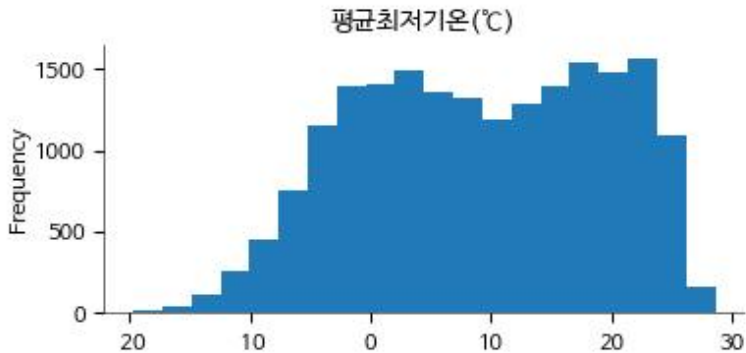
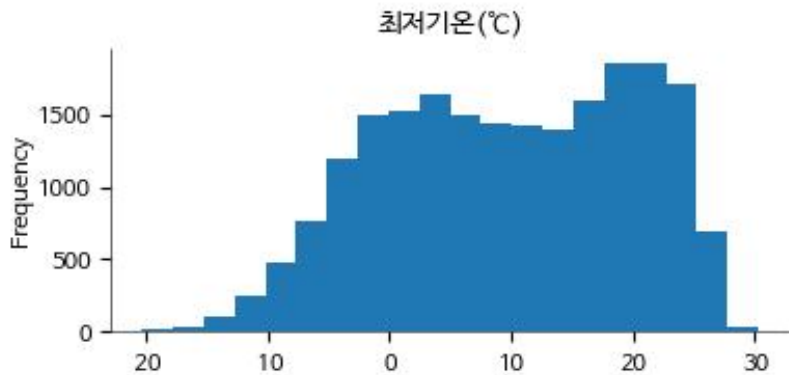


