

21st GSS Methodology Symposium

Data Warehousing

Developing an Integrated Business Survey System for Northern Ireland

Dr James Gillan
Northern Ireland Statistics and Research Agency
(NISRA)



The problem.....

- How to be agile enough to keep pace with ONS
- NISRA business surveys pre – 2011.....
 - Mixture of in-house and outsourced data collection
 - Almost exclusively paper based
 - 100% forms manually keyed and manually validated
 - Little or No cross survey coordination
 - Collection (contacts with business)
 - Use of data (Data sharing)
 - An array of stand alone databases using different software packages (collection & results production)
 - New surveys needed new databases



The requirements.....

- Easy development of new and existing surveys
- Reduce duplication and improve congruence
- Maximise use of administrative data sources
- Automated validation with better prioritisation
- Co-ordinate / integrate contact with businesses
- Manage survey progress and staff workloads
- Support flexible methods of data collection
- Reduce average time to capture business data
- Suitable for a small statistical office

The Solution

Integrated Business Survey System (IBSS)

April 2011 – build commences for new integrated business survey IT system for 11 business surveys.

Uses **commercial case management software**

An **external contractor** to build and provide a managed service

Phase 1:

- Print & Distribution
- Collection
- Edit & Validation

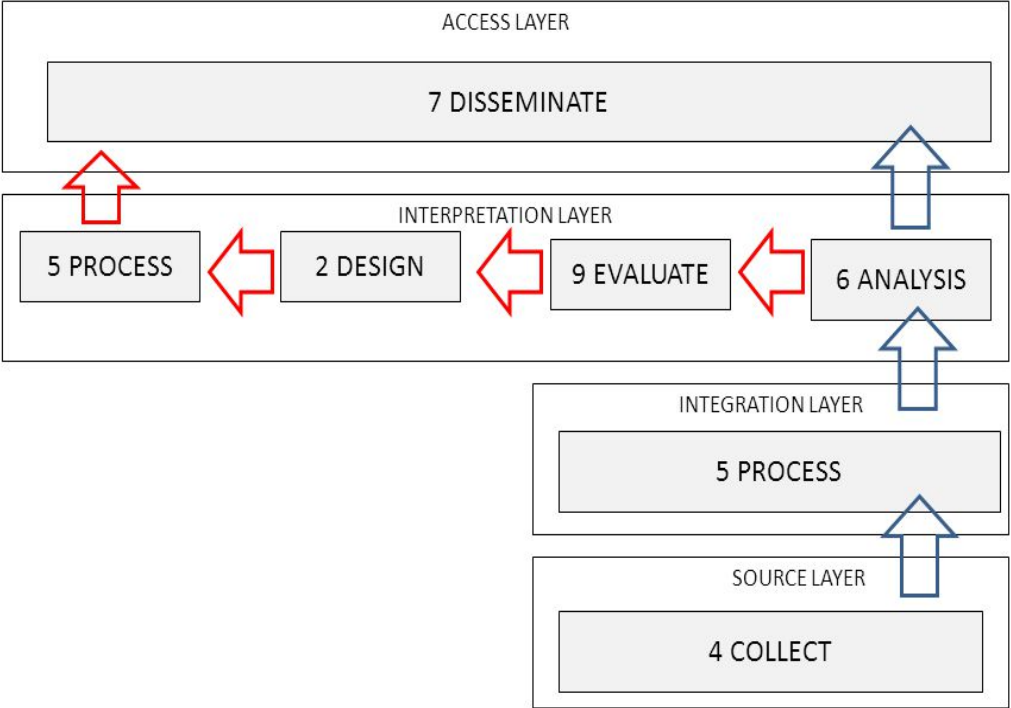
Phase 2:

