

# Master statistique pour l'évaluation et la prévision 2022-2023

## Outils pour le Big Data

### Analyse de l'inclusion financière des pays où les niveaux de l'indice de développement sont plus faibles

**Ismael Djoulde Diallo**

Dépôt git : [https://github.com/ISMAEL-DD/Outils\\_BigData.git](https://github.com/ISMAEL-DD/Outils_BigData.git)

#### Table des matières

Contexte .....	2
Condition d'accès aux données.....	2
1. Tendances générales de l'inclusion financière dans le monde .....	3
1. 1. Adultes détenant un compte dans le monde.....	3
1. 2. Influence du niveau d'éducation à l'accès aux services financiers .....	4
1. 3. Influence de l'urbanisation sur la détention d'un compte .....	5
Préparation des données.....	5
2. Indice de développement humain et inclusion financière.....	7
2. 1. Analyse de l'inclusion financière dans les pays aux plus bas niveaux d'IDH.....	8
Conclusion .....	10
Annexe .....	10
Références.....	11

## Contexte

L'inclusion financière est un enjeu majeur pour l'économie mondiale. Elle permet d'offrir des services financiers à des personnes exclues du système financier, et donc d'une certaine façon d'accroître le volume des transactions financières et tirer profit de ses effets multiplicateurs sur l'économie. L'inclusion financière permet aussi d'avoir une traçabilité des transactions financières effectuées pour lutter contre le blanchiment d'argent. Toutefois, le niveau d'inclusion financière n'est pas le même partout dans le monde, plusieurs facteurs expliquent cette disparité dans le monde.

Ce présent rapport se propose de faire une analyse de l'état de l'inclusion financière dans le monde et des liaisons entre quelques-uns de ses déterminants. Pour ce faire, nous utiliserons les données anonymisées récoltées dans le cadre du rapport sur l'inclusion financière Global Findex 2021<sup>1</sup> ainsi que la base de données sur l'indice de développement humain IDH<sup>2</sup> du programme des nations unies pour le développement PNUD.

Dans ce travail, un accent particulier sera mis sur le niveau de l'inclusion financière dans les pays ayant les plus faibles indices de développement humain.

Le producteur du rapport Global Findex 2021 est le groupe de recherche sur le développement de la banque mondiale. Actualisées le 16 décembre 2022, les données du Global Findex 2021 ont été recueillies pour 127 variables à partir d'entretiens téléphoniques ou en face à face auprès de 127854 personnes adultes choisies chacun au hasard au sein de ménages distincts et représentatifs eux-mêmes sélectionnés aléatoirement dans 123 économies représentant 91% de la population adulte mondiale. Les 127854 adultes ont été interrogés par Gallup Poll<sup>3</sup>, un cabinet spécialisé dans le conseil en stratégie et le sondage d'opinion.

## Condition d'accès aux données

Pour accéder aux données Findex au format JSON et les télécharger, il faut obligatoirement créer un compte sur le site [microdata.worldbank.org](https://microdata.worldbank.org/index.php/catalog/4607) qui est une collection jeu de données de la Banque mondiale et d'autres organisations internationales, régionales et nationales. Ainsi, avec les identifiants on peut utiliser toutes les données disponibles sur Microdata Library. En revanche, les données de l'IDH au format csv sont accessibles sans restriction aucune, il suffit de se rendre sur le site du PNUD pour télécharger les données. Pour le calcul de l'IDH, les données sont collectées auprès des institutions afférentes.

Les deux jeux de données brutes une fois téléchargés et mis au bon format ont été sauvegardé sur mon espace cloud MongoDB pour ne pas user de la capacité de stockage de mon ordinateur. Pour la préparation des données le langage python (voir [https://github.com/ISMAEL-DD/Outils\\_BigData.git](https://github.com/ISMAEL-DD/Outils_BigData.git)) a été utilisé et pour les requêtes Mongosh a été utilisé.

---

<sup>1</sup> <https://microdata.worldbank.org/index.php/catalog/4607>

<sup>2</sup> [https://hdr.undp.org/sites/default/files/2021-22\\_HDR/HDR21-22\\_Statistical\\_Annex\\_HDI\\_Table.xlsx](https://hdr.undp.org/sites/default/files/2021-22_HDR/HDR21-22_Statistical_Annex_HDI_Table.xlsx)

<sup>3</sup> <https://www.gallup.com/home.aspx>

Nous allons nous restreindre aux personnes interrogées ayant un compte Mobile Money<sup>4</sup> ou un compte au sein d'une institution financière, à des variables démographiques des individus comme le pays de résidence, l'âge et le sexe, mais aussi à des variables reflétant le niveau d'inclusion financière tels que l'utilisation des services financiers (comme les cartes bancaires, les services d'épargne), les paiements et les transferts en ligne ou mobile. Nous tiendrons aussi compte des occasions lors desquelles les services financiers sont utilisés par les individus retenus mais aussi des raisons de non-détention de compte pour certains. Nous prendrons en considération la représentativité de chaque adulte interrogé au moyen de sa pondération. Aussi, pour tenir compte des disparités du niveau de vie entre pays, nous utiliserons l'indice de développement humain du pays de chacun des individus. Enfin, les variables retenues seront décrites tout au long du rapport.

