



TAFRA

Proposition de projet : Mapping des populations vulnérables – Covid-19

Equipe du projet

Younes Benmoumen (coordination), Romain Ferrali (datascientist), Youssef Oulhote (épidémiologie et biostatistique), David Goeury (géographie et cartographie), Ahmed El Oufir (programmation et SIG), Laila Slassi (conseil juridique) et Othmane Bentaouzer (assistant de recherche).

Contact : yb@tafra.ma

Objectifs généraux du projet

1. Faciliter l'opération de confinement de la population, notamment des franges les plus vulnérables en termes socioéconomiques et épidémiologiques,
2. Orienter la mobilisation des ressources publiques (aides financières et en nature) vers les zones où les besoins sont les plus pressants,
3. Identifier les clusters potentiels de propagation de l'épidémie et orienter une éventuelle politique de testing ciblé.

Produits attendus

1. Une application mobile à l'usage des pouvoirs publics est disponible, permettant d'identifier à l'échelle des quartiers les zones de vulnérabilité socioéconomiques et épidémiologiques.
2. Des estimations chiffrées seront disponibles sur le nombre, la distribution statistique et géographique des ménages les plus vulnérables, en termes socio-économiques et épidémiologiques, à la propagation du virus.

Contexte et justification

Le gouvernement a ordonné la mise en place d'un confinement général de la population à partir du 20 mars 2020, pour une durée allant au moins jusqu'au 20 avril 2020. La réussite de cette opération de confinement se heurte à plusieurs contraintes, parmi lesquelles les deux principales sont :

- Une part significative de la population active marocaine est constituée de travailleurs informels dont les revenus sont générés sur une base quotidienne. Opérant dans l'économie informelle et n'étant pas bancarisés, ils ne disposent pas d'épargne/thésaurisation et sont difficilement atteignables pour une aide directe de l'Etat. D'autre part, une frange de la population marocaine n'a pas accès facilement à une information de qualité¹. Il s'agit notamment des personnes analphabètes. Ces obstacles ne permettent matériellement pas à cette frange de la population de suivre les consignes de confinement.
- Une part significative de la population marocaine est comprise dans la frange à risque, la plus vulnérable à l'épidémie. Il s'agit des plus de 60 ans et des personnes présentant

¹ Il ne s'agit pas de la problématique des « fake news », mais uniquement de la capacité à accéder à l'information écrite, la consulter et la comprendre.

des maladies chroniques (diabète, maladies cardio-vasculaires et respiratoires, immunodéficience).

Grâce au recensement et aux enquêtes menées par le HCP, le gouvernement détient des données dont l'exploitation à l'heure actuelle permettrait d'éclairer la décision publique et de faciliter l'organisation, planifiée au niveau central, coordonnée et mise en œuvre au niveau local, de l'action des pouvoirs publics et de la solidarité nationale.

Stratégie d'intervention

L'analyse des données du recensement général de la population et de l'habitat (RGPH 2014) et de la cartographie du territoire (au niveau le plus fin) permet d'identifier, au niveau des quartiers urbains, les zones de concentration des populations les plus vulnérables à la propagation de l'épidémie, à travers deux catégories se chevauchant :

1. La difficulté socio-économique à respecter les consignes de confinement (travailleurs journaliers),
2. La vulnérabilité épidémiologique (ménages comptant au moins ou exclusivement un/des membre(s) âgé(s) de plus de 60 ans ou présentant une/des maladies chroniques).

L'exploitation des micro-données du HCP (ménages et individus) géolocalisées sur la cartographie administrative permet d'identifier les zones de concentration des populations vulnérables. Les parties suivantes présentent la méthodologie pour arriver à ce résultat.

Étape 1 : collecte des bases de données

La première étape consiste à collecter les jeux de donnée suivants :

- La base de micro-données du RGPH (2014), fichier « individu » incluant la localisation GPS de l'îlot² du répondant, et métadonnées correspondantes (données HCP),
- La base de micro-données du RGPH (2014), fichier « ménages et logements » incluant la localisation GPS de l'îlot du ménage, et métadonnées correspondantes, (données HCP),
- La cartographie digitalisée du découpage administratif du territoire à l'échelle la plus fine. S'agissant du milieu urbain, le plus propice à la propagation de l'épidémie, il s'agit d'obtenir une cartographie à l'échelle infra-arrondissement, permettant de lier un territoire à l'autorité locale correspondante (données Ministère de l'Intérieur).

