

Using EG for Large and Complex Analytics

Can they really get out of control?

Paul Cripwell
NAV CANADA

Overview



- Context
- Creating behemoth projects
- Arranging the flow
- Processes vs Projects
- Sharing interim datasets
- Quick renames on source data



Context

- NAV CANADA air traffic analysis
- Extract working data from live systems
- “Freeze” data in monthly chunks
- Raw data to summary stats
- Report on a monthly basis
- Excel is preferred reporting method



Behemoth Projects

- How big is TOO BIG?
- Common to have 100 – 200 tasks
- Creation is one small step at a time
- EG keeps things straight for you
- Using PROC SQL instead of DATA steps



Arranging the Flow

- It makes sense
- Blocks of “work”
- Dead ends for tests
- Outputs collected in a common area



Processes and Projects

- A Project can consist of a number of Processes
- Why wide projects vs. multi-process
- Each on its own tab
- Watch for Run Project or Run Process
- Links between Processes



Sharing Interim Datasets

- Copy and re-use projects
- Different projects will use same interim files created in working directory
- EGTask folder for ease of control
- Can cause issues
- Be aware!



Quick Renames

- Several ways to do it
- Right click on source, Properties, then change files
- Within Query task, right click on source file on left panel, Properties, then select new file
- Within Join tab in Query task



Summary

- Let yourself go!
- SAS EG will look after you
- Build for future production efficiency



Paul Cripwell

Manager, Traffic Analysis and
Forecasting

Revenue and Pension Administration

Finance Department

NAV CANADA

Paul.cripwell@navcanada.ca

613-563-5876

