

PJT명	넷플릭스 주가 데이터 분석	
단계	02 PJT	
진행일자	2025.08.01	
예상 구현 시간	필수기능	6H
	심화기능	2H

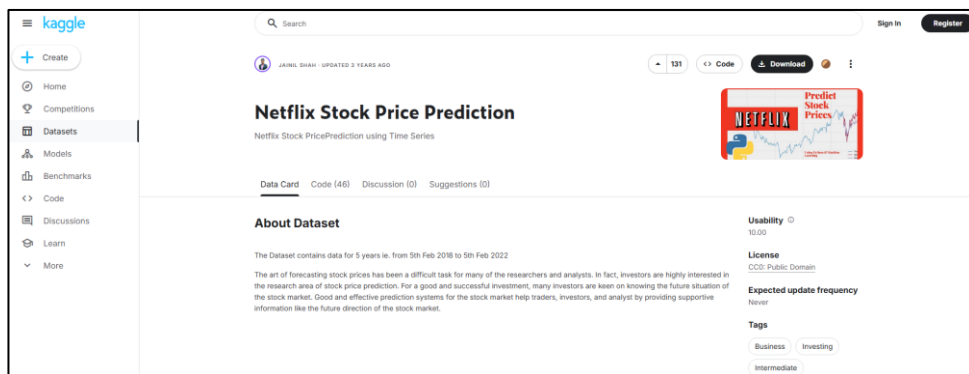
1. 목표

- 데이터 분석의 전체 흐름을 이해하고 응용할 수 있다.
- 요구사항에 맞는 데이터를 수집하고 전처리 할 수 있다.
- 시각화를 통해 데이터를 직관적으로 표현할 수 있다.

2. 준비사항

1) 필수 준비사항

- 데이터
 - Kaggle - [넷플릭스 주가 데이터](#)



2) 개발언어 및 툴

- Python
- Jupyter Notebook

3) 필수 라이브러리/오픈소스

- Numpy
- Pandas
- Matplotlib

3. 작업 순서

- 1) 페어와 같이 요구사항(기본/심화)을 분석한다.
- 2) 페어와 함께 개발 시 역할을 분담한다.
 - 팀장/네비게이터 등 역할 분담
- 3) [선택] 추가적인 아이디어를 적용하여 구현하고 프로젝트를 완성한다.
- 4) 코드 및 산출물을 정리하여 Git에 반영한다.

4. 요구사항

본 프로젝트는 넷플릭스 주가 데이터를 활용한 시계열 데이터 분석 프로젝트로, Pandas와 Matplotlib 등 데이터 분석에 필요한 핵심 라이브러리를 실습하며 데이터 사이언스의 기초를 학습하는 것을 목표로 한다.

외부 데이터를 활용한 전처리, 분석, 시각화, AI 기반 예측까지의 전체 흐름을 직접 수행함으로써 실무와 유사한 데이터 처리 경험을 제공한다.

이를 통해 현실 데이터를 다루는 실무 감각, 데이터 흐름에 대한 전반적 이해, 생성형 AI와의 결합 가능성 등을 학습하며, 향후 실사용 데이터 기반 애플리케이션 개발을 위한 기반 역량을 함양할 수 있다.

● 요구사항 명세

- 아래에 제시된 기능 명세에 맞는 기능을 구현한다.
- 필수 기능은 반드시 구현해야 하며, 심화 기능은 팀 내 기획에 따라 선택적으로 구현한다.

번호	분류	요구사항명	요구사항 상세	우선순위
기능적 요구사항				
F01	데이터 전처리	CSV 파일 읽기 및 필드 선택	Netflix 주가 데이터(NFLX.csv)를 Pandas를 사용해 DataFrame으로 읽어온다. 이 때, 'Date', 'Open', 'High', 'Low', 'Close' 필드만 선택한다.	필수

