



**Statistical Training
Framework based on the
GSBPM**

Abstract

The UN's Generic Statistical Business Process Model (GSBPM) defines the set of business processes required to produce Official Statistics. Ireland's Central Statistics Office (CSO) is currently developing a Statistical Training Framework which will align knowledge, skills and expertise for each role within a statistical organization based on the GSBPM.

To achieve this, the CSO established a working group to research and develop a Statistical Training Framework based on the GSBPM. Initially, up to 60 statistical skills were identified which was reduced to 15 high level headings. We are currently developing basic, intermediate and advanced descriptors under each heading. The next step will identify the relevant training interventions under each heading. The end result will be the development of Learning Paths for Official Statisticians. The design and application of this Framework can be utilised in any NSI.

This Framework will measure statistical capability based on the GSBPM, specify essential statistical infrastructure that will support statistical production and deliver an enhanced Statistical skillset. It will ensure Statisticians have the necessary knowledge, skills and expertise to work effectively in an end to end statistical production cycle. Furthermore, the Framework could potentially assist with the development of a mobility policy across NSIs, sharing knowledge, skills and expertise and identifying skills gaps in any Statistical organisation. The above Framework is designed as a fluid model and will be reviewed on a regular basis to reflect the ever changing statistical environment we work in.

Keywords: GSBPM, Statistical Framework, Capability, Knowledge, Skills, Mobility, Learning paths, Skills gaps.

Table of Contents

Summary.....	3
Introduction	4
Section one: GSBPM Model	5
Section two: Design	6
Section three: How the Framework will operate	12
Section four: The Irish Example: Training Interventions	13
Conclusion.....	16
<u>Appendices:</u>	
Appendix 3: Survey Data	17
Appendix 4: Sample of Descriptors Write up	18
Appendix 5: Employee Experience	19

Summary

This paper outlines how the GSBPM was utilised as a basis for building a Statistical Training framework encompassing the related elements of training interventions, skills register, learning paths, career paths, essential statistical infrastructure and capability development.

This framework is designed as a self-assessment tool for skills levels; each skill is linked to the GSBPM model and divided into three main levels basic, intermediate and advanced. This paper discusses the design elements of the framework, the benefits, how it will work (using work flow charts) and the overall impact of a statistical training framework within all NSI's using Ireland as a working example.

Introduction

The purpose of this project is to develop a Statistical Training Framework based on the Generic Statistical Business Process Model (GSBPM). Under EU Council regulation 223/2009, the Director General of NSI's are now responsible for all official statistics returned to Eurostat. It is imperative that NSI's can be assured of the quality and standards of all official statistics returned. The framework outlined in this document is a critical piece of work for NSI's to achieve this. The Central Statistics Office, Ireland commenced work on this model in 2016. We are collaborating with UNHLG-MOS, Task Team on Capabilities development to develop a model for use in the wider Statistical community.

This paper is divided into four main sections.

Section one, will outline the GSBPM model and why it was chosen as a model on which to base a Statistical Training framework.

Section two, will outline the design structure, the steps involved in creating the framework, the benefits of the framework in particular what this framework will mean for NSI and their staff, competency development, skills register, learning/career paths and related role profiles.

Section three will discuss how the framework will operate including some detailed work flows on how an individual and a manager will use the framework, it will also cover the supporting IT infrastructure and knowledge database creation for example *Communities of Practice*.

Finally, **section four** will cover the Irish example of training interventions which support the framework and a discussion on how this will operate.

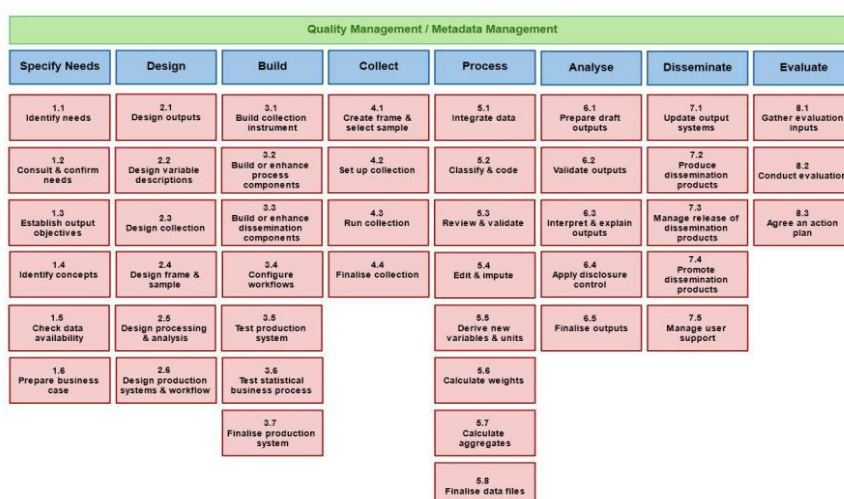
The framework is intended to be compatible in all NSI; it is accompanied by key supporting elements as mentioned above and discussed throughout this paper such as; a detailed policy document, skills register, learning paths/career paths, maturity model, training interventions and knowledge sharing forums. It is designed to be a fluid model adaptable to the ever changing modernisation of statistics including the specific needs of any NSI.

