Police Incident Data

CST8390 Final Report

Submitted by

Sebastian, Jules	040 983 119
Simpson, Skyler	040 951 590

Algonquin College

December 10th 2021

Table of Contents

1.0 Introduction

The data being used was sourced from preliminary reports given to the Dallas Police Department after June 1st 2014. Example sources are 911 calls, written reports, and word of mouth transcriptions from witnesses.

The data is frequently updated, and a more current version of the dataset we investigated can be found at the City of Dallas Open Data website.[1]

Due to known unreliability of the data in its current form, the Dallas Police Department has hosted this warning on its public website:

"*This information reflects crimes as reported to the Dallas Police Department as of the current date. Crime classifications are based upon preliminary information supplied to the Dallas Police Department by the reporting parties and the preliminary classifications may be changed at a later date based upon additional investigation. Therefore, the Dallas Police Department does not guarantee (either expressed or implied) the accuracy, completeness, timeliness, or correct sequencing of the information contained herein and the information should not be used for comparison purposes over time."*

Naturally, as police incident reports contain identifying information, this dataset is curated and incomplete. Some instances of omitted data pertain to sexual offenses, offenses pertaining to minors or offenses related to social services. In addition, property and vehicle information is circumstantially omitted.

In addition, typos are abundant throughout the dataset, including in the attribute names. This leads to many nominal fields with multiple variations of spelling the same values. These will need to be grouped.

1.1 Data Understanding

The original data is presented in a collection of Eighty-Six (86) Attributes. The following Table includes the descriptions of the data grouped in each of these categories.

[See Table 1.1.1 Attributes of Unmodified Data]

To summarize the contents, the original data contains various attributes which describe the ***circumstances surrounding the event***, such as the date and time the police was contacted, the type of incident reported, the means by which the incident was reported, the sector of Dallas in which the incident occured, the date and time the incident began, the date and time the incident ended, when the incident was inserted into the system, who was contacted and dispatched to resolve the incident and the officer and investigator information for each officer involved.

The original data also contains attributes that describe ***the incident itself***, notably the offense status, the modus operandi of the offender, whether the crime was hate crime, family, gang of drug related, details pertaining to the weapon(s) used, the offense type, the Criminal Justice Information service code and the penal code.

Lastly, the original data contains details that satisfy the requirements to be classified as a National Incident Based Reporting System (NIBRS). These fields classify the data existing in the aforementioned attributes according to the NIBRS Standards. These fields contain repeat information, most notably a recategorization of the same location and penal information.

The original data contains over Eight Hundred and Seventy-Four Thousand (874 000) rows of data, and includes hundreds of thousands of duplicate data points and incomplete fields.

1.2 Data Preprocessing

The data must be prepared to help our analysis. The raw data is provided as a Comma Separated Values (CSV) File.

1.2.1 Making the Data Weka-Readable

Due to formatting issues, missing data, and string fields that include special characters and commas, the raw data must be modified before it can be read by Weka.

Modifying the existing string Attributes in Microsoft Excel is not possible due to crashes that occur when attempting to modify the contents of the file.

For this reason, a Python Script was written to solve the CSV formatting issues, such that the data can be read in Weka.

The content of the data was not transformed during this process, only reformatted for compatibility.

1.3 Data Cleaning

Upon selecting which Attributes are meaningful, the data must be cleaned before it can be used to meaningfully perform analysis.

1.3.1 Duplicate Data

The dataset includes many duplicate fields that describe the same event. One such sample can be seen in the image below, where the same instance of a crime includes multiple reports;

[See Image 1.3.1.1 Duplication of Corpse Abuse Incident]

When duplicate rows of data of this nature are found in the data, they are removed.

1.3.2 Handle Missing Data

Initially instances having more than 6 missing values were removed from the analysis as they would skew the data by being filled up with Statistical values.

As the data was highly correlated, relation between missing values and present values could be figured, and thus some of the missing values could be filled up without affecting the data.

