



스마트 축산 대시보드 구현

“나는 착유한다,
고로 존재한다”

TEAM 12조 이김했조

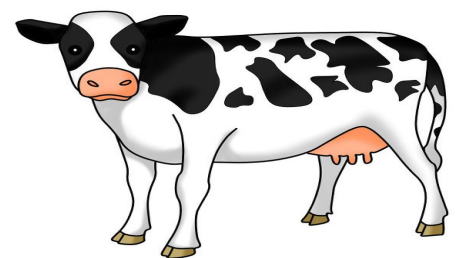
이하은, 김소영, 김우진, 이경민

튜터 : 김대영, 신한결

실전프로젝트

CONTENTS

- 01 데이터 소개
- 02 문제 정의 및 가설설정
- 03 분석
- 04 해석 및 적용



01 데이터 소개

▶ 컬럼 설명

종합 정보

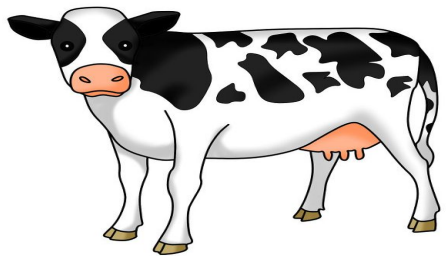
농장아이디	각 농장의 고유 식별 번호
개체번호	개별 소의 식별번호
축종코드	가축 종류를 나타내는 코드
제조사 아이디	착유기 제조사 정보

착유 기계 정보

착유회차	하루 중 몇 번째 착유인지 나타내는 값
착유시작일시, 착유종료일시	착유 시작 및 종료 시간
데이터 수집일자	데이터가 수집된 날짜
공기흐름	착유기 내부로 유입되는 공기량 또는 진공 흐름 상태

건강

착유량	착유된 우유의 양
전도도	<ul style="list-style-type: none">우유가 전기를 얼마나 잘 통하는지 나타내는 값유방염 진단 가능지표
혈액흐름	우유 내 혈액 혼합여부
온도	착유시 우유 온도 -> 소의 체온도 파악 가능
유지방	우유에 포함된 유지방 비율
유단백	우유에 포함된 유단백 비율



문제 정의 및 가설 설정

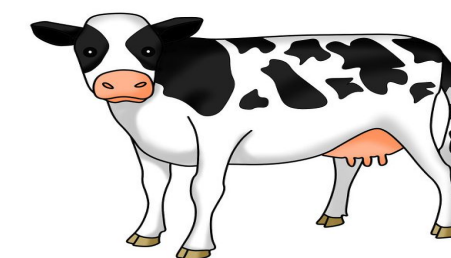
▶ 문제정의 & 목표 설정

문제정의

- 고온, 잦은 착유, 건강 이상(전도도 \uparrow 등) 등 다양한 요인이 착유량에 영향을 미침
- 데이터 분석 및 시각화 기반으로 건강 이상상태를 사전에 대응함

목표

- 착유량 최적화를 위한 범위 (온도, 전도도, 착유회차 등) 탐색
- 착유량 예측 머신러닝 모델을 구축
- 대시보드를 통해 농가가 실시간으로 생산성·품질을 모니터링하고 의사결정에 활용

