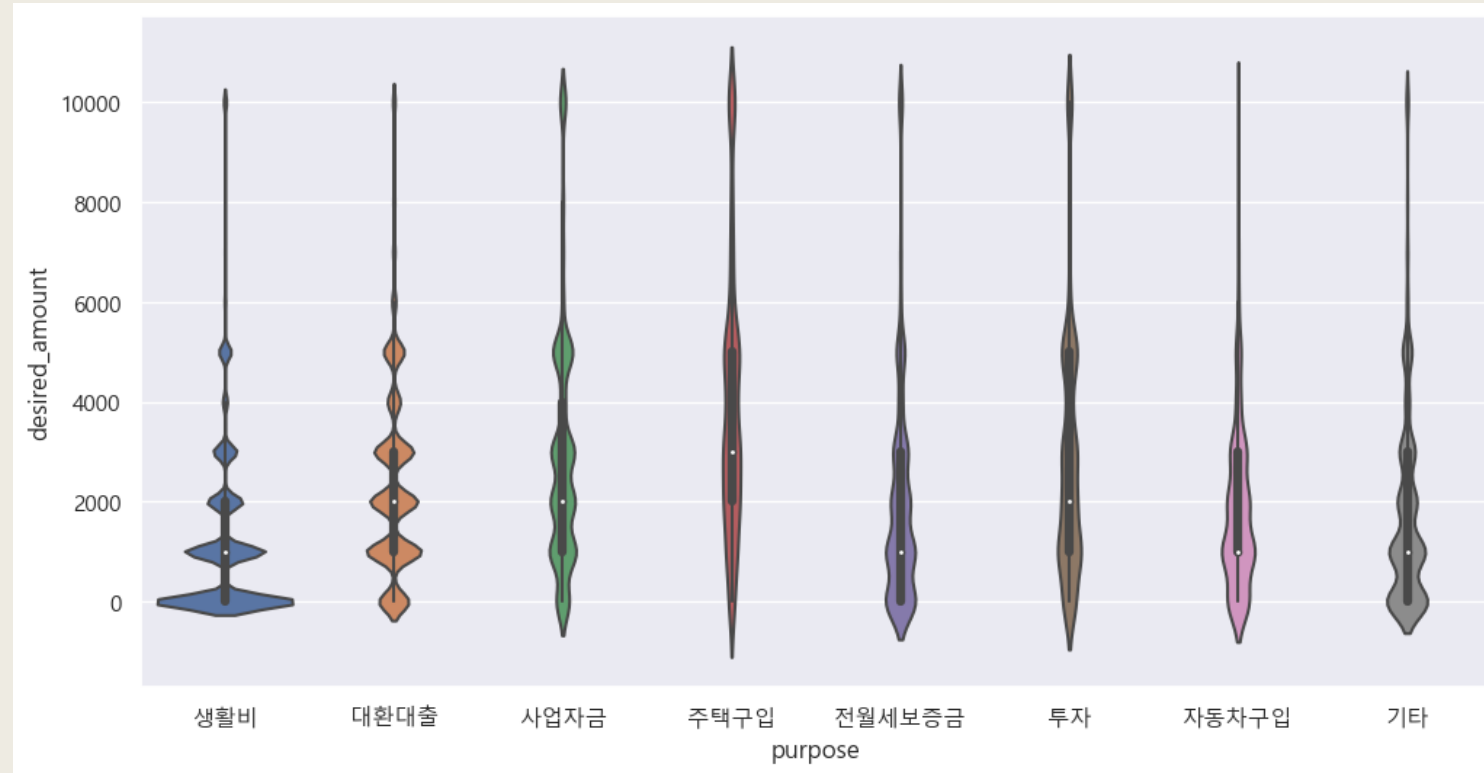


# 신청자 특성 분석 (신용등급 - 기대출금)



- 신용이 높을수록 기대출수와 기대출금액이 모두 낮은 추세를 보임.  
(8, 9, 10 등급 예외)

# 신청자 특성 분석 (목적 - 대출 희망 금액)

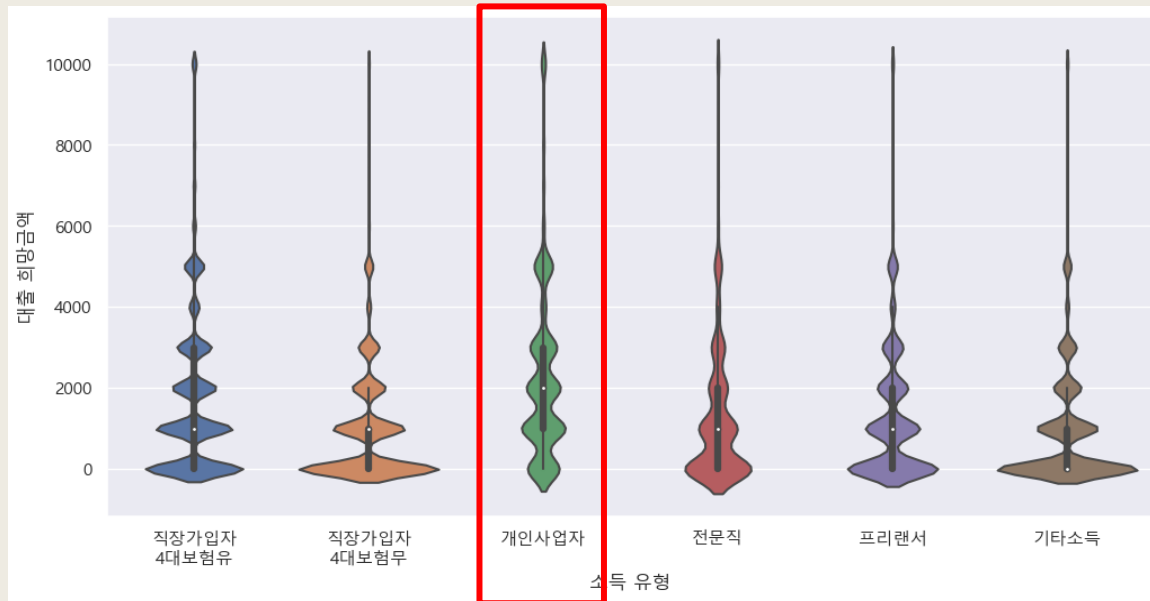


- 생활비가 목적인 경우 소액 대출을 희망함.

사업자금, 주택구입, 투자의 경우 타 목적에 비해 금액대가 골고루 존재함.