As the data was highly correlated, we did not proceed using statistical values to fill up the missing values, rather, opted to delete instances with the missing values. We came to this decision as we were working with a big dataset and had over 870,000 instances.

1.3.3 Outliers

Outliers in each attribute were analysed. Any information that did not make sense, and could not have been filled up based on the other values present, such instances were deleted. We tried to keep as much information as possible.
Beat with the value of 7, Sector with the value of 0 and Zipcode with the value of 0 were removed from the dataset as it would have negatively affected the analysis.

Eventually the dataset had two columns for years. One was dropped as it stated the year in which the incident report was created. The other column which had the year when the incident occurred, was providing the information on when the actual event happened and would have been more insightful for us as part of the analysis. Thus we decided to move on with this column over the other.

Year the incident was filled had outliers where value for the years dated back to 1914 and to the future where the year was 9999. This was clearly an outlier, and data were trimmed based on this value before dropping this column.

An interesting observation was how even after dropping the year column, we had some outliers in the second year column. On further analyzing the related papers, we came to the conclusion that these were not outliers as there were events reported on a later date, but it actually took place back in 1974. As these events did hold value to the analysis, we did keep these values for our analysis.

1.3.4 Grouping

Under the Drug Related Istevencident (Incident) unknown elements were noted as “UNK”, these elements were renamed as “Unknown”. Additionally, there were elements having numbers of 2 and 3. These were renamed to “Unknown” as well.

[See Image 1.3.4.1 Drug Related Istevencident (Incident) Before Attribute Grouping]

2.0 Data Analysis

2.1 K Nearest Neighbors

After translating nominal attributes to binary fields for vector comparison, the attribute number dramatically increased, signalling that further grouping is required. The load time on one instance of knn took well over 10 hours to compute.

Consider the following results for k equals Two (2):

[Image 2.1.1 kNN Results for k Equals Two]

Proceeding with this method does not seem favourable for our analysis with the model in the current state, and so, further refinement of the data is required.

2.2 Simple k Means

Clustering analysis made sense for our dataset, as it is less complex and would not take as long. The clustering was done using the train dataset. As we were interested in a simple “Yes” and “No” analysis, we stuck to the value of k as Two (2).

All the results of our analysis could be found in the attached folder named “ModelsandResultBuffer ”.

Improvements to the model were brought by ignoring some attributes that could have a negative impact on the model. The model initially had a percentage of incorrectly clustered instances of 40% and this was brought down to 38.79% by ignoring **Time of Occurrence** and **Type of Incidents.** Reason for ignoring these attributes was the number of unique values in these nominal attributes.

2.3 Farthest First

As the Farthest First is an improved version of the Simple kMeans algorithm, naturally this was our next choice as an algorithm. The value of k was set to Two (2). The percentage of incorrectly clustered instances from this algorithm without ignoring any attributes, was 11.37%. On ignoring **Time of Occurrence** and **Type of Incidents** the percentage of incorrectly clustered instances was reduced to 10.85%. Using the test dataset, we predicted 8% of the instances would involve drugs.

3.0 Discussion of Results

3.1 Possible Sources of Error

Though we cut down our attributes to 18 from 86 by removing the least suitable attributes, the model was still too complex. Upon being reprimanded for it during the presentation phase, we agreed to simplify the model by excluding additional attributes and grouping nominal attributes into smaller subcategories.

This topic is further discussed in section 4.1

3.2 Summary of Expected Results

The Training Dataset we selected contained a true positive rate of $3 < N < 4$ (%) when checking for whether or not it was a drug related incident. We expected the Test Dataset to align with these results. While it is possible that the test dataset would have vastly different tendencies than the Training set, we have no reason to believe that would be the case.

3.3 Summary of Findings

As part of the analysis, we decided to have one execution of kNN. On running this overnight, we got a kNN model predicting a true positive rate of 98.15% with an percentage of incorrectly clustered instances at 1.85%. Making predictions at such high accuracy generally means that the model was overfitted and would fail on another dataset. Thus we decided to bag and categorize the values in the attributes, to reduce variation.
Our doubts were confirmed when we were ignoring the attributes with a high number of values, in the Simple kMeans and Farthest First.

