



Targeted Review of Internal Models

Documentation request and list of mandatory data quality tests

Request to the institution for the review of the quality of the data used for IRB Models

Contents

1	REQUEST TO THE INSTITUTION FOR THE REVIEW OF THE QUALITY OF THE DATA USED FOR IRB MODELS	3
1.1	INTRODUCTION	3
1.2	MODULE 1 – GENERAL DOCUMENTATION REQUEST ON IRB DATA QUALITY MANAGEMENT FRAMEWORK AND IT INFRASTRUCTURE	4
1.3	MODULE 2 – PERFORMANCE OF DATA QUALITY CHECKS, ANALYSES AND RECONCILIATIONS	5
1.3.1	<i>Additional details on the scope of the mandatory tests</i>	7
A.	<i>PD Variables</i>	9
B.	<i>LGD Variables</i>	16
1.4	ANNEX 1 – ILLUSTRATIVE EXAMPLES	18

1 Request to the institution for the review of the quality of the data used for IRB Models

1.1 Introduction

The European Central Bank (ECB) has decided to perform a targeted review of internal models (TRIM), with the objective of reducing the unwarranted variability in RWAs by harmonising practices and checking the compliance of TRIM SSM approved Pillar 1 internal models for credit risk (CR), market risk (MR) and counterparty credit risk (CCR) with regulatory requirements.

In order to identify and assess the current practices of the institutions in terms of data quality as required in articles 144(1)(d), 174(b) and 176 of Regulation (EU) No575/2013 and in article 76(2) of the Final Draft RTS on assessment methodology for IRB and, in addition, to ensure efficient preparation by providing assessment teams with the same information before and during the on-site investigation (i.e., next phase of TRIM starting in Q3 2018), the present document provides instructions to be followed by the institutions.

In particular, the document lists the information and documentation the institution shall provide for an in-depth analysis of the IT infrastructure supporting IRB Approach and the data quality framework and processes in place at the institution and in particular applied for the selected model, together with the specifications of the tests and exercises to be performed by the institution regarding the quality of data used for the selected model.

Module 1 contains a documentation request mainly focusing on the institution IT infrastructure supporting the IRB Approach and the current data quality management processes. These instructions reflect the fact that a data quality review has already been performed, as part of past TRIM on-site investigations. However, even **if (part of) the documentation here requested has been already submitted, the institution is kindly asked to prepare this documentation**, with the aim to ensure the analysis of latest/more recent versions and to ease the operationalisation of the review.

The tests and exercises comprised in Module 2 of the data quality instructions are primarily aimed at assessing the quality of the data for a selection of variables within PD and LGD historical data to be used for modelling purposes within a reference period of three years, together with relevant variables for rating computation on current exposure for one annual snapshot¹, and mainly focusing on the data collection and storage stages of the data cycle.

In addition to the tests requested in Module 2 to be performed in advance by the institution, additional *ad hoc* tests or exercises can be requested by the assessment team during the on-site investigation (e.g., these may include data quality tests on historical data for PD/LGD model and segmentation drivers and/or collateral, specific analyses on certain databases/data processes or potentially additional tests aimed at assessing other data quality dimensions – as for example, for traceability by means of data lineage exercises for certain facilities/obligors).

All information, documentation, tests/exercises output and supporting evidence contained within this request is to be prepared by the institution in advance and be readily available at the start of the on-site investigation.

For additional clarifications, comments and any other issue regarding this request – in particular during the preparation period before the start of the on-site investigation – please use the attached

¹ For tests that leverage on exposure data, the institution shall use data that is representing the model under the current investigation. If doubts or questions arise, the institution should contact TRIM PMO through the FAQ process.

FAQ table and refer your questions to the following e-mail address: TRIM_FAQ_DQ@ecb.europa.eu. Please consider that a specific notation has been devised in order to identify the different tests contained in Module 2 for the purposes of this FAQ process (e.g. for the completeness test of the Default Flag variable within the PD mandatory tests is identified as PD_DF_COM).

1.2 Module 1 – General documentation request on IRB data quality management framework and IT infrastructure

In order to better understand data quality management practices within the institution and to provide for an effective starting point for the data quality on-site investigation which may serve to further calibrate the technical tests contained in the second module of this document, the following information/documents are expected to be available for the assessment team at the beginning of the on-site investigation²:

1. Organigram of the institution, with clear indication of the relevant parties involved in the management of data quality for the selected IRB model.
2. Overview of committees and other decision bodies involved in data quality management.
3. Policies and procedures on data management and data quality management (i.e., formalizing roles and responsibilities and governing/decision bodies, guidelines/principles and metric/indicator approach for the management of the quality of data, standardized procedures/workflows for improving data quality on an ongoing basis, data quality reporting procedures, etc.) and data processing along the IRB approach. In scope of this item are the Group policies and procedures and specific procedures for the IRB model under review.
4. IT infrastructure and architecture, illustrating the IRB data flow from the source to the output for the IRB model under review, both for modelling purposes and for rating computation on current exposure.
5. Related functional and technical documentation for the IRB model under review.
6. Overview of the data sources feeding the data bases used for the collection of historical data for modelling purposes and of current data for rating computation on current exposures, for the IRB model under review.
7. Overview of the data processes in place to incorporate external data into the model and of relevant data flows that incorporate inputs from third-parties involved in modelling or calibration (if any). Existing documentation related to these inputs (e.g., SLAs in place, data quality reports, related data quality procedures/ controls). Overview of (IT) systems used for the IRB model under review.
8. Data dictionary (for IRB related definitions) for the IRB model under review.

² Even if some of these documents may have already been delivered in the course of previous TRIM on-site investigations, these are kindly requested to be prepared, with the aim to ensure the analysis of latest/more recent versions and to ease the operationalisation of the review.

9. Layouts for all relevant databases/tables intervening in the IRB process, identifying the relevant key data fields for IRB purposes and main features for the IRB model under review (e.g. granularity).
10. Overview of the risk drivers for the IRB model under review.
11. Policy containing the definition of default.
12. Functional and technical documentation on the implementation of the default definition within the institutions systems/applications.
13. Documentation on validation and implementation tests on the definition of default. Identification of recent and projected changes to the definition of default and its technical implementation.
14. Policy describing the internal definition of obligor.
15. Documentation on technical implementation of the obligor identification within the institutions relevant systems/applications and data processes.
16. Any functional/technical traceability the institution may have documented specific to IRB Approach.
17. List of main data quality controls performed (i.e. technical controls, business controls, formatting controls, etc.) along the IRB process for the IRB model under review.
18. All reconciliations performed for IRB historical data and for data for rating computation on current exposure.
19. Data quality reports for IRB data for the IRB model under review produced in the last three years (2017, 2016 and 2015).
20. Audit and validation reports for exercises performed in the area of IRB models, on data quality related issues.
21. Existing documentation on IT testing procedures and tests results for IRB systems/applications for the IRB model under review.
22. Finally, a mapping table indicating for each one of the above items, the reference of the document or documents that contain the required information (include name and relevant sections where the information can be found).

1.3 Module 2 – Performance of data quality checks, analyses and reconciliations

In order to evaluate and assess the quality of the data currently used in the selected IRB model, the institution is requested to perform a set of checks/analyses/reconciliations and have results and supporting information/data to be readily available at the beginning of the on-site investigation.

The scope of the request is aimed at the historical data to be used for the modelling of the PD and LGD models (PD/LGD historical data), together with application data used for rating computation on

current exposure. A targeted set of variables for both PD and LGD historical data and for rating computation have been selected for the request and data quality will be assessed through the following data quality dimensions:

- (a) The values are present in the attributes that require them ('completeness').
- (b) The data is substantively error-free ('accuracy').
- (c) A given set of data can be matched across different data sources of the institution ('consistency').
- (d) The data values are up to date ('timeliness').
- (e) The aggregate data is free from any duplication given by filters or other transformations of source data ('uniqueness').
- (f) The data is founded on an adequate system of classification, rigorous enough to compel acceptance ('validity').
- (g) The history, processing and location of data under consideration can be easily traced ('traceability').

Besides the documentation requested to be readily available upon the on-site investigation, all relevant information and data used for the computation of the checks/analyses/reconciliations should also be available at the beginning of the investigation, as the assessment team will review and drill down the results. This includes:

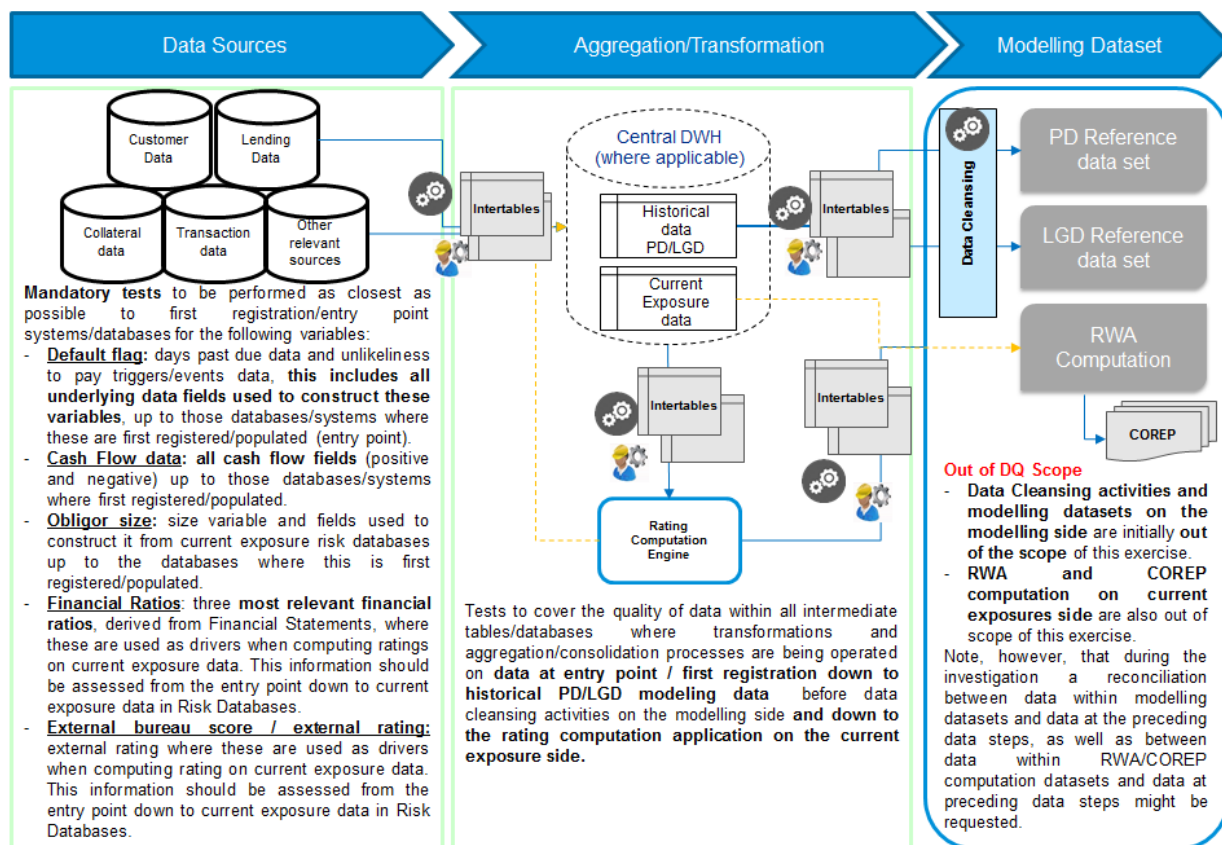
- The data(bases) and systems used;
- The data transformations and treatments operated before the performance of the checks and tests (for examples, queries/computer programs used).
- Test outcomes/results.
- The controls performed on the process or outcomes of the test.
- The relevant criteria and calculations used for the computation of the test (e.g., for missing analysis clarification on the criteria used to identify such missing cases – blanks, default values, etc. – is to be clearly stated and documented).
- Information on whether this test (or a similar one) is already performed internally in the institution on a regular basis outlining current measures in the institution for mitigating the issues arising from the tests, if any.
- Other supporting documentation/information (technical/functional documentation, databases layouts, etc.).

The institution should be ready to organise relevant walkthroughs in order to guide the assessment team through the process followed to perform the requested tests.

The detailed list of the tests on data quality to be performed and documented by the institution for each one of the variables for PD and LGD historical data, as well as variables for rating computation, can be found in sections A and B of this document.

1.3.1 Additional details on the scope of the mandatory tests

The ultimate goal of the tests is to assess the quality of data at entry point (first record/creation in the systems), before any manual treatment or discretionary enrichment of the data has been performed. A graphical representation of the intended scope for the mandatory tests along the PD/LGD historical data and current exposure dataflow is presented below:



Annex 1 provides an illustrative example of the intended scope and the extractions the institution is expected to perform for the purpose of the mandatory tests for the analyses of the data quality of the default flag.

More generally regarding the tests:

- The institution shall trace back the variables (and the respective fields used to construct them) up to the sources where the data used for the computation of the variables is first registered. It is from this entry point onwards that the tests shall be performed. In case the required tests were to start further along in the PD/LGD historical data and the current exposure dataflow, the institution shall evidence that no manual adjustments or discretionary treatments have been operated to the original/raw data down to that point. Where local entities exist with their own systems and applications which in turn feed data at the Group level, the institution shall perform the tests up to the sources of such local entities.

- Only exceptionally, when tracing the variables back to the sources is not feasible due to the multiplicity of sources, the tests could focus on a subset of those sources which is deemed representative.
- The tests shall provide a view into the quality of data for the exposures comprised within the selected model. If the institution is not able to isolate those facilities in scope of the model at a certain data step of the dataflow or database, this results in the tests being performed at this point for a different perimeter of exposures than the model perimeter. In this case, it shall provide evidence that the results of the test performed are still representative for the exposures under scope.
- Finally, when the same tests as those described in sections A and B below are already implemented by the institution and performed on a regular basis, these can be used for the purpose of this request.

In addition to this first list of tests, *ad hoc* tests, exercises and analyses can be requested by the assessment team for the institution to perform during the on-site inspection. Examples of such additional tests include the following:

- Analysis of other model specific variables: for example, PD and LGD segmentation data, collateral data, other current exposure drivers.
- Analysis of specific data processes or databases: for example, manual inputs, representative databases.
- Tests aimed at assessing specific data quality dimensions: for example, data lineage exercise on a sample of obligors/facilities for assessing traceability of data.

A. PD Variables

A set of mandatory tests will be performed on relevant variables on the PD side, concerning both historical data used for modelling purposes and data on current exposure, used for rating computation.

1. PD historical data

Listed below you will find the set of tests to be performed for the variables within PD historical data at the reference dates.

Reference dates, reference period: PD historical data for 2015, 2016 and 2017³, e.g. three annual snapshots for the tests described below, except for the cross consistency check of the default flag, where a continuous 3-year period should be considered.

1.1 Default flag

The checks described in this section deal with the quality of the data used⁴ for the computation of **default flag** of all exposures from data sources to PD historical data.

Tests to be performed by the institution include the following:

(i) **PD_DF_COM**: Completeness check

- Check specification: missing analysis of the (raw) data used for the computation of the default flag. Cases of missing information (for instance DPD missing) shall be reported by computing the proportion^{5 6} of obligors/facilities with non-populated fields: (blanks or similarly default values) or with erroneous ones (not present in the internal nomenclature).
- Scope: all fields used for the computation of the default flag within all databases along the IRB workflow, from source databases to PD historical data.
- Expected output: for the reference dates, report the proportion of missing for each database/dataset identified in the scope in terms of number and in terms of exposure. Information (name, identifier, description, particularities...) relative to the databases/datasets on which the test has been performed must be reported.

(ii) **PD_DF_ACU**: Accuracy check

- Check specification: statistical profiling (distribution analysis of obligors/facilities) of the (raw) data used for the computation of the default flag indicator and comparison of the distribution for the different databases. For distribution comparison of the categorical

³ In case this reference period cannot be applied, a different reference period will have to be formally requested by the institution and approved by TRIM PMO. Please contact TRIM PMO through the FAQ email address TRIM_FAQ_DQ@ecb.europa.eu as soon as this circumstance is detected.

⁴ E.g. 'Days Past Dues (DPD)' data and 'unlikeliness to pay triggers' data.

⁵ The proportion in Number = Card.(Tested fields) / Card.(Total fields).

⁶ The proportion in Exposure = Total Exposure (Tested fields) / Total Exposure (Total fields).

variables (default flag indicators and default event indicators), the diversity index (e.g. Shannon index) can be computed and compared in addition to the graphical comparison.

- Scope: all fields used for the computation of the default flag within all databases along the IRB workflow, from source databases to PD historical data.
- Expected output: for the reference dates, report the metrics (e.g. Shannon index) and the graphics (e.g. Histograms, box-plots etc.) computed for each database/dataset identified. Information (name, identifier, description, particularities...) relative to the databases/datasets on which the test has been performed must be reported.

(iii) **PD_DF_CC:** (Cross) Consistency check

- Check specification: reconciliation of obligor defaults and defaulted exposure amount between source databases and historical data. For the reference period, defaulted cases shall be tracked from the source databases identified in the IRB workflow, reconciled with the number of obligor defaults and defaulted exposures in the historical datasets. The proportion of inconsistencies should be reported (e.g., if n defaults were identified based on the information present in the source database during a given period, those defaults shall also be found in the related databases – PD historical data and LGD historical data). Additionally, for the reference dates, it shall be checked whether obligors flagged as defaulted in PD historical data are consistent with those in LGD historical data and report the proportion of inconsistencies.
- Expected output: for the reference dates, report the difference in terms of number and in terms of exposure between the reconciled databases and document and explain the differences observed. Information (name, identifier, description, particularities, etc.) relative to the databases/datasets on which the test has been performed must be reported.