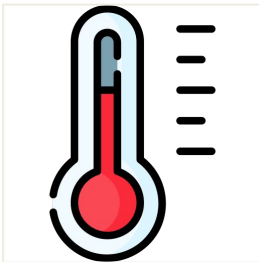


▶ 가설 설정

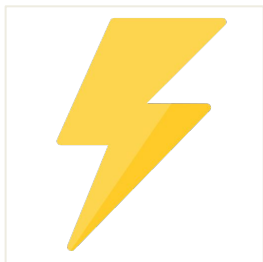
가설



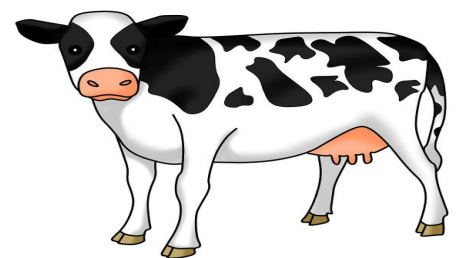
착유회차가 높을수록 평균 착유량이 감소할 것이다



온도가 높을수록 평균 착유량이 감소할 것이다



전도도가 높을수록 평균 착유량이 감소할 것이다



▶ 전처리



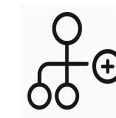
제거

- 사용하지 않는 컬럼 제거
- 결측치 존재하는 행 제거
- 중복 데이터 제거



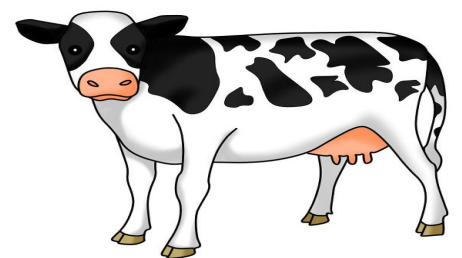
데이터 변환

- 농장아이디, 개체번호 카테고리화
- 혈액흐름 원-핫 인코딩



파생변수

- **유단백 대비 유지방 (PFR)**
 - 영양 상태를 진단 기준 중 하나
 - ['유단백'] / ['유지방']
- **착유 소요 시간 (착유소요시간 (분))**
 - ['착유종료일시'] - ['착유시작일시']
 - 두 시간의 차이를 분으로 환산



분석_가설검증

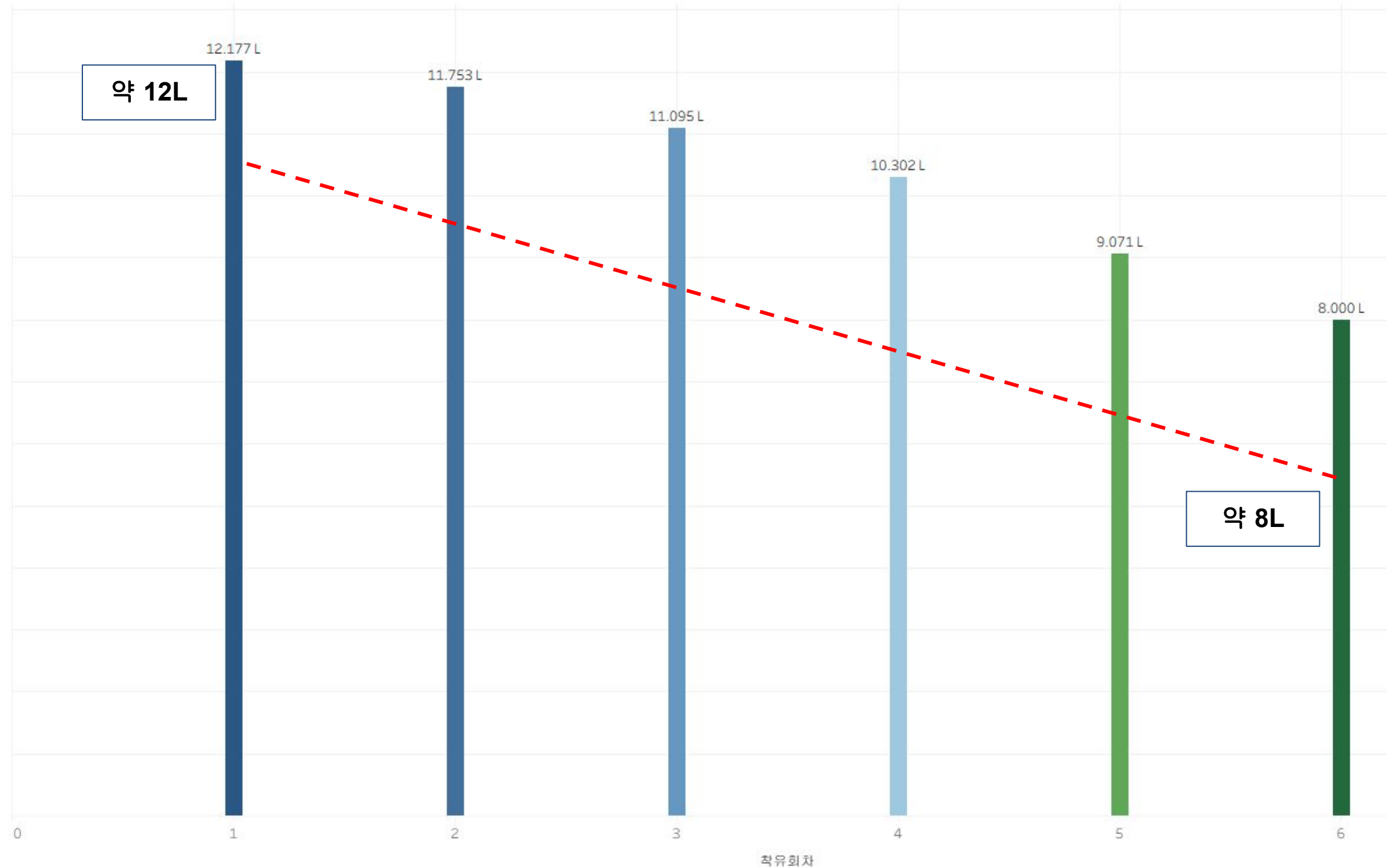
- ▶ 가설 1 (착유회차가 높을수록 평균 착유량이 감소할 것이다)

