description_data.md 2025-04-04

Data Description

1. Raw Data

국가 지역.csv 컬럼 이름 설명

- ALLSKY_SFC_UV_INDEX: 전체 하늘 조건에서의 자외선 지수 (W/m² × 40)
- ALLSKY_SFC_SW_DWN: 전체 하늘 조건에서 지표면에 도달한 단파 복사량 (kWh/m²/day)
- WS2M: 지상 2미터 높이의 평균 풍속 (m/s)
- CLRSKY SFC SW DWN: 맑은 하늘 조건에서의 지표면 단파 복사량 (kWh/m²/day)
- T2M: 지상 2미터에서의 일 평균 기온 (°C)
- T2M MAX: 지상 2미터에서의 일 최고 기온 (°C)
- T2M_MIN: 지상 2미터에서의 일 최저 기온 (°C)
- RH2M: 지상 2미터에서의 상대습도 (%)
- PRECTOTCORR: 보정된 일 누적 강수량 (mm/day)
- PS: 지표면 기압 (kPa)
- CLRSKY_SFC_PAR_TOT: 맑은 하늘 조건에서의 광합성 유효 복사(PAR) 총량 (kWh/m²/day)
- WS10M MAX: 지상 10미터에서의 최대 풍속 (m/s)
- WS10M MIN: 지상 10미터에서의 최소 풍속 (m/s)
- ALLSKY_KT: 전체 하늘 일사량 맑음 지수 (dimensionless)
- T2M_RANGE: 일일 기온 범위 = 최고 기온 최저 기온 (°C)
- WS10M: 지상 10미터에서의 평균 풍속 (m/s)
- TS: 지표면 온도 (지표 피부 온도, °C)
- WSC: 고도 보정이 적용된 바람 속도 (m/s)

coffee_c_price 컬럼 이름 설명

- Date: 날짜 (YYYY-MM-DD) 형식
- Coffee_Price: 커피 선물 가격 (센트/파운드, Coffee C Futures 심볼: KC=F)

market data 컬럼 이름 설명

- Date: 날짜 (YYYY-MM-DD) 형식
- Coffee Price: 커피 선물 가격 (센트/파운드, Coffee C Futures 심볼: KC=F)
- Crude Oil Price: 서부 텍사스산 원유(WTI) 가격 (달러/배럴 심볼: CL=F)
- USD KRW: 미국 달러(USD) 대비 한국 원(KRW) 환율 (원/달러 심볼: KRW=X)
- USD BRL: 미국 달러(USD) 대비 브라질 헤알(BRL) 환율 (헤알/달러 심볼: BRL=X)
- USD COP: 미국 달러(USD) 대비 콜롬비아 페소(COP) 환율 (페소/달러 심볼: COP=X)
- USD VND: 미국 달러(USD) 대비 베트남 동(VND) 환율 (동/달러 심볼: VND=X)

2. Processed1 Data

국가_지역.csv 컬럼 이름 설명

- raw와 같음
- 다음 컬럼 추가됨:

description data.md 2025-04-04

- o Coffee Price: 커피 선물 가격 (센트/파운드, Coffee C Futures 심볼: KC=F)
- Coffee_Return: 커피 선물 가격 변화율

coffee_c_price 컬럼 이름 설명

- raw와 같음
- 다음 컬럼 추가됨:
 - Coffee_Return: 커피 선물 가격 변화율

market data 컬럼 이름 설명

• raw와 같음

3. Processed2 Data

국가 지역.csv 컬럼 이름 설명

- raw와 같음
- 다음 컬럼 추가됨:
 - season_tag: 해당 날짜가 수확기(harvest), 수확 전(pre-harvest), 부수적 수확기 (sub_harvest), 비수확기(off-season) 중 어느 시기에 해당하는지 나타냄
 - o days_until_harvest: 해당 날짜로부터 주요 수확기(월 기준)의 시작일까지 남은 날짜 (단위: 일수)
 - T2M lag 1m ~ T2M lag 6m: 평균 기온의 1~6개월 전 시점 값을 lag feature로 생성
 - WS2M_lag_1m ~ WS2M_lag_6m: 2m 풍속의 1~6개월 전 시점 값을 lag feature로 생성
 - ALLSKY_SFC_SW_DWN_lag_1m ~ _6m: 단파 복사량의 1~6개월 전 시점 값을 lag feature로 생성
 - PRECTOTCORR_lag_1m ~ _6m: 강수량의 1~6개월 전 시점 값을 lag feature로 생성

