CoverMyMeds -- 2021 Bootcamp Project

Company Description

CoverMyMeds, part of McKesson's Prescription Technology Solutions, is a fast-growing healthcare technology company that has been recognized as a “Best Place to Work” by Glassdoor and a “Best Company to Work For” by FORTUNE. CoverMyMeds’ solutions help patients get the medications they need to live healthy lives by seamlessly connecting the healthcare network to improve medication access; thereby increasing speed to therapy and reducing prescription abandonment. CoverMyMeds’ network includes approximately 75 percent of electronic health record systems (EHRs), 96 percent of pharmacies, 750,000 providers and most health plans and PBMs. By facilitating appropriate access to medications, the company can help its customers avoid millions of dollars each year in administrative waste and avoidable medical spending caused by prescription abandonment. Visit www.covermymeds.com for more information.

Description of Project

Background

CoverMyMeds designs solutions that help patients get the medication they need to live healthy lives. When a doctor prescribes a therapy to a patient, they send the prescription to a pharmacy. The pharmacy, when going to fill the prescription, runs a claim against the patient’s insurance (in this context known as a Pharmacy Benefit Manager or PBM), to see if their insurance will cover the therapy as prescribed (correct drug, dosage, etc.). Thankfully, more times than not insurance companies, or payers, approve these claims and the pharmacy dispenses the medication as prescribed. The patient then picks it up and when relevant, pays the remaining cost and completes the transaction.

However, payers may not cover particular medications or dosages, and may reject the claim. Claims can be rejected for a variety of reasons. For example, the dosage or quantity dispensed may not be covered, or the drug might not be on formulary. A formulary is a list of the preferred drugs a payer has. Typically, formularies are tiered, with some drugs being the cheapest, some being more costly but covered, and some requiring a prior authorization for example.

When a claim is rejected, a “reject code” is provided which contextualizes the reason the claim was rejected. For our purposes, there are 3 rejection codes that we focus on: 70, 75, and 76. A code 70 implies that a drug is not covered by the plan and is not on

formulary, and typically implies that another course of therapy should be pursued. A code 75 implies that a drug is on the formulary but does not have preferred status and requires a prior authorization (PA). A PA is a form that care providers submit on behalf of their patient making the case that the therapy as described is a critical course of treatment. A code 76 simply means that the drug is covered, but that the plan limitations have been exceeded, which means that the limitations on the number of fills for that medication has been met. We might expect there to be variation in the type of reject codes we see for certain drugs by the payer.

If a claim is rejected, regardless of the reject code provided, a prior authorization can be started to prevent prescription abandonment and ensure a patient gets the therapy their provider thinks would work best for them. At CoverMyMeds, one of our products is an electronic PA (ePA) solution that replaces a largely manual process of looking for the relevant form, printing it, filling it out, and faxing it to a simple portal-based experience. When a provider is filling out an ePA, they may frequently be asked to provide information about a patient's diagnosis, lab values, contraindications (health-related reasons not to take certain medications), and if they have tried and failed other therapies. When reviewing prior authorizations, payers evaluate the information provided against their formulary and make a decision. That is to say, information contained on the PA and information contained in the original pharmacy claim can help us understand whether an ePA is likely to be approved or denied.

From a financial perspective, ePA volume is a predictor of revenue. We will often forecast future volume (number of ePAs a day/month/year) using time series analysis to help with budgeting. From a calendar perspective, we often see PA volume at its highest during workdays, and its lowest during holiday seasons. We also see peaks in volume around the start of the year, which is when many PAs expire and new ones will be resubmitted.

Application

From an analytics perspective, much can be learned with this data. In claims data, we can try to discover the formulary status of certain drugs for certain payers, which is very valuable data in this market. We can also try to predict whether a PA will be needed at the point of prescribing to prevent a patient from going to the pharmacy, only to return home empty handed. We can predict whether a PA will be successful based upon claims data, or data included on the ePA, to prevent patients from waiting on a decision, only to end up having to switch to another therapy if the PA is rejected. We can also forecast future monthly PA volume as a function of historical volume to generate more accurate budgets.

