TEAM 10(최준헌, 박해원, 정세화)

HOMEWORK: 2ND PROJECT

- Analysis of the association of established datasets based on infringement indicators

OUTLINE

> It checks whether normalized data sets using artifacts extracted through intrusion incident-based malicious code performed by applying the FP-growth algorithm can be used for algorithms such as correlation analysis, machine learning, and artificial intelligence.

- Based on the data set previously built in Project 1, the association analysis is performed focusing on the classification and year of infringement incidents.
- Through the FP-growth algorithm, each dataset or two or more datasets are combined to generate various rules.

TABLE SORTING ACCORDING TO THE SUPPORT VALUE & RULES ACCORDING TO THE TABLE

#1 Malware name: Downloader

support	itemsets
1.0	frozenset({'TCP'})
1.0	frozenset({'TCP', 'MD5'})
1.0	frozenset({'MD5', 'TCP', 'HTTP'})
1.0	frozenset({'TCP', 'HTTP'})
1.0	frozenset({'MD5', 'HTTP'})
1.0	frozenset({'MD5'})
1.0	frozenset({'HTTP'})
0.9444444444444440	frozenset({'c:\\', 'TCP', 'HTTP'})
0.9444444444444440	frozenset({'c:\\', 'TCP'})
0.9444444444444440	frozenset({'c:\\', 'MD5', 'TCP', 'HTTP'})
0.9444444444444440	frozenset({'c:\\', 'TCP', 'MD5'})
0.9444444444444440	frozenset({'c:\\', 'MD5', 'HTTP'})
0.9444444444444440	frozenset({'c:\\', 'MD5'})
0.9444444444444440	frozenset({'c:\\', 'HTTP'})
0.9444444444444440	frozenset({'c:\\'})
0.8333333333333330	frozenset({'REG_SZ', 'TCP', 'MD5'})
0.8333333333333330	frozenset({'MD5', '80.0', 'TCP', 'HTTP'})
0.8333333333333330	frozenset({'80.0', 'TCP', 'MD5'})

0.8333333333333330	frozenset({'MD5', '80.0', 'HTTP'})
0.8333333333333330	frozenset({'80.0', 'TCP'})
0.8333333333333330	frozenset({'80.0', 'MD5'})
0.8333333333333330	frozenset({'80.0', 'HTTP'})
0.8333333333333330	frozenset({'80.0', 'TCP', 'HTTP'})
0.8333333333333330	frozenset({'MD5', 'REG_SZ', 'TCP', 'HTTP'})
0.8333333333333330	frozenset({'80.0'})
0.8333333333333330	frozenset({'REG_SZ', 'TCP', 'HTTP'})
0.8333333333333330	frozenset({'MD5', 'REG_SZ', 'HTTP'})
0.8333333333333330	frozenset({'REG_SZ', 'TCP'})
0.8333333333333330	frozenset({'REG_SZ'})
0.8333333333333330	frozenset({'REG_SZ', 'MD5'})
0.8333333333333330	frozenset({'REG_SZ', 'HTTP'})
0.7777777777777780	frozenset({'c:\\', 'REG_SZ', 'MD5'})
0.7777777777777780	frozenset({'c:\\', 'TCP', 'HTTP', '80.0', 'MD5'})
0.7777777777777780	frozenset({'c:\\', 'MD5', '80.0', 'HTTP'})
0.7777777777777780	frozenset({'c:\\', '80.0', 'TCP', 'HTTP'})
0.7777777777777780	frozenset({'c:\\', '80.0', 'TCP', 'MD5'})
0.7777777777777780	frozenset({'c:\\', '80.0', 'HTTP'})
0.7777777777777780	frozenset({'c:\\', '80.0', 'MD5'})
0.7777777777777780	frozenset({'c:\\', '80.0', 'TCP'})
0.7777777777777780	frozenset({'c:\\', 'REG_SZ', 'HTTP'})
0.7777777777777780	frozenset({'c:\\', '80.0'})
0.7777777777777780	frozenset({'c:\\', 'REG_SZ'})
0.7777777777777780	frozenset({'c:\\', 'REG_SZ', 'TCP', 'HTTP', 'MD5'})
0.7777777777777780	frozenset({'c:\\', 'REG_SZ', 'TCP', 'MD5'})
0.7777777777777780	frozenset({'c:\\', 'REG_SZ', 'TCP', 'HTTP'})
0.7777777777777780	frozenset({'c:\\', 'MD5', 'REG_SZ', 'HTTP'})
0.7777777777777780	frozenset({'c:\\', 'REG_SZ', 'TCP'})
0.7222222222222220	frozenset({'80.0', 'REG_SZ'})
0.7222222222222220	frozenset({'80.0', 'REG_SZ', 'TCP'})
0.7222222222222220	frozenset({'80.0', 'REG_SZ', 'MD5'})
0.7222222222222220	frozenset({'80.0', 'REG_SZ', 'HTTP'})
0.7222222222222220	frozenset({'80.0', 'REG_SZ', 'TCP', 'MD5'})
0.7222222222222220	frozenset({'80.0', 'REG_SZ', 'TCP', 'HTTP'})
0.7222222222222220	frozenset({'80.0', 'MD5', 'REG_SZ', 'HTTP'})
0.7222222222222220	frozenset({'REG_SZ', 'TCP', 'HTTP', '80.0', 'MD5'})

#2 Malware name: Ransom

support	itemsets
1.0	frozenset({'TCP'})
1.0	frozenset({'MD5'})