Section one: GSBPM Model

The Generic Statistical Business Process Model (GSBPM) describes and defines the set of business processes needed to produce official statistics. It provides a standard framework and harmonised terminology to help statistical organisations to modernise their statistical production processes, as well as to share methods and components. The GSBPM can also be used for integrating data and metadata standards, as a template for process documentation, for harmonising statistical computing infrastructures, and to provide a framework for process quality assessment and improvement.

As discussed in the introduction the framework is based on the GSBPM model which was selected as a high European standard for processing meta-data and its related components. A recent survey¹ of National Statistical Institutions reported that 60% of NSI's currently use the model, of the 40% who are not using it 75% of those have plans to start implementing its use in the near future which would indicate that over 85% of all NSI's will use the GSBPM model in their organisations. This aided our decision to centre our Statistical training framework on the GSBPM, as we would be basing the framework on something which is respected and implemented in the majority of NSI's providing a standardised model for statistical business processes in an end to end statistical production cycle.

The aforementioned survey also highlighted the importance of the GSBPM to the standardisation of statistical business processes by NSI's. An interesting element of these responses included the GSBPM model aiding in "increasing the understanding of statistical business processes"² and "description and assessment of statistical business processes"³. These key components realised the importance of such a model and how it could be beneficial when adapted to a statistical training framework.



¹ Survey data available at <https://statswiki.unece.org/display/GSBPM/Uses+of+GSBPM>

² Summary survey data page 1 available at <https://statswiki.unece.org/display/GSBPM/Uses+of+GSBPM>

³ See appendix 3.

Section two: Design

Designing the Framework:

In January 2017 the GSBPM model was adapted to develop a statistical training framework in the Central Statistics Office, Ireland. Initial work identified a possible 60 headings across all stages of the GSBPM. In November 2016, a working group of experts was established in order to reduce those 60 headings to 12-15 key headings under which we could categorise certain topics and then build a training framework using an end to end statistical production cycle. This working group was set up as a structured project group with meetings held on a regular basis. The working group comprises of internal subject matter experts and Learning and Development professionals.

In April 2017, it was decided to hold a one day workshop where the working group could brainstorm the final 12-15 headings. This workshop was very successful and resulted in a concise model (see figure 2.1). This model outlines the final headings and assigns at what stage each heading fits into the GSBPM model. This ensures that the framework will be adaptable to any NSI and is designed as a fluid model to allow for modifications/additions in a fast changing statistical environment. With the modernisation in statistics this was an important element to build into the frameworks design.

The UNHLG- MOS has prioritised the importance of this work and established a task team on capabilities to work with CSO, Ireland to work on developing this model. This team led by the CSO, Ireland have reviewed the final high level headings and agreed on the overall 13 headings to be included, which are common to all NSI's and can be used to identify skills gaps across NSI's.

The below work flow (see figure 2.2) depicts the creation of the statistical training framework covering the 5 key elements involved in the creation of the training framework.

