



WORLD **TERROR** **ANALYSIS**

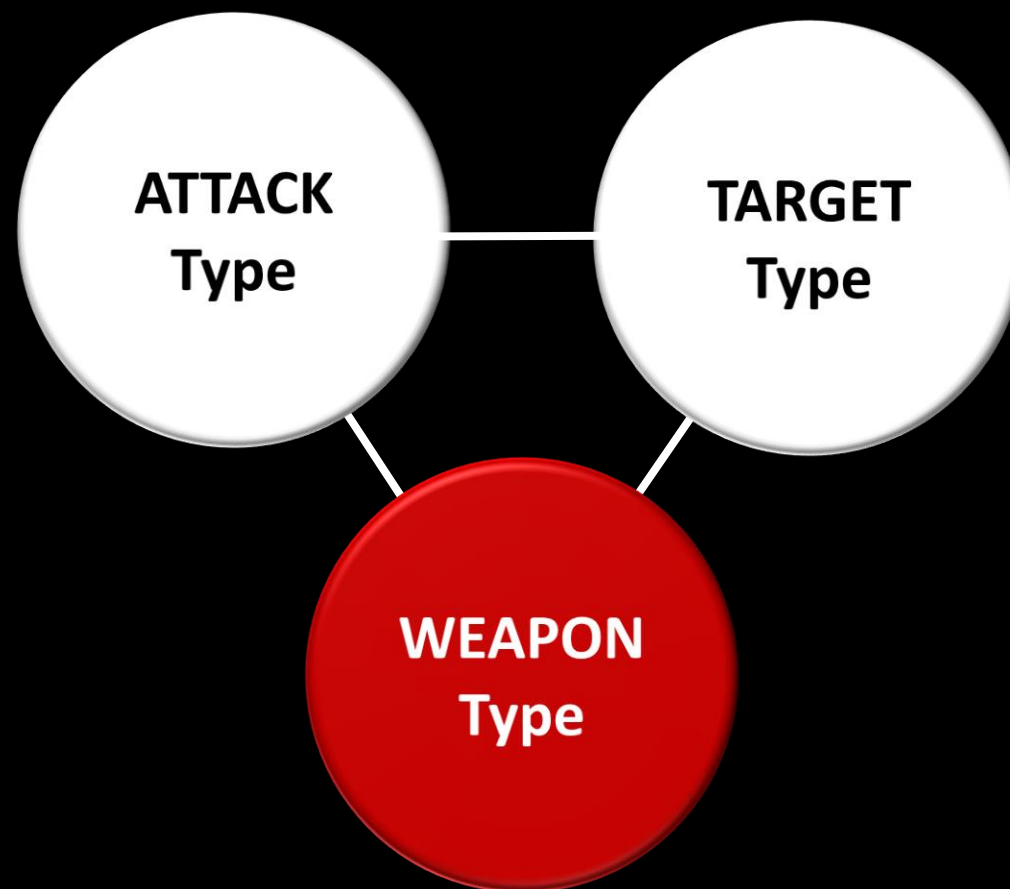
1970 - 2017

DATASCIENCE FINAL PROJECT

팀장 : 정재나 팀원 : 최수지, 이윤희, 박지영

목차

- 데이터 전처리
- 차원, 측정값
- 전 세계 테러 수 및 테러의 비율 관찰
- 주요테러그룹 그룹화
- 유명3 그룹의 Type(타겟, 무기) 분석
- 세부분석1 - 테러 집단별
- 세부분석2 - OECD와 NoOECD 와 테러 단체 연결
- 미국인 중심



데이터 전처리

- 필요없는 컬럼 삭제
- attacktype, gname, weapontype, woundus, dbsource 등 에서 값이 숫자이거나 널값인 레코드 삭제
- Lat, Lon, 합체_100km, 500km, 2500km 열 추가

차원

- 12continent : 전세계 대륙을 12개 지역으로 나눈 것
- targtype : 타겟/피해자의 큰 카테고리
- weaptype1_txt : 테러에 사용된 큰 무기 카테고리
- attacktype : 테러의 공격형태
- country : 사건이 발생한 국가 또는 위치
- city : 사건이 발생한 도시, 마을 또는 마을의 이름
- dbsource : 데이터를 수집한 기관
- gname : 공격한 단체의 이름
- nationality : 공격받은 타겟의 국적
- ID : 테러의 고유한 번호

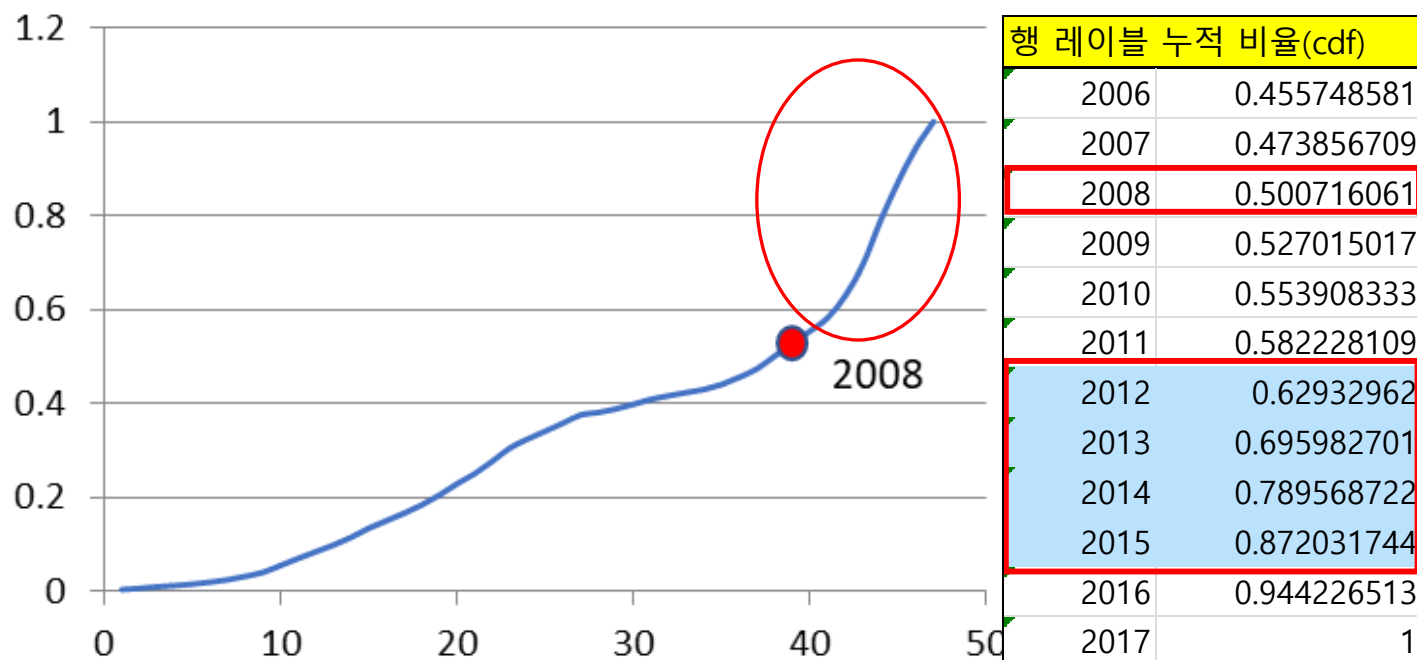
측정값

- latitude : 사건이 발생한 위도
- longitude : 사건이 발생한 경도
- kill : 사건의 확인된 총 사망자의 수
- killus : 사건으로 숨진 미국 시민 수
- wound : 가해자와 피해자에게 확인된 부상자 수
- woundus : 가해자와 피해자에게 확인된 미국인 부상자 수
- Lon,Lat_100, 500, 2500 : 100km, 500km, 2500km로 공간분할한 위도, 경도

전 세계 테러 수 및 테러의 비율 관찰



테러 발생 건 CDF (1970-2017)

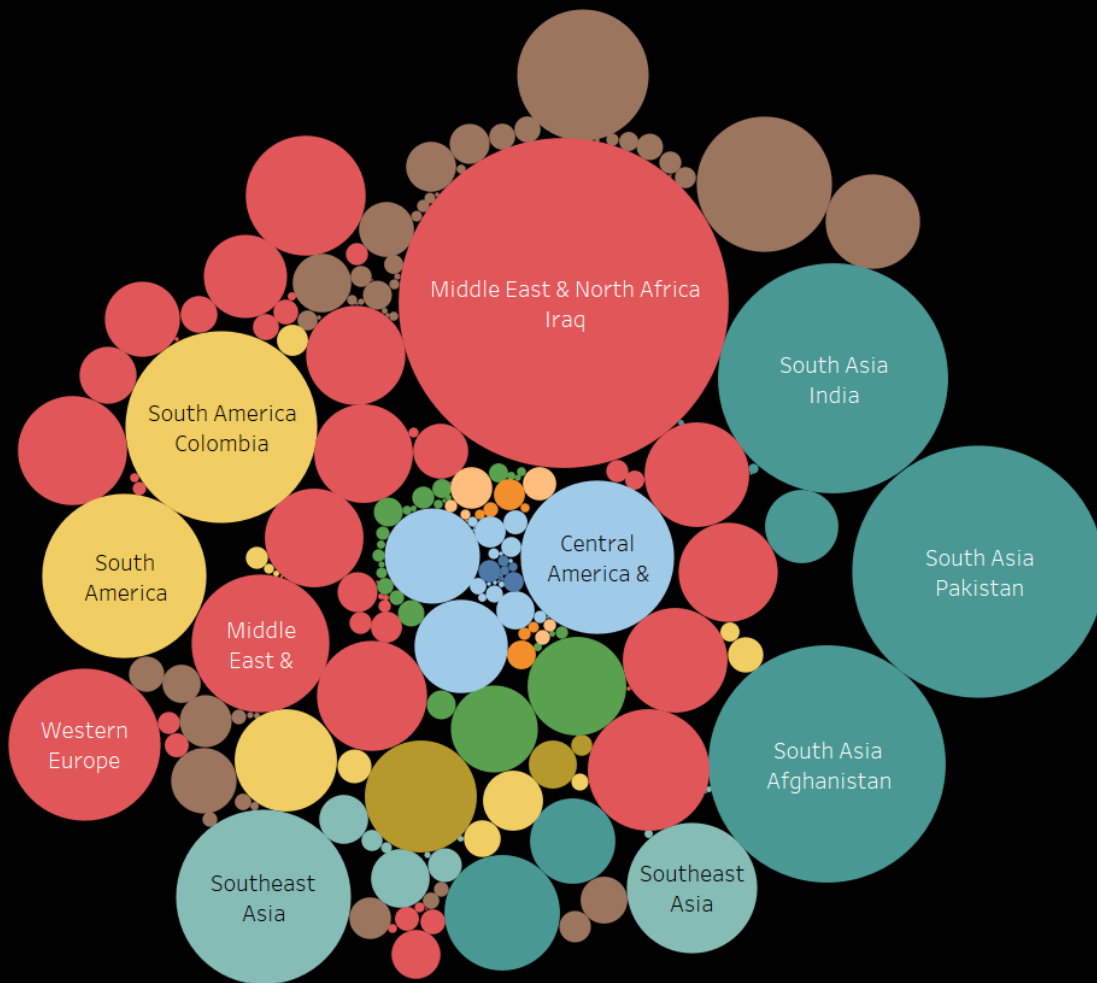


전세계 테러 건수 CDF(1970-2017)

전체 테러의 50% 이상이 2008년 이후에 발생했다는 것을 알 수 있다.

특히, 2012-2015년 사이에 급증한 것을 알 수 있다.

전체 국가별 시각화[버블]



Region_txt 및 country_txt. 색상이 region_txt의 세부 정보를 표시합니다. 크기에서 레코드 수의 합계(를) 표시합니다. 마크에 레이블이 region_txt 및 country_txt에 의해 지정되었습니다.

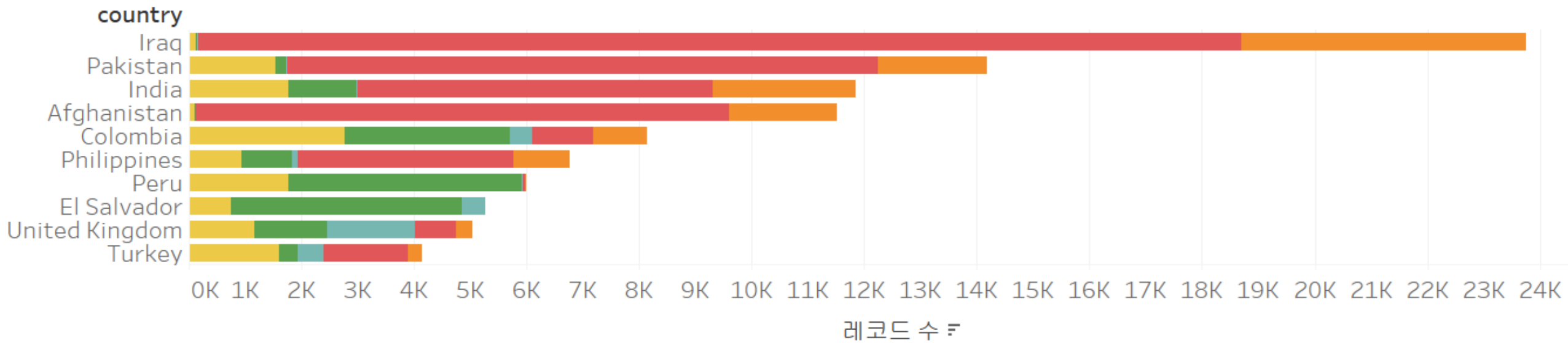
국가를 포함한 열 두 지역별 테러 수 버블 시각화

합계 (레코드수)를
region_txt의 색깔별로 나타냈다.

각 국가가 12지역의 어디에
속하면서 어느정도 크기인지
알 수 있다.

12지역으로 그룹화 된 나라 중
Irap, Afghanistan, Pakistan이
가장 테러 발생 수가 많다.

년도별 나라 테러 빈도 상위 10

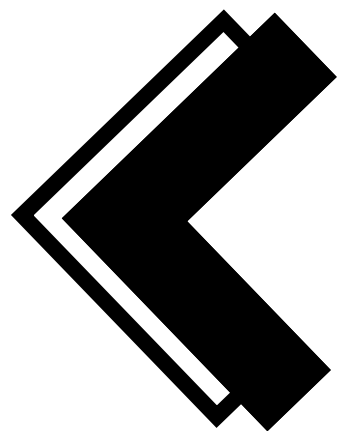


날짜 (그룹)

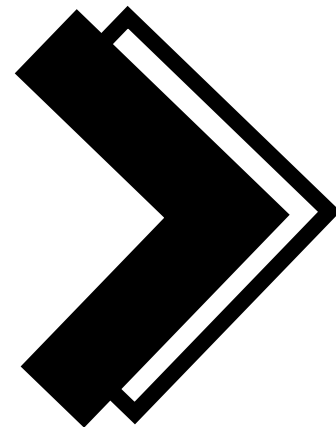
- 00s
- 10s
- 70s
- 80s
- 90s

년도별 나라 테러 빈도 상위 10개국

이라크, 파키스탄, 인도 순으로 테러가 많이 일어난다.
2010년대에 가장 테러가 많이 발생하였다.



주요 테러 그룹 그룹화



테러 그룹 (gname) 그룹화

- gname 필드의 taliban과 Al-qaida의 세부 그룹으로 나누어져 있어 Taliban, Al-qaida의 큰 그룹으로 묶음.

Taliban	
<input checked="" type="checkbox"/>	Punjabi Taliban
<input checked="" type="checkbox"/>	Taliban
<input checked="" type="checkbox"/>	Taliban (Pakistan)
<input checked="" type="checkbox"/>	Tehrik-e-Taliban Islami (TTI)
<input checked="" type="checkbox"/>	Tehrik-i-Taliban Pakistan (TTP)

Al-Qaida	
<input checked="" type="checkbox"/>	Al-Qaida
<input checked="" type="checkbox"/>	Al-Qaida in Iraq
<input checked="" type="checkbox"/>	Al-Qaida in Lebanon
<input checked="" type="checkbox"/>	Al-Qaida in Saudi Arabia
<input checked="" type="checkbox"/>	Al-Qaida in the Arabian Peninsula (AQAP)
<input checked="" type="checkbox"/>	Al-Qaida in the Indian Subcontinent
<input checked="" type="checkbox"/>	Al-Qaida in the Islamic Maghreb (AQIM)

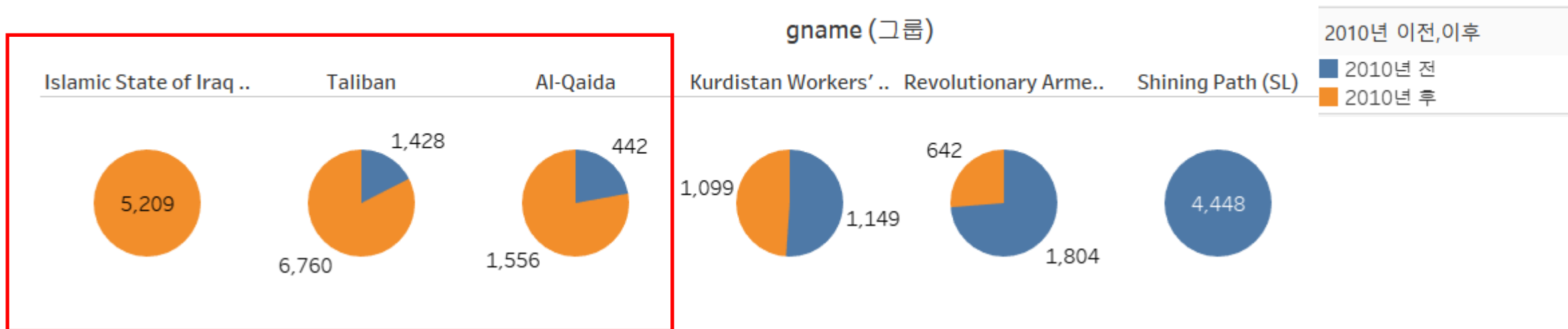
테러 그룹의 세력 분석(신흥, 도태 세력)

