Analysis & Clustering of Pharmaceutical Reviews

Presented by Lindsey Hunnicutt

Prescription Medication

In 2015 - 2018…

- *1 billion drugs* were prescribed
- *48%* of the U.S. population took *at least 1** prescribed drug in the last 30 days
- *34%* of the U.S. population took *at least 3** …
- *24%* of the U.S. population took 5 or more* …

*Source: FastStats - Therapeutic Drug Use. CDC. https://www.cdc.gov/nchs/fastats/drug-use-therapeutic.htm. (Reviewed Jan. 2023.)

Reviews of Pharmaceutical Drugs

- Analgesics[1] (pain medication[2]) most frequently prescribed class (2015-2018)
- Many options available & effects vary between people
- Drugs.com as a resource for patients & health professionals
- Reviews publicly available as a dataset
- As a data scientist with an interest in improving patient outcomes…

1. FastStats - Therapeutic Drug Use. CDC. https://www.cdc.gov/nchs/fastats/drug-use-therapeutic.htm. (Reviewed 2023.)
2. Analgesics: Use and Side Effects. Cleveland Clinic. https://my.clevelandclinic.org/health/drugs/21483-analgesics. (Reviewed 2021.)

Research Question

I undertook this study in order to answer the question:

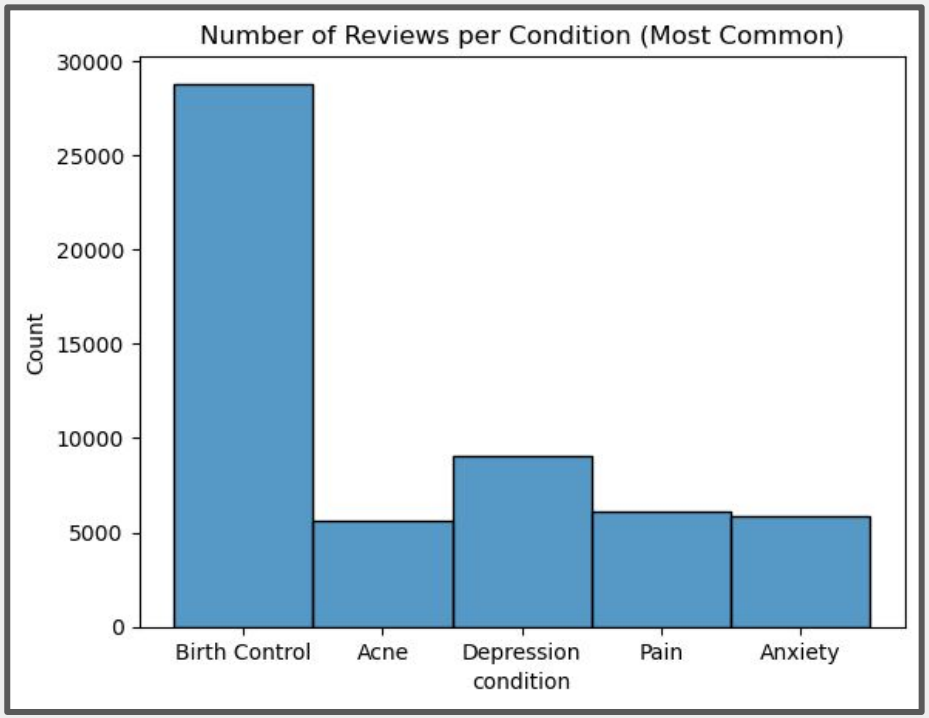
How can we use analysis and clustering of pharmaceutical drug reviews to improve patient outcomes?

