

산보연 Data check 6월 29일 ver

작성자 : 이은경

----- Update 사항 : data version2 에서 version4로 변경 -----

백혈병 Data

1. 사업장 별 “통합 누적 발생률”과 대조군의 “통합 누적 발생률” 비교

[대조군 생성 계획]

: UP2 상관없이, SEX, CAL, ECNY_AGE, YEAR가 똑같은 집단에 대한 백혈병 발생 건수, 추적인년합계 변수 aggregate

-> 이를 UP2, SEX, CAL, ECNY_AGE, YEAR에 따라 grouping한 데이터와 연결

① grouping key ; UP2, SEX, CAL, ECNY-AGE, YEAR

| UP2 | SEX | CAL | ECNY-AGE | YEAR | 통합누적발생률 (파란색) | Sum -FY-BZ | Sum -leukemia-BZ |
|-----|-----|-----|----------|------|---------------|---------------|---------------------|
| | | | | | | | |

② 대조군 ; grouping key는 ①의 grouping key에서 "UP2" 제외 전복!

| SEX | CAL | ECNY-AGE | YEAR | Sum -FY-BZ-2 | Sum -leukemia 2 | Weighted 통합누적발생률 |
|-----|-----|----------|------|-----------------|--------------------|--|
| | | | | | | $= w_i \times \frac{\text{Sum-leukemia 2}}{\text{Sum-FY-BZ 2}}$ <p>(이때 $w_i = \frac{\text{Sum-FY-BZ}}{\text{Sum-FY-BZ-2}}$)</p> |

①, ② 병합 후 생성하는 변수

⇒ ① +tblle 라 ② +tblle 을 SEX, CAL, ECNY-AGE, YEAR 변수 기준으로 병합, 그리고

"Weighted 통합누적발생률" 변수 생성

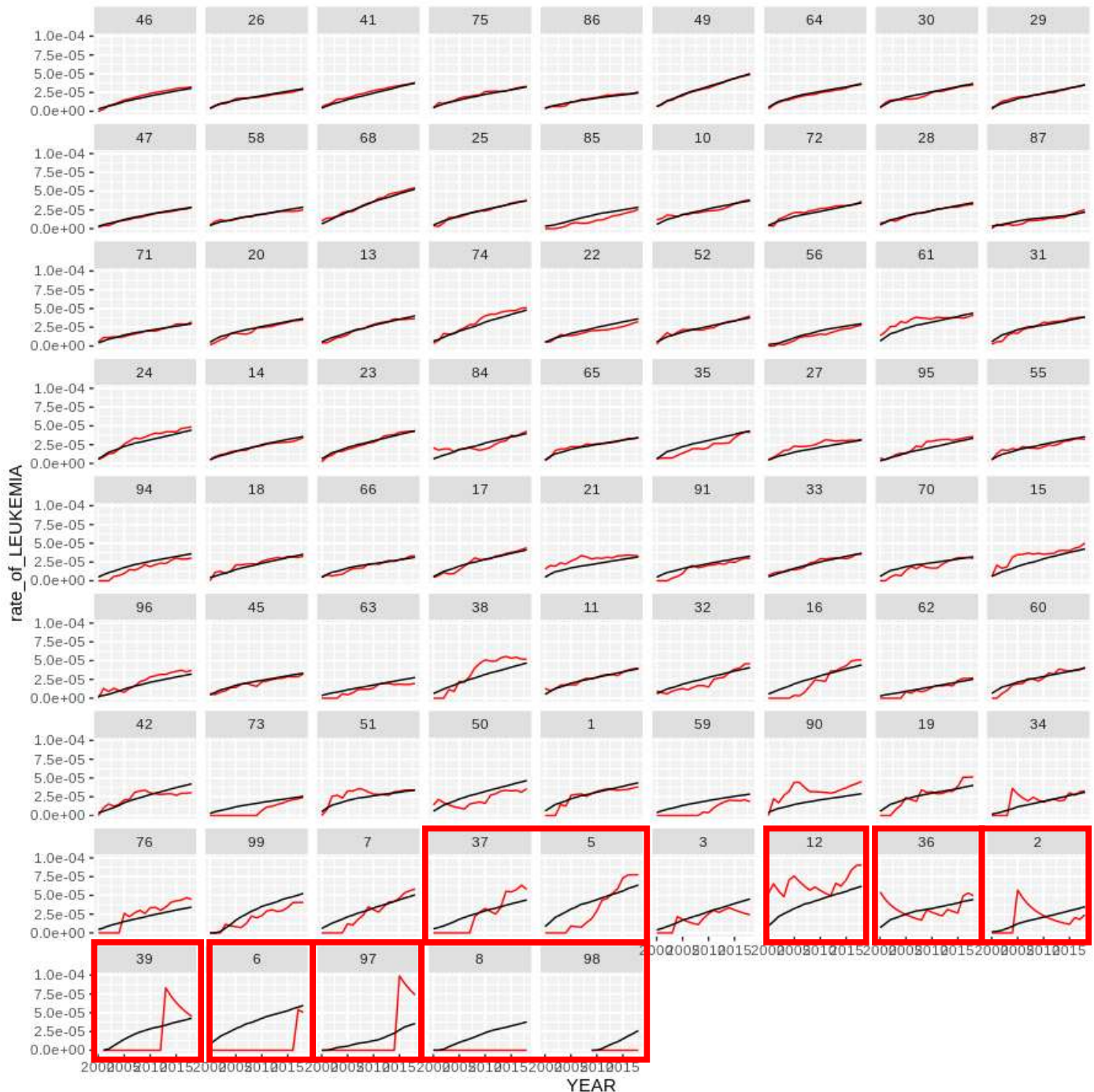
⇒ 위의 계획처럼, 대조군 생성 후에 통합 누적 발생률 계산. 각 사업장에 대한 통합 누적 발생률("origin_rate" 변수명)과 대조군의 통합 누적 발생률("weight_rate" 변수명) 겹쳐서 시각화

----- 다음 페이지로 -----

① y축의 scale을 조정하지 않은 경우

: 이때, 그래프의 순서는 2018년 기준 추적 인년 합계의 순위(오름차순 정렬)

: 빨간색이 사업장 별 통합 누적 발생률, 검은색이 대조군의 통합 누적 발생률 의미



⇒ 위 그래프에서 대조군(검은색)의 통합 누적 발생률과 큰 차이를 보이는 그래프가 빨간색 테두리가 있는 부분이다. 경향을 보면, 추적 인년 합계가 적은 사업장일수록 대조군과의 차이가 두드러져 보이는 것을 알 수 있다. 해당 사업장명은 다음과 같다.

