

Predictive Modeling of Opioid Prescription Fraud by CMS Providers

DATA606 Capstone Project - Part 2
William Rubin – Spring 2020

PROJECT TIMELINE

Part 1

Problem Definition

Predicting Opioid-Related
Prescription Fraud by CMS
Providers

Part 2

Existing Research / Initial EDA

Part 3

Final EDA / Initial Predictive Modeling

Part 4

Execution, Interpretation, and Results

UPDATES

Initial Identified CMS Data Source

- Detailed Data File
- Contains 24 columns with one record per prescription drug event
- Approx. 25M records per file (14.6GB)

Modified CMS Data Source

- Prescriber Summary Table
- Contains 84 columns with one record per registered provider
- Approx. 1M records per file (1.6GB)

Same CMS System origination with inclusion of additional provider-level aggregate detail and more manageable data volume

Part 2 Summary of Tasks

- Existing Research Efforts
 - HHS OIG Data Brief - 2017: Opioids in Medicare Part D: Concerns about Extreme Use and Questionable Prescribing
 - Medicare Part D Opioid Prescribing Mapping Tool
- Prepare CMS Data
 - Load Part D Prescriber Summary Tables from 2013-2017
 - Evaluate Available Data Fields / Categories of Data
- Perform Initial Exploratory Data Analysis
 - High-Level Review of Data
 - Analyze Trends Based on Data Categories

Existing Research Efforts

Opioids in Medicare Part D: Concerns about Extreme Use and Questionable Prescribing (*HHS OIG Data Brief – July 2017*)

- Evaluated prescription drug events records from 2016
 - Analysis of opioid dispensation and spending (utilization) across all beneficiaries
 - Individual beneficiary usage accounting for prescription strength and treatment duration
 - Identification of providers which exhibited extreme prescription patterns
- Key Results
 - 1/3 of Medicare Part D beneficiaries received at least one opioid prescription
 - Approximately 1/2 million beneficiaries received large opioid amounts
 - Around 400 providers exhibited potentially questionable prescription practices

