

빅데이터분석 텀프로젝트

AI 융합학부

20160348

강민성

INDEX

1. 서론
2. 연구 방법
3. 연구 결과
4. 결론

1. 서론

프로젝트의 목표

- 서울에서 발생한 교통사고 데이터를 수집
- 개별 사고를 제공된 여러 변수로 그룹화
- 사망사고 발생에 영향을 주는 요인 찾기
- 분석한 데이터를 한눈에 알아볼 수 있도록 그래프와 표로 시각화
- 지도 위에 통계 데이터를 시각화



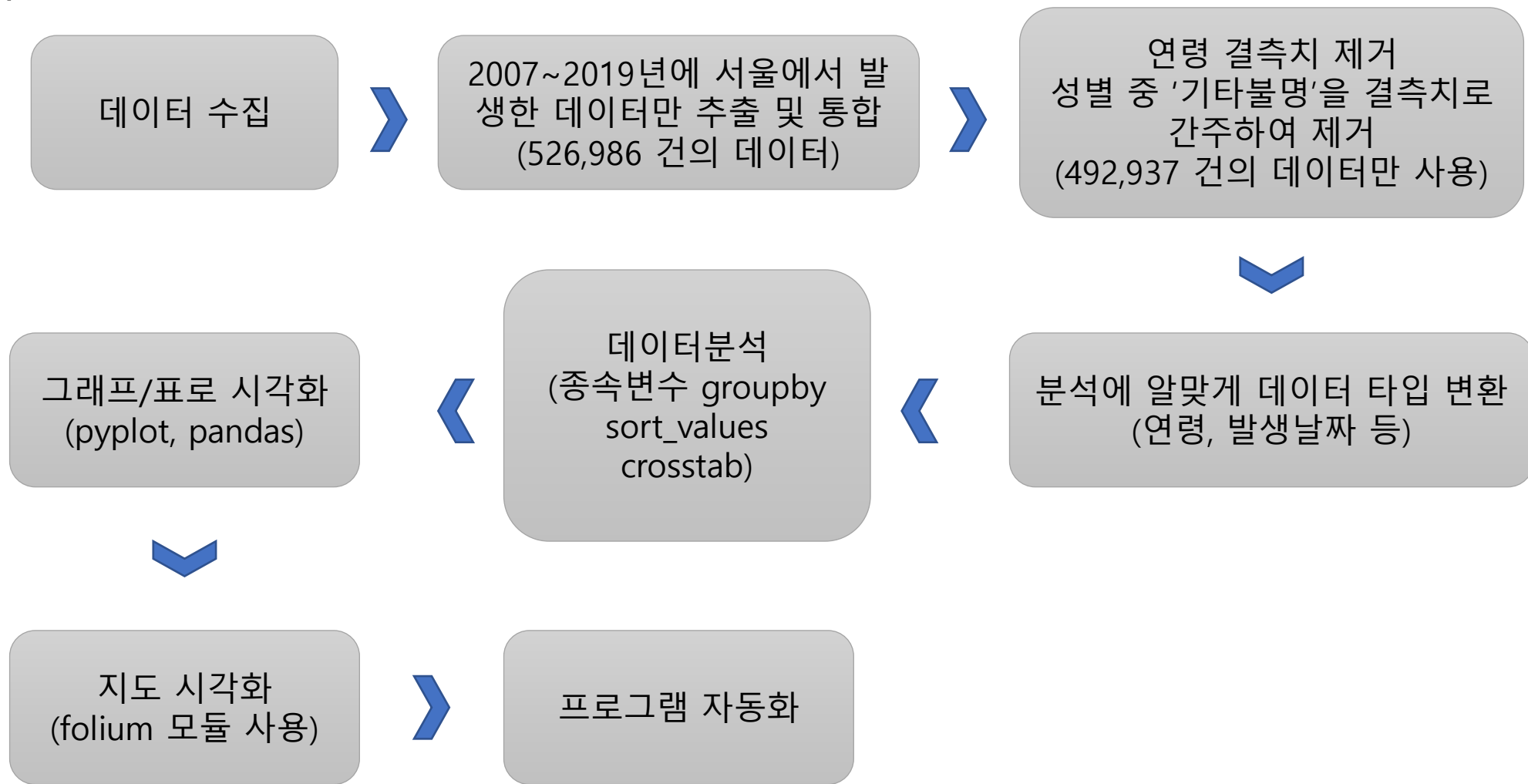
2. 연구 방법

1) 연구 데이터

- 교통사고분석시스템(TAAS)의 공공데이터 셋
 - 2014_19년 서울_경기 개별 교통사고 정보.xlsx (550,792 건)
 - 서울시 개별사고정보(2007_2009).xlsx (125,433 건)
 - 서울시 개별사고정보(2010_2018).xlsx (362,298 건)
- 주요 속성
 - ⇒ 발생일, 발생시간, 요일, 발생지(시군구), 사고내용, 사망/중상/경상/부상자수
사고유형, 가해자법규위반, 노면상태, 기상상태, 도로형태, 피/가해자 차종,
피/가해자 연령 및 성별
- 서울시 열린데이터 광장
 - Seoul_population.txt (2007~2019년 서울시 자치구별 인구 데이터)