- **Electronic Data Collection**
- **Integration of Statistical Processing**

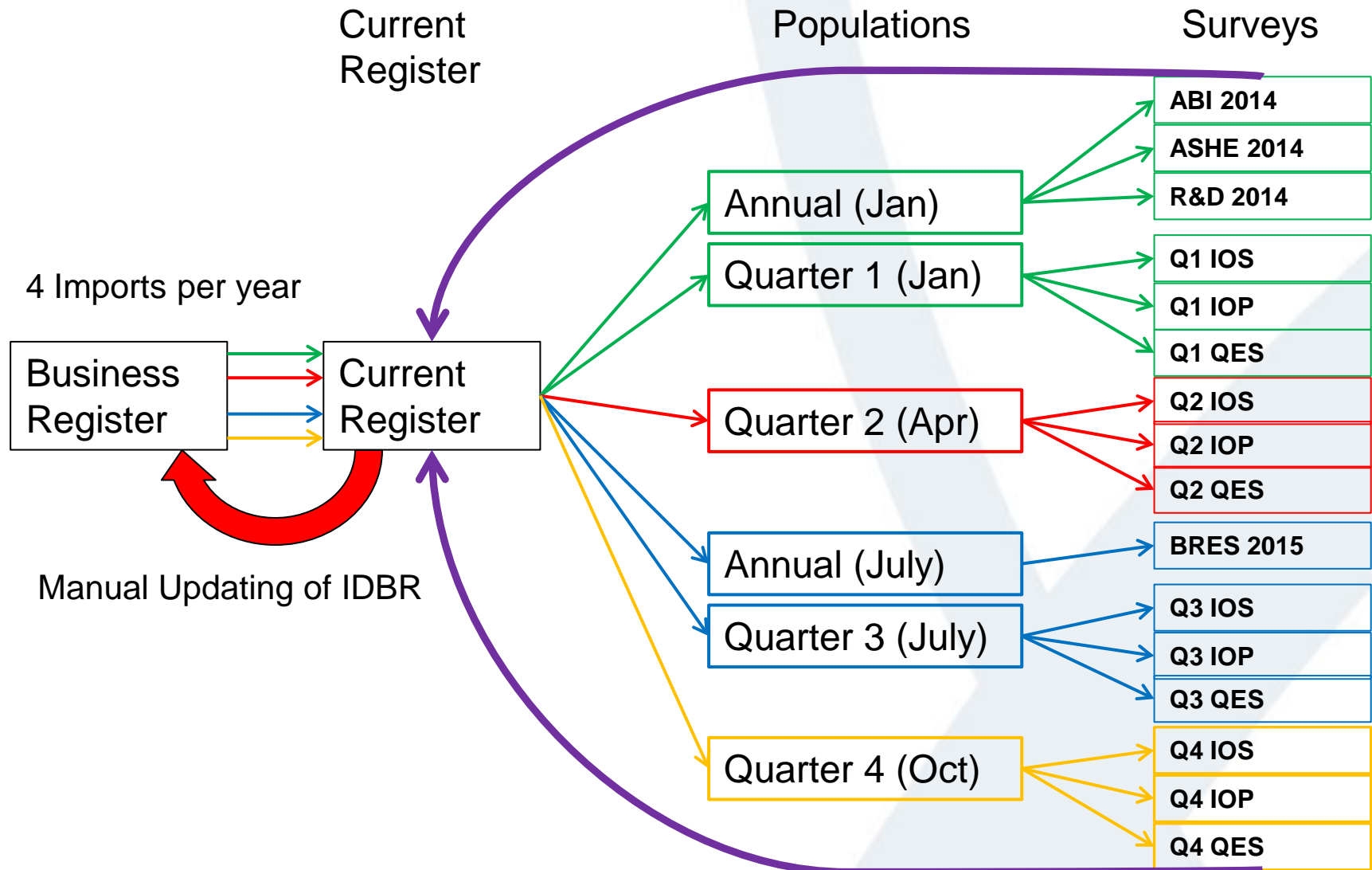
The GSBPM and Data Warehouse layers

Statistical Data Warehouse layered architecture

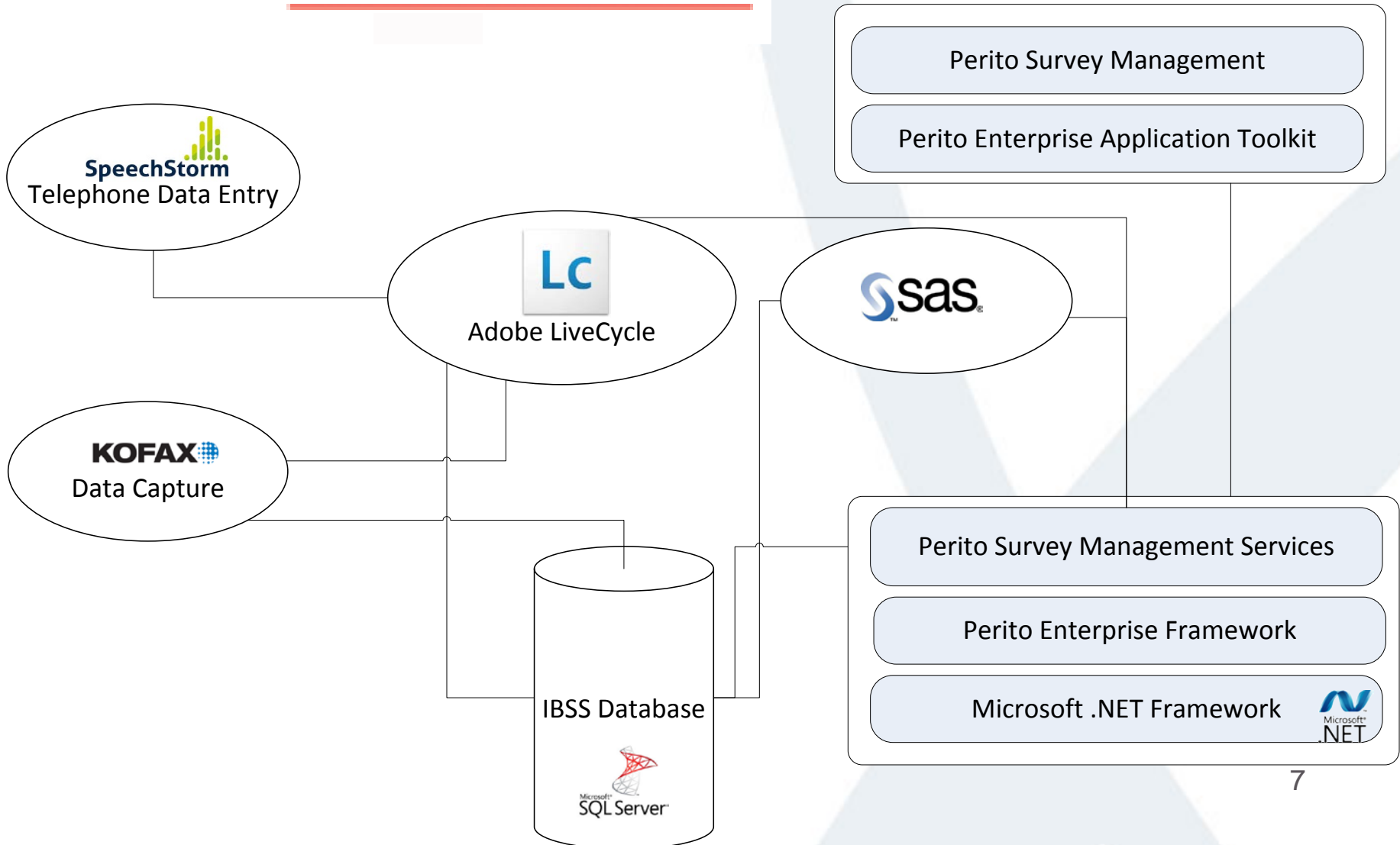
Case: re-use data to create new data



How information flows between Surveys and Register



Software design





Collection: IBSS Data capture methods

- Paper questionnaires
 - C.100,000 issued per year for 11 surveys, but scalable
 - Scanning with Optical character recognition
- Telephone Data Entry (TDE)
 - Option available on short questionnaires.
 - C. 25% of all short term survey responses.
- **Electronic Data Collection (EDC)**
 - e-Forms directly emailed to businesses from IBSS
 - Responses are encrypted with direct uploaded to application
 - Using a different EDC approach to many NSI's
 - Roll out plan across a range of surveys (uptake c.35%)