4.0 Further Analysis

After presenting our findings in the presentation, suspicions that our refined dataset was not apt for analysis was confirmed. We Modified the existing attributes to simplify the set. Some nominal fields that were incompressible were removed.

Other Attributes were regrouped. Consider Time, which was reclassified as Twelve (12) nominal groups, each representing Two (2) hour time windows. The grouping changes were made to correspond with the expectation of the evaluator, that all nominal fields should be represented by no more than Fifteen (15) different fields.

This inherently means we have less data in this model, and the data we have is less precise. Still, the new data model should not have the same issues with fitting that the previous model had.

This section discusses the changes made, the new findings, and briefly compares the findings with the results of the previous model.

4.1 Repreparing the Data

The way forward was to limit the attributes being used for analysis.
As a result, we decided that the fields containing nominal fields that could not be grouped into 15 or less meaningful groups without invalidating the data should be eliminated.

For this reason, the Attributes named **Type of Incident, Type of Location, Reporting Area, Beat, Division, Sector and Penal Code** were all removed. Any grouping of these data fields would bastardize the specific nature of these fields. Some of these attributes, most notably **Type of Location, Reporting Area, Beat, and Division** can be represented by the Council District, which is a nominal attribute representing which district the crime occurred in. This is a general way to categorize the specific location data into <15 nominal fields.

These omissions amount to a large and tangible loss of specific information.

In addition, of the remaining Attributes, Two (2) contain data issues:

Year of Occurrence is an attribute that mentions the year that an incident occurred. Keep in mind that this dataset contains data for incidents *REPORTED* after June 1st 2014. This means that reports of past incidents are included in this dataset.

Zip Code is kept as a way to subdivide locations, which are otherwise only denoted by the **Council District** Attribute. While this provides much needed detail to our location data, the values stored in the zip code attribute are nominal, and cannot be grouped in meaningful ways, just like almost every other nominal attribute in this dataset. This is because alike zip codes don't necessarily align to alike geographical regions.

These attributes are both included in the final dataset to meet the requirement of Ten (10) attributes [each with at least 100 instances of data], and are known sources of error.

4.2 Analysing the Adjusted Dataset

To analyze the new modified dataset, we abandoned the use of kNN. The processing took too long, and with Zip Code being as it was, preferred to simply evaluate Simple k Means and a Farthest First Clustering.

4.2.1 Simple K Means

Our simple K-means analysis of the more refined dataset led to more accurate results, with approximately half the number of incorrectly clustered data points, down to 20.7% from 40.0%.

For details pertaining to the results, see the files included in the Appendix of this report.

4.2.2 Farthest First

The results of the farthest first algorithm on the new dataset were extremely promising. With 6.13% of data incorrectly clustered, the model predicts that roughly 4% of instances are drug related. This aligns with our expectations.

Knowing that the **Year of Occurrence** attribute may be negatively impacting our results, we reattempted the farthest first algorithm, excluding that attribute. Without the attribute, the incorrectly clustered percentage is reduced to 5.6% and the model predicts that roughly 3% of incidents are drug related. This is a confirmation that this attribute is skewing our results.

For details pertaining to the results, see the files included in the Appendix of this report.

5.0 Conclusion

Our goal was to predict whether or not a given incident from the Dallas Police Department dataset was drug related. We expected our results to align with the existing dataset's tendencies, notably that $3 < N < 4$ (%) of incidents are drug related.

Our initial cleaning of the data was insufficient and led to overfitting the data. In addition, a lack of grouping caused faulty analysis to occur. Upon limiting the Attributes further, and grouping nominal fields to reduce their weights, a new analysis was performed.

