

Projet SEP0831:
Le nombre de bénéficiaires du programme
d'investissement de l'État français (IDEX) est t'il
influencé par le domaine de recherche et le
sexe des chercheurs ?

**Master 1 Mathématiques et
Applications**

UFR Sciences Exactes et Naturelles
Parcours Statistique pour l'Évaluation et la Prévision.

PAR
ADRIEN BIKEGNE.
OUSMANE FALL.
YANNICK GNAGO.

ENSEIGNANT: M. PHILIPPE REGNAULT

2021 / 2022

Résumé

Le programme d'investissement de l'État français dans l'enseignement supérieur et de recherche(IDEX) vise à optimiser les résultats dans le domaine de recherche scientifique et de l'innovation Français. On mettra en évidence les effets qui engendrent l'investissement de l'État français(IDEX) dans l'enseignement supérieur et la recherche. Nous avons donc une base de données relative aux bénéficiaires de la prime de l'excellence scientifique de 1993 à 2012 sur-le-site. On a d'abord faire le pré-traitement de la base, puis nous avons effectué des tests statistiques afin d'avoir une appréhension sur les variables qui influencent cet investissement.

Table des matières

1 L'étude univarié	1
1.1 Le programme d'investissement. idex	2
1.2 Le sexe des bénéficiaires. sexe	2
1.3 Le domaine de recherche des chercheurs. seteur_disciplinaire	2
1.4 La profession des bénéficiaires. groupe_de_corps	3
1.5 La région des bénéficiaires. region	3
1.6 Le lien entre le pôle de de recherche et l'établissement. distance_univ	4
2 L'étude bivarié	4
2.1 Le programme d'investissement et le sexe.	4
2.2 Le programme d'investissement et le domaine de recherche.	6
3 Analyse de la Problématique.	7
4 Bibliographie	8

1 L'étude univarié

Dans cette partie, nous essayerons au mieux d'expliquer les différentes variables et nous étudierons quelques-unes. Après le pré-traitement, notre jeu de donnée est constitué de 11 variables, dont voici ci-dessous quelques définitions :

- annee: l'année d'obtention de la prime(de 1993 à 2012).
- sexe: le sexe des bénéficiaires (**femmes** et **hommes**).
- seteur_disciplinaire: le secteur de section CNU(conseil national des universités) du bénéficiaire de la prime i.e. son domaine de recherche. Il y a dans l'ensemble le **Droit et sciences économiques**, les **Lettres et sciences humaines**, la **Santé**, les **Sciences** et les **STAPS**.
- etablissement: l'école des bénéficiaires.
- groupe_de_corps: la profession du bénéficiaire de la prime. Il y a : des **Professeurs et assimilés**, des **Maîtres de conférences et assimilés**, des **particuliers**(case vide dans la dataframe).
- region : la région du laboratoire de recherche des bénéficiaires de la prime.
- beneficiaires : Le nombre de bénéficiaires selon l' **etablissement**, le **sexe**, l' **année**,
- distance_univ: Le lien entre le pôle de de recherche et un etablissement supérieur. Il y a **PRES**¹ (pour les bénéficiaires dont le pôle de recherche est rattaché à l'établissement) et **HORS PRES**(pour les bénéficiaires dont le pôle de recherche n'est pas rattaché en dehors de l'établissement).
- idex²: il y a **HORS INDEX** pour les bénéficiaires sans le programme INDEX et **INDEX** pour les bénéficiaires avec le programme INDEX.

1. PRES sont les pôles de recherche et d'enseignement supérieur.

2. L'INDEX est un programme d'investissement de l'État français dans l'enseignement supérieur et de recherche.

1.1 Le programme d’investissement. idex

Il représente ici l’effectif des bénéficiaires suivant un le programme d’investissement IDEX ou pas. Parmi les établissements des bénéficiaires primés, on recense 75.4% **HORS IDEX** et 24.6% **IDEX**.

La répartition du programme d’investissement est illustrée par la **figure 1**.

