**1. Introduction**

The NYC Motor Vehicle Collisions Data Service is a web-based application designed to efficiently serve collision records to users. This documentation provides an overview of the system architecture, data model, implementation details, deployment instructions, and testing procedures.

**2. System Architecture**

The application follows a client-server architecture, with a Flask backend serving RESTful APIs and a frontend implemented using HTML, CSS, and JavaScript. The backend communicates with a MySQL database to retrieve collision records based on user queries.

**3. Data Model**

The MySQL database contains a single table named **crash**, which stores information about motor vehicle collisions. The table schema includes columns for attributes such as crash date, time, borough, location, and contributing factors.

Go to GitHub repository & navigate to (Database/DDL Script.sql)- provided to create database – nyccrashes in MySQL.

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

The "crash" table contains various columns to store information about motor vehicle collisions

1. **CRASH\_DATE (char(10))**: Stores the date of the crash in the format MM/DD/YYYY.
2. **CRASH\_TIME (char(5))**: Stores the time of the crash in the format HH:MM.
3. **BOROUGH (varchar(13))**: Represents the borough where the crash occurred.
4. **ZIP CODE (char(5))**: Stores the ZIP code of the location where the crash occurred.
5. **LATITUDE (double)**: Represents the latitude coordinate of the crash location.
6. **LONGITUDE (double)**: Represents the longitude coordinate of the crash location.
7. **LOCATION (char(25))**: Stores a brief description or address of the crash location.
8. **ON STREET NAME (varchar(32))**: Represents the name of the street where the crash occurred.
9. **CROSS STREET NAME (varchar(32))**: Represents the name of the cross street where the crash occurred.
10. **OFF STREET NAME (varchar(40))**: Represents the name of the off-street location where the crash occurred.
11. **NUMBER OF PERSONS INJURED (tinyint UNSIGNED)**: Stores the number of individuals injured in the crash.
12. **NUMBER OF PERSONS KILLED (tinyint UNSIGNED)**: Stores the number of individuals killed in the crash.
13. **NUMBER OF PEDESTRIANS INJURED (tinyint UNSIGNED)**: Stores the number of pedestrians injured in the crash.
14. **NUMBER OF PEDESTRIANS KILLED (tinyint UNSIGNED)**: Stores the number of pedestrians killed in the crash.
15. **NUMBER OF CYCLIST INJURED (tinyint UNSIGNED)**: Stores the number of cyclists injured in the crash.
16. **NUMBER OF CYCLIST KILLED (tinyint UNSIGNED)**: Stores the number of cyclists killed in the crash.
17. **NUMBER OF MOTORIST INJURED (tinyint UNSIGNED)**: Stores the number of motorists injured in the crash.
18. **NUMBER OF MOTORIST KILLED (tinyint UNSIGNED)**: Stores the number of motorists killed in the crash.
19. **CONTRIBUTING FACTOR VEHICLE 1 (varchar(53))**: Represents the primary contributing factor to the crash.
20. **CONTRIBUTING FACTOR VEHICLE 2 (varchar(53))**: Represents the secondary contributing factor to the crash.
21. **CONTRIBUTING FACTOR VEHICLE 3 (varchar(53))**: Represents the tertiary contributing factor to the crash.
22. **CONTRIBUTING FACTOR VEHICLE 4 (varchar(43))**: Represents the fourth contributing factor to the crash.
23. **CONTRIBUTING FACTOR VEHICLE 5 (varchar(43))**: Represents the fifth contributing factor to the crash.
24. **COLLISION\_ID (int, Primary Key)**: Unique identifier for each collision record.
25. **VEHICLE TYPE CODE 1 (varchar(38))**: Represents the type of vehicle involved in the crash (primary vehicle).
26. **VEHICLE TYPE CODE 2 (varchar(38))**: Represents the type of vehicle involved in the crash (secondary vehicle).
27. **VEHICLE TYPE CODE 3 (varchar(35))**: Represents the type of vehicle involved in the crash (third vehicle).
28. **VEHICLE TYPE CODE 4 (varchar(35))**: Represents the type of vehicle involved in the crash (fourth vehicle).
29. **VEHICLE TYPE CODE 5 (varchar(35))**: Represents the type of vehicle involved in the crash (fifth vehicle).