Our initial data did not align with our expectations, and was either overfit, or too inaccurate to be of any use. Upon reducing the specificity of our dataset, we can now be certain that the results are not overfit.

The new results align with our expectations, but we have no way of conclusively asserting that this is due to our analysis. The compression of the dataset is too extreme, and any assertions made on the dataset we used is borderline meaningless.

In addition, sources of error (notably the **Year of Occurrence** and **Zip Code** Attributes) are known to exist. They were included in their current for expediency or to avoid violating assignment requirements.

When these attributes are omitted, result accuracy increases, or at least aligns with our expectations.

In a more comprehensive analysis, the expense of a few hundred hours of analysis on the base dataset may allow us to employ an optimal selection for attributes and fields to reach our conclusion. As is, the time investment required to come to certainty that our data is as controlled as possible is not within the scope of reasonability.

Selection of a smaller and cleaner dataset for this assignment would have been preferable.

As an homage to the Dallas Police Department;

We do not guarantee (either expressed or implied) the accuracy, completeness, timeliness, or correct sequencing of the information contained herein.

We do not recommend the inclusion of our analysis in further research.

References

1. Police Incidents. Dallas Police Department (Online)
   https://www.dallasopendata.com/Public-Safety/Police-Incidents/qv6i-rri7
   (Last Visited 2021/12/10)

Appendix

Files

FarthestFirst-AllAtrtibutes-Self.txt

```
=== Run information ===

Scheme:       weka.clusterers.FarthestFirst -N 2 -S 1
Relation:     reWorked_PoliceIncident_Train-weka.filters.unsupervised.attribute.NumericToNominal-R2,10
Instances:    692519
Attributes:   10
              Council District
              Year of Occurrence
              Month of Occurence
              Day of the Week
              Time of Occurrence
              Person Involvement Type
              Victim Type
              Hate Crime Description
              Zip Code
Ignored:
              Drug Related Incident
Test mode:    Classes to clusters evaluation on training data

=== Clustering model (full training set) ===


FarthestFirst
==============

Cluster centroids:

Cluster 0
		D3 2014 September Thu 13:00 Victim Individual None 75232
Cluster 1
		D6 2017 July Mon 01:00 Other Business Unknown 75247



Time taken to build model (full training data) : 0.43 seconds

=== Model and evaluation on training set ===

Clustered Instances

0      674253 ( 97%)
1       18266 (  3%)


Class attribute: Drug Related Incident
Classes to Clusters:

      0      1  <-- assigned to cluster
 649691  17933 | No
  24562    333 | Yes

Cluster 0 <-- No
Cluster 1 <-- Yes

Incorrectly clustered instances :    42495.0       6.1363 %
```

FarthestFirst-AllAttributes-TestSet.txt

=== Run information ===

Scheme: weka.clusterers.FarthestFirst -N 2 -S 1
Relation: reWorked_PoliceIncident_Train-weka.filters.unsupervised.attribute.NumericToNominal-R2,10
Instances: 692519
Attributes: 10
            Council District
            Year of Occurrence
            Month of Occurence
            Day of the Week
            Time of Occurrence
            Person Involvement Type
            Victim Type
            Hate Crime Description
            Drug Related Incident
            Zip Code
Test mode: user supplied test set: 114490 instances

=== Clustering model (full training set) ===

FarthestFirst
==============

Cluster centroids:

Cluster 0
            D3 2014 September Thu 13:00 Victim Individual None No 75232
Cluster 1
            D6 2017 July Mon 01:00 Other Business Unknown No 75247

Time taken to build model (full training data) : 0.39 seconds

=== Evaluation on test set ===
Clustered Instances

0   109851 (96%)
1   4639 (4%)

