

대기오염이 질병에 미치는 영향

백 광 흡

1.질문

1. 대기오염이 질병발생 건수에유의미한 영향을 미칠까?
2. 미세먼지가 많이 발생하는 달은 언제일까?

2.참고 Data

서울시 월별 평균 대기오염정보.csv

시군구 지역코드.csv

실제진료정보_감기_시군구.csv

실제진료정보_눈병_시군구.csv

실제진료정보_천식_시군구.csv

실제진료정보_피부염_시군구.csv

- 출처 : 서울 열린데이터광장<http://data.seoul.go.kr/>

3.테이블 생성 스크립트

--대기정보 (ATMOSPHERE)

```
CREATE TABLE atmosphere
```

```
(day VARCHAR2(10), gu_name VARCHAR2(10), /*년월 ex) 201705*/
```

```
no2 number(10,4), /*이산화질소*/
```

```
ozon NUMBER(10,4), /*오존농도*/
```

```
co NUMBER(10,4), /*일산화탄소*/
```

```
so2 NUMBER(10,4) , /*아황산가스*/
```

```
fine_dust NUMBER(10), /*미세먼지*/
```

```
cho_fine_dust NUMBER(10) /*초미세먼지*/);--- 월별, 구별 평균 대기정보. 11120건
```

DELETE

FROM ATMOSPHERE

WHERE SUBSTR(day,1,4) NOT IN ('2014','2015','2016');

--- 2014, 2015, 2016 년도를 제외한 데이터 삭제. 1426건

DELETE

FROM ATMOSPHERE

WHERE gu_name NOT LIKE '%구';--- gu_name이 '구'로 끝나는 데이터만 남기고 삭제. 901건

DELETE

FROM ATMOSPHERE

WHERE gu_name IS NULL OR day IS NULL;--- null값을 삭제. 900건

--시군구 지역코드 (loc)

CREATE TABLE loc

(gu_num NUMBER(10), gu_name VARCHAR2(20)); --- 전국 시군구번호, 이름 데이터. 284건

DELETE FROM loc

WHERE gu_num > 11740;--- 서울이 데이터만 남기고 삭제. 64건

DELETE FROM loc

WHERE gu_num IS NULL OR gu_name IS NULL OR gu_name NOT LIKE '%구';

--- null값과 비정상적 데이터 삭제. 25건

--감기 (cold)

CREATE TABLE cold2

(day VARCHAR2(10), gu_num NUMBER(10), cnt NUMBER(10));---전국 일별, 구별 감기 데이터. 276793건

DELETE FROM cold2

WHERE gu_num > 11740;---서울 데이터만 남기고 삭제. 27401건

```
DELETE FROM cold2
```

```
WHERE gu_num IS NULL OR day IS NULL OR cnt IS NULL;--- null값을 삭제. 27400건
```

(ORA-24344 에러 발생)

하드로 필수)

```
CREATE TABLE COLD
```

```
AS
```

```
SELECT substr(day,1,6) day, gu_num, SUM(cnt) cnt
```

```
FROM cold2
```

```
GROUP BY substr(day,1,6), gu_num;--- 월별, 구별 감기 데이터. 900건
```

--눈병 (eye)

```
CREATE TABLE eye2
```

```
(day VARCHAR2(10), gu_num NUMBER(10), cnt NUMBER(10));---전국 일별, 구별 눈병 데이터.  
275800건
```

```
DELETE FROM eye2
```

```
WHERE gu_num > 11740;---서울 데이터만 남기고 삭제. 27401건
```

```
DELETE FROM EYE2
```

```
WHERE gu_num IS NULL OR day IS NULL OR cnt IS NULL; --- null값을 삭제. 27400건
```

```
CREATE TABLE eye
```

```
AS
```

```
SELECT substr(day,1,6) day, gu_num, SUM(cnt) cnt
```

```
FROM eye2
```

```
GROUP BY substr(day,1,6), gu_num;--- 월별, 구별 눈병 데이터. 900건
```

