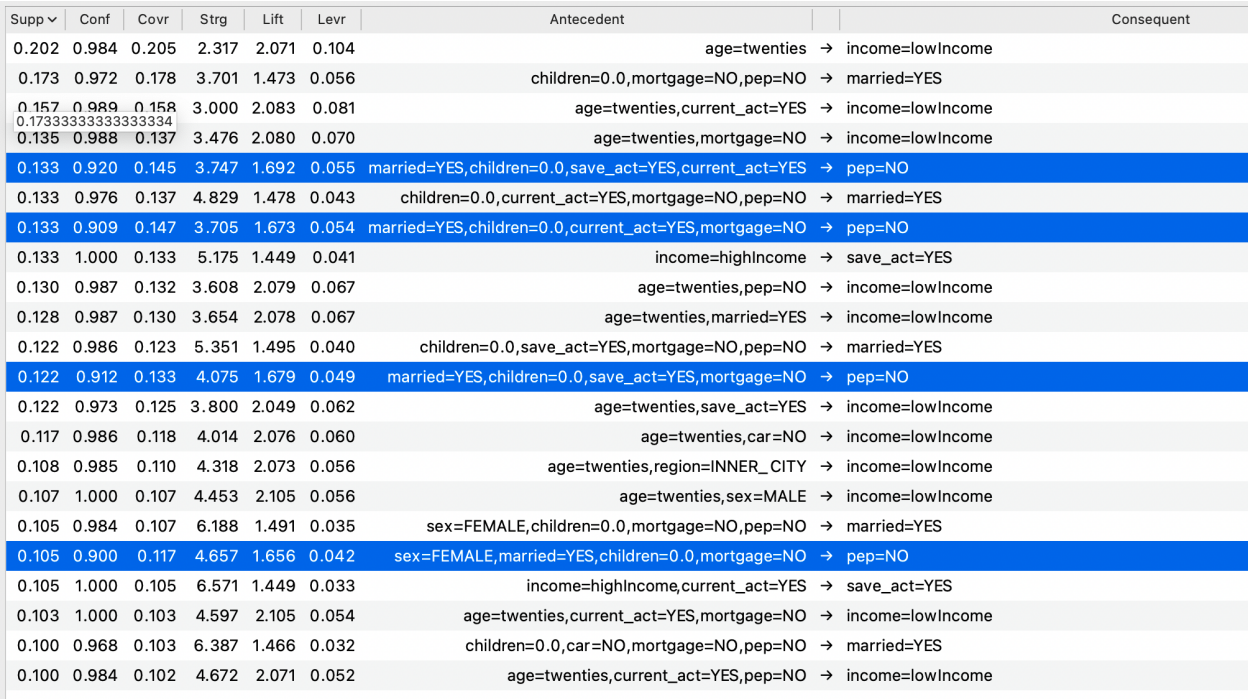
Homework 3: Association Rule Mining

When trying to find rules with support of 0.20 and Confidence of 0.9, we get:

Supp	Conf	Covr	Strg	Lift	Levr	Antecedent		
0.202	0.984	0.205	2.317	2.071	0.104	age=twenties	→	income=lowIncome

The lift is pretty good at 2.3. However, the result is not that interesting. It is obvious that in your twenties you may not earn a lot!

When trying to find rules with support of 0.10 and Confidence of 0.9, we get:



Supp	Conf	Covr	Strg	Lift	Levr	Antecedent		Consequent
0.202	0.984	0.205	2.317	2.071	0.104	age=twenties	→	income=lowIncome
0.173	0.972	0.178	3.701	1.473	0.056	children=0.0,mortgage=NO,pep=NO	→	married=YES
0.157	0.989	0.158	3.000	2.083	0.081	age=twenties,current_act=YES	→	income=lowIncome
0.135	0.988	0.137	3.476	2.080	0.070	age=twenties,mortgage=NO	→	income=lowIncome
0.133	0.920	0.145	3.747	1.692	0.055	married=YES,children=0.0,save_act=YES,current_act=YES	→	pep=NO
0.133	0.976	0.137	4.829	1.478	0.043	children=0.0,current_act=YES,mortgage=NO,pep=NO	→	married=YES
0.133	0.909	0.147	3.705	1.673	0.054	married=YES,children=0.0,current_act=YES,mortgage=NO	→	pep=NO
0.133	1.000	0.133	5.175	1.449	0.041	income=highIncome	→	save_act=YES
0.130	0.987	0.132	3.608	2.079	0.067	age=twenties,pep=NO	→	income=lowIncome
0.128	0.987	0.130	3.654	2.078	0.067	age=twenties,married=YES	→	income=lowIncome
0.122	0.986	0.123	5.351	1.495	0.040	children=0.0,save_act=YES,mortgage=NO,pep=NO	→	married=YES
0.122	0.912	0.133	4.075	1.679	0.049	married=YES,children=0.0,save_act=YES,mortgage=NO	→	pep=NO
0.122	0.973	0.125	3.800	2.049	0.062	age=twenties,save_act=YES	→	income=lowIncome
0.117	0.986	0.118	4.014	2.076	0.060	age=twenties,car=NO	→	income=lowIncome
0.108	0.985	0.110	4.318	2.073	0.056	age=twenties,region=INNER_CITY	→	income=lowIncome
0.107	1.000	0.107	4.453	2.105	0.056	age=twenties,sex=MALE	→	income=lowIncome
0.105	0.984	0.107	6.188	1.491	0.035	sex=FEMALE,children=0.0,mortgage=NO,pep=NO	→	married=YES
0.105	0.900	0.117	4.657	1.656	0.042	sex=FEMALE,married=YES,children=0.0,mortgage=NO	→	pep=NO
0.105	1.000	0.105	6.571	1.449	0.033	income=highIncome,current_act=YES	→	save_act=YES
0.103	1.000	0.103	4.597	2.105	0.054	age=twenties,current_act=YES,mortgage=NO	→	income=lowIncome
0.100	0.968	0.103	6.387	1.466	0.032	children=0.0,car=NO,mortgage=NO,pep=NO	→	married=YES
0.100	0.984	0.102	4.672	2.071	0.052	age=twenties,current_act=YES,pep=NO	→	income=lowIncome

For the highlighted rows, the lift is pretty good at 1.69, 1.67, 1.67, and 1.65. The result is interesting as we know these people may not obtain PEP. Thus, we won't waste our time and money on marketing for these people.

