Pranav Kaza

Professor Diego Klabjan

IEMS 308: Data Science and Analytics

8 February 2018

Association Rule Analysis of Dillard's Retail Data

Executive Summary

Dillard's is a retail chain with many stores. Although management has built up an abundance of point-of-sales data over time, it has yet to conduct any analysis on the effectiveness of the flooring and shelving of its stock keeping units (SKUs). To increase their revenues, Dillard's management would like to know the best way to rearrange the placement of their SKUs across their entire retail chain.

Association rule analysis is a method that examines items that are commonly bought in the same shopper baskets. Association rules between SKUs are found by using simple probability concepts on market basket data. In this case, a subset of the data was chosen based on date of purchase.

After the data was cleaned and organized to find each unique basket and the SKUs in those baskets, the algorithm was used to find the most powerful association rules within the data. The results showed 100 SKUs that are good candidates for re-shelving across the Dillard's chain. However, it is important to note that these 100 SKUs were chosen for short-term strategy, and that more robust analysis should be done with more data to find better SKUs for long-term strategy.

Problem Statement

Dillard's retail chain has an abundance of point-of-sales data stored in a remote database. The task was to find the 100 SKUs that are the best candidates for moving to different positions, of which 20 will be chosen to move across the chain due to budgetary constraints.

Assumptions

1. The biggest assumption made is that the sample taken is representative of all the data in the database. Since the subset of data is taken based on date, the assumption is that the date chosen is an accurate representation of all dates in terms of items bought. For example, if the date picked was hypothetically Black Friday, this would be an inaccurate representation of any ordinary date in what is bought.
2. Another assumption is that all the SKUs are relatively similar in profit margin. This is a large assumption to make, but is necessary because the association rule and apriori algorithms fail to account for high profit margin items that may have low support values, but are still important to analyze.

Methodology

The first step is to explore the data and determine which columns map to which features in the “trnsact” table. By examining the database diagram and column/attribute descriptions provided, as well as mapping to metadata provided in other tables, most of the features can be discerned. It is unclear which column maps to “INTERID” and which column maps to “SEQ.” To determine this, a query was run in Python selecting 10,000 rows and duplicates were removed. Since SEQ is a primary key, including it in the column subset (along with the other primary keys which are known) for duplicate removal should result in no duplicate removal. Using this method, the column “c12” was found to correspond to “SEQ.” All other relevant columns used and their corresponding features from the “trnsact” table are included in Appendix A.

The next step is to select a subset of the “trnsact” table in the database to have an amount of data that can be efficiently analyzed. Since random sampling is computationally inefficient, the sample was picked by choosing an arbitrary date, in this case May 5, 2005. This is a valid method because it includes data from all stores and does not favor one more than the other, given the assumptions included above. In addition, May 5 is not a special day in terms of holidays, so it is a relatively unbiased date to pick. Lastly, querying by date will guarantee that all unique baskets include every item that was bought because all items in a basket must be bought on the same day.

After cleaning the data and removing returns to isolate all purchase data, the sample data is uploaded back into the database as table “all_data.” In addition, another table “transactions” is uploaded with all unique transactions and a unique integer primary key for each transaction. An inner join is used to once again select all data, but this time including the integer transaction primary key “unique_tran” to identify the market basket to which the transaction belongs to. This data is then loaded back into Python for analysis.

Evidently, separating by date is still very computationally expensive, so a random sample of 100 stores was chosen from the data and the data from these 100 stores was used and the data is reorganized to have each unique basket as a row and the SKUs as binary columns indicating whether they were in that basket. A total of 37,974 baskets and 36,966 SKUs were used in the analysis. When running the apriori algorithm, a very low minimum support level of .034% was used to produce enough unique SKU combinations for analysis. Lastly, the association rules algorithm is used and all unique SKUs are found. A minimum threshold of 140 is used for lift because the lift values are extremely inflated, as is apparent in Appendix C and will be discussed further in the analysis. The confidence table produced with the association rules algorithm can be found in Appendix D.

Analysis

The SKUs that were chosen for re-location are shown in Appendix B. Some of the SKU numbers did not correspond with qualitative descriptions found in the metadata, so these SKUs should be more closely examined to discover specifically what items they correspond to. Although the 100 SKUs found may be good choices for initial re-location, it is important to understand the shortcomings of the results and methodology used so that a more robust analysis can be used for long-term strategy.

