

# Understanding New York Motor Vehicle Collision Crash

Y-Data Profiling

No of Columns: 48

No of Observations: 817723

CRASH\_RECORD\_ID

Description: This number can be used to link to the same crash in the Vehicles and People datasets. This number also serves as a unique ID in this dataset.

Type: STRING

DISTINCT: 817723

Missing: No

Natural Key

Max Length: 128

CRASH\_DATE\_EST\_I

Description: Crash date estimated by desk officer or reporting party (only used in cases where crash is reported at police station days after the crash)

Type: Boolean

DISTINCT: 2

Missing: 756594

CRASH\_DATE

Description: Date and time of crash as entered by the reporting officer

Type: Date

Min: 2013-03-03 16:48:00

Max: 2024-03-26 01:40:00

DISTINCT: 536888

Missing: No

#### POSTED\_SPEED\_LIMIT

Description: Posted speed limit, as determined by reporting officer

TYPE: Real Number

Distinct: 46

Min: 0

Max: 99

Missing: No

#### TRAFFIC\_CONTROL\_DEVICE

Description: Traffic control device present at crash location, as determined by reporting officer

TYPE: STRING

DISTINCT: 19

Missing: No

Max length: 24

#### DEVICE\_CONDITION

Description: Condition of traffic control device, as determined by reporting officer

TYPE: STRING

DISTINCT: 8

Missing: 0

Max length: 24

#### WEATHER\_CONDITION

Description: Weather condition at time of crash, as determined by reporting officer

TYPE: STRING

DISTINCT: 12

Missing: 0

Max length: 24

#### LIGHTNING\_CONDITION

Description: Light condition at time of crash, as determined by reporting officer

TYPE: STRING

DISTINCT: 6

Missing: 0

Max length: 22

#### FIRST\_CRASH\_TYPE

Description: Type of first collision in crash

TYPE: STRING

DISTINCT: 18

Missing: 0

Max length: 28

#### TRAFFICWAY\_TYPE

Description: Trafficway type, as determined by reporting officer

TYPE: STRING

DISTINCT: 20

Missing: 0

Max length: 31

#### LANE\_CNT

Description: Total number of through lanes in either direction, excluding turn lanes, as determined by reporting officer (0 = intersection)

TYPE: Real Number

DISTINCT: 41

Missing: 618714

Min: 0

Max: 1191625

Skewed

#### ALIGNMENT

Description: Street alignment at crash location, as determined by reporting officer

TYPE: STRING

DISTINCT: 6

Missing: 0

Max length: 21

#### ROADWAY\_SURFACE\_COND

Description: Road surface condition, as determined by reporting officer

TYPE: STRING

DISTINCT: 7

Missing: 0

Max length: 15

#### ROADWAY\_DEFECT

Description: Road defects, as determined by reporting officer

TYPE: STRING

DISTINCT: 7

Missing: 24314

Max length: 17

#### REPORT\_TYPE

Description: Administrative report type (at scene, at desk, amended)

TYPE: STRING

DISTINCT: 3 (Not on Scene (Desk Report), On Scene, Amended)

Missing: 0

Max length: 26

#### CRASH\_TYPE

Description: A general severity classification for the crash. Can be either Injury and/or Tow Due to Crash or No Injury / Drive Away

TYPE: STRING

DISTINCT: 2 (NO INJURY / DRIVE AWAY, INJURY AND / OR TOW DUE TO CRASH)

Missing: 0

Max length: 32

#### INTERSECTION\_RELATED\_I

Description: A field observation by the police officer whether an intersection played a role in the crash. Does not represent whether or not the crash occurred within the intersection.

TYPE: Boolean

DISTINCT: 2 (Y/N)

Missing: 630174 ~ 77.1%

#### NOT\_RIGHT\_OF\_WAY\_I

Description: Whether the crash begun or first contact was made outside of the public right-of-way.

TYPE: Boolean

DISTINCT: 2

Missing: 780015 ~ 95.4%

#### HIT\_AND\_RUN\_I

Description: Crash did/did not involve a driver who caused the crash and fled the scene without exchanging information and/or rendering aid

TYPE: Boolean

DISTINCT: 2

Missing: 561774 ~ 68.7%

#### DAMAGE

Description: A field observation of estimated damage.

TYPE: String

DISTINCT: 3 (OVER \$1,500, \$501 - \$1,500, \$500 OR LESS)

Missing: No

Max length: 13

#### DATE\_POLICE\_NOTIFIED

Description: Calendar date on which police were notified of the crash

TYPE: Date

DISTINCT: 620545

Missing: No

Min: 2013-06-01 20:31:00

Max: 2024-03-26 01:42:00

#### PRIM\_CONTRIBUTORY\_CAUSED

Description: The factor which was most significant in causing the crash, as determined by officer judgment

TYPE: String

DISTINCT: 40

Missing: No

Max Length: 80

#### SEC\_CONTRIBUTORY\_CAUSE

Description: The factor which was second most significant in causing the crash, as determined by officer judgment

TYPE: String

DISTINCT: 40

Missing: No

Max Length: 80

#### STREET\_NO

Description: Street address number of crash location, as determined by reporting officer

TYPE: Real Number

DISTINCT: 11728

Missing: No

Min: 0

Max: 451100

#### STREET\_DIRECTION

Description: Street address direction (N,E,S,W) of crash location, as determined by reporting officer

TYPE: String

DISTINCT: 4 (N,E,W,S)

Missing: 4

Max Length: 31

#### STREET\_NAME

Description: Street address name of crash location, as determined by reporting officer

TYPE: String

DISTINCT: 1641

Missing: 1

Max Length: 31

#### BEAT\_OF\_OCCURRENCE

Description: Chicago Police Department Beat ID. Boundaries available at <https://data.cityofchicago.org/d/aerh-rz74>

