

Données Massives :

l'évaluation de leur valeur
économique et l'impact sur
les performances des
entreprises



Projet financé par Grenoble Ecole de Management et la région

Auvergne-Rhône-Alpes

Auteurs : Claudio Vitari et Elisabetta Raguseo



Préface

La philosophie de ce projet

Ce rapport est issu d'un projet de recherche cofinancé par Grenoble École de Management et la région Auvergne-Rhône-Alpes. Cette étude a été réalisée dans le but de comprendre comment les technologies de l'information et de la communication (TIC) peuvent offrir de nouvelles opportunités aux entreprises, en mettant l'accent sur le rôle de Données Massives et des TIC sous-jacentes dans la création de valeur économique. Gartner définit les Données Massives comme les ensembles de données qui deviennent tellement volumineuses, rapides et variées qu'elles en deviennent difficiles à travailler avec des outils classiques de gestion de base de données ou de gestion de l'information et qui exigent donc des formes innovantes de traitement de l'information pour permettre une meilleure compréhension, prise de décision et automatisation des processus.

Sans forcément s'en apercevoir, les Données Massives sont produites autour de nous et elles changent la façon dont les entreprises offrent leurs services, les caractéristiques de leurs produits et la façon dont les professionnels travaillent. Nous sommes au sommet d'une révolution dans le traitement des données qui pourraient bien transformer la façon dont les organisations sont gérées, et aussi transformer les économies et les sociétés dans lesquelles elles opèrent. Pour les entreprises, cette révolution couvre plusieurs dimensions : 1) Les entreprises ont plus que jamais des données à traiter que jamais ; 2) Les entreprises comprennent de mieux en mieux les modèles comportementaux des consommateurs, en utilisant ces données massives internes et externes ; 3) Les entreprises utilisent de nouvelles solutions technologiques pour comprendre leurs propres opérations à un niveau de détail beaucoup plus fin.

En réalisant l'importance croissante de ces Données Massives, il devient nécessaire de comprendre si les entreprises sont prêtes à extraire de la valeur de Données Massives et comment elles extraient effectivement cette valeur.

Table des matières

Préface	1
La philosophie de ce projet	1
Table des matières	2
Liste des figures	3
Liste des tableaux	4
1. Introduction	5
2. Définition de Données Massives	6
3. De l'analyse de données traditionnelles à l'analyse de Données Massives.....	10
4. Les entreprises interrogées et la méthodologie	13
5. La création de valeur par les Données Massives	14
5.1 Modèle conceptuel	14
5.2 Définition de la valeur économique	15
5.3 Définition de la performance de l'entreprise.....	16
5.4 Explorations des tendances en France	17
5.5 De la valeur économique à la performance de l'entreprise	19
6. Groupes d'entreprises	21
7. Conclusions	22
8. Pistes de recherche future.....	23
Références.....	23

Liste des figures

Figure 1 Les 4V des Données Massives	7
Figure 2 Des données non structurées aux données structurées (Source : http://www.datameer.com/product/index.html)	9
Figure 3 La coexistence de solutions pour les Données Massives avec l'entrepôt traditionnel de données d'entreprise ou data warehouse (Source: SAS best practices)	13
Figure 4 Modèle conceptuel du projet	14
Figure 5 Entreprises avec et sans solutions d'analyse de Données Massives	17
Figure 6 La valeur informationnelle des Données Massives (la valeur moyenne est sur une échelle allant de -3 à +3).....	18
Figure 7 La performance des entreprises selon les solutions d'analyses de Données Massives (la valeur moyenne est sur une échelle allant de -3 à +3).....	19
Figure 8 Relation entre valeur économique et performance financière (base de départ : les entreprises qui investissent dans les solutions d'analyse de Données Massives).....	20
Figure 9 Relation entre valeur économique et compétitivité (base de départ : les entreprises qui investissent dans les solutions d'analyse de Données Massives)	21

Liste des tableaux

Tableau 1 Les deux types de Données Massives selon leurs origines.....	10
Tableau 2 Données traditionnelles et Données Massives (Davenport, 2014).....	10
Tableau 3 Les technologies pour les Données Massives (Davenport, 2014).....	12
Tableau 4 Composition de l'échantillon.....	13
Tableau 5 Groupes d'entreprises.....	21

1. Introduction

Les Données Massives sont un moteur du succès commercial dans un large éventail de secteurs industriels. Les organisations investissent des ressources considérables dans les Données Massives en quête d'opportunités de création de valeur, de nouvelles stratégies commerciales numériques et de décisions d'affaires mieux informées. Les Données Massives sont générées tous les jours par, entre autres, des tweets, des clics, des vidéos et la pléthore de capteurs intégrés sur les appareils électroniques. En outre, les instruments de mesure et les équipements industriels comme les compteurs intelligents, les capteurs de fabrication ou les systèmes de suivi des véhicules génèrent automatiquement et continuellement des données en abondance. Par exemple, lorsque les entreprises utilisent des technologies d'identification par radiofréquence (RFID) pour suivre les articles le long de leur chaîne d'approvisionnement, elles produisent des Données Massives. Encore, lorsque les entreprises suivent un lien vers un site Web, elles produisent également des Données Massives. Piccoli et Watson (2008) expliquent comment Caesars-Harrah's Entertainment (le plus grand groupe de casinos aux États-Unis) utilise son système de points de fidélité « Total Rewards », associé à des cartes RFID, pour recueillir des données détaillées sur les comportements de jeu de ses clients. Ce système intégré permet à Caesars-Harrah de surveiller le comportement au fil du temps et d'extraire de la valeur de ces données par l'adaptation de l'expérience de jeu à chaque client.

Au-delà de ces exemplarités, les organisations sont confrontées à d'énormes défis lors de l'accès, le traitement et l'analyse de telles Données Massives. En effet, de nombreuses entreprises sont submergées par d'énormes quantités de données très détaillées. Néanmoins, cette révolution récente axée sur les données peut offrir aux entreprises des occasions pour améliorer et accélérer leurs prises de décisions, si les entreprises arrivent à exploiter leur Données Massives. Les investissements dans le traitement et l'analyse des Données Massives couvrent l'ensemble des technologies qui accèdent, intègrent et élaborent les données disponibles en filtrant, en corrélant et en mettant en évidence des informations qui ne seraient pas accessibles par les technologies traditionnelles de traitement des données.

Sur la base de ces considérations, ce projet portera sur la valeur économique des Données Massives et sur son impact sur la performance des entreprises. Le projet étudiera la valeur économique en termes de valeur transactionnelle, stratégique, transformationnelle et informationnelle pour l'entreprise, ainsi que la performance de l'entreprise en termes de performance financière et de compétitivité.

