



내일주식시장은맑겠조

날씨와 주가의 상관관계 분석



목차

#1 프로젝트 소개

#2 데이터 설명

#3 데이터 분석

#4 분석 결과

#5 프로젝트 마무리

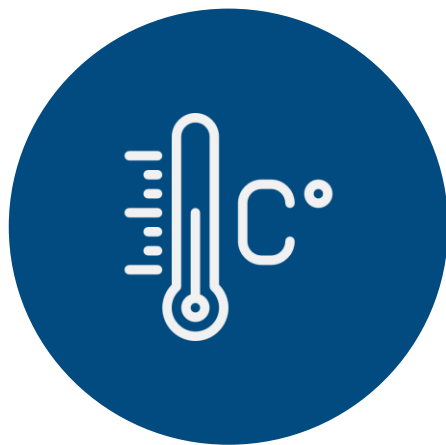


Part 1

프로젝트 소개

1. 팀원 소개
2. 주제 선정

팀원 소개



조휘

조장

기온 데이터

전처리, 분석, 시각화



조세익

불쾌지수 데이터

전처리, 분석, 시각화

발표



임채건

맑음, 흐림 데이터

전처리, 분석, 시각화

ppt 제작



한상준

태풍, 강수량 데이터

전처리, 분석, 시각화

주제 선정

기분이 좋은 상태에서는 긍정적인 방향으로 평가하려는 경향이 강하며,
이는 의사결정에 직접적인 영향을 미친다.

-
Bagozzi, Gopinath, and Nyer(1999) The role of emotions in marketing

뉴욕 시의 날씨와 뉴욕 증권 거래소의 주가 지수를 분석한 결과,
흐린 시간이 긴 날의 주식 수익률은 흐린 시간이 짧은 날에 비해서 유의적으로 낮다

-
Edward M. Saunders, Jr. (1993) Stock Prices and Wall Street Weather, The American Economic Review, Vol. 83, No. 5

일조량과 주식 연간 수익률 간의 관계를 분석한 결과,
뉴욕의 경우 하루 종일 해가 난 날의 연간 수익률은 평균 24%로
하루 종일 흐린 날의 연간 수익률인 8% 보다 유의적으로 높다

-
David Hirshleifer, Tyler Shumway (2003) Good Day Sunshine: Stock Returns and the Weather



행동투자론

주식을 구매하는데 있어, 이성적 판단에 근거하지 않고,
날씨, 기분 등이 소비자의 심리에 영향을 끼쳐 주식을 구매한다는 이론



Part 2

데이터 수집

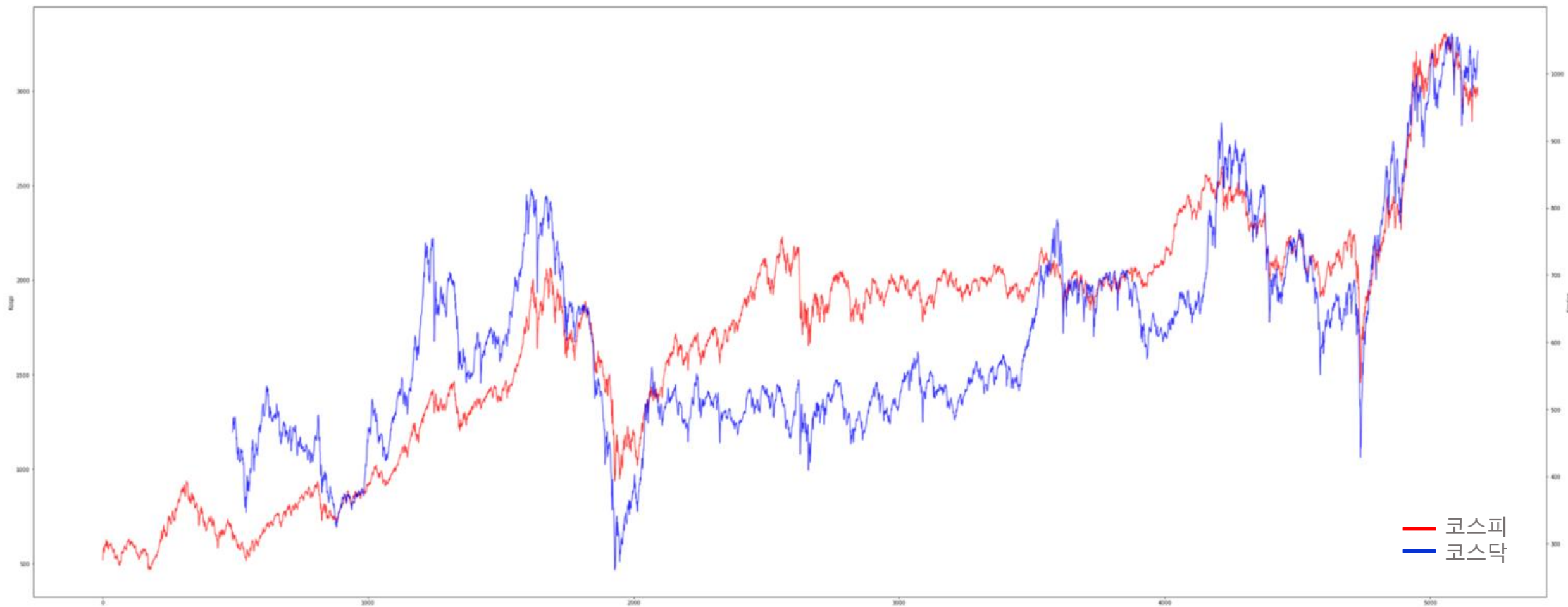
1. 데이터 출처 및 크기
2. 데이터

데이터 출처 및 크기

주가 데이터	KOSIS	5181rows x 7columns
기온 데이터	기상청	7670rows x 5columns
습도 데이터	기상청	7670rows x 5columns
흐림 데이터	기상청	7670rows x 5columns
강수량 데이터	기상청	7670rows x 3columns
태풍 데이터	기상청	576rows x 6columns

KOSIS (국가통계포털) – 주식시장(일별)
기상청 기상자료 개방포털 – 종관기상관측

데이터 — 일별 주가 데이터



데이터 — 일별 주가 데이터

	Date	kospi	volume_kospi	amount_kospi	kosdaq	volume_kosdaq	amount_kosdaq
0	2001-01-02	520.95	23101.0	12233.0	NaN	NaN	NaN
1	2001-01-03	521.43	32458.0	15834.0	NaN	NaN	NaN
2	2001-01-04	558.02	44454.0	29848.0	NaN	NaN	NaN
3	2001-01-05	580.85	57829.0	35696.0	NaN	NaN	NaN
4	2001-01-08	586.65	55865.0	29481.0	NaN	NaN	NaN
...
5177	2021-12-24	3012.43	53015.0	84694.0	1007.42	123576.0	108664.0
5178	2021-12-27	2999.55	46290.0	81341.0	1011.36	115153.0	100446.0
5179	2021-12-28	3020.24	59323.0	118684.0	1027.44	128675.0	132431.0
5180	2021-12-29	2993.29	53660.0	108018.0	1028.05	95094.0	116836.0
5181	2021-12-30	2977.65	45404.0	85951.0	1033.98	90742.0	101026.0

