# D212 Task 3

AUTHOR

Tyson Biegler. Student ID: 012170282

## Part I: Research Question

**A1:** What are the most commonly co-prescribed medications, and are there any strong association rules that suggest patterns in prescription behavior?

**A2:** My goal is to identify common medication combinations using market basket analysis, providing insights into prescription patters that can help executives better understand patient prescriptions.

## Part II: Market Basket Justification

**B1:** Market basket analysis finds patterns in the data by identifying frequently co-occurring variables, prescription in this case. In this case I am using the apriori algorithm which finds combinations of medications that are often prescribed together. As a result, the apriori algorithm will return the Support, Confidence, and Lift that I will explain in a later section.

After running market basket analysis using the apriori algorithm on this data I expect to see the following outcomes:

1. Groups of prescription that are frequently paired together.

2. The support value indicating the rate that these prescription combinations happen together.

3. Confidence values that determine the likelihood that groups will be prescribed together.

4. Lift, to determine if the prescriptions that are prescribed together is more likely than by chance alone.

**B2:** Not all transactions included any prescriptions. The following prescriptions are included in the second transaction. The first transaction did not include any prescriptions.

```
 [1] "abilify"             "albuterol aerosol"
 [3] "allopurinol"         "amlodipine"
 [5] "cialis"              "fluconozole"
 [7] "gabapentin"          "lorazepam"
 [9] "losartan"            "metoprolol succinate XL"
[11] "mometasone"          "omeprazole"
[13] "pantoprazole"        "pravastatin"
[15] "sulfamethoxazole"
```

Displayed below is a sample of the second transaction. The full transaction is very large so I included the first six values.

```
     abilify    acetaminophen        actonel albuterol aerosol
        TRUE            FALSE          FALSE              TRUE
```

```
    albuterol HFA          alendronate
          FALSE                FALSE
```

**B3:** One main assumption of market basket analysis is that certain products, prescriptions in this case, tend to appear together in multiple transactions more often than expected by chance. This implies that there is a consistent occurrence of two or more prescriptions across numerous baskets. In the context of this medical data the assumption is that specific medications are frequently prescribed together, indicating possible relationships or patters, or just common treatment combinations.

# Part III: Data Preparation and Analysis

**C1:** I will provide a cleaned csv file called ''*cleaned_basket.csv*" in the assessment submission.

**C2:** The following is the apriori algorithm code without errors. At the end of this code I print the line "Code ran successfully without errors" to demonstrate that there were no prior errors in the code.

```
# Analysis ---------------------------------------------------------
arules <- apriori(
  basket,
  control = list(verbose = FALSE),
  parameter = list(supp = 0.008, conf = 0.4, minlen = 2),
  appearance = list(rhs = "abilify", default = "lhs")
)

redundant_r <- is.redundant(arules) #removing redundant rules

refined_arules <- arules[!redundant_r]

inspect(sort(refined_arules, by = 'lift', decreasing = TRUE))
```

```
    lhs                         rhs        support      confidence coverage
[1] {metformin}              => {abilify} 0.011531796 0.4564644  0.02526330
[2] {carvedilol, lisinopril} => {abilify} 0.008532196 0.4353741  0.01959739
[3] {glipizide}              => {abilify} 0.013731502 0.4178499  0.03286229
[4] {lisinopril}             => {abilify} 0.020463938 0.4165536  0.04912678
    lift     count
[1] 3.829910 173
[2] 3.652955 128
[3] 3.505920 206
[4] 3.495043 307
```

```
# If this line runs then there were no previous errors
print("Code ran successfully without errors.")
```

```
[1] "Code ran successfully without errors."
```

**C3:** All values for this association rules table are listed below. With the confidence set to 0.4, five rules are returned.

```
     lhs                            rhs         support      confidence coverage
[1] {metformin}              => {abilify} 0.011531796 0.4564644  0.02526330
[2] {carvedilol, lisinopril} => {abilify} 0.008532196 0.4353741  0.01959739
[3] {glipizide}              => {abilify} 0.013731502 0.4178499  0.03286229
[4] {lisinopril}             => {abilify} 0.020463938 0.4165536  0.04912678
     lift     count
[1] 3.829910 173
[2] 3.652955 128
[3] 3.505920 206
[4] 3.495043 307
```