1.0	frozenset({'TCP', 'HTTP'})
1.0	frozenset({'MD5', 'HTTP'})
1.0	frozenset({'TCP', 'MD5'})
1.0	frozenset({'MD5', 'TCP', 'HTTP'})
1.0	frozenset({'HTTP'})
0.9333333333333330	frozenset({'MD5', 'SERVICE_WIN32_OWN_PROCESS', 'TCP', 'SERVICE_AUTO_START'})
0.9333333333333330	frozenset({'MD5', 'TCP', 'HTTP', 'SERVICE_AUTO_START'})
0.9333333333333330	frozenset({'SERVICE_AUTO_START', 'TCP', 'HTTP', 'SERVICE_WIN32_OWN_PROCESS', 'MD5'})
0.9333333333333330	frozenset({'MD5', 'SERVICE_AUTO_START'})
0.9333333333333330	frozenset({'SERVICE_WIN32_OWN_PROCESS', 'TCP', 'HTTP', 'SERVICE_AUTO_START'})
0.9333333333333330	frozenset({'MD5', 'SERVICE_WIN32_OWN_PROCESS', 'HTTP', 'SERVICE_AUTO_START'})
0.9333333333333330	frozenset({'MD5', 'TCP', 'SERVICE_AUTO_START'})
0.9333333333333330	frozenset({'TCP', 'HTTP', 'SERVICE_AUTO_START'})
0.9333333333333330	frozenset({'MD5', 'HTTP', 'SERVICE_AUTO_START'})
0.9333333333333330	frozenset({'SERVICE_WIN32_OWN_PROCESS', 'TCP', 'SERVICE_AUTO_START'})
0.9333333333333330	frozenset({'MD5', 'SERVICE_WIN32_OWN_PROCESS', 'SERVICE_AUTO_START'})
0.9333333333333330	frozenset({'SERVICE_WIN32_OWN_PROCESS', 'HTTP', 'SERVICE_AUTO_START'})
0.9333333333333330	frozenset({'TCP', 'SERVICE_AUTO_START'})
0.9333333333333330	frozenset({'MD5', 'SERVICE_WIN32_OWN_PROCESS', 'TCP', 'HTTP'})
0.9333333333333330	frozenset({'HTTP', 'SERVICE_AUTO_START'})
0.9333333333333330	frozenset({'SERVICE_WIN32_OWN_PROCESS', 'SERVICE_AUTO_START'})
0.9333333333333330	frozenset({'SERVICE_WIN32_OWN_PROCESS', 'TCP', 'MD5'})
0.9333333333333330	frozenset({'SERVICE_WIN32_OWN_PROCESS', 'TCP', 'HTTP'})
0.9333333333333330	frozenset({'MD5', 'SERVICE_WIN32_OWN_PROCESS', 'HTTP'})
0.9333333333333330	frozenset({'SERVICE_WIN32_OWN_PROCESS', 'TCP'})
0.9333333333333330	frozenset({'SERVICE_WIN32_OWN_PROCESS', 'MD5'})
0.9333333333333330	frozenset({'SERVICE_WIN32_OWN_PROCESS', 'HTTP'})
0.9333333333333330	frozenset({'SERVICE_AUTO_START'})
0.9333333333333330	frozenset({'SERVICE_WIN32_OWN_PROCESS'})
0.8666666666666670	frozenset({'c:\\', 'HTTP', 'MD5'})
0.8666666666666670	frozenset({'c:\\', 'TCP', 'HTTP', 'MD5'})
0.8666666666666670	frozenset({'c:\\', 'TCP', 'MD5'})
0.8666666666666670	frozenset({'c:\\', 'TCP', 'HTTP'})
0.8666666666666670	frozenset({'c:\\', 'HTTP'})
0.8666666666666670	frozenset({'c:\\', 'TCP'})
0.8666666666666670	frozenset({'c:\\', 'MD5'})
0.8666666666666670	frozenset({'c:\\'})
0.8	frozenset({'TCP', 'HTTP', '80.0', 'SERVICE_WIN32_OWN_PROCESS', 'MD5'})
0.8	frozenset({'MD5', '80.0', 'SERVICE_WIN32_OWN_PROCESS', 'SERVICE_AUTO_START'})
0.8	frozenset({'80.0', 'SERVICE_WIN32_OWN_PROCESS', 'SERVICE_AUTO_START'})
0.8	frozenset({'80.0', 'HTTP', 'SERVICE_AUTO_START'})
0.8	frozenset({'MD5', '80.0', 'SERVICE_AUTO_START'})
0.8	frozenset({'80.0', 'TCP', 'SERVICE_AUTO_START'})