One of the major shortcomings of the process was the inability to process a larger amount of data. This shortcoming is evident in the enormous lift values found in Appendix C. Since lift measures the dependence of the items purchased on each other, the numbers are likely this large because there is not enough data for those SKUs indicating how independent they are of each other. In addition, the confidence table in Appendix D also indicates inconsistencies in the confidence levels. Lastly, while there are over 1,000,000 total SKUs, only 36,966 of them were included in the analysis. Although these are likely more commonly bought items because of the random sampling methods used, this method likely eliminates high profit margin items that are bought less, but still may be good candidates for re-location. Therefore, an additional analysis may warrant use of a minimum item support criteria in generating the frequent item sets, instead of a static minimum support value.

Conclusion

Based on the association rules found, the 100 SKUs in Appendix B are the best candidates for re-location across the Dillard's retail chain.

Next Steps

In continuing a closer analysis of this problem, it is important to find a way to use more powerful processing and memory so that much larger amounts of data can be analyzed. In addition, using a minimum item support criteria for each individual SKU would prevent elimination of important high profit margin items that have most likely been excluded in this analysis.

APPENDIX A: Relevant Feature Mapping to “trnsact” Table Columns

Feature	Column in “trnsact” Table
SKU	c1
STORE	c2
REGISTER	c3
TRANNUM	c4
STYPE	c7
SEQUENCE	c12

APPENDIX B: Chosen SKUs

SKU	BRAND	STYLE	COLOR
5468528	D. G. EN	681 59551-	PINK
152307	MAIN KNI	10 Y25UR1	WHITE
768635	WESTPOIN	8 F30G61	SNOW
1100112	HENSON-K	R125	WHITE
5528349	LANCOME	2410	01-BLACK
1357342	ROYAL AL	NTRYROOLDCOU	15210698
1686435	WESTPOIN	9 F30G03	MULTI
1688364	SUMMER S	282 517SD6	YELLOW/RED
1756435	WESTPOIN	0 F30G04	NAVY/RED
1949753	LANCASTE	000 FP3009	EDP SPRAY
2419753	LANCASTE	250 FP0479	WATCH GWP
2088302	WESTPOIN	G131T0	WHITE
2927335	NOBLE EX	OTTON MICROC	WHEAT
3301573	GREAT AM	5572	WHITE
3680057	CLINIQUE	64MT	04 LT BEIGE
4142521	NOBLE EX	15 F11G75	CAMEO
4356618	CABERNET	12D028	WHITE CABE
5079905	MILCO IN	75415	BEIGE
5268597	CLINIQUE	6F6090	YELLOW/GREEN
9168271			
5453386	CLINIQUE	60C8	M LOTION
5600671	CABERNET	R111	WHITE
6308344	CLINIQUE	653901	M XTRA REFIL
6441416	HOT TOPI	22 ERX400	EDP SPRAY
6759905			
8579778			
9569904			
9686202			
9702306			
133101	NORITAKE	TCOBALPLCRES	4170-05M
208362	CLINIQUE	660601	BDY SMOOTHER
599839	PUIG USA	52 650120	ELEGANCE SET
726718	CLINIQUE	60MK01	TAKE DAY OFF
739830	LANCOME	4203	16-4PARTY
1138996	LENOX CH	SPANTRBUTLER	L6084875
748635	WESTPOIN	8 F30G61	SNOW
776350	LANCOME	8864	01-BLACK
803921	CLINIQUE	6.63E+04	RINSE OFF
944478	DESIGNER	5 7002-9	EDP SPRAY
1117865	MIKASA	ENT PARCHM	L3438-705
2176435	WESTPOIN	7 F30L01	RED

2316904	FRANCISC	ROSE DESERT	5260010232
2959603	LENOX CH	IREWHISOLITA	L6224158
3029412	MIKASA	NCOUNTITALIA	DD900-705
8159294			
1230703	NIKKO CE	DPEARLPBEADE	12346155
1346904	FRANCISC	APPLE	5270010233
1519759	CALVIN K	5 100120	EDT SPRAY
1539616	CALVIN K	4 110112	GWP
1588107	CLINIQUE	648G	04BISQUE
1832285	CLINIQUE	634C	01BLACK
2178731	CLINIQUE	6E9R	HAPPY TO BE
3998011	CLINIQUE	6126	CLARIFY
2726578	CLINIQUE	6EM6	DDM GEL
3161221	CLINIQUE	68A0	01BLACK
4992993	CLINIQUE	6CLW	LASH BUILD
3559555	CLINIQUE	667H	7DAY SCRUB
3589483	CLINIQUE	61EM	01LIPBLUSH
3690654	CLINIQUE	68MG	DDML TUBE
3898011	CLINIQUE	6101	MILD SOAP
3968011	CLINIQUE	6120	CLARIFY
3968356	CLINIQUE	6270	ANTI ACNE
3989735	LANCOME	3280	00-S'05
8129879			
3999735	LANCOME	3047	00-S'05
4138348	CLINIQUE	62A3	31JETBLACK
4370921	NORITAKE	OODPLACRESTW	4166-05M
4462521	NOBLE EX	11 F11G75	CAMEO
4512521	NOBLE EX	12 F11G75	CAMEO
4722472	WESTPOIN	8 F30G61	KHAKI
4772472	WESTPOIN	8 F30G61	KHAKI
4980033	LANCOME	2077	EMPTY
5108107	CLINIQUE	647J	02BEIGE
5172064	HUE	U6113	PALE BEIGE
5189905	MILCO IN	75416	BEIGE
5347422	MIKASA	NCOUNTITALIA	DD900-219
5369905	MILCO IN	75415	BEIGE
5749904	MILCO IN	75415	WHITE
6060702	NIKKO CE	UMPEARLPLATIN	12340155
6280019	LENOX CH	L ETERNA	L140190600
6318344	CLINIQUE	6034	MILD SOAP
6349904	MILCO IN	75415	WHITE
6949904			
7029904			
7202503			