When trying to find rules with support of 0.30 and Confidence of 0.7, we get:

Supp	Conf	Covr	Strg	Lift	Levr	Antecedent		Consequent
0.532	0.701	0.758	0.910	1.016	0.008	current_act=YES	→	save_act=YES
0.532	0.771	0.690	1.099	1.016	0.008	save_act=YES	→	current_act=YES
0.502	0.770	0.652	1.164	1.015	0.007	mortgage=NO	→	current_act=YES
0.488	0.740	0.660	1.149	0.976	-0.012	married=YES	→	current_act=YES
0.407	0.748	0.543	1.396	0.987	-0.005	pep=NO	→	current_act=YES
0.403	0.742	0.543	1.215	1.125	0.045	pep=NO	→	married=YES
0.392	0.721	0.543	1.270	1.045	0.017	pep=NO	→	save_act=YES
0.392	0.773	0.507	1.497	1.019	0.007	car=NO	→	current_act=YES
0.383	0.767	0.500	1.517	1.011	0.004	sex=FEMALE	→	current_act=YES
0.375	0.750	0.500	1.517	0.989	-0.004	sex=MALE	→	current_act=YES
0.367	0.743	0.493	1.537	0.980	-0.007	car=YES	→	current_act=YES
0.358	0.754	0.475	1.596	0.995	-0.002	income=lowIncome	→	current_act=YES
0.353	0.704	0.502	1.375	1.021	0.007	current_act=YES,mortgage=NO	→	save_act=YES
0.353	0.785	0.450	1.685	1.035	0.012	save_act=YES,mortgage=NO	→	current_act=YES
0.352	0.770	0.457	1.661	1.015	0.005	pep=YES	→	current_act=YES
0.348	0.706	0.493	1.399	1.023	0.008	car=YES	→	save_act=YES
0.343	0.703	0.488	1.413	1.019	0.006	married=YES,current_act=YES	→	save_act=YES
0.343	0.744	0.462	1.643	0.981	-0.007	married=YES,save_act=YES	→	current_act=YES
0.342	0.762	0.448	1.691	1.005	0.002	region=INNER_CITY	→	current_act=YES
0.332	0.762	0.435	1.743	1.005	0.002	married=YES,mortgage=NO	→	current_act=YES
0.332	0.757	0.438	1.730	0.998	-0.001	children=0.0	→	current_act=YES
0.307	0.705	0.435	1.586	1.022	0.007	married=YES,mortgage=NO	→	save_act=YES

There are some rules that are not interesting. Like, if someone obtained PEP, they would have a current account. It is obvious! However, some of the rules (the highlighted one) can be helpful for marketing purposes.

When trying to find rules with support of 0.08 and Confidence of 0.85, we get:

Supp	Conf	Covr	Strg	Lift	Levr	Antecedent		Consequent
0.105	0.863	0.122	3.753	1.890	0.049	children=1.0,save_act=YES,current_act=YES	→	pep=YES
0.095	0.877	0.108	4.215	1.920	0.046	married=YES,children=1.0,save_act=YES	→	pep=YES
0.093	0.862	0.108	4.215	1.887	0.044	married=YES,children=1.0,current_act=YES	→	pep=YES
0.082	0.925	0.088	5.170	2.025	0.041	income=midIncome,children=1.0	→	pep=YES
0.080	0.873	0.092	4.982	1.911	0.038	children=1.0,save_act=YES,mortgage=NO	→	pep=YES

These rules are interesting and can get used in marketing because the lift value is considerably high. We are 85% confident that people who satisfy the “Antecedent” requirement are the ones that will obtain the PEP. For example, if someone has a medium income and has one child, he/she probably will obtain PEP.

The interesting part about these rules is that people who have one child are probable to obtain pep. We can see all people who obtained pep had one child. Thus, we can guide the marketing campaign to focus on people who have a child.

When trying to find rules with support of 0.1 and Confidence of 0.90, we get:

Supp	Conf	Covr	Strg	Lift	Levr	Antecedent		Consequent
0.133	0.920	0.145	3.747	1.692	0.055	married=YES,children=0.0,save_act=YES,current_act=YES	→	pep=NO
0.133	0.909	0.147	3.705	1.673	0.054	married=YES,children=0.0,current_act=YES,mortgage=NO	→	pep=NO
0.122	0.912	0.133	4.075	1.679	0.049	married=YES,children=0.0,save_act=YES,mortgage=NO	→	pep=NO
0.105	0.900	0.117	4.657	1.656	0.042	sex=FEMALE,married=YES,children=0.0,mortgage=NO	→	pep=NO

These rules are interesting and can get used in marketing because the lift value is considerably high. We are 90% confident that people who satisfy the “Antecedent” requirement will not obtain the PEP. Thus, we won’t waste our time and money on marketing for these people. For example, if a female is married, has no children, and didn’t get a mortgage, she probably will NOT obtain PEP.

The interesting part about these rules is that people who have no child will NOT obtain pep.

A screenshot of the data flow for reference: