# 암 사망률과 흡연/음주 관계

이윤정

#### 1. 분석 목적

저는 '서울시의 암과 흡연, 음주 간의 관계'를 분석주제로 결정하였으며, 사망까지 도달 할수 있는 암이라는 병과 흡연, 음주의 분석을 통해 병에 대한 예방을 목적으로 분석주제를 설정하게 되었습니다. 또한 성별에 따라서 암 사망 발생 비율에 차이가 있는지, 음주와 흡연의 차이가 존재하는지 분석하고자 합니다.

#### 2. 분석 데이터 설명

'서울 열린데이터 광장'(<a href="https://data.seoul.go.kr/">https://data.seoul.go.kr/</a>) 의 총 3가지 데이터 '서울시 주요 암 사 망원인별 사망(성별/구별)통계(2017)', '서울시 고위험 음주율 통계(2017)', '서울시 현재흡연율 통계(2017)'를 활용하였습니다. 세 가지의 데이터는 모두 성별에 따른 데이터를 가지고 있으며 '구'라는 공통적인 행을 가지고 있습니다.

첫 번째 데이터인 '서울시 주요 암 사망원인별 사망(성별/구별)통계(2017)' 데이터는 각 암별 사망자수와 사망률, 각 암별/성별 사망자수와 사망률을 총 31개의 열로 가지고 있습니다. 두 번째 데 이터와 세 번째 데이터를 통합하여 '서울시 고위험 음주율, 흡연율 통계(2017)'라는 파일을 만들었으 며, 이는 음주율, 성별 음주율, 흡연율, 성별 흡연율 4가지 열로 이루어져 있습니다.

#### 3. 데이터분석 결과

1) 탐색적 자료 분석 - 막대그래프

폐암	간암	위암	대장암	췌장암	전체
강북구	강북구	종로구	종로구	중구	강북구
용산구	종로구	강북구	중구	강북구	종로구
서대문구	중랑구	금천구	강북구	용산구	중구

성별을 배제하고 사망률을 높은 순으로 정렬했을 때의 상위 3개 구 들입니다. 강북구는 모든 열에 속해 있을 만큼 사망률이 높다고 보이며, 종로구, 중구 또한 공통적으로 보입니다.

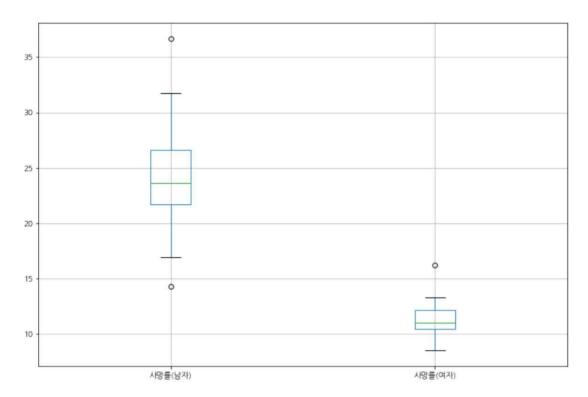
폐암(남자)	간암(남자)	위암(남자)	대장암(남자)	췌장암(남자)	전체(남자)
강북구	강북구	종로구	종로구	중구	강북구
중랑구	서대문구	강북구	도봉구	강북구	중구
용산구	중로구	중구	강북구	중랑구	종로구

위 표는 성별이 남자일 경우 각 사망률을 높은 순으로 정렬했을 때의 상위 3개 구입니다. 강북구는 이번에도 모든 열에 속해 있으며, 중랑구도 살짝 높아 보이며 중구와 종로구 역시 공통적으로 보입니다.

폐암(여자)	간암(여자)	위암(여자)	대장암(여자)	췌장암(여자)	전체(여자)
종로구	금천구	종로구	중구	용산구	종로구
영등포구	성동구	금천구	종로구	종로구	용산구

서대문구	용산구	양천구	강동구	중구	노원구

위 표는 성별이 여자일 경우 각 사망률을 높은 순으로 정렬했을 때의 상위 3개 구입니다. 앞서 주어진 두 표와는 차이가 있는 것으로 보입니다. 종로구, 용산구, 금천구가 사망률이 높은 것으로 보입니다. 전체 사망률 표가 성별이 남자인 사망률 표와는 구에 대한 차이가 없고, 성별이 여자인 사 망률 표와는 구에 대한 차이가 있는 것으로 보아 암 사망률의 비율은 남자가 더 높다는 추측이 가능합니다.



Boxplot 결과 추측과 일치합니다. 남자의 사망 비율이 여자의 사망 비율보다 현저히 높으며 둘의 평균 차이가 존재한다고 예상할 수 있습니다.

또한 암 사망률과 음주와 흡연 간의 관계파악을 위해 위에서 분석한 상위 3개 구의 음주 흡연율을 주목할 필요가 있습니다.