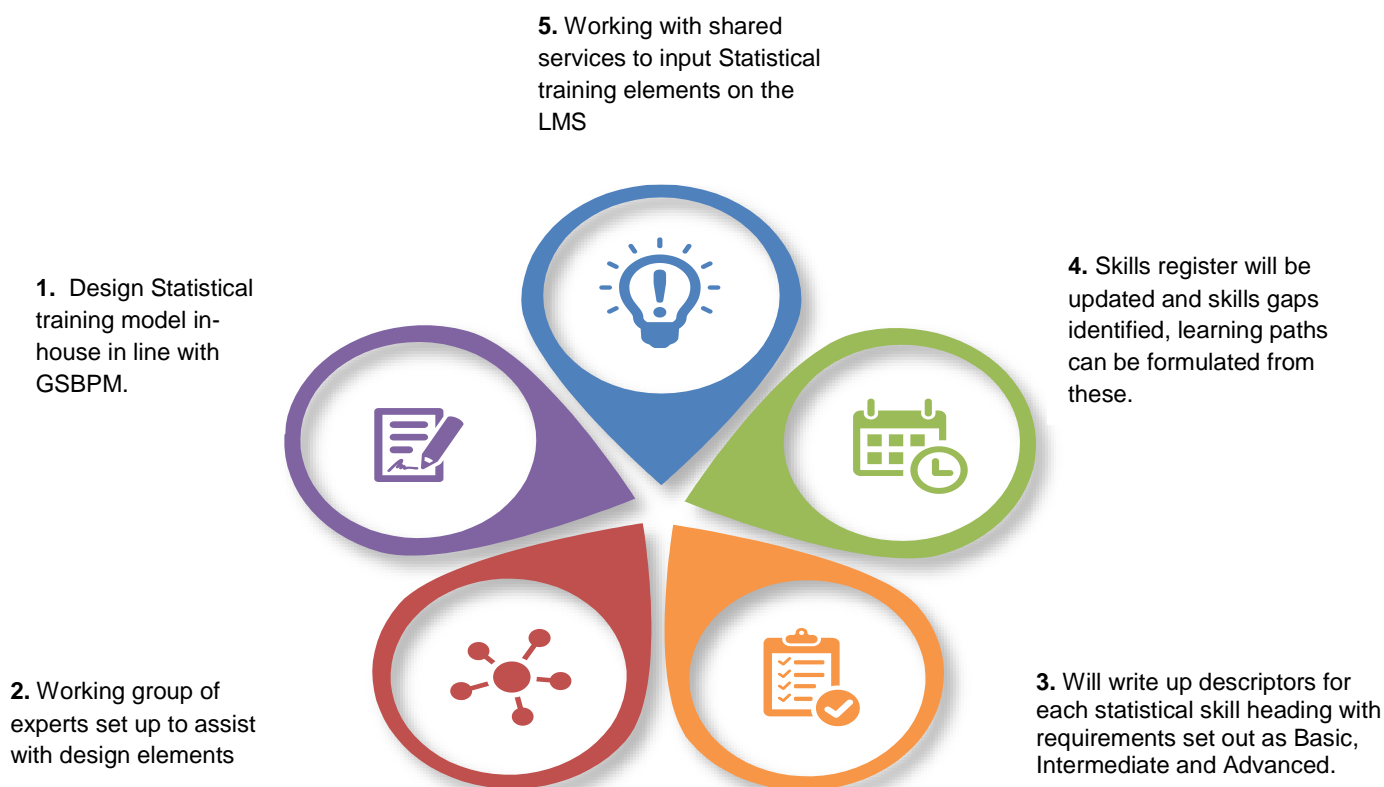
Figure 2.1

Table of 15 headings as aligned with GSBPM

Statistical Framework									
GSBPM								TOPICS	
Specify Needs	Design	Build	Collect	Process	Analyse	Disseminate	Evaluate		
								Data Governance	Ireland
								Fundamental Statistics (Introduction to Statistical Training)	Ireland
								Regulatory Statistical Frameworks (eg IESS, FRIBS)	Ireland
								1. Register Management	No one identified at present
								2. Sample & Estimation	Ireland
								3. Survey & Questionnaire Design	Ireland
								4. Imputation and non-response	Statistical data editing group.
								5. Statistical Data Editing	Statistical data editing group
								6. Evaluating user statistical needs	Discuss at next CapComm meeting
								7. Index Numbers	Ireland
								8. Regression	Ireland
								9. Time Series Analysis & Seasonal Adjustment	Ireland
								10. Statistical Disclosure control	Statistical Disclosure control group will be asked to assist. Webex meeting proposed.
								11. Visualisation and Presentation of Data	Discuss at next CapComm meeting
								12. Data Matching, Integration and Administrative Data	Data Integration Project will be asked to assist.
								13. National Accounts	iStat have agreed to assist.

Figure 2.2

Creation of a statistical training framework (Ireland)



The framework (Figure 2.1) is designed to work in any NSI and can be mapped to relevant training interventions, skills gap analysis, career paths and learning paths. NSI's could use the framework to assess current skill sets in their organisations and ensure that the necessary skills exist in order to produce high quality statistics. This framework also works well with a self-rated skills assessment questionnaire where the employee will self-rate their statistical skills level under the 13 headings, their manager will then sign off on this assessment and suggest required upskilling in areas as required. In order to ensure the framework was a workable document each heading is subsequently divided into three main levels, basic, intermediate and advanced.

These three levels (basic, intermediate and advanced) define the knowledge, skills and expertise required at each level in order to determine if an employee is at a basic level in a topic or intermediate etc. There is a prescriptive measure in which to assess the skills set by reading the detailed descriptors under each heading. An example of this is presented in appendix 4 (Sampling & Estimation).

In order to support this framework there will be a supporting detailed policy document which will be provided to each manager and Statistician working on Official Statistics. This policy document will outline in detail how the framework will operate with detailed flow charts outlining an "employees' experience" as supported by their manager. This will be discussed in further detail under section three of this paper.

