

Method of suicide:  
international suicide patterns  
derived from the WHO  
mortality database

# Papers



[Bull World Health Organ](#). 2008 Sep; 86(9): 726–732.

PMCID: PMC2649482

Published online 2008 Jun 2. doi: [10.2471/BLT.07.043489](https://doi.org/10.2471/BLT.07.043489)

PMID: [18797649](https://pubmed.ncbi.nlm.nih.gov/18797649/)

Language: English | [French](#) | [Spanish](#) | [Arabic](#)

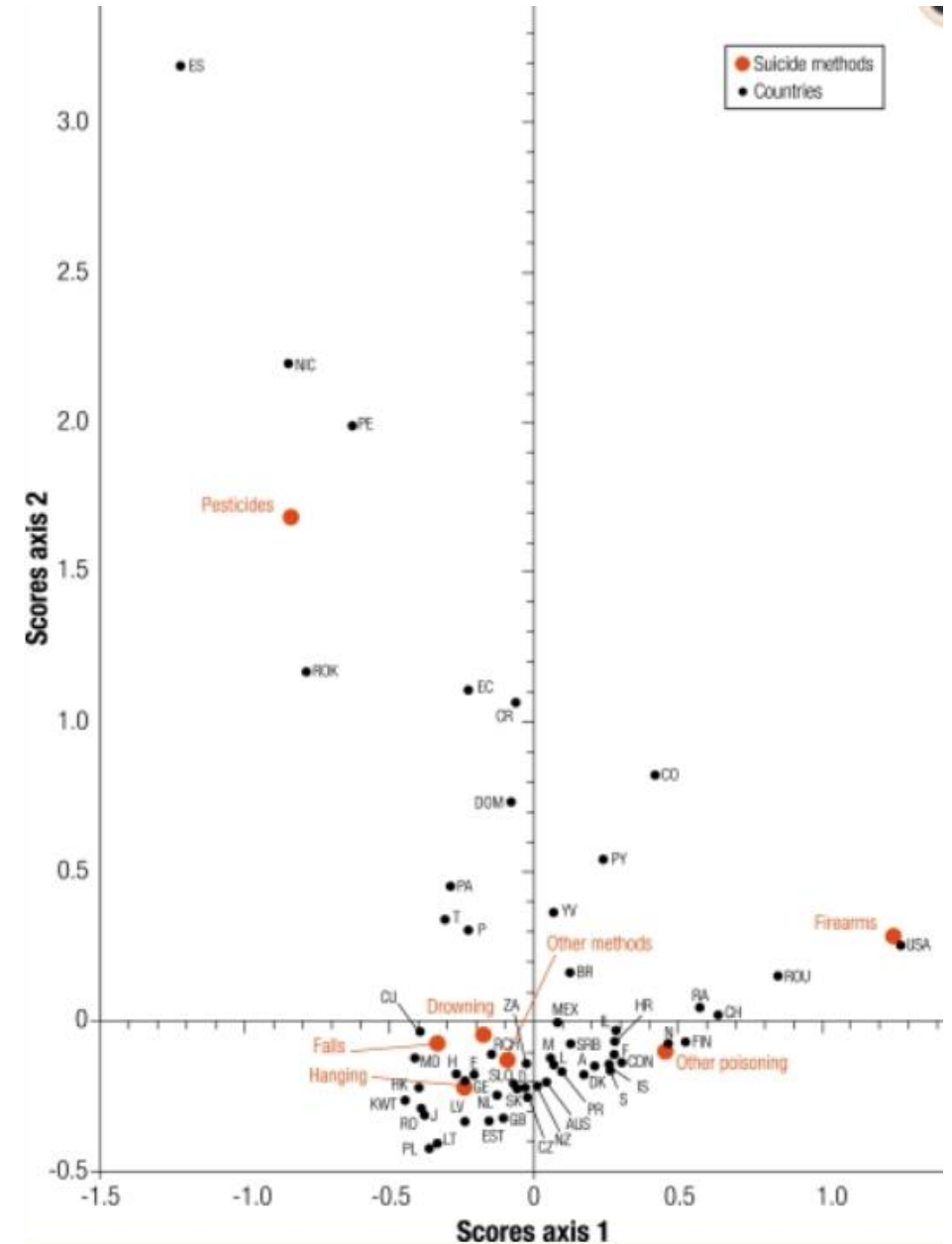
## Methods of suicide: international suicide patterns derived from the WHO mortality database

[Vladeta Ajdacic-Gross](#),<sup>a</sup> [Mitchell G Weiss](#),<sup>b</sup> [Mariann Ring](#),<sup>a</sup> [Urs Hepp](#),<sup>c</sup> [Matthias Bopp](#),<sup>d</sup> [Felix Gutzwiller](#),<sup>d</sup> and [Wulf Rössler](#)<sup>a</sup>

**FINDINGS:** Poisoning by pesticide was common in many Asian countries and in Latin America; poisoning by drugs was common in both Nordic countries and the United Kingdom. Hanging was the preferred method of suicide in eastern Europe, as was firearm suicide in the United States and jumping from a high place in cities and urban societies such as Hong Kong Special Administrative Region, China. Correspondence analysis demonstrated a polarization between pesticide suicide and firearm suicide at the expense of traditional methods, such as hanging and jumping from a high place, which lay in between.

# 논문 method

1. WHO mortality database에서  
국가, 자살 방법별 사망률에 대한 정보 뽑기
2. PCA 분석으로, 자살 패턴이 비슷한 나라 확인하기



WHO Mortality database 구조

# WHO Mortality Database



- World Health Organization
- WHO Mortality Database는  
연령, 성별 및 사망 원인 별 **사망률 데이터**를 집계 한 것으로,  
회원국이 시민 등록 시스템을 통해 매년 보고합니다.

Final Output

사망률 테이블(mortrate2)

국가	년도	ICD버전	사망원인	성별	나이0~4세 사망률						
Country_code	Year	List	Cause_code	Sex	Format	Age0to4	Age5to19	Age20to39	Age40to59	...	MortRate17
1125	1978	08A	A000	1	03	2997.8362	182.3809	319.1513	1143.4632	...	2299.4274
1125	1978	08A	A000	2	03	3053.4023	136.2138	198.1190	574.1872	...	1180.0894
1125	1978	08A	A001	1	03	0.0000	0.0000	0.0000	0.0000	...	0.0000

Country_code	Year	List	Cause_code	Sex	Format	Age0to4	Age5to19	Age20to39	Age40to59	...	MortRate17
1125	1978	08A	A000	1	03	2997.8362	182.3809	319.1513	1143.4632	...	2299.4274
1125	1978	08A	A000	2	03	3053.4023	136.2138	198.1190	574.1872	...	1180.0894
1125	1978	08A	A001	1	03	0.0000	0.0000	0.0000	0.0000	...	0.0000

사망률 = (사망자 수 / 전체 인구 수) \* 100,000  
 = 100,000명 당 \_\_\_ 명 사망

- 1125 국가에서 1978년에 A000 사망원인으로 죽은 남자 0-4세는 100,000명 당 2,997명이다.
- 1125 국가에서 1978년 A000 사망원인으로 죽은 여자 0-4세는 100,000명 당 3,053명이다.



# International Classification of disease (ICD)

- ICD: 질병 및 관련 건강 문제의 국제 통계 분류
- 세계 보건 기구에서 질병과 증상을 분류한 것
- ICD - version 7, 8, 9, 10

# International Classification of disease (ICD)

- ICD -10 질병의 코드 예시

챕터	구간	제목
I	A00-B99	특정 전염병 및 기생충병
II	C00-D48	종양
III	D50-D89	혈액과 조혈 기관의 질병 및 면역 체계 관련 특정 장애
IV	E00-E90	내분비, 영양 및 대사 질병
V	F00-F99	정신 및 행동 장애
VI	G00-G99	신경 계통의 질병
VII	H00-H59	눈과 부속기의 질병
VIII	H60-H95	귀와 <b>유양돌기</b> 의 질병
IX	<b>I00-I99</b>	순환 계통의 질병
X	<b>J00-J99</b>	호흡 계통의 질병
XI	<b>K00-K93</b>	소화 계통의 질병
XII	<b>L00-L99</b>	<b>피부</b> 와 <b>피하 조직</b> 의 질병

Final Output

사망률 테이블(mortrate2)

국가	년도	ICD버전	사망원인	성별		나이0~4세 사망률						
Country_code	Year	List	Cause_code	Sex	Format	Age0to4	Age5to19	Age20to39	Age40to59	...	MortRate17	
1125	1978	08A	A000	1	03	2997.8362	182.3809	319.1513	1143.4632	...	2299.4274	
1125	1978	08A	A000	2	03	3053.4023	136.2138	198.1190	574.1872	...	1180.0894	
1125	1978	08A	A001	1	03	0.0000	0.0000	0.0000	0.0000	...	0.0000	

ICD-10: <https://icd.who.int/browse10/2016/en>

# icdcodes table

List	Cause_code	Detailed_code	Cause
07A	A140	E870-E895	Accidental poisoning
07A	A141	E900-E904	Accidental falls
07A	A142	E912	Accident caused by machinery
07A	A143	E916	Accident caused by fire and explosion of combustible material

population1 – 각 국가의 해당 년도에 인구 수

Country_code	Year	Sex	Frmat	Pop1	Pop2	Pop3	Pop4	Pop5
1060	1980	1	7	137100	3400	15800	NA	NA
1060	1980	2	7	159000	4000	18400	NA	NA
1125	1955	1	2	5051500	150300	543400	NA	NA
1125	1955	2	2	5049400	145200	551000	NA	NA



mortality2 – 각 국가의 해당 년도에 특정 사망 원인으로 죽은 사람 수

Country_code	Year	List	Cause_code	Sex	Frmat	IM_Frmat	Deaths1	Deaths2	Deaths3	Deaths4	Deaths5
1060	1980	08A	A000	1	07	08	1087	375	132	NA	NA
1060	1980	08A	A000	2	07	08	1194	313	133	NA	NA
1060	1980	08A	A001	1	07	08	0	0	0	NA	NA
1060	1980	08A	A001	2	07	08	0	0	0	NA	NA

$$\text{mortrate} = 100000 * (\text{Deaths}/\text{Pop})$$

# Population table

Pop1 ↵	Population at all ages ↵
Pop2 ↵	Population at age 0 year ↵
Pop3 ↵	Population at age 1 year ↵
Pop4 ↵	Population at age 2 years ↵
Pop5 ↵	Population at age 3 years ↵
Pop6 ↵	Population at age 4 years ↵
Pop7 ↵	Population at age 5-9 years ↵
Pop8 ↵	Population at age 10-14 years ↵
Pop9 ↵	Population at age 15-19 years ↵
Pop10 ↵	Population at age 20-24 years ↵
Pop11 ↵	Population at age 25-29 years ↵
Pop12 ↵	Population at age 30-34 years ↵
Pop13 ↵	Population at age 35-39 years ↵
Pop14 ↵	Population at age 40-44 years ↵
Pop15 ↵	Population at age 45-49 years ↵
Pop16 ↵	Population at age 50-54 years ↵
Pop17 ↵	Population at age 55-59 years ↵
Pop18 ↵	Population at age 60-64 years ↵

Pop1 ↵	Population at all ages ↵
Pop2 ↵	Population at age 0 year ↵
Pop3 ↵	Population at age1 year ↵
Pop4 ↵	Population at age 2 years ↵
Pop5 ↵	Population at age 3 years ↵
Pop6 ↵	Population at age 4 years ↵
Pop7 ↵	Population at age 5-9 years ↵
Pop8 ↵	Population at age 10-14 years ↵

population1 – 각 국가의 해당 년도에 인구 수

Country_code	Year	Sex	Format	Pop1	Pop2	Pop3	Pop4	Pop5
1060	1980	1	7	137100	3400	15800	NA	NA
1060	1980	2	7	159000	4000	18400	NA	NA
1125	1955	1	2	5051500	150300	543400	NA	NA
1125	1955	2	2	5049400	145200	551000	NA	NA

- 첫번째 행의 Pop1: 1060 국가에 1980년 총 남자 인구 수
- 첫번째 행의 Pop2: 1060 국가에서 1980년 남자 0세 인구 수
- 첫번째 행의 Pop3: 1060 국가에서 1980년 남자 1세 인구 수

# Mortality table

Deaths1 ↴	Deaths at all ages ↴
Deaths2 ↴	Deaths at age 0 year ↴
Deaths3 ↴	Deaths at age 1 year ↴
Deaths4 ↴	Deaths at age 2 years ↴
Deaths5 ↴	Deaths at age 3 years ↴
Deaths6 ↴	Deaths at age 4 years ↴
Deaths7 ↴	Deaths at age 5-9 years ↴
Deaths8 ↴	Deaths at age 10-14 years ↴
Deaths9 ↴	Deaths at age 15-19 years ↴
Deaths10 ↴	Deaths at age 20-24 years ↴
Deaths11 ↴	Deaths at age 25-29 years ↴
Deaths12 ↴	Deaths at age 30-34 years ↴
Deaths13 ↴	Deaths at age 35-39 years ↴
Deaths14 ↴	Deaths at age 40-44 years ↴
Deaths15 ↴	Deaths at age 45-49 years ↴
Deaths16 ↴	Deaths at age 50-54 years ↴
Deaths17 ↴	Deaths at age 55-59 years ↴
Deaths18 ↴	Deaths at age 60-64 years ↴
Deaths19 ↴	Deaths at age 65-69 years ↴



Deaths1 ↵	Deaths at all ages ↵
Deaths2 ↵	Deaths at age 0 year ↵
Deaths3 ↵	Deaths at age 1 year ↵
Deaths4 ↵	Deaths at age 2 years ↵
Deaths5 ↵	Deaths at age 3 years ↵
Deaths6 ↵	Deaths at age 4 years ↵
Deaths7 ↵	Deaths at age 5-9 years ↵
Deaths8 ↵	Deaths at age 10-14 years ↵

mortality2 – 각 국가의 해당 년도에 특정 사망 원인으로 죽은 사람 수

Country_code	Year	List	Cause_code	Sex	Format	IM_Format	Deaths1	Deaths2	Deaths3	Deaths4	Deaths5
1060	1980	08A	A000	1	07	08	1087	375	132	NA	NA
1060	1980	08A	A000	2	07	08	1194	313	133	NA	NA
1060	1980	08A	A001	1	07	08	0	0	0	NA	NA

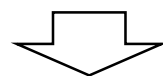
- 첫번째 행의 Deaths1: 1060 국가에 1980년 A000 사망원인으로 죽은 총 남자 사망자 수
- 첫번째 행의 Deaths2: 1060 국가에서 1980년 A000 사망원인으로 죽은 0세 남자 사망자 수
- 첫번째 행의 Deaths3: 1060 국가에서 1980년 A000 사망원인으로 죽은 1세 남자 사망자 수

population1 –인구수 정보

Country_code	Year	Sex	Format	Pop1	Pop2	Pop3	Pop4	Pop5
1060	1980	1	7	137100	3400	15800	NA	NA

mortality2 – 사망자 정보

Country_code	Year	List	Cause_code	Sex	Format	IM_Format	Deaths1	Deaths2	Deaths3	Deaths4	Deaths5
1060	1980	08A	A000	1	07	08	1087	375	132	NA	NA



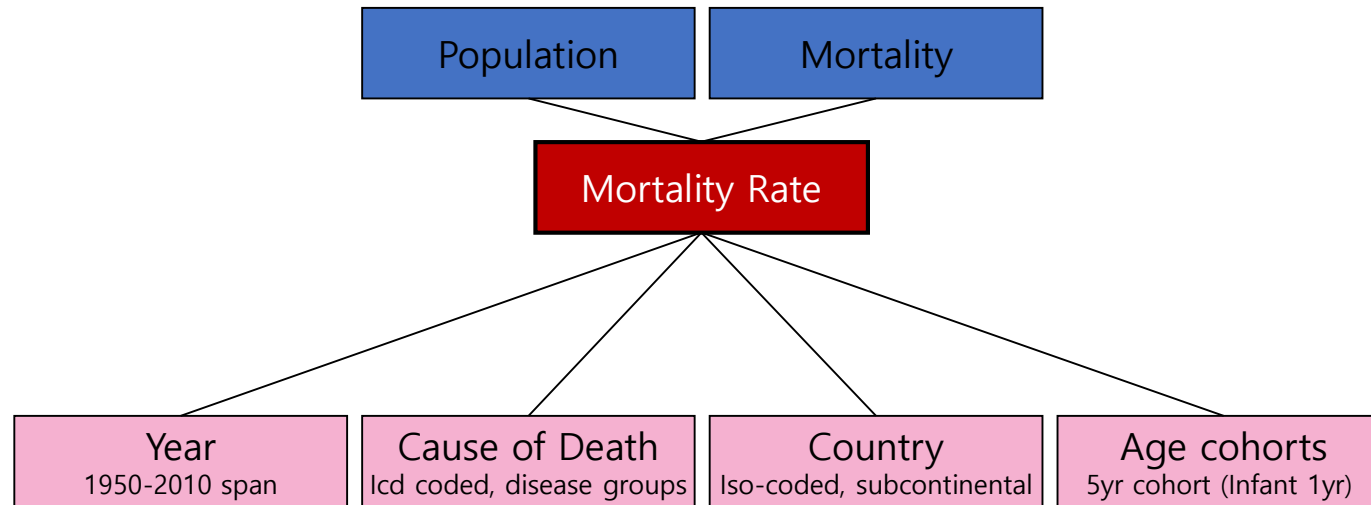
mortrate2 (mortrate1=100000 \* (Deaths1/Pop1))

Country_code	Year	List	Cause_code	Sex	Format	Age0to4	Age5to19	Age20to39	Age40to59	Age60to79	Age80andup	MortRate1	MortRate2
1060	1980	08A	A000	1	07	NA	NA	NA	NA	NA	NA	792.8519	11029.4117

$$\text{mortrate1} = (1087/137100) \times 100000 = 792.8519$$

mortrate1 = 1060 국가에서 1980년도에 A000의 사망 원인으로 죽은 남성은  
10만명당 792명

# WHO mortality database concept

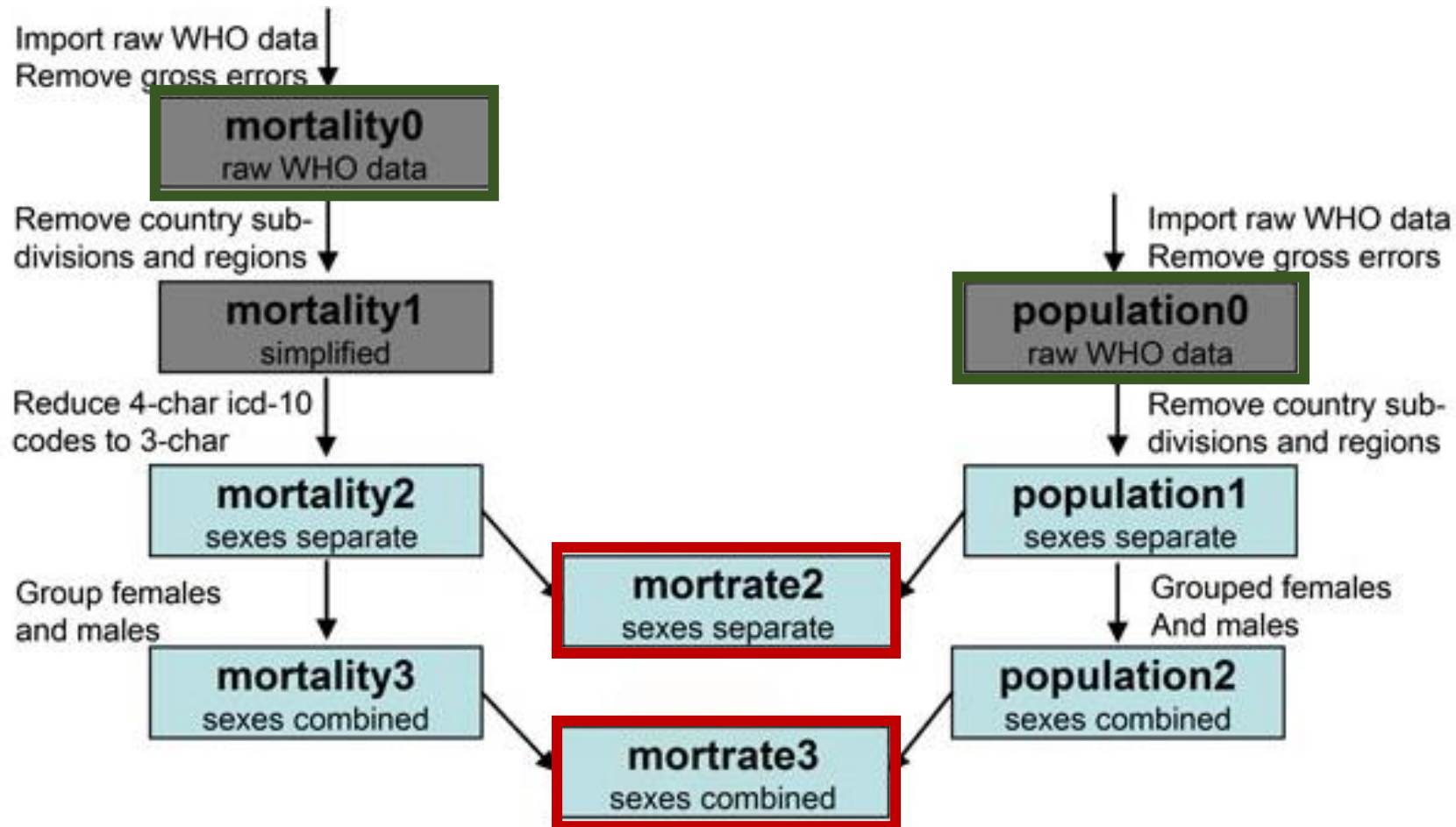


Mortality rates were calculated from the population and mortality data. Within the mortality dataset, one can look at the data from different perspectives using the variables year, cause of death, country, age cohort and sex.

