https://www.covermymeds.com/main/

ClusterMyMeds

Predicting Patient Pay, Insurance Formulary, and Drug Type.

Team Phoenix:

Mita Banik, Craig Franze, George Hauser, Haidong Tian

github.com/TeamPhoenix22/CMM_Challenge

Erdős Institute Data Science Bootcamp, May 2022

Problem Context & Overview

Companies like CoverMyMeds have access to a large quantity of pharmacy transaction data. Our goal is extrapolate from their data information about the expenses that patients should expect to see at the pharmacy. This information is relevant to the patients as well as their healthcare providers.

Out-of-pocket costs for prescription drugs are determined by complicated interactions between **drug manufacturers**, **insurance companies**, and **pharmacies**. Accurately predicting patient pay despite these complex factors is important for both **patients** and **healthcare providers**.

Detailed official coverage information is difficult to track down, but individual transaction data is abundant.

Problem: predict patients' **out-of-pocket pay** and drugs' **formulary status** for each insurance plan, directly from a large database of transaction data.

This helps patients save money on prescriptions, and also helps doctors to make an informed decision about medication costs when prescribing medications.

uspharmacist.com/article/a-pharmacists-primer-on-prescription-discount-cards

Our Approach

Data Exploration: The data consists of around 14 million pharmacy data-billing claims simulated by CoverMyMeds. Researched insurance plans and explored factors impacting the variation of copayments with respect to the time of year, pharmacy and insurance plans.

Copay prediction: Introduced new features by feature engineering and predicted copayments with an average 4% MAPE.

Formulary prediction: Predicted formulary status of each medication on an insurance plan based on discount offered on each medication and rejection rates.

Medication clustering : Clustered drugs based on similar rejection rates and copayment requirements on Type 1 and Type 2 insurance plans. Clustered drugs based on deductible and coinsurance for Type 3 insurance plans.

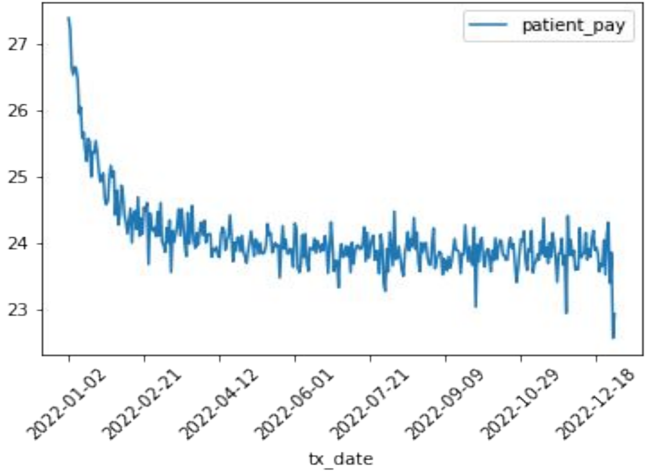
The Dataset

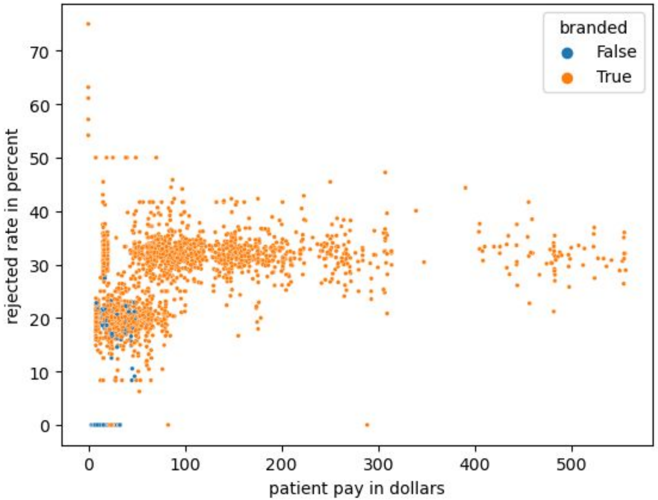
	tx_date	pharmacy	diagnosis	drug	bin	pcn	group	rejected	patient_pay
12173703	2022-11-16	Pharmacy #30	U60.52	generic keglusited	725700	327CKV	IOEAN1DWVV3Y	False	11.15
11715121	2022-11-07	Pharmacy #31	G51.87	branded prazinib	757349	MSCXSG	DGLGRYP	False	17.56
11017844	2022-10-20	Pharmacy #2	Q85.91	branded cupitelol	725700	327CKV	IOEAN1DWVV3Y	False	11.15
10859249	2022-10-16	Pharmacy #24	G99.93	generic rulfalol	664344	NaN	STGRDKR1J5RD	False	5.94
13314104	2022-12-13	Pharmacy #54	N55.01	branded dusin	664344	YFVIA	AJK5MZ25T9IA	True	0.00