FarthestFirst-YearPersonHate-Self.txt

```
=== Run information ===

Scheme:       weka.clusterers.FarthestFirst -N 2 -S 1
Relation:     reWorked_PoliceIncident_Train-weka.filters.unsupervised.attribute.NumericToNominal-R2,10
Instances:    692519
Attributes:   10
              Council District
              Month of Occurence
              Day of the Week
              Time of Occurrence
              Victim Type
              Zip Code
Ignored:
              Year of Occurrence
              Person Involvement Type
              Hate Crime Description
              Drug Related Incident
Test mode:    Classes to clusters evaluation on training data

=== Clustering model (full training set) ===


FarthestFirst
==============

Cluster centroids:

Cluster 0
                D3 September Thu 13:00 Individual 75232
Cluster 1
                D1 February Sun 01:00 Law Enforcement Offi 75217



Time taken to build model (full training data) : 1.23 seconds

=== Model and evaluation on training set ===

Clustered Instances

0      601391 ( 87%)
1       91128 ( 13%)


Class attribute: Drug Related Incident
Classes to Clusters:

      0      1  <-- assigned to cluster
 583062  84562 | No
  18329   6566 | Yes

Cluster 0 <-- No
Cluster 1 <-- Yes

Incorrectly clustered instances :    102891.0    14.8575 %
```

FarthestFirst-YearPersonHate-Test.txt

```
=== Run information ===

Scheme:       weka.clusterers.FarthestFirst -N 2 -S 1
Relation:     reWorked_PoliceIncident_Train-weka.filters.unsupervised.attribute.NumericToNominal-R2,10
Instances:    692519
Attributes:   10
              Council District
              Month of Occurence
              Day of the Week
              Time of Occurrence
              Victim Type
              Drug Related Incident
              Zip Code
Ignored:
              Year of Occurrence
              Person Involvement Type
              Hate Crime Description
Test mode:    user supplied test set: 114490 instances


=== Clustering model (full training set) ===


FarthestFirst
==============

Cluster centroids:

Cluster 0
         D3 September Thu 13:00 Individual No 75232
Cluster 1
         D7 November Sun 01:00 Society/Public Yes 75217



Time taken to build model (full training data) : 0.38 seconds

=== Evaluation on test set ===
Clustered Instances

0      110866 ( 97%)
1        3624 (  3%)


=== Re-evaluation on test set ===

User supplied test set
Relation:     reWorked_PoliceIncident_Test-weka.filters.unsupervised.attribute.NumericToNominal-R2,10-weka.filters.unsupervised.attribute.StringToNominal-R9
Instances:    114490
Attributes:   10


FarthestFirst
==============

Cluster centroids:

Cluster 0
         D3 September Thu 13:00 Individual No 75232
Cluster 1
         D7 November Sun 01:00 Society/Public Yes 75217

Clustered Instances

0      110866 ( 97%)
1        3624 (  3%)
```

FarthestFirst-YearPerson-Self.txt

```
=== Run information ===

Scheme:       weka.clusterers.FarthestFirst -N 2 -S 1
Relation:     reWorked_PoliceIncident_Train-weka.filters.unsupervised.attribute.NumericToNominal-R2,10
Instances:    692519
Attributes:   10
              Council District
              Month of Occurence
              Day of the Week
              Time of Occurrence
              Victim Type
              Hate Crime Description
              Zip Code
Ignored:
              Year of Occurrence
              Person Involvement Type
              Drug Related Incident
Test mode:    Classes to clusters evaluation on training data

=== Clustering model (full training set) ===


FarthestFirst
==============

Cluster centroids:

Cluster 0
		D3 September Thu 13:00 Individual None 75232
Cluster 1
		D1 July Tue 01:00 Government Unknown 75208



Time taken to build model (full training data) : 0.87 seconds

=== Model and evaluation on training set ===

Clustered Instances

0      651544 ( 94%)
1       40975 (  6%)


Class attribute: Drug Related Incident
Classes to Clusters:

      0      1  <-- assigned to cluster
 629882  37742 | No
  21662   3233 | Yes

Cluster 0 <-- No
Cluster 1 <-- Yes

Incorrectly clustered instances :	59404.0	  8.578  %
```

