

Des étoiles plein les yeux. Présenter des résultats d'analyse de régression avec R

Les « Vendredis quantis »

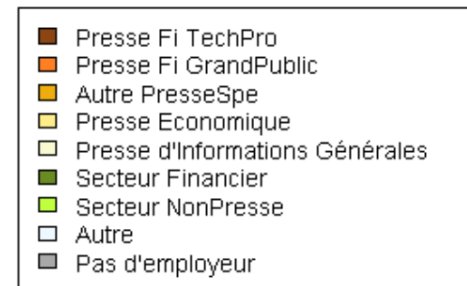
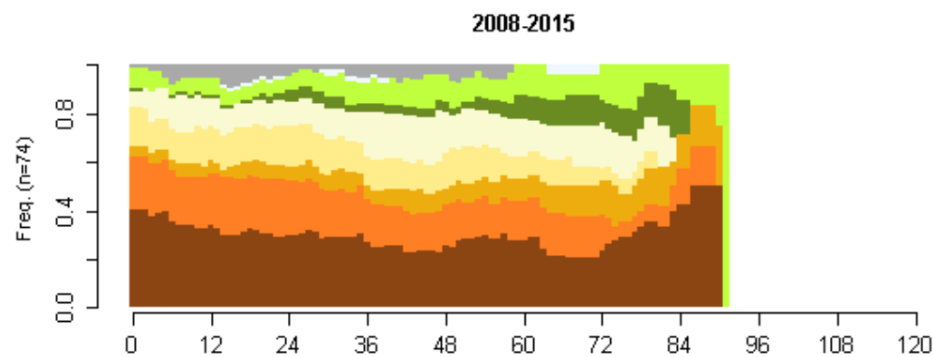