| UP2 | 사업장명 |
|-----|--------------------------|
| 37 | 하수, 폐기물 정화업 (하수, 폐수, 분뇨) |
| 5 | 광업(석탄, 원유, 채굴업) |
| 12 | 제조업(담배) |
| 36 | 용수 공급업 |
| 2 | 임업 |
| 39 | 환경 정화업 |
| 6 | 광업(금속, 철, 비금속) |
| 97 | 가구내 고용 활동 |
| 8 | 광업(광업, 자원, 원유) |
| 98 | 가사 생산 활동 |

② y축의 scale을 조정한 경우

: 이때, 그래프의 순서는 2018년 기준 추적인년합계의 순위(오름차순 정렬)

: 빨간색이 사업장 별 통합누적발생률, 검은색이 대조군의 통합누적발생률 의미



⇒ 위 그래프에도 대조군(검은색)의 통합 누적 발생률과 큰 차이를 보이는 그래프를 빨간색으로 표시했다. 여기서도 경향을 보면, 추적 인년 합계가 적은 사업장일수록 대조군과의 차이가 두드러져 보이는 것을 알 수 있다. 해당 사업장명은 다음과 같다. (파란색으로 표시한 부분은 y축의 scale을 고려하지 않았던 ①의 결과와 겹치는 사업장이다.)

| UP2 | 사업장명 |
|-----|----------------------------------|
| 21 | 의료용 물질 및 의약품 제조업 |
| 38 | 하수, 폐기물 정화업(방사성 물질, 폐기물, 건설 폐기물) |
| 73 | 수익업, 디자인업 |
| 51 | 항공 운수업 |
| 50 | 수상 운수업 |
| 59 | 영화, 비디오, 오디오물 출판, 상영업 |
| 90 | 예술, 스포츠업, 건설업 등 각종 사업 |
| 34 | 산업용 기계 및 장비 수리업 |
| 76 | 임대업; 부동산 제외 |
| 37 | 하수, 폐기물 정화업(하수, 폐수, 분뇨) |
| 5 | 광업(석탄, 원유, 채굴업) |

| | |
|----|----------------|
| 12 | 제조업(담배) |
| 36 | 용수 공급업 |
| 2 | 임업 |
| 39 | 환경 정화업 |
| 6 | 광업(금속, 철, 비금속) |
| 97 | 가구내 고용 활동 |
| 8 | 광업(광업, 자원, 원유) |
| 98 | 가사 생산 활동 |

2. 각 사업장 별 통합 누적 발생률과 대조군의 통합 누적 발생률 “차이”(Relative Risk)가 가장 큰 TOP10 사업장은?

| UP2 | 사업장명 | 발생률 차이 |
|-----|-------------------------|-----------|
| 97 | 가구 내 고용활동 | 0.0000380 |
| 8 | 광업(광업, 자원, 원유) | 0.0000378 |
| 12 | 제조업(담배) | 0.0000283 |
| 98 | 가사 생산 활동 | 0.0000262 |
| 3 | 어업 | 0.0000214 |
| 90 | 예술, 스포츠업, 건설업 등 각종 사업 | 0.0000163 |
| 37 | 하수, 폐기물 정화업(하수, 폐수, 분뇨) | 0.0000141 |
| 5 | 광업(석탄, 원유, 채굴업) | 0.0000137 |
| 99 | 국제 및 외국기관 | 0.0000119 |
| 42 | 건설업(건물 구축 관련) | 0.0000117 |

3. 각 사업장 별 통합 누적 발생률과 대조군의 통합 누적 발생률의 “비율”(간접 SIR)가 가장 큰 TOP10 사업장은?

| UP2 | 사업장명 | 간접 SIR |
|-----|-------------------------|--------|
| 97 | 가구 내 고용활동 | 2.06 |
| 90 | 예술, 스포츠업, 건설업 등 각종 사업 | 1.56 |
| 12 | 제조업(담배) | 1.45 |
| 37 | 하수, 폐기물 정화업(하수, 폐수, 분뇨) | 1.32 |
| 76 | 임대업; 부동산 제외 | 1.30 |
| 19 | 제조업(석유 정제품, 코크스 관련) | 1.28 |
| 5 | 광업(석탄, 원유, 채굴업) | 1.21 |
| 15 | 가족 관련 제조업 | 1.19 |
| 7 | 광업(비금속 광물, 소금 등) | 1.15 |
| 16 | 제조업(목재 관련) | 1.15 |

4. 백혈병 통합 누적 발생률 TOP 10인 사업장은?

