

## Gun Violence data

Team 1 | 국승지, 박효선, 이지인, 임아영, 조주영

**DATA1**      **Main dataset**      13,447KB      173,739 x 15

1	<b>incident_id</b>	int	primary key, 한 사건 당
2	year	int	연도 15,16,17
3	month	int	월
4	day	int	일
5	weekdays	factor	요일
6	<b>state</b>	factor	state
7	<b>city_or_country</b>	char	
8	n_killed	int	
9	n_injured	int	
10	latitude	num	
11	longitude	num	
12	state_house_distirct	int	Voting house district
13	state_senate_distirct	int	Territorial district from which a senator to legisture is elected
14	congressional_district	int	Congressional district

**DATA2**      **Participant dataset**      18,081KB      274,870 x 8

1	<b>incident_id</b>	int	key
2	<b>pct_id</b>	int	key, participant id
3	age	int	age of participant at the time of crime
4	age_group	char	Adult (18+) , Child (0-11), Teen (12-17)
5	gender	char	Female, Male
6	status	char	killed, injured, arrest, unharmed
7	type	char	Subject - Suspect, victom
8	Arrested	char	Subject - Suspect, victom

**DATA3**      **data2에서 뽑은columnr**      9,657KB      154,973 x 24

1	<b>incident_id</b>	int	primary key, 한 사건 당
2	n_participant	int	연루된 사람이 총 몇 명인지
3	n_Adult	int	어른의 수, 18세 이상
4	n_Child	int	어린이의 수, 0세 - 11세
5	n_Teen	int	청소년의 수, 12세 - 17세
6	n_NA_age_group	int	age_group에서 NA 값인 사람의 수 (age group 모르는 사람 수)
7	n_Female	int	여성의 수
8	n_Male	int	남성의 수
9	n_NA_gender	int	gender 에서 NA 값인 사람의 수 (gender 모르는 사람 수)
10	n_Injured	int	부상자 수
11	n_Killed	int	사망자 수
12	n_Unharmed	int	무사한 사람 수
13	n_NA_status	int	status 모르는 사람 수
14	n_Suspect	int	가해자 수
15	n_Victim	int	피해자 수
16	n_NA_type	int	type 모르는 사람 수

17	n_Arrest_FALSE	int	체포안된 사람 수 (victim 일 수도 있고, suspect 일 수 도 있음)
18	n_Arrest_TRUE	int	체포된 사람 수
19	n_Victim_Killed	int	피해자 중 사망자 수
20	n_Victim_Injured	int	피해자 중 부상자 수
21	n_Victim_Unharmed	int	피해자 중 무사한 사람 수
22	n_Suspect_Killed	int	용의자 중 사망자 수
23	n_Suspect_Injured	int	용의자 중 부상자 수
24	n_Suspect_Unharmed	int	용의자 중 무사한 사람 수

DATA4	text dataset	29,079KB	444,684 x 4
1	incident_id	int	primary key, 한 사건 당
2	state	factor	
3	city_or_country	char	
4	incident_characteristics	char	Characteristics of the incidence