

# 상관분석을 이용한 만성 폐 질환에 영향을 주는 요소 찾기

학번 : 201803018

이름 : 한장희

---

# Contents

|    |       |                 |
|----|-------|-----------------|
| 01 | _____ | 분석 목적           |
| 02 | _____ | 분석 방법           |
| 03 | _____ | 상관분석을 위한 데이터 준비 |
| 04 | _____ | 상관분석            |
| 05 | _____ | 상관분석 결과 표현      |
| 06 | _____ | 분석에 사용한 전체 스크립트 |
| 07 | _____ | 결론              |
| 08 | _____ | 참고 자료           |

## 01 분석 목적

### 만성 폐 질환 발병에 영향을 미치는 요인 분석

만성 폐 질환(chronic Lung Disease)에 영향을 미치는 요소는 무엇인가



나이(Age), 성별(Gender), 대기 오염(Air Pollution), 알코올 섭취(Alcohol use), 먼지 알레르기(Dust Allergy), 직업적 위험(Occupational Hazards), 유전적 위험(Genetic Risk), 비만(Obesity), 흡연(Smoking), 간접 흡연(Passive Smoker), 가슴 통증(Chest Pain), 객혈(Coughing of Blood), 피로(Fatigue), 체중 감소(Weight Loss), 호흡 곤란(Shortness of Breath), 천명(Wheezing), 삼킴 곤란(Swallowing Difficulty), 손톱 클럽 모양(Clubbing of Finger Nails), 잦은 감기(Frequent Cold), 마른 기침(Dry Cough), 코골이(Snoring) 등 21개의 변수

02 분석 방법

```
lung_cancerC <- lung_cancer[, -c(1, 2, 11, 26 )]
```

| Age        | Gender            | Air.Pollution | Alcohol.use | Dust.Allergy        | OccuPational.Hazards | Genetic.Risk          | chronic.Lung.Disease     | Obesity       | Smoking | Passive.Smoker |
|------------|-------------------|---------------|-------------|---------------------|----------------------|-----------------------|--------------------------|---------------|---------|----------------|
| 33         | 1                 | 2             | 4           | 5                   | 4                    | 3                     | 2                        | 4             | 3       | 2              |
| 17         | 1                 | 3             | 1           | 5                   | 3                    | 4                     | 2                        | 2             | 2       | 4              |
| 35         | 1                 | 4             | 5           | 6                   | 5                    | 5                     | 4                        | 7             | 2       | 3              |
| 37         | 1                 | 7             | 7           | 7                   | 7                    | 6                     | 7                        | 7             | 7       | 7              |
| 46         | 1                 | 6             | 8           | 7                   | 7                    | 7                     | 6                        | 7             | 8       | 7              |
| 35         | 1                 | 4             | 5           | 6                   | 5                    | 5                     | 4                        | 7             | 2       | 3              |
| Chest.Pain | Coughing.of.Blood | Fatigue       | Weight.Loss | Shortness.of.Breath | wheezing             | Swallowing.Difficulty | Clubbing.of.Finger.Nails | Frequent.Cold |         |                |
| 2          | 4                 | 3             | 4           | 2                   | 2                    | 3                     | 1                        | 2             |         |                |
| 2          | 3                 | 1             | 3           | 7                   | 8                    | 6                     | 2                        | 1             |         |                |
| 4          | 8                 | 8             | 7           | 9                   | 2                    | 1                     | 4                        | 6             |         |                |
| 7          | 8                 | 4             | 2           | 3                   | 1                    | 4                     | 5                        | 6             |         |                |
| 7          | 9                 | 3             | 2           | 4                   | 1                    | 4                     | 2                        | 4             |         |                |
| 4          | 8                 | 8             | 7           | 9                   | 2                    | 1                     | 4                        | 6             |         |                |
| Dry.Cough  | Snoring           |               |             |                     |                      |                       |                          |               |         |                |
| 3          | 4                 |               |             |                     |                      |                       |                          |               |         |                |
| 7          | 2                 |               |             |                     |                      |                       |                          |               |         |                |
| 7          | 2                 |               |             |                     |                      |                       |                          |               |         |                |
| 7          | 5                 |               |             |                     |                      |                       |                          |               |         |                |
| 2          | 3                 |               |             |                     |                      |                       |                          |               |         |                |
| 7          | 2                 |               |             |                     |                      |                       |                          |               |         |                |

.



## 02 분석 방법

lc\_cor <- cor(lung\_cancerC) # 상관분석

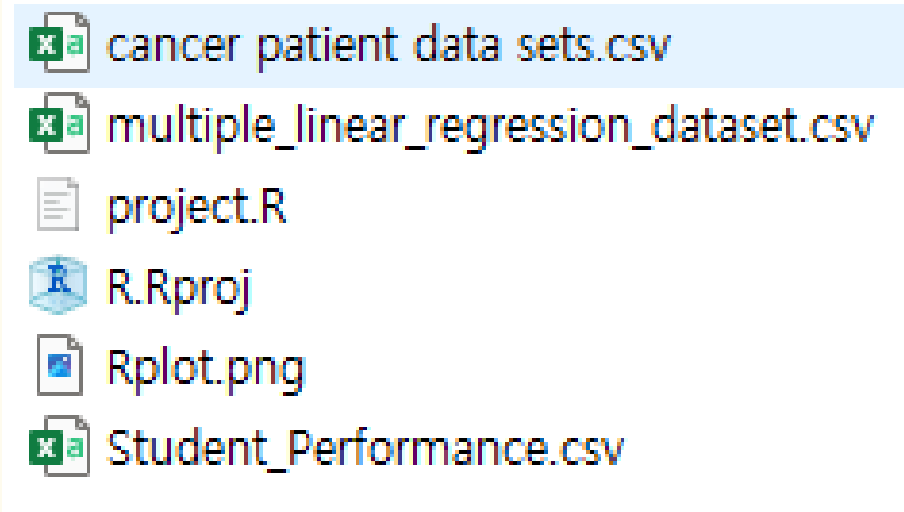
|                          | Age               | Gender        | Air.Pollution | Alcohol.use         | Dust.Allergy | OccuPational.Hazards  | Genetic.Risk             | chronic.Lung.Disease | Obesity     | Smoking      | Passive.Smoker | Chest.Pain   |
|--------------------------|-------------------|---------------|---------------|---------------------|--------------|-----------------------|--------------------------|----------------------|-------------|--------------|----------------|--------------|
| Age                      | 1.000000000       | -0.2020861307 | 0.09949419    | 0.1517417           | 0.03520170   | 0.062177375           | 0.07315054               | 0.128951642          | 0.03433716  | 0.07533258   | 0.004907618    | 0.012863549  |
| Gender                   | -0.202086131      | 1.000000000   | -0.24691184   | -0.2276359          | -0.20431216  | -0.192343411          | -0.22272747              | -0.205060621         | -0.12381258 | -0.20692427  | -0.184826132   | -0.218425815 |
| Air.Pollution            | 0.099494194       | -0.2469118436 | 1.000000000   | 0.7472926           | 0.63750350   | 0.608924458           | 0.70527606               | 0.626700905          | 0.60146750  | 0.48190161   | 0.606763704    | 0.585733513  |
| Alcohol.use              | 0.151741723       | -0.2276359165 | 0.74729261    | 1.00000000          | 0.81864352   | 0.878785921           | 0.87720989               | 0.763575773          | 0.66931156  | 0.54703463   | 0.592576385    | 0.717242276  |
| Dust.Allergy             | 0.035201697       | -0.2043121641 | 0.63750350    | 0.8186435           | 1.00000000   | 0.835859771           | 0.78790388               | 0.619555925          | 0.70067582  | 0.35869058   | 0.560002475    | 0.639983122  |
| OccuPational.Hazards     | 0.062177375       | -0.1923434108 | 0.60892446    | 0.8787859           | 0.83585977   | 1.00000000            | 0.89304852               | 0.858283853          | 0.72219074  | 0.49769258   | 0.555310666    | 0.775618729  |
| Genetic.Risk             | 0.073150538       | -0.2227274663 | 0.70527606    | 0.8772099           | 0.78790388   | 0.893048523           | 1.00000000               | 0.836230827          | 0.72982608  | 0.54325927   | 0.609071292    | 0.831750826  |
| chronic.Lung.Disease     | 0.128951642       | -0.2050606213 | 0.62670091    | 0.7635758           | 0.61955592   | 0.858283853           | 0.83623083               | 1.000000000          | 0.60175439  | 0.57858529   | 0.572698231    | 0.782646079  |
| Obesity                  | 0.034337163       | -0.1238125823 | 0.60146750    | 0.6693116           | 0.70067582   | 0.722190745           | 0.72982608               | 0.601754393          | 1.00000000  | 0.48679470   | 0.681888855    | 0.673150217  |
| Smoking                  | 0.075332578       | -0.2069242711 | 0.48190161    | 0.5470346           | 0.35869058   | 0.497692577           | 0.54325927               | 0.578585291          | 0.48679470  | 1.00000000   | 0.761621504    | 0.647926124  |
| Passive.Smoker           | 0.004907618       | -0.1848261320 | 0.60676370    | 0.5925764           | 0.56000248   | 0.555310666           | 0.60907129               | 0.572698231          | 0.68188885  | 0.76162150   | 1.000000000    | 0.696077265  |
| Chest.Pain               | 0.012863549       | -0.2184258147 | 0.58573351    | 0.7172423           | 0.63998312   | 0.775618729           | 0.83175083               | 0.782646079          | 0.67315022  | 0.64792612   | 0.696077265    | 1.000000000  |
| Coughing.of.Blood        | 0.053006399       | -0.1465053387 | 0.60782860    | 0.6676118           | 0.62529147   | 0.645946503           | 0.63223641               | 0.602986717          | 0.81480477  | 0.55528945   | 0.636222852    | 0.712158462  |
| Fatigue                  | 0.095058772       | -0.1164665840 | 0.21172390    | 0.2372451           | 0.33247156   | 0.267843992           | 0.23053044               | 0.247697060          | 0.55278817  | 0.20002932   | 0.377918692    | 0.251134875  |
| Weight.Loss              | 0.106945701       | -0.0579932590 | 0.25801612    | 0.2078511           | 0.32175619   | 0.176225579           | 0.27174268               | 0.104079762          | 0.31349493  | -0.21293705  | 0.058335615    | -0.001092243 |
| Shortness.of.Breath      | 0.035329285       | -0.0459715849 | 0.26955773    | 0.4357853           | 0.51868168   | 0.366481599           | 0.45820047               | 0.182425753          | 0.40620323  | -0.02325872  | 0.062948479    | 0.237044557  |
| Wheezing                 | -0.095354094      | -0.0763038662 | 0.05536764    | 0.1808170           | 0.30485003   | 0.178925472           | 0.20497278               | 0.057214055          | 0.09428706  | -0.04705954  | 0.200283185    | 0.107210569  |
| Swallowing.Difficulty    | -0.105832694      | -0.0583237829 | -0.08091767   | -0.1140732          | 0.03114127   | -0.002853115          | -0.06294835              | 0.007279413          | 0.12721284  | 0.23614119   | 0.348922307    | 0.071784297  |
| Clubbing.of.Finger.Nails | 0.039258302       | -0.0342191887 | 0.24106478    | 0.4149921           | 0.34571423   | 0.366446760           | 0.35781514               | 0.298022573          | 0.14909338  | -0.04114690  | -0.035535962   | 0.081386377  |
| Frequent.Cold            | -0.012706476      | -0.0005255951 | 0.17453909    | 0.1807778           | 0.21938921   | 0.077166008           | 0.08709161               | 0.028758662          | 0.28836813  | 0.03958531   | 0.104553265    | 0.042936926  |
| Dry.Cough                | 0.012127598       | -0.1230008344 | 0.26148864    | 0.2112772           | 0.30019510   | 0.159887039           | 0.19439933               | 0.114161406          | 0.20061819  | 0.01010135   | 0.120760846    | 0.142180146  |