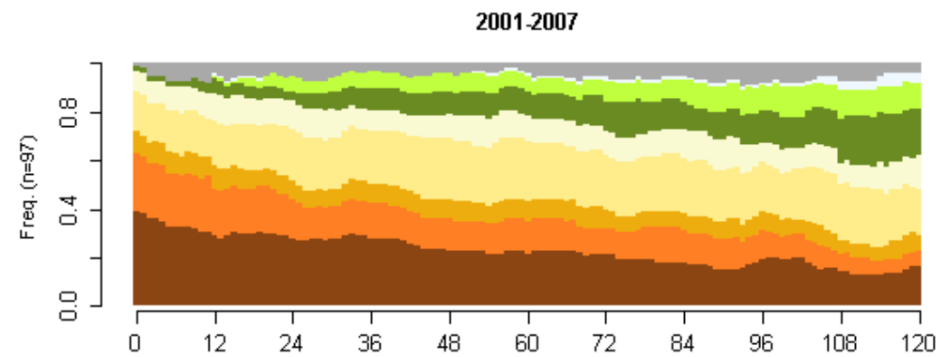
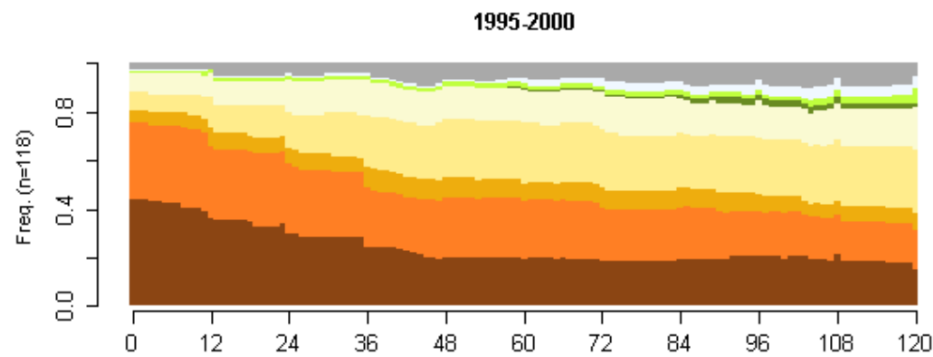
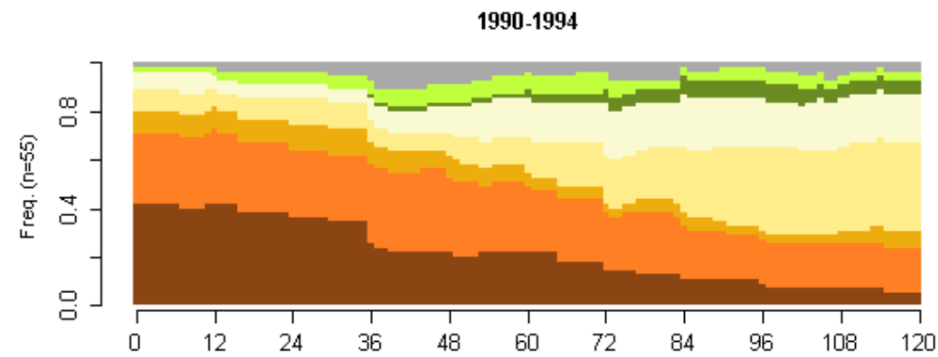
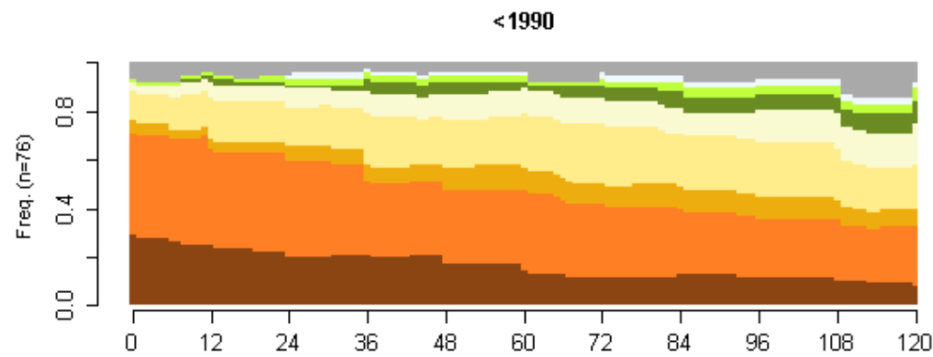
Antoine Machut