Key benefits of the framework:

The table below (Table 2.3) outlines the key benefits of the statistical training framework. This is broken down into four key elements. Organisational refer to the NSI as a whole, employee refers to the benefit to the individual, manager refers to the benefits to work outputs, and cost efficiency refers to strategic budget planning.

Table 2.3 Key benefits of Statistical Training Framework

Organisational	Better quality statistics, increased output, higher retention rates, upskilled workforce, and knowledge databases. Innovative work force, better staff engagement continuous learning culture. Gaps and Strengths identified, Effective Workforce Planning (right people, in the right place at the right time)
Employee	High work performance, self-rated skills assessment, knowledge of key statistical processes, career paths, learning paths, increase in expertise.
Manager	Aids with workforce planning, skills gaps are easily identified, on the job training becomes more transparent and quicker to identify and increased efficiency. Improved performance management.
Cost efficiency	Cost effective measures, in house expertise utilised for training, training interventions booked in advance leading to group rate discounts, strategic planning for training budgets, easy to see where money is spent and return on investment. Gaps identified and more focused training interventions. Workforce Planning, use of current resources (right people, right place, with the right skills).

What this framework will mean for NSI's & their Staff:

A Statistical Training Framework is a necessary resource which will form an integral part of strengthening NSI's overall statistical capability. While no one person is expected to possess all of the statistical skills listed in the framework to perform a particular job, a statistician should gain a broad range of statistical skills over time.

A Statistical Training Framework will ensure all areas within the statistical process will get the required focus and emphasis on training spend. On occasion it is the more public areas that have the 'spend' on the Specify Needs, Survey Design and the end element disseminate and evaluate.

A structured Statistical Training Framework will:

- Guide NSI's statistical capability development initiatives internally, but also for those working in the wider Statistical System.
- Provide essential statistical infrastructure that will support statistical production.
- Assist NSI's in Workforce Planning

- Assist with Career Paths and Succession Planning
- Assist mobility decisions, to ensure NSI's have the right people in the right place with the appropriate skills.

A structured Statistical Training Framework will give our staff:

- A greater understanding of the range of statistical skills, knowledge and expertise needed to work effectively throughout the end to end statistical production cycle, aligned to the GSBPM;
- A greater understanding of the role of NSI's in Official Statistics and the policy and legislative environment governing statistical work
- Structured statistical training path for the individuals
- Provide staff with career paths and succession planning
- Enable staff to share their knowledge and assist others in developing their statistical skills.

Competency based role profiles, skills register and Learning/Career Paths:

This framework will link to a set of technical competencies required to perform in the role of a Statistician in an NSI. Figure 2.4 presents a competency wheel which shows general technical competencies for the role of Statistician. Table 2.5 presents the technical competency and the headings each one encompasses in order to link the Competency wheel to the framework.

Figure 2.4 Technical Competency Wheel

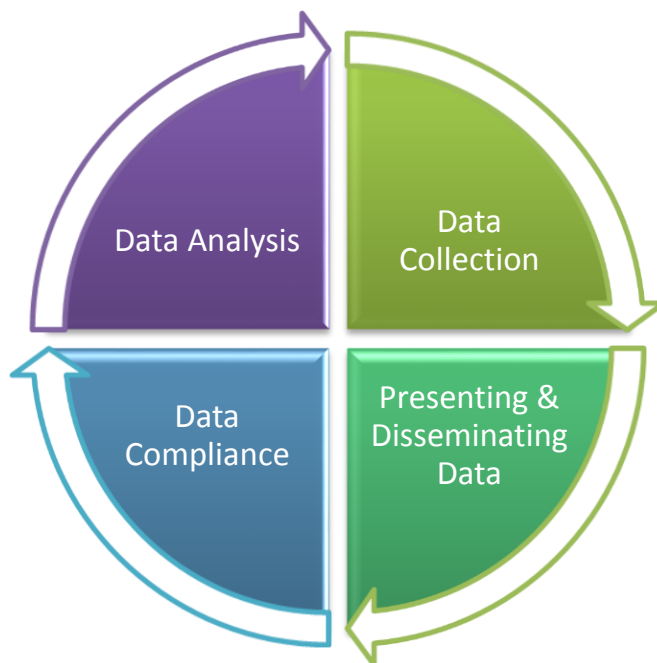


Table 2.5 Competency wheel linked to Statistical Training Framework.

Technical competency	Headings
Data Compliance	Data governance, Regulatory Statistical Frameworks,
Data collection	Register management, Imputation and Non-response Evaluating user statistical needs? Data matching, Integration & Admin Data** Survey & Questionnaire design Sampling & estimation***
Data Analysis	Index Numbers, Regression, Time series analysis & Adjustment, Data matching, Integration & Admin Data** Sampling & Estimation***
Presenting & Disseminating Data	Visualisation & presentation of data Statistical data editing Statistical disclosure control

*National accounts this particular heading sits outside the competency framework.