This is where you come in! We want you to take our data and do something exciting! You can use some of the examples previously listed to show off your skills in data analytics, predictive modeling, or time series analysis. This might include:

- Trying to make sense of the formulary for each payer from pharmacy claims data.
- Predicting whether a PA is going to be needed based upon pharmacy claims data.
- Predicting whether a PA is likely to be approved or denied.
- Analyzing which factors tend to influence whether PAs are likely to be approved or denied.
- Forecasting future PA volume at the monthly level using historical volume.

Ultimately, however, this is meant to be a "choose your own adventure" type exercise! You may answer as many or as few of those questions as you might like (paying attention to breadth v. depth) or ask and answer a series of questions that you might have that we haven't covered here! For example, you might want to focus on implementing a model, that's cool with us!

When presenting, your presentation should be geared towards an audience of business stakeholders and fellow data scientists. These business stakeholders will have some understanding of the data, but not necessarily much understanding of data science-specific concepts (precision v. recall, boosting v. bagging, etc.).

Data science professional "domain(s)"

- *Data Engineering* (e.g., database management, data warehousing)
- *Software Engineering* (e.g., software development, cross-platform software)
- *Machine Learning (ML) Engineering* (e.g., executing existing ML/AI algorithms)
- *Data Product* (e.g., ML/AI for a specific product, such as a smartphone app, app development and productionalizing)
- *Business Intelligence* (e.g., data science to understand the efficiency, profitability, and productivity of your company)

Description of Data

You have been provided dimension tables (claims, PAs, dates) stored in CSV files which link through a bridge table. You may process and analyze this data using your preferred language, although at CoverMyMeds we tend to use Python and R (in that order). Please feel free to use whatever tools and online resources you need to accomplish this task. We do want you to get all the help you need from the CoverMyMeds team, so asking tons of questions is highly encouraged.

Here is a brief data dictionary:

- dim_date.csv
  - This is a table containing date-level information that might be useful for producing transaction forecasts.
  - It contains the following columns:
    - dim_date_id: Primary key and index. Integer.
    - date_val: Date value corresponding to the record. Date.
    - calendar_year: Year for the record. Integer.
    - calendar_month: Month of the year for the record. Integer.
    - calendar_day: Day of the month for the record. Integer.
    - day_of_week: The Nth day of the week. Integer.
    - is_weekday: Flag for if the record is associated with Monday through Friday. Bit.
    - is_workday: Flag for if the record is associated with Monday through Friday and is not a holiday. Bit.
    - is_holiday: Flag for if the record is associated with a holiday. Bit.
- dim_claims.csv
  - This is a table containing pharmacy claim-level information.
  - It contains the following columns:
    - dim_claim_id: Primary key and index. Integer.
    - bin: The BIN of the payer for the claim. Numeric.
    - drug: The drug that was associated with the claim. Character. A, B, C
    - reject_code: If the claim was rejected, what was the associated rejection code. Character. 70, 75, 76, na
    - pharmacy_claim_approved: Flag for if the claim was approved by the payer. Bit.
- dim_pa.csv
  - This is a table containing PA-level for every rejected pharmacy claim in dim_claims.
  - It contains the following columns:
    - dim_pa_id: Primary key and index. Integer.
    - correct_diagnosis: Flag for information provided by the provider indicating that the patient has the correct diagnosis for the associated drug. Bit.
    - tried_and_failed: Flag for information provided by the provider indicating that the patient has tried and failed the relevant generic alternatives. Bit.

        - contraindication: Flag for information provided by the provider indicating that the patient has an associated contraindication for the medication requested. Bit.
        - pa_approved: Flag for if the payer favorably reviewed and approved the PA. Bit.
- bridge.csv
    - This is a bridge table that links the primary keys of all the tables to one another that is used for joining the tables.
    - It contains the following columns:
        - dim_claim_id: Primary key for dim_claims.
        - dim_pa_id: Primary key for dim_pa.
        - dim_date_id: Primary key for dim_date.