Emergency Department Drug Abuse Visits

Laura Eshee

Capstone Project 1

Background

- Addiction is a chronic disease characterized by drug seeking and use that is compulsive, or difficult to control, despite harmful consequences.

- Drug abuse occurs when a person takes a substance, whether illegal, prescribed or over the counter, for purposes other than those in which they are meant to be used, or when a person takes large quantities of the substance.

- Typically, the person is using the drug to alter his or her mood or feel better and not for a health reason.

Statistics

- Almost 21 million Americans have at least one addiction, yet only 10% of them receive treatment.[2]
- Drug overdose deaths have more than tripled since 1990.[2]
- From 1999 to 2017, more than 700,000 Americans died from overdosing on a drug.[2]
- More than 90% of people who have an addiction started to drink alcohol or use drugs before they were 18 years old.[2]
- Americans between the ages of 18 and 25 are most likely to use addictive drugs.[2]
- Alcohol and drug addiction cost the U.S. economy over $600 billion every year.[2]
- During 2008–2011, an average of 1.1 million emergency department (ED) visits were made each year for drug poisoning, with a visit rate of 35.4 per 10,000 persons.[3]
- About one-quarter (24.5%) of drug-poisoning ED visits resulted in hospital admission.[3]

Problem Statement

- Emergency Departments need to be prepared to deal with drug abuse cases

- Expected types and percentages of drug abuse related visits

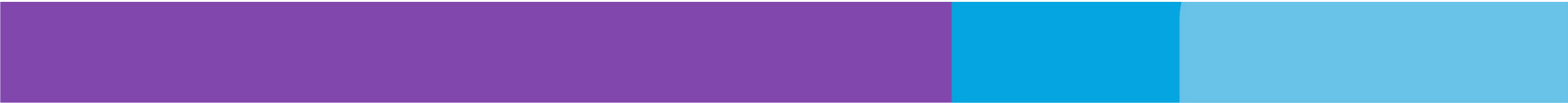- Staffing, medications, supplies
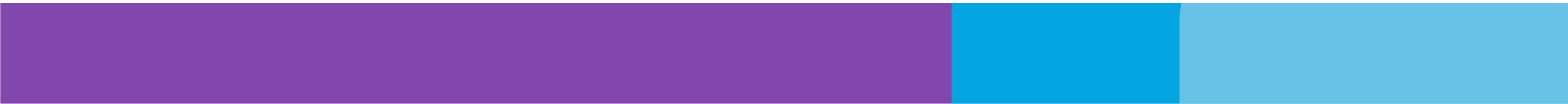
Data Set

- Drug Abuse Warning Network (DAWN)

  - A public health surveillance system that monitors drug abuse related visits to emergency departments in hospitals in large metro areas across the US

  - Monitors trends in drug misuse and abuse

  - Identifies the emergence of new substances and drug combinations

  - Estimates the impact of drug misuse and abuse on the Nation's health care system
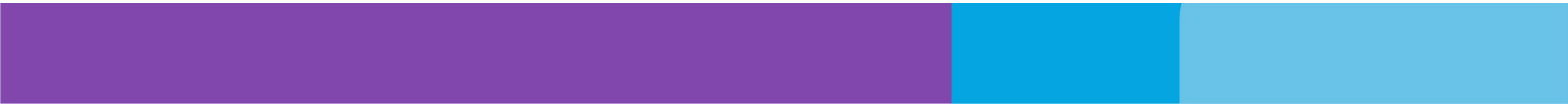
Data Set

- DAWN 2011 Data Set includes:
  - Demographics
  - Drugs involved (up to 22 in a case)
  - Toxicology confirmation
  - Route of administration
  - Type of case
  - Disposition of patient
  - Website: https://www.datafiles.samhsa.gov/study-dataset/drug-abuse-warning-network-2011-dawn-2011-ds0001-nid13747

Data Set Preparation (Data Wrangling)

Size and Columns

- The data set was examined to see its size and column names

  - 229,221 rows

  - 284 columns

  - It was decided to remove the 'sdled' and 'CATID' columns since their inclusion is beyond the scope of this project

  - Reduced data set has 84 columns
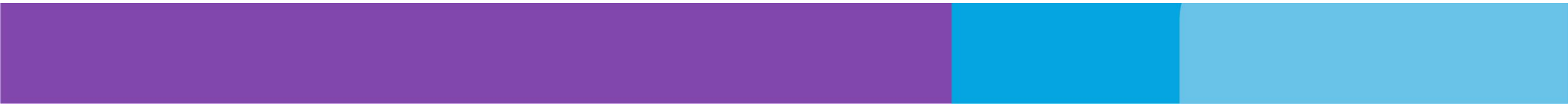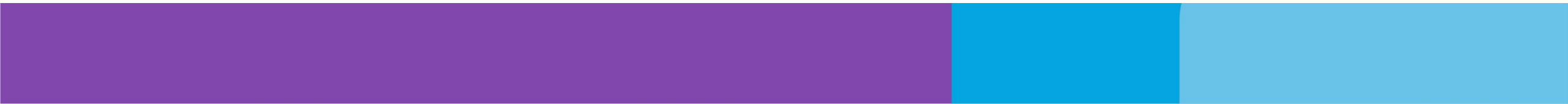
Descriptive Statistics

	STRATA	PSU	REPLICATE	CASEWGT	PSUFRAME	YEAR	QUARTER	NUMSUBS
count	229211.000000	229211.000000	229211.000000	229211.000000	229211.000000	229211.0	229211.000000	229211.000000
mean	24.681551	109.610839	1.500028	22.107901	73.348304	2011.0	2.495997	1.584104
std	13.331152	64.444097	0.500001	68.403862	211.852785	0.0	1.104806	1.163778
min	1.000000	1.000000	1.000000	0.938440	2.000000	2011.0	1.000000	1.000000
25%	13.000000	53.000000	1.000000	2.714999	8.000000	2011.0	2.000000	1.000000
50%	25.000000	109.000000	2.000000	4.190787	10.000000	2011.0	3.000000	1.000000
75%	35.000000	165.000000	2.000000	7.148615	17.000000	2011.0	3.000000	2.000000
max	51.000000	233.000000	2.000000	862.824350	1215.000000	2011.0	4.000000	22.000000

