Database architecture

- **The database plan to use for storage.**

  offline postgres

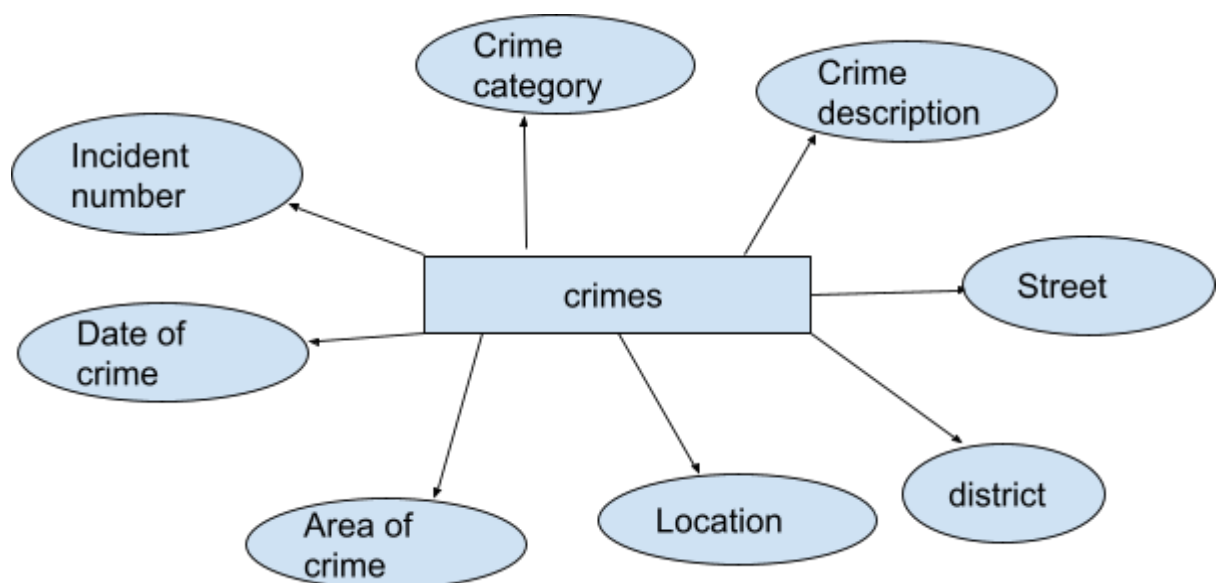- **Where the data is loaded from**

  Use a database-query of Database (underconsideration) to parse the data file and feed it in to the associated data base table

- **Data Cleaning(transformations done to load the data to the database)**
- ❖ Convert the 'Occurred_On_Date' column into a datetime format
- ❖ Fill 'N' in the 'Shooting' column,assume 'N' were not shooting.
- ❖ Removed the remaining Null/NaN values as it is small part of the crime data set
- ❖ Drop duplicates Rows in the dataset
- ❖ Replace " " in 'reporting_area' column to 0, means reporting_area not known

  After cleaning 304637 rows left

- **E-R diagram**

- **Connections to/from the database for Time Series Analysis module**
- ➔ Connected to the database with credential
- ➔ Use query to retrieve from database
- ➔ Fetch all row from the table
- ➔ Commit all changes in the database
- ➔ Close the connection