

DIPLOME UNIVERSITAIRE DATA ANALYST

UE RÉDACTION D' UN RAPPORT D' ANALYSE

Devoir n° 2 :
Analyse descriptive des thèses soutenues en France de
1984 à 2018

MOISE EHIMIGAYE SENGHOR



Année Universitaire : 2024-2025

Table des matières

Table des figures	1
Liste des tableaux	2
I Introduction	3
II Méthode	4
II.I Source des données	4
II.II Data Wrangling	6
II.III Méthodologie et outils d'analyse	7
III Résultats	8
III.I Évolution globale des thèses soutenues	8
III.II Analyse des thèses soutenues par Université	8
III.III Analyse des thèses soutenues par discipline	10
III.IV Analyse des thèses soutenues en fonction du genre	12
IV Discussion	14
IV.I Étude multivariée des soutenances de 1984-2018	15
IV.II Limites et perspectives	16

Table des figures

1 : Évolution du nombre de thèses soutenues annuellement en France de 1984 à 2018.	8
2 : Répartition du nombre total de thèses soutenues pour les universités ayant plus de 1% du nombre total de soutenances (1984-2018).	10
3 : Évolution du pourcentage de thèses soutenues par discipline et par année (1984-2018).	12

4 :	Évolution du nombre de thèses soutenues annuellement en fonction du genre de l' auteur (1984-2018).	13
5 :	Évolution du nombre thèses soutenues par discipline et par genre (1984-2018).	14

Liste des tableaux

1 :	Description des variables du jeu de données des thèses soutenues en France	6
2 :	Nombre total de thèses soutenues et pourcentage pour les 20 premières Universités Françaises (1984-2018)	9
3 :	Répartition des soutenances de doctorats par discipline en France de 1984 à 2018	11
4 :	Répartition des thèses soutenues entre 1984 et 2018 en fonction du genre de l' auteur	13

I Introduction

La France, pays des lumières, bénéficie d'une remarquable notoriété en matière de recherche et d'innovation. L'organisation Mondiale pour la Propriété Intellectuelle à travers son Indice Mondial de l'innovation la classe 11ème au niveau mondial en 2023 ¹.

La recherche constitue un levier économique important pour le développement et son amélioration est source d'externalités positives. L'Université, lieu d'acquisition et de perfection continu de la connaissance, en est l'un des principaux catalyseurs. Cette assertion est renforcée par un rapport de l'RNCD ² de la même année qui précise que "Les encadrants de doctorants estiment, à **74%** en Sciences et Technologies, **69%** en Sciences du Vivant et Environnement, et **24%** en Sciences Humaines et Sociales, qu'entre la moitié et la totalité de la production scientifique de leur laboratoire est associée à un projet doctoral."

Ce paradigme justifie l'importance de l'approfondissement de la maîtrise des contours de la recherche doctorale sous nos cieux. C'est à ce titre que nous nous proposons d'étudier l'évolution de cette dernière en concentrant notre laïus sur les thèses soutenues de 1984 à 2018.

Il faut reconnaître que la question du doctorat est à l'origine de plusieurs mutations du paysage normatif de l'enseignement supérieur. En effet le législateur à travers plusieurs réformes successives a essayé d'améliorer les conditions des doctorants. Nous pouvons citer la réforme LMD (Licence, Master, Doctorat) de 2022, l'Arrêté du 7 août 2006 relatif à la formation doctorale et plus précisément la Charte des thèses, l'Arrêté du 25 mai 2016 prenant en compte l'insertion des futurs diplômés, l'Arrêté du 26 août 2021 et un dernier arrêté du 26 août 2022.

En dehors de ces aspects juridiques, il existe un florilège d'articles de presse sur la question, des rapports ainsi que des articles scientifiques. Les angles d'analyse sont pléthores, le genre, la discipline, les écoles et universités d'origine, les langues d'écriture...

Il faut cependant reconnaître que la majorité de ces écrits examinent le sujet de façon parcellaire. C'est pour cette raison que nous essaierons humblement

1. OMPI, (2023, 27 septembre), Indice Mondial de l'innovation. https://www.wipo.int/pressroom/fr/articles/2023/article_0011.html

2. RNCD, (2023, 1 décembre) Le doctorat en France. Regards croisés des doctorants et de leurs encadrants. https://doctorat.u-bordeaux.fr/application/files/8917/0170/7124/Enquete_RNCD_2023_VF.pdf

d'étudier ce thème de manière allusif et plus global afin d'en avoir une représentation synthétique avec différents critères.

Ce processus nous permettra de faire un examen des thèses de doctorat en France avec une temporalité assez longue tout en la couplant à une vision multivariée. La combinaison de ces deux fils conducteurs devraient nous fournir un aperçu exhaustif et dynamique de l'écosystème de la recherche universitaire. A terme, cela pourra jouer positivement sur la significativité de nos futures hypothèses et conclusions.