Unique Values

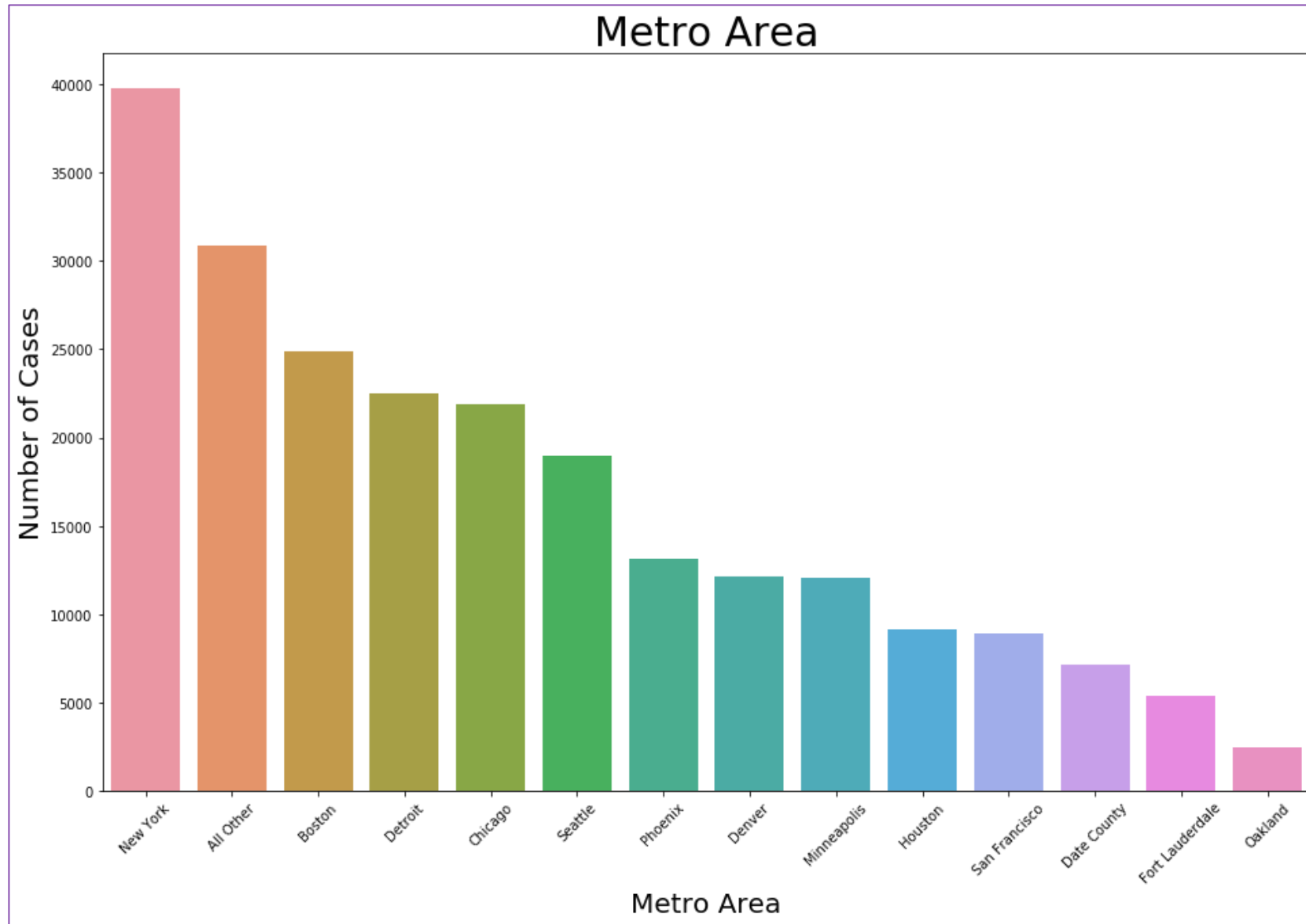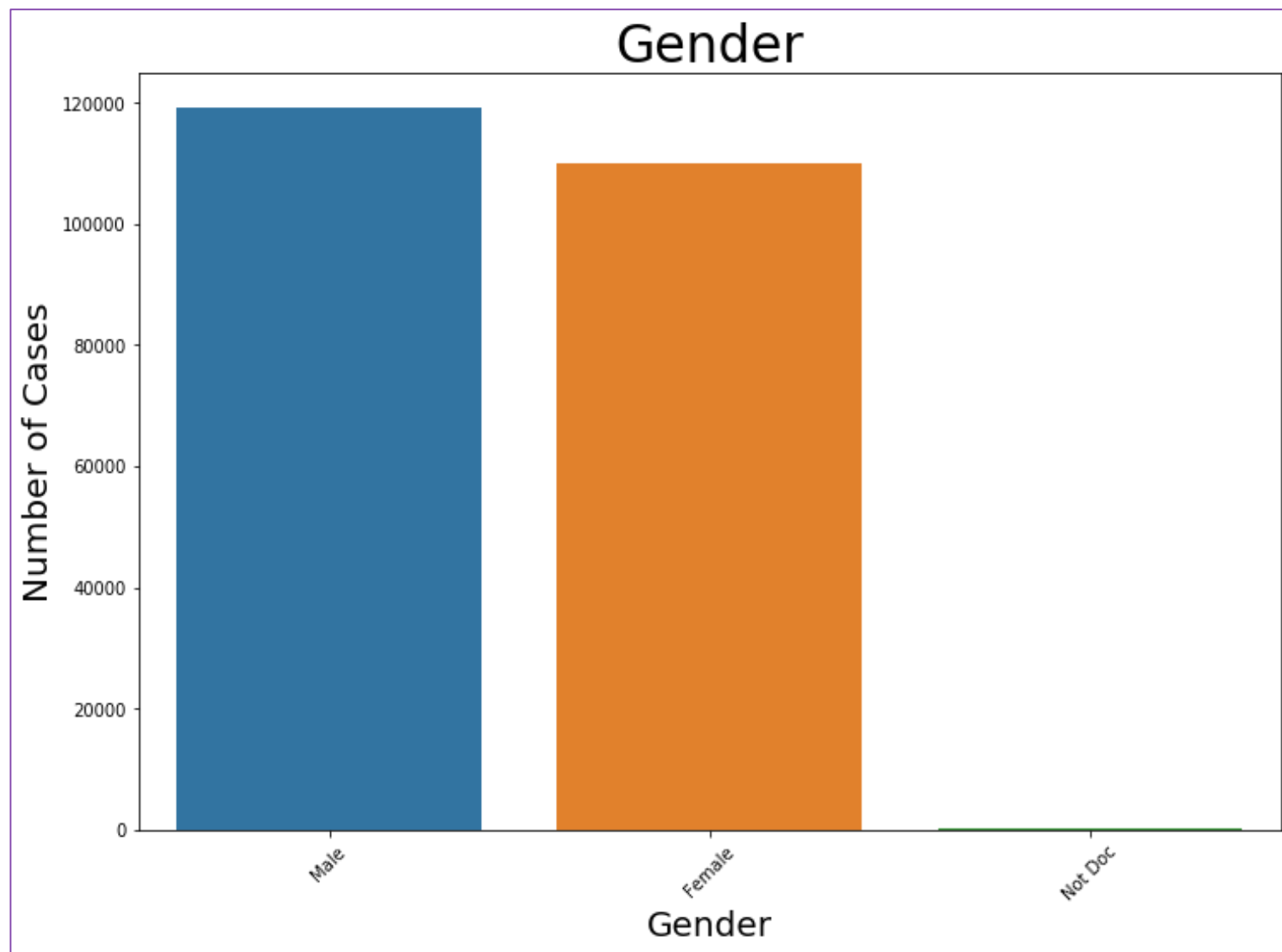| Feature	Count
CASEWGT	2,931
DRUGID_1	1,604
DRUGID_2	1,074
…	…
ALLABUSE	2
YEAR	1

