Making Audit Trails Accessible for the CE Quarterly Interview Survey

Brandon Kopp, Yezzi Angi Lee, Lucilla Tan

CE Methods Symposium
July 17, 2018
Making Audit Trails Accessible

- Record of all navigation and data transactions within the instrument
- Built into the Blaise programming language
- Raw audit trail files are provided for each case for a given month


Example Audit Trail

"6/12/2018 7:34:45 AM" "Leave Field: Sect03.Sect3A2.OTHERHOM", "Cause: Next Field", "Status: Normal", "Value: 2"

"6/12/2018 7:34:45 AM" "Enter Field: Sect03.Sect3A2.BUSPROP3", "Status: Normal", "Value: 2"

"6/12/2018 7:34:45 AM" "Leave Field: Sect03.Sect3A2.BUSPROP3", "Cause: Next Field", "Status: Normal", "Value: 2"

"6/12/2018 7:34:45 AM" "Enter Field: Sect03.Sect3A2.RESBUSPR", "Status: Normal", "Value: 2"

"6/12/2018 7:34:45 AM" "Leave Field: Sect03.Sect3A2.RESBUSPR", "Cause: Next Field", "Status: Normal", "Value: 2"

"6/12/2018 7:34:45 AM" "Enter Field: Sect03.Sect3A2.SECHOME", "Status: Normal", "Value: 1"

"6/12/2018 7:34:45 AM" "Leave Field: Sect03.Sect3A2.SECHOME", "Cause: Next Field", "Status: Normal", "Value: 1"

"6/12/2018 7:34:45 AM" "Enter Field: Sect03.Sect3A2.SEC_NUM", "Status: Normal", "Value: 1"

"6/12/2018 7:34:46 AM" "Leave Field: Sect03.Sect3A2.SEC_NUM", "Cause: Next Field", "Status: Normal", "Value: 1"

"6/12/2018 7:34:46 AM" "Enter Field: Sect03.Sect3A2.NOBUild", "Status: Normal", "Value: 2"

"6/12/2018 7:34:46 AM" "Leave Field: Sect03.Sect3A2.NOBUild", "Cause: Next Field", "Status: Normal", "Value: 2"

"6/12/2018 7:34:46 AM" "Enter Field: Sect03.Sect3A2.NOLONGER", "Status: Normal", "Value: 1"

"6/12/2018 7:34:46 AM" "Leave Field: Sect03.Sect3A2.NOLONGER", "Cause: Next Field", "Status: Normal", "Value: 1"

"6/12/2018 7:34:46 AM" "Enter Field: Sect03.Sect3A2.NUMPRTTY", "Status: Normal", "Value: 1"

*Does not contain actual respondent data*
In what month did you make the first payment on this mortgage?

- January
- February
- March
- April
- May
- June
- July
- August
- September
- October
- November
- December

Section 3, Part F: Mortgages and Lump Sum Home Equity Loans

<table>
<thead>
<tr>
<th>Prop #</th>
<th>Mortgage Term</th>
<th>Current Interest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Fixed Rate</td>
</tr>
</tbody>
</table>

Loan Number: 101
Continue/Delete: 1
Refinance: 2
Payment Year: 2018

Payment Month: 2
Original Mortgage: 

Payment Includes:
- Payments Specify
What was the amount of this mortgage when you obtained it, not including any interest?

<table>
<thead>
<tr>
<th>Which Field</th>
<th>Status</th>
<th>Value</th>
</tr>
</thead>
</table>

Datetime Stamp: 6/12/2018 7:24:12 AM
Is this a 30 year mortgage, a 15 year mortgage, or something else?

1. 30-year
2. 15-year
3. Something else

Data Were Recorded

Entered Field

Exited Field


Making Audit Trails Accessible

- How long did it take to answer the question?
- Was the answer changed?
- Were any errors triggered?
- Did the interviewer access any help materials?
- Did the interviewer come back to this question later?

35 s


TX"


Yes.
Value was added
Making Audit Trails Accessible

- Chan, Ho, Kang, Smith, & Tan (2004)
  - Used SAS to parse audit trail files into four tables
  - Prepared documentation for the audit trail tables
  - Conducted several analyses
  - Explored the idea of creating a long-term storage for future months of audit trail files

- Priorities changed and no further work was completed using the parsing program or the created datasets
Making Audit Trails Accessible

- This project looked to reprise the Chan et al. (2004) work with some modern updates
- The goal is to encourage researchers in the CE program to make use of the audit trail data to answer questions related to their branch
- Wrote a script that converts raw audit trail files into five rectangular datasets for easier analysis


