



le  
campus  
numérique  
in the ALPS



**Data + 8**

**Data Analyst**

Public cible : scientifiques en reconversion



## Data Analyst Un nouveau projet pour le Campus

Selon le même esprit que pour les formations existantes nous cherchons à identifier des cibles « non traditionnelles » pour les amener vers des métiers en forte tension. Il existe aujourd’hui entre 12 et 15% de docteurs des disciplines non informatiques au chômage 3 ans après leur doctorat.

Nous proposons une formation complémentaire accélérée en informatique et science de la donnée afin de les amener sur le marché du travail en tant que Data Analyst : il est important que les « Data Analyst » aient une forte compétence « métier » pour sourcer et qualifier les données (data) et une compétence complémentaire en informatique pour comprendre comment ces données sont traitées et exploitées, leur première compétence étant encore indispensable pour interpréter les résultats établis sur ces données.

### MÉTHODES PÉDAGOGIQUES

Notre pédagogie est centrée sur le « faire ». Tous nos apprentissages sont des exemples concrets rencontrés par nos équipes pédagogiques dans leurs métiers.

Nos formateurs ont avant tout une expertise métiers autour des métiers de la donnée ou du développement.

### RYTHME D’APPRENTISAGE

Les 2 grandes phases de notre programme sont :

- Un socle de 400h de formation en présentiel
- Une alternance de 6 mois (3 semaines en entreprise / 1 semaine en cours) comprenant 210h de formations

Avec le soutien de



# PROGRAMME PÉDAGOGIQUE

## Développement Python & Algorithmie (91h)

Les bases de la programmation par la pratique à l'aide du langage Python (instructions, structures de données, programmation orientée objet)

## Machine Learning (126h)

Implémentation et apprentissage sur des jeux de données réels pour appréhender les principaux outils de machine learning (Classification, Décomposition en composantes principales, K-Mean, K-Folds, Descentes de gradients, réseaux de neurones)

## Outils statistiques pour la donnée (77h)

Ensemble d'outils mathématiques pour l'analyse de données (statistiques descriptives, tests d'hypothèses, nettoyage de données, régressions, ...)

## Bonnes pratiques de développement logiciel en équipe (28h)

Méthodes de gestion de projet agile & Git

## Système d'exploitation et Bases de données (42h)

Notions de bases sur système Linux et manipulation de bases de données

## Outils de visualisation (21h)

Cartographier et adapter les outils de visualisation aux jeux de donnée

## Compétences douces (14h)

Communication interpersonnelle, codes de l'entreprise, présence numérique.

## Éthique de la donnée (35h)

Projet de sensibilisation aux 7 principes de la donnée (La Finalité / Les proportionnalités / La pertinence / La durée limitée de la conservation des données / La sécurité et la confidentialité / La transparence / Le respect du droit des personnes)

## R (35h)

Apprentissage du langage R et analyse de donnée à l'aide de ce dernier

## Deep Learning (35h)

Ensemble d'outils mathématiques pour l'analyse de données (statistiques descriptives, tests d'hypothèses, nettoyage de données, régressions, ...)

## Cas d'études sur des grands jeux de donnée (105h)

Mise en application des bonnes pratiques d'analyse de donnée sur des jeux de données de grande taille.