Dans la suite de ce rapport, nous utiliserons le terme « compte » qui englobera les comptes ouverts dans une institution financière et les comptes Mobile Money ouverts chez un opérateur téléphonique.

## 1. Tendances générales de l'inclusion financière dans le monde

Pour la suite de ce rapport, nous avons exclus les adultes résidents au Kosovo et à Taiwan parce que l'IDH de ces deux économies n'est pas fourni, nous retiendrons donc 125854 adultes.

### 1. 1. Adultes détenant un compte dans le monde

Parmi les 125854 adultes retenus, 29.3% n'ont pas de compte et 70.7% détiennent un compte. Ces taux reflètent les tendances mondiales dans la mesure où le poids<sup>5</sup> des individus a été pris en compte dans le calcul. Cette répartition est illustrée par la figure 1.

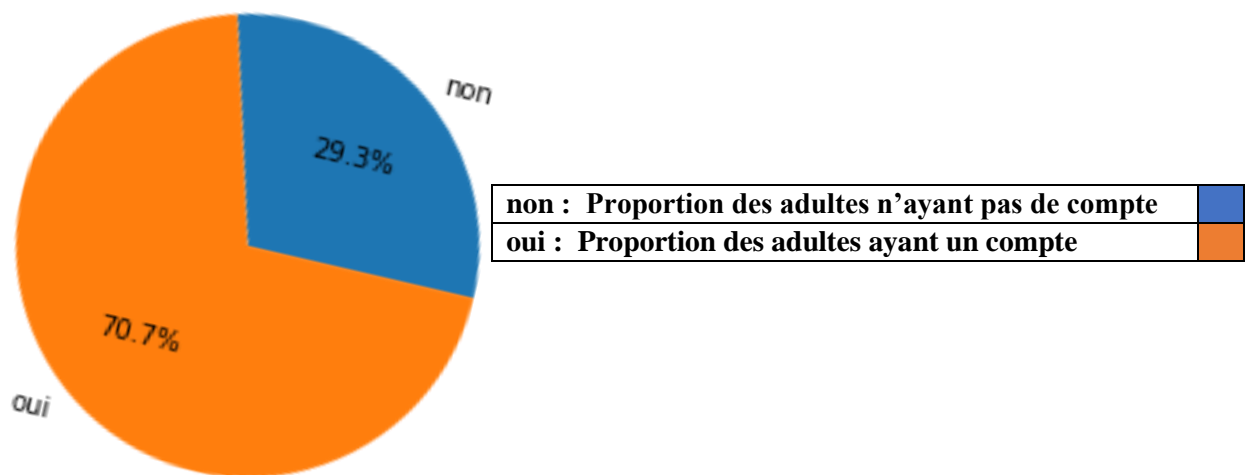


Figure 1 : Répartition des adultes selon qu'ils aient un compte ou non.

Dans le rapport FINDEX, la banque mondiale distingue six (6) régions qui sont : East Asia & Pacific, Europe & Central Asia, Latin America & Caribbean, Middle East & North Africa, South Asia et Sub-Saharan Africa. Un septième groupe d'adultes a été constitué pour regrouper tous les adultes interrogés ayant des revenus élevés, et ce, indépendamment de leur région, ce groupe c'est le High Income.

<sup>4</sup> Le mobile money (Banking), stricto sensu, désigne les services financiers par téléphone portable, offerts par les établissements de crédit. Il s'agit principalement de services de consultation de soldes, de paiement de factures, d'émission de cartes prépayées et de transfert d'argent. Ce concept, lato sensu, s'étend à l'ensemble des services financiers pouvant être offerts avec ou sans compte bancaire par tout établissement agréé à cet effet. [\[Références\]](#)

<sup>5</sup> Lors de l'enquête, un poids a été attribué à chacun des individus. Ce poids reflète la représentativité de l'individu en question dans la population mondiale et sert à inférer les statistiques de l'échantillon à la population mondiale.