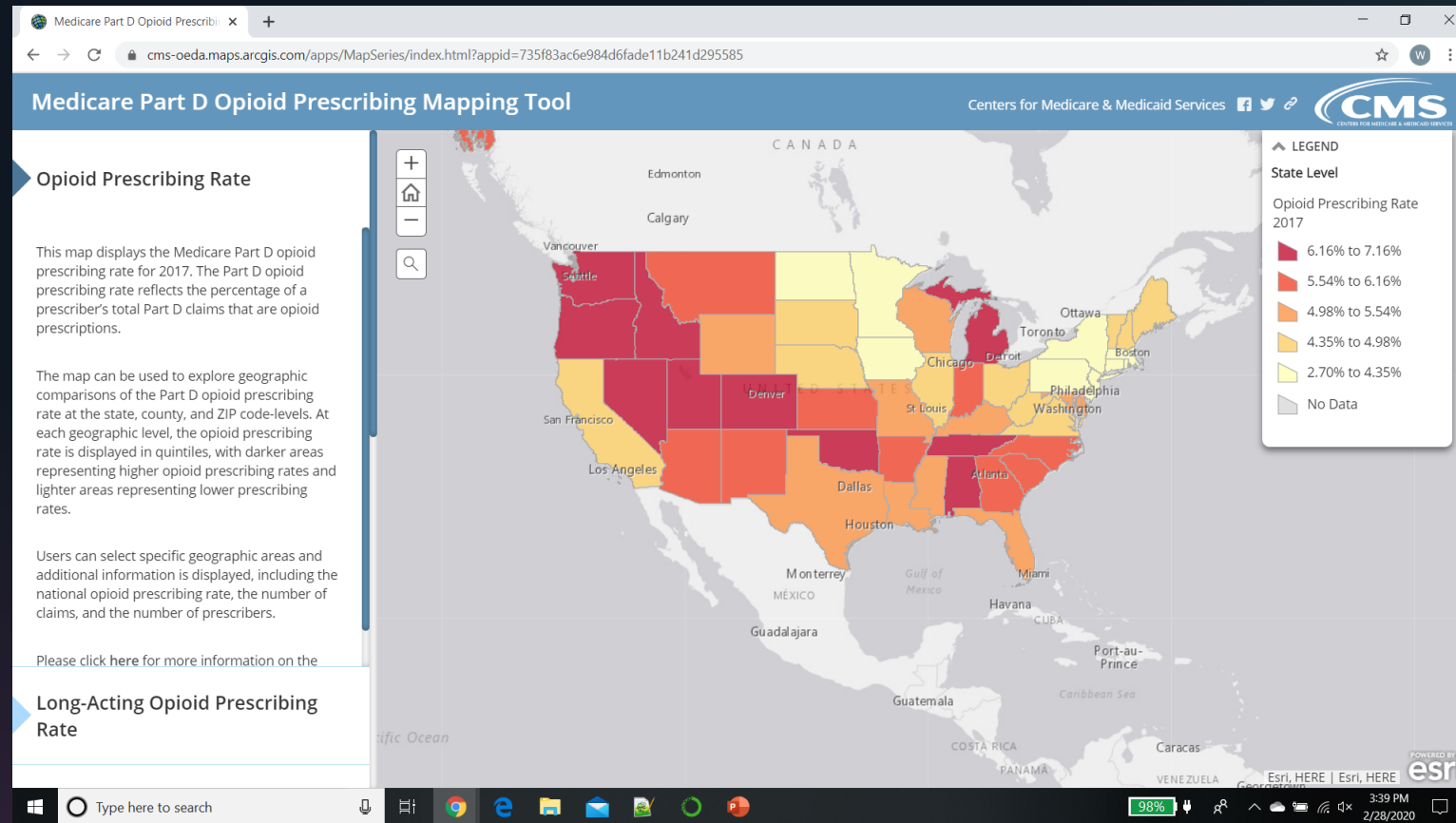
Existing Research Efforts

Medicare Part D Opioid Prescribing Mapping Tool

- Enables national-level scale comparisons of Medicare Part D opioid prescriptions
- Displays opioid and long-acting opioid claim totals and prescribing rates from 2013 through 2017
- Assists in visualization of both state and county level trends based upon percentage changes

Existing Research Efforts

Medicare Part D Opioid Prescribing Mapping Tool



Differentiator

- This projects' goal would expand upon scope of existing research
- Analysis based on known cases of provider opioid prescription related fraud
- Establish a predictive capability to actively identify suspect providers

Data Overview

Part D Prescriber Summary Tables

- 5 Source Extract Files (1 file per calendar year, 2013 – 2017)
- 1 Record per Registered Provider (NPI) per Calendar Year
 - Total Records: 5,518,978
- 84 Included Columns
 - 4 Categories of Data Columns
 - Provider Registration
 - Applicable Medicare Program