| Snoring                  | -0.004699822      | -0.1816184730 | -0.02134255   | 0.1226940           | 0.05284449   | 0.022916085           | -0.05683068              | 0.043374535          | 0.03942235  | 0.18905460   | 0.247943324    | 0.140035648  |
|                          | Coughing.of.Blood | Fatigue       | Weight.Loss   | Shortness.of.Breath | Wheezing     | Swallowing.Difficulty | Clubbing.of.Finger.Nails | Frequent.Cold        | Dry.Cough   | Snoring      |                |              |
| Age                      | 0.05300640        | 0.09505877    | 0.106945701   | 0.03532928          | -0.09535409  | -0.105832694          | -0.03925830              | -0.0127064759        | 0.01212760  | -0.004699822 |                |              |
| Gender                   | -0.14650534       | -0.11646658   | -0.057993259  | -0.04597158         | -0.07630387  | -0.058323783          | -0.03421919              | -0.0005255951        | -0.12300083 | -0.181618473 |                |              |
| Air.Pollution            | 0.60782860        | 0.21172390    | 0.258016123   | 0.26955773          | 0.05536764   | -0.080917667          | 0.24106478               | 0.1745390868         | 0.26148864  | -0.021342551 |                |              |
| Alcohol.use              | 0.66761184        | 0.23724513    | 0.207851084   | 0.43578528          | 0.18081698   | -0.114073172          | 0.41499215               | 0.1807778230         | 0.21127720  | 0.122693969  |                |              |
| Dust.Allergy             | 0.62529147        | 0.33247156    | 0.321756185   | 0.51868168          | 0.30485003   | 0.031141271           | 0.34571423               | 0.2193892111         | 0.30019510  | 0.052844486  |                |              |
| OccuPational.Hazards     | 0.64594650        | 0.26784399    | 0.176225579   | 0.36648160          | 0.17892547   | -0.002853115          | 0.36644676               | 0.0771660079         | 0.15988704  | 0.022916085  |                |              |
| Genetic.Risk             | 0.63223641        | 0.23053044    | 0.271742677   | 0.45820047          | 0.20497278   | -0.062948348          | 0.35781514               | 0.0870916062         | 0.19439933  | -0.056830680 |                |              |
| chronic.Lung.Disease     | 0.60298672        | 0.24769706    | 0.104079762   | 0.18242575          | 0.05721405   | 0.007279413           | 0.29802257               | 0.0287586624         | 0.11416141  | 0.043374535  |                |              |
| Obesity                  | 0.81480477        | 0.55278817    | 0.313494929   | 0.40620323          | 0.09428706   | 0.127212844           | 0.14909338               | 0.2883681320         | 0.20061819  | 0.039422347  |                |              |
| Smoking                  | 0.55528945        | 0.20002932    | -0.212937046  | -0.02325872         | -0.04705954  | 0.236141193           | -0.04114690              | 0.0395853062         | 0.01010135  | 0.189054603  |                |              |
| Passive.Smoker           | 0.63622285        | 0.37791869    | 0.058335615   | 0.06294848          | 0.20028319   | 0.348922307           | -0.03553596              | 0.1045532648         | 0.12076085  | 0.247943324  |                |              |
| Chest.Pain               | 0.71215846        | 0.25113487    | -0.001092243  | 0.23704456          | 0.10721057   | 0.071784297           | 0.08138638               | 0.0429369265         | 0.14218015  | 0.140035648  |                |              |
| Coughing.of.Blood        | 1.00000000        | 0.48153994    | 0.105857405   | 0.31877673          | -0.08569837  | 0.086289234           | -0.06644287              | 0.2442350363         | 0.14765857  | 0.087943617  |                |              |
| Fatigue                  | 0.48153994        | 1.00000000    | 0.469516784   | 0.39862499          | 0.17447749   | 0.149561513           | 0.04069401               | 0.4079150533         | 0.27116684  | 0.231748331  |                |              |
| Weight.Loss              | 0.10585741        | 0.46951678    | 1.000000000   | 0.57449652          | 0.33117897   | 0.053383884           | 0.37648356               | 0.1603483653         | 0.18859815  | -0.189106497 |                |              |
| Shortness.of.Breath      | 0.31877673        | 0.39862499    | 0.574496520   | 1.00000000          | 0.20756440   | -0.200477190          | 0.47427466               | 0.3514885104         | 0.49333053  | -0.159291086 |                |              |
| Wheezing                 | -0.08569837       | 0.17447749    | 0.331178974   | 0.20756440          | 1.00000000   | 0.393486838           | 0.33827094               | 0.0988549946         | 0.05438837  | 0.116182515  |                |              |
| Swallowing.Difficulty    | 0.08628923        | 0.14956151    | 0.053383884   | -0.20047719         | 0.39348684   | 1.00000000            | -0.11974052              | 0.1323632497         | -0.05542807 | 0.210539865  |                |              |
| Clubbing.of.Finger.Nails | -0.06644287       | 0.04069401    | 0.376483555   | 0.47427466          | 0.33827094   | -0.119740522          | 1.00000000               | 0.2425293894         | 0.30727080  | -0.017537310 |                |              |
| Frequent.Cold            | 0.24423504        | 0.40791505    | 0.160348365   | 0.35148851          | 0.09885499   | 0.132363250           | 0.24252939               | 1.0000000000         | 0.51591822  | 0.335843523  |                |              |
| Dry.Cough                | 0.14765857        | 0.27116684    | 0.188598147   | 0.49333053          | 0.05438837   | -0.055428066          | 0.30727080               | 0.5159182208         | 1.00000000  | 0.176145730  |                |              |
| Snoring                  | 0.08794362        | 0.23174833    | -0.189106497  | -0.15929109         | 0.11618251   | 0.210539865           | -0.01753731              | 0.3358435232         | 0.17614573  | 1.000000000  |                |              |

