

	A	B	C	AN	AO	AP	AQ	AR	AS	AT	AU	AV	AW	AX	AY	AZ	BA	BB	
1	ID	시술 시기	시술 당시	총 생성 배	미세주입된 배	미세주입 배	이식된 배	미세주입 배	저장된 배	미세주입 배	해동된 배	해동 난자	수집된 난자	저장된 난자	혼합된 난자	파트너 정자	기증자 정자	난자 출처	
2	TRAIN_001	TRZKPL	만18-34세	4	5	4	2	2	2	0	0	7	0	5	5	0	본인 제공		
3	TRAIN_001	TRYBLT	만45-50세	0	1	0	0	0	0	0	0	1	0	1	1	0	본인 제공		
4	TRAIN_001	TRVNRV	만18-34세	5	0	0	2	0	0	0	0	8	0	7	7	0	본인 제공		
5	TRAIN_001	TRJXFG	만35-37세	0	4	0	0	0	0	0	0	5	0	4	4	0	본인 제공		
6	TRAIN_001	TRVNRV	만18-34세	6	6	6	2	2	0	0	0	7	0	6	6	0	본인 제공		
7	TRAIN_001	TRVNRV	만38-39세	3	5	3	2	2	0	0	0	5	0	5	5	0	본인 제공		
8	TRAIN_001	TRJXFG	만18-34세	4	4	4	1	1	3	3	0	0	0	4	4	0	기증 제공		
9	TRAIN_001	TRXQMD	만40-42세	12	0	0	2	0	0	0	0	15	0	15	15	0	본인 제공		
10	TRAIN_001	TRYBLT	만38-39세	0	0	0	0	0	0	0	0	0	0	0	0	0	본인 제공		
11	TRAIN_001	TRYBLT	만18-34세	2	0	0	2	0	0	0	0	2	0	2	2	0	본인 제공		
12	TRAIN_001	TRCMWS	만40-42세	0	0	0	1	0	0	0	1	0	0	0	0	0	본인 제공		
13	TRAIN_001	TRZKPL	만40-42세	2	2	2	2	2	0	0	0	3	0	2	0	2	본인 제공		
14	TRAIN_001	TRYBLT	만38-39세	0	0	0	1	0	0	0	1	0	0	0	0	0	본인 제공		
15	TRAIN_001	TRDQAZ	만40-42세	13	0	0	2	0	3	0	0	15	0	15	15	0	본인 제공		
16	TRAIN_001	TRXQMD	만38-39세														알 수 없음		
17	TRAIN_001	TRDQAZ	만43-44세	3	3	3	2	2	0	0	0	3	0	3	3	0	본인 제공		
18	TRAIN_001	TRDQAZ	만43-44세	0	0	0	1	0	0	0	1	0	0	0	0	0	본인 제공		
19	TRAIN_001	TRZKPL	만18-34세	3	5	3	2	2	0	0	0	8	0	5	5	0	본인 제공		
20	TRAIN_001	TRDQAZ	만35-37세	3	5	3	1	1	0	0	0	5	0	5	5	0	본인 제공		
21	TRAIN_001	TRJXFG	만35-37세	0	0	0	1	0	0	0	6	0	0	0	0	0	본인 제공		
40	39	총 생성 배아 수					해당 시술에서 생성된 배아의 총 개수												0
41	40	미세주입된 난자 수					정자 미세주입(ICSI)으로 처리된 난자의 수												0
42	41	미세주입에서 생성된 배아 수					ICSI를 통해 생성된 배아의 수												0
43	42	이식된 배아 수					이식된 배아의 총 수												0
44	43	미세주입 배아 이식 수					ICSI 배아 중 이식된 배아의 수												0
45	44	저장된 배아 수					환자를 위해 저장된 배아의 수												0
46	45	미세주입 후 저장된 배아 수					ICSI 처리된 배아 중 저장된 배아의 수												0
47	46	해동된 배아 수					동결 해동된 배아의 수												0
48	47	해동 난자 수					동결 해동된 난자의 수												0
49	48	수집된 신선 난자 수					신선 난자가 수집된 수												0
50	49	저장된 신선 난자 수					저장된 신선 난자의 수												0
51	50	혼합된 난자 수					정자와 혼합된 총 난자 수												0
52	51	파트너 정자와 혼합된 난자 수					파트너 정자와 혼합된 난자의 수												0
