

데 이 터 시 각 화

# 각 시간대별 교통사고 현황

# CONTENTS

01

---

문제 제기

02

---

데이터 전처리

03

---

시각화 설명

04

---

가설검증

# 각 시간대별로 교통사고는 어디서 많이 일어날까?

출퇴근과 등하교 뿐만 아니라 우리가 놀러갈때 또한 자동차를 굉장히 많이 이용한다.

때문에 자동차를 이용할때는 항상 교통사고에 주의해야한다. 시간대 별로 교통사고 패턴을 알고있다면 사고가 많이나는 지역에서 조심할 수 있고 사고를 어느정도 예방 할 것이다. 그렇다면 각 시간대별로 교통사고는 어느곳에 많이 일어날까?

다음과 같은 가설을 세워두고 시각화를 해보려고 한다

**1.출퇴근시간에는 경기쪽에 많을 것이고 낮에는 고르게 분포되어 있을 것이다.**

**2.교통사고가 많이 일어나는 지역은 서울일것이다.**

한국교통사고 데이터를 두가지 형태로 가공했다

1. 각 시간대별 시 단위로 일어난 교통사고의 수
2. 각 시간대별 교통사고가 일어난 위도 경도

파이썬 csv모듈을 이용해 csv파일을 각 한줄씩 읽어가면서 만들어놓은 함수를 이용해 사전 두가지에 나누어 정보를 넣어놨다

각 시간대별 시 단위로 일어난 교통사고의 수의 정보는 results\_dic1에  
각 시간대별 교통사고가 일어난 위도경도는 results\_dic2에 넣어두었다

```
f = open('C:/Users/lg/Desktop/데이터사각화지도/도로교통공단_전국_사망교통사고정보(2018)_20190910 (1).csv', 'r')
rdr = csv.reader(f)

cnt = 0
for line in rdr:
    if(cnt == 0):
        cnt+=1
        continue
    #print(line[1][11:13], line[9], line[24], line[25])
    insertResultDic1(results_dic1, line)
    insertResultDic2(results_dic2, line)
    cnt += 1
```

results\_dic1에 각 시간대 지역별 사고수를 사전에 저장해주는 함수 insertResultDic1

results\_dic2에 각 시간대 교통사고가 일어난 위도 경도를 사전에 저장해주는 함수insertResultDic2

```
#각 시간대별 지역에서 일어난 사고수를 사전에 저장
def insertResultDic1(dic, line):
    time = int(line[1][11:13])//4 #4시간 단위로 나눈다
    province = line[9]
    #lng = line[24]
    #lat = line[25]
    if not time in dic:
        dic[time] = {}
    if not province in dic[time]:
        dic[time][province] = 1
    else:
        dic[time][province] += 1
```

```
#각 시간대별 사건이 일어난 위도 경도를 사전에 저장
def insertResultDic2(dic, line):
    time = int(line[1][11:13])//4
    province = line[9]
    lng = line[24]
    lat = line[25]
    tmp = [lat,lng,0,0]
    if not time in dic:
        dic[time] = [tmp]
    else:
        dic[time].append(tmp)
```