**C4:** The following are the first three rules from the association rules table that I will explain in more detail.

```
     lhs                            rhs         support      confidence coverage
[1] {metformin}              => {abilify} 0.011531796 0.4564644  0.02526330
[2] {carvedilol, lisinopril} => {abilify} 0.008532196 0.4353741  0.01959739
[3] {glipizide}              => {abilify} 0.013731502 0.4178499  0.03286229
     lift     count
[1] 3.829910 173
[2] 3.652955 128
[3] 3.505920 206
```

In the first rule `{metformin} => {abilify}`, the support is 0.0115, meaning 1.15% of all transactions include both Metformin and Abilify. The lift value is 3.83, making this the strongest association in the table. A lift of 3.83 indicates that patients prescribed Metformin are 3.83 times more likely to also be prescribed Abilify compared to random chance. The confidence is 0.4565, meaning that when Metformin appears in a transaction, there is a 45.65% probability that Abilify is also present.

The second rule `{carvedilol, lisinopril} => {abilify}` has the lowest support (0.0085), indicating that 0.85% of all transactions contain carvedilol, lisinopril, and Abilify. The lift is 3.65, meaning patients prescribed both Carvedilol and Lisinopril are 3.65 times more likely to also be prescribed Abilify compared to random chance. The confidence is 0.4354, indicating that when Carvedilol and lisinopril appear in a transaction, there is a 43.54% chance that Abilify is also present.

The third rule `{glipizide} => {abilify}` has the highest support (0.0137), meaning 1.37% of transactions contain both glipizide and Abilify. The lift is 3.51, suggesting that patients prescribed glipizide are 3.51 times more likely to also be prescribed Abilify compared to random chance. The confidence is 0.4178, meaning there is a 41.78% probability that Abilify is included when glipizide is prescribed.

# Part IV: Data Summary and Implications

**D1:** The highest support value among the top three rules was 0.0137 (1.37%), meaning that 1.37% of all transactions contained both glipizide and Abilify. Higher support values indicate more frequent prescribing patterns. The highest lift value was 3.83 for the rule `{metformin} => {abilify}`, meaning that patients prescribed Metformin were 3.83 times more likely to also be prescribed abilify compared to random chance. The highest confidence value was 0.4565 (45.65%), indicating that when Metformin was prescribed, there was a 45.65% probability that abilify was also included in the transaction.

**D2:** There are many practical implications of this analysis, such as understanding commodities among diagnosis. For example according to drugs.com, Metformin is primarily used to manage diabetes ((Drugs.com, 2023a) where as Abilify is an antipychotic (Drugs.com, 2023b). This appears to suggest that a subset of patients with diabetes is also being treated for mental health issues. Additionally, if a pharmacy knows these associations they can better stock medications if they know that patients that are prescribed Abilify (the most common prescription in the dataset) have a 43.54% probability that they will also need Metformin.

**D3:** Based on this analysis, the hospital executives should focus on medication management to ensure all frequently co-prescribed medications are available, assuming that this hospital chain provides the medications as well as prescribes them. Additionally, this information should be loaded into the electronic health record so that the health care providers can review common medication combinations and make more informed decisions.

# Part V: Attachments

E.  The Panopto video link will be provided in the submission files.

F. Code sources:

- Kamara, K. (2022, August 28). Data mining II - D212 [Video]. Panopto. https://wgu.hosted.panopto.com/Panopto/Pages/Viewer.aspx?id=5674b196-a9f1-4e85-a322-af0000021f3f

G. In text sources:

- Drugs.com. (2023a, August 22). Metformin: Uses, dosage, side effects, and warnings. https://www.drugs.com/metformin.html

- Drugs.com. (2023b, August 22). Abilify: Uses, dosage, side effects, and warnings. https://www.drugs.com/abilify.html