0.8	frozenset({'80.0', 'SERVICE_WIN32_OWN_PROCESS', 'HTTP'})
0.8	frozenset({'80.0', 'SERVICE_WIN32_OWN_PROCESS', 'MD5'})
0.8	frozenset({'TCP', '80.0', 'SERVICE_WIN32_OWN_PROCESS'})
0.8	frozenset({'MD5', '80.0', 'HTTP'})
0.8	frozenset({'80.0', 'TCP', 'HTTP'})
0.8	frozenset({'80.0', 'TCP', 'MD5'})
0.8	frozenset({'80.0', 'SERVICE_WIN32_OWN_PROCESS', 'HTTP', 'SERVICE_AUTO_START'})
0.8	frozenset({'MD5', '80.0', 'HTTP', 'SERVICE_AUTO_START'})
0.8	frozenset({'TCP', '80.0', 'SERVICE_WIN32_OWN_PROCESS', 'SERVICE_AUTO_START'})
0.8	frozenset({'SERVICE_AUTO_START', 'TCP', 'HTTP', '80.0', 'MD5'})
0.8	frozenset({'80.0', 'TCP', 'HTTP', 'SERVICE_AUTO_START'})
0.8	frozenset({'MD5', '80.0', 'TCP', 'SERVICE_AUTO_START'})
0.8	frozenset({'SERVICE_AUTO_START', 'TCP', 'HTTP', '80.0', 'SERVICE_WIN32_OWN_PROCESS', 'MD5'})
0.8	frozenset({'MD5', '80.0', 'SERVICE_WIN32_OWN_PROCESS', 'HTTP'})
0.8	frozenset({'TCP', '80.0', 'SERVICE_WIN32_OWN_PROCESS', 'HTTP'})
0.8	frozenset({'TCP', '80.0', 'SERVICE_WIN32_OWN_PROCESS', 'MD5'})
0.8	frozenset({'MD5', '80.0', 'TCP', 'HTTP'})
0.8	frozenset({'SERVICE_AUTO_START', 'HTTP', '80.0', 'SERVICE_WIN32_OWN_PROCESS', 'MD5'})
0.8	frozenset({'80.0', 'MD5'})
0.8	frozenset({'SERVICE_AUTO_START', 'TCP', '80.0', 'SERVICE_WIN32_OWN_PROCESS', 'HTTP'})
0.8	frozenset({'SERVICE_AUTO_START', 'TCP', '80.0', 'SERVICE_WIN32_OWN_PROCESS', 'MD5'})
0.8	frozenset({'80.0', 'TCP'})
0.8	frozenset({'c:\\', 'SERVICE_AUTO_START', 'TCP', 'HTTP', 'SERVICE_WIN32_OWN_PROCESS', 'MD5'})
0.8	frozenset({'80.0', 'HTTP'})
0.8	frozenset({'c:\\', 'MD5', 'TCP', 'SERVICE_AUTO_START'})
0.8	frozenset({'c:\\', 'SERVICE_WIN32_OWN_PROCESS', 'HTTP'})
0.8	frozenset({'c:\\', 'MD5', 'SERVICE_AUTO_START'})
0.8	frozenset({'c:\\', 'SERVICE_WIN32_OWN_PROCESS'})
0.8	frozenset({'c:\\', 'SERVICE_WIN32_OWN_PROCESS', 'MD5'})
0.8	frozenset({'c:\\', 'SERVICE_AUTO_START'})
0.8	frozenset({'c:\\', 'TCP', 'SERVICE_AUTO_START'})
0.8	frozenset({'c:\\', 'SERVICE_WIN32_OWN_PROCESS', 'TCP'})
0.8	frozenset({'c:\\', 'SERVICE_WIN32_OWN_PROCESS', 'SERVICE_AUTO_START'})
0.8	frozenset({'c:\\', 'MD5', 'HTTP', 'SERVICE_AUTO_START'})
0.8	frozenset({'c:\\', 'SERVICE_WIN32_OWN_PROCESS', 'HTTP', 'MD5'})
0.8	frozenset({'c:\\', 'TCP', 'HTTP', 'SERVICE_AUTO_START'})
0.8	frozenset({'c:\\', 'SERVICE_WIN32_OWN_PROCESS', 'TCP', 'HTTP'})
0.8	frozenset({'c:\\', 'SERVICE_WIN32_OWN_PROCESS', 'HTTP', 'SERVICE_AUTO_START'})
0.8	frozenset({'80.0'})
0.8	frozenset({'c:\\', 'TCP', 'HTTP', 'SERVICE_WIN32_OWN_PROCESS', 'MD5'})
0.8	frozenset({'c:\\', 'SERVICE_AUTO_START', 'TCP', 'HTTP', 'MD5'})
0.8	frozenset({'80.0', 'SERVICE_WIN32_OWN_PROCESS'})

0.8	frozenset({'80.0', 'SERVICE_AUTO_START'})
0.8	frozenset({'c:\\', 'SERVICE_AUTO_START', 'TCP', 'SERVICE_WIN32_OWN_PROCESS', 'MD5'})
0.8	frozenset({'c:\\', 'SERVICE_AUTO_START', 'TCP', 'SERVICE_WIN32_OWN_PROCESS', 'HTTP'})
0.8	frozenset({'c:\\', 'SERVICE_AUTO_START', 'HTTP', 'SERVICE_WIN32_OWN_PROCESS', 'MD5'})
0.8	frozenset({'c:\\', 'HTTP', 'SERVICE_AUTO_START'})
0.8	frozenset({'c:\\', 'MD5', 'SERVICE_WIN32_OWN_PROCESS', 'SERVICE_AUTO_START'})
0.8	frozenset({'c:\\', 'SERVICE_WIN32_OWN_PROCESS', 'TCP', 'MD5'})
0.8	frozenset({'c:\\', 'SERVICE_WIN32_OWN_PROCESS', 'TCP', 'SERVICE_AUTO_START'})
0.7333333333333330	frozenset({'c:\\', 'SERVICE_AUTO_START', 'TCP', 'HTTP', '80.0', 'MD5'})
0.7333333333333330	frozenset({'c:\\', 'SERVICE_AUTO_START', 'TCP', '80.0', 'SERVICE_WIN32_OWN_PROCESS'})
0.7333333333333330	frozenset({'c:\\', 'SERVICE_AUTO_START', 'TCP', '80.0', 'SERVICE_WIN32_OWN_PROCESS', 'MD5'})
0.7333333333333330	frozenset({'c:\\', 'SERVICE_AUTO_START', '80.0', 'SERVICE_WIN32_OWN_PROCESS', 'MD5'})
0.7333333333333330	frozenset({'c:\\', 'SERVICE_AUTO_START', 'HTTP', '80.0', 'MD5'})
0.7333333333333330	frozenset({'c:\\', 'SERVICE_AUTO_START', 'TCP', '80.0', 'HTTP'})
0.7333333333333330	frozenset({'c:\\', 'SERVICE_AUTO_START', 'TCP', '80.0', 'MD5'})
0.7333333333333330	frozenset({'c:\\', 'TCP', '80.0', 'SERVICE_WIN32_OWN_PROCESS', 'MD5'})
0.7333333333333330	frozenset({'c:\\', 'TCP', 'HTTP', '80.0', 'SERVICE_WIN32_OWN_PROCESS', 'MD5'})
0.7333333333333330	frozenset({'c:\\', 'HTTP', '80.0', 'SERVICE_WIN32_OWN_PROCESS', 'MD5'})
0.7333333333333330	frozenset({'c:\\', 'SERVICE_AUTO_START', 'TCP', '80.0', 'SERVICE_WIN32_OWN_PROCESS', 'HTTP'})
0.7333333333333330	frozenset({'c:\\', 'TCP', 'HTTP', '80.0', 'MD5'})
0.7333333333333330	frozenset({'c:\\', 'TCP', '80.0', 'SERVICE_WIN32_OWN_PROCESS', 'HTTP'})
0.7333333333333330	frozenset({'c:\\', 'SERVICE_AUTO_START', 'HTTP', '80.0', 'SERVICE_WIN32_OWN_PROCESS', 'MD5'})
0.7333333333333330	frozenset({'c:\\', '80.0', 'SERVICE_WIN32_OWN_PROCESS', 'HTTP'})
0.7333333333333330	frozenset({'c:\\', 'SERVICE_AUTO_START', '80.0', 'SERVICE_WIN32_OWN_PROCESS', 'HTTP'})
0.7333333333333330	frozenset({'c:\\', '80.0', 'HTTP', 'SERVICE_AUTO_START'})
0.7333333333333330	frozenset({'c:\\', '80.0', 'SERVICE_AUTO_START'})
0.7333333333333330	frozenset({'c:\\', '80.0', 'SERVICE_WIN32_OWN_PROCESS'})
0.7333333333333330	frozenset({'c:\\', '80.0', 'HTTP'})
0.7333333333333330	frozenset({'c:\\', '80.0', 'MD5'})
0.7333333333333330	frozenset({'c:\\', '80.0', 'TCP'})
0.7333333333333330	frozenset({'c:\\', '80.0', 'SERVICE_WIN32_OWN_PROCESS', 'SERVICE_AUTO_START'})
0.7333333333333330	frozenset({'MD5', 'c:\\', '80.0', 'SERVICE_AUTO_START'})
0.7333333333333330	frozenset({'c:\\', '80.0', 'TCP', 'MD5'})
0.7333333333333330	frozenset({'c:\\', '80.0', 'TCP', 'SERVICE_AUTO_START'})
0.7333333333333330	frozenset({'c:\\', '80.0'})
0.7333333333333330	frozenset({'c:\\', '80.0', 'SERVICE_WIN32_OWN_PROCESS', 'MD5'})
0.7333333333333330	frozenset({'c:\\', 'TCP', '80.0', 'SERVICE_WIN32_OWN_PROCESS'})
0.7333333333333330	frozenset({'MD5', 'c:\\', '80.0', 'HTTP'})
0.7333333333333330	frozenset({'c:\\', '80.0', 'TCP', 'HTTP'})
0.7333333333333330	frozenset({'c:\\', 'SERVICE_AUTO_START', 'TCP', 'HTTP', '80.0', 'SERVICE_WIN32_OWN_PROCESS', 'MD5'})

