

# 기후정보를 이용한 국내 주식 업종주가 예측

신은총, 안수현, 박수민, 최용호

# 목차

- I. 주제 선정
- II. 사전 정의
- III. 프로젝트 진행 계획
- IV. 데이터 수집 및 가공
- V. 분석 계획

# 주제 선정

- 주가는 폭염, 황사, 등 기후적인 요소에 의해 변동되기도 함
- 어떤 기후적인 요소가 주식 변동에 많은 영향을 주는 지 궁금함이 생김



# 참고: 계절에 따라 주가 오르내린다? 계절주

- **계절주**란 주식시장에 상장된 기업 가운데 회사의 매출 또는 수익이 계절에 따라 크게 변하는 기업의 주식을 뜻한다.

|   |  |
|---|--|
| <b>봄 &lt;황사 미세먼지&gt;</b><br><br>마스크, 공기청정기 등 제조 및 판매 업체:<br>위닉스, 코웨이, 웰크론 등 | <b>여름 &lt;지구온난화, 더위, 장마, 태풍&gt;</b><br><br>선풍기, 제습기 등 판매 회사:<br>신일산업, 파세코, 인션이엔티 |
| <b>가을 &lt;환절기, 질병&gt;</b><br><br>제약 백신주:<br>파루, 제일바이오, 대한뉴팜 등               | <b>겨울 &lt;추위&gt;</b><br><br>난방기구 생산 업체:<br>한국가스공사, 지역난방공사, 신일산업 등                |



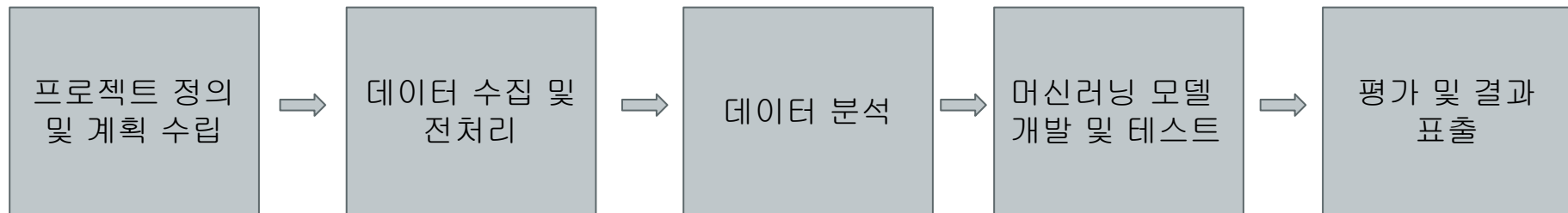
# 사전 정의

- 업종별: WICS(Wise Industry Classification Standard)의 산업 분류 기준을 따라 업종을 구획화함(현재 코스피의 업종 분류 개수는 74개)

| WICS 대분류 |     | WICS 중분류 |     | WICS 소분류 |           |
|----------|-----|----------|-----|----------|-----------|
| 10       | 에너지 | 1010     | 에너지 | 101010   | 에너지장비및서비스 |
|          |     |          |     | 101020   | 석유와가스     |
| 15       | 소재  | 1510     | 소재  | 151010   | 화학        |
|          |     |          |     | 151030   | 포장재       |
|          |     |          |     | 151040   | 비철금속      |
|          |     |          |     | 151050   | 철강        |
|          |     |          |     | 151060   | 종이와목재     |
| 20       | 산업재 | 2010     | 자본재 | 201010   | 우주항공과국방   |
|          |     |          |     | 201020   | 건축제품      |

- 기후적인 요소: 기상자료개방포털(<https://data.kma.go.kr/cmmn/main.do>)에서 얻을 수 있는 기상자료로서, 기준점은 서울로 둠
- 업종주가: 업종별 시가총액(=주식수\*당일가격)

# 프로젝트 진행 계획



# 데이터 수집 및 전처리

## 기후 데이터

- 기상자료개방포털에서 평균 기온/강수량/폭염/황사/미세먼지 농도 등의 데이터를 가져옴
- 2010.01.04 ~ 2019.09.02까지 수집함
- 데이터를 일별로 정리 및 각 기후 증상의 하나의 데이터프레임으로 합침

## KOSPI 데이터

- KRX(<https://kind.krx.co.kr/corpgeneral/corpList.do?method=loadInitPage>)에서 회사명과 종목코드를 크롤링
- Yahoo Finance(<https://finance.yahoo.com/>)에서 종목코드를 이용하여 KOSPI 주식 데이터를 일별로 가져옴(총 785개)
- 2010.01.04 ~ 2019.09.03까지 수집함
- 다음 금융에서 종목 코드를 이용하여 WICS 크롤링
- WICS 분류별로 데이터 정리

# 분석 계획

- 데이터 시각화
  - 기초 통계량 뿐만 아니라 시각화를 통해 어떤 패턴이 존재 하고 상관 관계를 파악
- 데이터 마이닝
  - 기상 데이터와 주가 데이터를 분석하여 알지 못했던 통계적 규칙이나 패턴을 찾기
- 데이터 모델링
  - 모델의 과적합과 일반화를 위해 분석용 데이터 셋을 훈련용과 테스트용으로 나눈 후 예측 모델을 만들



# Q&A