

2025

PROJECT
PORTFOLIO

ANALYSE RH

**Préférence pour la
flexibilité au travail**

Jean Martial GAHOUSSOU

SOMMAIRE

Introduction	1
Présentation de l'entreprise	2
Ma place à l'agence	3
Mes réalisations	4
Bilan	5
Conclusion & remerciements	6

INTRODUCTION

Ce projet propose une mise en pratique complète des approches d'apprentissage automatique supervisé appliquées à des problématiques réelles.

Il se structure en deux parties distinctes mais complémentaires, qui visent à reproduire des situations concrètes de travail analytique et décisionnel en entreprise. L'objectif global du projet est de permettre l'application rigoureuse de méthodes de classification afin d'extraire de la valeur à partir de données réelles, d'identifier des patterns et de proposer des recommandations opérationnelles éclairées.



CONTEXTE

L'évolution des modes de travail, marquée par l'essor du télétravail, des horaires variables et des semaines compressées, oblige les organisations à réviser leurs politiques de gestion des ressources humaines. Dans ce contexte, il devient impératif pour les directions RH de comprendre les facteurs déterminants qui influencent la propension des employés à adopter un mode de travail flexible.

Une telle compréhension constitue un levier stratégique permettant d'anticiper les préférences des employés, d'élaborer des politiques adaptées et de renforcer leur engagement tout en maintenant la performance organisationnelle.

Le jeu de données **WorkFlexSurvey** comprend **9 000** enregistrements représentant des employés issus de divers secteurs d'activité. Chaque enregistrement décrit un ensemble de **14 variables explicatives** couvrant des dimensions personnelles (âge, perception de la santé, situation familiale) et professionnelles (ancienneté, satisfaction au travail, temps de déplacement, confiance envers la hiérarchie). La **variable cible (Prefers_Flex_Work)** indique la préférence de l'employé pour un horaire flexible (1) ou non (0).



Offrir de la flexibilité à un salarié qui en a besoin.
C'est reconnaître son engagement et disponibilité lorsque l'entreprise en avait le plus besoin.

SOUAD LAAMOUI, Conseil RH & Entrepreneurial
(2024,déc)

CHOIX DU LANGAGE DE PROGRAMMATION

Nous avons choisi Python pour ce projet en raison de sa popularité et de sa robustesse en science des données. Python propose des bibliothèques puissantes comme Pandas, scikit-learn et Matplotlib, idéales pour la préparation des données, la modélisation (KNN, arbres de décision, forêts aléatoires, réseaux de neurones) et la visualisation. Il assure la reproductibilité et la transparence des analyses grâce aux notebooks, tout en respectant les consignes du projet qui autorisent d'autres outils que SPSS Modeler. Sa flexibilité permet de traiter efficacement les deux problématiques : prédire les préférences des employés pour le travail flexible et identifier les clients à fort potentiel d'achat. Python garantit ainsi une approche complète, rigoureuse et adaptée aux objectifs pédagogiques du projet.

OBJECTIFS

L'objectif de ce projet est de concevoir un modèle de classification supervisée permettant de prédire la préférence des employés en matière de flexibilité du travail à partir des variables observées. Pour atteindre cet objectif, plusieurs étapes méthodologiques seront mises en œuvre :

- Réalisation d'une analyse descriptive afin de caractériser la structure et les principales tendances du jeu de données.
- Préparation des données, incluant le nettoyage, la gestion des valeurs manquantes, l'encodage des variables catégorielles et la normalisation des variables numériques.
- Construction et évaluation de plusieurs modèles de classification : K-Nearest Neighbours (KNN), arbres de décision (CART et C5.0) et réseaux de neurones.
- Analyse comparative des modèles sur la base de métriques standardisées telles que l'accuracy, le rappel (sensibilité) et la précision.
- Application des modèles retenus pour générer des prédictions sur un jeu de données partiellement étiqueté (WorkFlexSurvey_Null).

MÉTHODOLOGIE

La démarche de modélisation prévoit une séparation stricte des données en ensembles d'apprentissage (70 %), de validation (15 %) et de test (15 %), afin de garantir une évaluation impartiale des performances des modèles. Elle inclut également la recherche des meilleurs hyperparamètres (comme la valeur optimale de k pour KNN), la présentation des matrices de confusion et des métriques essentielles telles que l'accuracy, la sensibilité (recall) et la précision (precision). De plus, les modèles seront appliqués à un second jeu de données « WorkFlexSurvey-null », dont la cible est absente, afin de démontrer leur capacité de généralisation et d'illustrer leur potentiel en contexte réel.

Cette démarche vise à fournir aux gestionnaires des ressources humaines un outil d'aide à la décision rigoureux et transparent, leur permettant d'anticiper les préférences des employés et d'adapter leurs politiques de flexibilité de manière éclairée et stratégique.

Le processus comprendra la compréhension et la préparation des données , la modélisation , l'évaluation des performances et le déploiement sur un second jeu de données. L'objectif final est de fournir à la direction une recommandation claire et justifiée sur la stratégie de ciblage, en s'appuyant sur l'approche de modélisation la plus performante.

Ce rapport documentera l'ensemble de la démarche analytique. Il visera à proposer une démarche rigoureuse, reproductible et compréhensible, permettant de répondre de façon convaincante aux problématiques posées tout en illustrant l'intérêt et la valeur ajoutée des techniques de forage de données dans la prise de décision stratégique en entreprise.

ANALYSE EXPLORATOIRE

Une brève analyse descriptive du jeu de données « WorkFlexSurvey ».
Et les premières conclusions que pouvons en tirer ?





WorkFlexSurvey

Nombre de variable : 15

✓ Variables explicatives : 14

✓ Variables cible: 1

Le jeu de données **WorkFlexSurvey** regroupe les informations de 9 000 employés issus de divers secteurs d'activité. L'objectif de cette analyse est d'identifier les **facteurs influençant la préférence pour un horaire flexible**, dans un contexte où les politiques de télétravail et d'horaires aménagés prennent de plus en plus d'importance. Une meilleure compréhension de ces déterminants permettra d'**adapter les politiques RH aux attentes réelles des employés**.

Exploratory Data Analysis can never be the whole story, but nothing else can serve as the foundation stone – as the first step.

Tukey, J. W. (1977, p. 3)

ANALYSE DESCRIPTIVE

L'analyse descriptive (aussi appelée statistique descriptive) est une étape fondamentale de tout processus analytique ou de data science. Elle consiste à résumer, organiser et présenter les caractéristiques principales d'un ensemble de données, souvent sous forme de tableaux, de graphiques et de mesures statistiques.

STATISTIQUES GLOBAL

Variable	Description	Type
ID	Identifiant unique de l'employé	Numérique (Identifiant)
Prefers_Flex_Work	Préférence pour un horaire flexible (1 = oui, 0 = non)	Binaire (Cible)
Age	Âge de l'employé	Numérique continu
Sector	Secteur d'emploi	Catégoriel nominal
Work_Life_Balance_Score	Score de l'équilibre vie pro/perso (1–10)	Numérique discret / Ordinal
Commute_Time	Temps de déplacement quotidien (minutes)	Numérique continu
Meetings_per_week	Nombre moyen de réunions par semaine	Numérique discret
Job_Satisfaction	Niveau de satisfaction au travail (Low, Medium, High)	Catégoriel ordinal
Manager_Trust_Score	Score de confiance envers la hiérarchie (1–10)	Numérique discret / Ordinal
Has_Dependent_Children	Présence d'enfants à charge (1 = oui, 0 = non)	Binaire
Tech_Comfort_Level	Niveau de confort avec les outils numériques (1–5)	Numérique discret / Ordinal
Health_Status	Perception de la santé générale (Good, Average, Poor)	Catégoriel ordinal
Annual_Salary_Bracket	Catégorie salariale (" <code><40k</code> ", " <code>40k–70k</code> ", " <code>>70k</code> ")	Catégoriel ordinal
Seniority_Years	Ancienneté en années	Numérique continu
Burnout_Risk_Level	Niveau de risque d'épuisement professionnel (Low, Medium, High)	Catégoriel ordinal

