Tarefa 1

Nome: Danilo Scutti Lima

RA: 263867

Leia o dataset em http://fimi.ua.ac.be/data/retail.dat (http://fimi.ua.ac.be/data/retail.dat) que é um dataset real de compras no varejo em uma loja na Belgica.

Descubra as regras de associação que tenham suporte mínimo de 0.005 e confiança minima de 0.9

Biblioteca utilizada

http://rasbt.github.io/mlxtend/ (http://rasbt.github.io/mlxtend/)

In [1]:

```
#!pip install mlxtend
```

Lendo conjunto de dados de compras

In [2]:

```
retail = [i.strip().split() for i in open("retail.dat").readlines()]
```

Suporte

In [ ]:

```
from mlxtend.preprocessing import TransactionEncoder
from mlxtend.frequent_patterns import apriori, fpmax, fpgrowth
import pandas as pd

pd.set_option('display.max_rows', 600)

te = TransactionEncoder()
te_ary = te.fit(retail).transform(retail)
df = pd.DataFrame(te_ary, columns=te.columns_)
```

In [7]:

```
frequent_itemsets = apriori(df, min_support=0.005, use_colnames=True)
print("[APRIORI] Quantidade de regras de associação com suporte mínimo de 0.005 =", fre
display(frequent_itemsets.sort_values("support"))
```

```
[APRIORI] Quantidade de regras de associação com suporte mínimo de 0.005 = 5
80
```

	support	itemsets
325	0.005002	(39, 269)
448	0.005013	(772, 48)
125	0.005013	(365)
359	0.005025	(334, 48)
547	0.005025	(39, 48, 589)
460	0.005025	(32, 38, 110)
272	0.005036	(48, 1600)
469	0.005036	(39, 48, 13041)
151	0.005036	(47)
158	0.005048	(490)
32	0.005059	(1404)
334	0.005059	(272, 39)
552	0.005059	(39, 48, 677)
352	0.005070	(32, 475)
410	0.005070	(789, 39)
579	0.005082	(32, 39, 38, 41, 48)
14	0.005082	(1113)
235	0.005093	(32, 110)
459	0.005093	(39, 38, 105)
45	0.005104	(1585)
55	0.005104	(1704)
572	0.005116	(32, 39, 48, 89)
293	0.005116	(1814, 48)
485	0.005116	(39, 48, 201)
501	0.005116	(41, 39, 271)
394	0.005127	(55, 39)
221	0.005127	(39, 10)
483	0.005127	(39, 179, 48)
498	0.005127	(39, 48, 258)
404	0.005138	(703, 39)
108	0.005138	(281)
203	0.005150	(831)

	support	itemsets
229	0.005161	(39, 105)
457	0.005172	(48, 976)
566	0.005195	(39, 48, 286, 38)
240	0.005195	(39, 1135)
51	0.005206	(1659)
255	0.005206	(41, 1327)
262	0.005206	(1435, 39)
74	0.005229	(2051)
79	0.005229	(2135)
81	0.005240	(2184)
475	0.005252	(39, 48, 16217)
234	0.005252	(11, 48)
452	0.005252	(812, 48)
197	0.005263	(80)
505	0.005263	(39, 48, 286)
3	0.005274	(1020)
291	0.005274	(18, 48)
488	0.005286	(32, 39, 225)
144	0.005286	(432)
275	0.005297	(41, 16010)
546	0.005297	(39, 48, 548)
66	0.005297	(1867)
232	0.005297	(48, 107)
444	0.005297	(703, 48)
576	0.005297	(39, 48, 89, 38)
30	0.005297	(136)
52	0.005308	(1677)
565	0.005320	(41, 39, 225, 48)
268	0.005320	(48, 1578)
251	0.005331	(39, 12929)
330	0.005331	(32, 271)
407	0.005331	(39, 766)
292	0.005342	(39, 1814)
555	0.005342	(824, 39, 48)
418	0.005342	(39, 976)
413	0.005342	(39, 812)
511	0.005354	(32, 38, 36)
233	0.005354	(11, 39)
486	0.005354	(41, 39, 2238)