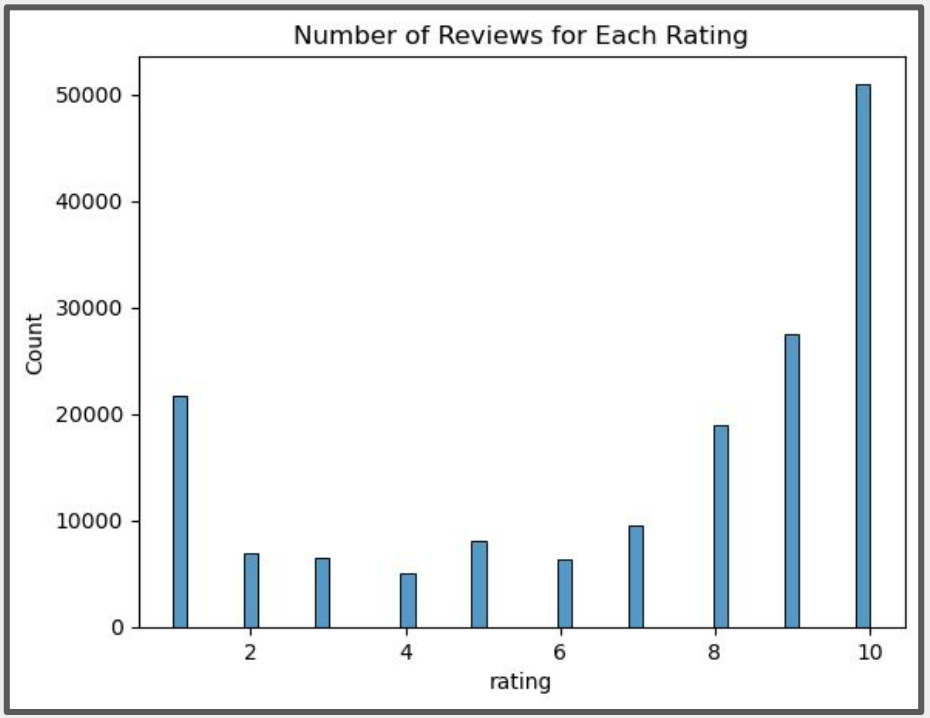
Data Processing

Datasets Utilized

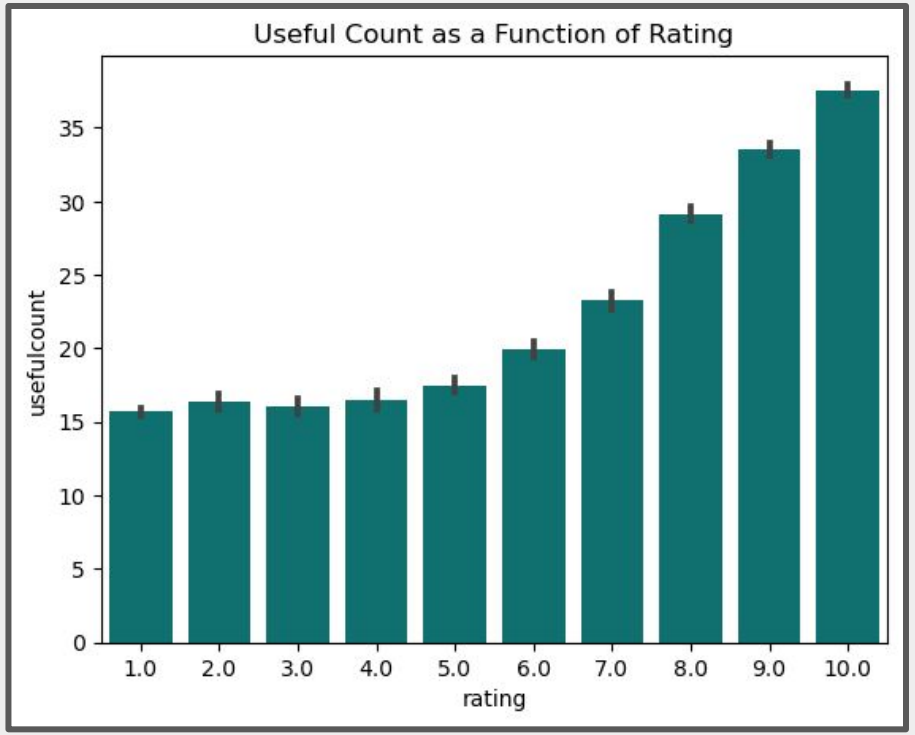
1. All Reviews (161,297)
2. Pain Medication Reviews (6,145)
    a. 10 Rating (2,108)
    b. 1 Rating (582)
    c. Of drugs frequently receiving ratings of 10 (1,208)
    d. Of drugs frequently receiving ratings of 1 (1,007)
    e. Of Oxycodone (281)
    f. Of Acetaminophen/codeine (94)
    g. Of Tramadol (357)

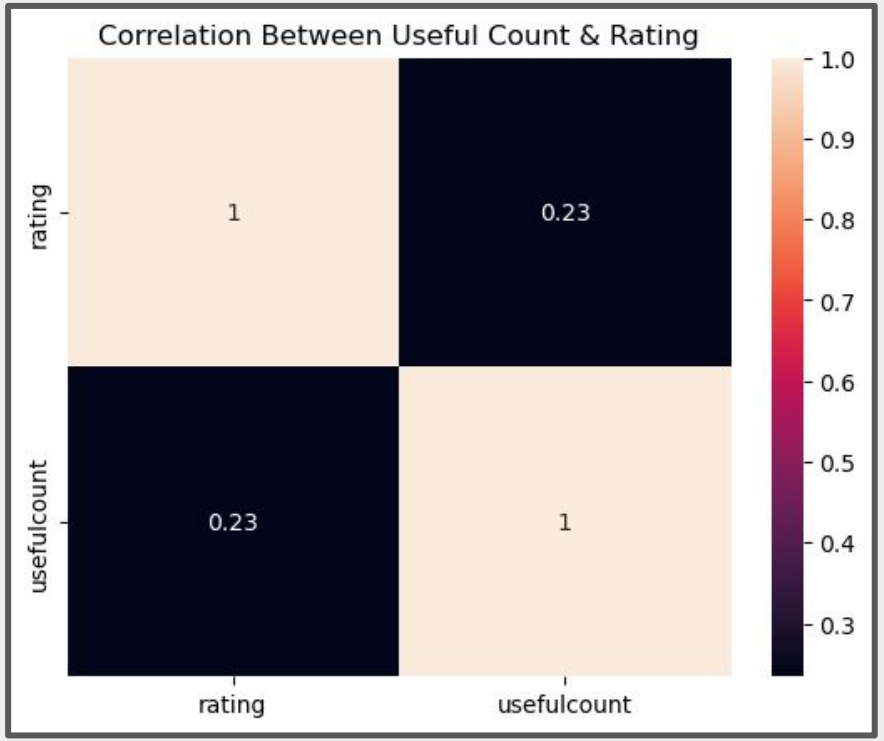
EDA - All Conditions



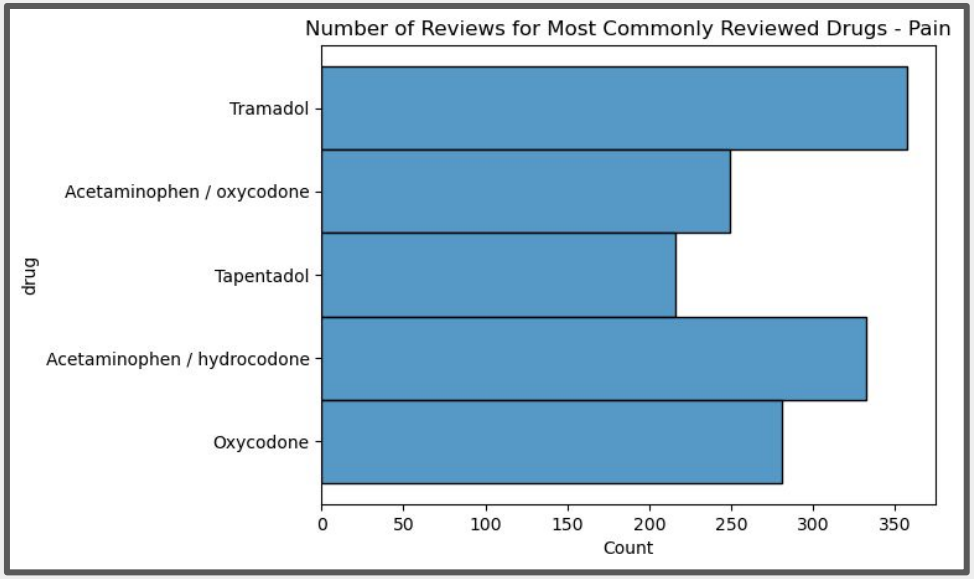


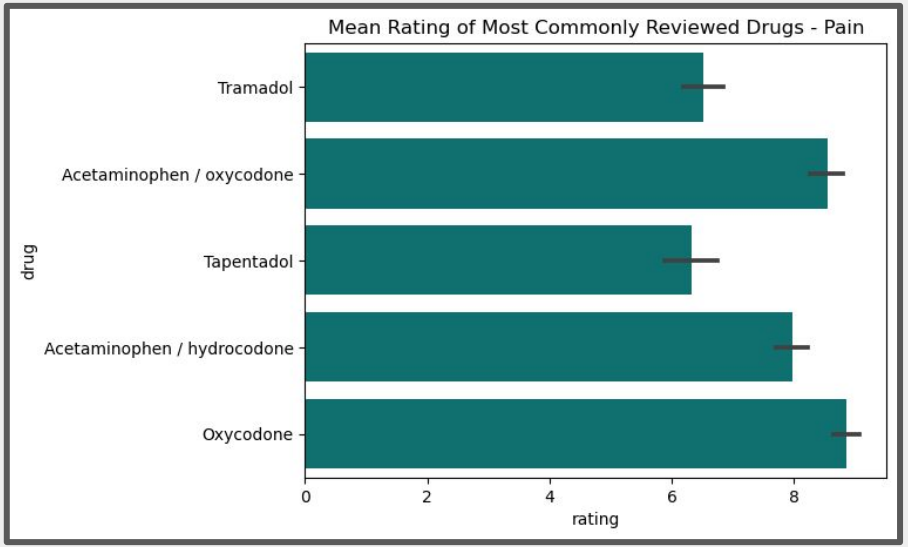
EDA - All Conditions



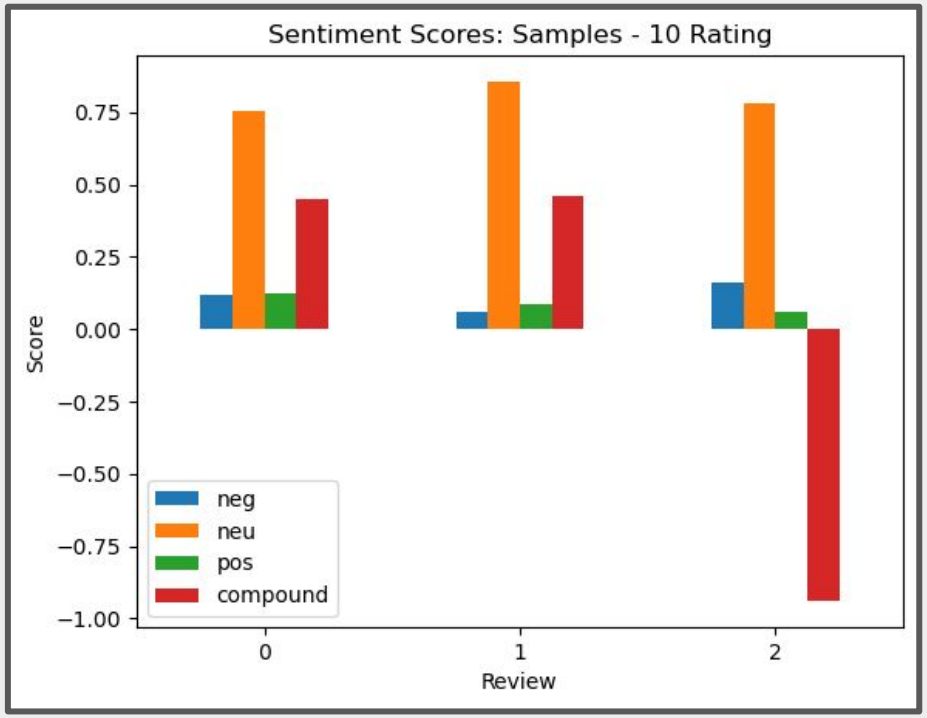


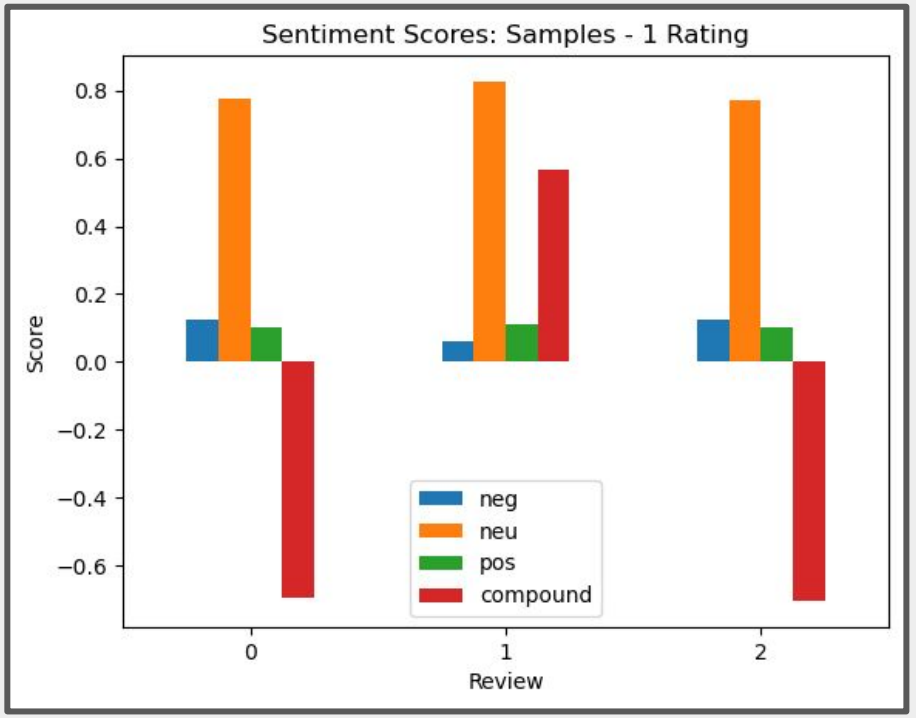
EDA - Pain


Number of Reviews for Most Commonly Reviewed Drugs - Pain
Tramadol
Acetaminophen / oxycodone
Tapentadol
Acetaminophen / hydrocodone
Oxycodone
drug
0
50
100
150
200
250
300
350
Count


Mean Rating of Most Commonly Reviewed Drugs - Pain
Tramadol
Acetaminophen / oxycodone
Tapentadol
Acetaminophen / hydrocodone
Oxycodone
drug
0
2
4
6
8
rating

EDA - Pain

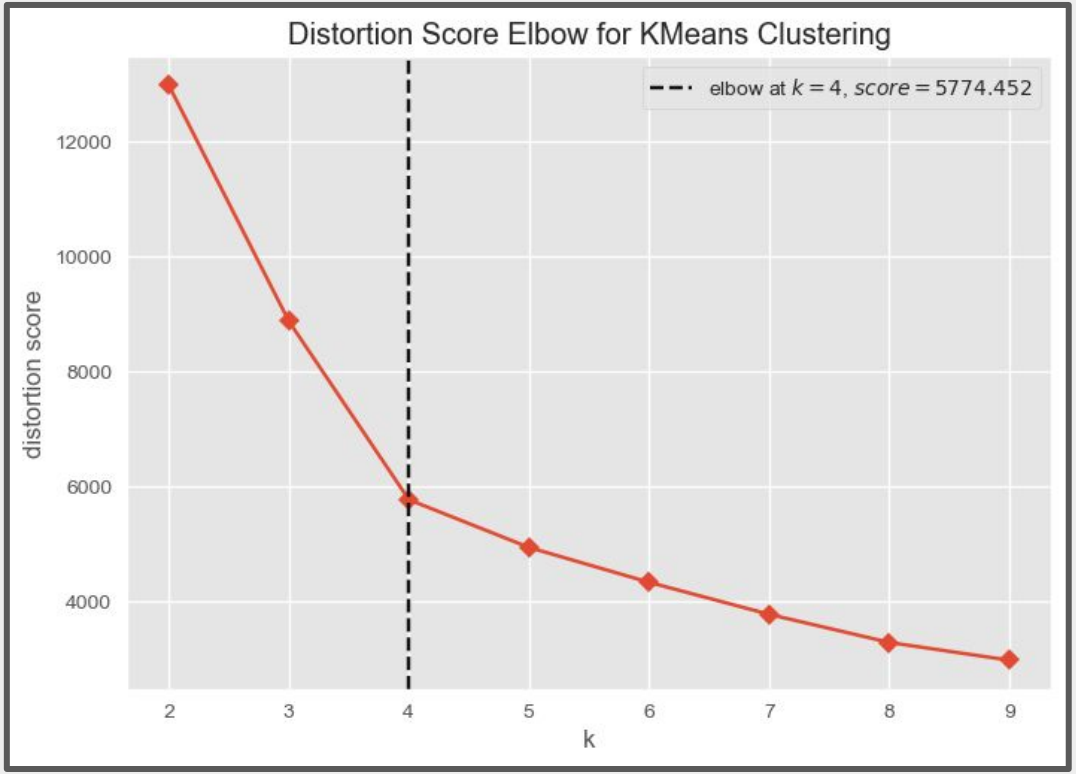

Sentiment Scores: Samples - 10 Rating
Score
0.75
0.50
0.25
0.00
−0.25
−0.50
−0.75
−1.00
neg
neu
pos
compound
0
1
2
Review


Sentiment Scores: Samples - 1 Rating
Score
0.8
0.6
0.4
0.2
0.0
−0.2
−0.4
−0.6
neg
neu
pos
compound
0
1
2
Review

K-Means - Clustering

Dataset	Optimal K	Highest Silhouette
10 & 1 Rated Reviews	3	0.538
10 & 1 Common	4	0.444
Oxycodone & Tramadol	4	0.477
All Pain	4	0.445
All Conditions	4	0.521

Elbow Plot - All Pain



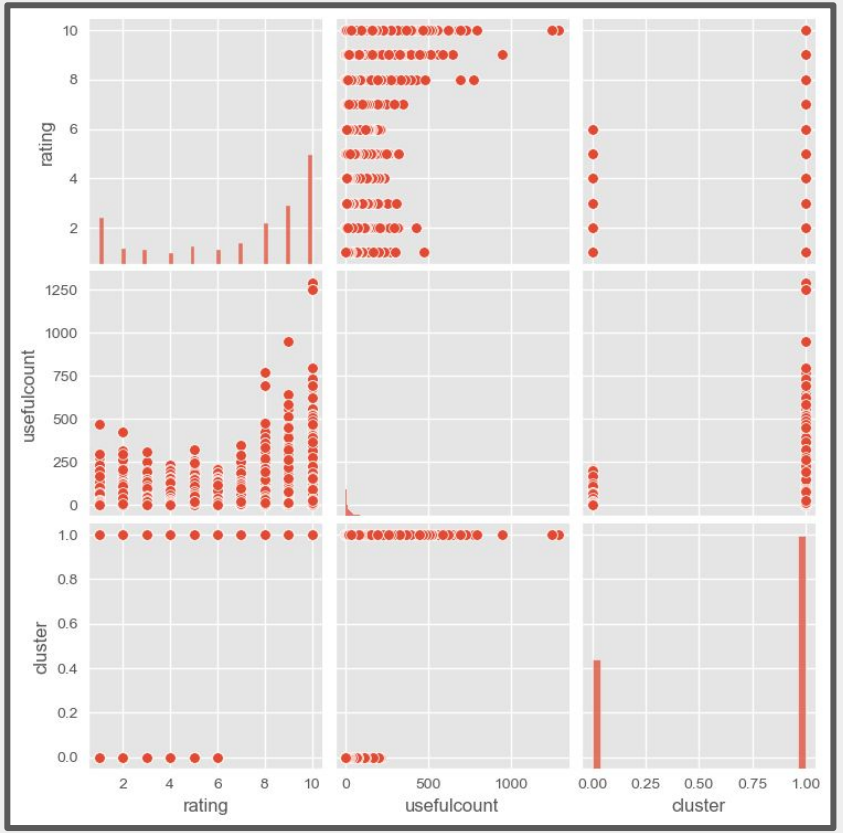
K - Means Clustering

Clusters often formed as described below…

Cluster	*Rating**	*Compound**	Useful Count
!	*High*	*Strongly (-)*	Low
@	*High*	*Slightly (-)*	High
#	*High*	*Moderately (+)*	Low
&	*Low*	*Moderately (-)*	Low

** Where applicable - not all datasets contained numerical ratings and/or sentiment scores*

Pair Plot - All Conditions



Conclusion

- EDA provides insights on a number of useful topics
  - Most/least commonly reviewed
  - Highest/lowest rated drugs
  - Average ratings
  - Average sentiment scores
  - And more…

- K-Means Clustering may provide opportunities to observe patterns in certain groups
  - Sentiment Scores
  - Ratings
  - Useful Count

Next Steps & Recommendations

- Process can be utilized to analyze review and other text
- Further analysis could provide insight regarding medication options, side effects, and other information patients, providers, and/or companies need
- Professional collaboration may generate findings useful for improving patient experience

Thank you!