Eric Yarger

Association Rules and Lift Analysis

Proposal of Question

Market basket analysis is an analytical process that identifies relationships between objects, items, or features that go well together. Our analyst team has been tasked with learning more about the typical prescriptions that our patients are prescribed. The dataset provided for this analysis contains 7,501 patient prescription histories spanning the past two years, and contains 20 features. The question posed for this analysis is:

'Using Market Basket Analysis, what are the top rules and insights discovered from our patient's prescription history, using a support minimum > .03, that can help guide our organization to providing better patient care and organizational profitability?'

Defined Goal

The top priority of this analysis, our teams' top goal, is to unearth viable, valuable insights from a dataset that could very likely be overlooked. The medical market basket dataset used for this analysis has just over seventy-five hundred patient prescription records, collected over the span of two years. This prescription history reflects all prescriptions prescribed at least once to the patient. According to the data dictionary, this includes Rx prescriptions, prescribed vitamins, and other medically prescribed solutions.

According to the 2021 US Prescription Drug Report, 75% of Americans don't think they receive good value for what they spend on healthcare, and 32% are unable to fill prescriptions or take less of prescribed medications than what is recommended because of its cost (The US prescription drug report 2021, n.d.). The goal of our team, to identify viable, actionable insights from this dataset, is both good for our organization and our patients.

If we can identify medications that are prescribed and taken together by other patients in our organization, from our dataset, our organization's physicians can better prescribe what future patients will require, reducing some of the ambiguity in treating patients. This benefits our patients directly through better care, and benefits our organization by providing a viable avenue for potentially reducing readmission rates at our hospitals.

Explanation of Market Basket

Market basket analysis is a technique used by organizations to find associations between different product or service offerings. It works by looking for combinations of items, in our case prescriptions, that frequently occur together. This is executed by looking for association rules. Association rules are used to qualify datasets by identifying strong rules using measures of inter-relatedness to identify offerings or products that share association (Li, 2017).

In these pairings the first feature is called the antecedent, and the second identified feature is called the consequent. The support metric in market basket analysis is defined as the frequency of the antecedent and consequent appearing in the dataset (Affinity analysis, 2022). Support and confidence levels, lift, leverage, and conviction are the rules and metrics that market basket pairing quality are measured with.

The expected outcome of this analysis is that our team will be able to discover a handful of medication combinations that occur together frequently in the dataset. Support, confidence, lift, leverage, and conviction will be calculated and used as metrics for assessing the discovered itemsets. It is expected that 15-20 medication pairings will be identified with a support threshold > .03.

Transaction Example

Provided below is a screenshot of the abridged unprepared dataset. An example of a transaction in this dataset is a full row, from column Presc01 to Presc20. Each row represents a unique transaction, consisting of the total and type of prescriptions the patient takes. In the unprepared dataset, the prescription name is listed in the transaction row, with 'NaN' used for a stand-in value that represents that there are no more prescriptions in that column for that transaction. Each row represents a separate transaction. Provided below is a screenshot from the unprepared dataset.

```
[84]: df
```

[84]:

	Presc01	Presc02	Presc03	Presc04	Presc05	Presc06	Presc07	Presc08	Presc09	Presc10	Presc11	Presc12	Presc13	Presc14	Presc15	Presc16	Presc17	Presc18	Presc19	Presc20
0	amlodipine	albuterol aerosol	allopurinol	pantoprazole	lorazepam	omeprazole	mometasone	fluconozole	gabapentin	pravastatin	cialis	losartan	metoprolol succinate XL	sulfamethoxazole	abilify	spironolactone	albuterol HFA	levofloxacin	promethazine	glipizide
1	citalopram	benicar	amphetamine salt combo xr	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
2	enalapril	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
3	paroxetine	allopurinol	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
4	abilify	atorvastatin	folic acid	naproxen	losartan	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...	...
7496	amphetamine	clotrimazole	lantus	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
7497	citalopram	metoprolol	amphetamine salt combo xr	glyburide	celebrex	losartan	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
7498	clopidogrel	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
7499	alprazolam	losartan	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN
7500	amphetamine salt combo xr	levofloxacin	diclofenac sodium	cialis	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN	NaN

The dataset needs to be prepared for Market Basket Analysis. Provided below is a screenshot showing an example of transactions in the dataset once it has been cleaned and prepared.

