

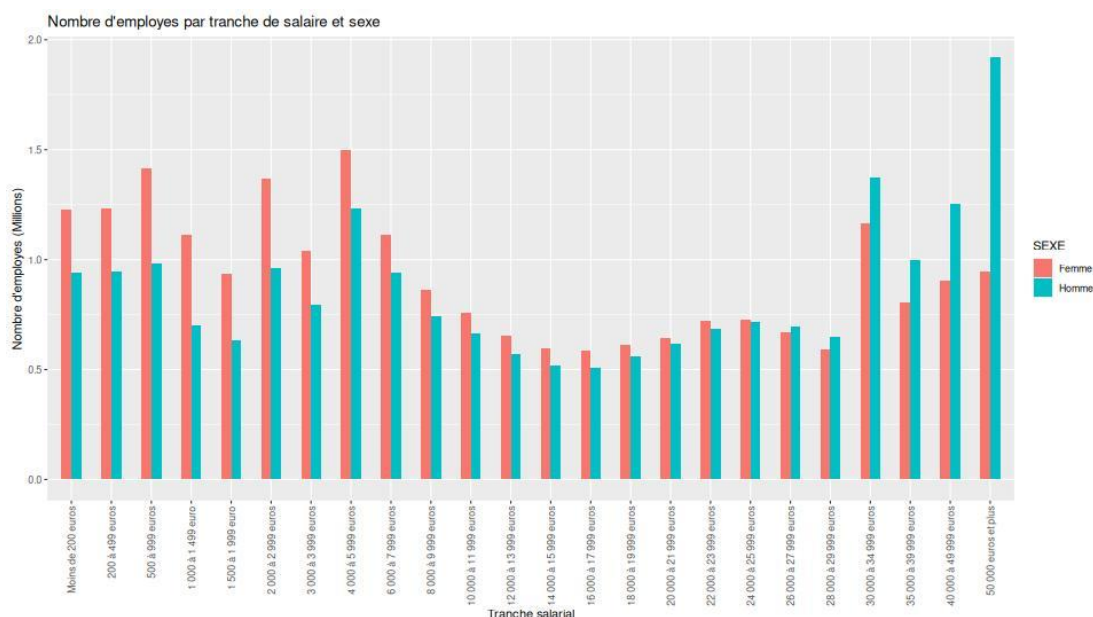
Rapport d'analyse exploratoire sur les postes et revenus français en 2020

Introduction

Pour cette analyse exploratoire, nous avons décidé d'étudier différents aspects des postes et revenus français sur l'année 2020. Ces données sont disponibles sur le site de l'INSEE à travers le fichier "postes" : <https://www.insee.fr/fr/statistiques/6524154>. Notre analyse a pour but de mettre en évidence les différentes répartitions que ce soit sur les revenus ou encore les catégories socio-professionnelles en fonction de différents critères. Le dataset représente un échantillon au plus proche de 1/12 de la population salariée en France (42.8 millions). Le code pour générer les graphiques est disponible sur ce github: <https://github.com/vitormaco/analyse-postes>.