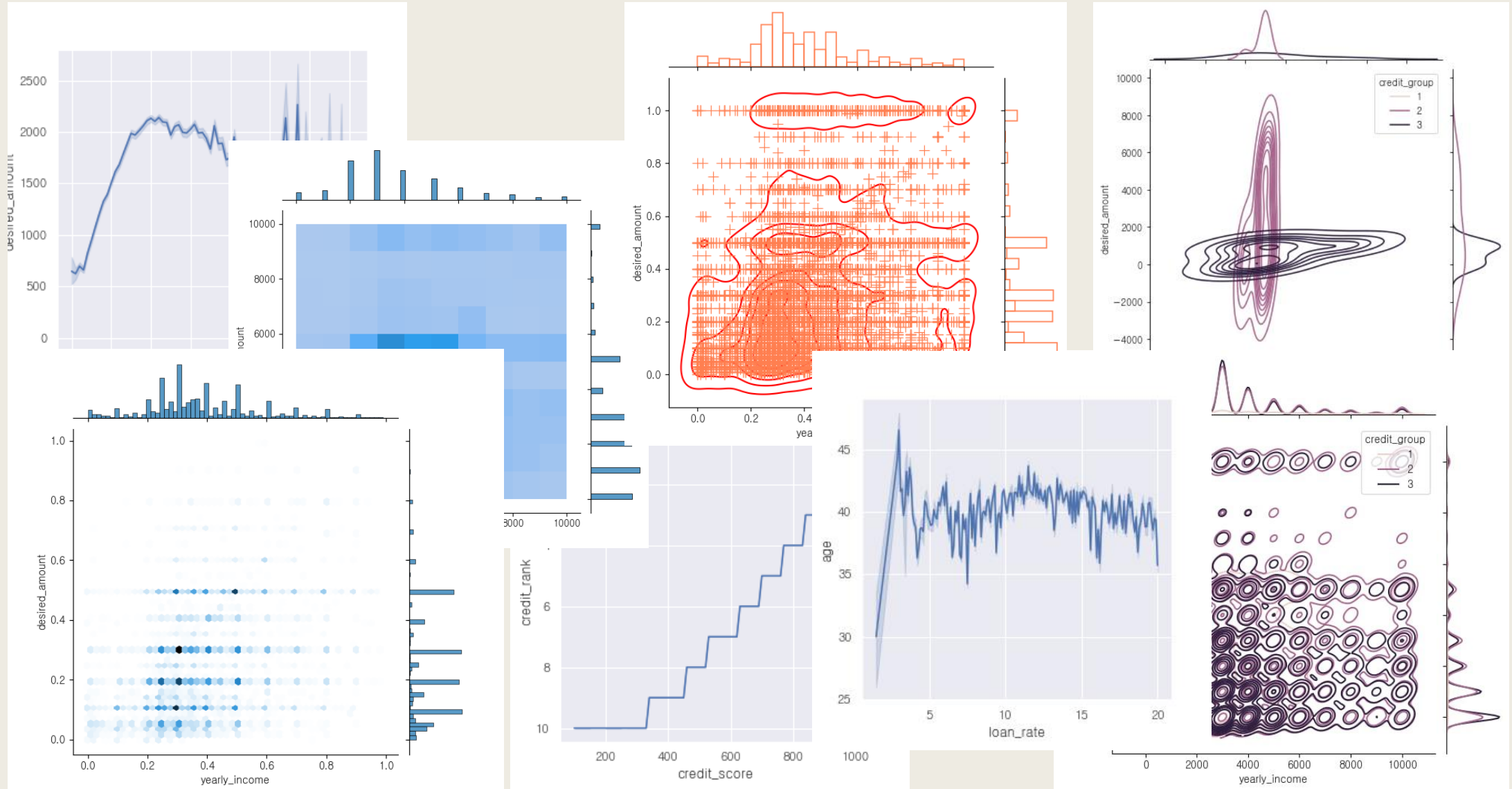
주택 구입의 경우 대출 희망 금액 평균치가 가장 높음.

# 신청자 특성 분석 (개인사업자)



- 개인사업자의 경우 대출 희망 금액이 골고루 존재하고, 평균이 높은 추세임.

