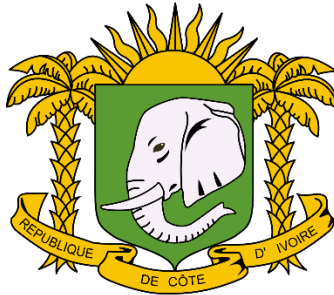


République de Côte d'Ivoire



Union – Discipline - Travail



Ecole Nationale Supérieure de Statistique et d'Economie Appliquée

**EMPLOI ET ECONOMIE : ANALYSE DES TENDANCES
ECONOMIQUES ET SOCIALES VIA LE WEB SCRAPING,
L'ANALYSE ET LA VISUALISATION DE DONNEES**

Réalisé par :

ATOUKOUI Komi Denis

Elève ISE 1 Eco

Enseignant :

Dr. N'GOLO Konaté

PhD in Computational Mathematics

Janvier 2025

Table des matières

Contexte et justification.....	2
Introduction	3
Partie 1 : web scraping	4
1. Choix de la plateforme scrappée	5
2. Méthodes et outils utilisés	5
Partie 2 : Analyse des données collectées.....	7
1. Nettoyage et traitement des données	8
2. Secteurs en vogue et compétences sollicités sur le marché.....	8
3. Répartition du type de contrat proposé par les employeurs.....	9
4. Analyse des Entreprises : nombres d’offres et notation	10
5. Analyse du salaire moyen proposé par les entreprises	11
6. Analyse de la répartition géographique des offres d’emploi	13
Conclusion et recommandations.....	15

Liste des figures

Figure 1: Nuage de mots montrant les secteurs en vogue et les compétences prisées	9
Figure 2: Répartition des types de contrats	10
Figure 3: Cartographie montrant les régions avec plus d'offres	13
Figure 4: Dashboard interactif créé sur Power BI	14

Liste des tableaux

Tableau 1: Top 5 des entreprises ayant publié le plus d’offre.....	10
Tableau 2: Top 10 entreprises les mieux notées	11
Tableau 3: Statistiques sommaire de la variable salaire moyen	12
Tableau 4: Top 10 entreprises qui payent le mieux	12

Contexte et justification

Dans un monde en constante évolution, où l'économie est de plus en plus guidée par l'innovation technologique et les dynamiques sociales, le marché de l'emploi joue un rôle central en tant que baromètre des transformations économiques. Comprendre les tendances de l'emploi n'est pas seulement une préoccupation pour les décideurs politiques ou les économistes, mais également un enjeu stratégique pour les entreprises, les chercheurs d'emploi et les institutions de formation. La France, en tant qu'une des principales économies européennes, fait face à des défis croissants en matière d'adéquation entre l'offre et la demande sur le marché du travail. Les entreprises recherchent des compétences spécifiques dans des secteurs en plein essor, tandis que certaines régions connaissent une plus forte concentration d'opportunités que d'autres. Parallèlement, des tendances émergentes comme le télétravail modifient les paradigmes traditionnels du travail, offrant de nouvelles perspectives tout en posant de nouveaux défis, notamment en matière d'organisation et d'accessibilité.

Ce travail s'inscrit dans ce contexte et vise à éclairer les tendances du marché de l'emploi en France à partir d'un ensemble de données collectées sur une plateforme de recrutement en ligne. L'objectif est de fournir une analyse rigoureuse, visuellement engageante et stratégiquement pertinente pour répondre aux questions suivantes : quelles sont les compétences les plus recherchées ? Quelles régions offrent le plus d'opportunités ? Quelles sont les entreprises qui offrent les salaires les plus attractifs, et comment ces salaires se comparent-ils à la perception de ces entreprises par leurs employés ?

Introduction

Dans un monde de plus en plus tourné vers la digitalisation, le marché de l'emploi constitue un indicateur clé des dynamiques économiques et sociales. L'essor des plateformes en ligne dédiées au recrutement offre une opportunité unique de recueillir des données riches et variées pour analyser les tendances du marché, les besoins des employeurs, et les évolutions des modes de travail, tels que le télétravail.