	Duloxetine	Premarin	Yaz	abilify	acetaminophen	actonel	albuterol HFA	albuterol aerosol	alendronate	allopurinol	...	trazodone HCI	triamcinolone Ace topical	triamterene	trimethoprim DS	valaciclovir	valsartan	venlafaxine XR	verapamil SR	viagra	zolpidem
0	False	False	False	True	False	False	True	True	False	True	...	False	False	False	False	False	False	False	False	False	False
1	False	False	False	False	False	False	False	False	False	False	...	False	False	False	False	False	False	False	False	False	False
2	False	False	False	False	False	False	False	False	False	False	...	False	False	False	False	False	False	False	False	False	False
3	False	False	False	False	False	False	False	False	False	True	...	False	False	False	False	False	False	False	False	False	False
4	False	False	False	True	False	False	False	False	False	False	...	False	False	False	False	False	False	False	False	False	False

As with the unprepared dataset, a single transaction is a row in the prepared dataset. However, the cleaned dataset has been transformed to be optimized for market basket analysis. Instead of 20 columns depicting the prescription # of the transaction, there are now 119 columns. One column for each prescription name. Each transaction now has a Boolean 'True' or 'False' for each prescription column.

Market Basket Assumption

One assumption of Market Basket Analysis (MBA) is the assumption of complementarity (Kamakura, 2012). Complementarity refers to the underlying assumption that joint occurrence of two or more products in most baskets implies that these products are complements in purchase, or in our case compliments in prescription. This can lead to the assumption that the purchase of one product will lead to the purchase of the other.

The algorithm used for this analysis is the Apriori algorithm. The Apriori algorithm uses frequent itemsets to get association rules, but on the assumptions that (Roshan, 2020):

- Subsets of frequent itemsets are frequent.
- In case of infrequent subsets a minimum support value is set to ensure that iteration of the algorithm occurs with frequent itemsets. Any itemset that falls below the specified minimum support threshold is ignored by the algorithm. In our analysis the minimum support threshold is set to .03.

Transforming the Dataset

The provided dataset was in need of cleaning and transformation to be made suitable for market basket analysis. This was accomplished by:

- Removing blank rows between transactions
- Using mlxtend.preprocessing TransactionEncoder to create a boolean array of the transactions.
- Assigning medication names to variable 'column_names'.
- Assigning the encoded array to a pandas DataFrame and assigning 'column_names' to the dataframe.

By following these steps the original dataset is transformed to be suitable for market basket analysis. A copy of the cleaned dataset is provided as an Excel file with this submission.

[96]:
```
# Apriori Market Basket Analysis
# Code Reference (Raschka, n.d.b)
itemsets = apriori(dataset, min_support=0.05, use_colnames=True)
```

[97]:
```
itemsets
```

[97]:

	support	itemsets
0	0.238368	(abilify)
1	0.079323	(alprazolam)
2	0.071457	(amlodipine)
3	0.068391	(amphetamine salt combo)
4	0.179709	(amphetamine salt combo xr)
5	0.129583	(atorvastatin)
6	0.174110	(carvedilol)
7	0.076523	(cialis)
8	0.087188	(citalopram)
9	0.059992	(clopidogrel)
10	0.081056	(dextroamphetamine XR)
11	0.163845	(diazepam)
12	0.095054	(doxycycline hyclate)
13	0.080389	(ezetimibe)
14	0.051060	(fenofibrate)
15	0.052393	(furosemide)
16	0.065858	(glipizide)
17	0.170911	(glyburide)
18	0.063325	(levofloxacin)
19	0.098254	(lisinopril)
20	0.132116	(losartan)
21	0.050527	(metformin)
22	0.095321	(metoprolol)
23	0.058526	(naproxen)
24	0.062525	(paroxetine)
25	0.050927	(abilify, amphetamine salt combo xr)
26	0.059725	(abilify, carvedilol)
27	0.052660	(abilify, diazepam)

```python
frequent_itemsets = apriori(dataset, min_support=0.03, use_colnames=True)
frequent_itemsets['length'] = frequent_itemsets['itemsets'].apply(lambda x: len(x))
frequent_itemsets
```

