UNIVERSITY OF MICHIGAN

Adding Business Intelligence to Paradata: The Blaise Audit Trail

Joel Devonshire
Gina-Qian Cheung

2013 International Blaise Users Conference
Washington, D.C.

Survey Research Operations ▪ Survey Research Center ▪ Institute for Social Research

First, I have some questions…..

- How does total interview length look across the sample and by key grouping variables?
- What questions in my data model may be problematic?
- Are there particular interviewers having problems or potentially falsifying data?
- Are the data models rules working as expected?
- Where do interviewers tend to suspend the instrument, enter remarks, or access help?
- Did a new version of our data model have any discernible effects on iwer behavior or data entry?
- How many days does it take to fully complete an iw?

Now, a bit of background…

- Simply putting audit trail data into a database doesn't facilitate exploratory analysis
- Queries are slow and data is "disconnected"
- We needed to design a tool that made ADT data as accessible as possible
- We chose the OLAP cube as a means of transforming the "raw" ADT data into something more meaningful
  - Create a multidimensional space where data is pre-aggregated and joined with other data

DimProject
PK_Project

DimField
PK_Field
Field_Name

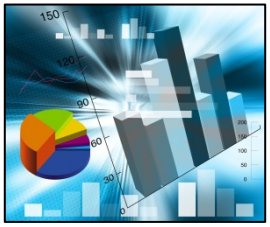
FldName	BlockName	EnterDate	EnterTime	LeaveDate	LeaveTime	Fld_SS
Interview.A006_	SecA	07/19/2012	11:04:22AM	07/19/2012	11:04:24AM	2.109
Interview.A007TRAlive_A	SecA	07/19/2012	11:04:24AM	07/19/2012	11:04:25AM	0.777
Interview.A002_IwBegin	SecA	07/19/2012	11:04:25AM	07/19/2012	11:04:26AM	0.620
Interview.A155_SelfPrxy	SecA	07/19/2012	11:04:26AM	07/19/2012	11:04:27AM	0.956
Interview.A012_LangSwitch	SecA	07/19/2012	11:04:27AM	07/19/2012	11:04:28AM	0.821
inuInterview.A165_A013_	SecA	07/19/2012	11:04:28AM	07/19/2012	11:05:38AM	70.432
inuInterview.A013_Continue	SecA	07/19/2012	11:05:39AM	07/19/2012	11:05:43AM	4.614
tions.A166_A020TSameSpP_A	SecA	07/19/2012	11:05:43AM	07/19/2012	11:05:51AM	8.191
tions.A167_A028_RInNHome	SecA	07/19/2012	11:05:51AM	07/19/2012	11:05:56AM	4.471
tions.A030_LivTogethr	SecA	07/19/2012	11:05:56AM	07/19/2012	11:06:00AM	3.718
ster.RGRID[1].X058AFName	SecA	07/19/2012	11:06:00AM	07/19/2012	11:06:08AM	8.194
ster.RGRID[1].X017ARLName	SecA	07/19/2012	11:06:08AM	07/19/2012	11:06:10AM	2.476