>>

시간이 지남에 따라 주가지수의 규모도 상승하는 추세

단순히 주가지수만으로 분석하기는 어려움

분석에 용이하도록 새로운 파생변수 필요

파생변수 추가

전일대비
증감수치

금일종가 - 전일종가
전일 대비 변동수치

전일대비
증감률

금일변동률 = (금일종가 - 전일종가) / 전일종가 X 100
pct.change() 함수를 이용하여 증감률 추가

전일대비
상승/하락

증감률을 이용하여 0보다 크면 상승, 0보다 작으면 하락 값 추가

그룹화

각 데이터들의 날짜들에 구간을 생성하여 구간별로 분석



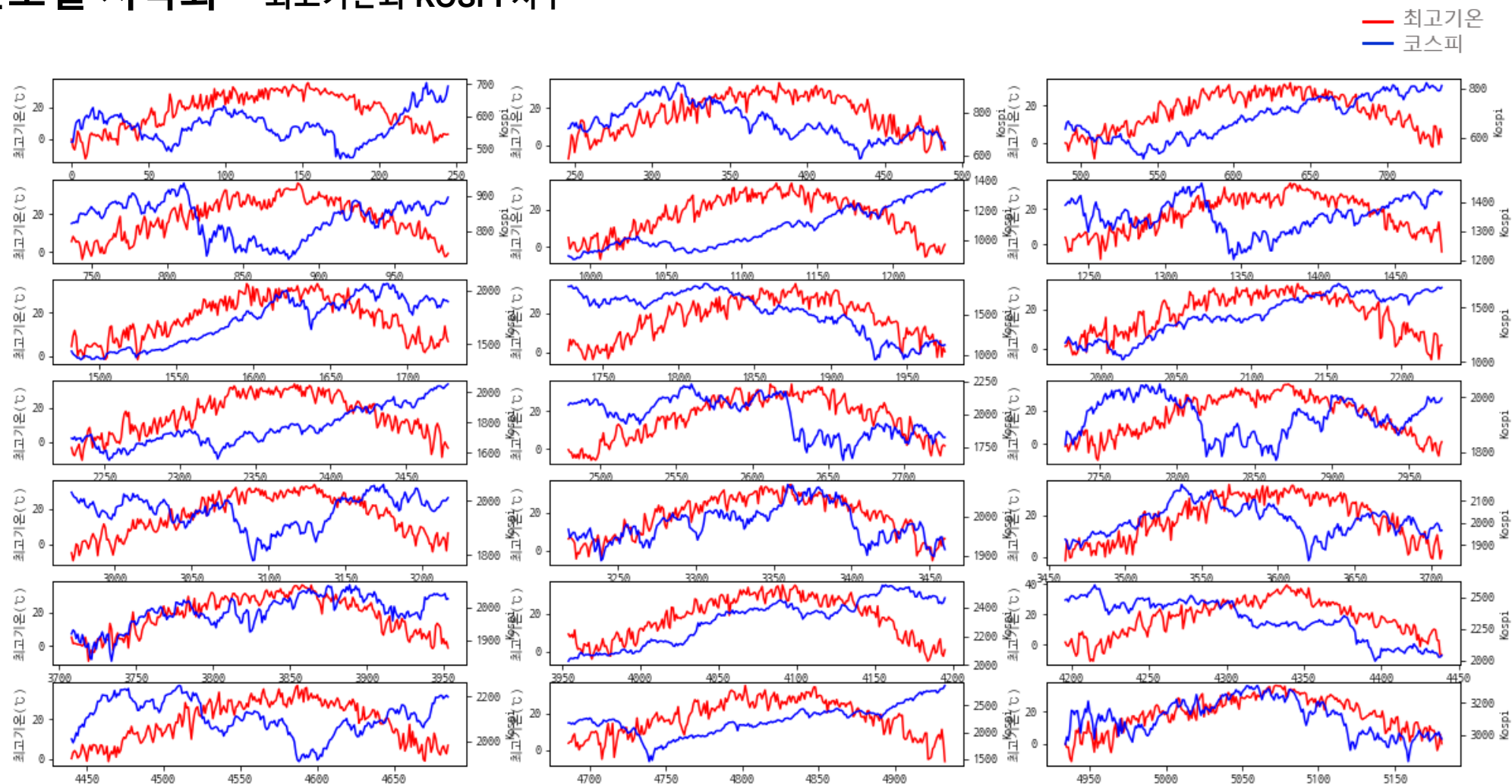
Part 3

데이터 분석

1. 기온 (최고기온)
2. 불쾌지수
3. 전운량
4. 태풍

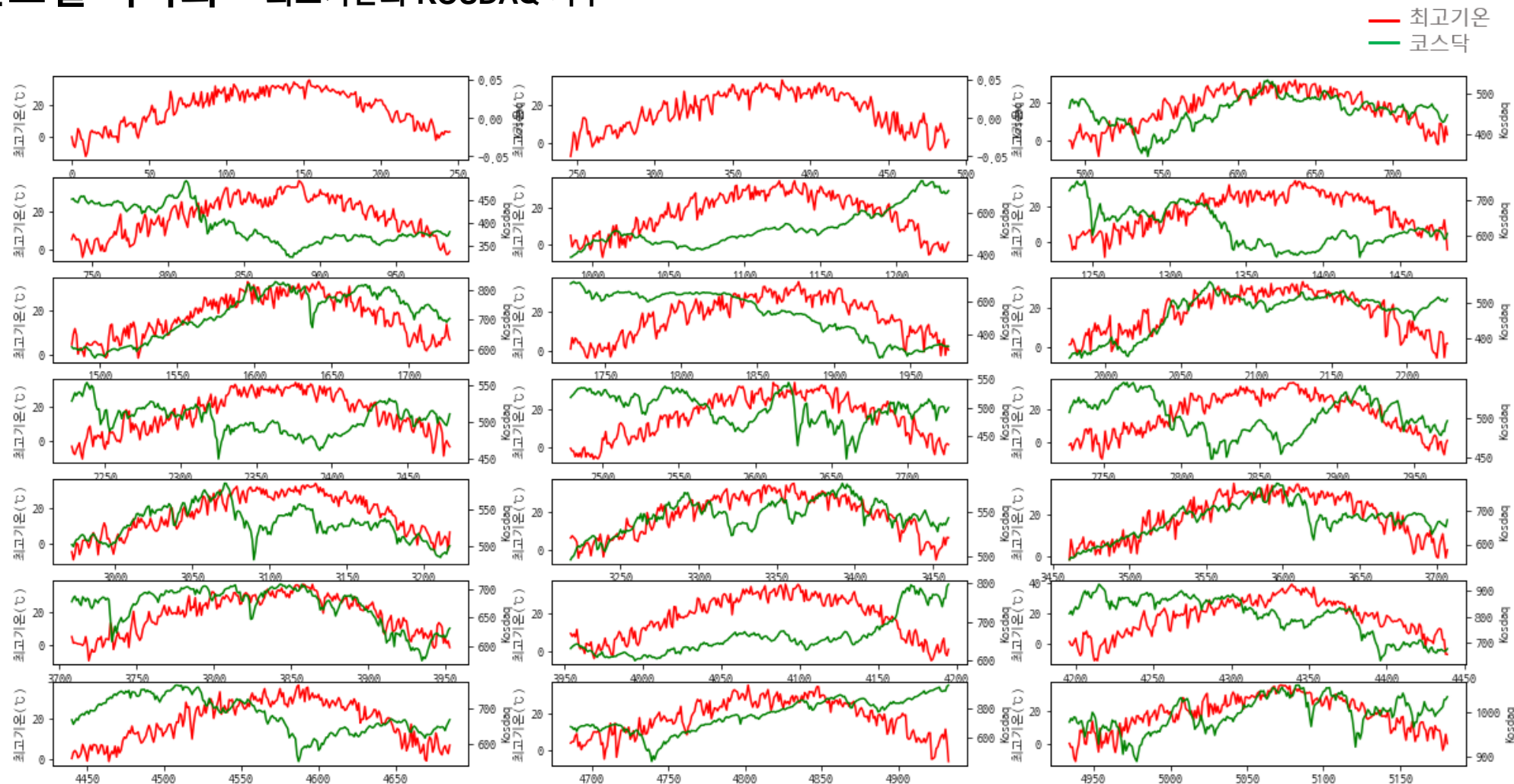
데이터 분석 — 최고 기온과 주가 분석

전체 연도별 시각화 — 최고기온과 KOSPI 지수



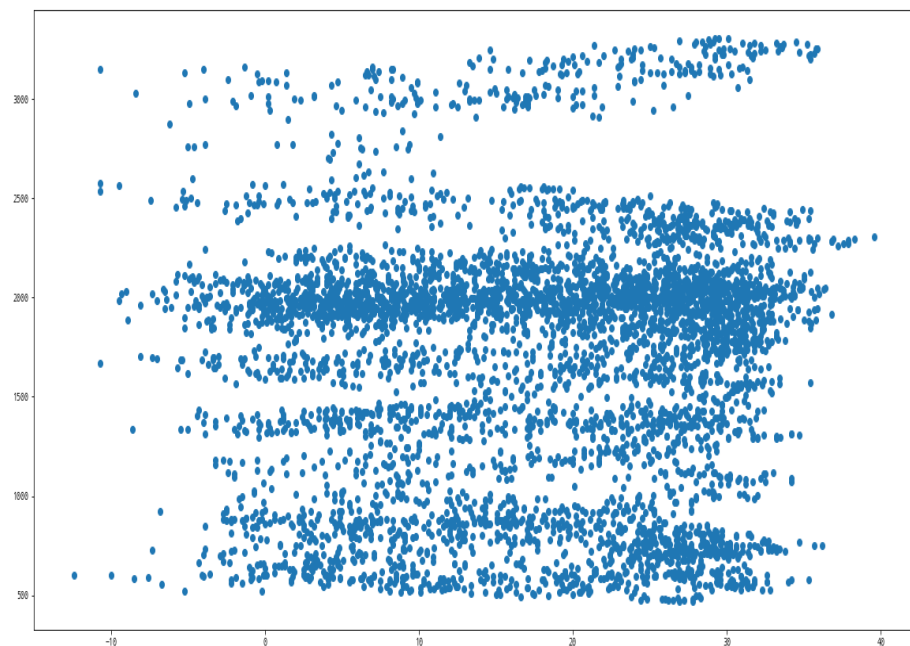
데이터 분석 — 최고 기온과 주가 분석

전체 연도별 시각화 — 최고기온과 KOSDAQ 지수

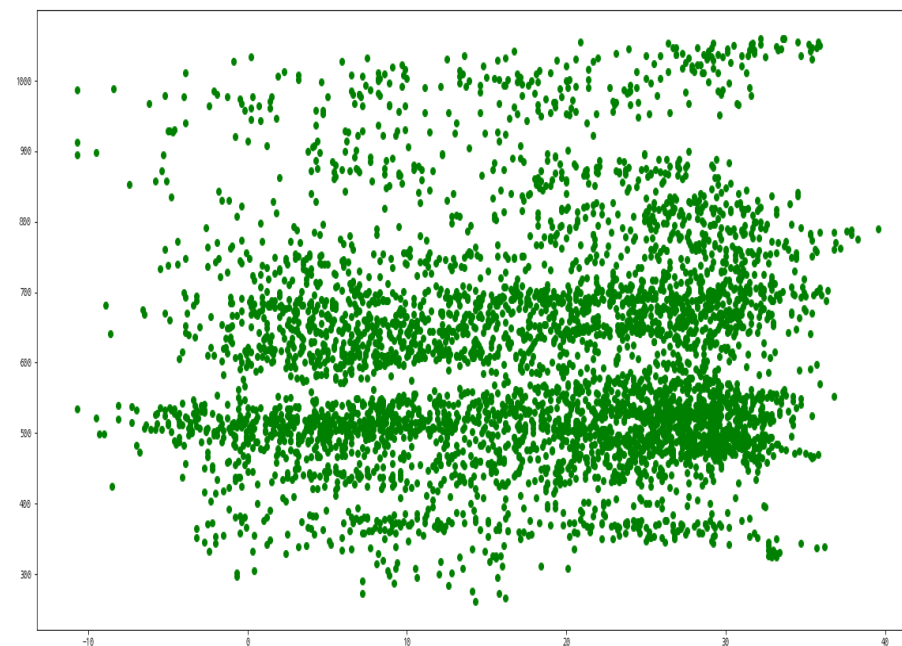


온도와 주가 간의 상관관계

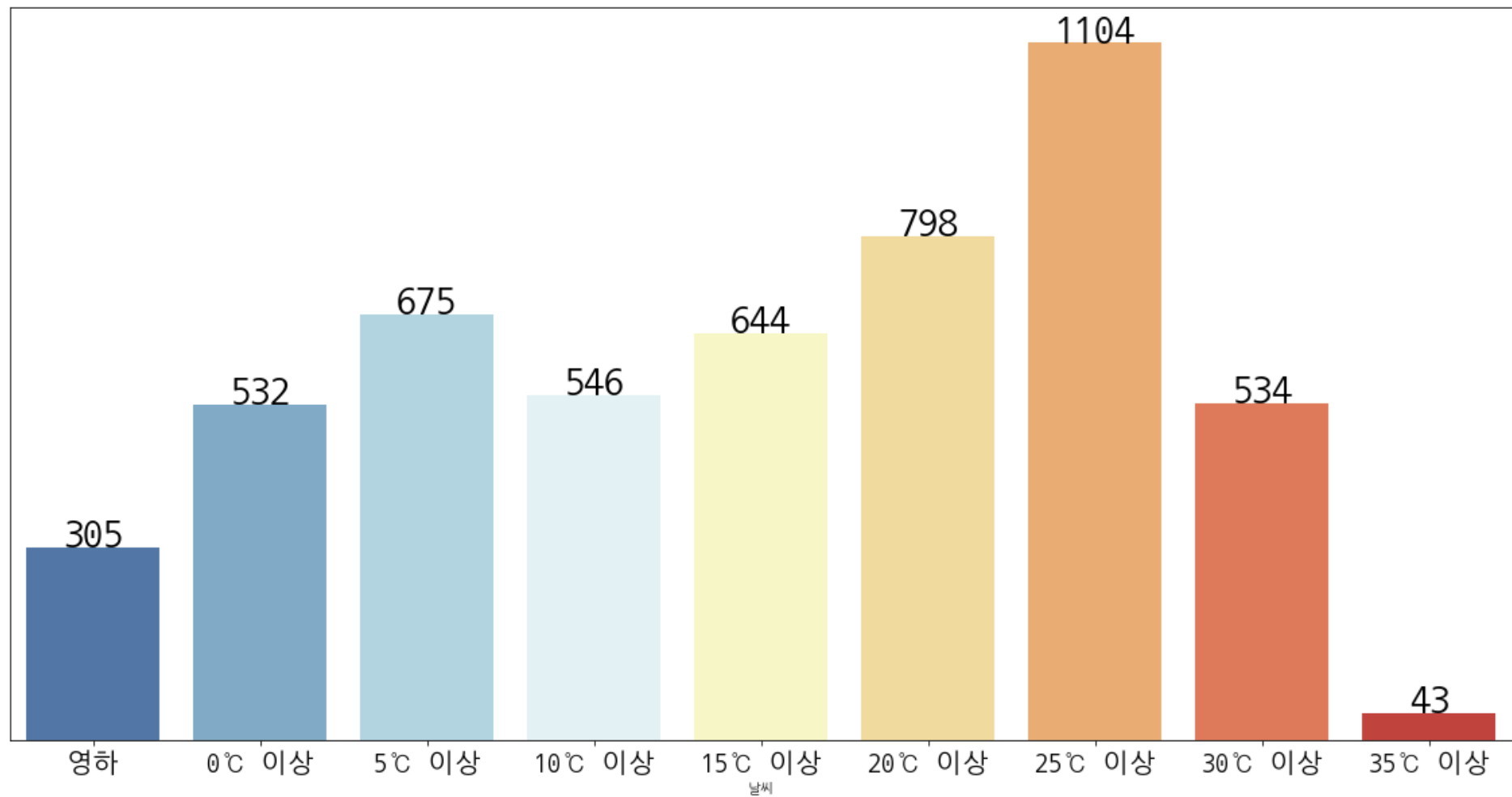
코스피



코스닥

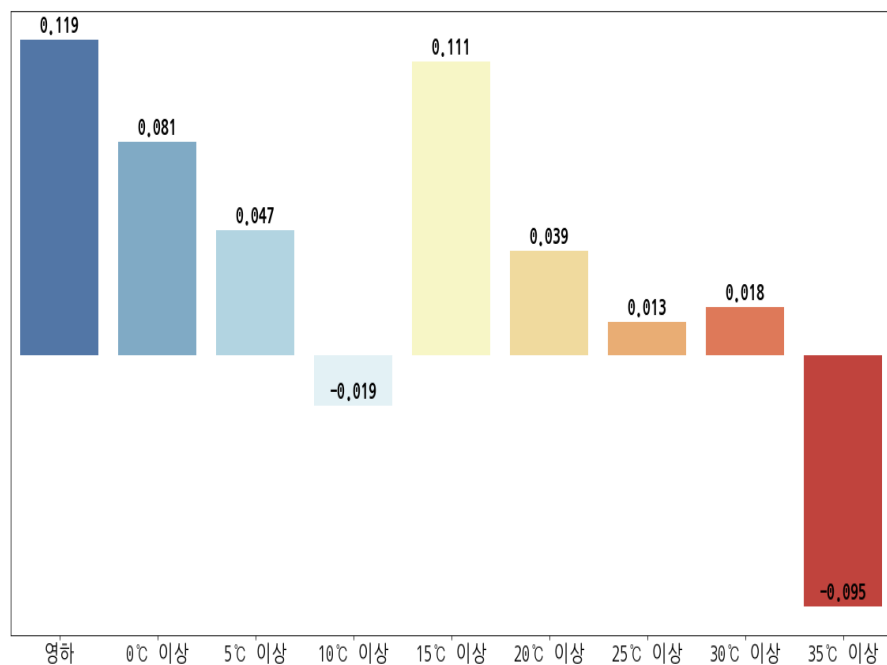


기온별 일 수

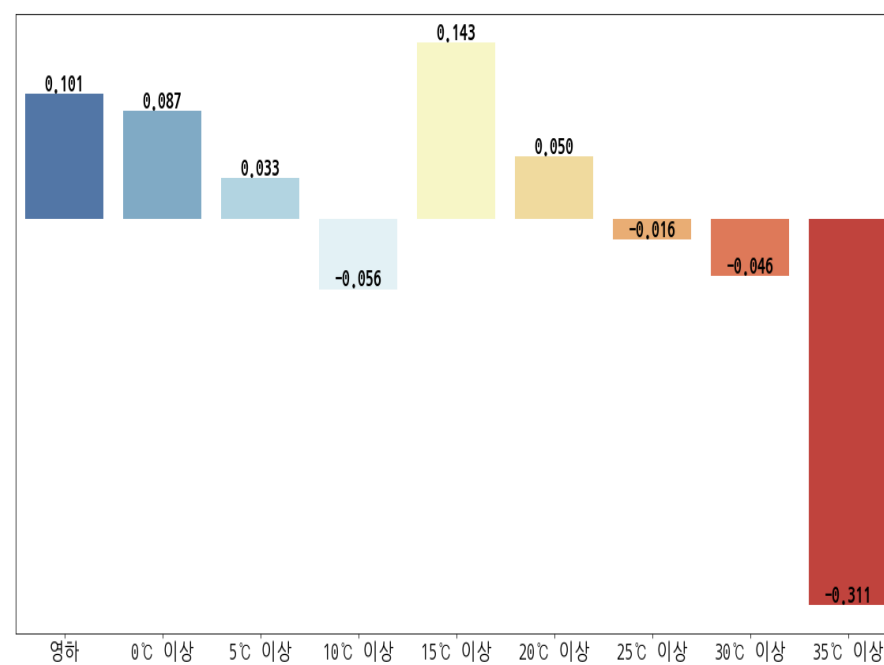


기온별 코스피, 코스닥 증감률

코스피

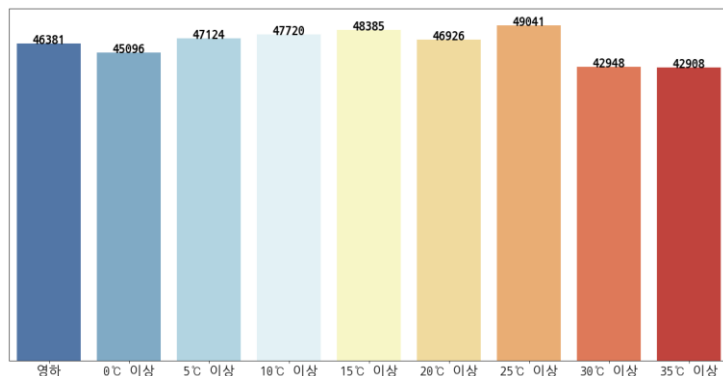


코스닥

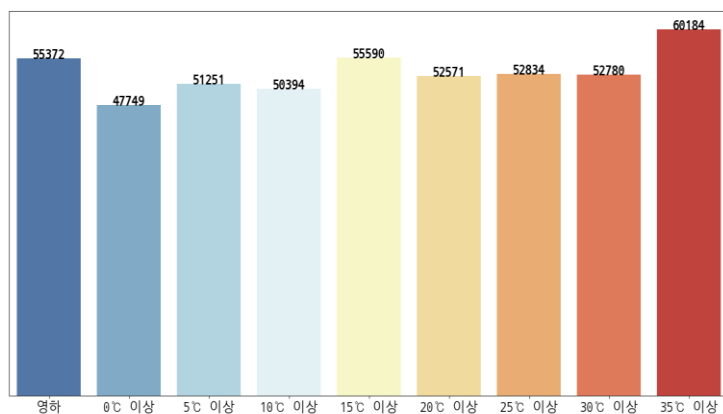


기온별 거래량과 거래대금

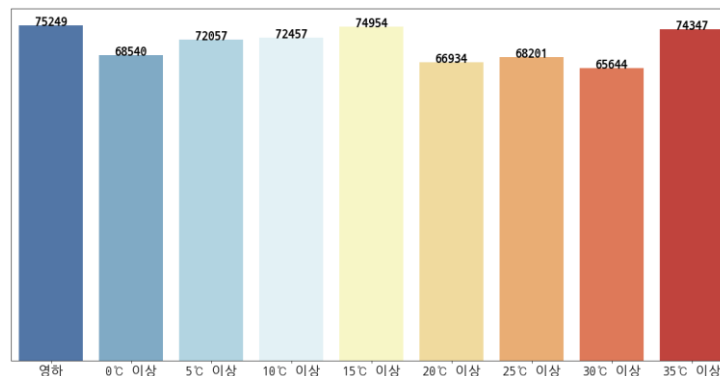
코스피 거래량



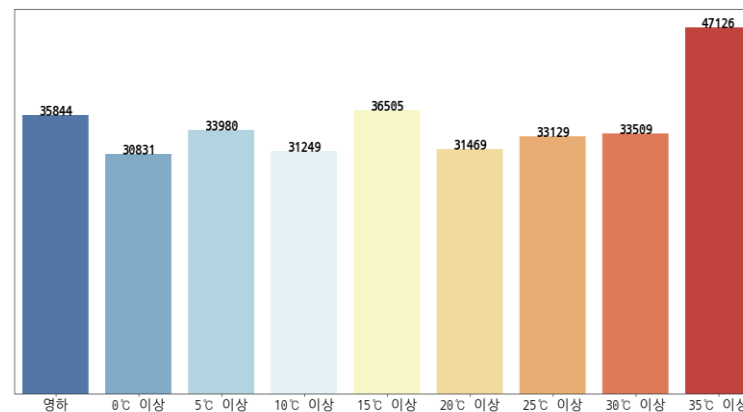
코스피 거래대금



코스닥 거래량



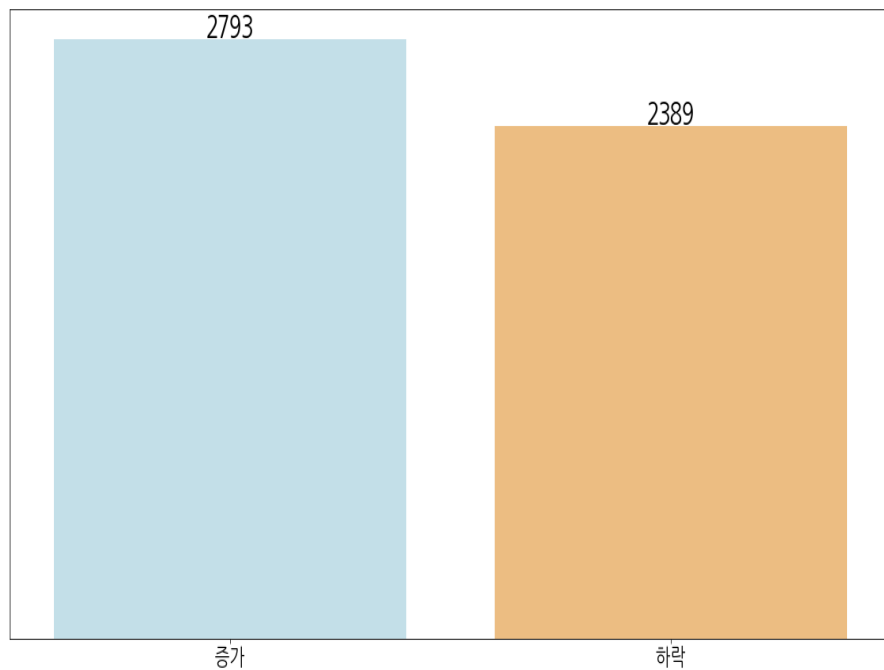
코스닥 거래대금



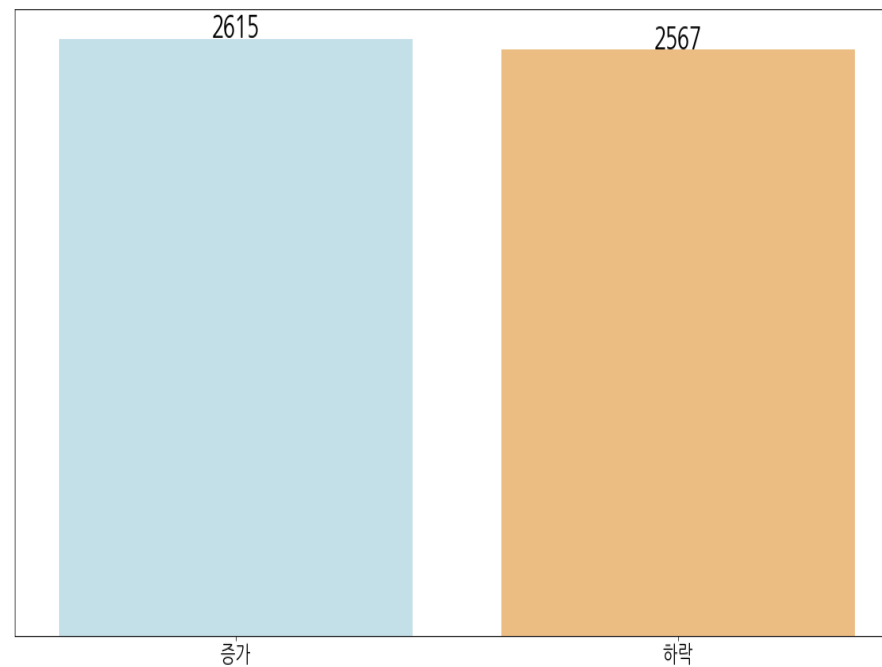
데이터 분석 — 최고 기온과 주가 분석

코스피, 코스닥의 상승, 하락 일 수

코스피



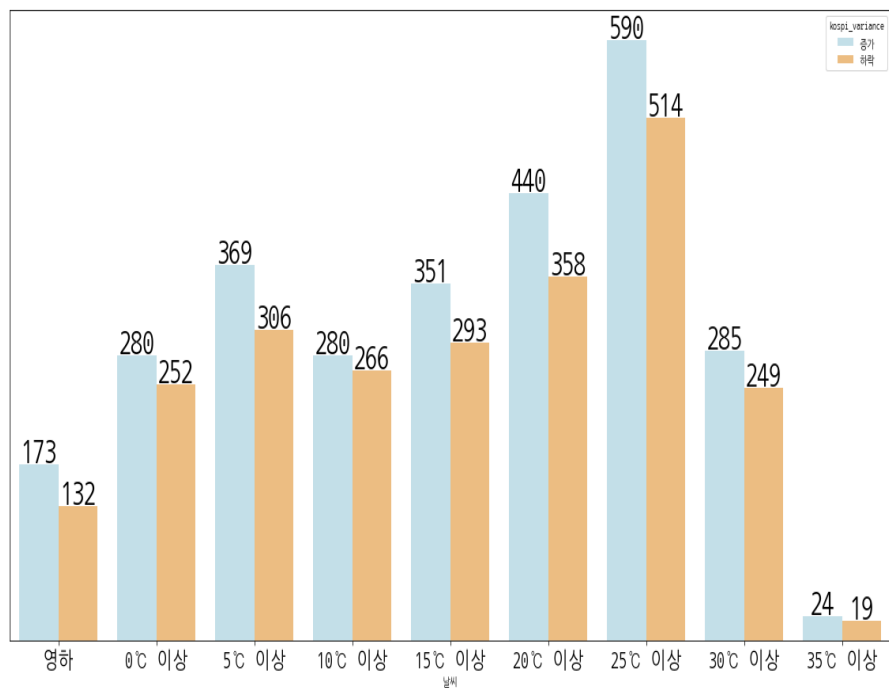
코스닥



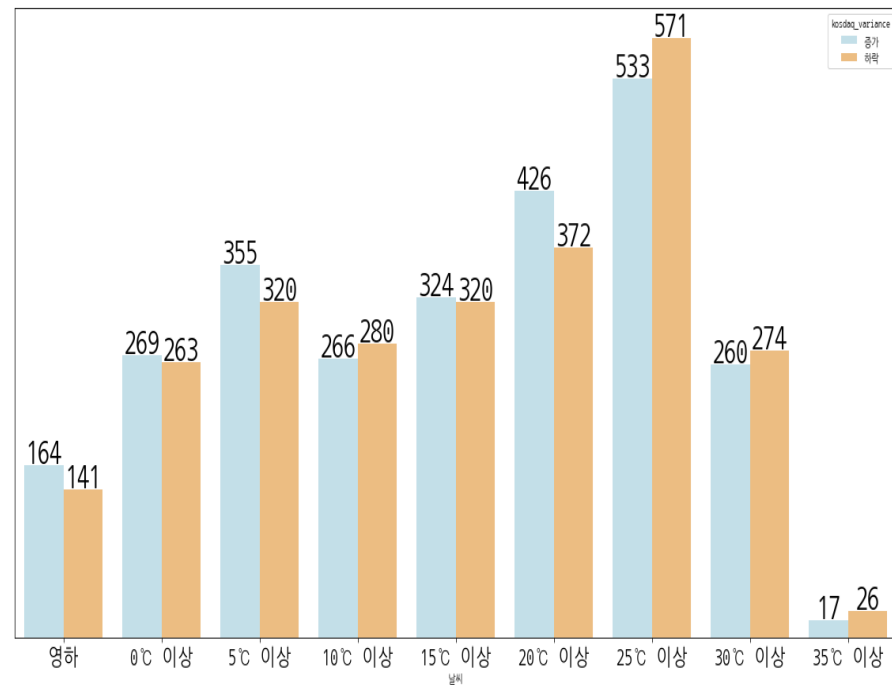
데이터 분석 — 최고 기온과 주가 분석

기온별 코스피, 코스닥 변동률 — 증감률이 0이상이면 증가, 0미만이면 하락

코스피



코스닥





Part 3

데이터 분석

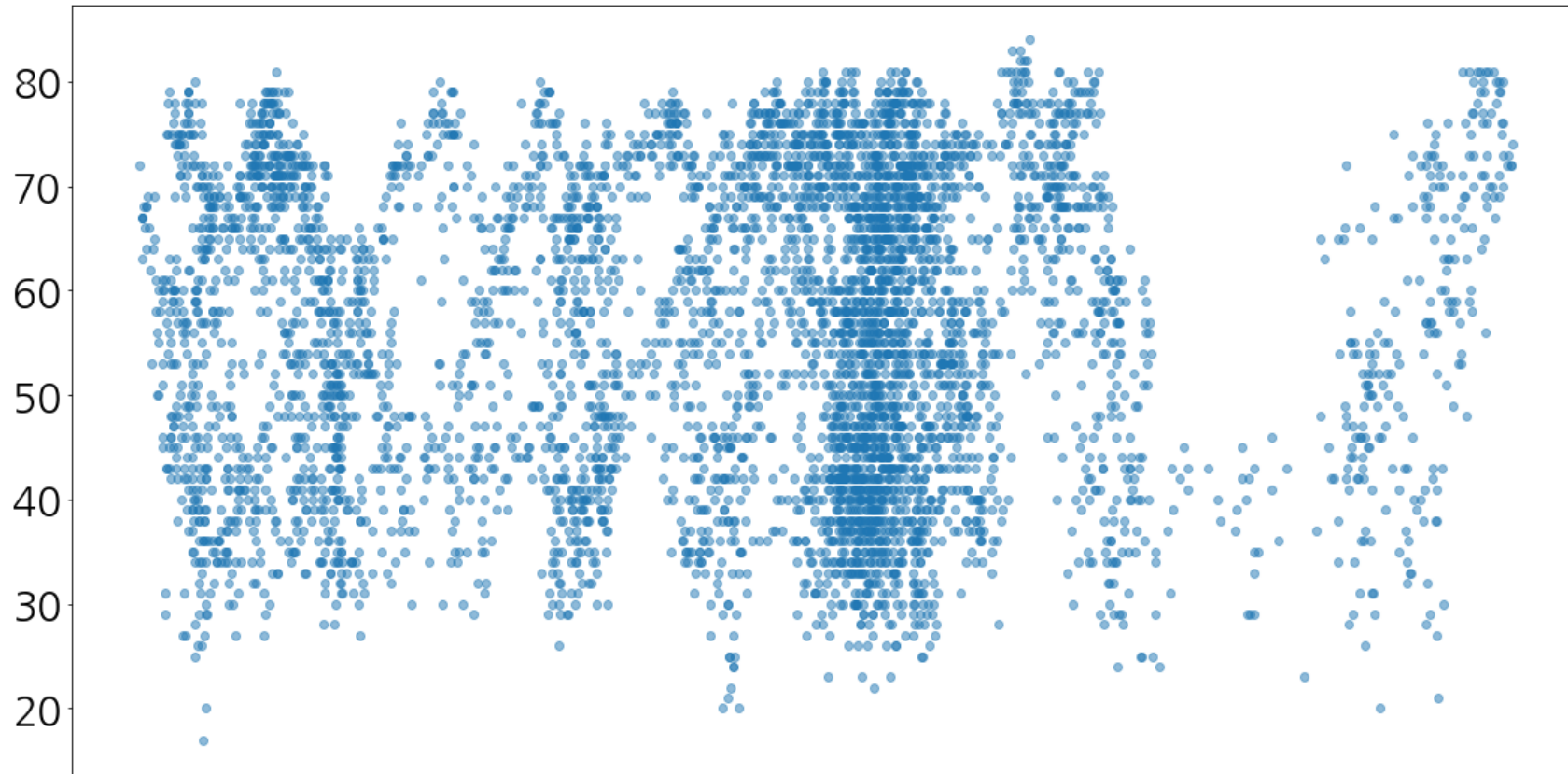
1. 기온 (최고기온)

2. 불쾌지수

3. 전운량

4. 태풍

불래지수와 KOSPI 지수 간의 상관관계



불쾌지수 구간 생성

$$\text{불쾌지수} = 1.8 \times \text{기온} - 0.55 \times (1 - \text{습도}) \times (1.8 \times \text{기온} - 26) + 32$$

불쾌지수 구분 기준

불쾌지수 68 미만 : 쾌적함

68 <= 불쾌지수 < 75 : 약간 불쾌함

75 <= 불쾌지수 < 80 : 불쾌함

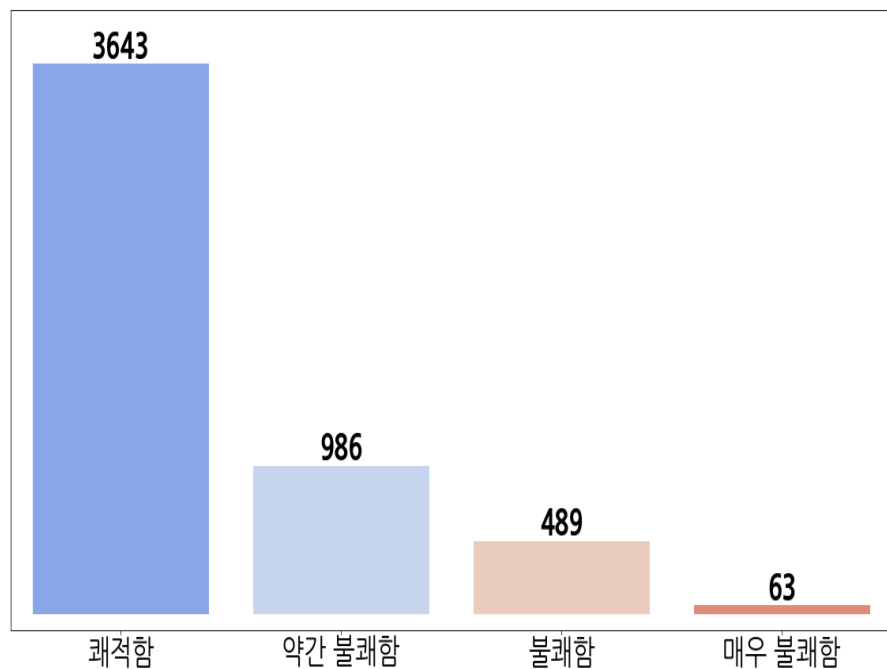
80 이상 : 매우 불쾌함

>>

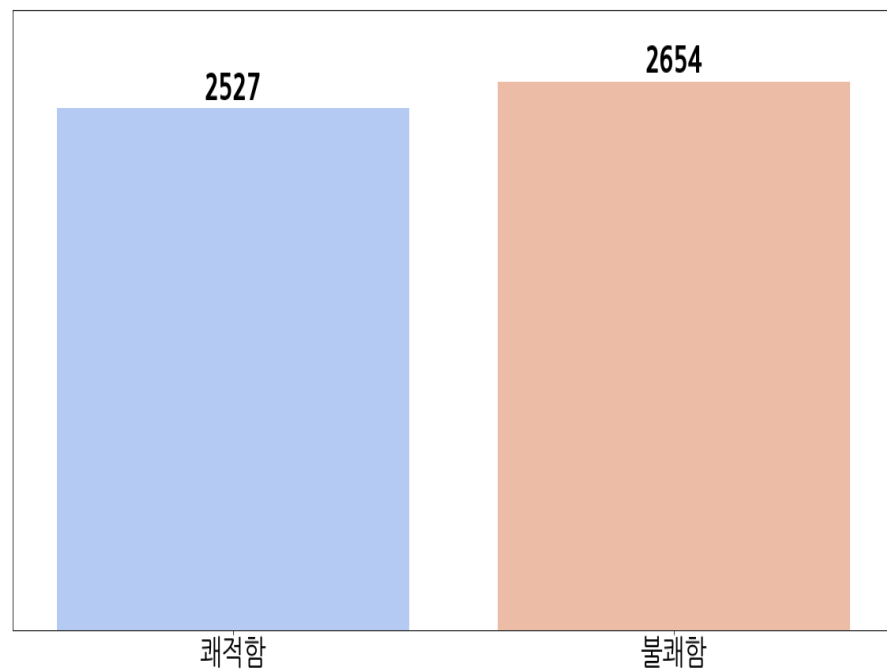
```
section = nomarly.copy()
section = section.assign(discomfort_section="0")
for i in range(len(section)):
    if section['discomfort_index'][i] < 68 :
        section['discomfort_section'][i] = '쾌적함'
    elif section['discomfort_index'][i] < 75 :
        section['discomfort_section'][i] = '약간 불쾌함'
    elif section['discomfort_index'][i] < 80 :
        section['discomfort_section'][i] = '불쾌함'
    else : section['discomfort_section'][i] = '매우 불쾌함'
```

