

휴먼-IT아카데미

암호화폐 가격과 외부 요인 간 상관관계 분석 및 시각화

TEAM 3조

임용균

목차

01. 프로젝트 개요

02. 프로젝트 팀 구성 및 역할

03. 프로젝트 수행 절차 및 방법

04. 프로젝트 수행 결과

05. 자체 평가 의견

01 프로젝트 개요

▶ 프로젝트 주제 및 선정 배경

최근 암호화폐는 전통적인 금융 시장과의 상관관계가 점점 뚜렷해지며 단순한 투기 자산이 아닌 하나의 금융 자산으로 주목받고 있다. 하지만 그 가격 변동의 원인을 명확히 설명할 수 있는 체계적인 분석은 부족한 실정이다.

이에 따라 본 프로젝트에서는 암호화폐가 정책금리나 주가지수처럼 경제 전반의 흐름을 나타내는 지표들과 어떠한 관계를 가지는지, 날씨, 달의 위상 같은 외부 요인들이 가격에 영향을 줄 수 있는지에 대한 궁금증에서 출발하였다.

📌 연구 가설 설정

📌 가설 1: 경제지표와 암호화폐 가격 간의 관계

- ☒ 귀무가설(H_{01}): 암호화폐 가격은 금리, 코스피, 나스닥 지수와 상관관계가 없다.
- ☐ 대립가설(H_{11}): 암호화폐 가격은 금리, 코스피, 나스닥 지수와 상관관계가 있다.

이 가설은 암호화폐 시장이 전통적인 거시경제 지표에 반응하는가를 검증하는 데 목적이 있음.

📌 가설 2: 비정형 요인과 암호화폐 가격 간의 관계

- ☒ 귀무가설(H_{02}): 암호화폐 가격은 날씨나 달의 위상과 상관관계가 없다.
- ☐ 대립가설(H_{12}): 암호화폐 가격은 날씨나 달의 위상과 상관관계가 있다.

이 가설은 날씨·천문학적 요인처럼 창의적이고 비정형적인 변수들이 가격에 영향을 미칠 수 있는지를 확인하는 데 목적이 있음.

01 프로젝트 개요

프로젝트 개요

본 프로젝트는 암호화폐(비트코인, 이더리움, 도지코인)의 가격을 중심으로 외부 요인들과의 상관관계를 분석하고 이를 시각화하는 웹 기반 분석 시스템을 구현한다.

본 프로젝트는 암호화폐(비트코인, 이더리움, 도지코인)의 가격을 중심으로 총 6개의 요인(한국 금리, 미국 금리, 코스피, 나스닥, 날씨, 달의 위상)과의 상관관계 분석을 고려하여 데이터를 수집 및 전처리하였고 Python 기반 통계 분석 및 시각화 도구를 통해 주요 요인과의 관계를 도출하였다.

분석 결과를 시각적으로 확인할 수 있는 웹 애플리케이션을 Flask 프레임워크로 제작하여 누구나 직접 확인하고 분석 결과를 볼 수 있도록 구현하였다.

01 프로젝트 개요

▶ 활용 장비 및 재료

- 사용한 데이터셋

| 파일명 | 기간 | 출처 |
|----------------------------|-------------------|----------------------|
| 비트코인_23.01.01~25.01.31.csv | 23.01.01~25.01.31 | 업비트API |
| 이더리움_23.01.01~25.01.31.csv | 23.01.01~25.01.31 | 업비트API |
| 도지코인_23.01.01~25.01.31.csv | 23.01.01~25.01.31 | 업비트API |
| 코스피_23.01.01~25.01.31.csv | 23.01.01~25.01.31 | KRX 정보데이터시스템 |
| 나스닥_23.01.01~25.01.31.csv | 23.01.01~25.01.31 | Investing.com |
| 한국금리_23.01.01~25.01.31.csv | 23.01.01~25.01.31 | 한국은행 경제통계시스템 |
| 미국금리_23.01.01~25.01.31.csv | 23.01.01~25.01.31 | FEDERAL RESERVE BANK |
| 월령_23.01.01~25.01.31.csv | 23.01.01~25.01.31 | 한국천문연구원 |
| 날씨_23.01.01~25.01.31.csv | 23.01.01~25.01.31 | Open-Meteo API |

01 프로젝트 개요

▶ 활용 장비 및 재료

- 프레임 워크 및 라이브러리

| 분류 | 라이브러리/프레임워크 | 주요 용도 및 기능 |
|-------|-------------------------------|--------------------------------|
| 프론트엔드 | HTML5, Tailwind | UI 구조 및 반응형 디자인 |
| | JavaScript, Cjart.js | 버튼 기능 및 데이터 시각화 |
| 백엔드 | Flask | 웹 서버 구축, 라우팅, 세션관리 |
| | Flask-Bcrypt | 비밀번호 암호화 |
| | MySQL, mysql-connector-python | 정보 저장/조회/연결 |
| 데이터분석 | Pandas | 데이터프레임 기반 정제 및 가공 |
| | scipy.stats | 정규성검정(Shapiro), 상관분석(Spearman) |
| | statsmodels | 회귀분석(ols), 다중공선성(VIF) |
| 기타 | os | 경로 처리 |

01 프로젝트 개요

▶ 프로젝트 구조



01 프로젝트 개요

▶ 기대효과

- **외부 요인과 암호화폐 가격 간의 관계 통찰 제공:** 금리, 주가지수, 날씨, 월령 등 다양한 외부 변수와 암호화폐 가격 간의 상관관계 및 영향도를 시각화하여 시장에 대한 정성적·정량적 통찰을 제공
- **창의적 변수 도입에 대한 실험적 접근 시도:** 기존 분석에 잘 사용되지 않는 달의 위상, 날씨와 같은 비정형적 외부 변수들을 포함시킴으로써 암호화폐 가격 분석의 범위를 확장하고자 하는 실험적 접근을 시도
- **웹 인터페이스를 통한 접근성 확보 및 공유 용이성:** 분석 결과를 웹 상에서 확인하고 실시간으로 실행 결과를 볼 수 있는 구조로 구현하여 공유할 수 있는 시스템을 완성함.

02 프로젝트 팀 구성 및 역할

▶ 프로젝트 팀 구성 및 역할

- 담당 업무

| 훈련생 | 역할 | 담당 업무 |
|-----|----|---|
| 임용균 | 팀장 | <ul style="list-style-type: none">▶ 기획 및 전체 구조 설계▶ 데이터 전처리 및 분석▶ 데이터 분석 페이지 백엔드 구현▶ 분석 결과 시각화 기능 구현(Chart.js)▶ 전체 배포 및 시스템 통합 테스트 |
| 박상우 | 팀원 | <ul style="list-style-type: none">▶ 데이터 수집▶ 메인 페이지 구축▶ 로그인/회원가입 프론트엔드 구축▶ 로그인/회원가입 백엔드 구현▶ MySQL 데이터 관리 |