2. Définition de Données Massives

La façon dont les organisations captent, créent et utilisent des données évolue. L'idée sous-jacente, qui gagne du terrain parmi les cadres, les universitaires et les consultants, est que nous sommes au sommet d'une révolution analytique qui peut bien transformer la façon dont les organisations sont gérées et transformer les économies et les sociétés dans lesquelles elles opèrent (Kiron, 2013). Pour les entreprises, cette révolution couvre plusieurs dimensions : 1) Les entreprises ont plus de données à traiter que jamais ; 2) Les entreprises comprennent de mieux en mieux les modèles comportementaux des consommateurs, en utilisant ces données massives internes et externes ; 3) Les entreprises utilisent de nouvelles solutions technologiques pour comprendre leurs propres opérations à un niveau de détail beaucoup plus fin.

Dans un monde envahi par le numérique comme celui dans lequel nous vivons aujourd'hui, la croissance du volume de données atteint facilement l'ordre de grandeur des téraoctets, voire des pétaoctets. Le terme de Données Massives est utilisé pour décrire ce phénomène. Plus précisément, Gartner définit les Données Massives comme les ensembles de données qui deviennent tellement volumineux, rapides et variées qu'ils deviennent difficiles à travailler avec des outils classiques de gestion de base de données ou de gestion de l'information et qui exigent donc des formes innovantes de traitement de l'information pour permettre une meilleure compréhension, prise de décision et automatisation des processus. De toute évidence, la taille n'est pas la seule nouveauté des Données Massives. D'autres caractéristiques décrivent les Données Massives. Laney (2001) a suggéré que le Volume, la Variété et la Vitesse (les trois V) sont trois dimensions qui peuvent décrire de manière plus complète les Données Massives. Une quatrième caractéristique critique des Données Massives, la Vérité, peut être ajoutée pour leur définition (Figure 1).

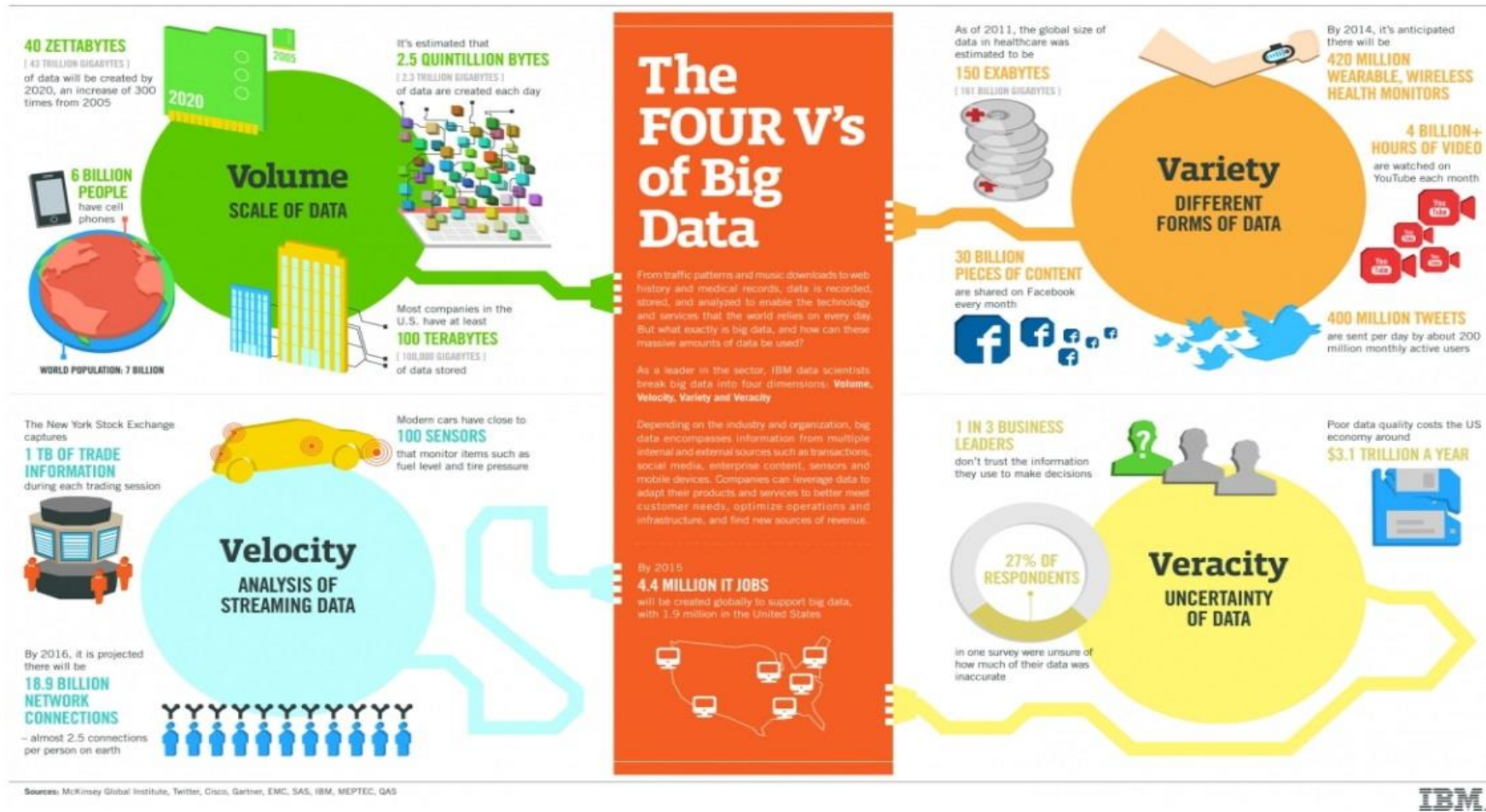


Figure 1 Les 4V des Données Massives

Le Volume indique la taille des Données Massives générées par les machines, les réseaux et les interactions humaines sur des systèmes comme les réseaux sociaux. Les mesures du Volume des Données Massives sont de l'ordre de grandeur des téraoctets et des pétaoctets.

La Vitesse traite du rythme auquel les données circulent à partir de leurs sources telles que les processus métiers, les machines, les réseaux et les interactions humaines, les appareils mobiles, etc. Le flux de données est massif, continu et en temps réel. Par exemple, la prolifération des appareils numériques comme les smartphones et les capteurs a entraîné un taux de création de données exponentiel et stimule un besoin croissant d'analyse en temps réel sur des données. Même des entreprises qui travaillent dans les secteurs plus traditionnels, comme Wal-Mart, traitent plus d'un million de transactions par heure. Enfin, les données générées par les appareils mobiles et les applications mobiles produisent des flux d'informations qui peuvent être utilisées pour générer des offres personnalisées en temps réel pour les clients.