La répartition des adultes ayant un compte ou non est décrit dans la figure 2. On décèle une dépendance significative entre la région des adultes et la détention d'un compte. Il existe également une dépendance significative entre le pays (economy) de résidence des adultes et la détention d'un compte.<sup>6</sup>

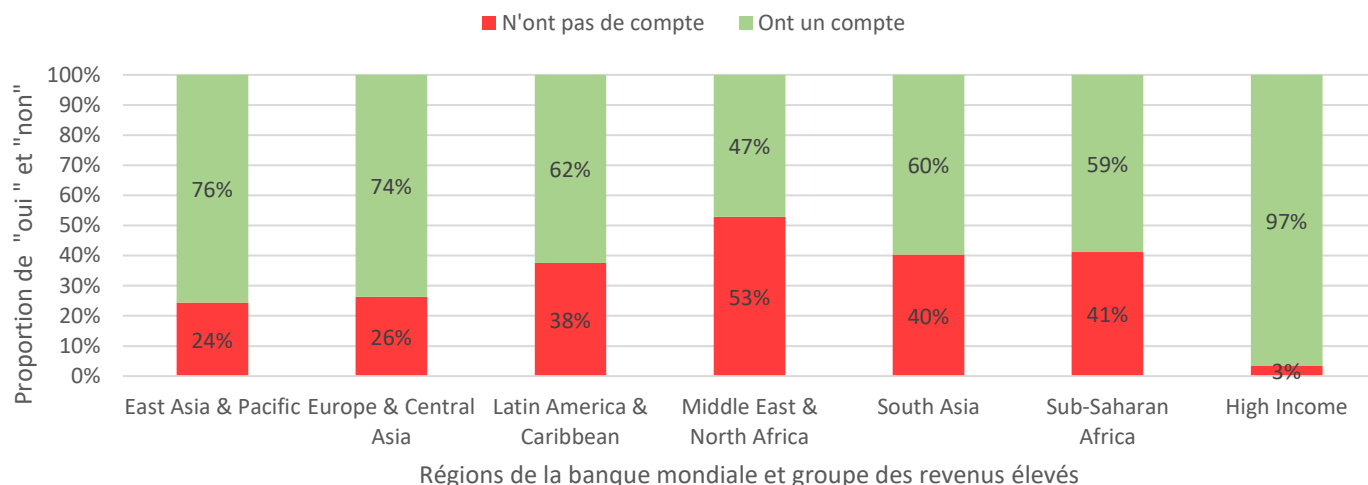


Figure 2 : Répartition des adultes selon la région et la détention (oui ou non ) d'un compte.

Seulement 3% des adultes aux revenus élevés n'ont pas de compte, ce taux se situe à 40% dans la région South Asia. En revanche, les taux de ceux qui ont un compte dans ces deux groupes valent respectivement 97% et 60%.

## 1. 2. Influence du niveau d'éducation à l'accès aux services financiers

Le niveau d'éducation a un grand impact sur l'accès aux services financiers. Les connaissances sur les produits et services financiers et les avantages associés à leur utilisation augmentent au fur et à mesure que le niveau d'éducation augmente. Les personnes ayant un niveau d'éducation élevé sont également plus susceptibles de disposer des compétences nécessaires pour utiliser les technologies financières.

La figure 3 ci-dessous illustre bien cela. La proportion des adultes interrogés ayant un compte est plus élevée chez ceux qui ont fait l'enseignement supérieur que les autres, ensuite suivent ceux qui ont fait un niveau secondaire puis ceux n'ayant que le niveau primaire.

