**개요**

**• 프로젝트 목표**

가지고 있는 Noom 고객 10,000명의 데이터(고객번호, 이름, 나이, 구매이력, 몸무게, 목표 몸무게 등)를 이용하여 이후 사내 운영, 기획, 마케팅, 코칭 팀에 도움이 될만한 인사이트를 도출한다. 개인적으로는 주어진 데이터를 활용한 데이터 분석과 시각화를 통한 데이터분석에 관한 전반적인 학습을 진행한다.

**• 가설과 프로젝트 해결 방향**

1. 현재 고객의 retention rate을 높인다
2. 잠재 고객을 유치하기 위해 효율적인 마케팅 target 집단을 찾기

**분석**

**• 서비스 현황**

* 코치 서비스: 가장 중요한 서비스로 꼽히는 것으로, 식단과 운동에 대한 기록을 남겨 본인이 건강한 식습관 생활을 갖추고 있는지 스스로 판단할 수 있다. 유료 모델에 가입하면 전문적인 코치가 채팅으로 1대 1로 식단과 운동을 코칭하여 주는 서비스를 운영 중

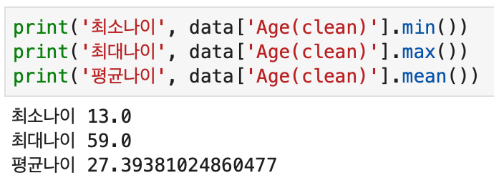
**스크린샷이(가) 표시된 사진

자동 생성된 설명• 데이터 로딩**

**• 칼럼의 데이터 정보**

1. Access code: 고객식별정보
2. Name: 고객 성명
3. Gender: 성별
4. Age: 나이
5. Height: 키
6. Initial Weight: 현재 몸무게
7. Lowest Weight: 눔 코치 서비스를 이용하던 중 입력했던 몸무게 중 가장 낮은 값
8. Target Weight: 목표 몸무게
9. Product Name: 눔 코치 상세 제품명
10. Status: 유료 서비스 결제내용 (complete: 결제중, cancelled: 결제취소, refunded: 환불완료)
11. Price: 서비스 구입시 가격 (₩)
12. Purchased At: 서비스 구입 시간
13. Payment Type: 결제 방식 (Recurring: 구독형)
14. Channel: 서비스 구입 경로

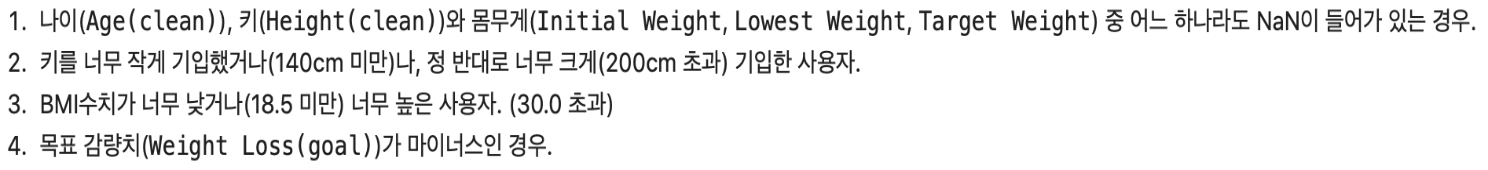
**• 전처리: 성별, 키, 나이**

1. 성별: 현재 성별은 FEMALE, MALE, NaN으로 구성되어 있다. 가독성을 높이기 위해 female, male로 바꿔준다.
2. 키: 가장 작은 키는 -1가 나온다. 잘못 입력된 값으로 생각하고 NaN으로 바꿔준다.
3. 나이: 최소 연령은 0세, 최고 연령은 173세라는 값이 나온다. 0세인 값은 NaN으로, 60세 이상의 값도 NaN으로 바꿔준다. 정리하고 난 다음 최소, 최고 연령을 확인한다.

