**Database Construction**

**Data Storage**

Two stages of data input were implemented. On the first stage, data were exactly entered as they appeared on 89 questionnaires, either qualitative or quantitative. An original Excel file and two back-up files were saved on separate computer devices (a laptop hard disk, a flash memory disk and a portable USB hard disk), considering that computer crash caused by virus is rather not unusual nowadays. In the second round, five external inspectors (three with bachelors’ degree and two with masters’ degree) were employed to check the correctness of the original input. It was hoped that unnecessary mistakes could be reduced to the minimum by doing so, if not none.

**Database Design**

In order to facilitate the following statistical analysis and econometric inferences, quantitative data were coded in a systematic format and entered in SPSS. For instance, the cities where firms were located were coded from 1 to 10 (e.g. the code for Guangzhou was 1). The two-digit CNSIC code was implemented to match each firm’s industry type (e.g. a leather producer was coded 19). And true or false questions were coded as 1 for yes and 0 for no. It was slightly more complicated to deal with questions that allowed multiple answers. In this case, each available option would be designed as a variable that had the value of 1 (chosen) and 0 (not chosen). Therefore, a multiple-answer question with eight options could be broken down into eight variables in the database.

As most data were collected at the time of interview, a cross-sectional database of 83 firms (excluding 6 public owned firms) was constructed with more than 250 variables and over 20,000 data points. Likewise, the database files were saved as the original and the backup files separately. As seen above, a number of steps were taken to ensure the applicability, credibility and safety of the data and to facilitate the data manipulation in next chapters.