Collection: OCR image available

Navigator

- Survey Data
- Validation Check
- Details
- Communication Log
- Documents
- Tasks
- Contact Information
- Alerts
- Variable Ratios

Summary

Survey Instance Reference:

BRES/2014

RU Reference:

49905135186

RU Name:

[Redacted]

Trading as:

Form type:

BRES Form Type 1

Strata:

Unknown

Data source:

Posted Form

Contact Name:

LOIS

Telephone:

[Redacted]

Mobile:

Email:

[Redacted]

Maintain Survey Response Documents

View Reprint Regenerate Amend Contact Details

Letter Type	Direction	Letter Format	Indicator	Date
reminder 1	Incoming	Full Form	Scanned Document	27/10/2014 15:31
reminder 1	Outgoing	Full Form	Reprinted Document	13/10/2014 09:54
reminder 1	Outgoing	Full Form	Document From Generate Print Run	03/10/2014 11:57
initial form	Outgoing	Full Form	Document From Generate Print Run	14/08/2014 10:52

00000658162.pdf - Adobe Reader

File Edit View Document Tools Window Help

2 / 9 66.7% Find

5. How many NI locations/workplaces, in total, did this business operate from on: 1 September 2014 ?

6. How many working owners are there in total in this business, as at 1 September 2014, who receive drawings or a share of profits but are NOT PAID via PAYE? Working owners include sole traders/proprietors, partners and directors (not applicable to limited companies, charities, churches, voluntary or public bodies)

Guidance Notes for completion of following employee questions

An employee is anyone aged 16 years or over that your organisation directly pays from its payroll(s), in return for carrying out a full-time or part-time job or being on a training scheme in Northern Ireland.

Please include: Please exclude:

1. All workers paid directly from this business's payroll(s), who are based in Northern Ireland.	1. All employees based outside Northern Ireland.
2. Those temporarily absent but still being paid, e.g. maternity leave.	2. Agency workers paid directly from the agency payroll.
3. Seconded / Agency staff paid directly by you.	3. Voluntary Workers.
	4. Former Employees only receiving a pension.
	5. Self-employed workers.
	6. Working owners who are not paid via PAYE.

401 00000658162 001 002 1 of 8

Details

Letter type:	reminder 1	Form type:	BRES Form Type 1
Batch no:	KOFAX	Direction:	Incoming
Date Received:	27/10/2014 15:31	Indicator:	Scanned Document
Is archived:	<input type="checkbox"/>	Format:	Full Form

Collection: Central Communications log

Navigator

- Survey Data
- Validation Check
- Details
- Communication Log
- Documents
- Tasks
- Contact Information
- Alerts
- Variable Ratios

Summary

Survey Instance Reference:
BRES/2014

RU Reference:
[Redacted]

RU Name:
[Redacted]

Trading as:
[Redacted]

Form type:
BRES Form Type 1

Strata:
Unknown

Data source:
Posted Form

Contact Name:
[Redacted]

Telephone:
[Redacted]

Mobile:
[Redacted]

Maintain Survey Response Communication Log

+ New | Edit | Save | Cancel | Email Contact

All Other Surveys Communication History

Survey Instance Ref.	Contact Type	Contact Outcome	Contact Name	Created On	Created By
ABI/2013	Letter	Other*	Keith Dinsmore	21/08/2014 1...	Ryan Dougal
ABI/2013	Phone	Other*	RODGER DALLAS	21/08/2014 1...	Ryan Dougal
ABI/2013	Phone	TRC - Spoke to Contact - Will Complete Form	RODGER DALLAS	21/08/2014 1...	Ryan Dougal

Log Details

Type: Letter

Contact Name: Keith Dinsmore

Created by: Ryan Dougal

Created on: 21/08/2014 14:31

Direction: Outgoing

Outcome: Other*

Notes: Have updated the new contact info on the system and issued a form for the new accountants to complete.