Pour se faire, nous essaierons de répondre à plusieurs questions :

- Comment a évolué le nombre de thèses soutenues de 1984 à 2018 ?
- Quelles sont les universités les plus prolifiques ?
- Comment le nombre de thèses a évolué selon les disciplines ?
- Quelle est l'influence du genre sur la progression du nombre de thèses soutenues ?

Après avoir défini notre feuille de route, il nous sera salutaire de circonscrire les moyens à notre disposition afin d'atteindre nos objectifs. Nous aurons donc le loisir de commencer par une revue de nos inputs, les données à notre disposition, qui seront le point de départ de notre travail. Ensuite, nous concentrerons notre propos sur les outils utilisés afin d'assainir notre base informationnelle et enfin la méthodologie choisie aussi bien pour l'épure de notre data que pour son exploitation.

II Méthode

II.I Source des données

Comme précisé précédemment la qualité de notre travail dépendra de la qualité de l'information à notre disposition. Nous baserons celui-ci sur un dataset fourni par le responsable de notre formation Mr Cisel, spécialiste des sciences de l'éducation et de la data science ayant à son effectif de nombreuses recherches et publications sur les soutenances de thèses en France. La base de données provient d'une extraction des données brutes qu'il a eu à utiliser avec son équipe dans le cadre de notre sujet.

Cette dernière a été effectuée grâce à un travail de sélection sur le site these.fr

(<https://theses.fr/>), une plate-forme nationale répertoriant les doctorats soutenus en France. On peut légitimement supposer que ce travail de collecte a dû se faire grâce à des techniques permettant de recueillir des renseignements à partir des sites internet comme le webscraping ¹.

Le jeu de données est composé de 448047 soutenances décrites par 23 variables relatives au thème de recherche, à l’auteur, l’encadreur et l’établissement de l’auteur.

La base de données est un fichier sous format CSV ² avec les différentes variables énumérées dans le tableau suivant. Après l’avoir consulté nous pourrions nous pencher sur la méthodologie utilisée dans nos travaux. Au préalable, nous insisterons sur le processus de pré-traitement des données.

1. Le web scraping, parfois appelé harvesting ou en français moissonnage, est une technique de « récupération et organisation automatisées des données Web. *Wikipedia*.

2. CSV : Comma-separated values est un format texte ouvert représentant des données tabulaires sous forme de valeurs séparées par des virgules. *Wikipedia*.

TABLE 1 – Description des variables du jeu de données des thèses soutenues en France

N°	Nom Variable	Type Variable
1	Unnamed :0	int64
2	Auteur	object
3	Identifiant auteur	object
4	Titre	object
5	Directeur de thèse	object
6	Directeur de thèse (nom prénom)	object
7	Identifiant directeur	object
8	Établissement de soutenance	object
9	Identifiant établissement	object
10	Discipline	object
11	Statut	object
12	Date de première inscription en doctorat	object
13	Date de soutenance	object
14	Year	float64
15	Langue de la thèse	object
16	Identifiant de la thèse	object
17	Accessible en ligne	object
18	Publication dans theses.fr	object
19	Mise à jour dans theses.fr	object
20	Discipline	object
21	Genre	object
22	établissement_rec	object
23	langue_rec	object

II.II Data Wrangling

Une première lecture succincte et un aperçu du dataframe mets en exergue l'inutilité de la colonne 'Unnamed :0' car elle représente un index. Nous avons donc fait le choix de l'élaguer.

A première vue, seules quelques variables ont un taux de données manquantes représentatif :

- Date de première inscription en doctorat 85.7%
- Identifiant auteur 29.1%

- Langue_rec 14.3%
- Year 12.7%
- Date de soutenance 12.7%
- Identifiant établissement 3.8%.

En corrélant, les variables précitées à nos axes de recherche nous remarquons que nous avons certaines d'entre elles qui peuvent être utilisées en toute sécurité car le taux de données manquantes est quasiment nul. Il s'agit des colonnes 'Etablissement de soutenance', 'Genre', 'Discipline_prÃdi'.

Par contre, le cas de la variable 'Date de soutenance' pourrait être source de problèmes car l'objet de l'étude concerne les thèses soutenues. De ce fait, toutes les thèses dont la date de soutenance ne sont pas connues sont à éliminer. Se faisant, nous nous retrouvons avec un nombre de 390961 thèses soutenues.

Les étapes suivantes ont été de transformer le format de la variable date en format datetime pour faciliter sa manipulation et de circonscrire les années de soutenance entre 1984 et 2018. Ces dernières manipulations nous ont permis d'obtenir une base nettoyée que nous pourrions utiliser pour créer du contenu.

II.III Méthodologie et outils d'analyse

Dans le cadre de notre travail, nous avons eu à utiliser le langage de programmation Python via l'environnement de travail Jupyter. Étant donné que nous sommes dans le champ de l'analyse de données, les packages de manipulation usuels tels que Pandas et Numpy ont été exploités de même que Plotly, Seaborn et Matplotlib pour les visualisations.

