Data Management – Audit Trail

Claire Osgood
November 2017
Audit Trail
Audit Trail

A collection of source files, intermediary files, and documentation that comprise a chronological record of the steps taken to create a published dataset. When followed, a clean audit trail allows every intermediary and final dataset to be replicated exactly.
Audit Trail

Source Data → Program/GIS Process → Intermediary Data → Program/GIS Process → Published Data

Output

Time/Date Stamp Chronological

Documentation
Audit Trail – Freeze Final Files

I am Read Only
Audit Trail – Program/GIS Process – Overwrite v. New

Re-run program, overwrite output:
- Revisions
- Corrections
- Updates
- Changes that bring us “closer to the truth”
- Similar to editing papers

ONY IF not finalized yet

New program, new output:
- Change in methodology
- Variations of methodology
- Alternate analyses
- Similar to editing papers

OR if already finalized

Same rules for GIS processes, processing and analysis programs.
Audit Trail – Overwrite v. New

Dataset

Program/ GIS Process
Audit Trail – Manual Data Entry and Verification

Source (Raw) → Manual → New File
Audit Trail – SAS programs

Specific Notes for SAS Programmers:
• PROC CONTENTS for every input file
• PROC CONTENTS for every output file
• Limited print of text, Excel, or other non-SAS input files
• SOURCE2 option if include code is used

Always Batch Submit for final run
Audit Trail – Analyses for Papers

Audit Trail: <Project Name>

Project description/goals:

Statistician: | GIS analyst:

Funding sources:

Data:

Sources:

Processing notes:

Key analytical decisions:

Resource: 

Status as of ________:

Submission history: