

Drug Consumption

2조

고정민 김관석 김수연

김유경 서상수

목차

STEP 1

문제 및 목표 정의

이번 분석의 배경과 문제점을 파악한 후 이를 해결하기 위한 목표 정의

>>

STEP 2

데이터 설명

표본 수, 변수 개수 등 기본적인 데이터 형태와 설명변수, 목적변수에 대한 설명

>>

STEP 3

데이터 분석

변수 전처리와 추출과 더불어 기본적인 EDA 및 시각화 진행

>>

STEP 4

향후 추진 계획

1주차 진행상황에 따른 보완점을 해결하기 위한 향후 과제 수립 및 2주차 모델링 방향 설명

문제 및 목표 정의

배경

- ✓ 2017년 트럼프의 opioid crisis 선언
마약성 진통제 '오피오이드' 남용에 대한 공중보건 위기사태를 선포함.
- ✓ 코로나로 인한 약물중독 사망자 수 30% 증가
방역 조치로 인해 약물 중독자를 위한 프로그램이 전혀 진행되지 못 함.
거리두기로 인한 외로움, 불안감 등이 커진 것도 주요 요인으로 생각됨.

과제 목표

- ✓ 주어진 정보를 바탕으로 한 사람의 약물 중독 여부를 예측
- ✓ 약물의 중독성을 파악하여 위험 대비
- ✓ 나아가 성격적인 특성이 약물 중독에 얼마나 영향을 미치는지 분석

데이터 설명

[illegible]

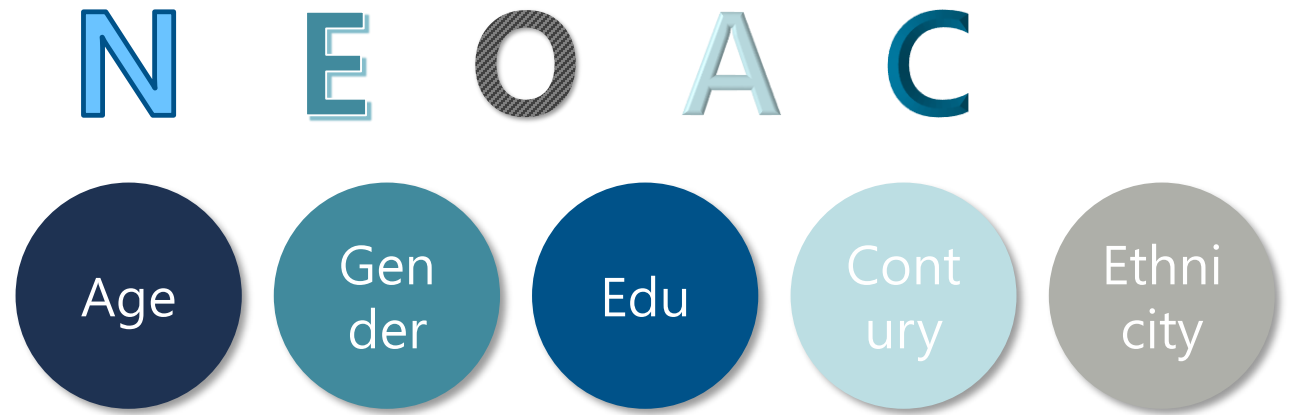
데이터 설명

설명변수

- Personal traits
- Demographic traits

목적변수

- Potential Overdose



데이터 분석

A. 인구통계

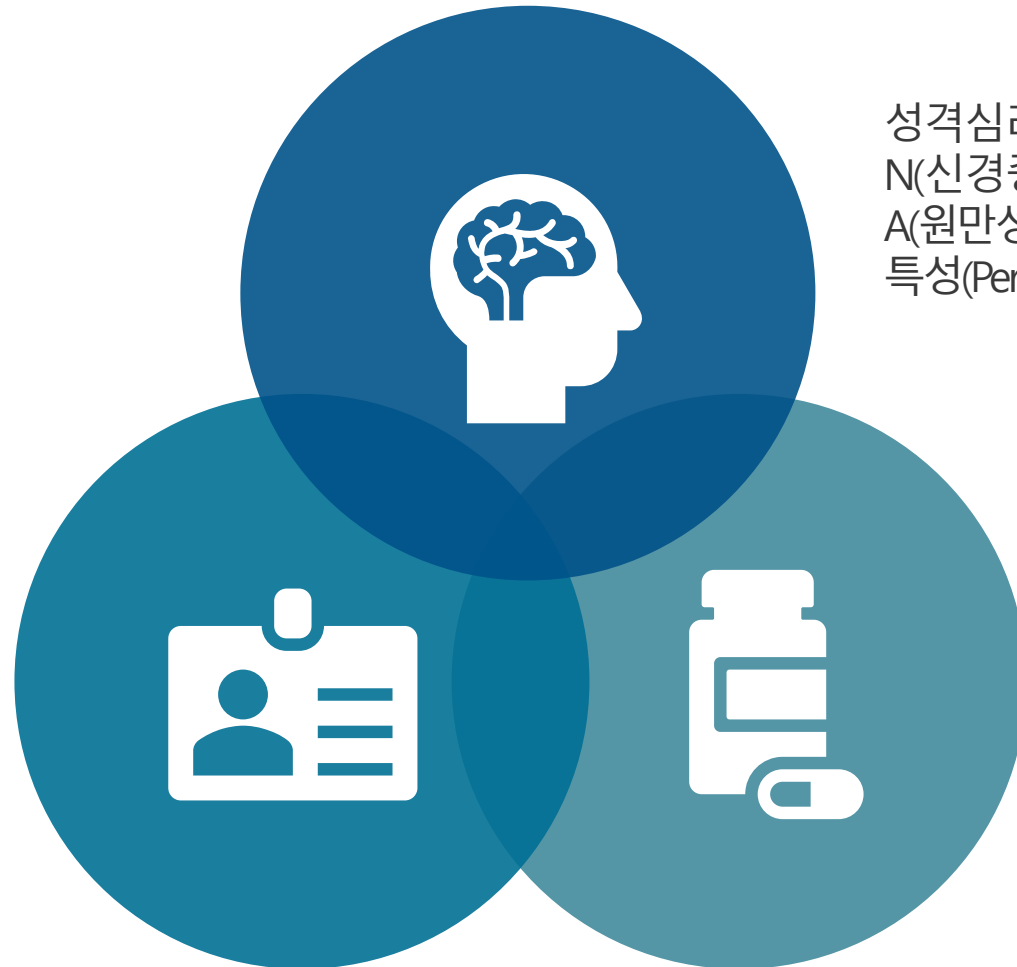
연령, 성별, 학력수준, 국가, 인종 등의 기본적인 인구통계적인 특성(Demographic traits)에 대한 분석