F02	데이터 전처리	날짜 필터링	2021년 이후 데이터 필터링	필수
F03	데이터 분석	최고/최저 증가 추출	2021년 이후 데이터에서 증가(Close)의 최대/최소 값을 추출	필수
F04	데이터 분석	월별 평균 증가 계산	2021년 이후 데이터를 월별로 그룹화하여 평균 증가를 계산하여 시각화	필수
F05	데이터 시각화	월별 최고/최저/증가 시각화	2022년 이후 데이터를 바탕으로 3가지 지표를 하나의 그래프에 시각화	필수
F06	생성형 AI	주가 예측 프롬프트 작성	생성형 AI에게 적절한 프롬프트를 설계하고 주가에 대한 예측 결과를 도출	심화
...
비 기능적 요구사항				
NF01	환경설정	라이브러리 설치	Pandas, Matplotlib 등 필요한 패키지를 설치하고 사용할 수 있도록 환경을 구성	필수
NF02	코드 품질	주석 및 가독성	주요 분석 로직에 주석을 작성하고, 직관적 변수명과 함수 구조로 코드를 구성	필수
NF03	효율성	함수 분리	중복 코드를 방지하고 기능별로 함수로 나누어 구조화	필수
NF04	문서화	README 작성	분석 흐름과 학습 내용, 어려웠던 점을 README.md에 기록하여 제출	필수
...

1) 기본(필수) 기능

본 프로젝트는 Kaggle에서 제공하는 Netflix 주가 데이터를 활용하여, Pandas 및 Matplotlib을 통해 시계열 데이터 전처리 → 분석 → 시각화의 전체 흐름을 실습하는 것을 목표로 한다.

각 기능을 함수 단위로 구현하며, 실제 데이터 처리와 해석의 기본 역량을 함양하게 된다.

A. 데이터 수집

- Kaggle에서 넷플릭스 주가 데이터를 다운로드한다.
 - 링크: [Netflix Stock Price Prediction](#)

B. 데이터 전처리

- Pandas를 활용하여 CSV 파일을 불러온다. (요구사항 F01)
 - CSV 구조를 확인하고 필요한 컬럼만 선별한다.
 - ✓ 사용 필드: Date, Open, High, Low, Close
 - 결과물 예시

```
[3]: # DataFrame 출력  
df
```

```
[3]:
```

	Date	Open	High	Low	Close
0	2018-02-05	262.000000	267.899994	250.029999	254.259995
1	2018-02-06	247.699997	266.700012	245.000000	265.720001
2	2018-02-07	266.579987	272.450012	264.329987	264.559998
3	2018-02-08	267.079987	267.619995	250.000000	250.100006
4	2018-02-09	253.850006	255.800003	236.110001	249.470001
...
1004	2022-01-31	401.970001	427.700012	398.200012	427.140015
1005	2022-02-01	432.959991	458.480011	425.540009	457.130005
1006	2022-02-02	448.250000	451.980011	426.480011	429.480011
1007	2022-02-03	421.440002	429.260010	404.279999	405.600006
1008	2022-02-04	407.309998	412.769989	396.640015	410.170013

```
1009 rows x 5 columns
```

- 2021년 이후의 데이터만 필터링한다. (요구사항 F02)
 - 필터링 완료된 DataFrame의 종가 데이터를 그래프로 시각화한다.
 - 결과물 예시



C. 데이터 분석

- 2021년 이후 데이터에서 종가(Close)의 최대/최소 값을 추출한다. (요구사항 F03)
 - 결과물 예시

```
[7]: print("최고 종가:", max_price)
      print("최저 종가:", min_price)
```