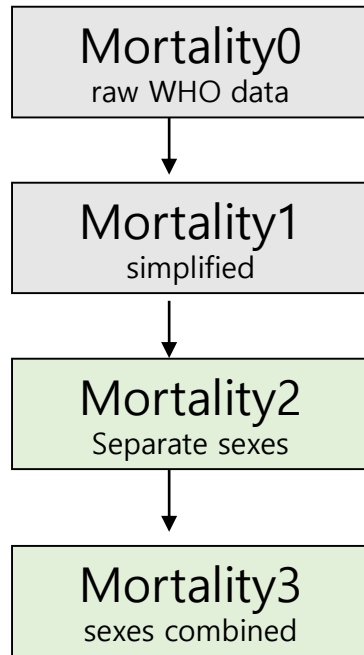
# WHO Mortality database Quality Control

Mortality QC

# Quality Control 과정



# Quality Control – mortality table



1. sub-divisions 또는 regions 정보가 있는 행은 모두 지우기

2. icd10 코드가 4글자로 되어있으면, 3글자로 치환

3. 남녀 정보 통합

# QC 1<sup>st</sup> step in mortality

1. mortality0 테이블에서 sub-divisions 또는 regions 정보가 있는 행은 모두 지우기
2. 1번의 결과를 mortality1 테이블로 저장

## ▪ Field “Admin1” ↵

Country ↵	Admin1 ↵	Description ↵
2070 ↵	901 ↵	North and North-East ↵
2070 ↵	902 ↵	South, South-East and Central West ↵
2350 ↵	901 ↵	Former Canal Zone ↵
3150 ↵	901 ↵	Jewish Population ↵

## ▪ Field “Subdiv” ↵

Subdiv ↵	Description ↵
A10 ↵	Survey ↵
A20 ↵	Reporting Areas ↵
A30 ↵	Part ↵
A35 ↵	Selected Urban and Rural Areas ↵
A41 ↵	Selected Rural Areas ↵
A51 ↵	Selected Urban Areas ↵
A70 ↵	Cities ↵
A80 ↵	Certified Deaths ↵



# QC 1<sup>st</sup> step in mortality

- mortality0 테이블에서 sub-divisions 또는 regions 정보가 있는 **행**은 모두 지우기

Country_code	Admin1	Sub-division	Cause_code	Death0
1101	NULL	NULL	A00	1000
1101	NULL	<b>North</b>	A00	50

첫번째 줄: 1101 국가의 A00 사망원인으로 죽은 전체 사망자 수는 1000명이다

두번째 줄: 1101 국가의 **북쪽 지역에서** A00 사망원인으로 죽은 전체 사망자 수는 100명이다

-> 두번째 줄은 첫번째 줄의 정보의 **부분집합**이다.

-> 우리는 국가별 사망률을 분석

-> Sub-division에 값이 있는 행 제거! ( 즉, 두번째 줄 제거)

# QC 1<sup>st</sup> step in mortality

- mortality0 테이블에서 sub-divisions 또는 regions 정보가 있는 **행**은 모두 지우기

Country_code	Admin1	Sub-division	Cause_code	Death0
1101	NULL	NULL	A00	1000
1101	<b>Cities</b>	NULL	A00	100

첫번째 줄: 1101 국가의 A00 사망원인으로 죽은 전체 사망자 수는 1000명이다

두번째 줄: 1101 국가에서 도시에서 A00 사망원인으로 죽은 전체 사망자 수는 100명이다

-> 두번째 줄은 첫번째 줄의 정보의 **부분집합**이다.

-> 우리는 국가별 사망률을 분석

-> Admin1에 값이 있는 행 제거! ( 즉, 두번째 줄 제거)

QC 1<sup>st</sup> step in raw mortality

## Mortality

Rows without subdivs/region	mort0	3920396
Rows with subdivs/regions	mort0	141236
		4061632
Mortality1	mort1	3920396

# QC 1<sup>st</sup> step in mortality

```
SELECT COUNT(*)  
FROM mortality1;
```

	COUNT(*)
▶	18114

- 1. mortality0 테이블에서 sub-divisions 또는 regions 정보가 있는 행은 모두 지우기
- 2. 1번의 결과를 mortality1 테이블로 저장

```
DROP TABLE IF EXISTS mortality1;  
CREATE TABLE mortality1 AS (  
SELECT Country_code,Year,List,Cause_code, Sex,  
FROM mortality0  
where (Admin1 IS NULL and SubDiv IS NULL)  
);
```

Country_code	Year	List	Cause_code	Sex	Frm	IM_Frm	Deaths1
4210	1996	10M	A020	1	00	08	2
4210	1996	10M	A020	2	00	08	2
....	....	....	....	.	--	--	-

Deaths26	IM_Deaths1	IM_Deaths2	IM_Deaths3	IM_Deaths4
0	0	NULL	NULL	NULL
0	0	NULL	NULL	NULL

# QC 2<sup>nd</sup> step in mortality

1. mortality table1에서 ICD-10 버전이면서,  
Cause\_code가 4글자로 되어 있으면 Cause\_code를 3글자로 바꾼다.
2. 동일한 Cause\_code끼리 묶어서 사망자 수 합산

예시)

mort1	4210	2000	104	C340	Male	0	1	26
mort1	4210	2000	104	C341	Male	0	1	164
mort1	4210	2000	104	C342	Male	0	1	8
mort1	4210	2000	104	C343	Male	0	1	54
mort1	4210	2000	104	C348	Male	0	1	9
mort1	4210	2000	104	C349	Male	0	1	6031
								6292
mort2	4210	2000	104	C34	Male	0	1	6292

C34 Malignant neoplasm of bronchus and lung

C340 Malignant neoplasm: Main bronchus

C341 Malignant neoplasm: Upper lobe, bronchus or lung

C342 Malignant neoplasm: Middle lobe, bronchus or lung

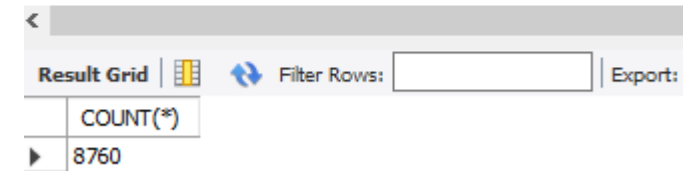
C343 Malignant neoplasm: Lower lobe, bronchus or lung

C348 Malignant neoplasm: Overlapping lesion of bronchus and lung

C349 Malignant neoplasm: Bronchus or lung, unspecified

# QC 2<sup>nd</sup> step in mortality

```
1 SELECT COUNT(*)
2 FROM mortality2;
```



COUNT(*)
8760

1. mortality table1에서 Cause\_code가 4글자로 되어 있으면 Cause\_code를 3글자로 바꾼다.
2. 동일한 Cause\_code끼리 묶어서 사망자 수 합산
3. Mortality2로 저장

Country_code	Year	List	Cause_code	Sex	Frmat	IM_Frmat	Deaths1
4210	1996	10M	A02	1	00	08	5
4210	1996	10M	A02	2	00	08	2
....	....	....	...	.	--	--	.