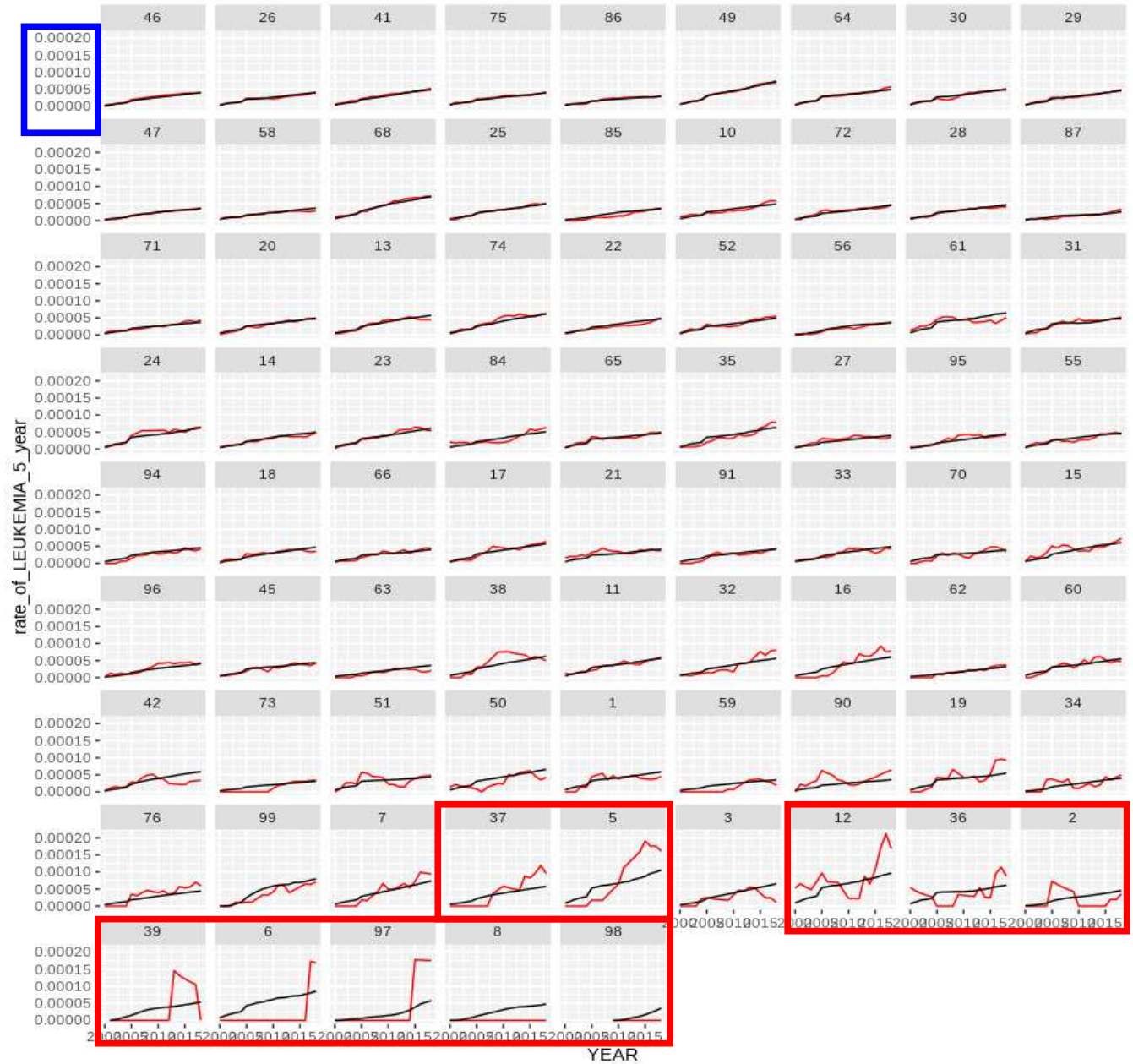
| UP2 | 사업장명 | 통합 누적 발생률 |
|-----|----------------------------------|-----------|
| 12 | 제조업(담배) | 0.0000905 |
| 5 | 광업(석탄, 원유, 채굴업) | 0.0000776 |
| 97 | 가구내 고용활동 | 0.0000738 |
| 7 | 광업(비금속 광물, 소금 등) | 0.0000584 |
| 37 | 하수, 폐기물 정화업(하수, 폐수, 분뇨) | 0.0000580 |
| 68 | 부동산업 | 0.0000546 |
| 38 | 하수, 폐기물 정화업(방사성 물질, 폐기물, 건설 폐기물) | 0.0000520 |
| 19 | 제조업(석유 정제품, 코크스 관련) | 0.0000515 |
| 74 | 사업 건물 관리, 조경 관리 서비스업 | 0.0000511 |
| 16 | 제조업(유리, 목재 등) | 0.0000509 |

5. 각 사업장 별로, 대조군 별로 5년 단위 통합 누적 발생률 계산 후 비교

: 계산 후, 추적 인년 합계 기준으로 오름차순 정렬 뒤, 시각화(사업장 별 5년 단위 통합 누적 발생률과 대응하는 대조군의 5년 단위 통합 누적 발생률 겹쳐서 그림)

: **빨간색이 사업장 별** 5년 단위 통합 누적 발생률, **검은색이 대조군의 5년 단위 통합 누적 발생률** 의미

① y축의 scale을 조정하지 않은 경우 : y축의 scale이 1번 버전 그래프(5년 단위로 나누지 않음)와 **다름**.



⇒ 위 그래프에서 대조군(검은색)의 통합 누적 발생률과 큰 차이를 보이는 그래프 그래프에 **빨간색 사각형**으로 표시하였다. 여기서도 경향을 보면, 추적 인년 합계가 적은 사업장일수록 대조군과의 차이가 두드러져 보이는 것을 알 수 있다. 해당 사업장명은 다음과 같다.

| UP2 | 사업장명 |
|-----|--------------------------|
| 37 | 하수, 폐기물 정화업 (하수, 폐수, 분뇨) |
| 5 | 광업(석탄, 원유, 채굴업) |
| 12 | 제조업(담배) |
| 36 | 용수 공급업 |
| 2 | 임업 |
| 39 | 환경 정화업 |
| 6 | 광업(금속, 철, 비금속) |
| 97 | 가구내 고용 활동 |
| 8 | 광업(광업, 자원, 원유) |
| 98 | 가사 생산 활동 |

② y축의 scale을 조정한 경우



⇒ 위 그래프에서 대조군(검은색)의 통합 누적 발생률과 큰 차이를 보이는 그래프에 빨간색 사각형으로 표시하였다.
(y축의 scale이 달라지는 점을 주목해서 볼 필요가 있다.) 여기서도 경향을 보면, 초적 인년 합계가 적은 사업장일수록 대조군과의 차이가 두드러져 보이는 것을 알 수 있다. 해당 사업장명은 다음과 같다. (파란색으로 표시한 부분은 y축의 scale을 고려하지 않았던 ①의 결과와 겹치는 사업장이다.)

| UP2 | 사업장명 |
|-----|----------------------------------|
| 21 | 의료용 물질 및 의약품 제조업 |
| 96 | 기타 개인 서비스업 |
| 38 | 하수, 폐기물 정화업(방사성 물질, 폐기물, 건설 폐기물) |
| 42 | 전문직별 공사업 |
| 73 | 수익업, 디자인업 |
| 51 | 항공 운수업 |
| 50 | 수상 운수업 |
| 1 | 농업 |
| 59 | 영화, 비디오, 오디오물 출판, 상영업 |
| 90 | 예술, 스포츠업, 건설업 등 각종 사업 |

| | |
|----|-------------------------|
| 19 | 제조업(석유정제품, 코크스 관련) |
| 34 | 산업용 기계 및 장비 수리업 |
| 76 | 임대업; 부동산 제외 |
| 37 | 하수, 폐기물 정화업(하수, 폐수, 분뇨) |
| 5 | 광업(석탄, 원유, 채굴업) |
| 3 | 어업 |
| 12 | 제조업(담배) |
| 36 | 용수 공급업 |
| 2 | 임업 |
| 39 | 환경 정화업 |
| 6 | 광업(금속, 철, 비금속) |
| 97 | 가구내 고용 활동 |
| 8 | 광업(광업, 자원, 원유) |
| 98 | 가사 생산 활동 |

6. 각 사업장 별 5년 단위 통합 누적 발생률과 대조군의 5년 단위 통합 누적 발생률의 “비율”(간접 SIR)가 가장 큰 TOP10 사업장은?

| UP2 | YEAR | 사업장명 | SIR |
|-----|------|-----------------|------|
| 36 | 2000 | 용수 공급업 | 7.87 |
| 12 | 2000 | 제조업(담배) | 5.64 |
| 2 | 2005 | 임업 | 4.80 |
| 97 | 2015 | 가구 내 고용 활동 | 4.61 |
| 12 | 2001 | 제조업(담배) | 4.27 |
| 34 | 2004 | 산업용 기계 및 장비 수리업 | 4.01 |
| 97 | 2016 | 가구 내 고용 활동 | 3.71 |
| 36 | 2001 | 용수 공급업 | 3.67 |
| 39 | 2013 | 환경 정화업 | 3.60 |
| 96 | 2001 | 개인 사업 | 3.60 |