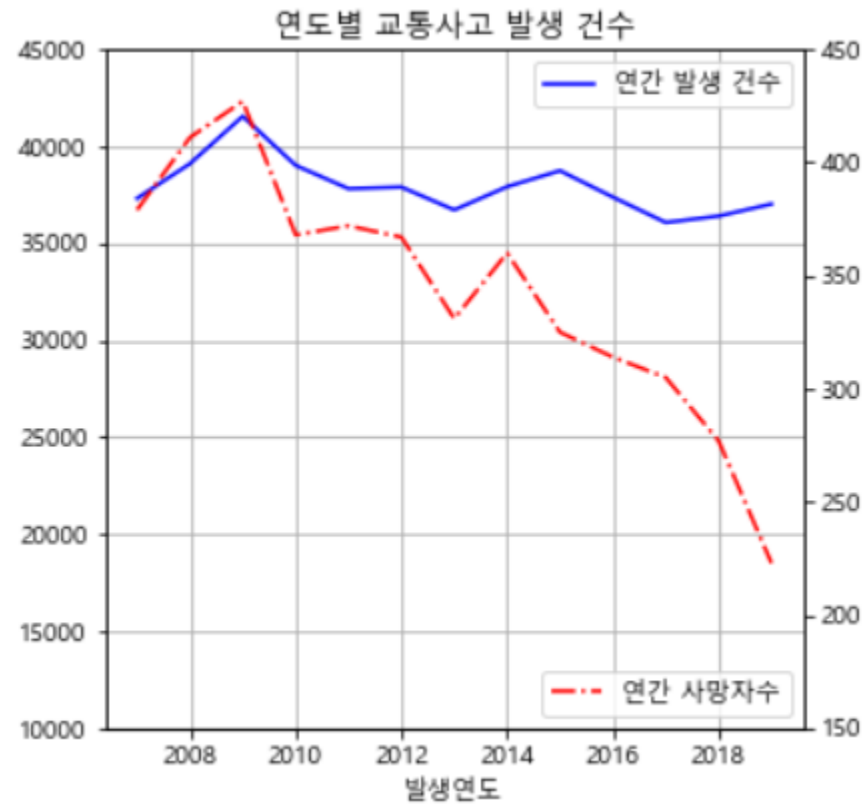
2. 연구 방법

2) 연구 과정



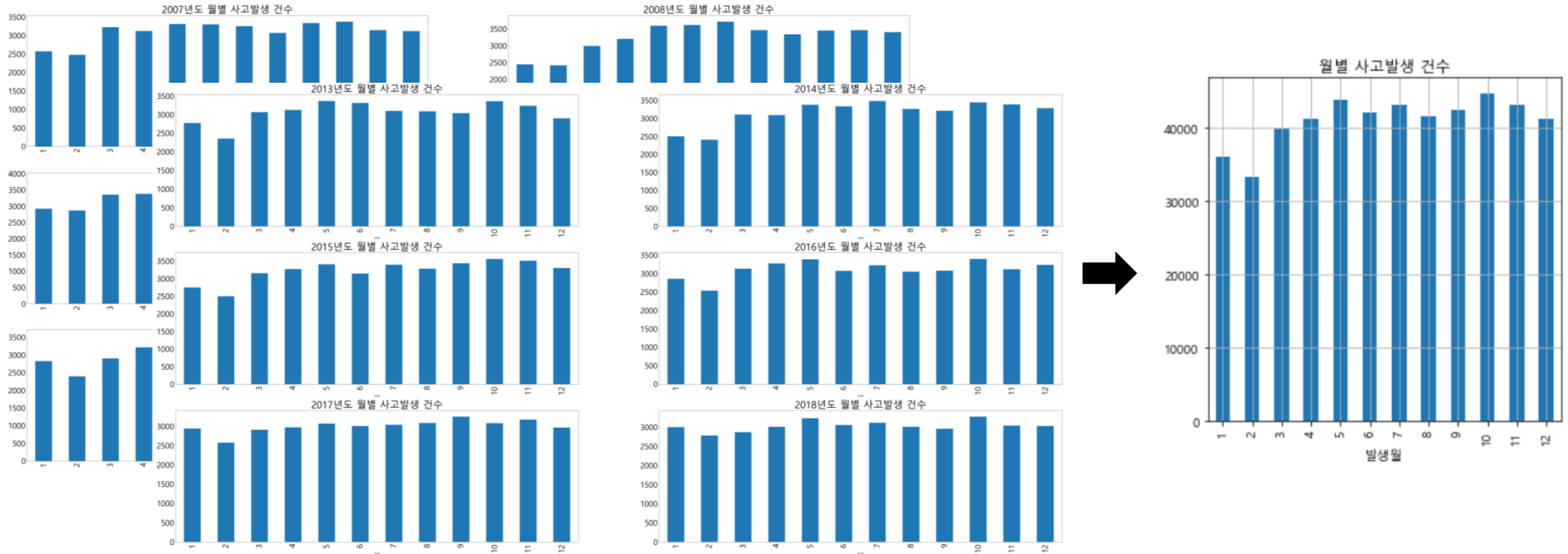
3. 연구 결과

