

## Précarité énergétique - Rapport 13/05/2019 (SEMAINE 2)

### 1. Tâches effectuées

#### 1.1. Problèmes d'encodage et de normes en vigueur

Tout d'abord, nous avons corrigé les problèmes d'affichage de différentes régions en fonction des NUTS choisis. En effet, dans l'enquête certaines régions ne respectaient pas la norme définie par les cartes comme Londres. Tous les problèmes d'encodage des régions présents dans les données publiques sont réglés à partir de dictionnaires de conversion stockés dans des fichiers csv. Par ailleurs, si de nouveaux problèmes s'additionnent dans les prochaines enquêtes, il suffira d'ajouter ces exceptions.

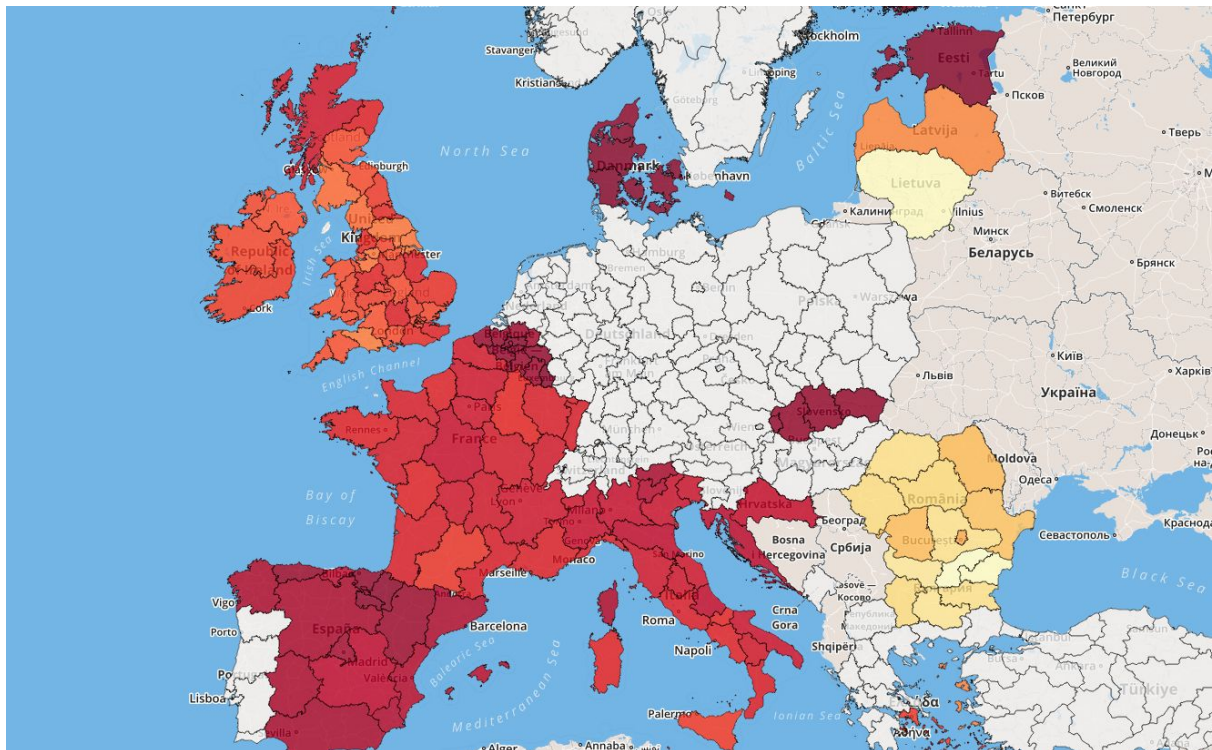
#### 1.2. Sécurisation des données et séparation calculs / représentations

En ce qui concerne l'idée d'un fichier final calculé à partir d'un code R indépendant et qui normalement ne sera lancé pas plus d'une fois par an, l'idée est d'avoir un dataframe qui contient les valeurs pour chaque pays, pour chaque niveau NUTS et pour chaque année. On aboutit alors à une table qui contient toutes les données nécessaires pour la cartographie et les statistiques descriptives. Ceci permet d'exécuter l'application sans lien avec les données micro.