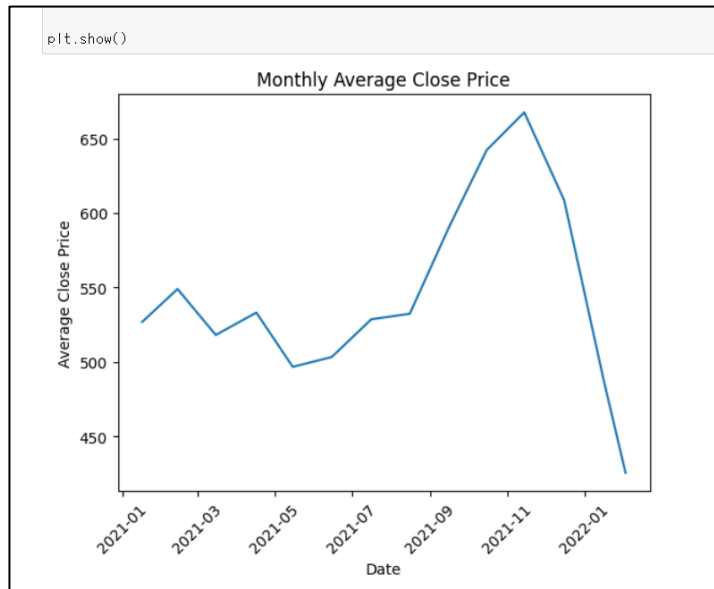
최고 종가: 691.690002

최저 종가: 359.700012

- 2021년 이후 데이터를 월별로 그룹화하여 평균 종가 계산한다.
(요구사항 F04)

■ 계산 결과를 DataFrame으로 저장하여 그래프로 시각화한다.

■ 결과물 예시



D. 시각화

- 2022년 이후의 데이터를 필터링하여, 월별 최고가, 최저가, 종가를 함께 표현하는 복합 라인 그래프를 작성한다. (요구사항 F05)

■ 결과물 예시



2) 심화 기능

본 심화 기능은 생성형 AI(OpenAI, Gemini 등)를 활용하여 넷플릭스 주가 데이터를 기반으로 자동 해석 또는 예측을 수행하는 기능을 중심으로 구성된다.


주가 시계열 데이터를 분석한 후, 생성형 AI에게 의미 있는 질문을 구성하고, AI의 응답을 통해 인사이트를 도출하는 흐름으로 진행된다.

A. 생성형 AI를 활용한 주가 예측

- OpenAI, Gemini 등 **생성형 AI**에게 주가 데이터를 설명하거나 예측하는 **질문 프롬프트**를 구성한다. (요구사항 F06)
- 매수/매도 판단 또는 데이터 해석이 가능하도록 프롬프트를 구성한다.
- 질문 예시

2022년 1월 이후 Netflix 주가는 상승세인가요? 주요 원인은 무엇인가요?

- 답변 예시

 **2022년 1월 이후 Netflix 주가는 상승세가 아닙니다. 오히려 크게 하락한 추세입니다.**

주요 하락 원인 (2022년 실제 배경 기반 요약)

1. 가입자 증가세 둔화

- 2022년 1분기 실적 발표 당시, 넷플릭스는 10여 년 만에 처음으로 **가입자 수가 감소**했다고 발표했으며 이는 투자 심리에 큰 타격을 주었습니다.

2. 스트리밍 시장 경쟁 심화

- 디즈니+, HBO Max, Amazon Prime Video 등 경쟁 서비스의 공격적인 콘텐츠 투자로 인해 넷플릭스의 점유율이 흔들리기 시작했습니다.

3. 금리 인상 우려 및 기술주 하락세

- 2022년 초부터 미국 연준(Fed)의 긴축 정책에 따라 **기술주 전반의 주가가 급락**했고, 넷플릭스도 그 영향을 받았습니다.

5. 참고자료

- 데이터셋 출처
 - Netflix Stock Price Prediction (Kaggle)
 - ◆ <https://www.kaggle.com/datasets/jainilcoder/netflix-stock-price-prediction>
- 공식 문서
 - Numpy 공식 문서
 - ◆ <https://numpy.org/doc/>
 - Pandas 공식 문서
 - ◆ <https://pandas.pydata.org/docs/>
 - Matplotlib 공식 문서
 - ◆ <https://matplotlib.org/stable/index.html>
 - ◆ [Matplotlib 시각화 예제 모음](#)

6. 결과

최종적으로 제출해야 할 항목은 아래와 같다

1. 구현 소스 코드

- Jupyter Notebook에서 작성한 ipynb 파일

2. 실행 결과 캡처본

- 콘솔 출력 결과 등

3. README.md

- 구현 기능 설명, 학습 내용, 느낀 점 포함

위 내용들을 모두 포함하여 GitLab에 업로드 한다.

GitLab 프로젝트 이름은 프로젝트 번호 + _pjt 로 지정한다 (02_pjt)