테러 그룹이 신흥, 도태 세력인지 알아보기 위해 2010년 기준으로 나눠 gname top11에서 6그룹 추출

2010년 기준인 이유

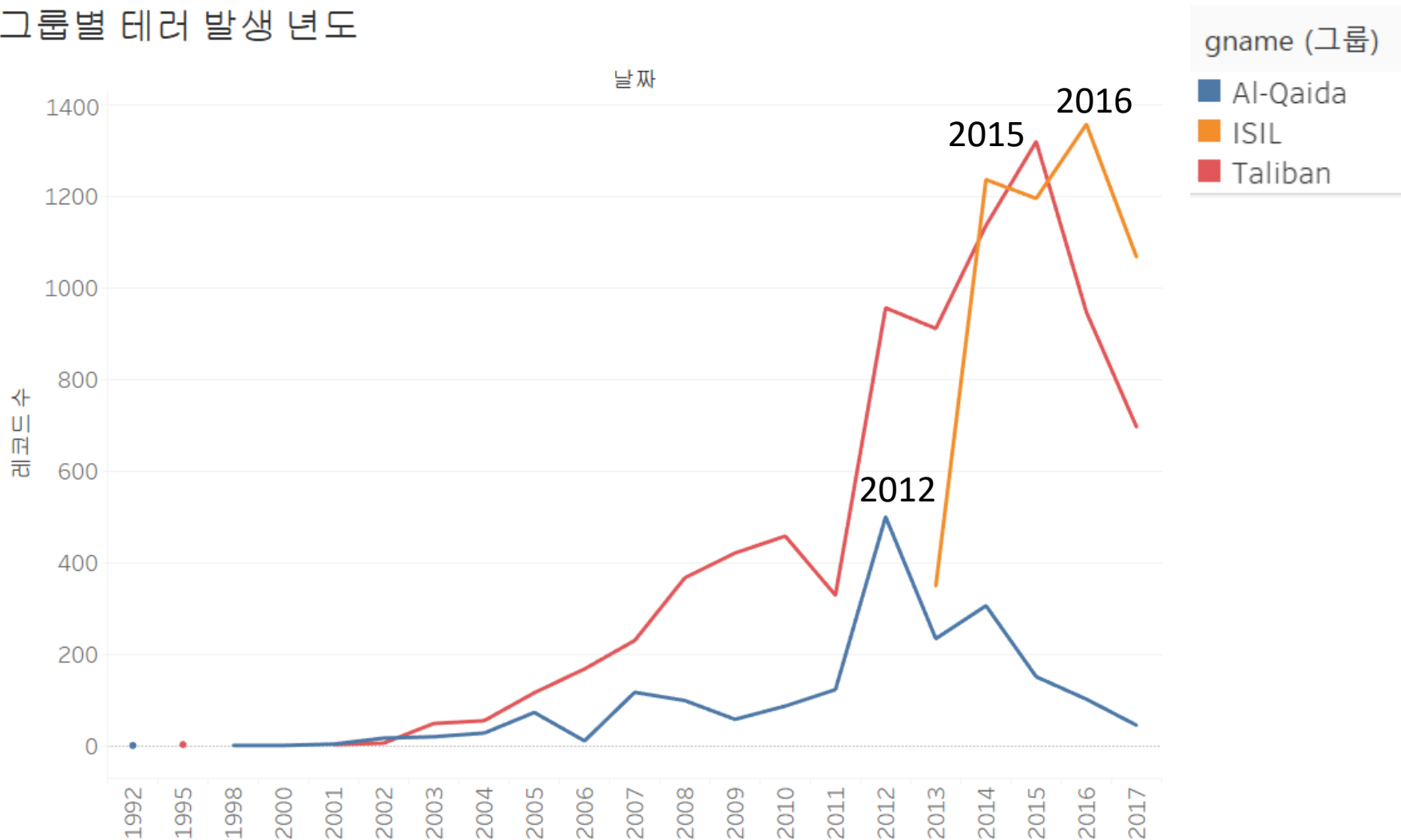
2010년쯤부터 전보다 테러수가 2배이상 급증하기 때문에 신흥 세력의 영향이 많을 거라고 예상

gname의 Top11 신흥,옛날세력



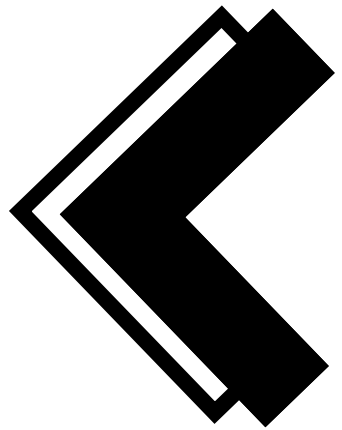
유명 3팀 : Taliban, ISIL, Al-Qaida으로 지정함

그룹별 테러 발생 년도

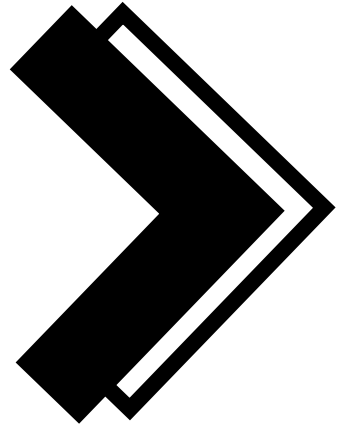


그룹별 테러 발생 년도

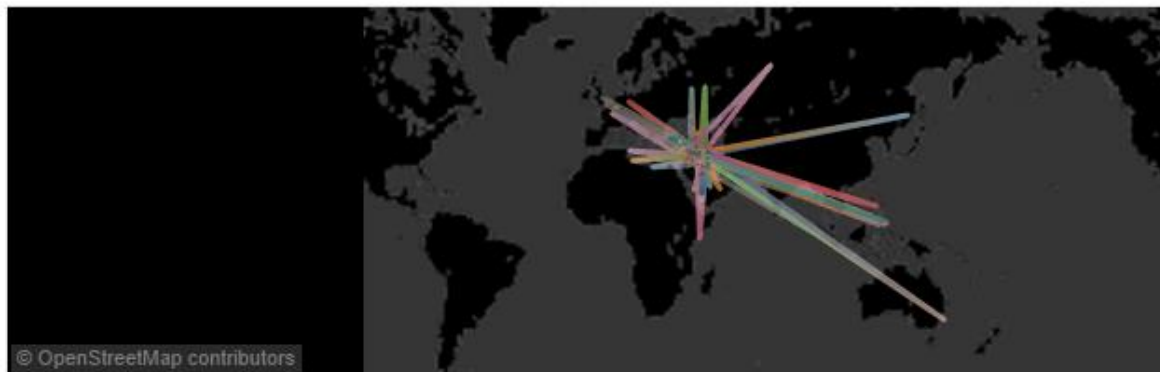
- Al-Qaida : 1992년에 그룹이 만들어졌고, 2012년에 테러가 가장 많이 발생하였다.
- ISIL : 2013년에 만들어졌고, 2016년에 테러가 가장 많이 발생하였다.
- Taliban : 1995년에 만들어졌고, 2015년에 테러가 가장 많이 발생하였다.