Cette "grande" table (qui ne fait pas plus de 5000 lignes finalement, donc rapidement importé dans l'application shiny) sera alors filtrée selon les entrées de l'utilisateur de l'application pour renvoyer une table filtrée qui sera représentée.

	PAYS	REGION	NUTS	ANNEE	warm	leaking
338	LV	LV	NUTS 0	2013	0.8293	0.1691
339	PL	PL	NUTS 0	2013	NA	NA
340	SE	SE	NUTS 0	2013	NA	NA
341	RO	RO	NUTS 0	2013	0.7454	0.1835
342	PT	PT	NUTS 0	2013	NA	NA
343	TR	TR	NUTS 0	2013	NA	NA
344	SK	SK	NUTS 0	2013	0.9798	0.5903
345	SI	SI	NUTS 0	2013	0.9454	0.2793
346	UK	UK	NUTS 0	2013	0.8809	0.1849
347	TR	TR2	NUTS 1	2004	NA	NA
348	ES	ES4	NUTS 1	2004	0.9354	0.1527
349	TR	TR7	NUTS 1	2004	NA	NA
350	TR	TR4	NUTS 1	2004	NA	NA
351	TR	TRA	NUTS 1	2004	NA	NA
352	GR	GR1	NUTS 1	2004	NA	NA
353	IT	ITF	NUTS 1	2004	0.9290	0.1879
354	ES	ES5	NUTS 1	2004	0.9438	0.1495
355	TR	TRB	NUTS 1	2004	NA	NA
356	TR	TR5	NUTS 1	2004	NA	NA