7710113				
8239752				
8503540				
8718362				
8889414				
9073382				
9277426				
9486925				
9552306				
5848345	LANCOME		6656	EDP SPR 3.4
6697047				
8899931				
6899800				
3524026	CLINIQUE	68LE		DDML PUMP
3978011	CLINIQUE		6121	CLARIFY

APPENDIX C: Lift-based Association Rule Table

antecedants	consequents	support	confidence	lift
frozenset({5468528})	frozenset({133101})	0.001395692	0.264150943	501.5433962
frozenset({133101})	frozenset({5468528})	0.000526676	0.7	501.5433962
frozenset({5468528})	frozenset({152307})	0.001395692	0.245283019	716.490566
frozenset({152307})	frozenset({5468528})	0.000342339	1	716.490566
frozenset({5468528})	frozenset({208362})	0.001395692	0.264150943	334.3622642
frozenset({208362})	frozenset({5468528})	0.000790014	0.466666667	334.3622642
frozenset({5468528})	frozenset({599839})	0.001395692	0.245283019	490.2303873
frozenset({599839})	frozenset({5468528})	0.000500342	0.684210526	490.2303873
frozenset({5468528})	frozenset({726718})	0.001395692	0.264150943	179.1226415
frozenset({726718})	frozenset({5468528})	0.001474693	0.25	179.1226415
frozenset({5468528})	frozenset({739830})	0.001395692	0.245283019	310.4792453
frozenset({739830})	frozenset({5468528})	0.000790014	0.433333333	310.4792453
frozenset({748635})	frozenset({1138996})	0.000684679	0.5	249.8289474
frozenset({1138996})	frozenset({748635})	0.002001369	0.171052632	249.8289474
frozenset({5468528})	frozenset({748635})	0.001395692	0.301886792	440.9172714
frozenset({748635})	frozenset({5468528})	0.000684679	0.615384615	440.9172714
frozenset({5468528})	frozenset({768635})	0.001395692	0.245283019	716.490566
frozenset({768635})	frozenset({5468528})	0.000342339	1	716.490566
frozenset({5468528})	frozenset({776350})	0.001395692	0.264150943	233.2759982
frozenset({776350})	frozenset({5468528})	0.001132354	0.325581395	233.2759982
frozenset({5468528})	frozenset({803921})	0.001395692	0.264150943	147.5127636
frozenset({803921})	frozenset({5468528})	0.001790699	0.205882353	147.5127636
frozenset({5468528})	frozenset({944478})	0.001395692	0.245283019	179.1226415
frozenset({944478})	frozenset({5468528})	0.001369358	0.25	179.1226415
frozenset({1100112})	frozenset({5468528})	0.000342339	1.076923077	771.605225
frozenset({5468528})	frozenset({1100112})	0.001395692	0.264150943	771.605225
frozenset({1100112})	frozenset({5528349})	0.000342339	1	214.5423729
frozenset({5528349})	frozenset({1100112})	0.004661084	0.073446328	214.5423729
frozenset({5468528})	frozenset({1117865})	0.001395692	0.264150943	401.234717
frozenset({1117865})	frozenset({5468528})	0.000658345	0.56	401.234717
frozenset({2176435})	frozenset({1138996})	0.000526676	0.65	324.7776316
frozenset({1138996})	frozenset({2176435})	0.002001369	0.171052632	324.7776316
frozenset({2316904})	frozenset({1138996})	0.00076368	0.448275862	223.9845735
frozenset({1138996})	frozenset({2316904})	0.002001369	0.171052632	223.9845735
frozenset({2959603})	frozenset({1138996})	0.000579344	0.590909091	295.2523923
frozenset({1138996})	frozenset({2959603})	0.002001369	0.171052632	295.2523923
frozenset({3029412})	frozenset({1138996})	0.000711013	0.481481481	240.5760234
frozenset({1138996})	frozenset({3029412})	0.002001369	0.171052632	240.5760234
frozenset({5079905})	frozenset({1138996})	0.000500342	0.684210526	341.8711911
frozenset({1138996})	frozenset({5079905})	0.002001369	0.171052632	341.8711911

