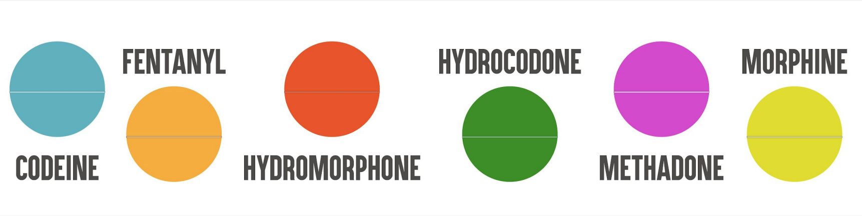
Opioids: Drivers and Anomalous Prescribing Patterns

Presented by: Steve Barry, Neeraj Tadur, Derek Wong

12/4/2018

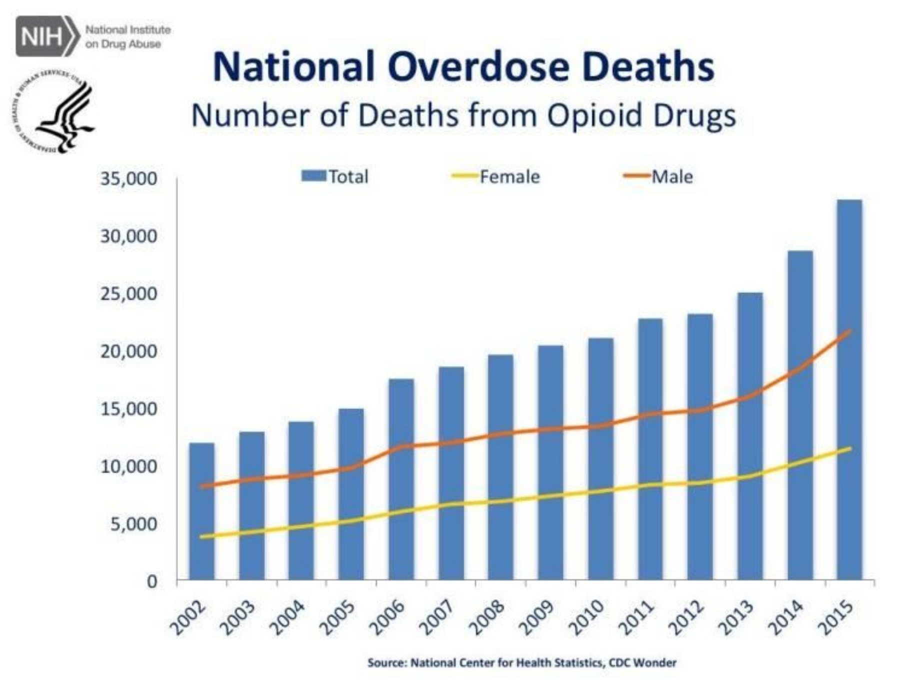
Opioids: What are they?

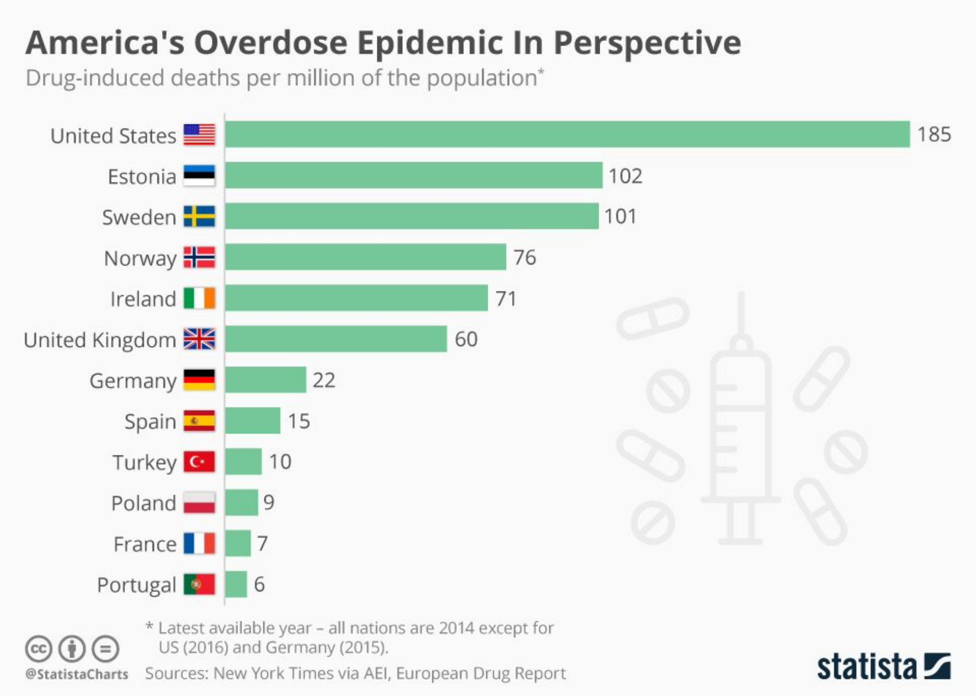


- Pain reliever drugs
- Physician prescribed and also illegal street drugs
- Varying kinds, varying strengths
- Some opioids have high potential for dependence and overdose
- “Man-made opioid epidemic”

Source: https://www.naabt.org/faq_answers.cfm?ID=4

Why is this important?





Data Used

- Original Source: Centers for Medicare and Medicaid Services (CMS)
- Sampled from Kaggle
  - https://www.kaggle.com/apryor6/us-opiate-prescriptions/home
- Prescriber Level (25,000 x 256)
  - 25,000 licensed medical professionals
  - By Gender, Specialty, and Credentials
  - 250 common non-opioid and opioid drugs - what they are prescribing, and how often
- State Level (50 x 4)
  - Aggregate population and overdose death rate statistics
- Drug Level (113 x 2)
  - Drug names, Generic

Data Used

```
od_df.head()
```

	State	Population	Deaths	Abbrev
0	Alabama	4,833,722	723	AL
1	Alaska	735,132	124	AK
2	Arizona	6,626,624	1,211	AZ
3	Arkansas	2,959,373	356	AR
4	California	38,332,521	4,521	CA

```
opioids_df.head()
```