**Figure 1.** Table résultant d'un fichier R indépendant



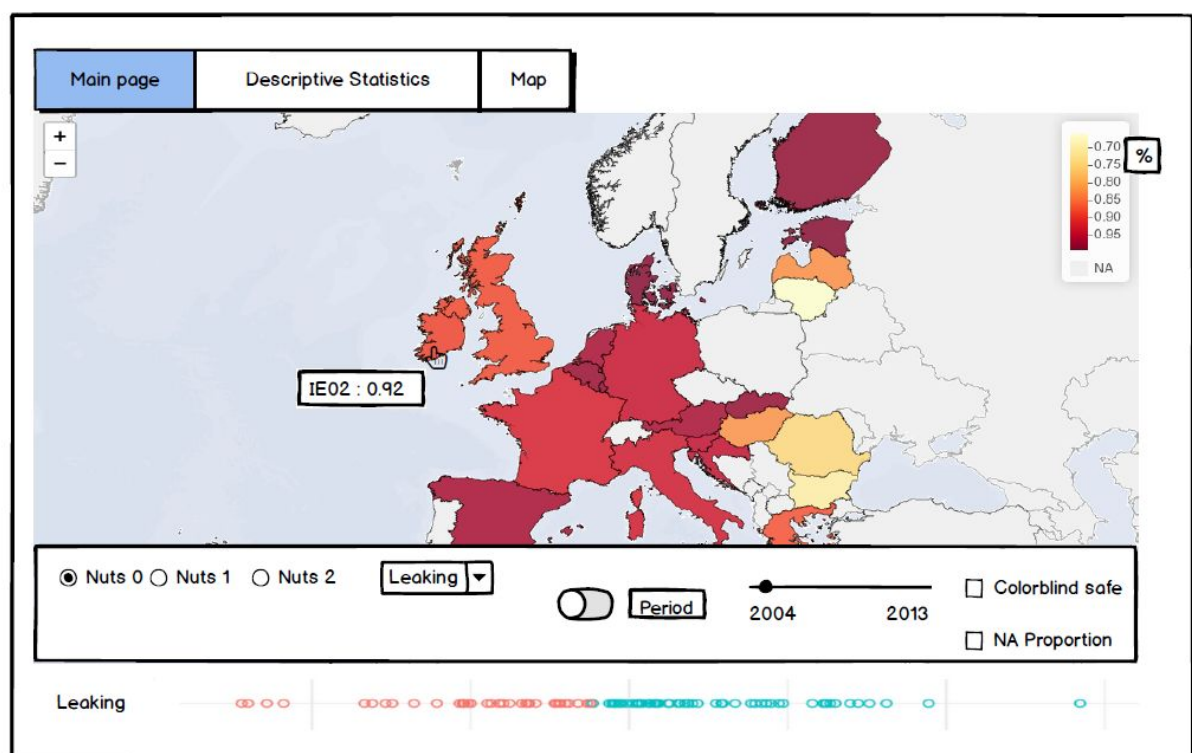
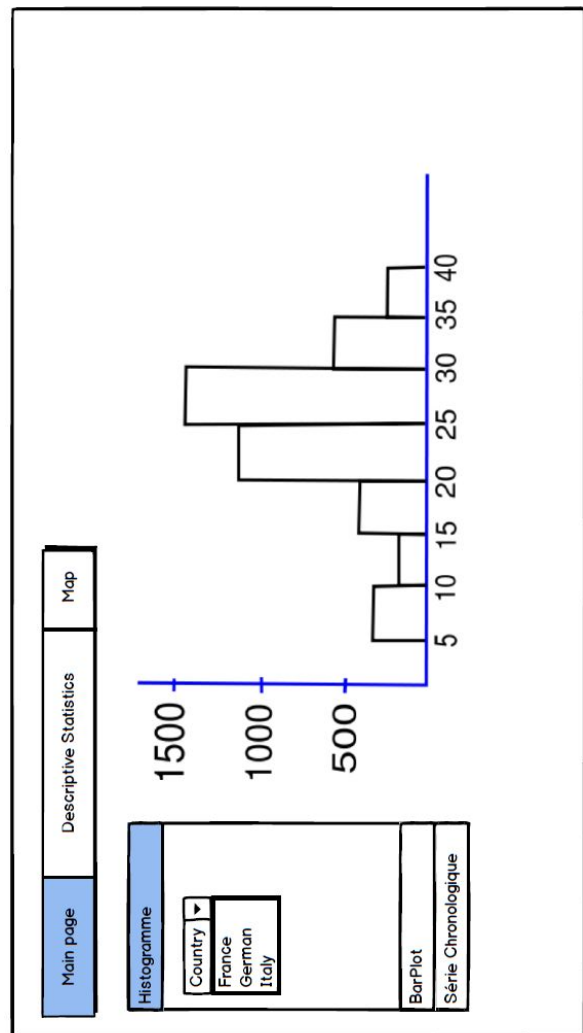
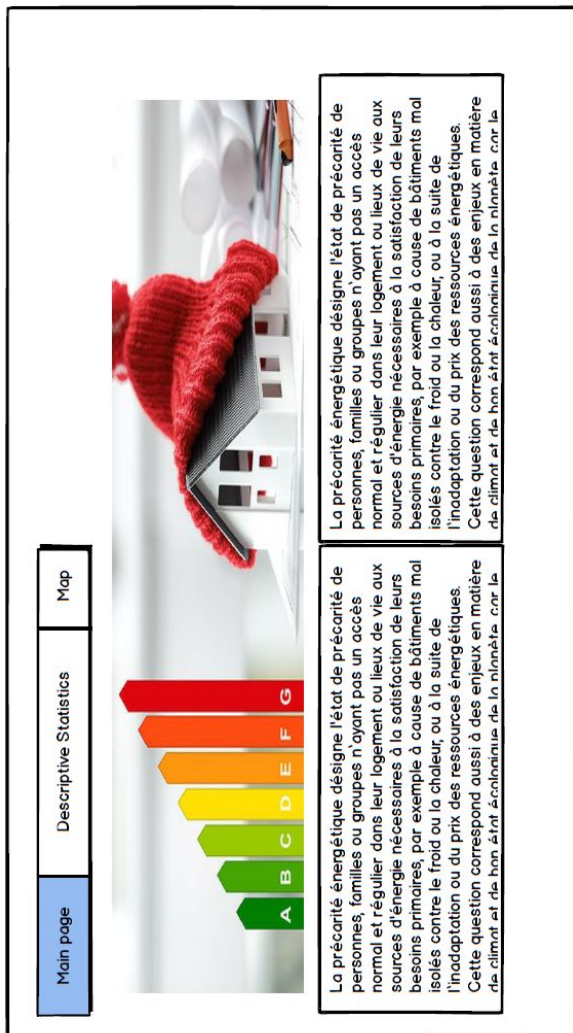
**Figure 2. Résultat cartographie**