Deaths26	IM_Deaths1	IM_Deaths2	IM_Deaths3	IM_Deaths4
0	0	NULL	NULL	NULL
0	0	NULL	NULL	NULL

# QC 3<sup>rd</sup> step in mortality

mortality2에는 남자, 여자 따로 사망자수가 기록되어 있다.

1. 남자 여자 합쳐서 사망자수가 기록되도록 한다.
2. 1번 결과값을 mortality3 table에 저장

## Mortality

mort2	4210	2000	104	C34	Male	6292
mort2	4210	2000	104	C34	Female	2261
						8553
mort3	4210	2000	104	C34		8553

# QC 3<sup>rd</sup> step in mortality

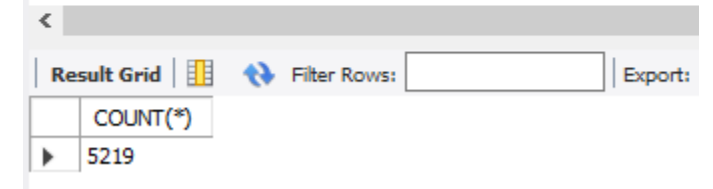
mortality2에는 남자, 여자 따로 사망자수가 기록되어 있다.

1. 남자 여자 합쳐서 사망자수가 기록되도록 한다.
2. 1번 결과값을 mortality3 table에 저장

Country_code	Year	List	Cause_code	Format	IM_Format	Deaths1
4210	1996	10M	A02	00	08	7
4210	1996	10M	A03	00	08	1

Deaths26	IM_Deaths1	IM_Deaths2	IM_Deaths3	IM_Deaths4
0	0	NULL	NULL	NULL
0	0	NULL	NULL	NULL
		NULL	NULL	NULL

```
1 SELECT COUNT(*)
2 FROM mortality3;
```



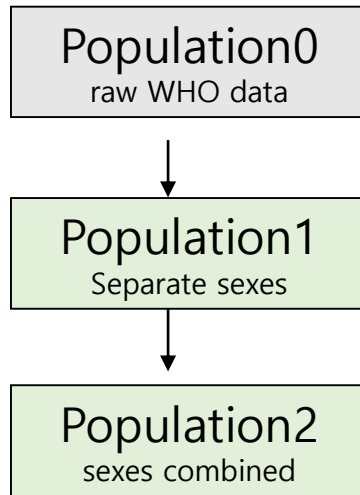
The screenshot shows a database interface with a query result grid. The query is 'SELECT COUNT(\*) FROM mortality3;'. The result grid displays a single row with the value '5219' under the column 'COUNT(\*)'. The interface includes a 'Result Grid' tab, a 'Filter Rows' input field, and an 'Export' button.

COUNT(*)
5219



# Population QC

# Quality Control – population table

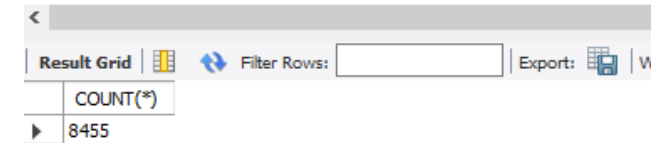


1. sub-divisions 또는 regions 정보가 있는 행은 모두 지우기

2. 남녀 정보 통합

# QC 1<sup>st</sup> step in population

```
1 SELECT COUNT(*)
2 FROM population1;
```



COUNT(*)
8455

1. population0 테이블에서 sub-divisions 또는 regions 정보가 있는 행은 모두 지우기
2. 1번의 결과를 population1 테이블로 저장

Country_code	Year	Sex	Frmat	Pop1	Pop2
1060	1980	1	7	137100	3400
1060	1980	2	7	159000	4000
1125	1955	1	2	5051500	150300

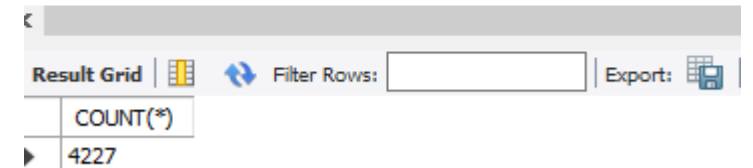
Pop25	Pop26	Lb
NULL	6500	5000
NULL	7500	6000
NULL	0	253329

# QC 2<sup>nd</sup> step in population

population1에는 남자, 여자 따로 사망자수가 기록되어 있다.

1. 남자 여자 합쳐서 사망자수가 기록되도록 한다.
2. 1번 결과값을 population2 table에 저장

```
1 SELECT COUNT(*)  
2 FROM population2;
```



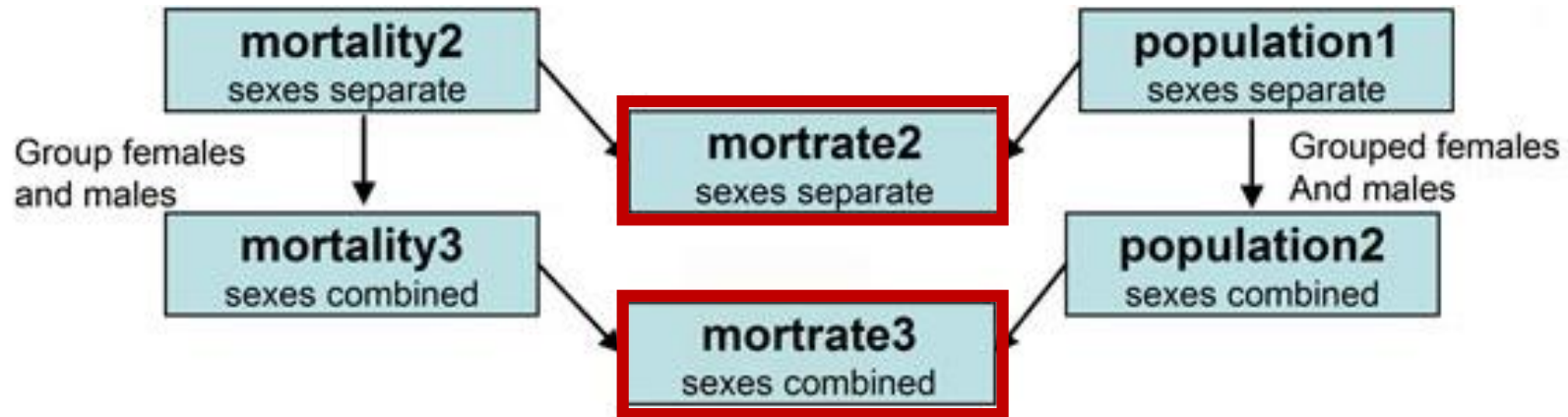
Result Grid | Filter Rows: | Export: |

COUNT(*)
4227

Country_code	Year	Format	Pop1	Pop2	Pop3
1060	1980	7	296100	7400	34200
1125	1955	2	10100900	295500	1094400
1125	1956	2	10705100	312300	1161400
1125	1957	2	10795000	315600	1170200
1125	1958	2	11001300	320600	1192100