2. PD current exposure data

Listed below you will find the set of tests to be performed on the relevant variables for rating computation at the reference date.

Reference date: Most recent rating assignment snapshot for each obligor⁷.

2.1 Obligor Size⁸

The checks described in this section deal with the quality of the data used for the computation of **obligor size**, from first registry down to current exposure data in risk databases, where obligor size is a relevant driver for the selected model.

Where the obligor size variable has not been identified as a relevant risk driver in the PD model of the institution, it should be replaced by an alternative variable that has been identified as a risk driver with the highest weight. This alternative risk driver shall be identified consistently with the information reported in the template PD Data request, as described in tab PD.3.Instructions.

This alternative variable is to be reviewed by means of tests on data completeness, accuracy, consistency and timeliness.

Test to be performed by the institution include the following:

(i) **CE_SIZ_COM**: Completeness check

- Check specification: missing analysis by computing the proportion of obligors without a size value (blank values or similarly default values) or with erroneous ones.
- Scope: obligor size field within all databases used in the IRB workflow from source databases down to current exposure data in risk databases.
- Expected output: for the reference date, report the proportion of missing for each database/dataset identified in the scope, in terms of number of obligors and in terms of exposure. Information (e.g., name, identifier, description, particularities) relative to the databases/datasets on which the test has been performed must be reported.

(ii) **CE_SIZ_ACU**: Accuracy check

- Check specification: statistical profiling of the obligor size variable data and comparison of the distributions observed across different databases and datasets. Anomalies in these distributions should be explained and documented.
- Scope: obligor size field within all databases used in the IRB workflow from source databases down to current exposure data in risk databases.
- Expected output: for the reference date, report the size at obligor level and the overall distributions of obligor size along different databases and datasets across the IRB workflow and identify outliers/anomalies. Unexpected values should be reported in terms

⁷ In case the lack of implementation of some elements of the model (e.g. for TRIMIX) impacts the scope or conditions of the tests, some modified scope or conditions will have to be formally requested by the institution and approved by TRIM PMO. Please contact TRIM PMO through the FAQ email address TRIM_FAQ_DQ@ecb.europa.eu as soon as this circumstance is detected.

⁸ For PD models covering exposures to financial institutions, define buckets of total assets. For PD models covering exposures to large corporates, define buckets of total turnover.

of number of obligors and in terms of exposure. Information (e.g., name, identifier, description, particularities) relative to the databases/datasets on which the test has been performed must be reported.

(iii) **CE_SIZ_CC:** (Cross) Consistency check

- Check specification: reconciliation, at obligor level, of obligor size data between sources and subsequent databases down to current exposure data in risk databases. Obligor size shall be tracked from the source databases identified in the IRB workflow, reconciled with size of the obligor stored in risk databases containing current exposure data. The proportion of inconsistencies should be reported.
- Expected output: for the reference date, report the differences between reconciled databases in terms of numbers and in terms of exposure, and provide differences explanation. Information (name identifier, description, particularities, etc.) relative to the databases/datasets on which the test has been performed must be reported.

(iv) **CE_SIZ_TIM:** Timeliness check

- Check specification: analysis of the reference date of the obligor size information.
- Scope: obligor size data used for rating computation within all databases and datasets used in the IRB workflow from the sources down to risk databases containing current exposure datasets.
- Expected output: report a table in which for each obligor at the reference date (most recent rating assignment), the following two dates are to be provided:
 - a. Reference date of obligor size data used for latest rating assignment.
 - b. Date of the latest rating assignment.

2.2 Financial ratios

The checks described in this section deal with the quality of the data used for the computation of the **three most relevant financial ratios** derived from Financial Statements, used by the institution when computing ratings, from first registry down to current exposure data in risk databases, where these three financial ratios are a relevant driver for the selected model.

Where the financial ratios have not been identified as a relevant risk driver in the PD model of the institution, it should be replaced by an alternative variable that has been identified as a risk driver with the highest weight. This alternative risk driver shall be identified consistently with the information reported in the template PD Data request, as described in tab PD.3.Instructions.

This alternative variable is to be reviewed by means of tests on data completeness, accuracy, consistency and timeliness.

Tests to be performed by the institution include the following:

(i) **CE_FR_COM:** Completeness check

- Check specification: missing analysis of the financial statements data used for the computation of the financial ratios as well as for the financial ratio itself. Cases of missing information shall be reported by computing the proportion of obligors with non-populated fields: (blanks or similarly default values) or with erroneous ones (not present in the internal nomenclature).