La Variété fait référence aux nombreuses sources et types de données disponibles à la fois structurées et non structurées. Gartner définit les données non structurées comme des contenus qui ne sont pas conformes à un modèle de données prédéfini et qui, en général, sont riches en texte. Toutefois, seules les 5% de toutes les données existantes sont considérées comme structurées (Cukier, 2010). En effet, les données se présentent plus souvent sous forme d'e-mails, photos, clickstreams, vidéos, fichiers PDF, audio et textes qui naissent dans un format non structuré. Ces données non structurées créent la nécessité de solutions technologiques adéquates pour les stocker et les analyser. Grâce aux technologies pour les Données Massives, les données non structurées peuvent devenir des données structurées (voir par exemple la Figure 2). Par exemple, les données des clickstreams fournissent une panoplie d'informations sur le comportement des clients et leurs modèles de navigation aux entreprises commerciales en ligne. Ces informations peuvent être exploitées pour la prise de décisions comme la promotion des produits ou à leur placement. Même les petites et moyennes entreprises, qui sont souvent plus contraintes financièrement que les grandes entreprises, peuvent extraire de la valeur des Données Massives pour améliorer la conception de leurs sites Web, les ventes croisées et les recommandations personnalisées.

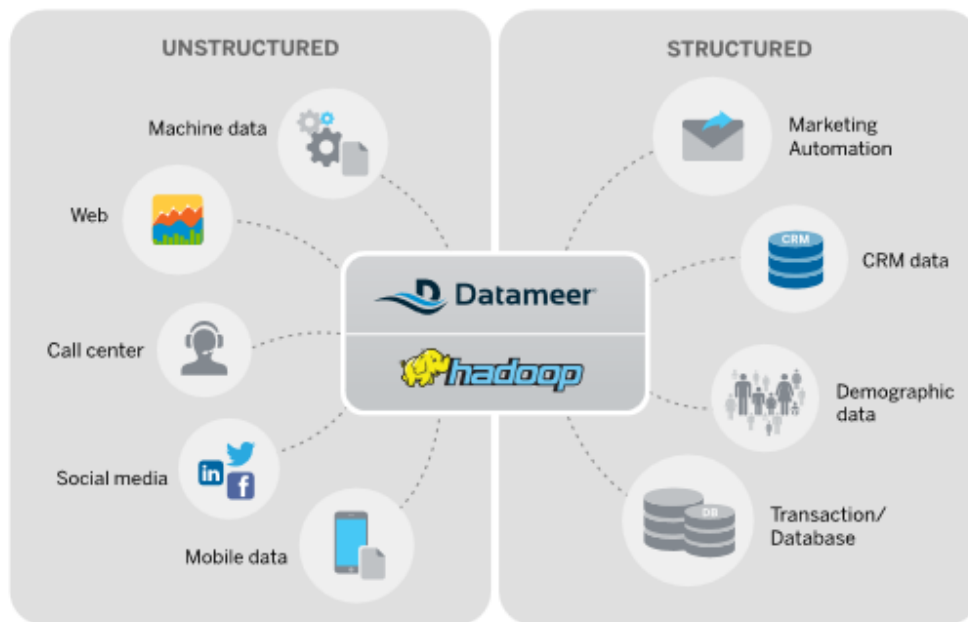


Figure 2 Des données non structurées aux données structurées (Source : <http://www.datameer.com/product/index.html>)

Enfin, la Véracité fait référence à l'incertitude des données en termes de biais, bruits et anomalies. Par exemple, les sentiments des clients dans les médias sociaux sont autant incertains dans leur nature propre que précieux pour de la communication personnalisée. Ainsi, la nécessité de traiter des données incertaines est une autre facette des Données Massives. Cette incertitude est traitée à l'aide d'outils d'analyse développés spécifiquement pour la gestion et l'exploitation des Données Massives.

Les Données Massives peuvent également être classées selon leur origine : Données Massives générées par les machines sans intervention humaine, et Données Massives générées par les humains en interaction avec les ordinateurs. Des exemples sont présentés dans le tableau 1.

Types	Origines
Données Massives générées par les machines	<ul style="list-style-type: none"> • Données des caméras, microphones, etc. • Données des capteurs d'identification par radiofréquences (RFID) • Données des capteurs de surveillance de conditions des marchandises dans la chaîne d'approvisionnement et les variables comme la température de la lumière, l'ouverture d'un colis, etc. (ILC) • Données des compteurs intelligents, des dispositifs médicaux et du

	<p>Système global de géolocalisation (GPS)</p> <ul style="list-style-type: none"> • Données de point de vente par la lecture de codes barres des produits achetés (POS)
Données Massives générées par les humains	<ul style="list-style-type: none"> • Tweets autres micro-blogging • Textes libres, comme livres, articles, corps de messages électroniques, documents, PDF • Articles dans les médias sociaux • Weblog • Données dans des feuille de calcul • Données des clickstream, chaque fois qu'on clique sur un lien sur un site Web • Contenus Web • Voix • Icones et autres espressions de l'humeur

Tableau 1 Les deux types de Données Massives selon leurs origines

3. De l'analyse de données traditionnelles à l'analyse de Données Massives

Les analyses des Données Massives diffèrent des analyses de données traditionnelles pour plusieurs aspects relatifs aux types, volumes, débits de données, méthodes d'analyse et but principal poursuivi (Tableau 2)

	Données traditionnelles	Données Massives
<i>Type de données</i>	Structurées en lignes et colonnes	Non-structurées
<i>Volume de données</i>	Moins de 10 téraoctets	Plus de 10 téraoctets
<i>Débit de données</i>	Données statiques	Flux continu de données
<i>Méthode</i>	Hypothético-déductif	Apprentissage automatique
<i>But</i>	Aide à la prise de décisions internes	Création de produits et services basés sur les données

Tableau 2 Données traditionnelles et Données Massives (Davenport, 2014)

Le problème principal dans l'analyse des Données Massives est l'absence de structure plutôt que leur taille. Par conséquent, des solutions technologiques adéquates et du personnel avec des compétences appropriées sont nécessaires pour réussir dans les tentatives d'analyse des Données Massives. En effet, la technologie, les données et les personnes sont les trois facteurs critiques de succès dans l'analyse des Données Massives (Davenport 2014). En soit, les données, les matériels informatiques et les logiciels sont souvent gratuits ou à bon marché. Ce

qui coûte cher sont les personnes car l'offre sur le marché du travail est très limitée et les compétences difficiles à évaluer.