FarthestFirst-YearPerson-Test.txt

```
=== Run information ===

Scheme:       weka.clusterers.FarthestFirst -N 2 -S 1
Relation:     reWorked_PoliceIncident_Train-weka.filters.unsupervised.attribute.NumericToNominal-R2,10
Instances:    692519
Attributes:   10
              Council District
              Month of Occurence
              Day of the Week
              Time of Occurrence
              Victim Type
              Hate Crime Description
              Drug Related Incident
              Zip Code
Ignored:
              Year of Occurrence
              Person Involvement Type
Test mode:    user supplied test set: 114490 instances


=== Clustering model (full training set) ===


FarthestFirst
==============

Cluster centroids:

Cluster 0
		D3 September Thu 13:00 Individual None No 75232
Cluster 1
		D7 October Sun 01:00 Business Unknown Yes 75227



Time taken to build model (full training data) : 0.85 seconds

=== Evaluation on test set ===
Clustered Instances

0      111131 ( 97%)
1        3359 (  3%)
```

FarthestFirst-Year-Self.txt

```
=== Run information ===

Scheme:       weka.clusterers.FarthestFirst -N 2 -S 1
Relation:     reWorked_PoliceIncident_Train-weka.filters.unsupervised.attribute.NumericToNominal-R2,10
Instances:    692519
Attributes:   10
              Council District
              Month of Occurence
              Day of the Week
              Time of Occurrence
              Person Involvement Type
              Victim Type
              Hate Crime Description
              Zip Code
Ignored:
              Year of Occurrence
              Drug Related Incident
Test mode:    Classes to clusters evaluation on training data

=== Clustering model (full training set) ===


FarthestFirst
==============

Cluster centroids:

Cluster 0
               D3 September Thu 13:00 Victim Individual None 75232
Cluster 1
               D6 July Mon 01:00 Other Business Unknown 75247



Time taken to build model (full training data) : 0.69 seconds

=== Model and evaluation on training set ===

Clustered Instances

0      678150 ( 98%)
1       14369 (  2%)


Class attribute: Drug Related Incident
Classes to Clusters:

      0      1  <-- assigned to cluster
 653499  14125 | No
  24651    244 | Yes

Cluster 0 <-- No
Cluster 1 <-- Yes

Incorrectly clustered instances :    38776.0        5.5993 %
```

Images

Image 1.3.1.1 Duplication of Corpse Abuse Incident

Year of Inci	Wa	Call (911) Problem	Type of Incident	Type Locati	Reporting A	Beat	Division	Sector	Council Dis	Year1 of Oc	Month1 of (	Day1 of the	Time1 of Oc	Person Invo	Victim Type	Hate Crime	Drug Rela
2019	1	11B - BURG OF BUS	ABUSE OF CORPSE	Other	1082	652	NORTH CENTRA	650	D13	2019	January	Tue	21:00	Victim	Individual	None	No
2019	1	11B - BURG OF BUS	ABUSE OF CORPSE	Other	1082	652	NORTH CENTRA	650	D13	2019	January	Tue	21:00	Victim	Individual	None	No
2019	1	11B - BURG OF BUS	ABUSE OF CORPSE	Other	1082	652	NORTH CENTRA	650	D13	2019	January	Tue	21:00	Victim	Individual	None	No
2019	1	11B - BURG OF BUS	ABUSE OF CORPSE	Other	1082	652	NORTH CENTRA	650	D13	2019	January	Tue	21:00	Victim	Individual	None	No
2019	1	11B - BURG OF BUS	ABUSE OF CORPSE	Other	1082	652	NORTH CENTRA	650	D13	2019	January	Tue	21:00	Victim	Individual	None	No
2019	1	11B - BURG OF BUS	ABUSE OF CORPSE	Other	1082	652	NORTH CENTRA	650	D13	2019	January	Tue	21:00	Victim	Individual	None	No
2019	1	11B - BURG OF BUS	ABUSE OF CORPSE	Other	1082	652	NORTH CENTRA	650	D13	2019	January	Tue	21:00	Victim	Individual	None	No
2019	1	11B - BURG OF BUS	ABUSE OF CORPSE	Other	1082	652	NORTH CENTRA	650	D13	2019	January	Tue	21:00	Victim	Individual	None	No
2019	1	11B - BURG OF BUS	ABUSE OF CORPSE	Other	1082	652	NORTH CENTRA	650	D13	2019	January	Tue	21:00	Victim	Individual	None	No
2019	1	11B - BURG OF BUS	ABUSE OF CORPSE	Other	1082	652	NORTH CENTRA	650	D13	2019	January	Tue	21:00	Victim	Individual	None	No
2019	1	11B - BURG OF BUS	ABUSE OF CORPSE	Other	1082	652	NORTH CENTRA	650	D13	2019	January	Tue	21:00	Victim	Individual	None	No
2019	1	11B - BURG OF BUS	ABUSE OF CORPSE	Other	1082	652	NORTH CENTRA	650	D13	2019	January	Tue	21:00	Victim	Individual	None	No
2017	2	58 - ROUTINE INVESTIGATION	ABUSE OF CORPSE	Outdoor Area F	1100	233	NORTHEAST	230	D9	2017	November	Sun	12:00	Victim	Government	None	UNK
2019	1	11B - BURG OF BUS	ABUSE OF CORPSE	Other	1082	652	NORTH CENTRA	650	D13	2019	January	Tue	21:00	Victim	Individual	None	No