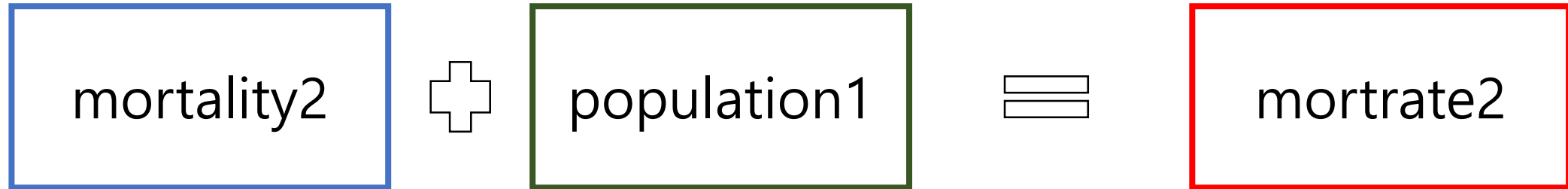
Pop22	Pop23	Pop24	Pop25	Pop26	Lb
NULL	NULL	NULL	NULL	14000	11000
29800	53800	NULL	NULL	0	491230
31400	56000	NULL	NULL	0	485909
31100	56300	NULL	NULL	0	447223
31300	56900	NULL	NULL	0	493709

# Quality Control 과정

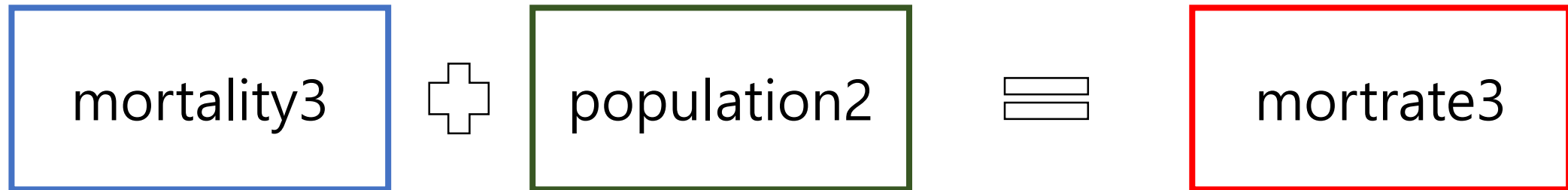


# Mortrate table 만들기 (join)

mortrate2: 국가, 연도, 사망원인, 성별, 나이별 사망률



mortrate3: 국가, 연도, 사망원인, 나이별 사망률



**mortrate=100000 \* (Deaths/Pop)**

mortrate2

country_code	year	List	Cause_code	Sex	Frmat	MortRate1
4210	1996	10M	A02	1	00	0.0651
4210	1996	10M	A03	1	00	0.0130
4210	1996	10M	A04	1	00	0.0130

MortRate23	MortRate24	MortRate25	MortRate26
0.0000	NULL	NULL	NULL
0.0000	NULL	NULL	NULL
0.0000	NULL	NULL	NULL

mortrate3

country_code	year	List	Cause_code	Frmat	MortRate1
4210	1996	10M	A02	00	0.0450
4210	1996	10M	A03	00	0.0064

MortRate23	MortRate24	MortRate25	MortRate26
0.4786	NULL	NULL	NULL
0.0000	NULL	NULL	NULL

```
1 SELECT COUNT(*)
2 FROM mortrate2;
3
```

Result Grid | Filter Rows: | Exports: | Wrap Cell Content: |

COUNT(*)
8760

```
4 SELECT COUNT(*)
5 FROM mortrate3;
```

Result Grid | Filter Rows: |

COUNT(*)
5219

# WHO DB 분석



# 논문 분석을 위한 Final Table

Mortrate3 를 통해 만들 Final table

whoregion	name	pesticide	other_poison	hanging	drowning	firearms	jumping
AFR	Mauritius	11.9502	0.8226	33.6238	0.4904	0.4939	0.1645
AFR	Rodrigues	8.0365	0.0000	23.9804	0.0000	0.0000	0.0000
AFR	South Africa	0.2154	0.3475	1.6405	0.0102	0.1840	0.0080
AMR	Canada	0.2814	17.1748	30.3154	1.9346	11.9188	3.4814
AMR	United States of America	0.3165	16.8653	20.8614	1.0947	52.4990	2.1488
EMR	Bahrain	0.1041	0.0962	30.2487	0.6169	0.0000	0.0000
EMR	Cyprus	3.5500	1.9875	8.1034	0.3599	4.6451	4.7333
EMR	Egypt	0.1311	0.0606	0.2309	0.0123	0.2999	0.0015
EMR	Jordan	0.0170	0.0671	0.1507	0.0000	0.0678	0.0000

1. Country, year, cause\_code별 사망률을  
Country, cause\_code별 사망률로 바꾸기  
(연도별 사망률 합하기)

```
USE who_mortality;

-- Combine annual mortality rates
DROP TABLE IF EXISTS analysis0;
CREATE TABLE analysis0 AS (
SELECT Country_code, Cause_code, SUM(MortRate1) AS MortRate1
FROM real_mortrate3
GROUP BY Country_code, Cause_code );|
```

analysis0 table

Country_code	Cause_code	MortRate1
1060	A000	770.3478
1060	A001	0.0000
1060	A002	6.0790
1060	A003	2.0263
1060	A004	1.6886
1060	A005	93.5494
1060	A006	12.1580
1060	A007	0.0000
1060	A008	0.0000
1060	A009	0.0000
1060	A010	0.6754
1060	A011	0.0000

## 2. 자살과 관련된 cause\_code만 추출

```
-- Extract cause_codes associated with suicide
DROP TABLE IF EXISTS analysis1;
CREATE TABLE analysis1 AS (
SELECT *
FROM analysis0
WHERE (Cause_code REGEXP 'X6[0-9]')
      OR (Cause_code REGEXP 'X7[0-5]')
      OR (Cause_code = 'X80')
);

SELECT DISTINCT Cause_code FROM analysis1;
```

analysis1 table

Country_code	Cause_code	MortRate1
1125	X60	0.0040
1125	X62	0.0094
1125	X65	0.0112
1125	X66	0.0097
1125	X67	0.0079

The ICD-10 codes for suicide are in the range X60–X849. We differentiated suicides due to a pesticide or an unspecified poison (X68–X699), other poisons (X60–X679), hanging (X70–X709), drowning (X71–X719), firearms and explosives (X72–X759), and jumping from a height (X80–X809), and other suicide methods. Some methods, such as hanging or firearm suicide, are reported more accurately than others, such as poisoning or drowning.<sup>12</sup> This should be kept in mind when interpreting frequency distributions, particularly in countries with incomplete data.