### 1.3. Maquette de l'application

Ensuite, nous avons réalisé une maquette de l'application Shiny avec l'outil Balzamiq mockup, disponible dans le dossier de partage dropbox. Nous avons pour idée de permettre à l'utilisateur de sélectionner dynamiquement les données à l'aide de boutons, de listes...

L'application sera alors découpée en deux onglets, un pour la cartographie et l'autre pour les statistiques descriptives.

Par ailleurs, en ce qui concerne la carte, nous avons pour idée de mettre une barre indiquant l'étendu (la distribution) de chaque variable, associée avec les mêmes couleurs que celle utilisées dans la carte. Cette distribution se fera sous la forme d'une projection des valeurs prises par les régions représentées sur un axe horizontal borné entre 0 et 1.



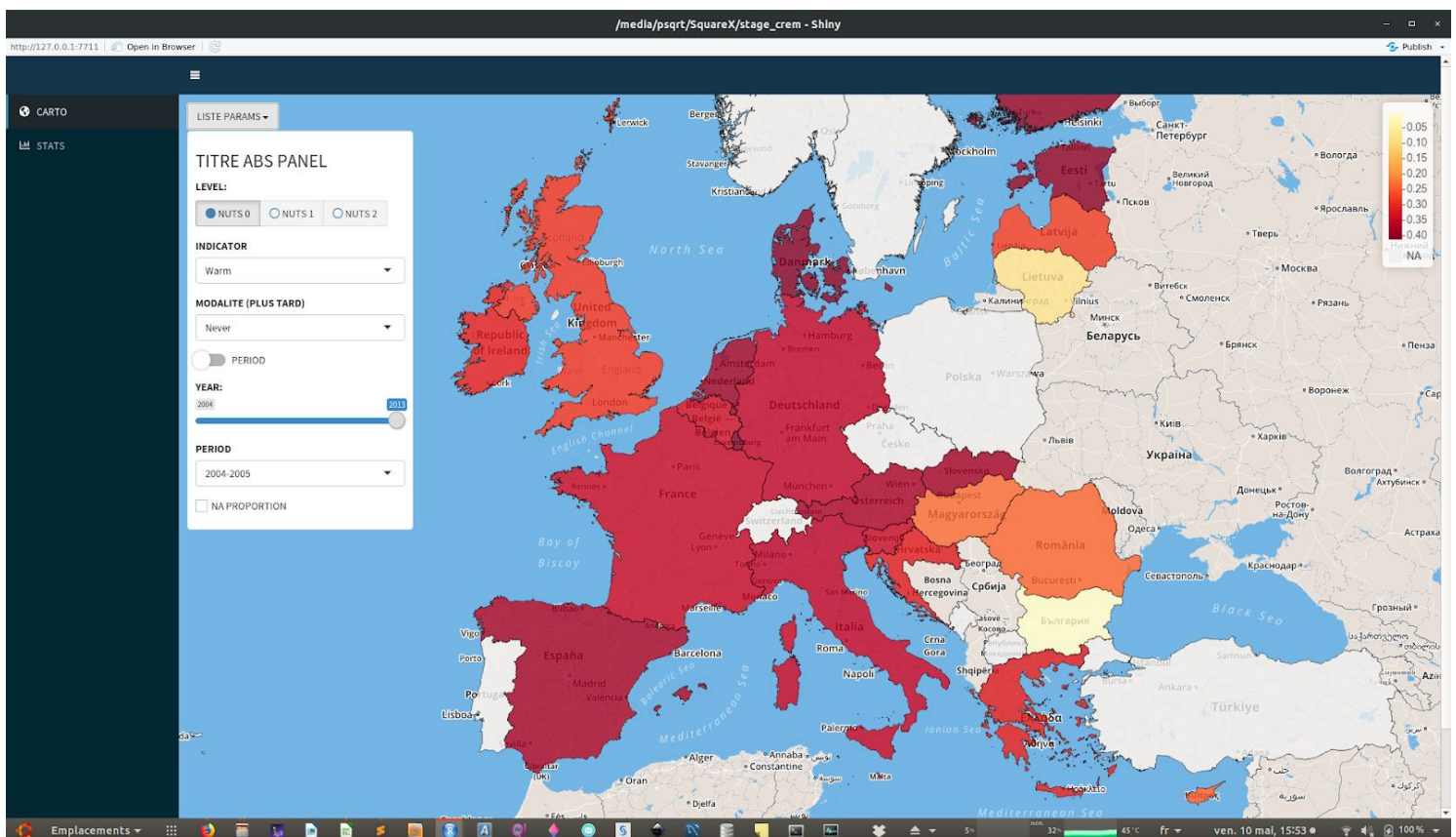


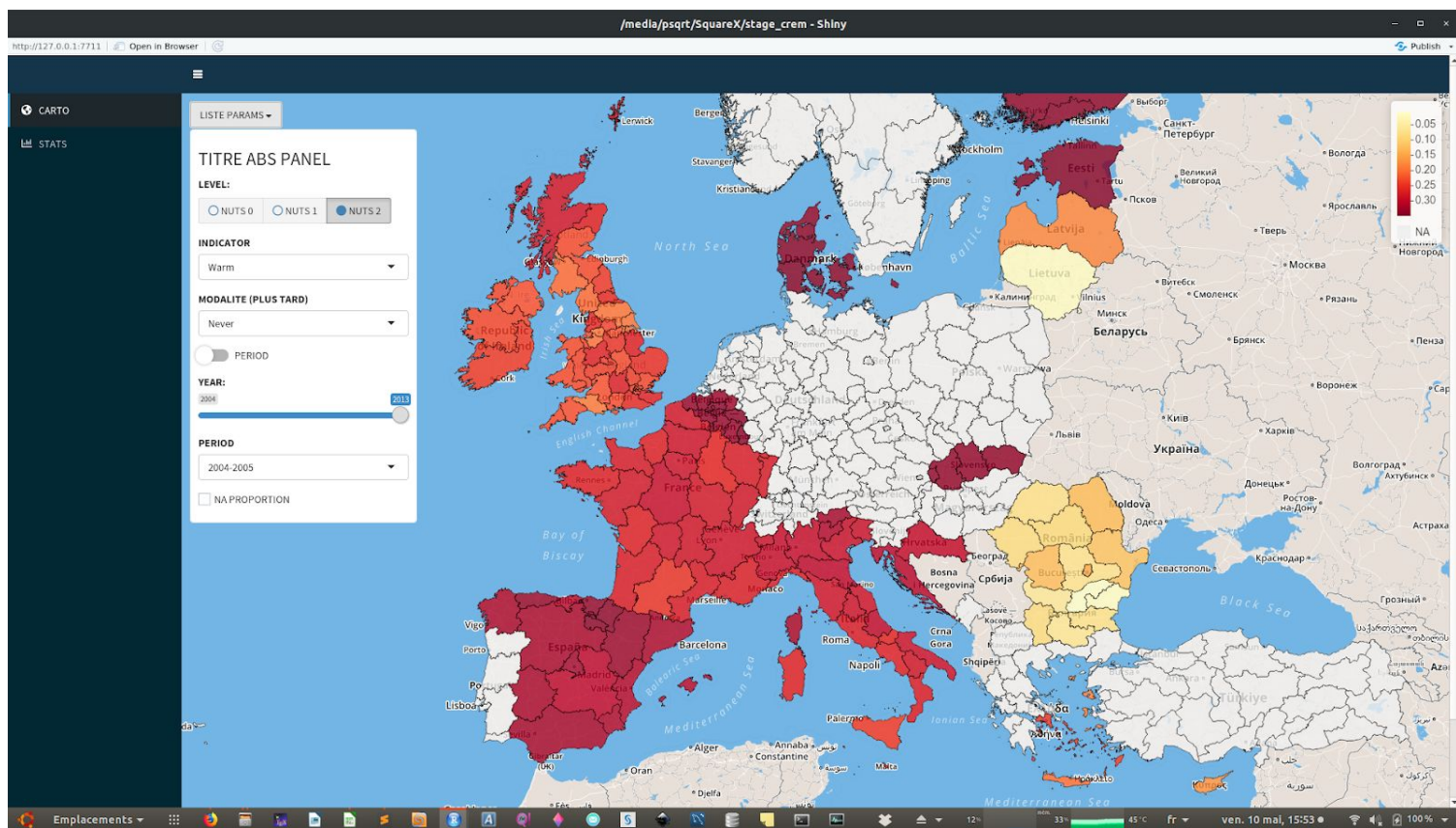
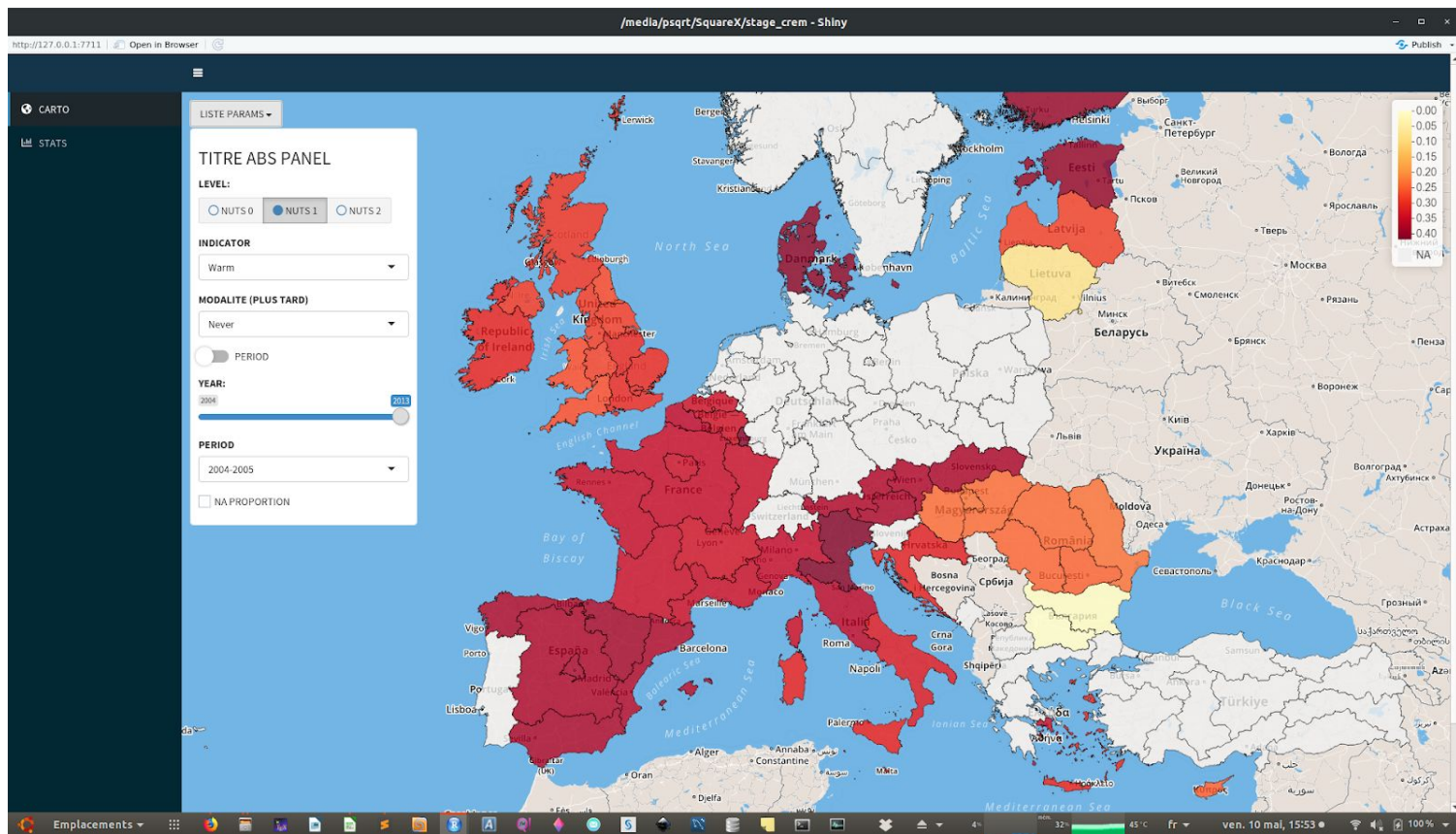
La maquette sera composée :

