Data Document and ERD

I have created 7 tables for my data of NYC shooting incidents.

Table 1: unique shooting

It is the first main table of my database. It contains the information of the incident key, the borough, the location, latitude and longitude where the incident occurred, the time when the incident occurred.

The create syntax of it is:

```
CREATE TABLE `unique_shooting` (
  `id` int NOT NULL AUTO_INCREMENT,
  `incident_key` int DEFAULT NULL,
  `borough_id` int DEFAULT NULL,
  `location_id` int DEFAULT NULL,
  `precinct` int DEFAULT NULL,
  `latitude` float DEFAULT NULL,
  `longitude` float DEFAULT NULL,
  `occur_time` datetime DEFAULT NULL,
  PRIMARY KEY (`id`),
  KEY `unique_shooting_borough_id_fk` (`borough_id`),
  KEY `unique_shooting_location_id_fk` (`location_id`),
  CONSTRAINT `unique_shooting_borough_id_fk` FOREIGN KEY (`borough_id`)
REFERENCES `borough` (`id`) ON DELETE CASCADE,
  CONSTRAINT `unique_shooting_location_id_fk` FOREIGN KEY (`location_id`)
REFERENCES `location` (`id`) ON DELETE CASCADE
) ENGINE=InnoDB AUTO_INCREMENT=512 DEFAULT CHARSET=utf8mb4
COLLATE=utf8mb4_0900_ai_ci
```

Table 2: perp_and_vict

It is the second main table of my database. In each row, it contains the information of a pair of perpetrator and victim like their race, sex and age. For a single shooting incident, there may be multiple pairs of perpetrator and victim, so in each row, the id of the incidents is also provided. Besides, for each pair of perpetrator and victim, a flag of 'is_murder' is given to indicate whether the shooting is considered as a murder.

The create syntax of it is:

```
CREATE TABLE `perp_and_vict` (
  `id` int NOT NULL AUTO_INCREMENT,
  `incident_id` int DEFAULT NULL,
  `perp_sex_id` int DEFAULT NULL,
  `perp_age_id` int DEFAULT NULL,
  `perp_race_id` int DEFAULT NULL,
  `vict_sex_id` int DEFAULT NULL,
  `vict_age_id` int DEFAULT NULL,
  `vict_race_id` int DEFAULT NULL,
  `is_murder` tinyint(1) DEFAULT NULL,
```

```
  PRIMARY KEY (`id`),
  KEY `perp_and_vict_age_group_id_fk` (`perp_age_id`),
  KEY `perp_and_vict_age_group_id_fk_2` (`vict_age_id`),
  KEY `perp_and_vict_race_id_fk` (`vict_race_id`),
  KEY `perp_and_vict_race_id_fk_2` (`perp_race_id`),
  KEY `perp_and_vict_sex_id_fk` (`perp_sex_id`),
  KEY `perp_and_vict_sex_id_fk_2` (`vict_sex_id`),
  KEY `perp_and_vict_unique_shooting_id_fk` (`incident_id`),
  CONSTRAINT `perp_and_vict_age_group_id_fk` FOREIGN KEY (`perp_age_id`)
REFERENCES `age_group` (`id`) ON DELETE CASCADE,
  CONSTRAINT `perp_and_vict_age_group_id_fk_2` FOREIGN KEY
(`vict_age_id`) REFERENCES `age_group` (`id`) ON DELETE CASCADE,
  CONSTRAINT `perp_and_vict_race_id_fk` FOREIGN KEY (`vict_race_id`)
REFERENCES `race` (`id`) ON DELETE CASCADE,
  CONSTRAINT `perp_and_vict_race_id_fk_2` FOREIGN KEY (`perp_race_id`)
REFERENCES `race` (`id`) ON DELETE CASCADE,
  CONSTRAINT `perp_and_vict_sex_id_fk` FOREIGN KEY (`perp_sex_id`)
REFERENCES `sex` (`id`) ON DELETE CASCADE,
  CONSTRAINT `perp_and_vict_sex_id_fk_2` FOREIGN KEY (`vict_sex_id`)
REFERENCES `sex` (`id`) ON DELETE CASCADE,
  CONSTRAINT `perp_and_vict_unique_shooting_id_fk` FOREIGN KEY
(`incident_id`) REFERENCES `unique_shooting` (`id`) ON DELETE CASCADE
) ENGINE=InnoDB AUTO_INCREMENT=512 DEFAULT CHARSET=utf8mb4
COLLATE=utf8mb4_0900_ai_ci
```

Table 3: sex

There are two columns in this table, which are id and the specific sex. It is created to avoid data duplication in the perp_and_vict table.

The create syntax of it is:

```
CREATE TABLE `sex` (
  `id` int NOT NULL AUTO_INCREMENT,
  `sex` varchar(255) DEFAULT NULL,
  PRIMARY KEY (`id`)
) ENGINE=InnoDB AUTO_INCREMENT=4 DEFAULT CHARSET=utf8mb4
COLLATE=utf8mb4_0900_ai_ci
```

Table 4:age_group

There are two columns in this table, which are id and the specific age_group. It is created to avoid data duplication in the perp_and_vict table.

The create syntax of it is:

```
CREATE TABLE `age_group` (
  `id` int NOT NULL AUTO_INCREMENT,
  `age_group` varchar(255) DEFAULT NULL,
  PRIMARY KEY (`id`)
```

```
) ENGINE=InnoDB AUTO_INCREMENT=8 DEFAULT CHARSET=utf8mb4
COLLATE=utf8mb4_0900_ai_ci
```

Table 5: race

There are two columns in this table, which are id and the specific race. It is created to avoid data duplication in the perp_and_vict table.