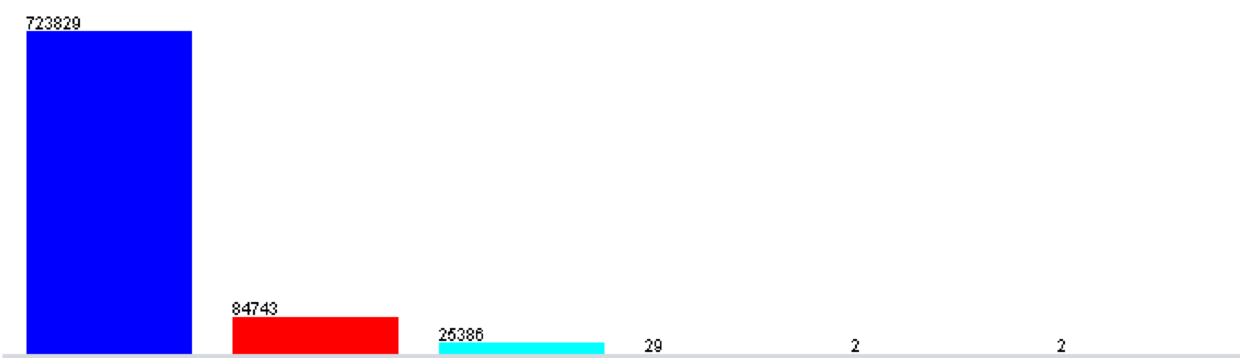
Image 1.3.4.1 Drug Related Istevencident (Incident) Before Attribute Grouping

Selected attribute

Name: Drug Related Istevencident Type: Nominal
Missing: 38358 (4%) Distinct: 6 Unique: 0 (0%)

No.	Label	Count	Weight
1	No	723829	723829.0
2	UNK	84743	84743.0
3	Yes	25386	25386.0
4	Unknown	29	29.0
5	2	2	2.0
6	3	2	2.0

Class: Drug Related Istevencident (Nom) Visualize All



Tables

Table 1.1.1 List of Unprocessed Attributes

Column Name	**Description**	**Type**
Incident Number w/year	An RMS generated incident number (report number) w	Plain Text
Year of Incident	Year associated with the incident number	Number
Service Number ID	Incident number plus year code plus offense number (Ex: -02 means there is two offense with this one incident) Internal use	Plain Text
Watch	Police watch 1st 2nd or 3rd (1st watch = Late Night, 2nd watch = Days and 3rd watch = Evenings)	Plain Text
Call (911) Problem	Police call signal generated by Communications (Type of 911 call dispacthed)	Plain Text
Type of Incident	Type of Incident	Plain Text
Type Location	Location type where incident took place For example, Apartment Parking, Residence	Plain Text
Type of Property	The target item… Parkinglot, Motor Vehicle	Plain Text