RÉSUMÉ STATISTIQUES

	count	mean	std	min	25%	50%	75%	max	mode	Range	IQR	Skewness	Kurtosis
Prefers_Flex_Work	9000.0	0.593889	0.491133	0.0	0.0	1.0	1.0	1.0	1	1	1.0	-0.382421	-1.854166
Age	9000.0	42.969667	12.400549	22.0	32.0	43.0	54.0	64.0	22	42	22.0	-0.006084	-1.191729
Work_Life_Balance_Score	9000.0	5.482778	2.863982	1.0	3.0	5.0	8.0	10.0	4	9	5.0	0.017279	-1.210042
Commute_Time	9000.0	71.260333	26.405859	0.0	49.0	75.0	92.0	120.0	90	120	43.0	-0.246417	-0.924459
Meetings_per_week	9000.0	4.982000	2.249399	0.0	3.0	5.0	6.0	17.0	4	17	3.0	0.446359	0.226346
Manager_Trust_Score	9000.0	5.522667	2.889131	1.0	3.0	6.0	8.0	10.0	10	9	5.0	-0.007752	-1.229439
Has_Dependent_Children	9000.0	0.401778	0.490285	0.0	0.0	0.0	1.0	1.0	0	1	1.0	0.400763	-1.839798
Tech_Comfort_Level	9000.0	3.015222	1.403206	1.0	2.0	3.0	4.0	5.0	2	4	2.0	-0.007490	-1.283989
Seniority_Years	9000.0	9.983333	5.202847	0.0	6.0	10.0	14.0	29.0	12	29	8.0	0.101758	-0.395599

L'analyse descriptive du jeu de données révèle plusieurs tendances clés sur les caractéristiques des employés. D'abord, une majorité des répondants (environ 59 %) expriment une préférence pour des modalités de travail flexibles. Cette variable binaire présente une distribution relativement équilibrée, bien que légèrement **asymétrique** vers la gauche, ce qui signifie que la préférence n'est pas unanime, mais relativement partagée dans l'ensemble.

ANALYSE EXPLORATOIRE

L'âge moyen des employés est d'environ **43 ans**, avec une dispersion importante allant de 22 à 64 ans, ce qui indique une population professionnelle **diversifiée en termes de générations**. La distribution des âges est presque parfaitement symétrique, confirmant une répartition équilibrée sans concentration extrême sur un groupe d'âge spécifique.

En ce qui concerne l'**équilibre entre la vie professionnelle et la vie personnelle**, les employés donnent une note moyenne de **5,48 sur 10**, suggérant une perception mitigée. La médiane étant de 5, la distribution semble relativement centrée, sans déséquilibre majeur.

Le **temps de déplacement quotidien**, quant à lui, se révèle particulièrement élevé, avec une moyenne de plus de **71 minutes**. Cette variable présente une forte variabilité, certaines personnes n'ayant aucun trajet tandis que d'autres peuvent passer jusqu'à deux heures par jour à se rendre au travail.

Le **nombre moyen de réunions par semaine** est de **5**, mais cette variable montre une dispersion notable, certaines personnes assistant à plus de quinze réunions hebdomadaires. Cette situation pourrait refléter des différences importantes selon les rôles, les départements ou les niveaux hiérarchiques.

Le **score moyen de confiance envers les gestionnaires** s'élève à **5,52 sur 10**, ce qui traduit une confiance modérée de la part des employés envers leur encadrement. Ce score est distribué de façon **relativement uniforme**, sans tendance marquée. D'un **point de vue familial**, environ **40 %** des employés déclarent avoir des enfants à charge, une proportion significative, mais qui ne constitue pas la majorité. Le **niveau de confort avec les outils technologiques** affiche une moyenne de **3 sur 5**, ce qui indique que les employés se sentent plutôt à l'aise avec la technologie, sans excès ni défi majeur apparent.