Glimpse

- 13.9 million entries
- 114 drugs
- 133 diagnosis codes
- 61 insurance plans
- 7.8% rejected claims

Higher out-of-pocket pay in the earlier months.





Branded drugs tend to be expensive and get rejected more often.

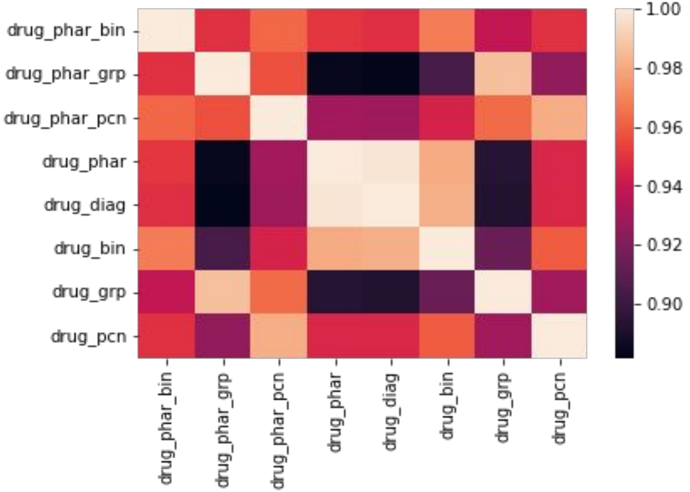
Predicting out-of-pocket payments

Observations :

- For each combination of *(drug*, *insurance plan, pharmacy)* there are either one or two values of *patient_pay*.
- Generally, we see the higher copay early in the year. This is due to the *deductible phase*.

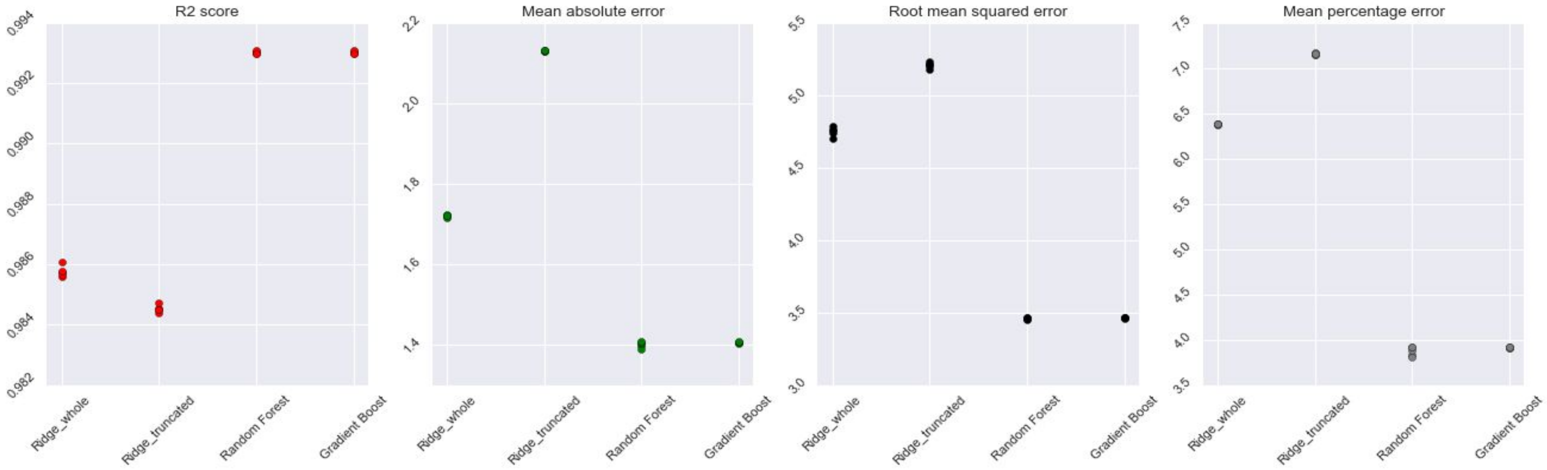
Preprocessing :