B. 성격지표

성격심리학의 대표적인 지표 - N(신경증), E(외향성), O(개방성), A(원만성), C(성실성) 등 성격적인 특성(Personality traits)에 대한 분석

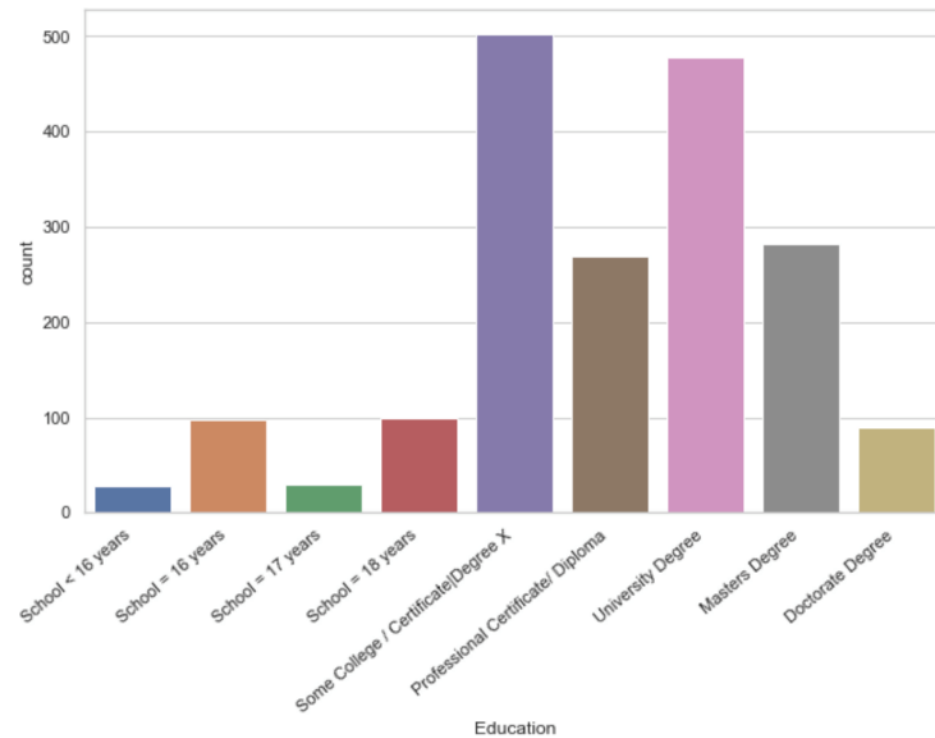
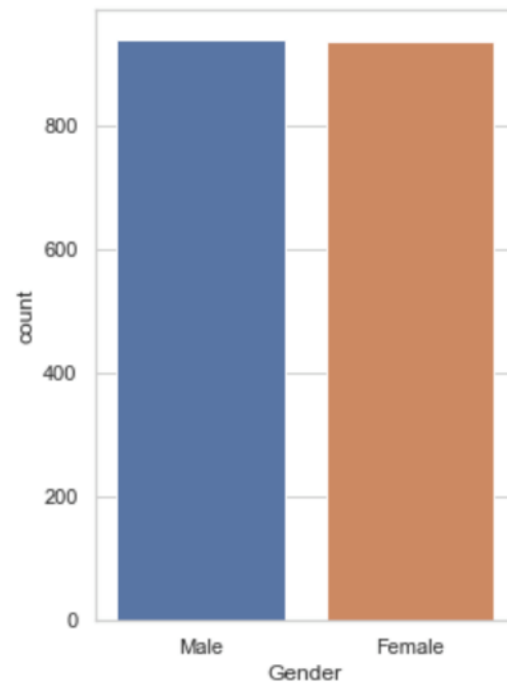
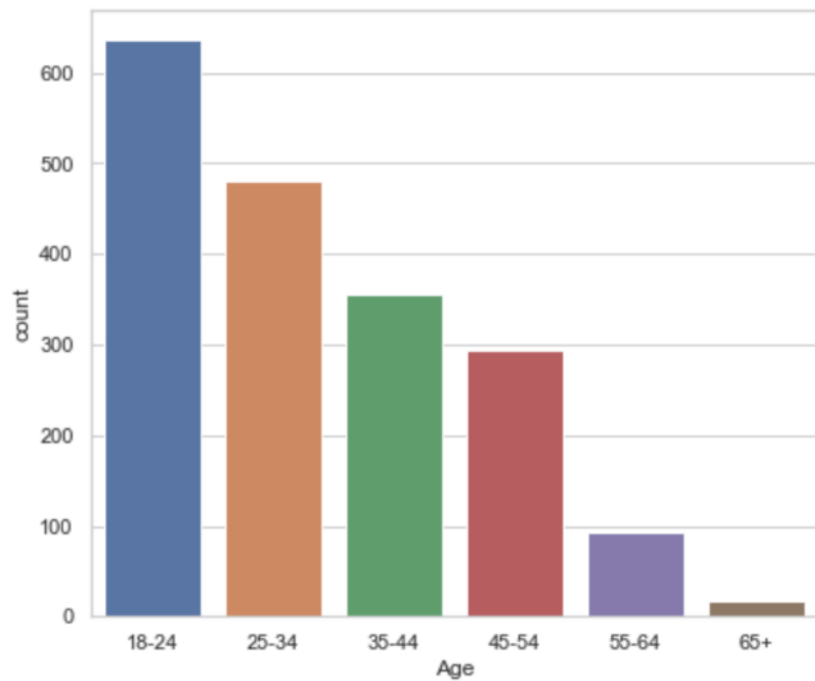
C. 목적변수