- Scope: financial statements data used for the computation of selected financial ratios within all databases and datasets used in the IRB workflow from financial statements first entry in the institution systems down to current exposure data in risk databases.
 - Expected output: for the reference date, report the proportion of missing for each database/dataset identified in the scope. Information (name, identifier, description, particularities...) relative to the databases/datasets on which the test has been performed must be reported.
- (ii) **CE_FR_ACU**: Accuracy check
- Check specification: statistical profiling of the financial statements data used for the computation of selected financial ratios and comparison of the distributions observed across different databases and datasets. Anomalies in these distributions should be explained and documented.
 - Scope: financial statements data used for the computation of selected financial ratios within all databases used in the IRB workflow from financial statements down to current exposure data in risk databases.
 - Expected output: for the reference date, report the financial statements data used for the computation of selected financial ratios at obligor level and the overall distributions of financial statements data along different databases and datasets across the IRB workflow and identify outliers/anomalies. Unexpected values should be reported in terms of numbers and in terms of exposure. Information (name, identifier, description, particularities, etc.) relative to the databases/datasets on which the test has been performed must be reported.
- (iii) **CE_FR_CC**: (Cross) Consistency check
- Check specification: reconciliation, at obligor level, of financial statements data used for the computation of selected financial ratios data between sources and current exposure data. Data used for the computation of selected financial ratios shall be tracked from the financial statements, and reconciled with the values stored risk databases containing current exposure data. The proportion of inconsistencies should be reported.
 - Expected output: for the reference date, reconcile the financial statements data used for the computation of selected financial ratios, between databases. Identified differences shall be reported in terms of number of obligors and in terms of exposure, and provide differences explanation. Information (name identifier, description, particularities, etc.) relative to the databases/datasets on which the test has been performed must be reported.
- (iv) **CE_FR_TIM**: Timeliness check
- Check specification: analysis of the reference date of the financial statements data used to compute the selected financial ratios at obligor level.
 - Scope: financial statements data used for the computation of selected financial ratios within all databases and datasets used in the IRB workflow from financial statements down to risk databases containing current exposure datasets (including accounting systems).
 - Expected output: report a table in which for each obligor at the reference date (most recent rating assignment), the following two dates are to be provided:
 - a. Reference date of financial statements data used to compute the selected financial ratios at latest rating assignment.

- b. Date of the latest rating assignment.

2.3 External bureau score / external rating

The checks described in this section deal with the quality of **external bureau score / external rating** data, used by the institution when computing ratings, from first registry down to current exposure data in risk databases, where external bureau score / external rating is a relevant driver (the main or one of the main input variables for the internal rating) for the selected model.

Where the external bureau score / external ratings have not been identified as a relevant risk driver in the PD model of the institution, it should be replaced by an alternative variable that has been identified as a risk driver with the highest weight. This alternative risk driver shall be identified consistently with the information reported in the template PD Data request, as described in tab PD.3.Instructions.

This alternative variable is to be reviewed by means of tests on data completeness, accuracy, consistency and timeliness.

Tests to be performed by the institution include the following:

(i) **CE_ER_COM:** Completeness check

- Check specification: missing analysis by computing the proportion of obligors without external bureau score / external rating information (blank values or similarly default values) or with erroneous ones.
- Scope: external bureau score / external rating values within all databases and datasets used in the IRB workflow from the external sources down to current exposure data in risk databases.
- Expected output: for the reference date, report the proportion of missing for each database/dataset identified in the scope. Information (name, identifier, description, particularities...) relative to the databases/datasets on which the test has been performed must be reported.

(ii) **CE_ER_ACU:** Accuracy check

- Check specification: statistical profiling of the external bureau score / external rating observed across different databases and datasets. Anomalies in these distributions should be explained and documented.
- Scope: external bureau score / external rating values within all databases and datasets used in the IRB workflow from the external sources down to current exposure data in risk databases.
- Expected output: for the reference date, report the external bureau score / external rating values at obligor level and the overall distributions of external bureau score / external rating values along different databases and datasets across the IRB workflow and identify outliers/anomalies. Unexpected values should be reported in terms of numbers and in terms of exposure. Information (name, identifier, description, particularities, etc.) relative to the databases/datasets on which the test has been performed must be reported.

