Deleting Duplicate Database Rows

By Roger Keays, 13 July 2008

Coming up with a query to find duplicate rows in a database is not as easy as it should be and always turns out dog slow. Here's a simple alternative which I used when I had to do just that:

- Dump the database.
- Delete all the records.
- Add a UNIQUE constraint on the natural key.
- Reload the records from the database dump.

If you have duplicates in the dump, the INSERT will fail and they won't make it back into the database.

Easy or WHAT?!

About Roger Keays



Roger Keays is an artist, an engineer, and a student of life. He has no fixed addressand has leftfootprints on 40-something different countries around the world. Roger is addicted to surfing. His other interests are music, psychology, languages, the proper use of semicolons, and finding good food.