알코올, 암페타민, 헤로인 등 총 18가지 약물 간의 관계 및 그에 대한 분석



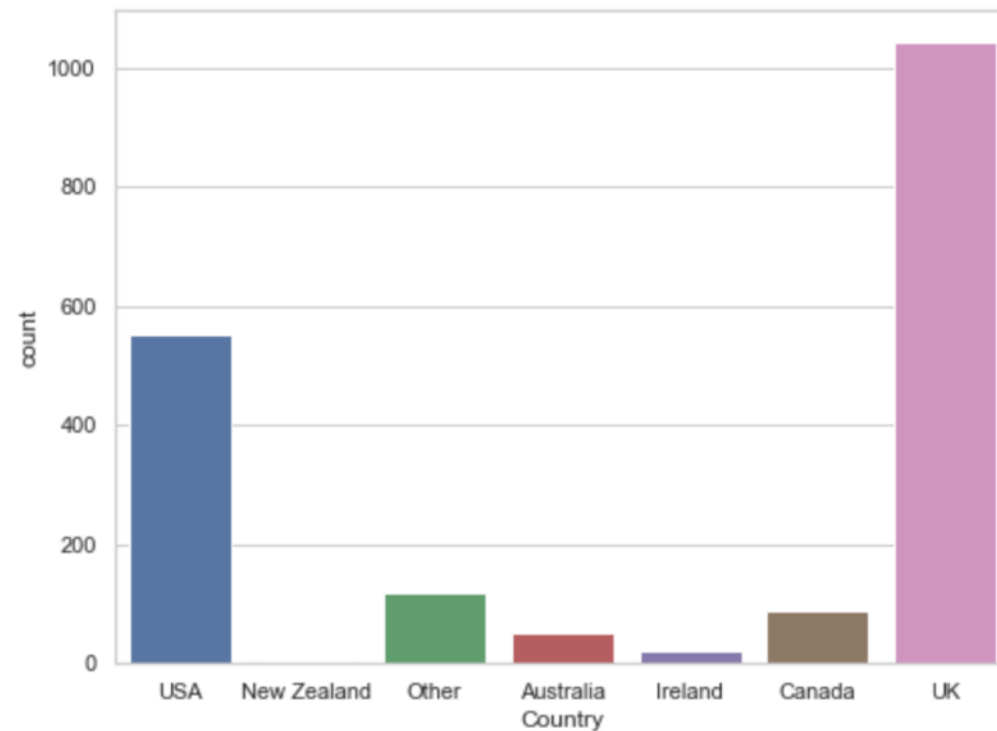
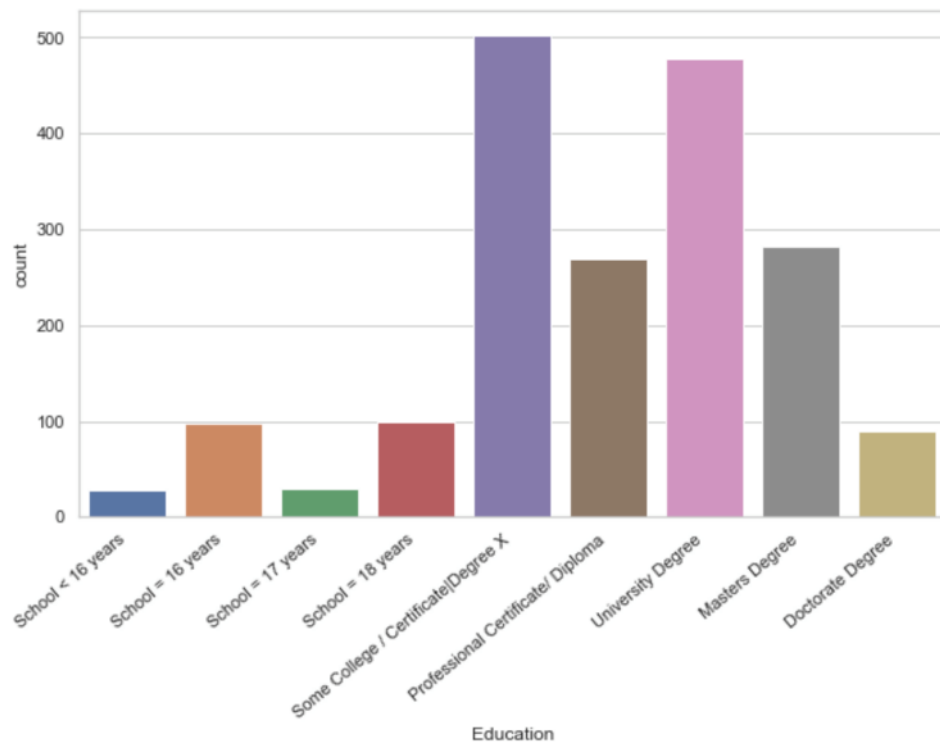
데이터 분석