Certaines des solutions technologiques pour le traitement des Données Massives sont nouvelles, mais d'autres solutions exploitent des technologies déjà affirmées, mais appliquées de façon différente. En effet, les Données Massives ne peuvent pas être traitées avec des bases de données et des systèmes de gestion des bases de données traditionnels et sur des serveurs uniques. Les pétaoctets de données non structurées dans des formats variés ne s'accroissent pas de ces solutions traditionnelles. Par conséquent, plusieurs aspects changent dans la gestion des Données Massives. Tout d'abord, de nouveaux logiciels de traitement de données doivent être utilisés pour gérer les Données Massives, comme Hadoop (tableau 3). La principale raison du succès d'Hadoop est sa capacité à traiter rapidement des données disponibles sur des serveurs différents. Hadoop fractionne une sur plusieurs serveurs et réduit ainsi le temps de traitement. Néanmoins, Hadoop n'est pas la seule solution technologique pour les Données Massives. Les ingénieurs ont su proposer de nouvelles formes de bases de données, des nouveaux langages de programmation de script interactifs tels que Python, Pig et Hive, ou encore des nouvelles architectures. Une autre limite de l'approche traditionnelle à l'analyse des données est liée aux pratiques de séparation des données dans des entrepôts de données distincts avant leur analyse. Toutefois, le volume et la vitesse des Données Massives contrastent avec cette approche de ségrégation des données, qui retarde les analyses et rend ses résultats obsolètes. Ainsi de plus en plus d'organisations utilisent Hadoop et des technologies similaires pour ne stocker que temporairement des grandes quantités de données et les remplacer régulièrement par des données plus récentes.

Technologies	Definition
Hadoop	Logiciel Open Source pour le traitement de Données Massives sur plusieurs serveurs en parallèle
MapReduce	L'architecture sur laquelle Hadoop s'appuie
Langages de script	Langages de programmation adaptés aux Données Massives (Python, Pig, Hive)
Apprentissage automatique	Logiciel pour trouver rapidement le modèle qui convient le mieux à un ensemble de données
Analyses visuelles (Visual analytics)	Affichage des résultats analytiques sous forme graphique ou visuelle

Traitement du langage naturel (Natural language processing)	Logiciels d'analyse de fréquences, de significations, etc.
Analyses en mémoire vive (In-memory analytics)	Traitement de Données Massives dans la mémoire vive de l'ordinateur pour une plus grande vitesse

Tableau 3 Les technologies pour les Données Massives (Davenport, 2014)

Une fois que les technologies pour les Données Massives ont transformé et structuré les données en lignes et colonnes typiques de textes et de nombres, la façon dont ces données sont analysées n'est pas nouvelle. Quand les données sont dans ces formats, les Données Massives peuvent être analysées comme tout autre ensemble de données et les algorithmes statistiques et mathématiques de base pour effectuer l'analyse sont tout à fait les mêmes que ceux employés pour les données traditionnelles. La différence est uniquement dans l'approche. L'approche traditionnelle est fondée sur la méthode hypothético-déductive de l'analyse statistique, impliquant la formulation d'hypothèses et le test de ces hypothèses avec les données disponibles. L'approche pour les Données Massive implique souvent l'apprentissage automatique. L'avantage de cette approche repose sur le fait que l'apprentissage automatique génère rapidement des modèles pour expliquer et prédire les relations dans les données en mouvement rapide. L'inconvénient, au contraire, est qu'il conduit généralement à des résultats qui sont difficiles à interpréter et à expliquer.

Les grandes entreprises suivent généralement une stratégie de coexistence entre l'utilisation de leur entrepôt traditionnel de données et les nouvelles solutions pour les Données Massives (voir la figure 3). La coexistence d'environnements différents réduit au minimum les perturbations des processus analytiques existants tout en ouvrant la possibilité à des processus nouveaux et stratégiques qui pourraient bénéficier d'une vitesse accrue. Cette coexistence comprend: 1) certaines sources de données qui alimentent l'entrepôt de données d'entreprise; 2) certaines solutions pour la structuration et l'exploration des Données Massives en vue de leur inclusion dans l'entrepôt de données d'entreprise; 3) certains traitements pour la transformation des données, l'exploration et la découverte des tendances, sans passer par l'entrepôt de données d'entreprise.

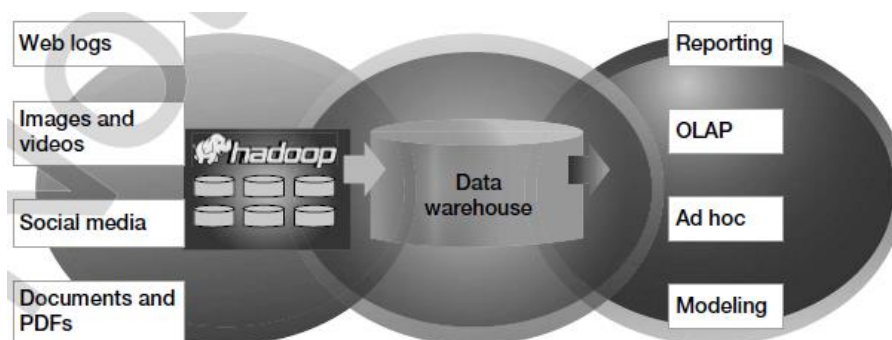


Figure 3 La coexistence de solutions pour les Données Massives avec l'entrepôt traditionnel de données d'entreprise ou data warehouse (Source: SAS best practices)

4. Les entreprises interrogées et la méthodologie

Les données utilisées dans ce projet de recherche sont issues d'une enquête réalisée en 2016 auprès de 200 entreprises implantées en France. L'enquête a été menée sur tous les secteurs économiques. Nous avons interrogé exclusivement les entreprises de taille moyenne (50 et 249 employés) ou grande (avec 250 employés ou plus), car elles sont pionnières dans les investissements en Données Massives. Le tableau 4 présente la composition de l'échantillon divisé par secteur et par taille des entreprises selon le nombre d'employés.

SECTEUR	NOM DU SECTEUR	Taille		
		Grande	Moyenne	Total
A	AGRICULTURE, FORESTERIE ET PÊCHE	0	2	2
B	MINES ET INDUSTRIE EXTRACTIVE	0	0	0
C	FABRICATION	7	60	67
E	APPROVISIONNEMENT EN EAU; ASSAINISSEMENT, GESTION DES DECHETS ET DEPOLLUTION	1	1	2
F	CONSTRUCTION	3	10	13
G	COMMERCE DE GROS ET DE DÉTAIL; RÉPARATION DES VÉHICULES AUTOMOBILES ET DES MOTOCYCLES	5	34	39
H	TRANSPORT ET STOCKAGE	4	9	13
I	HÉBERGEMENT ET ALIMENTATION	2	7	9
J	INFORMATION ET COMMUNICATION	0	6	6
K	FINANCE ET ASSURANCE	0	0	0
L	IMMOBILIER	1	3	4
M	ACTIVITÉS PROFESSIONNELLES, SCIENTIFIQUES ET TECHNIQUES	2	21	23
N	ADMINISTRATION	1	11	12
P	EDUCATION	0	1	1
Q	SANTÉ ET DE TRAVAIL SOCIAL	1	7	8
R	ARTS, DIVERTISSEMENT ET LOISIRS	0	1	1
S	AUTRES	0	0	0
	<i>Total</i>	27	173	200

Tableau 4 Composition de l'échantillon

Les questions de l'enquête adaptaient des questions formulées dans d'autres études de recherche à fin de garantir la fiabilité des réponses. Le questionnaire a été testé sur un sous-échantillon initial de 30 entreprises françaises. En général, un premier e-mail a été envoyé afin de sensibiliser les entreprises au projet. Ensuite, l'enquête a été effectuée via un entretien téléphonique assisté par ordinateur. Nous avons ciblé le directeur des Systèmes d'Information ou Informatique et, au cas où il n'était pas disponible, l'intervieweur a demandé d'interviewer le directeur général, le propriétaire ou une autre personne responsable du département informatique. Enfin, les données recueillies ont été analysées avec le logiciel statistique SPSS.

