Using Audit files to get the time of each question (TIEQ) in China Family Panel Study (CFPS)

Yongjian Zhang, Shuai Sun, Jiahui Yao , Peking university

Abstract

To verify the work of interviewers and ensure the high-quality interview of the surey,we want to know all the operations of interviewers during the interview .With the Blaise system all the interviewer's operations are recorded in audit trails files .We can get a lot of para data through analyze audit files, it contains a so many information that we can't use these files directly. So we should develop special softwares to help us analysis audit trails files.

In China Family Panel Study(CFPS),we use audit files to compute the answertime of each question(TIEQ). With TIEQ data we can do lots of analysis, such as verify the completion of the work of interviewers, determining the Interviewer's remuneration and labor costs, providing the basis for judging whether this sample need to re-visit and so on.

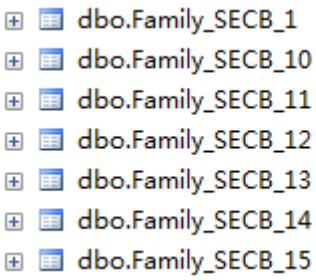
This paper describe the solution of getting TIEQ from audit files and challenges we encountered. And also discuss the future use of adt files in helping the analysis of para data.

1 Create Sqlserver Table

In China Family Panel Study(CFPS),our development environment is .net and Internet Information Services, So we had to use the sqlserver database.Now Our data analysts want to read data directly from sqlserver,but not bdb files or sas, so we need to copy the bdb data to sqlserver database. At the beginning, we tried to use the Blaise ole db workshop, but The effect is not very satisfactory, so we developed an application to implement this, fields in the sqlserver table and fields in the bdb files are one-to-one correspondence.As we know, in sqlserver the maximum number of a table is one thousand, Unfortunately, our questionnaires have several thousands, even more than ten thousands fields, in this case, the Ole db workshop will not be able to meet our needs. So we had to find something else to solve this problem

Now,we have several questionnaires in one project, we create the sql server tables named questionnaire name and '_'and block name,for example, the questionnaire named family,there are many fields named SecA.QAName, SecA.QAAge, then we can create a table named Famliy_SecA, QAName and QAAge are the properties of the table. In this way, it will be well organized.

picture p.1 show the sqlserver struct

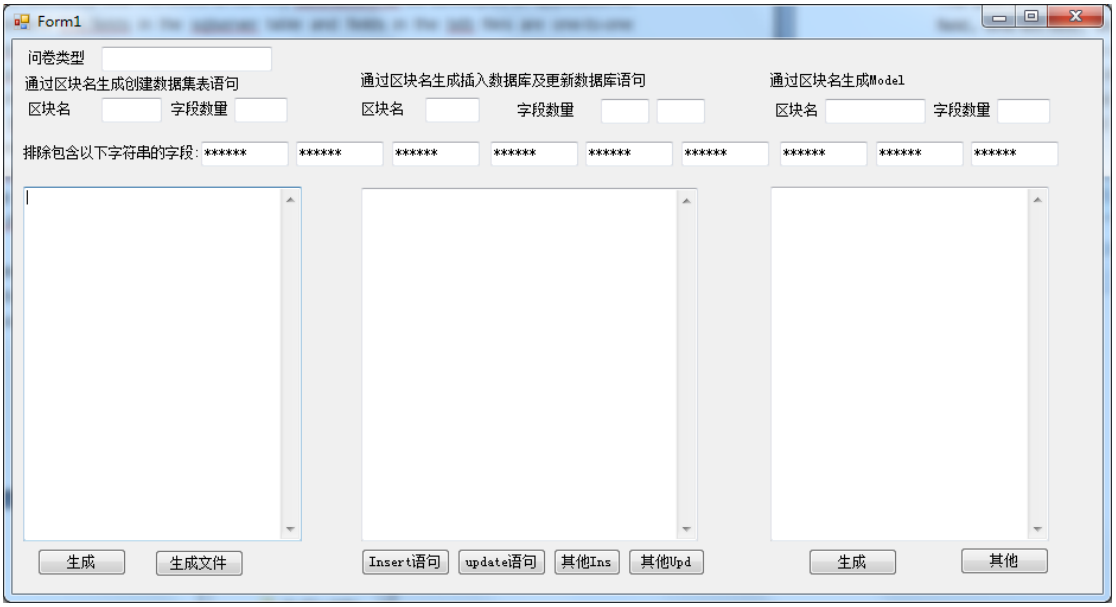


p.1.1

2 The Application named HelloBlaise

In developing, we use an application called hello blaise, this application can get all the fields contained in bmi files, and you can also list the fields that you did not want, the application will remove them. Despite this, we still have a problem, there are so many fields, should we write them

one by one? Absulately not, Use this tool , we can get the create table statement, the insert statement, also update statement,



p.1.2

3 Analysis audit trail files

```
"2011/6/24 15:05:49","Start Session"
"2011/6/24 15:05:49","Enter Form:0","Key:"
"2011/6/24 15:05:49","Enter Field:SampleID","Status:Normal","Value:"
"2011/6/24 15:05:53","Action:Store Field Data","Field:SampleID"
"2011/6/24 15:05:53","Leave Field:SampleID","Cause:Next
Field","Status:Normal","Value:1"
"2011/6/24 15:05:53","Enter Field:PreSecA.WAName","Status:Normal","Value:"
"2011/6/24 15:05:55","Action:Store Field Data","Field:PreSecA.WAName"
"2011/6/24 15:05:55","Leave Field:PreSecA.WAName","Cause:Next
Field","Status:Refusal","Value:"
"2011/6/24 15:05:56","Action:Error Jump","Field:PreSecA.WAName"
```

p.1.3

The data is stored as the picture ‘p.1.3’ shown. the contents contains entered time field, leave time field, final value and other information. The principle of our program is : to traverse the adt file for each row,first find “enter field :sampleID”, take the first part of this line(2011/6/24 15:05:49), we temporarily call it ‘enter-time’, second find ‘leave field:sampleid’, also take the first part of this line,(2011/6/24 15:05:53’),we temporarily call it ’leave-time’. The third use leave-time minus enter-time , wo get a time, that is what we need.