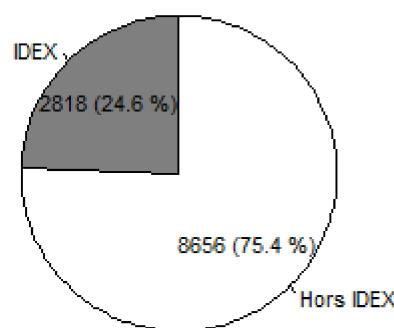


FIGURE 1 – Répartition des bénéficiaires selon le statut IDEX/HORS IDEX.

1.2 Le sexe des bénéficiaires. sexe

Il représente l’effectif du sexe des chercheurs durant cette période.

Parmi les chercheurs, on recense 62.9% d’hommes et 37.1% de femmes durant la période de 1993 à 2012 (**Voir la figure 2**).

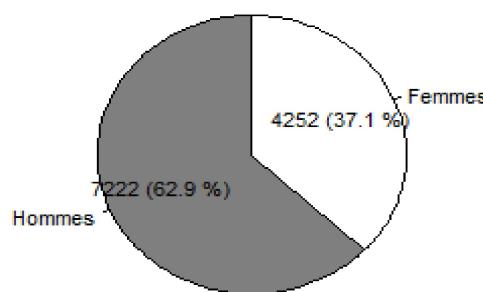


FIGURE 2 – Répartition du sexe des chercheurs.

1.3 Le domaine de recherche des chercheurs. seteur_disciplinaire

Il y a près de 50 % des chercheurs qui sont issus du domaine de la science, ensuite 21.9 % sont les lettres et sciences humaines (**Voir la figure 3**).

Il semble alors que les chercheurs en sciences sont les plus primés, cela est-il dû au programme d’investissement d’IDEX? (**Voir dans l’analyse de la problematique**)

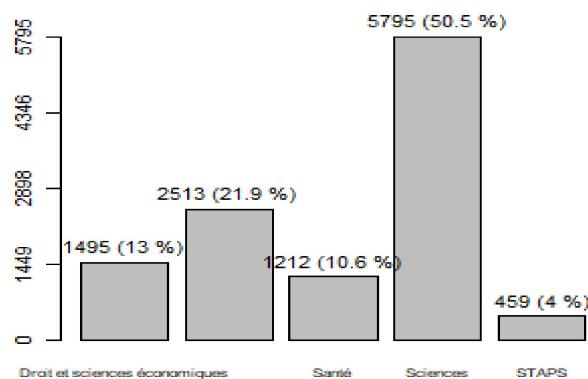


FIGURE 3 – Diagramme en barre des disciplines des bénéficiaires au prix d'excellence scientifique de 1993 à 2012.

1.4 La profession des bénéficiaires. groupe_de_corps

Nous constatons qu'il y a 63.8 % des bénéficiaires qui sont des Professeurs, 36.0 % Maîtres de conférences et 0.2 % particuliers. Dans l'ensemble, la majorité des bénéficiaires est dans le corps de l'enseignement supérieur, ce qui est d'autant plus logique. (**Voir la figure 4**)

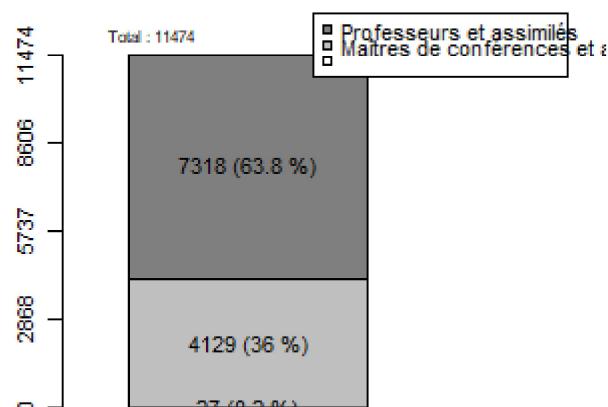


FIGURE 4 – Diagramme en Bande de la profession des bénéficiaires primés au prix d'excellence scientifique.

1.5 La région des bénéficiaires. region

La région des bénéficiaires la plus représentée est l'Île de france, puis celle de Rhône-Alpes(**Voir figure 5**). Cela s'explique par le fait qu'elles sont des régions de france avec le plus d'étudiants chercheurs d'après le site et qu'elles représentent les deux plus grandes régions contenant le plus de PRES³ d'après le site.