La procédure utilisée a été la même pour tous les axes de recherches. Nous avons commencé par travailler sur les variables pertinentes ensuite nous les avons transformées et enrichies pour obtenir des datasets personnalisées. Ces derniers ont ensuite servi à étudier le phénomène en fonction des différentes problématiques.

Une fois cette étape validée nous avons essayé de créer différentes visualisations ainsi que plusieurs tableaux en fonction de nos objectifs. Le but étant de faire le tri du contenu en choisissant les aspects les plus attrayants à mettre dans notre futur rapport. Lors de ce ciblage, nous avons identifié plusieurs clés de lecture. D'abord, une lecture dynamique ou statique de la réflexion, ensuite une lecture univariée ou bivariée et enfin, dans la mesure du possible, une corrélation entre deux ou plusieurs problématiques.

Cette démarche a abouti à plusieurs résultantes que nous essaierons de cerner dans la suite de notre propos.

III Résultats

III.I Évolution globale des thèses soutenues

Nous observons, dans la Figure 1, une augmentation continue et significative de 1984 à 1988 du nombre de thèses soutenues. Après cette ascension, il y a une certaine période de stabilité de relative jusqu'en 1991. Elle est suivie d'une évolution positive assez timide jusqu'en 1994, d'une baisse brusque en 1995 et d'une faible expansion jusqu'en 1997. De 1998 à 2000 une période de statu quo est attenante à une longue période d'augmentation continue et uniforme qui s'arrête en 2012. Par la suite, nous assistons à une dernière période de baisse uniforme et peu significative jusqu'en 2018.

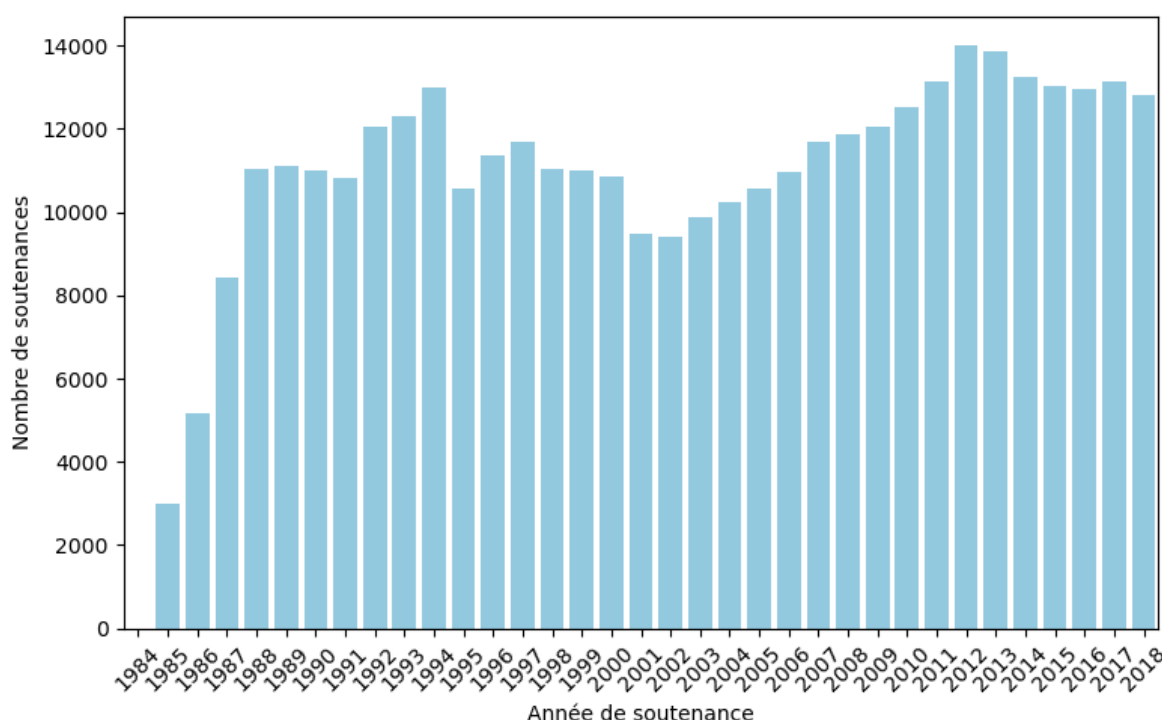


FIGURE 1 – Évolution du nombre de thèses soutenues annuellement en France de 1984 à 2018.

III.II Analyse des thèses soutenues par Université

Au total 535 établissements de l'enseignement supérieur sont concernés par les soutenances de doctorats. Les 20 premiers d'entre eux trustent 45.3%

des soutenances(Figure 2). On peut donc en conclure que moins de 4% des universités et établissements dédiées ont eu la charge de 45% des thèses soutenues.

TABLE 2 – Nombre total de thèses soutenues et pourcentage pour les 20 premières Universités Françaises (1984-2018)

N°	Université	Nombre de thèses soutenues	Pourcentage (%)
1	Paris 6	21 201	5.6
2	Paris 11	15 429	4.1
3	Toulouse 3	11 248	3.0
4	Paris 7	11 101	2.9
5	Lyon 1	10 522	2.8
6	Paris 1	10 457	2.8
7	Paris 4	8 301	2.2
8	Paris 5	7 682	2.0
9	Rennes 1	7 225	1.9
10	Nantes	7 102	1.9
11	Lille 1	6 657	1.8
12	Montpellier 1	6 642	1.8
13	Bordeaux 2	6 546	1.7
14	Nice	6 439	1.7
15	Aix-Marseille 2	6 254	1.6
16	Paris 10	6 144	1.6
17	Montpellier 2	5 974	1.6
18	Grenoble 1	5 792	1.5
19	Nancy 1	5 506	1.5
20	Paris, EHESS	5 048	1.3

Une autre grille de lecture, nous permet de mettre en exergue le fait que seulement 31 Universités ont un nombre de thèses soutenues supérieur à 1% du nombre total. Ces dernières représentent 58,4% des thèses soutenues, en d'autres termes moins de 6% des Universités ont produits à peu près 58% des thèses soutenues.

Un autre fait marquant, seuls deux des établissements ne sont pas des universités publiques mais des écoles supérieures spécialisées. Il s'agit de l'INPT de Toulouse ¹ et de l'EHESS de Paris ².

Il est aussi prégnant que Paris est la ville la plus présente, sur les 10 premières universités en terme de nombre de soutenances 6 sont parisiennes (Figure 2).

1. INPT Toulouse : INP-ENSIACET (Institut National Polytechnique de Toulouse - École Nationale Supérieure des Ingénieurs en Arts Chimiques et Technologiques)<https://www.inp-toulouse.fr/fr>

2. Paris, EHESS : École des Hautes Études en Sciences Sociales)<https://www.ehess.fr/fr>



FIGURE 2 – Répartition du nombre total de thèses soutenues pour les universités ayant plus de 1% du nombre total de soutenances (1984-2018).

III.III Analyse des thèses soutenues par discipline

Entre 1984 à 2018, les doctorats soutenus en France sont répartis en 15 disciplines, avec une prépondérance des sciences fondamentales et des disciplines à caractère scientifique. A titre d'exemple, 4 des cinq disciplines les plus représentatives sont la biologie, la chimie, la médecine, et l'ingénierat. Ces dernières représentent à elles seules 55,7% du total des thèses soutenues (Table 3).

Nous constatons aussi que 3 disciplines sont peu symboliques. Il s'agit des STAPS¹, des mathématiques appliquées à l'informatique et des sciences de l'éducation. Ces dernières à elles trois ne représentent que 0.5% des doctorats soutenus.

1. STAPS : Sciences et Techniques des Activités Physiques et Sportives

TABLE 3 – Répartition des soutenances de doctorats par discipline en France de 1984 à 2018

Discipline Agrégées	Soutenances	Pourcentage (%)
Biologie	86 194	22.7
Matériaux, Milieux et Chimie	65 157	17.2
SHS	43 249	11.4
Médecine	30 296	8.0
Science de l'ingénieur	29 620	7.8
Droit et Science Politique	21 659	5.7
Langues et Littératures	21 164	5.6
Mathématiques	18 411	4.9
Économie Gestion	17 824	4.7
Psychologie	15 182	4.0
Science de la Terre	14 316	3.8
Informatique	14 237	3.8
Sciences de l'éducation	804	0.2
Mathématiques et Informatique	776	0.2
STAPS	274	0.1

En approfondissant, nous avons essayer de nous concentrer sur l'évolution du nombre de thèses soutenues par discipline, par an et en pourcentage. La Figure 3 nous a permis d'observer certains faits.

D'abord, la prépondérance des trois principales disciplines reste a peu près constante sur toute la période. En effet, la biologie concerne plus de 15% des thèses sur une partie de la période et plus de 20% sur le reste. Le phénomène est le même pour la chimie dont le pourcentage varie de 10 à 15% sur le premier tiers de la phase étudiée et entre 15 et 20% sur le reste. Les sciences humaines et sociales sont aussi dans le même moule car à part un pic en 1986 nous remarquons une petite évolution de 5 à 10% à peu près sur les 12 années suivantes, par la suite une très légère ascension uniforme tournant entre 10 et 12%.

Certaines disciplines sont assez constantes avec peu de variabilité des pourcentages annuels de soutenances : le droit et les sciences politiques, l'économie et la gestion, et les sciences de la Terre. Pour d'autres comme l'informatique, il ressort une évolution positive dans le temps même si dans la plus part des cas elle est légère.

Une autre tendance témoigne de fluctuations avec des alternances de cycles d'augmentation et de réduction du taux de soutenances. C'est le cas des langues et de la littérature, et des mathématiques.