## 02

데이터 시각화를 하려면 지역 이름을 각 지역에 정해진 지역 코드로 바꿔야 했기 때문에 정해진 지역코드로 함수를 이용해 바꾸어 주었다

하단 좌측사진은  
results\_dic1을 이용해 각 시간대 지역에서 일어난 사고수를 지역코드에 맞춰 출력한 내용

하단 우측사진은  
results\_dic2을 이용해 각 시간대 교통사고가 일어난 위도 경도를 출력한 내용

```
#지방코드 판단해주는 함수
def defineRegionCode(n):
    if n == "서울":
        return "KR-11"

    elif n == "부산":
        return "KR-26"

    elif n == "대구":
        return "KR-27"

    elif n == "대전":
        return "KR-30"

    elif n == "광주":
        return "KR-29"

    elif n == "인천":
        return "KR-28"

    elif n == "울산":
        return "KR-31"
```

```
0
['KR-11', 57]
['KR-26', 20]
['KR-48', 30]
['KR-41', 85]
['KR-43', 25]
['KR-44', 28]
['KR-45', 18]
['KR-46', 25]
['KR-47', 32]
['KR-49', 15]
```

```
0
[['37.62279351', '127.0195897', 0, 0], ['37.52871253', '127.0357669', 0, 0],
.8800484', 0, 0], ['35.13109241', '129.0512513', 0, 0], ['35.20119156', '12
['35.19724042', '128.9661093', 0, 0], ['35.13215856', '128.9648647', 0, 0],
.0122611', 0, 0], ['36.98636278', '126.9381989', 0, 0], ['37.23617313', '12
['36.97692862', '127.4922812', 0, 0], ['36.75542861', '127.4798679', 0, 0],
.1343649', 0, 0], ['36.58739204', '127.4267114', 0, 0], ['35.94297686', '12
```