03 프로젝트 수행 절차 및 방법



| 구분 | 소요 기간(소요 시간:H) | 활동 | 비고 |
|--------------|----------------|---|---|
| 사전 기획 | ▶ 3H | ▶ 프로젝트 기획 및 주제 선정 ▶ 기획안 작성 | ▶ 아이디어 선정 |
| 연구 가설 설정 | ▶ 2H | ▶ 암호화폐와 외부 요인간의 상관관계를 추정할 가설 수립 | ▶ 기존 연구 사례, 선행 논문 참고 |
| 데이터 수집 및 전처리 | ▶ 10H | ▶ 데이터 정제 및 정규화 | |
| 데이터 분석 | ▶ 10H | ▶ 상관계수 분석 및 시각화 ▶ 변수간 관계 해석 ▶ 통계적 유의성 검정 | ▶ 팀 별 중간보고 실시 |
| 웹 구현 | ▶ 15H | ▶ 프론트 엔드 구축(메인페이지, 뷰페이지) ▶ 백엔드 구축(데이터베이스 생성 및 연결) ▶ 기능 구현(로그인, 로그아웃, 회원가입, 시각화) | |
| 퍼블리싱 및 테스트 | ▶ 10H | ▶ 라우터 모듈 구조 재정비, 중복 로직 제거 및 함수 모듈화 ▶ 우분투 기반 자체 서버 앱 배포 및 도메인 연결 ▶ 주요 기능 흐름 테스트, 분석 결과 시각화 오류 검증, 로그인/회원가입 예외 처리 테스트 | ▶ 코드 리팩터링 ▶ 웹 퍼블리싱 ▶ 통합 테스트 및 디버깅 |
| 총 개발기간 | ▶ 50H | - | - |

04 프로젝트 수행 결과

결과 제시 ① 데이터구조

▶ 데이터셋 구조 요약

| 파일명 | 컬럼수 | 결측치 컬럼 | 데이터 수 |
|------|-----|--------------|-------|
| 비트코인 | 13 | - | 762 |
| 이더리움 | 13 | change_price | 762 |
| 도지코인 | 13 | change_price | 762 |
| 코스피 | 10 | - | 507 |
| 나스닥 | 7 | - | 522 |
| 한국금리 | 2 | - | 509 |
| 미국금리 | 2 | - | 762 |
| 월령 | 3 | - | 762 |
| 날씨 | 2 | - | 762 |

04 프로젝트 수행 결과

결과 제시 ② 탐색적 분석 및 전처리

▶ 데이터 전처리(추출 및 병합)

- 전체 분석 기간은 2023년 1월 1일 ~ 2025년 1월 31일, 총 762일
 - 코스피, 나스닥, 한국 금리 < 762
- 이들 데이터셋에는 결측치가 없음에도 불구하고 데이터 수에 차이가 존재
- 코스피, 나스닥, 금리 데이터는 주말 및 공휴일에 시장이 열리지 않아 해당 일자의 데이터가 존재하지 않음
(미국 금리 데이터는 제공 사이트 측에서 전일 데이터로 보간된것으로 추정됨)
1. 각 데이터의 일부 컬럼들의 데이터타입 변환(object -> float)
 2. 각 데이터 날짜 타입 변환
 3. 각 데이터에서 필요한 컬럼 추출 및 컬럼명 변경
 4. 암호화폐데이터 날짜 기준 병합

| | |
|---|---|
| btc/eth/doge[candle_date_time_kst, trade_price] -> [date, btc_price/eth_price/doge_price] | |
| kospi[일자, 종가] -> [date, kospi] | nasdaq[날짜, 종가] -> [date, nasdaq] |
| kor_rate[일자, 금리] -> [date, kor_rate] | uas_rate[일자, 금리] -> [date, usa_rate] |
| moon[date, moon_phase] | weather[날짜, 날씨 상태] -> [date, condition] |

04 프로젝트 수행 결과

결과 제시 ② 탐색적 분석 및 전처리



데이터 전처리(결측치 처리)

| | date | btc_price | eth_price | doge_price | kospi | nasdaq | kor_rate | usa_rate | moon_phase | condition |
|---|------------|------------|-----------|------------|---------|----------|----------|----------|------------|-----------|
| 0 | 2023-01-01 | 21123000.0 | 1528000.0 | 89.3 | NaN | NaN | NaN | 4.33 | 9.1 | 구름 조금 |
| 1 | 2023-01-02 | 21215000.0 | 1543500.0 | 90.8 | 2225.67 | NaN | 3.25 | 4.33 | 10.1 | 맑음 |
| 2 | 2023-01-03 | 21351000.0 | 1554500.0 | 90.2 | 2218.68 | 10386.99 | 3.25 | 4.33 | 11.1 | 맑음 |
| 3 | 2023-01-04 | 21588000.0 | 1610000.0 | 93.7 | 2255.98 | 10458.76 | 3.25 | 4.33 | 12.1 | 맑음 |
| 4 | 2023-01-05 | 21575000.0 | 1602500.0 | 91.8 | 2264.65 | 10305.24 | 3.25 | 4.33 | 13.1 | 구름 조금 |
| 5 | 2023-01-06 | 21517000.0 | 1610000.0 | 91.9 | 2289.97 | 10569.29 | 3.25 | 4.33 | 14.1 | 눈 |
| 6 | 2023-01-07 | 21539000.0 | 1609000.0 | 91.9 | NaN | NaN | NaN | 4.33 | 14.6 | 눈 |
| 7 | 2023-01-08 | 21704000.0 | 1634500.0 | 93.3 | NaN | NaN | NaN | 4.33 | 13.6 | 흐림 |
| 8 | 2023-01-09 | 21655000.0 | 1663500.0 | 95.4 | 2350.19 | 10635.65 | 3.25 | 4.33 | 12.6 | 구름 조금 |
| 9 | 2023-01-10 | 21916000.0 | 1680500.0 | 97.9 | 2351.31 | 10742.63 | 3.25 | 4.33 | 11.6 | 구름 조금 |
| | date | btc_price | eth_price | doge_price | kospi | nasdaq | kor_rate | usa_rate | moon_phase | condition |
| 0 | 2023-01-01 | 21123000.0 | 1528000.0 | 89.3 | NaN | NaN | NaN | 4.33 | 9.1 | 구름 조금 |
| 1 | 2023-01-02 | 21215000.0 | 1543500.0 | 90.8 | 2225.67 | NaN | 3.25 | 4.33 | 10.1 | 맑음 |

- 병합으로 인한 kospi, nasdaq, kor_rate 결측치 처리하기 위해 전일 값으로 보간
- 분석 데이터의 기준 시작일은 23년 1월 1일
이전 일자인 22년 12월 데이터가 없기 때문에 전일 보간이 불가능함.
- 해당 결측값은 22년 12월 말 데이터를 직접 확인하여 수작업으로 입력함.
(KOSPI = 2236.40 | NASDAQ = 10466.48 | 한국 금리 = 3.25)

04 프로젝트 수행 결과

결과 제시 ② 탐색적 분석 및 전처리

▶ 데이터 전처리(날씨 변수 인코딩)

| | date | btc_price | eth_price | doge_price | kospi | nasdaq | kor_rate | usa_rate | moon_phase | partly_cloudy | snow | clear |
|---|------------|------------|-----------|------------|---------|----------|----------|----------|------------|---------------|------|-------|
| 0 | 2023-01-01 | 21123000.0 | 1528000.0 | 89.3 | 2236.40 | 10466.48 | 3.25 | 4.33 | 9.1 | 1 | 0 | 0 |
| 1 | 2023-01-02 | 21215000.0 | 1543500.0 | 90.8 | 2225.67 | 10466.48 | 3.25 | 4.33 | 10.1 | 0 | 0 | 1 |
| 2 | 2023-01-03 | 21351000.0 | 1554500.0 | 90.2 | 2218.68 | 10386.99 | 3.25 | 4.33 | 11.1 | 0 | 0 | 1 |

