

Précarité énergétique - Rapport 06/05/2019 (SEMAINE 1)

1. Tâches effectuées

Tout d'abord, nous avons lu la documentation qui nous a été mise à disposition et à l'aide de celle-ci nous avons sélectionné les variables pertinentes parmi l'enquête EU-SILC réalisée afin de visualiser les données concernant le thème de la précarité énergétique.

=> tableau Excel (en annexes).

Nous pourrions au travers de statistiques descriptives représenter ces données.

- **Proposition :**

Nous avons fait le choix de réaliser notre application de visualisation des données avec le langage R et le package shiny :

- Réalisation de graphiques traditionnels (boxplot, barplot, histogramme, nuage de point, série chronologique, ...)
- Réalisation de plusieurs cartes dynamiques.
 - Une carte de l'Europe à l'échelle des pays
 - Une carte nationale pour chaque pays à l'échelle des régions (NUTS2 au mieux)

Ainsi, pour cela, nous avons commencé à s'intéresser aux données

- Mise en place d'une importation des données de manière automatique et flexible. Nous avons 4 fichiers .csv pour chaque pays et chaque année.
 - Household Register (D),
 - Personal Register (R),
 - Household Data (H),
 - Personal Data (P).

Le formatage et la structuration des fichiers est harmonisée de façon à ce qu'on puisse importer l'ensemble des données (tous les pays et toutes les années) sans avoir à dézipper les archives.

- Nous nous sommes limités aux variables principales nous permettant de mesurer "la précarité énergétique". Il en ressort 3 :
 - ability to keep adequately warm (HH050)
 - arrears on utility bills (HS021)
 - Leaking roof, damp walls/floors/foundation or rot in window frames ... (HH040)

Nous nous sommes également limités à l'année 2013 pour l'instant.

Nous avons eu pour objectif de visualiser l'étendue de ces variables sur des cartes à l'aide du package leaflet :

- Moyenne d'une variable pour chaque région NUTS2 pour tester dans un premier temps.

Codage des variables :

Après avoir importé les données et après avoir choisi des variables d'étude, nous avons créé un tableau associant des données à projeter pour chaque région.

En ce qui concerne les variables d'études, certaines étaient binaires comme :

- HH050 : Ability to keep home adequately warm (YES = 1 / NO = 2).
- HH040 : Leaking roof, damp walls/floors/foundation, or rot in window frames or floor (YES = 1 / NO = 2).

Ces variables sont recodées de façon à avoir (YES = 1 et NO = 0).

D'autres, en revanche, qualitatives non binaires.

- HS021 : Arrears on utility bills.
 - o 1 = Yes, once
 - o 2 = Yes, twice or more
 - o 3 = No

Nous avons choisi de transformer cette variable multi-modale en variable binaire en regroupant les modalités "Yes, once" et "Yes, twice or more".

- (YES = 1 / NO = 0)

Pour chaque région, nous avons effectué la moyenne de ces variables afin d'avoir les proportions des ménages ayant répondu YES ou NO pour chaque question.

Selon la tournure de la question on choisira d'effectuer $1 - \text{MOYENNE}()$ ou juste $\text{MOYENNE}()$.

Partie Cartographie :

Après diverses recherches sur internet, nous avons décidé d'utiliser les packages leaflet et geojsonio pour faire nos cartes :

- Données GIS : les données GPS des régions ont été récupérées sur le site d'Eurostat, avec la nomenclature NUTS2 en vigueur en 2013 (la nomenclature en vigueur en 2016 n'est plus la même !).
 - o *Remarque 1 : différentes précisions des tracés des frontières sont proposées, on utilise, pour l'instant, celle la moins précise pour alléger les temps de calculs.*
 - o *Remarque 2 : l'ordre des régions diffère selon le choix de la précision, bien faire attention à bien relancer l'ensemble du code lié au traitement des données sinon des régions seront affectées de valeurs étranges.*

Nous avons remarqué qu'il y avait des normes différentes en ce qui concerne la représentation des régions. En effet, à partir de 2016, les labels des régions ont changé. Les études ayant été réalisées utilisent l'ancienne norme NUTS (celle en vigueur en 2013), elle contient les anciens labels. Nous avons donc utilisé les données cartographiques de 2013.