Étape 2 : nettoyage et harmonisation des bases de données

Cette étape consiste à faire communiquer les différents jeux de données, de façon à obtenir une base de données unique, comprenant les éléments suivants :

- Chaque individu (répondant) est affilié à un ménage,
- Chaque ménage est géolocalisé,
- Chaque ménage relève d'une division administrative,
- Les divisions administratives recoupent la totalité des ménages,
- Chaque ménage possède un indicateur binaire, prenant la valeur « 1 » si un cas de contamination ou de contact significatif avec une personne contaminée a été confirmé au sein de la même division administrative. Si un dataset comprenant les localisations (au moins un niveau d'une division administrative) des personnes contaminées est

² L'îlot est la subdivision la plus petite du niveau d'identification géographique pratiquée par le RGPH. Cette information permet d'assigner le répondant ou le ménage à une division administrative (donc à l'autorité locale compétente) mais également d'identifier des possibles clusters au niveau local.

fournie au démarrage du projet, cette donnée peut être incluse. Sinon, une des fonctionnalités de l'application permettra aux pouvoirs publics de la renseigner, et de corriger les indicateurs en conséquence.

Etape 3 : construction des indicateurs de vulnérabilité socioéconomique et épidémiologique

Cette étape consiste à développer une méthode algorithmique qui mesure le degré de vulnérabilité socioéconomique (Vse) et épidémiologique (Vep) de chaque ménage marocain. La méthodologie de construction des scores est basée sur le questionnaire « ménages et logements » du HCP, implémenté lors du dernier RGPH de 2014. Cette méthodologie pourrait être précisée grâce à l'accès au questionnaire « individu ». Les scores sont construits de la façon suivante.

1. Le score de vulnérabilité socio-économique du ménage (dite « Vse ») est construit sur 11 points de contrôle. Le score sera construit séparément pour les zones urbaines et rurales et sera basé sur les questions incluses dans la Figure 1 (d'autres variables pourront être incluses au besoin). Nous utiliserons une analyse en composantes principales qui nous permettra d'estimer une combinaison linéaire des indicateurs observés, et ainsi réduire le nombre d'items. En deuxième étape, nous allons combiner les scores estimés par la première et deuxième composantes principales qui expliquent la plus grande part de variance pour construire le score de vulnérabilité socio-économique finale. Le score final de vulnérabilité sera standardisé avec une moyenne de 0 et écart-type de 1.

Figure 1 : composition de l'indicateur de vulnérabilité socioéconomique (Ves)

Id. variable	Variable
Variable XX	Variable à définir à partir du questionnaire "individu" non disponible en ligne. Cette variable doit déterminer si le ménage a au moins un membre opérant dans le secteur formel (notamment inscrit à la CNSS). Si aucun=1
Variable 29	Avez-vous un emploi saisonnier ? Oui=1
Variable 31	Est-ce que vous êtes à la recherche d'un emploi ? Oui=1
Variable L01	Si « Sans abri (6) » = 10
Variable L02	Si « Maison sommaire ou bidonville (6) » ou « Local non destiné à l'origine à l'habitation (7) » = 1
Variable L03	Toute réponse à l'exception de « Béton armée, briques et mortier (1) » = 1
Variable L04	Toute réponse à l'exception de « Dalle (1) » =1
Variable L09	Si l'un des équipements « cuisine » ou « latrine » ou « bain », ou « réseau de distribution d'électricité » ou « réseau de distribution d'eau » est commun=1
Variable L09	Si l'un des équipements « cuisine » ou « latrine » ou « bain », ou « réseau de distribution d'électricité » ou « réseau de distribution d'eau » est inexistant =1
Variable L14	Si « mode d'évacuation des eaux usées » est autre que « réseau public (1) » =1
Variable L17	Si le ménage ne dispose pas d'un réfrigérateur =1
Variable combinée (à créer)	Si P1=1 (personne vivant seule) et tout autre variable=1**
Variable XX*	Le ménage ne compte aucune personne analphabétisée

* Cette variable est à identifier dans le questionnaire "individus" (non disponible en ligne)

Le score Vse prend une distribution normale avec des valeurs allant éventuellement de -5 à 5, ou -5 représente, par exemple, un ménage au risque le plus minimal, et 5 représente un ménage à fort risque. Nous envisageons également de catégoriser ce score en quintiles ou déciles pour stabiliser la combinaison avec le score de vulnérabilité épidémiologique et éviter les scores extrêmes.