: 파란색으로 표시한 부분은 3번 결과(연도 별로 나누지 않고 계산한 간접 SIR)와 겹치는 사업장 의미

7. 백혈병 5년 단위 통합 누적 발생률 TOP 10인 사업장은? (연도 고려하지 않음.)

| UP2 | YEAR | 사업장명 | 5년 단위 통합 누적 발생률 |
|-----|------|-----------------|-----------------|
| 12 | 2017 | 제조업(담배) | 0.000212 |
| 5 | 2015 | 광업(석탄, 원유, 채굴업) | 0.000191 |
| 97 | 2015 | 가구 내 고용 활동 | 0.000177 |
| 97 | 2016 | 가구 내 고용 활동 | 0.000177 |
| 97 | 2017 | 가구 내 고용 활동 | 0.000176 |
| 5 | 2017 | 광업(석탄, 원유, 채굴업) | 0.000176 |
| 5 | 2016 | 광업(석탄, 원유, 채굴업) | 0.000175 |
| 97 | 2018 | 가구 내 고용 활동 | 0.000175 |
| 6 | 2017 | 광업(금속, 철, 비금속) | 0.000173 |
| 12 | 2016 | 제조업(담배) | 0.000171 |

: 파란색으로 표시한 부분은 4번 결과(2018년 기준 통합 누적 발생률)와 겹치는 사업장 의미

8. 2018년 기준 백혈병 5년 단위 통합 누적 발생률 TOP10은?

| UP2 | 사업장명 | 2018년 기준 통합 누적 발생률 |
|-----|-------------------------|--------------------|
| 97 | 가구 내 고용 활동 | 0.000175 |
| 6 | 광업(금속, 철, 비금속) | 0.000169 |
| 12 | 제조업(담배) | 0.000168 |
| 5 | 광업(석탄, 원유, 채굴업) | 0.000161 |
| 37 | 하수, 폐기물 정화업(하수, 폐수, 분뇨) | 0.0000954 |
| 7 | 광업(비금속 광물, 소금) | 0.0000936 |
| 19 | 제조업(석유 정제품, 코크스 관련) | 0.0000928 |
| 36 | 용수 공급업 | 0.0000879 |
| 32 | 제조업(금속 가구, 목재 가구) | 0.0000802 |
| 35 | 전기, 가스, 증기 및 공기 조절 공급업 | 0.0000788 |

: 빨간색으로 표시한 부분은 7번 결과(연도 고려하지 않은 5년 단위 통합 누적 발생률)와 겹치는 사업장 의미

9. 2018년 기준 각 사업장 별 5년 단위 통합 누적 발생률과 대조군의 5년 단위 통합 누적 발생률의 “비율”(간접 SIR)가 가장 큰 TOP10 사업장은?

| UP2 | 사업장명 | SIR |
|-----|-------------------------|------|
| 97 | 가구 내 고용활동 | 3.05 |
| 6 | 광업(금속, 철, 비금속) | 1.98 |
| 90 | 예술, 스포츠업, 건설업 등 각종 사업 | 1.78 |
| 12 | 제조업(담배) | 1.74 |
| 19 | 제조업(석유 정제품, 코크스 관련) | 1.69 |
| 37 | 하수, 폐기물 정화업(하수, 폐수, 분뇨) | 1.65 |
| 5 | 광업(석탄, 원유, 채굴업) | 1.52 |
| 36 | 용수 공급업 | 1.44 |
| 32 | 가구 제조업 | 1.43 |
| 76 | 임대업; 부동산 제외 | 1.35 |