**• 데이터 분석: 현재 고객 분석**

1. 유료 사용자 중 사용자 정보를 잘못 기입한 사람을 찾기

서비스 이용의 만족도를 올리고, 이것을 통해 고객 retention rate을 높일 수 있도록 유료 사용자 중에 사용자 정보(나이, 키, 몸무게 등)를 잘못 기입한 사람을 찾아내기

* **사용자 정보 오기입 기준**

각각의 칼럼 정보를 이용하여, Height(Invalid), Weight(Invalid), Age(Invalid) 라는 칼럼 만들고, 하나라도 Invalid한 정보가 들어가 있다면 정보를 잘못 기입한 고객으로 분류

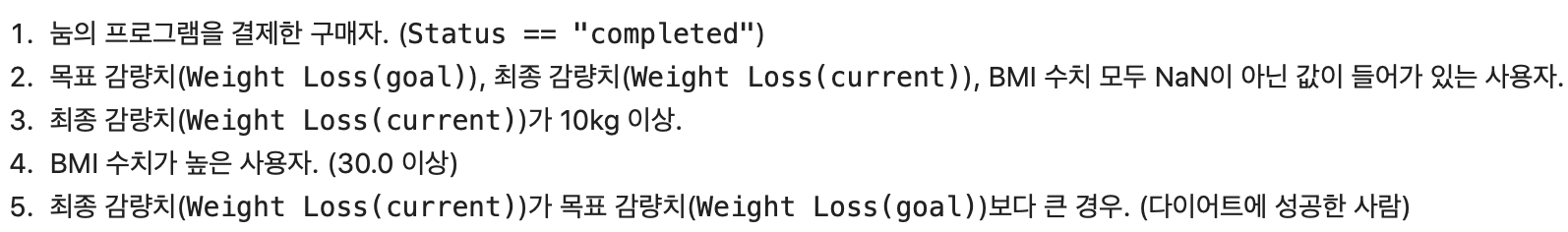
스크린샷이(가) 표시된 사진

자동 생성된 설명

1. VIP 사용자 찾아내기

VIP 사용자로 간주할 수 있는 고객을 찾아내어 홍보 수단으로 사용할 수 있는 가능성을 타진해보거나 추가 혜택을 제공하여 retention rate을 높인다.

* **VIP 사용자 선정 기준**



1. 기존에 만들어 놓은 weight\_history라는 Data set을 이용해서 vip\_list라는 새로운 data set만들기

개체이(가) 표시된 사진

자동 생성된 설명

1. vip를 선정하여 VIP인지 여부를 나타내는 column을 추가하기

스크린샷이(가) 표시된 사진

자동 생성된 설명

1. VIP 고객 명단 출력

스크린샷이(가) 표시된 사진

자동 생성된 설명

도출한 15명의 VIP 명단을 이용하여, 추가 혜택을 준다던지, 이후 marketing에 활용할 수 있는 방안을 고려할 수 있는 가능성이 있다.

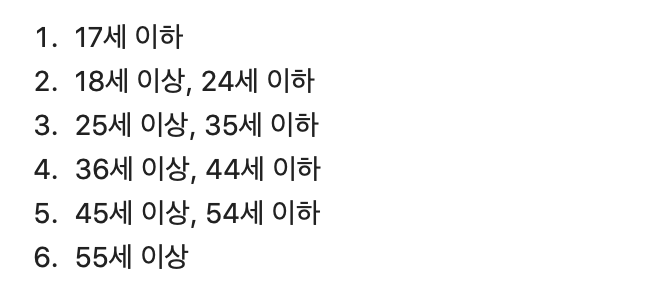
• 데이터 분석: 잠재 고객 유치

1. 성별, 연령으로 사용자가 많은 집단 알아내기

성별과 연령으로 고객을 분류하여, 사용자가 많은 집단을 알아낸다면 이후 잠재 고객을 유치하기 위해 마케팅 채널을 선정하는 등 마케팅 방법에 도움이 될 것.