--천식 (asthma)

```
CREATE TABLE asthma2
```

```
(day VARCHAR2(10), gu_num NUMBER(10), cnt NUMBER(10));---전국 일별, 구별 감기 데이터.  
275758건
```

```
DELETE FROM asthma2
```

```
WHERE gu_num > 11740;---서울 데이터만 남기고 삭제. 27401건
```

```
DELETE FROM asthma2
```

```
WHERE gu_num IS NULL OR day IS NULL OR cnt IS NULL;--- null값을 삭제. 27400건
```

```
CREATE TABLE asthma
```

```
AS
```

```
SELECT substr(day,1,6) day, gu_num, SUM(cnt) cnt
```

```
FROM asthma2
```

```
GROUP BY substr(day,1,6), gu_num;--- 월별, 구별 천식 데이터. 900건
```

--피부염 (dermatitis)

```
CREATE TABLE dermatitis2
```

```
(day VARCHAR2(10), gu_num NUMBER(10), cnt NUMBER(10));--- 전국 구별, 일별 피부염 데이터 276137건
```

```
DELETE FROM dermatitis2
```

```
WHERE gu_num > 11740; ---서울시 데이터만 남기고 삭제. 27401건
```

```
DELETE FROM dermatitis2
```

```
WHERE gu_num IS NULL OR day IS NULL OR cnt IS NULL;--- null값을 삭제. 27400건
```

```
CREATE TABLE dermatitis
```

```
AS
```

```
SELECT substr(day,1,6) day, gu_num, SUM(cnt) cnt
```

```
FROM dermatitis2
```

```
GROUP BY substr(day,1,6), gu_num;--- 월별, 구별 피부염 데이터. 900건
```