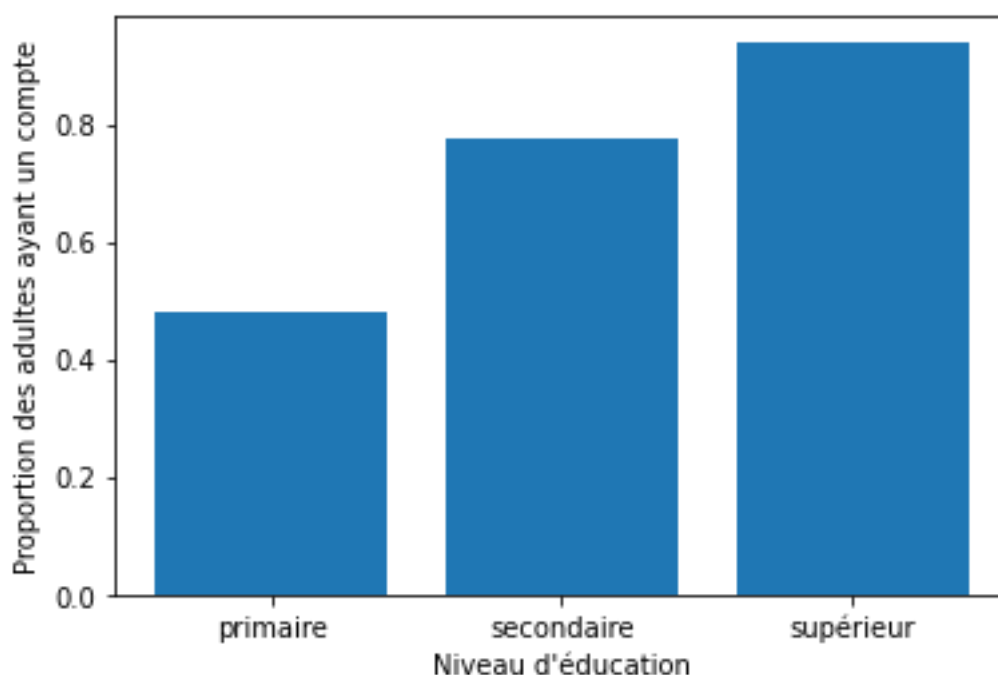


Figure 3 : Répartition des adultes interrogés ayant un compte selon leur niveau d'éducation.

<sup>6</sup> Un test d'indépendance du khi-deux a été réalisé et les détails sont disponibles dans la partie annexe [\[A1\]](#).

### 1. 3. Influence de l'urbanisation sur la détention d'un compte

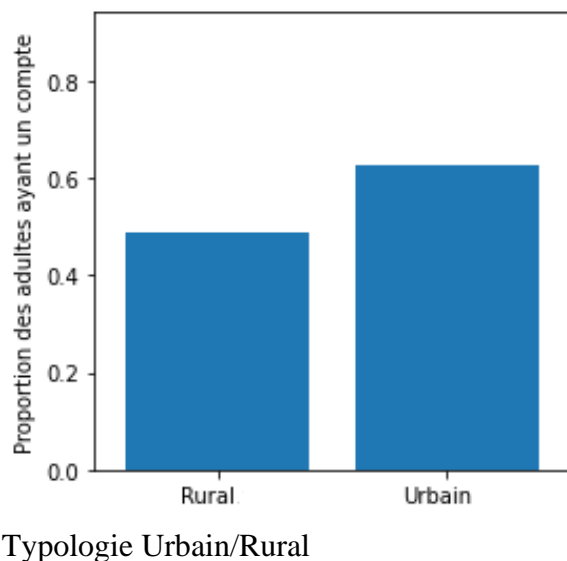


Figure 4 : Adultes ayant un compte selon leur urbanisation.

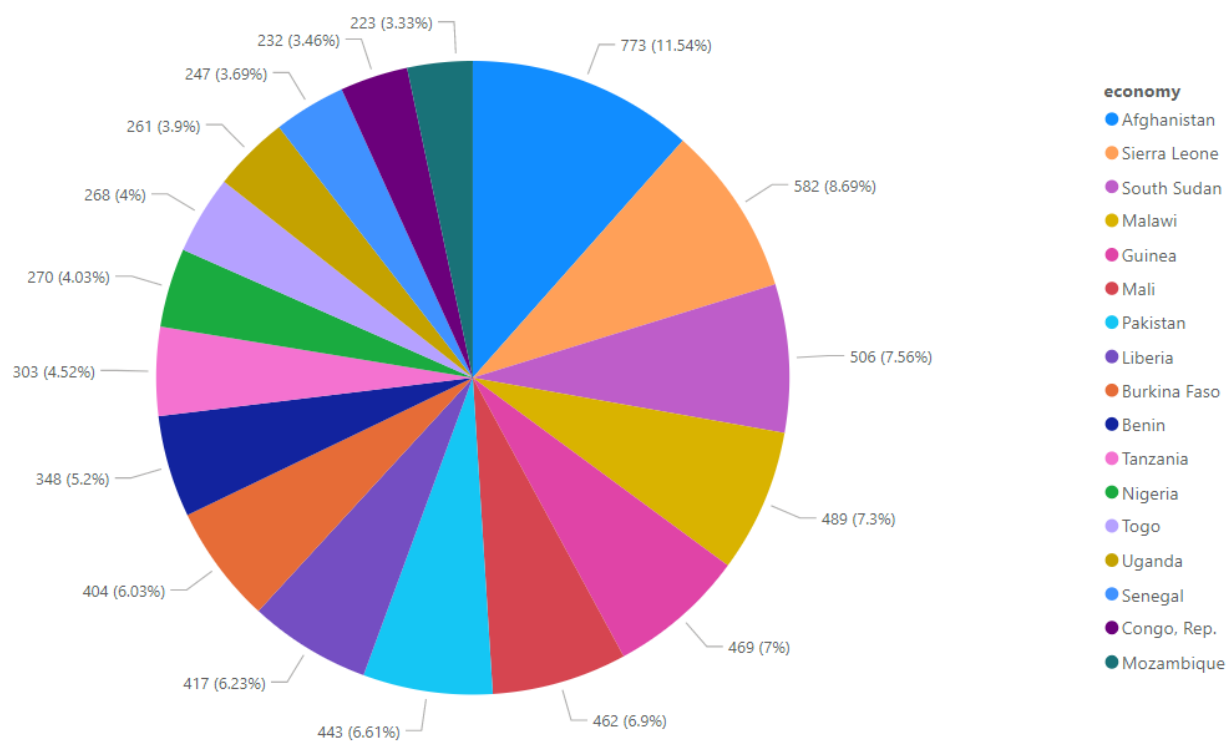
L'urbanisation a un impact important sur l'accès aux services financiers.

