

## Exploration des données

CATEGORIE : C

### Vue d'ensemble

Domaine(s) d'activité professionnel dans lequel(s) est utilisé la certification :

Transverse : ■ **Tous domaines concernés**

Code(s) NAF : **58.29A**, **49.10Z**, **64.11Z**, **65.11Z**,  
**61.10Z**

Code(s) NSF : **320**, **313**

Code(s) ROME : **M1403**, **M1804**, **C1102**, **C1101**, **M1701**

Formacode : **31054**

Date de création de la certification : **01/09/2015**

Mots clés : **QUALITE**, **Big data**, **exploration de données**,  
**Analyse de données**

### Identification

Identifiant : **2964**

Version du : **04/09/2017**

### Références

Consensus, reconnaissance ou recommandation :

Formalisé :

■ **Néant**

Non formalisé :

- [Gestion des risques dans l'exploration des données](#)
- [L'analyse de données, un enjeu capital](#)

### Descriptif

#### Objectifs de l'habilitation/certification

Aujourd'hui, les entreprises font face aux phénomènes d'explosion des données dues principalement à la multiplication des réseaux sociaux et à la prolifération des objets connectés. Les entreprises ont compris l'importance d'exploiter cette masse de données par une meilleure compréhension des clients permettant ainsi d'en retirer des avantages concurrentiels. Cela permet également de mieux prévenir et gérer le risque, de réduire les coûts ainsi que de générer des revenus additionnels.

L'exploration de données permet la mise en perspective des objectifs métiers au regard des données mises à disposition. Cela répond au besoin potentiel de détermination du périmètre de données et de sa profondeur d'analyse en ayant au préalable réalisé le travail de nettoyage et de mise en cohérence des données.

#### Lien avec les certifications professionnelles ou les CQP enregistrés au RNCP

■ Néant

#### Descriptif général des compétences constituant la certification

#### Définition des enjeux et objectifs relatifs aux travaux d'exploration des données.

Recueillir les informations de fonds sur la situation actuelle de l'entreprise pour appréhender ses enjeux.

Elaborer des organigrammes pour apprécier l'interdépendance entre les métiers.

Analyser l'environnement de la problématique posée pour la positionner dans un écosystème global.

Définir et valider la problématique pour identifier les moyens et ressources à mobiliser.

### Public visé par la certification

Tous publics

Définir les conditions préalables du projet d'exploration de données pour fixer ses facteurs clés de succès.  
Décrire la solution déjà utilisée pour adresser la problématique actuelle de l'entreprise.  
Mesurer le degré d'efficacité pour adapter les techniques d'exploration de données à proposer.  
Préciser les objectifs financiers pour mieux les décliner dans les techniques d'analyses.  
Réaliser un inventaire des sources de données et des connaissances disponibles pour en permettre l'analyse.  
Établir une cartographie des risques liés à l'exploration des données pour discerner celles qui peuvent être utilisées en respectant les restrictions d'ordre juridiques ou liées à la sécurité.

## **Conception du plan projet.**

Intégrer les objectifs préalablement fixés pour valider les sources de données.  
Construire la feuille de route pour planifier les axes d'exploration des données.  
Choisir et positionner les sources de données pour permettre la conception de la feuille de route.  
Produire un plan projet pour apporter les conditions optimales à sa conduite.  
Définir un plan de communication du projet pour sensibiliser et engager les différentes parties prenantes.

## **Identification et qualification des données.**

Classer les données préalablement listées en distinguant sources internes et externes pour qualifier leurs fiabilités en fonction de leurs typologies.  
Créer les liens entre les données mises à disposition et les différents métiers pour définir le périmètre de l'analyse.  
Appliquer les procédés statistiques pour dénombrer le contenu de l'échantillon des données.  
Formaliser un rapport de description des données pour restituer un état des lieux sur la qualité et la quantité des données.  
Déterminer les corrélations entre les données pour améliorer les enjeux des métiers existants.

## *Modalités générales*

*La formation se dispense en continue ou en alternance, en intra ou en inter.*

*3 modules de formation entre 7 et 10 jours. (49H à 70 H)*

*Module 1 : Définition des enjeux, objectifs et retour sur investissement des travaux d'exploration des données*

*Module 2 : Conception du plan projet*

*Module 3 : Identification et qualification des données*

## *Liens avec le développement durable*

niveau 1 : Certifications et métiers qui internalisent le développement durable. Les activités et compétences mobilisées mettent en oeuvre des matériaux et produits moins polluants

## Pour l'individu

L'individu bénéficie d'un apport en compétences lui permettant d'analyser à partir de données issues d'informations récoltées, une situation complète de l'entreprise cliente et de son environnement. Il sera à même de faire la synthèse des résultats et d'établir des rapports et présenter les conclusions. Ces compétences sont des facteurs clefs de son employabilité d'autant que ces offres en compétences manquent sur le marché au regard du besoin pourtant croissant.

L'individu pourra mettre en avant sa capacité d'appréciation de la qualité des données.