5. La création de valeur par les Données Massives

5.1 Modèle conceptuel

Afin d'étudier la relation entre la valeur économique apportée par les Données Massives au niveau de l'entreprise et de la performance de l'entreprise, le modèle conceptuel présenté dans la figure 4 a été appliqué dans ce projet. Plus précisément, deux blocs principaux constituent le modèle conceptuel à la base de ce projet :

1. Le premier bloc comprend les dimensions de la valeur économique liées à l'exploitation de Données Massives au niveau de l'entreprise: la valeur transactionnelle, la valeur stratégique, la valeur transformationnelle et la valeur informationnelle.
2. Le deuxième bloc comprend les dimensions de la performance des entreprises: la performance financière et la compétitivité.

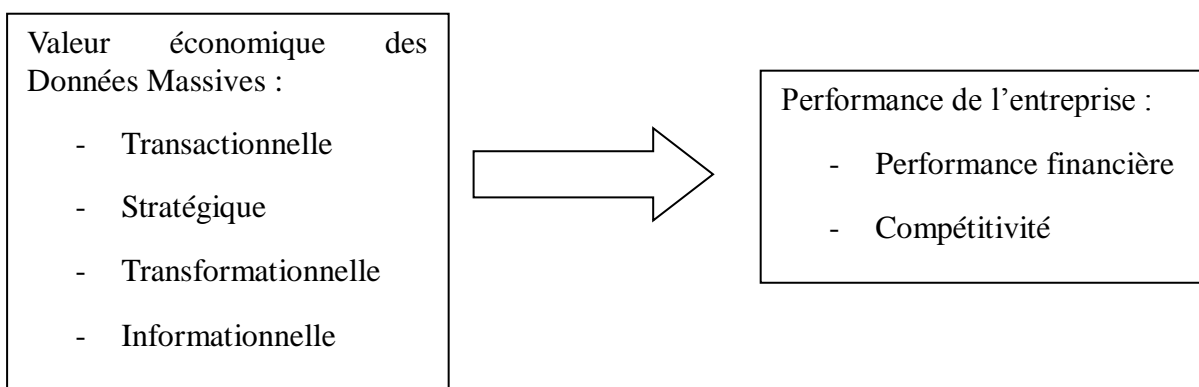


Figure 4 Modèle conceptuel du projet

5.2 Définition de la valeur économique

Les Données Massives peuvent fournir de la valeur économique aux entreprises de différentes façons. La valeur économique peut couvrir quatre dimensions, toutes étudiées dans ce projet. La première dimension se réfère à la «valeur transactionnelle», qui se concentre sur l'amélioration de l'efficacité et la réduction des coûts. Plus précisément, les entreprises ont été invitées à évaluer si les solutions d'analyse des Données Massives ont déterminé les avantages suivants :

1. Des économies dans la logistique ;
2. Des réductions des coûts opérationnels ;
3. Des baisses des coûts de communication ;
4. Des réductions du nombre des employés ;
5. Des augmentations des retours sur les investissements ;
6. Des augmentations de la productivité du personnel.

La deuxième dimension se réfère à la «valeur stratégique» qui a trait à l'obtention d'avantages concurrentiels à partir de l'exploitation des solutions d'analyse des Données Massives. Plus précisément, les entreprises ont été invitées à évaluer si les solutions d'analyse des Données Massives ont déterminé les avantages suivants :

1. Des élargissements de l'avantage concurrentiel ;
2. Des améliorations de l'alignement du Système d'information avec la stratégie d'entreprise ;
3. Des renforcements des liens avec d'autres organisations ;
4. Des accélérations des réponses aux changements ;
5. Des améliorations des relations avec les clients ;
6. Des améliorations des produits et services.

La troisième dimension se réfère à la «valeur transformationnelle» qui concerne la façon dont les Données Massives sont capables de transformer différents aspects de l'entreprise. Plus

précisément, au moyen du questionnaire, les entreprises ont été invitées à évaluer si les solutions d'analyse de Données Massives ont déterminé les avantages suivants :

1. Des renforcements des compétences du personnel ;
2. Des nouvelles opportunités d'affaires ;
3. Des élargissements des capacités ;
4. Des renforcements des modèles d'affaires ;
5. Des améliorations de l'organisation.

La quatrième et dernière dimension se réfère à la «valeur informationnelle» qui éclaire l'impact des solutions d'analyse des Données Massives sur la prise de décision en temps réel. Plus précisément, les entreprises ont été invitées à évaluer si les solutions d'analyse des Données Massives ont déterminé les avantages suivants :

1. Des accélérations dans l'accès aux données ;
2. Des simplifications dans l'accès aux données ;
3. Des améliorations dans la gestion des données ;
4. Des renforcements dans la précision des données ;
5. Des améliorations dans la mise en forme des données.

Par conséquent, grâce au questionnaire, la valeur économique des solutions d'analyse des Données Massives a été étudiée sur quatre dimensions complémentaires.

5.3 Définition de la performance de l'entreprise

A travers, la valeur économique générée par les Données Massives, les entreprises peuvent être en mesure d'obtenir de meilleures performances économiques. Afin d'évaluer cette relation, ce projet examine deux dimensions de la performance. La première est la « Compétitivité » qui se réfère à la capacité de l'entreprise d'acquérir et de conserver des clients. Plus précisément, les entreprises ont été invitées à évaluer si les solutions d'analyse des Données Massives ont déterminé les avantages suivants :

1. Une rapidité à entrer sur des nouveaux marchés supérieure à celle des concurrents ;

2. Une rapidité à introduire des nouveaux produits ou services sur le marché supérieure à celle des concurrents ;
3. Un taux de succès des produits et des services supérieur à celui des concurrents ;
4. Une part de marché supérieure à celle des concurrents.

La deuxième est la « Performance financière » qui se réfère à la capacité de l'entreprise d'améliorer les ventes, la rentabilité et le retour sur investissement. Plus précisément, les entreprises ont été invitées à évaluer si les solutions d'analyse des Données Massives ont déterminé les avantages suivants :

1. Des accroissements de la rétention des clients ;
2. Des accélérations dans la croissance des ventes ;
3. Des augmentations de la rentabilité.