Les zones urbaines, contrairement aux zones rurales, ont généralement un meilleur accès aux infrastructures, ce qui permet aux institutions financières de s'y installer et de fournir des services financiers à une population plus large. Aussi, les zones urbaines ont généralement un plus grand nombre de personnes éduquées et une plus grande capacité à utiliser les technologies financières, ce qui permet aux personnes d'accéder plus facilement aux services financiers et de les utiliser efficacement. La figure 4 ci-dessus montre la répartition des adultes ayant un compte selon leur zone d'habitation lors de l'enquête. Il y a une dépendance significative entre urbanisation et détention d'un compte [\[A2\]](#).

Contrairement à la région et le pays de résidence, le niveau du revenu (High Income), le niveau d'éducation et l'urbanisation des adultes dans le monde, il n'y a pas de dépendance significative entre le sexe (femme/homme) et la détention d'un compte [\[A3\]](#).

## Préparation des données

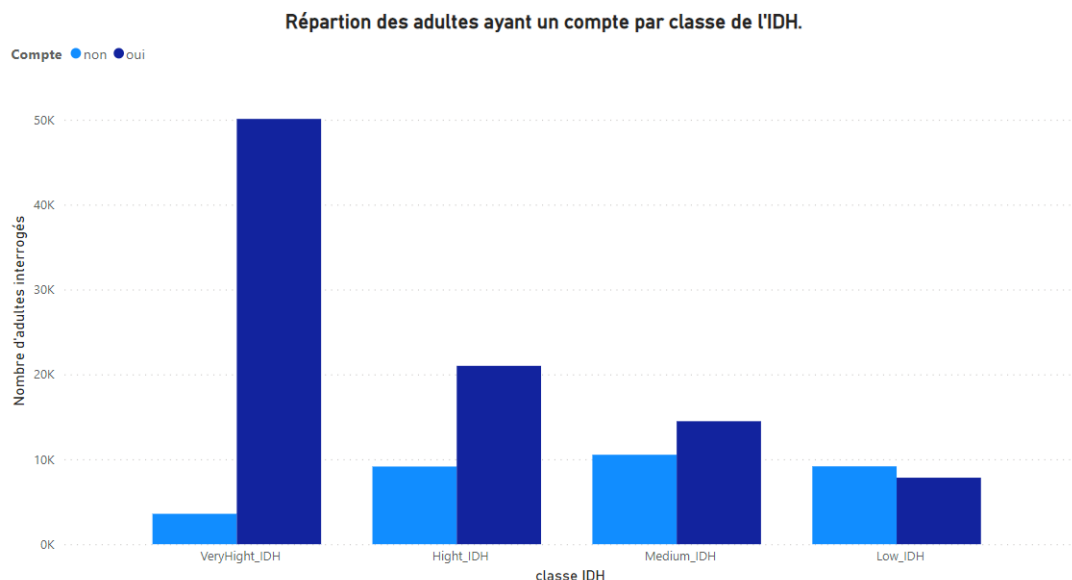
Nous avons fusionné les datasets Findex et IDH, la clé de jointure est la variable economy. Il y'avait plus d'économies dans l'IDH dataset que dans Findex, nous avons dû retenir les économies présentes dans Findex. Certaines économies n'ont pas le même nom dans les deux datasets, il a fallu créer une fonction pour renommage. Aussi, nous avons-nous même créé la variable classe IDH dans le dataset final. Pour les tests statistiques, des tableaux de contingences ont été construites. La majorité des variables dans Findex ont des valeurs définies entre 1 et 5, il a fallu donc recourir au dictionnaire des données de la Findex Database pour identifier à quoi correspond chaque chiffre pour ensuite le remplacer par sa signification. Enfin, comme toute enquête, les interviewés ne répondent pas toutes les questions qui leur sont posées, nous avons ainsi exclu les adultes n'ayant répondu à toutes les questions qui leur sont posées afin de constituer le jeu de données LowIDH. Et finalement, LowIDH n'est constitué que d'adultes n'ayant pas de compte.