 - Prescription Drug Utilization
 - Beneficiary Demographics / Health Characteristics
 - Column Added: Calendar Year (from filename)

Data Overview

Provider Registration

- NPI
- Full / Organization Name
- Address
- Medical Specialty

Applicable Medicare Program

- MAPD
- PDP

Prescription Drug Utilization

- Total Claims, Benes, Drug Costs
- Drug Type / Classifications
- Low Income Subsidy

Beneficiary Demographics / Health Characteristics

- Age Range
- Gender
- Ethnicity
- Average Risk Score

Data Limitations

CMS Part D Prescriber Summary Tables:

- Analysis limited to only providers serving individuals which are enrolled in Medicare Part D
- Data only publicly available for calendar years 2013 through 2017
- High opioid / long-acting opioid prescription patterns do not take into consideration quality of care or medical justification
- Aggregate provider records based on 10 or fewer claims are excluded

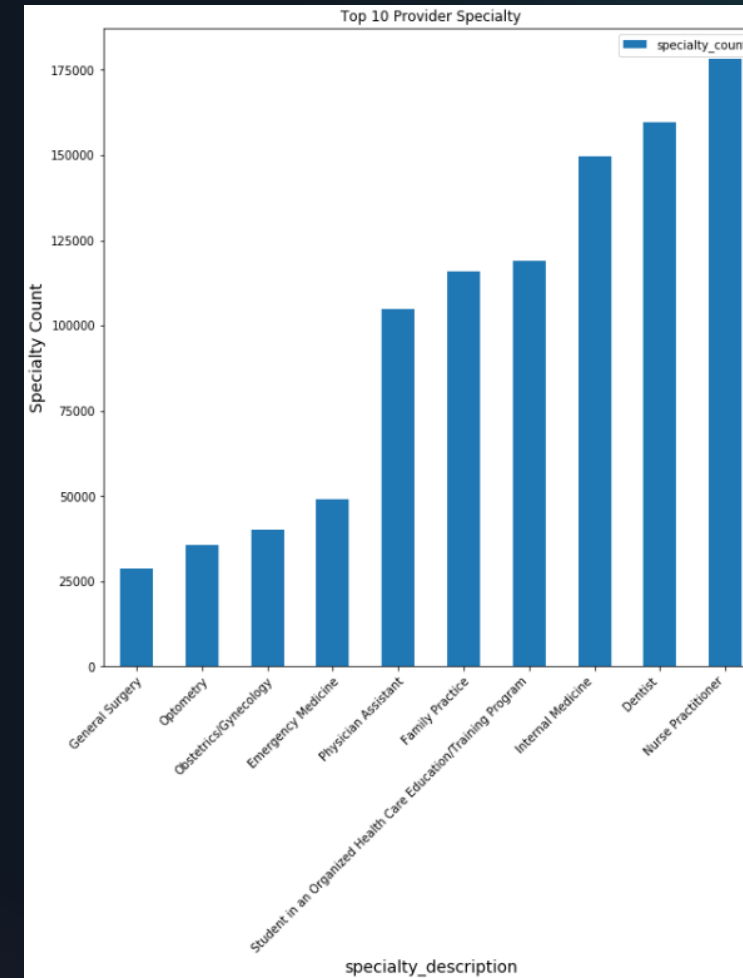
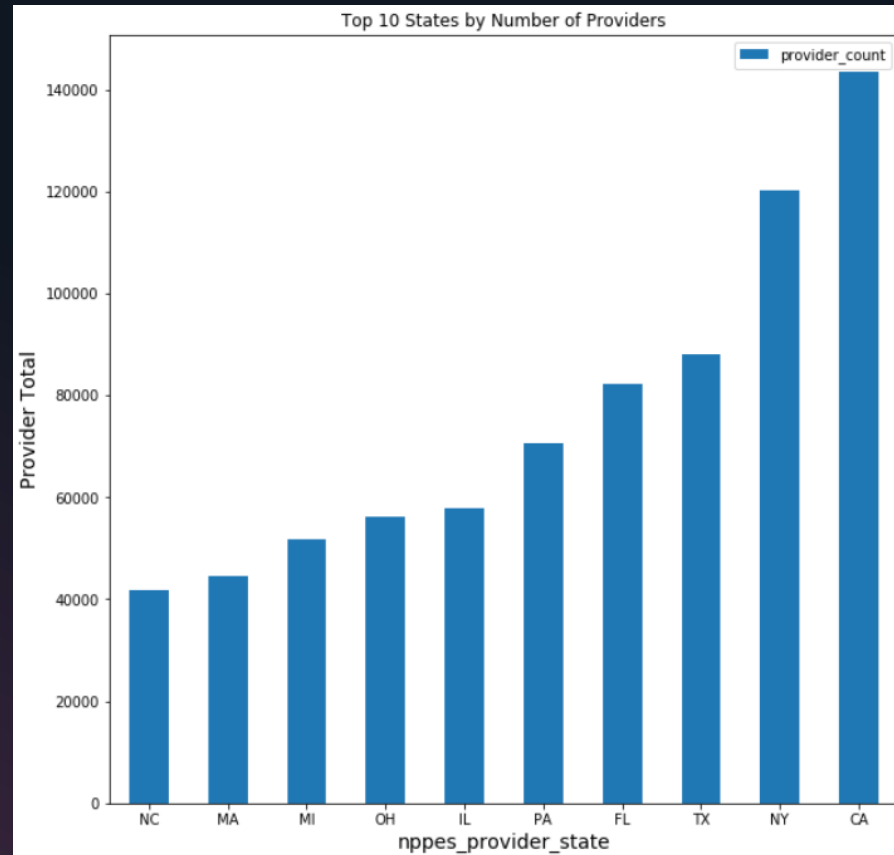
Exploratory Data Analysis

