



Démarrez votre formation de Data Scientist

**"DATA SCIENCE : LA FORMATION QUI TRACE LE FUTUR"**

**Data Scientist**

Samira mahjoub

12/01/2024

# SOMMAIRE

## 1. Introduction à la Formation

\*Présentation générale de la formation et objectifs personnels

## 2. Parcours de Formation

\*Aperçu des étapes clés de la formation

\*Description des projets et de la planification

## 3. Collaboration et Soutien

\*Outils de collaboration avec le mentor

# 1. INTRODUCTION À LA FORMATION



## Présentation générale de la formation et objectifs personnels

- **Description** : Cette formation vise à me fournir une expertise complète en Data Science, couvrant des sujets tels que le machine learning, l'analyse de données, et le big data.
- **Objectifs** : Mon objectif est de développer des compétences avancées en data science pour mener à bien des projets complexes et innovants.
- **Motivation** : J'ai choisi cette formation pour sa structure pratique, son approche par projet, et la flexibilité qu'elle offre.



## 2. PARCOURS DE FORMATION

### Aperçu des Étapes Clés de la Formation



**Inscription et  
Orientation**



**Webinaire de  
Bienvenue**



**Objectifs  
Professionnels**



**Parcours de  
Projets**



**Sessions de  
Mentorat**



**Évaluation et  
Soutenance de  
Projets**



**Collaboration  
et Réseau**



**Certification**

*"Les données sont le pétrole brut du 21e siècle, et l'analytique est le moteur de combustion." - Peter Sondergaard*

## 2. PARCOURS DE FORMATION

### Aperçu des Étapes Clés de la Formation

#### \* Bienvenue \* dans votre parcours ! \*

##### Étape n°1 : participez au webinaire de bienvenue

Listez 2 à 4 informations apprises pendant ce webinaire.

Information 1 /Déroulement de la Formation:La formation se compose de projets, de cours et d'évaluations structurés pour une immersion professionnelle.

Information 2 /Outils en Ligne et Ressources : Divers outils et ressources en ligne sont disponibles pour faciliter l'apprentissage et la collaboration.

##### Étape n°2 : définissez votre objectif professionnel

Décrivez de la façon la plus simple possible ce que vous aimeriez que cette formation apporte à votre parcours professionnel.

À court terme, je compte acquérir des compétences techniques approfondies en Data Science pour effectuer une transition professionnelle.

À moyen terme, je vise à intégrer la Data Science dans le secteur financier, en exploitant mes connaissances en contrôle de gestion pour apporter une valeur ajoutée à travers des analyses prédictives et des modélisations complexes.

##### Étape n° 3 : découvrez les projets de votre parcours

Quels projets vous motivent le plus ? Quels projets constitueront à priori des défis et pourraient être plus difficiles pour vous ?

###### Projets les plus motivants

- Projet 2
- Projet 3
- Projet 4
- Projet 5
- Projet 6
- Projet 8

###### Projets les plus difficiles à priori

- Projet 7
- Projet 9
- Projet 10

1

##### Étape n° 4 : fixez-vous des dates cibles pour terminer chaque projet

Indiquez une date cible de soutenance pour chaque projet, et rappelez-vous de votre date de fin.

"Mon calendrier de soutenances pour la formation Data Scientist est organisé de la manière suivante :

- Projet 1 : 12/01/2024
- Projet 2 : 06/02/2024
- Projet 3 : 05/03/2024
- Projet 4 : 11/04/2024
- Projet 5 : 13/05/2024

**Date de fin de formation : 26/09/2024**



Si OpenClassrooms vous a fourni un planning avec les dates prévisionnelles de soutenance, recopiez-les simplement ici.

Prenez le temps ici d'anticiper les difficultés éventuelles que vous pourriez rencontrer, et comment vous pourriez les surmonter.

Listez 2 à 4 bonnes pratiques que vous pensez appliquer durant votre formation pour bien gérer votre temps et vous tenir aux dates que vous avez fixées.

-Difficultés Anticipées:Les concepts théoriques les plus difficiles à maîtriser sont principalement les algorithmes de machine learning et le traitement des données Big Data.

Pour anticiper les défis, je m'engage à des sessions d'étude régulières et à utiliser des outils de gestion du temps.

##### Étape n° 5 : définissez un ou plusieurs outil(s) de collaboration avec votre mentor

Quels outils de collaboration avez-vous retenus avec votre mentor ?

-Les outils de collaboration avec mon mentor pour optimiser notre travail commun sont les suivants :

- Google Drive pour le stockage et le partage de documents, permettant un accès facile et une collaboration en temps réel sur les fichiers.

-Zoom pour les sessions de mentorat en visioconférence

2

# Description des projets et de la planification

## Déroulement de votre formation

Projet 1 10h	Démarrez votre formation de Data Scientist
Projet 2 60h	Analysez des données de systèmes éducatifs
Projet 3 70h	Préparez des données pour un organisme de santé publique
Projet 4 90h	Anticipez les besoins en consommation de bâtiments
Projet 5 70h	Segmentez des clients d'un site e-commerce
Projet 6 100h	Classifiez automatiquement des biens de consommation
Projet 7 80h	Implémentez un modèle de scoring
Projet 8 40h	Réalisez un dashboard et assurez une veille technique
Projet 9 40h	Réalisez un traitement dans un environnement Big Data sur le Cloud
Projet 10 40h	Réalisez le cadrage d'un projet IA

## Les 10 thématiques de votre formation

### Python

Pour nettoyer et analyser des datasets, vous devez maîtriser un langage de programmation. Python est le langage de plus en plus couramment privilégié en data science notamment parce qu'il est open-source.

### Notebooks

Un notebook est une interface interactive basée sur un navigateur web qui permet aux utilisateurs de créer, d'exécuter, de partager et de documenter des analyses de données en combinant du code, des visualisations, des notes et des résultats dans un même document.

### Apprentissage supervisé

Les algorithmes d'apprentissage supervisé s'entraînent sur des données déjà catégorisées afin de prédire ces mêmes catégories pour de nouvelles données.

### Apprentissage non-supervisé

L'apprentissage non supervisé permet d'identifier des structures cachées dans un ensemble de données non catégorisées, sans aucune intervention humaine.

### Données structurées

Les données structurées sont des données organisées dans un format tabulaire (avec des lignes et des colonnes), où chaque colonne représente une caractéristique (ou une variable) et chaque ligne représente une instance (ou une observation).

### Données non-structurées

Les données non structurées sont des données qui ne sont pas organisées dans un format tabulaire régulier. Elles peuvent être des données textuelles, audio, vidéo ou d'images, qui ne peuvent pas être facilement représentées dans un format standardisé.

### Feature engineering

Le feature engineering est le processus de sélection et de création de variables ou de caractéristiques (features) pertinentes à partir des données brutes, afin d'améliorer la performance des modèles d'apprentissage automatique.

### Deep Learning

Le deep learning utilise des réseaux de neurones profonds pour résoudre des problèmes complexes tels que la classification d'images, la reconnaissance vocale et la traduction automatique.

### Déploiement

Le déploiement réussi d'un modèle d'apprentissage automatique est essentiel pour la commercialisation d'une solution. Cela nécessite la création d'une API, la configuration de l'infrastructure ainsi que la surveillance de sa performance pour garantir son efficacité.

### Gestion de projet

Vous serez amené à estimer les besoins nécessaires pour mener à bien un projet, à planifier et prioriser vos tâches dans le cadre de vos projets, et à travailler en équipe.