# 그 외 특성 분석

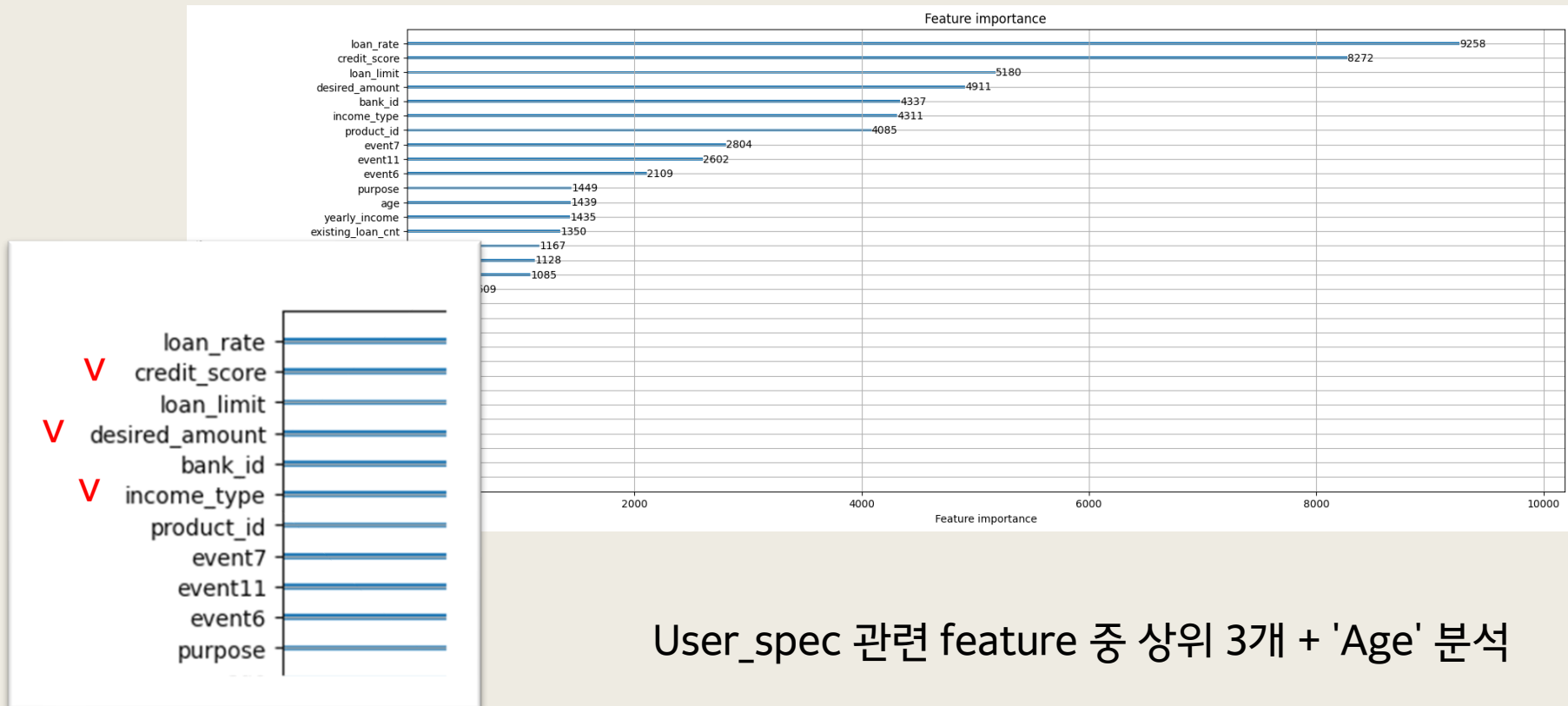




# 상품 분석

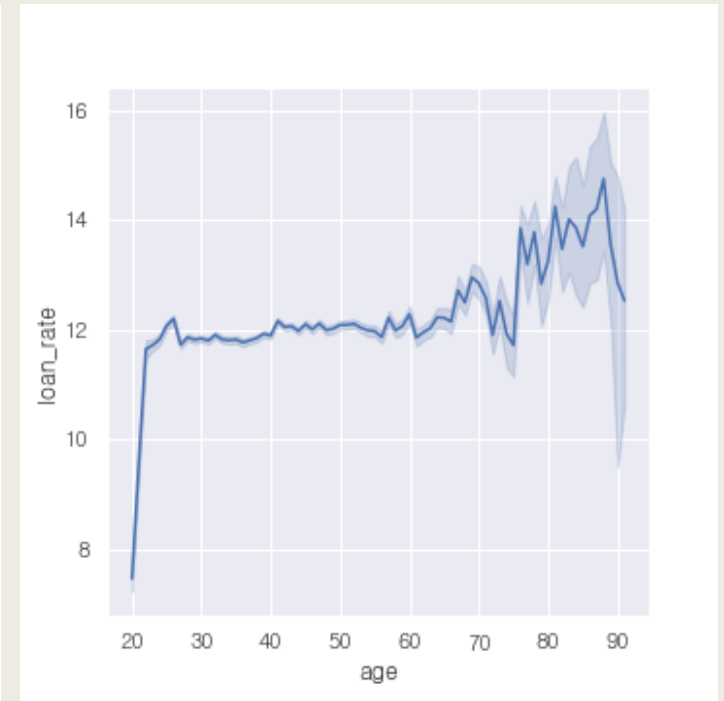
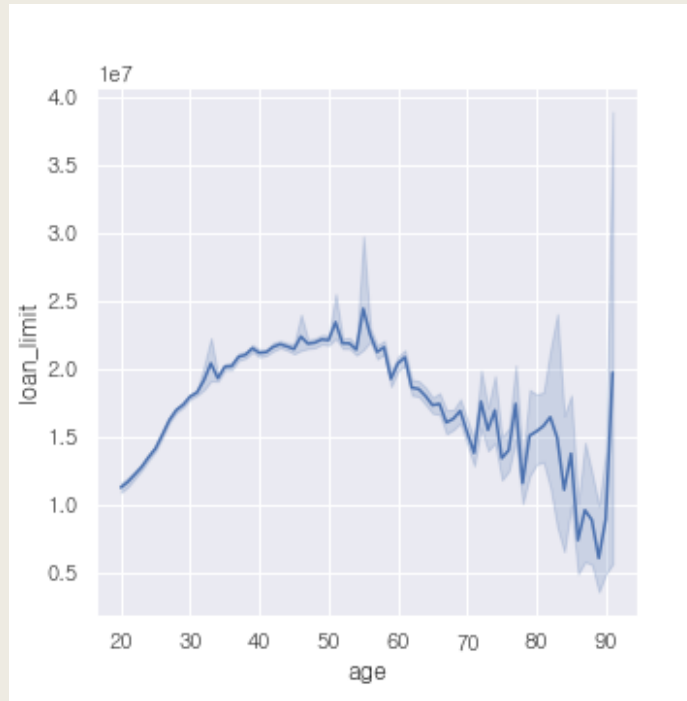
Seaborn 라이브러리를 활용한

# Feature importance 기반 분석



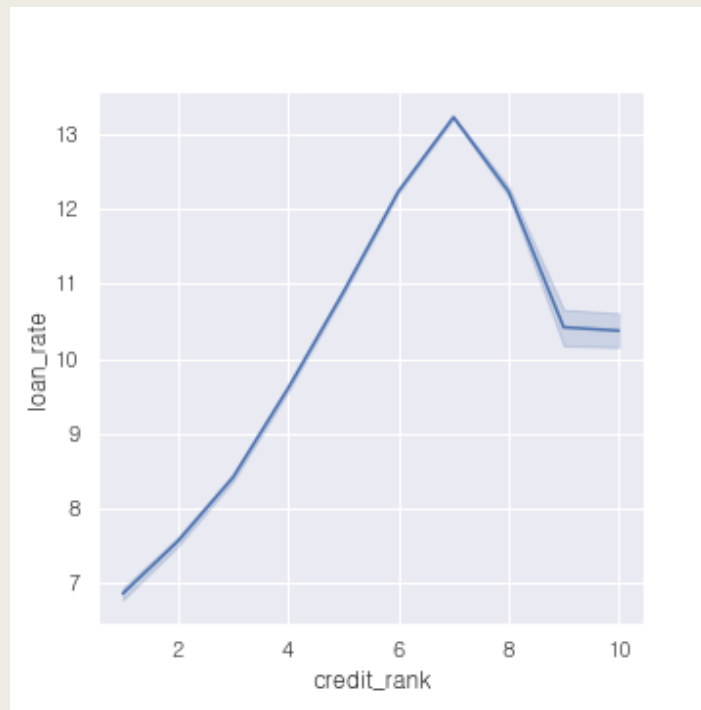
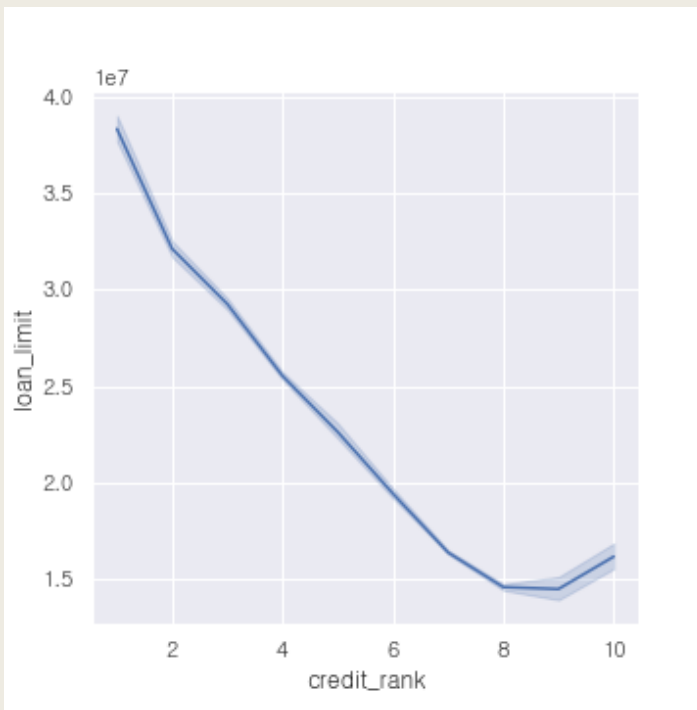
# 상품 분석 (나이)

나이	신청 대출 한도	신청 대출 금리
20대	1,505 만원	11.83 %
30대	1,942 만원	11.83 %
40대	2,164 만원	12.03 %
50대	2,233 만원	12.05 %
60대	1,880 만원	12.14 %
70대	1,562 만원	12.70 %
80대	1,430 만원	13.66 %
90대	1,104 만원	13.14 %