	Drug Name	Generic Name
0	ABSTRAL	FENTANYL CITRATE
1	ACETAMINOPHEN-CODEINE	ACETAMINOPHEN WITH CODEINE
2	ACTIQ	FENTANYL CITRATE
3	ASCOMP WITH CODEINE	CODEINE/BUTALBITAL/ASA/CAFFEIN
4	ASPIRIN-CAFFEINE-DIHYDROCODEIN	DIHYDROCODEINE/ASPIRIN/CAFFEIN

```
presciber_df.head()
```

	NPI	Gender	State	Credentials	Specialty	ABILIFY	ACETAMINOPHEN.CODEINE	ACYCLOVIR	ADVAIR.DISKUS	AGGRENOX	...	VERAPAI
0	1710982582	M	TX	DDS	Dentist	0	0	0	0	0	...	
1	1245278100	F	AL	MD	General Surgery	0	0	0	0	0	...	
2	1427182161	F	NY	M.D.	General Practice	0	0	0	0	0	...	
3	1669567541	M	AZ	MD	Internal Medicine	0	43	0	0	0	...	
4	1679650949	M	NV	M.D.	Hematology/Oncology	0	0	0	0	0	...	

Sections Today

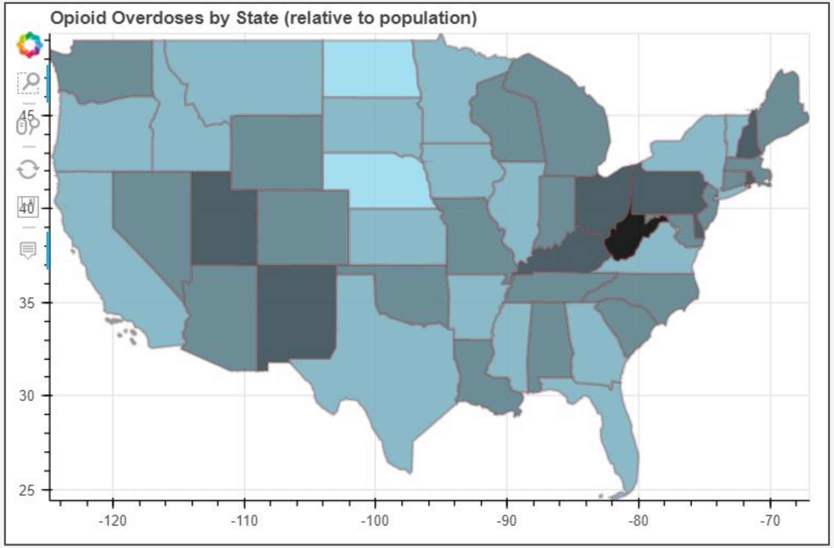
1. Exploration of the Data
2. Making Predictions
3. Identifying Prescriber Outliers through Cluster Analysis
4. Wrap-Up

Making Predictions

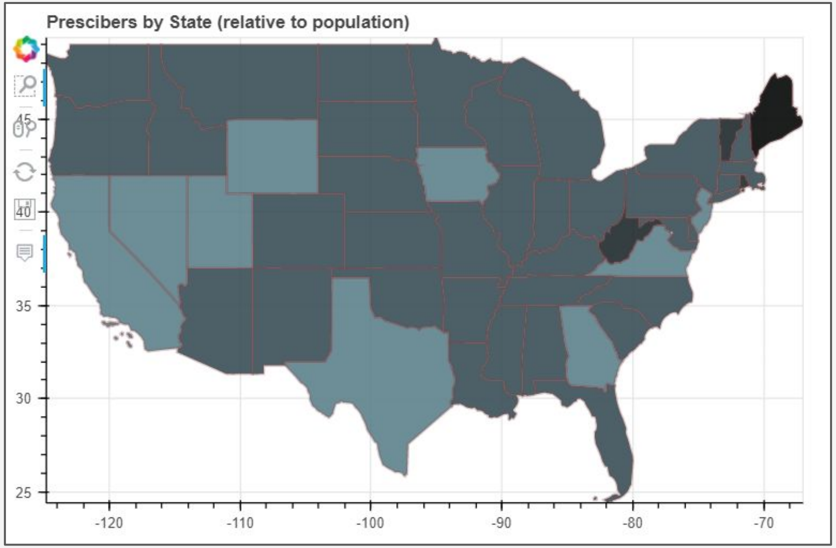
- Exploring the data
  - Geographic
    - Deaths from Opioids by State
    - Prescribers of Opioids by State
  - Demographic
    - Rates of Prescribers by Gender

- Predicting Opioid Prescribers
  - Using Categorical Supervised Learning
  - Preparing the data
  - Feature Selection
  - Which models worked the best
  - Using different features and filters

Overdoses vs Prescribers By State



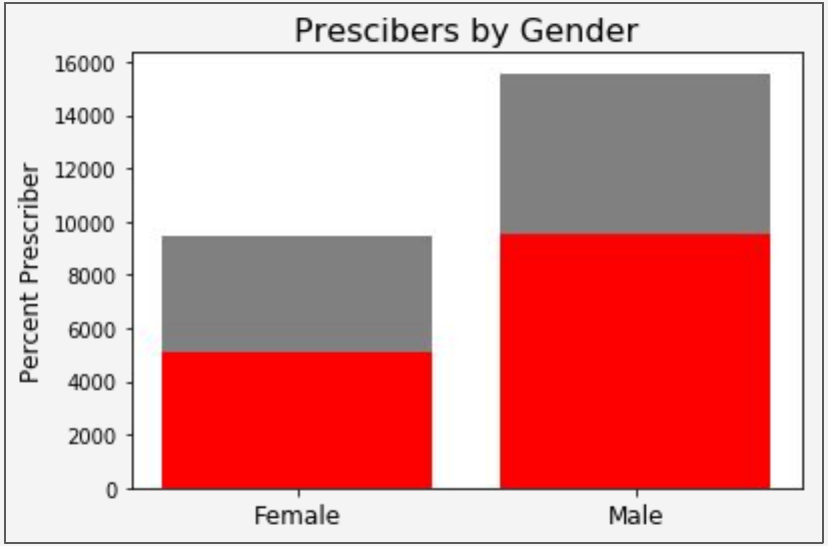
Overdose Rate by State per Capita



Opioid Prescriber Rate by State per Capita

- ❏ Prescriber rates by state show little variance, while overdose rates have higher variance
- ❏ There appears to be some weak regional correlation between prescribers rates and ODs
- ❏ Our data is limited in that we have only 25,000 prescribers and 256 drugs we are looking at