TYPE: Real Number

DISTINCT: 276

Missing: 5

Max: 6100

Min: 111

#### PHOTOS\_TAKEN\_I

Description: Whether the Chicago Police Department took photos at the location of the crash

TYPE: BOOLEAN

DISTINCT: 2 (True, False)

Missing: Yes (806948 ~ 98.7%)

#### STATEMENTS\_TAKEN\_I

Description: Whether statements were taken from unit(s) involved in crash

TYPE: BOOLEAN

DISTINCT: 2 (True, False)

Missing: Yes (799465 ~ 97.8%)

#### DOORING\_I

Description: Whether crash involved a motor vehicle occupant opening a door into the travel path of a bicyclist, causing a crash

TYPE: BOOLEAN

DISTINCT: 2 (True, False)

Missing: Yes (815211 ~ 99.7%)

#### WORK\_ZONE\_I

Description: Whether the crash occurred in an active work zone

TYPE: BOOLEAN

DISTINCT: 2 (True, False)

Missing: Yes (813053 ~ 99.4%)

#### WORK\_ZONE\_TYPE

Description: The type of work zone, if any

TYPE: STRING

Max length: 12

DISTINCT: 4 (

- CONSTRUCTION
- UNKNOWN
- MAINTENANCE
- UTILITY

)

Missing: Yes (814105 ~ 99.6%)

#### WORKERS\_PRESENT\_I

Description: Whether construction workers were present in an active work zone at crash location

TYPE: BOOLEAN

DISTINCT: 2 (True, False)

Missing: Yes (816529 ~ 99.9%)

#### NUM\_UNITS

Description: Number of units involved in the crash. A unit can be a motor vehicle, a pedestrian, a bicyclist, or another non-passenger roadway user. Each unit represents a mode of traffic with an independent trajectory

TYPE: INT

Min: 1

Max: 18

DISTINCT: 17

Missing: No



#### MOST\_SEVERE\_INJURY

Description: Most severe injury sustained by any person involved in the crash

TYPE: STRING

Max length: 24

DISTINCT: 5 (

NO INDICATION OF INJURY

NONINCAPACITATING INJURY

REPORTED, NOT EVIDENT

INCAPACITATING INJURY

FATAL

Missing: Yes (1792) ~ 0.2%

#### INJURIES\_TOTAL

Description: Total persons sustaining fatal, incapacitating, non-incapacitating, and possible injuries as determined by the reporting officer

TYPE: INT

Min: 0

Max: 21

DISTINCT: 20

Missing: Yes (1780) ~ 0.2%

#### INJURIES\_FATAL

Description: Total persons sustaining fatal injuries in the crash

TYPE: FLOAT

Min: 0.0

Max: 4.0

DISTINCT: 5

Missing: Yes (1780) ~ 0.2%

#### INJURIES\_INCAPACITATING

Description: Total persons sustaining incapacitating/serious injuries in the crash as determined by the reporting officer. Any injury other than fatal injury, which prevents the injured person from walking, driving, or normally continuing the activities they were capable of performing before the injury occurred. Includes severe lacerations, broken limbs, skull or chest injuries, and abdominal injuries.

TYPE: INT

Min: 0

Max: 10

DISTINCT: 10

Missing: Yes (1780) ~ 0.2%

#### INJURIES\_NON\_INCAPACITATING

Description: Total persons sustaining non-incapacitating injuries in the crash as determined by the reporting officer. Any injury, other than fatal or incapacitating injury, which is evident to observers at the scene of the crash. Includes lump on head, abrasions, bruises, and minor lacerations.

TYPE: INT

Min: 0

Max: 21

DISTINCT: 19

Missing: Yes (1780) ~ 0.2%

#### INJURIES\_REPORTED\_NOT\_EVIDENT

Description: Total persons sustaining possible injuries in the crash as determined by the reporting officer. Includes momentary unconsciousness, claims of injuries not evident, limping, complaint of pain, nausea, and hysteria.

TYPE: INT

Min: 0

Max: 15

DISTINCT: 13

Missing: Yes (1780) ~ 0.2%

#### INJURIES\_NO\_INDICATION

Description: Total persons sustaining no injuries in the crash as determined by the reporting officer

TYPE: INT

Min: 0

Max: 61

DISTINCT: 48

Missing: Yes (1780) ~ 0.2%

#### INJURIES\_UNKNOWN:

Description: Total persons for whom injuries sustained, if any, are unknown

TYPE: FLOAT

Min: 0.0

Max: 0.0

DISTINCT: 1

Missing: Yes (1780) ~ 0.2%

#### CRASH\_HOUR

Description: The hour of the day component of CRASH\_DATE

TYPE: INT

Min: 0

Max: 23

DISTINCT: 24

Missing: No

#### CRASH\_DAY\_OF\_WEEK

Description: The day of the week component of CRASH\_DATE. Sunday=1

TYPE: INT

Min: 1

Max: 7

DISTINCT: 12

Missing: No

#### CRASH\_MONTH

Description: The month component of CRASH\_DATE

TYPE: INT

Min: 1

Max: 12  
DISTINCT: 12  
Missing: No

LATITUDE  
TYPE: DOUBLE  
Min: 0 (Some entries have lat, long as 0,0 (49))  
Actual Min: 41.64467013  
Max: 42.0226457  
DISTINCT: 300091  
Missing: Yes (5615) ~ 0.7%

LONGITUDE  
TYPE: DOUBLE  
Min: -87.93619295  
Max: 0 (Some entries have lat, long as 0,0 (49))  
Actual Max: - -87.52458739  
DISTINCT: 300054  
Missing: Yes (5615) ~ 0.7%

LOCATION  
TYPE: STRING  
Column is not required as it is combination of Latitude and Longitude.