--데이터 종합 (result)

```
CREATE TABLE result
```

```
AS
```

```

SELECT at.DAY, I.GU_NUM, I.GU_NAME, at.no2, at.OZON, at.co, at.SO2, at.FINE_DUST,
at.CHO_FINE_DUST, c.CNT "감기 건수", e.cnt "눈병 건수", a.cnt "천식 건수",
d.CNT "피부염 건수",
      NTILE(5) OVER (ORDER BY c.cntdesc) "감기 등급" , NTILE(5) OVER (ORDER BY e.cntdesc)
"눈병 등급" , NTILE(5) OVER (ORDER BY a.cntdesc) "천식 등급",
      NTILE(5) OVER (ORDER BY d.cntdesc) "피부병 등급" , NTILE(5) OVER (ORDER BY at.no2 desc)
"이산화질소 등급", NTILE(5) OVER (ORDER BY at.ozon DESC) "오존 등급",
      NTILE(5) OVER (order BY at.co desc) "아황산가스 등급" , NTILE(5) OVER (ORDER BY at.so2)
"일산화탄소 등급", NTILE(5) OVER (ORDER BY at.fine_dust) "미세먼지 등급",
      NTILE(5) OVER (ORDER BY at.cho_fine_dust) "초미세먼지 등급", abs(NTILE(5) OVER
(ORDER BY c.cntdesc)-NTILE(5) OVER (ORDER BY at.no2 desc)) "감기-이산화질소",
abs(NTILE(5) OVER (ORDER BY c.cntdesc)-NTILE(5) OVER (ORDER BY at.ozon DESC)) "감기-오존",
abs(NTILE(5) OVER (ORDER BY c.cntdesc)- NTILE(5) OVER (ORDER BY at.co desc)) "감기-아황산가스",
abs(NTILE(5) OVER (ORDER BY c.cntdesc)- NTILE(5) OVER (ORDER BY at.so2)) "감기-일산화탄소",
abs(NTILE(5) OVER (ORDER BY c.cntdesc)- NTILE(5) OVER (ORDER BY at.fine_dust)) "감기-미세먼지",
abs(NTILE(5) OVER (ORDER BY c.cntdesc)- NTILE(5) OVER (ORDER BY at.cho_fine_dust)) "감기-초미세먼지",
abs(NTILE(5) OVER (ORDER BY e.cntdesc)- NTILE(5) OVER (ORDER BY at.no2 desc)) "눈병-이산화질소",
abs(NTILE(5) OVER (ORDER BY e.cntdesc)- NTILE(5) OVER (ORDER BY at.ozon DESC)) "눈병-오존",
abs(NTILE(5) OVER (ORDER BY e.cntdesc)- NTILE(5) OVER (ORDER BY at.co desc)) "눈병-아황산가스",
abs(NTILE(5) OVER (ORDER BY e.cntdesc)- NTILE(5) OVER (ORDER BY at.so2)) "눈병-일산화탄소",
abs(NTILE(5) OVER (ORDER BY e.cntdesc)- NTILE(5) OVER (ORDER BY at.fine_dust)) "눈병-미세먼지",
abs(NTILE(5) OVER (ORDER BY e.cntdesc)- NTILE(5) OVER (ORDER BY at.cho_fine_dust)) "눈병-초미세먼지",
abs(NTILE(5) OVER (ORDER BY a.cntdesc)- NTILE(5) OVER (ORDER BY at.no2 desc)) "천식-이산화질소",
abs(NTILE(5) OVER (ORDER BY a.cntdesc)- NTILE(5) OVER (ORDER BY at.ozon DESC)) "천식-오존",
abs(NTILE(5) OVER (ORDER BY a.cntdesc)-NTILE(5) OVER (order BY at.co desc)) "천식-아황산가스",
abs(NTILE(5) OVER (ORDER BY a.cntdesc)- NTILE(5) OVER (ORDER BY at.so2)) "천식-일산화탄소",
abs(NTILE(5) OVER (ORDER BY a.cntdesc)- NTILE(5) OVER (ORDER BY at.fine_dust)) "천식-미세먼지",
abs(NTILE(5) OVER (ORDER BY a.cntdesc)- NTILE(5) OVER (ORDER BY at.cho_fine_dust)) "천식-초미세먼지",
abs( NTILE(5) OVER (ORDER BY d.cntdesc)- NTILE(5) OVER (ORDER BY at.no2 desc)) "피부병-이산화질소",
abs( NTILE(5) OVER (ORDER BY d.cntdesc)- NTILE(5) OVER (ORDER BY at.ozon DESC)) "피부병-오존",
abs( NTILE(5) OVER (ORDER BY d.cntdesc)- NTILE(5) OVER (order BY at.co desc)) "피부병-아황산가스",
abs( NTILE(5) OVER (ORDER BY d.cntdesc)- NTILE(5) OVER (ORDER BY at.so2)) "피부병-일산화탄소",
abs( NTILE(5) OVER (ORDER BY d.cntdesc)- NTILE(5) OVER (ORDER BY at.fine_dust)) "피부병-미세먼지",
abs( NTILE(5) OVER (ORDER BY d.cntdesc)-NTILE(5) OVER (ORDER BY at.cho_fine_dust)) "피부병-초미세먼지"
FROM ATMOSPHERE at, LOC I, COLD c, EYE e, asthma a, dermatitis d
WHERE at.GU_NAME = I.GU_NAME
      AND c.GU_NUM = I.GU_NUM AND c.day = at.DAY
      AND e.GU_NUM = I.GU_NUM AND e.DAY = at.DAY
      AND a.gu_num = I.GU_NUM AND a.day = at.day
      AND d.GU_NUM = I.GU_NUM AND d.DAY = at.day
      ORDER BY at.day;

```

--대기오염~질병 연관관계 (result2) [abs(질병건수등급 - 오염등급)]

```

CREATE TABLE result2
AS
SELECT ROUND(AVG("감기-이산화질소"),3)"감기-이산화질소", ROUND(AVG("감기-오존"),3)"감기-오존",
ROUND(AVG("감기-아황산가스"),3)"감기-아황산가스",ROUND(AVG("감기-일산화탄소"),3)"감기-일산화탄소",
ROUND(AVG("감기-미세먼지"),3)"감기-미세먼지", ROUND(AVG("감기-초미세먼지"),3)"감기-초미세먼지",
ROUND(AVG("눈병-이산화질소"),3)"눈병-이산화질소", ROUND(AVG("눈병-오존"),3)"눈병-오존",
ROUND(AVG("눈병-아황산가스"),3)"눈병-아황산가스",ROUND(AVG("눈병-일산화탄소"),3)"눈병-일산화탄소",
ROUND(AVG("눈병-미세먼지"),3)"눈병-미세먼지", ROUND(AVG("눈병-초미세먼지"),3)"눈병-초미세먼지",
ROUND(AVG("천식-이산화질소"),3)"천식-이산화질소", ROUND(AVG("천식-오존"),3)"천식-오존",
ROUND(AVG("천식-아황산가스"),3)"천식-아황산가스",ROUND(AVG("천식-일산화탄소"),3)"천식-일산화탄소",
ROUND(AVG("천식-미세먼지"),3)"천식-미세먼지", ROUND(AVG("천식-초미세먼지"),3)"천식-초미세먼지",
ROUND(AVG("피부병-이산화질소"),3)"피부병-이산화질소",ROUND(AVG("피부병-오존"),3)"피부병-오존",
ROUND(AVG("피부병-아황산가스"),3)"피부병-아황산가스",ROUND(AVG("피부병-일산화탄소"),3)"피부병-일산화탄소",
ROUND(AVG("피부병-미세먼지"),3)"피부병-미세먼지", ROUND(AVG("피부병-초미세먼지"),3)"피부병-초미세먼지"
FROM result;

