

Introduction

Opioid and Stimulant toxicity has become a growing public health crisis in Canada, with significant impacts. The rates of overdose with Opioids such as Fentanyl and Stimulants such as Cocaine and MDMA are becoming an alarming issue, whether it results in death or hospitalization.

Additionally, challenges once associated with poorer nations, like homelessness and street violence, are now affecting countries like Canada, Norway, and Sweden. As a data student, this topic offers valuable opportunities for analysis, allowing us to use past and present data to identify patterns and trends for future insights.

Guiding Questions

Looking at the dataset, some questions arise rapidly, and exploratory data analysis will give us some ideas of paths to follow to extract valuable information from this historical data.

1. What is the growth percentage of incidents by Province from 2016 to 2023?

This answer will help us drive the next steps of the analysis and check if this is a particular situation of geography or if the behaviour is similar throughout the country.

2. Is there a specific group (Age, Sex) in which the growth was more prevalent?

This question aims to dive deep into the reasoning behind the crisis and elaborate possible motivations (e.g.: inexperience for young adults, work/family issues for middle-aged, pain/loneliness for the elderly, etc.) for this abuse.

3. Is the weather relevant for the number of incidents?

Most of Canada has distinct weather conditions throughout the year, mainly Q1 (usually extremely cold) and Q3 (usually hot). If there is a particular situation happening in a part of the year, we can derive some reasoning behind the crisis and complement the answers for question 2.

4. What is the share of intentional deaths/hospitalizations overall? Is it increasing or decreasing?

These substances are so powerful that accidents may happen even with “experienced” users. Investigating this will give us a direction for a possible conclusion to the study.

Datasets

To perform this task, we will split our dataset into two subsets:

1. Opioid and Stimulant-related Harms in Canada published by Health Canada (publicly available).

The dataset consists of a tabular file with 12 columns and approx. 28,000 rows that should be wrangled and distributed into other tables/data frames to perform the analysis in a correct manner.

Tasks

The dataset is comprehensive but structured in a way that we will need to perform some wrangling to obtain the best results, such as:

1. Column A (named Substance): contains two different categories - Opioids and Stimulants by Ruby.
2. Column G (Time_Period): contains data by Year or by Quarter that should not be analyzed together with the risk of duplication and misleading by Joao.
3. Column K (Unit): contains data related to the absolute number of events, the percentage of a specific population or the rate per 100,000 habitants for each Region. These three metrics should have a dedicated data frame for each by Ruby and Joao.

We intend to use visual exploration, such as bar and line charts (for time-series data), as part of the EDA to answer the guiding questions.

As part of the learning process, we will use libraries such as Matplotlib or Seaborn for the visual exploration, Pandas and NumPy for the data manipulation and Jupyter Notebook for calculations, trend identification and analysis.

References

Public Health Agency of Canada. *Apparent opioid and stimulant toxicity deaths: Surveillance of opioid- and stimulant-related harms in Canada.* Available at:

<https://health-infobase.canada.ca/substance-related-harms/opioids-stimulants/>

Libraries: Pandas, NumPy, Matplotlib, Seaborn