

# Référentiel de compétences du

B.U.T. *Statistique et informatique décisionnelle*

Parcours *Science des données : exploration et modélisation  
statistique*

Une *compétence* est un « *savoir-agir complexe*, prenant appui sur la mobilisation et la combinaison efficaces d'une variété de ressources à l'intérieur d'une famille de situations » (Tardif, 2006). Les ressources désignent ici les savoirs, savoir-faire et savoir-être dont dispose un individu et qui lui permettent de mettre en œuvre la compétence.

|  |  |   |
|--|--|---|
| <b>Traiter des données à des fins décisionnelles</b>           | <b>Traiter des données à des fins décisionnelles</b>           | <ul style="list-style-type: none"><li>en intervenant à toutes les étapes du cycle de vie de la donnée (insertion, modification, extraction, suppression)</li><li>en utilisant le modèle de données adapté aux besoins</li><li>en s'inscrivant dans une démarche de documentation des réalisations adaptée au public visé</li><li>en traduisant correctement les demandes métier en programmes, avec le respect du cahier des charges s'il existe</li><li>en écrivant un programme correctement structuré et documenté, respectant les bonnes pratiques</li><li>en identifiant les librairies et langages dédiés</li></ul> |
| <b>Analyser statistiquement les données</b>                    | <b>Analyser statistiquement les données</b>                    | <ul style="list-style-type: none"><li>en tenant compte du contexte de l'étude (économique, socio-démographique, commerciale, clinique...)</li><li>en mettant en évidence les grandes tendances et les informations principales</li><li>en identifiant et en mettant en œuvre les techniques adaptées aux attentes du client ou de l'instance décisionnaire</li><li>en identifiant et en mettant en œuvre les techniques adaptées aux données complexes (données massives, données mal structurées, flux de données...)</li><li>en tenant compte du contexte inférentiel (variabilité de l'échantillon)</li></ul>          |
| <b>Valoriser une production dans un contexte professionnel</b> | <b>Valoriser une production dans un contexte professionnel</b> | <ul style="list-style-type: none"><li>en s'adaptant au niveau d'expertise, à la culture et au statut du destinataire</li><li>en s'exprimant correctement, aussi bien en français qu'en anglais, à l'oral comme à l'écrit</li><li>en veillant aux aspects éthiques, déontologiques et réglementaires d'utilisation et de diffusion des données</li><li>en interprétant et contextualisant les résultats (citations, vérification des sources, esprit critique)</li><li>en utilisant la forme de restitution adaptée</li><li>en tenant compte des réalités économiques et managériales des entreprises</li></ul>            |
| <b>Modéliser les données dans un cadre statistique</b>         | <b>Modéliser les données dans un cadre statistique</b>         | <ul style="list-style-type: none"><li>en choisissant le modèle adapté à la situation</li><li>en maîtrisant la qualité du modèle</li><li>en s'adaptant aux spécificités (données, enjeux, méthodes) d'un domaine d'application particulier (santé, marketing, assurance, qualité, socio-démographie...)</li><li>en s'adaptant à la complexité des données (données massives, données mal structurées, flux de données...)</li></ul>  |