frozenset({5468528})	frozenset({1138996})	0.001395692	0.358490566	179.1226415
frozenset({1138996})	frozenset({5468528})	0.002001369	0.25	179.1226415
frozenset({1138996})	frozenset({8159294})	0.002001369	0.197368421	258.4437387
frozenset({8159294})	frozenset({1138996})	0.00076368	0.517241379	258.4437387
frozenset({5468528})	frozenset({1230703})	0.001395692	0.283018868	358.245283
frozenset({1230703})	frozenset({5468528})	0.000790014	0.5	358.245283
frozenset({1346904})	frozenset({5468528})	0.001132354	0.348837209	249.9385695
frozenset({5468528})	frozenset({1346904})	0.001395692	0.283018868	249.9385695
frozenset({5468528})	frozenset({1357342})	0.001395692	0.245283019	582.1485849
frozenset({1357342})	frozenset({5468528})	0.000421341	0.8125	582.1485849
frozenset({1539616})	frozenset({1519759})	0.001158688	0.295454545	224.3918182
frozenset({1519759})	frozenset({1539616})	0.00131669	0.26	224.3918182
frozenset({5468528})	frozenset({1519759})	0.001395692	0.245283019	186.2875472
frozenset({1519759})	frozenset({5468528})	0.00131669	0.26	186.2875472
frozenset({5468528})	frozenset({1588107})	0.001395692	0.283018868	565.6504469
frozenset({1588107})	frozenset({5468528})	0.000500342	0.789473684	565.6504469
frozenset({1588107})	frozenset({5528349})	0.000500342	0.736842105	158.0838537
frozenset({5528349})	frozenset({1588107})	0.004661084	0.079096045	158.0838537
frozenset({5468528})	frozenset({1686435})	0.001395692	0.264150943	716.490566
frozenset({1686435})	frozenset({5468528})	0.000368673	1	716.490566
frozenset({1686435})	frozenset({5528349})	0.000368673	0.928571429	199.2179177
frozenset({5528349})	frozenset({1686435})	0.004661084	0.073446328	199.2179177
frozenset({5468528})	frozenset({1688364})	0.001395692	0.245283019	665.3126685
frozenset({1688364})	frozenset({5468528})	0.000368673	0.928571429	665.3126685
frozenset({5468528})	frozenset({1756435})	0.001395692	0.283018868	826.7198839
frozenset({1756435})	frozenset({5468528})	0.000342339	1.153846154	826.7198839
frozenset({1756435})	frozenset({5528349})	0.000342339	1.076923077	231.0456323
frozenset({5528349})	frozenset({1756435})	0.004661084	0.079096045	231.0456323
frozenset({5468528})	frozenset({1832285})	0.001395692	0.320754717	171.5540792
frozenset({1832285})	frozenset({5468528})	0.0018697	0.23943662	171.5540792
frozenset({1832285})	frozenset({8159294})	0.0018697	0.183098592	239.758135
frozenset({8159294})	frozenset({1832285})	0.00076368	0.448275862	239.758135
frozenset({1949753})	frozenset({2419753})	0.000921683	0.657142857	891.2265306
frozenset({2419753})	frozenset({1949753})	0.000737347	0.821428571	891.2265306
frozenset({5468528})	frozenset({2088302})	0.001395692	0.283018868	767.6684636
frozenset({2088302})	frozenset({5468528})	0.000368673	1.071428571	767.6684636
frozenset({5528349})	frozenset({2088302})	0.004661084	0.079096045	214.5423729
frozenset({2088302})	frozenset({5528349})	0.000368673	1	214.5423729
frozenset({5468528})	frozenset({2176435})	0.001395692	0.301886792	573.1924528
frozenset({2176435})	frozenset({5468528})	0.000526676	0.8	573.1924528
frozenset({2176435})	frozenset({5528349})	0.000526676	0.75	160.9067797
frozenset({5528349})	frozenset({2176435})	0.004661084	0.084745763	160.9067797
frozenset({5468528})	frozenset({2178731})	0.001395692	0.245283019	258.7327044
frozenset({2178731})	frozenset({5468528})	0.000948017	0.361111111	258.7327044