Enfin, l'**ancienneté moyenne** dans l'organisation avoisine les dix ans, avec une variabilité importante, ce qui montre que l'échantillon inclut à la fois des employés très récents et d'autres en poste depuis plusieurs décennies.

Dans l'ensemble, ces statistiques brossent un portrait nuancé des employés : une population relativement expérimentée, diversifiée sur le plan démographique, et modérément satisfaite sur les plans organisationnels et professionnels. L'analyse descriptive permet ainsi de dégager les premières tendances utiles à l'élaboration d'analyses plus approfondies.

MA PLACE À L'AGENCE

Etiam velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

03

MON RÔLE

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

MON ORGANISATION

01. Vestibulum lectus mauris ultrices eros in cursus turpis massa tincidunt dui ut ornare lectus sit amet est placerat in egestas erat imperdiet

02. Vestibulum lectus mauris ultrices eros in cursus turpis massa tincidunt dui ut ornare lectus sit amet est placerat in egestas erat imperdiet

03. Vestibulum lectus mauris ultrices eros in cursus turpis massa tincidunt dui ut ornare lectus sit amet est placerat in egestas erat imperdiet

04. Vestibulum lectus mauris ultrices eros in cursus turpis massa tincidunt dui ut ornare lectus sit amet est placerat in egestas erat imperdiet

MES RÉALISATIONS

Etiam tempore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

A large graphic element consisting of two dark blue shapes on a light beige background. On the left is a circle containing a smaller, vertically oriented oval. To its right is a large, slanted rectangular shape that tapers to a point, with a white triangle cut out of its upper right corner.

PROJET 01

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.



3 jours

vestibulum lectus
mauris ultrices
eros in cursus
turpis massa
tincidunt dui ut
ornare lectus sit
amet est placerat
in egestas erat
imperdiet



L'idée

vestibulum lectus
mauris ultrices
eros in cursus
turpis massa
tincidunt dui ut
ornare lectus sit
amet est placerat
in egestas erat
imperdiet



Validation

vestibulum lectus
mauris ultrices
eros in cursus
turpis massa
tincidunt dui ut
ornare lectus sit
amet est placerat
in egestas erat
imperdiet



PROJET 02



3 jours



L'idée



Validation



Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

PROJET 03



Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.



3 jours



L'idée



Validation

BILAN

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

05

BILAN TECHNIQUE

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

BILAN HUMAIN

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

BILAN PROFESSIONNEL

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

CONCLUSION & RECOMMENDATION

Consectetur adipisci
et labore magna aliqua.
Ut enim ad minim
veniam, quis nostrud
exercitation ullamco laboris nisi
ut aliquip ex ea commodo
consequat. Duis aute irure
dolor in reprehenderit in
voluptate velit esse cillum
dolore eu fugiat nulla pariatur.
Excepteur sint occaecat
cupiditat non proident, sunt in
culpa qui officia deserunt mollit
anim id est laborum.

CONCLUSION

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.

REMERCIEMENTS

Lorem ipsum dolor sit amet, consectetur adipiscing elit, sed do eiusmod tempor incididunt ut labore et dolore magna aliqua. Ut enim ad minim veniam, quis nostrud exercitation ullamco laboris nisi ut aliquip ex ea commodo consequat. Duis aute irure dolor in reprehenderit in voluptate velit esse cillum dolore eu fugiat nulla pariatur. Excepteur sint occaecat cupidatat non proident, sunt in culpa qui officia deserunt mollit anim id est laborum.



**Vestibulum lectus
mauris ultrices eros in
cursus turpis massa
tincidunt dui ut ornare
lectus sit amet est
placerat in egestas
erat imperdiet**



BIBLIOGRAPHIE

- Souad LAAMOUI.(2024,déc). LinkedIn. Consulté le 05 Sept 2025 sur <https://www.linkedin.com/in/souadlaamoui/>
- Tukey, J. W. (1977). Exploratory Data Analysis. Addison-Wesley.
- ioh
-

2030

CONCORDIA

MERCI

Andréa Sanchez