(iii) **CE_ER_CC:** (Cross) Consistency check

- Check specification: reconciliation, at obligor level, of external bureau score / external rating values between external sources and current exposure data. External bureau score / external rating shall be tracked from the external source databases identified in the IRB workflow and reconciled with the values stored in risk databases containing current exposure data. The proportion of inconsistencies should be reported.
- Expected output: for the reference date, reconcile external bureau score / external rating values between databases. Identified differences shall be reported in terms of number of obligors and in terms of exposure, and provide differences explanation. Information (name identifier, description, particularities, etc.) relative to the databases/datasets on which the test has been performed must be reported.

(iv) **CE_ER_TIM**: Timeliness check

- Check specification: analysis of the reference date of the the external bureau score / external rating information, used in the ratings computation at obligor level.
- Scope: external bureau score / external rating values within all databases and datasets used in the IRB workflow from the external sources down to current exposure data in risk databases.
- Expected output: report a table in which for each obligor at the reference date (most recent rating assignment), the following two dates are to be provided:
 - a. Reference date of external bureau score / external rating information used to compute the selected financial ratios at latest rating assignment.
 - b. Date of the latest rating assignment.

B. LGD Variables

A set of mandatory tests will be performed on relevant variables on the LGD parameter that will focus only on historical data used for modelling purposes.

1. LGD historical data

Listed below you will find the set of tests to be performed for the variables within LGD historical data at the reference dates:

Reference period and dates: LGD historical data for 2015, 2016 and 2017⁹ (all cash flows related to any default present at some point in time during the 2015 to 2017 three year period).

1.1 Cash flow amount

The checks described in this section deal with cash flow amounts (both inflows and outflows, e.g., recoveries or additional drawings) recorded in the source systems during the recovery process and how these make their way to LGD historical data. This exercise shall also include cash flows derived from exposure variations.

Tests to be performed by the institution include the following:

(i) **LGD_CFA_COM**: Completeness check

- Check specification: checking the proportion of missing flows¹⁰ in the LGD historical data used for modelling.
- Scope: all cash flows (observed or calculated) related to any default present at some point in time between 2015 and 2017.
- Expected output: proportion of observations¹¹ missing (total and per year) in the historical risk database.

(ii) **LGD_CFA_ACU**: Accuracy check

- Check specification: distribution analysis of gross cash flow data (e.g. yearly recovery amounts / EaD).
- Scope: cash flow amounts fields within all databases used in the IRB workflow from the source databases to LGD historical data.
- Expected output: for the reference dates, report the distributions of annual gross recoveries / EaD along the IRB workflow and identify outliers/anomalies in terms of number and exposures.

(iii) **LGD_CFA_CC**: (Cross) Consistency check

⁹ In case this reference period cannot be applied, a different reference period will have to be formally requested by the institution and approved by TRIM PMO. Please contact TRIM PMO through the FAQ email address TRIM_FAQ_DQ@ecb.europa.eu as soon as this circumstance is detected.

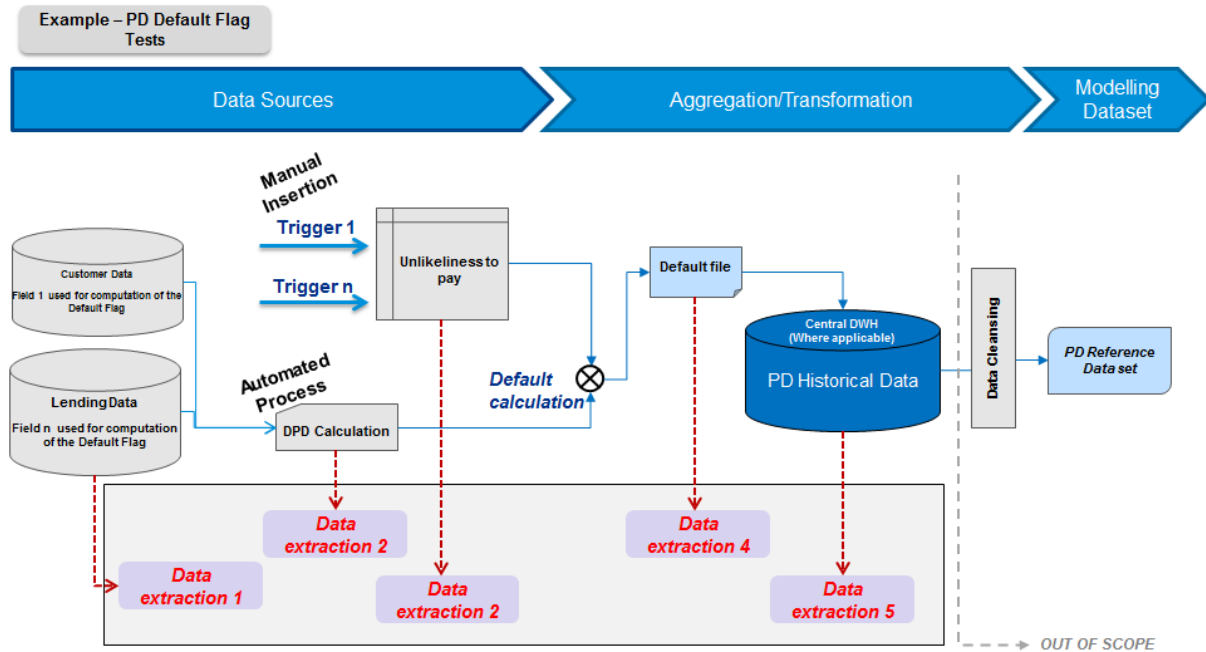
¹⁰ For instance, money or sale proceeds collected and recorded in the source application but not stored accordingly into the LGD historical data.