## 03

Google chart를 이용해 구현하였다

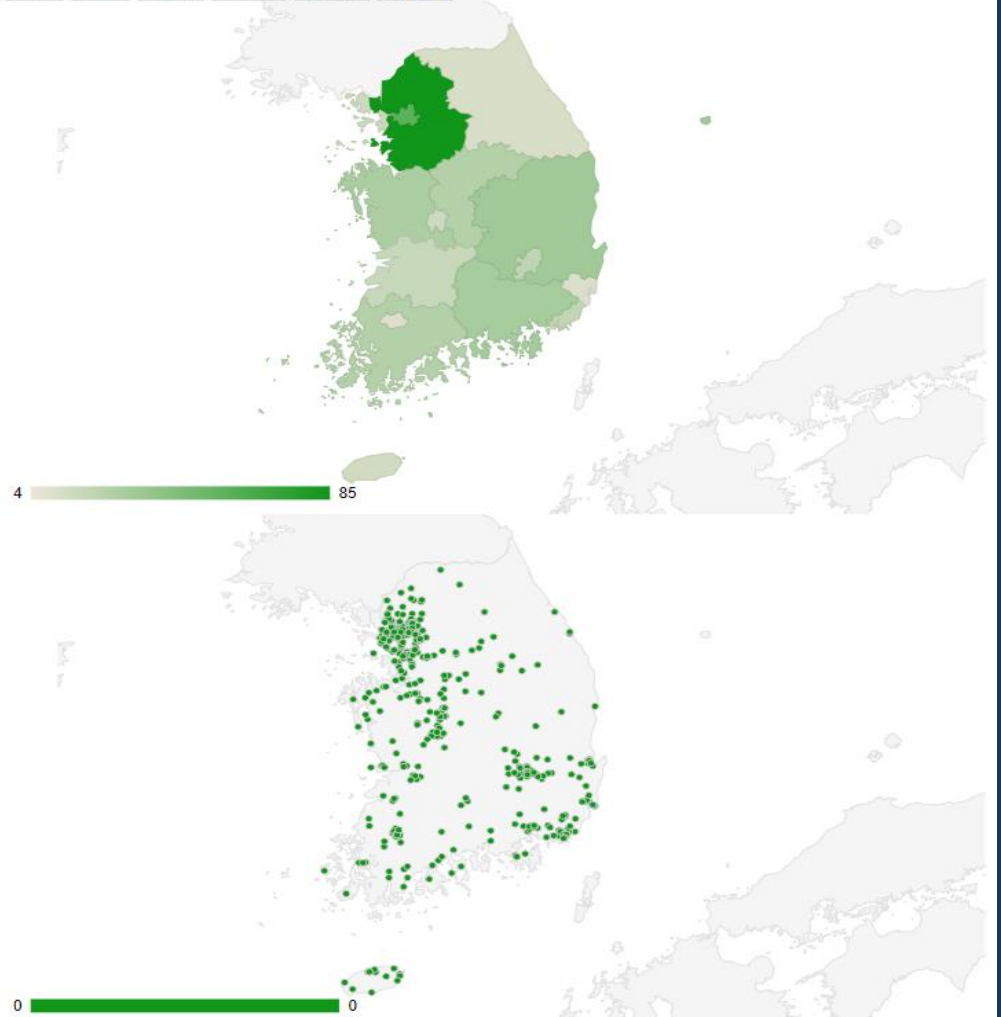
제목 밑 시간대 버튼을 누르면 두가지 그림이 동시에 바뀐다

첫번째 그림에서는 각 지역별로 그 시간대에 일어난 사고의 횟수가 색의 진한 정도로 표현된다

두번째 그림에서는 각 시간대별 교통사고가 일어난 위치가 지도에 점으로 찍힌다

## 시간대별 교통사고 그래프

0~4시 4~8시 8~12시 12~16시 16~20시 20~24시





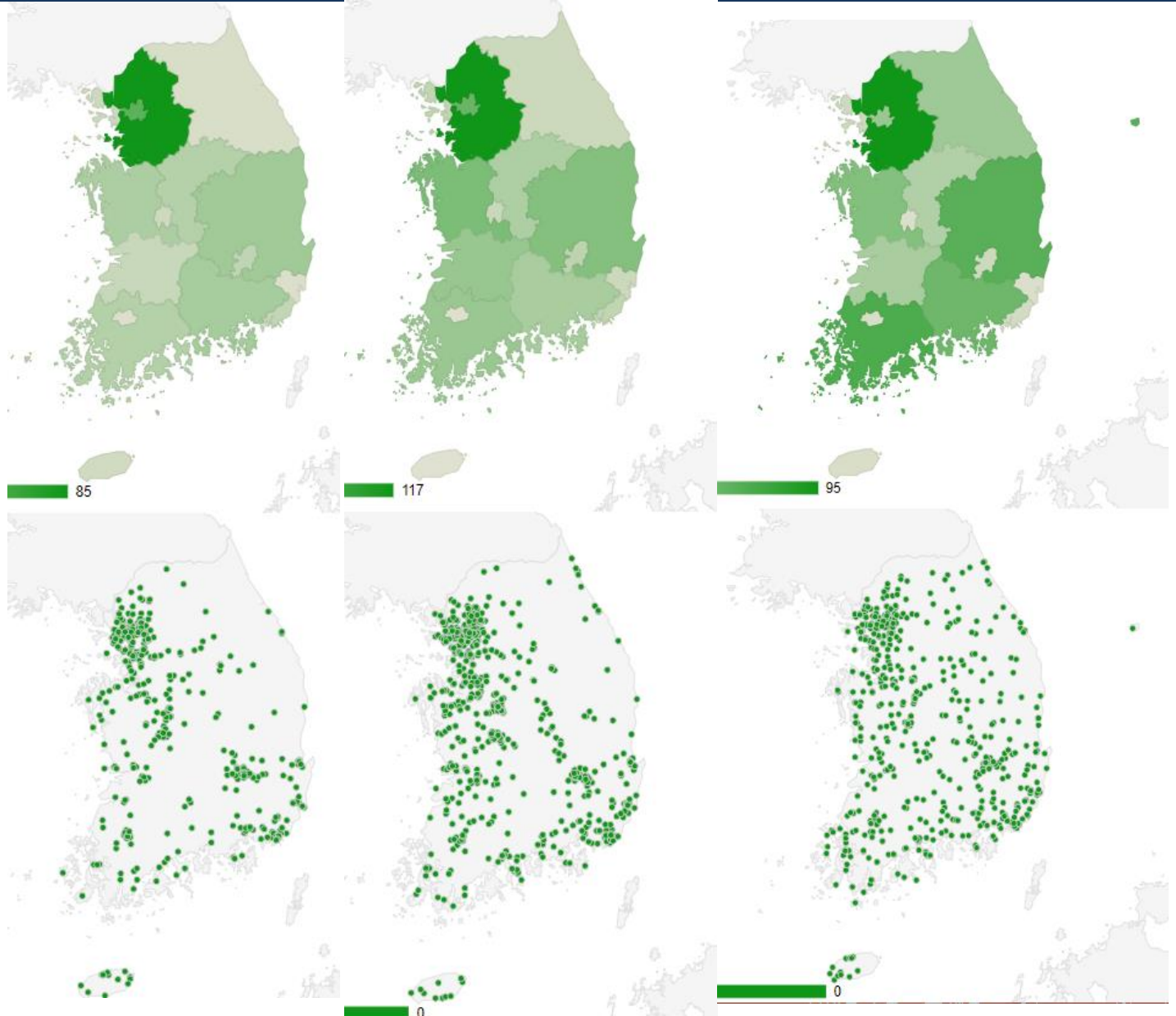
차례대로

0~4시

4~8시

8~12시

그림이다



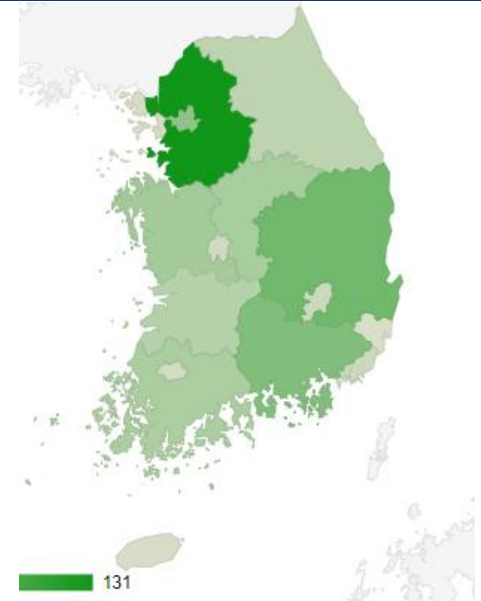
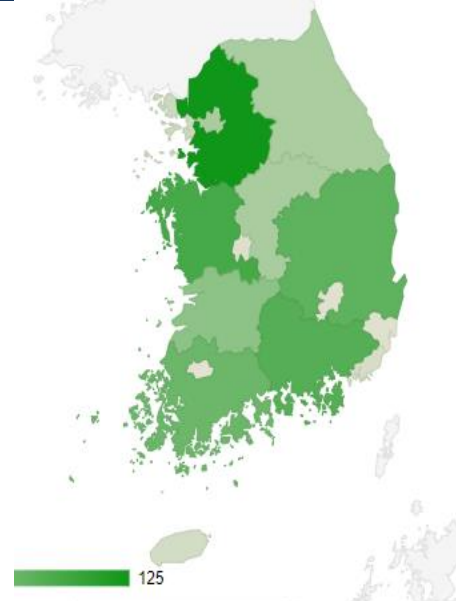