- Using exploratory analysis we extract 8 important features that would help predict patient pay.
- Truncating duplicates from training data to improve model performance.



Prediction Models

Metrics : We used a variety of metrics such as explained variance(r^2 score), MAE, RMSE and MAPE to compare our models.



Final Model : We chose **Random Forests** as our final model (with a 4% MAPE score and 0.99 explained variance).

Error Analysis: We observed more mismatch between our prediction and actual patient pay for copay-before-deductible entries. We recommend another model which predicts the copay-before-deductible if the doctor is interested in knowing out-of-pocket pay before the deductible phase starts.

Formulary Status

https://www.bcbsm.com/medicare/help/understanding-plans/pharmacy-prescription-drugs/tiers.html

Insurance plans determine drug prices based on a *formulary*: A classification of drugs into tiers which determine how coverage is applied.

Definition of formulary status:

High Tier (non-preferred branded/specialty):

Higher Costs,
Higher rejection rate,
Less coverage
Higher deductible

Low Tier (generic/preferred branded):

Lower cost
Infrequent rejection
High or 100% coverage.
Low or no deductible

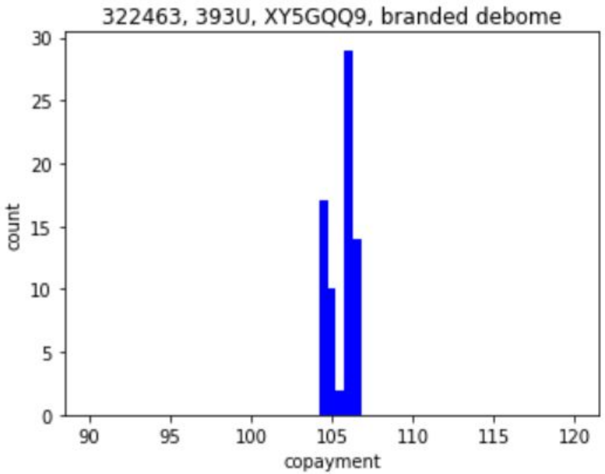
Different plans have different tiers, different numbers of tiers, and different assignments of drugs to tiers.

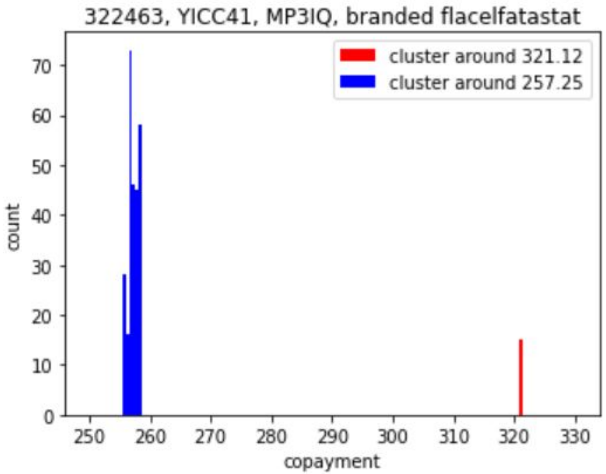
From Data alone, can we predict drug tiers for each plan? Clustering problem.