### 03 상관분석을 위한 데이터 준비

cancer patient data sets.csv 데이터 셋 : 폐암 예측 데이터

| Age | Gender | Air Pollutic | Alcohol us | Dust Allerg | OccuPatio | Genetic Ri | chronic Lu | Balanced I | Obesity | Smoking | Passive Sm | Chest Pain | Coughing | Fatigue | Weight Lo | Shortness | Wheezing | Swallowing | Clubbing c | Frequent C | Dry Cough | Snoring | Level  |
|-----|--------|--------------|------------|-------------|-----------|------------|------------|------------|---------|---------|------------|------------|----------|---------|-----------|-----------|----------|------------|------------|------------|-----------|---------|--------|
| 33  | 1      | 2            | 4          | 5           | 4         | 3          | 2          | 2          | 4       | 3       | 2          | 2          | 4        | 3       | 4         | 2         | 2        | 3          | 1          | 2          | 3         | 4       | Low    |
| 17  | 1      | 3            | 1          | 5           | 3         | 4          | 2          | 2          | 2       | 2       | 4          | 2          | 3        | 1       | 3         | 7         | 8        | 6          | 2          | 1          | 7         | 2       | Medium |
| 35  | 1      | 4            | 5          | 6           | 5         | 5          | 4          | 6          | 7       | 2       | 3          | 4          | 8        | 8       | 7         | 9         | 2        | 1          | 4          | 6          | 7         | 2       | High   |
| 37  | 1      | 7            | 7          | 7           | 7         | 6          | 7          | 7          | 7       | 7       | 7          | 7          | 8        | 4       | 2         | 3         | 1        | 4          | 5          | 6          | 7         | 5       | High   |
| 46  | 1      | 6            | 8          | 7           | 7         | 7          | 6          | 7          | 7       | 8       | 7          | 7          | 9        | 3       | 2         | 4         | 1        | 4          | 2          | 4          | 2         | 3       | High   |
| 35  | 1      | 4            | 5          | 6           | 5         | 5          | 4          | 6          | 7       | 2       | 3          | 4          | 8        | 8       | 7         | 9         | 2        | 1          | 4          | 6          | 7         | 2       | High   |
| 52  | 2      | 2            | 4          | 5           | 4         | 3          | 2          | 2          | 4       | 3       | 2          | 2          | 4        | 3       | 4         | 2         | 2        | 3          | 1          | 2          | 3         | 4       | Low    |
| 28  | 2      | 3            | 1          | 4           | 3         | 2          | 3          | 4          | 3       | 1       | 4          | 3          | 1        | 3       | 2         | 2         | 4        | 2          | 2          | 3          | 4         | 3       | Low    |
| 35  | 2      | 4            | 5          | 6           | 5         | 6          | 5          | 5          | 5       | 6       | 6          | 6          | 5        | 1       | 4         | 3         | 2        | 4          | 6          | 2          | 4         | 1       | Medium |
| 46  | 1      | 2            | 3          | 4           | 2         | 4          | 3          | 3          | 3       | 2       | 3          | 4          | 4        | 1       | 2         | 4         | 6        | 5          | 4          | 2          | 1         | 5       | Medium |
| 44  | 1      | 6            | 7          | 7           | 7         | 7          | 6          | 7          | 7       | 7       | 8          | 7          | 7        | 5       | 3         | 2         | 7        | 8          | 2          | 4          | 5         | 3       | High   |
| 64  | 2      | 6            | 8          | 7           | 7         | 7          | 6          | 7          | 7       | 7       | 8          | 7          | 7        | 9       | 6         | 5         | 7        | 2          | 4          | 3          | 1         | 4       | High   |

