Adding Business Intelligence to Paradata: The Blaise Audit Trail

Joel Devonshire
Gina-Qian Cheung

2013 International Blaise Users Conference
Washington, D.C.
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
Creating a cube…

Raw ADT

Data Warehouse ADT OLAP

ETL

Other Data Sources

Analysis/Reporting

2013 International Blaise Users Conference, Wash D.C.
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