Other Attributes

- Most of the variables are categorical—they can be divided into groups. Examples include race, gender, age group
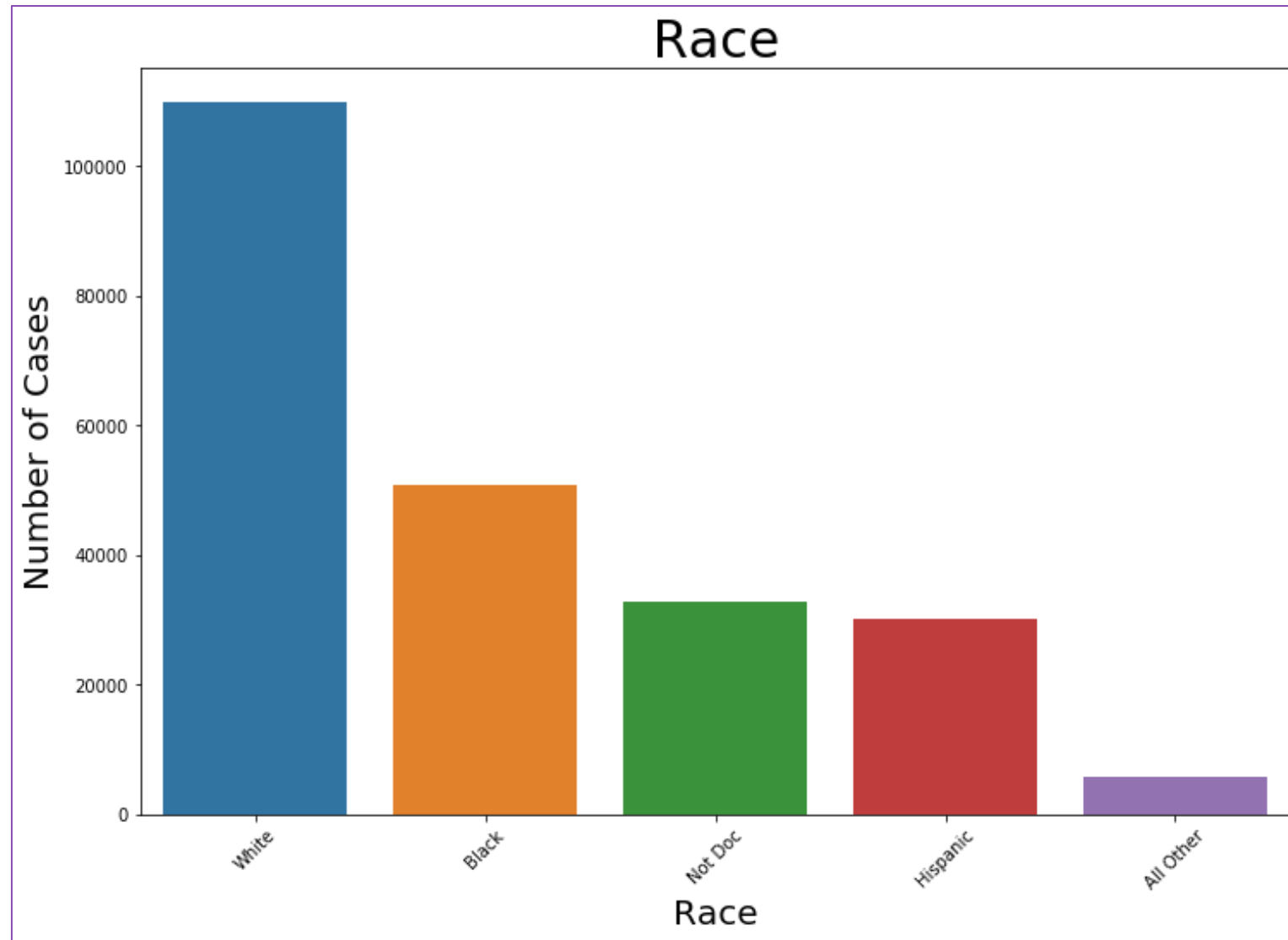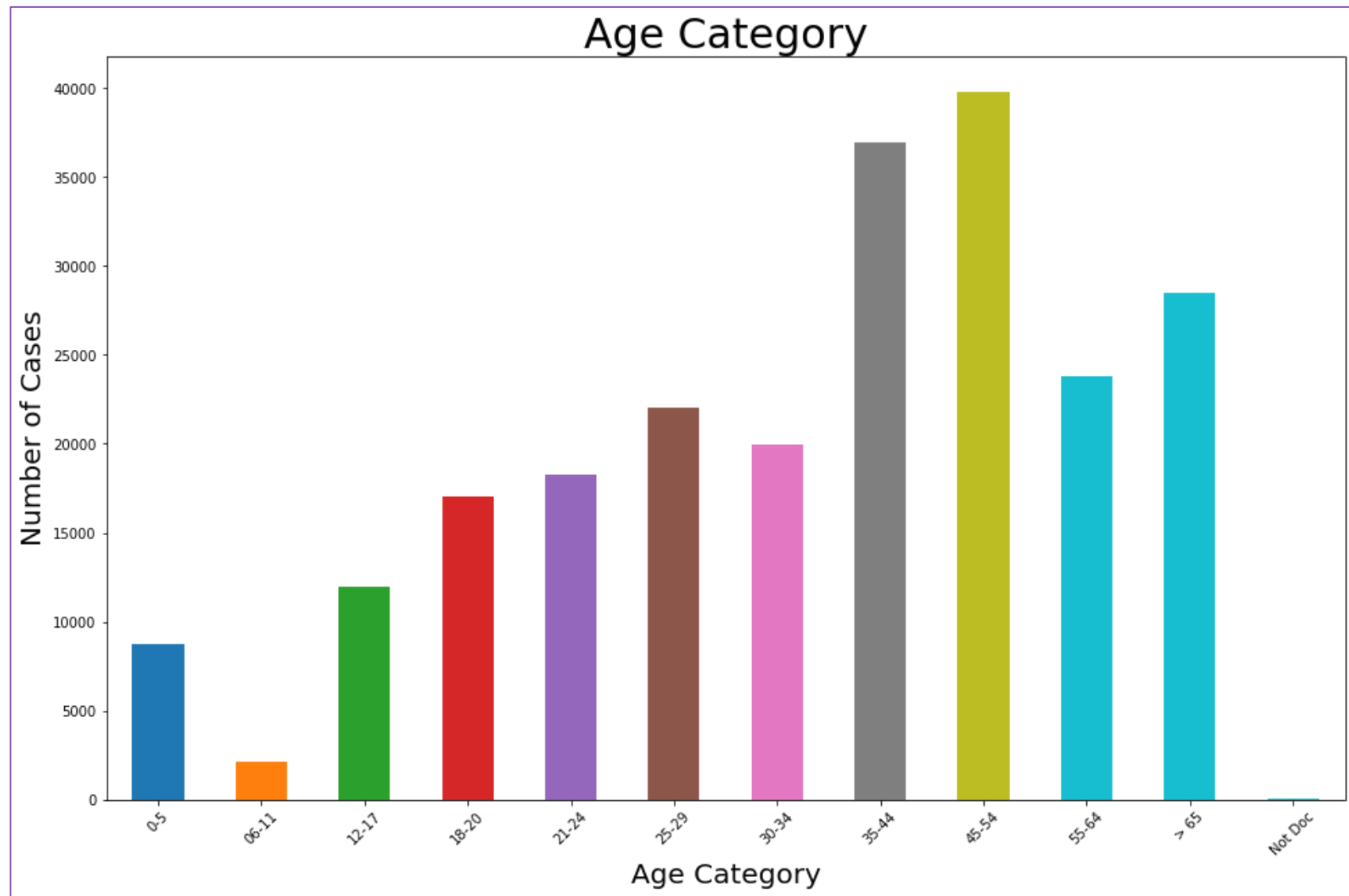