

Analyse Factorielle des Correspondances.

Travaux pratiques (statistiques exploratoires).

Les jeux de données étudiés sont disponibles sur <http://math.univ-lille1.fr/~marbaclo/>

1 Marketing sous R

La régie des tabacs voulait choisir un nom de marque pour une nouvelle cigarette brune qui devait paraître de qualité supérieure, prestigieuse, luxueuse, convenant à un fumeur viril, connaisseur, raffiné, distingué et niveau socio-économique élevé. Pour cela, elle avait sélectionné douze noms de cigarettes parmi lesquels il fallait sélectionner celui qui correspondait le mieux à la description précédente. Onze attributs pouvaient être associés à ces noms. On demanda alors à 100 fumeurs de désigner pour chacun des attributs la ou les cigarettes qui lui correspondaient le mieux. Les résultats se trouvent dans le fichier `tabac.txt`. Selon le cahier des charges décrit ci-dessus, quelle marque devrait être retenue? L'AFC se fait sous R en utilisant la fonction `CA()` du package `FactoMineR`.

2 Valeurs manquantes sous R

Charger le jeu de données `children` du package `FactoMineR` qui résume les réponses de différentes catégories de personnes (en colonne) à la question suivante : d'après vous, quelles sont les raisons qui peuvent faire hésiter un couple à avoir des enfants? Décrire la liaison entre les catégories de personnes et les réponses apportées.

3 Logements de vacances et activité professionnelle sous SAS

On a interrogé 31079 familles sur leur logement de vacances. Les données sont présentées dans le tableau de contingence `activite_logement.txt` croisant 10 types d'activités et 8 types de logements. La signification des abréviations est la suivante :

- hote-hôtel, habi-habitant, prop-propriété, pare-parents, amis-amis, teca-tente et caravane, viva-village de vacances et dive-divers.
- agri-agriculteurs, sala-salariés agricoles, part-partons ind. et com., cadS-cadres supérieurs et professions libérales, cadM-cadres moyens, empl-employés, ouvrv-ouvriers, pers-personnels de service, autr-autres actifs et nact-non actifs.

Expliquer le lien entre le type de logement de vacances et l'activité professionnelle s'il existe. L'AFC se fait sous SAS en utilisant la `proc corresp`.

4 Application en biologie sous SAS

Le micronecton est un ensemble d'organismes marins dont la capacité de nage est telle qu'il peut se déplacer contre les courants. On effectue un prélèvement¹ de micronecton dans 11 stations réparties le long d'une radiale Nice - Calvi. Des mesures hydrologiques associées ont montré que les stations 1 à 7 étaient situées dans la zone périphérique, les stations 8 à 10 dans la zone frontale et la station 12 dans la zone côtière. La station 11 n'a pas été échantillonnée. Les résultats sont présentés dans le tableau de comptage suivant :

Comment peut-on décrire la relation entre les associations entre les 12 espèces déterminées et la localisation des stations?

1. Distribution of crustacean micronekton across a Mediterranean front. JC Miquel & JPh Labat. 1992. ICSEM, Rapp. Com. Int. Mer Medit., 33, 261

Station	Meno	Neme	Stlo	Eukr	Thae	Euhe	Sear	Sero	Sesa	Pasi	Pamu	Geel	Total
S1	204	80	4	0	0	0	0	0	0	1	1	0	290
S2	272	88	12	0	0	0	0	0	0	3	0	1	376
S3	528	18	7	0	0	0	0	0	0	4	0	0	557
S4	208	186	12	2	0	0	1	0	0	5	0	0	414
S5	74	30	3	3	0	0	0	0	0	2	2	0	114
S6	49	77	0	5	0	0	0	1	1	1	1	0	135
S7	39	35	4	2	2	0	0	0	0	4	0	2	88
S8	25	155	33	17	2	0	0	0	0	0	0	2	234
S9	5	127	30	21	0	0	0	0	0	0	1	0	184
S10	3	107	31	38	0	2	0	0	0	1	1	0	183
S12	2	10	1	118	0	0	1	0	0	0	0	0	132
Total	1409	913	137	206	4	2	2	1	1	21	6	5	2707

TABLE 1 – Tableau de comptage de 12 espèces de micronecton pour les 11 stations

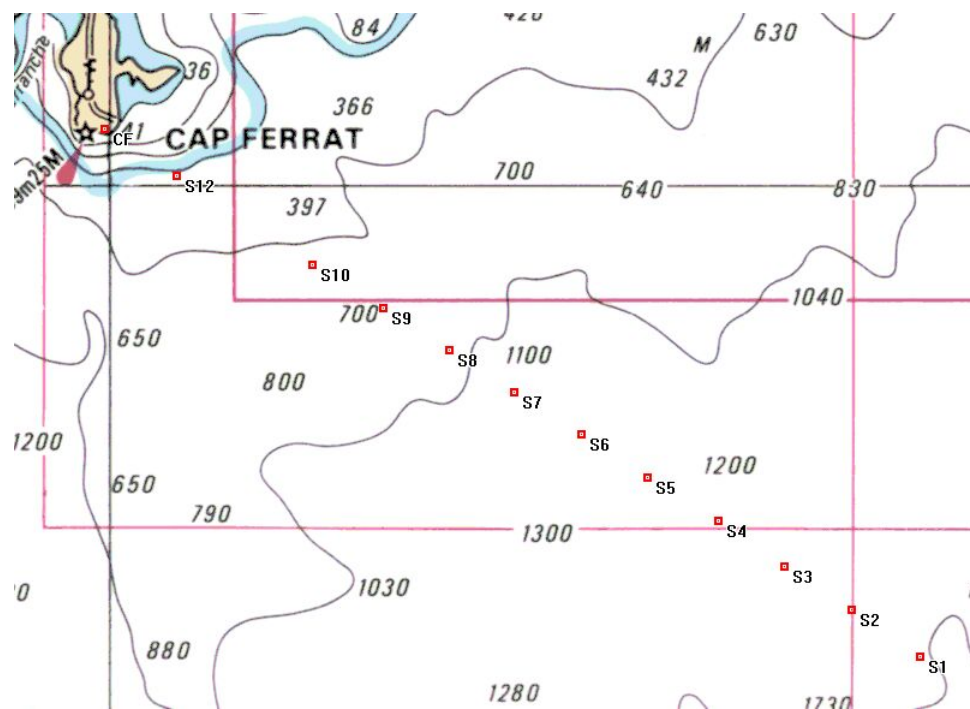


FIGURE 1 – Localisation des 11 stations