1) 연도별 교통사고 발생 건수



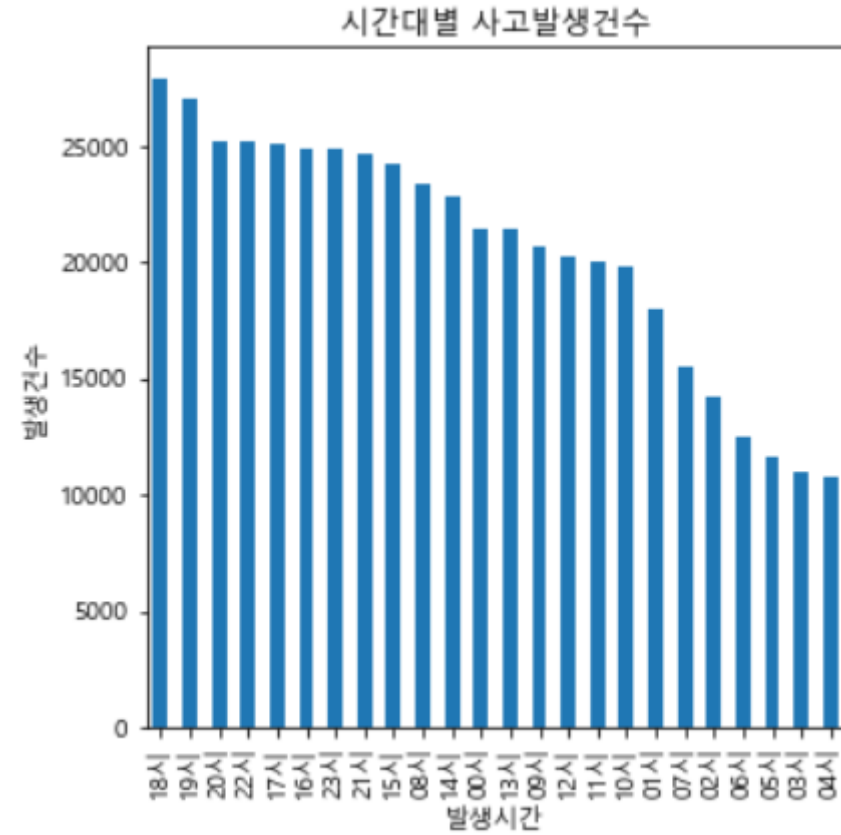
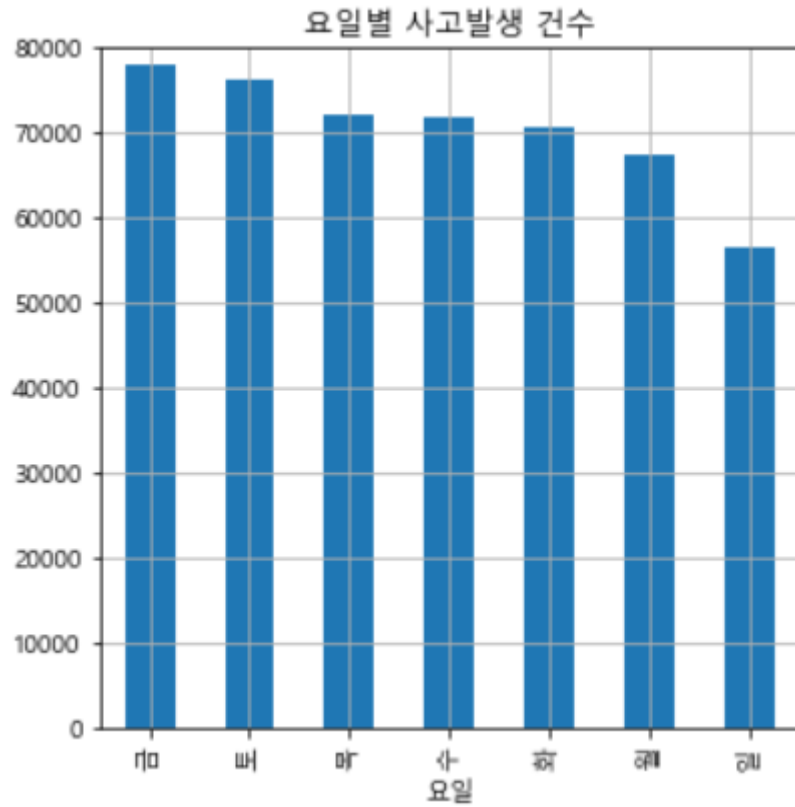
3. 연구 결과