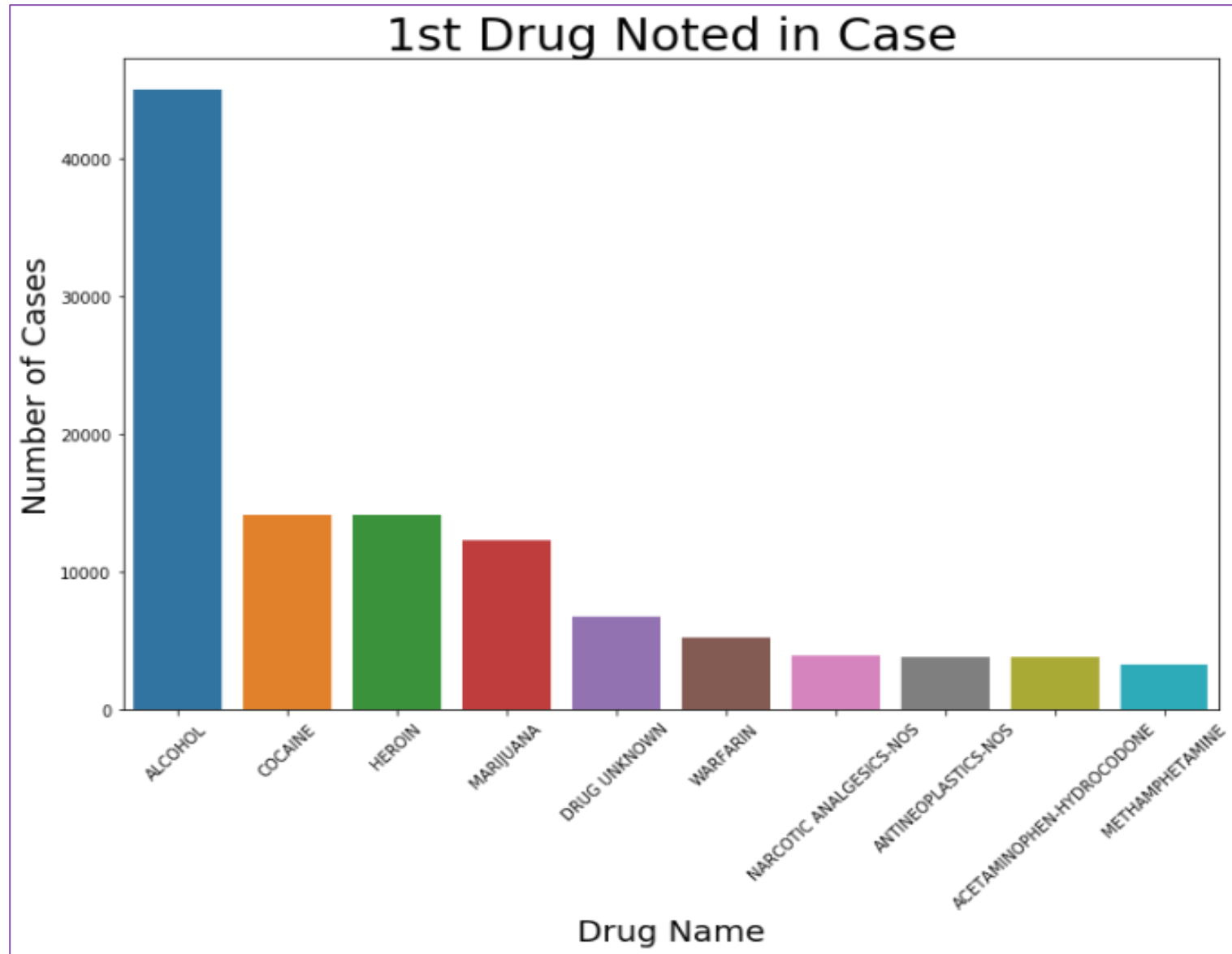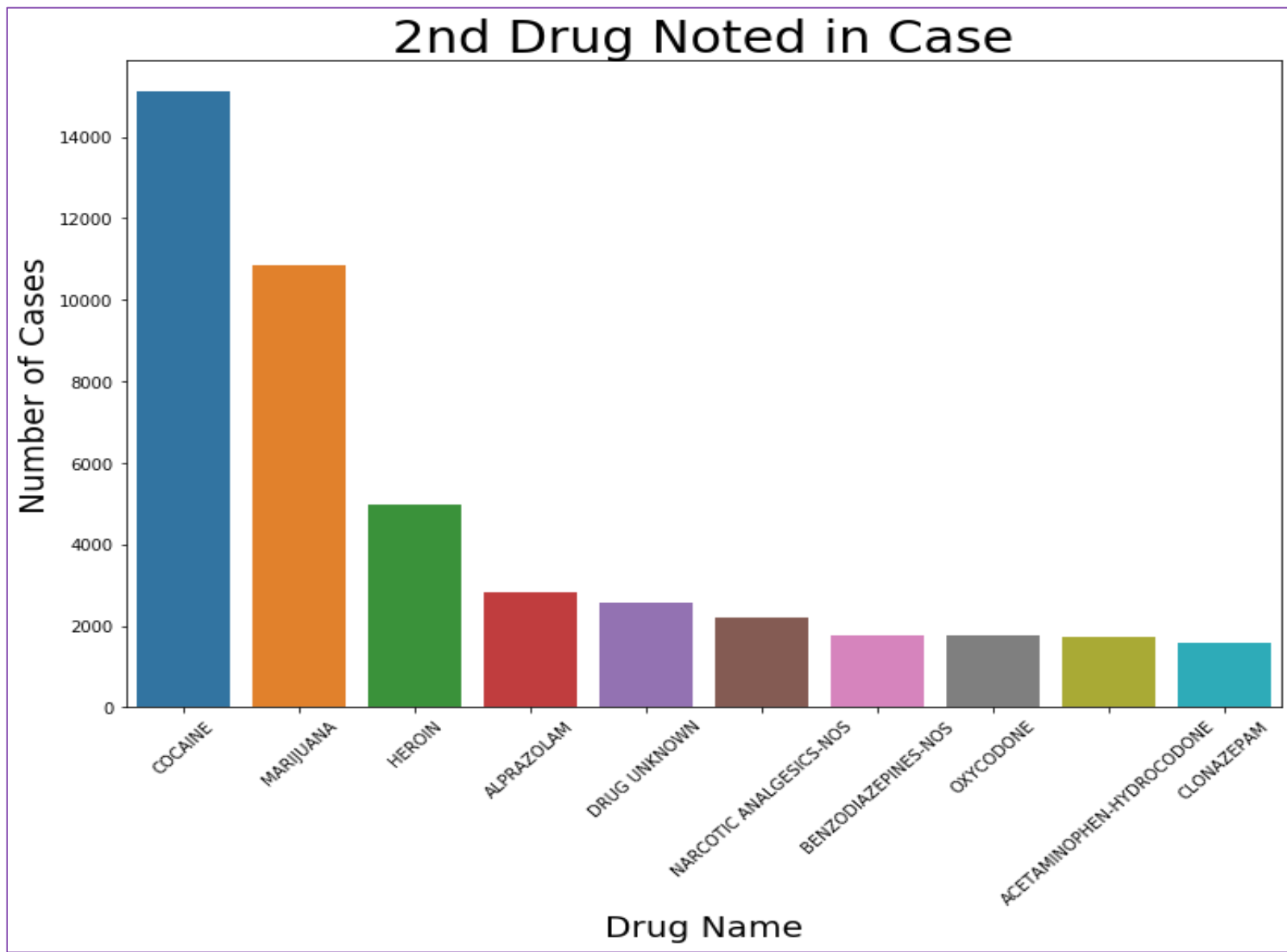
- No missing values