- condition 컬럼은 문자열 형태의 날씨 상태(예: 구름 조금, 눈, 맑음...)를 가지고 있었기 때문에 이를 분석에 사용할 수 있도록 one-hot encoding 실시.
- 인코딩 결과로 생성된 컬럼들은 기본적으로 boolean타입이므로 모델 학습 및 수치 분석을 위해 0/1 형태의 정수형으로 최종 변환

04 프로젝트 수행 결과

결과 제시 ② 탐색적 분석 및 전처리

▶ 데이터 전처리(정규화)

| | date | btc_price | eth_price | doge_price | kospi | nasdaq | kor_rate | usa_rate | moon_phase |
|-----|------------|-----------|-----------|------------|-----------|----------|-----------|-----------|------------|
| 759 | 2025-01-29 | 2.546962 | 1.208448 | 2.342932 | -0.214599 | 1.678511 | -2.926818 | -2.029105 | -1.725639 |
| 760 | 2025-01-30 | 2.526570 | 1.315362 | 2.367252 | -0.214599 | 1.697637 | -2.926818 | -2.029105 | -1.515155 |
| 761 | 2025-01-31 | 2.482415 | 1.441556 | 2.378306 | -0.377945 | 1.676623 | -2.926818 | -2.029105 | -1.281285 |

- 본 분석에서는 암호화폐 가격, 주가지수, 금리 등 서로 단위와 스케일이 상이한 변수들을 함께 분석함
ex) btc_price : 수천만 원 단위 | kor_rate : 0~5% 수준의 소수점 단위
- 이러한 스케일 차이를 그대로 사용할 경우 특정 변수 분석 결과에 과도한 영향을 주거나 왜곡이 발생 할 수 있음
- StandardScaler를 사용하여 모든 수치형 변수를 정규 분포 형태로 스케일링함
(date, 날씨 컬럼 제외)

04 프로젝트 수행 결과

결과 제시 ③ 상관관계분석

▶ 정규성 검정(Shapiro-Wilk)

```
btc_price → p-value: 0.0000 → ❌ 비정규
eth_price → p-value: 0.0000 → ❌ 비정규
doge_price → p-value: 0.0000 → ❌ 비정규
kospi → p-value: 0.0115 → ❌ 비정규
nasdaq → p-value: 0.0000 → ❌ 비정규
kor_rate → p-value: 0.0000 → ❌ 비정규
usa_rate → p-value: 0.0000 → ❌ 비정규
moon_phase → p-value: 0.0000 → ❌ 비정규
partly_cloudy → p-value: 0.0000 → ❌ 비정규
snow → p-value: 0.0000 → ❌ 비정규
clear → p-value: 0.0000 → ❌ 비정규
rain → p-value: 0.0000 → ❌ 비정규
cloudy → p-value: 0.0000 → ❌ 비정규
```

- 적절한 상관분석 방법을 선택하기 위해 정규성 검정 시행
- Shapiro-Wilk 검정은 35,000개 이하의 데이터에 적합

04 프로젝트 수행 결과

결과 제시 ③ 상관관계분석

▶ 상관관계 분석(Spearman)

▶ btc_price 와의 상관관계수 :

| | | | |
|---------------|----------------------------------|---|-----|
| kospi | → $p = 0.4156$ ($p = 0.0000$) | ✅ | 유의 |
| nasdaq | → $p = 0.9157$ ($p = 0.0000$) | ✅ | 유의 |
| kor_rate | → $p = -0.5022$ ($p = 0.0000$) | ✅ | 유의 |
| usa_rate | → $p = 0.0843$ ($p = 0.0199$) | ✅ | 유의 |
| moon_phase | → $p = -0.0183$ ($p = 0.6132$) | ⚠ | 유의x |
| partly_cloudy | → $p = -0.0027$ ($p = 0.9403$) | ⚠ | 유의x |
| snow | → $p = 0.0384$ ($p = 0.2895$) | ⚠ | 유의x |
| clear | → $p = 0.0753$ ($p = 0.0378$) | ✅ | 유의 |
| rain | → $p = -0.0560$ ($p = 0.1228$) | ⚠ | 유의x |
| cloudy | → $p = -0.0281$ ($p = 0.4390$) | ⚠ | 유의x |

▶ eth_price 와의 상관관계수 :

| | | | |
|---------------|----------------------------------|---|-----|
| kospi | → $p = 0.5227$ ($p = 0.0000$) | ✅ | 유의 |
| nasdaq | → $p = 0.8604$ ($p = 0.0000$) | ✅ | 유의 |
| kor_rate | → $p = -0.3674$ ($p = 0.0000$) | ✅ | 유의 |
| usa_rate | → $p = 0.1631$ ($p = 0.0000$) | ✅ | 유의 |
| moon_phase | → $p = -0.0098$ ($p = 0.7867$) | ⚠ | 유의x |
| partly_cloudy | → $p = -0.0186$ ($p = 0.6086$) | ⚠ | 유의x |
| snow | → $p = 0.0531$ ($p = 0.1431$) | ⚠ | 유의x |
| clear | → $p = 0.0604$ ($p = 0.0958$) | ⚠ | 유의x |
| rain | → $p = -0.0544$ ($p = 0.1336$) | ⚠ | 유의x |
| cloudy | → $p = -0.0105$ ($p = 0.7732$) | ⚠ | 유의x |

▶ doge_price 와의 상관관계수 :

| | | | |
|---------------|----------------------------------|---|-----|
| kospi | → $p = 0.3414$ ($p = 0.0000$) | ✅ | 유의 |
| nasdaq | → $p = 0.8056$ ($p = 0.0000$) | ✅ | 유의 |
| kor_rate | → $p = -0.5160$ ($p = 0.0000$) | ✅ | 유의 |
| usa_rate | → $p = -0.1190$ ($p = 0.0010$) | ✅ | 유의 |
| moon_phase | → $p = -0.0064$ ($p = 0.8590$) | ⚠ | 유의x |
| partly_cloudy | → $p = 0.0285$ ($p = 0.4318$) | ⚠ | 유의x |
| snow | → $p = 0.0887$ ($p = 0.0143$) | ✅ | 유의 |
| clear | → $p = 0.1092$ ($p = 0.0025$) | ✅ | 유의 |
| rain | → $p = -0.1542$ ($p = 0.0000$) | ✅ | 유의 |
| cloudy | → $p = 0.0002$ ($p = 0.9960$) | ⚠ | 유의x |

- 데이터 분포가 정규성을 만족하지 않기 때문에 비모수 분석 사용(Spearman)