ster.RGRID[1].X060ASex	SecA	07/19/2012	11:06:10AM	07/19/2012	11:06:13AM	2.422
ster.RGRID[2].X058AFName	SecA	07/19/2012	11:06:13AM	07/19/2012	11:06:19AM	6.156
ster.RGRID[2].X017ARLName	SecA	07/19/2012	11:06:19AM	07/19/2012	11:06:19AM	0.443
ster.RGRID[2].X060ASex	SecA	07/19/2012	11:06:20AM	07/19/2012	11:06:25AM	5.387
_	SecA	07/19/2012	11:06:25AM	07/19/2012	11:06:27AM	1.514
ncial.A053_	SecA	07/19/2012	11:06:27AM	07/19/2012	11:06:28AM	1.786
singLocation.A071_SameMainRes	SecA	07/19/2012	11:06:28AM	07/19/2012	11:06:45AM	16.467
singLocation.A079_	SecA	07/19/2012	11:06:45AM	07/19/2012	11:06:51AM	5.560
singLocation.A110TStophouseAssignments	SecA	07/19/2012	11:06:51AM	07/19/2012	11:06:52AM	1.343
_A2_Info1.A087TIntroChild_A	SecA2	07/19/2012	11:06:52AM	07/19/2012	11:07:11AM	18.526
dTab.Child[1].X058AFName	SecA2	07/19/2012	11:07:11AM	07/19/2012	11:07:14AM	3.634
dTab.Child[1].X061ARelateR	SecA2	07/19/2012	11:07:14AM	07/19/2012	11:07:23AM	8.404
dTab.Child[1].X063ARelateSp	SecA2	07/19/2012	11:07:23AM	07/19/2012	11:07:31AM	8.440
dTab.Child[1].X060ASex	SecA2	07/19/2012	11:07:31AM	07/19/2012	11:07:33AM	2.173
dTab.Child[1].X056AResStat	SecA2	07/19/2012	11:07:33AM	07/19/2012	11:07:37AM	3.115
dTab.Child[1].X095ASameSPP	SecA2	07/19/2012	11:07:37AM	07/19/2012	11:07:40AM	3.697
dTab.Child[1].X081AComment	SecA2	07/19/2012	11:07:40AM	07/19/2012	11:07:41AM	0.494
dTab.Child[2].X056AResStat	SecA2	07/19/2012	11:07:41AM	07/19/2012	11:07:44AM	3.007
dTab.Child[2].X081AComment	SecA2	07/19/2012	11:07:44AM	07/19/2012	11:07:44AM	0.497
dTab.Child[3].X058AFName	SecA2	07/19/2012	11:07:45AM	07/19/2012	11:07:45AM	0.458
dTab.Child[3].X061ARelateR	SecA2	07/19/2012	11:07:45AM	07/19/2012	11:07:49AM	3.785
dTab.Child[3].X063ARelateSp	SecA2	07/19/2012	11:07:49AM	07/19/2012	11:07:53AM	4.369
dTab.Child[3].X060ASex	SecA2	07/19/2012	11:07:54AM	07/19/2012	11:07:55AM	1.617
dTab.Child[3].X056AResStat	SecA2	07/19/2012	11:07:55AM	07/19/2012	11:07:58AM	3.044
dTab.Child[3].X096AMarried	SecA2	07/19/2012	11:07:58AM	07/19/2012	11:08:02AM	3.911
dTab.Child[3].X097APartnered	SecA2	07/19/2012	11:08:02AM	07/19/2012	11:08:07AM	4.656
dTab.Child[3].X081AComment	SecA2	07/19/2012	11:08:07AM	07/19/2012	11:08:07AM	0.388
dTab.Child[4].X058AFName	SecA2	07/19/2012	11:08:07AM	07/19/2012	11:08:08AM	0.782
dTab.Child[4].X061ARelateR	SecA2	07/19/2012	11:08:08AM	07/19/2012	11:08:13AM	4.762