A 인구통계



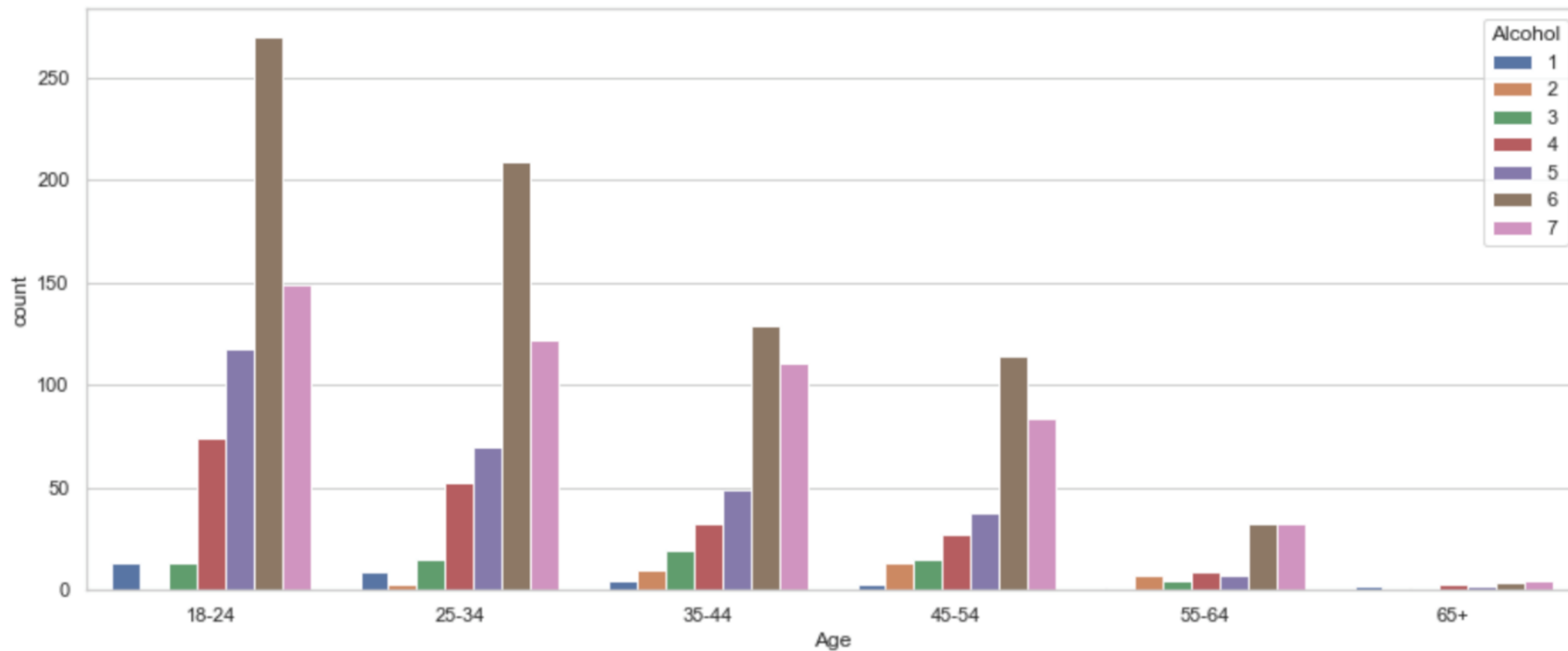
데이터 분석

A 인구통계



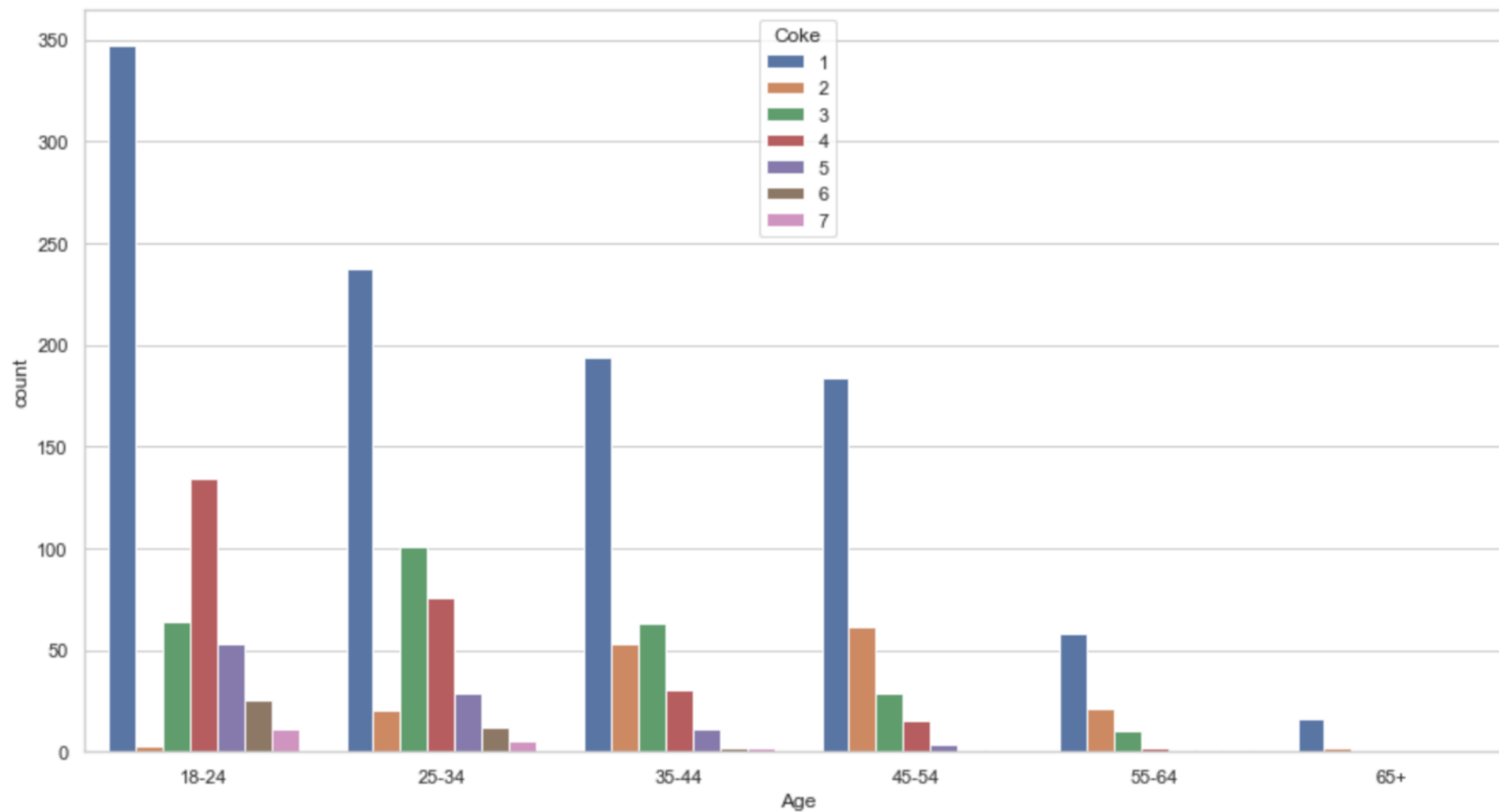
데이터 분석

A 인구통계



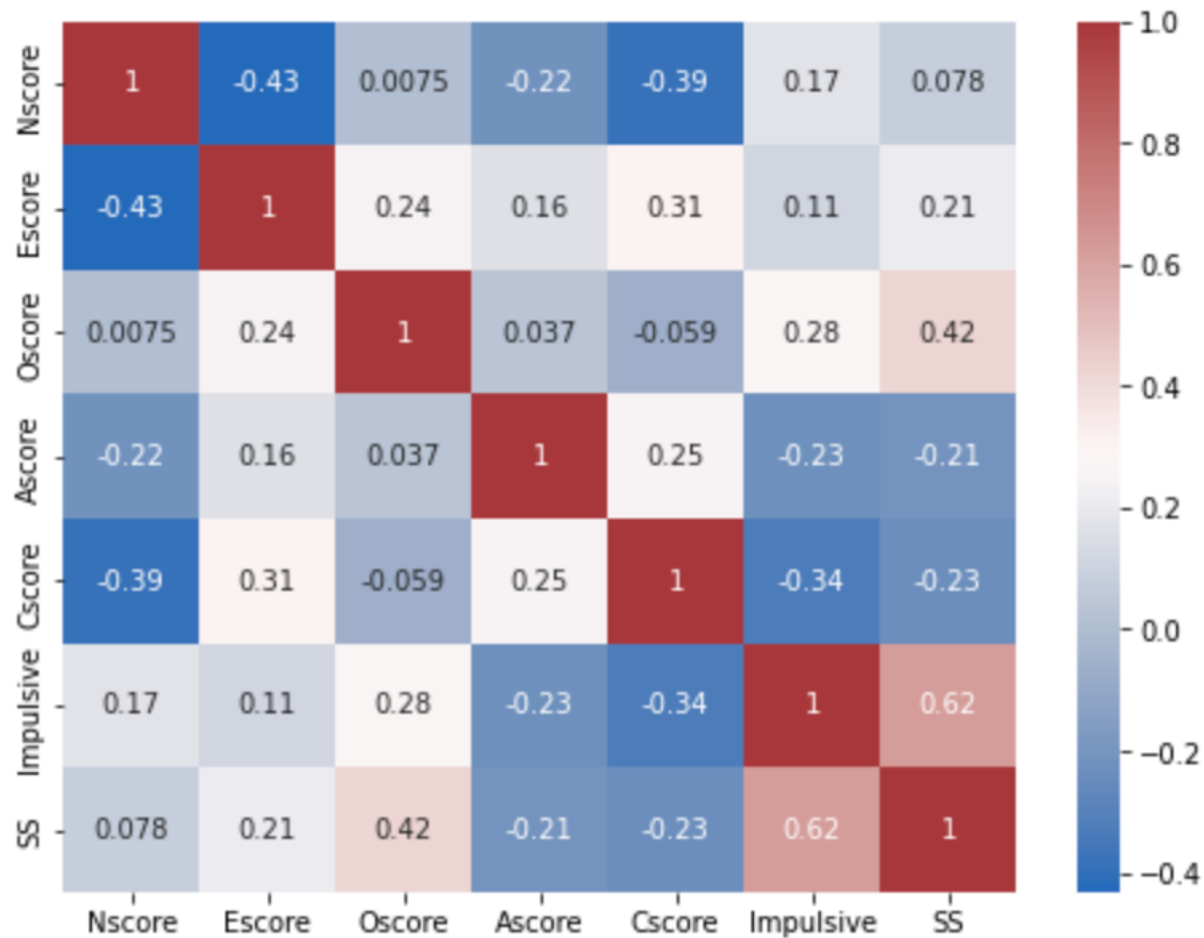
데이터 분석

A 인구통계



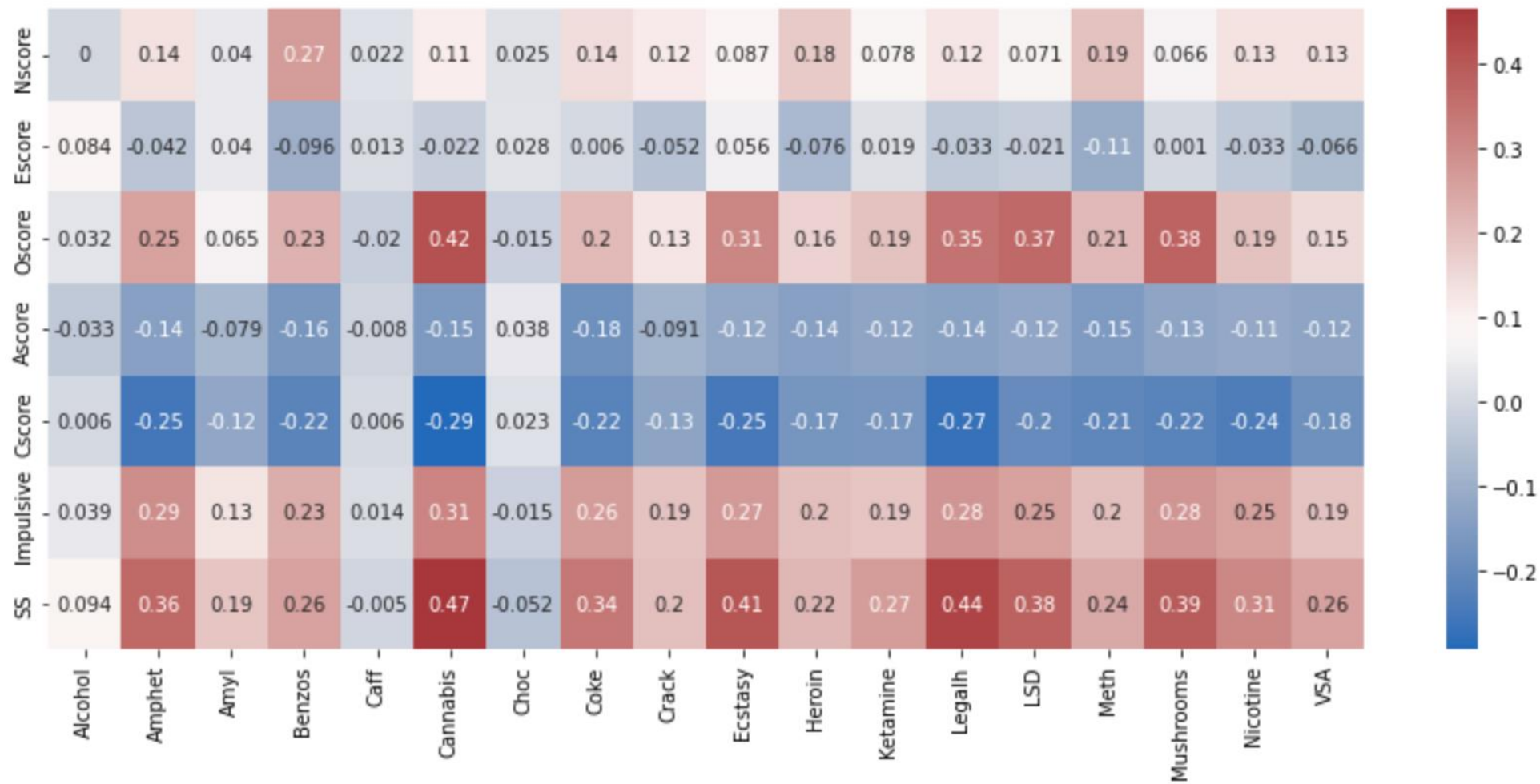
데이터 분석

B. 성격지표



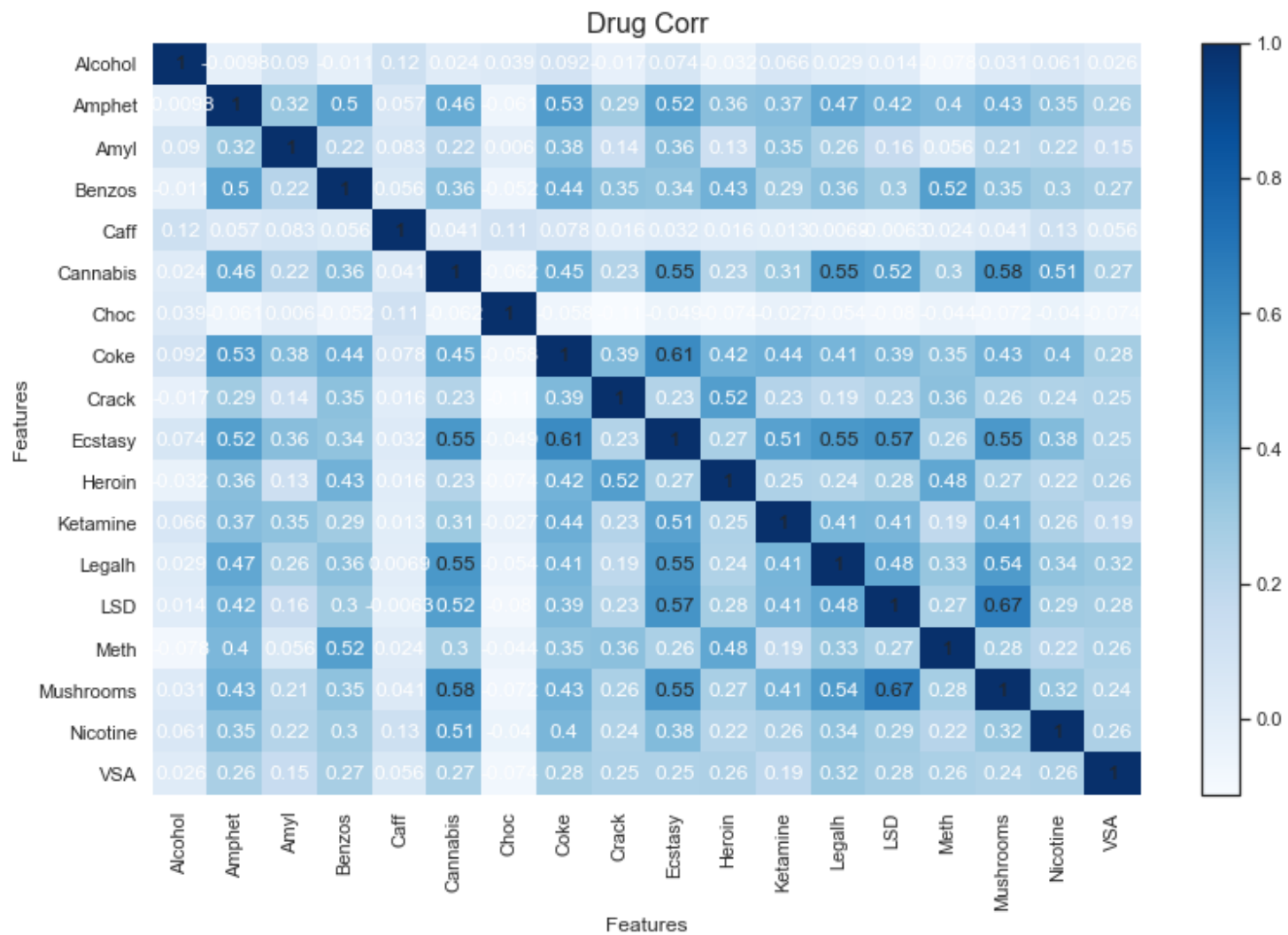
데이터 분석