유명 3의 Type(타겟, 무기) 분석



ISIL 활동범위



Al-Qaida 활동범위

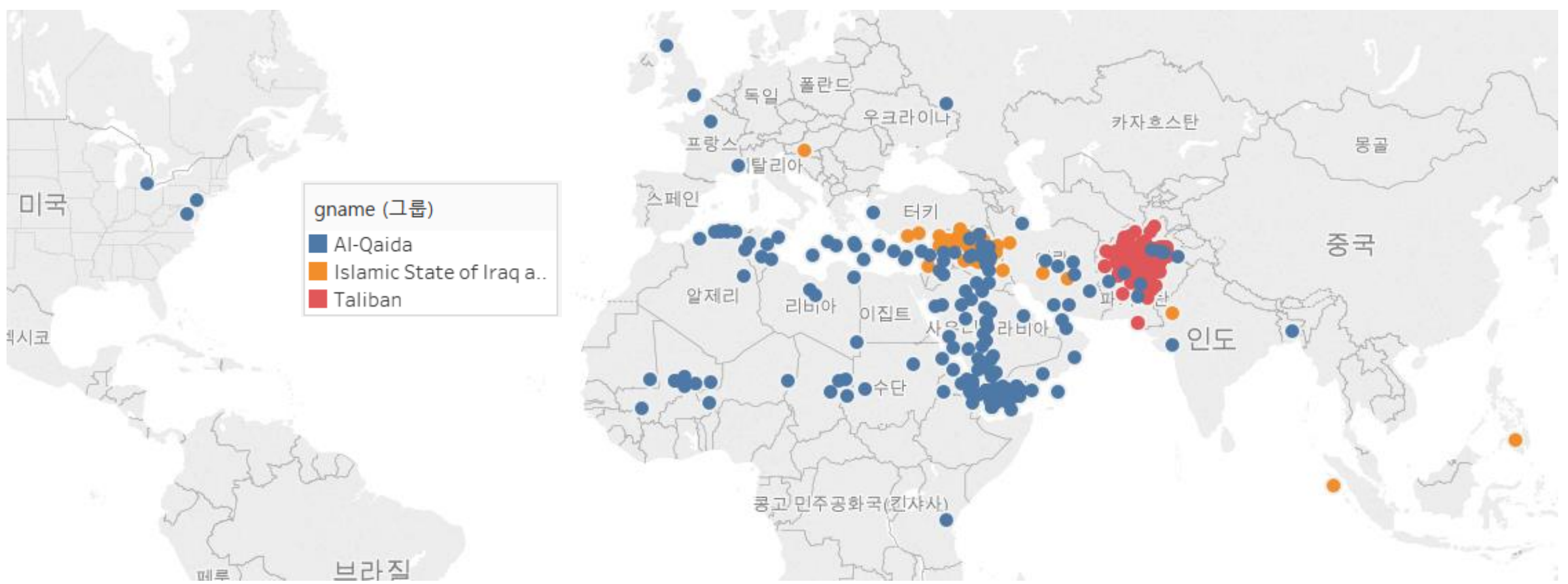


Taliban 활동범위



유명3팀의 활동 범위

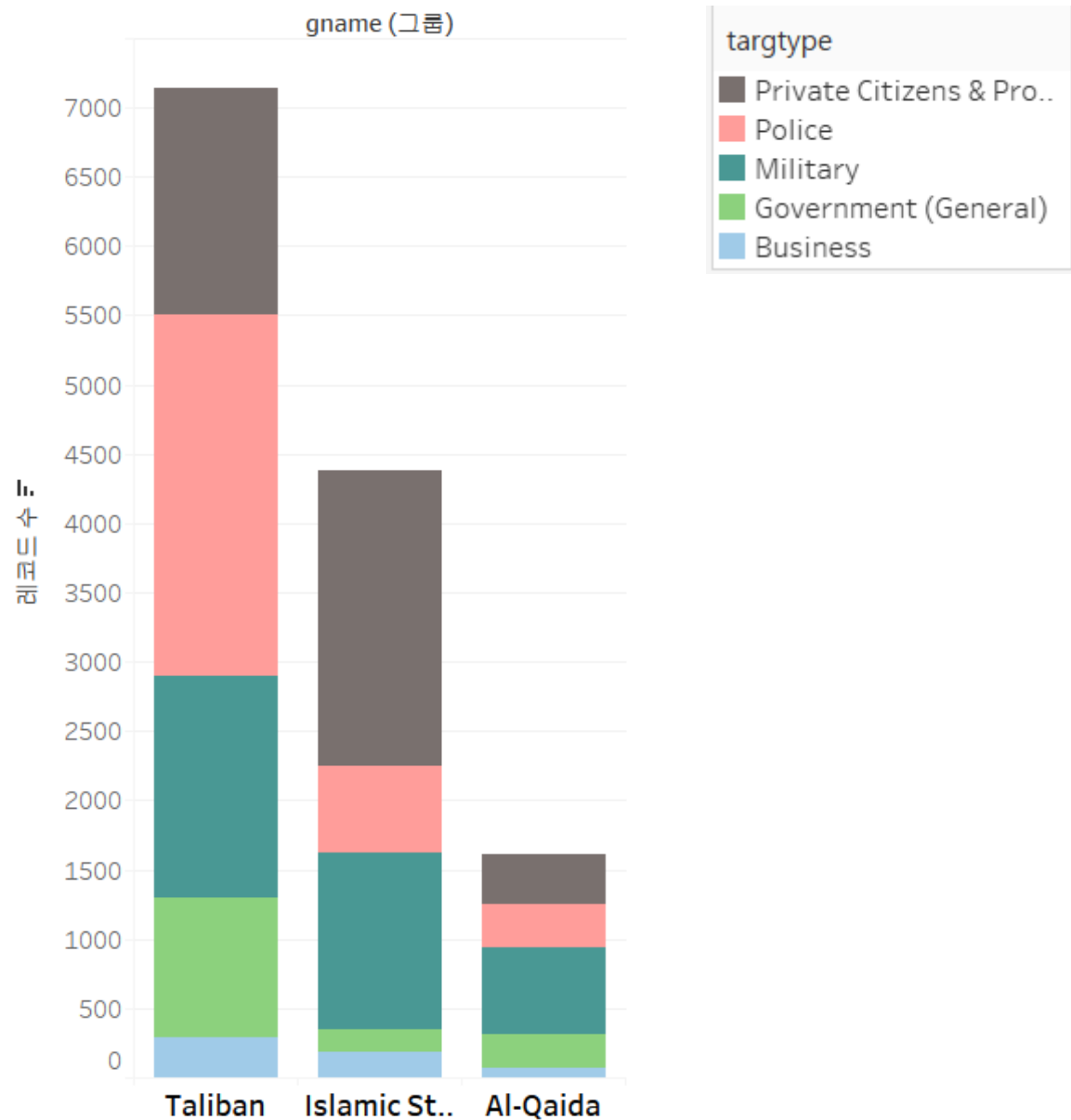
- ISIL은 이라크에서 출발하여, 여러 나라에서 활동하는 경향이 있다.
- Al-Qaida는 제일 넓게 활동 한다.
- Taliban은 주로 아프가니스탄에서만 활동한다.



유명3팀의 타겟 분포

Is, 탈리반은 이라크, 아프가니스탄에 모여있는 반면 알카이다는 여러 나라에 퍼져있는 형태이다.

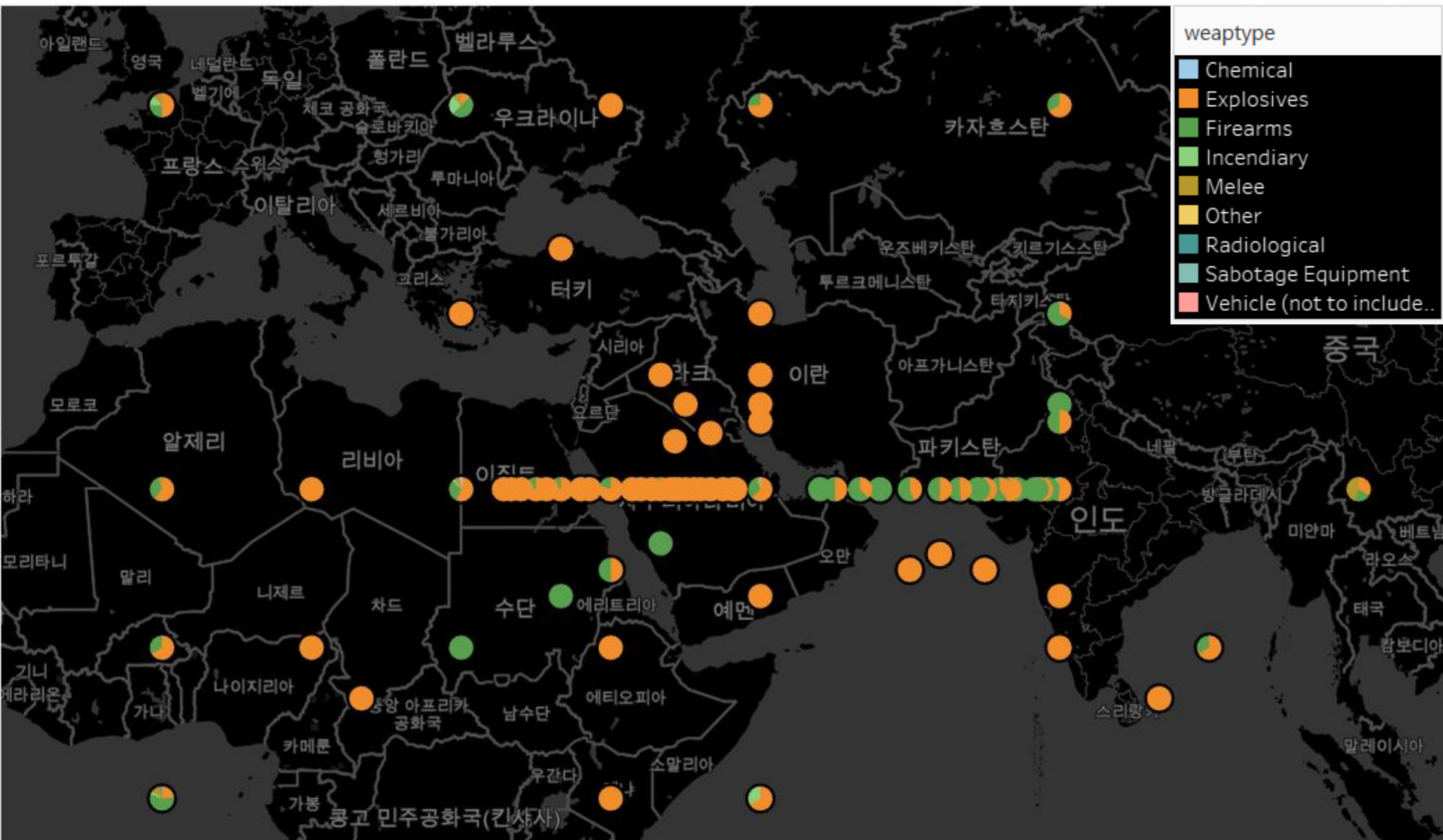
유명3의 타겟



유명3팀의 타겟 타입

유명 3팀은 타겟타입에서도 차이를 보인다.
탈리반과 알카이다는 민간인의 비율이 반을 차지하지 않는 반면 is는 민간인이 반절을 차지하고 있다.

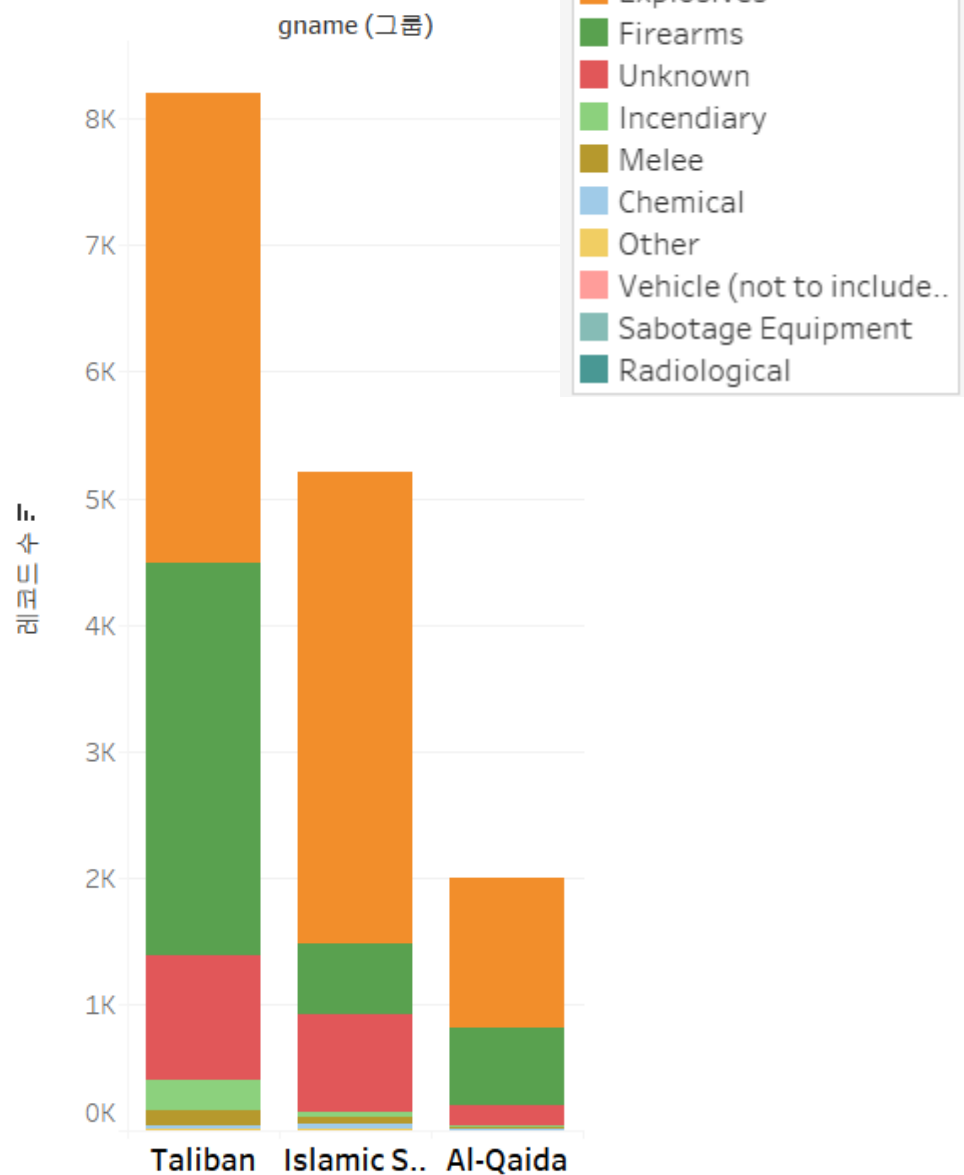
유명3의 무기분포



유명3팀의 무기 분포

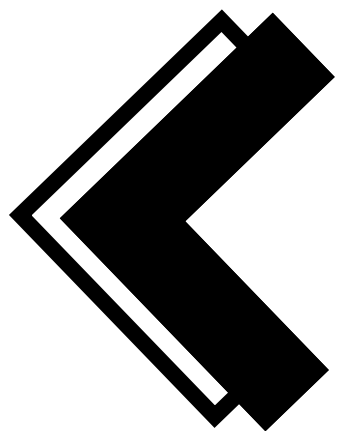
주로 Explosives와
Firearms가 사용되었다.