불쾌지수 구간별 일 수

4구간으로 나눈 불쾌지수



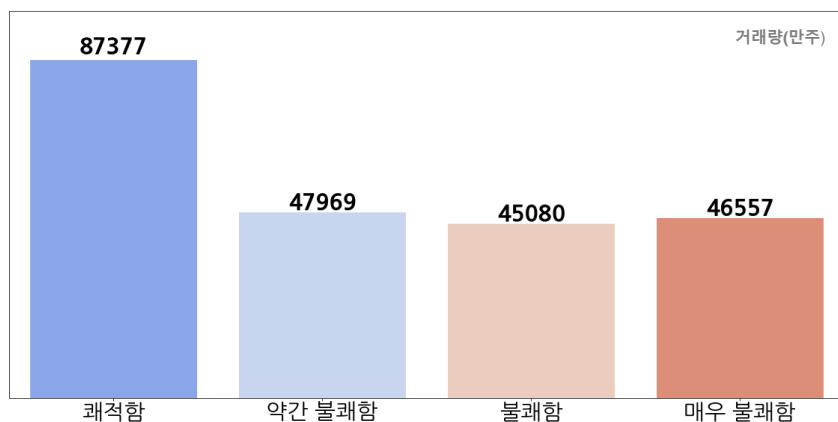
2구간으로 나눈 불쾌지수



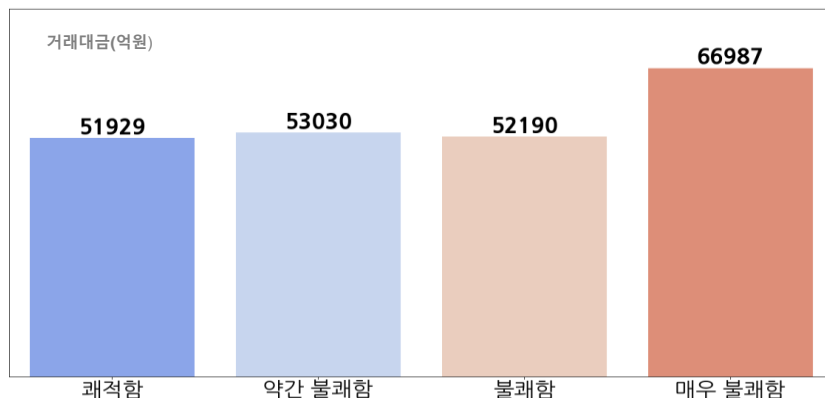
불쾌지수 구간별 거래량과 거래대금

코스피

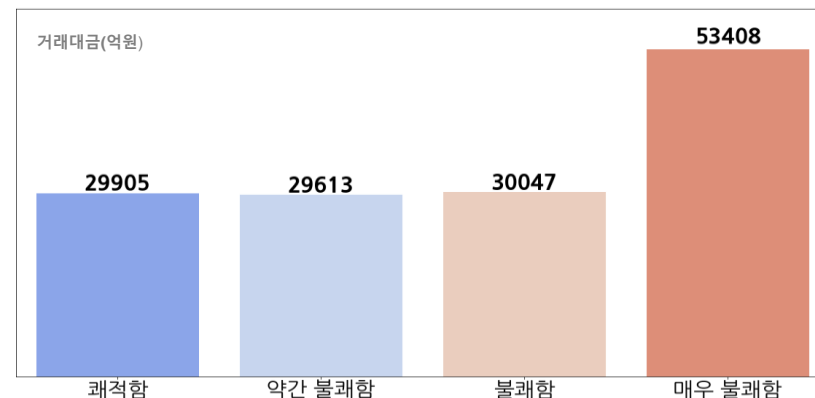
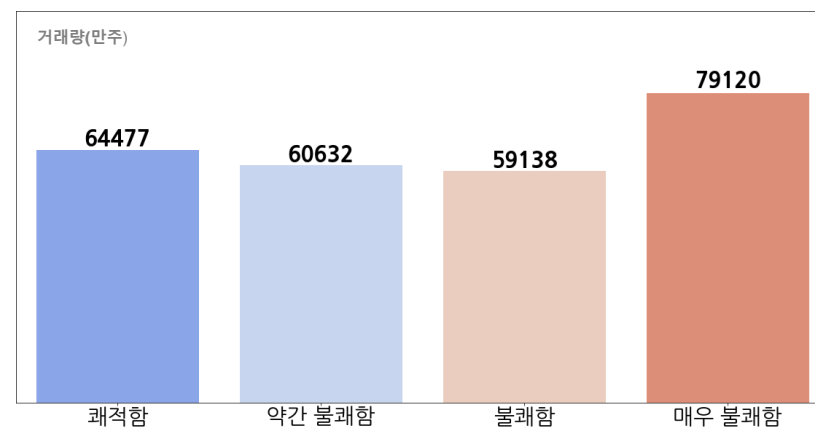
거래량



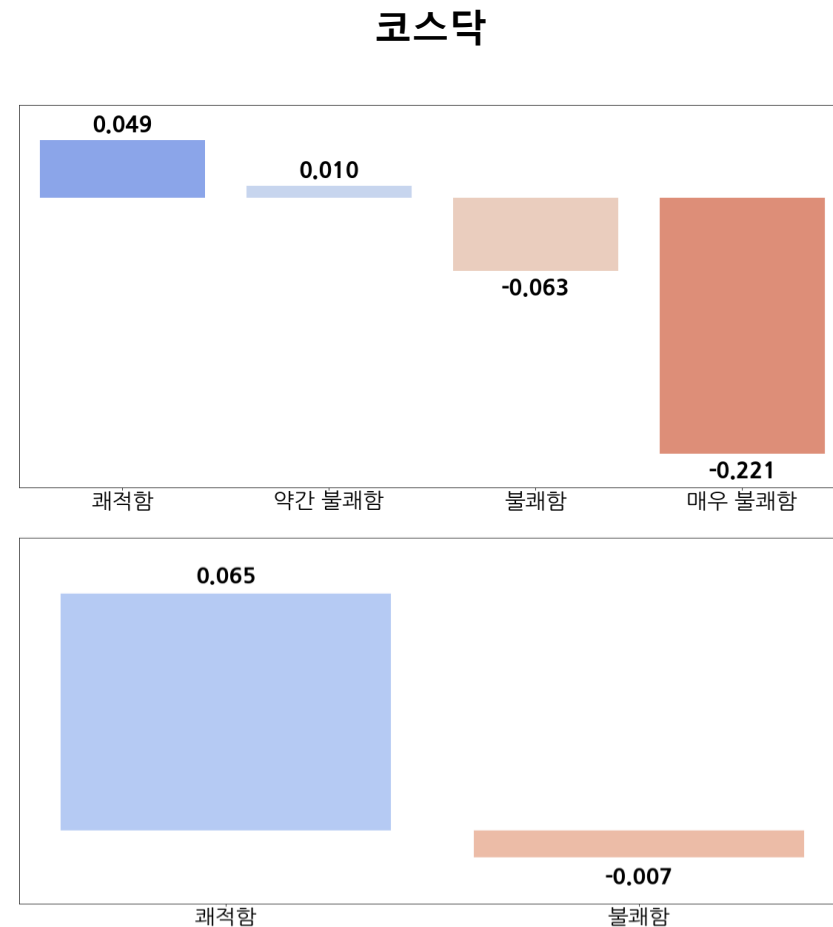
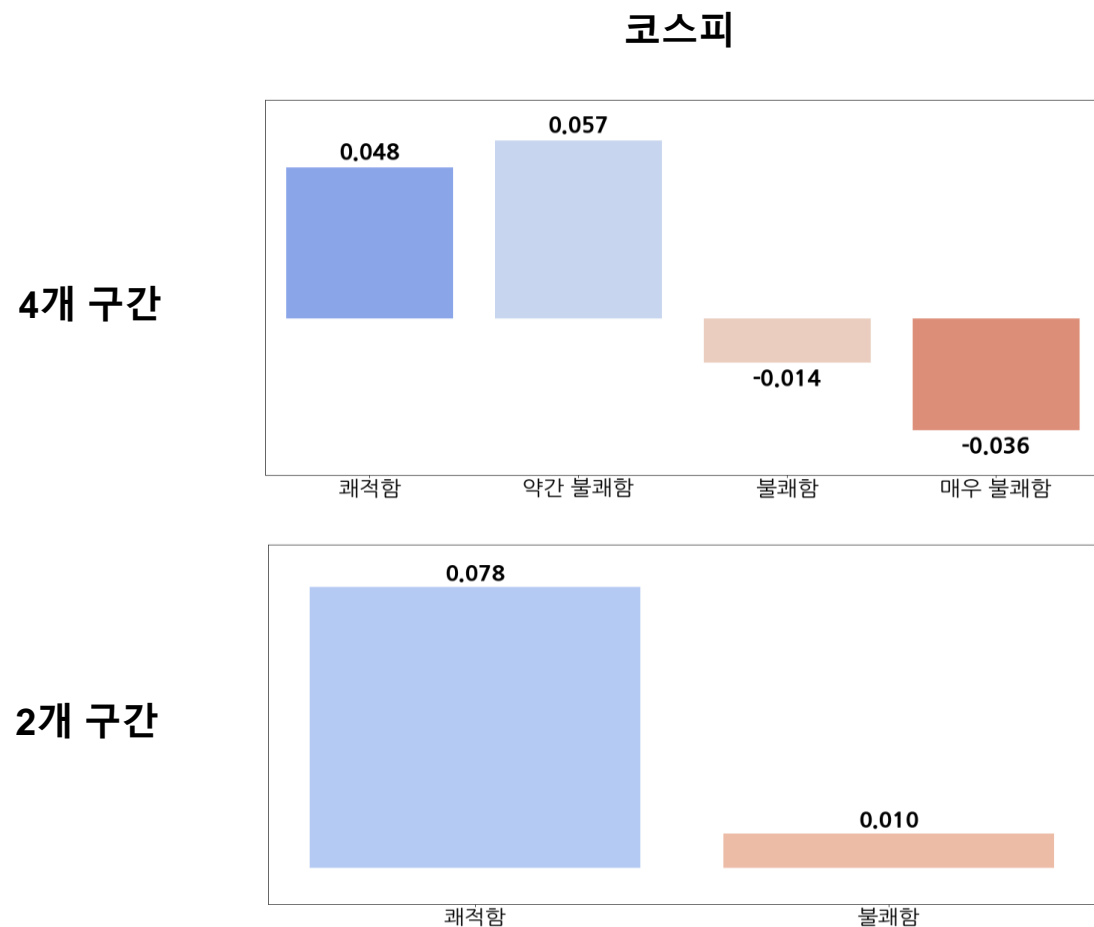
거래대금



코스닥



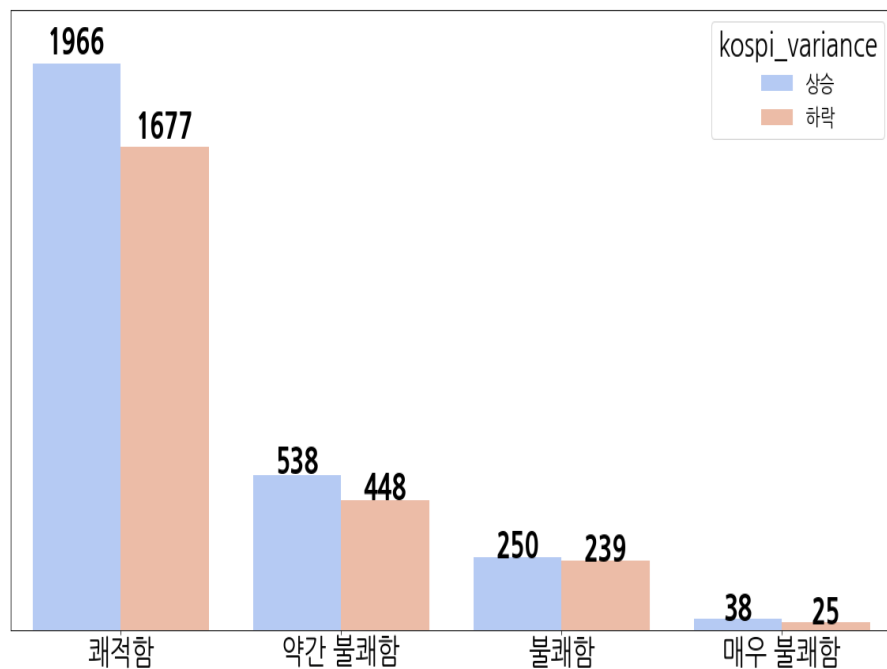
불쾌지수 구간별 증감률



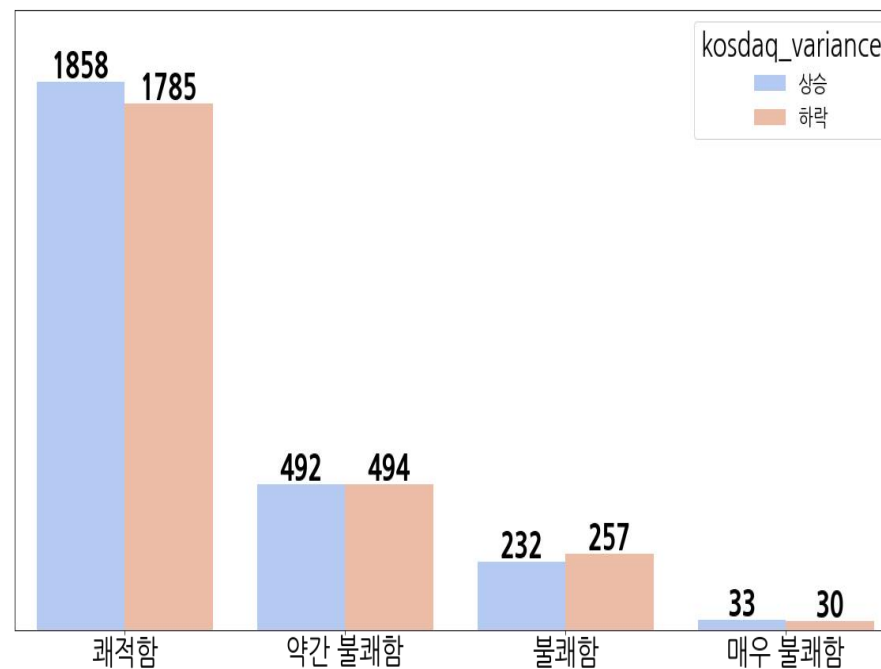
데이터 분석 — 불쾌지수와 주가 분석

불쾌지수 구간별 증감률 상승, 하락 일 수

코스피



코스닥





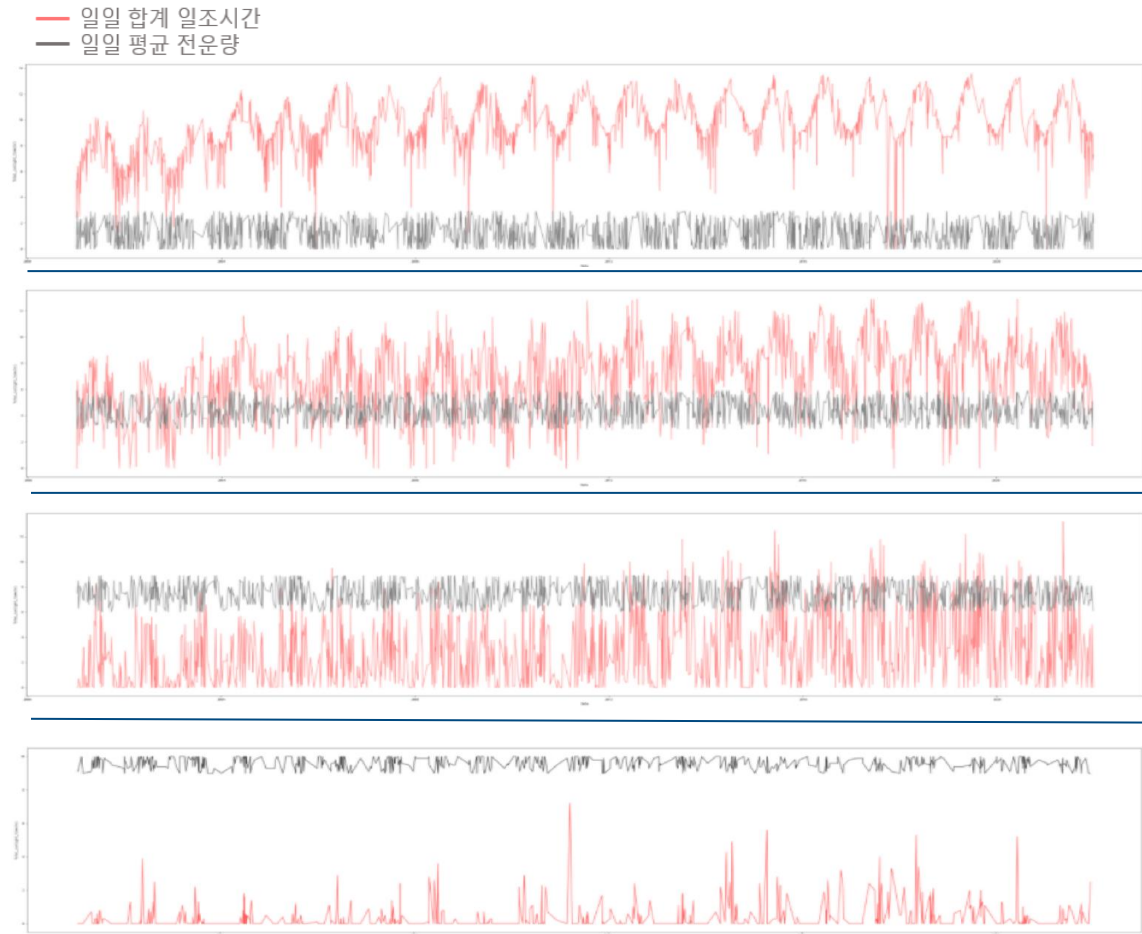
Part 3

데이터 분석

1. 기온 (최고기온)
2. 불쾌지수
- 3. 전운량**
4. 태풍

Part 3 데이터 분석 — 전운량과 주가 분석

맑음 / 흐림 구간 설정



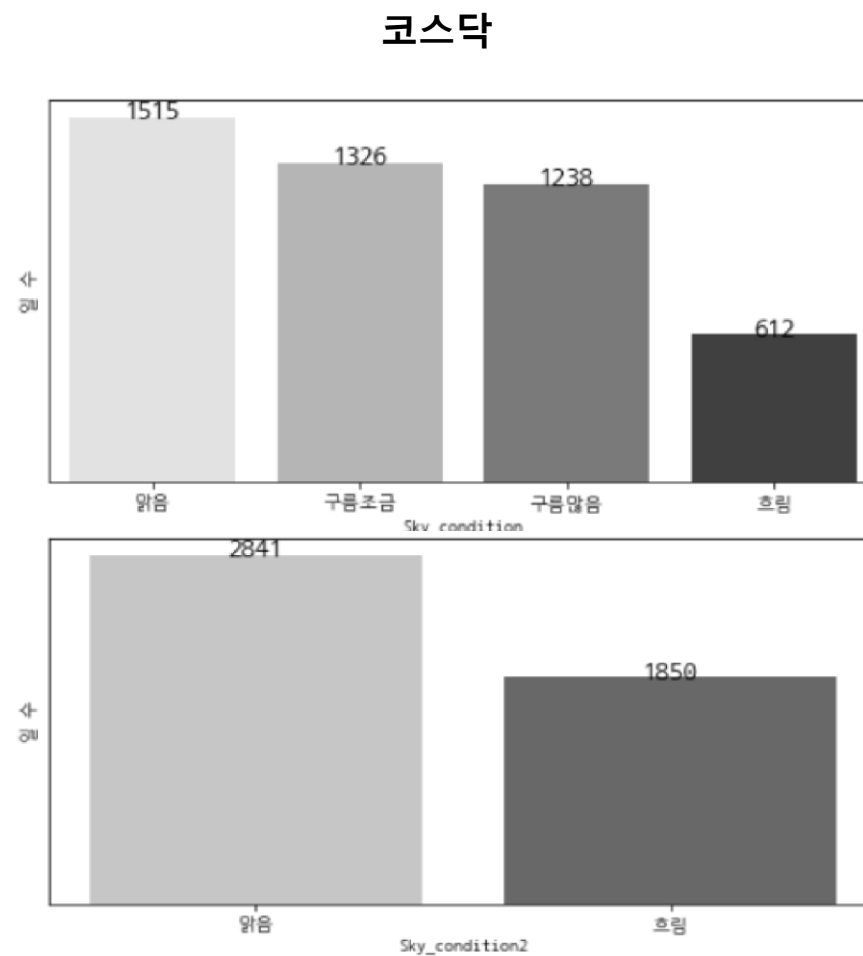
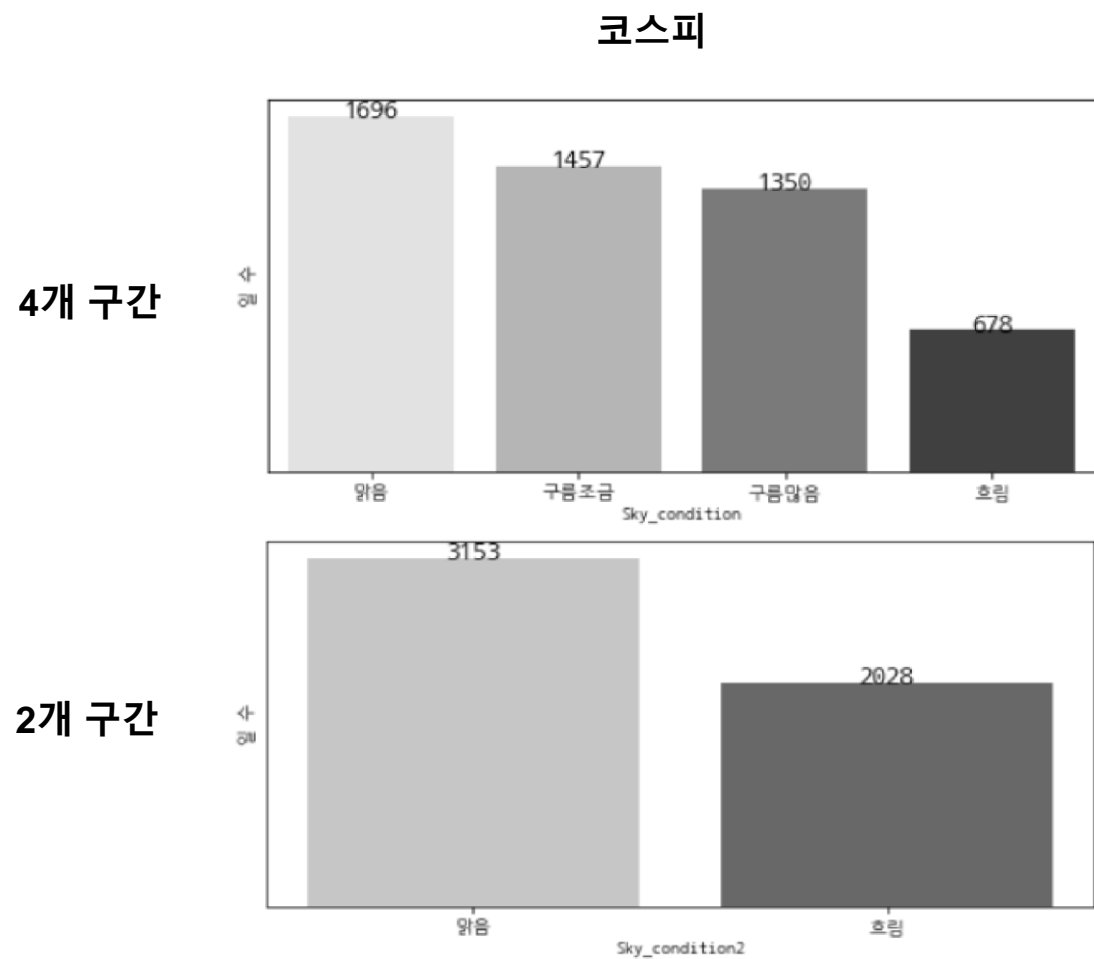
구간 기준 : 맑음(0~2) / 구름조금(3~5) / 구름많음(6~8) / 흐림(9~10)

↓

```
def func(x) :  
    if x >= 9 :  
        return "흐림"  
    elif x >= 6:  
        return "구름많음"  
    elif x >= 3:  
        return "구름조금"  
    else : return "맑음"  
cloudiness["Sky_condition"] = cloudiness["Average_total_cloudiness(1/10)"].apply(lambda x : func(x))
```

Part 3 데이터 분석 — 전운량과 주가 분석

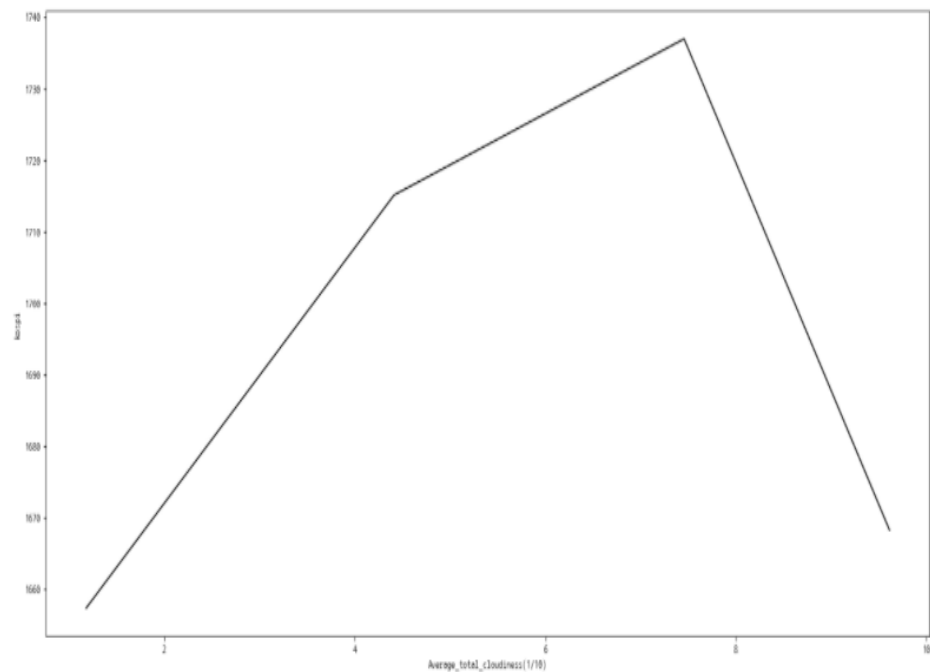
전운량 구간별 일 수



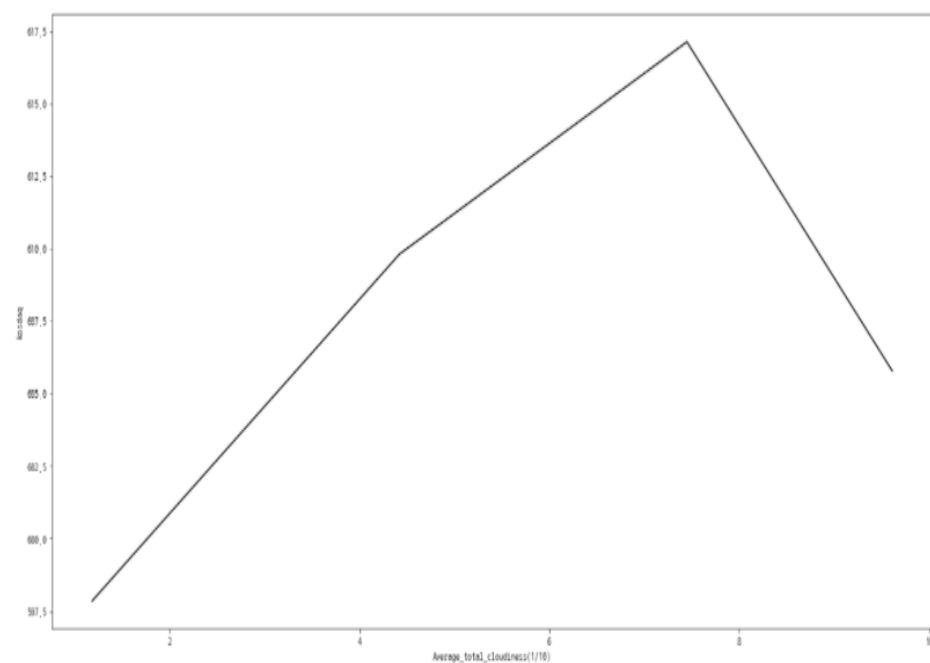
Part 3 데이터 분석 — 전운량과 주가 분석

전운량과 평균 코스피, 코스닥 지수

코스피



코스닥

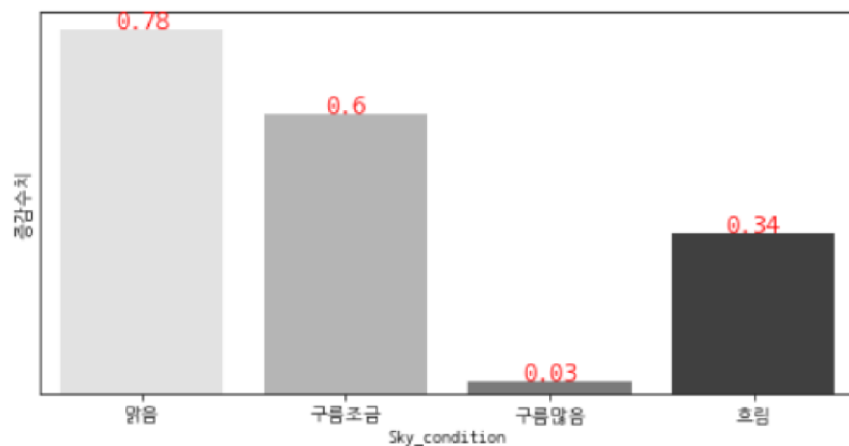


데이터 분석 — 전운량과 주가 분석

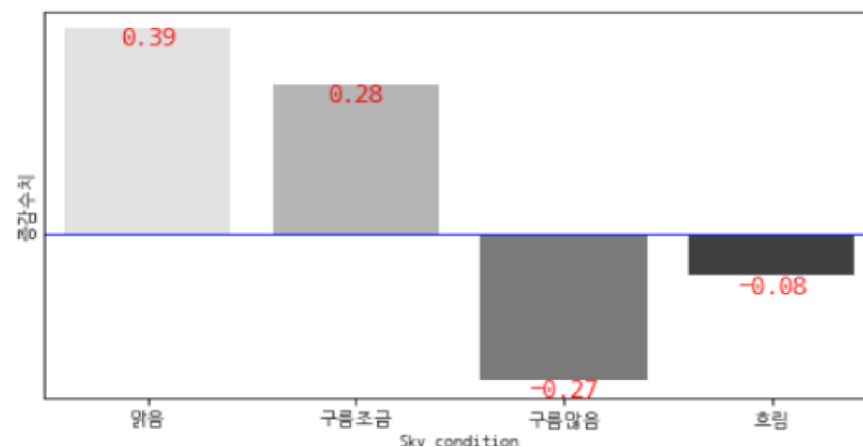
전운량 구간별 평균 증감수치

4개 구간

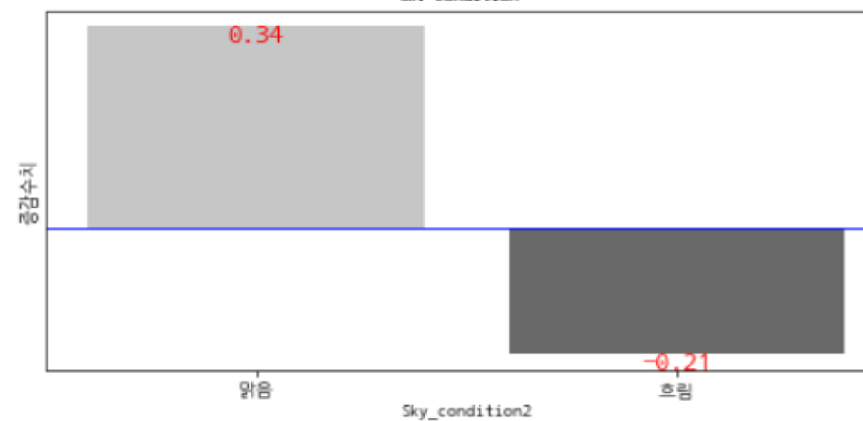
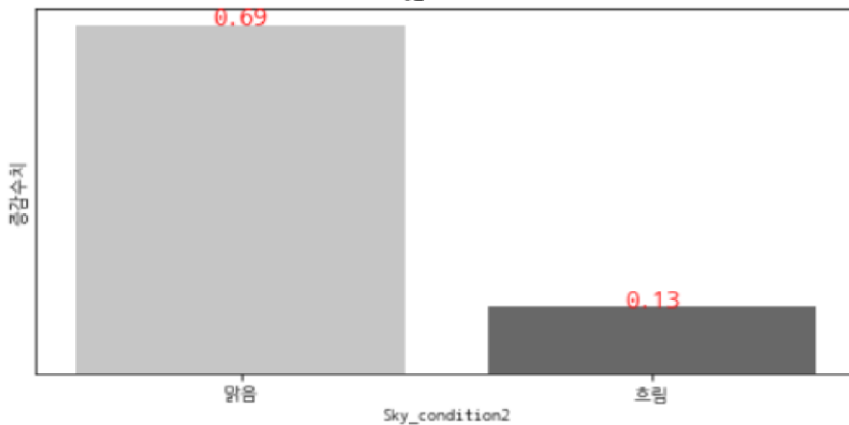
코스피



코스닥

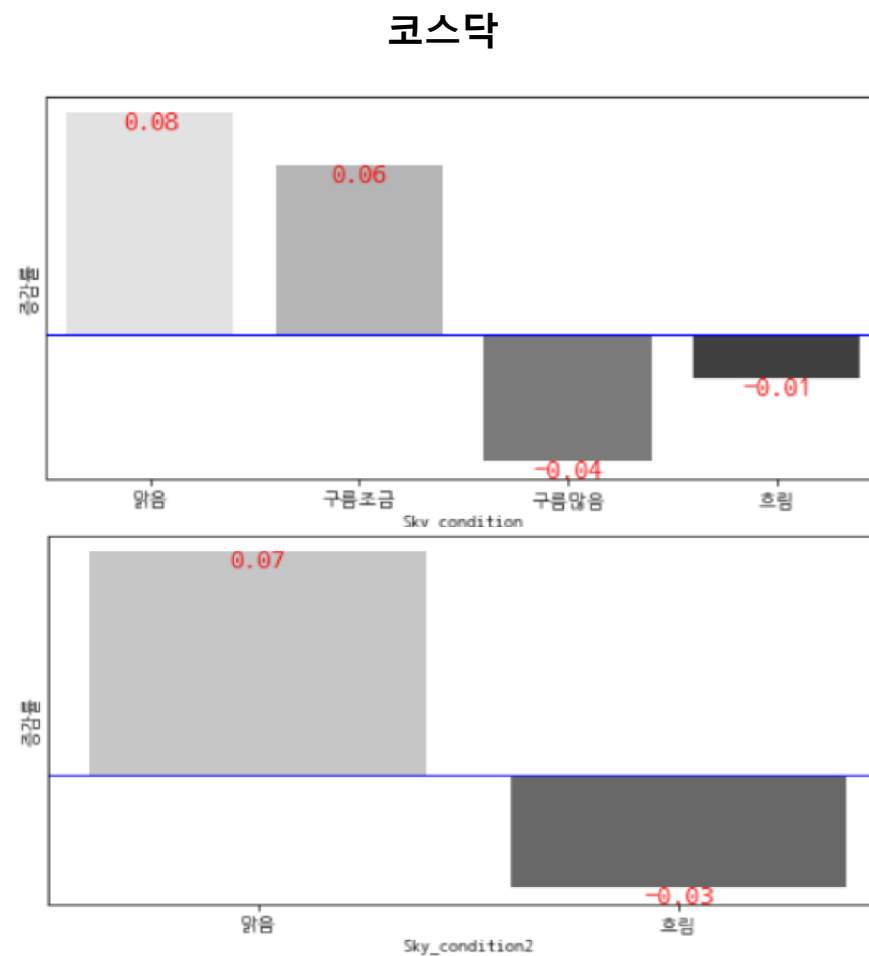
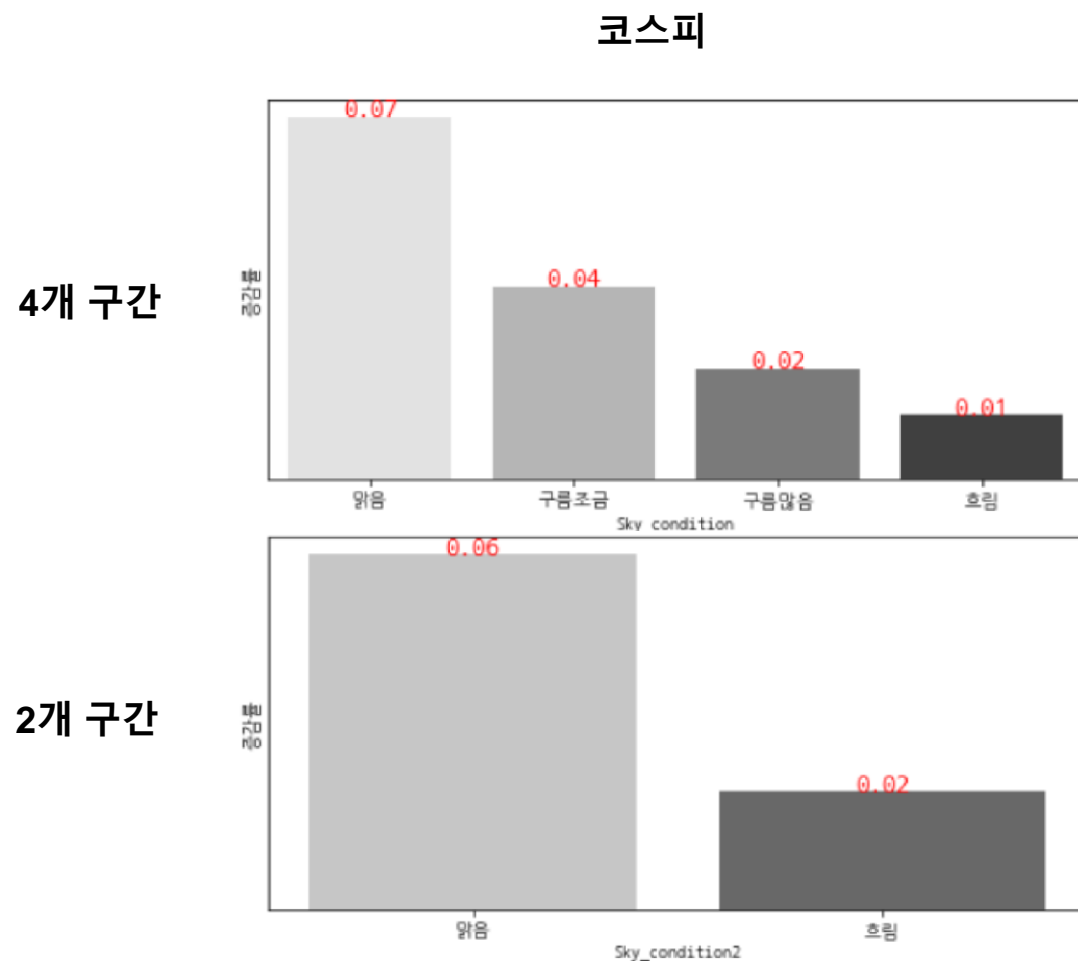


2개 구간



데이터 분석 — 전운량과 주가 분석

전운량 구간별 평균 증감률

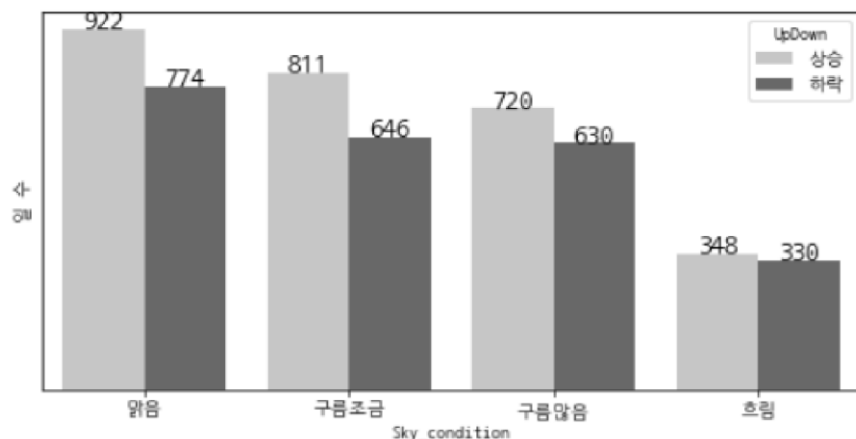


데이터 분석 — 전운량과 주가 분석

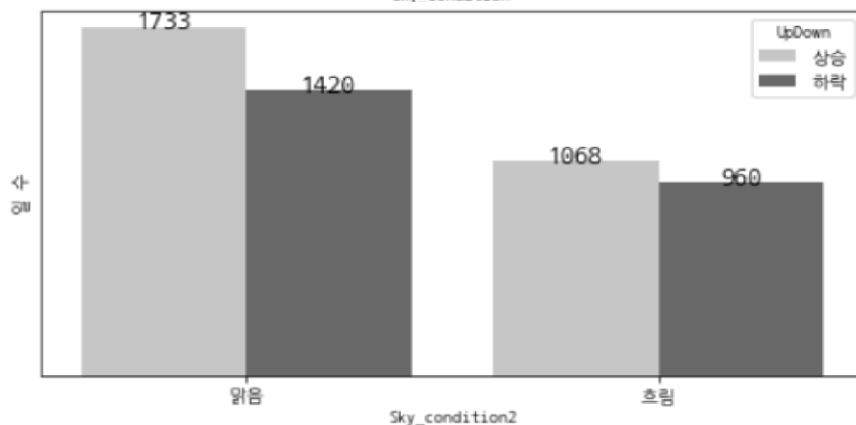
전운량 구간별 상승, 하락 일 수

코스피 상승, 하락 일 수

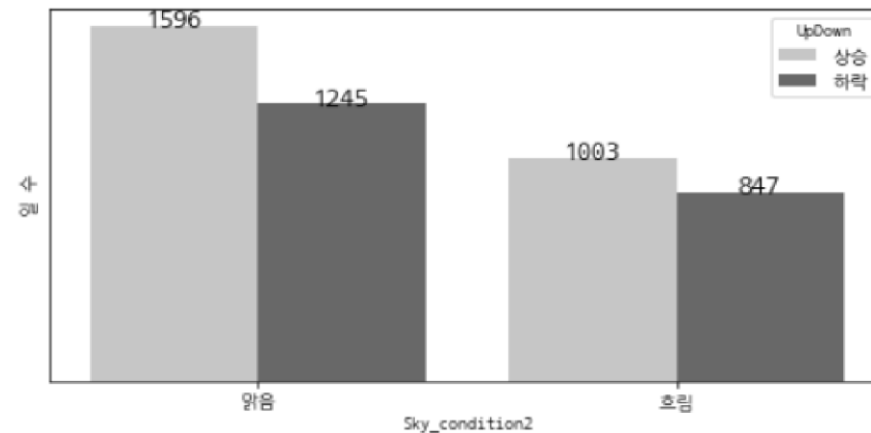
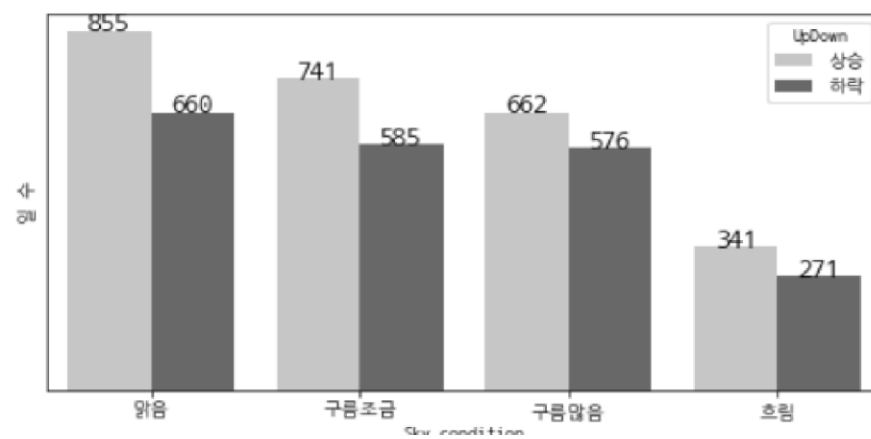
4개 구간