	support	itemsets	length
0	0.046794	(Premarin)	1
1	0.238368	(abilify)	1
2	0.033329	(allopurinol)	1
3	0.079323	(alprazolam)	1
4	0.071457	(amlodipine)	1
5	0.030129	(amphetamine)	1
6	0.068391	(amphetamine salt combo)	1
7	0.179709	(amphetamine salt combo xr)	1
8	0.129583	(atorvastatin)	1
9	0.174110	(carvedilol)	1
10	0.033729	(celecoxib)	1
11	0.076523	(cialis)	1
12	0.087188	(citalopram)	1
13	0.059992	(clopidogrel)	1
14	0.081056	(dextroamphetamine XR)	1
15	0.163845	(diazepam)	1
16	0.095054	(doxycycline hyclate)	1
17	0.080389	(ezetimibe)	1
18	0.051060	(fenofibrate)	1
19	0.031862	(fluconozole)	1
20	0.052393	(furosemide)	1
21	0.065858	(glipizide)	1
22	0.170911	(glyburide)	1
23	0.043061	(lantus)	1
24	0.063325	(levofloxacin)	1
25	0.098254	(lisinopril)	1
26	0.132116	(losartan)	1
27	0.050527	(metformin)	1
28	0.049460	(methylprednisone)	1
29	0.095321	(metoprolol)	1
30	0.047460	(metoprolol succinate XL)	1
31	0.032396	(metoprolol tartrate)	1
32	0.058526	(naproxen)	1
33	0.062525	(paroxetine)	1
34	0.030396	(pravastatin)	1
35	0.042528	(spironolactone)	1
36	0.050927	(abilify, amphetamine salt combo xr)	2

	support	itemsets	length
37	0.047994	(abilify, atorvastatin)	2
38	0.059725	(abilify, carvedilol)	2
39	0.052660	(abilify, diazepam)	2
40	0.033729	(abilify, doxycycline hyclate)	2
41	0.033729	(abilify, glyburide)	2
42	0.040928	(abilify, lisinopril)	2
43	0.031063	(abilify, losartan)	2
44	0.035729	(abilify, metoprolol)	2
45	0.030796	(amphetamine salt combo xr, atorvastatin)	2
46	0.036528	(amphetamine salt combo xr, carvedilol)	2
47	0.033196	(diazepam, amphetamine salt combo xr)	2
48	0.036395	(amphetamine salt combo xr, glyburide)	2
49	0.035462	(carvedilol, atorvastatin)	2
50	0.032129	(diazepam, atorvastatin)	2
51	0.039195	(diazepam, carvedilol)	2
52	0.039195	(lisinopril, carvedilol)	2
53	0.034395	(diazepam, glyburide)	2

0]:

```
frequent_itemsets[ (frequent_itemsets['length'] == 2) &
                   (frequent_itemsets['support'] >= 0.03) ]
```

0]:

	support	itemsets	length
36	0.050927	(abilify, amphetamine salt combo xr)	2
37	0.047994	(abilify, atorvastatin)	2
38	0.059725	(abilify, carvedilol)	2
39	0.052660	(abilify, diazepam)	2
40	0.033729	(abilify, doxycycline hyclate)	2
41	0.033729	(abilify, glyburide)	2
42	0.040928	(abilify, lisinopril)	2
43	0.031063	(abilify, losartan)	2
44	0.035729	(abilify, metoprolol)	2
45	0.030796	(amphetamine salt combo xr, atorvastatin)	2
46	0.036528	(amphetamine salt combo xr, carvedilol)	2
47	0.033196	(diazepam, amphetamine salt combo xr)	2
48	0.036395	(amphetamine salt combo xr, glyburide)	2
49	0.035462	(carvedilol, atorvastatin)	2
50	0.032129	(diazepam, atorvastatin)	2
51	0.039195	(diazepam, carvedilol)	2
52	0.039195	(lisinopril, carvedilol)	2
53	0.034395	(diazepam, glyburide)	2

```python
# Code Reference (Brown, 2019)
rules = association_rules(frequent_itemsets, min_threshold=.03)
rules
```