2. Le score de vulnérabilité épidémiologique (dite « Vep ») est calculé pour chaque ménage en prenant en compte les 4 points de contrôle ci-dessous. A noter que le RGPH n'inclut pas de questions sur l'état sanitaire et les pathologies des répondant(e)s, il est donc impossible de renseigner les vulnérabilités de ce type. Enfin, la vulnérabilité est mesurée au niveau des ménages car le mode de propagation du virus se fait principalement dans ce milieu comme rapporté dans les études chinoises, et comme prévu avec les mesures de confinement généralisées.

Figure 2 : composition de l'indicateur de vulnérabilité socioéconomique (Vep)

Id. variable	Variable
Variable combinée P1/L07 (à créer)	Cette variable indique le niveau de promiscuité au sein du ménage. Il est proposé de donner un score égal au ratio, puisque la vulnérabilité augmente avec la promiscuité. Par exemple : si 4 personnes habitent dans 2 pièces, le score est de $4/2=2$
AGE5*	Cette variable identifie les répondants de plus de 60 ans. Le "fichier technique" du dataset micro-données individuelles publiées en 2018 par le HCP indique qu'elle prend les valeurs suivantes, pour les plus de 55 ans : "55,60,65,70,75"
Var XX (à créer)	Indique si un cas de contamination au Covid-19 a été identifié dans la même unité administrative
Var XX (à créer)	Indique si un cas de contamination au Covid-19 s'est déplacé, durant les 14 jours précédant l'identification, au sein de la même unité administrative

*Le RGPH ayant été réalisé en 2014, on inclut la tranche d'âge qui avait 55 ans au moment du recensement, et qui a atteint les 60 ans aujourd'hui. La mortalité et les déménagements ne sont pas pris en compte

Le score Vep a une valeur minimale de 1 et une valeur maximale non définissable. Il a une valeur de 1 pour, par exemple, un ménage constitué d'une personne seule vivant dans un logement d'une pièce, habitant dans une division administrative n'ayant eu aucun cas détecté, et où aucune personne contaminée ne s'est rendue dans les 14 jours précédant la détection. Le score sera élevé pour, par exemple, un ménage de 4 personnes dont une a plus de 60 ans, vivant dans un logement de 2 pièces, habitant dans une division administrative où un cas a été détecté, et où un cas (le même ou un autre) a circulé durant les 14 jours précédents sa détection.

Ainsi, à l'issue de l'étape 3, nous disposerons d'un jeu de données unique indiquant, pour chaque ménage marocain, deux scores de vulnérabilité : socio-économique et épidémiologique.

Ces deux scores seront combinés en appliquant une multiplication pondérée, ainsi résultant en un score de vulnérabilité combinée, « Vc », où :

$$Vc = \beta_1 Vse + \beta_2 Vep.$$

Les valeurs de β_1 et β_2 seront définies ultérieurement.

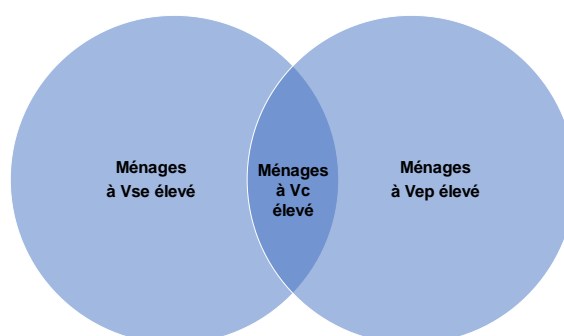
Etape 4 : identification des zones/clusters de vulnérabilité socio-économique et épidémiologique

Cette étape consiste à identifier les zones (chaque îlot et division administrative correspondante) qui concentrent le pourcentage de ménages ayant les plus grandes vulnérabilités socioéconomiques et/ou épidémiologiques à partir du score total de risque Vc. Un tableau peut être généré pour chaque division administrative, indiquant la distribution des ménages les plus vulnérables au sein de la division administrative. Il est également possible de générer des cartes indiquant les points où se situent les ménages les plus vulnérables, afin de diriger les efforts des autorités.