5.4 Explorations des tendances en France

Sur les 200 entreprises sondées, 76 entreprises ont déclaré disposer de solutions d'analyse de Données Massives tandis que les autres n'ont pas investi dans ce type de solutions (figure 5).

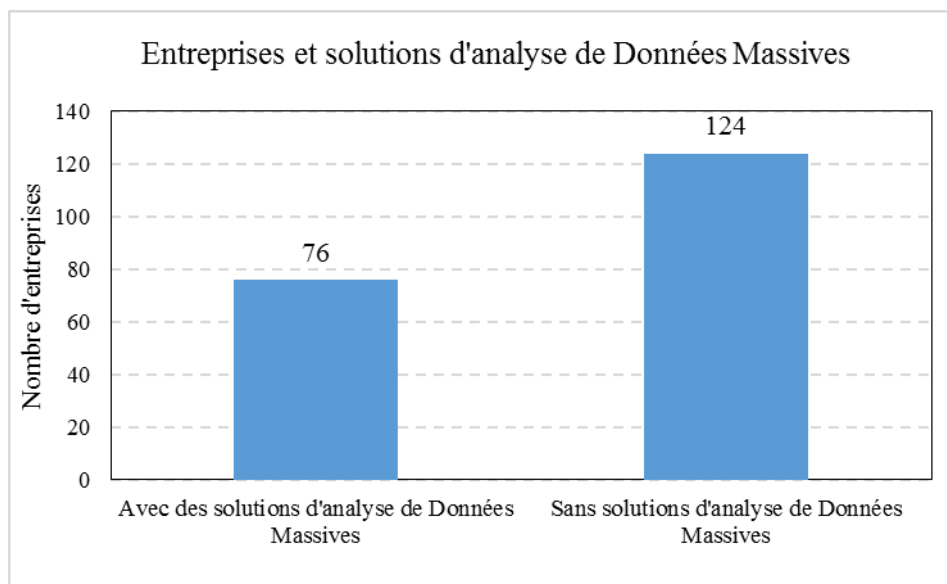


Figure 5 Entreprises avec et sans solutions d'analyse de Données Massives

La figure 6 montre comment les solutions d'analyse de Données Massives apportent de la valeur aux entreprises qui ont déjà investi dans ces solutions et comment les solutions d'analyse de Données Massives devraient apporter de la valeur aux entreprises qui n'ont pas

encore investi et l'opinion moyenne globale de toutes les entreprises confondues. Les résultats montrent que les sociétés qui ont investi dans les solutions d'analyse de Données Massives ont créé de valeur économique à un niveau supérieur que les attentes de création de valeur des entreprises qui n'ont pas investi dans ces solutions. Cela signifie que les entreprises qui n'investissent pas dans ces solutions sont sceptiques à propos de leurs potentialités et que leurs attentes sont inférieures par rapport à la valeur pratiquement atteignable.

En décomposant la valeur économique dans ses quatre dimensions, les résultats montrent que la valeur plus importante observée par les entreprises avec les solutions d'analyse de Données Massives est la « valeur informationnelle », suivie par la « valeur stratégique », la « valeur transformationnelle » et enfin la « valeur transactionnelle ». La prépondérance de la valeur informationnelle signifie que les entreprises bénéficient d'un accès plus rapide et facile aux données, l'amélioration de la gestion et de l'exactitude des données et la mise à disposition de données dans des formats plus utilisables.

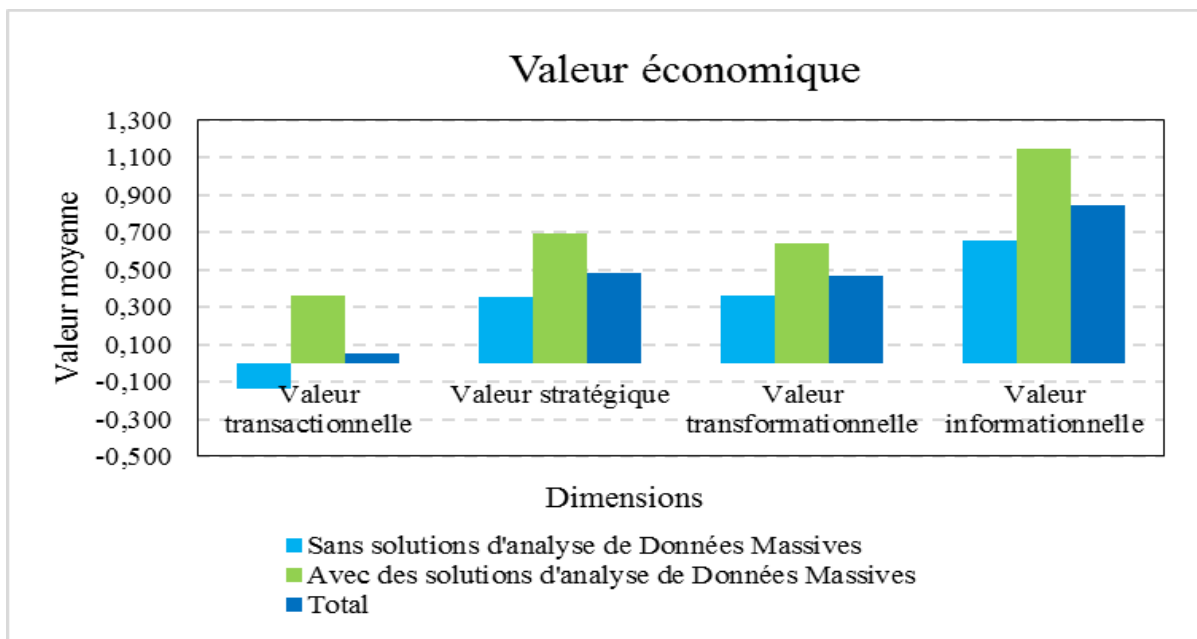


Figure 6 La valeur informationnelle des Données Massives (la valeur moyenne est sur une échelle allant de -3 à +3)

La figure 7 illustre la façon dont les solutions d'analyse de Données Massives mènent à une meilleure performance pour les entreprises qui ont investi dans ces solutions, les attentes en termes de performance pour les entreprises qui n'ont pas investi dans ces solutions et la moyenne globale. Les données montrent que les sociétés qui ont investi dans les solutions d'analyse de Données Massives ont une meilleure performance par rapport à la performance

attendue des entreprises qui n'ont pas investi dans ces solutions. Dans ce cas aussi, comme pour la valeur économique, les entreprises qui n'investissent pas dans ces solutions sont sceptiques quant aux potentialités des solutions d'analyse de Données Massives et leurs attentes sont inférieures aux avantages réels atteints par les entreprises ayant investi dans les solutions d'analyse de Données Massives.