Conducted five main initial analyses:

- Overall Provider Characteristics (Not Opioid Specific)
- Opioid Specific Provider Characteristics
- High-level Opioid-Specific Claims / Beneficiary Analysis
- Opioid and Long-acting Opioid Prescriber Rate Histogram
- Opioid and Long-acting Opioid Prescriber Cost Analysis

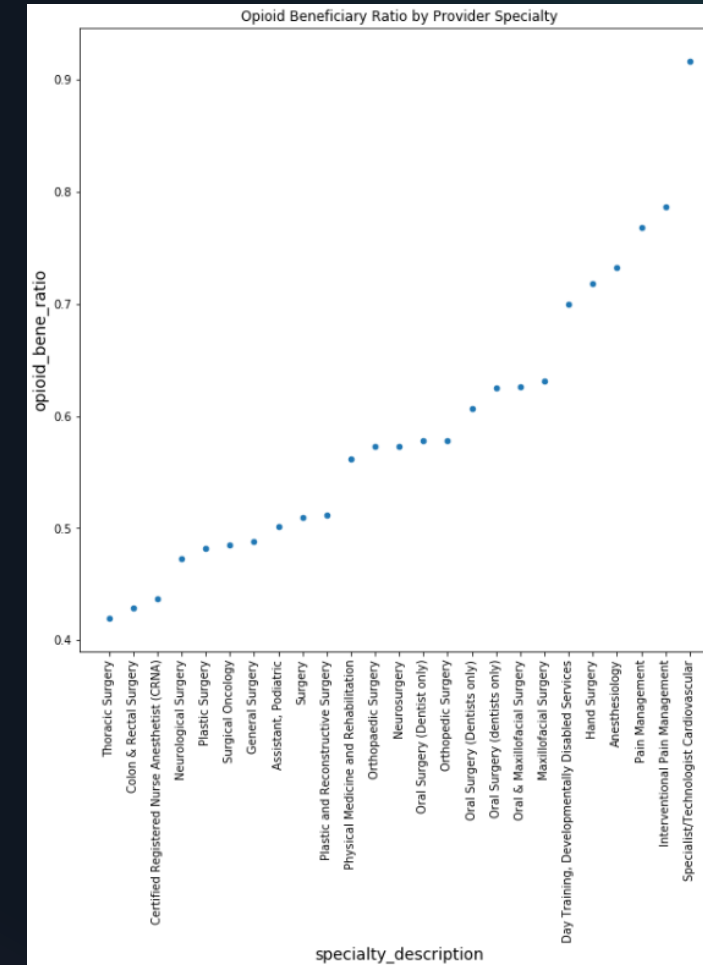
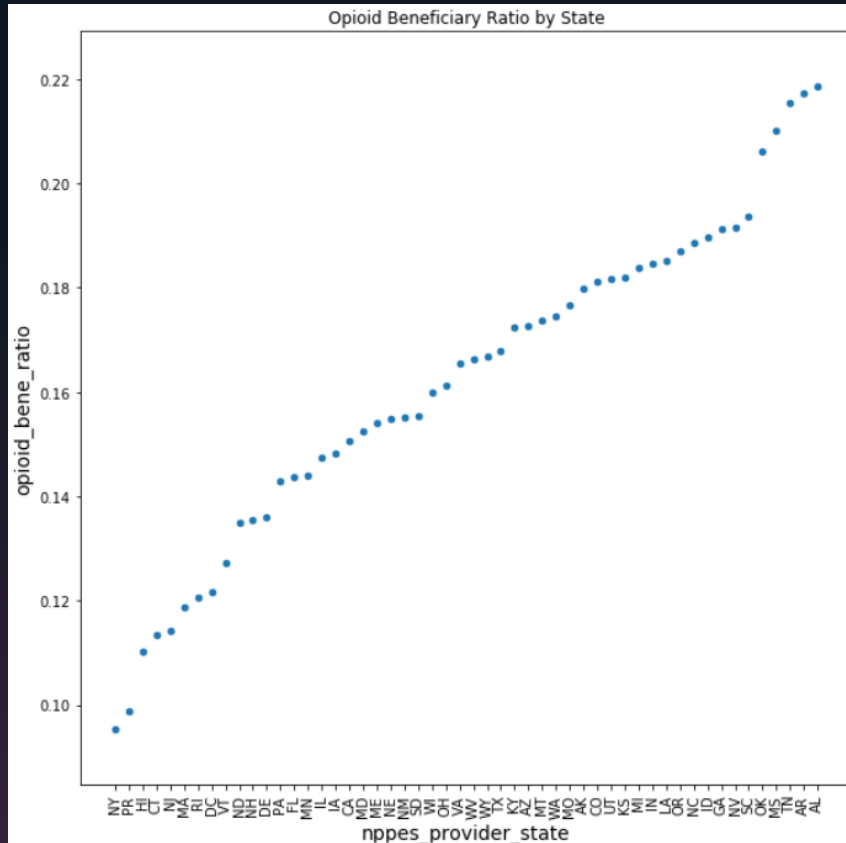
Exploratory Data Analysis