Des proportions différentes de salaires en fonctions du sexe

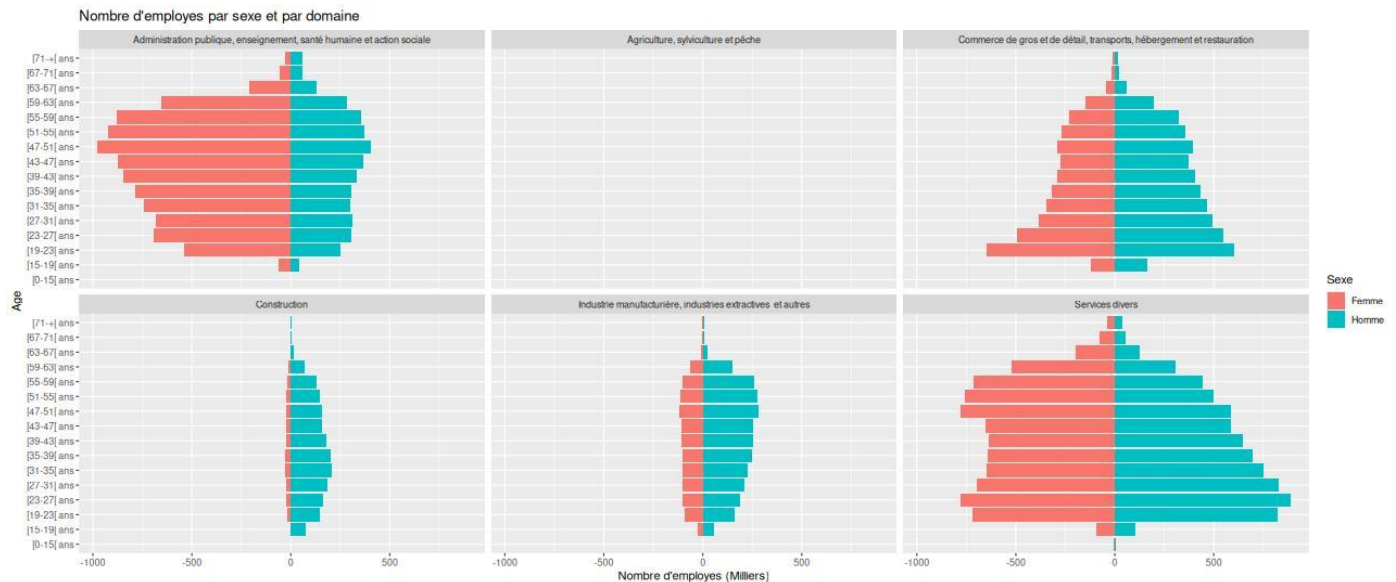
Le graphique suivant présente le nombre de personnes dans le dataset appartenant à une tranche de salaire. Les tranches de salaires permettent d'avoir une vision plus claire des répartitions. Cet histogramme différencie deux populations par un code couleur : les hommes et les femmes.



La première observation qui peut être faite est que le nombre d'hommes ayant des revenus supérieurs à 30 000 euros par an est clairement supérieur au nombre de femmes. Cette différence s'accroît dans les tranches de salaires de 50 000 euros par an et plus. Nous pouvons aussi observer le phénomène inverse pour toutes les tranches de salaires inférieures à 10 000 euros par an. Les proportions homme/femme sur toutes les tranches de salaire entre 10 000 euros et 30 000 euros par an restent à peu près équivalentes.

Une proportion différentes parmi domaines selon le sexe et l'age

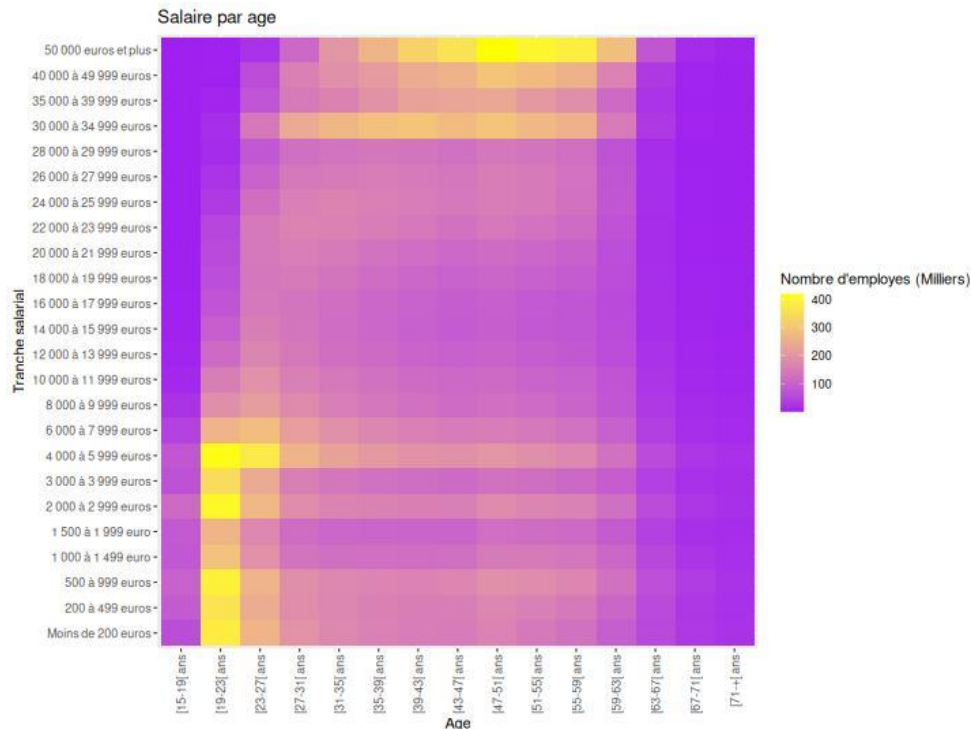
Le graphique suivant présente la quantité de personnes travaillant dans chaque domaine en fonction du sexe et de la tranche d'âge. La différenciation entre hommes et femme est faite par le code couleur du graphique. En ordonnées, nous avons une stratification en fonction de la tranche d'âge et en abscisse le nombre de personnes comptées.