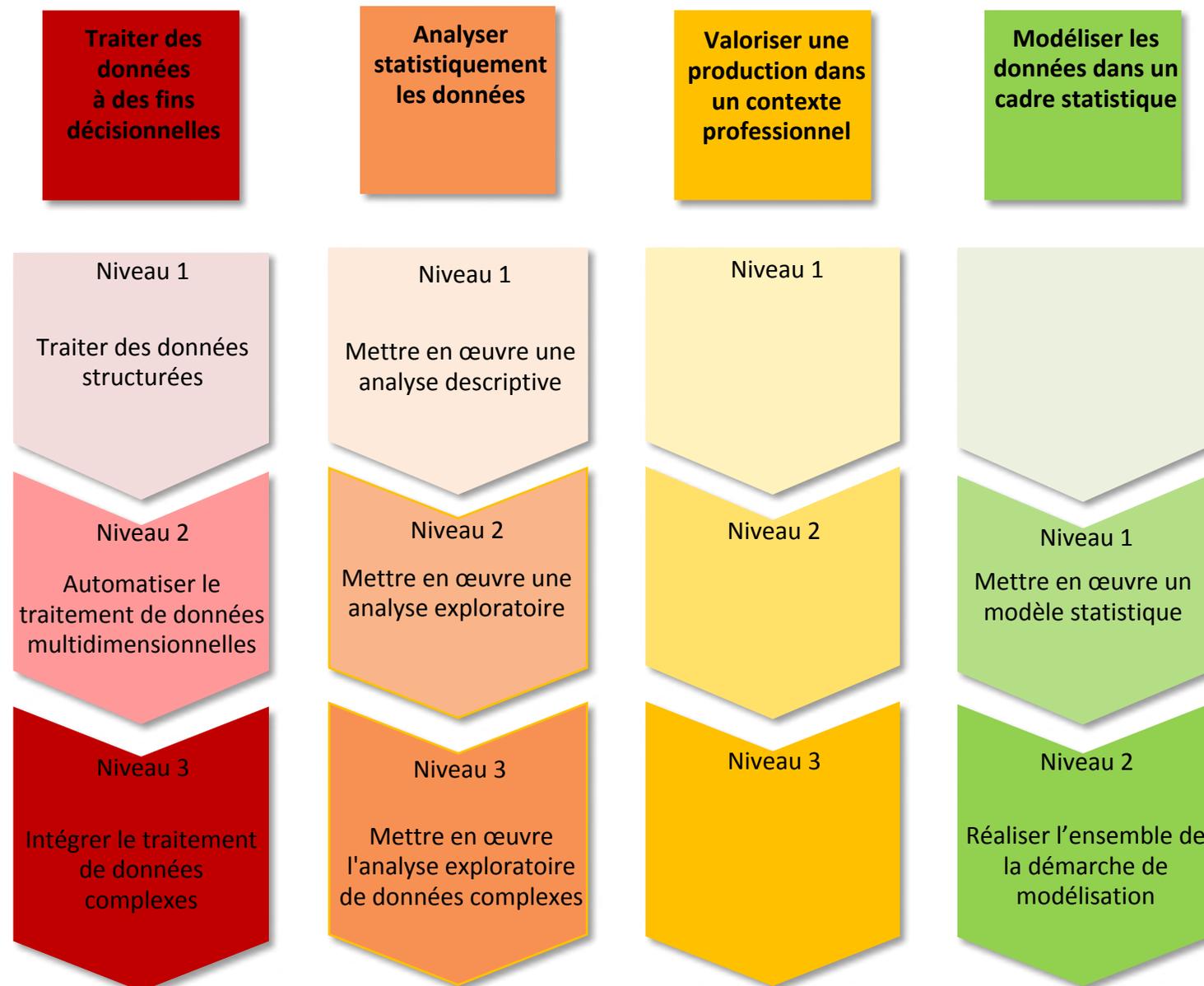
## Les situations professionnelles

Les situations professionnelles se réfèrent aux **contextes** dans lesquels les compétences sont mises en jeu. Ces situations varient selon la compétence ciblée.

|  |                             |  |
|--|-----------------------------|--|
| <b>Traiter des données à des fins décisionnelles</b>           | Situations professionnelles | Dans le contexte du développement d'un système d'information décisionnel<br>Dans le contexte de préparation des données à des fins d'analyse statistique |
| <b>Analyser statistiquement les données</b>                    | Situations professionnelles | Dans le contexte de programmation d'un outil d'aide à la décision<br>Dans le contexte d'un projet d'étude statistique                                    |
| <b>Valoriser une production dans un contexte professionnel</b> | Situations professionnelles | Dans le contexte du développement d'outils décisionnels<br>Dans le contexte d'une étude statistique  |
| <b>Modéliser les données dans un cadre statistique</b>         | Situations professionnelles | Dans le contexte d'une analyse statistique<br>Dans le contexte d'un développement statistique  |

## Les niveaux de développement des compétences

---



## Traiter des données à des fins décisionnelles

en intervenant à toutes les étapes du cycle de vie de la donnée (insertion, modification, extraction, suppression)  
en utilisant le modèle de données adapté aux besoins  
en s'inscrivant dans une démarche de documentation des réalisations adaptée au public visé  
en traduisant correctement les demandes métier en programmes, avec le respect du cahier des charges s'il existe  
en écrivant un programme correctement structuré et documenté, respectant les bonnes pratiques  
en identifiant les librairies et langages dédiés

Situations  
professionnelles

Dans le contexte du développement d'un système d'information décisionnel  
Dans le contexte de préparation des données à des fins d'analyse statistique

Niveaux de  
développement

### Apprentissages critiques

Niveau 1

Traiter des données  
structurées

Correctement interpréter et prendre en compte le besoin du commanditaire ou du client  
Respecter les formalismes de notation  
Connaître la syntaxe des langages et savoir l'utiliser  
Mesurer l'importance de maîtriser la structure des données à exploiter  
Comprendre les structures algorithmiques de base et leur contexte d'usage  
Prendre conscience de l'intérêt de la programmation

Niveau 2

Automatiser le  
traitement de données  
multidimensionnelles

Comprendre l'organisation des données de l'entreprise  
Réaliser le rôle central et spécifique de l'entrepôt de données dans la chaîne décisionnelle  
Identifier et résoudre les problèmes d'intégration de sources complémentaires et hétérogènes  
Comprendre la nécessité de tester, corriger et documenter un programme  
Apprécier l'intérêt de briques logicielles existantes et savoir les utiliser

Niveau 3

Intégrer le traitement  
de données  
complexes

Identifier les solutions technologiques permettant la collecte et la diffusion de données  
Comprendre les spécificités des données complexes et de leur exploitation  
Savoir mener une veille technologique

## Analyser statistiquement les données

en tenant compte du contexte de l'étude (économique, socio-démographique, commerciale, clinique...)  
en mettant en évidence les grandes tendances et les informations principales  
en identifiant et en mettant en œuvre les techniques adaptées aux attentes du client ou de l'instance décisionnaire  
en identifiant et en mettant en œuvre les techniques adaptées aux données complexes (données massives, données mal structurées, flux de données...)  
en tenant compte du contexte inférentiel (variabilité de l'échantillon)

Situations  
professionnelles

Dans le contexte de programmation d'un outil d'aide à la décision  
Dans le contexte d'un projet d'étude statistique

Niveaux de  
développement

### Apprentissages critiques

#### Niveau 1

Mettre en œuvre une  
analyse descriptive

Réaliser que les sources de données ont des caractéristiques propres à considérer (variation, précision, mise à jour...)  
Comprendre qu'une analyse correcte ne peut émaner que de données propres et préparées  
Comprendre l'intérêt des synthèses numériques et graphiques pour décrire une variable statistique  
Comprendre l'intérêt des synthèses numériques et graphiques pour mettre en évidence des liaisons entre variables.  
Comprendre l'intérêt de l'utilisation d'un modèle probabiliste  
Appréhender la notion de fluctuation d'échantillonnage, notamment à l'aide de simulations probabilistes

#### Niveau 2

Mettre en œuvre une  
analyse exploratoire

Prendre conscience de la différence entre modélisation statistique et analyse exploratoire  
Saisir la spécificité de l'analyse des données temporelles  
Comprendre l'intérêt des analyses multivariées pour synthétiser et résumer l'information portée par plusieurs variables  
Appréhender l'idée de confronter une hypothèse avec la réalité pour prendre une décision  
Apprécier les limites de validité et les conditions d'application d'une analyse

#### Niveau 3

Mettre en œuvre  
l'analyse exploratoire  
de données complexes

Prendre conscience des différences entre des outils statistiques pour choisir le plus adapté  
Saisir l'importance de la mise en œuvre de méthodes adaptées à des domaines et des données spécifiques (Marketing, Biostatistique, Statistique spatiale, Gestion ...)  
Prendre conscience des limites des méthodes classiques pour l'analyse des données complexes (données massives, données mal structurées...)  
Comprendre les mécanismes de bases de l'intelligence artificielle (apprentissage statistique supervisé, échantillons d'apprentissage et échantillons de test...)