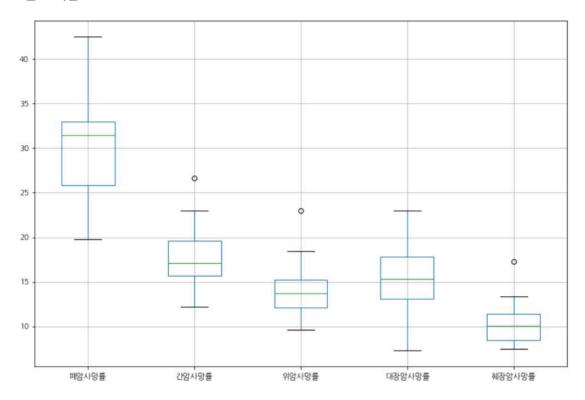
다음은 음주 흡연율 데이터에 관한 탐색적 자료 분석입니다.

음주율	흡연율	음주율(남자)	음주율(여자)	흡연율(남자)	흡연율(여자)
금천구	중랑구	금천구	동대문구	중랑구	은평구
동대문구	강북구	동대문구	강북구	강북구	동대문구
중랑구	동대문구	중랑구	중랑구	동대문구	강북구

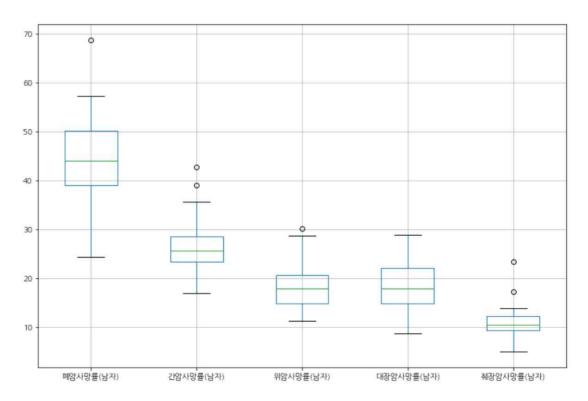
암 사망률이 높은 구인 강북구와 중랑구가 음주율과 흡연율에서도 공통적으로 보입니다. 그러나 음주와 흡연이 암 사망률과 상관관계가 존재하는지는 알 수 없습니다.

## 2) 시각화

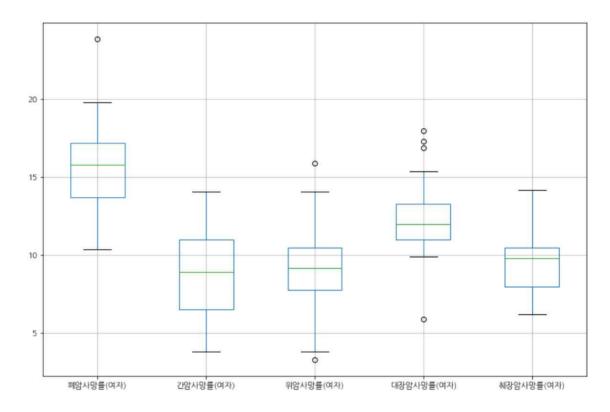
## # 분포 확인



위 그래프는 암 별 사망률의 Boxplot입니다. 폐암 사망률이 가장 높은 비율을 차지 하고 있으며 그 다음으로 간암, 대장암, 위암, 췌장암 순으로 이어짐을 확인할 수 있습니다.

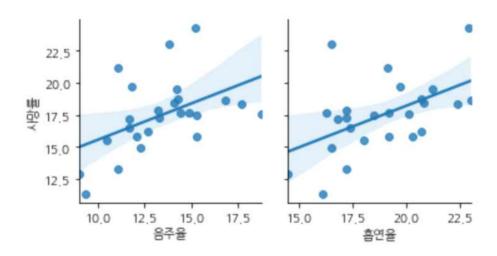


성별이 남자인 경우의 암 별 사망률 Boxplot입니다. 전체 사망률과 같이 폐암 사망률이 가장 높은 비율을 차지하며, 간암, 대장암, 위암, 췌장암 순으로 이어집니다.

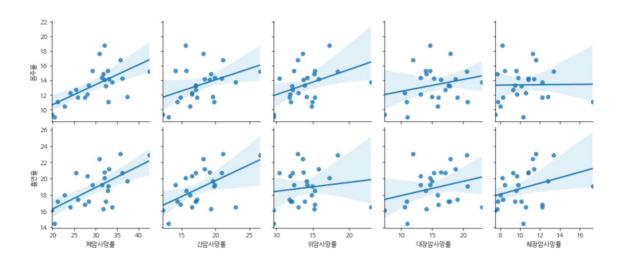


다음은 성별이 여자인 경우 암 별 사망률 Boxplot입니다. 이 역시 폐암이 가장 높으나, 대 장암이 두 번째, 간암, 위암, 췌장암이 비슷한 비율선상에 존재합니다.

