

Master Universitario in: "ANALISI DATI PER LA BUSINESS INTELLIGENCE E DATA SCIENCE"  
A.A. 2018-2019

Titolo della tesi: Data Quality per il calcolo della LGD: un approccio operativo a supporto del monitoraggio della qualità del dato.

Autrice: Rita Autieri

## **Abstract**

*Il presente lavoro ha come focus l'individuazione e la diagnostica di problematiche di Data Quality, con l'obiettivo di segnalare le eventuali anomalie riscontrate ai soggetti interessati dell'area Credit Risk Management di un importante istituto bancario, per valutare i potenziali interventi di bonifica.*

*Le valutazioni, in particolare, sono incentrate sull'ambito della Loss Given Default (LGD), uno degli indicatori necessari alla determinazione del patrimonio di vigilanza che le banche sono tenute a detenere (come richiesto dal Regulator).*

*I modelli di Risk Management, necessari al calcolo dei parametri delle varie categorie di rischio, possono essere definiti con diverse modalità (stime esterne e/o interne) e affidarsi a stime interne, puntuali e affidabili diventa un punto di forza e uno strumento di diversificazione che contraddistingue chi è leader nel mercato.*

*L'utilizzo di internal rating (IRB) richiede l'adozione di una strutturazione integrata di risorse umane, tecnologiche e di gestione del dato che diviene una sfida continua e complessa.*

*In questo contesto, la data challenge è focalizzata sulla qualità del dato e su quali strumenti possono essere utilizzati affinché ci si possa affidare al dato per consentire al top management di prendere decisioni con consapevolezza.*

*Si osserverà che il processo del flusso dei dati attraversa varie fasi e trasformazioni fino ad arrivare alla stima della LGD osservata. Le attività di data cleaning, posizionate sui vari livelli dell'architettura del flusso, consentono di controllare la qualità relazionale, l'integrazione e la coerenza, e la qualità andamentale del dato.*

*Unitamente a questi interventi, viene introdotta un'attività di monitoraggio ad hoc a supporto dei sistemi ingegnerizzati e automatizzati, che consente di intercettare in maniera puntuale delle anomalie conosciute e segnalate e/o non conosciute.*

*Il lavoro si è basato sulla conduzione di analisi su 9 tabelle di input al motore di calcolo della LGD che sono pre-trattate in modo da fornire informazioni sulla clientela che versa in stato di default.*

*Le evidenze scaturite hanno individuato che in media, con un controllo suppletivo integrato ai processi, viene intercettato circa il 5% di anomalie non gestite dai sistemi di data cleaning automatizzati.*