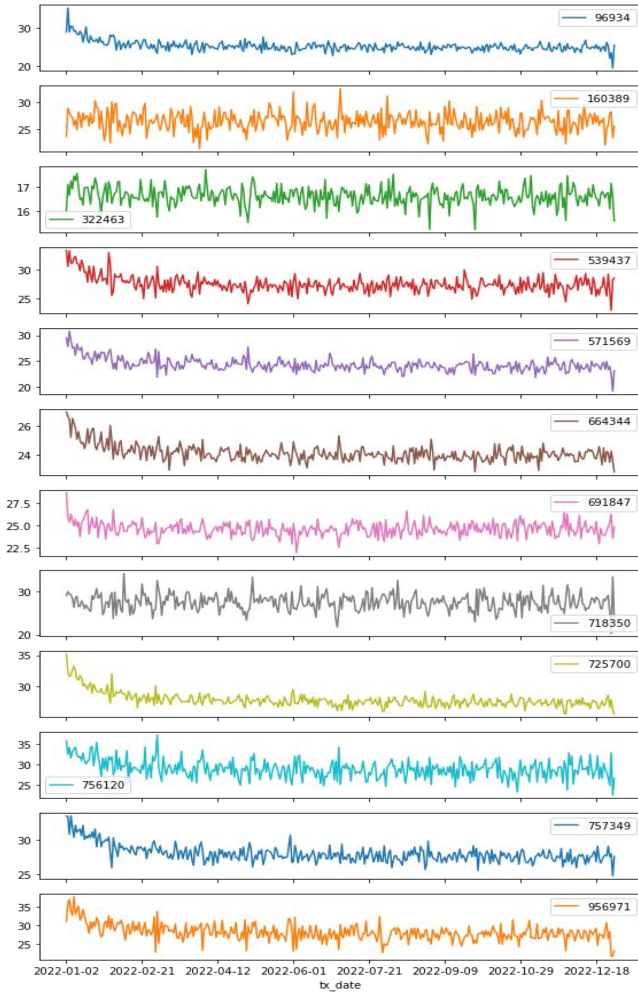
Deductible Phase

Patient pay skews higher early in the year (more for some BINs than others).

Gaussian mixture model to separate pre/post-deductible phases. Some drug/plan combinations have only one price, others have two.