cancer patient data sets.csv 파일을 프로젝트 폴더에 drag





### 03 상관분석을 위한 데이터 준비

```
lung_cancer <- read.csv("cancer patient data sets.csv", fileEncoding = "UTF-8")
```

```
> str(lung_cancer)
'data.frame': 1000 obs. of 26 variables:
 $ index      : int  0 1 2 3 4 5 6 7 8 9 ...
 $ Patient.Id : chr  "P1" "P10" "P100" "P1000" ...
 $ Age        : int  33 17 35 37 46 35 52 28 35 46 ...
 $ Gender     : int  1 1 1 1 1 1 2 2 2 1 ...
 $ Air.Pollution : int  2 3 4 7 6 4 2 3 4 2 ...
 $ Alcohol.use : int  4 1 5 7 8 5 4 1 5 3 ...
 $ Dust.Allergy : int  5 5 6 7 7 6 5 4 6 4 ...
 $ Occupational.Hazards : int  4 3 5 7 7 5 4 3 5 2 ...
 $ Genetic.Risk : int  3 4 5 6 7 5 3 2 6 4 ...
 $ chronic.Lung.Disease : int  2 2 4 7 6 4 2 3 5 3 ...
 $ Balanced.Diet : int  2 2 6 7 7 6 2 4 5 3 ...
 $ Obesity     : int  4 2 7 7 7 7 4 3 5 3 ...
 $ Smoking     : int  3 2 2 7 8 2 3 1 6 2 ...
 $ Passive.Smoker : int  2 4 3 7 7 3 2 4 6 3 ...
 $ Chest.Pain   : int  2 2 4 7 7 4 2 3 6 4 ...
 $ Coughing.of.Blood : int  4 3 8 8 9 8 4 1 5 4 ...
 $ Fatigue      : int  3 1 8 4 3 8 3 3 1 1 ...
 $ Weight.Loss  : int  4 3 7 2 2 7 4 2 4 2 ...
 $ Shortness.of.Breath : int  2 7 9 3 4 9 2 2 3 4 ...
 $ Wheezing     : int  2 8 2 1 1 2 2 4 2 6 ...
 $ Swallowing.Difficulty : int  3 6 1 4 4 1 3 2 4 5 ...
 $ Clubbing.of.Finger.Nails : int  1 2 4 5 2 4 1 2 6 4 ...
 $ Frequent.Cold : int  2 1 6 6 4 6 2 3 2 2 ...
 $ Dry.Cough    : int  3 7 7 7 2 7 3 4 4 1 ...
 $ Snoring     : int  4 2 2 5 3 2 4 3 1 5 ...
 $ Level       : chr  "Low" "Medium" "High" "High" ...
```

## 04 상관분석

1,2,11,26 열을 제외하고 별도의 데이터 셋 구성(11 : 역인과관계, 26 : 문자)

```
lung_cancerC <- lung_cancer[, -c(1, 2, 11, 26 )]
```

```
> str(lung_cancerC)
'data.frame': 1000 obs. of 22 variables:
 $ Age          : int  33 17 35 37 46 35 52 28 35 46 ...
 $ Gender       : int  1 1 1 1 1 1 2 2 2 1 ...
 $ Air.Pollution : int  2 3 4 7 6 4 2 3 4 2 ...
 $ Alcohol.use  : int  4 1 5 7 8 5 4 1 5 3 ...
 $ Dust.Allergy : int  5 5 6 7 7 6 5 4 6 4 ...
 $ Occupational.Hazards : int  4 3 5 7 7 5 4 3 5 2 ...
 $ Genetic.Risk : int  3 4 5 6 7 5 3 2 6 4 ...
 $ chronic.Lung.Disease : int  2 2 4 7 6 4 2 3 5 3 ...
 $ obesity      : int  4 2 7 7 7 7 4 3 5 3 ...
 $ Smoking      : int  3 2 2 7 8 2 3 1 6 2 ...
 $ Passive.Smoker : int  2 4 3 7 7 3 2 4 6 3 ...
 $ Chest.Pain   : int  2 2 4 7 7 4 2 3 6 4 ...
 $ Coughing.of.Blood : int  4 3 8 8 9 8 4 1 5 4 ...
 $ Fatigue      : int  3 1 8 4 3 8 3 3 1 1 ...
 $ weight.Loss  : int  4 3 7 2 2 7 4 2 4 2 ...
 $ Shortness.of.Breath : int  2 7 9 3 4 9 2 2 3 4 ...
 $ wheezing     : int  2 8 2 1 1 2 2 4 2 6 ...
 $ Swallowing.Difficulty : int  3 6 1 4 4 1 3 2 4 5 ...
 $ Clubbing.of.Finger.Nails : int  1 2 4 5 2 4 1 2 6 4 ...
 $ Frequent.Cold : int  2 1 6 6 4 6 2 3 2 2 ...
 $ Dry.Cough    : int  3 7 7 7 2 7 3 4 4 1 ...
 $ Snoring      : int  4 2 2 5 3 2 4 3 1 5 ...
```