04 프로젝트 수행 결과

결과 제시 ③ 상관관계분석

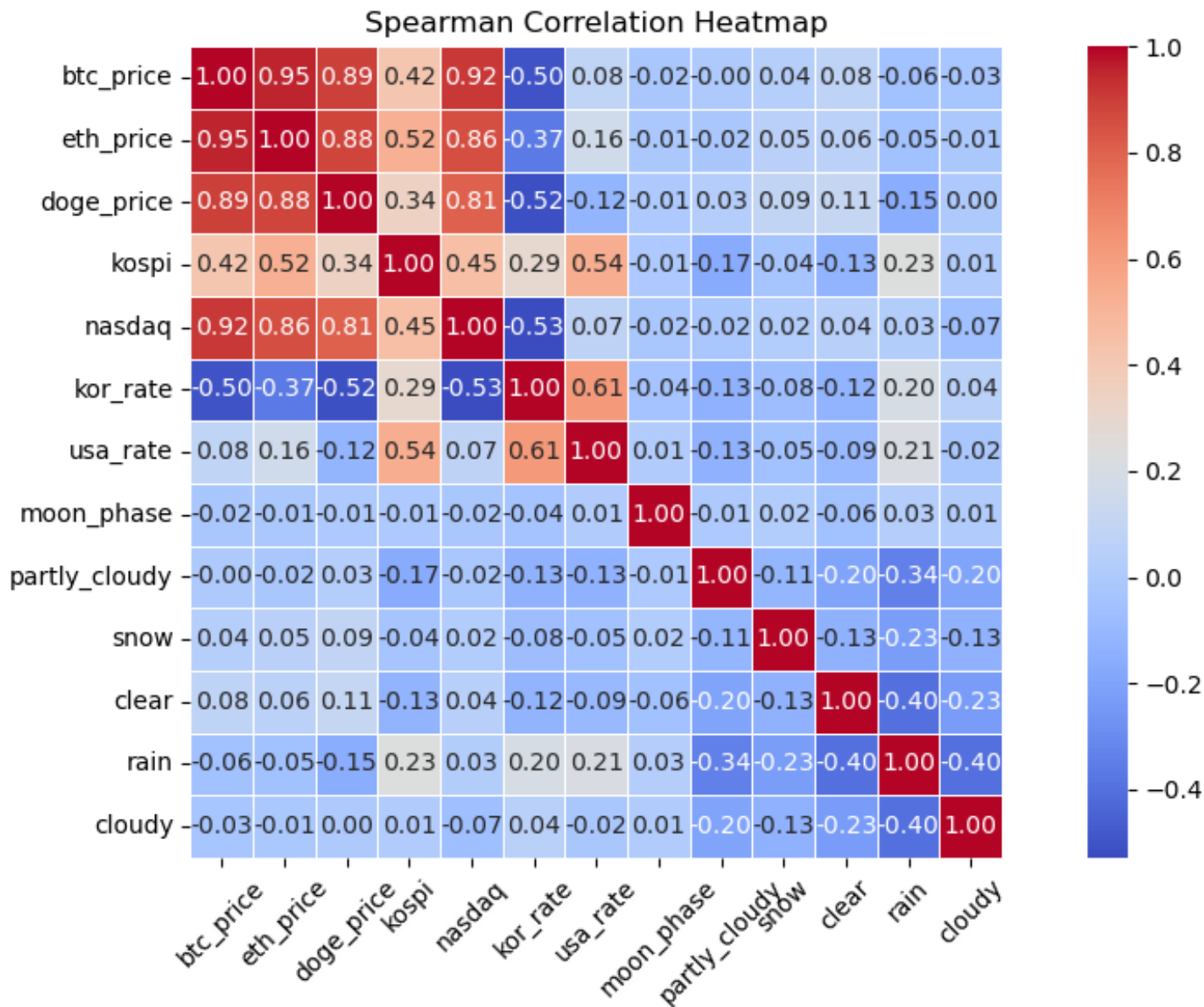
▶ 상관관계 분석(Heatmap)

경제 지표와 암호화폐 가격 간의 관계

- 모든 분석 결과에서 $p\text{-value} < 0.05$ 를 만족한 변수들이 다수 존재하였으며 이에 따라 암호화폐 가격은 경제지표와 상관관계가 있음을 확인.

비정형 요인과 암호화폐 가격 간의 관계

- 달의 위상 및 날씨 조건(맑음, 흐림)과의 관계는 대부분 $p\text{-value} > 0.05$ 로 나타났으며 이에 따라 암호화폐 가격은 비정형 요인과 상관관계가 없음을 확인.



04 프로젝트 수행 결과

결과 제시 ③ 상관관계분석

▶ 상관관계 분석

경제 지표와 암호화폐 가격 간의 관계

- 나스닥 지수와의 상관관계는 매우 강하게 나타남
이는 암호화폐 시장이 글로벌 증시 흐름과 밀접하게 연동되어 있음을 암시함
- 코스피와는 중간 정도의 양의 상관관계를 가짐
- 한국 금리와는 명확한 음의 상관관계를 가짐
금리가 상승하면 암호화폐 가격은 하락하는 경향
- 미국 금리와는 매우 약한 음의 상관관계를 보임
이는 분석에 사용된 업비트 원화 마켓 기준의 가격 데이터 특성
때문으로 판단됨

비정형 요인과 암호화폐 가격 간의 관계

- 날씨 요인은 대부분 통계적으로 유의하지 않음
예외적으로 doge의 경우 맑음, 눈, 비에서 약한 상관관계가 관찰됨
- 달의 위상은 세 코인 모두에서 유의하지 않음

추가 분석 고려

- 보다 정밀한 해석을 위해서는 달러마켓(Binance 등)
데이터와의 비교 분석 필요

04 프로젝트 수행 결과

결과 제시 ④ 선형 관계 분석

▶ 선형관계분석

앞서 수행한 상관관계 분석 결과를 바탕으로 암호화폐 가격에 영향을 미치는 다양한 외부 변수들을 보다 구체적으로 분석하고자 다중 선형 회귀분석을 수행. 비정형 요인 중 통계적으로 유의한 경향을 보였던 도지코인과 날씨 변수 간의 관계도 추가적으로 분석.

- 경제 지표가 암호화폐 가격에 미치는 영향력을 수치적으로 파악
- 독립 변수들의 조합이 종속 변수(암호화폐 가격)를 얼마나 잘 설명하는지 평가
- 비정형 요인인 날씨가 도지코인 가격에 실질적인 영향을 미치는지 회귀 분석을 통해 확인

04 프로젝트 수행 결과

결과 제시 ④ 선형 관계 분석

▶ 선형관계분석(경제지표 기반 회귀분석)

- 종속 변수 : btc_price, eth_price, doge_price
- 독립 변수 : kospi, nasdaq, kor_rate, usa_rate
- 도구 : ols

비트코인

| 변수 | 계수(coef) | P-value | 해석 |
|-------------|-------------------|---------|---------------------|
| kospi | 0.0505 | 0.008 | 양의 영향, 유의함 |
| nasdaq | 0.7593 | 0.000 | 가장 강한 양의 영향, 매우 유의함 |
| kor_rate | -0.2501 | 0.000 | 음의 영향, 금리 오르면 가격 하락 |
| usa_rate | -0.0494 | 0.009 | 약한 음의 영향, 유의 |
| R^2 Score | 0.902 → 설명력 매우 높음 | | |

04 프로젝트 수행 결과

결과 제시 ④ 선형 관계 분석



선형관계분석(경제지표 기반 회귀분석)

이더리움

| 변수 | 계수(coef) | P-value | 해석 |
|-------------|-------------------|---------|---------------|
| kospi | 0.3135 | 0.000 | 강한 양의 영향, 유의함 |
| nasdaq | 0.5563 | 0.000 | 가장 큰 영향력, 유의함 |
| kor_rate | -0.2491 | 0.000 | 음의 영향, 유의함 |
| usa_rate | 0.0471 | 0.127 | 통계적으로 유의하지 않음 |
| R^2 Score | 0.741 → 중간 이상 설명력 | | |