BRES Communication Log

Contact Type	Contact Outcome	Contact Name	Created On	Created By
Email	Other*	LOIS	17/12/2014 11:58	Liane McCreanor
Email	Other*	LOIS	17/12/2014 09:24	Liane McCreanor
Phone	Other*	FEMALE COLLEAGUE	28/11/2014 14:38	Liane McCreanor
Email	Other*	LOIS	21/11/2014 08:01	Liane McCreanor
Phone	Other*	LOIS	13/10/2014 09:55	Liane McCreanor



Validation – Savings from Automation

Table : Validation Failure rates by Survey

	QES	IOS	IOP	ABI
Cleared with failures manually accepted or edited	1198	1980	620	5665
Cleared With No Failures	2814	1568	442	191
Cleared with failures automatically accepted	974	0	0	172
Total Responses	4986	3548	1062	6028
Form cleared With No Failures	56%	44%	42%	3%
Validation Task issued	44%	56%	58%	97%

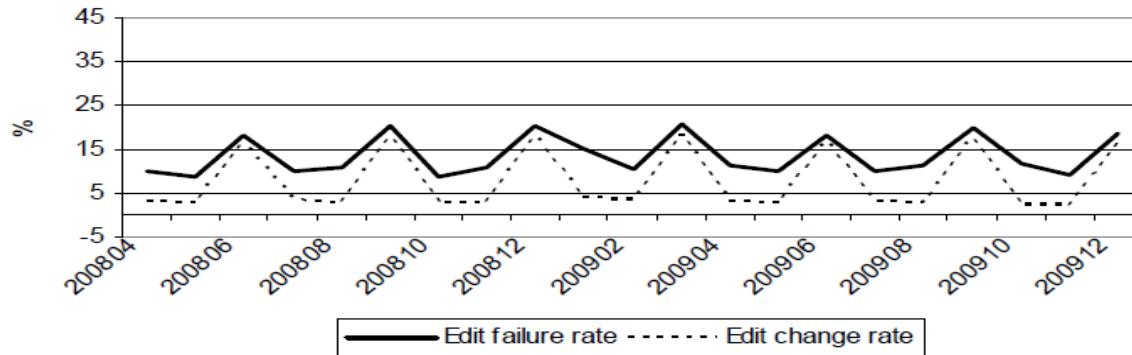
Approximately **ONE THIRD** reduction in fieldwork/validation staff

Validation: Failure and Change rates*

Graph 3. Edit failure rates and edit change rates after using edit rules to identify units to edit from the old MPI domains



Graph 4. Edit failure rates and edit change rates after using selective editing to identify units to edit from the old MPI domains