**Data matching, Integration & Admin data fits across two competencies

*** Sampling and Estimation fits across two competencies the heading is in bold where it fits.

Section three: How will the framework operate?

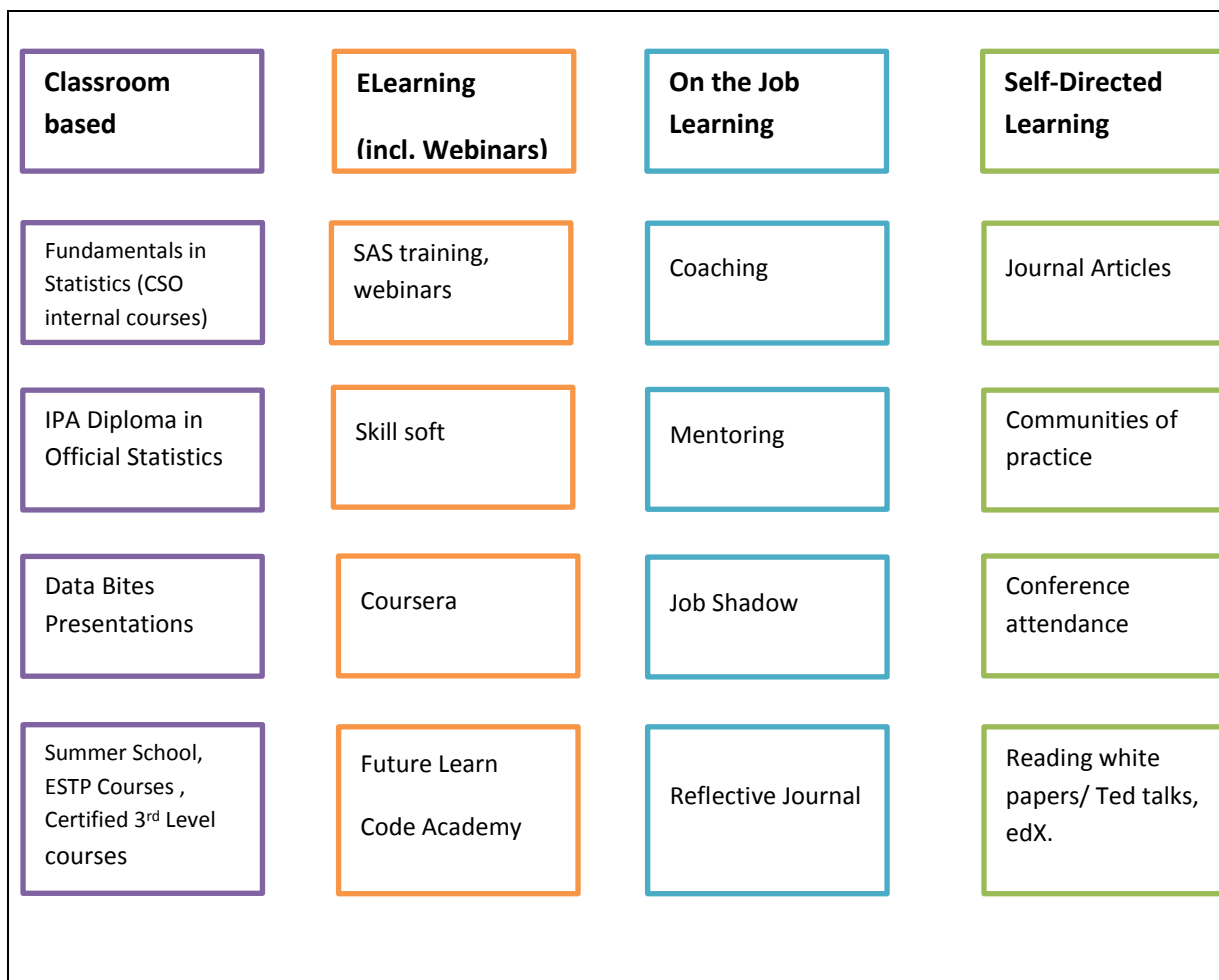
The framework is designed to operate in conjunction with other internal learning technologies. A number of training interventions have been identified which fit current Learning and Development practices. A detailed work flow of how this framework will operate is presented in appendix 5; this work flow goes through an example of a Statistician working in an area where there is a requirement to upskill for the particular role. The skills register, performance management system and learning paths are also utilised within the work flow to show how other learning technologies can be incorporated into the framework. Communities of practice is also noted as a knowledge management system where interactions are encouraged from colleagues in similar roles, this facilitates on the job learning and supports the individual through their learning journey.

The framework can also be utilised for a number of other key Learning and Development functions such as workforce planning, identification of required skills for role profiles and training interventions to meet skills gaps can be planned in a cost efficient manner. Statistical learning paths are easily identifiable so that individuals can map their performance goals to different methods of learning such as on the job, self-directed, classroom and web based. This framework will also assist with mobility such that if a Statistician has the required skill level they can be easily identified as suitable for a particular role this ensures that the right people with the right skills are in the correct roles performing to their full capability.

Section four: The Irish example: Training Interventions

In the CSO we are currently researching training interventions to meet the needs of the Framework. A specialised group has been set up internally to research best practice regarding online learning tools, classroom based options and delivery of courses in house. We are also researching gaining access to online journals to enable our knowledge network (Communities of Practice) to operate as a learning tool. The below diagram (4.1) depicts the key training interventions that we have identified to date.

Diagram 4.1 Key Training Interventions



The CSO has also developed an in house course entitled “Fundamentals in Statistics” to ensure both technical (Statisticians) and non-technical (general grades) have sufficient knowledge in basic statistical principles.

The course is run during working hours, 8 two hour modules will be delivered and assignments/course work will be handed to participants for completion during the course. The trainers on this course are internal experts in their field. The below table 4.2 outlines the key modules covered within this course. This course will set a minimum level for staff at all grades in developing statistical skills.

Table 4.2 Fundamentals in Statistics Course Outline

Week	Date	Time	Subject	Topic's covered
1	September 6, 2017	10.00 - 12.00	Data Summaries & graphs	Data measurements – sampling – visualisation - summary statistics
2	September 13, 2017	10.00 - 12.00	Statistical Models	Probability - Discrete probability distributions - Continuous probability distributions
3	September 20, 2017	10.00 - 12.00	Sampling Distributions	Sampling distribution of a mean – confidence interval of a single mean – sampling distribution of a proportion –Confidence interval of a single proportion
4	September 27, 2017	10.00 - 12.00	Simple Comparative Experiments	Hypothesis testing – 1 sample t-test/Z-test – independent t-test/Z-test – Paired t-test/z-test
5	October 4, 2017	10.00 - 12.00	Counted Data	1 sample proportion z-test – assumptions and examination of SAS output
6	October 11, 2017	10.00 - 12.00	Cross-Classified Frequency Data	Cross tabulations – chi-square distribution – expected vs observed frequencies –special cases – assumptions and examination of SAS output.
7	October 17, 2017	10.00 - 12.00	Introduction to Regression	Scatterplot & correlation – Notation & terminology – simple linear regression model – residuals and regression assumptions – multiple and logistic regressions
8	October 24, 2017	10.00 - 12.00	Introduction to ANOVA	When is ANOVA appropriate – ANOVA hypothesis – post-hoc analysis – assumptions – special cases

Other Activities:

- We are currently developing links with local Universities to provide modules on more advanced courses such as Sampling. This is under consideration and being researched by both Learning and Development and Methodology.
- We have trained over 15 internal Statistical trainers who will be available to deliver in house training in our three locations.

Communities of Practice:

Another training intervention currently being developed in house is a “*Community of Practice*” this consists of four key elements as linked to social learning theory. Diagram 4.3 depicts the key elements of how this will operate.

Diagram 4.3: Communities of Practice



Communities of Practice are defined as a group who share a passion for something they know how to do and who interact regularly to learn how to do it better. They are self-defined and self-managed, over time they establish their place in the organisation. In the beginning, they can utilise some light handed guidance regarding technology and infrastructure. Process support will also be provided by learning and development in terms of coaching and logistic assistance when setting up terms of the group, this enables a community that create examples of work to learn by doing. A core group will be gathered to initialise the launch process and identify potential leaders.

Learning and Development will also help organise initial value added activities through the Communities of practice portal. Active participation will be encouraged by all members and they will be encouraged to foster an increased responsibility for stewarding their knowledge.

Communities of practice create long term benefits for those involved by encouraging self-development, reputation enhancement, professional identity and networking opportunities. They also benefit the organisation by promoting innovative practices, creating knowledge databases and retaining talent.

Conclusion

This paper outlines how the GSBPM model was utilised as the basis for building a Statistical Training Framework; it ensures that each step in the statistical process is given the same focus when it comes to training. It also ensures that each staff member can get the best value from its resources, ensuring that we have the right people with the right skills in the right roles at the right time.

This paper also outlined the design structure of the framework, the benefits, how it links to competency development including role profiles, skills register and learning paths. We discussed how the framework will operate, the supporting learning elements taking Ireland as an example and showing an example of some of our planned training interventions.