“착유회차가 높을수록
평균 착유량이 감소할 것이다”



“가설과 일치”

착유회차에 따른 평균 착유량



분석_가설검증

- ▶ 가설2 (온도가 높을수록 평균 착유량이 감소할 것이다)

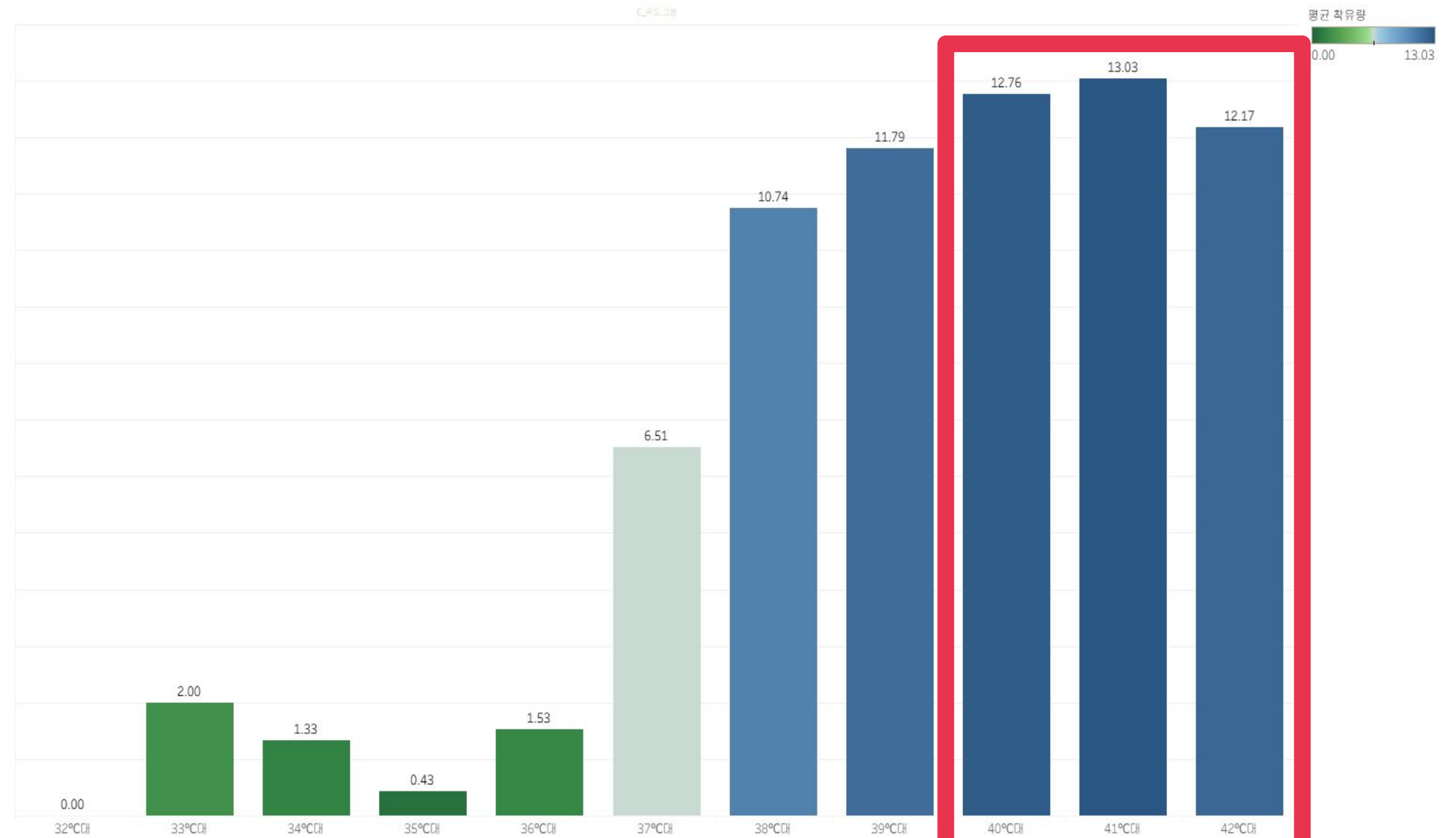
“온도가 높을수록
평균 착유량이 감소할 것이다”



“40~41도 온도에서
높은 착유량을 보임”