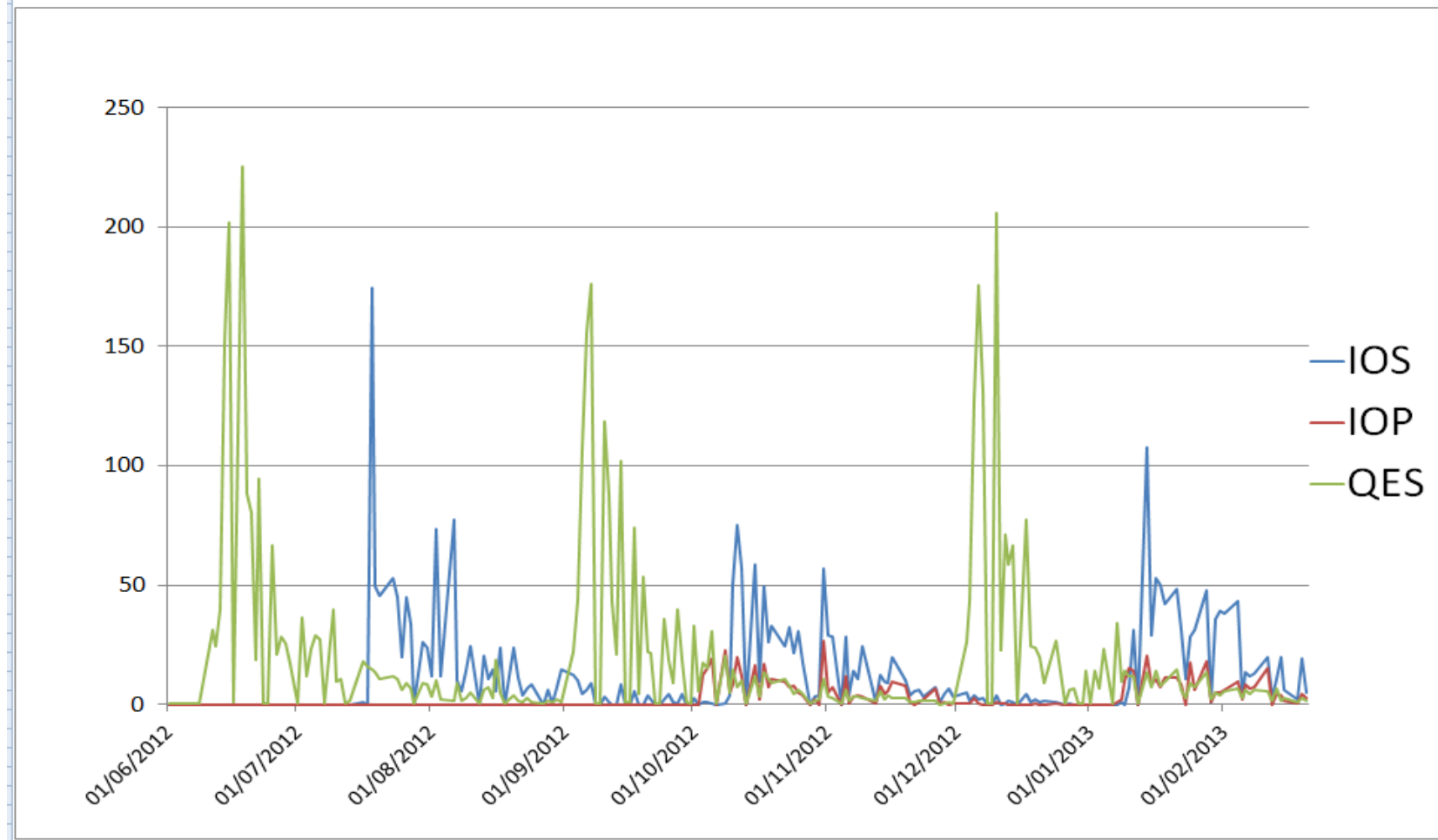


* Example from ONS Monthly Production Inquiry



Workflow: Restructuring Staffing Resource

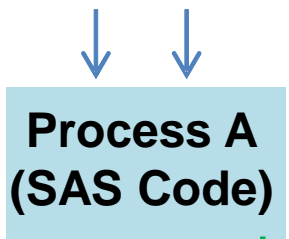
Chart: Daily Validation failure task generation by survey





Aggregation: Example Process Chain with 4 Processes

Validated Micro data



2 Inputs, 3 Outputs

Imported Input



Diagnostic Outputs



3 Inputs, 2 Outputs



2 Inputs, 3 Outputs

Imported Input



Diagnostic Outputs



3 inputs, 3 Outputs



Aggregated Macro Data



Aggregation – Video demo

- See short video demo of a process chain being executed within the IBSS.



BRES 3 Process Execution Chain.mp4



IBSS Contribution to Quality - Statistical Processes

Sound Methods & Procedures

- Stricter adherence to GSBPM
- Automated micro-data validation
- Automated imputation of non response

Reduced Burden

- Electronic data collection
- Integration of surveys & questionnaires

Cost effectiveness

- Manage survey progress and staff workloads
- Prioritised editing of most important “fails”
- Easy development of new and existing surveys



IBSS Contribution to Quality - Statistical Outputs

Accuracy & Reliability

- Diagnostics integrated into the processing
- Increased sample sizes (e.g. ABI / IOS)
- Full audit trail ensures reproducibility of estimates

Timeliness and Punctuality

- Reduce average time to capture business data
- Efficiency of automated validation and SAS integration
- Modular approach to statistical process chain

Coherence and Comparability

- Micro and Macro level coherence checks

Next Steps

- Due for contract renewal
- Test feasibility of 100% electronic first
- Extend scope to more data sources
 - administrative
 - social surveys
- Better integration of metadata
- Place within the context of a wider data management strategy
 - Relationship to economic accounts
 - Relationship to web dissemination