**Insertion professionnelle des licenciés (bac+3) en STAPS**

**Note méthodologique**

**Cette note vise à apporter des éléments complémentaires sur la méthode d’analyse mise en œuvre.**

Afin d’améliorer la représentativité des données, une **pondération** a d’abord été calculée, selon les étapes suivantes :

1. modélisation de la non-réponse totale (avec variables sexe, âge, série du bac, université) ;
2. repondération des répondants de l’inverse du taux de réponse du groupe homogène de réponse (GHR) auxquels ils appartiennent (définis à la suite du modèle de non réponse) ;
3. calage des poids finaux sur les marges par université et par genre des diplômés de STAPS. Ce calage permet d’avoir les mêmes répartitions et effectifs sur ces variables dans l’échantillon et dans la population mère.

Cette variable de pondération a été utilisée pour l’ensemble des analyses.

**La méthode d’analyse de données textuelles** exposée ci-dessous a ensuite été mise en œuvre pour analyser la satisfaction vis-à-vis de l’emploi occupé, en termes de compétences mises en œuvre et de missions, en rapport avec la nature de l’emploi appréhendée à travers l’intitulé de l’emploi occupé.

L’analyse des données textuelles est la seule méthode permettant d’exploiter systématiquement des « verbatim », comme l’intitulé de l’emploi occupé.

Le principe de la méthode est de recenser les formes lexicales « pleines » (i. e. significatives au sens ordinaire de « porteuses de sens ») associées aux modalités individuelles de la variable « intitulé de l’emploi occupé », les occurrences de ces formes lexicales sont recensées dans le tableau lexical entier (TLE, en l’occurrence il sera même binaire, ne comportant que zéro ou un, correspondant respectivement à la présence ou l’absence de la forme lexicale pleine dans la modalité), ensuite ce TLE est croisé avec les indicatrices des modalités des variables indiquant la satisfaction par rapport aux compétences ou aux missions.

Comme ce tableau est binaire (ne comprenant que zéro ou un) on peut lui appliquer la technique d’analyse des données la plus indiquée en pareil cas, l’analyse factorielle des correspondances (AFC). On obtient, pour chacune des deux variables, satisfaction par rapport aux compétences mises en œuvre et aux missions, une AFC qui se résume principalement à un tableau donnant le score ou indice signé de chaque modalité de la variable de satisfaction associée à chaque forme lexicale pleine, indiquant la plus ou moins grande[[1]](#footnote-1) proximité entre les deux, et un graphique représentant les formes lexicales pleines en projection sur le premier plan factoriel.

Si l’on considère par exemple la première AFC utilisant les modalités de la variable de « satisfaction en termes de compétences » (premier graphique), on constate sur le premier axe une opposition entre des termes tels que :

« assistant »,  « éducation », « employé »,  « vendeur », « polyvalent », « surveillant » correspondant plus ou moins au minimum de satisfaction d’une part (les termes regroupés à droite de l’axe) ;

« professeur », « EPS », « gendarme, « doctorant », « enseignant », « contractuel », « stagiaire », « école » se rapprochant du maximum de satisfaction (les termes regroupés à gauche de l’axe) d’autre part.

Cette opposition correspond au premier axe factoriel, celui qui correspond à l’opposition la plus forte au sein des données. Le deuxième axe en revanche ne fait pas apparaître de signification claire.

La seconde AFC, portant sur la variable de « satisfaction en termes de missions » (deuxième graphique), oppose les termes :

« assistant »,  « éducation », « employé »,  « vendeur », « polyvalent », dans le sens du minimum de satisfaction, à droite de l’axe, d’une part ;

« professeur », « EPS », « gendarme, « enseignant », « contractuel », « stagiaire », « doctorant » - donc à gauche de l’axe, d’autre part.

