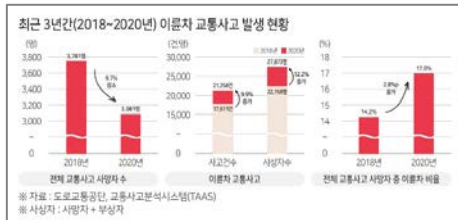


이륜차 교통사고, 멈춰!

1. 분석 배경

코로나19로 인한 **이륜차 배달 서비스** 폭증으로 인해 사고가 증가하고 있음.
사고 다발 지역을 미리 예측하고
 대비하는 방안의 필요성을 느껴 분석 진행함.



2. 분석 내용

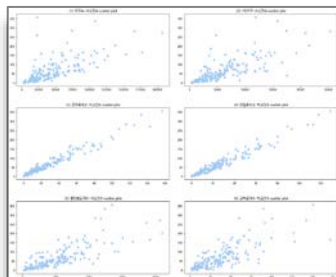
1) 데이터 전처리

시도별 인구 수	인구수	1인가구 수	도로면적	행정동수	행정리수	행정마을수	행정동수	행정리수
1. 경기도	3,113,200	388,000	3,271,242,000	340	130	363	23	340
2. 서울특별시	4,713,100	117,700	2,720,041,200	330	140	910	25	1670
3. 부산광역시	3,411,100	640,000	3,520,000,000	620	600	940	38	1530
4. 대전광역시	1,620,000	320,000	1,481,000	360	290	430	36	790
5. 광주광역시	1,480,000	280,000	1,481,000	320	470	880	43	990

사고건수
 = 강상자수 + 중상자수 + 부상신고자수

- 2020년 서울 이륜차 교통사고 데이터(출처:TAAS)를 활용
- 사고건수에 영향을 미칠 수 있는 변수인 **도로면적, 횡단보도 개수, 교차로 개수, 1인 가구수**를 법정동 기준으로 전처리를 진행

2) 상관관계 확인



- 사고건수와 인구수, 1인가구, 교차로사고건수, 단일로사고건수, 횡단보도개수, 교차로개수의 상관계수는 0.7 이상으로 아주 강한 양의 상관관계를 보임
 → 군집화에 변수로 사용
- 사고건수와 면적, 도로면적은 상관계수가 0.3 이하로 상관관계가 없음



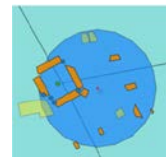
3) K-means Clustering

- 185개의 동을 3개의 군집으로 나눈 결과, 사고 다발 지역으로 **15개 동 선정**



4) 공간분석

- QGIS를 사용해 선정된 행정에 100m*100m의 그리드를 생성하고 50m의 버퍼 내부에 **교차로, 신호등, 횡단보도, 단독주택**을 표현



$$f(k) = \frac{x_{k1} + x_{k2} + x_{k3} + x_{k4}}{\sum_{i=1}^4 (x_{i1} + x_{i2} + x_{i3} + x_{i4})}$$

$f(k)$ = 제k번째 그리드의 정규화값
 x_{ij} = i번째 그리드의 특성j개수
 $(j=1: 교차로, j=2: 신호등, j=3: 횡단보도, j=4: 단독주택)$
 p = 해당 행정동의 그리드 개수

- 데이터를 정규화하여 히트맵으로 표현한 후 **사고 다발 예상 지역 선정**



3. 결론 및 기대효과

- 상계8동, 잠실7동, 구로1동, 신월6동이 사고 다발 지역으로 예측됨
- 예측을 통해 사고 발생 확률이 높은 구역은 자치구 단위에서 관리하여 사고 예방 가능