En détaillant performance financière et compétitivité, les résultats montrent que les solutions d'analyse des Données Massives apportent avant tout une meilleure performance financière et dans une moindre mesure de la compétitivité. Cela signifie que les investissements dans les solutions d'analyse des Données Massives sont un moyen d'accroître la fidélisation de la clientèle, d'atteindre une croissance des ventes et des niveaux élevés de rentabilité, d'entrer sur de nouveaux marchés plus rapidement que les concurrents, d'introduire de nouveaux produits ou services sur le marché plus rapidement que les concurrents, un taux de réussite plus élevé pour de nouveaux produits ou services et une plus grande part de marché par rapport aux concurrents.

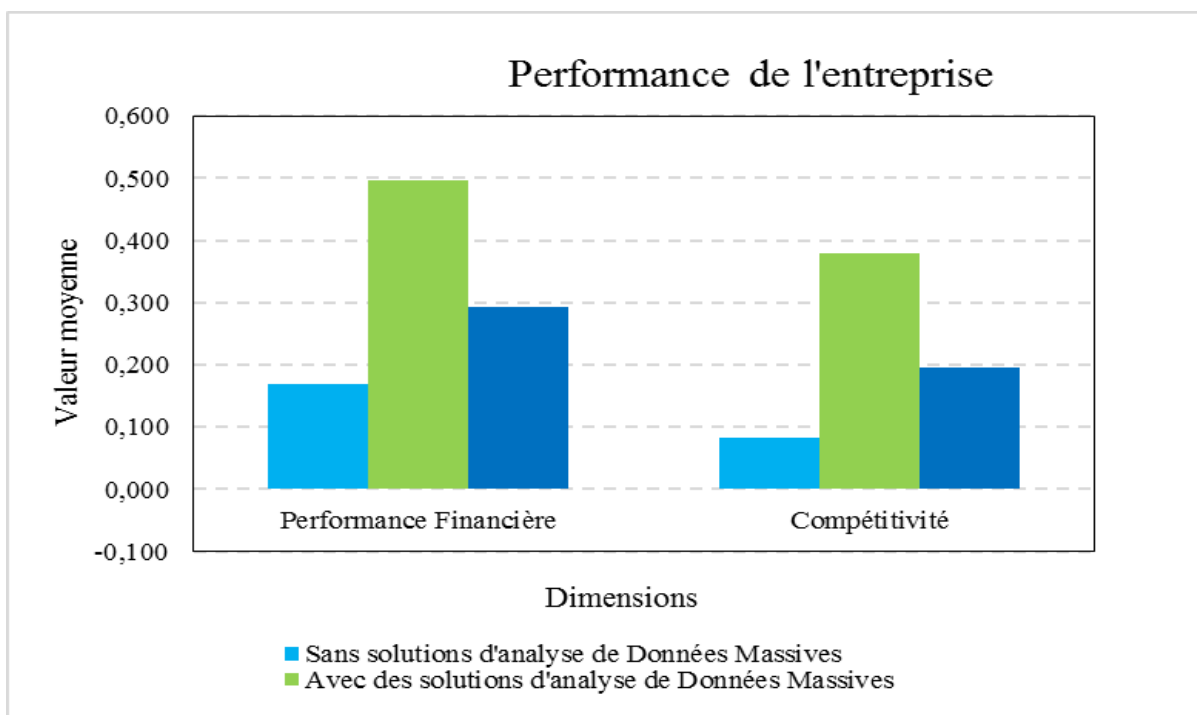


Figure 7 La performance des entreprises selon les solutions d'analyses de Données Massives (la valeur moyenne est sur une échelle allant de -3 à +3)

5.5 De la valeur économique à la performance de l'entreprise

Expliquer la variation dans le degré de succès des organisations est un problème permanent dans la gestion. C'est la raison pour laquelle nous avons examiné si la valeur économique

générée par les Données Massives permet d'atteindre une meilleure performance. A ce propos, les figures 8 et 9 montrent que les entreprises capables d'atteindre des niveaux plus élevés de valeur économique sont également en mesure d'atteindre des niveaux plus élevés de performance financière et de compétitivité. Cela signifie que les entreprises qui utilisent les solutions d'analyse de Données Massives économisent de l'argent d'une part et augmentent le rendement des actifs, d'autre part, ce qui mène à des meilleures performances globales.

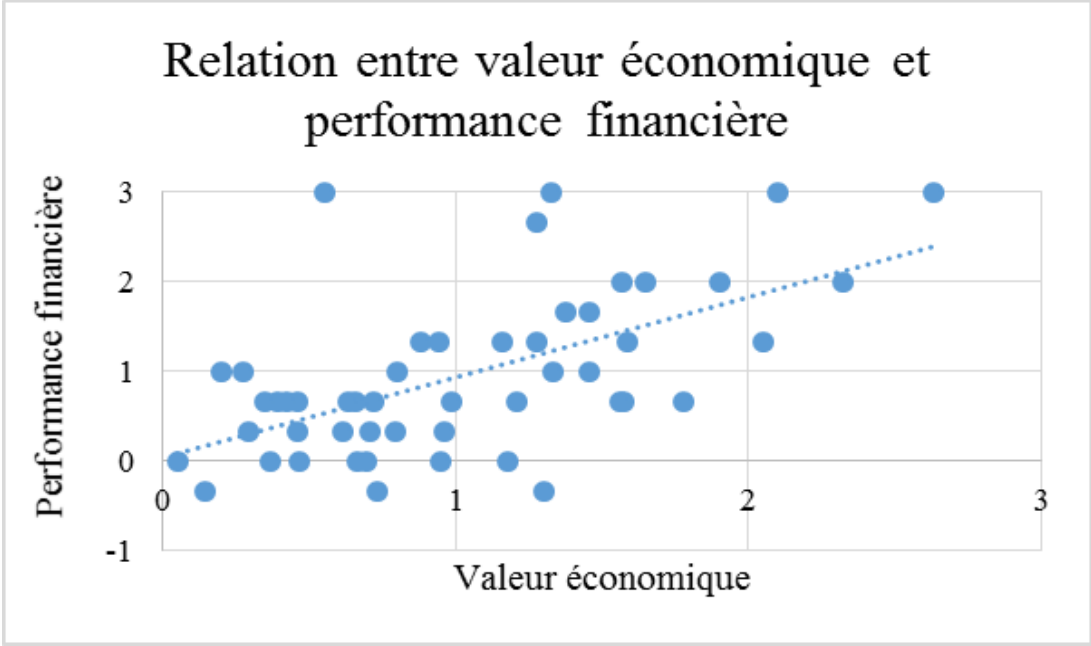


Figure 8 Relation entre valeur économique et performance financière (base de départ : les entreprises qui investissent dans les solutions d'analyse de Données Massives)