## Tâches

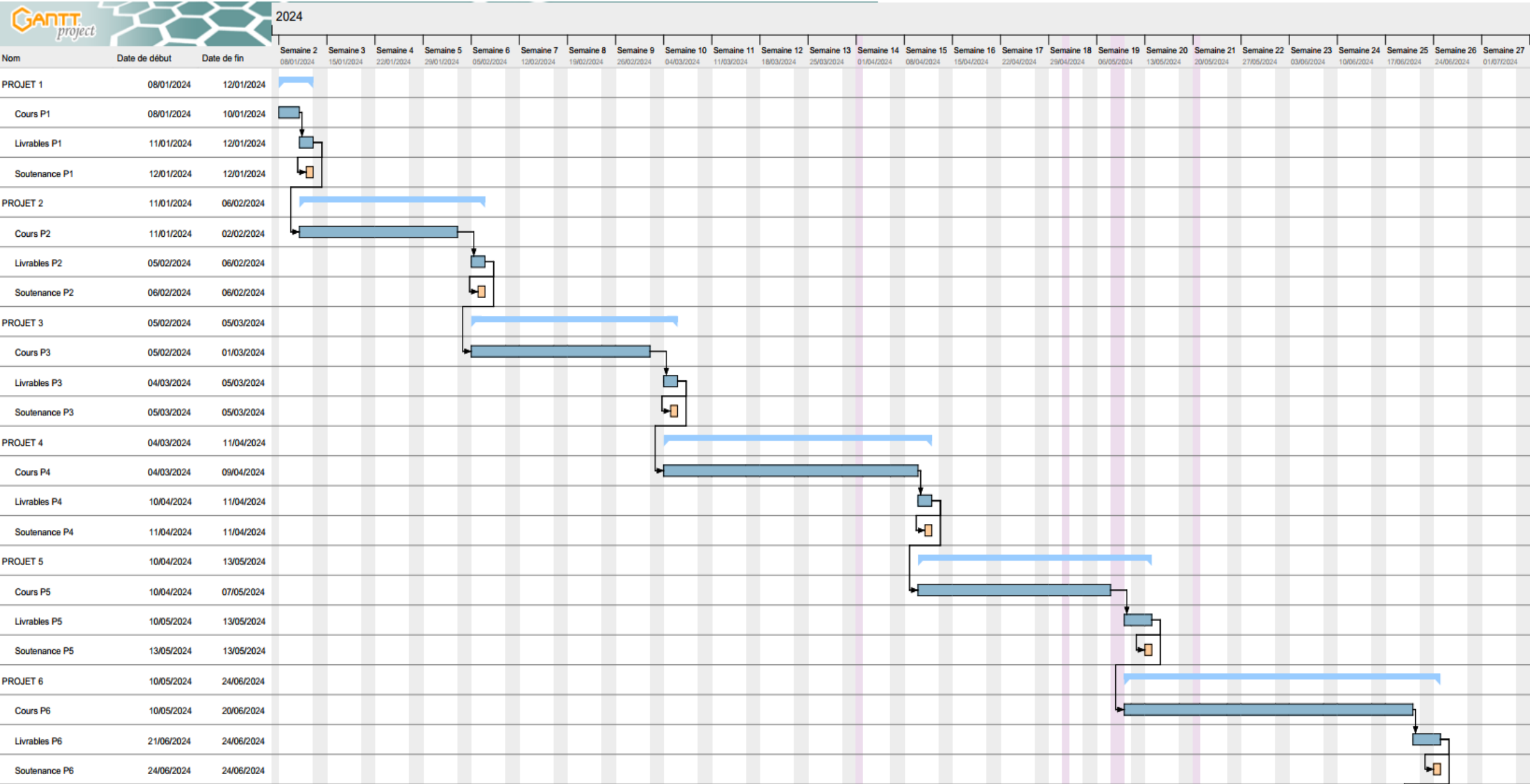
2

Nom	Date de début	Date de fin
PROJET 1	08/01/2024	12/01/2024
Cours P1	08/01/2024	10/01/2024
Livrables P1	11/01/2024	12/01/2024
Soutenance P1	12/01/2024	12/01/2024
PROJET 2	11/01/2024	06/02/2024
Cours P2	11/01/2024	02/02/2024
Livrables P2	05/02/2024	06/02/2024
Soutenance P2	06/02/2024	06/02/2024
PROJET 3	05/02/2024	05/03/2024
Cours P3	05/02/2024	01/03/2024
Livrables P3	04/03/2024	05/03/2024
Soutenance P3	05/03/2024	05/03/2024
PROJET 4	04/03/2024	11/04/2024
Cours P4	04/03/2024	09/04/2024
Livrables P4	10/04/2024	11/04/2024
Soutenance P4	11/04/2024	11/04/2024
PROJET 5	10/04/2024	13/05/2024
Cours P5	10/04/2024	07/05/2024
Livrables P5	10/05/2024	13/05/2024
Soutenance P5	13/05/2024	13/05/2024
PROJET 6	10/05/2024	24/06/2024
Cours P6	10/05/2024	20/06/2024
Livrables P6	21/06/2024	24/06/2024
Soutenance P6	24/06/2024	24/06/2024
PROJET 7	21/06/2024	24/07/2024
Cours P7	21/06/2024	22/07/2024
Livrables P7	23/07/2024	24/07/2024
Soutenance P7	24/07/2024	24/07/2024
PROJET 8	23/07/2024	07/08/2024
Cours P8	23/07/2024	05/08/2024
Livrables P8	06/08/2024	07/08/2024
Soutenance P8	07/08/2024	07/08/2024
PROJET 9	06/08/2024	22/08/2024
Cours P9	06/08/2024	20/08/2024
Livrables P9	21/08/2024	22/08/2024
Soutenance P9	22/08/2024	22/08/2024



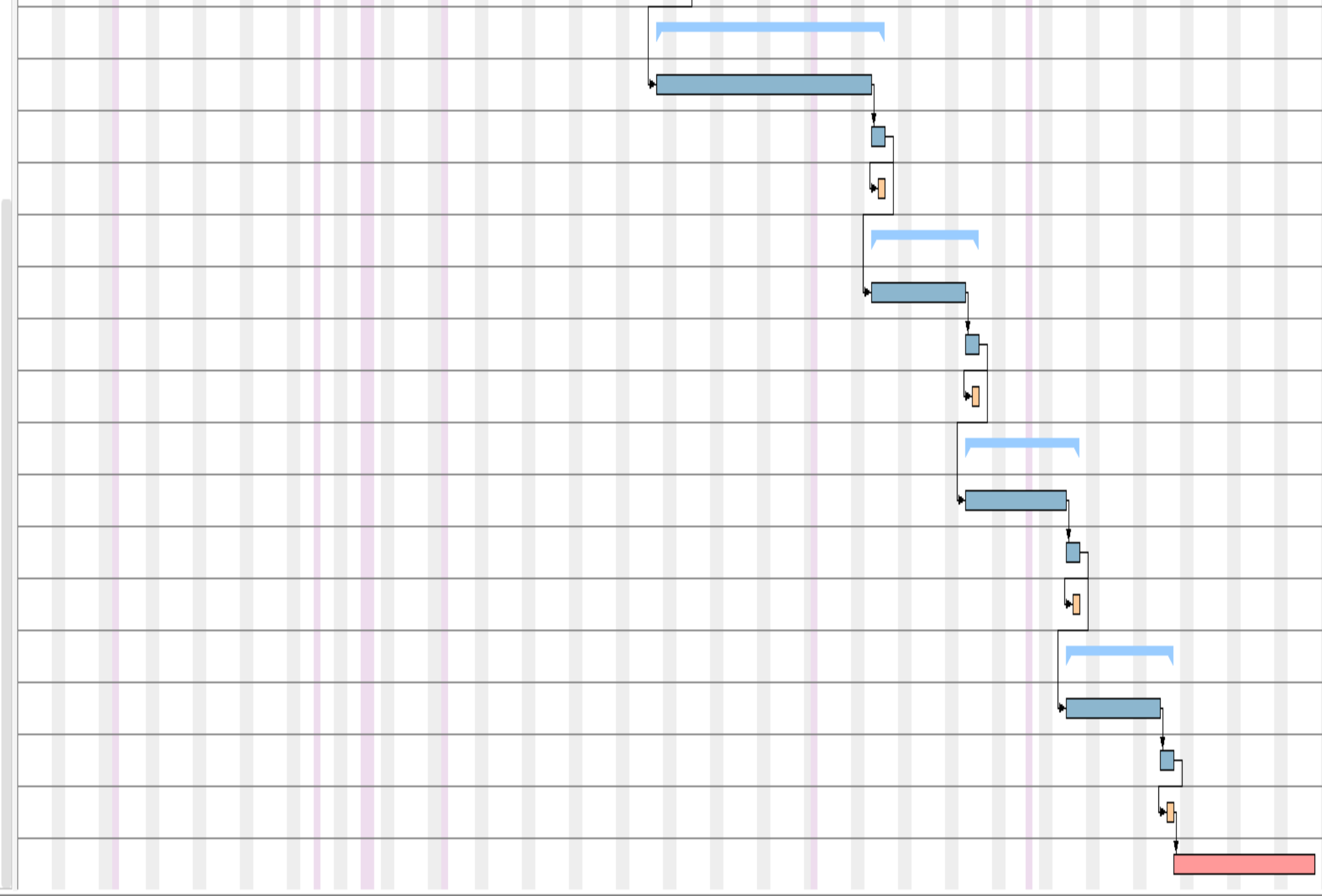
# Planification et Gestion de la Formation en Data Science

