

---

## Méthodologie d'échantillonnage de la future Enquête Emploi en Continu à Mayotte

*Julien Jamme (\*), Cyril Favre-Martinoz (\*\*), Edouard Fabre (\*\*\*), Joachim Timotéo (\*\*\*)*

*(\*) Insee, Direction de la méthodologie et de la coordination statistique et internationale*

*(\*\*) Insee, Département de la Démographie*

*(\*\*\*) Insee, Direction Interrégionale La Réunion-Mayotte*

`julien.jamme@insee.fr`

**Mots-clés.** (6 maximum) : Échantillonnage, équilibrage, collecte, variance

**Domaines.** Théorie des sondages amont.

---

## Résumé

Après une première expérience menée en 2009, l'enquête Emploi est réalisée tous les ans à Mayotte depuis 2013. Il s'agit d'une enquête annuelle réalisée de début mars à début juillet de chaque année. Aujourd'hui, le nouveau règlement européen (IESS), en vigueur depuis le 1er janvier 2021, impose que tout le territoire d'un pays soit couvert par une enquête Emploi en continu. La France a obtenu pour Mayotte une nouvelle dérogation pour la période 2021-2023. Elle est donc dans l'obligation de mettre en place une EEC opérationnelle au premier trimestre 2024.

Des premiers travaux relatifs à l'échantillonnage ont été menés afin de respecter certaines contraintes méthodologiques, réglementaires et de les concilier avec les questions d'organisation de la collecte sur le terrain (nombre d'enquêteurs, charge de travail par enquêteur, secteurs d'enquête, spécificité de la collecte à Mayotte, etc.). Cet article a pour objectif de décrire en détails ces travaux.

A Mayotte, la base de sondage retenue est la base cartographique. Des alternatives comme les sources administratives (Fidéli) et les enquêtes annuelles de recensement ont été envisagées mais écartées pour deux raisons : la première EAR date de 2021 et les sources administratives sur ce territoire ne sont pas exhaustives. La base cartographique de Mayotte est une base recensant l'ensemble des adresses présentes sur le territoire de Mayotte. Elle est mise à jour chaque année par l'intermédiaire d'une enquête cartographique qui a lieu sur un cinquième des îlots de Mayotte

(correspondant à un groupe de rotation<sup>1</sup> donné et attribué à chaque îlot<sup>2</sup>). Face à l'évolution très rapide du parc des logements, l'échantillonnage de l'enquête Emploi à l'année N se limite aux adresses qui ont été enquêtées par l'enquête cartographique à l'année N-1, donc sur un cinquième du territoire. La mise à jour de celle-ci ayant lieu fin octobre de l'année N-1, on procédera, en fin d'année N-1, au tirage de l'échantillon des trimestres T2, T3 et T4 de l'année N et du trimestre T1 de l'année N+1 dans la partie de la base cartographique qui aura été enquêtée au cours de l'année N-1.

Afin de répartir l'échantillon de logements à collecter dans l'espace et dans le temps, des Secteurs d'Activité Enquêteurs (SAE) ont été constitués par le Criem, Chaque secteur doit recevoir une partie égale de l'échantillon. La méthode consiste à regrouper les îlots en suivant un chemin issu de l'algorithme du voyageur du commerce. Cet algorithme permet de passer d'un îlot à l'un de ses voisins, chaque îlot étant visité une et une seule fois. Le regroupement des îlots est arrêté dès qu'un nombre de logements, fixé à l'avance, est atteint. L'objectif est de construire les SAE les plus homogènes possible en termes de nombre de logements sur chacun des groupes de rotation des îlots.

Ensuite, le plan de sondage proposé devra respecter un certain nombre de contraintes :

- 1. Tirer un échantillon dont la taille s'approche le plus possible des 936 logements entrants par an ;
- 2. Obtenir une égale répartition de l'échantillon entre les SAE, soit 36 +/- 4 logements entrants chaque année ;
- 3. Respecter la contrainte européenne de précision ;
- 4. Faire en sorte que le plan de sondage soit autopondéré ;
- 5. Obtenir la meilleure précision possible pour les estimateurs étudiés.

Ne disposant que d'une base de sondage d'adresses issue d'une enquête cartographique annuelle et de peu d'informations dans cette base, si ce n'est le nombre de logements à l'adresse, nous proposons un tirage stratifié des logements en deux étapes.

La variable de stratification retenue est la variable de type d'adresse définie comme suit :

- Les Monologements (Mono) : une adresse = un logement ;
- Les Petites Adresses (PA) : une adresse = 2 à X logements ;
- Les Grandes Adresses (GA) : une adresse = plus de X logements.

Le choix du seuil X fera l'objet d'une discussion spécifique.

Le tirage se décline en deux étapes :

- un tirage des logements en grandes adresses rassemblées en une seule strate sans distinction du SAE
- des tirages indépendants dans chaque croisement SAE\*type d'adresse avec des allocations égales par secteur et fixées à l'avance.

Les simulations effectuées montrent que le scénario retenu permet de contrôler la taille de l'échantillon et la charge par enquêteur. La construction des SAE assure empiriquement l'autopondération du plan de sondage. La taille de l'échantillon fixée en amont est suffisante pour respecter la

---

1. Les groupes de rotation des îlots ont été constitués séparément pour les petites communes par le Criem et les grandes communes par la Division des Méthodes et Traitements du Recensement (DMTR) de l'Insee, tout en suivant une démarche et des méthodes similaires. Ils sont constitués successivement par un tirage équilibré sur des variables auxiliaires telles que le nombre de logements, le nombre de logements en dur, le nombre de logements collectifs, la population par âge, la population selon le lieu de naissance et une typologie des villages.