Nous pouvons observer que les proportions sont complètement différentes selon le domaine et le sexe. En effet, des domaines comme "l'Agriculture, sylviculture et pêche" et reflètent la très faible quantité de personnes vivant en France et travaillant dans ces domaines. Par rapport au sexe, nous pouvons observer que des domaines tels que : "Industrie manufacturière, industries extractives et autres" et "Construction" sont, quel que soit l'âge, dominés par les hommes. Contrairement, les domaines comme "Administration publique, enseignement, santé humaine et action sociale" sont dominés par les femmes indépendamment de l'âge. Nous pouvons observer que la répartition selon les tranches d'âge dans quelques domaines est différente pour les deux sexes, les domaines des "Services divers" ont une tendance décroissante du côté des hommes alors que chez les femmes, la tendance est à peu près constante, par exemple. Une autre tendance notable est que le nombre de personnes et notamment de femmes présentes dans les domaines tels que : "Administration publique, enseignement, santé humaine et action sociale" est en baisse par rapport aux générations antérieures. L'inverse se produit pour les domaines de : "Commerce de gros et de détail, transports, hébergement et restauration" qui sont en croissance par rapport aux générations antérieures.

L'âge et le salaire

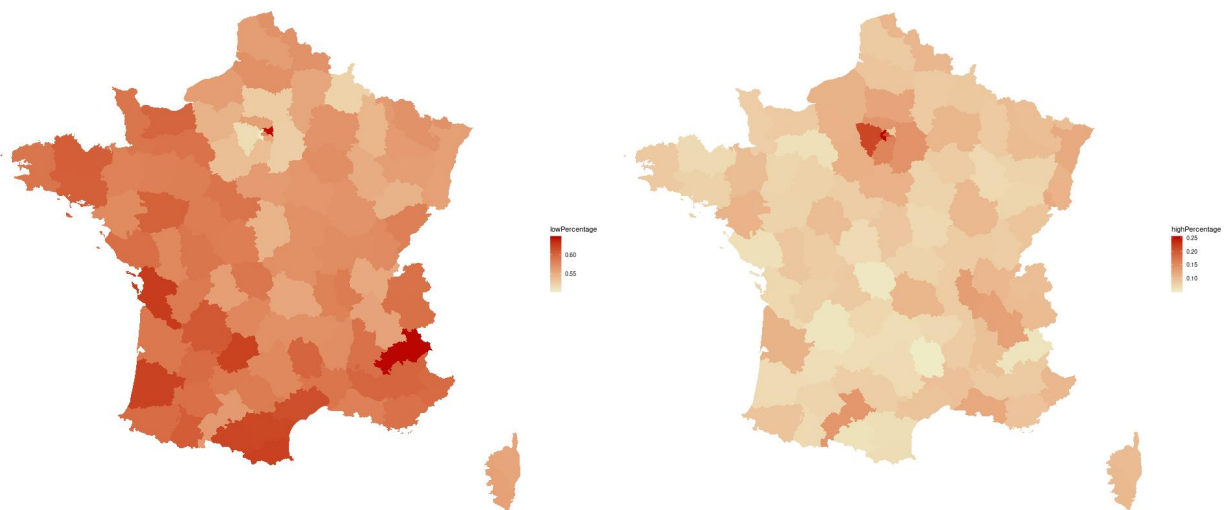
Le graphique de chaleur suivant présente le nombre de personnes en fonction du salaire et de leur tranche d'âge. En abscisse, nous avons la tranche d'âge et en ordonnées la tranche de salaire brut par an. Les couleurs vont du violet pour les nombres les plus bas au jaune pour les nombres les plus hauts.