2) 월별 교통사고 발생 건수



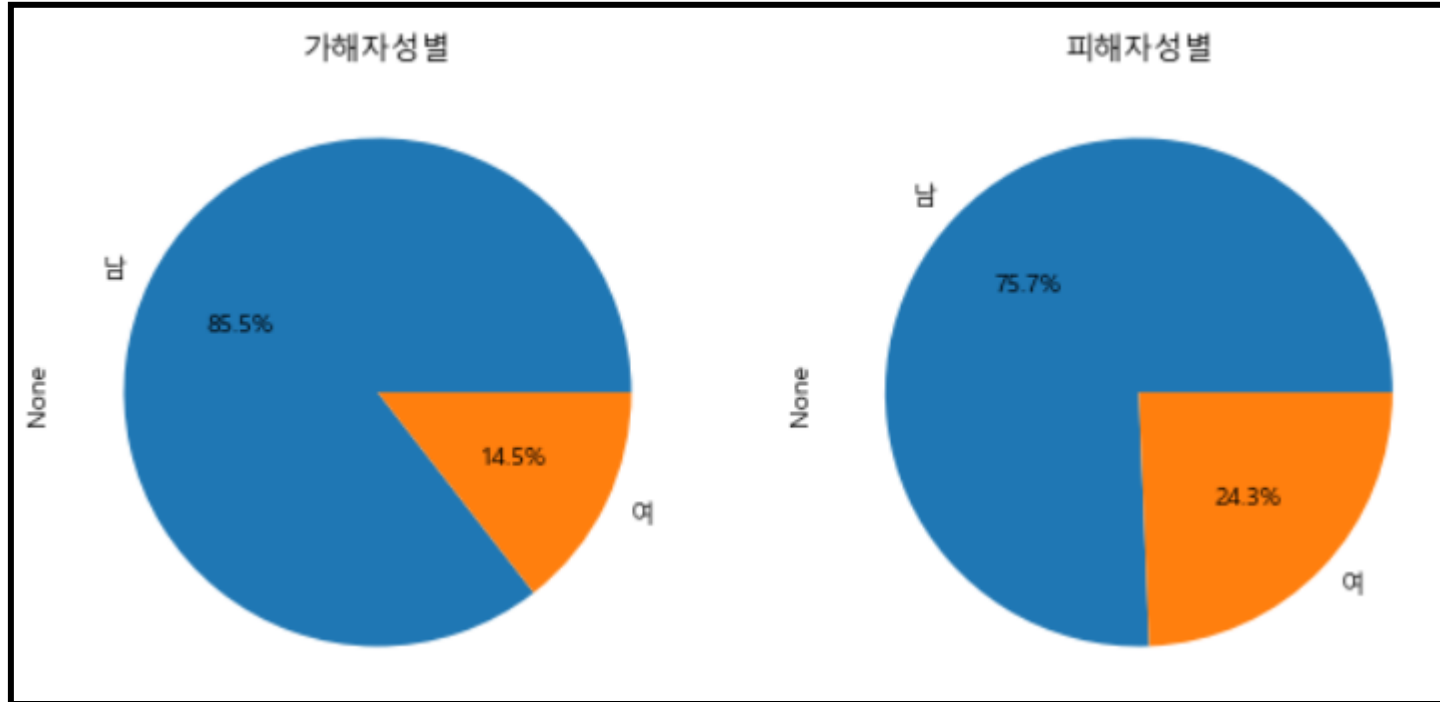
3. 연구 결과

3) 요일/시간대별 교통사고 발생 건수

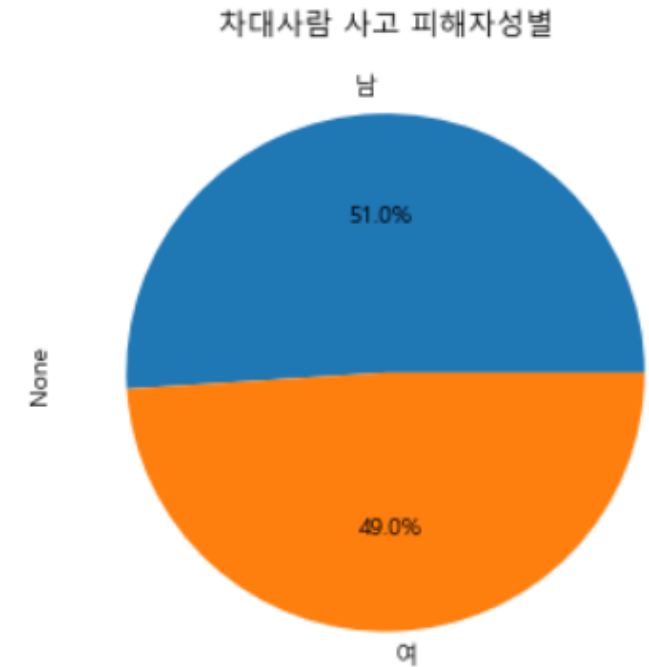


3. 연구 결과

4) 피/가해자 성별 비율



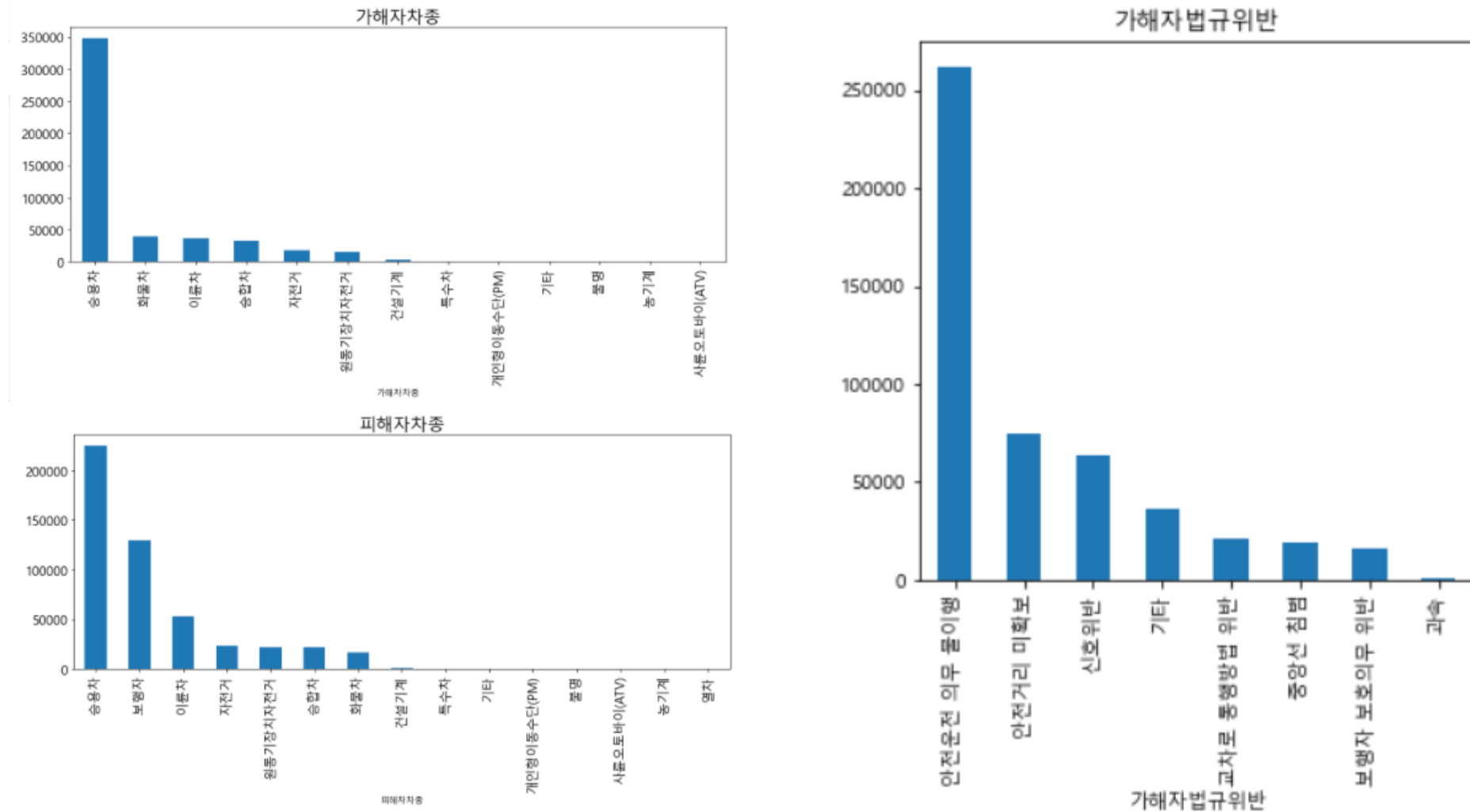
차대차, 차대사람 교통사고 전체에 대한 성비



* 남성 운전자가 많다는 점을 고려하여, 차대차 사고를 제외한 차대사람 교통사고 피해자의 성비를 보면 남녀의 차이가 크게 나지 않음

3. 연구 결과

5) 피/가해자 차종, 가해자 법규위반별 교통사고 발생 건수



3. 연구 결과

6) 각 항목별 사망사고 비율 비교

1. 주야별

사망자 발생 건수 사망자 미발생 건수 사고발생 건수 사망사고 비율(%)

주야

야	2448	239687	242135	1.0110
주	1927	248875	250802	0.7683

p-value: 1.2307750685998935e-19