온도에 따른 평균 착유량

온도에 따른 착유량



분석_가설검증

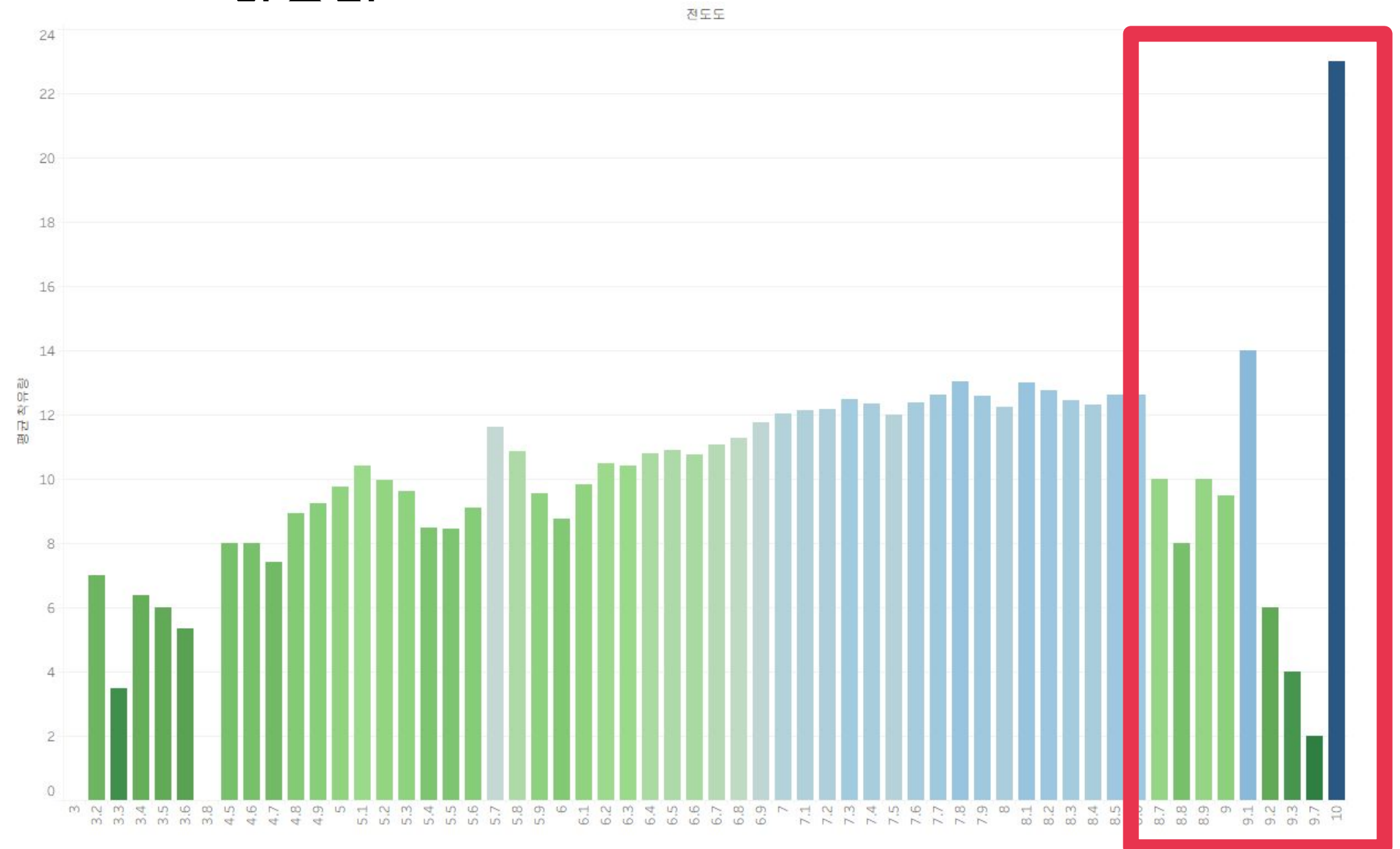
- ▶ 가설 3 (전도도가 높을수록 평균 착유량이 감소할 것이다)

“전도도가 높을수록
평균 착유량이 감소할 것이다”



“8.6이상부터 내려가는
추세이지만
값이 튀는 경우도 존재함”