```

4.분석쿼리

--1. 대기오염이 질병 발생 건수에 유의미한 영향을 미칠까?

--1.1 질병-대기오염 전체 연관도 순위 출력

```

SELECT item "질병-대기오염", cnt "연관도", RANK() OVER (ORDER BY cntasc) 순위
FROM result2
unpivot(cnt FOR item IN(
"감기-이산화질소","감기-오존","감기-아황산가스","감기-일산화탄소","감기-미세먼지","감기-초미세먼지",
"눈병-이산화질소","눈병-오존","눈병-아황산가스","눈병-일산화탄소","눈병-미세먼지","눈병-초미세먼지",
"천식-이산화질소","천식-오존","천식-아황산가스","천식-일산화탄소","천식-미세먼지","천식-초미세먼지",
"피부병-이산화질소","피부병-오존","피부병-아황산가스","피부병-일산화탄소","피부병-미세먼지","피부병-초미세먼지"));

```

--1.2 질병별 가장 영향을 끼치는 요인 2위까지 출력

```

SELECT substr("질병-대기오염",1,INSTR("질병-대기오염",-)-1) "질병"

```

```

,substr("질병-대기오염",INSTR("질병-대기오염",-)+1,length("질병-대기오염")) "대기오염", "
연관 순위"
FROM(SELECT "질병-대기오염",RANK() OVER(PARTITION BY substr("질병-대기오염",1,INSTR("질병
-대기오염",-)-1) ORDER BY 연관도 ) "연관 순위"
FROM (SELECT item "질병-대기오염", cnt "연관도", RANK() OVER (ORDER BY cntasc) 순
위
FROM result2
unpivot(cnt FOR item IN(
"감기-이산화질소", "감기-오존", "감기-아황산가스", "감기-일산화탄소", "감기-미세먼지", "감기-초
미세먼지",
"눈병-이산화질소", "눈병-오존", "눈병-아황산가스", "눈병-일산화탄소", "눈병-미세먼지", "눈병-초
미세먼지",
"천식-이산화질소", "천식-오존", "천식-아황산가스", "천식-일산화탄소", "천식-미세먼지", "천식-초
미세먼지",
"피부병-이산화질소", "피부병-오존", "피부병-아황산가스", "피부병-일산화탄소", "피부병-미세먼지
", "피부병-초미세먼지"))))
WHERE "연관 순위" <= 2 ;

```

--2. 미세먼지가 많이 발생하는 달은 언제일까?

```

SELECT SUBSTR(day,5,2)월, ROUND(AVG(fine_dust),2) "평균미세먼지",
round((AVG(fine_dust)-NVL(LAG(AVG(fine_dust),1) OVER (ORDER BY SUBSTR(day,5,2)),47.28))/
NVL(LAG(AVG(fine_dust),1) OVER (ORDER BY SUBSTR(day,5,2)),47.28)*100,2) "전월대비 상승률",
RANK() OVER(ORDER BY ROUND(AVG(fine_dust),2) desc) 순위
FROM result
group BY SUBSTR(day,5,2)
ORDER BY 월;

```

5.결과화면

--1. 대기오염이 질병 발생 건수에 유의미한 영향을 미칠까?

