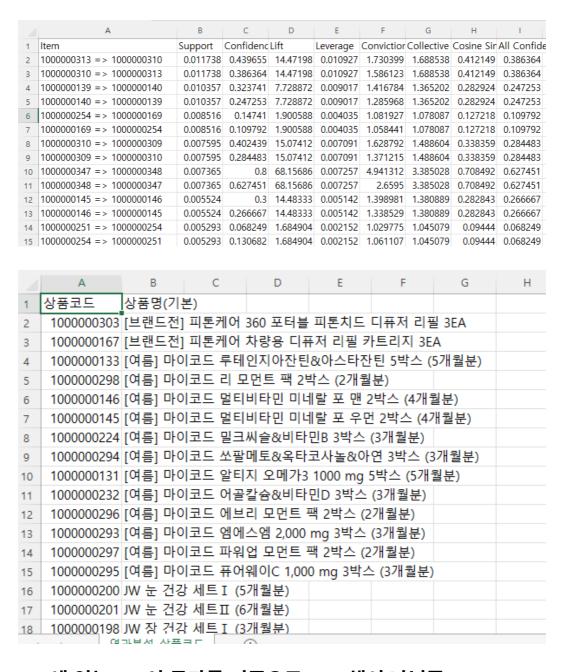


연관분석 결과 매칭

• 연관분석 결과로 나온 item에 있는 코드와 본자료에 있는 저 코드가 해당하는 값을 찾아 매칭시켜주려고 한다.



1. ITEM에 있는 '⇒'이 문자를 기준으로 split해서 나눠줌

code1값에 0번쨰 string값을 넣어줌

연관분석 결과 매칭 1

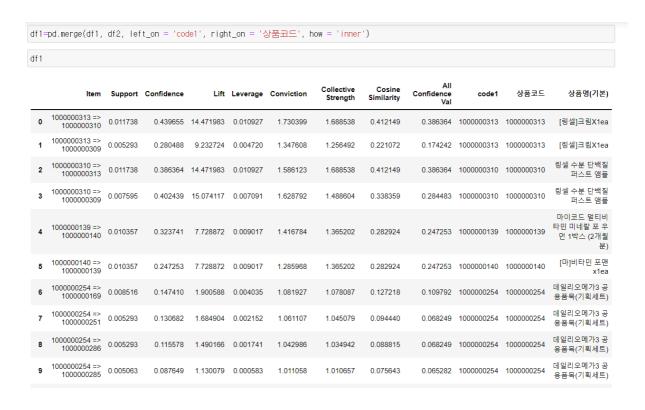
```
df1['code1']=df1.ltem.str.split('=>').str[0]
df1['code1']=df1['code1'].astype('int')

df1
```

	Item	Support	Confidence	Lift	Leverage	Conviction	Collective Strength	Cosine Similarity	All Confidence Val	code1
0	1000000313 => 1000000310	0.011738	0.439655	14.471983	0.010927	1.730399	1.688538	0.412149	0.386364	1000000313
1	1000000310 => 1000000313	0.011738	0.386364	14.471983	0.010927	1.586123	1.688538	0.412149	0.386364	1000000310
2	1000000139 => 1000000140	0.010357	0.323741	7.728872	0.009017	1.416784	1.365202	0.282924	0.247253	1000000139
3	1000000140 => 1000000139	0.010357	0.247253	7.728872	0.009017	1.285968	1.365202	0.282924	0.247253	1000000140
4	1000000254 => 1000000169	0.008516	0.147410	1.900588	0.004035	1.081927	1.078087	0.127218	0.109792	1000000254
5	1000000169 => 1000000254	0.008516	0.109792	1.900588	0.004035	1.058441	1.078087	0.127218	0.109792	1000000169
6	1000000310 => 1000000309	0.007595	0.402439	15.074117	0.007091	1.628792	1.488604	0.338359	0.284483	1000000310
7	1000000309 => 1000000310	0.007595	0.284483	15.074117	0.007091	1.371215	1.488604	0.338359	0.284483	1000000309
8	1000000347 => 1000000348	0.007365	0.800000	68.156863	0.007257	4.941312	3.385028	0.708492	0.627451	1000000347
9	1000000348 => 1000000347	0.007365	0.627451	68.156863	0.007257	2.659500	3.385028	0.708492	0.627451	1000000348
10	1000000145 => 1000000146	0.005524	0.300000	14.483333	0.005142	1.398981	1.380889	0.282843	0.266667	1000000145
11	1000000146 => 1000000145	0 005524	0 266667	14 483333	0 005142	1 338529	1 380889	0 282843	0 266667	1000000146

2. merge

• merge를 통해서 서로 다른 명칭의 이름을 가진 데이터 프레임을 how='inner'을 통해 존재하는 값는만 추출



연관분석 결과 매칭 2