유명3의 무기타입

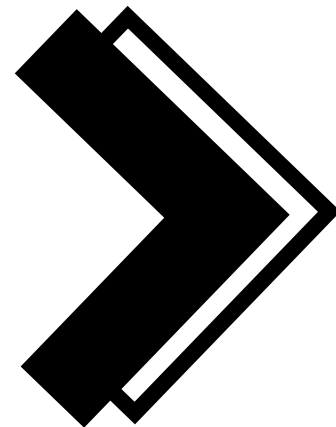


유명 3팀의 무기 타입

유명 3팀이 모두 비슷한 비율을 갖고있다.
Explosives, Firearms가 무기타입의 대부분을 차지하고 있다.

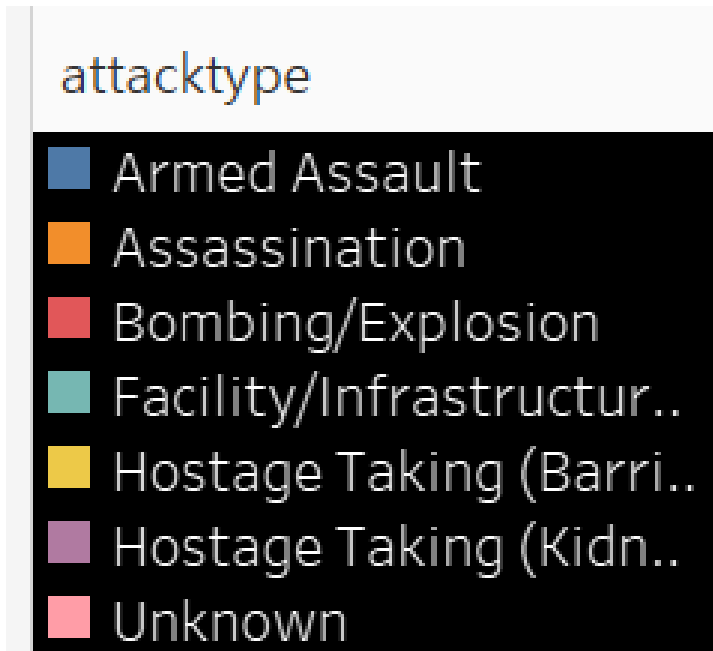


세부 분석1- 테러집단별



테러 집단별 가장 많은 테러를 발생했던 년도 세부분석

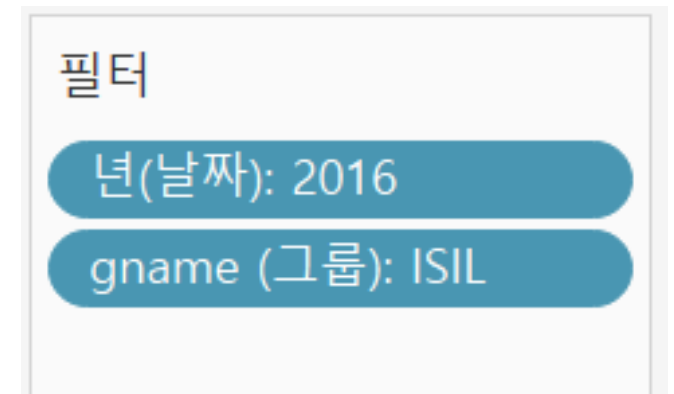
1. Attack type 그룹화



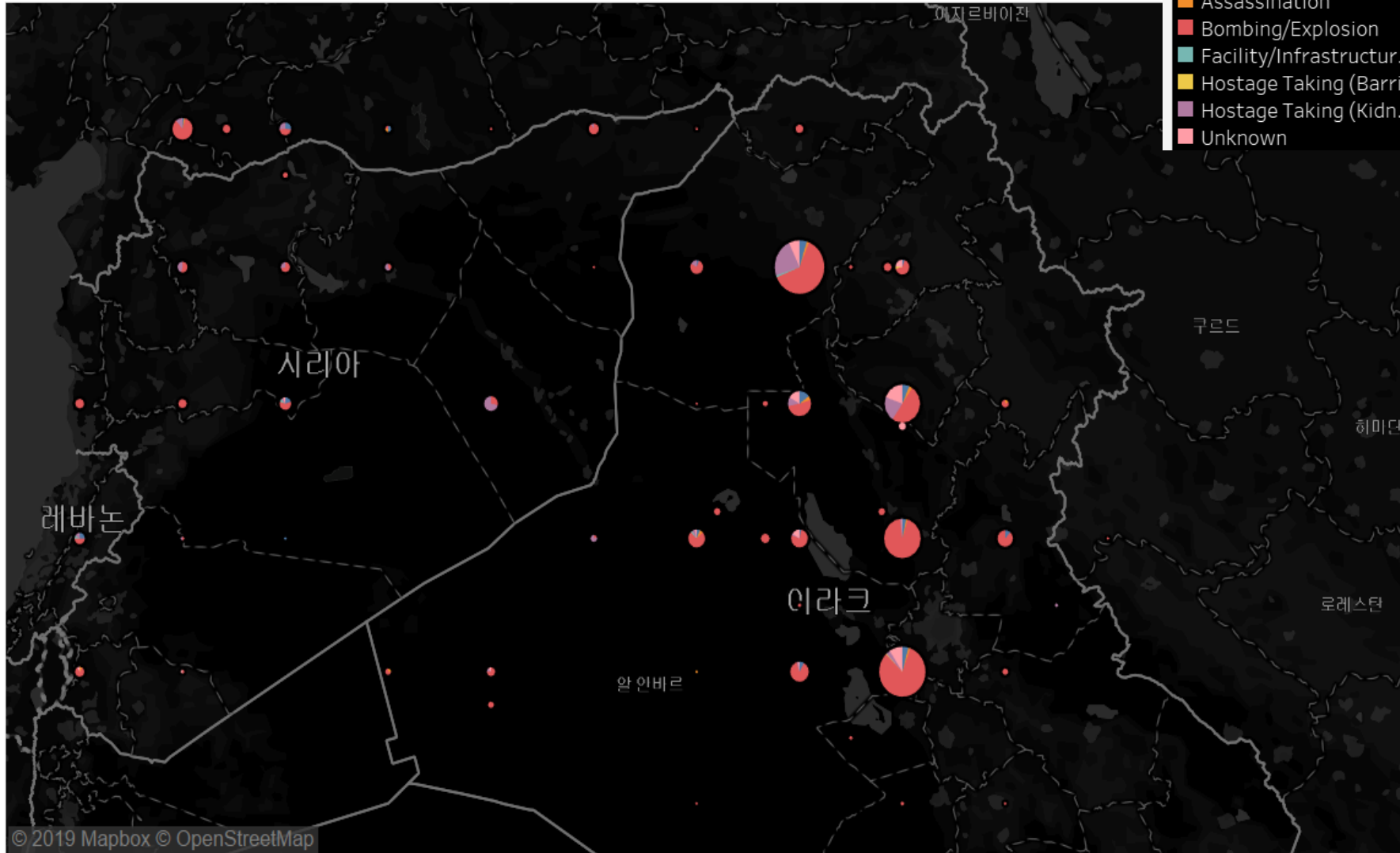
2. 맵에서 파이차트 구현



3. 특정년도와 그룹 필터링



IS 2016_attacktype



IS 2016_attacktype (100km 공간분할)

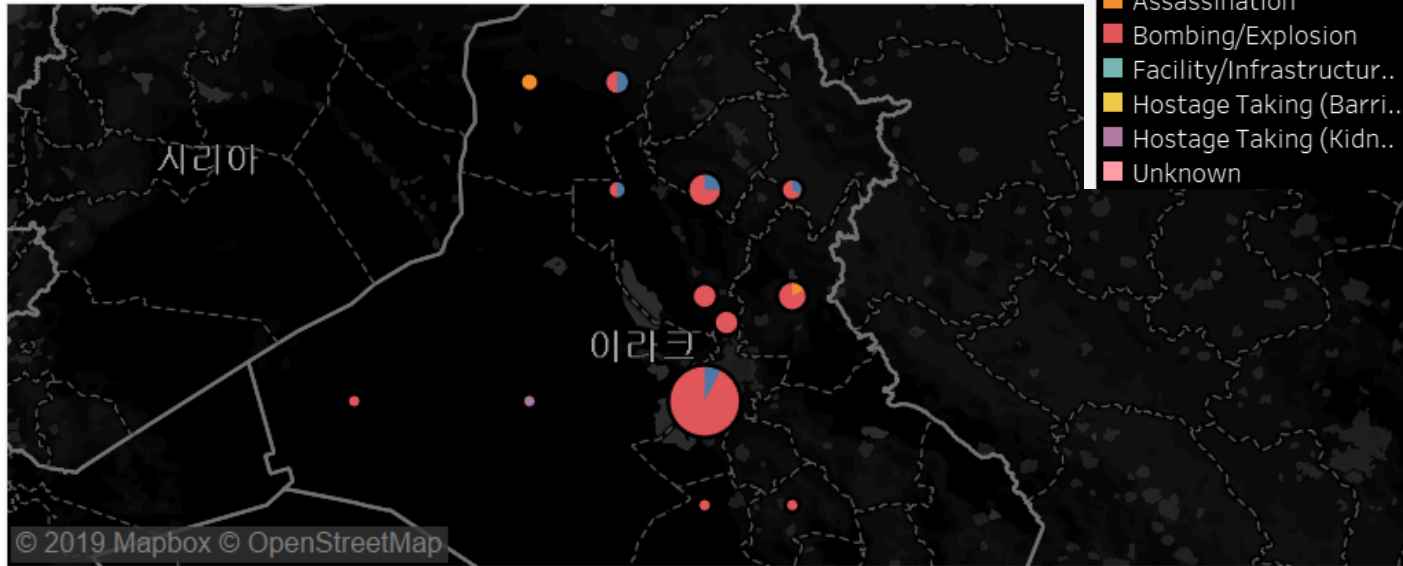
- Bombing/Explosion 타입이 가장 많다.
- 아래의 사진은 IS 배후세력의 폭탄 테러사건