#3 Malware name: Exploit

support	itemsets
1.0	frozenset({'TCP'})
1.0	frozenset({'80.0', 'REG_SZ', 'MD5'})
1.0	frozenset({'REG_SZ', 'TCP', 'MD5'})
1.0	frozenset({'REG_SZ', 'TCP', 'HTTP'})
1.0	frozenset({'80.0', 'TCP', 'REG_SZ'})
1.0	frozenset({'TCP', 'HTTP', 'MD5'})
1.0	frozenset({'80.0', 'TCP', 'MD5'})
1.0	frozenset({'80.0', 'TCP', 'HTTP'})
1.0	frozenset({'MD5', 'REG_SZ', 'HTTP'})
1.0	frozenset({'80.0', 'REG_SZ', 'HTTP'})
1.0	frozenset({'80.0', 'MD5'})
1.0	frozenset({'MD5', '80.0', 'HTTP'})
1.0	frozenset({'REG_SZ'})
1.0	frozenset({'80.0', 'TCP', 'REG_SZ', 'MD5'})
1.0	frozenset({'80.0', 'TCP', 'REG_SZ', 'HTTP'})
1.0	frozenset({'80.0', 'TCP', 'HTTP', 'MD5'})
1.0	frozenset({'MD5', '80.0', 'REG_SZ', 'HTTP'})
1.0	frozenset({'REG_SZ', 'TCP', 'HTTP', '80.0', 'MD5'})
1.0	frozenset({'80.0', 'HTTP'})
1.0	frozenset({'REG_SZ', 'TCP', 'HTTP', 'MD5'})
1.0	frozenset({'80.0'})
1.0	frozenset({'80.0', 'REG_SZ'})
1.0	frozenset({'TCP', 'HTTP'})
1.0	frozenset({'MD5'})
1.0	frozenset({'80.0', 'TCP'})
1.0	frozenset({'REG_SZ', 'TCP'})
1.0	frozenset({'REG_SZ', 'MD5'})
1.0	frozenset({'REG_SZ', 'HTTP'})
1.0	frozenset({'TCP', 'MD5'})
1.0	frozenset({'MD5', 'HTTP'})
1.0	frozenset({'HTTP'})
0.8333333333333330	frozenset({'winword.exe', '80.0', 'REG_SZ', 'MD5'})
0.8333333333333330	frozenset({'MD5', 'STD_INPUT_HANDLE', '80.0', 'HTTP'})
0.8333333333333330	frozenset({'REG_SZ', 'TCP', 'HTTP', 'STD_INPUT_HANDLE', 'MD5'})
0.8333333333333330	frozenset({'REG_SZ', 'winword.exe', 'TCP', 'HTTP', 'MD5'})
0.8333333333333330	frozenset({'STD_INPUT_HANDLE', '80.0', 'TCP', 'HTTP'})
0.8333333333333330	frozenset({'winword.exe', 'STD_INPUT_HANDLE', '80.0', 'HTTP'})
0.8333333333333330	frozenset({'winword.exe', 'STD_INPUT_HANDLE', 'TCP', 'HTTP'})
0.8333333333333330	frozenset({'winword.exe', 'STD_INPUT_HANDLE', '80.0', 'MD5'})
0.8333333333333330	frozenset({'MD5', 'STD_INPUT_HANDLE', 'winword.exe', 'HTTP'})
0.8333333333333330	frozenset({'MD5', '80.0', 'winword.exe', 'HTTP'})

