The data represents a piece of police report containing a thousand records all dated to the year of 2017. These records are incidents that are reported to or recorded by police, within the city. Each record is defined by 19 vectors, these vectors define 3 aspects of the incident:

1) When?

2) Where?

3) & What?

The "What" is defined by offense\_code\_extension, summarized\_offense\_description, offense\_type, summary, offense\_code, general\_offense\_number, census\_tract\_2000, offense\_code, rms\_cdw\_id.

The "When" is defined by year, date\_reported, occured\_date\_or\_date\_range\_start, month, occured\_date\_range\_end.

The "Where" aspects are defined by hundred\_block\_location, zone\_beat, latitude, district\_sector, longitude. The data also has a location data frame with 3 vectors, latitude, longitude and needs\_recording associated with each record in the data. The latitude vector is referenced from the location table into the police\_incidents main data frame as location.latitude.

[1] "year" - Describes the year of incidents, in this data frame all belong to the year 2017.

[2] "zone\_beat" - Describes the zone/ Beat of the incident into 53 categories.

[3] "latitude" - Describes the location of the incident in terms of latitude.

[4] "offense\_code\_extension" - Describes the extension of offense code of the incident into 17 categories.

[5] "summarized\_offense\_description" - Summarizes the offense description of the incident into 33 categories.

[6] "date\_reported" - Describes the reporting date of the incident in the form of a TIMESTAMP.

[7] "offense\_type" - Describes the offense type of the incident into roughly 70 categories.

[8] "occurred\_date\_or\_date\_range\_start" -Describes the occurred date or start date of the incident in the form of a TIMESTAMP.

[9] "summary\_offense\_code" - Describes the offense code of the incident into 21 categories.

[10] "month" - Describes the month of incidents, in this data frame, all belong to the 10th month (October).

[11] "general\_offense\_number" - Describes the offense general number of the incident. As the name suggests these are not unique as they are general. Existing records are attributed to 605 different types of general offense numbers.

[12] "census\_tract\_2000" - Describes the census tract for the area where the incident occurred.

The "Census Tract" is an area roughly equivalent to a neighborhood established by the Bureau of Census for analyzing populations. They generally encompass a population between 2,500 to 8,000 people.

[13] "location" - Data frame with 3 vectors, latitude, longitude and needs\_recording associated with each record in the data.

[14] "offense\_code" - Describes the code for the type of the incident into 51 categories.

[15] "hundred\_block\_location" - Describes the block location of the incident in the city.

[16] "rms\_cdw\_id" - Provides the UNIQUE identification for each incident that occurred and was recorded.

[17] "district\_sector" - Describes the district/sector of the incident into 19 categories.

[18] "longitude" - Describes the location of the incident in terms of longitude.

[19] "occurred\_date\_range\_end" - Describes the end date with time (TIMESTAMP) of the incident.