Parmi les catégories de ménages qui pourront être identifiées, figurent les suivantes :

- Les ménages ayant un score Vc élevé, c'est-à-dire combinant une vulnérabilité socio-économique Vse et épidémiologique Vep élevées. Les pouvoirs publics devront concentrer leurs efforts d'appui socio-économique au confinement et de testing dans les divisions/îlots-là comprenant la concentration la plus forte de cette catégorie de ménage. (*Intersection des deux diagrammes dans la figure ci-dessous*)
- Les ménages ayant un score de Vep élevé. Les autorités devront concentrer les efforts de testing dans ces divisions/îlots-là ayant la concentration la forte de cette catégorie.
- Les ménages ayant un score Ves élevé. Cette catégorie n'est pas prioritaire pour la politique de testing mais peut l'être pour un programme d'appui socio-économique.

Figure 3 : Ménages par type de vulnérabilité Vse, Vep et Vc



Etape 5 : développement de l'application web et mobile

Dans un premier temps, l'équipe développera une application web dont l'objectif est de visualiser les données collectées, en offrant, en temps réel, la cartographie et les indicateurs développés. Une fois l'application web développée et ses fonctionnalités opérationnelles, elle sera déclinée en application mobile, de façon à permettre l'accès à l'information sur le terrain.

Fonctionnalités de l'application :

1. Visualiser la cartographie du district administratif en faisant apparaître les îlots de vulnérabilité socio-économique (point vert), îlots de vulnérabilité épidémiologique (point bleu) et îlots de vulnérabilités combinées (points rouges), ces derniers étant des zones d'intervention prioritaires.
2. Recevoir des notifications dans le cas où des personnes contaminées ont été identifiées dans leur district ou lorsqu'il a été établi qu'elles ont été en contact significatif

avec autrui, dans le territoire du district. Cette information est géographique, et ne comporte aucune donnée personnelle relative à la personne contaminée.

Utilisateurs de l'application :

- Les utilisateurs "actifs" (à définir en concertation avec les pouvoirs publics), ont accès à l'information et sont en mesure d'alimenter la base de données en géolocalisant les lieux de contacts des personnes contaminées (ex : domicile, lieu de travail). Les utilisateurs "actifs" doivent nécessairement avoir accès à la fonctionnalité web pour entrer les données de localisation, sur la progression géographique de l'épidémie. Il est conseillé que les utilisateurs actifs soient des fonctionnaires ayant accès à ces données de manière quotidiennes et en mesure d'alimenter le système en temps réel.
- Les utilisateurs "passifs" ont accès à l'information relevant de leur district d'intervention. Il s'agit des agents d'autorité locaux (gouverneurs, pacha, caïd), en contact avec la population, et chargé d'affecter localement les ressources. Il est conseillé que les utilisateurs "passifs" soient équipés d'un smartphone, afin d'avoir accès à l'application et aux notifications en temps réel.

Note relative à la protection des données personnelles et de la vie privée : l'application permettra de renseigner uniquement des points géographiques en lien avec la propagation. Ces points seront ceux des domiciles, lieux de travail et lieux de contacts significatifs identifiés par les autorités. Les usagers de l'application auront donc accès à des points indiquant des zones de vulnérabilité épidémiologiques n'incluant aucune donnée personnelle sur les personnes infectées.

Etape 6 : Propositions d'appui de la société civile rendues possibles par ce projet

De nombreux citoyens ont fait part de leur volonté de participer à l'effort national de lutte contre cette épidémie. Les résultats du mapping et l'identification la plus fine possible des zones de concentration des vulnérabilités permettra d'orienter la solidarité nationale vers les zones où elle est la plus attendue.

Au niveau des autorités locales (pachalik, caïdat, cercles et annexes administratives), des rosters de volontaires âgés de 18 à 40 ans, en bonne condition physique, pourraient être créés, et se voir confier la gestion d'une zone géographique restreinte, entourant le périmètre de leur domicile. Pour que cette opération se fasse en sécurité, les volontaires devront s'inscrire auprès des autorités locales, disposer d'un téléphone portable, et de procédures simples, telles que :

- Faire des courses pour les personnes vulnérables et leur livrer sans entrer en contact avec elles,
- Faire régulièrement le tour d'un certain nombre d'habitations pour s'assurer, questionnaire en main, qu'aucune personne vulnérable n'est laissée dans le besoin,
- Appuyer les efforts des autorités dans d'éventuelles initiatives d'organisation logistique impliquant une connaissance des quartiers d'intervention

Si les élus communaux, et plus généralement les conseils de la commune manquent de moyens humains et techniques pour assurer cette fonction de coordination, les élu(e)s peuvent cependant être des moyens de transmission efficaces de l'information vis-à-vis des citoyens, en mesure de mobiliser leurs réseaux pour assurer le recrutement de volontaires.