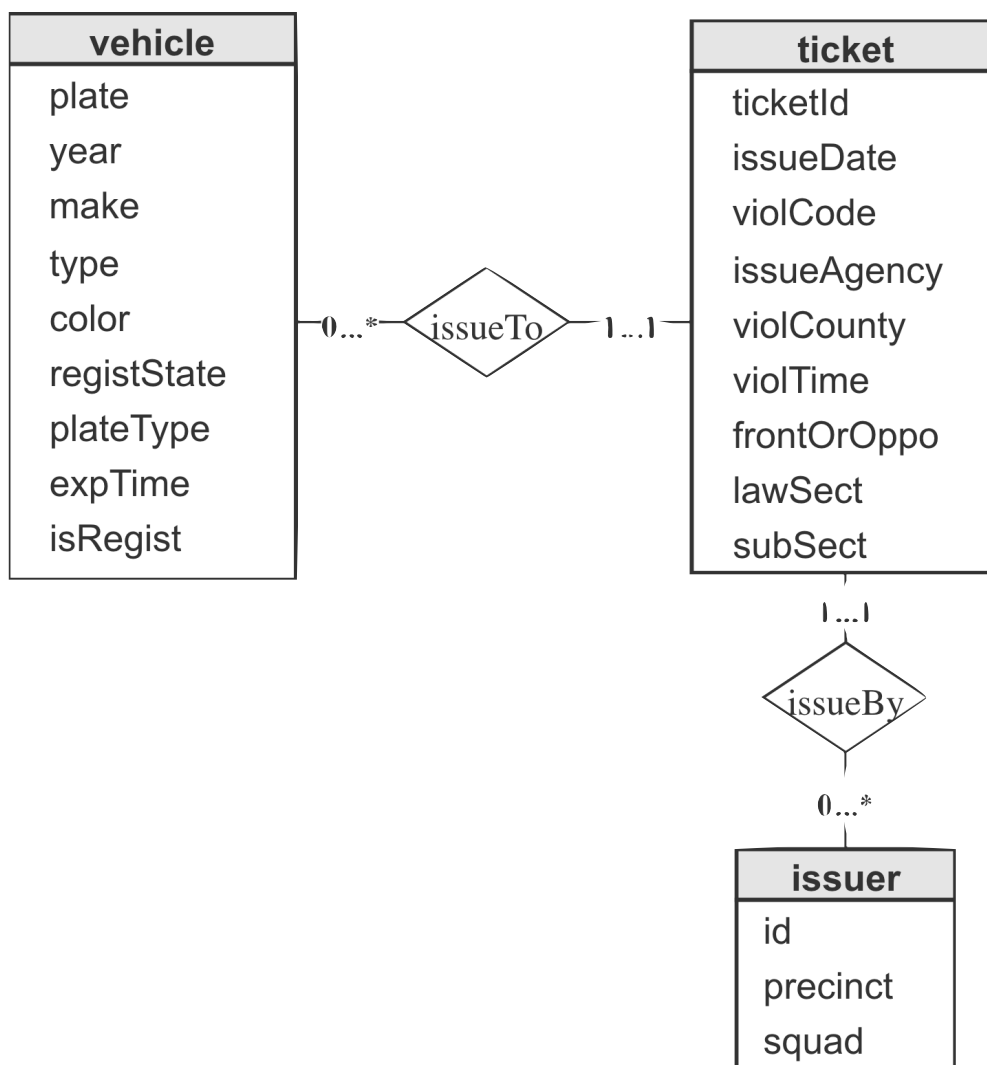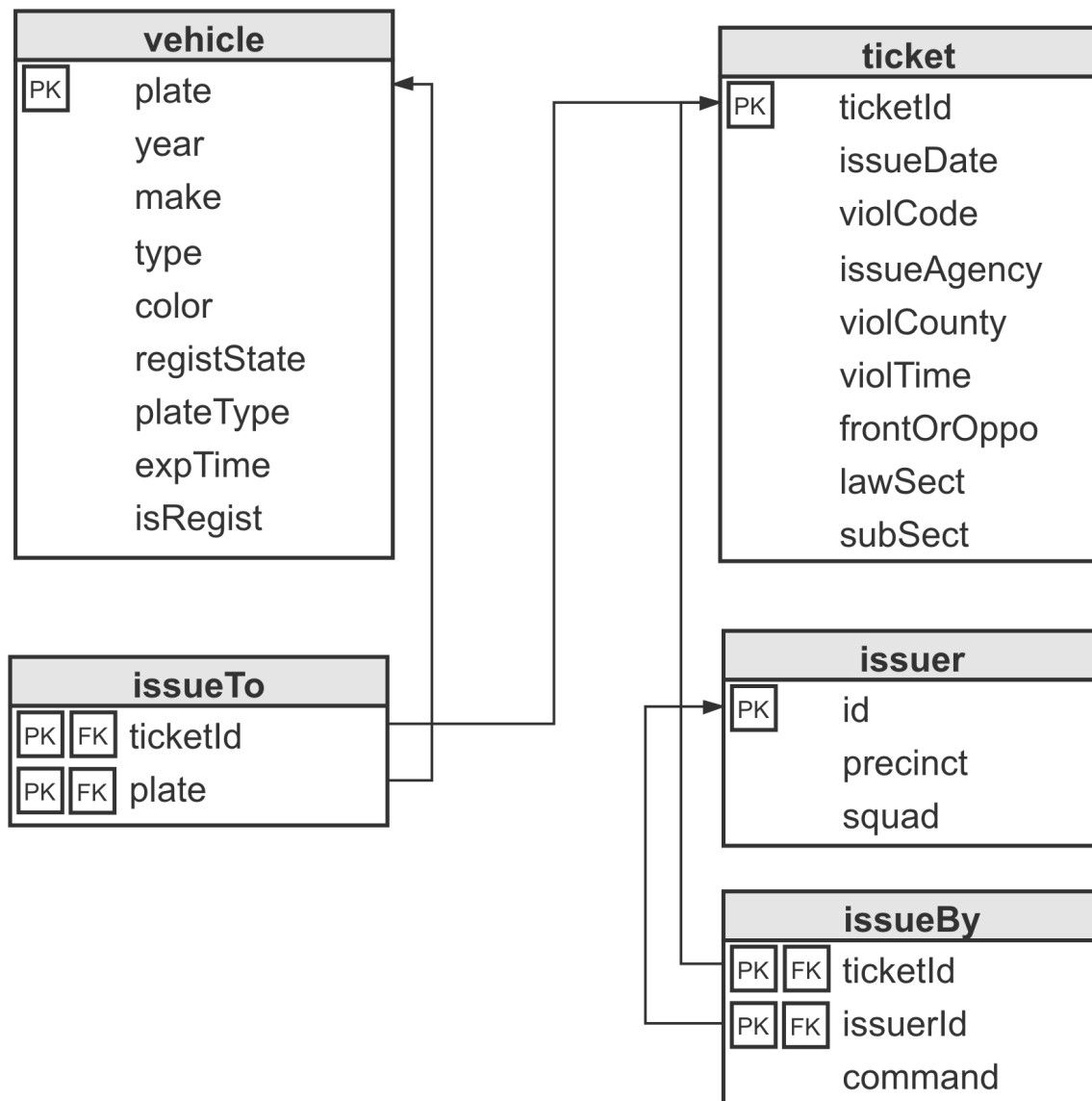# Phase 1 report
group 9

The dataset we use is NYC Parking Tickets, an open dataset. It contains every parking ticket issued in NYC from Aug 2013-June 2017. The main information contains a ticket issued by a officer to a vehicle.

As for our project, we abstract several entities and relationship among them from the original singe table. Then we create our own Entity-Relationship Diagram and Relational Model. The ER diagram and Relational Model are shown below.

ER diagram.

Relational Model:

When we deal with original data file, we separate it into several small data files. These smaller file matches our relational model, since the original file is too big and contains too much duplicate attributes to be loaded. In our program, we use JDBC batch insert to improve performance of our program.

Program source file and sql script files are in the same folder with this pdf.