Preventing Customer Churn, Identifying Hidden Segments and Improving Satisfaction of Pharmaceutical Drug Users

By: Yuan Tieming

Problem Statement

- Pharmaceutical companies are facing the challenge of customer churn. Unhappy customers are switching to other companies and giving bad reviews that might negatively impact their brand value. Therefore, there is a need to efficiently identify unhappy customers and quickly conduct customer recovery. Companies will also need to make improvements to their drugs after gathering more information from unsatisfied customers. Besides that, pharmaceutical companies will need to know their customer segments in order to provide tailored services to enhance customer experience and also to develop marketing strategies to boost sales. Lastly, companies will need to efficiently recommend medicines to customers based on their condition or switch drugs when customer is not satisfied with their existing drug.
- Therefore, our problem statement can be summed up to: preventing customer churn, identify hidden segments and improving satisfaction.
- Pharmaceutical companies, clinics and hospitals will benefit from this project. It will help them to reduce customer churn, improve their service offering and making their business more profitable. Patients will also benefit from better service and more effective medicines.

Contents

- Methodology
- Exploratory Data Analysis
- Sentiment Classification
- Clustering
- Recommender System
- Conclusion and recommendation



Data Inspection

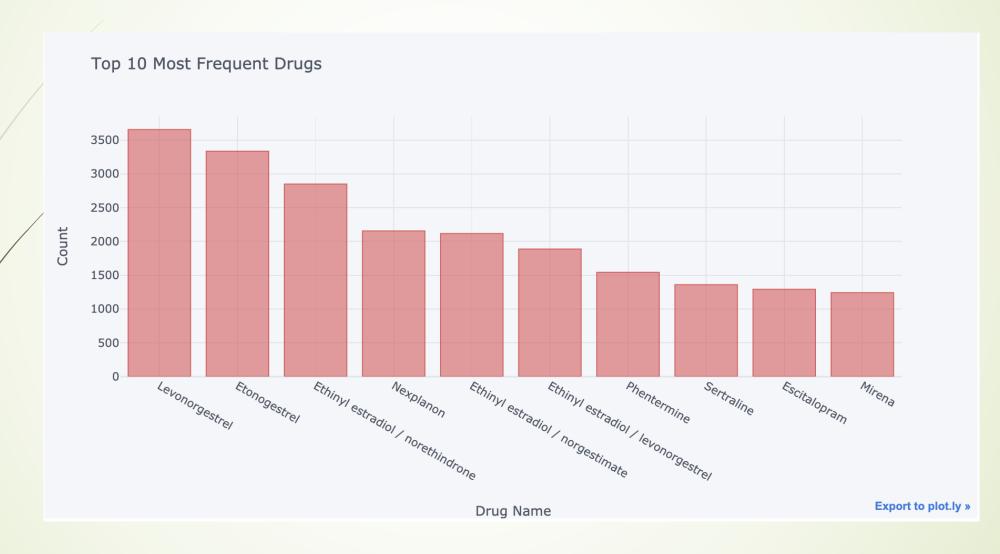
Exploratory
Data Analysis

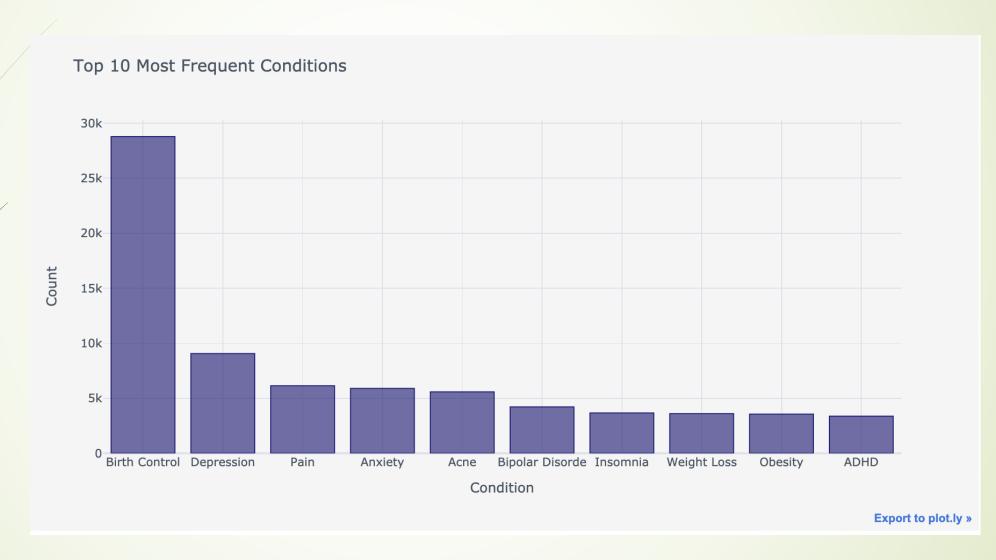
Data Preprocessing

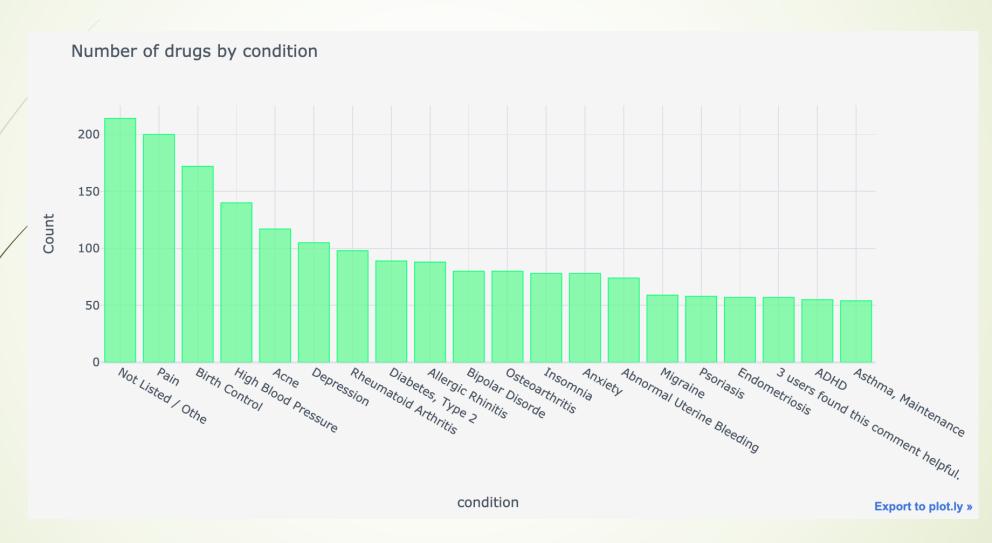
Classification
Modelling

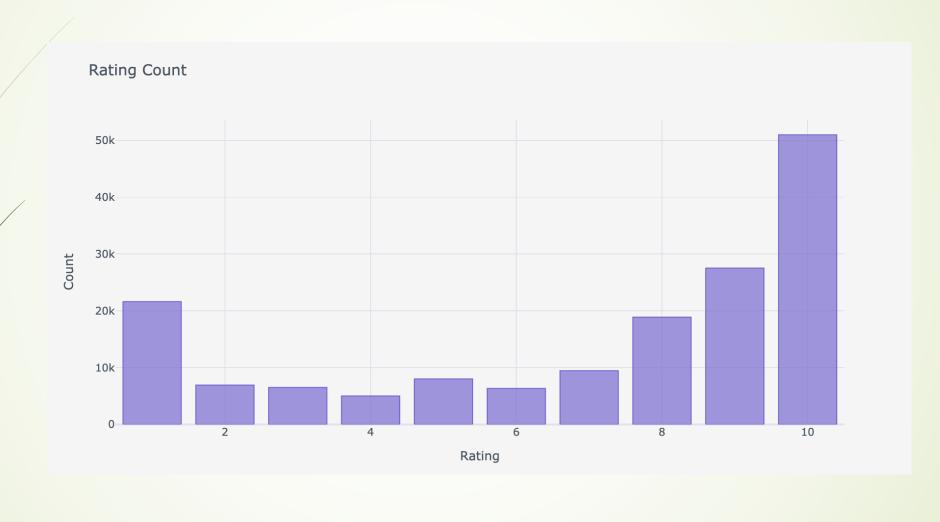
Clustering

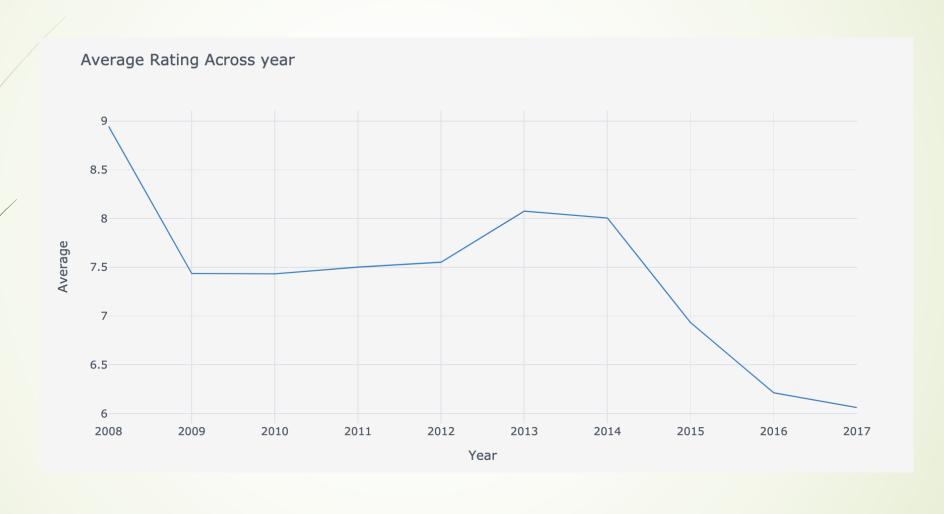
Recommender













Classification Models

Training Models	Accuracy	Precision	Recall
Base Line	0.5	-	-
Logistic Regression	0.82	0.83	0.81
Naïve Bayes	0.78	0.77	0.78
Random Forest	0.93	0.94	0.92
XGBoost	0.74	0.76	0.70
Neural Network	0.91	0.93	0.88

Testing Models	Accuracy	Precision	Recall
Logistic Regression	0.84	0.86	0.92
Random Forest	0.91	0.91	0.98

Logistic Coefficients

	coef	exp_coef
love	9.820405	18405.503769
amazing	8.745989	6285.427037
miracle	8.149585	3461.940519
changed life	7.387071	1614.969654
saved	7.004149	1101.192062
great	6.812076	908.755686
best	6.754044	857.519433
wonderful	6.144873	466.320538
complaint	5.749842	314.141105
highly recommend	5.564138	260.900160

	coef	exp_coef
worse	-6.596473	0.001365
disappointed	-6.450838	0.001579
wouldn recommend	-6.093831	0.002257
don recommend	-5.575770	0.003789
ruined	-5.477704	0.004179
worst	-5.471094	0.004207
doesn work	-4.861253	0.007741
stay away	-4.844076	0.007875
stopped taking	-4.693362	0.009156
waste	-4.577534	0.010280

Clustering

Cluster 1: The Image Conscious

- Uses Birth Control, Weight loss and anti-acne drugs
- Concern about the drug's side effect on their physical appearance
- Young to middle age female users
- Short to mid term drug user

"Fantastic! No bloating or weight gain!! Yippee!!! Hot flashes are gone, helped the first week, would have a few here and their but they were mild and by the second month- they are totally gone and sleeping so much better- the feeling of anxiety is gone- life grand again! \r\r\nImmmmmm mmm soooo happy!"



Clustering

Cluster 2: The Mental Well-Being

- Uses anti-depression, anxiety and ADHD drugs
- Concern about their mood, energy level and thoughts
- All age range and gender
- Long term drug user

"This is my miracle. I was put on it in Germany and had to go off it until I had insurance again. I suffer from depression and anxiety. Often I would find myself stuck in a daydream world because of all the things going on in my life, including my partner of years (15-25). I would have no energy and just feel so low. Finally I had insurance and was put on 1800mg a day. I feel so much better. I feel like I can handle life and set goals for myself. I don't feel anxiety. I don't sit here and think of the negative and focus on a positive life. I am more fun and interactive with my son. \r\n\r\nIt literately saved my life. I don't know why."



Clustering

Cluster 3: The Pain Killers

- Uses pain easing drugs
- Concern about mobility and level of pain
- All age range and gender
- Both short and long term users

"I have severe psoriatic arthritis and I' ve had it since I was 18. I' ve been on remicade for almost 10 years and I have almost no pain. My joints get a little stiff close to my infusion but that' s it. It is my miracle drug and I owe my life to it. I would be crippled without it."



Recommender

```
search = 'AbobotulinumtoxinA'
for drugName in df train rec.loc[df train rec['drugName'].str.contains(search), 'drugName']:
   print(drugName)
   print('Average rating', pivot.loc[drugName, :].mean())
   print('Number of ratings', pivot.T[drugName].count())
   print('')
   print('10 closest drugs')
   print(recommender_df[drugName].sort_values()[1:11])
   print('')
   print('')
AbobotulinumtoxinA
Average rating 6.5
Number of ratings 2
10 closest drugs
drugName
Dysport
                       0.018866
IncobotulinumtoxinA
                       0.152002
Juvederm Voluma XC
                       0.470001
Botox Cosmetic
                       0.470001
                       0.564689
Botox
OnabotulinumtoxinA
                       0.621743
Botulinum toxin type b
                       0.652011
Lioresal
                       0.670278
Tazarotene
                       0.685495
Dermal filler
                       0.693433
Name: AbobotulinumtoxinA, dtype: float64
```

help_dict['Birth Control'].iloc[:10]

	drug	average_rating
0	Plan B	8.674917
1	Lybrel	7.600000
2	Ortho Evra	7.470320
3	Kyleena	7.463576
4	Levonorgestrel	7.391031
5	Ortho-Cept	7.250000
6	Skyla	7.246359
7	Copper	7.235197
8	Beyaz	7.042553
9	Dienogest / estradiol	6.875000

Conclusion and Recommendation

- Marketing message should contain words such as love, amazing, miracle and life changing
- Differentiated Marketing Strategies
 - Marketing message for Image conscious segment could be better appearance, no weight gain or better skin
 - Marketing message for Mental well-being could be better mood, passivity and good energy
 - Marketing message for Pain killer could be disappearing pain
- Companies should quickly conduct customer recovery service when a negative sentiment was detected
 - Use the random forest model to classify positive and negative review
 - Conduct more R&D on drugs that generally have bad reviews
 - Recommend new drugs based on the recommender system built
 - New customers can recommend based on condition
- Companies should focus on developing drugs for birth control, weight loss, depression, anxiety and pain
 - Biggest customer base as shown in exploratory data analysis
 - Cluster analysis and topic modelling revealed three clusters

Questions?