Team17-Homework 5: Report

Team Members: Jay Patel, Chase Carroll

Breakdown of data.sql

The data generation process was broken up into roughly three segments: Input preparation, input processing, and output processing.

In the input processing section, all the data in the provided text files was pulled, formatted to have no extraneous characters, and then placed into arrays. There were no other notable functions of this portion of the code, nor anything particularly clever done to write it, except that the same code format was used to read and interpret each text file.

In the input processing section, functions were written for each table present in the database. The functions used whatever means were necessary for writing single line SQL value statements for their respective table. This often involved taking the length of an array of processed inputs from another function call in this section, and that length was then used to randomly generate values for foreign keys that would be valid in other tables. The modularity of the functions meant they could be tested individually, thus allowing us to easily debug the problems that appeared while trying to deal with foreign keys. At the end, each function returned an array with a single line of SQL for use as values to be inserted.

In the output processing section, the arrays created by the functions of section two were then written to data.sql, along with other necessary SQL code, to form complete table data insert statements. The file was progressively made longer as more of section two's functions were tested together, and a rule of "three times the charm" was used to determine if the SQL worked, generating a new data.sql file each time.

Afterthoughts

Following the completion of the script, some things came to light about our team organization, but very little in terms of code revelations was discovered. Aside from learning about the intricacies of foreign keys and their heinous existence, we found ourselves to have a hard time properly communicating and sharing code during the project. Our most useful cooperation occurred within the last few days before the due date, as we began using phone calls and Google Hangouts to code in tandem, but before then, we had even been separate enough to write two- partially completed- entirely different data.sql files. Because of this, we have begun to think about other ways to share and manage code than Google Drive, and GitHub has come to mind as a future solution to try.