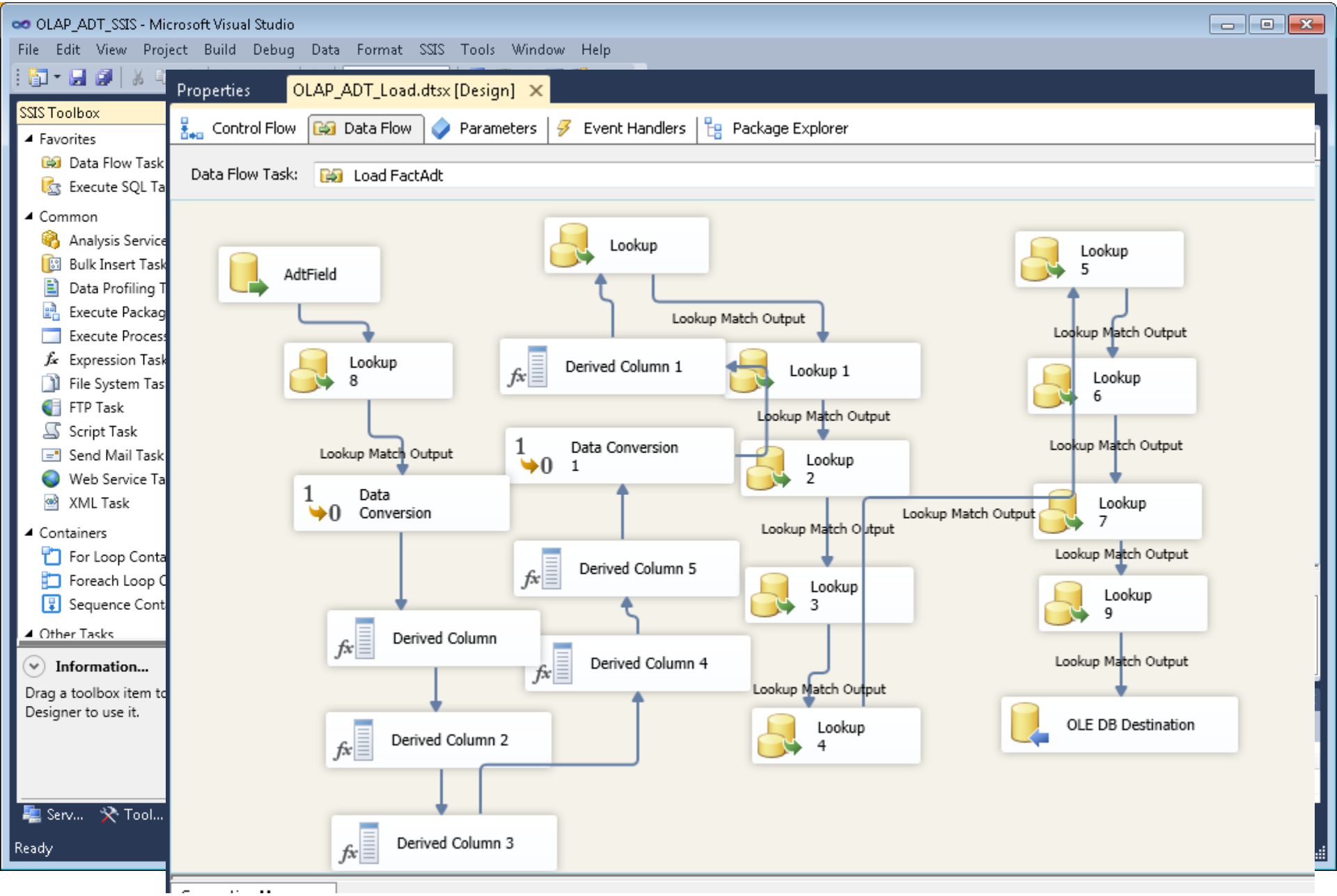
DimBlock
PK_Block
Block_Name

DimSample_Line
PK_Sample_Line
Sample_Line_Name
Result_Code
Release
Instrument_Type
Sample_Type
Resistance_Flag
Language
Mode
Cohort
DRI_Recorded
DRI_Consent
Day_Of_Week
Month
Year
Evaluation_Score
Interviewer_ID
Interviewer_Name
PM_ID
PM_Name
PC_ID
PC_Name
TL_ID
TL_Name
Iwer_Gender
Iwer_Experience
Iwer_Timezone
Iwer_Field_SSL
Project_Name

T OLAP

is/
ing




OLAP_ADT_SSIS - Microsoft Visual Studio
File Edit View Project Build Debug Data Format SSIS Tools Window Help
Properties
OLAP_ADT_Load.dtsx [Design]
SSIS Toolbox
Favorites
Data Flow Task
Execute SQL Ta
Common
Analysis Service
Bulk Insert Task
Data Profiling T
Execute Packag
Execute Process
Expression Task
File System Tas
FTP Task
Script Task
Send Mail Task
Web Service Ta
XML Task
Containers
For Loop Conta
Foreach Loop C
Sequence Cont
Other Tasks
Information...
Drag a toolbox item to
Designer to use it.
Serv...
Tool...
Ready
Control Flow
Data Flow
Parameters
Event Handlers
Package Explorer
Data Flow Task:
Load FactAdt
AdtField
Lookup
8
Lookup
Derived Column 1
Lookup 1
Lookup
5
Lookup
6
Lookup
2
Lookup
7
Lookup
3
Lookup
9
Lookup
4
Data Conversion
1
Data
Conversion
Derived Column 5
Derived Column
Derived Column 4
Derived Column 2
Derived Column 3
OLE DB Destination
Lookup Match Output

DEMO

OLAP Cube Advantages

- Extremely flexible and open; invites data exploration and new query ideas
- Pulls together data from many sources
- Very customizable, with ability to define limited views and data of interest
- Accessible with familiar tools
- Possible to use as basis of “dashboard” with key metrics and reports

Challenges/Constraints

- After initial design, changes are relatively easy, but still under centralized control
- Can be deceptively "easy" to use
  - Calculations and conclusions still need to be carefully thought through
  - Building an OLAP cube can be "overkill"
- Harmonization among studies

Thank You!

jdev@umich.edu

qianyang@umich.edu