도지코인

| 변수 | 계수(coef) | P-value | 해석 |
|-------------|-------------------|---------|---------------|
| kospi | 0.0257 | 0.321 | 유의하지 않음 |
| nasdaq | 0.3915 | 0.000 | 양의 영향, 유의함 |
| kor_rate | -0.5667 | 0.000 | 강한 양의 영향, 유의함 |
| usa_rate | -0.1099 | 0.000 | 약한 음의 영향, 유의함 |
| R^2 Score | 0.741 → 중간 이상 설명력 | | |

04 프로젝트 수행 결과

결과 제시 ④ 선형 관계 분석



선형관계분석(경제지표 기반 회귀분석)

종합비교

| 항목 | BTC | ETH | DOGE |
|--------------|-------|-------|-------|
| 가장 큰 영향 | 나스닥 | 나스닥 | 한국 금리 |
| 가장 큰 음의 영향 | 한국 금리 | 한국 금리 | 한국 금리 |
| 미국 금리 유의 | 있음 | 없음 | 있음 |
| 설명력(R^2) | 0.902 | 0.741 | 0.821 |

04 프로젝트 수행 결과

결과 제시 ④ 선형 관계 분석



선형관계분석(도지코인-날씨 회귀분석)

- 종속 변수 : doge_price
- 독립 변수 : clear, cloudy, rain, snow, partly_cloudy
- 도구 : ols

- 날씨 조건이 도지코인 가격에 일정 부분 영향을 미칠 수 있음을 시사
- 맑은 날과 눈 오는 날에는 도지코인 가격이 다소 상승하는 경향
- 반면, 비 오는 날에는 유의미한 하락 경향 존재
- 다만 전체 설명력이 낮기 때문에 날씨만으로 도지코인 가격을 설명하기 어려움

도지코인

| 변수 | 계수(coef) | P-value | 해석 |
|---------------|----------------|---------|---------------------|
| clear | 0.1664 | 0.026 | 양의 영향 |
| cloudy | -0.1260 | 0.094 | 유의수준 0.1 부근, 애매한 영향 |
| rain | -0.3212 | 0.000 | 음의 영향 |
| snow | 0.2285 | 0.045 | 양의 영향 |
| partly_cloudy | 0.1386 | 0.098 | 유의수준 0.1부근, 애매한 영향 |
| R^2 | 0.049 → 설명력 나쁨 | | |

04 프로젝트 수행 결과

결과 제시 ⑤ 다중 공선성 확인



해석 기준

VIF 값

해석

1 ~ 5 다중공선성 문제 없음 (안정적) 

5 이상 공선성 존재 가능성 있음 (주의 필요) 

10 이상 심각한 공선성 (모델 재설계 권장) 

다중 선형 회귀분석을 수행할 때, 독립 변수들 간에 서로 높은 상관관계가 존재할 경우 회귀 계수 해석이 왜곡되거나 통계적으로 유의하지 않게 나타날 수 있음

04 프로젝트 수행 결과

결과 제시 ⑤ 다중 공선성 확인



✅ 다중공선성 검사 결과 (VIF)

| 변수 | VIF 값 | 해석 |
|----------|-------|----------------------------|
| kospi | 2.82 | ✅ 다중공선성 문제 없음 |
| nasdaq | 4.47 | ✅ 다중공선성 문제 없음 |
| kor_rate | 5.69 | ⚠️ 약한 다중공선성 가능성 있음 (주의 요망) |
| usa_rate | 2.78 | ✅ 다중공선성 문제 없음 |

- 대부분의 변수는 $VIF < 5$ 로 다중공선성 문제는 크지 않음
- kor_rate 의 경우 $VIF \approx 5.69$ 로 약간의 공선성 가능성이 있으나 분석 해석에 큰 영향을 주지 않는 수준

04 프로젝트 수행 결과

결과 제시 ⑥ 정태적 분석 결론



경제 지표 분석 결과

- 나스닥 지수는 모든 암호화폐(BTC, ETH, DOGE)에 대해 가장 강한 양(+)의 영향력을 보였으며, 특히 비트코인 가격과의 관계가 가장 뚜렷했다.
- 금리(특히 한국 기준금리)는 이더리움, 도지코인에 음(-)의 영향으로 나타나, 금리 인상 시 가격이 하락하는 경향을 보임.
- 코스피 지수는 상관분석에서는 유의미했지만, 회귀분석에서는 도지코인에 유의하지 않은 변수로 나타나 다른 지표에 비해 영향력이 약함을 시사함.



날씨와 도지코인 가격

- 날씨 변수들은 낮은 설명력을 보였으나, 맑음(clear), 눈(snow)은 도지코인 가격 상승과 유의한 양의 관계를, 비(rain)는 유의한 음의 관계를 보임.
- 하지만 결정계수 $R^2 = 0.049$ 로 모델의 설명력이 낮아, 기후 변수 단독으로 도지코인 가격을 설명하기엔 한계가 있음.

04 프로젝트 수행 결과

결과 제시 ⑥ 정태적 분석 결론



다중공선성 검사

- 경제 지표 변수들에 대해 VIF를 확인한 결과, 모두 10 미만으로 다중공선성 문제는 없는 수준임을 확인함.
 - 따라서 회귀 계수 해석에 큰 무리는 없음.
-



정태적 분석 종합 정리

- 암호화폐 가격은 거시경제 지표의 영향을 크게 받으며,
특히 글로벌 금융시장의 흐름을 반영하는 나스닥 지수와의 관계가 가장 강함.
- 날씨 등의 외부 요인을 통해 도지코인과 같은 밈코인의 변동성을 보완적으로 설명할 수 있는 가능성도 확인됨.

04 프로젝트 수행 결과

결과 제시 ⑦ 시차 상관관계 분석



시계열 특성을 고려한 상관 분석

앞선 상관관계 분석은 시계열 데이터를 단순 수치형 변수로 처리하여, 시간 흐름에 따른 영향은 고려하지 않은 정태적 분석이었다. 그러나 암호화폐 가격과 외부 요인은 모두 시계열적 특성을 가지므로 시간의 흐름에 따른 관계성을 분석하는 것이 더 현실적이고 타당하다.