- pesticide or unspecified poison: X68-X69
- other poison: X60-X67
- hanging: X70
- drowning: X71
- firearms and explosives: X72-X75
- jumping from height: X80

### 3. 비슷한 방법끼리 같은 카테고리로 만들기

analysis2 table

Country_code	Cause_code	MortRate1	Suicide_method
1125	X60	0.0040	other_poison
1125	X62	0.0094	other_poison
1125	X65	0.0112	other_poison
1125	X66	0.0097	other_poison
1125	X67	0.0079	other_poison
1125	X68	0.0904	pesticide
1125	X69	0.0407	pesticide
1125	X70	0.2309	hanging
1125	X73	0.0205	firearms
1125	X74	0.2652	firearms
1125	X80	0.0015	jumping
1125	X64	0.0089	other_poison

```
-- Converting the same type of suicide method into one category
DROP TABLE IF EXISTS analysis2;
CREATE TABLE analysis2 AS (
SELECT *,
    CASE WHEN Cause_code REGEXP 'X6[89]' THEN 'pesticide'
         WHEN Cause_code REGEXP 'X6[0-7]' THEN 'other_poison'
         WHEN Cause_code = 'X70' THEN 'hanging'
         WHEN Cause_code = 'X71' THEN 'drowning'
         WHEN Cause_code REGEXP 'X7[2-5]' THEN 'firearms'
         WHEN Cause_code = 'X80' THEN 'jumping'
    END AS Suicide_method
FROM analysis1
);

SELECT * FROM analysis2;
```

- pesticide or unspecified poison: X68-X69
- other poison: X60-X67
- hanging: X70
- drowning: X71
- firearms and explosives: X72-X75
- jumping from height: X80

## 4. 자살 방법별 사망률 계산

같은 그룹의 사망률 합산

```
DROP TABLE IF EXISTS analysis3;  
CREATE TABLE analysis3 AS (  
SELECT a.Country_code, a.Suicide_method, SUM(MortRate1) AS MortRate1  
FROM analysis2 AS a  
GROUP BY a.Country_code, a.Suicide_method  
);  
SELECT * FROM analysis3;
```

analysis3 table

Country_code	Suicide_method	MortRate1
1125	other_poision	0.0606
1125	pesticide	0.1311
1125	hanging	0.2309
1125	firearms	0.2999
1125	jumping	0.0015
1125	drowning	0.0123
1300	other_poision	0.8226
1300	pesticide	11.9502

# 5. Reformat

```
-- final output
DROP TABLE IF EXISTS analysis4;
CREATE TABLE analysis4 AS (
SELECT Country_code,
       SUM(CASE WHEN Suicide_method='pesticide' THEN MortRate1 ELSE 0 END) AS pesticide,
       SUM(CASE WHEN Suicide_method='other_poison' THEN MortRate1 ELSE 0 END) AS other_poison,
       SUM(CASE WHEN Suicide_method='hanging' THEN MortRate1 ELSE 0 END) AS hanging,
       SUM(CASE WHEN Suicide_method='drowning' THEN MortRate1 ELSE 0 END) AS drowning,
       SUM(CASE WHEN Suicide_method='firearms' THEN MortRate1 ELSE 0 END) AS firearms,
       SUM(CASE WHEN Suicide_method='jumping' THEN MortRate1 ELSE 0 END) AS jumping
FROM analysis3
GROUP BY Country_code
);

SELECT * FROM analysis4;
```

## analysis4 table

Country_code	pesticide	other_poison	hanging	drowning	firearms	jumping
1125	0.1311	0.0606	0.2309	0.0123	0.2999	0.0015
1300	11.9502	0.8226	33.6238	0.4904	0.4939	0.1645
1365	8.0365	0.0000	23.9804	0.0000	0.0000	0.0000
1430	0.2154	0.3475	1.6405	0.0102	0.1840	0.0080

# 6. Final Output

- 국가의 whoregion 정보와 이름 추가
- whoregion, name 기준으로 table 정렬

```
DROP TABLE IF EXISTS analysis5;
CREATE TABLE analysis5 AS (
SELECT c.whoregion, c.name, a.pesticide, a.other_poison, a.hanging, a.drowning, a.firearms,
a.jumping
FROM analysis4 as a
LEFT JOIN country as c ON a.Country_code=c.Country_code
ORDER BY c.whoregion, c.name
);

SELECT *
FROM analysis5;

-- write table to csv file
SELECT * FROM analysis5 INTO OUTFILE 'C:\\ProgramData\\MySQL\\MySQL Server 8.0\\Uploads\\
\\analysis5.csv' FIELDS TERMINATED BY ',' LINES TERMINATED BY '\\n' ;
-- SHOW VARIABLES LIKE "secure_file_priv";
```

## analysis5 table

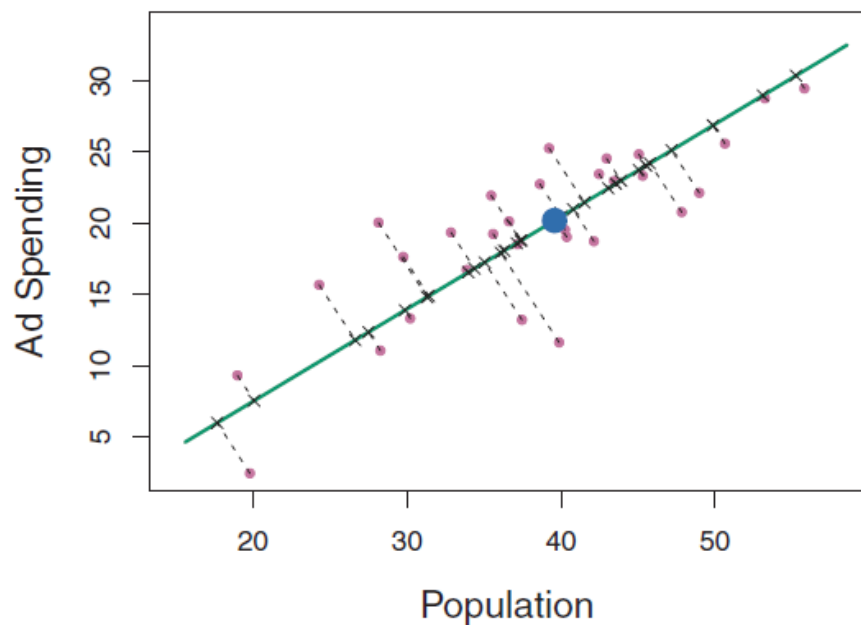
whoregion	name	pesticide	other_poison	hanging	drowning	firearms	jumping
AFR	Mauritius	11.9502	0.8226	33.6238	0.4904	0.4939	0.1645
AFR	Rodrigues	8.0365	0.0000	23.9804	0.0000	0.0000	0.0000
AFR	South Africa	0.2154	0.3475	1.6405	0.0102	0.1840	0.0080
AMR	Canada	0.2814	17.1748	30.3154	1.9346	11.9188	3.4814