53	52	기증자 정자와 혼합된 난자 수					기증자 정자와 혼합된 난자의 수												0
	A	B	C	AN	AO	AP	AQ	AR	AS	AT	AU	AV	AW	AX	AY	AZ	BA	BB	
1	시술 시기	시술 당시	총 생성 배	미세주입된 배	미세주입 배	이식된 배	미세주입 배	저장된 배	미세주입 배	해동된 난자	수집된 난자	저장된 난자	혼합된 난자	파트너 정자	기증자 정자	난자 출처	정자 출처	난자 기증자	
2	TRZKPL	만18-34세	4	5	4	2	2	2	0	0	7	0	5	5	0	본인 제공	배우자 제; 알 수 없음	알 수 없음	
3	TRYBLT	만45-50세	0	1	0	0	0	0	0	0	1	0	1	1	0	본인 제공	배우자 제; 알 수 없음	알 수 없음	
4	TRVNRV	만18-34세	5	0	0	2	0	0	0	0	8	0	7	7	0	본인 제공	배우자 제; 알 수 없음	알 수 없음	
5	TRJXFG	만35-37세	0	4	0	0	0	0	0	0	5	0	4	4	0	본인 제공	배우자 제; 알 수 없음	알 수 없음	
6	TRVNRV	만18-34세	6	6	6	2	2	0	0	0	7	0	6	6	0	본인 제공	배우자 제; 알 수 없음	알 수 없음	
7	TRVNRV	만38-39세	3	5	3	2	2	0	0	0	5	0	5	5	0	본인 제공	배우자 제; 알 수 없음	알 수 없음	
8	TRJXFG	만18-34세	4	4	4	1	1	3	3	0	0	0	4	4	0	기증 제공	배우자 제; 만21-25세	알 수 없음	
9	TRXQMD	만40-42세	12	0	0	2	0	0	0	0	15	0	15	15	0	본인 제공	배우자 제; 알 수 없음	알 수 없음	
10	TRYBLT	만38-39세	0	0	0	0	0	0	0	0	0	0	0	0	0	본인 제공	배우자 제; 알 수 없음	알 수 없음	
11	TRYBLT	만18-34세	2	0	0	2	0	0	0	0	2	0	2	2	0	본인 제공	배우자 제; 알 수 없음	알 수 없음	
12	TRCMWS	만40-42세	0	0	0	1	0	0	0	1	0	0	0	0	0	본인 제공	배우자 제; 알 수 없음	알 수 없음	
13	TRZKPL	만40-42세	2	2	2	2	2	0	0	0	3	0	2	0	2	본인 제공	기증 제공	알 수 없음	
14	TRYBLT	만38-39세	0	0	0	1	0	0	0	1	0	0	0	0	0	본인 제공	배우자 제; 알 수 없음	알 수 없음	
15	TRDQAZ	만40-42세	13	0	0	2	0	3	0	0	15	0	15	15	0	본인 제공	배우자 제; 알 수 없음	알 수 없음	
16	TRXQMD	만38-39세	5.327932	4.36803	2.992655	1.334678	0.711784	1.374455	0.717719	0.438947	0.045333	8.977774	0.086881	8.076885	7.416205	0.547026	알 수 없음	기증 제공	알 수 없음
17	TRDQAZ	만43-44세	3	3	3	2	2	0	0	0	3	0	3	3	0	본인 제공	배우자 제; 알 수 없음	알 수 없음	알 수 없음
18	TRDQAZ	만43-44세	0	0	0	1	0	0	0	1	0	0	0	0	0	본인 제공	배우자 제; 알 수 없음	알 수 없음	알 수 없음
19	TRZKPL	만18-34세	3	5	3	2	2	0	0	0	8	0	5	5	0	본인 제공	배우자 제; 알 수 없음	알 수 없음	알 수 없음
20	TRDQAZ	만35-37세	3	5	3	1	1	0	0	0	5	0	5	5	0	본인 제공	배우자 제; 알 수 없음	알 수 없음	알 수 없음

위 데이터는 난임 수술을 받은 난임 환자들의 임신 성공 여부를 예측하기 위한 난임 환자들의 데이터입니다. 실제로는 67개의 피쳐, 256352개의 데이터가 있지만 일부만 가져왔습니다. 총 생성 배아 수 ~ 기증자 정자와 혼합된 난자 수 피쳐들의 결측값들이 데이터에 존재했고 이 결측값들을 처리하고자 데이터 전처리 과정이 필요했습니다. 저는 표본의 결측값을 학습셋에서 찾은 가장 가까운 n_neighbors 이웃의 평균값을 사용해서 대체하는 KNNImputer 방식으로 결측치를 처리했습니다. 저는 n_neighbtors=5로 하여서 전처리를 하였고 위와 같은 결과가 나왔습니다.