Prescribers by Gender



Opioid Prescriber Rate

	Gender	Total.Prescriber	Opioid.Prescriber	Percent.Prescibers
0	F	9426	5135	54.477
1	M	15574	9553	61.339

- ❏ Male prescribers using this dataset do prescribe opioids at a higher rate than Female prescribers
- ❏ The type of practitioner might greatly influence this data
- ❏ Our data is limited in that we have only 25,000 prescribers and 256 drugs we are looking at

Making Predictions with Supervised Learning

Preparing and Choosing the data:

- Converting categorical fields to numeric
- Dropping uninformative data (NPI, Credentials)
- Tried predictions without all drugs (speciality/state)
- Predictions with all drugs
- Predictions without opioid drugs
- Predictions by specialty w/out opioid drugs

Categorical methods attempted:

- Logistic Regression
- K Nearest Neighbours
- Classification Tree
- Random Forest
- Bagging and Boosting

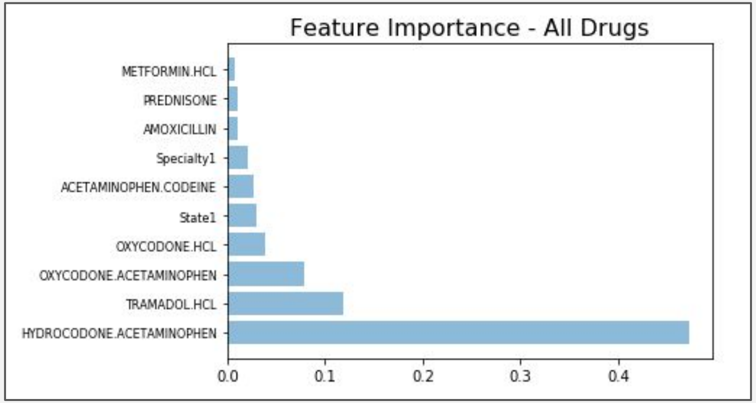
NPI	int64
Gender	object
State	object
Credentials	object
Specialty	object
ABILIFY	int64
ACETAMINOPHEN.CODEINE	int64
ACYCLOVIR	int64
ADVAIR.DISKUS	int64
AGGRENOX	int64
ALENDRONATE.SODIUM	int64
ALLOPURINOL	int64
ALPRAZOLAM	int64
AMIODARONE.HCL	int64
AMITRIPTYLINE.HCL	int64
AMLODIPINE.BESYLATE	int64
AMLODIPINE.BESYLATE.BENAZEPRIL	int64
AMOXICILLIN	int64
AMOX.TR.POTASSIUM.CLAVULANATE	int64
AMPHETAMINE.SALT.COMBO	int64
ATENOLOL	int64
ATORVASTATIN.CALCIUM	int64
AVODART	int64
AZITHROMYCIN	int64
BACLOFEN	int64
BD.ULTRA.FINE.PEN.NEEDLE	int64
BENAZEPRIL.HCL	int64
BENICAR	int64
BENICAR.HCT	int64
BENZTROPINE.MESYLATE	int64
...	
TIMOLOL.MALEATE	int64
TIZANIDINE.HCL	int64
TOLTERODINE.TARTRATE.ER	int64
TOPIRAMATE	int64
TOPROL.XL	int64
TORSEMIDE	int64
TRAMADOL.HCL	int64
TRAVATAN.Z	int64
TRAZODONE.HCL	int64
TRIAMCINOLONE.ACETONIDE	int64
TRIAMTERENE.HYDROCHLOROTHIAZID	int64
VALACYCLOVIR	int64
VALSARTAN	int64
VALSARTAN.HYDROCHLOROTHIAZIDE	int64
VENLAFAXINE.HCL	int64
VENLAFAXINE.HCL.ER	int64
VENTOLIN.HFA	int64
VERAPAMIL.ER	int64
VESICARE	int64
VOLTAREN	int64
VYTORIN	int64
WARFARIN.SODIUM	int64
XARELTO	int64
ZETIA	int64
ZIPRASIDONE.HCL	int64
ZOLPIDEM.TARTRATE	int64
Opioid.Prescriber	int64
Gender1	int64
State1	int64
Specialty1	int64

Length: 259, dtype: object

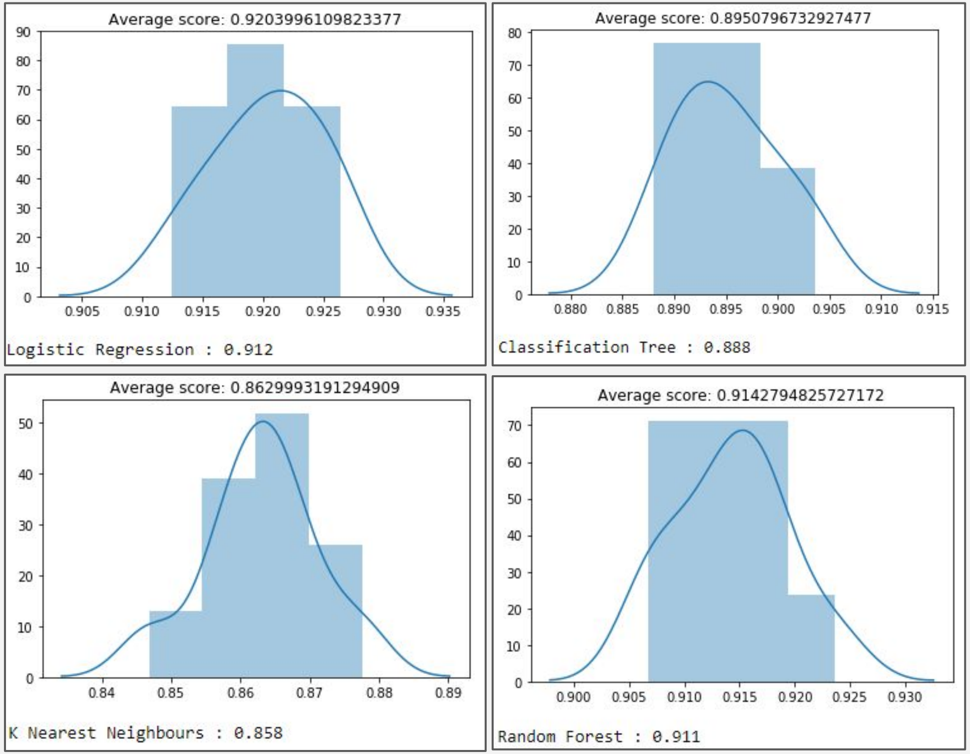
Model Performance using all drug columns

