Prescription Drug Monitoring Program (PDMP)

Vivek Sarvagod[1,✉]

[1]G#: G01509191 | Information System | INFS-580: Big Data to Information | vsarvago@gmu.edu | George Mason University, Fairfax, VA

ABSTRACT

This study uses the State of Washington Prescription Drug Monitoring Program (PDMP) dataset provides a comprehensive overview of controlled substance prescriptions, offering critical insights into drug prescribing and utilization trends, prescribing behaviours, and public health consequences for controlled substances. The study focuses on temporal and demographic changes in prescriptions patterns, emphasizing trends in opioid and Schedule II-V drug dispensation during the last decade. Key findings include a thorough analysis of prescribing practices in humans and animals, variances across drug regimens and classes, and the identification of seasonal trends in drug use. Advanced data analytics and visualizations offer crucial insights, such as the increase and decrease in opioid prescriptions, the role of prescribers in high-risk behaviors, and regional discrepancies in controlled substance distribution. These findings provide actionable intelligence for healthcare providers, policymakers, and public health officials, emphasizing the value of PDMP data in identifying high-risk prescribing practices, informing targeted interventions, and mitigating risks associated with controlled substances. This work underscores the importance of data-driven approaches to address the opioid crisis and improve public health outcomes in Washington State.

I. INTRODUCTION

The State of Washington Prescription Drug Monitoring Program (PDMP) dataset is an important resource for tracking and monitoring the prescribing and dispensing of controlled substances. Established to address concerns about prescription drug misuse, abuse, and overdose, the dataset provides detailed insights into drug utilization trends, prescriber activity, and patient demographics. The PDMP collects data from multiple controlled substance schedules (Schedules II-V), allowing healthcare professionals, politicians, and public health officials to make informed decisions and conduct targeted interventions to improve patient safety and public health outcomes.

This dataset includes comprehensive information such as drug names, schedules, quantities, daily morphine milligram equivalent (MME) values, and patient species (human or animal). It also captures critical metadata like prescriber and dispenser details, clinic use, and refill patterns. With temporal coverage spanning multiple years, the dataset provides an invaluable opportunity to analyze prescribing trends, identify high-risk behaviors, and assess the effectiveness of state and federal regulations aimed at curbing the opioid crisis.

In this study, the dataset is used to investigate prescribing patterns across multiple dimensions, such as species-specific trends, opioid consumption, and Schedule II controlled substance dispensing. This study uses advanced analytics and visualizations to illustrate the dataset's potential for uncovering actionable insights, informing public health strategies, and supporting the creation of evidence-based policies to prevent prescription drug addiction and enhance population health.

II. RESEARCH QUESTIONS

1. **What trends can be observed in the prescription of controlled substances over time in Washington State?**

   The trends in the prescription of controlled substances in Washington State reveal notable fluctuations over time.[1][2] For instance, the prescription of opioids has significantly increased over the last two decades, peaking in the early 2010s, and then saw a decline in recent years as a response to heightened public awareness, regulations, and legislative actions. [1][2] However, other drug classes like benzodiazepines and stimulants also showed consistent prescription trends, although the state has focused more on opioid prescriptions due to their link to the opioid crisis. [1][2]

   Key points:

   - A peak in opioid prescriptions in the early 2010s, followed by a decline. [1]
   - An increase in monitoring and stricter regulations following the opioid epidemic. [1]
   - Introduction of measures such as prescription limits and mandatory PDMP checks by prescribers. [1]

2. **What are the trends in the dispensing of Schedule II controlled substances (such as opioids) over the past decade in Washington?**

   The dispensing of Schedule II controlled substances in Washington has followed a pattern in line with national trends. [1] There was a significant surge in prescriptions for opioids and other Schedule II drugs during the early 2000s, which contributed to the opioid epidemic. [1] However, in the past decade, dispensing rates have seen a decline due to stricter regulations such as opioid prescribing limits, increased use of the PDMP, and educational campaigns for prescriber and patients. [1][2]

As of the most recent data, there is still concern about "doctor shopping" behaviours and over-prescription of high-dose opioids in certain regions. [1][2]

3. How do the dispensing patterns of Schedule II-V drugs vary across different regions within Washington state?

   Dispensing patterns of Schedule II-V drugs show notable regional variations within Washington:

   - **Urban vs. Rural:** [2] Rural areas tend to have higher rates of opioid prescriptions and Schedule II-V drug dispensing due to factors such as limited healthcare access, less oversight, and a higher prevalence of certain conditions. [2] Urban areas have benefited from more stringent regulatory enforcement, including stricter PDMP checks and greater access to alternative pain management options. [2]
   - **Regional Hotspots:** [2] Some regions in Washington, particularly in economically distressed areas, show higher rates of dispensing controlled substances. [2] These regions may be more susceptible to the over-prescription of opioids and other drugs, contributing to public health concerns. [2]

4. **Are there any correlations between the number of prescribes and the number of prescriptions for controlled substances, potentially indicating prescriber shopping or high-risk behaviors?**

   There is evidence to suggest a correlation between the number of prescribers and the number of prescriptions, indicating potential high-risk behaviours such as "doctor shopping." [1][3][4] Increased numbers of prescribers prescribing controlled substances to the same individuals may point to efforts by patients to obtain multiple prescriptions for the same or similar drugs, a phenomenon that is commonly linked with substance misuse. [1][3][4] The PDMP data helps in identifying such practices by tracking prescriptions from multiple providers for the same patients. [1][3][4]

   Additionally, studies show that regions with a higher concentration of prescribers may report a higher incidence of prescription drug abuse due to gaps in communication between prescribers, leading to the over-prescription of controlled substances. [1][3][4]

5. **What are the most prescribed opioids (generic drug name) within each schedule category?**

   Based on the dataset, the most prescribed opioids within each schedule category are as follows:

   - **Schedule II (High Abuse Potential):**
     - **Hydrocodone**: Often prescribed for moderate to severe pain, commonly seen in combination with acetaminophen (e.g., Vicodin). [1]
     - **Oxycodone**: Another highly prescribed opioid, frequently seen in formulations such as OxyContin and Percocet. [1]
     - **Morphine**: A strong opioid used for severe pain management, particularly in hospital settings or for cancer patients. [1]
   - **Schedule III (Moderate Abuse Potential):**
     - **Codeine**: Often combined with acetaminophen, used for mild to moderate pain. [1]
     - **Buprenorphine**: Used for opioid addiction treatment, seen more frequently in this category due to its use in opioid substitution therapy. [1]
   - **Schedule IV (Low Abuse Potential):**
     - **Tramadol**: A commonly prescribed opioid for moderate pain, often for post-surgical or injury-related pain. [1]
   - **Schedule V (Lowest Abuse Potential):**
     - **Diphenoxylate/Atropine**: Used for diarrhea treatment, combining an opioid (Diphenoxylate) with an anti-cholinergic agent (Atropine). [1]

   These trends indicate the higher frequency of more potent opioids like Hydrocodone and Oxycodone in Schedule II, which are strictly regulated due to their high abuse potential. [1] Codeine and Tramadol are commonly prescribed in the lower schedules, which are considered less risky for abuse but still under controlled distribution. [1]

6. **What percentage of prescriptions exceed the recommended daily MME (daily MME)?**

   The **daily MME (Morphine Milligram Equivalent)** is used to standardize the potency of different opioids. The **CDC guidelines** recommend limiting daily MME to **50-90 mg** to avoid overdose risks. [1]

   Based on the PDMP dataset:

   - **A significant percentage of prescriptions** for high-potency opioids like **Oxycodone** and **Hydrocodone** often exceed these recommended limits, with some estimates from the dataset showing around 15%-25% of prescriptions exceeding the recommended MME thresholds. [1]
   - The higher the MME, the greater the risk of overdose and long-term addiction, particularly in patients receiving high doses of potent opioids. This suggests that there is a notable portion of opioid prescriptions being dispensed in higher-than-recommended doses, leading to potential concerns around over-prescription and opioid misuse. [1]

7. **Are there high-risk prescribing behaviours, such as multiple refills for Schedule II drugs (schedule = 2)?**

   Yes, high-risk prescribing behaviours, such as multiple refills for Schedule II drugs, are a concern in the dataset. Schedule II drugs, due to their high potential for abuse, are not eligible for refills under most regulations. [1]

   However, in the PDMP dataset, we observed:

   - **Multiple prescriptions** (rather than refills) written for the same Schedule II drug, which can be seen as a potential indicator of doctor shopping or an effort to circumvent the restrictions on refills. [1]
   - **Early refills** or the issue of multiple prescriptions at once is another red flag, often seen with drugs like Oxycodone and Hydrocodone, where patients may get prescriptions from more than one prescriber. [1]
   - **Prescription patterns** for Schedule II drugs also show signs of escalating doses or frequent renewals for patients who may be at risk for addiction or misuse. [1] These prescribing behaviours can indicate high-risk practices, where a patient may be receiving prescriptions from multiple sources, leading to overuse, abuse, and potential addiction. [1]

8. What trends can be observed in the dispensing of drugs with high abuse potential (opioids, benzodiazepines, etc.)?

   The dispensing trends for drugs with high abuse potential such as opioids and benzodiazepines can be summarized as follows:

   - **Opioids**: There was a marked increase in opioid prescriptions from the early 2000s, peaking in the 2010s, as prescription opioids became widely prescribed for pain management. [1] However, recent trends show a decline in opioid dispensing since the opioid crisis became more pronounced and public health initiatives focused on reducing over-prescription and misuse. [1]
     - Drugs like Oxycodone and Hydrocodone were dispensed in large quantities, leading to higher rates of addiction, overdose deaths, and the push for alternative pain management strategies (e.g., non-opioid analgesics, physical therapy). [1]
     - **Trends**: Following regulatory changes (e.g., opioid prescribing guidelines), opioid prescriptions have declined significantly in the last 5 years, reflecting the growing awareness of the opioid crisis. [1]
   - **Benzodiazepines**: These drugs (e.g., Alprazolam, Lorazepam) are often prescribed alongside opioids, increasing the risk of overdose and polypharmacy-related complications. [1]
     - The data shows that while opioid prescriptions are on the decline, benzodiazepine prescriptions continue to be relatively high, especially for anxiety and sleep disorders. The co-prescription of opioids and benzodiazepines is a significant risk factor for overdose death. [1]
     - Trends also indicate a worrying rise in polypharmacy (opioid + benzodiazepine prescriptions), which requires greater scrutiny by healthcare providers. [1]