0.8333333333333330	frozenset({'MD5', 'STD_INPUT_HANDLE', 'REG_SZ', 'HTTP'})
0.8333333333333330	frozenset({'winword.exe', 'STD_INPUT_HANDLE', '80.0', 'REG_SZ'})
0.8333333333333330	frozenset({'winword.exe', 'STD_INPUT_HANDLE', '80.0', 'TCP'})
0.8333333333333330	frozenset({'winword.exe', 'STD_INPUT_HANDLE', 'REG_SZ', 'HTTP'})
0.8333333333333330	frozenset({'STD_INPUT_HANDLE', '80.0', 'REG_SZ', 'HTTP'})
0.8333333333333330	frozenset({'REG_SZ', 'winword.exe', 'TCP', '80.0', 'MD5'})
0.8333333333333330	frozenset({'MD5', 'REG_SZ', 'winword.exe', 'HTTP'})
0.8333333333333330	frozenset({'winword.exe', 'STD_INPUT_HANDLE', 'REG_SZ', 'MD5'})
0.8333333333333330	frozenset({'STD_INPUT_HANDLE', '80.0', 'REG_SZ', 'MD5'})
0.8333333333333330	frozenset({'winword.exe', '80.0', 'REG_SZ', 'HTTP'})
0.8333333333333330	frozenset({'REG_SZ', 'TCP', 'STD_INPUT_HANDLE', '80.0', 'HTTP'})
0.8333333333333330	frozenset({'REG_SZ', 'TCP', 'STD_INPUT_HANDLE', '80.0', 'MD5'})
0.8333333333333330	frozenset({'REG_SZ', 'winword.exe', 'HTTP', 'STD_INPUT_HANDLE', 'MD5'})
0.8333333333333330	frozenset({'REG_SZ', 'winword.exe', 'HTTP', 'STD_INPUT_HANDLE', '80.0', 'MD5'})
0.8333333333333330	frozenset({'winword.exe', 'TCP', 'HTTP', 'STD_INPUT_HANDLE', '80.0', 'MD5'})
0.8333333333333330	frozenset({'REG_SZ', 'winword.exe', 'TCP', 'STD_INPUT_HANDLE', '80.0', 'HTTP'})
0.8333333333333330	frozenset({'REG_SZ', 'winword.exe', 'TCP', 'STD_INPUT_HANDLE', '80.0', 'MD5'})
0.8333333333333330	frozenset({'REG_SZ', 'winword.exe', 'TCP', 'HTTP', 'STD_INPUT_HANDLE', 'MD5'})
0.8333333333333330	frozenset({'REG_SZ', 'TCP', 'HTTP', 'STD_INPUT_HANDLE', '80.0', 'MD5'})
0.8333333333333330	frozenset({'REG_SZ', 'winword.exe', 'TCP', 'HTTP', '80.0', 'MD5'})
0.8333333333333330	frozenset({'winword.exe', 'HTTP', 'STD_INPUT_HANDLE', '80.0', 'MD5'})
0.8333333333333330	frozenset({'REG_SZ', 'winword.exe', 'STD_INPUT_HANDLE', '80.0', 'HTTP'})
0.8333333333333330	frozenset({'REG_SZ', 'winword.exe', 'STD_INPUT_HANDLE', '80.0', 'MD5'})
0.8333333333333330	frozenset({'REG_SZ', 'HTTP', 'STD_INPUT_HANDLE', '80.0', 'MD5'})
0.8333333333333330	frozenset({'REG_SZ', 'winword.exe', 'TCP', 'STD_INPUT_HANDLE', 'MD5'})
0.8333333333333330	frozenset({'REG_SZ', 'winword.exe', 'HTTP', '80.0', 'MD5'})
0.8333333333333330	frozenset({'winword.exe', 'TCP', 'STD_INPUT_HANDLE', '80.0', 'HTTP'})
0.8333333333333330	frozenset({'winword.exe', 'TCP', 'STD_INPUT_HANDLE', '80.0', 'MD5'})
0.8333333333333330	frozenset({'winword.exe', 'TCP', 'HTTP', 'STD_INPUT_HANDLE', 'MD5'})
0.8333333333333330	frozenset({'TCP', 'HTTP', 'STD_INPUT_HANDLE', '80.0', 'MD5'})
0.8333333333333330	frozenset({'winword.exe', 'TCP', 'HTTP', '80.0', 'MD5'})
0.8333333333333330	frozenset({'REG_SZ', 'winword.exe', 'TCP', 'STD_INPUT_HANDLE', '80.0'})
0.8333333333333330	frozenset({'REG_SZ', 'winword.exe', 'TCP', 'STD_INPUT_HANDLE', 'HTTP'})
0.8333333333333330	frozenset({'winword.exe', 'STD_INPUT_HANDLE', 'TCP', 'MD5'})
0.8333333333333330	frozenset({'REG_SZ', 'winword.exe', 'TCP', '80.0', 'HTTP'})
0.8333333333333330	frozenset({'winword.exe', '80.0', 'TCP', 'HTTP'})
0.8333333333333330	frozenset({'winword.exe', 'MD5'})
0.8333333333333330	frozenset({'STD_INPUT_HANDLE', '80.0', 'TCP', 'MD5'})
0.8333333333333330	frozenset({'winword.exe', '80.0', 'REG_SZ'})
0.8333333333333330	frozenset({'winword.exe', 'REG_SZ', 'HTTP'})
0.8333333333333330	frozenset({'STD_INPUT_HANDLE', 'REG_SZ', 'MD5'})
0.8333333333333330	frozenset({'winword.exe', 'REG_SZ', 'MD5'})
0.8333333333333330	frozenset({'winword.exe', 'TCP'})