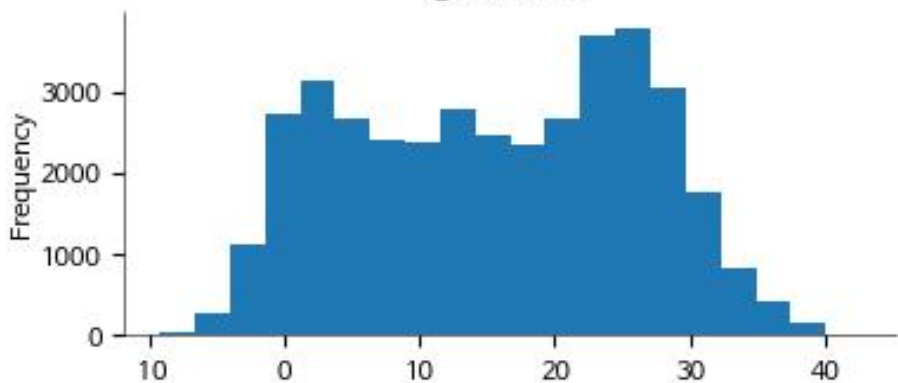
EDA

Distributions

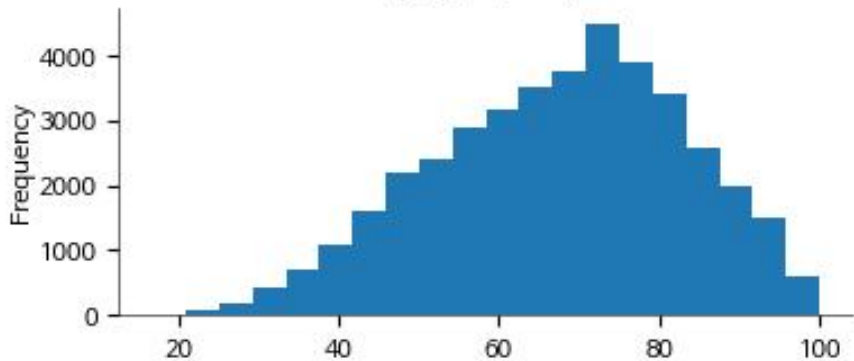


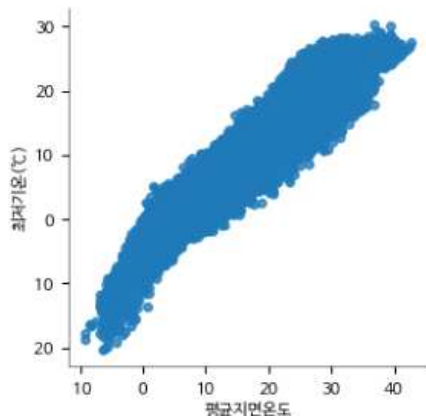


평균지면온도

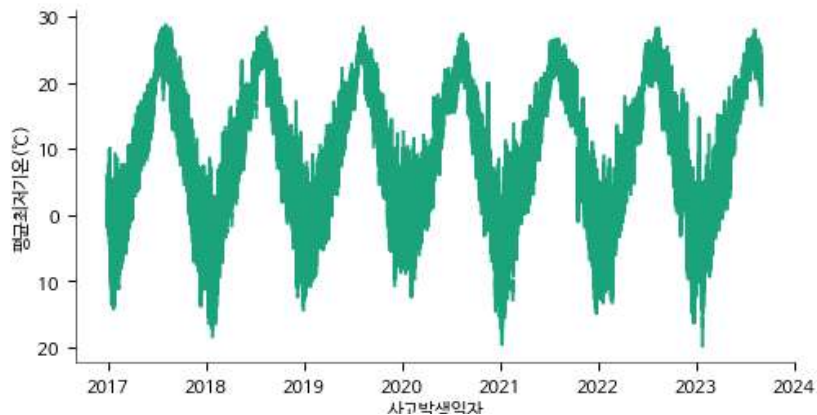


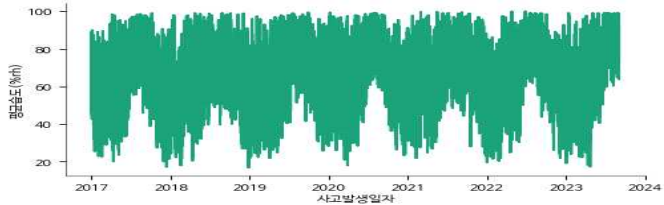
평균습도(%rh)



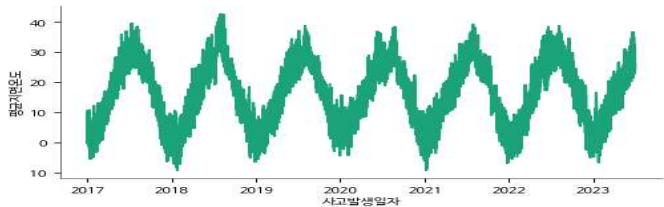


Time series





- (매우) 뚜렷한 주기성



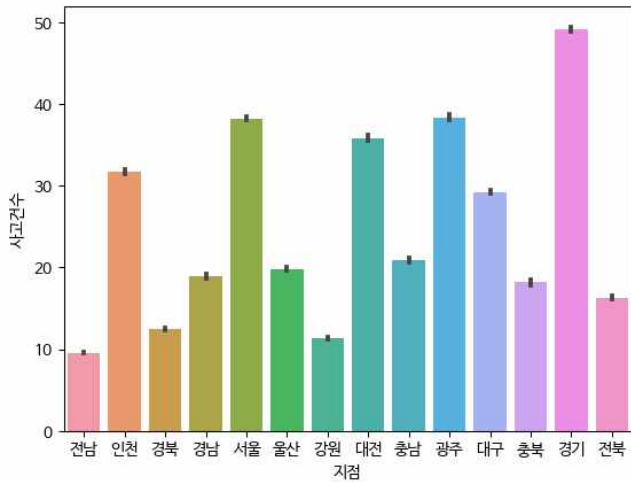
원본 데이터 사고건수 살펴보기 (5 - num summary)