: 빨간색으로 표시한 부분은 6번 결과(연도 상관없이 구한 간접 SIR)와 겹치는 사업장 의미

 ##### 폐암 data #####

1. 사업장 별 “통합 누적 발생률”과 대조군의 “통합 누적 발생률” 비교

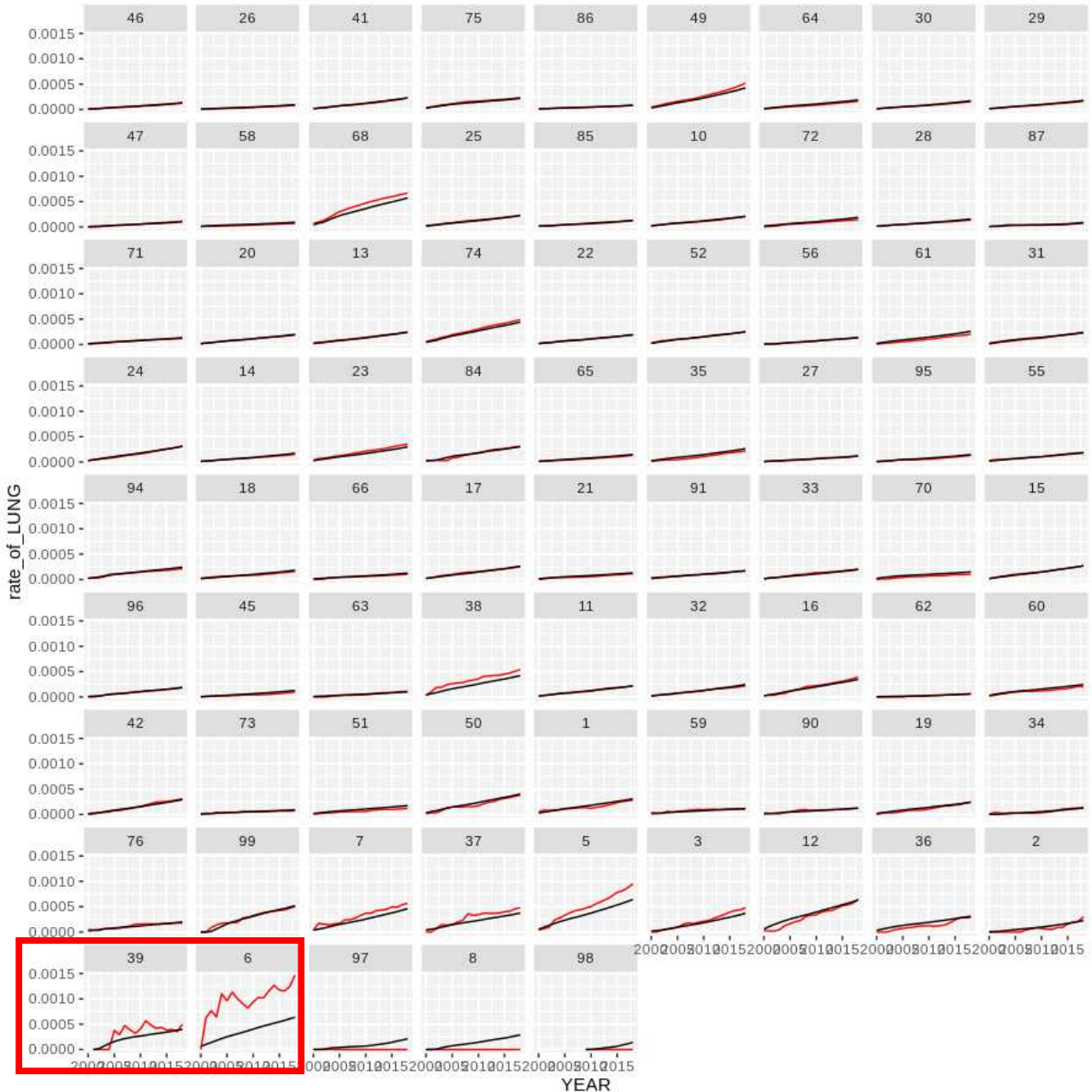
: 이전 백혈병 data에서 적용했던 방식과 동일한 방식 이용

⇒ 대조군 생성 후에 통합 누적 발생률 계산. 각 사업장에 대한 통합 누적 발생률(“origin_rate” 변수명)과 대조군의 통합 누적 발생률(“weight_rate” 변수명) 겹쳐서 시각화

① y축의 scale을 조정하지 않은 경우

: 이때, 그래프의 순서는 2018년 기준 추적 인년 합계의 순위(오름차순 정렬)

: 빨간색이 사업장 별 통합 누적 발생률, 검은색이 대조군의 통합 누적 발생률 의미



⇒ 백혈병 data보다는 추적 인년 합계가 적어도 안정된 추세를 보이는 것을 알 수 있다. 위 그래프에서 대조군(검은색)의 통합 누적 발생률과 큰 차이를 보이는 그래프에 빨간색 사각형으로 표시하였고, 해당 사업장 명은 아래 표에 제시하였다.

| | |
|-----|----------------|
| UP2 | 사업장명 |
| 39 | 환경 정화업 |
| 6 | 광업(금속, 철, 비금속) |

② y축의 scale을 조정한 경우

: 이때, 그래프의 순서는 2018년 기준 추적 인년 합계의 순위(오름차순 정렬)

: 빨간색이 사업장 별 통합 누적 발생률, 검은색이 대조군의 통합 누적 발생률 의미



⇒ 백혈병 data보다는 추적 인년 합계가 적어도 안정된 추세를 보이는 것을 알 수 있다. 위 그래프에서 대조군(검은색)의 통합 누적 발생률과 큰 차이를 보이는 그래프에 빨간색 사각형으로 표시하였고, 해당 사업장명은 아래 표에 제시하였다. (파란색으로 표시한 부분은 y축의 scale을 고려하지 않았던 ①의 결과와 겹치는 사업장이다.)

| UP2 | 사업장명 |
|-----|-------------------------|
| 37 | 하수, 폐기물 정화업(하수, 폐수, 분뇨) |
| 39 | 환경 정화업 |
| 6 | 광업(금속, 철, 비금속) |
| 97 | 가구내 고용 활동 |
| 8 | 광업(광업, 자원, 원유) |
| 98 | 가사 생산 활동 |

2. 각 사업장 별 통합 누적 발생률과 대조군의 통합 누적 발생률의 “비율”(간접 SIR)가 가장 큰 TOP10 사업장은?

| UP2 | 사업장명 | 간접 SIR |
|-----|----------------------------------|--------|
| 6 | 광업(금속, 철, 비금속) | 2.30 |
| 5 | 광업(석탄, 원유, 채굴업) | 1.48 |
| 2 | 임업 | 1.33 |
| 3 | 어업 | 1.30 |
| 37 | 하수, 폐기물 정화업(하수, 폐수, 분뇨) | 1.28 |
| 38 | 하수, 폐기물 정화업(방사성 물질, 폐기물, 건설 폐기물) | 1.28 |
| 39 | 환경 정화 및 복원업 | 1.25 |
| 7 | 광업(비금속 광물, 소금) | 1.23 |
| 49 | 육상 운송 및 파이프라인 운송업 | 1.21 |
| 23 | 비금속 광물 제품 제조업 | 1.18 |

3. 폐암 통합 누적 발생률 TOP 10인 사업장은?