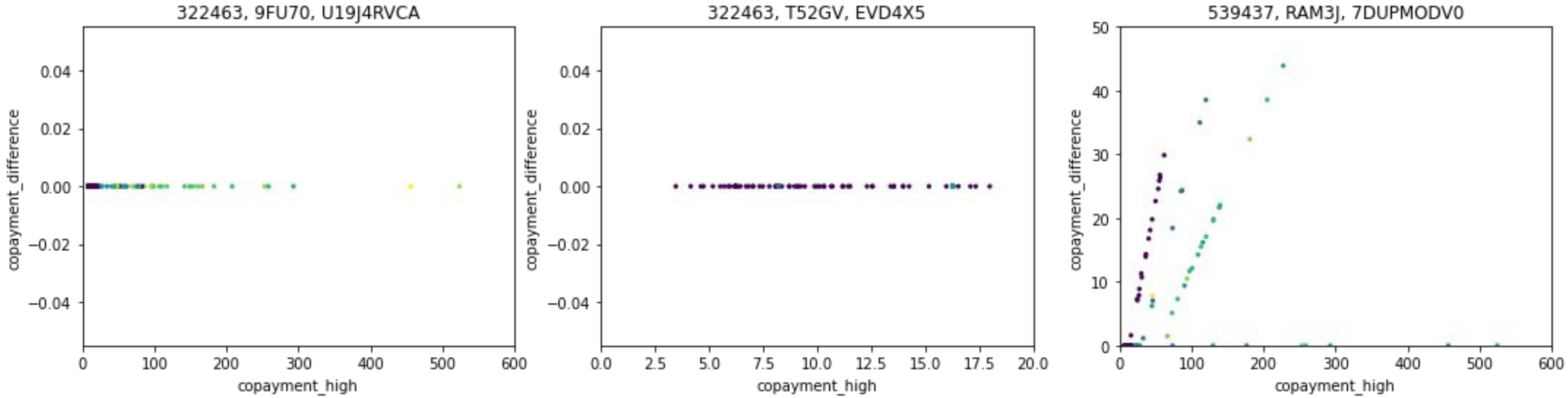






Three Main Types of Insurance Plans

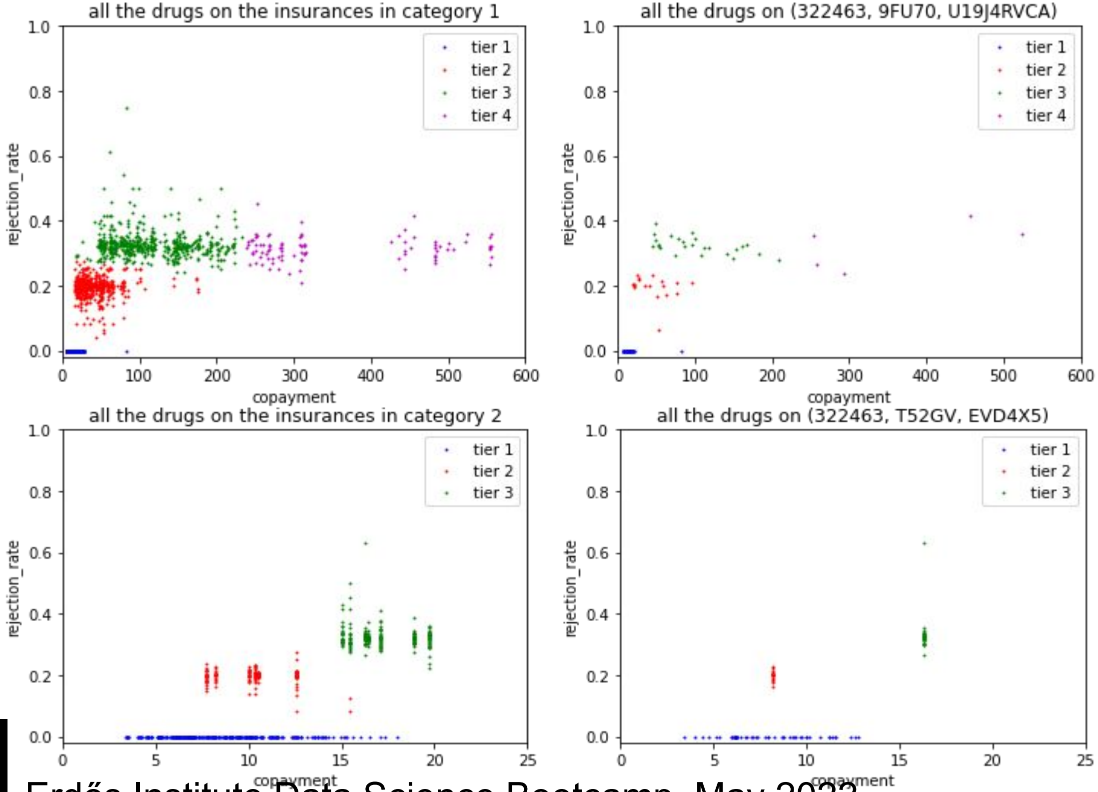
Three types of insurance plans, based on copay/deductible behaviour:



Type 1 (23 plans)	**Type 2** (7 plans)	**Type 3** (33 plans)
No deductible phase	No deductible phase	Deductible & tiered coinsurance
Large range of costs	Small range of costs	Large range of costs

Clustering Drugs – Type 1 & 2 Plans

For **drug tiers** in **Type 1 & 2 plans**, use ***rejection rate*** **&** ***copay***



4 tiers (Kmeans clustering):