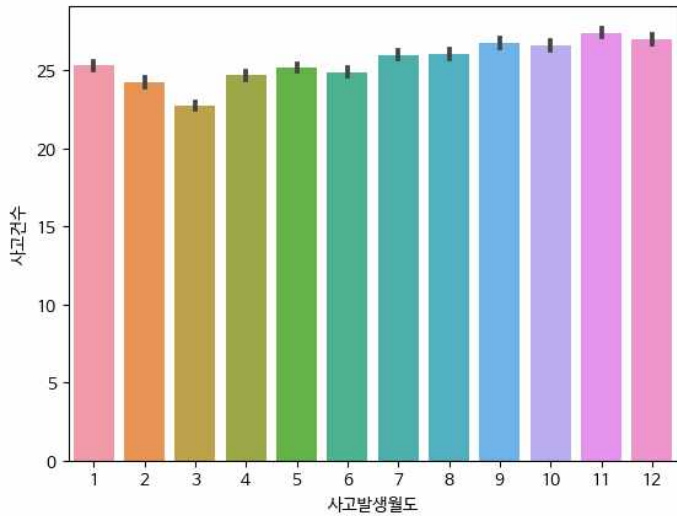
```
In [16]: Q1, Q2, Q3, Q4 = np.percentile(dat["사고건수"], [25, 50, 75, 100])  
         print(Q1)  
         print(Q2)  
         print(Q3)  
         print(Q4)
```

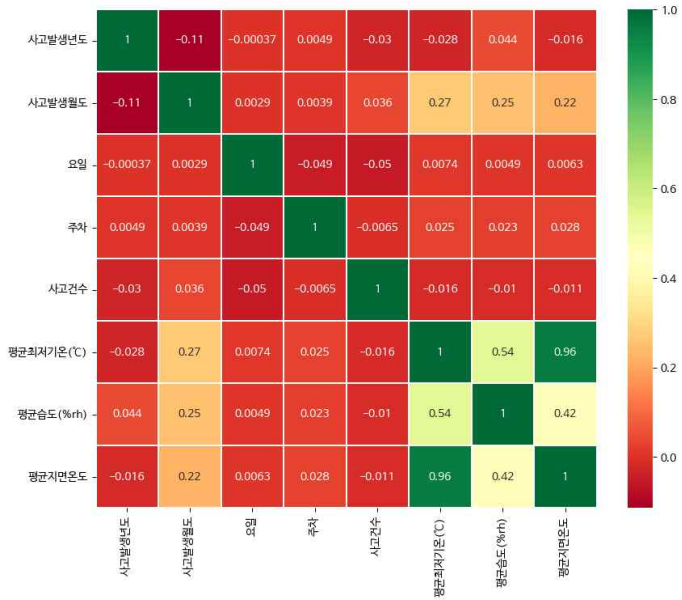
```
6.0  
16.0  
34.0  
1519.0
```

```
In [17]: dat['사고건수'].mean()
```

```
Out[17]: 24.963706888032416
```





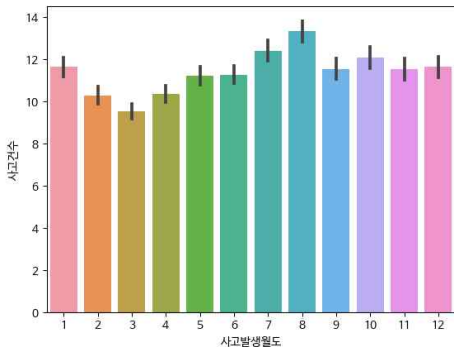




강원, 제주

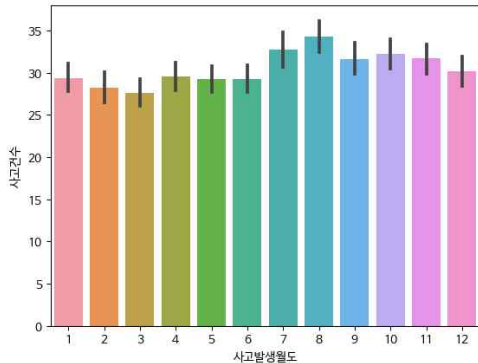
```
▶ gangwon=dat.loc[dat['지역(시도)']=='강원']  
sns.barplot(x='사고발생월도', y='사고건수', data=gangwon)
```

☞ <Axes: xlabel='사고발생월도', ylabel='사고건수'>



```
▶ jeju=dat.loc[dat['지역(시도)']=='제주']  
sns.barplot(x='사고발생월도', y='사고건수', data=jeju)
```

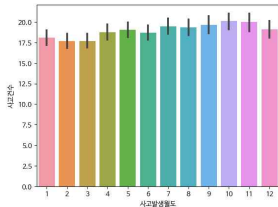
☞ <Axes: xlabel='사고발생월도', ylabel='사고건수'>