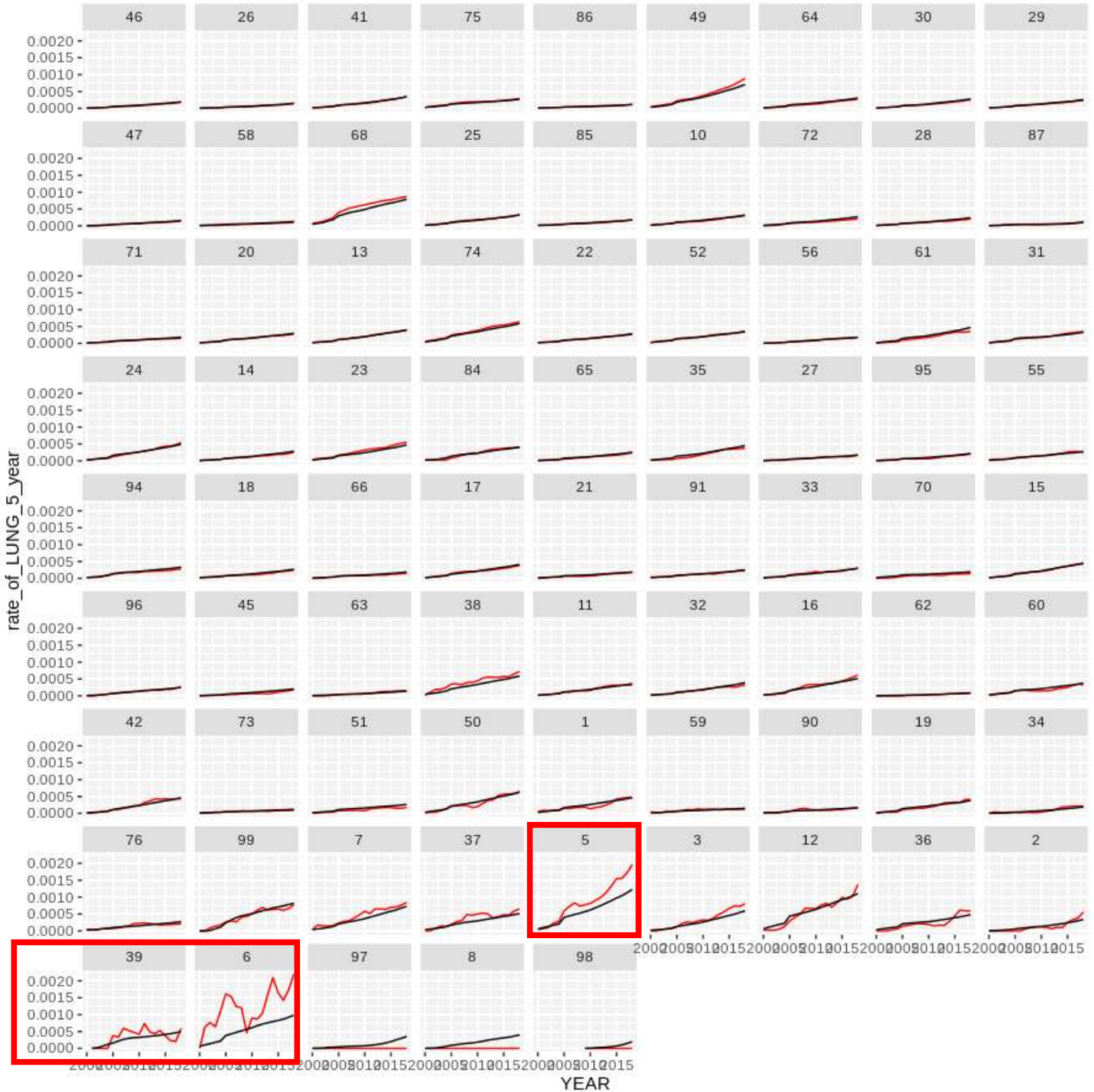
| UP2 | 사업장명 | 통합 누적 발생률 |
|-----|----------------------------------|-----------|
| 6 | 광업(금속, 철, 비금속) | 0.00147 |
| 5 | 광업(석탄, 원유, 채굴업) | 0.000951 |
| 68 | 부동산업 | 0.000670 |
| 12 | 제조업(담배) | 0.000650 |
| 7 | 광업(비금속 광물, 소금) | 0.000566 |
| 38 | 하수, 폐기물 정화업(방사성 물질, 폐기물, 건설 폐기물) | 0.000542 |
| 49 | 육상 운송 및 파이프라인 운송업 | 0.000517 |
| 99 | 국제 및 외국기관 | 0.000508 |
| 39 | 환경 정화 및 복원업 | 0.000497 |
| 74 | 사업시설 관리 및 조경 서비스업 | 0.000490 |

5. 각 사업장 별로, 대조군 별로 5년 단위 통합 누적 발생률 계산 후 비교

: 계산 후, 추적 인년 합계 기준으로 오름차순 정렬 뒤, 시각화(사업장 별 5년 단위 통합 누적 발생률과 대응하는 대조군의 5년 단위 통합 누적 발생률 겹쳐서 그림)

: **빨간색이 사업장 별** 5년 단위 통합 누적 발생률, **검은색이 대조군의 5년 단위 통합 누적 발생률** 의미

① y축의 scale을 조정하지 않은 경우 : y축의 scale이 1번 버전 그래프(5년 단위로 나누지 않음)와 **다름**.



⇒ 백혈병 data보다는 추적 인년 합계가 적어도 안정된 추세를 보이는 것을 알 수 있다. 위 그래프에서 대조군(검은색)의 통합 누적 발생률과 큰 차이를 보이는 그래프에 **빨간색 사각형**으로 표시하였고, 해당 사업장 명은 아래 표에 제시하였다.

| UP2 | 사업장명 |
|-----|-----------------|
| 5 | 광업(석탄, 원유, 채굴업) |
| 39 | 환경 정화업 |
| 6 | 광업(금속, 철, 비금속) |

② y축의 scale을 조정한 경우



→ 백혈병 data보다는 추적 인년 합계가 적어도 안정된 추세를 보이나, 대조군과 비교했을 때는 여전히 불안정한 추세를 보이는 것을 확인할 수 있다. 위 그래프에서 대조군(검은색)의 통합 누적 발생률과 큰 차이를 보이는 그래프에 빨간색 사각형으로 표시하였고, 해당 사업장 명은 아래 표에 제시하였다. (파란색으로 표시한 부분은 y축의 scale을 고려하지 않았던 ①의 결과와 겹치는 사업장이다.)

| | |
|-----|-------------------------|
| UP2 | 사업장명 |
| 59 | 영상·오디오 기록물 제작 및 배급업 |
| 90 | 예술, 스포츠업, 건설업 등 각종 사업 |
| 37 | 하수, 폐기물 정화업(하수, 폐수, 분뇨) |
| 36 | 용수 공급업 |
| 39 | 환경 정화업 |
| 6 | 광업(금속, 철, 비금속) |
| 97 | 가구내 고용 활동 |
| 8 | 광업(광업, 자원, 원유) |
| 98 | 가사 생산 활동 |

6. 각 사업장 별 5년 단위 통합 누적 발생률과 대조군의 5년 단위 통합 누적 발생률의 “비율”(간접 SIR)가 가장 큰 사업장 TOP10은?

| UP2 | YEAR | 사업장명 | 간접 SIR |
|-----|------|-----------------|--------|
| 34 | 2001 | 산업용 기계 및 장비 수리업 | 9.41 |
| 99 | 2002 | 국제 및 외국기관 | 7.07 |
| 6 | 2001 | 광업(금속, 철, 비금속) | 6.06 |
| 6 | 2002 | 광업(금속, 철, 비금속) | 5.55 |
| 6 | 2004 | 광업(금속, 철, 비금속) | 5.11 |
| 34 | 2002 | 산업용 기계 및 장비 수리업 | 4.38 |
| 6 | 2005 | 광업(금속, 철, 비금속) | 4.24 |
| 6 | 2006 | 광업(금속, 철, 비금속) | 3.61 |
| 6 | 2003 | 광업(금속, 철, 비금속) | 3.59 |
| 7 | 2001 | 광업(비금속 광물, 소금) | 2.94 |

7. 2018년 기준, 각 사업장 별 5년 단위 통합 누적 발생률과 대조군의 5년 단위 통합 누적 발생률의 “비율”(간접 SIR)가 가장 큰 사업장 TOP10은?

| UP2 | 사업장명 | 간접 SIR |
|-----|----------------------------------|--------|
| 6 | 광업(금속, 철, 비금속) | 2.24 |
| 2 | 임업 | 1.67 |
| 5 | 광업(석탄, 원유, 채굴업) | 1.60 |
| 3 | 어업 | 1.36 |
| 37 | 하수, 폐기물 정화업(하수, 폐수, 분뇨) | 1.28 |
| 49 | 육상 운송 및 파이프라인 운송업 | 1.26 |
| 36 | 용수 공급업 | 1.24 |
| 12 | 제조업(담배) | 1.24 |
| 38 | 하수, 폐기물 정화업(방사성 물질, 폐기물, 건설 폐기물) | 1.24 |
| 16 | 제조업(유리, 목재 등) | 1.20 |