기계학습을 이용한 분석

# Principal Component Analysis (PCA)

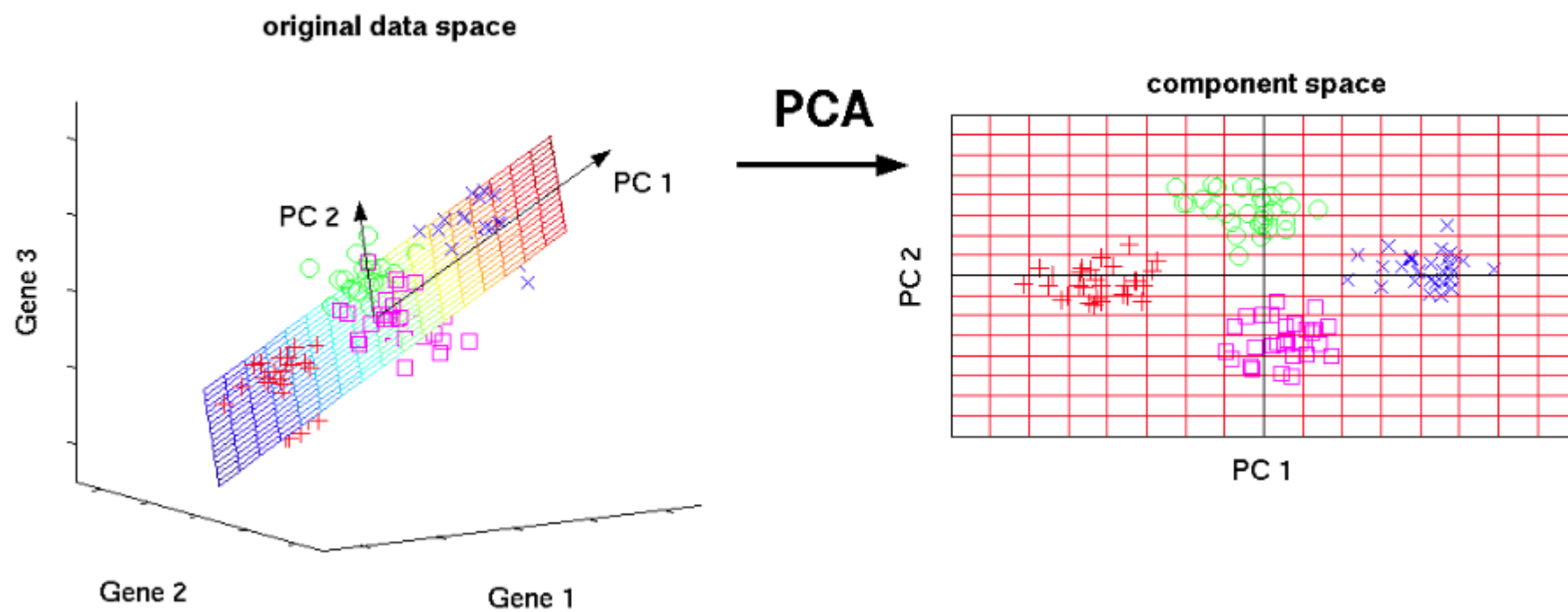
- 차원 축소
- 고차원에서 저차원으로 비슷한 패턴끼리 묶음
- 새로운 축을 만듦 -> PC1, PC2 ,...



예제)

- 2차원을 1차원으로 축소
- 초록색 선이 새로운 축(PC1)

# PCA



# Suicide 분석 – PCA

# suicide pattern 분석

- data: suicide\_mortrate2\_final.csv 이용

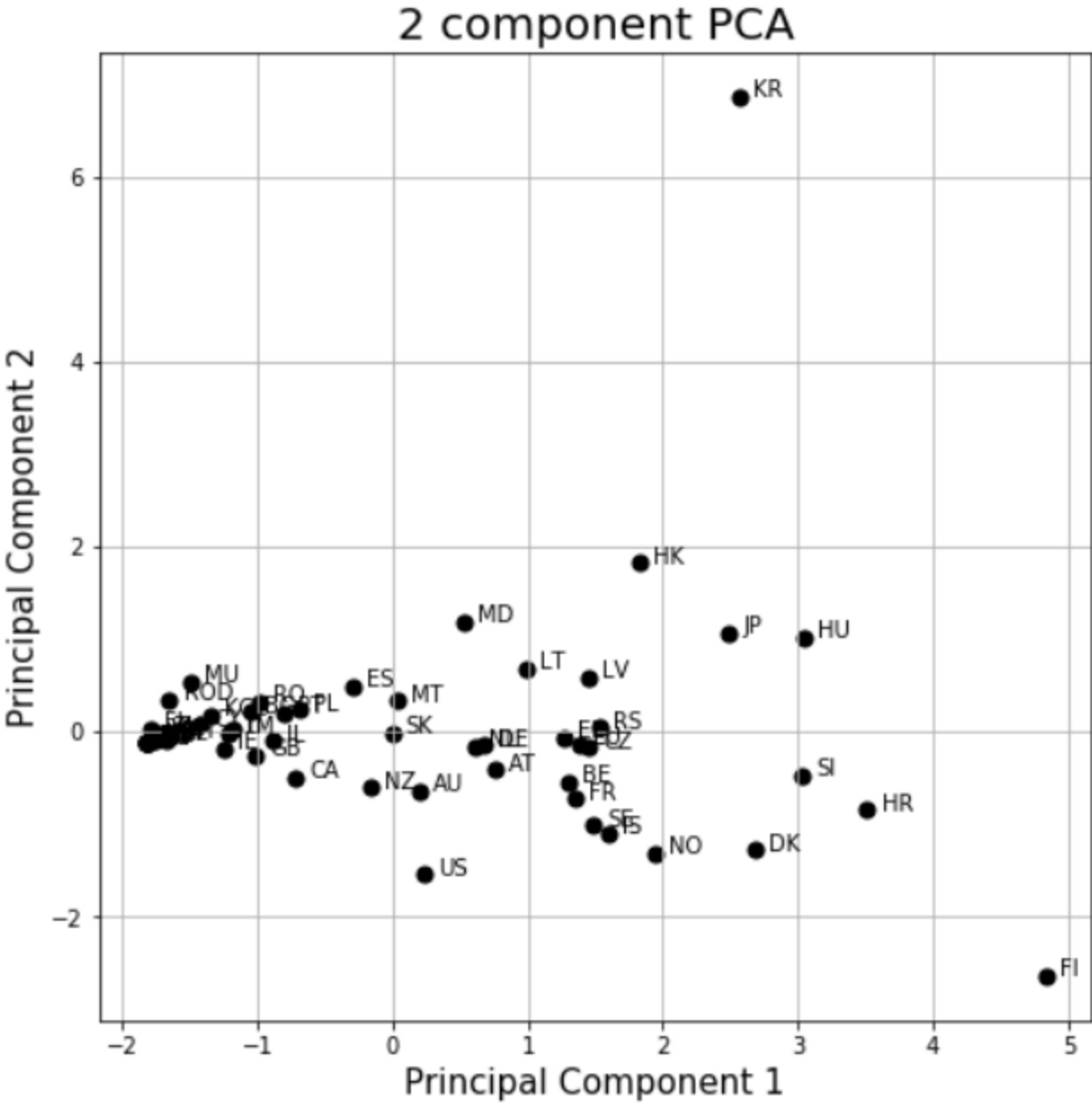
(전체 62개의 행( analysis5 )에서 동일 국가의 다른 지역 데이터 삭제 후,  
총 58개의 국가에 대한 table이다.)

- 분석 시, code 작성 위치와 csv 파일 위치 동일한 곳에서 하기
- Python으로 분석

# Python으로 분석

- 1) `pip install -U matplotlib`
- 2) `pip install -U scikit-learn`
- 3) `pip install pandas`

코드 [https://github.com/zlsisp54/WHO\\_Sucide\\_Analysis](https://github.com/zlsisp54/WHO_Sucide_Analysis)



Python으로 분석