Incident Address	Address where incident occurred	Plain Text
Apartment Number	Apartment number	Plain Text
Reporting Area	Geographic area comprised of reporting areas where incident occurred	Number
Beat	Geographic area comprised of beats where incident occurred	Number
Division	Geographic area comprised of census blocks where incident occurred (smallest police geography)	Plain Text
Sector	Geographic area comprised of Sectors where incident occurred	Number
Council District	Geographic area comprised of city council districts where incident occurred	Plain Text
Target Area Action Grids	Geographic areas targeted for higher than average crime	Plain Text
Community	Community Prosecution Areas as designated by the City Community Prosecutors	Plain Text

Date1 of Occurrence	The first date of the date occurrence of the incident (Ex: incident occurred between 1/1/2016 and 1/2/2016)	Plain Text
Year1 of Occurrence	Year of the indent based on the Date of Occurrence (Date1). Internal use	Number
Month1 of Occurence	Month (starting) of the indent based on the Date of Occurrence (Date1). Internal use	Plain Text
Day1 of the Week	Day of the indent based on the Date of Occurrence (Date1). Internal use	Plain Text
Time1 of Occurrence	The first (starting) time of the time occurrence of the incident (Ex: incident occurred between 8:00am and 5:00pm)	Plain Text
Day1 of the Year	The calender number of the year 1-365 based on Date1. Internal use	Number
Date2 of Occurrence	The second date of the date occurrence of the incident (Ex: incident occurred between 1/1/2016 and 1/2/2016)	Plain Text
Year2 of Occurrence	Year of the indent based on the Date of Occurrence (Date2)	Number

Month2 of Occurence	Month (end) of the indent based on the Date of Occurrence (Date2)	Plain Text
Day2 of the Week	Day of the indent based on the Date of Occurrence (Date1)	Plain Text
Time2 of Occurrence	The second(end) time of the time occurrence of the incident (Ex: incident occurred between 8:00am and 5:00pm)	Plain Text
Day2 of the Year	The calender number of the year 1-365 based on Date2. Internal use	Number
Date of Report	The date of the incident as reported to the police	Plain Text
Date incident created	The date the incident record was created. Internal use	Plain Text
Offense Entered Year	The year the offense was entered into the system . Internal use	Number
Offense Entered Month	The month the offense was entered into the system. Internal use	Plain Text
Offense Entered Day of the Week	The day the offense was entered into the system. Internal use	Plain Text

Offense Entered Time	The time the offense was entered into the system. Internal use	Plain Text
Offense Entered Date/Time	The calender number of the year the offense was entered. Internal use	Plain Text
CFS Number	CFS Number	Plain Text
Call Received Date Time	The date the related call was received	Plain Text
Call Date Time	Date and time of the related call	Plain Text
Call Cleared Date Time	Date and time related call was cleared	Plain Text
Call Dispatch Date Time	Date and time related call was dispatched	Plain Text
Special Report	No longer applies.. PreRMS	Plain Text
Person Involvement Type	Person can be; victim, reporting person, witness	Plain Text
Victim Type	Victim Type	Plain Text
Victim Race	Victim Race	Plain Text
Victim Ethnicity	Victim Ethnicity	Plain Text

Victim Gender	Victim Gender	Plain Text
Responding Officer #1 Badge No	Responding Officer #1 Badge No	Plain Text
Responding Officer #1 Name	Responding Officer #1 Name	Plain Text
Responding Officer #2 Badge No	Responding officer #2 Badge number	Plain Text
Responding Officer #2 Name	Responding officer #2 Name	Plain Text