	support	itemsets
37	0.005354	(1479)
266	0.005365	(156, 48)
6	0.005365	(10444)
507	0.005365	(39, 48, 301)
271	0.005376	(39, 1600)
25	0.005399	(12959)
225	0.005399	(41, 101)
537	0.005399	(41, 39, 475)
494	0.005410	(41, 39, 237)
59	0.005433	(178)
38	0.005433	(14933)
137	0.005433	(408)
76	0.005456	(2080)
110	0.005456	(2879)
402	0.005456	(39, 664)
91	0.005456	(2399)
140	0.005456	(418)
199	0.005467	(808)
70	0.005467	(1987)
499	0.005490	(39, 271, 270)
564	0.005490	(41, 38, 170, 48)
446	0.005501	(76, 48)
257	0.005501	(1344, 39)
567	0.005513	(41, 39, 48, 310)
88	0.005535	(2329)
533	0.005535	(39, 65, 38)
447	0.005547	(766, 48)
441	0.005558	(48, 649)
122	0.005558	(345)
520	0.005581	(32, 65, 48)
427	0.005581	(48, 45)
393	0.005581	(549, 39)
482	0.005581	(41, 170, 48)
412	0.005592	(39, 806)
287	0.005592	(48, 175)
343	0.005592	(48, 31)
195	0.005603	(793)
348	0.005603	(32, 36)
206	0.005603	(856)

	support	itemsets
80	0.005660	(2168)
423	0.005660	(41, 79)
380	0.005671	(39, 405)
243	0.005683	(41, 1146)
159	0.005683	(52)
308	0.005705	(39, 229)
196	0.005705	(798)
433	0.005717	(544, 48)
358	0.005739	(39, 334)
56	0.005751	(1714)
471	0.005751	(39, 1393, 48)
143	0.005762	(426)
377	0.005762	(790, 38)
328	0.005762	(41, 270)
85	0.005762	(2284)
467	0.005773	(39, 48, 123)
161	0.005773	(53)
496	0.005785	(39, 249, 48)
560	0.005796	(41, 39, 38, 110)
29	0.005807	(1355)
310	0.005807	(39, 23)
493	0.005807	(39, 237, 38)
227	0.005807	(39, 103)
517	0.005807	(32, 39, 65)
435	0.005819	(549, 48)
429	0.005819	(48, 479)
170	0.005819	(571)
436	0.005830	(48, 570)
375	0.005830	(56, 38)
548	0.005842	(592, 48, 39)
464	0.005842	(41, 39, 110)
134	0.005842	(398)
522	0.005853	(39, 48, 338)
216	0.005853	(94)
278	0.005864	(161, 48)
213	0.005887	(910)
205	0.005887	(855)
550	0.005898	(604, 39, 48)
445	0.005898	(48, 740)

	support	itemsets
156	0.005910	(4883)
578	0.005921	(41, 39, 48, 89)
194	0.005932	(790)
100	0.005932	(261)
231	0.005944	(48, 10515)
391	0.005944	(544, 39)
53	0.005944	(1693)
78	0.005955	(2118)
112	0.005955	(297)
177	0.005955	(623)
419	0.005955	(48, 405)
185	0.005966	(704)
530	0.005966	(39, 38, 371)
230	0.005989	(39, 10515)
474	0.006000	(39, 16010, 48)
75	0.006000	(208)
406	0.006000	(39, 76)
400	0.006023	(39, 649)
370	0.006034	(39, 371)
476	0.006034	(32, 170, 38)
518	0.006034	(32, 39, 89)
324	0.006057	(264, 48)
168	0.006068	(56)
24	0.006068	(12946)
489	0.006068	(225, 39, 38)
34	0.006080	(1417)
492	0.006091	(41, 225, 48)
342	0.006102	(39, 31)
290	0.006125	(18, 39)
113	0.006125	(30)
472	0.006125	(39, 48, 14098)
456	0.006125	(956, 48)
570	0.006125	(41, 32, 48, 38)
281	0.006125	(32, 170)
219	0.006125	(961)
344	0.006136	(32, 310)
241	0.006136	(48, 1135)
176	0.006159	(62)
215	0.006159	(916)