## Diagramme de Gantt





▼ PROJET 7	21/06/2024	24/07/2024
Cours P7	21/06/2024	22/07/2024
Livrables P7	23/07/2024	24/07/2024
Soutenance P7	24/07/2024	24/07/2024
▼ PROJET 8	23/07/2024	07/08/2024
Cours P8	23/07/2024	05/08/2024
Livrables P8	06/08/2024	07/08/2024
Soutenance P8	07/08/2024	07/08/2024
▼ PROJET 9	06/08/2024	22/08/2024
Cours P9	06/08/2024	20/08/2024
Livrables P9	21/08/2024	22/08/2024
Soutenance P9	22/08/2024	22/08/2024
▼ PROJET 10	21/08/2024	05/09/2024
Cours P10	21/08/2024	03/09/2024
Livrables P10	04/09/2024	05/09/2024
Soutenance P10	05/09/2024	05/09/2024
Jours d'avance	06/09/2024	26/09/2024



Planification et Gestion de la Formation en Data Science

OpenClassrooms  
Data Scientist

Début de Formation : 08/01/2024

PROJET	Nom	Cours (h)	Travail perso (h)	Charge prévue (h)	Delais livrables (J)	Charge totale (J)	DÉBUT	FIN	Soutenance
Projet 1	Démarrez votre formation de Data Scientist	10	10	20	2	5	08/01/2024	12/01/2024	12/01/2024
Projet 2	Analysez des données de systèmes éducatifs	60	60	120	2	19	11/01/2024	06/02/2024	06/02/2024
Projet 3	Préparez des données pour un organisme de santé publique	70	70	140	2	22	05/02/2024	05/03/2024	05/03/2024
Projet 4	Anticipez les besoins en consommation de bâtiments	90	90	180	2	28	04/03/2024	11/04/2024	11/04/2024
Projet 5	Segmentez des clients d'un site e-commerce	70	70	140	2	22	10/04/2024	13/05/2024	13/05/2024
Projet 6	Classifiez automatiquement des biens de consommation	100	100	200	2	31	10/05/2024	24/06/2024	24/06/2024
Projet 7	Implémentez un modèle de scoring	80	80	160	2	25	21/06/2024	24/07/2024	24/07/2024
Projet 8	Réalisez un dashboard et assurez une veille technique	40	40	80	2	13	23/07/2024	07/08/2024	07/08/2024
Projet 9	Réalisez un traitement dans un environnement Big Data sur le Cloud	40	40	80	2	13	06/08/2024	22/08/2024	22/08/2024
Projet 10	Réalisez le cadrage d'un projet IA	40	40	80	2	13	21/08/2024	05/09/2024	05/09/2024
TOTAL		600	600	1'200	20	191	21/08/2024	05/09/2024	05/09/2024
Jours d'avance							06/09/2024	26/09/2024	

### 3. COLLABORATION ET SOUTIEN

#### Outils de collaboration avec le mentor

- Afin d'assurer une collaboration efficace, nous utiliserons **Google Drive** pour le partage de documents et de ressources d'apprentissage, un **outil de visioconférence** intégré sur la plateforme d'OpenClassrooms pour les réunions hebdomadaires, et **Slack** pour les communications quotidiennes.

- **Trello** servira à la gestion des tâches et au suivi des progrès.

Ces outils sont essentiels pour une expérience d'apprentissage structurée et pour maintenir un lien constant avec mon mentor.





## Vers un Avenir en Data Science!

**"Chaque grand voyage commence par un premier pas. La formation en Data Science est mon horizon nouveau, une route vers des possibilités inexplorées et des connaissances qui attendent au tournant de chaque défi rencontré."**

**En avant toute!**



# "DATA SCIENCE : PRÉVOIR L'IMPACT DE DEMAIN"

"COMMENT ENVISAGEZ-VOUS L'AVENIR DE LA DATA SCIENCE ET SON INFLUENCE SUR NOTRE FAÇON DE PRENDRE DES DÉCISIONS AU QUOTIDIEN?«

PRENEZ UN MOMENT POUR RÉFLÉCHIR À CETTE QUESTION. NOUS DISCUTERONS DE VOS PERSPECTIVES DANS UN INSTANT.





**MERCI DE VOTRE ATTENTION**