Figure 1. Rendu-test de la carte à l'échelle des régions.

2. Problèmes

Après étude des fichiers excel, nous avons remarqué que les données fournies issues de l'enquête EU-SILC contiennent de nombreuses données manquantes qui vont nous gêner quant à la représentation de celle-ci :

1. Nous n'avons pas de données pour les pays suivants : Pologne, Portugal, République Tchèque et la Suède.
2. Lors du questionnaire, certaines régions ont été enregistrées sous la norme NUTS1 alors que la carte est en NUTS2 (Autriche et Hongrie).
3. En Allemagne, Slovaquie et Pays-Bas : aucune région n'a été précisée par les ménages (la variable identifiant la région du ménage existe mais n'est pas remplie du tout).

3. Solutions envisagées :

1. Nous ne pouvons rien faire.
2. La valeur prise par la région NUTS1 est injectée dans toutes les sous-régions NUTS2 (problème traité comme indiqué).
3. Nous pourrions représenter ces pays à l'échelle internationale (sur la carte mondiale). Toutefois, nous ne pourrions pas représenter ces pays à l'échelle nationale avec les valeurs de chaque variable pour chaque région.

4. Poursuite

Nous envisageons d'effectuer durant la SEMAINE 2 de :

- traiter les données après importation selon la décision qui sera prise suite à la réunion.
- réaliser une maquette de ce à quoi pourrait ressembler l'application.
- commencer les statistiques descriptives.

Fichier de référence : DESCRIPTION-OF-EU-SILC-TARGET-VARIABLES-2017.pdf (déposé dans le dropbox, dossier data)

Code fichier	Code variable	Page docu2017	Intitulé	Pertinence	Remarque
QUESTIONNAIRE					
D	DB110	111	Household status	-1	
D	DB130	115	Household questionnaire result	-1	
D	DB135	117	Household interview acceptance	-1	
R	RB100	134	Sample person or co-resident	-1	Qui a répondu ? L'intéressé ou un co-résident
R	RB245	149	Respondent status	-1	agé de plus de 16 ans ou non
R	RB260	151	Type of interview	-1	
H	HB060	167	Year of the household interview	-1	
P	PB100	245	Month of the personal interview	-1	
P	PB110	246	Year of the personal interview	-1	
VARIABLES D'IDENTIFICATION					
D	DB010	95	Year of the survey	-1	
D	DB020	96	Country	-1	
D	DB030	98	Household ID	-1	
R	RB010	120	Year of the survey	-1	
R	RB020	121	Country	-1	
R	RB030	123	Personal ID	-1	
R	RB040	125	Current household ID	-1	
H	HB010	162	Year of the survey	-1	
H	HB020	163	Country	-1	
H	HB030	164	Household ID	-1	
P	PB010	236	Year of the survey	-1	
P	PB020	237	Country	-1	
P	PB030	238	Personal ID	-1	
ZONE GEOGRAPHIQUE					
D	DB040	99	Region	2	
D	DB100	110	Degree of urbanisation	2	
D	DB120	114	Contact at address	1	servira à créer ou identifier une autre variable
R	RB120	137	Move to [Location]	0	pas sûr de l'horizon temporel de cette variable
LOGEMENT					
R	RB160	140	Number of months in household during the incoming reference	2	
R	RB200	144	Residential status	2	
H	HH010	171	Dwelling type	2	
H	HH021	172	Tenure status	2	
H	HH030	175	Number of rooms available to the household	2	
H	HH040	177	Leaking roof, dam walls/floors/foundation or rot in window frames or	3	
H	HH050	178	Ability to keep adequately warm	3	
H	HH060	179	Current rent related to occupied dwelling	2	
H	HH061	180	Subjective rent	2	estimation du loyer si le proprio le louait