	support	itemsets
214	0.006159	(913)
379	0.006170	(39, 389)
309	0.006182	(48, 229)
40	0.006193	(155)
557	0.006193	(39, 9, 48)
577	0.006204	(41, 39, 65, 48)
510	0.006204	(41, 48, 310)
71	0.006227	(2)
5	0.006250	(1043)
573	0.006273	(41, 39, 38, 36)
245	0.006284	(39, 117)
484	0.006295	(39, 48, 185)
529	0.006318	(37, 48, 38)
395	0.006329	(39, 570)
384	0.006341	(39, 45)
316	0.006397	(39, 242)
368	0.006409	(48, 37)
311	0.006443	(32, 237)
21	0.006454	(1239)
90	0.006454	(2383)
420	0.006465	(41, 475)
208	0.006465	(878)
506	0.006465	(2958, 39, 48)
301	0.006465	(41, 2238)
93	0.006477	(2437)
67	0.006488	(1872)
526	0.006488	(41, 39, 36)
521	0.006499	(32, 48, 89)
171	0.006511	(581)
209	0.006533	(885)
356	0.006533	(39, 3270)
267	0.006533	(39, 1578)
536	0.006556	(48, 89, 38)
386	0.006579	(39, 479)
504	0.006590	(48, 286, 38)
265	0.006601	(39, 156)
431	0.006601	(522, 48)
222	0.006613	(39, 1004)
450	0.006613	(783, 48)

	support	itemsets
332	0.006624	(41, 271)
10	0.006624	(1067)
128	0.006624	(374)
102	0.006624	(267)
270	0.006636	(48, 15832)
468	0.006670	(39, 48, 12925)
145	0.006681	(4336)
534	0.006681	(39, 38, 89)
57	0.006692	(1715)
297	0.006704	(48, 19)
217	0.006704	(947)
337	0.006704	(48, 286)
357	0.006715	(3270, 48)
296	0.006715	(39, 19)
204	0.006749	(846)
82	0.006772	(2199)
39	0.006772	(150)
314	0.006772	(41, 237)
323	0.006794	(264, 39)
16	0.006794	(1144)
141	0.006806	(420)
417	0.006806	(39, 956)
60	0.006806	(1783)
246	0.006817	(48, 117)
181	0.006817	(675)
253	0.006828	(48, 13041)
211	0.006874	(8978)
559	0.006874	(41, 48, 89)
142	0.006874	(425)
147	0.006908	(441)
549	0.006908	(60, 39, 48)
553	0.006919	(39, 78, 48)
77	0.006930	(209)
42	0.006964	(15618)
223	0.006964	(48, 1004)
69	0.006964	(1986)
562	0.006976	(41, 39, 170, 38)
544	0.006998	(39, 48, 49)
7	0.007010	(10446)

	support	itemsets
540	0.007021	(41, 39, 89)
95	0.007044	(251)
277	0.007055	(39, 161)
568	0.007055	(41, 39, 32, 38)
466	0.007067	(39, 48, 1146)
480	0.007078	(41, 39, 170)
289	0.007078	(179, 48)
73	0.007089	(2046)
508	0.007089	(41, 39, 310)
200	0.007101	(809)
437	0.007123	(48, 589)
321	0.007123	(39, 258)
286	0.007157	(39, 175)
207	0.007169	(865)
545	0.007169	(39, 48, 533)
107	0.007180	(279)
403	0.007203	(39, 677)
409	0.007259	(783, 39)
228	0.007293	(38, 105)
322	0.007293	(48, 258)
121	0.007293	(340)
376	0.007293	(65, 38)
280	0.007339	(48, 16217)
389	0.007350	(39, 522)
192	0.007350	(789)
106	0.007373	(272)
273	0.007384	(16010, 16011)
288	0.007407	(39, 179)
124	0.007418	(3616)
338	0.007430	(2958, 39)
303	0.007430	(32, 225)
8	0.007452	(105)
374	0.007452	(55, 38)
434	0.007464	(48, 548)
341	0.007464	(48, 301)
19	0.007486	(1198)
65	0.007486	(186)
392	0.007520	(39, 548)
558	0.007520	(41, 65, 48)