Data Story

Metro Area

1st Drug Noted in Case

Number of Cases (y-axis): 0, 10000, 20000, 30000, 40000

Drug Name (x-axis): ALCOHOL, COCAINE, HEROIN, MARIJUANA, DRUG UNKNOWN, WARFARIN, NARCOTIC ANALGESICS-NOS, ANTINEOPLASTICS-NOS, ACETAMINOPHEN-HYDROCODONE, METHAMPHETAMINE

2nd Route of Drug Ingestion Noted in Case
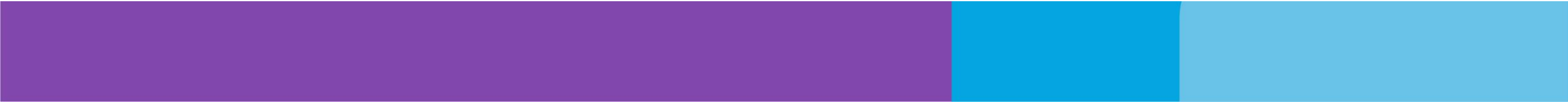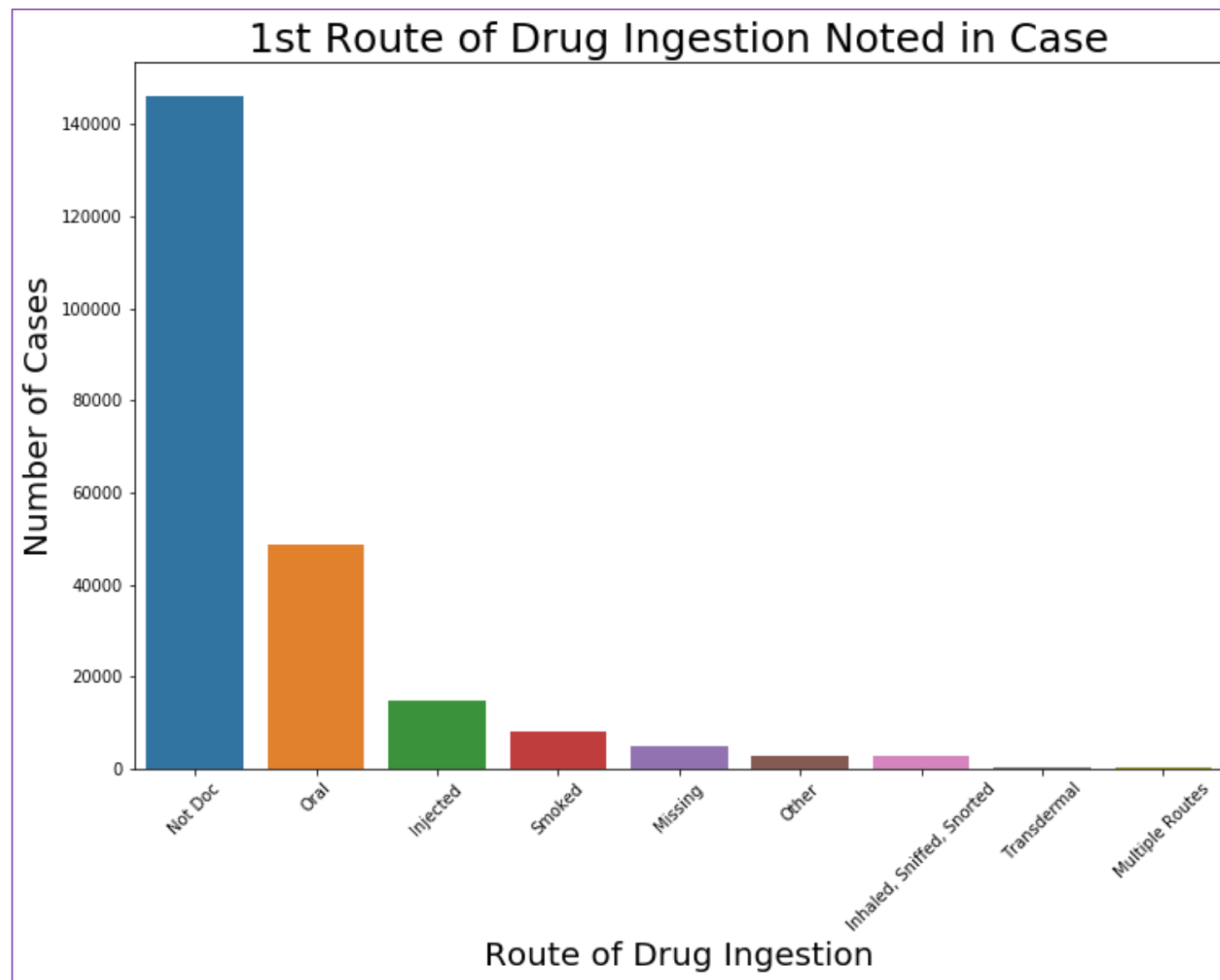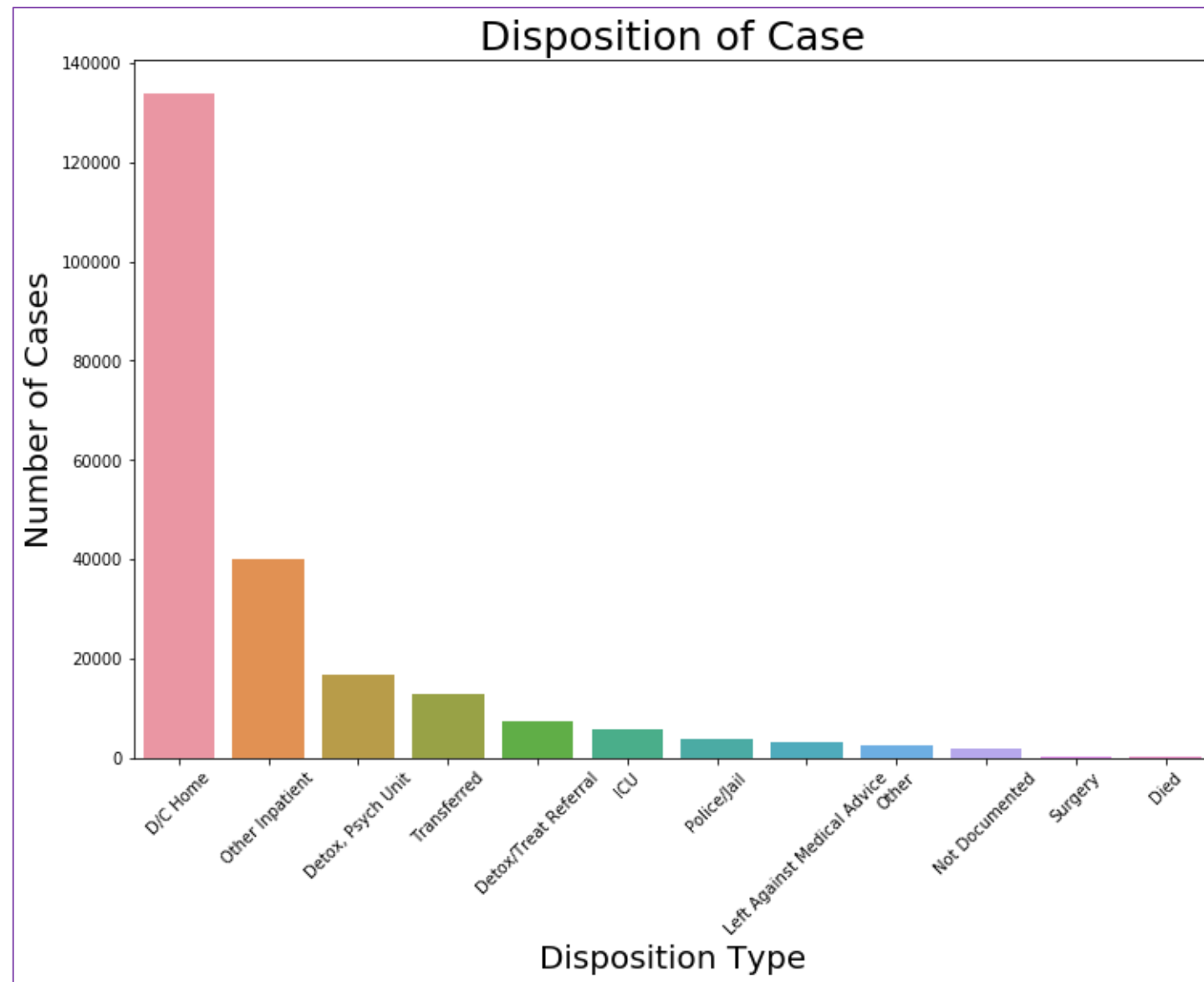
Data Story Summary