- D'un onglet d'accueil, présentant le but de l'enquête, et la façon dont les données ont été récoltées.
- D'un onglet composé de plusieurs pages d'analyses descriptives, avec des histogrammes, barplots, des séries chronologique et des tests de corrélation (l'essentiel de données étant qualitatives).
- D'une page cartographie où l'on représentera les différentes variables (en fonction de chaque modalité pour référence).

Enfin, après validation de la maquette, nous avons adapté notre code de façon dynamique sous Shiny en utilisant le package leaflet pour la cartographie. Nous avons fait le choix comme convenu de réaliser un fichier "données.csv" contenant le summary (ici moyennes) des valeurs de chaque variable pour chaque NUTS.

L'utilisateur est donc capable pour le moment d'utiliser les boutons radio pour changer le mode de représentation (NUTS 0, NUTS 1, NUTS 2). Il peut également utiliser le slider pour choisir une année à représenter.





### 3. Initiatives

Nous avons décidé de réaliser l'application Shiny en anglais étant donné qu'il s'agit d'une étude européenne. Ainsi, nous attendons la confirmation de l'équipe pédagogique.

### 4. Difficultés

Nous remarquons malgré le regroupement des données dans un fichier "données.csv" que l'affichage de la carte prend du temps (7 secondes en moyenne). Nous essayerons donc d'optimiser les calculs ou de trouver d'autres solutions pour rendre l'application la plus fluide possible.

Egalement, nous aurions souhaité écrire le nom des pays en anglais et les afficher au dessus des polygones pour avoir une "belle" carte. Toutefois leaflet nous prive de cette liberté. Nous pensons peut-être le faire grâce à un fichier personnalisé.

### 5. Poursuite

Nous pensons compléter la carte en implémentant l'option Période qui permettra de choisir une période de représentation des données. Toutefois, les normes NUTS changeant tous les 3 ans, nous ne pourrons représenter que les valeurs prélevées dans cet intervalle de 3 ans (2004-2005, 2006-2009, 2010-2012, 2013-2016). Nous avons donc opté pour une liste de choix possibles pour restreindre les possibilités de l'utilisateur.

Par ailleurs, nous aimerions réaliser comme convenu l'option "NA proportion" qui permettrait d'indiquer pour chaque pays le pourcentage de données manquantes.

Par ailleurs, ils nous restent à concevoir toute la partie statistique descriptive, le choix des variables étant effectuée lors de la réunion de la semaine 1.