2개 구간

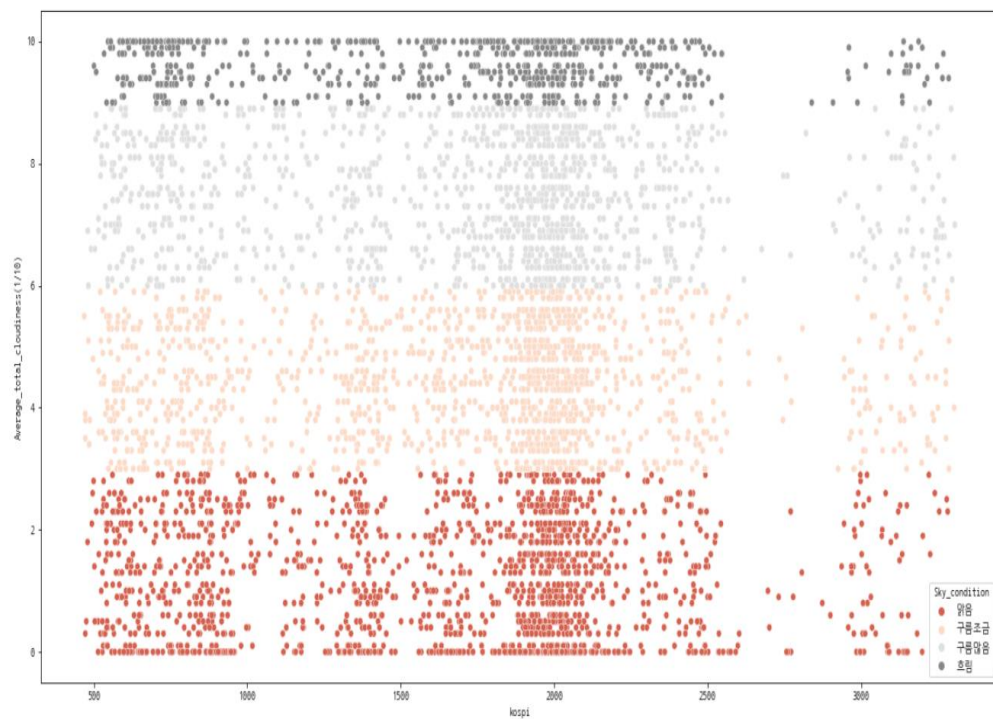


코스닥 상승, 하락 일 수

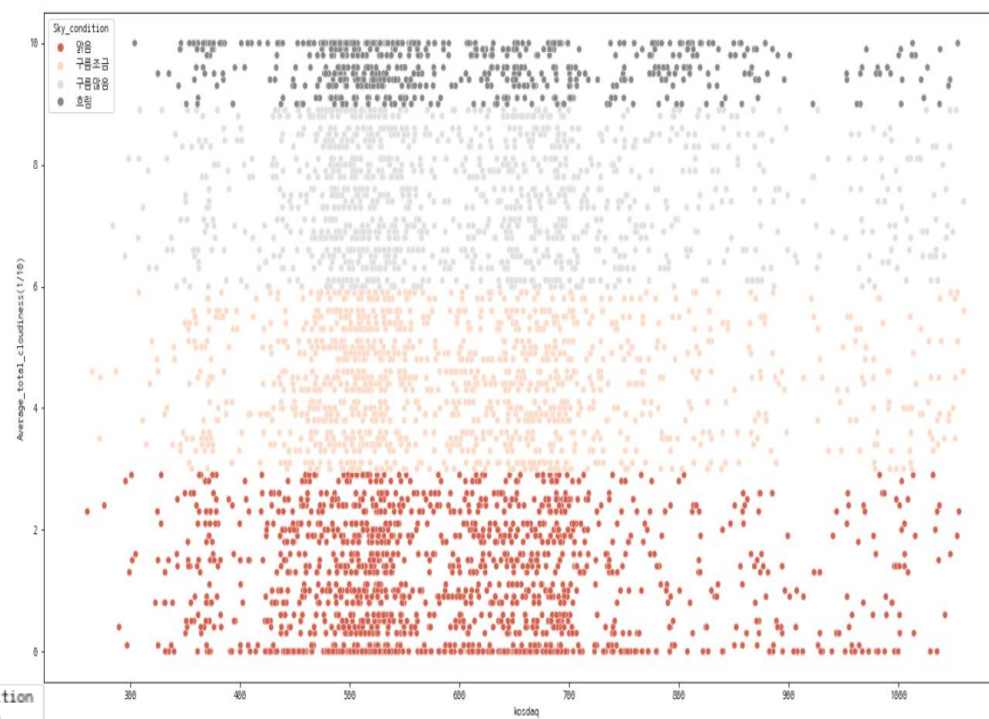


구간별 전운량과 주가지수 분포도

코스피



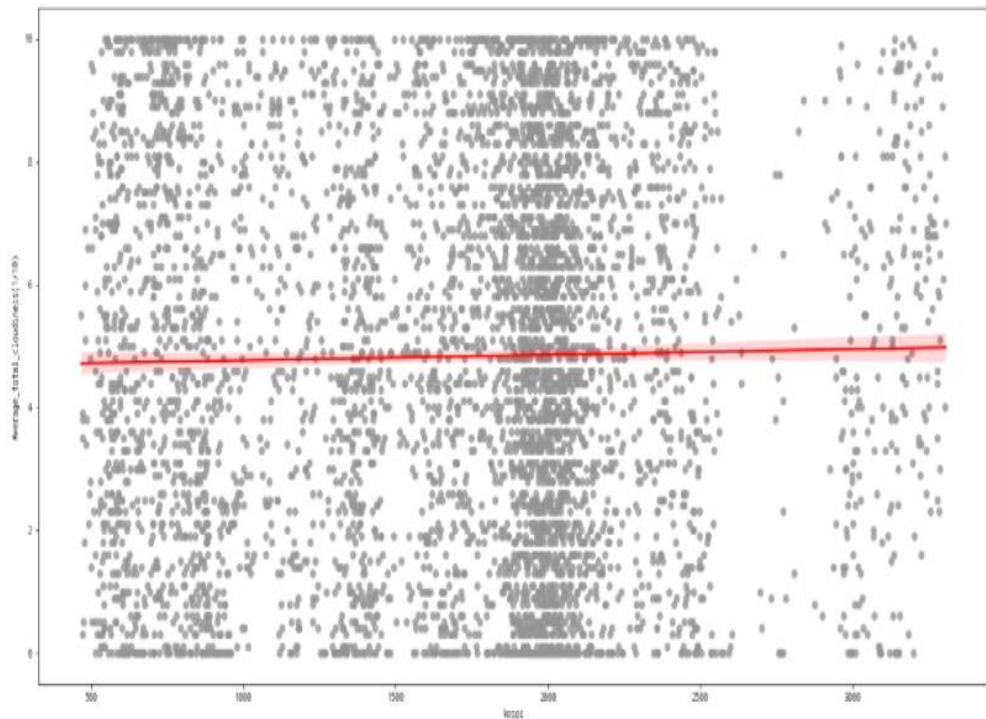
코스닥



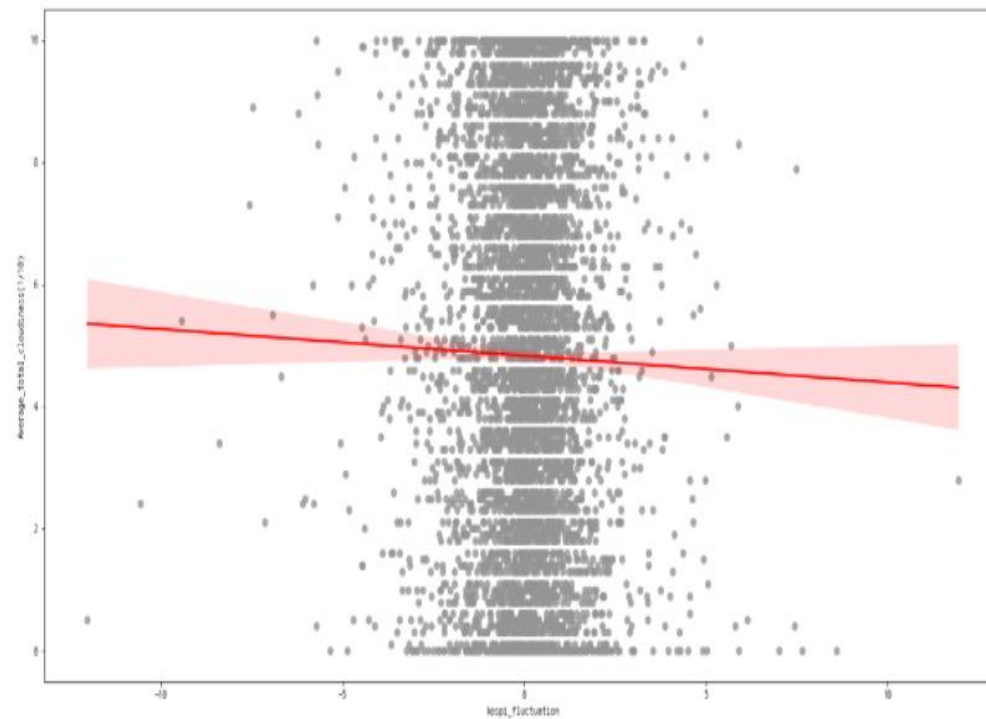
데이터 분석 — 전운량과 주가 분석

전운량과 주가지수, 증감률과의 관계 분포도

전운량과 코스피 지수

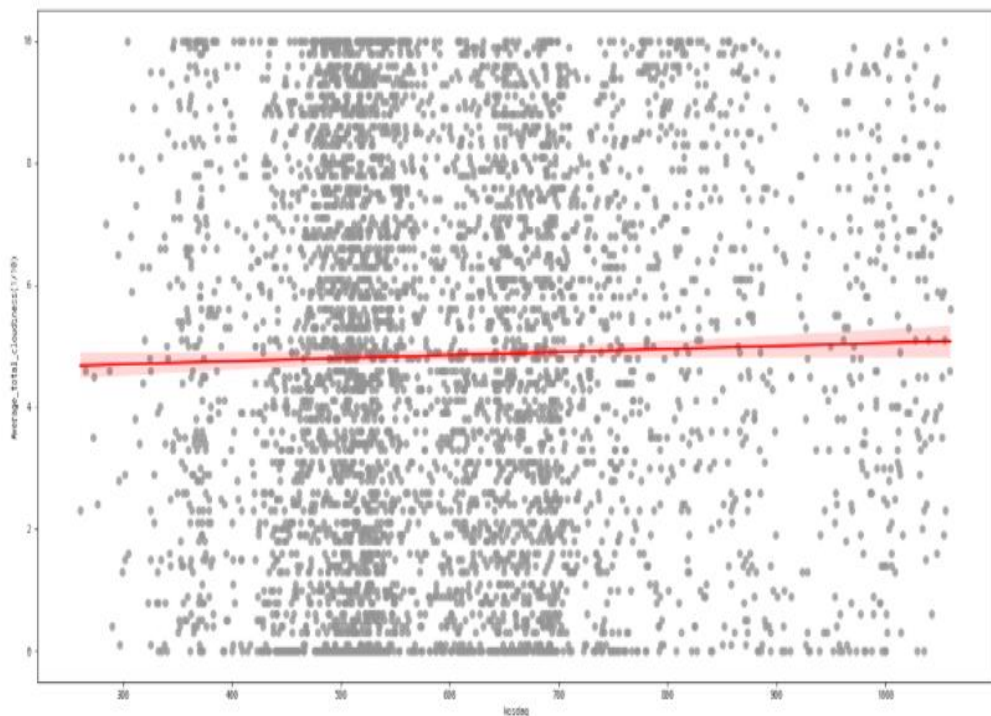


전운량과 코스피 증감률

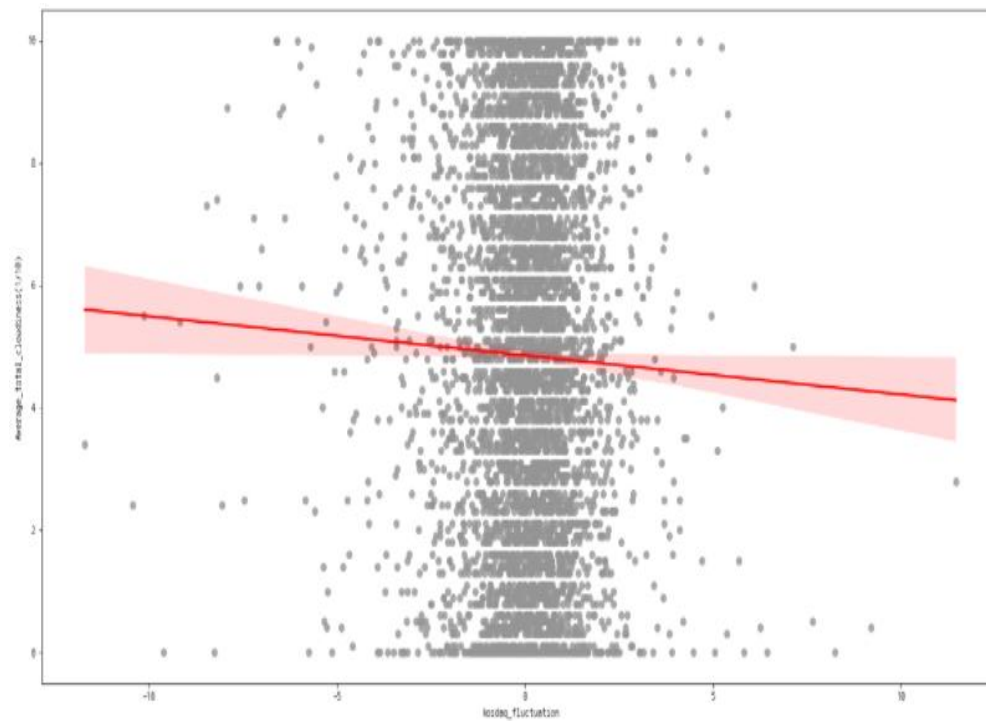


전운량과 주가지수, 증감률과의 관계 분포도

전운량과 코스닥 지수



전운량과 코스닥 증감률





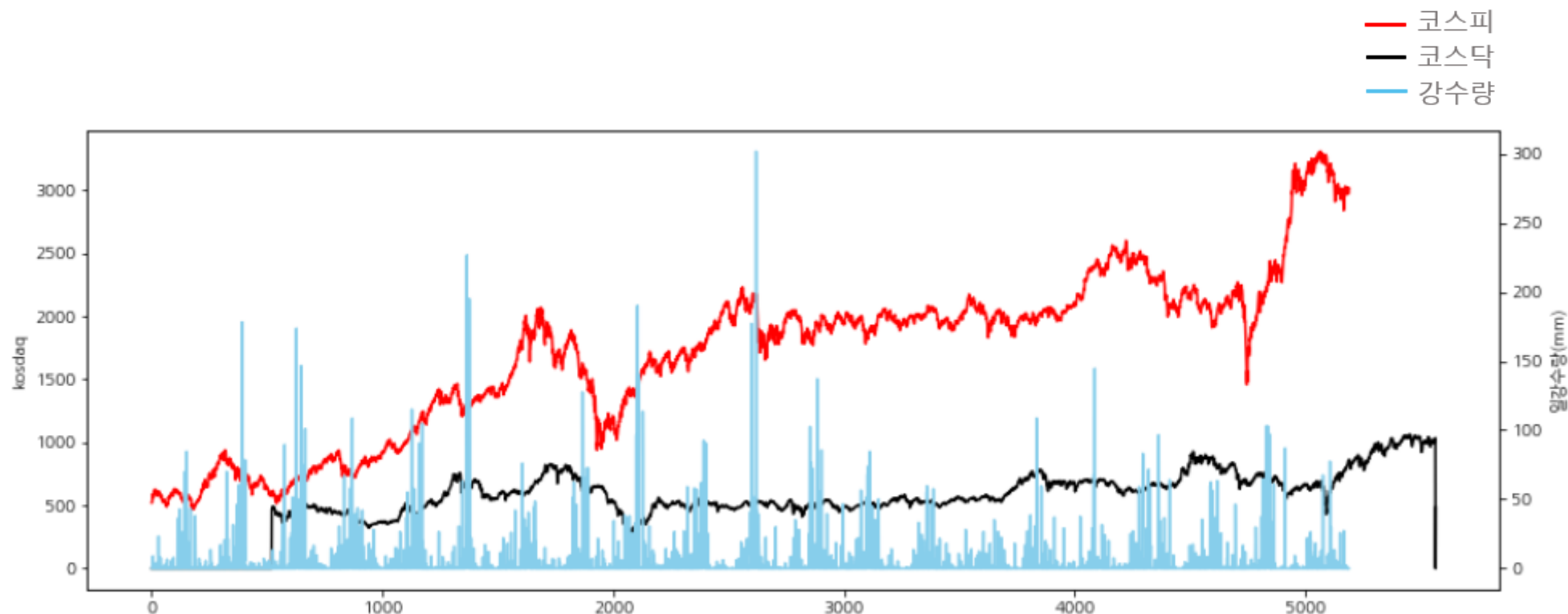
Part 3

데이터 분석

1. 기온 (최고기온)
2. 불쾌지수
3. 전운량
4.