Nous remarquons des cas particuliers tels que la médecine où nous avons une baisse très forte en trois ans ensuite on passe de moins de 5% en 1988 à plus de 30% en 90, par la suite une baisse continue est proportionnée jusqu'en

1997. En 1998 nous avons à un krach, le taux baisse jusqu'à 8%. Cette tendance baissière continue de manière équilibrée pendant 6 ans et les années restantes le taux est assez constant tournant autour de 3-4%. La psychologie aussi laisse perplexe, une phase expansionniste commence avec des apogées entre 7 et 9%, elle est brutalement abrégée par un effondrement en l'espace de 2 ans et sur le reste de la période un taux quasi égal à moins de 3%.

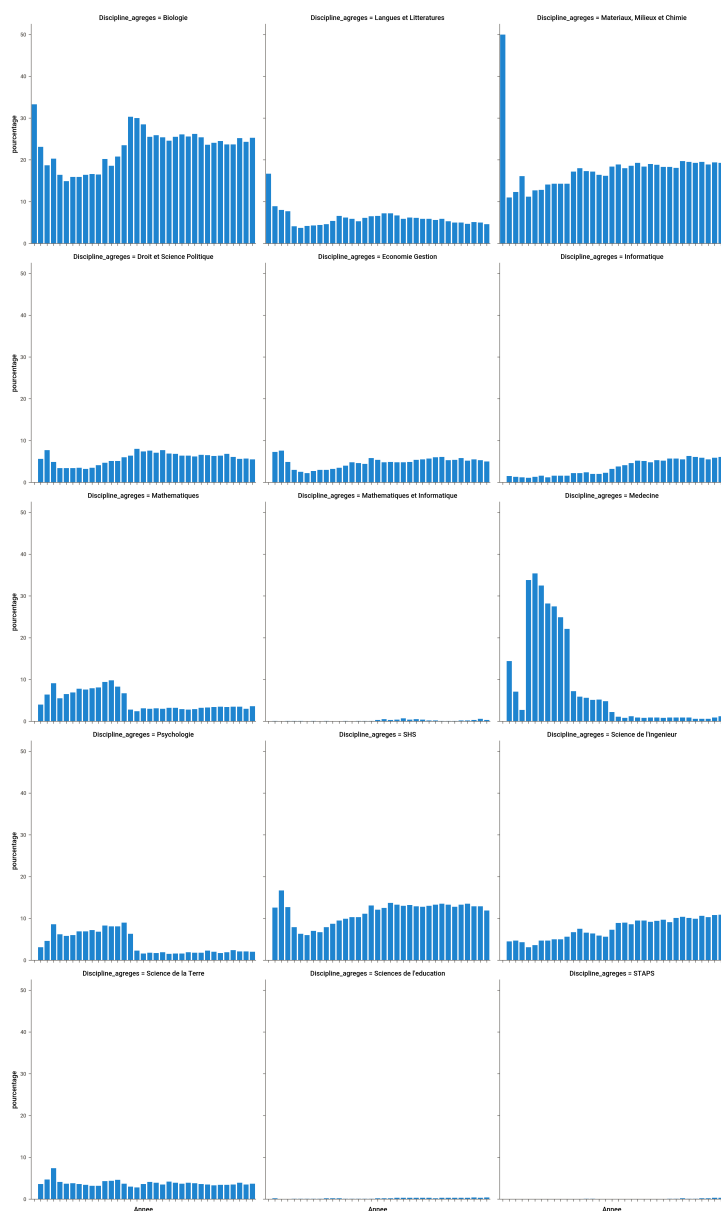


FIGURE 3 – Évolution du pourcentage de thèses soutenues par discipline et par année (1984-2018).

III.IV Analyse des thèses soutenues en fonction du genre

La Table 4 suivant mets en exergue une prépondérance des hommes dans la recherche doctorale. Nous avons choisi de garder l'itération Unknow par

soucis de fiabilité.

TABLE 4 – Répartition des thèses soutenues entre 1984 et 2018 en fonction du genre de l’auteur

Genre	Nombre de thèses	Pourcentage (%)
Female	134 797	35.6
Male	185 435	48.9
Unknown	51 578	13.6

La Figure 4 nous permet aussi de remarquer que cette primauté des auteurs de sexe masculin est assez significative et relativement constante dans toute la période d’étude. A partir de 1999, on voit des courbes assez similaires pour les hommes et les femmes. La distance entre les deux courbes est sensiblement la même sur le reste de la période.