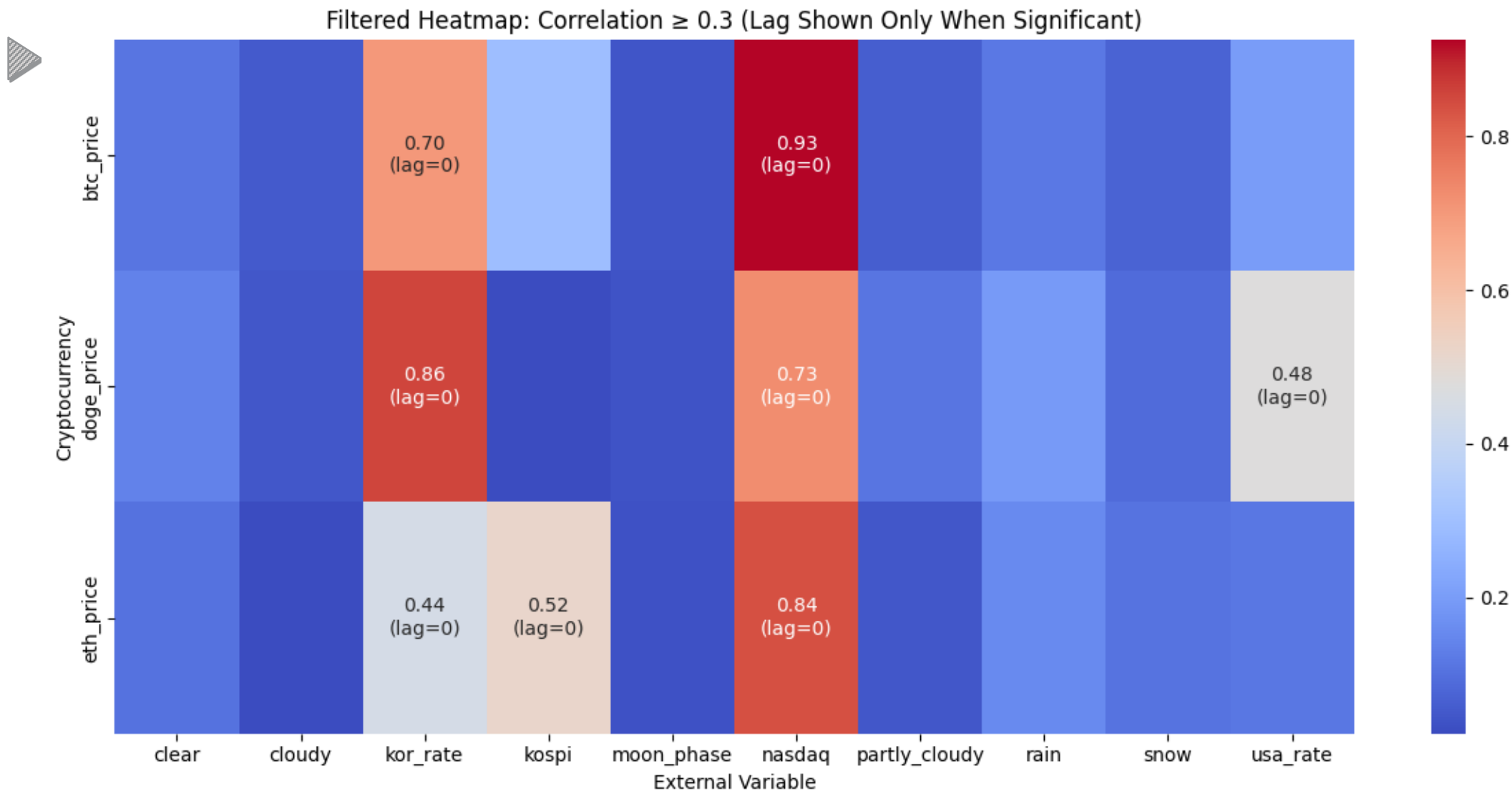


1. 시차 상관관계 분석 (Cross-Correlation Function)

- 암호화폐 가격과 외부 요인 간에 며칠 후에 영향을 주고받는가? 를 파악하기 위해 CCF 분석을 수행하였다.
- 각 변수 쌍에 대해 lag 0~29일 범위의 시차를 적용해 상관계수를 계산하였고 가장 강한 상관을 보인 lag 값을 기록하였다.

04 프로젝트 수행 결과

결과 제시 ⑦ 시차 상관관계 분석



04 프로젝트 수행 결과

결과 제시 ⑦ 시차 상관관계 분석



주요 해석 요약

✅ 전반적인 경향

- 대부분의 외부 요인은 **lag=0**에서 가장 높은 상관관계를 보임
- 이는 **동시 반응**이 대부분이며, 외부 요인이 가격에 즉각 반영된다는 의미
- 시차(lag) 효과가 일부 날씨 요인에서는 존재하지만, 상관계수는 전반적으로 낮음



변수별 상세 해석



`btc_price` (비트코인)

- **NASDAQ**: 0.93 (lag=0) → 매우 강한 동시 상관관계. 미국 증시 움직임이 즉각 반영됨
- **kor_rate**: 0.70 (lag=0) → 한국 기준금리와의 강한 즉시 반응
- 기타 요인(moon_phase, 날씨 등): 상관계수 0.2 이하, 시차 있음 → 영향 미미



`doge_price` (도지코인)

- **kor_rate**: 0.86 (lag=0) → 금리에 매우 민감한 반응. 구조적 요인 가능성
- **NASDAQ**: 0.73 (lag=0) → 주식시장과 강한 연동성
- **usa_rate**: 0.48 (lag=0) → 미국 금리와의 중간 정도 상관
- 그 외 요인들: 전반적으로 낮은 상관 (0.1 이하), 대부분 동시 반응

04 프로젝트 수행 결과

결과 제시 ⑦ 시차 상관관계 분석



◆ eth_price (이더리움)

- NASDAQ: 0.84 (lag=0) → 기술주 중심 시장과 강한 연동
 - kospi: 0.52 (lag=0) → 국내 주가 지수와도 중간 수준 동시 반응
 - kor_rate: 0.44 (lag=0) → 금리 영향 있음
 - 날씨 관련 변수는 모두 0.2 이하의 상관 → 영향 없음
-

📄 요약

- 주요 암호화폐들은 NASDAQ, 금리 등 금융 지표와 동시에 강하게 반응
- 날씨, 달의 위상과 같은 비정형 변수는 시차가 존재해도 상관이 매우 낮아 영향력 없음
- 다만, CCF는 단기 반응만을 측정하므로, 장기적인 관계는 별도로 공적분 분석을 통해 검토함

04 프로젝트 수행 결과

결과 제시 ⑧ 공적분 분석



2. 공적분 분석 (Cointegration Test)

- 비정상 시계열인 암호화폐 가격과 외부 요인 간에 장기적인 균형 관계가 존재하는지를 검증
- Engle-Granger 공적분 검정을 통해 3개 코인 × 외부 요인 전체 조합에 대해 p-value를 산출

| | coin | external_variable | coint_statistic | p_value | | | | | |
|----|------------|-------------------|-----------------|----------|----|------------|---------------|-----------|----------|
| 22 | doge_price | kor_rate | -3.501476 | 0.032265 | 26 | doge_price | snow | -1.132790 | 0.874633 |
| 2 | btc_price | kor_rate | -3.102715 | 0.087759 | 28 | doge_price | rain | -0.922329 | 0.916616 |
| 12 | eth_price | kor_rate | -2.512210 | 0.273711 | 20 | doge_price | kospi | -0.848982 | 0.927955 |
| 11 | eth_price | nasdaq | -2.191571 | 0.428714 | 24 | doge_price | moon_phase | -0.842725 | 0.928853 |
| 1 | btc_price | nasdaq | -2.079673 | 0.486996 | 25 | doge_price | partly_cloudy | -0.783375 | 0.936854 |
| 21 | doge_price | nasdaq | -2.060667 | 0.496934 | 29 | doge_price | cloudy | -0.700654 | 0.946568 |
| 23 | doge_price | usa_rate | -1.834915 | 0.612613 | 27 | doge_price | clear | -0.551988 | 0.960418 |
| 16 | eth_price | snow | -1.569859 | 0.733442 | 3 | btc_price | usa_rate | -0.272230 | 0.977120 |
| 10 | eth_price | kospi | -1.557166 | 0.738625 | 8 | btc_price | rain | 0.260365 | 0.990420 |
| 18 | eth_price | rain | -1.523708 | 0.751989 | 6 | btc_price | snow | 0.311581 | 0.991020 |
| 17 | eth_price | clear | -1.500560 | 0.760977 | 4 | btc_price | moon_phase | 0.605321 | 0.993239 |
| 14 | eth_price | moon_phase | -1.334608 | 0.819048 | 7 | btc_price | clear | 0.714853 | 0.993654 |
| 19 | eth_price | cloudy | -1.328706 | 0.820906 | 9 | btc_price | cloudy | 0.775844 | 0.993806 |
| 15 | eth_price | partly_cloudy | -1.320458 | 0.823477 | 0 | btc_price | kospi | 0.835307 | 0.993903 |
| 13 | eth price | usa rate | -1.211241 | 0.854933 | 5 | btc_price | partly_cloudy | 0.855418 | 0.993925 |

