# Contexte et objectif

**Contexte**

Le projet vise à automatiser la production des datapacks utilisés pour la titrisation synthétique, afin de répondre aux exigences réglementaires et de fournir des données fiables aux investisseurs. Des problèmes actuels incluent des erreurs dans l’identification des entités, des incohérences de segmentation, un taux élevé de données manquantes et la présence d’outliers.

**Objectifs**

Automatiser la collecte, le traitement et la production des indicateurs et datapacks, ainsi que la réalisation d’analyses de gap (comparaison entre résultats N et N-1). Le périmètre wholesale (gammes RACER et JUNON) est priorisé.

**Cible**

Utilisation d’un environnement de données fiable (UV MOD) et d’une solution basée sur Python/PySpark, intégrant les évolutions structurelles (nouvelle segmentation IFRS9[[1]](#footnote-1), fusion des périmètres CDN et BDDF, nouveaux modèles PD, LGD, EAD).

# Rappel des Indicateurs

## Matrices de transition estimées sur un an (Average One-Year Estimated Transition Matrices)

Les matrices de transition sont utilisées dans le cadre de la calibration IFRS 9 des provisions comptables. Elles reflètent les transitions de notation interne des portefeuilles SME et LC sur une période d’un an, en moyenne sur la période considérée.

* Pour RACER 3, la période de référence est T4 2007 à T3 2021.
* La définition du défaut utilisée est celle de Bâle (article 178 CRR), appliquée dans les systèmes internes de SG. Le défaut correspond à une notation interne de 8, 9 ou 10.
* Les matrices présentent, pour les segments SME et LC, le pourcentage annuel moyen de transitions par nombre de contreparties (hors transitions vers des contreparties non notées).

Chaque matrice montre :

* En ligne : la notation à la date d’observation initiale ;
* En colonne : la notation après un an ;
* Les notations vont de 1 à 20 (1 étant défaut) et correspondent aux notations internes

Veuillez noter que la taille de l’échantillon d’observation est plus faible pour les notations plus élevées.

## Matrices de transition estimées sur cinq ans (Average Five-Year Estimated Transition Matrices)

Des matrices de transition sont également fournies pour chaque segment (SME et LC) sur un horizon de cinq ans. Elles présentent les taux de transition moyens estimés sur cet horizon pour la période considérée.

## Matrices de transition observées sur un an (One-Year Observed Transition Matrices)

En complément des matrices estimées, des matrices de transition observées annuellement sont calculées pour chaque année de la période disponible.

Par exemple, si la période couvre 10 années, alors 10 matrices de transition observées (une par année) doivent être fournies.

Ces matrices permettent de comparer les données réelles observées à celles estimées et d’analyser la stabilité des transitions dans le temps.

## **Taux de défaut cumulés moyens par millésime d'origine (Average Cumulative Default Rates by Origination Vintage, by Number of Loans)**

Ces taux sont fournis pour les segments SME et LC. Ils correspondent aux taux de défaut cumulés historiques ventilés par année d'origine des prêts sur la période considérée (de 2008 à 2021 pour RACER 3).

Le taux est exprimé en nombre de prêts : il représente le nombre de prêts en défaut après *X* trimestres rapporté au nombre total de prêts originés durant l’année concernée.

L’ensemble des résultats est segmenté par type de portefeuille Bâle :

* P5 – SME
* P4 – LC

Puis, une ventilation par tranche de notation interne est appliquée :

* Toutes notations agrégées
* [1 à 3−]
* [4+ à 5−]
* [6+ à 7−]

## **Taux de défaut cumulés moyens par millésime d'origine, par montant (Average Cumulative Default Rates by Origination Vintage, by Amount)**

Les mêmes taux de défaut cumulés historiques que ceux décrits en 2.4 sont ici calculés par montant.  
Pour une année d’origine donnée et un trimestre donné, le taux correspond au rapport entre :

* la somme des expositions des prêts entrés en défaut sur l’horizon temporel considéré ;
* et (ii) la somme des expositions des prêts originés durant l’année concernée.

## **Taux de défaut cumulés moyens par millésime de notation, par nombre de contreparties (Average Cumulative Default Rates by Rating Vintage, by Number of Counterparties)**

Ces taux sont fournis pour chaque segment (SME et LC) et correspondent aux taux de défaut cumulés moyens sur différents horizons temporels (de 1 trimestre à 32 trimestres, soit 8 ans) observés sur la période disponible (de Q1 2008 à Q3 2021 pour RACER 3).

Le calcul est réalisé par millésime de notation, c’est-à-dire selon la notation initiale des contreparties au début de l’horizon temporel.