Répartition des adultes retenus n'ayant pas de compte par pays.

## 2. Indice de développement humain et inclusion financière

Il existe un lien avéré entre l'indice de développement humain (IDH) et l'inclusion financière. Les pays ayant un IDH plus élevé ont tendance à avoir des niveaux plus élevés d'inclusion financière, ce qui signifie qu'une plus grande proportion de la population a accès aux services financiers. Cela s'explique par le fait que les pays ayant un IDH élevé ont généralement des économies plus fortes et des infrastructures plus développées, ce qui peut faciliter la croissance des institutions financières et l'expansion des services financiers à un plus grand nombre de personnes. Un IDH élevé tend également à être associé à des niveaux d'éducation et d'alphabétisation plus élevés, ce qui peut également contribuer à l'inclusion financière en améliorant la compréhension des produits et services financiers par les populations.<sup>7</sup> Dans la partie suivante, nous allons nous intéresser au groupe d'adultes résidents dans les pays aux plus bas niveaux d'IDH (LowIDH) afin d'explorer plus en détail les facteurs qui expliquent le faible niveau d'inclusion financière dans ce groupe. Il s'agit d'un jeu de données de 6697 individus et 16 variables, dans ce jeu de données nous avons retenu que les données sans valeurs manquantes.



Le graphique ci-dessus nous montre clairement que le nombre d'adultes interrogés et ayant un compte est plus important dans les classes VeryHight\_IDH, Hight\_IDH et Medium\_IDH. On constate une tendance inverse dans la classe LowIDH, les adultes n'ayant pas de compte sont en effet majoritaires.

<sup>7</sup> Abdelghaffar, R.A., Emam, H.A. and Samak, N.A. (2022), "Financial inclusion and human development: is there a nexus?", Journal of Humanities and Applied Social Sciences, Vol. ahead-of-print No. ahead-of-print. <https://doi.org/10.1108/JHASS-11-2021-0178>.

## 2. 1. Analyse de l'inclusion financière dans les pays aux plus bas niveaux d'IDH

Dans le dataset LowIDH, il n'y a que des adultes qui n'ont pas de compte. Pour ces adultes, nous avons les raisons de non-détention d'un compte données par les adultes lors de l'enquête de la banque mondiale. Pour toutes les questions posées, l'adulte répondra oui ou non

-l'economy ou le pays de résidence de l'adulte

-Nous avons l'accès à internet : l'adulte répond Yes ou No si c'est un facteur qui l'empêche d'utiliser les services financiers.

- Les croyances religieuses

-NoAc\_Expensive : l'adulte n'a pas de compte parce qu'il est très coûteux d'en avoir de détenir un compte.

-NoAc\_Family : il n'a pas de compte parce qu'un membre de sa famille a déjà un compte

-NoAc\_Far: il est loin des institutions financières

-NoAc\_LackDoc : le fait de ne pas détenir des documents légaux.

NoAc\_Money : le fait de ne pas avoir d'argent

-NoAc\_NoFinServices: il n'a pas besoin de services financiers

-NoAc\_lackTrust : Manque de confiance au système financier

-AccountNoHelp: aura-t-il besoin d'aide pour utiliser son compte si tentait qu'il en détiendrait un.

- ReceiveWage : reçoit-il de l'aide ?

ci-dessous les requêtes réalisées sur les données LowIDH à l'aide MongoSH.

```
Atlas atlas-12rlxf-shard-0 [primary] findex> db.LowIDH.count()
DeprecationWarning: Collection.count() is deprecated. Use countDocuments instead.
6697
```

Nombre d'observations

```
Atlas atlas-12rlxf-shard-0 [primary] findex> db.LowIDH.distinct("economy")
[
  'Afghanistan', 'Benin',
  'Burkina Faso', 'Congo, Rep.',
  'Guinea', 'Liberia',
  'Malawi', 'Mali',
  'Mozambique', 'Nigeria',
  'Pakistan', 'Senegal',
  'Sierra Leone', 'South Sudan',
  'Tanzania', 'Togo',
  'Uganda'
]
```