## 04 상관분석

lc\_cor <- cor(lung\_cancerC) # cor() 함수를 이용해 상관분석 실시

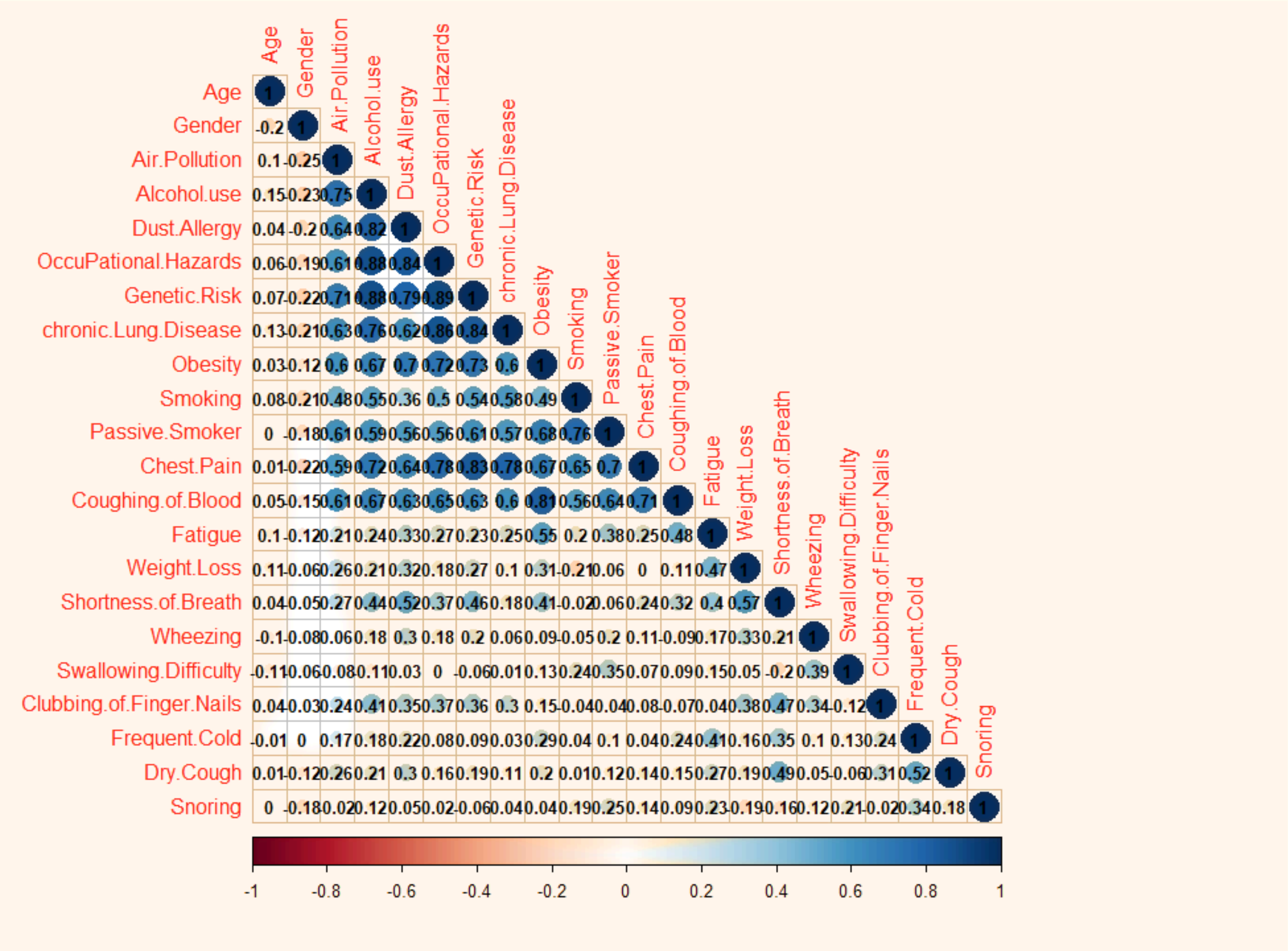
결론 : 만성 폐 질환과 상관관계(0.5이상)가 있는 것은 대기 오염, 알코올 섭취, 먼지 알레르기, 직업적 위험, 유전적 위험, 비만, 흡연, 간접흡연, 가슴 통증, 객혈 등이 있다.