: 빨간색으로 표시한 부분은 2번 결과(5년 단위로 나누지 않고 구한 간접 SIR)와 겹치는 사업장 의미

8. 폐암 5년 단위 통합 누적 발생률 TOP10인 사업장은?

| UP2 | YEAR | 사업장명 | 통합 누적 발생률 |
|-----|------|-----------------|-----------|
| 6 | 2018 | 광업(금속, 철, 비금속) | 0.00220 |
| 6 | 2014 | 광업(금속, 철, 비금속) | 0.00209 |
| 5 | 2018 | 광업(석탄, 원유, 채굴업) | 0.00197 |
| 6 | 2017 | 광업(금속, 철, 비금속) | 0.00174 |
| 5 | 2017 | 광업(석탄, 원유, 채굴업) | 0.00172 |
| 6 | 2015 | 광업(금속, 철, 비금속) | 0.00165 |
| 6 | 2005 | 광업(금속, 철, 비금속) | 0.00162 |
| 6 | 2013 | 광업(금속, 철, 비금속) | 0.00159 |
| 5 | 2016 | 광업(석탄, 원유, 채굴업) | 0.00155 |
| 5 | 2015 | 광업(석탄, 원유, 채굴업) | 0.00155 |

9. 2018년 기준, 폐암 5년 단위 통합 누적 발생률 TOP10인 사업장은?

| UP2 | 사업장명 | 통합 누적 발생률 |
|-----|----------------------------------|-----------|
| 6 | 광업(금속, 철, 비금속) | 0.00222 |
| 5 | 광업(석탄, 원유, 채굴업) | 0.00200 |
| 12 | 제조업(담배) | 0.00140 |
| 49 | 육상 운송 및 파이프라인 운송업 | 0.000898 |
| 68 | 부동산업 | 0.000881 |
| 7 | 광업(비금속 광물, 소금 등) | 0.000854 |
| 3 | 어업 | 0.000808 |
| 99 | 국제 및 외국기관 | 0.000792 |
| 38 | 하수, 폐기물 정화업(방사성 물질, 폐기물, 건설 폐기물) | 0.000731 |
| 37 | 하수, 폐기물 정화업(하수, 폐수, 분뇨) | 0.000660 |

: 빨간색으로 표시한 부분은 3번 결과(5년 단위로 나누지 않고 구한 통합 누적 발생률)와 겹치는 사업장 의미

LEUKEMIA Data와 LUNG CANCER Data 비교

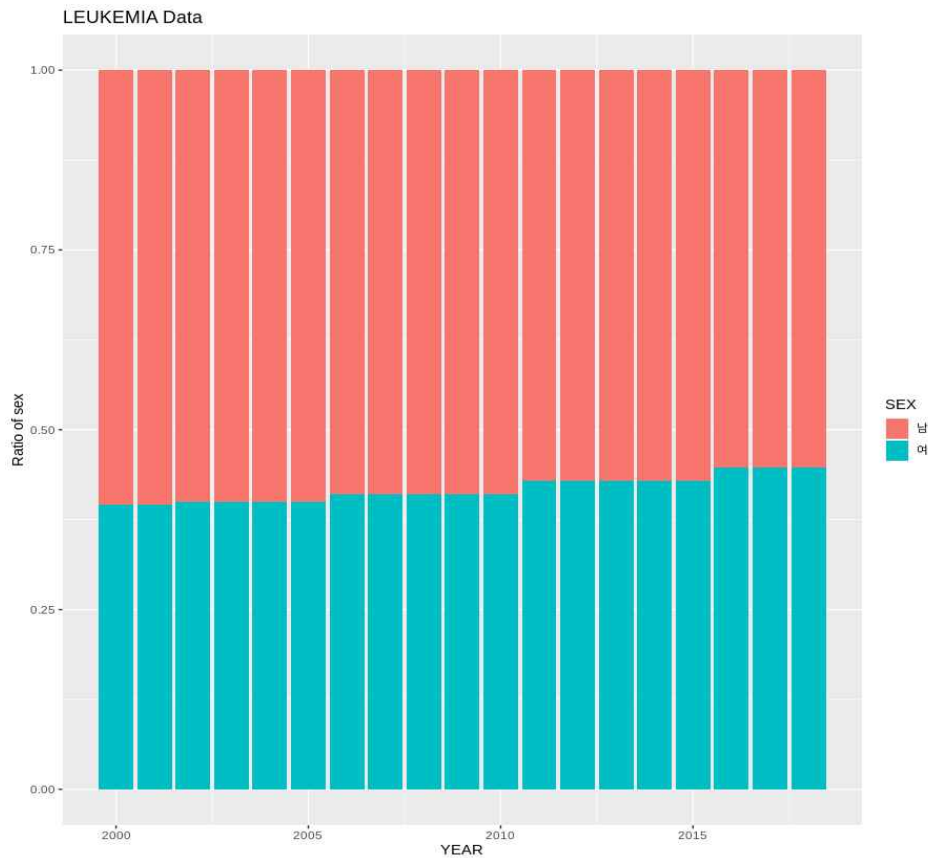
- ① 백혈병 데이터, 폐암 데이터를 시각화한 그래프들을 비교해보았을 때, 폐암의 경우 추적 인년 합계가 적어도 추세가 안정된 경우가 대부분이다.
- ② 폐암의 경우, “6번 업종(광업 - 금속, 철, 비금속)”이 모든 면(시각화, 간접 SIR, 통합 누적 발생률 파악 측면)에서 우위를 차지(대부분 높은 rank에 속함)하고 있다.
- ③ 통합 누적 발생률, 간접 SIR 값들이 전체적으로 백혈병의 경우보다 폐암의 경우가 더 크다.