¹¹ Monthly/quarterly/annual depending on the recovery process.

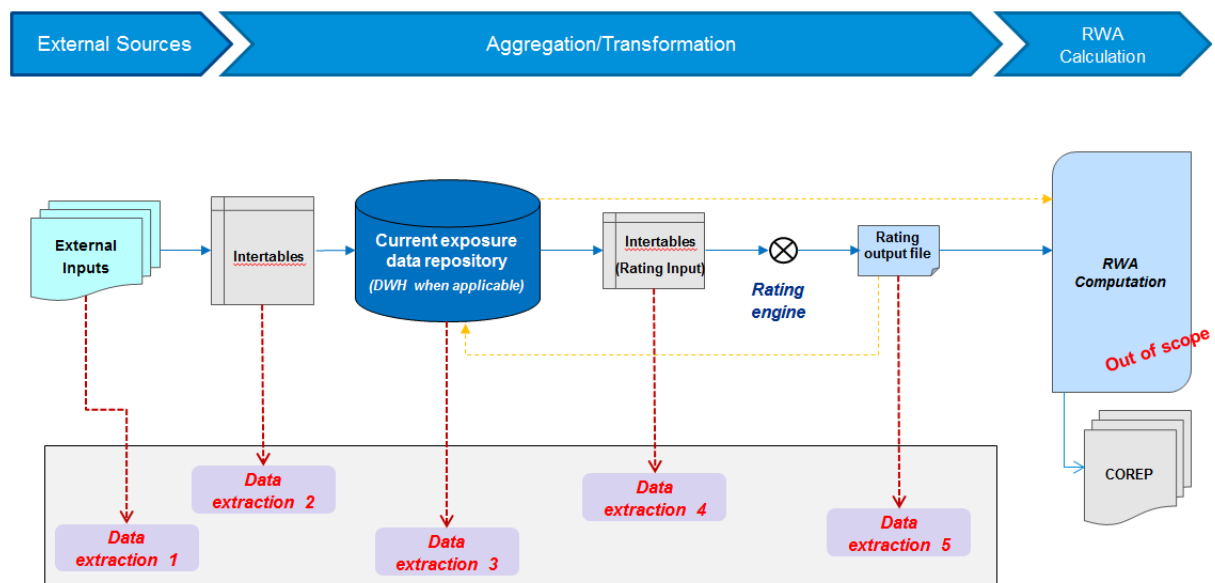
- Check specification: reconciliation between LGD historical data and the accounting data for the default cases closed during the reference period.
- Expected output: for the reference dates, report the difference in terms of number and in terms of exposure between the final recovery (whether observed or calculated) and accounting loss, and explain the differences observed.

1.4 Annex 1 – Illustrative examples

The dataflow chart below provides an illustrative example of the scope of some of the mandatory tests, namely the completeness and accuracy checks for the default flag calculation. This graphical representation identifies the relevant databases and steps that are targeted and for which the institution is expected to perform data extractions and present the assessment team with the test results.



In addition, an example of the scope of those mandatory tests that assess current exposure for rating computation data can be found below. In this case relevant databases and steps are portrayed from the external input down to the rating computation engine.



Acronyms

DPD	Days Past Due
DQ	Data Quality
EaD	Exposure at Default
EBA	European Banking Authority
IRB	Internal Ratings Based
LGD	Loss Given Default
PD	Probability of Default
Draft RTS on AM	Final Draft Regulatory Technical Standards on the IRB Assessment Methodology Approach (published by EBA on 21 July 2016)
SSM	Single Supervisory Mechanism
TRIM	Targeted Review of Internal Models