B. 성격지표



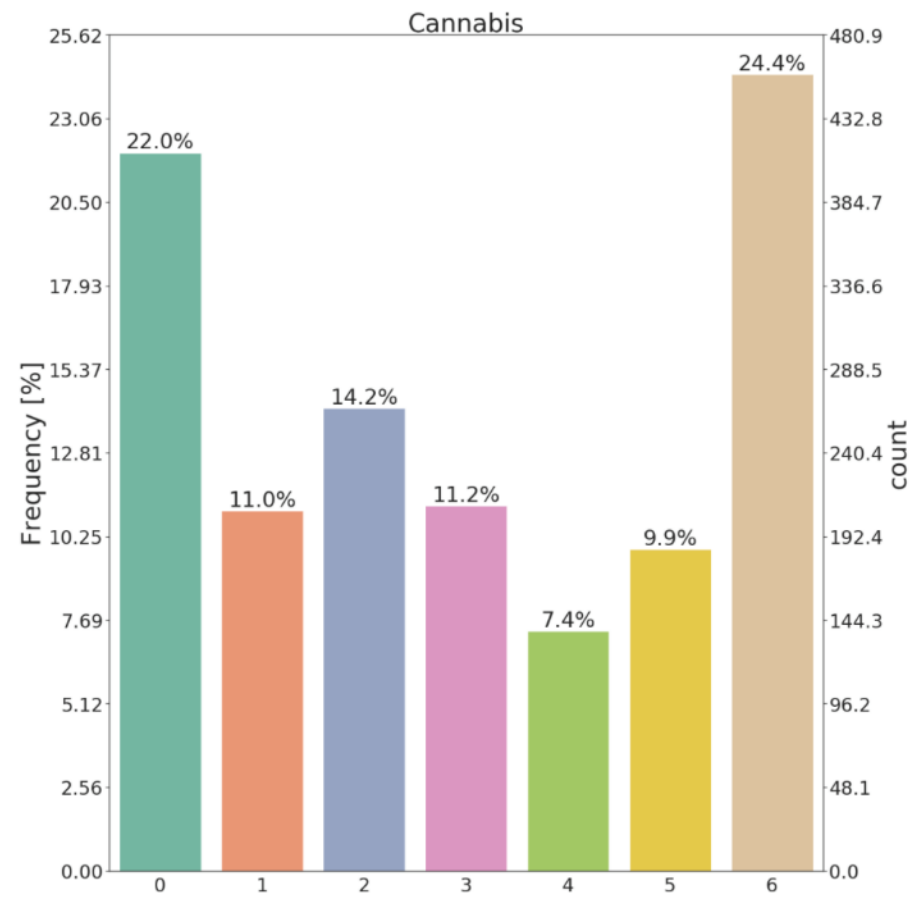
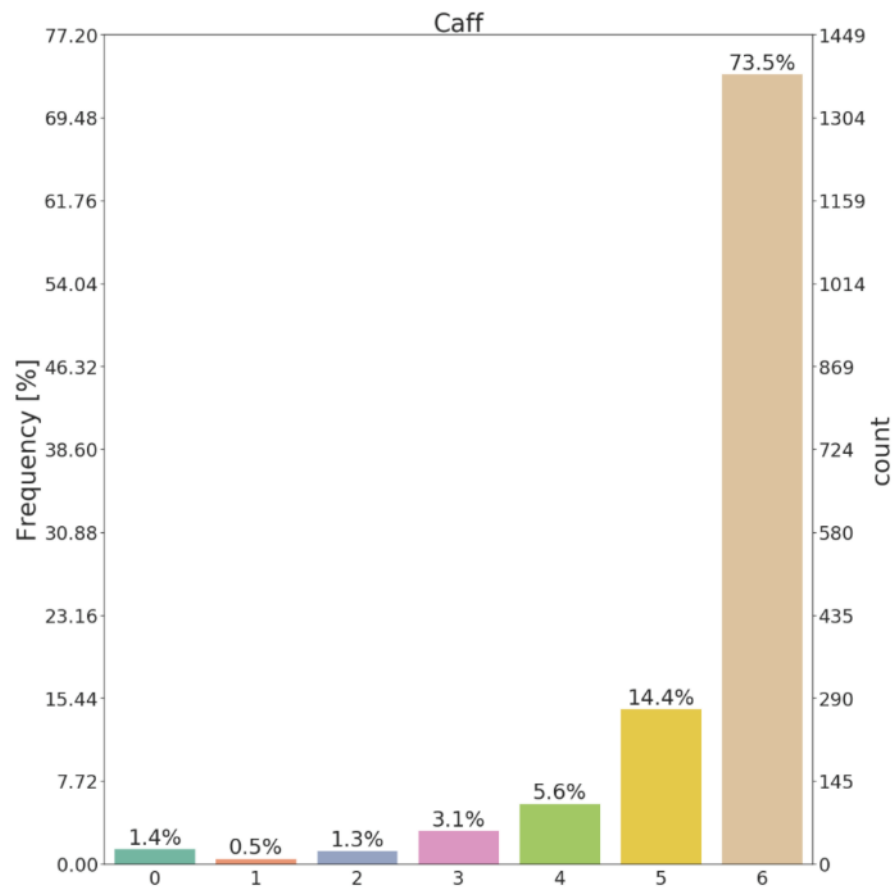
데이터 분석