전도도에 따른 평균

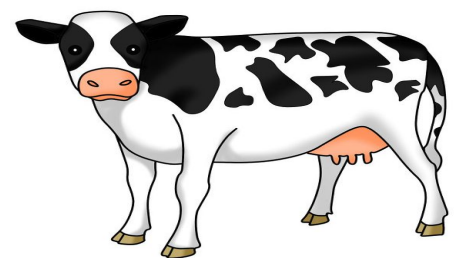
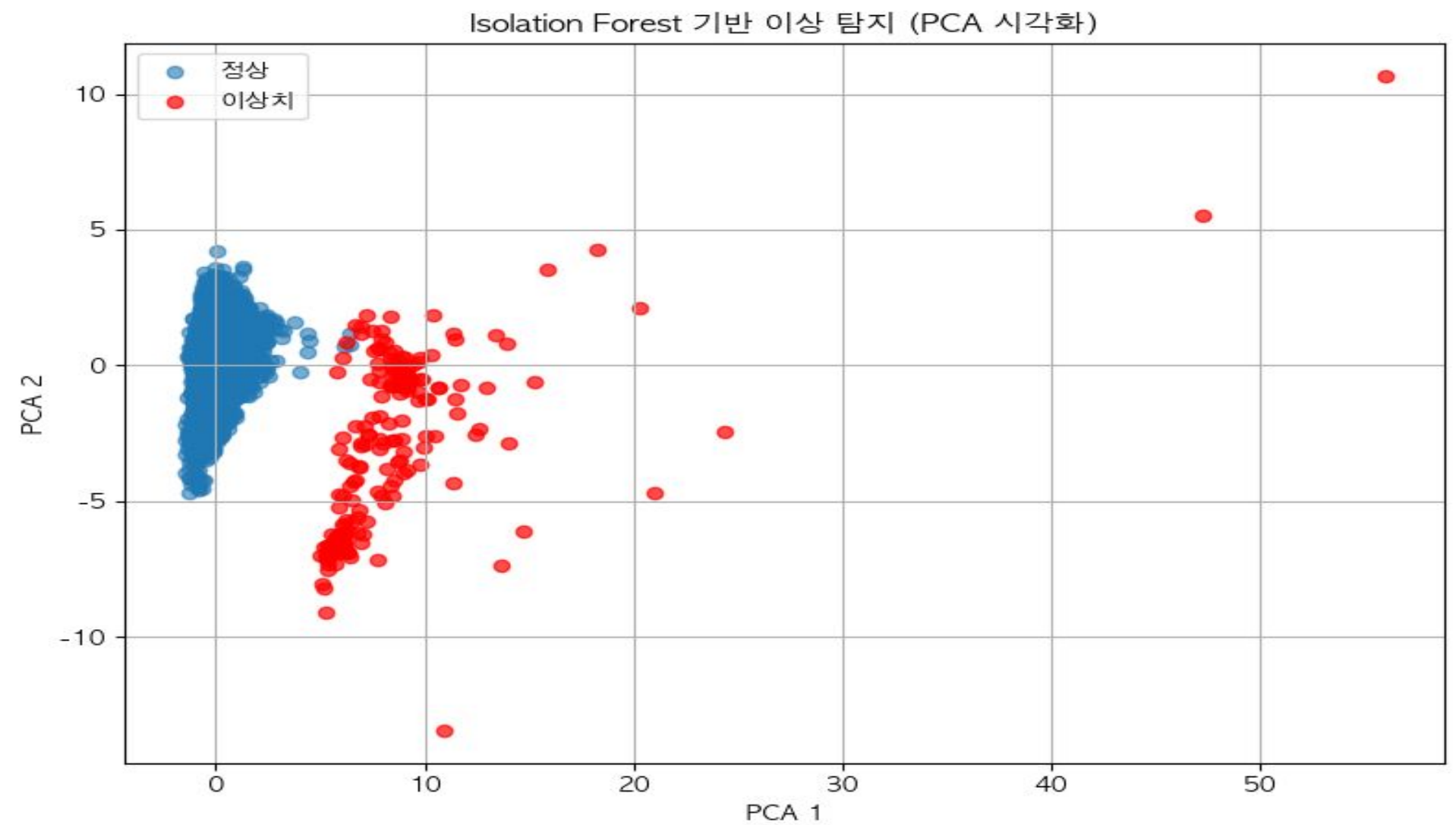


▶ 분석 (이상탐지_건강위험)

- 컬럼 : '온도', '전도도', 'PFR'
- 방법 : Isolation Forest



이상치로 탐지된 샘플 개수 : 189



03 분석_이상탐지

▶ 분석 (이상탐지_건강위험 결과)

	예측 : 이상	예측 : 정상
실제 : 이상	36	10
실제 : 정상	530	18298

	정밀도	재현율	F-1 score	실제 개수
이상	0.06	0.78	0.12	46
정상	1.00	0.97	0.99	18828
정확도			0.97	18874
macro avg (클래스별 평균)	0.53	0.88	0.55	18874
weighted avg (클래스 비중 반영한 평균)	1.00	0.97	0.98	18874