04 프로젝트 수행 결과

결과 제시 ⑧ 공적분 분석



분석 결과

- 대부분의 변수 쌍은 $p\text{-value} > 0.05$ 로 공적분 관계가 없는 것으로 나타남
- 단, 다음 조합에서 $p\text{-value} < 0.05$ 로 유의미한 공적분 관계가 존재

| 코인 | 외부 요인 | 통계량 | p-value | 해석 |
|------------|----------|-----------|----------|-------------------------------|
| doge_price | kor_rate | -3.501476 | 0.032265 | 도지코인과 한국 기준금리는 장기적으로 균형관계를 유지 |



해석

- doge_price 와 kor_rate 의 공적분 관계는,
→ 도지코인 가격과 한국 기준금리 간의 장기적 연동성이 존재할 수 있음을 시사
- 즉, 단기적으로 이탈이 있더라도 시간이 지나면서 다시 수렴하는 성향이 있음
- 반면, 대부분의 외부 요인(금리, 주가지수, 날씨 등)과 코인 가격 간에는
→ 장기적 균형관계 없음 → 가격이 따로 움직이는 독립적 경향

04 프로젝트 수행 결과

결과 제시 ⑨ 동태적 분석 결론



요약

- 공적분 분석 결과, 단 하나의 조합(`doge_price - kor_rate`)을 제외하고는 암호화폐 가격과 외부 요인 간에 장기적인 통계적 연관성은 확인되지 않음
 - 이는 암호화폐가 여전히 전통적 외부 요인과 구조적으로 연결되어 있지 않다는 점을 시사함
-



동태적 분석 종합 정리

- 암호화폐는 주요 금융지표(금리, 주가지수 등)와 즉시 반응하는 관계를 가지며,
- 날씨, 달의 위상 등의 변수는 단기/장기 모두에서 통계적으로 유의한 관계 없음
- 일부 금리 변수는 특정 코인과 장기적 구조적 연관성을 가질 수 있으며, 이는 향후 예측 모델 설계나 정책적 해석에 있어 중요한 단서가 될 수 있음

04 프로젝트 수행 결과

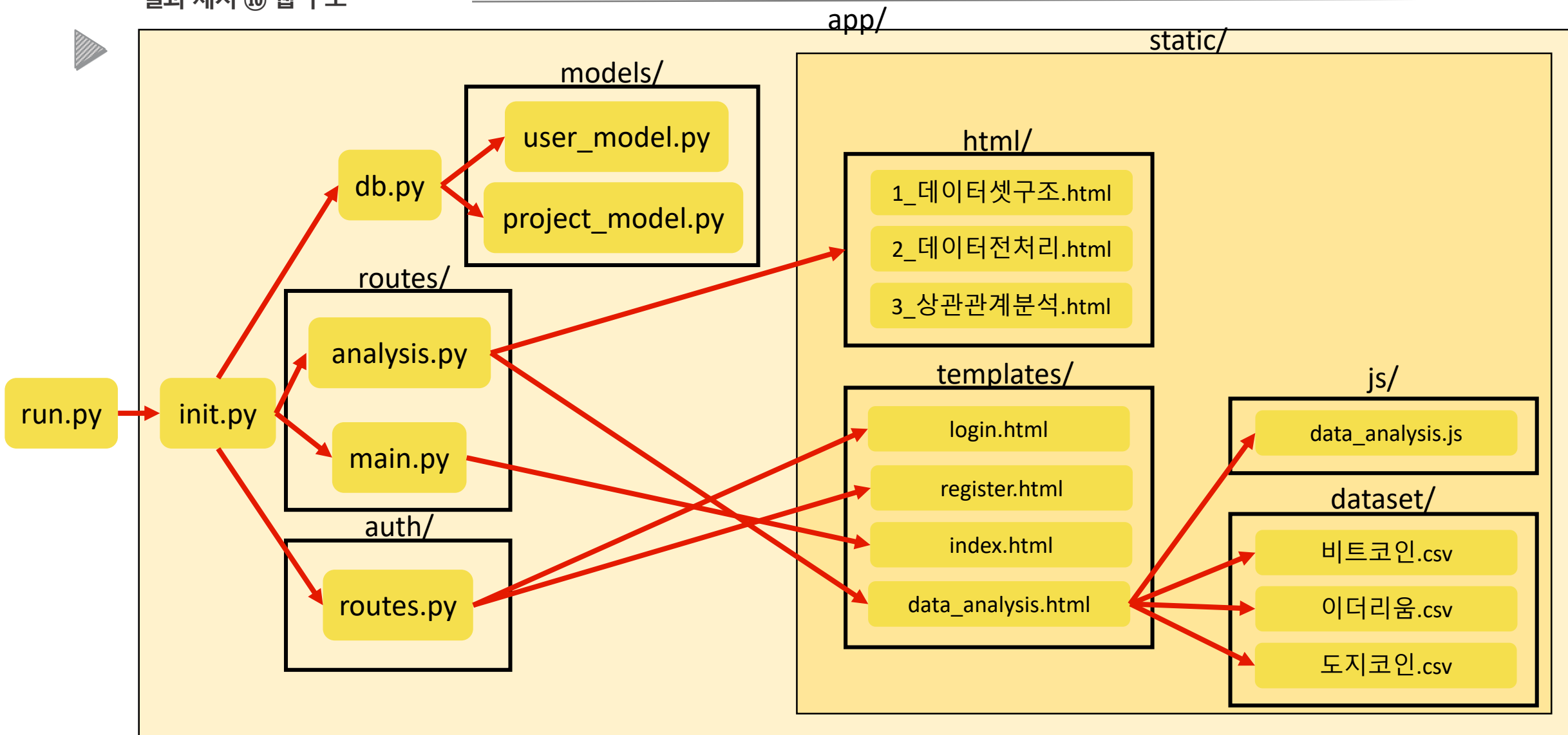
결과 제시 ⑨ 분석 요약



| 항목 | 요약 |
|--------|--|
| 무관한 요인 | 날씨, 달의 위상은 단기/장기 모두에서 상관 관계 없음 |
| 즉각적 요인 | kospi, nasdaq, 금리 등은 대부분 바로 암호화폐 가격에 바로 반응 |
| 장기적 요인 | doge_price - kor_rate 조합만 장기적 관계 확인 |
| 시사점 | 금리, 지수 기반의 변수는 추후 모델링 변수로 활용가치 존재, 외부 요인은 제외 해야함 |

04 프로젝트 수행 결과

결과 제시 ⑩ 웹 구조



04 프로젝트 수행 결과

결과 제시 ⑪ 주요 라우팅 기능



| 라우트 파일 | 주요기능 | 연결 템플릿 | 템플릿 구조 |
|-----------------------------|---|-----------------------------|-------------------------------------|
| main.py | 메인페이지 로딩 / 프로젝트 카드 조회 | index.html | 프로젝트 썸네일 카드 리스트 |
| analysis.py | 분석 결과 리포트 출력(HTML 시각화) | data_analysis.html | 정적 리포트(html) + 동적 그래프 시각화(Chart.js) |
| auth/routes.py | 로그인, 로그아웃, 회원가입 처리 (세션 기반) | login.html register.html | 입력 폼 및 오류 출력 구성 |
| app/db.py | MySQL 서버와 연결 | - | - |
| app/models/user_model.py | 회원가입 시 이메일/닉네임 중복 검사 Bcrypt를 사용한 비밀번호 암호화 / 검증 | - | - |
| app/models/project_model.py | 프로젝트 목록 불러오기 개별 프로젝트의 상세 정보(제목, 썸네일, 실행파일 경로) | - | - |

04 프로젝트 수행 결과

결과 제시 ⑫ 데이터베이스



| projects | |
|--------------------|-----------------|
| project_id | 프로젝트 ID |
| project_title | 프로젝트 제목 |
| project_image_path | 프로젝트 썸네일 경로 |
| project_model_path | 프로젝트 모델 경로 |
| project_html_path | 프로젝트 html리포트 경로 |
| project_created_at | 프로젝트 생성 시간 |
| users | |
| id | 회원 ID |
| email | 회원 이메일 |
| nickname | 회원 닉네임 |
| password_hash | 암호화 된 비밀번호 |
| created_at | 생성 시간 |

04 프로젝트 수행 결과

결과 제시 ⑬ 웹 퍼블리싱

▶ 퍼블리싱 환경 및 접근 구조

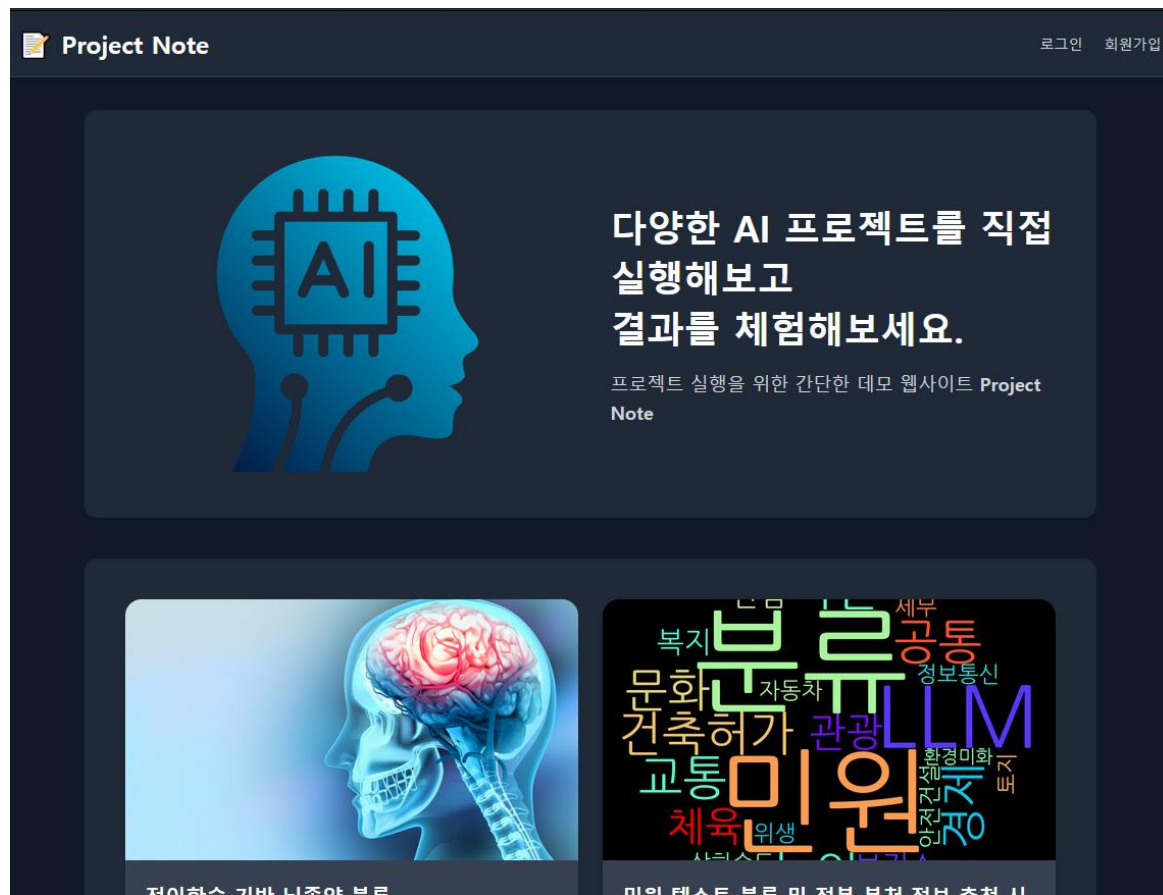
본 프로젝트는 Flask 기반 웹 애플리케이션으로 외부 사용자도 직접 접속하여 분석 결과를 확인할 수 있도록 퍼블리싱 환경을 구축하였다.

서버 환경 : Ubuntu 24.04(직접 운영 중인 PC 기반 서버)

- Gunicorn + Cloudflare Tunnel을 사용하여 외부에서 접속 가능
- Flask 웹 서버를 직접 실행한 후 포트를 외부로 노출하지 않고 보안 터널을 통해 도메인 연결

04 프로젝트 수행 결과

결과 제시 ⑭ 링크



johnnyim.ai.kr

05 자체 평가 의견



- 본 프로젝트는 암호화폐 가격과 외부 요인 간의 상관관계를 분석하고 이를 웹 기반 시각화 시스템으로 구현하는 것을 목표로 기획하였다. 기획 단계에서 설정한 데이터 분석 -> 가설 검증 -> 시각화 및 웹 서비스화의 전 과정을 실제로 구현하였으며 사용자 입력에 따라 실시간으로 분석 결과를 확인할 수 있는 구조로 기획 의도와 높은 부합도를 보였다.
- Flask 기반 웹 서버, MySQL 연동 Chartjs 기반 프론트엔드로 프로젝트를 구성하여 기술 스택을 종합적으로 활용하였으며 우분투 기반 개인 서버에 도메인을 연결해 외부 접속 가능한 상태로 퍼블리싱을 완료하였다. 데이터 분석부터 시각화, 모델 연동, 웹 구현까지 전 과정을 직접 수행함으로써 실무 개발의 전체 흐름을 경험해 볼 수 있었다.
- 분석 결과 암호화폐 가격은 달의 위상과는 유의한 관계가 없었고, 날씨와는 일부 암호화폐에서 상관관계를 보였으나 다중공선성 분석 결과 날씨 변수는 예측 모델의 독립 변수로 활용하기엔 부적합하다는 한계를 확인하였다. 이러한 점에서 향후 예측 모델 설계 시 변수 선택에 보다 신중한 접근이 필요함을 느낄 수 있었다.
- 팀 구성 면에서는 프로젝트 초기에는 2인이 함께 착수하였으나, 중도 이탈로 인해 대부분의 분석 및 개발 과정을 혼자서 수행하게 되었고 이로 인해 로그인/회원가입 외에 추가적인 사용자 기능을 확장하지 못한 점은 아쉬움으로 남는다.