## Pour l'entité utilisatrice

Pour l'entreprise, la certification « Exploration des données », professionnalise l'activité couvrant l'exploration de données, l'analyse et les conclusions. Cette certification permet aux entreprises d'avoir des collaborateurs aptes à créer un processus allant de la recherche d'informations à l'identification du périmètre des données, en passant par l'analyse et la communication liée aux conclusions, tout en restant dans un cadre légal. D'autre part, les collaborateurs pourront constituer une équipe assurant les travaux d'analyse du projet et de transformation du patrimoine des données en informations utiles pour l'entreprise. Il s'agit là, d'un processus global bénéfique à l'entreprise car elle se retrouve ainsi dotée de moyens plus efficaces. Face à des données de plus en plus massives, la démarche standardisée facilite l'analyse et apporte une valeur concurrentielle réelle.

## Evaluation / certification

### Pré-requis

Aucun prérequis particulier mais une connaissance en informatique est favorable

### Compétences évaluées

#### **Définition des enjeux et objectifs et retour sur investissement des travaux d'exploration de données.**

#### **Recueil des informations sur l'entreprise cible.**

Recueillir les informations de fonds sur la situation actuelle de l'entreprise pour appréhender ses enjeux.

Elaborer les organigrammes pour apprécier l'interdépendance entre les métiers.

Analyser l'environnement de la problématique pour la positionner dans un écosystème global.

#### **Définition de la problématique de l'entreprise cible.**

Définir et valider la problématique pour identifier les moyens et ressources à mobiliser.

Définir les conditions préalables du projet d'exploration des données pour fixer ses clés du succès.

Décrire la solution déjà utilisée pour adresser la problématique actuelle de l'entreprise.

Mesurer le degré d'efficacité de cette solution dans l'entreprise pour adapter les techniques d'exploration des données à proposer.

#### **Analyse des sources de données, des risques, coûts et bénéfices du projet.**

Réaliser un inventaire des sources de données et des connaissances disponibles pour en permettre l'analyse.

### Centre(s) de passage/certification

- Ecole Supérieure d'Assurances 13 rue Fernand Léger - 75020 Paris

Etablir une cartographie des risques liés à l'exploration des données en veillant à respecter les restrictions d'ordre juridiques ou liées à la sécurité.

Comparer les coûts du projet d'exploitation des données aux bénéfices attendus en exploitant les données disponibles et hypothèses prises pour analyser le coût-bénéfice du projet.

## **Détermination des objectifs du projet et du retour sur investissement.**

Préciser les objectifs financiers spécifiques définis par les décideurs de l'entreprise pour mieux les décliner dans les techniques d'analyses.

Définir les critères de succès au niveau financier, à utiliser en matière d'exploration des données pour permettre leur mesure.

Calculer le retour sur investissement pour prouver la pertinence et l'intérêt des travaux à réaliser.

## **Conception du plan projet.**

### **Conception de la feuille de route.**

Intégrer les objectifs fixés pour valider les sources de données en utiliser.

Construire la feuille de route en tenant compte de l'écosystème de l'entreprise pour planifier les axes d'exploitation des données.

Apprécier les méthodes d'évaluation des modèles d'analyse des données à utiliser pour choisir celle qui sera la plus appropriée.

Choisir et positionner les sources pour permettre la conception de la feuille de route.

## **Production du Plan Projet, exploitation des moyens mis en place pour le lancement du projet.**

Produire un plan projet incluant l'ensemble des éléments pour apporter les conditions optimales à sa conduite.

Elaborer le budget prévisionnel du projet pour créer les jalons financiers utiles à son pilotage.

Définir un plan de communication du projet en utilisant les outils à mobiliser pour sensibiliser et engager les différentes parties prenantes.

Constituer l'équipe à mobiliser pour assurer les travaux d'analyse du projet.

## **Identification et qualification des données.**

### **Identification du périmètre des données.**

Classer les données en distinguant les sources internes et externes pour qualifier leurs fiabilités en fonction de leurs typologies.

Appréhender les enjeux de chaque métier et des cas d'usage associés pour intégrer les objectifs à viser et les besoins à satisfaire.

Créer les liens entre les données mises à disposition et les différents métiers concernés pour définir le périmètre de l'analyse.

Exploiter la connaissance sur les données détenue par leurs propriétaires et gestionnaires pour les enrichir et mieux les interpréter.

### **Qualification des données.**

Appliquer les procédés statistiques pour dénombrer le contenu de l'échantillon.

Appliquer les techniques d'évaluation de la qualité des données pour apprécier celles de l'échantillon et y accorder un niveau de confiance.

Analyser les écarts constatés entre les résultats intermédiaires obtenus et les objectifs associés pour évaluer la nécessité de compléter ou non l'analyse, en y attribuant des moyens supplémentaires.

Formaliser un rapport de description des données incluant des représentations graphiques pour restituer un état des lieux sur la qualité et la quantité des données.

Communiquer sur le rapport d'analyse de données pour améliorer sa gouvernance.

Déterminer les corrélations entre les données pour améliorer les enjeux des métiers existants.

**Recommandation d'axes d'exploitation future des données.**

*Niveaux délivrés le cas échéant (hors nomenclature des niveaux de formation de 1969)*

Aucun

---

La validité est Permanente

**Possibilité de certification partielle :** non

Matérialisation officielle de la certification :

Certificat de compétences

## Plus d'informations

### *Statistiques*

120 stagiaires depuis 2015

### *Autres sources d'information*

[www.esassurances.com](http://www.esassurances.com)