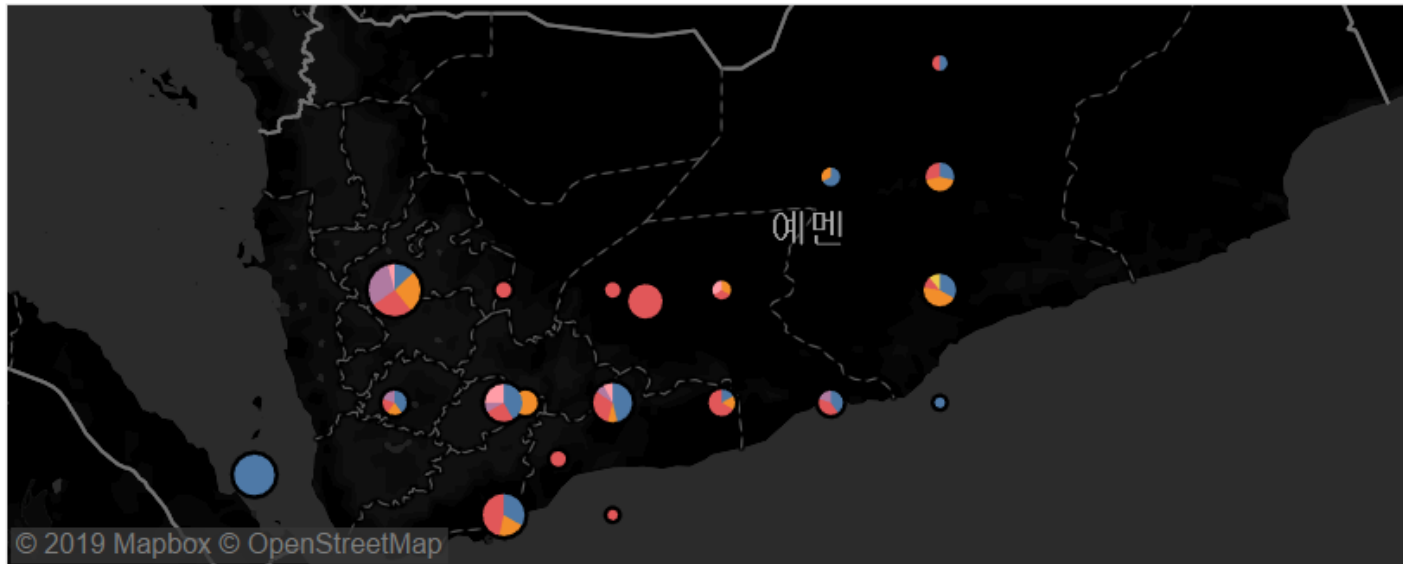
2016년 7월 3일 이라크의 바그다드에서 폭탄 테러가 발생했다. 이 테러로 290명 이상이 사망하고 225명 이상이 부상당했다.

- Hostage Taking(인질 납치)가 두번째 순위이다.

Al-Qaida 2012_attacktype_이라크



Al-Qaida 2012_attacktype_예멘



Al-Qaida_attacktype (100km 공간분할)

Al-Qaida 테러 집단은 크게 이라크와 예멘에서 활동

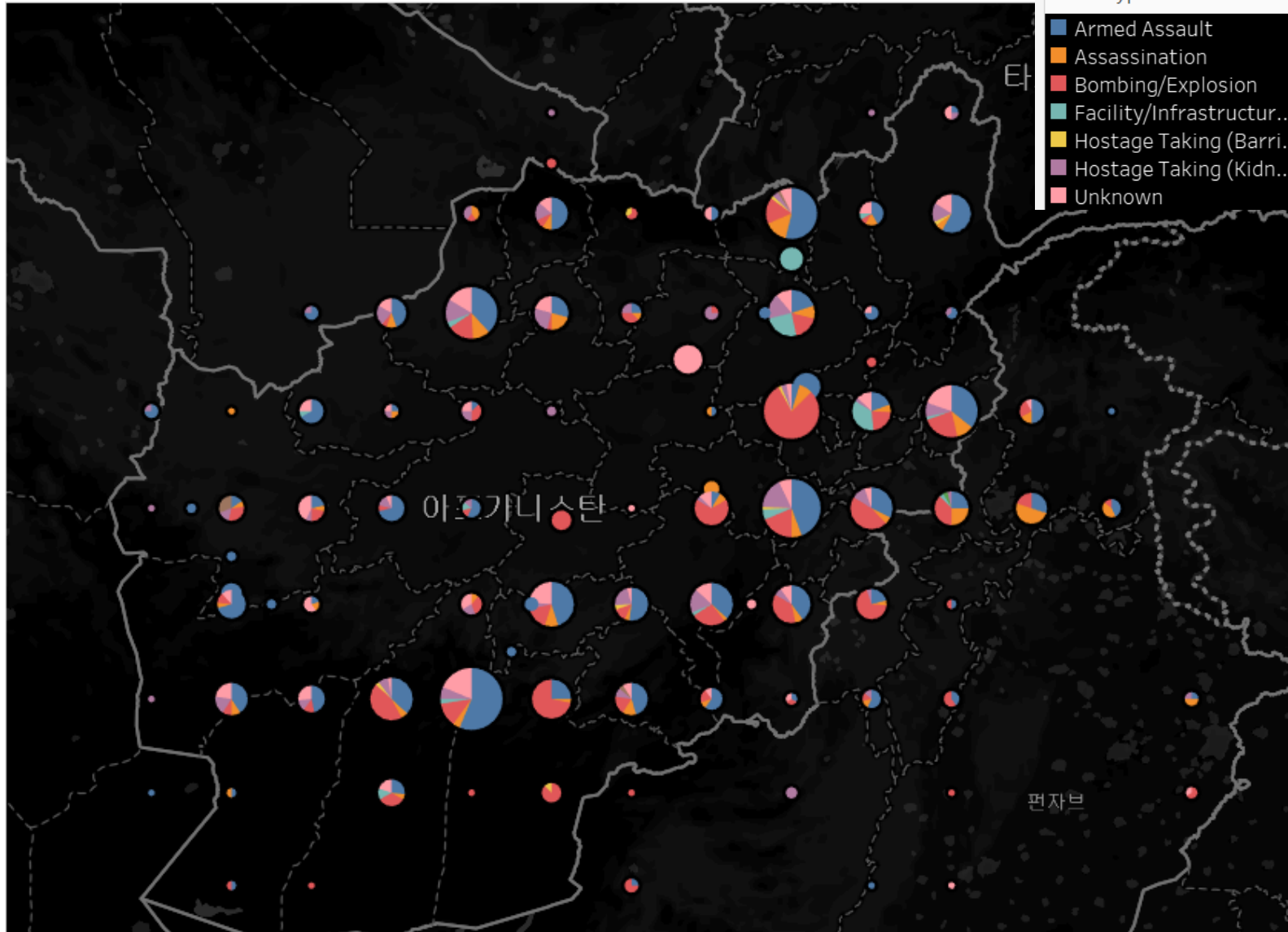
- Bombing/Explosion 타입이 가장 많다.
- Armed Assault(무장폭행)와 Assassination (암살)타입이 그 뒤를 따른다.