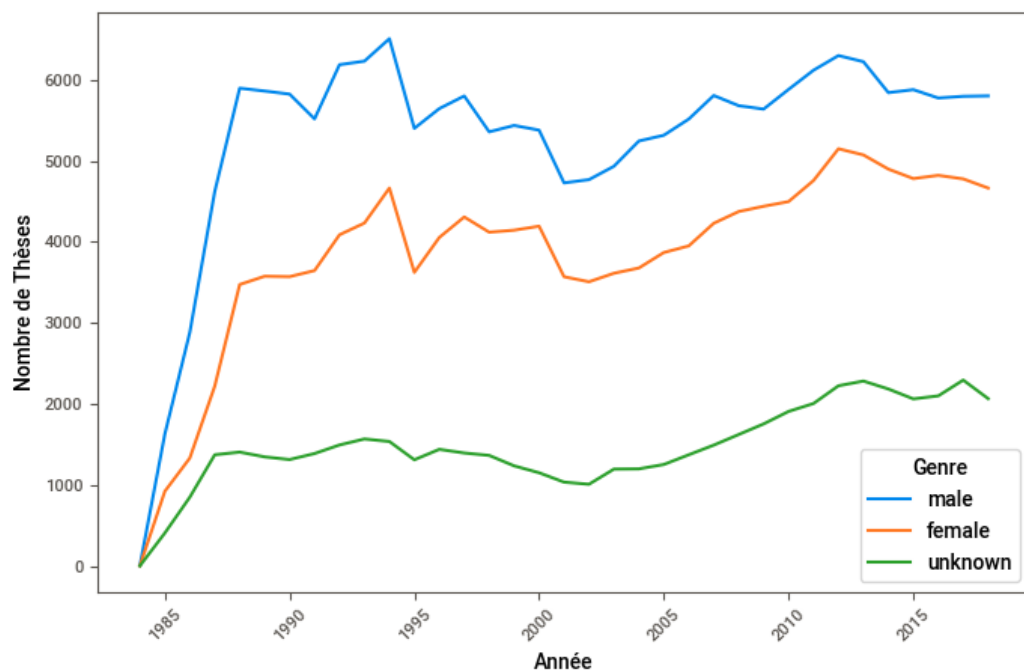


FIGURE 4 – Évolution du nombre de thèses soutenues annuellement en fonction du genre de l’auteur (1984-2018).

En observant la Figure 5, on se rend compte que la gent féminine est beaucoup moins représentée pour certaines matières telles que l’informatique, la chimie, les mathématiques, les mathématiques et l’informatique, les sciences de l’ingénieur et les sciences naturelles. Alors que pour les lettres et les langues, la psychologie ainsi que les sciences de l’éducation ces dernières années les femmes sont beaucoup plus nombreuses que les hommes.