Model Performance:

- Logistic Regression : 0.912
- K Nearest Neighbours : 0.858
- Classification Tree : 0.888
- Random Forest : 0.911
- DT with Bagging: 0.815
- DT with Boosting: 0.96



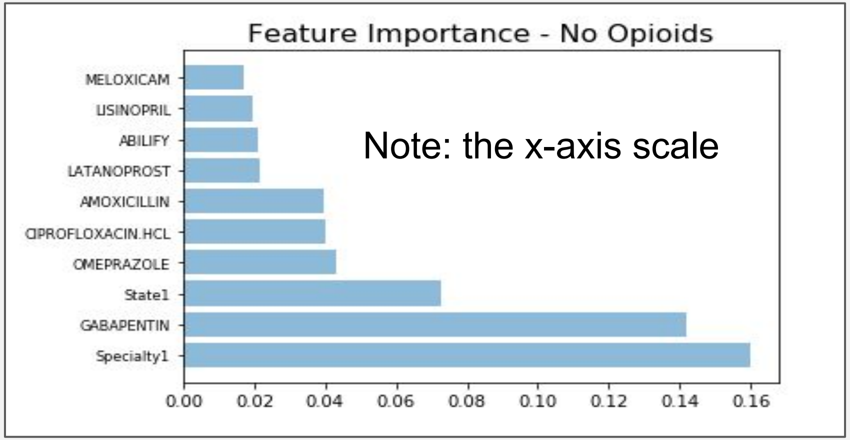
Cross Validation Plots/Scores



Model Performance without Opioid Drug Columns

Model Performance:

- Logistic Regression : 0.759
- K Nearest Neighbours : 0.778
- Classification Tree : 0.779
- Random Forest : 0.823
- DT with Bagging: 0.636
- DT with Boosting: 0.89



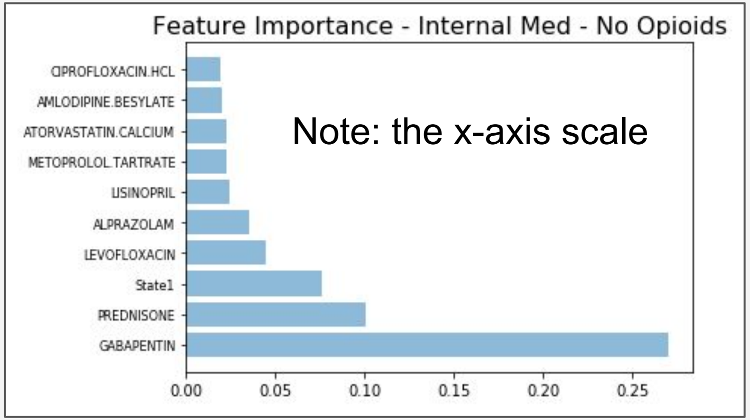
Opioids Dropped:

1 MORPHINE.SLFATE.ER
2 FENTANYL
3 OXYCODONE.HCL
4 OXYCONTIN
5 MORPHINE.SULFATE
6 OXYCODONE.ACETAMINOPHEN
7 HYDROMORPHONE.HCL
8 METHADONE.HCL
9 HYDROCODONE.ACETAMINOPHEN
10 TRAMADOL.HCL
11 ACETAMINOPHEN.CODEINE

Predicting Prescribers by Specialty (Internal Medicine)

Model Performance:

- Logistic Regression : 0.847
- K Nearest Neighbours : 0.812
- Classification Tree : 0.805
- Random Forest : 0.831
- DT with Bagging: 0.743
- DT with Boosting: 0.89



Notes:

- State1 informs this model the most of all features
- Remaining significant features are non-opioid drugs that may have "associative" relation with prescribed opioids

Why did Boosting using Decision Tree outperform?

Bagging vs Boosting

These are ensemble methods which combine several “weaker” models into a “stronger” ensemble.

Bagged models that are trained independently on data that is bootstrapped from the input data. Best used with limited data. We have 25K rows of prescribers.

Boosting creates a strong learner by iteratively adding “weak” learners and adjusting the weight of each weak learner to focus on misclassified examples. Worked best with our data.

When to use Boosting

- When predictors are categorical
- When the time taken to train a model is less of a concern

Source: https://quantdare.com/what-is-the-difference-between-bagging-and-boosting/

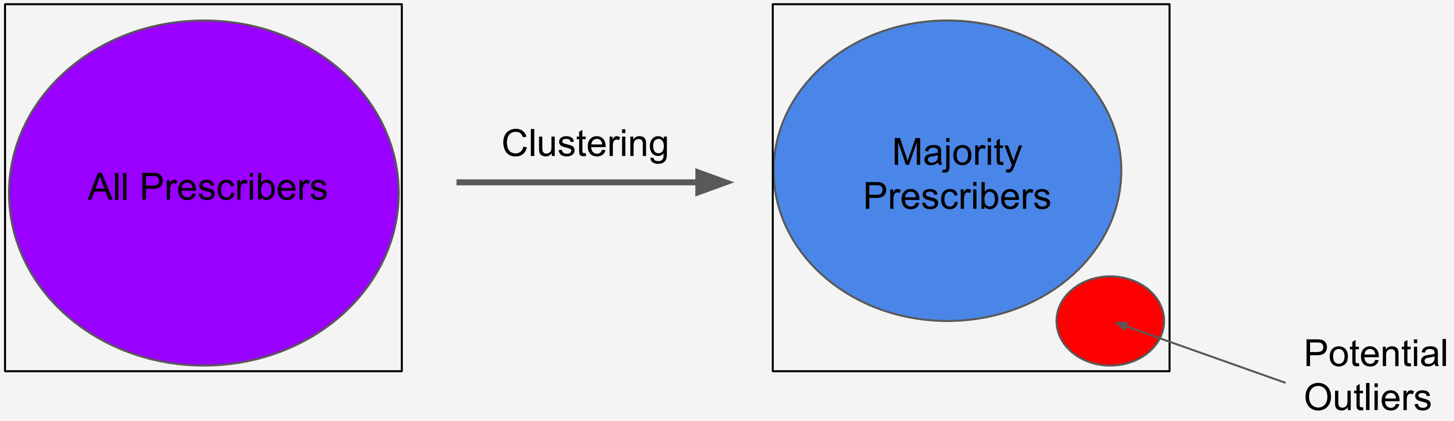
Identifying Prescriber Outliers through Cluster Analysis

- Motivation: Identify prescribers that are prescribing a different drug profile compared to their peers aka “outliers”

- Controlling Factors:
  - Specialty
  - State
  - Opioid Prescriber Flag

Prescriber Outliers - Approach

- We will use unsupervised clustering to create clusters of prescribers based on their drugs prescribed
- We expect to see a dominant cluster with most of the physicians, but also a smaller subset of physicians which we will consider the outliers



Prescriber Outliers - Scaling (TF-IDF Principles)

- But first...

