

NYC MOTOR VEHICLE COLLISION PROJECT

Data Source:

<https://opendata.cityofnewyork.us/>

Dataset:

Motor Vehicle Collisions – Crashes

https://data.cityofnewyork.us/Public-Safety/Motor-Vehicle-Collisions-Crashes/h9gi-nx95/about_data

Data Provided By: NYPD (Police Department)

Dataset Owner: NYC OpenData

Introduction:

The Motor Vehicle Collisions crash table contains details on the crash event.

Each row represents a crash event.

The Motor Vehicle Collisions data tables contain information from all police reported motor vehicle collisions in NYC.

The police report (MV104-AN) is required to be filled out for collisions where someone is injured or killed, or where there is at least \$1000 worth of damage.

(https://www.nhtsa.gov/sites/nhtsa.dot.gov/files/documents/ny_overlay_mv-104an_rev05_2004.pdf)

It should be noted that the data is preliminary and subject to change when the MV-104AN forms are amended based on revised crash details.

For the most accurate, up to date statistics on traffic fatalities, please refer to the NYPD Motor Vehicle Collisions page (updated weekly) or Vision Zero View (updated monthly).

Due to success of the CompStat program, NYPD began to ask how to apply the CompStat principles to other problems. Other than homicides, the fatal incidents with which police have the most contact with the public are fatal traffic collisions.

Therefore, in April 1998, the Department implemented TrafficStat, which uses the CompStat model to work towards improving traffic safety.

Police officers complete form MV-104AN for all vehicle collisions.

The MV-104AN is a New York State form that has all of the details of a traffic collision.

Before implementing TrafficStat, there was no uniform traffic safety data collection procedure for all of the NYPD precincts.

Therefore, the Police Department implemented the Traffic Accident Management System (TAMS) in July 1999 in order to collect traffic data in a uniform method across the City.

TAMS required the precincts manually enter a few selected MV-104AN fields to collect very basic intersection traffic crash statistics which included the number of accidents, injuries and fatalities.

As the years progressed, there grew a need for additional traffic data so that more detailed analyses could be conducted.

The Citywide traffic safety initiative, Vision Zero started in the year 2014.

Vision Zero further emphasized the need for the collection of more traffic data in order to work towards the Vision Zero goal, which is to eliminate traffic fatalities.

Therefore, the Department in March 2016 replaced the TAMS with the new Finest Online Records Management System (FORMS). FORMS enables the police officers to electronically, using a

Department cellphone or computer, enter all of the MV-104AN data fields and stores all of the MV-104AN data fields in the Department's crime data warehouse.

Since all of the MV-104AN data fields are now stored for each traffic collision, detailed traffic safety analyses can be conducted as applicable.

About the dataset:

Data Last Updated

July 18, 2025

(Automatically Updated Daily)

Date Created

April 28, 2014

What's in this Dataset?

Rows: 2.19M

Columns: 29

Each row is a Motor Vehicle Collision

Row Identifier: COLLISION_ID

Columns(29):

Column Name	Description	API Field Name	Data Type
CRASH DATE	Occurrence date of collision	crash_date	Floating Timestamp
CRASH TIME	Occurrence time of collision	crash_time	Text
BOROUGH	Borough where collision occurred	borough	Text
ZIP CODE	Postal code of incident occurrence	zip_code	Text
LATITUDE	Latitude coordinate for Global Coordinate System, WGS 1984, decimal degrees (EPSG 4326)	latitude	Number
LONGITUDE	Longitude coordinate for Global Coordinate System, WGS 1984, decimal degrees (EPSG 4326)	longitude	Number
LOCATION	Latitude , Longitude pair	location	Location
ON STREET NAME	Street on which the collision occurred	on_street_name	Text