## # 산점도

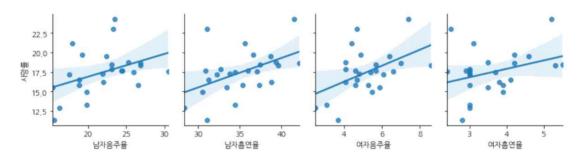


다음은 음주율과 흡연율 그리고 사망률 간의 상관관계입니다. 강한 상관관계를 띄고 있지는 않으나 약한 양의 상관관계를 보이고 있습니다. 더 세부적인 상관관계를 알아보기 위하여 각 암 별로 음주율과 흡연율 간의 상관관계를 분석했습니다.

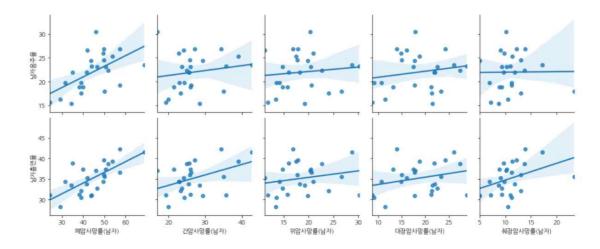


폐암사망률에서만 음주율과 흡연율 간의 양의 상관관계가 존재하며 간암은 흡연율과 약한 양의 상관관계를 보입니다. 그리고 나머지 암 들의 사망률은 상관관계가 없는 것으로 보입니다. 따라서 전체 사망률과 음주 흡연율 간의 약한 양의 상관관계가 나타난 이유는 폐암 사망률과 음주 흡연율 간의 양의 상관관계 때문이라고 볼 수 있습니다. 이 결과에 따라 폐암에 걸린 환자들은 회복을 위해 술과 담배를 멀리해야 할 필요가 있습니다.

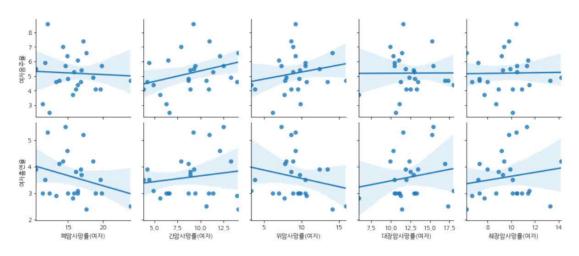
## # 성별에 따른 사망률과 흡연 음주율의 상관관계?



다음은 성별 음주 흡연율과 사망률 간의 상관관계입니다. 4개의 그래프 모두 약한 양의 상관 관계 혹은 아무 관계도 가지지 않는 것처럼 보입니다.

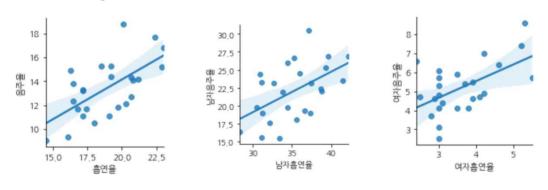


성별이 남자인 경우의 사망률과 음주 흡연율의 상관관계 역시 폐암 사망률만 양의 상관관계가 존재하는 것으로 보입니다. 간암 또한 흡연율과 약한 양의 상관관계를 보이고 나머지 암들은 모두 상관관계가 존재하지 않습니다.



성별이 여자인 경우의 사망률과 음주 흡연율의 상관관계는 전혀 없는 것처럼 보입니다.

# # 흡연과 음주의 상관관계



흡연과 음주율은 성별과 전체 모두 양의 상관관계를 가지고 있는 것으로 보입니다.

## # 히트맵

			사망자비율			
강북구 -	42,500000	26,600000	18,400000	20,600000	13,400000	
종로구 -	33,800000	23,000000	23,000000	23,000000	12,200000	- 40
중구 -	33,000000	19,800000	14,900000	20,700000	17,300000	
용산구 -	37,300000	19,600000	14,100000	15,000000	12,800000	
도봉구 -	33,100000	18,400000	15,800000	18,700000	11,400000	
서대문구 -	36,200000	20,900000	10,400000	15,300000	10,800000	- 35
중랑구 -	35,700000	21,100000	12,900000	11,900000	11,600000	
은평구 -	32,000000	19,100000	13,700000	16,400000	10,800000	
동대문구 -	30,900000	18,200000	13,600000	17,900000	11,300000	- 30
노원구 -	32,400000	17,100000	12,100000	17,800000	10,100000	
성동구 -	31,400000	19,900000	14,200000	13,100000	9,800000	
영등포구 -	31,800000	19,300000	15,200000	14,100000	7,700000	
금천구 -	32,000000	15,600000	17,300000	14,700000	8,200000	- 25
성북구 -	32,800000	15,700000	15,300000	13,900000	9,600000	
구로구 -	28,300000	17,100000	13,900000	17,600000	9,500000	
동작구 -	27,500000	14,800000	15,300000	16,000000	12,200000	- 20
강동구 -	25,800000	17,300000	11,800000	19,100000	8,300000	20
강서구 -	25,400000	17,100000	12,200000	17,100000	9,300000	
관악구 -	28,000000	16,500000	11,900000	13,100000	9,700000	
양천구 -	29,200000	14,100000	12,400000	14,700000	8,500000	- 15
광진구 -	22,700000	16,200000	14,800000	15,900000	8,000000	
마포구 -	24,100000	16,500000	13,000000	10,600000	10,600000	
송파구 -	21,200000	14,400000	12,000000	10,900000	7,800000	
서초구 -	20,400000	13,100000	9,600000	11,000000	10,300000	- 10
강남구 -	19,800000	12,200000	9,600000	7,300000	7,500000	
	폐암사망률	간암사망률	위암사망률	대장암사망률	췌장암사망률	

위의 그래프를 보면 강북구 종로구 중구 순으로 사망자 비율을 확인할 수 있으며 폐암의 사망자 비율이 가장 높은 것을 볼 수 있습니다.

	음주흡연비율	(음주기준)	
금천구 -	18,800000	20,100000	
동대문구 -	17,700000	22,400000	- 22
중광구 -	16,800000	23,000000	
성북구 -	15,300000	18,500000	
양천구 -	15,300000	19,200000	
강북구 -	15,200000	22,900000	- 20
영등포구 -	14,900000	16,300000	
성동구 -	14,400000	19,200000	
서대문구 -	14,300000	20,700000	- 18
도봉구 -	14,200000	21,200000	10
은평구 -	14.100000	20,800000	
종로구 -	13,800000	16,500000	
. 구로구 -	13,300000	17,200000	- 16
노원구 -	13,200000	17,200000	
강서구 -	12,700000	20,700000	
마포구 -	12,300000	16,500000	
관악구 -	12,100000	20,300000	- 14
용산구 -	11,800000	19,700000	
동작구 -	11,700000	16,800000	
강동구 -	11,700000	17,400000	- 12
중구 -	11,100000	19,100000	
송파구 -	11,100000	17,200000	
광진구 -	10,500000	18,000000	
강남구 -	9,300000	16,100000	- 10
서초구 -	9,000000	14,500000	
	음주율	- 흡연율	

위는 음주율을 기준으로 한 읍주 흡연 히트맵으로 금천구 동대문구 중랑구 순으로 음주율이 높은 것을 알 수 있습니다. 앞선 사망자 비율 히트맵과 공통되는 부분이 크지 않다고 보입니다.

음주흡연비율(흡연기준)

	BT B L T B	3/8 0/10/	
중광구 :	16,800000	23,000000	
강북구 ·	15,200000	22,900000	- 22
동대문구 -	17,700000	22,400000	
도봉구 -	14,200000	21,200000	
은평구 -	14,100000	20,800000	
서대문구 -	14,300000	20,700000	- 20
강서구 :	12,700000	20,700000	
관악구 -	12,100000	20,300000	
금천구 :	18,800000	20,100000	- 18
용산구 -	11,800000	19,700000	10
성동구 -	14,400000	19,200000	
양천구 -	15,300000	19,200000	
75 백	11,100000	19,100000	- 16
성북구	15,300000	18,500000	
광진구	10,500000	18,000000	
강동구 -	11,700000	17,400000	
구로구 -	13,300000	17,200000	- 14
노원구 -	13,200000	17,200000	
송파구	11,100000	17,200000	
동작구	11,700000	16,800000	- 12
마포구 -	12,300000	16,500000	
종로구 -	13,800000	16,500000	
영등포구 -	14,900000	16,300000	
강남구 :	9,300000	16,100000	- 10
서초구 :	9,000000	14,500000	
	용주율	- 홈연율	

위는 흡연율을 기준으로 한 음주흡연 히트맵으로 중랑구 강북구 동대문구 순으로 높은 흡연 율을 보입니다. 또한 사망자 비율의 상위 10구인 (강북구, 종로구, 중구, 용산구, 도봉구, 서대문구, 중 랑구, 은평구, 동대문구, 노원구) 와 (중랑구, 강북구, 동대문구, 도봉구, 은평구, 서대문구, 강서구, 관 악구, 금천구, 용산구)가 종로구, 중구, 노원구를 제외하고 일치하며 높은 상관성을 보입니다.