	support	itemsets
462	0.007554	(41, 38, 110)
443	0.007554	(48, 677)
15	0.007566	(1135)
298	0.007588	(39, 201)
48	0.007588	(16011)
259	0.007588	(1393, 48)
524	0.007611	(41, 38, 36)
299	0.007645	(201, 48)
133	0.007645	(396)
318	0.007656	(249, 48)
129	0.007679	(379)
238	0.007679	(41, 110)
326	0.007713	(271, 270)
304	0.007724	(225, 38)
312	0.007747	(237, 38)
98	0.007758	(259)
528	0.007758	(39, 38, 37)
440	0.007792	(604, 48)
63	0.007804	(1814)
252	0.007804	(39, 13041)
148	0.007827	(449)
99	0.007838	(260)
136	0.007861	(407)
150	0.007883	(464)
261	0.007917	(48, 14098)
554	0.007929	(39, 79, 48)
364	0.007940	(41, 36)
103	0.007951	(269)
453	0.007951	(824, 48)
414	0.007974	(824, 39)
167	0.007985	(55)
183	0.007985	(681)
367	0.008019	(39, 37)
396	0.008031	(39, 589)
12	0.008065	(11)
0	0.008076	(10)
355	0.008087	(32, 89)
340	0.008110	(39, 301)
163	0.008144	(535)

	support	itemsets
269	0.008144	(39, 15832)
346	0.008155	(41, 310)
155	0.008167	(488)
470	0.008167	(39, 48, 1327)
397	0.008201	(592, 39)
247	0.008235	(39, 123)
490	0.008235	(41, 39, 225)
503	0.008258	(39, 38, 286)
279	0.008258	(39, 16217)
276	0.008280	(16010, 48)
424	0.008303	(41, 89)
500	0.008314	(39, 48, 270)
198	0.008405	(806)
23	0.008416	(12929)
360	0.008416	(39, 338)
473	0.008416	(39, 147, 48)
438	0.008428	(592, 48)
336	0.008507	(39, 286)
180	0.008530	(664)
189	0.008530	(772)
317	0.008530	(39, 249)
28	0.008541	(1344)
87	0.008598	(23)
405	0.008609	(39, 740)
378	0.008666	(38, 89)
11	0.008677	(107)
35	0.008677	(1435)
369	0.008700	(38, 371)
388	0.008711	(39, 49)
354	0.008779	(32, 65)
119	0.008779	(334)
408	0.008779	(39, 78)
399	0.008791	(604, 39)
339	0.008836	(2958, 48)
361	0.008836	(48, 338)
542	0.008847	(39, 438, 48)
541	0.008859	(39, 413, 48)
127	0.008870	(371)
487	0.008938	(39, 2238, 48)

	support	itemsets
258	0.008949	(39, 1393)
455	0.008961	(9, 48)
169	0.008983	(570)
539	0.008983	(41, 39, 65)
478	0.009006	(41, 170, 38)
188	0.009040	(766)
248	0.009040	(48, 123)
260	0.009086	(39, 14098)
46	0.009086	(1600)
513	0.009131	(41, 32, 38)
284	0.009131	(41, 170)
184	0.009142	(703)
244	0.009199	(48, 1146)
4	0.009210	(103)
497	0.009222	(39, 48, 255)
439	0.009244	(60, 48)
295	0.009256	(48, 185)
220	0.009267	(976)
250	0.009267	(48, 12925)
86	0.009301	(229)
201	0.009312	(812)
449	0.009346	(78, 48)
502	0.009380	(39, 271, 48)
294	0.009392	(39, 185)
274	0.009403	(39, 16010)
41	0.009414	(156)
131	0.009426	(389)
416	0.009437	(39, 9)
166	0.009562	(549)
430	0.009562	(49, 48)
164	0.009653	(544)
178	0.009721	(649)
62	0.009755	(18)
432	0.009766	(48, 533)
187	0.009857	(76)
43	0.009914	(1578)
306	0.009948	(41, 225)
135	0.009970	(405)
9	0.010004	(10515)