Le taux est défini comme le rapport entre :

* le nombre de contreparties tombées en défaut pendant l’horizon ;
* et (ii) le nombre de contreparties saines initiales dans la même tranche de notation.

Les taux sont agrégés via une combinaison de cohortes mensuelles de contreparties et de leur notation initiale. Par exemple, pour un horizon de 32 trimestres, les données couvrent 36 cohortes (soit 3 ans) entre janvier 2008 et décembre 2010.

Les défauts observés sur ces cohortes jusqu’à 32 trimestres après leur date de référence permettent de calculer les taux finaux (ex. défauts de janvier 2011 à décembre 2018 pour la cohorte de décembre 2010).

À noter : plus l’horizon est long, plus la taille de l’échantillon est réduite, ce qui peut impacter la robustesse des résultats. Ces taux sont utilisés dans la calibration des matrices de transition estimées.

## **Taux de défaut cumulés moyens par millésime de notation, par montant (Average Cumulative Default Rates by Rating Vintage, by Amount)**

De manière analogue à la section précédente, les taux de défaut sont ici calculés par montant.  
Ils correspondent au rapport entre :

* la somme des expositions des contreparties en défaut sur l’horizon temporel considéré ;
* et (ii) la somme des expositions des contreparties saines initiales, classées par catégorie de notation au début de cet horizon.

## **Taux de défaut annuels sur un an, par nombre (Annual 1-Year Default Rates, by Numbers)**

Pour chaque segment (SME et LC), les taux de défaut observés sur un horizon d’un an sont présentés par notation interne pour différentes années.

Le taux est calculé comme suit :

* Nombre de contreparties en défaut sur une période d’un an ;
* divisé par (ii) le nombre de contreparties saines de la même notation au début de la période.

Les taux sont moyennés à partir de 12 cohortes mensuelles pour chaque année. Par exemple, pour 2008 : les cohortes de janvier à décembre 2008 sont chacune observées sur une période d’un an. Ainsi, la cohorte de janvier 2008 est observée jusqu’en janvier 2009.

**Remarque** : Lorsque le dénominateur est nul, la valeur retournée est NaN.

## **Taux de défaut annuels sur un an, par montant (Annual 1-Year Default Rates, by Amounts)**

Les mêmes taux de défaut annuels par notation sont également calculés par montant :

* Somme des expositions aux contreparties tombées en défaut sur une période d’un an ;
* divisé par (ii) la somme des expositions des contreparties saines de la même notation au début de la période.

## **LGD observées par année de défaut (Observed LGDs by Year of Default)**

Cette section vise à présenter :

* Le nombre d’obligors et de facilités en défaut (segments SME et LC) par année de défaut ;
* Les LGD observées sur les dossiers clôturés au 31/12/2021 (dans le cadre de RACER 3).

Les dossiers clôturés sont ceux pour lesquels le processus de recouvrement est terminé. Une répartition des facilités encore en cours de traitement est également présentée.

## **Back-in-Bonis**

Les statistiques agrègent les contreparties défaillantes sur l’ensemble des segments non-retail.

Les données incluent :

* Le nombre de dossiers de défaut fermés ayant été régularisés (c’est-à-dire revenus en statut “in bonis”) entre Q4 2007 et Q3 2021 (période RACER 3) ;
* La durée moyenne (en années) passée en défaut avant le retour à une situation saine.

Une contrepartie est considérée “in bonis” lorsque :

* Les conditions de défaut ne sont plus remplies pendant au moins 3 mois, et
* Tous les paiements sont à jour.

## **Taux de recouvrement cumulés par année de défaut (Cumulative Recovery Rates by Year of Default)**

Pour chaque segment (SME et LC), sont présentés :

* Le montant d’exposition des contreparties classées en défaut au cours d’une année donnée ;
* Et les taux de recouvrement cumulés trimestriels à partir de cette année.

Exemple : pour RACER 3, l’analyse couvre les années de défaut 2008 à 2020, avec une observation des recouvrements sur les trimestres suivants.

## **Taux de recouvrement cumulés moyens (Average Cumulative Recovery Rates)**

Pour chaque segment (SME et LC), cette section présente les taux de récupération cumulés moyens observés sur des horizons temporels allant de 1 à X trimestres après la date de défaut.

L’analyse porte par exemple sur les années 2008 à 2020, et permet d’observer l’évolution moyenne des recouvrements au fil du temps après le défaut.

## **Prépaiements (Prepayments)**

La première étape consiste à définir un périmètre restreint et représentatif du portefeuille titrisé, en appliquant des critères d’éligibilité pertinents et matériels, comme par exemple :

* PCRU = BDDF
* Segment = PME et Grandes Entreprises
* Notation interne à l’origine > 6+
* Méthode Bâle = IRBA
* Devise des prêts = EUR

Ces critères permettent de se rapprocher du portefeuille de référence, pour une vision plus fidèle du comportement de prépaiement.