frozenset({2316904})	frozenset({5468528})	0.00076368	0.551724138	395.3051399
frozenset({5468528})	frozenset({2316904})	0.001395692	0.301886792	395.3051399
frozenset({2726578})	frozenset({3998011})	0.001553695	0.338983051	195.0385208
frozenset({3998011})	frozenset({2726578})	0.001738031	0.303030303	195.0385208
frozenset({5468528})	frozenset({2927335})	0.001395692	0.283018868	716.490566
frozenset({2927335})	frozenset({5468528})	0.000395007	1	716.490566
frozenset({5528349})	frozenset({2927335})	0.004661084	0.079096045	200.239548
frozenset({2927335})	frozenset({5528349})	0.000395007	0.933333333	200.239548
frozenset({5468528})	frozenset({2959603})	0.001395692	0.301886792	521.084048
frozenset({2959603})	frozenset({5468528})	0.000579344	0.727272727	521.084048
frozenset({2959603})	frozenset({5528349})	0.000579344	0.681818182	146.2788906
frozenset({5528349})	frozenset({2959603})	0.004661084	0.084745763	146.2788906
frozenset({5468528})	frozenset({3029412})	0.001395692	0.301886792	424.5870021
frozenset({3029412})	frozenset({5468528})	0.000711013	0.592592593	424.5870021
frozenset({4992993})	frozenset({3161221})	0.00110602	0.333333333	162.2820513
frozenset({3161221})	frozenset({4992993})	0.002054037	0.179487179	162.2820513
frozenset({5468528})	frozenset({3161221})	0.001395692	0.301886792	146.9724238
frozenset({3161221})	frozenset({5468528})	0.002054037	0.205128205	146.9724238
frozenset({5468528})	frozenset({3301573})	0.001395692	0.264150943	590.0510544
frozenset({3301573})	frozenset({5468528})	0.000447675	0.823529412	590.0510544
frozenset({5528349})	frozenset({3301573})	0.004661084	0.073446328	164.0618146
frozenset({3301573})	frozenset({5528349})	0.000447675	0.764705882	164.0618146
frozenset({5468528})	frozenset({3559555})	0.001395692	0.245283019	194.0495283
frozenset({3559555})	frozenset({5468528})	0.001264023	0.270833333	194.0495283
frozenset({5468528})	frozenset({3589483})	0.001395692	0.245283019	517.4654088
frozenset({3589483})	frozenset({5468528})	0.000474009	0.722222222	517.4654088
frozenset({5468528})	frozenset({3680057})	0.001395692	0.245283019	665.3126685
frozenset({3680057})	frozenset({5468528})	0.000368673	0.928571429	665.3126685
frozenset({5468528})	frozenset({3690654})	0.001395692	0.245283019	194.0495283
frozenset({3690654})	frozenset({5468528})	0.001264023	0.270833333	194.0495283
frozenset({3968011})	frozenset({3898011})	0.001132354	0.418604651	252.3189369
frozenset({3898011})	frozenset({3968011})	0.00165903	0.285714286	252.3189369
frozenset({5468528})	frozenset({3968356})	0.001395692	0.245283019	404.9729286
frozenset({3968356})	frozenset({5468528})	0.000605678	0.565217391	404.9729286
frozenset({8129879})	frozenset({3989735})	0.001975036	0.266666667	163.3290323
frozenset({3989735})	frozenset({8129879})	0.001632696	0.322580645	163.3290323
frozenset({5468528})	frozenset({3999735})	0.001395692	0.245283019	465.7188679
frozenset({3999735})	frozenset({5468528})	0.000526676	0.65	465.7188679
frozenset({5468528})	frozenset({4138348})	0.001395692	0.264150943	271.1045385
frozenset({4138348})	frozenset({5468528})	0.000974351	0.378378378	271.1045385
frozenset({5468528})	frozenset({4142521})	0.001395692	0.264150943	668.7245283
frozenset({4142521})	frozenset({5468528})	0.000395007	0.933333333	668.7245283
frozenset({4142521})	frozenset({5528349})	0.000395007	0.866666667	185.9367232
frozenset({5528349})	frozenset({4142521})	0.004661084	0.073446328	185.9367232

