



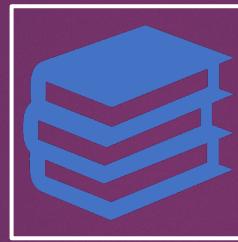
2016 MN Medicaid Prescription Data

OMAR ABDEL-KEREM, DATA SCIENCE, MS
NEUROSCIENCE, BS

Why Prescription Data?



I have been working in the pharmaceutical industry for the past 4 years



My job requires me to utilize my background in Neuroscience



This dataset includes information regarding medications that I interact with on a daily basis

Why Minnesota Medicaid Data?

- ▶ Gives insight on the common health issues that individuals who utilize Medicaid face
- ▶ Allows us to identify ways to help improve our pharmaceutical and healthcare systems for our community

Total Number of Prescriptions Filled in 2016

Cardiovascular Agents 12,301,511	Endocrine and Metabolic Drugs 6,353,762	Gastrointestinal Products 3,498,500	Respiratory Agents 3,408,997
	Analgesics and Anesthetics 4,744,143	Neuromuscular Drugs 3,239,925	Topical Products 2,963,496
Central Nervous System Drugs 8,045,808		Nutritional Products 1,446,193	
	Anti-Infective Agents 3,949,090	Hematological Agents 1,331,460	Miscellaneous Products 1,116,348
			Misc

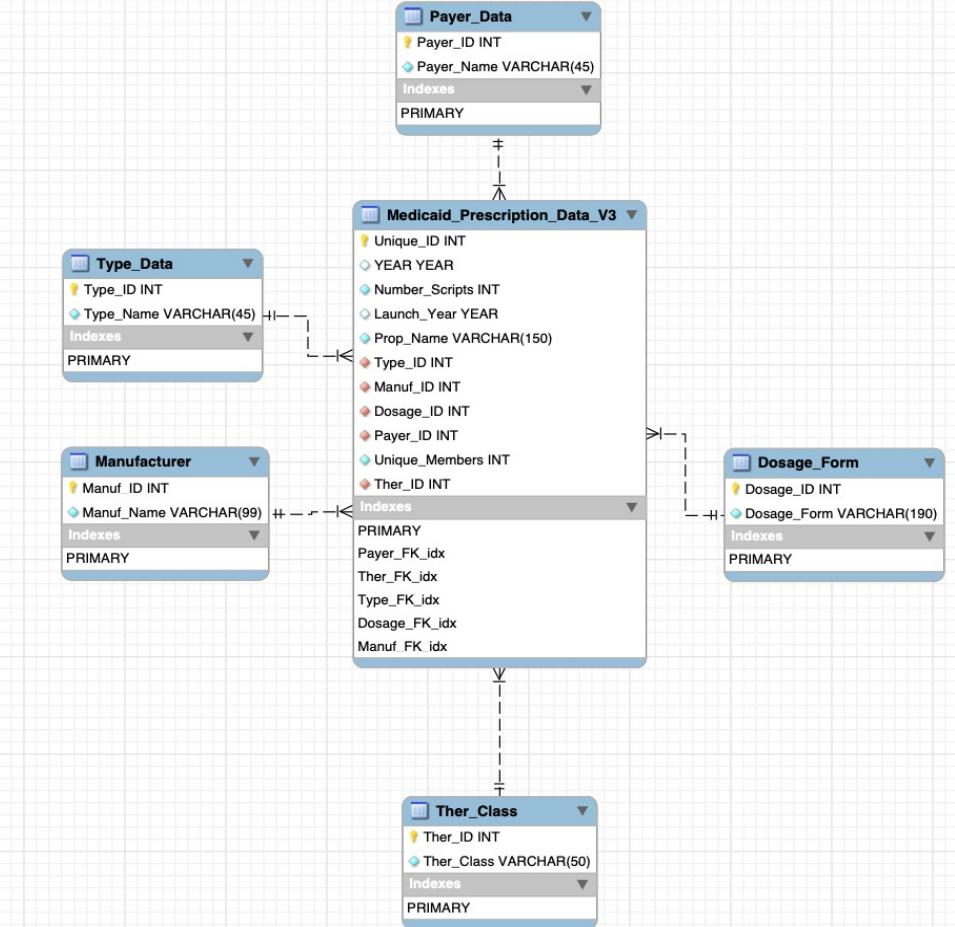
A Bit of Background About The Dataset

Unique_ID	YEAR	Number_Scri...	Launch_Year	Prop_Name	Type_ID	Ma...	Dosage_ID	Payer_ID	Unique_Me...	Ther_ID
637	2016	199642	2006	VENTOLIN	2	102	33	1	122975	10
686	2016	213171	2006	VENTOLIN	2	102	33	2	91757	10
2763	2016	98438	2005	AZITHROMYCIN	1	228	8	1	86112	4
661	2016	113446	2006	VENTOLIN	2	102	33	3	52586	10
1587	2016	65637	2010	AMOXICILLIN	1	66	37	1	52323	4

- Obtained from MN Department of Health
- Dataset contains ~30,000 rows
- Contains name of medication, number of prescriptions, number of unique members, and other information
- I used another dataset to pair each medication with their appropriate therapeutic class

My Project

- ▶ Utilized Data Modeler to create a logical and relational model
- ▶ I created a relational database using MySQL Server and MySQL Workbench
- ▶ Normalized my data before importing into database
- ▶ Created a parent table and 5 child tables



Topics that were incorporated

1. Data Preparation & Normalization
2. Data Definition Language (DDL)
3. Cardinality
4. Data Manipulation Language (DML)
5. Reverse Engineering

Informal Topics

1. Finding the right DB software and tools
2. Getting feedback and rebuilding

Challenges



Even though SQL and MySQL syntax are similar, however, they are a bit different



Working with large datasets, pinpointing issues with importing are a bit challenging but doable