|                          | Age          | Gender        | Air.Pollution | Alcohol.use | Dust.Allergy | OccuPational.Hazards | Genetic.Risk | chronic.Lung.Disease | Obesity       | Smoking     | Passive.Smoker | Chest.Pain   |
|--------------------------|--------------|---------------|---------------|-------------|--------------|----------------------|--------------|----------------------|---------------|-------------|----------------|--------------|
| Age                      | 1.000000000  | -0.2020861307 | 0.09949419    | 0.1517417   | 0.03520170   | 0.062177375          | 0.07315054   | 0.128951642          | 0.03433716    | 0.07533258  | 0.004907618    | 0.012863549  |
| Gender                   | -0.202086131 | 1.000000000   | -0.24691184   | -0.2276359  | -0.20431216  | -0.192343411         | -0.22272747  | -0.205060621         | -0.12381258   | -0.20692427 | -0.184826132   | -0.218425815 |
| Air.Pollution            | 0.099494194  | -0.2469118436 | 1.000000000   | 0.7472926   | 0.63750350   | 0.608924458          | 0.70527606   | 0.626700905          | 0.60146750    | 0.48190161  | 0.606763704    | 0.585733513  |
| Alcohol.use              | 0.151741723  | -0.2276359165 | 0.74729261    | 1.00000000  | 0.81864352   | 0.878785921          | 0.87720989   | 0.763575773          | 0.66931156    | 0.54703463  | 0.592576385    | 0.717242276  |
| Dust.Allergy             | 0.035201697  | -0.2043121641 | 0.63750350    | 0.8186435   | 1.00000000   | 0.835859771          | 0.78790388   | 0.619555925          | 0.70067582    | 0.35869058  | 0.560002475    | 0.639983122  |
| OccuPational.Hazards     | 0.062177375  | -0.1923434108 | 0.60892446    | 0.8787859   | 0.83585977   | 1.00000000           | 0.89304852   | 0.858283853          | 0.72219074    | 0.49769258  | 0.555310666    | 0.775618729  |
| Genetic.Risk             | 0.073150538  | -0.2227274663 | 0.70527606    | 0.8772099   | 0.78790388   | 0.893048523          | 1.00000000   | 0.836230827          | 0.72982608    | 0.54325927  | 0.609071292    | 0.831750826  |
| chronic.Lung.Disease     | 0.128951642  | -0.2050606213 | 0.62670091    | 0.7635758   | 0.61955592   | 0.858283853          | 0.83623083   | 1.000000000          | 0.60175439    | 0.57858529  | 0.572698231    | 0.782646079  |
| Obesity                  | 0.034337163  | -0.1238125823 | 0.60146750    | 0.6693116   | 0.70067582   | 0.722190745          | 0.72982608   | 0.601754393          | 1.00000000    | 0.48679470  | 0.681888855    | 0.673150217  |
| Smoking                  | 0.075332578  | -0.2069242711 | 0.48190161    | 0.5470346   | 0.35869058   | 0.497692577          | 0.54325927   | 0.578585291          | 0.48679470    | 1.00000000  | 0.761621504    | 0.647926124  |
| Passive.Smoker           | 0.004907618  | -0.1848261320 | 0.60676370    | 0.5925764   | 0.56000248   | 0.555310666          | 0.60907129   | 0.572698231          | 0.68188885    | 0.76162150  | 1.000000000    | 0.696077265  |
| Chest.Pain               | 0.012863549  | -0.2184258147 | 0.58573351    | 0.7172423   | 0.63998312   | 0.775618729          | 0.83175083   | 0.782646079          | 0.67315022    | 0.64792612  | 0.696077265    | 1.000000000  |
| Coughing.of.Blood        | 0.053006399  | -0.1465053387 | 0.60782860    | 0.6676118   | 0.62529147   | 0.645946503          | 0.63223641   | 0.602986717          | 0.81480477    | 0.55528945  | 0.636222852    | 0.712158462  |
| Fatigue                  | 0.095058772  | -0.1164665840 | 0.21172390    | 0.2372451   | 0.33247156   | 0.267843992          | 0.23053044   | 0.247697060          | 0.55278817    | 0.20002932  | 0.377918692    | 0.251134875  |
| Weight.Loss              | 0.106945701  | -0.0579932590 | 0.25801612    | 0.2078511   | 0.32175619   | 0.176225579          | 0.27174268   | 0.104079762          | 0.31349493    | -0.21293705 | 0.058335615    | -0.001092243 |
| Shortness.of.Breath      | 0.035329285  | -0.0459715849 | 0.26955773    | 0.4357853   | 0.51868168   | 0.366481599          | 0.45820047   | 0.182425753          | 0.40620323    | -0.02325872 | 0.062948479    | 0.237044557  |
| Wheezing                 | -0.095354094 | -0.0763038662 | 0.05536764    | 0.1808170   | 0.30485003   | 0.178925472          | 0.20497278   | 0.057214055          | 0.09428706    | -0.04705954 | 0.200283185    | 0.107210569  |
| Swallowing.Difficulty    | -0.105832694 | -0.0583237829 | -0.08091767   | -0.1140732  | 0.03114127   | -0.002853115         | -0.06294835  | 0.007279413          | 0.12721284    | 0.23614119  | 0.348922307    | 0.071784297  |
| Clubbing.of.Finger.Nails | 0.039258302  | -0.0342191887 | 0.24106478    | 0.4149921   | 0.34571423   | 0.366446760          | 0.35781514   | 0.298022573          | 0.14909338    | -0.04114690 | -0.035535962   | 0.081386377  |
| Frequent.Cold            | -0.012706476 | -0.0005255951 | 0.17453909    | 0.1807778   | 0.21938921   | 0.077166008          | 0.08709161   | 0.028758662          | 0.28836813    | 0.03958531  | 0.104553265    | 0.042936926  |
| Dry.Cough                | 0.012127598  | -0.1230008344 | 0.26148864    | 0.2112772   | 0.30019510   | 0.159887039          | 0.19439933   | 0.114161406          | 0.20061819    | 0.01010135  | 0.120760846    | 0.142180146  |
| Snoring                  | -0.004699822 | -0.1816184730 | -0.02134255   | 0.1226940   | 0.05284449   | 0.022916085          | -0.05683068  | 0.043374535          | 0.03942235    | 0.18905460  | 0.247943324    | 0.140035648  |
| Coughing.of.Blood        | 0.05300640   | 0.09505877    | 0.106945701   | 0.03532928  | -0.09535409  | -0.105832694         | -0.105832694 | 0.03925830           | -0.0127064759 | 0.01212760  | -0.004699822   |              |
| Gender                   | -0.14650534  | -0.11646658   | -0.057993259  | -0.04597158 | -0.07630387  | -0.058323783         | -0.058323783 | -0.03421919          | -0.0005255951 | -0.12300083 | -0.181618473   |              |
| Air.Pollution            | 0.60782860   | 0.21172390    | 0.258016123   | 0.26955773  | 0.05536764   | -0.080917667         | -0.080917667 | 0.24106478           | 0.1745390868  | 0.26148864  | -0.021342551   |              |
| Alcohol.use              | 0.66761184   | 0.23724513    | 0.207851084   | 0.43578528  | 0.18081698   | -0.114073172         | -0.114073172 | 0.41499215           | 0.1807778230  | 0.21127720  | 0.122693969    |              |
| Dust.Allergy             | 0.62529147   | 0.33247156    | 0.321756185   | 0.51868168  | 0.30485003   | 0.031141271          | 0.031141271  | 0.34571423           | 0.2193892111  | 0.30019510  | 0.052844486    |              |
| OccuPational.Hazards     | 0.64594650   | 0.26784399    | 0.176225579   | 0.36648160  | 0.17892547   | -0.002853115         | -0.002853115 | 0.36644676           | 0.0771660079  | 0.15988704  | 0.022916085    |              |
| Genetic.Risk             | 0.63223641   | 0.23053044    | 0.271742677   | 0.45820047  | 0.20497278   | -0.062948348         | -0.062948348 | 0.35781514           | 0.0870916062  | 0.19439933  | -0.056830680   |              |
| chronic.Lung.Disease     | 0.60298672   | 0.24769706    | 0.104079762   | 0.18242575  | 0.05721405   | 0.007279413          | 0.007279413  | 0.29802257           | 0.0287586624  | 0.11416141  | 0.043374535    |              |
| Obesity                  | 0.81480477   | 0.55278817    | 0.313494929   | 0.40620323  | 0.09428706   | 0.127212844          | 0.127212844  | 0.14909338           | 0.2883681320  | 0.20061819  | 0.039422347    |              |
| Smoking                  | 0.55528945   | 0.20002932    | -0.212937046  | -0.02325872 | -0.04705954  | 0.236141193          | 0.236141193  | -0.04114690          | 0.0395853062  | 0.01010135  | 0.189054603    |              |
| Passive.Smoker           | 0.63622285   | 0.37791869    | 0.058335615   | 0.06294848  | 0.20028319   | 0.348922307          | 0.348922307  | -0.03553596          | 0.1045532648  | 0.12076085  | 0.247943324    |              |
| Chest.Pain               | 0.71215846   | 0.25113487    | -0.001092243  | 0.23704456  | 0.10721057   | 0.071784297          | 0.071784297  | 0.08138638           | 0.0429369265  | 0.14218015  | 0.140035648    |              |
| Coughing.of.Blood        | 1.00000000   | 0.48153994    | 0.105857405   | 0.31877673  | -0.08569837  | 0.086289234          | 0.086289234  | -0.06644287          | 0.2442350363  | 0.14765857  | 0.087943617    |              |
| Fatigue                  | 0.48153994   | 1.00000000    | 0.469516784   | 0.39862499  | 0.17447749   | 0.149561513          | 0.149561513  | 0.04069401           | 0.4079150533  | 0.27116684  | 0.231748331    |              |
| Weight.Loss              | 0.10585741   | 0.46951678    | 1.000000000   | 0.57449652  | 0.33117897   | 0.053383884          | 0.053383884  | 0.37648356           | 0.1603483653  | 0.18859815  | -0.189106497   |              |
| Shortness.of.Breath      | 0.31877673   | 0.39862499    | 0.574496520   | 1.00000000  | 0.20756440   | -0.200477190         | -0.200477190 | 0.47427466           | 0.3514885104  | 0.49333053  | -0.159291086   |              |
| Wheezing                 | -0.08569837  | 0.17447749    | 0.331178974   | 0.20756440  | 1.00000000   | 0.393486838          | 0.393486838  | 0.33827094           | 0.0988549946  | 0.05438837  | 0.116182515    |              |
| Swallowing.Difficulty    | 0.08628923   | 0.14956151    | 0.053383884   | -0.20047719 | 0.39348684   | 1.000000000          | 1.000000000  | -0.11974052          | 0.1323632497  | -0.05542807 | 0.210539865    |              |
| Clubbing.of.Finger.Nails | -0.06644287  | 0.04069401    | 0.376483555   | 0.47427466  | 0.33827094   | -0.119740522         | -0.119740522 | 1.000000000          | 0.2425293894  | 0.30727080  | -0.017537310   |              |
| Frequent.Cold            | 0.24423504   | 0.40791505    | 0.160348365   | 0.35148851  | 0.09885499   | 0.132363250          | 0.132363250  | 0.24252939           | 1.0000000000  | 0.51591822  | 0.335843523    |              |
| Dry.Cough                | 0.14765857   | 0.27116684    | 0.188598147   | 0.49333053  | 0.05438837   | -0.055428066         | -0.055428066 | 0.30727080           | 0.5159182208  | 1.000000000 | 0.176145730    |              |
| Snoring                  | 0.08794362   | 0.23174833    | -0.189106497  | -0.15929109 | 0.11618251   | 0.210539865          | 0.210539865  | -0.01753731          | 0.3358435232  | 0.17614573  | 1.000000000    |              |