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
0	(abilify)	(amphetamine salt combo xr)	0.238368	0.179709	0.050927	0.213647	1.188845	0.008090	1.043158
1	(amphetamine salt combo xr)	(abilify)	0.179709	0.238368	0.050927	0.283383	1.188845	0.008090	1.062815
2	(abilify)	(atorvastatin)	0.238368	0.129583	0.047994	0.201342	1.553774	0.017105	1.089850
3	(atorvastatin)	(abilify)	0.129583	0.238368	0.047994	0.370370	1.553774	0.017105	1.209650
4	(abilify)	(carvedilol)	0.238368	0.174110	0.059725	0.250559	1.439085	0.018223	1.102008
5	(carvedilol)	(abilify)	0.174110	0.238368	0.059725	0.343032	1.439085	0.018223	1.159314
6	(abilify)	(diazepam)	0.238368	0.163845	0.052660	0.220917	1.348332	0.013604	1.073256
7	(diazepam)	(abilify)	0.163845	0.238368	0.052660	0.321400	1.348332	0.013604	1.122357
8	(abilify)	(doxycycline hyclate)	0.238368	0.095054	0.033729	0.141499	1.488616	0.011071	1.054100
9	(doxycycline hyclate)	(abilify)	0.095054	0.238368	0.033729	0.354839	1.488616	0.011071	1.180529
10	(abilify)	(glyburide)	0.238368	0.170911	0.033729	0.141499	0.827912	-0.007011	0.965741
11	(glyburide)	(abilify)	0.170911	0.238368	0.033729	0.197348	0.827912	-0.007011	0.948894
12	(abilify)	(lisinopril)	0.238368	0.098254	0.040928	0.171700	1.747522	0.017507	1.088672
13	(lisinopril)	(abilify)	0.098254	0.238368	0.040928	0.416554	1.747522	0.017507	1.305401
14	(abilify)	(losartan)	0.238368	0.132116	0.031063	0.130313	0.986357	-0.000430	0.997927
15	(losartan)	(abilify)	0.132116	0.238368	0.031063	0.235116	0.986357	-0.000430	0.995748
16	(abilify)	(metoprolol)	0.238368	0.095321	0.035729	0.149888	1.572463	0.013007	1.064189
17	(metoprolol)	(abilify)	0.095321	0.238368	0.035729	0.374825	1.572463	0.013007	1.218270
18	(amphetamine salt combo xr)	(atorvastatin)	0.179709	0.129583	0.030796	0.171365	1.322437	0.007509	1.050423
19	(atorvastatin)	(amphetamine salt combo xr)	0.129583	0.179709	0.030796	0.237654	1.322437	0.007509	1.076009
20	(amphetamine salt combo xr)	(carvedilol)	0.179709	0.174110	0.036528	0.203264	1.167446	0.005239	1.036592
21	(carvedilol)	(amphetamine salt combo xr)	0.174110	0.179709	0.036528	0.209801	1.167446	0.005239	1.038081
22	(diazepam)	(amphetamine salt combo xr)	0.163845	0.179709	0.033196	0.202604	1.127397	0.003751	1.028711
23	(amphetamine salt combo xr)	(diazepam)	0.179709	0.163845	0.033196	0.184718	1.127397	0.003751	1.025603
24	(amphetamine salt combo xr)	(glyburide)	0.179709	0.170911	0.036395	0.202522	1.184961	0.005681	1.039640
25	(glyburide)	(amphetamine salt combo xr)	0.170911	0.179709	0.036395	0.212949	1.184961	0.005681	1.042232
26	(carvedilol)	(atorvastatin)	0.174110	0.129583	0.035462	0.203675	1.571779	0.012900	1.093043
27	(atorvastatin)	(carvedilol)	0.129583	0.174110	0.035462	0.273663	1.571779	0.012900	1.137061
28	(diazepam)	(atorvastatin)	0.163845	0.129583	0.032129	0.196094	1.513276	0.010898	1.082736
29	(atorvastatin)	(diazepam)	0.129583	0.163845	0.032129	0.247942	1.513276	0.010898	1.111823
30	(diazepam)	(carvedilol)	0.163845	0.174110	0.039195	0.239219	1.373952	0.010668	1.085581
31	(carvedilol)	(diazepam)	0.174110	0.163845	0.039195	0.225115	1.373952	0.010668	1.079070
32	(lisinopril)	(carvedilol)	0.098254	0.174110	0.039195	0.398915	2.291162	0.022088	1.373997
33	(carvedilol)	(lisinopril)	0.174110	0.098254	0.039195	0.225115	2.291162	0.022088	1.163716
34	(diazepam)	(glyburide)	0.163845	0.170911	0.034395	0.209927	1.228284	0.006393	1.049383
35	(glyburide)	(diazepam)	0.170911	0.163845	0.034395	0.201248	1.228284	0.006393	1.046827

Association Rules Table

Provided below is a screenshot of the generated Association Rules. Values for support, lift, confidence, leverage, and conviction are generated.

```
# Code Reference (Brown, 2019)
rules = association_rules(frequent_itemsets, min_threshold=.03)
rules
```

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
0	(abilify)	(amphetamine salt combo xr)	0.238368	0.179709	0.050927	0.213647	1.188845	0.008090	1.043158
1	(amphetamine salt combo xr)	(abilify)	0.179709	0.238368	0.050927	0.283383	1.188845	0.008090	1.062815
2	(abilify)	(atorvastatin)	0.238368	0.129583	0.047994	0.201342	1.553774	0.017105	1.089850
3	(atorvastatin)	(abilify)	0.129583	0.238368	0.047994	0.370370	1.553774	0.017105	1.209650
4	(abilify)	(carvedilol)	0.238368	0.174110	0.059725	0.250559	1.439085	0.018223	1.102008
5	(carvedilol)	(abilify)	0.174110	0.238368	0.059725	0.343032	1.439085	0.018223	1.159314
6	(abilify)	(diazepam)	0.238368	0.163845	0.052660	0.220917	1.348332	0.013604	1.073256
7	(diazepam)	(abilify)	0.163845	0.238368	0.052660	0.321400	1.348332	0.013604	1.122357
8	(abilify)	(doxycycline hyclate)	0.238368	0.095054	0.033729	0.141499	1.488616	0.011071	1.054100
9	(doxycycline hyclate)	(abilify)	0.095054	0.238368	0.033729	0.354839	1.488616	0.011071	1.180529
10	(abilify)	(glyburide)	0.238368	0.170911	0.033729	0.141499	0.827912	-0.007011	0.965741
11	(glyburide)	(abilify)	0.170911	0.238368	0.033729	0.197348	0.827912	-0.007011	0.948894
12	(abilify)	(lisinopril)	0.238368	0.098254	0.040928	0.171700	1.747522	0.017507	1.088672
13	(lisinopril)	(abilify)	0.098254	0.238368	0.040928	0.416554	1.747522	0.017507	1.305401
14	(abilify)	(losartan)	0.238368	0.132116	0.031063	0.130313	0.986357	-0.000430	0.997927
15	(losartan)	(abilify)	0.132116	0.238368	0.031063	0.235116	0.986357	-0.000430	0.995748
16	(abilify)	(metoprolol)	0.238368	0.095321	0.035729	0.149888	1.572463	0.013007	1.064189
17	(metoprolol)	(abilify)	0.095321	0.238368	0.035729	0.374825	1.572463	0.013007	1.218270
18	(amphetamine salt combo xr)	(atorvastatin)	0.179709	0.129583	0.030796	0.171365	1.322437	0.007509	1.050423
19	(atorvastatin)	(amphetamine salt combo xr)	0.129583	0.179709	0.030796	0.237654	1.322437	0.007509	1.076009
20	(amphetamine salt combo xr)	(carvedilol)	0.179709	0.174110	0.036528	0.203264	1.167446	0.005239	1.036592
21	(carvedilol)	(amphetamine salt combo xr)	0.174110	0.179709	0.036528	0.209801	1.167446	0.005239	1.038081
22	(diazepam)	(amphetamine salt combo xr)	0.163845	0.179709	0.033196	0.202604	1.127397	0.003751	1.028711
23	(amphetamine salt combo xr)	(diazepam)	0.179709	0.163845	0.033196	0.184718	1.127397	0.003751	1.025603
24	(amphetamine salt combo xr)	(glyburide)	0.179709	0.170911	0.036395	0.202522	1.184961	0.005681	1.039640
25	(glyburide)	(amphetamine salt combo xr)	0.170911	0.179709	0.036395	0.212949	1.184961	0.005681	1.042232
26	(carvedilol)	(atorvastatin)	0.174110	0.129583	0.035462	0.203675	1.571779	0.012900	1.093043
27	(atorvastatin)	(carvedilol)	0.129583	0.174110	0.035462	0.273663	1.571779	0.012900	1.137061
28	(diazepam)	(atorvastatin)	0.163845	0.129583	0.032129	0.196094	1.513276	0.010898	1.082736
29	(atorvastatin)	(diazepam)	0.129583	0.163845	0.032129	0.247942	1.513276	0.010898	1.111823
30	(diazepam)	(carvedilol)	0.163845	0.174110	0.039195	0.239219	1.373952	0.010668	1.085581
31	(carvedilol)	(diazepam)	0.174110	0.163845	0.039195	0.225115	1.373952	0.010668	1.079070
32	(lisinopril)	(carvedilol)	0.098254	0.174110	0.039195	0.398915	2.291162	0.022088	1.373997
33	(carvedilol)	(lisinopril)	0.174110	0.098254	0.039195	0.225115	2.291162	0.022088	1.163716
34	(diazepam)	(glyburide)	0.163845	0.170911	0.034395	0.209927	1.228284	0.006393	1.049383
35	(glyburide)	(diazepam)	0.170911	0.163845	0.034395	0.201248	1.228284	0.006393	1.046827

Top Three Rules

Provided below is a screenshot identifying the top three rules generated by the Apriori algorithm.

```
[103]: rules.sort_values(by=['confidence'], ascending=False)
```

[103]:

	antecedents	consequents	antecedent support	consequent support	support	confidence	lift	leverage	conviction
13	(lisinopril)	(abilify)	0.098254	0.238368	0.040928	0.416554	1.747522	0.017507	1.305401
32	(lisinopril)	(carvedilol)	0.098254	0.174110	0.039195	0.398915	2.291162	0.022088	1.373997
17	(metoprolol)	(abilify)	0.095321	0.238368	0.035729	0.374825	1.572463	0.013007	1.218270

The top three rules are identified by sorting the 'rules' table by confidence in descending order. This lists each rule by confidence level, with the highest confidence level listed first.

To provide a quick summary for each rule:

- lisinopril → abilify
    - Confidence is .416554. This is a measurement of accuracy that tells us that for every transaction with the presence of lisinopril, abilify is also present.
    - Support is .040928. This measures how frequently the itemset occurs in the data.
    - Lift is 1.747522. This measures how much more likely an itemset is prescribed relative to the typical prescribe rate.
- lisinopril → carvedilol
    - Confidence is .398915. This is a measurement of accuracy that tells us that for every transaction with the presence of lisinopril, carvedilol is also present.
    - Support is .039195. This measures how frequently the itemset occurs in the data.
    - Lift is 2.291162. This measures how much more likely an itemset is prescribed relative to the typical prescribe rate.
- metoprolol → abilify
    - Confidence is .374825. This is a measurement of accuracy that tells us that for every transaction with the presence of metoprolol, abilify is also present.
    - Support is .035729. This measures how frequently the itemset occurs in the data.

  - Lift is 1.572463. This measures how much more likely an itemset is prescribed relative to the typical prescribe rate.

Significance of Support, Lift, and Confidence Summary

The significance of Support, Lift, and Confidence are summarized below. The results of this analysis were specific to having a minimum support > .03 specified in the Apriori algorithm.

- Support. Support is a measure of how frequently the itemset occurs in the data. This is significant in the analysis because it provides an initial basis for our analyst team to recognize how common the itemset is in the data. The more frequent an itemset, the more likely it is that it will be frequent in future patients. This makes Support a significant and applicable metric to prioritize and include in our research.
  - Support levels from the analysis ranged from a high of .059725 for itemset abilify → carvedilol to a low of .030796 for itemset amphetamine salt combo xr → atorvastatin.
- Lift. The Lift of a rule measures how much more likely an itemset is prescribed typical to the relative rate of prescription. If Lift > 1 for an itemset, it implies that the itemsets are found together more often than chance would dictate (Junaid, 2019). Large lift values are a strong indicator that an association rule is viable, and that the relationship between the antecedent and consequent is true.
  - Lift levels from the analysis ranged from a high of 2.291162 for itemset carvedilol → lisinopril to a low of .827912 for itemset abilify → glyburide.
- Confidence. Confidence is a measurement of accuracy in market basket analysis. Confidence tells our analyst team the proportion of prescriptions where the inclusion of the antecedent results in the presence of the consequent.
  - Confidence levels from the analysis ranged from a high of .416554 for lisinopril → abilify to a low of .130313 for abilify → losartan.