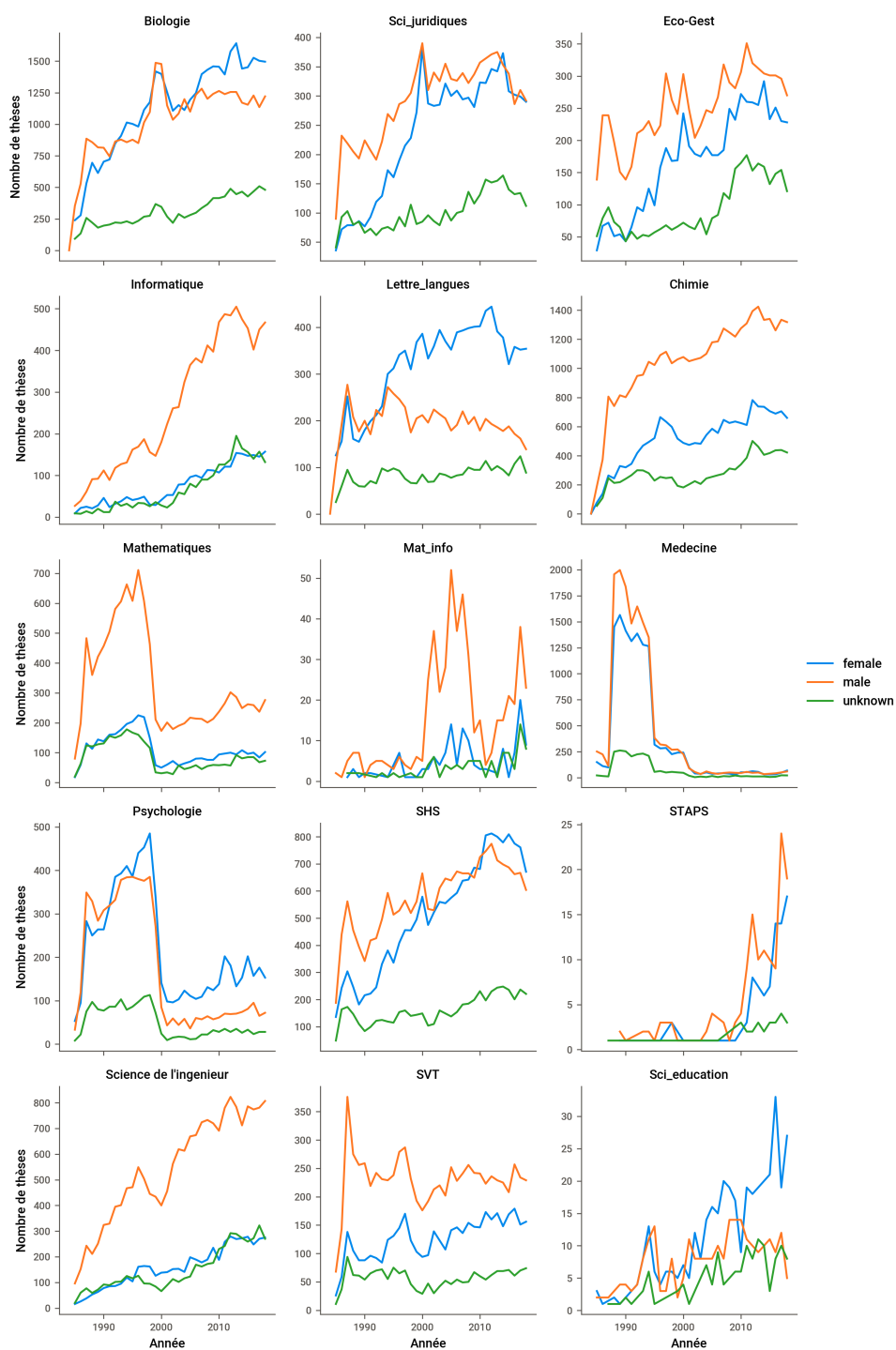


FIGURE 5 – Évolution du nombre thèses soutenues par discipline et par genre (1984-2018).

IV Discussion

Dans cette partie nous essaierons humblement de revenir sur l'interprétation des résultats obtenus avant de nous pencher sur les possibles perspectives de recherche futures.

IV.I Étude multivariée des soutenances de 1984-2018

Une lecture lacunaire des résultats obtenus sur l'évolution globale des soutenances peut sembler peu significative. Mais une vision plus globale de la question nous permet de faire des corrélations avec d'autres phénomènes sociaux. L'augmentation soutenue du nombre de doctorats de 1984 à 1988 peut être un des effets de la mise en place de la Loi Savary¹ du 26/08/1984. Cette réforme a eu pour conséquence de fusionner le doctorat de 3^{ème} cycle jugé trop court et peu qualitatif, et le doctorat d'état considéré trop long. En restant dans la même logique, nous pouvons aussi trouver un lien de causalité empirique entre l'accroissement du nombre thèses soutenues annuellement dans les années 2000 et la réforme LMD(Licence, Master, Doctorat) du début des années 2000.

Les Universités les plus anciennes datent du Moyen Age, les cinq premières sont l'Université de Paris, l'Université de Toulouse, l'Université de Montpellier, l'Université d'Avignon et l'Université d'Orléans. Les trois premières font partie des 20 universités les plus prolifiques en termes de soutenances de doctorat. La prépondérance des universités parisiennes saute aux yeux, 6 des dix premières sont de la région d'Ile de France. On peut facilement penser que cela découle de la décision de l'état de scinder l'Université de Paris en 13 universités autonomes et spécialisées. Cette mesure prise en 1970 fut une réponse aux événements de Mai 68. A l'image des résultats obtenus dans les analyses par discipline, la répartition de soutenances dans les universités parisiennes montre une présence marquante des thèses à consonance scientifique. En effet, les deux premiers établissements parisiens (Université Paris 6 Pierre et Marie Curie et Université Paris 11 Paris-Sud) sont les deux premiers au classement national. Ils ont pour domaine de spécialisation les sciences et la médecine, et les sciences et la technologie. Elles pèsent respectivement 5.6% et 4.1% soit près de 10% de total des thèses soutenues. Les quatre autres universités parisiennes sont les universités de Paris 7 Denis-Diderot, Paris 1 Panthéon-Sorbonne, Paris 4 Paris-Sorbonne, Paris 5 René Descartes, Paris 10 Nanterre. A l'instar des précédentes elles sont toutes issues de la décision de 1970.

Malgré cette scission, la région Parisienne demeure le creuset de la recherche doctorale Française, les huit universités de Paris les plus prolifiques pèsent

1. La loi du 26 janvier 1984 sur l'enseignement supérieur, dite loi Savary (du nom du ministre Alain Savary), procède à une large réforme de l'enseignement supérieur et notamment des universités.https://fr.wikipedia.org/wiki/Loi_Savary

22.5% de la totalité des doctorats.

La recherche française est fortement orientée vers les sciences fondamentales et les sciences appliquées. Les orientations de la recherche sont définies par l'état avec en tête le Ministère de l'Enseignement Supérieur, de la Recherche et de l'Innovation(MESRI¹) et d'autres institutions telles que Agence Nationale de la Recherche(ANR² ou encore le Centre National de la Recherche Scientifique(CNRS³). En général, ces orientations sont fixées en fonction d'objectifs environnementaux, technologiques et socio-économiques. Durant notre période d'étude ces dernières ont évoluées selon le contexte du contemporain, mais certains fils conducteurs restent constants. Il s'agit de l'environnement, la transition écologique, les technologies de l'information et l'informatique(révolution numérique et IA), les directives européennes, la technologie industrielle et enfin dans ces dernières années les problématiques sécuritaires et sanitaires. En somme ce sont des problématiques scientifiques qui ont été priorisées. On peut émettre l'hypothèse que ce sont ces faits qui sont à l'origine de la forte connotation scientifique des soutenances de thèses en France.