This Statistical training framework is designed to work in any NSI, linking it to the GSBPM gives it a European focus and enriches the framework to ensure that each heading is designed to work with the guidelines of the GSBPM. The training interventions and learning structures which support this framework are crucial to its success. For example within the Irish context the Central Statistics office has overall responsibility under the Irish Statistical system for all Statistics produced within Government departments. Through knowledge networks such as *Communities of Practice* as outlined within this document, we can create a knowledge database and share information between departments on “what works”. We can also ensure that ongoing training initiatives are made available to staff and furthermore that staff have opportunities for self-development, on the job training and upskilling on an ongoing basis. We also provide more structured training programmes under our refund of fees scheme where employees can undertake certified courses in courses that relate to their roles.

This framework provides the necessary tools to ensure learning is made available to all staff at all stages of the GSBPM.

Appendix 3: Survey Data

Question 2: How do you assess the contribution of the GSBPM to the following issues?

Issue	Important / Very important	Not all that important / not important at all	Non Respondent/ Don't know
Development of statistical metadata systems	9	-	1
Standardisation of statistical business processes	9	-	1
Quality assessment of statistical business processes	9	-	1
Description of statistical business processes	9	-	1
Impact on the organisational structure	3	6	1
Increase of understanding of statistical business processes	7	1	1
Provision of an input to high-level corporate work planning	6	2	2
Measurement of operational costs	5	3	1

Survey data available at <https://statswiki.unece.org/display/GSBPM/Uses+of+GSBPM>.

APPENDIX 4: Sample of descriptors write up

Sampling and Estimation

Definition: Sampling is concerned with the selection of a subset of units (e.g. people, households, organisations etc.) from within a population to estimate characteristics of the whole population. By studying the sample we can generalize our results back to the population from which they were chosen.

Basic:

- I understand the difference between probability and non-probability sampling, and the advantages and disadvantages of each.
- I am aware of issues that arise with sampling frames (over coverage, under coverage, duplication etc.)
- I understand the difference between sampling and non-sampling errors.
- I understand how weights are used to calculate population totals and means.
- I understand how weights are used to compensate for unequal selection probabilities or for nonresponse.
- For simple random sampling, I can draw the sample and estimate the population mean and total of a given variable, the variance of these estimates and confidence intervals for them.
- From a simple random sample, I can estimate a population proportion for a given variable, and calculate the variance of the estimate, and confidence intervals for it.

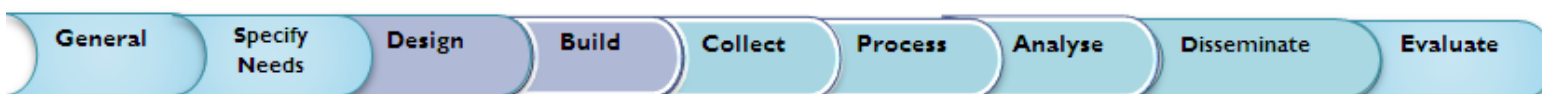
Intermediate:

- I understand the following four types of probability sampling: simple random sampling, stratified random sampling, cluster sampling and systematic sampling.
- I am aware of the strengths and weaknesses of each method, and can identify the most appropriate method for a given situation.
- For each type of probability sample, I can draw the sample and estimate the population mean and total of a given variable, the variance of these estimates, and confidence intervals for them.
- For each type of probability sample, I can estimate a population proportion of a given variable, and calculate the variance of the estimate, and confidence intervals for it.
- I understand and can implement ratio and regression estimation to improve the precision of sample estimates given good auxiliary data.
- For each type of probability sample, I can calculate the required sample size for estimating population means and totals, given a particular precision requirement.
- When designing stratified samples, I understand common methods of allocating the samples across the different strata, and the advantages and disadvantages of each method.
- I understand what the design effect of a cluster sample is, and how it affects the effective sample size of the survey, and the precision estimates of population parameters calculated from that sample.
- I know how to calculate the design effect of a cluster sample.
- I understand how post-stratification can be used to adjust a weighted sample for certain variables (e.g. age/sex) to make it conform to a known population distribution.