경남, 부산, 서울, 경기

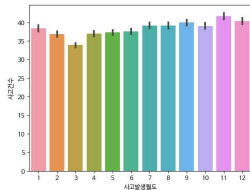
```
gyeongnae=dat.loc[dat['지역(시도)']=='경남']
sns.barplot(x='사고발생월도',y='사고건수',data=gyeongnae)
```

```
<axes: xlabel='사고발생월도', ylabel='사고건수'>
```



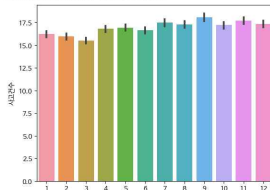
```
seoul=dat.loc[dat['지역(시도)']=='서울']
sns.barplot(x='사고발생월도',y='사고건수',data=seoul)
```

```
<axes: xlabel='사고발생월도', ylabel='사고건수'>
```



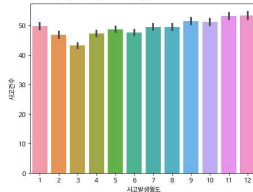
```
busan=dat.loc[dat['지역(시도)']=='부산']
sns.barplot(x='사고발생월도',y='사고건수',data=busan)
```

```
<axes: xlabel='사고발생월도', ylabel='사고건수'>
```



```
gyeonggi=dat.loc[dat['지역(시도)']=='경기']
sns.barplot(x='사고발생월도',y='사고건수',data=gyeonggi)
```

```
<axes: xlabel='사고발생월도', ylabel='사고건수'>
```



Issue

-일교차: 앞에 쉼표 붙은 데이터 있어서 숫자로 인식안됨

-평균최저기온과 평균지면온도는 서로 연관이 있는 데이터, 즉 함께 사용시 다중공선성 문제 발생가능->'사고건수'(y)와 상관계수가 더 높게 나오는 쪽 활용

-연장(결빙취약 도로의 길이)이 사고건수와 음(-)의 상관관계
결빙에 취약한 도로 길이가 길수록 사고가 적다?
Drop?

-	89	20170101	2017	1 경북	군위군	Y	Y	7	1	1
-	90	20170101	2017	1 경북	고령군	Y	Y	7	1	4
-	91	20170101	2017	1 경북	청송군	Y	Y	7	1	13
-	92			경북 요약						289
-	93	20170101	2017	1 광주	북구	Y	Y	7	1	27
-	94	20170101	2017	1 광주	남구	Y	Y	7	1	8
-	95	20170101	2017	1 광주	광산구	Y	Y	7	1	36
-	96	20170101	2017	1 광주	동구	Y	Y	7	1	8
-	97	20170101	2017	1 광주	서구	Y	Y	7	1	16
-	98			광주 요약						95
-	99	20170101	2017	1 대구	동구	Y	Y	7	1	33
-	100	20170101	2017	1 대구	남구	Y	Y	7	1	8
-	101	20170101	2017	1 대구	북구	Y	Y	7	1	22
-	102	20170101	2017	1 대구	수성구	Y	Y	7	1	29
-	103	20170101	2017	1 대구	서구	Y	Y	7	1	5
-	104	20170101	2017	1 대구	달성군	Y	Y	7	1	6
-	105	20170101	2017	1 대구	달서구	Y	Y	7	1	48
-	106	20170101	2017	1 대구	중구	Y	Y	7	1	3
-	107			대구 요약						154
-	108	20170101	2017	1 대전	대덕구	Y	Y	7	1	16
-	109	20170101	2017	1 대전	중구	Y	Y	7	1	15
-	110	20170101	2017	1 대전	동구	Y	Y	7	1	15
-	111	20170101	2017	1 대전	유성구	Y	Y	7	1	36
-	112	20170101	2017	1 대전	서구	Y	Y	7	1	30
-	113			대전 요약						112
-	114	20170101	2017	1 부산	금정구	Y	Y	7	1	16
-	115	20170101	2017	1 부산	사상구	Y	Y	7	1	10
-	116	20170101	2017	1 부산	기장군	Y	Y	7	1	20
-	117	20170101	2017	1 부산	강서구	Y	Y	7	1	15
-	118	20170101	2017	1 부산	서구	Y	Y	7	1	4
-	119	20170101	2017	1 부산	영도구	Y	Y	7	1	4
-	120	20170101	2017	1 부산	사하구	Y	Y	7	1	5

- <https://github.com/wrkurdrm/SamsungXPostech/tree/main>