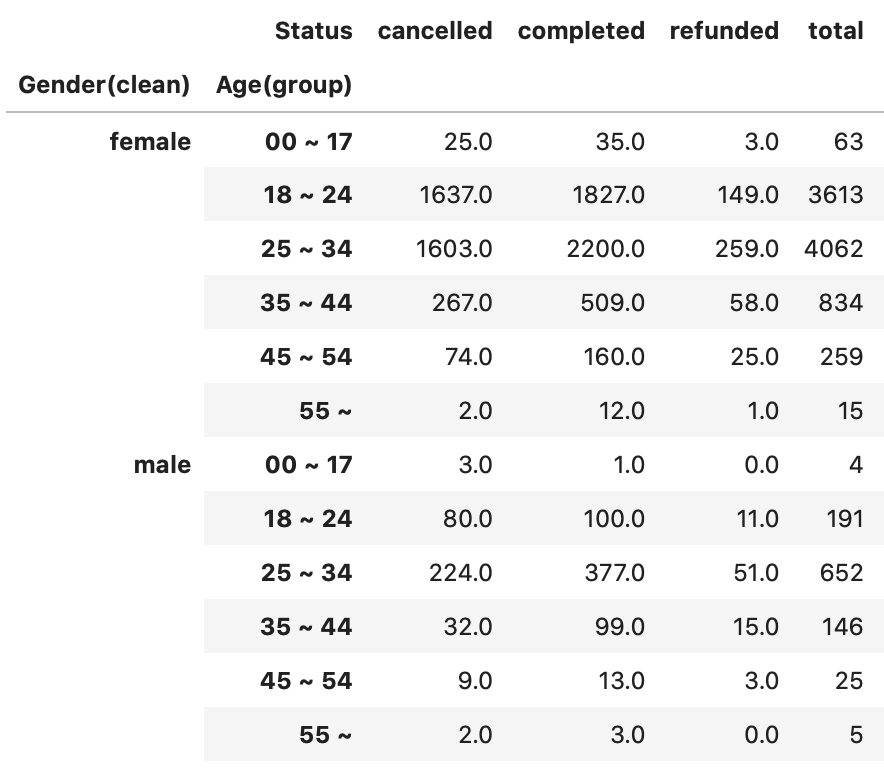
[성별, 연령] 구분 기준

[성별] 구분: [Female, Male]

[연령] 구분:

1. 스크린샷이(가) 표시된 사진

   자동 생성된 설명나이를 구간값으로 분류하기
2. 스크린샷이(가) 표시된 사진

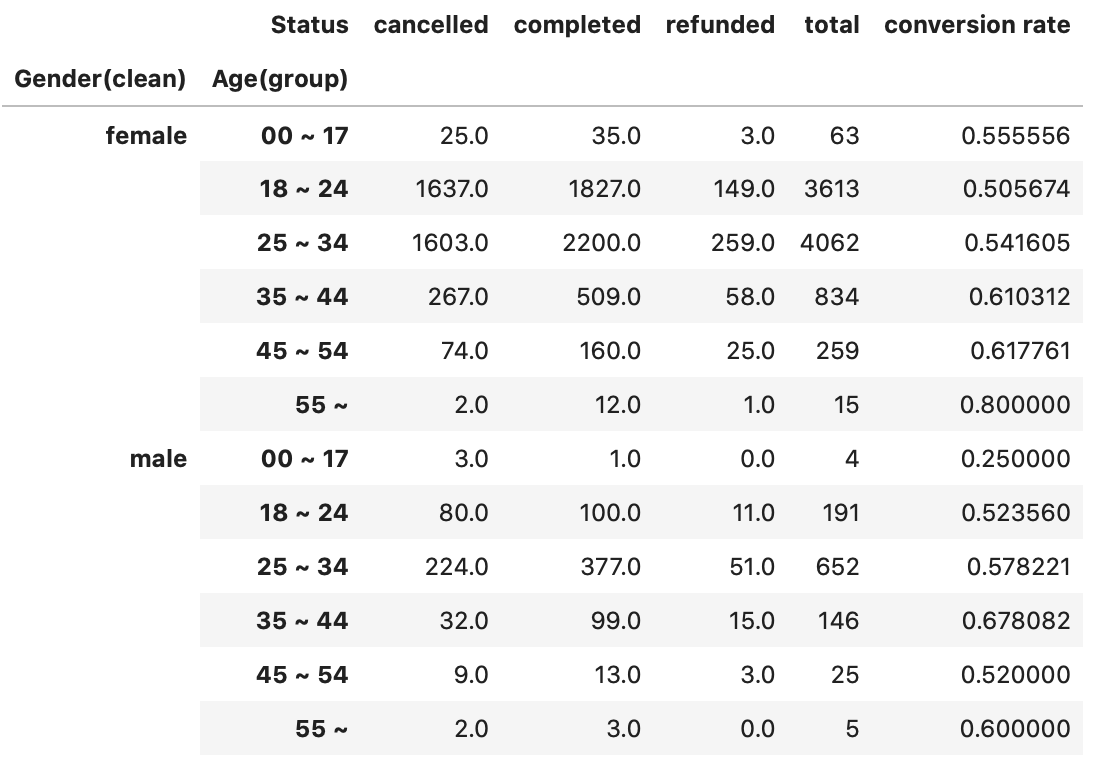
   자동 생성된 설명피벗테이블을 만들어 성별과 연령 구간값 별 Status (cancelled, completed, refunded) 정리
3. 전체 사용자의 수를 알기 위해 Status 총합 칼럼 추가하기
4. 스크린샷이(가) 표시된 사진