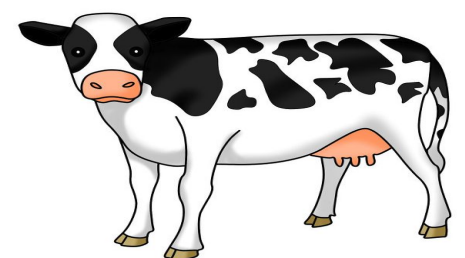
	온도	전도도	PFR
모델	36.9~42.6	3.2~10	0.38~3.54
실제	41도 이상	7.3 이상	0.7~1.0

▶ 지표 해석

모델이 **정상** 데이터는 거의 완벽하게 맞춘
하지만 **이상** 데이터는 찾을 때 오탐지가 많음(정밀도 6%)

반면, 이상을 놓치는 비율은 낮음(재현율 78%)

→ "이상이라면 일단 잡고 보는" 성향 존재

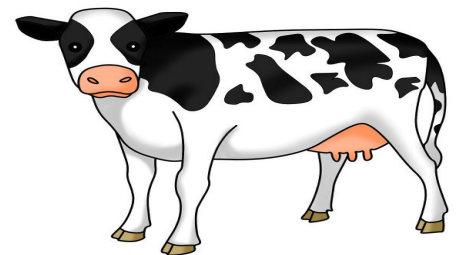
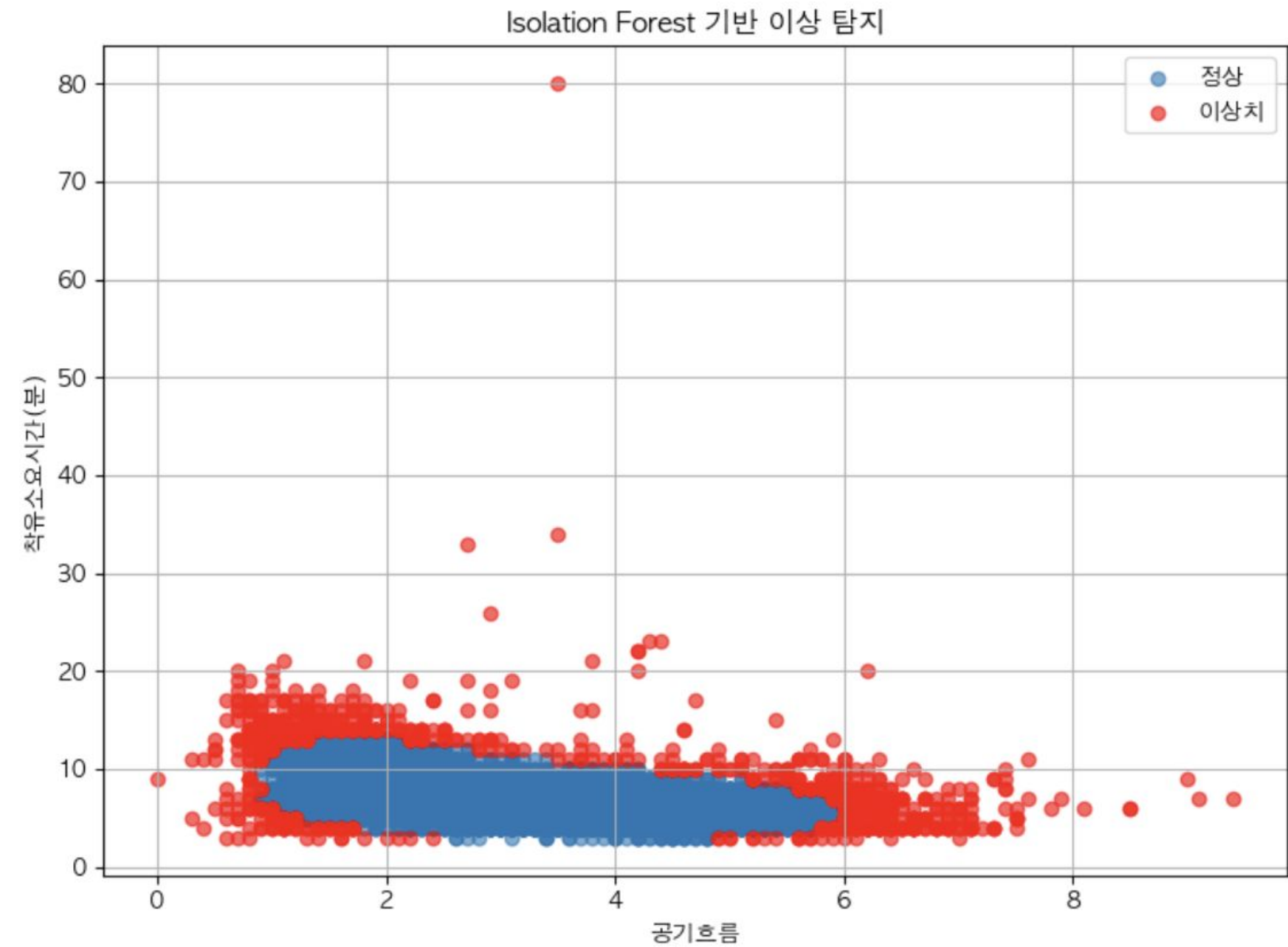