22 novembre 2019

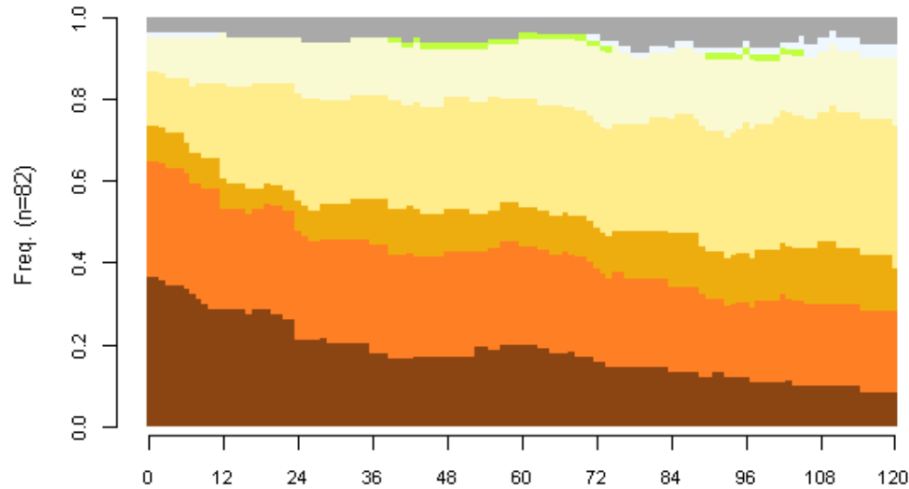
14h - 16h

Salle de réunion Pacte

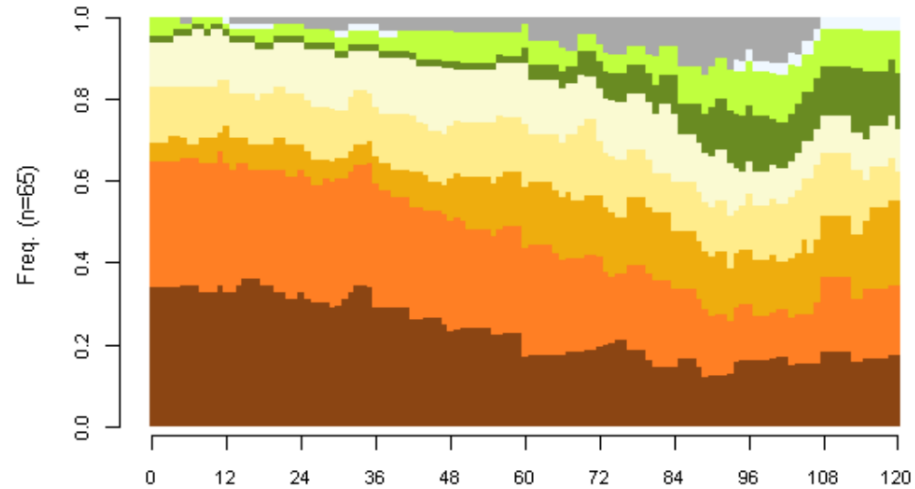




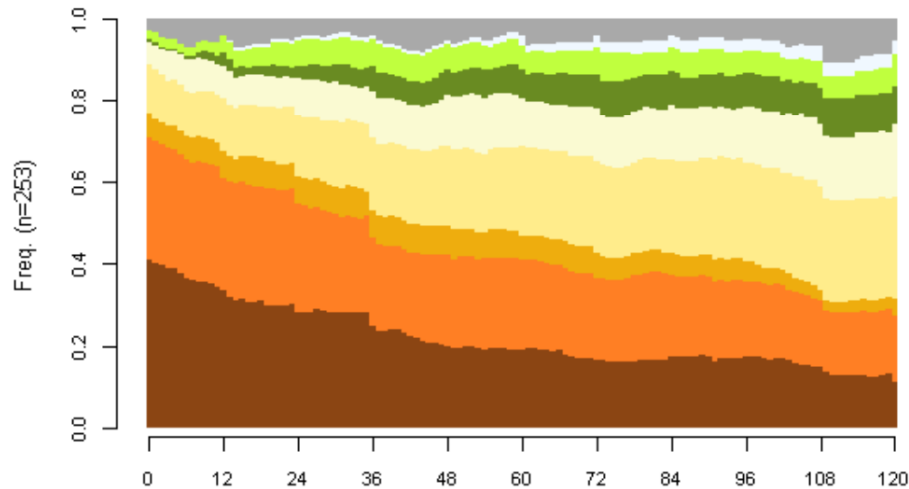
Ecoles de journalisme reconnues



Formations en journalisme non reconnues



Autres diplômes



- Presse Fi TechPro
- Presse Fi GrandPublic
- Autre PresseSpe
- Presse Economique
- Presse d'Informations Générales
- Secteur Financier
- Secteur NonPresse
- Autre
- Pas d'employeur

- Une fois prise en compte la « direction » prise par la trajectoire, plus ou moins spécialisée, subsiste-t-il un lien statistique entre le fait d'avoir suivi une formation en journalisme et la probabilité de quitter la presse ?
 - **Modèle n°1** : expliquer statistiquement tous les exits en fonction de la formation suivie et de la « direction » prise par la trajectoire.
 - **Modèle n°2** : expliquer les « exits » vers la finance.
 - **Modèle n°3** : expliquer les « exits » vers un autre secteur.