2. On appelle « îlot » une unité géographique regroupant un ou plusieurs pâtés de maisons et délimité par des voies (rues, routes, chemins) ou des barrières naturelles. Les 811 îlots actuels ont été constitués à partir des données du millésime 2017 du recensement de la population. Ce découpage a été réalisé en 2020.

contrainte européenne de précision. Le taux de chômage annuel serait estimé avec une précision de +/-3.4 points de pourcentage.

En parallèle, d'autres méthodes d'échantillonnage (tirage équilibré, spatialement équilibré et équilibré) ont été testées et les résultats sur les différents indicateurs d'intérêt (taille fixe, autopondération, précision) seront présentés.

## Références

- ARDILLY, Pascal (2006). *Les techniques de sondage*. 2e édition. Editions Technip. ISBN : 9782710808473.
- BENEDETTI, Roberto, Federica PIERSIMONI et Paolo POSTIGLIONE (2017). "Spatially Balanced Sampling : A Review and A Reappraisal". In : *International Statistical Review* 85.3, p. 439-454. DOI : <https://doi.org/10.1111/insr.12216>. URL : <https://onlinelibrary.wiley.com/doi/abs/10.1111/insr.12216>.
- CHAUVET, Guillaume (2009). "Stratified balanced sampling". In : *Survey Methodology*. Statistics Canada 35.1, p. 115-119. URL : [https://www.researchgate.net/profile/Guillaume\\_Chaudet/publication/288704033\\_Stratified\\_balanced\\_sampling/links/58d8e3d592851c44d4ada090/Stratified-balanced-sampling.pdf](https://www.researchgate.net/profile/Guillaume_Chaudet/publication/288704033_Stratified_balanced_sampling/links/58d8e3d592851c44d4ada090/Stratified-balanced-sampling.pdf).
- CHAUVET, Guillaume et Yves TILLÉ (2006). "A fast algorithm for balanced sampling". In : *Computational Statistics* 21.1, p. 53-62. DOI : 10.1007/s00180-006-0250-2. URL : <https://doi.org/10.1007/s00180-006-0250-2>.
- DEVILLE, Jean-Claude et Yves TILLÉ (2004). "Efficient balanced sampling : The cube method". In : *Biometrika* 91.4, p. 893-912. DOI : 10.1093/biomet/91.4.893. URL : <https://doi.org/10.1093/biomet/91.4.893>.
- (2005). "Variance approximation under balanced sampling". In : *Journal of Statistical Planning and Inference* 128.2, p. 569-591. DOI : 10.1016/j.jspi.2003.11.011. URL : <https://www.sciencedirect.com/science/article/abs/pii/S0378375804000035>.
- GRAFSTRÖM, Anton, Niklas L. P. LUNDSTRÖM et Lina SCHELIN (2012). "Spatially Balanced Sampling through the Pivotal Method". In : *Biometrics* 68.2, p. 514-520. DOI : <https://doi.org/10.1111/j.1541-0420.2011.01699.x>. URL : <https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1541-0420.2011.01699.x>.
- GRAFSTRÖM, Anton et Yves TILLÉ (2013). "Doubly balanced spatial sampling with spreading and restitution of auxiliary totals". In : *Environmetrics* 24.2, p. 120-131. DOI : <https://doi.org/10.1002/env.2194>. URL : <https://onlinelibrary.wiley.com/doi/abs/10.1002/env.2194>.
- INSEE-CRIEM (2013). *Mise en place d'un dispositif pérenne de sondage pour les enquêtes auprès des ménages à Mayotte*. N°797/DR974-DR974-SES.
- INSEE-DIVISION EMPLOI (2021). *Enquête Emploi, enquête sur l'emploi, le chômage et l'inactivité. Méthodologie*. <https://www.insee.fr/fr/metadonnees/source/operation/s2022/documentation-methodologique>.
- LOONIS, Vincent (dir.) et Marie-Pierre (coord.) BELLEFON (DE) (2018). *Manuel d'Analyse Spatiale*. Insee Méthodes. Insee. ISBN : 9782111396777.
- RIVEST, Louis-Paul et Sophie BAILLARGEON (2017). *stratification, package R*. URL : <https://cran.r-project.org/web/packages/stratification/>.
- THIBAULT, Pierre (août 2019). "Les villages de Mayotte en 2017". In : *Insee Analyses Mayotte* 22. URL : <https://www.insee.fr/fr/statistiques/4223807>.

- TILLÉ, Yves (2001). *Théorie des sondages : Échantillonnage et estimation en population finie*. Dunod. ISBN : 9782100054848.
- YATES, F. et P. M. GRUNDY (1953). "Selection Without Replacement from Within Strata with Probability Proportional to Size". In : *Journal of the Royal Statistical Society : Series B (Methodological)* 15.2, p. 253-261. DOI : <https://doi.org/10.1111/j.2517-6161.1953.tb00140.x>. URL : <https://rss.onlinelibrary.wiley.com/doi/abs/10.1111/j.2517-6161.1953.tb00140.x>.
- YATES, Frank et Henry Gerard THORNTON (1948). "Systematic sampling". In : *Philosophical Transactions of the Royal Society of London. Series A, Mathematical and Physical Sciences* 241.834, p. 345-377. DOI : 10.1098/rsta.1948.0023. URL : <https://royalsocietypublishing.org/doi/10.1098/rsta.1948.0023>.