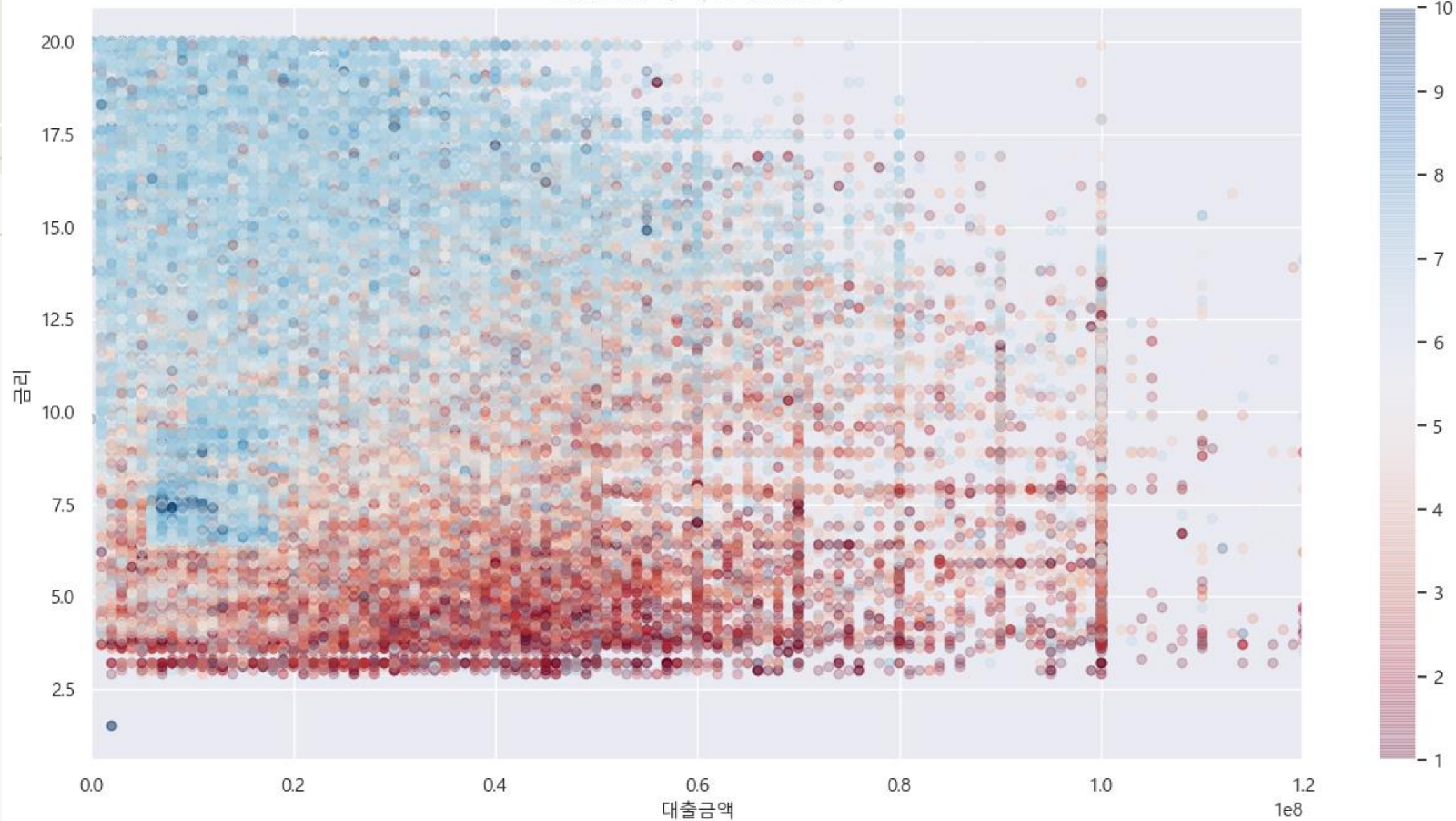
# 상품 분석 (신용 등급)

등급	신청 대출 한도	신청 대출 금리
1등급	3,833 만원	6.86 %
2등급	3,212 만원	7.56 %
3등급	2,927 만원	8.41 %
4등급	2,555 만원	9.60 %
5등급	2,265 만원	10.89 %
6등급	1,946 만원	12.24 %
7등급	1,614 만원	13.23 %
8등급	1,462 만원	12.24 %
9등급	1,452 만원	10.42 %
10등급	1,620 만원	10.37 %



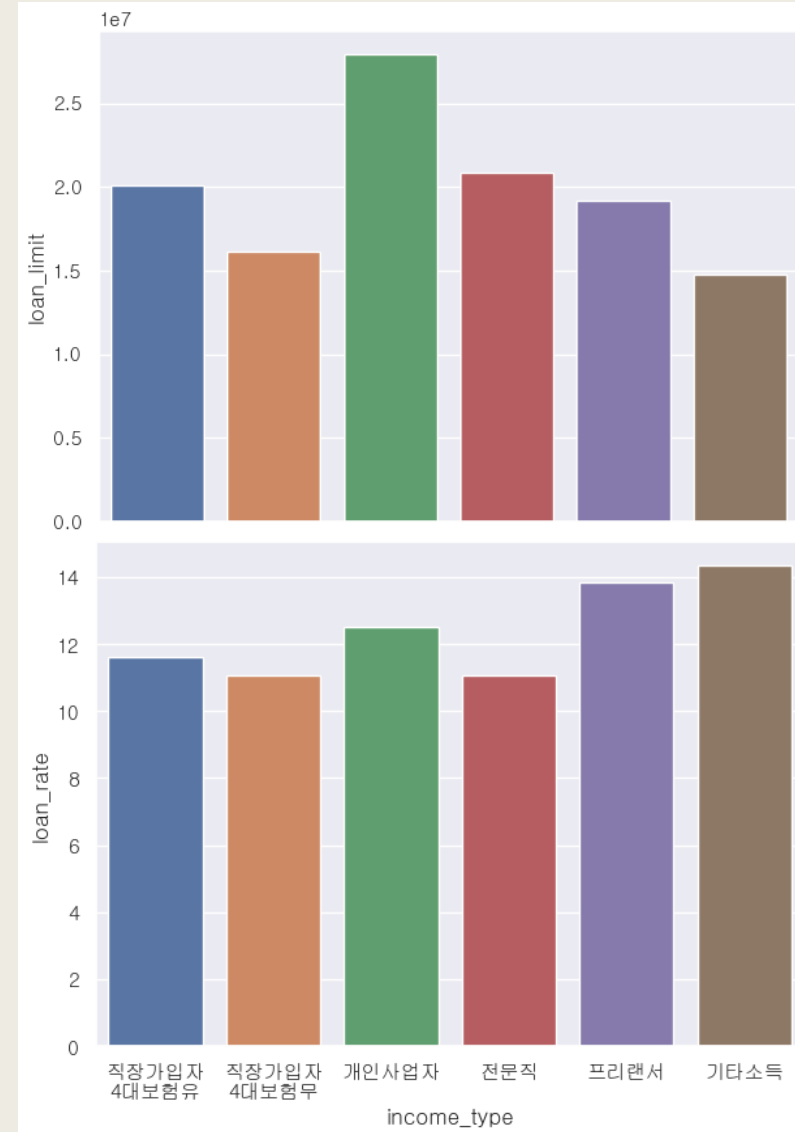


신용등급에 따른 상품분석



# 상품 분석 (근로 형태)

근로형태	신청 대출 한도	신청 대출 금리
직장가입자(4대보험O)	2,013 만원	11.60 %
직장가입자(4대보험X)	1,615 만원	11.08 %
개인사업자	2,793 만원	12.50 %
전문직	2,089 만원	11.08 %
프리랜서	1,920 만원	13.82 %
기타소득	1,481 만원	14.33 %





loan\_rate

20.0  
17.5  
15.0  
12.5  
10.0  
7.5  
5.0  
2.5

0.0

0.2

0.4

0.6

0.8

1.0

1.2

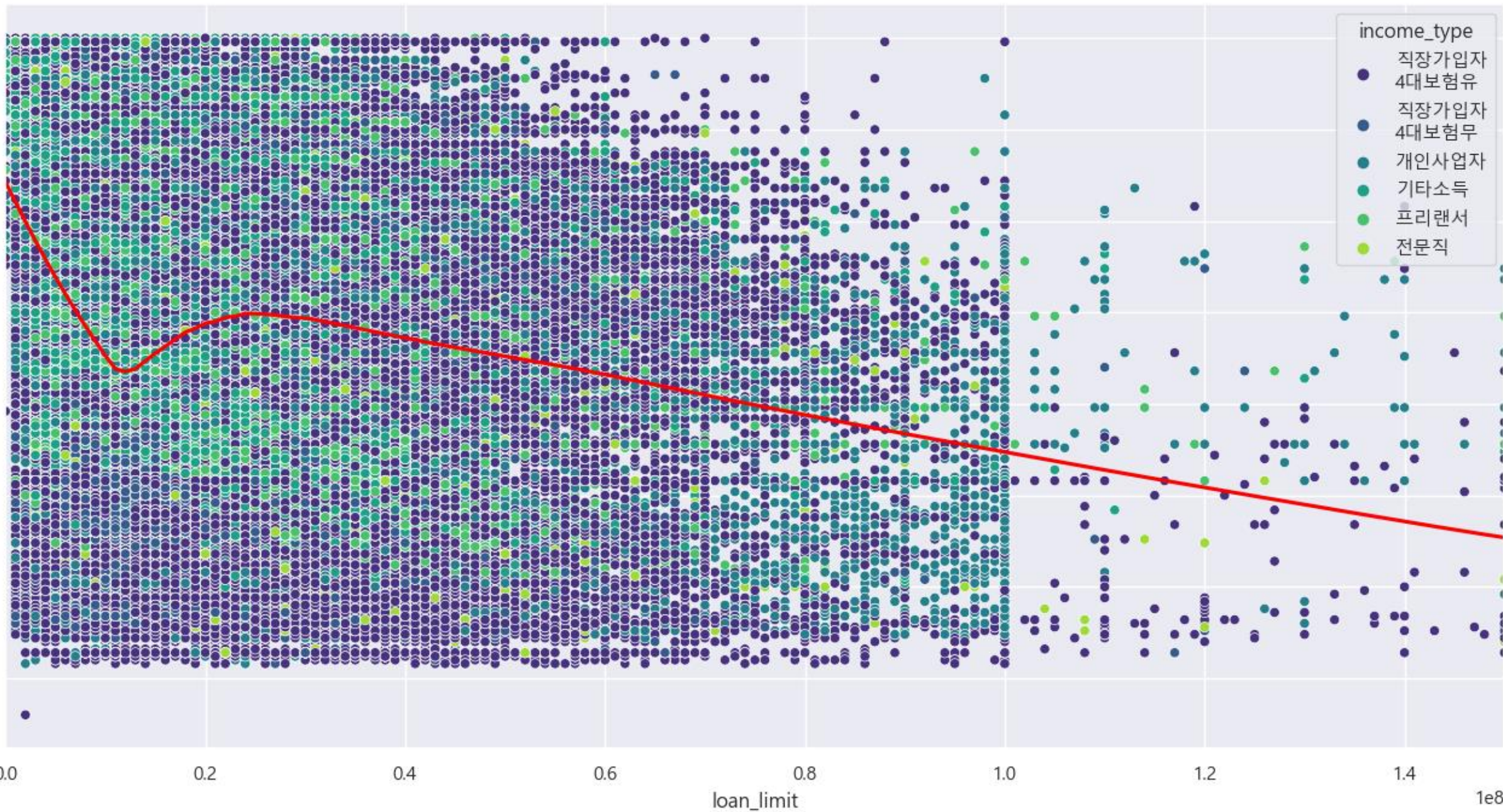
1.4

loan\_limit

1e8

income\_type

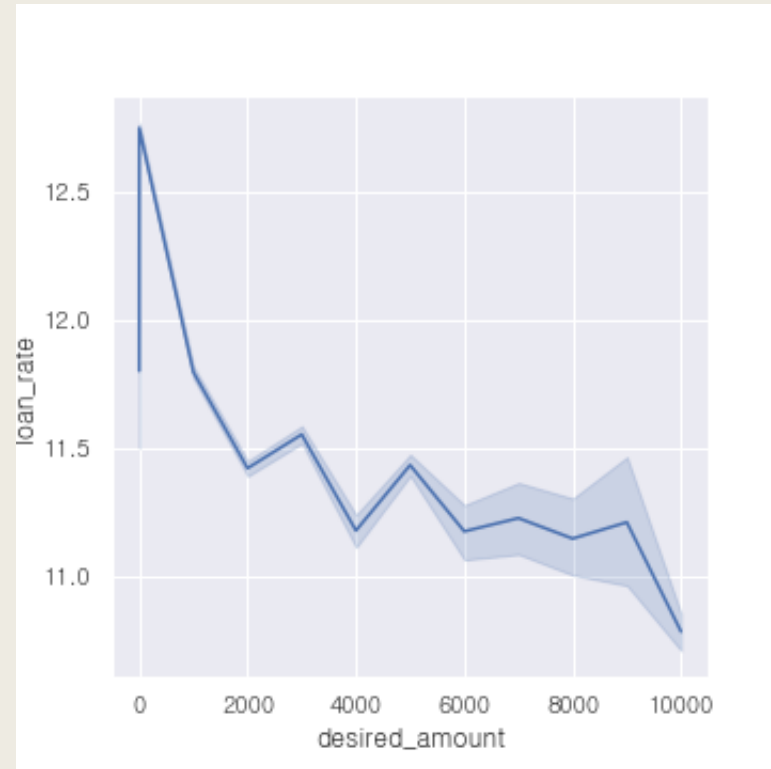
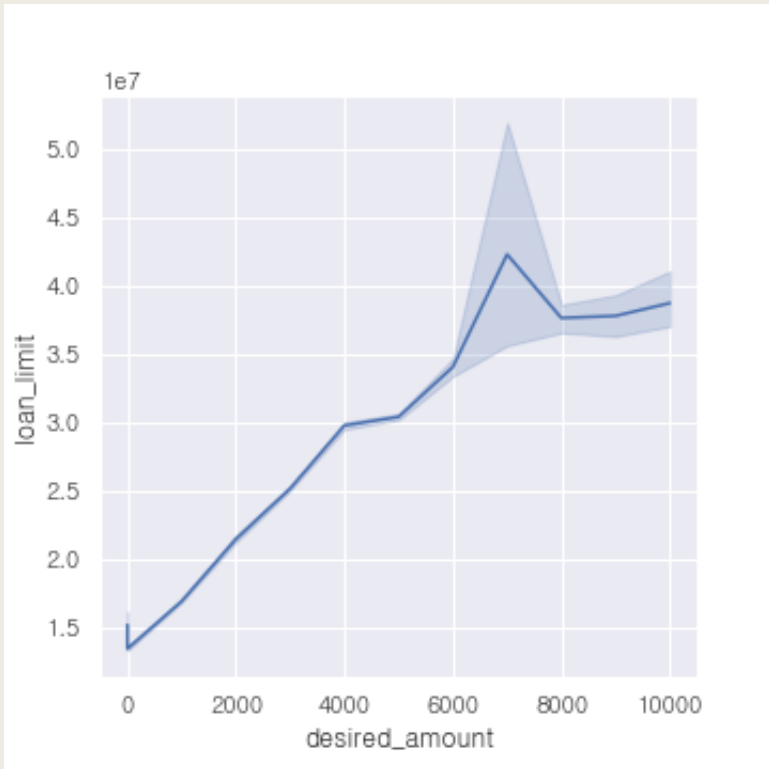
- 직장가입자
- 4대보험유
- 직장가입자
- 4대보험무
- 개인사업자
- 기타소득
- 프리랜서
- 전문직





# 상품 분석 (대출 희망 금액)

희망 금액	신청 대출 한도	신청 대출 금리
없음	1,525 만원	11.80 %
1,000 만원 미만	1,351 만원	12.75 %
1,000 만원 대	1,698 만원	11.80 %
2,000 만원 대	2,152 만원	11.42 %
3,000 만원 대	2,523 만원	11.55 %
4,000 만원 대	2,984 만원	11.18 %
5,000 만원 대	3,047 만원	11.44 %
6,000 만원 대	3,412 만원	11.18 %
7,000 만원 대	4,236 만원	11.23 %
8,000 만원 대	3,771 만원	11.15 %
9,000 만원 대	3,786 만원	11.21 %
1억 이상	3,879 만원	10.79 %





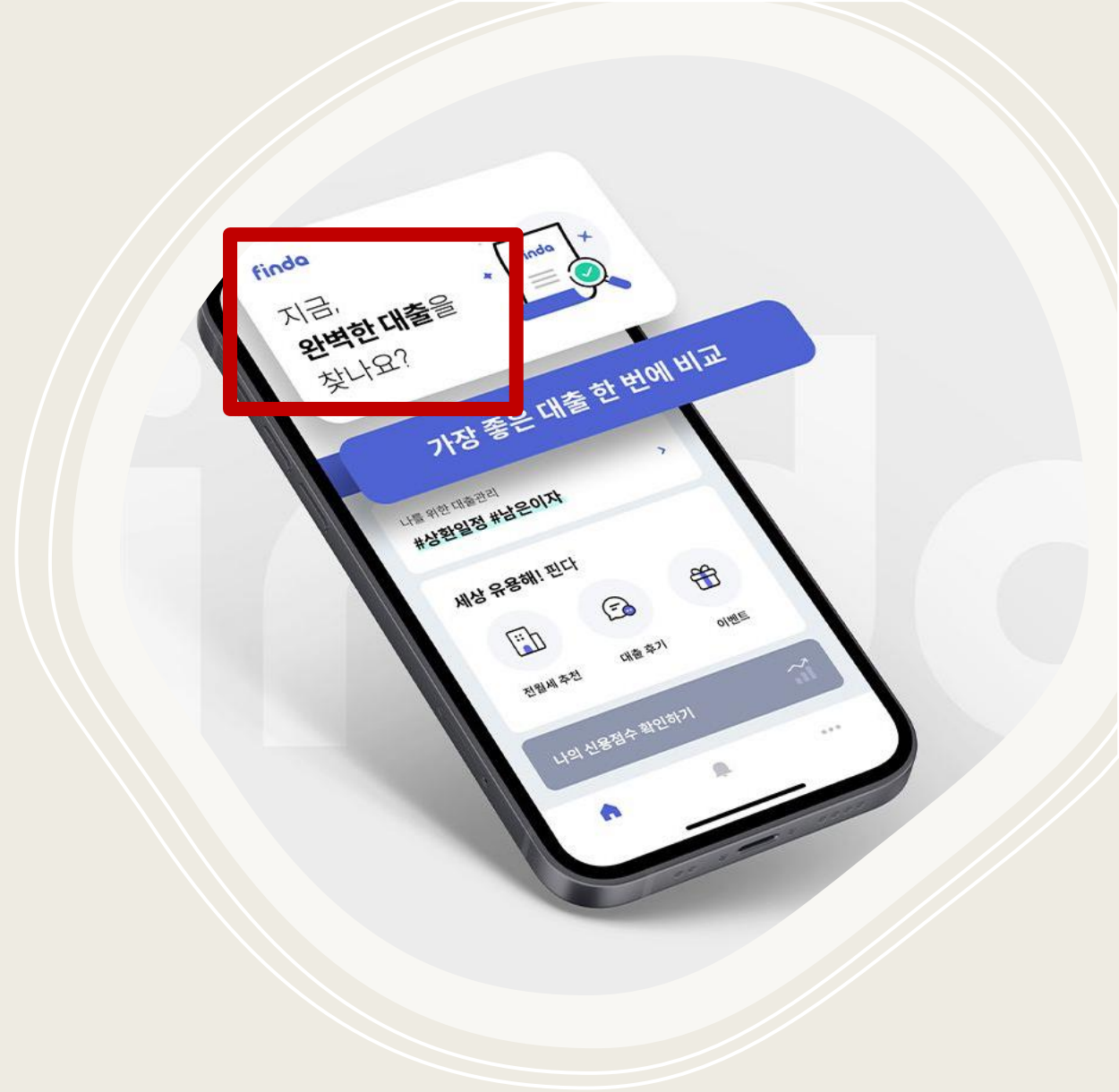
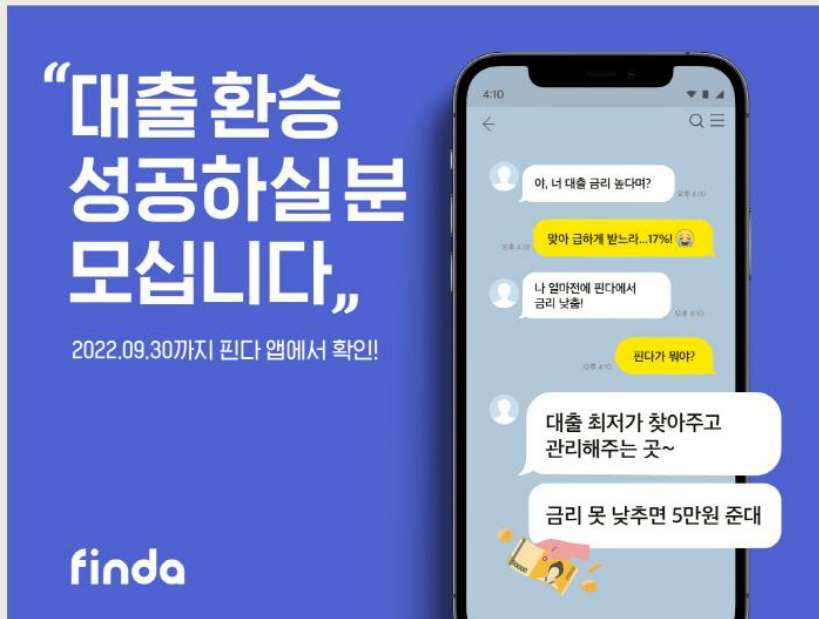
# 로그데이터 분석