The create syntax of it is:

```
CREATE TABLE `race` (
  `id` int NOT NULL AUTO_INCREMENT,
  `race` varchar(255) DEFAULT NULL,
  PRIMARY KEY (`id`)
) ENGINE=InnoDB AUTO_INCREMENT=7 DEFAULT CHARSET=utf8mb4
COLLATE=utf8mb4_0900_ai_ci
```

Table 6: borough

There are two columns in this table, which are id and the specific borough. It is created to avoid data duplication in the unique_shooting table.

The create syntax of it is:

```
CREATE TABLE `borough` (
  `id` int NOT NULL AUTO_INCREMENT,
  `borough` varchar(255) DEFAULT NULL,
  PRIMARY KEY (`id`)
) ENGINE=InnoDB AUTO_INCREMENT=6 DEFAULT CHARSET=utf8mb4
COLLATE=utf8mb4_0900_ai_ci
```

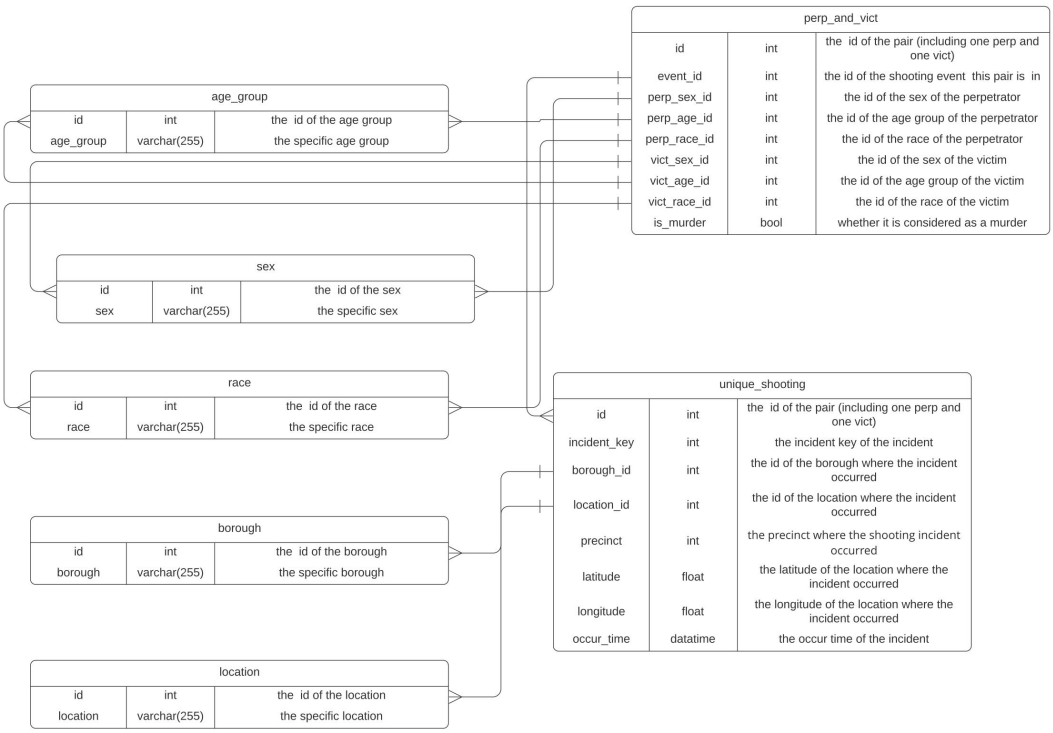
Table 7: location

There are two columns in this table, which are id and the specific location. It is created to avoid data duplication in the unique_shooting table.

The create syntax of it is:

```
CREATE TABLE `location` (
  `id` int NOT NULL AUTO_INCREMENT,
  `location` varchar(255) DEFAULT NULL,
  PRIMARY KEY (`id`)
) ENGINE=InnoDB AUTO_INCREMENT=12 DEFAULT CHARSET=utf8mb4
COLLATE=utf8mb4_0900_ai_ci
```

The ERD of my tables is:


perp_and_vict
id int the id of the pair (including one perp and one vict)
event_id int the id of the shooting event this pair is in
perp_sex_id int the id of the sex of the perpetrator
perp_age_id int the id of the age group of the perpetrator
perp_race_id int the id of the race of the perpetrator
vict_sex_id int the id of the sex of the victim
vict_age_id int the id of the age group of the victim
vict_race_id int the id of the race of the victim
is_murder bool whether it is considered as a murder
age_group
id int the id of the age group
age_group varchar(255) the specific age group
sex
id int the id of the sex
sex varchar(255) the specific sex
race
id int the id of the race
race varchar(255) the specific race
unique_shooting
id int the id of the pair (including one perp and one vict)
incident_key int the incident key of the incident
borough_id int the id of the borough where the incident occurred
location_id int the id of the location where the incident occurred
precinct int the precinct where the shooting incident occurred
latitude float the latitude of the location where the incident occurred
longitude float the longitude of the location where the incident occurred
occur_time datatime the occur time of the incident
borough
id int the id of the borough
borough varchar(255) the specific borough
location
id int the id of the location
location varchar(255) the specific location