frozenset({5468528})	frozenset({4356618})	0.001395692	0.283018868	671.7099057
frozenset({4356618})	frozenset({5468528})	0.000421341	0.9375	671.7099057
frozenset({4356618})	frozenset({5528349})	0.000421341	0.875	187.7245763
frozenset({5528349})	frozenset({4356618})	0.004661084	0.079096045	187.7245763
frozenset({5468528})	frozenset({4370921})	0.001395692	0.264150943	557.2704403
frozenset({4370921})	frozenset({5468528})	0.000474009	0.777777778	557.2704403
frozenset({4370921})	frozenset({5528349})	0.000474009	0.722222222	154.9472693
frozenset({5528349})	frozenset({4370921})	0.004661084	0.073446328	154.9472693
frozenset({5468528})	frozenset({4462521})	0.001395692	0.245283019	517.4654088
frozenset({4462521})	frozenset({5468528})	0.000474009	0.722222222	517.4654088
frozenset({5468528})	frozenset({4512521})	0.001395692	0.264150943	527.9404171
frozenset({4512521})	frozenset({5468528})	0.000500342	0.736842105	527.9404171
frozenset({4512521})	frozenset({5528349})	0.000500342	0.684210526	146.7921499
frozenset({5528349})	frozenset({4512521})	0.004661084	0.073446328	146.7921499
frozenset({4722472})	frozenset({5468528})	0.000500342	0.684210526	490.2303873
frozenset({5468528})	frozenset({4722472})	0.001395692	0.245283019	490.2303873
frozenset({4772472})	frozenset({5468528})	0.000474009	0.777777778	557.2704403
frozenset({5468528})	frozenset({4772472})	0.001395692	0.264150943	557.2704403
frozenset({4772472})	frozenset({5528349})	0.000474009	0.722222222	154.9472693
frozenset({5528349})	frozenset({4772472})	0.004661084	0.073446328	154.9472693
frozenset({5468528})	frozenset({4980033})	0.001395692	0.245283019	221.7708895
frozenset({4980033})	frozenset({5468528})	0.00110602	0.30952381	221.7708895
frozenset({5468528})	frozenset({5079905})	0.001395692	0.301886792	603.3604767
frozenset({5079905})	frozenset({5468528})	0.000500342	0.842105263	603.3604767
frozenset({5079905})	frozenset({5528349})	0.000500342	0.789473684	169.3755575
frozenset({5528349})	frozenset({5079905})	0.004661084	0.084745763	169.3755575
frozenset({5468528})	frozenset({5108107})	0.001395692	0.283018868	202.7803489
frozenset({5108107})	frozenset({5468528})	0.001395692	0.283018868	202.7803489
frozenset({5172064})	frozenset({5468528})	0.000526676	0.7	501.5433962
frozenset({5468528})	frozenset({5172064})	0.001395692	0.264150943	501.5433962
frozenset({5468528})	frozenset({5189905})	0.001395692	0.245283019	547.9045505
frozenset({5189905})	frozenset({5468528})	0.000447675	0.764705882	547.9045505
frozenset({5468528})	frozenset({5347422})	0.001395692	0.245283019	273.9522752
frozenset({5347422})	frozenset({5468528})	0.000895349	0.382352941	273.9522752
frozenset({5468528})	frozenset({5369905})	0.001395692	0.283018868	511.7789757
frozenset({5369905})	frozenset({5468528})	0.00055301	0.714285714	511.7789757
frozenset({5369905})	frozenset({5528349})	0.00055301	0.666666667	143.0282486
frozenset({5528349})	frozenset({5369905})	0.004661084	0.079096045	143.0282486
frozenset({5468528})	frozenset({5453386})	0.001395692	0.245283019	620.9584906
frozenset({5453386})	frozenset({5468528})	0.000395007	0.866666667	620.9584906
frozenset({5468528})	frozenset({5600671})	0.001395692	0.245283019	582.1485849
frozenset({5600671})	frozenset({5468528})	0.000421341	0.8125	582.1485849
frozenset({5468528})	frozenset({5749904})	0.001395692	0.245283019	388.0990566
frozenset({5749904})	frozenset({5468528})	0.000632011	0.541666667	388.0990566