▶ 분석 (이상탐지_기계오류)

- 컬럼 : '공기흐름', '착유소요시간 (분)'
- 방법 : Isolation Forest



이상치로 탐지된 샘플 개수 : 944



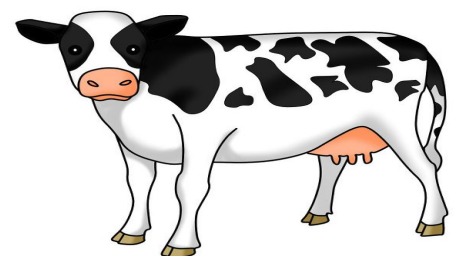
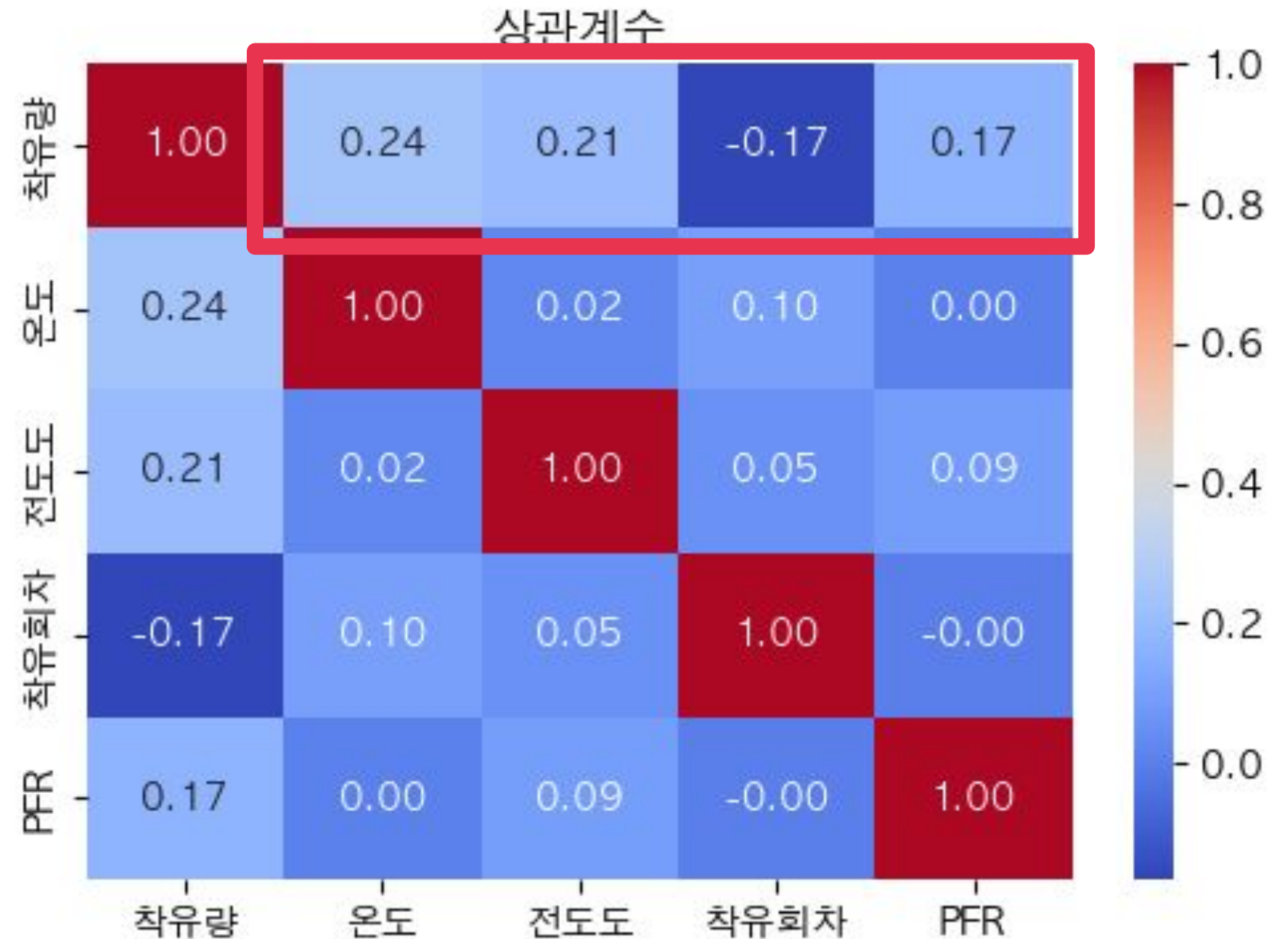
분석_상관관계