Pour chaque trimestre disponible :

* Le montant total des encours de prêts en début de trimestre est mesuré ;
* Ainsi que le montant total des prépaiements (amortissements anticipés partiels ou totaux) observés durant le trimestre.

Le CPR (Conditional Prepayment Rate) annualisé est alors calculé par la formule suivante :

## **Montants et nombre de prêts émis chaque année (Amounts and Number of Loans Originated per Year)**

L’objectif est de présenter :

* Le nombre de prêts émis ;
* Et le montant total des prêts octroyés chaque année.

Ces statistiques permettent de suivre l’évolution des volumes et valeurs des nouvelles émissions de prêts sur la période considérée, et ainsi d’analyser les dynamiques de croissance ou de contraction de l’activité de prêt.

## **Statistiques sur les pertes cumulées et provisions (Cumulative Losses and Provisions)**

Comme pour les autres analyses, la première étape consiste à définir un périmètre de données éligibles en appliquant des critères pertinents et matériels, représentatifs du portefeuille titrisé. Par exemple :

* PCRU = BDDF
* Segment = PME et Grandes Entreprises
* Notation interne à l’origine > 6+
* Méthode Bâle = IRBA
* Devise des prêts = EUR

Ces critères visent à constituer un répertoire de données reflétant au mieux le portefeuille de référence, afin de fournir une meilleure vision des pertes historiques.

Pour chaque date de clôture de fin d’année, de 2008 à 2021, on observe les éléments suivants :

* Le montant des prêts sains ;
* Le montant des prêts en défaut ;
* Le montant cumulé des pertes depuis l’origine ;
* Les provisions constituées sur les prêts en défaut.

Le **Coût du Risque (CoR)** est alors calculé comme suit :

Ce ratio reflète, à une date donnée, la perte cumulée estimée sur le stock de prêts en défaut, et représente une approximation de la perte finale attendue pour les encours les plus anciens.

## **Typologie des défauts observés (Typology of Defaults)**

L’objectif est d’analyser, par année, la répartition des causes de passage en défaut sur le périmètre sélectionné.

Cela inclut :

* L’identification des motifs de défaut (ex. : retard de paiement, procédure judiciaire, restructuration, etc.) codés en typologies 8, 9 et 10 ;
* Le volume de cas associés à chaque typologie pour chaque année de la période d’étude.

Ce suivi permet de mieux comprendre l’évolution des facteurs de risque associés aux défauts dans le temps.

## **Typologie des prêts en Watchlist et taux de transition vers le défaut (Typology of Loans under Watchlist and Transition Rates to Defaults)**

L’objectif est d’obtenir un historique des raisons de passage en Watchlist (WL), avec des informations permettant d’estimer la probabilité de transition vers un défaut et le temps écoulé entre le passage en WL et l’entrée en défaut.

Les éléments à documenter pour chaque année de la période d’étude :

* Les raisons de passage en WL, classées par typologie ;
* Le nombre de prêts en WL, par typologie et par segment (SME / LC) ;
* Le nombre et pourcentage de prêts en WL qui passent en défaut ;
* Le délai moyen entre WL et défaut (en trimestres).

## **Délai de paiement par millésime d’origine et par tranche de retard (Delinquency by Origination and Bucket)**

L’objectif est d’analyser les retards de paiement (delinquencies) des prêts en les classant :

* Par année de génération (millésime d’origine) ;
* Par tranche de retard (bucket), par exemple : 0–30 jours, 31–60 jours, 61–90 jours et 90+ jours.

Les données permettent de suivre l’évolution de la qualité des portefeuilles au fil des millésimes.

# Processus de production des datapacks

Une image contenant texte, capture d’écran, diagramme, ligne

Description générée automatiquement

Le nouveau processus de production des datapacks est automatisé et simplifié, grâce à une solution **Python** qui intègre :

* **Collecte des données** : automatisation de la collecte des données nécessaires à la production des datapacks, en utilisant les données préparées par SAS.
* **Traitement des données** : traitement des données collectées directement dans l’environnement Python.
* **Calcul des indicateurs** : utilisation de Python pour le calcul des indicateurs, remplaçant ainsi le besoin d'un programme R.
* **Automatisation des gap analyses** : mise en place d’un système pour automatiser l’analyse des écarts entre les transactions N et N-1, facilitant l’évaluation de la performance et des changements au fil du temps.

1. Intégration de la nouvelle segmentation IFRS9, validée par RISQ/DIR et attendue pour être mise en production d'ici fin 2024, après validation par le MRM. [↑](#footnote-ref-1)