Combien d'économies dans LowIDH : 17  
db.LowIDH.distinct("economy").length

```
Atlas atlas-12rlxf-shard-0 [primary] findex> db.LowIDH.aggregate([
  {
    $group: {
      _id: null,
      minAge: { $min: "$age" },
      maxAge: { $max: "$age" },
      avgAge: { $avg: "$age" },
      stdAge: { $stdDevSamp: "$age" },
      minWgt: { $min: "$wgt" },
      maxWgt: { $max: "$wgt" },
      avgWgt: { $avg: "$wgt" },
      stdWgt: { $stdDevSamp: "$wgt" }
    }
  }
])
{
  _id: null,
  minAge: '15',
  maxAge: '99',
  avgAge: null,
  stdAge: null,
  minWgt: '0.146237199',
  maxWgt: '4.853373766',
  avgWgt: null,
  stdWgt: null
}
```

Statistiques descriptives des variables Age et Weight



```
Atlas atlas-12rlxf-shard-0 [primary] findex> db.LowIDH.aggregate([
  { $group: { _id: "$Received_Wage", count: { $sum: 1 } } } ]])
[ { _id: 'No', count: 5983 }, { _id: 'Yes', count: 714 } ]
Atlas atlas-12rlxf-shard-0 [primary] findex>
```

Dans nos données 5983 adultes ne reçoivent pas de salaire.

```
Atlas atlas-12rlxf-shard-0 [primary] findex> db.LowIDH.aggregate([
  { $group: { _id: "$NoAc_Money", count: { $sum: 1 } } } ]])
[ { _id: 'No', count: 1287 }, { _id: 'Yes', count: 5410 } ]
Atlas atlas-12rlxf-shard-0 [primary] findex>
```

Dans nos données 5410 adultes disent qu'ils n'ont pas de compte parce qu'ils n'ont pas d'argent.

```
Atlas atlas-12rlxf-shard-0 [primary] findex> db.LowIDH.aggregate([
  { $group: { _id: "$NoAc_Expensive", count: { $sum: 1 } } } ]])
[ { _id: 'Yes', count: 2430 }, { _id: 'No', count: 4267 } ]
Atlas atlas-12rlxf-shard-0 [primary] findex>
```

Dans nos données 2430 adultes disent qu'ils n'ont pas de compte parce que les services financiers coûtent cher.

```
Atlas atlas-12rlxf-shard-0 [primary] findex> db.LowIDH.aggregate([
  { $group: { _id: "$NoAc_LackDoc", count: { $sum: 1 } } } ]])
[ { _id: 'No', count: 4067 }, { _id: 'Yes', count: 2630 } ]
Atlas atlas-12rlxf-shard-0 [primary] findex>
```

Dans nos données 2630 adultes disent qu'ils n'ont pas de compte parce qu'ils n'ont pas de documents officiels.

```
Atlas atlas-12rlxf-shard-0 [primary] findex> db.LowIDH.aggregate([
  { $group: { _id: "$NoAc_lackTrust", count: { $sum: 1 } } } ]])
[ { _id: 'Yes', count: 1613 }, { _id: 'No', count: 5084 } ]
Atlas atlas-12rlxf-shard-0 [primary] findex>
```

Dans nos données 1613 adultes disent qu'ils n'ont pas de compte parce qu'ils ne font pas confiance aux institutions financières.

Activate Windows  
Go to Settings to activate Windows.

```
Atlas atlas-12rlxf-shard-0 [primary] findex> db.LowIDH.aggregate([
  ... { $group: { _id: "$economy", count: { $sum: { $cond: [ { $eq: [ "$internetaccess", "yes" ] }, 1, 0 ] } } } },
  ... { $sort: { count: -1 } },
  ... { $limit: 1 }
  ... ])
[ { _id: 'Mali', count: 0 } ]
```

Le Mali est le pays où les gens ont moins access à internet.

```
... { $group: { _id: "$economy", count: { $sum: { $cond: [ { $eq: [ "$internetaccess", "no" ] }, 1, 0 ] } } } },
... { $sort: { count: -1 } },
... { $limit: 1 }
... ])
[ { _id: 'Liberia', count: 0 } ]
```

Le Liberia est le pays où les adultes répondent qu'ils ont moins access à internet.

```
Atlas atlas-12rlxf-shard-0 [primary] findex> db.LowIDH.aggregate([
  ... { $match: { economy: "Nigeria" } },
  ... { $group: { _id: "$NoAc_ReligiousRea", count: { $sum: 1 } } }
  ... ])
[ { _id: 'No', count: 251 }, { _id: 'Yes', count: 19 } ]
Atlas atlas-12rlxf-shard-0 [primary] findex>
```

Pour 19 Nigerians, la religion est ce qui justifie qu'il n'ont pas de compte et 251 autres, la religion n'est pas un obstacle

Pour 133 Afghan, la religion est ce qui justifie qu'il n'ont pas de compte et pour les 640 autres la religion n'est pas un obstacle.

```
Atlas atlas-12rlxf-shard-0 [primary] findex> db.LowIDH.aggregate([
  { $match: { economy: "Afghanistan" } },
  { $group: { _id: "$NoAc_ReligiousRea", count: { $sum: 1 } } }
  ... ])
[ { _id: 'No', count: 640 }, { _id: 'Yes', count: 133 } ]
```

