



연관분석_결과 매칭

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

	A	B	C	D	E	F	G	H	I
1	Item	Support	Confidenc	Lift	Leverage	Convictior	Collective	Cosine Sir	All Confide
2	1000000313 => 1000000310	0.011738	0.439655	14.47198	0.010927	1.730399	1.688538	0.412149	0.386364
3	1000000310 => 1000000313	0.011738	0.386364	14.47198	0.010927	1.586123	1.688538	0.412149	0.386364
4	1000000139 => 1000000140	0.010357	0.323741	7.728872	0.009017	1.416784	1.365202	0.282924	0.247253
5	1000000140 => 1000000139	0.010357	0.247253	7.728872	0.009017	1.285968	1.365202	0.282924	0.247253
6	1000000254 => 1000000169	0.008516	0.14741	1.900588	0.004035	1.081927	1.078087	0.127218	0.109792
7	1000000169 => 1000000254	0.008516	0.109792	1.900588	0.004035	1.058441	1.078087	0.127218	0.109792
8	1000000310 => 1000000309	0.007595	0.402439	15.07412	0.007091	1.628792	1.488604	0.338359	0.284483
9	1000000309 => 1000000310	0.007595	0.284483	15.07412	0.007091	1.371215	1.488604	0.338359	0.284483
10	1000000347 => 1000000348	0.007365	0.8	68.15686	0.007257	4.941312	3.385028	0.708492	0.627451
11	1000000348 => 1000000347	0.007365	0.627451	68.15686	0.007257	2.6595	3.385028	0.708492	0.627451
12	1000000145 => 1000000146	0.005524	0.3	14.48333	0.005142	1.398981	1.380889	0.282843	0.266667
13	1000000146 => 1000000145	0.005524	0.266667	14.48333	0.005142	1.338529	1.380889	0.282843	0.266667
14	1000000251 => 1000000254	0.005293	0.068249	1.684904	0.002152	1.029775	1.045079	0.09444	0.068249
15	1000000254 => 1000000251	0.005293	0.130682	1.684904	0.002152	1.061107	1.045079	0.09444	0.068249

	A	B	C	D	E	F	G	H
1	상품코드	상품명(기본)						
2	1000000303	[브랜드전] 피톤케어 360 포터블 피톤치드 디퓨저 리필 3EA						
3	1000000167	[브랜드전] 피톤케어 차량용 디퓨저 리필 카트리지 3EA						
4	1000000133	[여름] 마이코드 루테인지아잔틴&아스타잔틴 5박스 (5개월분)						
5	1000000298	[여름] 마이코드 리 모먼트 팩 2박스 (2개월분)						
6	1000000146	[여름] 마이코드 멀티비타민 미네랄 포 맨 2박스 (4개월분)						
7	1000000145	[여름] 마이코드 멀티비타민 미네랄 포 우먼 2박스 (4개월분)						
8	1000000224	[여름] 마이코드 밀크씨슬&비타민B 3박스 (3개월분)						
9	1000000294	[여름] 마이코드 쏘팔메토&옥타코사놀&아연 3박스 (3개월분)						
10	1000000131	[여름] 마이코드 알티지 오메가3 1000 mg 5박스 (5개월분)						
11	1000000232	[여름] 마이코드 어골칼슘&비타민D 3박스 (3개월분)						
12	1000000296	[여름] 마이코드 에브리 모먼트 팩 2박스 (2개월분)						
13	1000000293	[여름] 마이코드 엠에스엠 2,000 mg 3박스 (3개월분)						
14	1000000297	[여름] 마이코드 파워업 모먼트 팩 2박스 (2개월분)						
15	1000000295	[여름] 마이코드 퓨어웨이C 1,000 mg 3박스 (3개월분)						
16	1000000200	JW 눈 건강 세트 I (5개월분)						
17	1000000201	JW 눈 건강 세트II (6개월분)						
18	1000000198	JW 장 건강 세트 I (3개월분)						

1. ITEM에 있는 '⇒'이 문자를 기준으로 split해서 나눠줌

- code1값에 0번째 string값을 넣어줌

```
df1['code1']=df1.Item.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 값만 추출

```
df1=pd.merge(df1, df2, left_on = 'code1', right_on = '상품명', how = 'inner')
```

df1

	Item	Support	Confidence	Lift	Leverage	Conviction	Collective Strength	Cosine Similarity	All Confidence Val	code1	상품명	상품명(기본)
0	1000000313 => 1000000310	0.011738	0.439655	14.471983	0.010927	1.730399	1.688538	0.412149	0.386364	1000000313	1000000313	[링셀]크림X1ea
1	1000000313 => 1000000309	0.005293	0.280488	9.232724	0.004720	1.347608	1.256492	0.221072	0.174242	1000000313	1000000313	[링셀]크림X1ea
2	1000000310 => 1000000313	0.011738	0.386364	14.471983	0.010927	1.586123	1.688538	0.412149	0.386364	1000000310	1000000310	링셀 수분 단백질 퍼스트 앰플
3	1000000310 => 1000000309	0.007595	0.402439	15.074117	0.007091	1.628792	1.488604	0.338359	0.284483	1000000310	1000000310	링셀 수분 단백질 퍼스트 앰플
4	1000000139 => 1000000140	0.010357	0.323741	7.728872	0.009017	1.416784	1.365202	0.282924	0.247253	1000000139	1000000139	마이코드 멀티비타민 미네랄 포우먼 1박스 (2개월분)
5	1000000140 => 1000000139	0.010357	0.247253	7.728872	0.009017	1.285968	1.365202	0.282924	0.247253	1000000140	1000000140	[마]비타민 포먼 x1ea
6	1000000254 => 1000000169	0.008516	0.147410	1.900588	0.004035	1.081927	1.078087	0.127218	0.109792	1000000254	1000000254	데일리오메가3 공용품목(기획세트)
7	1000000254 => 1000000251	0.005293	0.130682	1.684904	0.002152	1.061107	1.045079	0.094440	0.068249	1000000254	1000000254	데일리오메가3 공용품목(기획세트)
8	1000000254 => 1000000286	0.005293	0.115578	1.490166	0.001741	1.042986	1.034942	0.088815	0.068249	1000000254	1000000254	데일리오메가3 공용품목(기획세트)
9	1000000254 => 1000000285	0.005063	0.087649	1.130079	0.000583	1.011058	1.010657	0.075643	0.065282	1000000254	1000000254	데일리오메가3 공용품목(기획세트)