▶ 분석 (상관관계)

- 착유량 ↔ 온도: 0.24
- 착유량 ↔ 전도도: 0.21
- 착유량 ↔ 착유회차: -0.17
- 착유량 ↔ PFR: 0.17



“변수별로 상관관계가 약하다”
비선형 모델 선택

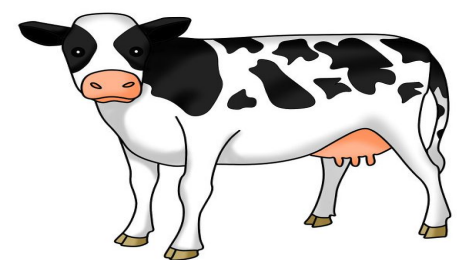


▶ 분석 (4개 모델 비교 결과)

	$R^2(\text{test})$	$R^2(\text{train})$
LightGBM	0.686	0.758
CatBoost	0.6882	0.764
RandomForest	0.677	0.956
XGBoost	0.676	0.825

▶ 모델 선정

$R^2(\text{test})$ 가 제일 높으면서
 $R^2(\text{train})$ 의 비율이 좋은 **CatBoost** 모델 선정



▶ 분석(착유량 예측 CatBoost 성능 결과)

● CatBoost 모델 결과

- Train R^2 : 0.764
- **Test R^2 : 0.688**
- Train MAE: 1.318
- **Test MAE: 1.508**



- Optuna를 활용한
CatBoost
하이퍼파라미터
최적화



● 최적화 모델 성능

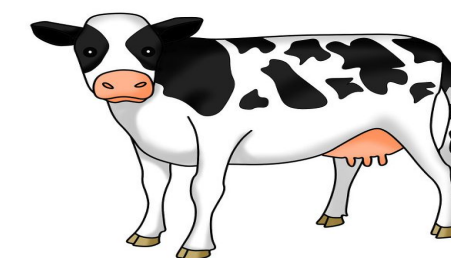
- Train R^2 : 0.833
- **Test R^2 : 0.703**
- Train MAE: 1.111
- **Test MAE : 1.466**

▶ 해석

Test R^2 : **0.688** → **0.703** 테스트 데이터 설명력이 **+0.015** 상승


Test MAE: **1.508** → **1.466** 평균 오차가 약 **0.042** 감소

→ Optuna로 찾은 하이퍼파라미터 세트가 기본값보다 전반적으로 성능 향상



04 해석 및 적용

▶ 분석(착유량 예측)

 **착유량 예측기 (CatBoost)**

학습된 모델에 x값을 입력해 착유량(L)을 예측하고, 유지방/유단백에 따른 우유 등급을 안내합니다.

온도(°C)	34.50	-	+
전도도	7.00	-	+
착유회차	3	-	+
유지방(%)	12.30	-	+
유단백(%)	3.80	-	+
공기흐름	1.10	-	+
착유소요시간(분)	7.50	-	+

☒ PFR 자동 계산 (유단백/유지방)

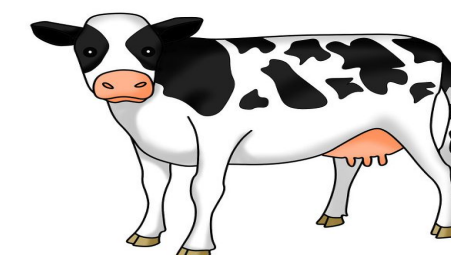
예측하기

예측된 착유량 (L)

3.94

우유 등급: **프리미엄 우유**

> 입력값 확인



04 해석 및 적용

▶ 적용

종류 / 범위 및 적용	범위	설명
기계 오류 범위(이상탐지)	공기흐름 : 0.9~5.9 벗어난 값 착유소요시간(분) : 3~13분 벗어난 값	<ul style="list-style-type: none">● 실제 오류 범위를 알 수 없어 이상탐지 모델 범위를 그대로 적용● 두 범위를 모두 벗어나면 기계 오류 의심
건강 위험 범위(실제지표)	온도 : 41도 이상 전도도 : 7.3 이상 PFR : 0.7~1.0 벗어난 값	<ul style="list-style-type: none">● 모두 정상 범위면 '정상'● 하나라도 벗어나다면 '관심 필요'● 모두 벗어나다면 '조치 필요'로 분류

▶ 제안

범위를 통해 분류를 진행하고 태블로를 이용하여

농장주인이 날마다 현황을 쉽게 파악할 수 있도록 대시보드 제작