0.8333333333333330	frozenset({'winword.exe', 'STD_INPUT_HANDLE', 'TCP'})
0.8333333333333330	frozenset({'STD_INPUT_HANDLE', '80.0', 'TCP'})
0.8333333333333330	frozenset({'winword.exe', '80.0', 'TCP'})
0.8333333333333330	frozenset({'STD_INPUT_HANDLE', 'TCP', 'HTTP'})
0.8333333333333330	frozenset({'winword.exe', 'TCP', 'HTTP'})
0.8333333333333330	frozenset({'STD_INPUT_HANDLE', 'TCP'})
0.8333333333333330	frozenset({'STD_INPUT_HANDLE', 'TCP', 'MD5'})
0.8333333333333330	frozenset({'winword.exe', 'TCP', 'MD5'})
0.8333333333333330	frozenset({'STD_INPUT_HANDLE', 'REG_SZ', 'TCP'})
0.8333333333333330	frozenset({'winword.exe', 'REG_SZ', 'TCP'})
0.8333333333333330	frozenset({'winword.exe', 'REG_SZ'})
0.8333333333333330	frozenset({'STD_INPUT_HANDLE', 'REG_SZ'})
0.8333333333333330	frozenset({'winword.exe', 'STD_INPUT_HANDLE'})
0.8333333333333330	frozenset({'STD_INPUT_HANDLE', '80.0'})
0.8333333333333330	frozenset({'winword.exe', '80.0'})
0.8333333333333330	frozenset({'STD_INPUT_HANDLE', 'HTTP'})
0.8333333333333330	frozenset({'winword.exe', 'HTTP'})
0.8333333333333330	frozenset({'STD_INPUT_HANDLE', 'REG_SZ', 'HTTP'})
0.8333333333333330	frozenset({'STD_INPUT_HANDLE', '80.0', 'REG_SZ'})
0.8333333333333330	frozenset({'winword.exe', '80.0', 'TCP', 'MD5'})
0.8333333333333330	frozenset({'winword.exe', 'STD_INPUT_HANDLE', 'REG_SZ'})
0.8333333333333330	frozenset({'STD_INPUT_HANDLE', 'TCP', 'HTTP', 'MD5'})
0.8333333333333330	frozenset({'winword.exe', 'TCP', 'HTTP', 'MD5'})
0.8333333333333330	frozenset({'STD_INPUT_HANDLE', 'MD5'})
0.8333333333333330	frozenset({'winword.exe', 'STD_INPUT_HANDLE', 'REG_SZ', 'TCP'})
0.8333333333333330	frozenset({'STD_INPUT_HANDLE', '80.0', 'TCP', 'REG_SZ'})
0.8333333333333330	frozenset({'winword.exe', '80.0', 'TCP', 'REG_SZ'})
0.8333333333333330	frozenset({'STD_INPUT_HANDLE', 'REG_SZ', 'TCP', 'HTTP'})
0.8333333333333330	frozenset({'winword.exe', 'REG_SZ', 'TCP', 'HTTP'})
0.8333333333333330	frozenset({'winword.exe'})
0.8333333333333330	frozenset({'STD_INPUT_HANDLE', 'REG_SZ', 'TCP', 'MD5'})
0.8333333333333330	frozenset({'winword.exe', 'REG_SZ', 'TCP', 'MD5'})
0.8333333333333330	frozenset({'STD_INPUT_HANDLE'})
0.8333333333333330	frozenset({'winword.exe', 'STD_INPUT_HANDLE', '80.0'})
0.8333333333333330	frozenset({'winword.exe', 'STD_INPUT_HANDLE', 'HTTP'})
0.8333333333333330	frozenset({'STD_INPUT_HANDLE', '80.0', 'HTTP'})
0.8333333333333330	frozenset({'winword.exe', '80.0', 'HTTP'})
0.8333333333333330	frozenset({'winword.exe', 'STD_INPUT_HANDLE', 'MD5'})
0.8333333333333330	frozenset({'STD_INPUT_HANDLE', '80.0', 'MD5'})
0.8333333333333330	frozenset({'winword.exe', '80.0', 'MD5'})
0.8333333333333330	frozenset({'MD5', 'STD_INPUT_HANDLE', 'HTTP'})
0.8333333333333330	frozenset({'MD5', 'winword.exe', 'HTTP'})
0.8333333333333330	frozenset({'REG_SZ', 'winword.exe', 'TCP', 'HTTP', 'STD_INPUT_HANDLE', '80.0', 'MD5'})

#4	Malware name: Trojan

support	itemsets
1.0	frozenset({'MD5'})
1.0	frozenset({'HTTP'})
1.0	frozenset({'MD5', 'HTTP'})
0.9696969696969700	frozenset({'c:\\'})
0.9696969696969700	frozenset({'c:\\', 'HTTP'})
0.9696969696969700	frozenset({'c:\\', 'MD5'})
0.9696969696969700	frozenset({'c:\\', 'MD5', 'HTTP'})
0.9393939393939390	frozenset({'MD5', 'TCP', 'HTTP'})
0.9393939393939390	frozenset({'TCP'})
0.9393939393939390	frozenset({'TCP', 'HTTP'})
0.9393939393939390	frozenset({'TCP', 'MD5'})
0.9090909090909090	frozenset({'c:\\', 'TCP', 'MD5'})
0.9090909090909090	frozenset({'c:\\', 'TCP', 'HTTP'})
0.9090909090909090	frozenset({'MD5', 'c:\\', 'TCP', 'HTTP'})
0.9090909090909090	frozenset({'c:\\', 'TCP'})
0.8787878787878790	frozenset({'c:\\', 'REG_SZ', 'HTTP'})
0.8787878787878790	frozenset({'REG_SZ'})
0.8787878787878790	frozenset({'MD5', '80.0', 'HTTP'})
0.8787878787878790	frozenset({'c:\\', 'MD5', 'REG_SZ', 'HTTP'})
0.8787878787878790	frozenset({'MD5', 'REG_SZ', 'HTTP'})
0.8787878787878790	frozenset({'c:\\', 'REG_SZ', 'MD5'})
0.8787878787878790	frozenset({'80.0', 'HTTP'})
0.8787878787878790	frozenset({'REG_SZ', 'MD5'})
0.8787878787878790	frozenset({'REG_SZ', 'HTTP'})
0.8787878787878790	frozenset({'c:\\', 'REG_SZ'})
0.8787878787878790	frozenset({'80.0', 'MD5'})
0.8787878787878790	frozenset({'80.0'})
0.8484848484848490	frozenset({'c:\\', '80.0', 'HTTP'})
0.8484848484848490	frozenset({'c:\\', 'MD5', '80.0', 'HTTP'})
0.8484848484848490	frozenset({'MD5', '80.0', 'TCP', 'HTTP'})
0.8484848484848490	frozenset({'80.0', 'TCP', 'HTTP'})
0.8484848484848490	frozenset({'80.0', 'TCP', 'MD5'})
0.8484848484848490	frozenset({'c:\\', '80.0'})
0.8484848484848490	frozenset({'c:\\', '80.0', 'MD5'})
0.8484848484848490	frozenset({'80.0', 'TCP'})
0.8181818181818180	frozenset({'c:\\', 'REG_SZ', 'TCP', 'HTTP', 'MD5'})
0.8181818181818180	frozenset({'c:\\', 'REG_SZ', 'TCP', 'MD5'})
0.8181818181818180	frozenset({'c:\\', 'REG_SZ', 'TCP', 'HTTP'})
0.8181818181818180	frozenset({'REG_SZ', 'TCP', 'MD5'})
0.8181818181818180	frozenset({'REG_SZ', 'TCP', 'HTTP'})