1.1 질병-대기오염 전체 연관도 순위 출력

구	질병-대기오염	연관도	순위
1	감기-아황산가스	1.153	1
2	감기-이산화질소	1.198	2
3	천식-아황산가스	1.238	3
4	천식-이산화질소	1.324	4
5	피부병-오존	1.333	5
6	눈병-오존	1.4	6
7	피부병-초미세먼지	1.404	7
8	눈병-초미세먼지	1.473	8
9	피부병-미세먼지	1.496	9
10	눈병-미세먼지	1.553	10
11	피부병-일산화탄소	1.618	11
12	눈병-일산화탄소	1.622	12
13	눈병-아황산가스	1.682	13
14	천식-초미세먼지	1.764	14
15	피부병-아황산가스	1.791	15
16	눈병-이산화질소	1.804	16
17	감기-초미세먼지	1.831	17
18	천식-일산화탄소	1.838	18
19	천식-오존	1.842	19
20	감기-일산화탄소	1.853	20
21	피부병-이산화질소	1.871	21
22	감기-오존	1.909	22
23	천식-미세먼지	1.982	23
24	감기-미세먼지	2	24

*연관도 : [대기오염 등급 - 질병 발생건수 등급]의 절대값. (최소 0 ~ 최대 4)
0에 가까울수록 대기오염의 정도와 질병 발생건수가 비례함.

1.2 질병별 가장 영향을 끼치는 요인 2위까지 출력

구	질병	대기오염	연관 순위
1	감기	아황산가스	1
2	감기	이산화질소	2
3	눈병	오존	1
4	눈병	초미세먼지	2
5	천식	아황산가스	1
6	천식	이산화질소	2
7	피부병	오존	1
8	피부병	초미세먼지	2

--2. 미세먼지가 많이 발생하는 달은 언제일까?

구	월	평균미세먼지	전월대비 상승률	순위
1	01	51.97	9.92	5
2	02	62.01	19.32	2
3	03	64.99	4.81	1
4	04	57.68	-11.25	3
5	05	54.56	-5.41	4
6	06	39.47	-27.66	8
7	07	33.87	-14.19	10
8	08	32.51	-4.02	11
9	09	31.63	-2.71	12
10	10	38.6	22.04	9
11	11	43.53	12.77	7
12	12	47.28	8.61	6

6.결론

--1. 대기오염이 질병 발생 건수에 유의미한 영향을 미칠까?

1. 질병과 대기오염의 연관도는 0~4의 값을 갖는데 모든 연관도가 2 이하의 값을 가졌다. (0에 가까울수록 영향 ↓)
이는 대기오염 농도와 질병 발생건수가 비례하고, 대기오염이 질병 발생 건수에 유의미한 영향을 끼친다고 생각할 수 있다.
2. 분석 전 미세먼지와 천식, 감시의 연관관계가 높을 것으로 예상했지만, 의외로 (초)미세먼지가 감기, 천식에 많은 영향을 미치지 않는 결과가 나타났다. 이 결과를 통해 서울 시민들이 미세먼지 관련 이슈에 반응하여 미세먼지 농도가 높은 날은 야외활동 자제 및 마스크 착용과 같은 대비를 했을 수도 있다는 생각을 해볼 수 있다.

구	질병	대기오염	연관 순위
1	감기	아황산가스	1
2	감기	이산화질소	2
3	눈병	오존	1
4	눈병	초미세먼지	2
5	천식	아황산가스	1
6	천식	이산화질소	2
7	피부병	오존	1
8	피부병	초미세먼지	2