CROSS STREET NAME	Nearest cross street to the collision	off_street_name	Text
OFF STREET NAME	Street address if known	cross_street_name	Text
NUMBER OF PERSONS INJURED	Number of persons injured	number_of_persons_injured	Number
NUMBER OF PERSONS KILLED	Number of persons killed	number_of_persons_killed	Number
NUMBER OF PEDESTRIANS INJURED	Number of pedestrians injured	number_of_pedestrians_injured	Number
NUMBER OF PEDESTRIANS KILLED	Number of pedestrians killed	number_of_pedestrians_killed	Number
NUMBER OF CYCLIST INJURED	Number of cyclists injured	number_of_cyclist_injured	Number
NUMBER OF CYCLIST KILLED	Number of cyclists killed	number_of_cyclist_killed	Number
NUMBER OF MOTORIST INJURED	Number of vehicle occupants injured	number_of_motorist_injured	Number
NUMBER OF MOTORIST KILLED	Number of vehicle occupants killed	number_of_motorist_killed	Number
CONTRIBUTING FACTOR VEHICLE 1	Factors contributing to the collision for designated vehicle	contributing_factor_vehicle_1	Text
CONTRIBUTING FACTOR VEHICLE 2	Factors contributing to the collision for designated vehicle	contributing_factor_vehicle_2	Text
CONTRIBUTING FACTOR VEHICLE 3	Factors contributing to the collision for designated vehicle	contributing_factor_vehicle_3	Text
CONTRIBUTING FACTOR VEHICLE 4	Factors contributing to the collision for designated vehicle	contributing_factor_vehicle_4	Text
CONTRIBUTING FACTOR VEHICLE 5	Factors contributing to the collision for designated vehicle	contributing_factor_vehicle_5	Text
COLLISION_ID	Unique record code generated by system. Primary Key for Crash table.	collision_id	Number

VEHICLE TYPE CODE 1	Type of vehicle based on the selected vehicle category (ATV, bicycle, car/suv, ebike, escooter, truck/bus, motorcycle, other)	vehicle_type_code1	Text
VEHICLE TYPE CODE 2	Type of vehicle based on the selected vehicle category (ATV, bicycle, car/suv, ebike, escooter, truck/bus, motorcycle, other)	vehicle_type_code2	Text
VEHICLE TYPE CODE 3	Type of vehicle based on the selected vehicle category (ATV, bicycle, car/suv, ebike, escooter, truck/bus, motorcycle, other)	vehicle_type_code_3	Text
VEHICLE TYPE CODE 4	Type of vehicle based on the selected vehicle category (ATV, bicycle, car/suv, ebike, escooter, truck/bus, motorcycle, other)	vehicle_type_code_4	Text
VEHICLE TYPE CODE 5	Type of vehicle based on the selected vehicle category (ATV, bicycle, car/suv, ebike, escooter, truck/bus, motorcycle, other)	vehicle_type_code_5	Text

Project Overview

Tools Used:

MySQL: Store and query the cleaned data

Python: Data cleaning, transformation, and loading into MySQL

Power BI: Build the final dashboard using MySQL as the data source

STEP-BY-STEP PLAN:

Step 1: Download the dataset.

- Go to NYC Open Data – Motor Vehicle Collisions (Crashes)
- Click Export → CSV
- Save it as `nyc_motor_collisions.csv`

Step 2: Clean & Load Using Python

- Clean the data.
- Connect to MySQL and Load the data.

Step 3: Connect PowerBI to MySQL.

- Open Power BI Desktop
- Click Home → Get Data → MySQL database
- Enter your server and database (e.g., `localhost`, `nyc_data`)
- Load the `collisions` table

Step 4: Add DAX Measures

Total Collisions = `COUNTROWS(collisions)`

Total Injuries = `SUM(collisions[persons_injured])`

Total Fatalities = `SUM(collisions[persons_killed])`

Step 5: Create PowerBI Visuals.

- Card KPIs: Total collisions, Total injuries, Total fatalities
- Bar Chart: Collisions by borough
- Line Chart: Monthly Collision trends
- Donut Chart: Distribution by Borough
- Heatmap Table: Collisions by Hour and Borough
- Stacked Column Chart - Accidents by Year

Add Filters:

- Slicers for Borough, Year, Month, Contributing Factor