Phase 5 - Process

Phase 5
Any sub-process

Phase / Metadata object	Process							
	5.1 Integrate data	5.2 Classify and code	5.3 Review, validate and edit	5.4 Impute	5.5 Derive new variables and statistical units	5.6 Calculate weights	5.7 Calculate aggregates	5.8 Finalize data files
User needs								
Survey metadata								
Survey objectives								
Legal framework								
Statistical outputs								
Quality measures for statistical outputs								
Concepts	10 Reuse						10 Reuse	
Data Sources	Reuse							
Variables (Observation)	Reuse						Reuse	
Variables (Derived)					Reuse/update		Reuse	
Statistical aggregates							Create	Approve → 8
Statistical units	Reuse				Reuse/update		Reuse	
Classifications/Code lists	Reuse	Reuse					Reuse	
Data collection methods								
Questionnaire								
Target population							Reuse	
Register (Universe)								
Sampling frame								
Sampling method								
Processing methodology	Apply	Apply	Apply	Apply	Apply	Apply		
Data protecton requirements	Apply							
Validation rules			Apply					
Imputation method				Apply				
Weighting method						Apply		
Disclosure control rules								
Operational processes	Apply	Apply	Apply	Apply	Apply	Apply		
Respons rates								
Imputation rates				Create				
Weights						Create →	Reuse	
Coefficients of variation							Create → 9	

Integrating data needs the knowledge of the **concepts, classifications** and **statistical units** behind the **variables** of the **data sources** that are beeing integrated to be sure of the result obtained.
The **operational processes** and the **processing methodologies** applied in this sub-process defined before shall be applied and validated. For some variables there is the need to **classify** or to reclassify them (administrative data, for instance), and code or recode them too. **Imputation** of missing data is done also.
The validation/revision of the data obtained, can determine some changes on the rules on the metadata objects used to create them.
This fact can determine several iterations on the sub-processes of this phase.
After stabilizing the microdata, **weights** are calculated applying the **weighting method** and the **statistical aggregates** are calculated.
Usually, **coefficients of variation** of the main variables are calculated too.
Macrodata are obtained which will be submitted to several analysis in order to be disseminated.