Overall Provider Characteristics (Not Opioid Specific)



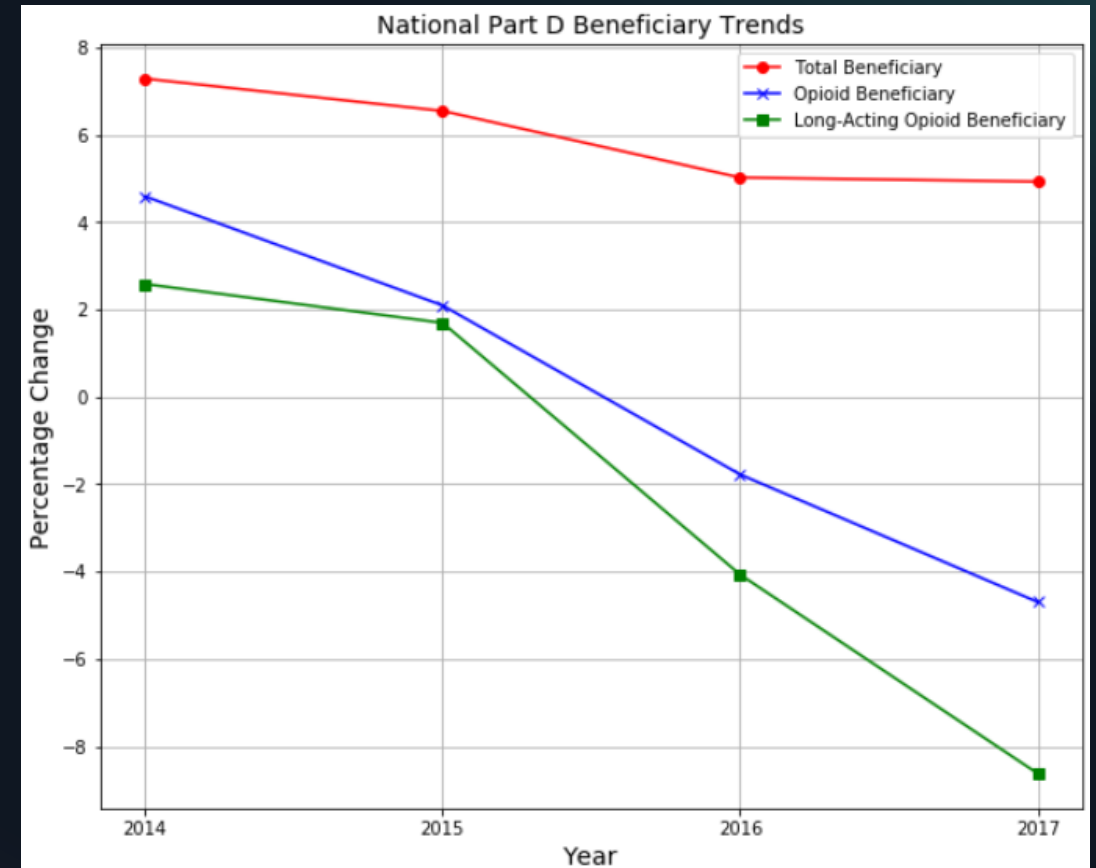
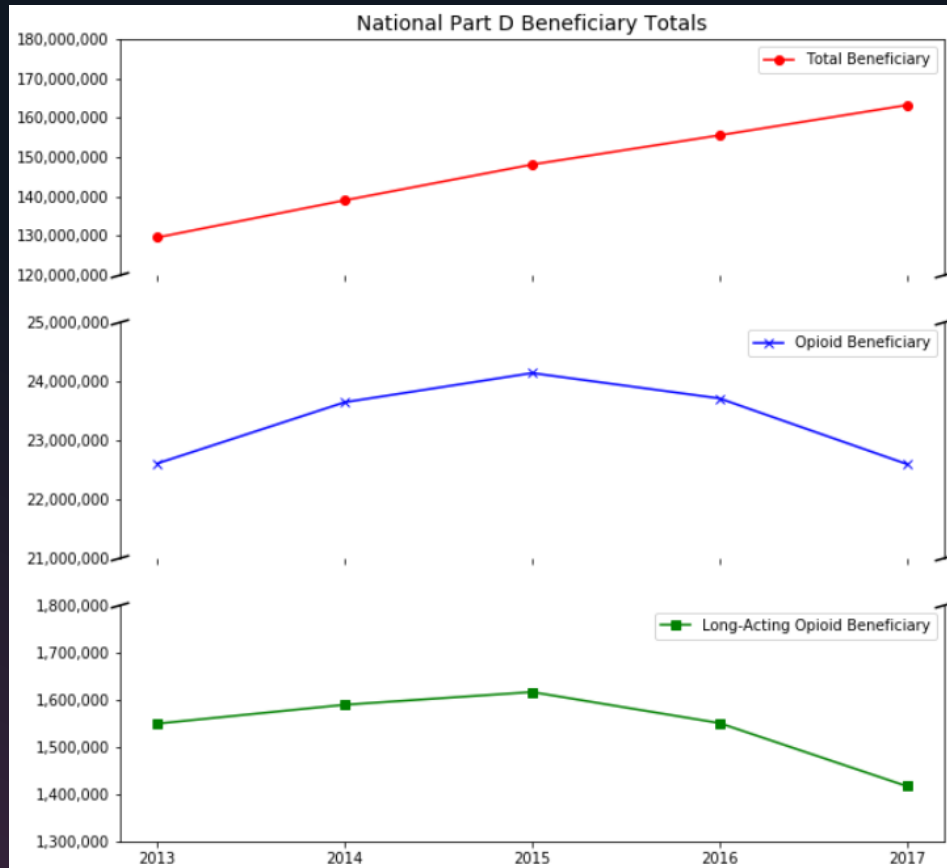
Exploratory Data Analysis