3. PRES sont les pôles de recherche et d'enseignement supérieur.

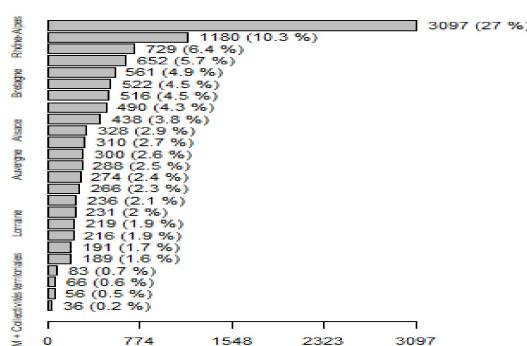


FIGURE 5 – Diagramme en barre représentant les régions des bénéficiaires.

1.6 Le lien entre le pôle de recherche et l'établissement. distance_univ

On observe qu'il y a environ 85 % des laboratoires de recherche qui sont rattachés aux établissements (**Voir la figure 6**), cela s'explique par le fait qu'il y a beaucoup de centres de recherche qui sont en partenariat avec les établissements publics comme la CNRS, l'EPST, l'INRA, l'INRIA, l'INSERM.

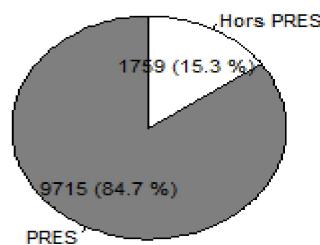


FIGURE 6 – Répartition des chercheurs en fonction du rattachement du laboratoire de recherche et l'établissement.

2 L'étude bivarié

Dans cette seconde partie, nous allons étudier le croisement de certaines variables que nous avons choisi. Parmi ces variables, nous avons effectué les croisements suivantes :

- eff_idex_sex: croisement entre les variables du programme d'investissement et le sexe.
- eff_idex_secteur : croisement entre les variables du programme d'investissement et le domaine de recherche.

2.1 Le programme d'investissement et le sexe.

2.1.1 Profil ligne

L'étude du profil ligne de ces deux variables nous montre que parmi les bénéficiaires du programme **IDeX**, il n'y a plus d' **hommes** que de **femmes** qui sont suivis par le programme d'investissement **IDeX**.

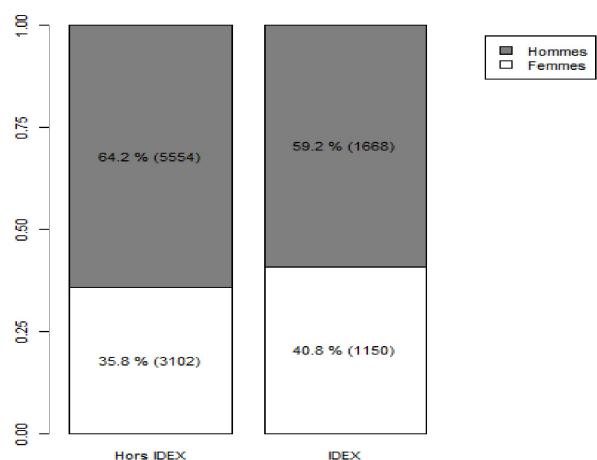


FIGURE 7 – Le profil ligne du sexe et le programme idex.

2.1.2 Profil colonne

Ici, on voit que la proportion des bénéficiaires du programme chez les **femmes** est supérieur à celle chez **hommes**.(voir figure 8)

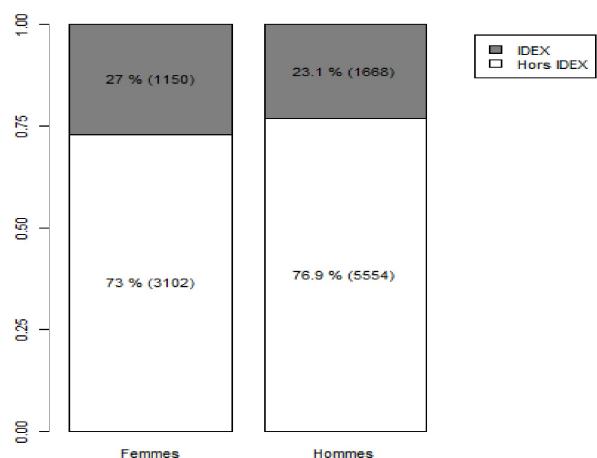


FIGURE 8 – Profil colonne des variables du sexe et la profession représentées sous forme de diagramme en barres