	support	itemsets
451	0.010129	(79, 48)
101	0.010152	(264)
111	0.010254	(2958)
92	0.010333	(242)
218	0.010333	(956)
149	0.010333	(45)
115	0.010435	(31)
390	0.010458	(39, 533)
153	0.010503	(479)
249	0.010640	(39, 12925)
458	0.010730	(39, 101, 48)
118	0.010776	(3270)
302	0.010832	(2238, 48)
329	0.010855	(48, 270)
191	0.010946	(783)
256	0.010980	(1327, 48)
58	0.011002	(175)
160	0.011048	(522)
398	0.011150	(39, 60)
242	0.011150	(39, 1146)
97	0.011195	(258)
422	0.011286	(41, 65)
61	0.011320	(179)
68	0.011399	(19)
49	0.011456	(161)
426	0.011626	(438, 48)
18	0.011638	(117)
561	0.011694	(39, 48, 110, 38)
264	0.011751	(147, 48)
465	0.011762	(39, 48, 110)
366	0.011865	(38, 37)
26	0.011921	(13041)
319	0.011989	(39, 255)
320	0.011989	(48, 255)
190	0.012023	(78)
126	0.012182	(37)
574	0.012250	(39, 48, 36, 38)
333	0.012364	(271, 48)
543	0.012386	(39, 48, 475)

	support	itemsets
1	0.012500	(1004)
182	0.012590	(677)
411	0.012602	(39, 79)
335	0.012659	(38, 286)
527	0.012659	(39, 48, 36)
172	0.012693	(589)
157	0.012704	(49)
382	0.012817	(39, 413)
72	0.012851	(201)
425	0.012874	(413, 48)
263	0.012897	(39, 147)
165	0.012897	(548)
44	0.012965	(15832)
254	0.013112	(39, 1327)
94	0.013158	(249)
31	0.013169	(1393)
50	0.013226	(16217)
186	0.013396	(740)
109	0.013418	(286)
563	0.013532	(39, 38, 48, 170)
327	0.013543	(39, 270)
114	0.013657	(301)
481	0.013679	(39, 48, 170)
175	0.013713	(604)
202	0.013725	(824)
173	0.013918	(592)
569	0.014020	(32, 39, 48, 38)
495	0.014110	(39, 237, 48)
383	0.014292	(39, 438)
120	0.014451	(338)
300	0.014598	(39, 2238)
33	0.014643	(14098)
20	0.014768	(123)
226	0.014870	(101, 48)
47	0.014927	(16010)
509	0.015279	(39, 48, 310)
525	0.015426	(48, 36, 38)
463	0.015437	(48, 110, 38)
212	0.015562	(9)

	support	itemsets
64	0.015608	(185)
239	0.015653	(48, 110)
491	0.015880	(225, 39, 48)
224	0.015880	(39, 101)
365	0.016061	(48, 36)
17	0.016175	(1146)
428	0.016197	(48, 475)
331	0.016266	(39, 271)
22	0.016640	(12925)
96	0.016719	(255)
162	0.016867	(533)
174	0.016889	(60)
385	0.017014	(39, 475)
479	0.017445	(38, 170, 48)
285	0.017661	(170, 48)
193	0.018148	(79)
571	0.018670	(41, 39, 32, 48)
514	0.018670	(32, 48, 38)
315	0.019079	(237, 48)
347	0.019192	(48, 310)
83	0.019453	(2238)
104	0.019668	(270)
307	0.019691	(225, 48)
461	0.019736	(39, 38, 110)
237	0.019952	(39, 110)
36	0.020179	(147)
27	0.020258	(1327)
551	0.020383	(65, 39, 48)
512	0.020871	(32, 39, 38)
345	0.021007	(39, 310)
146	0.021132	(438)
139	0.021324	(413)
313	0.021880	(39, 237)
523	0.022062	(39, 38, 36)
575	0.022583	(41, 39, 48, 38)
477	0.022901	(39, 170, 38)
363	0.023105	(39, 36)
283	0.023355	(39, 170)
519	0.023400	(41, 32, 48)