Opioid Specific Provider Characteristics



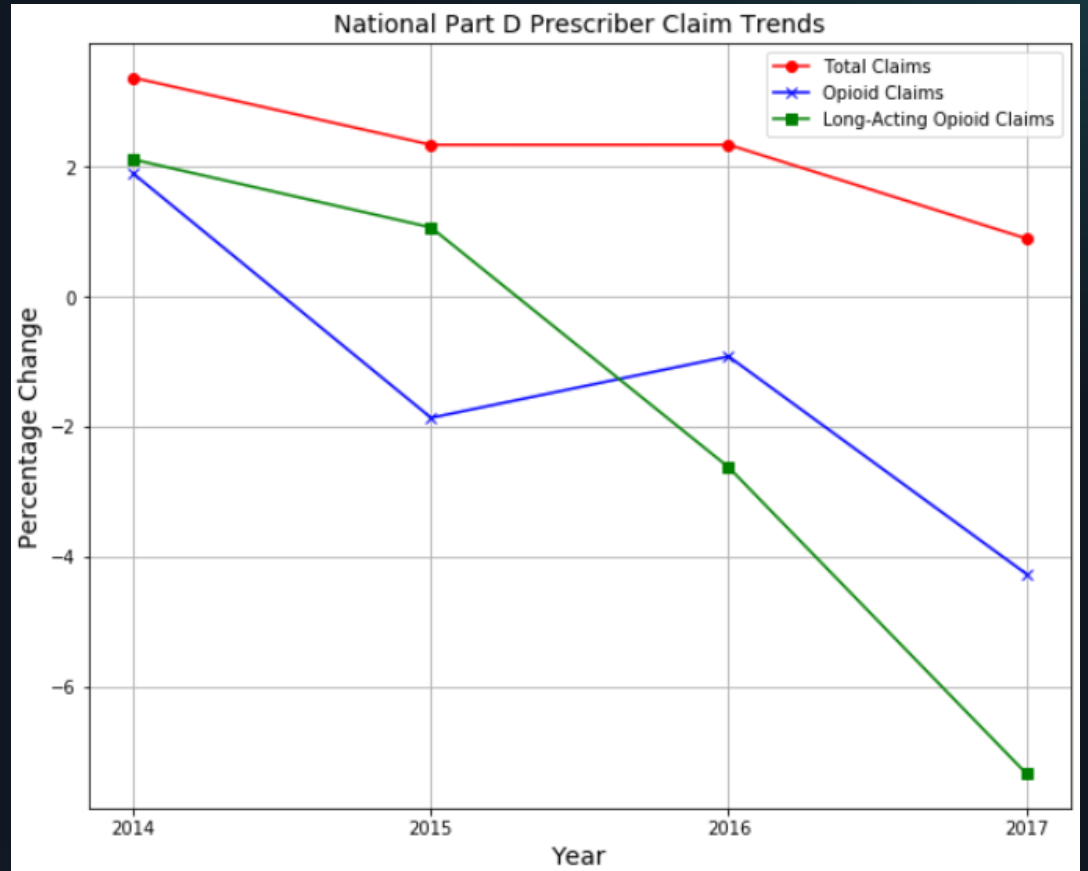
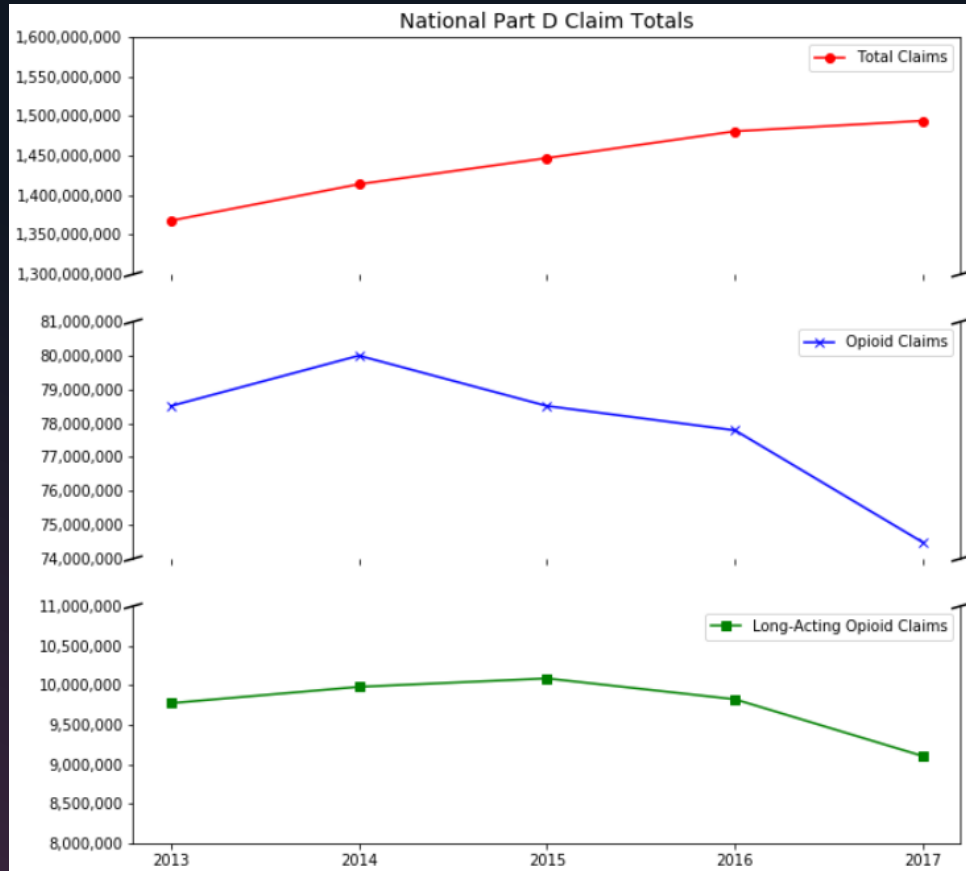
Exploratory Data Analysis

High-level Opioid-Specific Beneficiary Analysis



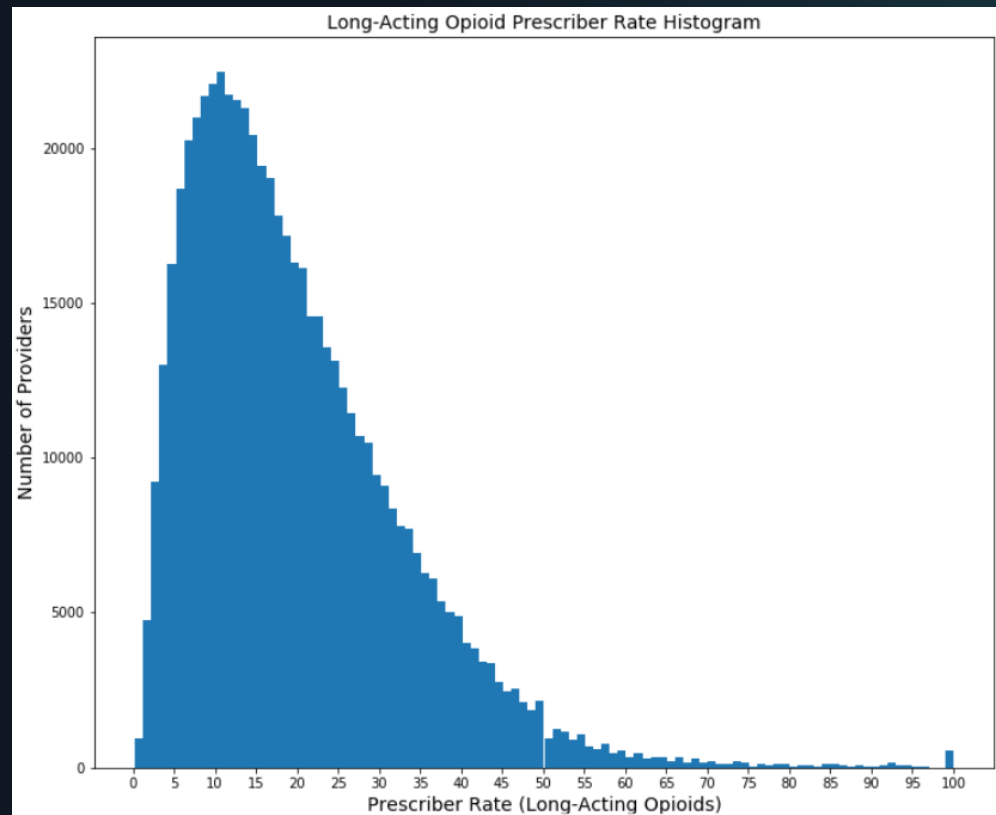
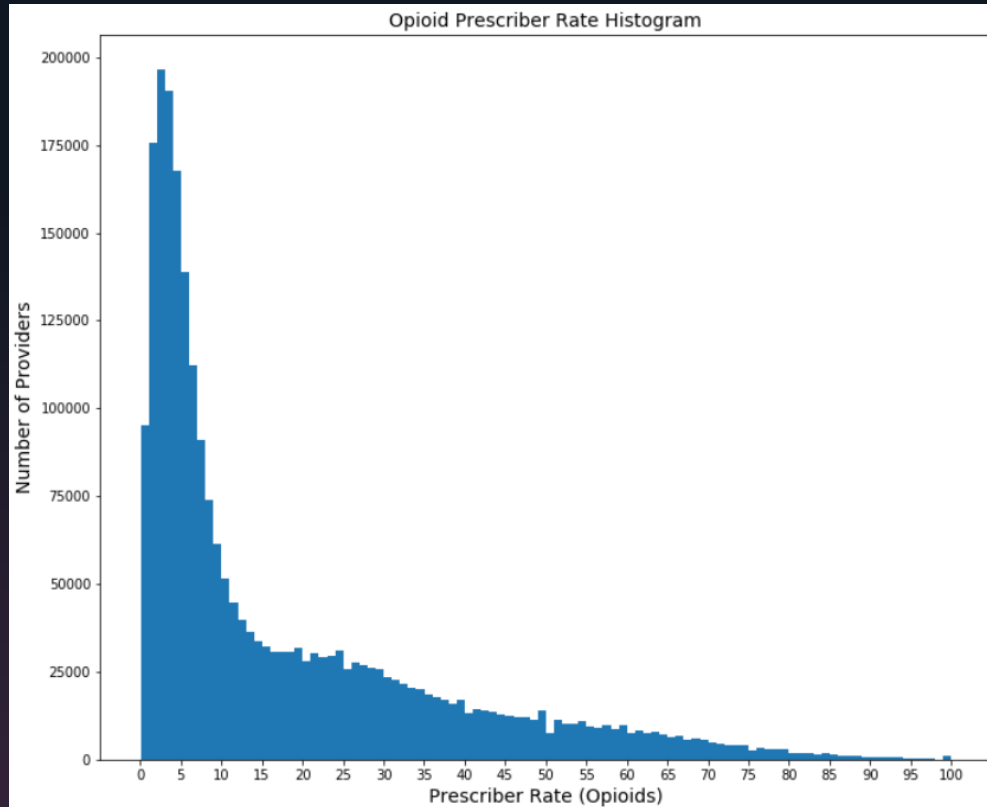
Exploratory Data Analysis