H	HH070	181 Total housing cost	2
H	HH081	184 Bath or shower in dwelling	1
H	HS160	202 Problems with the dwelling : too dark, not enough light	1
H	HS180	204 Pollution, grime or other environment problems	1
H	HS190	205 Crime, violence or vandalism in the area	1
PERSONNE			
R	RB031	124 Year of immigration	1
R	RB080	132 Year of birth	0 voir 3 lignes plus bas
R	RB090	133 Sex	0 voir 3 lignes plus bas
H	HY060G/N	215 Social exclusion not elsewhere classified	1
P	PB140	249 Year of birth	1
P	PB150	250 Sex	1
P	PB190	254 Marital status	2
P	PB220A	257 Citizenship 1	1 peu utile pour le modèle mais utile pour la stat desc
P	PB220B	257 Citizenship 2	1 peu utile pour le modèle mais utile pour la stat desc
P	PE010	258 Current education activity	1
P	PE020	260 ISCED level currently attended	1
P	PE040	263 Highest ISCED level attained	2
H	HC010T4	429 Food at home	1
P	PC010T4/HC020T	430 Food outside home	1
SANTE			
P	PH010	265 General health	2
P	PH020	267 Suffer from any chronic (long-standing) illness or condition	2
H	HS200	348 Financial burden of medical care	2
R	RC010T	374 General health (child)	1
LOISIR			
P	PD060	341 Regularly participate in a leisure activity	0,5
ACTIVITE PROFESSIONNELLE			
R	RB170	141 Main activity status during the incoming reference period	2
R	RB210	145 Basic activity status	2
P	PL015	280 Wheter person has ever worked	2
P	PL031	283 Self-defined current economic status	1
P	PL051	291 Occupation (ISCO-08 (COM))	2 permettra de construire une variable synthétique
P	PL060	292 Number of hours usually worked per week in main job	2 possibilité de sommer cette variable à PL100 pour synthétiser
P	PL100	296 Total number of hours usually worked in second, third, ... jobs	2
P	PL140	300 Type of contract	2
P	PL160	303 Change of job since last year	1
P	PL200	307 Number of years spent in paid work	1
SITUATION FAMILIALE			
R	RB220	146 Father ID	1 retrace-t-on le périmètre familial de l'individu ?
R	RB230	147 Mother ID	1 idem
R	RB240	148 Spouse/partner ID	1 idem

P	PB160	251 Father ID	1 idem
P	PB170	252 Mother ID	1 idem
P	PB180	253 Spouse/partner ID	1 idem

SITUATION FINANCIERE

H	HS060	191 Capacity to face unexpected financial expenses	2
H	HS120	198 Ability to make ends meet	2
H	HS140	200 Financial burden of the total housing cost	2
H	HY010	206 Total household gross income	2
H	HY020	208 Total disposable household income	2
H	HY022	208 Total disposable household income before social transfers other than old-age and survivor's benefits	0 à garder ?
H	HY023	208 Total disposable household income before social transfers including other than old-age and survivor's benefits	0 à garder ?
P	PY200G	337 Gross monthly earnings for employees	2
P	PC040T4/HC050T	434 Regular savings (in a typical month)	2
H	HV010T4	437 Value of main residence	2

AIDES

H	HY050G/N	215 Family/children-related allowances	1
H	HY070G/N	215 Housing allowances	1
H	HY081G/N	224 Alimonies received (compulsory + voluntary)	1
H	HY131G/N	230 Alimonies paid (compulsory + voluntary)	1
P	PY090G/N	325 Unemployment benefits	1
P	PY100G/N	325 Old-age benefits	1
P	PY110G/N	325 Survivor' benefits	1
P	PY120G/N	325 Sickness benefits	1
P	PY130G/N	325 Disability benefits	1
P	PY140G/N	325 Education-related allowances	1

DETTES / RETARDS

H	HH071	183 Mortgage principal repayment	1 arrears = retards
H	HS011	186 Arrears on mortgage or rental payments	1
H	HS021	187 Arrears on utility bills	3
H	HS031	188 Arrears on hire purchasing installments or other loan payments	1
H	HS150	201 Financial burden of the repayment of debts from hire purchases or	2
H	HI010T3	421 Total amount in arrears for utility (housing) bills/rent/mortgage repayments related to the main residence	1
H	HI020T3	422 Arrears on non-housing household bills	1
H	HI050T3	424 Household has debts from hire purchases or other loan payments (other than mortgage taken to buy the main dwelling) or credit cards	1