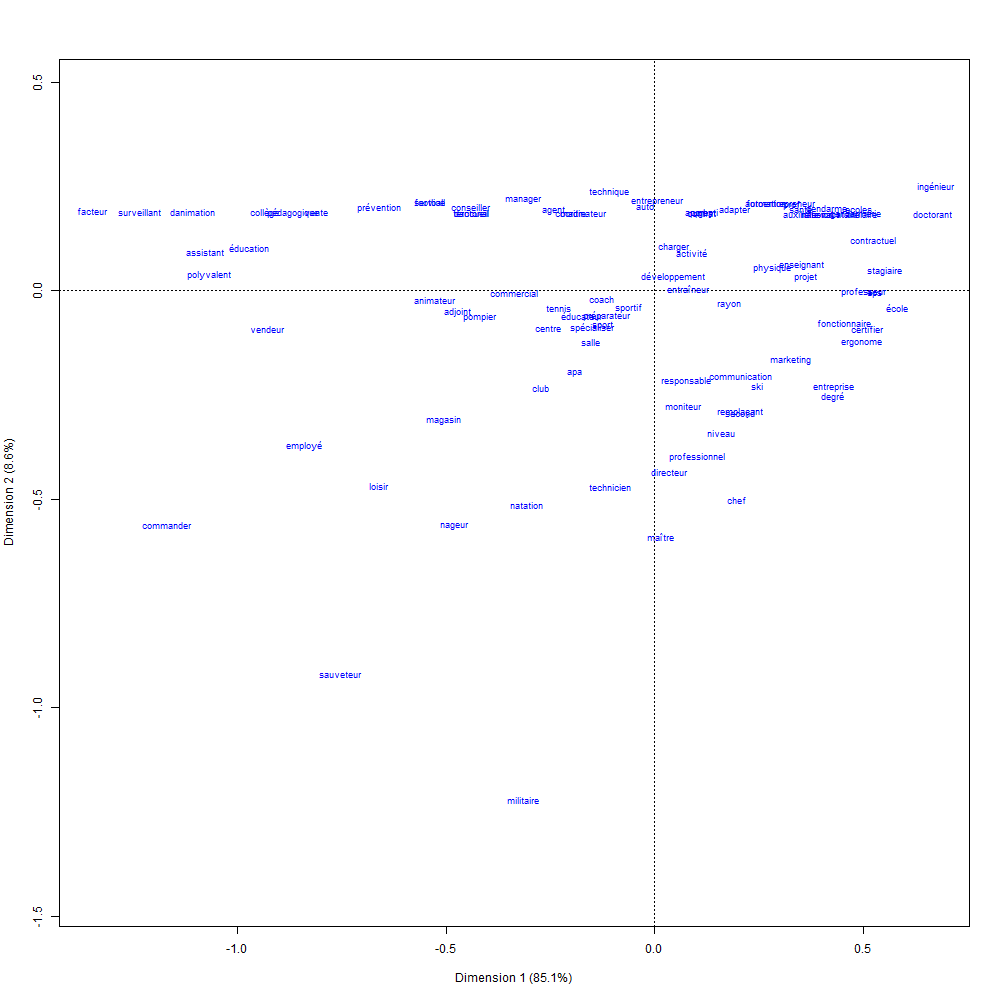
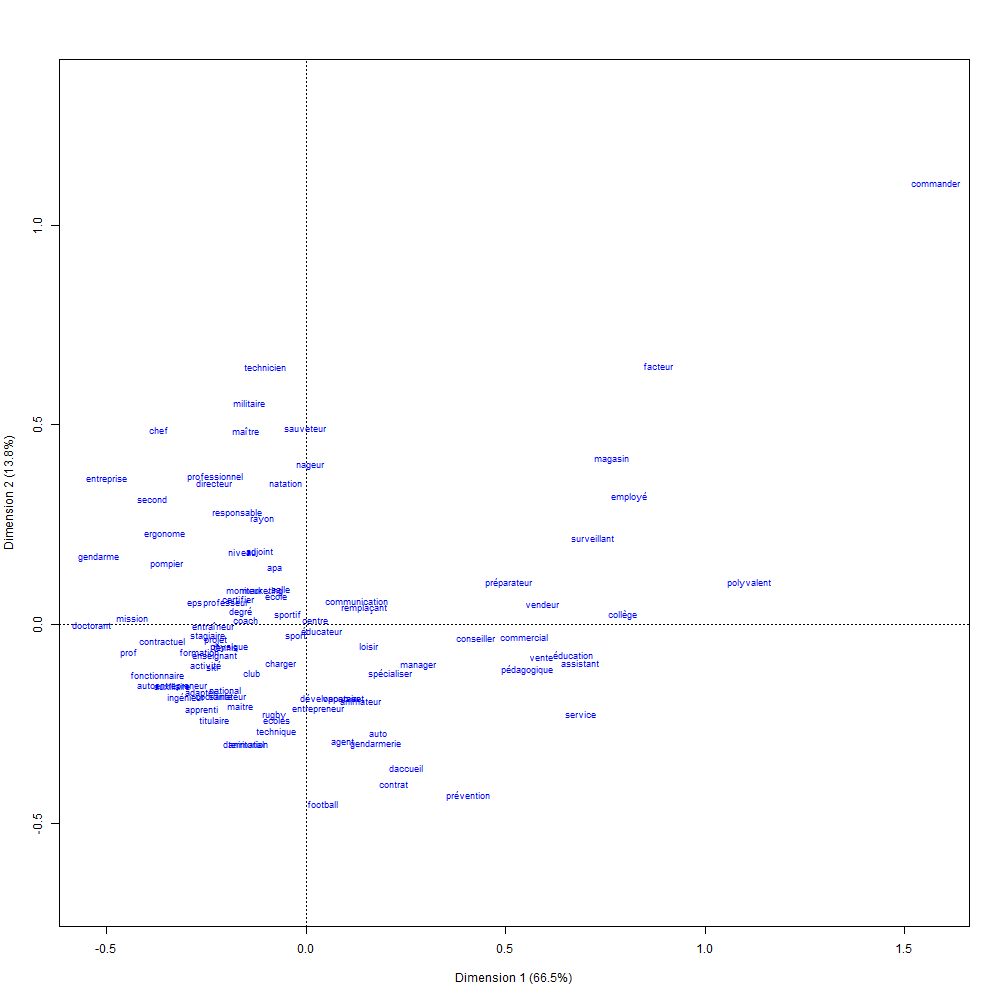
Cette opposition correspond au premier axe factoriel également, le deuxième axe ne faisant pas apparaître de signification claire.

Les deux notions de satisfaction se recouvrent donc en grande partie, et la leçon la plus claire de cette analyse est l’opposition entre les assistants d’éducation d’une part et les professeurs d’EPS d’autre part, correspondant presque exactement respectivement au minimum et au maximum de la fonction de satisfaction… résultat qui tend de plus à être corroboré si l’on tente des AFC équivalentes sur le corpus textuel constitué à partir des noms d’employeurs.

**Pour aller plus loin**

Sur l’analyse de données textuelles :

<http://www.jms-insee.fr/2015/S03_5_ACTE_CAVALLO_JMS2015.PDF>



1. Au sens de la distance associée à la loi du chi-2, ou à la loi hypergéométrique, qui a été retenue ici. [↑](#footnote-ref-1)