0.8181818181818180	frozenset({'c:\\', 'REG_SZ', 'TCP'})
0.8181818181818180	frozenset({'REG_SZ', 'TCP'})
0.8181818181818180	frozenset({'c:\\', '80.0', 'TCP'})
0.8181818181818180	frozenset({'c:\\', '80.0', 'TCP', 'MD5'})
0.8181818181818180	frozenset({'c:\\', '80.0', 'TCP', 'HTTP'})
0.8181818181818180	frozenset({'c:\\', 'TCP', 'HTTP', '80.0', 'MD5'})
0.8181818181818180	frozenset({'MD5', 'REG_SZ', 'TCP', 'HTTP'})
0.7878787878787880	frozenset({'80.0', 'REG_SZ', 'HTTP'})
0.7878787878787880	frozenset({'c:\\', 'MD5', 'REG_SZ', '80.0', 'HTTP'})
0.7878787878787880	frozenset({'80.0', 'MD5', 'REG_SZ', 'HTTP'})
0.7878787878787880	frozenset({'c:\\', 'REG_SZ', '80.0', 'HTTP'})
0.7878787878787880	frozenset({'c:\\', 'REG_SZ', '80.0', 'MD5'})
0.7878787878787880	frozenset({'80.0', 'REG_SZ', 'MD5'})
0.7878787878787880	frozenset({'c:\\', 'REG_SZ', '80.0'})
0.7878787878787880	frozenset({'80.0', 'REG_SZ'})
0.7575757575757580	frozenset({'80.0', 'REG_SZ', 'TCP'})
0.7575757575757580	frozenset({'MD5', 'd41d8cd98f00b204e9800998ecf8427e', 'HTTP'})
0.7575757575757580	frozenset({'d41d8cd98f00b204e9800998ecf8427e', 'MD5'})
0.7575757575757580	frozenset({'d41d8cd98f00b204e9800998ecf8427e', 'HTTP'})
0.7575757575757580	frozenset({'c:\\', 'MD5', 'REG_SZ', 'TCP', '80.0', 'HTTP'})
0.7575757575757580	frozenset({'REG_SZ', 'TCP', 'HTTP', '80.0', 'MD5'})
0.7575757575757580	frozenset({'c:\\', 'REG_SZ', 'TCP', '80.0', 'HTTP'})
0.7575757575757580	frozenset({'c:\\', 'REG_SZ', 'TCP', '80.0', 'MD5'})
0.7575757575757580	frozenset({'80.0', 'REG_SZ', 'TCP', 'HTTP'})
0.7575757575757580	frozenset({'80.0', 'REG_SZ', 'TCP', 'MD5'})
0.7575757575757580	frozenset({'d41d8cd98f00b204e9800998ecf8427e'})
0.7575757575757580	frozenset({'c:\\', 'REG_SZ', 'TCP', '80.0'})
0.7272727272727270	frozenset({'c:\\', 'd41d8cd98f00b204e9800998ecf8427e'})
0.7272727272727270	frozenset({'c:\\', 'd41d8cd98f00b204e9800998ecf8427e', 'HTTP'})
0.7272727272727270	frozenset({'c:\\', 'd41d8cd98f00b204e9800998ecf8427e', 'MD5'})
0.7272727272727270	frozenset({'MD5', 'c:\\', 'd41d8cd98f00b204e9800998ecf8427e', 'HTTP'})

#5 Malware name: Spy

support	itemsets
0.75	frozenset({'0.0'})

#6 Malware name: Dropper

support	itemsets
1.0	frozenset({'c:\\'})
1.0	frozenset({'TCP'})
1.0	frozenset({'REG_SZ'})

1.0	frozenset({'c:\\', 'TCP'})
1.0	frozenset({'c:\\', 'REG_SZ'})
1.0	frozenset({'REG_SZ', 'TCP'})
1.0	frozenset({'c:\\', 'REG_SZ', 'TCP'})