3. 노면상태별

사망자 발생 건수 사망자 미발생 건수 사고발생 건수 사망사고 비율(%)

노면상태

젖음/습기	613	50547	51160	1.1982
적설	14	1456	1470	0.9524
기타	69	7911	7980	0.8647
건조	3659	425901	429560	0.8518
서리/결빙	20	2739	2759	0.7249
침수	0	5	5	0.0000
해빙	0	3	3	0.0000

p-value: 9.210798378305468e-12

2. 기상상태별

사망자 발생 건수 사망자 미발생 건수 사고발생 건수 사망사고 비율(%)

기상상태

안개	2	118	120	1.6667
흐림	366	26691	27057	1.3527
비	392	35429	35821	1.0943
눈	29	2955	2984	0.9718
맑음	3565	418518	422083	0.8446
기타/불명	21	4851	4872	0.4310

p-value: 3.806708823673974e-21

3. 연구 결과

4. 도로형태별

사망자 발생 건수 사망자 미발생 건수 사고발생 건수 사망사고 비율(%)

도로형태				
불명	2	100	102	1.9608
횡단보도상	388	19758	20146	1.9259
터널안	20	1252	1272	1.5723
교량위	80	5089	5169	1.5477
지하차도(도로)내	16	1037	1053	1.5195
교차로횡단보도내	62	4114	4176	1.4847
횡단보도부근	142	11179	11321	1.2543
고가도로위	10	937	947	1.0560
기타단일로	2198	237803	240001	0.9158
교차로내	835	109724	110559	0.7553
교차로부근	535	81913	82448	0.6489
기타/불명	62	10128	10190	0.6084
기타	25	5525	5550	0.4505
철길건널목	0	3	3	0.0000

p-value: 2.549242475874938e-81

5. 발생시간별

사망자 발생 건수 사망자 미발생 건수 사고발생 건수 사망사고 비율(%)