D'après ces graphes, on voit bien qu'il y a une différence de résultats entre les partitions dans chaque profils. Alors, il semble émerger de tendance de dépendance entre le programme d'investissement et le sexe. Vu que ces variables sont catégorielles, on effectue le test de khi-deux:

```
Pearson's Chi-squared test with Yates' continuity correction  
data: eff_idex_sexe  
X-squared = 22.325, df = 1, p-value = 2.302e-06
```

Vu la p-value, ce test est significative, il y'a donc une relation de dépendance significative entre les variables **ideX** et le **sexe**.

2.2 Le programme d'investissement et le domaine de recherche.

2.2.1 Profil ligne

L'étude du profil ligne de ce croisement montre que parmi le domaine **sciences** est le plus représenté parmi **ideX**, en effet c'est le secteur où il y'a le plus de chercheurs primés. On peut constater qu'il y'a environ 47 % des bénéficiaires de la prime d'excellence en **Sciences**(voir figure 10).

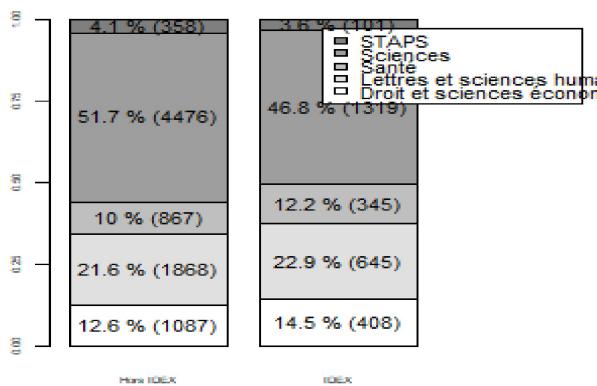


FIGURE 9 – Profil ligne représentée des variables ideX et le secteur disciplinaire des chercheurs.

2.2.2 Profil colonne

Ici, on voit que la proportion des bénéficiaires dans les différents secteurs disciplinaires est presque pareil. Il ne semble pas avoir de différence de résultats entre les partitions des profils(voir figure 11).

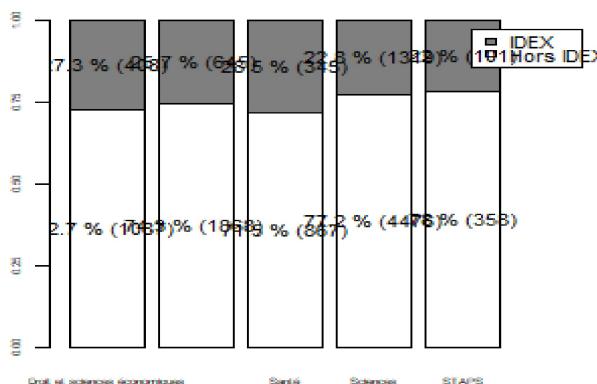


FIGURE 10 – Profil colonne représentée des variables sexe et secteur disciplinaires.

D'après ces graphes, on voit qu'il n'y a pas vraiment de différence de résultats entre les partitions des profils. Alors, il semble émerger de tendance de dépendance entre le programme d'investissement et le domaine de recherche des chercheurs. Vu que ces variables sont aussi catégorielles, on effectue le test de khi-deux:

```
Pearson's Chi-squared test  
data: eff_idex_secteur  
X-squared = 29.396, df = 4, p-value = 6.495e-06
```

FIGURE 11 – Le test de khi-2 entre le programme d'idex et le domaine de recherche des chercheurs

Vu la p-value, ce test est significative, il y'a aussi une relation de dépendance significative entre les variables **ideX** et le **secteur_disciplinaire**.