----- 다음 페이지로 -----

1. YEAR에 따른 SEX의 비율 변화 시각화

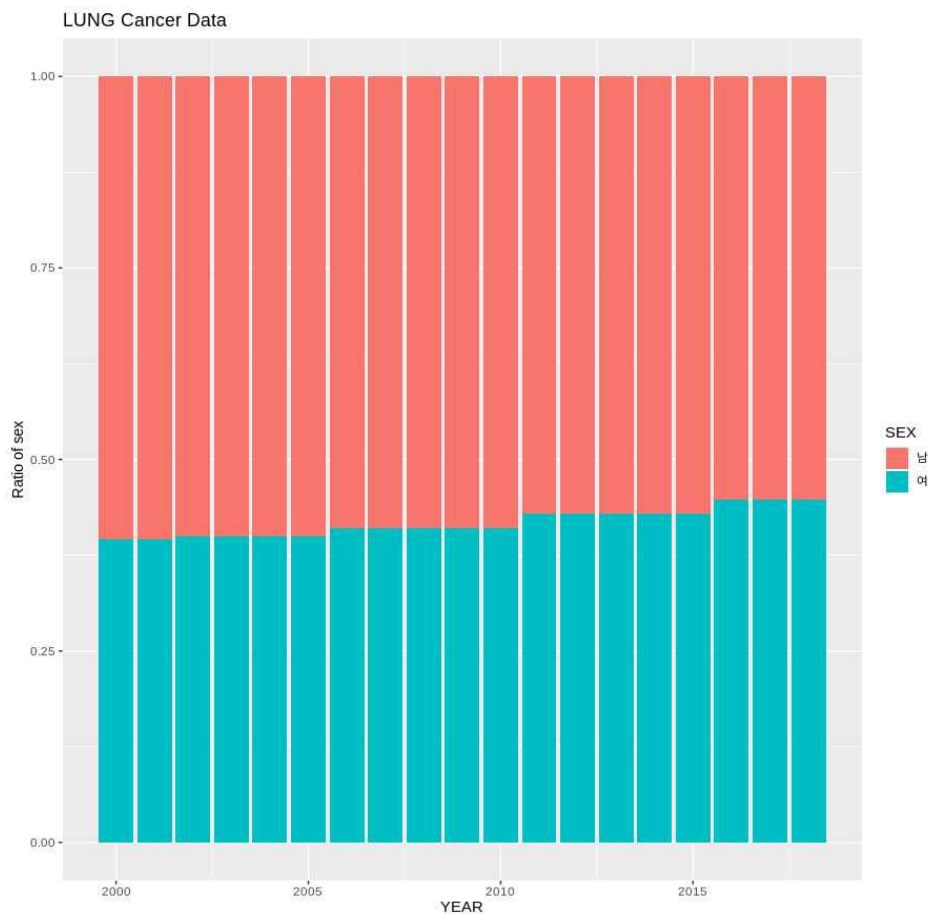
: 업종별로 나누지 않음. / **빨간색**이 “남성”, **파란색**이 “여성” 의미 / 성 비율 상대도수 표시 (전체 합이 1이 되도록)

① Leukemia data



LEUKEMIA Data와 LUNG CANCER Data의 연도별 성별 비율 변화를 살펴 보았을 때, 두 데이터에 따라 연도별 성별 비율이 크게 차이 나지 않는다. 모든 연도에 대해 남성의 비율이 더 크 나, 해가 지날수록 여성의 비율이 점차 커지는 것을 알 수 있다. 또한, 여성 비율에 해당하는 “파란색” 그래프만을 보았을 때, 증가하는 주기가 5년 단위 인 것을 알 수 있다.

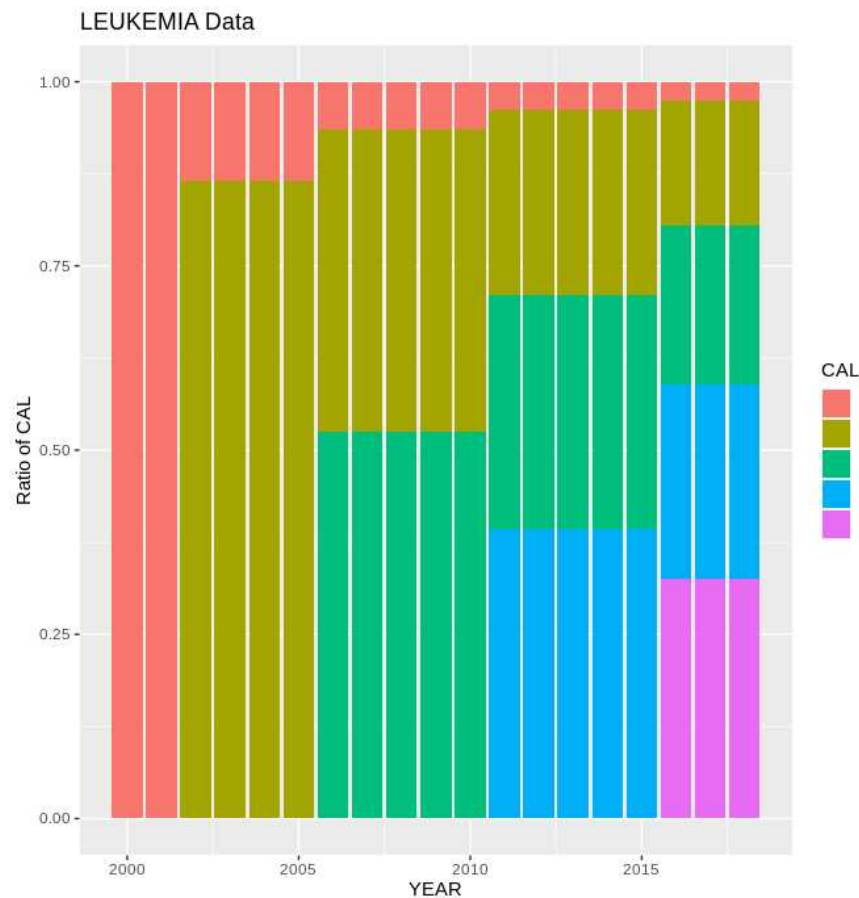
② Lung Cancer Data



2. YEAR에 따른 CAL(입사시기)의 비율 변화 시각화

: 업종별로 나누지 않음. / CAL(입사시기) 비율 상대도수 표시 (전체 합이 1이 되도록)

① Leukemia data



LEUKEMIA Data와 LUNG CANCER Data의 연도별 CAL(입사시기) 비율 변화를 살펴보았을 때, 두 데이터에 따라 연도별 CAL(입사시기) 비율이 크게 차이 나지 않는다. 2000년 ~ 2001년에는 CAL=0인, 즉 입사 시기가 1995년 이전인 근로자의 비율이 전부였고, 2002년 ~ 2005년까지는 CAL=1인, 입사 시기가 1996년 ~ 2000년인 근로자의 비율이 대다수이며, 2006년 ~ 2010년까지는 CAL=1인 근로자의 비율이 CAL=2(입사 시기가 01년 ~ 05년)인 근로자의 비율이 우위를 점한다. 2011년 ~ 2015년 사이에는 CAL=3(입사 시기가 06년 ~ 10년)인 근로자들의 비율이 급증하였으며, 2016년 이후에는 CAL=4(11년 ~ 15년)인 근로자들의 비율이 급증하였다. 성 비율 변화처럼 입사 시기도 5년 단위로 우위를 점하는 CAL의 범주가 변하는 것을 알 수 있다.

② Lung Cancer Data