## Conclusion

L'analyse que nous avons réalisée nous a permis d'avoir une vision particulière de l'inclusion financière dans le monde et dans les pays où les indices de développement humain sont bas. Nous avons également réussi à étudier les liaisons entre les déterminants de l'inclusion financière, ce qui caractérise certains pays et leurs similitudes avec d'autres pays. Nous avons réussi à identifier les facteurs qui justifient la faible inclusion financière dans certains pays ; faciliter l'accès à internet, l'urbanisation et l'éducation permettrait d'améliorer l'inclusion financière dans certains pays.

## Annexe

**A.1 :** La procédure `chi2_contingency` (Test d'indépendance khi -deux ) sous python appliqué au tableau de contingence du couple (account, regionwb) des 125854 adultes fournit le résultat ci-dessous. Le seuil de significativité par défaut, et donc 5%, a été retenu pour l'interprétation de la p-value des différents tests réalisés.

```
# On réalise un test de khi deux pour étudier la dépendance entre avoir un compte et être dans une région donnée.
chi2, p, dof, expected = chi2_contingency(table)

# la p-value du test
if p < 0.05:
    print("Il existe une dépendance significative entre les variables account et regionwb")
else:
    print("Il n'existe pas une dépendance significative entre les variables account et regionwb")
```

Il existe une dépendance significative entre les variables account et regionwb

Cas du couple (account, economy)

```
[26] # On réalise un test de khi deux pour étudier la dépendance entre avoir un compte et être dans une économie donnée
chi2, p, dof, expected = chi2_contingency(table)

# la p-value du test
if p < 0.05:
    print("Il existe une dépendance significative entre les variables account et economy")
else:
    print("Il n'existe pas une dépendance significative entre les variables account et economy")
```

Il existe une dépendance significative entre les variables account et economy

**A.2 :** Test d'indépendance khi -deux entre la zone d'habitation (urbanicity) et account.

```
# tableau de contingence de l'urbanisation vs account
table = pd.crosstab(data_hdi['urbanicity_f2f'], data_hdi['account'], values=data_hdi['wgt'], aggfunc='sum')

# test de khi 2
chi2, p, dof, expected = chi2_contingency(table)

# la p-value du test
if p < 0.05:
    print("Il existe une dépendance significative entre les variables account et la zone d'habitaion")
else:
    print("Il n'existe pas une dépendance significative entre les variables account et la zone d'habitaion")
```

Il existe une dépendance significative entre les variables account et la zone d'habitaion

**A.3 :** Test d'indépendance khi -deux entre le genre et account.

```
[36] # Cross-tabulation of 'account' and 'female' variables, taking weight into account
      replace_Colstring(data_hdi,"female","1","Femmes")
      replace_Colstring(data_hdi,"female","2","Hommes")
      data_hdi_clean = data_hdi.dropna(subset=['account', 'female', 'wgt'])
      account_by_gender = data_hdi_clean.groupby(['account', 'female'])['wgt'].sum()

      # la p-value du test
      if p < 0.05:
          print("Il existe une dépendance significative entre le genre du répondant et sa détention d'un compte")
      else:
          print("Il n'existe pas une dépendance significative entre le genre du répondant et sa détention d'un compte")
```

Il n'existe pas une dépendance significative entre les variables account et la zone d'habitation

## Références

-Asli Demirgüç-Kunt Leora Klapper Dorothe Singer Saniya Ansar, The Global Findex Database 2021, The World Bank Group. <https://www.worldbank.org/en/publication/globalfindex> [Consulté le 25/12/2022]

-PNUD : Human Development Report, Explore HDI, <https://hdr.undp.org/data-center/human-development-index> [Consulté le 25/12/2022].

-World Bank : Financial inclusion <https://www.worldbank.org/en/topic/financialinclusion/overview> [Consulté le 25/12/2022].

-Editions Harmattan : La mobile Banking ou Mobile Money ; vivement un instrument de lutte contre la corruption. [https://www.editions-harmattan.fr/auteurs/article\\_pop.asp?no=31662&no\\_artiste=31021](https://www.editions-harmattan.fr/auteurs/article_pop.asp?no=31662&no_artiste=31021) [Consulté le 25/12/2022].

-Geoconfluences : Critères, indicateurs et indices de comparaisons internationales : lesquels retenir ? <http://geoconfluences.ens-lyon.fr/doc/etpays/Medit/popup/Indicateurs.htm> [Consulté le 25/12/2022].