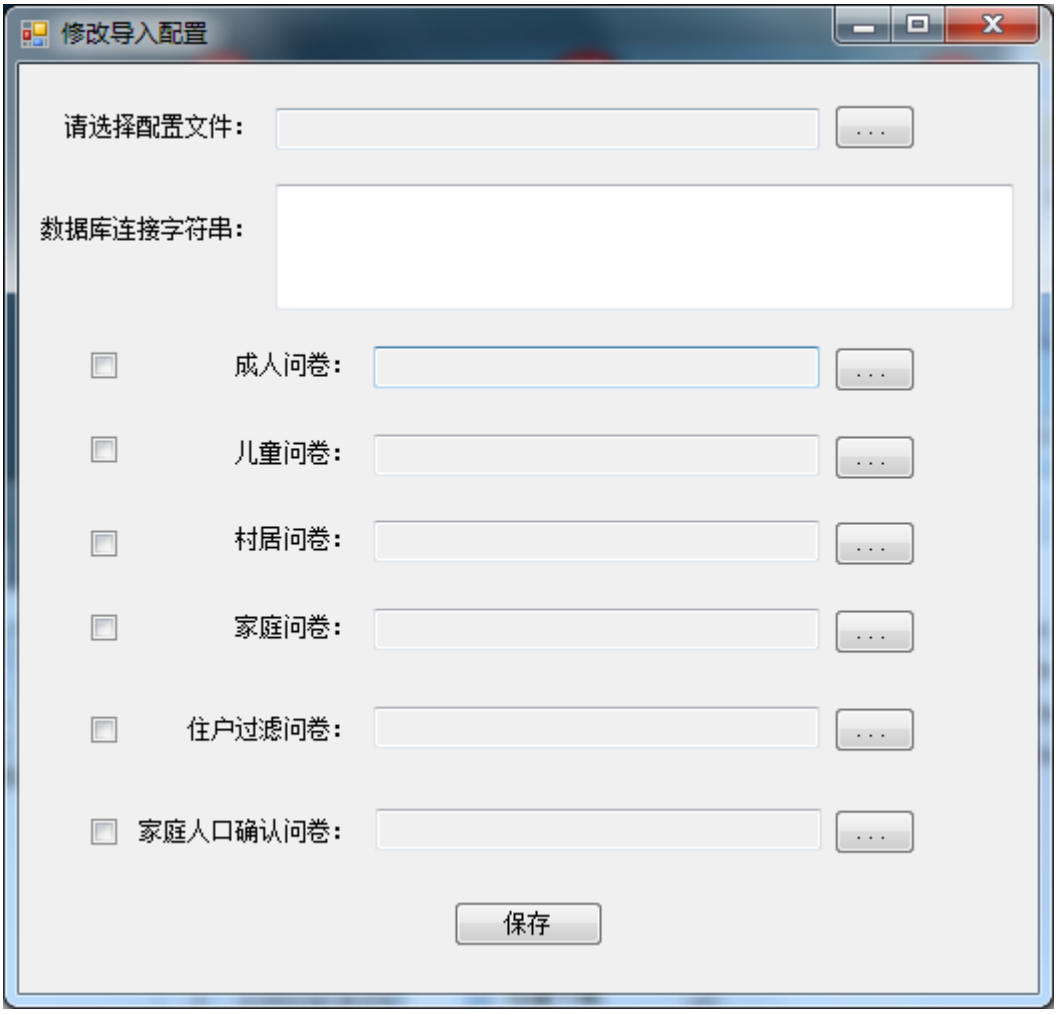
4 AnswerTimeSpanImport application

Since we get the time , wo can run our answerTimeSpanImport application, this application show the process of each questionnaire, during the run time of the application ,lots of information will be written into an txt file. such as the exception happened in the application ,the number of success imported ,also the number of fail imported and so on. we can select what should be written.
In our project , we set the application autorun one time everyday, we can use the Microsoft's operating system to realize the function. Click the [start menu], you will see find a button named [attachment],click ,you can see the button named [task scheduler] ,click it ,the system will teach you how to set the time when the application run , But a prerequisite is your computer should not be shut down

p.1.4

5 Audit trail files' path config tool

As shown in p.1.4, we have six questionnaires, the truth is the audit trail files are stored in the Hard disk in different paths, if we write these paths in the source codes, should we copy the files to the path that we write in the source codes when we use this import application? How do we control this? Don't worry , we has another tool, use this tool , we can chose the path that the audit trail files be stored, also we can set the sql server database.



P.1.5

6 Summary

The Blaise System open many interfaces,That we can use its dll files ,then we can do lots of things by ourselves,we can development lots of applications to assist our survey ,that's amazing.

The main purpose of our survey is to provide survey data with a national sample, regarding various aspects of social phenomena, different scholars or organizations may use this data to help them to do some research ,when the government can launch policies they can refer this data. As the Data gatherers, we have an obligation to ensure the survey data is maintained correctly.so we collect TIEQ, to assist the checks to check the interviewers' work.