4. Final Train Weather Data

🗐 컬럼 설명

train_weather의 Date 는 2015/01/01 - 2024/12/31까지 test_weather의 Date 는 2025/01/01 - 2025/04/01까지

컬럼명 	설명	
Date	관측 날짜 (YYYY-MM-DD)	
T2M	평균 기온 (2m 지면 기준, 단위: ℃)	
WS2M	평균 풍속 (2m 지면 기준, 단위: m/s)	
ALLSKY_SFC_SW_DWN	일사량 (단위: W/m²), 지표면에 도달한 단일 일평균 태양 복사량	
ALLSKY_SFC_UV_INDEX	자외선 지수 (UV Index)	
PRECTOTCORR	강수량 (단위: mm/day)	
RH2M	평균 상대 습도 (%), 2m 기준	
	0.10	

description_data.md 2025-04-04

컬럼명 	설명
PS	평균 지면 기압 (단위: hPa)
Coffee_Price	해당 날짜의 커피 가격 (cents/lb)
Coffee_Return	커피 가격 수익률 (전일 대비 변화율, pct_change)
season_tag	해당 날짜의 계절 구분 (예: harvest, pre-harvest, off-season)
days_until_harvest	다음 수확일까지 남은 일수

★ Lag Feature 설명

컬럼명	설명
T2M_lag_1m ~ T2M_lag_6m	1개월 ~ 6개월 전의 평균 기온
WS2M_lag_1m ~ WS2M_lag_6m	1개월 ~ 6개월 전의 평균 풍속
ALLSKY_SFC_SW_DWN_lag_1m ~ lag_6m	1개월 ~ 6개월 전의 일사량
PRECTOTCORR_lag_1m ~ lag_6m	1개월 ~ 6개월 전의 강수량

🔊 지역 정보

컬럼명	설명
locationName	데이터 수집 지역 이름 (예: brazil_varginha 등)

5. Final Market Data & Coffee Label

coffee_c_price 컬럼 이름 설명

- Date: 날짜 (YYYY-MM-DD) 형식
- Coffee_Price: 커피 선물 가격 (센트/파운드, Coffee C Futures 심볼: KC=F)

market_data 컬럼 이름 설명

- Date: 날짜 (YYYY-MM-DD) 형식
- Coffee Price: 커피 선물 가격 (센트/파운드, Coffee C Futures 심볼: KC=F)
- Crude Oil Price: 서부 텍사스산 원유(WTI) 가격 (달러/배럴 심볼: CL=F)
- USD KRW: 미국 달러(USD) 대비 한국 원(KRW) 환율 (원/달러 심볼: KRW=X)
- USD BRL: 미국 달러(USD) 대비 브라질 헤알(BRL) 환율 (헤알/달러 심볼: BRL=X)
- USD_COP: 미국 달러(USD) 대비 콜롬비아 페소(COP) 환율 (페소/달러 심볼: COP=X)
- USD ETB: 미국 달러(USD) 대비 에티오피아 비르(ETB) 환율 (비르/달러 심볼: ETB=X)
- USD_KRW_Return: USD/KRW 환율의 일간 변화율 (pct_change()로 계산)
- USD BRL Return: USD/BRL 환율의 일간 변화율
- USD_COP_Return: USD/COP 환율의 일간 변화율
- USD ETB Return: USD/ETB 환율의 일간 변화율