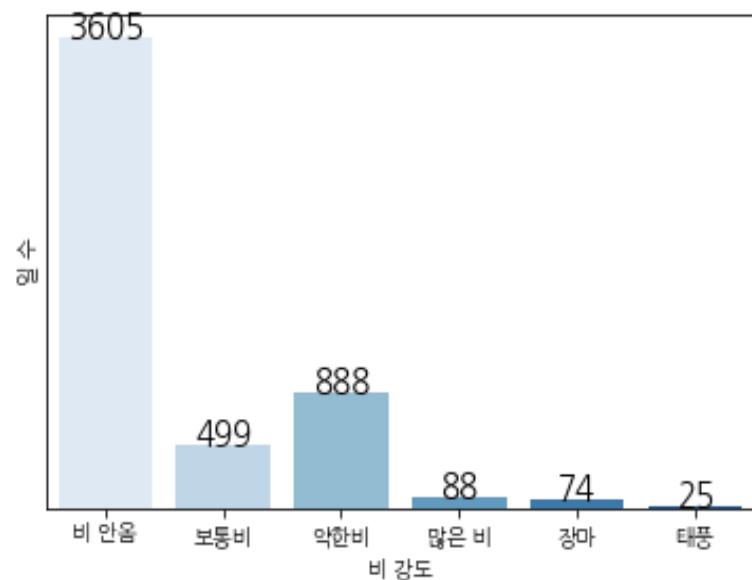
태풍

강수량에 따른 주가지수

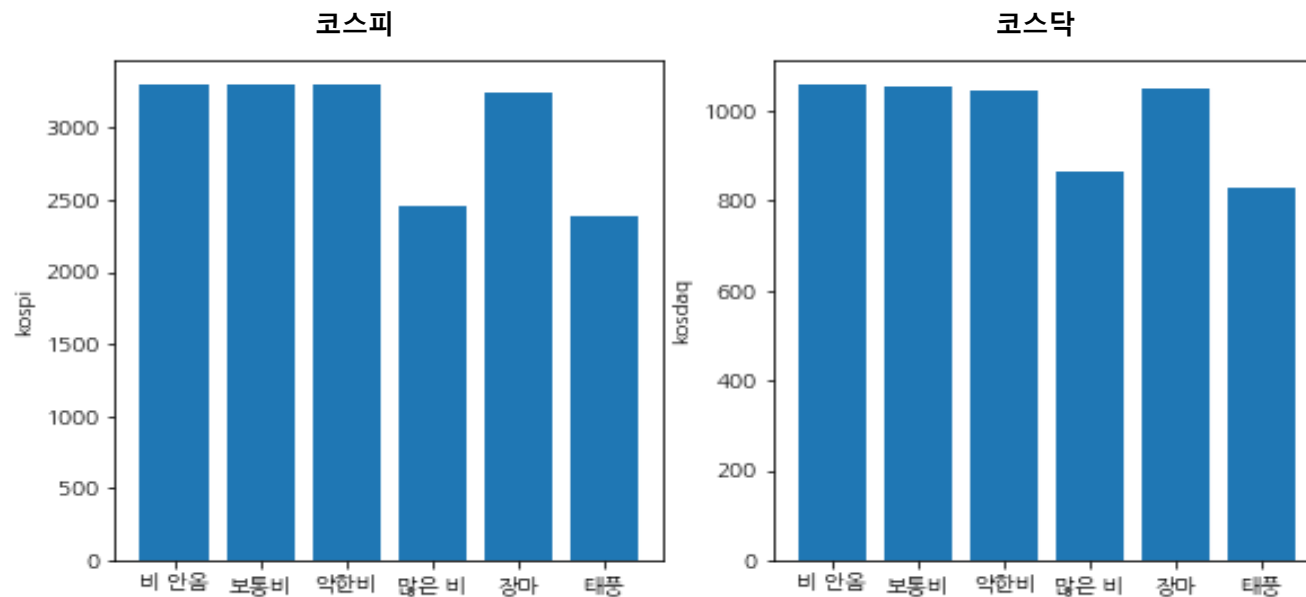


강수빈도와 강수 강도에 따른 주가지수

강수빈도

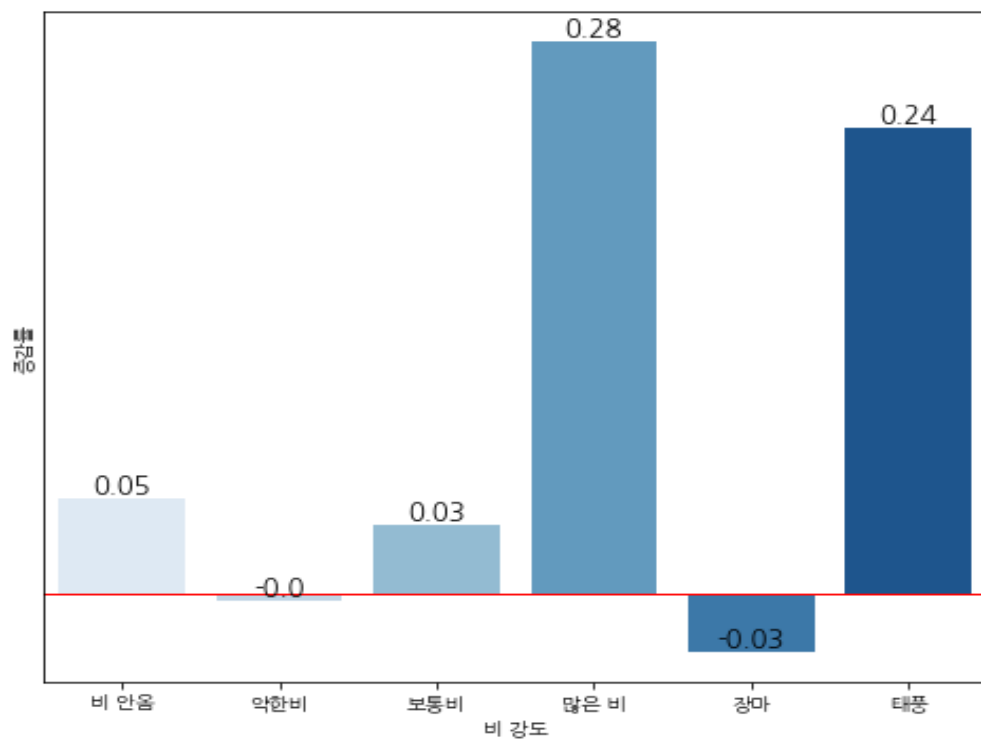


강수 강도에 따른 주가지수

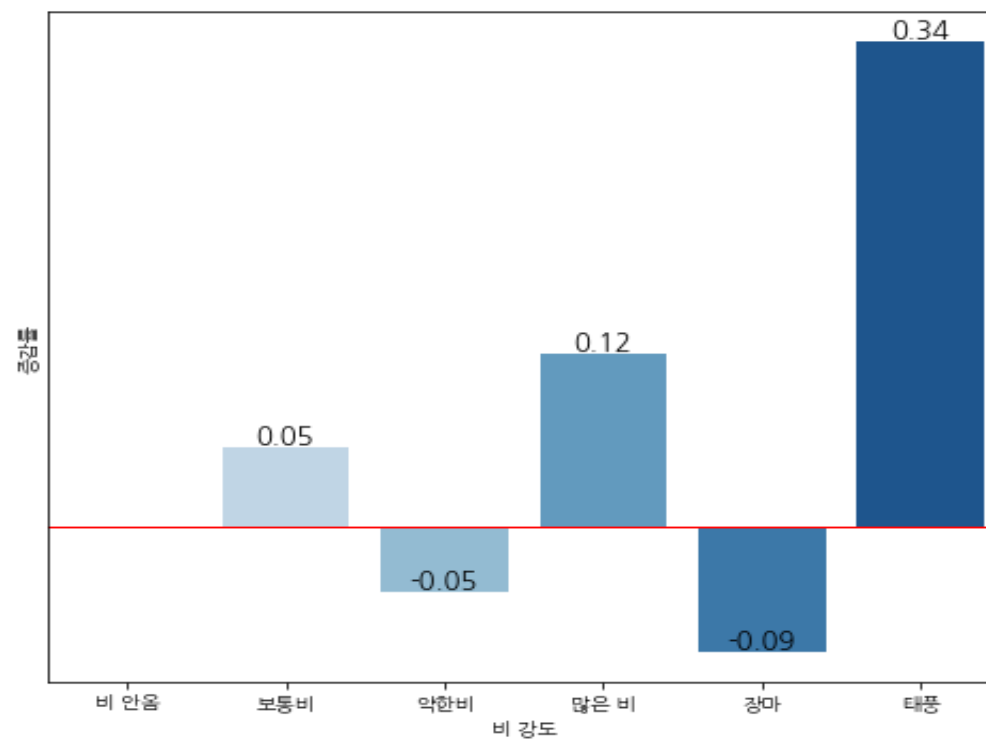


강수 강도에 따른 증감률

코스피 증감률



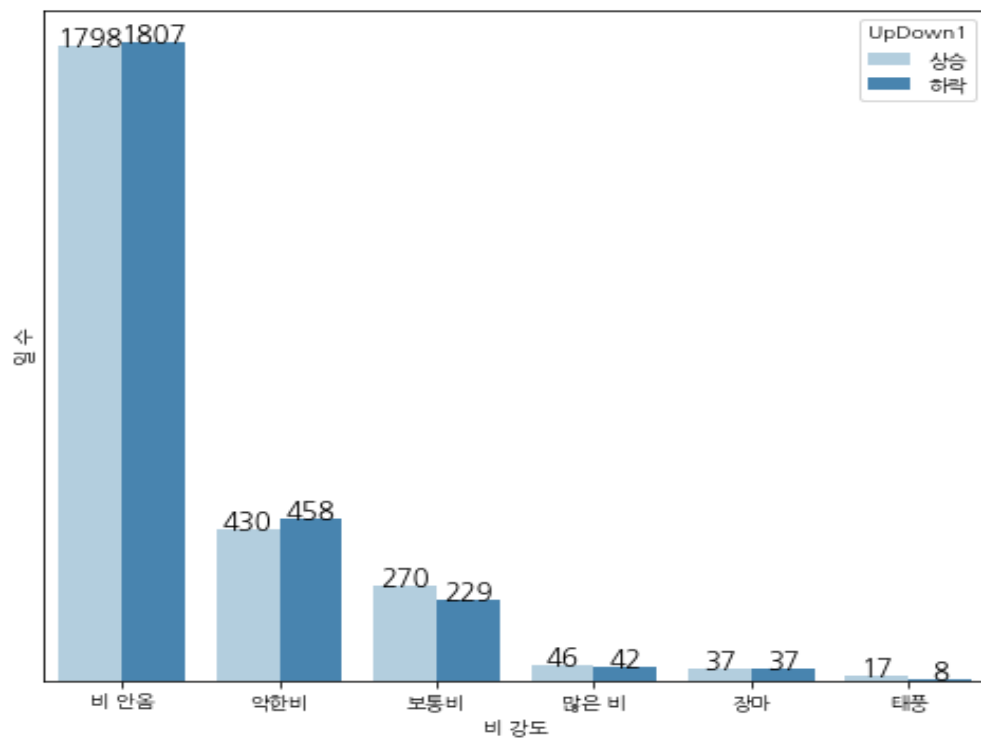
코스닥 증감률



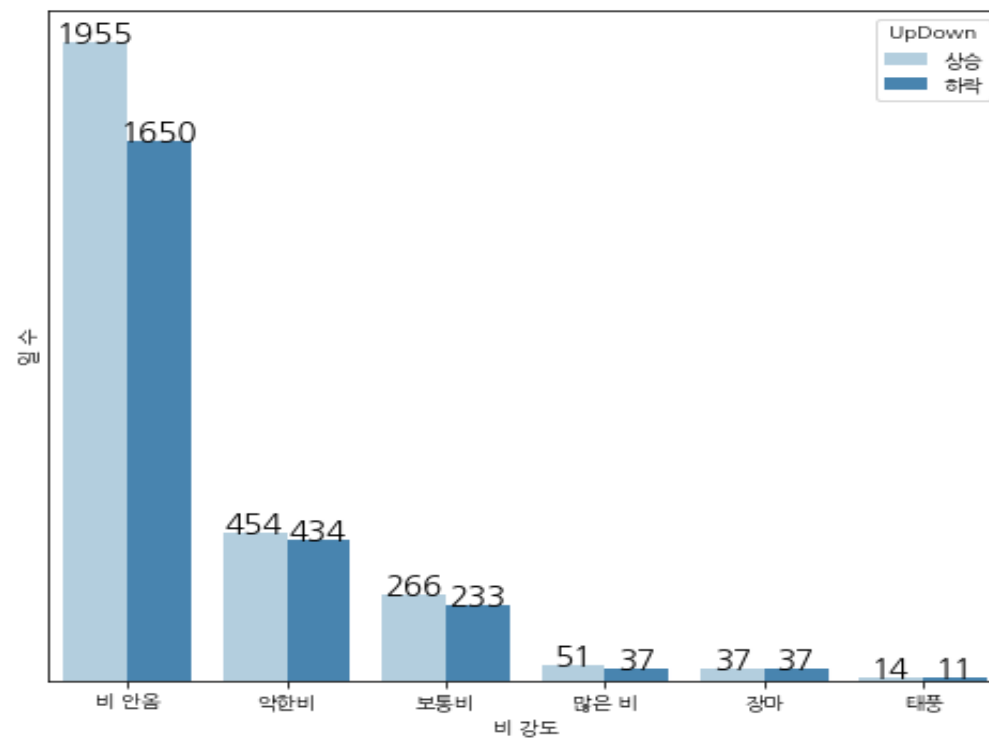
데이터 분석 — 전운량과 주가 분석

강수 강도에 따른 상승, 하락 일 수

코스피



코스닥





Part 3-2

데이터 분석

- 2010년 전후

1. 기온 (최고기온)

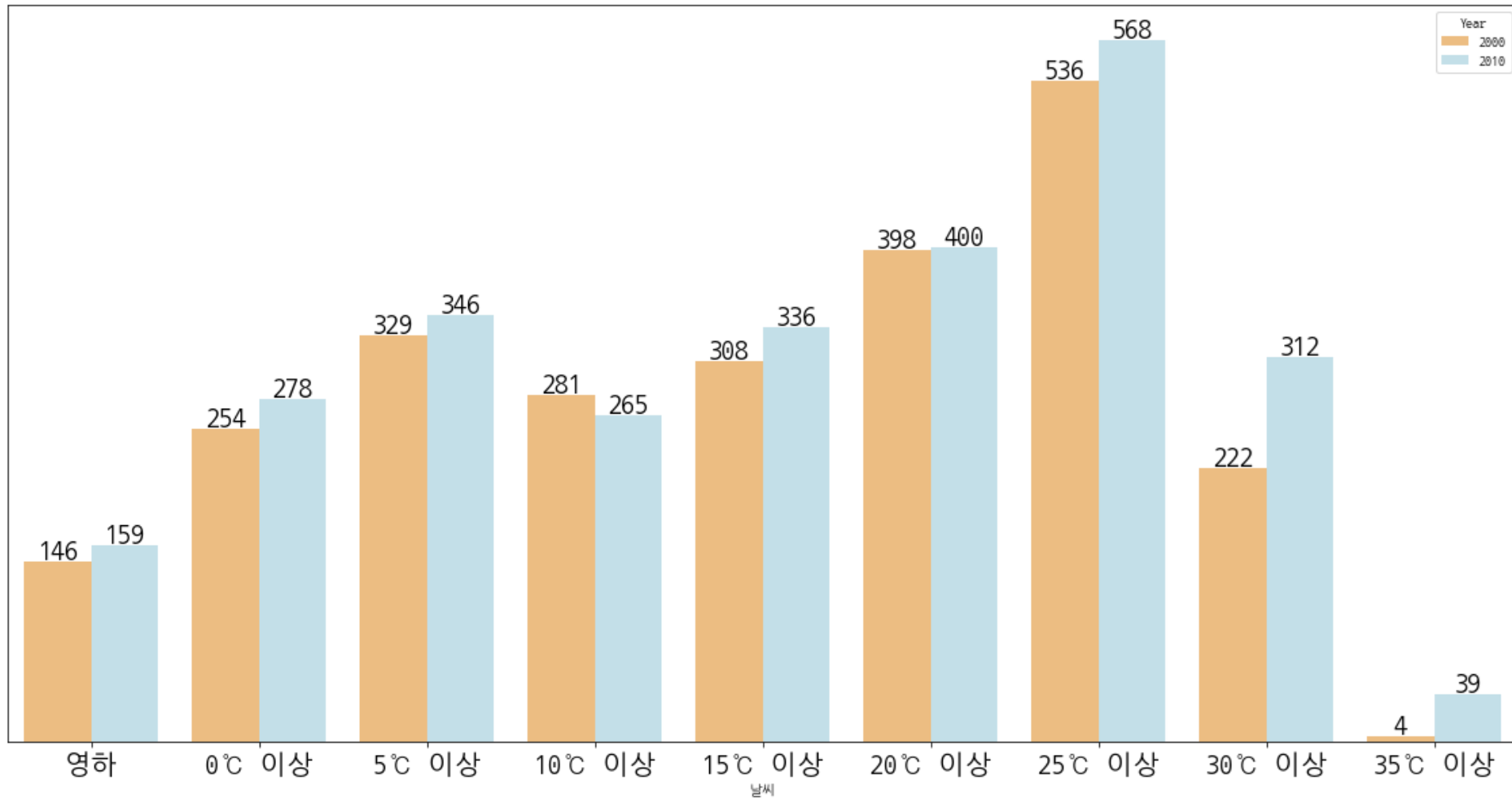
2. 불쾌지수

3. 전운량

4. 태풍

Part 3-2 데이터 분석 - 2010년 전후 분석

2010년 전후 기온 구간별 일 수

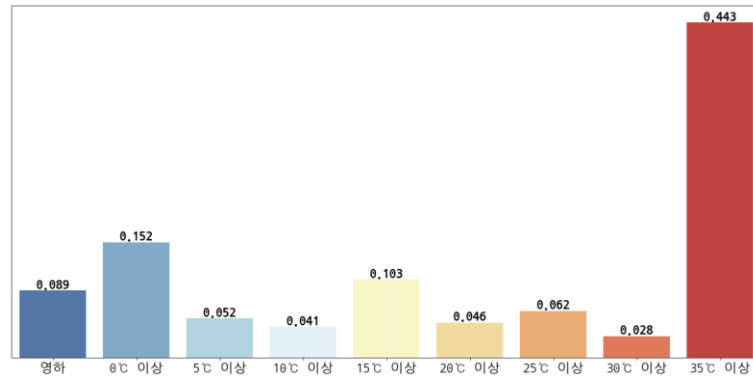


Part 3-2 데이터 분석 - 2010년 전후 분석

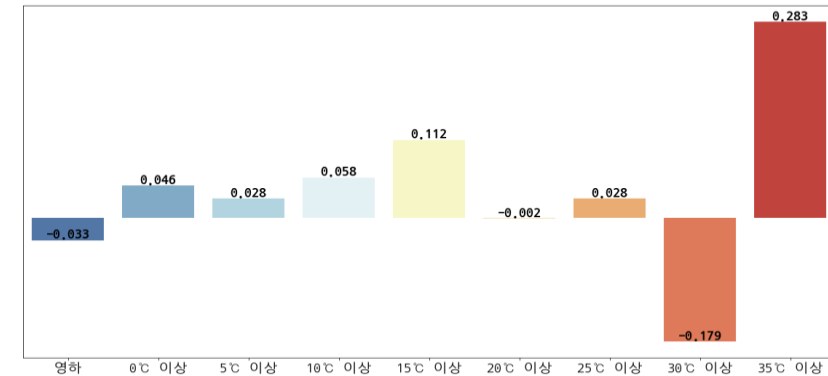
기온별 코스피, 코스닥 증감률

코스피 증감률

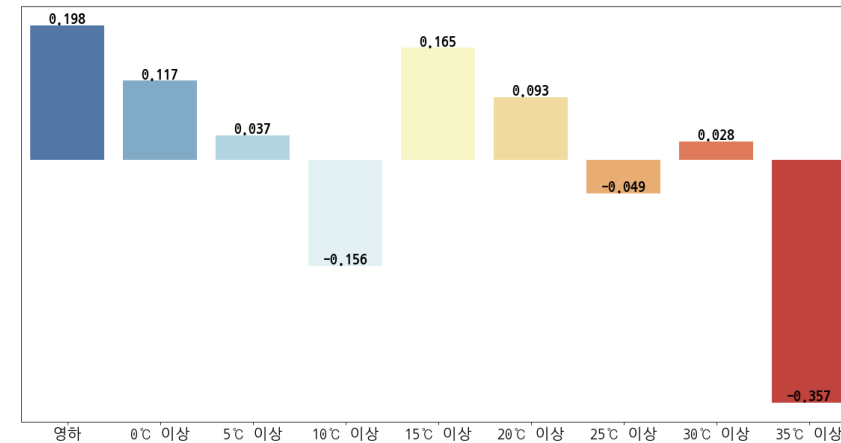
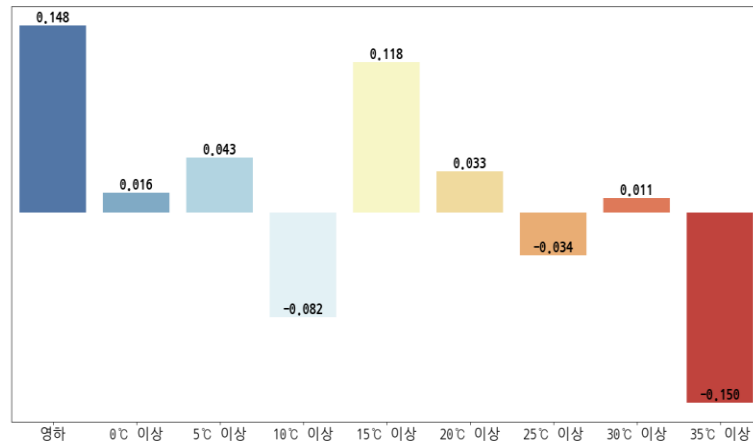
2001~2010년



코스닥 증감률



2011~2021년





Part 3-2

데이터 분석

- 2010년 전후

1. 기온 (최고기온)

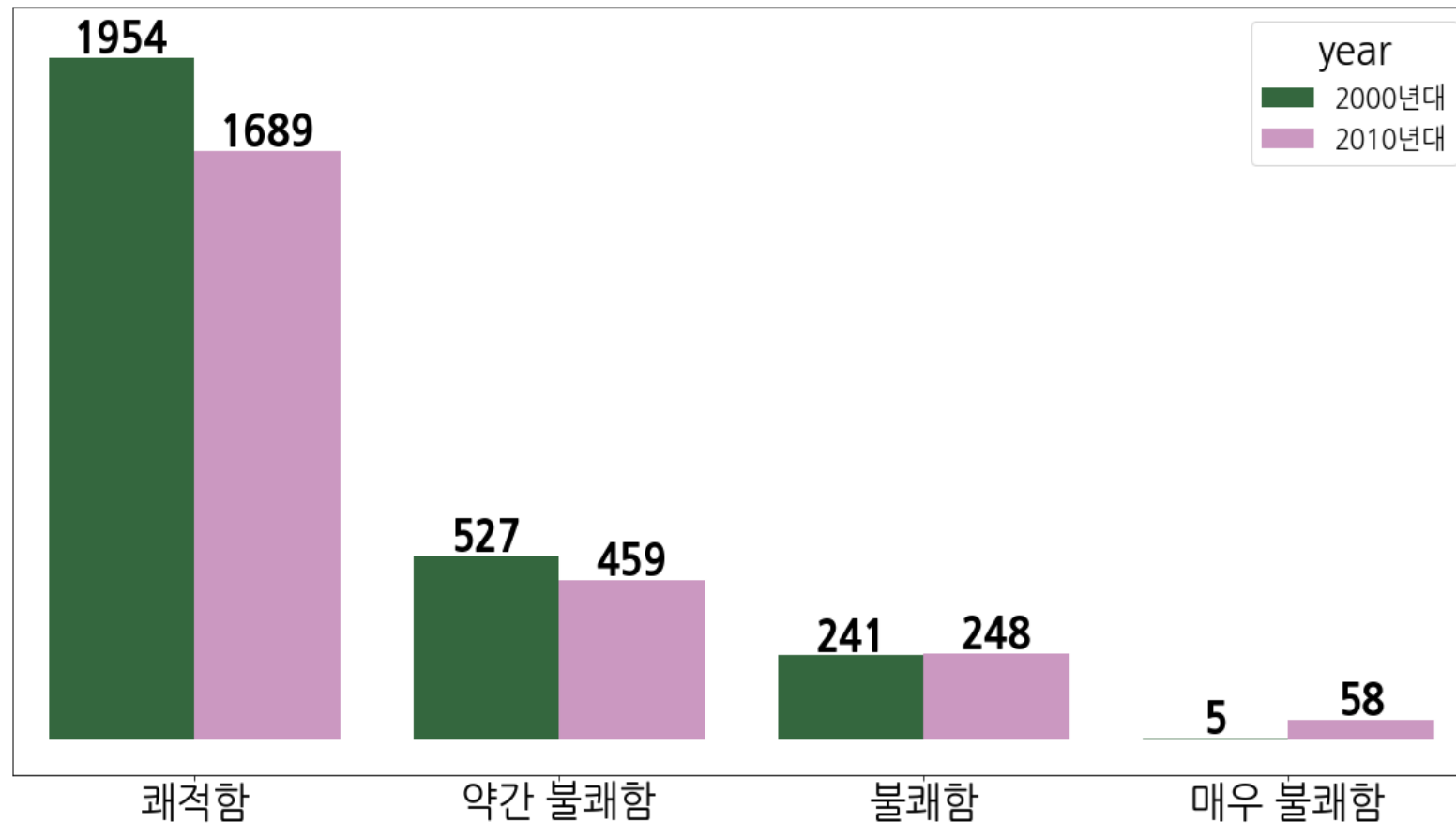
2. 불쾌지수

3. 전운량

4. 태풍

Part 3-2 데이터 분석 - 2010년 전후 분석

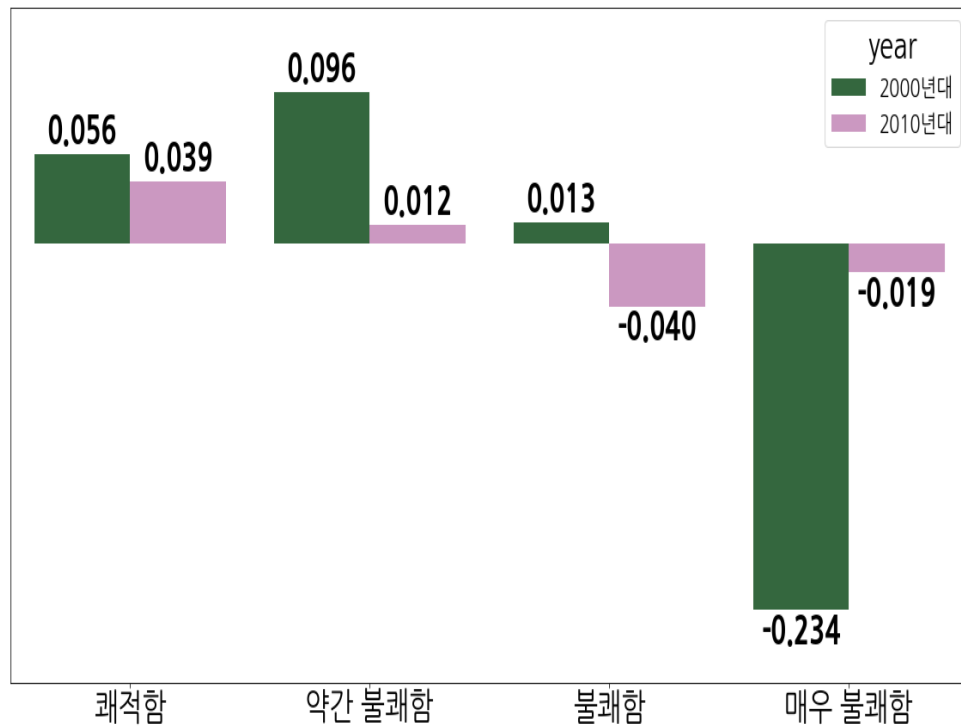
2010년 전후 불쾌지수 구간별 일 수



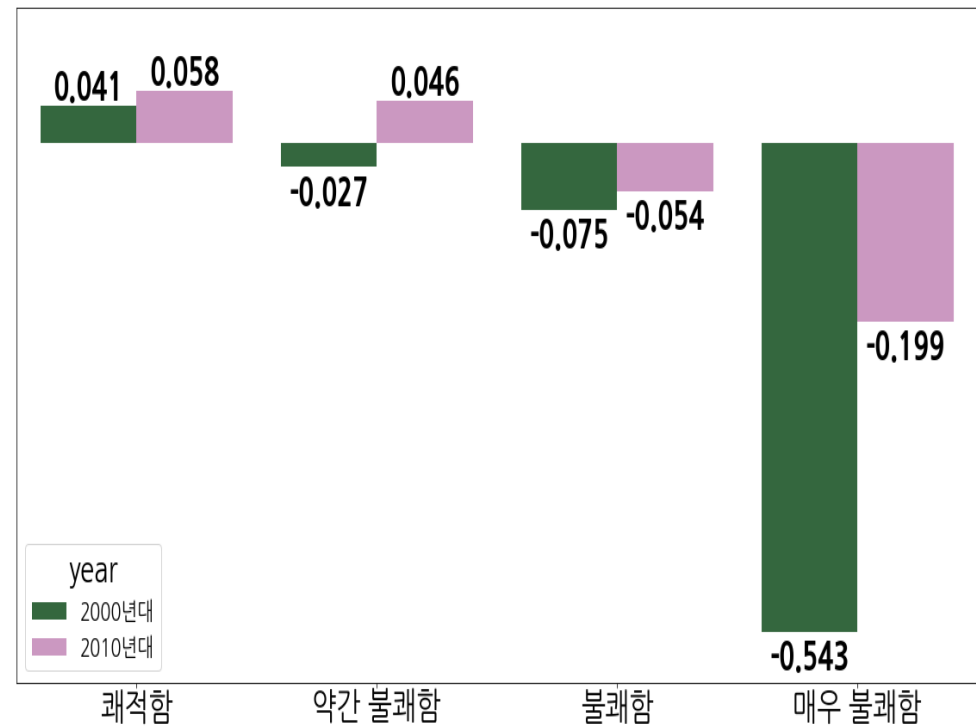
Part 3-2 데이터 분석 - 2010년 전후 분석

2010년 전후 불쾌지수 구간별 증감률

코스피



코스닥





Part 3-2

데이터 분석

- 2010년 전후

1. 기온 (최고기온)

2. 불쾌지수

3. 전운량

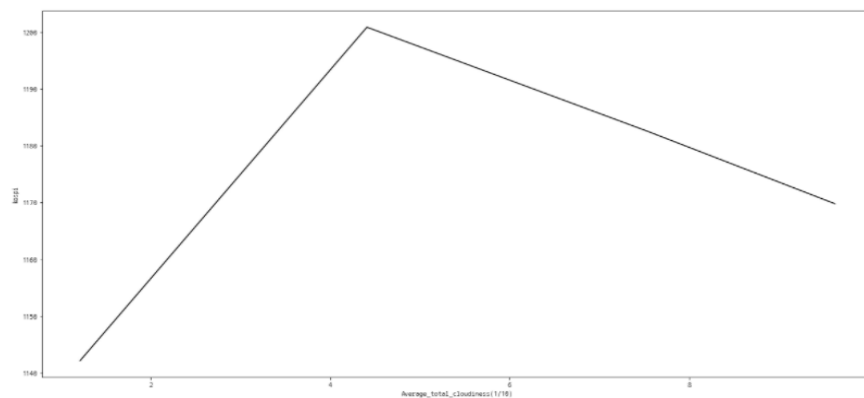
4. 태풍

Part 3-2 데이터 분석 - 2010년 전후 분석

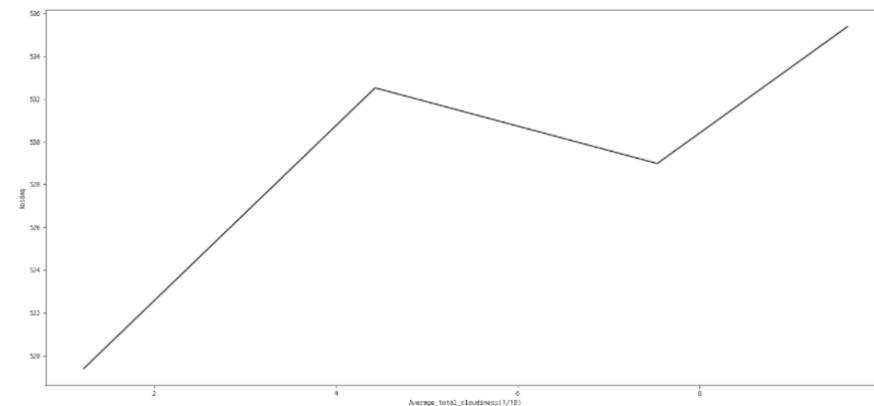
2010년 전후 전운량에 따른 평균 주가 지수

코스피 지수

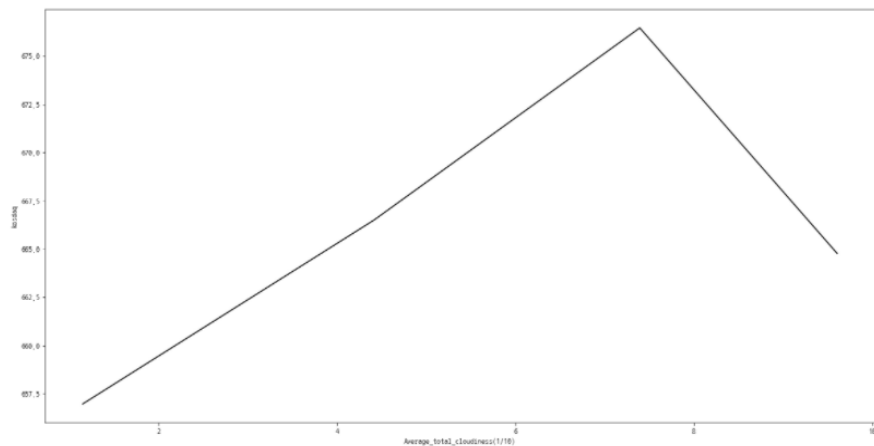
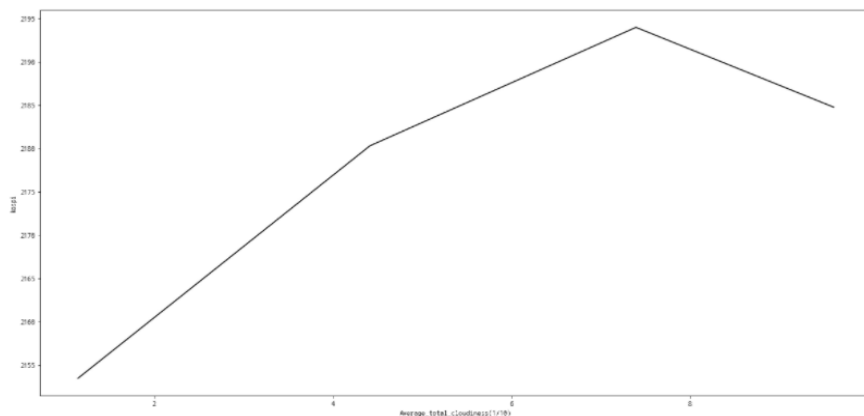
2001~2010



코스닥 지수



2011~2021

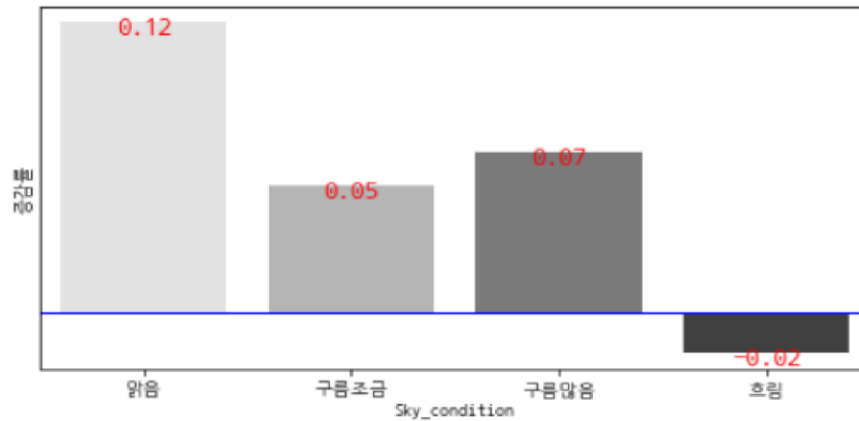


Part 3-2 데이터 분석 - 2010년 전후 분석

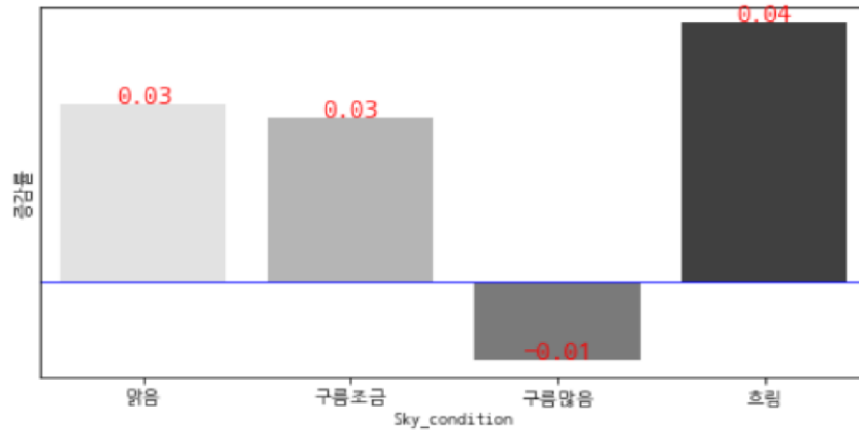
2010년 전후 전운량에 따른 평균 증감률

코스피 증감률

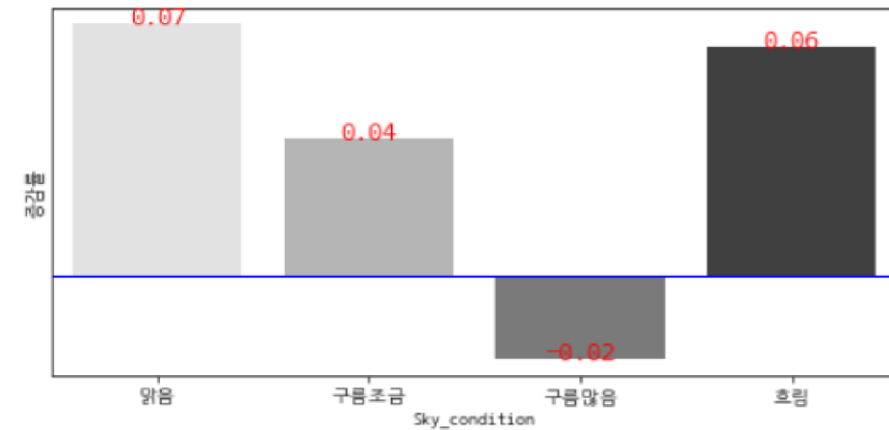
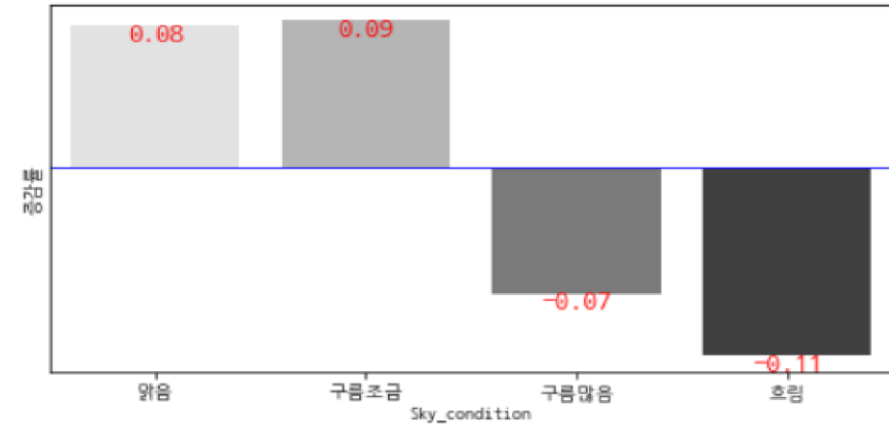
2001~2010



2011~2021



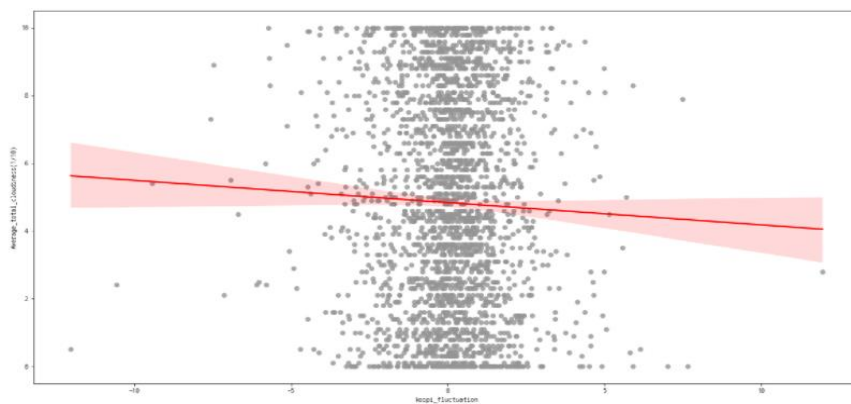
코스닥 증감률



Part 3-2 데이터 분석 - 2010년 전후 분석

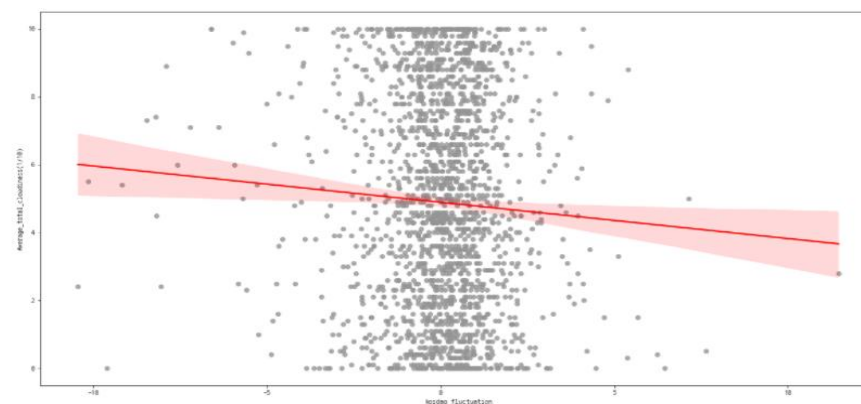
전운량과 증감률의 관계 분포도

코스피

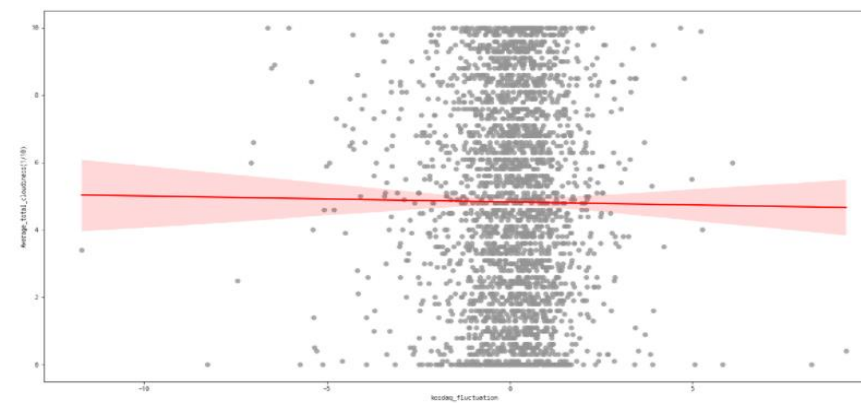
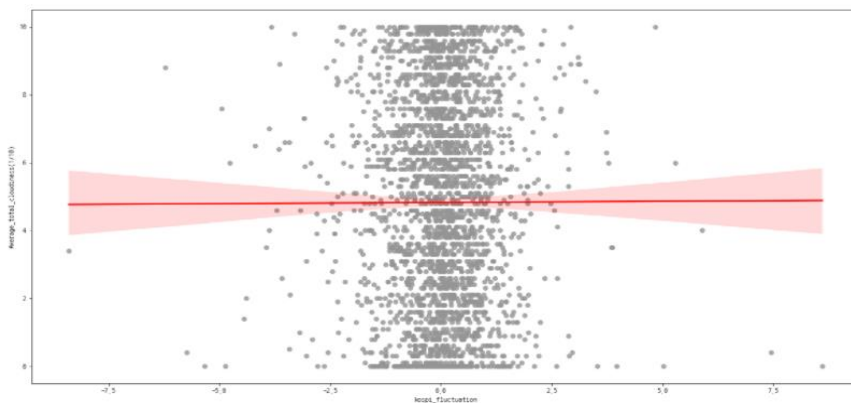


2001~2010

코스닥



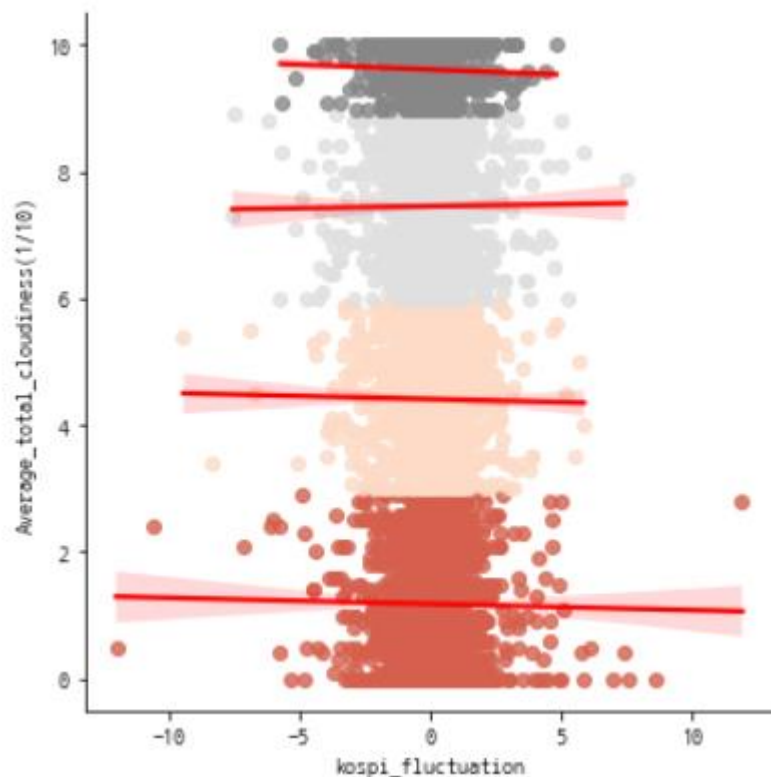
2011~2021



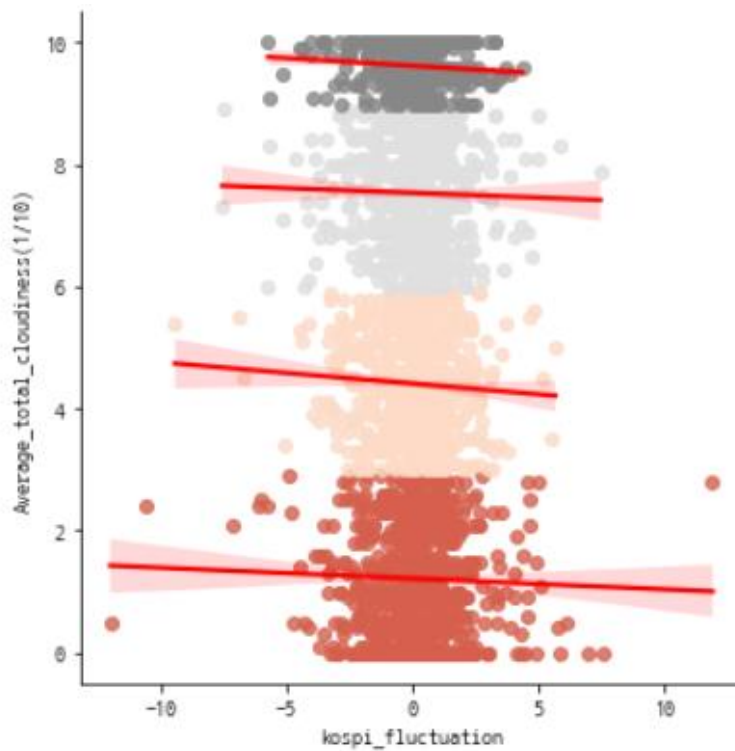
Part 3-2 데이터 분석 - 2010년 전후 분석

전운량과 증감률의 관계 분포도 - 전운량 구간별 코스피 증감률

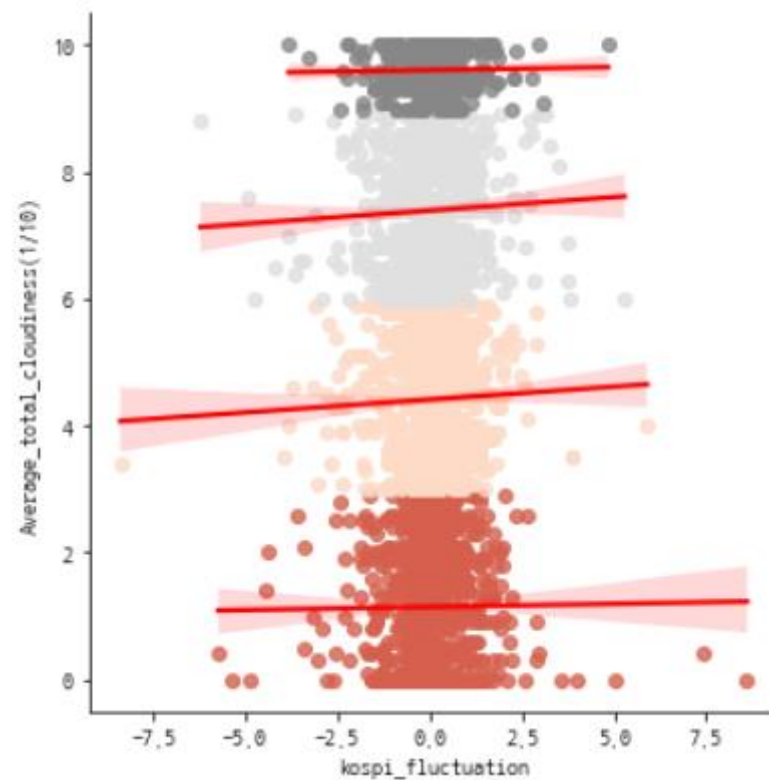
전체 일자의 분포도



2010년 까지의 분포도



2011년부터의 분포도



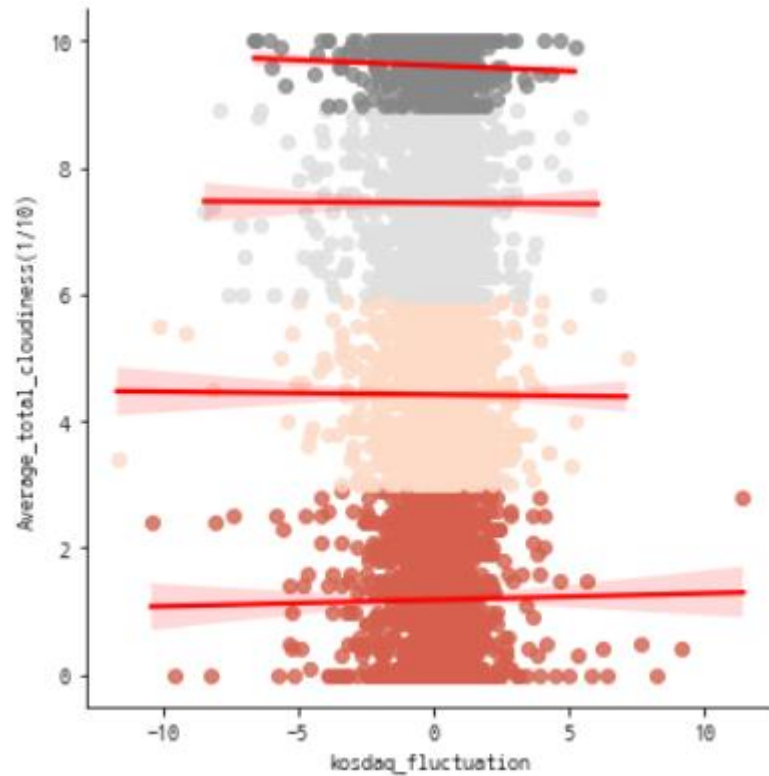
Sky_condition

- 구름조금
- 맑음
- 구름많음
- 흐림

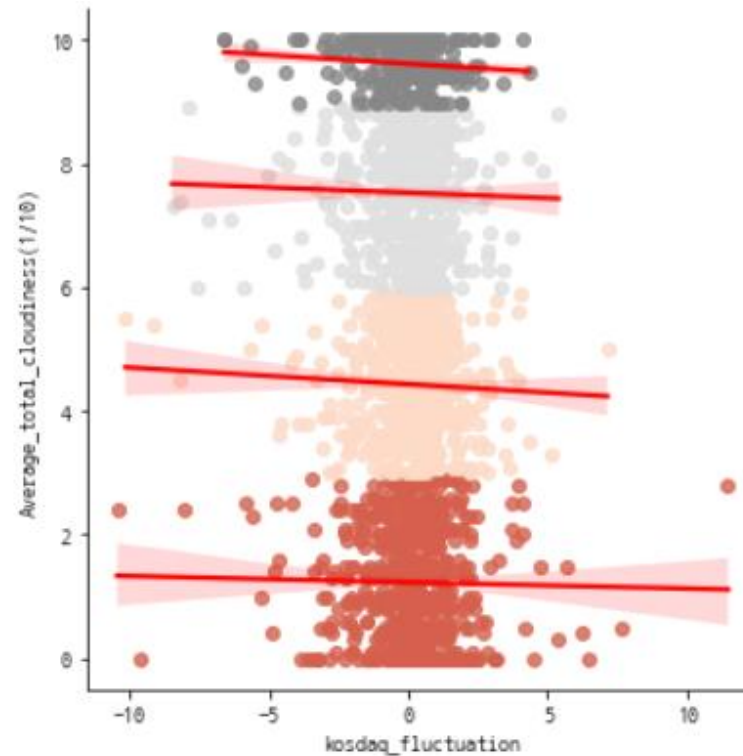
Part 3-2 데이터 분석 - 2010년 전후 분석

전운량과 증감률의 관계 분포도 - 전운량 구간별 코스닥 증감률

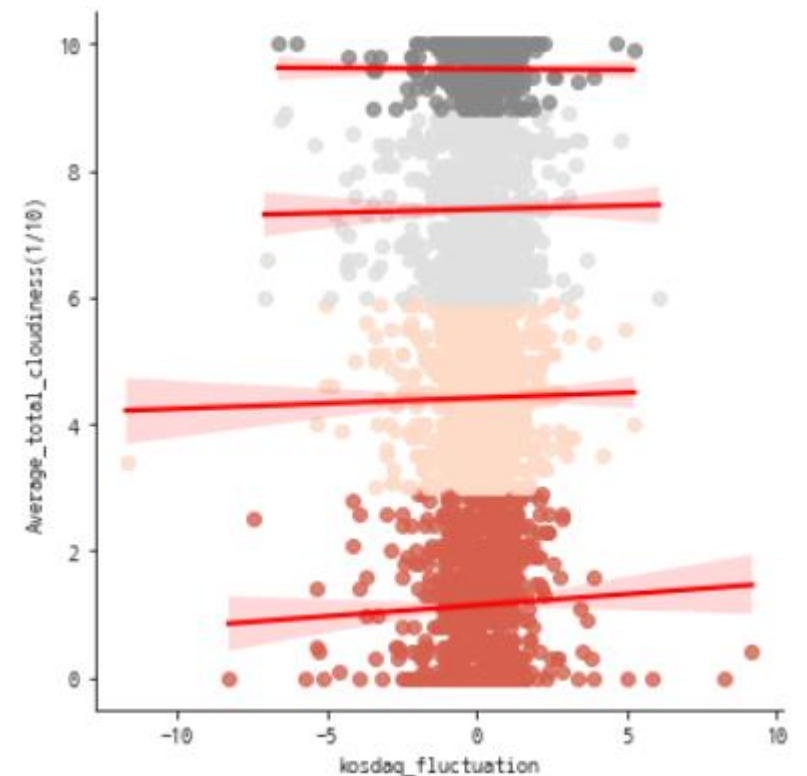
전체 일자의 분포도



2010년 까지의 분포도



2011년부터의 분포도



Sky_condition
● 구름조금
● 맑음
● 구름많음
● 흐림



Part 3-2

데이터 분석

- 2010년 전후

1. 기온 (최고기온)

2. 불쾌지수

3. 전운량

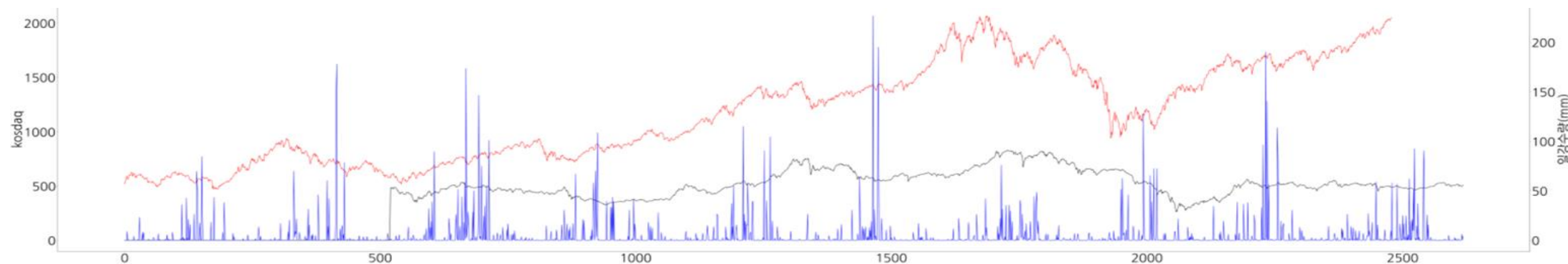
4. 태풍

Part 3-2 데이터 분석 - 2010년 전후 분석

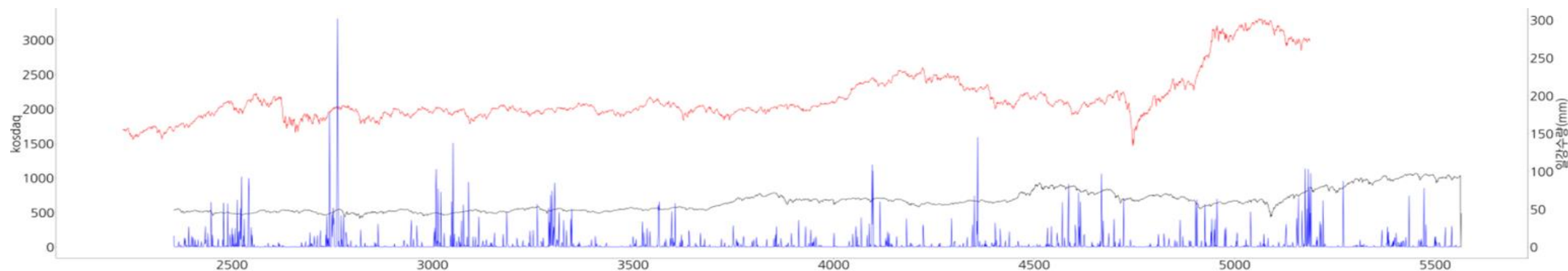
강수량에 따른 주가지수

코스피
코스닥
강수량

2001~2010



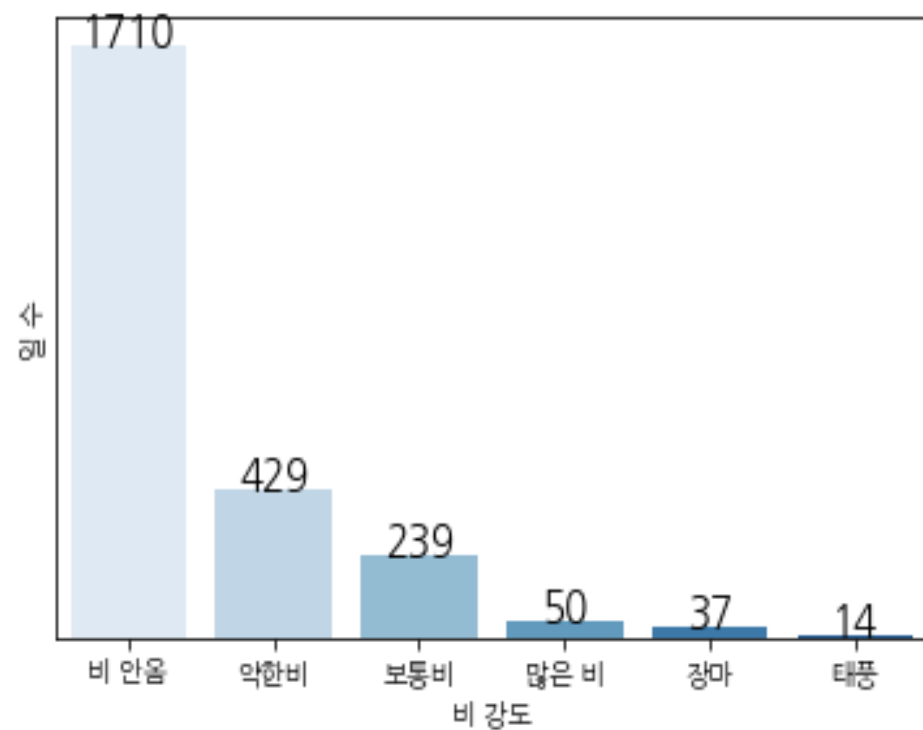
2011~2021



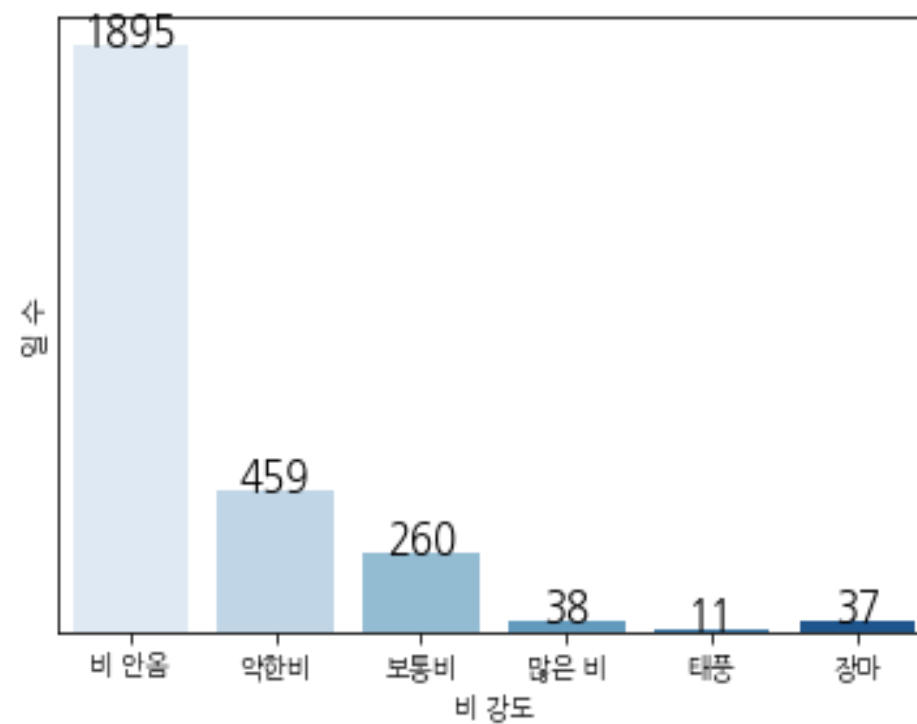
Part 3-2 데이터 분석 - 2010년 전후 분석

강수빈도

2010년 이전

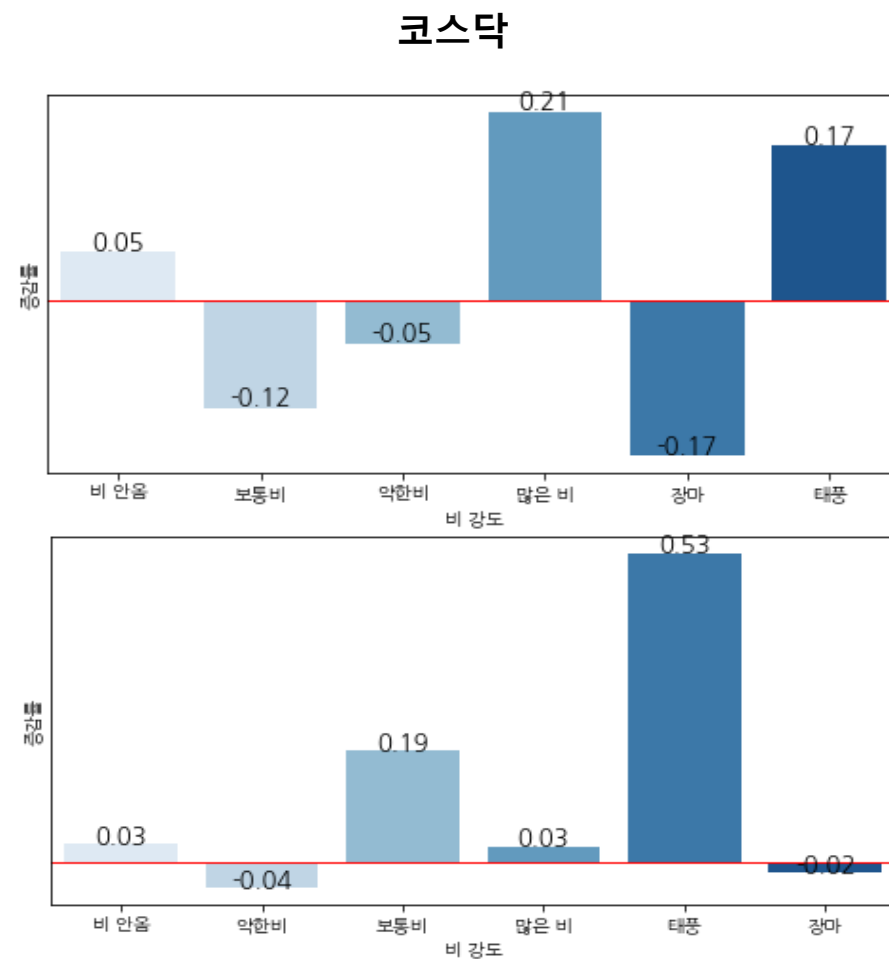
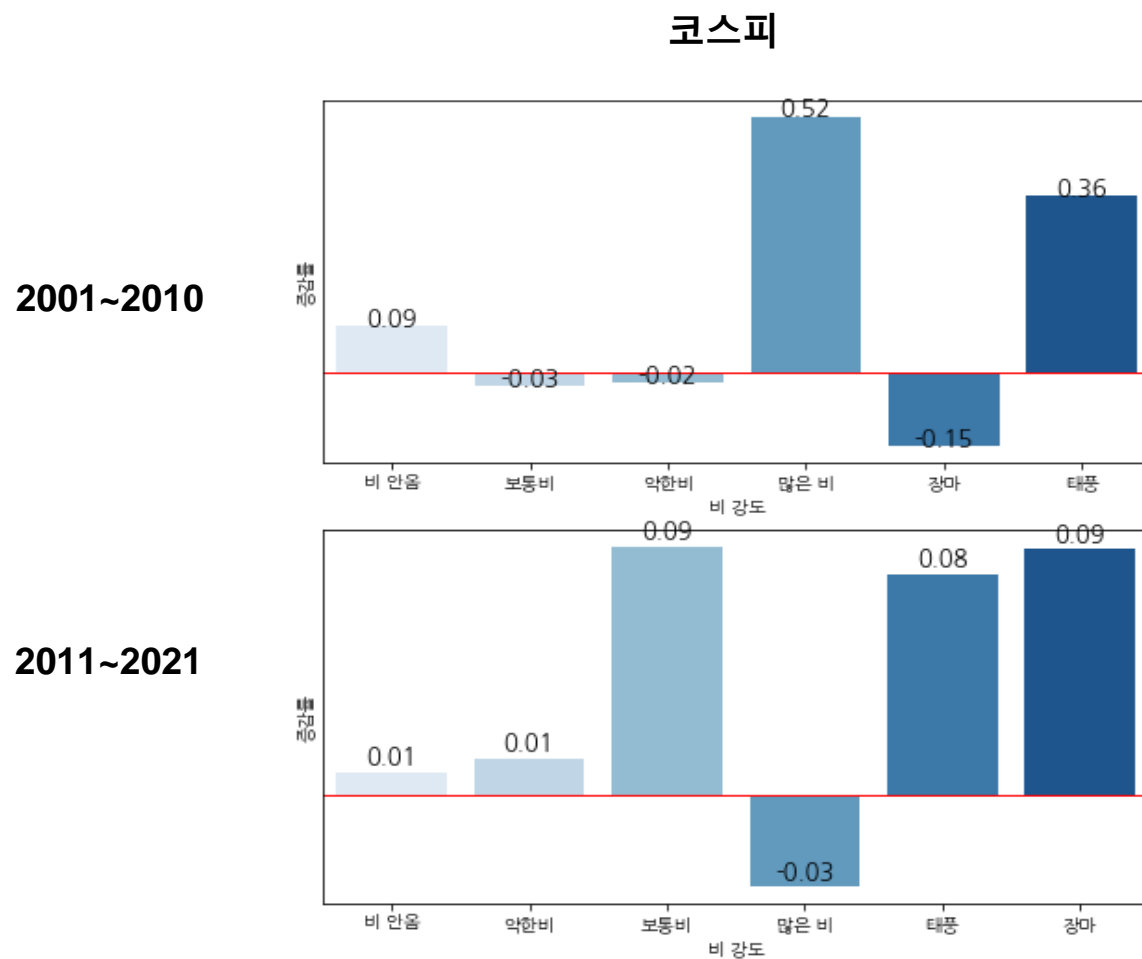


2011년 이후



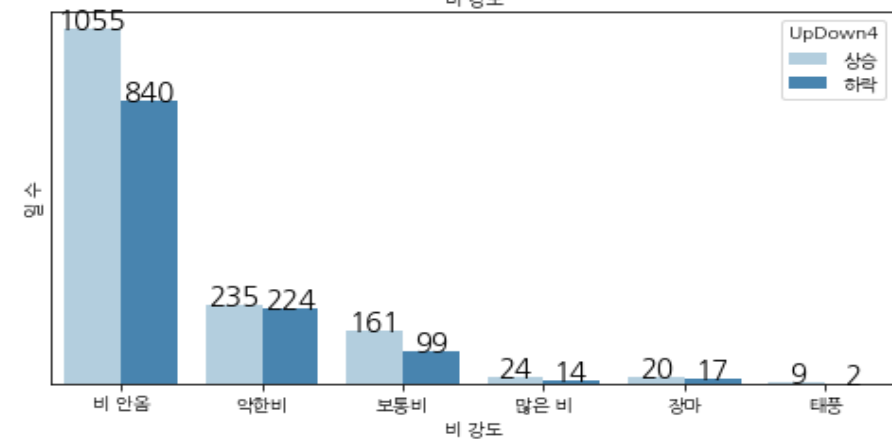
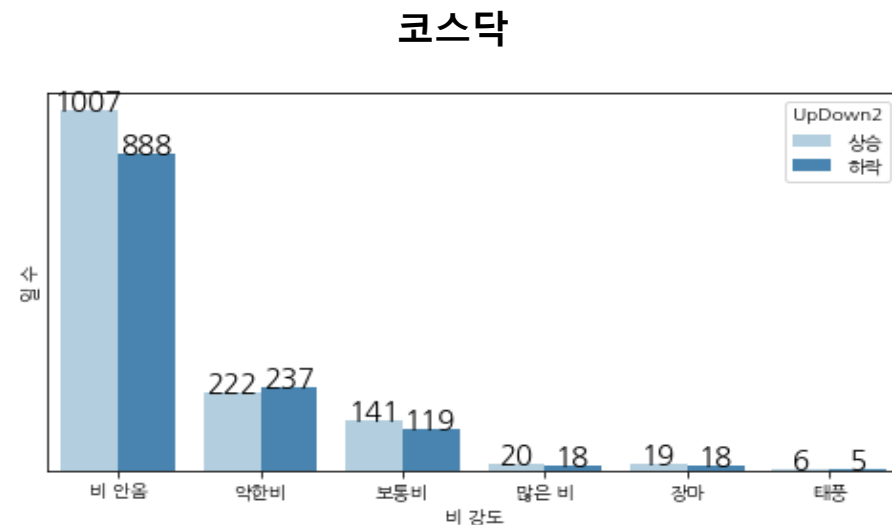
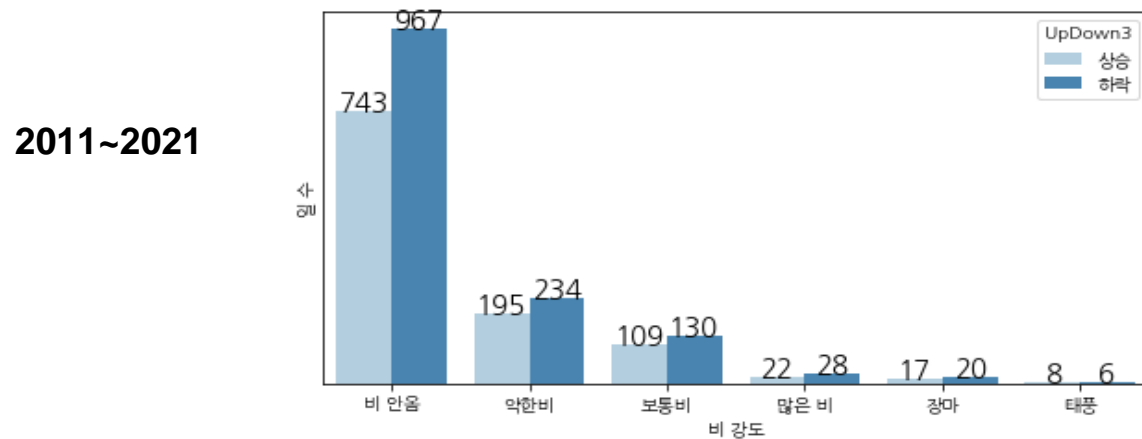
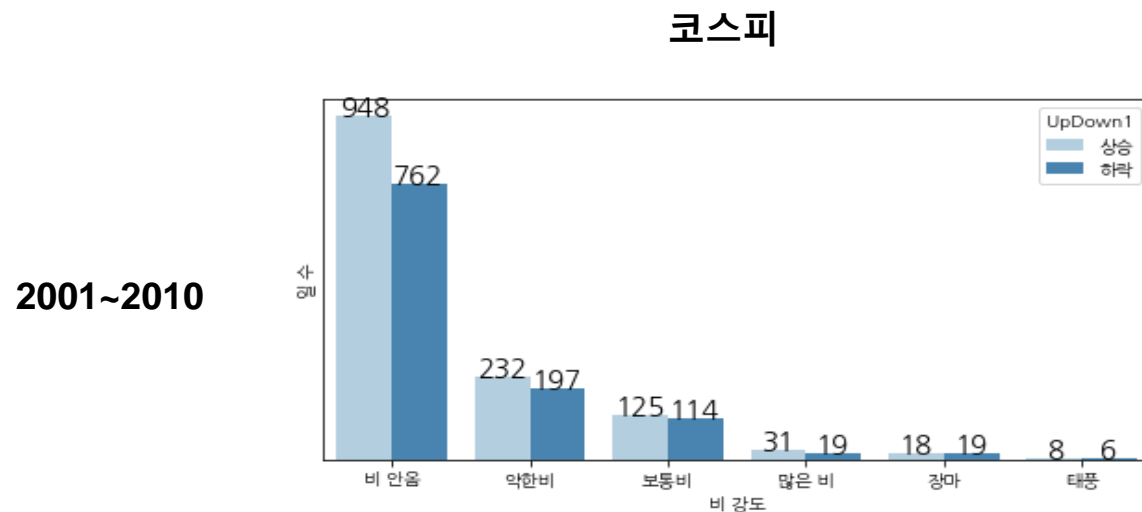
Part 3-2 데이터 분석 - 2010년 전후 분석

2010년 전후 강수 강도에 따른 평균 증감률



Part 3-2 데이터 분석 - 2010년 전후 분석

2010년 전후 강수 강도에 따른 상승, 하락 일 수





Part 4

분석 결과

1. 분석 결과

분석 결과



기온

>>

기온과 주가 지수의 증감률은 2010년 이전에는 기온이 올라갈 수록 조금씩 상승세를 보이고 2011년 이후부터는 기온이 올라갈 수록 전체적으로 하락세를 보인다. 하지만 그 차이가 의미 있을 정도로 크지 않아 기온과 주가와는 큰 관련이 없어 보인다.



불쾌지수

>>

불쾌지수는 결국 기온과 습도에서 파생된 변수라 기온의 결과와 비슷하게 나올 수 밖에 없다고 판단된다. 그래서 기온과 마찬가지로 불쾌지수역시 주가와는 큰 관련이 없어 보인다.



흐림

>>

2010년 전후로 분석해본 결과, 2010년까지의 주가지수의 증감률은 구름양이 많아질 수록 하락하는 경향을 보이는 반면, 스마트폰이 보급되기 시작한 2011년 이후부터는 전운량에 크게 영향을 받지 않는 것으로 보인다.



태풍

>>

전체, 그리고 스마트폰이 보급되기 시작한 2010년 기준으로 봤을 때, 큰 차이는 없는 것 같다.



Part 5

프로젝트 마무리

1. 프로젝트 후기
2. Q&A

프로젝트 후기

임채건

실력이 부족해서 데이터를 수집하는 과정부터도 어려웠고, 수집한 데이터를 데이터베이스에 저장하는 과정, 그리고 데이터를 가공하는 전처리부분까지 팀원분들이 정말 많이 도와주셨습니다. 중간에 팀장분이 나가면서 의욕이 저하되는 시기가 있었지만, 개의치 않고 프로젝트를 완성시켜 나가는 팀원들을 보며 덕분에 저도 끝까지 포기하지 않고 완료할 수 있었던 것 같습니다. 제 부족한 부분들을 진단할 수 있었고, 더 열심히 공부해야겠다는 의지가 생긴 프로젝트 경험이었습니다.

조세익

데이터 자체의 한계로 인해 좀 더 다양한 분석을 하면서 한계를 보완할만한 방법을 찾아보고 싶었지만, 다양한 한계로 인하여 충분한 분석을 행하지 못해 아쉽습니다. 그래도 남은 팀원들끼리 열심히 해서 마무리 지을 수 있어서 다행이라고 생각합니다.

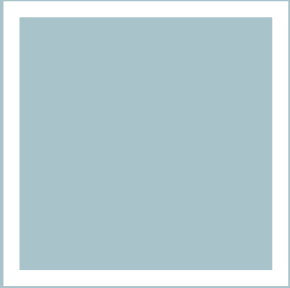
조휘

처음 접해보는 것들이 많아서 어려웠지만 조원 분들이 잘 도와주시고 서로 협력을 열심히 해 주셔서 잘 마무리 짓게 된 것 같습니다. 기획 단계에서 생각한 만큼의 의미있는 결과는 나오지 않아서 아쉽지만 그 자체로도 의미가 있는 조별 활동이었던 것 같습니다.

한상준

경험도 없고 생소해서 시간을 꽤 많이 투자한 것 같습니다. 신선한 주제인 만큼 어떤 데이터를 사용할지, 어떤 내용이 필요한지 조금이나마 느끼게 되었습니다. 중간에 조장이 나가서 방향을 잡는데 어려움을 겪었지만 조원 분들의 도움을 받아 마무리할 수 있었습니다. 제가 봐도 엉성하고 많이 부족하지만, 주말 밤까지 시간을 들여서 열심히 한 것에 만족합니다. 좋은 경험이었습니다.

Q&A



내일주식시장은맑겠조

감 사 합 니 다