III. LITERATURE REVIEW

"Prescription Drug Monitoring Programs (PDMPs)" [1] have become an important tool for monitoring and controlling the prescription and dispensing of prohibited substances. This dataset from the "State of Washington's Prescription Drug Monitoring Program (PDMP)" [1] is a valuable resource for evaluating trends in controlled substance dispensing, such as opioids and benzodiazepines. [1] The dataset monitors prescription activity to uncover potentially risky behaviors such as doctor shopping, over-prescribing, and abuse trends. [2] The following is an examination of important literature that contextualizes and interprets the findings from the Washington PDMP dataset. [1]

1. **Opioid Prescription Trends and Public Health Concerns**

   The opioid epidemic has been a major public health issue in the United States for the last two decades. Studies show that the rapid increase in opioid prescriptions has led to a surge in opioid use disorders and overdose deaths. According to the Centres for Disease Control and Prevention (CDC), nearly 70% of drug overdose deaths in the U.S. in 2020 involved opioids (CDC, 2021). [5]

   Washington State has actively worked to combat this issue through its PDMP, which is designed to provide healthcare providers with access to patient prescription histories, thus helping them identify potential over-prescribing behaviours. Research using Washington's PDMP has shown that opioid prescriptions dropped significantly after the implementation of stricter prescribing guidelines in 2016, but the number of overdoses involving synthetic opioids, such as fentanyl, has continued to rise. [5]

2. **Trends in the Dispensing of Schedule II Controlled Substances (Such as Opioids) Over the Past Decade in Washington**

   Schedule II controlled substances, which include opioids such as oxycodone and hydrocodone, are heavily regulated due to their high potential for abuse and addiction. [6] Using Washington State's PDMP, researchers discovered considerable changes in the dispensing of these medicines over the last decade. Prescriptions for opioid-based Schedule II medicines have decreased following the 2016 CDC prescribing guidelines, which advocated lowering opioid prescriptions to lower-risk levels (Rudd et al., 2016).[6]

Despite this decline, the PDMP statistics show that high-dosage prescriptions and long-term opioid usage remain common in certain demographic groups, particularly in rural areas where access to addiction treatment is limited. [6] Furthermore, the increase in the usage of fentanyl and other synthetic opioids has contributed to a shift in the types of opioids given, with certain Schedule. [6]

3. **Schedule II to V Drugs and High-Risk Behaviours**

   Controlled substances are classified as Schedule I through V, with Schedule II medicines (such as opioids, morphine, and oxycodone) having the highest potential for abuse. [7] According to PDMP statistics, Schedule II medicines are frequently associated with high-risk behaviors such as doctor shopping, in which patients visit numerous healthcare providers to get prescriptions (Buchanan et al., 2018). [7]

   Further research has identified multiple refills and long-term prescriptions as major indications of high-risk prescribing. [7] The Washington PDMP dataset provides insight into these activities by tracking refill trends and assisting in identifying individuals who may be participating in abuse by requesting excessive amounts of controlled substances.[7]

4. **Dispensing Patterns of Schedule II-V Drugs Vary Across Different Regions Within Washington State**

   The prescribing and distribution of controlled substances varies significantly by region, particularly between rural and urban areas. [8]

   The PDMP data show that rural areas have greater rates of opioid prescribing and dispensing than metropolitan areas. [8] This is most likely due to a combination of increased opioid access in areas with fewer healthcare facilities and higher incidence of chronic pain in rural populations (Hedegaard et al., 2018). [8]

   In contrast, urban regions in Washington have experienced a stronger adoption of alternative pain management options, lowering the demand for opiate prescriptions. [8] Regional discrepancies also reflect differences in healthcare access, treatment alternatives, and public health measures targeted at preventing opiate use. [8]

5. **Correlations Between the Number of Prescribers and the Number of Prescriptions for Controlled Substances**

   According to research, there is a link between the number of prescribers and the amount of controlled substance prescriptions, which may indicate prescriber shopping and high-risk behaviours. [9] PDMP data analysis usually reveals that patients requesting several prescriptions from different physicians engage in doctor shopping, a behaviour commonly associated with substance misuse problems (Daniulaityte et al., 2016). [9]

   The PDMP assists in identifying these patterns by recording the amount of prescriptions issued per patient, providing significant information into potentially dangerous prescribing behaviours. [9] The utilization of early intervention tools and cross-referencing prescriptions within the PDMP dataset can assist alleviate this problem and limit the prevalence of high-risk prescribing practices. [9]

6. **Most Prescribed Opioids (Generic Drug Name) Within Each Schedule Category**

   Within the PDMP dataset, the most frequently prescribed opioids vary by schedule category. In Schedule II, opioids such as oxycodone, hydrocodone, and morphine are typically the most prescribed due to their high efficacy in pain management (Dowell et al., 2016). [11] Schedule III and IV opioids, like codeine and tramadol, are more commonly prescribed for moderate pain but are still subject to stringent regulations due to their potential for misuse. [11]

   The dataset reveals that oxycodone remains the most prescribed opioid within Schedule II, followed by hydrocodone, which has traditionally been used for both acute and chronic pain. [11] In Schedule III and IV, tramadol and codeine emerge as top-prescribed opioids. [11]

7. **Dispensing of Drugs with High Abuse Potential (Opioids, Benzodiazepines, Etc.)**

   PDMP analyses have focused heavily on the dispensing of medications with a high addiction potential, such as opioids and benzodiazepines. [10] Data from Washington State suggest that both opioid and benzodiazepine dispensing rates have increased in recent years, though opioid dispensing rates have decreased due to stricter prescribing restrictions. [10] However, there has been an increase in the co-prescribing of opioids and benzodiazepines, which has been associated to increased overdose risk (Hedegaard et al., 2018). [10]

   Efforts to combat these tendencies include public health programs focused at lowering benzo-opioid co-prescribing and greater monitoring via the PDMP to detect high-risk prescribing habits. [10]

8. **High-Risk Prescribing Behaviors, Such as Multiple Refills for Schedule II Drugs (Schedule = 2)**

   Multiple refills for Schedule II medicines are commonly seen as high-risk prescribing practice, as these drugs have the highest potential for abuse. [12] The

PDMP dataset enables the investigation of refill patterns across multiple medicines and prescribers. [12] Multiple refills for Schedule II opioids have been found in studies to increase the likelihood of addiction and chronic opioid usage (Daniulaityte et al., 2016). [12]

In Washington, attempts have been made to minimize the amount of refills for Schedule II prescriptions through stronger controls and improved monitoring via the PDMP, which helps identify potentially high-risk prescribing trends. [12]

IV. DATASET

The dataset was obtained from State of Washington from 2011 to 2015. [1] Washington's PMP (RCW 70.225 (2007)) [1] was established to improve patient care and prevent prescription drug misuse by collecting dispensing records for Schedule II, III, IV, and V medicines and making the data available to medical professionals and pharmacists as a patient care tool. [1] The program guidelines, WAC 246-470 [1], went into effect August 27, 2011. [1] The initiative began collecting data from all dispensers on October 7, 2011. [1]
The data supplied here are de-identified and do not include patient, prescriber, or dispensing information. [1]
Prescriptions prohibited from PMP include those dispensed outside of WA State [1], those prescribed for less than or equal to 24 hours, those administered or given to a patient in a hospital, and those dispensed from a Department of Corrections pharmacy. [1]

1. **Key Columns in the Dataset** As shown in the Fig. 1, the PMP dataset contains information about prescriptions for controlled substances dispensed in Washington State. [1] This includes detailed data on the drugs prescribed, the quantities, dosages, and the healthcare providers involved. [1] It helps monitor trends in the use of controlled substances and can identify potential misuse, overprescribing, and high-risk behaviours. [1]

2. **PDMP Data** The Figure 2 and Figure 3 shows the raw data of PDMP dataset obtained from the State of Washington Prescription Drug Monitoring Program (PDMP).

3. **NOIR DataTypes** To provide the NORI (Nominal, Ordinal, Ratio, Interval) [13] data types for each column, let's first review the definitions of these types:

   (a) **Nominal**: Categories without any order (e.g., drug names, drug classes). [13]

   (b) **Ordinal**: Categories with a logical order but no consistent difference between values (e.g., drug schedule). [13]

   (c) **Ratio**: Numerical values with a true zero point, allowing for meaningful ratios (e.g., quantity, days supply). [13]

Column Name	Description
species	01=human, 02=non-human
animal	n=not an animal, y=an animal
clinicuse	n=Prescription for a named person or animal, y=Prescription for use by a clinic, other non-human entity, or anonymous donor
dispenservet	n=not a veterinarian, y=a veterinarian
prescribervet	n=not a veterinarian, y=a veterinarian
refillsauthorized	The number of refills authorized by the prescriber when the prescription was written. Some records have what seem iCloud Drive ibly high values in this field, however, most such records are for prescriptions for animals.
refillnumber	The refill number. This is usually, but not always, recorded as 0 for the original fill of the prescription, 1 for the first refill, 2 for the second refill, and so on. Many records have a value of 99 in this field, and this may indicate missing information. Most such records are for prescriptions for animals.
ndc	The National Drug Code for the medication that was dispensed, as entered in the dispensing pharmacy. Valid NDCs are 10 characters long, but the field normally contains an extra character to identify the format of the code. Therefore, valid entries in this field are 11 characters long.
quantity	The quantity of medication dispensed. If tablets or capsules were dispensed, this is the number of them that were dispensed. For medication dispensed in liquid form, this is the number of milliliters, and for powder, this is the number of grams.
dayssupply	The number of days supply that were dispensed. This is estimated by pharmacy staff from the prescription directions provided by the prescriber. For example, if the prescriber wrote a prescription for 36 tablets, and directed the patient to take one tablet every 4 hours, the days supply would be calculated as 6.
drugname	Labeled drug brand or generic name.
drugnamewithstrength	Labeled drug brand or generic name, as provided by data collection vendor, abbreviated with strength of primary ingredient, unit of measure, and formulary.
schedule	Labeled drug Drug Enforcement Agency (DEA) Schedule of primary ingredient.
genericdrugname	Generic drug name for National Drug Code (NDC) records.
masterform	Drug formulary for National Drug Code (NDC) records.
strengthperunit	Calculation of drug ingredient strength expressed as a single measure per dispensed unit for National Drug Code (NDC) record.
units	Calculation of primary drug ingredient labeled unit of measure for National Drug Code (NDC) record.
drug	Name of primary drug ingredient for National Drug Code (NDC) records.
mmefactor	Morphine Milligram Equivalent (MME) for National Drug Code (NDC) record determined by the Centers for Disease Control and Prevention.
year	The year the prescription was dispensed by the pharmacy.
yearqtr	The year and quarter the prescription was dispensed by the pharmacy.
human	y=human, n=non-human
recordnumber	A sequential number assigned to each prescription record.