C 목적변수



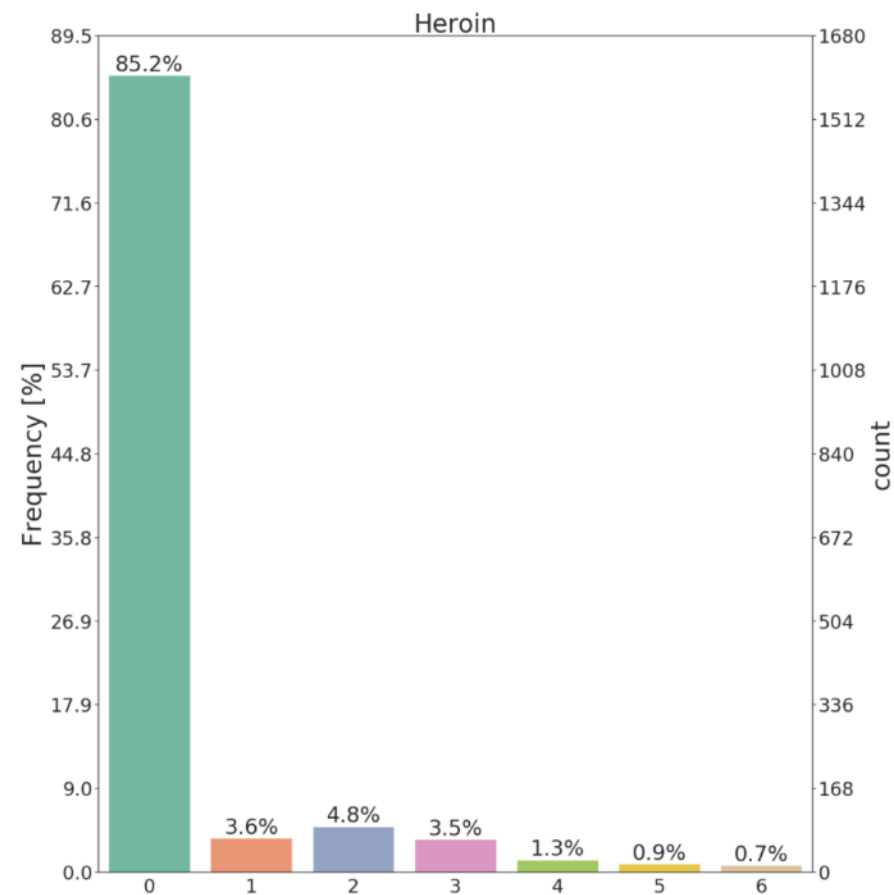
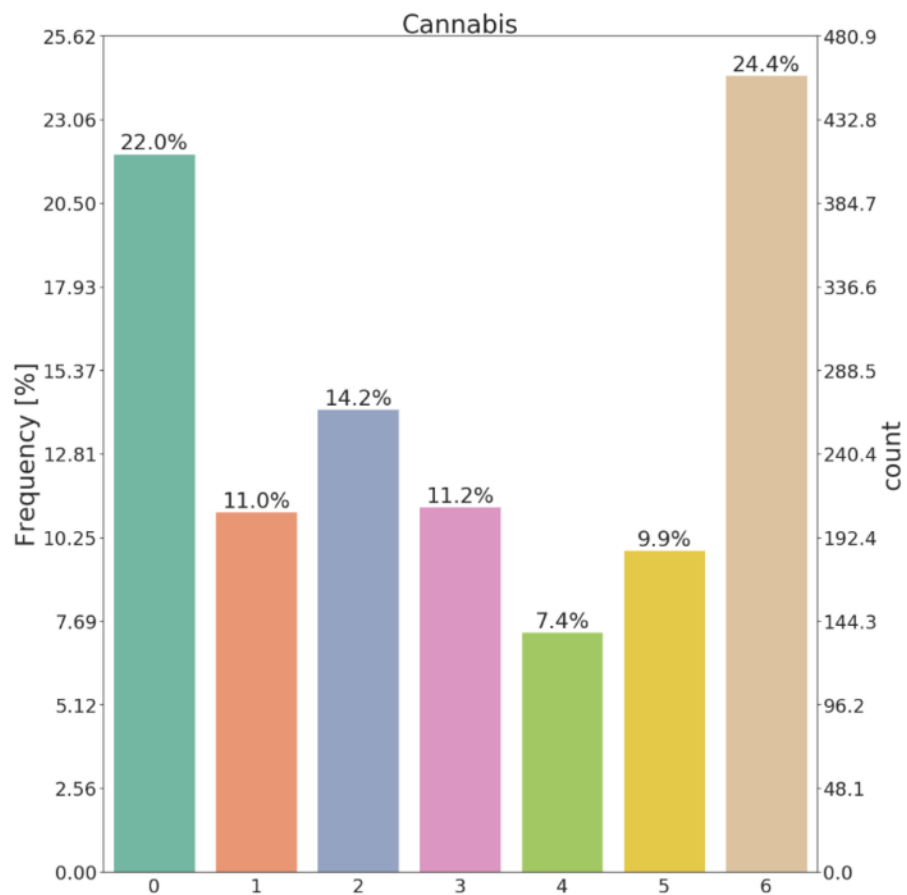
데이터 분석

C 목적변수



데이터 분석

C 목적변수



향후 추진 계획

설명변수 간의 독립성

- Personality & Demography 그리고 PCA

자료의 편향성

- Imbalanced data : White people, CL0

목적변수 처리 및 응용

- Clustering

모델링

- Individual model vs Softmax

thx