Cette contextualisation de la recherche est bien illustrée par une arrivée timide de nouvelles spécialités telles que les sciences de l'éducation et les STAPS dont l'entrée et la structuration dans l'enseignement datent des années 70-80. Cela explique la timide incursion qu'elles ont faites dans la recherche doctorale.

Les hommes sont plus présents dans la recherche doctorale en France. Ce paradigme est encore plus marqué pour les matières scientifiques. Les femmes par contre sont sur-représentées dans certaines disciplines comme les sciences de l'éducation, les lettres et les langues et légèrement la biologie.

IV.II Limites et perspectives

Notre travail avait pour but de faire un panorama de la recherche doctorale française dans une période déterminée en tenant de compte de différents critères. Cette approche globalisante bien que salubre à quand même certaines limites objectives.

1. <https://www.enseignementsup-recherche.gouv.fr/fr>

2. <https://anr.fr/>

3. <https://www.cnrs.fr/fr>

D'abord la source des données et son nettoyage peuvent être remises en question sur certains aspects. Le dataset provient de notre formateur ce qui le rend relativement fiable en tenant compte de l'aura de ce dernier. Cependant nous ne nous sommes fiés uniquement à lui sans avoir de certitude sur le processus d'extraction de la data. En outre, l'extraction a été faite sur le site **these.fr** qui n'est pas le seul site dédié. Il existe d'autres sources tels que les sites de HAL¹, du SUDOC² entre autres. Il faut aussi préciser que thèses.fr a été créé en 2007 soit 33 ans après le début de notre période de référence. Ces paramètres nous poussent à nous poser des questions sur l'exhaustivité des données ainsi que sur la représentativité du site d'origine de nos informations.

Le choix de la période de référence n'a pas été expliqué alors que cela nous aurait permis de mieux éclairer notre problématique. Lors du pré-traitement des données nous nous sommes bornés à reprendre le même processus que nous avons utilisé dans un exercice précédent, proposé par notre professeur, sur la même base de données. Objectivement, nous considérons que nous aurions dû aller plus loin dans cette phase d'autant plus que nous n'avons travaillé que sur les variables qui nous semblaient opportunes compte tenu de nos questions de recherche. Nous pensons qu'un approfondissement du processus était nécessaire afin de nous assurer de la qualité de notre dataset d'exploitation.

Ensuite, la problématique de recherche tournait autour de différentes questions. Ces questions ont peut-être été trop spécifiques et trop descriptives. Nos réponses se sont limitées à une description succincte du thème sans pour autant être dans une logique de résolution de problème. D'ailleurs lors de notre discussion la majeure partie de nos conclusions et hypothèses sont empiriques et se bornent à faire le lien entre le contexte historique de l'époque et les constats ressortis de notre travail. En terme de méthodologie et d'objectivité, il sera important de revoir notre procédé de réflexion.

Malgré tous ses impairs, nous pensons que le sujet reste très intéressant. Ce rapport bien qu'il n'apporte pas de plus-value scientifique peut être la base de nombreuses recherches et approfondissements de ce sujet. Il serait bien de ne pas limiter nos futures études sur la question aux sciences de l'éducation mais de les appliquer à d'autres sciences sociales, telles que la sociologie, l'économie, l'histoire (la liste n'est pas exhaustive). En prenant

1. Hyper Articles en Ligne <https://hal.science/>

2. Système Universitaire de Documentation <https://www.sudoc.abes.fr/cbs/xslt/>

du recul, on remarque que la recherche doctorale en dehors du carcan de l'enseignement supérieur est aussi un témoin de son temps et toutes les mutations sociales ont un effet sur elle.

Cela dit en le ramenant à son essence, l'enseignement universitaire, nous remarquons de nombreuses potentialités en termes d'axes de recherche sur ce thème. A titre d'exemple et sans être exhaustif nous pouvons convoquer l'attractivité du doctorat, les dynamiques migratoires des chercheurs, les évolutions du doctorat de certaines disciplines comme la médecine, les possibilités sont infinies...

Toujours dans cette logique en tant que résidant au Sénégal, nous pensons humblement qu'il serait intéressant que des recherches soit faites dans ce domaine. Cela permettrait de diagnostiquer notre enseignement supérieur pour le rendre plus attractif et plus en adéquation avec nos besoins à l'instar de la France. Les résultats de ces différentes prospections pourraient être le point de départ d'une amélioration de notre politique de recherche en rendant la recherche universitaire de notre pays plus impactante et plus attractive.