L'étude de ces deux croisements montre que l'investissement de l'État Français dans la recherche dépend fortement du domaine et du sexe des chercheurs, mais essayons maintenant de vérifier cet aspect avec les moyennes des bénéficiaires dans chaque secteurs.

3 Analyse de la Problématique.

Avant de répondre à la problématique, nous avons représenté la répartition des bénéficiaires d'**ideX** selon le sexe des chercheurs dans chaque domaine de recherche.

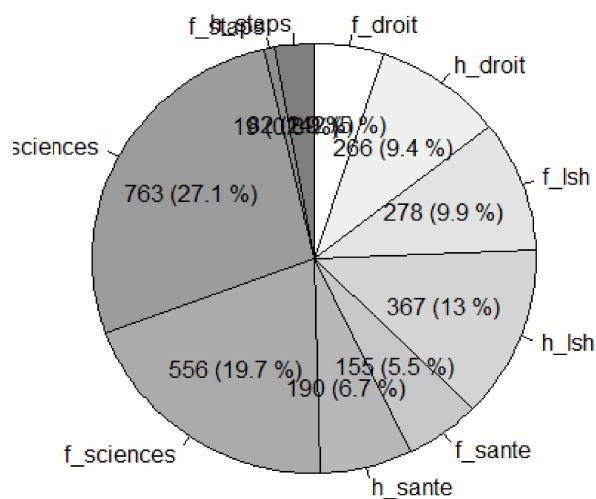


FIGURE 12 – La répartition des bénéficiaires d'IDEX selon le sexe dans chaque domaine de recherche.

On observe qu'il y'a effectivement plus de bénéficiaires qui sont dans le domaine scientifique, soit environ 27% sont des **hommes** et 27% des **femmes**. Maintenant, nous allons réorienté notre analyse en effectuant des tests sur les moyennes des bénéficiaires selon le sexe dans chaque domaine de recherches.

Pour cela, on a utilisé le test ANOVA factorielle afin de tester l'effet des variables indépendantes(**secteur_disciplinaire** et **sex**) sur une variable dépendante (**beneficiaires**) à l'intérieur d'une seule analyse (d'un seul modèle).

Anova Table (Type 3 tests)						
Response: beneficiaires						
	num	DF	den DF	MSE	F	Pr(>F)
secteur_disciplinaire		4	11464	48.765	123.150	
sex		1	11464	48.765	85.131	
secteur_disciplinaire:sex		4	11464	48.765	48.702	
ges						
secteur_disciplinaire		0.041199	< 2.2e-16	***		
sex		0.007371	< 2.2e-16	***		
secteur_disciplinaire:sex		0.016709	< 2.2e-16	***		

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1						

FIGURE 13 – le test ANOVA factorielle.

Les effets et pvalue trouvés pour **secteur_disciplinaire** et **sex** indiquent que le domaine de recherche et le sexe des chercheurs sont très significatif sur le programme d'investissement **ideX**, cela signifie qu'elles influencent réellement l'investissement **ideX** donné par l'état Français aux chercheurs. En effet, les moyennes des variables d'interactions ne sont pas équivalentes comme on peut le constater sur la **figure 12**. On pourrait voir qu'il y'a en moyenne plus de bénéficiaires du programme **ideX** chez les **Hommes**. Aussi,

les domaines les plus suivi par le programme **ideX** sont : **sciences** et **Lettres et sciences humaines**.

