e-Books

e-Stat All-hands (Bristol)
27th September 2010

e-Books - Proposal

- ".. Should have the expository and narrative advantages of a technical book and the interactive learning advantages of software."
- ".. Web-based documents which summarise the multiple processes within quantitative data analysis.."

e-Books - Proposal

- “.. document the research process where the scientific objects in the book, for example data, tables, graphs, diagrams and equations are live.. [meaning] that the code to produce and manipulate the objects is included in the book and can be altered and executed.”

Generic use

- Evolution into a paper
- Research tool
  - Electronic Lab notebooks
- Research objects

(Some) Requirements

- Styles of interaction
  - Tutorial
  - Exploratory
- Environment to do analysis
  - Worksheets
  - Workspaces, Logbook

(Some) Requirements (cont'd)

- Document with static and dynamic content
  - Computed
  - Interactive widgets/visualisations
- Automatically recalculates
  - Change data
  - Change inputs/selections/parameters
- Assemble all things together in one package

(Some) Requirements (cont'd)

- Easy to author
  - Writing text
  - Assembling components
  - Testing
- Rewrite an e-Book for a different domain
- Wrap up into a pre-calculated e-Book
- Be useful to Readers
  - Cut and paste into their papers

Existing Systems

- Mathematica “notebooks”
- Sage
- Sweave
- Microsoft Word + COM
- Matlab/R/SPSS?/Stata?/MLwiN?

Anatomy of an e-Book

- Static content
- Programs and scripts
  Templates
- Supporting objects
  Datasets
- Intermediate results and provenance
  Chains
- Outputs

Anatomy of an e-Stat e-Book

- Static content
- Programs and scripts
  - Templates
- Supporting objects
  - Datasets
- Intermediate results and provenance
  - Chains
- Outputs
  - Parameter Estimates

e-Book Prototype

- Wrap one or more templates into a document with static content
- Separate the display from the execution
- Understand the environment that templates need to run in
- Feasibility of using HTML / web browser

Why HTML

- Everyone has a browser
  - (and knows how to use it)
- A lot of recent developments to browsers
  - Faster (especially script performance)
  - Javascript libraries
  - Rich applications (using HTML5)
- Standards support for eg
  - SVG (diagrams, graphs), MathML (math)
- Easy to write
  - Tooling exists

Demo

Future

- Write some examples
  - Author feedback
  - Reader feedback
- Package format
- Authoring environment
- Execution environment
- Architectural specification
- Other requirements?
- Integration with other systems?
- Integration with myExperiment