## Valoriser une production dans un contexte professionnel

en s'adaptant au niveau d'expertise, à la culture et au statut du destinataire  
en s'exprimant correctement, aussi bien en français qu'en anglais, à l'oral comme à l'écrit  
en veillant aux aspects éthiques, déontologiques et réglementaires d'utilisation et de diffusion des données  
en interprétant et contextualisant les résultats (citations, vérification des sources, esprit critique)  
en utilisant la forme de restitution adaptée  
en tenant compte des réalités économiques et managériales des entreprises

Situations professionnelles

Dans le contexte du développement d'outils décisionnels  
Dans le contexte d'une étude statistique

Niveaux de développement

### Apprentissages critiques

Niveau 1

Prendre connaissance des biais rencontrés dans la mise en place d'une enquête  
Identifier l'importance de contextualiser ses données  
Mesurer l'importance de mettre en évidence des résultats clés par l'utilisation d'indicateurs pertinents  
Lors de la restitution des résultats, mesurer l'importance d'explicitement également la démarche suivie  
Comprendre les intérêts de la data visualisation et de l'infographie  
Mesurer l'importance d'une expression précise et nuancée dans la communication en français et dans une langue étrangère des résultats

Niveau 2

Saisir l'intérêt de mobiliser de manière proactive des ressources métiers liées à l'environnement (y compris économique, international...)  
Savoir défendre ses choix d'analyses  
Saisir la nécessité de choisir des indicateurs pertinents pour communiquer sur les résultats  
Prendre conscience de la rigueur requise dans ses productions et dans la communication à leur propos  
Comprendre les enjeux des relations en milieu professionnel adaptées à l'interlocuteur et à sa culture

Niveau 3

Savoir transformer la donnée pour la mettre en conformité avec des normes (anonymisation, normalisation)  
Mesurer l'impact d'un respect de la législation en terme de droit des données.  
Identifier les clés d'une bonne communication (procédure et techniques utilisées)  
Mesurer l'importance de comprendre et de répondre à l'ensemble des problématiques posées  
Être force de proposition  
Prendre conscience de la nécessité d'intégrer la vision de l'interlocuteur (transversalité, international, multiculturalité, niveau d'expertise...)

## Modéliser les données dans un cadre statistique

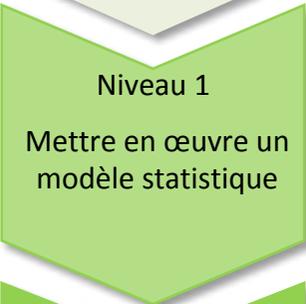
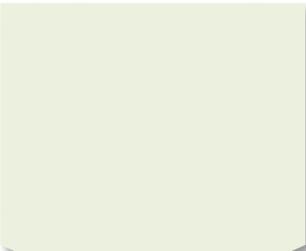
en choisissant le modèle adapté à la situation  
en maîtrisant la qualité du modèle  
en s'adaptant aux spécificités (données, enjeux, méthodes) d'un domaine d'application particulier (santé, marketing, assurance, qualité, socio-démographie...)  
en s'adaptant à la complexité des données (données massives, données mal structurées, flux de données...)

Situations  
professionnelles

Dans le contexte d'une analyse statistique  
Dans le contexte d'un développement statistique

Niveaux de  
développement

Apprentissages critiques



Niveau 1

Mettre en œuvre un  
modèle statistique

Comprendre l'intérêt de planifier le recueil des données  
Appréhender les difficultés et les limites rencontrées dans la mise en œuvre d'un terrain de collecte  
Comprendre l'impact du type de données sur le choix de la modélisation à mettre en œuvre  
Apprécier les limites de validité et les conditions d'application d'un modèle  
Réaliser l'importance de la mise en œuvre d'une procédure de test statistique pour valider ou non une hypothèse



Niveau 2

Réaliser l'ensemble de  
la démarche de  
modélisation

Comprendre l'intérêt des approches statistiques pour la fiabilisation, la validation, les incertitudes, les imprécisions des données  
Comprendre l'intérêt de la problématique métier pour réaliser la modélisation  
Viser la réalisation d'un processus de modélisation dans son ensemble  
Prendre conscience des différences entre les modèles pour choisir le plus adapté  
Prendre conscience de la nécessité d'utiliser des moyens spécifiques pour analyser les données massives ou les flux de données