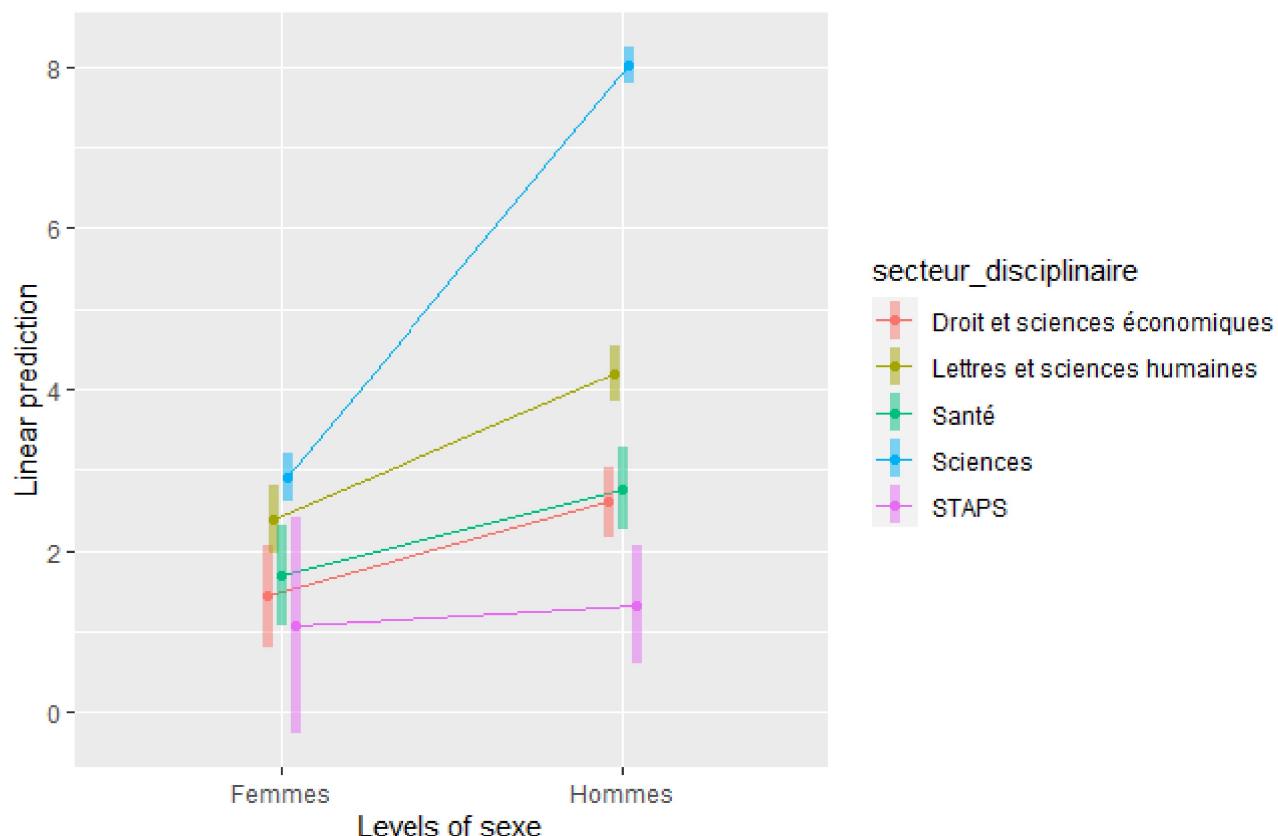


FIGURE 14 – Graphiques d'interaction pour les moyennes marginales estimées

Ainsi, sur la **figure 14**, on peut apercevoir que le domaine et le sexe de chercheurs ont une forte influence sur l'accompagnement financier de l'État (IDEX). En controverse, on dira que nous ne connaissons pas la véritable raison de la sur-représentation des interactions (**ideX, sciences**) et (**ideX, Hommes**) par rapport au financement de l'état dans la recherche.

Enfin, nous avons pu mettre en évidence l'influence qu'a le domaine et le sexe des chercheurs sur le suivi du programme d'investissement grâce aux données des bénéficiaires de la prime d'excellence scientifique de 1993 à 2012, mais nous ne pouvons pas affirmer avec certitude cette influence sachant que les résultats trouvés ne reflètent pas l'évolution de nombres de bénéficiaires par rapport au programme IDEX au fil des années. Cependant, il serait intéressant de traiter ces données par année afin d'avoir une étude plus pertinente sur les facteurs influents du programme d'investissement de l'État français dans l'enseignement supérieur et de recherche.

4 Bibliographie

1. Hamon Agnès et Jégou Nicolas *Presses universitaires de Rennes, 2008.* "Statistique descriptive."
2. Husson François et Lê, Sébastien *Pratique de la statistique* "Analyse de données avec R."
3. Les étapes statutaires Site
4. Ministère de l'Enseignement supérieur, de la Recherche et de l'Innovation Site
5. Les bénéficiaires de la prime d'excellence scientifique. Site.