발생시간				
05시	273	11397	11670	2.3393
04시	234	10508	10742	2.1784
06시	215	12279	12494	1.7208
02시	190	14044	14234	1.3348
03시	134	10845	10979	1.2205
01시	217	17773	17990	1.2062
00시	244	21256	21500	1.1349
16시	159	24748	24907	0.6384
19시	172	26923	27095	0.6348
17시	153	24986	25139	0.6086
18시	163	27748	27911	0.5840
15시	138	24133	24271	0.5686

p-value: 4.1760648315432355e-163

3. 연구 결과

6. 가해자법규위반별

사망자 발생 건수 사망자 미발생 건수 사고발생 건수 사망사고 비율(%)

가해자법규위반

과속	164	540	704	23.2955
중앙선 침범	270	18265	18535	1.4567
보행자 보호의무 위반	207	15705	15912	1.3009
안전운전 의무 불이행	2820	259022	261842	1.0770
신호위반	592	63249	63841	0.9273
기타	163	36015	36178	0.4506
안전거리 미확보	131	74533	74664	0.1755
교차로 통행방법 위반	28	21232	21260	0.1317

p-value: 0.0

7. 가해자차종별

사망자 발생 건수 사망자 미발생 건수 사고발생 건수 사망사고 비율(%)

가해자차종

건설기계	144	2274	2418	5.9553
특수차	33	950	983	3.3571
불명	4	144	148	2.7027
기타	3	147	150	2.0000
승합차	627	31795	32422	1.9339
화물차	636	39122	39758	1.5997
원동기장치자전거	204	15159	15363	1.3279
이륜차	410	36246	36656	1.1185
자전거	169	17095	17264	0.9789
승용차	2144	345377	347521	0.6169
개인형이동수단(PM)	1	199	200	0.5000
농기계	0	28	28	0.0000
사륜오토바이(ATV)	0	26	26	0.0000

p-value: 0.0

3. 연구 결과

8. 피해자차종별

사망자 발생 건수 사망자 미발생 건수 사고발생 건수 사망사고 비율(%)

피해자차종

기타	17	204	221	7.6923
건설기계	32	738	770	4.1558
농기계	1	24	25	4.0000
특수차	13	480	493	2.6369
보행자	2729	126595	129324	2.1102
불명	2	107	109	1.8349
화물차	199	16545	16744	1.1885
승합차	204	21592	21796	0.9360
자전거	191	23397	23588	0.8097
이륜차	266	52816	53082	0.5011
원동기장치자전거	105	21865	21970	0.4779
승용차	616	224007	224623	0.2742
개인형이동수단(PM)	0	156	156	0.0000
사륜오토바이(ATV)	0	33	33	0.0000
열차	0	3	3	0.0000

p-value: 0.0

9. 피해자연령대별

사망자 발생 건수 사망자 미발생 건수 사고발생 건수 사망사고 비율(%)

피해자연령대

90대	42	414	456	9.2105
80대	349	4130	4479	7.7919
70대	680	18975	19655	3.4597
60대	755	55358	56113	1.3455
10세 미만	66	7529	7595	0.8690
50대	882	103547	104429	0.8446
40대	718	102800	103518	0.6936
20대	358	70682	71040	0.5039
30대	417	96429	96846	0.4306
10대	108	28693	28801	0.3750
100세 이상	0	5	5	0.0000

p-value: 0.0

3. 연구 결과

10. 사고유형(대분류)별

사망자 발생 건수 사망자 미발생 건수 사고발생 건수 사망사고 비율(%)

사고유형_대분류

차대사람	2729	126593	129322	2.1102
차대차	1646	361962	363608	0.4527
차량단독	0	4	4	0.0000
철길건널목	0	3	3	0.0000

p-value: 0.0

11. 사고유형별

사망자 발생 건수 사망자 미발생 건수 사고발생 건수 사망사고 비율(%)