Ce projet s'articule autour d'un double objectif : démontrer la pertinence des techniques de web scraping pour extraire des données pertinentes et exploiter ces données afin de fournir une analyse approfondie et visuelle des offres d'emploi en France. La démarche est structurée en deux parties principales. La première partie se concentre sur le web scraping, en détaillant le choix de la plateforme étudiée, ainsi que les outils et méthodes utilisés pour collecter les données. Cette étape permet de constituer une base de données riche et fiable, essentielle à toute analyse. La seconde partie s'intéresse à l'analyse des données collectées. Elle explore d'abord les processus de nettoyage et de traitement des données, avant de mettre en lumière les secteurs en vogue, les compétences les plus recherchées, et les types de contrats proposés. Une attention particulière est accordée aux entreprises, à travers une analyse de leur notation et des salaires qu'elles offrent. Enfin, la répartition géographique des offres est examinée, avec un focus spécifique sur les emplois en télétravail et sur la ville de Paris.

Ce projet ambitionne de fournir une vision claire et exploitable des tendances actuelles du marché de l'emploi en France, tout en illustrant la puissance des outils numériques pour l'analyse des données économiques et sociales.

Partie 1 : web scraping

Dans cette partie, nous présenterons la plateforme scrappée et les méthodes et outils utilisés pour le web scraping

1.Choix de la plateforme scrappée

Pour ce projet, la plateforme **Glassdoor** a été choisie comme source principale de données. Glassdoor est l'une des plateformes de recherche d'emploi les plus populaires au monde, offrant une interface riche et une multitude d'informations pertinentes pour les chercheurs d'emploi et les analystes du marché.

Présentation de Glassdoor

Créée en 2008, Glassdoor se distingue par son approche unique qui combine des offres d'emploi avec des avis anonymes d'employés et des évaluations d'entreprises. Elle permet aux utilisateurs d'obtenir des informations non seulement sur les postes disponibles, mais aussi sur les environnements de travail, les salaires proposés et les avantages offerts par les employeurs. La plateforme couvre une large gamme de secteurs et d'industries, ce qui en fait une ressource précieuse pour analyser les tendances du marché du travail.

Pourquoi Glassdoor ?

Parmi les nombreuses plateformes de recherche d'emploi existantes, telles que LinkedIn, Indeed ou Monster, le choix de Glassdoor s'explique par plusieurs facteurs clés :

- **Données riches et variées** : Glassdoor offre des informations détaillées sur chaque offre, y compris le titre du poste, la localisation, le type de contrat, le salaire estimé, et parfois les compétences requises. De plus, la présence de notes attribuées aux entreprises par leurs employés constitue une donnée unique permettant une analyse plus qualitative des employeurs.
- **Focus sur les salaires et les entreprises** : Contrairement à d'autres plateformes qui se concentrent uniquement sur les offres d'emploi, Glassdoor inclut des estimations de salaires, permettant une exploration approfondie des rémunérations par secteur, région ou entreprise.
- **Représentativité du marché français** : Glassdoor est largement utilisé en France, ce qui garantit une couverture significative des offres d'emploi, y compris celles en télétravail, un mode de travail en pleine expansion.
- **Accessibilité des données** : La structure du site de Glassdoor facilite le scraping grâce à une organisation claire des informations dans le HTML, permettant une collecte efficace et ciblée.

2.Méthodes et outils utilisés

Pour collecter les données nécessaires à cette analyse, deux outils principaux ont été utilisés : **Selenium** et **BeautifulSoup**. Ces technologies, largement employées dans les projets de web

scraping, ont permis d'automatiser la collecte et le traitement des informations disponibles sur Glassdoor.

Selenium est un outil puissant permettant de simuler les interactions avec un navigateur web, ce qui est essentiel pour accéder aux contenus dynamiques de sites tels que Glassdoor. Il a été utilisé ici pour :

- Naviguer sur les différentes pages d'offres d'emploi en générant dynamiquement les URL à partir de mots-clés spécifiques.
- Gérer les délais de chargement des pages et éviter les blocages liés à des mécanismes anti-scraping en simulant un comportement humain grâce à des pauses aléatoires.
- Extraire les contenus HTML dynamiques, souvent chargés après le rendu initial de la page.

Une fois le contenu des pages récupéré grâce à Selenium, BeautifulSoup a été employé pour parcourir et analyser les structures HTML afin d'identifier et d'extraire les informations pertinentes, comme le titre du poste, le nom de l'entreprise, la localisation, le salaire et la notation des entreprises.

Approche de collecte des données

Les mots-clés utilisés pour le scraping ont été soigneusement sélectionnés pour couvrir une large gamme de métiers en demande, tels qu'ingénieur en informatique, ingénieur civil, ou data scientist. Un script Python automatisé a permis de collecter des données depuis des centaines d'annonces réparties sur plusieurs types de contrats. Un fichier CSV a été utilisé comme support pour sauvegarder progressivement les données, facilitant leur exploitation dans les analyses ultérieures.

Pourquoi cette combinaison ?

L'utilisation conjointe de Selenium et BeautifulSoup a offert une flexibilité et une efficacité optimales pour surmonter les défis techniques liés au scraping de sites complexes comme Glassdoor. Cette approche a permis de récupérer des données structurées et variées, tout en minimisant les risques de détection et de blocage.

Partie 2 : Analyse des données collectées

Après la phase de collecte, les données extraites nécessitent une analyse approfondie afin de répondre aux objectifs définis. Cette étape est essentielle pour transformer les informations brutes en insights exploitables, permettant ainsi de mieux comprendre les dynamiques du marché de l'emploi en France. Dans ce chapitre, nous détaillerons les différentes étapes de traitement et d'analyse des données.

1. Nettoyage et traitement des données

Pour garantir la qualité et la fiabilité des analyses, une étape préliminaire de nettoyage et de traitement des données a été réalisée. Cette phase a permis de corriger les incohérences, de gérer les valeurs manquantes et de standardiser les informations essentielles. Voici les principales opérations effectuées :

- Vérification et traitement des valeurs manquantes

Une inspection initiale des données a révélé la présence de valeurs manquantes dans certaines colonnes clés, telles que les informations sur les salaires. Ces lacunes ont été prises en compte pour éviter tout biais dans les analyses ultérieures.

- Suppression des doublons

Les doublons présents dans l'ensemble de données ont été supprimés afin d'éviter les redondances susceptibles de fausser les résultats. Cette étape a permis de garantir l'unicité et la pertinence des observations.

- Traitement de la colonne « Salaire »

La colonne des salaires, exprimée sous des formats variés (plage de salaires, montants uniques, ou encore valeurs non spécifiées), a été nettoyée et transformée pour une meilleure exploitation. Les mentions inutiles ont été supprimées. Les plages de salaires (ex. : « 36000 - 40000 € ») ont été décomposées en trois nouvelles colonnes : Salaire_Min, Salaire_Max et Salaire_Moyen. Les salaires exprimés en milliers d'euros (ex. : « 35K ») ont été convertis en montants numériques standards.

2. Secteurs en vogue et compétences sollicités sur le marché

L'analyse du nuage de mots généré à partir de la variable « titre » permet d'identifier les secteurs et compétences les plus prisés sur le marché de l'emploi en France. Voici les principaux enseignements tirés :

- **La prédominance des métiers d'ingénierie**

Le mot « ingénieur », particulièrement mis en avant, reflète une forte demande pour des profils techniques et spécialisés. Les domaines associés tels que génie civil, mécanique, informatique et sécurité confirment l'importance des secteurs industriels, technologiques et d'infrastructure dans le paysage économique actuel.

- **Montée en puissance des technologies émergentes**

Des termes comme Data Scientist, IA (intelligence artificielle), cloud, et développement logiciel mettent en évidence une demande croissante pour des compétences liées à la transformation

L'analyse des types de contrats montre une légère prédominance des Contrats à Durée Indéterminée (CDI), qui représentent 57,3 % des offres d'emploi analysées. Cette tendance souligne l'intérêt des employeurs pour des collaborations à long terme, garantissant une stabilité tant pour l'entreprise que pour les employés. Toutefois, une part significative des offres (42,7 %) concerne les Contrats à Durée Déterminée (CDD). Cela peut s'expliquer par la flexibilité qu'offrent ces contrats pour répondre à des besoins ponctuels, tels que des projets spécifiques, des remplacements temporaires, ou une adaptation à une demande fluctuante.

La prédominance des CDI peut refléter la volonté des entreprises de fidéliser des talents, notamment dans des secteurs où les compétences techniques et spécialisées sont rares. La proportion élevée de CDD témoigne également d'un marché de l'emploi dynamique, où les entreprises cherchent à s'adapter rapidement aux variations économiques ou à tester les candidats avant une éventuelle pérennisation du contrat.

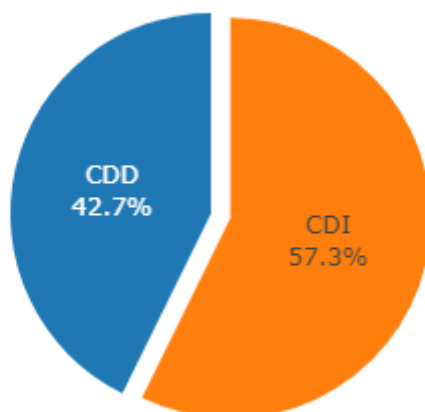


Figure 2: Répartition des types de contrats

4. Analyse des Entreprises : nombres d'offres et notation

Le tableau ci-après donne le top 5 des entreprises ayant publié le plus d'offres d'emploi dans l'échantillon analysé :

Entreprise	Nombre d'offre
CNRS	35
Safran	34
CEA	19
Inria	14
Inserm	11

Tableau 1: Top 5 des entreprises ayant publié le plus d'offre

Ces résultats montrent une forte présence d'entreprises issues de la recherche scientifique (CNRS, CEA, Inria, Inserm), ce qui reflète le dynamisme de ce secteur en France, particulièrement en termes de recrutements spécialisés. L'industrie aéronautique et technologique est également représentée avec Safran, un acteur clé dans ce domaine.

Le top 10 des entreprises avec les meilleures évaluations (note maximale de 5.0) met en avant des entreprises telles que SEE YOU SUN, Founders Hub France et Chanel, témoignant de la reconnaissance des employés pour des structures innovantes, flexibles ou prestigieuses. Des entreprises moins connues comme JUNGHANS T2M SAS, Rtone et Lenstra, qui se distinguent également par une excellente satisfaction des employés. Ces entreprises, bien que parfois moins prolifiques en termes d'offres, présentent une image de marque forte et un environnement de travail apprécié, attirant ainsi des profils talentueux.

Les entreprises les plus actives en termes de recrutements, notamment dans les secteurs public et industriel, illustrent les besoins importants dans la recherche et les technologies en France. En parallèle, les entreprises les mieux notées traduisent un engagement fort en faveur de la qualité de vie au travail et de la satisfaction des employés, ce qui peut être un atout pour attirer des candidats qualifiés

Entreprise	Note entreprise
SEE YOU SUN	5.0
Founders Hub France	5.0
Chanel	5.0
JUNGHANS T2M SAS	5.0
Rtone	5.0
Lenstra	5.0
Centre hospitalier Joseph Imbert	5.0
MYTTRA	5.0
Aketys	5.0
ENVEA	5.0

Tableau 2: Top 10 entreprises les mieux notées

5. Analyse du salaire moyen proposé par les entreprises

Les statistiques descriptives du salaire moyen révèlent une variabilité importante sur le marché de l'emploi. Le salaire moyen annuel proposé est d'environ 31 103 € avec une médiane de 38 000 €, ce qui montre que la moitié des emplois offrent un salaire supérieur ou égal à cette valeur. L'écart-type de 20 299 € illustre une grande dispersion des salaires, reflétant la diversité des postes et secteurs analysés. Ces chiffres mettent en évidence un marché hétérogène où les opportunités varient largement selon le niveau de qualification et le secteur d'activité.

mean	31103,344928
std	20299,365599
min	1,000000
25%	3000,000000

50%	38000,000000
75%	45000,000000
max	80000,000000

Tableau 3: Statistiques sommaire de la variable salaire moyen

Les postes les mieux rémunérés, avec un salaire moyen allant jusqu'à 80 000 €, concernent des métiers stratégiques comme l'ingénierie en virtualisation (Cost Solutions) ou le management régional des ventes (evo). Les CDI dominent, reflétant une volonté de stabiliser des talents dans des secteurs techniques et spécialisés. Cependant, certains CDD, comme ceux proposés par Messilat (Architecte Cloud, 66 500 €) ou SECURINFOR (Ingénieur système, 62 500 €), restent très compétitifs. Cela illustre l'importance croissante des compétences technologiques et managériales sur le marché.

Titre	Entreprise	Salaire Moyen (€)	Contrat
Ingénieur en virtualisation, stockage et sauvegarde	Cost Solutions	80000.0	CDI
Regional Sales Manager - Professional Haircare	evo	80000.0	CDI
Regional Sales Manager - Professional Haircare	evo	80000.0	CDD
Ingénieur virtualisation stockage H/F	Groum K	73000.0	CDI
Business Developer H/F - Ingénieur d'affaires	AFTERDATA	70000.0	CDI
Architecte cloud	Messilat	66500.0	CDD
Ingénieur Full Stack Java - (H/F/X)	Arolla	65000.0	CDI
Ingénieur Systèmes et Réseaux H/F	INFORMATIS	65000.0	CDI
Ingénieur logiciel .NET (F/H/X)	AROLLA	62500.0	CDI
Ingénieur système et réseau	SECURINFOR	62500.0	CDD

Tableau 4: Top 10 entreprises qui payent le mieux

Corrélation entre le salaire moyen et la note des entreprises

Le coefficient de corrélation entre le salaire moyen et la note des entreprises est de **-0,07**. Cette faible corrélation négative montre qu'il n'existe pratiquement aucune relation significative entre la satisfaction des employés (ou la réputation de l'entreprise) et les salaires moyens proposés. Cela pourrait suggérer que des facteurs tels que la culture d'entreprise ou les

conditions de travail jouent un rôle plus important dans la notation des entreprises que la rémunération directe.

6. Analyse de la répartition géographique des offres d'emploi

Le télétravail ne représente que 1,92 % des offres d'emploi, soulignant son faible développement malgré les évolutions récentes vers des modes de travail plus flexibles.

En ce qui concerne les autres opportunités, la répartition géographique montre une forte concentration dans les grandes métropoles, avec Paris en tête (111 offres), suivie par Toulouse (29 offres), Grenoble et Lyon. Comme le montre la cartographie ci-dessous, les offres se dispersent progressivement vers des villes secondaires et certaines localités plus petites, bien qu'elles y soient nettement moins nombreuses. Cette analyse reflète une centralisation des opportunités, particulièrement dans les grands pôles économiques.



Figure 3: Cartographie montrant les régions avec plus d'offres

Visualisation interactive à travers un tableau de bord

Dans le cadre de ce projet, un tableau de bord interactif a été développé sur Power BI afin de visualiser et d'analyser les données récoltées. Ce tableau de bord met en lumière plusieurs aspects clés du marché de l'emploi en France. Il offre une représentation claire de la distribution des différents types de contrats, permettant d'identifier les tendances en matière de flexibilité et de stabilité dans les offres d'emploi. De plus, une analyse des entreprises les plus actives sur le marché est proposée à travers un classement des 20 entreprises publiant le plus d'offres. Cette visualisation aide à mieux comprendre les acteurs dominants dans le secteur. En complément,

une exploration des postes les mieux rémunérés en France a été réalisée, fournissant des informations précieuses sur les opportunités professionnelles les plus lucratives. Par ailleurs, le tableau de bord met également en avant les entreprises offrant les salaires les plus élevés, ce qui permet de repérer celles qui valorisent le plus leurs collaborateurs. Enfin, une analyse géographique présente les salaires moyens par localité, offrant un aperçu des disparités salariales entre les différentes régions. Ces visualisations interactives permettent de croiser les données et de tirer des conclusions pertinentes, rendant le tableau de bord non seulement informatif, mais également stratégique pour une prise de décision éclairée.

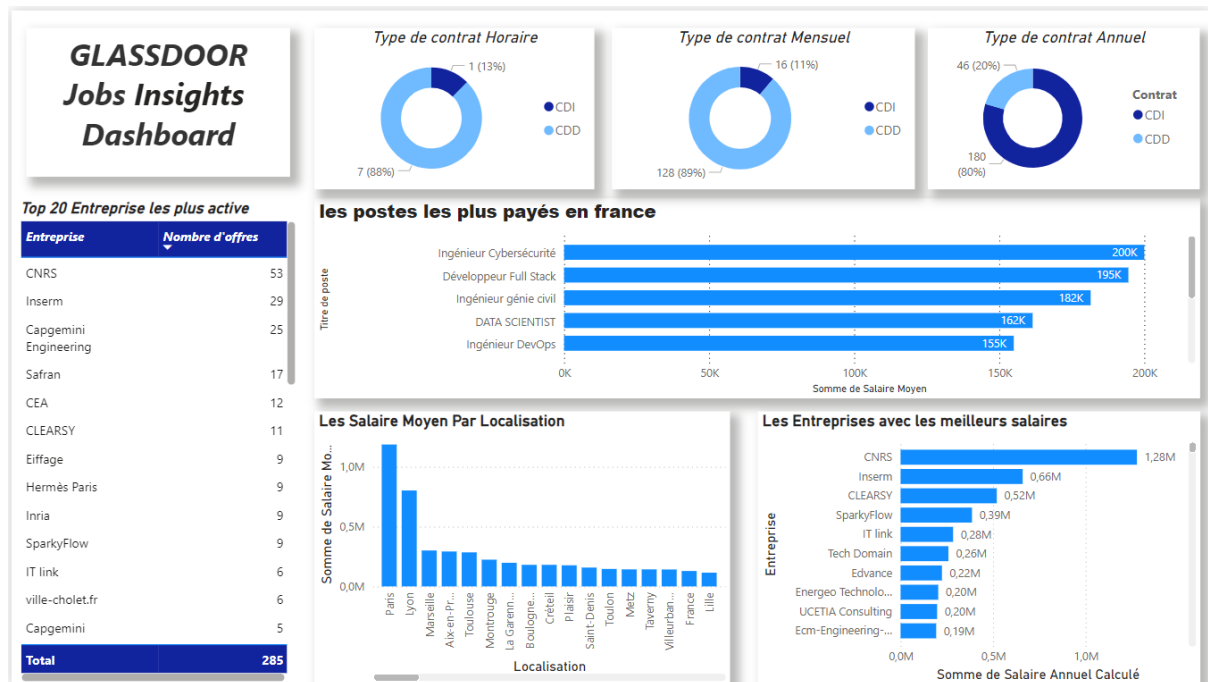


Figure 4: Dashboard interactif créé sur Power BI

Conclusion et recommandations

Ce projet a permis d'explorer le marché de l'emploi et les dynamiques économiques en France. Grâce à un processus de collecte, de traitement et d'analyse de données issues du site Glassdoor, nous avons pu extraire des informations précieuses sur les offres d'emploi, notamment sur les salaires, les entreprises, les localités et les compétences recherchées. La visualisation interactive réalisée avec Power BI a offert une perspective claire et engageante sur la distribution des types de contrats, le classement des entreprises les plus actives, les postes les mieux rémunérés, ainsi que les régions offrant les salaires les plus compétitifs. Ces résultats fournissent un outil précieux pour mieux comprendre les dynamiques du marché du travail et aider les décideurs, entreprises et chercheurs d'emploi à prendre des décisions éclairées. Ce travail a également mis en avant l'importance des technologies modernes telles que le web scraping et l'analyse de données pour répondre à des problématiques complexes. Cependant, nous avons également constaté les défis inhérents à ces méthodes, notamment le nettoyage des données, le respect des limites éthiques et juridiques, ainsi que l'interprétation des résultats dans un contexte économique et social plus large.

En se basant sur nos analyses, nous formulons les recommandations suivantes à l'endroit des offreurs et des demandeurs d'emploi en France :

✓ **Adapter les stratégies de recrutement et de recherche d'emploi aux réalités régionales**

Les disparités salariales entre localités observées dans les analyses invitent les entreprises à repenser leurs stratégies d'attractivité pour attirer les talents dans des zones moins compétitives. De même, les chercheurs d'emploi gagneraient à prioriser des régions offrant un meilleur rapport opportunités/rémunération.

✓ **Encourager la compétitivité des entreprises peu actives**

Le classement des entreprises les plus actives révèle que certaines structures dominent largement le marché. Les entreprises moins représentées pourraient tirer parti de ces informations pour ajuster leur positionnement, renforcer leur marque employeur et proposer des avantages compétitifs, notamment sur le plan des rémunérations.

✓ **Investir dans les postes les plus demandés et les mieux rémunérés**

L'identification des postes les mieux payés ou les plus proposés peut servir de guide stratégique aux professionnels en quête de reconversion ou de montée en compétence. Les résultats

suggèrent qu'un effort particulier devrait être fait pour former les talents locaux aux métiers en tension ou émergents.

✓ **Réévaluer les politiques de contrats de travail**

La répartition des types de contrats met en évidence une prédominance de certains types (CDI notamment). Les employeurs pourraient diversifier leurs propositions contractuelles (ex : contrats flexibles ou partiels) pour répondre aux besoins spécifiques de certaines catégories de travailleurs ou secteurs d'activité.