대기오염	인체에 미치는 영향
이산화질소	눈, 코 등의 점막에서 만성 기관지염, 호흡기 질환, 천식, 폐렴, 폐출혈, 폐수종의 발병으로까지 발전가능.
오존	가슴의 통증, 기침, 메스꺼움, 목 자극, 소화 불량 등의 증상을 유발하고 기관지염, 심장질환, 폐기종 및 천식을 악화, 폐활량을 감소 시키고 눈의 기능에 영향을 미친다. 오존에 의해 각질층 속의 불포화지방산이 산화되고 차단 기능이 손상돼 피부염 을 일으킬 가능성이 있다
아황산가스	인체의 점막을 자극하며, 고농도를 흡입하면 콧물, 담, 기침 등이 나오고 호흡곤란 을 초래한다. 이산화황 흡입을 통해 기관지염, 폐수종, 폐렴 등에 걸릴 가능성이 있다는 연구 결과가 있다.
일산화탄소	혈액순환 기능 중 산소의 운반기능을 저하시키며, 고농도의 일산화탄소는 유독성이 있어 건강한 사람에게도 치명적인 해를 입힌다.
미세먼지	천식과 같은 호흡기계 질병을 악화시키고, 폐 기능의 저하를 초래한다
초미세먼지	감기나 천식, 기관지염 등의 호흡기 질환을 비롯하여 심혈관 질환, 안구 질환, 피부 질환 등 각종 질병의 발병 확률을 높인다.

-출처

:

환경부

<http://www.me.go.kr/mamo/web/index.do?menuId=587>

--2. 미세먼지가 많이 발생하는 달은 언제일까?

1. 6~10월은 미세먼지 농도가 30대를 유지하며 비교적 쾌적한 상태를 유지한다.
2. 11월부터 1월까지 8~12%의 미세먼지 상승률을 보여주며 43~52의 미세먼지 농도를 보여준다.
3. 2월은 전월대비 19%라는 높은 상승률을 보여주며 3월달까지 올랐다가 9월달까지 꾸준히

감소한다. 특히 6월에는 전월대비 27%의 급격한감소율을 보여준다.

4. 분석결과를 통해 미세먼지는 봄, 겨울, 가을, 여름 순으로 높다는 것을 알 수 있다.
5. 여름인 6~9월까지의는 꾸준한 미세먼지 감소율을 보여주지만 10월부터는 3월까지의는 꾸준한 상승률을 보인다.
이는 여름에 발생하는 태풍의 영향과 남동계절풍이 주된 원인이라고 추론할 수 있다.
6. 미세먼지를 발생시키는 국내요인(교통,공사,시설물)의 급격한 변화가 없음에도 불구하고 계절별로 많은 차이를 보이는 것은"봄-겨울에 부는 북서계절풍, 편서풍으로 인해 중국의 영향을 많이 받는다"는 추론할 수 있다.
7. 즉, 미세먼지는 봄, 겨울, 가을, 여름 순으로 높고 계절풍으로 인한 국외요인의 영향을 상당히 받는다.

구분	월	평균미세먼지	전월대비 상승률	순위
1	01	51.97	9.92	5
2	02	62.01	19.32	2
3	03	64.99	4.81	1
4	04	57.68	-11.25	3
5	05	54.56	-5.41	4
6	06	39.47	-27.66	8
7	07	33.87	-14.19	10
8	08	32.51	-4.02	11
9	09	31.63	-2.71	12
10	10	38.6	22.04	9
11	11	43.53	12.77	7
12	12	47.28	8.61	6





7.연습문제

문제 1. 월,월별 감기 발생건수, 전월대비 상승률, 감기건수 순위를 출력 하시오.

	월	평균 감기건수	상승률	순위
1	01	87894.72	-18.3	6
2	02	97923.36	11.4	2
3	03	96136.04	-1.8	3
4	04	94366.67	-1.8	4
5	05	78204.83	-17.1	8
6	06	58679.31	-25	10
7	07	52880.15	-9.9	11
8	08	51753.16	-2.1	12
9	09	71465.89	38.1	9
10	10	82104.83	14.9	7
11	11	91413.8	11.3	5
12	12	107540.05	17.6	1

```

SELECT SUBSTR(day,5,2)월, ROUND(AVG("감기 건수"),2) "평균 감기건수",
round((AVG("감기 건수") - NVL(LAG(AVG("감기 건수"),1) OVER (ORDER BY SUBSTR(day,5,2)),107540.05))
/NVL(LAG(AVG("감기 건수"),1) OVER (ORDER BY SUBSTR(day,5,2)),107540.05),3)*100 "상승률",
RANK() OVER(ORDER BY AVG("감기 건수") desc) 순위
FROM result
Group BY SUBSTR(day,5,2)
ORDER BY 월;

```