frozenset({5468528})	frozenset({6060702})	0.001395692	0.283018868	398.0503145
frozenset({6060702})	frozenset({5468528})	0.000711013	0.555555556	398.0503145
frozenset({5468528})	frozenset({6280019})	0.001395692	0.245283019	517.4654088
frozenset({6280019})	frozenset({5468528})	0.000474009	0.722222222	517.4654088
frozenset({5468528})	frozenset({6308344})	0.001395692	0.264150943	771.605225
frozenset({6308344})	frozenset({5468528})	0.000342339	1.076923077	771.605225
frozenset({5468528})	frozenset({6318344})	0.001395692	0.264150943	179.1226415
frozenset({6318344})	frozenset({5468528})	0.001474693	0.25	179.1226415
frozenset({5468528})	frozenset({6349904})	0.001395692	0.245283019	423.380789
frozenset({6349904})	frozenset({5468528})	0.000579344	0.590909091	423.380789
frozenset({5468528})	frozenset({6441416})	0.001395692	0.245283019	716.490566
frozenset({6441416})	frozenset({5468528})	0.000342339	1	716.490566
frozenset({5468528})	frozenset({6759905})	0.001395692	0.245283019	716.490566
frozenset({6759905})	frozenset({5468528})	0.000342339	1	716.490566
frozenset({5468528})	frozenset({6949904})	0.001395692	0.245283019	238.8301887
frozenset({6949904})	frozenset({5468528})	0.001027018	0.333333333	238.8301887
frozenset({5468528})	frozenset({7029904})	0.001395692	0.264150943	501.5433962
frozenset({7029904})	frozenset({5468528})	0.000526676	0.7	501.5433962
frozenset({5468528})	frozenset({7202503})	0.001395692	0.245283019	358.245283
frozenset({7202503})	frozenset({5468528})	0.000684679	0.5	358.245283
frozenset({5468528})	frozenset({7710113})	0.001395692	0.245283019	232.859434
frozenset({7710113})	frozenset({5468528})	0.001053352	0.325	232.859434
frozenset({5468528})	frozenset({8159294})	0.001395692	0.339622642	444.7182824
frozenset({8159294})	frozenset({5468528})	0.00076368	0.620689655	444.7182824
frozenset({5468528})	frozenset({8239752})	0.001395692	0.245283019	145.5371462
frozenset({8239752})	frozenset({5468528})	0.001685364	0.203125	145.5371462
frozenset({5468528})	frozenset({8503540})	0.001395692	0.245283019	321.1854262
frozenset({8503540})	frozenset({5468528})	0.00076368	0.448275862	321.1854262
frozenset({5468528})	frozenset({8579778})	0.001395692	0.245283019	620.9584906
frozenset({8579778})	frozenset({5468528})	0.000395007	0.866666667	620.9584906
frozenset({5468528})	frozenset({8718362})	0.001395692	0.264150943	149.7144466
frozenset({8718362})	frozenset({5468528})	0.001764365	0.208955224	149.7144466
frozenset({5468528})	frozenset({8889414})	0.001395692	0.264150943	501.5433962
frozenset({8889414})	frozenset({5468528})	0.000526676	0.7	501.5433962
frozenset({5468528})	frozenset({9073382})	0.001395692	0.245283019	143.2981132
frozenset({9073382})	frozenset({5468528})	0.001711697	0.2	143.2981132
frozenset({5468528})	frozenset({9168271})	0.001395692	0.245283019	582.1485849
frozenset({9168271})	frozenset({5468528})	0.000421341	0.8125	582.1485849
frozenset({5468528})	frozenset({9277426})	0.001395692	0.245283019	141.1269297
frozenset({9277426})	frozenset({5468528})	0.001738031	0.196969697	141.1269297
frozenset({5468528})	frozenset({9486925})	0.001395692	0.245283019	388.0990566
frozenset({9486925})	frozenset({5468528})	0.000632011	0.541666667	388.0990566
frozenset({5468528})	frozenset({9552306})	0.001395692	0.264150943	477.6603774
frozenset({9552306})	frozenset({5468528})	0.00055301	0.666666667	477.6603774