  Term Frequency x Inverse Document Frequency

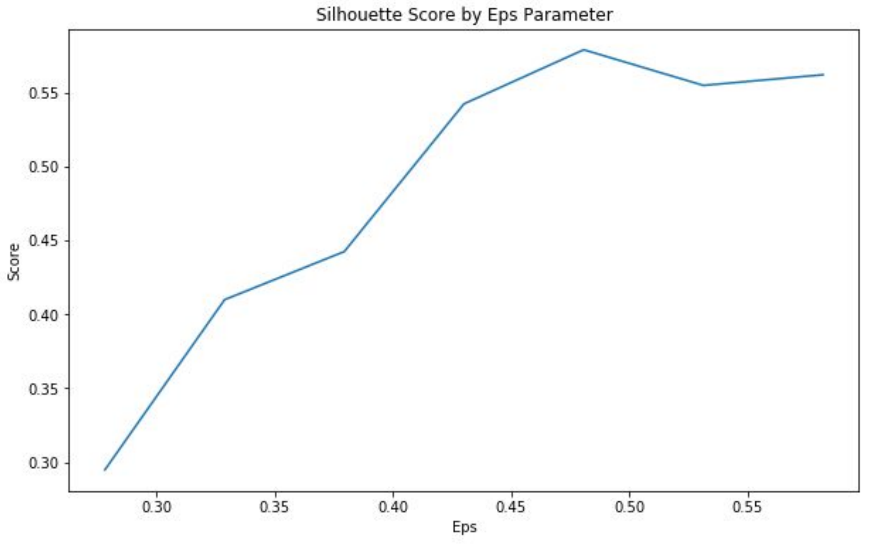
  - Commonly used in document classification, is meant to normalize the length of the documents and give more weight when documents share uncommon key words

  1) Normalize every prescriber by the total number of drugs prescribed. In other words, state each drug as % of total drugs prescribed

  2) Give varying weights to each of the drugs, where uncommon drugs receive more weight and common drugs receive less weight

Prescriber Outliers - Techniques Considered/Used

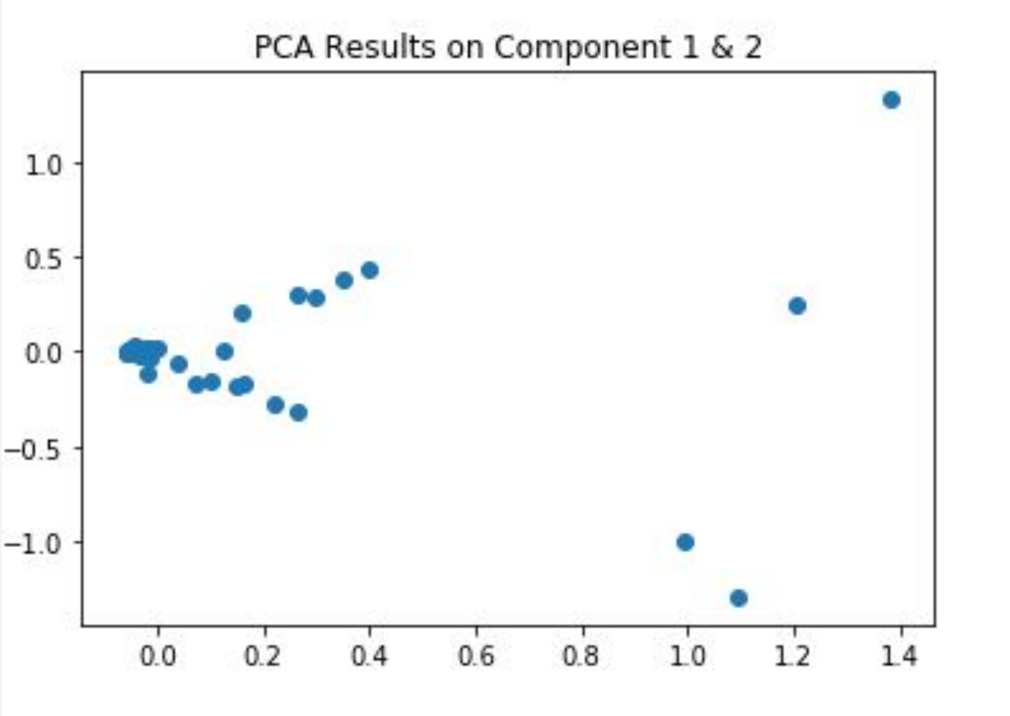
- Clustering
  - K-Means
  - Hierarchical Clustering
  - Density-based Clustering (DBSCAN)
- Dimensionality Reduction / Visualization
  - T-distributed Stochastic Neighbor Embedding (t-SNE)
  - Principal Component Analysis (PCA)



DBSCAN Parameter Eps: Silhouette Score is a score from -1 to +1, with +1 being the strongest evidence of distinct clusters.

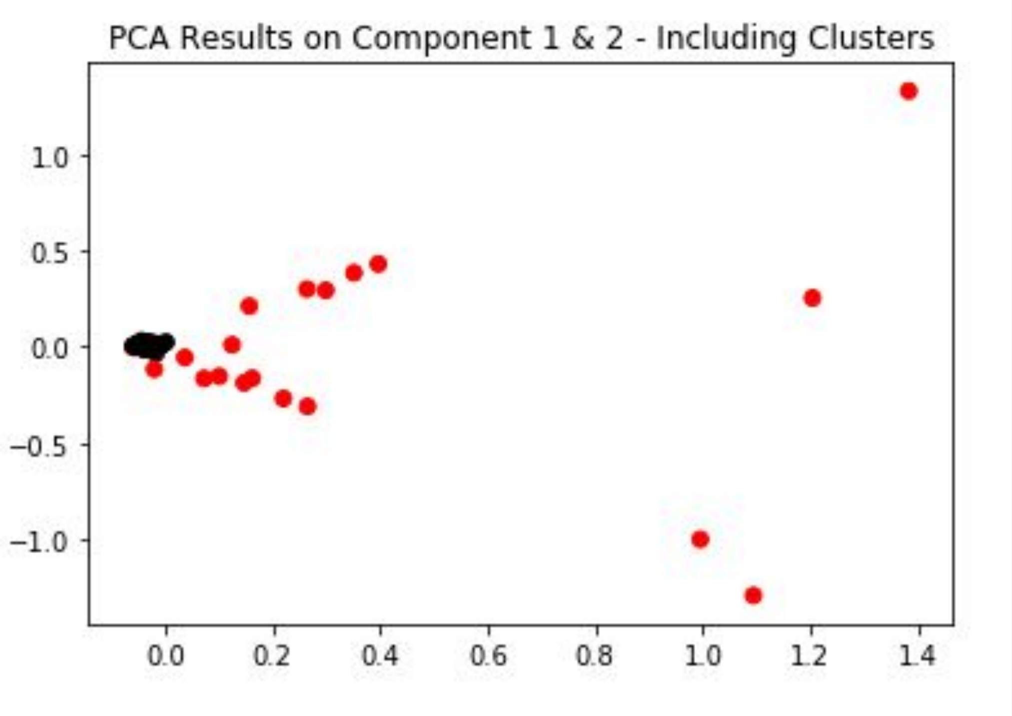
Prescriber Outliers - PCA

- Plotting data on PCA basis allows visualization of underlying data differences
- Filters
  - Internal Medicine
  - New York
  - Opioid Prescribers



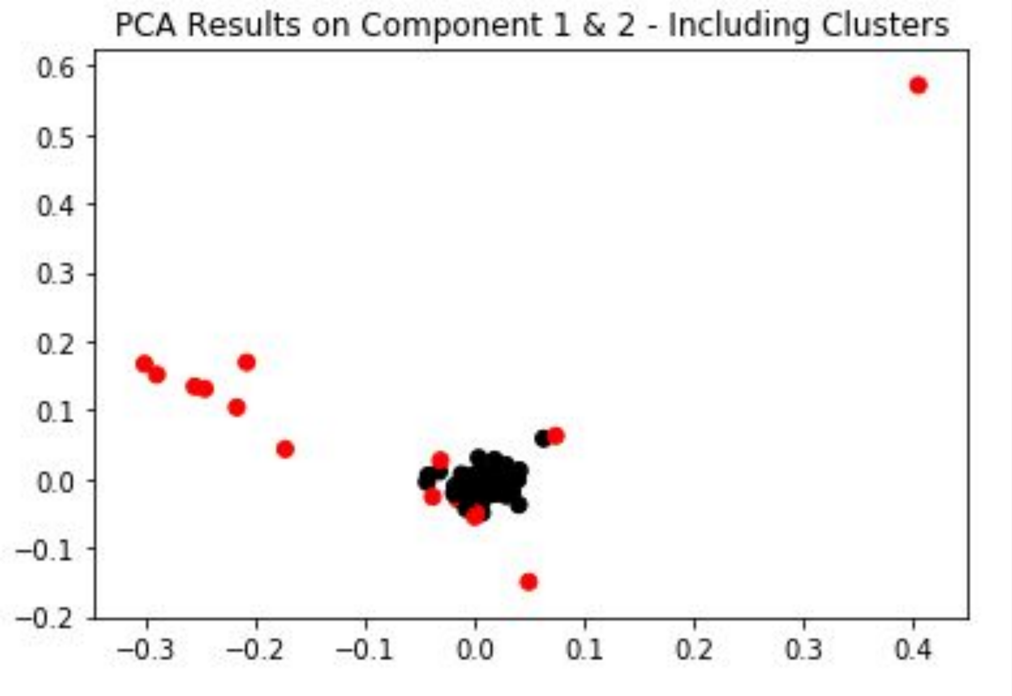
Prescriber Outliers - PCA with Cluster Results (NY)

- Plotting data on PCA basis allows visualization of underlying data differences
- Filters
  - Internal Medicine
  - New York
  - Opioid Prescribers
- Red represents "Outlier" cluster label



Prescriber Outliers - PCA with Cluster Results (PA)

- Plotting data on PCA basis allows visualization of underlying data differences
- Filters
  - Family Practice
  - Pennsylvania
  - Opioid Prescribers
- Red represents “Outlier” cluster label



Prescriber Outliers - Drug Patterns by Cluster

- Results shown for NY
- Prescribers in the Outlier group prescribed **4x** as many opioids as those in the Inlier group (NY, PA)

Differences - Opioids Only

Index	Inliers %	Outliers %	Difference	Is Opioid?
OXYCODONE.HCL	0.2	4.7	4.5	Yes
OXYCODONE.ACETAMINOPHEN	0.4	3.7	3.3	Yes
MORPHINE.SULFATE.ER	0	0.8	0.8	Yes
OXYCONTIN	0	0.7	0.7	Yes
HYDROMORPHONE.HCL	0	0.5	0.5	Yes
MORPHINE.SULFATE	0	0.3	0.3	Yes
ACETAMINOPHEN.CODEINE	0.2	0.5	0.3	Yes
METHADONE.HCL	0	0.2	0.2	Yes
FENTANYL	0.2	0.3	0.1	Yes
HYDROCODONE.ACETAMINOPHEN	0.7	0.9	0.1	Yes
TRAMADOL.HCL	0.5	0.3	-0.3	Yes

Differences - All Drugs

Index	Inliers %	Outliers %	Difference	Is Opioid?
OXYCODONE.HCL	0.2	4.7	4.5	Yes
OXYCODONE.ACETAMINOPHEN	0.4	3.7	3.3	Yes
PREDNISONE	0.6	3	2.4	No
LEVOFLOXACIN	0.2	2.5	2.2	No
ZOLPIDEM.TARTRATE	1.1	2.7	1.6	No
DOXYCYCLINE.HYCLATE	0.1	1.6	1.5	No
CEFUROXIME	0	1.4	1.4	No
SULFAMETHOXAZOLE.TRIMETHOPRIM	0.2	1.5	1.3	No
AZITHROMYCIN	0.6	1.7	1.1	No
IBUPROFEN	0.4	1.4	1.1	No
PROAIR.HFA	0.8	1.8	1.1	No
ACYCLOVIR	0	1.1	1.1	No
ALPRAZOLAM	0.6	1.6	1	No
NAPROXEN	0.3	1.3	1	No
LAMOTRIGINE	0	1.1	1	No
VOLTAREN	0.3	1	0.8	No
MORPHINE.SULFATE.ER	0	0.8	0.8	Yes
CIPROFLOXACIN.HCL	0.4	1.2	0.8	No
OXYCONTIN	0	0.7	0.7	Yes