La première tendance que nous pouvons observer à partir de ce graphique est qu'en fonction de l'âge, le salaire a tendance à augmenter. Cependant, ce n'est pas la seule observation qui peut être faite à partir de ce graphique. Nous pouvons observer une très grosse concentration de salaires brut annuels inférieurs à 6 000 euros pour les jeunes en dessous de 27 ans; une certaine concentration de personnes ayant entre 39 et 63 ans ayant des salaires supérieurs de 50 000 euros par an; ainsi qu'une grosse partie des personnes ayant entre 28 ans et 63 ans gagnent entre 30 000 et 35 000 euros par an. Cette concentration de personnes sur une certaine tranche de salaire indépendamment de l'âge est aussi présente pour les salaires entre 4 000 et 6 000 euros annuels. Nous pouvons conclure que même si la tendance du salaire en fonction de l'âge est haussière, beaucoup de personnes ne restent pas sur la même tranche salariale indépendamment de leur âge.

Répartition des revenus en fonctions des départements en France

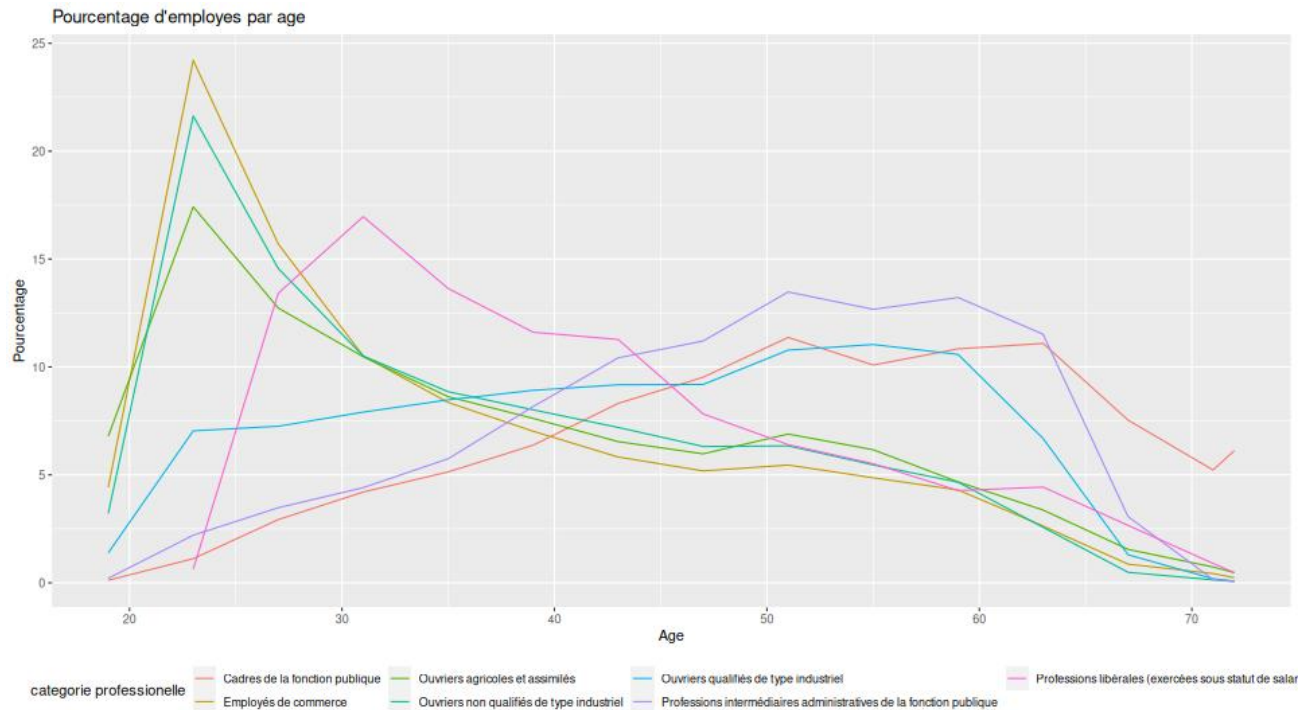
Nous avons utilisé pour dessiner ces cartes un fichier geojson disponible sur le site : <https://france-geojson.gregoireddavid.fr/>. Ces cartes représentent le pourcentage de personnes selon un critère lié au revenu. La carte de gauche présente le pourcentage de personnes par rapport aux habitants du département en dessous de 16 000 euros bruts annuels et la carte de droite ceux au-dessus de 40 000 euros bruts annuels.



