**Assignment - Reducing Datasets**

Shaun Pritchard

Rasmussen College

QMB4000

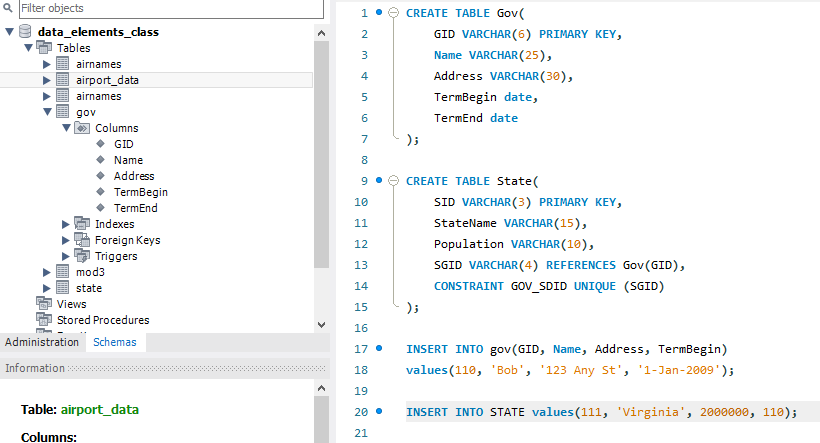
Benjamin Tasker

December 12, 2020

I created some datasets of me to express and show how relationships work based on the question. Using SQL to use to in reducing the dimensionality of data from Many-to-Many relationship to a One-to-One relationship or a One-to-many relationships as described below

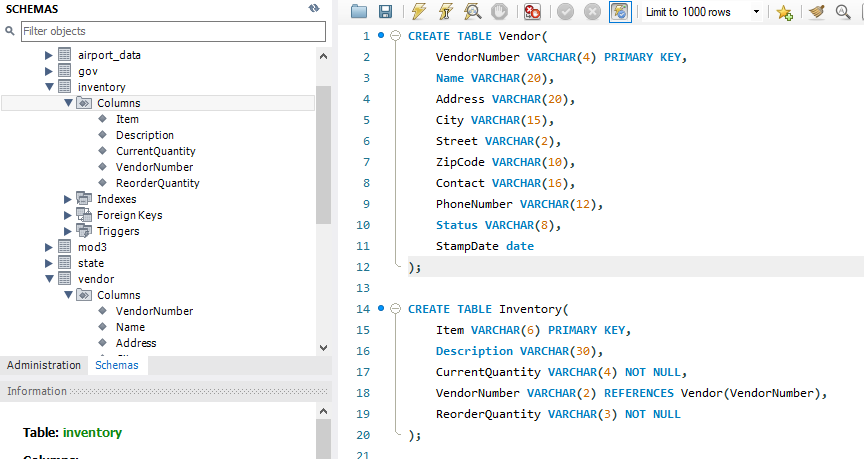
**One-to-one (1:1)**

A relationship is one-to-one if and only if one record from table A is related to a maximum of one record in table B. To establish a one-to-one relationship, the primary key of table B (with no orphan record) must be the secondary key of table A (with orphan records).



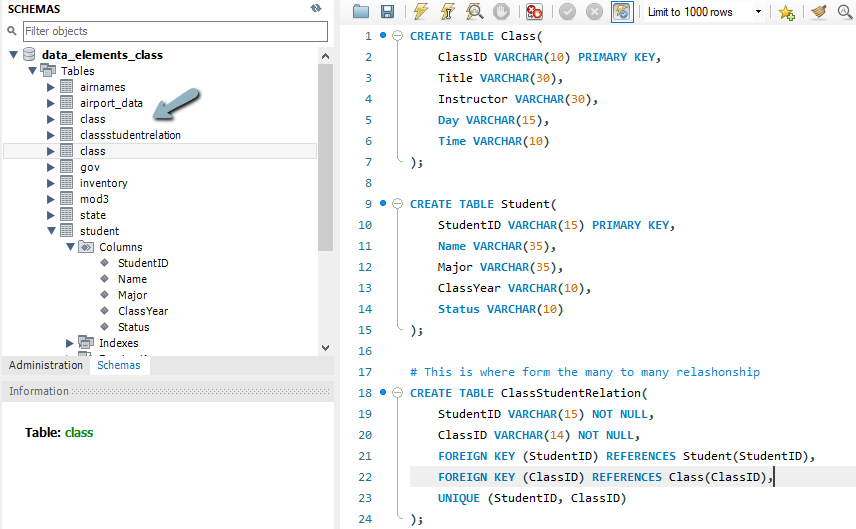
**One-to-many (1:M)**

A relationship is one-to-many if and only if one record from table A is related to one or more records in table B. However, one record in table B cannot be related to more than one record in table A. To establish a one-to-many relationship, the primary key of table A (the "one" table) must be the secondary key of table B (the "many" table).

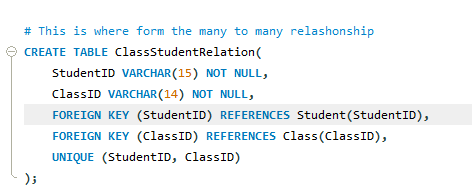


**Many-to-many (M:M)**

A relationship is many-to-many if and only if one record from table A is related to one or more records in table B and vice-versa. To establish a many-to-many relationship, create a third table called "***ClassStudentRelation***" which will have the primary keys of both table A and table B.



**Reducing Dimensionality:**



By subsiding another table to combine and join the references from the student and class table I have built a table that reduces the dimensions of both tables into just 1. 1:M one to many.

**How your data changed and how that could impact an analysis.**

A large-scale volumes of data dimensionality can cause many problems depending on the size of the dataset in question. Data reduction techniques are used to compression-based data reduction methods and algorithms for preprocessing, cluster-level data deduplication, redundancy elimination, and implementation of network (graph) theory concepts. Essentially the data reduction gives a pure sample of the data that makes running analysis more viable and accurate.