Al-Qaida

극단적 살라프파 무슬림에 의한 국제 무장 세력망

Taliban 2015_attacktype



Taliban_attacktype (100km 공간분할)

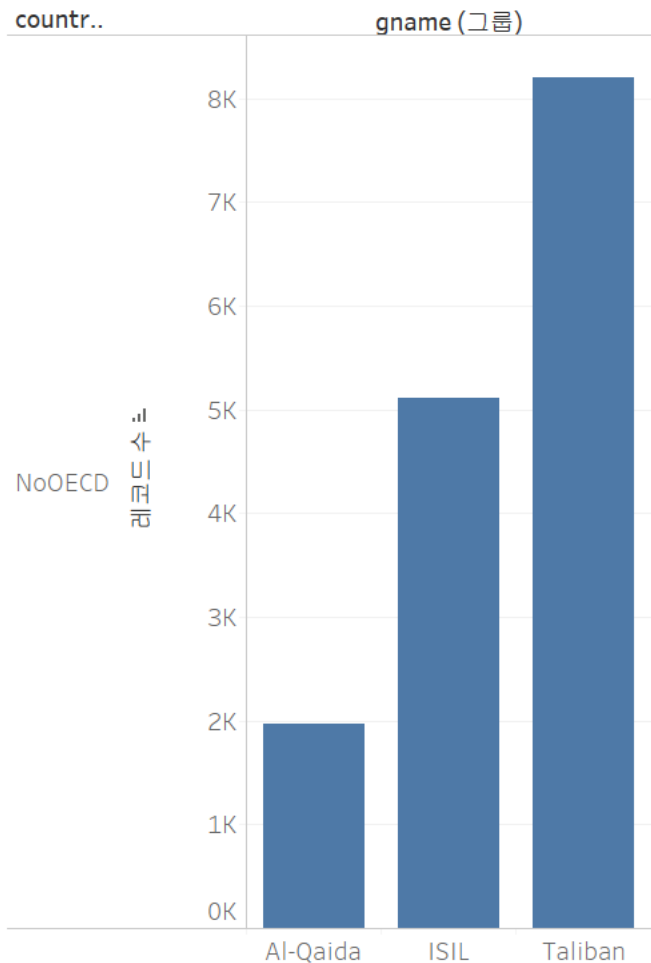
- Armed Assault(무장폭행)이 가장 많고, Bombing/Explosion, Hostage Taking(kidnapping), Assassination 순으로 양상을 보임
- 다른 테러 단체보다 다양한 양상을 보인다.
- 아래의 사진은 2015년 의사당 주변에서 탈레반의 연쇄 폭탄/차량 폭파 자살 테러



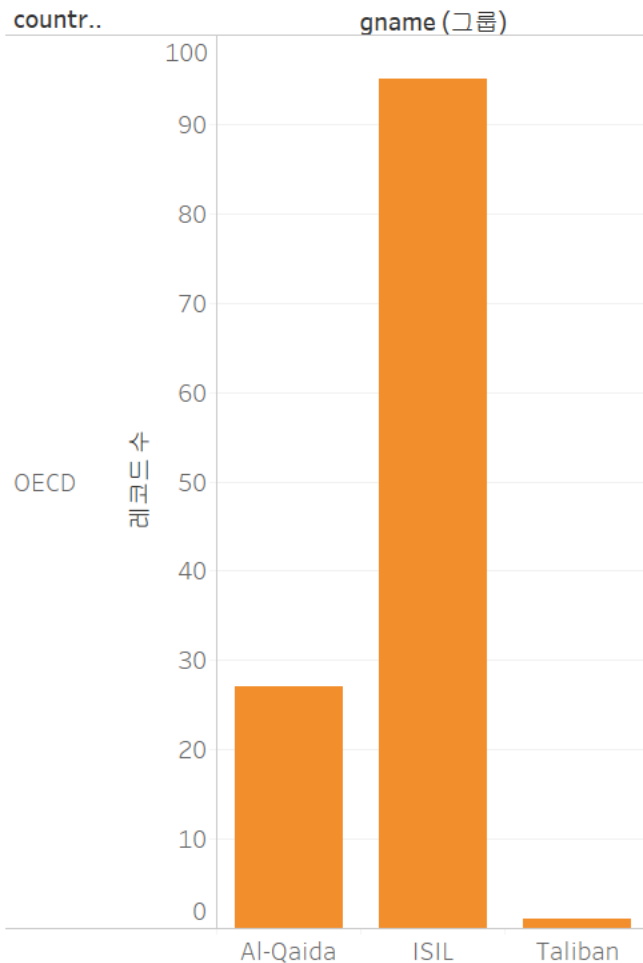
세부 분석2- OECD와 NoOECD 와 테러 단체



NoOECD

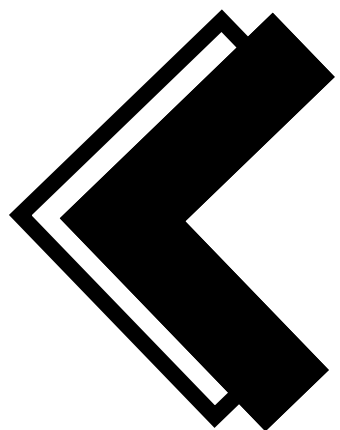


OECD

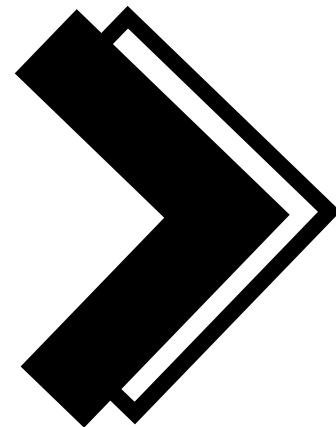


주요 테러 그룹의 OECD와 아닌 국가들에 미친 영향

- OECD 국가는 NoOECD 국가보다 레코드 수 차이가 많이 난다.
- OECD 국가
 - ISIL > Al-Qaida > Taliban
- NoOECD 국가
 - Al-Qaida > ISIL > Taliban
- - ISIL과 Al-Qaida은 Taliban에 비해 상대적으로 국가에 국한 되지않고 여러 국가에서 테러를 발생시켰고, 그에 반해 Taliban은 NoOECD에서 테러를 많이 발생시키는 양상을 보임

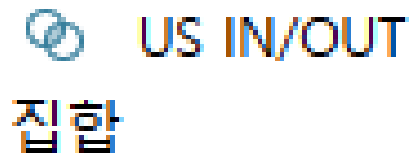


미국인 중심



집합 추가

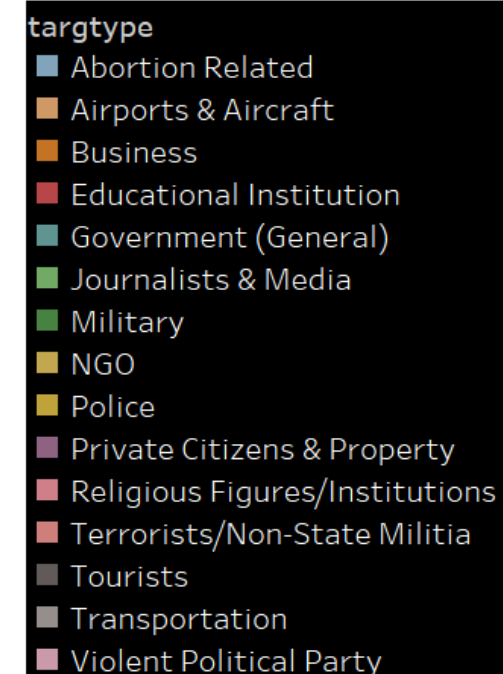
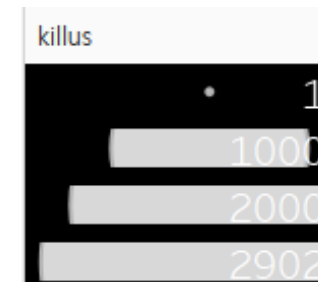
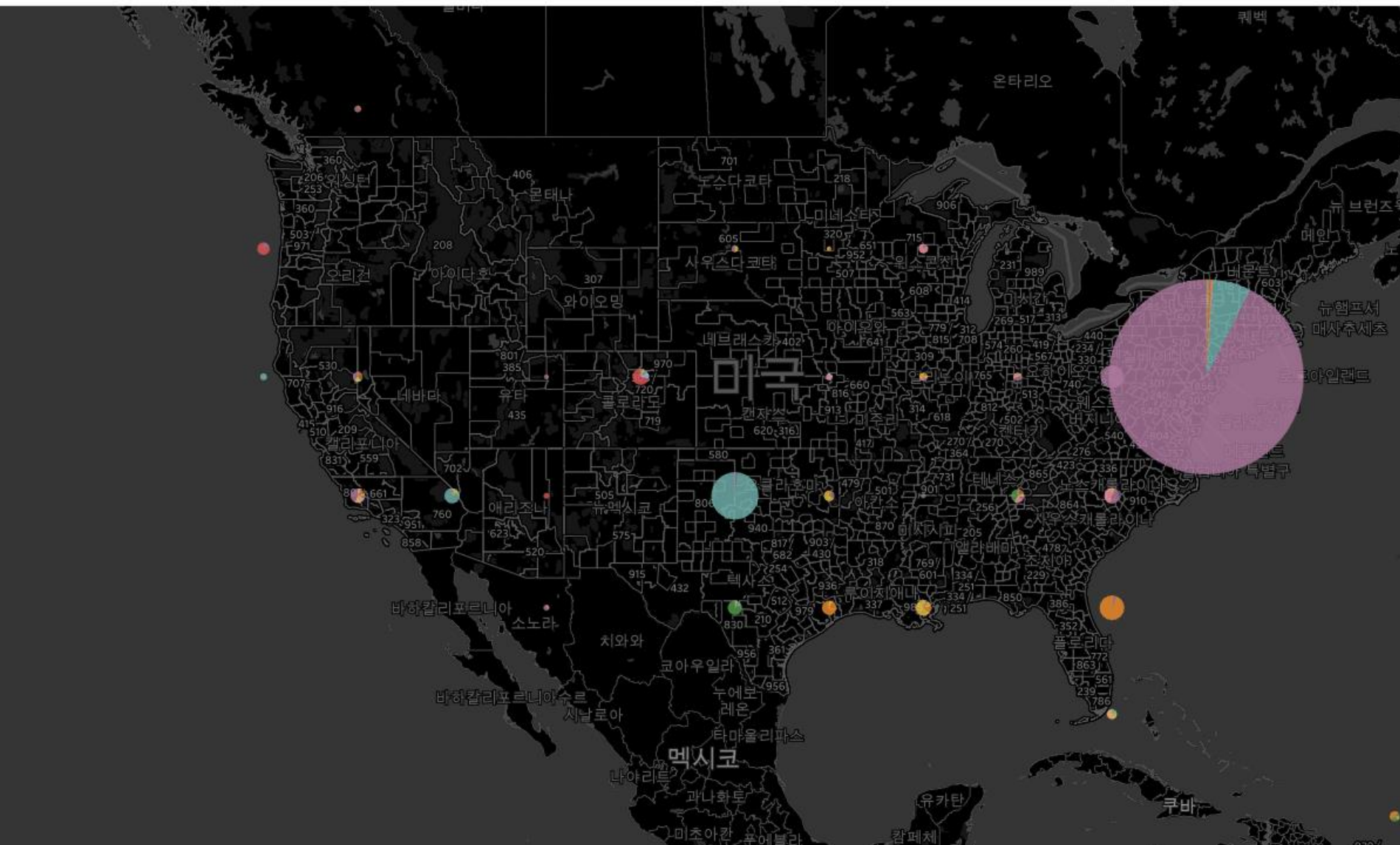
- US IN/OUT : 테러가 일어난 나라들 205개 중 United States를 제외한 204개의 나라의 집단이다.