05 상관분석 결과 표현

```
> install.packages("corrplot")
> library(corrplot)
> corrplot(lc_cor, method="circle", type="lower",
addCoef.col = "black", number.cex = 0.8)
```



corrplot 패키지 활용 상관분석 결과

## 06 분석에 사용한 전체 스크립트

---

```
lung_cancer <- read.csv("cancer patient data sets.csv", fileEncoding = "UTF-8")
str(lung_cancer)

lung_cancerC <- lung_cancer[, -c(1, 2, 11, 26 )]
lung_cancerC
str(lung_cancerC)

lc_cor <- cor(lung_cancerC)
lc_cor

install.packages("corrplot")
library(corrplot)
corrplot(lc_cor, method="circle", type="lower", addCoef.col = "black", number.cex = 0.8)
```

만성 폐 질환에 영향을 미치는 주요 요인들:

- 1. 대기 오염
- 2. 알코올 섭취
- 3. 먼지 알레르기
- 4. 직업적 위험 요인
- 5. 유전적 위험 요인
- 6. 비만
- 7. 흡연
- 8. 간접 흡연
- 9. 가슴 통증
- 10. 객혈

이 10개의 변수들은 만성 폐 질환과 0.5 이상의 상관계수를 보여, 이들 요인이 만성 폐 질환 발병에 중요한 영향을 미치는 것으로 나타났습니다.

이러한 결과는 만성 폐 질환 예방 및 관리를 위해 해당 요인들을 고려해야 함을 시사합니다.

예를 들어 대기 오염 관리, 금연 및 금주 캠페인, 직업 안전 강화, 유전적 위험군 관리 등의 정책 마련이 필요할 것으로 보입니다.

### 사용한 데이터

cancer patient data sets.csv

### 참고 사이트

<https://www.kaggle.com/datasets/thedevastator/cancer-patients-and-air-pollution-a-new-link>

# 감사합니다

---