Practical Significance of Findings

The results of this analysis provides our organization with a look at all itemsets with a minimum support threshold specified at > .03. This resulted in a total of 36 unique pairs, with each containing 1 antecedent and

1 consequent. Metrics and association rules are calculated for each itemset, as discussed and shown in sections C3 and D1 of this document.

The practical significance of the results of this analysis should be viewed from both a medical and business perspective. Medically, the findings of this analysis provide our organizations physicians, nurses, and full medical team with a potent tool. From a diagnosis perspective, if a patient is taking a medication that has a high confidence level with a medication from our findings, a physician can use this information in potentially forming a more thorough plan of treatment in our hospitals. From a post-treatment perspective, this analysis can help identify viable prescription pairings that can positively impact patient health outcomes after they leave our hospital.

Both diagnostically and post-discharge from the hospital, the results of this analysis have the potential to provide value for our medical staff and the patients they treat. This in turn has the potential to be a part of our organizations' strategy in lowering readmission rates. From a business perspective lowering readmission rates and providing the best care possible for our patients are the top priorities. Focusing on patient care above all else paves the road for our organizations' continued profitability, success, and reaching our targeted goal of lowering readmission rates.

Course of Action

To restate the question from section A1 of this document:

'Using Market Basket Analysis, what are the top rules and insights discovered from our patient's prescription history, using a support minimum = .03, that can help guide our organization to providing better patient care and organizational profitability?'

My recommended course of action for our organization is to make use of the results of this analysis as a screening tool for patient medication and medical history. This will be useful for our physicians in making medication recommendations for future patients that may not be obvious at first glance. Knowledge of the level of co-prescribed medication allows our physicians to more effectively recommend useful medications to patients that have the capacity to increase their quality of treatment.

Higher quality of treatment in-house provides our patients the chance of enjoying less medical problems post treatment. When used as a tool in our organizations' toolkit for effective patient treatment and prescription allocation, the results of this analysis have the potential to help lower readmission rates at our organizations.

Web Sources

Raschka, S. (n.d.a). TransactionEncoder - mlxtend. Retrieved from https://rasbt.github.io/mlxtend/user_guide/preprocessing/TransactionEncoder/

Boston, S. (2019, October 17). Applying transaction encoder on dataset. Retrieved from https://stackoverflow.com/questions/58435648/applying-transaction-encoder-on-dataset

Brown, E. (2019, December 26). Market basket analysis with Python and pandas. Retrieved from https://pythondata.com/market-basket-analysis-with-python-and-pandas/

Raschka, S. (n.d.b). Apriori - mlxtend. Retrieved from https://rasbt.github.io/mlxtend/user_guide/frequent_patterns/apriori/

G Sources

The US prescription drug report 2021. (n.d.). Retrieved from https://www.nicerx.com/prescription-drug-report-2021/

Li, S. (2017, September 27). A gentle introduction on market basket analysis — Association rules. https://towardsdatascience.com/a-gentle-introduction-on-market-basket-analysis-association-rules-fa4b986a40ce

Affinity analysis. (2022, January 28). Retrieved from https://en.wikipedia.org/wiki/Affinity_analysis

Kamakura, W. (2012, May 22). *Sequential market basket analysis.* Retrieved from wak2.web.rice.edu/bio/My%20Reprints/Sequential%20Market%20Basket%20Analysist.pdf

Junaid. (2019, November 18). Market basket analysis – The Apriori algorithm. Retrieved from https://scienceofdata.org/2019/10/27/market-basket-analysis-the-apriori-algorithm/

Roshan, B. (2020, December 22). Market basket analysis- Association rule mining with visualizations. Retrieved from https://medium.com/analytics-vidhya/market-basket-analysis-association-rule-mining-with-visualizations-cda24d537019