High-level Opioid-Specific Claims Analysis



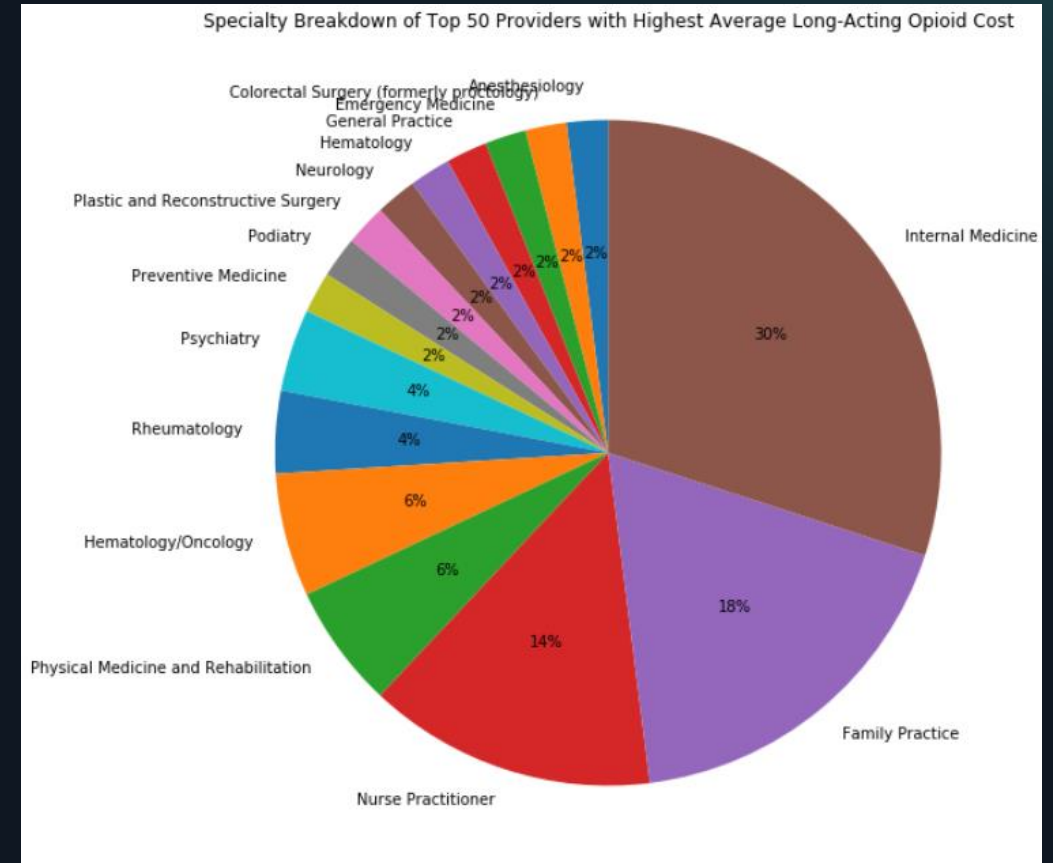
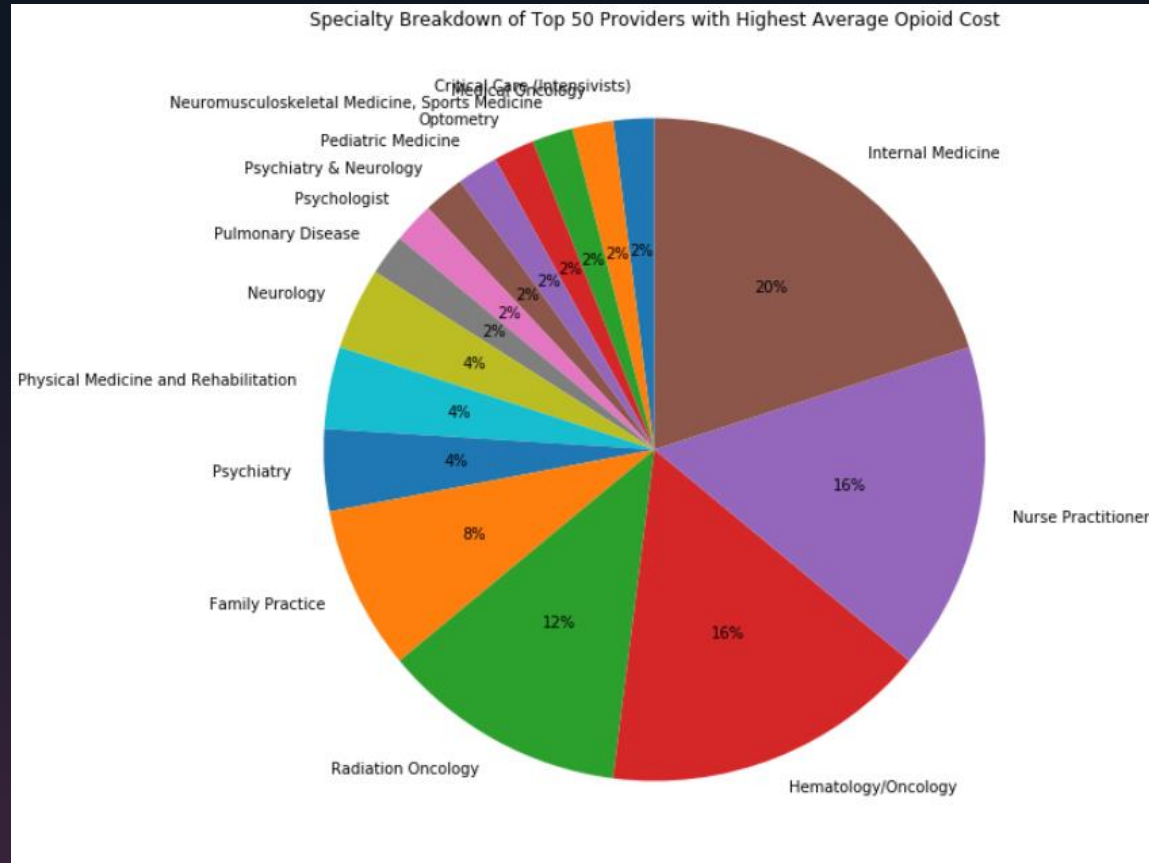
Exploratory Data Analysis

Opioid and Long-acting Opioid Prescriber Rate Histogram



Exploratory Data Analysis

Opioid and Long-acting Opioid Prescriber Cost Analysis



Next Steps

Overview of Planned Part 3 Tasks:

- Finalize EDA
- Perform Data Mining of DOJ Press Releases
- Construct Fraudulent Provider Dataset
- Build Initial Prediction Model

THANK YOU!

William Rubin

Email:

wrubin1@umbc.edu

GitHub:

<https://github.com/warubin410/DATA606Spring2020>

Google Site:

<https://sites.google.com/umbc.edu/data606/home/willie-rubin>