Advanced:

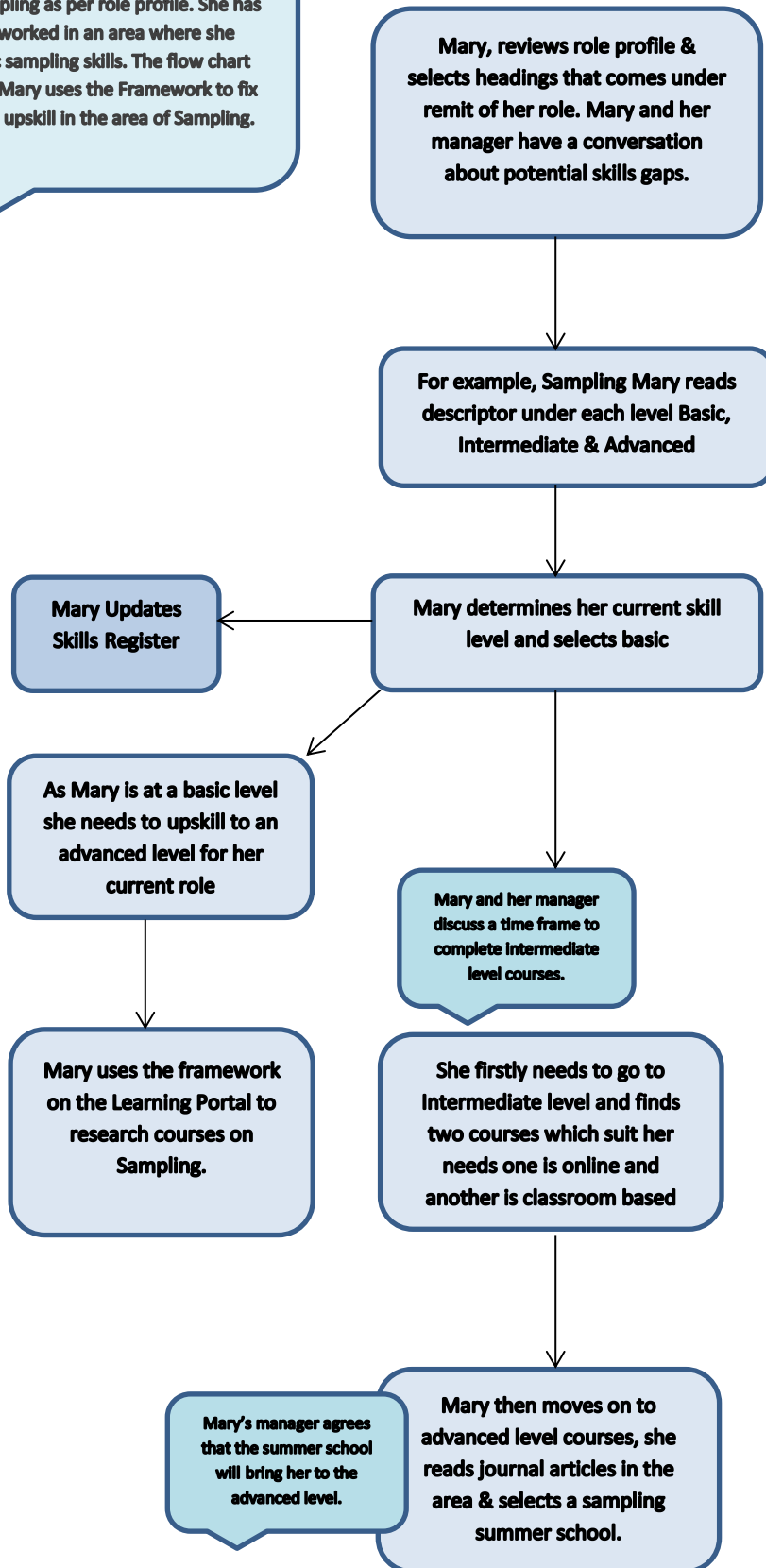
- I have a good understanding of multistage cluster samples and have experience in implementing them.
- I know how to analyse complex survey data and deal with issues of differential weighting, stratification and clustering.
- I am familiar with sample coordination methods which maximize or minimize the overlap between several samples drawn successfully in a population that changes with time.
- I am able to extend the design based sampling framework to include the model assisted approach, and understand the advantages and disadvantages of each.
- I understand the issues that arise from informative non-response, and how to design and implement strategies to compensate for it.
- I understand replication methods like balanced repeated replications, jackknife repeated replications and bootstrapping and can use them to estimate variance in complex surveys.
- I know how to design rotating surveys and multi-phase sample surveys.

Where do Sampling and Estimation skills fit into GSBPM model? (high-lighted in Purple)



Appendix 5: Employee Experience

Mary works in Business Statistics. She requires advanced sampling as per role profile. She has previously worked in an area where she needed basic sampling skills. The flow chart explains how Mary uses the Framework to fix this issue and upskill in the area of Sampling.



Within a short space of time Mary has upskilled in order to be competent in her role. The framework identified a skills gap the skills register logged this gap. The framework also helped to develop a learning path to bring Mary from Basic to advanced. PMDS was utilised to monitor and document progress. Mary also utilised Communities of Practice for on the job learning. This supported Mary through her learning Journey.

Performance Management & Learning Paths