	support	itemsets
105	0.023752	(271)
556	0.024103	(39, 48, 89)
152	0.024580	(475)
2	0.025374	(101)
305	0.026667	(225, 39)
515	0.026758	(41, 39, 32)
535	0.026928	(41, 48, 38)
442	0.028686	(65, 48)
116	0.029423	(310)
236	0.030909	(38, 110)
415	0.031181	(39, 89)
401	0.031612	(39, 65)
362	0.031646	(38, 36)
13	0.031692	(110)
454	0.031737	(48, 89)
349	0.032134	(32, 38)
123	0.033302	(36)
282	0.034380	(170, 38)
89	0.034391	(237)
531	0.034607	(41, 39, 38)
54	0.035151	(170)
351	0.036251	(41, 32)
84	0.036943	(225)
210	0.043522	(89)
372	0.044203	(41, 38)
179	0.050725	(65)
516	0.061274	(32, 39, 48)
532	0.069213	(39, 48, 38)
538	0.083551	(41, 39, 48)
373	0.090107	(48, 38)
353	0.091128	(32, 48)
350	0.095903	(32, 39)
421	0.102289	(41, 48)
371	0.117341	(39, 38)
381	0.129466	(41, 39)
138	0.169517	(41)
117	0.172036	(32)
130	0.176902	(38)
387	0.330551	(39, 48)

	support	itemsets
154	0.477927	(48)
133	0.534704	(39)

Confiança

In [8]:

```
from mlxtend.frequent_patterns import association_rules

rules = association_rules(frequent_itemsets, metric="confidence", min_threshold=0.9)
```

In [9]:

```
print("Quantidade de regras de associação com confiança minima de 0.9 =", rules.shape[0
display(rules[["antecedents", "consequents", "confidence"]].sort_values("confidence"))
```

```
Quantidade de regras de associação com confiança minima de 0.9 = 37
```

	antecedents	consequents	confidence
8	(55)	(38)	0.933239
4	(286)	(38)	0.943364
5	(36)	(38)	0.950272
23	(39, 36)	(38)	0.954836
22	(32, 36)	(38)	0.955466
24	(41, 36)	(38)	0.958571
25	(48, 36)	(38)	0.960452
9	(56)	(38)	0.960748
35	(41, 39, 36)	(38)	0.966783
26	(39, 37)	(38)	0.967468
36	(39, 48, 36)	(38)	0.967742
20	(39, 286)	(38)	0.970667
10	(790)	(38)	0.971319
2	(16011)	(16010)	0.973094
6	(37)	(38)	0.973929
1	(110)	(38)	0.975304
3	(170)	(38)	0.978057
0	(105)	(38)	0.978691
17	(39, 170)	(38)	0.980573
7	(371)	(38)	0.980818
21	(48, 286)	(38)	0.983080
33	(41, 170, 48)	(38)	0.983740
14	(41, 110)	(38)	0.983752
16	(32, 170)	(38)	0.985185
31	(41, 39, 170)	(38)	0.985577
27	(48, 37)	(38)	0.985841
15	(48, 110)	(38)	0.986232
18	(41, 170)	(38)	0.986335
12	(32, 110)	(38)	0.986637
11	(39, 105)	(38)	0.986813
34	(39, 48, 286)	(38)	0.987069
19	(170, 48)	(38)	0.987797
28	(39, 371)	(38)	0.988722
13	(39, 110)	(38)	0.989198

	antecedents	consequents	confidence
32	(39, 48, 170)	(38)	0.989221
29	(41, 39, 110)	(38)	0.992233
30	(39, 48, 110)	(38)	0.994214

In [ ]:

```
1
```