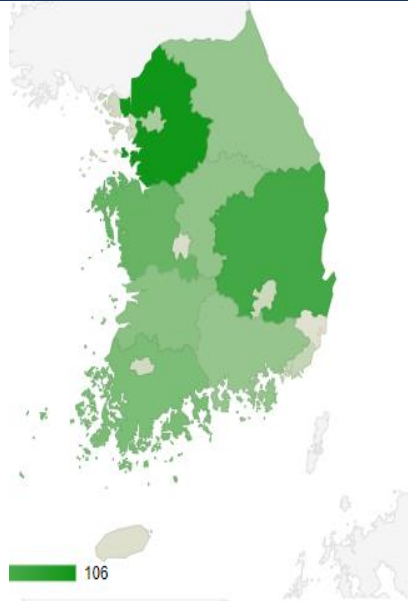
차례대로

12~16시

16~20시

20~24시

그림이다



경기도는 거의 전 시간대에서 가장 많이 사고가 일어나며 예상외로 부산쪽도 교통사고가 경기 만큼은 아니지만 자주 일어나는 것을 볼 수 있었다.

**가설 1.출퇴근시간에는 경기쪽에 많을 것이고 낮에는 고르게 분포되어 있을 것이다.**

**가설 2.교통사고가 많이 일어나는 지역은 서울일것이다.**

가설1 : 출퇴근 시간 뿐만 아니라 거의 모든시간대에 경기와 서울 지역에 교통사고가 몰려있었다. 또 낮에는 비교적 다른 시간대에 비해 고르게 분포되어 있는 모습을 볼 수 있었다.

가설2 : 교통사고가 가장 많이 일어나는 쪽은 경기도이며 그 다음이 서울과 부산쪽이었다.

**THANK  
YOU**