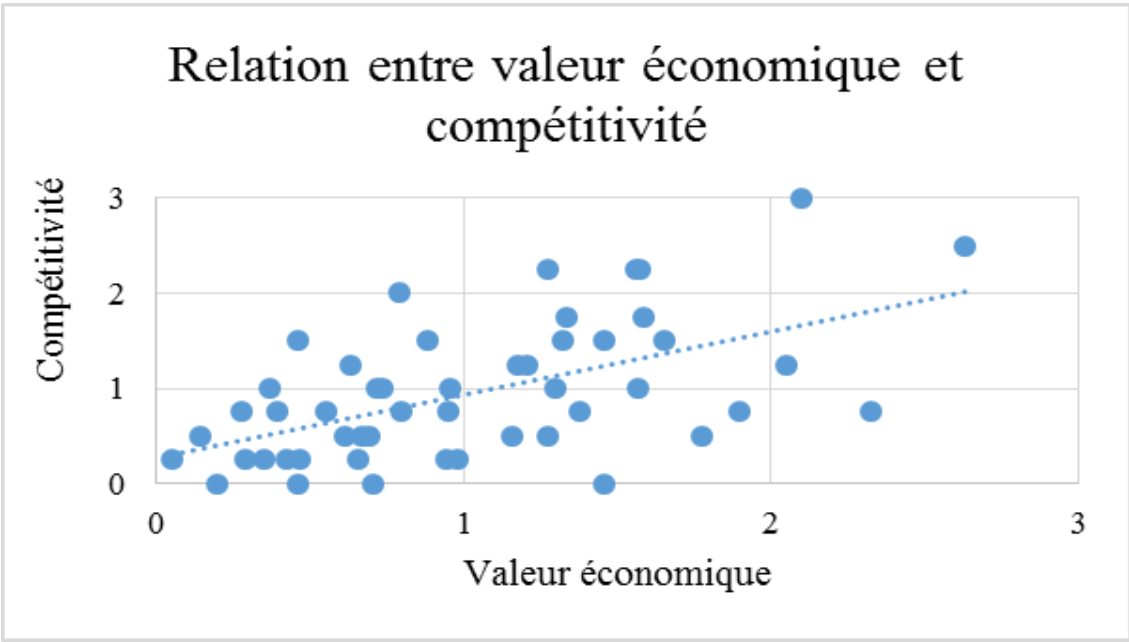


Figure 9 Relation entre valeur économique et compétitivité (base de départ : les entreprises qui investissent dans les solutions d'analyse de Données Massives)

6. Groupes d'entreprises

En étudiant les entreprises qui ont investi dans les solutions d'analyse de Données Massives, nous constatons que certaines d'entre elles excellent en termes de performance. Dans l'ensemble, nous identifions quatre types d'entreprises (voir le tableau 5).




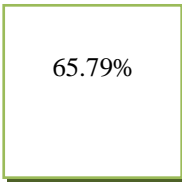
Performance de l'entreprise	Faible	Élevée
Valeur économique		
Sans solutions d'analyse de Données Massives	<i>Pas applicable</i>	<i>Pas applicable</i>
Solutions d'analyse de Données Massives rudimentaires	<i>Type A (26 entreprises)</i>  34.21%	<i>Type B (0 entreprises)</i>  0.00%
Solutions d'analyse de Données Massives sophistiquées	<i>Type C (0 entreprises)</i>  0.00%	<i>Type D (50 entreprises)</i>  65.79%

Tableau 5 Groupes d'entreprises

Type A

Les entreprises de «Type A» ont obtenu une faible valeur économique par les solutions d'analyse de Données Massives et des faibles performances. Ces entreprises doivent comprendre la valeur cachée des investissements dans les solutions d'analyse de Données Massives en termes de valeur stratégique, transactionnelle, transformationnelle et informationnelle. En réorientant leurs investissements dans les solutions d'analyse de Données Massives, ces entreprises pourraient améliorer leurs performances financières et leur compétitivité.

Type B

Le «Type B» inclurait les entreprises avec une faible valeur économique, mais des performances élevées. Comme on peut s'attendre, nous n'avons pas trouvé d'entreprise dans ce groupe, car il est improbable que les entreprises produisant une faible valeur économique soient capables d'atteindre une performance d'entreprise élevée.

Type C

Le «Type C» inclurait les entreprises avec des niveaux élevés de valeur économique et une faible performance. Ce groupe d'entreprises est aussi vide dans notre échantillon. Cela signifie qu'aucune entreprise n'a réalisé beaucoup de valeur économique sans être capable de la convertir en performance financière et compétitivité.

Type D

Les entreprises de «Type D» ont obtenu des niveaux élevés de valeur économique et des niveaux élevés de performance d'entreprise. Ces entreprises profitent complètement des avantages de leurs investissements dans les solutions d'analyse de Données Massives.

7. Conclusions

À mesure que les entreprises s'appuient de plus en plus sur les données pour mener leurs activités, les changements dans la disponibilité de ces données influencent la capacité des entreprises à créer de la valeur sous la forme de nouveaux produits, services ou processus. L'émergence de Données Massives crée des opportunités stratégiques pour les entreprises existantes et pour des nouvelles start-up.

La nouveauté des Données Massives exige de la recherche scientifique pour comprendre mieux ce phénomène. Dans ce projet nous nous sommes concentrés sur la compréhension de la valeur économique des Données Massives pour les entreprises qui ont investi dans les solutions d'analyse des Données Massives et les attentes économiques des entreprises qui n'ont pas investi dans ces solutions. De manière complémentaire, nous avons étudié la performance des entreprises en termes de performance financière et compétitivité. Les résultats montrent que les entreprises qui n'ont pas investi dans les Données Massives semblent sous-estimer le potentiel des Données Massives.

Ainsi, nous concluons que les dirigeants des entreprises ont besoin d'être informés sur les potentialités des Données Massives. Une fois conscients, à l'instar des entreprises qui se sont

déjà lancées dans les solutions d'analyse de Données Massives, ces dirigeants devraient être capables d'exploiter leur potentiel.

8. Pistes de recherche future

Des nouvelles recherches pourraient être menées de manière longitudinale pour comprendre la dynamique du phénomène d'utilisation de solutions d'analyse de Données Massives. Des données complémentaires provenant d'autres sources, telles que des entretiens semi-directifs, pourraient aussi être utilisées pour enrichir la compréhension de la dynamique derrière ces investissements.

Références

Cukier, K. (2010). Data, data everywhere: A special report on managing information. Economist Newspaper.

Davenport, T., Davenport, T. H., & Horváth, P. (2014). Big data@ work. Harvard Business Review Press, Boston.

Kiron, D., & Bean, R. (2013). Organizational alignment is key to big data success. MIT Sloan Management Review, 54(3), 1-7.

Laney, D. (2001). 3D data management: Controlling data volume, velocity and variety. META Group Research Note, 6, 70.

Piccoli, G., Watson, R.T. (2008). Profit From Customer Data by Identifying Strategic Opportunities and Adopting the “Born Digital” Approach. MIS Quarterly Executive, 7(3), 113-122.