

```

import pandas as pd

# CSV 파일을 읽어옴
file_path = 'C:/yelin/file/데이터/배양 전처리/배양검사결과.csv'
data = pd.read_csv(file_path)

# 필요한 열만 남기기
columns_to_keep = ['VTHOS_NO', 'RGNO', 'ANTBT_NM', 'RSLT_CONT']
data = data[columns_to_keep]

# null값 있는 행 제거
data = data.dropna()

# 'RSLT_CONT' 열에서 'NEG', 'POS', 'I', 'D', 'ETC', 'NULL' 값을 가진 행 제외
filtered_data = data[~data['RSLT_CONT'].isin(['NEG', 'POS', 'I', 'D', 'ETC', 'NULL'])]

# 결과를 CSV 파일로 저장
output_file_path = 'C:/yelin/file/데이터/배양 전처리/배양1.csv'
filtered_data.to_csv(output_file_path, index=False)

import pandas as pd

# CSV 파일을 읽어옴
file_path = 'C:/yelin/file/데이터/배양 전처리/배양1.csv'
data = pd.read_csv(file_path)

# ANTBT_NM 열에 해당 조건을 만족하는 행 제거
unwanted_antbt = ['ESBL', 'β-lactamase', 'Inducible Clindarmycin Resistance', 'Cefocxitin Screen']
data = data[~data['ANTBT_NM'].isin(unwanted_antbt)]

# 결과를 CSV 파일로 저장
output_file_path = 'C:/yelin/file/데이터/배양 전처리/배양2.csv'
data.to_csv(output_file_path, index=False)

import pandas as pd

# CSV 파일 경로 설정
csv_file_path = 'C:/yelin/file/데이터/배양 전처리/배양2.csv'

# CSV 파일 읽기
data = pd.read_csv(csv_file_path)

# "ANTBT_NM"와 "RSLT_CONT" 열만 선택
data = data[['ANTBT_NM', 'RSLT_CONT']]

# Pivot 테이블 생성 (약물 이름을 행, 결과를 열로)
pivot_table = pd.pivot_table(data, index='ANTBT_NM', columns='RSLT_CONT', aggfunc=len, fill_value=0)

# 행렬 출력
print(pivot_table)

RSLT_CONT
ANTBT_NM
Amikacin      1010  19358
Amoxicillin/Clavulanic acid  2940  6260
Amphotericin B      0    113
Ampicillin     14959  6558
Ampicillin/Sulbactam  2161  3613
...          ...    ...
Tobramycin     1155  3642
Trimethoprim/Sulfamethoxazole  6459  18643
Trimethoprim/sulfamethoxazole    7    237
Vancomycin     1396  11072
Voriconazole    7    101

[70 rows x 2 columns]

# 피벗 테이블을 CSV 파일로 출력
output_csv_file_path = 'C:/yelin/sp/배양/배양3.csv'
pivot_table.to_csv(output_csv_file_path)

```

=> support, confidence, lift는 엑셀에서 값 구하고 처리했음