   자동 생성된 설명시각화 모듈을 활용하여 성별과 연령 구간값 별로 사용자 수를 보기 쉽게 나타내기

여성사용자가 9할에 이르고, 여성 사용자 중에서는 18세 이상 24세 이하, 25세 이상 34세 이하 집단의 사용자가 많다.

1. 성별, 연령으로 conversion rate이 높은 집단 알아내기

Conversion rate이 높은 고객 집단을 알아내는 것은 단순히 사용자가 많은 집단을 알아내는 것보다 한발짝 더 나아간 targeted marketing에 도움이 될 것.

1. 앞에서 만들어놓은 피벗테이블에 conversion rate이라는 컬럼을 추가하기
2. Conversion rate이 추가된 테이블 확인
3. 지도, 텍스트이(가) 표시된 사진

   자동 생성된 설명시각화 모듈을 활용하여, 성별, 연령 구간별 conversion rate 보기 쉽게 나타내기

**시사점**

**• 현재 고객의 retention rate을 높일 수 있는 방법**

1. 5000명이 넘는 구매 고객 중 정보가 잘못 기입되었다고 판단되는 고객들에게 연락을 취하여, 올바른 정보를 기입하여 제대로된 서비스를 받을 수 있는 기회를 제공하고, 서비스에 대한 만족도를 높일 수 있도록 한다. 이러한 방법을 통해서 현재 고객의 retention rate을 높일 수 있을 것.
2. VIP 사용자로 선정된 15명의 고객 명단을 활용하여 추가 혜택을 주어, 충성 고객을 유지하고 해당 고객들의 정보를 활용하여 추후 마케팅에 활용할 수 있는 방안을 강구한다.

**• 잠재 고객을 끌어 들일 수 있는 방법**

1. 성별로 구분했을 때는 여성 사용자가 남성 사용자에 비해 월등히 많고(약 9:1 의 비율), 나이를 구간값으로 구분했을 때, 여성 사용자 중에서도 18~24세, 25~34세의 사용자가 총 사용자의 70%에 이른다. 사용자의 성별과 연령층을 고려하여 마케팅 채널과 방법을 선택할 수 있을 것
2. 성별과 연령 구간값 별로 conversion rate을 고려하면, 여성 25~34세, 35~44세, 남성 35~44세의 사용자 집단에서 높은 conversion rate을 보이므로 사용자 수와 conversion rate을 고려하여 마케팅 비용 지출에 대한 결정을 할 수 있을 것이다.