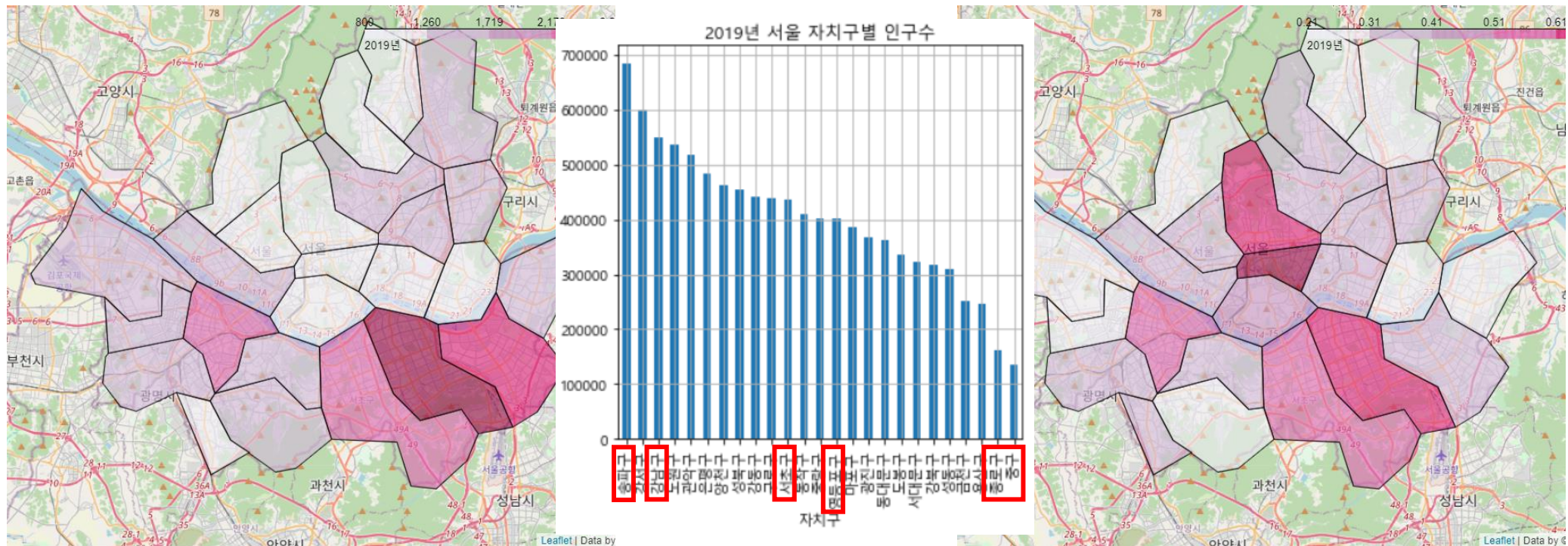
사고유형

횡단중	1523	51835	53358	2.8543
차도통행중	336	13097	13433	2.5013
정면충돌	228	16125	16353	1.3942
보도통행중	112	8074	8186	1.3682
기타	1028	114208	115236	0.8921
길가장자리구역통행중	73	9569	9642	0.7571
측면충돌	747	164383	165130	0.4524
진행중 추돌	172	55367	55539	0.3097
주정차중 추돌	109	35570	35679	0.3055
추돌	47	19128	19175	0.2451
도로이탈 추락	0	1	1	0.0000
주/정차차량 충돌	0	3	3	0.0000
직전진행	0	1	1	0.0000
차단기돌파	0	1	1	0.0000
후진중충돌	0	1200	1200	0.0000

p-value: 0.0

3. 연구 결과

7) 지도에 교통사고 통계 데이터 시각화



2019년 자치구별 교통사고 발생건수

2019년 자치구별 인구수 대비
교통사고 발생 비율

4. 결론

1) 각 항목별로 사망사고 비율을 비교하였을 때,

- 주간보다 야간에 발생한 사고에서 사망자 발생 확률이 높음
- 기상상태가 안개/흐림/비일 때 사망자 발생 확률이 높음
- 도로상태가 젖음/습기 또는 적설일 때 사망자 발생 확률이 높음
- 도로 종류가 횡단보도, 터널 안, 교량 위, 지하차도일 때 사망자 발생 확률이 높음
- 새벽 5시, 4시에 사망자 발생 확률이 가장 높고, 18시, 15시일 때 가장 낮음
- 가해자가 과속하여 사고가 발생하였을 때, 사망자 발생 확률이 높음
- 가해자 차종이 건설기계, 특수목적 차량일 때 사망자 발생 확률이 높음
- 사고 피해자가 고령자일수록 사망 확률이 높음

2) 지도에서 자치구별 사고 통계를 보았을 때,

- 강남3구(서초, 강남, 송파)와 영등포구의 사고발생 건수가 많음
- 종로구와 중구는 인구 대비 사고발생률이 높으며, 2018년 이후에는 서초구, 강남구, 영등포구의 인구 대비 사고발생률도 높아지는 추세를 보임

감사합니다