frozenset({5468528})	frozenset({9569904})	0.001395692	0.264150943	668.7245283
frozenset({9569904})	frozenset({5468528})	0.000395007	0.933333333	668.7245283
frozenset({5468528})	frozenset({9686202})	0.001395692	0.245283019	665.3126685
frozenset({9686202})	frozenset({5468528})	0.000368673	0.928571429	665.3126685
frozenset({5468528})	frozenset({9702306})	0.001395692	0.283018868	826.7198839
frozenset({9702306})	frozenset({5468528})	0.000342339	1.153846154	826.7198839
frozenset({6308344})	frozenset({5528349})	0.000342339	1	214.5423729
frozenset({5528349})	frozenset({6308344})	0.004661084	0.073446328	214.5423729
frozenset({9569904})	frozenset({5528349})	0.000395007	0.866666667	185.9367232
frozenset({5528349})	frozenset({9569904})	0.004661084	0.073446328	185.9367232
frozenset({9702306})	frozenset({5528349})	0.000342339	1.076923077	231.0456323
frozenset({5528349})	frozenset({9702306})	0.004661084	0.079096045	231.0456323
frozenset({5848345})	frozenset({8129879})	0.000632011	0.541666667	274.2566667
frozenset({8129879})	frozenset({5848345})	0.001975036	0.173333333	274.2566667
frozenset({8899931})	frozenset({6697047})	0.000737347	0.535714286	484.3622449
frozenset({6697047})	frozenset({8899931})	0.00110602	0.357142857	484.3622449
frozenset({6899800})	frozenset({8129879})	0.000974351	0.513513514	260.0021622
frozenset({8129879})	frozenset({6899800})	0.001975036	0.253333333	260.0021622
frozenset({3524026, 3978011})	frozenset({3898011})	0.000579344	0.681818182	410.974026
frozenset({3978011, 3898011})	frozenset({3524026})	0.000526676	0.75	181.4044586
frozenset({3524026, 3898011})	frozenset({3978011})	0.000605678	0.652173913	204.6748114
frozenset({3978011})	frozenset({3524026, 3898011})	0.003186391	0.123966942	204.6748114
frozenset({3524026})	frozenset({3978011, 3898011})	0.004134408	0.095541401	181.4044586
frozenset({3898011})	frozenset({3524026, 3978011})	0.00165903	0.238095238	410.974026

APPENDIX D: Confidence-based Association Rule Table

antecedants	consequents	support	confidence	lift
frozenset({152307})	frozenset({5468528})	0.000342339	1	716.490566
frozenset({768635})	frozenset({5468528})	0.000342339	1	716.490566
frozenset({1100112})	frozenset({5468528})	0.000342339	1.076923077	771.605225
frozenset({1100112})	frozenset({5528349})	0.000342339	1	214.5423729
frozenset({1357342})	frozenset({5468528})	0.000421341	0.8125	582.1485849
frozenset({1686435})	frozenset({5468528})	0.000368673	1	716.490566
frozenset({1686435})	frozenset({5528349})	0.000368673	0.928571429	199.2179177
frozenset({1688364})	frozenset({5468528})	0.000368673	0.928571429	665.3126685
frozenset({1756435})	frozenset({5468528})	0.000342339	1.153846154	826.7198839
frozenset({1756435})	frozenset({5528349})	0.000342339	1.076923077	231.0456323
frozenset({2419753})	frozenset({1949753})	0.000737347	0.821428571	891.2265306
frozenset({2088302})	frozenset({5468528})	0.000368673	1.071428571	767.6684636
frozenset({2088302})	frozenset({5528349})	0.000368673	1	214.5423729
frozenset({2927335})	frozenset({5468528})	0.000395007	1	716.490566
frozenset({2927335})	frozenset({5528349})	0.000395007	0.933333333	200.239548
frozenset({3301573})	frozenset({5468528})	0.000447675	0.823529412	590.0510544
frozenset({3680057})	frozenset({5468528})	0.000368673	0.928571429	665.3126685
frozenset({4142521})	frozenset({5468528})	0.000395007	0.933333333	668.7245283
frozenset({4142521})	frozenset({5528349})	0.000395007	0.866666667	185.9367232
frozenset({4356618})	frozenset({5468528})	0.000421341	0.9375	671.7099057
frozenset({4356618})	frozenset({5528349})	0.000421341	0.875	187.7245763
frozenset({5079905})	frozenset({5468528})	0.000500342	0.842105263	603.3604767
frozenset({9168271})	frozenset({5268597})	0.000421341	0.8125	116.8707386
frozenset({5453386})	frozenset({5468528})	0.000395007	0.866666667	620.9584906
frozenset({5600671})	frozenset({5468528})	0.000421341	0.8125	582.1485849
frozenset({6308344})	frozenset({5468528})	0.000342339	1.076923077	771.605225
frozenset({6441416})	frozenset({5468528})	0.000342339	1	716.490566
frozenset({6759905})	frozenset({5468528})	0.000342339	1	716.490566
frozenset({8579778})	frozenset({5468528})	0.000395007	0.866666667	620.9584906
frozenset({9168271})	frozenset({5468528})	0.000421341	0.8125	582.1485849
frozenset({9569904})	frozenset({5468528})	0.000395007	0.933333333	668.7245283
frozenset({9686202})	frozenset({5468528})	0.000368673	0.928571429	665.3126685
frozenset({9702306})	frozenset({5468528})	0.000342339	1.153846154	826.7198839
frozenset({6308344})	frozenset({5528349})	0.000342339	1	214.5423729
frozenset({9569904})	frozenset({5528349})	0.000395007	0.866666667	185.9367232
frozenset({9702306})	frozenset({5528349})	0.000342339	1.076923077	231.0456323