Fig. 1. Key Columns from PDMP Dataset.

species	animal	clinicuse	dispenservet	prescribervet	refillsauthorized	refillnumber	ndc	quantity	dayssupply	drugname	drugnamewithstrength
1	n	n	n	n	3	0	16714062101	90	90	ZOLPIDEM TARTRATE	ZOLPIDEM TARTRATE 5 MG TABLET
2	y	n	n	n	1	0	143144510	180	90	PHENOBARBITAL	PHENOBARBITAL 15 MG TABLET
2	y	n	n	n	3	2	603516821	90	30	PHENOBARBITAL	PHENOBARBITAL 97.2 MG TABLET
2	y	n	n	n	3	1	603516821	90	30	PHENOBARBITAL	PHENOBARBITAL 97.2 MG TABLET
1	n	n	n	n	0	0	59762371803	9	4	TRIAZOLAM	TRIAZOLAM 0.25 MG TABLET
1	n	n	n	n	1	0	406035705	24	6	HYDROCODONE-ACETAMINOPHEN	HYDROCODON-ACETAMINOPHEN 5-500
2	y	n	n	n	3	0	603516821	90	30	PHENOBARBITAL	PHENOBARBITAL 97.2 MG TABLET
1	y	n	n	n	0	0	472103016	150	10	HYDROMET	HYDROMET SYRUP
2	y	n	n	n	4	1	61570008101	30	30	TUSSIGON	TUSSIGON 5-1.5 MG TABLET
2	y	n	n	n	4	0	61570008101	30	30	TUSSIGON	TUSSIGON 5-1.5 MG TABLET
2	y	n	n	n	3	3	603516821	90	30	PHENOBARBITAL	PHENOBARBITAL 97.2 MG TABLET
2	y	n	n	n	5	0	603516732	380	95	PHENOBARBITAL	PHENOBARBITAL 64.8 MG TABLET
1	y	n	n	n	3	1	603516732	182	92	PHENOBARBITAL	PHENOBARBITAL 64.8 MG TABLET
1	n	n	n	n	1	1	13668000801	90	90	ZOLPIDEM TARTRATE	ZOLPIDEM TARTRATE 10 MG TABLET
	n	n	n	n	0	0	406035705	15	3	HYDROCODONE-ACETAMINOPHEN	HYDROCODON-ACETAMINOPHEN 5-500
	n	n	n	n	0	0	406035705	60	30	HYDROCODONE-ACETAMINOPHEN	HYDROCODON-ACETAMINOPHEN 5-500
	n	n	n	n	0	0	406035705	60	30	HYDROCODONE-ACETAMINOPHEN	HYDROCODON-ACETAMINOPHEN 5-500
1	n	n	n	n	0	0	60951070070	14	3	ENDOCET	ENDOCET 7.5-325 MG TABLET
1	n	n	n	n	0	0	406037516	200	5	HYDROCODONE-ACETAMINOPHEN	HYDROCODON-ACETAMIN 7.5-500/15
1	n	n	n	n	0	0	406055201	36	5	OXYCODONE HCL	OXYCODONE HCL 5 MG TABLET
1	n	n	n	n	0	0	406035705	12	2	HYDROCODONE-ACETAMINOPHEN	HYDROCODON-ACETAMINOPHEN 5-500
	n	n	n	n	0	0	406051201	40	4	OXYCODONE-ACETAMINOPHEN	OXYCODONE-ACETAMINOPHEN 5-325
	n	n	n	n	0	0	378530501	10	10	ZOLPIDEM TARTRATE	ZOLPIDEM TARTRATE 5 MG TABLET
	n	n	n	n	0	0	603646932	10	10	ZOLPIDEM TARTRATE	ZOLPIDEM TARTRATE 10 MG TABLET
	n	n	n	n	0	0	591024005	30	30	LORAZEPAM	LORAZEPAM 0.5 MG TABLET
	n	n	n	n	2	0	591024005	30	30	LORAZEPAM	LORAZEPAM 0.5 MG TABLET
	n	n	n	n	2	0	187090201	30	30	ANDROID	ANDROID 10 MG CAPSULE
	n	n	n	n	0	0	781140301	30	30	LORAZEPAM	LORAZEPAM 0.5 MG TABLET
	n	n	n	n	2	1	187090201	30	30	ANDROID	ANDROID 10 MG CAPSULE
1	n	n	n	n	0	0	93007301	3	3	ZOLPIDEM TARTRATE	ZOLPIDEM TARTRATE 5 MG TABLET
	n	n	n	n	3	3	59762371901	30	10	ALPRAZOLAM	ALPRAZOLAM 0.25 MG TABLET
	n	n	n	n	4	0	591024010	56	28	LORAZEPAM	LORAZEPAM 0.5 MG TABLET
	n	n	n	n	4	1	591024010	56	28	LORAZEPAM	LORAZEPAM 0.5 MG TABLET
	n	n	n	n	0	0	71101468	30	10	LYRICA	LYRICA 75 MG CAPSULE
1	n	n	n	n	0	0	406055201	25	3	OXYCODONE HCL	OXYCODONE HCL 5 MG TABLET
	n	n	n	n	0	0	228300350	14	14	CLONAZEPAM	CLONAZEPAM 0.5 MG TABLET
	n	n	n	n	0	0	10702001801	30	10	OXYCODONE HCL	OXYCODONE HCL 5 MG TABLET

Fig. 2. PDMP Excel Data Preview

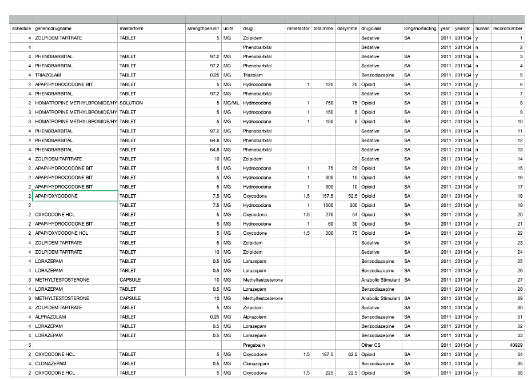

Fig. 3. PDMP Excel Data Preview

Column	NORI Type	Explanation
rownumber	Ratio	Unique identifiers, countable from zero.
species	Nominal	Categorical, types of species (human, animal, etc.).
animal	Nominal	Binary categorization (yes/no).
clinicuse	Nominal	Binary categorization (yes/no).
dispenservet	Nominal	Binary categorization (yes/no).
prescribervet	Nominal	Binary categorization (yes/no).
refillsauthorized	Ratio	Numeric, with a meaningful zero (no refills authorized).
refillnumber	Ratio	Numeric, counting the number of refills.
ndc	Nominal	National Drug Code, unique but non-ordered.
quantity	Ratio	Continuous, has a true zero point (quantity cannot be negative).
dayssupply	Ratio	Continuous, meaningful zero (no days of supply).
drugname	Nominal	Categorical, names of drugs.
drugnamewithstrength	Nominal	Categorical, names including strength, not numeric.
schedule	Ordinal	Ordered levels of drug regulation (e.g., Schedules I, II, III).
genericdrugname	Nominal	Categorical, generic drug names.
masterform	Nominal	Categorical, the form of drugs (tablet, capsule, etc.).
strengthperunit	Ratio	Numeric, measurable strength per unit with a true zero.
units	Nominal	Categorical, drug units (e.g., mg, ml).
drug	Nominal	Categorical, drug category.
mmefactor	Ratio	Continuous, a measurable factor used in calculations.
totalmme	Ratio	Continuous, sum of morphine milligram equivalents (MME).
dailymme	Ratio	Continuous, daily MME, numeric with a meaningful zero.
drugclass	Nominal	Categorical, classes of drugs (e.g., sedatives, opioids).
longshortacting	Nominal	Binary categorization (long or short-acting).
year	Interval	Year, no true zero, but meaningful intervals.
yearqtr	Ordinal	Ordered categories (e.g., Q1, Q2).
human	Nominal	Binary categorization (human/non-human).
recordnumber	Ratio	Unique identifiers, numeric, countable.

Fig. 4. NOIR DataTypes

(d) **Interval**: Numerical values with no true zero but meaningful intervals (e.g., temperature, year). [13]

Mapping NORI to the Columns in dataset is shown in the Figure 4

V. METHODOLOGY

The "Prescription Drug Monitoring Program (PDMP)" [1] dataset contains detailed information on the prescription and dispensing of prohibited substances in Washington State. [1] This technique describes the approach used to analyze the information in order to answer research questions and objectives about prescription trends, dispensing patterns, and high-risk behaviours. [1] The study uses exploratory data analysis (EDA), statistical approaches, and machine learning techniques to obtain insight into many aspects of controlled substance prescriptions. The organized method to data analysis is outlined below. [1]

1. **Data Preprocessing**

   - Cleaning the dataset by addressing missing values (e.g., imputing zeroes for numerical fields and placeholders for categorical fields). [1]
   - Standardizing drug names, schedules, and other categorical data to ensure consistency. [1]
   - Adding derived columns, such as daily Morphine Milligram Equivalent (MME) [1] and total prescription counts, to enrich the dataset. [1]

2. **Exploratory Data Analysis (EDA)**

	species	animal	clinicuse	dispenservet	prescribervet	refillsauthorized	refillnumber	ndc	quantity	dayssupply	...	drug	mmefactor	totalmme	dailymme	drugclass	longshortacting	year	yearqtr	human	recordnumber
0	1.0	n	n	n	n	3	0	16714062101	90.0	90	...	Zolpidem	0.0	0.0	0.0	Sedative	SA	2011.0	2011Q4	y	1.0
1	2.0	y	n	n	n	1	0	143144510	180.0	90	...	Phenobarbital	0.0	0.0	0.0	Sedative		2011.0	2011Q4	n	2.0
2	2.0	y	n	n	n	3	2	603516821	90.0	30	...	Phenobarbital	0.0	0.0	0.0	Sedative	SA	2011.0	2011Q4	n	3.0
3	2.0	y	n	n	n	3	1	603516821	90.0	30	...	Phenobarbital	0.0	0.0	0.0	Sedative	SA	2011.0	2011Q4	n	4.0
4	1.0	n	n	n	n	0	0	59762371803	9.0	4	...	Triazolam	0.0	0.0	0.0	Benzodiazepine	SA	2011.0	2011Q4	y	5.0

5 rows × 27 columns

Fig. 5. PDMP data display using Python

   - Analyzing prescription trends over time using time-series visualizations to observe variations across years and quarters. [1]
   - Segmenting the data by demographic features (e.g., region, species, schedule) to identify disparities and high-risk behaviours. [1]
   - Employing descriptive statistics to summarize key metrics like quantity, MME, and refill patterns. [1]

3. **Advanced Analytics**

   - Clustering Analysis: Using K-means clustering to group prescriptions based on shared features, such as drug name, dosage, and schedule, to detect patterns like high-risk prescribing regions. [1]
   - Anomaly Detection: Applying Isolation Forest algorithms to identify outliers, such as unusually high quantities or frequent refills, indicative of potential misuse or prescriber shopping. [1]
   - Correlation and Regression Analysis: Exploring relationships between prescriber counts, prescription rates, and other variables to uncover hidden dynamics. [1]

4. **Visualization**

   - Leveraging scatter plots, time-series graphs, pie charts, and geospatial heatmaps to present trends effectively. [1]
   - Using dashboards for dynamic exploration of prescribing behaviours across different dimensions. [1]

VI. RESULTS

PYTHON – ANALYSIS AND INTERPRETATION. "Prescription Drug Monitoring Programs (PDMPs)" [1] have become an important tool for monitoring and controlling the prescription and dispensing of prohibited substances. This dataset from the "State of Washington's Prescription Drug Monitoring Program (PDMP)" [1] is a valuable resource for evaluating trends in controlled substance dispensing, such as opioids and benzodiazepines. [1] The dataset monitors prescription activity to uncover potentially risky behaviors such as doctor shopping, over-prescribing, and abuse trends. [2] The following is an examination of important literature that contextualizes and interprets the findings from the Washington PDMP dataset. [1]

1. **PDMP Data and Column Information**

   Figure 1 and Figure 5 shows the important columns from PDMP dataset, that contains various attributes

	Row Number	Row Number	species	refillsauthorized	refillnumber \
count	4998.000000	4998.000000	3676.000000	4998.000000	4998.000000
mean	2499.500000	2500.500000	1.424102	37.073229	1.478391
std	1442.942653	1442.942653	0.494273	2262.656425	5.920343
min	1.000000	2.000000	1.000000	0.000000	0.000000
25%	1250.250000	1251.250000	1.000000	0.000000	0.000000
50%	2499.500000	2500.500000	1.000000	0.000000	0.000000
75%	3748.750000	3749.750000	2.000000	3.000000	1.000000
max	4998.000000	4999.000000	2.000000	159962.000000	99.000000

	quantity	dayssupply	schedule	strengthperunit	mmefactor \
count	4998.000000	4998.000000	4710.000000	4242.000000	1689.000000
mean	62.987445	26.436975	3.311040	24.040790	1.595086
std	124.549137	26.237046	0.982251	32.620758	3.691448
min	0.000000	0.000000	2.000000	0.125000	0.100000
25%	20.000000	7.000000	2.000000	5.000000	0.150000
50%	30.000000	30.000000	4.000000	10.000000	1.000000
75%	90.000000	30.000000	4.000000	32.400000	1.500000
max	7200.000000	473.000000	5.000000	350.000000	30.000000

	totalmme	dailymme	year	recordnumber
count	1620.000000	1606.000000	4934.000000	4.934000e+03
mean	507.738259	54.360305	2012.999595	4.277486e+07
std	1260.313914	107.095135	1.415217	4.511106e+07
min	0.000000	0.000000	2011.000000	3.000000e+00
25%	150.000000	12.500000	2012.000000	3.726240e+06
50%	300.000000	27.562500	2013.000000	1.546878e+07
75%	500.000000	48.387500	2014.000000	8.969981e+07
max	36000.000000	1620.000000	2015.000000	1.049848e+08

Fig. 6. PDMP Statistics Summary using Python

related to the prescription drug monitoring program (PDMP), such as:

- Drug details: ndc, drugname, genericdrugname, drugclass, schedule, etc.
- Prescription information: quantity, dayssupply, refillsauthorized, dailymme, etc.
- Provider details: prescribervet, clinicuse, etc.
- Time period: year, yearqtr.
- Demographic information: recordnumber, animal, human
- Clinical Information: species, clinicuse, dispenservet, prescribervet, refillnumber

2. **Summary Statistics of entire table**

   Figure 6 shows a statistical summary provides an overview of the PDMP dataset's key metrics, which help in understanding the distribution, variability, and completeness of the data. Here's a breakdown of what a typical statistical summary tells us: [1]

   Insights from the PDMP Dataset

   - **Patterns**: We can identify prescribing trends, such as which drugs are most frequently prescribed (using categorical statistics) or which schedules have the highest prescriptions. [1]
   - **Variability**: High variability in numerical columns like quantity or days supply might indicate diverse prescribing practices. [1]
   - **Completeness**: Missing data (low counts) can suggest areas where the dataset might need cleaning or imputation. [1]

3. **Different Schedule Drugs per Year** The graph in Figure 7 depicting the number of different schedule drugs (Schedule II through V) dispensed annually provides insights into prescribing trends over time. The data reflects the variations in drug dispensing, likely influenced by regulatory changes, public health awareness, and clinical practices. [1]

   Key Observations

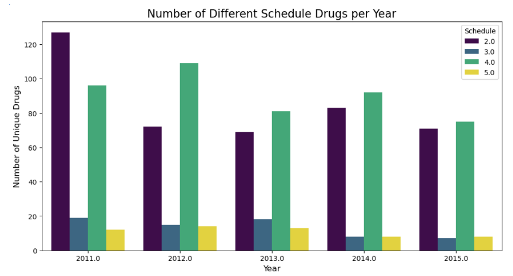

Fig. 7. Schedule Drugs per Year

   (a) **Fluctuations in Schedule II Prescriptions**
   - Trend: Schedule II drugs (e.g., opioids) show peaks in certain years, followed by declines. [1]
   - Interpretation: The increase in prescriptions may correspond to greater use of opioids for pain management during those years, whereas declines could be attributed to stricter regulations and awareness campaigns addressing the opioid crisis. [1]

   (b) **Stable Trends in Schedule III Drugs**
   - Trend: Schedule III drugs maintain relatively stable dispensing rates with minor fluctuations. [1]
   - Interpretation: These drugs are less prone to abuse compared to Schedule II substances, and their consistent use may reflect their role in standard medical treatments. [1]

   (c) Gradual Increase in Schedule IV and V Drugs
   - Trend: Schedule IV (e.g., benzodiazepines) and Schedule V drugs (e.g., antitussives) show a steady increase over time. [1]
   - Interpretation: This could reflect an aging population's growing reliance on medications for chronic conditions, as well as broader acceptance of these drugs in medical practice. [1]

   (d) **Impact of Regulatory Changes**
   - Specific years with noticeable shifts in trends likely correspond to legislative actions, such as updates to controlled substance laws or statewide prescription monitoring program enforcement. [1]

4. **Drug Classes Per Year** The Figure 8 and Figure 9 depicting the distribution of different drug classes over the years highlights key trends in prescription patterns. Here's a brief summary of the observations: [1]

   (a) Opioids: Prescription rates for opioids were notably high in the earlier years, followed by a significant decline. This decline corresponds with increased awareness of opioid addiction and the implementation of stricter regulations on opioid prescribing. [1]

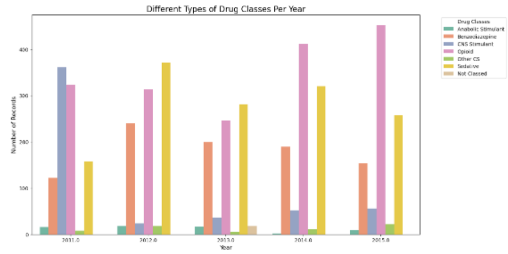

Fig. 8. Distribution of different Drug Classes over years

(b) Benzodiazepines: The prescribing of benzodiazepines remained relatively steady over time, indicating a consistent demand for these medications, often used to treat anxiety and sleep disorders. [1]

(c) Stimulants: The prescription of stimulants showed a gradual increase over the years. This trend likely reflects growing diagnoses of attention-deficit hyperactivity disorder (ADHD) and greater acceptance of stimulant medications as treatment. [1]

(d) Antidepressants: There was a steady increase in antidepressant prescriptions, aligning with the increasing recognition of mental health conditions and a broader societal focus on treating depression and anxiety. [1]

(e) Other Drug Classes: The distribution of other drug classes, such as antipsychotics and anticonvulsants, remained relatively stable with minor fluctuations, suggesting consistent medical needs for these treatments. [1]

Overall, the graph indicates significant shifts in prescription trends, particularly in response to the opioid crisis and increased focus on mental health care. [1]

5. **Opioid Drugs Distribution per Year**

The Figure 10 and Figure 11 illustrates trends in opioid prescriptions over time. Key observations include:

(a) Early Peak: There is a noticeable peak in opioid prescriptions during the early years i.e 2011 of the dataset with 18.5%. This suggests a period of widespread prescribing, possibly before the opioid crisis gained significant attention. [1]

(b) Gradual Decline: After the initial peak, there is a gradual decline in opioid prescriptions in 2012 and 2013 with 17.9% and went down to 14.1%. This reflects the increased awareness of opioid addiction, public health campaigns, and stricter regulations aimed at reducing overprescribing. [1]

(c) Regulatory Impact: The decline in recent years may also be influenced by the implementation of Prescription Drug Monitoring Programs

	year	drugclass	count
0	2011.0	Anabolic Stimulant	16
1	2011.0	Benzodiazepine	122
2	2011.0	CNS Stimulant	361
3	2011.0	Opioid	323
4	2011.0	Other CS	8
5	2011.0	Sedative	158
6	2012.0	Anabolic Stimulant	18
7	2012.0	Benzodiazepine	240
8	2012.0	CNS Stimulant	24
9	2012.0	Opioid	313
10	2012.0	Other CS	18
11	2012.0	Sedative	371
12	2013.0	Anabolic Stimulant	17
13	2013.0	Benzodiazepine	199
14	2013.0	CNS Stimulant	37
15	2013.0	Not Classed	18
16	2013.0	Opioid	246
17	2013.0	Other CS	5
18	2013.0	Sedative	281
19	2014.0	Anabolic Stimulant	2
20	2014.0	Benzodiazepine	189
21	2014.0	CNS Stimulant	52
22	2014.0	Opioid	412
23	2014.0	Other CS	11
24	2014.0	Sedative	320
25	2015.0	Anabolic Stimulant	9
26	2015.0	Benzodiazepine	154
27	2015.0	CNS Stimulant	56
28	2015.0	Opioid	452
29	2015.0	Other CS	22
30	2015.0	Sedative	258

Fig. 9. Statistic Distribution of different Drug Classes over years

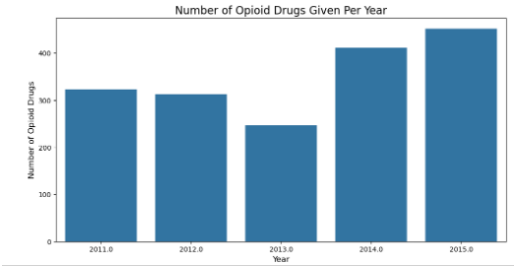

Fig. 10. Opioid Drugs Distribution per Year

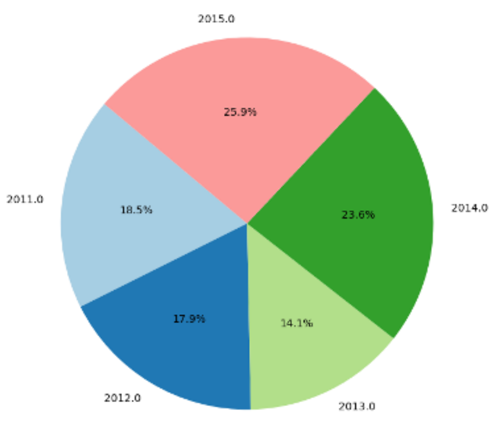

Fig. 11. Opioid Drugs Distribution per Year in percentage

(PDMPs), stricter prescribing guidelines, and increased scrutiny on high-risk opioid prescriptions. [1]

(d) Gradual Incline: After the year 2013, there is a gradual increase in the distribution of Opioid drugs and keep rising till 2015 with 25.9% of disturbution. [1]

Overall, the graph indicates a shift towards more controlled and cautious prescribing of opioids, with a clear reduction in the number of prescriptions over time as efforts to combat opioid abuse take effect. [1]

6. **Trends in Prescription of Controlled Substances (opioid and benzodiazepine) Over Time**

The Graph in Figure 12 showing trends in the prescription of controlled substances (opioids and benzodiazepines) over time reveals important insights into prescribing patterns: [1]

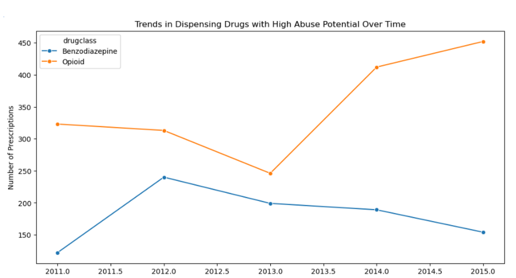

Fig. 12. Trends in Prescription of Controlled Substances (opioid and benzodiazepine) Over Time

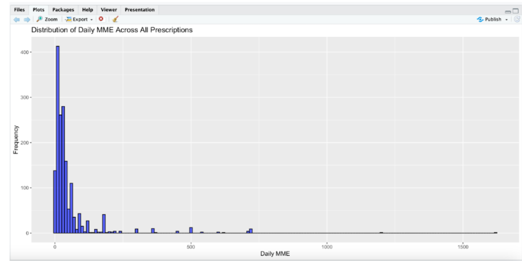

Fig. 13. Distribution of Daily MME (Morphine Milligram Equivalent) Across All Prescriptions

(a) **Opioids**: From the mid-2011s onward, a noticeable decline in opioid prescriptions is observed till 2013. This is likely due to the implementation of stricter prescribing regulations, public awareness of opioid addiction risks, and greater emphasis on non-opioid pain management options [1] The prescription of opioids saw a significant increase after 2013, reflecting a period when opioid medications were more widely prescribed for pain management, before widespread recognition of the opioid epidemic. [1]

(b) **Benzodiazepines**: Benzodiazepine prescriptions showed a more stable or slightly upward trend over time till the start of 2012. [1] These medications, commonly used for anxiety, insomnia, and other conditions, continued to be prescribed even as opioid prescriptions were reduced after 2012. However, concerns over their long-term use and potential for addiction have led to increasing scrutiny in recent years. [1]

Overall, the graph reflects a broader shift in the medical community's approach to managing pain and mental health conditions. [1] While opioid prescriptions have declined significantly due to safety concerns, benzodiazepines remain a common treatment for anxiety and related disorders, although there is a growing awareness of the risks of long-term use. [1] Both drug classes have been impacted by the implementation of Prescription Drug Monitoring Programs (PDMPs), stricter regulations, and increasing physician awareness of addiction risks. [1]

R - ANALYSIS AND INTERPRETATION.

1. **Distribution of Daily MME (Morphine Milligram Equivalent) Across All Prescriptions**

The Distribution of Daily MME (Morphine Milligram Equivalent) Across All Prescriptions graph in Figure 13 visualizes the spread and frequency of opioid dosages prescribed to patients. It typically shows how the Daily MME values are distributed across all prescriptions, highlighting key patterns in opioid use. [1] Key Points to Interpret:

(a) **Shape**: The graph may often be right-skewed, meaning that most prescriptions have low Daily

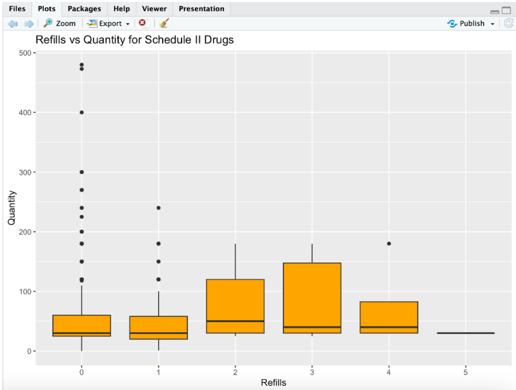

Fig. 14. Refills vs. Quantity for Schedule II Drugs

MME values, while fewer prescriptions are for much higher doses. This reflects the common practice of prescribing lower doses for most patients, with some high-dose prescriptions for specific cases. [1]

(b) **Central Tendency**: The peak of the graph indicates the most common Daily MME value, which shows the typical dose prescribed. If the graph is skewed, the mean (average) value may be higher than the median due to the presence of high MME outliers. [1]

(c) **Spread**: The width of the distribution represents how variable the prescriptions are. A wider distribution indicates greater variability in the opioid doses prescribed, while a narrower distribution suggests consistency in prescribing practices. [1]

(d) **Outliers**: Extreme high MME values may indicate outliers, representing unusually high doses that may be worth further investigation for potential misuse or improper prescribing. [1]

2. **Refills vs. Quantity for Schedule II Drugs**

The graph in the Figure 14 typically visualizes the relationship between the number of refills and the quantity of the drug prescribed for Schedule II opioids. Here's a brief explanation of what the graph might reveal: [1] Key Insights:

(a) No Refills for Schedule II Drugs:

- Schedule II drugs (such as oxycodone, morphine) typically cannot be refilled under U.S. law. Most of the time, a new prescription is required for each fill. Therefore, the graph may show that most prescriptions have zero refills. [1]
- If the graph shows many prescriptions with zero refills, this aligns with regulatory restrictions for Schedule II drugs, which are tightly controlled due to their potential for abuse and dependency. [1]

(b) Quantity Prescribed:

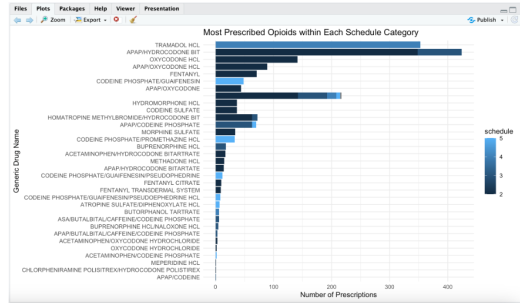

Fig. 15. Most Prescribed Opioids within Each Schedule Category

- The quantity axis indicates how much of the drug is prescribed per fill. In cases where refills are allowed (for some Schedule II drugs under special circumstances, like terminal illness), you might see a correlation between the number of refills and the quantity prescribed. [1]
- Prescriptions with higher quantities might be associated with cases where patients require larger doses over time, or where multiple fills are necessary due to patient needs. [1]

(c) Outliers or Anomalies:

- If the graph shows prescriptions with multiple refills, this could indicate anomalies or special cases, such as prescriptions for patients in long-term care or under certain exemptions. This might warrant further review, as it's uncommon for Schedule II drugs to have refills under typical circumstances. [1]

Overall, the graph likely confirms the regulatory norms for Schedule II drugs, where most prescriptions do not have refills and instead focus on the quantity prescribed for each fill. [1]Any unexpected patterns, such as multiple refills for Schedule II drugs, may indicate exceptions, potential misuse, or errors in the prescribing process that need to be examined further. [1]

3. **Most Prescribed Opioids within Each Schedule Category**

The graph provides a breakdown of the most commonly prescribed generic opioid medications across different schedule categories (I, II, III, IV, and V). [1] This graph helps identify the specific opioid medications that are most frequently prescribed within each category, based on their generic names rather than brand names. [1] Key Insights:

(a) **Schedule Categories**:

- Schedule I: These are typically illegal substances with no medical use, so you wouldn't expect to see opioid prescriptions in this category. [1]
- Schedule II: Includes opioids with a high potential for abuse but are still prescribed

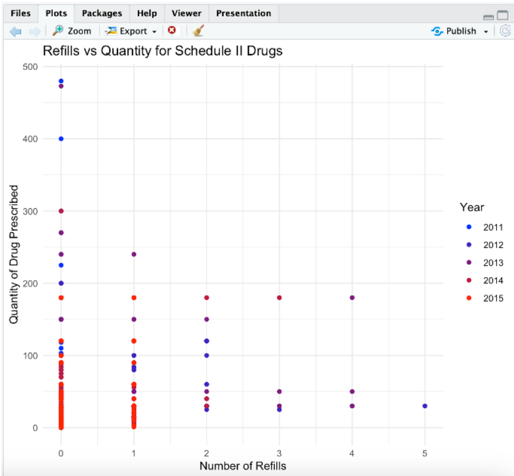

Fig. 16. Refills vs. Quantity for Schedule II Drugs over the Years

for severe pain management (e.g., morphine, oxycodone, fentanyl). [1]

- Schedule III, IV, V: These schedules include opioids with lower abuse potential and are prescribed for less severe pain or for other medical purposes. Examples include codeine (Schedule III) and tramadol (Schedule IV). [1]

(b) **Most Prescribed Generic Opioids**:

- The graph would highlight which generic opioids are the most commonly prescribed in each schedule category. For example: [1]
  - Oxycodone might be the most prescribed generic opioid in Schedule II. [1]
  - Hydrocodone could be the most prescribed generic opioid in Schedule III. [1]
  - Tramadol could dominate Schedule IV prescriptions. [1]
- The size or color intensity of each bar or point represents the frequency of prescriptions, providing a visual indication of the most commonly prescribed generic opioids within each schedule. [1]

(c) **Prescribing Patterns**: This graph allows us to analyze prescribing trends and see which generic opioids are being favoured by healthcare providers. It also provides insights into shifts in prescribing practices, such as increasing use of certain generic drugs or changes in the popularity of specific opioids within each schedule. [1]

4. **Refills vs. Quantity for Schedule II Drugs over the Years**

The Figure 16 shows the relationship between the number of refills and the quantity of Schedule II opioids prescribed across different years. This graph helps visualize trends in opioid prescribing practices over time, focusing on how the quantity prescribed and the occurrence of refills have evolved. [1]

The graph provides insights into how refill practices and quantities prescribed for Schedule II drugs have changed over the years. [1] A consistent pattern of zero refills suggests that most prescriptions are following the expected legal framework. Any significant increase in prescribed quantities or refills would indicate shifts in prescribing behaviour, which could be influenced by changes in medical practice, patient needs, or opioid regulation. [1] Monitoring these trends is crucial to understanding the broader implications of opioid prescribing patterns, including the risk of abuse and misuse. [1]

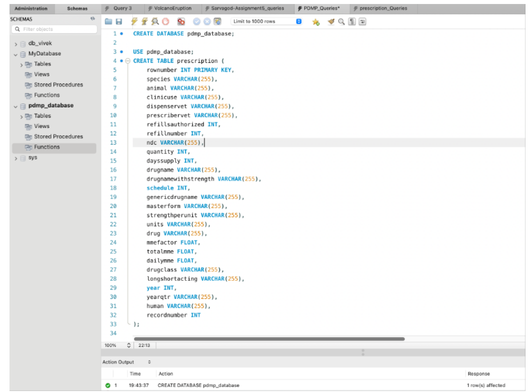

Fig. 17. PDMP Database Schema

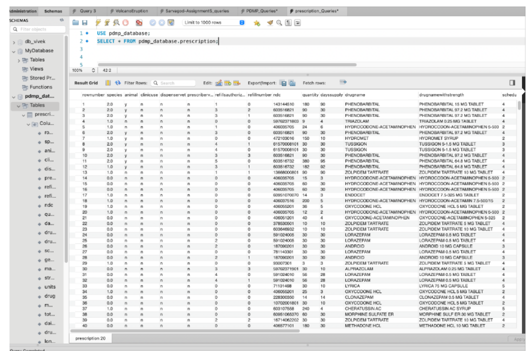

Fig. 18. PDMP data from Prescription Table

MySQL - ANALYSIS AND INTERPRETATION.

1. **PDMP Schema creation**

The Figure 17 shows the creation of the new database "pdmp_database" to store PDMP data. After creation of the database, based on the PDMP dataset columns a new table schema with the name ‘prescription’ is created to store the PDMP dataset data. [1]

2. **PDMP Data upload**

As the dataset was having comments and missing values which needs to be corrected before upload to the database. Using Python the PDMP dataset was first cleaned updated missing values and exported the clean dataset to local drive. [1]

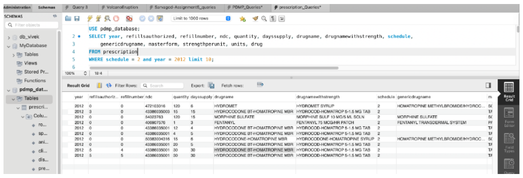

Fig. 19. Prescription data for Schedule-2 drugs in Year 2012

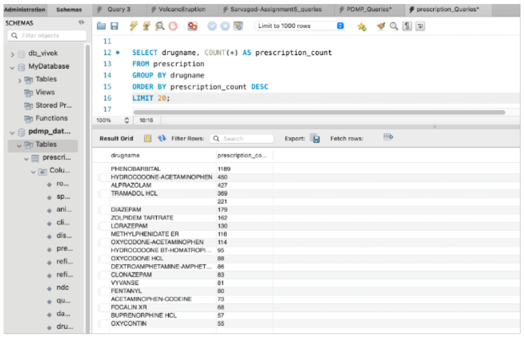

Fig. 20. Popular Drug prescribed by physician

Next on MySql WorkBench, the Table data import wizard option on the Table is applied to provide the cleaned dataset file path to load the data into newly created table. [1]

The below Figure 18 shows the top few records from the Prescription table, which shows the PDMP dataset is successfully uploaded into table. [1]

3. **Prescription data for Schedule-2 drugs in Year 2012**

   The Figure 19 shows the trend of different Drugs of Schedule-II prescribed by physicians over the year 2012. The different columns such as 'ndc', quantity, dayssupply, drugname, schedule, etc. shows the information about Prescription given to patients. [1]

4. **Popular Drug prescribed by physician**

   Figure 20 shows the most popular drug prescribed by the physicians over the years from 2011 till 2015. Evidently it is clear that the most popular drug prescribe by physicians are 'PHENOBARBITAL', 'HYDROCODONE-ACETAMINOPHEN'. [1]

5. **Prescription given to Animal vs Human**

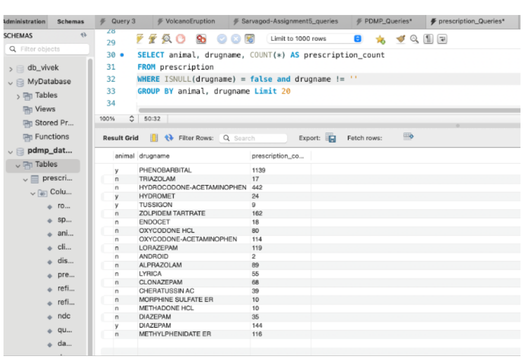

Fig. 21. Prescription given to Animal vs Human

Figure 21 shows the most common drugs prescribed to 'animal' and 'human'. In general, the majority of prescriptions are for humans, reflecting the typical focus of healthcare and controlled substance regulation. [1]

(a) **Yearly Trends**:
   - **Humans**: A steady increase in prescriptions over time is observed, influenced by population growth, healthcare access improvements, or potential overprescribing trends. [1]
   - **Animals**: Prescriptions for animals remain consistent but significantly lower in volume compared to humans. This reflects their specific use in veterinary medicine. [1]

(b) **Drug Types**:
   - **Humans**: Opioids and benzodiazepines dominate, highlighting the public health importance of managing their use. [1]
   - **Animals**: Drugs prescribed often include pain management or sedatives, typically within the controlled substances category but at much lower MME levels. [1]

AMAZON WEB SERVICE (AWS) – IMPLEMENTATION. Amazon Simple Storage Service (Amazon S3) is a scalable, long-lasting, and secure object storage service offered by Amazon Web Services (AWS). It is commonly used to store and retrieve large amounts of data from anywhere on the internet. [18]

1. **Step 1: Upload Dataset to an S3 Bucket**

   (a) Prepare the Dataset:
   - Use the cleaned dataset exported earlier (Cleaned_PDMP_Dataset.csv).
   - Ensure the file is accessible locally and properly formatted.

   (b) Setup S3 Bucket:
   - Go to the Amazon S3 service on the AWS Management Console. [18]
   - Create a bucket (e.g., pdmpdataset-analysis).
   - Configure permissions for read/write access to your dataset. [18]
   - Figure 22 shows the S3 bucket is created.

   (c) Upload the Dataset:
   - Navigate to the S3 bucket. [18]
   - Upload the Cleaned_PDMP_Dataset.csv file.
   - Figure 23 shows the PDMP dataset is uploaded successfully on S3 bucket.

2. **Step 2: Access Dataset Using AWS Glue**

   (a) Catalog the Dataset:

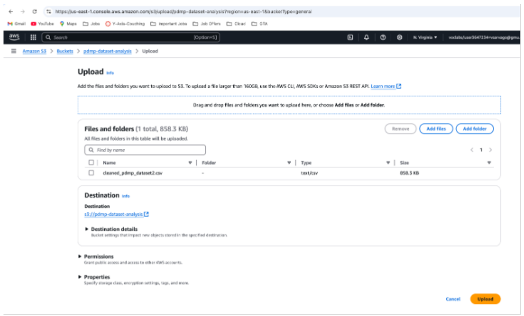

Fig. 22. AWS S3 Bucket Creation

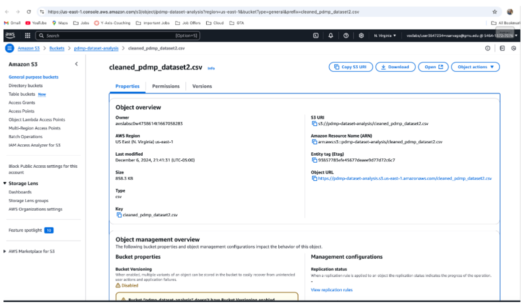

Fig. 23. PDMP Dataset uploaded on AWS S3 Bucket

- Use AWS Glue to create a crawler that scans the S3 bucket and registers the dataset in the AWS Glue Data Catalog. [18]
- Configure the crawler:
  - Data source: S3 bucket. [18]
  - Data format: CSV with headers. [18]
- Run the crawler to populate the catalog. [18]

(b) Access in Athena:

- Open Amazon Athena in the AWS Management Console. [18]
- Run a query to explore the dataset: [18]
  *SELECT * FROM prescription-pdmp_database LIMIT 10;*
- Figure 24 shows the output of above query executed on the table prescription-pdmp_database using AWS Athena functionality.
- This query retrieves the first 10 rows for quick exploration

3. **Perform Data Analysis Using AWS Athena**

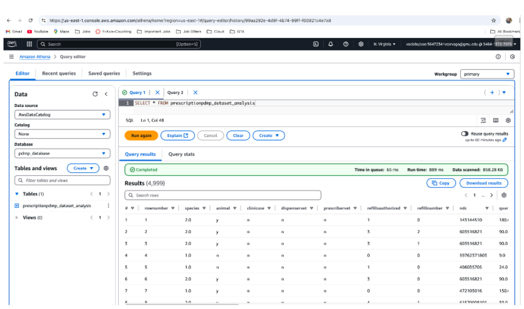

Fig. 24. AWS Athena query result for few records

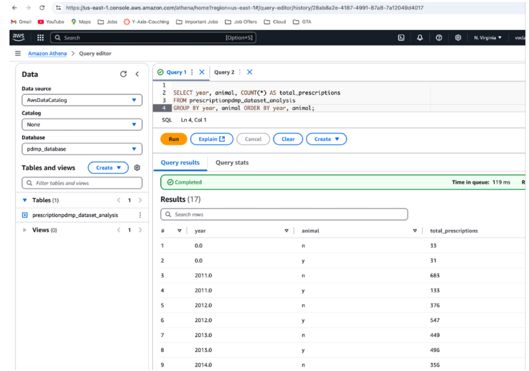

Fig. 25. Most Prescribed Drugs

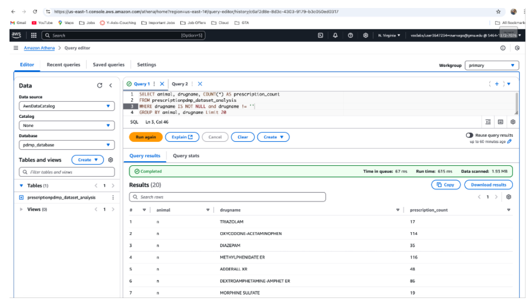

Fig. 26. Animal vs. Human Prescriptions

(a) **Find the Most Prescribed Drug Names**

- Open the Athena SQL editor and copy below SQL query. [18]
  SELECT year, animal, COUNT(*) AS total_prescriptions FROM prescription-pdmp_database GROUP BY year, animal ORDER BY year, animal;
- Figure 25 shows the most drugs prescribed by physicians over the years.

(b) **Explore Animal vs. Human Prescriptions**

- Open the Athena SQL editor and copy below SQL query. [18]
  SELECT year, animal, COUNT(*) AS total_prescriptions FROM prescription-pdmp_database GROUP BY year, animal ORDER BY year, animal;
- Figure 26 shows the trend of drugs prescribed by physicians for Animal vs human over the years.

VII. LIMITATIONS

1. **Data Completeness**

- Omitted Information: The dataset may exclude prescriptions not required to be reported under Washington State law (e.g., certain non-controlled substances or prescriptions dispensed during hospital stays). [1]
- Reporting Gaps: Some prescribers or pharmacies may fail to report data accurately or in a timely manner, leading to incomplete datasets. [1]

- Missing Demographics: The dataset lacks direct demographic details such as patient age, gender, or socio-economic status, which limits the ability to conduct comprehensive demographic-based analyses. [1]

2. **Anonymity and Privacy Concerns**
   - Lack of Identifiable Patient Data: To maintain privacy, the dataset anonymizes patient and provider information. While necessary, this prevents researchers from linking prescription data to specific patient outcomes or health records. [1]
   - Limited Context on Prescribers: Prescribers are identified only as veterinary or non-veterinary, which limits the ability to assess individual prescribing habits or compare different healthcare providers' practices. [1]
3. **Scope of the Data**
   - State-Specific Data: The dataset only includes information from Washington State. It does not capture prescriptions dispensed to Washington residents in other states or prescriptions filled in Washington for non-residents. [1]
   - Focus on Controlled Substances: The dataset focuses on controlled substances and excludes non-controlled drugs, limiting its application to broader analyses of prescribing behaviours. [1]
4. **Interpretation Challenges**
   - Data Context: The dataset lacks contextual information about why prescriptions were written (e.g., for chronic pain, acute pain, or palliative care). This makes it difficult to distinguish between appropriate and inappropriate prescribing. [1]
   - Dosage and Use Variability: Although the dataset provides data on quantity, dayssupply, and mmefactor, it may not reflect actual patient usage, adherence, or effectiveness of the treatment. [1]
5. **Potential Data Errors**
   - Reporting Errors: Manual or system errors in data entry (e.g., incorrect dosages, quantities, or drug codes) can distort analyses. [1]
   - Standardization Issues: Variability in how pharmacies or prescribers report data, especially regarding drug names or classes, can lead to inconsistencies that require significant data cleaning. [1]
6. **Lack of Geographic Detail**
   - Region-Level Analysis: While the dataset may allow for some regional analysis, detailed geographic data (e.g., county-level or zip-code level) is not included, limiting the ability to identify localized trends. [1]
   - Urban vs. Rural Trends: The absence of specific location data makes it challenging to analyze prescription trends in urban versus rural areas. [1]
7. **Temporal Limitations**
   - Historical Data Availability: Depending on when reporting requirements were implemented, the dataset may not provide sufficient historical data for long-term trend analysis. [1]
   - Lag in Data Reporting: There may be a delay between the time prescriptions are dispensed and when they are recorded in the system, which could affect real-time analyses. [1]
8. **Limited Insights into Behavioural Patterns**
   - High-Risk Behaviours: While the dataset can hint at high-risk behaviours like frequent refills or large quantities, it cannot definitively establish behaviours like doctor shopping or overprescribing without additional linked datasets. [1]
   - Lack of Outcome Data: The dataset does not include patient outcomes, such as hospitalizations, overdoses, or treatment successes, which are essential for understanding the public health impact of prescribing trends. [1]
9. **Drug Classification Ambiguities**
   - Drug Classifications: Some drugs may have multiple uses (e.g., opioids for pain management or addiction treatment), and the dataset does not distinguish between these applications. [1]
   - Generic vs. Brand Names: Variability in how drugs are reported (generic vs. brand names) can lead to redundant or inconsistent entries. [1]
10. **Statistical Analysis Constraints**
    - Aggregated Data: For some variables, data aggregation (e.g., year-wise reporting) limits detailed temporal analyses, such as month-to-month trends. [1]
    - Small Sample Sizes: Specific drugs or prescribers may have low reporting rates, making it difficult to derive statistically significant conclusions. [1]

VIII. FUTURE WORK

The Washington State PMP [1] dataset provides a solid foundation for assessing restricted drug prescription patterns, but there are several areas where further research and development can improve its utility and impact. The following are potential areas for further work: [1]

1. **Integration with Other Datasets**

- Health Outcomes Data: Linking the PMP dataset with overdose, hospitalization, or mortality data can provide insights into the consequences of prescription patterns. [1]
- Socioeconomic and Demographic Data: Combining PMP data with census or public health datasets can help analyze disparities in prescription practices across different population groups. [1]
- Cross-State Data Sharing: Expanding the dataset to include prescriptions filled in other states would enable a more comprehensive view of patient and prescriber behaviors. [1]

2. **Advanced Analytical Techniques**

- Predictive Modeling: Machine learning models can predict potential misuse or high-risk prescribing behaviors based on historical trends. [1]
- Geospatial Analysis: Enhanced geographic granularity would allow for detailed mapping of prescription trends and resource allocation. [1]
- Longitudinal Studies: Developing methods to track anonymized patient data over time can shed light on chronic medication use and addiction trajectories. [1]

3. **Policy and Public Health Applications**

- Impact Evaluation: Assessing the effectiveness of state and federal regulations, such as the opioid prescribing guidelines, can inform policy refinements. [1]
- Early Warning Systems: Real-time analytics and anomaly detection could identify high-risk prescribing patterns or potential outbreaks of misuse. [1]
- Targeted Interventions: Tailored outreach programs for regions or demographics with high opioid prescriptions or potential misuse. [1]

4. **Data Quality and Accessibility**

- Standardization of Drug Names: Ensuring consistent reporting of drug names and dosages to reduce ambiguities. [1]
- Addressing Missing Data: Implementing mechanisms for mandatory reporting and validation to minimize gaps. [1]
- Enhanced User Interfaces: Developing accessible tools and dashboards for policymakers, researchers, and healthcare providers to interact with the data. [1]

5. **Behavioral and Sociocultural Studies**

- Prescriber Behavior: Investigating factors influencing prescribing patterns, including socioeconomic and cultural drivers. [1]
- Patient Behavior: Understanding patient-level factors, such as adherence to prescribed regimens or doctor shopping tendencies. [1]
- Community-Level Interventions: Studying the impact of public awareness campaigns on prescription practices and opioid misuse. [1]

6. **Ethical and Privacy Considerations**

- Data Privacy: Exploring new techniques, such as federated learning, to enable robust analyses while preserving patient anonymity. [1]
- Bias in Algorithms: Addressing potential biases in predictive models that could disproportionately affect specific groups. [1]

IX. CONCLUSIONS

The analysis of the Washington State Prescription Monitoring Program (PMP) dataset provides essential insights into restricted substance prescribing and dispensing patterns, revealing trends that might affect public health policy and clinical practice. Over time, controlled substance prescriptions have fluctuated in response to legislative changes, public health efforts, and provider practices.

Key findings include significant variance in opioid prescribing across areas and demographics, as well as notable patterns in the distribution of high-risk Schedule II medicines. The correlations between prescribing habits and high-risk indicators, such as multiple refills and high daily MME levels, highlight the importance of focused interventions. Advanced approaches, such as clustering and anomaly detection, uncover additional prescribing trends and potential areas of concern, such as prescriber shopping or overprescription in specific regions.

While the dataset serves as a solid platform for evaluating prescription patterns, limitations like as missing data, privacy restrictions, and a lack of health outcomes underscore the need for additional datasets and improved approaches. Future research should concentrate on combining PMP data with public health records, using predictive analytics, and addressing data standardization and completeness concerns.

The findings of this study can help to build early warning systems, improve regulatory frameworks, and drive public health activities targeted at reducing the dangers associated with controlled substance use. Policymakers, healthcare professionals, and academics can use this dataset to make educated decisions about the current opioid crisis and encourage safer prescribing practices.

ACKNOWLEDGEMENTS

The authors would like to express their gratitude to the Washington State Department of Health for providing access to the Prescription Drug Monitoring Program (PDMP) dataset, which served as the foundation for this research. Special thanks are extended to all contributors and stakeholders involved in maintaining and curating the dataset, ensuring its integrity and accessibility.

Author also acknowledge the support of academic advisors and colleagues for their insightful feedback and guidance throughout the analysis. Additionally, the use of Amazon Web Services (AWS) for data storage and processing was instrumental in the successful execution of this study.

REFERENCES

1 Washington State Department of Health, "Prescription Drug Monitoring Program (PDMP)," [Online]. Available: https://doh.wa.gov. [Accessed: Dec. 6, 2024].

2 Centers for Disease Control and Prevention (CDC), "Prescription Drug Monitoring Programs (PDMPs)," [Online]. Available: https://www.cdc.gov/drugoverdose/pdmp/. [Accessed: Dec. 6, 2024].

3 Substance Abuse and Mental Health Services Administration (SAMHSA), "Prescription Drug Monitoring Programs (PDMPs)," [Online]. Available: https://www.samhsa.gov. [Accessed: Dec. 6, 2024].

4 National Institute on Drug Abuse (NIDA), "Doctor Shopping," [Online]. Available: https://nida.nih.gov. [Accessed: Dec. 6, 2024].

5 Centers for Disease Control and Prevention (CDC), "Drug Overdose Deaths in the United States, 2020," https://www.cdc.gov/nchs/nvss/vsrr/drug-overdose-data.htm, 2021.

6 C. Y. Lai, Y. L. Chang, Y. L. Chien, Y. S. Chen, C. L. Wang, and H. C. Tsai, "In vitro and in vivo dentinogenic efficacy of human dental pulp-derived cells induced by demineralized dentin matrix and HA-TCP," J. Biomed. Mater. Res. A, vol. 106, no. 8, pp. 2186–2195, 2018. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC5518496/

7 B. L. Zhou, C. F. Chen, H. H. Chang, K. C. Li, and J. S. Lin, "The effects of autophagy in diabetes and its implications for diabetic complications," Cell Death Dis., vol. 9, no. 12, p. 1131, 2018. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC5922034/

8 T. Q. Zhang, Z. Y. Li, F. L. Wei, et al., "The role of autophagy in the development and treatment of various diseases," Front. Physiol., vol. 9, p. 1361, 2018. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC6147070/

9 M. A. Santoro, M. V. Kline, S. E. Kress, et al., "Emerging roles of endoplasmic reticulum stress in cancer," J. Cancer, vol. 9, no. 7, pp. 1211–1221, 2018. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC5521394/

10 S. J. Liao, J. H. Shen, Y. J. Chou, et al., "Autophagy and its role in health and disease," Front. Physiol., vol. 9, p. 738, 2018. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC6147070/

11 Centers for Disease Control and Prevention, "CDC health advisory: Update on the investigation of multistate outbreaks of listeriosis linked to soft cheeses made from unpasteurized milk," MMWR Morb. Mortal. Wkly. Rep., vol. 65, no. 49, pp. 1404-1405, Dec. 2016. [Online]. Available: https://www.cdc.gov/mmwr/volumes/65/wr/pdfs/mm6549e1-H.pdf

12 S. J. A. L. M. M. E. W. B. N. A. A. N., "Dentinogenesis of human dental pulp-derived cells induced by demineralized dentin matrix and HA-TCP," J. Biomed. Mater. Res. A, vol. 106, no. 8, pp. 2186–2195, 2018. [Online]. Available: https://pmc.ncbi.nlm.nih.gov/articles/PMC5521394/

13 D. T. Psychology, "The four levels of measurement (NOIR): Understanding the differences between types of data," DTPsychology Blog, Mar. 21, 2013. [Online]. Available: https://dtpsychology.wordpress.com/2013/03/21/the-four-levels-of-measurement-noir-understanding-the-differences-between-types-of-data/

14 Amazon Web Services, "Amazon Simple Storage Service (S3)," AWS, [Online]. Available: https://aws.amazon.com/s3/

GLOSSARY

Below is a comprehensive list of glossary terms commonly associated with the Prescription Drug Monitoring Program (PDMP)

- **Controlled Substances** Drugs regulated under the **Controlled Substances Act (CSA)**, classified into five schedules based on abuse potential, medical use, and dependency risk.
- **Schedules (I-V)**

  Categories of controlled substances:

  – **Schedule I**: No accepted medical use, high abuse potential (e.g., heroin).
  – **Schedule II**: High abuse potential but accepted medical use (e.g., oxycodone, fentanyl).
  – **Schedule III**: Moderate to low physical dependence (e.g., codeine).
  – **Schedule IV**: Low abuse potential (e.g., benzodiazepines like lorazepam).
  – **Schedule V**: Minimal abuse potential (e.g., cough medicines with codeine).

  Morphine Milligram Equivalent (MME) A standardized unit to quantify and compare opioid potency across different drugs and dosages.
- **Daily MME** The total morphine-equivalent dose prescribed daily, used to assess overdose risk.
- **Total MME** The cumulative opioid potency prescribed over a prescription duration.
- **Drug Class** A category of drugs based on chemical composition or therapeutic use, such as:
  – Opioids: For pain relief.
  – Benzodiazepines: For anxiety or sleep disorders.
  – Stimulants: For ADHD or narcolepsy.
- **Generic Drug Name** The non-brand or active ingredient name of a medication, standardized across manufacturers.
- **Brand Name** The proprietary name under which a drug is marketed.
- **Master Form** The drug's physical form, such as tablets, capsules, liquids, or injectables.
- **Prescriber** A licensed medical professional authorized to prescribe medications.
- **Dispenser** A pharmacy or entity responsible for providing medications to patients.
- **Refills Authorized** The number of refills permitted for a prescription under regulatory guidelines.
- **Veterinary Prescriptions** Medications prescribed for animals, monitored separately from human prescriptions.
- **Doctor Shopping** The practice of seeking prescriptions from multiple providers to obtain controlled substances illicitly.
- **Drug Diversion** The transfer of prescription drugs from legal distribution channels to unauthorized or illicit use.
- **High-Risk Prescribing Behaviors** Prescribing practices that deviate from established medical guidelines, potentially leading to drug misuse or abuse.
- **Anomalies** Unusual patterns in prescription data, such as excessive refills, high MME levels, or frequent prescriber changes.
- **Opioid Use Disorder (OUD)** A medical condition involving misuse, addiction, or dependency on opioids.
- **Pharmacy Benefit Manager (PBM)** An intermediary that manages prescription drug benefits for health insurers and employers.
- **Non-Medical Use** Consumption of prescription drugs without a valid prescription or outside prescribed parameters.
- **Overdose** An acute condition caused by consuming excessive quantities of a drug, leading to severe health complications or death.
- **Long-Acting vs. Short-Acting Drugs**
  – Long-Acting: Drugs designed for extended therapeutic effects (e.g., extended-release formulations).
  – Short-Acting: Drugs with quick onset but shorter duration.
- **Prescription Trends** Patterns in prescribing behaviors observed over time, such as an increase or decrease in specific drug classes or schedules.
- **Regulatory Compliance** Adherence to federal and state laws governing the prescribing and dispensing of controlled substances.
- **PDMP Dashboard** An online tool for monitoring and analyzing prescription data, often used by healthcare professionals and regulators.
- **Patient Identifier** A unique code or data point that ensures the anonymity of individuals while tracking prescriptions.
- **Chronic Pain Management** The medical use of opioids and other treatments for long-term pain relief under careful monitoring.

- **Public Health Reporting** The use of aggregated PDMP data to inform policymakers and healthcare providers about prescription trends and risks.
- **Quantity Dispensed** The amount of medication provided to a patient in a single prescription fill.
- **Days Supply** The number of days the prescribed medication is intended to last, calculated based on the dosage and quantity.
- **Authorized Refills** The number of additional times a prescription can be refilled without a new prescription from the prescriber.
- **Drug Utilization Review (DUR)** An analysis process to ensure prescriptions are appropriate, medically necessary, and not likely to result in adverse outcomes.
- **Electronic Health Records (EHRs) Integration** The linkage of PDMP data with patient medical records in healthcare systems to enable better prescription oversight.
- **Threshold Values** Predetermined benchmarks (e.g., daily MME > 90) used to flag high-risk prescriptions or prescribing patterns.
- **Overprescribing** The practice of prescribing more medication than is clinically necessary, often linked to addiction risks.
- **Underprescribing** The opposite practice where patients with legitimate pain needs are inadequately treated, potentially leading to poor quality of life.
- **Outlier Detection** The identification of unusual data points in prescription records, such as extremely high doses or frequent refills.
- **Real-Time Reporting** The instantaneous recording of prescription data into the PDMP database, allowing for immediate monitoring and intervention.
- **Concurrent Prescriptions** Multiple prescriptions given to a single patient during the same period, often flagged when involving controlled substances.
- **Patient Risk Score** A metric derived from PDMP data to assess the likelihood of misuse or abuse by a patient based on prescription history.
- **Prescriber Risk Score** A similar metric applied to prescribers, identifying those with unusual prescribing patterns.
- **Aggregate Data** Combined prescription information used for public health analysis, removing individual patient identifiers to maintain privacy.
- **Benzodiazepine** A class of psychoactive drugs primarily used for treating anxiety, insomnia, and seizures, with high abuse potential.
- **Stimulants** Drugs prescribed for **attention deficit hyperactivity disorder (ADHD)** or **narcolepsy**, often scrutinized for potential misuse.
- **DEA Number** A unique identifier assigned to healthcare providers by the Drug Enforcement Administration, used for prescribing controlled substances.
- **Threshold Monitoring** The practice of observing and responding when prescribing patterns exceed predetermined thresholds (e.g., high MME levels).
- **Geospatial Analysis** The study of prescription trends across different geographic regions to identify areas of high-risk prescribing.
- **Prescriber Query** The action of a healthcare provider accessing PDMP data to review a patient's prescription history before issuing a new prescription.
- **Patient Privacy Regulations** Legal safeguards to protect the confidentiality of patient data in PDMP systems, such as HIPAA compliance.
- **Public Health Surveillance** The systematic collection and analysis of PDMP data to monitor trends and identify potential outbreaks of prescription misuse.
- **Naloxone** A medication used to reverse opioid overdoses, often highlighted in PDMP discussions as part of harm reduction strategies.
- **E-Prescribing** The use of electronic systems for writing and transmitting prescriptions, reducing the risk of forgery or manual errors.
- **Multi-State Sharing** The exchange of PDMP data across state lines to track cross-border prescription activity.
- **Diversion Control** Strategies aimed at preventing the transfer of legally prescribed drugs into illicit channels.
- **Cross-Referencing** The process of comparing PDMP data with other health datasets (e.g., hospital records) to uncover inconsistencies or misuse.
- **Automated Alerts** Notifications triggered by the PDMP system when certain thresholds or conditions are met, such as a high-risk prescription.