# 로그데이터 분석 (근로 형태)

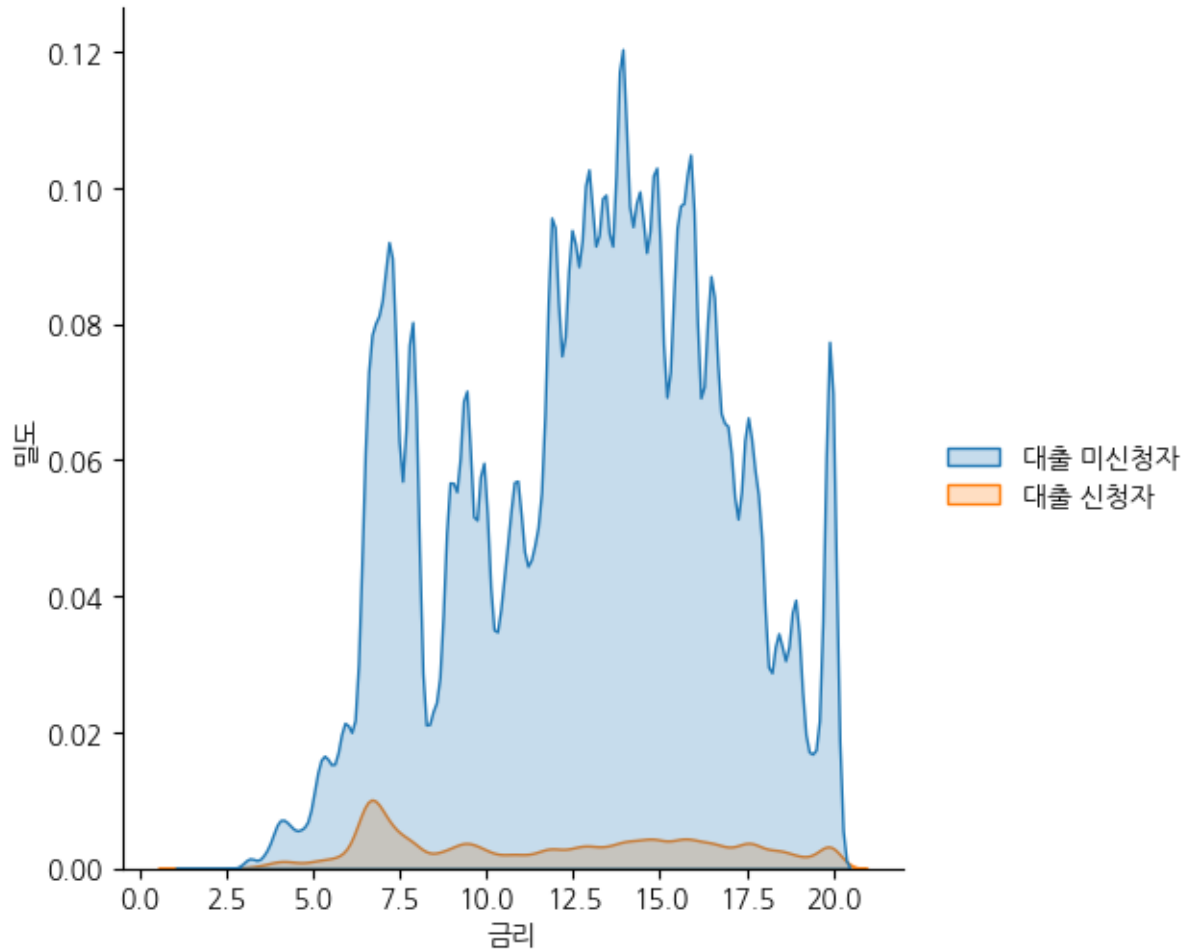
로그데이터 내에서

Importance	Action	Rate
Event 7	한도 조회 결과 확인	2804
Event 11	KCB 신용정보 조회	2602
Event 5	한도 조회 시작 버튼 클릭	2109
Event 6	본인 인증 완료	1085
Event 4	한도 조회 페이지 접속	609
Event 2	어플 실행	458
Event 3	어플 로그인	400
Event 8	대출 관리 서비스 이용	355
Event 1	회원 가입	188
Event 9	여윳돈 계산기 서비스 이용	52
Event 10	DSR 계산기 서비스 이용	46

# 제안 메시지



# 사용자 특성 분석 후 추천 멘트



※신청유도 : 6%대 금리 신청자가 가장 많음에도 불구하고 평균은 11.94%이므로 사용자를 현혹합니다.

**finda**

박재원님 안녕하세요

대출 한도 조회

현재이용자 평균 대출금리 11.94%

당신은 현재 높은 금리로 대출받고 있어요! 더 낮은 금리로 갈아타볼까요?

내 대출 한도 한번에 조회하기

내 대출

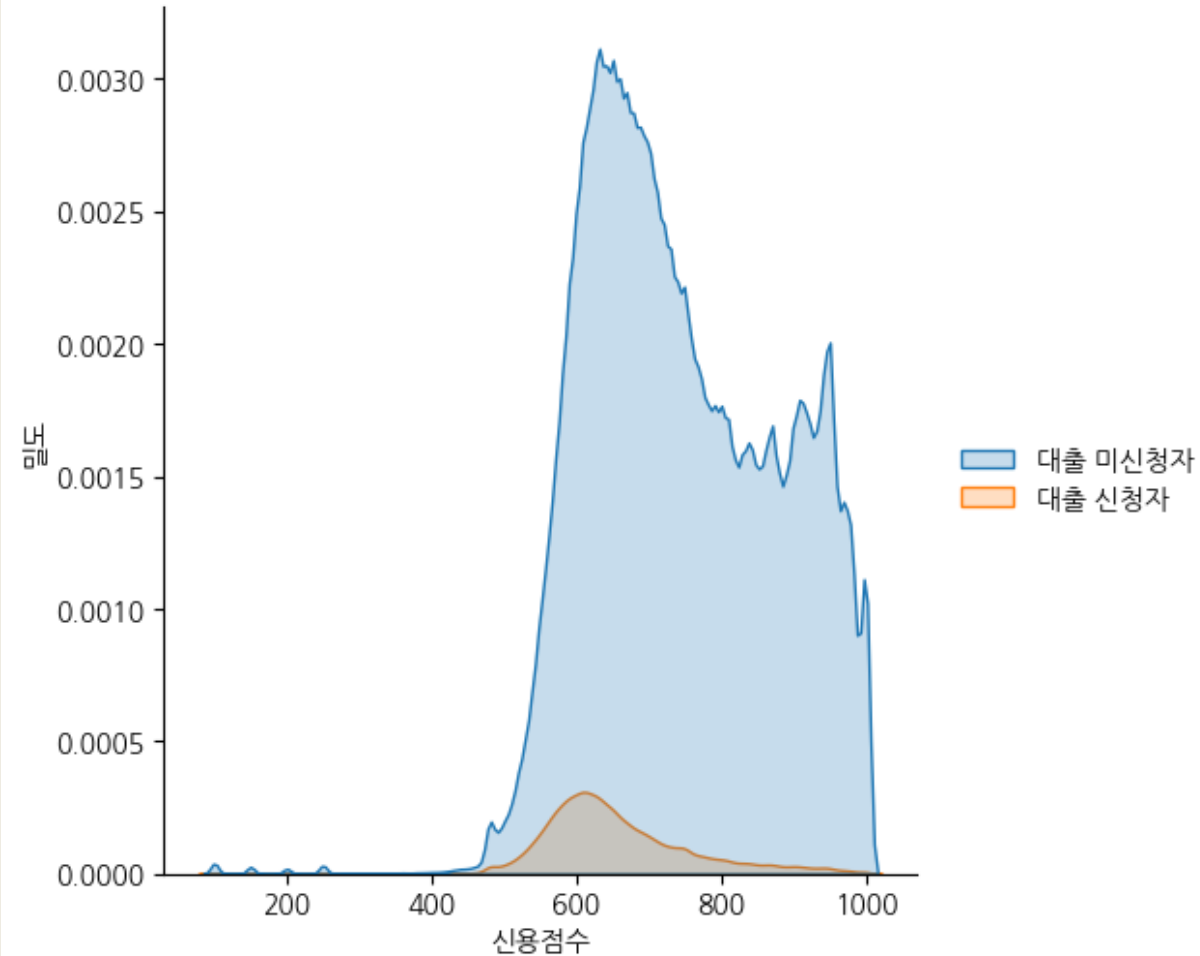
정확한 대출진단을 위해 대출 계좌 연결이 필요합니다.