Nous pouvons observer que les concentrations sont très différentes en fonction des départements. En effet, les départements ayant les plus grandes villes françaises ont tendance à avoir une plus forte concentration de personnes aisées que les autres. À l'opposé, des départements moins peuplés ont souvent un important pourcentage de personnes avec des salaires moins élevés comme par exemple les Hautes-Alpes. Nous observons aussi qu'il y a des inégalités très fortes aux alentours de Paris. En effet, les Yvelines, les Hauts-de-Seine ont un fort pourcentage de personnes aisées alors que la Seine-Saint-Denis a un très fort pourcentage de personnes avec de faibles revenus. Nous observons aussi que la concentration de personnes avec les revenus les plus faibles est majoritairement à l'Ouest, le Sud et le Sud-est de la France.

Evolution des catégories socio-professionnelles en fonction de l'âge

Le graphique ci-dessous représente le pourcentage de personnes appartenant à différentes catégories socioprofessionnelles en fonction de leur âge. Seulement quelques catégories ont été choisies pour faciliter la lecture et parce qu'elles laissaient apparaître des tendances.



Nous pouvons observer sur ce graphique que certaines catégories socio-professionnelles comme les employés de commerce, les ouvriers non qualifiés de type industriel ainsi que les ouvriers agricoles et assimilés ont tendance à être constitués de personnes jeunes et qui s'en vont avec l'âge. Les pourcentages d'employés par âge sur ces catégories diminuent de manière quasi-linéaire après un pic chez les personnes aux alentours de 23 ans. Nous pouvons observer le même phénomène, mais avec un pic aux alentours de 32 ans pour la catégorie représentant les professions libérales. Il est intéressant de noter que les ouvriers qualifiés de type industriel augmentent presque proportionnellement à la diminution des ouvriers non qualifiés de type industriel ce qui amène à penser que ces derniers passent des qualifications avec le temps. Les pourcentages des deux autres catégories socio-professionnelles augmentent selon l'âge des personnes. Le nombre de cadres de la fonction publique et les professions intermédiaires administratives de la fonction publique augmente avec l'âge.

Conclusion

En conclusion, nous pouvons observer que la répartition des salariés en France selon les domaines est fortement influencée par l'âge et par le sexe. Nous pouvons conclure aussi que quelques secteurs sont délaissés par les populations jeunes. Vis à vis des revenus des salariés en français, nous pouvons conclure que malgré des tendances haussières avec l'âge, beaucoup de salaires restent sur un même salaire indépendamment de leur âge. Un autre facteur qui influence les statistiques françaises est la concentration de personnes dans différentes régions de France selon leur salaire. Cela sera sûrement corrélé à la variation d'âge selon les différentes régions.