**Master 1 « Économie du Développement » et « Intelligence Économique »**

**UNIVERSITÉ DE BORDEAUX**

**Année 2017/2018**

**Gestion et analyse de données : Séance 7**

***Statistique exploratoire multidimensionnelle – Analyse factorielle des correspondances***

***Données initiales :***

Le tableau de la dernière page, tiré des données du World Values Survey 2015, ventile la population représentative de six pays d’Asie selon le pays d’origine et les quintiles de revenu des enquêtés d’une part et leur avis sur le principal problème économique et/ou social rencontré par leur pays d’autre part. On lit donc par exemple à la première case du tableau que 78 répondants de l’enquête sont des japonais du plus bas quintile de revenu considérant que le principal problème de leur pays est « l’économie ».

***Questions :***

1. A quel type de tableau statistique a-t-on à faire dans cet exercice ? Quelle est la population observée, quel est l’individu statistique et quels sont les caractères observés ?
2. Ouvrir le logiciel STATA®, puis ouvrir le do-file TD7\_tab.do.
3. Lancer la première ligne du do-file puis ouvrez le « data browser », que remarque-t-on ?
4. Lancer les lignes 2 à 5 du do\_file, qui permettent de nommer les modalités des deux variables, puis lancer la commande tabulate row col. Qu’obtient-on ? Quel est le problème ?
5. Lancer de nouveau la commande tabulate row col. Quel est le tableau obtenu ? Quels types de commentaires statistiques peut-on produire à partir d’un tel tableau ?
6. Etablissez maintenant le tableau des fréquences absolues. Quels types de commentaires statistiques peut-on dorénavant produire à partir de ce tableau ?
7. Etablissez ensuite séparément le tableau des profils-lignes (distributions conditionnelles en lignes), puis celui des profils-colonnes (distributions conditionnelles en colonnes). Quels types de commentaires statistiques peut-on produire à partir de ces deux tableaux ?
8. Quelle limite fondamentale présentent les commentaires issus de la question précédente ? Par extension, quelle(s) problématique(s) multidimensionnelle(s) seraient susceptibles d’être traitées, grâce aux outils appropriés, à partir des données fournies ? Quelle est alors la technique statistique devant être ici mise en œuvre ?
9. Pour commencer, grâce au test approprié, établissez statistiquement l’existence d’une dépendance significative entre les lignes et les colonnes du tableau.
10. Une fois son utilisation justifiée, lancez une analyse factorielle des correspondances sur les données considérées, puis utilisez les outils appropriés pour décider du nombre d’axes factoriels à examiner.
11. Utilisez les indicateurs nécessaires pour procéder au dépouillement puis à la description précise des axes retenus à la question précédente.
12. Utilisez ensuite l’outil approprié pour juger de la qualité de la représentation graphique des différents éléments.
13. Etablissez la représentation graphique de la projection des modalités sur les deux premiers axes factoriels (le premier plan factoriel de l’AFC) et proposez enfin sur la base de ce dernier une interprétation globale des résultats obtenus, visant à répondre à la problématique identifiée à la question 8

**Principale préoccupation des individus pour leur pays, selon la nationalité et le quintile de revenus en 2015 (Asie, six pays, source : WVS)**

