**Categorical data 정리**

**Categorical data로 보기 어려운 것 – 0,1,2 등의 값으로 판단하기 어려운 것 !**

Feature Target

-Personal identifiers(개인 식별자), TOO\_ID

-Unit data(단위 데이터), TOO\_DATA\_CLASS

-Survey date YYYYMM(조사일자), T01\_EDATE

-Age Years old(나이), T01\_AGE

-Amount of Rice wine drink per Drinking (막걸리 한 잔당 얼마나 마시는지) , T01\_TAKAM

-Amount of Sake drinking at once Glass (사케 한 잔당 얼마나 마시는지), T01\_RICEAM

-Amount of Wine Drinking at one Glass (와인 한 잔당 얼마나 마시는지) , T01\_WINEAM

-Amount of soju drinking at once Glass (소주 한 잔당 얼마나 마시는지), T01\_SOJUAM

-Amount of Beer drinking at once Bottle (맥주 한 병당 얼마나 마시는지), T01\_BEERAM

-Amount of spirits drinking Shot (독한 술 한 번에 얼마나 마시는지), T01\_HLIQAM

-Smoking starting age years old (흡연을 처음 시작한 나이), T01\_SMAG

-The total duration of Smoking in Total years (총 흡연 기간(년)), T01\_SMDU

-Amount of smoking per day pack (1일 흡연 팩의 개수), T01\_SMAM

-Age when first diagnosed with hypertension years old Diagnosed with Diabetes (고협압을 처음으로 진단받은 나이), T01\_HTNAG

-Age when diagnosed with hyperlipidemia years old (고지혈증 진단 시 연령), T01\_LIPAG

-Age when father was first diagnosed with hypertension years old (아버지가 고혈압 진단을 처음 받은 나이), T01\_FMFHTAG

-Age when mother was first diagnosed with hypertension years old (어머니가 고혈압 진단을 처음 받은 나이), T01\_FMMHTAG

-Age when father was first diagnosed with diabetes years old (아버지가 당뇨병 진단을 처음 받은 나이), T01\_FMFDMAG

-Age when mother was first diagnosed with diabetes years old (어머니가 당뇨병 진단을 처음 받은 나이) , T01\_FMFDMAG

-Menarche age years old (초경 경과 년도), T01\_MNSAG

-Age at first pregnancy years old (첫 임신 나이), T01\_FPREGAG

-Age at first childbirth years old (첫 아이 출산 나이), T01\_FLABAG

-Menopaused age years old (폐경 경과 년도), T01\_PMAG\_C

-Height cm (키), T01\_HEIGHT

-Weight kg (몸무게), T01\_WEIGHT

-The average waist circumference cm (평균 허리 둘레), T01\_WAIST

-The average hip circumference cm (평균 둔부 둘레) , T01\_HIP

-Average heart rate Times / min (평균 심박수 횟수/분) , T01\_PULSE

-Systolic blood pressure mmHg (수축기 혈압) , T01\_SBP

-Diastolic blood pressure mmHg (이완 혈압) , T01\_DBP

-HbA1C % (헤모글로빈 수치) , T01\_HBA1C

-Fasting Glucose mg / dL (단식 포도당) , T01\_GLU0

-Creatinine mg / dL (크레아니닌) , T01\_CREATININE

-AST IU / L (아스파르테이트아미노전달효소) , T01\_AST

-ALT IU / L (알라닌아미노전달효소) , T01\_ALT

-Total Cholesterol mg / dL (총 콜레스테롤) , T01\_TCHL

-HDL-Cholesterol mg / dL (HDL 콜레스테롤) , T01\_HDL

-Triglyceride mg / dL (트리글리세라이드) , T01\_TG

-Fasting Insulin µIU / m (단식 인슐린) , T01\_INS0

**Categorical data** – **Ordinal encoding이 이미 되어있는 데이터들 !**

**- Gender(성별)** ( 1 = Male, 2 = Female ), **T01\_SEX**

**- Level of education(교육 수준), T01\_EDU**

1 = elementary school or less(초등학교 이하),

2 = middle school(중학교),

3 = High school(고등학교),

4 = Bachelors Degree = 5(학사 학위),

6 = Graduate school(대학원)

- **Marriage status** **(혼인 여부), T01\_MARRY**

( 1 = Single, 2 = Married, 3 = Bereavement(사별), 4 = Separated(분리), 5 = Divorce(이혼), 6 = Other)

- **Alcohol drinking (음주 수준), T01\_DRINK** (1 = Non-drinker, 2 = formal drinker, 3 = regular drinker)

- **The total period of drinking** **(총 음주 경력)** , **T01\_DRDU**

(1 = 5 years or less, 2 = 6-10 years, 3 = 11-20 years, 4 = > 21 years)

- **Rice wine drinking frequency** **(막걸리 음주 빈도)**, **T01\_TAKFQ**

0 = None,

1 = once a month,

2 = two or three times a month,

3= once a week,

4 = 2-3 times a week,

5 = 4-6 times a week,

6 = Everyday

- **Glass Sake drinking frequency** **(사케 음주 빈도), T01\_RICEFQ**

0 = None.

1 = once a month,

2 = two or three times a month,

3= once a week,

4 = 2-3 times a week,

5 = 4-6 times a week,

6 = Everyday

- **Wine Drinking Frequency (와인 음주 빈도), T01\_WINEFQ**

0 = None.

1 = once a month,

2 = two or three times a month,

3= once a week,

4 = 2-3 times a week,

5 = 4-6 times a week,

6 = Everyday

- **Soju Drinking Frequency (소주 음주 빈도), T01\_SUJUFQ**

0 = None.

1 = once a month,

2 = two or three times a month,

3= once a week,

4 = 2-3 times a week,

5 = 4-6 times a week,

6 = Everyday

- **Beer Drinking Frequency (맥주 음주 빈도), T01\_BEERFQ**

0 = None.

1 = once a month,

2 = two or three times a month,

3= once a week,

4 = 2-3 times a week,

5 = 4-6 times a week,

6 = Everyday

- **Spirits Drinking Frequency (독한 술 음주 빈도), T01\_HLIQFQ**

0 = None.

1 = once a month,

2 = two or three times a month,

3= once a week,

4 = 2-3 times a week,

5 = 4-6 times a week,

1. = Everyday

- **Smoking status (흡연 여부), T01\_SMOKE** (1. non-smoker, 2 = formal smoker, 3 = regular smoker)

- **Secondhand smoke (간접 흡연 여부), T01\_PSM** (1 = no, 2 = yes)

- **Frequent high intensity exercise (고강도 운동), T01\_EXER** (1 = no, 2 = yes)

- **Hypertension history (고혈압 이력), T01\_HTN** (1 = no, 2 = yes)

- **Diagnosed with Diabetes** **(당뇨병 진단), T01,DM** (1 = no, 2 = yes)

- **Diagnosed with Hyperlipidemia** **(고지혈증 진단), T01\_LIP** (1 = no, 2 = yes)

- **Family history of hypertension (father, mother) (고혈압 아버지, 어머니 가족력), T01\_FMFHT, T01\_FMMHT** (1 = no, 2 = yes)

- **Family history of diabetes (father, mother)** **(당뇨병 아버지, 어머니 가족력), T01\_FMFDM, T01\_FMMDM (**1 = no; 2 = yes)

**- Pregnancy experience (임신 경험),** **T01\_PREG**(1 = no; 2 = yes)

**- Childbirth (출산), T01\_CHILD** (1 = no and 2 = yes)

- **Menopause (갱년기), T01\_PMYN\_C** (1 = no; 2 = yes)