**4. Backend Implementation**

The backend is implemented using Flask, a lightweight web framework for Python. It provides endpoints to fetch collision records based on user-defined parameters such as borough, crash date, and collision ID. Pagination is implemented to efficiently handle large datasets.

**5. Frontend Implementation**

The frontend is designed using HTML, CSS, and JavaScript. It includes a user-friendly form to input query parameters and dynamically fetches and displays collision records using asynchronous requests to the backend API.

**6. Deployment Instructions**

Instructions for setting up and deploying the application locally for testing purposes are provided in the README.md file in the project repository.

**7. Testing**

The application has been thoroughly tested to ensure functionality, performance, and responsiveness. Performance testing has been conducted to identify and optimize bottlenecks.

**8. Future Enhancements**

Future enhancements to the application may include:

* Implementation of caching mechanisms for improved performance.
* Enhancement of error handling and validation.
* Addition of user authentication and authorization.
* Integration with external APIs for additional data sources.
* Creating a dimensional model for optimizing the data storage & retrieval.

Purposed Dimensional Model –

**DateTime Dimension:**

* **Date\_Key (Primary Key)**: Unique identifier for each date.
* **Date**: Date of the crash (MM/DD/YYYY).
* **Time**: Time of the crash (HH:MM).
* **Day**: Day of the week (e.g., Monday, Tuesday).
* **Month**: Month of the year (e.g., January, February).
* **Year**: Year of the crash.
* **Quarter**: Quarter of the year (e.g., Q1, Q2).
* **Hour**: Hour of the day (24-hour format).
* **Minute**: Minute of the hour.

**Borough Dimension:**

* **Borough\_Key (Primary Key)**: Unique identifier for each borough.
* **Borough**: Name of the borough where the crash occurred.

**Contributing Factor Vehicle Dimension:**

* **Factor\_Vehicle\_Key (Primary Key)**: Unique identifier for each contributing factor vehicle.
* **Contributing Factor Vehicle**: Description of the contributing factor to the crash.

**Vehicle Type Code Dimension:**

* **Vehicle\_Type\_Key (Primary Key)**: Unique identifier for each vehicle type.
* **Vehicle Type Code**: Type code representing the vehicle involved in the crash.

**Fact Collision:**

* **Collision\_ID (Primary Key, Foreign Key)**: Unique identifier for each collision record.
* **Date\_Key (Foreign Key)**: Reference to the DateTime Dimension.
* **Borough\_Key (Foreign Key)**: Reference to the Borough Dimension.
* **ZIP Code**: ZIP code of the crash location.
* **Latitude**: Latitude coordinate of the crash location.
* **Longitude**: Longitude coordinate of the crash location.
* **On Street Name**: Name of the street where the crash occurred.
* **Cross Street Name**: Name of the cross street where the crash occurred.
* **Off Street Name**: Name of the off-street location where the crash occurred.
* **Number of Persons Injured**: Number of individuals injured in the crash.
* **Number of Persons Killed**: Number of individuals killed in the crash.
* **Number of Pedestrians Injured**: Number of pedestrians injured in the crash.
* **Number of Pedestrians Killed**: Number of pedestrians killed in the crash.
* **Number of Cyclists Injured**: Number of cyclists injured in the crash.
* **Number of Cyclists Killed**: Number of cyclists killed in the crash.
* **Number of Motorists Injured**: Number of motorists injured in the crash.
* **Number of Motorists Killed**: Number of motorists killed in the crash.
* **Factor\_Vehicle\_Key (Foreign Key)**: Reference to the Contributing Factor Vehicle Dimension.
* **Vehicle\_Type\_Key (Foreign Key)**: Reference to the Vehicle Type Code Dimension.
* **Vehicle Type Code 1**: Type of the primary vehicle involved in the crash.
* **Vehicle Type Code 2**: Type of the secondary vehicle involved in the crash.
* **Vehicle Type Code 3**: Type of the third vehicle involved in the crash.
* **Vehicle Type Code 4**: Type of the fourth vehicle involved in the crash.
* **Vehicle Type Code 5**: Type of the fifth vehicle involved in the crash.