Prescriber Outliers - Drug Patterns by Cluster

- Results shown for PA
- Prescribers in the Outlier group prescribed **4x** as many opioids as those in the Inlier group (NY, PA)

Differences - Opioids Only

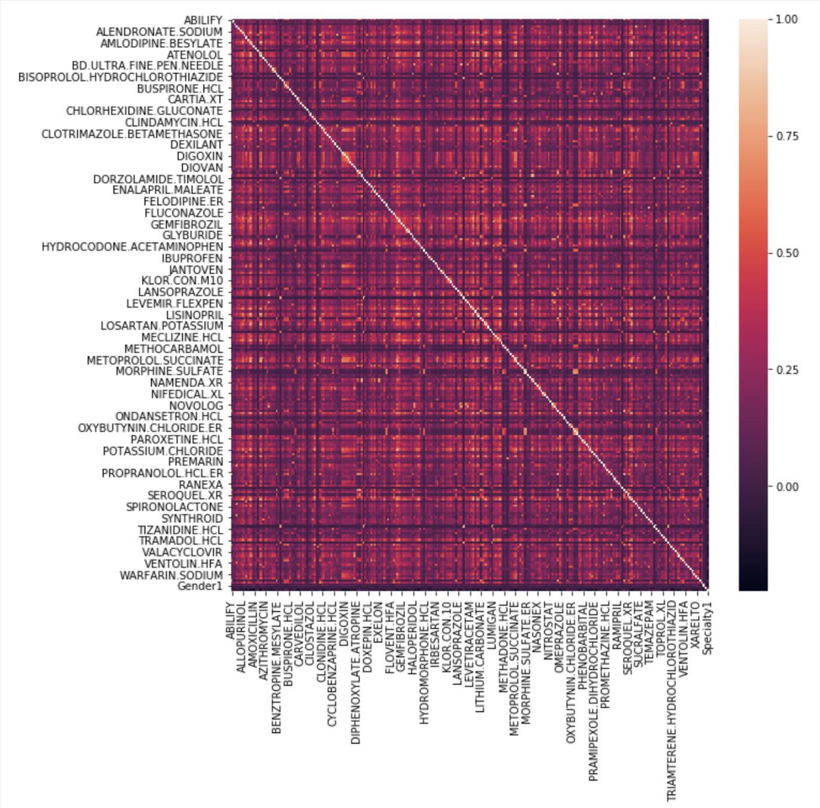
Index	Inliers %	Outliers %	Difference	Is Opioid?
OXYCODONE.HCL	0.5	11.7	11.2	Yes
OXYCODONE.ACETAMINOPHEN	0.6	4.3	3.7	Yes
HYDROCODONE.ACETAMINOPHEN	1.8	4.8	3	Yes
METHADONE.HCL	0	0.8	0.7	Yes
OXYCONTIN	0.2	0.6	0.4	Yes
MORPHINE.SULFATE.ER	0.2	0.5	0.3	Yes
HYDROMORPHONE.HCL	0	0	-0	Yes
MORPHINE.SULFATE	0.1	0	-0.1	Yes
FENTANYL	0.3	0.2	-0.1	Yes
ACETAMINOPHEN.CODEINE	0.1	0	-0.1	Yes
TRAMADOL.HCL	1.2	0.8	-0.5	Yes

Differences - All Drugs

Index	Inliers %	Outliers %	Difference	Is Opioid?
OXYCODONE.HCL	0.5	11.7	11.2	Yes
OXYCODONE.ACETAMINOPHEN	0.6	4.3	3.7	Yes
HYDROCODONE.ACETAMINOPHEN	1.8	4.8	3	Yes
ALPRAZOLAM	1.2	3.5	2.3	No
AZITHROMYCIN	0.6	2.3	1.8	No
PREDNISONE	0.7	2.5	1.8	No
CEPHALEXIN	0.2	1.9	1.7	No
DIAZEPAM	0.2	1.3	1.1	No
DOXYCYCLINE.HYCLATE	0.1	1.1	1	No
CIPROFLOXACIN.HCL	0.4	1.3	0.9	No
LISINOPRIL.HYDROCHLOROTHIAZIDE	0.8	1.7	0.9	No
IBUPROFEN	0.4	1.2	0.9	No
SULFAMETHOXAZOLE.TRIMETHOPRIM	0.2	1	0.8	No
AMOX.TR.POTASSIUM.CLAVULANATE	0.2	1	0.8	No
LEVOFLOXACIN	0.2	0.9	0.7	No
METHADONE.HCL	0	0.8	0.7	Yes

Methods Considered But Not Used for Prediction

- Dimensionality Reduction(PCA)
  - Interpretability would be lost since dimensions get compressed
  - Wanted to see the importance of Individual features and their combinations



Methods Considered But Not Used for Prediction Cont.

- SVM scores less than other models such as Boosting
  - Accuracy score: 0.73
- Why didn’t we use regression
  - Our data was categorical and not continuous
- Association Rule Mining or Graph / Social Network Analysis
  - Our data was not at the transaction level

Future Work

- Continue exploration on geographic drivers outside of data, including state regulatory and medical fee schedule differences
- Augment analysis with patient level data
  - Transaction data
  - Drug or medical history associations
  - Illicit drug data and consumption
- Determine if Outlier drug prescription patterns are indeed anomalous by engaging domain experts from medical field
- Analyze prescriber behavior relative to their peers on a transactional level

Opioid epidemic tracking Metrics(NC)

North Carolina Opioid action plan has been developed to combat the opioid crisis

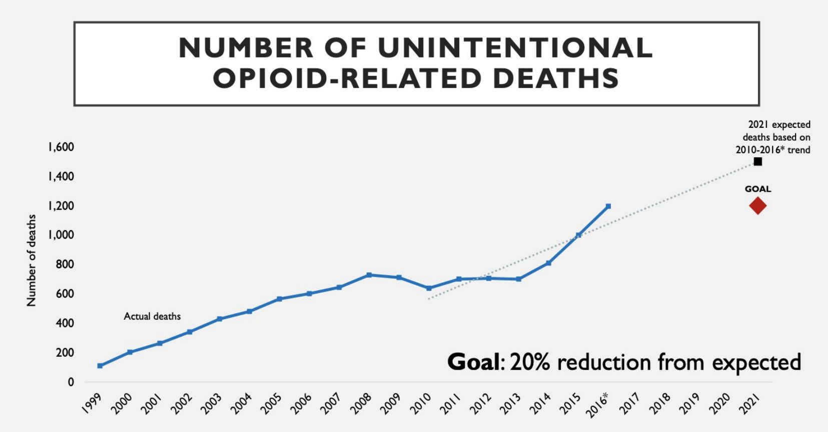
Metrics:

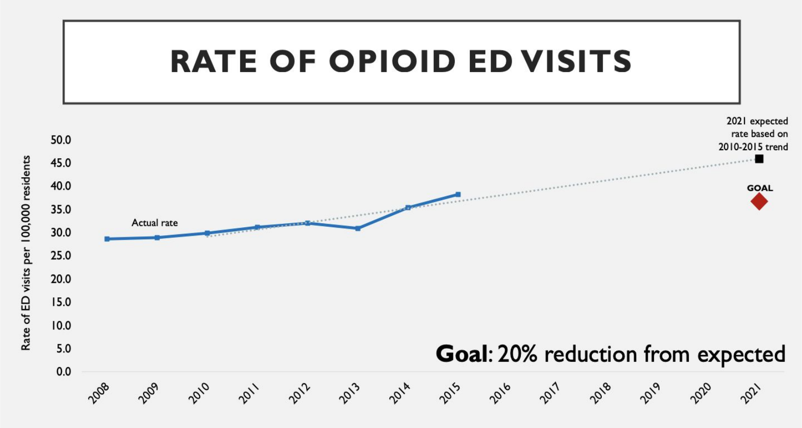
1. Treatment and Recovery
2. Reduce Oversupply of Prescription Opioids
3. Increase access to Naloxone
4. Reduce flow of Illicit Drugs

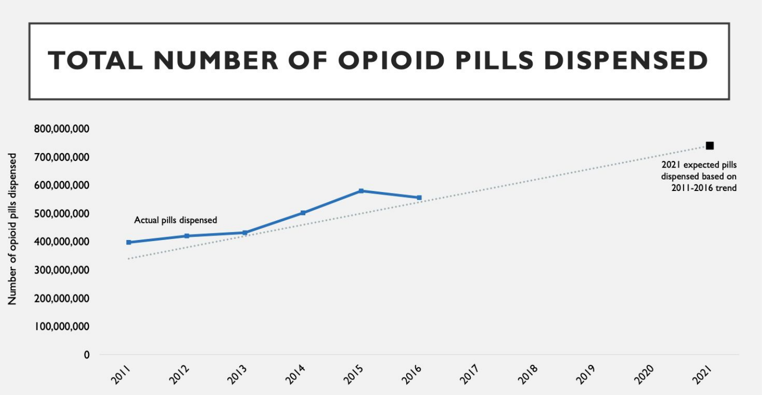
Metrics	Current Data	2021 Trend/Goal
OVERALL		
Number of unintentional opioid-related deaths (ICD10)	1,194 (2016, provisional)	20% reduction in expected 2021 number
Rate of opioid ED visits (all intents)	38.2 per 100,000 residents (2015)	20% reduction in expected 2021 rate
Reduce oversupply of prescription opioids		
Rate of multiple provider episodes for prescription opioids (times patients received opioids from ≥5 prescribers dispensed at ≥5 pharmacies in a six-month period), per 100,000 residents	27.3 per 100,000 residents (2016)	Decreasing trend
Total number of opioid pills dispensed	555,916,512 (2016)	Decreasing trend
Percent of patients receiving more than an average daily dose of >90 MME of opioid analgesics, per quarter	12.3% (Q1 2017)	Decreasing trend
Percent of prescription days any patient had at least one opioid AND at least one benzodiazepine prescription on the same day, per quarter	21.1% (Q1 2017)	Decreasing trend
Reduce Diversion/Flow of Illicit Drugs		
Percent of opioid deaths involving heroin or fentanyl/fentanyl analogues	58.4% (2016, provisional)	------------------------------
Number of acute Hepatitis C cases	182 (2016, provisional)	Decreasing trend
Increase Access to Naloxone		
Number of EMS naloxone administrations	13,069 (2016, provisional)	------------------------------
Number of community naloxone reversals	3,616 (2016)	Increasing trend
Treatment and Recovery		
Number of buprenorphine prescriptions dispensed	467,243 (2016)	Increasing trend
Number of uninsured individuals with an opioid use disorder served by treatment programs	12,248 (SFY16)	Increasing trend
Number of certified peer support specialists (CPSS) across NC	2,383 (2016)	Increasing trend

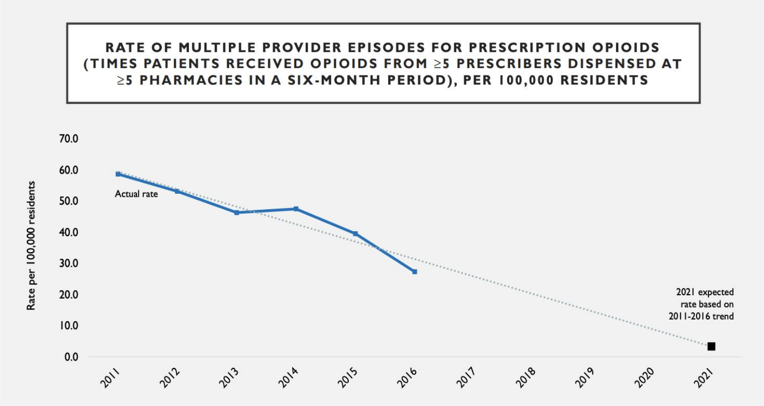
Source: https://www.ncdhhs.gov/about/department-initiatives/opioid-epidemic/north-carolinas-opioid-action-plan

Metric Plots(NC)









Source: https://www.ncdhhs.gov/about/department-initiatives/opioid-epidemic/north-carolinas-opioid-action-plan

Metrics Evaluation

- Analyze number of unintentional deaths and ED Visits by observing diagnosis codes and patient level transactional data
- Compute average rate of multiple provider episodes for prescription Opioids
- Analyze the diversion and flow of illicit drugs-
  Percent of opioid deaths involving heroin and fentanyl analogues by observing ICD - 10 codes
- Number of uninsured individuals and Medicaid beneficiaries with an opioid use disorder served by treatment programs