- The time of year is evenly distributed.
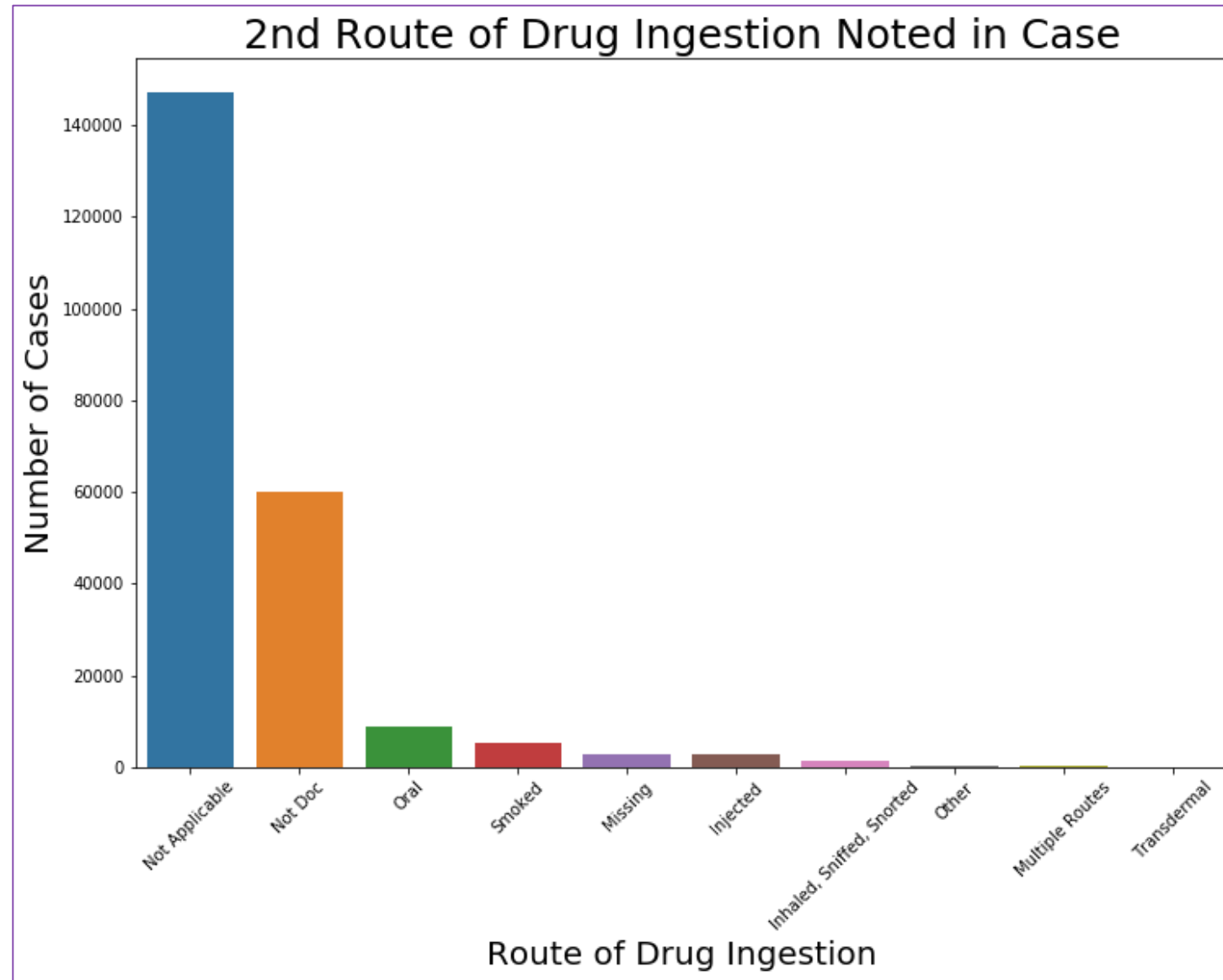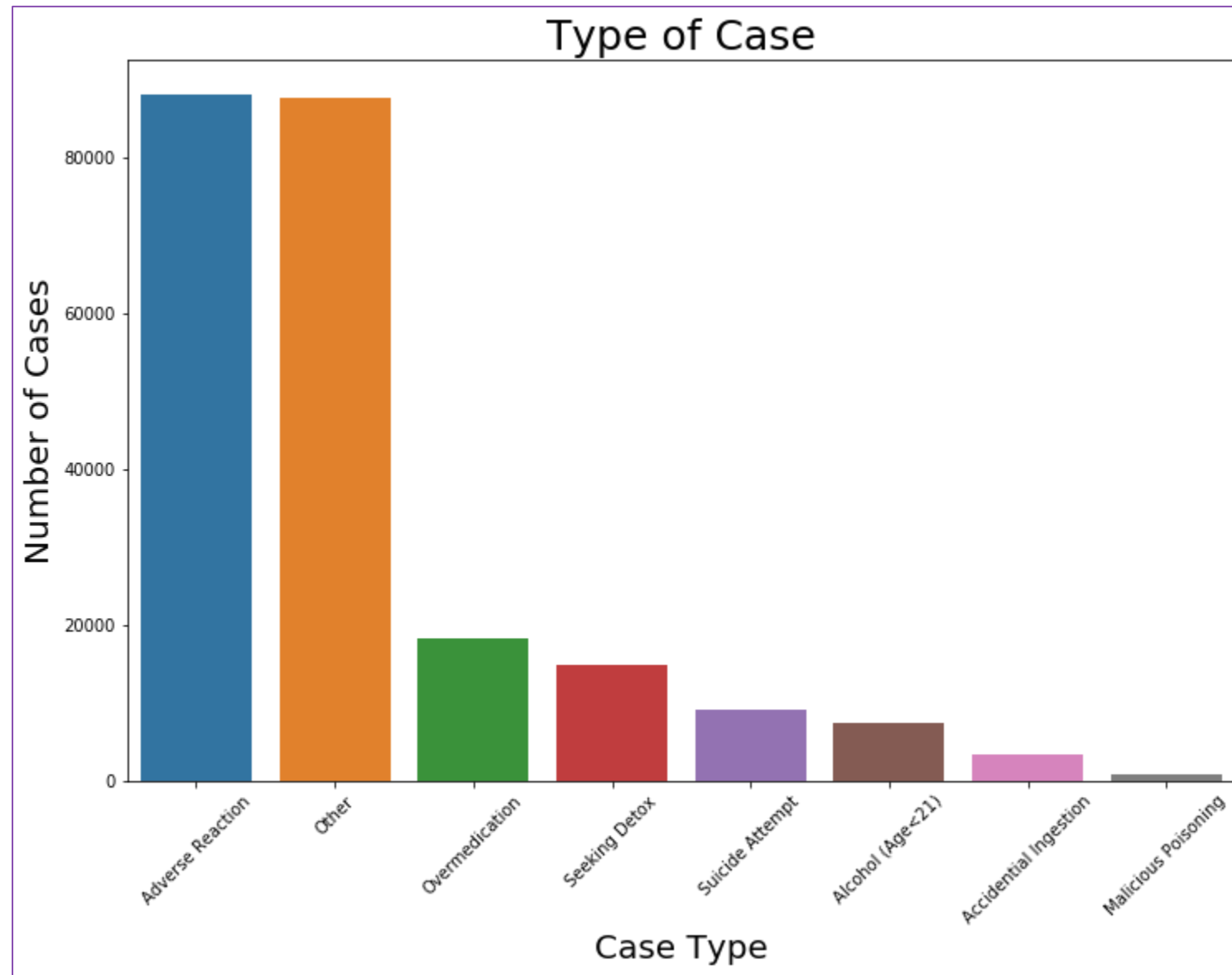- The time of day that most cases occur is between 12:00 pm - 11:59 pm.
- Alcohol, cocaine, heroin and marijuana are involved in many of the cases.
- Most of the drugs were taken orally.
- Most were not confirmed by a toxicology test.
- Most of the cases were due to adverse reactions to the drugs.
- Most cases were discharged home.
- Most of the drugs involved were pharmaceutical.

Inferential Statistics
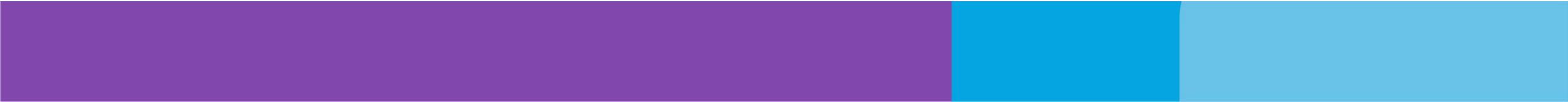
Is there a significant difference in the frequency of cases in each of the cities?

- H_o: p1 = p2... = p15
- H_a: p1 != p2... != p15

- The p-value was 0.0; therefore, H_o was rejected. It is reasonable to conclude that the distribution of the proportion of the location of the cases is not equal.

# Does the race of the patient vary significantly?

- H_o: p1 = p2... = p15
- H_a: p1 != p2... != p15

- The p-value was 0.0; therefore, H_o was rejected. It is reasonable to conclude that the distribution of the proportion of race is not equal.

Are the disposition of the case and the case type independent of each other?

- H_o: the population frequencies are equal to the expected frequencies

- H_a: the population frequencies are not equal to the expected frequencies.

- The p-value was 0.0; therefore, H_o was rejected. It is reasonable to conclude that the distribution of population frequencies are not equal to the expected frequencies.

Modeling
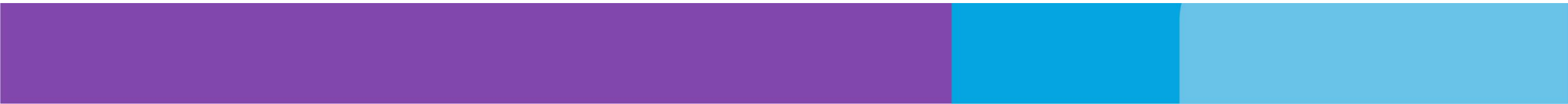
Logistic Regression Full Data Set

Class	Precision	Recall	F1-score	Support
Accidental Ingestion	0.88	0.01	0.02	1,301
Adverse Reaction	0.92	0.99	0.95	35,239
Alcohol (Age<21)	0.99	1.00	1.00	2,968
Malicious Poisoning	0.50	0.01	0.02	317
Other	0.87	0.93	0.90	35,051
Overmedication	0.74	0.72	0.73	7,259
Seeking Detox	0.77	0.51	0.61	5,937
Suicide Attempt	0.74	0.46	0.57	3,613

Logistic Regression Reduced Data Set

Class	Precision	Recall	F1-score	Support
Accidental Ingestion	0.22	0.68	0.33	814
Adverse Reaction	0.97	0.84	0.90	22,024
Alcohol (Age<21)	0.99	1.00	1.00	1,855
Malicious Poisoning	0.09	0.75	0.16	198
Other	0.95	0.65	0.77	21,907
Overmedication	0.47	0.75	0.58	3,710
Seeking Detox	0.47	0.75	0.58	3,710
Suicide Attempt	0.47	0.72	0.57	2,258

Random Forest Reduced Data Set

Class	Precision	Recall	F1-score	Support
Accidental Ingestion	0.82	0.04	0.08	814
Adverse Reaction	0.90	0.99	0.94	22,024
Alcohol (Age<21)	0.99	1.00	1.00	1,855
Malicious Poisoning	0.84	0.27	0.41	198
Other	0.84	0.93	0.89	21,907
Overmedication	0.71	0.66	0.68	4,537
Seeking Detox	0.80	0.44	0.57	3,710
Suicide Attempt	0.73	0.24	0.36	2,258

Random Forest Reduced Data Set—Hyperparameter Tuning

Class	Precision	Recall	F1-score	Support
Accidental Ingestion	0.22	0.30	0.25	814
Adverse Reaction	0.92	0.92	0.92	22,024
Alcohol (Age<21)	0.99	1.00	1.00	1,855
Malicious Poisoning	0.18	0.40	0.25	198
Other	0.91	0.75	0.82	21,907
Overmedication	0.53	0.94	0.67	4,537
Seeking Detox	0.57	0.60	0.59	3,710
Suicide Attempt	0.48	0.39	0.43	2,258

Precision Scores for Each Model

Precision Scores

Questions?