요약

필드:	[country]
선택 내용:	204개(205개 중)의 값을 선택함
와일드카드(W):	전체
조건(D):	없음
제한(M):	없음

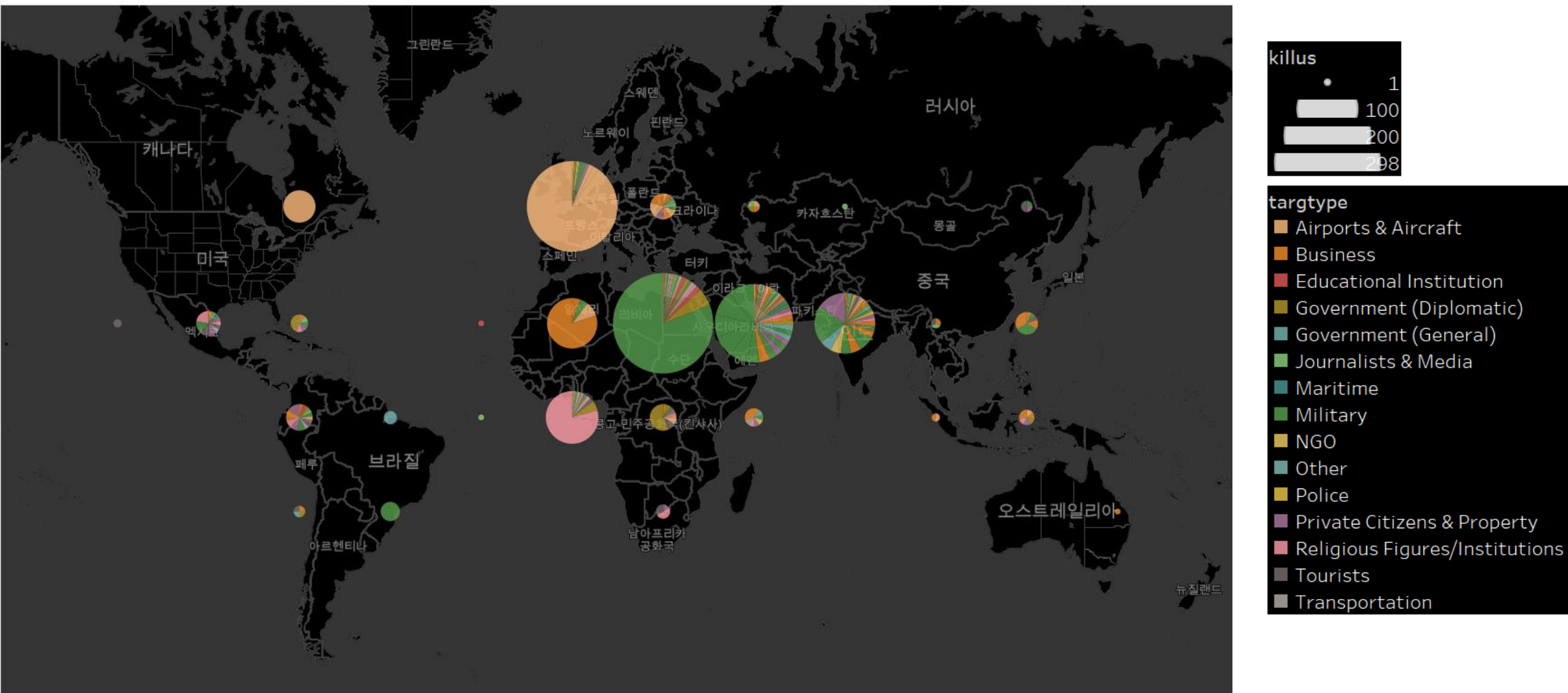
미국 내 미국인 사상자의 Target Type



미국 내 미국인 사상자의 Target Type(500km 공간 분할)

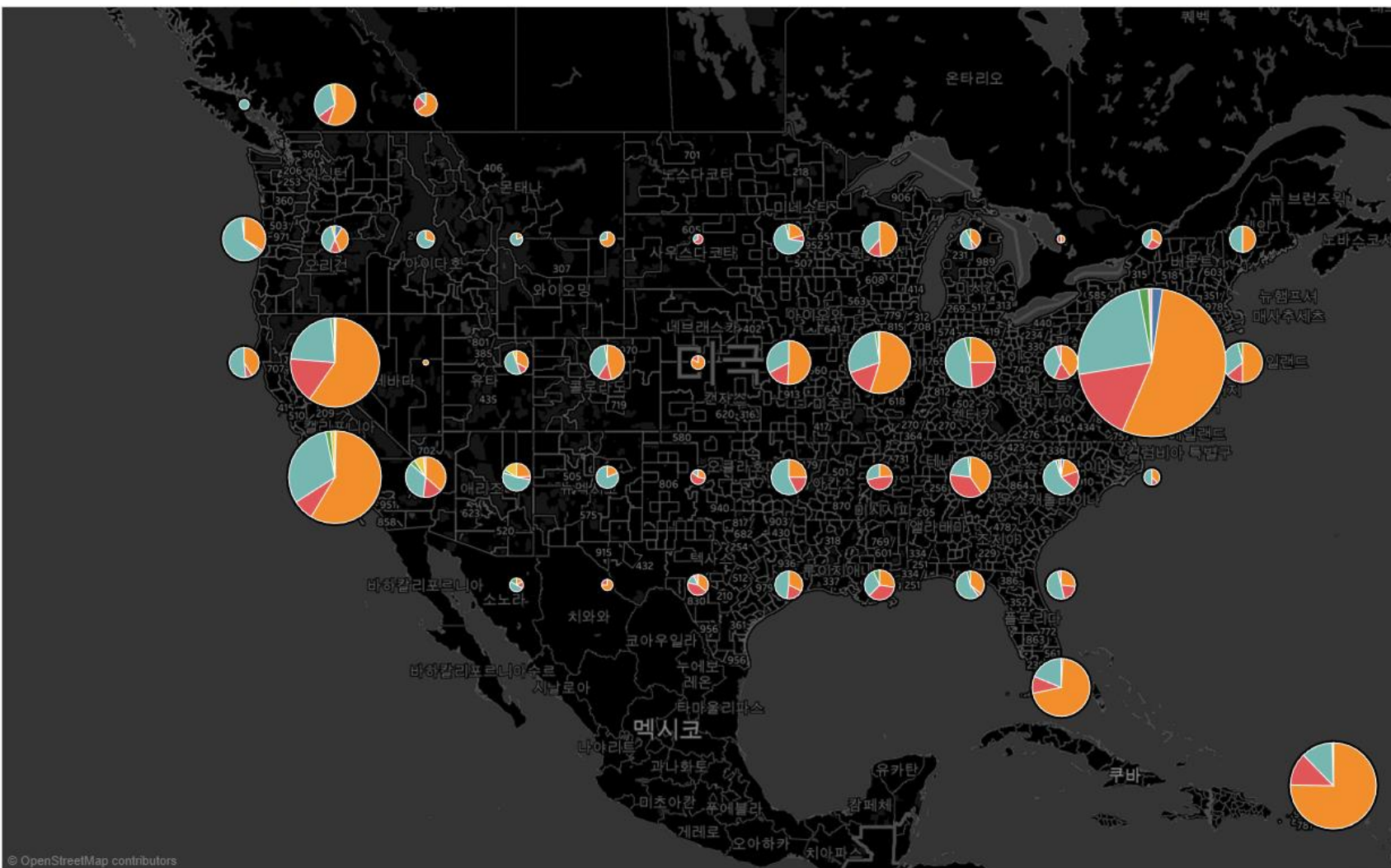
- 미국 내의 테러로 인한 미국인 사상자에는 Citizens가 유독 많았다.

미국 외 미국인 사상자의 Target Type



미국 외 미국인 사상자의 Target Type(2500km 공간 분할)

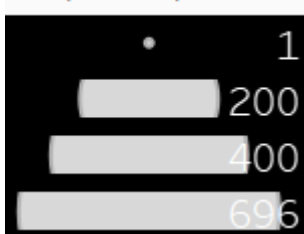
- 미국 외의 테러로 인한 미국인 사상자에는 주로 Military가 가장 많았고, 그 다음으로 Airports & Aircraft도 많다.



미국 내의 Weapon Type (500km 공간 분할)

- 여러가지 무기들이 사용되지만, 주로 Incendiary와 Explosive가 많은 비율을 차지하고 있다.

합계(레코드 수)



weaptype

- Biological
- Explosives
- Firearms
- Incendiary
- Melee
- Sabotage Equipment
- Vehicle (not to include vehicle-borne explosives, i.e., car or truck bombs)

고찰

데이터를 선정하는 것보다 그 선정한 데이터를 전처리하는 것이 더 어려웠다. 어떤 컬럼을 삭제해야 할지 고민됐지만 교수님의 조언을 듣고 눈 딱 감고 삭제하니 훨씬 더 수월하게 데이터를 가공할 수 있었다. 처음에 많은 컬럼이 있는 상태로 하다보니 거의 보고서가 80쪽이 넘을 정도로 많고 어떻게 보면 중구난방이었는데 교수님의 조언이 많은 도움이 되었다.

테러 그룹을 그룹화 하고 top3를 attacktype 과 연결 지으면서 테러 그룹에 대한 테러 사건 정보를 찾아보는 과정을 거쳤다. 데이터 분석 결과를 사회적 현상에 연결 짓고 추측하는 과정을 통해 데이터 분석에 흥미를 느꼈고 그 필요성도 알 수 있었던 활동이었다.