#7 Malware name: UDS

support	itemsets
1.0	frozenset({'winword.exe'})
1.0	frozenset({'winword.exe', 'c:\\program files\\microsoft office\\root\\office16\\winword.exe', '1112.0'})
1.0	frozenset({'"C:\\Program Files\\Microsoft Office\\Root\\Office16\\WINWORD.EXE" /n', 'c:\\program files\\microsoft office\\root\\office16\\winword.exe', 'C:\\Users\\aETAdzjz\\Desktop\\', '1112.0'})
1.0	frozenset({'"C:\\Program Files\\Microsoft Office\\Root\\Office16\\WINWORD.EXE" /n', 'C:\\Users\\aETAdzjz\\Desktop\\', 'winword.exe', '1112.0'})
1.0	frozenset({'"C:\\Program Files\\Microsoft Office\\Root\\Office16\\WINWORD.EXE" /n', 'c:\\program files\\microsoft office\\root\\office16\\winword.exe', 'winword.exe', '1112.0'})
1.0	frozenset({'"C:\\Program Files\\Microsoft Office\\Root\\Office16\\WINWORD.EXE" /n', 'c:\\program files\\microsoft office\\root\\office16\\winword.exe', 'C:\\Users\\aETAdzjz\\Desktop\\', 'winword.exe'})
1.0	frozenset({'winword.exe', 'c:\\program files\\microsoft office\\root\\office16\\winword.exe', 'C:\\Users\\aETAdzjz\\Desktop\\', '1112.0'})
1.0	frozenset({'"C:\\Program Files\\Microsoft Office\\Root\\Office16\\WINWORD.EXE" /n', 'C:\\Users\\aETAdzjz\\Desktop\\', '1112.0'})
1.0	frozenset({'"C:\\Program Files\\Microsoft Office\\Root\\Office16\\WINWORD.EXE" /n', 'c:\\program files\\microsoft office\\root\\office16\\winword.exe', '1112.0'})
1.0	frozenset({'"C:\\Program Files\\Microsoft Office\\Root\\Office16\\WINWORD.EXE" /n', 'c:\\program files\\microsoft office\\root\\office16\\winword.exe', 'C:\\Users\\aETAdzjz\\Desktop\\'})
1.0	frozenset({'c:\\program files\\microsoft office\\root\\office16\\winword.exe', 'C:\\Users\\aETAdzjz\\Desktop\\', '1112.0'})
1.0	frozenset({'"C:\\Program Files\\Microsoft Office\\Root\\Office16\\WINWORD.EXE" /n', 'winword.exe', '1112.0'})
1.0	frozenset({'"C:\\Program Files\\Microsoft Office\\Root\\Office16\\WINWORD.EXE" /n', 'C:\\Users\\aETAdzjz\\Desktop\\', 'winword.exe'})
1.0	frozenset({'winword.exe', 'C:\\Users\\aETAdzjz\\Desktop\\', '1112.0'})
1.0	frozenset({'"C:\\Program Files\\Microsoft Office\\Root\\Office16\\WINWORD.EXE" /n', 'c:\\program files\\microsoft office\\root\\office16\\winword.exe', 'winword.exe'})
1.0	frozenset({'winword.exe', 'c:\\program files\\microsoft office\\root\\office16\\winword.exe', 'C:\\Users\\aETAdzjz\\Desktop\\'})
1.0	frozenset({'c:\\program files\\microsoft office\\root\\office16\\winword.exe'})
1.0	frozenset({'"C:\\Program Files\\Microsoft Office\\Root\\Office16\\WINWORD.EXE" /n', '1112.0'})
1.0	frozenset({'"C:\\Program Files\\Microsoft Office\\Root\\Office16\\WINWORD.EXE" /n', 'C:\\Users\\aETAdzjz\\Desktop\\'})
1.0	frozenset({'C:\\Users\\aETAdzjz\\Desktop\\', '1112.0'})
1.0	frozenset({'"C:\\Program Files\\Microsoft Office\\Root\\Office16\\WINWORD.EXE" /n', 'c:\\program files\\microsoft office\\root\\office16\\winword.exe'})
1.0	frozenset({'c:\\program files\\microsoft office\\root\\office16\\winword.exe', '1112.0'})
1.0	frozenset({'c:\\program files\\microsoft office\\root\\office16\\winword.exe', 'C:\\Users\\aETAdzjz\\Desktop\\'})
1.0	frozenset({'"C:\\Program Files\\Microsoft Office\\Root\\Office16\\WINWORD.EXE" /n', 'winword.exe'})
1.0	frozenset({'winword.exe', '1112.0'})
1.0	frozenset({'winword.exe', 'C:\\Users\\aETAdzjz\\Desktop\\'})
1.0	frozenset({'winword.exe', 'c:\\program files\\microsoft office\\root\\office16\\winword.exe'})
1.0	frozenset({'"C:\\Program Files\\Microsoft Office\\Root\\Office16\\WINWORD.EXE" /n'})

1.0	frozenset({'1112.0'})
1.0	frozenset({'C:\\Users\\aETAdzjz\\Desktop\\'})
1.0	frozenset({'1112.0', '"C:\\Program Files\\Microsoft Office\\Root\\Office16\\WINWORD.EXE" /n', 'winword.exe', 'c:\\program files\\microsoft office\\root\\office16\\winword.exe', 'C:\\Users\\aETAdzjz\\Desktop\\'})

SOURCE CODES

```
import pandas as pd
import numpy as np
from mlxtend.preprocessing import TransactionEncoder
from mlxtend.frequent_patterns import association_rules
from mlxtend.frequent_patterns import fpgrowth

dataset_df = pd.read_csv('./dataset.csv', index_col='file_name')

malware_type_dict = {}

# 같은 family_name 을 가진 행끼리 모아서 dictionary 생성
# key: family_name, value: 해당 family_name 을 가진 row 들의 list
for row in dataset_df.iterrows():
    # 해당 family_name 과 일치하는 key 가 dictionary 에 존재하지 않는 경우, 빈
list 를 value 로 가지는 원소 생성
    if not row[1]['Family'] in malware_type_dict:
        malware_type_dict[row[1]['Family']] = pd.DataFrame()

    # 해당 family_name 과 일치하는 원소의 value 에 row 추가
    # row 의 type 은 tuple 이기 때문에, list 로 타입변경해줌
    malware_type_dict[row[1]['Family']] =
malware_type_dict[row[1]['Family']].append(row[1].drop('Family'),

ignore_index=True)

# print(malware_type_dict)

# 컬럼 전체 데이터가 NaN 경우 column drop
# 컬럼 데이터가 한 번씩 나올 경우 column drop

for key, df in malware_type_dict.items():
    for col in df.columns:
        # 컬럼 전체가 NaN 인 경우 or 컬럼 데이터가 한 번씩 나올 경우
        if df[col].isna().sum() == len(df) or
len(df[col].dropna().unique()) == len(df[col].dropna()):
            # if key == 'Downloader':
            #     print(col)
            df = df.drop(col, axis=1)
        else:
            df[col] = df[col].fillna(df[col].mode()[0])

    df = df.dropna(axis=0)
    malware_type_dict[key] = df
```

```
for key, value in malware_type_dict.items():

print("\n\n================================================================
================================")
    print("악성코드 명: {}".format(key)) # Family_name 출력
    print("해당하는 파일 개수: {}".format(len(value))) # 해당 Family_name 을
가진 row 의 개수 출력
    # print("\n")

    # 파일 한 개일 경우 연관성 분석 건너뜀
    if len(value) == 1:
        continue

    dataset_arr = value.to_numpy(dtype=str)

    # # row 생략 없이 출력
    # pd.set_option('display.max_rows', None)
    # # col 생략 없이 출력
    # pd.set_option('display.max_columns', None)

    pd.set_option('display.max_colwidth', 1)

    # 연관성 분석
    te = TransactionEncoder()
    te_ary = te.fit(dataset_arr).transform(dataset_arr)
    df = pd.DataFrame(te_ary, columns=te.columns_)
    result = fpgrowth(df, min_support=0.7, use_colnames=True)
    print("\n=====연관성 분석=====")
    print(result.sort_values(by=['support'], ascending=False))
    # print("\n")
    # result_chart = association_rules(result, metric="confidence",
min_threshold=0.5)
    # print('<result chart>')
    # print(result_chart)
    # print('\n')
```