상환 예정

코 앞으로 다가온 내 대출 상환일과 계좌,



# 사용자 특성 분석 후 추천 멘트



※신청자의 신용점수 평균이 낮다는 점을 이용해  
신용점수가 높은 대출 미신청자를 현혹합니다.

**finda**

박재원님 안녕하세요

대출 한도 조회

조회전

신청자 신용점수 평균 656점!

당신의 신용점수 800점으로  
대출 조건이 유리해요!

내 대출 한도 한번에 조회하기

내 대출

확인하기 >

정확한 대출진단을 위해 대출 계좌  
연결이 필요합니다.

상환 예정


확인하기 >


코 앞으로 다가온 내 대출 상환일과 계좌,

# 상품 분석 후 추천 멘트


## 나이

나이	신청 대출 한도	신청 대출 금리
20대	1,505 만원	11.83 %
30대	1,942 만원	11.83 %
40대	2,164 만원	12.03 %
50대	2,233 만원	12.05 %
60대	1,880 만원	12.14 %
70대	1,562 만원	12.70 %
80대	1,430 만원	13.66 %
90대	1,104 만원	13.14 %

**finda**



박재원님 안녕하세요

대출 한도 조회 

조회전


40대 신청대출금리는 12.03%  
대출 한도는 2,164 만원이예요!  
OOO님은 어떠실까요?

내 대출 한도 한번에 조회하기

내 대출

확인하기 >




정확한 대출진단을 위해 대출 계좌  
연결이 필요합니다.



상환 예정

확인하기 >


코 앞으로 다가온 내 대출 상환일과 계좌,




# 상품 분석 후 추천 멘트

## 신용 점수

등급	신청 대출 한도	신청 대출 금리
1등급	3,833 만원	6.86 %
2등급	3,212 만원	7.56 %
3등급	2,927 만원	8.41 %
4등급	2,555 만원	9.60 %
5등급	2,265 만원	10.89 %
6등급	1,946 만원	12.24 %
7등급	1,614 만원	13.23 %
8등급	1,462 만원	12.24 %
9등급	1,452 만원	10.42 %
10등급	1,620 만원	10.37 %

**finda**

박재원님 안녕하세요

대출 한도 조회 

조회전







비슷한 신용점수대의 회원들은  
대출금리 10.89%,  
대출한도 2265 만원의 상품을  
선택했어요!

내 대출 한도 한번에 조회하기

내 대출

확인하기 >




정확한 대출진단을 위해 대출 계좌  
연결이 필요합니다.



상환 예정

확인하기 >


코 앞으로 다가온 내 대출 상환일과 계좌,




# 상품 분석 후 추천 멘트


## 근로형태

근로형태	신청 대출 한도	신청 대출 금리
직장가입자(4대보험O)	2,013 만원	11.60 %
직장가입자(4대보험X)	1,615 만원	11.08 %
개인사업자	2,793 만원	12.50 %
전문직	2,089 만원	11.08 %
프리랜서	1,920 만원	13.82 %
기타소득	1,481 만원	14.33 %

**finda**



박재원님 안녕하세요

대출 한도 조회 

조회전


비슷한 직종의 회원들은  
대출금리 12.50%, 대출한도 2793 만원의  
상품을 선택했어요!

내 대출 한도 한번에 조회하기

내 대출

확인하기 >




정확한 대출진단을 위해 대출 계좌  
연결이 필요합니다.

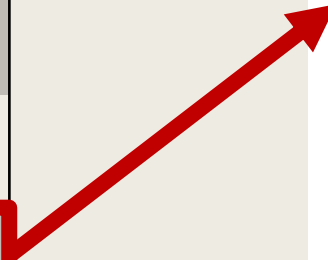


상환 예정

확인하기 >

코 앞으로 다가온 내 대출 상환일과 계좌,







# 상품 분석 후 추천 멘트

## 희망 금액

희망 금액	신청 대출 한도	신청 대출 금리
없음	1,525 만원	11.80 %
1,000 만원 미만	1,351 만원	12.75 %
1,000 만원 대	1,698 만원	11.80 %
2,000 만원 대	2,152 만원	11.42 %
3,000 만원 대	2,523 만원	11.55 %
4,000 만원 대	2,984 만원	11.18 %
5,000 만원 대	3,047 만원	11.44 %
6,000 만원 대	3,412 만원	11.18 %
7,000 만원 대	4,236 만원	11.23 %
8,000 만원 대	3,771 만원	11.15 %
9,000 만원 대	3,786 만원	11.21 %
1억 이상	3,879 만원	10.79 %

**finda**

박재원님 안녕하세요

대출 한도 조회 

조회전







5천만원대의 대출을 희망하는 회원들은 대출금리 11.44%, 대출한도 3047만 원의 상품을 선택했어요!

내 대출 한도 한번에 조회하기

내 대출

확인하기 >




정확한 대출진단을 위해 대출 계좌 연결이 필요합니다.



상환 예정

확인하기 >

코 앞으로 다가온 내 대출 상환일과 계좌,



# 광고멘트

---

"너 나보다 연봉은 높으면서 대출금리는 왜 그렇게 높아..?"

finda

나의 대출 스펙 비교하기 →

"은행 몇 군데 가봤어요? 우린 손가락 하나면 충분해요."

finda

손가락 하나로 수고 줄이기 →

"그 정도면 사기 아니야? 왜 그런 걸... 아니 클릭이나 해보지"

finda

현명한 대출자였는지 확인해보기 →