Nom de la variable	Type de variable et modalités	Signification
Ecole	Binaire, « oui » (N=82), « <u>non</u> » (N=318)	Mention ou non d'un diplôme en école de journalisme reconnue
Journalisme	Binaire, « oui » (N=71), « <u>non</u> » (N=329)	Mention ou non d'un cursus en journalisme, non reconnu par la profession.
Ecogestion	Binaire, « oui » (N=190), « <u>non</u> » (N=210)	Mention ou non d'un diplôme en économie et/ou gestion et/ou finance
Droit Sciences Po	Binaire, « oui » (N=139), « <u>non</u> » (N=261)	Mention ou non d'un diplôme en droit et/ou sciences politiques
Type de spécialisation	Nominale, « presse spe » (N=116), « presse ecoge » (N=69), « <u>autre</u> » (N=235)	« presse spe » : carrières journalistiques exclusivement dans la presse spécialisée « presse ecoge » : carrières journalistiques exclusivement en presse économique ou en presse généraliste « autre » : carrières journalistiques non-exclusives.
Sexe	Binaire, « <u>homme</u> » (N=235), « femme » (N=185)	
Cohortes_MIF	Nominale, « <1990 » (N=76), « 1990-1994 » (N=55), « 1995-2000 » (N=118), « 2001-2007 » (N=97), « 2008-2015 » (N=74)	Cohortes d'entrée dans le monde de l'information financière comme journaliste.
Origine.données	Binaire, « <u>LinkedIn</u> » (N=239), « Annuaire » (N=181)	Contrôle du support à partir duquel les informations de carrières ont été recodées.
Tps.journalisme	Continue, M=185,6, écart-type=106,48	Temps de carrière passé dans le journalisme, activités indépendantes et de rédaction en chef incluses (en mois).

Présenter les résultats des modèles avec R

- Le rapport $p/(1-p)$ est appelé *Odds* (« chance relative »).
 - Ex: parier sur un cheval A à 4 contre 1 signifie : « le cheval A a 4 fois plus de chances de perdre que de gagner » → $P_{\text{gagner}} = 20\%$ et $P_{\text{perdre}} = 80\%$.
- L'*Odds-ratio* est un rapport entre deux *Odds*, qui permet de comparer les « chances relatives » de connaître un événement pour deux catégories (a et b ci-dessous) de population.
 - $[P_a/(1-P_a)]/[P_b/(1-P_b)]$.
 - Régression logistique = estimer la probabilité de quitter la presse *plutôt que de ne pas la quitter* pour les personnes qui ont suivi des études de journalisme *par rapport* à ceux qui ne l'ont pas suivi.
- Pour aller plus loin...
 - Le billet de Pierre Mercklé sur la mesure des inégalités: <https://quanti.hypotheses.org/603>
 - Une défense de l'usage de l'odds-ratio par Louis-André Vallet dans courrier des statistiques (+ technique): <http://www.crest.fr/ckfinder/userfiles/files/pageperso/vallet/Courrier%20des%20Statistiques%20121-122.pdf>

- Représentations graphiques et tableaux
 - ggplot2 et l'extension Ggally
 - Présentation de ggplot2, cours de François Briatte
<http://larmarange.github.io/analyse-R/intro-ggplot2.html>
 - GGally: https://ggobi.github.io/ggally/#customizing_the_plot
 - Finalfit
 - Présentation: <https://www.datasurg.net/2018/05/16/elegant-regression-results-tables-and-plots-the-finalfit-package/>
 - En Français, dans un chapitre consacré à la régression logistique sur R
<http://larmarange.github.io/analyse-R/regression-logistique.html>
 - Texreg
 - Un ex. d'utilisation sur le lien entre prestige, formation et revenu
http://www.jkarreth.net/files/RPOS517_Day8_MultipleReg.html
 - Une revue des packages pour formater les tableaux de régression ici
<https://jrnold.github.io/intro-methods-notes/formatting-tables.html#overview-of-packages>