Five Linked Tables

Case Table: Consumer unit for one month

Form Table: Instance where the instrument was opened and closed for a case

Field Table: Interaction with a data field in the instrument

Action Table: Audit trail line containing the “Action” keyword

Error Table: Triggered error and its resolution
Making Audit Trails Accessible

### Form Table

<table>
<thead>
<tr>
<th>fprimary</th>
<th>month</th>
<th>year</th>
<th>order_enter</th>
<th>datetime_enter</th>
<th>order_leave</th>
<th>datetime_leave</th>
</tr>
</thead>
<tbody>
<tr>
<td>p00000102</td>
<td>1</td>
<td>2018</td>
<td>1</td>
<td>1/11/2018 10:06:57 AM</td>
<td>42</td>
<td>1/11/2018 10:09:56 AM</td>
</tr>
<tr>
<td>p00000103</td>
<td>1</td>
<td>2018</td>
<td>1</td>
<td>1/9/2018 06:19:56 PM</td>
<td>2832</td>
<td>1/9/2018 08:45:18 PM</td>
</tr>
</tbody>
</table>

---

00000102.adt

"1/11/2018 10:06:57 AM","Enter Form:1","Key:00000102"
... Stuff happening
... Stuff happening
"1/11/2018 10:09:56 AM","Leave Form:1","Key:00000102"

"1/13/2018 11:19:56 AM","Enter Form:1","Key:00000102"
... Stuff happening
... Stuff happening
"1/13/2018 12:45:18 AM","Leave Form:1","Key:00000102"

00000103.adt

"1/09/2018 06:19:56 PM","Enter Form:1","Key:00000103"
... Stuff happening
... Stuff happening
"1/09/2018 08:45:18 PM","Leave Form:1","Key:00000103"
Making Audit Trails Accessible

Some Quick Stats  (August 2017)

- Raw Audit Trail Files: 3,993
- Recorded Events: 3,756,625
- Mean Events per Case:
  - Complete: 1,662
  - Incomplete: 163
- Time to Parse: 8h 39m 24s

<table>
<thead>
<tr>
<th>Table Type</th>
<th>Number of Rows</th>
<th>File Size</th>
</tr>
</thead>
<tbody>
<tr>
<td>Case Table</td>
<td>3,993</td>
<td>0.7 MB</td>
</tr>
<tr>
<td>Form Table</td>
<td>16,950</td>
<td>1.4 MB</td>
</tr>
<tr>
<td>Field Table</td>
<td>1,247,006</td>
<td>181.6 MB</td>
</tr>
<tr>
<td>Action Table</td>
<td>853,618</td>
<td>73.2 MB</td>
</tr>
<tr>
<td>Error Table</td>
<td>11,652</td>
<td>2.1 MB</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td><strong>2,133,219</strong></td>
<td><strong>259.0 MB</strong></td>
</tr>
</tbody>
</table>
Making Audit Trails Accessible

- In addition to the parsing script, we have produced documentation for the newly created data tables.
- Five tables will be produced for each month going forward.
- Past months/years could be converted upon request or when time permits (more than 10 years’ worth).
- Tables will be stored as .csv files on secured share drive.
- Establish code repository:
  - Reading in files
  - Simple calculations (e.g., section/question timing)
Keeping Audit Trails Going

- We don’t want this project to suffer the same fate as Chan et al.
- We are building this pipeline for parsing audit trail files for CE Staff
- We are committed to knocking down barriers to CE Staff working with the parsed audit trails tables
  - Education/orientation
  - Providing sample analysis code
  - Generating research ideas
Making Audit Trails Useful
(Some Examples)
Field Table

Time Spent Reading Instructions
August 2017

- CC.Cu_intro
- Ce_intro.Ce_intro
- Front_genIntro
- Sect03.33_INTRO
- Sect04.34A_INTRO
- Sect04.34C_INTRO
- Sect05.55_INTRO
- Sect06.IntroA
- Sect06.IntroB
- Sect07.57_INTRO
- Sect08.LIs11.IntroA
- Sect09.IntroA
- Sect09.IntroB
- Sect10.IntroA
- Sect11.IntroA
- Sect12.IntroA
- Sect12.IntroB
- Sect14.S14AINTRO
- Sect14.S14BINTRO
- Sect15.S15A1_INTRO
- Sect16.EDA_INTRO
- Sect17.SUB_INTRO
- Sect19.S19A_INTRO
- Sect19.S19B_INTRO
- Sect22.Sect22.L1INTRO

• Do not read as worded below
  • Identify yourself - show I.D.
  • Ask for eligible respondent
  • Introduce survey - adjust introduction to last interview’s status and respondent
  • The Household address is:
    5034  Blue Ct

• Read if necessary
I am (your name) from the US Census Bureau. Here is my identification card. We are conducting a Consumer Expenditure survey for the Bureau of Labor Statistics. I have some questions I would like to ask you.

Did you receive our letter?

1. Yes
2. No
3. Non-interview

Weekly expense for other tobaccos

Average Time (in s)
Action Table

<table>
<thead>
<tr>
<th>Question</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sect20.GROCWX</td>
<td>12</td>
</tr>
<tr>
<td>Sect02.Rent[1].RENTX1</td>
<td>8</td>
</tr>
<tr>
<td>Sect12.vopgasx</td>
<td>7</td>
</tr>
<tr>
<td>Sect20.DINE_WKX</td>
<td>7</td>
</tr>
<tr>
<td>Sect20.GROCWX</td>
<td>22</td>
</tr>
<tr>
<td>Sect21.PartB.RETSURVX</td>
<td>16</td>
</tr>
<tr>
<td>Sect20.SECT20B.SALONX</td>
<td>14</td>
</tr>
</tbody>
</table>

Ford-Related Search Terms

<table>
<thead>
<tr>
<th>Term</th>
<th>Ford-Related Search Terms</th>
</tr>
</thead>
<tbody>
<tr>
<td>fo</td>
<td>ford f1502</td>
</tr>
<tr>
<td>for</td>
<td>ford focus</td>
</tr>
<tr>
<td>ford</td>
<td>ford free</td>
</tr>
<tr>
<td>ford 150</td>
<td>ford freestar</td>
</tr>
<tr>
<td>ford 500</td>
<td>ford fs</td>
</tr>
<tr>
<td>ford edg</td>
<td>ford fusio</td>
</tr>
<tr>
<td>ford escape</td>
<td>ford mus</td>
</tr>
<tr>
<td>ford ex</td>
<td>ford pick</td>
</tr>
<tr>
<td>ford excurs</td>
<td>ford ran</td>
</tr>
<tr>
<td>ford exp</td>
<td>ford suber</td>
</tr>
<tr>
<td>ford explor</td>
<td>ford t</td>
</tr>
<tr>
<td>ford f</td>
<td>ford ta</td>
</tr>
<tr>
<td>ford f-250</td>
<td>ford tau</td>
</tr>
<tr>
<td>ford f 250</td>
<td>ford u</td>
</tr>
</tbody>
</table>
### Error Table

<table>
<thead>
<tr>
<th>Question</th>
<th>Error Text</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sect20.DINE_WKX</td>
<td>THE VALUE ENTERED IS UNUSUALLY HIGH OR LOW PLEASE VERIFY</td>
<td>260</td>
</tr>
<tr>
<td>Sect09.TableA.Row[1].Amount</td>
<td>THE VALUE ENTERED IS UNUSUALLY HIGH OR LOW PLEASE VERIFY</td>
<td>254</td>
</tr>
<tr>
<td>Sect12.TableA.Row[1].Amount</td>
<td>THE VALUE ENTERED IS UNUSUALLY HIGH OR LOW PLEASE VERIFY</td>
<td>247</td>
</tr>
<tr>
<td>Sect20.SECT20B.SALONX</td>
<td>THE VALUE ENTERED IS UNUSUALLY HIGH OR LOW PLEASE VERIFY</td>
<td>162</td>
</tr>
<tr>
<td>Sect20.OTHSTUFX</td>
<td>THE VALUE ENTERED IS UNUSUALLY HIGH OR LOW PLEASE VERIFY</td>
<td>157</td>
</tr>
</tbody>
</table>

### Questions That Trigger the Most Errors

- **Sect20.DINE_WKX**: THE VALUE ENTERED IS UNUSUALLY HIGH OR LOW PLEASE VERIFY (260 errors)
- **Sect09.TableA.Row[1].Amount**: THE VALUE ENTERED IS UNUSUALLY HIGH OR LOW PLEASE VERIFY (254 errors)
- **Sect12.TableA.Row[1].Amount**: THE VALUE ENTERED IS UNUSUALLY HIGH OR LOW PLEASE VERIFY (247 errors)
- **Sect20.SECT20B.SALONX**: THE VALUE ENTERED IS UNUSUALLY HIGH OR LOW PLEASE VERIFY (162 errors)
- **Sect20.OTHSTUFX**: THE VALUE ENTERED IS UNUSUALLY HIGH OR LOW PLEASE VERIFY (157 errors)

### Range Errors for DINE_WKX

- **August 2017**
  - > $120: 150 errors
  - < $10: 100 errors
Contact Information

Brandon Kopp
Research Psychologist
Office of Survey Methods Research
www.bls.gov/osmr
202-691-7514
kopp.brandon@bls.gov