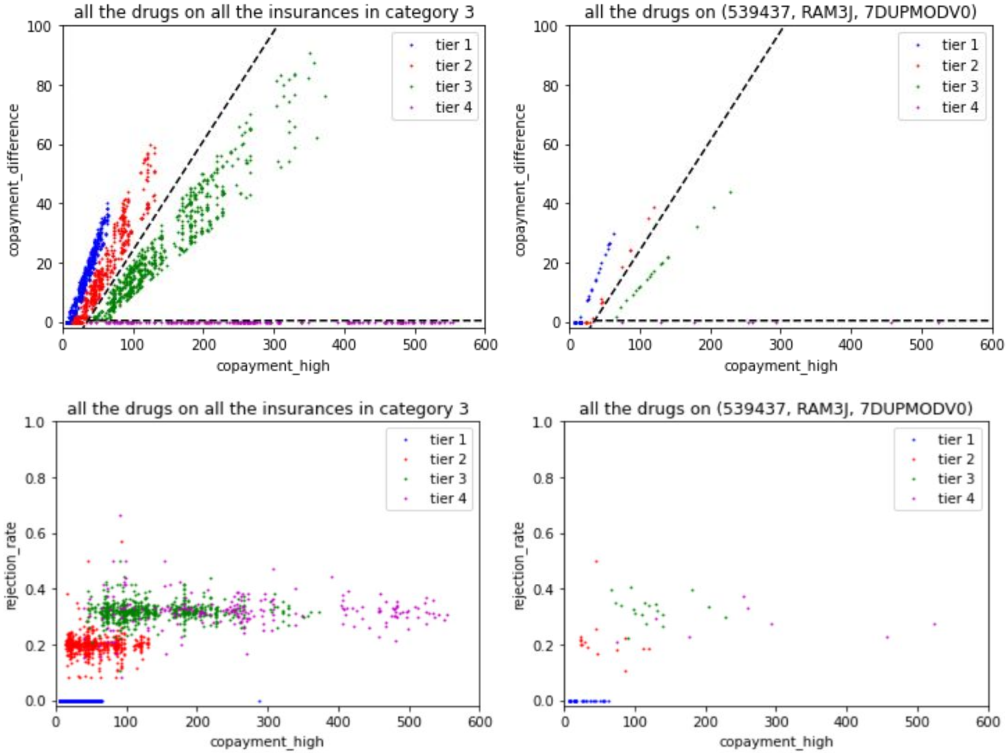
1. Zero rejection rate.
2. About 20% rejection & lower copays.
3. About 30% rejection & higher copays.
4. Significantly higher copays.

3 tiers (Kmeans clustering):

1. Zero rejection rate.
2. About 20% rejection & lower copays.
3. About 30% rejection & higher copays.

Clustering Drugs – Type 3 Plans

Drug tiers of **Type 3 plans** are determined by **deductible** & **coinsurance** (x-intercept & slope of lines in plot).



4 tiers (SVM and Gaussian Mixture):

1. Zero rejection rate.
2. About 20% rejection & lower copays.
3. About 30% rejection & higher copays.
4. Significantly higher copays.

Drugs in higher tiers have higher deductible and lower coinsurance

Further directions

- Make a supplemental model to predict out-of-pocket pay in the deductible phase for the early months of the year.

- Cluster preferred drugs based on the diagnoses they treat and compare how they differ across insurance plans. Help patients choose insurance plans based on their prescription needs.

- Identify preferred in-network and out-of-network pharmacies for each insurance plan.

- Deploy our model in an application that helps patients and doctors compare insurance plans, pharmacies, and medications with cost in mind.

References

1) https://www.uspharmacist.com/article/a-pharmacists-primer-on-prescription-discount-cards
2) https://www.bcbsm.com/medicare/help/understanding-plans/pharmacy-prescription-drugs/tiers.html
3) https://www.goodrx.com/insurance/health-insurance/medication-formulary

? https://www.patientadvocate.org/explore-our-resources/understanding-health-insurance/understanding-drug-tiers/

?https://www.healthinsurance.org/faqs/whats-the-difference-between-prescription-discount-plans-and-prescription-drug-insurance/

https://www.covermymeds.com/main/